Maximum likelihood estimation in mis-specified reliability distributions.

John, Andrea

Maximum Likelihood Estimation in Mis-Specified Reliability Distributions

by

Andrea John
BSc (University of Wales, Swansea)

Thesis

submitted to the University of Wales
in candidature for the degree of

PHILOSOPHIÆ DOCTOR

European Business Management School
University of Wales Swansea
Swansea SA2 8PP
United Kingdom

September 2003

Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

ANDREA JOHN
15 September 2003

Statement

This thesis is the result of my own investigation, except where acknowledgement of other sources is given.

ANDREA JOHN
15 September 2003

Statement

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan (subject to the law of copyright), and for the title and summary to be made available to outside organisations.

ANDREA JOHN
15 September 2003

Acknowledgements

This thesis would never have been possible without the constant and untiring help from my supervisor, Alan Watkins. Thank you so much for everything you've done throughout the six years I've spent in university, it's greatly appreciated.

Many thanks also go out to all my friends at the European Business Management School. You've made the environment a pleasure to work in.

Finally, I'd like to thank the EPSRC for their financial support over the last three years.

ANDREA JOHN

University of Wales
September 2003

This thesis was typeset with LaTeX 2_ε by the author. LaTeX 2_ε is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society. The macros used in formatting this thesis were written by Anton Merlushkin of the European Business Management School, University of Wales.

Summary

This thesis examines some effects of fitting the wrong distribution to reliability data. The parametric analysis of any data usually assumes that the form of the underlying distribution is known. In practice, however, the choice of distribution is subject to error, so the analysis could involve estimating parameters from a mis-specified model. In this thesis, we consider theoretical and practical aspects of maximum likelihood estimation under such mis-specification. Due to its popularity and wide use, we take the Weibull distribution to be the mis-specified model, and look at the effects of fitting this distribution to data from underlying Burr, Gamma and Lognormal models. We use entropy to obtain the theoretical counterparts to the Weibull maximum likelihood estimates, and obtain theoretical results on the distribution of the mis-specified Weibull maximum likelihood estimates and quantiles such as B_{10}.

Initially, these results are obtained for complete data, and then extended to type I and II censoring regimes, where consideration of terms in the likelihood and entropy functions leads to a detailed consideration of the properties of order statistics of the distributions. We also carry out a similar investigation on accelerated data sets, where there is additional complexity due to links between accelerating factors and scale parameters in reliability distributions. These links are also open to mis-specification, so allowing for various combinations of true and mis-specified models. We present theoretical results for general scale-stress relationships, but focus on practical results for the Log-linear and Arrhenius models, since these are the two relationships most widely used.

Finally, we link both acceleration and censoring, and obtain theoretical results for a type II censoring regime at the lowest stress level.

To Mam, Dad and Kieran

Contents

1 Introduction **1**

1.1 Key references and basic definitions . . . 2

1.1.1 General definitions . . . 2

1.1.2 Particular distributions . . . 3

1.1.3 Censored data . . . 10

1.1.4 Mathematical functions . . . 11

1.2 Numerical and computational aspects . . . 14

1.3 Sample procedures . . . 14

1.3.1 Hazard and cumulative hazard plots . . . 14

1.3.2 Kolmogorov-Smirnov distances . . . 15

1.3.3 Data simulation . . . 15

1.4 Structure of thesis . . . 16

2 Maximum Likelihood: Some Practical Considerations **17**

2.1 Introduction . . . 17

2.2 Fitting G_w . . . 18

2.3 Fitting G_b . . . 20

2.3.1 The limiting distribution and Δ . . . 22

2.3.2 Fitting the three parameter Burr distribution . . . 23

2.4 Fitting G_g . . . 26

2.5 Fitting $G_{\ln}$ . . . 29

2.6 Fitting G_w to G_w data . . . 30

2.7 Fitting G_w to G_b data . . . 31

2.7.1 Assessing the goodness of fit . . . 32

2.8 Fitting G_w to G_g data . . . 41

2.8.1 Assessing the goodness of fit . . . 42

2.9 Fitting G_w to $G_{\ln}$ data . . . 42

2.9.1 Assessing the goodness of fit . . . 48

2.10 Summary . . . 49

3 Maximum Likelihood: Some Theoretical Considerations **56**

3.1 Analysing data using the correct distribution . . . 56

3.1.1 The Weibull distribution . . . 57

3.1.2 The Burr distribution . . . 58

3.1.3 The Gamma distribution . . . 61

3.1.4 The Lognormal distribution . . . 62

3.2 Analysing data using the incorrect distribution . . . 65

3.2.1 Entropy for fitting G_w to data from G_b . . . 68

3.2.2 Entropy for fitting G_w to data from G_g . . . 81
3.2.3 Entropy for fitting G_w to data from $G_{\ln}$ . . . 88
3.3 Summary . . . 95

4 Censoring 96
4.1 Type I censoring . . . 96
4.1.1 ML estimation for G_w under type I censoring . . . 97
4.1.2 ML estimation for G_b under type I censoring . . . 99
4.1.3 Fitting G_w to G_b data . . . 101
4.1.4 Analysing data using the correct distribution . . . 103
4.1.5 Analysing data using the incorrect distribution . . . 108
4.2 Type II censoring . . . 121
4.2.1 ML estimation for G_w under type II censoring . . . 121
4.2.2 ML estimation for G_b under type II censoring . . . 123
4.2.3 Fitting G_w to G_b data . . . 126
4.2.4 Analysing data using the correct distribution . . . 126
4.2.5 Analysing data using the incorrect distribution . . . 137
4.3 Summary . . . 149

5 Acceleration In Life Testing 150
5.1 Introduction . . . 150
5.1.1 Scale-stress relationships . . . 151
5.1.2 Choice of sample size and stress levels . . . 153
5.2 Fitting G_{W*} . . . 155
5.2.1 The EFI matrix of the Weibull MLEs . . . 158
5.2.2 The asymptotic variance of $\widehat{B}_{W,10}$ . . . 162
5.2.3 Fitting G_{WP} . . . 162
5.2.4 Fitting G_{WA} . . . 167
5.2.5 Comparison between accelerated and non-accelerated Weibull distributions . . . 170
5.3 Fitting G_{B*} . . . 171
5.3.1 Fitting the accelerated Burr distribution and Δ . . . 176
5.3.2 The EFI matrix of the Burr MLEs . . . 178
5.3.3 The asymptotic variance of $\widehat{B}_{B,10}$ . . . 183
5.4 Fitting G_{G*} . . . 183
5.4.1 The EFI matrix of the Gamma MLEs . . . 186
5.5 Fitting G_{LN*} . . . 187
5.5.1 The EFI matrix of the Lognormal MLEs . . . 189
5.6 Summary . . . 190

6 Mis-specification In Accelerated Life Testing : Some Theoretical Considerations 192
6.1 The scope for mis-specification . . . 192
6.2 Entropy for Weibull to Weibull . . . 198
6.2.1 The variance structure of the mis-specified MLEs . . . 199
6.3 Entropy for Weibull to Burr . . . 202
6.3.1 Simplifications when $\rho_m(X_i) = \rho_t(X_i)$ . . . 204
6.3.2 The variance structure of the mis-specified MLEs . . . 206
6.4 Entropy for Weibull to Gamma . . . 208

6.4.1 The variance structure of the mis-specified MLEs 211
6.5 Entropy for Weibull to Lognormal 213
6.5.1 The variance structure of the mis-specified MLEs 215
6.6 Summary 218

7 Mis-specification In Accelerated Life Testing : Further Practical Considerations 219
7.1 Getting the scale-stress relationship wrong 219
7.1.1 Fitting G_{WA} to G_{WP} 219
7.1.2 Fitting G_{WP} to G_{WA} 234
7.2 Getting the distribution wrong 249
7.2.1 Fitting G_{WP} to G_{BP} 249
7.2.2 Fitting G_{WP} to G_{GP} 256
7.2.3 Fitting G_{WP} to G_{LNP} 265
7.3 Getting the distribution and scale-stress relationship wrong 275
7.3.1 Fitting G_{WP} to G_{BA} 275
7.3.2 Fitting G_{WA} to G_{BP} 286
7.3.3 Fitting G_{WP} to G_{GA} 300
7.3.4 Fitting G_{WA} to G_{GP} 308
7.3.5 Fitting G_{WP} to G_{LNA} 316
7.3.6 Fitting G_{WA} to G_{LNP} 325
7.4 Summary 333

8 Type II Censoring In Accelerated Life Testing 334
8.1 Introduction 334
8.2 The Weibull distribution 335
8.2.1 The EFI matrix for G_{W*} 340
8.3 Fitting G_W to G_W data 357
8.4 Fitting G_W to G_B data 358
8.5 The Burr distribution 364
8.5.1 Expectations 367
8.6 Summary 371

9 Summary 372
9.1 Chapter summary 372
9.2 Critical Appraisal 375
9.3 Areas for future research 377

Bibliography 379

Appendix A : Expectations of single order statistics 384

Appendix B : Sums of reciprocals and their powers 387

Appendix C : Expectations of joint order statistics 389

Appendix D : Δ for accelerated data 391

List of Figures

1.1 The function G_w for varying β. 4
1.2 The function G_b for $\alpha = 1$ and varying τ. 6
1.3 The function G_b for $\tau = 1$ and varying α. 6
1.4 The function G_g for varying τ. 8
1.5 The function $G_{\ln}$ for varying μ. 9
1.6 The function $G_{\ln}$ for varying σ. 10

2.1 Empirical (♦) and fitted (—) cdfs for G_w for the ball bearings data set. 25
2.2 Empirical (♦) and fitted (—) cdfs for G_b for the ball bearings data set. 26
2.3 Sample (♦) and fitted (—) cumulative hazard functions for G_w for the ball bearings data set. 27
2.4 Sample (♦) and fitted (—) cumulative hazard functions for G_b for the ball bearings data set. 27
2.5 $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 50$). 34
2.6 $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 100$). 34
2.7 $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 300$). 35
2.8 $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 500$). 35
2.9 $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 1000$). 36
2.10 Sample (♦) and theoretical (—) cumulative hazard functions for G_w with $\widehat{\beta} = 2.4171$ and $\widehat{\theta} = 66.9221$. We simulate 1000 data values from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$. 37
2.11 Sample (♦) and theoretical (—) cumulative hazard functions for G_b with $\widehat{\tau} = 2.9101$, $\widehat{\alpha} = 3.4480$ and $\widehat{\phi} = 94.6071$. We simulate 1000 data values from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$. 37
2.12 $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 50$). 38
2.13 $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 100$). 39
2.14 $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 300$). 39
2.15 $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 500$). 40
2.16 $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 1000$). 40

2.17 $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 50$). . . . 43
2.18 $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 100$). . . . 43
2.19 $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 300$). . . . 44
2.20 $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 500$). . . . 44
2.21 $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 1000$). . . . 45
2.22 $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 50$). . . . 45
2.23 $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 100$). . . . 46
2.24 $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 300$). . . . 46
2.25 $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 500$). . . . 47
2.26 $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 1000$). . . . 47
2.27 $D_w - D_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 50$). . . . 50
2.28 $D_w - D_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 100$). . . . 50
2.29 $D_w - D_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 300$). . . . 51
2.30 $D_w - D_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 500$). . . . 51
2.31 $D_w - D_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 1000$). . . . 52
2.32 $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 50$). . . . 52
2.33 $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 100$). . . . 53
2.34 $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 300$). . . . 53
2.35 $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 500$). . . . 54
2.36 $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 1000$). . . . 54

3.1 E_b^* versus m for $\alpha = 4$. . . . 73
3.2 $\frac{\partial E_b^*}{\partial m}$ versus m for $\alpha = 4$. . . . 74
3.3 Plot of m_0 against α. . . . 75
3.4 Comparison of Burr (—–) and Weibull (- - -) cdfs for Burr parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$, and entropy values $\beta_0 = 2.5528$, $\theta_0 = 67.2620$. . . . 77
3.5 Maximum absolute distance between G_b and G_w when $\tau = 3$, $\alpha = 4$, $\phi = 100$, and $\beta_0 = 2.5528$, $\theta_0 = 67.2620$. . . . 78

3.6 Maximum absolute distance between G_w and G_b for $0.5 \leq \alpha \leq 4$; $\tau = 0.5$ and $\phi = 1$. 79
3.7 The relationship between τ in G_g and β_0. 85
3.8 Comparison of Gamma (—–) and Weibull (- - -) cdfs for Gamma parameters $\tau = 3$, $\alpha = 50$ and entropy values $\beta_0 = 1.8328$, $\theta_0 = 169.3772$. 86
3.9 Maximum absolute distance between G_g and G_w when $\tau = 3$, $\alpha = 50$ and $\beta_0 = 1.8328$, $\theta_0 = 169.3772$. 86
3.10 Maximum absolute distance between G_g and G_w for $0 < \tau \leq 20$; $\alpha = 1$. . . . 87
3.11 Comparison of Lognormal (—–) and Weibull (- - -) cdfs for Lognormal parameters $\mu = 2$, $\sigma = 3$ and entropy values $\beta_0 = \frac{1}{3}$, $\theta_0 = 33.1155$. 93

4.1 β_0 versus y_c for $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 111
4.2 θ_0 versus y_c for $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 111
4.3 E_b^* vs β for data simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$, and subject to type II censoring with $n = 1000$ and $r = 800$. 140

7.1 Plots of the maximum absolute distance between cdfs for the true Weibull Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios and three stress levels. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).234
7.2 Plots of the maximum absolute distance between cdfs for the true Weibull Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios and four stress levels. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).237
7.3 Plots of the maximum absolute distance between cdfs for the true Weibull Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios and three stress levels. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).247
7.4 Plots of the maximum absolute distance between cdfs for the true Weibull Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios and four stress levels. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).248
7.5 Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios, three stress levels and $\tau = 1$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).284
7.6 Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios, three stress levels and $\tau = 2$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).285
7.7 Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios, three stress levels and $\tau = 3$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).285
7.8 Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios, three stress levels and $\tau = 4$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).286

7.9 Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios, four stress levels and $\tau = 1$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1). 287
7.10 Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios, four stress levels and $\tau = 2$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1). 287
7.11 Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios, four stress levels and $\tau = 3$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1). 288
7.12 Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios, four stress levels and $\tau = 4$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1). 288
7.13 Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios, three stress levels and $\tau = 1$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1). 296
7.14 Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios, three stress levels and $\tau = 2$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1). 296
7.15 Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios, three stress levels and $\tau = 3$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1). 297
7.16 Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios, three stress levels and $\tau = 4$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1). 297
7.17 Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios, four stress levels and $\tau = 1$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1). 298

7.18 Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios, four stress levels and $\tau = 2$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1). 298
7.19 Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios, four stress levels and $\tau = 3$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1). 299
7.20 Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios, four stress levels and $\tau = 4$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1). 299
7.21 Plots of the maximum absolute distance between cdfs for the true Gamma Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios and three stress levels. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).307
7.22 Plots of the maximum absolute distance between cdfs for the true Gamma Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios and four stress levels. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).308
7.23 Plots of the maximum absolute distance between cdfs for the true Gamma Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios and three stress levels. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).315
7.24 Plots of the maximum absolute distance between cdfs for the true Gamma Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios and four stress levels. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).316
7.25 Plots of the maximum absolute distance between cdfs for the true Lognormal Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios and three stress levels. 324
7.26 Plots of the maximum absolute distance between cdfs for the true Lognormal Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios and four stress levels. 324
7.27 Plots of the maximum absolute distance between cdfs for the true Lognormal Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios and three stress levels. 332
7.28 Plots of the maximum absolute distance between cdfs for the true Lognormal Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios and four stress levels. 332

List of Tables

1.1 Lifetimes (in millions of revolutions) of 23 ball bearings; from Lieblein and Zelen (1956); Dumonceaux and Antle (1973). 2

2.1 Summary of iterations for fitting G_w to the ball bearings data in Table 1.1. . 20
2.2 Summary of iterations for fitting G_b to the ball bearings data in Table 1.1. . 25
2.3 Summary of iterations for fitting G_g to the ball bearings data in Table 1.1. . 28
2.4 Comparison of maximised log-likelihoods and Kolmogorov-Smirnov tests for the Weibull, Burr, Gamma and Lognormal distributions when these are fitted to the ball bearings data set. 30
2.5 Summaries of the MLEs for G_w when fitted to Weibull data generated with $\beta = 2$, $\theta = 100$. Figures are based on at least 10000 replications. 31
2.6 Summary statistics of the MLEs for G_b and G_w for data generated from G_b with $\alpha = 4$, $\tau = 3$, $\phi = 100$. 32
2.7 Summary statistics for comparing D_w with D_b. Data is simulated from a Burr distribution with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ and for n =50,100,300,500 and 1000. 33
2.8 Summary statistics for comparing CH_w with CH_b. Data is simulated from a Burr distribution with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ and for n =50,100,300,500 and 1000. 38
2.9 Summaries of the MLEs for G_w and G_g, when fitted to Gamma data generated with $\tau = 3$, $\alpha = 50$. 41
2.10 Summary statistics for comparing Kolmogorov-Smirnov and hazard functions for G_w and G_g. Data is simulated from G_g with $\alpha = 50$, $\tau = 3$, and for n =50,100,300,500 and 1000. 42
2.11 Summaries of the MLEs for G_w and G_{ln}, when fitted to Lognormal data generated with $\mu = 2$, $\sigma = 3$. 48
2.12 Summary statistics for comparing Kolmogorov-Smirnov and hazard functions for G_w and G_{ln}. Data is simulated from G_{ln} with $\mu = 2$, $\sigma = 3$, and for n =50,100,300,500 and 1000 49

3.1 Theoretical standard errors for the MLEs of G_w for varying n. Data is simulated from G_w with $\beta = 2$, $\theta = 100$. 58
3.2 Theoretical standard errors for the MLEs of G_b for varying n. Data is simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 60
3.3 Theoretical standard errors for the MLEs of G_g for varying n. Data is simulated from G_g with $\tau = 3$, $\alpha = 50$. 62
3.4 Theoretical standard errors for the MLEs of $G_{\ln}$ for varying n. Data is simulated from $G_{\ln}$ with $\mu = 2$, $\sigma = 3$. 64

3.5 Theoretical standard errors for the MLEs of G_w for varying n. Data is simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 72
3.6 Values of m_0 for varying α. 76
3.7 Maximum absolute distance between G_b and G_w for varying τ and α. 78
3.8 Theoretical quantiles for G_b (top) and G_w (middle), with their corresponding relative percentage errors (bottom). 80
3.9 Theoretical standard errors for the MLEs of the Weibull distribution for varying n. Data is simulated from a Gamma model with $\tau = 3$, $\alpha = 50$. 84
3.10 Maximum absolute distance between G_g and G_w for varying τ; $\alpha = 1$. 87
3.11 Theoretical quantiles for G_g and G_w, with their corresponding relative percentage errors. 88
3.12 Theoretical standard errors for the MLEs of G_w for varying n. Data is simulated from $G_{\ln}$ with $\mu = 2$, $\sigma = 3$. 92
3.13 Theoretical quantiles for G_{ln} (top) and G_w (middle), with their corresponding relative percentage errors (bottom). 94

4.1 Summary statistics for G_w and G_b for varying sample size, when these distributions are fitted to data that has undergone a type I censoring regime from G_b with parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 102
4.2 Summary statistics for G_w and G_b for varying y_c, when these distributions are fitted to data that has undergone a type I censoring regime from G_b with parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 103
4.3 Theoretical standard errors for the MLEs from G_w and G_b for varying n. Data is subjected to type I censoring, and simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 119
4.4 Theoretical standard errors for the MLEs from G_w and G_b for varying y_c. Data is subjected to type I censoring, and simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 119
4.5 Summary statistics for G_w and G_b for varying n with $r = \frac{4n}{5}$. Both distributions are fitted to data that has undergone a type II censoring regime from G_b with parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 127
4.6 Summary statistics for G_w and G_b for varying r and $n = 1000$. Both distributions are fitted to data that has undergone a type II censoring regime from G_b with parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 127
4.7 Theoretical standard errors for the MLEs from G_b for varying n and $r = \frac{4n}{5}$. Data is subjected to type II censoring, and simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 149
4.8 Theoretical standard errors for the MLEs from G_b for varying r and $n = 1000$. Data is subjected to type II censoring, and simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$. 149

5.1 Hours to failure (y_{ij}) of $n = 40$ motorettes with a new type of insulation (Nelson, 1990), based on an experiment with $k = 4$, $n_i = 10$ for $1 \leq i \leq 4$, and the X_i as shown. 151
5.2 Fitting G_{WP} to G_{WP} for $k = 2$, $B_P = 2$. We show the sample means and standard deviations of parameters, where figures are based on at least 10000 replications. Throughout, sample standard errors are denoted by S, and their theoretical counterparts by T. 163

5.3 Fitting G_{WP} to G_{WP} for $k = 3$, $B_P = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 164
5.4 Fitting G_{WP} to G_{WP} for $k = 4$, $B_P = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 165
5.5 Breakdown of the denominator of the variance of $\widehat{\alpha}_{WP}$ and $\widehat{\beta}_{WP}$ for various n_i. 166
5.6 Fitting G_{WA} to G_{WA} for $k = 2$, $B_A = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 167
5.7 Fitting G_{WA} to G_{WA} for $k = 3$, $B_A = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 168
5.8 Fitting G_{WA} to G_{WA} for $k = 4$, $B_A = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 169
5.9 The theoretical standard errors of $\widehat{B}_{10}$ from accelerated and non-accelerated Weibull data for varying n and k (at design stress). Calculations use a shape parameter of 2 and a scale parameter of 100. 170

6.1 Types of mis-specification for accelerated data sets. 193

7.1 Fitting G_{WA} to G_{WP} for $k = 2$, $B_P = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 221
7.2 Fitting G_{WA} to G_{WP} for $k = 2$, $B_P = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 222
7.3 Fitting G_{WA} to G_{WP} for $k = 3$, $B_P = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 223
7.4 Fitting G_{WA} to G_{WP} for $k = 3$, $B_P = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 224
7.5 Fitting G_{WA} to G_{WP} for $k = 4$, $B_P = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 225
7.6 Fitting G_{WA} to G_{WP} for $k = 4$, $B_P = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 226
7.7 Comparison between the true Weibull Log-linear model, with parameters $B_P = 2$, $\alpha_{WP} = 8$, $\beta_{WP} = -0.02$, and the mis-specified Weibull Arrhenius distribution, with entropy values $\widetilde{B}_A = 1.9451$, $\widetilde{\alpha}_{WA} = -2.1558$, $\widetilde{\beta}_{WA} = 2969.7718$. Here, $k = 3$, $X_1 = 50$, $X_2 = 150$ and $X_3 = 200$. 228
7.8 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1:1:1)$. 229
7.9 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and $\beta_{WP} = -0.03, \alpha_{WP} = 11$; ratio is $(1:1:1)$. 230

7.10 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1 : 2 : 3)$. 231
7.11 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(3 : 2 : 1)$. 232
7.12 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(2 : 3 : 1)$. 232
7.13 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1 : 3 : 2)$. 232
7.14 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1 : 1 : 3)$. 233
7.15 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(3 : 1 : 1)$. 233
7.16 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1 : 3 : 1)$. 233
7.17 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(1 : 1 : 1 : 1)$. 235
7.18 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(1 : 1 : 1 : 4)$. 235
7.19 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(4 : 1 : 1 : 1)$. 235
7.20 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(1 : 2 : 2 : 1)$. 236
7.21 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(2 : 1 : 1 : 2)$. 236
7.22 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(1 : 2 : 3 : 4)$. 236
7.23 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(4 : 3 : 2 : 1)$. 237
7.24 Fitting G_{WP} to G_{WA} for $k = 2$, $B_A = 2$ with equal allocations. We show the sample means and standard error of parameters, where figures are based on at least 10000 replications. 238
7.25 Fitting G_{WP} to G_{WA} for $k = 2$, $B_A = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 239
7.26 Fitting G_{WP} to G_{WA} for $k = 3$, $B_A = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 240
7.27 Fitting G_{WP} to G_{WA} for $k = 3$, $B_A = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 241
7.28 Fitting G_{WP} to G_{WA} for $k = 4$, $B_A = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 242
7.29 Fitting G_{WP} to G_{WA} for $k = 4$, $B_A = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 243

7.30 Comparison between the true Weibull Arrhenius model, with parameters $B_A = 2$, $\alpha_{WA} = 0.3$, $\beta_{WA} = 2347$, and the mis-specified Weibull Log-linear distribution, with entropy values $\widetilde{B}_P = 1.9685$, $\widetilde{\alpha}_{WP} = -0.0156$, $\widetilde{\beta}_{WP} = 8.3024$. Here, $k = 3$, $X_1 + c = 323.16$, $X_2 + c = 423.16$ and $X_3 + c = 473.16$. . 245
7.31 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WA} with $k = 3$ and the ratio $(1:1:1)$. . . . 246
7.32 Maximum distances for G_{WP} and G_{WA}. Data is from G_{WA} with $k = 4$ and the ratio $(1:1:1:1)$. . . . 248
7.33 Fitting G_{WP} to G_{BP} for $k = 2$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 250
7.34 Fitting G_{WP} to G_{BP} for $k = 2$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 251
7.35 Fitting G_{WP} to G_{BP} for $k = 3$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 252
7.36 Fitting G_{WP} to G_{BP} for $k = 3$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 253
7.37 Fitting G_{WP} to G_{BP} for $k = 4$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 254
7.38 Fitting G_{WP} to G_{BP} for $k = 4$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 255
7.39 Comparison between the true Burr Log-linear model, with parameters $\tau = 3$, $a = 4$, $\alpha_B = 8$, $\beta_B = -0.02$, and the mis-specified Weibull Log-linear distribution, with entropy values $\widetilde{B}_P = 2.55279$, $\widetilde{\alpha}_{WP} = 7.60343$, $\widetilde{\beta}_{WP} = -0.02$. Here, $k = 3$, $X_1 = 50$, $X_2 = 150$, $X_3 = 200$. . . . 257
7.40 Fitting G_{WP} to G_{GP} for $k = 2$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 259
7.41 Fitting G_{WP} to G_{GP} for $k = 2$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 260
7.42 Fitting G_{WP} to G_{GP} for $k = 3$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 261
7.43 Fitting G_{WP} to G_{GP} for $k = 3$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 262
7.44 Fitting G_{WP} to G_{GP} for $k = 4$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 263
7.45 Fitting G_{WP} to G_{GP} for $k = 4$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. . . . 264

7.46 Comparison between the true Gamma Log-linear model, with parameters $\tau = 3$, $\alpha_G = 8$, $\beta_G = -0.02$, and the mis-specified Weibull Log-linear distribution, with entropy values $\widetilde{B}_P = 1.8328$, $\widetilde{\alpha}_{WP} = 9.2201$, $\widetilde{\beta}_{WP} = -0.02$. Here, $k = 3$, $X_1 = 50$, $X_2 = 150$, $X_3 = 200$. 266
7.47 Data from a Weibull Log-linear distribution for varying B_P. 267
7.48 Data from a Lognormal Log-linear distribution for varying σ. 268
7.49 Fitting G_{WP} to G_{LNP} for $k = 2$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 269
7.50 Fitting G_{WP} to G_{LNP} for $k = 2$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 270
7.51 Fitting G_{WP} to G_{LNP} for $k = 3$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 271
7.52 Fitting G_{WP} to G_{LNP} for $k = 3$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 272
7.53 Fitting G_{WP} to G_{LNP} for $k = 4$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 273
7.54 Fitting G_{WP} to G_{LNP} for $k = 4$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 274
7.55 Comparison between the true Lognormal Log-linear model, with parameters $\sigma = 1$, $\alpha_{LN} = 8$, $\beta_{LN} = -0.02$, and the mis-specified Weibull Log-linear distribution, with entropy values $\widetilde{B}_{WP} = 1$, $\widetilde{\alpha}_{WP} = 8.5$, $\widetilde{\beta}_{WP} = -0.02$. Here, $k = 3$, $X_1 = 50$, $X_2 = 150$, $X_3 = 200$. 276
7.56 Fitting G_{WP} to G_{BA} for $k = 2$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 278
7.57 Fitting G_{WP} to G_{BA} for $k = 2$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 279
7.58 Fitting G_{WP} to G_{BA} for $k = 3$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 280
7.59 Fitting G_{WP} to G_{BA} for $k = 3$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 281
7.60 Fitting G_{WP} to G_{BA} for $k = 4$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 282
7.61 Fitting G_{WP} to G_{BA} for $k = 4$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 283

7.62 Fitting G_{WA} to G_{BP} for $k = 2$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 290
7.63 Fitting G_{WA} to G_{BP} for $k = 2$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 291
7.64 Fitting G_{WA} to G_{BP} for $k = 3$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 292
7.65 Fitting G_{WA} to G_{BP} for $k = 3$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 293
7.66 Fitting G_{WA} to G_{BP} for $k = 4$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 294
7.67 Fitting G_{WA} to G_{BP} for $k = 4$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 295
7.68 Fitting G_{WP} to G_{GA} for $k = 2$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 301
7.69 Fitting G_{WP} to G_{GA} for $k = 2$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 302
7.70 Fitting G_{WP} to G_{GA} for $k = 3$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 303
7.71 Fitting G_{WP} to G_{GA} for $k = 3$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 304
7.72 Fitting G_{WP} to G_{GA} for $k = 4$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 305
7.73 Fitting G_{WP} to G_{GA} for $k = 4$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 306
7.74 Fitting G_{WA} to G_{GP} for $k = 2$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 309
7.75 Fitting G_{WA} to G_{GP} for $k = 2$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 310
7.76 Fitting G_{WA} to G_{GP} for $k = 3$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 311
7.77 Fitting G_{WA} to G_{GP} for $k = 3$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 312

7.78 Fitting G_{WA} to G_{GP} for $k = 4$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 313
7.79 Fitting G_{WA} to G_{GP} for $k = 4$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 314
7.80 Fitting G_{WP} to G_{LNA} for $k = 2$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 318
7.81 Fitting G_{WP} to G_{LNA} for $k = 2$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 319
7.82 Fitting G_{WP} to G_{LNA} for $k = 3$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 320
7.83 Fitting G_{WP} to G_{LNA} for $k = 3$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 321
7.84 Fitting G_{WP} to G_{LNA} for $k = 4$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 322
7.85 Fitting G_{WP} to G_{LNA} for $k = 4$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 323
7.86 Fitting G_{WA} to G_{LNP} for $k = 2$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 326
7.87 Fitting G_{WA} to G_{LNP} for $k = 2$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 327
7.88 Fitting G_{WA} to G_{LNP} for $k = 3$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 328
7.89 Fitting G_{WA} to G_{LNP} for $k = 3$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 329
7.90 Fitting G_{WA} to G_{LNP} for $k = 4$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 330
7.91 Fitting G_{WA} to G_{LNP} for $k = 4$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications. 331

8.1 MLEs for G_{WP} for varying k, n, when data is generated from G_{WP} with $B_P = 2$, $\alpha_{WP} = 8$, $\beta_{WP} = -0.02$. In all cases, the n_i are equal, with $\frac{r}{n_1} = \frac{1}{2}$. 359
8.2 MLEs for G_{WP} for varying r, when data is generated from G_{WP} with $k = 3$, $B_P = 2$, $\alpha_{WP} = 8$, $\beta_{WP} = -0.02$. In all cases, $n = 1500$ and $n_i = 500$ for all $1 \leq i \leq 3$, but $\frac{r}{n_1}$ varies. 360

8.3 MLEs for G_{WP} and G_{BP} for $k = 2, 3$ and varying n, when data is generated from G_{BP} with $\tau = 3$, $a = 4$, $\alpha_B = 8$, $\beta_B = -0.02$. In all cases, the n_i are equal, with $\frac{r}{n_1} = \frac{1}{2}$. 361
8.4 MLEs for G_{WP} and G_{BP} for $k = 4$ and varying n, when data is generated from G_{BP} with $\tau = 3$, $a = 4$, $\alpha_B = 8$, $\beta_B = -0.02$. In all cases, the n_i are equal, with $\frac{r}{n_1} = \frac{1}{2}$. 362
8.5 MLEs for G_{WP} and G_{BP} for varying r, when data is generated from G_{BP} with $k = 3$, $\tau = 3$, $a = 4$, $\alpha_B = 8$, $\beta_B = -0.02$. In all cases, $n = 1500$ and $n_i = 500$ for all $1 \leq i \leq 3$, but $\frac{r}{n_1}$ varies. 363

Chapter 1

Introduction

The study of the reliability of electrical components, biological systems or any other item with a life span is a well established and specialised area in the field of statistical investigation, with many texts covering solely this area; see, for instance, Nelson (1982) or Crowder, Kimber, Smith and Sweeting (1991). Engineers and scientists use reliability distributions to model lifetimes of items in order to make inferences concerning such items. They are also widely used in the field of medical statistics to model survival times of people and animals. For example, we might be interested in the time at which 10% of light bulbs fail, the probability that an electrical component fails in a given time interval, or the chances of a certain number of people surviving past age 40. These calculations generally involve using a mathematical model to represent the data set, and assume that the data follows some underlying distribution. In practice, however, the identification of the correct distribution is subject to error, and consequently, the analysis may involve estimating parameters from a mis-specified reliability distribution. This thesis considers various aspects of fitting a mis-specified distribution to a data set.

Our approach thus involves assumptions concerning underlying distributions. From the literature, we see that the Weibull distribution (Weibull, 1951) seems to be the distribution most commonly fitted to survival data. Reasons for this include relative ease of fitting, and the fact that its variety of different shapes provides a good fit to many types of data. Thus, most statisticians tend to choose the Weibull distribution over any other reliability distribution and so will usually fit this to a set of survival data. However, this might not always be the best distribution to use, and many other reliability distributions might prove a better fit. For instance, consider the ball bearings data (Lieblein and Zelen, 1956; Dumonceaux and Antle, 1973), subsequently discussed with proposed corrections in Caroni (2002). Table 1.1 shows the usual $n = 23$ lifetimes (in millions of revolutions), and Dumonceaux and Antle (1973) model this data set using both Weibull and Lognormal distributions. However, they make no attempt to fit any other reliability distribution, although, on the basis of maximised likelihoods, the Burr distribution provides an even better fit for this particular set of data.

This is just one example, of many, where the Weibull distribution is chosen to model a

17.88	28.92	33.00	41.52	42.12
45.60	48.48	51.84	51.96	54.12
55.56	67.80	68.64	68.64	68.88
84.12	93.12	98.64	105.12	105.84
127.92	128.04	173.40		

Table 1.1: Lifetimes (in millions of revolutions) of 23 ball bearings; from Lieblein and Zelen (1956); Dumonceaux and Antle (1973).

data set, when, in fact, another distribution represents it better, and many statistical tests, such as those in Watkins (2001b) and Cain (2002), determine if other distributions such as the Burr and Lognormal represent a data set more adequately than the Weibull. Mackisack and Stillman (1996) also outline some of the possible perils of fitting Weibull distributions to data, and illustrate these with a published data set. We examine the effects of fitting the wrong distribution to a data set, when, in fact, that data has some kind of other underlying distribution. We first outline some necessary background.

1.1 Key references and basic definitions

1.1.1 General definitions

In this section, we summarise basic properties of probability density functions (pdf), cumulative distribution functions (cdf), hazard functions, survivor functions and quantiles, and then give specific examples in the next section. We let the continuous random variable $Y \geq 0$ represent the time to failure. The pdf g of Y defines the probability of a failure in a very small interval. It is given by

$$P(t < Y < t + dt) = \int_t^{t+dt} g(y)\,dy \simeq g(t)\,dt.$$

The cdf is based on cumulating probabilities, and is defined as

$$G(t) = \int_0^t g(y)\,dy.$$

The hazard or failure rate function indicates the proneness to failure of a unit at time t, given successful operation up to this time. It is thus based on conditional probabilities and is defined as

$$h(t) = \frac{g(t)}{1 - G(t)}.$$

Many authors also write

$$h(t) = \frac{g(t)}{S(t)},$$

where $S(t)$ is called the survivor function and gives the probability of an item surviving past time t. The cumulative hazard function is given by

$$H(t) = \int_0^t h(u)\,du,$$

and it can be easily shown that

$$S(t) = \exp\{-H(t)\}. \tag{1.1}$$

When examining lifetime data, we are often interested in estimating a percentile of the lifetime distribution, which is the time at which a specified percentage or proportion of the items fail. We use B_p to denote $100p^{th}$ percentile of a distribution, the time by which a proportion p $(0 < p < 1)$ of the population will fail. If a distribution has cdf $G(y)$, then the $100p^{th}$ percentile is defined by

$$p = G(B_p),$$

and, on rearranging, we have

$$B_p = G^{-1}(p) = Q(p),$$

where the quantile function $Q(p)$ is the inverse of the cdf. These percentiles are usually used to determine a warranty period for the items under consideration, since a balance is needed between the proportion of items failing within the warranty period and the length of the warranty period itself. Generally, we do not want too many items failing during the period, otherwise the company that sells the goods have to face considerable costs in repairing them. The quantile function is also used in the simulation of random variables from a specific distribution, as illustrated below.

The 10^{th} percentile B_{10}, commonly used in reliability analysis, concentrates on early failures and is of particular relevance in electrical and mechanical engineering. Other widely used percentiles include the median or the 50th percentile, which gives the time at which half the observations or items have failed; this is more commonly used in medical statistics.

1.1.2 Particular distributions

There are numerous descriptions of different reliability distribution functions and probability density functions, along with their properties. For example, see Ansell and Phillips (1994), Nelson (1982) and Leitch (1995); Richards and McDonald (1987) also summarise the relationships that exist between various distributions. Throughout this thesis, we will require the following distribution functions.

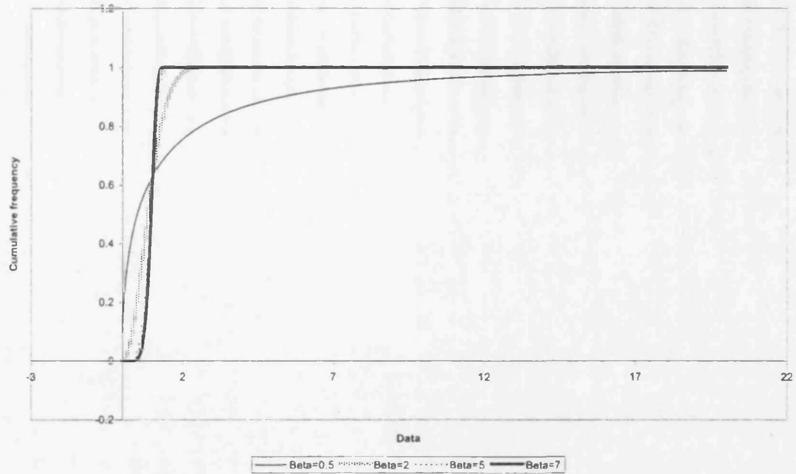


Figure 1.1: The function G_w for varying β.

The Weibull distribution

The Weibull distribution is the most common and well known of all the reliability distribution functions, and reasons for this include the fact that it is a relatively robust distribution that can be fitted to data very easily without too many numerical problems. It was introduced by Weibull (1951), and has pdf defined by

$$g_w(y;\beta,\theta) = \frac{\beta y^{\beta-1}}{\theta^\beta}\exp\left\{-\left(\frac{y}{\theta}\right)^\beta\right\} \quad \text{for } y>0, \tag{1.2}$$

and cdf given by

$$G_w(y;\beta,\theta) = 1-\exp\left\{-\left(\frac{y}{\theta}\right)^\beta\right\} \quad \text{for } y>0. \tag{1.3}$$

The Weibull distribution has two positive parameters; β is the shape parameter and θ the scale parameter. Figure 1.1 shows how varying β affects the shape of the Weibull cdf; larger values of β correspond to steeper distribution functions that tend to 1 more rapidly. The hazard function for the Weibull distribution is given by

$$h_w(t;\beta,\theta) = \frac{\beta t^{\beta-1}}{\theta^\beta},$$

and, by integration or via (1.1), the cumulative hazard function takes the form

$$H_w(t;\beta,\theta) = \left(\frac{t}{\theta}\right)^\beta.$$

The quantile function for the Weibull distribution is defined as

$$B_{w,p} = G_w^{-1}(p) = \theta \{-\ln(1-p)\}^{\frac{1}{\beta}}.$$

The Burr XII distribution

The Burr XII distribution (from now on abbreviated to Burr distribution), was first introduced by Burr (1942), and has received considerable attention by Wingo (1983) and Tadikamalla (1980). It has pdf given by

$$g_b(y;\tau,\alpha,\phi) = \frac{\alpha\tau y^{\tau-1}}{\phi^\tau}\left\{1+\left(\frac{y}{\phi}\right)^\tau\right\}^{-(\alpha+1)} \qquad \text{for } y>0,$$

and cdf

$$G_b(y;\tau,\alpha,\phi) = 1-\left\{1+\left(\frac{y}{\phi}\right)^\tau\right\}^{-\alpha} \qquad \text{for } y>0, \tag{1.4}$$

where the positive parameters α and τ control the shape of the distribution, and $\phi > 0$ is a scale parameter. We see that it is a three parameter distribution, but can be reduced to just two parameters, α and τ, by rescaling the data by ϕ. Thus, the two parameter Burr distribution is a special case of the Burr distribution with ϕ equal to 1; it is often convenient to derive results for this special case and then generalise to the three parameter Burr distribution. The effects of changing the shape parameters from the Burr distribution can be seen by examining plots of G_b for varying α and τ. Since ϕ represents a scale parameter, it will not affect the shape in any way. Figure 1.2 and Figure 1.3 show distribution functions for $\phi = 1$. Figure 1.2 shows the effect of changing τ when $\alpha = 1$, whilst Figure 1.3 gives a similar comparison for varying α, with $\tau = 1$. Increasing τ produces a steeper distribution function, so that most of the probability is contained at the smaller data values. The function also tends to 1 much more quickly with large values of τ. A similar pattern is observed when α is allowed to vary, and we see, for larger values of α, a much steeper distribution function that tends to 1 extremely quickly. The hazard function for the Burr distribution is given by

$$h_b(t;\tau,\alpha,\phi) = \frac{\alpha\tau t^{\tau-1}}{\phi^\tau}\left\{1+\left(\frac{t}{\phi}\right)^\tau\right\}^{-1},$$

with cumulative hazard function

$$H_b(t;\tau,\alpha,\phi) = \alpha \ln\left\{1+\left(\frac{t}{\phi}\right)^\tau\right\}.$$

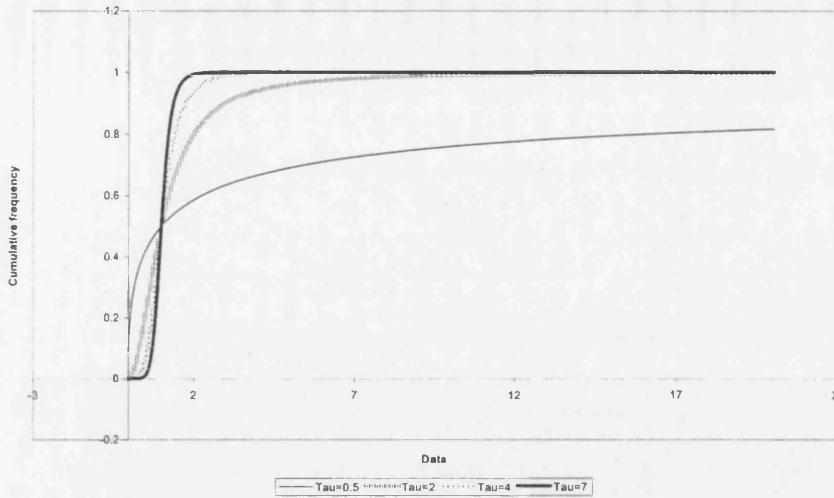


Figure 1.2: The function G_b for $\alpha = 1$ and varying τ.

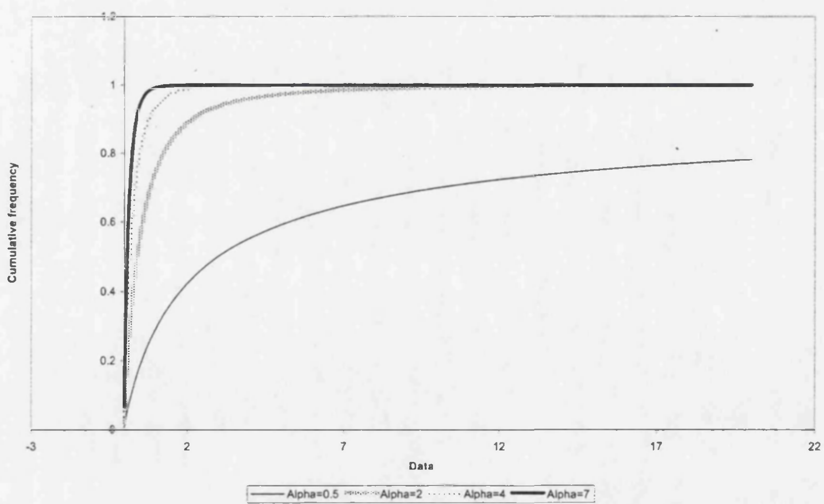


Figure 1.3: The function G_b for $\tau = 1$ and varying α.

The quantile function for the Burr distribution is

$$B_{b,p} = G_b^{-1}(p) = \phi\left\{(1-p)^{-\frac{1}{\alpha}} - 1\right\}^{\frac{1}{\tau}}.$$

The Burr distribution has the important property of including the Weibull model as a limiting case. This property, and the consequences of it, will be considered in detail in Chapter 2.

The Gamma distribution

The Gamma distribution is not as widely used as the Weibull and Burr distributions, since the form of its hazard function makes it less suitable to use in some cases; see Nelson (1982) and below. It has pdf defined by

$$g_g(y;\tau,\alpha) = \frac{y^{\tau-1}\exp\left(-\frac{y}{\alpha}\right)}{\alpha^{\tau}\Gamma(\tau)} \qquad \text{for } y > 0, \tag{1.5}$$

where $\Gamma(.)$ is the gamma function given by

$$\Gamma(\tau) = \int_0^{\infty} z^{\tau-1}\exp(-z)\,dz. \tag{1.6}$$

By integration, we see that the cdf of the Gamma distribution is

$$G_g(y;\tau,\alpha) = \frac{\Gamma\left(\frac{y}{\alpha},\tau\right)}{\Gamma(\tau)} \qquad \text{for } y > 0,$$

where

$$\Gamma(z,\tau) = \int_0^{z} u^{\tau-1}\exp(-u)\,du \tag{1.7}$$

denotes the incomplete gamma function. Here, α represents a scale parameter, whilst τ is the shape parameter. Figure 1.4 shows that, as τ decreases, the distribution function generally becomes much steeper for smaller data values. The effects of increasing τ seem to shift the distribution curve along the horizontal axis. The hazard function for the Gamma distribution is given by

$$h_g(t;\tau,\alpha) = \frac{\left(\frac{t}{\alpha}\right)^{\tau-1}\exp\left(-\frac{t}{\alpha}\right)}{\alpha\left\{\Gamma(\tau) - \Gamma\left(\frac{t}{\alpha},\tau\right)\right\}},$$

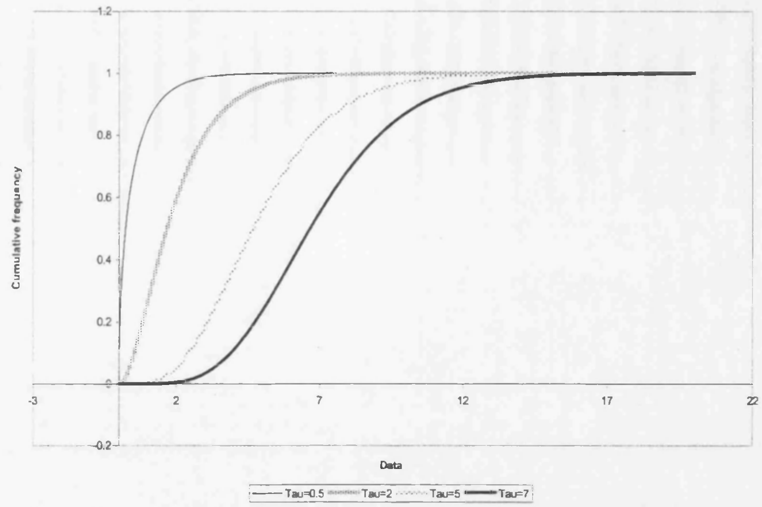


Figure 1.4: The function G_g for varying τ.

and, therefore,

$$H_g\left(t;\tau,\alpha\right) = -\ln\left\{1 - \frac{\Gamma\left(\frac{t}{\alpha},\tau\right)}{\Gamma\left(\tau\right)}\right\}.$$

The fact that these functions cannot be expressed in closed form is a considerable practical barrier to the use of this distribution. A further related disadvantage is that the quantile function must be obtained numerically, since we cannot explicitly write down the inverse of G_g.

The Lognormal distribution

The Lognormal distribution empirically fits many types of data adequately, because it has a great variety of shapes; see Nelson (1982). The distribution is often used when the range of data is extremely large, as is sometimes the case for data on metal fatigue and electrical insulation life, and, away from a reliability setting, economic data and responses of biological material to stimulus. This distribution function is closely related to the Normal, in that, if Y follows a Lognormal distribution, then $\ln(Y)$ has a Normal distribution with mean μ and standard deviation σ. Thus, the Lognormal cdf is defined to be

$$G_{\ln}\left(y;\mu,\sigma\right) = \Phi\left(\frac{\ln y - \mu}{\sigma}\right) \quad \text{for } y \geq 0, \tag{1.8}$$

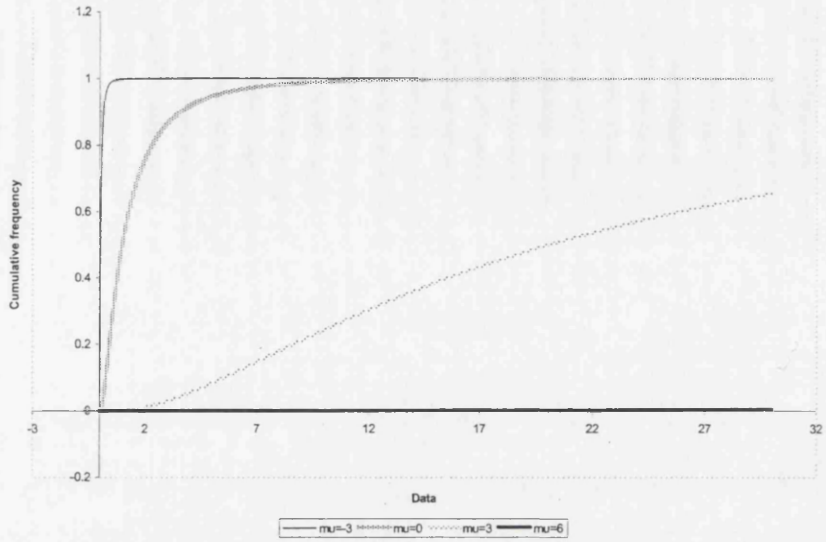


Figure 1.5: The function $G_{\ln}$ for varying μ.

where Φ is the usual standard Normal cdf; from this, the pdf is

$$g_{\ln}(y;\mu,\sigma) = \frac{1}{\sigma y\sqrt{2\pi}} \exp\left(\frac{-(\ln y-\mu)^2}{2\sigma^2}\right) \qquad \text{for } y \geq 0, \tag{1.9}$$

where the parameter μ is the mean of the log of life and may take any value; again, see Nelson (1982). The parameter σ is called the log standard deviation and must be positive; it is the standard deviation of the log of life. The value of σ determines the shape of the distribution, whilst μ determines the 50% point and the spread. Figures 1.5 and 1.6 illustrate how the parameters of the Lognormal distribution affect the shape of the distribution function. We see that increasing μ (keeping σ fixed at 1), even by just a small amount, flattens the cdf and prolongs the period it takes for the function to tend to one. Increasing σ (and keeping μ fixed at 1) has a similar effect but there are also high probabilities associated with smaller data values. The hazard function for the Lognormal distribution is given by

$$h_{\ln}(t;\mu,\sigma) = \frac{\frac{1}{\sigma t\sqrt{2\pi}} \exp\left(\frac{-(\ln t-\mu)^2}{2\sigma^2}\right)}{1-\Phi\left(\frac{\ln t-\mu}{\sigma}\right)},$$

with cumulative hazard function of the form

$$H_{\ln}(t;\mu,\sigma) = -\ln\left\{1-\Phi\left(\frac{\ln t-\mu}{\sigma}\right)\right\},$$

on using (1.1). We see that the cumulative hazard function involves a function that must be evaluated numerically.

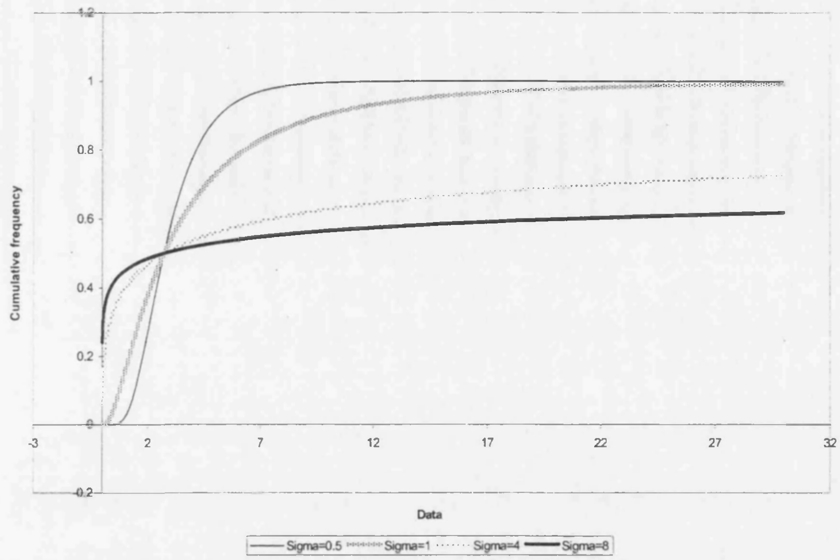


Figure 1.6: The function $G_{\ln}$ for varying σ.

1.1.3 Censored data

More often than not, observed reliability data have the complicating feature of containing censored values. Censoring occurs when the exact lifetime of an item is not observed, but, for example, is only known to exceed a certain time, possibly the lifetime of another item that has failed. Such a situation arises when an experiment is terminated before all items fail, leading to an incomplete set of data. Reasons for termination may include the fact that waiting for all items to fail may take several days, weeks or even longer. Thus, we may observe the first few failures, but then stop the experiment as the other items continue to function. If we do observe all the failures, then we have a complete data set, which contains more information than a data set of the same size that has undergone some form of censoring. There are many ways in which observations can be censored. We list some of these below, but just consider the first two in this thesis.

Type I censoring

Items are said to have undergone a type I censoring regime if the experiment is terminated after some specified fixed time y_c, also called the stopping time. As a result, the number of observed failures N $(0 \leq N \leq n)$ is a random variable, and the remaining $(n - N)$ items are censored at the stopping time. This type of censoring has the practical advantages of known experimental duration, but the statistical disadvantage of prior uncertainty in the precise number of failure times available for analysis.

Type II censoring

Type II censoring occurs when n items are tested, and the experiment is halted after the r^{th} $(r \leq n)$ failure. Thus, the remaining $(n-r)$ items will yield censored times in service. Type II censoring has the statistical advantage of ensuring a precise number of failure times for analysis, but experimental duration is not known precisely in advance, and it is possible for an experiment to continue for long periods until r failures are observed.

Interval censoring

This occurs when units are put on test, but only checked for failure every hour, every day, or at other specified time points. If a unit is found to have failed during the last period, then we do not know precisely its time to failure, and can only estimate the failure time as a point in this period.

Progressive censoring

Progressive censoring occurs when test units are removed at different stages during the experiment for various reasons. These may include the removal of items for more thorough inspection, or use elsewhere. There are many different methods of specifying the removal times and the number of items withdrawn at each removal time. For instance, Tse, Yang and Yuen (2000), assume a progressive censoring regime with Binomial removals.

1.1.4 Mathematical functions

We have already seen that the basic reliability distributions introduce the gamma function. Our analysis will require the consideration of other related functions, and we now list these; we refer to Abramowitz and Stegun (1972) for further details.

The gamma and incomplete gamma functions

The gamma function, at (1.6), satisfies the following recursive relationship

$$\Gamma(r+1) = r\Gamma(r),$$

and for n, an integer, we write

$$\Gamma(n+1) = n!$$

The incomplete gamma function arises in connection with type I and II censoring, where we may also use

$$\Gamma(z,a) = z^a U_1(z,a), \tag{1.10}$$

where

$$U_i(z,a) = \sum_{n=0}^{\infty} \frac{(-z)^n}{(n+a)^i n!}$$

with the property that

$$\frac{\partial U_i}{\partial a} = -iU_{i+1}.$$

The incomplete gamma function also has an important recurrence relationship given by

$$\Gamma(z, a+1) = a\Gamma(z,a) - z^a \exp(-z). \tag{1.11}$$

Derivatives of the gamma function

The psi function is defined as the derivative of the log of the gamma function, given by

$$\Psi(z) = \frac{d \ln \Gamma(z)}{dz} = \frac{\Gamma'(z)}{\Gamma(z)},$$

and satisfies the recursive relationship

$$\Psi(z+1) = \psi(z) + z^{-1}. \tag{1.12}$$

We write

$$\Psi(1) = -\gamma,$$

where $\gamma = 0.57721\cdots$ is known as Euler's constant. We can also differentiate the psi function to obtain the trigamma function defined as

$$\frac{d\Psi(z)}{dz} = \frac{\Gamma(z)\Gamma''(z) - \Gamma'(z)^2}{\{\Gamma(z)\}^2}.$$

Further derivatives of the gamma function can be found in Abramowitz and Stegun (1972). In particular, we have

$$\Psi^n(z) = (-1)^{n+1} \int_0^{\infty} \frac{u^n \exp(-zu)}{1 - \exp(-u)} du.$$

Beta and incomplete Beta functions

The Beta function is defined in terms of gamma functions, and is given by

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt = \int_0^{\infty} \frac{t^{a-1}}{(1+t)^{a+b}} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}; \tag{1.13}$$

recursive relationships also exist for this function. In particular, we have

$$abB(a,b) = b(a+b)B(a+1,b) = a(a+b)B(a,b+1). \tag{1.14}$$

The incomplete beta function is defined as

$$B_z(a,b) = \int_0^z t^{a-1}(1-t)^{b-1}\,dt. \tag{1.15}$$

Hypergeometric functions

These arise when considering type I censoring, but also appear in results for joint expectations of random variables that have undergone type II censoring. The generalised hypergeometric function is defined as

$$F_{p,q}(\{a_1,a_2,...,a_p\},\{b_1,b_2,...,b_q\};z) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k \cdots (a_p)_k}{(b_1)_k (b_2)_k \cdots (b_q)_k} \frac{z^k}{k!},$$

where $(x)_r$ is called Pochhammers symbol, given by

$$(x)_r = \frac{\Gamma(x+r)}{\Gamma(x)}.$$

For later use, we note that

$$(x)_r = x(x+1)_{r-1},$$

and

$$(x+1)_r = \frac{(x+r)}{x}(x)_r.$$

We can also express the incomplete Beta function in terms of hypergeometric expressions, and note that

$$\begin{aligned} B_z(a,b) &= a^{-1}z^a F_{2,1}(\{a,1-b\};\{a+1\};z) \\ &= a^{-1}z^a \sum_{n=0}^{\infty} \frac{(a)_n(1-b)_n}{(a+1)_n}\frac{z^n}{n!}. \end{aligned} \tag{1.16}$$

See Slater (1966) for other results.

Order statistics

To analyse observations that have undergone type II censoring requires ordering the data set from the smallest to the largest. The i^{th} largest item in a sample of size n is usually denoted by

$$y_{(i:n)},$$

and, by ordering the data set, we will always have

$$y_{(1:n)} \leq y_{(2:n)} \leq \cdots \leq y_{(n:n)}.$$

See David (1981) for results on order statistics and their properties.

1.2 Numerical and computational aspects

In order to check theoretical results developed, this thesis will rely heavily on programming and computational packages such as SAS and Mathematica; we refer to Der and Everitt (2002) and Wolfram (1988), respectively, for further details. We use Excel for simpler calculations and graphs.

Throughout this thesis, we will need to maximise functions $l(\pi)$ based on likelihoods with respect to model parameters π; in most cases, only limited analytical progress is possible, so that maximisation must be performed numerically. Our approach is as follows: we discount the possibility of multiple stationary points, and regard the maximisation of $l(\pi)$ as equivalent to finding the roots of $\frac{\partial l}{\partial \pi}$. We generally locate these roots using the Newton-Raphson computational procedure. This is the quickest and probably most straightforward method to program in SAS, although other procedures, such as the Bisection, Iterative and Secant methods, can be employed; see Kennedy and Gentle (1980) for further details. We will provide additional references as necessary; for instance, when we need to employ more sophisticated algorithms. In certain situations, any numerical method may fail to locate the roots of $\frac{\partial l}{\partial \pi}$; we only accept proposed solutions $\widetilde{\pi}$ for which each element of $\frac{\partial l}{\partial \pi}$ is less than 10^{-9} in absolute value; this cut-off value is regarded as sufficiently close to zero to indicate convergence, and in simulations, we usually observed values considerably smaller than this (generally less than 10^{-15}).

1.3 Sample procedures

We list some standard sample procedures used throughout this thesis below. Most will be used to compare models based on both true and mis-specified distributions for data.

1.3.1 Hazard and cumulative hazard plots

As mentioned above, the hazard function indicates the proneness to failure of a unit at time t, given successful operation up to this time. From the literature, we see that the empirical hazard and cumulative hazard functions can be calculated in different ways. Crowder, Kimber, Smith and Sweeting (1991) suggest that the empirical survivor function $\widehat{S}(t)$, defined as

$$\widehat{S}(t) = \frac{\text{number of observations greater than or equal to } t}{n}$$

should first be calculated. The empirical cumulative hazard function is then defined as

$$\widehat{H}(t) = -\ln \widehat{S}(t).$$

On the other hand, Newton (1991) counted the number of observations greater than or equal to each data point, and then used this number to estimate $h(t)$. Cumulative values then estimate $H(t)$. As both methods give very similar results, we use the method given in Crowder, Kimber, Smith and Sweeting (1991). We note, but do not use, the Kaplan-Meier estimate, which provides another way of estimating the survivor function, and is particularly useful when the data set has undergone some form of censoring.

1.3.2 Kolmogorov-Smirnov distances

These procedures are used when a theoretical distribution, usually with unknown parameters replaced by maximum likelihood estimates, are used to model data. The distance between the fitted theoretical cumulative distribution and the empirical distribution functions is then calculated to produce the Kolmogorov-Smirnov test statistic. We denote this by D, with appropriate subscripts for Weibull, Burr, Gamma and Lognormal distributions. Since the test statistics is based on the maximum of D, then the larger this statistic is, the worse the fit between the theoretical distribution and data set. If D is significantly large, then we reject the hypothesis that the underlying data is adequately modelled using the distribution specified. For further details on this distance, we refer to Lawless (1982).

1.3.3 Data simulation

Since this thesis will involve fitting the Weibull distribution to data with other underlying distribution functions, then we will need to simulate data from such distributions. To generate data y_i from a distribution with cdf G, we compute

$$y_i = Q(u_i) = G^{-1}(u_i),$$

where the u_i are independently and Uniformly distributed on $[0,1]$. Thus, to simulate a set of data from a Weibull distribution with known, specified parameters β and θ, we use (1.3) and compute

$$y_i = \theta\left(-\ln\left[1-u_i\right]\right)^{\frac{1}{\beta}}. \tag{1.17}$$

For the Burr, we use (1.4) and calculate

$$y_i = \phi \left\{ (1 - u_i)^{-\frac{1}{\alpha}} - 1^{\frac{1}{\tau}} \right\}.$$

As previously noted, quantile functions for the Gamma and Lognormal distributions cannot be expressed explicitly. Consequently, we rely on the quantile functions defined in SAS to compute a set of data from both distributions. For the Gamma distribution, the RANGAM command in SAS produces random numbers from G_g with specified shape parameter τ and scale parameter $\alpha = 1$. Thus, we use $\alpha\times$RANGAM[τ] to simulate a set of data with general shape and scale parameters. To simulate data from a Lognormal distribution, we again exploit the link between this distribution and the Normal; see (1.8) and (1.9). Thus, if we simulate t_i from $N(\mu, \sigma^2)$, then $y_i = \exp(t_i)$ will be a random sample from G_{ln}.

1.4 Structure of thesis

The remaining chapters of this thesis are as follows. Chapter 2 examines the effects of fitting the Weibull distribution to data from the Weibull, Burr, Gamma and Lognormal distributions, and assessing the goodness of fit. Chapter 3 considers theoretical counterparts to the parameter estimates from the mis-specified Weibull distribution; this involves entropy functions, and we also derive the variance covariance matrix of the estimates under mis-specification. Chapter 4 derives similar results for data that has undergone type I and type II censoring regimes. In Chapters 5, 6, 7 and 8, we extend the ideas of mis-specification to deal with accelerated data sets. Finally, Chapter 9 finishes with a summary of our work, together with a brief outline of any future research.

Chapter 2

Maximum Likelihood: Some Practical Considerations

2.1 Introduction

As previously noted, the process of modelling a set of reliability data usually involves three main steps. These are

- Identifying a suitable model for the data.
- Estimating the parameters contained in the model.
- Assessing the goodness of fit to ascertain if an adequate fit has been achieved. If not, a different model is chosen and the three steps are repeated.

Numerous methods have been devised for selecting models, obtaining parameter estimates, and for assessing the goodness of fit of the proposed model. Model selection is based on graphical methods; see Chapter 1 above. Methods for fitting a distribution to a data set include the widely used maximum likelihood (ML) approach, where we obtain maximum likelihood estimates (MLEs) for the true parameter values. We denote these by a caret, so, for example, the MLE for α is $\widehat{\alpha}$. For more details on MLE, and on the asymptotic properties of the MLEs, we refer to Cox and Hinkley (1974). Other approaches include least squares and the method of moments. To assess the adequacy of the fitted distribution usually involves graphical techniques such as plots of the fitted cdf against the empirical cdf obtained from the data set. The agreement between the two functions can then be summarised by measures of functional distances, such as the Kolmogorov-Smirnov distance. We may also consider examining plots of theoretical and empirical hazard and cumulative hazard functions, and observe whether any large discrepancies occur between the fitted distribution and data set in this case. Other plots, such as Kaplan-Meier plots, can consider distances between sample and theoretical cdfs for data containing censored observations. For further details on such sample procedures, we refer to Chapter 1.

This thesis is concerned with the effects of mis-specification and choosing an incorrect distribution function to represent a particular data set. Most statistical analyses implicitly assume that the underlying model is correctly specified, and the question of model choice has received relatively little attention. However, Marshall, Meza and Olkin (2001) address this problem briefly, and examine maximised likelihoods and Kolmogorov-Smirnov distances between true and mis-specified distributions, when a range of underlying models are used. White (1982) looks at the properties of MLEs of mis-specified models, whilst Hutton and Monaghan (2002) examine mis-specification of accelerated life and proportional hazard models for survival data. We begin to consider the problem of model mis-specification in this chapter, but first summarise the theory for fitting the various distributions given in Chapter 1. Thus, we consider fitting the Weibull, Burr, Gamma and Lognormal distributions to data using ML techniques. We then address the problem of mis-specification, and fitting the incorrect distribution function. From Chapter 1, and the references given there, we saw how the Weibull distribution was the most common reliability distribution fitted to data, and, in some cases, is wrongly chosen to represent a particular data set although another distribution provides a better fit. Thus, due to this wide use, we always take this as our mis-specified model, and look at the effects of fitting it to data with an underlying Burr, Gamma and Lognormal model. Of course, there are many other possible variants we could consider, since any other distribution could be mis-specified, so that, for example, we could try to fit the Lognormal distribution to data from an underlying Burr model. Such scenarios will be considered elsewhere. We first derive results for complete data, and give corresponding results on type I and type II censoring in Chapter 4.

2.2 Fitting G_w

Likelihood may be defined as the joint pdf based on a specific distribution at the observed sample points. The MLEs of the parameters in the specified distribution are then the values that maximise the likelihood function. Equivalently, (since the likelihood function for independent observations involves taking products of terms) we usually consider the natural logarithm of the likelihood function, and obtain the maximum of the log-likelihood. This approach converts products to sums, which are easier to manipulate when we consider locating maximum turning points. Since ln is a monotonic increasing function, the log-likelihood has the same stationary points as the original likelihood. The likelihood for the Weibull distribution is given in Cohen (1965), and expectations of terms that appear in the log-likelihood and score functions are summarised in, for example, Watkins (1998). We briefly outline the important functions below for a complete data set $y_1, y_2, ..., y_n$. The likelihood and log-likelihood based on (1.3) are

$$L_w(y_1, y_2, .., y_n; \beta, \theta) = \prod_{i=1}^{n} g_w(y_i; \beta, \theta) = \left(\frac{\beta}{\theta^\beta}\right)^n \left(\prod_{i=1}^{n} y_i\right)^{\beta-1} \exp\left(-\sum_{i=1}^{n} \frac{y_i^\beta}{\theta^\beta}\right)$$

and

$$l_w = n \ln \beta - n\beta \ln \theta + (\beta - 1) \sum_{i=1}^{n} \ln y_i - \theta^{-\beta} \sum_{i=1}^{n} y_i^{\beta}. \tag{2.1}$$

We will find it convenient to define

$$S_e = \sum_{i=1}^{n} \ln y_i$$

and

$$S_j(r) = \sum_{i=1}^{n} y_i^r (\ln y_i)^j,$$

for real $r > 0$ and integer $j \geq 0$, taking $0^0 = 1$ if necessary; note that for $j \geq 1$

$$S_j(r) = \frac{d^j S_0(r)}{dr^j} = \frac{dS_{j-1}(r)}{dr}.$$

Then (2.1) can be written

$$l_w = n \ln \beta - n\beta \ln \theta + (\beta - 1) S_e - \theta^{-\beta} S_0(\beta);$$

MLEs are obtained by maximising l_w or, equivalently, finding the roots of the score function, based on the two partial derivatives given by

$$\frac{\partial l_w}{\partial \beta} = n\beta^{-1} - n \ln \theta + S_e + \theta^{-\beta} \ln \theta S_0(\beta) - \theta^{-\beta} S_1(\beta), \tag{2.2}$$

$$\frac{\partial l_w}{\partial \theta} = -n\beta\theta^{-1} + \beta\theta^{-(\beta+1)} S_0(\beta). \tag{2.3}$$

There are no analytic expressions for these roots. However, we note that if we equate (2.3) to zero, then we can express θ in terms of the data and the shape parameter β; we obtain

$$\theta = \left\{ \frac{S_0(\beta)}{n} \right\}^{\frac{1}{\beta}}. \tag{2.4}$$

By substituting (2.4) into (2.1) and (2.2), we obtain our profile log-likelihood

$$l_w^* = n \ln \beta - n \ln S_0(\beta) + n(\ln n - 1) + (\beta - 1) S_e \tag{2.5}$$

with first derivative

$$\frac{dl_w^*}{d\beta} = n\beta^{-1} + S_e - \frac{nS_1(\beta)}{S_0(\beta)}. \tag{2.6}$$

Iteration	$\widehat{\beta}$	$\frac{dl^*_w}{d\beta}$	$\frac{d^2 l^*_w}{d\beta^2}$
1	1.989760	1.0796821	-9.989062
2	2.097846	0.0390496	-9.281992
3	2.102053	0.0000541	-9.256282
4	2.102059	1.045×10^{-10}	-9.256246
5	2.102059	0	-9.256246

Table 2.1: Summary of iterations for fitting G_w to the ball bearings data in Table 1.1.

Locating the root of (2.6) is now considerably easier, since it is a univariate function of β. Our MLE for θ is then obtained by substituting the MLE for β into (2.4). As noted in Chapter 1, many methods have been established to locate roots of functions like (2.6). We use the Newton-Raphson approach, which requires the derivative

$$\frac{d^2 l^*_w}{d\beta^2} = -n\beta^{-2} - n\left\{\frac{S_0(\beta)\,S_2(\beta) - S_1(\beta)^2}{S_0(\beta)^2}\right\},$$

and an initial starting value. This starting value should be close to $\widehat{\beta}$, otherwise the Newton-Raphson process may fail to converge; Farnum and Booth (1997) discuss initial starting values; for a complete data set, the quantity

$$V = \ln\left(y_{(n:n)}\right) - n^{-1}S_e$$

measures variation in data, and a good starting point for locating $\widehat{\beta}$ is $2V^{-1}$, since convergence is then guaranteed. We locate the MLEs for G_w for the ball bearings data given in Table 1.1. Here, $n = 23$, $S_e = 95.4605$ and $y_{(n:n)} = 173.40$, so we have $V = 1.0051$. Using $2V^{-1} = 1.9898$ as a starting value for $\widehat{\beta}$, the Newton-Raphson process converged in just five iterations to $\widehat{\beta} = 2.1021$ for which $\widehat{\theta} = 81.8783$. We summarise these iterations in Table 2.1, and observe how the profile score function for β tends to zero very quickly; note that the second derivative is negative, indicating that we have found a maximum of l^*_w. This maximum value is $l^*_w = -113.6913$.

2.3 Fitting G_b

We now fit the Burr distribution to $y_1, y_2, ..., y_n$. This requires the likelihood

$$L_b(y_1, y_2, ..., y_n; \tau, \alpha, \phi) = \prod_{i=1}^{n} g_b(y_i; \tau, \alpha, \phi) = \left(\frac{\alpha\tau}{\phi^\tau}\right)^n \prod_{i=1}^{n} y_i^{\tau-1}\left\{1 + \left(\frac{y_i}{\phi}\right)^\tau\right\}^{-(\alpha+1)},$$

from which the log-likelihood is

$$l_b = n\ln\alpha + n\ln\tau - n\tau\ln\phi + (\tau-1)S_e - (\alpha+1)\sum_{i=1}^{n}\ln\left\{1+\left(\frac{y_i}{\phi}\right)^{\tau}\right\}. \tag{2.7}$$

We will find it convenient to write

$$T = \sum_{i=1}^{n}\ln\left\{1+\left(\frac{y_i}{\phi}\right)^{\tau}\right\},$$

with partial derivatives given by

$$\begin{aligned}
\frac{\partial T}{\partial\tau} &= T_{1,0} = \sum_{i=1}^{n}\frac{\left(\frac{y_i}{\phi}\right)^{\tau}\ln\left(\frac{y_i}{\phi}\right)}{1+\left(\frac{y_i}{\phi}\right)^{\tau}},\\
\frac{\partial T_{1,0}}{\partial\tau} &= T_{2,0} = \sum_{i=1}^{n}\frac{\left(\frac{y_i}{\phi}\right)^{\tau}\left\{\ln\left(\frac{y_i}{\phi}\right)\right\}^2}{\left\{1+\left(\frac{y_i}{\phi}\right)^{\tau}\right\}^2},
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial T}{\partial\phi} &= T_{0,1} = \frac{-\tau}{\phi}\sum_{i=1}^{n}\frac{\left(\frac{y_i}{\phi}\right)^{\tau}}{1+\left(\frac{y_i}{\phi}\right)^{\tau}},\\
\frac{\partial T_{0,1}}{\partial\phi} &= T_{0,2} = \frac{\tau^2}{\phi^2}\sum_{i=1}^{n}\frac{\left(\frac{y_i}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{y_i}{\phi}\right)^{\tau}\right\}^2} + \frac{\tau}{\phi^2}\sum_{i=1}^{n}\frac{\left(\frac{y_i}{\phi}\right)^{\tau}}{1+\left(\frac{y_i}{\phi}\right)^{\tau}},\\
T_{1,1} &= \frac{\partial T_{0,1}}{\partial\tau} = \frac{-1}{\phi}\sum_{i=1}^{n}\frac{\left(\frac{y_i}{\phi}\right)^{\tau}}{1+\left(\frac{y_i}{\phi}\right)^{\tau}} - \frac{\tau}{\phi}\sum_{i=1}^{n}\frac{\left(\frac{y_i}{\phi}\right)^{\tau}\ln\left(\frac{y_i}{\phi}\right)}{\left\{1+\left(\frac{y_i}{\phi}\right)^{\tau}\right\}^2}.
\end{aligned}$$

The subscripts in this notation indicate differentiation with respect to τ and ϕ; note that Watkins (1999) used a slightly different convention, linking subscripts to powers of terms in summations. By differentiating (2.7) with respect to the three parameters, we obtain the score function; its components are

$$\begin{aligned}
\frac{\partial l_b}{\partial\tau} &= n\tau^{-1} - n\ln\phi + S_e - (\alpha+1)\,T_{1,0},\\
\frac{\partial l_b}{\partial\alpha} &= n\alpha^{-1} - T,\\
\frac{\partial l_b}{\partial\phi} &= -n\tau\phi^{-1} - (\alpha+1)\,T_{0,1},
\end{aligned} \tag{2.8}$$

while the six second partial derivatives are given by

$$\left.\begin{aligned}
\frac{\partial^2 l_b}{\partial\tau^2} &= -n\tau^{-2} - (\alpha+1)\,T_{2,0}\\
\frac{\partial^2 l_b}{\partial\alpha^2} &= -n\alpha^{-2}\\
\frac{\partial^2 l_b}{\partial\phi^2} &= n\tau\phi^{-2} - (\alpha+1)\,T_{0,2}\\
\frac{\partial^2 l_b}{\partial\tau\partial\alpha} &= -T_{1,0}\\
\frac{\partial^2 l_b}{\partial\phi\partial\alpha} &= -T_{0,1}\\
\frac{\partial^2 l_b}{\partial\tau\partial\phi} &= -n\phi^{-1} - (\alpha+1)\,T_{1,1}
\end{aligned}\right\} \tag{2.9}$$

The components of the score function have no analytic roots, so numerical techniques must be used to locate them. Some simplifications occur on equating (2.8) to zero, yielding

$$\alpha = nT^{-1}. \tag{2.10}$$

However, maximising the resulting profile log-likelihood is problematical; for instance, the Newton-Raphson process is sensitive to starting values, and iterations diverge if inappropriate initial values are chosen. This divergence is linked to the fact that the Burr has a limiting Weibull distribution, and, if this distribution provides an improved fit over the Burr to a data set, then the Burr cannot be fitted by ML. We also note that, unlike the Weibull, which seems a relatively robust distribution to fit, we may encounter further numerical problems (such as the selection of initial values and speed of convergence) when fitting the Burr to data. We therefore adopt the method outlined by Watkins (1999) to fit the three parameter Burr distribution, and briefly outline the steps in the next section.

2.3.1 The limiting distribution and Δ

Tadikamalla (1980) showed that the Weibull distribution is a limiting case of the Burr distribution. We reparameterise (1.4), using

$$\lambda = \phi^\tau$$

so that

$$G_b(y;\tau,\alpha,\lambda) = 1 - \left\{1 + \left(\frac{y^\tau}{\lambda}\right)\right\}^{-\alpha} = 1 - \left[\left\{1 + \left(\frac{y^\tau}{\lambda}\right)\right\}^{\lambda}\right]^{\frac{-\alpha}{\lambda}}.$$

If we let $\alpha, \lambda \to \infty$, so that $\frac{\alpha}{\lambda} = \frac{\alpha}{\phi^\tau}$ remains constant, then the distribution function becomes

$$1 - \left[\exp\left(y^\tau\right)\right]^{\frac{-\alpha}{\lambda}},$$

on recognising the limiting form of the exponential function, given by

$$\lim_{m\to\infty}\left(1+\frac{z}{m}\right)^m = \exp(z).$$

Then, by rearranging, we see that the limiting form of the Burr is

$$G_b(y;\tau,\alpha,\lambda) = 1-\exp\left(-\frac{\alpha y^\tau}{\lambda}\right) = 1-\exp\left\{-\left(\frac{y}{\alpha^{-\frac{1}{\tau}}\phi}\right)^\tau\right\},$$

which is G_w with shape parameter τ, and scale parameter $\alpha^{\frac{-1}{\tau}}\phi$.

This results suggests that, if a Weibull distribution provides a better fit to a set of data than the Burr, and we try to fit the Burr distribution, then we will observe ϕ and α becoming very large, rather than converging on finite numbers. Such observations prompted Watkins (2001b) to derive a function to determine which of the Burr or Weibull provides a better fit to a set of data. We mention it here since we require it for our simulation studies, and consider it in detail later. For complete data, the function Δ is given by

$$\frac{S_0\left(2\widehat{\beta}\right)}{2} - \frac{S_0\left(\widehat{\beta}\right)^2}{n};$$

if $\Delta > 0$, then the Burr distribution provides the better fit, in terms of maximised log-likelihood, to a set of data, whilst a negative Δ suggests that the Weibull distribution models the data set more appropriately, and in fact, the Burr distribution cannot be fitted by ML in this case. Other methods have been derived to assess whether a limiting distribution (sometimes called an embedded distribution) should be chosen in favour of an underlying model. In particular, Crowder and Kimber (1997) derive a score test to determine whether the multivariate Burr should be used in preference to the multivariate Weibull model. Cheng and Iles (1990) also provide formal tests to decide if the embedded model should be fitted instead of the three parameter distribution. They do this for a general distribution that contains an embedded model, but also consider specific cases as examples. In particular, they show that the three parameter Weibull model contains the Extreme Value distribution as an embedded model, and derive statistical tests to deduce if the corresponding embedded model should be chosen in preference to the three parameter distribution.

2.3.2 Fitting the three parameter Burr distribution

Watkins (1999) presents an algorithm for fitting the three parameter Burr distribution to a data set, using Δ to eliminate the possibilities of fitting the Burr with $\phi \to \infty$ to data better modelled by the Weibull distribution. The method utilizes the Weibull MLEs, with $\widehat{\theta}$ scaling the data. The algorithm is also described in detail in Johnson (2003). We outline the main steps below.

STEP 1. We first fit the Weibull distribution to the data, using the profile log-likelihood given by (2.5). This yields estimates for β and θ.

STEP 2. The original data is then rescaled using $\widehat{\theta}$, so that

$$y_i \longrightarrow \frac{y_i}{\widehat{\theta}},$$

and Δ is calculated for these scaled values. This rescaling introduces some numerical stability in the calculation of Δ, especially in the computation of $S_0\left(2\widehat{\beta}\right)$. Also, the rescaled data values are now centered around one, thus providing us with a starting value for $\widehat{\phi}$, denoted by ϕ_1, if necessary.

STEP 3. If $\Delta > 0$, then we proceed to locate the MLEs that maximise l_b using the Newton-Raphson method. We set $\phi_1 = 1$, and fit the two parameter Burr distribution using a profile approach. Thus, we iterate on τ, setting our initial estimate equal to $\widehat{\beta}$. This provides us with a new estimate for τ, which we denote by τ_1; we then obtain α_1 using τ_1 in (2.10).

STEP 4. Using α_1 and τ_1, we calculate the three score functions and six second partial derivatives at (2.9), and use these values to obtain a new estimate of ϕ, ϕ_2. There are two ways we can update the estimate for ϕ in the Newton-Raphson approach. We either use the full-matrix of second derivatives, given by (2.9), or take the ratio of $\frac{\partial l_b}{\partial \phi}$ with its corresponding second derivative; the former usually results in faster convergence, but we illustrate both approaches in examples below. We then further scale the rescaled data by ϕ_2, and obtain new estimates of α and τ, now denoted by α_2, τ_2.

STEP 5. Step 4 is repeated until we converge onto $\widehat{\phi}$, $\widehat{\alpha}$ and $\widehat{\tau}$. We finally undo the effect of the initial rescaling by multiplying $\widehat{\phi}$ by $\widehat{\theta}$ to obtain $\widehat{\phi}$ for the original unscaled data.

Example

To illustrate this method, we consider the ball bearings data given in Table 1.1. The first step is to fit the Weibull distribution to this data set in order to obtain a value for Δ. From the above section, we have $\widehat{\beta} = 2.1021$, and so $S_0\left(\widehat{\beta}\right) = 241725.08$ and $S_0\left(2\widehat{\beta}\right) = 5.55078 \times 10^9$; this gives a value of 234911677 for Δ. Since Δ is positive, the Burr distribution will provide an improved fit, in terms of maximised likelihoods, over the Weibull. We rescale the data by $\widehat{\theta}$, and to this rescaled data, fit the three parameter Burr model. Table 2.2 shows results of the first 7 iterations for the rescaled ball bearings data. The first line shows starting values for fitting the two parameter Burr distribution. In this case, we set $\tau = \widehat{\beta}$ and give corresponding results for α and $\frac{\partial l_b}{\partial \phi}$. The remaining lines shown the iterations on the three parameters when fitting the Burr distribution. Note that we have used the full matrix of second derivatives to iterate on ϕ, as the MLEs converge much more quickly in this case. The second approach of using the reciprocal of the second derivative with respect to ϕ also eventually results in convergence, but this takes place well after the 100^{th} iteration. The final step is to undo the effects of scaling by multiplying $\widehat{\phi}$ by $\widehat{\theta}$. Thus, the Burr MLEs for

Iteration	$\widehat{\alpha}$	$\widehat{\tau}$	$\widehat{\phi}$	$\frac{\partial l_b}{\partial \phi}$
1	1.7062008	2.1020589	1	3.5692577
2	1.6570158	2.8899107	1	0.0920201
3	1.7859946	2.8368636	1.0407981	0.0113776
4	1.8071976	2.8288123	1.0474051	0.0002448
5	1.8076752	2.8286330	1.0475536	1.20×10^{-7}
6	1.8076754	2.8286329	1.0475536	3.84×10^{-14}
7	1.8076754	2.8289329	1.0475536	-9.59^{-15}

Table 2.2: Summary of iterations for fitting G_b to the ball bearings data in Table 1.1.

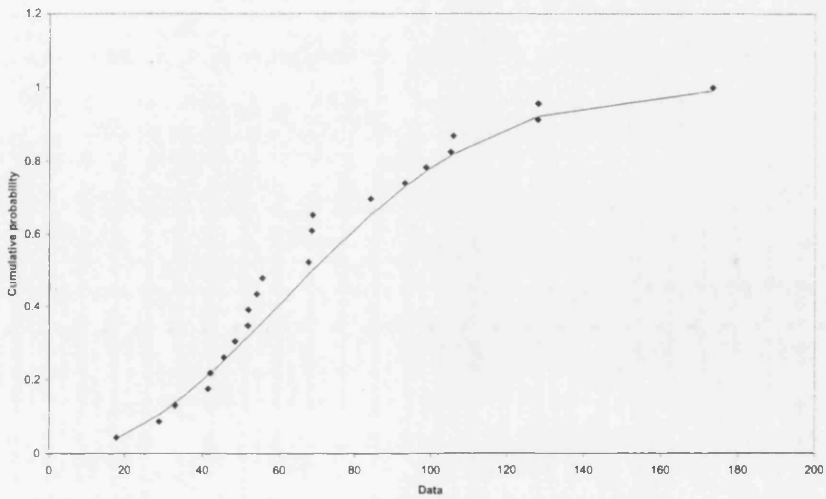


Figure 2.1: Empirical (♦) and fitted (—) cdfs for G_w for the ball bearings data set.

the ball bearings data set are

$$\widehat{\alpha} = 1.8077,\ \widehat{\tau} = 2.8286,\ \widehat{\phi} = 85.7719,$$

which gives a maximised log-likelihood of -113.2498. Although we know that the Burr distribution provides an improved fit over the Weibull for this particular data set, we still assess this goodness of fit for both distribution functions. Plots of the empirical cdf (ecdf) and fitted Weibull and Burr cdfs, given by Figures 2.1 and 2.2, clearly show larger discrepancies for the Weibull model. This fact is further strengthened when we compare the Kolmogorov-Smirnov distance for the Weibull, with that of the Burr. In the notation of Chapter 1, we have $D_b = 0.1116$, which compares with $D_w = 0.1511$. Thus, the Kolmogorov-Smirnov distance also supports the fact that the Burr distribution provides a better fit than the Weibull model. Plots of the cumulative hazard functions for the Weibull and Burr (given by Figures 2.3 and 2.4, respectively) are also consistent with this, and we observe a larger maximum absolute distance between sample and theoretical cumulative hazard functions

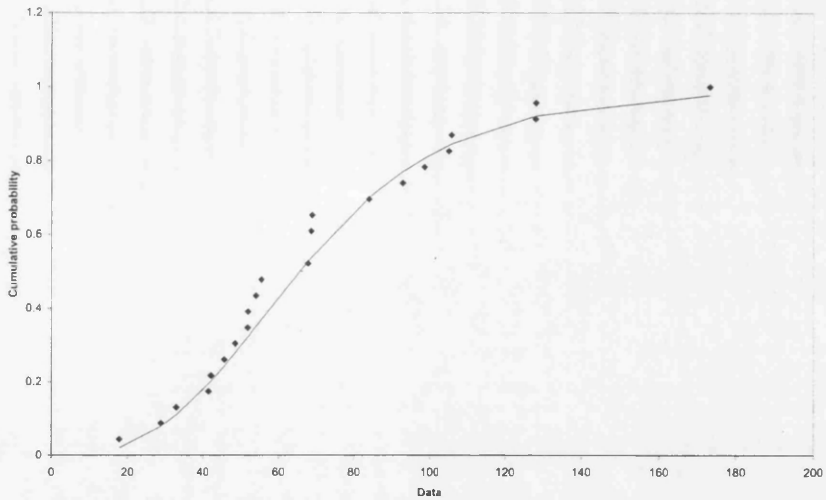


Figure 2.2: Empirical (♦) and fitted (—–) cdfs for G_b for the ball bearings data set.

for the Weibull distribution.

2.4 Fitting G_g

Using (1.5), we see that the likelihood function for data $y_1, y_2, ..., y_n$ is given by

$$L_g(y_1, y_2, ..., y_n; \tau, \alpha) = \prod_{i=1}^{n} g_g(y_i; \tau, \alpha) = \prod_{i=1}^{n} \frac{y_i^{\tau-1} \exp\left(-\frac{y_i}{\alpha}\right)}{\alpha^{\tau}\Gamma(\tau)},$$

from which

$$l_g = (\tau - 1) S_e - \alpha^{-1} S_0(1) - n\tau \ln\alpha - n \ln\Gamma(\tau). \tag{2.11}$$

The maximum of this function clearly has no analytic form, and must therefore be found numerically using iterative techniques. Equivalently, we seek the roots of

$$\frac{\partial l_g}{\partial \alpha} = \alpha^{-2} S_0(1) - n\tau\alpha^{-1}, \tag{2.12}$$

and

$$\frac{\partial l_g}{\partial \tau} = S_e - n\ln\alpha - n\Psi(\tau). \tag{2.13}$$

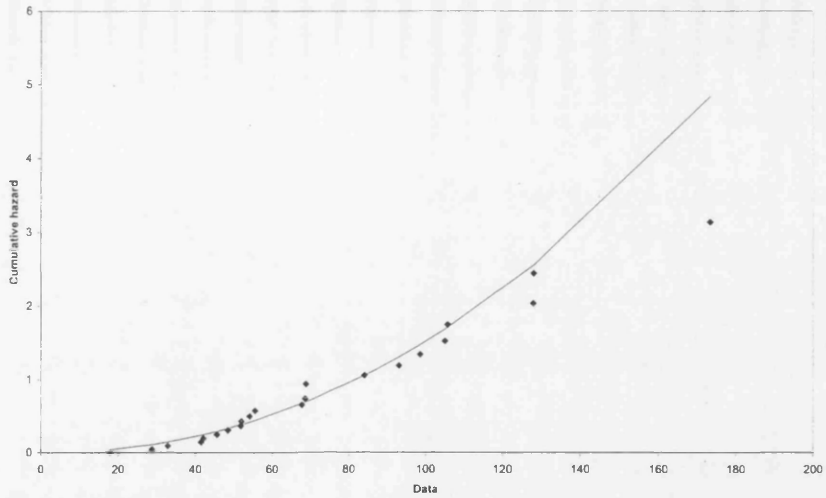


Figure 2.3: Sample (♦) and fitted (—) cumulative hazard functions for G_w for the ball bearings data set.

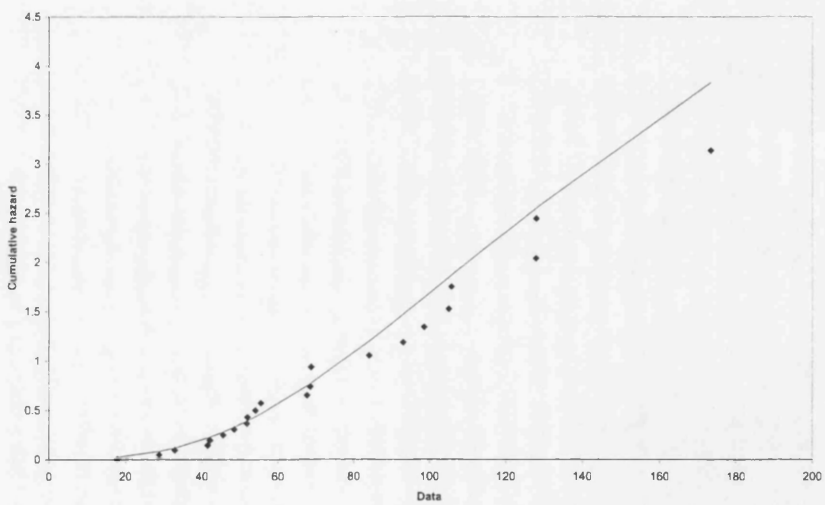


Figure 2.4: Sample (♦) and fitted (—) cumulative hazard functions for G_b for the ball bearings data set.

Iteration	τ	$\frac{dl_g^*}{d\tau}$	$\frac{d^2l_g^*}{d\tau^2}$
1	3	1.069620	-1.416817
2	3.754946	0.223208	-0.887050
3	4.006576	0.014534	-0.775282
4	4.025323	0.000070	-0.767812
5	4.025415	1.65×10^{-9}	-0.767776
6	4.025415	3.55×10^{-15}	-0.767776

Table 2.3: Summary of iterations for fitting G_g to the ball bearings data in Table 1.1.

Equating (2.12) to zero yields α in terms of the data and τ; we have

$$\alpha = \frac{S_0(1)}{n\tau}, \tag{2.14}$$

and substituting (2.14) into (2.11) yields the profile log-likelihood, which again offers the practical benefit of depending only on the single parameter, τ. We have

$$l_g^* = (\tau - 1) S_e - n\tau - n\tau \ln S_0(1) + n\tau \ln n + n\tau \ln \tau - n \ln \Gamma(\tau),$$

and

$$\frac{dl_g^*}{d\tau} = S_e - n \ln S_0(1) + n \ln n + n \ln \tau - n\Psi(\tau). \tag{2.15}$$

In order to use Newton-Raphson to locate the root of (2.15), we also require the second derivative,

$$\frac{d^2 l_g^*}{d\tau^2} = n\tau^{-1} - n\Psi'(\tau).$$

We now have all the necessary terms to locate MLEs of parameters from a Gamma distribution, and, with previous distributions, illustrate this with a worked example. Thus, we fit this distribution to the ball bearings data set given in Table 1.1, and find it convenient to set our initial value for τ equal to 3; this is consistent with MLEs of shape parameters of previous distributions. We note, but do not use, a method established in Hirose (1998) to locate MLEs for the three parameter Gamma distribution. Details on the iterations are shown in Table 2.3; we see convergence after just 6 iterations, and the MLEs for this particular data set are

$$\widehat{\tau} = 4.0254, \ \widehat{\alpha} = 17.9421,$$

which gives a maximum log-likelihood of -113.0293. Thus, in terms of maximised log-likelihoods, the Gamma distribution provides a better fit to the ball bearings data than the Weibull, and is also a further improvement over the Burr. We illustrate this by computing the Kolmogorov-Smirnov distance when we assume that the theoretical distribution is

Gamma with the above parameters. We have

$$D_g = 0.1230,$$

and note that $D_w > D_g > D_b$. Thus, in terms of maximised likelihoods the Gamma distribution provides the best fit to the data set, but when we consider the Kolmogorov-Smirnov statistic, then we favour the Burr. This example highlights some of the inconsistencies between methods for assessing the goodness of fit of a distribution to data.

2.5 Fitting $G_{\ln}$

The final distribution we consider is the Lognormal distribution. Using (1.9), we write the likelihood function as

$$L_{\ln}(y_1, y_2, ..., y_n; \mu, \sigma) = \prod_{i=1}^{n} g_{\ln}(y_i; \mu, \sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma y_i} \exp\left\{ \frac{-(\ln y_i - \mu)^2}{2\sigma^2} \right\},$$

with log-likelihood

$$l_{\ln} = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - S_e - \sum_{i=1}^{n} \frac{(\ln y_i - \mu)^2}{2\sigma^2}. \tag{2.16}$$

We exploit the link between the Lognormal and Normal distributions described in Chapter 1, to write down explicit expressions for the MLEs from the Lognormal model. Using the fact that $\ln Y\tilde{}N(\mu, \sigma^2)$, we have

$$\widehat{\mu} = n^{-1} S_e,$$

and

$$\widehat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(\ln y_i - \mu)^2}{n}};$$

thus, in contrast to the other distributions studied, we do not need to use numerical methods, since explicit expressions for the MLEs of this distribution are available. We illustrate this by fitting the Lognormal distribution to the ball bearings data given in Table 1.1, and assess the goodness of fit by calculating the Kolmogorov-Smirnov distance. Our MLEs are

$$\widehat{\mu} = 4.1505, \ \widehat{\sigma} = 0.5216,$$

which gives a maximum log-likelihood of -113.1286. This suggests that the Lognormal provides an improved fit over the Weibull on the basis of maximised likelihoods. This is further strengthened by the fact that $D_{\ln} = 0.0898 < D_w$. If we consider all four distributions

	l	D
Weibull	-113.6913	0.1511
Burr	-113.2498	0.1116
Gamma	-113.0293	0.1230
Lognormal	-113.1286	0.0898

Table 2.4: Comparison of maximised log-likelihoods and Kolmogorov-Smirnov tests for the Weibull, Burr, Gamma and Lognormal distributions when these are fitted to the ball bearings data set.

simultaneously, then we see that fitting the Weibull model to the ball bearings data actually results in the worst fit, when we use both maximised log-likelihoods and Kolmogorov-Smirnov distances; see Table 2.4. This example illustrates some of the problems of choosing a distribution based on popularity.

Now that we have all the theory developed to fit our four main distribution functions to data, we extend the results by examining the effects of mis-specifying the Weibull distribution and using this to model data with underlying Burr, Gamma and Lognormal models. We address this below, but first consider the results when we fit the correct distribution. This is, from a statistical perspective, the best scenario, and provides a standard against which all other analyses can subsequently be compared.

2.6 Fitting G_w to G_w data

We simulate data from G_w using (1.17) with

$$\beta = 2,\ \theta = 100,$$

for n = 50, 100, 300, 500 and 1000, and summarise the behaviour of the MLEs and $\widehat{B}_{w,10}$ when we fit the correct distribution to this data. Note that we can compare our sample values of $\widehat{B}_{w,10}$ with the true value given by

$$B_{w,10} = 100\{-\ln(0.9)\}^{\frac{1}{2}} = 32.4593.$$

Our results are summarised in Table 2.5. We observe excellent agreement between true and estimated results, even for small sample sizes, and this improves as n increases. We also see the sample standard errors for the Weibull MLEs and $\widehat{B}_{w,10}$ decrease as the sample size increases. A natural extension would be to examine how the Weibull MLEs are distributed, especially for large sample sizes. This will not be considered here, since the asymptotic distribution of these MLEs is well known; see, for instance, Bain (1978).

n	50	100	300	500	1000
$\widehat{\beta}$: mean	2.0558	2.0286	2.0099	2.0051	2.0017
(st.err.)	(0.2357)	(0.1606)	(0.0901)	(0.0704)	(0.0493)
$\widehat{\theta}$: mean	99.8239	99.9358	99.9847	99.9735	99.9629
(st.err.)	(7.4090)	(5.2494)	(3.0497)	(2.3615)	(1.6880)
$\widehat{B}_{w,10}$: mean	33.2909	32.9027	32.6175	32.5341	32.4731
(st.err.)	(5.4256)	(3.7708)	(2.1497)	(1.6924)	(1.1959)

Table 2.5: Summaries of the MLEs for G_w when fitted to Weibull data generated with $\beta = 2$, $\theta = 100$. Figures are based on at least 10000 replications.

2.7 Fitting G_w to G_b data

We now simulate data from G_b with

$$\alpha = 4, \tau = 3, \phi = 100;$$

as above, these values are somewhat arbitrary. For a set of data, we first calculate the Weibull MLEs, and find the sign of Δ to determine whether the Burr distribution can provide a better fit. If it does, then we fit this distribution using the algorithm outlined above, and obtain the Burr MLEs. We repeat this procedure at least 10000 times, and summarise the MLEs for both distributions; see Table 2.6 for n = 50, 100, 300, 500 and 1000. We also show the average value of Δ and the probability of fitting the Weibull distribution based on the proportion of times this function is negative. Finally, we include the estimate of B_{10} from both true and mis-specified distributions. Since we have used simulated data, then we can compare these estimates with the true value given by

$$B_{b,10} = 100\left\{(1-0.1)^{-\frac{1}{4}} - 1\right\}^{\frac{1}{3}} = 29.8848.$$

The results show that for small sample sizes, the Burr MLEs, especially for α and ϕ, do not agree with the true values very well at all, and it is only really for a sample size of 500 that we begin to obtain similarities between observed and expected results; even this agreement is not particularly good. We further note, for small sample sizes, particular replications with large estimates of α and ϕ; such values will clearly affect the sample mean and standard error. The reason for this seems to be linked to the fact that the Weibull is the limiting distribution for the Burr. The smaller the sample size, the less information we have, thus increasing the chance of choosing the incorrect distribution. Thus, if we almost, but not quite, prefer the Weibull distribution in favour of the Burr then the limiting arguments concerning α and ϕ apply and we see their estimates becoming extremely large. This occurs far more often for smaller sample sizes because of the lack of data, and leads, on occasion, to considerably larger estimates for α and ϕ. We note that these estimates do not affect the estimate of $B_{b,10}$ for the true distribution, and across all sample sizes, the agreement

n	50	100	300	500	1000
$\widehat{\beta}$: mean	2.6531	2.6035	2.5706	2.5649	2.5592
(st.err.)	(0.3222)	(0.2232)	(0.1304)	(0.1014)	(0.0726)
$\widehat{\theta}$: mean	67.1717	67.2201	67.2263	67.2715	67.2597
(st.err.)	(3.9092)	(2.7712)	(1.6188)	(1.2433)	(0.8809)
$\widehat{B}_{w,10}$: mean	28.5667	28.2215	27.9774	27.9549	27.9062
(st.err.)	(3.4209)	(2.4139)	(1.4219)	(1.1089)	(0.7908)
Δ : mean	3.2630	9.3855	35.1285	61.7992	129.9680
(st.err.)	(5.7515)	(11.0927)	(26.5352)	(40.2960)	(68.9834)
Pr (Fit Weibull)	0.3120	0.1696	0.0241	0.0051	0
$\widehat{\tau}$:mean	3.2453	3.1210	3.0205	3.0085	3.0048
(st.err.)	(0.5356)	(0.3907)	(0.2273)	(0.1782)	(0.1272)
$\widehat{\alpha}$: mean	5.8481	6.0846	5.6085	4.9535	4.3944
(st.err.)	(7.6847)	(7.3206)	(6.0028)	(3.4441)	(1.6074)
$\widehat{\phi}$: mean	104.4379	108.5870	108.8744	106.1578	102.8393
(st.err.)	(45.8092)	(42.4754)	(32.1425)	(23.6247)	(14.5185)
$\widehat{B}_{b,10}$: mean	30.7167	30.2949	29.9458	29.9275	29.9059
(st.err.)	(3.5238)	(2.5300)	(1.4970)	(1.1670)	(0.8302)

Table 2.6: Summary statistics of the MLEs for G_b and G_w for data generated from G_b with $\alpha = 4$, $\tau = 3$, $\phi = 100$.

between $\widehat{B}_{b,10}$ and its true value is relatively good. When we estimate this quantile using the mis-specified distribution, we generally seem to be under-estimating the time to which 10% of the observations fail. In all cases, the sample standard errors of the MLEs from both Weibull and Burr distributions decrease. The MLEs for the Weibull also seem to be centering around some limiting values as n increases ($\beta \simeq 2.55$, $\theta \simeq 67.25$).

2.7.1 Assessing the goodness of fit

When we consider mis-specification using the Burr distribution to generate the underlying data set, we have a numerical method to determine which is the preferred distribution. When we consider using other distributions as the underlying model, then, as we see below, there are no discriminating functions. Consequently, we should assess the goodness of fit of Burr and Weibull models to Burr data using alternative techniques, since such procedures must be used for the other distribution functions. It is also of interest to assess whether these procedures give results consistent with Δ. We examine two such procedures below, based on Kolmogorov-Smirnov distances and cumulative hazard plots.

The Kolmogorov-Smirnov distance

We use this test to determine whether the true distribution or mis-specified Weibull provides a better fit to a set of data, and calculate this statistic for the 10000 replications of data, with n = 50, 100, 300, 500, 1000, in each case using the MLEs as the parameter values in

n	50	100	300	500	1000
D_w : mean (st.dev.)	0.0897 (0.0213)	0.0680 (0.0168)	0.0461 (0.0112)	0.0398 (0.0094)	0.0339 (0.0071)
D_b : mean (st.dev.)	0.0781 (0.0172)	0.0564 (0.0125)	0.0332 (0.0073)	0.0259 (0.0057)	0.0183 (0.0040)
$D_w - D_b$: mean (st.dev.)	0.0147 (0.0178)	0.0135 (0.0149)	0.0132 (0.0108)	0.0140 (0.0093)	0.0156 (0.0073)
Δ : mean (st.dev.)	3.2630 (5.7515)	9.3855 (11.0927)	35.1285 (26.5352)	61.7992 (40.9260)	129.9680 (68.9834)
$\Pr(\Delta > 0)$	0.6880	0.8304	0.9759	0.9949	1
$\Pr(D_w > D_b \mid \Delta > 0)$	0.7929	0.8247	0.8999	0.9420	0.9848

Table 2.7: Summary statistics for comparing D_w with D_b. Data is simulated from a Burr distribution with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ and for n =50,100,300,500 and 1000.

both distributions. From Chapter 1, we noted how this test could be used to accept or reject the hypothesis that a data set is adequately modelled using a distribution specified. If the probability of observing a value of D, derived from this test, was small (that is, less than some specified significance level), then we reject the null hypothesis that the distribution is a good representation of the data set. We do not use this interpretation here, but as above, base our choice between true and mis-specified distributions on the size of the test statistic. For example, if $D_w < D_b$ then we prefer the mis-specified distribution over the true. We summarise the results in Table 2.7, and also include the summary statistics for Δ, after this has been rescaled using $\widehat{\theta}$ from the Weibull distribution, to gauge the consistency between the discriminating function and the Kolmogorov-Smirnov distance. We note that results for D_b are only recorded when $\Delta > 0$, since, if this condition is not satisfied, we cannot fit the Burr distribution, and so do not obtain results for the Kolmogorov-Smirnov distance for this model. An alternative approach would be to set $D_b = D_w$ since the Weibull is embedded in the Burr, thus resulting in no difference between true and mis-specified distributions. The results show a decrease in the Kolmogorov-Smirnov statistics for both distributions as the sample size increases. In all cases, the average of D_b is always less than the average of D_w so, on the whole, we prefer to fit the true distribution. We also see the number of times both methods reach the same conclusion increases with n; this suggests the tests are increasingly consistent for larger samples. Figures 2.5, 2.6, 2.7, 2.8 and 2.9 illustrate the relationship between $D_w - D_b$ and Δ for n = 50, 100, 300, 500 and 1000 respectively. Generally across all sample sizes, large values of Δ correspond to large differences. We also see more extreme values of Δ for larger n.

Hazard plots

We also assess the goodness of fit between the Burr and mis-specified Weibull using cumulative hazard functions. We compare these continuous functions for both distributions with the empirical counterparts obtained from the data. To illustrate this procedure, we consider

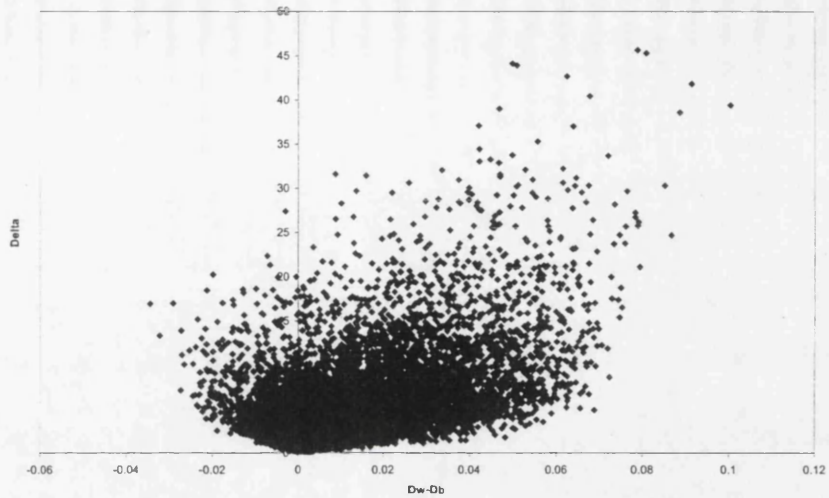


Figure 2.5: $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 50$).

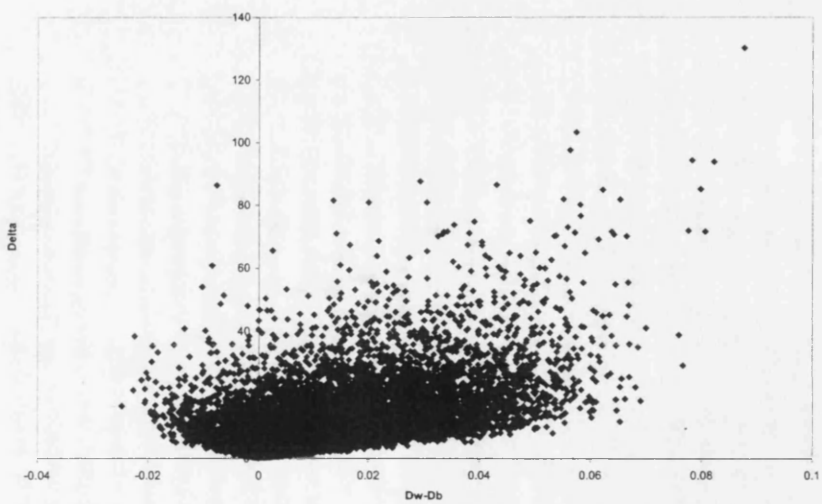


Figure 2.6: $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 100$).

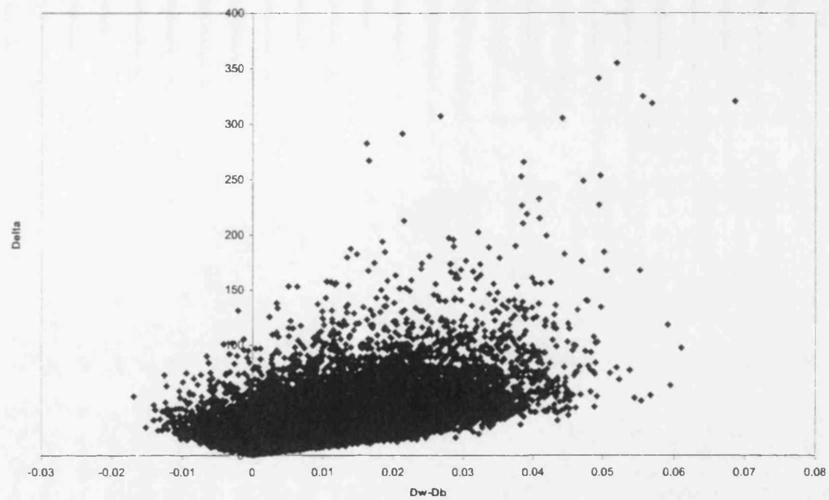


Figure 2.7: $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 300$).

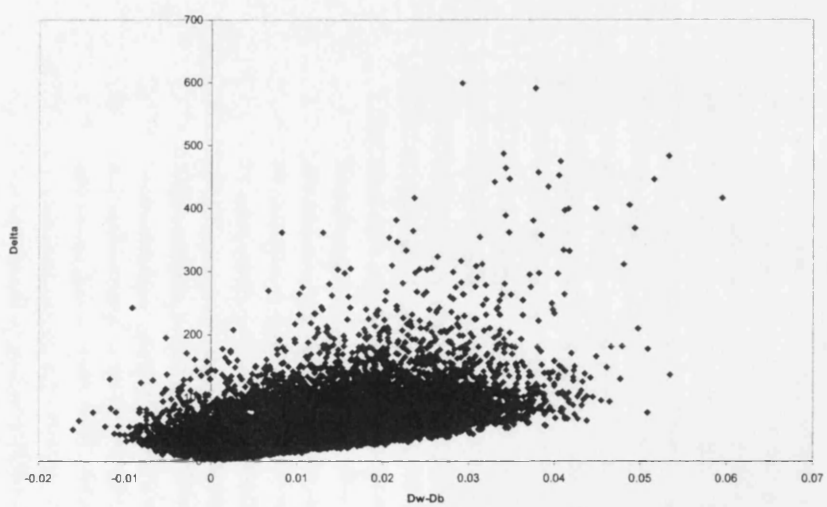


Figure 2.8: $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 500$).

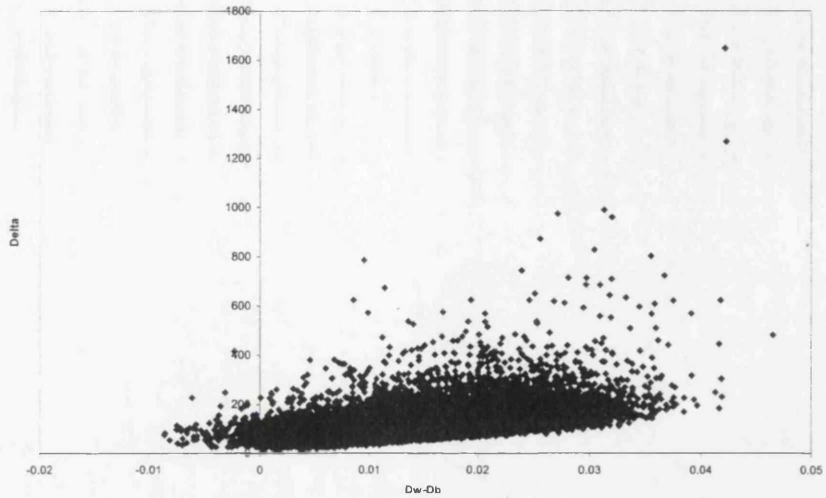


Figure 2.9: $D_w - D_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 1000$).

1000 observations simulated from the Burr with the above parameter values. To this, we fit both the Weibull and Burr distributions, and obtain the following MLEs

$$\widehat{\beta} = 2.4171,\ \widehat{\theta} = 66.9221,\ \widehat{\tau} = 2.9101,\ \widehat{\alpha} = 3.4480,\ \widehat{\phi} = 94.6071$$

We then construct the sample cumulative hazard function for data, and compare this to $H_w\left(y; \widehat{\beta}, \widehat{\theta}\right)$ and $H_b\left(y; \widehat{\alpha}, \widehat{\tau}, \widehat{\phi}\right)$; see Figures 2.10 and 2.11. These show generally good agreement between sample and theoretical results, although there are some discrepancies for the Weibull distribution, especially for large y. When we compute the maximum absolute distance between the sample cumulative hazard function and the theoretical hazard function for both the Weibull and Burr, we see that the larger distance occurs for the Weibull distribution; we denote this distance by CH, with corresponding sub-scripts for the appropriate distributions. Thus, for the Weibull we have $CH_w = 3.8455$, which we compare to $CH_b = 0.8780$ for the Burr. So, for this particular data set, we conclude, by examining hazard plots, that the Burr distribution provides a better fit over the Weibull. Note that this is consistent with a positive value of Δ (137.8498), and $D_w > D_b$.

This procedure is repeated at least 10000 times each for varying values of n, and we record the percentage of times the three methods for comparing goodness of fit are consistent. This is achieved by examining, for $\Delta > 0$, the number of times $D_w > D_b$ and $CH_w > CH_b$; the results are summarised in Table 2.8. The table shows CH_w increasing as the sample size increases; the distance for the Burr remains constant at around 1. The average difference between the two distances also increases for larger n, so suggesting that the Burr is favoured over the Weibull for such sample sizes; this is consistent with observing small probabilities

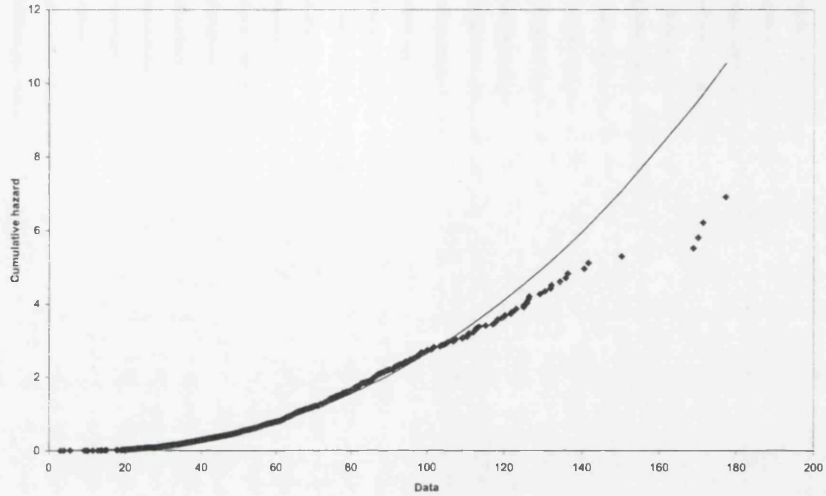


Figure 2.10: Sample (♦) and theoretical (—) cumulative hazard functions for G_w with $\widehat{\beta} = 2.4171$ and $\widehat{\theta} = 66.9221$. We simulate 1000 data values from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$.

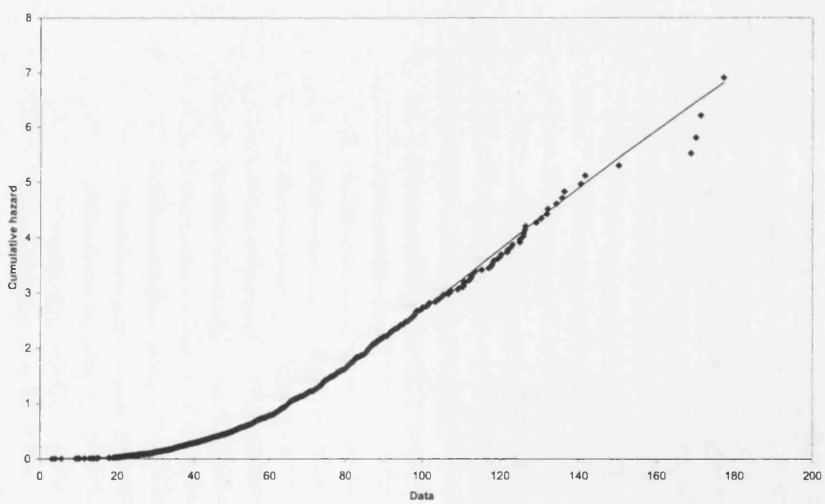


Figure 2.11: Sample (♦) and theoretical (—) cumulative hazard functions for G_b with $\widehat{\tau} = 2.9101$, $\widehat{\alpha} = 3.4480$ and $\widehat{\phi} = 94.6071$. We simulate 1000 data values from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$.

n	50	100	300	500	1000
CH_w : mean	1.6247	2.1566	3.3522	4.1039	5.3849
(st.dev.)	(1.1714)	(1.6574)	(2.5253)	(3.0929)	(3.9183)
CH_b : mean	1.0057	0.9872	1.0078	1.0250	1.0624
(st.dev.)	(0.4547)	(0.5476)	(0.6804)	(0.7251)	(0.7778)
$CH_b - CH_w$: mean	1.0527	1.4748	2.4099	3.0964	4.3225
(st.dev.)	(0.8987)	(1.2989)	(2.0316)	(2.5414)	(3.3019)
$\Pr(CH_w > CH_b \| \Delta > 0)$	0.9985	0.9971	0.9814	0.9797	0.9875
$\Pr(CH_w > CH_b$ and $D_w > D_b \| \Delta > 0)$	0.7917	0.8219	0.8825	0.9223	0.9725

Table 2.8: Summary statistics for comparing CH_w with CH_b. Data is simulated from a Burr distribution with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ and for n =50,100,300,500 and 1000.

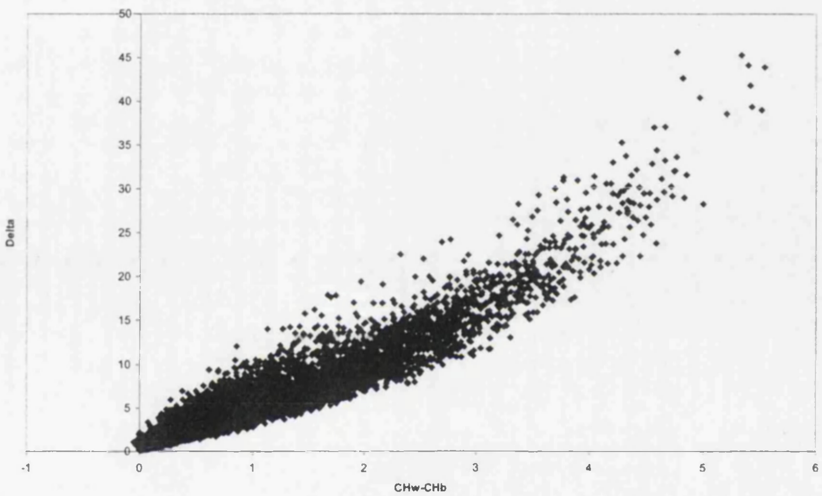


Figure 2.12: $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 50$).

for fitting the Weibull distribution in our simulation studies. The agreement between the number of times $\Delta > 0$ and $CH_w > CH_b$ is very good across all sample sizes, and never goes below 97%. There is also a 20% improvement over the Kolmogorov-Smirnov distance, when we compare this method with Δ for a sample size of 50. As n increases, we see the agreement between all three methods improve. Figures 2.12, 2.13, 2.14, 2.15 and 2.16 show $CH_w - CH_b$ against Δ for $n = 50$, 100, 300, 500 and 1000. The plots have a distinct linear structure, with some discrepancies at the tails, usually corresponding to the more extreme values of Δ.

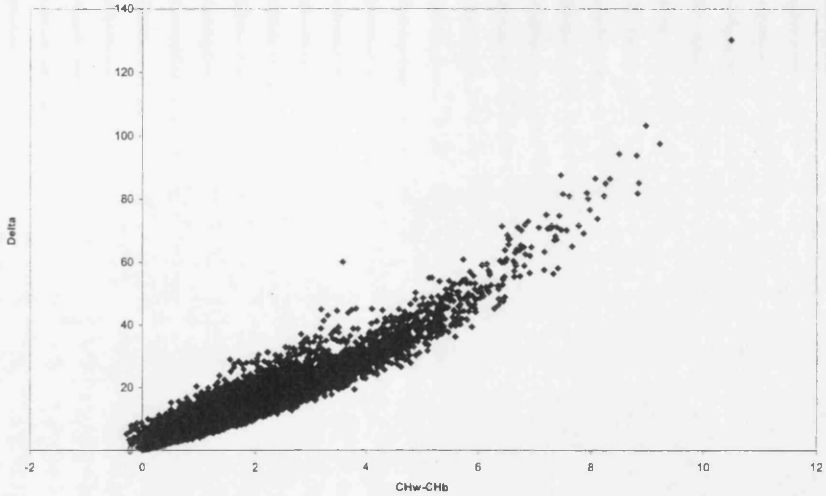


Figure 2.13: $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 100$).

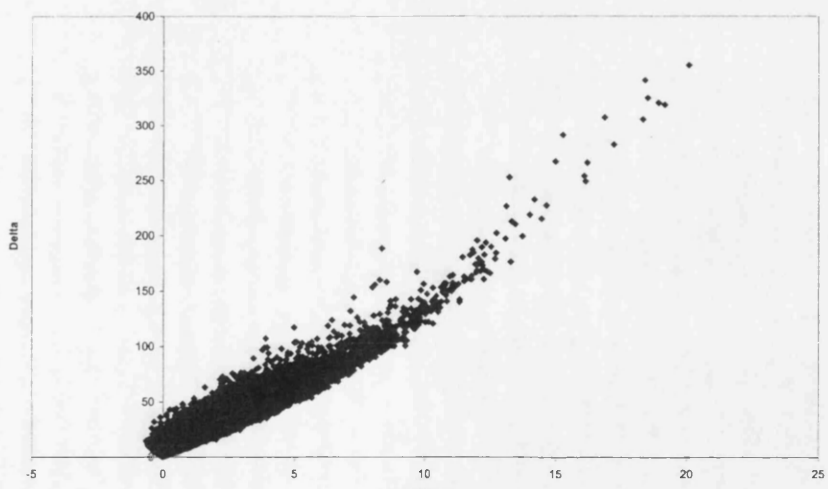


Figure 2.14: $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 300$).

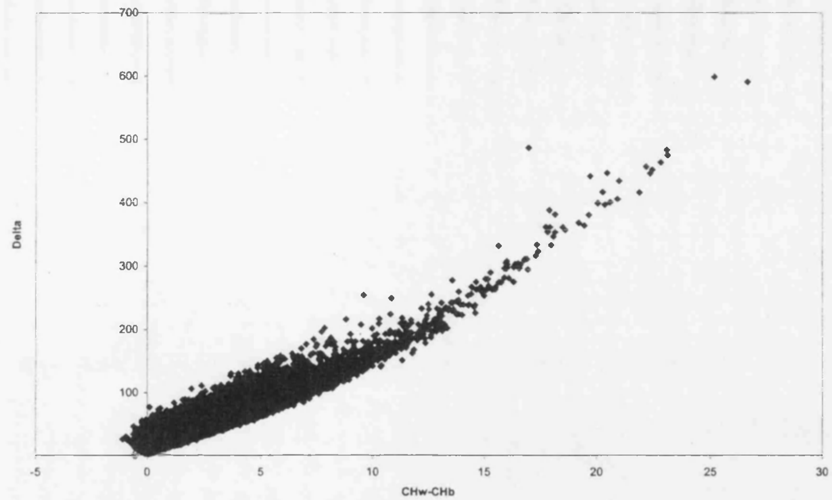


Figure 2.15: $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 500$).

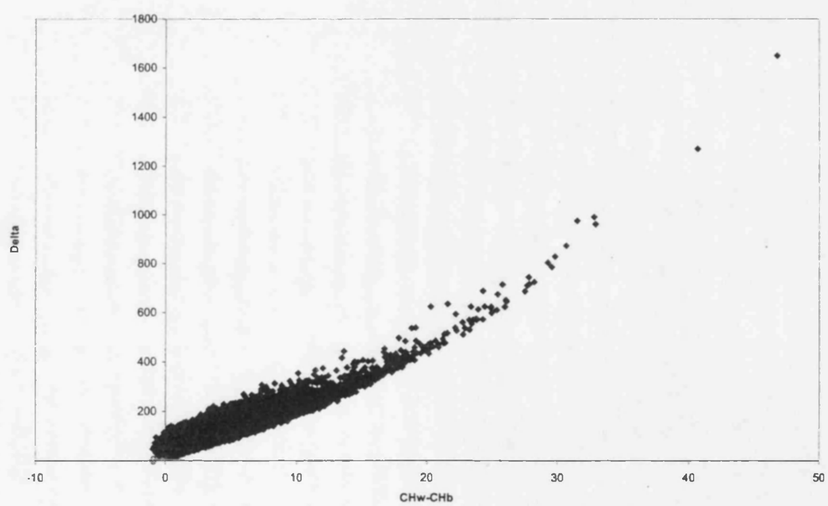


Figure 2.16: $CH_w - CH_b$ against Δ for data generated from G_b with $\alpha = 4$, $\tau = 3$ and $\phi = 100$ ($n = 1000$).

n	50	100	300	500	1000
$\widehat{\beta}$: mean (st.err.)	1.8980 (0.2130)	1.8679 (0.1461)	1.8439 (0.0828)	1.8391 (0.0644)	1.8368 (0.0454)
$\widehat{\theta}$: mean (st.err.)	169.3207 (13.8920)	169.2524 (9.7800)	169.3936 (5.6231)	169.3192 (4.3701)	169.3600 (3.1471)
$\widehat{B}_{w,10}$: mean (st.err.)	51.4429 (8.0651)	50.6017 (5.6769)	49.9438 (3.2634)	49.7806 (2.5433)	49.7313 (1.8006)
$\widehat{\tau}$: mean (st.err.)	3.1849 (0.6401)	3.0962 (0.4238)	3.0286 (0.2341)	3.0155 (0.1840)	3.0104 (0.1288)
$\widehat{\alpha}$: mean (st.err.)	48.9377 (10.2611)	49.3171 (7.2266)	49.8312 (4.1609)	49.9124 (3.2788)	49.9143 (2.3188)
$\widehat{B}_{g,10}$: mean (st.err.)	56.5015 (8.4512)	55.8505 (5.9162)	55.3384 (3.3787)	55.2008 (2.6646)	55.1875 (1.8817)
$\Pr(l_w > l_g)$	0.2975	0.2084	0.0724	0.0303	0.0040

Table 2.9: Summaries of the MLEs for G_w and G_g, when fitted to Gamma data generated with $\tau = 3$, $\alpha = 50$.

2.8 Fitting G_w to G_g data

We now carry out a similar investigation when the data set is simulated from G_g, and we fit G_w and G_g to data for a selection of sample sizes. Unlike the Burr distribution, however, we can always fit G_g to data, but cannot determine in advance whether this distribution or Weibull distribution is preferred, since there is no counterpart to Δ here. We rely on other techniques to assess the goodness of fit between the distributions, and can also use maximised log-likelihoods as a basis for determining the better fit. If $l_g > l_w$ then we conclude that G_g is to be preferred to G_w, whilst $l_w > l_g$ gives the opposite conclusion. For illustration, we simulate data from a Gamma distribution with

$$\tau = 3,\ \alpha = 50;$$

we consider the effects of changing parameter values below. As before, we also compare our estimates for B_{10} from true and mis-specified distribution functions, comparing these estimates with the true value $B_{g,10} = 55.1033$. The results from the simulations are summarised in Table 2.9 for varying sample sizes. As expected, we see the MLEs for the Gamma distribution converging towards their true values, and the standard errors of these MLEs decrease. The estimates for the Weibull again seem to be centering around specific values; here $\widehat{\beta} \simeq 1.84$, $\widehat{\theta} \simeq 169.4$. When examining log-likelihoods, we see that the probability of choosing the Weibull distribution over the Gamma decreases with increasing sample size. As with the case of the Burr distribution, when we mis-specify the Weibull model and fit this to data with an underlying Gamma distribution, we always seem to under-estimate B_{10}. When we fit the true distribution, there is excellent agreement between estimates of this quantile and its true value.

n	50	100	300	500	1000
$D_w - D_g$: mean	0.0050	0.0079	0.0123	0.0145	0.0168
(st.dev.)	(0.0178)	(0.0155)	(0.0117)	(0.0097)	(0.0075)
$CH_w - CH_g$: mean	0.5142	0.7947	1.3516	1.6879	2.2007
(st.dev.)	(0.4460)	(0.6405)	(1.0043)	(1.1780)	(1.3791)
$\Pr(l_g > l_w)$	0.7025	0.7916	0.9276	0.9697	0.9960
$\Pr(D_g < D_w)$	0.6401	0.7146	0.8479	0.9066	0.9667
$\Pr(CH_g < CH_w)$	0.8966	0.8934	0.9096	0.9226	0.9510
$\Pr(l_g > l_w$ and $D_g < D_w)$	0.5658	0.6596	0.8276	0.8975	0.9653
$\Pr(l_g > l_w$ and $CH_g < CH_w)$	0.6882	0.7619	0.8642	0.9041	0.9488
$\Pr(D_g < D_w$ and $CH_g < CH_w)$	0.6060	0.6653	0.7823	0.8427	0.9209
Pr (Consistent conclusions)	0.5548	0.6339	0.7712	0.8382	0.9205

Table 2.10: Summary statistics for comparing Kolmogorov-Smirnov and hazard functions for G_w and G_g. Data is simulated from G_g with $\alpha = 50$, $\tau = 3$, and for n =50,100,300,500 and 1000.

2.8.1 Assessing the goodness of fit

As in the case of the Burr distribution, we use Kolmogorov-Smirnov and cumulative hazard distances to assess how often we would prefer to fit the mis-specified distribution over the true Gamma. We summarise results in Table 2.10 for varying sample sizes, and note the probability that the tests give the same result. This is achieved by recording the number of times $D_w > D_g$, $CH_w > CH_g$ and $l_w < l_g$. We also include the number of times when two of the three tests result in the same conclusion. The table shows the difference between both the Kolmogorov-Smirnov statistics and cumulative hazards increase as the sample size increases, and always remains positive. Such results suggest that the true distribution is preferred over the mis-specified for all sample sizes considered, but particularly so as n increases. When looking at agreement between the methods, we see these proportions also increase for larger n. For smaller sample sizes, maximum absolute distances between hazards agrees far more with maximised likelihoods, than the Kolmogorov-Smirnov distance. We also construct plots of $D_w - D_g$ against $l_g - l_w$; see Figures 2.17, 2.18, 2.19, 2.20 and 2.21, and $CH_w - CH_g$ against $l_g - l_w$, shown by Figures 2.22, 2.23, 2.24, 2.25 and 2.26, for varying sample sizes. They both show that as the distances between either the hazard functions or cdfs increase, so does the difference between the maximised log-likelihoods. This is true across all sample sizes.

2.9 Fitting G_w to G_{ln} data

We simulate data from a Lognormal distribution with $\mu = 2$, $\sigma = 3$, and fit the Weibull distribution to this data set. As in previous studies, these parameter values are arbitrary, and are used only to illustrate some effects of mis-specification; we examine the effects of varying such parameter values in Chapter 7. This process is repeated for at least 10000

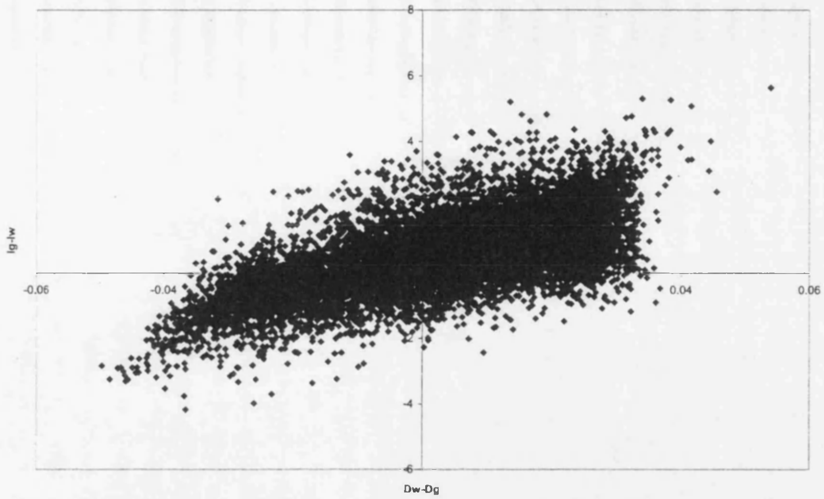


Figure 2.17: $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ $(n = 50)$.

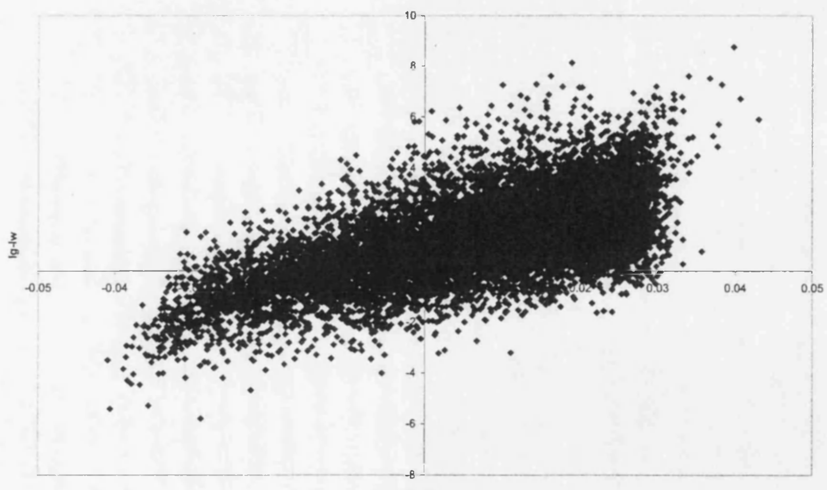


Figure 2.18: $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ $(n = 100)$.

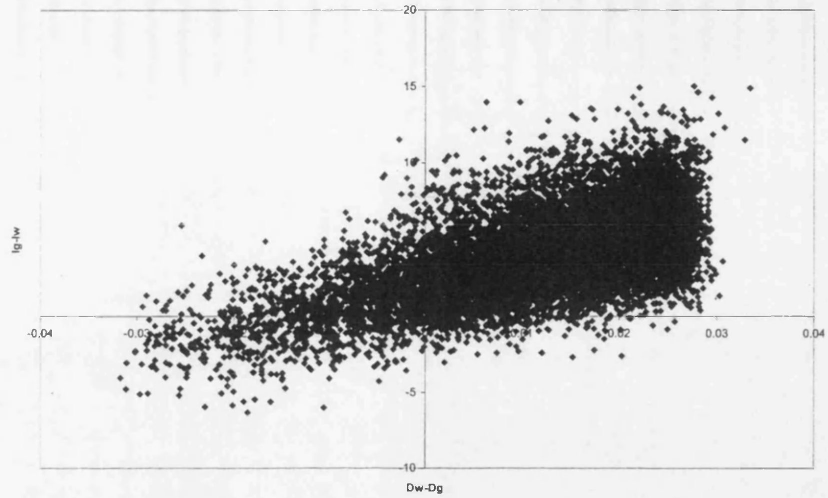


Figure 2.19: $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 300$).

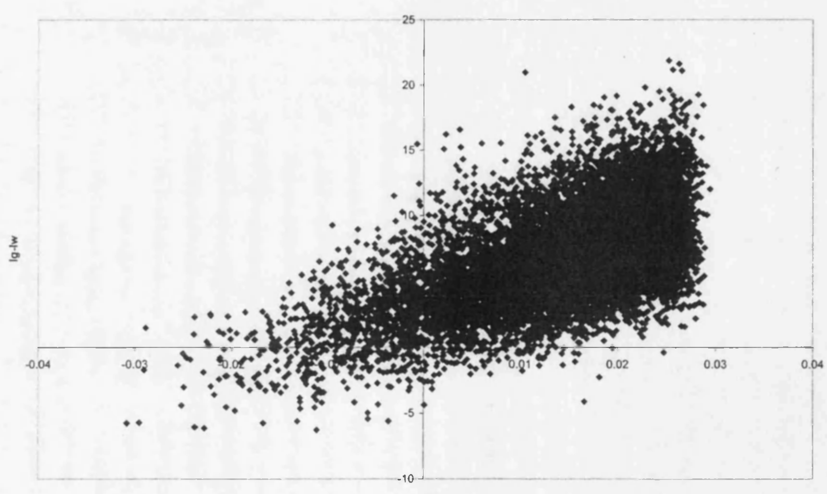


Figure 2.20: $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 500$).

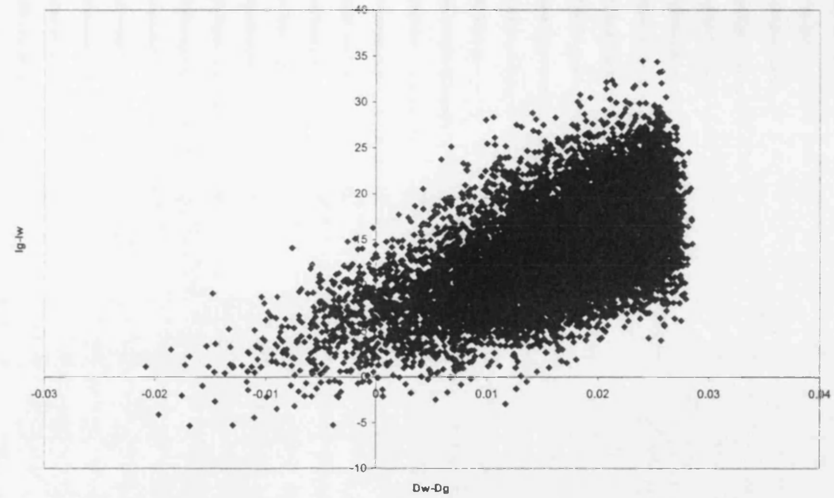


Figure 2.21: $D_w - D_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 1000$).

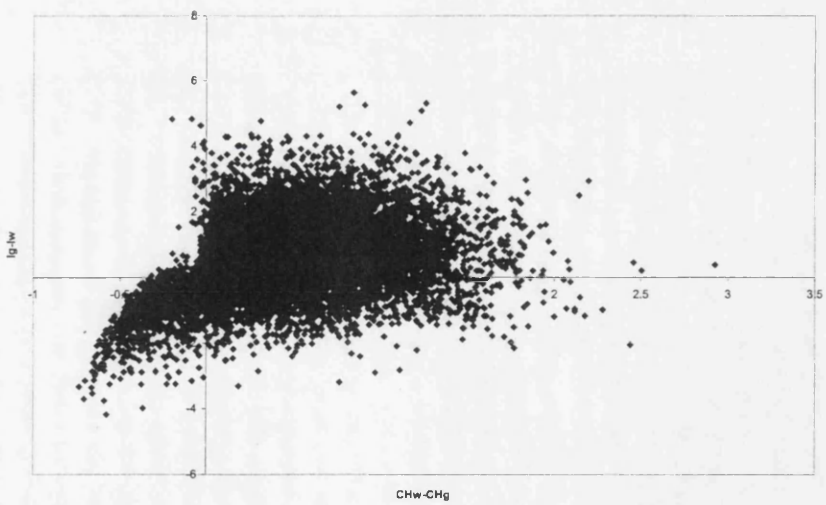


Figure 2.22: $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 50$).

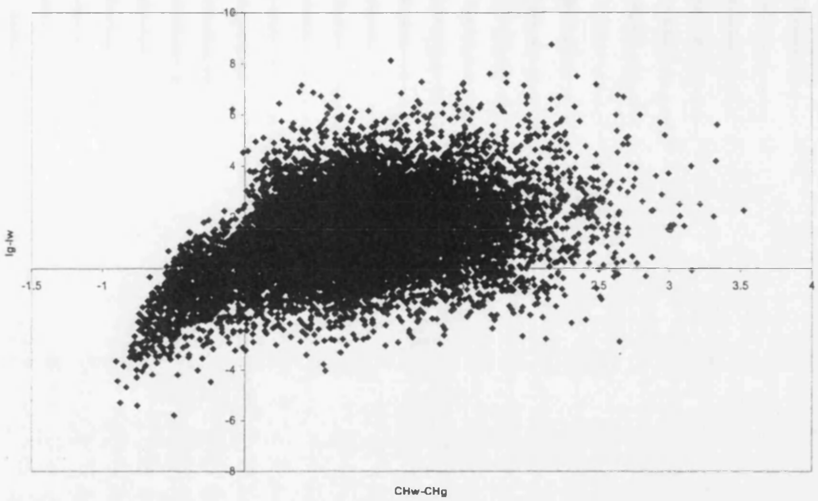


Figure 2.23: $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 100$).

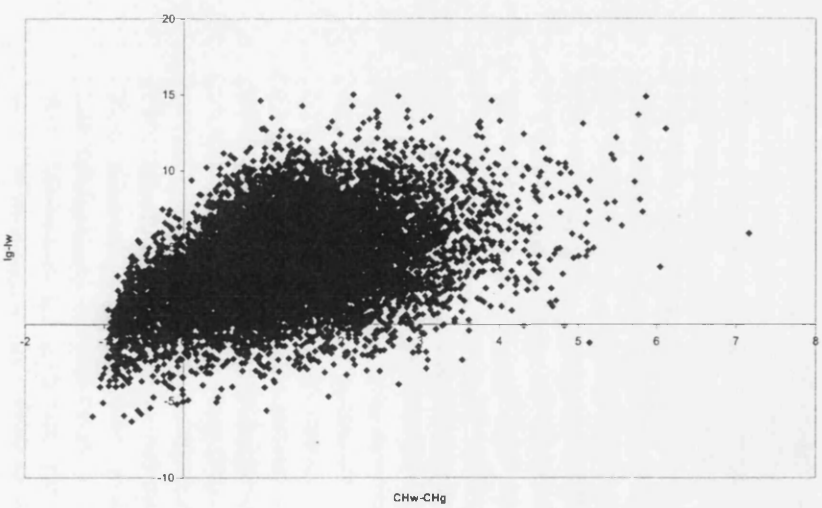


Figure 2.24: $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 300$).

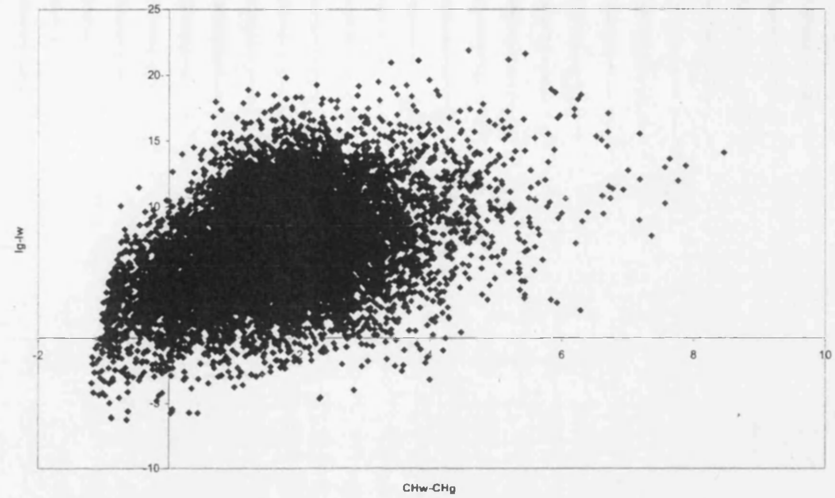


Figure 2.25: $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 500$).

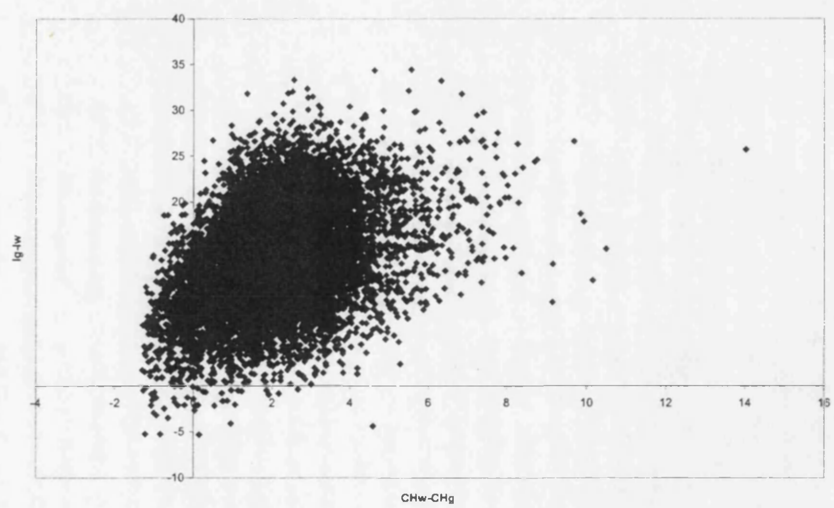


Figure 2.26: $CH_w - CH_g$ against $l_g - l_w$ for data generated from G_g with $\tau = 3$ and $\alpha = 50$ ($n = 1000$).

n	50	100	300	500	1000
$\widehat{\beta}$: mean	0.3526	0.3437	0.3372	0.335	0.3347
(st.err.)	(0.0465)	(0.0323)	(0.0192)	(0.0151)	(0.0108)
$\widehat{\theta}$: mean	36.1582	34.5009	33.5641	33.4747	33.2101
(st.err.)	(17.4014)	(11.2337)	(6.2663)	(4.8275)	(3.3867)
$\widehat{B}_{w,10}$	0.0678	0.0527	0.0437	0.0419	0.0403
(st.err.)	(0.0597)	(0.0320)	(0.0161)	(0.0122)	(0.0085)
$\widehat{\mu}$: mean	2.0054	2.0010	1.9998	2.0020	1.9988
(st.err.)	(0.4279)	(0.2964)	(0.1738)	(0.1343)	(0.0948)
$\widehat{\sigma}$: mean	2.9554	2.9783	2.9918	2.9962	2.9976
(st.err.)	(0.3023)	(0.2120)	(0.1223)	(0.0962)	(0.0678)
$\widehat{B}_{\text{ln},10}$: mean	0.1980	0.1761	0.1642	0.1618	0.1597
(st.err.)	(0.1220)	(0.0722)	(0.0390)	(0.0295)	(0.0204)
$\Pr(l_w > l_{\text{ln}})$	0.0930	0.0275	0.0005	0	0

Table 2.11: Summaries of the MLEs for G_w and G_{ln}, when fitted to Lognormal data generated with $\mu = 2$, $\sigma = 3$.

replications for each value of n, and, for each replication, we calculate the MLEs for both distributions. We also compute the estimates for B_{10} and compare these with the true value of 0.1581. The results for varying sample sizes are shown in Table 2.11. As expected, the estimates for the parameters from the Lognormal distribution converge to the true values as the sample size is increased. The Weibull MLEs also seem to be converging to some fixed value for large n ($\beta \simeq 0.335$, $\theta \simeq 33.21$). With respect to the log-likelihoods, we see much smaller probabilities associated with fitting the Weibull distribution than previously. Thus, we are less likely to fit the Weibull distribution, if the underlying data set is Lognormal with parameters similar to the ones used in this simulation. When we compare estimates of B_{10}, we see good agreement for the true distribution. However, the estimate of this quantile from the mis-specified model is particularly bad, and the time to which 10% of observations fail is grossly under-estimated.

2.9.1 Assessing the goodness of fit

This final section of Chapter 2 will consider just how good a fit the Weibull distribution is to data simulated from the Lognormal. Some attention has been received in this area, and Croes, Manca, De Ceuninck, De Schepper and Molenberghs (1998) derive a method for choosing between the Weibull or Lognormal distribution. They calculate the correlation coefficient of the points on the Weibull and Lognormal probability plots of the experiment under consideration, and then consider the ratio of these coefficients to determine whether to choose the Weibull or Lognormal distribution. Cain (2002) further discusses this test, but we do not consider such procedures here. As with previous distributions studied, we compare sample and theoretical cdfs and cumulative hazard functions for varying sample sizes, and when we set $\mu = 2$, $\sigma = 3$. We also use maximised log-likelihoods from the Weibull and

n	50	100	300	500	1000
$D_w - D_{\ln}$: mean	0.0281	0.0350	0.0438	0.0467	0.0502
(st.dev.)	0.0281	0.0223	0.0142	0.0112	0.0082
$CH_w - CH_{\ln}$: mean	1.4982	2.4269	4.6655	6.0091	8.3187
(st.dev.)	0.9497	1.5248	2.8354	3.4914	4.8120
$\Pr(l_{\ln} > l_w)$	0.9070	0.9725	0.9995	1	1
$\Pr(D_{\ln} < D_w)$	0.8377	0.9214	0.9932	0.9992	1
$\Pr(CH_{\ln} < CH_w)$	0.9553	0.9674	0.9894	0.9959	0.9997
$\Pr(l_{\ln} > l_w$ and $D_{\ln} < D_w)$	0.8115	0.9135	0.9930	0.9992	1
$\Pr(l_{\ln} > l_w$ and $CH_{\ln} < CH_w)$	0.8929	0.9496	0.9892	0.9959	0.9997
$\Pr(D_{\ln} < D_w$ and $CH_{\ln} < CH_w)$	0.8160	0.8968	0.9829	0.9951	0.9997
Pr (Consistent conclusions)	0.7997	0.8925	0.9829	0.9951	0.9997

Table 2.12: Summary statistics for comparing Kolmogorov-Smirnov and hazard functions for G_w and G_{ln}. Data is simulated from G_{ln} with $\mu = 2$, $\sigma = 3$, and for n =50,100,300,500 and 1000

Lognormal distributions to determine the better fit. The results are summarised in Table 2.12. We include details on the number of times test methods agree with one another. The final row shows the percentage of consistent results. This is calculated by recording the number of times all three test methods agree out of the total number of simulations; this always exceeds 10000. The results show an increase in the difference between D_w and $D_{\ln}$ as n increases, which suggests, for larger sample sizes, the Lognormal is preferred. The standard deviation for this function also decreases, so the chances of having a negative value is much less for larger sample sizes. This is consistent with the probabilities associated with choosing the Weibull over the Lognormal when we examine maximised log-likelihood. When we look at differences between cumulative hazard functions, we prefer the true distribution over the Weibull far more often for smaller sample sizes (there is a 5% increase when compared to maximised log-likelihoods, and a 10% rise when compared to the Kolmogorov-Smirnov distance). The percentage of consistent results between all three methods is very good, even for small sample sizes. This tends to one for larger values of n. We construct plots of $D_w - D_{\ln}$ against $l_{\ln} - l_w$; see Figures 2.27, 2.28, 2.29, 2.30 and 2.31, and $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$, shown by Figures 2.32, 2.33, 2.34, 2.35 and 2.36, for varying sample sizes. They both show that as the distances between either the hazard functions or cdfs increase, so does the difference between the maximised log-likelihoods. This holds across all sample sizes.

2.10 Summary

In this chapter, we outlined the theory necessary to fit Weibull, Burr, Gamma and Lognormal distributions to data using ML techniques. We then considered mis-specifying the Weibull distribution, fitting this to data with an underlying Burr, Gamma and Lognormal model. In each case, we constructed a set of simulations for varying sample sizes, each time

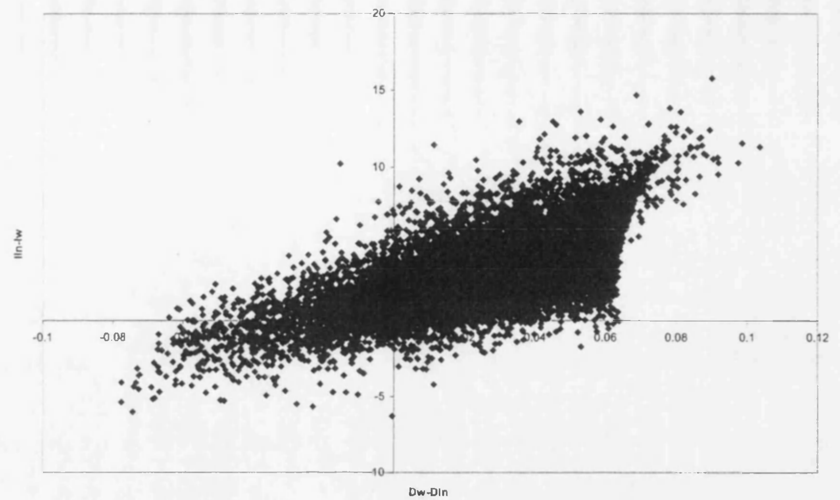


Figure 2.27: $D_w - D_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 50$).

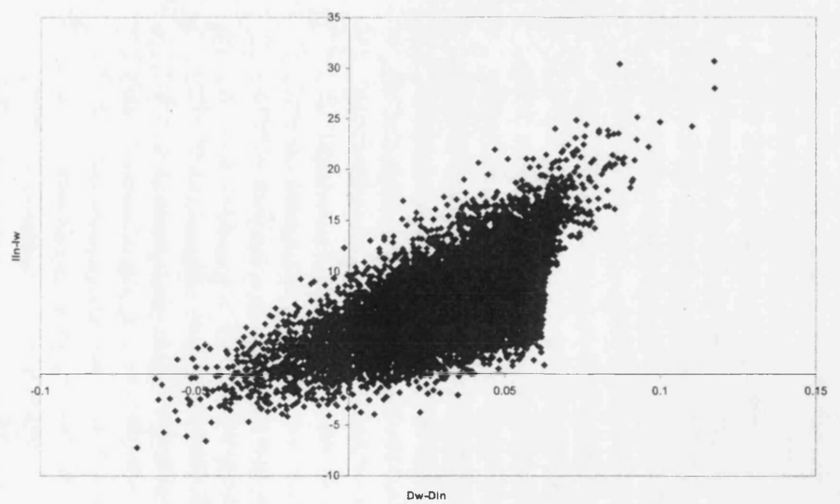


Figure 2.28: $D_w - D_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 100$).

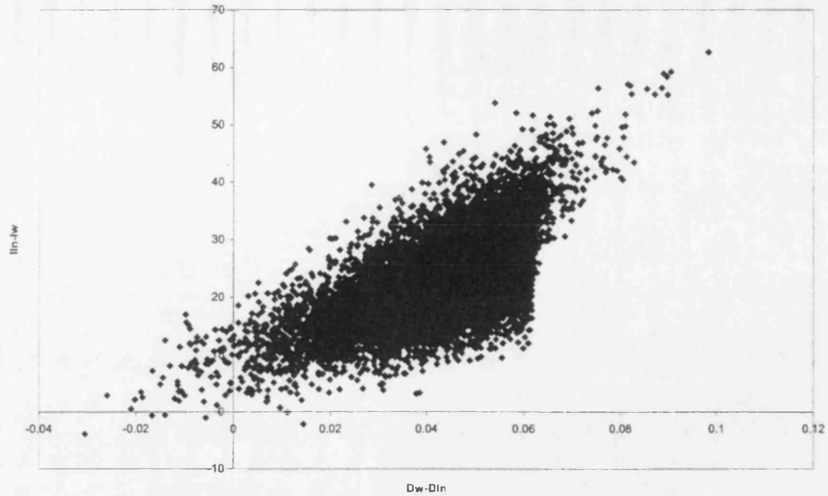


Figure 2.29: $D_w - D_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 300$).

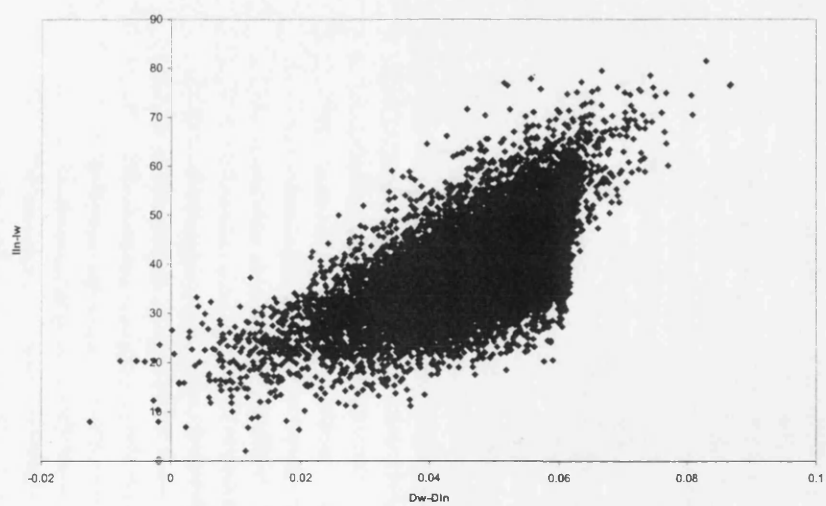


Figure 2.30: $D_w - D_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 500$).

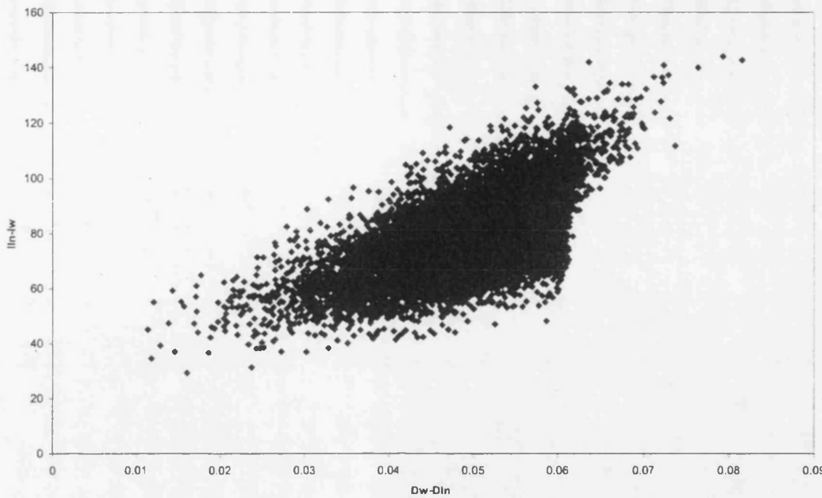


Figure 2.31: $D_w - D_{\text{ln}}$ against $l_{\text{ln}} - l_w$ for data generated from G_{ln} with $\mu = 2$ and $\sigma = 3$ ($n = 1000$).

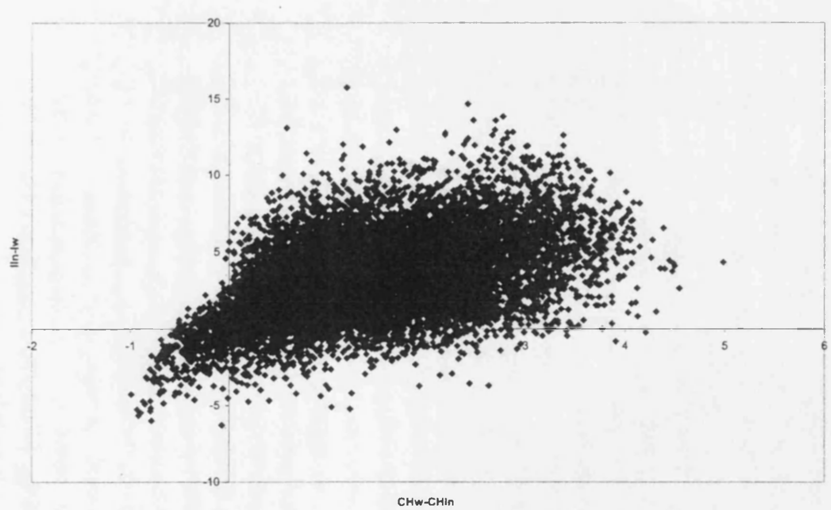


Figure 2.32: $CH_w - CH_{\text{ln}}$ against $l_{\text{ln}} - l_w$ for data generated from G_{ln} with $\mu = 2$ and $\sigma = 3$ ($n = 50$).

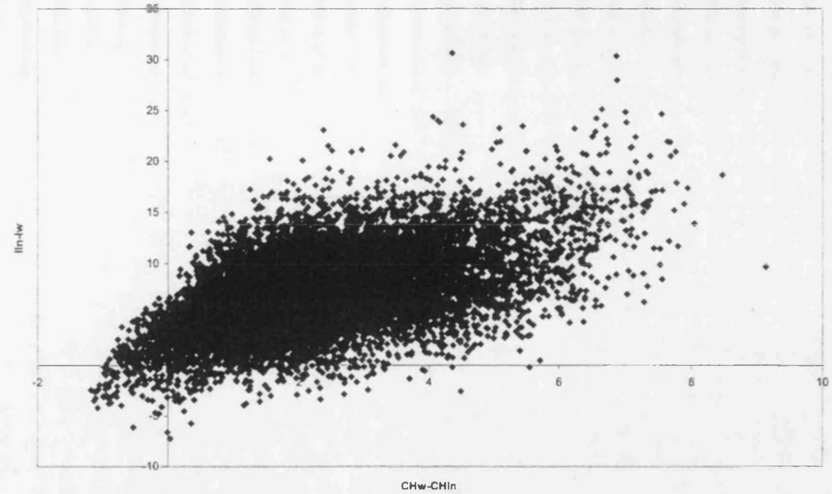


Figure 2.33: $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 100$).

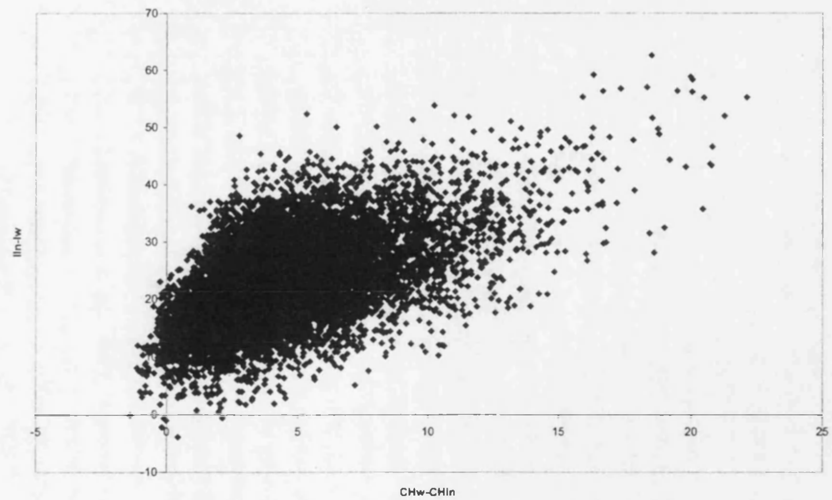


Figure 2.34: $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 300$).

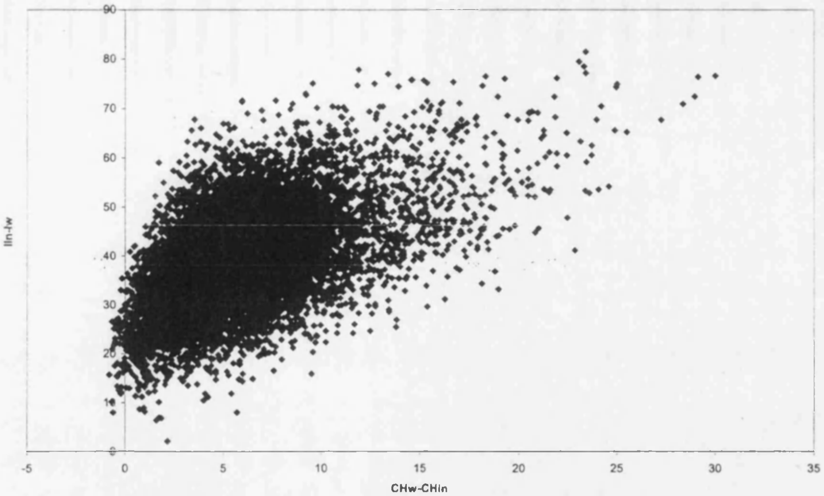


Figure 2.35: $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 500$).

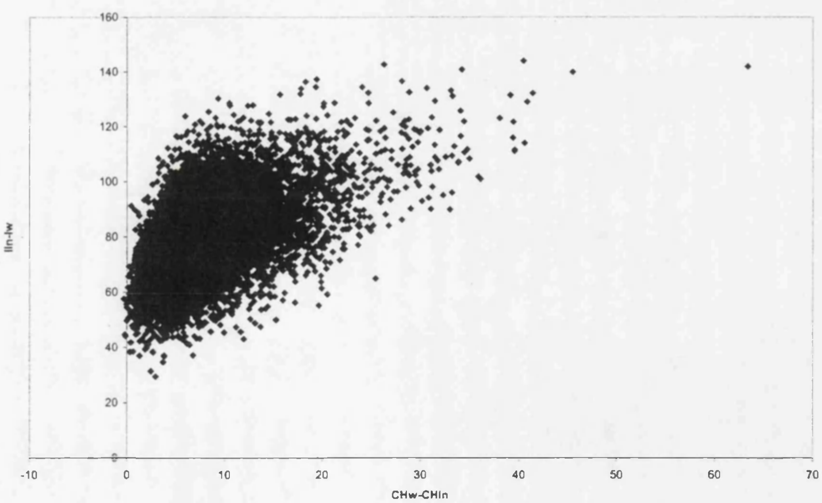


Figure 2.36: $CH_w - CH_{\ln}$ against $l_{\ln} - l_w$ for data generated from $G_{\ln}$ with $\mu = 2$ and $\sigma = 3$ ($n = 1000$).

noting the average MLEs from true and mis-specified distributions, and their corresponding standard errors. We also included details on B_{10}, and compared estimates with the true value from running simulations. Finally, we assessed the goodness of fit of both true and mis-specified distributions, via a number of techniques. These included numerical tests such as Kolmogorov-Smirnov distances, and we also examined the maximum absolute distances between fitted and empirical cumulative hazard functions.

In the next chapter, we consider the entropy function, which provides an explanation for the values onto which the MLEs from the mis-specified Weibull distribution converge as the sample size increased. This is considered first for complete data; results for censored data sets are given in Chapter 4.

Chapter 3

Maximum Likelihood: Some Theoretical Considerations

The simulation studies reported in Chapter 2 for the Burr, Gamma and Lognormal distributions indicate that, as sample size increases, the MLEs for the mis-specified Weibull distribution converge to some fixed parameter values, with decreasing standard errors. In this chapter, we derive theoretical counterparts for both mean and standard errors of the MLEs for parameters in the mis-specified model. We also consider this for the correct distribution function, where results are known. We first outline some general theory; results for specific distributions will then follow.

3.1 Analysing data using the correct distribution

Asymptotic properties of the distribution of MLEs when we assume the correct distribution function is specified are relatively well known; see, for example, Cox and Hinkley (1974). In summary, for model parameters $\pi = (\pi_1, \pi_2, \ldots, \pi_k)'$, the asymptotic distribution of $\widehat{\pi}$ will be Normal with mean π, and variance covariance matrix equal to the inverse of the Expected Fisher Information matrix (from now on, abbreviated to EFI matrix); this is symmetric, with $(i,j)^{th}$ element

$$A = -E\left[\frac{\partial^2 l(\pi)}{\partial \pi_i \partial \pi_j}\right]; \tag{3.1}$$

we exploit this symmetry, and give only the lower triangle of elements. We use this matrix to derive the asymptotic distribution of $\widehat{B}_{10}$, which, in general, is a non-linear function of $\widehat{\pi}$. Consequently, a first order Taylor series about the true parameter π (for which the mean and variance can be computed) is used to approximate the quantile; we have

$$\widehat{B}_{10} \doteqdot B_{10} + c_\pi' (\widehat{\pi} - \pi)$$

where

$$c_\pi = \frac{\partial B_{10}}{\partial \pi}.$$

For large samples, we therefore have

$$E\left[\widehat{B}_{10}\right] \doteqdot B_{10},$$

since, asymptotically, the expected values of the MLEs tend to their true values. Similarly, for the variance, we have

$$Var\left(\widehat{B}_{10}\right) \doteqdot c'_\pi A^{-1} c_\pi, \tag{3.2}$$

and further note that, in this limit, the distribution of $\widehat{B}_{10}$ is Normal with the above mean and variance. This follows from approximating the quantile as a linear combination of the MLEs. We refer to Mardia, Kent and Bibby (1995) for a further discussion on the asymptotic distribution of non-linear functions of MLEs.

We now consider the form of the EFI matrix and the distribution of $\widehat{B}_{10}$ below, for the Weibull, Burr, Gamma and Lognormal distributions.

3.1.1 The Weibull distribution

Watkins (1998) computes expectations of second derivatives from G_w given by

$$\begin{aligned}
\frac{\partial^2 l_w}{\partial \beta^2} &= -n\beta^{-2} + 2\theta^{-\beta}\ln\theta S_1(\beta) - \theta^{-\beta}(\ln\theta)^2 S_0(\beta) - \theta^{-\beta}S_2(\beta), && (3.3)\\
\frac{\partial^2 l_w}{\partial \theta^2} &= n\beta\theta^{-2} - \beta(\beta+1)\theta^{-(\beta+2)}S_0(\beta), && (3.4)\\
\frac{\partial^2 l_w}{\partial \beta \partial \theta} &= -n\theta^{-1} + \theta^{-(\beta+1)}S_0(\beta)\{1-\beta\ln\theta\} + \beta\theta^{-(\beta+1)}S_1(\beta). && (3.5)
\end{aligned}$$

Using these, we state that the distribution of $\left(\widehat{\beta},\widehat{\theta}\right)'$ is Bivariate Normal with mean $(\beta,\theta)'$ and variance covariance matrix A^{-1}, where

$$A = n\begin{bmatrix} \beta^{-2}\left\{\frac{\pi^2}{6} + (\gamma-1)^2\right\} & \\ -\theta^{-1}(1-\gamma) & \beta^2\theta^{-2} \end{bmatrix},$$

and use this result to compute the variance of $\widehat{B}_{w,10}$, from (3.2) with

$$c_\pi = \begin{pmatrix} \frac{\partial B_{w,10}}{\partial \beta} \\ \frac{\partial B_{w,10}}{\partial \theta} \end{pmatrix} = \begin{pmatrix} c_\beta \\ c_\theta \end{pmatrix} = \begin{pmatrix} \frac{-\theta(-\ln 0.9)^{\frac{1}{\beta}}\ln(-\ln 0.9)}{\beta^2} \\ (-\ln 0.9)^{\frac{1}{\beta}} \end{pmatrix}. \tag{3.6}$$

n	50	100	300	500	1000
St.err. $\left(\widehat{\beta}\right)$	0.2205	0.1559	0.0900	0.0697	0.0493
St.err. $\left(\widehat{\theta}\right)$	7.4454	5.2647	3.0396	2.3544	1.6648
St.err. $\left(\widehat{B}_{w,10}\right)$	5.3059	3.7519	2.1661	1.6779	1.1864

Table 3.1: Theoretical standard errors for the MLEs of G_w for varying n. Data is simulated from G_w with $\beta = 2$, $\theta = 100$.

We compare sample results, summarised in Table 2.5, with theoretical counterparts, when $\beta = 2$, $\theta = 100$; for this particular set of Weibull parameters, the variance covariance matrix is

$$n^{-1}\begin{bmatrix} 2^{-2}\left\{\frac{\pi^2}{6}+(\gamma-1)^2\right\} & \\ -100^{-1}(1-\gamma) & \left(\frac{2}{100}\right)^2 \end{bmatrix}^{-1} = n^{-1}\begin{bmatrix} 2.4317 & \\ 25.7022 & 2771.6622 \end{bmatrix},$$

and this matrix yields the theoretical standard errors of the Weibull MLEs and $\widehat{B}_{w,10}$ for varying sample sizes. The results are shown in Table 3.1; we see excellent agreement between theoretical and sample values of the Weibull MLEs and $\widehat{B}_{w,10}$. This is true across all sample sizes.

3.1.2 The Burr distribution

To calculate the EFI matrix from G_b, we need the expected values of the six second partial derivatives given by (2.9). Again, we refer to Watkins (1997) for details on this, and list the required expectations below. From (2.9), we require

$$E\left[\frac{\left(\frac{Y}{\phi}\right)^{\tau}\ln\left(\frac{Y}{\phi}\right)}{1+\left(\frac{Y}{\phi}\right)^{\tau}}\right] = \frac{1-\gamma-\Psi(\alpha)}{\tau(\alpha+1)}, \tag{3.7}$$

$$E\left[\frac{\left(\frac{Y}{\phi}\right)^{\tau}\left\{\ln\left(\frac{Y}{\phi}\right)\right\}^2}{\left\{1+\left(\frac{Y}{\phi}\right)^{\tau}\right\}^2}\right] = \frac{\alpha\left\{\frac{\pi^2}{6}+\gamma^2-2\gamma+2(\gamma-1)\Psi(\alpha+1)+\Psi(\alpha+1)^2+\Psi'(\alpha+1)\right\}}{\tau^2(\alpha+1)(\alpha+2)}, \tag{3.8}$$

$$E\left[\frac{\left(\frac{Y}{\phi}\right)^{\tau}}{1+\left(\frac{Y}{\phi}\right)^{\tau}}\right] = \frac{1}{\alpha+1}, \tag{3.9}$$

$$E\left[\frac{\left(\frac{Y}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{Y}{\phi}\right)^{\tau}\right\}^{2}}\right]=\frac{\alpha}{(\alpha+1)(\alpha+2)}, \tag{3.10}$$

and

$$E\left[\frac{\left(\frac{Y}{\phi}\right)^{\tau}\ln\left(\frac{Y}{\phi}\right)}{\left\{1+\left(\frac{Y}{\phi}\right)^{\tau}\right\}^{2}}\right]=\frac{\alpha\left\{1-\gamma-\Psi(\alpha+1)\right\}}{\tau(\alpha+1)(\alpha+2)}. \tag{3.11}$$

Using these, we write the EFI matrix for the Burr distribution as

$$A=n\begin{bmatrix} \tau^{-2}+\frac{n\alpha\left[\frac{\pi^2}{6}+\gamma^2-2\gamma+2(\gamma-1)\Psi(\alpha+1)+\{\Psi(\alpha+1)\}^2+\Psi'(\alpha+1)\right]}{\tau^2(\alpha+2)} & & \\ \frac{\{1-\gamma-\Psi(\alpha)\}}{\tau(\alpha+1)} & \alpha^{-2} & \\ -\frac{\alpha\{1-\gamma-\Psi(\alpha+1)\}}{\phi(\alpha+2)} & -\frac{\tau}{\phi(\alpha+1)} & \frac{\alpha\tau^2}{\phi^2(\alpha+2)} \end{bmatrix}.$$

We use this to compute the theoretical variance of $\widehat{B}_{b,10}$ from (3.2) with

$$c_{\pi}=\begin{pmatrix} c_{\tau}\\ c_{\alpha}\\ c_{\phi}\end{pmatrix}=\begin{pmatrix} \frac{-\phi\left(0.9^{-\frac{1}{\alpha}}-1\right)^{\frac{1}{\tau}}\ln\left(0.9^{-\frac{1}{\alpha}}-1\right)}{\tau^2}\\ \frac{\phi\left(0.9^{-\frac{1}{\alpha}}-1\right)^{\frac{1}{\tau}-1}\left(0.9^{-\frac{1}{\alpha}}\ln 0.9\right)}{\tau\alpha^2}\\ \left(0.9^{-\frac{1}{\alpha}}-1\right)^{\frac{1}{\tau}}\end{pmatrix}.$$

We compute the theoretical values of the asymptotic variance covariance matrix of the Burr MLEs, and compare with simulated counterparts; the theoretical values can be obtained using Mathematica or SAS. We take $\tau=3$, $\alpha=4$ and $\phi=100$, so that

$$\begin{aligned} A &= n\begin{bmatrix} 3^{-2}+\frac{4\begin{bmatrix}\frac{\pi^2}{6}+\gamma^2-2\gamma\\ +2(\gamma-1)\Psi(4+1)\\ +\{\Psi(4+1)\}^2+\Psi'(4+1)\end{bmatrix}}{\tau^2(4+2)} & & \\ \frac{\{1-\gamma-\Psi(4)\}}{3(4+1)} & 4^{-2} & \\ -\frac{4\{1-\gamma-\Psi(4+1)\}}{100(4+2)} & -\frac{3}{100(4+1)} & \frac{4\times 3^2}{100^2(4+2)} \end{bmatrix}\\ &= n\begin{bmatrix} 0.2622 & & \\ -0.0556 & 0.0625 & \\ 0.0072 & -0.0060 & 0.0006 \end{bmatrix}, \end{aligned}$$

n	50	100	300	500	1000
St.err.$(\widehat{\tau})$	0.5583	0.3948	0.2279	0.1765	0.1248
St.err.$(\widehat{\alpha})$	4.6747	3.3055	1.9084	1.4783	1.0453
St.err.$\left(\widehat{\phi}\right)$	52.5738	37.1753	21.4631	16.6253	11.7558
St.err.$\left(\widehat{B}_{b,10}\right)$	3.6844	2.6053	1.5041	1.1651	0.8239

Table 3.2: Theoretical standard errors for the MLEs of G_b for varying n. Data is simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

and so

$$Var\begin{pmatrix}\widehat{\tau}\\ \widehat{\alpha}\\ \widehat{\phi}\end{pmatrix} = n^{-1}\begin{bmatrix} 15.5843 & & \\ -103.8953 & 1092.6351 & \\ -1226.5413 & 12176.9419 & 138200.0084 \end{bmatrix}.$$

Table 3.2 summarises the theoretical standard errors of the MLEs and $\widehat{B}_{b,10}$ when $n =$ 50, 100, 300, 500 and 1000; we may compare these to sample counterparts shown in Table 2.6. We observe considerable differences between sample and theoretical standard errors of Burr MLEs, especially for $\widehat{\alpha}$ and $\widehat{\phi}$, and small sample sizes; some intuitive explanation for this has been provided in Chapter 2, when we compared average MLEs with their true values, and observed averages for $\widehat{\alpha}$ and $\widehat{\phi}$ considerably larger than expected. This will also affect the corresponding standard errors. We run a further set of simulations, this time with $n = 2000$. The theoretical variance covariance matrix is given by

$$\begin{bmatrix} 0.0078 & & \\ -0.0581 & 0.7592 & \\ -0.6502 & 7.7668 & 81.6661 \end{bmatrix},$$

which we compare to the sample counterpart

$$\begin{bmatrix} 0.0078 & & \\ -0.0515 & 0.5417 & \\ -0.6089 & 6.0373 & 68.5361 \end{bmatrix}.$$

Increasing the sample size has improved the agreement between theoretical and sample variances and covariances for the parameters quite considerably. When we examine results for the quantile $\widehat{B}_{b,10}$, we observe sample means approaching the true value of 29.8848 as the sample size increases. The agreement between theoretical and observed standard errors is generally quite good, even for small sample sizes. This is somewhat surprising, given the poor agreement between sample and theoretical standard errors of the Burr MLEs, which then contribute to the standard error of this quantile.

3.1.3 The Gamma distribution

We consider (3.1) for G_g. Differentiating (2.12) with respect to α gives

$$\frac{\partial^2 l_g}{\partial \alpha^2} = -2\alpha^{-3} S_0(1) + n\tau\alpha^{-2};$$

similarly, we differentiate (2.13) with respect to τ to obtain

$$\frac{\partial^2 l_g}{\partial \tau^2} = -n\Psi'(\tau);$$

finally, we see that

$$\frac{\partial^2 l_g}{\partial \alpha \partial \tau} = -n\alpha^{-1}.$$

Hence, by taking expected values of these second derivatives we have

$$\begin{aligned} -E\left[\frac{\partial^2 l_g}{\partial \tau^2}\right] &= n\Psi'(\tau), \\ -E\left[\frac{\partial^2 l_g}{\partial \alpha \partial \tau}\right] &= n\tau^{-1}, \end{aligned}$$

and

$$-E\left[\frac{\partial^2 l_g}{\partial \alpha^2}\right] = 2n\alpha^{-3}E[Y] - n\tau\alpha^{-2}.$$

We note that

$$E[Y^m] = \frac{\alpha^m \Gamma(m+\tau)}{\Gamma(\tau)}, \tag{3.12}$$

and so

$$E[Y] = \alpha\tau.$$

Hence,

$$-E\left[\frac{\partial^2 l_g}{\partial \alpha^2}\right] = n\tau\alpha^{-2}.$$

Thus, for large sample sizes, $(\widehat{\tau}, \widehat{\alpha})'$ will be Normally distributed with mean $(\tau, \alpha)'$ and variance covariance matrix

$$A^{-1} = n^{-1}\begin{bmatrix} \Psi'(\tau) & \\ \alpha^{-1} & \tau\alpha^{-2} \end{bmatrix}^{-1}.$$

n	50	100	300	500	1000
St.err.$(\widehat{\tau})$	0.5678	0.4029	0.2326	0.1802	0.1274
St.err.$(\widehat{\alpha})$	10.3370	7.3094	4.2201	3.2688	2.3114
St.err.$\left(\widehat{B}_{g,10}\right)$	8.3188	5.8823	3.3961	2.6306	1.8601

Table 3.3: Theoretical standard errors for the MLEs of G_g for varying n. Data is simulated from G_g with $\tau = 3$, $\alpha = 50$.

Unlike the Weibull and Burr distributions, we cannot write down analytic expressions for the mean and variance of $\widehat{B}_{g,10}$. However, we can compute this quantile using Mathematica to solve

$$0.1 = \frac{\Gamma\left(\frac{y}{\alpha}, \tau\right)}{\Gamma(\tau)}$$

for y, using the Inverse Gamma Regularised function, for given τ, α. Thus, $B_{g,10}$ is obtained from

$$\texttt{InverseGammaRegularized}[\tau, 0, 0.1] * \alpha.$$

Mathematica can also compute derivatives of $B_{g,10}$, required for (3.2). Although further analytical progress is possible, ultimately, their evaluation must be numerical, and we therefore omit these simplifications. We compute the theoretical standard errors of the MLEs and $\widehat{B}_{g,10}$ for varying n. For $\tau = 3$, $\alpha = 50$, we have

$$A = n \begin{bmatrix} \Psi'(3) & 50^{-1} \\ 50^{-1} & 3 \times 50^{-2} \end{bmatrix} = n \begin{bmatrix} 0.3949 & \\ 0.02 & 0.0012 \end{bmatrix},$$

so that

$$Var \begin{pmatrix} \widehat{\tau} \\ \widehat{\alpha} \end{pmatrix} = n^{-1} \begin{bmatrix} 16.2336 & \\ -270.5595 & 5342.6591 \end{bmatrix}.$$

We summarise theoretical results in Table 3.3, and compare with simulated counterparts shown in Table 2.9. We see good agreement between observed and expected results, especially for larger sample sizes. This is true for both the Gamma MLEs and $\widehat{B}_{g,10}$. We also note that the theoretical mean of 55.1033 is close to the sample means, even for small sample sizes.

3.1.4 The Lognormal distribution

The Lognormal distribution differs from other reliability distributions, in that explicit expressions exist both for the MLEs of this distribution, and the elements in (3.1). The score for this distribution is obtained on differentiating (2.16), which gives

$$\frac{\partial l_{\ln}}{\partial \mu} = \frac{\sum_{i=1}^{n} (\ln Y_i - \mu)}{\sigma^2},$$

and

$$\frac{\partial l_{\ln}}{\partial \sigma} = \frac{-n}{\sigma} + \frac{\sum_{i=1}^{n} (\ln Y_i - \mu)^2}{\sigma^3}.$$

The second derivatives are then

$$\frac{\partial^2 l_{\ln}}{\partial \mu^2} = \frac{-n}{\sigma^2},$$

$$\frac{\partial^2 l_{\ln}}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3\sum_{i=1}^{n} (\ln Y_i - \mu)^2}{\sigma^4},$$

and

$$\frac{\partial^2 l_{\ln}}{\partial \mu \partial \sigma} = \frac{-2\sum_{i=1}^{n} (\ln Y_i - \mu)}{\sigma^3}.$$

We take expected values of these to obtain the variance covariance matrix for the MLEs of the Lognormal distribution. These are given by

$$-E\left[\frac{\partial^2 l_{\ln}}{\partial \mu^2}\right] = \frac{n}{\sigma^2},$$

and

$$-E\left[\frac{\partial^2 l_{\ln}}{\partial \sigma^2}\right] = \frac{-n}{\sigma^2} + \frac{3E\left[\sum_{i=1}^{n} (\ln Y_i - \mu)^2\right]}{\sigma^4},$$

where, on exploiting the link between Normal and Lognormal distributions, we write

$$E\left[\sum_{i=1}^{n} (\ln Y_i - \mu)^2\right] = n\sigma^2.$$

Thus,

$$-E\left[\frac{\partial^2 l_{\ln}}{\partial \sigma^2}\right] = \frac{2n}{\sigma^2}.$$

Finally, we have

$$-E\left[\frac{\partial^2 l_{\ln}}{\partial \mu \partial \sigma}\right] = \frac{2E\left[\sum_{i=1}^{n} (\ln Y_i - \mu)\right]}{\sigma^3} = 0.$$

n	50	100	300	500	1000
St.err.$(\widehat{\mu})$	0.4243	0.3	0.1732	0.1342	0.0949
St.err.$(\widehat{\sigma})$	0.3	0.2121	0.1225	0.0949	0.0671
St.err.$\left(\widehat{B}_{\ln,10}\right)$	0.0905	0.0640	0.0370	0.0286	0.0202

Table 3.4: Theoretical standard errors for the MLEs of $G_{\ln}$ for varying n. Data is simulated from $G_{\ln}$ with $\mu = 2$, $\sigma = 3$.

Thus, $(\widehat{\mu}, \widehat{\sigma})'$ will have a bivariate Normal distribution with mean vector $(\mu, \sigma)'$ and variance covariance matrix

$$A^{-1} = n^{-1} \begin{bmatrix} \sigma^2 & \\ 0 & \frac{\sigma^2}{2} \end{bmatrix}.$$

When we compute the mean and variance of $\widehat{B}_{\ln,10}$, then, as with the Gamma distribution, we are not able to write down a theoretical expression for this quantile. We can, however, use Mathematica to compute numerical values, by solving

$$0.1 = \Phi\left(\frac{\ln y - \mu}{\sigma}\right)$$

for y. The command `Quantile[LogNormalDistribution]` provides solutions to such an equation for a particular set of Lognormal parameters. Thus, $\widehat{B}_{\ln,10}$ will be Normally distributed with mean $E\left[\widehat{B}_{\ln,10}\right] = B_{\ln,10}$ and variance given by (3.2) with

$$c_{\pi} = \begin{pmatrix} c_{\mu} \\ c_{\sigma} \end{pmatrix} = \begin{pmatrix} \frac{\partial B_{\ln,10}}{\partial \mu} \\ \frac{\partial B_{\ln,10}}{\partial \sigma} \end{pmatrix},$$

where these derivatives can also be obtained using Mathematica. Again, as with the Gamma distribution, some simplification for these derivatives is possible, but will not be considered here. We compare theoretical results with our simulated values in Chapter 2, noting that the variance covariance matrix for our particular set of parameters is

$$n^{-1} \begin{bmatrix} 3^2 & \\ 0 & \frac{3^2}{2} \end{bmatrix} = n^{-1} \begin{bmatrix} 9 & \\ 0 & \frac{9}{2} \end{bmatrix}.$$

We use this matrix to summarise the results in Table 3.4 for varying n; we compare this with the values in Table 2.11. As previously noted, the results show very good agreement between observed and expected values, and this improves as the sample size increases. The average of $\widehat{B}_{\ln,10}$ also seems to be tending to its true value of 0.1581 as n increases.

3.2 Analysing data using the incorrect distribution

When we fit the incorrect distribution to data, the distribution of the mis-specified MLEs will change to compensate for this. Work in this important area can be traced back at least as far as Cox (1961), who considers the expected score equations, and shows how to obtain asymptotic parameter means from the mis-specified model; these values can be defined in terms of the true parameter values. The asymptotic distribution of the mis-specified MLEs is also considered. As previously discussed, we always consider the Weibull distribution as the mis-specified model. Consequently, now the asymptotic distribution of $\left(\widehat{\beta}, \widehat{\theta}\right)'$ will be Normal with mean $(\beta_0, \theta_0)'$, where (β_0, θ_0) are the roots of the expected score equations

$$\begin{aligned} E_t \left[\frac{\partial l_w}{\partial \beta} \right] &= 0, \\ E_t \left[\frac{\partial l_w}{\partial \theta} \right] &= 0. \end{aligned}$$

Here, E_t indicates that expectations are taken with respect to the true distribution function with known parameter values. Equivalently, we maximise $E_t\left[l_w\right]$, the entropy function; see Shannon (1948) and Jaynes (1957). From (2.1), we obtain

$$E_t = E_t[l_w] = n \ln \beta - n\beta \ln \theta + (\beta - 1) E_t[S_e] - \theta^{-\beta} E_t[S_0\left(\beta\right)],$$

and, via independence, this can be expressed as

$$n \left\{ \ln \beta - \beta \ln \theta + (\beta - 1)\, E_t \left[\ln Y\right] - \theta^{-\beta} E_t \left[Y^{\beta} \right] \right\}. \tag{3.13}$$

We differentiate this function to obtain the entropy score function with elements

$$\frac{\partial E_t}{\partial \beta} = n \left[\beta^{-1} - \ln \theta + E_t \left[\ln Y\right] + \theta^{-\beta} \left\{ \ln \theta E_t \left[Y^{\beta} \right] - E_t \left[Y^{\beta} \ln Y \right] \right\} \right],$$

and

$$\frac{\partial E_t}{\partial \theta} = n \left\{ -\beta \theta^{-1} + \beta \theta^{-\beta - 1} E_t \left[Y^{\beta} \right] \right\}, \tag{3.14}$$

and note that we can equate (3.14) to zero to obtain

$$\theta = \left\{ E_t \left[Y^{\beta} \right] \right\}^{\frac{1}{\beta}}. \tag{3.15}$$

Thus, the profile entropy function is

$$E_t^* = n \left\{ \ln \beta - \ln E_t \left[Y^{\beta} \right] + (\beta - 1)\, E_t \left[\ln Y\right] - 1 \right\}, \tag{3.16}$$

and the profile entropy score function is given by

$$\frac{dE_t^*}{d\beta} = n\left\{\beta^{-1} - \frac{E_t\left[Y^\beta \ln Y\right]}{E_t\left[Y^\beta\right]} + E_t\left[\ln Y\right]\right\}, \tag{3.17}$$

with derivative

$$\frac{d^2 E_t^*}{d\beta^2} = -n\left\{\beta^{-2} + \frac{E_t\left[Y^\beta (\ln Y)^2\right] E_t\left[Y^\beta\right] - \left(E_t\left[Y^\beta \ln Y\right]\right)^2}{\left(E_t\left[Y^\beta\right]\right)^2}\right\}.$$

We can locate the root of (3.17) either numerically (for instance, using Newton-Raphson) or graphically; we denote this root by β_0, and obtain θ_0 on taking $\beta = \beta_0$ in (3.15). We use these entropy values to derive the variance covariance matrix of the mis-specified MLEs. This matrix is given by

$$A^{-1} V A^{-1}, \tag{3.18}$$

(evaluated at β_0, θ_0) where, now,

$$A = \begin{bmatrix} -E_t\left[\frac{\partial^2 l_w}{\partial \beta^2}\right] & \\ -E_t\left[\frac{\partial^2 l_w}{\partial \beta \partial \theta}\right] & -E_t\left[\frac{\partial^2 l_w}{\partial \theta^2}\right] \end{bmatrix},$$

and

$$V = \begin{bmatrix} Var_t\left(\frac{\partial l_w}{\partial \beta}\right) & \\ Cov_t\left(\frac{\partial l_w}{\partial \beta}, \frac{\partial l_w}{\partial \theta}\right) & Var_t\left(\frac{\partial l_w}{\partial \theta}\right) \end{bmatrix}.$$

This variance covariance matrix reduces to (3.1) when $t = w$. We use (3.3), (3.4) and (3.5) to write the elements of the matrix A as

$$-E_t\left[\frac{\partial^2 l_w}{\partial \beta^2}\right] = n\left\{\begin{array}{c} \beta^{-2} - 2\theta^{-\beta} \ln\theta E_t\left[Y^\beta \ln Y\right] + \theta^{-\beta} (\ln\theta)^2 E_t\left[Y^\beta\right] \\ +\theta^{-\beta} E_t\left[Y^\beta (\ln Y)^2\right] \end{array}\right\}, \tag{3.19}$$

$$-E_t\left[\frac{\partial^2 l_w}{\partial \theta^2}\right] = -n\beta\theta^{-2}\left\{1 - (\beta+1)\theta^{-\beta} E_t\left[Y^\beta\right]\right\}, \tag{3.20}$$

$$-E_t\left[\frac{\partial^2 l_w}{\partial \beta \partial \theta}\right] = n\theta^{-1}\left\{1 - \theta^{-\beta}(1 - \beta\ln\theta) E_t\left[Y^\beta\right] - \beta\theta^{-\beta} E_t\left[Y^\beta \ln Y\right]\right\}. \tag{3.21}$$

From (2.2) and (2.3), we evaluate the elements of V as follows: we first need

$$Var_t\left(\frac{\partial l_w}{\partial \beta}\right) = M' Var_t \begin{pmatrix} S_e \\ S_0(\beta) \\ S_1(\beta) \end{pmatrix} M, \tag{3.22}$$

where

$$M' = \left(1 \quad \theta^{-\beta} \ln \theta \quad -\theta^{-\beta} \right),$$

and the elements of the matrix in this variance are

$$Var_t \left(S_e \right) = n \left\{ E_t \left[\left(\ln Y \right)^2 \right] - E_t \left[\ln Y \right]^2 \right\},$$

$$Var_t \left\{ S_0 \left(\beta \right) \right\} = n \left\{ E_t \left[Y^{2\beta} \right] - E_t \left[Y^{\beta} \right]^2 \right\},$$

$$Var_t \left\{ S_1 \left(\beta \right) \right\} = n \left\{ E_t \left[Y^{2\beta} \left(\ln Y \right)^2 \right] - E_t \left[Y^{\beta} \ln Y \right]^2 \right\},$$

$$Cov_t \left\{ S_e, S_0 \left(\beta \right) \right\} = n \left\{ E_t \left[Y^{\beta} \ln Y \right] - E_t \left[Y^{\beta} \right] E_t \left[\ln Y \right] \right\},$$

$$Cov_t \left\{ S_e, S_1 \left(\beta \right) \right\} = n \left\{ E_t \left[Y^{\beta} \left(\ln Y \right)^2 \right] - E_t \left[\ln Y \right] E_t \left[Y^{\beta} \ln Y \right] \right\},$$

and

$$Cov_t \left\{ S_0 \left(\beta \right), S_1 \left(\beta \right) \right\} = n \left\{ E_t \left[Y^{2\beta} \ln Y \right] - E_t \left[Y^{\beta} \right] E_t \left[Y^{\beta} \ln Y \right] \right\}.$$

For V, we also need

$$Var_t \left(\frac{\partial l_w}{\partial \theta} \right) = \beta^2 \theta^{-2(\beta+1)} Var_t \left\{ S_0 \left(\beta \right) \right\}, \tag{3.23}$$

and

$$Cov_t \left(\frac{\partial l_w}{\partial \beta}, \frac{\partial l_w}{\partial \theta} \right) = \beta \theta^{-\beta-1} M' \begin{pmatrix} Cov_t \left\{ S_e, S_0 \left(\beta \right) \right\} \\ Var_t \left\{ S_0 \left(\beta \right) \right\} \\ Cov_t \left\{ S_1 \left(\beta \right), S_0 \left(\beta \right) \right\} \end{pmatrix}, \tag{3.24}$$

where the elements in this vector are given above. Thus, we require four different expectations to evaluate the entropy function and variance covariance matrix of the mis-specified MLEs. These are $E_t \left[\left(\ln Y \right)^m \right]$, $E_t \left[Y^{m\beta} \right]$, $E_t \left[Y^{m\beta} \ln Y \right]$ and $E_t \left[Y^{m\beta} \left(\ln Y \right)^2 \right]$. We also consider the asymptotic distribution of $\widehat{B}_{w,10}$; again, we use a linear approximation to the quantile based on a first order Taylor series centered on $(\beta_0, \theta_0)'$. Hence, we have

$$\widehat{B}_{w,10} \simeq B_{w,10}(\beta_0, \theta_0) + \begin{pmatrix} c_\beta & c_\theta \end{pmatrix} \begin{pmatrix} \widehat{\beta} - \beta_0 \\ \widehat{\theta} - \theta_0 \end{pmatrix},$$

where $B_{w,10}(\beta_0, \theta_0)$ denotes the quantile evaluated at the entropy values, and $\begin{pmatrix} c_\beta & c_\theta \end{pmatrix}'$ is given by (3.6). Thus, we have

$$E\left[\widehat{B}_{w,10}\right] \simeq B_{w,10}(\beta_0, \theta_0) = \theta_0 (-\ln 0.9)^{\frac{1}{\beta_0}},$$

and

$$Var\left(\widehat{B}_{w,10}\right) \simeq \begin{pmatrix} c_\beta & c_\theta \end{pmatrix} A^{-1} V A^{-1} \begin{pmatrix} c_\beta \\ c_\theta \end{pmatrix}. \tag{3.25}$$

3.2.1 Entropy for fitting G_w to data from G_b

We first list the required expectations to compute entropy values from the Weibull distribution. From Watkins (1997), we have

$$E_b[Y^{m\beta}] = \frac{\phi^{m\beta} P_m}{\Gamma(\alpha)}, \tag{3.26}$$

where

$$P_j = \Gamma\left(\frac{j\beta}{\tau} + 1\right) \Gamma\left(\alpha - \frac{j\beta}{\tau}\right) \tag{3.27}$$

is defined for $\alpha > \beta\tau^{-1}$. This condition must be satisfied to ensure the fourth moment exists; see Tadikamalla (1980) for further details. For future reference, we note that

$$P_j' = \frac{dPj}{d\beta} = j\tau^{-1} P_j \left\{ \Psi\left(\frac{j\beta}{\tau} + 1\right) - \Psi\left(\alpha - \frac{j\beta}{\tau}\right) \right\}, \tag{3.28}$$

and

$$P_j'' = \frac{dP_j'}{d\beta} = \left(j\tau^{-1}\right)^2 P_j \left[\begin{array}{c} \left\{ \Psi\left(\frac{j\beta}{\tau} + 1\right) - \Psi\left(\alpha - \frac{j\beta}{\tau}\right) \right\}^2 \\ + \Psi'\left(\frac{j\beta}{\tau} + 1\right) + \Psi'\left(\alpha - \frac{j\beta}{\tau}\right) \end{array} \right]. \tag{3.29}$$

We differentiate (3.26) with respect to β to obtain

$$E_b\left[Y^{m\beta} \ln Y\right] = \frac{\phi^{m\beta}}{\Gamma(\alpha)} \left\{ \ln \phi P_m + \frac{P_m'}{m} \right\}, \tag{3.30}$$

and setting $\beta = 0$ in (3.30), we have

$$E_b[\ln Y] = \ln\phi + \tau^{-1}\{\Psi(1) - \Psi(\alpha)\} = \ln\phi - \tau^{-1}\{\gamma + \Psi(\alpha)\}. \tag{3.31}$$

We also require

$$E_b\left[Y^{m\beta}(\ln Y)^2\right] = \frac{\phi^{m\beta}}{\Gamma(\alpha)}\left\{(\ln\phi)^2 P_m + \frac{2\ln\phi P_m'}{m} + \frac{P_m''}{m^2}\right\}, \tag{3.32}$$

and we see that if we put $\beta = 0$ in (3.32) then we obtain an expression for $E_b\left[(\ln Y)^2\right]$, given by

$$\left[\ln\phi - \tau^{-1}\{\gamma + \Psi(\alpha)\}\right]^2 + \tau^{-2}\left\{\Psi'(1) + \Psi'(\alpha)\right\}. \tag{3.33}$$

Using these expectations and (3.13), we can derive the entropy function for the Weibull distribution. This is given by

$$E_b = E_b[l_w] = n\left[\begin{array}{c} \ln\beta + \beta\ln\left(\frac{\phi}{\theta}\right) - \frac{\beta}{\tau}\{\gamma + \Psi(\alpha)\} - \left\{\frac{\left(\frac{\phi}{\theta}\right)^\beta P_1}{\Gamma(\alpha)}\right\} \\ -\ln\phi + \tau^{-1}\{\gamma + \Psi(\alpha)\} \end{array}\right],$$

which, we note, is a function of β and θ; α, τ and ϕ are effectively constants. On using (3.15) and (3.26) with $m = 1$, we have

$$\theta = \phi\left\{\frac{P_1}{\Gamma(\alpha)}\right\}^{\frac{1}{\beta}}, \tag{3.34}$$

and, from (3.16), we obtain

$$E_b^* = n\left\{\begin{array}{c} \ln\beta - \ln P_1 - \frac{\beta}{\tau}\{\gamma + \Psi(\alpha)\} + \ln\Gamma(\alpha) \\ -\ln\phi + \tau^{-1}\{\gamma + \Psi(\alpha)\} - 1 \end{array}\right\}.$$

As in the general case, the maximum of this function can be located from a plot; equivalently, we can compute the first and second derivatives

$$\frac{dE_b^*}{d\beta} = n\left\{\beta^{-1} - \frac{P_1'}{P_1} - \tau^{-1}\{\gamma + \Psi(\alpha)\}\right\}, \tag{3.35}$$

and

$$\frac{d^2E_b^*}{d\beta^2} = -n\left\{\beta^{-2} + \frac{P_1P_1'' - \left(P_1'\right)^2}{P_1^2}\right\}, \tag{3.36}$$

and use Newton-Raphson to locate the root of (3.35). Using the latter approach, we now compute β_0, θ_0 for $\tau = 3$, $\alpha = 4$, $\phi = 100$; this enables comparison with Table 2.6. After just five iterations, we obtain $\beta_0 = 2.5528$, $\theta_0 = 67.2620$; we note that these values are very close to the sample means of the MLEs obtained in Table 2.6, especially for large sample sizes.

Now that we are able to obtain theoretical counterparts to the MLEs of the Weibull distribution under mis-specification, we can obtain the asymptotic distribution of these estimates. We address this in the next section.

The variance structure of the mis-specified MLEs

We now derive the distribution of the Weibull MLEs, after this distribution has been wrongly fitted to data with an underlying Burr model. We first consider expected values of second derivatives given by (3.19), (3.20) and (3.21). Using the appropriate expectations, we can write the matrix A in (3.18) as

$$n\left[\begin{array}{cc} \beta_0^{-2} + \frac{\theta_0^{-\beta_0}\phi^{\beta_0}}{\Gamma(\alpha)}\left\{\rho^2 P_1 + 2\rho P_1' + P_1''\right\} & \\ \theta_0^{-1}\left\{1 - \frac{\theta_0^{-\beta_0}\phi^{\beta_0}(1+\beta_0\rho)P_1}{\Gamma(\alpha)} - \frac{\beta_0\theta^{-\beta_0}\phi^{\beta_0}P_1'}{\Gamma(\alpha)}\right\} & \frac{\beta_0(\beta_0+1)\theta_0^{-\beta_0-2}\phi^{\beta_0}P_1}{\Gamma(\alpha)} - \beta_0\theta_0^{-2} \end{array}\right],$$

where

$$\rho = \ln\left(\frac{\phi}{\theta}\right).$$

In order to calculate the distribution of the mis-specified MLEs, we also require the variance covariance matrix of the Weibull score. By examining (3.22), (3.23) and (3.24), we list the functions that appear in the elements of this matrix:

$$\begin{aligned}
Var_b\left(S_e\right) &= n\left\{\frac{\Psi'(1)+\Psi'(\alpha)}{\tau^2}\right\}, \\
Var_b\left\{S_0\left(\beta\right)\right\} &= \frac{n\phi^{2\beta}}{\Gamma(\alpha)}\left\{P_2 - \frac{(P_1)^2}{\Gamma(\alpha)}\right\}, \\
Var_b\left\{S_1\left(\beta\right)\right\} &= \frac{n\phi^{2\beta}}{\Gamma(\alpha)}\left[\begin{array}{c}(\ln\phi)^2\left\{P_2 - \frac{(P_1)^2}{\Gamma(\alpha)}\right\} + \ln\phi\left\{P_2' - \frac{2P_1P_1'}{\Gamma(\alpha)}\right\} \\ +\frac{P_2''}{4} - \frac{\left(P_1'\right)^2}{\Gamma(\alpha)}\end{array}\right], \\
Cov_b\left\{S_e, S_0\left(\beta\right)\right\} &= \frac{n\phi^{\beta}}{\Gamma(\alpha)}\left[P_1' + P_1\left\{\frac{\gamma+\Psi(\alpha)}{\tau}\right\}\right], \\
Cov_b\left\{S_e, S_1\left(\beta\right)\right\} &= \frac{n\phi^{\beta}}{\Gamma(\alpha)}\left[\begin{array}{c}\ln\phi P_1' + P_1'' + \ln\phi\left\{\frac{\gamma+\Psi(\alpha)}{\tau}\right\}P_1 \\ +\left\{\frac{\gamma+\Psi(\alpha)}{\tau}\right\}P_1'\end{array}\right], \\
Cov_b\left\{S_0\left(\beta\right), S_1\left(\beta\right)\right\} &= \frac{n\phi^{2\beta}}{\Gamma(\alpha)}\left\{\ln\phi P_2 + \frac{P_2'}{2} - \frac{\ln\phi(P_1)^2}{\Gamma(\alpha)} - \frac{P_1P_1'}{\Gamma(\alpha)}\right\}.
\end{aligned}$$

We can now list the elements of the variance covariance matrix of the Weibull score functions. Using (3.22), we first consider $Var_b\left(\frac{\partial l_w}{\partial\beta}\right)$. This is given by

$$n\left\{\begin{array}{c} \left(\frac{\Psi'(1)+\Psi'(\alpha)}{\tau^2}\right)+\frac{\left(\frac{\phi}{\theta}\right)^{2\beta}\rho^2P_2}{\Gamma(\alpha)}-\frac{\left(\frac{\phi}{\theta}\right)^{2\beta}\rho^2\{P_1\}^2}{\{\Gamma(\alpha)\}^2} \\ +\frac{\left(\frac{\phi}{\theta}\right)^{2\beta}\rho P_2'}{\Gamma(\alpha)}+\frac{\left(\frac{\phi}{\theta}\right)^{2\beta}P_2''}{4\Gamma(\alpha)}-\frac{2\left(\frac{\phi}{\theta}\right)^{\beta}P_1''}{\Gamma(\alpha)}-\frac{2\left(\frac{\phi}{\theta}\right)^{2\beta}\rho P_1P_1'}{\{\Gamma(\alpha)\}^2} \\ -\frac{\left(\frac{\phi}{\theta}\right)^{2\beta}\left\{P_1'\right\}^2}{\{\Gamma(\alpha)\}^2}-\frac{2\left(\frac{\phi}{\theta}\right)^{\beta}\left\{\rho+\frac{\gamma+\Psi(\alpha)}{\tau}\right\}P_1'}{\Gamma(\alpha)} \\ -\frac{2\left(\frac{\phi}{\theta}\right)^{\beta}\rho\left\{\frac{\gamma+\Psi(\alpha)}{\tau}\right\}P_1}{\Gamma(\alpha)} \end{array}\right\}.$$

Next, we use (3.23) to write

$$Var_b\left(\frac{\partial l_w}{\partial\theta}\right)=\frac{n\beta^2\theta^{-2}\left(\frac{\phi}{\theta}\right)^{2\beta}}{\Gamma(\alpha)}\left\{P_2-\frac{P_1^2}{\Gamma(\alpha)}\right\};$$

finally, on using (3.24), we write $Cov_b\left(\frac{\partial l_w}{\partial\beta},\frac{\partial l_w}{\partial\theta}\right)$ as

$$\frac{n\beta\theta^{-1}\left(\frac{\phi}{\theta}\right)^{\beta}}{\Gamma(\alpha)}\left\{\begin{array}{c} P_1'-\frac{\left(\frac{\phi}{\theta}\right)^{\beta}P_2'}{2}+\left\{\frac{\gamma+\Psi(\alpha)}{\tau}\right\}P_1 \\ +\frac{\left(\frac{\phi}{\theta}\right)^{\beta}P_1P_1'}{\Gamma(\alpha)}-\left(\frac{\phi}{\theta}\right)^{\beta}\rho P_2 \\ +\frac{\left(\frac{\phi}{\theta}\right)^{\beta}\rho P_1^2}{\Gamma(\alpha)} \end{array}\right\}.$$

We can now compute theoretical standard errors of the MLEs and $\widehat{B}_{w,10}$, for data generated from a Burr model with our usual set of parameter values. When $\tau=3$, $\alpha=4$ and $\phi=100$, we have $\beta_0=2.5528$, $\theta_0=67.2620$, and these correspond (to 4 decimal places) to $P_1=2.1801, P_2=1.8029, P_1'=-0.4783, P_2'=0.2387$, $P_1''=0.3676$, $P_2''=0.8248$ and $\rho=0.3966$. As a result, A simplifies to

$$n\begin{bmatrix} 0.3053 & \\ -0.0067 & 0.0014 \end{bmatrix},$$

and V becomes

$$n\begin{bmatrix} 0.5408 & \\ -0.0184 & 0.0018 \end{bmatrix}.$$

Hence,

$$Var\begin{pmatrix}\widehat{\beta}\\ \widehat{\theta}\end{pmatrix}=A^{-1}VA^{-1}$$

n	50	100	300	500	1000
St.err. $\left(\widehat{\beta}\right)$	0.3303	0.2335	0.1348	0.1044	0.0738
St.err. $\left(\widehat{\theta}\right)$	3.9313	2.7798	1.6049	1.2432	0.8791
St.err. $\left(\widehat{B}_{w,10}\right)$	3.5838	2.5341	1.4631	1.1333	0.8014

Table 3.5: Theoretical standard errors for the MLEs of G_w for varying n. Data is simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

becomes

$$\begin{aligned} & n^{-1}\begin{bmatrix} 0.3053 & \\ -0.0067 & 0.0014 \end{bmatrix}^{-1}\begin{bmatrix} 0.5408 & \\ -0.0184 & 0.0018 \end{bmatrix}\begin{bmatrix} 0.3053 & \\ -0.0067 & 0.0014 \end{bmatrix}^{-1} \\ = \; & n^{-1}\begin{bmatrix} 5.4536 & \\ 0.6263 & 772.7363 \end{bmatrix}. \end{aligned}$$

We use this to construct Table 3.5, which lists the standard errors of the MLEs and $\widehat{B}_{w,10}$ for $n = 50$, 100, 300, 500 and 1000. We compare these theoretical values to their simulated counterparts shown in Table 2.6. The sample standard errors are very close to the theoretical values, even for small sample sizes. We compare sample values of $\widehat{B}_{w,10}$ with the theoretical estimate given by

$$B_{w,10} = \theta_0 \left(-\ln 0.9\right)^{\frac{1}{\beta_0}} = 67.2620 \left(\ln 0.9\right)^{\frac{1}{2.5528}} = 27.8565,$$

and observe good agreement between observed and theoretical results, especially for large sample sizes. Since the true distribution is Burr, we must also compare these results with the true value given by

$$B_{b,10} = \phi \left\{0.9^{\frac{-1}{\alpha}} - 1\right\}^{\frac{1}{\tau}} = 29.8848.$$

In all cases, the average of the sample values of $\widehat{B}_{w,10}$ under-estimate $B_{b,10}$.

Relationships between β_0, θ_0 and the parameters in G_b

In this section, we consider how β_0, θ_0 vary with parameters in G_b, and determine the extent to which there is a relationship between parameters of the two distributions. Given parameter values from the Burr distribution, we wish to find β_0 and θ_0 as easily as possible. We begin by varying each parameter from the Burr.

When ϕ varies There is no relationship between ϕ and β_0, as is clear from the profile entropy score function (3.35), which is independent of ϕ. Thus, its maximising value β_0 does not depend on the scale parameter of the Burr distribution. When we look at how ϕ affects θ_0, (3.34) shows that this entropy value is linearly related to ϕ.

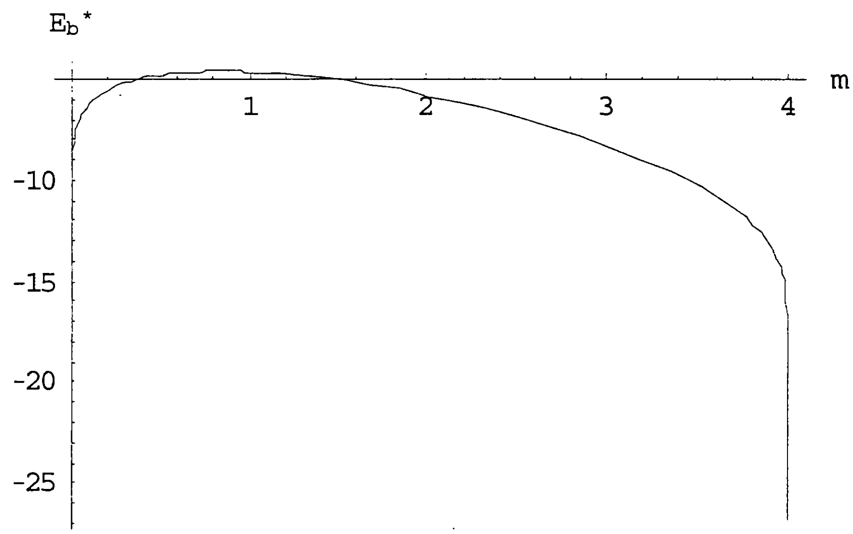


Figure 3.1: E_b^* versus m for $\alpha = 4$.

When τ and α vary We still have to find β_0 for given α and τ before we can calculate θ_0. We seek some relationship between β_0 and the shape parameters in G_b. We first note that (3.35) can be expressed as

$$\frac{dE_b^*}{d\beta} = nm\beta^{-1}\left\{m^{-1} - \gamma - \Psi(\alpha) - \Psi(m+1) + \Psi(\alpha - m)\right\},$$

where $m = \beta\tau^{-1} < \alpha$. For fixed α, we see that changes in τ induce a corresponding change in β_0, with $\beta\tau^{-1}$ constant. Hence, there is a linear relationship between τ and β_0, and we therefore have $\beta_0 = m_0\tau$, where m_0 is the root of

$$m^{-1} - \Psi(m+1) + \Psi(\alpha - m) - \gamma - \Psi(\alpha). \tag{3.37}$$

We next show that (3.37) will have a unique root in the interval $(0, \alpha)$, and then give a simple example. As $m \to 0^+$, $m^{-1} \to \infty$, and so (3.37) is positive here; as $m \to \alpha^-$, $\Psi(\alpha - m) \to -\infty$, since $\Psi(0^+) = -\infty$ and so (3.37) is negative here. Since (3.37) is continuous, this change of sign establishes the existence of a root in the interval. We can show that this root maximises E_b^*, since the second derivative at (3.36) is negative, as $P_1 P_1'' > (P_1')^2$.

Example Using Mathematica, and the usual parameter values for α, τ and ϕ, we see that the root of (3.37) is $m_0 = 0.8509$. This is shown both in Figure 3.1, which illustrates the maximum of the profile entropy function, and in Figure 3.2 which shows that this maximum corresponds to a unique root of (3.37). Using this value of m_0, we have

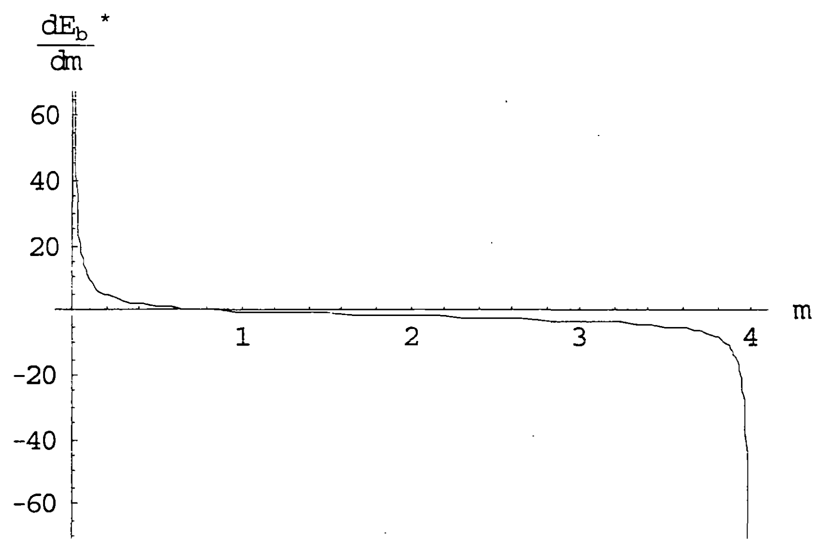


Figure 3.2: $\frac{\partial E_b^*}{\partial m}$ versus m for $\alpha = 4$.

$$\beta_0 = 0.8509\tau = 2.5528,$$

and

$$\theta_0 = \phi \left\{ \frac{\Gamma(1.8509)\Gamma(\alpha - 0.8509)}{\Gamma(\alpha)} \right\}^{\frac{1}{0.8509\tau}} = 67.2620,$$

which are the values obtained from maximising the entropy function. Hence, we see that if we specify parameter values from the Burr distribution, we can obtain, with very little computation, the theoretical equivalents of the MLEs of the Weibull, under mis-specification; all we require is the root of (3.37). We can further improve on this, and show that the root of (3.37) always lies in the interval $(0, 1)$ for $\alpha > 1$. At $m = 1$, (3.37) reduces to

$$-\frac{1}{\alpha - 1} < 0.$$

At $m = 0^+$ this derivative is positive, so the root must occur somewhere in this interval. This is illustrated in Figure 3.3, which shows m_0 as a function of α. We see that as α increases, m_0 tends to 1. This implies that for a large value of α, τ and β_0 are equal, which is consistent with the asymptotic theory on the Burr distribution, since we know that as α and ϕ tend to infinity, the Burr distribution tends to a Weibull with shape parameter equal to τ.

Note that, we now have a relationship between parameters from both distributions, although we do not have explicit expressions for β_0 and θ_0. We still have to find the root of (3.37) in order to relate τ and β_0. One exception is when $\alpha = 1$. In this case, we need

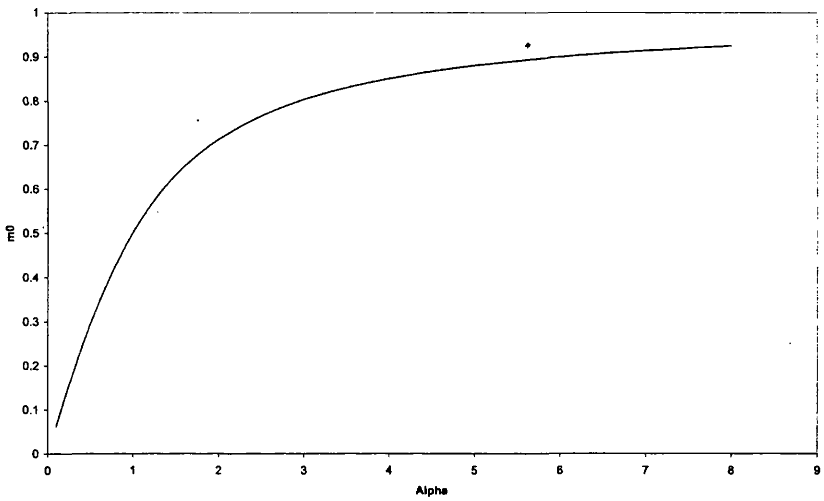


Figure 3.3: Plot of m_0 against α.

the root of

$$m^{-1} - \Psi(m+1) + \Psi(1-m),$$

and, using the fact that we can rewrite $\Psi(m+1)$ using (1.12), and (6.3.7) of Abramowitz and Stegun (1972) to write $\Psi(1-m)$ as

$$\Psi(m) + \pi \cot \pi m, \tag{3.38}$$

we see that we must solve

$$\pi \cot \pi m = 0.$$

for m. As we require $0 < m < 1$, the only root occurs when

$$\pi m = \frac{\pi}{2} \Longrightarrow m_0 = \frac{1}{2}.$$

Thus, when $\alpha = 1$, $\tau = 3$ and $\phi = 100$,

$$\beta_0 = \frac{\tau}{2} = 1.5$$

and

$$\theta_0 = 100 \left\{ \frac{\Gamma\left(\frac{1.5}{3}+1\right)\Gamma\left(1-\frac{1.5}{3}\right)}{\Gamma(1)} \right\}^{\frac{1}{1.5}} = 135.128.$$

α	1	2	3	4	5	6	7	8
m_0	0.5	0.7129	0.8032	0.8509	0.8801	0.8998	0.9140	0.9246

Table 3.6: Values of m_0 for varying α.

Simplifications can be made to (3.37) when α is a positive integer, but numerical methods are still needed to locate the root of the resulting equation. For example, if we put $\alpha = 2$, then using (1.12) and (3.38), we see that (3.37) reduces to

$$\frac{1}{1-m} + \pi \cot \pi \left(m-1\right) - 1,$$

so we must solve

$$\frac{1}{1-m} + \pi \cot \pi \left(m-1\right) = 1$$

for m. In general, for $\alpha \geqslant 2$ the equation becomes

$$\frac{1}{1-m} + \frac{1}{2-m} + ... + \frac{1}{(\alpha-1)-m} + \pi \cot \pi \left\{m - (\alpha - 1)\right\} = \Psi\left(\alpha\right) + \gamma,$$

which involves numerical techniques to find m. Table 3.6 contains some values m_0 for different values of α; as noted, we see m_0 tending to one as α increases.

The effects of changing the parameter values from G_b

So far, we have used the same parameter values in the Burr distribution to simulate data. This has provided us with parameter estimates from the Weibull distribution or theoretical counterparts to the MLEs under the assumption of mis-specification. In order to assess the agreement between this distribution and the Burr, we can examine plots of both cdfs and observe whether any significant differences occur, and measures based on functional distance would be one procedure to summarise such differences. Since we have two theoretical distribution functions, we will, for given sets of τ and α, calculate the entropy values, and then find the largest absolute distance between the two cdfs.

Note that we look at distances between cdfs. Since the general discussion is on reliability distributions, we could consider hazard functions or cumulative hazard functions, since these indicate the probability of failure after a given time has elapsed. We, however, choose to examine cdfs, since the maximum difference can never exceed one. Also, if we allow time to substantially exceed the scaling parameter from the Burr distribution, we observe quite large discrepancies between both cumulative hazard functions, and it becomes very difficult to locate the maximum difference. With cdfs, we know that as we increase time, the maximum absolute distance will have to tend to zero eventually, as both functions tend to one.

As an example, we will carry out this procedure below for the usual set of Burr parameter values and entropy estimates, and then construct a grid of maximum absolute distances for

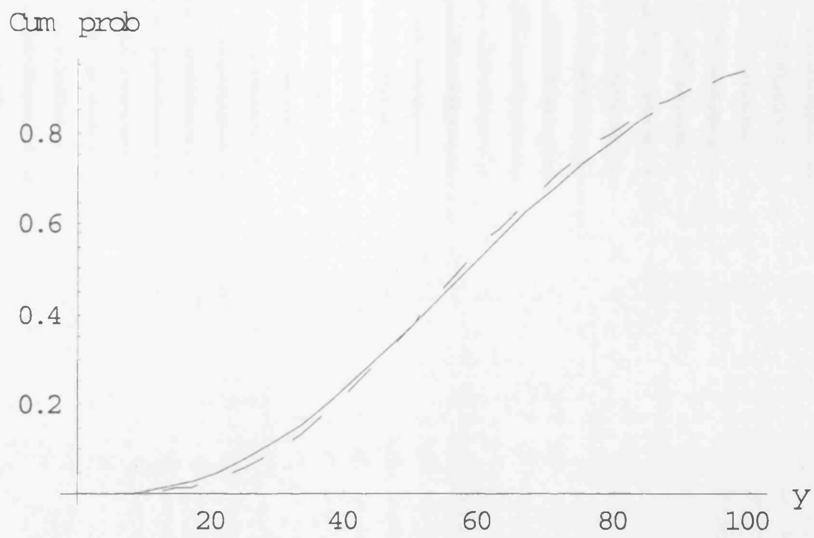


Figure 3.4: Comparison of Burr (—) and Weibull (- - -) cdfs for Burr parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$, and entropy values $\beta_0 = 2.5528$, $\theta_0 = 67.2620$.

appropriate ranges of α and τ. We first construct plots of the two cdfs, in order to illustrate how close a match the Weibull is. Figure 3.4 shows good agreement between the true distribution and fitted Weibull, and it seems surprising that during the simulations, for a sample size of $n = 1000$, not once did we choose to fit the Weibull over the Burr. In order to calculate the maximum distance between the two functions, we consider

$$|G_w(y; \beta_0, \theta_0) - G_b(y; \tau, \alpha, \phi)|.$$

Plotting this function using the above parameter estimates yields Figure 3.5, and, using Mathematica, we locate its maximum at 71.0614, with a distance of 0.0231.

We use this approach for values of α and τ ranging between 0.5 and 4 in steps of 0.5, keeping ϕ fixed at one. Each time, we record the maximum absolute distance between the two cdfs. Note that keeping the scale parameter of the distribution fixed at 1 has no effect on the maximum absolute distances between true and mis-specified distribution functions, since it does not alter the shape of the distribution function in any way at all. The results are shown in Table 3.7. There are a number of points to note when considering this table :

1. We present the maximum absolute distance to 14 decimal places. This reflects the fact that when we allow τ to vary for fixed α, there is very little difference between the distances of the cdfs. It is quite surprising that, for such contrasting distribution functions, we have virtually the same distance between true and entropy distributions (for example, when $\tau = 0.5$ and $\alpha = 0.5$, $\beta_0 = 0.14469$ and $\theta_0 = 236.11452$; this gives the same distance to 5 decimal places when we use the same value for α, but take $\tau = 4$, so $\beta_0 = 1.15749$ and $\theta_0 = 1.97989$). We conclude that varying τ has little effect

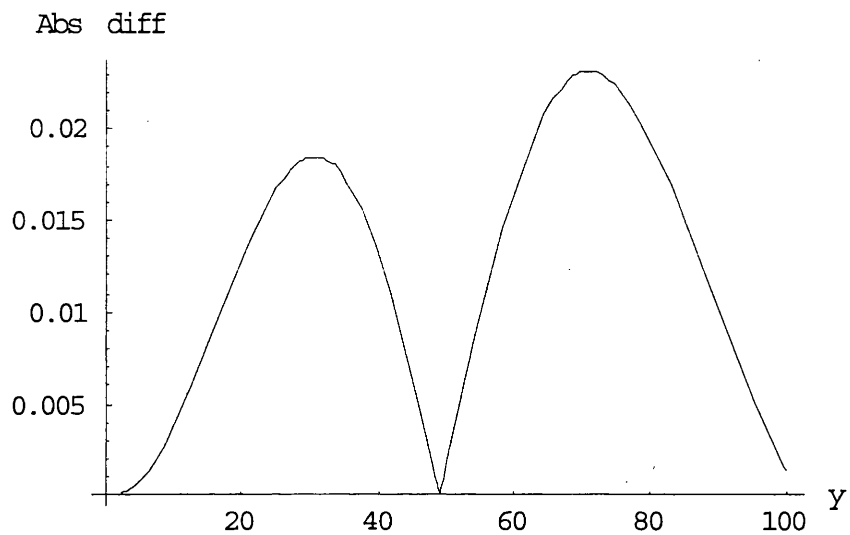


Figure 3.5: Maximum absolute distance between G_b and G_w when $\tau = 3$, $\alpha = 4$, $\phi = 100$, and $\beta_0 = 2.5528$, $\theta_0 = 67.2620$.

		α			
		0.5	1	1.5	2
τ	0.5	0.16304477172366	0.09143926829390	0.05909502750956	0.04546627471506
	1	0.16304477173675	0.09143926832861	0.05909504694886	0.04546627471309
	1.5	0.16304477173539	0.09143926742652	0.05909504170759	0.04546628640850
	2	0.16304477173669	0.09143926832867	0.05909503816060	0.04546628348530
	2.5	0.16304477173282	0.09143926832872	0.05909503602986	0.04546628172821
	3	0.16304477173706	0.09143926832869	0.05909504170620	0.04546628056191
	3.5	0.16304477173672	0.09143926832748	0.05909503968020	0.04546628473824
	4	0.16304477173707	0.09143926832873	0.05909503816014	0.04546628348525
		α			
		2.5	3	3.5	4
τ	0.5	0.03674306450371	0.03075737624454	0.02641969879506	0.02314093454044
	1	0.03674308012891	0.03075736167293	0.02641971265422	0.02314093455289
	1.5	0.03674307495140	0.03075736653089	0.02641970803745	0.02314094345737
	2	0.03674308017500	0.03075736167201	0.02641971265080	0.02314094112818
	2.5	0.03674307704095	0.03075736457341	0.02641970988491	0.02314093989572
	3	0.03674308017509	0.03075736167328	0.02641971254615	0.02314093900651
	3.5	0.03674307793576	0.03075736375422	0.02641971066981	0.02314093836659
	4	0.03674308017420	0.03075736166972	0.02641971265612	0.02314094123198

Table 3.7: Maximum absolute distance between G_b and G_w for varying τ and α.

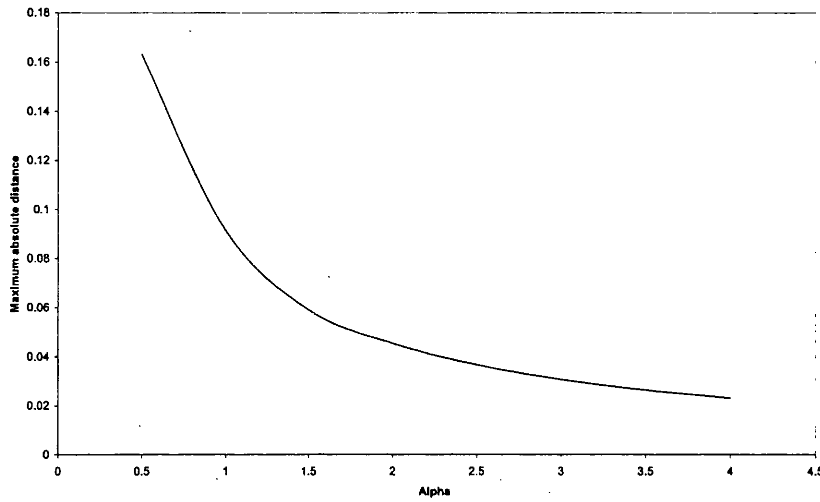


Figure 3.6: Maximum absolute distance between G_w and G_b for $0.5 \leq \alpha \leq 4$; $\tau = 0.5$ and $\phi = 1$.

on the maximum absolute distance between the true and mis-specified distribution functions.

2. The largest value for the maximum absolute difference occurs for the smallest value of α, and we generally see a decrease in the maximum distance between Burr and Weibull, as α is increased. We suspect that if we allowed α to get quite large, but still kept ϕ fixed at 1, then θ_0 and the maximum absolute difference would tend to zero, as the Burr tended to a Weibull distribution. In fact, by considering the structure of (3.34), if we let $\alpha \longrightarrow \infty$, then $\theta_0 \longrightarrow 0$. As an example, we let $\alpha = 100$, and set τ at 0.5. These correspond to $\beta_0 = 0.496962$, $\theta_0 = 0.000101501$; the maximum distance between the true distribution and the fitted Weibull is just 0.000922399.

Figure 3.6 shows a plot of the maximum absolute distance between the two cdfs and varying α, with $\tau = 0.5$ and $\phi = 1$. We generally see a decrease in the maximum absolute distance for increasing α. This might have been expected, since increasing α quite substantially results in the Weibull distribution emerging as the limiting distribution of the Burr. Thus the distances between them will become less, as the Burr distribution becomes more and more like the Weibull. We also examine how changes in parameter values from the Burr distribution affect $B_{w,10}(\beta_0, \theta_0)$ when this quantile is compared to the true value. Table 3.8 summarises the results for α and τ ranging between 0.5 and 4. We observe the largest relative errors for small values of α and τ; this corresponds to the largest maximum absolute distance between the cdfs of the two distributions. Generally, as we increase both shape parameters from the Burr distribution, we observe the Weibull and Burr quantiles becoming more alike.

		τ							
		0.5	1	1.5	2	2.5	3	3.5	4
		0.0550	0.2346	0.3803	0.4843	0.5599	0.6167	0.6608	0.6959
	0.5	4.15×10^{-5}	0.0064	0.0346	0.0803	0.1329	0.1861	0.2366	0.2833
		99.9	97.3	90.9	83.4	76.3	69.8	64.2	59.3
		0.0123	0.1111	0.2311	0.3333	0.4152	0.4807	0.5338	0.5774
	1	0.0008	0.0274	0.0909	0.1655	0.2372	0.3014	0.3578	0.4068
		93.9	75.3	60.7	50.4	42.9	37.3	33.0	29.5
		0.0053	0.0728	0.1743	0.2698	0.3506	0.4175	0.4730	0.5194
	1.5	0.0011	0.0334	0.1037	0.1827	0.2567	0.3220	0.3786	0.4275
		78.9	54.1	40.5	32.3	26.8	22.9	20.0	17.7
		0.0029	0.0541	0.1430	0.2326	0.3114	0.3782	0.4345	0.4823
	2	0.0010	0.0322	0.1012	0.1795	0.2530	0.3182	0.3747	0.4236
		64.6	40.5	29.2	22.8	18.7	15.9	13.8	12.2
α		0.0019	0.0430	0.1228	0.2075	0.2842	0.3505	0.4071	0.4555
	2.5	0.0009	0.0293	0.0951	0.1713	0.2438	0.3084	0.3649	0.4139
		53.5	31.8	22.5	17.4	14.2	12.0	10.4	9.1
		0.0013	0.0357	0.1085	0.1891	0.2638	0.3294	0.3860	0.4348
	3	0.0007	0.0264	0.0887	0.1626	0.2338	0.2979	0.3541	0.4032
		45.3	26.1	18.2	14.0	11.4	9.6	8.3	7.3
		0.0009	0.0306	0.0977	0.1748	0.2478	0.3126	0.3691	0.4181
	3.5	0.0006	0.0238	0.0828	0.1544	0.2243	0.2878	0.3438	0.3929
		39.2	22.0	15.3	11.7	9.5	8.0	6.9	6.0
		0.0007	0.0267	0.0893	0.1634	0.2347	0.2988	0.3551	0.4042
	4	0.0005	0.0216	0.0776	0.1470	0.2157	0.2786	0.3344	0.3834
		34.4	19.0	13.1	10.0	8.1	6.8	5.8	5.1

Table 3.8: Theoretical quantiles for G_b (top) and G_w (middle), with their corresponding relative percentage errors (bottom).

3.2.2 Entropy for fitting G_w to data from G_g

We consider (3.13) when the underlying distribution of the data is Gamma. In order to obtain expectations such as $E_g\left[Y^{m\beta}\ln Y\right]$, we apply similar techniques to those used with G_b. The easiest approach is to use

$$E_g\left[Y^{m\beta}\ln Y\right] = \frac{1}{m}\frac{d}{d\beta}E_g\left[Y^{m\beta}\right] = \frac{\alpha^{m\beta}\Gamma(m\beta+\tau)}{\Gamma(\tau)}\left\{\ln\alpha + \Psi(m\beta+\tau)\right\}, \qquad (3.39)$$

where $E_g\left[Y^{m\beta}\right]$ can be derived from (3.12). To obtain higher order expectations of the form $E_g\left[Y^{m\beta}(\ln Y)^w\right]$, we just differentiate (3.39) the appropriate number of times. In particular, we see that

$$\begin{aligned} E_g\left[Y^{m\beta}(\ln Y)^2\right] &= \frac{1}{m}\frac{d}{d\beta}E_g\left[Y^{m\beta}\ln Y\right] \\ &= \frac{\alpha^{m\beta}\Gamma(m\beta+\tau)}{\Gamma(\tau)}\left[\left\{\ln\alpha + \Psi(m\beta+\tau)\right\}^2 + \Psi'(m\beta+\tau)\right]. \end{aligned} \qquad (3.40)$$

Substituting $\beta = 0$ in (3.39) yields

$$E_g[\ln Y] = \ln\alpha + \Psi(\tau), \qquad (3.41)$$

and, similarly, taking $\beta = 0$ in (3.40) gives

$$E_g\left[(\ln Y)^2\right] = \left\{\ln\alpha + \Psi(\tau)\right\}^2 + \Psi'(\tau).$$

Hence, using (3.13), (3.12) and (3.41), we have

$$E_g = E_g[l_w] = n\left[\ln\beta + \beta\ln\left(\frac{\alpha}{\theta}\right) + \beta\Psi(\tau) - \frac{\left(\frac{\alpha}{\theta}\right)^\beta\Gamma(\beta+\tau)}{\Gamma(\tau)} - \ln\alpha - \Psi(\tau)\right].$$

On using (3.15), we obtain

$$\theta = \alpha\left\{\frac{\Gamma(\beta+\tau)}{\Gamma(\tau)}\right\}^{\frac{1}{\beta}}, \qquad (3.42)$$

from which we obtain the profile entropy

$$E_g^* = n\left[\ln\beta - \ln\left\{\frac{\alpha^\beta\Gamma(\beta+\tau)}{\Gamma(\tau)}\right\} + (\beta-1)\left\{\ln\alpha + \Psi(\tau)\right\} - 1\right], \qquad (3.43)$$

with first and second derivatives

$$\frac{dE_g^*}{d\beta} = n\left\{\beta^{-1} + \Psi(\tau) - \Psi(\beta+\tau)\right\}, \qquad (3.44)$$

and

$$\frac{d^2 E_g^*}{d\beta^2} = -n\left\{\beta^{-2} + \Psi'(\beta+\tau)\right\}.$$

As in the general case, the maximum of (3.43) can be located from a plot. Equivalently, we use Newton-Raphson to locate the root of (3.44); for example, with $\tau = 3$, $\alpha = 50$, we obtain

$$\beta_0 = 1.8328,\ \theta_0 = 169.3772.$$

From Table 2.9, we observe excellent agreement between the MLEs of θ and β, and the theoretical estimated counterparts.

Now that we are able to obtain the entropy values, we continue by deriving the distribution of the Weibull MLEs under mis-specification. We begin this in the next section.

The variance structure of the mis-specified MLEs

We derive the distribution of the Weibull MLEs under the assumption that this distribution has been mis-specified and fitted to data with an underlying Gamma model. Using (3.18), we require expectations of second derivatives of the parameters from the Weibull distribution, where these expected values are taken with respect to the Gamma distribution, and the variance covariance structure of the score functions. We begin with examining second derivatives. Using (3.19), (3.20), and (3.21), we write the matrix A as

$$n\begin{bmatrix} \beta^{-2} + \frac{\left(\frac{\alpha}{\theta}\right)^{\beta}\Gamma(\beta+\tau)}{\Gamma(\tau)}\left[\begin{array}{c}\left\{\ln\left(\frac{\alpha}{\theta}\right) + \Psi(\beta+\tau)\right\}^2 \\ +\Psi'(\beta+\tau)\end{array}\right] & \\ \theta^{-1} - \frac{\theta^{-1}\left(\frac{\alpha}{\theta}\right)^{\beta}\Gamma(\beta+\tau)}{\Gamma(\tau)}\left[1 + \beta\left\{\begin{array}{c}\ln\left(\frac{\alpha}{\theta}\right) + \\ \Psi(\beta+\tau)\end{array}\right\}\right] & \beta\theta^{-2}\left\{(\beta+1)\left(\frac{\alpha}{\theta}\right)^{\beta}\frac{\Gamma(\beta+\tau)}{\Gamma(\tau)} - 1\right\} \end{bmatrix},$$

where the required expectations with respect to the Gamma distribution are obtained using (3.12), (3.39) and (3.40), with appropriate substitutions for m. We now consider the elements that make up the variance covariance matrix of the score functions from the Weibull distribution. We first list functions that make up the elements of this matrix below.

$$\begin{aligned}
Var_g(S_e) &= n\Psi'(\tau), \\
Var_g\{S_0(\beta)\} &= \tfrac{n\alpha^{2\beta}}{\Gamma(\tau)}\left\{\Gamma(2\beta+\tau) - \tfrac{\Gamma(\beta+\tau)^2}{\Gamma(\tau)}\right\}, \\
Var_g\{S_1(\beta)\} &= \tfrac{n\alpha^{2\beta}}{\Gamma(\tau)}\left[\begin{array}{c}\Gamma(2\beta+\tau)\left[\{\ln\alpha+\Psi(2\beta+\tau)\}^2+\Psi'(2\beta+\tau)\right] \\ -\tfrac{\Gamma(\beta+\tau)^2}{\Gamma(\tau)}\{\ln\alpha+\Psi(\beta+\tau)\}^2\end{array}\right], \\
Cov_g\{S_e, S_0(\beta)\} &= \tfrac{n\alpha^{\beta}\Gamma(\beta+\tau)}{\Gamma(\tau)}\{\Psi(\beta+\tau)-\Psi(\tau)\}, \\
Cov_g\{S_e, S_1(\beta)\} &= \tfrac{n\alpha^{\beta}\Gamma(\beta+\tau)}{\Gamma(\tau)}\left[\begin{array}{c}\{\Psi(\beta+\tau)-\Psi(\tau)\}\{\ln\alpha+\Psi(\beta+\tau)\} \\ +\Psi'(\beta+\tau)\end{array}\right], \\
Cov_g\{S_0(\beta), S_1(\beta)\} &= \tfrac{n\alpha^{2\beta}}{\Gamma(\tau)}\left[\begin{array}{c}\Gamma(2\beta+\tau)\{\ln\alpha+\Psi(2\beta+\tau)\} \\ -\tfrac{\Gamma(\beta+\tau)^2}{\Gamma(\tau)}\{\ln\alpha+\Psi(\beta+\tau)\}\end{array}\right].
\end{aligned}$$

We use (3.22) and the above expectations to write

$$Var_g\left(\frac{\partial l_w}{\partial\beta}\right) = n\left[\begin{array}{c}\Psi'(\tau) + \frac{\left(\frac{\alpha}{\theta}\right)^{2\beta}\Gamma(2\beta+\tau)\left\{\ln\left(\frac{\alpha}{\theta}\right)+\Psi(2\beta+\tau)\right\}^2}{\Gamma(\tau)} \\ -\frac{\left(\frac{\alpha}{\theta}\right)^{2\beta}\Gamma(\beta+\tau)^2\left\{\ln\left(\frac{\alpha}{\theta}\right)+\Psi(\beta+\tau)\right\}^2}{\Gamma(\tau)^2} \\ -\frac{2\left(\frac{\alpha}{\theta}\right)^{\beta}\Gamma(\beta+\tau)\{\Psi(\beta+\tau)-\Psi(\tau)\}\left\{\ln\left(\frac{\alpha}{\theta}\right)+\Psi(\beta+\tau)\right\}}{\Gamma(\tau)} \\ +\frac{\left(\frac{\alpha}{\theta}\right)^{\beta}}{\Gamma(\tau)}\left\{\begin{array}{c}\left(\frac{\alpha}{\theta}\right)^{\beta}\Gamma(2\beta+\tau)\Psi'(2\beta+\tau) \\ -2\Gamma(\beta+\tau)\Psi'(\beta+\tau)\end{array}\right\}\end{array}\right].$$

Now, using (3.23), we see that

$$Var_g\left(\frac{\partial l_w}{\partial\theta}\right) = \frac{n\beta^2\theta^{-2}\left(\frac{\alpha}{\theta}\right)^{2\beta}}{\Gamma(\tau)}\left\{\Gamma(2\beta+\tau) - \frac{\Gamma(\beta+\tau)^2}{\Gamma(\tau)}\right\},$$

and finally, by using (3.24), we have

$$Cov_g\left(\frac{\partial l_w}{\partial\beta}, \frac{\partial l_w}{\partial\theta}\right) = \frac{n\beta\theta^{-1}\left(\frac{\alpha}{\theta}\right)^{\beta}}{\Gamma(\tau)}\left[\begin{array}{c}\Gamma(\beta+\tau)\{\Psi(\beta+\tau)-\Psi(\tau)\} - \\ \left(\frac{\alpha}{\theta}\right)^{\beta}\Gamma(2\beta+\tau)\left\{\ln\left(\frac{\alpha}{\theta}\right)+\Psi(2\beta+\tau)\right\} \\ +\frac{\left(\frac{\alpha}{\theta}\right)^{\beta}\Gamma(\beta+\tau)^2\left\{\ln\left(\frac{\alpha}{\theta}\right)+\Psi(\beta+\tau)\right\}}{\Gamma(\tau)}\end{array}\right].$$

We compute theoretical standard-errors for the MLEs of the Weibull distribution for varying sample size, and when $\tau = 3$ and $\alpha = 50$; this corresponds to entropy values of $\beta_0 = 1.8328$, $\theta_0 = 169.3772$. We note that

$$A = n\begin{bmatrix} 0.5892 & \\ -0.0027 & 0.0001 \end{bmatrix},$$

n	50	100	300	500	1000
St.err.$\left(\widehat{\beta}\right)$	0.2021	0.1429	0.0825	0.0639	0.0452
St.err.$\left(\widehat{\theta}\right)$	13.8474	9.7916	5.6532	4.3789	3.0964
St.err.$\left(\widehat{B}_{w,10}\right)$	7.9918	5.6510	3.2626	2.5272	1.7870

Table 3.9: Theoretical standard errors for the MLEs of the Weibull distribution for varying n. Data is simulated from a Gamma model with $\tau = 3$, $\alpha = 50$.

$$V = n \begin{bmatrix} 0.7595 & \\ -0.0058 & 0.0001 \end{bmatrix},$$

which gives

$$Var\begin{pmatrix} \widehat{\beta} \\ \widehat{\theta} \end{pmatrix} = A^{-1}VA^{-1} = n^{-1}\begin{bmatrix} 2.0424 & \\ 5.8722 & 9587.5795 \end{bmatrix}.$$

The results are summarised in Table 3.9, and we compare these to the simulated counterparts shown in Table 2.9. The results show excellent agreement between observed and expected values across all sample sizes. We also include details on $\widehat{B}_{w,10}$, with

$$B_{w,10}\left(\beta_0, \theta_0\right) = \theta_0\left(-\ln 0.9\right)^{\frac{1}{\beta_0}} = 169.3772(-\ln 0.9)^{\frac{1}{1.8328}} = 49.7779.$$

The sample means of this quantile are relatively close to this value, even for small sample sizes. We also compare this theoretical estimate with the true value of B_{10} given by

$$B_{g,10} = 55.1033,$$

and note that the Weibull estimate under-estimates this quantile for this particular set of Gamma parameter values.

Relationships between β_0, θ_0 and the parameters in G_g

In this section, we consider how β_0, θ_0 vary with parameters in G_g, and determine the extent to which there is a relationship between parameters of the two distributions. We first note that, just like the Burr distribution, β_0 is independent of the scale parameter from the Gamma distribution. This becomes evident from examining the profile entropy score function, given by (3.44), and noting that the function is independent of α. We use (3.42) to conclude that θ_0 is linearly related to α. Unlike G_b, however, we cannot really say much more about the relationship between β_0 and τ. Figure 3.7 illustrates this relationship, but, from (3.44), it is clear that we can only locate the root of this function numerically.

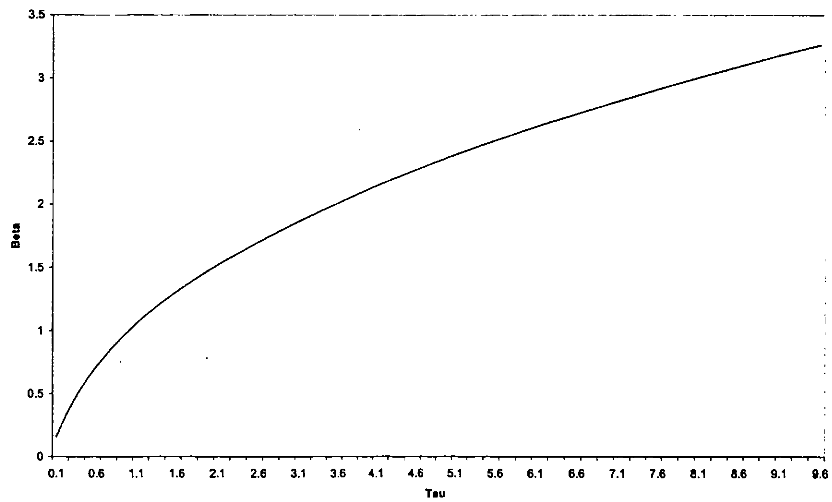


Figure 3.7: The relationship between τ in G_g and β_0.

The effects of changing the parameter values from G_g

In this section we assess the goodness of fit between a mis-specified Weibull distribution (where the parameter values are obtained by maximising the entropy function) and the corresponding Gamma distribution. We will begin by specifying particular parameter ranges for the Gamma distribution, and will first set α equal to 1, since it represents a scale parameter and so does not alter the shape of the distribution function. We have seen, with G_b, that particularly small values for the shape parameters resulted in a relatively poor fit between Weibull and Burr, and this was where the largest values of the maximum absolute distance between both cdfs occurred. We will examine the extent to which a similar pattern holds in this case, for an appropriate range for τ.

As an example, consider $\tau = 3$, $\alpha = 50$. Figure 3.8 shows the Gamma and Weibull cdfs, when we use entropy values for the parameters in the Weibull. We see, for this particular case, that there is very little difference between the Weibull and Gamma distributions. To get a summary measure of how different they are, we plot the absolute distance between them; that is we calculate

$$\left|G_g\left(y;\tau,\alpha\right) - G_w\left(y;\beta_0,\theta_0\right)\right|.$$

Figure 3.9 shows this distance and, we see that the maximum absolute difference between the two cdfs is 0.0263.

The above process is repeated for a range of values for τ; Table 3.10, and Figure 3.10 summarise these results. As with G_b, we observe the maximum absolute difference between the two cdfs for small values of τ. At $\tau = 1$, there is no difference between the cdfs of the

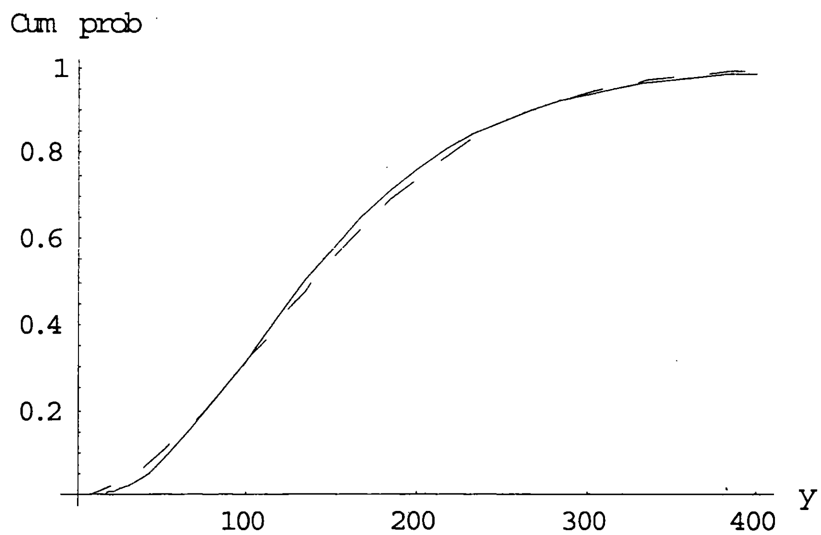


Figure 3.8: Comparison of Gamma (—) and Weibull (- - -) cdfs for Gamma parameters $\tau = 3$, $\alpha = 50$ and entropy values $\beta_0 = 1.8328$, $\theta_0 = 169.3772$.

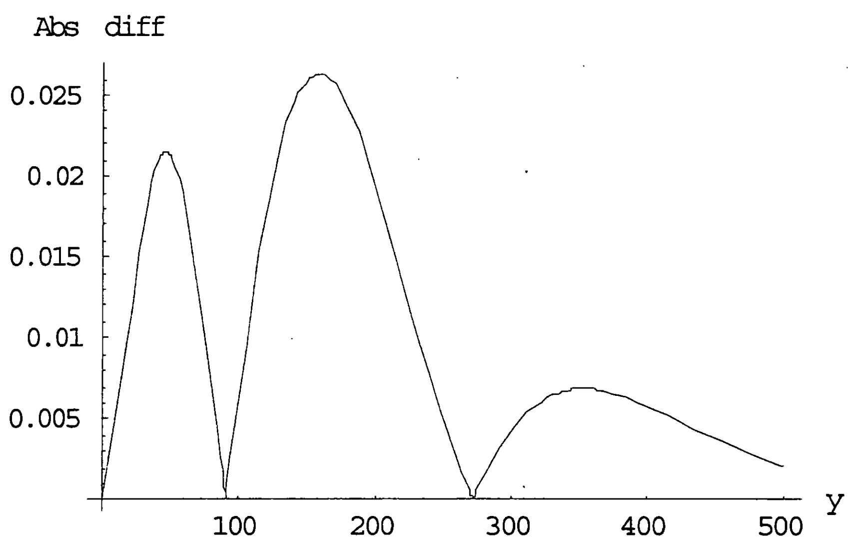


Figure 3.9: Maximum absolute distance between G_g and G_w when $\tau = 3$, $\alpha = 50$ and $\beta_0 = 1.8328$, $\theta_0 = 169.3772$.

τ	Maximum distance	τ	Maximum distance	τ	Maximum distance
0.1	0.0737794	2.5	0.0229684	6.5	0.0373599
0.2	0.0584186	3	0.0263246	7	0.0381965
0.3	0.044574	3.5	0.0289099	7.5	0.0389465
0.4	0.033503	4	0.0309793	8	0.0396239
0.5	0.0247552	4.5	0.0326836	20	0.0471572
1	0	5	0.0341189	100	0.0542719
1.5	0.0115462	5.5	0.0353491	500	0.0574571
2	0.0183709	6	0.0364187	1000	0.0582134

Table 3.10: Maximum absolute distance between G_g and G_w for varying τ; $\alpha = 1$.

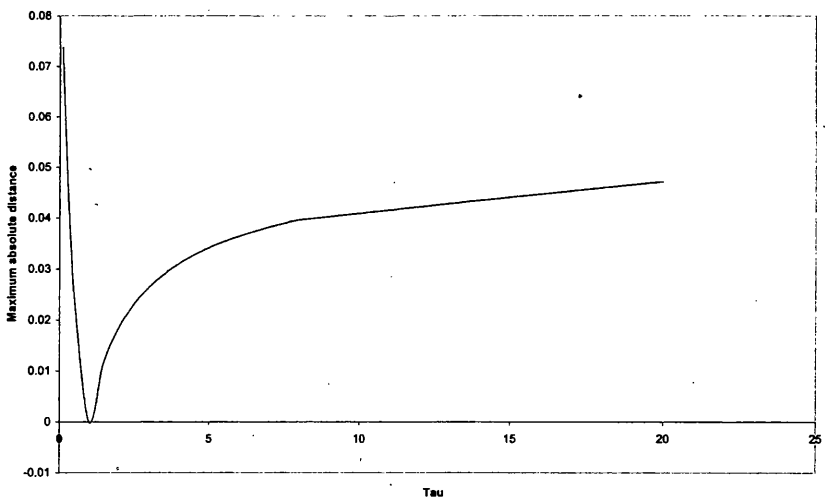


Figure 3.10: Maximum absolute distance between G_g and G_w for $0 < \tau \leq 20$; $\alpha = 1$.

	τ							
	0.5	1	1.5	2	2.5	3	3.5	4
$B_{g,10}$	0.0079	0.1054	0.2922	0.5318	0.8052	1.1021	1.4166	1.7448
$B_{w,10}(\beta_0, \theta_0)$	0.0104	0.1054	0.2737	0.4863	0.7285	0.9923	1.2728	1.5668
% Rel error	31.2	0	6.3	8.6	9.5	10.0	10.1	10.2

Table 3.11: Theoretical quantiles for G_g and G_w, with their corresponding relative percentage errors.

Gamma and Weibull distributions; here, the Gamma distribution reduces to the Negative Exponential distribution, which is also a special case of G_w. Thus, the best Weibull fit to the Negative Exponential is the Negative Exponential itself, and leads to unit scale and shape parameter estimates in this case. As we allow τ to become very large, we see the maximum absolute distance begin to level off at approximately 0.06. In practice, however, we do not expect to observe such extreme values of τ.

We also summarise the effects of changing parameter values from the Gamma distribution on B_{10} from both models, and compare $B_{g,10}$ with $B_{w,10}(\beta_0, \theta_0)$. The results for varying τ are shown in Table 3.11, along with the relative percentage error. We observe similar results to those obtained when we examined the maximum absolute distance between theoretical cdfs. The largest relative error occurs for smaller values of τ, and levels off at 10% as τ increases. We observe no difference between the quantiles when $\tau = 1$.

3.2.3 Entropy for fitting G_w to data from $G_{\ln}$

We consider (3.13) when the underlying distribution of the data is Lognormal. We obtain an expression for $E_{\ln}\left[Y^{m\beta}\right]$ by making use of the relationship between Normal and Lognormal distributions, and the moment generating function from the Normal distribution. From Mann, Schafer and Singpurwalla (1974), we know that when X is Normally distributed with mean μ and variance σ^2, the moment generating function is given by

$$E\left[\exp(tX)\right] = \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}.$$

Now, using

$$Y = \exp(X) \Longleftrightarrow \ln(Y) = X,$$

we have

$$E_{\ln}\left[Y^{m\beta}\right] = E\left[\exp(m\beta X)\right] = \exp(\delta_m), \tag{3.45}$$

where

$$\delta_m = \mu m\beta + \frac{\sigma^2 (m\beta)^2}{2}.$$

We differentiate (3.45) with respect to β to obtain

$$E_{\ln}\left[Y^{m\beta}\ln Y\right] = \left(\mu + \sigma^2 m\beta\right)\exp\left(\delta_m\right), \tag{3.46}$$

and also differentiate this expectation to get

$$E_{\ln}\left[Y^{m\beta}(\ln Y)^2\right] = \left\{\left(\mu + \sigma^2 m\beta\right)^2 + \sigma^2\right\}\exp\left(\delta_m\right). \tag{3.47}$$

By setting $\beta = 0$ in (3.46) we have

$$E_{\ln}[\ln Y] = E[X] = \mu,$$

and an equivalent substitution in (3.47) yields

$$E_{\ln}\left[(\ln Y)^2\right] = E\left[X^2\right] = \mu^2 + \sigma^2.$$

Thus, we see that the entropy function can now be expressed as

$$E_{\ln} = E_{\ln}[l_w] = n\ln\beta - n\beta\ln\theta + n(\beta-1)\mu - n\theta^{-\beta}\exp\left(\beta\mu + \frac{\sigma^2\beta^2}{2}\right).$$

Using (3.15), we have

$$\theta = \exp\left(\mu + \frac{\sigma^2\beta}{2}\right), \tag{3.48}$$

and using (3.16) we obtain

$$E_{\ln}^* = n\ln\beta - n\beta\left(\mu + \frac{\sigma^2\beta}{2}\right) + n\mu(\beta-1) - n,$$

with first derivative

$$\frac{dE_{\ln}^*}{d\beta} = n\left(\beta^{-1} - \sigma^2\beta\right). \tag{3.49}$$

The root of (3.49) is

$$\beta_0 = \sigma^{-1},$$

and, inserting this into (3.48), we obtain

$$\theta_0 = \exp\left(\mu + \frac{\sigma}{2}\right).$$

We compare our entropy values to the simulated results summarised in Table 2.11. Here, we set $\mu = 2$, $\sigma = 3$; this yields

$$\theta_0 = \exp\left(2 + \frac{3}{2}\right) = 33.1155,$$

and

$$\beta_0 = \frac{1}{3}.$$

The table shows how the sample means for $\widehat{\beta}$ and $\widehat{\theta}$ seem to be tending to their entropy values for larger n. We use these entropy values in the next section, where we derive the distribution of the MLEs from the mis-specified model.

The variance structure of the mis-specified MLEs We evaluate each element of the matrix A in (3.18) below, and first consider $-E_{\ln}\left[\frac{\partial^2 l_w}{\partial\theta^2}\right]$; using (3.20) and (3.45), we see that this simplifies to

$$-n\beta\theta^{-2}\left\{1 - (\beta+1)\,\theta^{-\beta}\exp(\delta_1)\right\}.$$

Next, we examine $E_{\ln}\left[\frac{\partial^2 l_w}{\partial\beta^2}\right]$; using (3.19), we have

$$-E_{\ln}\left[\frac{\partial^2 l_w}{\partial\beta^2}\right] = n\left[\beta^{-2} + \theta^{-\beta}\exp(\delta_1)\left\{\left(\mu + \beta\sigma^2 - \ln\theta\right)^2 + \sigma^2\right\}\right],$$

and finally, by making use of (3.21), we write

$$-E_{\ln}\left[\frac{\partial^2 l_w}{\partial\beta\partial\theta}\right] = n\theta^{-1}\left[1 - \theta^{-\beta}\exp(\delta_1)\left\{1 + \beta\left(\mu + \beta\sigma^2 - \ln\theta\right)\right\}\right].$$

We now consider elements which make up the variance covariance matrix of score functions from the Weibull distribution. As with previous distributions, we first list variances and covariances of functions which make up these elements.

$$\begin{aligned}
Var_{\ln}(S_e) &= n\sigma^2, \\
Var_{\ln}\{S_0(\beta)\} &= n\{\exp(\delta_2) - \exp(2\delta_1)\}, \\
Var_{\ln}\{S_1(\beta)\} &= n\left[\left\{(\mu + 2\beta\sigma^2)^2 + \sigma^2\right\}\exp(\delta_2) - (\mu + \beta\sigma^2)^2\exp(2\delta_1)\right], \\
Cov_{\ln}\{S_e, S_0(\beta)\} &= n\beta\sigma^2\exp(\delta_1), \\
Cov_{\ln}\{S_e, S_1(\beta)\} &= n\left\{\beta\sigma^2(\mu + \beta\sigma^2) + \sigma^2\right\}\exp(\delta_1), \\
Cov_{\ln}\{S_0(\beta), S_1(\beta)\} &= n\left[(\mu + 2\beta\sigma^2)\exp(\delta_2) - (\mu + \beta\sigma^2)\exp(2\delta_1)\right].
\end{aligned}$$

Using (3.22), we have

$$Var_{\ln}\left(\frac{\partial l_w}{\partial \beta}\right) = n\left\{\begin{array}{c} \sigma^2 + \theta^{-2\beta}\exp(\delta_2)\left[\left\{\mu + 2\beta\sigma^2 - \ln\theta\right\}^2 + \sigma^2\right] \\ -\theta^{-2\beta}\exp(2\delta_1)\left\{\mu + \beta\sigma^2 - \ln\theta\right\}^2 \\ -2\theta^{-\beta}\exp(\delta_1)\left[\beta\sigma^2\left\{\mu + \beta\sigma^2 - \ln\theta\right\} + \sigma^2\right] \end{array}\right\}.$$

Next, we derive $Var_{\ln}\left(\frac{\partial l_w}{\partial \theta}\right)$, which, by using (3.23), takes the form

$$n\beta^2\theta^{-2\beta-2}\{\exp(\delta_2) - \exp(2\delta_1)\}.$$

Finally, using (3.24), we see that

$$Cov_{\ln}\left(\frac{\partial l_w}{\partial \beta}, \frac{\partial l_w}{\partial \theta}\right) = n\beta\theta^{-\beta-1}\left[\begin{array}{c} \beta\sigma^2\exp(\delta_1) - \theta^{-\beta}\exp(\delta_2)\left\{\mu + 2\beta\sigma^2 - \ln\theta\right\} \\ +\theta^{-\beta}\exp(2\delta_1)\left\{\mu + \beta\sigma^2 - \ln\theta\right\} \end{array}\right].$$

We compute theoretical values for the standard errors of the Weibull MLEs, and compare these to simulated counterparts shown in Table 2.11. We also compute the theoretical standard error of $\widehat{B}_{w,10}$, and note that this will take the same form as (3.25), but with the entropy values and variances of MLEs replaced by expressions derived in this section. As in previous cases, we evaluate

$$A = n\left[\begin{array}{cc} 20.25 & \\ -0.0151 & 0.0001 \end{array}\right],$$

and

$$V = n\left[\begin{array}{cc} 59.2597 & \\ -0.0778 & 0.0002 \end{array}\right],$$

n	50	100	300	500	1000
St.err. $\left(\widehat{\beta}\right)$	0.0496	0.0351	0.0203	0.0157	0.0111
St.err. $\left(\widehat{\theta}\right)$	14.9774	10.5906	6.1145	4.7363	3.3491
St.err. $\left(\widehat{B}_{w,10}\right)$	0.0381	0.0270	0.0156	0.0121	0.0085

Table 3.12: Theoretical standard errors for the MLEs of G_w for varying n. Data is simulated from G_{ln} with $\mu = 2$, $\sigma = 3$.

which gives, for this particular set of Lognormal parameters, a variance covariance matrix for $\left(\widehat{\beta}, \widehat{\theta}\right)'$ of the form

$$n^{-1}\begin{bmatrix} 0.1232 & \\ -10.0860 & 11216.1833 \end{bmatrix}.$$

The results are summarised in Table 3.12; we note that

$$B_{w,10}\left(\beta_0, \theta_0\right) = \theta_0\left(-\ln 0.9\right)^{\frac{1}{\beta_0}} = \exp\left(\mu + \frac{\sigma}{2}\right)\{-\ln 0.9\}^{\sigma} = 0.0387,$$

which we can compare with the true quantile from the Lognormal distribution given by

$$B_{\text{ln},10} = 0.1581.$$

We compare Table 3.12 with Table 2.11, and observe that the MLEs for the Weibull distribution seem to be tending towards their corresponding entropy values for large sample sizes. The agreement between observed and expected standard errors also improves for larger n, but is very good even for small sample sizes. When we examine $B_{w,10}$, the sample mean matches up reasonably well to the estimated value from the Weibull distribution. When we compare these to the true value from the Lognormal, we see the time to which 10% of observations fail is very much under-estimated when we fit the wrong distribution function.

The effects of changing the parameter values from G_{ln}

To assess the agreement between the Weibull and Lognormal distributions, we will use similar techniques to those established for the Burr and Gamma distributions. That is, we choose an appropriate range of values for the Lognormal parameters, calculate entropy values for parameters in the Weibull, and then assess, by plots, if the fitted Weibull is close to the Lognormal. As an example, we will carry out this process using the usual set of parameter values. Figure 3.11 shows the agreement between true and mis-specified distributions. By computing the maximum of

$$\left|G_{\text{ln}}\left(y; \mu, \sigma\right) - G_w\left(y; \beta_0, \theta_0\right)\right|,$$

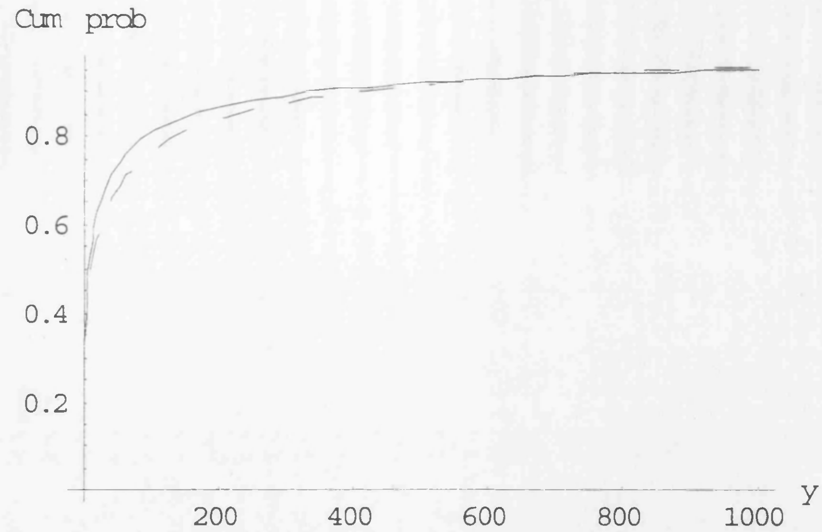


Figure 3.11: Comparison of Lognormal (—) and Weibull (- - -) cdfs for Lognormal parameters $\mu = 2$, $\sigma = 3$ and entropy values $\beta_0 = \frac{1}{3}$, $\theta_0 = 33.1155$.

we see that the largest distance between the two cdfs is 0.0605. This process can be repeated for a range of parameters from the Lognormal distribution, where, we recall that the parameter μ can take any value, whether it is negative or positive, but σ must be larger than zero. However, it is not necessary to construct the table of values showing the maximum distance, since the same maximum absolute distance is observed no matter what parameters we use from the Lognormal distribution. Thus, we can choose any values from the Lognormal distribution, fit a Weibull to this, and be sure that we will always observe an adequate fit, the largest distance between the two distribution functions being just 0.0605. We can further simplify this maximum distance, and write β_0, θ_0 in terms of the parameters from the Lognormal distribution. Thus, we seek

$$\max_{y>0}\left[\left|\exp\left\{-\left(\frac{y}{\exp\left(\mu+\frac{\sigma}{2}\right)}\right)^{\frac{1}{\sigma}}\right\}-1+\Phi\left\{\ln\left(\frac{y}{\exp\left(\mu\right)}\right)^{\frac{1}{\sigma}}\right\}\right|\right],$$

a function that gives the same answer no matter what parameter values we use. We note that this maximum is a function of

$$\left(\frac{y}{\exp\left(\mu\right)}\right)^{\frac{1}{\sigma}};$$

further examination of this is possible, but omitted here.

We also summarise the effects changing parameter values from the Lognormal distribution has on the theoretical quantiles from both models. The results are shown in Table 3.13 for varying μ and σ. We see that, although varying μ changes the values of both quantiles,

		σ				
		1	2	3	4	5
μ	-3	0.0138 0.0086 37.4	0.0038 0.0015 60.8	0.0011 0.0003 75.5	0.0003 4.53×10^{-5} 84.7	8.20×10^{-5} 7.87×10^{-6} 90.4
	-2	0.0376 0.0235 37.4	0.0104 0.0041 60.8	0.0029 0.0007 75.5	0.0008 0.0001 84.7	0.0002 2.14×10^{-5} 90.4
	-1	0.1021 0.0639 37.4	0.0284 0.0111 60.8	0.0079 0.0019 75.5	0.0022 0.0003 84.7	0.0006 5.82×10^{-5} 90.4
	0	0.2776 0.1737 37.4	0.0771 0.0302 60.8	0.0214 0.0052 75.5	0.0059 0.0009 84.7	0.0016 0.0002 90.4
	1	0.7546 0.4722 37.4	0.2095 0.0820 60.8	0.0582 0.0142 75.5	0.0161 0.0025 84.7	0.0045 0.0004 90.4
	2	2.0513 1.2836 37.4	0.5694 0.2230 60.8	0.1581 0.0387 75.5	0.0439 0.0067 84.7	0.0122 0.0012 90.4
	3	5.5759 3.4891 37.4	1.5479 0.6061 60.8	0.4297 0.1053 75.5	0.1193 0.0183 84.7	0.0331 0.0032 90.4

Table 3.13: Theoretical quantiles for G_{ln} (top) and G_w (middle), with their corresponding relative percentage errors (bottom).

it has no effect on the relative error between them. As we increase σ, the error between the two quantiles increases considerably.

3.3 Summary

In this chapter, we obtained the EFI matrix for the MLEs from the Weibull, Burr, Gamma and Lognormal distributions. We then considered mis-specifying the Weibull distribution, and developed general results to obtain theoretical counterparts to the MLEs from this mis-specified distribution using the entropy function. We used these estimates to derive the theoretical variance covariance matrix for the MLEs, and also considered the mean and variance of $\widehat{B}_{w,10}$. Results were then obtained when we assumed the underlying distribution was Burr, Gamma and Lognormal. In all cases, we also examined relationships between parameters from the true distribution, and our entropy values. Finally, we considered sets of parameter values from the true distribution, where the Weibull did not provide an adequate fit.

The following chapter extends these results when we assume the data set has undergone censoring.

Chapter 4

Censoring

In Chapters 2 and 3, we examined the effects of mis-specifying the Weibull distribution to a complete set of data. In practice, life data rarely contains all observations that have failed. Instead, the running time is shortened using a technique such as censoring. This chapter examines the effects of mis-specifying the Weibull distribution when the data set has undergone a type I and type II censoring regime. Unlike Chapter 3, where we varied the underlying distribution of the data, this chapter will focus on only using the Burr distribution as the true underlying model. Thus, censoring will not be considered for the Gamma or Lognormal distributions. The Burr does have asymptotic links with the Weibull, and so is of interest for this reason. We begin with type I censoring, and consider this below.

4.1 Type I censoring

Recall, from Chapter 1, that if we subject data to type I censoring, then all items start in service at the same time, and there is some pre-specified (fixed) time y_c after which the experiment is terminated. The number of failures N is random. We begin by first deriving the theory necessary to fit the Weibull and Burr distributions to a set of data that has been censored using a type I regime, and also examine the variance covariance structure of their MLEs, under the assumption that the models have been correctly specified. Next, we examine the effects of mis-specifying the Weibull distribution, and, as in Chapter 2, use simulations to study MLEs from true and mis-specified distributions, and the preference for fitting the Weibull over the Burr. The final sections examine the theory to explain our simulated values, and discuss the agreement between the true Burr distribution and mis-specified Weibull.

4.1.1 ML estimation for G_w under type I censoring

Here, we assume that the lifetime of items follow G_w, so that the probability of an item failing in the interval $(0, y_c)$ is given by

$$1 - \exp\left\{-\left(\frac{y_c}{\theta}\right)^\beta\right\} = 1 - \exp(-z_c),$$

where

$$z_c = y_c^\beta \theta^{-\beta}.$$

Here, N will follow a Binomial distribution with sample size n, and probability to failure $1 - \exp(-z_c)$, so that

$$E[N] = n\{1 - \exp(-z_c)\}.$$

The observed times to failure follow the truncated Weibull distribution with pdf

$$\frac{\frac{\beta y^{\beta-1}}{\theta^\beta}\exp\left\{-\left(\frac{y}{\theta}\right)^\beta\right\}}{1 - \exp(-z_c)} \tag{4.1}$$

for $0 < y < y_c$. We note that $Z = \left(\frac{Y}{\theta}\right)^\beta$, follows the truncated Negative Exponential distribution with pdf

$$\frac{\exp(-z)}{1 - \exp(-z_c)}, \tag{4.2}$$

for $0 < z < z_c$. Without loss of generality, the likelihood for data under this censoring regime is

$$L_w = \prod_{i=1}^{N} \frac{\beta y_i^{\beta-1}}{\theta^\beta} \exp\left\{-\left(\frac{y_i}{\theta}\right)^\beta\right\} \prod_{i=1}^{n-N} \exp\left\{-\left(\frac{y_c}{\theta}\right)^\beta\right\},$$

where $y_1, y_2, ..., y_N$ are observed times to failure, and $y_{N+1}, y_{N+2}, ..., y_n = y_c$ are censored times in service of items still operational at y_c. The log-likelihood is

$$l_w = N \ln\beta - N\beta \ln\theta + (\beta - 1) S_e - \theta^{-\beta} S_0(\beta), \tag{4.3}$$

where

$$S_e = \sum_{i=1}^{N} \ln y_i,$$

and

$$S_j(\beta) = \sum_{i=1}^{N} y_i^\beta (\ln y_i)^j + (n-N)\, y_c^\beta (\ln y_c)^j ,$$

with the property that

$$S_j(\beta) = \frac{d^j S_0(\beta)}{d\beta^j} = \frac{dS_{j-1}(\beta)}{d\beta}.$$

In what follows, we assume $N > 0$. The score function will contain the elements

$$\frac{\partial l_w}{\partial \theta} = -N\beta\theta^{-1} + \beta\theta^{-\beta-1} S_0(\beta), \tag{4.4}$$

$$\frac{\partial l_w}{\partial \beta} = N\beta^{-1} - N\ln\theta + S_e - \theta^{-\beta}\{S_1(\beta) - \ln\theta S_0(\beta)\}. \tag{4.5}$$

We equate (4.4) to zero, to obtain

$$\theta = \left\{\frac{S_0(\beta)}{N}\right\}^{\frac{1}{\beta}};$$

substituting this into (4.3) gives the profile log-likelihood l_w^*,

$$l_w^* = N\ln\beta + N\ln N - N\ln S_0(\beta) + (\beta-1)S_e - N,$$

with first derivative

$$\frac{dl_w^*}{d\beta} = N\beta^{-1} + S_e - \frac{NS_1(\beta)}{S_0(\beta)},$$

and second derivative

$$\frac{d^2 l_w^*}{d\beta^2} = -N\beta^{-2} - N\left\{\frac{S_2(\beta)\, S_0(\beta) - S_1(\beta)^2}{S_0(\beta)^2}\right\}.$$

We also note the derivatives of the score function with respect to β and θ, which will be used in the computation of the EFI matrix. We have

$$\frac{\partial^2 l_w}{\partial \theta^2} = N\beta\theta^{-2} - \beta(\beta+1)\theta^{-\beta-2} S_0(\beta),$$

$$\frac{\partial^2 l_w}{\partial \beta^2} = -N\beta^{-2} - \theta^{-\beta}\left\{S_2(\beta) - 2\ln\theta S_1(\beta) + (\ln\theta)^2 S_0(\beta)\right\},$$

$$\frac{\partial^2 l_w}{\partial\theta\partial\beta} = -N\theta^{-1} + \theta^{-\beta-1}\left[\beta S_1(\beta) + \{1 - \beta\ln\theta\} S_0(\beta)\right].$$

Thus, we have all the necessary terms to compute the MLEs for the Weibull distribution, which, as before, must be obtained numerically. Convergence requires a suitable starting value for β; see Farnum and Booth (1997) for details on an initial starting value for β when data has undergone any form of censoring (including type I). This method will be used when we compute MLEs from the Weibull distribution in later sections.

We carry out a similar analysis for G_b below.

4.1.2 ML estimation for G_b under type I censoring

We outline the theory necessary to fit G_b, using properties of the two parameter Burr distribution. We include only the main points in the following sections, and refer to Johnson (2003) for further details. We now assume that the lifetime of items follow G_b, and are subject to type I censoring at y_c. The probability that an item fails in $(0, y_c)$ is given by

$$q_{z_c,\alpha} = 1 - \left\{1 + \left(\frac{y_c}{\phi}\right)^{\tau}\right\}^{-\alpha} = 1 - (1 + z_c^{\tau})^{-\alpha},$$

where, now, $z_c = \frac{y_c}{\phi}$. The random variable N will have a Binomial distribution with parameters n and $q_{z_c,\alpha}$, so that

$$E[N] = nq_{z_c,\alpha}.$$

The observed times to failure are from a truncated Burr distribution with pdf

$$\frac{\frac{\alpha\tau y^{\tau-1}}{\phi^{\tau}}\left\{1 + \left(\frac{y}{\phi}\right)^{\tau}\right\}^{-(\alpha+1)}}{q_{z_c,\alpha}}$$

for $0 < y < y_c$. The likelihood is

$$L_b = \prod_{i=1}^{N} \frac{\alpha\tau y_i^{\tau-1}}{\phi^{\tau}} \left\{1 + \left(\frac{y_i}{\phi}\right)^{\tau}\right\}^{-(\alpha+1)} \prod_{i=1}^{n-N} \left\{1 + \left(\frac{y_c}{\phi}\right)^{\tau}\right\}^{-\alpha},$$

where $y_1, y_2, ..., y_N$ are observed times to failure, and $y_{N+1}, y_{N+2}, ..., y_n = y_c$ are censored times in service of items still operational at y_c; the corresponding log-likelihood is

$$l_b = N\ln\alpha + N\ln\tau - N\tau\ln\phi + (\tau - 1)S_e - (\alpha + 1)T_f - \alpha T_c,$$

where

$$T_f = \sum_{i=1}^{N} \ln\left\{1 + \left(\frac{y_i}{\phi}\right)^{\tau}\right\},$$

and

$$T_c = (n - N)\ln\left\{1 + \left(\frac{y_c}{\phi}\right)^{\tau}\right\}.$$

For future use, we also introduce the following notation:

$$\begin{aligned}
T_{1,0f} &= \frac{\partial T_f}{\partial \tau} = \sum_{i=1}^{N} \frac{\left(\frac{y_i}{\phi}\right)^{\tau} \ln\left(\frac{y_i}{\phi}\right)}{1+\left(\frac{y_i}{\phi}\right)^{\tau}}, \\
T_{1,0c} &= \frac{\partial T_c}{\partial \tau} = \frac{(n-N)\left(\frac{y_c}{\phi}\right)^{\tau} \ln\left(\frac{y_c}{\phi}\right)}{1+\left(\frac{y_c}{\phi}\right)^{\tau}}, \\
T_{2,0f} &= \frac{\partial T_{1,0c}}{\partial \tau} = \sum_{i=1}^{N} \frac{\left(\frac{y_i}{\phi}\right)^{\tau} \left\{\ln\left(\frac{y_i}{\phi}\right)\right\}^2}{\left\{1+\left(\frac{y_i}{\phi}\right)^{\tau}\right\}^2}, \\
T_{2,0c} &= \frac{\partial T_{1,0c}}{\partial \tau} = \frac{(n-N)\left(\frac{y_c}{\phi}\right)^{\tau} \left\{\ln\left(\frac{y_c}{\phi}\right)\right\}^2}{\left\{1+\left(\frac{y_c}{\phi}\right)^{\tau}\right\}^2},
\end{aligned}$$

and

$$\begin{aligned}
T_{0,1f} &= \frac{\partial T_f}{\partial \phi} = -\tau\phi^{-1} \sum_{i=1}^{N} \frac{\left(\frac{y_i}{\phi}\right)^{\tau}}{1+\left(\frac{y_i}{\phi}\right)^{\tau}}, \\
T_{0,1c} &= \frac{\partial T_c}{\partial \phi} = \frac{-\tau(n-N)\left(\frac{y_c}{\phi}\right)^{\tau}}{1+\left(\frac{y_c}{\phi}\right)^{\tau}}, \\
T_{0,2f} &= \frac{\partial T_{0,1f}}{\partial \phi} = \tau\phi^{-2} \sum_{i=1}^{N} \frac{\left(\frac{y_i}{\phi}\right)^{\tau}}{1+\left(\frac{y_i}{\phi}\right)^{\tau}} + \tau^2\phi^{-2} \sum_{i=1}^{N} \frac{\left(\frac{y_i}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{y_i}{\phi}\right)^{\tau}\right\}^2}, \\
T_{0,2c} &= \frac{\tau(n-N)\left(\frac{y_c}{\phi}\right)^{\tau}}{\phi^2\left\{1+\left(\frac{y_c}{\phi}\right)^{\tau}\right\}} + \frac{\tau^2(n-N)\left(\frac{y_c}{\phi}\right)^{\tau}}{\phi^2\left\{1+\left(\frac{y_c}{\phi}\right)^{\tau}\right\}^2},
\end{aligned}$$

$$\begin{aligned}
T_{1,1f} &= \frac{\partial T_{0,1f}}{\partial \tau} = -\phi^{-1} \sum_{i=1}^{N} \frac{\left(\frac{y_i}{\phi}\right)^{\tau}}{1+\left(\frac{y_i}{\phi}\right)^{\tau}} - \tau\phi^{-1} \sum_{i=1}^{N} \frac{\left(\frac{y_i}{\phi}\right)^{\tau} \ln\left(\frac{y_i}{\phi}\right)}{\left\{1+\left(\frac{y_i}{\phi}\right)^{\tau}\right\}^2}, \\
T_{1,1c} &= \frac{\partial T_{0,1c}}{\partial \tau} = -\frac{(n-N)\left(\frac{y_c}{\phi}\right)^{\tau}}{\phi\left\{1+\left(\frac{y_c}{\phi}\right)^{\tau}\right\}} - \frac{\tau(n-N)\left(\frac{y_c}{\phi}\right)^{\tau} \ln\left(\frac{y_c}{\phi}\right)}{\phi\left\{1+\left(\frac{y_c}{\phi}\right)^{\tau}\right\}^2},
\end{aligned}$$

where sub-scripts indicate differentiation with respect to τ or ϕ, and f and c denote failed and censored items, respectively. Using this, the score function for l_b is

$$
\begin{aligned}
\frac{\partial l_b}{\partial \alpha} &= N\alpha^{-1} - T_f - T_c, && (4.6)\\
\frac{\partial l_b}{\partial \tau} &= N\tau^{-1} - N\ln\phi + S_e - (\alpha+1)T_{1,0f} - \alpha T_{1,0c}, && (4.7)\\
\frac{\partial l_b}{\partial \phi} &= -N\tau\phi^{-1} - (\alpha+1)T_{0,1f} - \alpha T_{0,1c}. && (4.8)
\end{aligned}
$$

We can equate (4.6) to zero and insert the expression into (4.7) and (4.8) to obtain a profile score function. However, this function is not required, and so will not be considered here; we use the same approach to fit the Burr distribution as with complete data, outlined in Section 2.3.2. To compute the EFI matrix, we require the expectations of the following second derivatives

$$
\begin{aligned}
\frac{\partial^2 l_b}{\partial \alpha^2} &= -N\alpha^{-2},\\
\frac{\partial^2 l_b}{\partial \tau^2} &= -N\tau^{-2} - (\alpha+1)T_{2,0f} - \alpha T_{2,0c},\\
\frac{\partial^2 l_b}{\partial \phi^2} &= N\tau\phi^{-2} - (\alpha+1)T_{0,2f} - \alpha T_{0,2c},\\
\frac{\partial^2 l_b}{\partial \alpha\partial \tau} &= -T_{1,0f} - T_{1,0c},\\
\frac{\partial^2 l_b}{\partial \alpha\partial \phi} &= -T_{0,1f} - T_{0,1c},\\
\frac{\partial^2 l_b}{\partial \tau\partial \phi} &= -N\phi^{-1} - (\alpha+1)T_{1,1f} - \alpha T_{1,1c},
\end{aligned}
$$

which will be considered in later sections.

We continue by looking at the effects of fitting G_w to type I censored data from G_b, via simulations.

4.1.3 Fitting G_w to G_b data

Our analysis of mis-specified distributions fitted to data subjected to censoring follows a similar structure to the complete scenario. Since we are interested in fitting the Weibull distribution to Burr data, we begin by simulating sets of data from a Burr distribution with appropriate parameter values and stopping times. Once this is done, we fit the Weibull distribution and obtain MLEs. We compute the sign of Δ to deduce if the Burr distribution can be fitted, and, if $\Delta > 0$, we fit this distribution to the data. The form of Δ will be slightly different from the complete case, since we have to take censoring into consideration. Watkins (1999) computes Δ for data that has undergone type I censoring, as

n	100	300	500	1000
y_c	80	80	80	80
$\widehat{\beta}$	2.8046	2.7786	2.7715	2.7686
(st.err.)	(0.2589)	(0.1458)	(0.1133)	(0.0794)
$\widehat{\theta}$	66.1870	66.2424	66.2488	66.2561
(st.err.)	(2.7074)	(1.5609)	(1.2023)	(0.8429)
$\widehat{B}_{w,10}$	29.5507	29.4334	29.3902	29.3803
(st.err.)	(2.5156)	(1.4430)	(1.1195)	(0.7886)
Δ	0.9362	3.6542	6.1797	12.2826
(st.err.)	(1.9718)	(3.6568)	(4.7207)	(6.4424)
Pr (Fit G_w)	0.3145	0.1823	0.1197	0.0473
$\widehat{\tau}$	3.2541	3.1414	3.0891	3.0439
(st.err.)	(0.4107)	(0.2691)	(0.2151)	(0.1545)
$\widehat{\alpha}$	3.9590	4.2679	4.4653	4.5488
(st.err.)	(5.1291)	(5.1448)	(4.6451)	(4.3040)
$\widehat{\phi}$	89.4824	94.8136	98.0404	100.6558
(st.err.)	(32.6317)	(29.9797)	(28.0450)	(23.9417)
$\widehat{B}_{b,10}$	30.4023	30.1723	30.0509	29.9689
(st.err.)	(2.5520)	(1.4849)	(1.1596)	(0.8236)

Table 4.1: Summary statistics for G_w and G_b for varying sample size, when these distributions are fitted to data that has undergone a type I censoring regime from G_b with parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

$$\Delta = \frac{S_0\left(2\widehat{\beta}\right)}{2} - \frac{\left\{S_0\left(\widehat{\beta}\right)\right\}\sum_{i=1}^{N} y_i^{\widehat{\beta}}}{N}.$$

We run a series of simulations for varying stopping times and sample sizes, and obtain MLEs and the probability of choosing the Burr distribution over the Weibull. As usual, we set

$$\alpha = 4,\ \tau = 3,\ \phi = 100,$$

and for each particular set of parameters, run the simulation at least 10000 times to ensure accurate average MLEs are computed. Table 4.1 summarises the results for varying sample sizes, keeping the stopping time fixed at 80, whilst Table 4.2 shows the results for varying stopping times, keeping the sample size fixed at 1000. When we vary the sample size, the standard errors for both the Weibull and Burr MLEs, and their quantile functions, decrease as the sample size increases. However, when compared to complete counterparts in Table 2.6, the standard errors for $\widehat{\alpha}$ and $\widehat{\phi}$ do appear smaller than expected for lower stopping times. In these cases, however, we do prefer to fit the Weibull distribution far more often than in the complete counterpart; for example, for $n = 100$ complete data values, we preferred to fit the Weibull 16% of the time, whilst this figure doubled for an equivalent sample size and a stopping time of 80 time units. Thus, one possible explanation for observing this decrease

n	1000	1000	1000	1000	1000
y_c	50	70	100	120	200
$\widehat{\beta}$	2.9251	2.8229	2.6786	2.6188	2.5611
(st.err.)	(0.1408)	(0.0920)	(0.0675)	(0.0645)	(0.0710)
$\widehat{\theta}$	64.7115	65.8055	66.8310	67.0847	67.2608
(st.err.)	(1.4894)	(0.9027)	(0.8398)	(0.8541)	(0.8744)
$\widehat{B}_{w,10}$	29.9329	29.6351	28.8398	28.3998	27.9251
(st.err.)	(0.8528)	(0.8227)	(0.7476)	(0.7387)	(0.7774)
Δ	0.1938	5.7325	34.2733	62.3454	122.5729
(st.err.)	(1.1996)	(4.4817)	(11.4573)	(17.3012)	(46.2282)
Pr (Fit G_w)	0.4624	0.1381	0.0013	0.0002	0
$\widehat{\tau}$	3.1466	3.0840	3.0148	3.0072	3.0026
(st.err.)	(0.1820)	(0.1666)	(0.1425)	(0.1328)	(0.1252)
$\widehat{\alpha}$	2.0402	3.8614	4.6429	4.4702	4.4050
(st.err.)	(2.0585)	(3.4429)	(3.0844)	(2.0114)	(1.7421)
$\widehat{\phi}$	73.6130	93.6691	103.6084	103.1065	102.9414
(st.err.)	(17.1919)	(23.0911)	(20.3540)	(16.6039)	(14.5409)
$\widehat{B}_{b,10}$	29.8380	29.9943	29.9276	29.9009	29.8966
(st.err.)	(0.8483)	(0.8363)	(0.8277)	(0.8272)	(0.8253)

Table 4.2: Summary statistics for G_w and G_b for varying y_c, when these distributions are fitted to data that has undergone a type I censoring regime from G_b with parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

in sample standard errors from G_b, is that the figures are calculated after conditioning on $\Delta > 0$. The probability of fitting G_w decreases as n increases, and is as high as 31% for a sample size of 100. When we compare estimates of the quantile functions with a true value of 29.8848, we see that $\widehat{B}_{w,10}$ is closer to $B_{b,10}$ for smaller n, but the standard errors are then larger. On the whole $\widehat{B}_{b,10}$ matches up very well to its true value across all sample sizes. When we vary the stopping times, again the standard errors for the MLEs of G_w and $\widehat{B}_{w,10}$ decrease as y_c increases, but this is not true for $\widehat{\alpha}$ and $\widehat{\phi}$ from G_b. In fact, we observe smaller standard errors for these estimates for lower stopping times. The probability of choosing G_w over G_b is very high for small y_c, and for $y_c = 50$, we prefer to fit Weibull over Burr 46% of the time, even though the sample size is as large as 1000.

We continue by obtaining the theory necessary to explain these simulated values, and, as in Chapter 3, first assume that no mis-specification has taken place. Thus, we derive the EFI matrix for G_w and G_b for type I censored data.

4.1.4 Analysing data using the correct distribution

Asymptotic results for the distribution of the MLEs when no mis-specification has taken place are well known, and have been outlined in Chapter 3. We use these results to obtain the EFI matrices for the Weibull and Burr distributions below.

The Weibull distribution

On examining second derivatives from G_w, we see that we will require $E\left[Y^\beta\right]$, $E\left[Y^\beta \ln Y\right]$ and $E\left[Y^\beta (\ln Y)^2\right]$. We use the relationship between truncated Weibull and Negative Exponential distributions to compute these, and first consider

$$E\left[Y^\beta\right] = \theta^\beta E\left[Z\right],$$

for Y, Z following (4.1) and (4.2), respectively. We first note that we can write expectations with respect to Z in terms of the incomplete gamma function given by (1.7). Thus, we have

$$E\left[Z^r\right] = \frac{\Gamma\left(z_c, r+1\right)}{1-\exp\left(-z_c\right)},$$

which simplifies to

$$\frac{r\Gamma\left(z_c, r\right) - z_c^r \exp\left(-z_c\right)}{1-\exp\left(-z_c\right)},$$

on using (1.11). We further exploit properties of the incomplete gamma function, and use (1.10) to write

$$E\left[Z^r\right] = \frac{r z_c^r U_1\left(z_c, r\right) - z_c^r \exp\left(-z_c\right)}{1-\exp\left(-z_c\right)}. \tag{4.9}$$

Thus, on substituting $r = 1$ into (4.9), we have

$$E\left[Y^\beta\right] = \theta^\beta \left\{\frac{1-\exp\left(-z_c\right) - z_c \exp\left(-z_c\right)}{1-\exp\left(-z_c\right)}\right\}. \tag{4.10}$$

We differentiate (4.9) with respect to r to obtain

$$E\left[Z^r \ln Z\right] = \frac{z_c^r}{1-\exp\left(-z_c\right)} \left\{\left(1 + r\ln z_c\right) U_1\left(z_c, r\right) - rU_2\left(z_c, r\right) - \ln z_c \exp\left(-z_c\right)\right\}, \tag{4.11}$$

and use this to write

$$E\left[Y^\beta \ln Y\right] = \theta^\beta \left\{\ln\theta E\left[Z\right] + \beta^{-1} E\left[Z \ln Z\right]\right\}$$

as

$$\frac{\theta^\beta}{1-\exp\left(-z_c\right)} \left[\begin{array}{c} \left\{1-\exp\left(-z_c\right) - z_c\exp\left(-z_c\right)\right\}\left\{\ln\theta + \beta^{-1}\ln z_c\right\} \\ +\beta^{-1}\left\{1-\exp\left(-z_c\right)\right\} - z_c\beta^{-1}U_2\left(z_c, 1\right) \end{array}\right].$$

Finally, we consider $E\left[Y^{\beta}(\ln Y)^2\right]$, and differentiate (4.11) with respect to r, to write

$$E\left[Z^r(\ln Z)^2\right] = \frac{z_c^r}{1-\exp(-z_c)}\left\{\begin{array}{c}\ln z_c(2+r\ln z_c)U_1(z_c,r)-2(1+r\ln z_c)U_2(z_c,r)\\ +2rU_3(z_c,r)-(\ln z_c)^2\exp(-z_c)\end{array}\right\}.$$

Thus, we have

$$E\left[Y^{\beta}(\ln Y)^2\right] = \frac{\theta^{\beta}z_c}{1-\exp(-z_c)}\left[\begin{array}{c}\left(\left\{\ln\theta+\beta^{-1}(1+\ln z_c)\right\}^2-\beta^{-2}\right)\frac{\{1-\exp(-z_c)\}}{z_c}\\ -2\beta^{-1}\left\{\ln\theta+\beta^{-1}(1+\ln z_c)\right\}U_2(z_c,1)\\ +2\beta^{-2}U_3(z_c,1)\\ -\exp(-z_c)\left(\ln\theta+\beta^{-1}\ln z_c\right)^2\end{array}\right].$$

Using these expectations, we now compute the elements in the EFI matrix for G_w. We consider $E\left[\frac{\partial^2 l_w}{\partial\theta^2}\right]$, and note that we must first condition on the random variable N, before taking expectations with respect to Y. Thus, with a slight abuse of notation, we write

$$-E\left[\left.\frac{\partial^2 l_w}{\partial\theta^2}\right|N\right] = -\beta\theta^{-2}\left[\begin{array}{c}N-N(\beta+1)\left\{\frac{1-\exp(-z_c)-z_c\exp(-z_c)}{1-\exp(-z_c)}\right\}\\ -(n-N)(\beta+1)z_c\end{array}\right].$$

Thus, on taking expectations with respect to N, we have

$$-E\left[\frac{\partial^2 l_w}{\partial\theta^2}\right] = n\beta^{-2}\theta^{-2}\left\{1-\exp(-z_c)\right\}.$$

We use similar arguments to derive the other elements of the EFI matrix; these are given by

$$-E\left[\frac{\partial^2 l_w}{\partial\beta^2}\right] = n\beta^{-2}\left[\begin{array}{c}1-\exp(-z_c)+\ln z_c\{1-\exp(-z_c)\}\{2+\ln z_c\}\\ -2z_c(1+\ln z_c)U_2(z_c,1)+2z_cU_3(z_c,1)\end{array}\right],$$

and

$$-E\left[\frac{\partial^2 l_w}{\partial\beta\partial\theta}\right] = -n\theta^{-1}\left[\begin{array}{c}1-\exp(-z_c)-z_c\exp(-z_c)+\ln z_c\{1-\exp(-z_c)\}\\ -z_cU_2(z_c,1)+\beta\ln\theta z_c\exp(-z_c)\end{array}\right].$$

This list provides us with the elements required to compute the EFI matrix from the Weibull distribution, after the data has undergone type I censoring. We carry out a similar analysis for the Burr distribution below.

The Burr distribution

To compute the elements of (3.1), we make use of the relationship between the Burr and its two parameter counterpart. Note that $Z = \frac{Y}{\phi}$ has a two parameter Burr distribution, and we can compute expectations in terms of this variable. We also use a similar technique as that established in Watkins (1997), who writes expectations as E_α to emphasize the role of the parameter α. Thus, we write

$$E\left[\frac{Z^\tau}{1+Z^\tau}\right] \equiv E_\alpha\left[\frac{Z^\tau}{1+Z^\tau}\right],$$

and, by exploiting the role of α in G_b, can write this as

$$\frac{\alpha}{\alpha+1}E_{\alpha+1}\left[Z^\tau\right].$$

On closer examination of the second derivatives, we see that we require $E\left[Z^m\right]$, $E\left[Z^m \ln Z\right]$ and $E\left[Z^m \left(\ln Z\right)^2\right]$. We consider these next, and first compute $E\left[Z^m\right]$. This is given by

$$E[Z^m] = \frac{\alpha}{q_{z_c,\alpha}}B_{\frac{z_c^\tau}{1+z_c^\tau}}\left(m_1, \alpha - m_0\right),$$

where $B_z\left(a,b\right)$ is the incomplete Beta function (1.15), and

$$m_i = \frac{m}{\tau} + i.$$

We obtain an expression for $E[Z^m(\ln Z)^r]$ by differentiating this expectation r times with respect to m. In order to obtain the derivative of the incomplete Beta function, we re-write it using its hypergeometric counterpart given by (1.16). Then, using (15.3.4) of Abramowitz and Stegun (1972), we have

$$E[Z^m] = \frac{\alpha z_c^{m+\tau}}{m_1 q_{z_c,\alpha}}F_{2,1}(\{m_1, \alpha+1\};\{m_2\};-z_c^\tau). \tag{4.12}$$

To differentiate this function with respect to m, we first note that the arguments m_1, m_2 are separated by unity. Watkins and Johnson (2002) consider hypergeometric functions with this property, and prove that a hypergeometric function of the form

$$f_q(a,b,z) = F_{q+1,q}(\{a,...,a,b\};\{a+1,...,a+1\};z)$$

has derivative

$$\frac{qa^{q-1}bz}{(a+1)^{q+1}}f_{q+1}\left(a+1,b+1,z\right).$$

They also examine relationships between neighbouring hypergeometric functions, and establish two results which help simplify such functions. We use these results to simplify our expectations. Thus, the derivative of $E[Z^m]$ with respect to m is given by

$$E\left[Z^m \ln Z\right] = \frac{\alpha z_c^{\tau+m}}{q_{z_c,\alpha} m_1}\left\{\ln z_c f_1(m_1, \alpha+1, -z_c^\tau) - \frac{f_2(m_1, \alpha+1, -z_c^\tau)}{\tau m_1}\right\}. \quad (4.13)$$

We differentiate this function again to obtain

$$E[Z^m (\ln Z)^2] = \frac{\alpha z_c^{\tau+m}}{q_{z_c,\alpha} m_1}\left[\begin{array}{c}\{\ln z_c\}^2 f_1(m_1, \alpha+1, -z_c^\tau) - \frac{2\ln z_c f_2(m_1,\alpha+1,-z_c^\tau)}{\tau m_1} \\ + \frac{2 f_3(m_1,\alpha+1,-z_c^\tau)}{(\tau m_1)^2}\end{array}\right]. \quad (4.14)$$

We use these results for the two parameter Burr distribution to list the expectations required for the EFI matrix from the three parameter model :

$$\begin{aligned}
E\left[\frac{\left(\frac{Y}{\phi}\right)^\tau}{1+\left(\frac{Y}{\phi}\right)^\tau}\right] &= \frac{\alpha z_c^{2\tau}}{2 q_{z_c,\alpha}} f_1\left(2, \alpha+2, -z_c^\tau\right), \\
E\left[\frac{\left(\frac{Y}{\phi}\right)^\tau}{\left\{1+\left(\frac{Y}{\phi}\right)^\tau\right\}^2}\right] &= \frac{\alpha z_c^{2\tau}}{2 q_{z_c,\alpha}} f_1\left(2, \alpha+3, -z_c^\tau\right), \\
E\left[\frac{\left(\frac{Y}{\phi}\right)^\tau \ln\left(\frac{Y}{\phi}\right)}{1+\left(\frac{Y}{\phi}\right)^\tau}\right] &= \frac{\alpha z_c^{2\tau}}{2 q_{z_c,\alpha}}\left\{\begin{array}{c}\ln z_c f_1\left(2, \alpha+2, -z_c^\tau\right) \\ -\frac{1}{2\tau} f_2\left(2, \alpha+2, -z_c^\tau\right)\end{array}\right\}, \\
E\left[\frac{\left(\frac{Y}{\phi}\right)^\tau \ln\left(\frac{Y}{\phi}\right)}{\left\{1+\left(\frac{Y}{\phi}\right)^\tau\right\}^2}\right] &= \frac{\alpha z_c^{2\tau}}{2 q_{z_c,\alpha}}\left\{\begin{array}{c}\ln z_c f_1\left(2, \alpha+3, -z_c^\tau\right) - \\ \frac{f_2(2,\alpha+3,-z_c^\tau)}{2\tau}\end{array}\right\}, \\
E\left[\frac{\left(\frac{Y}{\phi}\right)^\tau \left\{\ln\left(\frac{Y}{\phi}\right)\right\}^2}{\left\{1+\left(\frac{Y}{\phi}\right)^\tau\right\}^2}\right] &= \frac{\alpha z_c^{2\tau}}{2 q_{z_c,\alpha}}\left[\begin{array}{c}\{\ln z_c\}^2 f_1\left(2, \alpha+3, -z_c^\tau\right) \\ -\tau^{-1} \ln z_c f_2\left(2, \alpha+3, -z_c^\tau\right) \\ + \frac{f_3(2,\alpha+3,-z_c^\tau)}{2\tau^2}\end{array}\right].
\end{aligned}$$

We now list the elements in the EFI matrix for G_b:

$$\begin{aligned}
-E\left[\frac{\partial^2 l_b}{\partial \alpha^2}\right] &= n\alpha^{-2} q_{z_c,\alpha}, \\
-E\left[\frac{\partial^2 l_b}{\partial \tau^2}\right] &= n\left[\begin{array}{c}\tau^{-2} q_{z_c,\alpha} + \frac{\alpha(1-q_{z_c,\alpha}) z_c^\tau (\ln z_c)^2}{\{1+z_c^\tau\}^2} + \\ \frac{\alpha(\alpha+1) z_c^{2\tau}}{2}\left\{\begin{array}{c}(\ln z_c)^2 f_1\left(2, \alpha+3, -z_c^\tau\right) \\ -\tau^{-1} \ln z_c f_2\left(2, \alpha+3, -z_c^\tau\right) \\ + \frac{f_3(2,\alpha+3,-z_c^\tau)}{2\tau^2}\end{array}\right\}\end{array}\right], \\
-E\left[\frac{\partial^2 l_b}{\partial \phi^2}\right] &= n\tau\phi^{-2}\left[\begin{array}{c}\frac{\alpha\{1-q_{z_c,\alpha}\} z_c^\tau}{1+z_c^\tau}\left\{1 + \frac{\tau}{1+z_c^\tau}\right\} - q_{z_c,\alpha} \\ + \frac{\alpha(\alpha+1) z_c^{2\tau} f_1(2,\alpha+2,-z_c^\tau)}{2} + \frac{\tau\alpha(\alpha+1) z_c^{2\tau} f_1(2,\alpha+3,-z_c^\tau)}{2}\end{array}\right], \\
-E\left[\frac{\partial^2 l_b}{\partial \alpha \partial \tau}\right] &= n z_c^\tau\left[\begin{array}{c}\frac{\alpha z_c^\tau}{2}\left\{\ln z_c f_1\left(2, \alpha+2, -z_c^\tau\right) - \frac{f_2(2,\alpha+2,-z_c^\tau)}{2\tau}\right\} \\ + \frac{\{1-q_{z_c,\alpha}\}\ln z_c}{1+z_c^\tau}\end{array}\right], \\
-E\left[\frac{\partial^2 l_b}{\partial \alpha \partial \phi}\right] &= -n\tau\phi^{-1} z_c^\tau\left[\frac{\alpha z_c^\tau f_1(2,\alpha+2,-z_c^\tau)}{2} + \left\{\frac{1-q_{z_c,\alpha}}{1+z_c^\tau}\right\}\right], \\
-E\left[\frac{\partial^2 l_b}{\partial \tau \partial \phi}\right] &= n\phi^{-1}\left[\begin{array}{c}q_{z_c,\alpha} - \frac{\alpha\{1-q_{z_c,\alpha}\} z_c^\tau}{1+z_c^\tau} - \frac{\alpha\tau\{1-q_{z_c,\alpha}\} z_c^\tau \ln z_c}{\{1+z_c^\tau\}^2} \\ - \frac{\alpha(\alpha+1) z_c^{2\tau}}{2} f_1\left(2, \alpha+2, -z_c^\tau\right) \\ - \frac{\tau\alpha(\alpha+1) z_c^{2\tau}}{2}\left\{\begin{array}{c}\ln z_c f_1\left(2, \alpha+3, -z_c^\tau\right) \\ - \frac{f_2(2,\alpha+3,-z_c^\tau)}{2\tau}\end{array}\right\}\end{array}\right].
\end{aligned}$$

This matrix will be used to derive the mean and variance of $\widehat{B}_{b,10}$; see (3.2).

We continue by examining the effects of fitting the incorrect distribution to data from G_b.

4.1.5 Analysing data using the incorrect distribution

In Chapter 3, we used the entropy function to obtain theoretical counterparts to $\widehat{\beta}$ and $\widehat{\theta}$ of G_w, when this distribution was fitted to a set of data simulated from G_b, and no censoring took place. We now examine similar results for data that has undergone type I censoring.

Recall that the theoretical counterparts to $\widehat{\beta}$ and $\widehat{\theta}$ are obtained by maximising the expected value of the log-likelihood l_w, with respect to G_b. Since type I censoring introduces another random variable N into the analysis, we must first condition on this. Using (4.3), we see that

$$\begin{aligned} E\left[l_w\right| N] &= E\left[\begin{array}{c} N\ln\beta - N\beta\ln\theta + (\beta-1)\sum_{i=1}^{N}\ln Y_i \\ -\theta^{-\beta}\left\{\sum_{i=1}^{N} Y_i^{\beta} + (n-N)\, y_c^{\beta}\right\} \end{array}\right] \\ &= N\ln\beta - N\beta\ln\theta + N(\beta-1)\,E\left[\ln Y\right] - N\theta^{-\beta}E\left[Y^{\beta}\right] - (n-N)\,\theta^{-\beta}y_c^{\beta}. \end{aligned}$$

We use (4.12) to write

$$E[Y^m] = \frac{\alpha\phi^m z_c^{\tau+m}}{m_1 q_{z_c,\alpha}} f_1\left(m_1, \alpha+1, -z_c^{\tau}\right), \tag{4.15}$$

and note that

$$E[\ln Y] = \ln\phi + E[\ln Z],$$

where $E\left[\ln Z\right]$ is obtained by setting $m = 0$ in (4.13). Thus,

$$E[\ln Y] = \ln\phi + \frac{\alpha z_c^{\tau}}{q_{z_c,\alpha}}\left\{\begin{array}{c} \ln z_c f_1\left(1, \alpha+1, -z_c^{\tau}\right) \\ -\tau^{-1} f_2\left(1, \alpha+1, -z_c^{\tau}\right) \end{array}\right\}. \tag{4.16}$$

The conditional entropy now becomes

$$\begin{aligned} & N\ln\beta - N\beta\ln\theta - (n-N)\left(\frac{\phi}{\theta}\right)^{\beta} z_c^{\beta} \\ & +N(\beta-1)\left[\ln\phi + \frac{\alpha z_c^{\tau}}{q_{z_c,\alpha}}\left\{\begin{array}{c} \ln z_c f_1\left(1, \alpha+1, -z_c^{\tau}\right) - \\ \tau^{-1} f_2\left(1, \alpha+1, -z_c^{\tau}\right) \end{array}\right\}\right] \\ & -\frac{N\alpha\theta^{-\beta}\phi^{\beta} z_c^{\tau+\beta}}{q_{z_c,\alpha}\left(\frac{\beta}{\tau}+1\right)} f_1\left(\frac{\beta}{\tau}+1, \alpha+1, -z_c^{\tau}\right), \end{aligned}$$

and, taking expectations with respect to N, we see that

$$\begin{aligned} E_b &= n(\ln\beta)\,q_{z_c,\alpha} - n\beta(\ln\theta)\,q_{z_c,\alpha} - nz_c^{\beta} + nq_{z_c,\alpha}\left(\frac{\phi}{\theta}\right)^{\beta} z_c^{\frac{\beta}{\tau}} \\ &\quad + nq_{z_c,\alpha}(\beta-1)\left[\ln\phi + \frac{\alpha z_c^{\tau}}{q_{z_c,\alpha}}\left\{\begin{array}{c} \ln z_c f_1(1,\alpha+1,-z_c^{\tau}) - \\ \tau^{-1} f_2(1,\alpha+1,-z_c^{\tau}) \end{array}\right\}\right] \\ &\quad - \frac{n\alpha z_c^{\tau+\frac{\beta}{\tau}}\left(\frac{\phi}{\theta}\right)^{\beta}}{\left(\frac{\beta}{\tau}+1\right)} f_1\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^{\tau}\right). \end{aligned} \tag{4.17}$$

We differentiate (4.17) with respect to θ to obtain

$$\frac{\partial E_b}{\partial\theta} = n\beta\theta^{-1}\left\{\begin{array}{c} \left(\frac{\phi}{\theta}\right)^{\beta} z_c^{\beta} - q_{z_c,\alpha} - \left(\frac{\phi}{\theta}\right)^{\beta} z_c^{\beta} q_{z_c,\alpha} \\ + \frac{\alpha\left(\frac{\phi}{\theta}\right)^{\beta} z_c^{\tau+\beta}}{\left(\frac{\beta}{\tau}+1\right)} f_1\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^{\tau}\right) \end{array}\right\},$$

and equate this to zero to get

$$\theta = \left\{\frac{\phi^{\beta} z_c^{\beta}(1-q_{z_c,\alpha}) + \alpha z_c^{\tau+\beta}\phi^{\beta}\left(\frac{\beta}{\tau}+1\right)^{-1} f_1\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^{\tau}\right)}{q_{z_c,\alpha}}\right\}^{\frac{1}{\beta}}.$$

We let

$$g_1 = \phi^{\beta} z_c^{\beta}(1-q_{z_c,\alpha}) + \alpha z_c^{\tau+\beta}\phi^{\beta}\left(\frac{\beta}{\tau}+1\right)^{-1} f_1\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^{\tau}\right),$$

so

$$\theta = \left\{\frac{g_1}{q_{z_c,\alpha}}\right\}^{\frac{1}{\beta}}. \tag{4.18}$$

For future use, we also let

$$\begin{aligned} g_2 &= \frac{\partial g_1}{\partial\beta} = \phi^{\beta} z_c^{\beta}\ln(\phi z_c)\{1-q_{z_c,\alpha}\} + \\ &\quad \frac{\alpha\phi^{\beta} z_c^{\tau+\beta}}{\left(\frac{\beta}{\tau}+1\right)}\left\{\begin{array}{c} \ln(\phi z_c) f_1\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^{\tau}\right) \\ - \frac{f_2\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^{\tau}\right)}{\tau\left(\frac{\beta}{\tau}+1\right)} \end{array}\right\}, \end{aligned}$$

and

$$g_3 = \frac{\partial g_2}{\partial \beta} = \phi^\beta z_c^\beta \{\ln(\phi z_c)\}^2 (1 - q_{z_c,\alpha}) + \frac{\alpha\phi^\beta z_c^{\tau+\beta}}{\left(\frac{\beta}{\tau}+1\right)} \left[\begin{array}{c} \{\ln(\phi z_c)\}^2 f_1\left(\frac{\beta}{\tau}+1, \alpha+1, -z_c^\tau\right) \\ -2\left\{\frac{\ln(\phi z_c)}{\tau\left(\frac{\beta}{\tau}+1\right)} - \frac{1}{\tau^2\left(\frac{\beta}{\tau}+1\right)^2} + 1\right\} f_2\left(\frac{\beta}{\tau}+1, \alpha+1, -z_c^\tau\right) \\ +2f_3\left(\frac{\beta}{\tau}+1, \alpha+1, -z_c^\tau\right) \end{array} \right].$$

Inserting (4.18) into (4.17) will give the profile entropy

$$E_b^* = nq_{z_c,\alpha} \left[\begin{array}{c} \ln\beta - \ln g_1 + \ln q_{z_c,\alpha} - 1 + \\ (\beta - 1)\left[\ln\phi + \frac{\alpha z_c^\tau}{q_{z_c,\alpha}} \left\{ \begin{array}{c} \ln z_c f_1(1, \alpha+1, -z_c^\tau) - \\ \tau^{-1} f_2(1, \alpha+1, -z_c^\tau) \end{array} \right\}\right] \end{array} \right], \tag{4.19}$$

where here, we are using the fact that

$$\frac{nq_{z_c,\alpha}\left\{g_1 - \phi^\beta z_c^\beta(1 - q_{z_c,\alpha})\right\}}{g_1} = \frac{n\alpha\phi^\beta z_c^{\tau+\beta} q_{z_c,\alpha} f_1\left(\frac{\beta}{\tau}+1, \alpha+1, -z_c^\tau\right)}{g_1\left(\frac{\beta}{\tau}+1\right)}.$$

We could plot (4.19) with respect to β and locate the maximum, or use the Newton-Raphson process to obtain the root of

$$\frac{dE_b^*}{d\beta} = nq_{z_c,\alpha} \left[\begin{array}{c} \beta^{-1} - \frac{g_2}{g_1} + \ln\phi + \\ \frac{\alpha z_c^\tau}{q_{z_c,\alpha}} \left\{\ln z_c f_1(1, \alpha+1, -z_c^\tau) - \tau^{-1} f_2(1, \alpha+1, -z_c^\tau)\right\} \end{array} \right], \tag{4.20}$$

with derivative

$$\frac{d^2 E_b^*}{d\beta^2} = -nq_{z_c,\alpha}\left\{\beta^{-2} + \left(\frac{g_1 g_3 - g_2^2}{g_1^2}\right)\right\}.$$

We are now in a position to obtain theoretical counterparts to the MLEs from G_w. We first note, by examining (4.19), that maximising the entropy function is not affected by the sample size. Thus, entropy values for $n = 100$, $y_c = 80$ are the same as those for $n = 576$, $y_c = 80$. They are, however, affected by the stopping time. Figures 4.1 and 4.2 show how β_0 and θ_0 vary with y_c, when we set $\tau = 3$, $\alpha = 4$ and $\phi = 100$. We observe how the entropy values tend to their complete counterparts as y_c increases. We also note how well the entropy values match up with simulated counterparts in Table 4.2. We could, at this point, examine relationships that exist between parameters from G_b and β_0 and θ_0. The above investigation shows how β_0 and θ_0 vary with y_c, and, on examining (4.20), we also note a relationship between β_0 and ϕ; this is contrary to what we observed in our investigations on complete data. Due to this added complication, we omit any further details on relationships between true and mis-specified parameters.

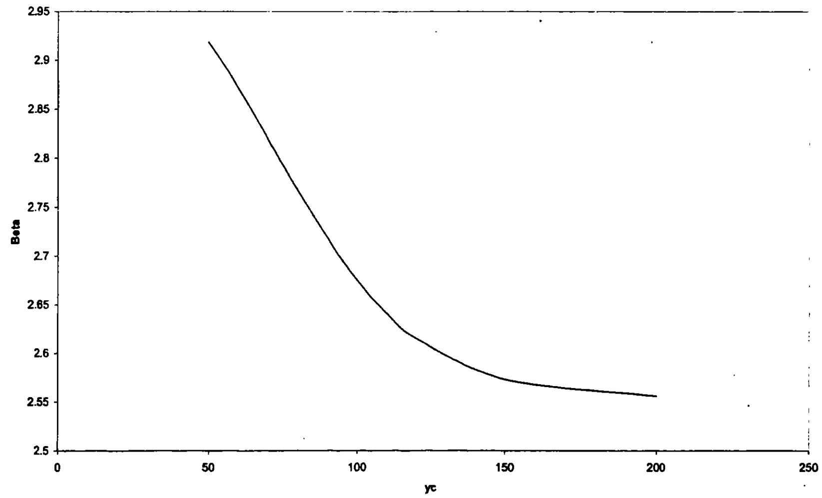


Figure 4.1: β_0 versus y_c for $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

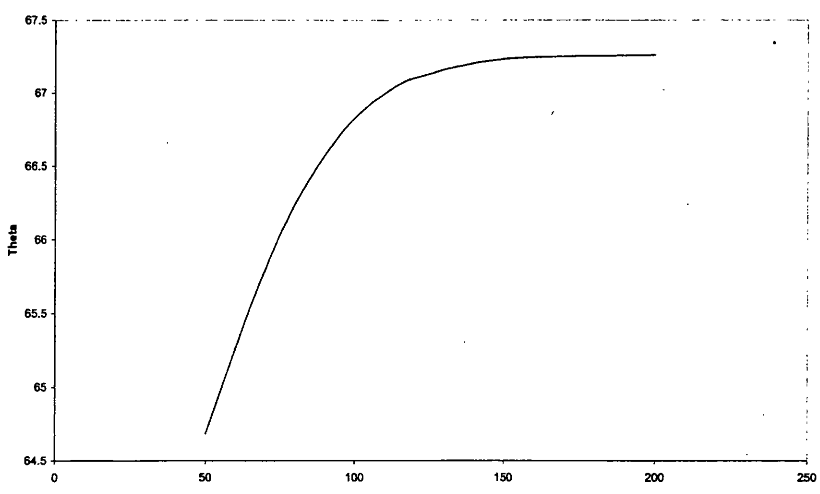


Figure 4.2: θ_0 versus y_c for $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

Now that we have theoretical counterparts to the MLEs from the Weibull distribution, we can use these to derive the asymptotic distribution of these estimates. We consider this below.

The variance structure of the mis-specified MLEs

From our work on the distribution of the mis-specified Weibull MLEs for complete data, we know that, asymptotically, $\left(\widehat{\beta}, \widehat{\theta}\right)'$ will be Normally distributed with mean vector $(\beta_0, \theta_0)'$, and variance covariance matrix given by (3.18). As a result, we require expected values of second derivatives from the Weibull distribution, and variances and covariances between score functions. On closer examination of these functions, we require expressions for $E[Y^m]$, $E[Y^m \ln Y]$ and $E\left[Y^m (\ln Y)^2\right]$, where Y is a random variable from G_b. $E[Y^m]$ is given by (4.15). We compute the remaining two expected values by using the relationship between G_b and its two parameter counterpart. Thus, we have

$$E[Y^m \ln Y] = \frac{\alpha\phi^m z_c^{\tau+m}}{q_{z_c,\alpha}\left(\frac{m}{\tau}+1\right)} \left\{ \begin{array}{c} \ln(\phi z_c) f_1\left(\frac{m}{\tau}+1, \alpha+1, -z_c^\tau\right) \\ -\frac{f_2\left(\frac{m}{\tau}+1,\alpha+1,-z_c^\tau\right)}{\tau\left(\frac{m}{\tau}+1\right)} \end{array} \right\}, \tag{4.21}$$

and

$$E\left[Y^m (\ln Y)^2\right] = \frac{\alpha\phi^m z_c^{\tau+m}}{q_{z_c,\alpha}\left(\frac{m}{\tau}+1\right)} \left\{ \begin{array}{c} \{\ln(\phi z_c)\}^2 f_1\left(\frac{m}{\tau}+1, \alpha+1, -z_c^\tau\right) \\ -\frac{2\ln(\phi z_c) f_2\left(\frac{m}{\tau}+1,\alpha+1,-z_c^\tau\right)}{\tau\left(\frac{m}{\tau}+1\right)} \\ +\frac{2 f_3\left(\frac{m}{\tau}+1,\alpha+1,-z_c^\tau\right)}{\tau^2\left(\frac{m}{\tau}+1\right)^2} \end{array} \right\}. \tag{4.22}$$

Using these, we first list the elements of the matrix A:

$$-E\left[\frac{\partial^2 l_w}{\partial\theta^2}\right] = n\beta\theta^{-2} \left[\begin{array}{c} \frac{\alpha(\beta+1)\left(\frac{\phi}{\theta}\right)^\beta z_c^{\tau+\beta} f_1\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^\tau\right)}{\frac{\beta}{\tau}+1} \\ +z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta (\beta+1)\{1-q_{z_c,\alpha}\} - q_{z_c,\alpha} \end{array} \right],$$

$$-E\left[\frac{\partial^2 l_w}{\partial\beta^2}\right] = n \left[\begin{array}{c} \beta^{-2} q_{z_c,\alpha} + z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta (1-q_{z_c,\alpha}) \left\{\ln\left(\frac{z_c\phi}{\theta}\right)\right\}^2 \\ +\frac{\alpha z_c^{\tau+\beta}\left(\frac{\phi}{\theta}\right)^\beta}{\left(\frac{\beta}{\tau}+1\right)} \left\{ \begin{array}{c} \left\{\ln\left(\frac{z_c\phi}{\theta}\right)\right\}^2 f_1\left(\frac{\beta}{\tau}+1, \alpha+1, -z_c^\tau\right) \\ -\frac{2\ln\left(\frac{z_c\phi}{\theta}\right) f_2\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^\tau\right)}{\tau\left(\frac{\beta}{\tau}+1\right)} \\ +\frac{2 f_3\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^\tau\right)}{\tau^2\left(\frac{\beta}{\tau}+1\right)^2} \end{array} \right\} \end{array} \right],$$

and

$$-E\left[\frac{\partial^2 l_w}{\partial\beta\partial\theta}\right] = n\theta^{-1}\left[\begin{array}{c} q_{z_c,\alpha} - \frac{\alpha z_c^{\tau+\beta}\left(\frac{\phi}{\theta}\right)^\beta}{\left(\frac{\beta}{\tau}+1\right)}\left(\begin{array}{c}\left\{1+\beta\ln\left(\frac{z_c\phi}{\theta}\right)\right\} f_1\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^\tau\right) \\ -\frac{\beta}{\tau\left(\frac{\beta}{\tau}+1\right)} f_2\left(\frac{\beta}{\tau}+1,\alpha+1,-z_c^\tau\right)\end{array}\right) \\ -z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\left\{1-q_{z_c,\alpha}\right\}\left\{1+\beta\ln\left(\frac{z_c\phi}{\theta}\right)\right\}\end{array}\right].$$

Next, we consider the elements in the matrix V, and first examine $Var\left(\frac{\partial l_w}{\partial\theta}\right)$. We have

$$Var\left(\frac{\partial l_w}{\partial\theta}\right) = E_N\left[E_Y\left[\left(\frac{\partial l}{\partial\theta}\right)^2\middle| N\right]\right] - \left(E_N\left[E_Y\left[\frac{\partial l}{\partial\theta}\middle| N\right]\right]\right)^2.$$

We consider

$$E_Y\left[\left(\frac{\partial l_w}{\partial\theta}\right)^2\middle| N\right] = \beta^2\theta^{-2}\left[\begin{array}{c} N^2-(n-N)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\left\{2N-(n-N)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\right\} \\ -2N\theta^{-\beta}\left\{N-(n-N)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\right\}E\left[Y^\beta\right]+N\theta^{-2\beta}E\left[Y^{2\beta}\right] \\ +N(N-1)\theta^{-2\beta}\left(E\left[Y^\beta\right]\right)^2\end{array}\right],$$

and note that substituting expressions for expectations in the above function produce no simplifications; we just end up with higher order hypergeometric functions with different arguments. Thus, we will keep the variances and covariance between the score functions from the Weibull distribution in terms of their expectations. After taking expectations with respect to N, we see that

$$E\left[\left(\frac{\partial l_w}{\partial\theta}\right)^2\right] = \beta^2\theta^{-2}\left\{\begin{array}{c} E\left[N^2\right]-2z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta E\left[N(n-N)\right] \\ +z_c^{2\beta}\left(\frac{\phi}{\theta}\right)^{2\beta}E\left[(n-N)^2\right]-2\theta^{-\beta}E\left[Y^\beta\right]E\left[N^2\right] \\ +2\theta^{-\beta}z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta E\left[Y^\beta\right]E\left[N(n-N)\right]+\theta^{-2\beta}E\left[Y^{2\beta}\right]E[N] \\ +\theta^{-2\beta}\left(E\left[Y^\beta\right]\right)^2E\left[N(N-1)\right]\end{array}\right\}.$$

Since

$$E[N] = nq_{z_c,\alpha},$$

we have

$$Var(N) = nq_{z_c,\alpha}\left\{1-q_{z_c,\alpha}\right\},$$

and

$$E\left[N^2\right] = nq_{z_c,\alpha}\left\{1 - q_{z_c,\alpha}\right\} + n^2 q_{z_c,\alpha}^2,$$

and hence can obtain the necessary expectations with respect to N. Using these, we have

$$E\left[\left(\frac{\partial l_w}{\partial \theta}\right)^2\right] = n\beta^2\theta^{-2}\left[\begin{array}{c} q_{z_c,\alpha}\left\{1 - q_{z_c,\alpha} + nq_{z_c,\alpha}\right\} - \\ 2(n-1)\,q_{z_c,\alpha}\left\{1 - q_{z_c,\alpha}\right\} z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta \\ +z_c^{2\beta}\left(\frac{\phi}{\theta}\right)^{2\beta}\left\{\begin{array}{c} n\left\{1 - q_{z_c,\alpha}\right\}^2 + \\ q_{z_c,\alpha}\left\{1 - q_{z_c,\alpha}\right\}\end{array}\right\} \\ -2q_{z_c,\alpha}\theta^{-\beta}\left\{\begin{array}{c} 1 - q_{z_c,\alpha} + nq_{z_c,\alpha} - \\ (n-1)\,z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\left\{1 - q_{z_c,\alpha}\right\}\end{array}\right\} E\left[Y^\beta\right] \\ +q_{z_c,\alpha}\theta^{-2\beta}E\left[Y^{2\beta}\right] + \\ (n-1)\,q_{z_c,\alpha}^2\theta^{-2\beta}\left(E\left[Y^\beta\right]\right)^2 \end{array}\right].$$

We also require

$$\left(E_N\left[E_Y\left[\left.\frac{\partial l_w}{\partial \theta}\right| N\right]\right]\right)^2$$

in order to calculate $Var\left(\frac{\partial l_w}{\partial \theta}\right)$, and see that

$$E\left[\frac{\partial l_w}{\partial \theta}\right] = n\beta\theta^{-1}\left\{q_{z_c,\alpha}\theta^{-\beta}E\left[Y^\beta\right] + z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\left(1 - q_{z_c,\alpha}\right) - q_{z_c,\alpha}\right\}.$$

Thus,

$$\left(E\left[\frac{\partial l_w}{\partial \theta}\right]\right)^2 = n^2\beta^2\theta^{-2}\left[\begin{array}{c} \theta^{-2\beta}q_{z_c,\alpha}^2\left(E\left[Y^\beta\right]\right)^2 + \\ 2z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta q_{z_c,\alpha}\left\{1 - q_{z_c,\alpha}\right\}\theta^{-\beta}E\left[Y^\beta\right] \\ -2z_c^{2\beta}\left(\frac{\phi}{\theta}\right)^{2\beta}\theta^{-\beta}q_{z_c,\alpha}^2 E\left[Y^\beta\right] + \left\{1 - q_{z_c,\alpha}\right\}^2 \\ -2z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta q_{z_c,\alpha}\left\{1 - q_{z_c,\alpha}\right\} + q_{z_c,\alpha}^2 \end{array}\right],$$

and so

$$Var\left(\frac{\partial l_w}{\partial \theta}\right) = nq_{z_c,\alpha}\beta^2\theta^{-2}\left\{\begin{array}{c} 1 - q_{z_c,\alpha} + 2z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\left\{1 - q_{z_c,\alpha}\right\} + \\ z_c^{2\beta}\left(\frac{\phi}{\theta}\right)^{2\beta}\left\{1 - q_{z_c,\alpha}\right\} \\ -2\theta^{-\beta}\left\{1 - q_{z_c,\alpha}\right\}\left\{1 + z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\right\} E\left[Y^\beta\right] \\ +\theta^{-2\beta}E\left[Y^{2\beta}\right] - q_{z_c,\alpha}\theta^{-2\beta}\left(E\left[Y^\beta\right]\right)^2 \end{array}\right\}.$$

We use similar techniques to calculate the variance of $\frac{\partial l_w}{\partial \beta}$, and first consider

$$E_Y\left[\left(\frac{\partial l_w}{\partial \beta}\right)^2 \middle| N\right].$$

Using (4.5), this simplifies to

$$\begin{aligned}
&P^2 + 2NPE[\ln Y] - 2N\theta^{-\beta}(P - \ln\theta)E\left[Y^\beta \ln Y\right] + 2NP\theta^{-\beta}\ln\theta E\left[Y^\beta\right] \\
&+NE\left[(\ln Y)^2\right] + N(N-1)(E[\ln Y])^2 - 2N\theta^{-\beta}E\left[Y^\beta(\ln Y)^2\right] \\
&-2N(N-1)\theta^{-\beta}E\left[Y^\beta \ln Y\right]E[\ln Y] + 2N(N-1)\theta^{-\beta}\ln\theta E\left[Y^\beta\right]E[\ln Y] \\
&+N\theta^{-2\beta}E\left[Y^{2\beta}(\ln Y)^2\right] + N(N-1)\theta^{-2\beta}\left(E\left[Y^\beta \ln Y\right]\right)^2 \\
&-2N\theta^{-2\beta}\ln\theta E\left[Y^{2\beta}\ln Y\right] - 2N(N-1)\theta^{-2\beta}\ln\theta E\left[Y^\beta \ln Y\right]E\left[Y^\beta\right] \\
&+N\theta^{-2\beta}(\ln\theta)^2 E\left[Y^{2\beta}\right] + N(N-1)\theta^{-2\beta}(\ln\theta)^2\left(E\left[Y^\beta\right]\right)^2,
\end{aligned}$$

where

$$P = N\beta^{-1} - N\ln\theta - (n-N)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta \ln\left(\frac{z_c\phi}{\theta}\right).$$

We now take expectations with respect to N, and see that we require

$$\begin{aligned}
c_1 &= E[P] = n\left[\begin{array}{c}\beta^{-1}q_{z_c,\alpha} - q_{z_c,\alpha}\ln\theta - \\ z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\{1-q_{z_c,\alpha}\}\ln\left(\frac{z_c\phi}{\theta}\right)\end{array}\right], \\
c_2 &= E[NP] = nq_{z_c,\alpha}\left[\begin{array}{c}\{1-q_{z_c,\alpha}+nq_{z_c,\alpha}\}\left(\beta^{-1}-\ln\theta\right) - \\ (n-1)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\{1-q_{z_c,\alpha}\}\ln\left(\frac{z_c\phi}{\theta}\right)\end{array}\right],
\end{aligned}$$

and

$$c_3 = E\left[P^2\right] = n\left[\begin{array}{c} q_{z_c,\alpha}\{1-q_{z_c,\alpha}+nq_{z_c,\alpha}\}\left(\beta^{-1}-\ln\theta\right)^2 - \\ 2(n-1)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta q_{z_c,\alpha}\{1-q_{z_c,\alpha}\}\ln\left(\frac{z_c\phi}{\theta}\right)\left\{\beta^{-1}-\ln\theta\right\} \\ +z_c^{2\beta}\left(\frac{\phi}{\theta}\right)^{2\beta}\{1-q_{z_c,\alpha}\}\{n(1-q_{z_c,\alpha})+q_{z_c,\alpha}\}\left\{\ln\left(\frac{z_c\phi}{\theta}\right)\right\}^2\end{array}\right].$$

Hence, $E\left[\left(\frac{\partial l_w}{\partial \beta}\right)^2\right]$ becomes

$$\begin{gathered}
c_3 + 2c_2 E\left[\ln Y\right] - 2\theta^{-\beta}\left(c_2 - nq_{z_c,\alpha}\ln\theta\right)E\left[Y^\beta \ln Y\right] + 2c_2\theta^{-\beta}\ln\theta E\left[Y^\beta\right] \\
+nq_{z_c,\alpha}E\left[(\ln Y)^2\right] + n\left(n-1\right)q_{z_c,\alpha}^2 E\left[\ln Y\right]^2 - 2n\theta^{-\beta}q_{z_c,\alpha}E\left[Y^\beta\left(\ln Y\right)^2\right] \\
-2n\left(n-1\right)\theta^{-\beta}q_{z_c,\alpha}^2 E\left[Y^\beta \ln Y\right]E\left[\ln Y\right] + 2n\left(n-1\right)\theta^{-\beta}\ln\theta q_{z_c,\alpha}^2 E\left[Y^\beta\right]E\left[\ln Y\right] \\
+n\theta^{-2\beta}q_{z_c,\alpha}E\left[Y^{2\beta}\left(\ln Y\right)^2\right] + n\left(n-1\right)\theta^{-2\beta}q_{z_c,\alpha}^2 E\left[Y^\beta \ln Y\right]^2 \\
-2n\theta^{-2\beta}\ln\theta q_{z_c,\alpha}E\left[Y^{2\beta}\ln Y\right] - 2n\left(n-1\right)\theta^{-2\beta}\ln\theta q_{z_c,\alpha}^2 E\left[Y^\beta \ln Y\right]E\left[Y^\beta\right] \\
+n\theta^{-2\beta}\left(\ln\theta\right)^2 q_{z_c,\alpha}E\left[Y^{2\beta}\right] + n\left(n-1\right)\theta^{-2\beta}\left(\ln\theta\right)^2 q_{z_c,\alpha}^2 E\left[Y^\beta\right]^2
\end{gathered}$$

In order to calculate $Var\left(\frac{\partial l_w}{\partial \beta}\right)$, we also require

$$\left(E\left[\frac{\partial l_w}{\partial \beta}\right]\right)^2 = \left(E_N\left[E_Y\left[\left.\frac{\partial l_w}{\partial \beta}\right| N\right]\right]\right)^2,$$

and note that

$$E\left[\frac{\partial l_w}{\partial \beta}\right] = c_1 + nq_{z_c,\alpha}E\left[\ln Y\right] - n\theta^{-\beta}q_{z_c,\alpha}E\left[Y^\beta \ln Y\right] + n\theta^{-\beta}\ln\theta q_{z_c,\alpha}E\left[Y^\beta\right].$$

Hence, $\left(E\left[\frac{\partial l_w}{\partial \beta}\right]\right)^2$ equates to

$$\begin{aligned}
&c_1^2 + 2nc_1 q_{z_c,\alpha}E\left[\ln Y\right] - 2nc_1\theta^{-\beta}q_{z_c,\alpha}E\left[Y^\beta \ln Y\right] \\
&+2nc_1\theta^{-\beta}\ln\theta q_{z_c,\alpha}E\left[Y^\beta\right] + n^2 q_{z_c,\alpha}^2 E\left[\ln Y\right]^2 \\
&-2n^2\theta^{-\beta}q_{z_c,\alpha}^2 E\left[\ln Y\right]E\left[Y^\beta \ln Y\right] + 2n^2\theta^{-\beta}\ln\theta q_{z_c,\alpha}^2 E\left[\ln Y\right]E\left[Y^\beta\right] \\
&+n^2\theta^{-2\beta}q_{z_c,\alpha}^2 E\left[Y^\beta \ln Y\right]^2 - 2n^2\theta^{-2\beta}\ln\theta q_{z_c,\alpha}^2 E\left[Y^\beta \ln Y\right]E\left[Y^\beta\right] \\
&+n^2\theta^{-2\beta}\left(\ln\theta\right)^2 q_{z_c,\alpha}E\left[Y^\beta\right]^2 + n^2 q_{z_c,\alpha}^2\left(\ln\theta\right)^2,
\end{aligned}$$

and, after much simplification, we see that $Var\left(\frac{\partial l_w}{\partial \beta}\right)$ becomes

$$nq_{z_c,\alpha}\left[\begin{array}{c} \lambda^2\left\{1-q_{z_c,\alpha}\right\}+2\lambda\left\{1-q_{z_c,\alpha}\right\}E\left[\ln Y\right]-2\theta^{-\beta}\left\{\lambda\left(1-q_{z_c,\alpha}\right)-\ln\theta\right\}E\left[Y^\beta\ln Y\right] \\ +2\lambda\theta^{-\beta}\ln\theta\left\{1-q_{z_c,\alpha}\right\}E\left[Y^\beta\right]+E\left[\left(\ln Y\right)^2\right]-q_{z_c,\alpha}E\left[\ln Y\right]^2 \\ -2\theta^{-\beta}E\left[Y^\beta\left(\ln Y\right)^2\right]+2\theta^{-\beta}q_{z_c,\alpha}E\left[Y^\beta\ln Y\right]E\left[\ln Y\right] \\ -2\theta^{-\beta}\ln\theta q_{z_c,\alpha}E\left[Y^\beta\right]E\left[\ln Y\right]+\theta^{-2\beta}E\left[Y^{2\beta}\left(\ln Y\right)^2\right] \\ -\theta^{-2\beta}q_{z_c,\alpha}E\left[Y^\beta\ln Y\right]^2-2\theta^{-2\beta}\ln\theta E\left[Y^{2\beta}\ln Y\right] \\ +2\theta^{-2\beta}\ln\theta q_{z_c,\alpha}E\left[Y^\beta\ln Y\right]E\left[Y^\beta\right]+\theta^{-2\beta}\left(\ln\theta\right)^2E\left[Y^{2\beta}\right] \\ -q_{z_c,\alpha}\theta^{-2\beta}\left(\ln\theta\right)^2E\left[Y^\beta\right]^2 \end{array}\right],$$

where

$$\lambda=\beta^{-1}-\ln\theta+z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\ln\left(\frac{z_c\phi}{\theta}\right).$$

The final function we require is

$$Cov\left(\frac{\partial l_w}{\partial\theta},\frac{\partial l_w}{\partial\beta}\right)=E_N\left[E_Y\left[\frac{\partial l_w}{\partial\theta}\frac{\partial l_w}{\partial\beta}\middle|N\right]\right]-E_N\left[E_Y\left[\frac{\partial l_w}{\partial\theta}\middle|N\right]\right]E_N\left[E_Y\left[\frac{\partial l_w}{\partial\beta}\middle|N\right]\right].$$

From the calculation of the variances of both score functions, we see that we already have expressions for

$$E\left[\frac{\partial l_w}{\partial\beta}\middle|N\right]\text{ and }E\left[\frac{\partial l_w}{\partial\theta}\middle|N\right].$$

We write $E\left[\frac{\partial l_w}{\partial\theta}\frac{\partial l_w}{\partial\beta}\right]$ as

$$\beta\theta^{-1}\left[\begin{array}{c} \left(nc_1-c_2\right)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta-c_2 \\ +nq_{z_c,\alpha}\left\{\left(n-1\right)\left(1-q_{z_c,\alpha}\right)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta-\left(1-q_{z_c,\alpha}\right)-nq_{z_c,\alpha}\right\}E\left[\ln Y\right] \\ -n\theta^{-\beta}q_{z_c,\alpha}\left\{\left(n-1\right)\left(1-q_{z_c,\alpha}\right)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta-\left(1-q_{z_c,\alpha}\right)-nq_{z_c,\alpha}-1\right\}E\left[Y^\beta\ln Y\right] \\ +\theta^{-\beta}\left\{\begin{array}{c} n\left(n-1\right)q_{z_c,\alpha}\left(1-q_{z_c,\alpha}\right)z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\ln\theta \\ -nq_{z_c,\alpha}\ln\theta\left(1-q_{z_c,\alpha}+nq_{z_c,\alpha}\right)+c_2 \end{array}\right\}E\left[Y^\beta\right] \\ +n\left(n-1\right)\theta^{-\beta}q_{z_c,\alpha}E\left[Y^\beta\right]E\left[\ln Y\right]-n\theta^{-2\beta}q_{z_c,\alpha}E\left[Y^{2\beta}\ln Y\right] \\ -n\left(n-1\right)\theta^{-2\beta}q_{z_c,\alpha}E\left[Y^\beta\right]E\left[Y^\beta\ln Y\right]+n\theta^{-2\beta}\ln\theta q_{z_c,\alpha}E\left[Y^{2\beta}\right] \\ +n\left(n-1\right)\theta^{-2\beta}\ln\theta q_{z_c,\alpha}\left(E\left[Y^\beta\right]\right)^2 \end{array}\right],$$

and $E\left[\frac{\partial l_w}{\partial \theta}\right]E\left[\frac{\partial l_w}{\partial \beta}\right]$ as

$$n\beta\theta^{-1}\left[\begin{array}{c} c_1\left\{z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta(1-q_{z_c,\alpha})-q_{z_c,\alpha}\right\} \\ +nq_{z_c,\alpha}\left\{z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta(1-q_{z_c,\alpha})-q_{z_c,\alpha}\right\}E[\ln Y] \\ -n\theta^{-\beta}q_{z_c,\alpha}\left\{z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta(1-q_{z_c,\alpha})-q_{z_c,\alpha}\right\}E\left[Y^\beta\ln Y\right] \\ +\theta^{-\beta}q_{z_c,\alpha}\left\{nz_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\ln\theta(1-q_{z_c,\alpha})-n\ln\theta q_{z_c,\alpha}+c_1\right\}E\left[Y^\beta\right] \\ +n\theta^{-\beta}q_{z_c,\alpha}^2E\left[Y^\beta\right]E[\ln Y]-n\theta^{-2\beta}q_{z_c,\alpha}^2E\left[Y^\beta\right]E\left[Y^\beta\ln Y\right] \\ +n\theta^{-2\beta}\ln\theta q_{z_c,\alpha}^2E\left[Y^\beta\right]^2 \end{array}\right].$$

Hence, we see that

$$Cov\left(\frac{\partial l_w}{\partial \theta},\frac{\partial l_w}{\partial \beta}\right)$$

can be expressed as

$$n\beta\theta^{-1}q_{z_c,\alpha}\left[\begin{array}{c} -\{1-q_{z_c,\alpha}\}(\beta^{-1}-\ln\theta)-z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\{1-q_{z_c,\alpha}\}\left\{\ln\left(\frac{z_c\phi}{\theta^2}\right)+\beta^{-1}\right\} \\ -z_c^{2\beta}\left(\frac{\phi}{\theta}\right)^{2\beta}\{1-q_{z_c,\alpha}\}\ln\left(\frac{z_c\phi}{\theta}\right)-\{1-q_{z_c,\alpha}\}\left\{1+z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\right\}E[\ln Y] \\ +\theta^{-\beta}\left[2-q_{z_c,\alpha}+z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\{1-q_{z_c,\alpha}\}\right]E\left[Y^\beta\ln Y\right] \\ +\theta^{-\beta}\{1-q_{z_c,\alpha}\}\left\{\beta^{-1}-2\ln\theta+z_c^\beta\left(\frac{\phi}{\theta}\right)^\beta\ln\left(\frac{z_c\phi}{\theta^2}\right)\right\}E\left[Y^\beta\right] \\ -\theta^{-\beta}q_{z_c,\alpha}E[\ln Y]E\left[Y^\beta\right]-\theta^{-2\beta}E\left[Y^{2\beta}\ln Y\right]+\theta^{-2\beta}q_{z_c,\alpha}E\left[Y^\beta\right]E\left[Y^\beta\ln Y\right] \\ +\theta^{-2\beta}\ln\theta E\left[Y^{2\beta}\right]-\theta^{-2\beta}\ln\theta q_{z_c,\alpha}E\left[Y^\beta\right]^2 \end{array}\right].$$

We now have all the required terms and expectations to obtain the variance covariance matrix for the MLEs of the Weibull distribution, when this distribution is fitted to Burr data that has undergone type I censoring. This matrix will be used to calculate the variance of $\widehat{B}_{w,10}$, using (3.25). We do not give explicit results for these functions since the algebra is very complicated, especially with the addition of hypergeometric functions that appear in the expectations. Thus, as already stated, we leave all functions in terms of expected values.

In the next section, we check our theoretical results by comparing them with simulated values.

Agreement between theoretical and sample results

We present the theoretical standard errors of the MLEs from the mis-specified Weibull distribution for varying sample sizes and stopping times, and compare these to the sample

n	100	300	500	1000
y_c	80	80	80	80
St.err. $\left(\widehat{\beta}\right)$	0.2527	0.1459	0.1130	0.0799
St.err. $\left(\widehat{\theta}\right)$	2.6995	1.5586	1.2073	0.8537
$E\left[\widehat{B}_{w,10}\right]$	29.3690	29.3690	29.3690	29.3690
St.err. $\left(\widehat{B}_{w,10}\right)$	2.5129	1.4508	1.1238	0.7946
St.err. $(\widehat{\tau})$	0.5364	0.3097	0.2399	0.1696
St.err. $(\widehat{\alpha})$	7.8599	4.5379	3.5151	2.4855
St.err. $\left(\widehat{\phi}\right)$	81.8466	47.2542	36.6029	25.8822
St.err. $\left(\widehat{B}_{b,10}\right)$	2.6700	1.5415	1.1941	0.8443

Table 4.3: Theoretical standard errors for the MLEs from G_w and G_b for varying n. Data is subjected to type I censoring, and simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

n	1000	1000	1000	1000	1000
y_c	50	70	100	120	200
St.err. $\left(\widehat{\beta}\right)$	0.1405	0.0919	0.0677	0.0643	0.0703
St.err. $\left(\widehat{\theta}\right)$	1.4777	0.9017	0.8484	0.8621	0.8787
$E\left[\widehat{B}_{w,10}\right]$	29.9160	29.6233	28.8172	28.3809	27.8857
St.err. $\left(\widehat{B}_{w,10}\right)$	0.8606	0.8202	0.7552	0.7379	0.7730
St.err. $(\widehat{\tau})$	0.2875	0.1929	0.1440	0.1327	0.1250
St.err. $(\widehat{\alpha})$	12.1483	3.6561	1.5340	1.2191	1.0482
St.err. $\left(\widehat{\phi}\right)$	112.9568	36.8359	16.7029	13.5529	11.7869
St.err. $\left(\widehat{B}_{b,10}\right)$	0.8632	0.7145	0.8378	0.6907	0.6791

Table 4.4: Theoretical standard errors for the MLEs from G_w and G_b for varying y_c. Data is subjected to type I censoring, and simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

results presented in Tables 4.1 and 4.2. We also include details of theoretical results from G_b, for both the Burr MLEs and $\widehat{B}_{b,10}$. Note that we are comparing all values of B_{10} with the true value given by

$$\phi\left(0.9^{\frac{-1}{\alpha}} - 1\right)^{\frac{1}{\tau}} = 100\left(0.9^{\frac{-1}{4}} - 1\right)^{\frac{1}{3}} = 29.8848.$$

The results for varying sample sizes are summarised in Table 4.3, and for varying y_c we present the results in Table 4.4. We outline the main points when varying n and y_c below.

When n varies For G_w, there is excellent agreement between observed and expected results across all sample sizes, and the theoretical standard errors decrease as n increases. When we compare sample and theoretical values from G_b, we observe some surprising results. Firstly, the sample standard errors for $\widehat{\alpha}$ and $\widehat{\phi}$ are smaller than their theoretical counterparts when n is small, and also less than their complete counterparts. An intuitive explanation for this has been provided when we examined simulated values. When we examine $\widehat{B}_{b,10}$, the agreement between expected and simulated results are good, even when n is relatively small.

When y_c varies When we examine results for G_w, we note that for large stopping times, the theoretical standard errors of $\widehat{\beta}$, $\widehat{\theta}$ and $\widehat{B}_{w,10}$ tend to their counterparts from the complete scenario. There is one surprising outcome; when we vary the censoring time, and keep the sample size fixed at 1000, there is a decrease in the theoretical standard errors of both $\widehat{\beta}$ and $\widehat{\theta}$ for smaller stopping times. This result seems counter-intuitive, since smaller stopping times result in a larger proportion of censored observations. Thus, we would expect standard errors to increase. We also examine the effects of how the theoretical mean and standard errors for $\widehat{B}_{w,10}$ change with varying stopping times, and note that, for smaller values of y_c, larger values for the mean of $\widehat{B}_{w,10}$ are observed. For example, when $y_c = 50$, the theoretical mean of $\widehat{B}_{w,10}$ is 29.9160; this figure decreases to 27.8857 for large stopping times. Such results imply that when we have more censoring, the time to which 10% of the data values fail is much higher than if we had a complete data set. Since we are comparing this mean with the true value of 29.8848 from the Burr distribution, then we see an improved agreement between means when we have more censoring. We do not pay a very significant penalty with regard to the standard error for censored data, and just observe a small rise for equivalent sample sizes from complete and censored data when $y_c = 50$. In fact, for $y_c = 120$, we actually observe a smaller standard error and a closer mean than the complete counterpart. An intuitive explanation for this could be linked to the fact that when we censor, we only have to match the lower tail of the distribution function which contains the estimate for B_{10}. When we have complete data, we have to match both tails, so the estimate might not be as good as the censored counterpart.

A further point concerns the stopping time of 100 time units. This choice of censoring time causes problems with convergence of hypergeometric functions, since the stopping time is the same as the scale parameter from the Burr distribution, and the final argument in the hypergeometric function

$$F_{p,q}\left(\{\alpha_1,\alpha_2,...,\alpha_p\},\{\rho_1,\rho_2,...,\rho_q\},z\right),$$

then has a modulus of unity, so that we are on the boundary for convergence. There are rules governing the convergence of this type of function if the final argument is of unit size, and if $q = p + 1$; see, for example, Luke (1969). However, theoretical results still match up very well with simulated counterparts for such stopping times.

Below, we carry out a similar investigation for type II censored data.

4.2 Type II censoring

This section extends the discussion in previous chapters to deal with type II censored data. Its structure is similar to the above section on type I censoring, and we examine the theory required to fit Weibull and Burr distributions to data from a type II censoring regime. We then look at the effects of fitting G_w and G_b to data from G_b via simulations, and consider changes as we vary the observation number that we stop the experiment at. We continue by deriving the EFI matrices for G_w and G_b when we assume no mis-specification has taken place, and finish with a discussion on the entropy function and the agreement between theoretical and sample results.

4.2.1 ML estimation for G_w under type II censoring

We begin by considering the Weibull distribution, and derive the necessary theory to fit this distribution to a set of data that has undergone a type II censoring regime. We assume that the data $y_1, y_2, ..., y_n$ comes from a Weibull distribution with pdf given by (1.2). The experiment is terminated after the r^{th} component has failed. Thus, the first r observations will each have their own distribution, depending on their order, whilst the remaining $(n-r)$ will be censored with distribution function equal to $Y_{(r:n)}$, where we refer to Chapter 1 for our notation representing order statistics. The likelihood and log-likelihood are reasonably simple to construct; we first derive the likelihood, given, without loss of generality, by

$$
\begin{aligned}
L_w &= \prod_{i=1}^{r} g_w\left(y_i;\beta,\theta\right) \prod_{i=r+1}^{n} \left[1 - G_w\left(y_i;\beta,\theta\right)\right] \\
&= \prod_{i=1}^{r} \frac{\beta y_{(i:n)}^{\beta-1}}{\theta^{\beta}} \exp\left\{-\left(\frac{y_{(i:n)}}{\theta}\right)^{\beta}\right\} \prod_{i=r+1}^{n} \exp\left\{-\left(\frac{y_{(r:n)}}{\theta}\right)^{\beta}\right\},
\end{aligned}
$$

from which the log-likelihood is

$$
l_w = r\ln\beta + (\beta-1)\,S_{f,1}\left(0\right) - r\beta\ln\theta - \theta^{-\beta}\left\{S_{f,0}\left(\beta\right) + S_{c,0}\left(\beta\right)\right\}, \tag{4.23}
$$

where

$$
S_{f,j}\left(k\right) = \sum_{i=1}^{r} y_{(i:n)}^{k}\left(\ln y_{(i:n)}\right)^{j},
$$

$$
S_{c,j}\left(k\right) = \sum_{i=r+1}^{n} y_{(i:n)}^{k}\left(\ln y_{(i:n)}\right)^{j} = \left(n-r\right) y_{(r:n)}^{k}\left(\ln y_{(r:n)}\right)^{j},
$$

with the property that

$$\frac{\partial S_{*,j}(k)}{\partial k} = S_{*,j+1}(k) \quad \text{for } * = f \text{ or } c.$$

The score function is based on the elements

$$\frac{\partial l_w}{\partial \theta} = -r\beta\theta^{-1} + \beta\theta^{-\beta-1}\left\{S_{f,0}(\beta) + S_{c,0}(\beta)\right\}, \tag{4.24}$$

and

$$\frac{\partial l_w}{\partial \beta} = r\beta^{-1} + S_{f,1}(0) - r\ln\theta - \theta^{-\beta}\left[S_{f,1}(\beta) + S_{c,1}(\beta) - \ln\theta\left\{S_{f,0}(\beta) + S_{c,0}(\beta)\right\}\right]. \tag{4.25}$$

We equate (4.24) to zero, to obtain an expression for θ in terms of β. This is given by

$$\theta = \left\{\frac{S_{f,0}(\beta) + S_{c,0}(\beta)}{r}\right\}^{\frac{1}{\beta}}. \tag{4.26}$$

Inserting (4.26) into (4.23) yields a profile log-likelihood given by

$$l_w^* = r\ln\beta + (\beta - 1)S_{f,1}(0) - r\ln\left\{S_{f,0}(\beta) + S_{c,0}(\beta)\right\} + r\ln r - r,$$

and a profile score function with respect to β of the form

$$\frac{dl_w^*}{d\beta} = r\beta^{-1} + S_{f,1}(0) - r\left\{\frac{S_{f,1}(\beta) + S_{c,1}(\beta)}{S_{f,0}(\beta) + S_{c,0}(\beta)}\right\}. \tag{4.27}$$

We use the Newton-Raphson method to locate the root of (4.27), and so require

$$\frac{d^2 l_w^*}{d\beta^2} = -r\beta^{-2} - r\left[\frac{\left\{S_{f,2}(\beta) + S_{c,2}(\beta)\right\}\left\{S_{f,0}(\beta) + S_{c,0}(\beta)\right\} - \left\{S_{f,1}(\beta) + S_{c,1}(\beta)\right\}^2}{\left\{S_{f,0}(\beta) + S_{c,0}(\beta)\right\}^2}\right].$$

In later sections, we will compute the EFI matrix of the Weibull MLEs. This requires results on second derivatives, which we list below:

$$\frac{\partial^2 l_w}{\partial \theta^2} = r\beta\theta^{-2} - \beta(\beta+1)\theta^{-\beta-2}\left\{S_{f,0}(\beta) - S_{c,0}(\beta)\right\}, \tag{4.28}$$

$$\frac{\partial^2 l_w}{\partial \beta^2} = -r\beta^{-2} - \theta^{-\beta}\left[\begin{array}{c} S_{f,2}(\beta) + S_{c,2}(\beta) - 2\ln\theta\left\{S_{f,1}(\beta) + S_{c,1}(\beta)\right\} \\ +(\ln\theta)^2\left\{S_{f,0}(\beta) + S_{c,0}(\beta)\right\}\end{array}\right], \tag{4.29}$$

and

$$\frac{\partial^2 l_w}{\partial\beta\partial\theta} = -r\theta^{-1} + \theta^{-\beta-1}\left[\begin{array}{c} \beta\left\{S_{f,1}(\beta) + S_{c,1}(\beta)\right\} \\ +(1-\beta\ln\theta)\left\{S_{f,0}(\beta) + S_{c,0}(\beta)\right\} \end{array}\right]. \tag{4.30}$$

We continue by deriving a similar set of results for G_b

4.2.2 ML estimation for G_b under type II censoring

As with the Weibull distribution, we construct likelihood and log-likelihood functions when the data has undergone type II censoring. We assume that the experiment has been stopped after the r^{th} observation has failed, and that the data has an underlying Burr distribution. Wingo (1993) has considered ML estimation of parameters from G_b when the data has undergone type II censoring. We write the likelihood as

$$\begin{aligned} L_b &= \prod_{i=1}^{r} g_b(y_i;\tau,\alpha,\phi) \prod_{i=r+1}^{n} \left\{1 - G_b(y_i;\tau,\alpha,\phi)\right\} \\ &= \prod_{i=1}^{r} \frac{\alpha\tau y_{(i:n)}^{\tau-1}}{\phi^{\tau}} \left\{1 + \left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}\right\}^{-\alpha-1} \prod_{i=r+1}^{n} \left\{1 + \left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}\right\}^{-\alpha}, \end{aligned}$$

from which the log-likelihood is

$$l_b = r\ln\alpha + r\ln\tau - r\tau\ln\phi + (\tau-1)S_{f,1}(0) - (\alpha+1)T_f - \alpha T_c, \tag{4.31}$$

where

$$\begin{aligned} T_f &= \sum_{i=1}^{r} \ln\left\{1 + \left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}\right\}, \\ T_c &= \sum_{i=r+1}^{n} \ln\left\{1 + \left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}\right\} = (n-r)\ln\left\{1 + \left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}\right\}. \end{aligned}$$

For future use, we introduce the following notation; note that f and c indicate a failed and censored item, whilst the other sub-scripts represent differentiation with respect to τ or ϕ :

$$T_{f,1,0} = \frac{\partial T_f}{\partial\tau} = \sum_{i=1}^{r} \left\{ \frac{\left(\frac{y_{(i:n)}}{\phi}\right)^{\tau} \ln\left(\frac{y_{(i:n)}}{\phi}\right)}{1 + \left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}} \right\},$$

and, in general,

$$T_{f,m,0} = \frac{\partial^m T_f}{\partial\tau^m} = \sum_{i=1}^{r} \left[\frac{\left(\frac{y_{(i:n)}}{\phi}\right)^{\tau} \left\{\ln\left(\frac{y_{(i:n)}}{\phi}\right)\right\}^{m}}{\left\{1 + \left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}\right\}^{m}} \right].$$

Similarly, for the censored data

$$T_{c,1,0} = \frac{\partial T_c}{\partial \tau},$$

and

$$T_{c,m,0} = (n-r)\frac{\left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}\left\{\ln\left(\frac{y_{(r:n)}}{\phi}\right)\right\}^{m}}{\left\{1+\left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}\right\}^{m}}.$$

If we consider derivatives with respect to the parameter ϕ, then we see that

$$\begin{aligned} T_{f,0,1} &= \frac{\partial T_f}{\partial \phi} = \frac{-\tau}{\phi}\sum_{i=1}^{r}\left\{\frac{\left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}}\right\}, \\ T_{f,0,2} &= \frac{\tau}{\phi^2}\sum_{i=1}^{r}\left\{\frac{\left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}}\right\} + \frac{\tau^2}{\phi^2}\sum_{i=1}^{r}\left[\frac{\left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{y_{(i:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right], \end{aligned}$$

and

$$\begin{aligned} T_{c,0,1} &= \frac{-(n-r)\left(\frac{\tau}{\phi}\right)\left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}}, \\ T_{c,0,2} &= -\frac{\tau(n-r)}{\phi^2}\left[\frac{\left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}} + \frac{\tau\left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{y_{(r:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right]. \end{aligned}$$

The score function contains the following elements

$$\begin{aligned} \frac{\partial l_b}{\partial \alpha} &= r\alpha^{-1} - T_f - T_c, \\ \frac{\partial l_b}{\partial \tau} &= r\tau^{-1} - r\ln\phi + S_{f,1}(0) - (\alpha+1)T_{f,1,0} - \alpha T_{c,1,0}, \end{aligned} \tag{4.32}$$

and

$$\frac{\partial l_b}{\partial \phi} = -r\tau\phi^{-1} - (\alpha+1)T_{f,0,1} - \alpha T_{c,0,1}.$$

Equating (4.32) to zero yields an expression for α in terms of the other two parameters. We have

$$\alpha = \frac{r}{T_f + T_c},$$

and inserting this into (4.31) gives the profile log-likelihood

$$\begin{aligned} l_b^* &= r\ln\left(\frac{r}{T_f+T_c}\right)+r\ln\tau-r\tau\ln\phi+(\tau-1)\,S_{f,1}(0) \\ &\quad -T_f\left(1+\frac{r}{T_f+T_c}\right)-T_c\left(\frac{r}{T_f+T_c}\right). \end{aligned}$$

We will not need the two profile score functions with respect to τ and ϕ, since we use the approach outlined in Chapter 2 to compute MLEs from the Burr distribution; this requires the full score vector. At this point, we also give second derivatives for the three parameters of the Burr distribution, since they will be used in the derivation of the EFI matrix. We list these below:

$$\frac{\partial^2 l_b}{\partial\alpha^2}=-r\alpha^{-2},$$

$$\frac{\partial^2 l_b}{\partial\tau^2}=-r\tau^{-2}-(\alpha+1)\,T_{f,2,0}-\alpha T_{c,2,0},$$

$$\frac{\partial^2 l_b}{\partial\phi^2}=r\tau\phi^{-2}-(\alpha+1)\,T_{f,0,2}-\alpha T_{c,0,2},$$

$$\frac{\partial^2 l_b}{\partial\alpha\partial\tau}=-T_{f,1,0}-T_{c,1,0},$$

$$\frac{\partial^2 l_b}{\partial\alpha\partial\phi}=-T_{f,0,1}-T_{c,0,1},$$

and

$$\frac{\partial^2 l_b}{\partial\tau\partial\phi}=-r\phi+(\alpha+1)\,T_{f,1,1}+\alpha T_{c,1,1},$$

where

$$T_{f,1,1}=\frac{-1}{\phi}\sum_{i=1}^{r}\left\{\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right\}-\frac{\tau}{\phi}\sum_{i=1}^{r}\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(i:n)}}{\phi}\right)}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^2}\right],$$

and

$$T_{c,1,1}=\frac{-(n-r)\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}{\phi\left\{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\right\}}-\frac{(n-r)\,\tau\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(r:n)}}{\phi}\right)}{\phi\left\{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\right\}^2}.$$

Now that we have the results to fit both the Weibull and Burr distributions to type II censored data, we continue by examining the effects of fitting G_w and G_b to data from G_b.

4.2.3 Fitting G_w to G_b data

In this section, we summarise a series of simulations that fit G_w and G_b to a set of data from G_b that has undergone a type II censoring regime. As in previous chapters, we use the discriminating Δ to determine which distribution offers the better fit, and if $\Delta < 0$ conclude that the Weibull distribution is an improvement over the Burr. We use the algorithm outlined in Watkins (1999) to fit the Burr distribution, with one additional feature to allow for observations exceeding the r^{th} item to be censored. Summary statistics for varying sample sizes are shown in Table 4.5, and for varying r, these are summarised in Table 4.6; both correspond to the usual set of Burr parameter values given by

$$\tau = 3,\ \alpha = 4,\ \phi = 100.$$

We also show the average value of Δ for each set of simulations, and the probability of choosing the Weibull over the Burr. Note that the censoring time chosen when we vary the sample size remains the same, and we censor 20% of the observations. This figure has been chosen arbitrarily; we just wish to see how MLEs change for a censored sample, as we lower the overall sample size. The tables show an increase in the probability of fitting the Weibull distribution over the Burr, as the number of censored observations increases. We also see the scale parameter of the Weibull distribution decrease, and the shape increase, as r is lowered. For smaller sample sizes, just as in the complete case, the probability of fitting the Weibull is higher. We also observe a surprising high probability for fitting the Weibull over the Burr, when we have a large sample size, but censor more observations. For example, when $n = 1000$, $r = 600$, the probability for fitting the Weibull distribution is 0.2526, slightly higher that the equivalent figure for $n = 300$, $r = 240$. We also note, as with type I censoring, surprisingly small standard errors associated with $\widehat{\alpha}$ and $\widehat{\phi}$ for lower sample sizes; these are less than the complete counterparts in Table 2.6.

We continue by deriving the theory necessary to explain these results, and first consider the best scenario, where no mis-specification has taken place. The effects of getting the distribution wrong will be considered in later parts.

4.2.4 Analysing data using the correct distribution

We present the theory necessary to compute the EFI matrices for G_w and G_b, and give some new results on expectations of order statistics. We include these when they are needed.

n	100	300	500	1000
r	80	240	400	800
$\widehat{\beta}$	2.8373	2.7898	2.7822	2.7750
(st.err.)	(0.2743)	(0.1551)	(0.1174)	(0.0825)
$\widehat{\theta}$	66.0652	66.1864	66.2064	66.2332
(st.err.)	(2.8509)	(1.6351)	(1.2639)	(0.8956)
$\widehat{B}_{w,10}$	29.7497	29.4965	29.4602	29.4231
(st.err.)	(2.5155)	(1.4508)	(1.1139)	(0.7874)
Δ	0.5503	3.0045	5.4573	11.3106
(st.err.)	(1.8942)	(3.6759)	(4.7326)	(6.4684)
Pr (Fit G_w)	0.3925	0.2388	0.1517	0.0634
$\widehat{\tau}$	3.2739	3.1391	3.0892	3.0446
(st.err.)	(0.4135)	(0.2707)	(0.2098)	(0.1514)
$\widehat{\alpha}$	3.6849	4.1201	4.3167	4.4157
(st.err.)	(3.7204)	(3.7840)	(3.1696)	(2.6624)
$\widehat{\phi}$	88.5262	94.7945	97.8293	100.3893
(st.err.)	(29.9935)	(27.9583)	(25.4579)	(21.7425)
$\widehat{B}_{b,10}$	30.5910	30.2008	30.0917	29.9914
(st.err.)	(2.5434)	(1.4974)	(1.1497)	(0.8154)

Table 4.5: Summary statistics for G_w and G_b for varying n with $r = \frac{4n}{5}$. Both distributions are fitted to data that has undergone a type II censoring regime from G_b with parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

n	1000	1000	1000	1000
r	500	600	700	900
$\widehat{\beta}$	2.8960	2.8609	2.8218	2.7136
(st.err.)	(0.1188)	(0.1049)	(0.0933)	(0.0750)
$\widehat{\theta}$	65.0710	65.4556	65.8338	66.6368
(st.err.)	(1.1822)	(1.0340)	(0.9609)	(0.8864)
$\widehat{B}_{w,10}$	29.8853	29.7851	29.6373	29.0672
(st.err.)	(0.8486)	(0.8481)	(0.8177)	(0.7775)
Δ	1.1557	2.8633	5.7853	23.9459
(st.err.)	(2.2855)	(3.3930)	(4.7861)	(9.7901)
Pr (Fit G_w)	0.3690	0.2526	0.1585	0.0117
$\widehat{\tau}$	3.1475	3.1189	3.0826	3.0205
(st.err.)	(0.1777)	(0.1732)	(0.1645)	(0.1441)
$\widehat{\alpha}$	2.4560	3.0699	3.7438	4.6757
(st.err.)	(1.5860)	(1.9146)	(2.5077)	(2.7572)
$\widehat{\phi}$	79.4887	86.5411	93.5683	103.5577
(st.err.)	(17.1498)	(19.0298)	(20.6538)	(20.9846)
$\widehat{B}_{b,10}$	29.9585	30.0115	30.0145	29.9492
(st.err.)	(0.8417)	(0.8502)	(0.8342)	(0.8348)

Table 4.6: Summary statistics for G_w and G_b for varying r and $n = 1000$. Both distributions are fitted to data that has undergone a type II censoring regime from G_b with parameters $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

The Weibull distribution

On examining the structure of the second derivatives from G_w, we see that we require expectations of the form

$$E\left[h\left(Y_{(r:n)}\right)\right]$$

for an arbitrary function h. We use the recursive result linking expectations of order statistics with different sample sizes given by

$$iE\left[h\left(Y_{(i+1:n+1)}\right)\right]+(n-i+1)\,E\left[h\left(Y_{(i:n+1)}\right)\right]=(n+1)\,E\left[h\left(Y_{(i:n)}\right)\right], \tag{4.33}$$

found, for example, in Balakrishnan and Rao (1998a) and David (1981), to write this expectation as

$$r\binom{n}{r}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i)}E\left[h\left(Y_{(1:n-i)}\right)\right], \tag{4.34}$$

where expectations of the first order statistic are usually the most straightforward to calculate; we refer to Appendix A for the proof of (4.34). We extend this result by deriving an expression for the expected value of the sum of the first r order statistics, given by

$$\sum_{i=1}^{r}E\left[h\left(Y_{(i:n)}\right)\right]=\begin{cases}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i-1)(n-i)}E\left[h\left(Y_{(1:n-i)}\right)\right] & \text{for } 1\le r\le n-1\\ nE\left[h\left(Y\right)\right] & \text{for } r=n\end{cases}, \tag{4.35}$$

the proof of which can again be found in Appendix A. We also note that summations such as

$$\sum_{i=1}^{r}E\left[h\left(Y_{(i:n)}\right)\right]+(n-r)\,E\left[h\left(Y_{(r:n)}\right)\right],$$

which appear on numerous occasions when examining second derivatives, can be simplified as follows

$$\begin{aligned}&(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-i-1}\binom{r-1}{i}}{(n-i-1)(n-i)}E\left[h\left(Y_{(1:n-i)}\right)\right]\\&+(n-r)\,r\binom{n}{r}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{n-i}E\left[h\left(Y_{(1:n-i)}\right)\right]\\=\;&(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{n-i-1}E\left[h\left(Y_{(1:n-i)}\right)\right].\end{aligned} \tag{4.36}$$

We use these results to derive expected values of (4.28), (4.29) and (4.30). On examining these second derivatives, we list expectations required, and first consider

$$E\left[S_{f,0}\left(\beta\right)+S_{c,0}\left(\beta\right)\right]=\sum_{i=1}^{r}E\left[Y_{(i:n)}^{\beta}\right]+\left(n-r\right)E\left[Y_{(r:n)}^{\beta}\right].$$

On using (4.36), we write this as

$$\left(n-r+1\right)\left(n-r\right)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{\left(-1\right)^{r-1-i}\binom{r-1}{i}}{n-i-1}E\left[Y_{(1:n-i)}^{\beta}\right],$$

and note that

$$E\left[Y_{(1:n)}^{m}\right]=n^{-\frac{m}{\beta}}\theta^{m}\Gamma\left(\frac{m}{\beta}+1\right), \tag{4.37}$$

since $Y_{(1:n)}$ is from $G_w\left(\beta,\theta n^{-\frac{1}{\beta}}\right)$. Hence, we have

$$E\left[S_{f,0}\left(\beta\right)+S_{c,0}\left(\beta\right)\right]=\left(n-r+1\right)\left(n-r\right)\binom{n}{r-1}\theta^{\beta}\sum_{i=0}^{r-1}\frac{\left(-1\right)^{r-1-i}\binom{r-1}{i}}{\left(n-i-1\right)\left(n-i\right)}.$$

We use a series of results from John, Johnson & Watkins (2003) to simplify the summation in this expected value, and list these below; we refer to Appendix B for the necessary proofs. Let k and m be non-negative integers and a a constant, where $a>m$. We define two forms of summations of reciprocals and their powers. These are given by

$$F_{m,k}\left(a\right)=\sum_{i=0}^{m}\left(a-i\right)^{-k},$$

and

$$A_{m,k}\left(a\right)=\sum_{i=0}^{m}\left(-1\right)^{m-i}\binom{m}{i}\left(a-i\right)^{-k},$$

and note that, when $k=1$,

$$A_{m,1}\left(a\right)=B\left(a-m,m+1\right), \tag{4.38}$$

where B is the usual Beta function given by (1.13). For $k=2$, we have

$$A_{m,2}\left(a\right)=B\left(a-m,m+1\right)F_{m,1}\left(a\right)=A_{m,1}\left(a\right)F_{m,1}\left(a\right). \tag{4.39}$$

We use these results to simplify our expected values, and first use partial fractions to split the denominator of the summation. Hence, we see that

$$\sum_{i=0}^{r-1} \frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i-1)(n-i)} = A_{r-1,1}(n-1) - A_{r-1,1}(n),$$

which we can write as

$$\frac{(n-r-1)!r!}{n!}.$$

Thus, we obtain

$$E\left[S_{f,0}(\beta) + S_{c,0}(\beta)\right] = r\theta^{\beta}.$$

Next, we consider

$$E\left[S_{f,1}(\beta) + S_{c,1}(\beta)\right],$$

and so require an expression for $E\left[Y_{(1:n)}^{m} \ln Y_{(1:n)}\right]$. On differentiating (4.37), with respect to m, we have

$$E\left[Y_{(1:n)}^{m} \ln Y_{(1:n)}\right] = \theta^{m} n^{-\frac{m}{\beta}} \Gamma\left(\frac{m}{\beta}+1\right)\left\{\beta^{-1}\Psi\left(\frac{m}{\beta}+1\right) + \ln\theta - \beta^{-1}\ln n\right\}, \tag{4.40}$$

and, on simplifying, we see that

$$\begin{aligned} E\left[S_{f,1}(\beta) + S_{c,1}(\beta)\right] &= r\theta^{\beta}\left\{\ln\theta + \beta^{-1}(1-\gamma)\right\} - \theta^{\beta}\beta^{-1}(n-r+1)(n-r)\binom{n}{r-1} \\ &\quad \times \sum_{i=0}^{r-1} \frac{(-1)^{r-1-i}\binom{r-1}{i}\ln(n-i)}{(n-i-1)(n-i)}. \end{aligned}$$

Finally, we consider

$$E\left[S_{f,2}(\beta) + S_{c,2}(\beta)\right];$$

we use the fact that

$$E\left[Y_{(1:n)}^{m}\left(\ln Y_{(1:n)}\right)^{2}\right] = \theta^{m} n^{-\frac{m}{\beta}} \Gamma\left(\frac{m}{\beta}+1\right)\left[\begin{array}{c} \left\{\ln\theta + \beta^{-1}\Psi\left(\frac{m}{\beta}+1\right) - \beta^{-1}\ln n\right\}^{2} \\ +\beta^{-2}\Psi'\left(\frac{m}{\beta}+1\right) \end{array}\right]$$

to write this expectation as

$$r\theta^{\beta}\left[\left\{\ln\theta+\beta^{-1}(1-\gamma)\right\}^{2}+\beta^{-2}\Psi'(2)\right]$$
$$-2(n-r+1)(n-r)\binom{n}{r-1}\theta^{\beta}\beta^{-1}\left\{\ln\theta+\beta^{-1}(1-\gamma)\right\}\times\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\ln(n-i)}{(n-i-1)(n-i)}$$
$$+(n-r+1)(n-r)\binom{n}{r-1}\theta^{\beta}\beta^{-2}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\{\ln(n-i)\}^{2}}{(n-i-1)(n-i)}.$$

We now list the elements of the EFI matrix as follows:

$$-E\left[\frac{\partial^2 l_w}{\partial\theta^2}\right]=r\beta^2\theta^{-2},$$

$$\begin{aligned}-E\left[\frac{\partial^2 l_w}{\partial\beta^2}\right] &= r\beta^{-2}\left\{1+(1-\gamma)^2+\Psi'(2)\right\}\\ &\quad+(n-r+1)(n-r)\binom{n}{r-1}\beta^{-2}\\ &\quad\times\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\ln(n-i)\{\ln(n-i)-2(1-\gamma)\}}{(n-i-1)(n-i)},\end{aligned}$$

and

$$\begin{aligned}-E\left[\frac{\partial^2 l_w}{\partial\beta\partial\theta}\right] &= -r\theta^{-1}(1-\gamma)+(n-r+1)(n-r)\binom{n}{r-1}\theta^{-1}\\ &\quad\times\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\ln(n-i)}{(n-i-1)(n-i)}.\end{aligned}$$

We now give a similar analysis for G_b.

The Burr distribution

To compute the EFI matrix of the Burr MLEs, we use expectations listed in the complete scenario, and the fact that $Y_{(1:n)}$ is from $G_b(\tau, n\alpha, \phi)$. We first consider

$$\begin{aligned}-E\left[\frac{\partial^2 l_b}{\partial\tau^2}\right] &= r\tau^{-2}+\alpha\left\{\begin{array}{c}\sum_{i=1}^{r}E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\left\{\ln\left(\frac{Y_{(i:n)}}{\phi}\right)\right\}^{2}}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right]\\ +(n-r)E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\left\{\ln\left(\frac{Y_{(r:n)}}{\phi}\right)\right\}^{2}}{\left\{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right]\end{array}\right\}\\ &\quad-\sum_{i=1}^{r}E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\left\{\ln\left(\frac{Y_{(i:n)}}{\phi}\right)\right\}^{2}}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right].\end{aligned}$$

We replace α with $n\alpha$ in (3.8) to derive an expression for

$$E\left[\frac{\left(\frac{Y_{(1:n)}}{\phi}\right)^{\tau}\left\{\ln\left(\frac{Y_{(1:n)}}{\phi}\right)\right\}^{2}}{\left\{1+\left(\frac{Y_{(1:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right],$$

and use (4.36) to write

$$\sum_{i=1}^{r} E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\left\{\ln\left(\frac{Y_{(i:n)}}{\phi}\right)\right\}^{2}}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right]+(n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\left\{\ln\left(\frac{Y_{(r:n)}}{\phi}\right)\right\}^{2}}{\left\{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right]$$

as

$$\tau^{-2}\alpha^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\{n-i\}}{(n-i-1)\left(n-i+\alpha^{-1}\right)\left(n-i+2\alpha^{-1}\right)}$$
$$\times\left\{\frac{\pi^{2}}{6}+\gamma-2\gamma+2(\gamma-1)\,\Psi\{(n-i)\,\alpha+1\}+\Psi\{(n-i)\,\alpha+1\}^{2}+\Psi'\{(n-i)\,\alpha+1\}\right\}.$$

Now, we use (4.35) to obtain

$$\begin{aligned}&\sum_{i=1}^{r} E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\left\{\ln\left(\frac{Y_{(i:n)}}{\phi}\right)\right\}^{2}}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right]\\ =\;& \tau^{-2}\alpha^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i-1)\left(n-i+\alpha^{-1}\right)\left(n-i+2\alpha^{-1}\right)}\\ &\times\left\{\frac{\pi^{2}}{6}+\gamma-2\gamma+2(\gamma-1)\,\Psi\{(n-i)\,\alpha+1\}+\Psi\{(n-i)\,\alpha+1\}^{2}+\Psi'\{(n-i)\,\alpha+1\}\right\}.\end{aligned}$$

Hence,

$$\begin{aligned}-E\left[\frac{\partial^{2} l_{b}}{\partial\tau^{2}}\right] &= r\tau^{-2}+\tau^{-2}(n-r+1)(r-1)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i-1)\left(n-i+2\alpha^{-1}\right)}\\ &\quad\times\left\{\begin{array}{c}\frac{\pi^{2}}{6}+\gamma^{2}-2\gamma+2(\gamma-1)\,\Psi\{(n-i)\,\alpha+1\}+\Psi\{(n-i)\,\alpha+1\}^{2}\\ +\Psi'\{(n-i)\,\alpha+1\}\end{array}\right\}.\end{aligned}$$

Next, we have

$$-E\left[\frac{\partial^{2} l_{b}}{\partial\alpha^{2}}\right]=r\alpha^{-2}.$$

Now, we derive

$$
\begin{aligned}
-E\left[\frac{\partial^2 l_b}{\partial\phi^2}\right] &= -r\tau\phi^{-2}+\tau\alpha\phi^{-2}\left\{\sum_{i=1}^{r}E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right]+(n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}\right]\right\}\\
&\quad+\tau\phi^{-2}\sum_{i=1}^{r}E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right]\\
&\quad+\tau^2\alpha\phi^{-2}\left\{\sum_{i=1}^{r}E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^2}\right]+(n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\right\}^2}\right]\right\}\\
&\quad+\tau^2\phi^{-2}\sum_{i=1}^{r}E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^2}\right].
\end{aligned}
$$

Using (3.9) from the complete scenario, and (4.36), we see that

$$
\begin{aligned}
&E\left[\sum_{i=1}^{r}\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right]+(n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}\right]\\
&= \frac{n}{\alpha+1}-\frac{n!\Gamma\left(n+\alpha^{-1}-r+1\right)}{(n-r-1)!\,(\alpha+1)\,\Gamma\left(n+\alpha^{-1}+1\right)}.
\end{aligned}
\tag{4.41}
$$

We then use (4.35) to write

$$
\begin{aligned}
&\sum_{i=1}^{r}E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right]\\
&= \frac{n}{\alpha+1}-(n-r)+\frac{n!\alpha\Gamma\left(n+\alpha^{-1}-r+1\right)}{(n-r-1)!\,(\alpha+1)\,\Gamma\left(n+\alpha^{-1}+1\right)}.
\end{aligned}
$$

To derive the remaining expectations, we use (3.10) with α replaced by $n\alpha$, and (4.36) to write

$$
\sum_{i=1}^{r}E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^2}\right]+(n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\right\}^2}\right]
$$

as

$$
\alpha^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{n}{r-1}(n-i)}{(n-i-1)\,(n-i+\alpha^{-1})\,(n-i+2\alpha^{-1})}.
$$

Finally, we use (4.35) to obtain an expression for

$$\sum_{i=1}^{r} E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^{2}}\right];$$

this is given by

$$\alpha^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{n}{r-1}}{(n-i-1)(n-i+\alpha^{-1})(n-i+2\alpha^{-1})}.$$

To simplify this, we write

$$\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{n}{r-1}}{(n-i-1)(n-i+\alpha^{-1})(n-i+2\alpha^{-1})}$$

as

$$\alpha^{2}(r-1)!\left\{\begin{array}{c}\frac{(n-r-1)!}{(\alpha+1)(\alpha+2)(n-1)!}-\frac{\Gamma(n+\alpha^{-1}-r+1)}{(\alpha+1)\Gamma(n+\alpha^{-1}+1)}\\+\frac{\Gamma(n+2\alpha^{-1}-r+1)}{(\alpha+2)\Gamma(n+2\alpha^{-1}+1)}\end{array}\right\},$$

and so the required expectation becomes

$$\frac{n\alpha}{(\alpha+1)(\alpha+2)}-\frac{n!\alpha\Gamma\left(n+\alpha^{-1}-r+1\right)}{(\alpha+1)(n-r-1)!\Gamma\left(n+\alpha^{-1}+1\right)}$$
$$+\frac{n!\alpha\Gamma\left(n+2\alpha^{-1}-r+1\right)}{(\alpha+2)(n-r-1)!\Gamma\left(n+2\alpha^{-1}+1\right)}.$$

Hence, we write $-E\left[\frac{\partial^2 l_b}{\partial\phi^2}\right]$ as

$$\frac{n\tau^{2}\alpha}{\phi^{2}(\alpha+1)(\alpha+2)}-\frac{n!\tau^{2}\alpha\Gamma\left(n+\alpha^{-1}-r+1\right)}{\phi^{2}(\alpha+1)(n-r-1)!\Gamma\left(n+\alpha^{-1}+1\right)}$$
$$+\frac{n!\tau^{2}\alpha\Gamma\left(n+2\alpha^{-1}-r+1\right)}{\phi^{2}(\alpha+2)(n-r-1)!\Gamma\left(n+2\alpha^{-1}+1\right)}$$
$$+\tau^{2}\phi^{-2}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}(n-i)}{(n-i-1)(n-i+\alpha^{-1})(n-i+2\alpha^{-1})}.$$

The next expected value we compute is

$$-E\left[\frac{\partial^{2}l_{b}}{\partial\alpha\partial\tau}\right]=E\left[\sum_{i=1}^{r}\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(i:n)}}{\phi}\right)}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right]+(n-r)E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(r:n)}}{\phi}\right)}{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}\right];$$

we use (3.7) and (4.36) to write

$$E\left[\sum_{i=1}^{r}\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(i:n)}}{\phi}\right)}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right]+(n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(r:n)}}{\phi}\right)}{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}\right]$$

as

$$\begin{aligned}&\alpha^{-1}\tau^{-1}(1-\gamma)(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i-1)(n-i+\alpha^{-1})}\\&-\alpha^{-1}\tau^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Psi\{(n-i)\alpha\}}{(n-i-1)(n-i+\alpha^{-1})}.\end{aligned}$$

Now, using partial fractions, we write

$$\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i-1)(n-i+\alpha^{-1})}=\frac{\alpha(r-1)!}{\alpha+1}\left\{\begin{array}{c}\frac{(n-r-1)!}{(n-1)!}-\\\frac{\Gamma(n+\alpha^{-1}-r+1)}{\Gamma(n+\alpha^{-1}+1)}\end{array}\right\}.$$

Thus,

$$E\left[\sum_{i=1}^{r}\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(i:n)}}{\phi}\right)}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right]+(n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(r:n)}}{\phi}\right)}{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}\right]$$

becomes

$$\begin{aligned}&\frac{n(1-\gamma)}{\tau(\alpha+1)}-\frac{n!(1-\gamma)\Gamma\left(n+\alpha^{-1}-r+1\right)}{\tau(a+1)(n-r-1)!\Gamma\left(n+\alpha^{-1}+1\right)}\\&-\alpha^{-1}\tau^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Psi\{(n-i)\alpha\}}{(n-i-1)(n-i+\alpha^{-1})},\end{aligned}$$

and so we have

$$\begin{aligned}-E\left[\frac{\partial^2 l_w}{\partial\alpha\partial\tau}\right]&=\frac{-n!(1-\gamma)\Gamma\left(n+\alpha^{-1}-r+1\right)}{\tau(\alpha+1)(n-r-1)!\Gamma\left(n+\alpha^{-1}+1\right)}+\frac{n(1-\gamma)}{\tau(\alpha+1)}\\&\quad-\alpha^{-1}\tau^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Psi\{(n-i)\alpha\}}{(n-i-1)(n-i+\alpha^{-1})}.\end{aligned}$$

We now derive

$$\begin{aligned}-E\left[\frac{\partial^2 l_b}{\partial\alpha\partial\phi}\right]&=-\tau\phi^{-1}\left\{E\left[\sum_{i=1}^{r}\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right]+(n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}\right]\right\}\\&=-\tau\phi^{-1}\left\{\frac{n}{\alpha+1}-\frac{n!\Gamma\left(n+\alpha^{-1}-r+1\right)}{(n-r-1)!(\alpha+1)\Gamma\left(n+\alpha^{-1}+1\right)}\right\},\end{aligned}$$

which is obtained using (4.41). Finally, we compute

$$
\begin{aligned}
-E\left[\frac{\partial^2 l_b}{\partial\tau\partial\phi}\right] &= r\phi^{-1} - \alpha\phi^{-1}\left\{\sum_{i=1}^{r} E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right] + (n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}}\right]\right\} \\
&\quad -\phi^{-1}\sum_{i=1}^{r} E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}}\right] \\
&\quad -\tau\alpha\phi^{-1}\left\{\sum_{i=1}^{r} E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(i:n)}}{\phi}\right)}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^2}\right] + (n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(r:n)}}{\phi}\right)}{\left\{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\right\}^2}\right]\right\} \\
&\quad -\tau\phi^{-1}\sum_{i=1}^{r} E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(i:n)}}{\phi}\right)}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^2}\right].
\end{aligned}
$$

We use (3.11) with α replaced by $n\alpha$ to obtain an expression for

$$
E\left[\frac{\left(\frac{Y_{(1:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(1:n)}}{\phi}\right)}{\left\{1+\left(\frac{Y_{(1:n)}}{\phi}\right)^{\tau}\right\}^2}\right],
$$

and (4.36) to write

$$
\sum_{i=1}^{r} E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(i:n)}}{\phi}\right)}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^2}\right] + (n-r)\,E\left[\frac{\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(r:n)}}{\phi}\right)}{\left\{1+\left(\frac{Y_{(r:n)}}{\phi}\right)^{\tau}\right\}^2}\right]
$$

as

$$
\tau^{-1}\alpha^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}(n-i)\left[1-\gamma-\Psi\{(n-i)\alpha+1\}\right]}{(n-i-1)(n-i+\alpha^{-1})(n-i+2\alpha^{-1})}.
$$

Finally, using (4.35), we express

$$
\sum_{i=1}^{r} E\left[\frac{\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\ln\left(\frac{Y_{(i:n)}}{\phi}\right)}{\left\{1+\left(\frac{Y_{(i:n)}}{\phi}\right)^{\tau}\right\}^2}\right]
$$

as

$$\frac{n\alpha(1-\gamma)}{\tau(\alpha+1)(\alpha+2)} - \frac{n!\alpha(1-\gamma)\Gamma(n+\alpha^{-1}-r+1)}{\tau(n-r-1)!(\alpha+1)\Gamma(n+\alpha^{-1}+1)}$$
$$+\frac{n!\alpha(1-\gamma)\Gamma(n+2\alpha^{-1}-r+1)}{\tau(n-r-1)!(\alpha+2)\Gamma(n+2\alpha^{-1}+1)}$$
$$-\tau^{-1}\alpha^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Psi\{(n-i)\alpha+1\}}{(n-i-1)(n-i+\alpha^{-1})(n-i+2\alpha^{-1})}.$$

Thus, $-E\left[\frac{\partial^2 l_b}{\partial\tau\partial\phi}\right]$ becomes

$$\frac{-n\alpha(1-\gamma)}{\phi(\alpha+1)(\alpha+2)} + \frac{n!\alpha(1-\gamma)\Gamma(n+\alpha^{-1}-r+1)}{\phi(n-r-1)!(\alpha+1)\Gamma(n+\alpha^{-1}+1)}$$
$$-\frac{n!\alpha(1-\gamma)\Gamma(n+2\alpha^{-1}-r+1)}{\phi(n-r-1)!(\alpha+2)\Gamma(n+2\alpha^{-1}+1)}$$
$$-\phi^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i-1)(n-i+\alpha^{-1})(n-i+2\alpha^{-1})}$$
$$\times\left[(n-i)(1-\gamma)-\Psi\{(n-i)\alpha+1\}\left(n-i+\alpha^{-1}\right)\right].$$

We now have all the elements that make up the variance covariance matrix for the MLEs of the Burr distribution. We check the results in later sections when we compare simulated values with their theoretical counterparts. These theoretical values are computed relatively quickly using Mathematica, even for large sample sizes. We continue by examining the effects of mis-specifying G_w to type II censored data from G_b.

4.2.5 Analysing data using the incorrect distribution

We compute theoretical counterparts to the MLEs from G_w; this will allow comparison with Tables 4.5 and 4.6. Using (4.23), we write the entropy function as

$$\begin{aligned} E_b &= r\ln\beta + (\beta-1)E\left[\sum_{i=1}^{r}\ln Y_{(i:n)}\right] - r\beta\ln\theta \\ &\quad -\theta^{-\beta}\left\{E\left[\sum_{i=1}^{r}Y_{(i:n)}^{\beta}\right] + (n-r)E\left[Y_{(r:n)}^{\beta}\right]\right\}. \end{aligned}$$

We use (3.31) with α replaced by $n\alpha$ to write down an expression for $E\left[\ln Y_{(1:n)}\right]$, and (4.35) to write

$$\begin{aligned} E\left[\sum_{i=1}^{r}\ln Y_{(i:n)}\right] &= r\left\{\ln\phi+\tau^{-1}\Psi(1)\right\} - \tau^{-1}(n-r+1)(n-r)\binom{n}{r-1} \\ &\quad \times\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Psi\{(n-i)\alpha\}}{(n-i-1)(n-i)}. \end{aligned} \tag{4.42}$$

To compute the remaining expected value, we use (3.26) from the complete scenario, and (4.36). Thus

$$E\left[\sum_{i=1}^{r} Y_{(i:n)}^{m}\right] + (n-r)\,E\left[Y_{(r:n)}^{m}\right],$$

simplifies to

$$\phi^{m}\Gamma\left(\frac{m}{\tau}+1\right)(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Gamma\left\{(n-i)\alpha-\frac{m}{\tau}\right\}}{(n-i-1)\Gamma\left\{(n-i)\alpha\right\}},$$

and we have

$$\begin{aligned} E_b &= r\ln\beta - r\beta\ln\theta + r\beta\ln\phi - r\ln\phi + r(\beta-1)\tau^{-1}\Psi(1) \\ &\quad -(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{n-i-1} \\ &\quad \times\left[\frac{(\beta-1)\Psi\{(n-i)\alpha\}}{\tau(n-i)} + \frac{\phi^{\beta}\Gamma\left(\frac{\beta}{\tau}+1\right)\Gamma\left\{(n-i)\alpha-\frac{\beta}{\tau}\right\}}{\theta^{\beta}\Gamma\{(n-i)\alpha\}}\right]. \end{aligned} \tag{4.43}$$

We can either maximise this entropy function, or compute the roots of the entropy score functions. These are given by

$$\begin{aligned} \frac{\partial E_b}{\partial\theta} &= -r\beta\theta^{-1} + \beta\theta^{-\beta-1}\phi^{\beta}\Gamma\left(\frac{\beta}{\tau}+1\right)(n-r+1)(n-r)\binom{n}{r-1} \\ &\quad \times\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Gamma\left\{(n-i)\alpha-\frac{\beta}{\tau}\right\}}{(n-i-1)\Gamma(\{n-i\}\alpha)}, \end{aligned} \tag{4.44}$$

and

$$\begin{aligned} \frac{\partial E_b}{\partial\beta} &= r\beta^{-1} - r\ln\theta + r\ln\phi + r\tau^{-1}\Psi(1) - \\ &\quad (n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{n-i-1} \\ &\quad \times\left[\begin{array}{c} \frac{\Psi\{(n-i)\alpha\}}{\tau(n-i)} + \\ \frac{\phi^{\beta}\Gamma\left(\frac{\beta}{\tau}+1\right)\Gamma\left\{(n-i)\alpha-\frac{\beta}{\tau}\right\}}{\theta^{\beta}\Gamma\{(n-i)\alpha\}}\left\{\begin{array}{c}\ln\phi-\ln\theta+\tau^{-1}\Psi\left(\frac{\beta}{\tau}+1\right) \\ -\tau^{-1}\Psi\left\{(n-i)\alpha-\frac{\beta}{\tau}\right\}\end{array}\right\}\end{array}\right]. \end{aligned}$$

We can equate (4.44) to zero to get

$$\theta = \left[\begin{array}{c} r^{-1}\phi^{\beta}\Gamma\left(\frac{\beta}{\tau}+1\right)(n-r+1)(n-r)\binom{n}{r-1} \\ \times\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Gamma\left\{(n-i)\alpha-\frac{\beta}{\tau}\right\}}{(n-i-1)\Gamma\{(n-i)\alpha\}}\end{array}\right]^{\frac{1}{\beta}},$$

and insert this into (4.43) to obtain the profile entropy and profile entropy score function, here given by

$$r\ln\beta - r\beta^{-1}\ln\phi - r\ln\Gamma\left(\frac{\beta}{\tau}+1\right) - r\ln\left[(n-r+1)(n-r)\binom{n}{r-1}\right]$$
$$+r(\beta-1)\left\{\ln\phi+\tau^{-1}\Psi(1)\right\} - r\ln\left[\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Gamma\left\{(n-i)\alpha-\frac{\beta}{\tau}\right\}}{(n-i-1)\Gamma\{(n-i)\alpha\}}\right]$$
$$-\tau^{-1}(\beta-1)(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Psi\{(n-i)\alpha\}}{(n-i-1)(n-i)},$$

and

$$r\beta^{-1} + r\beta^{-2}\ln\phi - r\tau^{-1}\Psi\left(\frac{\beta}{\tau}+1\right) + r\left\{\ln\phi+\tau^{-1}\Psi(1)\right\}$$
$$\tau^{-1}(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Psi\{(n-i)\alpha\}}{(n-i-1)(n-i)}$$
$$+\frac{r\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Psi\left\{(n-i)\alpha-\frac{\beta}{\tau}\right\}\Gamma\left\{(n-i)\alpha-\frac{\beta}{\tau}\right\}}{(n-i-1)\Gamma\{(n-i)\alpha\}}}{\tau\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}\Gamma\left\{(n-i)\alpha-\frac{\beta}{\tau}\right\}}{(n-i)\Gamma\{(n-i)\alpha\}}},$$

respectively. We try and maximise the profile entropy function, or equivalently, find the root of the profile score, using Mathematica, for a particular set of Burr parameters. Further investigations in Mathematica indicate that commands currently used to locate roots, or find the maximum of functions, fail to locate the maximum of the profile entropy function, or to locate the root of the profile entropy score. Thus, we are currently limited to constructing a grid of values of the profile entropy function for varying β. We do this using our usual set of Burr parameters, given by

$$\tau = 3,\ \alpha = 4,\ \phi = 100,$$

and a sample size of $n = 1000$. We also censor values that exceed the $r = 800^{th}$ data point. In Mathematica, we construct a table of values for the entropy function, for $2 < \beta < 3.5$, a range based on the sample average MLE of β for this set of parameters. A plot of these values is shown in Figure 4.3. A clear maximum exists at around $\beta = 2.77$, so we refine the search in Mathematica, and calculate values for the profile entropy function around a smaller interval containing this point. We see that the maximum still occurs when $\beta = 2.770$ (to three decimal places) which corresponds to a value of 66.217 for θ. Comparing these values with the simulations run in the previous section, we see that the MLEs for β and θ are close to their theoretical counterparts, considering that one fifth of the data has been censored. Unfortunately, we cannot easily improve on this technique for locating the maximum of the entropy function, and the problem seems to arise from evaluating summations in E_b^*,

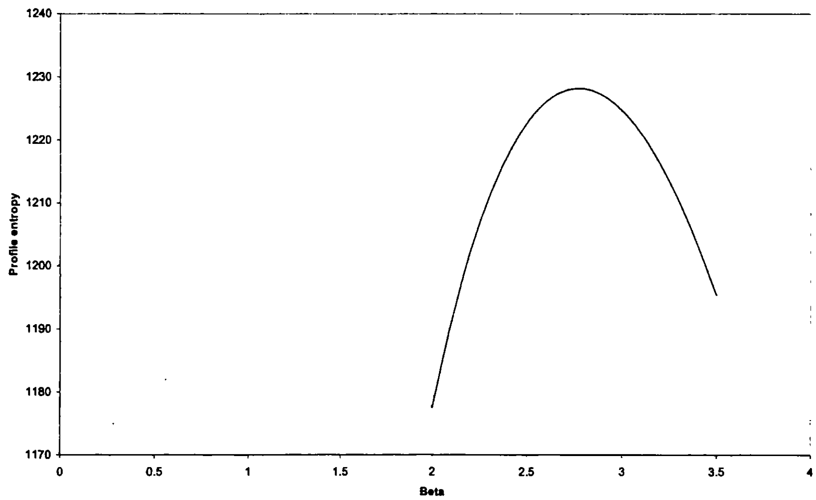


Figure 4.3: E_b^* vs β for data simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$, and subject to type II censoring with $n = 1000$ and $r = 800$.

especially for large sample sizes, where the computational demands on Mathematica cause problems. We must also choose an appropriate range for β when constructing the grid search. For a large enough sample size, this task can be simplified on exploiting the fact that we expect the MLE for this parameter to be close to its theoretical counterpart. Consequently, we may choose the range of possible values of β close to the sample average of the MLE. For smaller sample sizes, we do not expect such good agreement between observed and expected counterparts, and as a result, have to consider increasing the range of values that we take for β. However, for small sample sizes, the Mathematica program is reasonably quick at computing the values of the entropy function, so increasing the range does not really lengthen the running time.

Now that we are able to compute these theoretical counterparts, we use these in the computation of the variance covariance matrix of the mis-specified Weibull MLEs. We consider this below.

The variance structure of the mis-specified MLEs

We consider the variance covariance matrix of the Weibull MLEs, after this distribution has been fitted to data from a Burr model that has undergone a type II censoring regime. We refer to our previous work on complete data to state that the asymptotic distribution of the mis-specified MLEs from the Weibull distribution will be Normally distributed with mean vector $(\beta_0, \theta_0)'$, and variance covariance matrix given by (3.18). Thus, we require expected values of second derivatives and variances of score functions from the Weibull distribution, when the data is simulated from a Burr. From our work on obtaining the variance covariance matrix of the MLEs from G_b, we have seen that deriving the matrix of second derivatives is relatively straightforward, and will be so even when we have mis-specified the distribution. However, we encounter problems when evaluating the variance covariance matrix of the score functions from the Weibull distribution. For example, if we consider $Var\left(\frac{\partial l_w}{\partial \theta}\right)$, then using (4.24), we see that this is equal to

$$\begin{aligned} & Var\left\{-r\beta\theta^{-1} + \beta\theta^{-\beta-1}\left(\sum_{i=1}^{r} Y_{(i:n)}^{\beta} + (n-r)\, Y_{(r:n)}^{\beta}\right)\right\} \\ = \; & \beta^2\theta^{-2\beta-2} Var\left(\sum_{i=1}^{r} Y_{(i:n)}^{\beta} + (n-r)\, Y_{(r:n)}^{\beta}\right). \end{aligned}$$

We split this variance into

$$E\left[\left(\sum_{i=1}^{r} Y_{(i:n)}^{\beta} + (n-r)\, Y_{(r:n)}^{\beta}\right)^2\right] - E\left[\sum_{i=1}^{r} Y_{(i:n)}^{\beta} + (n-r)\, Y_{(r:n)}^{\beta}\right]^2,$$

and observe that the second term can be evaluated relatively easily using (4.36). The term that causes the problem is

$$\begin{aligned} E\left[\left(\sum_{i=1}^{r} Y_{(i:n)}^{\beta} + (n-r)\, Y_{(r:n)}^{\beta}\right)^2\right] = \; & \sum_{i=1}^{r} E\left[Y_{(i:n)}^{2\beta}\right] + (n-r)^2 E\left[Y_{(r:n)}^{2\beta}\right] \\ & +2\,(n-r)\sum_{i=1}^{r} E\left[Y_{(i:n)}^{\beta} Y_{(r:n)}^{\beta}\right] + 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r} E\left[Y_{(i:n)}^{\beta} Y_{(j:n)}^{\beta}\right], \end{aligned}$$

and in particular, taking expectations of products of order statistics. We extend our results on expectations of single order statistics, to allow for the product of two order statistics, and to do so, require the recursive relationship between expectations of joint order statistics. We use Balakrishnan and Rao (1998a) to write

$$\begin{aligned} nE\left[Y_{(i-1:n-1)}^{k} Y_{(j-1:n-1)}^{l}\right] = \; & (i-1)\, E\left[Y_{(i:n)}^{k} Y_{(j:n)}^{l}\right] + (j-i)\, E\left[Y_{(i-1:n)}^{k} Y_{(j:n)}^{l}\right] \\ & + (n-j+1)\, E\left[Y_{(i-1:n)}^{k} Y_{(j-1:n)}^{l}\right], \end{aligned} \tag{4.45}$$

and use this to state that

$$\begin{aligned} E\left[Y_{(i:n)}^{k} Y_{(j:n)}^{l}\right] = \; & \frac{n!}{(j-i-1)!}\sum_{s=1}^{i}\sum_{q=0}^{i-s} \frac{(-1)^{s+q-1}\,(n+q-j)!\,(s+j-i-2)!}{q!\,(n-j)!\,(i-q-s)!\,(s-1)!\,(n+q+s-i)!} \\ & \times E\left[Y_{(1:n-i+s+q)}^{k} Y_{(j-i+s:n-i+s+q)}^{l}\right], \end{aligned} \tag{4.46}$$

where the proof can be found in Appendix C. Obtaining expectations of joint order statistics simplifies considerably when considering specific distribution functions. We outline our approach for the Weibull and Burr distributions, and refer to Balakrishnan and Rao (1998b) for equivalent details on the Negative Exponential distribution. Dyer and Whisenand (1973) derive joint expectations when the underlying distribution has a Rayleigh model, and Balasooriya and Hapuarachchi (1992) compute these expectations for the Gamma distribution.

Joint expectations for the Weibull distribution We start by deriving a useful result with direct applications to the problem under consideration; the lemma concerns a particular integral that appears in the product moment of the first order statistic with the j^{th} (for $j \geq 2$) for the Weibull distribution. We define, for arbitrary positive β, p, q, k and l,

$$I_{kl}^{pq}(\beta) = \int_{t=0}^{\infty}\int_{s=0}^{t} t^{p+\beta-1}e^{-kt^{\beta}}s^{q+\beta-1}e^{-ls^{\beta}}dsdt,$$

and now prove the following.

Lemma 1 *We have*

$$I_{kl}^{pq}(\beta) = \frac{\theta^{p+q+2\beta}\Gamma\left(\frac{p+q}{\beta}+2\right)}{\beta^2 k^{2+(p+q)/\beta}\left(\frac{q}{\beta}+1\right)} \times F_{2,1}\left(\frac{q}{\beta}+1, \frac{p+q}{\beta}+2; \frac{q}{\beta}+2; -\frac{l}{k}\right).$$

Proof. From its definition, we may write

$$I_{kl}^{pq}(\beta) = \int_{t=0}^{\infty} t^{p+\beta-1}e^{-k\left(\frac{t}{\theta}\right)^{\beta}}\left[\int_{s=0}^{t} s^{q+\beta-1}e^{-l\left(\frac{s}{\theta}\right)^{\beta}}ds\right]dt,$$

in which

$$\int_{s=0}^{t} s^{q+\beta-1}e^{-l\left(\frac{s}{\theta}\right)^{\beta}}ds = \frac{\theta^{q+\beta}}{\beta l^{1+q/\beta}}\int_{u=0}^{l\left(\frac{t}{\theta}\right)^{\beta}} u^{q/\beta}e^{-u}du = \frac{\theta^{q+\beta}}{\beta l^{1+q/\beta}}\Gamma\left(l\left(\frac{t}{\theta}\right)^{\beta}, \frac{q}{\beta}+1\right),$$

obtained by writing $u = l\left(\frac{s}{\theta}\right)^{\beta}$, so $0 \leq s \leq t \Leftrightarrow 0 \leq u \leq l\left(\frac{t}{\theta}\right)^{\beta}$, and $s = \theta\left(\frac{u}{l}\right)^{1/\beta}$ so $ds = \theta\beta^{-1}l^{-1/\beta}u^{1/\beta-1}du$. Thus, we have

$$I_{kl}^{pq}(\beta) = \frac{\theta^{q+\beta}}{\beta l^{1+q/\beta}}\int_{t=0}^{\infty} t^{p+\beta-1}e^{-k\left(\frac{t}{\theta}\right)^{\beta}}\Gamma\left(l\left(\frac{t}{\theta}\right)^{\beta}, \frac{q}{\beta}+1\right)dt$$

Using (1.10), we now write

$$\Gamma\left(l\left(\frac{t}{\theta}\right)^{\beta}, \frac{q}{\beta}+1\right) = l^{q/\beta+1}\left(\frac{t}{\theta}\right)^{q+\beta}\sum_{m=0}^{\infty}\frac{(-l)^m\left(\frac{t}{\theta}\right)^{m\beta}}{m!\left(\frac{q}{\beta}+1+m\right)},$$

so that

$$
\begin{aligned}
I_{kl}^{pq}(\beta) &= \frac{1}{\beta}\int_{t=0}^{\infty} t^{p+q+2\beta-1} e^{-k\left(\frac{t}{\theta}\right)^{\beta}} \sum_{m=0}^{\infty} \frac{(-l)^m \left(\frac{t}{\theta}\right)^{m\beta}}{m!\left(\frac{q}{\beta}+1+m\right)} dt \\
&= \frac{1}{\beta}\int_{t=0}^{\infty} \left[\sum_{m=0}^{\infty} \frac{(-l)^m\, t^{p+q+(m+2)\beta-1} e^{-k\left(\frac{t}{\theta}\right)^{\beta}}}{m!\left(\frac{q}{\beta}+1+m\right)}\right] dt \\
&= \frac{1}{\beta}\sum_{m=0}^{\infty}\left[\int_{t=0}^{\infty} \frac{(-l)^m\, t^{p+q+(m+2)\beta-1} e^{-kt^{\beta}}}{\theta^{m\beta} m!\left(\frac{q}{\beta}+1+m\right)} dt\right],
\end{aligned}
$$

on reversing the order of integration and summation. We thus have

$$
\begin{aligned}
I_{kl}^{pq}(\beta) &= \frac{1}{\beta}\sum_{m=0}^{\infty}\left[\frac{(-l)^m}{\theta^{m\beta} m!\left(\frac{q}{\beta}+1+m\right)}\int_{t=0}^{\infty} t^{p+q+(m+2)\beta-1} e^{-k\left(\frac{t}{\theta}\right)^{\beta}} dt\right] \\
&= \frac{1}{\beta}\sum_{m=0}^{\infty}\left[\frac{(-l)^m\, \theta^{p+q+2\beta}}{m!\left(\frac{q}{\beta}+1+m\right)} \times \frac{\Gamma\left(\frac{p+q+(m+1)\beta}{\beta}+1\right)}{\beta k^{1+\{p+q+(m+1)\beta\}/\beta}}\right].
\end{aligned}
$$

Therefore,

$$
I_{kl}^{pq}(\beta) = \frac{\theta^{p+q+2\beta}}{\beta^2 k^{2+(p+q)/\beta}} \sum_{m=0}^{\infty}\left[\frac{\left(-\frac{l}{k}\right)^m \Gamma\left(\frac{p+q}{\beta}+m+2\right)}{m!\left(\frac{q}{\beta}+1+m\right)}\right]
$$

We now introduce a hypergeometric function, writing the summation as

$$
\begin{aligned}
&\sum_{m=0}^{\infty}\left[\frac{\Gamma\left(\frac{q}{\beta}+1+m\right)\Gamma\left(\frac{p+q}{\beta}+m+2\right)}{\Gamma\left(\frac{q}{\beta}+2+m\right)} \times \frac{\left(-\frac{l}{k}\right)^m}{m!}\right] \\
&= \frac{\Gamma\left(\frac{q}{\beta}+1\right)\Gamma\left(\frac{p+q}{\beta}+2\right)}{\Gamma\left(\frac{q}{\beta}+2\right)} \times F_{2,1}\left(\frac{q}{\beta}+1, \frac{p+q}{\beta}+2; \frac{q}{\beta}+2; -\frac{l}{k}\right) \\
&= \frac{\Gamma\left(\frac{p+q}{\beta}+2\right)}{\frac{q}{\beta}+1} \times F_{2,1}\left(\frac{q}{\beta}+1, \frac{p+q}{\beta}+2; \frac{q}{\beta}+2; -\frac{l}{k}\right)
\end{aligned}
$$

Hence we have

$$
I_{kl}^{pq}(\beta) = \frac{\theta^{p+q+2\beta}\Gamma\left(\frac{p+q}{\beta}+2\right)}{\beta^2 k^{2+(p+q)/\beta}\left(\frac{q}{\beta}+1\right)} \times F_{2,1}\left(\frac{q}{\beta}+1, \frac{p+q}{\beta}+2; \frac{q}{\beta}+2; -\frac{l}{k}\right),
$$

as required. ■

We now present a general result on the expectations of products of arbitrary powers of $Y_{(i:n)}$ and of $Y_{(j:n)}$. We note the possibility of using our recursive result, given by (4.46), to obtain this joint expectation, where the lemma above can be used to compute the expected value of $Y_{(1:n)}$ with $Y_{(j:n)}$. However, this joint expectation can be computed directly without having to use the recursive form. This approach also results in a simplified expression containing one less summation. The joint pdf of $Y_{(i:n)}$ and $Y_{(j:n)}$ is

$$\beta^2\theta^{-2\beta}c_{i,j:n}s^{\beta-1}t^{\beta-1}\left\{1-e^{-\left(\frac{s}{\theta}\right)^\beta}\right\}^{i-1}\left\{e^{-\left(\frac{s}{\theta}\right)^\beta}-e^{-\left(\frac{t}{\theta}\right)^\beta}\right\}^{j-i-1}e^{-(n-j+1)\left(\frac{t}{\theta}\right)^\beta}e^{-\left(\frac{s}{\theta}\right)^\beta},$$

for $0\leq s<t<\infty$, where

$$c_{i,j:n}=\frac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!}.$$

The expectation $E\left[Y_{(i:n)}^kY_{(j:n)}^l\right]$ thus takes the form

$$\beta^2\theta^{-2\beta}c_{i,j:n}\int_{t=0}^{\infty}\int_{s=0}^{t}s^{k+\beta-1}t^{l+\beta-1}\left\{1-e^{-\left(\frac{s}{\theta}\right)^\beta}\right\}^{i-1}\left\{e^{-\left(\frac{s}{\theta}\right)^\beta}-e^{-\left(\frac{t}{\theta}\right)^\beta}\right\}^{j-i-1}$$
$$\times e^{-(n-j+1)\left(\frac{t}{\theta}\right)^\beta}e^{-\left(\frac{s}{\theta}\right)^\beta}dsdt.$$

We now expand both brackets inside the integral, writing

$$\left\{1-e^{-\left(\frac{s}{\theta}\right)^\beta}\right\}^{i-1}=\sum_{q=0}^{i-1}\binom{i-1}{q}(-1)^{i-1-q}\,e^{-(i-1-q)\left(\frac{s}{\theta}\right)^\beta},$$

and

$$\left\{e^{-\left(\frac{s}{\theta}\right)^\beta}-e^{-\left(\frac{t}{\theta}\right)^\beta}\right\}^{j-i-1}=\sum_{p=0}^{j-i-1}\binom{j-i-1}{p}(-1)^{j-i-1-p}\,e^{-p\left(\frac{s}{\theta}\right)^\beta}e^{-(j-i-1-p)\left(\frac{t}{\theta}\right)^\beta},$$

so that the expectation takes the form

$$\beta^2\theta^{-2\beta}c_{i,j:n}\int_{t=0}^{\infty}\int_{s=0}^{t}s^{k+\beta-1}t^{l+\beta-1}\left[\sum_{q=0}^{i-1}\binom{i-1}{q}(-1)^{i-1-q}\,e^{-(i-1-q)\left(\frac{s}{\theta}\right)^\beta}\right]$$
$$\times\left[\sum_{p=0}^{j-i-1}\binom{j-i-1}{p}(-1)^{j-i-1-p}\,e^{-p\left(\frac{s}{\theta}\right)^\beta}e^{-(j-i-1-p)\left(\frac{t}{\theta}\right)^\beta}\right]e^{-(n-j+1)\left(\frac{t}{\theta}\right)^\beta}e^{-\left(\frac{s}{\theta}\right)^\beta}dsdt.$$

This reduces to

$$\beta^2\theta^{-2\beta}c_{i,j:n}\sum_{q=0}^{i-1}\sum_{p=0}^{j-i-1}\binom{i-1}{q}\binom{j-i-1}{p}(-1)^{j-p-q}$$
$$\times\int_{t=0}^{\infty}\int_{s=0}^{t}s^{k+\beta-1}t^{l+\beta-1}e^{-(i+p-q)\left(\frac{s}{\theta}\right)^{\beta}}e^{-(n-i-p)\left(\frac{t}{\theta}\right)^{\beta}}dsdt$$
$$=\beta^2\theta^{-2\beta}c_{i,j:n}\sum_{q=0}^{i-1}\sum_{p=0}^{j-i-1}\binom{i-1}{q}\binom{j-i-1}{p}(-1)^{j-p-q}I^{l,k}_{n-i-p,i+p-q}(\beta).$$

Using the Lemma, and simplifying the Binomial coefficients, we see that the expectation can be written as

$$\frac{\theta^{l+k}c_{i,j:n}\Gamma\left(\frac{l}{\beta}+\frac{k}{\beta}+2\right)}{\left(\frac{k}{\beta}+1\right)}\sum_{q=0}^{i-1}\sum_{p=0}^{j-i-1}\left[\begin{array}{c}\frac{(-1)^{j-p-q}\binom{i-1}{q}\binom{j-i-1}{p}}{(n-i-p)^{\frac{l}{\beta}+\frac{k}{\beta}+2}}\\ \times F_{2,1}\left(\frac{k}{\beta}+1,\frac{k}{\beta}+\frac{l}{\beta}+2;\frac{k}{\beta}+2;\frac{-(i+p-q)}{n-i-p}\right)\end{array}\right].$$

We note, but do not here exploit, some scope for simplification, through symmetry in the binomial coefficients, the alternating signs of terms in the summation, and the structure in the arguments of the hypergeometric function, particularly in the important case $k=l=1$.

Joint expectations for the Burr distribution Unlike G_w, we cannot compute a direct expression for $E\left[Y^k_{(i:n)}Y^l_{(j:n)}\right]$ when the underlying distribution is G_b. As a result, we use (4.46), and note that we require the expected value of the first order statistic with the j^{th}. Thus, we require the joint pdf between such order statistics. This is given by

$$\begin{aligned}f_{Y_{(1:n)},Y_{(j:n)}}(s,t) &= n(n-1)\binom{n-2}{j-2}\left[\left\{1+\left(\frac{s}{\phi}\right)^{\tau}\right\}^{-\alpha}-\left\{1+\left(\frac{t}{\phi}\right)^{\tau}\right\}^{-\alpha}\right]^{j-2}\\ &\quad\left\{1+\left(\frac{t}{\phi}\right)^{\tau}\right\}^{-\alpha(n-j)}\frac{\alpha\tau s^{\tau-1}}{\phi^{\tau}}\left\{1+\left(\frac{s}{\phi}\right)^{\tau}\right\}^{-\alpha-1}\frac{\alpha\tau t^{\tau-1}}{\phi^{\tau}}\left\{1+\left(\frac{t}{\phi}\right)^{\tau}\right\}^{-\alpha-1}.\end{aligned}$$

Using this, we have

$$\begin{aligned}E\left[Y^k_{(1:n)}Y^l_{(j:n)}\right] &= n(n-1)\binom{n-2}{j-2}\int_0^{\infty}\frac{\alpha\tau t^{l+\tau-1}}{\phi^{\tau}}\left\{1+\left(\frac{t}{\phi}\right)^{\tau}\right\}^{-\alpha(n-j+1)-1}dt\\ &\quad\int_0^t\frac{\alpha\tau s^{k+\tau-1}}{\phi^{\tau}}\left\{1+\left(\frac{s}{\phi}\right)^{\tau}\right\}^{-\alpha-1}\left[\left\{1+\left(\frac{s}{\phi}\right)^{\tau}\right\}^{-\alpha}-\left\{1+\left(\frac{t}{\phi}\right)^{\tau}\right\}^{-\alpha}\right]^{j-2}ds\end{aligned}$$

Just as in the case of the Weibull distribution, we use the Binomial Theorem to write

$$\left[\left\{1+\left(\frac{s}{\phi}\right)^{\tau}\right\}^{-\alpha}-\left\{1+\left(\frac{t}{\phi}\right)^{\tau}\right\}^{-\alpha}\right]^{j-2}$$

as

$$\sum_{p=0}^{j-2}(-1)^{j-2-p}\binom{j-2}{p}\left\{1+\left(\frac{s}{\phi}\right)^{\tau}\right\}^{-\alpha p}\left\{1+\left(\frac{t}{\phi}\right)^{\tau}\right\}^{-\alpha(j-2-p)},$$

so $E\left[Y_{(1:n)}^{k}Y_{(j:n)}^{l}\right]$ becomes

$$n\,(n-1)\binom{n-2}{j-2}\sum_{p=0}^{j-2}(-1)^{j-2-p}\binom{j-2}{p}\int_0^{\infty}\frac{\alpha\tau t^{l+\tau-1}}{\phi^{\tau}}\left\{1+\left(\frac{t}{\phi}\right)^{\tau}\right\}^{-\alpha(n-p-1)-1}dt$$
$$\int_0^{t}\frac{\alpha\tau s^{k+\tau-1}}{\phi^{\tau}}\left\{1+\left(\frac{s}{\phi}\right)^{\tau}\right\}^{-\alpha(1+p)-1}ds.$$

We now set

$$u=\left(\frac{s}{\phi}\right)^{\tau},$$

so that

$$\int_0^{t}\frac{\alpha\tau s^{k+\tau-1}}{\phi^{\tau}}\left\{1+\left(\frac{s}{\phi}\right)^{\tau}\right\}^{-\alpha(1+p)-1}ds=\int_0^{\left(\frac{t}{\phi}\right)^{\tau}}\alpha\phi^{k}u^{\frac{k}{\tau}}\left(1+u\right)^{-\alpha(1+p)-1}du,$$

and if we now put

$$z=\frac{u}{1+u},$$

then this integral becomes

$$\int_0^{\frac{\left(\frac{t}{\phi}\right)^{\tau}}{1+\left(\frac{t}{\phi}\right)^{\tau}}}\alpha\phi^{k}z^{\frac{k}{\tau}}\left(1-z\right)^{\alpha(1+p)-\frac{k}{\tau}-1}dz=\alpha\phi^{k}B_{\frac{\left(\frac{t}{\phi}\right)^{\tau}}{1+\left(\frac{t}{\phi}\right)^{\tau}}}\left(\frac{k}{\tau}+1,\alpha\left\{1+p\right\}-\frac{k}{\tau}\right).$$

If we also set $u=\left(\frac{t}{\phi}\right)^{\tau}$, then we have

$$\begin{aligned}E\left[Y_{(1:n)}^{k}Y_{(j:n)}^{l}\right] &= n\,(n-1)\binom{n-2}{j-2}\sum_{p=0}^{j-2}(-1)^{j-2-p}\binom{j-2}{p}\phi^{k+l}\alpha^{2}\\ &\int_0^{\infty}u^{\frac{l}{\tau}}\left(1+u\right)^{-\alpha(n-p-1)-1}B_{\frac{u}{1+u}}\left(\frac{k}{\tau}+1,\alpha\left\{1+p\right\}-\frac{k}{\tau}\right).\end{aligned}$$

In order to evaluate this integral, we rewrite the incomplete Beta function in terms of a hypergeometric function, given by (1.16). Therefore

$$B_{\frac{u}{1+u}}\left(\frac{k}{\tau}+1,\alpha\{1+p\}-\frac{k}{\tau}\right) = \frac{1}{\Gamma\left(\frac{k}{\tau}-\alpha\{1+p\}+1\right)} \times\sum_{m=0}^{\infty}\frac{\Gamma\left(\frac{k}{\tau}-\alpha\{1+p\}+1+m\right)}{\left(\frac{k}{\tau}+1+m\right)m!}\left(\frac{u}{1+u}\right)^{m+\frac{k}{\tau}+1},$$

and we have

$$E\left[Y_{(1:n)}^{k}Y_{(j:n)}^{l}\right] = n(n-1)\binom{n-2}{j-2}\sum_{p=0}^{j-2}\sum_{m=0}^{\infty}\frac{(-1)^{j-2-p}\binom{j-2}{p}\phi^{k+l}\alpha^{2}\Gamma\left(\frac{k}{\tau}-\alpha\{1+p\}+1+m\right)}{m!\left(\frac{k}{\tau}+1+m\right)\Gamma\left(\frac{k}{\tau}-\alpha\{1+p\}+1\right)} \int_{0}^{\infty}u^{\frac{l}{\tau}+\frac{k}{\tau}+m+1}(1+u)^{-\alpha(n-p-1)-m-\frac{k}{\tau}-2}\,du.$$

We now see that

$$\begin{aligned}\int_{0}^{\infty}u^{\frac{l}{\tau}+\frac{k}{\tau}+m+1}(1+u)^{-\alpha(n-p-1)-m-\frac{k}{\tau}-2}\,du &= \int_{0}^{\infty}\frac{u^{\frac{l}{\tau}+\frac{k}{\tau}+m+1}}{(1+u)^{\alpha(n-p-1)+m+\frac{k}{\tau}+2}}du\\ &= \int_{0}^{\infty}\frac{u^{\frac{l}{\tau}+\frac{k}{\tau}+m+2-1}}{(1+u)^{\frac{l}{\tau}+\frac{k}{\tau}+m+2+\alpha(n-p-1)-\frac{l}{\tau}}}du\\ &= B\left(\frac{l}{\tau}+\frac{k}{\tau}+m+2,\alpha\{n-p-1\}-\frac{l}{\tau}\right)\\ &= \frac{\Gamma\left(\frac{l}{\tau}+\frac{k}{\tau}+m+2\right)\Gamma\left(\alpha\{n-p-1\}-\frac{l}{\tau}\right)}{\Gamma\left(\frac{k}{\tau}+m+2+\alpha\{n-p-1\}\right)}.\end{aligned}$$

Hence, $E\left[Y_{(1:n)}^{k}Y_{(j:n)}^{l}\right]$ becomes

$$n(n-1)\binom{n-2}{j-2}\phi^{k+l}\alpha^{2}\sum_{p=0}^{j-2}\frac{(-1)^{j-2-p}\binom{j-2}{p}\Gamma\left(\alpha\{n-p-1\}-\frac{l}{\tau}\right)}{\Gamma\left(\frac{k}{\tau}-\alpha\{1+p\}+1\right)} \sum_{m=0}^{\infty}\frac{\Gamma\left(\frac{k}{\tau}-\alpha\{1+p\}+1+m\right)\Gamma\left(\frac{l}{\tau}+\frac{k}{\tau}+2+m\right)}{m!\left(\frac{k}{\tau}+1+m\right)\Gamma\left(\frac{k}{\tau}+\alpha\{n-p-1\}+2+m\right)},$$

and finally, using the definition for a hypergeometric function, we see that

$$\sum_{m=0}^{\infty}\frac{\Gamma\left(\frac{k}{\tau}-\alpha\{1+p\}+1+m\right)\Gamma\left(\frac{l}{\tau}+\frac{k}{\tau}+2+m\right)}{m!\left(\frac{k}{\tau}+1+m\right)\Gamma\left(\frac{k}{\tau}+\alpha\{n-p-1\}+2+m\right)},$$

simplifies to

$$\sum_{m=0}^{\infty} \frac{\Gamma\left(\frac{k}{\tau}+1+m\right)\Gamma\left(\frac{k}{\tau}-\alpha\{1+p\}+1+m\right)\Gamma\left(\frac{l}{\tau}+\frac{k}{\tau}+2+m\right)}{\Gamma\left(\frac{k}{\tau}+2+m\right)\Gamma\left(\frac{k}{\tau}+\alpha\{n-p-1\}+2+m\right)} \frac{1^m}{m!}$$

$$= \frac{\Gamma\left(\frac{k}{\tau}-\alpha\{1+p\}+1\right)\Gamma\left(\frac{l}{\tau}+\frac{k}{\tau}+2\right)}{\left(\frac{k}{\tau}+1\right)\Gamma\left(\frac{k}{\tau}+\alpha\{n-p-1\}+2\right)}$$

$$F_{3,2}\left(\left\{\frac{k}{\tau}+1, \frac{k}{\tau}-\alpha(1+p)+1, \frac{l}{\tau}+\frac{k}{\tau}+2\right\};\left\{\frac{k}{\tau}+2, \frac{k}{\tau}+\alpha(n-p-1)+2\right\};1\right),$$

giving

$$\begin{aligned} E\left[Y_{(1:n)}^{k} Y_{(j:n)}^{l}\right] &= n(n-1)\binom{n-2}{j-2}\phi^{k+l}\alpha^2 \sum_{p=0}^{j-2}(-1)^{j-2-p}\binom{j-2}{p} \\ &\times \frac{\Gamma\left(\alpha\{n-p-1\}-\frac{l}{\tau}\right)\Gamma\left(\frac{l}{\tau}+\frac{k}{\tau}+2\right)}{\left(\frac{k}{\tau}+1\right)\Gamma\left(\frac{k}{\tau}+\alpha\{n-p-1\}+2\right)} \\ &F_{3,2}\left(\begin{array}{c}\left\{\frac{k}{\tau}+1, \frac{k}{\tau}-\alpha(1+p)+1, \frac{l}{\tau}+\frac{k}{\tau}+2\right\}; \\ \left\{\frac{k}{\tau}+2, \frac{k}{\tau}+\alpha(n-p-1)+2\right\};1\end{array}\right). \end{aligned}$$

Thus, using (4.46), we can compute an expression for $E\left[Y_{(i:n)}^{k} Y_{(j:n)}^{l}\right]$. We note that the form of this will contain a triple summation and a hypergeometric function which itself is an infinite sum. The current capabilities of Mathematica will be severely tested when evaluating numerical values for these joint expectations, especially for large sample sizes. Thus, due to limited numerical progress for joint expectations of the Burr, we omit any further details on the variance covariance matrix of the mis-specified Weibull MLEs. In fact, if we were to examine (4.25), then we see that we would have to derive joint expectations of functions of order statistics from the Burr distribution, for example, $E\left[Y_{(i:n)}^{\beta} \ln Y_{(i:n)} Y_{(j:n)}^{\beta} \ln Y_{(j:n)}\right]$. The theoretical and analytical progress possible with such functions is currently limited, and hence will be considered elsewhere.

Agreement between theoretical and sample results

We compute theoretical standard errors for the MLEs from G_b for varying sample sizes and censoring values, and compare these to sample counterparts shown in Tables 4.5 and 4.6. We also include theoretical standard errors for $\widehat{B}_{b,10}$, using our work on the complete scenario for results on the mean and variance of this quantile. We note that equivalent results for the MLEs from G_w can not be obtained due to reasons given above. The results for varying sample sizes are summarised in Table 4.7, and equivalent figures for varying r are shown in Table 4.8. We observe similar outcomes to our type I investigations. Surprisingly, we see small sample standard errors for $\widehat{\alpha}$ and $\widehat{\phi}$ when compared to the theoretical values, across most sample sizes. As in our type I scenario, we can only provide an intuitive explanation for this occurrence.

n	100	300	500	1000
r	80	240	400	800
Sd.err.$(\widehat{\tau})$	0.5339	0.3112	0.2415	0.1710
Sd.err.$(\widehat{\alpha})$	7.8364	4.6215	3.5956	2.5509
Sd.err.$\left(\widehat{\phi}\right)$	81.5245	48.0255	37.3559	26.4983
Sd.err.$\left(\widehat{B}_{b,10}\right)$	2.6690	1.5415	1.1942	0.8444

Table 4.7: Theoretical standard errors for the MLEs from G_b for varying n and $r = \frac{4n}{5}$. Data is subjected to type II censoring, and simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

n	1000	1000	1000	1000
r	500	600	700	900
Sd.err.$(\widehat{\tau})$	0.2408	0.2136	0.1910	0.1517
Sd.err.$(\widehat{\alpha})$	7.1393	4.9722	3.5524	1.7863
Sd.err.$\left(\widehat{\phi}\right)$	68.4824	48.9090	35.8706	19.1727
Sd.err.$\left(\widehat{B}_{b,10}\right)$	0.8479	0.8455	0.8453	0.8407

Table 4.8: Theoretical standard errors for the MLEs from G_b for varying r and $n = 1000$. Data is subjected to type II censoring, and simulated from G_b with $\tau = 3$, $\alpha = 4$ and $\phi = 100$.

4.3 Summary

This chapter examined aspects of mis-specifying G_w when the underlying distribution was G_b, and the data was subjected to censoring. We first considered type I censoring, and derived the EFI matrices for both Weibull and Burr distributions when we assumed no mis-specification took place; we extended these results to calculate the variance covariance matrix of the mis-specified Weibull MLEs. This then allowed us to examine theoretical properties of $\widehat{B}_{w,10}$, and enabled us to compare true and mis-specified quantiles. All results were checked using simulations, and the effects of varying the stopping time was also considered. Our results on type II censoring followed a similar structure, and we also examined aspects of mis-specifying G_w when the underlying distribution was G_b. In doing so, we derived a series of new results on expectations of single and joint order statistics, and used these to compute the EFI matrix from G_w and G_b under the assumption that the distributions were correctly chosen. We used simulations to examine how mis-specified MLEs and $\widehat{B}_{w,10}$ changed with varying r, and completed the chapter with a brief introduction on computing the variance covariance matrix of the mis-specified MLEs. This brought in expectations of joint order statistics, and due to the complexity of such joint expectations from G_b, we omitted any further details on the variance covariance matrix of the mis-specified MLEs.

We continue by examining another technique for speeding up the running time of an experiment, and consider mis-specification for accelerated data sets.

Chapter 5

Acceleration In Life Testing

5.1 Introduction

We extend our ideas of mis-specification to data sets obtained from experiments that have undergone some form of acceleration. Like censoring, acceleration is used to speed up the running time of an experiment, and results in a data set that typically contains a larger number of failed observations than would be obtained by conducting the experiment under normal conditions. Acceleration thus subjects items to higher levels of stress, which in turn induces early failures. For example, if we consider modelling the lifetimes of ball bearings, then, under normal operating conditions (known as design stress), each bearing will operate in oil at a temperature of $50°C$. Here, temperature is the stress, denoted by X, and $50°C$ is the design stress, denoted by X_s. In practice, we may have to wait months or years for an appropriate number of ball bearings to fail. Thus, we accelerate the failure times by increasing the temperature of the oil in which the ball bearings operate; we then expect a reduction in the lifetime of each bearing. Typically, (for practical reasons) accelerated tests use a set number of stress levels (usually 3 or 4), which we denote as X_i for $1 \leq i \leq k$, at which failure times are observed. We denote these failures as y_{ij}, for $1 \leq j \leq n_i$. We thus have $n = \sum_{i=1}^{k} n_i$, the total number of failures across all stress levels. In all accelerated tests, we have $X_s \leq X_1 < X_2 < ... < X_k$; if $X_s = X_1$, the experiment is described as partially accelerated. For example, consider the accelerated data set from Nelson (1990) shown in Table 5.1, which comprises the hours to failure of $n = 40$ motorettes with a new type of insulation. The experiment was conducted at 190, 220, 240 and $260°C$ with 10 observations at each stress level. Thus here, $k = 4$, $X_1 = 190$, $X_2 = 220$, $X_3 = 240$ and $X_4 = 260$, and $n_i = 10$ for all $1 \leq i \leq 4$. The test is not partially accelerated, since $X_s = 180°C < X_1$.

The aim of an accelerated life test is to make inferences about the lifetimes at design stress - usually, to describe B_{10} at normal operating conditions. This requires a scale-stress relationship, which links the scale parameter θ_i (or ϕ_i) of the underlying distribution to its corresponding stress level X_i. This implies that stress only affects the scale of the distribution; the other remaining parameters are then assumed to remain constant across

i	1	2	3	4
X_i	190°C	220°C	240°C	260°C
y_{ij}	7228	1764	1175	600
	7228	2436	1175	744
	7228	2436	1521	744
	8448	2436	1569	744
	9167	2436	1617	912
	9167	2436	1665	1128
	9167	3108	1665	1320
	9167	3108	1713	1464
	10511	3108	1761	1608
	10511	3108	1953	1896

Table 5.1: Hours to failure (y_{ij}) of $n = 40$ motorettes with a new type of insulation (Nelson, 1990), based on an experiment with $k = 4$, $n_i = 10$ for $1 \le i \le 4$, and the X_i as shown.

stress levels. Nelson (1990) describes models and methods for analysing data when this assumption needs to be relaxed, and Hirose (1993) considers likelihood ratio tests on the use of fixed shape parameters, against the alternative of allowing such parameters to depend on stress levels. Further recent research in accelerated testing is also discussed in Meeker and Escobar (1993), and Johnson (2003) considers numerous examples of accelerated data sets, and the possibility of using separate analyses at each stress level as a basis for identifying and fitting the accelerated model. Here, however, we will always assume that the underlying data set can be modelled using fixed shape parameters. Thus, the set of observations from each stress level are assumed to have the same underlying distribution but with varying scale parameters. We can link θ_i to the stress level X_i in various ways.

5.1.1 Scale-stress relationships

Log-linear model

This links stress and scale by

$$\theta_i = \exp(\alpha + \beta X_i)\,; \tag{5.1}$$

note that

$$\ln \theta_i = \alpha + \beta X_i,$$

so the logarithms of the scale parameter are linearly related to stress.

Arrhenius model

Nelson (1990) states that the Arrhenius relationship is widely used to model life times as a function of temperature, where temperature is usually measured in degrees kelvin (which is

273.16 plus the temperature in degrees centigrade). This model is most commonly used in these circumstances, but can also be used elsewhere. The scale is linked to stress by

$$\theta_i = \exp\left(\alpha + \frac{\beta}{X_i + c}\right), \tag{5.2}$$

where $c = 273.16$ converts temperature from degrees centigrade to degrees kelvin. Equivalently, this relationship can be expressed as

$$\ln\theta_i = \alpha + \frac{\beta}{X_i + c}.$$

Inverse power model

Here, we assume that scale is linked to stress by

$$\theta_i = \frac{\exp(\alpha)}{X_i^{\beta}},$$

or, equivalently,

$$\ln\theta_i = \alpha - \beta\ln X_i.$$

Exponential power model

This relationship is most often used when the stress X denotes voltage or the inverse of absolute temperature. It is defined as

$$\theta_i = \exp\left(\alpha - \beta X_i^{\lambda}\right),$$

or

$$\ln\theta_i = \alpha - \beta X_i^{\lambda};$$

unlike previous scale-stress relationships, this model has three parameters. We note that the Log-linear and Arrhenius models are special cases of this relationship.

Quadratic and Polynomial models

These types of relationships are often used when a linear form (such as (5.1)) does not provide an adequate fit; the quadratic form is

$$\ln\theta_i = \alpha + \beta_1 X_i + \beta_2 X_i^2,$$

and the extension to higher powers is clear.

Extensions to two or more stresses

The relationships listed above link scale to a single stress level X. However, some accelerated tests involve more than one accelerating factor. For example, electrical items such as capacitors are affected both by high voltage and high temperature. More general mathematical formula can describe such relationships, but generally at the expense of increasing the number of parameters in the model.

Here, we only consider using the Log-linear and Arrhenius scale-stress relationships; the literature shows these to be the most common models used. For convenience, we will also refer to stress as temperature when we run simulations. For the Log-linear model, temperature will be in terms of degrees centigrade, although it makes no difference whether we work in degrees centigrade or degrees kelvin, whilst, for reasons described above, we convert temperature for the Arrhenius model to degrees kelvin. In the next section, we discuss possible choices for the number of stress levels, the corresponding values of the X_i, and the number of observations at each level; such values will then be used in our simulation experiments.

5.1.2 Choice of sample size and stress levels

We consider the choice of k, the X_i and n_i for Log-linear and Arrhenius models. We also derive some sensible parameter values in each scale-stress relationship, calculated on the basis of what we expect to see in practice. We list some possible choices for these below.

- We allow k to range from 2 to 4; these seem to be the number of stress levels most commonly used.

- For each value of k, we can then allow the sample size to vary, first keeping equal proportions of observations at each level. For $k = 2$, we consider total sample sizes of 100, 200, 500 and 1000, so, for example, when $n = 500$, $n_1 = n_2 = 250$; for $k = 3$, we let $n = 75$, 150, 300, 900 and 1500. Finally, for $k = 4$, we set $n = 100$, 200, 400, 800 and 2000.

- Again, for each value of k, we can keep n fixed but vary the proportion of observations at each X_i. Based on practical considerations of actually conducting experiments, we will assume that $n_i \geq 25$. Thus, for $k = 2$, we consider a total sample size of 200, and so, for example, can consider $n_1 = 25$ (so $n_2 = 175$), $n_1 = 50$ (so $n_2 = 150$), and so on. For $k = 3$, we take $n = 300$; finally for $k = 4$, we have $n = 200$. Different arrangements of n_i for varying k will be listed in our tables of simulations.

- We also allow the parameter values of the underlying distribution to vary, but fix the values taken in the scale-stress relationship. Values for these parameters will rely on the distribution function used, and will be listed as appropriate.

- Many other aspects of the experiment can also be varied; for example, Johnson (2003) examines the effect of varying the middle stress on the standard error of $\widehat{B}_{10}$. We could also vary the parameters of the scale-stress relationship, and examine cases where we have more extreme levels of acceleration. However, such experiments will be considered elsewhere.

We consider how to choose X_i, α and β for Log-linear and Arrhenius models below.

The Log-linear model

Reviewing examples from Nelson (1990), we see that temperatures, especially of industrial experiments, are typically quite high, and range from $50°C$ to $250°C$. As a result, we take $X_1 = 50°C$, $X_k = 200°C$, and let the scale parameters at these stress levels be 2000 and 200, respectively. These figures (and their units) are arbitrary, but can be related to practical situations, and we would not generally see items placed under higher stress having longer lifetimes, and hence larger scale parameters, than their lower stress counterparts. These figures correspond to an acceleration factor of 10; published data sets show that this figure is quite reasonable, and the accelerated data set in Table 5.1 has a similar acceleration factor. For these scale parameters and stress levels, and the Log-linear relationship (5.1), we therefore need

$$\alpha = 8.36843,\ \beta = -0.01535;$$

on rounding, we take

$$\alpha = 8,\ \beta = -0.02.$$

We will also assume that $X_s = 50°C$, so we have a partially accelerated experiment. For simulations with $k = 3$, we have

$$X_1 = 50,\ X_2 = 150,\ X_3 = 200;$$

for $k = 4$, we use

$$X_1 = 50,\ X_2 = 150,\ X_3 = 180,\ X_4 = 200.$$

The Arrhenius model

We adapt the above approach to obtain parameter values for the Arrhenius relationship. Thus, $X_1 + c = 323.16°K$, $X_k + c = 473.16°K$, and again let the scale parameters at the lowest and highest stress levels be 2000 and 200 respectively. From (5.2), and after rounding, we have

$$\alpha = 0.3,\ \beta = 2347.$$

For $k = 3$, we have

$$X_1 + c = 323.16,\ X_2 + c = 423.16,\ X_3 + c = 473.16;$$

for $k = 4$

$$X_1 + c = 323.16,\ X_2 + c = 423.16,\ X_3 + c = 453.16,\ X_4 + c = 473.16.$$

Again, we consider a partially accelerated experiment with $X_s + c = X_1 + c = 323.16°K$.

We first derive the theory necessary to fit the Weibull, Burr, Gamma and Lognormal distributions to an accelerated data set containing all failures, when we have either the Log-linear or Arrhenius model linking scale to stress. We also consider the EFI matrix of the MLEs, when we assume the correct model is specified; in later chapters, we discuss the effects of mis-specification and censoring. We use similar notation to that previously established, but now denote pdfs and cdfs with capital subscripts. Thus, for example, the cdf from the accelerated Weibull Arrhenius distribution is denoted by G_{WA}; we also introduce further notation to include the Log-linear and Arrhenius scale-stress relationships simultaneously, writing G_{W*}, where we set $* = P$ if we are using the Log-linear model to link scale to stress, and $* = A$ for the Arrhenius relationship.

5.2 Fitting G_{W*}

We assume that the data y_{ij} represents the j^{th} observation from stress level X_i, where the underlying distribution is $G_{W*}(y_{ij}; B_*, \theta_{i*})$. We also let Y_{ij} denote the corresponding random variable, and Y_i represent a random variable from stress level X_i. Watkins (1991) outlines ML estimation for the accelerated Weibull distribution; we summarise the main points here, where now, using the above notation,

$$\theta_{iP} = \exp\left(\alpha_{WP} + \beta_{WP} X_i\right) \qquad \text{for } * = P,$$

and

$$\theta_{iA} = \exp\left(\alpha_{WA} + \frac{\beta_{WA}}{X_i + c}\right) \qquad \text{for } * = A.$$

Using this notation, the pdf of G_{W*} is

$$\frac{B_* y_{ij}^{B_*-1}}{\exp\left\{\alpha_{W*} + \beta_{W*}\rho\left(X_i\right)\right\}^{B_*}} \exp\left\{-\left(\frac{y_{ij}}{\exp\left\{\alpha_{W*} + \beta_{W*}\rho\left(X_i\right)\right\}}\right)^{B_*}\right\} \quad \text{for } 1 \leq j \leq n_i$$

where

$$\rho(X_i) = \begin{cases} X_i & \text{for } * = P \\ (X_i + c)^{-1} & \text{for } * = A \end{cases} \tag{5.3}$$

Thus, the theory developed can cover both scale-stress relationships simultaneously. The likelihood is

$$\begin{aligned} L_{W*}(B_*, \alpha_{W*}, \beta_{W*}) &= \prod_{i=1}^{k}\prod_{j=1}^{n_i} \frac{B_* y_{ij}^{B_*-1}}{\exp\{\alpha_{W*} + \beta_{W*}\rho(X_i)\}^{B_*}} \\ &\times \exp\left\{-\left(\frac{y_{ij}}{\exp\{\alpha_{W*} + \beta_{W*}\rho(X_i)\}}\right)^{B_*}\right\}, \end{aligned}$$

from which the log-likelihood is

$$l_{W*} = n\ln B_* + (B_* - 1)S_e - nB_*\alpha_{W*} - B_*\beta_{W*}\sum_{i=1}^{k} n_i\rho(X_i) - \exp(-B_*\alpha_{W*})S(B_*, \beta_{W*}), \tag{5.4}$$

where, now,

$$S_e = \sum_{i=1}^{k}\sum_{j=1}^{n_i} \ln y_{ij},$$

and

$$S(B_*, \beta_{W*}) = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^{B_*} \exp\{-B_*\beta_{W*}\rho(X_i)\}.$$

The three partial derivatives are

$$\begin{aligned} \frac{\partial l_{W*}}{\partial B_*} &= \frac{n}{B_*} + S_e - n\alpha_{W*} - \beta_{W*}\sum_{i=1}^{k} n_i\rho(X_i) - \\ &\quad \exp(-B_*\alpha_{W*})\{S_{1,0}(B_*, \beta_{W*}) - \alpha_{W*}S(B_*, \beta_{W*})\}, & (5.5) \\ \frac{\partial l_{W*}}{\partial \alpha_{W*}} &= -nB_* + B_*\exp(-B_*\alpha_{W*})S(B_*, \beta_{W*}), & (5.6) \end{aligned}$$

and

$$\frac{\partial l_{W*}}{\partial \beta_{W*}} = -B_*\sum_{i=1}^{k} n_i\rho(X_i) - \exp(-B_*\alpha_{W*})S_{0,1}(B_*, \beta_{W*}), \tag{5.7}$$

with

$$
\begin{aligned}
S_{1,0}(B_*, \beta_{W*}) &= \frac{\partial}{\partial B_*} S(B_*, \beta_{W*}) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}^{B_*} \exp\{-B_* \beta_{W*} \rho(X_i)\} \{\ln y_{ij} - \beta_{W*} \rho(X_i)\}, \\
S_{0,1}(B_*, \beta_{W*}) &= \frac{\partial}{\partial \beta_{W*}} S(B_*, \beta_{W*}) \\
&= -\sum_{i=1}^{k} \sum_{j=1}^{n_i} B_* \rho(X_i) y_{ij}^{B_*} \exp\{-B_* \beta_{W*} \rho(X_i)\}.
\end{aligned}
$$

As above, we can reduce the number of parameters we estimate, and make use of a profile log-likelihood in order to obtain the remaining MLEs. By equating (5.6) to zero, we see that

$$
\alpha_{W*} = B_*^{-1} \ln \left\{ \frac{S(B_*, \beta_{W*})}{n} \right\};
$$

using this, we can write the profile log-likelihood as

$$
l_{W*}^{+} = n \ln B_* + (B_* - 1) S_e - n \ln S(B_*, \beta_{W*}) - B_* \beta_{W*} \sum_{i=1}^{k} n_i \rho(X_i), \tag{5.8}
$$

and the two profile score functions are

$$
\begin{aligned}
\frac{\partial l_{W*}^{+}}{\partial B_*} &= \frac{n}{B_*} + S_e - \frac{n S_{1,0}(B_*, \beta_{W*})}{S(B_*, \beta_{W*})} - \beta_{W*} \sum_{i=1}^{k} n_i \rho(X_i), \\
\frac{\partial l_{W*}^{+}}{\partial \beta_{W*}} &= -\frac{n S_{0,1}(B_*, \beta_{W*})}{S(B_*, \beta_{W*})} - B_* \sum_{i=1}^{k} n_i \rho(X_i).
\end{aligned}
$$

We use the Newton-Raphson method to locate the roots of these derivatives, and so also require second partial derivatives of (5.8). For completeness, these are given by

$$
\frac{\partial^2 l_{W*}^{+}}{\partial \beta_{W*}^2} = -n \left\{ \frac{S_{0,2}(B_*, \beta_{W*}) S(B_*, \beta_{W*}) - S_{0,1}(B_*, \beta_{W*})^2}{S(B_*, \beta_{W*})^2} \right\},
$$

where

$$
\begin{aligned}
S_{0,2}(B_*, \beta_{W*}) &= \frac{\partial}{\partial \beta_{W*}} S_{0,1}(B_*, \beta_{W*}) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{n_i} B_*^2 \rho(X_i)^2 y_{ij}^{B_*} \exp\{-B_* \beta_{W*} \rho(X_i)\},
\end{aligned}
$$

$$\frac{\partial^2 l^+_{W*}}{\partial B_*^2} = -\frac{n}{B_*^2} - n\left\{\frac{S_{2,0}(B_*,\beta_{W*})\,S(B_*,\beta_{W*}) - S_{1,0}(B_*,\beta_{W*})^2}{S(B_p,\beta_{W*})^2}\right\},$$

where

$$\begin{aligned} \dot{S}_{2,0}(B_*,\beta_{W*}) &= \frac{\partial}{\partial B_*} S_{1,0}(B_*,\beta_{W*}) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^{B_*}\exp\{-B_*\beta_{W*}\rho(X_i)\}\{\ln y_{ij} - \beta_{W*}\rho(X_i)\}^2, \end{aligned}$$

and finally,

$$\frac{\partial^2 l^+_{W*}}{\partial B_*\partial\beta_{W*}} = -\sum_{i=1}^{k} n_i\rho(X_i) - n\left\{\frac{S_{1,1}(B_*,\beta_{W*})\,S(B_*,\beta_{W*}) - S_{1,0}(B_*,\beta_{W*})\,S_{0,1}(B_*,\beta_{W*})}{S(B_*,\beta_{W*})^2}\right\},$$

with $S_{1,1}(B_*,\beta_{W*})$ equal to

$$-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\rho(X_i)\,y_{ij}^{B_*}\exp\{-B_*\beta_{W*}\rho(X_i)\}\left[1 + B_*\{\ln y_{ij} - \beta_{W*}\rho(X_i)\}\right].$$

To obtain starting values for B_* and β_{W*}, we can fit separate Weibull distributions to each stress level, and chose a value of B_* based on these. To find an initial value for β_{W*}, we take the first and last estimates, θ_{1*} and θ_{k*}, and use the fact that

$$\frac{\theta_{1*}}{\theta_{k*}} = \frac{\exp\{\alpha_{W*} + \beta_{W*}\rho(X_1)\}}{\exp\{\alpha_{W*} + \beta_{W*}\rho(X_k)\}} = \frac{\exp\{\beta_{W*}\rho(X_1)\}}{\exp\{\beta_{W*}\rho(X_k)\}}$$

So, on solving, an appropriate starting value for β_{W*} is given by

$$\frac{\ln\widehat{\theta}_{1*} - \ln\widehat{\theta}_{k*}}{\rho(X_1) - \rho(X_k)}.$$

Next, we consider the EFI matrix of the Weibull MLEs, and include a discussion on the asymptotic variance of $\widehat{B}_{W,10}$, since this is of considerable use when we extrapolate back to design stress, and make inferences concerning the reliability of items at normal operating conditions. Again, this is carried out simultaneously for both Log-linear and Arrhenius relationships.

5.2.1 The EFI matrix of the Weibull MLEs

From our previous work, we state that, asymptotically, the Weibull MLEs, under the assumption that this distribution has been chosen correctly to model the data, will have a Normal distribution with mean $(B_*,\alpha_{W*},\beta_{W*})'$ and variance covariance matrix equal to (3.1). Thus, we require expected values of second derivatives. Using (5.5), (5.6) and (5.7), we obtain the six partial second derivatives

$$
\begin{aligned}
\frac{\partial^2 l_{W*}}{\partial B_*^2} &= \frac{-n}{B_*^2} - \exp\left(-B_*\alpha_{W*}\right)\left\{\begin{array}{c} S_{2,0}\left(B_*,\beta_{W*}\right) \\ -2\alpha_{W*}S_{1,0}\left(B_*,\beta_{W*}\right) \\ +\alpha_{W*}^2 S\left(B_*,\beta_{W*}\right)\end{array}\right\}, \\
\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}^2} &= -B_*^2\exp\left(-B_*\alpha_{W*}\right)S\left(B_*,\beta_{W*}\right), \\
\frac{\partial^2 l_{W*}}{\partial \beta_{W*}^2} &= -\exp\left(-B_*\alpha_{W*}\right)S_{0,2}\left(B_*,\beta_{W*}\right), \\
\frac{\partial^2 l_{W*}}{\partial B_*\partial \alpha_{W*}} &= -n + \exp\left(-B_*\alpha_{W*}\right)\left[\begin{array}{c} B_*S_{1,0}\left(B_*,\beta_{W*}\right) + \\ \left\{1 - B_*\alpha_{W*}\right\}S\left(B_*,\beta_{W*}\right)\end{array}\right], \\
\frac{\partial^2 l_{W*}}{\partial B_*\partial \beta_{W*}} &= -\sum_{i=1}^{k} n_i\rho\left(X_i\right) - \exp\left(-B_*\alpha_{W*}\right)\left\{\begin{array}{c} S_{1,1}\left(B_*,\beta_{W*}\right) - \\ \alpha_{W*}S_{0,1}\left(B_*,\beta_{W*}\right)\end{array}\right\}, \\
\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}\partial \beta_{W*}} &= B_*\exp\left(-B_*\alpha_{W*}\right)S_{0,1}\left(B_*,\beta_{W*}\right).
\end{aligned}
$$

To obtain the variance covariance matrix of the MLEs, we take expected values of these second derivatives. We take into account that, across stress levels, observations are not identically distributed, but within stress levels they are independently and identically distributed. On examining the second derivatives, we see that we require expected values of $S\left(B_*,\beta_{W*}\right)$, $S_{1,0}\left(B_*,\beta_{W*}\right)$, $S_{0,1}\left(B_*,\beta_{W*}\right)$, $S_{2,0}\left(B_*,\beta_{W*}\right)$, $S_{0,2}\left(B_*,\beta_{W*}\right)$ and $S_{1,1}\left(B_*,\beta_{W*}\right)$. We first consider

$$
\begin{aligned}
E\left[S\left(B_*,\beta_{W*}\right)\right] &= E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i} Y_{ij}^{B_*}\exp\left\{-B_*\beta_{W*}\rho\left(X_i\right)\right\}\right] \\
&= E\left[\sum_{i=1}^{k} n_i Y_i^{B_*}\exp\left\{-B_*\beta_{W*}\rho\left(X_i\right)\right\}\right] \\
&= \sum_{i=1}^{k} n_i\exp\left\{-B_*\beta_{W*}\rho\left(X_i\right)\right\}E\left[Y_i^{B_*}\right].
\end{aligned}
$$

Using results in, for example, Watkins (1998), we may immediately state that if Y_i is from $G_{W*}\left(y;B_*,\theta_{i*}\right)$ then

$$
E\left[Y_i^m\right] = \theta_{i*}^m\Gamma\left(\frac{m}{B_*}+1\right) = \exp\left(m\alpha_{W*}\right)\exp\left\{m\beta_{W*}\rho\left(X_i\right)\right\}\Gamma\left(\frac{m}{B_*}+1\right), \tag{5.9}
$$

and hence

$$
E\left[S\left(B_*,\beta_{W*}\right)\right] = n\exp\left(B_*\alpha_{W*}\right).
$$

Next, we consider

$$E\left[S_{1,0}\left(B_*,\beta_{W*}\right)\right]=\sum_{i=1}^{k} n_i \exp\left(-B_*\beta_{W*}\rho\left(X_i\right)\right)\left\{E\left[Y_i^{B_*}\ln Y_i\right]-\beta_{W*}\rho\left(X_i\right)E\left[Y_i^{B_*}\right]\right\}.$$

We differentiate (5.9) with respect to m to obtain

$$\begin{aligned} E\left[Y_i^m \ln Y_i\right] &= \exp\left(m\alpha_{W*}\right)\exp\left\{m\beta_{W*}\rho\left(X_i\right)\right\}\Gamma\left(\frac{m}{B_*}+1\right) \\ &\quad \times\left\{\alpha_{W*}+\beta_{W*}\rho\left(X_i\right)+\frac{\Psi\left(\frac{m}{B_*}+1\right)}{B_*}\right\}; \end{aligned} \tag{5.10}$$

we therefore have

$$E\left[S_{1,0}\left(B_*,\beta_{W*}\right)\right]=n\exp\left(B_*\alpha_{W*}\right)\left\{\alpha_{W*}+\frac{\Psi\left(2\right)}{B_*}\right\}.$$

Now, we examine

$$\begin{aligned} E\left[S_{0,1}\left(B_*,\beta_{W*}\right)\right] &= -\sum_{i=1}^{k} n_i B_*\rho\left(X_i\right)\exp\left\{-B_*\beta_{W*}\rho\left(X_i\right)\right\}E\left[Y_i^{B_*}\right] \\ &= -B_*\exp\left(B_*\alpha_{W*}\right)\sum_{i=1}^{k} n_i\rho\left(X_i\right). \end{aligned}$$

Next, we consider

$$E\left[S_{2,0}\left(B_*,\beta_{W*}\right)\right]=\sum_{i=1}^{k} n_i\exp\left\{-B_*\beta_{W*}\rho\left(X_i\right)\right\}\left\{\begin{array}{c} E\left[Y_i^{B_*}\left(\ln Y_i\right)^2\right] \\ -2\beta_{W*}\rho\left(X_i\right)E\left[Y_i^{B_*}\ln Y_i\right] \\ +\beta_{W*}^2\rho\left(X_i\right)^2 E\left[Y_i^{B_*}\right] \end{array}\right\}.$$

We differentiate (5.10) (again with respect to m) to get

$$\begin{aligned} E\left[Y_i^m\left(\ln Y_i\right)^2\right] &= \exp\left(m\alpha_{W*}\right)\exp\left\{m\beta_{W*}\rho\left(X_i\right)\right\}\Gamma\left(\frac{m}{B_*}+1\right) \\ &\quad \times\left[\left\{\alpha_{W*}+\beta_{W*}\rho\left(X_i\right)+\frac{\Psi\left(\frac{m}{B_*}+1\right)}{B_*}\right\}^2+\frac{\Psi'\left(\frac{m}{B_*}+1\right)}{B_*^2}\right] \end{aligned} \tag{5.11}$$

and we see that

$$E\left[S_{2,0}\left(B_*,\beta_{W*}\right)\right]=n\exp\left(B_*\alpha_{W*}\right)\left[\left\{\alpha_{W*}+\frac{\Psi\left(2\right)}{B_*}\right\}^2+\frac{\Psi'\left(2\right)}{B_*^2}\right].$$

Next, we need

$$\begin{aligned} E\left[S_{0,2}\left(B_*, \beta_{W*}\right)\right] &= \sum_{i=1}^{k} n_i B_*^2 \rho\left(X_i\right)^2 \exp\left\{-B_* \beta_{W*} \rho\left(X_i\right)\right\} E\left[Y_i^{B_*}\right] \\ &= B_*^2 \exp\left(B_* \alpha_{W*}\right) \sum_{i=1}^{k} n_i \rho\left(X_i\right)^2 . \end{aligned}$$

Finally, we examine

$$\begin{aligned} E\left[S_{1,1}\left(B_*, \beta_{W*}\right)\right] &= -\sum_{i=1}^{k} n_i \rho\left(X_i\right) \exp\left\{-B_* \beta_{W*} \rho\left(X_i\right)\right\} \\ &\quad \times \left\{B_* E\left[Y_i^{B_*} \ln Y_i\right] - \left(B_* \beta_{W*} \rho\left(X_i\right) - 1\right) E\left[Y_i^{B_*}\right]\right\} \\ &= -\left\{1 + B_* \alpha_{W*} + \Psi(2)\right\} \exp\left(B_* \alpha_{W*}\right) \sum_{i=1}^{k} n_i \rho\left(X_i\right). \end{aligned}$$

Now that we have these functions, we can derive the elements of the EFI matrix. The diagonal elements are

$$\begin{aligned} -E\left[\frac{\partial^2 l_{W*}}{\partial B_*^2}\right] &= \frac{n}{B_*^2} + \exp\left(-B_* \alpha_{W*}\right) \left\{ \begin{array}{c} E\left[S_{2,0}\left(B_*, \beta_{W*}\right)\right] - \\ 2\alpha_{W*} E\left[S_{1,0}\left(B_*, \beta_{W*}\right)\right] + \alpha_{W*}^2 E\left[S\left(B_*, \beta_{W*}\right)\right] \end{array} \right\} \\ &= \frac{n}{B_*^2} \left\{1 + \Psi(2)^2 + \Psi'(2)\right\}, \\ -E\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}^2}\right] &= B_*^2 \exp\left(-B_* \alpha_{W*}\right) E\left[S\left(B_*, \beta_{W*}\right)\right] = n B_*^2, \\ -E\left[\frac{\partial^2 l_{W*}}{\partial \beta_{W*}^2}\right] &= \exp\left(-B_* \alpha_{W*}\right) E\left[S_{2,0}\left(B_*, \beta_{W*}\right)\right] = B_*^2 \sum_{i=1}^{k} n_i \rho\left(X_i\right)^2 , \end{aligned}$$

while the off diagonal elements are

$$\begin{aligned} -E\left[\frac{\partial^2 l_{W*}}{\partial B_* \partial \alpha_{W*}}\right] &= n - \exp\left(-B_* \alpha_{W*}\right) \left\{B_* E\left[S_{1,0}\left(B_*, \beta_{W*}\right)\right] + \left(1 - B_* \alpha_{W*}\right) E\left[S\left(B_*, \beta_{W*}\right)\right]\right\} \\ &= -n\Psi(2), \\ -E\left[\frac{\partial^2 l_{W*}}{\partial B_* \partial \beta_{W*}}\right] &= \sum_{i=1}^{k} n_i \rho\left(X_i\right) + \exp\left(-B_* \alpha_{W*}\right) \left\{E\left[S_{1,1}\left(B_*, \beta_{W*}\right)\right] - \alpha_{W*} E\left[S_{0,1}\left(B_*, \beta_{W*}\right)\right]\right\} \\ &= -\Psi(2) \sum_{i=1}^{k} n_i \rho\left(X_i\right), \\ -E\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*} \partial \beta_{W*}}\right] &= -B_* \exp\left(-B_* \alpha_{W*}\right) E\left[S_{0,1}\left(B_*, \beta_{W*}\right)\right] = B_*^2 \sum_{i=1}^{k} n_i \rho\left(X_i\right). \end{aligned}$$

Using these in (3.1), we see that, asymptotically, $\left(\widehat{B}_*, \widehat{\alpha}_{W*}, \widehat{\beta}_{W*}\right)'$ has a Normal distribution with mean vector $(B_*, \alpha_{W*}, \beta_{W*})'$ and variance covariance matrix A^{-1}, where

$$A = \begin{bmatrix} nB_*^{-2}\left\{1+\Psi(2)^2+\Psi'(2)\right\} & & \\ -n\Psi(2) & nB_*^2 & \\ -\Psi(2)\sum_{i=1}^k n_i\rho(X_i) & B_*^2\sum_{i=1}^k n_i\rho(X_i) & B_*^2\sum_{i=1}^k n_i\rho(X_i)^2 \end{bmatrix}.$$

5.2.2 The asymptotic variance of $\widehat{B}_{W,10}$

Estimation of the quantile B_{10} at design stress plays a significant role in determining and assessing the reliability of items, since the experimenter is really interested in the performance of the items at X_s. Thus, we extrapolate back to assess the reliability of the components, using the relationship between scale and stress. At X_s, $\widehat{B}_{W,10}$ is given by

$$\exp\left\{\widehat{\alpha}_{W*} + \widehat{\beta}_{W*}\rho(X_s)\right\}(-\ln 0.9)^{\frac{1}{\widehat{B}_*}};$$

we obtain a linear approximation to this quantile using a Taylor series centered on the true values α_{W*}, β_{W*} and B_*; we have

$$\widehat{B}_{W,10} \simeq B_{W,10} + \begin{pmatrix} c_{B_*} & c_{\alpha_{W*}} & c_{\beta_{W*}} \end{pmatrix} \begin{pmatrix} \widehat{B}_* - B_* \\ \widehat{\alpha}_{W*} - \alpha_{W*} \\ \widehat{\beta}_{W*} - \beta_{W*} \end{pmatrix},$$

where

$$\begin{pmatrix} c_{B_*} \\ c_{\alpha_{W*}} \\ c_{\beta_{W*}} \end{pmatrix} = \begin{pmatrix} \frac{-\exp\{\alpha_{W*}+\beta_{W*}\rho(X_s)\}(-\ln 0.9)^{\frac{1}{B_*}}\ln(-\ln 0.9)}{B_*^2} \\ B_{W,10} \\ \rho(X_s)\exp\{\alpha_{W*}+\beta_{W*}\rho(X_s)\}(-\ln 0.9)^{\frac{1}{B_*}} \end{pmatrix}. \tag{5.12}$$

On taking expected values, we have $E\left[\widehat{B}_{W,10}\right] \simeq B_{W,10}$; the variance of the estimator of this quantile is given by the appropriate application of (3.2). We check these results, and those based on the EFI matrix of the Weibull MLEs, in the next section, where we summarise a series of simulations for various sets of parameter values, sample sizes and stress levels.

5.2.3 Fitting G_{WP}

Possible choices for k, X_i, n and n_i are outlined in Section 5.1.2; we use these values throughout our simulations. As discussed, we can also vary the parameters of the underlying distribution. We run experiments with B_P equal to 0.5, 1, 2 and 3, but here only report results for $B_P = 2$, since they show the same features for all values of this parameter. The simulations are summarised in Tables 5.2, 5.3 and 5.4, for $k = 2$, 3 and 4 respectively. For each set of stress levels, we vary the overall sample size, and the number of observations

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_P$	2.0381	2.0197	2.0078	2.0040
S	0.1623	0.1140	0.0706	0.0490
T	0.1559	0.1103	0.0697	0.0493
$\widehat{\alpha}_{WP}$	7.9956	7.9991	7.9993	7.9994
S	0.0987	0.0693	0.0437	0.0309
T	0.0986	0.0697	0.0441	0.0312
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0007	0.0005	0.0003	0.0002
T	0.0007	0.0005	0.0003	0.0002
$\widehat{B}_{W,10}$	362.5252	359.6669	357.2740	356.6435
S	45.6550	31.9996	19.9137	13.9800
T	44.8286	31.6986	20.0479	14.1760
n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_P$	2.0205	2.0207	2.0201	2.0206
S	0.1116	0.1133	0.1120	0.1132
T	0.1103	0.1103	0.1103	0.1103
$\widehat{\alpha}_{WP}$	7.9873	8.0024	7.9941	8.0004
S	0.1363	0.0619	0.0964	0.0607
T	0.1344	0.0615	0.0960	0.0605
$\widehat{\beta}_{WP}$	-0.0199	-0.0201	-0.0200	-0.0200
S	0.0007	0.0007	0.0006	0.0005
T	0.0007	0.0007	0.0005	0.0005
$\widehat{B}_{W,10}$	358.1386	359.9149	358.9636	359.8369
S	44.3080	29.9109	36.2418	30.5286
T	44.2165	29.4795	36.3534	29.9869

Table 5.2: Fitting G_{WP} to G_{WP} for $k = 2$, $B_P = 2$. We show the sample means and standard deviations of parameters, where figures are based on at least 10000 replications. Throughout, sample standard errors are denoted by S, and their theoretical counterparts by T.

used at each stress. We include details on the sample means of the MLEs, the sample standard errors of these estimates, the value of $\widehat{B}_{W,10}$ (which we compare to a true value of 355.9593 for $B_P = 2$) and the standard error for this quantile. We also include theoretical counterparts for all estimates, since this will verify the results established above.

The effect of varying n

As expected, when we increase the overall sample size, we observe an improved agreement between sample and theoretical values of the standard errors of the MLEs, and these standard errors for $\widehat{B}_{W,10}$. These standard errors also decrease as n increases. We observe the sample means of the MLEs and $\widehat{B}_{W,10}$ tend to their true values for larger sample sizes; this is true across all values of k.

n_1, n_2 / n_3	25,25 / 25	50,50 / 50	100,100 / 100	300,300 / 300	500,500 / 500
$\widehat{B}_P$	2.0566	2.0268	2.0145	2.0038	2.0025
S	0.1938	0.1321	0.0913	0.0521	0.0402
T	0.1801	0.1273	0.0900	0.0520	0.0403
$\widehat{\alpha}_{WP}$	7.9933	7.9970	7.9989	7.9996	7.9998
S	0.1384	0.0978	0.0687	0.0397	0.0307
T	0.1376	0.0973	0.0688	0.0397	0.0308
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0009	0.0007	0.0005	0.0003	0.0002
T	0.0009	0.0007	0.0005	0.0003	0.0002
$\widehat{B}_{W,10}$	366.0449	360.8025	358.7481	356.6466	356.4166
S	56.7415	39.5388	27.3622	15.8955	12.2724
T	54.8755	38.8028	27.4377	15.8412	12.2705

n_1, n_2 / n_3	50,100 / 150	25,200 / 75	25,25 / 250	200,50 / 50
$\widehat{B}_P$	2.0123	2.0120	2.0143	2.0122
S	0.0903	0.0903	0.0916	0.0911
T	0.0900	0.0900	0.0900	0.0900
$\widehat{\alpha}_{WP}$	7.9968	7.9955	7.9888	8.0001
S	0.0889	0.1231	0.1296	0.0541
T	0.0909	0.1211	0.1282	0.0532
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0199	-0.0200
S	0.0005	0.0008	0.0007	0.0005
T	0.0005	0.0008	0.0007	0.0005
$\widehat{B}_{W,10}$	358.0819	358.3118	357.2576	358.2118
S	31.2101	37.3768	40.5615	25.0714
T	31.6150	36.8697	40.0490	24.7807

Table 5.3: Fitting G_{WP} to G_{WP} for $k = 3$, $B_P = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 25,25	50,50 50,50	100,100 100,100	200,200 200,200	500,500 500,500	25,25, 75,75
$\widehat{B}_P$	2.0392	2.0208	2.0112	2.0054	2.0021	2.0201
S	0.1624	0.1128	0.0793	0.0559	0.0351	0.1112
T	0.1559	0.1103	0.0780	0.0551	0.0349	0.1103
$\widehat{\alpha}_{WP}$	7.9907	7.9953	7.9968	7.9994	7.9995	7.9897
S	0.1368	0.0973	0.0686	0.0477	0.0301	0.1330
T	0.1363	0.0964	0.0682	0.0482	0.0305	0.1311
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0199
S	0.0009	0.0006	0.0004	0.0003	0.0002	0.0008
T	0.0009	0.0006	0.0004	0.0003	0.0002	0.0008
$\widehat{B}_{W,10}$	362.1429	359.3485	357.6686	356.9730	356.3001	358.5559
S	51.0878	36.1488	25.4415	17.8975	11.2579	43.0575
T	50.5236	35.7256	25.2618	17.8628	11.2974	42.8015
n_1, n_2 n_3, n_4	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_P$	2.0203	2.0232	2.0212	2.0215	2.0184	2.0189
S	0.1134	0.1142	0.1124	0.1123	0.1111	0.1129
T	0.1103	0.1103	0.1103	0.1103	0.1103	0.1103
$\widehat{\alpha}_{WP}$	7.9971	7.9932	7.9960	7.9991	7.9897	7.9897
S	0.0835	0.1284	0.0818	0.0736	0.1307	0.1331
T	0.0829	0.1262	0.0817	0.0731	0.1301	0.1319
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0199	-0.0199
S	0.0006	0.0008	0.0006	0.0006	0.0007	0.0008
T	0.0006	0.0008	0.0006	0.0006	0.0007	0.0008
$\widehat{B}_{W,10}$	359.5264	359.9914	359.3948	360.0602	358.2204	358.3142
S	33.5322	42.2586	33.2908	31.7788	42.7243	42.4734
T	33.0141	41.4431	33.1637	31.5513	42.7478	42.1629

Table 5.4: Fitting G_{WP} to G_{WP} for $k = 4$, $B_P = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2, n_3	100,100,100	200,50,50	25,25,250
$n\sum_{i=1}^{k} n_i X_i^2$	1.95×10^9	1.09×10^9	3.19×10^9
$\left(\sum_{i=1}^{k} n_i X_i\right)^2$	1.60×10^9	7.56×10^8	3.03×10^9
$n\sum_{i=1}^{k} n_i X_i^2 - \left(\sum_{i=1}^{k} n_i X_i\right)^2$	3.50×10^8	3.31×10^8	1.63×10^8

Table 5.5: Breakdown of the denominator of the variance of $\widehat{\alpha}_{WP}$ and $\widehat{\beta}_{WP}$ for various n_i.

The effect of varying n_i

When we vary the allocation of items at each stress level, the theoretical standard error of $\widehat{B}_P$ is not affected, and remains constant. This is consistent with the EFI matrix A above, since, from the inverse of this matrix, we have

$$Var\left(\widehat{B}_*\right) \simeq \frac{B_*^2}{n\left\{1+\Psi'(2)\right\}},$$

which only depends on n, and not the allocations n_i. In contrast, the theoretical standard errors of $\widehat{\alpha}_{WP}$ and $\widehat{\beta}_{WP}$ are greatly affected by this allocation of the total sample size, and the more observations we test at the higher stress levels, the larger the theoretical and sample standard errors of this parameter. Again, from the inverse of the EFI matrix, we have

$$\begin{aligned} Var\left(\widehat{\alpha}_{W*}\right) \simeq & \frac{\sum_{i=1}^{k} n_i \rho\left(X_i\right)^2}{B_*^2\left\{n\sum_{i=1}^{k} n_i\rho\left(X_i\right)^2 - \left(\sum_{i=1}^{k} n_i\rho\left(X_i\right)\right)^2\right\}} \\ & + \frac{\Psi\left(2\right)^2}{nB_*^2\left\{1+\Psi'(2)\right\}}, \end{aligned}$$

and

$$Var\left(\widehat{\beta}_{W*}\right) \simeq \frac{n}{B_*^2\left\{n\sum_{i=1}^{k} n_i\rho\left(X_i\right)^2 - \left(\sum_{i=1}^{k} n_i\rho\left(X_i\right)\right)^2\right\}}.$$

If we examine the variance of $\widehat{\alpha}_{W*}$, then the numerator of the first term of this function will increase if more observations are allocated to the higher stress levels, since the $\rho\left(X_i\right)$ increases with X_i. It is the denominator of this variance, which also appears in the variance of $\widehat{\beta}_{W*}$, that we next investigate. We consider specific examples to illustrate the behaviour of the denominator as the number of observations allocated to the higher stress levels increase. Table 5.5 summarises numerical results for terms in this denominator, together with the denominator itself, for $k = 3$, $* = P$, and various n_i. We observe that the two functions which make up the denominator of both variances are closer to one another for larger n_k; the difference then decreases, and so the variance, and hence the standard error, increase as we divide by a smaller denominator. The theoretical standard error of $\widehat{B}_{W,10}$ is also greatly

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_A$	2.0394	2.0201	2.0076	2.0038
S	0.1630	0.1142	0.0710	0.0494
T	0.1455	0.1029	0.0651	0.0460
$\widehat{\alpha}_{WA}$	0.2959	0.2963	0.2992	0.3000
S	0.2734	0.1925	0.1221	0.0858
T	0.2705	0.1913	0.1210	0.0856
$\widehat{\beta}_{WA}$	2347.4312	2347.6175	2346.9822	2346.9073
S	103.4943	72.5992	45.9857	32.2951
T	101.9376	72.0808	45.5879	32.2355
$\widehat{B}_{W,10}$	637.4767	631.0271	627.1464	626.0630
S	80.6346	56.6498	35.8977	24.8922
T	74.1278	53.1234	33.5982	23.7575
n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_A$	2.0195	2.0195	2.0222	2.0207
S	0.1132	0.1119	0.1122	0.1128
T	0.1029	0.1029	0.1029	0.1029
$\widehat{\alpha}_{WA}$	0.3199	0.2720	0.3063	0.2917
S	0.2460	0.3290	0.2002	0.2448
T	0.2465	0.3260	0.1998	0.2400
$\widehat{\beta}_{WA}$	2337.4794	2355.8518	2343.9445	2349.5583
S	108.8290	109.9314	83.4924	84.8016
T	108.9759	108.9759	83.2317	83.2317
$\widehat{B}_{W,10}$	628.2301	631.2518	631.7653	631.8486
S	77.6324	52.1920	64.4974	53.0745
T	75.8297	49.0286	61.6288	49.9675

Table 5.6: Fitting G_{WA} to G_{WA} for $k = 2$, $B_A = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

affected by how we allocate the overall sample, and generally, the more observations we have at the lower stresses, the more accurate this function becomes as the standard error decreases. However, in practice, we will need to strike a balance between the number of observations we have at the lower stress level, and the length of time the experiment takes to run.

5.2.4 Fitting G_{WA}

We consider a similar scenario for the Weibull Arrhenius combination. Again, we report results only for $B_A = 2$, although results for the remaining values of the shape parameter show similar patterns, with $k = 2$, 3 and 4. The results for these are summarised in Tables 5.6, 5.7 and 5.8 respectively. This time, we compare the sample means for $B_{W,10}$ with a true value of 624.8256.

We observe similar outcomes to those observed with G_{WP}. Across all stress levels, we see an improved agreement between MLEs and true values as the overall sample size is

n_1, n_2 / n_3	25,25 / 25	50,50 / 50	100,100 / 100	300,300 / 300	500,500 / 500
$\widehat{B}_A$	2.0530	2.0264	2.0130	2.0048	2.0025
S	0.1907	0.1310	0.0921	0.0518	0.0401
T	0.1680	0.1188	0.0840	0.0485	0.0376
$\widehat{\alpha}_{WA}$	0.3049	0.3027	0.3009	0.3010	0.2996
S	0.3610	0.2527	0.1778	0.1032	0.0798
T	0.3552	0.2512	0.1776	0.1025	0.0794
$\widehat{\beta}_{WA}$	2343.5976	2344.8688	2346.3270	2346.3993	2347.0150
S	140.8301	98.5410	69.3635	40.3990	31.1881
T	138.6989	98.0749	69.3495	40.0389	31.0140
$\widehat{B}_{W,10}$	641.1912	632.3986	629.0613	626.1164	625.6107
S	98.3663	68.9104	48.6823	28.1455	21.7993
T	93.1172	65.8438	46.5586	26.8806	20.8216

n_1, n_2 / n_3	50,100 / 150	25,200 / 75	25,25 / 250	200,50 / 50
$\widehat{B}_A$	2.0132	2.0133	2.0143	2.0137
S	0.0909	0.0918	0.0909	0.0902
T	0.0840	0.0840	0.0840	0.0840
$\widehat{\alpha}_{WA}$	0.3035	0.2848	0.2414	0.2948
S	0.1999	0.2848	0.2414	0.2002
T	0.1987	0.2796	0.2357	0.2001
$\widehat{\beta}_{WA}$	2344.9974	2340.5408	2340.1997	2348.5658
S	83.5548	119.7570	108.3593	70.4635
T	83.2302	117.7030	105.4980	70.4134
$\widehat{B}_{W,10}$	628.9102	628.3078	628.3623	629.4926
S	56.6169	68.8487	73.4151	43.6691
T	54.9008	66.7847	70.0641	41.3968

Table 5.7: Fitting G_{WA} to G_{WA} for $k = 3$, $B_A = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 25,25	50,50 50,50	100,100 100,100	200,200 200,200	500,500 500,500	25,25, 75,75
$\widehat{B}_A$	2.0428	2.0193	2.0088	2.0053	2.0016	2.0218
S	0.1634	0.1116	0.0787	0.0551	0.0348	0.1126
T	0.1455	0.1029	0.0728	0.0514	0.0325	0.1029
$\widehat{\alpha}_{WA}$	0.3054	0.3075	0.3031	0.3021	0.2993	0.3147
S	0.3246	0.2311	0.1617	0.1158	0.0719	0.2641
T	0.3211	0.2271	0.1606	0.1135	0.0718	0.2657
$\widehat{\beta}_{WA}$	2343.7594	2343.0682	2345.2921	2346.0596	2347.1858	2340.1936
S	131.0662	93.0792	65.1852	46.7073	29.0391	113.8765
T	129.6222	91.6568	64.8111	45.8284	28.9844	114.2882
$\widehat{B}_{W,10}$	638.8891	629.6424	627.0709	626.4082	625.4034	630.8959
S	91.3794	62.9155	44.5901	31.4284	20.1260	76.5852
T	86.2528	60.9900	43.1264	30.4950	19.2867	74.3920
n_1, n_2 n_3, n_4	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_A$	2.0185	2.0193	2.0186	2.0188	2.0204	2.0203
S	0.1119	0.1109	0.1110	0.1128	0.1118	0.1125
T	0.1029	0.1029	0.1029	0.1029	0.1029	0.1029
$\widehat{\alpha}_{WA}$	0.3010	0.3173	0.3021	0.2958	0.3155	0.3132
S	0.2328	0.2767	0.2145	0.2276	0.2619	0.2927
T	0.2306	0.2742	0.2135	0.2268	0.2605	0.2911
$\widehat{\beta}_{WA}$	2345.7314	2338.8762	2345.5608	2347.8035	2339.7362	2340.6945
S	88.9141	117.1067	83.0034	83.9079	113.3066	123.0327
T	87.9890	116.0506	82.5293	83.5440	112.4357	122.1216
$\widehat{B}_{W,10}$	630.0202	629.0101	630.3646	630.6534	630.1513	630.4483
S	58.7288	74.7967	57.8297	56.4989	79.9088	76.4958
T	55.7990	72.9099	55.9419	52.9770	74.3177	73.7202

Table 5.8: Fitting G_{WA} to G_{WA} for $k = 4$, $B_A = 2$. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n	Non-accelerated	Accelerated			
		k	n_i	Log-linear	Arrhenius
75	4.3323	3	25	5.0040	4.8374
100	3.7519	2	50	4.0878	3.8509
		4	25	4.6072	4.4808
150	3.0634	3	50	3.5384	3.4205
200	2.6530	2	100	2.8905	2.7597
		4	50	3.2578	3.1684
300	2.1661	3	100	2.5020	2.4187
400	1.8759	4	100	2.3036	2.2404
500	1.6779	2	250	1.8281	1.7454
800	1.3265	4	200	1.6289	1.5842
900	1.2506	3	300	1.4445	1.3964
1000	1.1864	2	500	1.2927	1.2342
1500	0.9687	3	500	1.1189	1.0817
2000	0.8389	4	500	1.0302	1.0019

Table 5.9: The theoretical standard errors of $\widehat{B}_{10}$ from accelerated and non-accelerated Weibull data for varying n and k (at design stress). Calculations use a shape parameter of 2 and a scale parameter of 100.

increased. When we begin to vary n_i, we observe smaller standard errors for the MLEs and $\widehat{B}_{W,10}$ if we have more observations at the lower stress levels.

5.2.5 Comparison between accelerated and non-accelerated Weibull distributions

We compare theoretical standard errors of $\widehat{B}_{10}$ from accelerated and non-accelerated Weibull distributions; we have observed throughout, that the agreement between theoretical and sample results for the Weibull distribution (both in the accelerated and non-accelerated case) always match up very well, even for small sample sizes. Thus, we compare theoretical quantities, making the necessary adjustments to give a scale of 100 at design stress, and taking $\theta = 100$ in the non-accelerated case. Table 5.9 summarises the theoretical standard errors for the non-accelerated Weibull distribution, and the accelerated Weibull model with both Log-linear and Arrhenius relationships. We list the results for $k = 2$, 3 and 4, and for equal n_i. We compute the corresponding non-accelerated counterpart for equivalent values of n. So, for example, if we take $n = 100$, then we compare the theoretical standard error of $\widehat{B}_{10}$ for the non-accelerated case, with the equivalent value for $k = 2$, $n_i = 50$, and $k = 4$, $n_i = 25$; the n_i are as used in our simulations. Due to the extensive number of ways we can allocate the n_i, we do not include results for different loadings. When we have an equal allocation of items at each stress level, we always observe larger theoretical standard errors of $\widehat{B}_{10}$ for the accelerated case, although figures for the Arrhenius relationship are slightly smaller than the Log-linear counterparts. Thus, if we accelerate an experiment, and keep the allocation of items equal at each stress level, then we pay the penalty by observing larger standard errors of $\widehat{B}_{10}$. There are a few exceptions to this; for example, if we take $k = 2$, $n = 200$ and $n_1 = 175$, then the theoretical standard error of this estimate with an Arrhenius relationship is slightly smaller than its non-accelerated counterpart. However, for the majority of cases considered above with unequal allocations, we observe a rise in the theoretical standard error.

5.3 Fitting G_{B*}

As with previous work, we now present the theory necessary to fit the accelerated Burr distribution to a set of data, and derive the EFI matrix of the Burr MLEs. As above, the theory covers both Log-linear and Arrhenius models simultaneously. Unlike the Weibull distribution, however, we do not present results from simulations in this section to verify our theory, since these will be considered when we examine the effects of mis-specification in the next chapter. Now, y_{ij} represents the j^{th} observation from the i^{th} stress level, for which the pdf is

$$g_{B*}\left(y_{ij};\tau,a,\phi_{i*}\right)=\frac{a\tau y_{ij}^{\tau-1}}{\phi_{i*}^{\tau}}\left\{1+\left(\frac{y_{ij}}{\phi_{i*}}\right)^{\tau}\right\}^{-(a+1)},$$

for $*=A$ or P, with

$$\phi_{i*}=\exp\left\{\alpha_B+\beta_B\rho\left(X_i\right)\right\}.$$

Here, the shape parameters a and τ, remain constant across stress levels, ϕ_{i*} is the scale parameter at stress X_i and α_B and β_B are the parameters of the Log-linear or Arrhenius model linking stress X_i with ϕ_{i*}. Assuming y_{ij} form a complete set of observations, the likelihood and log-likelihood are given by

$$L_B\left(\tau,a,\alpha_B,\beta_B\right)=\prod_{i=1}^{k}\prod_{j=1}^{n_i}\frac{a\tau y_{ij}^{\tau-1}}{\phi_{i*}^{\tau}}\left\{1+\left(\frac{y_{ij}}{\phi_{i*}}\right)^{\tau}\right\}^{-(a+1)},$$

and

$$\begin{aligned} l_B &= n\ln a+n\ln\tau+(\tau-1)S_e-n\tau\alpha_B-\tau\beta_B\sum_{i=1}^{k}n_i\rho\left(X_i\right) \\ &\quad -(a+1)F\left(\tau,\alpha_B,\beta_B\right), \end{aligned} \tag{5.13}$$

where

$$F\left(\tau,\alpha_B,\beta_B\right)=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln\left[1+y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_i\right)\right\}\right].$$

To evaluate derivatives of (5.13), we note

$$\begin{aligned} F_{1,0,0} &= \frac{\partial F(\tau, \alpha_B, \beta_B)}{\partial \tau} \\ &= \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{y_{ij}^{\tau} \exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}\{\ln y_{ij} - \alpha_B - \beta_B\rho(X_i)\}}{1 + y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}, \end{aligned}$$

$$\begin{aligned} F_{0,1,0} &= \frac{\partial F(\tau, \alpha_B, \beta_B)}{\partial \alpha_B} \\ &= -\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\tau y_{ij}^{\tau} \exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}{1 + y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}, \end{aligned}$$

and

$$\begin{aligned} F_{0,0,1} &= \frac{\partial F(\tau, \alpha_B, \beta_B)}{\partial \beta_B} \\ &= -\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\tau \rho(X_i) y_{ij}^{\tau} \exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}{1 + y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}. \end{aligned}$$

We can now evaluate the four score functions. These are given by

$$\frac{\partial l_B}{\partial a} = \frac{n}{a} - F(\tau, \alpha_B, \beta_B), \tag{5.14}$$

$$\frac{\partial l_B}{\partial \tau} = \frac{n}{\tau} + S_e - n\alpha_B - \beta_B \sum_{i=1}^{k} n_i \rho(X_i) - (a+1) F_{1,0,0}, \tag{5.15}$$

$$\frac{\partial l_B}{\partial \alpha_B} = -n\tau - (a+1) F_{0,1,0}, \tag{5.16}$$

and

$$\frac{\partial l_B}{\partial \beta_B} = -\tau \sum_{i=1}^{k} n_i \rho(X_i) - (a+1) F_{0,0,1}. \tag{5.17}$$

For later use, we also evaluate second derivatives required for the variance covariance matrix of the Burr MLEs. We write

$$\begin{aligned} F_{2,0,0} &= \frac{\partial F_{1,0,0}}{\partial \tau} \\ &= \sum_{i=1}^{k}\sum_{i=1}^{n_i} \frac{y_{ij}^{\tau} \exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}\{\ln y_{ij} - \alpha_B - \beta_B\rho(X_i)\}^2}{\left[1 + y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}\right]^2}, \end{aligned}$$

$$
\begin{aligned}
F_{0,2,0} &= \frac{\partial F_{0,1,0}}{\partial \alpha_B} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\tau^2 y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\alpha\beta_B \rho\left(X_i\right)\right\}}{\left[1 + y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\tau\beta_B \rho\left(X_i\right)\right\}\right]^2},
\end{aligned}
$$

$$
\begin{aligned}
F_{0,0,2} &= \frac{\partial F_{0,0,1}}{\partial \beta_B} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\tau^2 \rho\left(X_i\right)^2 y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left(-\tau\beta_B \rho\left(X_i\right)\right)}{\left[1 + y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left(-\tau\beta_B \rho\left(X_i\right)\right)\right]^2},
\end{aligned}
$$

$$
\begin{aligned}
F_{1,1,0} &= \frac{\partial F_{1,0,0}}{\partial \alpha_B} \\
&= -\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\tau\beta_B \rho\left(X_i\right)\right\}}{\left[1 + y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\tau\beta_B \rho\left(X_i\right)\right\}\right]^2} \\
&\quad \times \left[1 + y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\tau\beta_B \rho\left(X_i\right)\right\} + \tau\left\{\ln y_{ij} - \alpha_B - \beta_B \rho\left(X_i\right)\right\}\right],
\end{aligned}
$$

$$
\begin{aligned}
F_{1,0,1} &= \frac{\partial F_{1,0,0}}{\partial \beta_B} \\
&= -\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\rho\left(X_i\right) y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\tau\beta_B \rho\left(X_i\right)\right\}}{\left[1 + y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\tau\beta_B \rho\left(X_i\right)\right\}\right]^2} \\
&\quad \times \left[1 + y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\tau\beta_B \rho\left(X_i\right)\right\} + \tau\left\{\ln y_{ij} - \alpha_B - \beta_B \rho\left(X_i\right)\right\}\right],
\end{aligned}
$$

and, finally,

$$
\begin{aligned}
F_{0,1,1} &= \frac{\partial F_{0,0,1}}{\partial \alpha_B} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\tau^2 \rho\left(X_i\right) y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\tau\beta_B \rho\left(X_i\right)\right\}}{\left[1 + y_{ij}^{\tau} \exp\left(-\tau\alpha_B\right) \exp\left\{-\tau\beta_B \rho\left(X_i\right)\right\}\right]^2}.
\end{aligned}
$$

Using (5.14), (5.15), (5.16) and (5.17), we now list the second derivatives for the accelerated Burr distribution as follows.

$$
\frac{\partial^2 l_B}{\partial a^2} = -\frac{n}{a^2},
$$

$$
\frac{\partial^2 l_B}{\partial \tau^2} = -\frac{n}{\tau^2} - (a+1)\, F_{2,0,0},
$$

$$\frac{\partial^2 l_B}{\partial \alpha_B^2} = -(a+1)\,F_{0,2,0},$$

$$\frac{\partial^2 l_B}{\partial \beta_B^2} = -(a+1)\,F_{0,0,2},$$

$$\frac{\partial^2 l_B}{\partial a \partial \tau} = -F_{1,0,0},$$

$$\frac{\partial^2 l_B}{\partial a \partial \alpha_B} = -F_{0,1,0},$$

$$\frac{\partial^2 l_B}{\partial a \partial \beta_B} = -F_{0,0,1},$$

$$\frac{\partial^2 l_B}{\partial \tau \partial \alpha_B} = -n - (a+1)\,F_{1,1,0},$$

$$\frac{\partial^2 l_B}{\partial \tau \partial \beta_B} = -\sum_{i=1}^{k} n_i \rho(X_i) - (a+1)\,F_{1,0,1},$$

and

$$\frac{\partial^2 l_B}{\partial \alpha_B \partial \beta_B} = -(a+1)\,F_{0,1,1}.$$

As previously, we can make limited algebraical progress, and reduce the number of model parameters under active consideration by one. Here, we can equate (5.14) to zero to obtain

$$a = \frac{n}{F(\tau, \alpha_B, \beta_B)},$$

insert this into (5.13), and derive the profile log-likelihood as

$$\begin{aligned} l_B^+ &= n \ln\left\{\frac{n}{F(\tau, \alpha_B, \beta_B)}\right\} + n \ln \tau + (\tau - 1)\,S_e - n\tau\alpha_B \\ &\quad -\tau\beta_B \sum_{i=1}^{k} n_i \rho(X_i) - F(\tau, \alpha_B, \beta_B). \end{aligned}$$

The three profile score functions for τ, α_B and β_B are then

$$\frac{\partial l_B^+}{\partial \tau} = \frac{-nF_{1,0,0}}{F(\tau, \alpha_B, \beta_B)} + \frac{n}{\tau} + S_e - n\alpha_B - \beta_B \sum_{i=1}^{k} n_i \rho(X_i) - F_{1,0,0},$$

$$\frac{\partial l_B^+}{\partial \alpha_B} = \frac{-nF_{0,1,0}}{F(\tau, \alpha_B, \beta_B)} - n\tau - F_{0,1,0},$$

and

$$\frac{\partial l_B^+}{\partial \beta_B} = \frac{-nF_{0,0,1}}{F(\tau, \alpha_B, \beta_B)} - \tau \sum_{i=1}^{k} n_i \rho(X_i) - F_{0,0,1}.$$

To fit the accelerated Burr distribution to a set of data, we will use the Newton-Raphson method to iterate on the three parameters until they converge onto their MLEs. Thus, we require profile second derivatives, which, for completeness, we list below.

$$\frac{\partial^2 l_B^+}{\partial \tau^2} = -n \left\{ \frac{F(\tau, \alpha_B, \beta_B) F_{2,0,0} - F_{1,0,0}^2}{F(\tau, \alpha_B, \beta_B)^2} \right\} - \frac{n}{\tau^2} - F_{2,0,0},$$

$$\frac{\partial^2 l_B^+}{\partial \alpha_B^2} = -n \left\{ \frac{F(\tau, \alpha_B, \beta_B) F_{0,2,0} - F_{0,1,0}^2}{F(\tau, \alpha_B, \beta_B)^2} \right\} - F_{0,2,0},$$

$$\frac{\partial^2 l_B^+}{\partial \beta_B^2} = -n \left\{ \frac{F(\tau, \alpha_B, \beta_B) F_{0,0,2} - F_{0,0,1}^2}{F(\tau, \alpha_B, \beta_B)^2} \right\} - F_{0,0,2},$$

$$\frac{\partial^2 l_B^+}{\partial \tau \partial \alpha_B} = -n \left\{ \frac{F(\tau, \alpha_B, \beta_B) F_{1,1,0} - F_{0,1,0} F_{1,0,0}}{F(\tau, \alpha_B, \beta_B)^2} \right\} - n - F_{1,1,0},$$

$$\frac{\partial^2 l_B^+}{\partial \tau \partial \beta_B} = -n \left\{ \frac{F(\tau, \alpha_B, \beta_B) F_{1,0,1} - F_{0,0,1} F_{1,0,0}}{F(\tau, \alpha_B, \beta_B)^2} \right\} - \sum_{i=1}^{k} n_i \rho(X_i) - F_{1,0,1},$$

and, finally,

$$\frac{\partial^2 l_B^+}{\partial \alpha_B \partial \beta_B} = -n \left\{ \frac{F(\tau, \alpha_B, \beta_B) F_{0,1,1} - F_{0,1,0} F_{0,0,1}}{F(\tau, \alpha_B, \beta_B)^2} \right\} - F_{0,1,1}.$$

MLEs may be obtained in the usual way, provided sensible initial starting values are chosen for the three parameters. In the next section, we consider how to fit an accelerated Burr distribution to a set of data, and in keeping with the non-accelerated scenario, derive a discriminating Δ to determine whether the Burr or Weibull distributions provide a better fit. This will be carried out for a general scale-stress relationship, with the Log-linear and Arrhenius forms as special cases.

5.3.1 Fitting the accelerated Burr distribution and Δ

In this section, we derive the form of Δ for a general scale-stress relationship linking the scale parameters from the Weibull and Burr distributions with stress. The work is a generalisation of Johnson (2003), who presents results for the Log-linear relationship. Given the form of many scale-stress relationships listed in Section 5.1.1, we write a general relationship as

$$\phi_i = \phi_s h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right) = Qf\left(X_s\right) h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right),$$

for $m = 1, 2, .., p$, where Q is a constant independent of stress, $f\left(X_s\right)$ is a function of design stress and $h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right)$ a function of design and accelerated stress levels and the parameters β_{Bm}, with the property that

$$h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right) = 1 \text{ for } X_i = X_s,$$

so that the scale ϕ_s at design stress is $Qf\left(X_s\right)$. Thus, for example, we write the Log-linear and Arrhenius relationships as

$$\phi_{i*} = \exp\left(\alpha_B\right) \exp\left\{\beta_B \rho\left(X_s\right)\right\} \exp\left[\beta_B\left\{\rho\left(X_i\right) - \rho\left(X_s\right)\right\}\right],$$

giving $m = 1$ and

$$\begin{aligned} Q &= \exp\left(\alpha_B\right), \\ f\left(X_s\right) &= \exp\left\{\beta_B \rho\left(X_s\right)\right\}, \\ h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right) &= \exp\left[\beta_B\left\{\rho\left(X_i\right) - \rho\left(X_s\right)\right\}\right]; \end{aligned}$$

this argument extends to more complex relationships such as the general polynomial. We refer to Appendix D for an outline of the proof for obtaining the form of the discriminating function Δ, given by

$$\Delta = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left\{ \frac{y_{ij}}{f(X_s) h\left(X_i, X_s, \widehat{\beta}_{W1}, ..., \widehat{\beta}_{Wm}\right)} \right\}^{2\widehat{B}}}{2} - \frac{\left[\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left\{ \frac{y_{ij}}{f(X_s) h\left(X_i, X_s, \widehat{\beta}_{W1}, ..., \widehat{\beta}_{Wm}\right)} \right\}^{\widehat{B}} \right]^2}{n}. \tag{5.18}$$

Thus, we have

$$\Delta_* = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{\exp\{\widehat{\beta}_{W*}\rho(X_i)\}}\right\}^{2\widehat{B}_*}}{2} - \frac{\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{\exp\{\widehat{\beta}_{W*}\rho(X_i)\}}\right\}^{\widehat{B}_*}\right]^2}{n}.$$

We next consider fitting an accelerated Burr distribution to a set of data; the key is to start by fitting the corresponding Weibull distribution with the same scale-stress relationship (here, either Log-linear or Arrhenius). We refer to Watkins and Johnson (1999), which outlines fitting the three parameter accelerated Burr Log-linear distribution, and now adapt that approach for both Log-linear and Arrhenius models. We list the main points below.

- We first fit the accelerated Weibull distribution to the set of data using the profile log-likelihood given by (5.8). For real life data, this will usually involve fitting the Weibull distribution to each subset of data, and then use the MLEs from these to obtain initial starting values for the MLEs from the accelerated model. When we consider simulated data from a Burr distribution with known parameters, initial starting values for the Weibull MLEs can be obtained relatively easily from these, and we set $B_* = \tau$ and $\beta_{W*} = \beta_B$. Alternatively, we can maximise the entropy function for the accelerated Weibull distribution, and use entropy values as initial estimates.

- We next rescale the subsets of data by their appropriate scale estimates from the Weibull distribution. Thus, the data becomes

$$y_{ij} \longrightarrow \frac{y_{ij}}{\exp\left\{\widehat{\alpha}_{W*} + \widehat{\beta}_{W*}\rho\left(X_i\right)\right\}}$$

  This rescaling effectively removes the effects of the stress factor, with all data centered around one; the notion of acceleration is largely reduced.

- Since the Burr distribution is the limiting distribution of the Weibull, then we use the appropriate discriminating Δ_*, which takes into account the small amount of remaining acceleration, to determine which distribution function provides the better fit. If $\Delta_* > 0$, then we proceed to the next stage of fitting the accelerated Burr distribution.

- This stage of the algorithm involves fitting the non-accelerated Burr distribution to the rescaled y_{ij}, following Watkins (1999), as outlined in Chapter 2. This is equivalent to fitting the accelerated model with $\beta_B = 0$ to the rescaled data, and leads to estimates of the single scale parameter ϕ ($\simeq 1$) and shape parameters a and τ. These MLEs then provide us with suitable starting values to fit the accelerated Burr distribution. We then move to the final stage of the algorithm.

- We fit the accelerated Burr distribution using the profile log-likelihood to the rescaled data, with starting values based on estimates found in the previous stage. Thus, we start with τ in the accelerated model equal to the same shape parameter from the non-accelerated case, $\beta_B = 0$ and $\alpha_B = \ln(\phi)$. We then reverse the effects of rescaling, noting that estimates of the two shape parameters a and τ are not affected by scaling, and so remain the same. The scale estimates are found by adding $\widehat{\alpha}_{W*}$ and $\widehat{\beta}_{W*}$ to the estimates of α_B and β_B respectively.

We see that the above algorithm accommodates both Log-linear and Arrhenius relationships, and further generalises. Note that the limiting link between Weibull and Burr distributions means that Δ can only be calculated when using the two distributions with the same scale-stress relationship.

5.3.2 The EFI matrix of the Burr MLEs

In this section, we derive the asymptotic variance covariance matrix of the Burr MLEs for a set of accelerated observations, and consider both the Arrhenius and Log-linear models simultaneously. Such results will enable us to compare theoretical and simulated values, and also deduce how accurate the MLEs are for various sample sizes, stress levels and parameter values. The derivation of expected values of such derivatives will be simplified considerably if we first obtain the expectations of the functions that make up these derivatives. For instance, we consider

$$E\left[F_{1,0,0}\right] = E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_i\right)\right\}\left\{\ln Y_{ij}-\alpha_B-\beta_B\rho\left(X_i\right)\right\}}{1+Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_i\right)\right\}}\right].$$

To derive such expectations, we convert back to ϕ_{i*} and use work on the non-accelerated Burr distribution. Thus,

$$E\left[F_{1,0,0}\right] = \sum_{i=1}^{k}\sum_{j=1}^{n_i} E\left[\frac{\left(\frac{Y_{ij}}{\phi_{i*}}\right)^{\tau}\ln\left(\frac{Y_{ij}}{\phi_{i*}}\right)}{1+\left(\frac{Y_{ij}}{\phi_{i*}}\right)^{\tau}}\right].$$

For stress level i, the Y_{ij} are independently and identically distributed. Therefore, this expectation becomes

$$\sum_{i=1}^{k} n_i E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\ln\left(\frac{Y_i}{\phi_{i*}}\right)}{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}\right],$$

where Y_i is a random variable from $G_{B*}(\tau, a, \phi_{i*})$. We now use (3.7) to write

$$E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\ln\left(\frac{Y_i}{\phi_{i*}}\right)}{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}\right]=\frac{1-\gamma-\Psi(a)}{\tau(a+1)},$$

and hence,

$$E[F_{1,0,0}]=\frac{n\{1-\gamma-\Psi(a)\}}{\tau(a+1)}.$$

Similar procedures are used for eight other expectations, which we now list.

$$\begin{aligned}
E[F_{0,1,0}] &= E\left[-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\tau Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}{1+Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}\right]\\
&= -\tau\sum_{i=1}^{k}n_iE\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}\right]=-\frac{n\tau}{a+1},
\end{aligned}$$

on using (3.9); while

$$\begin{aligned}
E[F_{0,0,1}] &= E\left[-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\tau\rho(X_i)Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}{1+Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}\right]\\
&= -\tau\sum_{i=1}^{k}n_i\rho(X_i)E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}\right]=-\frac{\tau}{a+1}\sum_{i=1}^{k}n_i\rho(X_i).
\end{aligned}$$

Next, we compute

$$\begin{aligned}
E[F_{2,0,0}] &= E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}\{\ln Y_{ij}-\alpha_B-\beta_B\rho(X_i)\}^2}{\left[1+Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}\right]^2}\right]\\
&= \sum_{i=1}^{k}n_iE\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\left\{\ln\left(\frac{Y_i}{\phi_{i*}}\right)\right\}^2}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right]\\
&= \frac{na\left\{\frac{\pi^2}{6}+\gamma^2-2\gamma+2(\gamma-1)\Psi(a+1)+\Psi(a+1)^2+\Psi'(a+1)\right\}}{\tau^2(a+1)(a+2)},
\end{aligned}$$

from (3.8). Next, we look at

$$\begin{aligned} E\left[F_{0,2,0}\right] &= E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\tau^2 Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_i\right)\right\}}{\left[1+Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_i\right)\right\}\right]^2}\right] \\ &= \tau^2\sum_{i=1}^{k}n_i E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right], \end{aligned}$$

which, on using (3.10) becomes

$$\frac{na\tau^2}{(a+1)(a+2)};$$

we now examine

$$\begin{aligned} E\left[F_{0,0,2}\right] &= E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\tau^2\rho\left(X_i\right)^2 Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_i\right)\right\}}{\left[1+Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_i\right)\right\}\right]^2}\right] \\ &= \tau^2\sum_{i=1}^{k}n_i\rho\left(X_i\right)^2 E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right] \\ &= \frac{a\tau^2}{(a+1)(a+2)}\sum_{i=1}^{k}n_i\rho\left(X_i\right)^2. \end{aligned}$$

Next, we consider

$$\begin{aligned} E\left[F_{1,1,0}\right] &= E\left[\begin{array}{c} -\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}{\left[1+Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}\right]^2} \\ \times\left[1+Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_i\right)\right\}+\tau\left\{\ln Y_{ij}-\alpha_B-\beta_B\rho\left(X_i\right)\right\}\right] \end{array}\right] \\ &= -\sum_{i=1}^{k}n_i\left(\begin{array}{c} E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right]+E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{2\tau}}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right]+ \\ \tau E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\ln\left(\frac{Y_i}{\phi_{i*}}\right)}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right] \end{array}\right). \end{aligned}$$

We note that

$$E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{2\tau}}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right]=\frac{2}{(a+1)(a+2)},$$

and that

$$E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\ln\left(\frac{Y_i}{\phi_{i*}}\right)}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right]$$

is given by (3.11). Thus,

$$E\left[F_{1,1,0}\right]=\frac{-n\left\{2\left(a+1\right)-a\gamma-a\Psi\left(a+1\right)\right\}}{\left(a+1\right)\left(a+2\right)}.$$

Now, we examine

$$\begin{aligned}
E\left[F_{1,0,1}\right] &= E\left[\begin{array}{c} -\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\rho(X_i)Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}}{\left[1+Y_{ij}^{\tau}\exp(-\tau\alpha_B)\exp\{-\tau\beta_B\rho(X_i)\}\right]^2} \\ \times\left[1+Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_i\right)\right\}+\tau\left\{\ln Y_{ij}-\alpha_B-\beta_B\rho\left(X_i\right)\right\}\right]\end{array}\right] \\
&= -\sum_{i=1}^{k}n_iX_i\left(\begin{array}{c} E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right]+E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{2\tau}}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right]+ \\ \tau E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\ln\left(\frac{Y_i}{\phi_{i*}}\right)}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right]\end{array}\right) \\
&= \frac{\left\{a\gamma+a\Psi\left(a+1\right)-2\left(a+1\right)\right\}}{\left(a+1\right)\left(a+2\right)}\sum_{i=1}^{k}n_i\rho\left(X_i\right).
\end{aligned}$$

Finally, we compute

$$\begin{aligned}
E\left[F_{0,1,1}\right] &= E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\tau^2\rho\left(X_i\right)Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left(-\tau\beta_B\rho\left(X_i\right)\right)}{\left\{1+Y_{ij}^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left(-\tau\beta_B\rho\left(X_i\right)\right)\right\}^2}\right] \\
&= \tau^2\sum_{i=1}^{k}n_i\rho\left(X_i\right)E\left[\frac{\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}}{\left\{1+\left(\frac{Y_i}{\phi_{i*}}\right)^{\tau}\right\}^2}\right] \\
&= \frac{a\tau^2}{\left(a+1\right)\left(a+2\right)}\sum_{i=1}^{k}n_i\rho\left(X_i\right).
\end{aligned}$$

The expected values of the second derivatives are

$$-E\left[\frac{\partial^2 l_B}{\partial a^2}\right]=E\left[\frac{n}{a^2}\right]=\frac{n}{a^2},$$

$$
\begin{aligned}
-E\left[\frac{\partial^2 l_B}{\partial \tau^2}\right] &= \frac{n}{\tau^2} + (a+1)\,E\left[F_{2,0,0}\right] \\
&= \frac{n}{\tau^2}\left[1 + \frac{a\left\{\frac{\pi^2}{6} + \gamma^2 - 2\gamma + 2(\gamma-1)\,\Psi(a+1) + \Psi(a+1)^2 + \Psi'(a+1)\right\}}{(a+2)}\right],
\end{aligned}
$$

$$
-E\left[\frac{\partial^2 l_B}{\partial \alpha_B^2}\right] = (a+1)\,E\left[F_{0,2,0}\right] = \frac{na\tau^2}{a+2},
$$

$$
-E\left[\frac{\partial^2 l_B}{\partial \beta_B^2}\right] = (a+1)\,E\left[F_{0,0,2}\right] = \frac{a\tau^2}{a+2}\sum_{i=1}^{k} n_i \rho\,(X_i)^2,
$$

$$
-E\left[\frac{\partial^2 l_B}{\partial a \partial \tau}\right] = E\left[F_{1,0,0}\right] = \frac{n\left\{1 - \gamma - \Psi(a)\right\}}{\tau\,(a+1)},
$$

$$
-E\left[\frac{\partial^2 l_B}{\partial a \partial \alpha_B}\right] = E\left[F_{0,1,0}\right] = -\frac{n\tau}{a+1},
$$

$$
-E\left[\frac{\partial^2 l_B}{\partial a \partial \beta_B}\right] = E\left[F_{0,0,1}\right] = -\frac{\tau}{a+1}\sum_{i=1}^{k} n_i \rho\,(X_i),
$$

$$
-E\left[\frac{\partial^2 l_B}{\partial \tau \partial \alpha_B}\right] = n + (a+1)\,E\left[F_{1,1,0}\right] = -\frac{na}{a+2}\left\{1 + \gamma + \Psi(a+1)\right\},
$$

$$
\begin{aligned}
-E\left[\frac{\partial^2 l_B}{\partial \tau \partial \beta_B}\right] &= \sum_{i=1}^{k} n_i \rho\,(X_i) + (a+1)\,E\left[F_{1,0,1}\right] \\
&= -\frac{a}{a+2}\left\{1 + \gamma + \Psi(a+1)\right\}\sum_{i=1}^{k} n_i \rho\,(X_i),
\end{aligned}
$$

and, finally,

$$
-E\left[\frac{\partial^2 l_B}{\partial \alpha_B \partial \beta_B}\right] = (a+1)\,E\left[F_{0,1,1}\right] = \frac{a\tau^2}{a+2}\sum_{i=1}^{k} n_i \rho\,(X_i).
$$

We use these results to derive the asymptotic variance of $\widehat{B}_{B,10}$ in the next section.

5.3.3 The asymptotic variance of $\widehat{B}_{B,10}$

In this section, we derive the asymptotic variance of $\widehat{B}_{B,10}$, again for both Arrhenius and Log-linear models simultaneously. If ϕ_{s*} is the scale parameter of the Burr distribution at design stress X_s, then

$$\phi_{s*} = \exp\{\alpha_B + \beta_B \rho(X_s)\}.$$

Thus, at X_s, we have

$$\widehat{B}_{B,10} = \exp(\widehat{\alpha}_B)\exp\left\{\widehat{\beta}_B \rho(X_s)\right\}\left\{0.9^{\frac{-1}{\widehat{a}}} - 1\right\}^{\frac{1}{\widehat{\tau}}},$$

and this is, asymptotically, Normally distributed with mean $B_{B,10}$ and variance given by (3.2) with $c_{\pi} = \begin{pmatrix} c_{\tau} & c_a & c_{\alpha_B} & c_{\beta_B} \end{pmatrix}'$, where

$$c_{\tau} = \frac{-\exp(\alpha_B)\exp\{\beta_B \rho(X_s)\}\left\{0.9^{\frac{-1}{a}} - 1\right\}^{\frac{1}{\tau}} \ln\left(0.9^{\frac{-1}{a}} - 1\right)}{\tau^2},$$

$$c_a = \frac{\exp(\alpha_B)\exp\{\beta_B \rho(X_s)\}\left\{0.9^{\frac{-1}{a}} - 1\right\}^{\frac{1}{\tau}-1} 0.9^{\frac{-1}{a}} \ln 0.9}{\tau a^2},$$

$$c_{\alpha_B} = \exp(\alpha_B)\exp\{\beta_B \rho(X_s)\}\left\{0.9^{\frac{-1}{a}} - 1\right\}^{\frac{1}{\tau}},$$

and

$$c_{\beta_B} = \rho(X_s)\exp(\alpha_B)\exp\{\beta_B \rho(X_s)\}\left\{0.9^{\frac{-1}{a}} - 1\right\}^{\frac{1}{\tau}}.$$

We check these results in later chapters, when we examine the effects of mis-specification.

5.4 Fitting G_{G*}

In the calculations below, we include how to fit both the Gamma Log-linear and Gamma Arrhenius distributions at once. Using our usual notation, the pdf of the Gamma distribution is $G_{G*}(y_{ij}; \tau, a_{i*})$ for $* = A$ or P, with

$$a_{iP} = \exp(\alpha_G + \beta_G X_i)$$

if we are using the Log-linear relationship, and

$$a_{iA} = \exp\left(\alpha_G + \frac{\beta_G}{X_i + c}\right)$$

for the Arrhenius model. The likelihood and log-likelihood are given by

$$L_G(\tau, \alpha_G, \beta_G) = \prod_{i=1}^{k}\prod_{j=1}^{n_i} \frac{y_{ij}^{\tau-1}\exp\left(-\frac{y_{ij}}{\exp(\alpha_G)\exp\{\beta_G\rho(X_i)\}}\right)}{\exp(\tau\alpha_G)\exp\{\tau\beta_G\rho(X_i)\}\Gamma(\tau)},$$

and

$$l_G = (\tau - 1) S_e - \exp(-\alpha_G) F(\beta_G) - n\tau\alpha_G - \tau\beta_G \sum_{i=1}^{k} n_i\rho(X_i) - n\ln\Gamma(\tau), \qquad (5.19)$$

where

$$F(\beta_G) = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}\exp\{-\beta_G\rho(X_i)\}.$$

The three score functions are given by

$$\begin{aligned} \frac{\partial l_G}{\partial \tau} &= S_e - n\alpha_G - \beta_G \sum_{i=1}^{k} n_i\rho(X_i) - n\Psi(\tau), \\ \frac{\partial l_G}{\partial \alpha_G} &= \exp(-\alpha_G) F(\beta_G) - n\tau, \end{aligned} \qquad (5.20)$$

and

$$\frac{\partial l_G}{\partial \beta_G} = -\exp(-\alpha_G) F_1(\beta_G) - \tau \sum_{i=1}^{k} n_i\rho(X_i),$$

where

$$F_m(\beta_G) = (-1)^m \sum_{i=1}^{k}\sum_{j=1}^{n_i} \rho(X_i)^m y_{ij}\exp\{-\beta_G\rho(X_i)\},$$

and, for future reference, the second partial derivatives are

$$\frac{\partial^2 l_G}{\partial \tau^2} = -n\Psi'(\tau),$$

$$\frac{\partial^2 l_G}{\partial \alpha_G^2} = -\exp(-\alpha_G) F(\beta_G),$$

$$\frac{\partial^2 l_G}{\partial \beta_G^2} = -\exp(-\alpha_G) F_2(\beta_G),$$

$$\frac{\partial^2 l_G}{\partial \tau \partial \alpha_G} = -n,$$

$$\frac{\partial^2 l_G}{\partial \tau \partial \beta_G} = -\sum_{i=1}^{k} n_i \rho(X_i),$$

and

$$\frac{\partial^2 l_G}{\partial \alpha_G \partial \beta_G} = \exp(-\alpha_G) F_1(\beta_G).$$

We can equate (5.20) to zero to obtain

$$\alpha_G = \ln\left\{\frac{F(\beta_G)}{n\tau}\right\},$$

and substitute this into (5.19) to derive the profile log-likelihood given by

$$l_G^+ = (\tau - 1) S_e - n\tau - n\tau \ln F(\beta_G) + n\tau \ln n\tau - \tau\beta_G \sum_{i=1}^{k} n_i \rho(X_i) - n \ln \Gamma(\tau).$$

The two profile score functions are then given by

$$\frac{\partial l_G^+}{\partial \tau} = S_e - n - n \ln F(\beta_G) + n(\ln n\tau + 1) - \beta_G \sum_{i=1}^{k} n_i \rho(X_i) - n\Psi(\tau),$$

and

$$\frac{\partial l_G^+}{\partial \beta_G} = \frac{-n\tau F_1(\beta_G)}{F(\beta_G)} - \tau \sum_{i=1}^{k} n_i \rho(X_i),$$

and, to use the Newton-Raphson process to obtain the roots of these profile functions, we also include, for completeness, second derivatives given by

$$\frac{\partial^2 l_G^+}{\partial \tau^2} = \frac{n}{\tau} - n\Psi'(\tau),$$

$$\frac{\partial^2 l_G^+}{\partial \beta_G} = -n\tau \left\{\frac{F(\beta_G) F_2(\beta_G) - F_1(\beta_G)^2}{F(\beta_G)^2}\right\},$$

and

$$\frac{\partial^2 l_G^+}{\partial\tau\partial\beta_G} = \frac{-nF_1(\beta_G)}{F(\beta_G)} - \sum_{i=1}^{k} n_i\rho(X_i).$$

So we now have all the functions to fit a Gamma distribution with either a Log-linear or Arrhenius relationship to an accelerated data set. We continue by deriving the EFI matrix for the Gamma MLEs.

5.4.1 The EFI matrix of the Gamma MLEs

Before deriving the EFI matrix of the Gamma MLEs, we first compute

$$\begin{aligned} E[F(\beta_G)] &= E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i} Y_{ij}\exp\{-\beta_G\rho(X_i)\}\right] \\ &= \sum_{i=1}^{k} n_i\exp\{-\beta_G\rho(X_i)\}E[Y_i], \end{aligned}$$

where Y_i is a random variable from $G_{G*}(y_{ij};\tau,\alpha_G,\beta_G)$. We therefore have

$$E[Y_i^m] = \frac{\exp\{\alpha_G+\beta_G\rho(X_i)\}^m\Gamma(m+\tau)}{\Gamma(\tau)}, \tag{5.21}$$

so

$$E[Y_i] = \tau\exp\{\alpha_G+\beta_G\rho(X_i)\},$$

and

$$E[F(\beta_G)] = n\tau\exp(\alpha_G).$$

Next, we consider

$$\begin{aligned} E[F_1(\beta_G)] &= E\left[-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\rho(X_i)Y_{ij}\exp\{-\beta_G\rho(X_i)\}\right] \\ &= -\sum_{i=1}^{k} n_i\rho(X_i)\exp\{-\beta_G\rho(X_i)\}E[Y_i] \\ &= -\tau\exp(\alpha_G)\sum_{i=1}^{k} n_i\rho(X_i), \end{aligned}$$

and, finally

$$\begin{aligned} E\left[F_2\left(\beta_G\right)\right] &= E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\rho\left(X_i\right)^2 Y_{ij}\exp\left\{-\beta_G\rho\left(X_i\right)\right\}\right] \\ &= \sum_{i=1}^{k} n_i\rho\left(X_i\right)^2\exp\left\{-\beta_G\rho\left(X_i\right)\right\}E\left[Y_i\right] \\ &= \tau\exp\left(\alpha_G\right)\sum_{i=1}^{k}n_i\rho\left(X_i\right)^2. \end{aligned}$$

Using these, the EFI matrix from the Gamma distribution is

$$A = \begin{bmatrix} n\Psi'\left(\tau\right) & & \\ n & n\tau & \\ \sum_{i=1}^{k} n_i\rho\left(X_i\right) & \tau\sum_{i=1}^{k} n_i\rho\left(X_i\right) & \tau\sum_{i=1}^{k} n_i\rho\left(X_i\right)^2 \end{bmatrix}.$$

We check these results in later chapters, when we begin to examine the effects of misspecification in accelerated data sets.

At this point we would usually proceed by considering the asymptotic variance of $\widehat{B}_{G,10}$. Due to the form of the Gamma cdf, we cannot write down a theoretical expression for this quantile. The function can only be computed, for various parameter values, using SAS or Mathematica. For instance, Mathematica not only has the ability to compute values for $\widehat{B}_{G,10}$, but also differentiate this function with respect to the distributional parameters, so that a value for the theoretical standard error can be computed. Thus, we may state that $B_{G,10}$ is given by

$$\texttt{InverseGammaRegularized}\left[\tau, 0, 0.1\right] * \exp\left\{\alpha_B + \beta_B\rho\left(X_s\right)\right\}.$$

$\widehat{B}_{G,10}$ will be Normally distributed with mean $B_{G,10}$ and variance (3.2), where numerical results for

$$c'_\pi = \left(\begin{array}{ccc} \frac{\partial B_{G,10}}{\partial\tau} & \frac{\partial B_{G,10}}{\partial\alpha_G} & \frac{\partial B_{G,10}}{\partial\beta_G} \end{array}\right),$$

are obtained using Mathematica.

5.5 Fitting G_{LN*}

We assume that the underlying distribution is now Lognormal, and derive the theory necessary to fit this to a data set when either the Arrhenius or Log-linear scale-stress relationship is used. The accelerated Lognormal pdf is $g_{LN}\left(y_{ij};\mu_i,\sigma\right)$ where

$$\mu_{i*} = \alpha_{LN} + \beta_{LN}\rho\left(X_i\right).$$

This parameterisation suggests that the mean of log life is a linear function of the stress level. We recall (using Nelson, 1982) that the parameter μ_{i*} really determines the scale of the data, whilst plots from Nelson (1982) show that σ influences the shape of the distribution function and typically ranges from 0.5 to 5. Thus, if we wish to have a set of data with values which lie around 2000, then we would have to set $\mu_{i*} = \ln 2000$. To see why this is so, recall the link between Lognormal and Normal distributions. Namely, that if Y has a Lognormal distribution with parameters μ_{i*} and σ, then $\ln Y$ has a Normal distribution with mean μ_{i*} and standard deviation σ. Thus, if we want an average value of around 2000, then we have to set μ_{i*} equal to the log of this number. This fact will be used in later sections when we begin to run simulations on the Lognormal distribution, and we have to chose sensible parameter values that mimic real life experiments. By substituting μ_{i*} in terms of α_{LN} and β_{LN}, the pdf of the Lognormal distribution becomes

$$g_{LN*}\left(y_{ij};\sigma,\alpha_{LN},\beta_{LN}\right) = \frac{1}{\sqrt{2\pi}\sigma y_{ij}} \exp\left[\frac{-\left\{\ln y_{ij} - \alpha_{LN} - \beta_{LN}\rho\left(X_i\right)\right\}^2}{2\sigma^2}\right].$$

The likelihood and log-likelihood are given by

$$L_{LN}\left(\sigma,\alpha_{LN},\beta_{LN}\right) = \prod_{i=1}^{k}\prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi}\sigma y_{ij}} \exp\left[\frac{-\left\{\ln y_{ij} - \alpha_{LN} - \beta_{LN}\rho\left(X_i\right)\right\}^2}{2\sigma^2}\right],$$

and so

$$l_{LN} = -n\ln\sqrt{2\pi} - n\ln\sigma - S_e - \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\ln y_{ij} - \alpha_{LN} - \beta_{LN}\rho\left(X_i\right)\right\}^2}{2\sigma^2}.$$

The three score functions are given by

$$\frac{\partial l_{LN}}{\partial \alpha_{LN}} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\ln y_{ij} - \alpha_{LN} - \beta_{LN}\rho\left(X_i\right)\right\}}{\sigma^2},$$

$$\frac{\partial l_{LN}}{\partial \beta_{LN}} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\rho\left(X_i\right)\left\{\ln y_{ij} - \alpha_{LN} - \beta_{LN}\rho\left(X_i\right)\right\}}{\sigma^2},$$

and

$$\frac{\partial l_{LN}}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\ln y_{ij} - \alpha_{LN} - \beta_{LN}\rho\left(X_i\right)\right\}^2}{\sigma^3}.$$

As previously, equating these score functions to zero and solving yields explicit parameter estimates. We note that the first two score functions effectively yield the Normal equations; see, for instance, Montgomery (1997). We can write the equations based on equating these derivatives to zero as

$$\begin{pmatrix} \sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln y_{ij} \\ \sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln y_{ij}\rho(X_i) \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^{k} n_i\rho(X_i) \\ \sum_{i=1}^{k} n_i\rho(X_i) & \sum_{i=1}^{k} n_i\rho(X_i)^2 \end{pmatrix} \begin{pmatrix} \alpha_{LN} \\ \beta_{LN} \end{pmatrix},$$

and then solve for α_{LN} and β_{LN}. Substituting these solutions into $\frac{\partial l_{LN}}{\partial\sigma} = 0$ will then yield the MLE for σ, so that all MLEs can be obtained explicitly.

5.5.1 The EFI matrix of the Lognormal MLEs

To obtain the asymptotic variance covariance matrix of the Lognormal MLEs, we consider the expectations of second derivatives below, and on differentiating the Lognormal score function, we have

$$\frac{\partial^2 l_{LN}}{\partial\sigma^2} = \frac{n}{\sigma^2} - \frac{3\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\ln y_{ij} - \alpha_{LN} - \beta_{LN}\rho(X_i)\right\}^2}{\sigma^4},$$

$$\frac{\partial^2 l_{LN}}{\partial\alpha_{LN}^2} = -\frac{n}{\sigma^2},$$

$$\frac{\partial^2 l_{LN}}{\partial\beta_{LN}^2} = -\frac{\sum_{i=1}^{k} n_i\rho(X_i)^2}{\sigma^2},$$

$$\frac{\partial^2 l_{LN}}{\partial\sigma\partial\alpha_{LN}} = \frac{-2\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\ln y_{ij} - \alpha_{LN} - \beta_{LN}\rho(X_i)\right\}}{\sigma^3},$$

$$\frac{\partial^2 l_{LN}}{\partial\sigma\partial\beta_{LN}} = \frac{-2\sum_{i=1}^{k}\sum_{j=1}^{n_i}\rho(X_i)\left\{\ln y_{ij} - \alpha_{LN} - \beta_{LN}\rho(X_i)\right\}}{\sigma^3},$$

and

$$\frac{\partial^2 l_{LN}}{\partial\alpha_{LN}\partial\beta_{LN}} = -\frac{\sum_{i=1}^{k} n_i\rho(X_i)}{\sigma^2}.$$

On taking expectations of these second derivatives, we see that we require expressions for $E[\ln Y_i]$ and $E\left[(\ln Y_i)^2\right]$. We again use our work on the non-accelerated Lognormal distribution to write

$$E[Y_i^m] = \exp\left[m\left\{\alpha_{LN} + \beta_{LN}\rho(X_i)\right\} + \frac{\sigma^2 m^2}{2}\right]. \tag{5.22}$$

Thus, on differentiating with respect to m, we have

$$E\left[Y_i^m \ln Y_i\right] = \left(\alpha_{LN} + \beta_{LN}\rho\left(X_i\right) + \sigma^2 m\right) \exp\left[m\left\{\alpha_{LN} + \beta_{LN}\rho\left(X_i\right)\right\} + \frac{\sigma^2 m^2}{2}\right], \quad (5.23)$$

so

$$E\left[\ln Y_i\right] = \alpha_{LN} + \beta_{LN}\rho\left(X_i\right).$$

By differentiating $E\left[Y_i^m\right]$ twice, we obtain an expression for

$$\begin{aligned} E\left[Y_i^m \left(\ln Y_i\right)^2\right] &= \exp\left[m\left\{\alpha_{LN} + \beta_{LN}\rho\left(X_i\right)\right\} + \frac{\sigma^2 m^2}{2}\right] \\ &\quad \times \left\{\sigma^2 + \left(\alpha_{LN} + \beta_{LN}\rho\left(X_i\right) + \sigma^2 m\right)^2\right\}, \end{aligned} \quad (5.24)$$

thus giving

$$E\left[\left(\ln Y_i\right)^2\right] = \sigma^2 + \left\{\alpha_{LN} + \beta_{LN}\rho\left(X_i\right)\right\}^2.$$

Using these expectations, we write the EFI matrix as

$$A = \sigma^{-2}\begin{bmatrix} 2n & & \\ 0 & n & \\ 0 & \sum_{i=1}^k n_i\rho\left(X_i\right) & \sum_{i=1}^k n_i\rho\left(X_i\right)^2 \end{bmatrix}$$

We use this matrix to compute the asymptotic variance of $\widehat{B}_{LN,10}$. From above, we have seen that Mathematica can compute numerical values for the theoretical mean and variance of this quantile function. Thus, $\widehat{B}_{LN,10}$ will be Normally distributed with mean $B_{LN,10}$ and variance given by (3.2), where

$$c'_\pi = \left(\begin{array}{ccc} \frac{\partial B_{LN,10}}{\partial \sigma} & \frac{\partial B_{LN,10}}{\partial \alpha_G} & \frac{\partial B_{LN,10}}{\partial \beta_{LN}} \end{array}\right),$$

is evaluated numerically using Mathematica.

5.6 Summary

This chapter outlined the theory necessary to fit accelerated Weibull, Burr, Gamma and Lognormal distributions to data sets, under the assumption that the distributions were correctly chosen. Our notation allowed us to include details for both Log-linear and Arrhenius scale-stress relationships simultaneously. We also considered the EFI matrix, and discussed the distribution of B_{10}. When we examined the Weibull distribution, we reported on a series of simulations to assess the effects of fitting this model to data, when no mis-specification had taken place. We did this for various parameter values and sample sizes, and outlined how changes in various aspects of the experiment affect the standard errors of the Weibull MLEs and $\widehat{B}_{W,10}$. We also compared theoretical standard errors of $\widehat{B}_{10}$ for the accelerated and non-accelerated Weibull distributions for varying sample sizes, in order to deduce if acceleration greatly increased this quantity. The results will be used in the next chapter as a benchmark for mis-specified scenarios, where we assess the effects of mis-specification for the accelerated Weibull distribution.

Chapter 6

Mis-specification In Accelerated Life Testing : Some Theoretical Considerations

6.1 The scope for mis-specification

In this chapter, we examine the theoretical aspects of mis-specification in accelerated life testing. In keeping with our work on non-accelerated data sets, we always have the Weibull distribution as the mis-specified model, and fit this to data with an underlying Burr, Gamma or Lognormal distribution. However, when we consider accelerated models, there are other aspects of the model which can be mis-specified. We can also choose the wrong relationship between stress level and scale parameter, and so, for example, fit the Weibull Log-linear model to data from an underlying Burr Arrhenius model. Thus, there are many possible combinations to take when considering the effects of mis-specification in accelerated distributions. Of course, our best case scenario would be no mis-specification, so that the distribution we fit is the same as the true distribution of the data; this has been covered in the last chapter, and provides a benchmark for the results here. We limit ourselves to examining the following cases of mis-specification :

- We mis-specify the scale-stress relationship, but choose the correct underlying distribution. We keep the Weibull distribution as the true underlying model, and look at the effects of mis-specifying the Log-linear and Arrhenius relationships.

- We correctly specify the underlying scale-stress relationship, but mis-specify the distribution function. We keep the Log-linear model as the true relationship, and fit the Weibull distribution to data with an underlying Burr, Gamma and Lognormal model.

- The final scenario involves mis-specifying both the scale-stress relationship and underlying distribution function. The cases we consider are shown in Table 6.1.

True	Mis-specified
Burr Log-linear	Weibull Arrhenius
Burr Arrhenius	Weibull Log-linear
Gamma Log-linear	Weibull Arrhenius
Gamma Arrhenius	Weibull Log-linear
Lognormal Log-linear	Weibull Arrhenius
Lognormal Arrhenius	Weibull Log-linear

Table 6.1: Types of mis-specification for accelerated data sets.

In all the above cases, we examine the effects of mis-specification using methods established for the non-accelerated case. Thus, in this chapter, we derive the entropy function, and the asymptotic variance covariance matrix of the mis-specified MLEs, from which we can obtain the asymptotic variance of $\widehat{B}_{10}$. A more practical approach of running simulations to assess the effects of using the mis-specified model is discussed in Chapter 7. Since we always take the Weibull as the mis-specified model, we can generalise the form of the entropy function, and the asymptotic variance covariance matrix of the mis-specified MLEs for any true underlying distribution function and scale-stress relationship. We introduce some further notation, and write

$$\begin{aligned} \rho_t(X_i) &= \rho(X_i) \text{ from the true distribution} \\ \rho_m(X_i) &= \rho(X_i) \text{ from the mis-specified distribution} \end{aligned}$$

So, for example, if we are fitting the Weibull Arrhenius to data with an underlying Burr Log-linear distribution, then

$$\begin{aligned} \rho_t(X_i) &= X_i \\ \rho_m(X_i) &= (X_i + c)^{-1} \end{aligned}$$

Using (5.4), we write the entropy as

$$\begin{aligned} E_t &= E_t[l_{W*}] = n \ln B_* + (B_* - 1) \sum_{i=1}^{k} n_i E_t[\ln Y_i] - nB_*\alpha_{W*} - B_*\beta_{W*} \sum_{i=1}^{k} n_i \rho_m(X_i) \\ &\quad - \exp(-B_*\alpha_{W*}) \sum_{i=1}^{k} n_i \exp\{-B_*\beta_{W*}\rho_m(X_i)\} E_t\left[Y_i^{B_*}\right], \end{aligned} \tag{6.1}$$

where the expectations $E_t[.]$ are with respect to the true model, and involve parameters from the underlying distribution and $\rho_t(X_i)$. If we differentiate E_t with respect to α_{W*}, and equate to zero, then we obtain

$$\alpha_{W*} = B_*^{-1} \ln \left\{ \frac{\sum_{i=1}^{k} n_i \exp\{-B_*\beta_{W*}\rho_m(X_i)\} E_t\left[Y_i^{B_*}\right]}{n} \right\}. \tag{6.2}$$

Inserting this into (6.1) yields the profile entropy

$$\begin{aligned} E_t^+ &= n \ln B_* + (B_* - 1) \sum_{i=1}^{k} n_i E_t [\ln Y_i] - B_* \beta_{W*} \sum_{i=1}^{k} n_i \rho_m (X_i) \\ &\quad - n \ln \left\{ \sum_{i=1}^{k} n_i \exp\{-B_* \beta_{W*} \rho_m (X_i)\} E_t \left[Y_i^{B_*} \right] \right\}, \end{aligned} \tag{6.3}$$

with score functions

$$\begin{aligned} \frac{\partial E_t^+}{\partial B_*} &= n B_*^{-1} + \sum_{i=1}^{k} n_i E_t [\ln Y_i] - \beta_{W*} \sum_{i=1}^{k} n_i \rho_m (X_i) \\ &\quad - \frac{n \sum_{i=1}^{k} n_i \exp\{-B_* \beta_{W*} \rho_m (X_i)\} \left\{ E_t \left[Y_i^{B_*} \ln Y_i \right] - \beta_{W*} \rho_m (X_i) E_t \left[Y_i^{B_*} \right] \right\}}{\sum_{i=1}^{k} n_i \exp\{-B_* \beta_{W*} \rho_m (X_i)\} E_t \left[Y_i^{B_*} \right]}, \end{aligned}$$

and

$$\frac{\partial E_t^+}{\partial \beta_{W*}} = -B_* \sum_{i=1}^{k} n_i \rho_m (X_i) + \frac{n B_* \sum_{i=1}^{k} n_i \rho_m (X_i) \exp\{-B_* \beta_{W*} \rho_m (X_i)\} E_t \left[Y_i^{B_*} \right]}{\sum_{i=1}^{k} n_i \exp\{-B_* \beta_{W*} \rho_m (X_i)\} E_t \left[Y_i^{B_*} \right]}.$$

Second derivatives can also be written down but are omitted here. The roots of these profile entropy score functions are obtained by using the iterative Newton-Raphson process. For convergence, we choose appropriate starting values for the Weibull parameters B_* and β_{W*}. In simulation experiments, these are obtained by making use of the true relationship between scale and stress, and the true parameter values set by the experimenter. For example, suppose we fit G_{WA} to data from G_{BP}. Choices for stress levels and values in the scale-stress relationship (either Log-linear or Arrhenius) are outlined in Section 5.1.2. Thus, here we set

$$\alpha_B = 8, \ \beta_B = -0.02;$$

this results in first and k^{th} scale parameters from G_{BP} given by

$$\begin{aligned} \phi_{1P} &= \exp\{8 - 0.02\,(50)\} = 1096.6332, \\ \phi_{kP} &= \exp\{8 - 0.02\,(200)\} = 54.5982. \end{aligned}$$

To obtain starting values for B_A and β_{WA}, we set $B_A = \tau$, and use the scale parameters above to derive an initial estimate for β_{WA} given by

$$\frac{\ln\left(\frac{1096.6332}{54.5982}\right)}{\frac{1}{323.16} - \frac{1}{473.16}} \simeq 3058.$$

A similar approach can be used for any other forms of mis-specification.

We extend the details on the distribution of Weibull MLEs for non-accelerated data sets, to incorporate the extra parameter for the accelerated model. This enables us to state that the asymptotic distribution of $(B_*, \alpha_{W*}, \beta_{W*})'$ is Normal with mean $\left(\widetilde{B}_*, \widetilde{\alpha}_{W*}, \widetilde{\beta}_{W*}\right)'$, the entropy values obtained from maximising (6.1), and variance covariance matrix based on (3.18), with

$$A = \begin{bmatrix} -E_t\left[\frac{\partial^2 l_{W*}}{\partial B_*^2}\right] & & \\ -E_t\left[\frac{\partial^2 l_{W*}}{\partial B_* \partial \alpha_{W*}}\right] & -E_t\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}^2}\right] & \\ -E_t\left[\frac{\partial^2 l_{W*}}{\partial B_* \partial \beta_{W*}}\right] & -E_t\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*} \partial \beta_{W*}}\right] & -E_t\left[\frac{\partial^2 l_{W*}}{\partial \beta_{W*}^2}\right] \end{bmatrix},$$

and

$$V = \begin{bmatrix} Var_t\left(\frac{\partial l_{W*}}{\partial B_*}\right) & & \\ Cov_t\left(\frac{\partial l_{W*}}{\partial B_*}, \frac{\partial l_{W*}}{\partial \alpha_{W*}}\right) & Var_t\left(\frac{\partial l_{W*}}{\partial \alpha_{W*}}\right) & \\ Cov_t\left(\frac{\partial l_{W*}}{\partial B_*}, \frac{\partial l_{W*}}{\partial \beta_{W*}}\right) & Cov_t\left(\frac{\partial l_{W*}}{\partial \alpha_{W*}}, \frac{\partial l_{W*}}{\partial \beta_{W*}}\right) & Var_t\left(\frac{\partial l_{W*}}{\partial \beta_{W*}}\right) \end{bmatrix}.$$

We first list the elements that make up the matrix A :

$$\begin{aligned} -E_t\left[\frac{\partial^2 l_{W*}}{\partial B_*^2}\right] &= nB_*^{-2} + \exp\left(-B_*\alpha_{W*}\right)\sum_{i=1}^{k} n_i \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\} \\ &\quad \times \left\{ \begin{array}{c} E_t\left[Y_i^{B_*}\left(\ln Y_i\right)^2\right] \\ -2\left\{\alpha_{W*} + \beta_{W*}\rho_m\left(X_i\right)\right\} E_t\left[Y_i^{B_*}\ln Y_i\right] \\ +\left\{\alpha_{W*} + \beta_{W*}\rho_m\left(X_i\right)\right\}^2 E_t\left[Y_i^{B_*}\right] \end{array} \right\}, \end{aligned} \tag{6.4}$$

$$-E_t\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}^2}\right] = B_*^2 \exp\left(-B_*\alpha_{W*}\right)\sum_{i=1}^{k} n_i \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\} E_t\left[Y_i^{B_*}\right], \tag{6.5}$$

$$-E_t\left[\frac{\partial^2 l_{W*}}{\partial \beta_{W*}^2}\right] = B_*^2 \exp\left(-B_*\alpha_{W*}\right)\sum_{i=1}^{k} n_i\rho_m\left(X_i\right)^2 \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\} E_t\left[Y_i^{B_*}\right], \tag{6.6}$$

$$\begin{aligned} -E_t\left[\frac{\partial^2 l_{W*}}{\partial B_*\partial \alpha_{W*}}\right] &= n - \exp\left(-B_*\alpha_{W*}\right)\sum_{i=1}^{k} n_i \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\} \\ &\quad \times \left[\begin{array}{c} B_* E_t\left[Y_i^{B_*}\ln Y_i\right] \\ +\left\{1 - B_*\left\{\alpha_{W*} + \beta_{W*}\rho_m\left(X_i\right)\right\} E_t\left[Y_i^{B_*}\right]\right\} \end{array} \right], \end{aligned} \tag{6.7}$$

$$-E_t\left[\frac{\partial^2 l_{W*}}{\partial B_*\partial\beta_{W*}}\right] = \sum_{i=1}^{k} n_i\rho_m(X_i) - \exp(-B_*\alpha_{W*})\sum_{i=1}^{k} n_i\rho_m(X_i)\exp\{-B_*\beta_{W*}\rho_m(X_i)\} \times\left[\begin{array}{c} B_*E_t\left[Y_i^{B_*}\ln Y_i\right] \\ +\left\{1-B_*\{\alpha_{W*}+\beta_{W*}\rho_m(X_i)\}E_t\left[Y_i^{B_*}\right]\right\}\end{array}\right], \tag{6.8}$$

and, finally,

$$-E_t\left[\frac{\partial^2 l_{W*}}{\partial\alpha_{W*}\partial\beta_{W*}}\right] = B_*^2\exp(-B_*\alpha_{W*})\sum_{i=1}^{k} n_i\rho_m(X_i)\exp\{-B_*\beta_{W*}\rho_m(X_i)\}E_t\left[Y_i^{B_*}\right]. \tag{6.9}$$

Now, we list the elements that make up the matrix V :

$$\begin{aligned}Var_t\left(\frac{\partial l_{W*}}{\partial B_*}\right) =& \sum_{i=1}^{k} n_i\left\{E_t\left[(\ln Y_i)^2\right]-(E_t[\ln Y_i])^2\right\} \\ &+\exp(-2B_*\alpha_{W*})\sum_{i=1}^{k} n_i\exp\{-2B_*\beta_{W*}\rho_m(X_i)\} \\ &\times\left[\begin{array}{l} E_t\left[Y_i^{2B_*}(\ln Y_i)^2\right]-\left\{E_t\left[Y_i^{B_*}\ln Y_i\right]\right\}^2 \\ +\{\alpha_{W*}+\beta_{W*}\rho_m(X_i)\}^2\left\{E_t\left[Y_i^{2B_*}\right]-\left(E_t\left[Y_i^{B_*}\right]\right)^2\right\} \\ +2\{\alpha_{W*}+\beta_{W*}\rho_m(X_i)\}\left\{\begin{array}{c} E_t\left[Y_i^{B_*}\ln Y_i\right]E_t\left[Y_i^{B_*}\right] \\ -E_t\left[Y_i^{2B_*}\ln Y_i\right]\end{array}\right\}\end{array}\right] \\ &-2\exp(-B_*\alpha_{W*})\sum_{i=1}^{k} n_i\exp\{-B_*\beta_{W*}\rho_m(X_i)\} \\ &\times\left[\begin{array}{l} E_t\left[Y_i^{B_*}(\ln Y_i)^2\right]-E_t\left[Y_i^{B_i}\ln Y_i\right]E_t[\ln Y_i] \\ -\{\alpha_{W*}+\beta_{W*}\rho_m(X_i)\}\left\{\begin{array}{c} E_t\left[Y_i^{B_*}\ln Y_i\right] \\ -E_t\left[Y_i^{B_*}\right]E_t[\ln Y_i]\end{array}\right\}\end{array}\right],\end{aligned} \tag{6.10}$$

$$\begin{aligned}Var_t\left(\frac{\partial l_{W*}}{\partial\alpha_{W*}}\right) =& B_*^2\exp(-2B_*\alpha_{W*})\sum_{i=1}^{k} n_i\exp\{-2B_*\beta_{W*}\rho_m(X_i)\} \\ &\times\left\{E_t\left[Y_i^{2B_*}\right]-\left(E_t\left[Y_i^{B_*}\right]\right)^2\right\},\end{aligned} \tag{6.11}$$

$$\begin{aligned}Var_t\left(\frac{\partial l_{W*}}{\partial\beta_{W*}}\right) =& B_*^2\exp(-2B_*\alpha_{W*})\sum_{i=1}^{k} n_i\rho_m(X_i)^2\exp\{-2B_*\beta_{W*}\rho_m(X_i)\} \\ &\times\left\{E_t\left[Y_i^{2B_*}\right]-\left(E_t\left[Y_i^{B_*}\right]\right)^2\right\},\end{aligned} \tag{6.12}$$

$$
\begin{aligned}
Cov_t\left(\frac{\partial l_{W*}}{\partial B_*}, \frac{\partial l_{W*}}{\partial \alpha_{W*}}\right) &= B_* \exp\left(-B_*\alpha_{W*}\right) \sum_{i=1}^{k} n_i \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\} \\
&\times \left\{E_t\left[Y_i^{B_*}\ln Y_i\right] - E_t\left[Y_i^{B_*}\right] E_t\left[\ln Y_i\right]\right\} \\
&- B_* \exp\left(-2B_*\alpha_{W*}\right) \sum_{i=1}^{k} n_i \exp\left\{-2B_*\beta_{W*}\rho_m\left(X_i\right)\right\} \\
&\times \left\{\begin{array}{c} E_t\left[Y_i^{2B_*}\ln Y_i\right] - E_t\left[Y_i^{B_*}\right] E_t\left[Y_i^{B_*}\ln Y_i\right] \\ -\left\{\alpha_{W*} + \beta_{W*}\rho_m\left(X_i\right)\right\} \\ \times\left\{E_t\left[Y_i^{2B_*}\right] - \left(E_t\left[Y_i^{B_*}\right]\right)^2\right\} \end{array}\right\}, \quad (6.13)
\end{aligned}
$$

$$
\begin{aligned}
Cov_t\left(\frac{\partial l_{W*}}{\partial B_*}, \frac{\partial l_{W*}}{\partial \beta_{W*}}\right) &= B_* \exp\left(-B_*\alpha_{W*}\right) \sum_{i=1}^{k} n_i\rho_m\left(X_i\right) \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\} \\
&\times \left\{E_t\left[Y_i^{B_*}\ln Y_i\right] - E_t\left[Y_i^{B_*}\right] E_t\left[\ln Y_i\right]\right\} \\
&- B_* \exp\left(-2B_*\alpha_{W*}\right) \sum_{i=1}^{k} n_i\rho_m\left(X_i\right) \exp\left\{-2B_*\beta_{W*}\rho_m\left(X_i\right)\right\} \\
&\times \left\{\begin{array}{c} E_t\left[Y_i^{2B_*}\ln Y_i\right] - E_t\left[Y_i^{B_*}\right] E_t\left[Y_i^{B_*}\ln Y_i\right] \\ -\left\{\alpha_{W*} + \beta_{W*}\rho_m\left(X_i\right)\right\} \\ \times\left\{E_t\left[Y_i^{2B_*}\right] - \left(E_t\left[Y_i^{B_*}\right]\right)^2\right\} \end{array}\right\}, \quad (6.14)
\end{aligned}
$$

and

$$
\begin{aligned}
Cov_t\left(\frac{\partial l_{W*}}{\partial \alpha_{W*}}, \frac{\partial l_{W*}}{\partial \beta_{W*}}\right) &= B_*^2 \exp\left(-2B_*\alpha_{W*}\right) \sum_{i=1}^{k} n_i\rho_m\left(X_i\right) \exp\left\{-2B_*\beta_{W*}\rho_m\left(X_i\right)\right\} \\
&\times \left\{E_t\left[Y_i^{2B_*}\right] - \left(E_t\left[Y_i^{B_*}\right]\right)^2\right\}. \quad (6.15)
\end{aligned}
$$

We can also write the approximation to the variance of $\widehat{B}_{W,10}$ in terms of these matrices. Again we adapt the results from the non-accelerated Weibull distribution to state that, asymptotically, $\widehat{B}_{W,10}$ is Normally distributed with mean

$$
B_{W,10}\left(\widetilde{B}_*, \widetilde{\alpha}_{W*}, \widetilde{\beta}_{W*}\right) = \exp\left\{\widetilde{\alpha}_{W*} + \widetilde{\beta}_{W*}\rho_m\left(X_s\right)\right\}\left(-\ln 0.9\right)^{\frac{1}{B_*}},
$$

and variance

$$
\left(\begin{array}{ccc} c_{B_*} & c_{\alpha_{W*}} & c_{\beta_{W*}} \end{array}\right) A^{-1}VA^{-1} \left(\begin{array}{c} c_{B_*} \\ c_{\alpha_{W*}} \\ c_{\beta_{W*}} \end{array}\right), \quad (6.16)
$$

where $\left(\begin{array}{ccc} c_{B_*} & c_{\alpha_{W*}} & c_{\beta_{W*}} \end{array} \right)$ is given by (5.12). Thus, to derive entropy values, the variance covariance structure of the mis-specified MLEs, and the asymptotic variance of $\widehat{B}_{W,10}$, we require $E_t\left[Y_i^m\right]$, $E_t\left[Y_i^m \ln Y_i\right]$ and $E_t\left[Y_i^m (\ln Y_i)^2\right]$. Using these results, we consider the above three scenarios of mis-specification; to do so, we first derive explicit results for the entropy function of our four distribution functions.

6.2 Entropy for Weibull to Weibull

We assume data has an underlying Weibull distribution, and mis-specify the scale-stress relationship. We extend notation established to distinguish between the $\rho(X_i)$ in true and mis-specified models to distinguish between the parameters from the Weibull distributions. Thus, we denote parameters from the mis-specified model as B_m, β_{Wm} and α_{Wm}; for the true, these become B_t, β_{Wt} and α_{Wt}. So, for example, if we were fitting G_{WA} to data from G_{WP}, then we would set $m = A$ and $t = P$. On examining our results for the general case above, we require $E_{Wt}\left[Y_i^m\right]$, $E_{Wt}\left[Y_i^m \ln Y_i\right]$ and $E_{Wt}\left[Y_i^m (\ln Y_i)^2\right]$; these are given by (5.9), (5.10) and (5.11) respectively, with $*$ replaced by t. We use (6.1) to write

$$\begin{aligned} E_{Wt} &= E_{Wt}\left[l_{Wm}\right] = n \ln B_m + (B_m - 1)\left\{ n\alpha_{Wt} + nB_t^{-1}\Psi(1) + \beta_{Wt}\sum_{i=1}^{k} n_i \rho_t(X_i) \right\} \\ &\quad -nB_m\alpha_{Wm} - B_m\beta_{Wm}\sum_{i=1}^{k} n_i \rho_m(X_i) - \\ &\quad \exp(-B_m\alpha_{Wm})\exp(B_m\alpha_{Wt})\Gamma\left(\frac{B_m}{B_t}+1\right) R(B_m, \beta_{Wm}), \end{aligned}$$

where

$$R(B_m, \beta_{Wm}) = \sum_{i=1}^{k} n_i \exp\left\{-B_m\beta_{Wm}\rho_m(X_i)\right\} \exp\left\{B_m\beta_{Wt}\rho_t(X_i)\right\}. \tag{6.17}$$

We also require derivatives of this function; these are given by

$$R_{1,0}(B_m, \beta_{Wm}) = \sum_{i=1}^{k} n_i \exp\left\{-B_m\beta_{Wm}\rho_m(X_i)\right\} \exp\left\{B_m\beta_{Wt}\rho_t(X_i)\right\}\left[\beta_{Wt}\rho_t(X_i) - \beta_{Wm}\rho_m(X_i)\right],$$

$$R_{0,1}(B_m, \beta_{Wm}) = -\sum_{i=1}^{k} n_i B_m \rho_m(X_i) \exp\left\{-B_m\beta_{Wm}\rho_m(X_i)\right\} \exp\left\{B_m\beta_{Wt}\rho_t(X_i)\right\},$$

$$R_{2,0}(B_m, \beta_{Wm}) = \sum_{i=1}^{k} n_i \exp\{-B_m\beta_{Wm}\rho_m(X_i)\} \exp\{B_m\beta_{Wt}\rho_t(X_i)\} \times [\beta_{Wt}\rho_t(X_i) - \beta_{Wm}\rho_m(X_i)]^2,$$

$$R_{0,2}(B_m, \beta_{Wm}) = \sum_{i=1}^{k} n_i B_m^2 \rho_m(X_i)^2 \exp\{-B_m\beta_{Wm}\rho_m(X_i)\} \exp\{B_m\beta_{Wt}\rho_t(X_i)\},$$

and

$$R_{1,1}(B_m, \beta_{Wm}) = -\sum_{i=1}^{k} n_i \rho_m(X_i) \exp\{-B_m\beta_{Wm}\rho_m(X_i)\} \exp\{B_m\beta_{Wt}\rho_t(X_i)\} \times [1 + B_m\{\beta_{Wt}\rho_t(X_i) - \beta_{Wm}\rho_m(X_i)\}].$$

We use (6.2) to write

$$\alpha_{Wm} = B_m^{-1} \ln\left\{\frac{\exp(B_m\alpha_{Wt})\Gamma\left(\frac{B_m}{B_t}+1\right) R(B_m, \beta_{Wm})}{n}\right\},$$

and now use (6.3), with appropriate substitutions to write the profile entropy as

$$\begin{aligned} E_{Wt}^{+} &= n\ln B_m + (B_m - 1)\left\{nB_t^{-1}\Psi(1) + \beta_{Wt}\sum_{i=1}^{k} n_i\rho_t(X_i)\right\} \\ &\quad -n\ln\Gamma\left(\frac{B_m}{B_t}+1\right) - n\ln R(B_m, \beta_{Wm}) \\ &\quad -B_m\beta_{Wm}\sum_{i=1}^{k} n_i\rho_m(X_i) + n\{\ln n - \alpha_{Wt} - 1\}, \end{aligned}$$

with first derivatives given by

$$\begin{aligned} \frac{\partial E_{Wt}^{+}}{\partial \beta_{Wm}} &= \frac{-nR_{0,1}(B_m, \beta_{Wm})}{R(B_m, \beta_{Wm})} - B_m\sum_{i=1}^{k} n_i\rho_m(X_i), \\ \frac{\partial E_{Wt}^{+}}{\partial B_m} &= nB_m^{-1} + nB_t^{-1}\Psi(1) + \beta_{Wt}\sum_{i=1}^{k} n_i\rho_t(X_i) - nB_t^{-1}\Psi\left(\frac{B_m}{B_t}+1\right) \\ &\quad -\frac{nR_{1,0}(B_m, \beta_{Wm})}{R(B_m, \beta_{Wm})} - \beta_{Wm}\sum_{i=1}^{k} n_i\rho_m(X_i). \end{aligned}$$

6.2.1 The variance structure of the mis-specified MLEs

We derive the variance covariance structure of the mis-specified Weibull MLEs, when this model is fitted to data arising from Weibull distributions with a different scale-stress relationship. We first consider the elements that make up the matrix A, and use (6.4) to write $-E_{Wt}\left[\frac{\partial^2 l_{Wm}}{\partial B_m^2}\right]$ as

$$nB_m^{-2} + \exp\left(-B_m\alpha_{Wm}\right)\exp\left(B_m\alpha_{Wt}\right)\Gamma\left(\frac{B_m}{B_t}+1\right)$$
$$\times\left[\begin{array}{c}\left(\left\{\alpha_{Wt}-\alpha_{Wm}+B_t^{-1}\Psi\left(\frac{B_m}{B_t}+1\right)\right\}^2+B_t^{-2}\Psi'\left(\frac{B_m}{B_t}+1\right)\right)R\left(B_m,\beta_{Wm}\right)\\+2\left\{\alpha_{Wt}-\alpha_{Wm}+B_t^{-1}\Psi\left(\frac{B_m}{B_t}+1\right)\right\}R_{1,0}\left(B_m,\beta_{Wm}\right)+R_{2,0}\left(B_m,\beta_{Wm}\right)\end{array}\right].$$

Next, using (6.5), we consider

$$-E_{Wt}\left[\frac{\partial^2 l_{Wm}}{\partial\alpha_{Wm}^2}\right] = B_m^2\exp\left(-B_m\alpha_{Wm}\right)\exp\left(B_{Wm}\alpha_{Wt}\right)\Gamma\left(\frac{B_m}{B_t}+1\right)R\left(B_m,\beta_{Wm}\right).$$

Now, we examine

$$-E_{Wt}\left[\frac{\partial^2 l_{Wm}}{\partial\beta_{Wm}^2}\right] = \exp\left(-B_m\alpha_{Wm}\right)\exp\left(B_m\alpha_{Wt}\right)\Gamma\left(\frac{B_m}{B_t}+1\right)R_{0,2}\left(B_m,\beta_{Wm}\right),$$

obtained from (6.6). Next, we derive $-E_{Wt}\left[\frac{\partial^2 l_{Wm}}{\partial B_m\partial\alpha_{Wm}}\right]$; from (6.7), this is equal to

$$n - \exp\left(-B_m\alpha_{Wm}\right)\exp\left(B_m\alpha_{Wt}\right)\Gamma\left(\frac{B_m}{B_t}+1\right)$$
$$\times\left[\begin{array}{c}B_mR_{1,0}\left(B_m,\beta_{Wm}\right)\\+\left(1+B_m\left\{\alpha_{Wt}-\alpha_{Wm}+B_t^{-1}\Psi\left(\frac{B_m}{B_t}+1\right)\right\}\right)R\left(B_m,\beta_{Wm}\right)\end{array}\right].$$

Now, on using (6.8), we consider

$$\begin{aligned}-E_{Wt}\left[\frac{\partial^2 l_{Wm}}{\partial B_m\partial\beta_{Wm}}\right] &= \sum_{i=1}^{k} n_i\rho_m\left(X_i\right)+\exp\left(-B_m\alpha_{Wm}\right)\exp\left(B_m\alpha_{Wt}\right)\Gamma\left(\frac{B_m}{B_t}+1\right)\\&\quad\times\left[\begin{array}{c}\left\{\alpha_{Wt}-\alpha_{Wm}+B_t^{-1}\Psi\left(\frac{B_m}{B_t}+1\right)\right\}R_{0,1}\left(B_m,\beta_{Wm}\right)\\+R_{1,1}\left(B_m,\beta_{Wm}\right)\end{array}\right],\end{aligned}$$

and finally, using (6.9), we have

$$-E_{Wt}\left[\frac{\partial^2 l_{Wm}}{\partial\alpha_{Wm}\partial\beta_{Wm}}\right] = -B_m\exp\left(-B_m\alpha_{Wm}\right)\exp\left(B_m\alpha_{Wt}\right)\Gamma\left(\frac{B_m}{B_t}+1\right)R_{0,1}\left(B_m,\beta_{Wm}\right).$$

We now list the elements which make up the matrix V, and to simplify our results write

$$\lambda_j = \exp\left(jB_m\alpha_{Wt}\right)\sum_{i=1}^{k} n_i\exp\left\{-jB_m\beta_{Wm}\rho_m\left(X_i\right)\right\}\exp\left\{jB_m\beta_{Wt}\rho_t\left(X_i\right)\right\},$$

$$\eta_j = \alpha_{Wt} + \beta_{Wt}\rho_t(X_i) + \frac{\Psi\left(\frac{jB_m}{B_t}+1\right)}{B_t},$$

and

$$\delta_j = \eta_j^2 + \{\beta_{Wm}\rho_m(X_i)\}^2 - 2\eta_j\beta_{Wm}\rho_m(X_i) + \alpha_{Wm}^2 - 2\alpha_{Wm}\{\eta_j - \beta_{Wm}\rho_m(X_i)\},$$

for $j = 1, 2$. Using this notation, and (6.10), we see that $Var_{Wt}\left(\frac{\partial l_{Wm}}{\partial B_m}\right)$ becomes

$$\begin{aligned}
& nB_t^{-2}\Psi'(1) + \lambda_2 \exp(-2B_m\alpha_{Wm}) \\
& \times\left[\Gamma\left(\frac{2B_m}{B_t}+1\right)\left\{\delta_2 + B_t^{-2}\Psi'\left(\frac{2B_m}{B_t}+1\right)\right\} - \Gamma\left(\frac{B_m}{B_t}+1\right)^2\delta_1\right] \\
& -2\lambda_1\exp(-B_m\alpha_{Wm})\Gamma\left(\frac{B_m}{B_t}+1\right) \\
& \times\left[\begin{array}{c} B_t^{-1}\left\{\Psi\left(\frac{B_m}{B_t}+1\right) - \Psi(1)\right\}\{\eta_1 - \beta_{Wm}\rho_m(X_i) - \alpha_{Wm}\} \\ +B_t^{-2}\Psi'\left(\frac{B_m}{B_t}+1\right) \end{array}\right].
\end{aligned}$$

We now have

$$Var_{Wt}\left(\frac{\partial l_{Wm}}{\partial \alpha_{Wm}}\right) = \lambda_2 B_m^2 \exp(-2B_m\alpha_{Wm})\left\{\Gamma\left(\frac{2B_m}{B_t}+1\right) - \Gamma\left(\frac{B_m}{B_t}+1\right)^2\right\},$$

on using (6.11). Next, we use (6.12) to write

$$\begin{aligned}
Var_{Wt}\left(\frac{\partial l_{Wm}}{\partial \beta_{Wm}}\right) &= \lambda_2\exp(-2B_m\alpha_{Wm})\{B_m\rho_m(X_i)\}^2 \\
&\quad \times\left\{\Gamma\left(\frac{2B_m}{B_t}+1\right) - \Gamma\left(\frac{B_m}{B_t}+1\right)^2\right\}.
\end{aligned}$$

We now consider covariances, and use (6.13) to obtain

$$\begin{aligned}
Cov_{Wt}\left(\frac{\partial l_{Wm}}{\partial B_m}, \frac{\partial l_{Wm}}{\partial \alpha_{Wm}}\right) &= \lambda_1 B_m\exp(-B_m\alpha_{Wm})\Gamma\left(\frac{B_m}{B_t}+1\right) \\
&\quad \times B_t^{-1}\left\{\Psi\left(\frac{B_m}{B_t}+1\right) - \Psi(1)\right\} \\
&\quad -\lambda_2 B_m\exp(-2B_m\alpha_{Wm}) \\
&\quad \times\left[\begin{array}{c} \Gamma\left(\frac{2B_m}{B_t}+1\right)\{\eta_2 - \beta_{Wm}\rho_m(X_i) - \alpha_{Wm}\} - \\ \Gamma\left(\frac{B_m}{B_t}+1\right)^2\{\eta_1 - \beta_{Wm}\rho_m(X_i) - \alpha_{Wm}\} \end{array}\right].
\end{aligned}$$

Next, we use (6.14) to derive

$$Cov_{Wt}\left(\frac{\partial l_{Wm}}{\partial B_m},\frac{\partial l_{Wm}}{\partial \beta_{Wm}}\right) = \lambda_1 \exp\left(-B_m\alpha_{Wm}\right)\left\{B_m\rho_m\left(X_i\right)\right\}\Gamma\left(\frac{B_m}{B_t}+1\right) \times B_t^{-1}\left\{\Psi\left(\frac{B_m}{B_t}+1\right)-\Psi\left(1\right)\right\} -\lambda_2\exp\left(-2B_m\alpha_{Wm}\right)\left\{B_m\rho_m\left(X_i\right)\right\} \times\left[\begin{array}{c}\Gamma\left(\frac{2B_m}{B_t}+1\right)\left\{\eta_2-\beta_{Wm}\rho_m\left(X_i\right)-\alpha_{Wm}\right\}\\ -\Gamma\left(\frac{B_m}{B_t}+1\right)^2\left\{\eta_1-\beta_{Wm}\rho_m\left(X_i\right)-\alpha_{Wm}\right\}\end{array}\right].$$

Finally, we use (6.15) to write

$$Cov_{Wt}\left(\frac{\partial l_{Wm}}{\partial \alpha_{Wm}},\frac{\partial l_{Wm}}{\partial \beta_{Wm}}\right) = \lambda_2 B_m\exp\left(-2B_m\alpha_{Wm}\right)\left\{B_m\rho_m\left(X_i\right)\right\} \times\left\{\Gamma\left(\frac{2B_m}{B_t}+1\right)-\Gamma\left(\frac{B_m}{B_t}+1\right)^2\right\}.$$

This list provides us with all the elements required to obtain the variance covariance matrix of the MLEs from the Weibull distribution, after this has been fitted to data also with an underlying Weibull distribution, but different scale-stress relationship. It also enables us to compute the distribution of $B_{W,10}$ from the mis-specified model; (6.16) gives this asymptotic variance.

6.3 Entropy for Weibull to Burr

We generalise results for the Arrhenius and Log-linear relationships, and derive the entropy function for the Weibull distribution, when this model is fitted to data with an underlying Burr distribution. Using previous results, we have

$$E_B\left[Y_i^{mB_*}\right]=\frac{\exp\left(mB_*\alpha_B\right)\exp\left\{mB_*\beta_B\rho_t\left(X_i\right)\right\}P_m}{\Gamma\left(a\right)}, \tag{6.18}$$

$$E_B\left[Y_i^{mB_*}\ln Y_i\right] = \frac{\exp\left(mB_*\alpha_B\right)\exp\left\{mB_*\beta_B\rho_t\left(X_i\right)\right\}}{\Gamma\left(a\right)} \times\left[\left\{\alpha_B+\beta_B\rho_t\left(X_i\right)\right\}P_m+\frac{P_m'}{m}\right], \tag{6.19}$$

and

$$E_B\left[Y_i^{mB_*}\left(\ln Y_i\right)^2\right] = \frac{\exp\left(mB_*\alpha_B\right)\exp\left\{mB_*\beta_B\rho_t\left(X_i\right)\right\}}{\Gamma\left(a\right)} \times\left[\begin{array}{c}\left\{\alpha_B+\beta_B\rho_t\left(X_i\right)\right\}^2P_m+\\ \frac{2\left\{\alpha_B+\beta_B\rho_t\left(X_i\right)\right\}P_m'}{m}+\frac{P_m''}{m^2}\end{array}\right]. \tag{6.20}$$

where P_j, P_j' and P_j'' are given by (3.27), (3.28) and (3.29), with β replaced by B_* and α replaced by a. Using (6.1), we write the entropy function as

$$\begin{aligned} E_B &= E\left[l_{W*}\right] = n \ln B_* + n\left(B_* - 1\right)\left\{\alpha_B + \frac{\Psi(1) - \Psi(a)}{\tau}\right\} + \beta_B\left(B_* - 1\right)\sum_{i=1}^{k} n_i \rho_t\left(X_i\right) \\ &\quad - nB_*\alpha_{W*} - B_*\beta_{W*}\sum_{i=1}^{k} n_i \rho_m\left(X_i\right) \\ &\quad - \frac{\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_B\right)P_1}{\Gamma(a)} R\left(B_*, \beta_{W*}\right), \end{aligned}$$

where now

$$R\left(B_*, \beta_{W*}\right) = \sum_{i=1}^{k} n_i \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\}\exp\left\{B_*\beta_B\rho_t\left(X_i\right)\right\}.$$

Note that, with our usual convention for partial derivatives,

$$\begin{aligned} R_{1,0}\left(B_*, \beta_{W*}\right) &= \frac{\partial R\left(B_*, \beta_{W*}\right)}{\partial B_*} \\ &= \sum_{i=1}^{k} n_i \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\}\exp\left\{B_*\beta_B\rho_t\left(X_i\right)\right\}\left[\beta_B\rho_t\left(X_i\right) - \beta_{W*}\rho_m\left(X_i\right)\right], \end{aligned}$$

$$\begin{aligned} R_{0,1}\left(B_*, \beta_{W*}\right) &= \frac{\partial R\left(B_*, \beta_{W*}\right)}{\partial \beta_{W*}} \\ &= -\sum_{i=1}^{k} B_* n_i \rho_m\left(X_i\right)\exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\}\exp\left\{B_*\beta_B\rho_t\left(X_i\right)\right\}, \end{aligned}$$

and, for later use,

$$\begin{aligned} R_{2,0}\left(B_*, \beta_{W*}\right) &= \frac{\partial R_{1,0}\left(B_*, \beta_{W*}\right)}{\partial B_*} \\ &= \sum_{i=1}^{k} n_i \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\}\exp\left\{B_*\beta_B\rho_t\left(X_i\right)\right\}\left[\beta_B\rho_t\left(X_i\right) - \beta_{W*}\rho_m\left(X_i\right)\right]^2, \end{aligned}$$

$$\begin{aligned} R_{0,2}\left(B_*, \beta_{W*}\right) &= \frac{\partial R_{0,1}\left(B_*, \beta_{W*}\right)}{\partial \beta_{W*}} \\ &= \sum_{i=1}^{k} B_*^2 n_i \rho_m\left(X_i\right)^2 \exp\left\{-B_*\beta_{W*}\rho_m\left(X_i\right)\right\}\exp\left\{B_*\beta_B\rho_t\left(X_i\right)\right\}, \end{aligned}$$

and

$$R_{1,1}(B_*, \beta_{W*}) = \frac{\partial R_{1,0}(B_*, \beta_{W*})}{\partial \beta_{W*}}$$
$$= -\sum_{i=1}^{k} n_i \rho_m(X_i) \exp(-B_* \beta_{W*} \rho_m(X_i)) \exp(B_* \beta_B \rho_t(X_i))$$
$$\times [1 + B_* \{\beta_B \rho_t(X_i) - \beta_{W*} \rho_m(X_i)\}].$$

We use (6.2) and (6.18) to write

$$\alpha_{W*} = B_*^{-1} \ln \left\{ \frac{\exp(B_* \alpha_B) P_1 R(B_*, \beta_{W*})}{n\Gamma(a)} \right\},$$

and (6.3) to derive the profile entropy function; this is given by

$$E_B^+ = n \ln B_* + n(B_* - 1) \left\{ \alpha_B + \frac{\Psi(1) - \Psi(a)}{\tau} \right\} + \beta_B (B_* - 1) \sum_{i=1}^{k} n_i \rho_t(X_i)$$
$$-nB_* \alpha_B - n \ln P_1 - n \ln R(B_*, \beta_{W*})$$
$$+n \ln \{n\Gamma(a)\} - B_* \beta_{W*} \sum_{i=1}^{k} n_i \rho_m(X_i) - n,$$

and has first derivatives

$$\frac{\partial E_B^+}{\partial B_*} = nB_*^{-1} + n\tau^{-1} \left\{ \Psi(1) - \Psi(a) - \Psi\left(\frac{B_*}{\tau} + 1\right) + \Psi\left(a - \frac{B_*}{\tau}\right) \right\}$$
$$+\sum_{i=1}^{k} n_i \{\beta_B \rho_t(X_i) - \beta_{W*} \rho_m(X_i)\} - \frac{nR_{1,0}(B_*, \beta_{W*})}{R(B_*, \beta_{W*})},$$

and

$$\frac{\partial E_B^+}{\partial \beta_{W*}} = \frac{-nR_{0,1}(B_*, \beta_{W*})}{R(B_*, \beta_{W*})} - B_* \sum_{i=1}^{k} n_i \rho_m(X_i).$$

6.3.1 Simplifications when $\rho_m(X_i) = \rho_t(X_i)$

If we set $\rho_m(X_i) = \rho_t(X_i)$, then, with equal scale-stress relationships, the entropy values will have the following properties:

Property 1

The first property is

$$\widetilde{\beta}_{W*} = \beta_B.$$

We prove this by examining the form of $\frac{\partial E_B^+}{\partial \beta_{W*}}$ for $\rho_m(X_i) = \rho_t(X_i)$; this is given by

$$\frac{\partial E_B^+}{\partial \beta_{W*}} = \frac{nB_* \sum_{i=1}^k n_i \rho_t(X_i) \exp\{-B_* \rho_t(X_i)(\beta_{W*} - \beta_B)\}}{\sum_{i=1}^k n_i \exp\{-B_* \rho_t(X_i)(\beta_{W*} - \beta_B)\}} - B_* \sum_{i=1}^k n_i \rho_t(X_i),$$

which, if $\beta_{W*} = \beta_B$, reduces to zero. Hence, β_B is a root of this function. To show that this is the only root, we consider the gradient of this function for $\beta_{W*} > \beta_B$ and $\beta_{W*} < \beta_B$. If the gradient is always increasing or always decreasing, then this will prove that the root is unique, since the function will never cross the horizontal axis again for values greater than or less that β_B. The gradient of this profile score function is given by $\frac{\partial^2 E_B^+}{\partial \beta_{W*}^2}$ which now, with $\widetilde{\beta}_{W*} = \beta_B$, simplifies to

$$-\frac{B_*^2}{n}\left\{n\sum_{i=1}^k n_i \rho_t(X_i)^2 - \left(\sum_{i=1}^k n_i \rho_t(X_i)\right)^2\right\},$$

a term which is now independent of β_{W*}. Hence, this gradient will always have the same sign, thus proving that β_B is the only root of the profile entropy score function. In fact, the gradient is negative, since the term in brackets summarises the spread of $\rho_t(X_i)$, which is positive. The function also appears in the denominator for the variance of α_{W*} and β_{W*} above, which, as Table 5.5 shows, is positive. This property will simplify matters considerably, since functions like (6.17) then reduce to n. We also have one less entropy value to estimate, so simplifying the search for entropy values to one dimension. The algebra for obtaining the variance covariance matrix of the mis-specified Weibull MLEs will also be greatly simplified.

Property 2

The profile entropy function is directly proportional to the sample size, and its maximising value is independent of k, X_i and n_i. We prove this second property by using the first, so that $\widetilde{\beta}_{W*} = \beta_B$. Using this, the profile entropy function becomes

$$n\left[\ln B_* + (B_* - 1)\left\{\alpha_B + \frac{\Psi(1) - \Psi(a)}{\tau}\right\} - B_*\alpha_B - \ln\left\{\Gamma\left(\frac{B_*}{\tau} + 1\right)\Gamma\left(a - \frac{B_*}{\tau}\right)\right\}\right] + C,$$

where

$$C = n \ln \Gamma(a) - n - \beta_B \sum_{i=1}^k n_i \rho_t(X_i).$$

If we ignore this term, since this is independent of B_*, and hence does not contribute to $\widetilde{B}_*$, then we see that the maximising value of this function is independent of the sample size, how this sample is arranged within stress levels, and the values and number of stress levels used.

Property 3

The entropy value for the Weibull shape parameter B_* is identical to the entropy value for the shape parameter of the Weibull distribution when no acceleration takes place and the same shape parameters from the Burr distribution are used (remember that $\widetilde{B}_*$ is not affected by the value of ϕ_{i*}). This final point does not really require a proof, since we use our work from the non-accelerated case. Here, we saw that the entropy value of the shape parameter from the Weibull distribution was independent of the scale parameter from the Burr. Thus, no matter what value of ϕ was chosen, the value of $\widetilde{B}_*$ (denoted by β_0 in the non-accelerated case) always remained the same. Since the process of acceleration is equivalent to fitting a distribution with varying scale parameter at each stress level, then the entropy value for B_* will be the same as the non-accelerated parameter, simply because it is not affected by the value of the Burr scale parameter chosen and set at each stress level.

6.3.2 The variance structure of the mis-specified MLEs

We list the elements that make up the variance covariance matrix of the mis-specified MLEs. We first consider expected values of second derivatives, and use (6.4) to write

$$-E_B\left[\frac{\partial^2 l_{W*}}{\partial B_*^2}\right] = nB_*^{-2} + \frac{\exp(-B_*\alpha_{W*})\exp(B_*\alpha_B)}{\Gamma(a)} \times \left[\begin{array}{c} R_{2,0}(B_*,\beta_{W*})P_1 + \\ 2\{(\alpha_B-\alpha_{W*})P_1 + P_1'\}R_{1,0}(B_*,\beta_{W*}) \\ +\left\{(\alpha_B-\alpha_{W*})^2 P_1 + 2(\alpha_B-\alpha_{W*})P_1' + P_1''\right\}R(B_*,\beta_{W*}) \end{array}\right].$$

We now make use of (6.5) to write

$$-E_B\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}^2}\right] = \frac{B_*^2\exp(-B_*\alpha_{W*})\exp(B_*\alpha_B)P_1R(B_*,\beta_{W*})}{\Gamma(a)}.$$

Using (6.6), we see that

$$-E_B\left[\frac{\partial^2 l_{W*}}{\partial \beta_{W*}^2}\right] = \frac{\exp(-B_*\alpha_{W*})\exp(B_*\alpha_B)P_1R_{0,2}(B_*,\beta_{W*})}{\Gamma(a)}.$$

Next, with (6.7), we have

$$-E_B\left[\frac{\partial^2 l_{W*}}{\partial B_*\partial \alpha_{W*}}\right] = n - \frac{\exp(-B_*\alpha_{W*})\exp(B_*\alpha_B)}{\Gamma(a)} \times \left\{\begin{array}{c} P_1B_*R_{1,0}(B_*,\beta_{W*}) + \\ \left[P_1 + B_*\{(\alpha_B-\alpha_{W*})P_1 + P_1'\}\right]R(B_*,\beta_{W*}) \end{array}\right\}.$$

Now we use (6.8) to write

$$-E_B\left[\frac{\partial^2 l_{W*}}{\partial B_*\partial\beta_{W*}}\right] = \sum_{i=1}^{k} n_i\rho_m(X_i) + \frac{\exp(-B_*\alpha_{W*})\exp(B_*\alpha_B)}{\Gamma(a)} \times\left[\begin{array}{c}\{(\alpha_B-\alpha_{W*})P_1+P_1'\}R_{0,1}(B_*,\beta_{W*})\\ +P_1R_{1,1}(B_*,\beta_{W*})\end{array}\right].$$

Finally, using (6.9), we consider

$$-E_B\left[\frac{\partial^2 l_{W*}}{\partial\alpha_{W*}\partial\beta_{W*}}\right] = \frac{-B_*\exp(-B_*\alpha_{W*})\exp(B_*\alpha_B)P_1R_{0,1}(B_*,\beta_{W*})}{\Gamma(a)}.$$

We introduce further notation to write the elements which make up the matrix V, and write

$$\Gamma_j = \frac{P_j^{3-j}}{\Gamma(a)^{2-j}},$$

and

$$g_j = \alpha_B + \frac{\Psi\left(\frac{jB_*}{\tau}+1\right)-\Psi\left(a-\frac{jB_*}{\tau}\right)}{\tau}.$$

Thus, using (6.10), we have

$$\begin{aligned} Var_B\left(\frac{\partial l_{W*}}{\partial B_*}\right) &= n\left\{\frac{\Psi'(1)-\Psi'(a)}{\tau^2}\right\} + \frac{\exp(-2B_*\alpha_{W*})\exp(2B_*\alpha_B)}{\Gamma(a)} \\ &\times\left[\begin{array}{c}\left(\begin{array}{c}\Gamma_2\{g_2-\alpha_{W*}\}^2-\Gamma_1\{g_1-\alpha_{W*}\}^2\\ +\Gamma_2\left\{\frac{\Psi'\left(\frac{2B_*}{\tau}+1\right)+\Psi'\left(a-\frac{2B_*}{\tau}\right)}{\tau^2}\right\}\end{array}\right)R(2B_*,\beta_{W*})\\ +2\{\Gamma_2(g_2-\alpha_{W*})-\Gamma_1(g_1-\alpha_{W*})\}R_{1,0}(2B_*,\beta_{W*})\\ +\{\Gamma_2-\Gamma_1\}R_{2,0}(2B_*,\beta_{W*})\end{array}\right] \\ &-\frac{2\exp(-B_*\alpha_{W*})\exp(B_*\alpha_B)\Gamma\left(\frac{B_*}{\tau}+1\right)\Gamma\left(a-\frac{B_*}{\tau}\right)}{\Gamma(a)} \\ &\times\left[\begin{array}{c}\left\{(g_1-\alpha_{W*})(g_1-g_0)+\frac{\Psi'\left(\frac{B_*}{\tau}+1\right)+\Psi'\left(a-\frac{B_*}{\tau}\right)}{\tau^2}\right\}R(B_*,\beta_{W*})\\ +(g_1-g_0)R_{1,0}(B_*,\beta_{W*})\end{array}\right].\end{aligned}$$

Now, using (6.11), we see that

$$Var_B\left(\frac{\partial l_{W*}}{\partial\alpha_{W*}}\right) = \frac{B_*^2\exp(-2B_*\alpha_{W*})\exp(2B_*\alpha_B)(\Gamma_2-\Gamma_1)R(2B_*,\beta_{W*})}{\Gamma(a)}.$$

Next we consider $Var_B\left(\frac{\partial l_{W*}}{\partial \beta_{W*}}\right)$, which, using (6.12) equates to

$$\frac{\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_B\right)\left(\Gamma_2-\Gamma_1\right)R_{0,2}\left(2B_*,\beta_{W*}\right)}{4\Gamma\left(a\right)}.$$

We now start to examine the covariances; from (6.13)

$$\begin{aligned} Cov_B\left(\frac{\partial l_{W*}}{\partial B_*},\frac{\partial l_{W*}}{\partial \alpha_{W*}}\right) &= \frac{B_*\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_B\right)\Gamma\left(\frac{B_*}{\tau}+1\right)\Gamma\left(a-\frac{B_*}{\tau}\right)}{\Gamma\left(a\right)} \\ &\times\left\{g_1-g_0\right\}R\left(B_*,\beta_{W*}\right)-\frac{B_*\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_B\right)}{\Gamma\left(a\right)} \\ &\times\left[\begin{array}{c}\left\{\Gamma_2\left(g_2-\alpha_{W*}\right)-\Gamma_1\left(g_1-\alpha_{W*}\right)\right\}R\left(2B_*,\beta_{W*}\right) \\ +\left\{\Gamma_2-\Gamma_1\right\}R_{1,0}\left(2B_*,\beta_{W*}\right)\end{array}\right].\end{aligned}$$

Next, we consider (6.14)

$$\begin{aligned} Cov_B\left(\frac{\partial l_{W*}}{\partial B_*},\frac{\partial l_{W*}}{\partial \beta_{W*}}\right) &= \frac{-\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_B\right)\Gamma\left(\frac{B_*}{\tau}+1\right)\Gamma\left(a-\frac{B_*}{\tau}\right)}{\Gamma\left(a\right)} \\ &+\times\left\{g_1-g_0\right\}R_{0,1}\left(B_*,\beta_{W*}\right)-\frac{\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_B\right)}{2\Gamma\left(a\right)} \\ &\times\left[\begin{array}{c}\left\{\Gamma_2-\Gamma_1\right\}R_{1,1}\left(2B_*,\beta_{W*}\right)+ \\ \left\{\Gamma_2\left(g_2-\alpha_{W*}\right)-\Gamma_1\left(g_1-\alpha_{W*}\right)-\frac{\Gamma_2-\Gamma_1}{2B_*}\right\}R_{0,1}\left(2B_*,\beta_{W*}\right)\end{array}\right].\end{aligned}$$

and finally, from (6.15), we derive

$$Cov_B\left(\frac{\partial l_{W*}}{\partial \alpha_{W*}},\frac{\partial l_{W*}}{\partial \beta_{W*}}\right)=\frac{-B_*\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_B\right)\left\{\Gamma_2-\Gamma_1\right\}R_{0,1}\left(2B_*,\beta_{W*}\right)}{2\Gamma\left(a\right)}.$$

We now have all the elements required to compute the variance covariance matrix of the MLEs from the mis-specified Weibull distribution, when this is fitted to data with an underlying Burr model and either Arrhenius or Log-linear relationships are used in both cases. These elements are used in the derivation of the mean and variance of $\widehat{B}_{W,10}$. The form of this has been considered in previous scenarios, and so we just note that, asymptotically, $\widehat{B}_{W,10}$ will be Normally distributed with mean $B_{W,10}\left(\widetilde{B}_*,\widetilde{\alpha}_{W*},\widetilde{\beta}_{W*}\right)$ and variance given by (6.16).

6.4 Entropy for Weibull to Gamma

We derive results to obtain theoretical counterparts to the Weibull MLEs when this distribution is fitted to data with an underlying Gamma model; as in previous cases, we do this simultaneously for both Arrhenius and Log-linear scale stress relationships. We first derive expectations required to compute the entropy function and its derivatives, and use (5.21) to obtain an expression for $E_G[Y_i^m]$. The remaining expectations are given by

$$E_G[Y_i^m \ln Y_i] = \frac{\exp(m\alpha_G)\exp\{m\beta_G\rho_t(X_i)\}\Gamma(m+\tau)}{\Gamma(\tau)}\{\alpha_G + \beta_G\rho_t(X_i) + \Psi(m+\tau)\}, \tag{6.21}$$

and

$$\begin{aligned} E_G\left[Y_i^m(\ln Y_i)^2\right] &= \frac{\exp(m\alpha_G)\exp\{m\beta_G\rho_t(X_i)\}\Gamma(m+\tau)}{\Gamma(\tau)} \\ &\quad\times\left[\{\alpha_G + \beta_G\rho_t(X_i) + \Psi(m+\tau)\}^2 + \Psi'(m+\tau)\right]. \end{aligned} \tag{6.22}$$

Thus, with appropriate substitutions for m, and (6.1), we have

$$\begin{aligned} E_G &= E[l_{W*}] = n\ln B_* + n(B_*-1)\{\alpha_G + \Psi(\tau)\} + \beta_G(B_*-1)\sum_{i=1}^{k} n_i\rho_t(X_i) - nB_*\alpha_{W*} \\ &\quad -B_*\beta_{W*}\sum_{i=1}^{k} n_i\rho_m(X_i) - \frac{\exp(-B_*\alpha_{W*})\exp(B_*\alpha_G)\Gamma(B_*+\tau)R(B_*,\beta_{W*})}{\Gamma(\tau)}, \end{aligned}$$

where

$$R(B_*,\beta_{W*}) = \sum_{i=1}^{k} n_i \exp\{-B_*\beta_{W*}\rho_m(X_i)\}\exp\{B_*\beta_G\rho_t(X_i)\}.$$

A slightly more compact notation is possible here; we have

$$\begin{aligned} R_{j,0}(B_*,\beta_{W*}) &= \frac{\partial^j R(B_*,\beta_{W*})}{\partial B_*^j} \\ &= \sum_{i=1}^{k} n_i \exp\{-B_*\beta_{W*}\rho_m(X_i)\}\exp\{B_*\beta_G\rho_t(X_i)\} \\ &\quad \times\{\beta_G\rho_t(X_i) - \beta_{W*}\rho_m(X_i)\}^j, \end{aligned}$$

$$\begin{aligned} R_{0,j}(B_*,\beta_{W*}) &= \frac{\partial^j R(B_*,\beta_{W*})}{\partial \beta_{W*}^j} \\ &= (-1)^j\sum_{i=1}^{k} n_i B_*^j \rho_m(X_i)^j \exp\{-B_*\beta_{W*}\rho_m(X_i)\}\exp\{B_*\beta_G\rho_t(X_i)\}, \end{aligned}$$

and, for later use,

$$\begin{aligned} R_{1,1}(B_*, \beta_{W*}) &= \frac{\partial R_{1,0}(B_*, \beta_{W*})}{\partial \beta_{W*}} \\ &= -\sum_{i=1}^{k} n_i \rho_m(X_i) \exp\{-B_* \beta_{W*} \rho_m(X_i)\} \exp\{B_* \beta_G \rho_t(X_i)\} \\ &\quad \times [1 + B_* \{\beta_G \rho_t(X_i) - \beta_{W*} \rho_m(X_i)\}]. \end{aligned}$$

We use (6.2) to write the parameter α_{W*} as

$$B_*^{-1} \ln \left\{ \frac{\exp(B_* \alpha_G) \Gamma(B_* + \tau) R(B_*, \beta_{W*})}{n\Gamma(\tau)} \right\},$$

and (6.3) to derive the profile entropy; this is given by

$$\begin{aligned} E_G^+ &= n \ln B_* + n(B_* - 1)\Psi(\tau) - n\alpha_G + \beta_G(B_* - 1)\sum_{i=1}^{k} n_i \rho_t(X_i) \\ &\quad -n \ln \Gamma(B_* + \tau) + n \ln n\Gamma(\tau) - n \ln R(B_*, \beta_{W*}) - B_* \beta_{W*} \sum_{i=1}^{k} n_i \rho_m(X_i), \end{aligned}$$

which has profile score functions

$$\begin{aligned} \frac{\partial E_G^+}{\partial B_*} &= nB_*^{-1} + n\Psi(\tau) + \beta_G \sum_{i=1}^{k} n_i \rho_t(X_i) - n\Psi(B_* + \tau) - \\ &\quad \frac{nR_{1,0}(B_*, \beta_{W*})}{R(B_*, \beta_{W*})} - \beta_{W*} \sum_{i=1}^{k} n_i \rho_m(X_i), \end{aligned}$$

and

$$\frac{\partial E_G^+}{\partial \beta_{W*}} = \frac{-nR_{0,1}(B_*, \beta_{W*})}{R(B_*, \beta_{W*})} - B_* \sum_{i=1}^{k} n_i \rho_m(X_i).$$

As for the Burr distribution, we note that considerable simplifications take place if the same scale-stress relationship is used in the true and mis-specified distributions. These include the fact that the theoretical counterpart to $\widehat{\beta}_{W*}$ is β_G, while the entropy values are not affected by the overall sample size, how this sample is arranged among the stress levels, the number of stress levels we take and how we chose the stress values. The proofs for such results are analogous to those for the Burr distribution, and hence are omitted. We further note that the value of $\widetilde{B}_*$ is not influenced by the scale parameters chosen from the Gamma distribution, and that entropy values for this parameter are the same as the non-accelerated counterparts provided the same shape parameters from the Gamma distribution are taken.

We continue by deriving the variance covariance matrix of the mis-specified Weibull MLEs. Again, considerable simplifications take place if the same scale-stress relationships are taken for the true and mis-specified distributions.

6.4.1 The variance structure of the mis-specified MLEs

Below we list the expectations and variances that make up the variance covariance matrix of the mis-specified Weibull distribution. We first consider the elements which make up the matrix A; these are given by

$$
\begin{aligned}
-E_G\left[\frac{\partial^2 l_{W*}}{\partial B_*^2}\right] &= nB_*^{-2} + \frac{\exp(-B_*\alpha_{W*})\exp(B_*\alpha_G)\Gamma(B_*+\tau)}{\Gamma(\tau)} \\
&\quad \times \left[\begin{array}{c} \left[\{\alpha_G - \alpha_{W*} + \Psi(B_*+\tau)\}^2 + \Psi'(B_*+\tau)\right] R(B_*, \beta_{W*}) \\ +2\{\alpha_G - \alpha_{W*} + \Psi(B_*+\tau)\} R_{1,0}(B_*, \beta_{W*}) \\ +R_{2,0}(B_*, \beta_{W*}) \end{array}\right],
\end{aligned}
$$

$$
-E_G\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}^2}\right] = \frac{B_*^2 \exp(-B_*\alpha_{W*})\exp(B_*\alpha_G)\Gamma(B_*+\tau) R(B_*, \beta_{W*})}{\Gamma(\tau)},
$$

$$
-E_G\left[\frac{\partial^2 l_{W*}}{\partial \beta_{W*}^2}\right] = \frac{\exp(-B_*\alpha_{W*})\exp(B_*\alpha_G)\Gamma(B_*+\tau) R_{0,2}(B_*, \beta_{W*})}{\Gamma(\tau)},
$$

$$
\begin{aligned}
-E_G\left[\frac{\partial^2 l_{W*}}{\partial B_* \partial \alpha_{W*}}\right] &= n - \frac{\exp(-B_*\alpha_{W*})\exp(B_*\alpha_G)\Gamma(B_*+\tau)}{\Gamma(\tau)} \\
&\quad \times \left[\begin{array}{c} B_* R_{1,0}(B_*, \beta_{W*}) + \\ 1 + B_*\{\alpha_G - \alpha_{W*} + \Psi(B_*+\tau)\} R(B_*, \beta_{W*}) \end{array}\right],
\end{aligned}
$$

$$
\begin{aligned}
-E_G\left[\frac{\partial^2 l_{W*}}{\partial B_* \partial \beta_{W*}}\right] &= \sum_{i=1}^{k} n_i \rho_m(X_i) + \frac{\exp(-B_*\alpha_{W*})\exp(B_*\alpha_G)\Gamma(B_*+\tau)}{\Gamma(\tau)} \\
&\quad \times \left[\{\alpha_G - \alpha_{W*} + \Psi(B_*+\tau)\} R_{0,1}(B_*, \beta_{W*}) + R_{1,1}(B_*, \beta_{W*})\right],
\end{aligned}
$$

and

$$
-E_G\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*} \partial \beta_{W*}}\right] = \frac{-B_* \exp(-B_*\alpha_{W*})\exp(B_*\alpha_G)\Gamma(B_*+\tau) R_{0,1}(B_*, \beta_{W*})}{\Gamma(\tau)}.
$$

We next consider the variance covariance structure of the scores; with

$$
\Gamma_j = \frac{\Gamma(jB_*+\tau)^{3-j}}{\Gamma(\tau)^{2-j}},
$$

and

$$h_j = \alpha_G + \Psi\left(jB_* + \tau\right),$$

we can list the elements which make up the matrix V below : the three variances are

$$\begin{aligned} Var_G\left(\frac{\partial l_{W*}}{\partial B_*}\right) &= n\Psi'\left(\tau\right) + \frac{\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_G\right)}{\Gamma\left(\tau\right)} \\ &\times \left[\begin{array}{c} \left\{\Gamma_2\left(h_2-\alpha_{W*}\right)^2 - \Gamma_1\left(h_1-\alpha_{W*}\right)^2 + \Gamma_2\Psi'\left(2B_*+\tau\right)\right\} R\left(2B_*,\beta_{W*}\right) \\ +2\left\{\Gamma_2\left(h_2-\alpha_{W*}\right) - \Gamma_1\left(h_1-\alpha_{W*}\right)\right\} R_{1,0}\left(2B_*,\beta_{W*}\right) \\ +\left(\Gamma_2-\Gamma_1\right) R_{2,0}\left(2B_*,\beta_{W*}\right) \end{array}\right] \\ &- \frac{2\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_G\right)\Gamma\left(B_*+\tau\right)}{\Gamma\left(\tau\right)} \\ &\times \left[\begin{array}{c} \left\{\left(h_1-h_0\right)\left(h_1-\alpha_{W*}\right) + \Psi'\left(B_*+\tau\right)\right\} R\left(B_*,\beta_{W*}\right) \\ +\left(h_1-h_0\right) R_{1,0}\left(B_*,\beta_{W*}\right) \end{array}\right], \end{aligned}$$

$$Var_G\left(\frac{\partial l_{W*}}{\partial \alpha_{W*}}\right) = \frac{B_*^2\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_G\right)\left\{\Gamma_2-\Gamma_1\right\} R\left(2B_*,\beta_{W*}\right)}{\Gamma\left(\tau\right)},$$

and

$$Var_G\left(\frac{\partial l_{W*}}{\partial \beta_{W*}}\right) = \frac{\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_G\right)\left\{\Gamma_2-\Gamma_1\right\} R_{0,2}\left(2B_*,\beta_{W*}\right)}{4\Gamma\left(\tau\right)},$$

while the three covariances are

$$\begin{aligned} Cov_G\left(\frac{\partial l_{W*}}{\partial B_*},\frac{\partial l_{W*}}{\partial \alpha_{W*}}\right) &= \frac{B_*\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_G\right)\Gamma\left(B_*+\tau\right)\left\{h_1-h_0\right\} R\left(B_*,\beta_{W*}\right)}{\Gamma\left(\tau\right)} \\ &- \frac{B_*\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_G\right)}{\Gamma\left(\tau\right)} \\ &\times \left[\begin{array}{c} \left\{\Gamma_2\left(h_2-\alpha_{W*}\right) - \Gamma_1\left(h_1-\alpha_{W*}\right)\right\} R\left(2B_*,\beta_{W*}\right) \\ +\left(\Gamma_2-\Gamma_1\right) R_{1,0}\left(2B_*,\beta_{W*}\right) \end{array}\right], \end{aligned}$$

$$\begin{aligned} Cov_G\left(\frac{\partial l_{W*}}{\partial B_*},\frac{\partial l_{W*}}{\partial \beta_{W*}}\right) &= \frac{\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_G\right)}{2\Gamma\left(\tau\right)} \\ &\times \left[\begin{array}{c} \left\{\Gamma_2\left(h_2-\alpha_{W*}\right) - \Gamma_1\left(h_1-\alpha_{W*}\right) - \frac{\Gamma_2-\Gamma_1}{2B_*}\right\} R_{0,1}\left(2B_*,\beta_{W*}\right) \\ +\left(\Gamma_2-\Gamma_1\right) R_{1,1}\left(2B_*,\beta_{W*}\right) \end{array}\right] \\ &- \frac{\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_G\right)\Gamma\left(B_*+\tau\right)\left\{h_1-h_0\right\} R_{0,1}\left(B_*,\beta_{W*}\right)}{\Gamma\left(\tau\right)}, \end{aligned}$$

and

$$Cov_G\left(\frac{\partial l_{W*}}{\partial \alpha_{W*}}, \frac{\partial l_{W*}}{\partial \beta_{W*}}\right) = \frac{-B_* \exp(-2B_*\alpha_{W*}) \exp(2B_*\alpha_G) \{\Gamma_2 - \Gamma_1\} R_{0,1}(2B_*, \beta_{W*})}{2\Gamma(\tau)}.$$

Using all these elements, we are now able to construct an asymptotic theoretical correlation matrix for the MLEs from a mis-specified accelerated Weibull distribution, when this has been fitted to data with an underlying accelerated Gamma model. We can also use this to compute the theoretical mean and variance of $\widehat{B}_{W,10}$.

6.5 Entropy for Weibull to Lognormal

We derive entropy values for the Weibull distribution when this model has been fitted to data with an underlying Lognormal distribution. The expectations required to do this are given by (5.22), (5.23) and (5.24). Using (6.1), we write the entropy function as

$$\begin{aligned} E_{LN} &= n\ln B_* + n\alpha_{LN}(B_* - 1) + \beta_{LN}(B_* - 1)\sum_{i=1}^{k} n_i\rho_t(X_i) \\ &\quad -nB_*\alpha_{W*} - B_*\beta_{W*}\sum_{i=1}^{k} n_i\rho_m(X_i) \\ &\quad -\exp(-B_*\alpha_{W*})\exp\left(B_*\alpha_{LN} + \frac{\sigma^2 B_*^2}{2}\right) R(B_*, \beta_{W*}), \end{aligned}$$

where

$$R(B_*, \beta_{W*}) = \sum_{i=1}^{k} n_i \exp\{-B_*\beta_{W*}\rho_m(X_i)\} \exp\{B_*\beta_{LN}\rho_t(X_i)\}.$$

For future reference, we also note that

$$\begin{aligned} R_{1,0}(B_*, \beta_{W*}) &= \frac{\partial R(B_*, \beta_{W*})}{\partial B_*} \\ &= \sum_{i=1}^{k} n_i \exp\{-B_*\beta_{W*}\rho_m(X_i)\} \exp\{B_*\beta_{LN}\rho_t(X_i)\} \\ &\quad \times \{\beta_{LN}\rho_t(X_i) - \beta_{W*}\rho_m(X_i)\}, \end{aligned}$$

$$\begin{aligned} R_{0,1}(B_*, \beta_{W*}) &= \frac{\partial R(B_*, \beta_{W*})}{\partial \beta_{W*}} \\ &= -B_*\sum_{i=1}^{k} n_i\rho_m(X_i) \exp\{-B_*\beta_{W*}\rho_m(X_i)\} \exp\{B_*\beta_{LN}\rho_t(X_i)\}, \end{aligned}$$

$$\begin{aligned} R_{2,0}(B_*, \beta_{W*}) &= \frac{\partial^2 R(B_*, \beta_{W*})}{\partial B_*^2} \\ &= \sum_{i=1}^{k} n_i \exp\{-B_* \beta_{W*} \rho_m(X_i)\} \exp\{B_* \beta_{LN} \rho_t(X_i)\} \\ &\quad \times \{\beta_{LN} \rho_t(X_i) - \beta_{W*} \rho_m(X_i)\}^2, \end{aligned}$$

$$\begin{aligned} R_{0,2}(B_*, \beta_{W*}) &= \frac{\partial^2 R(B_*, \beta_{W*})}{\partial \beta_{W*}^2} \\ &= B_*^2 \sum_{i=1}^{k} n_i \rho_m(X_i)^2 \exp\{-B_* \beta_{W*} \rho_m(X_i)\} \exp\{B_* \beta_{LN} \rho_t(X_i)\}, \end{aligned}$$

and

$$\begin{aligned} R_{1,1}(B_*, \beta_{W*}) &= \frac{\partial R_{1,0}(B_*, \beta_{W*})}{\partial \beta_{W*}} \\ &= -\sum_{i=1}^{k} n_i \rho_m(X_i) \exp\{-B_* \beta_{W*} \rho_m(X_i)\} \exp\{B_* \beta_{LN} \rho_t(X_i)\} \\ &\quad \times [1 + B_* \{\beta_{LN} \rho_t(X_i) - \beta_{W*} \rho_m(X_i)\}]. \end{aligned}$$

We use (6.2) to write

$$\alpha_{W*} = B_*^{-1} \ln \left\{ \frac{\exp\left(B_* \alpha_{LN} + \frac{\sigma^2 B_*^2}{2}\right) R(B_*, \beta_{W*})}{n} \right\},$$

and insert this into the entropy function to obtain the profile entropy given by

$$\begin{aligned} E_{LN}^{+} &= n \ln B_* - n\alpha_{LN} + \beta_{LN}(B_* - 1) \sum_{i=1}^{k} n_i \rho_t(X_i) - \frac{n\sigma^2 B_*^2}{2} \\ &\quad - n \ln R(B_*, \beta_{W*}) - B_* \beta_{W*} \sum_{i=1}^{k} n_i \rho_m(X_i). \end{aligned}$$

This has first derivatives

$$\begin{aligned} \frac{\partial E_{LN}^{+}}{\partial B_*} &= n B_*^{-1} + \beta_{LN} \sum_{i=1}^{k} n_i \rho_t(X_i) - n\sigma^2 B_* - \frac{n R_{1,0}(B_*, \beta_{W*})}{R(B_*, \beta_{W*})} \\ &\quad - \beta_{W*} \sum_{i=1}^{k} n_i \rho_m(X_i), \end{aligned}$$

and

$$\frac{\partial E_{LN}^{+}}{\partial \beta_{W*}} = \frac{-nR_{0,1}(B_*, \beta_{W*})}{R(B_*, \beta_{W*})} - B_* \sum_{i=1}^{k} n_i \rho_m(X_i).$$

Using these functions, we are now able to derive theoretical counterparts to the MLEs of the Weibull distribution, when this has been fitted to data with an underlying Lognormal distribution, using either Arrhenius or Log-linear scale-stress relationships. As in previous cases, considerable simplifications take place when deriving entropy values for the Weibull distribution if the same scale-stress relationship is used for this mis-specified distribution and the true Lognormal. If we set

$$\rho_m(X_i) = \rho_t(X_i)$$

then $\widetilde{\beta}_{W*} = \beta_{LN}$. Using this, we see that the profile entropy score function with respect to B_* becomes

$$\frac{\partial E_{LN}^{+}}{\partial B_*} = nB_*^{-1} - n\sigma^2 B_*,$$

giving

$$\widetilde{B}_* = \sigma^{-1},$$

a result analogous to that in the non-accelerated scenario. Also, if we have the same scale-stress relationship for true and mis-specified models, then entropy values are independent of the sample size, number of stress levels and the values that these take, with proofs as for the Burr distribution. If different scale-stress relationships are used for the true and mis-specified distributions, then entropy values do not undergo such simplifications, and we only note that they are independent of the total sample size if the sample is arranged equally amongst the stress levels.

Now that we can derive theoretical counterparts to the MLEs from the Weibull distribution, the next step is to use these to compute the variance covariance matrix of the mis-specified Weibull MLEs. We consider this in the next section.

6.5.1 The variance structure of the mis-specified MLEs

In order to obtain the distribution of Weibull MLEs when this model has been subjected to mis-specification and fitted to data with an underlying Lognormal distribution, we must obtain expected values of second derivatives and the variance covariance matrix of score functions, where variances and expectations are taken with respect to the Lognormal distribution. We first list the elements which make up the matrix A :

$$-E_{LN}\left[\frac{\partial^2 l_{W*}}{\partial B_*^2}\right] = nB_*^{-2}+\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_{LN}+\frac{\sigma^2 B_*^2}{2}\right) \times\left[\begin{array}{c}\left\{\sigma^2+\left(\alpha_{LN}-\alpha_{W*}+\sigma^2 B_*\right)^2\right\}R\left(B_*,\beta_{W*}\right)\\ +2\left(\alpha_{LN}-\alpha_{W*}+\sigma^2 B_*\right)R_{1,0}\left(B_*,\beta_{W*}\right)\\ +R_{2,0}\left(B_*,\beta_{W*}\right)\end{array}\right],$$

$$-E_{LN}\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}^2}\right] = B_*^2\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_{LN}+\frac{\sigma^2 B_*^2}{2}\right)R\left(B_*,\beta_{W*}\right),$$

$$-E_{LN}\left[\frac{\partial^2 l_{W*}}{\partial \beta_{W*}^2}\right] = \exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_{LN}+\frac{\sigma^2 B_*^2}{2}\right)R_{0,2}\left(B_*,\beta_{W*}\right),$$

$$-E_{LN}\left[\frac{\partial^2 l_{W*}}{\partial B_*\partial \alpha_{W*}}\right] = n-\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_{LN}+\frac{\sigma^2 B_*^2}{2}\right) \times\left[\begin{array}{c}\left\{1+B_*\left(\alpha_{LN}-\alpha_{W*}+\sigma^2 B_*\right)\right\}R\left(B_*,\beta_{W*}\right)\\ +B_*R_{1,0}\left(B_*,\beta_{W*}\right)\end{array}\right],$$

$$-E_{LN}\left[\frac{\partial^2 l_{W*}}{\partial B_*\partial \beta_{W*}}\right] = \sum_{i=1}^{k} n_i\rho_m\left(X_i\right)+\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_{LN}+\frac{\sigma^2 B_*^2}{2}\right) \times\left\{\left(\alpha_{LN}-\alpha_{W*}+\sigma^2 B_*\right)R_{0,1}\left(B_*,\beta_{W*}\right)+R_{1,1}\left(B_*,\beta_{W*}\right)\right\},$$

and, finally, we have

$$-E_{LN}\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}\partial \beta_{W*}}\right] = -B_*\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_{LN}+\frac{\sigma^2 B_*^2}{2}\right)R_{0,1}\left(B_*,\beta_{W*}\right).$$

We now list the elements that make up the matrix V, writing

$$E_j = \exp\left(j\sigma^2 B_*^2\right),$$

and

$$h_j = \alpha_{LN}+j\sigma^2 B_*.$$

Thus, the three variances are

$$\begin{aligned}Var_{LN}\left(\frac{\partial l_{W*}}{\partial B_*}\right) &= n\sigma^2 + \exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_{LN}\right) \\ &\times \left[\begin{array}{c}\left\{E_2\sigma^2 + E_2\left(h_2 - \alpha_{W*}\right)^2 - E_1\left(h_1 - \alpha_{W*}\right)^2\right\} R\left(2B_*, \beta_{W*}\right) \\ +2\left\{E_2\left(h_2 - \alpha_{W*}\right) - E_1\left(h_1 - \alpha_{W*}\right)\right\} R_{1,0}\left(2B_*, \beta_{W*}\right) \\ +\left\{E_2 - E_1\right\} R_{2,0}\left(2B_*, \beta_{W*}\right)\end{array}\right] \\ &-2\sigma^2\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_{LN} + \frac{\sigma^2 B_*^2}{2}\right) \\ &\times \left[\begin{array}{c}\left\{1 + B_*\left(h_1 - \alpha_{W*}\right)\right\} R\left(B_*, \beta_{W*}\right) \\ +B_* R_{1,0}\left(B_*, \beta_{W*}\right)\end{array}\right],\end{aligned}$$

$$Var_{LN}\left(\frac{\partial l_{W*}}{\partial \alpha_{W*}}\right) = B_*^2\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_{LN}\right)\left\{E_2 - E_1\right\} R\left(2B_*, \beta_{W*}\right),$$

and

$$Var_{LN}\left(\frac{\partial l_{W*}}{\partial \beta_{W*}}\right) = \frac{\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_{LN}\right)\left\{E_2 - E_1\right\} R_{0,2}\left(2B_*, \beta_{W*}\right)}{4},$$

while the three covariances are

$$\begin{aligned}Cov_{LN}\left(\frac{\partial l_{W*}}{\partial B_*}, \frac{\partial l_{W*}}{\partial \alpha_{W*}}\right) &= B_*^2\sigma^2\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_{LN} + \frac{\sigma^2 B_*^2}{2}\right) R\left(B_*, \beta_{W*}\right) \\ &-B_*\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_{LN}\right) \\ &\times \left[\begin{array}{c}\left\{E_2\left(h_2 - \alpha_{W*}\right) - E_1\left(h_1 - \alpha_{W*}\right)\right\} R\left(2B_*, \beta_{W*}\right) \\ +\left\{E_2 - E_1\right\} R_{1,0}\left(2B_*, \beta_{W*}\right)\end{array}\right],\end{aligned}$$

$$\begin{aligned}Cov_{LN}\left(\frac{\partial l_{W*}}{\partial B_*}, \frac{\partial l_{W*}}{\partial \beta_{W*}}\right) &= -B_*\sigma^2\exp\left(-B_*\alpha_{W*}\right)\exp\left(B_*\alpha_{LN} + \frac{\sigma^2 B_*^2}{2}\right) R_{0,1}\left(B_*, \beta_{W*}\right) \\ &+\frac{\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_{LN}\right)}{2} \\ &\times \left[\begin{array}{c}\left\{E_2\left(h_2 - \alpha_{W*}\right) - E_1\left(h_1 - \alpha_{W*}\right) - \frac{E_2 - E_1}{2B_*}\right\} R_{0,1}\left(2B_*, \beta_{W*}\right) \\ +\left(E_2 - E_1\right) R_{1,1}\left(2B_*, \beta_{W*}\right)\end{array}\right],\end{aligned}$$

and

$$Cov_{LN}\left(\frac{\partial l_{W*}}{\partial \alpha_{W*}}, \frac{\partial l_{W*}}{\partial \beta_{W*}}\right) = \frac{-B_*\exp\left(-2B_*\alpha_{W*}\right)\exp\left(2B_*\alpha_{LN}\right)\left\{E_2 - E_1\right\} R_{0,1}\left(2B_*, \beta_{W*}\right)}{2}.$$

These expectations, variances and covariances provide us with all the elements required to compute the variance covariance matrix of the Weibull MLEs when this distribution has been fitted to a set of data with an underlying Lognormal distribution. We check our results in the next chapter.

6.6 Summary

This chapter outlined the theory required to compute entropy values from the Weibull distribution, and the variance covariance structure of the mis-specified MLEs. The added relationship linking scale to stress further complicated mis-specification, and we had to allow for mis-specifying this relationship also. We derived the theory necessary to obtain entropy values when we mis-specified just the scale-stress relationship (either Log-linear or Arrhenius), when we mis-specified the Weibull distribution and fitted this to data from another model (either the Burr, Gamma or Lognormal), and finally, when we mis-specified both the distribution function and scale-stress relationship. The theory we developed allowed us to consider the three scenarios simultaneously. In the next chapter, we take a more practical approach to mis-specification in accelerated life testing, and use the theory outlined in this chapter to assess the effects of using the Weibull distribution to model a data set with some other underlying distribution.

Chapter 7

Mis-specification In Accelerated Life Testing : Further Practical Considerations

In the last chapter, we outlined the theory required to compute entropy values and standard errors of the mis-specified MLEs. We now use this theory to assess some of the effects of mis-specification, and to consider how well a mis-specified distribution fits the data. We take a similar approach to the non-accelerated case, and first use simulations to examine the effects of mis-specification. This also provides us with a check on the theory developed above, and enables us to assess how often we prefer the mis-specified distribution over the true, using a criterion based on maximised likelihoods. Running simulations also allows comparison between estimates of B_{10} from true and mis-specified distributions. We then assess the agreement between the mis-specified distribution and the true underlying model, and compute maximum absolute distances between the cdfs of true and mis-specified distributions, across all stress levels, and for varying true parameter values. This approach is carried out for the scenarios described above; thus, we begin by first mis-specifying the scale-stress relationship.

7.1 Getting the scale-stress relationship wrong

7.1.1 Fitting G_{WA} to G_{WP}

Simulation studies

We report simulations to assess how well sample MLEs and their standard errors agree with theoretical counterparts, when we vary the true parameter values, number of stress levels and sample sizes set at each stress. As in cases where we specify the correct underlying distribution and scale-stress relationship, we split the results into three main parts corresponding to k = 2, 3 and 4 stress levels. We then set B_P equal to 0.5, 1, 2 and 3, since in practice, these values cover most of the cases of interest, but here only report results for $B_P = 2$; possible variations for k, X_i, n and n_i are discussed in Section 5.1.2 above. When we run simulations for all sets of stress levels, we record the MLEs from both the true and mis-specified distributions, and their standard errors, the entropy values corresponding to the MLEs from the Weibull Arrhenius model, and the theoretical standard errors from both distributions. We also include $\widehat{B}_{W,10}$ for both cases and the corresponding standard errors of these estimates. We compare this quantile with a true value of 355.9593 for $B_P = 2$. The results are summarised in Tables 7.1 and 7.2 for $k = 2$, Tables 7.3 and 7.4 for $k = 3$, and Tables 7.5 and 7.6 for $k = 4$. Generally, we see excellent agreement between sample and theoretical standard errors of MLEs from both true and mis-specified distributions across all values of k. As the sample size increases, this agreement improves. We also observe good agreement between MLEs from G_{WA} and the corresponding entropy values, even for relatively small sample sizes. The probability of fitting the mis-specified model for $k \geq 3$, is as high as 18% for small sample sizes, but decreases to zero as n increases. If we compare these probabilities for equal and non equal allocations, then we generally see a rise if we place more observations at the higher stress levels. For example, when $k = 3$ and we have 100 observations at each stress level, then we prefer G_{WA} approximately 3% of the time. In contrast, if we allocate 250 observations to X_3 and 25 to the remainder, then this figure increases to 14%. We also observe an increase if we allocate 200 observations to X_1 and 50 to X_2 and X_3. The standard error of $\widehat{B}_{W,10}$ for both distribution functions increases as the number of observations in the middle and higher stress levels increases. On the whole, there is generally good agreement between $\widehat{B}_{W,10}$ from true and mis-specified distributions, and when we carry out simulations for 2 stress levels, there is no difference at all. This reflects the fact that with $k = 2$, we effectively have a reparameterisation of the model, so whether we express this in terms of α_{WP} and β_{WP} or α_{WA} and β_{WA} makes no difference.

We conclude that theoretical results match up with simulations, and that we have derived the correct entropy values and standard errors of the mis-specified MLEs, and proceed, in the next section, by examining the penalty we pay for fitting the wrong distribution function.

The effects of mis-specification

By maximising the entropy function, we obtain the parameter values for the mis-specified distribution that provides the best possible approximation, in the circumstances, to the true underlying model. It is appropriate to assess just how well this mis-specified distribution does this, and if, for a particular set of true parameter values and stress levels, we may pay a large penalty by fitting the incorrect distribution. We can approach this problem in a number of ways; firstly, by comparing theoretical hazard functions of true and mis-specified distributions, and seeing whether significant distances occur between these. Since we also have a relationship linking scale and stress, then we could also compare these functions for both distributions, or report the relative error between the scale parameters of true and mis-specified models. However, we take a similar approach to that used in our non-accelerated

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_A$	2.0412	2.0179	2.0076	2.0031
$\widetilde{B}_A$	2	2	2	2
S	0.1644	0.1124	0.0699	0.0491
T	0.1559	0.1103	0.0697	0.0493
$\widehat{\alpha}_{WA}$	-2.4647	-2.4603	-2.4638	-2.4638
$\widetilde{\alpha}_{WA}$	-2.4632	-2.4632	-2.4632	-2.4632
S	0.2729	0.1837	0.1269	0.0859
T	0.2706	0.1914	0.1210	0.0856
$\widehat{\beta}_{WA}$	3057.5631	3056.5117	3058.0804	3058.2086
$\widetilde{\beta}_{WA}$	3058.1277	3058.1277	3058.1277	3058.1277
S	103.0099	69.0940	45.8022	32.3376
T	101.9376	72.0808	45.5879	32.2355
$\widehat{B}_{W,10(A)}$	363.3199	358.9180	357.2736	356.4884
S	46.2130	31.5229	19.9463	14.0966
T	44.8286	31.6986	20.0479	14.1760
Pr (Fit G_{WA})	-	-	-	-
B_P	2.0142	2.0179	2.0076	2.0032
S	0.1644	0.1124	0.0699	0.0491
T	0.1599	0.1103	0.0697	0.0493
$\widehat{\alpha}_{WP}$	7.9966	7.9974	7.9992	7.9997
S	0.1002	0.0678	0.0442	0.0312
T	0.0986	0.0697	0.0441	0.0312
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0007	0.0005	0.0003	0.0002
T	0.0007	0.0005	0.0003	0.0002
$\widehat{B}_{W,10(P)}$	363.3199	358.9180	357.2736	356.4884
S	46.2130	31.5229	19.9463	14.0966
T	44.8286	31.6986	20.0479	14.1760

Table 7.1: Fitting G_{WA} to G_{WP} for $k = 2$, $B_P = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_A$	2.0200	2.0228	2.0194	2.0195
$\widetilde{B}_A$	2	2	2	2
S	0.1149	0.1140	0.1117	0.1136
T	0.1103	0.1103	0.1103	0.1103
$\widehat{\alpha}_{WA}$	-2.4440	-2.4876	-2.4553	-2.4760
$\widetilde{\alpha}_{WA}$	-2.4632	-2.4632	-2.4632	-2.4632
S	0.2472	0.3302	0.2007	0.2403
T	0.2465	0.3260	0.1998	0.2400
$\widehat{\beta}_{WA}$	3048.9370	3065.9691	3054.3836	3062.1271
$\widetilde{\beta}_{WA}$	3058.1277	3058.1277	3058.1277	3058.1277
S	109.2660	110.2739	83.5282	83.2521
T	108.9759	108.9759	83.2317	83.2317
$\widehat{B}_{W,10(A)}$	358.1367	360.4713	359.1105	359.7035
S	44.5876	30.1261	36.1343	30.4144
T	44.2165	29.4795	36.3534	29.9869
Pr (Fit G_{WA})	-	-	-	-
$\widehat{B}_P$	2.0200	2.0228	2.0194	2.0195
S	0.1149	0.1140	0.1117	0.1136
T	0.1103	0.1103	0.1103	0.1103
$\widehat{\alpha}_{WP}$	7.9878	8.0025	7.9951	8.0008
S	0.1348	0.0618	0.0959	0.0604
T	0.1344	0.0615	0.0960	0.0605
$\widehat{\beta}_{WP}$	-0.0199	-0.0201	-0.0200	-0.0200
S	0.0007	0.0007	0.0005	0.0005
T	0.0007	0.0007	0.0005	0.0005
$\widehat{B}_{W,10(P)}$	358.1367	360.4713	359.1105	359.7035
S	44.5876	30.1261	36.1343	30.4144
T	44.2165	29.4795	36.3534	29.9869

Table 7.2: Fitting G_{WA} to G_{WP} for $k = 2$, $B_P = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	300,300	500,500
n_3	25	50	100	300	500
$\widehat{B}_A$	1.9955	1.9699	1.9575	1.9499	1.9481
$\widetilde{B}_A$	1.9451	1.9451	1.9451	1.9451	1.9451
S	0.1858	0.1279	0.0895	0.0510	0.0395
T	0.1757	0.1242	0.0878	0.0507	0.0393
$\widehat{\alpha}_{WA}$	-2.1442	-2.1513	-2.1509	-2.1551	-2.1552
$\widetilde{\alpha}_{WA}$	-2.1558	-2.1558	-2.1558	-2.1558	-2.1558
S	0.3626	0.2582	0.1808	0.1037	0.0809
T	0.3589	0.2537	0.1794	0.1036	0.0802
$\widehat{\beta}_{WA}$	2963.7244	2967.2885	2967.5143	2969.3653	2969.4507
$\widetilde{\beta}_{WA}$	2969.7718	2969.7718	2969.7718	2969.7718	2969.7718
S	141.2752	100.3784	70.3140	40.3386	31.5862
T	139.7392	98.8105	69.8696	40.3392	31.2466
$\widehat{B}_{W,10(A)}$	365.7774	361.3433	358.8949	357.6428	357.3055
S	57.8430	40.1802	28.2605	16.2582	12.6514
T	56.2254	39.7573	28.1127	16.2309	12.5724
Pr (Fit G_{WA})	0.1738	0.0923	0.0299	0.0009	0
$\widehat{B}_P$	2.0545	2.0269	2.0135	2.0048	2.0032
S	0.1903	0.1314	0.0920	0.0522	0.0405
T	0.1801	0.1273	0.0900	0.0520	0.0403
$\widehat{\alpha}_{WP}$	7.9929	7.9971	7.9977	7.9994	7.9996
S	0.1390	0.0980	0.0691	0.0396	0.0312
T	0.1376	0.0973	0.0688	0.0397	0.0308
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0009	0.0007	0.0005	0.0003	0.0002
T	0.0009	0.0007	0.0005	0.0003	0.0002
$\widehat{B}_{W,10(P)}$	365.5324	360.8675	358.2848	356.8197	356.5389
S	56.2596	39.1775	27.6762	15.7984	12.3582
T	54.8755	38.8028	27.4377	15.8412	12.2705

Table 7.3: Fitting G_{WA} to G_{WP} for $k = 3$, $B_P = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	50,100 150	25,200 75	25,25 250	200,50 50
$\widehat{B}_A$	1.9513	1.9578	1.9891	1.9844
$\widetilde{B}_A$	1.9392	1.9473	1.9763	1.9710
S	0.0900	0.0888	0.0906	0.0903
T	0.0878	0.0872	0.0897	0.0889
$\widehat{\alpha}_{WA}$	-2.3742	-2.3049	-2.5369	-2.0766
$\widetilde{\alpha}_{WA}$	-2.3842	-2.3116	-2.5531	-2.0726
S	0.2001	0.2841	0.2381	0.2041
T	0.1988	0.2797	0.2369	0.2050
$\widehat{\beta}_{WA}$	3056.0636	3059.8002	3100.9965	2935.3286
$\widetilde{\beta}_{WA}$	3060.6083	3063.0995	3108.5634	2934.1459
S	83.8062	119.6057	106.6882	71.7443
T	83.2419	117.7164	106.1167	71.9604
$\widehat{B}_{W,10(A)}$	376.5582	410.6235	376.8391	355.2084
S	34.2895	45.0224	43.0858	25.1451
T	34.1132	44.5839	42.8641	24.8686
Pr (Fit G_{WA})	0.0248	0.0205	0.1374	0.0885
$\widehat{B}_P$	2.0138	2.0125	2.0134	2.0139
S	0.0923	0.0916	0.0909	0.0914
T	0.0900	0.0900	0.0900	0.0900
$\widehat{\alpha}_{WP}$	7.9952	7.9976	7.9902	8.0004
S	0.0916	0.1228	0.1289	0.0534
T	0.0909	0.1211	0.1282	0.0532
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0199	-0.0200
S	0.0005	0.0008	0.0007	0.0005
T	0.0005	0.0008	0.0007	0.0005
$\widehat{B}_{W,10(P)}$	358.0148	358.8638	357.4852	358.7051
S	31.8111	37.2861	40.2532	25.0573
T	31.6150	36.8697	40.0490	24.7807

Table 7.4: Fitting G_{WA} to G_{WP} for $k = 3$, $B_P = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_A$	1.9978	1.9762	1.9670	1.9632	1.9600
$\widetilde{B}_A$	1.9583	1.9583	1.9583	1.9583	1.9583
S	0.1611	0.1114	0.0778	0.0538	0.0341
T	0.1532	0.1083	0.0766	0.0542	0.0343
$\widehat{\alpha}_{WA}$	-2.1466	-2.1512	-2.1555	-2.1572	-2.1580
$\widetilde{\alpha}_{WA}$	-2.1588	-2.1588	-2.1588	-2.1588	-2.1588
S	0.3278	0.2307	0.1613	0.1152	0.0722
T	0.3227	0.2282	0.1613	0.1141	0.0722
$\widehat{\beta}_{WA}$	2964.8558	2967.4456	2969.3863	2970.4403	2970.8105
$\widetilde{\beta}_{WA}$	2971.2285	2971.2285	2971.2285	2971.2285	2971.2285
S	131.7773	92.9593	65.0521	46.4406	29.0535
T	130.0480	91.9578	65.0240	45.9789	29.0796
$\widehat{B}_{W,10(A)}$	366.7286	363.0111	361.3588	360.9709	360.3351
S	52.5689	37.3911	26.1470	18.3668	11.5104
T	52.0054	36.7734	26.0027	18.3867	11.6288
Pr (Fit G_{WA})	0.1785	0.0957	0.0306	0.0051	0
$\widehat{B}_P$	2.0413	2.0195	2.0091	2.0051	2.0017
S	0.1647	0.1130	0.0792	0.0550	0.0347
T	0.1559	0.1103	0.0780	0.0551	0.0349
$\widehat{\alpha}_{WP}$	7.9914	7.9955	7.9972	7.9990	7.9994
S	0.1375	0.0971	0.0684	0.0487	0.0304
T	0.1363	0.0964	0.0682	0.0482	0.0305
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0009	0.0006	0.0004	0.0003	0.0002
T	0.0009	0.0006	0.0004	0.0003	0.0002
$\widehat{B}_{W,10(P)}$	362.7668	359.1653	357.2686	356.8567	356.2038
S	51.3522	36.1960	25.4149	17.9461	11.1902
T	50.5236	35.7256	25.2618	17.8628	11.2974

Table 7.5: Fitting G_{WA} to G_{WP} for $k = 4$, $B_P = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 75,75	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_A$	1.9850	1.9810	1.9612	1.9997	1.9900	1.9852	1.9806
$\widetilde{B}_A$	1.9671	1.9642	1.9438	1.9781	1.9692	1.9647	1.9630
S	0.1110	0.1104	0.1100	0.1117	0.1109	0.1111	0.1105
T	0.1089	0.1082	0.1074	0.1094	0.1086	0.1090	0.1083
$\widehat{\alpha}_{WA}$	-2.3134	-1.9092	-2.3463	-2.0893	-1.9671	-2.3919	-2.1012
$\widetilde{\alpha}_{WA}$	-2.3330	-1.9096	-2.3617	-2.0920	-1.9659	-2.4022	-2.1232
S	0.2712	0.2319	0.2771	0.2173	0.2298	0.2655	0.2959
T	0.2662	0.2335	0.2745	0.2146	0.2302	0.2606	0.2917
$\widehat{\beta}_{WA}$	3025.3077	2884.6033	3054.1205	2940.4122	2901.0047	3053.4610	2961.7871
$\widetilde{\beta}_{WA}$	3034.4916	2885.3982	3061.2082	2941.7754	2901.0464	3058.4162	2971.6532
S	116.7780	88.2932	117.5091	83.9001	84.5961	115.0169	124.3637
T	114.4705	88.9126	116.1886	82.8637	84.6132	112.6270	122.3030
$\widehat{B}_{W,10(A)}$	371.9306	358.5057	387.9019	359.6745	357.6604	375.1829	376.8334
S	45.9648	33.6998	46.7418	33.9304	32.1787	46.4254	46.4861
T	45.2440	33.4826	46.2290	33.4293	31.7638	45.3332	45.4895
Pr(Fit G_{WA})	0.1242	0.0968	0.0570	0.1720	0.1212	0.1252	0.0931
$\widehat{B}_P$	2.0193	2.0180	2.0197	2.0222	2.0214	2.0216	2.0194
S	0.1124	0.1125	0.1128	0.1125	0.1125	0.1123	0.1127
T	0.1103	0.1103	0.1103	0.1103	0.1103	0.1103	0.1103
$\widehat{\alpha}_{WP}$	7.9889	7.9985	7.9924	7.9984	7.9991	7.9938	7.9899
S	0.1338	0.0823	0.1276	0.0824	0.0735	0.1331	0.1346
T	0.1311	0.0829	0.1262	0.0817	0.0731	0.1301	0.1319
$\widehat{\beta}_{WP}$	-0.0199	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0199
S	0.0008	0.0006	0.0008	0.0006	0.0006	0.0007	0.0008
T	0.0008	0.0006	0.0008	0.0006	0.0006	0.0007	0.0008
$\widehat{B}_{W,10(P)}$	358.2741	359.2730	359.0701	360.2508	360.0243	360.0105	358.5769
S	43.6121	33.1719	41.8448	33.6557	31.9490	43.7611	43.1995
T	42.8015	33.0141	41.4431	33.1637	31.5513	42.7478	42.1629

Table 7.6: Fitting G_{WA} to G_{WP} for $k = 4$, $B_P = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

work, and examine distances between cdfs of true and mis-specified distributions across all stress levels; comparisons based on other approaches will be considered elsewhere. We illustrate this approach using the true parameter values

$$B_P = 2,\ \alpha_{WP} = 8,\ \beta_{WP} = -0.02,$$

with $k = 3$ and the usual X_i. When we have equal sample sizes at each stress level (irrespective of the overall sample size), the entropy values are

$$\widetilde{B}_A = 1.9451,\ \widetilde{\alpha}_{WA} = -2.1558,\ \widetilde{\beta}_{WA} = 2969.7718.$$

As with our work on non-accelerated data, we examine the maximum absolute distance between true and mis-specified cdfs across the three stress levels. Thus, we will have three different distributions to consider. With the true Weibull Log-linear model, we have scale parameters

$$\theta_{1P} = 1096.6332,\ \theta_{2P} = 148.4132,\ \theta_{3P} = 54.5982,$$

which we compare to the mis-specified Weibull Arrhenius entropy values

$$\widetilde{\theta}_{1A} = 1134.5459,\ \widetilde{\theta}_{2A} = 129.3190,\ \widetilde{\theta}_{3A} = 61.6001,$$

where the shape parameters from each distribution remain constant across stress levels. The table of figures, Table 7.7, shows the three different distribution functions corresponding to each stress level for the true and mis-specified models. The fit between true and mis-specified distributions for the lowest stress level is the best, in terms of having the smallest maximum absolute distance. The second stress level seems to give the worst fit, and we observe a maximum distance of 0.10018 between the two distribution functions.

We now carry out a similar procedure for varying k, n_i and B_P.

Two stress levels We first consider results for 2 stress levels, and set the stresses and scale-stress parameters to values used in simulation experiments. When we allow B_P to vary, we observe that $\widetilde{\beta}_{WA}$ and $\widetilde{\alpha}_{WA}$ are unchanged, and $\widetilde{B}_A = B_P$. Also, neither of the three parameters depend on how the sample is arranged across stress levels. Thus, for any sample size, and any value of B_P, we have

$$B_P = \widetilde{B}_A$$

and

$$\widetilde{\alpha}_{WA} = -2.4632,\ \widetilde{\beta}_{WA} = 3058.1277.$$

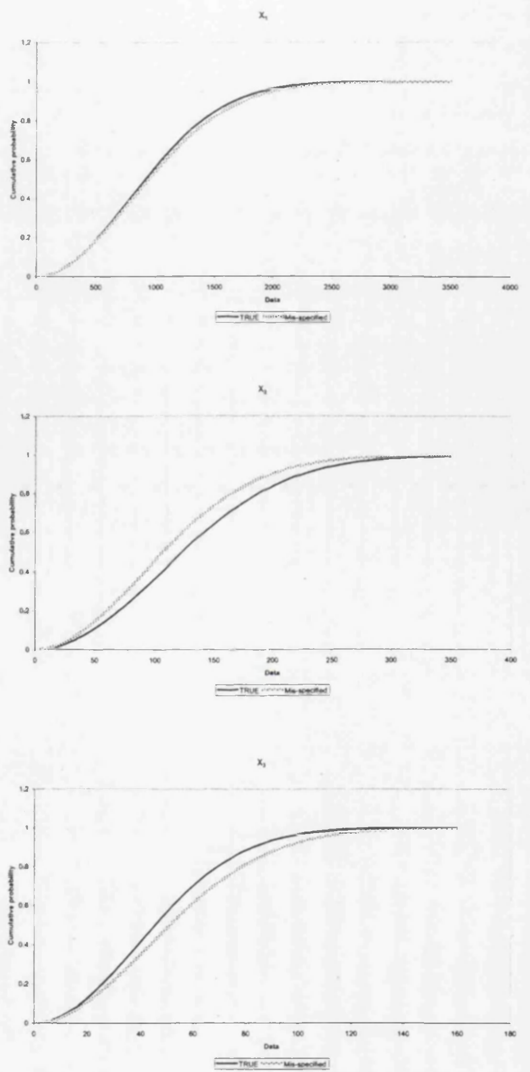


Table 7.7: Comparison between the true Weibull Log-linear model, with parameters $B_P = 2$, $\alpha_{WP} = 8$, $\beta_{WP} = -0.02$, and the mis-specified Weibull Arrhenius distribution, with entropy values $\widetilde{B}_A = 1.9451$, $\widetilde{\alpha}_{WA} = -2.1558$, $\widetilde{\beta}_{WA} = 2969.7718$. Here, $k = 3$, $X_1 = 50$, $X_2 = 150$ and $X_3 = 200$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.09999	2984.43251	-2.19816	0.00535
0.5	0.49912	2981.47438	-2.18958	0.02644
1	0.99296	2977.65074	-2.17852	0.05201
1.5	1.47644	2973.73258	-2.16723	0.07663
2	1.94510	2969.77184	-2.15584	0.10021
2.5	2.39535	2965.81951	-2.14449	0.12272
3	2.82468	2961.92164	-2.13330	0.14417
3.5	3.23156	2958.11677	-2.12237	0.16456
4	3.61539	2954.43483	-2.11178	0.18578
4.5	3.97623	2950.89723	-2.10159	0.21046
5	4.31468	2947.51777	-2.09184	0.23483

Table 7.8: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1 : 1 : 1)$.

The scale parameters for both distributions always match up, so a perfect fit is observed between the true and mis-specified models. This will be true regardless of the sample size, true parameter values and how the sample is arranged across the two stress levels. Thus, if we mis-specify the Log-linear scale-stress relationship and incorrectly use the Arrhenius model, then no penalties are paid as a result of this, and an equivalent fit will always be observed. This is because we can always write the parameters from the mis-specified scale-stress relationship in terms of the parameters from the true model for $k = 2$.

Three stress levels We continue by considering distances for three stress levels, and first derive results when we arrange the sample equally amongst stress levels. The maximum absolute distances for varying B_P are summarised in Table 7.8. We observe the maximum absolute distance between true and mis-specified distribution functions increase as B_P becomes larger. We conclude, for these particular parameter values and stress levels, the larger B_P becomes, the worse the fit between true and mis-specified distributions, as shape parameters increase, but at different rates. We examine this behaviour for a different set of stress levels and parameter values, and try to choose them so they are not like the ones used in this example. In this case, the acceleration factor for the above set of stress levels and parameter values was around 10. We further increase the rate of acceleration to see if similar results are observed for such cases. We set

$$\beta_{WP} = -0.03, \alpha_{WP} = 11,$$

and have three stress levels with values

$$X_1 = 50,\ X_2 = 150,\ X_3 = 250.$$

The entropy values and maximum distances for varying B_P are shown in Table 7.9. Again

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.09933	4976.02818	-5.72024	0.01696
0.5	0.49155	4962.02028	-5.67506	0.08196
1	0.93653	4943.00758	-5.61510	0.15536
1.5	1.30960	4924.14342	-5.55624	0.21927
2	1.61038	4906.77559	-5.50217	0.27437
2.5	1.85017	4981.46041	-5.45438	0.32187
3	2.04188	4878.25240	-5.41298	0.36725
3.5	2.19645	4866.97790	-5.37741	0.41563
4	2.32234	4857.38818	-5.34695	0.45827
4.5	2.42591	4849.23004	-5.32083	0.49466
5	2.51195	4842.27378	-5.29836	0.52710

Table 7.9: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and $\beta_{WP} = -0.03, \alpha_{WP} = 11$; ratio is $(1 : 1 : 1)$.

we see a similar pattern to the previous case, and increasing the value of B_P considerably affects the fit between true and mis-specified distribution functions. Generally, the larger B_P becomes, the worse the fit between both models. This investigation shows that if we have three stress levels with equal allocations of items, then for small values of B_P, an excellent fit occurs between true and mis-specified distributions, and this seems true regardless of the underlying stress levels and acceleration factor. As B_P increases, the fit becomes gradually worse, and this is further affected by the acceleration factor. From examining both examples, we see that for experiments with higher rates of acceleration, the fit is generally worse for larger B_P than it is for similar experiments but with a smaller acceleration factor. Thus, we would have a very poor fit if we were to consider using the Weibull Arrhenius model to represent data with an underlying Weibull Log-linear distribution and parameter values and stress levels equal to

$$B_P = 10,\ \beta_{WP} = -0.03, \alpha_{WP} = 11,\ X_1 = 50,\ X_2 = 150,\ X_2 = 250.$$

In fact, the largest absolute distance between true and mis-specified distributions over the three stress levels is 0.692158.

We now allow the allocation of observations at each stress level to vary. We have seen how the overall sample size does not affect the entropy values, only the way in which we arrange this sample amongst the levels. Thus, entropy values for a sample of $(100, 200, 300)$ are the same as entropy values for $(6, 12, 18)$, $(1, 2, 3)$, and so on. We now examine the effects of changing the proportion of observations across stress levels, where we set stress levels and parameter values for G_{WP} as outlined in Section 5.1.2. We tabulate the results for varying B_P and ratios, representing, for example, an experiment with twice as many observations at the second stress level, and three times as many at the third as $(1 : 2 : 3)$. We tabulate results for this ratio, and $(3 : 2 : 1)$, $(2 : 3 : 1)$, $(1 : 3 : 2)$, $(1 : 1 : 3)$, $(3 : 1 : 1)$ and $(1 : 3 : 1)$. We consider whether having a higher proportion of observations at a certain stress level improves

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.09999	3060.10425	-2.38879	0.00578
0.5	0.49904	3060.20125	-2.38781	0.02867
1	0.99229	3060.33006	-2.38658	0.05664
1.5	1.47405	3060.46622	-2.38538	0.08376
2	1.93921	3060.60827	-2.38419	0.10994
2.5	2.38360	3060.75458	-2.38303	0.13512
3	2.80428	3060.90343	-2.38189	0.15926
3.5	3.19947	3061.05319	-2.38079	0.18234
4	3.56851	3061.20239	-2.37972	0.20437
4.5	3.91156	3061.34977	-2.37868	0.22538
5	4.22945	3061.49431	-2.37768	0.24538

Table 7.10: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1 : 2 : 3)$.

the fit between true and mis-specified distributions. Tables 7.10, 7.11, 7.12, 7.13, 7.14, 7.15 and 7.16, respectively, show entropy values and maximum distances for varying B_P with the above ratios, and Figure 7.1 shows these maximum absolute distances for the seven different ratios at once. We see the best scenario (in terms of having the minimum absolute distance between the two distribution functions) when we have equal observations across the three stress level, or a slightly higher proportion at the lowest stress. The worse case occurs for the ratio $(2 : 3 : 1)$ implying that if we put the majority of observations at the middle stress level, then we compensate for this by observing a bad fit between true and mis-specified distribution functions. We can further illustrate this point by computing entropy values when we have quite a considerable acceleration factor, large value of B_P and most of the sample contained at the middle stress level. In fact, if we put

$$X_1 = 50,\ X_2 = 150,\ X_3 = 250,$$

and

$$B_P = 10,\ \alpha_{WP} = 11,\ \beta_{WP} = -0.03,$$

and consider a ratio where we have 5 times as many observations in the middle stress level, then we obtain entropy values of

$$\widetilde{B}_A = 4.5721,\ \widetilde{\beta}_{WA} = 4465.7630,\ \widetilde{\alpha}_{WA} = -4.1405$$

and the largest absolute distance between true and mis-specified distribution functions is 0.95993.

Four stress levels We summarise a similar investigation for $k = 4$; Table 7.17 shows how the entropy values and the maximum absolute distances vary when we have an equal pro-

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.09999	2909.31381	-1.98433	0.00605
0.5	0.49935	2904.90486	-1.97164	0.03083
1	0.99491	2899.38669	-1.95578	0.06304
1.5	1.48327	2893.92100	-1.94007	0.09642
2	1.96158	2888.56569	-1.92467	0.13073
2.5	2.42768	2883.36866	-1.90971	0.16569
3	2.88005	2878.36636	-1.89531	0.20101
3.5	3.31777	2873.58396	-1.88151	0.23641
4	3.74043	2869.03648	-1.86838	0.27161
4.5	4.14795	2864.73051	-1.85593	0.30636
5	4.54057	2860.66609	-1.84415	0.34045

Table 7.11: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(3 : 2 : 1)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.09999	2895.39686	-1.92734	0.00706
0.5	0.49925	2890.92111	-1.91489	0.03582
1	0.99420	2885.36499	-1.89944	0.07277
1.5	1.48118	2879.91470	-1.88428	0.11057
2	1.95732	2874.62635	-1.86958	0.14891
2.5	2.42056	2869.54237	-1.85543	0.18745
3	2.86958	2864.69162	-1.84192	0.22589
3.5	3.30370	2860.09096	-1.82908	0.26392
4	3.72266	2855.74743	-1.81695	0.30128
4.5	4.12657	2851.66051	-1.80552	0.33786
5	4.51577	2847.82431	-1.79477	0.37340

Table 7.12: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(2 : 3 : 1)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.09999	3015.56485	-2.24248	0.00481
0.5	0.49896	3013.69144	-2.23664	0.02438
1	0.99181	3011.24351	-2.22910	0.04951
1.5	1.47300	3008.70490	-2.22137	0.07526
2	1.93799	3006.10678	-2.21352	0.10146
2.5	2.38343	3003.47986	-2.20563	0.12792
3	2.80715	3000.85232	-2.19779	0.15449
3.5	3.20805	2998.24850	-2.19005	0.18097
4	3.58585	2995.68841	-2.18247	0.20719
4.5	3.94092	2993.18777	-2.17509	0.23300
5	4.27402	2990.75834	-2.16793	0.25823

Table 7.13: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1 : 3 : 2)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.10000	3059.31363	-2.41852	0.00695
0.5	0.49930	3059.37177	-2.41778	0.03454
1	0.99424	3059.44918	-2.41684	0.06847
1.5	1.48021	3059.53146	-2.41590	0.10160
2	1.95259	3059.61793	-2.41497	0.13376
2.5	2.40716	3059.70773	-2.41403	0.16481
3	2.84040	3059.79987	-2.41310	0.19463
3.5	3.24973	3059.89327	-2.41218	0.22317
4	3.63354	3059.98691	-2.41127	0.25038
4.5	3.99121	3060.07983	-2.41038	0.27626
5	4.32287	3060.17122	-2.40949	0.30082

Table 7.14: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1 : 1 : 3)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.10000	2956.32648	-2.13561	0.00549
0.5	0.49946	2952.61959	-2.12472	0.02705
1	0.99563	2947.86824	-2.11077	0.05300
1.5	1.48529	2943.02993	-2.09659	0.07775
2	1.96544	2938.15449	-2.08230	0.10123
2.5	2.43351	2933.29141	-2.06804	0.12342
3	2.88745	2928.48652	-2.05396	0.14600
3.5	3.32579	2923.77962	-2.04015	0.17370
4	3.74761	2919.20311	-2.02671	0.20185
4.5	4.15246	2914.78156	-2.01371	0.23025
5	4.54031	2910.53212	-2.00120	0.25869

Table 7.15: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(3 : 1 : 1)$.

B_p	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.09999	2933.12724	-2.01369	0.00682
0.5	0.49913	2928.98783	-2.00238	0.03452
1	0.99328	2923.75742	-1.98813	0.06999
1.5	1.47819	2918.52871	-1.97391	0.10615
2	1.95058	2913.36334	-1.95989	0.14271
2.5	2.40811	2908.31265	-1.94619	0.17939
3	2.84933	2903.41643	-1.93292	0.21591
3.5	3.27350	2898.70326	-1.92014	0.25198
4	3.68043	2894.19181	-1.90790	0.28739
4.5	4.07032	2889.89248	-1.89624	0.32205
5	4.44360	2885.80922	-1.88516	0.35573

Table 7.16: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 3$ and the ratio $(1 : 3 : 1)$.

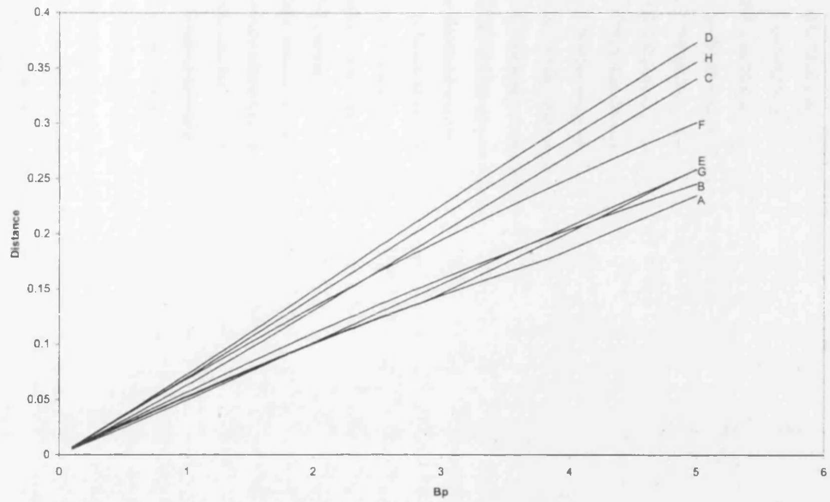


Figure 7.1: Plots of the maximum absolute distance between cdfs for the true Weibull Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios and three stress levels. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

portion of observations at each stress level. We observe similar results to when we examined three stress level, and the maximum absolute distance between the four distribution functions increases, as the parameter B_P also increases. Tables 7.18, 7.19, 7.20, 7.21, 7.22 and 7.23, respectively, summarise similar investigations for the ratios (1 : 1 : 1 : 4), (4 : 1 : 1 : 1), (1 : 2 : 2 : 1), (2 : 1 : 1 : 2), (1 : 2 : 3 : 4) and (4 : 3 : 2 : 1). We summarise these distances for all ratios in Figure 7.2. The largest distances are observed for the ratio (4 : 3 : 2 : 1) and following closely from this (1 : 1 : 1 : 4), whilst the smallest distances occur when we have equal sample sizes across stress levels or considerably higher proportions at the lower stress level, and then equal proportions across the remainder. In practice, we would prefer to have a much larger proportion of observations at higher stresses, since failure times are generally observed earlier, thus reducing the time the experiment takes to run.

7.1.2 Fitting G_{WP} to G_{WA}

Simulation studies

We run simulations to observe the behaviour of the MLEs from the true Weibull Arrhenius and mis-specified Weibull Log-linear distributions, when we allow sample sizes, parameter values and stress levels to vary; choices for these have been outlined in Section 5.1.2 above. We also set B_A equal to 0.5, 1, 2 and 3, but include results only for $B_A = 2$. These are summarised in Tables 7.24 and 7.25 for $k = 2$, Tables 7.26 and 7.27 for $k = 3$, and Tables 7.28 and 7.29 for $k = 4$. The tables show similar results to those obtained when we

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Max dist1
0.1	0.09999	2979.59854	-2.18326	0.00522
0.5	0.49934	2977.87563	-2.17821	0.02591
1	0.99470	2975.68096	-2.17179	0.05127
1.5	1.48221	2973.45862	-2.16530	0.07598
2	1.95828	2971.22852	-2.15880	0.09996
2.5	2.41980	2969.01015	-2.15234	0.12317
3	2.86429	2966.82152	-2.14596	0.14556
3.5	3.28990	2964.67833	-2.13971	0.16712
4	3.69544	2962.59358	-2.13362	0.18785
4.5	4.08028	2960.57739	-2.12772	0.20774
5	4.44430	2958.63707	-2.12203	0.22681

Table 7.17: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(1:1:1:1)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.10000	3062.49878	-2.42293	0.00683
0.5	0.49943	3063.07211	-2.42353	0.03399
1	0.99531	3063.85252	-2.42442	0.06746
1.5	1.48392	3064.70324	-2.42545	0.10023
2	1.96149	3065.62157	-2.42664	0.13215
2.5	2.42449	3066.60245	-2.42796	0.16309
3	2.86981	3067.63868	-2.42941	0.19293
3.5	3.29495	3068.72142	-2.43097	0.22161
4	3.69811	3069.84073	-2.43262	0.24907
4.5	4.07822	3070.98623	-2.43434	0.27528
5	4.43487	3072.14172	-2.43611	0.30026

Table 7.18: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(1:1:1:4)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.10000	2948.17058	-2.11351	0.00539
0.5	0.49961	2945.71694	-2.10633	0.02667
1	0.99686	2942.60080	-2.09722	0.05262
1.5	1.48940	2939.44850	-2.08801	0.07776
2	1.97496	2936.28066	-2.07876	0.10202
2.5	2.45145	2933.11802	-2.06953	0.12536
3	2.91704	2929.98049	-2.06036	0.14776
3.5	3.37019	2926.88640	-2.05132	0.16920
4	3.80970	2923.85188	-2.04244	0.19478
4.5	4.23469	2920.89052	-2.03377	0.22090
5	4.64460	2918.01323	-2.02534	0.24698

Table 7.19: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(4:1:1:1)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.09999	2971.03954	-2.13772	0.00520
0.5	0.49935	2969.71297	-2.13372	0.02622
1	0.99488	2968.04240	-2.12870	0.05287
1.5	1.48292	2966.37035	-2.12367	0.07983
2	1.96021	2964.70946	-2.11869	0.10696
2.5	2.42396	2963.07127	-2.11377	0.13412
3	2.87196	2961.46582	-2.10895	0.16117
3.5	3.30257	2959.90142	-2.10425	0.18800
4	3.71473	2958.38458	-2.09968	0.21447
4.5	4.10783	2956.92006	-2.09525	0.24049
5	4.48169	2955.51109	-2.09099	0.26594

Table 7.20: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(1:2:2:1)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.10000	2999.05357	-2.25696	0.00624
0.5	0.49942	2997.44597	-2.25222	0.03101
1	0.99532	2995.37518	-2.24613	0.06144
1.5	1.48409	2993.25109	-2.23990	0.09115
2	1.96220	2991.09178	-2.23358	0.12002
2.5	2.42642	2988.91667	-2.22722	0.14796
3	2.87398	2986.74536	-2.22087	0.17488
3.5	3.30271	2984.59649	-2.21458	0.20074
4	3.71107	2982.48675	-2.20840	0.22552
4.5	4.09812	2980.43030	-2.20237	0.24921
5	4.46350	2978.43844	-2.19651	0.27182

Table 7.21: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(2:1:1:2)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.09999	3066.33461	-2.40010	0.00566
0.5	0.49935	3067.63534	-2.40225	0.02810
1	0.99478	3069.40663	-2.40527	0.05566
1.5	1.48241	3071.33820	-2.40866	0.08254
2	1.95857	3073.42471	-2.41240	0.10866
2.5	2.42005	3075.65689	-2.41648	0.13392
3	2.86422	3078.02210	-2.42088	0.15829
3.5	3.28913	3080.50521	-2.42556	0.18172
4	3.69348	3083.08954	-2.43049	0.20419
4.5	4.07658	3085.75771	-2.43563	0.22571
5	4.43831	3088.49248	-2.44094	0.24627

Table 7.22: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(1:2:3:4)$.

B_P	$\widetilde{B}_A$	$\widetilde{\beta}_{WA}$	$\widetilde{\alpha}_{WA}$	Distance
0.1	0.10000	2905.23726	-1.97177	0.00619
0.5	0.49950	2902.59972	-1.96427	0.03129
1	0.99603	2899.30190	-1.95489	0.06332
1.5	1.48681	2896.02619	-1.94558	0.09593
2	1.96935	2892.79551	-1.93639	0.12893
2.5	2.44156	2889.63008	-1.92739	0.16213
3	2.90172	2886.54680	-1.91861	0.19533
3.5	3.34854	2883.55906	-1.91009	0.22833
4	3.78113	2880.67675	-1.90187	0.26096
4.5	4.19892	2877.90657	-1.89395	0.29302
5	4.60162	2875.25239	-1.88635	0.32443

Table 7.23: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WP} with $k = 4$ and the ratio $(4:3:2:1)$.

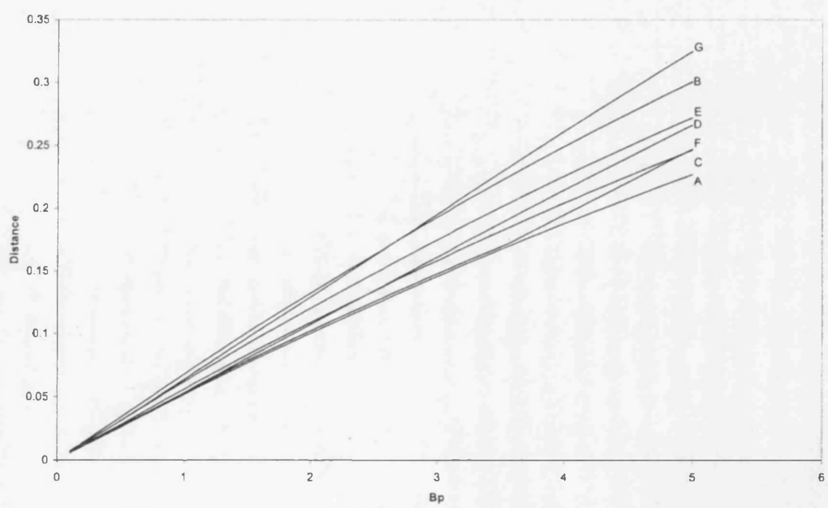


Figure 7.2: Plots of the maximum absolute distance between cdfs for the true Weibull Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios and four stress levels. Here, A=$(1:1:1:1)$, B=$(1:1:1:4)$, C=$(4:1:1:1)$, D=$(1:2:2:1)$, E=$(2:1:1:2)$, F=$(1:2:3:4)$, G=$(4:3:2:1)$.

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_P$	2.0381	2.0193	2.0077	2.0044
$\widetilde{B}_P$	2	2	2	2
S	0.1621	0.1138	0.0701	0.0495
T	0.1559	0.1103	0.0697	0.0493
$\widehat{\alpha}_{WP}$	8.3265	8.3286	8.3304	8.3300
$\widetilde{\alpha}_{WP}$	8.3301	8.3301	8.3301	8.3301
S	0.0985	0.0698	0.0445	0.0309
T	0.0986	0.0697	0.0441	0.0312
$\widehat{\beta}_{WP}$	-0.0153	-0.0154	-0.0154	-0.0154
$\widetilde{\beta}_{WP}$	-0.0153	-0.0153	-0.0153	-0.0153
S	0.0007	0.0005	0.0003	0.0002
T	0.0007	0.0005	0.0003	0.0002
$\widehat{B}_{W,10(P)}$	636.4742	630.8779	627.6892	626.3308
S	79.6747	56.3265	35.3239	24.9241
T	78.6871	55.6402	35.1899	24.8830
Pr (Fit G_{WP})	-	-	-	-
B_A	2.0381	2.0193	2.0077	2.0044
S	0.1621	0.1138	0.0701	0.0495
T	0.1455	0.1029	0.0651	0.0460
$\widehat{\alpha}_{WA}$	0.2979	0.2978	0.2977	0.2993
S	0.2740	0.1917	0.1209	0.0855
T	0.2705	0.1913	0.1210	0.0856
$\widehat{\beta}_{WA}$	2346.5665	2347.2060	2347.7471	2347.1541
S	102.9408	72.1963	45.7100	32.1537
T	101.9376	72.0808	45.5879	32.2355
$\widehat{B}_{W,10(A)}$	636.4742	630.8779	627.6892	626.3308
S	79.6747	56.3265	35.3239	24.9241
T	74.1218	53.1234	33.5982	23.7575

Table 7.24: Fitting G_{WP} to G_{WA} for $k = 2$, $B_A = 2$ with equal allocations. We show the sample means and standard error of parameters, where figures are based on at least 10000 replications.

mis-specified G_{WA}, and fitted this to data from G_{WP}. As expected, we see an improved agreement between sample and theoretical standard errors of true and mis-specified MLEs as the sample size increases; the average MLEs from both distributions also tend to their theoretical counterparts as n increases. Generally, there is good agreement between the estimates for B_{10} from the true and mis-specified models, and there is no difference between these estimates for $k = 2$. This is consistent with other results on mis-specification, as two stress levels is just a reparameterisation of the original model. The probability of favouring the mis-specified distribution is as high as 24% for small sample sizes, and this figure increases for unequal allocations with more observations placed at the higher stress levels.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_P$	2.0202	2.0212	2.0202	2.0211
$\widetilde{B}_P$	2	2	2	2
S	0.1132	0.1122	0.1133	0.1131
T	0.1103	0.1103	0.1103	0.1103
$\widehat{\alpha}_{WP}$	8.3180	8.3324	8.3255	8.3317
$\widetilde{\alpha}_{WP}$	8.3301	8.3301	8.3301	8.3301
S	0.1355	0.0617	0.0964	0.0612
T	0.1344	0.0615	0.0960	0.0605
$\widehat{\beta}_{WP}$	-0.0153	-0.0154	-0.0153	-0.0154
$\widetilde{\beta}_{WP}$	-0.0153	-0.0153	-0.0153	-0.0153
S	0.0007	0.0007	0.0005	0.0006
T	0.0007	0.0007	0.0005	0.0005
$\widehat{B}_{W,10(P)}$	628.7441	631.6704	630.5538	632.1334
S	77.7826	51.8646	64.0498	53.4358
T	77.6137	51.7454	63.8109	52.6358
Pr (Fit G_{WP})	-	-	-	-
$\widehat{B}_A$	2.0202	2.0212	2.0202	2.0211
S	0.1132	0.1122	0.1133	0.1131
T	0.1029	0.1029	0.1029	0.1029
$\widehat{\alpha}_{WA}$	0.3184	0.2689	0.3042	0.2832
S	0.2489	0.3295	0.2004	0.2440
T	0.2465	0.3260	0.1998	0.2400
$\widehat{\beta}_{WA}$	2338.0900	2356.7734	2344.3978	2352.3661
S	110.0396	110.1228	83.5051	84.5065
T	108.9759	108.9759	83.2317	83.2317
$\widehat{B}_{W,10(A)}$	628.7441	631.6704	630.5538	632.1334
S	77.7826	51.8646	64.0498	53.4358
T	75.8297	49.0286	61.6288	49.9675

Table 7.25: Fitting G_{WP} to G_{WA} for $k = 2$, $B_A = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	25,25 25	50,50 50	100,100 100	300,300 300	500,500 500
$\widehat{B}_P$	2.0228	1.9980	1.9804	1.9734	1.9710
$\widetilde{B}_P$	1.9685	1.9685	1.9685	1.9685	1.9685
S	0.1888	0.1301	0.0877	0.0514	0.0394
T	0.1770	0.1252	0.0885	0.0511	0.0396
$\widehat{\alpha}_{WP}$	8.2976	8.2983	8.2997	8.3020	8.3020
$\widetilde{\alpha}_{WP}$	8.3024	8.3024	8.3024	8.3024	8.3024
S	0.1389	0.0985	0.0695	0.0395	0.0306
T	0.1376	0.0973	0.0688	0.0397	0.0308
$\widehat{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
$\widetilde{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
S	0.0009	0.0007	0.0005	0.0003	0.0002
T	0.0009	0.0007	0.0005	0.0003	0.0002
$\widehat{B}_{W,10(P)}$	607.2656	599.2758	593.5625	591.9563	591.0739
S	93.6898	66.0333	45.6817	26.3826	20.1385
T	91.5024	64.7020	45.7512	26.4145	20.4606
Pr (Fit G_{WP})	0.2379	0.1582	0.0690	0.0051	0.0004
$\widehat{B}_A$	2.0535	2.0287	2.0119	2.0046	2.0023
S	0.1914	0.1322	0.0893	0.0519	0.0404
T	0.1680	0.1188	0.0840	0.0485	0.0376
$\widehat{\alpha}_{WA}$	0.2999	0.3084	0.3034	0.3005	0.3008
S	0.3588	0.2545	0.1786	0.1031	0.0794
T	0.3552	0.2512	0.1776	0.1025	0.0794
$\widehat{\beta}_{WA}$	2345.1428	2343.0559	2345.1118	2346.6902	2346.5941
S	140.0301	99.2835	69.8535	40.1048	30.9767
T	138.6989	98.0749	69.3495	40.0389	31.0140
$\widehat{B}_{W,10(A)}$	641.3109	633.3076	627.9611	626.3098	625.4740
S	98.9828	69.9013	48.3880	27.8577	21.3928
T	93.1172	65.8438	46.5586	26.8806	20.8216

Table 7.26: Fitting G_{WP} to G_{WA} for $k = 3$, $B_A = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	50,100 150	25,200 75	25,25 250	200,50 50
$\widehat{B}_P$	1.9813	1.9791	2.0020	1.9960
$\widetilde{B}_P$	1.9680	1.9644	1.9900	1.9833
S	0.0892	0.0902	0.0915	0.0914
T	0.0884	0.0888	0.0895	0.0892
$\widehat{\alpha}_{WP}$	8.2453	8.1606	8.2619	8.3392
$\widetilde{\alpha}_{WP}$	8.2497	8.1679	8.2736	8.3395
S	0.0913	0.1217	0.1320	0.0534
T	0.0911	0.1214	0.1287	0.0533
$\widehat{\beta}_{WP}$	-0.0152	-0.0150	-0.0151	-0.0158
$\widetilde{\beta}_{WP}$	-0.0152	-0.0151	-0.0151	-0.0158
S	0.0005	0.0008	0.0007	0.0005
T	0.0005	0.0008	0.0007	0.0005
$\widehat{B}_{W,10(P)}$	573.6234	531.3995	595.3167	615.9095
S	50.3970	54.9215	68.0977	43.2393
T	50.6985	54.7462	66.8762	42.8053
Pr (Fit G_{WP})	0.0719	0.0840	0.1904	0.1412
$\widehat{B}_A$	2.0128	2.0129	2.0118	2.0126
S	0.0909	0.0911	0.0922	0.0920
T	0.0840	0.0840	0.0840	0.0840
$\widehat{\alpha}_{WA}$	0.3072	0.3168	0.3184	0.2943
S	0.2005	0.2807	0.2424	0.1996
T	0.1987	0.2796	0.2357	0.2001
$\widehat{\beta}_{WA}$	2343.4641	2339.3615	2338.1588	2348.5166
S	83.8749	118.1342	108.4598	70.2937
T	83.2302	117.7630	105.4986	70.4134
$\widehat{B}_{W,10(A)}$	628.0381	627.3125	626.0745	628.6493
S	56.2532	68.0375	72.6598	43.8610
T	54.9008	66.7847	70.0641	41.3968

Table 7.27: Fitting G_{WP} to G_{WA} for $k = 3$, $B_A = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_P$	2.0197	1.9954	1.9864	1.9811	1.9777
$\widetilde{B}_P$	1.9757	1.9757	1.9757	1.9757	1.9757
S	0.1625	0.1113	0.0785	0.0546	0.0345
T	0.1540	0.1089	0.0770	0.0545	0.0344
$\widehat{\alpha}_{WP}$	8.2978	8.3023	8.3031	8.3048	8.3053
$\widetilde{\alpha}_{WP}$	8.3056	8.3056	8.3056	8.3056	8.3056
S	0.1382	0.0972	0.0686	0.0478	0.0301
T	0.1364	0.0964	0.0682	0.0482	0.0305
$\widehat{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
$\widetilde{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
S	0.0009	0.0006	0.0004	0.0003	0.0002
T	0.0009	0.0006	0.0004	0.0003	0.0002
$\widehat{B}_{W,10(P)}$	606.1152	599.1682	596.0140	594.6784	593.6328
S	87.4284	60.1932	42.7263	29.7821	18.8531
T	84.5166	59.7623	42.2583	29.8811	18.8985
Pr (Fit G_{WP})	0.2419	0.1512	0.0766	0.0182	0.0010
$\widehat{B}_A$	2.0428	2.0192	2.0104	2.0054	2.0020
S	0.1640	0.1122	0.0791	0.0552	0.0349
T	0.1455	0.1029	0.0728	0.0514	0.0325
$\widehat{\alpha}_{WA}$	0.3122	0.3064	0.3043	0.3007	0.3006
S	0.3239	0.2284	0.1617	0.1132	0.0712
T	0.3211	0.2271	0.1606	0.1135	0.0718
$\widehat{\beta}_{WA}$	2340.7306	2343.9868	2344.8986	2346.4937	2346.7080
S	131.0333	92.2486	65.2735	45.5811	28.6823
T	129.6222	91.6568	64.8111	45.8284	28.9844
$\widehat{B}_{W,10(A)}$	637.2949	630.6991	627.6161	626.4867	625.4289
S	92.0613	63.3911	45.0615	31.5047	19.9627
T	86.2528	60.9900	43.1264	30.4950	19.2867

Table 7.28: Fitting G_{WP} to G_{WA} for $k = 4$, $B_A = 2$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 75,75	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_P$	2.0031	1.9970	1.9895	2.0078	2.0001	2.0044	1.9973
$\widetilde{B}_P$	1.9821	1.9748	1.9679	1.9869	1.9799	1.9820	1.9766
S	0.1110	0.1120	0.1113	0.1115	0.1106	0.1112	0.1118
T	0.1093	0.1091	0.1085	0.1096	0.1092	0.1092	0.1091
$\widehat{\alpha}_{WP}$	8.2707	8.3352	8.2204	8.3373	8.3417	8.2640	8.2630
$\widetilde{\alpha}_{WP}$	8.2798	8.3352	8.2276	8.3409	8.3434	8.2742	8.2721
S	0.1319	0.0831	0.1264	0.0816	0.0736	0.1313	0.1320
T	0.1313	0.0831	0.1266	0.0817	0.0732	0.1303	0.1321
$\widehat{\beta}_{WP}$	-0.0153	-0.0160	-0.0151	-0.0158	-0.0160	-0.0152	-0.0156
$\widetilde{\beta}_{WP}$	-0.0154	-0.0160	-0.0152	-0.0159	-0.0160	-0.0153	-0.0156
S	0.0008	0.0006	0.0008	0.0006	0.0006	0.0007	0.0008
T	0.0008	0.0006	0.0008	0.0006	0.0006	0.0007	0.0008
$\widehat{B}_{W,10(P)}$	592.7054	606.5866	564.0093	617.3849	613.4048	592.4093	579.1054
S	70.9571	56.6725	65.2500	57.4491	54.5141	70.8271	68.3876
T	70.7376	56.0200	65.0727	57.1777	54.1759	70.5617	68.1976
Pr(Fit G_{WP})	0.1837	0.1645	0.1263	0.2241	0.1744	0.1901	0.1665
$\widehat{B}_A$	2.0206	2.0214	2.0204	2.0210	2.0199	2.0218	2.0199
S	0.1119	0.1134	0.1128	0.1124	0.1116	0.1123	0.1128
T	0.1029	0.1029	0.1029	0.1029	0.1029	0.1029	0.1029
$\widehat{\alpha}_{WA}$	0.3152	0.2974	0.3127	0.3036	0.3013	0.3200	0.3170
S	0.2663	0.2320	0.2736	0.2136	0.2283	0.2631	0.2926
T	0.2657	0.2306	0.2742	0.2135	0.2268	0.2605	0.2911
$\widehat{\beta}_{WA}$	2339.6592	2347.4699	2340.8172	2344.8152	2345.9673	2337.9543	2338.9902
S	114.6104	88.3982	115.8465	82.5483	84.1094	113.6432	122.5874
T	114.2882	87.9890	116.0506	82.5293	83.5440	112.4357	122.1216
$\widehat{B}_{W,10(A)}$	629.8182	632.1414	630.2064	630.7727	630.9409	629.8806	629.2586
S	76.2945	58.8319	74.7962	58.5944	55.8907	76.3671	75.7513
T	74.3920	55.7990	72.9099	55.9419	52.9770	74.3177	73.7202

Table 7.29: Fitting G_{WP} to G_{WA} for $k = 4$, $B_A = 2$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

The effects of mis-specification

We now investigate whether, for a particular set of Weibull Arrhenius parameters and stress levels, we pay large penalties by mis-specifying the scale-stress model. We do this in our usual way, and the consequences of mis-specification are assessed by deriving the maximum absolute distance between true and mis-specified distribution functions across all stress levels. The larger the distance, the greater the penalty we pay. For example, if we consider our usual set of Weibull Arrhenius parameters and stress levels for $k = 3$, then we have true scale parameters given by

$$\theta_{1A} = 1924.9519,\ \theta_{2A} = 345.9746,\ \theta_{3A} = 192.5329.$$

The entropy values for the Weibull Log-linear parameters are then

$$\widetilde{B}_P = 1.9685,\ \widetilde{\alpha}_{WP} = -0.0156,\ \widetilde{\beta}_{WP} = 8.3024,$$

which correspond to entropy scale parameters given by

$$\widetilde{\theta}_{1P} = 1851.8122,\ \widetilde{\theta}_{2P} = 390.3017,\ \widetilde{\theta}_{3P} = 179.1852.$$

The table of figures, Table 7.30, shows the three plots corresponding to true and mis-specified distribution functions, across the three stress levels. Across all stress levels, there seems to be a reasonable fit between the two distribution functions. The worst case is for the second stress level where we observe a maximum absolute distance of 0.0880. We repeat this investigation for $k = 2$, 3 and 4 stress levels.

Two stress levels When we have two stress levels, the entropy values are not affected by how we arrange the sample or the total sample size used. Thus, entropy values for $(50, 150)$ are the same as those for $(234, 567)$. Also, for all values of B_A, $B_A = \widetilde{B}_P$. In fact, for any value of B_A, using our usual set of two stress levels gives entropy values

$$\widetilde{B}_P = B_A,\ \widetilde{\beta}_{WP} = -0.0154,\ \widetilde{\alpha}_{WP} = 8.3301.$$

The scale parameter for both distributions always match up so a perfect fit is observed between the true and mis-specified models. This will be true regardless of the sample size, true parameter values and how the sample is arranged across the two stress levels. Thus, if we mis-specify the Arrhenius scale-stress relationship and incorrectly use the Log-linear model, then no penalties are paid as a result of this, and an equivalent fit will always be observed.

Three stress levels We now examine the effects of mis-specifying the scale-stress model for three stress levels. We first tabulate the results when we have equal numbers of obser-

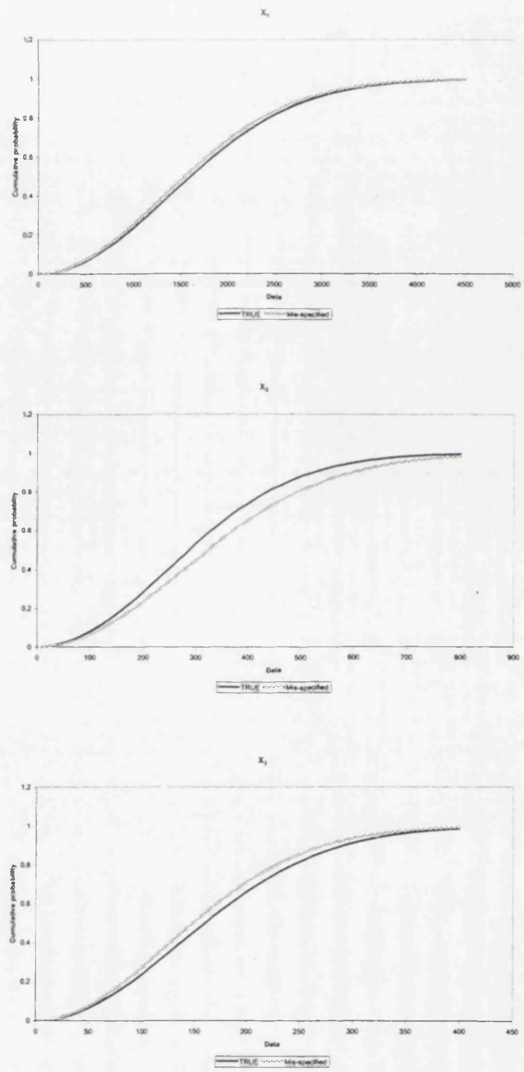


Table 7.30: Comparison between the true Weibull Arrhenius model, with parameters $B_A = 2$, $\alpha_{WA} = 0.3$, $\beta_{WA} = 2347$, and the mis-specified Weibull Log-linear distribution, with entropy values $\widetilde{B}_P = 1.9685$, $\widetilde{\alpha}_{WP} = -0.0156$, $\widetilde{\beta}_{WP} = 8.3024$. Here, $k = 3$, $X_1 + c = 323.16$, $X_2 + c = 423.16$ and $X_3 + c = 473.16$.

B_A	$\widetilde{B}_P$	$\widetilde{\beta}_{WP}$	$\widetilde{\alpha}_{WP}$	Distance
0.1	0.10000	-0.01561	8.30409	0.00430
0.5	0.49947	-0.01560	8.30366	0.02162
1	0.99586	-0.01559	8.30318	0.04355
1.5	1.48636	-0.01558	8.30277	0.06570
2	1.96854	-0.01557	8.30242	0.08798
2.5	2.44043	-0.01556	8.30214	0.11031
3	2.90047	-0.01555	8.30191	0.13260
3.5	3.34753	-0.01555	8.30174	0.15478
4	3.78085	-0.01554	8.30162	0.17678
4.5	4.19998	-0.01553	8.30154	0.19853
5	4.60473	-0.01553	8.30150	0.21997

Table 7.31: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WA} with $k = 3$ and the ratio $(1 : 1 : 1)$.

vations at each stress level and allow B_A to vary from 0.1 to 5 in appropriate increments; these maximum absolute distances are shown in Table 7.31. We observe similar results for when we fitted the Weibull Arrhenius distribution to data with an underlying Weibull Log-linear model. As we increase the shape parameter from the Weibull Arrhenius distribution, the maximum absolute distance between the two distribution functions also increases. If we further increase the acceleration factor then we expect to observe even larger distances between the two distribution functions. For example, if we set

$$B_A = 3,\ \alpha_{WA} = -7,\ \beta_{WA} = 4694,$$

and use our usual set of three stress levels, then entropy values for such a set of parameters are

$$\widetilde{B}_P = 2.6856,\ \widetilde{\alpha}_{WP} = 9.0030,\ \widetilde{\beta}_{WP} = -0.0310,$$

and the maximum absolute distance is 0.2617. When we compare this to a figure of 0.1326 for our usual acceleration factor of around 10, then we see a large increase in the distance between the two distribution functions as the acceleration factor is highered. We also examine the effects that varying the allocation of items at each stress level has on the maximum absolute distance between the two distribution functions. Figure 7.3 shows how the distance varies as B_A is changed, for the ratios $(1 : 2 : 3)$, $(3 : 2 : 1)$, $(2 : 3 : 1)$, $(1 : 3 : 2)$, $(1 : 1 : 3)$, $(3 : 1 : 1)$ and $(1 : 3 : 1)$. Surprisingly, the results are quite different to the previous scenario whereby we mis-specified the Weibull Arrhenius distribution and fitted this to data from a Weibull Log-linear model. The worst case scenario occurs when we have most of the observations at the highest stress level. The best line, in terms of having the smallest maximum absolute distance across all stress levels, occurs for the ratio $(1 : 3 : 2)$ or when most of the observations lie in the middle stress. This is quite different to the results

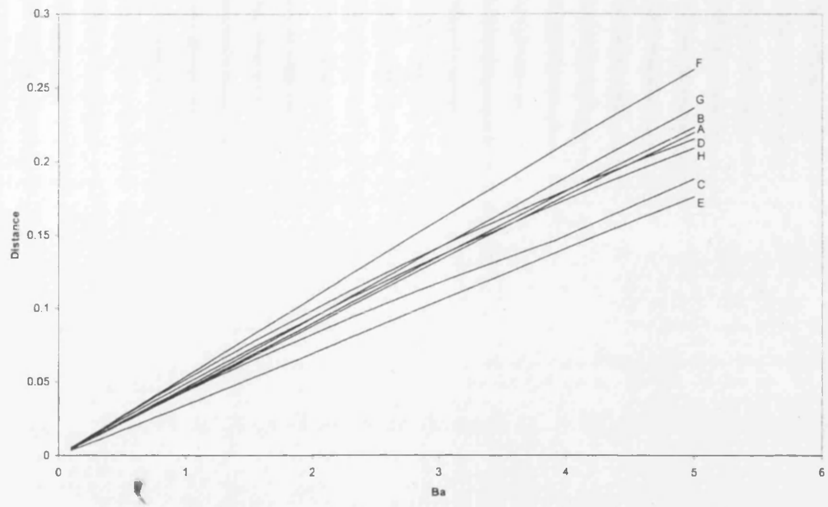


Figure 7.3: Plots of the maximum absolute distance between cdfs for the true Weibull Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios and three stress levels. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

obtained for the previous scenario whereby the best results were observed when we had equal ratios or most of the observations at the lowest stress level.

Four stress levels We finally consider examining results for our usual set of four stress levels, and tabulate the maximum absolute distance when we have equal sample sizes across all levels. Table 7.32 shows similar results as those above, and we see an increase in the maximum absolute distance between the two cdfs as B_A increases. We now construct a plot of the maximum absolute distances between the cdfs as we allow the proportion of observations at each stress level to vary. Figure 7.4 shows these distances for the ratios (1 : 1 : 1 : 1), (1 : 1 : 1 : 4), (4 : 1 : 1 : 1), (1 : 2 : 2 : 1), (2 : 1 : 1 : 2), (1 : 2 : 3 : 4) and (4 : 3 : 2 : 1). We again see that the ratio (1 : 2 : 2 : 1) gives the best fit in terms of having the smallest maximum absolute distance between the two distribution functions across the four stress levels. The worst case scenario is when we have most of the observations at the highest stress level. The outcome of this investigation provides quite contrasting results to the scenario when we fitted Weibull Arrhenius to data with an underlying Weibull Log-linear model. In that case we saw that the worst penalties were paid with most of our observations at the middle stress levels. The results for this case show quite the contrary, and we see the best fit (in terms of having the smallest maximum absolute distance) when we have larger proportions of observations at the middle levels. This is not consistent with simulation studies where we had larger standard errors for simulations with more observations at the middle and highest stress levels, and may require further investigation.

B_A	$\widetilde{B}_P$	$\widetilde{\beta}_{WP}$	$\widetilde{\alpha}_{WP}$	Distance
0.1	0.10000	-0.01565	8.30621	0.00414
0.5	0.49960	-0.01564	8.30606	0.02081
1	0.99684	-0.01564	8.30589	0.04181
1.5	1.48952	-0.01563	8.30574	0.06294
2	1.97566	-0.01563	8.30560	0.08414
2.5	2.45352	-0.01562	8.30549	0.10534
3	2.92166	-0.01562	8.30540	0.12649
3.5	3.37890	-0.01561	8.30533	0.14752
4	3.82433	-0.01561	8.30527	0.16838
4.5	4.25729	-0.01560	8.30524	0.18903
5	4.67733	-0.01560	8.30522	0.20940

Table 7.32: Maximum distances for G_{WP} and G_{WA}. Data is from G_{WA} with $k = 4$ and the ratio $(1 : 1 : 1 : 1)$.

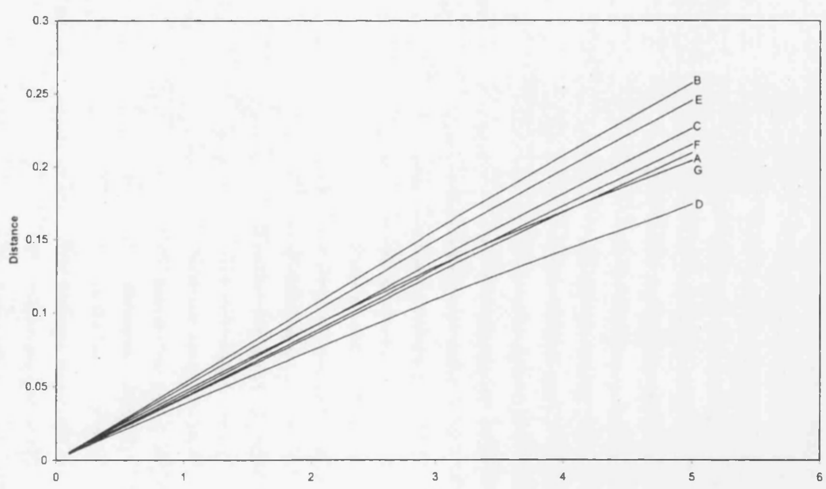


Figure 7.4: Plots of the maximum absolute distance between cdfs for the true Weibull Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios and four stress levels. Here, A=$(1 : 1 : 1 : 1)$, B=$(1 : 1 : 1 : 4)$, C=$(4 : 1 : 1 : 1)$, D=$(1 : 2 : 2 : 1)$, E=$(2 : 1 : 1 : 2)$, F=$(1 : 2 : 3 : 4)$, G=$(4 : 3 : 2 : 1)$.

7.2 Getting the distribution wrong

We now assume that we have chosen the correct relationship linking scale and stress, but mis-specify the underlying distribution function. We use the Log-linear scale-stress relationship, and, due to its popularity, always keep the Weibull as the mis-specified distribution. We look at the effects of fitting this, to data from a Burr, Gamma or Lognormal distribution.

7.2.1 Fitting G_{WP} to G_{BP}

Simulation studies

In order to check theoretical results, and provide an idea of how the mis-specified MLEs respond to changes in n, n_i and k, we run simulations for varying τ and a. We choose 2, 3 and 4 stress values, identical to those used before for previous examples, allow the sample sizes to vary in a similar manner, and set the parameters in the scale-stress relationship equal to the values outlined in Section 5.1.2 above. Since we now have two shape parameters that can vary, we run simulations for $\tau = 1$ and $a = 4$, $\tau = 4$ and $a = 3$, $\tau = 3$ and $a = 4$, and finally $\tau = 4$ and $a = 1$, although, due to the size of the tables, only include results for $\tau = 3$ and $a = 4$. We note that the true value of B_{10} for this set of parameters is 327.7264, and the entropy values are $\widetilde{B}_P = 2.55279$, $\widetilde{\alpha}_{WP} = 7.60343$, which, as we have previously shown, do not depend on stress levels or sample sizes. Tables 7.33 and 7.34 summarise the results for $k = 2$, Tables 7.35 and 7.36 for $k = 3$, and Tables 7.37 and 7.38 for $k = 4$. We note a number of points concerning our results. Firstly, for small sample sizes, across all numbers of stress levels, and in particular for the first three sets of parameter values, the estimate for a and standard error of $\widehat{a}$ is large, and does not come close to its theoretical counterpart. The figures do decrease as the sample size is highered, but even for large samples like 2000, this agreement is not excellent. Just as in the non-accelerated case, the reason for such an occurrence may be linked to the fact that the Weibull Log-linear distribution is embedded in the Burr Log-linear model, and, if the Weibull almost provides an improved fit over the Burr, we observe $\widehat{a}$ tending to infinity. This occurs far more often for small sample sizes, as we have less information to base our analyses on. When we examine results for the mis-specified distribution, we see that for small sample sizes, the agreement between the Weibull MLEs and their theoretical counterparts is very good. Sample standard errors also match up well with their theoretical counterparts. True and mis-specified estimates for B_{10} never really come close, and we see that $\widehat{B}_{W,10}$ is always less than the same estimate from the Burr. Thus, we always under estimate the time to which 10% of observations fail if we mis-specify the distribution function. Generally, for both true and mis-specified distributions, the standard error for $\widehat{B}_{10}$ increases as the number of observations at the middle and higher stress levels is raised. Patterns in the results for other choices of shape parameters are more variable than in the Weibull case, but details of the results are omitted here due to lack of space. There are also some theoretical issues with some choices of parameter values, for

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_P$	2.6298	2.5914	2.5698	2.5614
S	0.2282	0.1608	0.1027	0.0723
T	0.2335	0.1651	0.1044	0.0738
$\widehat{\alpha}_{WP}$	7.6010	7.6019	7.6035	7.6031
S	0.0833	0.0595	0.0378	0.0266
T	0.0845	0.0598	0.0378	0.0267
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0006	0.0004	0.0003	0.0002
T	0.0006	0.0004	0.0003	0.0002
$\widehat{B}_{W,10}$	312.1009	308.7977	307.1199	306.2268
S	29.9839	21.3281	13.6488	9.6123
T	30.9027	21.8515	13.8201	9.7723
Pr (Fit G_{WP})	0.1855	0.0663	0.0049	0.0001
$\widehat{\tau}$	3.1359	3.0444	3.0082	3.0033
S	0.4083	0.2835	0.1796	0.1258
T	0.3948	0.2791	0.1765	0.1248
$\widehat{a}$	11.8646	10.5032	5.4076	4.4472
S	67.5547	65.4313	8.1228	2.0506
T	3.3055	2.3373	1.4783	1.0453
$\widehat{\alpha}_B$	8.0555	8.0798	8.0473	8.0218
S	0.4130	0.3496	0.2144	0.1355
T	0.3779	0.2672	0.1690	0.1195
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0005	0.0004	0.0002	0.0002
T	0.0005	0.0004	0.0002	0.0002
$\widehat{B}_{B,10}$	333.6839	329.7841	328.3594	327.9286
S	31.5683	22.2561	14.2607	9.9447
T	31.5477	22.3076	14.1085	9.9762

Table 7.33: Fitting G_{WP} to G_{BP} for $k = 2$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

which the computation of certain terms requires further attention.

The effects of mis-specification

We examine the effects of wrongly using an accelerated Weibull distribution to represent a data set with an underlying Burr model. As in previous cases, we illustrate this with one particular example, and then go on and examine the fit between the two distribution functions for a wider range of Burr parameter values. In keeping with previous examples, we set $\tau = 3$, $a = 4$ and $\phi_{iP} = \exp(8 - 0.02X_i)$ and use our usual set of three stress levels. The corresponding entropy values are

$$\widetilde{B}_P = 2.55279,\ \widetilde{\alpha}_{WP} = 7.60343,\ \widetilde{\beta}_{WP} = -0.02.$$

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_P$	2.5918	2.5926	2.5926	2.5924
S	0.1613	0.1624	0.1638	0.1593
T	0.1651	0.1651	0.1651	0.1651
$\widehat{\alpha}_{WP}$	7.5928	7.6056	7.5985	7.6039
S	0.1154	0.0514	0.0835	0.0510
T	0.1180	0.0523	0.0836	0.0513
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0006	0.0006	0.0005	0.0005
T	0.0006	0.0006	0.0005	0.0005
$\widehat{B}_{W,10}$	307.6902	309.1381	308.4625	309.0870
S	31.5696	19.5316	25.7218	19.6756
T	32.0250	19.9796	25.6942	20.4104
Pr (Fit G_{WP})	0.0665	0.0685	0.0614	0.0674
$\widehat{\tau}$	3.0463	3.0401	3.0488	3.0462
S	0.2811	0.2820	0.2862	0.2824
T	0.2791	0.2791	0.2791	0.2791
$\widehat{a}$	9.3728	10.0749	9.9448	9.6323
S	48.2151	51.8504	72.4974	52.1763
T	2.3373	2.3373	2.3373	2.3373
$\widehat{\alpha}_B$	8.0726	8.0860	8.0743	8.0779
S	0.3542	0.3471	0.3510	0.3425
T	0.2832	0.2659	0.2726	0.2657
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0006	0.0006	0.0005	0.0005
T	0.0006	0.0006	0.0004	0.0004
$\widehat{B}_{B,10}$	329.9409	329.7020	329.8102	329.9301
S	32.5149	20.3483	26.5231	20.6494
T	32.1660	20.5161	26.0122	20.9274

Table 7.34: Fitting G_{WP} to G_{BP} for $k = 2$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	25,25 25	50,50 50	100,100 100	300,300 300	500,500 500
$\widehat{B}_P$	2.6542	2.6046	2.5795	2.5625	2.5582
S	0.2671	0.1858	0.1321	0.0762	0.0588
T	0.2697	0.1907	0.1348	0.0778	0.0603
$\widehat{\alpha}_{WP}$	7.5978	7.6009	7.6027	7.6028	7.6032
S	0.1169	0.0830	0.0586	0.0342	0.0266
T	0.1192	0.0843	0.0596	0.0344	0.0267
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0008	0.0006	0.0004	0.0002	0.0002
T	0.0008	0.0006	0.0004	0.0002	0.0002
$\widehat{B}_{W,10}$	313.9109	310.0093	307.9299	306.3005	305.9356
S	37.1631	26.3129	18.6825	10.9318	8.4108
T	38.2714	27.0620	19.1357	11.0480	8.5578
Pr (Fit G_{WP})	0.2457	0.0996	0.0250	0.0005	0
$\widehat{\tau}$	3.1908	3.0727	3.0173	3.0040	3.0026
S	0.4669	0.3282	0.2262	0.1345	0.1033
T	0.4558	0.3223	0.2279	0.1316	0.1019
$\widehat{a}$	13.1306	11.4254	7.5900	4.5614	4.2541
S	115.6727	59.1836	32.7202	5.9265	1.1380
T	3.8169	2.6989	1.9084	1.1018	0.8535
$\widehat{\alpha}_B$	8.0340	8.0789	8.0704	8.0243	8.0133
S	0.4365	0.3853	0.2897	0.1471	0.1074
T	0.4409	0.3118	0.2205	0.1273	0.0956
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0008	0.0005	0.0004	0.0002	0.0002
T	0.0008	0.0005	0.0004	0.0002	0.0002
$\widehat{B}_{B,10}$	335.9790	331.1162	328.7905	327.9630	327.8464
S	38.9279	27.3915	19.2843	11.2415	8.7046
T	38.9171	27.5185	19.4585	11.2344	8.7021

Table 7.35: Fitting G_{WP} to G_{BP} for $k = 3$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	50,100	25,200	25,25	200,50
n_3	150	75	250	50
$\widehat{B}_P$	2.5805	2.5770	2.5783	2.5814
S	0.1315	0.1311	0.1306	0.1318
T	0.1348	0.1348	0.1348	0.1348
$\widehat{\alpha}_{WP}$	7.5995	7.5991	7.5924	7.6036
S	0.0778	0.1046	0.1112	0.0454
T	0.0795	0.1065	0.1128	0.0454
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0199	-0.0200
S	0.0005	0.0007	0.0006	0.0004
T	0.0005	0.0007	0.0006	0.0004
$\widehat{B}_{W,10}$	307.5922	307.4559	306.3173	308.1406
S	22.0056	26.1821	29.0284	16.4577
T	22.5490	26.7585	29.2757	16.9160
Pr (Fit G_{WP})	0.0255	0.0265	0.0257	0.0245
$\widehat{\tau}$	3.0222	3.0195	3.0207	3.0219
S	0.2318	0.2291	0.2305	0.2331
T	0.2279	0.2279	0.2279	0.2279
$\widehat{a}$	7.8323	7.1670	8.0072	7.8872
S	37.8120	27.3575	58.6574	38.5970
T	1.9084	1.9084	1.9084	1.9084
$\widehat{\alpha}_B$	8.0679	8.0662	8.0642	8.0729
S	0.2977	0.2975	0.3077	0.2934
T	0.2257	0.2350	0.2375	0.2176
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0004	0.0006	0.0006	0.0004
T	0.0004	0.0006	0.0006	0.0004
$\widehat{B}_{B,10}$	328.8534	329.0663	328.3528	329.0491
S	22.7787	26.9757	29.5692	17.3485
T	22.7604	26.8589	29.3188	17.3271

Table 7.36: Fitting G_{WP} to G_{BP} for $k = 3$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_P$	2.6291	2.5914	2.5739	2.5628	2.5578
S	0.2264	0.1596	0.1127	0.0813	0.0514
T	0.2335	0.1651	0.1168	0.0826	0.0522
$\widehat{\alpha}_{WP}$	7.5967	7.5998	7.6010	7.6026	7.6031
S	0.1167	0.0834	0.0582	0.0416	0.0264
T	0.1187	0.0839	0.0593	0.0420	0.0265
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0008	0.0005	0.0004	0.0003	0.0002
T	0.0008	0.0005	0.0004	0.0003	0.0002
$\widehat{B}_{W,10}$	311.7542	308.6489	307.0921	306.3461	305.9219
S	34.7724	24.4734	17.3044	12.4251	7.9090
T	35.6115	25.1812	17.8058	12.5906	7.9630
Pr (Fit G_{WP})	0.1821	0.0657	0.0128	0.0005	0
$\widehat{\tau}$	3.1353	3.0432	3.0095	3.0057	3.0002
S	0.4048	0.2810	0.1979	0.1422	0.0887
T	0.3948	0.2791	0.1974	0.1396	0.0883
$\widehat{a}$	11.4334	9.8036	6.3429	4.5801	4.1969
S	57.7325	65.0065	24.2344	2.9274	0.8851
T	3.3055	2.3373	1.6528	1.1687	0.7391
$\widehat{\alpha}_B$	8.0542	8.0779	8.0579	8.0260	8.0120
S	0.4200	0.3461	0.2491	0.1566	0.0907
T	0.3857	0.2727	0.1928	0.1364	0.0862
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0007	0.0005	0.0004	0.0002	0.0002
T	0.0007	0.0005	0.0004	0.0003	0.0002
$\widehat{B}_{B,10}$	333.9442	329.7909	328.2041	328.1436	327.8415
S	36.1414	25.2620	17.9675	12.7704	8.0941
T	36.0850	25.5159	18.0425	12.7580	8.0688

Table 7.37: Fitting G_{WP} to G_{BP} for $k = 4$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 75,75	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_P$	2.5936	2.5938	2.5918	2.5917	2.5946	2.5883	2.5912
S	0.1585	0.1603	0.1601	0.1607	0.1605	0.1598	0.1599
T	0.1651	0.1651	0.1651	0.1651	0.1651	0.1651	0.1651
$\widehat{\alpha}_{WP}$	7.5946	7.6007	7.5975	7.6019	7.6025	7.5944	7.5963
S	0.1124	0.0705	0.1083	0.0698	0.0619	0.1127	0.1142
T	0.1150	0.0718	0.1107	0.0706	0.0629	0.1141	0.1157
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0007	0.0005	0.0006	0.0005	0.0005	0.0006	0.0007
T	0.0007	0.0005	0.0007	0.0005	0.0005	0.0006	0.0007
$\widehat{B}_{W,10}$	308.1957	308.8979	308.5593	308.9323	309.2557	307.6599	308.3556
S	30.1921	22.3660	29.2077	22.4267	21.2181	30.2668	29.9532
T	30.8962	22.9475	29.8089	23.0716	21.7282	30.8533	30.3856
Pr(Fit G_{WP})	0.0681	0.0666	0.0645	0.0638	0.0689	0.0638	0.0656
$\widehat{\tau}$	3.0479	3.0430	3.0471	3.0478	3.0465	3.0405	3.0432
S	0.2834	0.2832	0.2837	0.2819	0.2841	0.2811	0.2796
T	0.2791	0.2791	0.2791	0.2791	0.2791	0.2791	0.2791
$\widehat{a}$	9.4311	10.2111	11.1259	9.4127	10.3612	8.9349	9.8802
S	51.7178	77.6502	264.6415	51.4263	83.3490	39.5936	59.0093
T	2.3373	2.3373	2.3373	2.3373	2.3373	2.3373	2.3373
$\widehat{\alpha}_B$	8.0727	8.0813	8.0740	8.0746	8.0805	8.0729	8.0760
S	0.3547	0.3458	0.3519	0.3444	0.3487	0.3495	0.3558
T	0.2822	0.2697	0.2807	0.2695	0.2678	0.2819	0.2825
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0006	0.0005	0.0006	0.0005	0.0005	0.0006	0.0007
T	0.0006	0.0005	0.0006	0.0005	0.0005	0.0006	0.0007
$\widehat{B}_{B,10}$	329.8489	329.7177	330.1545	330.1400	330.1124	329.4424	329.8482
S	30.9942	23.3929	30.3286	23.2790	22.1138	31.3658	30.8668
T	31.0655	23.3609	30.0066	23.4804	22.1892	31.0237	30.5681

Table 7.38: Fitting G_{WP} to G_{BP} for $k = 4$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

As with our other mis-specified scenarios, we compare the fit of both distributions by examining the maximum absolute distance between the distribution functions across all stress levels. The true scale parameters are

$$\phi_{1P} = 1096.63316,\ \phi_{2P} = 148.41316,\ \phi_{3P} = 54.59815,$$

and the mis-specified counterparts from G_{WP} are given by

$$\widetilde{\theta}_{1P} = 737.62089,\ \widetilde{\theta}_{2P} = 99.82613,\ \widetilde{\theta}_{3P} = 36.72398.$$

So we are always under-estimating the scale parameters from the mis-specified distribution. We construct three plots of the true and mis-specified distribution functions, corresponding to the three stress levels. The table of figures, Table 7.39, illustrate how well the two distribution functions match up. In fact, the maximum absolute distance between true and mis-specified distribution functions is just 0.023145, a figure which remains constant across all three stress levels, and is also identical to the equivalent non-accelerated scenario. This fact can be explained as follows : we have seen in the non-accelerated case, that the maximum absolute distance between the Burr and mis-specified Weibull models is independent of the scale parameter. We have also seen (and proved in the non-accelerated case) that $\widetilde{B}_P$ is independent of the scale parameter from the Burr, and that this entropy value remains constant across stress levels. Thus, if we have the same Burr parameters, but varying scale parameter then this will give the same entropy shape parameter, but, of course, varying entropy scale parameter. When we consider acceleration, the only parameter that changes is the scale parameter, so we obtain equivalent results for the accelerated and non-accelerated cases. Thus, using our work on non-accelerated data sets, we can immediately conclude that, regardless of the number of stress levels we chose, and how we arrange these stress levels and corresponding sample sizes, in terms of maximised entropy, we observe an improved fit as we increase a. An alternative statement is as follows : when we have the same scale-stress model, we reduce the number of entropy values to just 2, since we know that $\widetilde{\beta}_{W*} = \beta_B$. B_* remains the same, α_{W*} is just a reparameterised θ_{i*}. Thus we are back to the non-accelerated scenario, and so all the results apply in this situation.

7.2.2 Fitting G_{WP} to G_{GP}

Simulation studies

In keeping with previous examples, we construct tables for our usual sets of 2, 3 and 4 stress levels. We vary τ from the Gamma distribution and set the parameter equal to 0.5, 1, 2 and 3, but only include results for $\tau = 3$. We also fix α_G and β_G at 8 and -0.02, respectively. When $\tau = 3$, we have entropy values $\widetilde{B}_P = 1.8328$ and $\widetilde{\alpha}_{WP} = 9.2201$, with $\widetilde{\beta}_{WP} = \beta_G$, and a true value of B_{10} given by 1208.5614. We summarise our results for $k = 2$ in Tables 7.40 and 7.41, for $k = 3$ in Tables 7.42 and 7.43, and for $k = 4$, these can be found in Tables

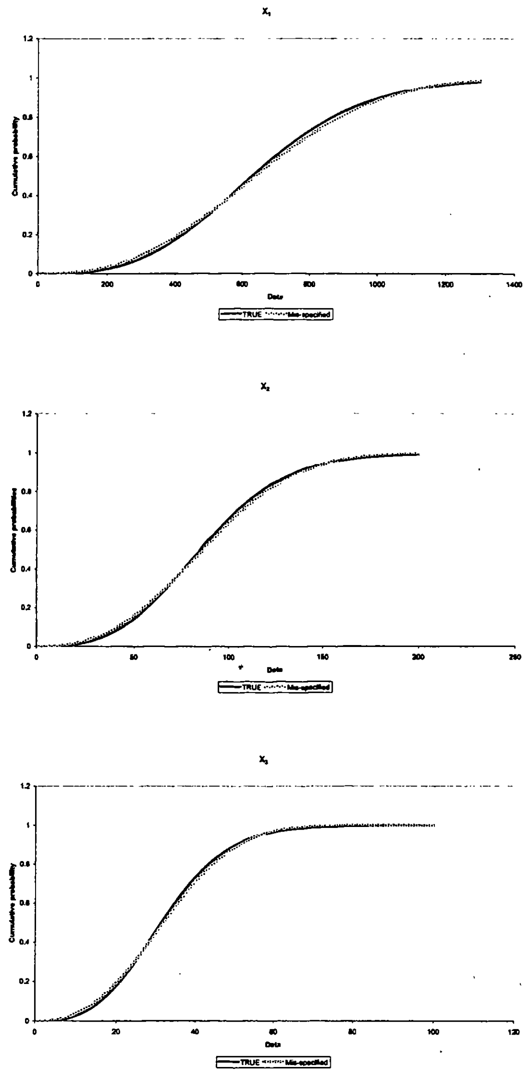


Table 7.39: Comparison between the true Burr Log-linear model, with parameters $\tau = 3$, $a = 4$, $\alpha_B = 8$, $\beta_B = -0.02$, and the mis-specified Weibull Log-linear distribution, with entropy values $\widetilde{B}_P = 2.55279$, $\widetilde{\alpha}_{WP} = 7.60343$, $\widetilde{\beta}_{WP} = -0.02$. Here, $k = 3$, $X_1 = 50$, $X_2 = 150$, $X_3 = 200$.

7.44 and 7.45. We list a number of points concerning results from these tables. Firstly, the Gamma and Weibull MLEs match up well with their theoretical counterparts even for small sample sizes, and, as expected, the standard errors for all MLEs and $\widehat{B}_{10}$ from true and mis-specified distributions decrease as the sample size increases. The theoretical standard errors for $\widehat{\tau}$ and $\widehat{B}_P$ remain constant as we vary the n_i, and larger standard errors are observed for the remaining MLEs and $\widehat{B}_{10}$, as more observations are assigned to the higher stress levels. For $\tau = 0.5$, 2 and 3, and small samples (say less that 100), the probability of choosing the incorrect distribution is as high as 20-30%. We also see quite a difference in the estimates for B_{10} from true and mis-specified distributions for this particular set of Gamma parameters. With $\tau = 1$, we are fitting, and simulating data from, a Negative Exponential distribution. Since this distribution is also a special case of the Weibull distribution, we cannot discriminate between Weibull and Gamma. This is why we can expect, across all stress levels, the probability of choosing the incorrect distribution function fluctuating around 0.5, and should also see excellent agreement between the estimates for B_{10} for true and mis-specified models.

The effects of mis-specification

This section examines the consequences of fitting the Weibull Log-linear distribution to data with an underlying Gamma Log-linear model. We do this in our usual way, and compare theoretical true and fitted distributions across stress levels by examining maximum absolute distances between cdfs. We do this in detail for one example, and then tabulate the results for varying parameters thereafter. We set the parameters from the Gamma Log-linear distribution to

$$\tau = 3, \alpha_G = 8, \beta_G = -0.02,$$

which correspond to scale parameters for our usual three stress levels of

$$a_1 = 1096.6332, a_2 = 148.4132, a_3 = 54.5982.$$

When we fit the Weibull distribution to data with this underlying Gamma distribution, we obtain entropy values of

$$\widetilde{B}_P = 1.8328, \widetilde{\alpha}_{WP} = 9.2201, \widetilde{\beta}_{WP} = -0.02,$$

which correspond to Weibull scale parameters equal to

$$\widetilde{\theta}_{1P} = 3714.8739, \widetilde{\theta}_{2P} = 502.7535, \widetilde{\theta}_{3P} = 184.9527.$$

We construct plots comparing the cdfs of true and mis-specified distributions across the three stress levels, each time noting the maximum absolute distance between the two distribution

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_P$	1.8834	1.8566	1.8423	1.8372
S	0.1459	0.1026	0.0647	0.0450
T	0.1429	0.1011	0.0639	0.0452
$\widehat{\alpha}_{WP}$	9.2161	9.2188	9.2192	9.2194
S	0.1143	0.0817	0.0521	0.0364
T	0.1158	0.0819	0.0518	0.0366
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0008	0.0006	0.0004	0.0003
T	0.0008	0.0006	0.0004	0.0003
$\widehat{B}_{W,10}$	1120.8493	1104.3137	1094.3840	1090.7757
S	139.6749	99.2357	63.2593	44.1047
T	140.1841	99.1251	62.6922	44.3301
Pr (Fit G_{WP})	0.2300	0.1337	0.0353	0.0045
$\widehat{\tau}$	3.1244	3.0563	3.0220	3.0093
S	0.4282	0.2921	0.1838	0.1273
T	0.4029	0.2849	0.1802	0.1274
$\widehat{\alpha}_G$	7.9647	7.9844	7.9937	7.9971
S	0.1761	0.1235	0.0788	0.0550
T	0.1750	0.1238	0.0783	0.0553
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0008	0.0005	0.0003	0.0002
T	0.0008	0.0005	0.0003	0.0002
$\widehat{B}_{G,10}$	1231.9567	1219.6071	1212.6143	1209.9148
S	146.9338	103.8495	66.2977	46.0933
T	146.6743	103.7144	65.5947	46.3825

Table 7.40: Fitting G_{WP} to G_{GP} for $k = 2$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_P$	1.8552	1.8571	1.8560	1.8578
S	0.1025	0.1036	0.1032	0.1034
T	0.1011	0.1011	0.1011	0.1011
$\widehat{\alpha}_{WP}$	9.2043	9.2235	9.2124	9.2208
S	0.1600	0.0712	0.1125	0.0708
T	0.1608	0.0718	0.1141	0.0704
$\widehat{\beta}_{WP}$	-0.0199	-0.0201	-0.0200	-0.0200
S	0.0009	0.0009	0.0006	0.0007
T	0.0009	0.0009	0.0007	0.0007
$\widehat{B}_{W,10}$	1097.8089	1104.8552	1100.9326	1105.1424
S	151.0026	89.9034	118.8285	91.9691
T	150.6503	89.3715	118.8095	91.6287
Pr (Fit G_{WP})	0.1294	0.1317	0.1290	0.1346
$\widehat{\tau}$	3.0550	3.0594	3.0578	3.0592
S	0.2951	0.3000	0.2953	0.2959
T	0.2849	0.2849	0.2849	0.2849
$\widehat{\alpha}_G$	7.9771	7.9861	7.9807	7.9844
S	0.1828	0.1201	0.1452	0.1182
T	0.1815	0.1178	0.1453	0.1171
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0008	0.0008	0.0006	0.0006
T	0.0008	0.0008	0.0006	0.0006
$\widehat{B}_{G,10}$	1218.2638	1219.7885	1218.9797	1219.6914
S	159.7946	94.4010	125.0095	96.3917
T	159.2573	93.1132	125.0016	95.5707

Table 7.41: Fitting G_{WP} to G_{GP} for $k = 2$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	25,25 25	50,50 50	100,100 100	300,300 300	500,500 500
$\widehat{B}_P$	1.9025	1.8634	1.8484	1.8384	1.8356
S	0.1741	0.1187	0.0833	0.0484	0.0369
T	0.1650	0.1167	0.0825	0.0476	0.0369
$\widehat{\alpha}_{WP}$	9.2099	9.2158	9.2185	9.2191	9.2198
S	0.1562	0.1148	0.0809	0.0467	0.0364
T	0.1629	0.1152	0.0815	0.0470	0.0364
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0011	0.0008	0.0006	0.0003	0.0002
T	0.0011	0.0008	0.0006	0.0003	0.0002
$\widehat{B}_{W,10}$	1131.7639	1108.1852	1098.9021	1091.8549	1090.1168
S	177.4187	125.4105	88.3920	50.8594	39.3979
T	175.2036	123.8876	87.6018	50.5769	39.1767
Pr (Fit G_{WP})	0.2686	0.1716	0.0803	0.0058	0.0005
$\widehat{\tau}$	3.1747	3.0762	3.0380	3.0132	3.0067
S	0.5149	0.3433	0.2372	0.1358	0.1051
T	0.4652	0.3290	0.2326	0.1343	0.1040
$\widehat{\alpha}_G$	7.9484	7.9776	7.9894	7.9959	7.9981
S	0.2206	0.1565	0.1099	0.0634	0.0493
T	0.2209	0.1562	0.1105	0.0638	0.0494
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0010	0.0008	0.0005	0.0003	0.0002
T	0.0011	0.0008	0.0005	0.0003	0.0002
$\widehat{B}_{G,10}$	1241.6525	1223.4006	1216.7104	1211.3198	1209.9008
S	186.8707	131.9524	92.9665	53.1138	41.4149
T	183.8071	129.9712	91.9035	53.0605	41.1005

Table 7.42: Fitting G_{WP} to G_{GP} for $k = 3$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	50,100	25,200	25,25	200,50
n_3	150	75	250	50
$\widehat{B}_P$	1.8486	1.8480	1.8481	1.8483
S	0.0837	0.0839	0.0830	0.0842
T	0.0825	0.0825	0.0825	0.0825
$\widehat{\alpha}_{WP}$	9.2162	9.2118	9.2060	9.2218
S	0.1081	0.1454	0.1541	0.0617
T	0.1084	0.1450	0.1536	0.0622
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0199	-0.0200
S	0.0006	0.0009	0.0008	0.0006
T	0.0007	0.0009	0.0008	0.0006
$\widehat{B}_{W,10}$	1098.5563	1097.1009	1093.8586	1099.6147
S	105.2205	127.5779	139.2290	77.1511
T	104.9515	126.0492	138.5650	76.1254
Pr (Fit G_{WP})	0.0826	0.0812	0.0818	0.0789
$\widehat{\tau}$	3.0361	3.0346	3.0363	3.0372
S	0.2388	0.2380	0.2357	0.2382
T	0.2326	0.2326	0.2326	0.2326
$\widehat{\alpha}_G$	7.9887	7.9858	7.9829	7.9918
S	0.1311	0.1611	0.1685	0.0978
T	0.1300	0.1595	0.1668	0.0983
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0006	0.0009	0.0008	0.0005
T	0.0006	0.0009	0.0008	0.0005
$\widehat{B}_{G,10}$	1216.1851	1215.4418	1215.3498	1216.6478
S	110.2100	135.4477	147.7182	80.1565
T	110.6244	133.3001	146.7230	79.4590

Table 7.43: Fitting G_{WP} to G_{GP} for $k = 3$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_P$	1.8816	1.8585	1.8450	1.8380	1.8351
S	0.1464	0.1018	0.0720	0.0508	0.0320
T	0.1429	0.1011	0.0715	0.0505	0.0320
$\widehat{\alpha}_{WP}$	9.2078	9.2147	9.2184	9.2186	9.2197
S	0.1592	0.1151	0.0813	0.0577	0.0362
T	0.1620	0.1146	0.0810	0.0573	0.0362
$\widehat{\beta}_{WP}$	-0.0199	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0010	0.0007	0.0005	0.0004	0.0002
T	0.0010	0.0007	0.0005	0.0004	0.0002
$\widehat{B}_{W,10}$	1116.9103	1104.5101	1096.4632	1091.4256	1089.7244
S	165.4157	116.3267	82.4688	58.5031	36.9845
T	164.3341	116.2018	82.1671	58.1009	36.7462
Pr (Fit G_{WP})	0.2356	0.1342	0.0482	0.0083	0.0002
$\widehat{\tau}$	3.1170	3.0611	3.0296	3.0127	3.0051
S	0.4278	0.2902	0.2046	0.1444	0.0905
T	0.4029	0.2849	0.2015	0.1425	0.0901
$\widehat{\alpha}_G$	7.9626	7.9803	7.9914	7.9959	7.9985
S	0.2065	0.1476	0.1029	0.0735	0.0458
T	0.2060	0.1457	0.1030	0.0728	0.0461
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0010	0.0007	0.0005	0.0004	0.0002
T	0.0010	0.0007	0.0005	0.0004	0.0002
$\widehat{B}_{G,10}$	1230.0131	1220.7140	1214.9867	1210.9377	1209.6635
S	175.0763	122.0486	87.2290	61.5564	38.9390
T	172.7997	122.1879	86.3999	61.0939	38.6392

Table 7.44: Fitting G_{WP} to G_{GP} for $k = 4$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	75,75	25,75	75,25	100,50	25,25	25,75
n_3, n_4	75,75	25,25	25,75	75,25	25,25	50,100	75,25
$\widehat{B}_P$	1.8563	1.8561	1.8586	1.8556	1.8546	1.8574	1.8575
S	0.1033	0.1024	0.1028	0.1030	0.1014	0.1036	0.1031
T	0.1011	0.1011	0.1011	0.1011	0.1011	0.1011	0.1011
$\widehat{\alpha}_{WP}$	9.2089	9.2180	9.2111	9.2163	9.2183	9.2093	9.2100
S	0.1575	0.0982	0.1485	0.0955	0.0860	0.1554	0.1488
T	0.1568	0.0981	0.1508	0.0966	0.0860	0.1555	0.1577
$\widehat{\beta}_{WP}$	-0.0199	-0.0200	-0.0200	-0.0200	-0.0200	-0.0199	-0.0199
S	0.0009	0.0007	0.0009	0.0007	0.0007	0.0009	0.0009
T	0.0009	0.0007	0.0009	0.0007	0.0007	0.0009	0.0010
$\widehat{B}_{W,10}$	1101.7270	1104.2874	1104.2801	1102.5499	1102.4953	1102.8239	1102.7546
S	147.0875	105.5711	139.8414	105.9428	98.4938	146.9636	140.0436
T	145.0100	104.7791	139.5639	105.4171	98.4866	144.7954	142.4541
Pr(Fit G_{WP})	0.1340	0.1381	0.1322	0.1303	0.1285	0.1309	0.1340
$\widehat{\tau}$	3.0577	3.0542	3.0641	3.0551	3.0530	3.0600	3.0594
S	0.2953	0.2933	0.2973	0.2954	0.2933	0.2965	0.2961
T	0.2849	0.2849	0.2849	0.2849	0.2849	0.2849	0.2849
$\widehat{\alpha}_G$	7.9793	7.9855	7.9784	7.9838	7.9853	7.9792	7.9784
S	0.1807	0.1344	0.1720	0.1330	0.1271	0.1785	0.1730
T	0.1782	0.1342	0.1734	0.1331	0.1263	0.1772	0.1790
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0009	0.0007	0.0009	0.0006	0.0007	0.0009	0.0009
T	0.0009	0.0007	0.0009	0.0006	0.0007	0.0008	0.0009
$\widehat{B}_{G,10}$	1221.2248	1220.0362	1222.8640	1218.9286	1218.2290	1222.2051	1220.4060
S	156.0030	110.7813	148.2242	110.6514	103.6939	154.8323	147.5701
T	153.2000	109.8412	147.3476	110.5319	103.0217	152.9695	150.4539

Table 7.45: Fitting G_{WP} to G_{GP} for $k = 4$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

functions. The table of figures in Table 7.46 shows that across all stress levels, we see an equivalent fit between true and mis-specified distributions, and a maximum absolute distance of 0.0263 is observed at each case. This figure is identical to the non-accelerated counterpart, a result that can be expected since we know that the entropy value for the shape parameter from the Weibull distribution is independent of the Gamma scale parameter. Since, with acceleration, it is only the scale parameter that changes at each stress level, we will observe identical results to the non-accelerated scenario. Thus, we no longer need to tabulate results for varying values of τ, since we use our results for the non-accelerated Gamma distribution. Here we established that the worst fit between Weibull and Gamma occurred for very small values of τ, whilst the best fit was when $\tau = 1$, which was equivalent to fitting the Negative Exponential distribution. As τ increased, we saw the maximum absolute distance between the cdfs of the Gamma and Weibull distributions level off at about 0.06.

7.2.3 Fitting G_{WP} to G_{LNP}

Simulation studies

We run simulations for varying parameter values and stress levels from the Lognormal distribution. As in previous cases, we allow the number of stress levels to vary from 2 to 4, and use the same sets of levels and sample sizes. However, the way in which we chose parameter values from this distribution must be consistent with the methods established for the Burr and Gamma distributions. For these cases, it was relatively straightforward to obtain values for the parameters in the scale-stress relationship since we had obvious scale parameters. It is slightly more complicated for the Log-normal distribution since we do not have an unique scale parameter. The parameter μ roughly determines where the data points lie but is not an actual scale parameter. Thus, if we want the data points at the first stress level to lie roughly around 2000 time units, then this implies we must take $\mu_1 = \ln 2000$; this really means that the middle value of the data set is 2000 since the median for the Lognormal distribution is $\exp(\mu)$. Similarly, at the k^{th} stress level, if we expect the lifetime of items to have decreased to about 200 time units, then we must set $\mu_k = \ln 200$. Inserting these values into the log-linear scale stress relationship, solving for α_{LN} and β_{LN}, and rounding up implies that we should set

$$\alpha_{LN} = 8,\ \beta_{LN} = -0.02.$$

To determine values for the remaining parameter σ, we consider the following simple example. Since we are interested in how the Weibull distribution represents a set of data with another underlying model, in practice, we should not fit this distribution to a set of data that has quite extreme values and which looks quite different from data that has originated from a Weibull distribution. To illustrate data from a Weibull distribution, we display plots

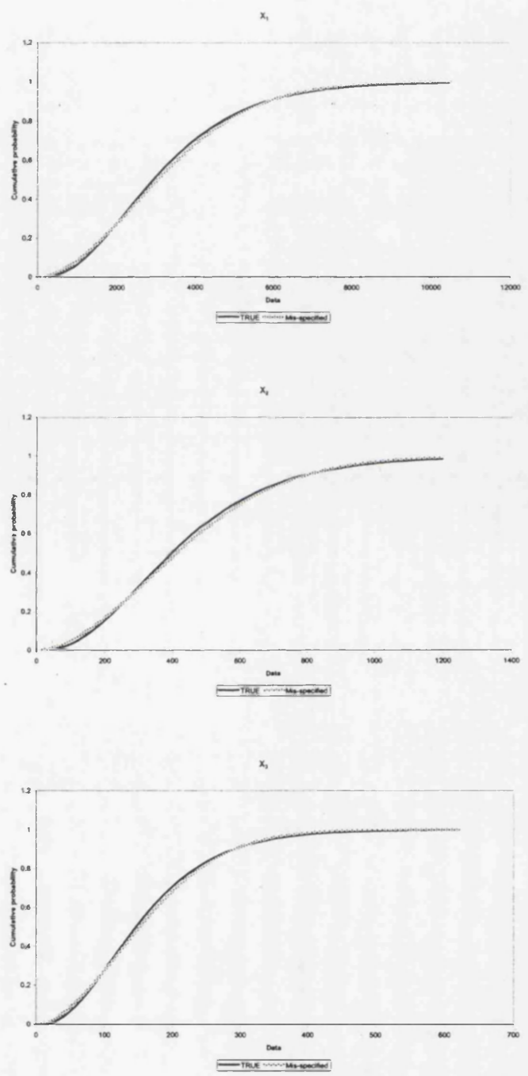
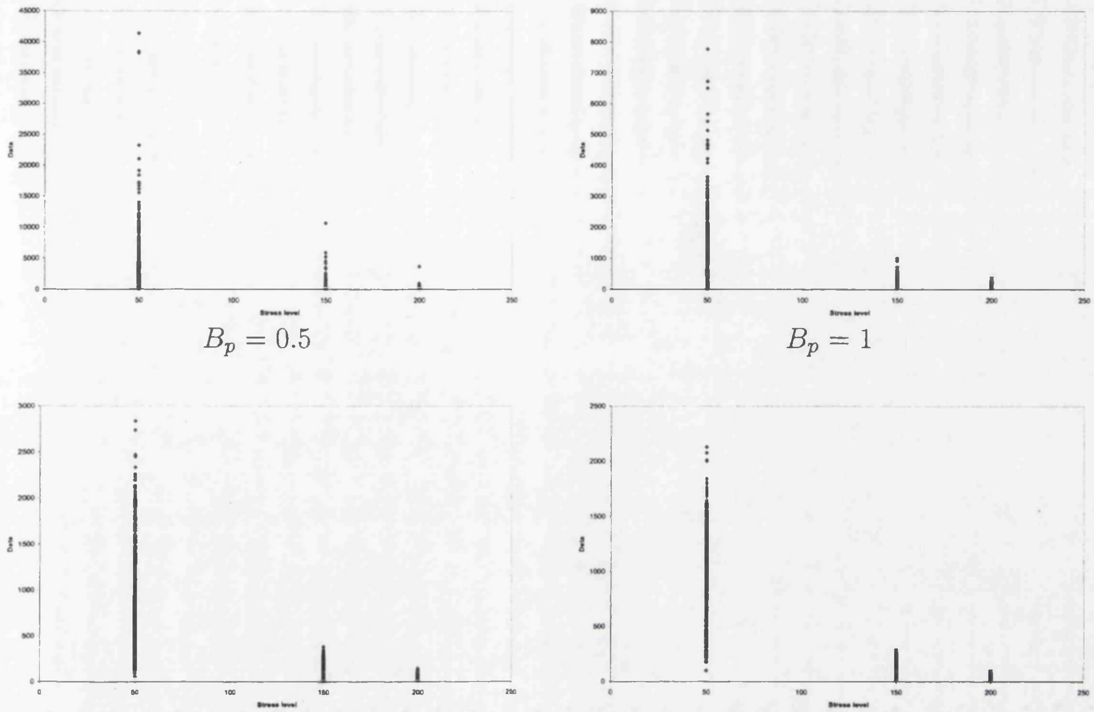


Table 7.46: Comparison between the true Gamma Log-linear model, with parameters $\tau = 3$, $\alpha_G = 8$, $\beta_G = -0.02$, and the mis-specified Weibull Log-linear distribution, with entropy values $\underset{\sim}{B}_P = 1.8328$, $\widetilde{\alpha}_{WP} = 9.2201$, $\widetilde{\beta}_{WP} = -0.02$. Here, $k = 3$, $X_1 = 50$, $X_2 = 150$, $X_3 = 200$.



$B_p = 0.5$ $B_p = 1$

$B_p = 2$ $B_p = 3$

Table 7.47: Data from a Weibull Log-linear distribution for varying B_P.

of data from a Weibull Log-linear model with parameters

$$\alpha_{WP} = 8,\ \beta_{WP} = -0.02,$$

and $B_p = 0.5, 1, 2$ and 3. The scale stress parameters correspond to similar scaling values from the Lognormal distribution. We can then compare this data to that from a Lognormal Log-linear model, as we allow σ to vary also from 0.5 to 3. Table 7.47 shows four Weibull plots corresponding to each value of B_P. Note we have three stress levels with 500 observations at each level. The corresponding Lognormal plots are shown in Table 7.48. We see that small values of B_P correspond to quite large data values (and hence a large spread), whilst larger values of σ have the same effect on data from the Lognormal distribution. The effect however, is intensified in this case, and for $\sigma = 3$, the largest data point is a very large 13000000 time units. As a result, it is only small values of σ, say between 0.5 to 1.5, that generate data that might be mistakenly regarded as from a Weibull distribution. These are the range of values we consider when running simulations, and take $\sigma = 0.5$, 1 and 1.5, although only summarise results for $\sigma = 1$; these can be found in Tables 7.49 and 7.50 for $k = 2$, Tables 7.51 and 7.52 for $k = 3$, and Tables 7.53 and 7.54 for $k = 4$. We compare MLEs from the Weibull distribution with entropy values given by $\widetilde{B}_P = 1$, $\widetilde{\beta}_{WP} = -0.02$, and $\widetilde{\alpha}_{WP} = 8.5$, and all estimates of B_{10} with a true value given by $B_{LN,10} = 304.4323$. The tables show similar results to those already observed in previous cases. We see good

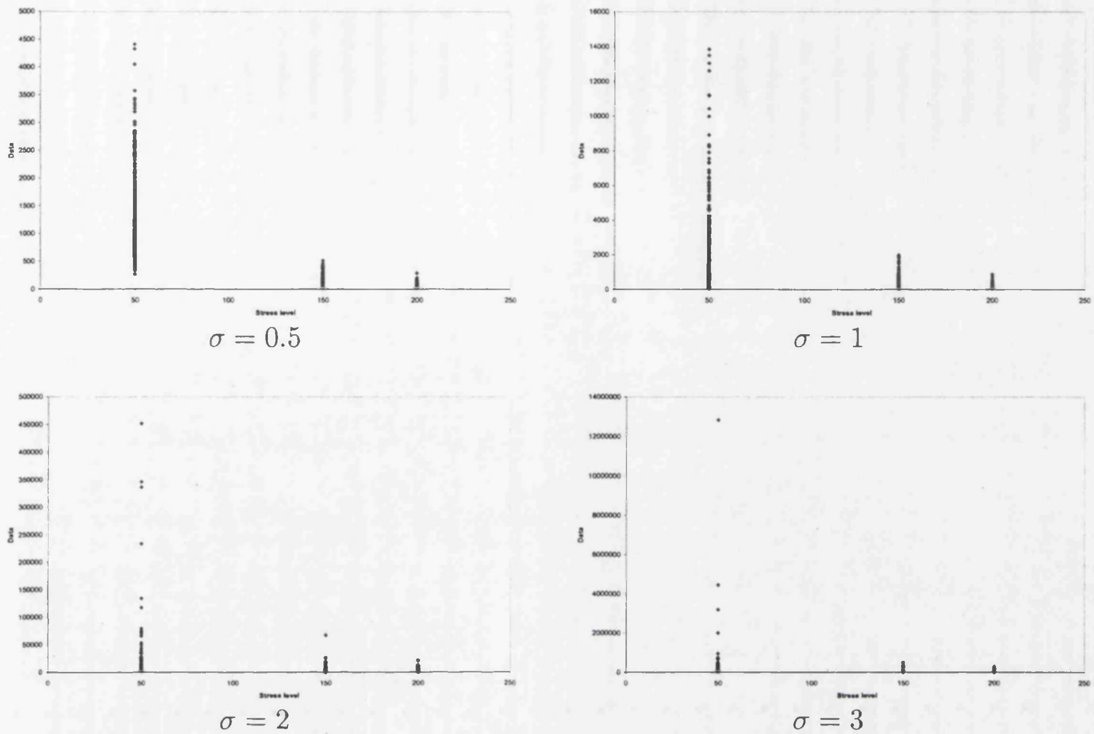


Table 7.48: Data from a Lognormal Log-linear distribution for varying σ.

agreement between MLEs and their theoretical counterparts for both distribution functions, and this agreement improves as the sample size increases. The standard errors also decrease, indicating more accurate estimates. The theoretical standard errors for $\widehat{\sigma}$ and $\widehat{B}_P$ are not affected by how we arrange the sample sizes across stress levels, but generally these values for the remaining MLEs and $\widehat{B}_{10}$ increase as more observations are allocated to the middle and higher stress levels. When we examine B_{10} from true and mis-specified distributions, we see quite considerable differences, and this quantile is always under-estimated when we fit the Weibull distribution. We also see much smaller probabilities associated with preferring the Weibull distribution over the true, and these rarely exceed 5%.

The effects of mis-specification.

We show how we compare true and mis-specified distributions for the Lognormal distribution for a particular set of parameters and stress levels. We then tabulate the results for the remaining parameter values. We set

$$\sigma = 1,\ \alpha_{LN} = 8,\ \beta_{LN} = -0.02,$$

and use our usual set of three stress levels. These correspond to

$$\mu_1 = 7,\ \mu_2 = 5,\ \mu_3 = 4.$$

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_P$	1.0419	1.0221	1.0101	1.0053
S	0.0973	0.0703	0.0448	0.0320
T	0.1053	0.0745	0.0471	0.0333
$\widehat{\alpha}_{WP}$	8.4922	8.4948	8.4978	8.4995
S	0.2359	0.1671	0.1079	0.0797
T	0.2431	0.1719	0.1087	0.0769
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0017	0.0012	0.0008	0.0006
T	0.0017	0.0012	0.0008	0.0006
$\widehat{B}_{W,10}$	208.7436	200.0274	194.8248	192.8877
S	48.2857	34.4702	21.9088	15.7381
T	50.7621	35.8942	22.7015	16.0524
Pr (Fit G_{WP})	0.0293	0.0035	0	0
$\widehat{\sigma}$	0.9876	0.9933	0.9970	0.9986
S	0.0706	0.0499	0.0314	0.0223
T	0.0707	0.0500	0.0316	0.0224
$\widehat{\alpha}_{LN}$	7.9977	7.9989	7.9988	8.0000
S	0.1931	0.1351	0.0870	0.0621
T	0.1944	0.1374	0.0869	0.0615
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0013	0.0009	0.0006	0.0004
T	0.0013	0.0009	0.0006	0.0004
$\widehat{B}_{LN,10}$	313.1263	308.8949	306.2157	305.4202
S	52.2090	36.1446	22.9694	16.4383
T	51.1336	36.1569	22.8677	16.1699

Table 7.49: Fitting G_{WP} to G_{LNP} for $k = 2$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_P$	1.0234	1.0230	1.0230	1.0240
S	0.0701	0.0703	0.0704	0.0698
T	0.0745	0.0745	0.0745	0.0745
$\widehat{\alpha}_{WP}$	8.4681	8.5106	8.4867	8.5043
S	0.3353	0.1424	0.2356	0.1518
T	0.3469	0.1489	0.2438	0.1458
$\widehat{\beta}_{WP}$	-0.0198	-0.0202	-0.0199	-0.0201
S	0.0018	0.0018	0.0014	0.0014
T	0.0019	0.0019	0.0014	0.0014
$\widehat{B}_{W,10}$	201.6528	200.9562	200.9491	201.1952
S	58.0354	30.1682	43.6099	30.9560
T	56.2052	31.9556	43.7257	32.8717
Pr (Fit G_{WP})	0.0035	0.0029	0.0041	0.0032
$\widehat{\sigma}$	0.9933	0.9930	0.9936	0.9932
S	0.0501	0.0501	0.0502	0.0498
T	0.0500	0.0500	0.0500	0.0500
$\widehat{\alpha}_{LN}$	8.0024	8.0017	7.9999	8.0012
S	0.2704	0.1203	0.1901	0.1178
T	0.2679	0.1208	0.1905	0.1186
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0014	0.0014	0.0011	0.0011
T	0.0014	0.0014	0.0011	0.0011
$\widehat{B}_{LN10}$	314.4705	309.2148	310.6762	308.9610
S	67.2498	30.7549	48.7640	31.8840
T	63.9351	30.1683	47.2664	31.5974

Table 7.50: Fitting G_{WP} to G_{LNP} for $k = 2$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	25,25 25	50,50 50	100,100 100	300,300 300	500,500 500
$\widehat{B}_P$	1.0555	1.0306	1.0162	1.0058	1.0032
S	0.1126	0.0816	0.0572	0.0340	0.0265
T	0.1216	0.0860	0.0608	0.0351	0.0272
$\widehat{\alpha}_{WP}$	8.4847	8.4901	8.4957	8.4976	8.5003
S	0.3349	0.2384	0.1714	0.0993	0.0777
T	0.3462	0.2448	0.1731	0.1000	0.0774
$\widehat{\beta}_{WP}$	-0.0199	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0023	0.0017	0.0012	0.0007	0.0005
T	0.0024	0.0017	0.0012	0.0007	0.0005
$\widehat{B}_{W,10}$	215.8159	204.2198	197.7850	192.9754	192.1208
S	63.8147	44.6255	31.0770	17.9222	14.0380
T	63.9449	45.2159	31.9725	18.4593	14.2985
Pr (Fit G_{WP})	0.0557	0.0097	0.0001	0	0
$\widehat{\sigma}$	0.9829	0.9908	0.9956	0.9985	0.9993
S	0.0808	0.0577	0.0407	0.0234	0.0181
T	0.0816	0.0577	0.0408	0.0236	0.0183
$\widehat{\alpha}_{LN}$	8.0003	8.0005	7.9999	7.9987	8.0007
S	0.2745	0.1932	0.1387	0.0791	0.0613
T	0.2726	0.1927	0.1363	0.0787	0.0609
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0019	0.0013	0.0009	0.0005	0.0004
T	0.0019	0.0013	0.0009	0.0005	0.0004
$\widehat{B}_{LN,10}$	318.6971	312.0114	308.0314	305.3694	305.2527
S	70.2666	48.7916	34.4553	19.2691	14.9626
T	66.7617	47.2076	33.3808	19.2724	14.9284

Table 7.51: Fitting G_{WP} to G_{LNP} for $k = 3$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	50,100 150	25,200 75	25,25 250	200,50 50
$\widehat{B}_P$	1.0151	1.0161	1.0149	1.0154
S	0.0574	0.0576	0.0578	0.0575
T	0.0608	0.0608	0.0608	0.0608
$\widehat{\alpha}_{WP}$	8.4885	8.4838	8.4698	8.5013
S	0.2277	0.3049	0.3205	0.1283
T	0.2329	0.3135	0.3323	0.1299
$\widehat{\beta}_{WP}$	-0.0199	-0.0199	-0.0198	-0.0200
S	0.0014	0.0019	0.0017	0.0012
T	0.0014	0.0020	0.0018	0.0012
$\widehat{B}_{W,10}$	197.6669	199.2142	197.8795	197.3409
S	38.5590	47.4881	53.2655	26.0495
T	38.8307	47.0766	51.9383	27.3701
Pr (Fit G_{WP})	0.0004	0.0002	0.0005	0.0002
$\widehat{\sigma}$	0.9963	0.9954	0.9962	0.9955
S	0.0408	0.0407	0.0406	0.0407
T	0.0408	0.0408	0.0408	0.0408
$\widehat{\alpha}_{LN}$	7.9997	8.0002	8.0022	8.0006
S	0.1827	0.2419	0.2576	0.1039
T	0.1808	0.2415	0.2557	0.1046
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0011	0.0015	0.0014	0.0009
T	0.0011	0.0015	0.0014	0.0010
$\widehat{B}_{LN,10}$	308.9417	311.0516	312.4970	307.4714
S	44.0338	55.1882	62.5759	26.9258
T	42.8487	53.7485	60.0366	26.6147

Table 7.52: Fitting G_{WP} to G_{LNP} for $k = 3$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_P$	1.0431	1.0232	1.0117	1.0062	1.0027
S	0.0983	0.0706	0.0500	0.0364	0.0230
T	0.1053	0.0745	0.0527	0.0372	0.0235
$\widehat{\alpha}_{WP}$	8.4783	8.4846	8.4925	8.4983	8.4988
S	0.3323	0.2393	0.1719	0.1218	0.0771
T	0.3464	0.2450	0.1732	0.1225	0.0775
$\widehat{\beta}_{WP}$	-0.0199	-0.0199	-0.0200	-0.0199	-0.0200
S	0.0022	0.0016	0.0011	0.0008	0.0005
T	0.0023	0.0016	0.0011	0.0008	0.0005
$\widehat{B}_{W,10}$	210.1290	200.5018	195.7326	193.4745	191.7031
S	60.6362	41.9601	29.8985	20.9874	13.2869
T	60.3798	42.6950	30.1899	21.3475	13.5013
Pr (Fit G_{WP})	0.0309	0.0032	0	0	0
$\widehat{\sigma}$	0.9868	0.9934	0.9969	0.9986	0.9991
S	0.0706	0.0505	0.0351	0.0254	0.0159
T	0.0707	0.0500	0.0354	0.0250	0.0158
$\widehat{\alpha}_{LN}$	7.9981	7.9976	7.9985	7.9996	8.0002
S	0.2713	0.1923	0.1365	0.0952	0.0605
T	0.2706	0.1914	0.1353	0.0957	0.0605
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0017	0.0012	0.0009	0.0006	0.0004
T	0.0017	0.0012	0.0009	0.0006	0.0004
$\widehat{B}_{LN,10}$	316.1574	309.9421	307.1367	305.7418	305.1553
S	67.8121	46.7294	33.0353	23.0122	14.5163
T	64.8339	45.8445	32.4169	22.9222	14.4973

Table 7.53: Fitting G_{WP} to G_{LNP} for $k = 4$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	75,75	25,75	75,25	100,50	25,25	25,75
n_3, n_4	75,75	25,25	25,75	75,25	25,25	50,100	75,25
$\widehat{B}_P$	1.0234	1.0223	1.0232	1.0227	1.0238	1.0232	1.0230
S	0.0695	0.0706	0.0707	0.0703	0.0695	0.0702	0.0708
T	0.0745	0.0745	0.0745	0.0745	0.0745	0.0745	0.0745
$\widehat{\alpha}_{WP}$	8.4774	8.4925	8.4768	8.4971	8.5013	8.4752	8.4807
S	0.3282	0.2031	0.3074	0.1999	0.1753	0.3227	0.3213
T	0.3381	0.2083	0.3250	0.2049	0.1813	0.3354	0.3402
$\widehat{\beta}_{WP}$	-0.0199	-0.0200	-0.0199	-0.0200	-0.0200	-0.0199	-0.0199
S	0.0019	0.0015	0.0019	0.0014	0.0014	0.0018	0.0020
T	0.0020	0.0016	0.0020	0.0015	0.0015	0.0019	0.0021
$\widehat{B}_{W,10}$	202.5715	200.2485	201.5248	201.1068	201.4161	201.9580	202.1969
S	55.7509	37.1060	51.8538	37.3813	33.7359	55.0082	52.9871
T	54.0059	38.1571	51.8784	38.4117	35.6378	53.9222	53.0080
Pr(Fit G_{WP})	0.0038	0.0049	0.0038	0.0034	0.0038	0.0033	0.0041
$\widehat{\sigma}$	0.9929	0.9942	0.9932	0.9940	0.9934	0.9931	0.9929
S	0.0497	0.0500	0.0498	0.0509	0.0497	0.0499	0.0505
T	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
$\widehat{\alpha}_{LN}$	8.0016	7.9992	7.9981	8.0012	8.0039	8.0001	8.0046
S	0.2626	0.1642	0.2499	0.1618	0.1432	0.2604	0.2598
T	0.2612	0.1642	0.2514	0.1616	0.1443	0.2591	0.2628
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0015	0.0012	0.0015	0.0011	0.0011	0.0015	0.0016
T	0.0015	0.0012	0.0015	0.0011	0.0011	0.0015	0.0016
$\widehat{B}_{LN,10}$	313.9232	309.0568	312.3394	309.6813	309.8247	313.5410	314.0248
S	64.0900	40.2039	60.1883	40.9125	36.3069	63.7197	61.3595
T	61.0526	39.4511	58.2460	39.8166	35.7780	60.9425	59.7385

Table 7.54: Fitting G_{WP} to G_{LNP} for $k = 4$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

The entropy values for this set of parameters are

$$\widetilde{B}_{WP} = 1, \widetilde{\alpha}_{WP} = 8.5, \widetilde{\beta}_{WP} = -0.02,$$

which then correspond to entropy scale parameters of

$$\widetilde{\theta}_{1P} = 1808.0424, \widetilde{\theta}_{2P} = 244.6919, \widetilde{\theta}_{3P} = 90.0171.$$

We plot the three distribution functions corresponding to each stress level, shown by the table of figures, Table 7.55. The plots look very similar across all three stress levels, and in fact, the maximum absolute distance at each level remains the same and peaks at 0.060497. This can be expected since we know that if $\rho_m(X_i) = \rho_t(X_i)$, then μ_i does not influence the entropy shape parameter from the Weibull distribution (since $B_* = \sigma^{-1}$). When we allow the data set to undergo acceleration, only the μ_i and the shape parameters from both distributions remain constant. Thus, we see identical plots across stress levels, since the shapes of both Weibull and Lognormal distribution functions do not change as we vary μ_i. We also note that the maximum absolute distance obtained is consistent with the non-accelerated scenario since the entropy value for the shape parameter from the Weibull distribution is the same for both acceleration and non-acceleration. Since we observe such similarities between both cases, we are able to use the results from our non-acceleration work to conclude that the maximum absolute distance between true and mis-specified distribution functions remains the same across all values of μ_i and σ and hence, α_{LN}, β_{LN} and σ.

7.3 Getting the distribution and scale-stress relationship wrong

We now consider the final scenario, and mis-specify the underlying distribution function and scale-stress relationship. There are six possible combinations we could take, and list these below. In all cases, we assess the effects of fitting the incorrect distribution in two ways - simulation studies of parameter estimates, and maximum distances between cdfs across all stress levels.

7.3.1 Fitting G_{WP} to G_{BA}

Simulation studies

We summarise results of our simulations when we fit G_{WP} to data from G_{BA}. We run simulations for our usual four sets of Burr parameters, but, as previously, due to the size of each table, only include results for $\tau = 3$, $a = 4$; these are shown in Tables 7.56 and 7.57 for $k = 2$, Tables 7.58 and 7.59 for $k = 3$, and Tables 7.60 and 7.61 for $k = 4$. We include details on entropy values, average MLEs, and sample and theoretical standard errors of MLEs for both true and mis-specified distributions. We also summarise results for B_{10}, and compare sample results to a true value of $B_{B,10} = 575.2676$. The tables show similar

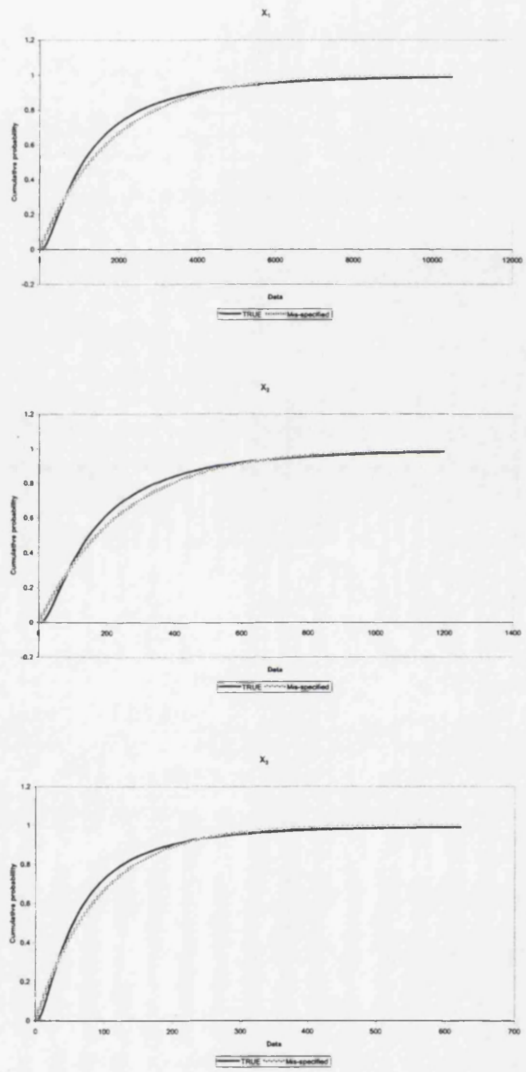


Table 7.55: Comparison between the true Lognormal Log-linear model, with parameters $\sigma = 1$, $\alpha_{LN} = 8$, $\beta_{LN} = -0.02$, and the mis-specified Weibull Log-linear distribution, with entropy values $\widetilde{B}_{WP} = 1$, $\widetilde{\alpha}_{WP} = 8.5$, $\widetilde{\beta}_{WP} = -0.02$. Here, $k = 3$, $X_1 = 50$, $X_2 = 150$, $X_3 = 200$.

results to previous investigations. When examining sample standard errors of the MLEs from the true distribution, we see that this statistic for $\widehat{a}$ is large for small sample sizes, and at times, we see the standard error rise as the sample size also increases; we have observed similar results for our non-accelerated counterparts. Otherwise, the sample standard errors for the remaining MLEs from the Burr are generally close to their theoretical counterparts and decrease as n increases. Our theoretical and sample values for $\widehat{B}_{B,10}$ also match up well, and the standard errors decrease as the total sample size increases. This is true even when we observe large estimates for the standard error of $\widehat{a}$. Entropy values are generally close to the Weibull MLEs and do not vary at all when $k = 2$ or when there are equal loadings at each stress level. The standard errors of the mis-specified MLEs also decrease as the overall sample size increases, and we see good agreement between the sample values and their theoretical counterparts. Generally, sample values for B_{10} from true and mis-specified distribution functions do not match up, and, as in previous scenarios, we always see this quantile being under-estimated when we fit the wrong distribution function.

The effects of mis-specification

We examine if a particular set of Burr Arrhenius parameter values results in a poor fit from the Weibull Log-linear model in the usual way. That is, we derive the maximum absolute distance between true and mis-specified distribution functions for different sets of Burr parameters, stress levels and sample sizes, and observe if we have any obvious differences. Since we now have two shape parameters from the Burr distribution that can change, we must accommodate for this and allow both to vary. To do this, we fix τ at 1, 2, 3 and 4 and allow a to vary each time from 1 to 4 in steps of 1. We first examine the results for $k = 2$.

Two stress levels We consider results for 2 stress levels, and choose the stresses and scale-stress parameters in a similar fashion to when we run simulations. We found equivalent results to the non-accelerated scenario and try to provide an explanation for this below. When we have two stress levels, we can completely determine the scale-stress parameters from the Weibull distribution, since we have two equations with 2 unknowns. This then just leaves $\widetilde{B}_*$ to estimate, and so we have reduced the problem to the non-accelerated scenario. Further increasing the number of stress levels leads to the system of simultaneous equations having no solutions and so we cannot used our work on non-acceleration in such cases. Our results for two stress levels strengthen this fact. When we allowed τ and a to vary, we observed equivalent values to the non-accelerated scenario for $\widetilde{B}_P$ and the maximum absolute distance between distribution functions across the two stress levels was also the same. Thus, we can conclude that we will see an improved fit between Weibull and Burr as we increase a, and varying τ has little if no effect. Such results are also consistent with those, when we fixed the scale-stress relationship and fitted Weibull to Burr.

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_P$	2.6312	2.5904	2.5695	2.5610
$\widetilde{B}_p$	2.5528	2.5528	2.5528	2.5528
S	0.2270	0.1613	0.1020	0.0730
T	0.2335	0.1651	0.1044	0.0738
$\widehat{\alpha}_{WP}$	7.9313	7.9310	7.9331	7.9331
$\widetilde{\alpha}_{WP}$	7.9335	7.9335	7.9335	7.9335
S	0.0827	0.0592	0.0376	0.0265
T	0.0845	0.0598	0.0378	0.0267
$\widehat{\beta}_{WP}$	-0.0153	-0.0153	-0.0153	-0.0153
$\widetilde{\beta}_{WP}$	-0.0153	-0.0153	-0.0153	-0.0153
S	0.0006	0.0004	0.0003	0.0002
T	0.0006	0.0004	0.0003	0.0002
$\widehat{B}_{W,10}$	548.3657	541.4616	538.8914	537.4582
S	52.5781	37.6093	23.9523	17.0559
T	54.2418	38.3547	24.2577	17.1528
Pr (Fit G_{WP})	0.0001	0.0002	0	0
$\widehat{\tau}$	3.1350	3.0411	3.0092	3.0037
S	0.4034	0.2820	0.1805	0.1284
T	0.3948	0.2791	0.1765	0.1248
$\widehat{a}$	12.9394	14.7534	5.9043	4.4710
S	147.1365	303.5799	31.9980	2.7715
T	3.3055	2.3373	1.4783	1.0453
$\widehat{\alpha}_B$	0.3563	0.3847	0.3466	0.3223
S	0.4554	0.3804	0.2376	0.1538
T	0.4303	0.3043	0.1924	0.1361
$\widehat{\beta}_B$	2346.3631	2346.0663	2346.9072	2346.9638
S	83.9091	59.5998	37.0641	26.4132
T	83.2317	58.8537	37.2223	26.3202
$\widehat{B}_{B,10}$	586.0124	578.1394	576.3102	575.6993
S	55.1458	39.1531	25.0217	17.6731
T	55.3765	39.1571	24.7651	17.5116

Table 7.56: Fitting G_{WP} to G_{BA} for $k = 2$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_P$	2.5940	2.5924	2.5938	2.5943
$\widetilde{B}_p$	2.5528	2.5528	2.5528	2.5528
S	0.1596	0.1585	0.1594	0.1602
T	0.1651	0.1651	0.1651	0.1651
$\widehat{\alpha}_{WP}$	7.9229	7.9362	7.9296	7.9343
$\widetilde{\alpha}_{WP}$	7.9335	7.9335	7.9335	7.9335
S	0.1157	0.0517	0.0824	0.0510
T	0.1180	0.0523	0.0836	0.0513
$\widehat{\beta}_{WP}$	-0.0153	-0.0154	-0.0153	-0.0154
$\widetilde{\beta}_{WP}$	-0.0153	-0.0153	-0.0153	-0.0153
S	0.0006	0.0006	0.0005	0.0005
T	0.0006	0.0006	0.0005	0.0005
$\widehat{B}_{W,10}$	540.5169	542.8663	542.0010	542.9140
S	55.5907	33.5568	43.7642	34.5321
T	56.2130	35.0696	45.0999	35.8256
Pr (Fit G_{WP})	0.0001	0.0001	0.0002	0.0005
$\widehat{\tau}$	3.0481	3.0479	3.0456	3.0469
S	0.2825	0.2826	0.2823	0.2833
T	0.2791	0.2791	0.2791	0.2791
$\widehat{a}$	10.3314	10.0699	10.0979	12.1062
S	133.5666	50.2903	87.0543	140.5326
T	2.3373	2.3373	2.3373	2.3373
$\widehat{\alpha}_B$	0.3867	0.3681	0.3805	0.3772
S	0.3980	0.4326	0.3771	0.3979
T	0.3297	0.3729	0.3079	0.3265
$\widehat{\beta}_B$	2343.5494	2350.2164	2346.3447	2348.4343
S	89.5832	89.8283	68.7511	68.9709
T	88.9784	88.9784	67.9584	67.9584
$\widehat{B}_{B,10}$	579.3728	579.6950	579.2306	579.3802
S	57.1014	35.7353	45.4451	36.4035
T	56.4619	36.0125	45.6600	36.7346

Table 7.57: Fitting G_{WP} to G_{BA} for $k = 2$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	25,25 25	50,50 50	100,100 100	300,300 300	500,500 500
$\widehat{B}_P$	2.5946	2.5482	2.5222	2.5034	2.4996
$\widetilde{B}_p$	2.4942	2.4942	2.4942	2.4942	2.4942
S	0.2581	0.1791	0.1272	0.0748	0.0573
T	0.2611	0.1846	0.1306	0.0754	0.0584
$\widehat{\alpha}_{WP}$	7.9007	7.9031	7.9041	7.9051	7.9057
$\widetilde{\alpha}_{WP}$	7.9056	7.9056	7.9056	7.9056	7.9056
S	0.1173	0.0829	0.0588	0.0338	0.0266
T	0.1186	0.0839	0.0593	0.0342	0.0265
$\widehat{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
$\widetilde{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
S	0.0008	0.0006	0.0004	0.0002	0.0002
T	0.0008	0.0006	0.0004	0.0002	0.0002
$\widehat{B}_{W,10}$	520.3669	513.7325	509.6496	506.7288	506.3396
S	62.5444	43.6757	31.2509	18.2903	14.1261
T	63.8032	45.1156	31.9016	18.4184	14.2668
Pr (Fit G_{WP})	0.1117	0.0465	0.0114	0	0
$\widehat{\tau}$	3.1973	3.0766	3.0216	3.0034	3.0012
S	0.4753	0.3270	0.2314	0.1344	0.1029
T	0.4558	0.3223	0.2279	0.1316	0.1019
$\widehat{a}$	20.1934	12.7870	10.7892	4.5524	4.2617
S	371.7213	164.1102	228.9296	4.3853	1.1220
T	3.8169	2.6989	1.9084	1.1018	0.8535
$\widehat{\alpha}_B$	0.3336	0.3798	0.3734	0.3254	0.3139
S	0.5197	0.4263	0.3276	0.1659	0.1219
T	0.5157	0.3647	0.2579	0.1489	0.1153
$\widehat{\beta}_B$	2345.0224	2345.6118	2346.1908	2346.9081	2347.2876
S	116.6829	80.9464	57.1038	32.7096	25.4786
T	113.2472	80.0779	56.6236	32.6916	25.3228
$\widehat{B}_{B,10}$	589.6606	581.8787	577.3128	575.5894	575.5796
S	69.0346	48.2571	34.5350	20.0247	15.4666
T	68.8181	48.6617	34.4090	19.8661	15.3882

Table 7.58: Fitting G_{WP} to G_{BA} for $k = 3$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	50,100 150	25,200 75	25,25 250	200,50 50
$\widehat{B}_P$	2.5207	2.5137	2.5604	2.5474
$\widetilde{B}_p$	2.4933	2.4848	2.5343	2.5217
S	0.1283	0.1286	0.1301	0.1283
T	0.1304	0.1313	0.1333	0.1381
$\widehat{\alpha}_{WP}$	7.8514	7.7707	7.8691	7.9421
$\widetilde{\alpha}_{WP}$	7.8551	7.7744	7.8793	7.9419
S	0.0786	0.1043	0.1110	0.0451
T	0.0792	0.1060	0.1131	0.0458
$\widehat{\beta}_{WP}$	-0.0152	-0.0151	-0.0151	-0.0158
$\widetilde{\beta}_{WP}$	-0.0152	-0.0151	-0.0151	-0.0157
S	0.0005	0.0007	0.0006	0.0004
T	0.0005	0.0007	0.0006	0.0004
$\widehat{B}_{W,10}$	492.4346	456.7643	512.1062	528.5095
S	35.3798	38.8611	48.2584	28.2749
T	35.9937	39.4955	48.8454	30.6059
Pr (Fit G_{WP})	0.0116	0.0099	0.0309	0.0231
$\widehat{\tau}$	3.0222	3.0168	3.0197	3.0189
S	0.2330	0.2317	0.2318	0.2334
T	0.2279	0.2279	0.2279	0.2279
$\widehat{a}$	8.0718	9.7601	7.6543	8.6059
S	47.9911	178.3940	45.1066	58.8513
T	1.9084	1.9084	1.9084	1.9084
$\widehat{\alpha}_B$	0.3741	0.3791	0.3783	0.3718
S	0.3354	0.3728	0.3469	0.3409
T	0.2679	0.3124	0.2872	0.2686
$\widehat{\beta}_B$	2345.7289	2345.2295	2343.1043	2347.5361
S	68.8901	97.1708	87.0496	58.1261
T	67.9571	96.1041	86.1388	57.4923
$\widehat{B}_{B,10}$	577.3160	577.4678	576.7046	576.8893
S	41.0020	49.9687	52.8274	30.3178
T	40.7711	49.7945	52.2790	30.4535

Table 7.59: Fitting G_{WP} to G_{BA} for $k = 3$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_P$	2.5861	2.5466	2.5282	2.5181	2.5113
$\widetilde{B}_p$	2.5071	2.5071	2.5071	2.5071	2.5071
S	0.2212	0.1565	0.1108	0.0787	0.0502
T	0.2282	0.1614	0.1141	0.0807	0.0510
$\widehat{\alpha}_{WP}$	7.9034	7.9059	7.9071	7.9080	7.9085
$\widetilde{\alpha}_{WP}$	7.9089	7.9089	7.9089	7.9089	7.9089
S	0.1155	0.0832	0.0580	0.0413	0.0265
T	0.1183	0.0836	0.0591	0.0418	0.0264
$\widehat{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
$\widetilde{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
S	0.0007	0.0005	0.0004	0.0003	0.0002
T	0.0008	0.0005	0.0004	0.0003	0.0002
$\widehat{B}_{W,10}$	519.5776	513.7450	510.9342	509.5064	508.4659
S	57.5765	41.4355	29.1137	20.7478	13.2211
T	59.5577	42.1137	29.7788	21.0568	13.3175
Pr (Fit G_{WP})	0.0983	0.0373	0.0076	0.0005	0
$\widehat{\tau}$	3.1380	3.0434	3.0111	3.0051	3.0024
S	0.4021	0.2817	0.2014	0.1417	0.0892
T	0.3948	0.2791	0.1974	0.1396	0.0883
$\widehat{a}$	17.1850	9.3225	9.8330	4.6567	4.1740
S	296.6904	43.3573	308.3059	5.2556	0.8773
T	3.3055	2.3373	1.6528	1.1687	0.7391
$\widehat{\alpha}_B$	0.3606	0.3851	0.3619	0.3288	0.3098
S	0.4962	0.3919	0.2844	0.1793	0.1050
T	0.4529	0.3203	0.2265	0.1601	0.1013
$\widehat{\beta}_B$	2345.6578	2345.9635	2346.1252	2346.4153	2346.8412
S	107.2704	75.8845	52.9006	37.3110	23.6892
T	105.8361	74.8374	52.9181	37.4187	23.6657
$\widehat{B}_{B,10}$	586.7590	579.2420	576.3225	575.7671	575.4764
S	63.7145	45.5256	31.8928	22.5367	14.2937
T	63.8837	45.1726	31.9419	22.5863	14.2848

Table 7.60: Fitting G_{WP} to G_{BA} for $k = 4$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 75,75	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_P$	2.5592	2.5468	2.5330	2.5664	2.5544	2.5599	2.5488
$\widetilde{B}_p$	2.5192	2.5049	2.4925	2.5281	2.5148	2.5192	2.5083
S	0.1584	0.1577	0.1553	0.1591	0.1574	0.1589	0.1577
T	0.1623	0.1617	0.1602	0.1632	0.1622	0.1621	0.1619
$\widehat{\alpha}_{WP}$	7.8754	7.9360	7.8262	7.9423	7.9459	7.8700	7.8672
$\widetilde{\alpha}_{WP}$	7.8842	7.9373	7.8335	7.9438	7.9458	7.8791	7.8762
S	0.1130	0.0710	0.1088	0.0692	0.0618	0.1112	0.1141
T	0.1148	0.0717	0.1103	0.0706	0.0629	0.1140	0.1154
$\widehat{\beta}_{WP}$	-0.0153	-0.0160	-0.0152	-0.0158	-0.0160	-0.0152	-0.0156
$\widetilde{\beta}_{WP}$	-0.0154	-0.0160	-0.0152	-0.0158	-0.0160	-0.0153	-0.0156
S	0.0007	0.0005	0.0007	0.0005	0.0005	0.0006	0.0007
T	0.0007	0.0005	0.0007	0.0005	0.0005	0.0006	0.0007
$\widehat{B}_{W,10}$	508.4472	519.0361	483.6255	529.9895	526.3659	508.2451	496.5792
S	50.5107	38.4658	45.7582	39.1159	36.0818	49.8127	48.4508
T	50.9475	39.1293	46.5849	39.8599	37.4682	50.7946	49.0453
Pr(Fit G_{WP})	0.0487	0.0391	0.0313	0.0583	0.0449	0.0474	0.0418
$\widehat{\tau}$	3.0494	3.0522	3.0442	3.0417	3.0465	3.0458	3.0453
S	0.2835	0.2860	0.2819	0.2804	0.2812	0.2822	0.2818
T	0.2791	0.2791	0.2791	0.2791	0.2791	0.2791	0.2791
$\widehat{a}$	11.4374	10.8228	10.6876	10.3732	10.9175	9.8033	13.8485
S	156.9435	78.2220	93.7003	84.4515	89.5536	66.2633	317.3796
T	2.3373	2.3373	2.3373	2.3373	2.3373	2.3373	2.3373
$\widehat{\alpha}_B$	0.3829	0.3771	0.3886	0.3810	0.3765	0.3871	0.3920
S	0.4075	0.3913	0.4118	0.3845	0.3923	0.4006	0.4195
T	0.3395	0.3219	0.3440	0.3140	0.3201	0.3366	0.3531
$\widehat{\beta}_B$	2343.5137	2346.4473	2343.6289	2346.9479	2347.5737	2343.0063	2341.4368
S	94.8757	72.7243	95.6842	67.4172	68.9976	92.0924	101.5984
T	93.3160	71.8427	94.7549	67.3849	68.2134	91.8034	99.7119
$\widehat{B}_{B,10}$	579.3744	579.9322	578.9435	579.0198	579.4777	579.0498	578.3467
S	56.1285	41.5589	54.3195	41.1362	38.6766	55.3059	55.4409
T	55.3705	41.2064	54.2450	41.3158	39.0449	55.3141	54.8604

Table 7.61: Fitting G_{WP} to G_{BA} for $k = 4$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

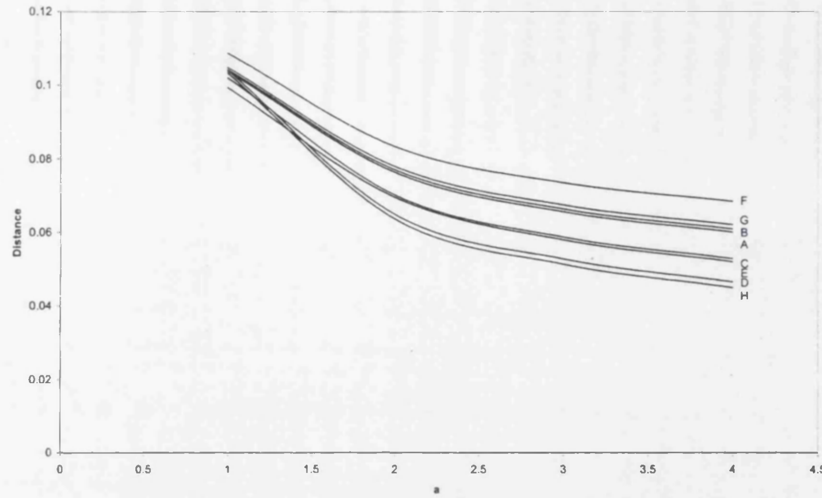


Figure 7.5: Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios, three stress levels and $\tau = 1$. Here, A=$(1:1:1)$, B=$(1:2:3)$, C=$(3:2:1)$, D=$(2:3:1)$, E=$(1:3:2)$, F=$(1:1:3)$, G=$(3:1:1)$, H=$(1:3:1)$.

Three stress levels We now consider results for three stress levels, and again use the same true parameter values and stress levels as those established when running simulations. We also allow τ and a to vary in the same way as above. We not only consider results when we arrange the sample equally amongst stress, but also run simulations for varying ratios; these will be the same as those used in previous investigations. We summarise the results using Figures 7.5, 7.6, 7.7 and 7.8, corresponding to $\tau = 1$, 2, 3 and 4 respectively. For each value of τ, we vary a from 1 to 4, and plot the absolute distance on the vertical axis. The plots clearly show that for small values of τ, as we allow a to increase we see a decrease in the maximum absolute distance between the two cdfs; this is true for all ratios. As we increase τ, the maximum absolute distances for varying a begin to level off, and at times, we observe an increase in the distances as a is highered. Such a fact seems counter-intuitive since we expect there to be an improved fit between Weibull and Burr as a tends to infinity. However, we must now remember that the Weibull Log-linear distribution is not the limiting distribution of the Burr Arrhenius model, and so is perhaps not unusual for particular values of τ, to see an improved fit between these two accelerated distributions with lower a. With regard to how we arrange the samples amongst the three stress levels, we observe the largest distances for varying a across all values of τ for the ratio $(1:1:3)$. The best case is when we have the most observations in the middle stress; that is, for $(1:3:1)$.

Four stress levels We finally consider the affects of fitting the Weibull Log-linear distribution to data with an underlying Burr Arrhenius model when we have 4 stress levels.

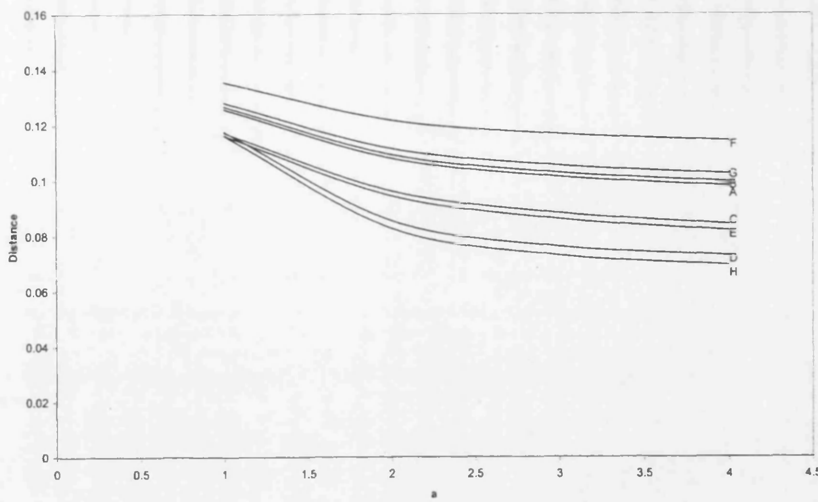


Figure 7.6: Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios, three stress levels and $\tau = 2$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

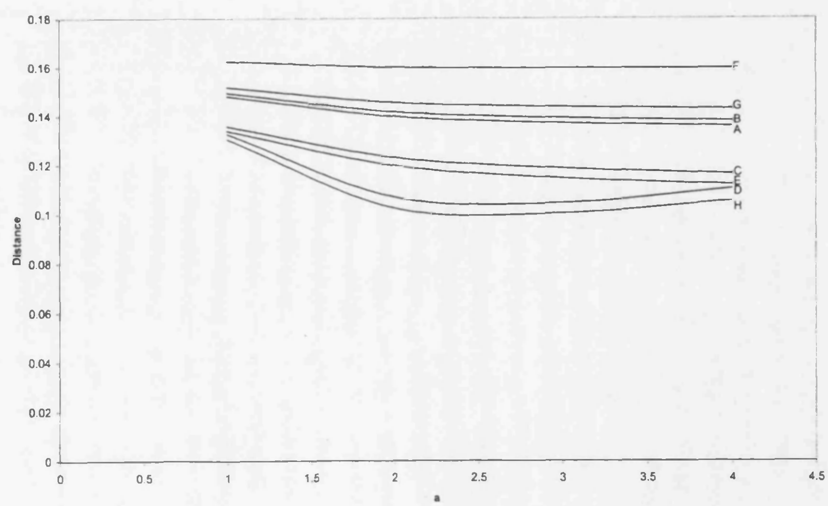


Figure 7.7: Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios, three stress levels and $\tau = 3$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

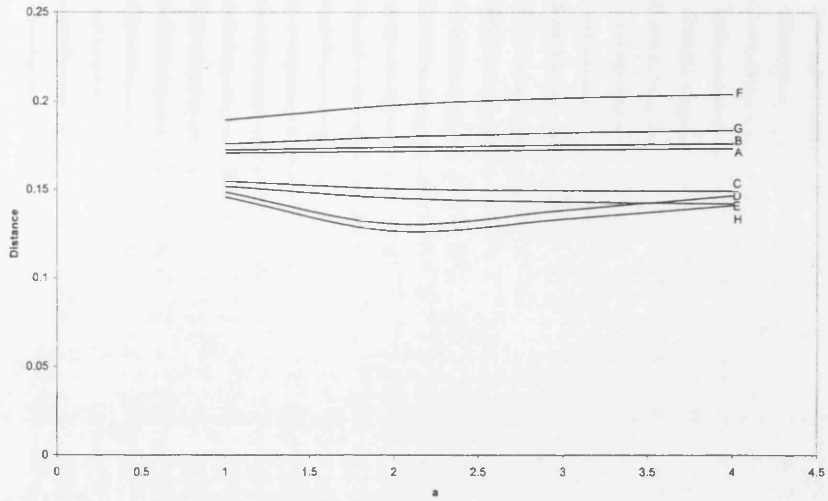


Figure 7.8: Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios, three stress levels and $\tau = 4$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

We approach this in a similar way, and compute the maximum absolute distance between the true and mis-specified distribution functions across the 4 stress levels for varying ratios. Figures 7.9, 7.10, 7.11 and 7.12, summarise the results for τ = 1, 2, 3 and 4 respectively. Again we see similar results with regards to varying τ and a, and for small values of τ, as a is increased, we observe a reduction in the maximum absolute distance between true and mis-specified distribution functions. For larger values of τ, we do not observe this, and again for some ratios, we see the maximum absolute distance increase (slightly) as a increases. On the whole, however, we see the larger distances for larger τ; a result that was not witnessed in the non-accelerated scenario. When we begin to look at the effect that varying ratios has on this distance, we see that the best case occurs for (4 : 3 : 2 : 1) and then, following closely, (1 : 2 : 2 : 1). So again, we see that the lower and middle stresses results in the "best fit" between true and mis-specified distribution functions. As in the case with three stress levels, the largest absolute distance is observed for the ratio (1 : 1 : 1 : 4) so having most of the observations at the highest stress level will result in the worst fit.

7.3.2 Fitting G_{WA} to G_{BP}

We move on to our final scenario concerning the Burr distribution, and examine the effects of fitting the Weibull distribution, now with the Arrhenius scale-stress relationship, to data with an underlying Burr Log-linear model.

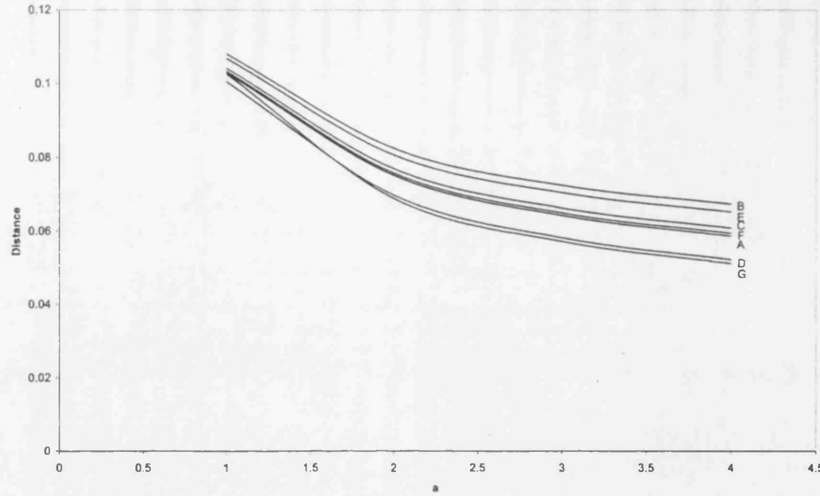


Figure 7.9: Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios, four stress levels and $\tau = 1$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).

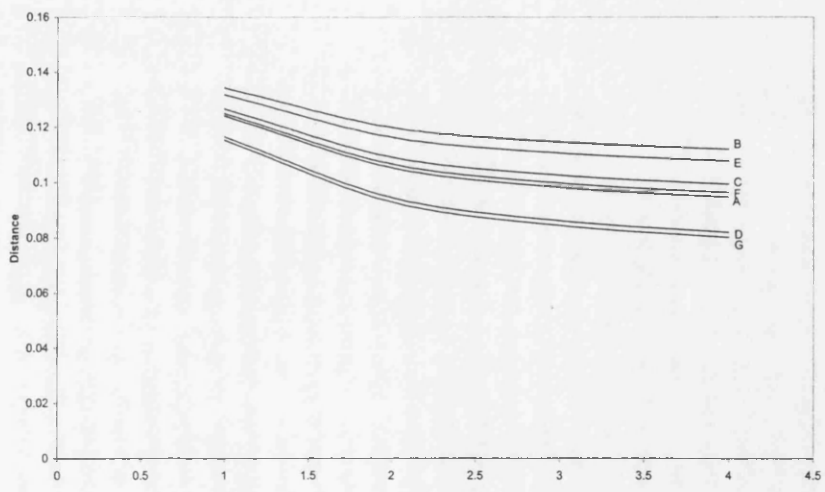


Figure 7.10: Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios, four stress levels and $\tau = 2$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).

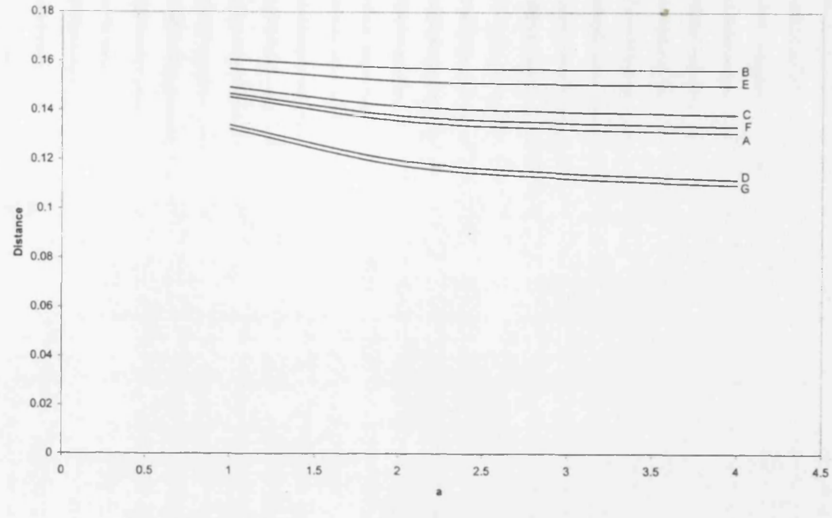


Figure 7.11: Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios, four stress levels and $\tau = 3$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).

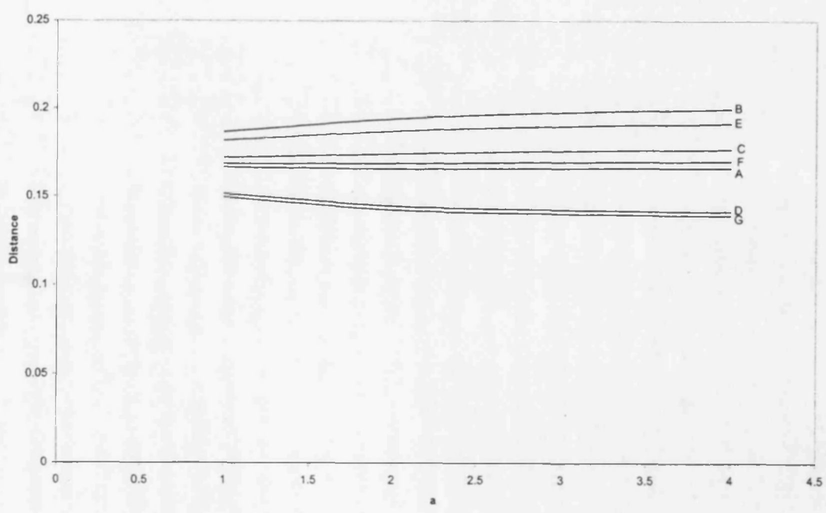


Figure 7.12: Plots of the maximum absolute distance between cdfs for the true Burr Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios, four stress levels and $\tau = 4$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).

Simulation studies

We summarise results for $\tau = 3$, $a = 4$; these are found in Tables 7.62 and 7.63 for $k = 2$, Tables 7.64 and 7.65 for $k = 3$, and Tables 7.66 and 7.67 for $k = 4$. We compare the estimates for B_{10} from true and mis-specified distributions with a true value of 327.7264. We note a number of points concerning the mis-specified distribution; results for fitting the Burr Log-linear distribution have already been considered in previous sections of this chapter. Firstly, the MLEs for the Weibull model approach the corresponding entropy values as the sample size increases. However, we observe poor agreement between sample and theoretical standard errors for small sample sizes, although this agreement does become very good for samples over about 1000. The standard errors of the mis-specified MLEs and $\widehat{B}_{W,10}$ decrease as more observations are placed at the lowest stress levels. We note that even though the summary statistics for the MLEs of G_{WA} do not match up with theoretical counterparts for small sample sizes, we still see good agreement between observed and theoretical standard errors of $\widehat{B}_{W,10}$. We always under-estimate the time to which 10% of observations fail for the mis-specified model.

The effects of mis-specification

Just as in previous cases, we also consider how theoretical true and mis-specified distribution functions match up for varying Burr parameters, stress levels and sample sizes. We do not need to consider results for 2 stress levels since they are covered by the non-accelerated case (for reasons given previously). Thus, we consider results for 3 and 4 stress levels below.

Three stress levels We construct plots of the maximum absolute distance between G_{WA} and G_{BP} for varying τ and a. Figures 7.13, 7.14, 7.15 and 7.16, show these distances for $\tau = 1$, 2, 3 and 4 respectively, and for varying ratios. The plots show that, on the whole, increasing τ also increases the maximum absolute distance between true and mis-specified distribution functions, and this seems true regardless of how we allocate observations to each stress level. Generally, for small values of τ, the maximum absolute distance decreases as a is highered. Contrary to this, for larger values of τ, we observe this distance increasing (slightly) as a is increased. When we begin to vary the ratios, we see the best cases, in terms of minimum absolute distances, when we have larger numbers of observations at the higher stress levels. The worst cases occur for the ratios $(1:3:1)$ and $(2:3:1)$, or when we have most of the observations at the middle and lower stress levels.

Four stress levels Four stress levels result in the same conclusions as those for $k = 3$. The results are illustrated by Figures 7.17, 7.18, 7.19 and 7.20.

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_A$	2.6311	2.5937	2.5697	2.5606
$\widetilde{B}_A$	2.5528	2.5528	2.5528	2.5528
S	0.2269	0.1590	0.1024	0.0731
T	0.2335	0.1651	0.1044	0.0738
$\widehat{\alpha}_{WA}$	-0.9183	-2.1916	-2.8125	-2.8593
$\widetilde{\alpha}_{WA}$	-2.8598	-2.8598	-2.8598	-2.8598
S	4.0709	2.5650	0.7108	0.0753
T	0.2385	0.1687	0.1067	0.0754
$\widehat{\beta}_{WA}$	2490.0473	2862.4936	3044.0955	3057.8655
$\widetilde{\beta}_{WA}$	3058.1277	3058.1277	3058.1277	3058.1277
S	1191.8408	751.1804	209.3050	28.4964
T	90.2095	63.7878	40.3429	28.5268
$\widehat{B}_{W,10}$	312.4045	309.0989	306.9090	306.1368
S	29.9556	21.0803	13.5991	9.7070
T	30.9027	21.8515	13.8201	9.7723
Pr (Fit G_{WA})	0.0003	0.0001	0	0
$\widehat{\tau}$	3.1310	3.0488	3.0067	3.0025
S	0.4064	0.2836	0.1804	0.1266
T	0.3948	0.2791	0.1765	0.1248
$\widehat{a}$	13.0137	9.3411	5.6571	4.4205
S	141.4013	47.4201	13.0892	1.9932
T	3.3055	2.3373	1.4783	1.0453
$\widehat{\alpha}_B$	8.0613	8.0754	8.0495	8.0204
S	0.4149	0.3415	0.2215	0.1338
T	0.3779	0.2672	0.1690	0.1195
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0006	0.0004	0.0002	0.0002
T	0.0005	0.0004	0.0002	0.0002
$\widehat{B}_{B,10}$	333.6675	330.0769	328.0808	327.8812
S	31.3518	22.0852	14.1332	9.9808
T	31.5477	22.3076	14.1085	9.9762

Table 7.62: Fitting G_{WA} to G_{BP} for $k = 2$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_A$	2.5929	2.5939	2.5915	2.5914
$\widetilde{B}_A$	2.5528	2.5528	2.5528	2.5528
S	0.1623	0.1617	0.1609	0.1600
T	0.1651	0.1651	0.1651	0.1651
$\widehat{\alpha}_{WA}$	-2.1419	-2.1872	-2.2154	-2.1972
$\widetilde{\alpha}_{WA}$	-2.8598	-2.8598	-2.8598	-2.8598
S	2.6123	2.6304	2.5036	2.5727
T	0.2176	0.2881	0.1762	0.2119
$\widehat{\beta}_{WA}$	2844.5732	2862.2349	2868.1802	2864.6804
$\widetilde{\beta}_{WA}$	3058.1277	3058.1277	3058.1277	3058.1277
S	767.2886	770.7415	764.0120	753.5567
T	96.4381	96.4381	73.6558	73.6558
$\widehat{B}_{W,10}$	307.2960	309.2422	308.2258	309.0636
S	31.6516	19.4696	25.1324	19.7492
T	32.0250	19.9796	25.6942	20.4104
Pr (Fit G_{WA})	0.0001	0.0002	0.0001	0.0001
$\widehat{\tau}$	3.0436	3.0486	3.0401	3.0431
S	0.2830	0.2826	0.2816	0.2820
T	0.2791	0.2791	0.2791	0.2791
$\widehat{a}$	10.1524	10.9695	9.6912	10.6845
S	67.3746	118.3992	48.4110	86.7535
T	2.3373	2.3373	2.3373	2.3373
$\widehat{\alpha}_B$	8.0762	8.0784	8.0797	8.0811
S	0.3608	0.3429	0.3470	0.3469
T	0.2832	0.2659	0.2726	0.2657
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0006	0.0006	0.0005	0.0005
T	0.0006	0.0006	0.0004	0.0004
$\widehat{B}_{B,10}$	329.3066	330.1567	329.3965	329.8428
S	32.6550	20.4139	26.0266	20.8049
T	32.1660	20.5161	26.0122	20.9274

Table 7.63: Fitting G_{WA} to G_{BP} for $k = 2$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	25,25 25	50,50 50	100,100 100	300,300 300	500,500 500
$\widehat{B}_A$	2.5743	2.5137	2.4758	2.4571	2.4531
$\widetilde{B}_A$	2.4494	2.4494	2.4494	2.4494	2.4494
S	0.2741	0.1884	0.1282	0.0737	0.0565
T	0.2570	0.1818	0.1285	0.0742	0.0575
$\widehat{\alpha}_{WA}$	0.0044	-1.3902	-2.2849	-2.5362	-2.5379
$\widetilde{\alpha}_{WA}$	-2.5396	-2.5396	-2.5396	-2.5396	-2.5396
S	4.3912	3.2140	1.5888	0.1475	0.0705
T	0.3145	0.2224	0.1573	0.0908	0.0703
$\widehat{\beta}_{WA}$	2219.4157	2628.4535	2890.3614	2964.0800	2964.6830
$\widetilde{\beta}_{WA}$	2965.3355	2965.3355	2965.3355	2965.3355	2965.3355
S	1286.6805	941.9352	466.5547	49.0066	27.5515
T	122.7847	86.8219	61.3923	35.4449	27.4555
$\widehat{B}_{W,10}$	312.9113	309.2226	306.3581	304.8551	304.5105
S	38.7522	27.2227	19.1025	11.3288	8.7576
T	39.4804	27.9169	19.7402	11.3970	8.8281
Pr (Fit G_{WA})	0.0792	0.0238	0.0036	0	0
$\widehat{\tau}$	3.1896	3.0772	3.0188	3.0025	3.0019
S	0.4708	0.3266	0.2264	0.1334	0.1032
T	0.4558	0.3223	0.2279	0.1316	0.1019
$\widehat{a}$	12.1471	10.8864	7.4228	4.5363	4.2454
S	73.4434	70.0025	28.2443	2.9100	1.1104
T	3.8169	2.6989	1.9084	1.1018	0.8535
$\widehat{\alpha}_B$	8.0328	8.0757	8.0699	8.0248	8.0128
S	0.4362	0.3780	0.2905	0.1463	0.1057
T	0.4409	0.3118	0.2205	0.1273	0.0956
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0008	0.0005	0.0004	0.0002	0.0002
T	0.0008	0.0005	0.0004	0.0002	0.0002
$\widehat{B}_{B,10}$	335.8273	331.7626	328.8023	327.8628	327.7955
S	39.1739	27.3566	19.0032	11.2311	8.7806
T	38.9171	27.5185	19.4585	11.2344	8.7021

Table 7.64: Fitting G_{WA} to G_{BP} for $k = 3$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	50,100	25,200	25,25	200,50
n_3	150	75	250	50
$\widehat{B}_A$	2.4644	2.4812	2.5326	2.5275
$\widetilde{B}_A$	2.4376	2.4562	2.5058	2.4976
S	0.1300	0.1260	0.1321	0.1298
T	0.1284	0.1271	0.1341	0.1317
$\widehat{\alpha}_{WA}$	-2.5118	-2.4334	-2.6432	-2.1916
$\widetilde{\alpha}_{WA}$	-2.7794	-2.7074	-2.9579	-2.4527
S	1.6354	1.6350	1.7499	1.6225
T	0.1733	0.2445	0.2091	0.1823
$\widehat{\beta}_{WA}$	2980.7833	2980.0484	3016.6529	2852.7812
$\widetilde{\beta}_{WA}$	3060.7731	3063.4394	3112.5449	2928.5931
S	485.2240	491.0923	520.6420	474.0085
T	72.7128	103.0354	93.8625	64.1217
$\widehat{B}_{W,10}$	321.8117	349.9540	322.8335	304.4550
S	23.9837	32.2922	30.5449	16.6461
T	24.2486	32.1971	31.2657	17.1469
Pr (Fit G_{WA})	0.0025	0.0023	0.0221	0.0138
$\widehat{\tau}$	3.0166	3.0185	3.0209	3.0234
S	0.2305	0.2300	0.2308	0.2305
T	0.2279	0.2279	0.2279	0.2279
$\widehat{a}$	7.4995	7.8507	7.8015	8.2502
S	36.5875	47.2980	53.3888	56.5583
T	1.9084	1.9084	1.9084	1.9084
$\widehat{\alpha}_B$	8.0729	8.0693	8.0636	8.0707
S	0.2937	0.3032	0.3045	0.2952
T	0.2257	0.2350	0.2375	0.2176
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0004	0.0006	0.0006	0.0004
T	0.0004	0.0006	0.0006	0.0004
$\widehat{B}_{B,10}$	328.7831	328.9442	328.3455	329.2200
S	23.1166	27.2262	29.3489	17.1109
T	22.7604	26.8589	29.3188	17.3271

Table 7.65: Fitting G_{WA} to G_{BP} for $k = 3$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_A$	2.5659	2.5160	2.4935	2.4824	2.4769
$\widetilde{B}_A$	2.4737	2.4737	2.4737	2.4737	2.4737
S	0.2328	0.1608	0.1120	0.0791	0.0501
T	0.2258	0.1597	0.1129	0.0798	0.0505
$\widehat{\alpha}_{WA}$	-0.6445	-1.8273-	2.4308	-2.5409	-2.5488
$\widetilde{\alpha}_{WA}$	-2.5481	-2.5481	-2.5481	-2.5481	-2.5481
S	3.9483	2.6031	1.0822	0.2545	0.0630
T	0.2828	0.2000	0.1414	0.1000	0.0632
$\widehat{\beta}_{WA}$	2409.3113	2756.9971	2933.8747	2966.3413	2968.9940
$\widetilde{\beta}_{WA}$	2968.7477	2968.7477	2968.7477	2968.7477	2968.7477
S	1157.6457	763.4110	319.2386	79.4281	25.4474
T	114.2677	80.7995	57.1338	40.3997	25.5510
$\widehat{B}_{W,10}$	313.3485	310.6879	309.2190	308.3278	307.9659
S	35.5795	25.3977	18.2469	12.9908	8.2022
T	36.7644	25.9963	18.3822	12.9982	8.2208
Pr (Fit G_{WP})	0.0680	0.0204	0.0024	0.0001	0
$\widehat{\tau}$	3.1356	3.0405	3.0129	3.0022	3.0008
S	0.4042	0.2793	0.2014	0.1417	0.0894
T	0.3948	0.2791	0.1974	0.1396	0.0883
$\widehat{a}$	12.4597	12.1503	6.3180	4.6940	4.1819
S	81.1487	310.4447	27.1407	6.5445	0.8834
T	3.3055	2.3373	1.6528	1.1687	0.7391
$\widehat{\alpha}_B$	8.0535	8.0821	8.0569	8.0288	8.0108
S	0.4236	0.3535	0.2525	0.1612	0.0907
T	0.3857	0.2727	0.1928	0.1364	0.0862
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0007	0.0005	0.0004	0.0002	0.0002
T	0.0007	0.0005	0.0004	0.0003	0.0002
$\widehat{B}_{B,10}$	333.5788	329.8369	328.5211	327.8278	327.8664
S	36.2003	25.4794	18.1793	12.7792	8.0859
T	36.0850	25.5159	18.0425	12.7580	8.0688

Table 7.66: Fitting G_{WA} to G_{BP} for $k = 4$, $\tau = 3$, $a = 4$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 75,75	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_A$	2.5321	2.5284	2.4889	2.5528	2.5408	2.5257	2.5256
$\widetilde{B}_A$	2.4896	2.4863	2.4472	2.5108	2.4950	2.4847	2.4832
S	0.1637	0.1605	0.1625	0.1607	0.1609	0.1620	0.1614
T	0.1614	0.1592	0.1570	0.1626	0.1605	0.1616	0.1597
$\widehat{\alpha}_{WA}$	-2.0216	-1.6171	-2.0324	-1.8312	-1.6974	-2.1027	-1.7991
$\widetilde{\alpha}_{WA}$	-2.7293	-2.2939	-2.7590	-2.4832	-2.3503	-2.8001	-2.5164
S	2.5830	2.5013	2.6303	2.4811	2.4790	2.5770	2.5767
T	0.2336	0.2058	0.2398	0.1889	0.2033	0.2286	0.2559
$\widehat{\beta}_{WA}$	2823.9574	2683.6207	2845.5104	2749.2517	2706.6986	2851.4458	2757.6568
$\widetilde{\beta}_{WA}$	3034.7899	2881.0723	3062.2480	2939.8958	2896.8553	3059.4486	2970.6783
S	762.0345	729.9979	780.5498	725.2429	723.5422	761.3916	760.2387
T	100.6804	78.5299	101.6756	73.1489	74.9213	99.0196	107.4571
$\widehat{B}_{W,10}$	318.2669	307.3930	330.5742	307.8014	306.6119	319.1909	322.2519
S	32.0760	22.8371	32.6357	22.5970	21.5958	31.7510	32.2603
T	32.5733	23.3869	33.0817	23.3735	22.0279	32.6185	32.7027
Pr(Fit G_{WA})	0.0293	0.0204	0.0112	0.0454	0.0302	0.0316	0.0226
$\widehat{\tau}$	3.0436	3.0422	3.0463	3.0465	3.0447	3.0449	3.0435
S	0.2786	0.2818	0.2860	0.2848	0.2819	0.2816	0.2841
T	0.2791	0.2791	0.2791	0.2791	0.2791	0.2791	0.2791
$\widehat{a}$	9.7433	10.7785	9.2639	9.4210	9.6930	9.3998	8.9060
S	55.4236	108.8897	54.2815	39.1534	79.5123	56.2659	33.7493
T	2.3373	2.3373	2.3373	2.3373	2.3373	2.3373	2.3373
$\widehat{\alpha}_B$	8.0755	8.0831	8.0750	8.0803	8.0821	8.0734	8.0792
S	0.3548	0.3493	0.3544	0.3477	0.3414	0.3527	0.3542
T	0.2822	0.2697	0.2807	0.2695	0.2678	0.2819	0.2825
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0006	0.0005	0.0006	0.0005	0.0005	0.0006	0.0007
T	0.0006	0.0005	0.0006	0.0005	0.0005	0.0006	0.0007
$\widehat{B}_{B,10}$	329.6546	329.9942	330.0711	330.1087	330.2417	329.7167	330.0628
S	31.2791	23.1895	30.2095	23.4570	22.2289	31.0719	30.9744
T	31.0655	23.3609	30.0066	23.4804	22.1892	31.0237	30.5681

Table 7.67: Fitting G_{WA} to G_{BP} for $k = 4$, $\tau = 3$, $a = 4$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

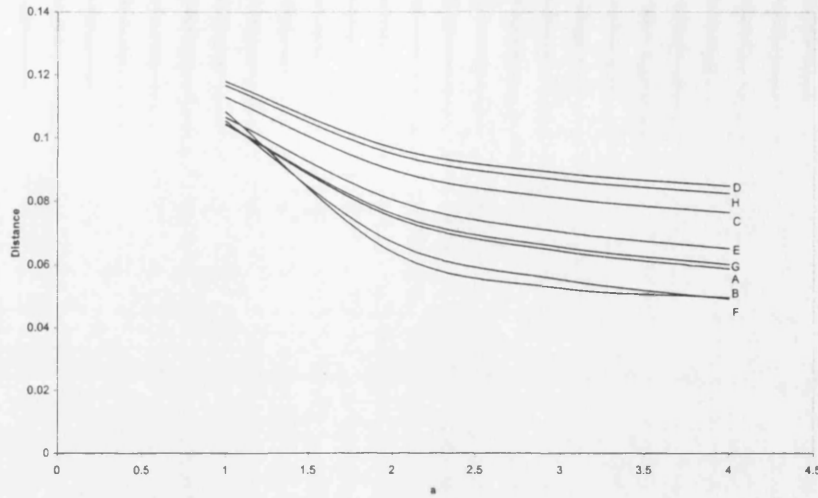


Figure 7.13: Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios, three stress levels and $\tau = 1$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

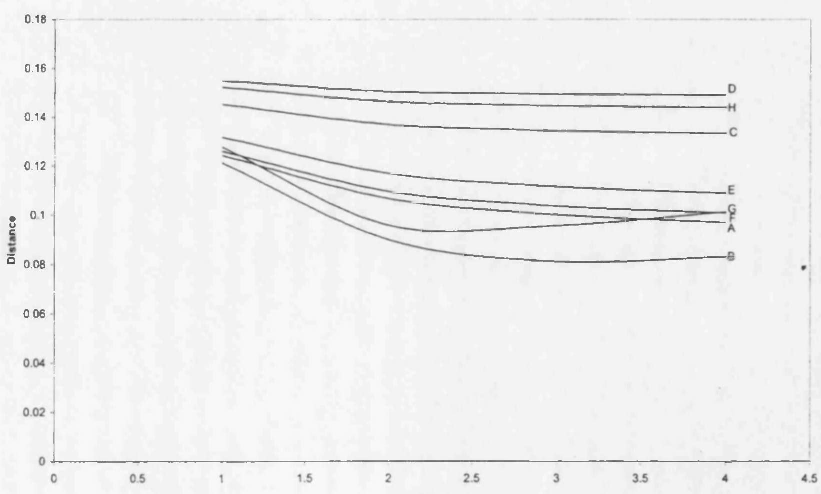


Figure 7.14: Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios, three stress levels and $\tau = 2$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

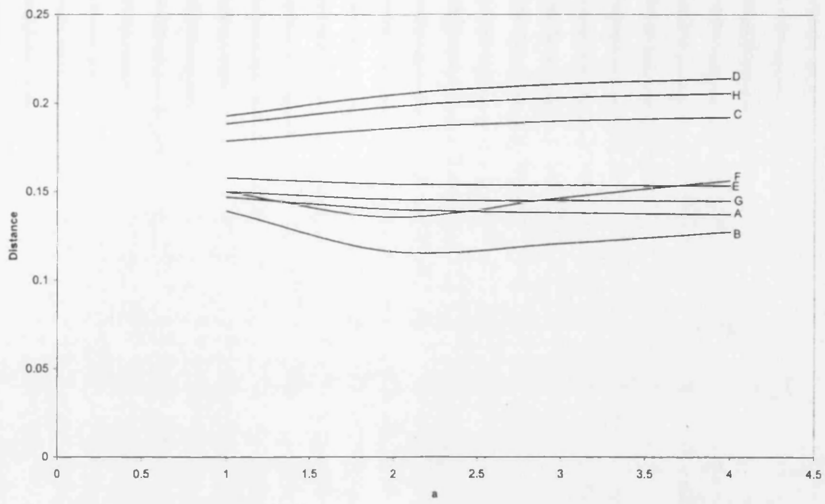


Figure 7.15: Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios, three stress levels and $\tau = 3$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

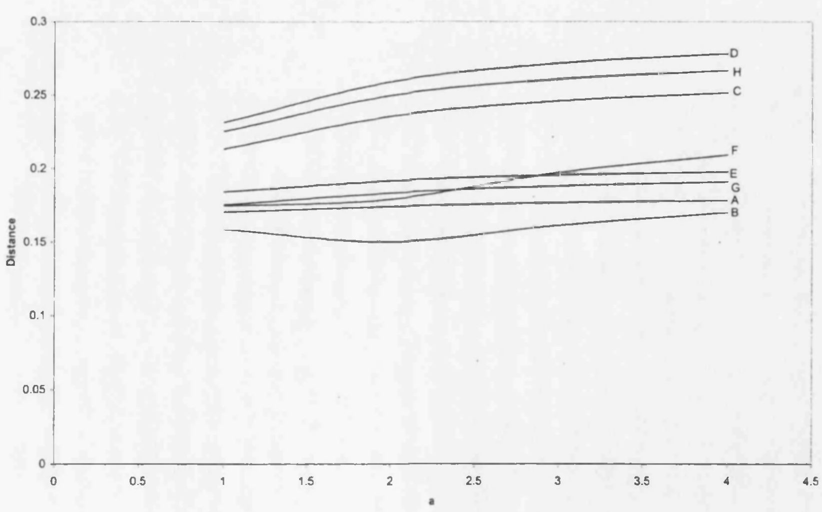


Figure 7.16: Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios, three stress levels and $\tau = 4$. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

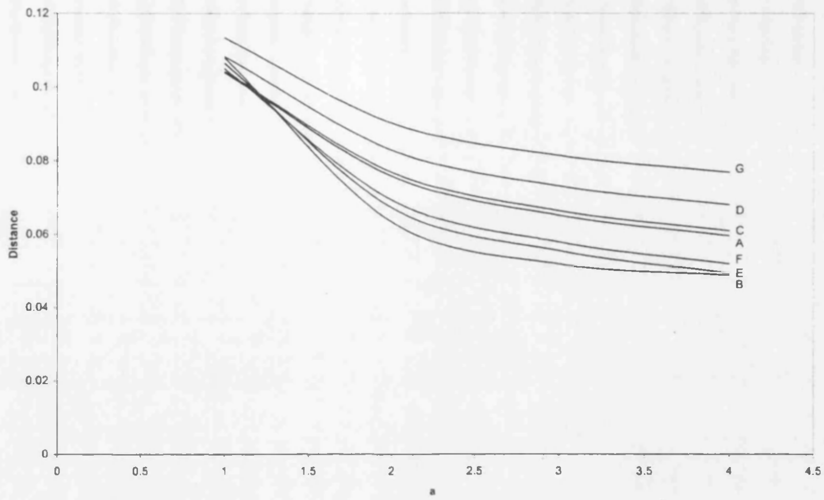


Figure 7.17: Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios, four stress levels and $\tau = 1$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).

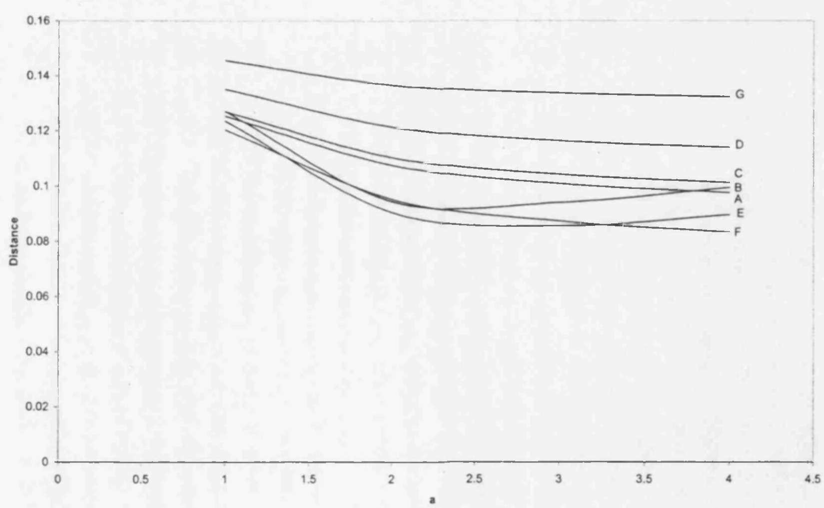


Figure 7.18: Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios, four stress levels and $\tau = 2$. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).

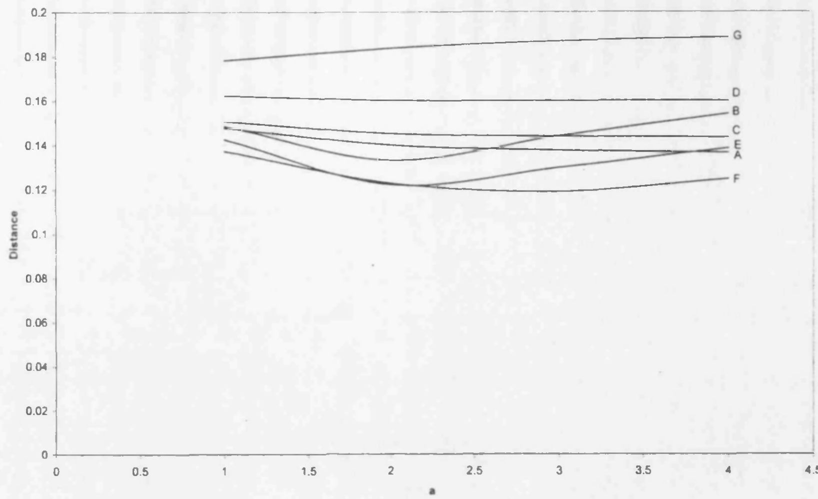


Figure 7.19: Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios, four stress levels and $\tau = 3$. Here, A=$(1:1:1:1)$, B=$(1:1:1:4)$, C=$(4:1:1:1)$, D=$(1:2:2:1)$, E=$(2:1:1:2)$, F=$(1:2:3:4)$, G=$(4:3:2:1)$.

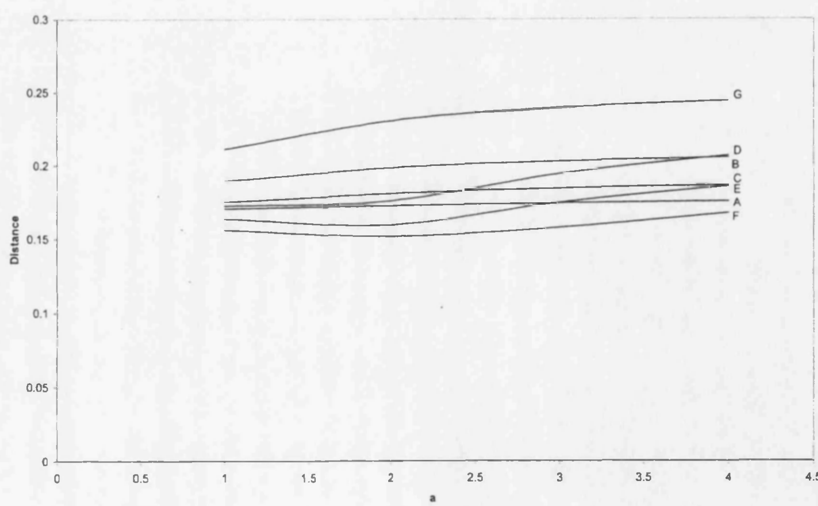


Figure 7.20: Plots of the maximum absolute distance between cdfs for the true Burr Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios, four stress levels and $\tau = 4$. Here, A=$(1:1:1:1)$, B=$(1:1:1:4)$, C=$(4:1:1:1)$, D=$(1:2:2:1)$, E=$(2:1:1:2)$, F=$(1:2:3:4)$, G=$(4:3:2:1)$.

7.3.3 Fitting G_{WP} to G_{GA}

We continue with mis-specifying both scale-stress relationship and distribution function, and consider the Gamma distribution as the underlying model.

Simulation Studies

We run simulations to assess the effects of fitting G_{WP} to data from G_{GA} for $\tau = 0.5$, 1, 2 and 3, although only summarise the results for $\tau = 3$. These are shown in Tables 7.68 and 7.69 for $k = 2$, Tables 7.70 and 7.71 for $k = 3$, and Tables 7.72 and 7.73 for $k = 4$. We compare sample values of B_{10} with a true value of $B_{G,10} = 2121.4228$, and list a number of key results. Firstly, when we fit G_{GA}, we observe good agreement between observed and theoretical results across all stress levels and all values of τ. $\widehat{B}_{G,10}$ is also very close to its true value, even for small sample sizes. The theoretical standard error for $\widehat{\tau}$ remains constant when varying loadings at each stress level. For the remaining Gamma MLEs and $\widehat{B}_{10}$, larger standard errors are observed when we have more observations in the middle and higher stress levels. Entropy values do not vary when $k = 2$, and when the n_i are equal. We also see good agreement between the MLEs and entropy values for large sample sizes. The standard errors of the MLEs from the mis-specified distribution match up well, even for small sample sizes. This is also true for $\widehat{B}_{W,10}$. When $\tau = 0.5$, we always over-estimate B_{10} from the mis-specified distribution. This quantile is then under-estimated when $\tau = 2$ and 3. For $\tau = 1$, we are just fitting the Negative Exponential model, which is a special case of the Weibull. Thus, we observe much better agreement between $\widehat{B}_{10}$ from both distributions. The agreement for $k = 2$ is excellent, since we are fitting the same model. For $k = 3$ and 4, $\widehat{B}_{W,10}$ is slightly under-estimated, since we are just mis-specifying the scale-stress relationship. For $\tau = 0.5, 2$ and 3, the probability for fitting the wrong distribution function is as high as 21%. For $\tau = 1$ and $k = 2$, we fit the wrong distribution half the time. This does decrease when we have more stress levels, but for small sample sizes, is still as high as 36%.

The effects of mis-specification

We examine the effects of mis-specifying the Weibull Log-linear model in the usual way. Thus, we construct tables of the maximum absolute distance between the cdf of the true distribution function, and the cdf of the mis-specified distribution, with parameter values replaced by entropy values. We do this for a range of true parameters, varying stress levels and loadings. As in previous scenarios, we do not examine the effects for $k = 2$, since only the distribution function is mis-specified. Results from the non-accelerated case then hold for such cases. We consider three and four levels below.

Three stress levels We construct plots of the maximum absolute distances for τ varying between 0.1 and 5 in steps of 0.1. This is done for varying ratios at each of the three

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_P$	1.8858	1.8567	1.8436	1.8367
$\widetilde{B}_P$	1.8328	1.8328	1.8328	1.8328
S	0.1480	0.1022	0.0639	0.0458
T	1.4291	0.1011	0.0639	0.0452
$\widehat{\alpha}_{WP}$	9.5442	9.5488	9.5504	9.5498
$\widetilde{\alpha}_{WP}$	9.5502	9.5502	9.5502	9.5502
S	0.1119	0.0810	0.0512	0.0369
T	0.1158	0.0819	0.0518	0.0366
$\widehat{\beta}_{WP}$	-0.0153	-0.0154	-0.0154	-0.0153
$\widetilde{\beta}_{WP}$	-0.0153	-0.0153	-0.0153	-0.0153
S	0.0008	0.0006	0.0004	0.0003
T	0.0008	0.0006	0.0004	0.0003
$\widehat{B}_{W,10}$	1969.4289	1938.4223	1923.9501	1914.5792
S	250.7819	173.9423	109.0894	79.0387
T	245.0317	173.2636	109.5815	77.4858
Pr (Fit G_{WP})	0.2353	0.1333	0.0332	0.0041
$\widehat{\tau}$	3.1313	3.0597	3.0257	3.0087
S	0.4364	0.2946	0.1820	0.1293
T	0.4029	0.2849	0.1802	0.1274
$\widehat{\alpha}_G$	0.2717	0.2842	0.2899	0.2971
S	0.3350	0.2399	0.1517	0.1072
T	0.3396	0.2401	0.1519	0.1074
$\widehat{\beta}_G$	2343.9531	2347.0013	2348.0354	2347.2253
S	113.8542	82.8970	52.3818	37.2533
T	117.7074	83.2317	52.6403	37.2223
$\widehat{B}_{G,10}$	2164.2842	2140.4505	2131.7348	2124.1355
S	262.0951	182.2863	114.7782	82.3657
T	257.4616	182.0528	115.1403	81.4165

Table 7.68: Fitting G_{WP} to G_{GA} for $k = 2$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_P$	1.8585	1.8542	1.8574	1.8580
$\widetilde{B}_P$	1.8328	1.8328	1.8328	1.8328
S	0.1033	0.1035	0.1032	0.1018
T	0.1010	0.1011	0.1010	0.1011
$\widehat{\alpha}_{WP}$	9.5339	9.5532	9.5442	9.5511
$\widetilde{\alpha}_{WP}$	9.5502	9.5502	9.5502	9.5502
S	0.1588	0.0716	0.1134	0.0700
T	0.1608	0.0718	0.1141	0.0704
$\widehat{\beta}_{WP}$	-0.0153	-0.0154	-0.0153	-0.0154
$\widetilde{\beta}_{WP}$	-0.0153	-0.0153	-0.0153	-0.0153
S	0.0008	0.0009	0.0007	0.0007
T	0.0008	0.0009	0.0007	0.0007
$\widehat{B}_{W,10}$	1930.1502	1935.3409	1936.6403	1940.8099
S	264.6225	157.8412	211.2009	160.5647
T	264.0931	156.3722	207.8950	160.2365
Pr (Fit G_{WP})	0.1364	0.1254	0.1310	0.1314
$\widehat{\tau}$	3.0633	3.0546	3.0595	3.0580
S	0.2968	0.2959	0.2980	0.2916
T	0.2849	0.2849	0.2849	0.2849
$\widehat{\alpha}_G$	0.2979	0.2692	0.2878	0.2779
S	0.2991	0.3889	0.2490	0.2934
T	0.2998	0.3880	0.2491	0.2927
$\widehat{\beta}_G$	2340.0995	2352.2132	2345.0756	2349.0671
S	125.3577	126.3607	95.6845	96.1282
T	125.8345	125.8345	96.1077	96.1077
$\widehat{B}_{G,10}$	2142.1233	2138.6576	2141.8074	2141.3908
S	282.0263	164.1849	221.4711	168.0294
T	279.5489	163.4443	219.4190	167.7580

Table 7.69: Fitting G_{WP} to G_{GA} for $k = 2$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	25,25 25	50,50 50	100,100 100	300,300 300	500,500 500
$\widehat{B}_P$	1.8886	1.8445	1.8271	1.8152	1.8126
$\widetilde{B}_P$	1.8100	1.8100	1.8100	1.8100	1.8100
S	0.1701	0.1173	0.0833	0.0475	0.0365
T	0.1637	0.1157	0.0818	0.0472	0.0366
$\widehat{\alpha}_{WP}$	9.5045	9.5148	9.5207	9.5224	9.5224
$\widetilde{\alpha}_{WP}$	9.5226	9.5226	9.5226	9.5226	9.5226
S	0.1483	0.1121	0.0812	0.0472	0.0363
T	0.1626	0.1150	0.0813	0.0470	0.0364
$\widehat{\beta}_{WP}$	-0.0155	-0.0155	-0.0156	-0.0156	-0.0156
$\widetilde{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
S	0.0010	0.0008	0.0006	0.0003	0.0002
T	0.0011	0.0008	0.0006	0.0003	0.0002
$\widehat{B}_{W,10}$	1886.0999	1843.9178	1829.1602	1815.9601	1812.3937
S	294.1220	209.5183	148.4150	85.9323	65.3027
T	294.2341	208.0549	147.1171	84.9381	65.7928
Pr (Fit G_{WP})	0.2037	0.0958	0.0287	0.0003	0
$\widehat{\tau}$	3.1947	3.0832	3.0398	3.0126	3.0048
S	0.5120	0.3429	0.2394	0.1350	0.1042
T	0.4652	0.3290	0.2326	0.1343	0.1040
$\widehat{\alpha}_G$	0.2811	0.2893	0.2918	0.2973	0.2992
S	0.4146	0.3060	0.2192	0.1272	0.0976
T	0.4380	0.3097	0.2190	0.1264	0.0979
$\widehat{\beta}_G$	2333.0177	2341.8482	2346.0061	2346.7524	2346.8404
S	147.8304	110.8130	80.1104	46.3866	35.6948
T	160.1557	113.2472	80.0779	46.2330	35.8119
$\widehat{B}_{G,10}$	2177.7044	2145.6503	2135.9864	2126.3441	2122.7842
S	326.6067	232.0168	164.6649	94.7214	72.4352
T	325.5523	230.2003	162.7762	93.9789	72.7957

Table 7.70: Fitting G_{WP} to G_{GA} for $k = 3$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	50,100 150	25,200 75	25,25 250	200,50 50
$\widehat{B}_P$	1.8265	1.8245	1.8413	1.8374
$\widetilde{B}_P$	1.8096	1.8073	1.8255	1.8207
S	0.0823	0.0826	0.0833	0.0831
T	0.0818	0.0822	0.0822	0.0821
$\widehat{\alpha}_{WP}$	9.4645	9.3792	9.4785	9.5598
$\widetilde{\alpha}_{WP}$	9.4692	9.3871	9.4930	9.5599
S	0.1078	0.1419	0.1515	0.0618
T	0.1083	0.1450	0.1540	0.0623
$\widehat{\beta}_{WP}$	-0.0152	-0.0150	-0.0150	-0.0158
$\widetilde{\beta}_{WP}$	-0.0152	-0.0151	-0.0151	-0.0158
S	0.0006	0.0009	0.0008	0.0006
T	0.0007	0.0009	0.0008	0.0006
$\widehat{B}_{W,10}$	1764.7723	1634.7786	1825.6502	1892.5930
S	169.0839	185.3710	229.6118	132.7864
T	158.6162	187.6101	231.2229	132.4201
Pr (Fit G_{WP})	0.0290	0.0320	0.0584	0.0413
$\widehat{\tau}$	3.0381	3.0360	3.0373	3.0410
S	0.2356	0.2371	0.2386	0.2382
T	0.2326	0.2326	0.2326	0.2326
$\widehat{\alpha}_G$	0.2951	0.3022	0.3019	0.2856
S	0.2416	0.3287	0.2818	0.2407
T	0.2420	0.3319	0.2828	0.2435
$\widehat{\beta}_G$	2344.6815	2342.0123	2341.5036	2347.7067
S	95.8303	134.3377	121.3654	80.4228
T	96.1059	135.9118	121.8186	81.3064
$\widehat{B}_{G,10}$	2135.6691	2135.4372	2134.5045	2133.8017
S	198.8266	245.5949	260.9138	140.1467
T	198.7592	248.4590	261.9726	139.7066

Table 7.71: Fitting G_{WP} to G_{GA} for $k = 3$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_P$	1.8713	1.8401	1.8279	1.8213	1.8176
$\widetilde{B}_P$	1.8152	1.8152	1.8152	1.8152	1.8152
S	0.1454	0.1007	0.0716	0.0504	0.0320
T	0.1421	0.1005	0.0711	0.0502	0.0318
$\widehat{\alpha}_{WP}$	9.5042	9.5164	9.5233	9.5242	9.5249
$\widetilde{\alpha}_{WP}$	9.5258	9.5258	9.5258	9.5258	9.5258
S	0.1518	0.1130	0.0807	0.0568	0.0363
T	0.1619	0.1145	0.0809	0.0572	0.0362
$\widehat{\beta}_{WP}$	-0.0155	-0.0156	-0.0156	-0.0156	-0.0156
$\widetilde{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
S	0.0010	0.0007	0.0005	0.0004	0.0002
T	0.0010	0.0007	0.0005	0.0004	0.0002
$\widehat{B}_{W,10}$	1863.0284	1838.3739	1830.1088	1822.6616	1818.4440
S	274.4307	194.2362	137.5725	98.5035	62.1574
T	275.2209	194.6106	137.6104	97.3053	61.5413
Pr (Fit G_{WP})	0.1773	0.0767	0.0183	0.0015	0
$\widehat{\tau}$	3.1342	3.0589	3.0305	3.0140	3.0049
S	0.4296	0.2936	0.2042	0.1437	0.0908
T	0.4029	0.2849	0.2015	0.1425	0.0901
$\widehat{\alpha}_G$	0.2979	0.2988	0.2924	0.2973	0.2996
S	0.3786	0.2786	0.1968	0.1388	0.0883
T	0.3940	0.2786	0.1970	0.1393	0.0881
$\widehat{\beta}_G$	2332.2639	2340.7806	2346.4281	2346.4471	2346.5451
S	141.0717	105.1124	74.7431	52.6190	33.4613
T	149.6749	105.8361	74.8374	52.9181	33.4683
$\widehat{B}_{G,10}$	2153.2608	2136.6187	2132.2618	2126.0382	2122.4586
S	304.8050	218.1766	153.4075	109.4524	68.9412
T	306.4006	216.6579	153.2003	108.3290	68.5133

Table 7.72: Fitting G_{WP} to G_{GA} for $k = 4$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	75,75	25,75	75,25	100,50	25,25	25,75
n_3, n_4	75,75	25,25	25,75	75,25	25,25	50,100	75,25
$\widehat{B}_P$	1.8440	1.8399	1.8348	1.8501	1.8447	1.8453	1.8424
$\widetilde{B}_P$	1.8199	1.8147	1.8096	1.8234	1.8183	1.8198	1.8159
S	0.1028	0.1020	0.1022	0.1034	0.1029	0.1029	0.1017
T	0.1006	0.1007	0.1004	0.1008	0.1007	0.1006	0.1007
$\widehat{\alpha}_{WP}$	9.4841	9.5506	9.4358	9.5552	9.5618	9.4769	9.4766
$\widetilde{\alpha}_{WP}$	9.4996	9.5558	9.4470	9.5612	9.5639	9.4939	9.4919
S	0.1527	0.0949	0.1474	0.0942	0.0838	0.1538	0.1535
T	0.1568	0.0982	0.1508	0.0966	0.0861	0.1556	0.1577
$\widehat{\beta}_{WP}$	-0.0153	-0.0160	-0.0151	-0.0158	-0.0160	-0.0152	-0.0155
$\widetilde{\beta}_{WP}$	-0.0154	-0.0160	-0.0152	-0.0159	-0.0160	-0.0153	-0.0156
S	0.0009	0.0007	0.0009	0.0007	0.0007	0.0009	0.0010
T	0.0009	0.0007	0.0009	0.0007	0.0007	0.0009	0.0010
$\widehat{B}_{W,10}$	1815.6624	1860.6075	1733.3769	1897.8381	1889.3409	1813.5826	1779.0288
S	239.6311	178.3196	220.4087	184.8332	170.5141	236.5810	229.4992
T	240.1318	178.4758	219.6320	182.9625	170.4034	239.4139	230.6695
Pr(Fit G_{WP})	0.0926	0.0763	0.0662	0.1017	0.0868	0.0857	0.0765
$\widehat{\tau}$	3.0573	3.0589	3.0571	3.0641	3.0642	3.0607	3.0602
S	0.2945	0.2933	0.2964	0.2965	0.2975	0.2967	0.2947
T	0.2849	0.2849	0.2849	0.2849	0.2849	0.2849	0.2849
$\widehat{\alpha}_G$	0.2990	0.2923	0.2984	0.2894	0.2834	0.3036	0.3074
S	0.3157	0.2767	0.3268	0.2633	0.2755	0.3134	0.3450
T	0.3209	0.2824	0.3303	0.2639	0.2783	0.3148	0.3491
$\widehat{\beta}_G$	2340.2838	2343.4594	2341.1489	2343.6901	2346.4331	2337.8969	2336.9745
S	129.4203	98.5632	131.8498	93.9749	94.5504	129.5721	138.7792
T	131.9687	101.6009	134.0036	95.2966	96.4683	129.8296	141.0139
$\widehat{B}_{G,10}$	2137.9506	2137.9188	2141.2763	2139.1750	2143.3175	2135.4239	2136.9128
S	273.0520	193.3161	268.3202	195.9158	183.0201	270.5914	268.8167
T	273.5473	193.9655	267.3435	194.5971	181.3961	273.2364	270.7373

Table 7.73: Fitting G_{WP} to G_{GA} for $k = 4$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

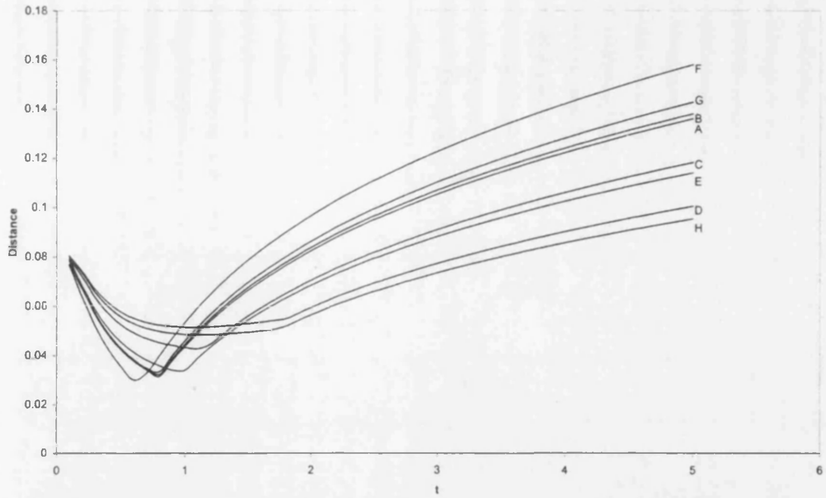


Figure 7.21: Plots of the maximum absolute distance between cdfs for the true Gamma Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios and three stress levels. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

stress levels. The results are shown in Figure 7.21. The plots for the varying ratios take roughly the same shape, although some are more extreme than others. For small values of τ, the maximum absolute distance remains quite constant across all ratios, and all begin at around 0.08. As τ increases, there is a decrease in the maximum absolute distance and all plots obtain their minimum absolute distance at around $0.5 < \tau < 1.5$. This coincides with a Negative Exponential distribution, which is of course a special case of both Gamma and Weibull distributions. As τ increases beyond one, there is a gradual increase in the maximum absolute distance for all ratios. The plots also show that varying the allocation of observations at each stress level does have an impact on the distances. For example, the ratio (1 : 1 : 3) has the minimum absolute distance across all other ratios which is attained at $\tau \simeq 0.6$. However, as we allow τ to increase, we see this distance increase much more rapidly than any others, and for large values of τ, this ratio actually yields the largest distances. The best case scenario with the most gradual increase seems to be when we have most of the observations allocated to the middle and lower stress levels.

Four stress levels We carry out a similar investigation for four stress levels, the results of which are shown in Figure 7.22. The plots show a similar picture to when we had three stress levels. All have roughly the same shape with a minimum absolute distance for $0.5 < \tau < 1.5$. This distance then increases with τ. The ratio (1 : 1 : 1 : 4) yields the largest and smallest distances depending on the value of τ. The best case scenario again seems to be when we have larger proportions of observations in the middle and lower stress levels.

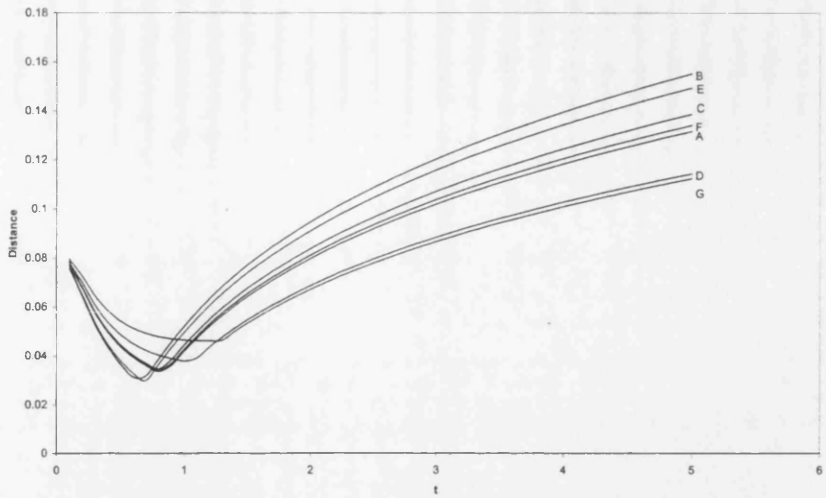


Figure 7.22: Plots of the maximum absolute distance between cdfs for the true Gamma Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios and four stress levels. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).

7.3.4 Fitting G_{WA} to G_{GP}

We investigate our final scenario for the Gamma distribution, and look at the effects of fitting the Weibull Arrhenius distribution to data with an underlying Gamma Log-linear model. We present the results below.

Simulation Studies

The results from running simulations for $\tau = 3$ are shown in Tables 7.74 and 7.75 for $k = 2$, Tables 7.76 and 7.77 for $k = 3$, and Tables 7.78 and 7.79 for $k = 4$. We compare sample values of B_{10} from both distributions with a true value of 1208.5614. Generally, across all values of τ and all numbers of stress levels, there is good agreement between sample and theoretical results for both distribution functions. The MLEs from the Gamma Log-linear distribution tend to their true values as the sample size increases, and the MLEs from the mis-specified Weibull Arrhenius model tend to their corresponding entropy values. We observe larger standard errors of the MLEs from both distributions, when we have more observations at the higher stress levels. When we examine $\widehat{B}_{10}$, the agreement between true and mis-specified quantiles is very good when $\tau = 1$ and for $k = 2$. This is because the mis-specified model reduces to the true distribution. For $k = 3$ and 4, the quantile is slightly over-estimated. This is also true when $\tau = 0.5$. For the remaining values of τ, we always under-estimate the time to which 10% of the population fail when we fit the incorrect distribution function.

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_A$	1.8856	1.8571	1.8418	1.8380
$\widetilde{B}_A$	1.8328	1.8328	1.8328	1.8328
S	0.1469	0.1037	0.0647	0.0449
T	0.1429	0.1011	0.0639	0.0452
$\widehat{\alpha}_{WA}$	-1.2453	-1.2411	-1.2426	-1.2433
$\widetilde{\alpha}_{WA}$	-1.2431	-1.2431	-1.2431	-1.2431
S	0.3167	0.2280	0.1458	0.1021
T	0.3247	0.2296	0.1452	0.1027
$\widehat{\beta}_{WA}$	3057.5647	3056.9597	3057.7876	3058.0730
$\widetilde{\beta}_{WA}$	3058.1277	3058.1277	3058.1277	3058.1277
S	119.6892	86.2061	55.1358	38.5120
T	122.7130	86.7712	54.8789	38.8053
$\widehat{B}_{W,10}$	1122.6514	1104.3492	1094.2600	1091.6789
S	140.9600	99.7191	63.1503	43.9326
T	140.1841	99.1251	62.6922	44.3301
Pr (Fit G_{WA})	0.2404	0.1325	0.0333	0.0039
$\widehat{\tau}$	3.1274	3.0597	3.0216	3.0124
S	0.4296	0.2964	0.1830	0.1273
T	0.4029	0.2849	0.1802	0.1274
$\widehat{\alpha}_G$	7.9636	7.9828	7.9940	7.9963
S	0.1759	0.1252	0.0790	0.0552
T	0.1750	0.1238	0.0783	0.0553
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0008	0.0005	0.0003	0.0002
T	0.0008	0.0005	0.0003	0.0002
$\widehat{B}_{G,10}$	1232.6946	1220.0448	1212.7921	1210.9356
S	147.1780	104.6885	65.8840	45.8542
T	146.6743	103.7144	65.5947	46.3825

Table 7.74: Fitting G_{WA} to G_{GP} for $k = 2$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_A$	1.8550	1.8577	1.8569	1.8575
$\widetilde{B}_A$	1.8328	1.8328	1.8328	1.8328
S	0.1028	0.1024	0.1034	0.1023
T	0.1011	0.1011	0.1011	0.1011
$\widehat{\alpha}_{WA}$	-1.2200	-1.2863	-1.2326	-1.2555
$\widetilde{\alpha}_{WA}$	-1.2431	-1.2431	-1.2431	-1.2431
S	0.2941	0.3865	0.2342	0.2844
T	0.2962	0.3920	0.2398	0.2884
$\widehat{\beta}_{WA}$	3047.0152	3072.2359	3052.7294	3061.9110
$\widetilde{\beta}_{WA}$	3058.1277	3058.1277	3058.1277	3058.1277
S	130.0473	129.4292	98.0891	98.8212
T	131.1857	131.1857	100.1948	100.1948
$\widehat{B}_{W,10}$	1098.1200	1106.2703	1101.2083	1105.0737
S	150.7366	89.5220	119.3913	91.1490
T	150.6503	89.3715	118.8095	91.6287
Pr (Fit G_{WA})	0.1338	0.2672	0.1313	0.1337
$\widehat{\tau}$	3.0536	3.0607	3.0586	3.3060
S	0.2927	0.2947	0.2948	0.2930
T	0.2849	0.2849	0.2849	0.2849
$\widehat{\alpha}_G$	7.9789	7.9870	7.9795	7.9840
S	0.1822	0.1187	0.1442	0.1174
T	0.1815	0.1178	0.1453	0.1171
$\widehat{\beta}_G$	-0.0200	-0.0201	-0.0200	-0.0200
S	0.0008	0.0008	0.0006	0.0006
T	0.0008	0.0008	0.0006	0.0006
$\widehat{B}_{G,10}$	1219.0383	1221.3014	1218.2901	1220.0069
S	159.8087	94.0139	125.2039	95.4184
T	159.2573	93.1132	125.0016	95.5707

Table 7.75: Fitting G_{WA} to G_{GP} for $k = 2$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	300,300	500,500
n_3	25	50	100	300	500
$\widehat{B}_A$	1.8605	1.8244	1.8082	1.7991	1.7957
$\widetilde{B}_A$	1.7931	1.7931	1.7931	1.7931	1.7931
S	0.1687	0.1183	0.0836	0.0476	0.0370
T	0.1635	0.1156	0.0817	0.0472	0.0366
$\widehat{\alpha}_{WA}$	-0.9348	-0.9321	-0.9384	-0.9385	-0.9403
$\widetilde{\alpha}_{WA}$	-0.9395	-0.9395	-0.9395	-0.9395	-0.9395
S	0.4131	0.3007	0.2165	0.1225	0.0964
T	0.4284	0.3029	0.2142	0.1237	0.0958
$\widehat{\beta}_{WA}$	2967.2556	2967.4161	2970.1071	2970.5786	2971.2855
$\widetilde{\beta}_{WA}$	2971.0736	2971.0736	2971.0736	2971.0736	2971.0736
S	161.8115	117.6425	84.6239	47.6824	37.6719
T	167.3822	118.3571	83.6911	48.3191	37.4278
$\widehat{B}_{W,10}$	1141.8291	1117.0242	1106.2964	1100.0985	1097.8885
S	181.3771	129.3399	92.7705	52.4590	41.0959
T	181.7621	128.5252	90.8810	52.4702	40.6432
Pr (Fit G_{WA})	0.1636	0.0647	0.0133	0.0001	0
$\widehat{\tau}$	3.1710	3.0809	3.0383	3.0154	3.0066
S	0.5053	0.3463	0.2402	0.1360	0.1048
T	0.4652	0.3290	0.2326	0.1343	0.1040
$\widehat{\alpha}_G$	7.9512	7.9768	7.9894	7.9955	7.9985
S	0.2225	0.1584	0.1123	0.0630	0.0497
T	0.2209	0.1562	0.1105	0.0638	0.0494
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0010	0.0008	0.0005	0.0003	0.0002
T	0.0011	0.0008	0.0005	0.0003	0.0002
$\widehat{B}_{G,10}$	1242.3270	1224.9177	1216.4244	1211.8857	1210.0788
S	183.6267	130.9203	93.2557	53.0352	41.3926
T	183.8071	129.9712	91.9035	53.0605	41.1005

Table 7.76: Fitting G_{WA} to G_{GP} for $k = 3$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	50,100 150	25,200 75	25,25 250	200,50 50
$\widehat{B}_A$	1.8024	1.8072	1.8316	1.8274
$\widetilde{B}_A$	1.7889	1.7944	1.8160	1.8119
S	0.0826	0.0823	0.0838	0.0830
T	0.0819	0.0811	0.0828	0.0822
$\widehat{\alpha}_{WA}$	-1.1578	-1.0683	-1.3051	-0.8648
$\widetilde{\alpha}_{WA}$	-1.1645	-1.0917	-1.3307	-0.8574
S	0.2379	0.3366	0.2831	0.2435
T	0.2379	0.3351	0.2841	0.2447
$\widehat{\beta}_{WA}$	3057.0607	3052.6526	3095.2113	2937.9808
$\widetilde{\beta}_{WA}$	3060.5612	3063.0030	3107.4381	2935.7989
S	99.6904	141.9072	127.0517	85.7120
T	99.7863	141.1880	127.3972	86.1560
$\widehat{B}_{W,10}$	1159.6545	1258.7120	1152.7896	1091.3791
S	114.9676	155.6147	148.1501	77.8165
T	114.2950	153.9334	148.5343	77.0955
Pr (Fit G_{WA})	0.0119	0.0117	0.0433	0.0324
$\widehat{\tau}$	3.0353	3.0361	3.0374	3.0368
S	0.2391	0.2397	0.2384	0.2364
T	0.2326	0.2326	0.2326	0.2326
$\widehat{\alpha}_G$	7.9885	7.9848	7.9804	7.9905
S	0.1296	0.1611	0.1678	0.0980
T	0.1300	0.1595	0.1668	0.0983
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0006	0.0009	0.0008	0.0005
T	0.0006	0.0009	0.0008	0.0005
$\widehat{B}_{G,10}$	1215.6134	1215.5636	1213.3868	1215.5996
S	111.3115	134.6617	145.9992	79.5347
T	110.6244	133.3001	146.7230	79.4590

Table 7.77: Fitting G_{WA} to G_{GP} for $k = 3$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_A$	1.8507	1.8247	1.8150	1.8090	1.8057
$\widetilde{B}_A$	1.8027	1.8027	1.8027	1.8027	1.8027
S	0.1450	0.1009	0.0716	0.0506	0.0315
T	0.1421	0.1005	0.0710	0.0502	0.0318
$\widehat{\alpha}_{WA}$	-0.9267	-0.9288	-0.9351	-0.9379	-0.9386
$\widetilde{\alpha}_{WA}$	-0.9408	-0.9408	-0.9408	-0.9408	-0.9408
S	0.3764	0.2754	0.1918	0.1378	0.0868
T	0.3861	0.2730	0.1930	0.1365	0.0863
$\widehat{\beta}_{WA}$	2965.0472	2966.1955	2969.2897	2970.5019	2971.1264
$\widetilde{\beta}_{WA}$	2971.9636	2971.9636	2971.9636	2971.9636	2971.9636
S	151.6310	111.3052	77.4824	55.6639	35.1235
T	156.0255	110.3267	78.0128	55.1633	34.8884
$\widehat{B}_{W,10}$	1137.1416	1117.4061	1112.5512	1108.4294	1106.8631
S	170.0578	120.8897	86.0458	60.4583	37.7752
T	170.5652	120.6078	85.2826	60.3039	38.1395
Pr (Fit G_{WA})	0.1399	0.0515	0.0105	0.0005	0
$\widehat{\tau}$	3.1205	3.0566	3.0308	3.0148	3.0068
S	0.4270	0.2932	0.2042	0.1446	0.0896
T	0.4029	0.2849	0.2015	0.1425	0.0901
$\widehat{\alpha}_G$	7.9646	7.9814	7.9897	7.9949	7.9978
S	0.2050	0.1472	0.1021	0.0739	0.0466
T	0.2060	0.1457	0.1030	0.0728	0.0461
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0010	0.0007	0.0005	0.0004	0.0002
T	0.0010	0.0007	0.0005	0.0004	0.0002
$\widehat{B}_{G,10}$	1233.4919	1218.7350	1214.4187	1211.1343	1210.1311
S	171.9631	123.3116	87.0517	61.2300	38.3066
T	172.7997	122.1879	86.3999	61.0939	38.6392

Table 7.78: Fitting G_{WA} to G_{GP} for $k = 4$, $\tau = 3$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 75,75	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_A$	1.8312	1.8310	1.8128	1.8420	1.8339	1.8301	1.8287
$\widetilde{B}_A$	1.8092	1.8069	1.7921	1.8171	1.8106	1.8075	1.8061
S	0.1041	0.1021	0.1019	0.1024	0.1020	0.1036	0.1012
T	0.1008	0.1001	0.1000	0.1008	0.1004	0.1009	0.1003
$\widehat{\alpha}_{WA}$	-1.0865	-0.6900	-1.1234	-0.8736	-0.7562	-1.1580	-0.8845
$\widetilde{\alpha}_{WA}$	-1.1130	-0.6932	-1.1413	-0.8736	-0.7495	-1.1817	-0.9041
S	0.3203	0.2773	0.3268	0.2551	0.2708	0.3100	0.3421
T	0.3191	0.2792	0.3287	0.2571	0.2752	0.3124	0.3497
$\widehat{\beta}_{WA}$	3022.1488	2884.6928	3052.6703	2941.3158	2904.5040	3047.1441	2962.7361
$\widetilde{\beta}_{WA}$	3034.4162	2886.7064	3060.9169	2942.3389	2902.3087	3058.1276	2971.9491
S	137.8304	105.8716	138.4523	98.7267	99.9054	134.3569	143.6176
T	137.4530	106.6049	139.2987	99.5405	101.4471	135.2137	146.8229
$\widehat{B}_{W,10}$	1143.4503	1105.8853	1195.5727	1104.8289	1101.9375	1149.2855	1162.0406
S	156.3906	107.1856	158.6687	107.5005	99.2304	156.7745	153.6715
T	153.9960	107.2505	156.9930	106.8190	100.0072	154.2533	154.7114
Pr(Fit G_{WA})	0.0655	0.0542	0.0123	0.0830	0.0641	0.0628	0.0560
$\widehat{\tau}$	3.0562	3.0592	3.0534	3.0603	3.0590	3.0560	3.0584
S	0.2987	0.2966	0.2946	0.2964	0.2948	0.2955	0.2943
T	0.2849	0.2849	0.2849	0.2849	0.2849	0.2849	0.2849
$\widehat{\alpha}_G$	7.9777	7.9823	7.9815	7.9817	7.9859	7.9784	7.9802
S	0.1791	0.1353	0.1735	0.1333	0.1283	0.1770	0.1777
T	0.1782	0.1342	0.1734	0.1331	0.1263	0.1772	0.1790
$\widehat{\beta}_G$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0009	0.0007	0.0009	0.0006	0.0007	0.0008	0.0009
T	0.0009	0.0007	0.0009	0.0006	0.0007	0.0008	0.0009
$\widehat{B}_{G,10}$	1218.6557	1219.4611	1220.0385	1219.4715	1221.9183	1219.4439	1221.5690
S	155.6946	110.0319	149.1390	111.5769	102.6705	155.0127	150.4987
T	153.2000	109.8412	147.3476	110.5319	103.0217	152.9695	150.4539

Table 7.79: Fitting G_{WA} to G_{GP} for $k = 4$, $\tau = 3$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

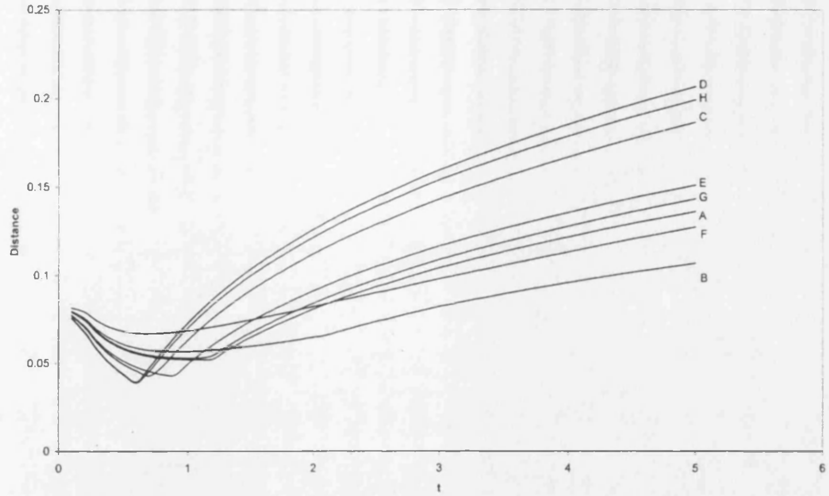


Figure 7.23: Plots of the maximum absolute distance between cdfs for the true Gamma Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios and three stress levels. Here, A=(1 : 1 : 1), B=(1 : 2 : 3), C=(3 : 2 : 1), D=(2 : 3 : 1), E=(1 : 3 : 2), F=(1 : 1 : 3), G=(3 : 1 : 1), H=(1 : 3 : 1).

The effects of mis-specification

We examine the effects of mis-specifying the Weibull Arrhenius distribution, when we assume that the underlying model is Gamma Log-linear. We have already done this using simulations, so extend the analysis to include a theoretical approach. We do this in our usual way, and examine maximum absolute distances between true and mis-specified cdfs, when entropy values are used in the distribution function for the mis-specified model. The results are shown below for $k = 3$ and 4; we do not include details for two stress levels, since this results in the same scale-stress relationship, and so is covered in the above sections.

Three stress levels We plot the maximum absolute distance between true and mis-specified distribution functions for $k = 3$, and for varying ratios. The results are shown in Figure 7.23. The plots take roughly the same shape as those produced when we mis-specified the Weibull Log-linear distribution and fitted this to data with an underlying Gamma Arrhenius model. They all begin at around 0.08 and gradually decrease to attain their minimum distances at around $0.5 < \tau < 1.5$. However, the way in which we arrange the observations across the stress levels is now reversed. Although we observe the minimum distances for the ratios (2 : 3 : 1) and (1 : 3 : 1), as τ increases above 1, these become the worst cases, and the absolute maximum distance increases much more rapidly, and to a larger degree than any other ratio. With the previous case of mis-specification we observed the exact opposite to this. The best cases now seem to be when we have more observations allocated to the middle and higher stress levels, and very little at the lower.

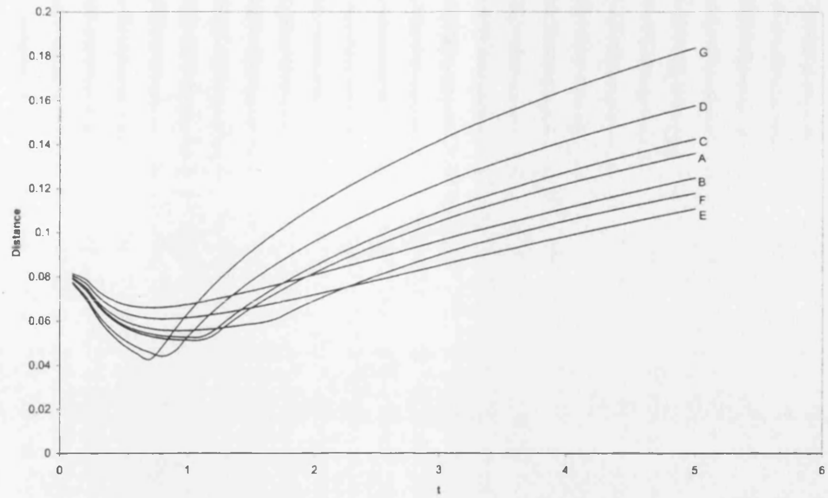


Figure 7.24: Plots of the maximum absolute distance between cdfs for the true Gamma Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios and four stress levels. Here, A=(1 : 1 : 1 : 1), B=(1 : 1 : 1 : 4), C=(4 : 1 : 1 : 1), D=(1 : 2 : 2 : 1), E=(2 : 1 : 1 : 2), F=(1 : 2 : 3 : 4), G=(4 : 3 : 2 : 1).

We note that a similar outcome was witnessed when the underlying distribution was Burr. If we fitted the Weibull Log-linear to Burr Arrhenius, the smallest absolute maximum distances occurred when we had most of the observations at the lower stress levels. The situation was reversed when we took the Weibull Arrhenius as the mis-specified model and fitted this to data with an underlying Burr Log-linear distribution.

Four stress levels We consider a similar investigation for four stress levels, the results of which are shown in Figure 7.24. Just as in three stress levels, we see the best fit between Weibull and Gamma when we have most of the observations at the middle and higher stress levels. The poorest fit, in terms of having the largest maximum absolute distance, occurs for the ratio (4 : 3 : 2 : 1), although this ratio does also yield the smallest distance for $\tau < 1$.

7.3.5 Fitting G_{WP} to G_{LNA}

We consider our final distribution function - the Lognormal distribution - and examine the effects of fitting the Weibull Log-linear distribution to data with an underlying Lognormal Arrhenius model.

Simulation Studies

We run simulations for $\sigma = 0.5$, 1 and 1.5, but only include results for $\sigma = 1$. These are summarised in Tables 7.80 and 7.81 for $k = 2$, Tables 7.82 and 7.83 for $k = 3$, and Tables

7.84 and 7.85 for $k = 4$. Note that we compare our sample values of B_{10} with a true value of $B_{LN,10} = 534.3787$. We note a number of important points. Firstly, there is excellent agreement between sample and theoretical standard errors for the true distribution, even when the sample size is relatively small; the sample means also tend to their true values very quickly. The theoretical standard error for $\widehat{\sigma}$ remains constant when we vary the loadings at each stress levels. The remaining standard errors for the Lognormal distribution increase if we have more observations at higher stress levels. Entropy values and average MLEs from the Weibull distribution are generally close to one another, especially for larger sample sizes. The agreement between sample and theoretical standard errors of the MLEs from the Weibull distribution and $\widehat{B}_{W,10}$ is generally quite good, and improves for larger sample sizes. When we compare B_{10} for both distributions, across all values of σ, $\widehat{B}_{W,10}$ and $\widehat{B}_{LN,10}$ are very different, and this quantile is always under-estimated when the wrong distribution function is fitted. The probabilities associated with fitting the Weibull distribution are generally very small, and tend to zero rapidly as n increases.

The effects of mis-specification

We examine the effects of mis-specifying the Weibull distribution in the usual way. Thus, we compare the true cdf from the Lognormal Arrhenius distribution with the mis-specified Weibull Log-linear model, when entropy values are used for the parameters in this distribution. We do not consider results for two stress levels, for reasons discussed in previous sections, but outline the results for three and four levels below.

Three stress levels We plot the maximum absolute distance between true and mis-specified distribution functions, shown by Figure 7.25, for varying ratios. We allow σ to vary between 0.1 and 5. When $\sigma > 1$, the plots are very similar across the eight ratios, and all decrease quite rapidly to around 0.06. As a result, we have been unable to distinguish between them on the plot. We note, however, that small values of σ, especially those less than 1, result in quite substantial differences in the maximum absolute distance between varying ratios. The worst case is observed for $(1:1:3)$ where the maximum distance begins at 0.5107 (this corresponds to $\sigma = 0.1$). The best cases occur when we have larger proportions of observations at the lower stresses, and for $(1:3:1)$ this distances decreases to 0.3593 when $\sigma = 0.1$. On the whole, across all stress levels, we see a better fit between Weibull Log-linear and Lognormal Arrhenius as σ is increased above 1.

Four stress levels We construct similar plots for $k = 4$. These are shown in Figure 7.26. We observe similar results to when $k = 3$, and an increase in σ results in the maximum absolute distance decreasing very quickly to around 0.06; this occurs across all ratios. We observe the largest distances when we have most of the observations at the higher stress levels.

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_P$	1.0433	1.0227	1.0099	1.0053
$\widetilde{B}_P$	1	1	1	1
S	0.0986	0.0701	0.0457	0.0322
T	0.1053	0.0745	0.0471	0.0333
$\widehat{\alpha}_{WP}$	8.8228	8.8289	8.8288	8.8300
$\widetilde{\alpha}_{WP}$	8.8301	8.8301	8.8301	8.8301
S	0.2345	0.1687	0.1080	0.0762
T	0.2431	0.1719	0.1087	0.0769
$\widehat{\beta}_{WP}$	-0.0153	-0.0154	-0.0153	-0.0154
$\widetilde{\beta}_{WP}$	-0.0153	-0.0153	-0.0153	-0.0153
S	0.0017	0.0012	0.0008	0.0005
T	0.0017	0.0012	0.0008	0.0006
$\widehat{B}_{W,10}$	367.6454	352.3847	342.0978	338.6204
S	86.2282	60.2485	39.0772	27.5260
T	89.1027	63.0051	39.8479	28.1768
Pr (Fit G_{WP})	0.0349	0.0037	0	0
$\widehat{\sigma}$	0.9879	0.9941	0.9976	0.9990
S	0.0709	0.0500	0.0319	0.0223
T	0.0707	0.0500	0.0316	0.0224
$\widehat{\alpha}_{LN}$	0.3021	0.2978	0.3008	0.2993
S	0.5365	0.3775	0.2448	0.1718
T	0.5402	0.3820	0.2416	0.1708
$\widehat{\beta}_{LN}$	2345.7852	2347.8332	2346.6869	2347.2015
S	202.7045	142.2640	92.2331	64.7739
T	203.8752	144.1615	91.1758	64.4710
$\widehat{B}_{LN,10}$	549.6744	542.3806	537.4622	535.8211
S	93.6811	64.1601	40.6534	28.6193
T	89.7563	63.4673	40.1402	28.3834

Table 7.80: Fitting G_{WP} to G_{LNA} for $k = 2$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_P$	1.0235	1.0224	1.0241	1.0230
$\widetilde{B}_P$	1	1	1	1
S	0.0697	0.0703	0.0700	0.0694
T	0.0745	0.0745	0.0745	0.0745
$\widehat{\alpha}_{WP}$	8.7981	8.8403	8.8186	8.8349
$\widetilde{\alpha}_{WP}$	8.8301	8.8301	8.8301	8.8301
S	0.3295	0.1433	0.2386	0.1428
T	0.3469	0.1489	0.2438	0.1459
$\widehat{\beta}_{WP}$	-0.0152	-0.0155	-0.0153	-0.0154
$\widetilde{\beta}_{WP}$	-0.0153	-0.0153	-0.0153	-0.0153
S	0.0018	0.0018	0.0014	0.0014
T	0.0019	0.0019	0.0014	0.0014
$\widehat{B}_{W,10}$	353.5087	351.7807	354.0608	352.5013
S	98.9840	52.5452	76.4279	53.6767
T	98.6584	56.0920	76.7522	57.6998
Pr (Fit G_{WP})	0.0031	0.0028	0.0049	0.0035
$\widehat{\sigma}$	0.9930	0.9939	0.9928	0.9936
S	0.0497	0.0498	0.0499	0.0499
T	0.0500	0.0500	0.0500	0.0500
$\widehat{\alpha}_{LN}$	0.3000	0.3000	0.2966	0.2956
S	0.4938	0.6505	0.3993	0.4810
T	0.4925	0.6516	0.3990	0.4795
$\widehat{\beta}_{LN}$	2347.1270	2346.9713	2348.9406	2348.6658
S	218.5020	217.3970	166.5251	166.7964
T	217.9517	217.9517	166.4634	166.4634
$\widehat{B}_{LN,10}$	551.3096	541.1779	547.2480	542.0442
S	115.7162	53.7037	85.1775	55.8286
T	112.2271	52.9553	82.9681	55.4638

Table 7.81: Fitting G_{WP} to G_{LNA} for $k = 2$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	25,25 25	50,50 50	100,100 100	300,300 300	500,500 500
$\widehat{B}_P$	1.0528	1.0257	1.0124	1.0022	1.0001
$\widetilde{B}_P$	0.9966	0.9966	0.9966	0.9966	0.9966
S	0.1130	0.0795	0.0576	0.0338	0.0264
T	0.1211	0.0857	0.0606	0.0350	0.0271
$\widehat{\alpha}_{WP}$	8.7809	8.7956	8.8005	8.8004	8.8024
$\widetilde{\alpha}_{WP}$	8.8032	8.8032	8.8032	8.8032	8.8032
S	0.3253	0.2354	0.1691	0.0982	0.0764
T	0.3457	0.2444	0.1729	0.0998	0.0773
$\widehat{\beta}_{WP}$	-0.0155	-0.0156	-0.0156	-0.0156	-0.0156
$\widetilde{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
S	0.0023	0.0016	0.0012	0.0007	0.0005
T	0.0024	0.0017	0.0012	0.0007	0.0005
$\widehat{B}_{W,10}$	360.3237	341.5497	331.5227	323.2544	321.8662
S	105.0198	72.8813	51.7531	30.4054	23.4683
T	107.3744	75.9252	53.6872	30.9963	24.0097
Pr (Fit G_{WP})	0.0530	0.0086	0.0003	0	0
$\widehat{\sigma}$	0.9822	0.9917	0.9959	0.9989	0.9993
S	0.0813	0.0573	0.0413	0.0235	0.0183
T	0.0816	0.0577	0.0408	0.0236	0.0183
$\widehat{\alpha}_{LN}$	0.3051	0.3007	0.3018	0.3022	0.2998
S	0.7009	0.5012	0.3543	0.2044	0.1589
T	0.7095	0.5017	0.3548	0.2048	0.1587
$\widehat{\beta}_{LN}$	2344.4702	2346.8793	2346.3912	2346.0714	2347.1509
S	273.8754	195.8352	138.3057	79.8327	61.9265
T	277.3978	196.1499	138.6989	80.0779	62.0280
$\widehat{B}_{LN,10}$	558.5333	546.9109	540.4427	535.9439	535.6677
S	123.2248	85.6054	59.6890	34.6738	26.2536
T	118.7110	83.9414	59.3555	34.2689	26.5446

Table 7.82: Fitting G_{WP} to G_{LNA} for $k = 3$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	50,100	25,200	25,25	200,50
n_3	150	75	250	50
$\widehat{B}_P$	1.0125	1.0130	1.0143	1.0137
$\widetilde{B}_P$	0.9965	0.9963	0.9989	0.9982
S	0.0581	0.0578	0.0574	0.0575
T	0.0606	0.0606	0.0607	0.0607
$\widehat{\alpha}_{WP}$	8.7348	8.6459	8.7402	8.8432
$\widetilde{\alpha}_{WP}$	8.7459	8.6620	8.7691	8.8415
S	0.2273	0.3003	0.3226	0.1286
T	0.2325	0.3130	0.3324	0.1299
$\widehat{\beta}_{WP}$	-0.0151	-0.0149	-0.0150	-0.0158
$\widetilde{\beta}_{WP}$	-0.0152	-0.0150	-0.0151	-0.0158
S	0.0014	0.0019	0.0017	0.0012
T	0.0014	0.0020	0.0018	0.0012
$\widehat{B}_{W,10}$	319.7978	298.3271	330.8914	341.5052
S	62.4170	70.9259	89.2606	45.4301
T	62.6170	70.2799	86.6041	47.4616
Pr (Fit G_{WP})	0.0007	0.0004	0.0004	0.0004
$\widehat{\sigma}$	0.9954	0.9957	0.9961	0.9958
S	0.0410	0.0407	0.0406	0.0407
T	0.0408	0.0408	0.0408	0.0408
$\widehat{\alpha}_{LN}$	0.3037	0.3014	0.2971	0.2998
S	0.3971	0.5597	0.4745	0.4022
T	0.3971	0.5589	0.4711	0.3998
$\widehat{\beta}_{LN}$	2345.7372	2346.6101	2347.9753	2347.0424
S	166.6359	235.4138	212.1588	141.7603
T	166.4603	235.4061	210.9961	140.8268
$\widehat{B}_{LN,10}$	543.0352	547.1208	548.2230	539.3323
S	79.0457	103.9662	111.6127	47.7647
T	77.4560	101.1454	107.4403	46.8482

Table 7.83: Fitting G_{WP} to G_{LNA} for $k = 3$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_P$	1.0413	1.0200	1.0097	1.0033	0.9999
$\widetilde{B}_P$	0.9974	0.9974	0.9974	0.9974	0.9974
S	0.0981	0.0696	0.0504	0.0365	0.0233
T	0.1050	0.0743	0.0525	0.0371	0.0235
$\widehat{\alpha}_{WP}$	8.7892	8.7957	8.7985	8.8028	8.8047
$\widetilde{\alpha}_{WP}$	8.8059	8.8059	8.8059	8.8059	8.8059
S	0.3289	0.2406	0.1703	0.1204	0.0768
T	0.3460	0.2447	0.1730	0.1223	0.0774
$\widehat{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
$\widetilde{\beta}_{WP}$	-0.0156	-0.0156	-0.0156	-0.0156	-0.0156
S	0.0021	0.0016	0.0011	0.0008	0.0005
T	0.0023	0.0016	0.0011	0.0008	0.0005
$\widehat{B}_{W,10}$	354.9154	337.7138	328.9694	324.3081	321.7643
S	101.2519	71.1927	49.7405	35.2995	22.6802
T	101.4948	71.7677	50.7474	35.8838	22.6949
Pr (Fit G_{WP})	0.0283	0.0027	0	0	0
$\widehat{\sigma}$	0.9879	0.9937	0.9963	0.9984	0.9994
S	0.0719	0.0500	0.0353	0.0254	0.0158
T	0.0707	0.0500	0.0354	0.0250	0.0158
$\widehat{\alpha}_{LN}$	0.3027	0.3009	0.3019	0.3038	0.2995
S	0.6180	0.4485	0.3225	0.2262	0.1442
T	0.6416	0.4536	0.3208	0.2268	0.1435
$\widehat{\beta}_{LN}$	2345.8166	2346.6624	2346.3232	2345.4771	2347.1908
S	250.4320	181.8830	130.5351	91.5649	58.1744
T	259.2445	183.3135	129.6222	91.6568	57.9688
$\widehat{B}_{LN,10}$	554.8329	545.0316	539.9604	536.5170	535.4572
S	119.7795	83.6426	58.6043	41.2788	25.9685
T	115.3647	81.5752	57.6824	40.7876	25.7963

Table 7.84: Fitting G_{WP} to G_{LNA} for $k = 4$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 75,75	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_P$	1.0200	1.0212	1.0189	1.2026	1.0202	1.0194	1.0196
$\widetilde{B}_P$	0.9981	0.9974	0.9966	0.9986	0.9979	0.9981	0.9976
S	0.0690	0.0696	0.0699	0.0695	0.0701	0.0697	0.0696
T	0.0743	0.0743	0.0742	0.0744	0.0743	0.0743	0.0743
$\widehat{\alpha}_{WP}$	8.7504	8.8305	8.7008	8.8361	8.8449	8.7454	8.7424
$\widetilde{\alpha}_{WP}$	8.7778	8.8379	8.7230	8.8419	8.8457	8.7715	8.7707
S	0.3217	0.2024	0.3100	0.1999	0.1757	0.3170	0.3232
T	0.3379	0.2082	0.3246	0.2048	0.1813	0.3352	0.3399
$\widehat{\beta}_{WP}$	-0.0152	-0.0160	-0.0150	-0.0158	-0.0160	-0.0151	-0.0154
$\widetilde{\beta}_{WP}$	-0.0154	-0.0160	-0.0152	-0.0159	-0.0160	-0.0153	-0.0156
S	0.0019	0.0015	0.0019	0.0014	0.0014	0.0018	0.0020
T	0.0020	0.0016	0.0020	0.0015	0.0015	0.0019	0.0021
$\widehat{B}_{W,10}$	333.3059	341.4297	318.2138	345.8053	344.6918	332.5655	326.1469
S	89.6783	62.9837	82.2157	63.9470	58.4201	89.2605	86.3192
T	89.4869	65.0311	81.8546	68.3838	61.4258	89.2104	86.0656
Pr(Fit G_{WP})	0.0022	0.0025	0.0025	0.0030	0.0034	0.0036	0.0037
$\widehat{\sigma}$	0.9937	0.9930	0.9946	0.9943	0.9939	0.9939	0.9943
S	0.0496	0.0498	0.0505	0.0501	0.0501	0.0499	0.0496
T	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
$\widehat{\alpha}_{LN}$	0.3100	0.3027	0.2975	0.2981	0.3007	0.3019	0.3059
S	0.5279	0.4571	0.5453	0.4230	0.4474	0.5193	0.5853
T	0.5310	0.4606	0.5480	0.4264	0.4530	0.5199	0.5819
$\widehat{\beta}_{LN}$	2343.0410	2346.0435	2348.1026	2347.3381	2346.6390	2346.4531	2344.8485
S	227.1646	174.7511	230.6164	163.4952	164.6902	224.7995	245.7112
T	228.5765	175.9779	232.1011	165.0586	167.0881	224.8714	244.2433
$\widehat{B}_{LN,10}$	548.7869	543.6662	549.2960	542.4522	542.0060	549.8931	549.0941
S	112.9433	71.7141	109.9624	71.1497	63.5010	113.1828	112.4278
T	109.3757	69.8625	106.4156	70.1961	63.1083	109.2276	108.0366

Table 7.85: Fitting G_{WP} to G_{LNA} for $k = 4$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

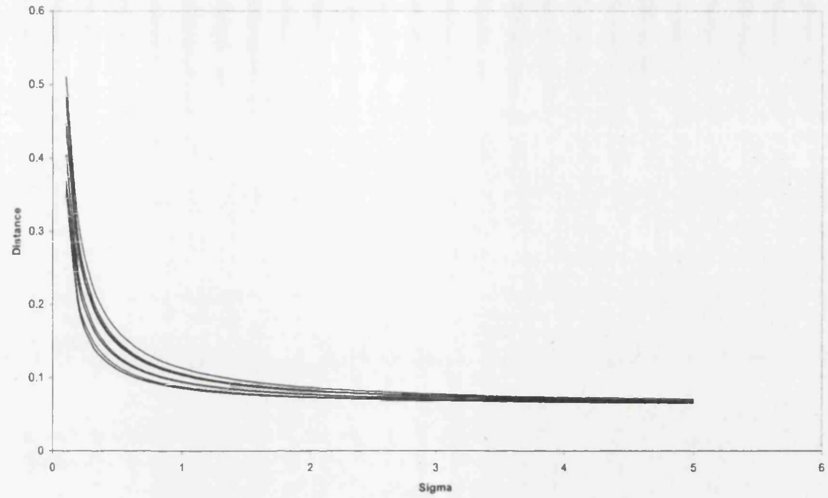


Figure 7.25: Plots of the maximum absolute distance between cdfs for the true Lognormal Arrhenius distribution and mis-specified Weibull Log-linear model for the eight different ratios and three stress levels.

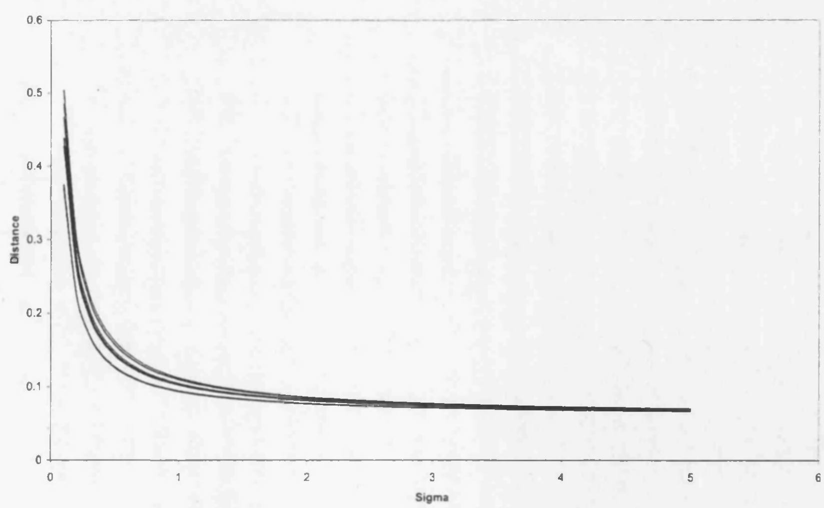


Figure 7.26: Plots of the maximum absolute distance between cdfs for the true Lognormal Arrhenius distribution and mis-specified Weibull Log-linear model for the seven different ratios and four stress levels.

7.3.6 Fitting G_{WA} to G_{LNP}

We consider our final scenario, and examine the effects of fitting the Weibull Arrhenius distribution to data with an underlying Lognormal Log-linear model. We first do this by constructing our usual sets of simulations.

Simulation Studies

We run simulations for $\sigma = 0.5$, 1 and 1.5, although only include results for $\sigma = 1$. These are shown in Tables 7.86 and 7.87 for $k = 2$, Tables 7.88 and 7.89 for $k = 3$, and Tables 7.90 and 7.91 for $k = 4$. We compare estimates for B_{10} from both distributions with a true value of 304.4323. We note the following points when mis-specifying the Weibull Arrhenius model. Firstly, we generally see good agreement between average MLEs and entropy values, which, as expected, improves as we increase the sample size. The agreement between sample and theoretical standard errors of the mis-specified MLEs differs slightly for smaller samples, but again improves as n is increased. When examining $\widehat{B}_{W,10}$ we observe considerable differences between sample values for the true and mis-specified distributions, and we always under-estimate this quantile when we fit the wrong distribution function.

The effects of mis-specification

We compare true and mis-specified theoretical cdfs, and examine the maximum absolute distances between them for varying values of σ and stress loadings. We do not do this for $k = 2$, since we reduce the scenario to just mis-specifying the distribution function, results of which are covered by the non-accelerated case. We present the results for 3 and 4 stress levels below.

Three stress levels We construct plots of the maximum absolute distance between true and mis-specified cdfs, when we allow σ to vary from 0.1 to 5, and for varying ratios. The results are shown in Figure 7.27. Again, since the plots are so similar for larger values of σ, we have been unable to distinguish between them in our usual way. Like the Burr and Gamma distributions, we see an improved fit between Weibull and Lognormal distributions if we have more observations in the higher and middle stress levels, whilst a larger proportion of observations at the lower stress levels results in the largest maximum absolute distance across all values of σ. This seems to be a common fact if the Arrhenius scale-stress relationship is used as the mis-specified model. On the whole, however, we seen an improved fit between Weibull and Lognormal as we increase σ; this is not consistent with previous plots used to determine values of σ for simulations, and may require further investigation.

Four stress levels We carry out a similar investigation for $k = 4$, the results of which are shown by Figure 7.28. Again we observe similar results to three stress levels. Having most

n_1, n_2	50,50	100,100	250,250	500,500
$\widehat{B}_A$	1.0435	1.0223	1.0094	1.0051
$\widetilde{B}_A$	1	1	1	1
S	0.0975	0.0701	0.0451	0.0323
T	0.1053	0.0745	0.0471	0.0333
$\widehat{\alpha}_{WA}$	-1.9759	-1.9616	-1.9638	-1.9643
$\widetilde{\alpha}_{WA}$	-1.9632	-1.9632	-1.9632	-1.9632
S	0.6685	0.4878	0.3081	0.2204
T	0.7040	0.4978	0.3148	0.2226
$\widehat{\beta}_{WA}$	3060.0058	3057.0430	3058.1583	3058.4712
$\widetilde{\beta}_{WA}$	3058.1277	3058.1277	3058.1277	3058.1277
S	253.6566	185.2951	116.9927	83.6310
T	267.2462	188.9716	119.5161	84.5107
$\widehat{B}_{W,10}$	209.5574	200.6027	194.8001	192.8663
S	48.8257	34.3142	21.9072	15.8009
T	50.7621	35.8942	22.7015	16.0524
Pr (Fit G_{WA})	0.0344	0.0025	0	0
$\widehat{\sigma}$	0.9878	0.9937	0.9973	0.9985
S	0.0703	0.0506	0.0316	0.0222
T	0.0707	0.0500	0.0316	0.0224
$\widehat{\alpha}_{LN}$	8.0005	8.0013	8.0005	8.0011
S	0.1925	0.1390	0.0873	0.0614
T	0.1944	0.1374	0.0869	0.0615
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0013	0.0010	0.0006	0.0004
T	0.0013	0.0009	0.0006	0.0004
$\widehat{B}_{LN,10}$	313.4327	309.5085	306.5089	305.6861
S	52.6416	36.9892	23.1519	16.2358
T	51.1336	36.1569	22.8677	16.1699

Table 7.86: Fitting G_{WA} to G_{LNP} for $k = 2$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,175	175,25	50,150	150,50
$\widehat{B}_A$	1.0216	1.0229	1.0233	1.0234
$\widetilde{B}_A$	1	1	1	1
S	0.0708	0.0708	0.0698	0.0704
T	0.0745	0.0745	0.0745	0.0745
$\widehat{\alpha}_{WA}$	-1.9029	-2.0520	-1.9446	-1.9947
$\widetilde{\alpha}_{WA}$	-1.9632	-1.9632	-1.9632	-1.9632
S	0.6163	0.8125	0.5096	0.6149
T	0.6433	0.8524	0.5202	0.6263
$\widehat{\beta}_{WA}$	3029.6838	3086.6910	3048.9493	3068.3063
$\widetilde{\beta}_{WA}$	3058.1277	3058.1277	3058.1277	3058.1277
S	273.9361	272.4366	213.5898	214.3042
T	285.6982	285.6982	218.2056	218.2056
$\widehat{B}_{W,10}$	200.1366	200.4436	201.1750	200.8748
S	57.4617	30.1992	43.5740	31.0770
T	56.2052	31.9556	43.7257	32.8717
Pr (Fit G_{WA})	0.0030	0.0046	0.0028	0.0040
$\widehat{\sigma}$	0.9941	0.9938	0.9933	0.9935
S	0.0500	0.0501	0.0501	0.0505
T	0.0500	0.0500	0.0500	0.0500
$\widehat{\alpha}_{LN}$	8.0002	7.9995	8.0014	8.0005
S	0.2677	0.1207	0.1936	0.1189
T	0.2679	0.1208	0.1905	0.1186
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0014	0.0014	0.0011	0.0011
T	0.0014	0.0014	0.0011	0.0011
$\widehat{B}_{LN,10}$	313.5527	308.0851	311.1656	308.8016
S	66.5431	30.5418	49.2174	32.4347
T	63.9351	30.1683	47.2664	31.5974

Table 7.87: Fitting G_{WA} to G_{LNP} for $k = 2$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	300,300	500,500
n_3	25	50	100	300	500
$\widehat{B}_A$	1.0480	1.0209	1.0088	0.9995	0.9971
$\widetilde{B}_A$	0.9942	0.9942	0.9942	0.9942	0.9942
S	0.1130	0.0806	0.0569	0.0336	0.0268
T	0.1209	0.0855	0.0605	0.0349	0.0270
$\widehat{\alpha}_{WA}$	-1.6487	-1.6695	-1.6780	-1.6781	-1.6756
$\widetilde{\alpha}_{WA}$	-1.6784	-1.6784	-1.6784	-1.6784	-1.6784
S	0.8711	0.6352	0.4530	0.2649	0.2053
T	0.9247	0.6539	0.4623	0.2669	0.2068
$\widehat{\beta}_{WA}$	2962.2511	2973.3737	2977.1837	2977.4288	2976.4416
$\widetilde{\beta}_{WA}$	2977.6410	2977.6410	2977.6410	2977.6410	2977.6410
S	340.6653	249.3648	177.8096	103.9772	80.6030
T	363.0372	256.7061	181.5186	104.7998	81.1776
$\widehat{B}_{W,10}$	219.2351	208.1781	202.2312	197.5301	196.1653
S	65.2999	46.5173	32.0349	18.6442	14.6537
T	66.4162	46.9634	33.2081	19.1727	14.8511
Pr (Fit G_{WA})	0.0483	0.0083	0.0003	0	0
$\widehat{\sigma}$	0.9833	0.9927	0.9960	0.9989	0.9992
S	0.0812	0.0585	0.0408	0.0235	0.0184
T	0.0816	0.0577	0.0408	0.0236	0.0183
$\widehat{\alpha}_{LN}$	7.9989	8.0034	8.0030	8.0014	8.0000
S	0.2674	0.1926	0.1349	0.0788	0.0606
T	0.2726	0.1927	0.1363	0.0787	0.0609
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0018	0.0013	0.0009	0.0005	0.0004
T	0.0019	0.0013	0.0009	0.0005	0.0004
$\widehat{B}_{LN,10}$	318.0077	311.8307	308.4703	305.7706	305.0962
S	69.5326	48.6155	33.1913	19.3234	14.8605
T	66.7617	47.2076	33.3808	19.2724	14.9284

Table 7.88: Fitting G_{WA} to G_{LNP} for $k = 3$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3	50,100 150	25,200 75	25,25 250	200,50 50
$\widehat{B}_A$	1.0092	1.0086	1.0124	1.0124
$\widetilde{B}_A$	0.9936	0.9941	0.9977	0.9970
S	0.0575	0.0575	0.0576	0.0576
T	0.0605	0.0604	0.0607	0.0606
$\widehat{\alpha}_{WA}$	-1.8707	-1.7623	-1.9827	-1.6207
$\widetilde{\alpha}_{WA}$	-1.8865	-1.8130	-2.0401	-1.6019
S	0.5004	0.7007	0.5994	0.5183
T	0.5168	0.7289	0.6158	0.5242
$\widehat{\beta}_{WA}$	3052.7985	3040.2663	3075.6126	2949.9621
$\widetilde{\beta}_{WA}$	3060.3304	3062.5355	3102.2102	2944.0648
S	210.4623	295.6562	269.2069	183.5438
T	217.3983	307.5264	276.5303	185.2584
$\widehat{B}_{W,10}$	212.3745	230.1716	208.6896	197.7144
S	42.2178	58.9291	57.2523	26.3553
T	42.6335	58.0986	55.7381	27.6190
Pr (Fit G_{WA})	0.0005	0.0002	0.0002	0.0002
$\widehat{\sigma}$	0.9960	0.9955	0.9959	0.9956
S	0.0409	0.0405	0.0405	0.0409
T	0.0408	0.0408	0.0408	0.0408
$\widehat{\alpha}_{LN}$	8.0011	7.9994	7.9999	8.0009
S	0.1806	0.2424	0.2603	0.1050
T	0.1808	0.2415	0.2557	0.1046
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0011	0.0015	0.0014	0.0010
T	0.0011	0.0015	0.0014	0.0010
$\widehat{B}_{LN,10}$	309.2218	310.9053	312.2341	307.4003
S	43.5359	55.4095	63.0973	26.8305
T	42.8487	53.7485	60.0366	26.6147

Table 7.89: Fitting G_{WA} to G_{LNP} for $k = 3$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2	25,25	50,50	100,100	200,200	500,500
n_3, n_4	25,25	50,50	100,100	200,200	500,500
$\widehat{B}_A$	1.0372	1.0182	1.0066	1.0019	0.9982
$\widetilde{B}_A$	0.9956	0.9956	0.9956	0.9956	0.9956
S	0.0981	0.0697	0.0497	0.0359	0.0231
T	0.1049	0.0742	0.0524	0.0371	0.0235
$\widehat{\alpha}_{WA}$	-1.6318	-1.6528	-1.6541	-1.6653	-1.6690
$\widetilde{\alpha}_{WA}$	-1.6717	-1.6717	-1.6717	-1.6717	-1.6717
S	0.8000	0.5780	0.4081	0.2932	0.1861
T	0.8362	0.5913	0.4181	0.2956	0.1870
$\widehat{\beta}_{WA}$	2956.2597	2966.5559	2967.9469	2972.7472	2974.4273
$\widetilde{\beta}_{WA}$	2975.6768	2975.6768	2975.6768	2975.6768	2975.6768
S	324.1256	234.3145	165.5726	118.7898	75.4877
T	339.2414	239.8799	169.6207	119.9399	75.8567
$\widehat{B}_{W,10}$	214.7652	206.3222	200.4878	198.5053	196.7719
S	63.0394	44.2816	30.7687	21.7923	13.9362
T	62.8443	44.4377	31.4222	22.2188	14.0524
Pr (Fit G_{WA})	0.0283	0.0025	0	0	0
$\widehat{\sigma}$	0.9867	0.9934	0.9970	0.9984	0.9994
S	0.0702	0.0499	0.0353	0.0249	0.0160
T	0.0707	0.0500	0.0354	0.0250	0.0158
$\widehat{\alpha}_{LN}$	7.9983	8.0009	7.9984	7.9995	8.0004
S	0.2702	0.1913	0.1348	0.0959	0.0605
T	0.2706	0.1914	0.1353	0.0957	0.0605
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0017	0.0012	0.0009	0.0006	0.0004
T	0.0017	0.0012	0.0009	0.0006	0.0004
$\widehat{B}_{LN,10}$	316.2514	310.6645	306.9925	305.8119	305.0591
S	68.0536	47.3479	32.5882	23.0638	14.5934
T	64.8339	45.8445	32.4169	22.9222	14.4973

Table 7.90: Fitting G_{WA} to G_{LNP} for $k = 4$, $\sigma = 1$ with equal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

n_1, n_2 n_3, n_4	25,25 75,75	75,75 25,25	25,75 25,75	75,25 75,25	100,50 25,25	25,25 50,100	25,75 75,25
$\widehat{B}_A$	1.0185	1.0177	1.0161	1.0208	1.0189	1.0184	1.0170
$\widetilde{B}_A$	0.9966	0.9961	0.9940	0.9977	0.9967	0.9964	0.9961
S	0.0708	0.0692	0.0697	0.0690	0.0705	0.0704	0.0698
T	0.0742	0.0741	0.0740	0.0743	0.0742	0.0742	0.0742
$\widehat{\alpha}_{WA}$	-1.7751	-1.4281	-1.8096	-1.5894	-1.4989	-1.8573	-1.5604
$\widetilde{\alpha}_{WA}$	-1.8337	-1.4323	-1.8605	-1.6018	-1.4881	-1.9002	-1.6294
S	0.6728	0.5914	0.6802	0.5360	0.5820	0.6486	0.7258
T	0.6928	0.6019	0.7140	0.5562	0.5924	0.6782	0.7594
$\widehat{\beta}_{WA}$	3007.5998	2890.6056	3036.4749	2939.1065	2911.6969	3037.0447	2943.1041
$\widetilde{\beta}_{WA}$	3034.1261	2893.4598	3059.5354	2945.1729	2908.7500	3056.7943	2973.4895
S	290.2091	227.0293	289.0230	208.7564	215.1624	281.3290	305.4818
T	299.1429	230.7964	303.2618	216.3045	219.3304	294.2601	319.5533
$\widehat{B}_{W,10}$	210.5640	202.9429	220.7533	202.0183	201.7651	212.2706	212.7894
S	59.2297	38.0232	59.7081	37.7160	34.8207	59.0917	58.6101
T	57.6292	39.0837	58.8811	38.9010	36.1383	57.7282	57.9291
Pr(Fit G_{WA})	0.0024	0.0025	0.0023	0.0031	0.0027	0.0030	0.0030
$\widehat{\sigma}$	0.9934	0.9943	0.9936	0.9939	0.9940	0.9935	0.9943
S	0.0502	0.0500	0.0500	0.0496	0.0500	0.0500	0.0495
T	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
$\widehat{\alpha}_{LN}$	7.9979	7.9994	7.9974	7.9993	8.0020	8.0049	7.9956
S	0.2639	0.1651	0.2508	0.1615	0.1441	0.2579	0.2619
T	0.2612	0.1642	0.2514	0.1616	0.1443	0.2591	0.2628
$\widehat{\beta}_{LN}$	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200	-0.0200
S	0.0015	0.0012	0.0015	0.0011	0.0011	0.0015	0.0016
T	0.0015	0.0012	0.0015	0.0011	0.0011	0.0015	0.0016
$\widehat{B}_{LN,10}$	312.9590	309.0118	312.0999	309.3090	309.1646	314.3367	311.5735
S	63.9339	39.9803	60.5766	40.8721	36.2195	63.7586	61.0967
T	61.0526	39.4511	58.2460	39.8166	35.7780	60.9425	59.7385

Table 7.91: Fitting G_{WA} to G_{LNP} for $k = 4$, $\sigma = 1$ with unequal allocations. We show the sample means and standard errors of parameters, where figures are based on at least 10000 replications.

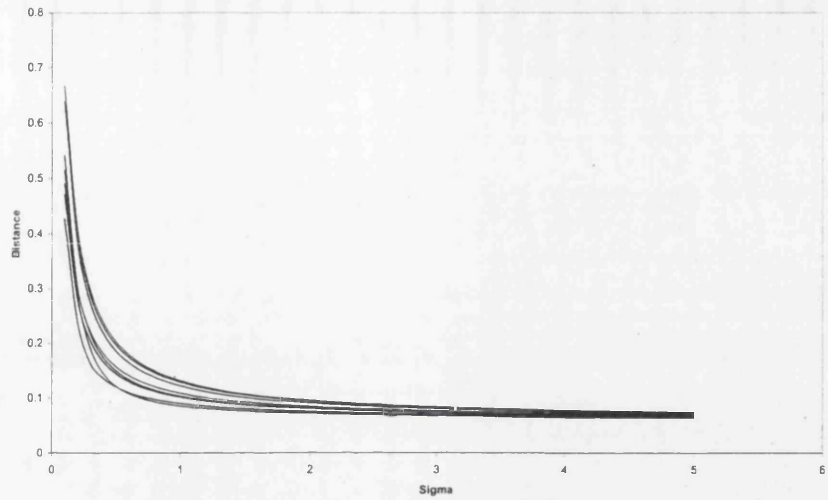


Figure 7.27: Plots of the maximum absolute distance between cdfs for the true Lognormal Log-linear distribution and mis-specified Weibull Arrhenius model for the eight different ratios and three stress levels.

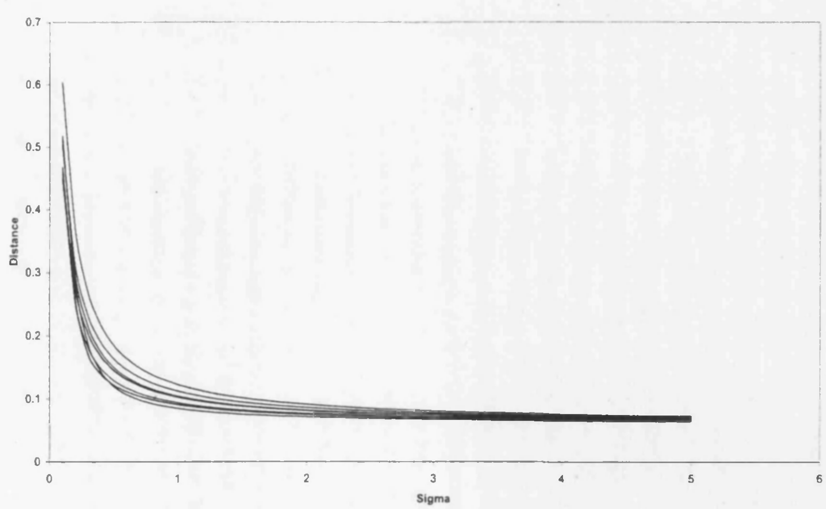


Figure 7.28: Plots of the maximum absolute distance between cdfs for the true Lognormal Log-linear distribution and mis-specified Weibull Arrhenius model for the seven different ratios and four stress levels.

of the observations at the lower stresses give the worst fit, whilst large proportions at the higher stress levels give the best, in terms of having smaller maximum absolute distances.

7.4 Summary

In this chapter, we examined some effects of mis-specifying the accelerated Weibull distribution. Since acceleration introduces a relationship between scale parameter and stress level, then we could also mis-specify this relationship. We considered three scenarios of mis-specification; mis-specifying the scale-stress relationship and keeping the distribution fixed, mis-specifying the distribution but choosing the correct scale-stress relationship, and finally mis-specifying both models. As in our work on non-accelerated data sets, we carried out the analysis for the Burr, Gamma and Lognormal distributions, and considered the Log-linear and Arrhenius scale-stress relationships. We verified our theoretical results by running simulations, and assessed the goodness of fit by examining distances between true and estimated cdfs across all stress levels.

In the next chapter, we consider a combination of both acceleration and censoring.

Chapter 8

Type II Censoring In Accelerated Life Testing

8.1 Introduction

We finally examine accelerated data sets under a type II censoring regime. It is clear that, without acceleration, the practical application of type II censoring is straightforward. Out of our total sample size n, we terminate the experiment once the r^{th} item has failed; the stopping time is then random. Thus, we have a known number of failures r, and a known number of censored items $n - r$. However, several extensions in an accelerated framework are possible. For illustration, consider an experiment with $k = 3$ stress levels. We could stop the experiment once a certain number of items have failed at the lowest stress level, that is, until r_1 have failed out of n_1, or we could wait until a certain number of items have failed at the highest stress level, that is, until r_3 have failed out of n_3. Other possibilities include stopping the experiment once a certain number of items have failed out of n, no matter at which stress level the failures occur. Our approach will involve using the first censoring regime, when we stop the experiment after r failures are observed at the lowest stress level. This approach provides the experimenter with some idea on the number of failures, since we are certain of a fixed number of failures at X_1, and (generally larger) proportions of failures at higher stress levels. However, as in type II censoring for non-accelerated data sets, we cannot be sure of the running time of the experiment; this remains a practical drawback of the method. Structuring the experiment in this way results in ordering only observations at the lowest stress level. We use $y_{(j:n_1,1)}$ to denote the ordered observed failures at X_1 for $1 \leq j \leq r$, with corresponding random variables $Y_{(j:n_1,1)}$ (for simplicity, we write $Z = Y_{(r:n_1,1)}$). All observations at higher stress levels will be censored if they exceed this failure time; thus, we need no ordering at these stress levels, but regard data as subject to a type I censoring regime where the stopping time Z is now random. We denote observed failures at higher stress levels as $y_{i,j}$ with corresponding random variable $Y_{i,j}$, for $2 \leq i \leq k$, $1 \leq j \leq M_i$, where M_i is the random variable representing the number of failures at the i^{th} stress level. Note that for $i = 1$, we can also write $M_1 = r$, a specified number.

We have not been able to find references on this censoring regime. Various authors (for example, Menzefricke, 1992) consider accelerated life tests when data has undergone a type II censoring regime, but usually assume a fixed number of failures r_i at each stress level, so that there is no connection with type I censoring. Tseng and Hsu (1994) compare type I and type II censoring regimes in accelerated life tests, but again, for type II censoring, assume a fixed number of failures at each stress level. Censoring in accelerated life tests is briefly mentioned in Meeker and Escobar (1993), but without reference to type I or type II censoring; Nelson (1990) also discusses ML for right censored data (which includes type I and type II as special cases), but does not consider type II specifically, or even describe how observations are censored at each stress level.

8.2 The Weibull distribution

We begin by considering the Weibull distribution, and, as previously, use a general scale-stress relationship which includes both the Arrhenius and Log-linear models as special cases. Thus, we have data from a Weibull distribution with shape parameter B_* and scale parameter

$$\theta_{i*} = \exp\left\{\alpha_{W*} + \beta_{W*}\rho\left(X_i\right)\right\},$$

where $\rho\left(X_i\right)$ is given by (5.3). As discussed, we terminate the experiment after the r^{th} failure at X_1; the running time of the experiment is $Z = Y_{(r:n_1,1)}$, and any observations that exceed this time are censored. We refer to Balakrishnan and Rao (1998a) to write the pdf of Z as

$$\frac{n_1!}{(n_1-r)!\,(r-1)!}\left[1-\exp\left\{-\left(\frac{z}{\theta_{1*}}\right)^{B_*}\right\}\right]^{r-1}\exp\left\{-\left(\frac{z}{\theta_{1*}}\right)^{B_*}\right\}^{n_1-r}\frac{B_* z^{B_*-1}}{\theta_{1*}^{B_*}}\exp\left\{-\left(\frac{z}{\theta_{1*}}\right)^{B_*}\right\}$$

which can also be written as

$$\frac{B_* n_1!}{\theta_{1*}^{B_*}\,(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}(-1)^{r-1-j}\binom{r-1}{j}z^{B_*-1}\exp\left\{-\left(n_1-j\right)\left(\frac{z}{\theta_{1*}}\right)^{B_*}\right\}. \tag{8.1}$$

The likelihood can be obtained from the product of the likelihood for the first stress level, which features the ordering of data there, together with the likelihood at the remaining stress levels, given by

$$L_i = \prod_{j=1}^{M_i}\frac{B_* Y_{i,j}^{B_*-1}}{\theta_{i*}^{B_*}}\exp\left\{-\left(\frac{Y_{i,j}}{\theta_{i*}}\right)^{B_*}\right\}\prod_{j=1}^{n_i-M_i}\exp\left\{-\left(\frac{Z}{\theta_{i*}}\right)^{B_*}\right\},$$

for $i \geq 2$. The corresponding log-likelihoods are

$$\begin{aligned} l_i &= M_i \ln B_* - M_i B_* \alpha_{W*} - M_i B_* \beta_{W*} \rho(X_i) + (B_* - 1) \sum_{j=1}^{M_i} \ln Y_{i,j} \\ &- \exp(-B_* \alpha_{W*}) \sum_{j=1}^{M_i} \exp\{-B_* \beta_{W*} \rho(X_i)\} Y_{i,j}^{B_*} - \\ &(n_i - M_i) \exp(-B_* \alpha_{W*}) \exp\{-B_* \beta_{W*} \rho(X_i)\} Z^{B_*}, \end{aligned}$$

again for $i \geq 2$. At the first stress level, we have, without loss of generality,

$$L_1 = \prod_{j=1}^{r} \frac{B_* Y_{(j:n_1,1)}^{B_*-1}}{\theta_{1*}^{B_*}} \exp\left\{-\left(\frac{Y_{(j:n_1,1)}}{\theta_{1*}}\right)^{B_*}\right\} \prod_{j=r+1}^{n_1} \exp\left\{-\left(\frac{Z}{\theta_{1*}}\right)^{B_*}\right\},$$

and so

$$\begin{aligned} l_1 &= r \ln B_* - r B_* \alpha_{W*} - r B_* \beta_{W*} \rho(X_1) + (B_* - 1) \sum_{j=1}^{r} \ln Y_{(j:n_1,1)} \\ &- \exp(-B_* \alpha_{W*}) \exp\{-B_* \beta_{W*} \rho(X_1)\} \sum_{j=1}^{r} Y_{(j:n_1,1)}^{B_*} - \\ &(n_1 - r) \exp(-B_* \alpha_{W*}) \exp\{B_* \beta_{W*} \rho(X_1)\} Z^{B_*}. \end{aligned}$$

To simplify matters, we introduce some further notation, and write

$$S_{e,i} = \sum_{j=1}^{M_i} \ln Y_{i,j}$$

for $2 \leq i \leq k$; at the first stress level, we have

$$S_{e,1} = \sum_{j=1}^{r} \ln Y_{(j:n_1,1)}.$$

We also write, for $i \geq 2$,

$$\begin{aligned} S_{i,f} &= \sum_{j=1}^{M_i} Y_{i,j}^{B_*}, \\ S_{i,f}^{(n)} &= \frac{\partial^n S_{i,f}}{\partial B_*^n} = \sum_{j=1}^{M_i} Y_{i,j}^{B_*} (\ln Y_{i,j})^n, \end{aligned}$$

and

$$\begin{aligned} S_{1,f} &= \sum_{j=1}^{r} Y_{(j:n_1,1)}^{B_*}, \\ S_{1,f}^{(n)} &= \frac{\partial^n S_{1,f}}{\partial B_*^n} = \sum_{j=1}^{r} Y_{(j:n_1,1)}^{B_*} \left\{ \ln Y_{(j:n_1,1)} \right\}^n, \end{aligned}$$

while, for the censored data, we have

$$\begin{aligned} S_{i,c} &= (n_i - M_i) Z^{B_*}, \\ S_{i,c}^{(n)} &= \frac{\partial^n S_{i,c}}{\partial B_*^n} = (n_i - M_i) Z^{B_*} (\ln Z)^n, \end{aligned}$$

again, for $i \geq 2$, and

$$\begin{aligned} S_{1,c} &= (n_1 - r) Z^{B_*}, \\ S_{1,c}^{(n)} &= \frac{\partial^n S_{1,c}}{\partial B_*^n} = (n_1 - r) Z^{B_*} (\ln Z)^n. \end{aligned}$$

We also introduce further functions of the data, which appear in the full log-likelihood. These are given by

$$\begin{aligned} R^1 (B_*, \beta_{W*}) &= \exp \{-B_* \beta_{W*} \rho (X_1)\} \{S_{1,f} + S_{1,c}\}, \\ R (B_*, \beta_{W*}) &= \sum_{i=2}^{k} \exp \{-B_* \beta_{W*} \rho (X_i)\} \{S_{i,f} + S_{i,c}\}, \\ R_{1,0}^1 (B_*, \beta_{W*}) &= \frac{\partial R^1 (B_*, \beta_{W*})}{\partial B_*} \\ &= \exp \{-B_* \beta_{W*} \rho (X_1)\} \left[\begin{array}{c} S_{1,f}^{(1)} + S_{1,c}^{(1)} - \\ \beta_{W*} \rho (X_1) \{S_{1,f} + S_{1,c}\} \end{array} \right], \\ R_{1,0} (B_*, \beta_{W*}) &= \frac{\partial R (B_*, \beta_{W*})}{\partial B_*} \\ &= \sum_{i=2}^{k} \exp \{-B_* \beta_{W*} \rho (X_i)\} \left[\begin{array}{c} S_{i,f}^{(1)} + S_{i,c}^{(1)} - \\ \beta_{W*} \rho (X_i) \{S_{i,f} + S_{i,c}\} \end{array} \right], \\ R_{0,1}^1 (B_*, \beta_{W*}) &= \frac{\partial R^1 (B_*, \beta_{W*})}{\partial \beta_{W*}} \\ &= -B_* \rho (X_1) \exp \{-B_* \beta_{W*} \rho (X_1)\} \{S_{1,f} + S_{1,c}\}, \\ R_{0,1} (B_*, \beta_{W*}) &= \frac{\partial R (B_*, \beta_{W*})}{\partial \beta_{W*}} \\ &= -B_* \sum_{i=2}^{k} \rho (X_i) \exp \{-B_* \beta_{W*} \rho (X_i)\} \{S_{i,f} + S_{i,c}\}, \end{aligned}$$

$$
\begin{aligned}
R^1_{2,0}(B_*,\beta_{W*}) &= \frac{\partial R^1_{1,0}(B_*,\beta_{W*})}{\partial B_*} \\
&= \exp\{-B_*\beta_{W*}\rho(X_1)\}\left[\begin{array}{c} S^{(2)}_{1,f}+S^{(2)}_{1,c}-2\beta_{W*}\rho(X_1)\left\{S^{(1)}_{1,f}+S^{(1)}_{1,c}\right\} \\ +\beta^2_{W*}\rho(X_1)^2\{S_{1,f}+S_{1,c}\} \end{array}\right], \\
R_{2,0}(B_*,\beta_{W*}) &= \frac{\partial R_{1,0}(B_*,\beta_{W*})}{\partial B_*} \\
&= \sum_{i=2}^{k}\exp\{-B_*\beta_{W*}\rho(X_i)\}\left[\begin{array}{c} S^{(2)}_{i,f}+S^{(2)}_{i,c}-2\beta_{W*}\rho(X_i)\left\{S^{(1)}_{i,f}+S^{(1)}_{i,c}\right\} \\ +\beta^2_{W*}\rho(X_i)^2\{S_{i,f}+S_{i,c}\} \end{array}\right], \\
R^1_{0,2}(B_*,\beta_{W*}) &= \frac{\partial R^1_{0,1}(B_*,\beta_{W*})}{\partial \beta_{W*}} \\
&= B^2_*\rho(X_1)^2\exp\{-B_*\beta_{W*}\rho(X_1)\}\{S_{1,f}+S_{1,c}\}, \\
R_{0,2}(B_*,\beta_{W*}) &= \frac{\partial R_{0,1}(B_*,\beta_{W*})}{\partial \beta_{W*}} \\
&= B^2_*\sum_{i=2}^{k}\rho(X_i)^2\exp\{-B_*\beta_{W*}\rho(X_i)\}\{S_{i,f}+S_{i,c}\},
\end{aligned}
$$

and, finally,

$$
\begin{aligned}
R^1_{1,1}(B_*,\beta_{W*}) &= \frac{\partial R^1_{1,0}(B_*,\beta_{W*})}{\partial \beta_{W*}} \\
&= -\rho(X_1)\exp\{-B_*\beta_{W*}\rho(X_1)\}\left[\begin{array}{c} B_*\left(S^{(1)}_{1,f}+S^{(1)}_{1,c}\right) \\ +(S_{1,f}+S_{1,c})\{1-B_*\beta_{W*}\rho(X_1)\} \end{array}\right], \\
R_{1,1}(B_*,\beta_{W*}) &= \frac{\partial R_{1,0}(B_*,\beta_{W*})}{\partial \beta_{W*}} \\
&= -\sum_{i=2}^{k}\rho(X_i)\exp\{-B_*\beta_{W*}\rho(X_i)\}\left[\begin{array}{c} B_*\left(S^{(1)}_{i,f}+S^{(1)}_{i,c}\right) \\ +(S_{i,f}+S_{i,c})\{1-B_*\beta_{W*}\rho(X_i)\} \end{array}\right].
\end{aligned}
$$

The full log-likelihood is the sum of the log-likelihoods at each stress level, and we have

$$
\begin{aligned}
l_{W*} &= \sum_{i=1}^{k} l_i = \sum_{i=1}^{k} M_i \ln B_* - B_*\alpha_{W*}\sum_{i=1}^{k} M_i \\
&\quad -B_*\beta_{W*}\sum_{i=1}^{k} M_i\rho(X_i) + (B_*-1)\sum_{i=1}^{k} S_{e,i} \\
&\quad -\exp(-B_*\alpha_{W*})\left\{R^1(B_*,\beta_{W*}) + R(B_*,\beta_{W*})\right\}.
\end{aligned}
$$

The three score functions are

$$\begin{aligned}\frac{\partial l_{W*}}{\partial B_*} &= \sum_{i=1}^{k}\frac{M_i}{B_*} - \alpha_{W*}\sum_{i=1}^{k} M_i - \beta_{W*}\sum_{i=1}^{k} M_i\rho\left(X_i\right) + \sum_{i=1}^{k} S_{e,i} \\ &\quad - \exp\left(-B_*\alpha_{W*}\right)\left\{R_{1,0}^{1}\left(B_*,\beta_{W*}\right) - \alpha_{W*}R^{1}\left(B_*,\beta_{W*}\right)\right\} \\ &\quad - \exp\left(-B_*\alpha_{W*}\right)\left\{R_{1,0}\left(B_*,\beta_{W*}\right) - \alpha_{W*}R\left(B_*,\beta_{W*}\right)\right\},\end{aligned}$$

$$\frac{\partial l_{W*}}{\partial \beta_{W*}} = -B_*\sum_{i=1}^{k} M_i\rho\left(X_i\right) - \exp\left(-B_*\alpha_{W*}\right)\left\{R_{0,1}^{1}\left(B_*,\beta_{W*}\right) + R_{0,1}\left(B_*,\beta_{W*}\right)\right\},$$

and

$$\frac{\partial l_{W*}}{\partial \alpha_{W*}} = -B_*\sum_{i=1}^{k} M_i + B_*\exp\left(-B_*\alpha_{W*}\right)\left\{R^{1}\left(B_*,\beta_{W*}\right) + R\left(B_*,\beta_{W*}\right)\right\}.$$

If we equate $\frac{\partial l_{W*}}{\partial \alpha_{W*}}$ to zero, and solve, then we obtain

$$\alpha_{W*} = \frac{1}{B_*}\ln\left\{\frac{R^{1}\left(B_*,\beta_{W*}\right) + R\left(B_*,\beta_{W*}\right)}{\sum_{i=1}^{k} M_i}\right\},$$

so that, without loss of generality, the profile log-likelihood then becomes

$$\begin{aligned}l_{W*}^{+} &= \sum_{i=1}^{k} M_i\ln B_* - B_*\beta_{W*}\sum_{i=1}^{k} M_i\rho\left(X_i\right) + \left(B_* - 1\right)\left\{\sum_{i=1}^{k} S_{e,i}\right\} \\ &\quad - \sum_{i=1}^{k} M_i\ln\left\{R^{1}\left(B_*,\beta_{W*}\right) + R\left(B_*,\beta_{W*}\right)\right\}.\end{aligned}$$

The profile score functions are

$$\begin{aligned}\frac{\partial l_{W*}^{+}}{\partial B_*} &= \frac{\sum_{i=1}^{k} M_i}{B_*} - \beta_{W*}\sum_{i=1}^{k} M_i\rho\left(X_i\right) + \sum_{i=1}^{k} S_{e,i} \\ &\quad - \frac{\sum_{i=1}^{k} M_i\left\{R_{1,0}^{1}\left(B_*,\beta_{W*}\right) + R_{1,0}\left(B_*,\beta_{W*}\right)\right\}}{R^{1}\left(B_*,\beta_{W*}\right) + R\left(B_*,\beta_{W*}\right)},\end{aligned}$$

and

$$\frac{\partial l_{W*}^{+}}{\partial \beta_{W*}} = -\frac{\sum_{i=1}^{k} M_i\left\{R_{0,1}^{1}\left(B_*,\beta_{W*}\right) + R_{0,1}\left(B_*,\beta_{W*}\right)\right\}}{R^{1}\left(B_*,\beta_{W*}\right) + R\left(B_*,\beta_{W*}\right)} - B_*\sum_{i=1}^{k} M_i\rho\left(X_i\right).$$

In order to find the MLEs, we also require second derivatives, given by

$$\frac{\partial^2 l^+_{W*}}{\partial B_*^2} = -\frac{\sum_{i=1}^k M_i}{B_*^2} - \frac{\sum_{i=1}^k M_i \left[\begin{array}{c} \left\{R^1(B_*,\beta_{W*}) + R(B_*,\beta_{W*})\right\} \\ \times \left\{R^1_{2,0}(B_*,\beta_{W*}) + R_{2,0}(B_*,\beta_{W*})\right\} \\ - \left\{R^1_{1,0}(B_*,\beta_{W*}) + R_{1,0}(B_*,\beta_{W*})\right\}^2 \end{array} \right]}{\left\{R^1(B_*,\beta_{W*}) - R(B_*,\beta_{W*})\right\}^2},$$

$$\frac{\partial^2 l^+_{W*}}{\partial \beta_{W*}^2} = -\frac{\sum_{i=1}^k M_i \left[\begin{array}{c} \left\{R^1(B_*,\beta_{W*}) + R(B_*,\beta_{W*})\right\} \\ \times \left\{R^1_{0,2}(B_*,\beta_{W*}) + R_{0,2}(B_*,\beta_{W*})\right\} \\ - \left\{R^1_{0,1}(B_*,\beta_{W*}) + R_{0,1}(B_*,\beta_{W*})\right\}^2 \end{array} \right]}{\left\{R^1(B_*,\beta_{W*}) + R(B_*,\beta_{W*})\right\}^2},$$

and

$$\frac{\partial^2 l^+_{W*}}{\partial B_* \partial \beta_{W*}} = -\frac{\sum_{i=1}^k M_i \left[\begin{array}{c} \left\{R^1(B_*,\beta_{W*}) + R(B_*,\beta_{W*})\right\} \\ \times \left\{R^1_{1,1}(B_*,\beta_{W*}) + R_{1,1}(B_*,\beta_{W*})\right\} \\ - \left\{R^1_{0,1}(B_*,\beta_{W*}) + R_{0,1}(B_*,\beta_{W*})\right\} \\ \times \left\{R^1_{1,0}(B_*,\beta_{W*}) + R_{1,0}(B_*,\beta_{W*})\right\} \end{array} \right]}{\left\{R^1(B_*,\beta_{W*}) + R(B_*,\beta_{W*})\right\}^2} - \sum_{i=1}^k M_i \rho(X_i).$$

Next, we consider expectations required for the EFI matrix from G_{W*}.

8.2.1 The EFI matrix for G_{W*}

We begin with various expectations required for the EFI matrix from G_{W*}, and recall that these are obtained through an argument based on conditioning on the random stopping time Z. We also have different forms of expectations depending on whether we consider the first stress level or any other.

Expectations at X_1

We first consider $E[Z]$. The pdf for Z is given at (8.1); thus, for general m

$$E[Z^m] = \frac{B_* n_1!}{\theta_{1*}^{B_*}(n_1 - r)!(r-1)!} \sum_{j=0}^{r-1} (-1)^{r-1-j} \binom{r-1}{j} \times \int_0^\infty z^{m+B_*-1} \exp\left\{-(n_1 - j)\left(\frac{z}{\theta_{1*}}\right)^{B_*}\right\} dz.$$

We write $u = (n_1 - j)\left(\frac{z}{\theta_{1*}}\right)^{B_*}$ in the integrand, so that the integral form of the gamma function appears, and

$$E\left[Z^m\right] = \frac{n_1!\theta_{1*}^m \Gamma\left(\frac{m}{B_*}+1\right)}{(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)^{\frac{m}{B_*}+1}}, \tag{8.2}$$

with $(m = 1)$

$$E\left[Z\right] = \frac{n_1!\theta_{1*}\Gamma\left(\frac{1}{B_*}+1\right)}{(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)^{\frac{1}{B_*}+1}};$$

this expectation will only simplify further if $B_* = 1$. Next, we consider $E\left[S_{1,f}\right]$, and use (4.35) to write

$$E\left[S_{1,f}\right] = (n_1-r+1)(n_1-r)\binom{n_1}{r-1}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}E\left[Y_{(1:n_1-j,1)}^{B_*}\right]}{(n_1-1-j)(n_1-j)}.$$

We know that $Y_{(1:n_1,1)}$ is a random variable from $G_w\left(B_*, \theta_{1*}n_1^{-B_*^{-1}}\right)$. Thus,

$$E\left[Y_{(1:n_1,1)}^m\right] = \left(\frac{\theta_{1*}}{n_1^{\frac{1}{B_*}}}\right)^m \Gamma\left(\frac{m}{B_*}+1\right),$$

and so using this,

$$\begin{aligned} E\left[\sum_{j=1}^{r} Y_{(j:n_1,1)}^m\right] &= (n_1-r+1)(n_1-r)\binom{n_1}{r-1}\theta_{1*}^m\Gamma\left(\frac{m}{B_*}+1\right) \\ &\quad\times\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-1-j)(n_1-j)^{\frac{m}{B_*}+1}}. \end{aligned} \tag{8.3}$$

Thus, if we set $m = B_*$ in (8.3), we have

$$E\left[S_{1,f}\right] = (n_1-r+1)(n_1-r)\binom{n_1}{r-1}\theta_{1*}^{B_*}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-1-j)(n_1-j)^2},$$

which simplifies to

$$\theta_{1*}^{B_*}\left[r - (n_1-r)\sum_{j=0}^{r-1}(n_1-j)^{-1}\right].$$

Next, we consider $E\left[S_{1,f}^{(1)}\right]$; from (4.35), we write this as

$$(n_1-r+1)(n_1-r)\binom{n_1}{r-1}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}E\left[Y_{(1:n_1-j,1)}^{B_*}\ln Y_{(1:n_1-j,1)}\right]}{(n_1-1-j)(n_1-j)}.$$

For general m, we have

$$\begin{aligned} E\left[\sum_{j=1}^{r}Y_{(j:n_1,1)}^{m}\ln Y_{(j:n_1,1)}\right] &= (n_1-r+1)(n_1-r)\binom{n_1}{r-1}\theta_{1*}^{m}\Gamma\left(\frac{m}{B_*}+1\right) \\ &\quad\times\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\begin{bmatrix}\ln\theta_{1*}\\ -B_*^{-1}\ln(n_1-j)\\ +B_*^{-1}\Psi\left(\frac{m}{B_*}+1\right)\end{bmatrix}}{(n_1-1-j)(n_1-j)^{\frac{m}{B_*}+1}}, \end{aligned} \tag{8.4}$$

and so

$$\begin{aligned} E\left[S_{1,f}^{(1)}\right] &= (n_1-r+1)(n_1-r)\binom{n_1}{r-1}\theta_{1*}^{B_*} \\ &\quad\times\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\left\{\ln\theta_{1*}-B_*^{-1}\ln(n_1-j)+B_*^{-1}\Psi(2)\right\}}{(n_1-1-j)(n_1-j)^2} \\ &= (n_1-r+1)(n_1-r)\binom{n_1}{r-1}\theta_{1*}^{B_*} \\ &\quad\times\begin{bmatrix}\left\{\ln\theta_{1*}+B_*^{-1}\Psi(2)\right\}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-1-j)(n_1-j)^2}\\ -B_*^{-1}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln(n_1-j)}{(n_1-1-j)(n_1-j)^2}\end{bmatrix}; \end{aligned}$$

we use partial fractions and (4.39) to write

$$\begin{aligned} \sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-1-j)(n_1-j)^2} &= B(n_1-r,r)-B(n_1-r+1,r) \\ &\quad -B(n_1-r+1,r)\sum_{j=0}^{r-1}(n_1-j)^{-1}, \end{aligned}$$

and so $E\left[S_{1,f}^{(1)}\right]$ becomes

$$\begin{aligned} &\theta_{1*}^{B_*}\left\{\ln\theta_{1*}+B_*^{-1}\Psi(2)\right\}\left[r-(n_1-r)\sum_{j=0}^{r-1}(n_1-j)^{-1}\right] \\ &-\frac{n_1!\,\theta_{1*}^{B_*}}{B_*(r-1)!\,(n_1-r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln(n_1-j)}{(n_1-1-j)(n_1-j)^2}. \end{aligned}$$

Next, we consider $E\left[S_{1,f}^{(2)}\right]$. This is given by

$$(n_1 - r + 1)(n_1 - r)\binom{n_1}{r-1}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}E\left[Y_{(1:n_1-j,1)}^{B_*}\left(\ln Y_{(1:n_1-j,1)}\right)^2\right]}{(n_1-1-j)(n_1-j)}.$$

Our general expression for $E\left[Y_{(1:n_1,1)}^{m}\left\{\ln Y_{(1:n_1,1)}\right\}^2\right]$ is

$$n_1^{-\frac{m}{B_*}}\theta_{1*}^{m}\Gamma\left(\frac{m}{B_*}+1\right)\left[\begin{array}{c} B_*^{-2}\Psi'\left(\frac{m}{B_*}+1\right) \\ +\left\{\ln\theta_{1*} - B_*^{-1}\ln n_1 + B_*^{-1}\Psi\left(\frac{m}{B_*}+1\right)\right\}^2 \end{array}\right].$$

Therefore

$$\begin{aligned} E\left[\sum_{j=1}^{r} Y_{(j:n_1,1)}^{m}\left\{\ln Y_{(j:n_1,1)}\right\}^2\right] &= (n_1-r+1)(n_1-r)\binom{n_1}{r-1}\theta_{1*}^{m}\Gamma\left(\frac{m}{B_*}+1\right) \\ &\quad\times\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-1-j)(n_1-j)^{\frac{m}{B_*}+1}} \\ &\quad\times\left[B_*^{-2}\Psi'\left(\frac{m}{B_*}+1\right)+\left\{\begin{array}{c}\ln\theta_{1*} - B_*^{-1}\ln(n_1-j)+ \\ B_*^{-1}\Psi\left(\frac{m}{B_*}+1\right)\end{array}\right\}^2\right], \end{aligned}$$

and so

$$\begin{aligned} E\left[S_{1,f}^{(2)}\right] &= (n_1-r+1)(n_1-r)\binom{n_1}{r-1}\theta_{1*}^{B_*} \\ &\quad\times\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-1-j)(n_1-j)^2} \\ &\quad\times\left[B_*^{-2}\Psi'(2)+\left\{\begin{array}{c}\ln\theta_{1*} - B_*^{-1}\ln(n_1-j)+ \\ B_*^{-1}\Psi(2)\end{array}\right\}^2\right]. \end{aligned}$$

After some simplification, this becomes

$$\begin{aligned} &\theta_{1*}^{B_*}\left[B_*^{-2}\Psi'(2)+\left\{\ln\theta_{1*}+B_*^{-1}\Psi(2)\right\}^2\right]\left\{r-(n_1-r)\sum_{j=0}^{r-1}(n_1-j)^{-1}\right\} \\ &-\frac{n_1!\theta_{1*}^{B_*}}{B_*(n_1-r-1)!(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln(n_1-j)}{(n_1-1-j)(n_1-j)^2} \\ &\times\left[2\left\{\ln\theta_{1*}+B_*^{-1}\Psi(2)\right\}-B_*^{-1}\ln(n_1-j)\right]. \end{aligned}$$

We next obtain $E\left[S_{e,1}\right]$ by setting $m=0$ in (8.4); this expectation is

$$(n_1-r+1)(n_1-r)\binom{n_1}{r-1}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\left\{\ln\theta_{1*}-B_*^{-1}\ln(n_1-j)+B_*^{-1}\Psi(1)\right\}}{(n_1-1-j)(n_1-j)},$$

which, after using partial fractions and multiplying out the bracket, simplifies to

$$r\left\{\ln\theta_{1*}+B_*^{-1}\Psi(1)\right\}-\frac{n_1!}{B_*(r-1)!(n_1-r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln(n_1-j)}{(n_1-j-1)(n_1-j)}.$$

We now consider expectations for censored data at X_1, and first derive $E\left[S_{1,c}\right]$. This is given by

$$(n_1-r)E\left[Z^{B_*}\right]=\frac{(n_1-r)\,n_1!\,\theta_{1*}^{B_*}}{(n_1-r)!(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)^2},$$

from (8.2), which, on using (4.39), simplifies to

$$(n_1-r)\,\theta_{1*}^{B_*}\sum_{j=0}^{r-1}(n_1-j)^{-1}.$$

Next, we consider $E\left[S_{1,c}^{(1)}\right]$, differentiating (8.2) to obtain

$$\begin{aligned}E\left[Z^m\ln Z\right] &= \frac{\partial}{\partial m}E\left[Z^m\right]=\frac{n_1!\,\theta_{1*}^m\Gamma\left(\frac{m}{B_*}+1\right)}{(n_1-r)!(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)^{\frac{m}{B_*}+1}}\\ &\quad\times\left\{\ln\theta_{1*}+B_*^{-1}\Psi\left(\frac{m}{B_*}+1\right)-B_*^{-1}\ln(n_1-j)\right\},\end{aligned}$$

so that (with $m=B_*$)

$$\begin{aligned}E\left[S_{1,c}^{(1)}\right] &= \frac{(n_1-r)\,n_1!\,\theta_{1*}^{B_*}}{(n_1-r)!(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)^2}\\ &\quad\times\left\{\ln\theta_{1*}+B_*^{-1}\Psi(2)-B_*^{-1}\ln(n_1-j)\right\},\end{aligned}$$

which is equivalent to

$$\theta_{1*}^{B_*}(n_1-r)\left\{\ln\theta_{1*}+B_*^{-1}\Psi(2)\right\}\sum_{j=0}^{r-1}(n_1-j)^{-1}$$
$$-\frac{(n_1-r)\,n_1!\,\theta_{1*}^{B_*}}{B_*(n_1-r)!(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln(n_1-j)}{(n_1-j)^2}.$$

We consider

$$E\left[S_{1,c}^{(2)}\right] = (n_1 - r)\, E\left[Z^{B_*} (\ln Z)^2\right];$$

the final expectation at X_1. We differentiate

$$E\left[Z^m \ln Z\right]$$

with respect to m to obtain

$$\begin{aligned} E\left[Z^m (\ln Z)^2\right] &= \frac{n_1!\theta_{1*}^m \Gamma\left(\frac{m}{B_*}+1\right)}{(r-1)!\,(n_1-r)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)^{\frac{m}{B_*}+1}} \\ &\quad \times \left[\begin{array}{c} B_*^{-2}\Psi'\left(\frac{m}{B_*}+1\right) \\ +\left\{\ln\theta_{1*} + B_*^{-1}\Psi\left(\frac{m}{B_*}+1\right) - B_*^{-1}\ln(n_1-j)\right\}^2 \end{array} \right]. \end{aligned}$$

Hence, by substituting B_* for m in this expression, we have

$$\begin{aligned} E\left[S_{1,c}^{(2)}\right] &= (n_1 - r)\,\theta_{1*}^{B_*} B_*^{-2}\Psi'(2) \sum_{j=0}^{r-1} (n_1 - j)^{-1} \\ &\quad + \frac{n_1!\theta_{1*}^{B_*}}{(r-1)!\,(n_1-r-1)!} \sum_{j=0}^{r-1-j} \frac{(-1)^{r-1-j}\binom{r-1}{j} \left\{ \begin{array}{c} \ln\theta_{1*} + B_*^{-1}\Psi(2) \\ -B_*^{-1}\ln(n_1-j) \end{array} \right\}^2}{(n_1-j)^2}. \end{aligned}$$

Expectations at X_i $(i \geq 2)$

Now we derive $E[M_i]$. If we first condition on $Z = z$, then M_i will have a Binomial distribution with parameters n_i and

$$q_i(z) = 1 - \exp\left\{-\left(\frac{z}{\theta_{i*}}\right)^{B_*}\right\},$$

since a failure at the i^{th} stress level has to occur in the interval $(0, z)$; hence,

$$E\left[M_i \mid Z = z\right] = n_i q_i(z).$$

We now take expectations with respect to Z to obtain

$$E[M_i] = n_i\left(1 - E\left[\exp\left\{-\left(\frac{Z}{\theta_{i*}}\right)^{B_*}\right\}\right]\right),$$

where

$$E\left[\exp\left\{-\left(\frac{Z}{\theta_{i*}}\right)^{B_*}\right\}\right] = \frac{B_* n_1!}{\theta_{1*}^{B_*}(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}(-1)^{r-1-j}\binom{r-1}{j} \times\int_0^\infty z^{B_*-1}\exp\left[-\left\{\frac{1}{\theta_{i*}^{B_*}}+\frac{(n_1-j)}{\theta_{1*}^{B_*}}\right\}z^{B_*}\right]dz,$$

which, after setting

$$u=\left\{\frac{1}{\theta_{i*}^{B_*}}+\frac{(n_1-j)}{\theta_{1*}^{B_*}}\right\}z^{B_*}$$

reduces to

$$\frac{n_1!\theta_{i*}^{B_*}}{(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{\theta_{1*}^{B_*}+\theta_{i*}^{B_*}(n_1-j)}.$$

Unlike the expression for $E[Z]$, this expectation will simplify quite considerably. We write

$$\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{\theta_{1*}^{B_*}+\theta_{i*}^{B_*}(n_1-j)}=\theta_{i*}^{-B_*}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j},$$

and use (4.38) to write the summation in terms of the Beta function

$$\theta_{i*}^{-B_*}B\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1,r\right\}.$$

Thus, we have

$$E[M_i]=n_i\left[1-\frac{n_1!\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{(n_1-r)!\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}}\right].$$

Next, we consider $E[S_{i,f}]$, which requires consideration of the distribution of the failures at X_i. We denote Y_i as a failure at X_i for $i\geq 2$, and note that the truncated pdf of this random variable is given by

$$\frac{B_* y^{B_*-1}\exp\left\{-\left(\frac{y}{\theta_{i*}}\right)^{B_*}\right\}}{\theta_{i*}^{B_*}q_i(z)}.$$

We first condition on the values of M_i and Z to obtain

$$E\left[\sum_{j=1}^{M_i} Y_{i,j}^m \middle| M_i = m_i, Z = z\right] = m_i E\left[Y_i^m \middle| Z = z\right],$$

where

$$\begin{aligned} E\left[Y_i^m \middle| Z = z\right] &= \int_0^z \frac{B_* y^{m+B_*-1} \exp\left\{-\left(\frac{y}{\theta_{i*}}\right)^{B_*}\right\}}{\theta_{i*}^{B_*} q_i(z)} dy \\ &= \frac{\theta_{i*}^m}{q_i(z)} \Gamma\left\{\left(\frac{z}{\theta_{i*}}\right)^{B_*}, \frac{m}{B_*} + 1\right\}. \end{aligned}$$

Therefore

$$\begin{aligned} E\left[\sum_{j=1}^{M_i} Y_{i,j}^m \middle| Z = z\right] &= E\left[M_i\right] \frac{\theta_{i*}^m}{q_i(z)} \Gamma\left\{\left(\frac{z}{\theta_{i*}}\right)^{B_*}, \frac{m}{B_*} + 1\right\} \\ &= n_i \theta_{i*}^m \Gamma\left\{\left(\frac{z}{\theta_{i*}}\right)^{B_*}, \frac{m}{B_*} + 1\right\}. \end{aligned}$$

We now use (1.10) to write

$$\Gamma\left\{\left(\frac{z}{\theta_{i*}}\right)^{B_*}, \frac{m}{B_*} + 1\right\} = \sum_{p=0}^{\infty} \frac{(-1)^p \left(\frac{z}{\theta_{i*}}\right)^{B_*\left(\frac{m}{B_*}+1+p\right)}}{\left(\frac{m}{B_*} + 1 + p\right) p!},$$

so that

$$E\left[\sum_{j=1}^{M_i} Y_{i,j}^m \middle| Z = z\right] = n_i \sum_{p=0}^{\infty} \frac{(-1)^p z^{m+B_*(p+1)}}{\theta_{i*}^{B_*(p+1)} \left(\frac{m}{B_*} + 1 + p\right) p!}.$$

We now take expectations with respect to Z; on moving the expectation through the infinite summation, we have

$$\begin{aligned} E\left[\sum_{j=1}^{M_i} Y_{i,j}^m\right] &= n_i \sum_{p=0}^{\infty} \frac{(-1)^p E\left[Z^{m+B_*(p+1)}\right]}{\theta_{i*}^{B_*(p+1)} \left(\frac{m}{B_*} + 1 + p\right) p!} \\ &= \frac{n_i n_1! \theta_{1*}^m}{(n_1 - r)!(r-1)!} \sum_{p=0}^{\infty} \frac{(-1)^p \left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*(p+1)} \Gamma\left(\frac{m}{B_*} + p + 1\right)}{p!} \\ &\quad \times \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j} \binom{r-1}{j}}{(n_1 - j)^{\frac{m}{B_*}+p+2}}, \end{aligned}$$

and can simplify this further by interchanging both summations and introducing hypergeometric functions. Thus, we write

$$E\left[\sum_{j=1}^{M_i} Y_{i,j}^m\right] = \frac{n_i n_1! \theta_{1*}^m}{(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}\Gamma\left(\frac{m}{B_*}+1\right)}{(n_1-j)^{\frac{m}{B_*}+2}} F_{1,0}\left(\frac{m}{B_*}+1;\frac{-\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}}{n_1-j}\right),$$

where (since $F_{1,0}(u,z) = (1-z)^{-u}$) we obtain

$$E\left[\sum_{j=1}^{M_i} Y_{i,j}^m\right] = \frac{n_i n_1! \theta_{1*}^m \left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}\Gamma\left(\frac{m}{B_*}+1\right)}{(n_1-r)!\,(r-1)!} \times\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{\frac{m}{B_*}+1}}.$$

Therefore, we have $(m = B_*)$

$$E\left[S_{i,f}\right] = \frac{n_i n_1! \theta_{1*}^{B_*}\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}}{(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^2},$$

which, on using (4.39), simplifies to

$$\frac{n_i n_1! \theta_{1*}^{B_*}}{(n_1-r)!}\left[\begin{array}{c} \frac{(n_1-r)!}{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} n_1!} - \frac{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}} \\ -\frac{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}}\sum_{j=0}^{r-1}\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{-1} \end{array}\right].$$

Next, we consider calculating $E\left[S_{i,f}^{(1)}\right]$. We have

$$\begin{aligned} E\left[\sum_{j=1}^{M_i} Y_{i,j}^m \ln Y_{i,j}\right] &= \frac{\partial}{\partial m} E\left[\sum_{j=1}^{M_i} Y_{i,j}^m\right] \\ &= \frac{n_i n_1! \theta_{1*}^m \left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} \Gamma\left(\frac{m}{B_*}+1\right)}{(n_1-r)!\,(r-1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{\frac{m}{B_*}+1}} \\ &\quad \times \left[\ln\theta_{1*} + B_*^{-1}\Psi\left(\frac{m}{B_*}+1\right) - B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}\right], \quad (8.5) \end{aligned}$$

and so, with $m = B_*$, we obtain

$$\begin{aligned} E\left[S_{i,f}^{(1)}\right] &= \frac{n_i n_1! \theta_{1*}^{B_*} \left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}}{(n_1-r)!\,(r-1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{2}} \\ &\quad \times \left[\ln\theta_{1*} + B_*^{-1}\Psi(2) - B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}\right], \end{aligned}$$

which simplifies to

$$\begin{aligned} &\frac{n_i n_1! \theta_{1*}^{B_*}\left(\ln\theta_{1*} + B_*^{-1}\Psi(2)\right)}{(n_1-r)!} \left[\begin{array}{c} \frac{(n_1-r)!}{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} n_1!} - \frac{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}} \\ -\frac{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}} \sum_{j=0}^{r-1}\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{-1} \end{array} \right] \\ &- \frac{n_i n_1! \theta_{1*}^{B_*}\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}}{B_*(n_1-r)!\,(r-1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{2}}. \end{aligned}$$

We next derive an expression for $E\left[S_{i,f}^{(2)}\right]$, using

$$
\begin{aligned}
E\left[\sum_{j=1}^{M_i} Y_{i,j}^m (\ln Y_{i,j})^2\right] &= \frac{\partial}{\partial m} E\left[\sum_{j=1}^{M_i} Y_{i,j}^m \ln Y_{i,j}\right] \\
&= \frac{n_i n_1! \theta_{1*}^m \left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} \Gamma\left(\frac{m}{B_*}+1\right)}{(n_1-r)!\,(r-1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{\frac{m}{B_*}+1}} \\
&\quad\times\left[B_*^{-2}\Psi'\left(\frac{m}{B_*}+1\right)+\left(\begin{array}{c}\ln\theta_{1*}+B_*^{-1}\Psi\left(\frac{m}{B_*}+1\right)\\ -B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}\end{array}\right)^2\right].
\end{aligned}
$$

Thus, (with $m = B_*$)

$$
E\left[S_{i,f}^{(2)}\right] = \frac{n_i n_1! \theta_{1*}^{B_*} \left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}}{(n_1-r)!\,(r-1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{2}}
$$
$$
\times\left[B_*^{-2}\Psi'(2)+\left(\begin{array}{c}\ln\theta_{1*}+B_*^{-1}\Psi(2)\\ -B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}\end{array}\right)^2\right],
$$

which simplifies to

$$
\frac{n_i n_1! \theta_{1*}^{B_*}\left[\left\{\ln\theta_{1*}+B_*^{-1}\Psi(2)\right\}^2+B_*^{-2}\Psi'(2)\right]}{(n_1-r)!}
$$
$$
\times\left[\begin{array}{c}\frac{(n_1-r)!}{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} n_1!}-\frac{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}}-\\ \frac{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}}\sum_{j=0}^{r-1}\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{-1}\end{array}\right]
$$
$$
-\frac{n_i n_1! \theta_{1*}^{B_*}\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}}{B_*(r-1)!\,(n_1-r)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^2}
$$
$$
\times\left[2\left\{\ln\theta_{1*}+B_*^{-1}\Psi(2)\right\}-B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}\right].
$$

Next, we consider

$$
E\left[S_{e,i}\right] = E\left[\sum_{j=1}^{M_i}\ln Y_{i,j}\right],
$$

which is obtained by setting $m = 0$ in (8.5); thus

$$\begin{aligned} E\left[S_{e,i}\right] &= \frac{n_i n_1! \left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}}{(n_1 - r)!\,(r-1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}}{(n_1 - j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}} \\ &\quad \times \left[\ln\theta_{1*} + B_*^{-1}\Psi(1) - B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}\right], \end{aligned}$$

which simplifies to

$$\begin{aligned} &\frac{n_i n_1! \left(\ln\theta_{1*} + B_*^{-1}\Psi(1)\right)}{(n_1 - r)!} \left[\frac{(n_1 - r)!}{n_1!} - \frac{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - r + 1\right\}}{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 + 1\right\}}\right] \\ &- \frac{n_i n_1! \left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}}{B_*\,(n_1 - r)!\,(r-1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}}{(n_1 - j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}}. \end{aligned}$$

We now consider expectations for censored data, and first derive $E\left[S_{i,c}\right]$. We consider

$$E\left[(n_i - M_i)\, Z^m\right],$$

and after first conditioning on the value of the random variable Z, we have

$$E\left[(n_i - M_i)\, Z^m \middle| Z = z\right] = n_i z^m \exp\left\{-\left(\frac{z}{\theta_{i*}}\right)^{B_*}\right\}.$$

We therefore consider

$$\begin{aligned} E\left[Z^m \exp\left\{-\left(\frac{Z}{\theta_{i*}}\right)^{B_*}\right\}\right] &= \frac{B_* n_1!}{\theta_{1*}^{B_*}\,(n_1 - r)!\,(r-1)!} \sum_{j=0}^{r-1} (-1)^{r-1-j}\binom{r-1}{j} \\ &\quad \times \int_0^\infty z^{m+B_*-1} \exp\left[-\left\{\frac{\theta_{i*}^{B_*}(n_1 - j) + \theta_{1*}^{B_*}}{\theta_{i*}^{B_*}\theta_{1*}^{B_*}}\right\} z^{B_*}\right] dz, \end{aligned}$$

and, on using

$$u = \left\{\frac{\theta_{i*}^{B_*}(n_1 - j) + \theta_{1*}^{B_*}}{\theta_{i*}^{B_*}\theta_{1*}^{B_*}}\right\} z^{B_*}$$

may write this expectation as

$$\frac{n_1!\theta_{1*}^m\Gamma\left(\frac{m}{B_*}+1\right)}{(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{\frac{m}{B_*}+1}}.$$

Hence, we have

$$E\left[(n_i-M_i)\,Z^m\right]=\frac{n_in_1!\theta_{1*}^m\Gamma\left(\frac{m}{B_*}+1\right)}{(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{\frac{m}{B_*}+1}},\tag{8.6}$$

and so (with $m=B_*$)

$$E\left[S_{i,c}\right]=\frac{n_in_1!\theta_{1*}^{B_*}}{(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{2}}.$$

Finally, we use (4.39) to write this as

$$\frac{n_in_1!\theta_{1*}^{B_*}\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{(n_1-r)!\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}}\sum_{j=0}^{r-1}\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{-1}.$$

Next, we consider $E\left[S_{i,c}^{(1)}\right]$; we compute

$$\frac{\partial}{\partial m}E\left[(n_i-M_i)\,Z^m\right],$$

and then set $m=B_*$. From (8.6), we have

$$\begin{aligned}E\left[(n_i-M_i)\,Z^m\ln Z\right] &= \frac{n_in_1!\theta_{1*}^m\Gamma\left(\frac{m}{B_*}+1\right)}{(n_1-r)!\,(r-1)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}^{\frac{m}{B_*}+1}}\\ &\quad\times\left[\begin{array}{c}\ln\theta_{1*}+B_*^{-1}\Psi\left(\frac{m}{B_*}+1\right)\\ -B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}\end{array}\right],\end{aligned}$$

and so $(m = B_*)$

$$E\left[S_{i,c}^{(1)}\right] = \frac{n_i n_1! \theta_{1*}^{B_*}}{(n_1 - r)!\,(r-1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}}{\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}^2} \times \left[\begin{array}{c} \ln\theta_{1*} + B_*^{-1}\Psi(2) \\ -B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}\end{array}\right].$$

This, in turn, simplifies to

$$\frac{n_i n_1! \theta_{1*}^{B_*}\left\{\ln\theta_{1*} + B_*^{-1}\Psi(2)\right\}\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - r + 1\right\}}{(n_1 - r)!\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 + 1\right\}} \sum_{j=0}^{r-1}\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}^{-1}$$
$$-\frac{n_i n_1! \theta_{1*}^{B_*}}{B_*(n_1 - r)!\,(r-1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}}{\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}^2}.$$

Finally, we compute $E\left[S_{i,c}^{(2)}\right]$; we again differentiate $E\left[(n_i - M_i)\, Z^m \ln Z\right]$ with respect to m, yielding

$$\frac{n_i n_1! \theta_{1*}^{m}\Gamma\left(\frac{m}{B_*} + 1\right)}{(r-1)!\,(n_1 - r)!.} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}}{\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}^{\frac{m}{B_*}+1}} \times \left[\begin{array}{c}\left(\ln\theta_{1*} + B_*^{-1}\Psi\left(\frac{m}{B_*} + 1\right) - B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} n_1 - j\right\}\right)^2 \\ +B_*^{-2}\Psi'\left(\frac{m}{B_*} + 1\right)\end{array}\right],$$

and, with $m = B_*$, we obtain

$$E\left[S_{i,c}^{(2)}\right] = \frac{n_i n_1! \theta_{1*}^{B_*}\Psi'(2)\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - r + 1\right\}}{B_*^2 (n_1 - r)!\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 + 1\right\}} \sum_{j=0}^{r-1}\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}^{-1}$$
$$+\frac{n_i n_1! \theta_{1*}^{B_*}}{(r-1)!\,(n_1 - r)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j}\binom{r-1}{j}}{\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}^2}$$
$$\times\left[\ln\theta_{1*} + B_*^{-1}\Psi(2) - B_*^{-1}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}\right]^2.$$

We use now use these results to compute the EFI matrix for the Weibull MLEs. The six second partial derivatives are given by

$$\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}^2} = -B_*^2 \exp\left(-B_*\alpha_{W*}\right)\left\{R^1\left(B_*, \beta_{W*}\right) + R\left(B_*, \beta_{W*}\right)\right\},$$

$$\frac{\partial^2 l_{W*}}{\partial \beta_{W*}^2} = -\exp\left(-B_*\alpha_{W*}\right)\left\{R_{0,2}^1\left(B_*, \beta_{W*}\right) + R_{0,2}\left(B_*, \beta_{W*}\right)\right\},$$

$$\begin{aligned}\frac{\partial^2 l_{W*}}{\partial B_*^2} &= -B_*^{-2}\sum_{i=1}^{k} M_i \\ &\quad - \exp\left(-B_*\alpha_{W*}\right)\left\{\begin{array}{c} R_{2,0}^1\left(B_*, \beta_{W*}\right) - 2\alpha_{W*}R_{1,0}^1\left(B_*, \beta_{W*}\right) \\ +\alpha_{W*}^2 R^1\left(B_*, \beta_{W*}\right) \\ +R_{2,0}\left(B_*, \beta_{W*}\right) - 2\alpha_{W*}R_{1,0}\left(B_*, \beta_{W*}\right) \\ +\alpha_{W*}^2 R\left(B_*, \beta_{W*}\right)\end{array}\right\},\end{aligned}$$

$$\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}\partial \beta_{W*}} = B_* \exp\left(-B_*\alpha_{W*}\right)\left\{R_{0,1}^1\left(B_*, \beta_{W*}\right) + R_{0,1}\left(B_*, \beta_{W*}\right)\right\},$$

$$\begin{aligned}\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}\partial B_*} &= -\sum_{i=1}^{k} M_i \\ &\quad + \exp\left(-B_*\alpha_{W*}\right)\left[\begin{array}{c} B_*R_{1,0}^1\left(B_*, \beta_{W*}\right) + \left\{1 - B_*\alpha_{W*}\right\}R^1\left(B_*, \beta_{W*}\right) \\ +B_*R_{1,0}\left(B_*, \beta_{W*}\right) + \left\{1 - B_*\alpha_{W*}\right\}R\left(B_*, \beta_{W*}\right)\end{array}\right],\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2 l_{W*}}{\partial \beta_{W*}\partial B_*} &= -\sum_{i=1}^{k} M_i \rho\left(X_i\right) \\ &\quad - \exp\left(-B_*\alpha_{W*}\right)\left[\begin{array}{c} R_{1,1}^1\left(B_*, \beta_{W*}\right) - \alpha_{W*}R_{0,1}^1\left(B_*, \beta_{W*}\right) \\ +R_{1,1}\left(B_*, \beta_{W*}\right) - \alpha_{W*}R_{0,1}\left(B_*, \beta_{W*}\right)\end{array}\right].\end{aligned}$$

We first consider the three diagonal elements, and have

$$\begin{aligned}E\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}^2}\right] &= -B_*^2\theta_{1*}^{-B_*}\left\{E\left[S_{1,f}\right] + E\left[S_{1,c}\right]\right\} - B_*^2\sum_{i=2}^{k}\theta_{i*}^{-B_*}\left\{E\left[S_{i,f}\right] + E\left[S_{i,c}\right]\right\} \\ &= -rB_*^2 - B_*^2\sum_{i=2}^{k} n_i + \frac{n_1! B_*^2}{(n_1 - r)!}\sum_{i=2}^{k}\frac{n_i\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - r + 1\right\}}{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 + 1\right\}}.\end{aligned}$$

Next, we derive

$$
\begin{aligned}
E\left[\frac{\partial^2 l_{W*}}{\partial \beta_{W*}^2}\right] &= -B_*^2 \rho (X_1)^2 \theta_{1*}^{-B_*} \{E[S_{1,f}] + E[S_{1,c}]\} \\
&\quad - B_*^2 \sum_{i=2}^{k} \rho (X_i)^2 \theta_{i*}^{-B_*} \{E[S_{i,f}] + E[S_{i,c}]\} \\
&= -rB_*^2 \rho (X_i)^2 - B_*^2 \sum_{i=2}^{k} n_i \rho (X_i)^2 \\
&\quad + \frac{n_1! B_*^2}{(n_1 - r)!} \sum_{i=2}^{k} \frac{n_i \rho (X_i)^2 \Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - r + 1\right\}}{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 + 1\right\}}.
\end{aligned}
$$

The third diagonal element is $E\left[\frac{\partial^2 l_{W*}}{\partial B_*^2}\right]$, which, after some simplification, becomes

$$
\begin{aligned}
& -rB_*^{-2} - B_*^{-2} \sum_{i=2}^{k} E[M_i] \\
& -\theta_{1*}^{-B_*} \left\{ \begin{array}{c} E\left[S_{1,f}^{(2)}\right] + E\left[S_{1,c}^{(2)}\right] - 2\ln\theta_{1*} E\left[S_{1,f}^{(1)}\right] - 2\ln\theta_{1*} E\left[S_{1,c}^{(1)}\right] \\ + (\ln\theta_{1*})^2 E[S_{1,f}] + (\ln\theta_{1*})^2 E[S_{1,c}] \end{array} \right\} \\
& -\sum_{i=2}^{k} \theta_{i*}^{-B_*} \left\{ \begin{array}{c} E\left[S_{i,f}^{(2)}\right] + E\left[S_{i,c}^{(2)}\right] - 2\ln\theta_{i*} E\left[S_{i,f}^{(1)}\right] - 2\ln\theta_{i*} E\left[S_{i,c}^{(1)}\right] \\ + (\ln\theta_{i*})^2 E[S_{i,f}] + (\ln\theta_{i*})^2 E[S_{i,c}] \end{array} \right\}.
\end{aligned}
$$

After substituting the relevant expectations into this expression, it becomes

$$
\begin{aligned}
& -rB_*^{-2}\left\{1 + \Psi'(2) + \Psi(2)^2\right\} - \sum_{i=2}^{k} n_i \left[B_*^{-2} + B_*^{-2}\Psi'(2) + \left\{\ln\left(\frac{\theta_{1*}}{\theta_{i*}}\right) + B_*^{-1}\Psi(2)\right\}^2\right] \\
& \times \left[1 - \frac{n_1! \Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - r + 1\right\}}{(n_1 - r)! \Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 + 1\right\}}\right] \\
& + \frac{n_1!}{B_*^2 (r-1)! (n_1 - r - 1)!} \sum_{j=0}^{r-1} \frac{(-1)^{r-1-j} \binom{r-1}{j} \ln\{n_1 - j\} [2\Psi(2) - \ln\{n_1 - j\}]}{(n_1 - 1 - j)(n_1 - j)} \\
& + \frac{n_1!}{B_* (r-1)! (n_1 - r)!} \sum_{i=2}^{k} n_i \left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} \sum_{j=0}^{r-1-j} \frac{(-1)^{r-1-j} \binom{r-1}{j} \ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}}{(n_1 - j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}} \\
& \times \left[2\left\{\ln\left(\frac{\theta_{1*}}{\theta_{i*}}\right) + B_*^{-1}\Psi(2)\right\} - B_*^{-1} \ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*} + n_1 - j\right\}\right],
\end{aligned}
$$

which does not seem to simplify further. Next, we consider the three off-diagonal elements, and have

$$
\begin{aligned}
E\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}\partial \beta_{W*}}\right] &= -B_*^2 \rho(X_1)\,\theta_{1*}^{-B_*}\left\{E\left[S_{1,f}\right]+E\left[S_{1,c}\right]\right\} \\
&\quad -B_*^2\sum_{i=2}^{k}\rho(X_i)\,\theta_{i*}^{-B_*}\left\{E\left[S_{i,f}\right]+E\left[S_{i,c}\right]\right\} \\
&= -rB_*^2\rho(X_1)-B_*^2\sum_{i=2}^{k}n_i\rho(X_i) \\
&\quad +\frac{n_1!B_*^2}{(n_1-r)!}\sum_{i=2}^{k}\frac{n_i\rho(X_i)\,\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}}.
\end{aligned}
$$

Now we examine $E\left[\frac{\partial^2 l_{W*}}{\partial \alpha_{W*}\partial B_*}\right]$, which is

$$
\begin{aligned}
&-r-\sum_{i=2}^{k}E\left[M_i\right]+B_*\theta_{1*}^{-B_*}E\left[S_{1,f}^{(1)}\right]+B_*\theta_{1*}^{-B_*}E\left[S_{1,c}^{(1)}\right] \\
&+\theta_{1*}^{-B_*}\left(1-B_*\ln\theta_{1*}\right)E\left[S_{1,f}\right]+\theta_{1*}^{-B_*}\left(1-B_*\ln\theta_{1*}\right)E\left[S_{1,c}\right] \\
&+B_*\sum_{i=2}^{k}\theta_{i*}^{-B_*}E\left[S_{i,f}^{(1)}\right]+B_*\sum_{i=2}^{k}\theta_{i*}^{-B_*}E\left[S_{i,c}^{(1)}\right] \\
&+\sum_{i=2}^{k}\theta_{i*}^{-B_*}\left(1-B_*\ln\theta_{i*}\right)E\left[S_{i,f}\right]+\sum_{i=2}^{k}\theta_{i*}^{-B_*}\left(1-B_*\ln\theta_{i*}\right)E\left[S_{i,c}\right],
\end{aligned}
$$

and simplifies to

$$
\begin{aligned}
&r\Psi(2)+B_*\sum_{i=2}^{k}n_i\left\{\ln\left(\frac{\theta_{1*}}{\theta_{i*}}\right)+B_*^{-1}\Psi(2)\right\}\left[1-\frac{n_1!\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{(n_1-r)!\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}}\right] \\
&-\frac{n_1!}{(r-1)!\,(n_1-r-1)!}\sum_{j=0}^{r-1-j}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln\left\{n_1-j\right\}}{(n_1-1-j)\,(n_1-j)} \\
&-\frac{n_1!}{(r-1)!\,(n_1-r)!}\sum_{i=2}^{k}n_i\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}\sum_{j=0}^{r-1-j}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}}.
\end{aligned}
$$

Finally, we compute $E\left[\frac{\partial^2 l_{W*}}{\partial\beta_{W*}\partial B_*}\right]$, which is

$$-r\rho(X_1)-\sum_{i=2}^{k}\rho(X_i)E[M_i]+B_*\rho(X_1)\theta_{1*}^{-B_*}E\left[S_{1,f}^{(1)}\right]+B_*\rho(X_1)\theta_{1*}^{-B_*}E\left[S_{1,c}^{(1)}\right]$$
$$+\rho(X_1)\theta_{1*}^{-B_*}(1-B_*\ln\theta_{1*})E[S_{1,f}]+\rho(X_1)\theta_{1*}^{-B_*}(1-B_*\ln\theta_{1*})E[S_{1,c}]$$
$$+B_*\sum_{i=2}^{k}\rho(X_i)\theta_{i*}^{-B_*}E\left[S_{i,f}^{(1)}\right]+B_*\sum_{i=2}^{k}\rho(X_i)\theta_{i*}^{-B_*}E\left[S_{i,c}^{(1)}\right]$$
$$+\sum_{i=2}^{k}\rho(X_i)\theta_{i*}^{-B_*}(1-B_*\ln\theta_{i*})E[S_{i,f}]+\sum_{i=2}^{k}\rho(X_i)\theta_{i*}^{-B_*}(1-B_*\ln\theta_{i*})E[S_{i,c}],$$

and reduces to

$$r\rho(X_1)\Psi(2)+B_*\sum_{i=2}^{k}n_i\rho(X_i)\left\{\ln\left(\frac{\theta_{1*}}{\theta_{i*}}\right)+B_*^{-1}\Psi(2)\right\}\left[1-\frac{n_1!\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-r+1\right\}}{(n_1-r)!\Gamma\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1+1\right\}}\right]$$
$$-\frac{n_1!\rho(X_1)}{(r-1)!(n_1-r-1)!}\sum_{j=0}^{r-1-j}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln\{n_1-j\}}{(n_1-1-j)(n_1-j)}$$
$$-\frac{n_1!}{(r-1)!(n_1-r)!}\sum_{i=2}^{k}n_i\rho(X_i)\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}\sum_{j=0}^{r-1-j}\frac{(-1)^{r-1-j}\binom{r-1}{j}\ln\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}}{(n_1-j)\left\{\left(\frac{\theta_{1*}}{\theta_{i*}}\right)^{B_*}+n_1-j\right\}}.$$

Using these elements which make up the EFI matrix for the Weibull MLEs, we can also compute the asymptotic theoretical mean and variance of $\widehat{B}_{W,10}$; the variance is given by the appropriate application of (3.2). In the following section, we compare these theoretical results for both the MLEs and $\widehat{B}_{W,10}$ with simulated counterparts.

8.3 Fitting G_W to G_W data

We run a series of simulations to observe how the Weibull MLEs and their standard errors are affected by values for the sample size, stopping times and number of stress levels taken, and how they compare with theoretical values. As in other simulation studies, there are numerous aspects of the experiment that can vary; for example, we could change the values taken for the parameters in G_W, the number of stress levels, the overall sample size, and how we allocate this sample across the stress levels. Type II censoring also introduces a further aspect that can vary, and we can now consider how changing r affects the MLEs and their standard errors. Here, we assume that the underlying distribution of the data is Weibull with Log-linear scale-stress relationship, and set the parameter values to

$$B_P=2,\ \alpha_{WP}=8,\ \beta_{WP}=-0.02;$$

for this particular set of Weibull parameters, the true value of B_{10} is 355.9593. We run simulations for equal n_i, and for our usual set of 2, 3 and 4 stress levels; different allocations will be considered elsewhere. We first consider the effects of varying the overall sample size, keeping the number of observations censored at X_1 fixed at 50%. We then allocate 500 observations to $k = 3$ stress levels, and vary r. The results for varying n are summarised in Table 8.1, and for varying r, the results are shown in Table 8.2. We observe excellent agreement between sample and theoretical standard errors even for small sample sizes, and the MLEs are always very close to their true values. When we keep the number of censored observations at the lowest stress level fixed at 50%, and increase the sample size, the standard errors of the MLEs and $\widehat{B}_{W,10}$ decreases. Increasing the number of failures for fixed sample sizes also has the same effect.

8.4 Fitting G_W to G_B data

We also consider the effects of fitting the Weibull distribution to data from a Burr model via simulations, with this type II censoring regime. We run simulations when we assume data is from G_{BP}, with parameters

$$\tau = 3,\ a = 4,\ \alpha_B = 8,\ \beta_B = -0.02,$$

and allow n and r to vary as above. Tables 8.3 and 8.4 summarise MLEs and their sample standard errors from G_{WP} and G_{BP} for varying n and k, when we censor half of the observations at X_1, and keep the allocation of observations across stress levels equal, and Table 8.5 contains results for varying r, when we fix 500 observations at three stress levels; different arrangements of n_i will be considered elsewhere. We see the MLEs for G_{WP} centre around some fixed values as the sample size increases. The probability of fitting the incorrect distribution is generally very low, even for small sample sizes, and as in previous cases, we always under-estimate the time to which 10% of observations fail. When we examine MLEs from G_{BP}, we have no theoretical results to compare these with, but do observe some large standard errors for $\widehat{a}$. Generally, the standard errors for both $\widehat{B}_{10}$ and MLEs from Weibull and Burr decrease as r increases. The only exception is $\widehat{a}$, where we observe that standard errors decrease as we censor more observations. However, these are known problems in the estimation of this parameter in other simulation studies; in particular, large estimates can occur.

We attempt to explain some of our results in the next section, where we consider the theory behind fitting the Burr distribution, and also expectations required to compute theoretical standard errors of the Burr MLEs and $\widehat{B}_{B,10}$.

n_1, n_2	$(50, 50)$	$(250, 250)$	$(500, 500)$
r	25	125	250
$\widehat{B}_P$	2.0605	2.0100	2.0054
S	0.2000	0.0847	0.0607
T	0.1888	0.0847	0.0599
$\widehat{\beta}_{WP}$	-0.0199	-0.0200	-0.0200
S	0.0009	0.0004	0.0003
T	0.0009	0.0004	0.0003
$\widehat{\alpha}_{WP}$	7.9789	7.9967	7.9972
S	0.1437	0.0633	0.0450
T	0.1412	0.0633	0.0447
$\widehat{B}_{W,10}$	361.2443	356.8717	356.2391
S	46.5852	20.4588	14.5903
T	45.8434	20.4718	14.4731

n_1, n_2, n_3	$(50, 50, 50)$	$(300, 300, 300)$	$(500, 500, 500)$
r	25	150	250
$\widehat{B}_P$	2.0368	2.0040	2.0038
S	0.1473	0.0585	0.0454
T	0.1423	0.0582	0.0451
$\widehat{\beta}_{WP}$	-0.0199	-0.0200	-0.0200
S	0.0008	0.0003	0.0003
T	0.0008	0.0003	0.0003
$\widehat{\alpha}_{WP}$	7.9878	7.9982	7.9982
S	0.1334	0.0539	0.0422
T	0.1329	0.0543	0.0421
$\widehat{B}_{W,10}$	360.2651	356.3123	356.2975
S	41.0526	16.6021	12.8163
T	40.5591	16.5360	12.8074

n_1, n_2, n_3, n_4	$(50, 50, 50, 50)$	$(200, 200, 200, 200)$	$(500, 500, 500, 500)$
r	25	100	250
$\widehat{B}_P$	2.0216	2.0067	2.0024
S	0.1211	0.0602	0.0380
T	0.1196	0.0598	0.0379
$\widehat{\beta}_{WP}$	-0.0199	-0.0200	-0.0200
S	0.0008	0.0004	0.0003
T	0.0008	0.0004	0.0003
$\widehat{\alpha}_{WP}$	7.9882	7.9965	7.9990
S	0.1335	0.0652	0.0419
T	0.1317	0.0659	0.0417
$\widehat{B}_{W,10}$	357.7398	356.5216	356.2395
S	38.7536	19.2011	12.1228
T	38.5258	19.2427	12.1676

Table 8.1: MLEs for G_{WP} for varying k, n, when data is generated from G_{WP} with $B_P = 2$, $\alpha_{WP} = 8$, $\beta_{WP} = -0.02$. In all cases, the n_i are equal, with $\frac{r}{n_1} = \frac{1}{2}$.

n_1, n_2, n_3	$(500, 500, 500)$	$(500, 500, 500)$	$(500, 500, 500)$
r	30	50	100
$\widehat{B}_P$	2.0034	2.0036	2.0029
S	0.0476	0.0470	0.0466
T	0.0479	0.0466	0.0464
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200
S	0.0005	0.0005	0.0004
T	0.0005	0.0004	0.0004
$\widehat{\alpha}_{WP}$	7.9963	7.9962	7.9965
S	0.0872	0.0773	0.0621
T	0.0870	0.0765	0.0618
$\widehat{B}_{W,10}$	356.1449	356.0329	355.7694
S	22.2824	19.7482	15.9750
T	22.1952	19.5150	15.9769
n_1, n_2, n_3	$(500, 500, 500)$	$(500, 500, 500)$	$(500, 500, 500)$
r	200	300	400
$\widehat{B}_P$	2.0035	2.0029	2.0029
S	0.0459	0.0451	0.0426
T	0.0457	0.0444	0.0427
$\widehat{\beta}_{WP}$	-0.0200	-0.0200	-0.0200
S	0.0003	0.0003	0.0002
T	0.0003	0.0002	0.0002
$\widehat{\alpha}_{WP}$	7.9988	7.9986	7.9996
S	0.0462	0.0392	0.0340
T	0.0467	0.0385	0.0336
$\widehat{B}_{W,10}$	356.3603	356.1944	356.4624
S	13.2901	12.7648	12.3347
T	13.3275	12.5191	12.2979

Table 8.2: MLEs for G_{WP} for varying r, when data is generated from G_{WP} with $k = 3$, $B_P = 2$, $\alpha_{WP} = 8$, $\beta_{WP} = -0.02$. In all cases, $n = 1500$ and $n_i = 500$ for all $1 \leq i \leq 3$, but $\frac{r}{n_1}$ varies.

n_1, n_2	(50, 50)	(250, 250)	(500, 500)
r	25	125	250
$\widehat{B}_P$	2.7039	2.6464	2.6373
(st.dev.)	0.2799	0.1258	0.0887
$\widehat{\beta}_{WP}$	-0.0198	-0.0199	-0.0199
(st.dev.)	0.0006	0.0003	0.0002
$\widehat{\alpha}_{WP}$	7.5691	7.5807	7.5818
(st.dev.)	0.1027	0.0461	0.0326
$\widehat{B}_{W,10}$	311.9977	309.9339	309.4228
(st.dev.)	30.1728	13.6444	9.6561
Pr (Fit Weibull)	0.0001	0	0
$\widehat{\tau}$	3.2087	3.0184	3.0082
(st.dev.)	0.4453	0.1936	0.1382
$\widehat{a}$	10.4641	6.7503	4.9014
(st.dev.)	46.5110	37.3109	5.5601
$\widehat{\alpha}_B$	8.0023	8.0585	8.0305
(st.dev.)	0.4128	0.2617	0.1754
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200
(st.dev.)	0.0006	0.0003	0.0002
$\widehat{B}_{B,10}$	334.8402	328.2800	327.8940
(st.dev.)	32.0946	14.3817	10.2301

n_1, n_2, n_3	(50, 50, 50)	(300, 300, 300)	(500, 500, 500)
r	25	150	250
$\widehat{B}_P$	2.6348	2.6023	2.5990
(st.dev.)	0.2059	0.0870	0.0673
$\widehat{\beta}_{WP}$	-0.0199	-0.0199	-0.0199
(st.dev.)	0.0007	0.0003	0.0002
$\widehat{\alpha}_{WP}$	7.5813	7.5878	7.5890
(st.dev.)	0.1012	0.0411	0.0317
$\widehat{B}_{W,10}$	308.7775	307.2931	307.2477
(st.dev.)	26.3459	10.9509	8.4241
Pr (Fit Weibull)	0.0001	0	0
$\widehat{\tau}$	3.1042	3.0044	3.0035
(st.dev.)	0.3424	0.1417	0.1078
$\widehat{a}$	10.1444	4.7589	4.3298
(st.dev.)	50.2562	3.6225	1.3888
$\widehat{\alpha}_B$	8.0482	8.0296	8.0155
(st.dev.)	0.3837	0.1678	0.1187
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200
(st.dev.)	0.0006	0.0003	0.0002
$\widehat{B}_{B,10}$	331.0552	327.6073	327.7757
(st.dev.)	28.2535	11.7132	8.9756

Table 8.3: MLEs for G_{WP} and G_{BP} for $k = 2, 3$ and varying n, when data is generated from G_{BP} with $\tau = 3$, $a = 4$, $\alpha_B = 8$, $\beta_B = -0.02$. In all cases, the n_i are equal, with $\frac{r}{n_1} = \frac{1}{2}$.

n_i	$(50, 50, 50, 50)$	$(200, 200, 200, 200)$	$(500, 500, 500, 500)$
r	25	100	250
$\widehat{B}_P$	2.6170	2.5908	2.5853
(st.dev.)	0.1749	0.0882	0.0573
$\widehat{\beta}_{WP}$	-0.0199	-0.0199	-0.0199
(st.dev.)	0.0006	0.0003	0.0002
$\widehat{\alpha}_{WP}$	7.5824	7.5897	7.5903
(st.dev.)	0.1003	0.0496	0.0313
$\widehat{B}_{W,10}$	307.4827	306.5038	306.1050
(st.dev.)	25.1351	12.5559	8.0394
Pr (Fit Weibull)	0.0001	0	0
$\widehat{\tau}$	3.0696	3.0056	3.0021
(st.dev.)	0.2947	0.1475	0.0926
$\widehat{a}$	9.4972	4.8168	4.2203
(st.dev.)	50.1835	4.4667	1.0492
$\widehat{\alpha}_B$	8.0589	8.0306	8.0107
(st.dev.)	0.3580	0.1763	0.0994
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200
(st.dev.)	0.0006	0.0003	0.0002
$\widehat{B}_{B,10}$	329.7360	327.8151	327.6925
(st.dev.)	26.9522	13.4652	8.5312

Table 8.4: MLEs for G_{WP} and G_{BP} for $k = 4$ and varying n, when data is generated from G_{BP} with $\tau = 3$, $a = 4$, $\alpha_B = 8$, $\beta_B = -0.02$. In all cases, the n_i are equal, with $\frac{r}{n_1} = \frac{1}{2}$.

n_1, n_2, n_3	$(500, 500, 500)$	$(500, 500, 500)$	$(500, 500, 500)$
r	30	50	100
$\widehat{B}_P$	2.5853	2.5830	2.5843
(st.dev.)	0.0678	0.0696	0.0693
$\widehat{\beta}_{WP}$	-0.0203	-0.0203	-0.0201
(st.dev.)	0.0004	0.0004	0.0003
$\widehat{\alpha}_{WP}$	7.6559	7.6518	7.6300
(st.dev.)	0.0693	0.0609	0.0472
$\widehat{B}_{W,10}$	321.1359	319.8272	314.8895
(st.dev.)	15.5133	13.2264	10.3852
Pr (Fit Weibull)	0	0	0
$\widehat{\tau}$	3.0039	3.0027	3.0038
(st.dev.)	0.1181	0.1179	0.1158
$\widehat{a}$	4.3830	4.3816	4.3735
(st.dev.)	1.7552	1.5334	2.1715
$\widehat{\alpha}_B$	8.0153	8.0173	8.0160
(st.dev.)	0.1540	0.1467	0.1369
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200
(st.dev.)	0.0004	0.0003	0.0003
$\widehat{B}_{B,10}$	327.4704	327.5877	327.6400
(st.dev.)	15.0531	12.9708	10.4995
n_1, n_2, n_3	$(500, 500, 500)$	$(500, 500, 500)$	$(500, 500, 500)$
r	200	300	400
$\widehat{B}_P$	2.5938	2.6036	2.6048
(st.dev.)	0.0679	0.0672	0.0640
$\widehat{\beta}_{WP}$	-0.0200	-0.0199	-0.0199
	0.0002	0.0002	0.0002
$\widehat{\alpha}_{WP}$	7.5987	7.5830	7.5809
(st.dev.)	0.0353	0.0292	0.0262
$\widehat{B}_{W,10}$	308.8331	306.4322	306.1270
(st.dev.)	8.6992	8.4815	8.4966
Pr (Fit Weibull)	0	0	0
$\widehat{\tau}$	3.0042	3.0035	3.0023
(st.dev.)	0.1108	0.1068	0.1063
$\widehat{a}$	4.3417	4.3320	4.3502
(st.dev.)	1.4933	1.3372	1.5261
$\widehat{\alpha}_B$	8.0155	8.0160	8.0173
(st.dev.)	0.1240	0.1166	0.1178
$\widehat{\beta}_B$	-0.0200	-0.0200	-0.0200
(st.dev.)	0.0002	0.0002	0.0002
$\widehat{B}_{B,10}$	327.8037	327.8103	327.8163
(st.dev.)	9.2460	8.9400	8.4966

Table 8.5: MLEs for G_{WP} and G_{BP} for varying r, when data is generated from G_{BP} with $k = 3$, $\tau = 3$, $a = 4$, $\alpha_B = 8$, $\beta_B = -0.02$. In all cases, $n = 1500$ and $n_i = 500$ for all $1 \leq i \leq 3$, but $\frac{r}{n_1}$ varies.

8.5 The Burr distribution

We finally consider our type II censoring regime when the underlying distribution is Burr. We generalise, and derive the theory for both Log-linear and Arrhenius scale-stress relationships using our usual notation. The experiment is terminated after the r^{th} failure at the lowest stress level. The random stopping time $Z = Y_{(r:n_1,1)}$ here has pdf given by

$$\frac{n_1!}{(r-1)!\,(n_1-r)!}\left[1-\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{-a}\right]^{r-1}\left[\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{-a}\right]^{n_1-r}$$
$$\times\frac{a\tau z^{\tau-1}}{\phi_{1*}^{\tau}}\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{-a-1}.$$

We use the Binomial theorem to expand

$$\left[1-\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{-a}\right]^{r-1}$$

as

$$\sum_{j=0}^{r-1}(-1)^{r-1-j}\binom{r-1}{j}\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{-a(r-1-j)},$$

so the pdf for this random variable becomes

$$\frac{a\tau n_1!}{\phi_{1*}^{\tau}\,(r-1)!\,(n_1-r)!}\sum_{j=0}^{r-1}(-1)^{r-1-j}\binom{r-1}{j}z^{\tau-1}\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{-a(n_1-j)-1}. \tag{8.7}$$

We split the likelihood into the product of two terms. The first is the likelihood for observations at the first stress level, which relies on the ordering of the data, and the second involves the likelihood at the remaining stress levels, which depends on the random stopping time Z. Before doing so, we introduce some notation to simplify such functions, and write

$$\begin{aligned}
F_{p,f} &= \sum_{j=1}^{N}\ln\left[1+Y^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_p\right)\right\}\right],\\
F_{p,c} &= \left(n_p-N\right)\ln\left[1+Z^{\tau}\exp\left(-\tau\alpha_B\right)\exp\left\{-\tau\beta_B\rho\left(X_p\right)\right\}\right],
\end{aligned}$$

where, if $p = 1$, $Y = Y_{(j:n_1,1)}$, $N = r$, and $p = i \geq 2$ implies that $Y = Y_{i,j}$, $N = M_i$. Derivatives of these functions are given by

$$\begin{aligned}
F_{p,f}^{(1,0,0)} &= \frac{\partial F_{p,f}}{\partial \tau} = \sum_{j=1}^{N} \frac{\left(\frac{Y}{\phi_{p*}}\right)^{\tau} \ln\left(\frac{Y}{\phi_{p*}}\right)}{1+\left(\frac{Y}{\phi_{p*}}\right)^{\tau}}, \\
F_{p,c}^{(1,0,0)} &= \frac{\partial F_{p,c}}{\partial \tau} = \frac{(n_p - N)\left(\frac{Z}{\phi_{p*}}\right)^{\tau} \ln\left(\frac{Z}{\phi_{p*}}\right)}{1+\left(\frac{Z}{\phi_{p*}}\right)^{\tau}}, \\
F_{p,f}^{(0,1,0)} &= \frac{\partial F_{p,f}}{\partial \alpha_B} = -\tau \sum_{j=1}^{N} \frac{\left(\frac{Y}{\phi_{p*}}\right)^{\tau}}{1+\left(\frac{Y}{\phi_{p*}}\right)^{\tau}}, \\
F_{p,c}^{(0,1,0)} &= \frac{\partial F_{p,c}}{\partial \alpha_B} = \frac{-\tau (n_p - N)\left(\frac{Z}{\phi_{p*}}\right)^{\tau}}{1+\left(\frac{Z}{\phi_{p*}}\right)^{\tau}}, \\
F_{p,f}^{(0,0,1)} &= \frac{\partial F_{p,f}}{\partial \beta_B} = -\tau \sum_{i=1}^{N} \frac{\rho(X_p)\left(\frac{Y}{\phi_{p*}}\right)^{\tau}}{1+\left(\frac{Y}{\phi_{p*}}\right)^{\tau}}, \\
F_{p,c}^{(0,0,1)} &= \frac{\partial F_{p,c}}{\partial \beta_B} = \frac{-\tau (n_p - N)\rho(X_p)\left(\frac{Z}{\phi_{p*}}\right)^{\tau}}{1+\left(\frac{Z}{\phi_{p*}}\right)^{\tau}}.
\end{aligned}$$

At X_1, the likelihood, without loss of generality, is given by

$$\prod_{j=1}^{r} \frac{a\tau Y_{(j:n_1,1)}^{\tau-1}}{\phi_{1*}^{\tau}} \left\{1+\left(\frac{Y_{(j:n_1,1)}}{\phi_{1*}}\right)^{\tau}\right\}^{-a-1} \prod_{j=r+1}^{n_1} \left\{1+\left(\frac{Z}{\phi_{1*}}\right)^{\tau}\right\}^{-a},$$

with log-likelihood

$$r \ln a + r \ln \tau + (\tau - 1) S_{e,1} - r\tau\alpha_B - r\tau\beta_B (X_1) - (a+1) F_{1,f} - aF_{1,c}.$$

At the i^{th} stress level ($i \geq 2$), the likelihood is given by

$$\prod_{j=1}^{M_i} \frac{a\tau Y_{i,j}^{\tau-1}}{\phi_{i*}^{\tau}} \left\{1+\left(\frac{Y_{i,j}}{\phi_{i*}}\right)^{\tau}\right\}^{-a-1} \prod_{j=1}^{n_i-M_i} \left\{1+\left(\frac{Z}{\phi_{i*}}\right)^{\tau}\right\}^{-a},$$

with log-likelihood

$$M_i \ln a + M_i \ln \tau + (\tau - 1) S_{e,i} - M_i\tau\alpha_B - M_i\tau\beta_B\rho(X_i) - (a+1) F_{i,f} - aF_{i,c}.$$

The full log-likelihood is the sum of the log-likelihoods over all stress levels, and is given by

$$\begin{aligned} l_{B*} &= \sum_{i=1}^{k} M_i \left\{ \ln a + \ln \tau - \tau \alpha_B \right\} + (\tau - 1) \sum_{i=1}^{k} S_{e,i} - \tau \beta_B \sum_{i=1}^{k} M_i \rho \left(X_i \right) \\ &\quad - (a+1) \sum_{i=1}^{k} F_{i,f} - a \sum_{i=1}^{k} F_{i,c}. \end{aligned} \tag{8.8}$$

The elements of the score vector are then given by

$$\begin{aligned} \frac{\partial l_{B*}}{\partial \tau} &= \left(\tau^{-1} - \alpha_B \right) \sum_{i=1}^{k} M + \sum_{i=1}^{k} S_{e,i} - \beta_B \sum_{i=1}^{k} M_i \rho \left(X_i \right) \\ &\quad - (a+1) \sum_{i=1}^{k} F_{i,f}^{(1,0,0)} - a \sum_{i=1}^{k} F_{i,c}^{(1,0,0)}, \end{aligned}$$

$$\frac{\partial l_{B*}}{\partial a} = a^{-1} \sum_{i=1}^{k} M_i - \sum_{i=1}^{k} F_{i,f} - \sum_{i=1}^{k} F_{i,c},$$

$$\frac{\partial l_{B*}}{\partial \alpha_B} = -\tau \sum_{i=1}^{k} M_i - (a+1) \sum_{i=1}^{k} F_{i,f}^{(0,1,0)} - a \sum_{i=1}^{k} F_{i,c}^{(0,1,0)},$$

and

$$\frac{\partial l_{B*}}{\partial \beta_B} = -\tau \sum_{i=1}^{k} M_i \rho \left(X_i \right) - (a+1) \sum_{i=1}^{k} F_{i,f}^{(0,0,1)} - a \sum_{i=1}^{k} F_{i,c}^{(0,0,1)}.$$

We can equate $\frac{\partial l_{B*}}{\partial a}$ to zero to obtain

$$a = \frac{r + \sum_{i=2}^{k} M_i}{F_{1,f} + F_{1,c} + \sum_{i=2}^{k} \left\{ F_{i,f} + F_{i,c} \right\}},$$

and insert this into (8.8) to obtain a profile log-likelihood. The parameters are estimated in the usual way. We first fit the non-accelerated Burr to the censored data set, obtain appropriate starting values for τ, α_B and β_B, then fit the accelerated Burr using the profile log-likelihood. Below, we consider expectations which appear in the EFI matrix from the Burr distribution, and the progress possible in their calculation.

8.5.1 Expectations

We consider the expected values that will be required in the derivation of the EFI matrix. We first compute $E\left[Z^m\right]$. From (8.7), we have

$$E\left[Z^m\right]=\frac{a\tau n_1!}{\phi_{1*}^{\tau}\left(r-1\right)!\left(n_1-r\right)!}\sum_{j=0}^{r-1}\left(-1\right)^{r-1-j}\binom{r-1}{j}\int_0^{\infty}\frac{z^{m+\tau-1}dz}{\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{a(n_1-j)+1}};$$

with $u=\left(\frac{z}{\phi_{1*}}\right)^{\tau}$, this expectation becomes

$$\frac{an_1!\phi_{1*}^{m}}{\left(r-1\right)!\left(n_1-r\right)!}\sum_{j=0}^{r-1}\left(-1\right)^{r-1-j}\binom{r-1}{j}\int_0^{\infty}\frac{u^{\frac{m}{\tau}}}{\left(1+u\right)^{a(n_1-j)+1}}du,$$

where we recognise the integral as a form of the Beta function (1.13). Thus,

$$E\left[Z^m\right]=\frac{an_1!\phi_{1*}^{m}\Gamma\left(\frac{m}{\tau}+1\right)}{\left(r-1\right)!\left(n_1-r\right)!}\sum_{j=0}^{r-1}\frac{\left(-1\right)^{r-1-j}\binom{r-1}{j}\Gamma\left\{a\left(n_1-j\right)-\frac{m}{\tau}\right\}}{\Gamma\left\{a\left(n_1-j\right)+1\right\}}. \tag{8.9}$$

Next, we consider $E\left[M_i\right]$. The number of failures at the i^{th} stress level depends on the random stopping time Z. Thus, if we first condition on $Z=z$, then M_i will have a Binomial distribution with parameters n_i and

$$q_i\left(z\right)=\Pr\left(\text{failure at } X_i\right)=1-\left\{1+\left(\frac{z}{\phi_{i*}}\right)^{\tau}\right\}^{-a}.$$

Thus,

$$E\left[M_i\right|Z=z]=n_iq_i\left(z\right).$$

We now take expectations with respect to Z to obtain

$$E\left[M_i\right]=n_i-n_iE\left[\left\{1+\left(\frac{Z}{\phi_{i*}}\right)^{\tau}\right\}^{-a}\right],$$

and write

$$\begin{aligned}E\left[\left\{1+\left(\frac{Z}{\phi_{i*}}\right)^{\tau}\right\}^{-a}\right] &= \frac{a\tau n_1!}{\phi_{1*}^{\tau}\left(r-1\right)!\left(n_1-r\right)!}\sum_{j=0}^{r-1}\left(-1\right)^{r-1-j}\binom{r-1}{j}\\ &\quad\times\int_0^{\infty}z^{\tau-1}\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{-a(n_1-j)-1}\left\{1+\left(\frac{z}{\phi_{i*}}\right)^{\tau}\right\}^{-a}dz.\end{aligned}$$

We set $u = \left(\frac{z}{\phi_{1*}}\right)^{\tau}$, so that

$$\int_0^\infty z^{\tau-1}\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{-a(n_1-j)-1}\left\{1+\left(\frac{z}{\phi_{i*}}\right)^{\tau}\right\}^{-a} dz$$

becomes

$$\phi_{1*}^{\tau}\tau^{-1}\int_0^\infty u^0\,(1+u)^{-a(n_1-j)-1}\left\{1+\left(\frac{\phi_{1*}}{\phi_{i*}}\right)^{\tau}u\right\}^{-a} du, \tag{8.10}$$

and rewrite the integral into a form which resembles Euler's Hypergeometric Transformation, given by

$$\frac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\int_0^\infty t^{c-b-1}\,(1+t)^{a-c}\,(1+t-z)^{-a}\,dt = F_{2,1}(a,b;c;z)\,.$$

Thus, (8.10) is given by

$$\begin{aligned} & \tau^{-1}\phi_{1*}^{\tau}\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{a\tau}\int_0^\infty u^{a(n_1-j+1)+1-a(n_1-j+1)-1}\,(1+u)^{a-\{a(n_1-j+1)+1\}} \\ & \times\left\{1+u-\left(1-\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{\tau}\right)\right\}^{-a} du \\ = & \frac{\phi_{1*}^{\tau}}{\tau a\,(n_1-j+1)}\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{a\tau} F_{2,1}\left(a, a\,(n_1-j+1)\,; a\,(n_1-j+1)+1; 1-\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{\tau}\right), \end{aligned}$$

and so we have

$$\begin{aligned} E\left[\left\{1+\left(\frac{Z}{\phi_{i*}}\right)^{\tau}\right\}^{-a}\right] &= \frac{n_1!\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{a\tau}}{(r-1)!\,(n_1-r)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{n_1-j+1} \\ & \times F_{2,1}\left(a, a\,(n_1-j+1)\,; a\,(n_1-j+1)+1; 1-\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{\tau}\right). \end{aligned}$$

Thus,

$$\begin{aligned} E\,[M_i] &= n_i-\frac{n_i n_1!\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{a\tau}}{(r-1)!\,(n_1-r)!}\sum_{j=0}^{r-1}\frac{(-1)^{r-1-j}\binom{r-1}{j}}{n_1-j+1} \\ & \times F_{2,1}\left(a, a\,(n_1-j+1)\,; a\,(n_1-j+1)+1; 1-\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{\tau}\right). \end{aligned}$$

We now consider general moments, and first compute $E\left[\sum_{j=1}^{M_i} Y_{i,j}^{p}\right]$ for some arbitrary power p. This requires consideration of the distribution of the failures at X_i, for $i \geq 2$. We first condition on M_i and Z to get

$$E\left[\sum_{j=1}^{M_i} Y_{i,j}^p \middle| M_i = m_i, Z = z\right] = m_i E\left[Y_{i,j}^p \middle| Z = z\right],$$

and denote Y_i as the random variable representing a failure in $(0, z)$. This variable will have a truncated Burr pdf given by

$$\frac{a\tau y^{\tau-1}}{\phi_{i*}^{\tau} q_i(z)} \left\{1 + \left(\frac{y}{\phi_{i*}}\right)^{\tau}\right\}^{-a-1}.$$

We use this to write

$$E\left[Y_{i,j}^p \middle| Z = z\right] = \int_0^z \frac{a\tau y^{p+\tau-1}}{\phi_{i*}^{\tau} q_i(z)} \left\{1 + \left(\frac{y}{\phi_{i*}}\right)^{\tau}\right\}^{-a-1} dy,$$

and set $u = \left(\frac{y}{\phi_{i*}}\right)^{\tau}$, to write this conditional expectation as

$$\frac{a\phi_{i*}^p}{q_i(z)} \int_0^{\left(\frac{z}{\phi_{i*}}\right)^{\tau}} u^{\frac{p}{\tau}} (1+u)^{-a-1}\, du.$$

By further setting $v = \frac{u}{1+u}$, and $z_c = \frac{\left(\frac{z}{\phi_{i*}}\right)^{\tau}}{1+\left(\frac{z}{\phi_{i*}}\right)^{\tau}}$, we can write this expectation in terms of the incomplete Beta function as follows

$$\frac{a\phi_{i*}^p}{q_i(z)} B_{z_c}\left(\frac{p}{\tau} + 1, a - \frac{p}{\tau}\right),$$

which, when expressed using hypergeometric functions, is given by

$$\frac{a\phi_{i*}^p z_c^{\frac{p}{\tau}+1}}{q_i(z)\left\{\frac{p}{\tau}+1\right\}} F_{2,1}\left(\left\{\frac{p}{\tau}+1, \frac{p}{\tau}+1-a\right\}; \left\{\frac{p}{\tau}+2\right\}; z_c\right).$$

We now take expectations with respect to M_i to get

$$E\left[\sum_{j=1}^{M_i} Y_{i,j}^p \middle| Z = z\right] = \frac{n_i a\phi_{i*}^p z_c^{\frac{p}{\tau}+1}}{\frac{p}{\tau}+1} F_{2,1}\left(\left\{\frac{p}{\tau}+1, \frac{p}{\tau}+1-a\right\}; \left\{\frac{p}{\tau}+2\right\}; z_c\right),$$

and, after taking expectations with respect to Z, we have

$$E\left[\sum_{j=1}^{M_i} Y_{i,j}^p\right] = \frac{n_i a\phi_{i*}^p}{\frac{p}{\tau}+1} E\left[\sum_{m=0}^{\infty} \frac{\left(\frac{p}{\tau}+1\right)_m \left(1+\frac{p}{\tau}-a\right)_m}{m!\left(\frac{p}{\tau}+2\right)_m} \left\{\frac{\left(\frac{Z}{\phi_{i*}}\right)^{\tau}}{1+\left(\frac{Z}{\phi_{i*}}\right)^{\tau}}\right\}^{m+\frac{p}{\tau}+1}\right].$$

We obtain an expression for

$$E\left[\left\{\frac{\left(\frac{Z}{\phi_{i*}}\right)^{\tau}}{1+\left(\frac{Z}{\phi_{i*}}\right)^{\tau}}\right\}^{m+\frac{p}{\tau}+1}\right],$$

given by

$$\frac{a\tau n_1!\phi_{i*}^{-\tau\left(m+\frac{p}{\tau}+1\right)}}{\phi_{1*}^{\tau}(r-1)!(n_1-r)!}\sum_{j=0}^{r-1}(-1)^{r-1-j}\binom{r-1}{j}$$
$$\times\int_0^{\infty} z^{\tau(m+2)+p-1}\left\{1+\left(\frac{z}{\phi_{1*}}\right)^{\tau}\right\}^{-a(n_1-j)-1}\left\{1+\left(\frac{z}{\phi_{i*}}\right)^{\tau}\right\}^{-\left(m+\frac{p}{\tau}+1\right)}dz,$$

and if we set $u=\left(\frac{z}{\phi_{1*}}\right)^{\tau}$, then this expectation becomes

$$\frac{an_1!}{(r-1)!(n_1-r)!}\sum_{j=0}^{r-1}(-1)^{r-1-j}\binom{r-1}{j}$$
$$\times\int_0^{\infty} u^{m+\frac{p}{\tau}+1}(1+u)^{-a(n_1-j)-1}\left[1+u-\left\{1-\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{\tau}\right\}\right]^{-\left(m+\frac{p}{\tau}+1\right)}du,$$

the integral in which we recognise Euler's Hypergeometric Transformation. Therefore, the expected value becomes

$$\frac{an_1!}{(r-1)!(n_1-r)!}\sum_{j=0}^{r-1}(-1)^{r-1-j}\binom{r-1}{j}\frac{\Gamma\{a(n_1-j)\}\Gamma\left\{m+\frac{p}{\tau}+2\right\}}{\Gamma\left\{a(n_1-j)+m+\frac{p}{\tau}+2\right\}}$$
$$\times F_{2,1}\left(\begin{array}{c}\left\{m+\frac{p}{\tau}+1,a(n_1-j)\right\};\\ \left\{a(n_1-j)+m+\frac{p}{\tau}+2\right\};1-\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{\tau}\end{array}\right),$$

and so

$$E\left[\sum_{j=1}^{M_i}Y_{i,j}^{p}\right]$$

is equal to

$$\frac{a^2\phi_{i*}^{p}n_in_1!}{\left(\frac{p}{\tau}+1\right)(r-1)!(n_1-r)!}\sum_{j=0}^{r-1}(-1)^{r-1-j}\binom{r-1}{j}$$
$$\times\sum_{m=0}^{\infty}\frac{\left(\frac{p}{\tau}+1\right)_m\left(\frac{p}{\tau}+1-a\right)_m\Gamma\{a(n_1-j)\}\Gamma\left(m+\frac{p}{\tau}+2\right)}{m!\left(\frac{p}{\tau}+2\right)_m\Gamma\left\{a(n_1-j)+m+\frac{p}{\tau}+2\right\}}$$
$$\times F_{2,1}\left(\left\{m+\frac{p}{\tau}+1,a(n_1-j)\right\};\left\{a(n_1-j)+m+\frac{p}{\tau}+2\right\};1-\left(\frac{\phi_{i*}}{\phi_{1*}}\right)^{\tau}\right).$$

Further simplification of this expectation does not seem possible, and, because of its form, numerical results are currently not available from Mathematica. Thus, we omit any further results on expectations from the Burr distribution, since most will involve derivatives of functions, like the one above. We note, but do not consider here, the possibility of truncating the infinite summation in this expectation; such an approach is different from those previously outlined in this thesis. Thus, there is considerable scope for further research in this area.

8.6 Summary

This chapter outlined results on fitting the accelerated Weibull and Burr distributions to data that had undergone a type II censoring regime. We assumed a fixed number of failures at X_1, which in turn induced a type I censoring regime, now with random stopping times, at the higher stress levels. We derived the EFI matrix for G_{W*}, under the assumption that no mis-specification had taken place, and verified these results with simulations. We also included a brief summary of the effects of fitting G_{WP} and G_{BP} to data from G_{BP}, via simulations. A similar investigation was attempted for G_{B*}, but due to the complicated form of expectations that made up elements of the EFI matrix, this matrix was omitted.

We conclude our work in the final chapter by briefly summarising previous chapters. We end with a discussion on areas of future research that have arisen as a result of our work.

Chapter 9

Summary

In this final chapter, we summarise our work in previous chapters, outline the practical implications of our work, and end with a discussion on areas of future research. We begin by listing each chapter in turn.

9.1 Chapter summary

Chapter 1

In Chapter 1, we outlined the main focus of this thesis, and discussed the wide use of the Weibull distribution to model reliability data. We then summarised the four reliability distributions used in our work, namely the Weibull, Burr, Gamma and Lognormal, and included details of their pdfs, cdfs, hazard, cumulative hazard and quantile functions. We continued by considering various censoring regimes, stressing that type I and II censoring would be used throughout our work, and also discussed various mathematical functions required, mainly as a result of censoring. We concluded with a summary of sample procedures used to study data sets, such as the Kolmogorov-Smirnov distance, and sample hazard and cumulative hazard functions.

Chapter 2

We began Chapter 2 by outlining results for fitting Weibull, Burr, Gamma and Lognormal distributions to complete data sets via ML, and illustrated our techniques with the ball bearings data set given by Table 1.1. For the Burr distribution, we discussed various problems faced when fitting this model to data, and outlined the limiting link between this distribution and the Weibull. We then considered a practical approach to illustrate the effects of fitting the Weibull distribution to data sets with other underlying models, and examined the consequences of fitting this mis-specified distribution to data from a Burr, Gamma and Lognormal distribution. We used simulation studies to illustrate how the MLEs from the Weibull distribution varied as the sample size of the data changed, and used sample procedures discussed in Chapter 1 to assess the goodness of fit that the true and mis-specified distributions provided to the simulated data sets. This involved two different approaches; firstly we examined the difference between the ecdf from the simulated data, and the cdf of the underlying distribution, when MLEs were used as parameter values. The second approach was similar, but involved comparing sample and theoretical cumulative hazard functions. The two approaches were compared for true and mis-specified models for varying sample sizes.

Chapter 3

Chapter 3 outlined the theory necessary to explain our simulations in Chapter 2. We first discussed the computation of the EFI matrices for Weibull, Burr, Gamma and Lognormal distributions, when we assumed no mis-specification had taken place. Next, we considered explaining our simulated values for the mis-specified Weibull distribution. We used the entropy function to derive theoretical counterparts to the MLEs from the Weibull model, and used these to obtain the asymptotic variance covariance matrix of the mis-specified MLEs. We did this for a general underlying true distribution, summarising our results in terms of expectations from this model, and then gave details for Burr, Gamma and Lognormal distributions. We also included an investigation into how entropy values for the Weibull distribution varied with parameters from the true model. We concluded with a summary of how the true and mis-specified models compared for varying parameter values from the true distribution, and computed maximum absolute distances between true and mis-specified cdfs. We also considered how the quantile B_{10} varied between true and mis-specified models, and tabulated figures for this quantile for both distributions, together with the relative error, for varying true parameter values.

Chapter 4

Chapter 4 extended our results for complete data to allow for censoring to take place. However, unlike Chapters 2 and 3, where we allowed for data to be generated from a variety of true distributions, we only considered the Burr model in this case. We first considered type I censoring, and outlined the theory required to fit Weibull and Burr distributions to data sets that had undergone a type I censoring regime. As in Chapter 2, we examined the effects of fitting Weibull to Burr via simulations, this time varying the stopping time as well as the sample size. We continued by deriving the theory necessary to explain these simulated values, first listing results for the EFI matrices for Weibull and Burr models when no mis-specification had taken place. We extended these results to allow for mis-specifying the Weibull distribution, and obtained entropy values and the variance covariance matrix of the mis-specified MLEs. A similar investigation was constructed for type II censored data sets, where now we had the added complication of obtaining expectations of order statistics and sums of order statistics in order to compute the EFI matrices from Weibull and Burr. When we began to derive the variance covariance matrix of the mis-specified MLEs, we found we had to compute expectations of joint order statistics and sums of expectations of joint order statistics. We constructed a general result that linked the expected value of the i^{th} and j^{th} order statistic with the first and some other, and also derived a series of new results on expectations of joint order statistics from the Weibull and Burr distributions. Weibull expectations were slightly simpler to compute, and so we could compare theoretical values with simulated counterparts for large sample sizes. However, due to the complicated structure of the Burr expectations, obtaining theoretical results were limited to only very small sample sizes. Consequently, the theoretical variance covariance matrix of the mis-specified Weibull MLEs was omitted.

Chapter 5

This chapter introduced the concept of acceleration, and listed the main relationships linking the scale parameter of a distribution to a stress level. Throughout our work, we considered the two most popular relationships - the Log-linear and Arrhenius models. We outlined the theory required to fit Weibull, Burr, Gamma and Lognormal distributions to accelerated complete data sets, for either Log-linear or Arrhenius models, and also derived the EFI matrix. When examining the Weibull distribution, we used simulations to assess how well theoretical results matched up with sample values, and how the Weibull MLEs and their standard errors changed, as we varied the sample size, the number of stress levels, and the allocation of the sample across stress levels.

Chapter 6

Chapter 6 extended Chapter 5 to deal with mis-specification. Here, we allowed for the possibility of not only mis-specifying the Weibull distribution, but also the scale-stress relationship. We discussed three forms of mis-specification; in particular, mis-specifying the scale-stress relationship, mis-specifying the distribution function, and mis-specifying both the distribution and scale-stress relationship. We outlined the theory necessary to compute entropy values from the Weibull distribution, and to compute the asymptotic variance covariance matrix of the mis-specified MLEs for the three scenarios.

Chapter 7

Chapter 7 used the theory developed in Chapter 6 to assess the effects of mis-specification for the scenarios considered there. We did this via two approaches; firstly we used extensive sets of simulations, whereby we simulated data from the true model with varying parameter values, stress levels and sample sizes, and then fitted both the mis-specified and true distributions. We compared our sample results with theoretical values, and also compared estimates of B_{10} from true and mis-specified models. The second approach was more theoretical, and compared the cdfs of true and mis-specified distributions across all stress levels, where entropy values were used for the mis-specified model. As in our work in Chapter 3, we computed the maximum absolute distance between both cdfs across all stress levels, to deduce where the worst fit between true and mis-specified distributions occurred. We used both approaches for the three scenarios of mis-specification listed in Chapter 6. Thus, we had two sections for mis-specifying the scale-stress relationship, where we kept the Weibull distribution fixed as the true model, and mis-specified either the Log-linear or Arrhenius relationship; we had three sections for mis-specifying the distribution function, and examined the effects of mis-specifying the Weibull distribution to data from a Burr, Gamma or Lognormal distribution, when the Log-linear model was used as the scale-stress relationship. Finally, we had 6 scenarios for mis-specifying both the distribution and scale-stress relationship. In each case, we used the two approaches to assess the effects of mis-specification.

Chapter 8

In this chapter, we considered one combination of censoring and acceleration. We considered likelihood and the derivation of MLEs for Weibull and Burr distributions when we assumed that the data was subjected to a type II censoring regime at the lowest stress level. This type of censoring regime induced a type I process at the higher stress levels, where the stopping time was now random. We computed the EFI matrix for the Weibull MLEs, and attempted a similar investigation for the Burr. However, as with our work on non-accelerated data sets, some expectations for the Burr for this censoring regime were very complicated, and only limited progress with the computation of numerical values was possible. Therefore, complete results are omitted, and the chapter concluded with a brief introduction into mis-specifying the Weibull distribution, where we considered effects using simulation studies.

9.2 Critical Appraisal

Throughout this thesis, our primary focus has been on the development of theoretical results, and the validation of these through extensive simulation experiments. However, given the practical basis of our investigations, we are also able to interpret these results, and this section summarises some of the main practical consequences of our work.

We first review the question of mis-specification: Due to its wide use, the problem of mis-specifying models based around the Weibull distribution is, we believe, an important aspect in reliability analysis. In practice, the use of model identification techniques means that we are unlikely to mis-specify the Weibull distribution if the underlying data had no similarities with observations from this model (for example, if the data arose from the Lognormal distribution with a large value of σ, for which the underlying distribution looks nothing like any Weibull distribution). The simulation experiments reported in this thesis are intended to reflect this: we have specified parameter values in the true model precisely in order to generate data that could well be regarded as arising from a Weibull distribution. Intuitively, the penalty paid for mis-specifying a model is smaller when the true distribution is closer to some Weibull distribution, and our results allow us to quantify this penalty. In particular, we assessed how well the Weibull model represented this data in several ways. These included a practical approach, whereby we used simulation studies and Kolmogorov Smirnov distances to assess how often we would prefer the mis-specified model in favour of the true. A more theoretical approach of comparing true and mis-specified quantiles was also considered. In particular, an examination of B_{10} from true and mis-specified distributions was carried out, and on numerous occasions we saw this quantile continuously being underestimated if the mis-specified Weibull model was chosen in favour of the true. Such results have potentially important practical implications in the determination of warranty periods for items; for instance, if the distribution is mis-specified, then the corresponding warranty period sees rather fewer failures than might be expected.

These methods were extended to deal with accelerated data sets. The additional relationship between stress level with scale further complicates mis-specification, where we must now allow for the possibility of not only mis-specifying the Weibull distribution, but also the scale-stress relationship. Here, we emphasise that practical model identification techniques for the scale-stress relationship will be based on very few sample points (perhaps as few as two, but rarely more than five), and, unless there is considerable data at each stress, there is likely to be considerable deviation from any true relationship, making mis-specification even more likely. We assessed the possibility of choosing the mis-specified Weibull distribution in favour of the true using similar techniques to the non-accelerated scenario. As before, we considered the estimation of B_{10} (at design stress) using both true and mis-specified distributions. However, in this instance, we allowed not only for mis-specifying the underlying distribution, but also the scale-stress relationship and a combination of both. Again, as with the non-accelerated scenario, we observed, on numerous occasions, this quantile being underestimated, which, as before, has considerable practical implications.

Our results for the accelerated case feature the allocation of items across stress levels, so that we are in a position to consider the effect of varying this allocation. The approach adopted here is reasonably straightforward, and reflects the underlying assumption of a fixed capacity in testing facilities. Thus, for a given number of items, we assume that there is a minimum number (or batch size) to be tested at any stress level, and that the number tested at other levels will be a multiple of this minimum batch size. When the stressing factor is temperature, we can start with temperature at its minimum value for the first batch, and then raise it in suitable steps for subsequent batches; alternatively, we can set temperature to its maximum for the first batch, and control its cooling for subsequent batches. Although a detailed discussion of such practical issues is beyond the scope of this thesis, our work provides the fundamental results needed to assess the relative efficiency of such plans in some generality. Across the rather restricted range of plans we have considered in this thesis, we have generally observed higher standard errors for B_{10} (at design stress) when we allocated higher proportions of items to higher stress levels. At the same time, the estimated value of B_{10} (at design stress) for the mis-specified model was usually closer to the true value, so practically, a compromise of having large variation but with small bias can be attained.

Finally, we remark that the censoring regime considered in Chapter 8 is again governed by some practical considerations, in that it guarantees a given number of failures at one (the lowest) stress, which then generally tends to mean higher proportions of failures at higher stresses. The practical consequences of our results are as follows: with heavy censoring at the lowest stress level, we see (as expected, given the relative lack of information) larger standard errors for B_{10}, although the mis-specified estimate of the quantile was relatively close to the true value. As the censoring decreased at the lowest stress, we observed smaller standard errors, but larger differences between the value and mis-specified estimate of B_{10}. The trade-off can be formalised by considering a single mean square error, which can then be used as the criterion for assessing competing experimental plans. Again, our work yields the initial results needed to consider such plans, although we note that there is no straightforward extension using Type I censoring at one stress level to induce a censoring regime at others.

9.3 Areas for future research

Throughout our work, we have always taken the Weibull distribution as the mis-specified model, since it is widely used by most practitioners, and, in some circumstances, used when other distributions provide an improved fit. An obvious area for future research would be to mis-specify other reliability distributions, and so, for example, examine the effects of fitting the Gamma distribution to data from a Weibull model. This also applies to accelerated distributions where the added relationship linking scale parameter to stress level allows for further areas of mis-specification. In our work, we considered the Log-linear and Arrhenius scale-stress relationships. Chapter 5 listed other available models, so we could examine the effects of mis-specifying these also.

We considered two approaches to speed up the running time of experiments - censoring and acceleration. We could discuss other forms of censoring, like those outlined in Chapter 1, or even examine other approaches. For example, Meeker, Escobar and Lu (1998) consider accelerated degradation tests, where degradation measures are taken over time rather than recording actual failure times. Such approaches are suitable if the product tested is extremely reliable, and few or no failures occur, even with acceleration. Similarly, Watkins (2001a) briefly introduces the concept of step-stress testing, where the stress level is progressively increased if items are found not to fail. Mis-specification of the Weibull distribution could be considered for both approaches, so allowing considerable scope for research in this area.

When we examined type II censoring for non-accelerated data sets, we derived new results on joint expectations of order statistics, particularly for Weibull and Burr distributions. Numerical progress for the Burr was limited here, and as a result, no asymptotic variance covariance matrix for the mis-specified Weibull MLEs was derived. Further research in this area could centre around obtaining numerical results for these joint expectations, and functions of joint expectations, for significant sample sizes, at a reasonable speed. We were able to obtain some joint expectations, but the computational effort required to do so was significant. If this could be improved, then making inferences about the mis-specified Weibull MLEs, with type II censoring could be considered.

Finally, we examined accelerated data sets which had undergone a type II censoring regime at the lowest stress level. For large samples, Mathematica was unable to compute the EFI matrix for the Burr MLEs for such a censoring regime, again because some expectations from the Burr distribution were complicated. We noted the possibility of truncating infinite summations in these expectations, so presenting a possible area for future research here. We chose our regime as it seemed a practically sensible approach, which guaranteed a minimum number of failures. However, many other approaches could be considered.

Bibliography

Abramowitz, M. and Stegun, I.A. (1972), *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* John Wiley & Sons.

Ansell, J.I. and Phillips, M.J. (1994), *Practical Methods for Reliability Data Analysis.* Oxford University Press.

Bain, L.J. (1978), *Statistical Analysis of Reliability and Life-Testing Models.* Marcel Dekker, Inc.

Balakrishnan, N. and Rao, C.R. (1998a), *Order Statistics: Theory and Methods.* Handbook of Statistics Volume 16. Elsevier.

Balakrishnan, N. and Rao, C.R. (1998b), *Order Statistics: Applications.* Handbook of Statistics Volume 17. Elsevier.

Balasooriya, U. and Hapuarachchi, K.P. (1992), Extended Tables for the Moments of Gamma-Distribution Order Statistics. *IEEE Transactions on Reliability*, **41** : 256-271.

Burr, I.W. (1942), Cumulative Frequency Functions. *Annals of Mathematical Statistics*, **13** : 215-232.

Cain, S.R. (2002), Distinguishing Between Lognormal and Weibull Distributions. *IEEE Transactions on Reliability*, **51** : 32-38.

Caroni, C. (2002), The Correct "Ball Bearings" Data. *Lifetime Data Analysis*, **8** : 395-399.

Cheng, R.C.H. and Iles, T.C. (1990), Embedded Models in Three-Parameter Distributions and their Estimation. *Journal of the Royal Statistical Society Series B*, **52** : 135-149.

Cohen, A.C. (1965), Maximum Likelihood Estimation in the Weibull Distribution Based On Complete and On Censored Samples. *Technometrics*, **7** : 579-588.

Cox, D.R. (1961), Tests of Separate Families of Hypotheses. *Proceedings of the Fourth Berkeley Symposium*, **1** : 105-123.

Cox, D.R. and Hinkley, D.V. (1974), *Theoretical Statistics.* Chapman and Hall.

Croes, K., Manca, J.V., De Ceuninck, W., De Schepper, L. and Molenberghs, G. (1998), The Time of "Guessing" Your Failure Distribution is Over! *Microelectronics Reliability*, **38** : 1187-1191.

Crowder, M. and Kimber, A. (1997), A Score Test for the Multivariate Burr and Other Weibull Mixture Distributions. *Scandinavian Journal of Statistics*, **24** : 419-432.

Crowder, M.J., Kimber, A.C., Smith, R.L. and Sweeting, T.J. (1991), *Statistical Analysis of Reliability Data.* Chapman and Hall.

David, H.A. (1981), *Order Statistics 2^{nd} edition.* John Wiley and Sons.

Der, G. and Everitt, B.S. (2002), *A Handbook of Statistical Analyses Using SAS.* Chapman and Hall.

Dumonceaux, R. and Antle, C.E. (1973), Discrimination Between the Lognormal and Weibull Distributions, *Technometrics*, **15** : 923-926.

Dyer, D.D. and Whisenand, C.W. (1973), Best Linear Unbiased Estimator of the Parameter of the Rayleigh Distribution - Part I: Small Sample Theory for Censored Order Statistics. *IEEE Transactions on Reliability*, **22** : 27-34.

Farnum, N.R. and Booth, P. (1997), Uniqueness of Maximum Likelihood Estimators of the 2-Parameter Weibull Distribution. *IEEE Transactions on Reliability*, **46** : 523-525.

Hirose, H. (1993), Estimation of Threshold Stress in Accelerated Life-Testing. *IEEE Transactions on Reliability*, **42** : 650-657.

Hirose, H. (1998), Parameter Estimation for the 3-Parameter Gamma Distribution Using the Continuation Method. *IEEE Transactions on Reliability*, **47** : 188-196.

Hutton, J.L. and Monaghan, P.F. (2002), Choice of Parametric Accelerated Life and Proportional Hazards Models for Survival Data: Asymptotic Results. *Lifetime Data Analysis*, **8** : 375-393.

Jaynes, E.T. (1957), Information Theory and Statistical Mechanics. *Physical Review*, **106** : 620-630.

John, A.M., Johnson, R. and Watkins, A.J. (2003), On Finite Sums of Powers of Reciprocals. *International Journal of Pure and Applied Mathematics*, **7** : 7-17.

Johnson, R. (2003), *Accelerated Life Testing and the Burr XII Distribution.* PhD Thesis, University of Wales Swansea.

Kennedy, W.J. and Gentle, J.E., (1980), *Statistical Computing.* New York Publishing, Statistics: Textbooks and Monographs, **33** : 73-74.

Lawless, J.F. (1982), *Statistical Models and Methods for Lifetime data.* Wiley series in Probability and Mathematical Statistics.

Leitch, R.D. (1995), *Reliability Analysis for Engineers.* Oxford University Press.

Lieblein, J. and Zelen, M. (1956), Statistical Investigation of the Fatigue Life of Deep-Grove Ball Bearings. *Journal of Research of the National Bureau of Standards*, **57** : 273-316.

Luke, Y.L. (1969), *The Special Functions and Their Approximations.* Volume 53-I, Mathematics in Science and Engineering. Academic Press.

Mackisack, M.S. and Stillman, R.H. (1996), A Cautionary Tale About Weibull Analysis. *IEEE Transactions on Reliability*, **45** : 244-248.

Mann, N.R., Schafer, R.E. and Singpurwalla, N.D. (1974), *Methods for Statistical Analysis of Reliability and Life Data.* John Wiley & Sons.

Mardia, K.V., Kent, J.T. and Bibby, J.M. (1995), *Multivariate Analysis.* Academic Press.

Marshall, A.W., Meza, J.C. and Olkin, I. (2001), Can Data Recognise Its Parent Distribution?. *Journal of Computational and Graphical Statistics*, **10** : 555-580.

Meeker, W.Q. and Escobar, L.A. (1993), A Review of Recent Research and Current Issues in Accelerated Testing. *International Statistical Review*, **61** : 147-168.

Meeker, W.Q., Escobar, L.A. and Lu, C.J. (1998), Accelerated Degradation Tests: Modelling and Analysis. *Technometrics*, **40** : 89-99.

Menzefricke, U. (1992), Designing Accelerated Life Tests when there is Type II Censoring. *Communications in Statistics - Theory and Methods*, **21** : 2569-2589.

Montgomery, D.C. (1997), *Design and Analysis of Experiments.* John Wiley & Sons.

Nelson, W. (1982), *Applied Life Data Analysis.* John Wiley & Sons.

Nelson, W. (1990), *Accelerated Testing - Statistical Models, Test Plans, and Data Analyses.* John Wiley & Sons.

Newton, D.W (1991), Some Pitfalls in Reliability Data Analysis. *Reliability Engineering and System Safety*, **34** : 7-21.

Richards, D.O. and McDonald, J.B. (1987), A General Methodology for Determining Distributional Forms with Applications in Reliability. *Journal of Statistical Planning and Inference*, **16** : 365-376.

Shannon, C.E. (1948), A Mathematical Theory of Communication. *Bell System Technical Journal*, **27** : 379-423; 623-656.

Slater, L. (1966), *Generalised Hypergeometric Functions.* Cambridge University Press.

Tadikamalla, P.R. (1980), A Look at the Burr and Related Distributions. *International Statistical Review*, **48** : 337-344.

Tse, S.K., Yang, C. and Yuen, H.K. (2000), Statistical Analysis of Weibull Distributed Lifetime Data Under Type II Progressive Censoring with Binomial Removals. *Journal of Applied Statistics*, **27** : 1033-1043.

Tseng, S.T. and Hsu, C.H. (1994), Comparison of Type-I & Type-II Accelerated Life Tests for Selecting the Most Reliable Product. *IEEE Transactions on Reliability*, **43** : 503-510.

Watkins, A.J. (1991), On the Analysis of Accelerated Life-Testing Experiments. *IEEE Transactions on Reliability*, **40** : 98-101.

Watkins, A.J. (1997), A Note on Expected Fisher Information for the Burr XII Distribution. *Microelectronics Reliability*, **37** : 1849-1852.

Watkins, A.J. (1998), On Expectations Associated with Maximum Likelihood Estimation in the Weibull Distribution. *Journal of the Italian Statistical Society*, **7** : 15-26.

Watkins, A.J. (1999), An Algorithm for Maximum Likelihood Estimation in the Three Parameter Burr XII Distribution. *Computational Statistics & Data Analysis*, **32** : 19-27.

Watkins, A.J. (2001a), Commentary: Inference in Simple Step-Stress Models. *IEEE Transactions on Reliability*, **50** : 36-37.

Watkins, A.J. (2001b), On the Likelihood Function for the Three Parameter Burr XII Distribution. *International Journal of Reliability, Quality and Safety Engineering*, **8** : 173-188.

Watkins, A.J. and Johnson, R. (1999), Using SAS/IML to Fit Embedded Models to Accelerated Life Test Data. *Proceedings of SEUGI 17, Den Haag, The Netherlands.*

Watkins, A.J. and Johnson, R. (2002), Two Results on Hypergeometric Functions with Applications in Reliability Analysis. *International Journal of Pure and Applied Mathematics*, **3** : 71-89.

Weibull, W. (1951), A Statistical Distribution Function of Wide Applicability. *Journal of Applied Mechanics*, **18** : 293-297.

White, H. (1982), Maximum Likelihood Estimation of Misspecified Models. *Econometrica*, **50** : 1-25.

Wingo, D.R. (1983), Maximum Likelihood Methods for Fitting the Burr Type XII Distribution to Life Test Data. *Biometrical Journal*, **25** : 77-84.

Wingo, D.,R. (1993), Maximum Likelihood Estimation of Burr XII Distribution Parameters Under Type II Censoring. *Microelectronics Reliability*, **33** : 1251-1257.

Wolfram, S. (1988), *Mathematica: A System for Doing Mathematics by Computer.* Addison-Wesley.

Appendix A : Expectations of single order statistics

In this appendix, we prove our result linking the expected value of the r^{th} order statistic with the expected value of the first, given by (4.34), using induction. We first show that (4.34) holds for $r = 1$:

$$\begin{aligned} E\left[h\left(Y_{(1:n)}\right)\right] &= 1\binom{n}{1}\sum_{i=0}^{0}\frac{(-1)^{1-1-i}\binom{1-1}{i}}{n-i}E\left[h\left(Y_{(1:n-i)}\right)\right] \\ &= \frac{n!}{(n-1)!}\left(\frac{1}{n}E\left[h\left(Y_{(1:n)}\right)\right]\right) \\ &= E\left[h\left(Y_{(1:n)}\right)\right]. \end{aligned}$$

Next, we assume that the identity holds for $1 \leq r \leq n$. We must show that the same is true for $1 \leq r+1 \leq n$. Using (4.33), we write

$$\begin{aligned} rE\left[h\left(Y_{(r+1:n)}\right)\right] &= nE\left[h\left(Y_{(r:n-1)}\right)\right] - (n-r)E\left[h\left(Y_{(r:n)}\right)\right] \\ &= nr\binom{n-1}{r}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{n-1-i}E\left[h\left(Y_{(1:n-1-i)}\right)\right] - \\ &\quad (n-r)\,r\binom{n}{r}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{n-i}E\left[h\left(Y_{(1:n-i)}\right)\right] \\ &= \frac{n!}{(r-1)!\,(n-(r+1))!}\left(\begin{array}{c} \sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{n-1-i}E\left[h\left(Y_{(1:n-1-i)}\right)\right] \\ -\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{n-i}E\left[h\left(Y_{(1:n-i)}\right)\right] \end{array}\right). \end{aligned}$$

Therefore

$$\begin{aligned} E\left[h\left(Y_{(r+1:n)}\right)\right] &= \frac{n!}{r!\,(n-(r+1))!} \\ &\quad \times\left(\begin{array}{c} \frac{(-1)^r}{n}E\left[h\left(Y_{(1:n)}\right)\right] + \frac{(-1)^{r-1}}{n-1}\left(\binom{r-1}{0}+\binom{r-1}{1}\right)E\left[h\left(Y_{(1:n-1)}\right)\right] \\ +\frac{(-1)^{r-2}}{n-2}\left(\binom{r-1}{1}+\binom{r-1}{2}\right)E\left[h\left(Y_{(1:n-2)}\right)\right] \\ +\ldots+\frac{(-1)}{n-r+1}\left(\binom{r-1}{r-2}+\binom{r-1}{r-1}\right)E\left[h\left(Y_{(1:n-r+1)}\right)\right] + \\ \frac{1}{n-r}E\left[h\left(Y_{(1:n-r)}\right)\right] \end{array}\right). \end{aligned}$$

Now, we make use of the fact that

$$\binom{r-1}{i} + \binom{r-1}{i-1} = \binom{r}{i}$$

to obtain

$$E\left[h\left(Y_{(r+1:n)}\right)\right] = (r+1)\binom{n}{r+1}\sum_{i=0}^{r}\frac{(-1)^{r-i}\binom{r}{i}}{n-i}E\left[h\left(Y_{(1:n-i)}\right)\right],$$

as required. Hence by Mathematical Induction, (4.34) holds for all $1 \leq r \leq n$.

We also prove, by induction, our expression for the sum of the first r order statistics, given by (4.35). For the Basis, we let $r = 1$ and see that

$$\begin{aligned} E\left[h\left(Y_{(1:n)}\right)\right] &= \frac{n(n-1)}{n(n-1)}E\left[h\left(Y_{(1:n)}\right)\right] \\ &= E\left[h\left(Y_{(1:n)}\right)\right]. \end{aligned}$$

Next we assume that (4.35) holds for r. For the sum of the first $r+1$ expectations we have

$$\sum_{i=1}^{r+1} E\left[h\left(Y_{(i:n)}\right)\right] = \sum_{i=1}^{r} E\left[h\left(Y_{(i:n)}\right)\right] + E\left[h\left(Y_{(r+1:n)}\right)\right],$$

and to obtain an expression for $E\left[h\left(Y_{(r+1:n)}\right)\right]$, we use (4.34). Thus,

$$\sum_{i=1}^{r+1} E\left[f\left(X_{(i:n)}\right)\right]$$

equates to

$$\begin{aligned} &(n-r+1)(n-r)\binom{n}{r-1}\sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i-1)(n-i)}E\left[h\left(Y_{(1:n-i)}\right)\right] + \\ &(r+1)\binom{n}{r+1}\sum_{i=0}^{r}\frac{(-1)^{r-1-i}\binom{r}{i}}{n-i}E\left[h\left(Y_{(1:n-i)}\right)\right] \\ = \; & \frac{n!}{(r-1)!(n-r-1)!}\left[\begin{array}{c} \sum_{i=0}^{r-1}\frac{(-1)^{r-1-i}\binom{r-1}{i}}{(n-i-1)(n-i)}E\left[h\left(Y_{(1:n-i)}\right)\right] \\ +\frac{1}{r}\sum_{i=0}^{r}\frac{(-1)^{r-i}\binom{r}{i}}{n-i}E\left[h\left(Y_{(1:n-i)}\right)\right] \end{array}\right] \\ = \; & \frac{n!}{(r-1)!(n-r-1)!}\left[\begin{array}{c} \sum_{i=0}^{r-1}\frac{(-1)^{r-i}}{n-i}E\left[h\left(Y_{(1:n-i)}\right)\right]\left\{\frac{\binom{r}{i}}{r} - \frac{\binom{r-1}{i}}{n-i-1}\right\} \\ +\frac{1}{r(n-r)}E\left[h\left(Y_{(1:n-r)}\right)\right] \end{array}\right]. \end{aligned}$$

We simplify

$$\frac{\binom{r}{i}}{r} - \frac{\binom{r-1}{i}}{n-i-1}$$

by writing

$$\frac{\binom{r}{i}(n-r-1)}{r(n-i-1)},$$

so the expected value becomes

$$\begin{aligned} &\frac{n!}{r!\,(n-r-1)!}\left[\begin{array}{c}\sum_{i=0}^{r-1}\frac{(-1)^{r-i}\binom{r}{i}(n-r-1)}{(n-i-1)(n-i)}E\left[h\left(Y_{(1:n-i)}\right)\right]\\ +\frac{1}{n-r}E\left[h\left(Y_{1:n-r}\right)\right]\end{array}\right] \\ = \; &(n-r)(n-r-1)\binom{n}{r}\sum_{i=0}^{r}\frac{(-1)^{r-i}\binom{r}{i}}{(n-i-1)(n-i)}E\left[h\left(Y_{(1:n-i)}\right)\right], \end{aligned}$$

as required.

Appendix B : Sums of reciprocals and their powers

We present proofs for (4.38) and (4.39). These were used in our work on type II censoring to simplify sums of reciprocals with powers. We first consider (4.38), and prove this result by induction. We start with $m = 0$, which is straightforward, since

$$A_{0,1}(a) = a^{-1}$$

while

$$B(a,1) = \frac{\Gamma(a)\Gamma(1)}{\Gamma(a+1)} = a^{-1}$$

We now consider $A_{m+1,1}(a)$, and use the identity

$$A_{m+1,k}(a) = A_{m,k}(a-1) - A_{m,k}(a)$$

to write this as

$$A_{m,1}(a-1) - A_{m,1}(a)$$

We use the inductive assumption on both terms to express $A_{m+1,1}(a)$ as

$$B(a-1-m, m+1) - B(a-m, m+1)$$

and now use the recursive relationship, given by (1.14), to state that

$$A_{m+1,1}(a) = \frac{a}{m+1} B(a-1-m, m+2) - \frac{(a-1-m)}{m+1} B(a-1-m, m+2).$$

This simplifies to

$$B(a-1-m, m+2),$$

as we require.

We next prove (4.39). Again we start with $m = 0$, and write

$$A_{0,2}(a) = a^{-2},$$

while

$$B(a,1)\,F_{0,1}(a) = a^{-2}.$$

We now consider $A_{m+1,2}(a)$, and write this as

$$A_{m,2}(a-1) - A_{m,2}(a).$$

We use the inductive assumption on both terms to write this as

$$B(a-1-m, m+1)\,F_{m,1}(a-1) - B(a-m, m+1)\,F_{m,1}(a),$$

and use the recursive relationships given by (1.14) to write

$$\frac{B(a-1-m, m+2)}{m+1}\left\{aF_{m,1}(a-1) - (a-1-m)\,F_{m,1}(a)\right\}.$$

We now use the fact that

$$F_{m+1,k}(a) = F_{m,k}(a-1) + a^{-k} = F_{m,k}(a) + (a-m-1)^{-k},$$

to express both F_m functions in terms of F_{m+1}, so that $A_{m+1,2}(a)$ is

$$\frac{B(a-1-m, m+2)}{m+1}\left[a\left\{F_{m+1,1}(a) - a^{-1}\right\} - (a-1-m)\left\{F_{m+1,1}(a) - \frac{1}{a-1-m}\right\}\right],$$

which simplifies to

$$B(a-1-m, m+2)\,F_{m+1,1}(a),$$

as required.

Appendix C : Expectations of joint order statistics

We present a proof based on recognising patterns for our result on the expected value of joint order statistics, given by (4.46). We illustrate the method by taking $k = l = 1$, and, for convenience, write $E_n^{i,j} = E\left[Y_{(i:n)}Y_{(j:n)}\right]$. By taking $i = 2$ in (4.45), we have

$$E_n^{2,j} = nE_{n-1}^{1,j-1} - (j-2)\,E_n^{1,j} - (n-j+1)\,E_n^{1,j-1}. \qquad \text{(C.1)}$$

Next, we take $i = 3$ in (4.45), and consider

$$2E_n^{3,j} = nE_{n-1}^{2,j-1} - (j-3)\,E_n^{2,j} - (n-j+1)\,E_n^{2,j-1},$$

in which we now substitute appropriate expressions for $E_n^{2,j}$, $E_{n-1}^{2,j-1}$ and $E_n^{2,j-1}$. On replacing j by $j-1$ in (C.1), we see that

$$E_n^{2,j-1} = nE_{n-1}^{1,j-2} - (j-3)\,E_n^{1,j-1} - (n-j+2)\,E_n^{1,j-2},$$

and similarly, by replacing j by $j-1$ and n by $n-1$, we also have

$$E_{n-1}^{2,j-1} = (n-1)\,E_{n-2}^{1,j-2} - (j-3)\,E_{n-1}^{1,j-1} - (n-j+1)\,E_{n-1}^{1,j-2}.$$

Thus,

$$\begin{aligned}2E_n^{3,j} &= n\,(n-1)\,E_{n-2}^{1,j-2} - 2n\,(j-3)\,E_{n-1}^{1,j-1} - 2n\,(n-j+1)\,E_{n-1}^{1,j-2}\\ &\quad + (j-3)\,(j-2)\,E_n^{1,j} + 2\,(j-3)\,(n-j+1)\,E_n^{1,j-1}\\ &\quad + (n-j+1)\,(n-j+2)\,E_n^{1,j-2}.\end{aligned}$$

This expression gives $E_n^{3,j}$ in terms of $E_{n-q}^{1,j-p}$ for $p, q = 0, 1, 2$; by taking $i \geq 4$ in (4.45) and extending this approach, a pattern develops and, generally, we have

$$E_n^{i,j} = \frac{n!}{(j-i-1)!} \sum_{s=1}^{i} \sum_{q=0}^{i-s} \left[\begin{array}{c} \frac{(-1)^{s+q-1}(n+q-j)!(s+j-i-2)!}{q!(n-j)!(i-q-s)!(s-1)!(n+q+s-i)!} \\ \times E_{n-i+s+q}^{1,j-i+s} \end{array} \right],$$

for $1 \leq i \leq j-1$.

Appendix D : Δ for accelerated data

We derive Δ for data sets that have undergone acceleration, given by (5.18). We take a similar approach to the non-accelerated case, and first establish a link between Weibull and Burr distributions. We begin by proving that the Weibull distribution emerges as the limiting distribution of the Burr as Q, $a \longrightarrow \infty$, and put

$$\lambda = Q^{\tau},$$

so as $Q \longrightarrow \infty$, so does λ. Using this reparameterisation, we write the cdf from the Burr distribution as

$$\begin{aligned} & 1-\left\{1+\frac{y_{ij}^{\tau}}{\lambda f\left(X_s\right)^{\tau} h\left(X_i, X_s, \beta_{B1}, .., \beta_{Bm}\right)^{\tau}}\right\}^{-a} \\ = \; & 1-\left[\left\{\begin{array}{c} 1+ \\ \frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},..,\beta_{Bm})^{\tau}} \end{array}\right\}^{\lambda f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},..,\beta_{Bm})^{\tau}}\right]^{\frac{-a}{\lambda f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},..,\beta_{Bm})^{\tau}}}. \end{aligned}$$

Assuming that $\frac{a}{\lambda f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},..,\beta_{Bm})^{\tau}}$ remains constant as $a, \lambda \longrightarrow \infty$ we have

$$\begin{aligned} G_B &= 1-\exp\left(y_{ij}^{\tau}\right)^{\frac{-a}{\lambda f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},..,\beta_{Bm})^{\tau}}} \\ &= 1-\exp\left\{-\left(\frac{y_{ij}}{a^{-\frac{1}{\tau}} Q f\left(X_s\right) h\left(X_i, X_s, \beta_{B1}, .., \beta_{Bm}\right)}\right)^{\tau}\right\}, \end{aligned}$$

which is the cdf for a Weibull distribution with $B=\tau$, $\beta_{Wm}=\beta_{Bm}$ for all $m=1,2,...,p$ and $\theta_i = a^{-\frac{1}{\tau}} Q f\left(X_s\right) h\left(X_i, X_s, \beta_{B1}, .., \beta_{Bm}\right)$. Such a result implies that the limiting distribution for the Burr with general scale-stress relationship given by $Q f\left(X_s\right) h\left(X_i, X_s, \beta_{B1}, .., \beta_{Bm}\right)$, is the Weibull distribution with the same scale-stress relationship, the only difference being that the constant term Q is replaced by

$$Q' = \frac{Q}{a^{\frac{1}{\tau}}}.$$

We derive Δ for this general form, following Watkins (1999) for the structure and underlying steps. First note that as $Q, a \longrightarrow \infty$, τ and β_{Bm} remain bounded, so we leave both parameters in their original form. Next, we replace Q by $-\ln w$ so as $Q \longrightarrow \infty$, $w \longrightarrow 0$ and is contained within the bounds $(0,1)$. We also use the relationship between the constant from the Burr distribution and the constant from the Weibull to introduce our final reparameterisation given by

$$v = \frac{Q}{a^{\frac{1}{\tau}}} \Longrightarrow a = \left(\frac{-\ln w}{v}\right)^{\tau},$$

so as $Q, a \longrightarrow \infty$, $v \longrightarrow Q'$.

The limiting distribution argument implies that we must study the behaviour of the four derivatives from (reparameterised) l_B at $\tau = \widehat{B}$, $\beta_{Bm} = \widehat{\beta}_{Wm}$, $v = \widehat{Q'}$ and $w \longrightarrow 0$. To do this, we first note that the log-likelihood for the Burr distribution with general scale-stress relationship is

$$\begin{aligned} l_B &= n \ln a + n \ln \tau + (\tau - 1) S_e - n\tau \ln Q - n\tau \ln f(X_s) - \tau \sum_{i=1}^{k} n_i \ln h(X_i, X_s, \beta_{B1}, .., \beta_{Bm}) \\ &\quad - (a+1) F(\tau, \beta_{Bm}, Q), \end{aligned}$$

where

$$F(\tau, \beta_{Bm}, Q) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \ln \left\{ 1 + \left(\frac{y_{ij}}{Q f(X_s) h(X_i, X_s, \beta_{B1}, .., \beta_{Bm})} \right)^{\tau} \right\}.$$

Thus, the reparameterised log-likelihood becomes

$$\begin{aligned} l_B &= n \ln \tau - n\tau \ln v + (\tau - 1) S_e - n\tau \ln f(X_s) \\ &\quad - \tau \sum_{i=1}^{k} n_i \ln h(X_i, X_s, \beta_{B1}, .., \beta_{Bm}) - \left\{ \left(\frac{-\ln w}{v} \right)^{\tau} + 1 \right\} F(\tau, \beta_{Bm}, -\ln w). \end{aligned}$$

We note that

$$\begin{aligned} F_{1,0,0} &= \frac{\partial F(\tau, \beta_{Bm}, -\ln w)}{\partial \tau} \\ &= \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{\left\{ \frac{y_{ij}}{-\ln w f(X_s) h(X_i, X_s, \beta_{B1}, .., \beta_{Bm})} \right\}^{\tau} \ln \left\{ \frac{y_{ij}}{-\ln w f(X_s) h(X_i, X_s, \beta_{B1}, .., \beta_{Bm})} \right\}}{1 + \left\{ \frac{y_{ij}}{-\ln w f(X_s) h(X_i, X_s, \beta_{B1}, .., \beta_{Bm})} \right\}^{\tau}}, \end{aligned}$$

$$F_{0,1,0} = \frac{\partial F(\tau, \beta_{Bm}, -\ln w)}{\partial \beta_{Bm}}$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{-\tau \frac{\partial h(X_i, X_s, \beta_{B1},..,\beta_{Bm})}{\partial \beta_{Bm}} \left\{\frac{y_{ij}}{Af(X_s)h(X_i,X_s,\beta_{B1},..,\beta_{Bm})}\right\}^{\tau}}{h(X_i, X_s, \beta_{B1},..,\beta_{Bm})\left[1+\left\{\frac{y_{ij}}{Af(X_s)h(X_i,X_s,\beta_{B1},..,\beta_{Bm})}\right\}^{\tau}\right]},$$

and

$$F_{0,0,1} = \frac{\partial F(\tau, \beta_{Bm}, -\ln w)}{\partial w}$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\frac{\tau}{w(-\ln w)}\left\{\frac{y_{ij}}{-\ln w f(X_s)h(X_i,X_s,\beta_{B1},..,\beta_{Bm})}\right\}^{\tau}}{1+\left\{\frac{y_{ij}}{-\ln w f(X_s)h(X_i,X_s,\beta_{B1},..,\beta_{Bm})}\right\}^{\tau}}.$$

Using this notation, we derive the four reparameterised score functions as follows

$$\frac{\partial l_B}{\partial \tau} = \frac{n}{\tau} - n\ln v + S_e - n\ln f(X_s) - \sum_{i=1}^{k} n_i \ln h(X_i, X_s, \beta_{B1},..,\beta_{Bm})$$
$$-\left(\frac{-\ln w}{v}\right)^{\tau} \ln\left(\frac{-\ln w}{v}\right) F(\tau, \beta_{Bm}, -\ln w)$$
$$-\left\{\left(\frac{-\ln w}{v}\right)^{\tau}+1\right\} F_{1,0,0},$$

$$\frac{\partial l_B}{\partial \beta_{Bm}} = -\tau \sum_{i=1}^{k} \frac{n_i \frac{\partial h(X_i,X_s,\beta_{B1},..,\beta_{Bm})}{\partial \beta_{Bm}}}{h(X_i, X_s, \beta_{B1},..,\beta_{Bm})} - \left\{\left(\frac{-\ln w}{v}\right)^{\tau}+1\right\} F_{0,1,0},$$

$$\frac{\partial l_B}{\partial v} = \frac{-n\tau}{v} + \frac{\tau}{v}\left(\frac{-\ln w}{v}\right)^{\tau} F(\tau, \beta_{Bm}, -\ln w),$$

and

$$\frac{\partial l_B}{\partial w} = \frac{\tau}{v^{\tau} w}(-\ln w)^{\tau-1} F(\tau, \beta_{Bm}, -\ln w) - \left\{\left(\frac{-\ln w}{v}\right)^{\tau}+1\right\} F_{0,0,1}.$$

We must examine the limits of these derivatives as $v \longrightarrow \widehat{Q'}$, $\tau \longrightarrow \widehat{B}$, $\beta_{Bm} \longrightarrow \widehat{\beta}_{Wm}$ and $w \longrightarrow 0$, and, as in the non-accelerated case, require some preliminary results to do so. When considering them, we find it convenient to introduce a further parameterisation and write

$$\lambda = (-\ln w)^{\tau},$$

so as $w \longrightarrow 0$, $\lambda \longrightarrow \infty$.

RESULT 1.

The first result is derived from the reparameterised Weibull distribution. The log-likelihood for this distribution, when we have a general scale-stress relationship is given by

$$\begin{aligned} l_W &= n\ln B + (B-1)\,S_e - nB\ln Q' - nB\ln f(X_s) \\ &\quad - B\sum_{i=1}^{k} n_i \ln h(X_i, X_s, \beta_{W1},...,\beta_{Wm}) - \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{Q' f(X_s)\, h(X_i, X_s, \beta_{W1},...,\beta_{Wm})}\right\}^B, \end{aligned}$$

with score functions

$$\begin{aligned} \frac{\partial l_W}{\partial B} &= \frac{n}{B} + S_e - n\ln Q' - n\ln f(X_s) - \sum_{i=1}^{k} n_i \ln h(X_i, X_s, \beta_{W1},...,\beta_{Wm}) \\ &\quad - \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{Q' f(X_s)\, h(X_i, X_s, \beta_{W1},...,\beta_{Wm})}\right\}^B \ln\left\{\frac{y_{ij}}{Q' f(X_s)\, h(X_i, X_s, \beta_{W1},...,\beta_{Wm})}\right\} \end{aligned}$$

$$\begin{aligned} \frac{\partial l_W}{\partial \beta_{Wm}} &= -B\sum_{i=1}^{k}\frac{n_i \frac{\partial h(X_i, X_s, \beta_{B1},..,\beta_{Bm})}{\partial \beta_{Bm}}}{h(X_i, X_s, \beta_{W1},...,\beta_{Wm})} + \\ &\quad \sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{B\frac{\partial h(X_i, X_s, \beta_{B1},..,\beta_{Bm})}{\partial \beta_{Bm}}}{h(X_i, X_s, \beta_{W1},...,\beta_{Wm})}\left\{\frac{y_{ij}}{Q' f(X_s)\, h(X_i, X_s, \beta_{W1},...,\beta_{Wm})}\right\}^B, \end{aligned}$$

and

$$\frac{\partial l_W}{\partial Q'} = \frac{-nB}{Q'} + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{B}{Q'}\left\{\frac{y_{ij}}{Q' f(X_s)\, h(X_i, X_s, \beta_{W1},...,\beta_{Wm})}\right\}^B.$$

If we equate $\frac{\partial l_W}{\partial Q'}$ to zero then we obtain

$$Q' = \left[\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},...,\beta_{Wm})}\right\}^B}{n}\right]^{\frac{1}{B}}, \tag{D.1}$$

a profile score function of the form

$$\begin{aligned} l_W^+ &= n\ln B + (B-1)\,S_e - n\ln\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)\, h(X_i, X_s, \beta_{W1},...,\beta_{Wm})}\right\}^B\right] \\ &\quad - nB\ln f(X_s) - B\sum_{i=1}^{k} n_i \ln h(X_i, X_s, \beta_{W1},...,\beta_{Wm}), \end{aligned}$$

and profile score functions for B and β_{Wm} given by

$$\frac{\partial l_W^+}{\partial B} = \frac{n}{B} + S_e - n\ln f(X_s) - \sum_{i=1}^{k} n_i \ln h(X_i, X_s, \beta_{W1}, ..., \beta_{Wm})$$
$$-n\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},...,\beta_{Wm})}\right\}^B \ln\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},...,\beta_{Wm})}\right\}}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},...,\beta_{Wm})}\right\}^B} \quad \text{(D.2)}$$

and

$$\frac{\partial l_W^+}{\partial \beta_{Wm}} = \frac{n\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{B\frac{\partial h(X_i,X_s,\beta_{B1},..,\beta_{Bm})}{\partial \beta_{Bm}}}{h(X_i,X_s,\beta_{W1},...,\beta_{Wm})}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},...,\beta_{Wm})}\right\}^B}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},...,\beta_{Wm})}\right\}^B} -$$
$$B\sum_{i=1}^{k}\frac{n_i\frac{\partial h(X_i,X_s,\beta_{B1},..,\beta_{Bm})}{\partial \beta_{Bm}}}{h(X_i, X_s, \beta_{W1}, ..., \beta_{Wm})}. \quad \text{(D.3)}$$

RESULT 2.

The second result is analogous to the non-accelerated scenario, and we note that as $\xi \longrightarrow \infty$,

$$\ln\left(1+\frac{z}{\xi}\right) = \frac{z}{\xi} + o\left(\xi^{-2}\right) \quad \text{(D.4)}$$
$$= \frac{z}{\xi} - \frac{z^2}{2\xi^2} + o\left(\xi^{-3}\right), \quad \text{(D.5)}$$

both of which can be obtain respectively using a first and second order Taylor approximation.

RESULT 3.

$$(-\ln w)^{\tau} F(\tau, \beta_{Bm}, -\ln w) = \lambda\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln\left\{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau} h(X_i, X_s, \beta_{B1}, ..., \beta_{Bm})^{\tau}}\right\}.$$

Using (D.4), we write this as

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\tau}}{f(X_s)^{\tau} h(X_i, X_s, \beta_{B1}, ..., \beta_{Bm})^{\tau}} + o\left(\lambda^{-1}\right), \quad \text{(D.6)}$$

or, equivalently, using (D.5) we have

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}} \\ -\frac{1}{2\lambda}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{2\tau}}{f(X_s)^{2\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{2\tau}}+o\left(\lambda^{-2}\right). \tag{D.7}$$

RESULT 4.

$$\begin{aligned} F_{1,0,0} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}\ln\left\{\frac{y_{ij}}{\lambda^{\frac{1}{\tau}}f(X_s)h(X_i,X_s,\beta_{B1},...,\beta_{Bm})}\right\}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}} \\ &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{y_{ij}^{\tau}\ln y_{ij}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}-\frac{y_{ij}^{\tau}\ln\left(\lambda^{\frac{1}{\tau}}f(X_s)h(X_i,X_s,\beta_{B1},...,\beta_{Bm})\right)}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}. \end{aligned}$$

In the limit (since $\frac{\ln\lambda^{\frac{1}{\tau}}}{\lambda}\longrightarrow 0$ as $\lambda\longrightarrow\infty$) all summands tend to zero, and we can write

$$F_{1,0,0}=o\left(\lambda^{-1}\right). \tag{D.8}$$

RESULT 5.

$$\begin{aligned} (-\ln w)^{\tau}F_{1,0,0} &= \lambda\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}\ln\left\{\frac{y_{ij}}{\lambda^{\frac{1}{\tau}}f(X_s)h(X_i,X_s,\beta_{B1},...,\beta_{Bm})}\right\}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}} \\ &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}\ln\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{B1},...,\beta_{Bm})}\right\}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}} \\ &\quad -\tau^{-1}\ln\lambda\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}. \end{aligned}$$

We now use the fact that

$$\frac{1}{1+\xi}=1-\xi+\frac{\xi^2}{2!}-...$$

to write both summations as

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}\ln\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{B1},...,\beta_{Bm})}\right\}$$
$$-\tau^{-1}\ln\lambda\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}+o\left(\lambda^{-1}\right). \qquad \text{(D.9)}$$

RESULT 6.

$$\begin{aligned}(-\ln w)^{\tau}F_{0,0,1} &= \frac{\lambda\tau}{\exp\left(-\lambda^{\frac{1}{\tau}}\right)\lambda^{\frac{1}{\tau}}}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}\\ &= \frac{\tau}{\exp\left(-\lambda^{\frac{1}{\tau}}\right)\lambda^{\frac{1}{\tau}}}\\ &\quad\times\left\{\begin{array}{c}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}\\ -\lambda^{-1}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{2\tau}}{f(X_s)^{2\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{2\tau}}\\ +o\left(\lambda^{-2}\right)\end{array}\right\}. \qquad \text{(D.10)}\end{aligned}$$

RESULT 7.

$$\begin{aligned}F_{0,0,1} &= \frac{\tau}{\lambda^{\frac{1}{\tau}+1}\exp\left(-\lambda^{\frac{1}{\tau}}\right)}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}\\ &= \frac{\tau}{\lambda^{\frac{1}{\tau}+1}\exp\left(-\lambda^{\frac{1}{\tau}}\right)}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}+o\left(\lambda^{-2}\right) \qquad \text{(D.11)}\end{aligned}$$

RESULT 8.

$$\begin{aligned}(-\ln w)^{\tau}F_{0,1,0} &= \lambda\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{-\tau\frac{\partial h(X_i,X_s,\beta_{B1},..,\beta_{Bm})^{\tau}}{\partial\beta_{Bm}}}{h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}}\\ &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{-\tau\frac{\partial h(X_i,X_s,\beta_{B1},..,\beta_{Bm})^{\tau}}{\partial\beta_{Bm}}y_{ij}^{\tau}}{h(X_i,X_s,\beta_{B1},...,\beta_{Bm})f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}\\ &\quad+o\left(\lambda^{-1}\right). \qquad \text{(D.12)}\end{aligned}$$

RESULT 9.

$$\begin{aligned} F_{0,1,0} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\frac{-\tau \frac{\partial h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}}{\partial \beta_{Bm}}}{h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}} \frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}}} \\ &= o\left(\lambda^{-1}\right). \end{aligned} \tag{D.13}$$

We now consider the behaviour of the derivatives for the reparameterised log-likelihood as $\tau \longrightarrow \widehat{B}$, $\beta_{Bm} \longrightarrow \widehat{\beta}_{Wm}$, $v \longrightarrow \widehat{Q'}$ and $w \longrightarrow 0$. We first examine $\frac{\partial l_B}{\partial v}$ and use (D.6) to write

$$\frac{\partial l_B}{\partial v} = \frac{-n\tau}{v} + \frac{\tau}{v^{\tau+1}} \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{y_{ij}^{\tau}}{f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}} + o\left(\lambda^{-1}\right).$$

We note that

$$\begin{aligned} &\frac{1}{v^{\tau}} \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{y_{ij}^{\tau}}{f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}}\Bigg|_{\tau=\widehat{B},\, \beta_{Bm}=\widehat{\beta}_{Wm},\, v=\widehat{Q'}} \\ =\ &\frac{1}{\widehat{Q'}^{\widehat{B}}} \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{y_{ij}^{\widehat{B}}}{f(X_s)^{\widehat{B}} h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})^{\widehat{B}}}, \end{aligned}$$

which, using (D.1) equates to

$$\begin{aligned} &\frac{n}{\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{y_{ij}^{\widehat{B}}}{f(X_s)^{\widehat{B}} h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})^{\widehat{B}}}} \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{y_{ij}^{\widehat{B}}}{f(X_s)^{\widehat{B}} h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})^{\widehat{B}}} \\ =\ & n. \end{aligned} \tag{D.14}$$

Thus,

$$\lim_{\lambda\longrightarrow\infty} \frac{\partial l_B}{\partial v}\Bigg|_{\tau=\widehat{B},\, \beta_{Bm}=\widehat{\beta}_{Wm},\, v=\widehat{Q'}} = \lim_{w\longrightarrow 0} \frac{\partial l_B}{\partial v}\Bigg|_{\tau=\widehat{B},\, \beta_{Bm}=\widehat{\beta}_{Wm},\, v=\widehat{Q'}} = 0.$$

We now use (D.6), (D.8) and (D.9) to write

$$\begin{aligned} \frac{\partial l_B}{\partial \tau} &= \frac{n}{\tau} - n\ln v + S_e - n\ln f(X_s) - \sum_{i=1}^{k} n_i \ln h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm}) \\ &\quad + \frac{\ln v}{v^{\tau}} \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{y_{ij}^{\tau}}{f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}} \\ &\quad - \frac{1}{v^{\tau}} \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{y_{ij}^{\tau}}{f(X_s)^{\tau} h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}} \ln\left\{\frac{y_{ij}}{f(X_s)\, h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})}\right\}. \end{aligned}$$

We use (D.14) to write

$$\frac{\ln v}{v^{\tau}}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\tau}}{f\left(X_s\right)^{\tau} h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right)^{\tau}} = n\ln v,$$

and note, using (D.1), that

$$\left.\begin{array}{l}\frac{1}{v^{\tau}}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},...,\beta_{Bm})^{\tau}}\\ \quad\times\ln\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{B1},...,\beta_{Bm})}\right\}\end{array}\right|_{\tau=\widehat{B},\,\beta_{Bm}=\widehat{\beta}_{Wm},\,v=\widehat{Q'}}$$

equates to

$$\frac{n\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\widehat{B}}}{f(X_s)^{\widehat{B}}h(X_i,X_s,\beta_{W1},...,\beta_{Wm})^{\widehat{B}}}\ln\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},...,\beta_{Wm})}\right\}}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{B}}{f(X_s)^{B}h(X_i,X_s,\beta_{W1},...,\beta_{Wm})^{\widehat{B}}}},$$

which, using (D.2), is equal to

$$\frac{n}{B}+S_e-n\ln f\left(X_s\right)-\sum_{i=1}^{k}n_i\ln h\left(X_i, X_s, \beta_{W1}, ..., \beta_{Wm}\right).$$

Thus,

$$\lim_{\lambda\longrightarrow\infty}\left.\frac{\partial l_B}{\partial\tau}\right|_{\tau=\widehat{B},\,\beta_{Bm}=\widehat{\beta}_{Wm},\,v=\widehat{Q'}}=0.$$

Next, we look at

$$\frac{\partial l_B}{\partial\beta_{Bm}}=-\tau\sum_{i=1}^{k}\frac{n_i\frac{\partial h(X_i,X_s,\beta_{B1},..,\beta_{Bm})}{\partial\beta_{Bm}}}{h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right)}-\left(\frac{\lambda}{v^{\tau}}+1\right)F_{0,1,0}.$$

We use (D.12) and (D.13) to write this as

$$-\tau\sum_{i=1}^{k}\frac{n_i\frac{\partial h(X_i,X_s,\beta_{B1},..,\beta_{Bm})}{\partial\beta_{Bm}}}{h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right)}+$$

$$\frac{\tau}{v^{\tau}}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{\partial h(X_i,X_s,\beta_{B1},...,\beta_{Bm})}{\partial\beta_{Bm}}y_{ij}^{\tau}}{h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right)f\left(X_s\right)^{\tau}h\left(X_i, X_s, \beta_{B1}, ..., \beta_{Bm}\right)^{\tau}},$$

and use the fact that

$$\left.\frac{\tau}{v^{\tau}}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{\partial h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})}{\partial\beta_{Bm}}y_{ij}^{\tau}}{h\left(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm}\right)f\left(X_s\right)^{\tau}h\left(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm}\right)^{\tau}}\right|_{\tau=\widehat{B},\,\beta_{Bm}=\widehat{\beta}_{Wm},\,v=\widehat{Q'}}$$

$$=\frac{n\widehat{B}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{\partial h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})}{\partial\beta_{Bm}}}{h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})}\frac{y_{ij}^{\widehat{B}}}{f(X_s)^{\widehat{B}}h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})^{\widehat{B}}}}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\widehat{B}}}{f(X_s)^{\widehat{B}}h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})^{\widehat{B}}}},$$

which, using (D.3), is

$$B\sum_{i=1}^{k}\frac{n_i\frac{\partial h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})}{\partial\beta_{Bm}}}{h\left(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm}\right)}.$$

Thus,

$$\lim_{\lambda\longrightarrow\infty}\left.\frac{\partial l_B}{\partial\beta_{Bm}}\right|_{\tau=\widehat{B},\,\beta_{Bm}=\widehat{\beta}_{Wm},\,v=\widehat{Q'}}=0.$$

Finally, we examine

$$\begin{aligned}\frac{\partial l_B}{\partial w} &= \frac{\tau}{v^{\tau}\exp\left(-\lambda^{\frac{1}{\tau}}\right)\lambda^{\frac{1}{\tau}}}\lambda\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln\left\{1+\frac{y_{ij}^{\tau}}{\lambda f\left(X_s\right)^{\tau}h\left(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm}\right)^{\tau}}\right\}\\ &\quad-\left(\frac{\lambda}{v^{\tau}}+1\right)\frac{\tau}{\exp\left(-\lambda^{\frac{1}{\tau}}\right)\lambda^{\frac{1}{\tau}}}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}}}{1+\frac{y_{ij}^{\tau}}{\lambda f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}}}.\end{aligned}$$

We use (D.7), (D.10) and (D.11) to write this as

$$\frac{\tau}{\lambda^{\frac{1}{\tau}+1}\exp\left(-\lambda^{\frac{1}{\tau}}\right)}\left\{\begin{array}{c}\frac{1}{2v^{\tau}}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{2\tau}}{f(X_s)^{2\tau}h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{2\tau}}\\ -\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{y_{ij}^{\tau}}{f(X_s)^{\tau}h(X_i,X_s,\beta_{B1},\ldots,\beta_{Bm})^{\tau}}\end{array}\right\}+o\left(\lambda^{-2}\right).$$

We now let $\tau \longrightarrow \widehat{B}$, $\beta_{Bm} \longrightarrow \widehat{\beta}_{Wm}$ and $v \longrightarrow \widehat{Q'}$ to write this as

$$\frac{\widehat{B}}{\lambda^{\frac{1}{\widehat{B}}+1}\exp\left(-\lambda^{\frac{1}{\widehat{B}}}\right)}\left[\begin{array}{c}\frac{1}{2}\frac{n}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})}\right\}^{\widehat{B}}} \\ \times\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})}\right\}^{2\widehat{B}} \\ -\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})}\right\}^{\widehat{B}}+o\left(\lambda^{-2}\right)\end{array}\right]$$

$$= \frac{\widehat{B}}{\lambda^{\frac{1}{\widehat{B}}+1}\exp\left(-\lambda^{\frac{1}{\widehat{B}}}\right)}\frac{n}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})}\right\}^{\widehat{B}}}$$

$$\times\left[\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})}\right\}^{2\widehat{B}}}{2}-\frac{\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})}\right\}^{\widehat{B}}\right]^2}{n}\right]$$

and see that the sign of $\frac{\partial l_B}{\partial w}$ depends on

$$\Delta = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})}\right\}^{2\widehat{B}}}{2}-\frac{\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{\frac{y_{ij}}{f(X_s)h(X_i,X_s,\beta_{W1},\ldots,\beta_{Wm})}\right\}^{\widehat{B}}\right]^2}{n}.$$