

Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

Modelos de regressão quando a função de taxa de falha não é monótona e o
modelo probabilístico beta Weibull modificada

Giovana Oliveira Silva

Tese apresentada para obtenção do título de Doutor
em Agronomia. Área de concentração: Estatística e
Experimentação Agronômica

Piracicaba
2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Giovana Oliveira Silva
Bacharel em Estatística

**Modelos de regressão quando a função de taxa de falha não é monótona e o
modelo probabilístico beta Weibull modificada**

Orientador:

Prof. Dr. **EDWIN M. M. ORTEGA**

Tese apresentada para obtenção do título de Doutor
em Agronomia. Área de concentração: Estatística e
Experimentação Agronômica

Piracicaba
2008

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP**

Silva, Giovana Oliveira

Modelos de regressão quando a função de taxa de falha não é monótona e o modelo probabilístico beta Weibull modificada / Giovana Oliveira Silva. - - Piracicaba, 2008.
207 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 2008.
Bibliografia.

1. Análise de regressão e correlação 2. Análise de sobrevivência 3. Bootstrap Jackknife
remonstragem 4. Dados censurados 5. Distribuições (Probabilidade) 6. Inferência
bayesiana 7. Verossimilhança I. Título

CDD 634.945
S729e

“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”

DEDICATÓRIA

Aos meus pais,

Siloé Angelo da Silva e Maria Aparecida Oliveira Silva, pessoas fundamentais nesta conquista. Por todo o amor, o carinho, a paciência e o apoio incondicional dedicados a mim.

AGRADECIMENTOS

À Deus, por me dar forças, iluminar meu caminho e pela graça de uma família unida e solidária.

Ao meu orientador, Prof. Edwin Moises Marcos Ortega, pela orientação e colaboração dispensada ao trabalho, pela convivência enriquecedora dos últimos quatro anos e por ter se tornado um verdadeiro amigo.

Ao Prof. Gauss Moutinho Cordeiro e ao Prof. Vicente Garibay Cancho pelas importantes sugestões neste trabalho. Agradeço, também, por todo apoio e a amizade.

Aos participantes da banca da qualificação, Prof^a Clarice Garcia Borges Demétrio, Prof. Vicente Garibay Cancho, Prof. Francisco Cribari Neto, pelas sugestões.

Aos meus irmãos, Siloé Jr e Morgana, por todo amor, incentivo e carinho.

A toda minha família, em especial, a meus cunhados Margareth e Leonardo, pela amizade e carinho.

As minhas muitíssimas amadas sobrinhas, Alice e Melina, pelos momentos de descontração durante a realização deste trabalho.

A Andréia da Silva Meyer, amiga querida, pelo apoio, incentivo e bons momentos e por termos conseguido vencer juntas, desde o início, todos os desafios desse doutorado.

Aos professores do Departamento de Estatística UFBA, pelo incentivo, amizade e confiança em mim depositada.

Aos professores e funcionários do Departamento de Ciências Exatas da ESALQ/USP por suas contribuições à minha formação.

Aos amigos Afrânio Márcio Corrêa Vieira, Ana Maria Souza de Araújo, Cesar Augusto Taconeli, Denise Nunes Viola, Genevile Carife Bergamo, Idemarauro Antonio Rodrigues de Lara, Osmar Jesus Macedo, pelo acolhimento, amizade sincera e por todos os momentos de alegria.

A Elizabeth Mie Hashimoto, Fernanda Bühner Rizzato e Juliana Betini Fachini pela convivência fraterna, repleta de boas conversas e compreensão. Obrigada por todo carinho.

Aos amigos dos cursos de mestrado e doutorado do Departamento de Ciências

Exatas da ESALQ/USP, pela amizade, apoio e partilha de conhecimentos.

À Coordenadoria para o Aperfeiçoamento de Pessoal de Nível Superior (CAPES)/PICDT-UFBA, pelo apoio financeiro em forma de bolsa de estudos.

A todas as pessoas que de forma direta ou indireta, contribuíram para a realização deste trabalho.

”Deus, por meio do seu poder que age em nós, pode realizar muito mais do que pedimos ou imaginamos.”

Eféso 3, 20

SUMÁRIO

RESUMO	10
ABSTRACT	11
LISTA DE FIGURAS	12
LISTA DE TABELAS	18
1 INTRODUÇÃO	20
2 MODELOS PROBABILÍSTICOS PARA O TEMPO DE SOBREVIVÊNCIA	28
2.1 Conceitos básicos em análise de sobrevivência	28
2.2 Distribuição Weibull estendida	29
2.3 Distribuição Burr XII	32
2.4 Determinação da função de taxa de falha empírica	34
3 MODELOS DE REGRESSÃO	36
3.1 Modelo de locação e escala	36
3.2 Inferência para os modelos de regressão	37
3.2.1 Método de máxima verossimilhança	38
3.2.2 Análise bayesiana	39
3.2.3 Método de jackknife	43
3.3 Modelos de regressão quando a função de taxa de falha é não-monótona	44
3.3.1 Modelo de regressão log-Weibull estendida	45
3.3.1.1 Estudo de Simulação	49
3.3.2 Modelo de regressão log-Burr XII	52
3.3.2.1 Estudo de simulação	57
4 ANÁLISE DE SENSIBILIDADE	61
4.1 Influência global	62
4.2 Influência local	62
4.3 Esquemas de perturbação	64
4.3.1 Perturbação de casos	65
4.3.2 Perturbação da resposta	65
4.3.3 Perturbação da variável explicativa	65
4.4 “Leverage” generalizado	65

4.5	Análise de sensibilidade para os modelos de regressão	66
4.5.1	Esquemas de perturbação para o modelo log-Weibull estendida	66
4.5.1.1	Perturbação de casos	67
4.5.1.2	Perturbação da resposta	68
4.5.1.3	Perturbação da variável explicativa	69
4.5.2	Esquemas de perturbação para o modelo de regressão log-Burr XII	70
4.5.2.1	Perturbação de casos	71
4.5.2.2	Perturbação da resposta	71
4.5.2.3	Perturbação da variável explicativa	72
4.5.3	“Leverage” generalizado para o modelo de regressão log-Burr XII	73
5	ANÁLISE DE RESÍDUOS	74
5.1	Resíduo martingale	74
5.2	Resíduo tipo martingale	75
5.3	Resíduos para o modelo de regressão log-Weibull estendida	76
5.3.1	Estudo de simulação	76
5.4	Resíduos para o modelo de regressão log-Burr XII	79
5.4.1	Estudo de simulação	79
5.4.2	Resíduo tipo martingale modificado	87
5.4.2.1	Estudo de simulação	87
6	APLICAÇÕES	94
6.1	Análise dos dados de câncer de pulmão em pacientes que receberam terapia prévia	94
6.2	Análise dos dados de câncer de pulmão em pacientes que não receberam terapia prévia	99
6.3	Análise dos dados de peixes da espécie <i>Golden shiner</i>	110
6.4	Análise dos dados de suplementação de vitamina A	122
7	MODELO DE LOCAÇÃO ESCALA PARA DADOS AGRUPADOS	131
7.1	Modelo de regressão com efeito aleatório	131
7.1.1	Verossimilhança Aumentada	132
7.1.2	Verossimilhança Marginal	133
7.2	Algoritmo EM por Monte Carlo	134

7.3 Algoritmo EM por Monte Carlo para dados com censura	137
7.4 Modelo de Regressão log-Burr XII com Efeito Aleatório	141
7.5 Exemplo Simulado	143
8 NOVA DISTRIBUIÇÃO: BETA WEIBULL MODIFICADA	149
8.1 Definição do novo modelo probabilístico	149
8.2 Expansões para as funções de densidade, de distribuição acumulada e de sobrevivência	154
8.3 Confiabilidade	156
8.4 Fórmula geral para os momentos	156
8.5 Estimação de Máxima Verossimilhança	158
8.6 Análise dos dados de Aarset	160
9 CONSIDERAÇÕES FINAIS	163
9.1 Conclusões	163
9.2 Pesquisas futuras	164
REFERÊNCIAS	165
APÊNDICES	175

RESUMO

Modelos de regressão quando a função de taxa de falha não é monótona e o modelo probabilístico beta Weibull modificada

Em aplicações na área de análise de sobrevivência, é frequente a ocorrência de função de taxa de falha em forma de U ou unimodal, isto é, funções não-monótonas. Os modelos de regressão comumente usados para dados de sobrevivência são log-Weibull, função de taxa de falha monótona, e log-logística, função de taxa de falha decrescente ou unimodal. Um dos objetivos deste trabalho é propor os modelos de regressão, em forma de locação e escala, log-Weibull estendida que apresenta função de taxa de falha em forma de U e log-Burr XII que tem como caso particular o modelo de regressão log-logística. Considerando dados censurados, foram utilizados três métodos para estimação dos parâmetros, a saber, máxima verossimilhança, bayesiana e jackknife. Para esses modelos foram calculadas algumas medidas de diagnósticos de influência local e global. Adicionalmente, desenvolveu-se uma análise de resíduos baseada no resíduo tipo martingale. Para diferentes parâmetros fixados, tamanhos de amostra e porcentagens de censuras, várias simulações foram feitas para avaliar a distribuição empírica do resíduo tipo martingale e compará-la com a distribuição normal padrão. Esses estudos sugerem que a distribuição empírica do resíduo tipo martingale para o modelo de regressão log-Weibull estendida com dados censurados aproxima-se de uma distribuição normal padrão quando comparados com outros resíduos considerados neste estudo. Para o modelo de regressão log-Burr XII, foi proposta uma modificação no resíduo tipo martingale baseada no estudo de simulação para obter concordância com a distribuição normal padrão. Conjuntos de dados reais foram utilizados para ilustrar a metodologia desenvolvida. Também pode ocorrer que em algumas aplicações a suposição de independência dos tempos de sobrevivência não é válida. Assim, outro objetivo deste trabalho é introduzir um modelo de regressão log-Burr XII com efeito aleatório para o qual foi proposto um método de estimação para os parâmetros baseado no algoritmo EM por Monte Carlo. Por fim, foi desenvolvido um novo modelo probabilístico denominado de beta Weibull modificado que apresenta cinco parâmetros. A vantagem desse novo modelo é a flexibilidade em acomodar várias formas da função de taxa de falha, por exemplo, U e unimodal, e mostrou-se útil na discriminação entre alguns modelos probabilísticos alternativos. O método de máxima verossimilhança é proposto para estimar os parâmetros desta distribuição. A matriz de informação observada foi calculada. Um conjunto de dados reais é usado para ilustrar a aplicação da nova distribuição.

Palavras-chave: Análise de resíduo; Análise de sensibilidade; Distribuição beta Weibull modificada; Regressão log-Burr XII; Regressão log-Weibull estendida

ABSTRACT

Regression models when the failure rate function is no monotone and the new beta modified Weibull model

In survival analysis applications, the failure rate function may have frequently unimodal or bathtub shape, that is, non-monotone functions. The regression models commonly used for survival studies are log-Weibull, monotone failure rate function shape, and log-logistic, decreased or unimodal failure rate function shape. In the first part of this thesis, we propose location-scale regression models based on an extended Weibull distribution for modeling data with bathtub-shaped failure rate function and on a Burr XII distribution as an alternative to the log-logistic regression model. Assuming censored data, we consider a classical analysis, a Bayesian analysis and a jackknife estimator for the parameters of the proposed models. For these models, we derived the appropriate matrices for assessing the local influence on the parameter estimates under different perturbation schemes, and we also presented some ways to perform global influence. Additionally, we developed residual analysis based on the martingale-type residual. For different parameter settings, sample sizes and censoring percentages, various simulation studies were performed and the empirical distribution of the martingale-type residual was displayed and compared with the standard normal distribution. These studies suggest that the empirical distribution of the martingale-type residual for the log-extended Weibull regression model with data censored present a high agreement with the standard normal distribution when compared with other residuals considered in these studies. For the log-Burr XII regression model, it was proposed a change in the martingale-type residual based on some studies of simulation in order to obtain an agreement with the standard normal distribution. Some applications to real data illustrate the usefulness of the methodology developed. It can also happen in some applications that the assumption of independence of the times of survival is not valid, so it was added to the log-Burr XII regression model of random effects for which an estimate method was proposed for the parameters based on the EM algorithm for Monte Carlo simulation. Finally, a five-parameter distribution so called the beta modified Weibull distribution is defined and studied. The advantage of that new distribution is its flexibility in accommodating several forms of the failure rate function, for instance, bathtub-shaped and unimodal shape, and it is also suitable for testing goodness-of-fit of some special sub-models. The method of maximum likelihood is used for estimating the model parameters. We calculate the observed information matrix. A real data set is used to illustrate the application of the new distribution.

Keywords: Beta modified Weibull; Log-extended Weibull Regression; Log-Burr XII Regression; Residual analysis; Sensibility analysis

LISTA DE FIGURAS

Figura 1 - Função de taxa de falha para a distribuição Weibull estendida com $\lambda = 2$, $\alpha = 100$ e τ assumindo os valores 0,4, 0,6, 0,8, 1,0 e 1,2	31
Figura 2 - Exemplos da função de taxa de falha para a distribuição Burr XII	33
Figura 3 - Gráficos ilustrativos de algumas curvas TTT	35
Figura 4 - Função de densidade para a distribuição log-Weibull estendida com valores de σ e μ indicados e $\lambda = 1$	45
Figura 5 - Função de densidade da distribuição log-Burr XII com valores de σ e μ indicados e $k = 2$	53
Figura 6 - Gráfico normal de probabilidades para os resíduos baseados no martingale para o modelo de regressão log-Weibull estendida. Tamanho da amostra $n=50$, porcentagens de censura = 10 e 30, parâmetros $\sigma=2$ e 0,4 e $\lambda = 0,1$.	77
Figura 7 - Gráfico normal de probabilidades para os resíduos baseados no martingale para o modelo de regressão log-Weibull estendida. Tamanho da amostra $n=100$, porcentagens de censura = 10 e 30, parâmetros $\sigma=2$ e 0,4 e $\lambda = 0,1$	78
Figura 8 - Gráfico normal de probabilidade para resíduos baseados no martingale: resíduo martingale (r_M) e resíduo tipo martingale (r_D) para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura= 0, 10 e 30 e valores de parâmetros $k = 0,27, 1,2$ e $\sigma = 0,5$	83
Figura 9 - Gráfico normal de probabilidade para resíduos baseados no martingale: resíduo martingale (r_M) e resíduo tipo martingale (r_D) para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura= 0, 10 e 30 e valores de parâmetros $k = 0,27, 1,2$ e $\sigma = 0,8$	84
Figura 10 - Gráfico normal de probabilidade para resíduos baseados no martingale: resíduo martingale (r_M) e resíduo tipo martingale (r_D) para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 100$, porcentagens de censura= 0, 10 e 30 e valores de parâmetros $k = 0,27, 1,2$ e $\sigma = 0,5$	85

- Figura 11 - Gráfico normal de probabilidade para resíduos baseados no martingale: resíduo martingale (r_M) e resíduo tipo martingale (r_D) para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 100$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0, 27, 1, 2$ e $\sigma = 0, 8$ 86
- Figura 12 - Gráfico normal de probabilidade para resíduo tipo martingale modificado para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0, 27, 1, 2$ e $\sigma = 0, 5$ 90
- Figura 13 - Gráfico normal de probabilidade para resíduo tipo martingale modificado para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0, 27, 1, 2$ e $\sigma = 0, 8$ 91
- Figura 14 - Gráfico normal de probabilidade para resíduo tipo martingale modificado para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0, 27, 1, 2$ e $\sigma = 0, 5$ 92
- Figura 15 - Gráfico normal de probabilidade para resíduo tipo martingale modificado para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0, 27, 1, 2$ e $\sigma = 0, 8$ 93
- Figura 16 - Curva TTT para os dados câncer do pulmão em pacientes que receberam terapia prévia 95
- Figura 17 - Gráfico do resíduo tipo martingale contra os valores ajustados nos dados de sobrevivência de câncer de pulmão em pacientes que receberam terapia prévia para os modelos de regressão (a) log-Weibull (b) log-Weibull estendida 98
- Figura 18 - Gráfico normal de probabilidade para o resíduo tipo martingale com envelope nos dados de sobrevivência de câncer de pulmão em pacientes que receberam terapia prévia para os modelos de regressão (a) log-Weibull (b) log-Weibull estendida 98

Figura 19 - Curva TTT para o conjunto dos dados de câncer de pulmão que não receberam terapia prévia: (a) todo conjunto, (b) <i>performance status</i> com alto escore e (c) <i>performance status</i> com baixo escore	100
Figura 20 - Estimativa por histograma da função de densidade a posteriori dos parâmetros β , σ e k do modelo log-Burr XII para os dados de câncer em pacientes que não receberam terapia prévia	103
Figura 21 - Gráficos das medidas de influência global para o modelo log-Burr XII ajustado aos dados de câncer para os pacientes que não receberam terapia prévia (a) $ GD_i $ (b) $ LD_i $	105
Figura 22 - Gráficos de medidas de influência para θ do modelo de regressão log-Burr XII usando perturbação de casos nos dados de câncer para os pacientes que não receberam terapia prévia (a) Influência local d_{max} (b) Influência local total C_i	105
Figura 23 - Gráficos de medidas de influência para θ do modelo de regressão log-Burr XII usando perturbação de resposta nos dados de câncer para os pacientes que não receberam terapia prévia (a) Influência local d_{max} (b) Influência local total C_i	105
Figura 24 - Gráfico do leverage generalizado para o modelo de regressão log-Burr XII para os dados de câncer para os pacientes que não receberam terapia prévia	106
Figura 25 - (a) Gráfico normal de probabilidade para o resíduo tipo martingale modificado com envelope simulado e (b) gráfico dos valores ajustados contra o resíduo tipo martingale modificado para o modelo de regressão log-Burr XII aos dados de câncer para os pacientes que não receberam terapia prévia . .	107
Figura 26 - Curva TTT para dados dos peixes <i>Golden shiner</i>	111
Figura 27 - Estimativa por histograma da função de densidade a posteriori dos parâmetros β , σ e λ do modelo log-Weibull estendida para os dados de peixe da espécie <i>Golden shiner</i>	114
Figura 28 - Gráficos das medidas de influência global para o modelo log-Weibull estendida ajustado aos dados de sobrevivência da espécie de peixe <i>Golden shiner</i> . (a) GD_i (b) LD_i	116

- Figura 29 - Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação de casos dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total \mathbf{d}_{max} (b) Influência local total C_i . 117
- Figura 30 - Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação das resposta dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total \mathbf{d}_{max} (b) Influência local total C_i 117
- Figura 31 - Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_2 (distância) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total \mathbf{d}_{max} (b) Influência local total C_i 117
- Figura 32 - Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_3 (tamanho) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total \mathbf{d}_{max} (b) Influência local total C_i 118
- Figura 33 - Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_4 (profundidade) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total \mathbf{d}_{max} (b) Influência local total C_i 118
- Figura 34 - Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_5 (abundância) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total \mathbf{d}_{max} (b) Influência local total C_i 118
- Figura 35 - Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_6 (transparência) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total \mathbf{d}_{max} (b) Influência local total C_i 119
- Figura 36 - Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_7 (tempo inicial) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total \mathbf{d}_{max} (b) Influência local total C_i 119

Figura 37 - Gráfico do resíduo tipo martingale para o modelo log-Weibull estendida nos dados da espécie de peixe <i>Golden shiner</i>	120
Figura 38 - Gráfico normal de probabilidade para o resíduo tipo martingale com envelope sob o modelo log-Weibull estendida para os dados de sobrevivência da espécie de peixe <i>Golden shiner</i>	120
Figura 39 - Curva TTT para os dados de suplementação de vitamina A	123
Figura 40 - Estimativas de Kaplan-Meier das funções de sobrevivência para os dados de suplementação de vitamina A (a) por tratamento (b) por sexo	124
Figura 41 - Estimativa por histograma da função de densidade a posteriori dos parâmetros $\beta_0, \beta_1, \beta_2, \beta_3, \sigma$ e k do modelo log-Burr XII para os dados de suplementação de vitamina A	126
Figura 42 - Gráficos das medidas de influência global para o modelo log-Burr XII ajustado aos dados de suplementação de vitamina A (a) GD_i (b) LD_i	127
Figura 43 - Gráficos de medidas de influência para θ do modelo de regressão log-Burr XII usando perturbação de casos nos dados de sobrevivência de suplementação de vitamina A (a) Influência local d_{max} (b) Influência local total C_i	128
Figura 44 - Gráficos de medidas de influência para θ do modelo de regressão log-Burr XII usando perturbação de resposta nos dados de sobrevivência de suplementação de vitamina A (a) Influência local d_{max} (b) Influência local total C_i	128
Figura 45 - (a) Gráfico das medidas de influência para d_{max} do modelo de regressão log-Burr XII usando perturbação na variável explicativa idade nos dados de sobrevivência de suplementação de vitamina A (a) influência local d_{max} (b) influência local total C_i	128
Figura 46 - Gráfico do leverage generalizado para o modelo de regressão log-Burr XII para os dados de suplementação de Vitamina A	129
Figura 47 - (a) Gráfico dos resíduos para o modelo log-Burr XII ajustado aos dados de suplementação de Vitamina A (a) martingale (b) tipo martingale modificado	130
Figura 48 - Traço para os efeitos aleatório gerados na 100 ^a iteração	145
Figura 49 - Traço para os efeitos aleatório gerados na última iteração	145

Figura 50 - Estimativas dos parâmetros a cada iteração do algoritmo EM por Monte Carlo usando o algoritmo de Metropolis-Hasting com $N=1001$ e 2001 para as iterações 1-100 e 101-200; a linha horizontal representa os verdadeiros valores dos parâmetros que foram os valores iniciais - modelo log-Burr XII com efeito aleatório gama aplicado em dados simulados	146
Figura 51 - Diferenças entre os valores sucessivos dos parâmetros a cada iteração do algoritmo EM por Monte Carlo usando o algoritmo de Metropolis-Hasting com $N=1001$ e 2001 para as iterações 1-100 e 101-200 - modelo log-Burr XII com efeito aleatório gama aplicado em dados simulados	147
Figura 52 - Diagrama das distribuições obtidas como casos particulares da distribuição BWM.	152
Figura 53 - Gráficos da função de densidade exata da distribuição BWM com histograma para dados simulados com: (a) $a = 1,5, b = 0,5, \alpha = 0,5, \lambda = 1,5, \gamma = 3$ e (b) $a = 8, b = 1, \alpha = 0,43, \lambda = 0, \gamma = 0,5$	152
Figura 54 - Gráfico da função de densidade da distribuição BWM para alguns valores dos parâmetros	153
Figura 55 - Gráfico da função de taxa de falha da distribuição BWM para alguns valores dos parâmetros	153
Figura 56 - Assimetria e curtose da distribuição BWM como função do parâmetro a , para alguns valores de b	157
Figura 57 - Assimetria e curtose da distribuição BWM como função do parâmetro b , para alguns valores de a	158
Figura 58 - Curva TTT para os dados de Aarset	160
Figura 59 - Função de sobrevivência estimada para alguns modelos ajustados e a função de sobrevivência estimada pelo método de Kaplan-Meier dos dados de Aarset	161
Figura 60 - Função de densidade estimada de alguns modelos para os dados de Aarset .	161

LISTA DE TABELAS

Tabela 1 - Vício e Erro quadrático médio (EQM) de β_1 e o critério de informação AIC_{SUR} para os modelo de regressão Weibull e log-Weibull estendida com $\sigma = 0,5$ (função de taxa de falha crescente)	51
Tabela 2 - Vício e Erro quadrático médio (EQM) de β_1 e o critério de informação AIC_{SUR} para os modelo de regressão Weibull e log-Weibull estendida com $\sigma = 2,0$ (função de taxa de falha em forma de U)	52
Tabela 3 - Tamanhos e potência simulados para o teste da razão de verossimilhança para as hipóteses $H_0 : k = 1$ (modelo de regressão log-logística) <i>versus</i> $H_1 : k \neq 1$ (modelo de regressão log-Burr XII)	59
Tabela 4 - Variância e vício de β_1 para os modelos de regressão log-logística e log-Burr quando $k = 0,15, 0,27$ e $1,00$	60
Tabela 5 - Estatísticas descritivas para o resíduo martingale no modelo de regressão log-Burr XII	81
Tabela 6 - Estatísticas descritivas para o resíduo tipo martingale no modelo de regressão log-Burr XII	82
Tabela 7 - Estatísticas descritivas para o resíduo tipo martingale modificado no modelo de regressão log-Burr XII	89
Tabela 8 - Estimativas de máxima verossimilhança dos parâmetros dos modelos de regressão log-Weibull estendida e log-Weibull para os dados de câncer de pulmão em pacientes que receberam terapia prévia	97
Tabela 9 - Estimativas de máxima verossimilhança para os parâmetros do modelo de regressão log-Burr XII para os dados de câncer em pacientes que não receberam terapia prévia	102
Tabela 10 - Sumários a posteriori dos parâmetros do modelo log-Burr XII para os dados de câncer em pacientes que não receberam terapia prévia	103
Tabela 11 - Estimativas jackknife para os parâmetros do modelo de regressão log-Burr XII para os dados de câncer em pacientes que não recebem terapia prévia .	104

Tabela 12 -Estimativas de máxima verossimilhança e p-valor entre parênteses para os parâmetros do modelo de regressão log-Burr XII para os conjuntos de dados indicados dos dados de câncer referente a pacientes que não receberam terapia prévia	108
Tabela 13 -Variação percentual da estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-Burr XII para o conjunto de dados indicados dos dados de câncer referente a pacientes que não receberam terapia prévia	109
Tabela 14 -Estimativas de máxima verossimilhança dos parâmetros do modelo log-Weibull estendida para os dados da espécie de peixe <i>Golden Shiner</i>	112
Tabela 15 -Sumários a posteriori dos parâmetros do modelo log-Weibull estendida para os dados da espécie de peixe <i>Golden Shiner</i>	113
Tabela 16 -Estimativas de jackknife para os parâmetros do modelo log-Weibull estendida para os dados da espécie de peixe <i>Golden Shiner</i>	115
Tabela 17 -Estimativas de máxima verossimilhança para os dados da espécie de peixe <i>Golden Shiner</i> para o modelo de regressão log-Weibull estendida - modelo final	121
Tabela 18 -Estimativas de máxima verossimilhança dos parâmetros do modelo log-Burr XII para os dados de Suplementação de Vitamina A	124
Tabela 19 -Sumários a posteriori dos parâmetros do modelo de regressão log-Burr XII para os dados de suplementação de vitamina A	125
Tabela 20 -Estimativas jackknife dos parâmetros do modelo log-Burr XII para os dados de Suplementação de Vitamina A	126
Tabela 21 -Estimativas dos parâmetros do modelo de regressão log-Burr XII com efeito aleatório gama para dados simulados usando a verossimilhança marginal (SAS)	148
Tabela 22 -Estimativas de máxima verossimilhança para alguns sub-modelos para os dados de Aarset, os correspondentes erros padrão dados entre parênteses e os valores do critério AIC	162

1 INTRODUÇÃO

Considere um conjunto de dados cuja variável resposta é o tempo até a ocorrência de um evento de interesse. Esse tempo é denotado por tempo de sobrevivência ou tempo de falha e o conjunto de observações por dados de sobrevivência.

Os dados de sobrevivência apresentam características especiais. A primeira é que essa variável possui suporte \mathbb{R}_+ , portanto, não é razoável assumir que tem distribuição normal e geralmente, apresenta distribuição assimétrica positiva. A segunda característica muito freqüente em dados de sobrevivência é a presença de observações censuradas, isto é, para alguns elementos em estudo não se conhece o tempo de interesse exato, mas apenas que ocorre à direita ou à esquerda de um certo valor. Usualmente, dados censurados ocorrem, pois nem sempre é possível esperar que o evento de interesse ocorra para todos os elementos em teste. Existem diversos tipos de censura, neste trabalho considerou-se censura à direita e mecanismo aleatório, além de que a censura é não informativa. A área da estatística que engloba metodologias para esse tipo de dados é denominada análise de sobrevivência, em geral, um termo usado na área médica, ou confiabilidade, na pesquisa industrial.

Além disso, é comum em situações práticas que o tempo de sobrevivência esteja relacionado com covariáveis que explicam a variabilidade do tempo de sobrevivência. Para estudar o efeito dessas covariáveis no tempo de sobrevivência deve-se utilizar um modelo de regressão apropriado para dados censurados e tempo de sobrevivência com distribuição assimétrica.

Para escrever modelos de regressão paramétricos é necessário associar ao tempo de sobrevivência um modelo probabilístico. O tempo de sobrevivência é uma variável aleatória não negativa, e, neste trabalho, contínua e tem seu comportamento representado pela função de taxa de falha. Essa função pode ser: constante, crescente, decrescente ou não-monótona, por exemplo, forma de U ou unimodal. A forma da função de taxa de falha define um possível modelo probabilístico para o tempo de sobrevivência. Por exemplo, quando a função de taxa de falha é monótona as distribuições de probabilidade tradicionalmente associadas ao tempo de sobrevivência são exponencial, Weibull e gama e quando a função de taxa de falha é unimodal comumente utilizam-se as distribuições log-logística ou log-normal. Para

mais detalhes ver, por exemplo, Collett (2003) e Colosimo e Giolo (2006).

Apesar de nenhum dos modelos citados anteriormente apresentar função de taxa de falha na forma de U , modelos com este tipo de função são úteis em análise de sobrevivência. Mas, segundo Nelson (1982), as distribuições apresentadas na literatura para modelar este tipo de dados, como a distribuição IDB proposta por Hjorth (1980), são bastante complexas e, portanto, difíceis de serem tratadas.

Posteriormente, outros trabalhos introduziram novas distribuições para o tempo de sobrevivência cuja função de taxa de falha tem forma de U , por exemplo, Rajarshi e Rajarshi (1988) apresentaram uma revisão dessas distribuições e Haupt e Schabe (1992) propuseram um novo modelo. Mas, esses modelos não apresentam muita praticidade para serem usados.

Entretanto, na última década foi proposta uma nova classe de distribuições que é baseada em modificações da distribuição Weibull, e que podem apresentar função de taxa de falha na forma de U . Entre essas, destacam-se, a distribuição Weibull-exponenciada introduzida por Mudholkar e Srivastava (1993), que também apresenta função de taxa de falha unimodal, a distribuição Weibull aditiva proposta por Xie e Lai (1996), distribuição Weibull estendida apresentada no trabalho de Xie; Tang e Goh (2002), a distribuição Weibull modificada foi introduzida por Lai; Xie e Murthy (2003) e a distribuição Weibull modificada generalizada apresentada por Carrasco; Ortega e Cordeiro (2008). Nesses trabalhos são apresentadas aplicações dessas distribuições em análise de sobrevivência. Nadarajah e Kotz (2005) e Pham e Lai (2007) apresentam algumas das modificações mais recentes da distribuição Weibull.

Por outro lado, apesar de existirem muitas distribuições de probabilidade com função de taxa de falha em forma de U não existem muitos modelos de regressão para este tipo de função de taxa, os modelos de regressão mais conhecidos são os modelos de regressão log-Weibull exponenciada (CANCHO; BOLFARINE; ACHCAR, 1999) e log-F generalizada (KALBFLEISCH; PRENTICE, 2002).

A distribuição F generalizada tem a vantagem de ter diversas distribuições como casos especiais, entretanto não apresenta forma fechada para a sua função de sobrevivência e

função de taxa de falha, o que torna difícil a sua aplicação. A distribuição F generalizada tem como caso especial a distribuição log-logística e como caso especial no limite as distribuições exponencial, Weibull e gama (KALBFLEISCH; PRENTICE, 2002).

Diante do exposto até o momento, um dos objetivos deste trabalho é propor o modelo de regressão baseado na distribuição Weibull estendida que modela funções de taxa de falha em forma de U . Essa distribuição apresenta função de sobrevivência e função de taxa de falha analiticamente explícita o que favorece o seu uso para estudar o comportamento do tempo de sobrevivência. Esse modelo é chamado de modelo de regressão log-Weibull estendida.

Por outro lado, quando a função de taxa de falha apresenta forma unimodal, usualmente, usa-se o modelo de regressão log-normal ou log-logística, sendo que esta última distribuição apresenta função de sobrevivência e função de taxa de falha na forma fechada, além de ser uma boa aproximação para a distribuição log-normal (KALBFLEISCH; PRENTICE, 2002). Assim, outro objetivo deste trabalho é propor um modelo de regressão quando a variável tempo de sobrevivência tem distribuição Burr XII (ZIMMER; KEATS; WANG, 1998), que é uma viável alternativa às distribuições log-normal ou log-logística. Esse modelo é chamado de modelo de regressão log-Burr XII. A distribuição Burr XII tem as vantagens de apresentar função de sobrevivência e função de taxa de falha na forma fechada e ter a distribuição log-logística como um caso especial. Como o modelo de regressão log-logística é um dos mais usados em aplicações de análise de sobrevivência a proposta de um modelo mais geral é importante.

Para estimar os parâmetros desses modelos de regressão usou-se o método de máxima verossimilhança em que o método numérico do tipo quase-Newton foi necessário. Também realizaram-se testes de hipóteses a partir da distribuição assintótica dos estimadores de máxima verossimilhança. Entretanto, usar a distribuição assintótica dos estimadores de máxima verossimilhança em situações que a amostra é pequena, situação comum na prática, ou que as condições de regularidade não sejam satisfeitas pode ser inadequado e pode conduzir a resultados difíceis de serem justificados. Assim, como alternativa ao método de máxima verossimilhança, foi utilizada uma abordagem bayesiana (IBRAHIM; CHEN; SINHA, 2001)

para a estimação dos parâmetros dos modelos propostos, que utiliza a estimativa das quantidades de interesse a posteriori e intervalos de credibilidade, e foi usado também o método de jackknife (MANLY, 2005).

Numa etapa posterior ao ajuste de um modelo a um conjunto de dados, é necessário verificar as suposições deste modelo. Caso o modelo não seja adequado ele pode levar a conclusões errôneas. É importante ainda verificar a presença de observações extremas ou observações influentes que podem causar distorções nos resultados das estimativas dos parâmetros do modelo.

Para detectar observações influentes nos modelos propostos foram usadas medidas de diagnóstico baseadas no caso deleção (COOK, 1977; CHATTERJEE; HADI, 1988), denotada por influência global, e medidas de influência baseadas em pequenas perturbações nos dados ou no modelo (COOK, 1986), chamada de influência local. Além dessas medidas, foi também utilizada a medida de influência local total desenvolvida por Lesaffre e Verbeke (1998).

Além disso, para verificar se existe afastamento das suposições feitas ao modelo e detectar pontos atípicos, foram usados dois tipos de resíduos para dados de sobrevivência: martingale e tipo martingale, veja, por exemplo, Collett (2003), Klein e Moeschberger (1997), Tableman e Kim (2004), bem como foi realizado um estudo por meio da simulação Monte Carlo para identificar a distribuição empírica do resíduo tipo martingale para ambos os modelos de regressão com o intuito de utilizar as bandas de confiança construídas a partir da geração de envelopes (ATKINSON, 1987) como uma medida a mais para verificar a qualidade de ajuste do modelo de regressão. Também aplicou-se a metodologia de “leverage” generalizado desenvolvida por Wei; Hu e Fung (1998) para avaliar a influência da resposta no valor predito. Finalmente, os resultados são aplicados a conjuntos de dados reais.

Nos modelos de regressão propostos foi suposto que os tempos de sobrevivência de elementos distintos são independentes, entretanto essa suposição pode não ser válida em algumas aplicações.

Às vezes, acontece que os elementos observados formam algum agrupamento e assim os tempos de sobrevivência, dentro de cada grupo, podem não ser mutuamente in-

dependentes. Esse é o caso, por exemplo, de dados observados em grupo de elementos, por exemplo, em pessoas da mesma família ou em animais da mesma ninhada. Além disso, pode acontecer que quando alguns indivíduos têm características comuns não medidas os tempos de sobrevivência podem estar correlacionados, por exemplo, pessoas com pré-disposição a alguma doença.

Também pode ocorrer que para um único indivíduo observa-se o mesmo evento várias vezes (medidas repetidas ou eventos recorrentes), e isto faz com que os tempos de sobrevivência para o mesmo elemento apresentem correlação.

Além desses casos, pode ocorrer que não é possível incluir todos os fatores de risco relevantes no modelo de regressão, devido à limitação de conhecimento de quais fatores são importantes ou à falta dos dados necessários, o que pode ocasionar uma correlação entre as observações de um determinado grupo. Segundo Keiding; Andersen e Klein (1997), diversos trabalhos da década de 80 deram ênfase ao estudo do possível viés na estimativa dos coeficientes das covariáveis quando alguma covariáveis importantes são omitidas. Assim, caso essa correlação ou heterogeneidade não sejam consideradas pode-se chegar a resultados não esperados.

Nos casos citados, os tempos de sobrevivência não são independentes e apresentam uma estrutura de grupos, portanto, os modelos tradicionais para análise de sobrevivência não podem ser usados diretamente.

Então, pode-se assumir que todos os elementos em um dado grupo apresentam respostas correlacionadas/dependentes porque estas são dependentes de variáveis causais. Portanto, a inclusão dessas variáveis como covariáveis no modelo de regressão poderia proporcionar independência (condicional) entre as respostas do mesmo grupo. Assim, um modo de representar essa correlação dos dados é incluir ao modelo uma variável aleatória não observável, denotada por efeito aleatório, para cada indivíduo. Dessa maneira, espera-se que um modelo que considera as variáveis não observadas como um efeito aleatório desconhecido produza outros resultados que o modelo que não considera esta correlação/dependência entre as respostas. Assim, além de propor modelos de regressão para dados censurados é importante adicionar a esses modelos o efeito aleatório.

Muitos autores têm usado o modelo de locação e escala Weibull, em geral, assumindo distribuição gama para o efeito aleatório (KEIDING; ANDERSEN e KLEIN, 1997; MORRIS; CHRISTIANSEN, 1995). Para mais detalhes, ver Valença (2003). Diante da necessidade de considerar modelo de regressão com efeito aleatório, outro objetivo desta pesquisa, é propor modelos de regressão com efeito aleatório associando ao tempo de sobrevivência a distribuição Burr XII e para o efeito aleatório considerar a distribuição gama. A motivação para escolha da distribuição gama é a sua conveniência algébrica e computacional, essa última, na prática, é mais importante do que a complexidade da distribuição em casos que a forma da distribuição não é objeto de pesquisa (COLOSIMO; GIOLO, 2006).

Para estimar os parâmetros dos modelos de regressão com efeito aleatório, usualmente, usa-se o método de máxima verossimilhança baseado na verossimilhança marginal (ver por exemplo Valença, 2003). Entretanto, neste trabalho foi proposto usar o algoritmo EM por Monte Carlo (WEI; TANNER, 1990) em que para simular amostras foi adotado o algoritmo de Metropolis-Hasting (CHIB; GREENBERG, 1995). A idéia de usar essa abordagem para estimação é em pesquisa futura propor medidas de influência baseadas na função Q-afastamento (ZHU; LEE, 2001).

Outra linha de pesquisa importante na área de análise de sobrevivência é propor novas distribuições que modelem dados de sobrevivência com função de taxa de falha não-monótona.

Além das modificações das distribuições Weibull citadas anteriormente, Famoye, Lee e Olumolade (2005) propuseram a distribuição beta Weibull que apresenta como casos particulares a própria distribuição Weibull e também recentes distribuições desenvolvidas, por exemplo, a distribuição beta exponencial introduzida por Nadarajah e Kotz (2006). Usando a mesma idéia desse artigo, outro objetivo deste trabalho é propor a distribuição beta Weibull modificada que apresenta flexibilidade para acomodar várias formas da função de taxa de falha, por exemplo, forma de U e unimodal, e é apropriada na discriminação entre alguns modelos probabilísticos alternativos, tais como, a distribuição Weibull exponenciada, a distribuição Rayleigh generalizada proposta por Kundu e Rakab (2005) e a distribuição Weibull modificada generalizada apresentada por Carrasco, Ortega e Cordeiro (2008). Assim, espera-

se que esta distribuição seja útil em uma variedade de aplicações nas áreas de confiabilidade, medicina e biologia, bem como em outras áreas de investigação.

Como existem várias distribuições que podem ser associadas ao tempo de sobrevivência é necessário usar uma metodologia para selecionar um modelo adequado. Neste trabalho foi usado o gráfico do tempo total em teste (curva TTT) proposto por Aarset (1987).

O objetivo principal deste trabalho é apresentar os modelos de regressão log-Weibull estendida e log-Burr XII bem como desenvolver vários aspectos relacionados à qualidade de ajuste desses modelos. Outros objetivos referem-se a procedimento de estimação para modelos de regressão com efeito aleatório e apresentar novo modelo probabilístico. O trabalho está organizado da seguinte maneira. No Capítulo 2 são apresentadas as principais características das distribuições Weibull estendida e Burr XII. Também, discute-se a curva TTT como uma metodologia para determinação da função de taxa de falha empírica. No Capítulo 3, os modelos de regressão log-Weibull estendida e o modelo de regressão log-Burr XII são apresentados e utilizam-se três métodos de estimação para os parâmetros desses modelos: máxima verossimilhança, bayesiana e jackknife. Para estudar o desempenho dos modelos de regressão log-Weibull estendida comparado com o modelo de regressão log-Weibull e comparar o modelo de regressão log-Burr XII com o modelo de regressão log-logística foram realizados estudos de simulação. No Capítulo 4 são obtidas as curvaturas normais para a influência local sob três esquemas de perturbação, a saber, perturbação de casos, perturbação da resposta e perturbação da variável explicativa, para ambos os modelos. O “leverage” generalizado para o modelo log-Burr XII é desenvolvido. No Capítulo 5 são apresentados dois resíduos: martingale e tipo martingale, para investigar possíveis afastamentos da suposição da distribuição dos erros bem como para detectar possíveis observações atípicas. Além disso, apresenta-se um estudo de simulação via Monte Carlo com o objetivo de analisar o comportamento da distribuição empírica do resíduo tipo martingale para os modelos de regressão log-Weibull estendida e log-Burr XII. No capítulo 6 as metodologias propostas são aplicadas a conjuntos de dados reais. No sétimo capítulo foi desenvolvido um procedimento para estimação dos parâmetros do modelo de regressão com efeito aleatório. O capítulo 8 apresenta um novo modelo probabilístico, juntamente com um estudo gráfico da função de taxa de falha, expressão

para os momentos e estimação dos parâmetros o mediante o método de máxima verossimilhança. Também apresenta-se algumas distribuições que são derivadas a partir dessa nova distribuição e uma aplicação em análise de dados de sobrevivência. Por fim, o Capítulo 9 relata as principais conclusões e descreve o direcionamento da continuidade deste trabalho.

2 MODELOS PROBABILÍSTICOS PARA O TEMPO DE SOBREVIVÊNCIA

2.1 Conceitos básicos em análise de sobrevivência

Os conjuntos de dados de sobrevivência são compostos pelos tempos de sobrevivência e, muito freqüentemente, pelas censuras e as covariáveis.

Os tipos de censura são: à direita, se o tempo de ocorrência do evento de interesse está à direita do tempo registrado, à esquerda, se o tempo registrado é maior do que o tempo de falha, e intervalar, que acontece quando se sabe somente que o evento de interesse ocorreu em um certo intervalo de tempo. Para mais detalhes ver Lawless (2003). Censura à direita, a única abordada neste trabalho, é a que acontece com maior freqüência nos dados de sobrevivência.

As censuras à direita podem ocorrer de várias maneiras, de acordo com os diferenciados mecanismos de censuras, dentre os quais destacam-se: censura do tipo I, aquela em que o estudo será terminado após um período pré-estabelecido de tempo; censura do tipo II, aquela em que o estudo será terminado após ter ocorrido o evento de interesse em um número pré-estabelecido de elementos. Um terceiro mecanismo de censura, o tipo aleatório, é aquele que acontece quando o elemento deixa o estudo sem que o evento de interesse tenha ocorrido. Neste trabalho foi abordado, unicamente, mecanismo de censura aleatória.

Seja T uma variável aleatória não negativa e contínua que representa o tempo de sobrevivência de um elemento e seja E uma variável aleatória, independente de T , que representa o tempo de censura associado a este elemento. Assim, os dados observados são $t = \min(T, E)$ e

$$\delta = \begin{cases} 1 & \text{se } T \leq E \\ 0 & \text{se } T > E. \end{cases}$$

A distribuição de probabilidade da variável aleatória T pode ser especificada por meio da função de densidade, função de sobrevivência (ou função de confiabilidade) ou função de taxa de falha, sendo as três maneiras equivalentes.

A função de sobrevivência é dada por

$$S(t) = P(T \geq t) = \int_t^{\infty} f(x)dx$$

e representa a probabilidade de o elemento sobreviver pelo menos até o tempo t . Além disso, pode-se mostrar que $S(t)$ é uma função monótona decrescente, contínua, com $S(0) = 1$ e $\lim_{t \rightarrow \infty} S(t) = 0$.

A função de taxa de falha especifica a taxa instantânea de falha ou morte no tempo t condicional à sobrevivência do elemento até o tempo t . Essa função é definida por

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}.$$

Esta é a função que foi usada neste trabalho para estudar o comportamento da variável aleatória T .

Obtidos os dados, toda a informação dos parâmetros de um modelo está contida na função de verossimilhança, sendo importante escrevê-la. Considere uma amostra aleatória $(T_1, \delta_1), (T_2, \delta_2), \dots, (T_n, \delta_n)$ em que todos os T_i são provenientes de uma distribuição com vetor de parâmetro θ e que os tempos de sobrevivência e de censura são independentes. Além disso, considere que o tipo de censura é à direita e que a censura é não informativa, ou seja, sua distribuição não depende dos parâmetros de interesse. Nesse contexto, a função de verossimilhança para todos os mecanismos de censura assume a seguinte forma

$$L(\theta) \propto \prod_{i \in F} f(t_i; \theta) \prod_{i \in C} S(t_i; \theta),$$

em que F denota o conjunto de observações não censuradas e C denota o conjunto de observações censuradas. Observe que a contribuição de cada observação não censurada é a sua função de densidade e que cada observação censurada contribui com a função de sobrevivência.

2.2 Distribuição Weibull estendida

A distribuição Weibull estendida apresentada em Xie; Tang e Goh (2002) com parâmetros λ , α e τ considera que o tempo de sobrevivência T tem função de densidade dada por

$$f(t; \lambda, \tau, \alpha) = \lambda \tau \left(\frac{t}{\alpha}\right)^{\tau-1} \exp \left\{ \left(\frac{t}{\alpha}\right)^{\tau} + \lambda \alpha \left[1 - \exp \left(\left(\frac{t}{\alpha}\right)^{\tau} \right) \right] \right\}, \quad t > 0, \quad (1)$$

em que $\lambda > 0$, $\alpha > 0$ é parâmetro de escala e $\tau > 0$ é o parâmetro de forma.

A função de sobrevivência para a variável aleatória T com densidade Weibull estendida é dada por

$$S(t; \lambda, \tau, \alpha) = P(T \geq t) = \exp \left\{ \lambda \alpha \left[1 - \exp \left(\left(\frac{t}{\alpha} \right)^\tau \right) \right] \right\}.$$

A função de taxa de falha correspondente tem a seguinte forma:

$$h(t; \lambda, \tau, \alpha) = \lambda \tau \left(\frac{t}{\alpha} \right)^{\tau-1} \exp \left[\left(\frac{t}{\alpha} \right)^\tau \right].$$

Para estudar o comportamento da função taxa de falha obtém-se a derivada de primeira ordem, que é dada por

$$h'(t; \lambda, \tau, \alpha) = \frac{\lambda \tau}{\alpha} \left(\frac{t}{\alpha} \right)^{\tau-2} \exp \left[\left(\frac{t}{\alpha} \right)^\tau \right] \left[\tau \left(\frac{t}{\alpha} \right)^\tau + (\tau - 1) \right].$$

Dois situações devem ser consideradas (XIE; TANG; GOH, 2002):

i) $\tau \geq 1$.

Para qualquer $t > 0$, $h'(t) > 0$ e, portanto, $h(t)$ é uma função crescente.

ii) $\tau < 1$.

Fazendo $h'(t^*) = 0$ tem-se que $\tau \left(\frac{t^*}{\alpha} \right)^\tau + \tau - 1 = 0$; assim, o ponto crítico é dado por $t^* = \alpha \left(\frac{1}{\tau} - 1 \right)^{\frac{1}{\tau}}$. Quando $t < t^*$, $h'(t) < 0$ e, portanto, a função de taxa de falha é decrescente e quando $t > t^*$, $h'(t) > 0$ e, portanto, a função de taxa de falha é uma função crescente. Então, t^* é ponto de inflexão e a função de taxa de falha tem forma U . Além disso, $h(t) \rightarrow \infty$ quando $t \rightarrow 0$ ou $t \rightarrow \infty$.

A Figura 1 mostra o gráfico da função de taxa de falha para algumas combinações dos parâmetros. Observa-se que a função de taxa de falha é crescente quando $\tau \geq 1$ e $h(t)$ apresenta forma de U quando $\tau < 1$.

O q -ésimo momento para o tempo de sobrevivência seguindo a distribuição nova Weibull estendida é dado por

$$E(T^q) = q \int_0^{+\infty} t^{q-1} \exp \left\{ \lambda \alpha \left[1 - \exp \left(\left(\frac{t}{\alpha} \right)^\tau \right) \right] \right\} dt.$$

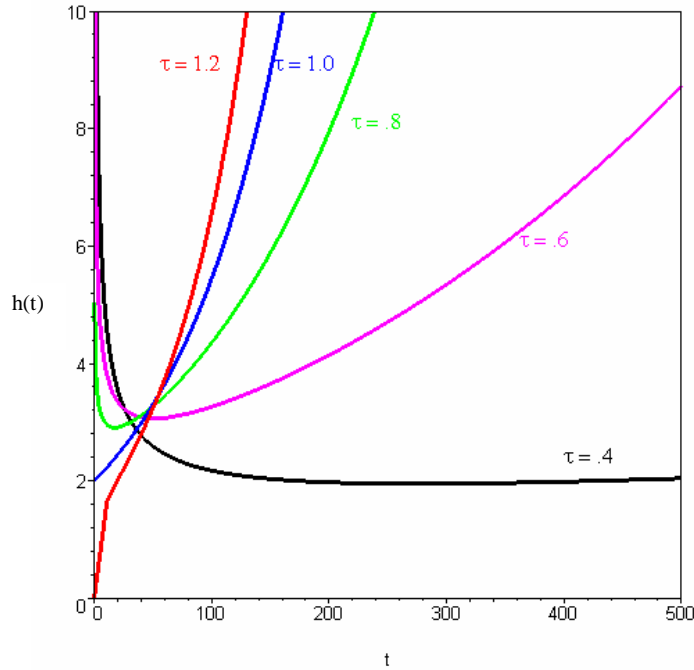


Figura 1 – Função de taxa de falha para a distribuição Weibull estendida com $\lambda = 2$, $\alpha = 100$ e τ assumindo os valores 0,4, 0,6, 0,8, 1,0 e 1,2

A integral mencionada é resolvida numericamente, podendo usar o *software* MAPLE.

Outra maneira de calcular esse q -ésimo momento é dada por Nadarajah (2005) quando $s/\tau = m$ é um número inteiro como

$$\mu'_s = m\alpha^s \exp(\lambda\alpha) \frac{\partial^{m-1}(\lambda\alpha)^{-\nu} \gamma(\nu, \lambda\alpha)}{\partial \nu^{m-1}} \Big|_{\nu=0},$$

em que $\gamma(\nu, \lambda\alpha) = \int_0^{\lambda\alpha} w^{\nu-1} e^{-w} dw$ é a função gama incompleta. Alguns casos especiais da equação (2) para $\tau = 1/2, 1/3$ e $s = 1$ e 2 são dados por Nadarajah (2005).

A partir da função de densidade da distribuição Weibull estendida dada na equação (1), nota-se que quando $\alpha = 1$, o modelo é reduzido para o modelo proposto por Chen (2000) em que a função de sobrevivência pode ser escrita como $S(t; \lambda, \tau) = \exp \left\{ \lambda \left[1 - \exp \left(t^\tau \right) \right] \right\}$.

Além disso, quando $\alpha \rightarrow \infty$ tem-se que $1 - \exp[(t/\alpha)^\tau] \approx -(t/\alpha)^\tau$, assim a função de sobrevivência reduz-se a $S(t; \lambda, \tau, \alpha) \approx \exp \left(-\lambda \alpha^{1-\tau} t^\tau \right)$, que é a distribuição Weibull com dois parâmetros sendo τ o parâmetro de forma e $\frac{\alpha^{\tau-1}}{\lambda}$ o parâmetro de escala. Se adicionalmente a isso, $\tau = 1$ e $\frac{\alpha^{\tau-1}}{\lambda}$ é constante, o modelo se reduz à distribuição exponencial

com parâmetro $\frac{\alpha^{\tau-1}}{\lambda}$. Para mais detalhes ver Xie; Tang e Goh (2002).

2.3 Distribuição Burr XII

A distribuição Burr XII, considerada em Zimmer; Keats e Wang (1998), com parâmetros s , c e k considera que o tempo de sobrevivência T tem função de densidade dada por

$$f(t; s, k, c) = ck \left[1 + \left(\frac{t}{s} \right)^c \right]^{-(k-1)} \frac{t^{c-1}}{s^c}, \quad t > 0, \quad (2)$$

em que $s > 0$ é o parâmetro de escala e $k > 0$ e $c > 0$ são os parâmetros de forma, o que a torna mais flexível. A função de sobrevivência correspondente para a variável aleatória T com densidade Burr XII é dada por

$$S(t; s, k, c) = P(T \geq t) = \left[1 + \left(\frac{t}{s} \right)^c \right]^{-k}.$$

A função de taxa de falha correspondente tem a seguinte forma

$$h(t; s, k, c) = \frac{ck \left(\frac{t}{s} \right)^{c-1}}{s \left[1 + \left(\frac{t}{s} \right)^c \right]}.$$

Com a finalidade de descrever o comportamento da função de taxa de falha, foi calculada a derivada de primeira ordem que é dada por

$$h'(t; c, k, s) = \frac{ckt^{c-2}}{s^c \left[1 + \left(\frac{t}{s} \right)^c \right]^2} \left[c - 1 - \left(\frac{t}{s} \right)^c \right].$$

Podem ser consideradas duas situações (ZIMMER; KEATS e WANG, 1998):

i) $c \leq 1$.

Para qualquer $t > 0$, $h'(t) < 0$ e, portanto, $h(t)$ é uma função decrescente.

ii) $c > 1$.

Fazendo $h'(t^*) = 0$ tem-se que $c - 1 - \left(\frac{t^*}{s} \right)^c = 0$ e o ponto crítico é dado por $t^* = s(c-1)^{\frac{1}{c}}$.

Quando $t < t^*$, $h'(t) > 0$ e, portanto, a função de taxa de falha é crescente e quando $t > t^*$, $h'(t) < 0$ e, portanto, a função de taxa de falha é uma função decrescente. Então,

t^* é ponto de inflexão e a função de taxa de falha tem forma unimodal. Além disso, $h(t) \rightarrow 0$ quando $t \rightarrow 0$ ou $t \rightarrow \infty$.

A Figura 2 mostra o gráfico da função de taxa de falha da distribuição Burr XII para algumas combinações dos parâmetros. Observa-se que a função de taxa de falha é decrescente quando $c \leq 1$ e $h(t)$ é unimodal quando $c > 1$. Além disso, para valores de c entre 1 e 2, a função de taxa de falha pode ser praticamente constante sob quase toda amplitude da função, isto depende do valor de s .

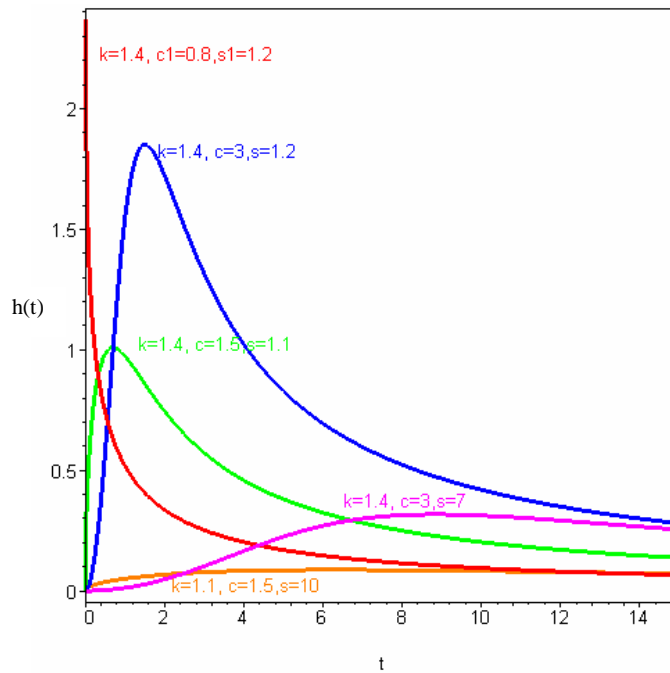


Figura 2 – Exemplos da função de taxa de falha para a distribuição Burr XII

O q -ésimo momento para o tempo de sobrevivência é dado por

$$E(T^q) = s^q k B \left[\frac{q}{c} + 1, k - \frac{q}{c} \right], \quad \text{se } ck > q,$$

em que $B(a, b)$ é a função beta completa (LAWLESS, 2003).

A partir da função de densidade da distribuição Burr XII dada na equação (2), nota-se que quando $\frac{1}{s} = m$ e $k = 1$, a distribuição Burr XII é reduzida à distribuição log-logística com função de sobrevivência dada por $S(t; s, c) = \frac{1}{1 + (tm)^c}$.

Além disso, Rodriguez (1977) mostrou que a área de cobertura da distribuição Burr XII em um determinado plano é ocupada por várias distribuições conhecidas e úteis, incluindo as distribuições normal, log-normal, gama, logística e valor extremo tipo I.

2.4 Determinação da função de taxa de falha empírica

Como existem várias formas que o gráfico da função de taxa de falha da variável T pode assumir é importante utilizar uma metodologia para identificar o modelo mais apropriado para esta variável. Em aplicações existem informações qualitativas sobre a função de taxa de falha que podem ser obtidas a partir de uma análise gráfica. Nesse contexto, o gráfico do tempo total em teste (curva TTT), proposto por Aarset (1987), pode ser usado. A curva TTT é obtida construindo um gráfico de $G(r/n) = [(\sum_{i=1}^r T_{i:n}) + (n-r)T_{r:n}]/(\sum_{i=1}^n T_{i:n})$ por r/n , sendo $r = 1, \dots, n$, e $T_{i:n}$, $i = 1, \dots, n$, as estatísticas de ordem da amostra (MUDHOLKAR; SRIVASTAVA; KOLLIA, 1996).

A curva TTT pode apresentar várias formas. Aarset (1987) mostrou que caso a curva TTT apresente uma reta diagonal (curva A), a função de taxa de falha constante é adequada; quando a curva é convexa (curva B) ou côncava (curva C), a função de taxa de falha é monotonicamente decrescente ou crescente, respectivamente; e se a curva é convexa e então côncava (curva D), a função de taxa de falha tem forma de U , e no caso reverso (curva E) é unimodal. A Figura 3 ilustra as diversas formas que a curva TTT pode apresentar.

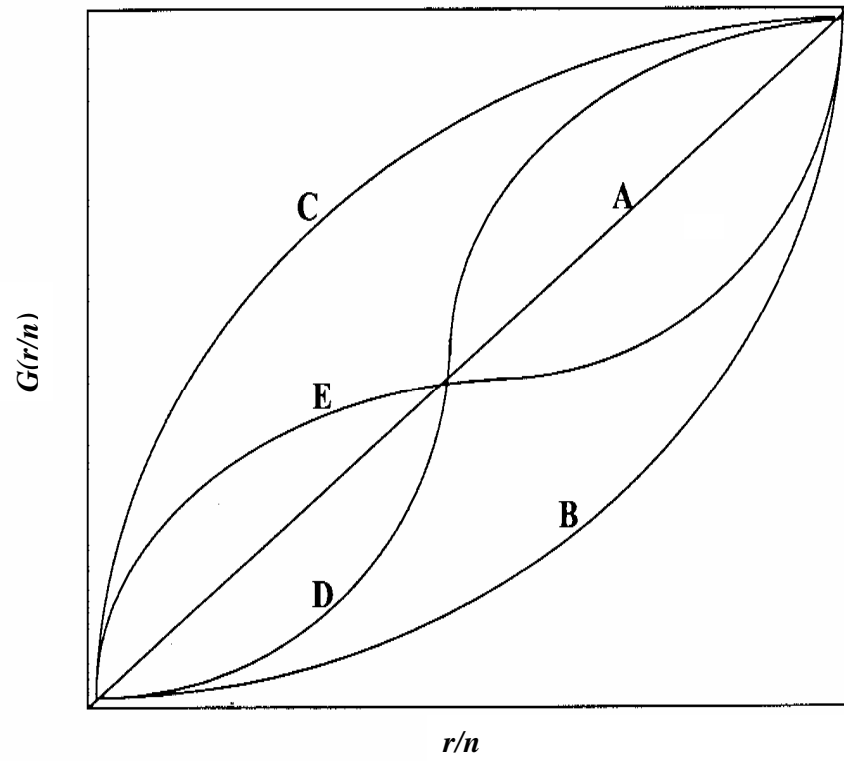


Figura 3 – Gráficos ilustrativos de algumas curvas TTT

3 MODELOS DE REGRESSÃO

É comum, na prática, a ocorrência de situações em que existam uma ou mais covariáveis associadas ao tempo de sobrevivência. Por exemplo, na indústria, o tempo de sobrevivência de um determinado equipamento pode ser influenciado pelo nível de voltagem ao qual o equipamento é submetido e na área médica o tempo de sobrevivência de um paciente pode estar relacionado com tipo de tumor ou quantidade de hemoglobina no sangue do paciente. Tais covariáveis explicam parte da heterogeneidade do tempo de sobrevivência.

Assim, faça \mathbf{x} denota um conjunto de covariáveis (independentes do tempo) que pode influenciar o tempo de sobrevivência T . Então pode-se considerar a densidade de T , dado \mathbf{x} , que é denotada por $f(t|\mathbf{x})$ e a correspondente função de sobrevivência por $S(t|\mathbf{x})$.

Uma maneira eficiente de estudar o efeito dessas covariáveis no tempo de sobrevivência é por meio de um modelo de regressão. Duas classes de modelos de regressão podem ser consideradas em análise de sobrevivência: os modelos paramétricos, também conhecidos como modelos de tempo de vida acelerado ou modelos de locação e escala, e os semi-paramétricos, também conhecidos como modelos de riscos proporcionais ou modelo de regressão de Cox. Existe uma vasta literatura sobre ambos os modelos de regressão, por exemplo, Cox (1972), Cox e Oakes (1984); Kalbfleisch e Prentice (1980), Lawless (2003), dentre outros.

Neste trabalho, foi considerado unicamente o modelo de locação e escala que é descrito a seguir.

3.1 Modelo de locação e escala

Um modelo log-linear é escrito da seguinte forma:

$$Y = \mu + \sigma Z.$$

Considerando que Y pertence a família de distribuições que se caracteriza pelo fato de ter um parâmetro de locação μ ($-\infty < \mu < \infty$) e um parâmetro de escala σ ($0 < \sigma < \infty$).

As distribuições que pertencem a essa família têm função de densidade de pro-

babilidade da forma

$$f(y; \mu, \sigma) = \frac{1}{\sigma} g\left(\frac{y - \mu}{\sigma}\right), \quad -\infty < y < \infty$$

e função de sobrevivência $G\left(\frac{y - \mu}{\sigma}\right)$.

Agora, considere que $Y = \log(T)$ e que o parâmetro de locação μ depende do vetor de covariáveis \mathbf{x} e σ é o parâmetro de escala constante. Geralmente, o parâmetro de locação é escrito como $\mu(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\beta}$, em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ é o vetor de parâmetros desconhecidos.

Nesse caso, um modelo de regressão que relaciona Y e o vetor de covariáveis \mathbf{x} é o modelo de locação e escala que é escrito da seguinte forma

$$Y = \mu(\mathbf{x}) + \sigma Z, \tag{3}$$

em que $Y = \log(T)$, Z é um erro aleatório e $\mu(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\beta}$.

É importante observar que esse modelo é log linear para T , logo é um modelo de regressão linear para Y . Outra característica desse modelo é que o vetor de covariáveis \mathbf{x} tem efeito multiplicativo em T ($T = \exp\{\mu(\mathbf{x})\} \exp\{\sigma Z\}$), logo tem efeito linear em Y . Além disso, a função de sobrevivência para Y , dado \mathbf{x} , tem a forma de $G\left(\frac{y - \mu(\mathbf{x})}{\sigma}\right)$, em que $G(z)$ é a função de sobrevivência de Z .

Lawless (2003) apresenta mais detalhes dessa classe de modelos, além de vários modelos usando distribuições de probabilidade para T comumente usadas em análise de sobrevivência.

3.2 Inferência para os modelos de regressão

Os modelos de regressão são caracterizados por quantidades desconhecidas, chamadas de parâmetros, que devem ser estimados a partir de uma amostra aleatória. O objetivo desta seção é apresentar, de maneira geral, três metodologias para a parte inferencial dos modelos de regressão.

3.2.1 Método de máxima verossimilhança

O estimador de máxima verossimilhança $\hat{\boldsymbol{\theta}}$ para o vetor de parâmetros $\boldsymbol{\theta}$ é calculado maximizando a função de verossimilhança ou, equivalentemente, maximizando o logaritmo da função de verossimilhança, $l(\boldsymbol{\theta})$. Portanto, esses estimadores são obtidos da solução do sistema de equações:

$$U(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0.$$

Como, geralmente, esse sistema é não linear, torna-se necessário usar um algoritmo de otimização. Por exemplo, um algoritmo Newton-Raphson ou quase-Newton (NO-CEDAL; WRIGHT, 2006; PRESS et al., 1992).

No caso em que o tamanho da amostra é grande e sob certas condições de regularidade para a função de verossimilhança, intervalos de confiança e testes de hipóteses podem ser obtidos usando o fato de que a distribuição assintótica para os estimadores de máxima verossimilhança é a distribuição normal com média $\boldsymbol{\theta}$ e matriz de covariância dada pelo inverso da matriz de informação de Fisher (SEN; SINGER, 1993). Portanto, a matriz de covariância assintótica é dada por $\mathbf{I}^{-1}(\boldsymbol{\theta})$ com $\mathbf{I}(\boldsymbol{\theta}) = E[\ddot{\mathbf{L}}(\boldsymbol{\theta})]$, em que $\ddot{\mathbf{L}}(\boldsymbol{\theta}) = -\left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right\}$.

Como é complicado calcular a matriz de informação de Fisher $\mathbf{I}(\boldsymbol{\theta})$ devido às observações censuradas, pode-se usar a matriz hessiana, $-\ddot{\mathbf{L}}(\boldsymbol{\theta})$, avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, que é um estimador consistente para matriz de covariância assintótica (MUDHOLKAR; SRIVASTAVA; KOLLIA, 1996). Assim, a distribuição assintótica para $\hat{\boldsymbol{\theta}}$ é dada por $\hat{\boldsymbol{\theta}}^T \sim N_{(q)}\{\boldsymbol{\theta}^T; \ddot{\mathbf{L}}(\boldsymbol{\theta})^{-1}\}$, em que $\ddot{\mathbf{L}}(\boldsymbol{\theta})$ é matriz de informação observada e q é o número de parâmetros do modelo de regressão.

Quanto à interpretação dos coeficientes estimados, uma possível proposta é baseada na razão de tempos medianos (HOSMER Jr.; LEMESHOW, 1999). Desta forma, quando a covariável é binária (1 ou 0) e considerando a razão dos tempos medianos com $x = 1$ no numerador, se $\hat{\beta}$ é negativo (positivo) implica que indivíduos com $x = 1$ apresentam tempo mediano de sobrevivência reduzido (aumentado) em $[e^{\hat{\beta}} \times 100\%]$ relativamente aos indivíduos no grupo com $x = 0$, fixando as demais covariáveis. Essa interpretação pode ser estendida para covariáveis contínuas ou categóricas.

Alternativamente, foram usados uma abordagem bayesiana e o método de jack-knife, que não pressupõem grandes amostras ou condições de regularidade para a função de verossimilhança com o objetivo de realizar inferências nos modelos propostos.

3.2.2 Análise bayesiana

Um dos objetivos da inferência bayesiana é obter estimativas pontuais ou intervalares para o vetor de parâmetros não observável $\boldsymbol{\theta}$, que também é um dos objetivos da inferência clássica. Assim, a análise bayesiana pode ser considerada uma análise alternativa à clássica.

Na abordagem bayesiana, o vetor de parâmetros $\boldsymbol{\theta}$ é assumido aleatório e tem uma distribuição a priori denotada por $\pi(\boldsymbol{\theta})$. Dessa forma, a análise bayesiana permite a incorporação de conhecimento prévio sobre os parâmetros por meio de distribuições a priori informativas. Quando não existe informação prévia sobre o parâmetro de interesse utilizam-se prioris não informativas ou vagas. A inferência sobre $\boldsymbol{\theta}$ é baseada na distribuição a posteriori, que é obtida por meio do teorema de Bayes.

Considere que $\mathbf{t} = (t_1, t_2, \dots, t_n)^T$ seja um vetor de n observações em que a distribuição de probabilidade $\pi(\mathbf{t}|\boldsymbol{\theta})$ depende de q parâmetros, isto é, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)$. Observe que $\pi(\mathbf{t}|\boldsymbol{\theta})$ é a função de verossimilhança $L(\boldsymbol{\theta}|\mathbf{t})$.

A distribuição conjunta a posteriori para $\boldsymbol{\theta}$ é dada por

$$\pi(\boldsymbol{\theta}|\mathbf{t}) = \frac{\pi(\mathbf{t}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\Theta} \pi(\mathbf{t}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}, \quad (4)$$

em que Θ denota o espaço paramétrico de $\boldsymbol{\theta}$. A quantidade $\int_{\Theta} \pi(\mathbf{t}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$, que é chamada de distribuição marginal dos dados, é a constante normalizadora de $\pi(\boldsymbol{\theta}|\mathbf{t})$.

Da equação (4), observa-se que $\pi(\boldsymbol{\theta}|\mathbf{t})$ é proporcional à função de verossimilhança multiplicada pela priori

$$\pi(\boldsymbol{\theta}|\mathbf{t}) \propto L(\boldsymbol{\theta}|\mathbf{t})\pi(\boldsymbol{\theta}).$$

Portanto, a distribuição a posteriori envolve uma contribuição das observações por meio da função de verossimilhança e uma contribuição da informação a priori para $\boldsymbol{\theta}$ por meio de $\pi(\boldsymbol{\theta})$.

A informação referente aos parâmetros do modelo é obtida por meio dos momentos de interesse da distribuição marginal a posteriori. Nesse sentido, surgem duas dificuldades: a primeira refere-se à obtenção da distribuição marginal a posteriori e a segunda ao cálculo de momentos de interesse. Em ambos os casos são necessárias resoluções de integrais que muitas vezes não apresentam solução analítica. Uma possível solução é utilizar os métodos de Monte Carlo via Cadeias de Markov - MCMC.

A combinação dos algoritmos para a simulação de distribuições com o método de Monte Carlo para a aproximação de integrais conduz aos métodos de MCMC. Dentre os vários métodos de Monte Carlo via Cadeias de Markov destacam-se o algoritmo de Gibbs (GELFAND; SMITH, 1990) e o algoritmo de Metropolis-Hastings (CHIB; GREENBERG, 1995).

O algoritmo de Gibbs usa as distribuições condicionais completas a posteriori $\pi(\theta_j | \mathbf{t}, \boldsymbol{\theta}_{(-j)})$, $j = 1, \dots, q$ em que $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)$ e $\boldsymbol{\theta}_{-j}$ representa o vetor $\boldsymbol{\theta}$ sem a j -ésima componente para obter os passos do algoritmo de Gibbs:

- atribua um valor inicial para $\boldsymbol{\theta}^0 = (\theta_1^0, \dots, \theta_q^0)$,
- amostre θ_1^1 de $\pi(\theta_1^1 | \theta_2^0, \dots, \theta_q^0, \mathbf{t})$,
- amostre θ_2^1 de $\pi(\theta_2^1 | \theta_1^1, \theta_3^0, \dots, \theta_q^0, \mathbf{t})$,
- amostre θ_3^1 de $\pi(\theta_3^1 | \theta_1^1, \theta_2^1, \dots, \theta_q^0, \mathbf{t})$,
- ⋮
- amostre θ_p^1 de $\pi(\theta_p^1 | \theta_1^1, \theta_2^1, \dots, \theta_{q-1}^1, \mathbf{t})$.

Completa-se assim uma iteração do esquema e uma transição de $\boldsymbol{\theta}^0$ para $\boldsymbol{\theta}^1 = (\theta_1^1, \theta_2^1, \dots, \theta_p^1)^T$. Repetem-se os passos anteriores com $\boldsymbol{\theta}^1$ obtido anteriormente sendo o vetor inicial e obtém-se $\boldsymbol{\theta}^2$. Repetem-se m vezes os passos acima, até que a convergência da cadeia seja alcançada. Sob certas condições de regularidade, para m suficientemente grande, $(\theta_1^m, \theta_2^m, \dots, \theta_p^m)$ tem distribuição aproximada igual a $\pi(\theta_1, \theta_2, \dots, \theta_p | \mathbf{t})$ (CASELLA; GEORGE, 1992). Além disso, segundo Geman e Geman (1984), θ_j^m pode ser considerada uma observação simulada de $\pi(\theta_j | \mathbf{t})$, que é a distribuição marginal a posteriori de θ_j .

Entretanto, existem situações em que as distribuições condicionais completas não assumem forma conhecida ou são especificadas a menos da constante de proporcionalidade, assim o problema de simulação torna-se uma tarefa complicada. Nesse contexto, uma solução é utilizar o algoritmo de Metropolis-Hastings para simular de $\pi(\theta_j | \boldsymbol{\theta}_{(-j)}, \mathbf{t})$ nos passos do algoritmo de Gibbs em cada iteração. O amostrador de Gibbs é um caso particular do amostrador de Metropolis-Hastings.

Para inicializar o algoritmo de Metropolis-Hastings, um valor inicial para $\boldsymbol{\theta}$ é dado, $\boldsymbol{\theta}^0$. Agora, suponha que o algoritmo encontra-se na iteração m e que se deseja obter uma nova observação θ_j^m de $f(x) = \pi(x | \boldsymbol{\theta}_{-j}^m, \mathbf{t})$ em que x representa um valor de θ_j . Proceda-se da seguinte maneira:

- i) amostre um ponto y de $q(z_m, y)$ em que z_m denota o valor atual de θ_j e $q(z_m, y)$ é uma distribuição proposta (função de transição apropriada),
- ii) amostre um ponto u da distribuição $U[0, 1]$ e caso $u \leq \alpha(z_m; y)$ faça $z_{m+1} = y$, caso contrário faça $z_{m+1} = z_m$ em que, $\alpha(z_m; y) = \min \left\{ 1, \frac{f(y)q(z_m|y)}{f(z_m)q(y|z_m)} \right\}$.

Repita esses passos até que a convergência da cadeia seja atingida. A distribuição proposta $q(z, y)$ deve ser de fácil simulação e garantir uma convergência rápida para a distribuição alvo $\pi(\theta_j^m | \boldsymbol{\theta}_{-j}^m, \mathbf{t})$ (CHIB; GREENBERG, 1995).

Um fator importante nos métodos de MCMC é o estudo da convergência da cadeia de Markov para o estado de equilíbrio. Existem várias maneiras propostas para o estudo da convergência e uma possibilidade é usar o método que usa componentes de variância de sequências múltiplas da cadeia proposto por Gelman e Rubin (1992).

Esses autores propuseram um método de simulação baseado em várias cadeias de Markov simuladas a partir de vários pontos iniciais. Assim, o método de Gelman e Rubin está baseado na construção de l cadeias independentes, cada uma partindo de valores iniciais distintos. Além disso, para cada parâmetro, devem-se descartar as primeiras iterações para eliminar o efeito do valor inicial e para evitar os problemas de correlação deve-se considerar um espaçamento entre as iterações sucessivas. Desse modo, obtém-se uma amostra final $\boldsymbol{\theta}^{(v,s)} = (\theta_1^{(v,s)}, \theta_2^{(v,s)}, \dots, \theta_p^{(v,s)})'$ em que v denota a v -ésima cadeia e s denota a s -ésima réplica

na v -ésima cadeia, $s = (1, \dots, m)$ e $v = (1, \dots, l)$.

Para verificar a convergência a partir do método de Gelman e Rubin, considere que U represente qualquer um dos componentes do vetor de parâmetro θ . Para cada parâmetro de interesse, calcula-se

$$\bar{u}_v = \frac{1}{m} \sum_{s=1}^m u_{vs} \quad \text{e} \quad s_v^2 = \frac{1}{m-1} \sum_{s=1}^m (u_{vs} - \bar{u}_v)^2$$

a média e a variância, respectivamente, de U para cada cadeia $v = 1, \dots, l$.

Agora, calcula-se as seguintes componentes da variância, a variância entre as cadeias e a variância dentro das cadeias, respectivamente, dadas por

$$B = \frac{m}{l-1} \sum_{v=1}^l (\bar{u}_v - \bar{u}_{..})^2 \quad \text{e} \quad W = \sum_{v=1}^l \frac{s_v^2}{l}$$

em que $\bar{u}_{..} = \frac{1}{l} \sum_{v=1}^l \bar{u}_v$.

A variância de U é estimada como a média ponderada de W e B

$$\hat{\sigma}^2 = \frac{m-1}{m} W + \frac{1}{m} B.$$

Tem-se que a distribuição de U é t de Student com parâmetro de locação $\hat{\mu} = \bar{u}_{..}$, parâmetro de escala $\sqrt{\hat{V}} = \sqrt{\hat{\sigma}^2 + B/lm}$ e graus de liberdade $gl = 2\hat{V}^2/\widehat{\text{var}}(\hat{V})$ em que

$$\begin{aligned} \widehat{\text{var}}(\hat{V}) &= \left(\frac{m-1}{m}\right)^2 \frac{1}{l} \widehat{\text{var}}(s_v^2) + \left(\frac{l+1}{vs}\right)^2 \frac{2}{l-1} B^2 \\ &\quad + 2 \frac{(l-1)(m-1)m}{lm^2} \frac{1}{l} \left[\widehat{\text{cov}}(s_v^2, \bar{u}_v^2) - 2\bar{u}_{..} \widehat{\text{cov}}(s_v^2, \bar{u}_v^2) \right] \end{aligned}$$

e as variâncias e as covariâncias são estimadas a partir dos l valores amostrais de s_v^2 , \bar{u}_v e \bar{u}_v^2 .

A convergência é monitorada calculando o fator de redução potencial de escala que é definido como:

$$\sqrt{\hat{R}} = \sqrt{\left(\frac{\hat{V}}{W}\right) \frac{gl}{df-2}} \quad (5)$$

em que \hat{R} mede o fator de redução que sofreria o parâmetro de escala da distribuição t de Student para U se a simulação continuasse.

À medida que n cresce, \hat{R} converge para 1. Na prática assume-se convergência quando \hat{R} está próximo de 1 e as iterações selecionadas para compor a amostra final são

independentes e identicamente distribuídas (GELMAN; RUBIN, 1992). Portanto, quando o fator de redução potencial de escala apresenta um valor elevado, haveria razão para acreditar que um maior número de simulações proporcionaria uma melhora na inferência sobre a distribuição desejada.

Considerando que se tem uma amostra aleatória da distribuição marginal a posteriori para θ que se pretende resumir por meio de estimação pontual ou do intervalo de credibilidade, a estimativa pontual pode ser dada pela média ou mediana amostrais. O intervalo de credibilidade central a $100\% \gamma$ para θ é dado por $\left(\theta_{\frac{1-\gamma}{2}}, \theta_{\frac{1+\gamma}{2}}\right)$, que são definidos pelos quantis da distribuição marginal a posteriori $\frac{1-\gamma}{2}$ e $\frac{1+\gamma}{2}$, respectivamente, e uma aproximação Monte Carlo para esse intervalo é obtida ordenando-se a amostra aleatória e calculando os quantis empíricos.

Em análise de sobrevivência, existem vários trabalhos utilizando a abordagem bayesiana, por exemplo Ibrahim, Chen e Sinha (2001). Além disso, Cancho, Bolfarine e Achcar (1999) apresentaram uma análise bayesiana para a distribuição Weibull exponenciada.

3.2.3 Método de jackknife

A idéia do método jackknife é pensar em termos do que é feito usualmente quando se deseja estimar um valor médio, mas de um ponto de vista não usual.

Seja T_1, \dots, T_n uma amostra aleatória de tamanho n . A média amostral é dada por

$$\bar{T} = \sum_{i=1}^n \frac{T_i}{n},$$

que é usada para estimar a média populacional.

Pode-se calcular a média amostral sem a l -ésima observação, ou seja,

$$\bar{T}_{-l} = \frac{(\sum_{i=1}^n T_i) - T_l}{n - 1}. \quad (6)$$

Então, da equação (6) obtém-se que

$$T_l = n\bar{T} - (n - 1)\bar{T}_{-l}. \quad (7)$$

Em uma situação geral, suponha que um parâmetro θ é estimado por alguma função dos n valores amostrais. Denote esse estimador por $\hat{E}(T_1, \dots, T_n)$; para simplificar a notação será omitido (T_1, \dots, T_n) . Remova T_l e obtenha o estimador parcial \hat{E}_{-l} . Pela analogia com a equação (7), existe um conjunto de pseudo-valores que podem ser calculados:

$$\hat{E}_i^* = n\hat{E} - (n-1)\hat{E}_{-i}, \quad i = 1, \dots, n. \quad (8)$$

Esses pseudo-valores preservam a mesma lei dos valores T_i em estimar a média. A média amostral dos pseudo-valores é dada por

$$\hat{E}^* = \frac{\sum_{i=1}^n \hat{E}_i^*}{n},$$

que é o estimador jackknife de θ . Considerando os pseudo-valores como uma amostra aleatória de estimativas independentes, então, sugere-se que a variância desse estimador jackknife pode ser estimada por s^2/n , em que s^2 é a variância amostral dos pseudo-valores. Assim, pode-se sugerir que um intervalo de confiança $100(1 - \alpha)\%$ para θ é dado por $\hat{E}^* \pm t_{\alpha/2, n-1}s/\sqrt{n}$, em que $t_{\alpha/2, n-1}$ é o valor que é excedido com probabilidade $\alpha/2$ para a distribuição t com $(n-1)$ graus de liberdade. Para mais detalhes, leia Manly (2005).

Uma vantagem de substituir um estimador pela sua versão jackknife é que esse último remove o viés de ordem $1/n$ (MANLY, 2005). Isso pode ser visto supondo que o valor esperado do estimador \hat{E} de θ baseado na amostra completa de tamanho n é $\theta(1 + A/n)$. Então o valor esperado do estimador parcial \hat{E}_{-j} com a j -ésima observação removida é $\theta(1 + A/(n-1))$. Segue que o valor esperado do j -ésimo pseudo-valor como definido na equação (8) é

$$n\left[\theta\left(1 + \frac{A}{n}\right)\right] - (n-1)\left[\theta\left(1 + \frac{A}{n-1}\right)\right] = \theta.$$

Isso ocorre para cada valor de j e conseqüentemente o valor esperado de \hat{E}^* , a média dos pseudo-valores, é também θ .

3.3 Modelos de regressão quando a função de taxa de falha é não-monótona

A seguir são descritos modelos de locação e escala considerando as distribuições Weibull estendida e Burr XII, além de apresentar a parte inferencial desses modelos.

3.3.1 Modelo de regressão log-Weibull estendida

Considere que T seja uma variável aleatória com distribuição Weibull estendida, com função de densidade dada na equação (1). Também considere as seguintes reparametrizações $\tau = \frac{1}{\sigma}$ e $\alpha = \exp(\mu)$. Nesse caso, por meio do método jacobiano (CASELLA; BERGER, 2002), a função de densidade de $Y = \log(T)$ é dada por

$$f(y; \lambda, \sigma, \mu) = \frac{\lambda}{\sigma} \exp\left(\frac{y - \mu}{\sigma}\right) \exp\left\{\mu + \exp\left(\frac{y - \mu}{\sigma}\right) + \lambda \exp(\mu)\right\} \left[1 - \exp\left(\exp\left(\frac{y - \mu}{\sigma}\right)\right)\right], \quad (9)$$

em que $-\infty < y < \infty$, $\lambda > 0$, $\sigma > 0$, e $-\infty < \mu < \infty$. A Figura 4 mostra a função de densidade para a variável aleatória Y , que é chamada de distribuição log-Weibull estendida.

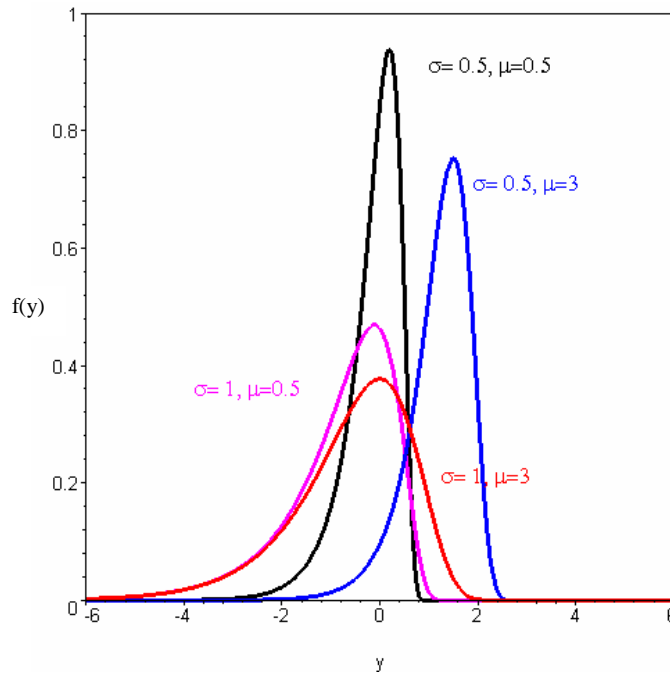


Figura 4 – Função de densidade para a distribuição log-Weibull estendida com valores de σ e μ indicados e $\lambda = 1$

A função de sobrevivência correspondente é dada por

$$S(y; \lambda, \sigma, \mu) = \exp\left\{\lambda \exp(\mu) \left[1 - \exp\left(\exp\left(\frac{y - \mu}{\sigma}\right)\right)\right]\right\}.$$

O modelo de locação e escala dado na equação (3), considerando que Y , dado

\mathbf{x} , tem distribuição log-Weibull estendida, dada na equação (9), pode ser representado por

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \sigma z_i, \quad i = 1, \dots, n, \quad (10)$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$, $\sigma > 0$ e $\lambda > 0$ são parâmetros desconhecidos, $\mathbf{x}_i^T = (x_{i1}, \dots, x_{ip})$ é o vetor de variáveis explicativas e z_i tem função de densidade dada por

$$f(z; \lambda, \mu) = \lambda \exp \left\{ z + \mu + \exp(z) + \lambda \exp(\mu) [1 - \exp(\exp(z))] \right\}, \quad \forall -\infty < z < \infty, \quad (11)$$

que também foi obtida pelo método do jacobiano.

Nesse caso, a função de sobrevivência de $Y|\mathbf{x}$ é dada por

$$S(y|\mathbf{x}; \lambda, \sigma, \boldsymbol{\beta}) = \exp \left\{ \lambda \exp\{\mathbf{x}^T \boldsymbol{\beta}\} \left[1 - \exp \left(\exp \left(\frac{y - \mathbf{x}^T \boldsymbol{\beta}}{\sigma} \right) \right) \right] \right\}. \quad (12)$$

Além disso, para os valores amostrais $(y_1, \delta_1, \mathbf{x}_1), \dots, (y_n, \delta_n, \mathbf{x}_n)$, em que $y_i = \min(\log(t_i), \log(E_i))$, δ_i é o indicador de censura e \mathbf{x}_i o vetor de covariáveis associado ao i -ésimo indivíduo, o logaritmo da função de verossimilhança de λ , σ e $\boldsymbol{\beta}$ considerando dados com censura à direita e mecanismo aleatório para o modelo de localização e escala na forma da equação (10) é dado por

$$\begin{aligned} l(\boldsymbol{\theta}) \propto & r \log(\lambda) - r \log(\sigma) + \sum_{i \in F} \mathbf{x}_i^T \boldsymbol{\beta} + \sum_{i \in F} z_i + \sum_{i \in F} \exp(z_i) \\ & + \sum_{i \in F} \lambda \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \left\{ 1 - \exp[\exp(z_i)] \right\} + \sum_{i \in C} \lambda \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \left\{ 1 - \exp[\exp(z_i)] \right\}, \end{aligned} \quad (13)$$

em que r é o número de observações não censuradas, F denota o conjunto de observações não censuradas, C denota o conjunto de observações censuradas e $z_i = \frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}$.

Da seção 3.2.1, sabe-se que os estimadores de máxima verossimilhança $\hat{\lambda}$, $\hat{\sigma}$ e $\hat{\boldsymbol{\beta}}$ de λ , σ e $\boldsymbol{\beta}$ são obtidos a partir da resolução do sistema de equações

$$\begin{cases} \frac{\partial l(\lambda, \sigma, \boldsymbol{\beta})}{\partial \lambda} = 0 \\ \frac{\partial l(\lambda, \sigma, \boldsymbol{\beta})}{\partial \sigma} = 0 \\ \frac{\partial l(\lambda, \sigma, \boldsymbol{\beta})}{\partial \beta_l} = 0, \quad l = 0, 1, \dots, p, \end{cases}$$

ou, equivalentemente,

$$\left\{ \begin{array}{l} \frac{r}{\lambda} + \sum_{i=1}^n \exp(\mathbf{x}_i^T \boldsymbol{\beta}) [1 - \exp(z_i)] = 0 \\ -\frac{r}{\sigma} - \frac{1}{\sigma} \sum_{i \in F} z_i - \frac{1}{\sigma} \sum_{i \in F} z_i \exp(z_i) + \sum_{i=1}^n \frac{\lambda \exp(\mathbf{x}_i^T \boldsymbol{\beta})}{\sigma} z_i \exp(z_i) \exp[\exp(z_i)] = 0 \\ \sum_{i \in F} x_{il} - \sum_{i \in F} \frac{x_{il}}{\sigma} - \sum_{i \in F} \frac{x_{il}}{\sigma} \exp(z_i) + \sum_{i=1}^n \lambda x_{il} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) - \sum_{i=1}^n \lambda x_{il} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \\ \exp[\exp(z_i)] + \sum_{i=1}^n \lambda \frac{x_{il}}{\sigma} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \exp(z_i) \exp[\exp(z_i)] = 0. \end{array} \right.$$

Como não existe uma forma analítica fechada para encontrar esses estimadores tem-se que recorrer a métodos numéricos para resolver o sistema de equações. Dessa forma, as estimativas desses parâmetros foram obtidas por meio de maximização numérica do logaritmo da função de verossimilhança, utilizando o algoritmo quase-Newton baseado no método BFGS (NOCEDAL; WRIGHT, 2006; PRESS et al., 1992).

A linguagem matricial de programação Ox (DOORNIK, 2006), por meio da subrotina MaxBFGS, foi usada para obter as estimativas de máxima verossimilhança para o vetor de parâmetros $\boldsymbol{\theta} = (\lambda, \sigma, \boldsymbol{\beta}^T)^T$ e os valores iniciais para os coeficientes de regressão foram dados pelas estimativas destes parâmetros sob o modelo de regressão log-Weibull ou valor extremo (LAWLESS, 2003) e para os outros parâmetros foram dados valores próximos às estimativas dos parâmetros de locação e escala do modelo de regressão log-Weibull. Essa linguagem está disponível em www.doornik.com de onde pode ser obtida gratuitamente para uso acadêmico.

Intervalos de confiança e testes de hipóteses podem ser obtidos usando a distribuição para grandes amostras dos estimadores de máxima verossimilhança, como descrito na seção 3.2.1. Assim, a aproximação normal assintótica para $\hat{\boldsymbol{\theta}}$ pode ser expressa por $\hat{\boldsymbol{\theta}}^T \sim N_{(p+2)}\{\boldsymbol{\theta}^T; \ddot{\mathbf{L}}(\boldsymbol{\theta})^{-1}\}$, em que $\ddot{\mathbf{L}}(\boldsymbol{\theta})$ é a matriz de informação observada com dimensão

$(p+2)(p+2)$, obtida de

$$\ddot{\mathbf{L}}(\boldsymbol{\theta}) = - \begin{pmatrix} \mathbf{L}_{\lambda\lambda} & \mathbf{L}_{\lambda\sigma} & \mathbf{L}_{\lambda\beta_j} \\ \cdot & \mathbf{L}_{\sigma\sigma} & \mathbf{L}_{\sigma\beta_j} \\ \cdot & \cdot & \mathbf{L}_{\beta_j\beta_s} \end{pmatrix}$$

com sub-matrizes dadas no apêndice A.

Para a abordagem bayesiana no modelo dado na equação (10), foram consideradas as seguintes densidades a priori para σ , λ e $\boldsymbol{\beta}^T$:

- $\lambda \sim \Gamma(c_1, d_1)$, c_1 e d_1 conhecidos;
- $\sigma \sim \Gamma(c_2, d_2)$, c_2 e d_2 conhecidos;
- $\beta_j \sim N(\mu_{0j}, \sigma_{0j}^2)$, μ_{0j} e σ_{0j}^2 conhecidos, $j = 1, \dots, p$.

Prioris não informativas, considerando independência entre os parâmetros são obtidas considerando $c_1 = c_2 = d_1 = d_2 = 0$ e σ_{0j}^2 grande.

Assumindo independência entre os parâmetros, a distribuição conjunta a posteriori para σ , λ e $\boldsymbol{\beta}$ é dada por

$$\begin{aligned} \pi(\sigma, \lambda, \boldsymbol{\beta}^T | \mathbf{t}) &\propto \lambda^{c_1-1} \exp(-\lambda d_1) \sigma^{c_2-1} \exp(-\sigma d_2) \exp \left[-\frac{1}{2} \sum_{j=1}^p \left(\frac{\beta_j - \mu_{0j}}{\sigma_{0j}} \right)^2 \right] \left(\frac{\lambda}{\sigma} \right)^r \\ &\exp \left(\sum_{i \in F} \mathbf{x}_i^T \boldsymbol{\beta} \right) \exp \left(\sum_{i \in F} z_i \right) \exp \left(\sum_{i \in F} \exp(z_i) \right) \\ &\exp \left\{ \sum_{i=1}^n \lambda \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \left[1 - \exp \left(-\exp(z_i) \right) \right] \right\}, \end{aligned} \quad (14)$$

em que $z_i = \frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}$.

As densidades marginais a posteriori dos parâmetros $(\lambda, \sigma, \boldsymbol{\beta}^T)$ não são facilmente obtidas, isto porque a integração da densidade posteriori conjunta, equação (14), é muito complexa. Uma alternativa para situações desse tipo é o uso do algoritmo Metropolis-Hastings, descrito na seção 3.2.2.

O algoritmo Metropolis-Hastings permite simular observações de distribuições conjuntas a partir das densidades condicionais a posteriori. Assim, para o uso de tal algoritmo consideram-se as seguintes distribuições condicionais a posteriori completas para os

parâmetros λ , σ e $\boldsymbol{\beta}$:

$$\pi(\lambda|\sigma, \boldsymbol{\beta}^T, \mathbf{t}) \propto \lambda^{c_1+r-1} \exp \left\{ -\lambda d_1 + \sum_{i=1}^n \left[\lambda \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \left[1 - \exp \left(\exp(z_i) \right) \right] \right] \right\},$$

$$\begin{aligned} \pi(\sigma|\lambda, \boldsymbol{\beta}^T, \mathbf{t}) &\propto \sigma^{c_2-r-1} \exp \left\{ -\sigma d_2 + \left(\sum_{i \in F} z_i \right) + \left(\sum_{i \in F} \exp(z_i) \right) \right. \\ &\quad \left. + \sum_{i=1}^n \left[\lambda \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \left[1 - \exp \left(\exp(z_i) \right) \right] \right] \right\}, \end{aligned}$$

$$\begin{aligned} \pi(\beta_j|\lambda, \sigma, \boldsymbol{\beta}_{-j}^T, \mathbf{t}) &\propto \exp \left[-\frac{1}{2} \left(\frac{\beta_j - \mu_{0j}}{\sigma_{0j}} \right)^2 \right] \exp \left(\sum_{i \in F} x_{ij} \beta_j \right) \exp \left[\sum_{i \in F} \left(\frac{y_i - x_{ij} \beta_j}{\sigma} \right) \right] \\ &\quad \exp \left\{ \sum_{i \in F} \left[\exp \left(\frac{y_i - x_{ij} \beta_j}{\sigma} \right) \right] \right\} \\ &\quad \exp \left\{ \sum_{i=1}^n \lambda \exp^{x_{ij} \beta_j} \left[1 - \exp \left(\exp \left(\frac{y_i - x_{ij} \beta_j}{\sigma} \right) \right) \right] \right\}. \end{aligned}$$

Seguindo a teoria descrita na seção 3.2.3, as estimativas para o vetor de parâmetros $\boldsymbol{\theta}$ do modelo de regressão log-Weibull estendida usando o método de jackknife podem ser obtidas usando a linguagem matricial de programação Ox.

3.3.1.1 Estudo de Simulação

Para examinar o desempenho dos modelos de regressão log-Weibull e log-Weibull estendida, esta seção apresenta um estudo de simulação para diferentes valores de n (tamanho da amostra) e porcentagens de censura.

Os tempos de sobrevivência, denotados por T_1, \dots, T_n , foram gerados da distribuição Weibull estendida, função de densidade dada na equação (1), por meio da simulação de Monte Carlo e consideraram-se, novamente, as seguintes reparametrizações: $\sigma = \frac{1}{\tau}$ e $\alpha = \exp(\mu)$. Os tempos de censura, denotados por E_1, \dots, E_n , foram gerados de uma distribuição uniforme no intervalo $[0, \eta]$, em que η foi escolhido de maneira que as porcentagens de censura, 10 ou 30, fossem alcançadas. Os tempos de sobrevivência considerados foram obtidos fazendo $t_i = \min\{E_i, T_i\}$ e o indicador de censura é $\delta_i = I_{T_i \leq E_i}$, sendo $I(\cdot)$ a função indicadora.

Foram geradas amostras com tamanhos 50 e 100 e os valores dos parâmetros foram fixados em $\sigma = 0,5$ e $2,0$, que estão relacionados a função de taxa de falha crescente e em forma de U , respectivamente, $\lambda = 0,15, 0,5$ e 2 e assumindo $\mu_i = \beta_0 + \beta_1 x_i$, em que x_i foi obtido a partir de uma distribuição uniforme no intervalo $(0,1)$, $\beta_0 = -3$ e $\beta_1 = 3$.

Para cada combinação dos valores de n , λ , σ e porcentagem de censura, 1.000 amostras foram simuladas. Para cada amostra, o modelo de regressão log-Weibull e log-Weibull estendida, equação (10), com $\mu_i = \beta_0 + \beta_1 x_i$, foram ajustados e o vício e o erro quadrático médio (EQM) dos estimadores de máxima verossimilhança (COX; HINDLEY, 1974) e os critérios de informação de Akaike (AIC) e de Bayes (BIC) e um AIC melhorado proposto por Liang and Zou (2008), denotado por AIC_{SUR} , foram calculados.

Os critérios de informação AIC e BIC podem ser calculados por

$$AIC = -2l(\boldsymbol{\theta}) + 2(p + 2 + u) \quad \text{e} \quad BIC = -2l(\boldsymbol{\theta}) + (p + u) \log(n),$$

em que $u = 1$ para o modelo Weibull e $u = 2$ para o modelo Weibull estendida. Para mais detalhes, leia Klein e Moeschberger (1997). Liang e Zou (2008) propuseram um AIC melhorado que é dado por

$$AIC_{SUR} = AIC + \frac{2(p+2)(p+3)}{n-p-3}.$$

Menores valores de AIC, BIC e AIC_{SUR} indicam modelos mais adequados. Nesse estudo de simulação, esses três critérios apresentaram a mesma conclusão, portanto, unicamente, o AIC_{SUR} foi apresentado nesta sub-seção.

Dos resultados da simulação, mostrados nas Tabelas 1-2, foi observado que quando os tempos de sobrevivência são simulados da distribuição Weibull estendida, o modelo de regressão log-Weibull estendida apresenta razoáveis estimativas para β_1 . Além disso, EQM diminui quando n aumenta ou a porcentagem de censura diminui. Em alguns casos, quando o tamanho da amostra é pequeno e\ou a porcentagem de censura é alto, o EQM para Weibull estendida é maior que para Weibull. Em geral, o EQM mostra melhores resultados quando a função de taxa de falha é crescente que quando a função de taxa de falha tem forma de U .

Mais importante, em todos os casos, os valores do AIC_{SUR} para o modelo de regressão log-Weibull estendida, apresentados nas Tabelas 1-2, são menores que esses valores

para o modelo de regressão log-Weibull. Isto indica que o modelo de regressão log-Weibull estendida pode melhorar significativamente o ajuste sobre a Weibull se esse é o modelo correto, o que indica as vantagens do modelo de regressão log-Weibull estendida.

Tabela 1 – Vício e Erro quadrático médio (EQM) de β_1 e o critério de informação AIC_{SUR} para os modelo de regressão Weibull e log-Weibull estendida com $\sigma = 0,5$ (função de taxa de falha crescente)

λ	n	% de censura	Vício		EQM		AIC_{SUR}	
			Weibull	Weibull Estendida	Weibull	Weibull Estendida	Weibull	Weibull Estendida
0,15	50	10	-0,4262	-0,0560	0,1992	1,5673	5,2935	-3,2924
		30	-0,4362	0,0035	0,2155	0,0306	11,2620	3,1068
	100	10	-0,4159	-0,0013	0,1807	0,0102	-2,4467	-20,566
		30	-0,4279	-0,0308	0,1944	0,7992	9,9607	-5,9404
0,50	50	10	-0,5701	0,0139	0,3558	0,0173	29,1930	20,9440
		30	-0,5777	0,0199	0,3767	0,0608	30,232	23,238
	100	10	-0,5683	-0,0009	0,3373	0,0198	49,9930	31,9620
		30	-0,5752	-0,0270	0,3530	0,8430	51,5530	36,2730
2,00	50	10	-0,8230	0,0288	0,7379	0,1341	58,157	52,844
		30	-0,8344	0,0175	0,7755	0,1599	52,496	48,167
	100	10	-0,8252	-0,0018	0,7089	0,0537	106,28	94,621
		30	-0,8384	-0,0088	0,7422	0,0792	101,53	91,661

Tabela 2 – Vício e Erro quadrático médio (EQM) de β_1 e o critério de informação AIC_{SUR} para os modelo de regressão Weibull e log-Weibull estendida com $\sigma = 2,0$ (função de taxa de falha em forma de U)

λ	n	% de censura	Vício		EQM		AIC_{SUR}	
			Weibull	Weibull Estendida	Weibull	Weibull Estendida	Weibull	Weibull Estendida
0,15	50	10	-1,7047	0,0829	3,1726	1,1377	134,93	125,02
		30	-1,8079	0,1572	3,6719	1,9820	121,75	115,80
	100	10	-1,6680	0,0456	2,9044	0,5176	239,72	258,41
		30	-1,7193	0,0642	3,1315	0,7628	238,11	228,70
0,50	50	10	-2,3069	0,3423	5,8210	5,8500	162,26	153,34
		30	-2,4660	0,2697	6,8075	8,9306	147,86	143,99
	100	10	-2,2966	0,1258	5,5058	1,7795	315,10	317,61
		30	6,3861	2,3386	282,33	270,16	278,93	263,95
2,00	50	10	-3,3186	0,4226	11,9710	16,5380	190,04	189,91
		30	-3,5187	0,3299	13,6730	18,0210	166,68	162,41
	100	10	11,2950	5,0585	373,89	365,27	370,50	359,05
		30	-3,5544	0,3343	13,2760	7,5188	335,25	327,30

3.3.2 Modelo de regressão log-Burr XII

Considere que T seja uma variável aleatória com distribuição Burr XII, com densidade dada na equação (2). Também considere as seguintes reparametrizações: $c = \frac{1}{\sigma}$ e $s = \exp(\mu)$. Nesse caso, por meio do método jacobiano, a função de densidade de $Y = \log(T)$ é dada por

$$f(y; k, \sigma, \mu) = \frac{k}{\sigma} \left[1 + \exp\left(\frac{y - \mu}{\sigma}\right) \right]^{-(k+1)} \exp\left(\frac{y - \mu}{\sigma}\right), \quad (15)$$

em que $-\infty < y < \infty$, $k > 0$, $\sigma > 0$, e $-\infty < \mu < \infty$. A Figura 5 mostra a função de densidade para a variável aleatória Y , que é chamada de distribuição log-Burr XII.

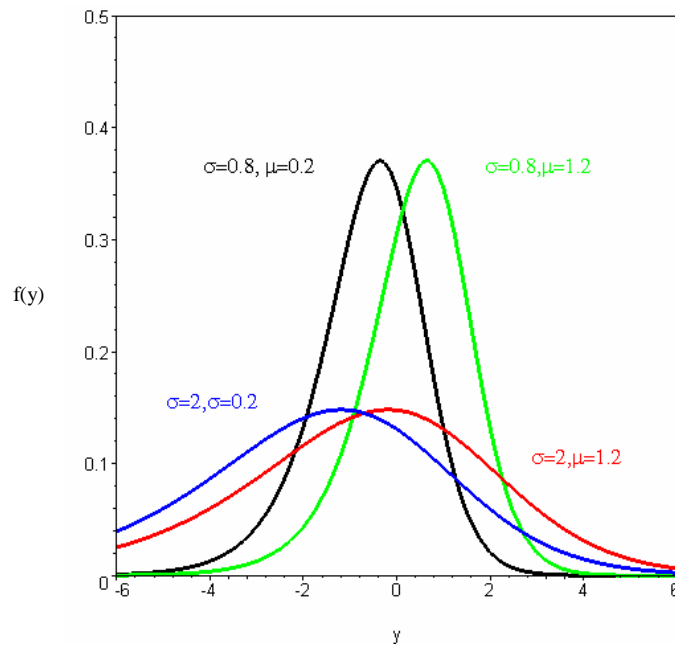


Figura 5 – Função de densidade da distribuição log-Burr XII com valores de σ e μ indicados e $k = 2$

A função de sobrevivência correspondente é dada por

$$S(y; k, \sigma, \mu) = \left[1 + \exp\left(\frac{y - \mu}{\sigma}\right) \right]^{-k}. \quad (16)$$

Além disso, o seguinte teorema apresenta uma característica importante da variável Y .

Teorema 1: Para a variável aleatória Y a função geradora de momentos é dada por

$$M_Y(t) = ks^t B\left(\frac{t}{c} + 1, k - \frac{t}{c}\right), \quad \text{se } kc > t,$$

em que $B(a, b)$ é a função beta completa (LAWLESS, 2003). A demonstração é dada no apêndice B.

Portanto, a média de Y é dada por

$$E(Y) = s + \sigma[\psi(1) - \psi(k)], \quad \text{se } kc > t,$$

em que $\psi(a)$ é a função digama (LAWLESS, 2003).

O modelo de locação e escala dado na equação (3), considerando que Y , dado

\mathbf{x} , tem distribuição log-Burr XII, função de densidade dada na equação (15), pode ser representado por

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \sigma z_i, \quad i = 1, \dots, n, \quad (17)$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$, $\sigma > 0$ e $k > 0$ são os parâmetros desconhecidos, $\mathbf{x}_i^T = (x_{i1}, \dots, x_{ip})$ é o vetor das variáveis explicativas e Z tem a seguinte função de densidade:

$$f(z; k) = k[1 + \exp(z)]^{-(k-1)} \exp(z), \quad -\infty < z < \infty. \quad (18)$$

Nesse caso, a função de sobrevivência de $Y|\mathbf{x}$ é dada por

$$S(y|\mathbf{x}; k, \sigma, \boldsymbol{\beta},) = \left[1 + \exp\left(\frac{y - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}\right) \right]^{-k}.$$

Além disso, para os valores da amostra $(y_1, \delta_1, \mathbf{x}_1), (y_2, \delta_2, \mathbf{x}_2), \dots, (y_n, \delta_n, \mathbf{x}_n)$, em que $y_i = \min\{\log(t_i), \log(E_i)\}$, δ_i é o indicador de censura e \mathbf{x}_i o vetor de covariáveis associado ao i -ésimo indivíduo, o logaritmo da função de verossimilhança de k , σ e $\boldsymbol{\beta}$ considerando dados com censura à direita e mecanismo aleatório para o modelo de locação e escala, equação (17), é dado por

$$l(\boldsymbol{\theta}) \propto r \log(k) - r \log(\sigma) + \sum_{i \in F} z_i - (k+1) \sum_{i \in F} \log[1 + \exp(z_i)] - k \sum_{i \in C} \log[1 + \exp(z_i)], \quad (19)$$

em que r é o número de observações não censuradas, F denota o conjunto de observações não censuradas, C denota o conjunto de observações censuradas e $z_i = \frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}$.

Da seção 3.2.1, sabe-se que os estimadores de máxima verossimilhança \hat{k} , $\hat{\sigma}$ e $\hat{\boldsymbol{\beta}}$ de k , σ e $\boldsymbol{\beta}$ são obtidos a partir da resolução do sistema de equações

$$\begin{cases} \frac{\partial l(k, \sigma, \boldsymbol{\beta})}{\partial k} = 0 \\ \frac{\partial l(k, \sigma, \boldsymbol{\beta})}{\partial \sigma} = 0 \\ \frac{\partial l(k, \sigma, \boldsymbol{\beta})}{\partial \beta_l} = 0, \quad l = 0, 1, \dots, p, \end{cases}$$

ou equivalentemente

$$\left\{ \begin{array}{l} \frac{r}{k} - \sum_{i=1}^n \log(1 + \exp(z_i)) = 0 \\ -\frac{r}{\sigma} - \frac{1}{\sigma} \sum_{i \in F} z_i + \frac{1}{\sigma} \sum_{i \in F} \frac{z_i \exp(z_i)}{1 + \exp(z_i)} + \frac{k}{\sigma} \sum_{i=1}^n \frac{z_i \exp(z_i)}{1 + \exp(z_i)} = 0 \\ \sum_{i \in F} -\frac{x_{il}}{\sigma} + \frac{1}{\sigma} \sum_{i \in F} \frac{x_{il} \exp(z_i)}{1 + \exp(z_i)} + \frac{k}{\sigma} \sum_{i=1}^n \frac{x_{il} \exp(z_i)}{1 + \exp(z_i)} = 0. \end{array} \right.$$

Como não existe uma forma analítica fechada para encontrar esses estimadores tem-se que recorrer a métodos numéricos para resolver o sistema de equações. Dessa forma, as estimativas desses parâmetros foram obtidas por meio de maximização numérica do logaritmo da função de verossimilhança, utilizando um processo iterativo, neste trabalho, o algoritmo quase-Newton baseado no método BFGS (NOCEDAL; WRIGHT, 2006; PRESS et al., 1992).

A linguagem matricial de programação Ox (DOORNIK, 2006), por meio da sub-rotina MaxBFGS, foi usada para obter as estimativas de máxima verossimilhança para o vetor de parâmetros $\boldsymbol{\theta} = (k, \sigma, \boldsymbol{\beta}^T)^T$ e os valores iniciais para os coeficientes de regressão foram dados pelas estimativas destes parâmetros sob o modelo de regressão log-logística e para os outros parâmetros foram dados valores próximos as estimativas dos parâmetros de locação e escala do modelo de regressão log-logística.

Intervalos de confiança e testes de hipóteses podem ser obtidos usando a distribuição para grandes amostras dos estimadores de máxima verossimilhança, como descrito na seção 3.2.1. Assim, a distribuição normal assintótica para $\hat{\boldsymbol{\theta}}$ pode ser expressa por $\hat{\boldsymbol{\theta}}^T \sim N_{(p+2)}\{\boldsymbol{\theta}^T; \ddot{\mathbf{L}}(\boldsymbol{\theta})^{-1}\}$, em que $\ddot{\mathbf{L}}(\boldsymbol{\theta})$ é a matriz de informação observada com dimensão $(p+2)(p+2)$, obtida de

$$\ddot{\mathbf{L}}(\boldsymbol{\theta}) = - \begin{pmatrix} \mathbf{L}_{kk} & \mathbf{L}_{k\sigma} & \mathbf{L}_{k\beta_j} \\ \cdot & \mathbf{L}_{\sigma\sigma} & \mathbf{L}_{\sigma\beta_j} \\ \cdot & \cdot & \mathbf{L}_{\beta_j\beta_s} \end{pmatrix}$$

com sub-matrizes dadas no apêndice C.

Um possível interesse é investigar a utilização do modelo de regressão log-logística que é um modelo mais simples do que o modelo proposto. Como os modelos de

regressão log-Burr e log-logística são encaixados pode-se utilizar o teste da razão da verossimilhanças (CASELLA; BERGER, 2002) para discriminar estes modelos. Nesse caso, as hipóteses são dadas por $H_0 : k = 1$ versus $H_1 : k \neq 1$. A estatística de teste é dada por $TRV = -2 \times \log \left(\frac{L(\hat{\boldsymbol{\theta}}_0)}{L(\hat{\boldsymbol{\theta}})} \right)$, em que $\hat{\boldsymbol{\theta}}_0$ é o estimador de máxima verossimilhança para $\boldsymbol{\theta}$ especificado pela hipótese nula e rejeita-se a hipótese nula quando $TRV > \chi_{1-\alpha}^2(1)$ que é o quantil $(1 - \alpha)$ da distribuição qui-quadrado com 1 grau de liberdade.

Para a abordagem bayesiana no modelo dado na equação (17), foram consideradas as seguintes densidades a priori para σ , k e $\boldsymbol{\beta}^T$ que são dadas por

- $k \sim \Gamma(c_1, d_1)$, c_1 e d_1 conhecidos;
- $\sigma \sim$ gama inversa (c_2, d_2) , c_2 e d_2 conhecidos;
- $\beta_j \sim N(\mu_{0j}, \sigma_{0j}^2)$, μ_{0j} e σ_{0j}^2 conhecidos, $j = 0, \dots, p$.

Prioris não informativas, considerando independência entre os parâmetros, são obtidas fazendo $c_1 = c_2 = d_1 = d_2 = 0$ e σ_{0j}^2 grande.

Assumindo independência entre os parâmetros, a distribuição conjunta a posteriori para σ , k e $\boldsymbol{\beta}$ é dada por

$$\begin{aligned} \pi(\sigma, k, \boldsymbol{\beta}^T | D) \propto & k^{c_1-1} \exp(-kd_1) \sigma^{-(c_2+1)} \exp\left(-\frac{d_2}{\sigma}\right) \exp\left[-\frac{1}{2} \sum_{j=0}^p \left(\frac{\beta_j - \mu_{0j}}{\sigma_{0j}}\right)^2\right] \\ & \left(\frac{k}{\sigma}\right)^r \exp\left(\sum_{i \in F} z_i\right) \prod_{i \in F} \{[1 + \exp(z_i)]^{-(k+1)}\} \prod_{i \in c} \{[1 + \exp(z_i)]^{-k}\}, \quad (20) \end{aligned}$$

em que $z_i = \frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}$.

Analogamente à seção anterior, as densidades marginais a posteriori dos parâmetros $(k, \sigma, \boldsymbol{\beta}^T)^T$, obtidas da expressão dada na equação (20), não são facilmente obtidas e para analisar os parâmetros de interesse utiliza-se o algoritmo de Metropolis-Hastings, descrito na seção 3.2.2.

Para o uso do algoritmo Metropolis-Hastings, pode-se mostrar que as distribuições condicionais a posteriori completas são dadas por

$$\pi(k | \sigma, \boldsymbol{\beta}^T, \mathbf{t}) \propto k^{c_1+r-1} \exp\{-kd_1\} \exp \prod_{i \in F} \{[1 + \exp(z_i)]^{-(k+1)}\} \prod_{i \in c} \{[1 + \exp(z_i)]^{-k}\},$$

$$\pi(\sigma|k, \boldsymbol{\beta}^T, \mathbf{t}) \propto \sigma^{-(c_2+r+1)} \exp\left(-\frac{d_2}{\sigma}\right) \exp\left(\sum_{i \in F} z_i\right) \prod_{i \in F} \{[1 + \exp(z_i)]^{-(k+1)}\} \\ \prod_{i \in c} \{[1 + \exp(z_i)]^{-k}\},$$

$$\pi(\beta_j|k, \sigma, \boldsymbol{\beta}_{-j}, \mathbf{t}) \propto \exp\left[-\frac{1}{2} \sum_{j=0}^p \left(\frac{\beta_j - \mu_{0j}}{\sigma_{0j}}\right)^2\right] \\ \exp\left(\sum_{i \in F} z_i\right) \prod_{i \in F} \{[1 + \exp(z_i)]^{-(k+1)}\} \prod_{i \in c} \{[1 + \exp(z_i)]^{-k}\}.$$

Seguindo a teoria descrita na seção 3.2.3, as estimativas para o vetor de parâmetros $\boldsymbol{\theta}$ do modelo de regressão log-Burr XII usando o método de jackknife podem ser obtidas usando a linguagem matricial de programação Ox.

3.3.2.1 Estudo de simulação

Nesta seção é apresentado um estudo de simulação para investigar o desempenho do modelo de regressão log-Burr XII para diferentes valores de n (tamanho da amostra) e porcentagens de censura.

Os tempos de sobrevivência, denotados por T_1, \dots, T_n , foram gerados da distribuição Burr XII, função de densidade dada na equação (2), por meio da simulação de Monte Carlo e consideraram-se, novamente, as seguintes reparametrizações $c = \frac{1}{\sigma}$ e $s = \exp(\mu)$. Os tempos de censura, denotados por E_1, \dots, E_n , foram gerados de uma distribuição uniforme no intervalo $[0, \eta]$, em que η foi escolhido de maneira que as porcentagens de censura, 10 ou 30, fossem alcançadas. Os tempos de sobrevivência considerados foram obtidos fazendo $t_i = \min\{E_i, T_i\}$ e o indicador de censura é $\delta_i = I_{T_i \leq E_i}$, sendo $I(\cdot)$ a função indicadora.

Para esse estudo, geraram-se amostras com tamanhos 50 e 100 e os valores dos parâmetros foram fixados em $\sigma = 0,36$ e $0,8$ (relacionados a dados com função de taxa de falha unimodal), $k = 0,15, 0,27$ e $1,00$ e assumindo $\mu_i = \beta_0 + \beta_1 x_i$ em que x_i foi gerado a partir de uma distribuição uniforme no intervalo $(0, 1)$ e β_0 e β_1 fixados.

Para cada situação avaliada, combinação dos valores de n, k, σ e porcentagem de censura, 1.000 amostras aleatórias foram simuladas. Para cada amostra, o modelo de regressão log-Burr XII, equação (17), e o modelo de regressão log-logística, com $\mu_i = \mathbf{x}_i^T \boldsymbol{\beta} = \beta_0 + \beta_1 x_i$,

foram ajustados e o teste da razão de verossimilhança para as hipóteses $H_0 : k = 1$ versus $H_1 : k \neq 1$ foi realizado. Então, os valores da estatística de teste, descrita em 3.3.2, foram comparados com o valor crítico dado pelo percentil 0,95 da distribuição χ^2 com 1 grau de liberdade, ou seja, $\chi_1^2 = 3,84$. Em seguida, calculou-se a proporção de amostras que H_0 foi rejeitada que para amostras geradas com $k = 1$ é o tamanho do teste e para amostras geradas com $k \neq 1$ é poder do teste. O apêndice D apresenta o programa feito na linguagem matricial de programação Ox para obter o tamanho e o poder do teste.

Os resultados mostrados na Tabela 3 correspondem aos valores de poder e tamanho para os testes realizados. Observa-se dessa tabela, que os tamanhos dos testes foram próximos de 5% e a potência do teste apresenta valores superiores a 65%, que é uma boa potência. Vale destacar que a potência aumenta quando k e a porcentagem de censura diminui e/ou n aumenta. Quando os tempos de sobrevivência são simulados a partir da distribuição Burr XII ($k=0,15$ e $0,27$) a probabilidade de não rejeitar este modelo é alta, portanto, o modelo de regressão log-Burr XII parece ser o mais preferível em relação ao modelo de regressão log-logística quando k apresenta valores menores do que 1. Além disso, o valor de σ parece não influenciar os resultados da simulação.

A variância e o vício (EQM) (COX; HINKLEY, 1974) foram também calculados para amostras com 10 e 30% de censuras considerando as mesmas combinações citadas anteriormente dos valores dos parâmetros e tamanhos de amostra. Para cada situação, foram geradas 1.000 amostras e os modelos de regressão log-logística e log-Burr XII ajustados as estas amostras. Dos resultados das simulações, dados na Tabela 4, observa-se que quando $k = 1$ a variância e o vício apresentam valores similares para ambos os modelos e quando $k < 1$ os valores para variância e para o EQM são menores para o modelo de regressão log-Burr XII. Note também que, a variância e o vício diminuem quando n aumenta, são maiores quando σ se aproxima de 1 e aumentam quando a porcentagem de censura aumenta. Também observa-se para ambos modelos que, em geral, as estimativas para o parâmetro β_1 são razoáveis (valor do vício aproximadamente zero).

Tabela 3 – Tamanhos e potência simulados para o teste da razão de verossimilhança para as hipóteses $H_0 : k = 1$ (modelo de regressão log-logística) *versus* $H_1 : k \neq 1$ (modelo de regressão log-Burr XII)

k	σ	n	% de censura	% de resultados significativos para o teste ao nível de 5%
0,15	0,36	50	10	95,4
			30	89,2
		100	10	99,9
	0,80	50	30	99,3
			10	95,2
		100	10	99,9
0,27	0,36	50	30	99,3
			10	75,5
		100	10	96,1
	0,80	50	30	92,6
			10	76,0
		100	10	96,3
1,00	0,36	50	30	93,5
			10	4,30
		100	10	5,90
	0,80	50	30	5,40
			10	4,60
		100	10	5,70
			30	5,30

Tabela 4 – Variância e vício de β_1 para os modelos de regressão log-logística e log-Burr quando $k = 0,15, 0,27$ e $1,00$

k	σ	n	% de censura	Variância		Vício	
				Log-logística	Log-Burr	Log-logística	Log-Burr
0,15	0,36	50	10	1,06	0,52	0,04	0,05
			30	1,15	0,56	0,18	0,05
		100	10	0,53	0,23	-0,00	0,02
			30	0,50	0,24	0,01	0,01
	0,80	50	10	5,13	2,52	0,10	0,10
			30	5,08	2,61	0,21	0,12
		100	10	2,53	1,15	-0,01	0,02
			30	2,51	1,16	0,15	0,02
0,27	0,36	50	10	0,41	0,28	0,01	0,03
			30	0,47	0,33	0,12	0,04
		100	10	0,20	0,14	0,00	0,01
			30	0,22	0,15	0,04	0,01
	0,80	50	10	1,97	1,36	0,03	0,06
			30	2,18	1,48	0,19	0,07
		100	10	0,96	0,66	-0,01	0,01
			30	0,95	0,68	0,04	0,02
1,00	0,36	50	10	0,10	0,10	0,01	0,01
			30	0,13	0,13	0,02	0,02
		100	10	0,05	0,05	0,00	0,00
			30	0,07	0,07	0,00	0,00
	0,80	50	10	0,49	0,50	0,02	0,03
			30	0,59	0,60	0,04	0,05
		100	10	0,24	0,24	0,01	0,01
			30	0,29	0,29	0,02	0,02

4 ANÁLISE DE SENSIBILIDADE

Um etapa importante após a modelagem é detectar observações influentes que podem causar importantes distorções nos resultados da análise. Tal etapa é conhecida como análise de sensibilidade.

Denote por i -ésimo caso a combinação das observações (y_i, δ_i) com os respectivos valores das variáveis explicativas. Cook (1977) introduziu a medida de influência denotada por distância de Cook para avaliar a influência de um caso nas estimativas dos parâmetros do modelo de regressão linear. Essa distância é baseada na deleção de casos que é um enfoque comum para estudar o efeito de remover um caso completamente da análise e assim é possível determinar quais casos podem influenciar nos resultados da análise. Essa é a base para a influência global. Uma desvantagem da deleção de casos é que ela não permite dizer em que aspecto específico do modelo o caso tem influência. Para mais detalhes, veja Chatterjee e Hadi (1988); Cook, Peña e Weisberg (1988).

Medidas de diagnóstico baseadas na deleção de casos, tais como distância de Cook ou o afastamento de verossimilhança, têm sido aplicadas em vários modelos estatísticos, veja, por exemplo, Christensen, Pearson e Johnson (1992), Davison e Tsai (1992), Xie e Wei (2007).

Mas o trabalho mais inovador na análise de sensibilidade foi feito por Cook (1986), que propõe avaliar se um modelo é relativamente estável sob pequenas perturbações incorporadas ao modelo, ao invés de usar deleção de casos. Essa metodologia é denominada de influência local. Nesse contexto, investiga-se como os resultados da análise são modificados sob pequenas perturbações no modelo e essas perturbações têm interpretações específicas.

Diversos autores têm estudado influência local em modelos de análise de sobrevivência. Por exemplo, Pettit e Bin Daud (1989) e Weissfeld (1990) investigaram a influência local em modelo de regressão com riscos proporcionais, Escobar e Meeker (1992) adaptaram métodos de influência local para modelos de regressão paramétricos com dados censurados. Para o modelo de regressão log-Weibull com dados censurados, Weissfeld e Schneider (1990) compararam numericamente os métodos de influência local e deleção de casos. Além desses, existem vários trabalhos que consideraram influência local em modelos de regressão com dados

censurados para uma específica distribuição, por exemplo, Ortega, Bolfarine e Paula (2003) que consideraram a distribuição log-gama generalizada, Ortega, Cancho e Bolfarine (2006) a distribuição log-Weibull exponenciada e, mais recentemente, Leiva et al. (2007) a distribuição log-Birnbaum-Saunders e Fachini, Ortega e Louzada-Neto (2008) consideraram influência local nos modelos de riscos múltiplos.

4.1 Influência global

O modelo caso deleção para o modelo, equação (3), é dado por

$$y_l = \mathbf{x}_l^T \boldsymbol{\beta} + \sigma z_l, \quad l = 1, 2, \dots, n, \quad l \neq i. \quad (21)$$

Agora considere que uma quantidade com subscrito i significa a quantidade original com o i -ésimo caso deletado. Para o modelo (21), o logaritmo da função de verossimilhança de $\boldsymbol{\theta}$ excluído o i -ésimo caso é denotada por $l_{(i)}(\boldsymbol{\theta})$ e $\hat{\boldsymbol{\theta}}_{(i)}$ é o estimador de máxima verossimilhança de $\boldsymbol{\theta}$ a partir de $l_{(i)}(\boldsymbol{\theta})$. Para avaliar a influência do i -ésimo caso nas estimativas de máxima verossimilhança de $\hat{\boldsymbol{\theta}}$, a idéia básica é usar a diferença entre $\hat{\boldsymbol{\theta}}_{(i)}$ e $\hat{\boldsymbol{\theta}}$. Então, se $\hat{\boldsymbol{\theta}}_{(i)}$ está distante de $\hat{\boldsymbol{\theta}}$ o i -ésimo caso pode ser considerado como uma observação influente.

As medidas de diagnóstico utilizadas para avaliar a influência de um caso na análise dos dados, baseadas na deleção de casos, foram a distância de Cook generalizada (XIE; WEI, 2007) e o afastamento de verossimilhança (COOK; WEISBERG, 1982). Essas medidas são definidas, respectivamente, por

$$GD_i(\boldsymbol{\theta}) = (\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}})^T [\ddot{\mathbf{L}}(\boldsymbol{\theta})] (\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}), \quad (22)$$

$$LD_i(\boldsymbol{\theta}) = 2 \left[l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{(i)}) \right]. \quad (23)$$

4.2 Influência local

Influência local foi proposta por Cook (1986) com o objetivo de avaliar a mudança nos resultados da análise quando se incorporam “pequenas perturbações” ao modelo.

Seja $\boldsymbol{\omega}$ o vetor de perturbação de maneira que $\boldsymbol{\omega} \in \boldsymbol{\Omega} \subseteq \mathbf{R}^s$, $\boldsymbol{\Omega}$ é um conjunto aberto e $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ denota o estimador de máxima verossimilhança de $\boldsymbol{\theta}$ sob o modelo perturbado,

ou seja, o estimador de máxima verossimilhança baseado em $l(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})$, que é o logaritmo da função de verossimilhança para o modelo perturbado.

Cook (1986) sugere medidas de influência depois do modelo ou observações serem perturbados através do vetor $\boldsymbol{\omega}$. O enfoque de influência local compara $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ com $\hat{\boldsymbol{\theta}}$ em que similaridade entre estes vetores indica que as estimativas dos parâmetros são robustas com respeito à perturbação no modelo. Por outro lado, diferenças relevantes sugerem que as estimativas produzidas são sensíveis a tais perturbações.

A abordagem original baseia-se na medida de afastamento da verossimilhança definida por

$$LD(\boldsymbol{\omega}) = 2\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\},$$

em que $l(\boldsymbol{\theta})$ é o logaritmo da função de verossimilhança do modelo postulado.

Considere que existe o vetor que representa a ausência de perturbação, denotado por $\boldsymbol{\omega}_0 \in \Omega \subseteq \mathbf{R}^s$, ou seja, $l(\boldsymbol{\theta}_{\boldsymbol{\omega}_0}) = l(\boldsymbol{\theta})$, para todo $\boldsymbol{\theta} \in \Theta$ e que $l(\boldsymbol{\theta}_{\boldsymbol{\omega}})$ tem as duas primeiras derivadas contínuas em $(\hat{\boldsymbol{\theta}}^T, \boldsymbol{\omega}_0^T)$.

A idéia consiste em estudar a maior variação de LD em torno de $\boldsymbol{\omega}_0$. Com esse fim, considere em \mathbf{R}^{s+1} a seguinte superfície:

$$\alpha(\boldsymbol{\omega}) = \begin{pmatrix} \boldsymbol{\omega} \\ LD(\boldsymbol{\omega}) \end{pmatrix}, \quad (24)$$

que contém informação essencial sobre a influência da perturbação em questão e o comportamento dessa função é analisado em uma vizinhança do ponto mínimo local $\boldsymbol{\omega}_0$.

Para medir a sensibilidade do afastamento da verossimilhança, Cook (1986) utilizou a curvatura normal dada na equação (24) ao redor de $\boldsymbol{\omega}_0$ na direção de um vetor \mathbf{d} ($q \times 1$) de norma unitária. Cook (1986) mostrou que a curvatura normal da superfície dada na equação (24) na direção \mathbf{d} é dada por

$$C_{\mathbf{d}}(\boldsymbol{\theta}) = 2|\mathbf{d}^T \boldsymbol{\Delta}^T [\ddot{\mathbf{L}}(\boldsymbol{\theta})]^{-1} \boldsymbol{\Delta} \mathbf{d}|, \quad (25)$$

em que $\boldsymbol{\Delta}$ é a matriz $p \times n$ que depende do esquema de perturbação e cujos elementos são dados por

$$\Delta_{ji} = \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \theta_j \partial \omega_i}, \quad i = 1, \dots, n \text{ e } j = 1, \dots, p, \quad (26)$$

avaliados em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega}_0$.

Usando a equação (25), pode-se caracterizar a superfície α e, conseqüentemente, analisar a função $LD(\boldsymbol{\omega})$. Um procedimento que é freqüentemente empregado para análise é calcular a curvatura máxima $C_{\mathbf{d}_{max}}$ na direção correspondente \mathbf{d}_{max} que indica qual tipo de perturbação no modelo produz a maior mudança local no afastamento da verossimilhança, $LD(\boldsymbol{\omega})$. A curvatura normal, equação (25), assume seu valor máximo quando $\mathbf{d} = \mathbf{d}_{max}$, com \mathbf{d}_{max} representando o autovetor normalizado associado ao maior autovalor da matriz $\mathbf{B} = \boldsymbol{\Delta}^T \ddot{\mathbf{L}}(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}$. Os elementos de \mathbf{d}_{max} que são comparativamente discrepantes indicam que os respectivos elementos de $\boldsymbol{\omega}$ podem produzir mudanças substanciais nas estimativas de máxima verossimilhança e os respectivos pontos da amostra são possíveis pontos influentes. O gráfico dos elementos de \mathbf{d}_{max} contra a ordem das observações pode ser usado para revelar elementos influentes.

Por outro lado, Lesaffre e Verbeke (1998) sugerem calcular $C_{\mathbf{d}_i}(\boldsymbol{\theta})$ em que \mathbf{d}_i denota um vetor, $n \times 1$, de zeros com 1 na i -ésima posição que é denominada de influência local total. Nesse caso, a curvatura na direção \mathbf{d}_i assume a forma $C_i = 2|\boldsymbol{\Delta}_i^T \ddot{\mathbf{L}}(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}_i|$, em que $\boldsymbol{\Delta}_i^T$ denota o i -ésima linha de $\boldsymbol{\Delta}$. O gráfico dos C_i *versus* a ordem das observações pode ser usado para mostrar elementos influentes. Também é sugerido que as observações tais que $C_i > 2\bar{C}$, em que $\bar{C} = \frac{\sum_{i=1}^n C_i}{n}$, tenham uma atenção especial.

4.3 Esquemas de perturbação

As medidas de influência são calculadas depois de um esquema de perturbação ser adotado no logaritmo da função de verossimilhança. Diferentes esquemas de perturbação podem ser considerados de acordo com o objetivo principal da análise. A seguir apresenta-se o logaritmo da função de verossimilhança perturbada para três esquemas de perturbação. Considere o vetor de pesos sendo $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_n)^T$.

4.3.1 Perturbação de casos

Nesse caso, o logaritmo da função de verossimilhança perturbada incorpora diferentes pesos para as diferentes observações e é dado por

$$l(\boldsymbol{\theta}, \boldsymbol{\omega}) = \sum_{i=1}^n \omega_i l_i(\boldsymbol{\theta}),$$

em que $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$ e $0 \leq \omega_i \leq 1$.

4.3.2 Perturbação da resposta

Nesta seção, o interesse é detectar a sensibilidade do modelo quando a variável resposta é submetida a uma perturbação aditiva dada por

$$y_{iw} = y_i + \omega_i S_y,$$

em que S_y é um fator escala que pode ser o desvio padrão amostral de Y . Assim, o logaritmo da função de verossimilhança perturbada é dado substituindo y_i por y_{iw} em $l(\boldsymbol{\theta})$ e o vetor de não perturbação é $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$.

4.3.3 Perturbação da variável explicativa

Agora, o objetivo é avaliar a sensibilidade do modelo a pequenas perturbações em uma particular variável explicativa contínua, X_t . Considere uma perturbação aditiva para a variável explicativa do mesmo modo como foi feito na seção 4.3.2 para a variável resposta, isto é,

$$x_{it\omega} = x_{it} + \omega_i S_t,$$

em que S_t é um fator escala que pode ser o desvio padrão amostral da variável explicativa X_t .

Nesse caso, o logaritmo da função de verossimilhança perturbada é dado substituindo x_{it} por $x_{it\omega}$ e o vetor de não perturbação é $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$.

4.4 “Leverage” generalizado

Sejam $l(\boldsymbol{\theta})$ o logaritmo da função de verossimilhança do modelo postulado, $\hat{\boldsymbol{\theta}}$ o vetor dos estimadores de máxima verossimilhança de $\boldsymbol{\theta}$ e $\boldsymbol{\mu}$ o valor esperado de $\mathbf{Y} =$

$(Y_1, \dots, Y_n)^T$, então, $\hat{\mathbf{y}} = \boldsymbol{\mu}(\hat{\boldsymbol{\theta}})$ é o vetor das respostas previstas.

A idéia principal por trás do conceito de “leverage” (vide, por exemplo, COOK; WEISBERG, 1982; WEI; HU e FUNG, 1998) é avaliar a influência de y_i em seu próprio valor predito. Esta influência pode ser representada pela derivada $\frac{\partial \hat{y}_i}{\partial y_i}$. Wei; Hu e Fung (1998) mostraram que a matriz $\left(\frac{\partial \hat{\mathbf{y}}}{\partial \mathbf{y}}\right)$ de “leverage” generalizado, $n \times n$, pode ser expressa como

$$\mathbf{GL}(\hat{\boldsymbol{\theta}}) = \left[\mathbf{D}_{\boldsymbol{\theta}} \left(\ddot{\mathbf{L}}(\boldsymbol{\theta}) \right)^{-1} \ddot{\mathbf{L}}_{\boldsymbol{\theta}\mathbf{y}} \right] \quad (27)$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ em que $\mathbf{D}_{\boldsymbol{\theta}} = \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}}$ e $\ddot{\mathbf{L}}_{\boldsymbol{\theta}\mathbf{y}} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \mathbf{y}^T}$.

Para identificar pontos de alavanca, faz-se o gráfico dos elementos da diagonal de $\mathbf{GL}(\hat{\boldsymbol{\theta}})$ *versus* a ordem das observações.

4.5 Análise de sensibilidade para os modelos de regressão

Nesta seção, são apresentados os cálculos necessários para obtenção das medidas de diagnósticos apresentadas neste capítulo para os modelos de regressão log-Weibull estendida e log-Burr XII, descritos no capítulo 3. Para calcular as medidas de influência global foi usada a linguagem matricial de programação Ox.

4.5.1 Esquemas de perturbação para o modelo log-Weibull estendida

Para o modelo de regressão log-Weibull estendida, equação (10), o logaritmo da função de verossimilhança é dado na equação (13). Nesse contexto, a seguir apresentam-se os elementos da matriz $\boldsymbol{\Delta}$ definida na equação (26), denotada por $(\boldsymbol{\Delta}_1, \dots, \boldsymbol{\Delta}_{p+2})^T$, para os três esquemas de perturbação apresentados na seção 4.3. Por outro lado, nota-se da expressão do “leverage” generalizado definido na equação (27) que é necessário obter o valor esperado da distribuição log-Weibull estendida e isto não foi possível, portanto o “leverage” generalizado para o modelo de regressão log-Weibull estendida não foi utilizado neste trabalho.

4.5.1.1 Perturbação de casos

No esquema de perturbação de casos, o logaritmo da função de verossimilhança perturbada é dado por

$$\begin{aligned}
l(\boldsymbol{\theta}, \boldsymbol{\omega}) &= [\log(\lambda) - \log(\sigma)] \sum_{i \in F} \omega_i + \sum_{i \in F} \omega_i \left[\mathbf{x}_i^T \boldsymbol{\beta} + z_i + \exp(z_i) \right] + \lambda \sum_{i \in F} \omega_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \\
&\quad - \lambda \sum_{i \in F} \omega_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \exp \left[\exp(z_i) \right] + \lambda \sum_{i \in C} \omega_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \\
&\quad - \lambda \sum_{i \in C} \omega_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \exp \left[\exp(z_i) \right], \tag{28}
\end{aligned}$$

em que $0 \leq \omega_i \leq 1$ e $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$. Usando (28) e depois de manipulações algébricas os elementos do vetor $\boldsymbol{\Delta}_1$ são dados por

$$\Delta_{1i} = \begin{cases} \hat{\lambda}^{-1} + \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) - \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \exp \left[\exp(\hat{z}_i) \right] & \text{se } i \in F \\ \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) - \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \exp \left[\exp(\hat{z}_i) \right] & \text{se } i \in C, \end{cases}$$

os elementos do vetor $\boldsymbol{\Delta}_2$ são

$$\Delta_{2i} = \begin{cases} -\hat{\sigma}^{-1} - \hat{z}_i \hat{\sigma}^{-1} - \hat{z}_i \hat{\sigma}^{-1} \exp(\hat{z}_i) + \hat{\lambda} \hat{z}_i \hat{\sigma}^{-1} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right] & \text{se } i \in F \\ \hat{\lambda} \hat{z}_i \hat{\sigma}^{-1} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right] & \text{se } i \in C \end{cases}$$

e os elementos do vetor $\boldsymbol{\Delta}_j$ para $j = 3, \dots, p+2$ são

$$\Delta_{ji} = \begin{cases} x_{ij} - x_{ij} \hat{\sigma}^{-1} - x_{ij} \hat{\sigma}^{-1} \exp(\hat{z}_i) + \hat{\lambda} x_{ij} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \times \\ \left\{ 1 - \exp \left[\exp(\hat{z}_i) \right] + \hat{\sigma}^{-1} \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right] \right\} & \text{se } i \in F \\ \hat{\lambda} x_{ij} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \left\{ 1 - \exp \left[\exp(\hat{z}_i) \right] + \hat{\sigma}^{-1} \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right] \right\} & \text{se } i \in C. \end{cases}$$

4.5.1.2 Perturbação da resposta

Quando a perturbação da resposta é usada, o logaritmo da função de verossimilhança perturbada é expresso por

$$\begin{aligned}
l(\boldsymbol{\theta}, \boldsymbol{\omega}) &= r[\log(\lambda) - \log(\sigma)] + \sum_{i \in F} \left[\mathbf{x}_i^T \boldsymbol{\beta} + z_i^* + \exp(z_i^*) \right] + \lambda \sum_{i \in F} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \\
&\quad - \lambda \sum_{i \in F} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \exp \left[\exp(z_i^*) \right] + \lambda \sum_{i \in C} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \\
&\quad - \lambda \sum_{i \in C} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \exp \left[\exp(z_i^*) \right], \tag{29}
\end{aligned}$$

em que $z_i^* = \frac{(y_i + \omega_i S_y) - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}$ e $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$. Usando a equação (29) e depois de manipulações algébricas os elementos do vetor $\boldsymbol{\Delta}_1$ são dados por

$$\Delta_{1i} = \begin{cases} -\exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) S_y \hat{\sigma}^{-1} \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right] & \text{se } i \in F \\ -\exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) S_y \hat{\sigma}^{-1} \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right] & \text{se } i \in C, \end{cases}$$

os elementos do vetor $\boldsymbol{\Delta}_2$ são

$$\Delta_{2i} = \begin{cases} -S_y \sigma^{-2} \left[1 + \exp(\hat{z}_i) + \hat{z}_i \exp(\hat{z}_i) \right] + \hat{\lambda} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) S_y \hat{\sigma}^{-2} \hat{g}_i \times \\ \left[1 + \hat{z}_i + \hat{z}_i \exp(\hat{z}_i) \right] & \text{se } i \in F \\ \hat{\lambda} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) S_y \hat{\sigma}^{-2} \hat{g}_i \left[1 + \hat{z}_i + \hat{z}_i \exp(\hat{z}_i) \right] & \text{se } i \in C, \end{cases}$$

e os elementos do vetor $\boldsymbol{\Delta}_j$, para $j = 3, \dots, p + 2$, são expressos por

$$\Delta_{ji} = \begin{cases} -x_{ij} S_y \hat{\sigma}^{-2} \exp(\hat{z}_i) - \hat{\lambda} x_{ij} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) S_y \hat{\sigma}^{-1} \hat{g}_i \left[1 - \hat{\sigma}^{-1} - \hat{\sigma}^{-1} \exp(\hat{z}_i) \right] & \text{se } i \in F \\ -\hat{\lambda} x_{ij} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) S_y \hat{\sigma}^{-1} \hat{g}_i \left[1 - \hat{\sigma}^{-1} - \hat{\sigma}^{-1} \exp(\hat{z}_i) \right] & \text{se } i \in C, \end{cases}$$

em que $\hat{g}_i = \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right]$.

4.5.1.3 Perturbação da variável explicativa

Para o esquema de perturbação da variável explicativa, o logaritmo da função de verossimilhança perturbada é dado por

$$\begin{aligned}
l(\boldsymbol{\theta}\boldsymbol{\omega}) &= r[\log(\lambda) - \log(\sigma)] + \sum_{i \in F} \left[\mathbf{x}_i^{*T} \boldsymbol{\beta} + z_i^* + \exp(z_i^*) \right] + \lambda \sum_{i \in F} \exp(\mathbf{x}_i^{*T} \boldsymbol{\beta}) \\
&\quad - \lambda \sum_{i \in F} \exp(\mathbf{x}_i^{*T} \boldsymbol{\beta}) \exp \left[\exp(z_i^*) \right] + \lambda \sum_{i \in C} \exp(\mathbf{x}_i^{*T} \boldsymbol{\beta}) \\
&\quad - \lambda \sum_{i \in C} \exp(\mathbf{x}_i^{*T} \boldsymbol{\beta}) \exp \left[\exp(z_i^*) \right], \tag{30}
\end{aligned}$$

em que $z_i^* = \frac{y_i - \mathbf{x}_i^{*T} \boldsymbol{\beta}}{\sigma}$, $\mathbf{x}_i^{*T} \boldsymbol{\beta} = \beta_1 + \beta_2 x_{i2} + \dots + \beta_t (x_{it} + \omega_i S_t) + \dots + \beta_p x_{ip}$ e $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$.

Usando a equação (30) e depois de manipulações algébricas os elementos do vetor $\boldsymbol{\Delta}_1$ são dados por

$$\Delta_{1i} = \begin{cases} S_x \hat{\beta}_t \exp(\hat{\mathbf{x}}_i^T \hat{\boldsymbol{\beta}}) \left\{ 1 - \exp \left[\exp(\hat{z}_i) \right] + \hat{\sigma}^{-1} \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right] \right\} & \text{se } i \in F \\ S_x \hat{\beta}_t \exp(\hat{\mathbf{x}}_i^T \hat{\boldsymbol{\beta}}) \left\{ 1 - \exp \left[\exp(\hat{z}_i) \right] + \hat{\sigma}^{-1} \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right] \right\} & \text{se } i \in C, \end{cases}$$

os elementos do vetor $\boldsymbol{\Delta}_2$ são

$$\Delta_{2i} = \begin{cases} S_x \hat{\beta}_t \hat{\sigma}^{-2} \left[1 + \exp(\hat{z}_i) + \hat{z}_i \exp(\hat{z}_i) \right] + \hat{\lambda} S_x \hat{\beta}_t \hat{\sigma}^{-1} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \hat{g}_i \times \\ \left[\hat{z}_i - \hat{\sigma}^{-1} - \hat{\sigma}^{-1} \hat{z}_i - \hat{\sigma}^{-1} \hat{z}_i \exp(\hat{z}_i) \right] & \text{se } i \in F \\ \hat{\lambda} S_x \hat{\beta}_t \hat{\sigma} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \hat{g}_i \left[\hat{z}_i - \hat{\sigma}^{-1} - \hat{\sigma}^{-1} \hat{z}_i - \hat{\sigma}^{-1} \hat{z}_i \exp(\hat{z}_i) \right] & \text{se } i \in C, \end{cases}$$

os elementos do vetor Δ_j , para $j = 3, \dots, p + 2$ e $j \neq t$, assumem a seguinte expressão

$$\Delta_{ji} = \begin{cases} x_{ij} S_x \hat{\beta}_t \left\{ \sigma^{-2} \exp(\hat{z}_i) + \hat{\lambda} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) - \hat{\lambda} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \exp \left[\exp(\hat{z}_i) \right] - \right. \\ \left. - \hat{\lambda} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \hat{\sigma}^{-1} \hat{g}_i \left[-2 + \hat{\sigma}^{-1} + \hat{\sigma}^{-1} \exp(\hat{z}_i) \right] \right\} & \text{se } i \in F \\ x_{ij} S_x \hat{\beta}_t \hat{\lambda} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \left\{ 1 - \exp \left[\exp(\hat{z}_i) \right] - \hat{\sigma}^{-1} \hat{g}_i \times \right. \\ \left. \left[-2 + \hat{\sigma}^{-1} + \hat{\sigma}^{-1} \exp(\hat{z}_i) \right] \right\} & \text{se } i \in C \end{cases}$$

e os elementos do vetor Δ_t são dados por

$$\Delta_{ti} = \begin{cases} S_x \left[1 - \hat{\sigma}^{-1} - \hat{\sigma}^{-1} \exp(\hat{z}_i) \right] + \hat{\lambda} S_x \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \left\{ 1 - \exp \left[\exp(\hat{z}_i) \right] \times \right. \\ \left. \left[-1 - \hat{\sigma}^{-1} \exp(\hat{z}_i) \right] \right\} + x_{it} \hat{\beta}_t S_x \left\{ \exp(\hat{z}_i) \hat{\sigma}^{-2} + \hat{\lambda} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \times \right. \\ \left. \left\{ 1 - \exp \left[\exp(\hat{z}_i) \right] \right\} - \hat{\lambda} \hat{\sigma}^{-1} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \hat{g}_i \left[-2 + \hat{\sigma}^{-1} + \hat{\sigma}^{-1} \exp(\hat{z}_i) \right] \right\} & \text{se } i \in F \\ \hat{\lambda} S_x \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \left\{ 1 - \exp \left[\exp(\hat{z}_i) \right] \left[1 - \hat{\sigma}^{-1} \exp(\hat{z}_i) \right] \right\} + x_{it} \hat{\beta}_t S_x \times \\ \left\{ \exp(\hat{z}_i) - \hat{\lambda} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \exp \left[\exp(\hat{z}_i) \right] - \right. \\ \left. \hat{\lambda} \hat{\sigma}^{-1} \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \hat{g}_i \left[-2 + \hat{\sigma}^{-1} + \hat{\sigma}^{-1} \exp(\hat{z}_i) \right] \right\} & \text{se } i \in C, \end{cases}$$

em que $\hat{g}_i = \exp(\hat{z}_i) \exp \left[\exp(\hat{z}_i) \right]$.

4.5.2 Esquemas de perturbação para o modelo de regressão log-Burr XII

Para o modelo de regressão log-Burr XII, equação (17), o logaritmo da função de verossimilhança é dado na equação (19). Nesse contexto, a seguir apresentam-se os elementos da matriz Δ definida em (26), denotada por $(\Delta_1, \dots, \Delta_{p+2})^T$, para os três esquemas de

perturbação apresentados na seção 4.3, bem como apresentam-se os cálculos necessários para a obtenção do “leverage” generalizado definido na equação (27).

4.5.2.1 Perturbação de casos

No esquema de perturbação de casos, o logaritmo da função de verossimilhança perturbada é dado por

$$l(\theta_{\omega}) = [\log(k) - \log(\sigma)] \sum_{i \in F} \omega_i + \sum_{i \in F} \omega_i z_i + (k+1) \sum_{i \in F} \omega_i \log [1 + \exp(z_i)] - \sum_{i \in C} \omega_i \log [1 + \exp(z_i)], \quad (31)$$

em que $0 \leq \omega_i \leq 1$ e $\omega_0 = (1, \dots, 1)^T$. Usando a equação (31) e depois de manipulações algébricas os elementos do vetor Δ_1 são dados por

$$\Delta_{1i} = \begin{cases} \hat{k}^{-1} + \log [1 + \exp(\hat{z}_i)] & \text{if } i \in F \\ \log [1 + \exp(\hat{z}_i)] & \text{se } i \in C, \end{cases}$$

os elementos do vetor Δ_2 são

$$\Delta_{2i} = \begin{cases} -\hat{\sigma}^{-1} \left\{ 1 + \hat{z}_i + (\hat{k} + 1) \hat{z}_i \exp(\hat{z}_i) [1 + \exp(z_i)]^{-1} \right\} & \text{se } i \in F \\ \hat{k} \hat{\sigma}^{-1} \hat{z}_i \exp(\hat{z}_i) [1 + \exp(z_i)]^{-1} & \text{se } i \in C \end{cases}$$

e os elementos do vetor Δ_j , para $j = 3, \dots, p+2$, podem ser expressos por

$$\Delta_{ji} = \begin{cases} -x_{ij} \hat{\sigma}^{-1} \left\{ 1 + (\hat{k} + 1) \exp(\hat{z}_i) [1 + \exp(\hat{z}_i)]^{-1} \right\} & \text{se } i \in F \\ x_{ij} \hat{k} \hat{\sigma}^{-1} \exp(z_i) [1 + \exp(z_i)]^{-1} & \text{se } i \in C. \end{cases}$$

4.5.2.2 Perturbação da resposta

Quando o esquema de perturbação da resposta é adotado, o logaritmo da função de verossimilhança perturbada é dado por

$$l(\theta_{\omega}) = r[\log(k) - \log(\sigma)] + \sum_{i \in F} z_i^* - (k+1) \sum_{i \in F} \log [1 + \exp(z_i^*)] - k \sum_{i \in C} \log [1 + \exp(z_i^*)], \quad (32)$$

em que $z_i^* = \frac{(y_i + \omega_i S_y) - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}$. Usando a equação (32) e depois de manipulações algébricas os elementos do vetor $\boldsymbol{\Delta}_1$ são dados por

$$\Delta_{1i} = \begin{cases} -S_y \hat{\sigma}^{-1} \hat{z}_i [1 + \exp(z_i)]^{-1} & \text{se } i \in F \\ -S_y \hat{\sigma}^{-1} \hat{z}_i [1 + \exp(z_i)]^{-1} & \text{se } i \in C, \end{cases}$$

elementos do vetor $\boldsymbol{\Delta}_2$ são

$$\Delta_{2i} = \begin{cases} -S_y \hat{\sigma}^{-2} \left\{ 1 - (\hat{k} + 1) \exp(\hat{z}_i) [1 + \exp(\hat{z}_i)]^{-1} \left(\hat{z}_i [1 + \exp(\hat{z}_i)]^{-1} + 1 \right) \right\} & \text{se } i \in F \\ S_y \hat{k} \hat{\sigma}^{-2} \exp(\hat{z}_i) [1 + \exp(\hat{z}_i)]^{-1} \left\{ \hat{z}_i [1 + \exp(\hat{z}_i)]^{-1} + 1 \right\} & \text{se } i \in C \end{cases}$$

e os elementos do vetor $\boldsymbol{\Delta}_j$, para $j = 3, \dots, p + 2$, podem ser expressos como

$$\Delta_{ji} = \begin{cases} x_{ij} S_y (\hat{k} + 1) \hat{\sigma}^{-2} \exp(\hat{z}_i) [1 + \exp(\hat{z}_i)]^{-2} & \text{se } i \in F \\ x_{ij} S_y \hat{k} \hat{\sigma}^{-2} \exp(\hat{z}_i) [1 + \exp(\hat{z}_i)]^{-2} & \text{se } i \in C. \end{cases}$$

4.5.2.3 Perturbação da variável explicativa

Quando o esquema de perturbação da variável explicativa é adotado, o logaritmo da função de verossimilhança perturbada é dado pela expressão

$$\begin{aligned} l(\boldsymbol{\theta}, \boldsymbol{\omega}) &= r[\log(k) - \log(\sigma)] + \sum_{i \in F} z_i^* - (k + 1) \sum_{i \in F} \log[1 + \exp(z_i^*)] \\ &\quad - k \sum_{i \in C} \log[1 + \exp(z_i^*)], \end{aligned} \quad (33)$$

em que $z_i^* = \frac{y_i - \mathbf{x}_i^{*T} \boldsymbol{\beta}}{\sigma}$ e $\mathbf{x}_i^{*T} = \beta_1 + \beta_2 x_{i2} + \dots + \beta_t (x_{it} + \omega_i S_t) + \dots + \beta_p x_{ip}$. Usando a equação (33) e depois de manipulações algébricas os elementos do vetor $\boldsymbol{\Delta}_1$ são dados por

$$\Delta_{1i} = \begin{cases} S_x \hat{\beta}_t \hat{\sigma}^{-1} \exp(\hat{z}_i) [1 + \exp(\hat{z}_i)]^{-1} & \text{se } i \in F \\ S_x \hat{\beta}_t \hat{\sigma}^{-1} \exp(\hat{z}_i) [1 + \exp(\hat{z}_i)]^{-1} & \text{se } i \in C, \end{cases}$$

os elementos do vetor $\boldsymbol{\Delta}_2$ são,

$$\Delta_{2i} = \begin{cases} \hat{\beta}_t S_x \hat{\sigma}^{-2} \left\{ 1 - (k + 1) \exp(\hat{z}_i) [1 + \exp(\hat{z}_i)]^{-1} \left[1 + z_i (1 + \exp(\hat{z}_i))^{-1} \right] \right\} & \text{se } i \in F \\ -\hat{\beta}_t \hat{k} S_x \hat{\sigma}^{-2} \exp(\hat{z}_i) [1 + \exp(\hat{z}_i)]^{-1} \left[1 + \hat{z}_i (1 + \exp(\hat{z}_i))^{-1} \right] & \text{se } i \in C, \end{cases}$$

os elementos do vetor Δ_j , para $j = 3, \dots, p + 2$ e $j \neq t$, assumem a seguinte expressão

$$\Delta_{ji} = \begin{cases} -x_{ij}S_x\beta_t(\hat{k} + 1)\sigma^{-2} \exp(\hat{z}_i)[1 + \exp(\hat{z}_i)]^{-2} & \text{se } i \in F \\ -x_{ij}S_x\beta_t\hat{k}\sigma^{-2} \exp(\hat{z}_i)[1 + \exp(\hat{z}_i)]^{-2} & \text{se } i \in C \end{cases}$$

e os elementos do vetor Δ_t são dados por

$$\Delta_{ti} = \begin{cases} S_x\hat{\sigma}^{-1} + (\hat{k} + 1)S_x\sigma^{-1} \exp\{z_i\}[1 + \exp(\hat{z}_i)]^{-1} [x_{it}\hat{\beta}_t - 1] & \text{se } i \in F \\ \hat{k}S_x\sigma^{-1} \exp(\hat{z}_i)[1 + \exp(\hat{z}_i)]^{-1} [x_{it}\hat{\beta}_t - 1] & \text{se } i \in C. \end{cases}$$

4.5.3 “Leverage” generalizado para o modelo de regressão log-Burr XII

A matriz $\mathbf{GL}(\hat{\boldsymbol{\theta}})$ definida na equação (27) para o modelo log-Burr XII, equação (17), apresenta os seguintes elementos:

$$\mathbf{D}_{\boldsymbol{\theta}} = \begin{cases} \frac{\partial E(Y)}{\partial k} = -\frac{1}{\hat{c}}\Psi'(\hat{k}) \\ \frac{\partial E(Y)}{\partial \sigma} = \Psi(1) - \Psi(\hat{k}) \\ \frac{\partial E(Y)}{\partial \beta_j} = x_{ij} \quad j = 0, 1, \dots, p, \end{cases}$$

em que $\Psi(\cdot)$ é a função digama (LAWLESS, 2003) e os componentes da matriz $\mathbf{D}_{\boldsymbol{\theta}}$ são dados por

$$\begin{aligned} \ddot{\mathbf{L}}_{ky_i} &= \begin{cases} -\hat{\sigma}^{-1}\hat{h}_i & \text{se } i \in F \\ -\hat{\sigma}^{-1} \exp(\hat{h}_i) & \text{se } i \in C, \end{cases} \\ \ddot{\mathbf{L}}_{\hat{\sigma}y_i} &= \begin{cases} \hat{\sigma}^{-2} \left\{ -1 + (\hat{k} + 1)\hat{h}_i[1 + \hat{z}_i + \exp(\hat{z}_i)][1 + \exp(\hat{z}_i)]^{-1} \right\} & \text{se } i \in F \\ \hat{\sigma}^{-2}\hat{k}\hat{h}_i[1 + \hat{z}_i + \exp(\hat{z}_i)][1 + \exp(\hat{z}_i)]^{-1} & \text{se } i \in C, \end{cases} \\ \ddot{\mathbf{L}}_{\hat{\beta}y_i} &= \begin{cases} x_{ij}\hat{\sigma}^{-2}(\hat{k} + 1)\hat{h}_i[1 + \exp(\hat{z}_i)]^{-1} & \text{se } i \in F \\ x_{ij}\hat{\sigma}^{-2}\hat{k}\hat{h}_i[1 + \exp(\hat{z}_i)]^{-1} & \text{se } i \in C, \end{cases} \end{aligned}$$

em que $\hat{h}_i = \exp(\hat{z}_i)[1 + \exp(\hat{z}_i)]^{-1}$.

5 ANÁLISE DE RESÍDUOS

Quando se adota um modelo de regressão e este é ajustado a um conjunto de dados, a avaliação desse ajuste pode ser realizada por meio da análise dos resíduos. A partir dessa análise, identificam-se observações discrepantes, além de avaliar se existem afastamentos sérios das suposições feitas ao modelo. Se o modelo for adequado, os gráficos de resíduos *versus* a ordem das observações ou os valores preditos devem apresentar um comportamento aleatório em torno do zero.

De uma maneira geral, pode-se definir o resíduo referente à i -ésima observação por meio de uma função qualquer não trivial (COX; SNELL, 1968)

$$\hat{e}_i = g(y_i, \hat{\boldsymbol{\theta}}, \mathbf{x}_i),$$

em que $\hat{\boldsymbol{\theta}}$ é o estimador de máxima verossimilhança de $\boldsymbol{\theta}$ e $(y_i, \delta_i, \mathbf{x}_i)$, $i = 1, \dots, n$, são os dados de sobrevivência.

Existem diversos resíduos propostos na literatura de análise de sobrevivência (COLOSIMO; GIOLO, 2006; LAWLESS, 2003; COLLETT, 2003). Entre esses destacam-se os resíduos: Martingale e tipo Martingale. Para aplicações desses resíduos, veja, por exemplo, Ortega, Paula e Bolfarine (2008) que propuseram um resíduo para o modelo de regressão log-gama generalizada para dados censurados e Leiva et al. (2007) que estudaram resíduos para o modelo de regressão log-Birnbaum-Saunders com dados censurados.

5.1 Resíduo martingale

Este resíduo foi introduzido em processos de contagem, para mais detalhes leia Fleming e Harrington (1991), e pode ser escrito como

$$r_{M_i} = \delta_i + \log[S(y_i, \hat{\boldsymbol{\theta}})], \quad (34)$$

em que $\delta_i = 0$ denota observações censuradas, $\delta_i = 1$ observações não censuradas e $S(y_i, \hat{\boldsymbol{\theta}})$ é a função de sobrevivência para $Y = \log(T)$. Esse resíduo pode ser usado para identificar a melhor forma funcional para uma dada covariável no modelo de regressão, além de auxiliar na identificação de pontos atípicos (COLOSIMO; GIOLO, 2006).

A distribuição de r_{M_i} é assimétrica, tem máximo no valor $+1$ e mínimo no valor $-\infty$ e, portanto, o resíduo Martingale não é simetricamente distribuído ao redor do zero (COLLETT, 2003).

5.2 Resíduo tipo martingale

Em modelos lineares generalizados, o resíduo componente do desvio (resíduo *deviance*) é dado por

$$r_{D_i} = \text{sinal}(y_i - \mu_i) \sqrt{2} \{l_i(y_i; \tilde{\mu}_i) - l_i(y_i; \hat{\mu}_i)\}^{\frac{1}{2}},$$

em que $l_i(y_i; \cdot)$ é a contribuição de y_i no logaritmo da função de verossimilhança, $\tilde{\mu}_i$ é o estimador de máxima verossimilhança para μ_i baseado unicamente em y_i e $\hat{\mu}_i$ é o estimador de máxima verossimilhança para μ_i sob o modelo proposto. Para mais detalhes ver, por exemplo, McCullagh e Nelder (1989) e Tableman e Kim (2004).

Therneau, Grambsch e Fleming (1990) introduziram o resíduo componente do desvio em processos de contagem e este resíduo para os modelos de riscos proporcionais com variáveis explicativas não dependendo do tempo é definido como

$$r_{D_i} = \text{sign}(r_{M_i}) \left\{ -2 \left[r_{M_i} + \delta_i \log(\delta_i - r_{M_i}) \right] \right\}^{\frac{1}{2}}, \quad (35)$$

em que r_{M_i} é o resíduo martingale.

Esse resíduo pode ser considerado como uma modificação feita no resíduo martingale com a finalidade de torná-lo mais simétrico em torno do zero. Nesse contexto, o resíduo definido na equação (35) é denotado por resíduo tipo Martingale.

Gráficos dos resíduos martingale e tipo martingale fornecem uma maneira de verificar a adequação do modelo ajustado e a presença de observações atípicas. A seguir são descritos esses resíduos para os modelos de regressão propostos na seção 3.3.

5.3 Resíduos para o modelo de regressão log-Weibull estendida

Os resíduos martingale, equação (34), e tipo martingale, equação (35), para o modelo log-Weibull estendida, equação (10), são, respectivamente, dados por

$$r_{M_i} = \begin{cases} 1 + \lambda\alpha \left\{ 1 - \exp \left[\left(\frac{t}{\alpha} \right)^\tau \right] \right\} & \text{se } i \in F \\ \lambda\alpha \left\{ 1 - \exp \left[\left(\frac{t}{\alpha} \right)^\tau \right] \right\} & \text{se } i \in C, \end{cases}$$

$$r_{D_i} = \begin{cases} \operatorname{sinal} \left\{ 1 + \hat{\lambda} \exp(\hat{\mu}) \left[1 - \exp(\exp(\hat{z}_i)) \right] \right\} (2)^{\frac{1}{2}} \left\{ -1 - \hat{\lambda} \exp(\hat{\mu}) \times \right. \\ \left. \left[1 - \exp(\exp(\hat{z}_i)) \right] - \log \left\{ -\hat{\lambda} \exp(\hat{\mu}) \left[1 - \exp(\exp(\hat{z}_i)) \right] \right\} \right\}^{\frac{1}{2}} & \text{se } i \in F \\ \operatorname{sinal} \left\{ \hat{\lambda} \exp(\hat{\mu}) \left[1 - \exp(\exp(\hat{z}_i)) \right] \right\} (2)^{\frac{1}{2}} \times \\ \left\{ -\hat{\lambda} \exp(\hat{\mu}) \left[1 - \exp(\exp(\hat{z}_i)) \right] \right\}^{\frac{1}{2}} & \text{se } i \in C. \end{cases}$$

5.3.1 Estudo de simulação

Um estudo de simulação de Monte Carlo com 1.000 amostras geradas foi realizado para investigar o comportamento da distribuição empírica dos resíduos martingale; seção (5.3), e tipo martingale; seção (5.3) do modelo de regressão log-Weibull estendida (10).

Os tempos de sobrevivência T_1, \dots, T_n foram gerados, por meio da simulação de Monte Carlo, da distribuição Weibull estendida, função de densidade dada na equação (1), considerando as seguintes reparametrizações $\tau = \frac{1}{\sigma}$ e $\alpha = \exp(\mu)$. Simularam-se amostras de tamanhos $n = 50$ e $n = 100$, para valores de $\sigma = 2$ (relacionado a dados com função de taxa de falha em forma de U) e $0,4$ (relacionado a dados com função de taxa de falha crescente), $\lambda = 0, 1$ e assumindo $\mu_i = \beta_0 + \beta_1 x_i$, com x_i sendo gerada da distribuição uniforme $(0, 1)$ e β_0 e β_1 fixados. Os tempos de censura denotados por E_1, \dots, E_n foram gerados da distribuição uniforme $(0, \eta)$, em que η foi ajustado até as porcentagens de censura 10 ou 30 serem atingidas. Os tempos de sobrevivência considerados em cada ajuste foram obtidos de $t_i = \min\{E_i, T_i\}$ e o indicador de censura é $\delta_i = I_{T_i \leq E_i}$, sendo $I(\cdot)$ a função indicadora.

Para cada combinação de n, λ, σ e porcentagens de censuras, amostras aleatórias foram geradas e para cada uma ajustou-se o modelo de regressão log-Weibull estendida na

equação (10) com $\mu_i = \beta_0 + \beta_1 x_i$ (programa no apêndice E). As Figuras 6-7 apresentam o gráfico dos resíduos *versus* os valores esperados das estatísticas de ordem da distribuição normal padrão; esse é o gráfico normal de probabilidades que serve para avaliar o afastamento da suposição de normalidade para esse resíduo (WEISBERG, 1985). O apêndice E também apresenta um programa feito no *software* R para obter o gráfico de probabilidade normal para as amostras geradas. Desses gráficos as seguintes interpretações são obtidas:

- i. Pode-se observar que a distribuição empírica do resíduo tipo martingale apresenta concordância com a distribuição normal padrão.
- ii. Quando a porcentagem de censura na amostra diminui ou o tamanho da amostra aumenta, a distribuição empírica do resíduo tipo martingale apresenta uma melhor concordância com a distribuição normal padrão.
- iii. A distribuição empírica do resíduo tipo martingale quando a função de taxa de falha é crescente ($\sigma = 0,4$) apresenta uma melhor concordância com a distribuição normal padrão do que quando a função de taxa de falha tem forma de U ($\sigma = 2,0$).

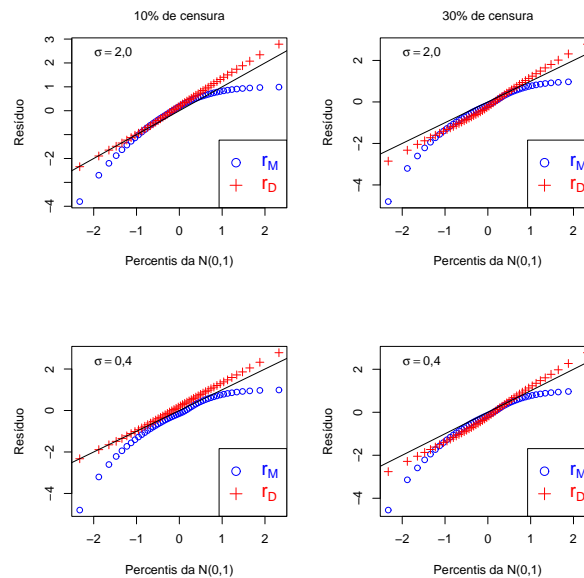


Figura 6 – Gráfico normal de probabilidades para os resíduos baseados no martingale para o modelo de regressão log-Weibull estendida. Tamanho da amostra $n=50$, porcentagens de censura = 10 e 30, parâmetros $\sigma=2$ e $0,4$ e $\lambda = 0, 1$

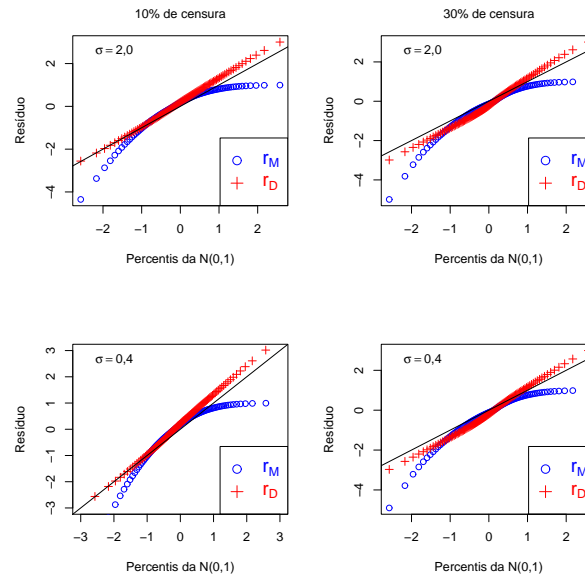


Figura 7 – Gráfico normal de probabilidades para os resíduos baseados no martingale para o modelo de regressão log-Weibull estendida. Tamanho da amostra $n=100$, porcentagens de censura = 10 e 30, parâmetros $\sigma=2$ e 0,4 e $\lambda = 0, 1$

Portanto, o gráfico de probabilidade normal para o resíduo tipo martingale com envelope (ATKINSON, 1987) pode ser usado para auxiliar na verificação de afastamento da suposição para a distribuição do erro referente ao modelo de regressão log-Weibull estendida, bem como na identificação de pontos atípicos. Esse gráfico pode ser produzido das seguinte maneira: (i) ajuste o modelo e simule uma amostra de n observações independentes usando este modelo como se fosse o modelo verdadeiro; (ii) ajuste o modelo para as amostras geradas usando (δ_i, \mathbf{x}_i) do conjunto de dados, e calcule os valores do resíduo; (iii) repita os passos (i) e (ii) m vezes; (iv) ordene os valores dos resíduos, $r_{(i)v}^*$, $i = 1, \dots, n$ e $v = 1, \dots, m$; (v) considere os n conjuntos das m estatísticas de ordem; para cada conjunto calcule a média, o mínimo e o máximo; (vi) faça o gráfico desses valores e dos resíduos ordenados da amostra original contra os quantis da distribuição normal padrão. O valores do máximo e do mínimo das m estatísticas de ordem produzem o envelope. Observações correspondentes aos resíduos fora dos limites dados pelo simulado envelope devem ser investigadas. Além disso, se uma porção considerável de pontos situam-se fora do envelope, então existe evidência de que o modelo ajustado não é adequado.

5.4 Resíduos para o modelo de regressão log-Burr XII

Os resíduos martingale, equação (34), e tipo martingale, equação (35), para o modelo log-Burr XII na equação (17) são, respectivamente, dados por

$$r_{M_i} = \begin{cases} 1 + \log \left\{ \left[1 + \left(\frac{t}{s} \right)^c \right]^{-k} \right\} & \text{se } i \in F \\ \log \left\{ \left[1 + \left(\frac{t}{s} \right)^c \right]^{-k} \right\} & \text{se } i \in C, \end{cases}$$

$$r_{D_i} = \begin{cases} - \left\{ -2 \left[-\hat{k} \log(1 + \exp(\hat{z}_i)) + \log(1 + \hat{k} \log(1 + \exp(\hat{z}_i))) \right] \right\}^{\frac{1}{2}} & \text{se } i \in F \\ \text{sinal} \left\{ 1 - \hat{k} \log[1 + \exp(\hat{z}_i)] \right\} \left\{ -2 + 2\hat{k} \log[1 + \exp(\hat{z}_i)] \right\}^{\frac{1}{2}} & \text{se } i \in C. \end{cases}$$

5.4.1 Estudo de simulação

Para investigar o comportamento da distribuição empírica dos resíduos martingale; seção (5.4), e tipo martingale; seção (5.4) do modelo de regressão log-Burr XII, equação (17), foi realizado um estudo de simulação de Monte Carlo com 1.000 amostras geradas que é descrito a seguir.

Os tempos de sobrevivência T_1, \dots, T_n foram gerados, por meio da simulação de Monte Carlo, da distribuição Burr XII, função de densidade dada na equação (2), considerando as seguintes reparametrizações $c = \frac{1}{\sigma}$ e $s = \exp(\mu)$. Simularam-se amostras de tamanhos $n = 50$ e $n = 100$, para valores de $\sigma = 0,5$ e $0,8$ (relacionados a dados com função de taxa de falha unimodal), $k = 0,27, 1,00$ e $2,00$ e assumindo $\mu_i = \beta_0 + \beta_1 x_i$, com x_i sendo gerada da distribuição uniforme $(0, 1)$ e β_0 e β_1 fixados. Os tempos de censura denotados por E_1, \dots, E_n foram gerados da distribuição uniforme $(0, \eta)$, em que η foi ajustado até as porcentagens de censura, 0, 10 e 30 serem atingidas. Os tempos de sobrevivência considerados em cada ajuste foram obtidos de $t_i = \min\{E_i, T_i\}$ e o indicador de censura é $\delta_i = I_{T_i \leq E_i}$, sendo $I(\cdot)$ a função indicadora.

Para cada combinação de n , k , σ e porcentagens de censuras, foram geradas amostras aleatórias e para cada uma ajustou-se o modelo de regressão log-Burr XII, equação (17), com $\mu_i = \beta_0 + \beta_1 x_i$. As Tabelas 5-6 fornecem um sumário numérico das simulações.

Pode-se notar que o r_{M_i} apresenta curtose que difere de 3 e assimetria diferente de zero. Então, a forma da distribuição de r_{M_i} é assimétrica, tem valor máximo +1 e valor mínimo $-\infty$. Por outro lado, o resíduo r_{D_i} tem assimetria negligenciável, pequena curtose e desvio padrão aproximadamente 1, mas não tem média aproximadamente 0. Assim, o resíduo r_{D_i} não apresenta concordância com a distribuição normal padrão devido, unicamente, devido ao fato de sua média não ser zero. Além disso, a distribuição empírica desse resíduo apresenta concordância similar com a distribuição normal padrão para valores de $\sigma < 1, 0$, para cada n , k , σ e porcentagem de censuras fixados.

As Figuras 8-11 apresentam o gráfico dos resíduos *versus* os valores esperados das estatísticas de ordem da distribuição normal padrão; esse é o gráfico normal de probabilidades que serve para avaliar o afastamento da suposição de normalidade para esse resíduo (WEISBERG, 1985). Essas figuras confirmam as tendências observadas nas Tabelas 5-6.

A partir disso, é proposto uma mudança no resíduo tipo martingale para observações não censuradas com o intuito de obter média aproximadamente zero.

Tabela 5 – Estatísticas descritivas para o resíduo martingale no modelo de regressão log-Burr XII

σ	k	média	Resíduo martingale (r_M)					
			0% de censura		10% de censura		30% de censura	
			$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$
0,50	0,27	Média	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
		Desvio Padrão	0,9779	0,9887	0,8774	0,8892	0,6797	0,6863
		Assimetria	-1,7962	-1,9025	-1,2478	-1,4025	-0,7258	-0,9190
		Curtose	7,4183	8,1942	4,5822	5,7999	2,8766	3,9953
	1,00	Média	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
		Desvio Padrão	0,9744	0,9902	0,8808	0,8911	0,6791	0,6917
		Assimetria	-1,7237	-1,8778	-1,1385	-1,4522	-0,8608	-0,8200
		Curtose	6,6780	7,8461	4,8780	5,6602	3,1532	3,0237
	2,00	Média	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
		Desvio Padrão	0,9793	0,9913	0,8811	0,8927	0,6820	0,6937
		Assimetria	-1,7311	-1,8669	-1,3912	-1,4983	-0,8742	-0,8807
		Curtose	6,5616	7,6331	5,0676	5,8461	3,1151	3,2560
0,80	0,27	Média	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
		Desvio Padrão	0,9764	0,9895	0,8748	0,8872	0,6793	0,6862
		Assimetria	-1,7912	-1,9080	-1,2692	-1,3372	-0,7083	-0,7100
		Curtose	7,3948	8,2353	4,8867	5,4181	2,8978	3,0452
	1,00	Média	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
		Desvio Padrão	0,9743	0,9904	0,8797	0,8916	0,6805	0,6904
		Assimetria	-1,7230	-1,8799	-1,3035	-1,4246	-0,7295	-0,7460
		Curtose	6,6731	7,8619	4,7302	5,6377	2,6435	2,8459
	2,00	Média	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
		Desvio Padrão	0,9791	0,9913	0,8814	0,8924	0,6825	0,6924
		Assimetria	-1,7302	-1,8672	-1,3299	-1,4451	-0,7386	-0,8326
		Curtose	6,5575	7,6363	4,7674	5,6021	2,5888	3,2125

Tabela 6 – Estatísticas descritivas para o resíduo tipo martingale no modelo de regressão log-Burr XII

σ	k	Média	Resíduo tipo martingale (r_D)					
			0% de censura		10% de censura		30% de censura	
			$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$
0,50	0,27	Média	0,3405	0,3408	0,2713	0,2721	0,1435	0,1480
		Desvio Padrão	1,0340	1,0352	1,2006	1,1829	1,2415	1,1635
		Assimetria	-0,0186	-0,0201	-0,18690	-0,1267	0,2820	0,3967
		Curtose	3,0743	3,1418	3,8918	3,7468	3,4828	3,2464
	1,00	Média	0,3404	0,3413	0,2770	0,2773	0,1589	0,1528
		Desvio Padrão	1,0327	1,0358	1,1217	1,1257	1,0908	1,1880
		Assimetria	-0,0247	-0,0260	-0,0588	-0,0539	0,3526	0,2212
		Curtose	3,0240	3,1292	3,4946	3,5723	3,0836	3,5948
	2,00	Média	0,3407	0,3414	0,2790	0,2789	0,1612	0,1564
		Desvio Padrão	1,0332	1,0358	1,0881	1,1002	1,0793	1,1498
		Assimetria	-0,0300	-0,0285	0,0117	0,0013	0,3338	0,2517
		Curtose	3,0408	3,1303	3,2538	3,4227	3,1005	3,5015
0,80	0,27	Média	0,3402	0,3407	0,2699	0,2702	0,1366	0,1346
		Desvio Padrão	1,0332	1,0348	1,1931	1,2115	1,2755	1,3173
		Assimetria	-0,0172	-0,0217	-0,1100	-0,1814	0,3390	0,2883
		Curtose	3,0702	3,1412	3,6318	3,8860	3,4694	3,6413
	1,00	Média	0,3404	0,3413	0,2734	0,0039	0,1441	0,1420
		Desvio Padrão	1,0327	1,0359	1,1532	1,0761	1,2158	1,2604
		Assimetria	-0,0241	-0,0258	-0,0719	0,0239	0,3167	0,2566
		Curtose	3,0253	3,1313	3,5148	3,1586	3,4139	3,6398
	2,00	Média	0,3407	0,3414	0,2746	0,2749	0,1461	0,1458
		Desvio Padrão	1,0332	1,0358	1,1345	1,1386	1,2033	1,2090
		Assimetria	-0,0298	-0,0285	-0,0416	-0,0322	0,2984	0,3153
		Curtose	3,0406	3,1299	3,4366	3,5178	3,4308	3,5002

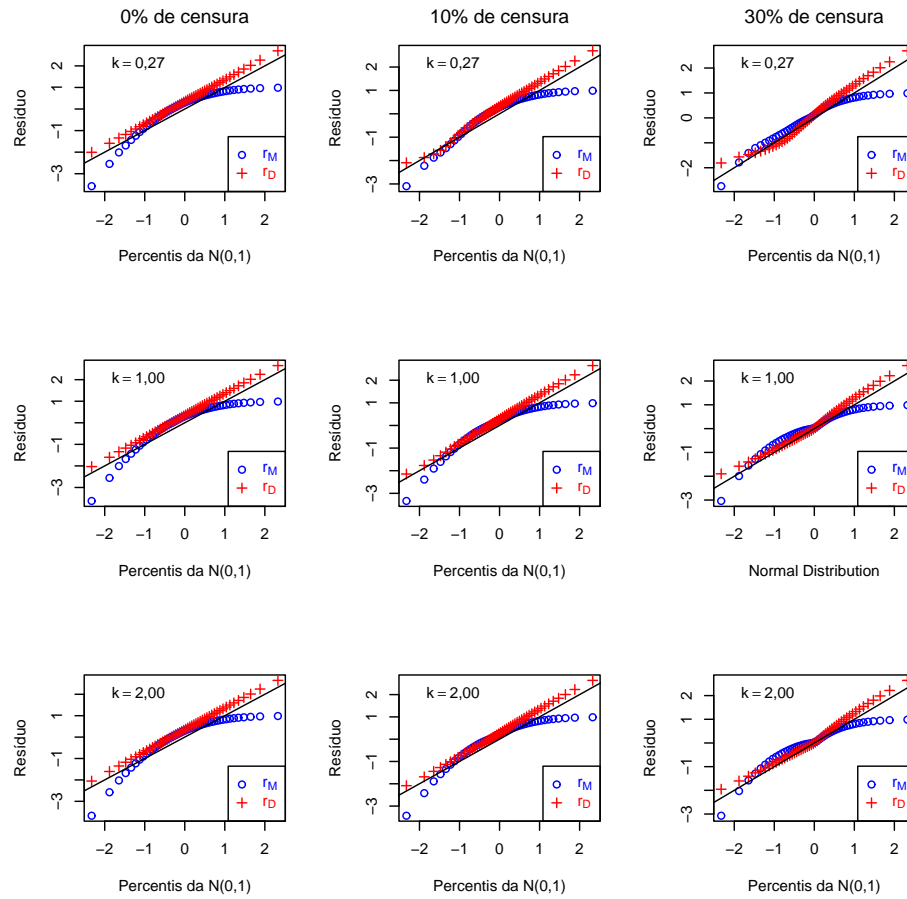


Figura 8 – Gráfico normal de probabilidade para resíduos baseados no martingale: resíduo martingale (r_M) e resíduo tipo martingale (r_D) para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0,27, 1, 2$ e $\sigma = 0,5$

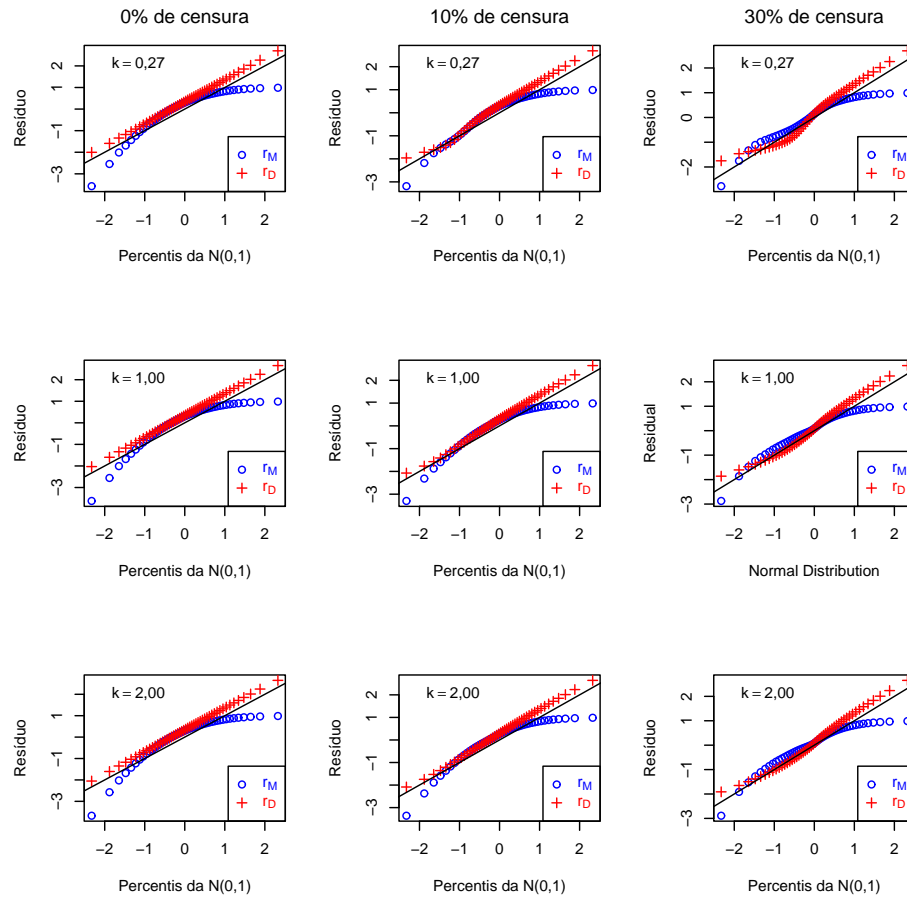


Figura 9 – Gráfico normal de probabilidade para resíduos baseados no martingale: resíduo martingale (r_M) e resíduo tipo martingale (r_D) para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0,27, 1, 2$ e $\sigma = 0,8$

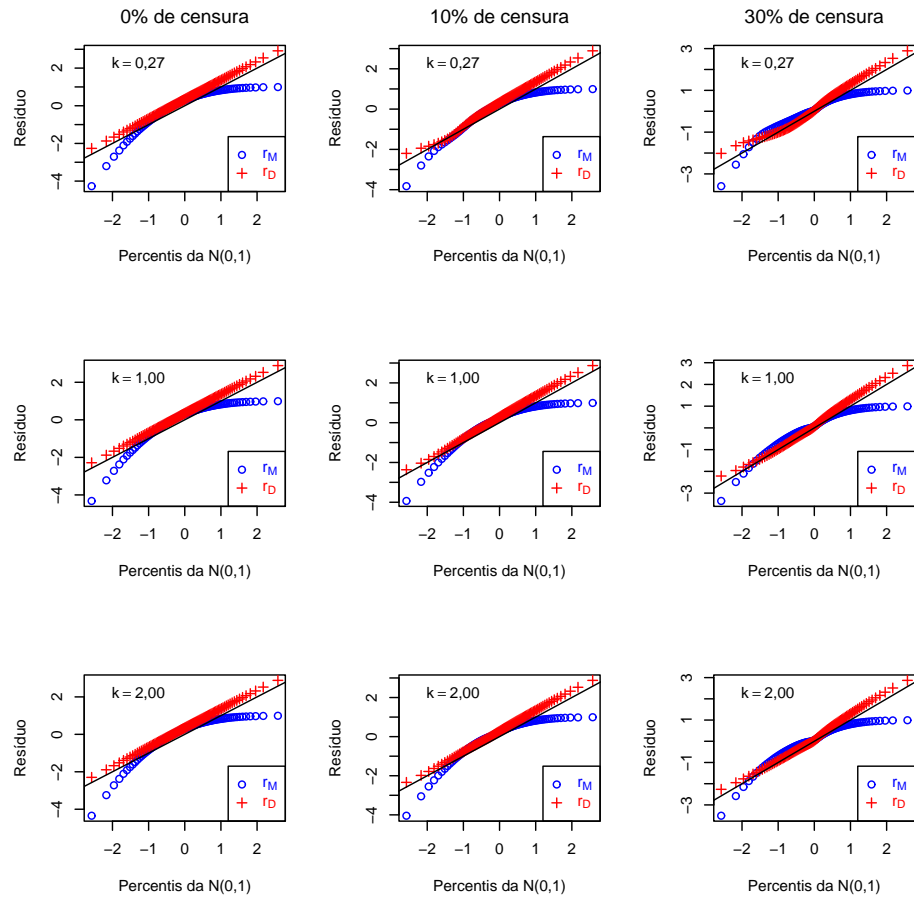


Figura 10 – Gráfico normal de probabilidade para resíduos baseados no martingale: resíduo martingale (r_M) e resíduo tipo martingale (r_D) para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 100$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0,27, 1, 2$ e $\sigma = 0,5$

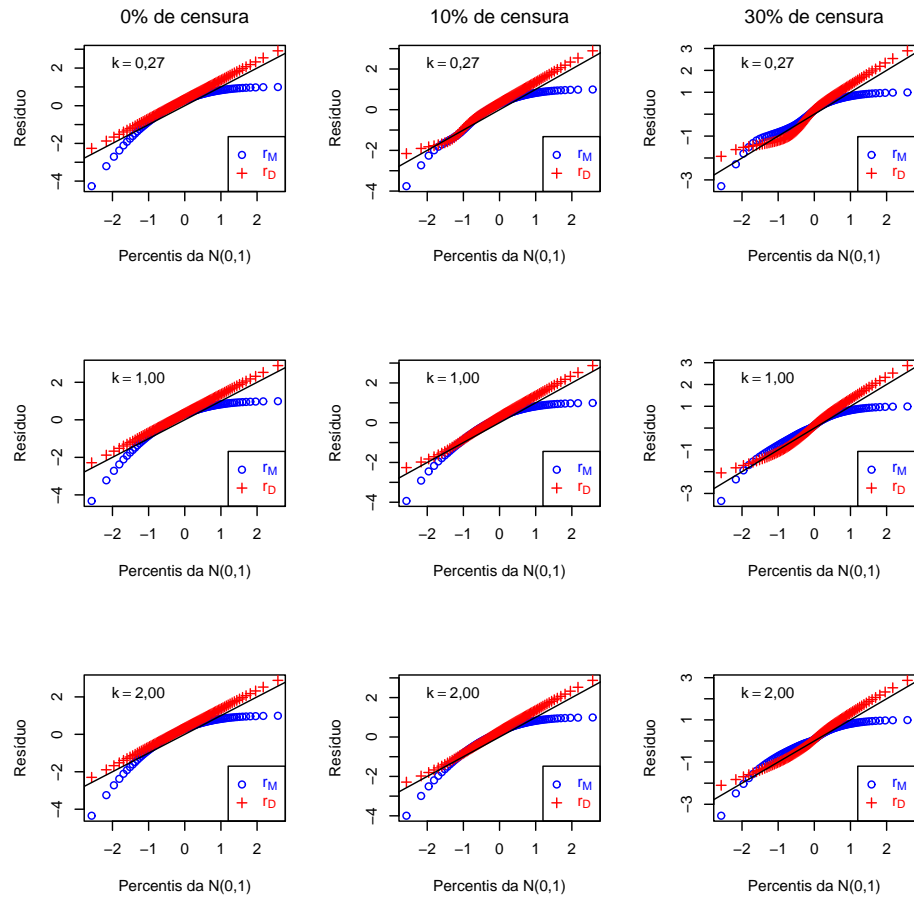


Figura 11 – Gráfico normal de probabilidade para resíduos baseados no martingale: resíduo martingale (r_M) e resíduo tipo martingale (r_D) para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 100$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0,27, 1, 2$ e $\sigma = 0,8$

5.4.2 Resíduo tipo martingale modificado

O resíduo tipo martingale, seção 5.4, foi modificado subtraindo o valor esperado do resíduo tipo martingale, que baseado em estudos de simulação é de aproximadamente 0,34. Assim, o resíduo tipo martingale modificado é dado por

$$r_{MD_i} = \begin{cases} \text{sinal}\{1 - \hat{k} \log[1 + \exp(\hat{z}_i)]\} \times \\ \left\{ -2 \left[1 - \hat{k} \log(1 + \exp(\hat{z}_i)) + \log[1 - \hat{k} \log(1 + \exp(\hat{z}_i))] \right] \right\}^{\frac{1}{2}} - 0.34, & \text{if } i \in F, \\ \text{sinal}\{ -\hat{k} \log[1 + \exp(\hat{z}_i)] \} \left[2\hat{k} \log(1 + \exp(\hat{z}_i)) \right]^{\frac{1}{2}} & \text{if } i \in C. \end{cases}$$

5.4.2.1 Estudo de simulação

Para o resíduo tipo martingale modificado foi realizado um estudo de simulação como conduzido na seção 5.4.1. Os resultados que estão na Tabela 7 indicam que esse resíduo tem uma concordância superior com a distribuição normal padrão que o resíduo tipo martingale.

Resumidamente, pode-se extrair as seguintes interpretações das simulações:

- Observa-se que a distribuição empírica do resíduo tipo martingale modificado apresenta concordância com a distribuição normal padrão para cada n , k , σ e porcentagem de censura.
- Quando a porcentagem de censura diminui, a distribuição empírica do resíduo tipo martingale modificado apresenta uma melhor concordância com a distribuição normal padrão.
- A distribuição empírica do resíduo tipo martingale modificado apresenta concordância similar com a distribuição normal padrão para valores de $\sigma < 1.0$, para cada n , k , σ e porcentagem de censura fornecido.

As Figuras 12-15 apresentam o gráfico do resíduo *versus* os valores esperados das estatísticas de ordem da distribuição normal padrão; esse é o gráfico normal de probabi-

lidades que serve para avaliar o afastamento da suposição de normalidade para esse resíduo (WEISBERG, 1985). Essas figuras confirmam as tendências observadas na Tabela 7.

Baseado nos resultados das simulações, o gráfico de probabilidade normal para r_{MD_i} com envelope simulado como sugerido por Atkinson (1981) pode ser usado para auxiliar na verificação de afastamento da suposição para a distribuição do erro referente ao modelo de regressão log-Burr XII, bem como na identificação de pontos atípicos. Os passos para a construção desse gráfico estão descritos na sub-seção 5.3.1.

Tabela 7 – Estatísticas descritivas para o resíduo tipo martingale modificado no modelo de regressão log-Burr XII

σ	k	Média	Resíduo tipo martingale modificado (r_{MD})					
			0% de censura		10% de censura		30% de censura	
			$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$
0,50	0,27	Média	0,0005	0,0008	-0,0346	-0,0339	-0,0945	-0,0900
		Desvio Padrão	1,0340	1,0352	1,0917	1,0819	1,0407	0,9916
		Assimetria	-0,0186	-0,0201	-0,0131	0,0107	0,2965	0,3379
		Curtose	3,0743	3,1418	3,1772	3,1596	2,7256	2,6794
	1,00	Média	0,0004	0,0013	-0,0290	-0,0287	-0,0792	-0,0852
		Desvio Padrão	1,0327	1,0358	1,0423	1,0459	0,9426	1,0047
		Assimetria	-0,0247	-0,0260	0,0194	0,0241	0,2747	0,2332
		Curtose	3,0240	3,1292	3,0320	3,1112	2,5660	2,8119
	2,00	Média	0,0007	0,0014	-0,0270	-0,0271	-0,0768	-0,0816
		Desvio Padrão	1,0332	1,0358	1,0216	1,0303	0,9348	0,9803
		Assimetria	-0,0300	-0,0285	0,0422	0,0438	0,2564	0,2362
		Curtose	3,0408	3,1303	2,9407	3,0690	2,5781	2,7885
0,80	0,27	Média	0,0002	0,0007	-0,0361	-0,0358	-0,1015	-0,1034
		Desvio Padrão	1,0332	1,0348	1,0885	1,1020	1,0639	1,0910
		Assimetria	-0,0172	-0,0217	0,0288	-0,0058	0,3492	0,3329
		Curtose	3,0702	3,1412	3,0643	3,1894	2,7248	2,8112
	1,00	Média	0,0004	0,0013	-0,0326	-0,0321	-0,0939	-0,0960
		Desvio Padrão	1,0327	1,0359	1,0629	1,0665	1,0239	1,0529
		Assimetria	-0,0241	-0,0258	0,0279	0,0319	0,3087	0,2875
		Curtose	3,0253	3,1313	3,0264	3,1137	2,7051	2,8235
	2,00	Média	0,0007	0,0014	-0,0314	-0,0311	-0,0918	-0,0922
		Desvio Padrão	1,0332	1,0358	1,0511	1,0550	1,0153	1,202
		Assimetria	-0,0298	-0,0285	0,0360	0,0439	0,2904	0,3040
		Curtose	3,0406	3,1299	3,0050	3,0939	2,7147	2,7908

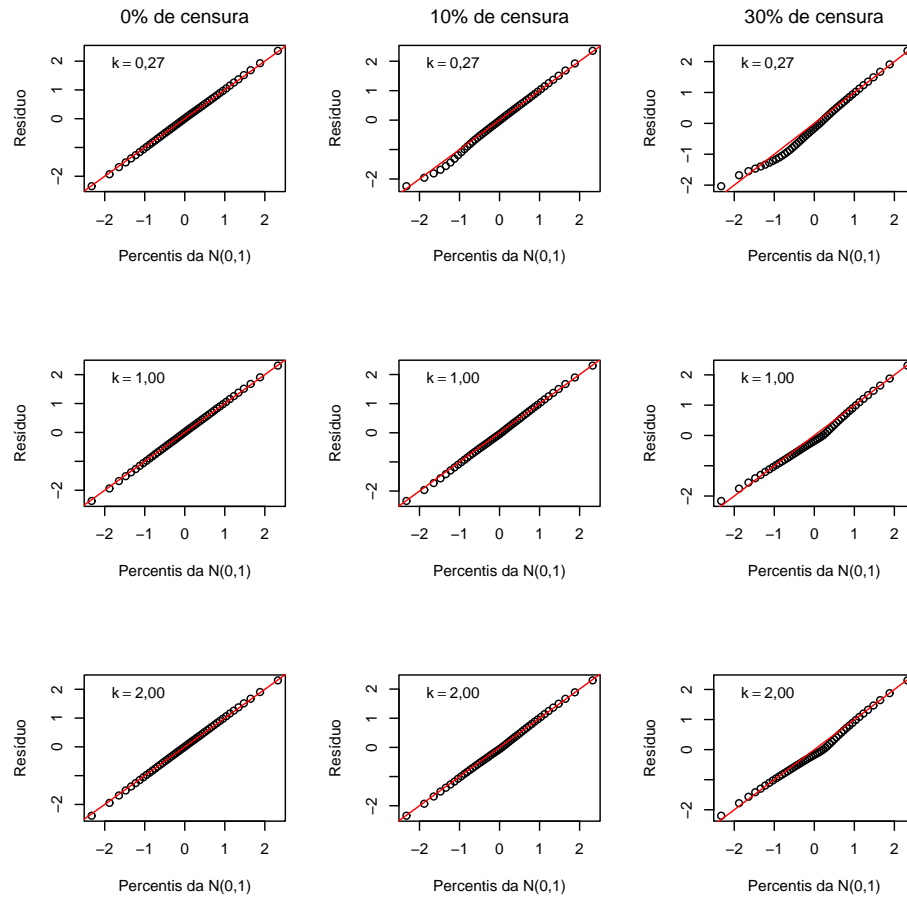


Figura 12 – Gráfico normal de probabilidade para resíduo tipo martingale modificado para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0,27, 1,2$ e $\sigma = 0,5$

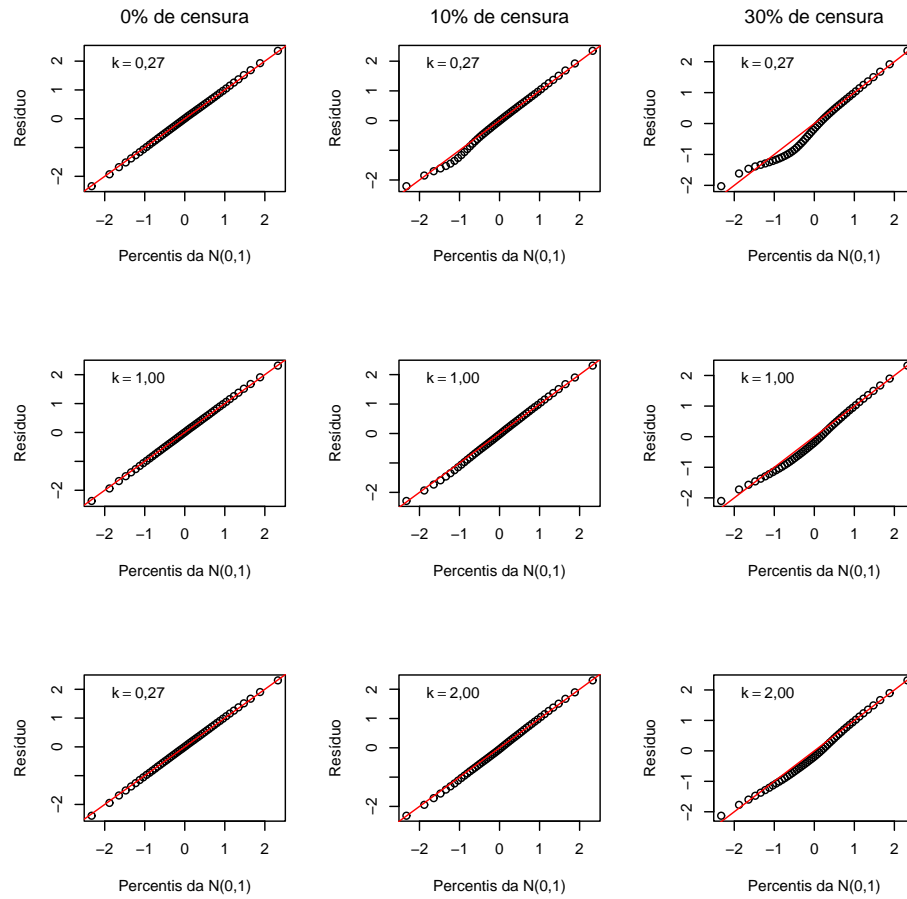


Figura 13 – Gráfico normal de probabilidade para resíduo tipo martingale modificado para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0,27, 1,2$ e $\sigma = 0,8$

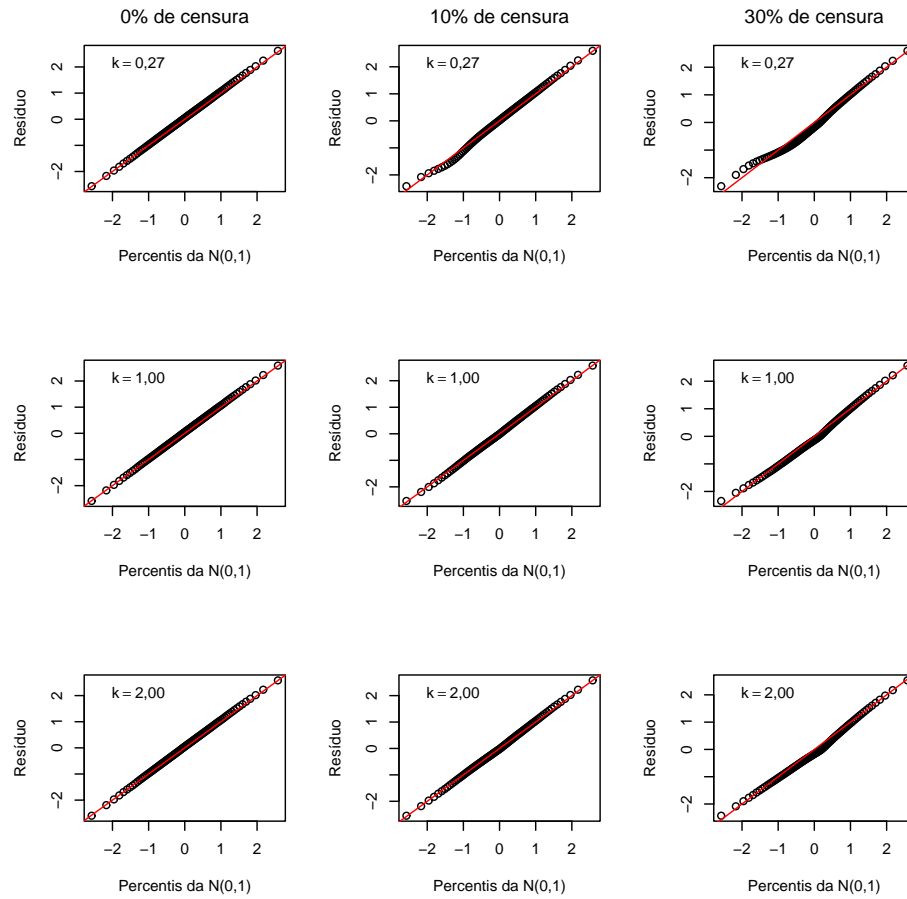


Figura 14 – Gráfico normal de probabilidade para resíduo tipo martingale modificado para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0,27, 1,2$ e $\sigma = 0,5$

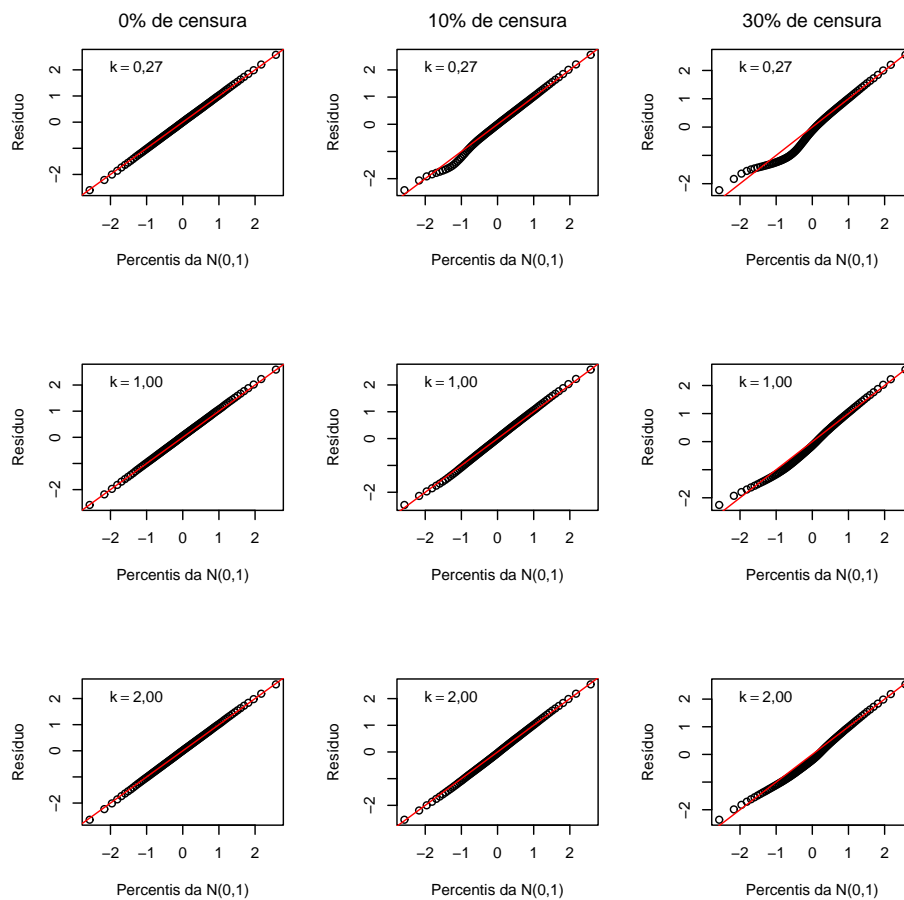


Figura 15 – Gráfico normal de probabilidade para resíduo tipo martingale modificado para o modelo de regressão log-Burr XII. Tamanho da amostra $n = 50$, porcentagens de censura = 0, 10 e 30 e valores de parâmetros $k = 0,27, 1,2$ e $\sigma = 0,8$

6 APLICAÇÕES

As aplicações descritas a seguir têm como objetivo mostrar a utilidade dos modelos de regressão log-Weibull estendida e log-Burr XII. São apresentadas quatro aplicações baseadas em dados reais com censura à direita.

As duas primeiras aplicações foram obtidas de Prentice (1973) e se referem ao conjunto de dados de pacientes com câncer de pulmão avançado e inoperável. Seguindo sugestão de Kalbfleisch e Prentice (2002) esse conjunto deve ser analisado separadamente para dois grupos, um formado por pacientes que receberam terapia prévia antes do início do estudo e outro por pacientes que não receberam terapia prévia. Bennett (1983) reanalisou os dados dos pacientes que não receberam terapia prévia usando o modelo de regressão log-logística e Lawlees (2003) usou o modelo de regressão log-Weibull para reanalisar o outro grupo.

A primeira aplicação tem como objetivo usar o modelo de regressão log-Weibull estendida como uma alternativa ao modelo de regressão log-Weibull no conjunto de dados dos pacientes que receberam terapia prévia. Por outro lado, a segunda aplicação apresenta um exemplo em que o modelo de regressão log-Burr pode melhorar o ajuste dado pelo modelo de regressão log-logística.

As outras duas aplicações apresentam função de taxa de falha não-monótona, sendo que na primeira esta função tem forma de U e na segunda é unimodal. Dessa forma, esses conjuntos de dados foram usados para ilustrar a aplicabilidade dos modelos propostos na seção (3.3). Em ambas aplicações seguiu-se a mesma rotina. Utilizou-se os três métodos de estimação descritos na seção 3.2 e, em seguida, foram realizadas análises de sensibilidade e de resíduos apresentadas nos capítulos 4 e 5, respectivamente.

6.1 Análise dos dados de câncer de pulmão em pacientes que receberam terapia prévia

Os dados se referem ao estudo de $n = 40$ pacientes com câncer de pulmão em estado avançado e inoperável, dos quais 21 receberam o tratamento padrão e 19 o tratamento em teste. O objetivo principal do estudo é comparar o efeito dos dois tratamentos de quimioterapia (padrão ou teste) sobre o tempo de sobrevivência (t), em dias, a partir do início

do tratamento. Das $n = 40$ observações dos tempos de sobrevivência 3 foram censuradas.

Existem outros fatores que também podem influenciar o tempo de sobrevivência, estes fatores são a *performance status* (x_1), que é uma medida do estado médico do paciente dada no diagnóstico em uma escala de 10 a 90, em que 10, 20, 30 significa completamente hospitalizado, 40, 50, 60 confinamento parcial no hospital e 70, 80, 90 capaz de cuidar de si mesmo, a idade do paciente em anos (x_2), o número de meses desde o diagnóstico de câncer de pulmão a entrada no estudo (x_3) e os tumores que foram classificados em quatro tipos: grande, adeno, pequeno e escamoso. Assim, o que foi estudado é o efeito das variáveis independentes (tratamento e fatores) registradas durante o experimento no tempo de sobrevivência.

Para identificar um modelo apropriado para o tempo de sobrevivência desse conjunto de dados foi utilizada a curva TTT, definida na seção 2.4, que é apresentada na Figura 16. Essa figura deixa dúvida sobre a forma empírica da função de taxa de falha pois indica uma função de taxa de falha na forma de U ou decrescente. Então, para esse conjunto de dados pode-se utilizar o modelo de regressão log-Weibull estendida com parâmetro de forma maior que 1 ou modelo de regressão log-Weibull com parâmetro de forma menor que 1.

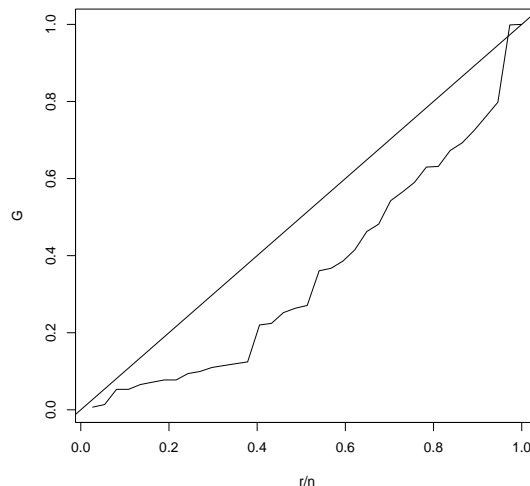


Figura 16 – Curva TTT para os dados câncer do pulmão em pacientes que receberam terapia prévia

As variáveis x_1 , x_2 e x_3 foram centradas como em Lawless (2003). Portanto, um

possível modelo de locação escala, apresentado no capítulo 3, é dado por

$$y_i = \beta_0 + \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_2) + \beta_3(x_{i3} - \bar{x}_3) + \sum_{p=4}^7 \beta_p x_{ip} + \sigma z_i, \quad i = 1, \dots, 40, \quad (36)$$

em que

$x_{i4} = 1$ se o tipo de tumor é escamoso, 0 caso contrário;

$x_{i5} = 1$ se o tipo de tumor é pequeno, 0 caso contrário;

$x_{i6} = 1$ se o tipo de tumor é adeno, 0 caso contrário;

$x_{i7} = 0$ se o tratamento é o teste, 1 caso contrário.

Lawless (2003) usou o modelo de regressão log-Weibull para analisar esse conjunto de dados. Uma possível alternativa a esse modelo é usar o modelo log-Weibull estendida, proposto na seção 3.3.1.1. Do estudo de simulação realizado nessa seção, tem-se que esse modelo pode melhorar de maneira significativa o ajuste dado pela distribuição Weibull.

Para examinar o desempenho dos modelos de regressão log-Weibull e log-Weibull estendida, ambos modelos foram ajustados aos dados de câncer de pulmão em pacientes que receberam terapia prévia e o modelo mais adequado foi selecionado por meio da medida AIC_{SUR} , descrito na seção 3.3.1.1. Usando os modelos Weibull e Weibull estendida, $AIC_{SUR} = 138,43$ (Weibull) e $128,05$ (Weibull estendida), respectivamente. A partir desses valores, nota-se que o modelo log-Weibull estendida é mais apropriado para esses dados. A Tabela 8 mostra as estimativas de máxima verossimilhança e os respectivos erros padrão para os parâmetros dos dois modelos.

As estimativas dos parâmetros de regressão e os respectivos erros padrão são diferentes para os dois modelos e, conseqüentemente, as conclusões obtidas por ambos modelos são diferentes. Usando um nível de significância de 5%, se o modelo de regressão log-Weibull estendida é ajustado as variáveis x_1 , x_2 , x_3 e x_6 são significativas, por outro lado, nenhuma variável é significativa se o modelo de regressão log-Weibull é ajustado. Além disso, para alguns parâmetros os erros padrão dos parâmetros para o modelo de regressão log-Weibull estendida são maiores que esses valores para o modelo de regressão log-Weibull.

Tabela 8 – Estimativas de máxima verossimilhança dos parâmetros dos modelos de regressão log-Weibull estendida e log-Weibull para os dados de câncer de pulmão em pacientes que receberam terapia prévia

Parâmetros	Weibull Estendida			Weibull		
	Estimativas	E.P.	p-valor	Estimativas	E.P.	p-valor
λ	0,0020	0,0007	-	-	-	-
σ	1,3151	0,2797	-	0,8728	0,1152	-
β_0	4,3305	0,8474	< 0,0001	4,7398	0,3605	< 0,0001
β_1	0,0648	0,0065	< 0,0001	0,0538	0,0096	< 0,0001
β_2	0,0432	0,0180	0,0161	0,0097	0,0175	0,5788
β_3	0,0973	0,0210	< 0,0001	0,0044	0,0104	0,6932
β_4	3,3805	1,9710	0,0863	0,3995	0,3945	0,3112
β_5	-0,5811	0,2978	0,0511	-0,1317	0,4252	0,7568
β_6	-2,3097	0,5050	< 0,0001	-0,8809	0,5134	0,0862
β_7	-0,4693	0,3205	0,1431	-0,2570	0,3465	0,4583

Nota: E.P.=Erro Padrão

Foi realizada uma análise de resíduos, que é útil para avaliar a adequabilidade dos modelos ajustados, por meio de gráficos do resíduo tipo martingale, dado na seção 5.3. As Figuras 17-18 apresentam, para ambos os modelos, os gráficos dos resíduos contra os valores ajustados e o gráfico normal de probabilidades com envelope simulado, respectivamente.

O gráfico do resíduo tipo martingale *versus* os valores ajustados para o modelo de regressão log-Weibull, Figura 17(a), sugere que os resíduos não estão dispostos aleatoriamente em torno do zero e que o caso 1 apresenta valor do resíduo maior que 3. Essa observação não é censurada e apresenta um alto valor para o tempo de sobrevivência. Adicionalmente, da Figura 18(a) observa-se que a maioria dos pontos localiza-se fora do envelope. Por outro lado, da análise de resíduo para o modelo de regressão log-Weibull estendida observa-se na Figura 17(b) um comportamento aleatório em torno do zero dos resíduos e a Figura 18(b) mostra que todos os pontos situam-se dentro do envelope. Portanto, o modelo de regressão log-Weibull estendida parece ser mais adequado para modelar os dados de câncer em pacientes

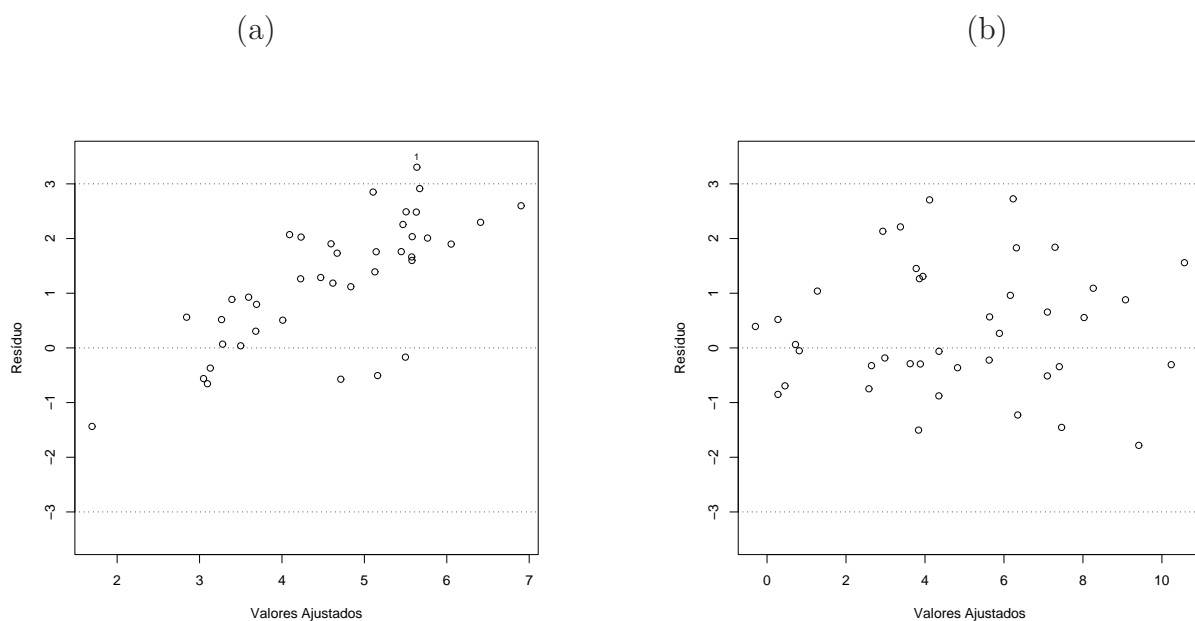


Figura 17 – Gráfico do resíduo tipo martingale contra os valores ajustados nos dados de sobrevivência de câncer de pulmão em pacientes que receberam terapia prévia para os modelos de regressão (a) log-Weibull (b) log-Weibull estendida

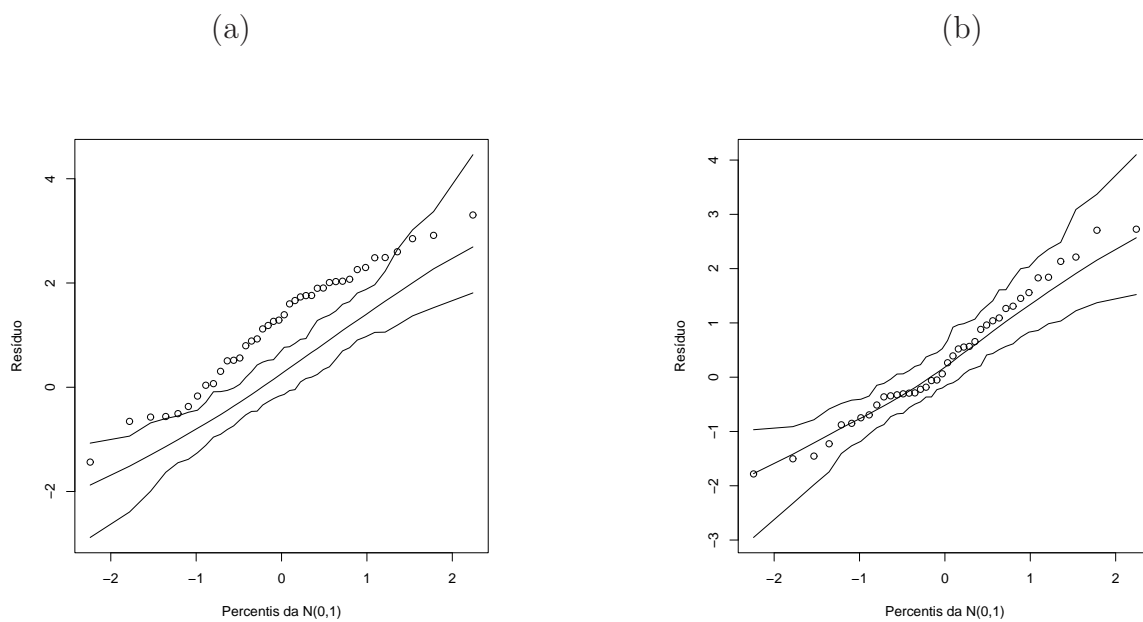


Figura 18 – Gráfico normal de probabilidade para o resíduo tipo martingale com envelope nos dados de sobrevivência de câncer de pulmão em pacientes que receberam terapia prévia para os modelos de regressão (a) log-Weibull (b) log-Weibull estendida

que receberam terapia prévia.

Em resumo, recomenda-se o uso do modelo de regressão log-Weibull estendida baseado na análise descrita anteriormente. Assim, os resultados dados na Tabela 8 sugerem que x_1 , x_2 e x_3 são significativas, e pode-se interpretar os coeficientes estimados da seguinte maneira: o tempo de sobrevivência estimado deve aumentar de aproximadamente $4\%[e^{0,0432} \times 100\%]$ quando a variável idade centrada aumenta de uma unidade, sendo as outras variáveis fixadas, da mesma forma pode-se fazer interpretações para as variáveis x_1 e x_3 . Os tratamentos, padrão e teste, não apresentam diferença significativa em relação ao efeito no tempo de sobrevivência. Quanto ao tipo de tumor, apenas o tumor do tipo adeno foi significativo comparado com tumor do tipo grande. Como a estimativa do parâmetro β_6 é negativa, o paciente cujo tipo de tumor é adeno apresenta menor probabilidade de sobrevivência que o paciente com tumor do tipo grande.

6.2 Análise dos dados de câncer de pulmão em pacientes que não receberam terapia prévia

Novamente utiliza-se o conjunto de dados de pacientes com câncer de pulmão avançado discutido por Prentice (1973), que também foi considerado na aplicação da seção 6.1, mas agora utiliza-se os pacientes que não receberam terapia prévia. Essa parte do conjunto de dados apresenta os tempos de sobrevivência, em dias, para 97 pacientes e destes 6 são observações censuradas à direita.

Bennett (1983) considerou unicamente as covariáveis mais importantes: tipo de tumor e *performance status* (PS), ambas descritas na seção 6.1, para uma aplicação do modelo de regressão log-logística para dados de sobrevivência. O objetivo foi estudar a relação entre o tempo de sobrevivência (t), em dias, a partir do início do tratamento de câncer pulmonar avançado e inoperável com um número de variáveis prognósticos. Nesse artigo não foi considerado a diferença entre os tratamentos, que mostrou ser não significativa. Kalbfleisch e Prentice (2002) na seção 3.8.2, ajustou os modelos de regressão log-Weibull, log-normal e log-F para esse mesmo sub-conjunto de pacientes e covariáveis.

A curva TTT, definida na seção 2.4, é apresentada na Figura 19 considerando

todos os 97 pacientes e para os pacientes classificados pela covariável *performance status* como alto escore (acima de 50) ou baixo escore (menor ou igual 50), respectivamente. Esses gráficos indicam uma função de taxa de falha unimodal unicamente para os pacientes classificados com escore alto para a variável *performance status*.

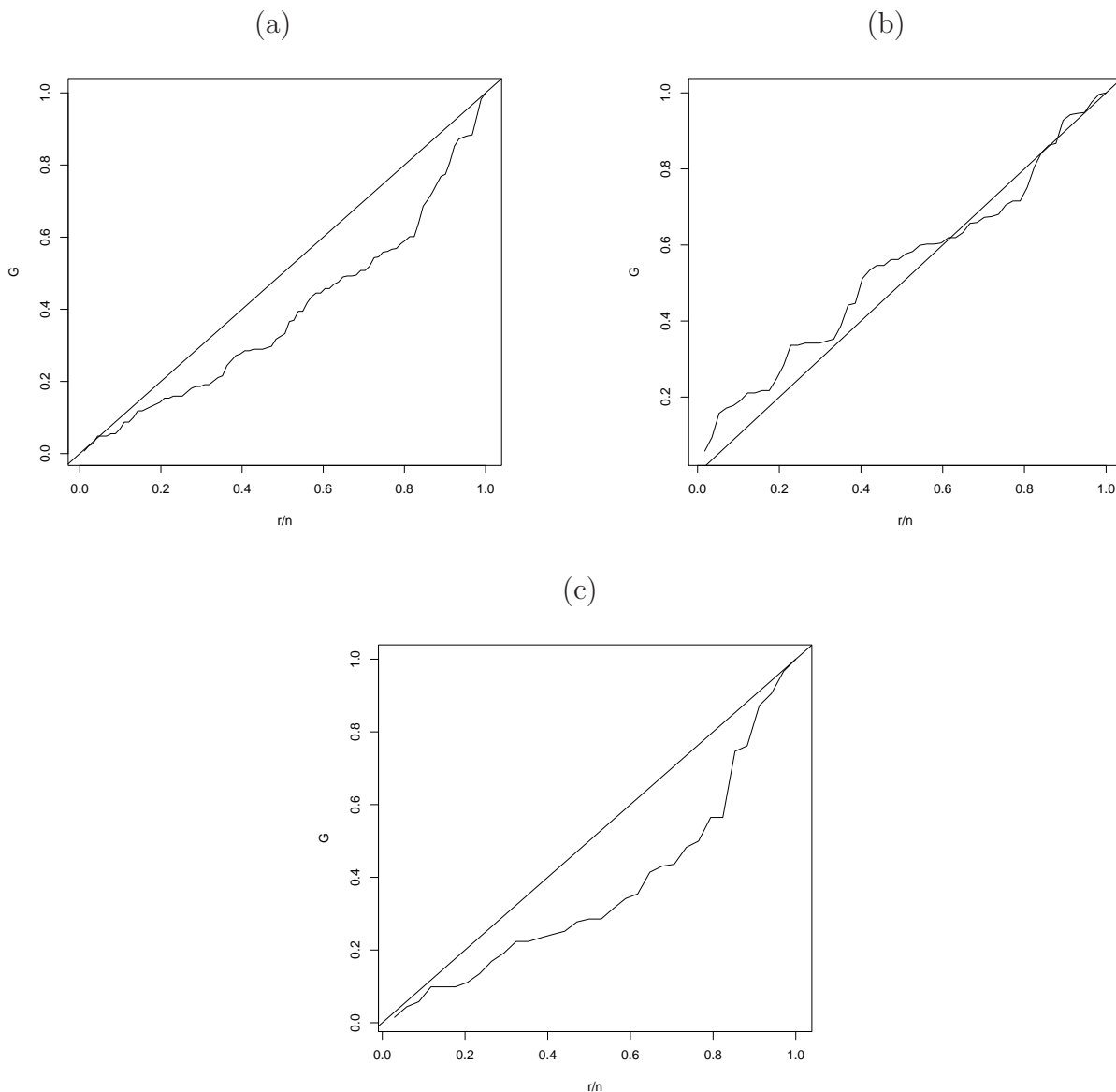


Figura 19 – Curva TTT para o conjunto dos dados de câncer de pulmão que não receberam terapia prévia: (a) todo conjunto, (b) *performance status* com alto escore e (c) *performance status* com baixo escore

Entretanto, Bennett (1983) mostrou por meio da estimativa empírica suavizada para a função de taxa de falha para os pacientes agrupados pela *performance status* que esta

função é unimodal. Assim, tanto o modelo de regressão log-Burr XII como o modelo de regressão log-logística são possíveis candidatas para analisar os dados desta aplicação.

Para discriminar esses modelos probabilísticos foi realizado o teste da razão de verossimilhança (seção 3.3.2) para as hipóteses $H_0 : k = 1$ ou modelo log-logístico é adequado *versus* $H_1 : k \neq 1$ ou modelo log-Burr XII é adequado. A estatística de teste resultou em $TRV = 2 \times (-130,876 + 131,202) = 0,652$ (p-valor=0,4194) e este resultado fornece indicações favoráveis ao modelo de regressão log-logística.

Um critério que pode ser usado para selecionar um modelo é a medida AIC_{SUR} definida na seção 3.3.1.1, em que $u = 1$ para o modelo logístico e $u = 2$ para o modelo Burr XII. O modelo que apresenta o menor valor para a medida AIC_{SUR} é selecionado com o mais adequado. Para o modelo de regressão log-logística, $AIC_{SUR} = 263,662$, enquanto que $AIC_{SUR} = 263,010$ para o log-Burr XII. Assim, esse critério indica que ambos modelos de regressão são apropriados para ajustar o conjunto de dados de pacientes com câncer no pulmão que não receberam terapia prévia.

Apesar disso, esse conjunto foi utilizado para dar uma aplicação dos resultados obtidos previamente para o modelo de regressão log-Burr XII pois o modelo de regressão log-Burr XII é uma alternativa para o modelo de regressão log-logística.

Nesta aplicação, o modelo de regressão log-Burr XII, apresentado na seção 3.3.2, é dado por

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \sigma z_i, \quad (37)$$

em que a variável Y_i dado \mathbf{x}_i tem distribuição log-Burr XII com densidade dada na equação (15), $i = 1, \dots, 97$. Para os parâmetros desse modelo, as estimativas de máxima verossimilhança foram obtidas usando a sub-routine MaxBFGS da linguagem matricial de programação Ox, cujos resultados são dados na Tabela 9.

Note que as variáveis *performance status* e tipo de tumor são fatores significativos. As estimativas dos parâmetros e os respectivos erros padrão são similares aos obtidos pelo modelo de regressão log-logística. As conclusões que podem ser extraídas são semelhantes às obtidas por Bennett (1983).

Considerando agora uma análise bayesiana no modelo, equação (37), foram

Tabela 9 – Estimativas de máxima verossimilhança para os parâmetros do modelo de regressão log-Burr XII para os dados de câncer em pacientes que não receberam terapia prévia

Parâmetro	Estimativa	E.P.	p-valor	
k	1,4287	0,6812	-	
σ	0,6023	0,0930	-	
β_0	3,2603	0,6611	< 0,00001	
β_1	0,1147	0,3153	0,71615	escamoso vs. grande
β_2	-0,7406	0,2775	0,0076	pequeno vs. grande
β_3	-0,7195	0,2960	0,0151	adeno vs. grande
β_4	0,0282	0,0052	< 0,0001	PS

Nota: E.P.=Erro Padrão

adotadas as seguintes densidades a priori vagas para os parâmetros desse modelo

- $\beta_j \sim N(0, 10^2), j = 0, \dots, 4;$
- $k \sim \Gamma(0, 01; 0, 01);$
- $\sigma \sim \text{gama inversa}(0, 1; 0, 1).$

As densidades a posteriori são dadas na seção 3.3.2. A partir delas foram geradas duas cadeias paralelas e independentes com tamanho 110.000 para cada parâmetro usando o algoritmo de Metropolis-Hasting, descrito na seção 3.2.2. As primeiras 10.000 iterações foram descartadas para eliminar o efeito dos valores iniciais e para evitar problemas de correlação foi considerado um espaço de tamanho 40 entre os valores gerados. Desse modo foi obtida uma amostra de tamanho 5.000. Para monitorar a convergência foi usado o fator de redução da escala, \hat{R} (GELMAN; RUBIN, 1992).

A Tabela 10 apresenta os sumários a posteriori para os parâmetros do modelo de regressão log-Burr XII, além de mostrar os valores estimados do fator de redução da escala. Nota-se, em todos os casos, que os valores de $\sqrt{\hat{R}}$, definido na equação (5), estão próximos de 1, o que indica convergência das cadeias.

A Figura 20 apresenta as densidades marginais a posteriori aproximadas pelo histograma, considerando as 5.000 observações amostrais geradas.

Tabela 10 – Sumários a posteriori dos parâmetros do modelo log-Burr XII para os dados de câncer em pacientes que não receberam terapia prévia

Parâmetro	Média	Desvio Padrão	2,5%	97,5%	\hat{R}
k	1,3034	0,4848	0,6154	2,5454	1,0149
σ	0,5942	0,0797	0,4515	0,7624	1,0059
β_0	3,1089	0,5744	1,9902	4,2492	1,0098
β_1	0,0974	0,3164	-0,5388	0,7010	1,0040
β_2	-0,7436	0,2795	-1,2894	-0,1851	1,0027
β_3	-0,7123	0,3055	-1,3285	-0,1162	1,0054
β_4	0,0289	0,0052	0,0189	0,0392	1,0015

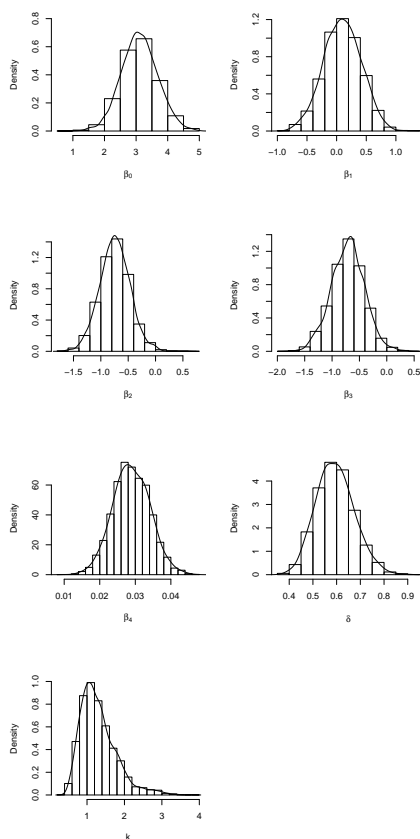


Figura 20 – Estimativa por histograma da função de densidade a posteriori dos parâmetros β , σ e k do modelo log-Burr XII para os dados de câncer em pacientes que não receberam terapia prévia

Dessa figura, observa-se que as covariáveis x_2 , x_3 e x_4 são importantes para explicar a variabilidade do tempo de sobrevivência e os valores menores do que um para o parâmetro σ indicam que a forma da função de taxa de falha deve ser unimodal.

Novamente com respeito ao objetivo do estudo, nota-se que as variáveis x_2 $(-1, 2894; -0, 1851)$, x_3 $(-1, 3285; -0, 1162)$ e x_4 $(0, 0189; 0, 0392)$ são significativas usando a abordagem bayesiana.

Para o método de jackknife, descrito na seção 3.2.3, as estimativas para os parâmetros do modelo de regressão log-Burr XII ajustado aos dados câncer para pacientes que não receberam terapia prévia estão registradas na Tabela 11.

Tabela 11 – Estimativas jackknife para os parâmetros do modelo de regressão log-Burr XII para os dados de câncer em pacientes que não recebem terapia prévia

Parâmetro	Estimativa	E.P.	Intervalo de Confiança de 95%	
k	0,8353	1,0319	(-1,1872;2,8578)	
σ	0,6115	0,1203	(0,3758;0,8473)	
β_0	3,0421	0,8659	(1,3450;4,7392)	
β_1	0,1197	0,3309	(-0,5289;0,7683)	escamoso vs. grande
β_2	-0,7375	0,2635	(-1,2539;-0,2212)	pequeno vs. grande
β_3	-0,6944	0,2644	(-1,2126;-0,1762)	adeno vs. grande
β_4	0,0285	0,0057	(0,0174; 0,0397)	PS

Nota: E.P.=Erro Padrão

Note que as variáveis *performance status* e tipo de tumor são fatores significativos. Mas as estimativas e os erros padrão dos parâmetros k , σ e β_0 não são similares aos obtidos pelo método de máxima verossimilhança, entretanto mesmo para estes parâmetros as conclusões são as mesmas considerando os dois métodos de estimação.

Com o objetivo de verificar a existência de pontos influentes, as Figuras 21(a) e (b) mostram os gráficos das medidas de influência global, $GD_i(\boldsymbol{\theta})$, equação (22), e $LD_i(\boldsymbol{\theta})$, equação (23), respectivamente. Desses gráficos nota-se que dois casos, 31 e 59, se destacam.

As Figuras 22-23 apresentam os gráficos para a medida de influência local, $|\mathbf{d}_{max}|$, e para a medida de influência local total, C_i , para $\boldsymbol{\theta} = (k, \sigma, \boldsymbol{\beta}^T)^T$ considerando os esquemas de

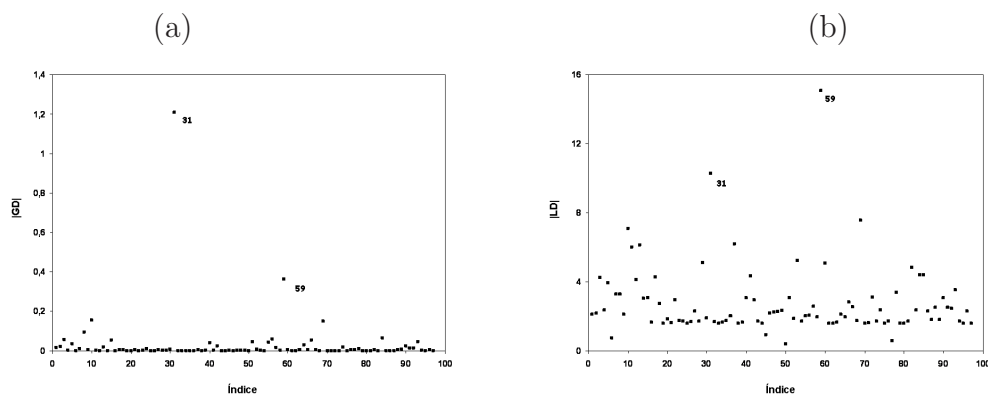


Figura 21 – Gráficos das medidas de influência global para o modelo log-Burr XII ajustado aos dados de câncer para os pacientes que não receberam terapia prévia (a) $|GD_i|$ (b) $|LD_i|$

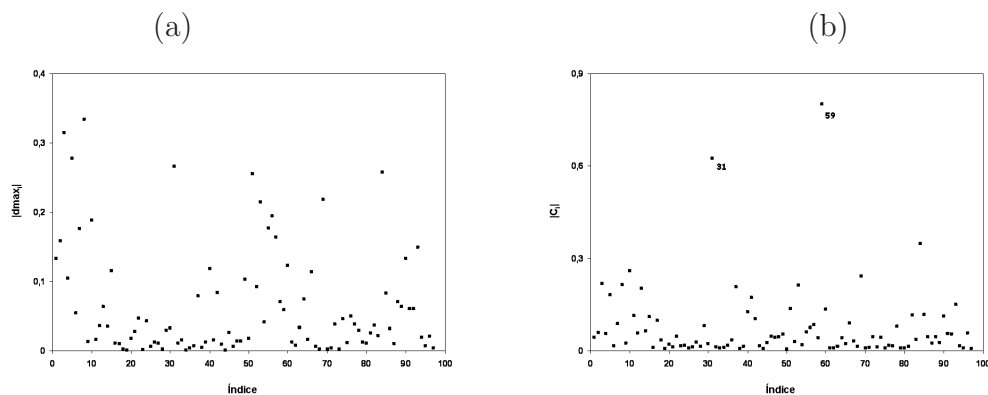


Figura 22 – Gráficos de medidas de influência para θ do modelo de regressão log-Burr XII usando perturbação de casos nos dados de câncer para os pacientes que não receberam terapia prévia (a) Influência local d_{max} (b) Influência local total C_i

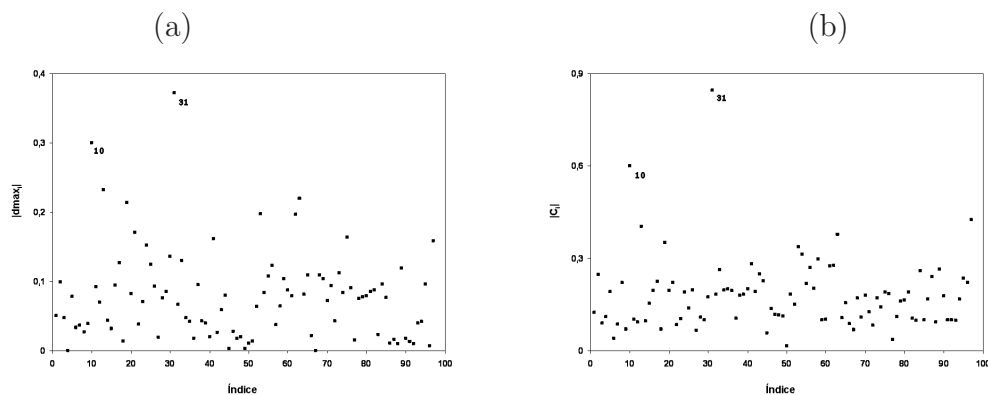


Figura 23 – Gráficos de medidas de influência para θ do modelo de regressão log-Burr XII usando perturbação de resposta nos dados de câncer para os pacientes que não receberam terapia prévia (a) Influência local d_{max} (b) Influência local total C_i

perturbação estudados na seção 4.5.2, estas medidas estão descritas na seção 4.2. Entretanto, para esse conjunto de dados, o esquema de perturbação da covariável é inviável, pois as covariáveis são categóricas. Considerando a perturbação de casos, da Figura 22(b) observa-se que os valores de C_i das observações 31 e 59 destacam-se das demais e, portanto, são possíveis candidatas a pontos influentes. Por outro lado, quando considera-se a perturbação da variável resposta, as Figuras 23(a) e (b), respectivamente, para as medidas d_{max} e C_i , apresentam como possíveis pontos influentes as observações 10 e 31.

A Figura 24 exibe o gráfico do “leverage” generalizado para o modelo de regressão log-Burr XII, descrito na seção 4.5.3, ajustado aos dados de câncer para os pacientes que não receberam terapia prévia e não mostra nenhum ponto como possível candidato a ponto de alavanca.

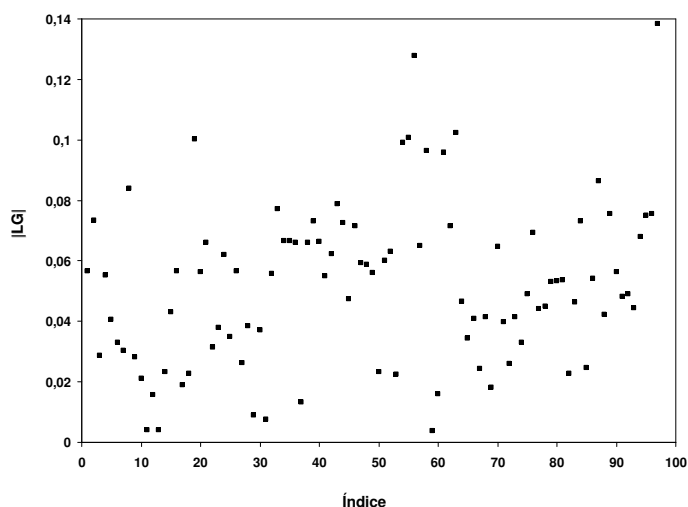


Figura 24 – Gráfico do leverage generalizado para o modelo de regressão log-Burr XII para os dados de câncer para os pacientes que não receberam terapia prévia

Para avaliar possíveis observações discrepantes bem como afastamentos das suposições feitas para o modelo de regressão log-Burr XII, apresenta-se na Figura 25(a) o gráfico normal de probabilidade para o resíduo tipo martingale modificado, dado na Seção 5.4.2, com envelope simulado. Nota-se da Figura 25(a) que a distribuição log-Burr XII com $k = 1,4287$ é adequada para ajustar os dados de câncer para os pacientes que não receberam terapia prévia. Adicionalmente, o gráfico do resíduo tipo martingale modificado contra os valores ajustados, Figura 25(b), sugere que os resíduos estão aleatoriamente dispostos so redor de zero, apesar

do caso 59 apresentar valor de resíduo maior que 3.

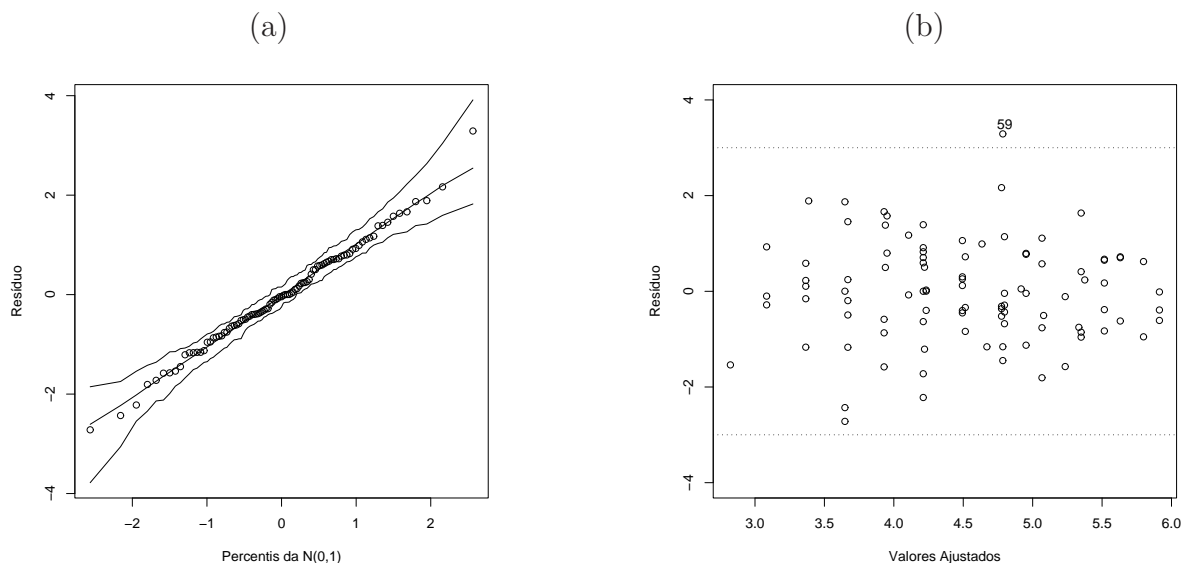


Figura 25 – (a) Gráfico normal de probabilidade para o resíduo tipo martingale modificado com envelope simulado e (b) gráfico dos valores ajustados contra o resíduo tipo martingale modificado para o modelo de regressão log-Burr XII aos dados de câncer para os pacientes que não receberam terapia prévia

De acordo com a análise de diagnóstico, os casos 10, 31 e 59 são possíveis pontos influentes. Os três casos apresentam tempos de sobrevivência não censurados. Os tempos de sobrevivência para os casos 10 e 31 correspondem aos maiores valores para os tempos de sobrevivência para os pacientes com tumor do tipo pequeno. Por outro lado, o caso 59 apresenta o menor tempo de sobrevivência (1 dia).

Agora é necessário avaliar se impacto das observações detectadas como possíveis pontos influentes é significativo. Foram realizadas várias combinações de exclusões das candidatas à observações influentes $A = \{10, 31 \text{ e } 59\}$, e novamente os parâmetros do modelo foram reestimados (Tabela 12). Na Tabela 13, encontram-se as variações percentuais de cada parâmetro estimado, definida por: $[(\theta_j - \theta_{j(I)})/\theta_j] \times 100$ em que $\theta_{j(I)}$ denota a estimativa de máxima verossimilhança de θ_j obtida depois de remover o conjunto I de observações. Nota-se a partir da Tabela 13 que o maior impacto ocorre ao excluir conjuntamente as observações 10 e 31, no entanto, as conclusões sobre a significância, ao nível de 5%, dos parâmetros não mudam. Portanto, decidiu-se que nenhuma observação deve ser removida.

Tabela 12 – Estimativas de máxima verossimilhança e p-valor entre parênteses para os parâmetros do modelo de regressão log-Burr XII para os conjuntos de dados indicados dos dados de câncer referente a pacientes que não receberam terapia prévia

Conjunto	Parâmetros						
	k	σ	β_0	β_1	β_2	β_3	β_4
A	1,4287	0,6023	3,2603	0,1147	-0,7406	-0,7195	0,0282
			< 0,0001	0,7161	0,0076	0,0151	< 0,0001
A-{caso 10}	1,6728	0,6106	3,4525	0,1294	-0,7973	-0,7272	0,0272
			< 0,0001	0,6741	0,0033	0,0121	< 0,0001
A-{caso 31}	2,2153	0,6394	3,6777	0,1556	-0,7865	-0,7358	0,0274
			< 0,0001	0,6058	0,0033	0,0101	< 0,0001
A-{caso 59}	1,0601	0,5227	2,9711	0,1689	-0,7494	-0,7005	0,0289
			< 0,0001	0,5786	0,0046	0,0135	< 0,0001
A-{casos 10 e 31}	5,5710	0,6932	4,6486	0,1872	-0,8780	-0,7849	0,0248
			0,0050	0,4989	0,0005	0,0043	< 0,0001
A-{casos 10 e 59}	1,2346	0,5343	3,1450	0,1830	-0,7971	-0,7060	0,0279
			< 0,0001	0,5405	0,0022	0,0112	< 0,0001
A-{casos 31 e 59}	1,4975	0,5565	3,2817	0,2030	-0,7895	-0,7111	0,0281
			< 0,0001	0,4924	0,0022	0,0100	< 0,0001
A-{casos 10, 31 e 59}	2,6200	0,6029	3,8837	0,2308	-0,8679	-0,7459	0,0259
			0,0002	0,4081	0,0005	0,0056	< 0,0001

Tabela 13 – Variação percentual da estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-Burr XII para o conjunto de dados indicados dos dados de câncer referente a pacientes que não receberam terapia prévia

Conjunto	Parâmetro						
	k	σ	β_0	β_1	β_2	β_3	β_4
A-{caso 10}	-17,086	-1,376	-5,895	-12,844	-7,658	-1,079	3,436
A-{caso 31}	25,800	13,221	8,870	-47,306	-1,187	2,639	-2,316
A-{caso 59}	-55,057	-6,161	-12,802	-35,691	-6,205	-2,268	2,674
A-{casos 10 e 31}	-289,935	-15,097	-42,582	-63,237	-18,557	-9,094	12,081
A-{casos 10 e 59}	13,586	11,285	3,537	-59,592	-7,633	1,872	0,904
A-{casos 31 e 59}	-4,816	7,602	-0,656	-77,014	-6,610	1,165	0,231
A-{casos 10, 31 e 59}	-83,384	-0,091	-19,121	-101,238	-17,193	-3,669	8,106

Então, o modelo final para os dados de câncer com pacientes que não receberam terapia prévia é composto pelas covariáveis PS e tipo de tumor, isto é, nenhuma variável explicativa foi removida. De acordo com o modelo final, a interpretação, de acordo com a seção 3.2.1, é que o tempo mediano de sobrevivência fica aumentado em 2,86% ($[\exp\{0,0282\}] \times 100\%$) quando a variável PS aumenta em uma unidade, sendo as outras variáveis fixadas. Como as estimativas dos parâmetros β_2 e β_3 assumem valores negativos, então os pacientes com tumor do tipo pequeno ou adeno apresentam probabilidade de sobrevivência menor comparado com pacientes que têm tumor grande ou escamoso.

A próxima aplicação apresenta curva TTT que indica que a função de taxa de falha apresenta forma de U . Desta forma, serve para ilustrar o modelo de regressão log-Weibull estendida proposto na seção (3.3.1) bem como foram realizadas as análises de sensibilidade e de resíduos para este modelo.

6.3 Análise dos dados de peixes da espécie *Golden shiner*

Em um estudo conduzido no Lago Saint Pierre, Quebec, foram observados os tempos de sobrevivência para a espécie de peixe *Golden shiner* (LAPLANTE-ALBERT, 2008). O estudo incluiu 106 peixes, sendo que cada peixe foi unido por meio de uma corda a um dispositivo permitindo ao peixe nadar em *midwater* e um cronômetro foi acoplado a este dispositivo de maneira que a contagem do tempo é iniciada quando o peixe é capturado por um predador. Das $n = 106$ observações dos tempos de sobrevivência 15 foram censuradas.

O pesquisador observou dois tempos, o primeiro é o tempo entre o início do experimento e a recuperação desse dispositivo e o segundo é o tempo marcado no cronômetro. O tempo de sobrevivência é obtido a partir da diferença entre esses tempos e caso o peixe não seja capturado por um predador em um período de aproximadamente 24 horas o dispositivo é recuperado e a observação é considerada censurada. Assim, a resposta de interesse é o tempo que o peixe levou para ser capturado por um predador.

O objetivo do estudo é identificar que variáveis podem influenciar o tempo de captura do peixe da espécie *Golden shiner*. As variáveis consideradas nesse experimento foram:

- t_i : tempo de sobrevivência observado (em horas);
- δ_i : indicador de censura (0=observação censurada, 1= tempo de sobrevivência observado);
- x_{i1} : margem sul ou norte do lago (0=norte, 1=sul);
- x_{i2} : distância sobre o eixo longitudinal do lago (em km);
- x_{i3} : tamanho do peixe (em cm);
- x_{i4} : profundidade do lago (em cm);
- x_{i5} : abundância de plantas *macro – thin* (em porcentagem);
- x_{i6} : transparência da água (em cm);

- x_{i7} : tempo de início.

Para identificar um modelo apropriado para o tempo de sobrevivência do conjunto de dados *Golden shiner* foi utilizada a curva TTT, definida na seção 2.4, que é apresentada na Figura 26 a qual indica uma função de taxa de falha na forma de U . Então, para esse conjunto de dados pode-se utilizar o modelo de regressão log-Weibull estendida.

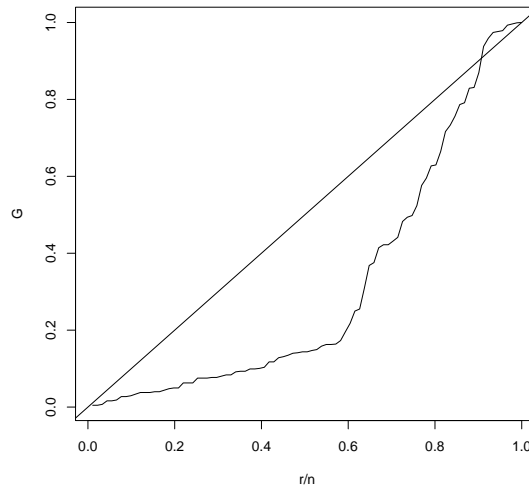


Figura 26 – Curva TTT para dados dos peixes *Golden shiner*

Nesta aplicação, o modelo log-Weibull estendida, apresentado na seção 3.3.1, é dado por

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \sigma z_i, \quad (38)$$

em que a variável Y_i dado x_i tem distribuição log-Weibull estendida com densidade dada na equação (9), $i = 1, \dots, 106$. Para os parâmetros desse modelo, as estimativas de máxima verossimilhança foram obtidas usando a sub-routine MaxBFGS da linguagem matricial de programação Ox (programa no apêndice F), cujos resultados são dados na tabela 14.

Com respeito ao objetivo do estudo, nota-se que as variáveis x_1 ($\hat{\beta}_1 = -3,3214$ e p-valor $< 0,01$), x_2 ($\hat{\beta}_2 = -0,2371$ e p-valor $< 0,01$), x_3 ($\hat{\beta}_3 = 0,2309$ e p-valor $= 0,013$), x_4 ($\hat{\beta}_4 = -0,0915$ e p-valor $< 0,001$) e x_5 ($\hat{\beta}_5 = -0,1120$ e p-valor $< 0,01$) são significativas usando o teste Z .

Tabela 14 – Estimativas de máxima verossimilhança dos parâmetros do modelo log-Weibull estendida para os dados da espécie de peixe *Golden Shiner*

Parâmetro	Estimativa	E.P.	p-valor
λ	0,0189	0,0060	-
σ	1,5969	0,1383	-
β_0	7,3999	5,2890	0,1617
β_1	-3,3214	0,9752	< 0,01
β_2	-0,2371	0,0634	< 0,01
β_3	0,2309	0,0604	0,013
β_4	-0,0915	0,0201	< 0,001
β_5	-0,1120	0,0449	< 0,01
β_6	-0,1840	0,3809	0,629
β_7	-0,0944	0,2668	0,724

Nota: E.P.=Erro Padrão

Considerando agora uma análise bayesiana no modelo, equação (38), foram adotadas as seguintes densidades a priori vagas para os parâmetros desse modelo

- $\beta_j \sim N(0, 10^2), j = 0, \dots, 7;$
- $\lambda \sim \Gamma(1, 1);$
- $\sigma \sim \Gamma(0, 1, 0, 1).$

As densidades a posteriori são dadas na seção 3.3.1. A partir delas foram geradas duas cadeias paralelas e independentes com tamanho 110.000 para cada parâmetro usando o algoritmo de Metropolis-Hasting, descrito na seção 3.3.1. As primeiras 10.000 iterações foram descartadas para eliminar o efeito dos valores iniciais e para evitar problemas de correlação foi considerado um espaço de tamanho 40 entre os valores gerados. Desse modo foi obtida uma amostra de tamanho 5.000. Para monitorar a convergência foi usado o fator de redução da escala, \hat{R} (GELMAN; RUBIN, 1992). A Tabela 15 apresenta os sumários a posteriori para os parâmetros do modelo de regressão log-Weibull estendida, além de mostrar os valores estimados do fator de redução da escala.

Tabela 15 – Sumários a posteriori dos parâmetros do modelo log-Weibull estendida para os dados da espécie de peixe *Golden Shiner*

Parâmetro	Média	Mediana	Desvio Padrão	2,5%	97,5%	\hat{R}
λ	0,0149	0,0150	0,0073	0,0016	0,0293	0,999
σ	1,5490	1,2980	0,1421	1,2980	1,8550	0,997
β_0	7,1270	7,1720	5,4130	1,5320	12,6800	1,006
β_1	-3,4350	-3,3490	1,0580	-5,8270	-1,6100	1,001
β_2	-0,2316	-0,2241	0,0847	-0,4321	-0,0858	1,001
β_3	0,2248	0,2137	0,0764	0,1020	0,4155	1,005
β_4	-0,0843	-0,0819	0,0264	-0,1441	-0,0388	1,003
β_5	-0,0978	-0,0949	0,0377	-0,1832	-0,0352	1,004
β_6	-0,3770	-0,3539	0,4117	-1,2770	0,3845	1,001
β_7	0,0278	0,0231	0,2129	-0,3821	0,4660	1,001

Nota-se, em todos os casos, que os valores de $\sqrt{\hat{R}}$, definido na equação (5), estão próximos de 1, o que indica convergência das cadeias. Além disso, a Figura 27 apresenta as densidades marginais a posteriori aproximadas pelo histograma, considerando as 5.000 observações amostrais geradas. Dessa figura, observa-se que as covariáveis, com exceção de x_6 e x_7 , são importantes para explicar a variabilidade do tempo de sobrevivência e os valores maiores do que um para o parâmetro σ indicam que a função de taxa de falha deve ser em forma de U .

Novamente com respeito ao objetivo do estudo, nota-se que as variáveis x_1 $(-5,8270; -1,6100)$, x_2 $(-0,4321; -0,0858)$, x_3 $(0,1020; 0,4155)$, x_4 $(-0,1441; -0,0388)$ e x_5 $(-0,1832; -0,0352)$ são significativas usando a abordagem bayesiana.

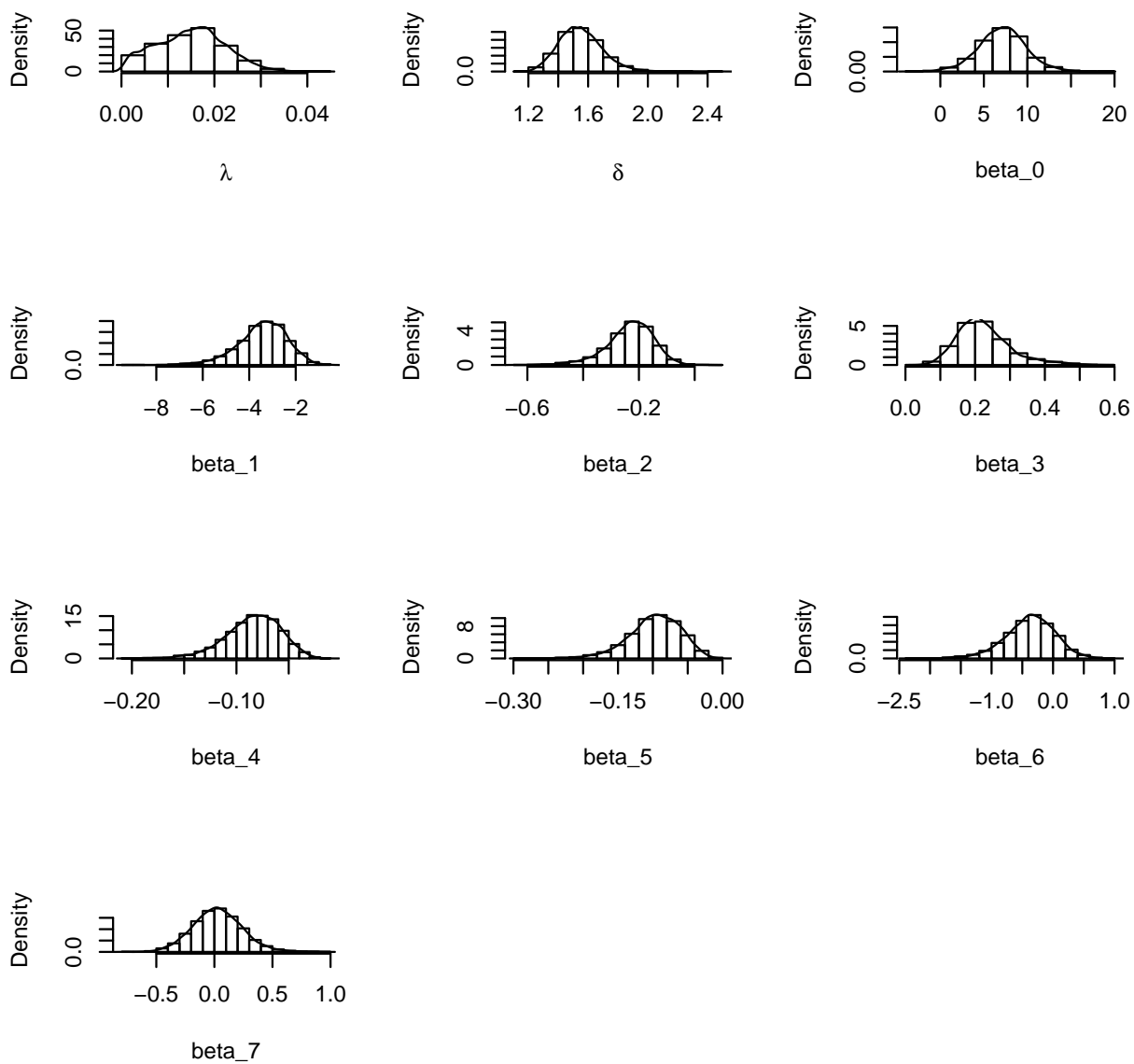


Figura 27 – Estimativa por histograma da função de densidade a posteriori dos parâmetros β , σ e λ do modelo log-Weibull estendida para os dados de peixe da espécie *Golden shiner*

Usando o método de jackknife, descrito na seção 3.2.3, para estimar os parâmetros do modelo, equação (38), ocorreu que, quando a observação 5, que é uma observação censurada e corresponde ao maior tempo de sobrevivência, foi retirada, não se obteve convergência no algoritmo quase-Newton e o mesmo ocorreu na abordagem bayesiana. Assim, foi realizada a aproximação de primeira ordem (COOK, WEISBERG, 1982)

$$\hat{\theta}_i^1 = \hat{\theta} + \{\ddot{\mathbf{L}}(\hat{\theta})\}^{-1} \dot{\mathbf{L}}(\hat{\theta}), \quad (39)$$

em que $\dot{\mathbf{L}}_i(\hat{\theta}) = \left. \frac{\partial \mathbf{L}_i(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}}$. Os resultados estão na Tabela 16 e o programa feito na linguagem matricial de programação Ox encontra-se no apêndice G.

Tabela 16 – Estimativas de jackknife para os parâmetros do modelo log-Weibull estendida para os dados da espécie de peixe *Golden Shiner*.

Parâmetro	Estimativa	E.P.	95% Intervalo de Confiança
λ	0,0190	0,0075	(0,0018; 0,0361)
σ	1,5969	0,1379	(1,2833; 1,9105)
β_0	7,4001	6,8398	(-8,1515; 22,9517)
β_1	-3,3222	1,1250	(-5,8802; -0,7642)
β_2	-0,2370	0,0586	(-0,3704; -0,1037)
β_3	0,2309	0,0633	(0,0868; 0,3749)
β_4	-0,0915	0,0205	(-0,1380; -0,0450)
β_5	-0,1120	0,1018	(-0,3435; 0,1195)
β_6	-0,1838	0,4743	(-1,2623; 0,8947)
β_7	-0,0944	0,3326	(-0,8507; 0,6618)

Nota: E.P.=Erro Padrão

Novamente com respeito ao objetivo do estudo, nota-se da Tabela 16 que as variáveis x_1 (-5,8802; -0,7642), x_2 (-0,3704; -0,1037), x_3 (0,0868; 0,3749), x_4 (-0,1380; -0,0450) são significativas usando o método de jackknife.

Os três métodos de estimação apresentam estimativas similares para os parâmetros, entretanto o método de jackknife não apresenta as mesmas covariáveis como sendo significativas, o que pode ser justificado pela aproximação utilizada.

Para detectar possíveis casos influentes no modelo de regressão log-Weibull estendida para os dados de peixe da espécie *Golden shiner*, a Figura 28 mostra os gráficos das medidas de influência global, $GD_i(\boldsymbol{\theta})$, equação (22), e $LD_i(\boldsymbol{\theta})$, equação (23). Como no método jackknife ocorreu problema com a observação 5 e, portanto, essas medidas foram obtidas usando a aproximação dada na equação (39) por meio da linguagem matricial de programação Ox, veja programa no apêndice G. Nota-se uma certa discrepância dos valores para os casos 5, 40 e 101 que os indicam como possíveis observações influentes.

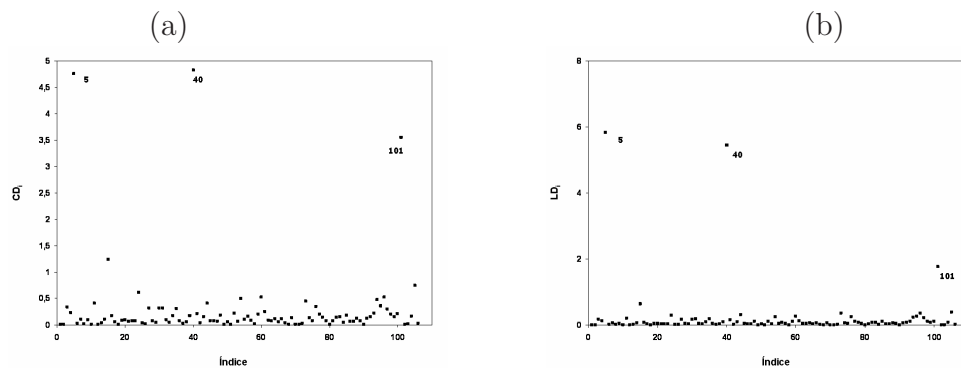


Figura 28 – Gráficos das medidas de influência global para o modelo log-Weibull estendida ajustado aos dados de sobrevivência da espécie de peixe *Golden shiner*. (a) GD_i (b) LD_i

As Figuras 29-36 apresentam os gráficos para a medida de influência local, $|\mathbf{d}_{max}|$, e para a medida de influência local total, C_i , para $\boldsymbol{\theta} = (\lambda, \sigma, \boldsymbol{\beta}^T)^T$ considerando os três esquemas de perturbação estudados na seção 4.5.1; ambas medidas estão descritas na seção 4.2 (programa no apêndice H). Nessas figuras, as observações 5, 17, 40 e 101 destacam-se das demais e, portanto, são possíveis candidatas a pontos influentes.

Portanto, os casos 5, 17, 40 e 101 são possíveis pontos influentes. Os casos 5 e 101 apresentam tempos de sobrevivência censurados. O tempo de sobrevivência para o caso 5 corresponde ao valor de 44,4 horas que é o maior da amostra e o caso 17 apresenta um tempo alto (19,2 horas). Por outro lado, o caso 101 apresenta o menor tempo de sobrevivência (18,7 horas) para os casos censurados. Por fim, o caso 40 apresenta o maior valor para a variável abundância de plantas *macro – thin* (64,16%) , x_5 .

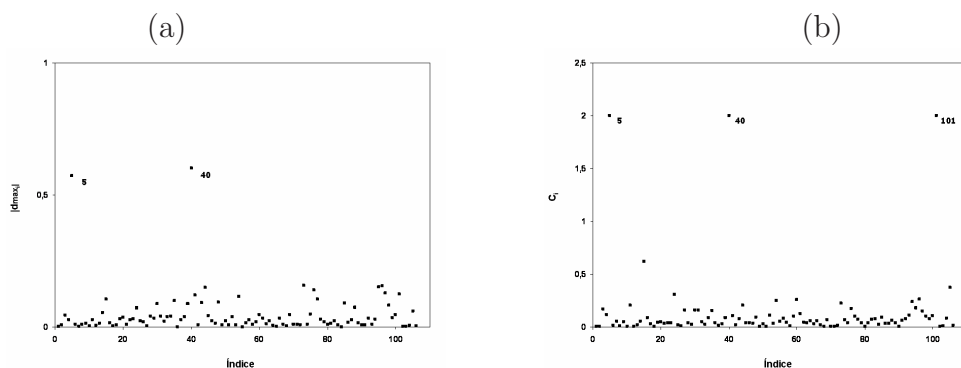


Figura 29 – Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação de casos dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total d_{max} (b) Influência local total C_i

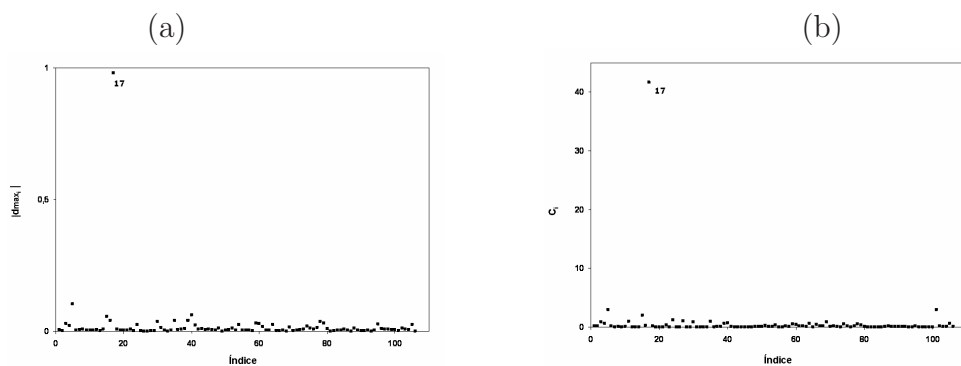


Figura 30 – Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação das resposta dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total d_{max} (b) Influência local total C_i

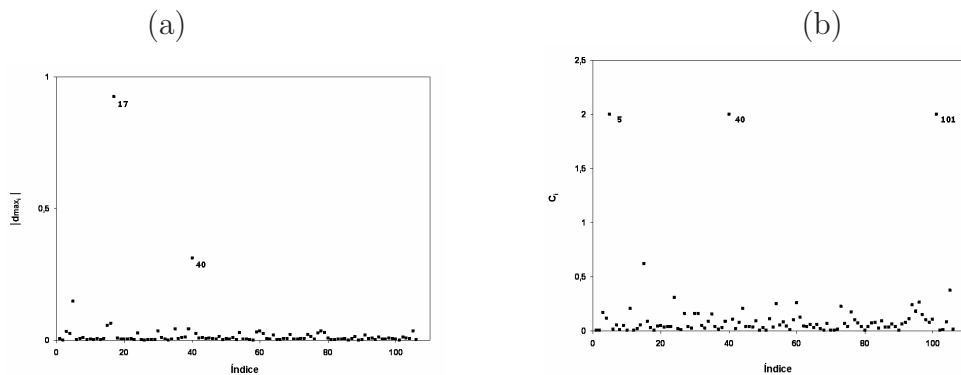


Figura 31 – Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_2 (distância) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total d_{max} (b) Influência local total C_i

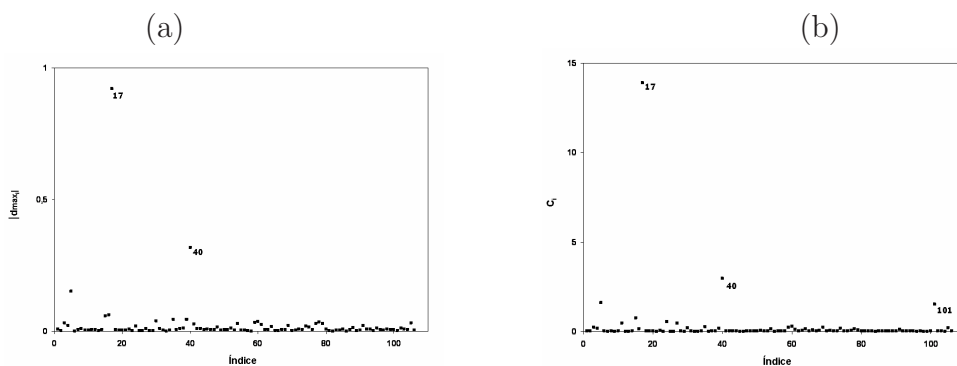


Figura 32 – Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_3 (tamanho) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total d_{max} (b) Influência local total C_i

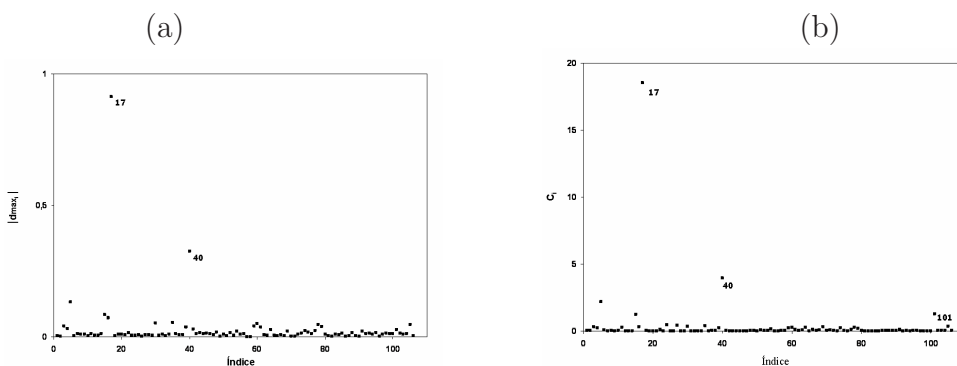


Figura 33 – Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_4 (profundidade) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total d_{max} (b) Influência local total C_i

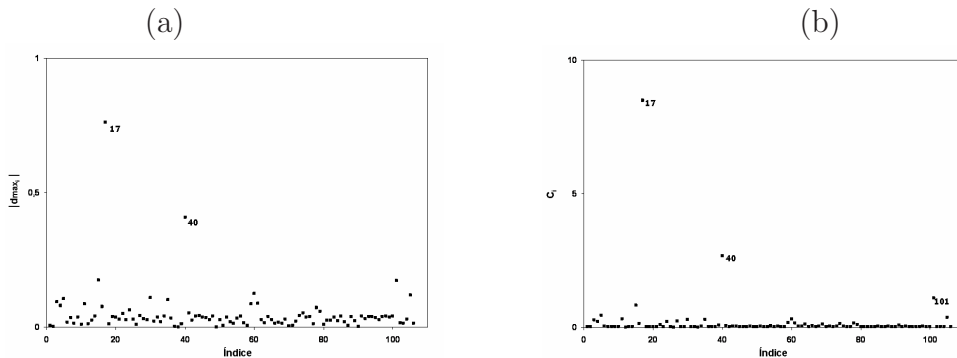


Figura 34 – Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_5 (abundância) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total d_{max} (b) Influência local total C_i

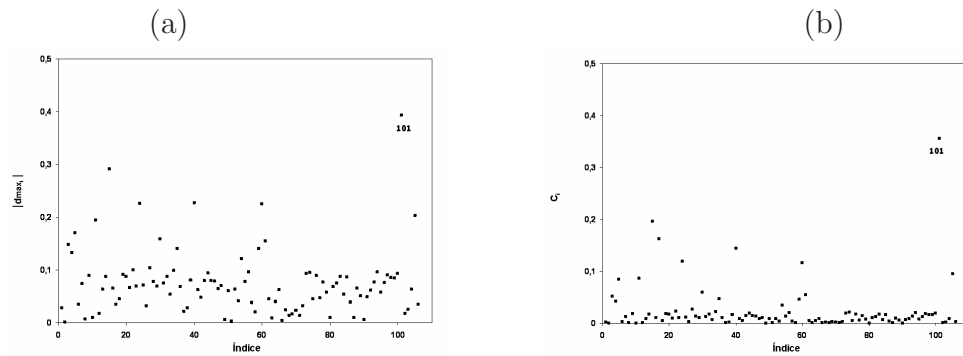


Figura 35 – Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_6 (transparência) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total d_{max} (b) Influência local total C_i

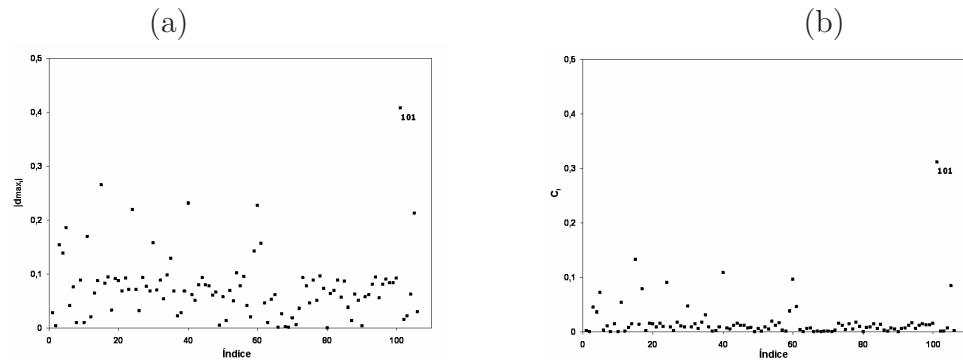


Figura 36 – Gráficos de medidas de influência para θ do modelo log-Weibull estendida usando perturbação na variável explicativa x_7 (tempo inicial) dos dados de sobrevivência dos peixes da espécie *Golden shiner* (a) Influência total d_{max} (b) Influência local total C_i

Agora é necessário avaliar o impacto das observações detectadas como possíveis pontos influentes. Uma possibilidade é estimar novamente os parâmetros do modelo retirando o conjunto de observações influentes $\{5, 17, 40 \text{ e } 101\}$, ou qualquer subconjunto deste, e analisar se as estimativas dos parâmetros mudam de maneira significativa. Entretanto, quando retira-se qualquer subconjunto que contenha a observação 5, o algoritmo quase-Newton não converge. Assim, como não existe nenhum ponto que apareça em todos os gráficos decidiu-se continuar com a amostra total.

Além disso, a Figura 37 apresenta graficamente o resíduo tipo martingale, dado na seção 5.3, para os dados de sobrevivência da espécie de peixe *Golden shiner* para o modelo log-Weibull estendida em que se pode notar que todas as observações estão contidas no intervalo $(-3, 3)$ e os valores dos resíduos apresentam comportamento aleatório, então,

não existem evidências de que as suposições do modelo ajustado sejam inadequadas e não se observa a presença de pontos atípicos.

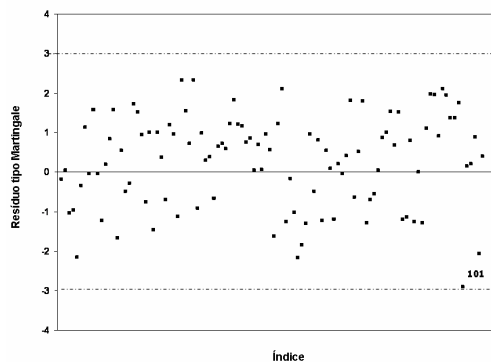


Figura 37 – Gráfico do resíduo tipo martingale para o modelo log-Weibull estendida nos dados da espécie de peixe *Golden shiner*.

Como uma medida de qualidade de ajuste, a Figura 38 mostra o gráfico normal de probabilidades para o resíduo tipo martingale com envelope simulado (o programa *R* é apresentado no apêndice I). Nota-se, pelo gráfico, que não há indícios de afastamentos sérios da suposição de distribuição Weibull estendida para os tempos de sobrevivência dos peixes da espécie *Golden shiner* e não existe nenhum possível valor atípico.

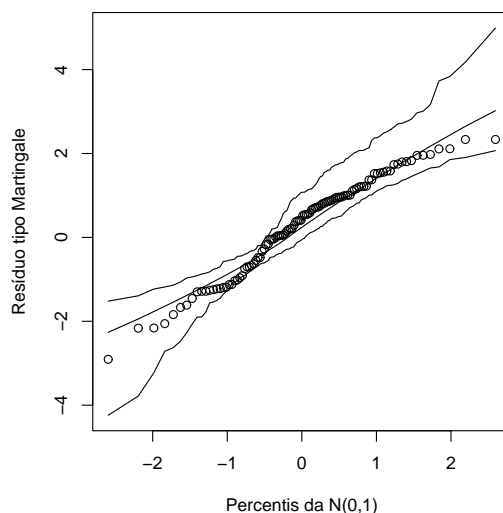


Figura 38 – Gráfico normal de probabilidade para o resíduo tipo martingale com envelope sob o modelo log-Weibull estendida para os dados de sobrevivência da espécie de peixe *Golden shiner*

Portanto, o modelo final é dado por

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \sigma z_i, \quad i = 1, \dots, 106. \quad (40)$$

As estimativas de máxima verossimilhança para os parâmetros do modelo final, equação (40), são dadas na Tabela 17 e nota-se que as variáveis x_1, x_2, x_3, x_4 e x_5 são significativas.

Tabela 17 – Estimativas de máxima verossimilhança para os dados da espécie de peixe *Golden Shiner* para o modelo de regressão log-Weibull estendida - modelo final

Parâmetro	Estimativa	E.P.	p-valor
λ	0,0193	0,0057	-
σ	1,6107	0,1337	-
β_0	6,7917	2,9448	0,0211
β_1	-3,3201	0,9067	< 0,001
β_2	-0,2275	0,0614	< 0,001
β_3	0,2158	0,0520	< 0,0001
β_4	-0,0916	0,0199	< 0,0001
β_5	-0,1041	0,0377	0,00571

Nota: E.P.=Erro Padrão

De acordo com o modelo final, a interpretação, de acordo com a seção 3.2.1, é que o tempo mediano de sobrevivência fica reduzido em 80% ($[\exp\{-0.2275\}] \times 100\%$) quando a variável x_2 aumenta em uma unidade, sendo as outras variáveis fixadas. Como a estimativa do parâmetro β_1 é negativa, então os peixes da margem sul apresentam probabilidade de sobrevivência menor do que a dos peixes da margem norte.

A próxima aplicação apresenta curva TTT que indica que a função de taxa de falha apresenta forma unimodal. Desta forma, serve para ilustrar o modelo de regressão log-Burr XII proposto na seção (3.3.2) bem como foram realizadas as análises de sensibilidade e de resíduos para este modelo.

6.4 Análise dos dados de suplementação de vitamina A

O modelo de regressão log-Burr XII proposto na seção (3.3.2) é ilustrado nesta seção usando dados de um experimento que foi realizado para avaliar o efeito da suplementação da vitamina A em episódios de diarreia. A amostra é composta de 1207 crianças, com idade entre 6 e 48 meses no início do experimento, que foram alocadas para receber placebo ou vitamina A. Este experimento foi realizado em uma pequena cidade no nordeste do Brasil entre dezembro de 1990 e dezembro de 1991. A dosagem de vitamina A foi 100.000 IU para crianças com menos de 12 meses e 200.000 IU para crianças com mais de 12 meses que são as dosagens estabelecidas pela Organização Mundial de Saúde para a prevenção de deficiência em vitamina A. Para mais detalhes sobre os dados, veja Barreto et al. (1994).

O tempo de sobrevivência foi definido como o tempo da primeira dose de vitamina A até a ocorrência de um episódio de diarreia, que foi definido como uma seqüência de dias com diarreia e um dia com diarreia ocorre quando três ou mais líquidos ou semi-líquidos foram registrados em um período de 24 horas. A informação de ocorrência de diarreia foi coletada em cada visita que ocorre no período de 48 a 72 horas. O número de líquidos e semi-líquidos ocorrido, num período de 24 horas, foi registrado. As variáveis consideradas no modelo foram:

- t_i : tempo de sobrevivência (em horas);
- δ_i : indicador de censura (0=observação censurada, 1= tempo de sobrevivência observado);
- x_{i1} : idade no início do experimento (em meses);
- x_{i2} : tratamento (0 = placebo, 1 = vitamina A);
- x_{i3} : sexo (0 = feminino, 1 = masculino).

O objetivo do estudo é identificar que variáveis influenciam o tempo até ocorrência de um episódio de diarreia.

Para identificar um modelo apropriado para o tempo de sobrevivência do conjunto de dados de suplementação de vitamina A, foi utilizada a curva TTT, definida na seção

2.4, que é apresentada na Figura 39 e indica uma função de taxa de falha unimodal. Então, para esse conjunto de dados pode-se utilizar o modelo de regressão log-Burr XII.

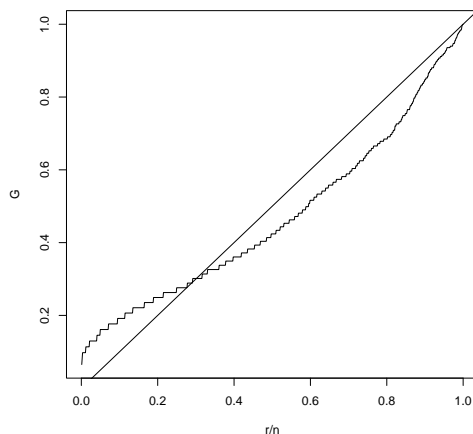


Figura 39 – Curva TTT para os dados de suplementação de vitamina A

Nesta aplicação, o modelo de regressão log-Burr XII, apresentado na seção 3.3.2, é dado por

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \sigma z_i, \quad (41)$$

em que a variável Y_i dado \mathbf{x}_i tem distribuição log-Burr XII com densidade dada na equação (15), $i = 1, \dots, 1207$. Para os parâmetros desse modelo, as estimativas de máxima verossimilhança foram obtidas usando a sub-routine MaxBFGS da linguagem matricial de programação Ox, cujos resultados são dados na Tabela 18.

Com respeito ao objetivo do estudo, observa-se que a variável x_1 ($\hat{\beta}_1=0,0221$ e p-valor $< 0,01$) é significativa para o modelo. Assim, unicamente a variável idade influencia o tempo de interesse.

Apenas para ilustrar, a Figura 40 mostra as estimativas, por meio do estimador de Kaplan-Meier, para a função de sobrevivência por tratamento e por sexo. Pode-se observar que não existe evidência de diferença das funções de sobrevivência empírica entre os tratamentos, vitamina A e placebo, bem como entre os sexos. Então, tais variáveis não parecem influenciar o tempo de sobrevivência, o que é compatível com os resultados do método de máxima verossimilhança encontrados na Tabela 18.

Tabela 18 – Estimativas de máxima verossimilhança dos parâmetros do modelo log-Burr XII para os dados de Suplementação de Vitamina A

Parâmetro	Estimativa	E.P.	p-valor
k	0,2764	0,0387	-
σ	0,3567	0,0305	-
β_0	2,2522	0,0927	0
β_1	0,0221	0,0029	< 0,01
β_2	0,0898	0,05999	0,1346
β_3	0,0441	0,0598	0,4601

Nota: E.P.=Erro Padrão

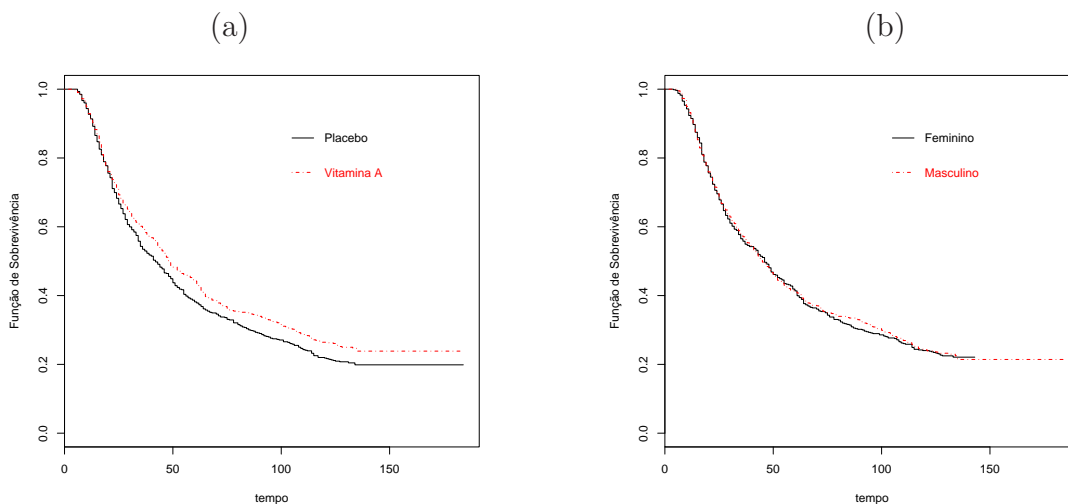


Figura 40 – Estimativas de Kaplan-Meier das funções de sobrevivência para os dados de suplementação de vitamina A (a) por tratamento (b) por sexo

Considerando agora uma análise bayesiana do modelo, equação (41), foram adotadas as seguintes densidades a priori vagas para os parâmetros do modelo:

- $\beta_j \sim N(0, 1000)$, $j = 0, 1, 2, 3$;
- $\sigma \sim \text{gama inversa}(0, 01; 0, 01)$;
- $k \sim \Gamma(0, 01; 0, 01)$.

As densidades a posteriori são dadas na seção 3.3.2. A partir delas foram geradas

duas cadeias paralelas e independentes com tamanho 35.000 para cada parâmetro usando o algoritmo de Metropolis-Hasting, descrito na seção 3.2.2. As primeiras 5.000 iterações foram descartadas para eliminar o efeito dos valores iniciais e para evitar problemas de correlação foi adotado um espaçamento de tamanho 10 entre os valores gerados. Desse modo, foi obtida uma amostra de tamanho 6.000. Para monitorar a convergência, foi usado o fator de redução da escala, \hat{R} , definido na equação (5). A Tabela 19 apresenta os sumários a posteriori para os parâmetros do modelo de regressão log-Burr XII, além de mostrar os valores estimados do fator de redução da escala.

Tabela 19 – Sumários a posteriori dos parâmetros do modelo de regressão log-Burr XII para os dados de suplementação de vitamina A

Parâmetro	Média	Mediana	Desvio Padrão	2,5%	97.5%	\hat{R}
k	0,2853	0,2817	0,04085	0,2171	0,3763	1,001
σ	0,3628	0,3617	0,0308	0,3064	0,4271	1,009
β_0	2,2551	2,2534	0,0948	2,0693	2,4451	1,000
β_1	0,0224	0,0223	0,0028	0,0169	0,0281	1,002
β_2	0,0905	0,0904	0,0602	-0,0267	0,2098	1,006
β_3	0,0461	0,0458	0,0605	-0,0743	0,16428	1,004

Nota-se, em todos os casos, que os valores de $\sqrt{\hat{R}}$ estão próximos de 1, o que indica convergência das cadeias. Além disso, a Figura 41 apresenta as densidades marginais a posteriori aproximadas pelo histograma, considerando as 6.000 observações amostrais geradas. Dessa figura, observa-se que apenas a covariável x_1 é importante para explicar a variabilidade do tempo de sobrevivência e os valores menores que um para o parâmetro σ indicam que a função de taxa de falha deve ser unimodal.

Retornando ao objetivo do estudo, nota-se que a variável x_1 (0,0169; 0,0281) é significativa para o modelo quando é usada a abordagem bayesiana.

Para o método de jackknife, as estimativas para os parâmetros do modelo de regressão log-Burr XII ajustado aos dados de suplementação de vitamina A estão registradas na Tabela 20. Dessa Tabela, observa-se que a variável x_1 (0,0186; 0,0324) é significativa para o modelo quando é usado o estimador de jackknife. Portanto, para os três métodos de estimação abordados as estimativas são similares.

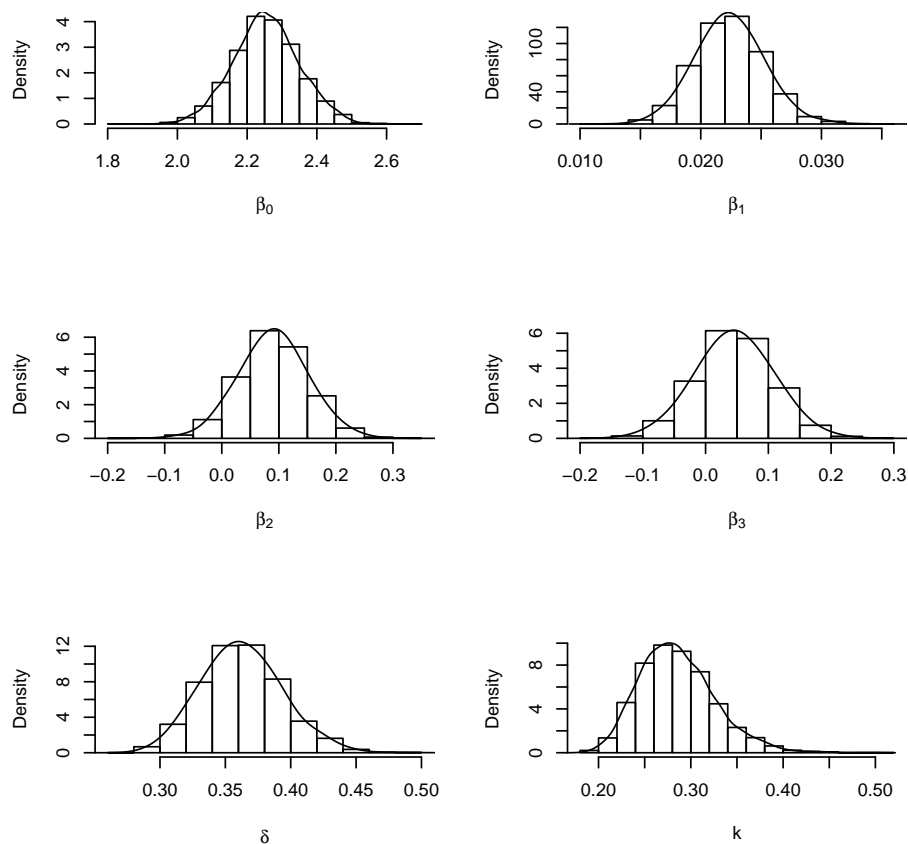


Figura 41 – Estimativa por histograma da função de densidade a posteriori dos parâmetros β_0 , β_1 , β_2 , β_3 , σ e k do modelo log-Burr XII para os dados de suplementação de vitamina A

Tabela 20 – Estimativas jackknife dos parâmetros do modelo log-Burr XII para os dados de Suplementação de Vitamina A

Parâmetros	Estimativas	E.P.	95% Intervalo de Confiança
k	0,2664	0,0477	(0,1728;0,3600)
σ	0,3599	0,0364	(0,2885;0,4313)
β_0	2,2464	0,0879	(2,0739;2,4189)
β_1	0,0255	0,0035	(0,0186;0,0324)
β_2	0,0921	0,0622	(-0,0299; 0,2141)
β_3	0,0482	0,0616	(-0,0727; 0,1691)

Nota: E.P.=Erro Padrão

Para detectar possíveis casos influentes no modelo de regressão log-Burr XII para os dados de suplementação de vitamina A. A Figura 42 mostra os gráficos das medidas de influência global, $GD_i(\boldsymbol{\theta})$, equação (22), e $LD_i(\boldsymbol{\theta})$, equação (23). Essa medidas foram obtidas usando a linguagem matricial de programação Ox.

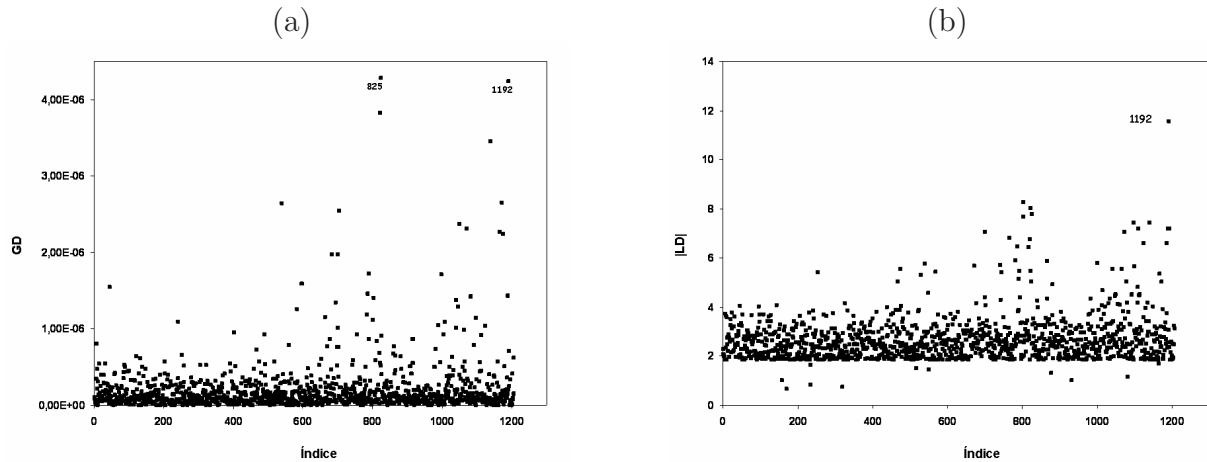


Figura 42 – Gráficos das medidas de influência global para o modelo log-Burr XII ajustado aos dados de suplementação de vitamina A (a) GD_i (b) LD_i

Da Figura 42(a) pode-se notar que dois casos, 825 e 1192, se destacam. Por outro lado, a Figura 42(b) mostra, unicamente, o caso 1192 como uma possível observação influente.

As Figuras 43-45 apresentam os gráficos para a medida de influência local, $|\mathbf{d}_{max}|$, e para a medida de influência local total, C_i , para $\boldsymbol{\theta} = (\lambda, \sigma, \boldsymbol{\beta}^T)^T$ considerando os três esquemas de perturbação estudados na seção 4.5.2, ambas medidas estão descritas na seção 4.2. Da Figura 43(b), o valor de C_i da observação 1192 destaca-se das demais e, portanto, é um possível candidato a ponto influente. As outras figuras não apresentam nenhuma observação candidata a ponto influente.

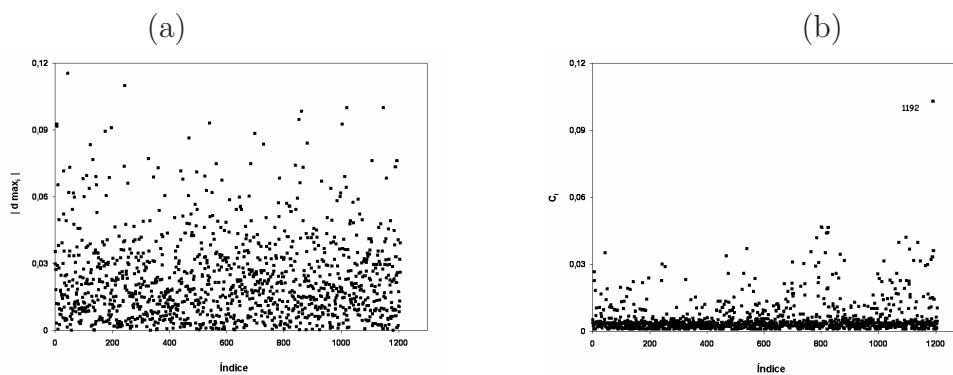


Figura 43 – Gráficos de medidas de influência para θ do modelo de regressão log-Burr XII usando perturbação de casos nos dados de sobrevivência de suplementação de vitamina A (a) Influência local d_{max} (b) Influência local total C_i

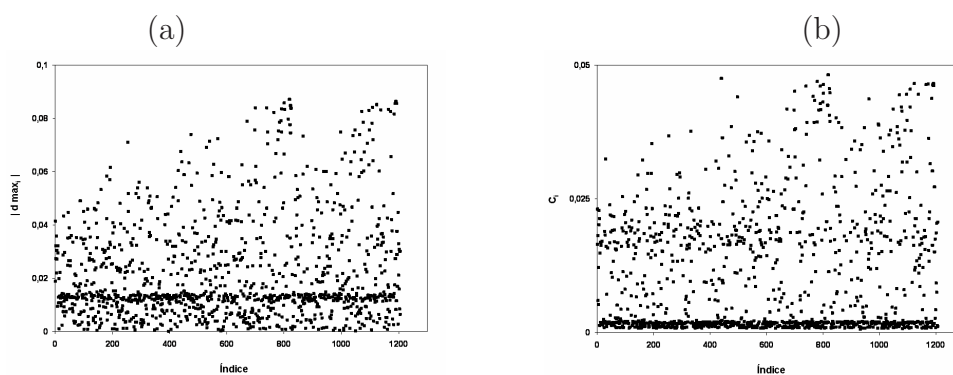


Figura 44 – Gráficos de medidas de influência para θ do modelo de regressão log-Burr XII usando perturbação de resposta nos dados de sobrevivência de suplementação de vitamina A (a) Influência local d_{max} (b) Influência local total C_i

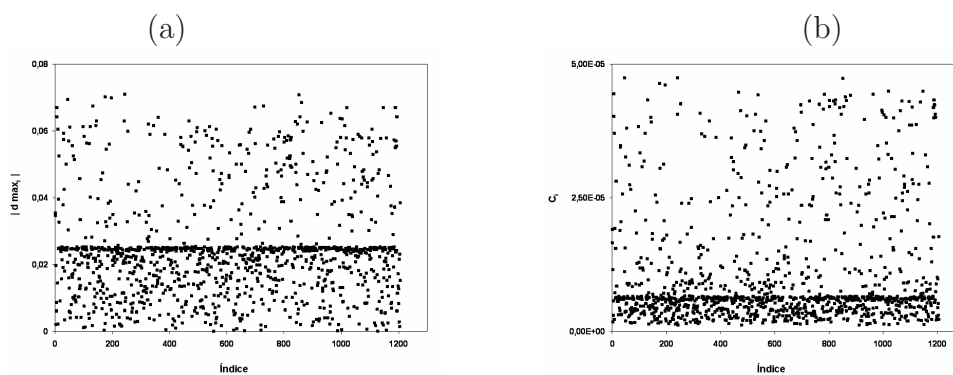


Figura 45 – (a) Gráfico das medidas de influência para d_{max} do modelo de regressão log-Burr XII usando perturbação na variável explicativa idade nos dados de sobrevivência de suplementação de vitamina A (a) influência local d_{max} (b) influência local total C_i

A Figura 46 exibe o gráfico do “leverage” generalizado para o modelo de regressão log-Burr XII, descrito na seção 4.5.3, ajustado aos dados de suplementação de vitamina A e não mostra nenhum ponto como possível candidato a ponto de alavanca.

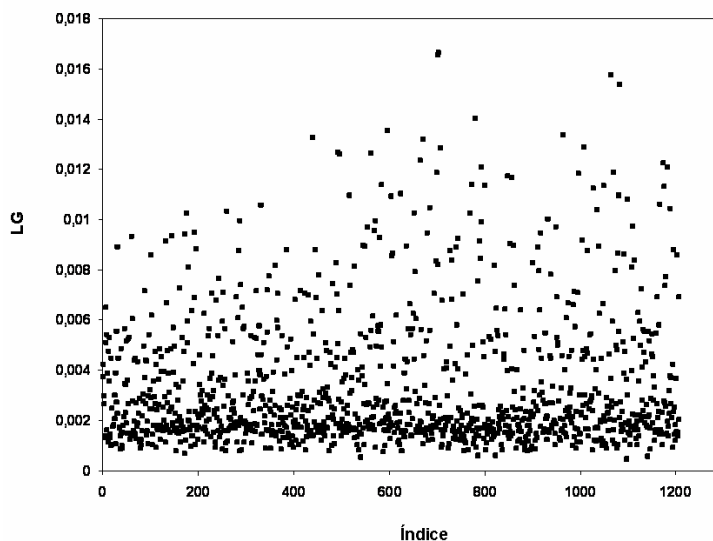


Figura 46 – Gráfico do leverage generalizado para o modelo de regressão log-Burr XII para os dados de suplementação de Vitamina A

Finalmente, as Figuras 47(a) e (b) apresentam os gráficos dos resíduos tipo martingale modificado, dado na seção 5.4. A Figura 47(a) apresenta graficamente o resíduo tipo martingale modificado, para os dados de sobrevivência de suplementação de Vitamina A para o modelo log-Burr XII em que se pode notar assimetria, então, existem evidências de que as suposições do modelo ajustado sejam inadequadas.

Como uma medida de qualidade de ajuste, a Figura 47(b) mostra o gráfico normal de probabilidades para o resíduo tipo martingale modificado com envelope simulado. Nota-se, pelo gráfico, que há indícios de afastamentos sérios da suposição de distribuição Burr XII para os tempos de sobrevivência dos dados de suplementação de Vitamina A.

Portanto, a análise de resíduos mostra evidências que o modelo de regressão log-Burr XII não é adequado para os dados de suplementação de vitamina A. Observando os dados nota-se a ocorrência de muitos empates nos tempos, o que pode justificar a inadequabilidade do referido modelo. Colosimo e Giolo (2006) sugerem que nessa situação um modelo de regressão com dados agrupados é mais conveniente. Vale ressaltar que o resíduo proposto

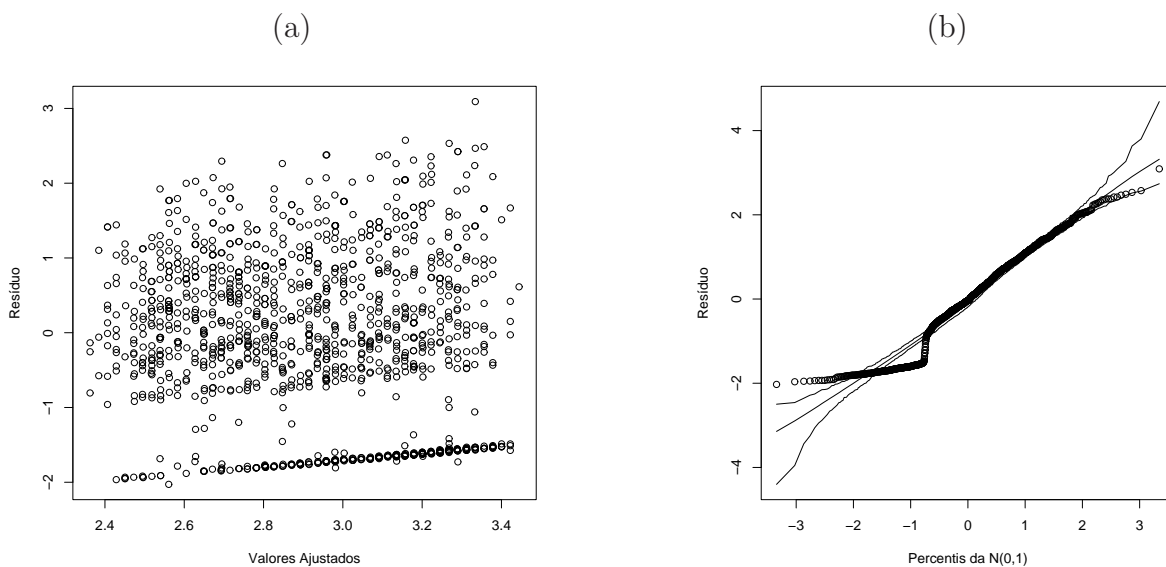


Figura 47 – (a) Gráfico dos resíduos para o modelo log-Burr XII ajustado aos dados de suplementação de Vitamina A (a) martingale (b) tipo martingale modificado

conseguiu detectar uma possibilidade de melhora na modelagem dos dados de suplementação de vitamina A. Assim, sugere-se para a análise desses dados um modelo para dados agrupados considerando a distribuição Burr XII. Diante do exposto, análise desse conjunto de dados não foi concluída.

7 MODELO DE LOCAÇÃO ESCALA PARA DADOS AGRUPADOS

No capítulo 3 foram apresentados modelos de regressão que assumem independência entre os tempos de sobrevivência. Apesar dessa suposição ser válida para algumas aplicações, ela pode ser inadequada para outras aplicações. Assim, é razoável construir modelos de regressão que considerem a existência dessa possível associação entre os tempos e nesta situação um modelo frequentemente usado é o modelo de regressão com efeito aleatório. Este capítulo tem por objetivo apresentar um método de estimação para os parâmetros do modelo de locação e escala, descrito na seção 3.1, estendido pela adição de um efeito aleatório.

7.1 Modelo de regressão com efeito aleatório

Admita, novamente, que os tempos de sobrevivência estão sujeitos a censura à direita, que o mecanismo de censura é aleatório e que a censura é não informativa. Considere uma amostra dividida em m grupos e seja T_{ji} o tempo de sobrevivência do i -ésimo indivíduo no j -ésimo grupo, com $i = 1, \dots, n_j$ e $j = 1, \dots, m$. Sejam \mathbf{x}_{ji} o vetor de covariáveis $p \times 1$ e E_{ji} uma variável aleatória, independente de T_{ji} , que representa o tempo de censura associado a este indivíduo. Assim, os dados observados são $t_{ji} = \min\{T_{ji}, E_{ji}\}$ e

$$\delta_{ji} = \begin{cases} 1 & \text{se } T_{ji} \leq E_{ji} \\ 0 & \text{se } T_{ji} > E_{ji}, \end{cases}$$

que é o indicador de censura.

No modelo de locação e escala com efeito aleatório, todos os indivíduos do mesmo grupo tem um efeito aleatório comum, denotado por W_j , que é uma variável aleatória não observável. Nesse contexto, o modelo de locação e escala com efeito aleatório é representado da seguinte forma

$$Y_{ji} = W_j + \mathbf{x}_{ji}^T \boldsymbol{\beta} + \sigma Z_{ji}, \quad (42)$$

em que $Y_{ji} = \log(T_{ji})$, Z_{ji} 's são os erros aleatórios independentes e identicamente distribuídos e os elementos do vetor $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma)^T$ são parâmetros desconhecidos. Cada grupo tem um

efeito aleatório W_j , que é representado por variáveis aleatórias independentes e identicamente distribuídas com densidade $g(w_j; \mathbf{D})$ com variância ξ e \mathbf{D} é o vetor de parâmetros.

Assuma que $\text{Cov}(W_j, Z_{ji}) = 0$, e que, condicionado ao efeito aleatório W_j , as respostas dentro do grupo j são independentes. Assuma também que os efeitos aleatórios são independentes dos tempos de censura. A covariância entre o logaritmo dos tempos de sobrevivência dentro um grupo é dada por ξ , portanto o modelo reduz-se ao modelo de locação escala usual quando $\xi = 0$.

Para estimar o vetor dos parâmetros $\Psi = (\boldsymbol{\theta}^T, \mathbf{D}^T)^T$ pode-se utilizar o método de máxima verossimilhança que baseia-se na função de verossimilhança que é descrita a seguir.

7.1.1 Verossimilhança Aumentada

Considerando o modelo dado na equação (42), a função de verossimilhança condicional ao efeito aleatório para o i -ésimo indivíduo no j -ésimo grupo é dada por

$$f(y_{ji} | w_j, \mathbf{x}_{ji})^{\delta_{ji}} S(y_{ji} | w_j, \mathbf{x}_{ji})^{1-\delta_{ji}}$$

em que $f(\cdot)$ e $S(\cdot)$ denotam, respectivamente, as funções de densidade e sobrevivência de Y_{ji} condicional ao efeito aleatório W_j .

Para o grupo j , o vetor da variável resposta é representada por $\mathbf{Y}_j = (Y_{j1}, Y_{j2}, \dots, Y_{jn_j})^T$ e de maneira análoga defina $\boldsymbol{\delta}_j$ o vetor contendo o indicador de censuras. A função de verossimilhança condicional ao efeito aleatório (independência dentro o grupo) para os indivíduos do grupo j é dada por

$$L_j(\boldsymbol{\theta} | w_j) = \prod_{i=1}^{n_j} f(y_{ji} | w_j, \mathbf{x}_{ji})^{\delta_{ji}} S(y_{ji} | w_j, \mathbf{x}_{ji})^{1-\delta_{ji}}. \quad (43)$$

Então, a verossimilhança aumentada para Ψ no grupo j é

$$L_j(\Psi; \mathbf{y}_j, \boldsymbol{\delta}_j, w_j) = L_j(\boldsymbol{\theta} | w_j) g(w_j; \mathbf{D}), \quad (44)$$

em que L_j está definida na equação (43).

Assumindo que existe independência entre os vetores $(\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_m)$, a função de verossimilhança aumentada para toda amostra é dada por

$$L(\Psi; \mathbf{y}, \boldsymbol{\delta}, \mathbf{w}) = \prod_{j=1}^m L_j(\Psi; \mathbf{y}_j, \boldsymbol{\delta}_j, w_j), \quad (45)$$

em que $\mathbf{y} = (\mathbf{y}_1^T, \dots, \mathbf{y}_m^T)^T$, $\boldsymbol{\delta} = (\boldsymbol{\delta}_1^T, \dots, \boldsymbol{\delta}_m^T)^T$, $\mathbf{w} = (w_1, \dots, w_m)^T$ e L_j definida na equação (44).

O logaritmo da função de verossimilhança correspondente é

$$l(\boldsymbol{\Psi}; \mathbf{y}, \boldsymbol{\delta}, \mathbf{w}) = \sum_{j=1}^m \log [L_j(\boldsymbol{\theta} | w_j)] + \sum_{j=1}^m \log [g(w_j; \mathbf{D})] \quad (46)$$

esta expressão pode ser separada em duas partes, uma que depende somente do vetor de parâmetros \mathbf{D} e outra que depende do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma)^T$.

7.1.2 Verossimilhança Marginal

A função de verossimilhança marginal para $\boldsymbol{\Psi}$ no grupo j é dada por

$$L_j(\boldsymbol{\Psi}; \mathbf{y}_j, \boldsymbol{\delta}_j, w_j) = \int L_j(\boldsymbol{\theta} | w_j) g(w_j; \mathbf{D}) dw_j,$$

em que L_j está definida na equação (43).

Como os vetores $\mathbf{Y}_j = (Y_{j1}, Y_{j2}, \dots, Y_{jn_j})^T$, $j = 1, \dots, m$, são independentes, o logaritmo da função de verossimilhança marginal para toda amostra é

$$l(\boldsymbol{\Psi}; \mathbf{y}, \boldsymbol{\delta}, \mathbf{w}) = \sum_{j=1}^m \log \left\{ \int L_j(\boldsymbol{\theta} | w_j) g(w_j; \mathbf{D}) dw_j \right\}. \quad (47)$$

Inferência para o vetor de parâmetros $\boldsymbol{\Psi}$ pode ser baseada na verossimilhança marginal para os dados $\mathbf{Y} = (\mathbf{Y}_1^T, \mathbf{Y}_2^T, \dots, \mathbf{Y}_m^T)^T$ que é obtida resolvendo a integral em (47) e maximizando-a com respeito a $\boldsymbol{\Psi}$. Dessa maneira, a verossimilhança marginal depende das distribuições do efeito aleatório e do tempo de sobrevivência.

Entretanto, essa integral, usualmente, não tem solução analítica. Uma possível aproximação da integral na equação (47) pode ser feita abordando, por exemplo, a quadratura gaussiana adaptada (LIU; PIERCE, 1994). A realização desta metodologia pode ser feita por meio do PROC NLMIXED do SAS versão 9 (SAS, 2004). Para mais detalhes ver, por exemplo, Valença (2003), Nelson et al. (2006) ou Liu e Yu (2008).

Neste trabalho é proposto usar o algoritmo EM por Monte Carlo (WEI; TANNER, 1990) considerando censura e efeito aleatório como dados perdidos (*missing*). Na próxima seção é apresentado, de maneira geral, algoritmo EM por Monte Carlo.

7.2 Algoritmo EM por Monte Carlo

Considere que $\mathbf{Y}_c = (\mathbf{Y}, \mathbf{W})$, em que \mathbf{Y} e \mathbf{W} são os dados observados e os dados perdidos (*missing*), respectivamente, representa o conjunto de dados completos com densidade $f(\mathbf{y}_c; \Psi)$ parametrizada por um vetor de parâmetros $\Psi \in \Theta \subseteq R^q$.

O algoritmo EM (DEMPSTER, LAIRD e RUBIN, 1977) em cada iteração é formado pelos passos: Esperança (passo E) e Maximização (passo M). O logaritmo da função de verossimilhança dos dados completos é denotada por $l(\Psi; \mathbf{y}_c) = \log\{f(\mathbf{y}, \mathbf{w}; \Psi)\}$. Então, os passos desse algoritmo para dados incompletos são

Passo E para $\hat{\Psi} = \Psi^{(s)}$, calcular

$$Q(\Psi|\hat{\Psi}^{(s)}) = E[l(\Psi; \mathbf{y}_c)|\mathbf{y}, \hat{\Psi}^{(s)}] = \int l(\Psi; \mathbf{y}_c)g(\mathbf{w}|\mathbf{y}, \hat{\Psi}^{(s)})d\mathbf{w}.$$

Passo M escolher $\Psi^{(s+1)}$ que é um valor de $\Psi \in \Theta$ que maximize $Q(\Psi|\hat{\Psi}^{(s)})$.

Alterna-se os passos E e M até atingir convergência. Sob certas condições apropriadas, a sequência $\{\Psi^{(s)}\}$ obtida por meio do algoritmo EM converge para a estimativa de máxima verossimilhança de Ψ . Para mais detalhes ver Wu (1983) ou Little e Rubin (2002).

Para monitorar a convergência pode-se utilizar a mudança nos valores das estimativas dos parâmetros. Assim, o algoritmo pára quando a mudança nos valores das estimativas dos parâmetros é insignificante (pequena), mas não é claro, em todos os casos, que métrica deve ser usada. Por exemplo, a convergência pode ser verificada quando

$$\max\left(\frac{|\Psi_i^{(s+1)} - \Psi_i^{(s)}|}{|\Psi_i^{(s)}| + \delta_1}\right) < \delta_2$$

em que δ_1 e δ_2 são constantes pré-determinadas. Searle et al. (1992) comentam que vários pesquisadores usam $\delta_1 = 0,001$ e $\delta_2 = 0,0001$. Alternativamente, a convergência pode ser verificada quando $|Q(\Psi^{(s+1)}|\hat{\Psi}^{(s)}) - Q(\Psi^{(s)}|\hat{\Psi}^{(s)})|$ é suficientemente pequena (TANNER, 1991).

Existem casos que a integral no passo E do algoritmo EM não tem forma fechada. Wei e Tanner (1990) introduziram o algoritmo EM por Monte Carlo (EMMC) que substitui a esperança matemática no passo E por uma aproximação Monte Carlo.

Assim, se uma amostra aleatória de $[\mathbf{w}_1^{(s)}, \dots, \mathbf{w}_N^{(s)}]$ é gerada a partir da distribuição $g(\mathbf{w}|\mathbf{y}, \hat{\Psi}^{(s)})$, a esperança matemática do passo E pode ser estimada pela soma de

Monte Carlo

$$Q_N(\Psi|\hat{\Psi}^{(s)}) = E\{l(\Psi; \mathbf{y}_c)|\mathbf{y}, \hat{\Psi}^{(s)}\} = \frac{1}{N} \sum_{n=1}^N l(\Psi; \mathbf{y}, \mathbf{w}_n^{(s)}), \quad (48)$$

em que o sub-escrito N denota a dependência do estimador em relação ao tamanho amostral empregada no método de Monte Carlo e $l(\Psi; \mathbf{y}_c) = \sum_i \log[f(y_i|\mathbf{w})] + \log[g(\mathbf{w}; \mathbf{D})]$. Pela lei dos grandes números, o estimador na equação (48) converge para a esperança teórica no passo E (LEVINE; CASELLA, 2001). Portanto o algoritmo EM por Monte Carlo envolve o uso da quantidade estimada na equação (48) no lugar de Q . Então, o passo M maximiza a soma dada na equação (48) em relação a Ψ . Para mais detalhes ver Chan e Ledolter (1995). Além disso, o passo M pode requerer procedimentos numéricos para maximização.

Então, no passo E do algoritmo EM por Monte Carlo é necessário simular amostras a partir de $g(\mathbf{w}|\mathbf{y}, \Psi^{(s)})$. Existem diversos métodos na literatura para simular tal amostra, dentre eles, destacam-se amostragens a partir do método de Monte Carlo baseado em cadeias de Markov (MCMC), tais como, o algoritmo de Metropolis-Hasting, amostragem de importância ou amostragem com rejeição.

Baseado no trabalho de McCulloch (1997) foi escolhido o algoritmo de Metropolis-Hasting, que para ser usado é necessário especificar uma distribuição proposta, $h(\mathbf{u})$ da qual novos valores são gerados e uma função de aceitação fornece a probabilidade de aceitação do novo valor (e com probabilidade complementar o valor anterior é mantido).

Faça \mathbf{w} denotar uma amostra prévia gerada da distribuição condicional de $\mathbf{w}|\mathbf{y}$ e gere um novo valor \mathbf{w}^* usando a distribuição proposta. Então aceita-se \mathbf{w}^* como novo valor com probabilidade $A_q(\mathbf{w}, \mathbf{w}^*)$; caso contrário, retenha \mathbf{w} . Sendo que, $A_q(\mathbf{w}, \mathbf{w}^*)$ é dado por

$$A_q(\mathbf{w}, \mathbf{w}^*) = \min \left\{ 1, \frac{g(\mathbf{w}^*|\mathbf{y}; \Psi)h(\mathbf{w})}{g(\mathbf{w}|\mathbf{y}; \Psi)h(\mathbf{w}^*)} \right\}. \quad (49)$$

McCulloch (1997) escolheu $h = g$, assim o segundo termo entre as chaves em (49) simplifica para

$$\frac{g(\mathbf{w}^*|\mathbf{y}; \Psi)h(\mathbf{w})}{g(\mathbf{w}|\mathbf{y}; \Psi)h(\mathbf{w}^*)} = \frac{\prod_{i=1}^n f(y_i|\mathbf{w}^*; \Psi)g(\mathbf{w}^*; \mathbf{D})g(\mathbf{w}; \mathbf{D})}{\prod_{i=1}^n f(y_i|\mathbf{w}; \Psi)g(\mathbf{w}; \mathbf{D})g(\mathbf{w}^*; \mathbf{D})} = \frac{\prod_{i=1}^n f(y_i|\mathbf{w}^*; \Psi)}{\prod_{i=1}^n f(y_i|\mathbf{w}; \Psi)} \quad (50)$$

Quando o algoritmo de Metropolis-Hasting é incorporado no algoritmo EM obtém-se o algoritmo EMMC com os seguintes passos:

1. escolha os valores iniciais $\Psi^{(0)}$. Faça $s=0$.
2. Gere N valores, $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N$, a partir de $g(\mathbf{w}|\mathbf{y}; \boldsymbol{\beta}^{(s)}, \sigma^{(s)})$ usando o algoritmo de Metropolis:
 - a. escolha $\boldsymbol{\beta}^{(s+1)}$ e $\sigma^{(s+1)}$ que maximize $\frac{1}{N} \sum_{n=1}^N \log[f(\mathbf{y}|\mathbf{w}_n; \Psi)]$
 - b. escolha $\mathbf{D}^{(s+1)}$ que maximize $\frac{1}{N} \sum_{n=1}^N \log[g(\mathbf{w}_n; \mathbf{D})]$
 - c. Faça $s = s + 1$.
3. Se convergência é obtida, assumo que $\boldsymbol{\beta}^{(s+1)}, \sigma^{(s+1)}$ e $\mathbf{D}^{(s+1)}$ são as estimativas de máxima verossimilhança, caso contrário, retorne ao passo 2.

Duas importantes considerações devem ser levadas em conta na implementação do algoritmo EMMC que são monitoramento da convergência do algoritmo e a especificação do tamanho da amostra, N , do método de Monte Carlo.

Wei e Tanner (1990) notaram que é ineficiente iniciar o algoritmo EMMC com elevados valores de N quando o valor atual da aproximação da estimativa de máxima verossimilhança não está próxima do verdadeiro valor do parâmetro. Então, sugeriram aumentar o valor de N com o número de iterações.

Devido a introdução do erro da simulação de Monte Carlo no passo E, não se pode garantir que o logaritmo da função de verossimilhança seja crescente a cada iteração. Mas, o algoritmo EMMC ainda converge para as estimativas de máxima verossimilhança sob certas condições de regularidade (veja, por exemplo, Chan e Ledolter, 1995). Vaida, Fitzgerald e DeGruttola (2007) propuseram um critério de parada para algoritmo EMMC para modelos de efeitos mistos não-linear com dados censurados usando o desvio padrão empírico do logaritmo da função de verossimilhança.

Alternativamente, convergência pode ser verificada quando a mudança nos valores das estimativas dos parâmetros é insignificante (pequena), como no algoritmo EM, e novamente não é claro que métrica deve ser usada.

Existem diversos trabalhos sobre convergência e de como selecionar o tamanho da amostra do método de Monte Carlo para o algoritmo EMMC, por exemplo, Chan e Ledolter (1995) para modelos de série temporais envolvendo contagens usando o logaritmo da função de verossimilhança e Booth e Hobert (1999) implementaram um algoritmo EMMC para modelos lineares generalizados mixtos de maneira que o tamanho da amostra do Monte Carlo aumenta automaticamente de acordo com a avaliação do erro de Monte Carlo associado com $\Psi^{(s+1)}$.

Entretanto, na situação de interesse a variável resposta y_{ji} não é observada completamente para todo par (j, i) , devido à ocorrência de observações censuradas. Assim, é necessário incorporar ao algoritmo EMMC descrito nesta seção as observações censuradas como dados perdidos. Na próxima seção é apresentado, de maneira geral, o algoritmo EMMC considerando o efeito aleatório e os tempos censurados como dados perdidos.

7.3 Algoritmo EM por Monte Carlo para dados com censura

Os dados observados para o i -ésimo elemento no j -ésimo grupo é $(y_{ji}, T_{ji}, E_{ji}, \delta_{ji})$ em que $y_{ji} = \min\{\log(T_{ji}), \log(E_{ji})\}$ e

$$\delta_{ji} = \begin{cases} 1 & \text{se } T_{ji} \leq E_{ji} \\ 0 & \text{se } T_{ji} > E_{ji}. \end{cases}$$

Para aplicar o algoritmo EMMC, é necessário escrever a função de verossimilhança para os dados completos em vez da função de verossimilhança para modelo com efeito aleatório com observações censuradas, equação (45). Isto é feito aumentando os dados com duas classes de variáveis latentes. Uma é a variável aleatória truncada denotada por v que nos permite considerar uma verossimilhança para observações independente e identicamente distribuída (i.i.d.) sem censura. A outra é o efeito aleatório, w , que converte o modelo com dependência (grupos) para um modelo com elementos amostrais independentes. Assim, pode-se escrever

- dados completos: $(v_{ji}, w_j, y_{ji}, E_{ji}, \delta_{ji})$, $j = 1, \dots, m$ e $i = 1, \dots, n_j$, em que

$$\begin{cases} v_{ji} = y_{ji} & \text{se } \delta_{ji} = 1 \\ v_{ji} > y_{ji} & \text{se } \delta_{ji} = 0. \end{cases}$$

O logaritmo da função de verossimilhança para dados completos do j -ésimo grupo é dado por

$$l_j(\Psi; \mathbf{y}_c) = \sum_{i=1}^{n_j} \left\{ \delta_{ji} \log[f(y_{ji}|w_j; \boldsymbol{\theta})] + (1 - \delta_{ji}) \log[f(v_{ji}|w_j; \boldsymbol{\theta})] \right\} + \log[g(w_j|\mathbf{D})].$$

O, correspondente, logaritmo da função de verossimilhança para dados completos e para toda amostra é dado por

$$l(\Psi; \mathbf{y}_c) = \sum_{j=1}^m l_j(\Psi; \mathbf{y}_c) = \sum_{j=1}^m \sum_{i=1}^{n_j} \left\{ \delta_{ji} \log[f(y_{ji}|w_j; \boldsymbol{\theta})] + (1 - \delta_{ji}) \log[f(v_{ji}|w_j; \boldsymbol{\theta})] \right\} + \sum_{j=1}^m \log[g(w_j|\mathbf{D})]. \quad (51)$$

esta expressão pode ser separada em duas partes, uma que depende somente de \mathbf{D} e outra que depende do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma)^T$. Dessa forma, tem-se que $l(\Psi; \mathbf{y}_c) = l_1(\boldsymbol{\theta}) + l_2(\mathbf{D})$, em que

$$l_1(\boldsymbol{\theta}) = \sum_{j=1}^m \sum_{i=1}^{n_j} \left\{ \delta_{ji} \log[f(y_{ji}|w_j, \boldsymbol{\theta})] + (1 - \delta_{ji}) \log[f(v_{ji}|w_j, \boldsymbol{\theta})] \right\} \quad \text{e}$$

$$l_2(\mathbf{D}) = \sum_{j=1}^m \log[g(w_j|\mathbf{D})].$$

Para cada grupo j , a esperança matemática no passo E do algoritmo EMMC na iteração s passa a ser escrita como

$$Q^j(\Psi|\Psi^{(s)}) = E[l^j(\Psi; \mathbf{y}_c)] = \int \int \left\{ \sum_{i=1}^{n_j} \left\{ \delta_{ji} \log[f(y_{ji}|w_j; \boldsymbol{\theta})] + (1 - \delta_{ji}) \log[f(v_{ji}|w_j; \boldsymbol{\theta})] \right\} + \log g(w_j|\mathbf{D}) \right\} f(\mathbf{v}_j, w_j | \mathbf{y}_j, \boldsymbol{\delta}_j) d\mathbf{v}_j dw_j.$$

Seguindo a mesma idéia do algoritmo EMMC, descrito na sub-seção 7.2, essa esperança matemática é estimada pela soma de Monte Carlo. Então, para implementar o passo E do algoritmo EMMC uma amostra da distribuição condicional de w_j dado y_{ji} avaliada em $\Psi^{(s)}$ para cada j é necessária mas primeiro é preciso uma amostra de \mathbf{v}_j .

Para cada grupo j , gera-se amostras dessas variáveis a partir de $(\mathbf{v}_j, w_j | \mathbf{y}_j, \boldsymbol{\delta}_j, \Psi^{(s)})$ usando as condicionais $(v_{ji} | y_{ji}, \delta_{ji}, w_j, \Psi^{(s)})$ e $(w_j | y_{ji}, v_{ji}, \delta_{ji}, \Psi^{(s)}) = (w_j | v_{ji}, \Psi^{(s)})$ (efeito aleatório independente do tempo de censura). Assim, iterativamente $\mathbf{v} = (\mathbf{v}_1^T, \dots, \mathbf{v}_j^T)^T$ e \mathbf{w} são obtidas a partir dos seguintes passos

- Amostre $v_{ji} \sim f(v_{ji}|w_j, \delta_{ji}, y_{ji}, \Psi)$, para $i = 1, 2, \dots, n_j$. Condicional em w_j , \mathbf{v}_j é um vetor de observações independentes cuja distribuição é truncada.

Para cada elemento (j,i) , $j = 1, \dots, m$ e $i = 1, \dots, n_j$, a variável latente v_{ji} é truncada em t_{ji} se é censurada à direita e é t_{ji} se o tempo de falha é observado exatamente. Então a densidade conjunta dos dados e da variável latente v 's tem componentes i.i.d., cada uma com densidade $f(t)$. Portanto, se o i -ésimo elemento do j -ésimo grupo é censurado à direita em t_{ji} , então a variável latente v_{ji} é gerada da densidade truncada $f(v_{ji})/[1 - F(t_{ji})]\mathbf{I}(t_{ji} < v_{ji} < \infty)$ em que f e F são as funções de densidade e de distribuição, respectivamente, da variável t .

- Para cada grupo j , amostre $w_j \sim g(w_j|\mathbf{v}_j, \Psi)$, $j = 1, \dots, m$, usando o algoritmo de Metropolis descrito na sub-seção 7.2 em que a equação (50) é obtida usando

$$g(w_j|\mathbf{v}_j; \Psi^{(s)}) = \frac{\prod_{i=1}^{n_j} f(v_{ji}|w_j; \Psi^{(s)})g(w_j; \Psi^{(s)})}{\int \prod_{i=1}^{n_j} f(v_{ji}|w_j; \Psi^{(s)})g(w_j; \Psi^{(s)})dw_j} \propto \prod_{i=1}^{n_j} f(v_{ji}|w_j; \Psi^{(s)})g(w_j; \Psi^{(s)}).$$

As amostras são simuladas separadas por grupo devido ao fato que foi suposto que as observações de grupos diferentes são independentes.

Para cada par (j, i) , gera-se v_{ji} e usando o algoritmo de Metropolis-Hasting, como descrito anteriormente, obtém-se $[w_j^{(1)}, \dots, w_j^{(N)}]$, $j = 1, \dots, m$, assim o passo E do algoritmo EMMC na iteração $(s + 1)$ pode ser escrito como

$$\begin{aligned} Q_N(\Psi|\Psi^{(s)}) &= \sum_{j=1}^m Q^j(\Psi|\Psi^{(s)}) = \sum_{j=1}^m \left[\frac{1}{N} \sum_{n=1}^N l^j(\Psi; \mathbf{v}_j, w_j^{(n)}) \right] \\ &= \sum_{j=1}^m \left\{ \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^{n_j} \left[\delta_{ji} \log[f(y_{ji}|w_j^{(n)}, \boldsymbol{\theta})] + (1 - \delta_{ji}) \log[f(v_{ji}|w_j^{(n)}, \boldsymbol{\theta})] \right] \right\} + \\ &\quad \sum_{j=1}^m \left[\frac{1}{N} \sum_{n=1}^N \log g(w_j^{(n)}|\mathbf{D}) \right], \end{aligned}$$

observe que esta expressão pode ser separada em duas partes, uma que depende somente de \mathbf{D} e outra que depende do vetor de parâmetros $\boldsymbol{\theta}$.

Em resumo, o algoritmo EM por Monte Carlo para modelo de regressão com efeito aleatório e dados censurados segue os seguintes passos:

1. Escolha os valores iniciais $\Psi^{(0)}$. Faça $s=0$.

2. Dado $\Psi^{(0)}$, \mathbf{t} , $\boldsymbol{\delta}$, gere \mathbf{v} .
3. Gere N valores, $[w_j^{(1)}, \dots, w_j^{(N)}]$ usando o algoritmo de Metropolis-Hasting, descrito anteriormente:
 - a. escolha $\boldsymbol{\beta}^{(s+1)}$, $\sigma^{(s+1)}$ que maximize $\sum_{j=1}^m \left\{ \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^{n_j} \left[\delta_{ji} \log[f(y_{ji}|w_j^{(n)}; \boldsymbol{\theta})] + (1 - \delta_{ji}) \log[f(v_{ji}^{(n)}|w_j^{(n)}, \boldsymbol{\theta})] \right] \right\}$
 - b. escolha $\mathbf{D}^{(s+1)}$ para maximizar $\sum_{j=1}^m \frac{1}{N} \sum_{n=1}^N \log[g(w_j^{(n)}|\mathbf{D})]$
 - c. Faça $s = s + 1$.
4. Se convergência é obtida, assuma que $\boldsymbol{\beta}^{(s+1)}$, $\sigma^{(s+1)}$ e $\mathbf{D}^{(s+1)}$ são as estimativas de máxima verossimilhança, caso contrário, retorne ao passo 2.

Neste trabalho, a convergência do algoritmo EMMC para dados censurados foi verificada quando a mudança nas estimativas dos parâmetros em iterações sucessivas é pequena. O algoritmo é executado um pré-determinado número de iterações até a convergência ser obtida, veja, por exemplo, McCulloch (1994,1997) e Chan e Kuk (1997).

A matriz de variância e covariância de $\widehat{\Psi}$, obtida pelo algoritmo EMMC, do vetor de parâmetros Ψ do modelo dado na equação (42), pode ser estimada usando a inversa da matriz de informação de Fisher. Usando resultado de Louis (1982), a matriz de informação observada para Ψ é dada pela diferença entre duas matrizes definidas positivas

$$\mathbf{I}_{\mathbf{Y}}(\Psi) = -E \left[\frac{\partial^2 l(\Psi|\mathbf{y}_c)}{\partial \Psi \Psi'} \middle| \mathbf{y}, \widehat{\Psi} \right] - V \left[\frac{\partial l(\Psi|\mathbf{y}_c)}{\partial \Psi} \middle| \mathbf{y}, \widehat{\Psi} \right]. \quad (52)$$

em que ambos são avaliados em $\widehat{\Psi}$.

O primeiro termo do lado direito da equação (52) pode ser aproximado pela estimativa de Monte Carlo dada para $\left[-\frac{\partial^2 l(\Psi|\mathbf{y}_c)}{\partial \Psi \Psi'} \middle| \mathbf{y}, \widehat{\Psi} \right]$ na última interação do algoritmo e a estimativa de Monte Carlo para o segundo termo pode ser obtida de simulações a partir da última interação (TANNER, 1996).

Assim, o erro padrão de $\widehat{\Psi}$ pode ser estimado a partir de $1/\sqrt{\widehat{\mathbf{I}}}$ avaliada na última interação do algoritmo EMMC. Mas devido ao erro de Monte Carlo, essa diferença

pode nem sempre ser uma matriz definida positiva na prática. Nesse caso, pode-se aproximar a matriz de informação de Fisher observada diretamente via Monte Carlo depois que a convergência é obtida ou usar o erro padrão empírico que é dado pelo desvio padrão das estimativas obtidas nas iterações. Para, mais detalhes, veja, Chen, Zhang e Davidian (2002).

A seguir descreve-se as expressões para o algoritmo EMMC para dados censurados considerando que o tempo de sobrevivência tenha distribuição Burr XII.

7.4 Modelo de Regressão log-Burr XII com Efeito Aleatório

Considere que $Y = \log(T)$ seja uma variável aleatória com distribuição log-Burr XII, com densidade dada na equação (15), e função de sobrevivência na equação (16). Neste caso, para o modelo de locação escala com efeito aleatório dado na equação (42), a densidade de Y para o i -ésimo indivíduo no j -ésimo grupo, dado \mathbf{x}_{ji} e w_j é definida como

$$f(y_{ji}|w_j; k, \sigma, \boldsymbol{\beta}, \mathbf{x}_{ji}) = \frac{k}{\sigma} \left[1 + \exp\left(\frac{y_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma}\right) \right]^{-(k+1)} \exp\left(\frac{y_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma}\right). \quad (53)$$

A função de sobrevivência correspondente é dada por

$$S(y_{ji}|w_j; k, \sigma, \boldsymbol{\beta}, \mathbf{x}_{ji}) = \left[1 + \exp\left(\frac{y_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma}\right) \right]^{-k}. \quad (54)$$

Assuma para as variáveis aleatórias W_j , $j = 1, \dots, m$, a distribuição $\Gamma(\frac{1}{\xi}, \frac{1}{\xi})$ que apresenta função de densidade dada por:

$$g\left(w; \frac{1}{\xi}, \frac{1}{\xi}\right) = \frac{1}{\xi} w^{\frac{1}{\xi}-1} \exp\left(-\frac{w}{\xi}\right), \quad (55)$$

em que $w \geq 0$. Segue, então, que $E(W_j) = 1$ e $V(W_j) = \xi$. Se $\xi = 0$ então obtém-se o modelo usual de locação e escala. Para mais detalhes, veja, Colosimo e Giolo (2006) ou Li (2006).

De acordo com a seção (7.3), o procedimento de estimação é baseado na construção do logaritmo da função de verossimilhança para os dados completos. Para o modelo de regressão log-Burr XII com efeito aleatório seguindo distribuição dada na equação (55), o logaritmo da função de verossimilhança para os dados completos, definida na equação (51), é dado por:

$$l(\boldsymbol{\beta}, k, \sigma, \xi; \mathbf{y}_c) = \sum_{j=1}^m \left\{ \sum_{i=1}^{n_j} \delta_{ji} \left[\log\left(\frac{k}{\sigma}\right) - (k+1) \log\left(1 + \exp\left(\frac{y_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma}\right)\right) \right] + \right.$$

$$\begin{aligned} & \left(\frac{y_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma} \right) \Big] - (1 - \delta_{ji}) \left[\log \left(\frac{k}{\sigma} \right) - (k + 1) \times \right. \\ & \left. \log \left(1 + \exp \left(\frac{v_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma} \right) \right) + \left(\frac{v_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma} \right) \right] \Big\} + \\ & \left(\frac{1}{\xi} - 1 \right) \sum_{j=1}^m \log(w_j) - \frac{1}{\xi} \sum_{j=1}^m w_j - m \left\{ \frac{1}{\xi} \log(\xi) + \log \left[\Gamma \left(\frac{1}{\xi} \right) \right] \right\} \end{aligned}$$

esta expressão pode ser separada em duas partes, uma que depende somente de ξ e outra que depende dos parâmetros $\boldsymbol{\beta}^T, k, \sigma$. Dessa forma, tem-se que $l(\boldsymbol{\beta}, k, \sigma, \xi; \mathbf{y}_c) = l_1(\boldsymbol{\beta}, k, \sigma; \mathbf{y}_c) + l_2(\xi; \mathbf{y}_c)$ em que

$$\begin{aligned} l_1(\boldsymbol{\beta}, k, \sigma) &= \sum_{j=1}^m \left\{ \sum_{i=1}^{n_j} \delta_{ji} \left[\log \left(\frac{k}{\sigma} \right) - (k + 1) \log \left(1 + \exp \left(\frac{y_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma} \right) \right) \right. \right. \\ & \left. \left. \left(\frac{y_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma} \right) \right] - (1 - \delta_{ji}) \left[\log \left(\frac{k}{\sigma} \right) - (k + 1) \times \right. \right. \\ & \left. \left. \log \left(1 + \exp \left(\frac{v_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma} \right) \right) + \left(\frac{v_{ji} - \mathbf{x}_{ji}^T \boldsymbol{\beta} - w_j}{\sigma} \right) \right] \right\} \end{aligned} \quad (56)$$

$$l_2(\xi) = \left(\frac{1}{\xi} - 1 \right) \sum_{j=1}^m \log(w_j) - \frac{1}{\xi} \sum_{j=1}^m w_j - m \left\{ \frac{1}{\xi} \log(\xi) + \log \left[\Gamma \left(\frac{1}{\xi} \right) \right] \right\}. \quad (57)$$

Para esse modelo, a implementação do algoritmo EMMC para dados censurados depende da distribuição Burr truncada cuja densidade é dada por

$$f(v|v > t; k, c, s) = kc \left[1 + \left(\frac{v}{s} \right)^c \right]^{-(k+1)} \left(\frac{v^{(c-1)}}{s^c} \right) \left[1 + \left(\frac{t}{s} \right)^c \right]^k.$$

A correpondente função de distribuição é dada por

$$F(v|v > t; k, c, s) = 1 - \left[1 + \left(\frac{v}{s} \right)^c \right]^{-(k+1)} \left[1 + \left(\frac{t}{s} \right)^c \right]^k.$$

Amostra aleatória da distribuição Burr truncada pode ser gerada de

$$v_{ji} = s \left\{ \left[(1 - u)^{-1/k} \left(\frac{t_{ji}}{s} \right)^c \right] - 1 \right\}^{1/c}, \quad (58)$$

em que u tem distribuição uniforme no intervalo $[0, 1]$, $\sigma = \frac{1}{c}$, $s = \exp(\mathbf{x}_{ji}^T \boldsymbol{\beta} + w_j)$ e w_j tem distribuição $\Gamma(1/\xi, 1/\xi)$. Essa expressão foi obtida do método da transformação integral.

Assim, o algoritmo EMMC para dados com censura, descrito na seção 7.3, assume a seguinte forma:

1. Escolha os valores iniciais $\beta^{(0)}$, $k^{(0)}$, $\sigma^{(0)}$ e $\xi^{(0)}$. Faça $s=0$.
2. Impute os tempos censurados a partir da equação (58) e obtenha o vetor \mathbf{v} .
3. Gere N valores, $[w_j^{(1)}, \dots, w_j^{(N)}]$ e calcule (avaliadas em $\beta^{(s)}$, $k^{(s)}$, $\sigma^{(s)}$ e $\xi^{(s)}$):
 - a. $\beta^{(s+1)}$, $k^{(s+1)}$, $\sigma^{(s+1)}$ que maximiza $\frac{1}{N} \sum_{n=1}^N l_1(\beta, k, \sigma; \mathbf{v}, \mathbf{w}^{(n)})$ em que l_1 é dado em (56).
 - b. $\xi^{(s+1)}$ que maximiza $\frac{1}{N} \sum_{n=1}^N l_2(\xi; \mathbf{v}, \mathbf{w}^{(n)})$ em que l_2 é dado em (57).
 - c. Faça $s = s + 1$.
4. Se convergência é obtida, assuma que $\beta^{(s+1)}$, $k^{(s+1)}$, $\sigma^{(s+1)}$ e $\xi^{(s+1)}$ são as estimativas de máxima verossimilhança, caso contrário, retorne ao passo 2.

O efeito aleatório w_j é gerado a partir do algoritmo de Metropolis-Hasting, descrito na seção 7.2, em que a função f é dada na equação (53) e a função g na equação (55). As maximizações no item 3 foram feitas utilizando o algoritmo quase-Newton baseado no método BFGS (NOCEDAL; WRIGHT, 2006; PRESS et al., 1992).

7.5 Exemplo Simulado

Para investigar o desempenho do algoritmo EMMC proposto na sub-seção 7.4 foi simulado uma amostra em que os dados são gerados com estrutura de grupos, ou seja, com correlação dentre os grupos, e observações censuradas.

Os erros Z_{ji} foram gerados dentro de cada grupo a partir da densidade dada na equação (18) e as variáveis W_j foram geradas como variáveis gama, densidade dada na equação (55) e $\xi = 3$, independentes e identicamente distribuídas. Os tempos de censura E_{ji} foram gerados conforme variáveis independentes com distribuição uniforme definida no intervalo $[0, \eta]$ em que η foi escolhido de maneira que a porcentagem de censura, 10, dentro cada grupo fosse alcançada.

Os logaritmos dos tempos de sobrevivência, Y_{ji} , foram obtidos do modelo de locação e escala, dado na equação (42), com uma única covariável x_{ji} , cujos valores foram

extraídos da distribuição uniforme definida no intervalo $[0, 1]$. Os verdadeiros valores para os coeficientes do modelo foram $\beta_0 = 1,5$ e $\beta_1 = 1,5$.

Para esse exemplo, foi gerada uma amostra com quatro grupos sendo cada grupo com tamanho de amostra igual a 30. Os valores dos parâmetros foram fixados em $\sigma = 0,5$ que determina função de taxa de falha unimodal, $k = 0,25$ e assumindo $y_{ji} = \min\{\log(E_{ji}), \log(T_{ji})\}$ em que o indicador de censura é $\delta_{ji} = I_{T_{ji} \leq E_{ji}}$, sendo $I(\cdot)$ a função indicadora.

Os procedimentos para as simulações foram construídos e executados na linguagem de programação matricial Ox (DOORNIK, 2006). Para implementar o algoritmo EMMC descrito na sub-seção 7.4, inicialmente, foram imputadas as observações censuradas usando a equação (58).

A partir da amostra com os dados censurados imputados foi gerada uma cadeia com tamanho 20.000 para cada efeito aleatório usando o algoritmo de Metropolis-Hasting, descrito na seção 7.2. As primeiras 2.000 iterações foram descartadas para eliminar o efeito dos valores iniciais e para evitar problemas de correlação foi considerado um espaço de tamanho 18 entre os valores gerados. Desse modo foi obtida uma amostra de tamanho 1.001. Para monitorar a convergência do algoritmo de Metropolis-Hasting foram usados gráficos. Foram executadas 100 iterações e o valor inicial para todos os W 's no algoritmo EMMC foi zero. O programa encontra-se no apêndice J.

Os valores dos parâmetros fornecidos pela última iteração foram usados para reinicializar novamente o algoritmo EMMC mas, agora, o tamanho da cadeia de Markov dada pelo algoritmo de Metropolis-Hasting foi aumentado para 2001 e os W 's foram inicializados com o último valor dado à eles pelas cadeias geradas na última iteração. O algoritmo EMMC foi executado mais 100 vezes.

Para o algoritmo de Metropolis-Hasting, a taxa de aceitação (porcentagem de amostras aceitas) está entre 15 e 60 por cento e a convergência das cadeias para cada efeito aleatório foi verificada graficamente, como se verifica da visualização gráfica do traço, dada nas Figuras 48 e 49 para as iterações 100 e 200.

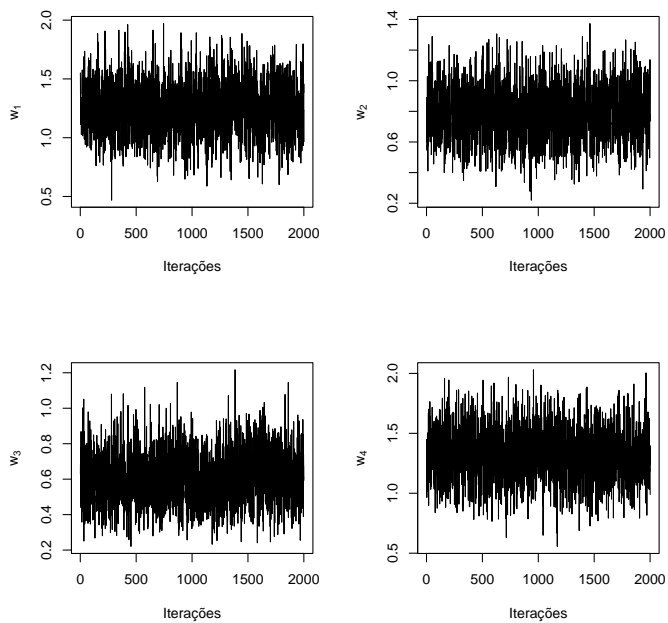


Figura 48 – Traço para os efeitos aleatório gerados na 100^a iteração

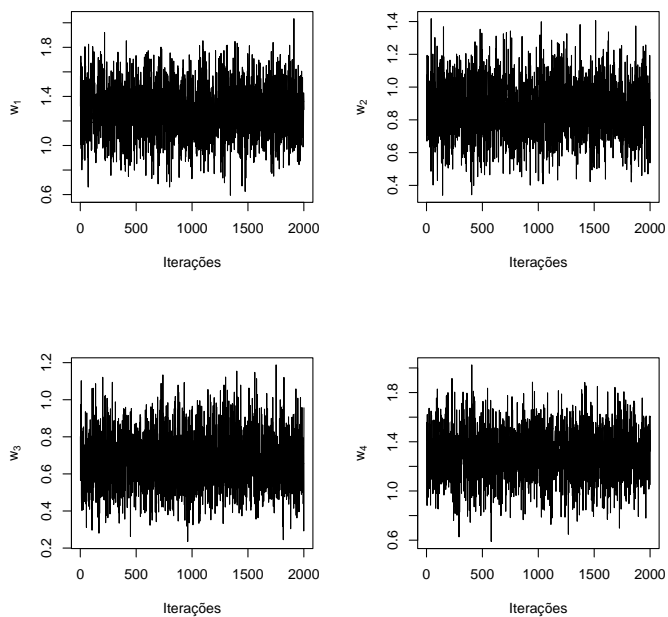


Figura 49 – Traço para os efeitos aleatório gerados na última iteração

A Figura 50 mostra os gráficos contendo os valores dos parâmetros para cada iteração. Nota-se uma certa estabilização dos valores para cada parâmetro. Para verificar convergência, a Figura 51 exibe a diferença entre os sucessivos valores dos parâmetros e observa-se que esses valores não se estabilizam devido à mudança de valor na segunda casa decimal, exceto para o parâmetro ξ , mas as mudanças são aleatórias em torno do zero. Essa situação não melhora mesmo que o tamanho da cadeia dada pelo algoritmo de Metropolis-Hasting seja aumentado, o método BFGS que foi usado para a maximização nos passos do algoritmo EMMC passa a ter problema de convergência, ou mesmo aumentando o número de iterações. Um estudo para o cálculo do erro padrão das estimativas dos parâmetros não foi realizado devido à falta de convergência.

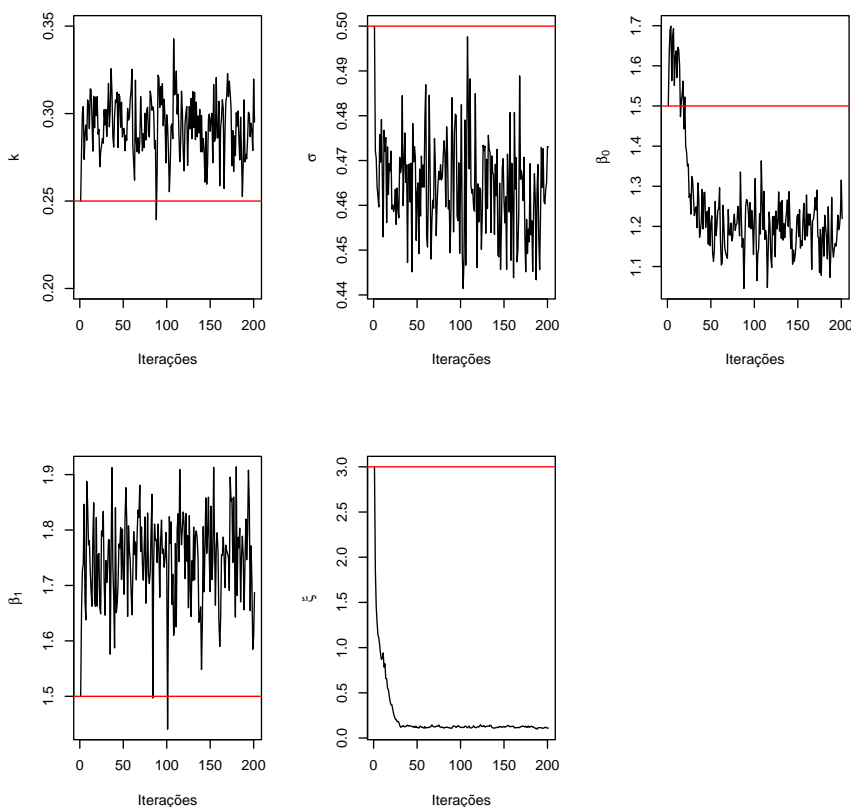


Figura 50 – Estimativas dos parâmetros a cada iteração do algoritmo EM por Monte Carlo usando o algoritmo de Metropolis-Hasting com $N=1001$ e 2001 para as iterações 1-100 e 101-200; a linha horizontal representa os verdadeiros valores dos parâmetros que foram os valores iniciais - modelo log-Burr XII com efeito aleatório gama aplicado em dados simulados

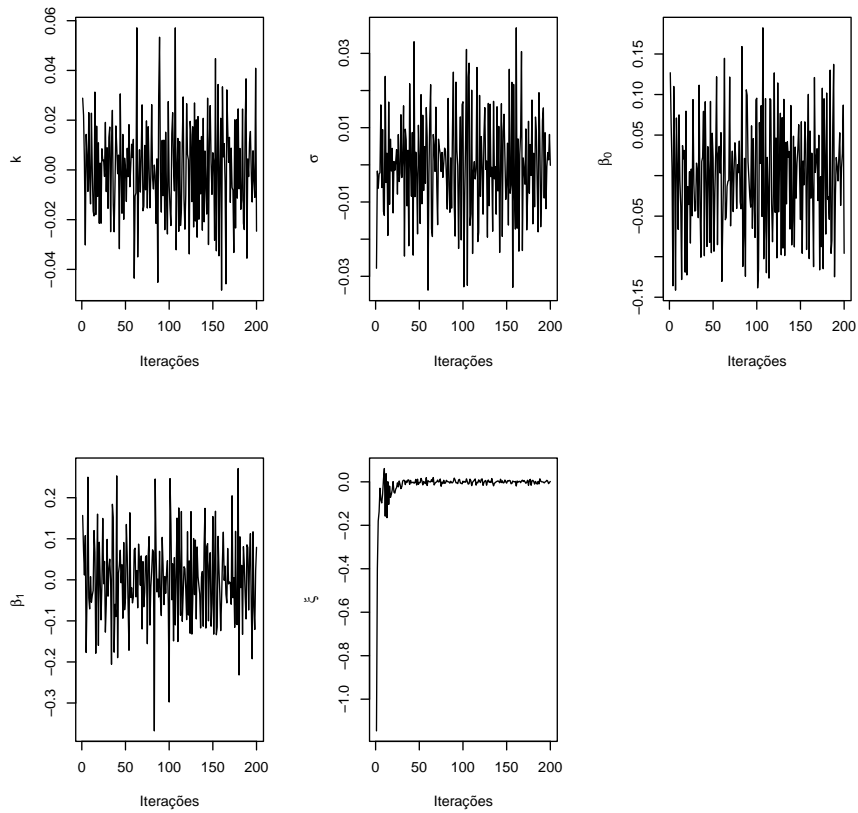


Figura 51 – Diferenças entre os valores sucessivos dos parâmetros a cada iteração do algoritmo EM por Monte Carlo usando o algoritmo de Metropolis-Hasting com $N=1001$ e 2001 para as iterações 1-100 e 101-200 - modelo log-Burr XII com efeito aleatório gama aplicado em dados simulados

Para verificar os resultados obtidos pelo algoritmo EMMC proposto neste capítulo, estimativas para os parâmetros do modelo de regressão log-Burr XII com efeito aleatório gama usando a verossimilhança marginal por meio do PROC NLMIXED SAS (NELSON et al., 2006) também foram obtidas (Tabela 21). O programa está no apêndice K.

Nota-se que as estimativas dadas pelo SAS estão próximas dos valores usados para simular os dados. Se a convergência fosse alcançada usando o algoritmo EMMC proposto neste trabalho os parâmetros provavelmente seriam estimados por valores similares aos dados pelo SAS, exceto para o β_0 e ξ .

O algoritmo EMMC precisa ser melhorado. De acordo com Booth e Hobert (1999), para estimar os parâmetros do modelo linear misto generalizado usando o algoritmo EMMC tem-se que, um método de amostragem com rejeição ou amostragem de importância

Tabela 21 – Estimativas dos parâmetros do modelo de regressão log-Burr XII com efeito aleatório gama para dados simulados usando a verossimilhança marginal (SAS)

Parâmetros	Estimativas	Erro Padrão
k	0,2950	0,0715
σ	0,4580	0,0756
β_0	1,7689	0,3317
β_1	1,7464	0,4075
ξ	2,9238	1,8415

precisam de menos iterações que o algoritmo de Metropolis-Hasting como proposto McCulloch (1997), que foi o usado neste trabalho, para se obter convergência. Assim, usando outro método de amostragem a convergência pode ser que seja obtida.

8 NOVA DISTRIBUIÇÃO: BETA WEIBULL MODIFICADA

Nesta seção, é proposta uma nova distribuição com cinco parâmetros, denotada por distribuição beta Weibull modificada (BWM). Essa nova distribuição apresenta flexibilidade para acomodar várias formas da função de taxa de falha e é apropriada na discriminação entre alguns modelos probabilísticos alternativos, tais como, Weibull (W), Weibull exponenciada (WE), e Weibull modificada generalizada (WMG).

8.1 Definição do novo modelo probabilístico

A idéia na qual a distribuição BWM tem origem é uma classe geral de distribuições definida da seguinte maneira: se G denota a função distribuição acumulada (FDA) de uma variável aleatória então uma classe de distribuições generalizadas pode ser definida por

$$F(x) = I_{G(x)}(a, b) = \frac{1}{B(a, b)} \int_0^{G(x)} \omega^{a-1} (1 - \omega)^{b-1} d\omega \quad (59)$$

para $a > 0$ e $b > 0$, em que $I_y(a, b) = B_y(a, b)/B(a, b)$ denota a função razão beta incompleta e $B_y(a, b) = \int_0^y w^{a-1} (1 - w)^{b-1} dw$ que denota a função beta incompleta (ABRAMOWITZ; SEGUN, 1968)

Esta classe de generalizações de distribuições tem recebido considerável atenção nos últimos anos, em particular depois dos recentes trabalhos de Eugene, Lee e Famoye (2002) e Jones (2004). Eugene, Lee e Famoye (2002) introduziram o que é conhecido como a distribuição beta normal (BN) por considerar $G(x)$ na equação (59) como sendo a FDA da normal e calcularam alguns dos seus primeiros momentos. Expressões mais gerais para esses momentos foram obtidas por Gupta e Nadarajah (2004). Nadarajah e Kotz (2004) introduziram a distribuição beta Gumbel (BG) por considerar $G(x)$ como sendo a FDA da distribuição Gumbel e forneceram expressões de forma fechada para os momentos, a distribuição assintótica das estatísticas ordinais extremas e discutiram o procedimento de estimação de máxima verossimilhança. Nadarajah e Gupta (2004) introduziram a distribuição beta Fréchet (BF) por considerar $G(x)$ com sendo a distribuição de Fréchet, obtiveram a forma analítica da função densidade de probabilidade e da função de taxa de falha, além de calcularem a distribuição

assintótica das estatísticas ordinais extremas. Também, Nadarajah e Kotz (2006) apresentaram a distribuição beta exponencial (BE) e obtiveram a função geradora dos momentos, os quatro primeiros cumulantes, a distribuição assintótica das estatísticas ordinais extremas e discutiram o procedimento de estimação de máxima verossimilhança. Outra distribuição que passa a pertencer a equação (59) é a distribuição beta logística, estudada por Brown, Floyd e Levy (2002), mesmo que não tenha sido obtida diretamente desta equação.

A motivação para introduzir a distribuição BWM é devido as generalizações citadas acima, o amplo uso da distribuição Weibull e o fato de que a atual generalização fornece meios para a sua contínua extensão à situações ainda mais complexas.

Lai, Xie e Murthy (2003) introduziram a distribuição Weibull modificada (WM) que tem três parâmetros $\alpha > 0$, $\lambda > 0$ e $\gamma \geq 0$ com FDA e função de densidade de probabilidade dadas por

$$G_{\alpha,\gamma,\lambda}(x) = 1 - \exp[-\alpha x^\gamma \exp(\lambda x)] \quad (60)$$

e

$$g_{\alpha,\gamma,\lambda}(x) = \alpha x^{\gamma-1}(\gamma + \lambda x) \exp[\lambda x - \alpha x^\gamma \exp(\lambda x)], \quad x > 0, \quad (61)$$

respectivamente. Os parâmetros α e γ controlam a escala e a forma da distribuição, respectivamente. O parâmetro λ é uma espécie de fator de aceleração do tempo de sobrevivência e funciona como um fator de fragilidade na sobrevivência do indivíduo quando o tempo aumenta. A distribuição Weibull é um caso especial de (61) quando $\lambda = 0$. Se, adicionalmente a $\lambda = 0$, $\gamma = 1$ e $\gamma = 2$, obtem-se, respectivamente, as distribuições exponencial (E) e Rayleigh (R).

A função de densidade de probabilidade correspondente a equação (59) pode ser escrita da seguinte maneira

$$f(x) = \frac{1}{B(a, b)} G(x)^{a-1} [1 - G(x)]^{b-1} g(x), \quad (62)$$

em que $g(x) = dG(x)/dx$ é a densidade da distribuição geradora. A função de densidade de probabilidade $f(x)$ será mais tratável quando as funções $G(x)$ e $g(x)$ têm expressões analíticas

simples como é o caso da distribuição WM. Exceto para alguns casos especiais escolhidos para $G(x)$ na equação (59), poderá acontecer que a equação (62) seja difícil de tratar.

Substituindo $G(x)$ na equação (59) pela FDA da distribuição WM, dada na equação (60), obtem-se a distribuição BWM. Então, a forma da função de distribuição acumulada da distribuição BWM é dada por

$$F(x) = \frac{1}{B(a, b)} \int_0^{1 - \exp\{-\alpha x^\gamma \exp(\lambda x)\}} \omega^{a-1} (1 - \omega)^{b-1} d\omega. \quad (63)$$

A função de densidade para BWM é obtida substituindo as equações (60) e (61) na equação (62), e pode ser escrita como

$$f(x) = \frac{\alpha x^{\gamma-1} (\gamma + \lambda x) \exp(\lambda x)}{B(a, b)} [1 - \exp\{-\alpha x^\gamma \exp(\lambda x)\}]^{a-1} \exp\{-b\alpha x^\gamma \exp(\lambda x)\}, \quad x > 0. \quad (64)$$

A notação utilizada será $X \sim \text{BWM}(a, b, \alpha, \lambda, \gamma)$.

A distribuição BWM contém como caso especiais várias distribuições conhecidas. Por exemplo, quando $\lambda = 0$ simplifica para a distribuição beta Weibull (BW). Se $\gamma = 1$ além de $\lambda = 0$, obtém-se a distribuição beta exponencial (BE). A distribuição WMG é um caso especial quando $b = 1$. Se $a = 1$ além de $b = 1$, a distribuição WM é obtida. Para $b = 1$ e $\lambda = 0$, a distribuição BWM reduz-se a distribuição WE. Se $\gamma = 1$ além de $b = 1$ e $\lambda = 0$, a distribuição BWM torna-se a distribuição exponencial exponenciada (EE) apresentada por Gupta e Kundu (2001). Para $\gamma = 2$, $\lambda = 0$ e $b = 1$, a distribuição BWM torna-se distribuição Rayleigh generalizada (RG). A distribuição Weibull é um caso especial quando $a = b = 1$ e $\lambda = 0$. Os casos particulares da distribuição BWM são ilustrados na Figura 52, em que os casos conhecidos da literatura e não definidos anteriormente são: beta Rayleigh modificada (BRM), beta exponencial modificada (BEM), Rayleigh modificada generalizada (RMG), exponencial modificada generalizada (EMG), beta Rayleigh (BR), Rayleigh modificada (RM) e exponencial modificada (EM).

A distribuição BWM pode ser simulada a partir da equação (63) como segue: se V tem função de densidade de probabilidade beta com parâmetros a e b , então a solução da equação não linear $X^\gamma \exp(\lambda X) = -\alpha^{-1} \log(1 - V)$ tem distribuição $\text{BWM}(a, b, \alpha, \gamma, \lambda)$. Para simular dados a partir dessa equação não linear, foi usada a linguagem de programação ma-

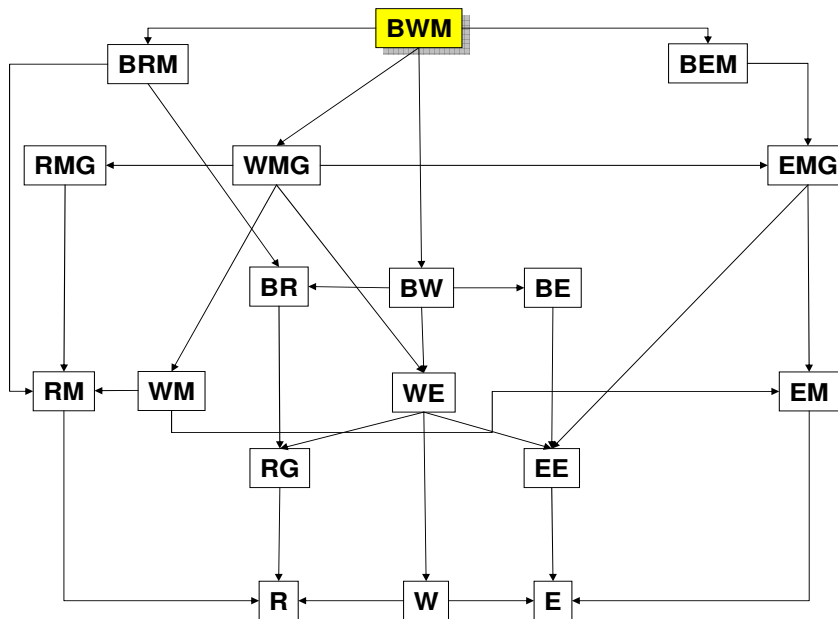


Figura 52: Diagrama das distribuições obtidas como casos particulares da distribuição BWM.

tricial Ox por meio da sub-rotina *SolveNLE* (programa no apêndice L). Gráficos comparando a densidade exata da distribuição BWM e o histograma do conjunto de dados simulado para alguns valores dos parâmetros são dados na Figura 53. Nota-se dessa figura que os valores simulados são consistentes com a distribuição BWM.

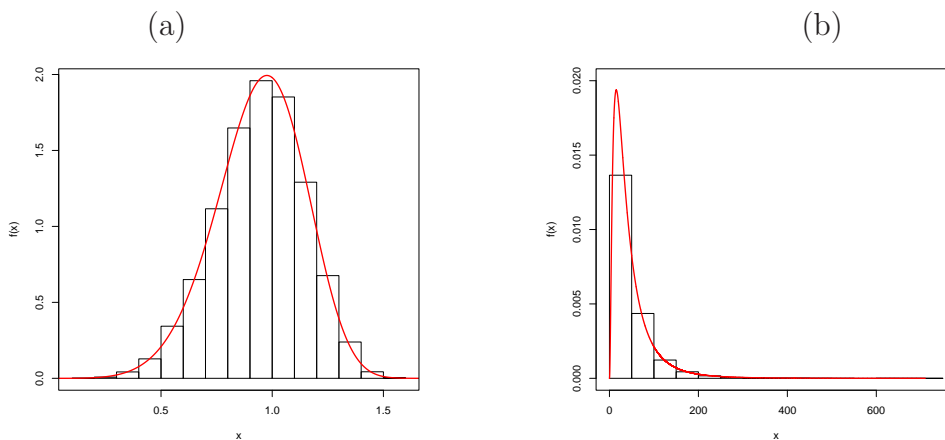


Figura 53 – Gráficos da função de densidade exata da distribuição BWM com histograma para dados simulados com: (a) $a = 1,5, b = 0,5, \alpha = 0,5, \lambda = 1,5, \gamma = 3$ e (b) $a = 8, b = 1, \alpha = 0,43, \lambda = 0, \gamma = 0,5$

A função de taxa de falha da distribuição BWM depende da função razão beta

incompleta e é dada por

$$h(x) = \frac{\alpha x^{\gamma-1}(\gamma + \lambda x) \exp(\lambda x)}{B(a, b)[1 - I_{1-\exp(-\alpha x^\gamma \exp(\lambda x))}(a, b)]} [1 - \exp(-\alpha x^\gamma \exp(\lambda x))]^{a-1} \exp(-b\alpha x^\gamma \exp(\lambda x)), \quad x > 0. \quad (65)$$

As Figuras 54 e 55 ilustram algumas das possíveis formas da função de densidade (64) e da função de taxa de falha (65), respectivamente, para determinados valores de parâmetros, incluindo alguns modelos particulares. Uma característica da distribuição BWM é que sua função de taxa de falha pode ser monotonicamente crescente ou decrescente, em forma de U ou unimodal dependendo unicamente dos valores dos parâmetros.

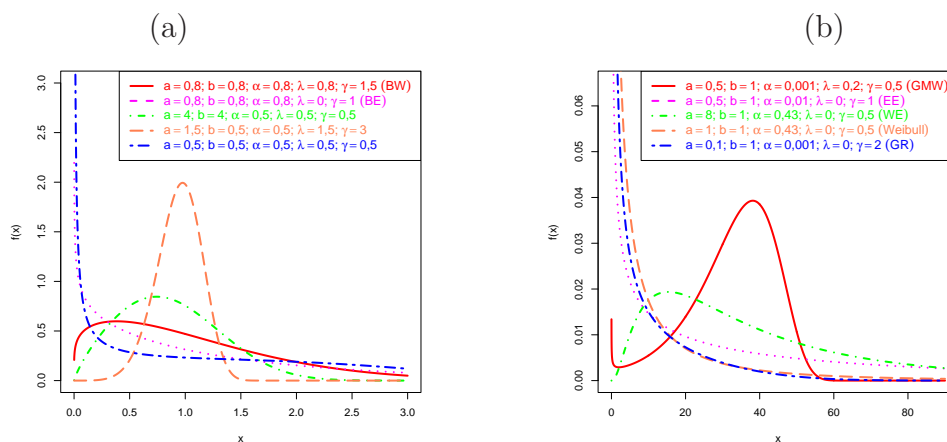


Figura 54 – Gráfico da função de densidade da distribuição BWM para alguns valores dos parâmetros

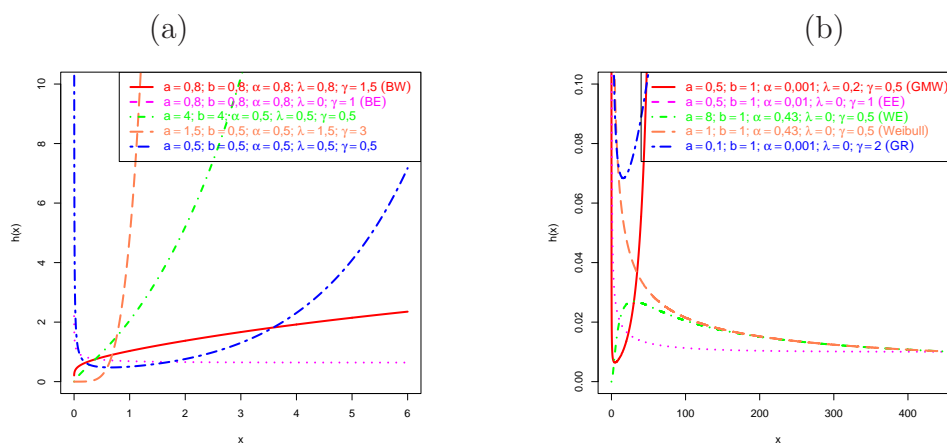


Figura 55 – Gráfico da função de taxa de falha da distribuição BWM para alguns valores dos parâmetros

8.2 Expansões para as funções de densidade, de distribuição acumulada e de sobrevivência

Expansões simples para a FDA da distribuição BWM depende se o parâmetro b (ou a) é um número real inteiro ou não-inteiro. A densidade dessa distribuição, dada em (64), é simples de trabalhar usando qualquer *software* estatístico ou matemático. Mas, essa densidade pode ser expressa como uma combinação linear ponderada infinita (ou finita) da função de densidade da distribuição WM, que é importante para obter algumas propriedades matemáticas da distribuição BWM diretamente das correspondentes propriedades da distribuição WM.

Usando a representação em série (assumindo $a > 0$ um número real e não-inteiro) tem-se que (GRADSHTEYN; RYZHIK, 2007)

$$(1+z)^{a-1} = \sum_{j=0}^{\infty} \frac{\Gamma(a)z^j}{\Gamma(a-j)j!}, \quad (66)$$

em que $\Gamma(\cdot)$ é a função gama. Assim, pode-se escrever que

$$\omega^{a-1} = (1 - 1 + \omega)^{a-1} = \{1 + [(-1)(1 - \omega)]\}^{a-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(a)(1 - \omega)^j}{\Gamma(a-j)j!}. \quad (67)$$

Então,

$$\begin{aligned} \int_0^x \omega^{a-1}(1-\omega)^{b-1}d\omega &= \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(a)}{\Gamma(a-j)j!} \int_0^x (1-\omega)^{b+j-1}d\omega \\ &= \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(a)}{\Gamma(a-j)(b+j)j!} [1 - (1-x)^{b+j}], \end{aligned} \quad (68)$$

a última integral é resolvida usando a seguinte mudança de variável $u = 1 - \omega$.

Substituindo a equação (68) na equação (63), obtem-se que

$$F(x) = \frac{1}{B(a,b)} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(a)}{\Gamma(a-j)(b+j)j!} \{1 - [1 - G_{\alpha,\gamma,\lambda}(x)]^{b+j}\}$$

e, então

$$F(x) = \sum_{j=0}^{\infty} \omega_j G_{\alpha(b+j),\gamma,\lambda}(x), \quad (69)$$

em que

$$\omega_j = \omega_j(a,b) = \frac{(-1)^j \Gamma(a)}{B(a,b)\Gamma(a-j)(b+j)j!}$$

é constante de maneira que $\sum_{j=0}^{\infty} \omega_j = 1$ e $G_{\alpha(b+j),\gamma,\lambda}(x)$ é a FDA da distribuição WM com parâmetro de escala $\alpha(b+j)$, parâmetro de forma γ e o parâmetro de aceleração λ . Se a é um inteiro, o indexador j na soma anterior pára em $a-1$. A expansão (69) é usada no restante deste capítulo para qualquer a . Para $\sum_{j=0}^{\infty} \omega_j = 1$, a função de sobrevivência para a distribuição BWM tem a seguinte expansão

$$\begin{aligned} S(x) &= 1 - F(x) = 1 - \sum_{j=0}^{\infty} \omega_j G_{\alpha(b+j),\gamma,\lambda}(x) = 1 - \sum_{j=0}^{\infty} \omega_j [1 - S_{\alpha(b+j),\gamma,\lambda}(x)] \\ &= 1 - \sum_{j=0}^{\infty} \omega_j + \sum_{j=0}^{\infty} \omega_j S_{\alpha(b+j),\gamma,\lambda}(x) = \sum_{j=0}^{\infty} \omega_j S_{\alpha(b+j),\gamma,\lambda}(x), \end{aligned}$$

em que $S_{\alpha(b+j),\gamma,\lambda}(x) = \exp\{-\alpha(b+j)x^\gamma \exp(\lambda x)\}$ é a função de sobrevivência da distribuição WM com parâmetros $\alpha(b+j)$, γ e λ . A função de densidade da distribuição BWM segue imediatamente da expressão (69) como

$$f(x) = \sum_{j=0}^{\infty} \omega_j g_{\alpha(b+j),\gamma,\lambda}(x). \quad (70)$$

Portanto, os momentos ordinários, centrais e fatoriais e a função geradora de momentos da distribuição BWM podem, a princípio, serem obtidos dessa combinação linear ponderada infinita (ou finita se a é um número inteiro) das quantidades correspondentes da distribuição WM.

A função de densidade da distribuição WM representa um caso particular de (70) quando $a = b = 1$. A densidade da distribuição GMW é também um caso particular de (70) quando unicamente $b = 1$. Além disso, enquanto a equação (59) não é analiticamente tratável no caso geral, a fórmula relacionada a distribuição BWM é manipulável (como é mostrado no restante deste capítulo), e com o uso de computadores com recursos modernos essa distribuição pode-se transformar em uma ferramenta útil para aplicações na área da estatística.

As equações (69) e (70) representam os principais resultados desta sub-seção. A vantagem dessas expressões é que podem ser usadas para determinar momentos sem quaisquer restrições ou condições para os cinco parâmetros.

8.3 Confiabilidade

Existem muitos trabalhos interessados na estimação da confiabilidade $R = Pr(X_2 < X_1)$ quando X_1 e X_2 são variáveis aleatórias independentes que pertencem a mesma família univariada de distribuições. A forma algébrica para R tem sido obtida para a maioria das distribuições padrões bem conhecidas. Mas, existem ainda muitas outras distribuições (incluindo generalizações das distribuições bem-conhecidas) para as quais a forma de R não foi obtida. Quando X_1 e X_2 são independentes e têm a mesma distribuição, sabe-se que a forma de R pode ser expressa como

$$R = \int_0^{\infty} f(x)F(x)dx. \quad (71)$$

Para mais detalhes leia Kotz; Lumelskii e Pensky (2003).

Substituindo as equações (64) e (69) na equação (71) obtem-se a forma de R para X_1 e X_2 independentes e com a mesma distribuição BMW

$$R = \frac{\alpha}{B(a, b)} \sum_{j=0}^{\infty} \omega_j(a, b) \int_0^{\infty} x^{\gamma-1} (\gamma + \lambda x) e^{\lambda x} (1 - u^{b+j}) (1 - u)^{a-1} u^b dx,$$

em que $u = \exp\{-\alpha x^\gamma e^{\lambda x}\}$ e $\omega_j(a, b)$ foi definido na sub-seção 8.2. Assim, $du = u \log(u)[(\gamma + \lambda x)/x]dx$ e então

$$R = \frac{1}{B(a, b)} \sum_{j=0}^{\infty} \omega_j(a, b) \int_0^1 (1 - u^{b+j}) (1 - u)^{a-1} u^{b-1} dx.$$

A última integral foi calculada usando o Maple e o resultado obtido é

$$R = \frac{1}{B(a, b)} \sum_{j=0}^{\infty} \omega_j(a, b) \left[B(a, b) - \frac{\Gamma(a)\Gamma(2b + j)}{\Gamma(2b + j + a)} \right]. \quad (72)$$

Note que a confiabilidade dada na equação (72) depende unicamente dos parâmetros a e b .

8.4 Fórmula geral para os momentos

Algumas dos principais aspectos e características de uma distribuição podem ser estudadas por meio dos momentos (por exemplo, tendência, dispersão, assimetria e curtose).

O s -ésimo momento ordinário da distribuição BWM, μ'_s , usando soma infinita, pode ser obtido da equação (70) da seguinte maneira

$$\mu'_s = E(X^s) = \int_0^\infty x^s \sum_{j=0}^{\infty} \omega_j g_{\alpha(b+j), \gamma, \lambda}(x) dx = \sum_{j=0}^{\infty} \omega_j \tau_s(j), \quad (73)$$

em que $\tau_s(j) = \int_0^\infty x^s g_{\alpha(b+j), \gamma, \lambda}(x) dx$ é o s -ésimo momento da distribuição WM com parâmetros $\alpha(b+j)$, γ e λ .

Diretamente da equação (73), os momentos das distribuições BWM podem ser obtidos utilizando o Maple para resolver a integral que fornece os momentos da distribuição WM. Representações gráficas da assimetria e da curtose quando $\alpha = 1, 2$, $\lambda = 4, 2$ e $\gamma = 3, 2$, como função do parâmetro a para alguns valores escolhidos do parâmetro b e como uma função do parâmetro b para alguns valores escolhidos do parâmetro a , são dadas nas Figuras 56 e 57, respectivamente. Essas figuras mostram que a assimetria e a curtose aumentam quando b diminui e quando a aumenta.

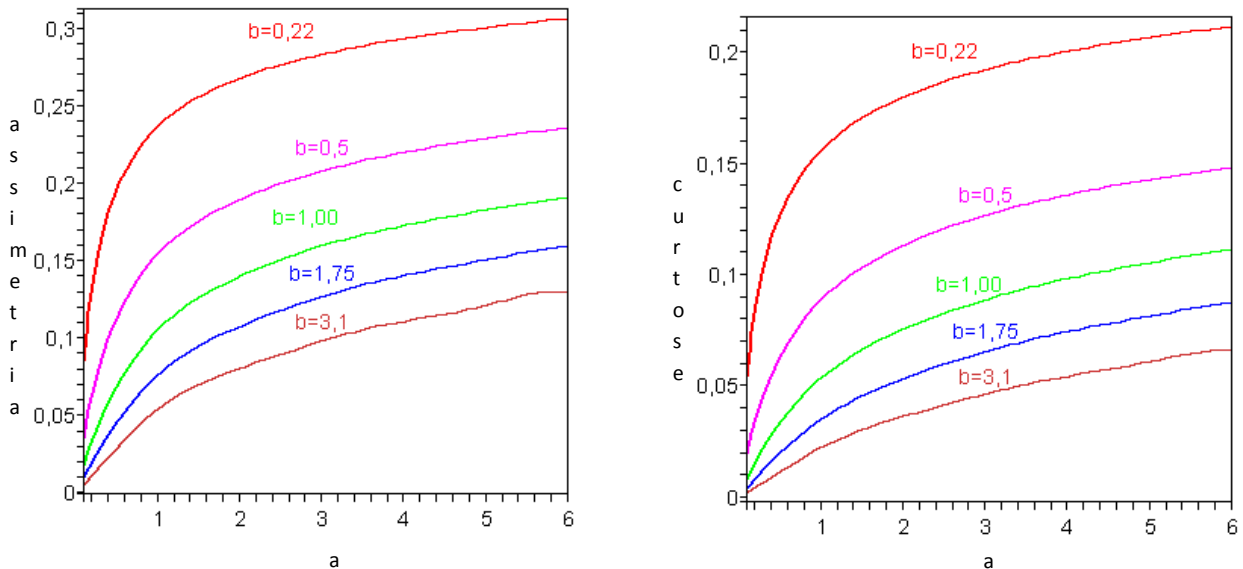


Figura 56 – Assimetria e curtose da distribuição BWM como função do parâmetro a , para alguns valores de b

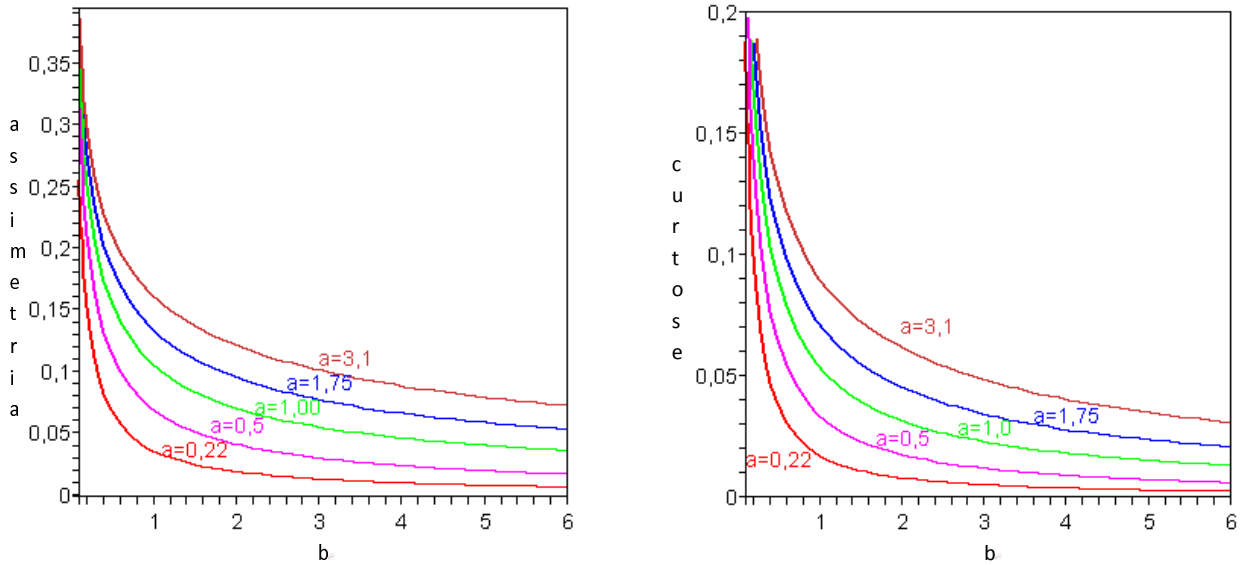


Figura 57 – Assimetria e curtose da distribuição BWM como função do parâmetro b , para alguns valores de a

8.5 Estimação de Máxima Verossimilhança

Nesta seção, apresenta-se como obter as estimativas de máxima verossimilhança dos parâmetros da distribuição BWM. Seja x_1, \dots, x_n uma amostra aleatória de tamanho n da distribuição $BWM(a, b, \alpha, \gamma, \lambda)$. Então, o logaritmo da função de verossimilhança para o vetor de parâmetros $\boldsymbol{\theta} = (a, b, \alpha, \gamma, \lambda)^T$ pode ser escrito como

$$\begin{aligned}
 l(\boldsymbol{\theta}) &= -n \log [B(a, b)] + \sum_{i=1}^n \left\{ [\log(\gamma + \lambda x_i)] + \log(v_i) - \log(x_i) \right\} \\
 &\quad + (a - 1) \sum_{i=1}^n \log \left\{ 1 - \exp(-v_i) \right\} - b \sum_{i=1}^n v_i,
 \end{aligned} \tag{74}$$

em que $v_i = \alpha x_i^\gamma \exp(\lambda x_i)$ é uma observação transformada.

Os componentes do vetor escore $U(\boldsymbol{\theta})$, definidos na seção 3.2.1, são dados por

$$\begin{aligned}
 U_a(\boldsymbol{\theta}) &= -n[\psi(a) + \psi(a + b)] + \sum_{i=1}^n \log [1 - \exp(-v_i)], \\
 U_b(\boldsymbol{\theta}) &= -n[\psi(b) + \psi(a + b)] - \sum_{i=1}^n v_i, \\
 U_\alpha(\boldsymbol{\theta}) &= n\alpha^{-1} + (a - 1) \sum_{i=1}^n \frac{\exp(-v_i) x_i^\gamma \exp(\lambda x_i)}{[1 - \exp(-v_i)]} - b \sum_{i=1}^n x_i^\gamma \exp(\lambda x_i),
 \end{aligned}$$

$$\begin{aligned}
U_\gamma(\boldsymbol{\theta}) &= \sum_{i=1}^n \left[\frac{1}{\gamma + \lambda x_i} + \log(x_i) \right] + (a-1) \sum_{i=1}^n \frac{v_i \exp(-v_i) \log(x_i)}{[1 - \exp(-v_i)]} - b \sum_{i=1}^n v_i \log(x_i), \\
U_\lambda(\boldsymbol{\theta}) &= \sum_{i=1}^n \left(\frac{x_i}{\gamma + \lambda x_i} + x_i \right) + (a-1) \sum_{i=1}^n \frac{v_i x_i \exp(-v_i)}{[1 - \exp(-v_i)]} - b \sum_{i=1}^n v_i x_i,
\end{aligned}$$

em que $\psi(\cdot)$ é a função digama. Na obtenção de $U_a(\boldsymbol{\theta})$ e $U_b(\boldsymbol{\theta})$ foi utilizado que

$$\frac{\partial \Gamma(a+b)}{\partial a} = \int_0^\infty \frac{\partial t^a}{\partial a} t^{b-1} e^{-t} dt = \int_0^\infty \log(t) t^{a+b-1} e^{-t} dt = \Psi(a+b) \Gamma(a+b)$$

em que a última expressão foi obtida pelo Maple.

Intervalos de confiança e teste de hipóteses para os parâmetros da distribuição podem ser obtidos como descrito na seção 3.2.1. Assim, sob certas condições de regularidade da função de verossimilhança, a distribuição assintótica aproximada de $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ é $N_5(0, \ddot{L}(\boldsymbol{\theta})^{-1})$, em que $\ddot{L}(\boldsymbol{\theta})$ é a matriz de informação observada, obtida de

$$\ddot{L}(\boldsymbol{\theta}) = - \begin{pmatrix} \mathbf{L}_{aa} & \mathbf{L}_{ab} & \mathbf{L}_{a\alpha} & \mathbf{L}_{a\gamma} & \mathbf{L}_{a\lambda} \\ \cdot & \mathbf{L}_{bb} & \mathbf{L}_{b\alpha} & \mathbf{L}_{b\gamma} & \mathbf{L}_{b\lambda} \\ \cdot & \cdot & \mathbf{L}_{\alpha\alpha} & \mathbf{L}_{\alpha\gamma} & \mathbf{L}_{\alpha\lambda} \\ \cdot & \cdot & \cdot & \mathbf{L}_{\gamma\gamma} & \mathbf{L}_{\gamma\lambda} \\ \cdot & \cdot & \cdot & \cdot & \mathbf{L}_{\lambda\lambda} \end{pmatrix}$$

com sub-matrizes dadas no apêndice M.

Para testar a qualidade de ajuste da distribuição BWM pode ser usado o teste da razão de verossimilhança (CASELLA; BERGER, 2002) para verificar se o ajuste usando a distribuição BWM é estatisticamente “superior” ao ajuste usando quaisquer das distribuições mencionadas na Figura 52 para um determinado conjunto de dados. Também podem ser utilizadas as estatísticas de Wald e escore de Rao (SEN; SINGER, 1993). Essas três estatísticas de teste são equivalentes assintoticamente.

Portanto, para testar hipóteses do tipo $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ versus $H : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ pode-se usar qualquer uma das três estatísticas citadas anteriormente. Por exemplo, o teste $H_0 : a = b = 1$ versus $H : H_0$ não é verdade é equivalente a comparar a distribuição BWM com a distribuição WM e a estatística de teste da razão de verossimilhança é dada por

$$TRV = 2\{\ell(\hat{a}, \hat{b}, \hat{\alpha}, \hat{\gamma}, \hat{\lambda}) - \ell(1, 1, \tilde{\alpha}, \tilde{\gamma}, \tilde{\lambda})\},$$

em que \hat{a} , \hat{b} , $\hat{\alpha}$, $\hat{\gamma}$ and $\hat{\lambda}$ são as estimativas de máxima verossimilhança sob H e $\tilde{\alpha}$, $\tilde{\gamma}$ e $\tilde{\lambda}$ são as estimativas sob H_0 .

8.6 Análise dos dados de Aarset

Para ilustrar a aplicação da distribuição BWM foi utilizado um conjunto analisado em Aarset (1987). Os dados se referem a 50 dispositivos e também foram analisados por Mudholkar e Srivastava (1993), Mudholkar, Srivastava e Kollia (1996) e Wang (2000).

A curva TTT, definida na seção 2.4, para o conjunto de dados de Aarset está na Figura 58 e indica uma função de taxa de falha na forma de U . Assim, para analisar esse conjunto de dados a distribuição BWM pode ser usada.

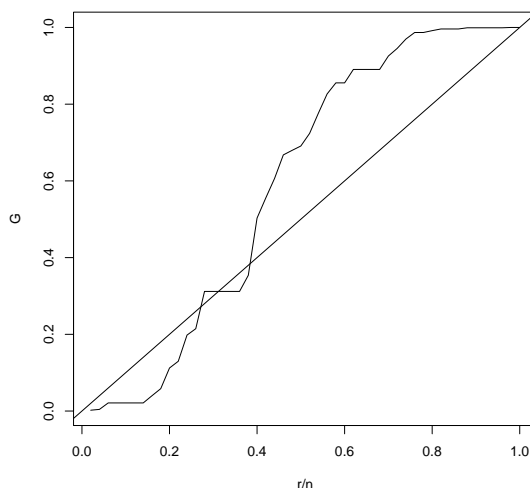


Figura 58 – Curva TTT para os dados de Aarset

Além disso, as funções de sobrevivência e de densidade estimadas para os diferentes modelos ajustados nos dados de Aarset são representadas graficamente nas Figuras 59 e 60, respectivamente. Dessas figuras, pode ser observado que para as diferentes distribuições ajustadas aos dados de Aarset a distribuição BWM apresenta um melhor ajuste se comparado aos outros modelos probabilísticos alternativos.

Para obter uma breve avaliação da qualidade de ajuste dos dados de Aarset podem ser realizados testes da razão de verossimilhança descritos na seção (8.5) para testar

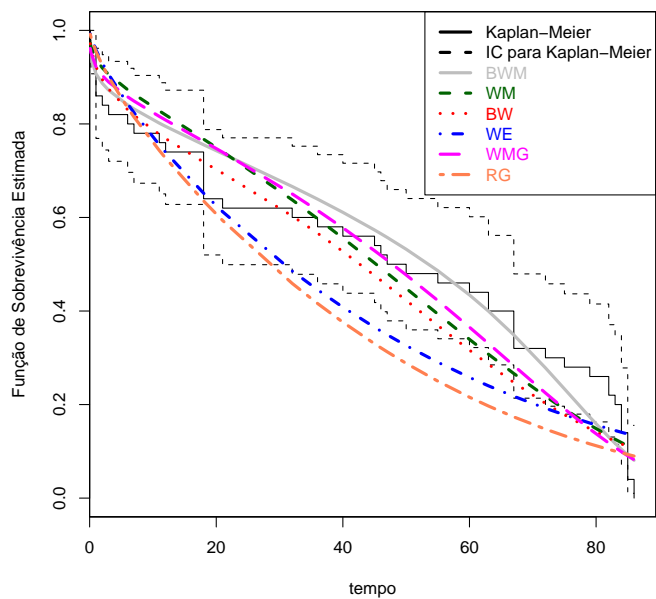


Figura 59 – Função de sobrevivência estimada para alguns modelos ajustados e a função de sobrevivência estimada pelo método de Kaplan-Meier dos dados de Aarset

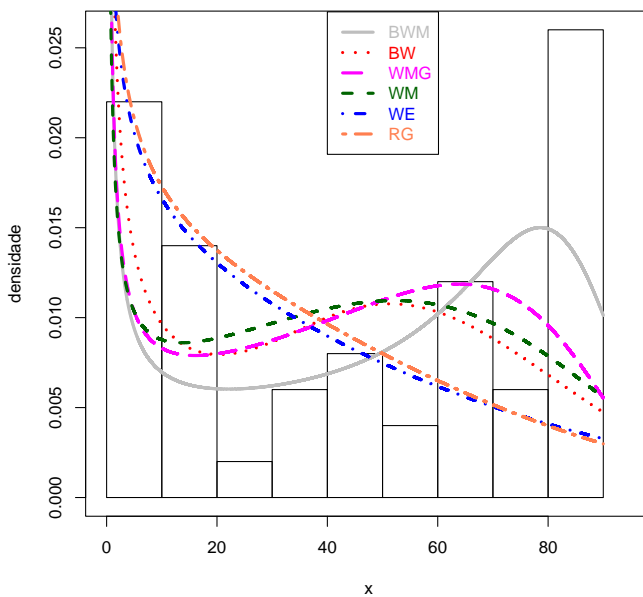


Figura 60 – Função de densidade estimada de alguns modelos para os dados de Aarset

alguns sub-modelos da distribuição BWM apresentados na seção 8.1. Por exemplo, a estatística do teste da razão de verossimilhança foi obtida para testar as hipóteses $H_0:a=b=1$ versus $H_1:H_0$ não é verdade, isto é, compara o modelo WM com o modelo BWM. Nesse contexto, a referida estatística é $TRV = 2\{-220,817 - (-227,982)\} = 14,3293$ (p-valor $< 0,01$) que fornece indicações favoráveis ao modelo BWM. As estimativas de máxima verossimilhança para os parâmetros dos modelos foram obtidas na linguagem de programação matricial Ox usando a sub-rotina MaxBFGS.

Essas estimativas com os correspondentes erros padrão, entre parênteses, são dados na Tabela 22 para diversos modelos ajustados, além de fornecer os valores da medida AIC, definida na sub-seção (3.3.1.1). Esses resultados indicam que o modelo BWM apresenta o menor valor para o AIC entre todos os modelos ajustados, e portanto pode ser escolhido como o melhor modelo para analisar os dados de Aarset.

Tabela 22 – Estimativas de máxima verossimilhança para alguns sub-modelos para os dados de Aarset, os correspondentes erros padrão dados entre parênteses e os valores do critério AIC

Modelo	Parâmetro					AIC
	a	b	α	λ	γ	
BWM	0,1975 (0,0462)	0,1647 (0,0830)	0,0002 (6,6931e-005)	0,0541 (0,0157)	1,3771 (0,3387)	451,635
BW	0,18356 (0,0509)	0,0748 (0,0353)	0,0007 (0,0004)	0 0	2,3615 (0,1715)	463,964
GMW	0,2975 (0,0613)	1	0,0002 (0,0001)	0,0529 (0,0138)	0,9942 (0,2396)	455,765
MW	1	1	0,0624 (0,0267)	0,0233 (0,0048)	0,3548 (0,1127)	460,310
EW	0,4668 (0,0889)	1	0,0011 (0,0010)	0	1,5936 (0,1858)	480,447
GR	0,3643 (0,0624)	1	0,0002 (4,8738e-005)	0	2	475,904

9 CONSIDERAÇÕES FINAIS

9.1 Conclusões

Neste trabalho foram propostos dois modelos de regressão, log-Weibull estendida e log-Burr XII, para dados considerando censura à direita e com mecanismo aleatório. As vantagens desses modelos é que são para dados de sobrevivência com função de taxa de falha não-monótona, respectivamente, em forma de U e unimodal e apresentam função de sobrevivência e função de taxa de falha analiticamente explícitas. Assim, tais modelos de regressão podem ser úteis em análise de sobrevivência. Para estimar os parâmetros desses modelos foi utilizado o método de máxima verossimilhança e realizados testes de hipóteses utilizando a distribuição assintótica dos estimadores de máxima verossimilhança que exigem certas condições de regularidade para a função de verossimilhança. Diante disso, foram também propostos uma abordagem bayesiana e o método de jackknife para estimação dos parâmetros.

Além disso, para cada um dos modelos foram derivadas as curvaturas normais para se obterem as medidas diagnósticos, baseadas em três esquemas de perturbação, de influência local e de influência local total, bem como foram dadas medidas diagnóstico baseadas em deleção de casos (influência global). Um outro aspecto abordado foi a análise de resíduos, sendo propostos resíduos para ambos os modelos cuja distribuição empírica sugere uma concordância razoável com a distribuição normal padrão.

As metodologias propostas foram ilustradas por meio de dados reais, mostrando que os métodos de estimação apresentam resultados similares e que as análises de resíduos e sensibilidade auxiliam na escolha do modelo adequado.

Outro enfoque abordado foi modelo de regressão com efeito aleatório em que foi proposto um método de estimação baseado no algoritmo EM por Monte Carlo. Constatou-se com base em um conjunto de dados simulados que o desempenho desse algoritmo proposto não é adequado.

Finalmente, foi definido um novo modelo probabilístico com cinco parâmetros, denominado de distribuição beta Weibull modificada, que estende várias distribuições amplamente abordadas na literatura de análise de sobrevivência. A principal vantagem do novo

modelo é a flexibilidade em acomodar diferentes formas da função de taxa de falha. Também foram explicitadas expressões para calcular o s -ésimo momento usando soma infinita. Por meio de exemplo com dados reais mostrou-se a utilidade desse novo modelo probabilístico.

9.2 Pesquisas futuras

Dando continuidade a esta pesquisa, pretende-se melhorar o método de estimação baseado no algoritmo EM por Monte Carlo para o modelo de regressão com efeito aleatório substituindo o algoritmo de Metropolis-Hasting por outro método de amostragem por rejeição. Conseguindo melhorar esse método de estimação, pode-se propor medidas de influência baseadas na função Q-afastamento (ZHU; LEE, 2001) para estudar medidas de influência local para os modelos de regressão com efeito aleatório.

Outra pesquisa será propor um método de estimação para os parâmetros da distribuição beta Weibull modificada para dados com observações censuradas, além de apresentar um modelo de regressão para a distribuição beta Weibull modificada. Também podem ser introduzida uma análise de diagnóstico (influência local e resíduos) para esse novo modelo de regressão.

REFERÊNCIAS

- AARSET, M.V. How to identify bathtub hazard rate. **IEEE Transactions on Reliability**, New York, v. 36, n. 1, p. 106-108, 1987.
- ABRAMOWITZ, M.; SEGUN, I. A. **Handbook of mathematical functions**. New York: Dover Publications, 1972. 1046 p.
- ATKINSON, A. C. **Plots, transformations and regression: an introduction to graphical methods of diagnostics regression analysis**. 2nd ed. Oxford: Clarendon Press, 1987. 282 p.
- BARRETO, M.L.; SANTOS, L.M.P.; ASSIS, A.M.O.; ARAÚJO, M.P.N.; FARENZENA, G.G.; SANTOS, P.A.B.; FIACCONE, R.L. Effect of vitamin A supplementation on diarrhoea and acute lower-respiratory-tract infections in young children in Brazil. **Lancet**, London, v. 344, p. 228-231, 1994.
- BENNETT, S. Log-logistic regression models for survival data. **Applied Statistics**, London, v. 32, n. 2, p. 165-171, 1983.
- BOOTH, J.G.; HOBERT, J.H. Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. **Journal of the Royal Statistical Society: Series B, Statistical Methodology**, Oxford, v. 61, n. 2, p. 265-285, 1999.
- BROWN, B.W.; FLOYD, M. S.; LEVY, L.B. The log F: a distribution for all seasons. **Computational Statistics**, Berlin, v. 17, n. 1, p. 47-58, 2002.
- CANCHO, V.; BOLFARINE, H.; ACHCAR, J. A. A Bayesian analysis for the exponentiated-Weibull distribution. **Journal of Applied Statistical Science**, New York, v. 8, n. 4, p. 227-242, 1999.
- CARRASCO, J. M. F.; ORTEGA, E. M. M.; CORDEIRO, G. M. (2008). A generalized modified Weibull distribution for lifetime modeling. **Computational Statistics and Data Analysis**: doi 10.1016/j.CSDA.2008.08.023.
- CASELLA, G.; BERGER, R. L. **Statistical inference**. 2nd ed. Pacific Grove: Thomson Learning, 2002. 660 p.

- CASELLA, G.; GEORGE, E. I. Explaining the gibbs sampler. **The American Statistician**, Washington, v. 46, n. 3, p. 167-174, 1992.
- CHAN, J.S.K.; KUK, A.Y.C. Maximum likelihood estimation for probit-linear mixed models with correlated random effects. **Biometrics**, Washington, v. 53, p. 86-97, 1997.
- CHAN, K. S.; LEDOLTER, J. Monte Carlo EM Estimation for Time Series Models Involving Counts. **Journal of the American Statistical Association**, Boston, v. 90, n. 429, p. 242-252, 1995.
- CHATTERJEE, S.; HADI, A. S. **Sensitivity analysis in linear regression**. New York: Wiley, 1994. 315 p.
- CHEN, Z.A. A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. **Statistics and Probability Letters**, Amsterdam, v. 49, p. 155-161, 2000.
- CHEN, J; ZHANG, D.; DAVIDIAN, M. A Monte Carlo EM algorithm for generalized linear mixed models with flexible random effects distribution. **Biostatistics**, v. 3, n. 3, p. 347-360, 2002.
- CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hastings algorithm. **The American Statistician**, Washington, v. 49, n. 4, p. 327-335, 1995.
- CHRISTENSEN, R.; PEARSON, L.; JOHNSON, W. Case-deletion diagnostics for mixed models. **Technometrics**, Alexandria, v. 34, p. 38-45, 1992.
- COLLET, D. **Modeling survival data in medical research**. 2nd ed. London: Chapman and Hall, 2003. p.
- COLOSIMO, E. A.; GIOLO, S. R. **Análise de sobrevivência aplicada**. São Paulo: Edgard Blücher, 2006. 392 p.
- COOK, R.D. Assesment of local influence (with discussion). **Journal of the Royal Statistical Society: Series B, Statistical Methodology**, Oxford, v. 48, n. 2, p. 133-169, 1986.

———. Detection of influential observations in linear regression. **Technometrics**, Alexandria, v. 19, p. 15-18, 1977.

COOK, R. D.; WEISBERG, S. **Residuals and influence in regression**. New York: Chapman and Hill, 1982. 230 p.

COOK, R. D.; PEÑA, D.; WEISBERG, S. The likelihood displacement: a unifying principle for influence. **Communications in Statistics: Part Theory and Methods**, New York, v. 17, n. 3, p. 623-640, 1988.

COX, D.R. Regression models and life-tables (with discussion). **Journal of the Royal Statistical Society: Series B, Methodological**, London, v. 34, n. 2, p. 187-220, 1972.

COX, D. R.; HINKLEY, D. V. **Theoretical Statistics**. Chapman and Hall: London, 1974.

COX, D.R.; OAKES, D. **Analysis of survival data**. London: Chapman and Hall, 1984. 201 p.

COX, D.R.; SNELL, E.J. A general definition of residuals. **Journal of the Royal Statistical Society: Series B, Methodological**, London, v. 30, n. 2, p. 248-275, 1968.

DAVISON, A.C.; TSAI, C. L. Regression model diagnostics. **International Statistical Review**, Hague, v. 60, p. 337-355, 1992.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm (with discussion). **Journal of the Royal Statistical Society: Series B, Methodological**, London, v. 39, n. 1, p. 1-38, 1977.

DOORNIK, J. **Ox 5: Object-oriented matrix programming language**. 5th ed. London : Timberlake Consultants, 2007. 461p.

ESCOBAR, L.A.; MEEKER, W.Q. Assessing influence in regression analysis with censored data. **Biometrics**, Washington, v. 48, n.2, p. 507-528, 1992.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal Distribution and its applications. **Communication in Statistics - Theory and Methods**, New York, v. 31, 497-512, 2002.

- FACHINI, J. B.; ORTEGA, E. M. M.; LOUZADA-NETO, F. Influence diagnostics for poly-hazard models in the presence of covariates. **Statistical Methods and Applications**, New York, v. 17, p. 413-433, 2008.
- FAMOYE, F.; LEE, C.; OLUMALADE, O. The Beta-Weibull Distribution. **Journal of Statistical Theory and Applications**, Holland, v. 4, p. 121-136, 2005.
- FLEMING, T.R.; HARRINGTON, D.P. **Counting process and survival analysis**. New York: Wiley, 1991. 429 p.
- GELFAND, A.E.; SMITH, A.F.M. Sampling based approaches to calculating marginal densities. **Journal of the American Statistical Association**, Boston, v. 85 , p. 398-409, 1990.
- GELMAN, A.; RUBIN, D.B. Inference from iterative simulation using multiple sequences (with discussion). **Statistical Science**, Hayward, v. 7, n. 4, p. 457-472, 1992.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions and the Bayesian restoration of images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v. 6, p. 721-741, 1984.
- GRADSHTEYN, I. S.; RYZHIK, I. M. **Table of integrals, series and products**. 7th ed. New York: Academic Press, 2007. 1171 p.
- GUPTA, R. D.; KUNDU, D. Exponentiated exponential distribution: an alternative to gamma and Weibull distributions. **Biometrical Journal**, Berlin, v. 43, p. 117-130, 2001.
- GUPTA, A. K.; NADARAJAH, S. **Handbook of beta distribution and its applications**. New York: Marcel Dekker, 2004.
- HAUPT, E.; SCHABE, H. A new model for a lifetime distribution with bathtub shaped failure rate. **Microelectronic and Reliability**, Kidlington, v. 32, p. 633-639, 1992.
- HJORTH, U. A realibility distributions with increasing, decreasing, constant and bathtub failure rates. **Technometrics**, Alexandria, v. 22, n. 1, p. 99-107, 1980.
- HOSMER Jr., D. W.; LEMESHOW, S. **Applied survival analysis: regression modeling of time to event data**. New York: John Wiley and Sons, 1999. 386 p.

- IBRAHIM, J. G.; CHEN, M. H.; SINHA, D. **Bayesian survival analysis**. New York: Springer, 2001. 479 p.
- JONES, M. C. Family of distributions arising from distribution of order statistics. **Test**, Madri, v. 13, p. 1-43, 2004.
- KALBFLEISCH, J.D.; PRENTICE, R.L. **The Statistical analysis of Failure Time Data**. 2nd ed. New York: John Wiley, 2002. 439 p.
- KEIDING, N.; ANDERSEN, P.K.; KLEIN, J.P. The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates. **Statistics in Medicine**, New York, v. 16, p. 215-224, 1997.
- KLEIN, J. P.; MOESCHBERGER, M. L. **Survival analysis: techniques for censored and truncated data**. New York : Springer. 1997.
- KOTZ, S.; LUMELSKII, Y.; PENSKY, M. **The stress-strength model and its generalizations: theory and applications**. New Jersey: World Scientific Publishing. 2003. 253 p.
- KUNDU, D.; RAKAB, M. Z. Generalized Rayleigh distribution: different methods of estimation. **Computational Statistics and Data Analysis**, Amsterdam, v. 49, p. 187-200, 2005.
- KUO, L; PENG, F. **A mixture-model approach to the analysis of survival data**. Storrs: Department of Statistics, University of Connecticut, 1995, p. 95-31. Statistics Technical Reports.
- LAI, C. D.; XIE, M.; MURTHY, D. N. P. A modified Weibull distribution. **IEEE Transactions on Reliability**, New York, v. 52, n. 1, p. 33-37, 2003.
- LAPLANTE-ALBERT, K. A. **Habitat-dependent mortality risk in lacustrine fish**. 2008. 123 p. M. Sc. Science . Université du Québec à Trois-Rivières, Québec, 2008.
- LAWLESS, J. F. **Statistical models and methods for lifetime data**. 2nd ed. New York: Wiley, 2003. 630 p.

- LESAFFRE, E.; VERBEKE, G. Local influence in linear mixed models. **Biometrics**, Washington, v. 54, n. 2, p. 570-582, 1998.
- LEIVA, V.; BARROS, M.; PAULA, G.A.; GALEA, M. Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. **Computational Statistics and Data Analysis**, Amsterdam, v. 51, n. 12, p.5694-5707, 2007.
- LEVINE, R. A.; CASELLA G. Implementations of the Monte Carlo EM Algorithm. **Journal of Computational and Graphical Statistics**, Alexandria, v. 10, n. 3, p. 422- 439, 2001.
- LI, H. **Survival analysis methods in genetic epidemiology**. Upenn Biostatistics Working Papers. Working Paper 7. Feb. 2006. <Disponível em: <http://biostats.bepress.com/upennbiostat/papers/art7>>. Acesso em: 18 nov. 2008.
- LIU, L.; YU, Z. A likelihood reformulation method in non-normal random effects models. **Statistics in Medicine**, New York, v. 27, p. 3105-3124, 2008.
- LIANG, H.; ZOU, G. Improved AIC selection strategy for survival analysis. **Computational Statistics and Data Analysis**, New York, v. 52, p. 2538-2548, 2003.
- LITTLE, R. J.; RUBIN, D. B. **Statistical analysis with missing data**. 2nd ed., New York: Wiley, 2002. 408 p.
- LIU, Q.; PIERCE, D. A. A note on Gauss-Hermite quadrature. **Biometrika**, London, v.81, p. 624-629, 1994.
- LOUIS, T. A. Finding the observed information matrix when using the EM algorithm. **Journal of the Royal Statistical Society: Series B, Statistical Methodology**, Oxford, v. 44, n. 2, p. 226-233, 1982.
- MANLY, B.F.J. **Randomization, bootstrap and Monte Carlo methods in biology**. 2nd ed., New York : Chapman and Hall, 1997. 399 p.
- McCULLAGH, P.; NELDER, J. A. **Generalized linear models**. 2nd ed., London: Chapman and Hall, 1989. 511 p.

- McCULLOCH, C. E. Maximum likelihood variance components estimation for binary data. **Journal of the American Statistical Association**, Boston, v. 89, n. 425, p. 330-335, 1994.
- McCULLOCH, C. E. Maximum likelihood algorithms for generalized linear mixed models. **Journal of the American Statistical Association**, Boston, v. 92, n. 437, p. 162-170, 1997.
- MORRIS, C.; CHRISTIANSEN, C. Fitting Weibull duration model with random effects. **Lifetime Data Analysis**, Hingham, v. 1, p. 347-359, 1995.
- MUDHOLKAR, G. S.; SRIVASTAVA, D. K. Exponentiated Weibull family for analyzing bathtub failure-real data. **IEEE Transaction on Reliability**, New York, v. 42, p. 299-302, 1993.
- MUDHOLKAR, G.S; SRIVASTAVA, D.K.; KOLLIA, G.D. A generalization of the Weibull distribution with application to the analysis of survival data. **Journal of the American Statistical Association**, Boston, v. 91, n. 436, p. 1575-1583, 1996.
- NADARAJAH, S. On the moments of the modified Weibull distribution. **Reliability Engineering and System Safety**, Barking, v. 90, p. 114-117, 2005.
- NADARAJAH, S.; GUPTA, A. K. The beta Fréchet distribution. **Far East Journal of Theoretical Statistics**, Allahabad, v.14, p. 15-24, 2004.
- NADARAJAH, S.; GUPTA, A. K. On the moments of the exponentiated Weibull distribution. **Communications in Statistics: Part Theory and Methods**, New York, v. 34, 253-256, 2005.
- NADARAJAH, S.; KOTZ, S. The beta Gumbel distribution. **Mathematical Problems in Engineering**, New York, v. 2004, p. 323-332, 2004.
- NADARAJAH, S.; KOTZ, S. On some recent modifications of Weibull distribution. **IEEE Transaction on Reliability**, New York, v. 54, n.4, p. 689-697, 2005.
- NADARAJAH, S.; KOTZ, S. The beta exponential distribution. **Reliability Engineering and System Safety**, Barking, v. 91, p. 689-697, 2006

NELSON, K. P.; LIPSITZ, S. R.; FITZMAURICE, G. M.; IBRAHIM, J.; PARZEN, M.; STRAWDERMAN, R. Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. **Journal Computational and Graphical Statistics**, Alexandria, v. 15, p. 39 - 57, 2006.

NELSON, W. **Applied life data analysis**. New York: Wiley, 1982. 634 p.

NOCEDAL, J.; WRIGHT, S. J. **Numerical optimization**. 2nd ed. New York: Springer, 2006. 664 p.

ORTEGA, E. M. M.; BOLFARINE, H.; PAULA, G. A. Influence diagnostics in generalized log-gamma regression models. **Computational Statistics and Data Analysis**, New York, v. 42, p. 165-186, 2003.

ORTEGA, E. M. M.; CANCHO, V. G.; BOLFARINE, H. Influence diagnostics in exponentiated-Weibull regression models with censored data. **Statistics and Operation Reserch Transactions**, Catalunya, v. 30, n. 2, p. 171-192, 2006.

ORTEGA, E. M. M. ; PAULA, G. A. ; BOLFARINE, H. Deviance residuals in generalized log-gamma regression models with censored observations. **Journal of Statistical Computation and Simulation**, New York, v. 78, p. 747-764, 2008.

PETTITT, A. N.; BIN DAUD, I. Case-weight measures of influence for proportional hazards regression. **Applied Statistics**, New York, v. 38, p. 51-67, 1989.

PHAM, H.; LAI, C. D. On Recent Generalizations of the Weibull Distribution. **IEEE Transactions on Reliability**, New York, v. 56, p. 454-458, 2007.

PRENTICE, R. L. Exponential survival with censoring and explanatory variables. **Biometrika**, London, v. 60, p. 279-288, 1973.

PRESS, W. H.; TEULOSKY, S.A.; VETERLING, W. T.; FLANNERY, B. P. **Numerical recipes in C: the art of scientific computing**. 2nd ed. London: Prentice Hall, 1992. 630 p.

RAJARSHI, S.; RAJARSHI, M. B. Bathtub distribution: a review. **Communications in Statistics: Part A, Theory and Methods**, New York, v. 17, p. 2597-2621, 1988.

- RODRIGUEZ, R.N. A Guide to the Burr Type XII distributions. **Biometrika**, London, v. 64, n.1, p. 129-134, 1977.
- SEARLE, S. R.; CASELLA, G.; McCULLOCH, C. E. **Variance components**. New York: John Wiley e Sons, 1992. 501p.
- SEN, P.K.; SINGER, J.M. **Large sample methods in statistics: An introduction with applications**. New York: Chapman and Hall, 1993. 382 p.
- SAS Institute. **SAS/STAT User's Guide: Version 9**. Cary: SAS Institute. 2004. 5121 p.
- TABLEMAN, M.; KIM, J. S. **Survival analysis using S: analysis of time-to-data**. New York: Chapman and Hall, 2004. 260 p.
- TANNER, M. A. **Tools for statistical inference: observed data and augmentation methods**. 3rd ed. New York: Springer-Verlag, 1991. 111 p.
- TANNER, M. A. **Tools for statistical inference: methods for the exploration of posteriori distributions and likelihood functions**. 3rd ed. New York: Springer-Verlag, 1996. 221 p.
- THERNEAU, T.M.; GRAMBSCHE, P.M.; FLEMING, T.R. Martingale-based residuals for survival models. **Biometrika**, London, v. 77, p. 147-60, 1990.
- VAIDA, F.; FITZGERALD, A. P.; DeGRUTTOLA, V. Efficient hybrid EM for linear and nonlinear mixed effects models with censored response. **Computational Statistics and Data Analysis**, Amsterdam, v. 51 , n. 12, p. 5718-5730, Aug. 2007.
- VALENÇA, D.M. **Teste de homogeneidade e estimação para dados de sobrevivência agrupados e com erros de medidas**. 2003. 151p. Tese (Doutorado em Estatística) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2003.
- WANG, F. K. A new model with bathtub-shaped failure rate using an additive Burr XII distribution. **Reliability Engineering and System Safety**, Barking, v. 70, p. 305-312, 2000.

WEI, B.; HU, Y.Q.; FUNG, W.K. Generalized leverage and its applications. **Scandinavian Journal of Statistics**, Oxford, v. 25, p. 25-37, 1998.

WEI, G. C. G.; TANNER, M. A. Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithms **Journal of the American Statistical Association**, Boston, v. 85, p. 699- 704, Sept. 1990.

WEISSFELD, L. A. Influence diagnostics for the proportional hazards models. **Statistics and Probability Letters**, Amsterdam, v. 10, p. 411-417, 1990.

WEISSFELD, L. A.; SCHNEIDER, H. Influence diagnostics for the Weibull model fit to censored data. **Statistics and Probability Letters**, Amsterdam, v. 9, p.67-73, 1990.

WEISBERG, S. **Applied linear regression**. 2nd ed. New York: Wiley, 1985. 324 p.

WU, C. F. J. On the convergence properties of the EM algorithm. **Annals of Statistics**, Hayward, v. 11, p. 95-103, 1983.

XIE, M.; LAI, C.D. Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function. **Reliability Engineering and System Safety**, Barking, v. 52, p. 87-93, 1996.

XIE, M.; TANG, Y.; GOH, T.N. A modified Weibull extension with bathtub failure rate function. **Reliability Engineering and System Safety**, Barking, v. 76, p. 279-285, 2002.

XIE, F.; WEI, B. Diagnostics analysis for log-Birnbaum-Saunders regression models. **Computational Statistics and Data Analysis**, Amsterdam, v. 51, n. 9, p. 4692-4706, 2007.

ZIMMER, W.J.; KEATS, J.B.; WANG F.K. The Burr XII distribution in reliability analysis. **Journal of Quality Technology**, Milwaukee, v. 30, p. 389-394, 1998.

ZHU, H.; LEE, S. Local influence for incomplete-data models. **Journal of the Royal Statistical Society: Series B, Statistical Methodology**, Oxford, v. 63, n. 1, p. 111-126, 2001.

APÊNDICES

Apêndice A - Elementos da matriz de informação observada para o modelo de regressão log-Weibull estendida

$$\mathbf{L}_{\lambda\lambda} = -\frac{r}{\lambda^2} \quad \mathbf{L}_{\lambda\sigma} = \frac{1}{\sigma} \sum_{i=1}^n z_i h_i \exp(\mathbf{x}_i^T \boldsymbol{\beta})$$

$$\mathbf{L}_{\lambda\beta_j} = \sum_{i=1}^n x_{ij} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) + \frac{1}{\sigma} \sum_{i=1}^n x_{ij} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) h_i [1 - \sigma \exp(-z_i)]$$

$$\begin{aligned} \mathbf{L}_{\sigma\sigma} &= \frac{r}{\sigma^2} + \frac{1}{\sigma^2} \sum_{i \in F} \left\{ z_i [2 + \exp(z_i)(z_i + 2)] \right\} + \frac{\lambda}{\sigma^2} \\ &\quad \sum_{i=1}^n z_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}) h_i [-2 - z_i(1 + \exp(z_i))] \end{aligned}$$

$$\begin{aligned} \mathbf{L}_{\sigma\beta_j} &= \frac{1}{\sigma^2} \sum_{i \in F} x_{ij} [1 + \exp(z_i)(1 + z_i)] + \frac{\lambda}{\sigma^2} \\ &\quad \sum_{i=1}^n x_{ij} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) h_i [-1 + z_i(\sigma - 1 - \exp(z_i))] \end{aligned}$$

$$\begin{aligned} \mathbf{L}_{\beta_j\beta_s} &= \frac{1}{\sigma^2} \sum_{i \in F} x_{ij} x_{is} \exp(z_i) + \lambda \sum_{i=1}^n x_{ij} x_{is} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \\ &\quad + \frac{\lambda}{\sigma^2} \sum_{i=1}^n x_{ij} x_{is} \exp(\mathbf{x}_i^T \boldsymbol{\beta}) h_i [2\sigma - 1 - \sigma^2 \exp(-z_i) - \exp(z_i)] \end{aligned}$$

em que $j, s = 1, 2, \dots, p$, $h_i = \exp(z_i + \exp(z_i))$ e $z_i = \frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}$

Apêndice B: Demonstração do teorema 1

Para a distribuição log-Burr XII (15), a função geradora de momentos é dada pela resolução da seguinte equação

$$M_Y(T) = E(\exp(ty)) = \int_{-\infty}^{\infty} \exp(ty) kc \left[1 + \left(\frac{\exp(y)}{s} \right)^c \right]^{-k-1} \left(\frac{\exp(y)}{s} \right)^c dt$$

Faça a seguinte mudança de variável $u = \left(\frac{\exp(y)}{s}\right)^c$, então, $du = c\left(\frac{\exp(y)}{s}\right)^{c-1} dy$.

Portanto,

$$M_Y(T) = \int_{-\infty}^{\infty} (su^{\frac{1}{c}})^t k(1+u)^{-k-1} du$$

Agora, faça outra mudança de variável, considere que $v = \frac{1}{1+u}$, então $dv = -(1+u)^{-2} du$. E reescreva a integral

$$M_Y(T) = \int_0^1 -s^t k(1-v)^{\frac{t}{c}} v^{-\frac{t}{c}+k-1} dv$$

$$ks^t B\left[\frac{t}{c} + 1, k - \frac{t}{c}\right], \quad \text{se } kc > t$$

em que $B(a, b)$ é a função beta completa (LAWLESS, 2003). Para obter a segunda identidade, reorganize o integrando como o núcleo da função de densidade da distribuição beta.

Apêndice C - Elementos da matriz de informação observada para o modelo de regressão log-Burr XII

$$\mathbf{L}_{kk} = -\frac{r}{k^2} \quad \mathbf{L}_{k\sigma} = \sum_{i=1}^n \frac{z_i}{\sigma} h_i \quad \mathbf{L}_{k\beta_j} = \sum_{i=1}^n \frac{x_{ij}}{\sigma} h_i$$

$$\mathbf{L}_{\sigma\sigma} = \frac{r}{\sigma^2} + \frac{2}{\sigma^2} \sum_{i \in F} z_i - 2(k+1) \sum_{i \in F} \frac{z_i}{\sigma^2} h_i - (k+1) \sum_{i \in F} \left(\frac{z_i}{\sigma}\right)^2 h_i +$$

$$\sum_{i \in F} \left(\frac{z_i}{\sigma} h_i\right)^2 - k \sum_{i \in C} \frac{z_i}{\sigma^2} h_i - k \sum_{i \in C} \left(\frac{z_i}{\sigma}\right)^2 h_i + k \sum_{i=1}^n \left(\frac{z_i}{\sigma} h_i\right)^2$$

$$\mathbf{L}_{\sigma\beta_j} = \sum_{i \in F} \frac{x_{ij}}{\sigma^2} - (k+1) \sum_{i \in F} \left(\frac{x_{ij}}{\sigma^2} + \frac{x_{ij}}{z_i} \sigma^2\right) h_i + \sum_{i \in F} \frac{x_{ij}}{\sigma} \frac{z_i}{\sigma} h_i^2$$

$$- k \sum_{i \in C} \left(\frac{x_{ij}}{\sigma^2} + \frac{x_{ij}}{z_i} \sigma^2\right) h_i + k \sum_{i=1}^n \frac{x_{ij}}{\sigma} \frac{z_i}{\sigma} h_i^2$$

$$\mathbf{L}_{\beta_j\beta_s} = -(k+1) \sum_{i \in F} \frac{x_{ij}x_{is}}{\sigma^2} h_i + (k+1) \sum_{i \in F} \frac{x_{ij}x_{is}}{\sigma^2} h_i^2 -$$

$$k \sum_{i \in C} \frac{x_{ij}x_{is}}{\sigma^2} h_i + k \sum_{i \in C} \frac{x_{ij}x_{is}}{\sigma^2} h_i^2$$

em que $j, s = 1, 2, \dots, p$, $h_i = \frac{\exp(z_i)}{1+\exp(z_i)}$ e $z_i = \frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}$

Apêndice D - Programa na linguagem de programação matricial Ox para verificar o desempenho do modelo de regressão log-Burr XII por meio do teste da razão de verossimilhança

```

/**PROGRAMA para simular de New Weibull e testar logistica**/
#include<oxstd.h> #include<oxdraw.h> #include<oxfloat.h>
#include<maximize.h> #include<simula.h> #import<maximize>

#pragma link("maximize.oxo")
static decl s_mT; //dados T simulados
static decl s_mX; //covariável simulada

static decl s_mC; //censura simulada

static decl dados; // valores de y e da censura

log_vero(const vP,const adFunc, const avScore,const amHessian) {
decl n=rows(s_mT); decl cont,y,cens,xx0, xx1 ; decl uns=ones(n,1);
decl vero=zeros(1,n); decl a1=fabs(vP[0][0]); //k

decl a2=fabs(vP[1][0]); //sigma decl a3=vP[2][0]; //beta 0

decl a4=vP[3][0]; //beta 1

for(cont=0;cont<n;++cont) {
//y=log(dados[cont][0]);
y=log(dados[cont][0]); cens=s_mC[cont][0]; xx0=1; xx1=s_mX[cont][0];

decl xbeta=xx0*a3+xx1*a4; decl z=(y-xbeta)/a2;

if(dados[cont][1]==1)
vero[0][cont]=log(a1)-log(a2)+z-(a1+1)*log(1+exp(z));
if(dados[cont][1]==0) vero[0][cont]= -a1*log(1+exp(z)); }

adFunc[0]=double(vero*uns); return 1; }
/////////////////////////////////////////////////////////////////
log_vero1(const vP1,const adFunc1, const avScore,const amHessian) {
decl n=rows(s_mT); decl cont,y,cens,xx0, xx1 ; decl uns=ones(n,1);
decl vero1=zeros(1,n); decl a2=fabs(vP1[0][0]); //sigma

decl a3=vP1[1][0]; //beta 0 decl a4=vP1[2][0]; //beta 1

```

```

for(cont=0;cont<n;++cont) {
//y=log(dados[cont][0]);
y=log(dados[cont][0]); cens=s_mC[cont][0]; xx0=1; xx1=s_mX[cont][0];

decl xbeta=xx0*a3+xx1*a4; decl z=(y-xbeta)/a2;
if(dados[cont][1]==1)
vero1[0][cont]=-log(a2)+z-(1+1)*log(1+exp(z));
if(dados[cont][1]==0)
vero1[0][cont]= -1*log(1+exp(z)); }
adFunc1[0]=double(vero1*uns);
return 1; }
////////////////////////////////////
main() { decl nc=30; decl dfunc; decl dfunc1; decl rep=1000; decl
inter=0; decl ncensur=zeros(rep,1); decl lambda=zeros(rep,1);

do { decl unif; unif=ranu(nc,1); s_mX=ranu(nc,1); print(" os valores
da covariável \n", s_mX);

//gerando T

decl beta0=-3; decl beta1=3; decl uns1=ones(nc,1);

decl alpha=exp(beta0.*uns1+beta1.*s_mX[][0]); decl k=0.27;

decl sigma=0.36;

s_mT=alpha.*(((1-unif).^(-1/k))-1).^sigma;
decl m; decl scensur;
scensur=0; dados=ones(nc,2); s_mC=0.4*ranu(nc,1);

for (m=0; m<nc; m++) { if (s_mC[m][0]-s_mT[m][0] < 0 && scensur <3)
{ dados[m][0]= s_mC[m][0];
dados[m][1]=0;
scensur=scensur+1; }
else
{dados[m][0]= s_mT[m][0]; }
}

print(" os valores de T e da censura \n", dados);

decl vP=<0.27;0.36;-3;3>; log_vero(vP,&dfunc,0,0);
decl mhess=0.00001*unit(4);

```

```

decl ir; MaxControl(-1,20);
ir=MaxBFGS(log_vero,&vP,&dfunc,&mhess,1);
////////////////////////////////////
decl vP1=<0.36;-3;3>; log_vero1(vP1,&dfunc1,0,0); decl ir1;
MaxControl(100,20);
ir1=MaxBFGS(log_vero1,&vP1,&dfunc1,&mhess,1);
////////////////////////////////////
if (ir == MAX_CONV && ir1 == MAX_CONV )

{ print(" os valores dos parametros \n", vP);

print(" os valores dos parametros \n", vP1);

println("vero=" ,dfunc); println("vero1=" ,dfunc1);

decl verologistica=dfunc1; decl veroburr=dfunc;
ncensur[inter][0]=sumc(dados[][1]);

//teste da razão de verossimilhança
lambda[inter][0]=-2*(verologistica-veroburr); ++inter; } } while
(inter<rep);

print(" teste da razao de verossimilhanca \n", lambda); print("
número de censuras na amostra \n", ncensur);

decl j; decl pvalor=ones(rep,1); for (j=0; j<rep; j++) { if
(lambda[j][0] > 3.841459)
    { pvalor[j][0]= 1; }
else
    {pvalor[j][0]= 0; }
}

decl contagem=sumc(pvalor); print("numero de amostras em que H0 é
rejeitada \n", contagem); }

```

Apêndice E - Programa na linguagem de programação matricial OX para simular resíduo tipo martingale a partir do modelo de regressão log-Weibull estendida e programa no *software* R para obter gráfico de probabilidade normal dos resíduos simulados

```

    /**PROGRAMA PARA CRIAR GERAR RESÍDUO TIPO MARTINGALE **//
        /**Modelo log-Weibull estendida**//

#include<oxstd.h> #include<oxdraw.h> #include<oxfloat.h>
#include<maximize.h> #include<simula.h> #import<maximize>

#pragma link("maximize.oxo")
static decl s_mT; //dados T simulados
static decl s_mX; //covariável simulada

static decl s_mC; //censura simulada

static decl dados; // valores de y e da censura

log_vero(const vP,const adFunc, const avScore,const amHessian) {
decl n=rows(s_mT); decl cont,y,cens,xx0, xx1 ; decl uns=ones(n,1);
decl vero=zeros(1,n); decl a1=fabs(vP[0][0]); //lambda

decl a2=fabs(vP[1][0]); //sigma decl a3=vP[2][0]; //beta 0

decl a4=vP[3][0]; //beta 1

for(cont=0;cont<n;++cont) { y=log(dados[cont][0]);
cens=s_mC[cont][0]; xx0=1; xx1=s_mX[cont][0];

decl xbeta=xx0*a3+xx1*a4; decl z=(y-xbeta)/a2;

if(dados[cont][1]==1) vero[0][cont]=log(a1+0.0000000001)
-log(a2+0.0000000001)+xbeta+z+exp(z)+a1*exp(xbeta)*(1-exp(exp(z)));
if(dados[cont][1]==0) vero[0][cont]= a1*exp(xbeta)*(1-exp(exp(z)));
}

adFunc[0]=double(vero*uns);
return 1; }

main() { decl nc=30; decl dfunc; decl rep=1000; decl
res=zeros(nc,rep); decl residuos=zeros(nc,rep); decl

```

```

ordelinha=zeros(nc,rep); decl inter=0; decl ncensur=zeros(rep,1); do
{

decl unif; unif=ranu(nc,1); print(" os valores DA UNIFORME \n",
unif); s_mX=ranu(nc,1); print(" os valores da covariável \n", s_mX);

//gerando T

decl beta0=-3; decl beta1=3; decl uns1=ones(nc,1); decl
alpha=exp(beta0.*uns1+beta1.*s_mX[] [0]); decl lambda=0.1; decl
sigma=2;
s_mT=( ((alpha).^(1/sigma)).*log(1-(log(1-unif)./(lambda.*alpha)))
).^ (sigma);

print(" os valores simulados \n", s_mT);

decl m; decl scensur; scensur=0; dados=ones(nc,2);
s_mC=0.5*ranu(nc,1); print(" os valores simulados da censura \n",
s_mC);

for (m=0; m<nc; m++)
{
if (s_mC[m] [0]-s_mT[m] [0] < 0 && scensur < 3)
{ dados[m] [0]= s_mC[m] [0];
dados[m] [1]=0;
scensur=scensur+1; }
else
{dados[m] [0]= s_mT[m] [0]; }
}

print("número de censurados \n", scensur); print(" os valores de T e
da censura \n", dados);

decl vP=<0.1;2;-3;3>; log_vero(vP,&dfunc,0,0); println("vero="
,dfunc);

decl mhess=0.00001*unit(4); decl ir; MaxControl(-1,20);
ir=MaxBFGS(log_vero,&vP,&dfunc,&mhess,1); if (ir == MAX_CONV) {

print(" os valores dos parametros \n", vP);
ncensur[inter] [0]=sumc(dados[] [1]);
decl xbeta1=(vP[2] [0]*uns1+
vP[3] [0].*s_mX[] [0]);

```



```

decl zz=(log(s_mT[][0])-xbeta1); decl zzz=zz./fabs(vP[1][0]); decl
sob2=1-exp(exp(zzz)); decl
sob3=exp(fabs(vP[0][0]).*exp(xbeta1).*sob2);

print(" os valores de sob3 \n", sob3);

decl martingal= dados[][1]+log(sob3+0.000000000001);

res[][inter]= (martingal./fabs(martingal)).*
(1.4142*(-martingal-(dados[][1].*log(dados[][1]-martingal))).^(1/2));;

decl resorde=(sortc(res[][inter])); residuos[][inter]=resorde;

++inter; } } while (inter<rep);

print(" os valores dos residuos simulados \n", res); print(" número
de censuras na amostra \n", ncensur);

decl j; decl mresults = zeros (nc,3); for (j=0; j < nc; ++j) {
// calculando estatísticas
mresults[j][0]=min(residuos[j][]);
mresults[j][1]=meanr(residuos[j][]);
mresults[j][2]=max(residuos[j][]); } print(" estatísticas dos
valores simulados \n", mresults); print(" número de iterações \n",
inter); }

library(RODBC)
x1<-seq(-3,3,0.01) y1<-x1

# Read EXCEL sheet containing spatial data
channel <-
odbcConnectExcel("MARTI_50_SC.xls") MARTISC <- sqlFetch(channel,
"MARTI_50_SC") odbcClose(channel)

channel <- odbcConnectExcel("DEVI_50_SC.xls") DEVISC <-
sqlFetch(channel, "DEVI_50_SC") odbcClose(channel)

q1<-qqnorm(MARTISC$res50ESC,plot.it = FALSE)
q2<-qqnorm(DEVISC$res50ESC,plot.it = FALSE)

plot(range(q1$x, q2$x), range(q1$y, q2$y), type = "n",,main = "",
xlab="Percentis da N(0,1)",ylab="Resíduo") points(q1,col =

```

```
"blue",pch=1) points(q2, col = "red", pch = 3) lines(x1,y1)

texto<-expression(r[M],r[D])

legend("bottomright",texto,text.col=c("blue","red"),col=c("blue","red"),pch
= c(1,3), cex = 1,bty = "o" )
mtext("0% de censura", cex = 0.8, line =1,side=3)
legend("topleft",expression(k=="0,27"), cex = 1,bty = "n")
```

Apêndice F - Programa na linguagem de programação matricial OX para estimar os parâmetros do modelo de regressão log-Weibull estendida ajustado aos dados dos peixes da espécie *Golden Shiner* usando o método de máxima verossimilhança e calcular os resíduos

```
/**PROGRAMA PARA ESTIMAR OS PARÂMETROS DO MODELO LOG NOVA WEIBULL
    ESTENDIDA**//
    /**DADOS REAIS - PEIXES DA EXPÉCIE GOLDEN SHINER**//
#include<oxstd.h> #include<oxdraw.h> #include<oxfloat.h>
#include<maximize.h> #include<simula.h> #pragma link("maximize.oxo")
static decl g_mY; static decl g_mX;

log_vero(const vP,const adFunc, const avScore,const amHessian) {
decl n=rows(g_mY); decl cont,y,xx0, xx1, xx2, xx3, xx4,xx5, xx6,
xx7; decl uns=ones(n,1); decl vero=zeros(1,n); decl
a1=fabs(vP[0][0]); //lambda decl a2=fabs(vP[1][0]); //sigma

decl a3=vP[2][0]; //beta 0 decl a4=vP[3][0]; //beta 1

decl a5=vP[4][0]; //beta 2 decl a6=vP[5][0]; //beta 3

decl a7=vP[6][0]; //beta 4 decl a8=vP[7][0]; //beta 5

decl a9=vP[8][0]; //beta 6 decl a10=vP[9][0]; //beta 7

for(cont=0;cont<n;++cont) { y=g_mY[cont][0]; xx0=g_mY[cont][2];
xx1=g_mY[cont][3]; xx2=g_mY[cont][4]; xx3=g_mY[cont][5];
xx4=g_mY[cont][6]; xx5=g_mY[cont][7]; xx6=g_mY[cont][8];
xx7=g_mY[cont][9];

decl xbeta=xx0*a3+xx1*a4+xx2*a5+xx3*a6+xx4*a7+xx5*a8+xx6*a9+xx7*a10;
```

```

decl z=(y-(xbeta))/a2;

if(g_mY[cont][1]==1) vero[0][cont]=log(a1+0.00000000001)
-log(a2+0.00000000001)+xbeta+z+exp(z)+a1*exp(xbeta)*(1-exp(exp(z)));
if(g_mY[cont][1]==0) vero[0][cont]= a1*exp(xbeta)*(1-exp(exp(z))); }

adFunc[0]=double(vero*uns);
return 1; }

main() {
  g_mY = loadmat('c:\TESE\New_Weibull\especie22.txt') ;

  //print(" Dados",g_mY);
  decl nc=rows(g_mY); decl dfunc; decl avScore; decl
  vP=<0.5;3;5.3439;1.9119;0.1054;-0.1229;0.0367;0.0364;0.2627;-0.09434>;
  log_vero(vP,&dfunc,&avScore,0); println("vero=" ,dfunc);

  MaxControl(-1,20); decl mhess=0.00001*unit(4); decl ir, var;
  ir=MaxBFGS(log_vero,&vP,&dfunc,&mhess,1);
  Num1Derivative(log_vero,vP,&avScore);
  Num2Derivative(log_vero,vP,&mhess); var=invertsym((-1)*mhess);
  print(" os valores das variâncias \n",var); print(" os valores dos
  parametros \n", vP); print(" os valores score \n", avScore);

  decl zz=(g_mY[] [0]-(vP[2] [0].*g_mY[] [2]+
                vP[3] [0].*g_mY[] [3]+
                vP[4] [0].*g_mY[] [4]+
                vP[5] [0].*g_mY[] [5]+
                vP[6] [0].*g_mY[] [6]+
                vP[7] [0].*g_mY[] [7]+
                vP[8] [0].*g_mY[] [8]+
                vP[9] [0].*g_mY[] [9]));
  print(" os valores de zz \n", zz);

  decl xbeta1=(vP[2] [0].*g_mY[] [2]+
                vP[3] [0].*g_mY[] [3]+
                vP[4] [0].*g_mY[] [4]+
                vP[5] [0].*g_mY[] [5]+
                vP[6] [0].*g_mY[] [6]+
                vP[7] [0].*g_mY[] [7]+
                vP[8] [0].*g_mY[] [8]+
                vP[9] [0].*g_mY[] [9]);
  print(" os valores de xbeta1 \n", xbeta1);

```

```

decl res=zz./vP[1][0]; print(" os valores dos residuos \n", res);

decl sob1=exp(res); print(" os valores de exp(res) \n", sob1);

decl sob2=1-exp(sob1);

print(" os valores de 1-exp(exp(res)) \n", sob2);

decl sob3=exp(vP[0][0].*exp(xbeta1).*sob2);

print(" os valores de sob estimada para y\n", sob3);

decl martingale= g_mY[][1]+log(sob3); print(" os valores do residuo
martingale \n", martingale);

decl tipmartin= (martingale./fabs(martingale)).*
(1.4142*(-martingale-(g_mY[][1].*log(g_mY[][1]-martingale))).^(1/2));
print(" os valores do residuo tipo martingale \n", tipmartin);
}

```

Apêndice G - Programa na linguagem de programação matricial OX para obter as estimativas de jackknife e medidas de influência global para o modelo de regressão log-Weibull estendida ajustado aos dados dos peixes da espécie *Golden Shiner*

```

/**PROGRAMA PARA ESTIMAR OS PARÂMETROS DO MODELO DE REGRESSÃO**\
  /**LOG-NOVA WEIBULL ESTENDIDA - Método de Jackknife**//
    /**MEDIDAS DE INFLUÊNCIA GLOBAL **//
      /**DADOS PEIXES DA ESPÉCIE GOLDEN SHINER**//
#include<oxstd.h> #include<oxdraw.h> #include<oxfloat.h>
#include<maximize.h> #include<simula.h> #pragma link("maximize.oxo")
static decl g_mY; static decl g_mX;

log_vero(const vP,const adFunc, const avScore,const amHessian) {
decl n=rows(g_mY);

decl cont,t,xx0, xx1, xx2, xx3, xx4,xx5, xx6, xx7;

decl uns=ones(n,1); decl vero=zeros(1,n);

decl a1=fabs(vP[0][0]);//lambda decl a2=fabs(vP[1][0]); //sigma

```

```

decl a3=vP[2][0]; //beta 0 decl a4=vP[3][0]; //beta 1

decl a5=vP[4][0]; //beta 2 decl a6=vP[5][0]; //beta 3

decl a7=vP[6][0]; //beta 4 decl a8=vP[7][0]; //beta 5

decl a9=vP[8][0]; //beta 6 decl a10=vP[9][0]; //beta 7

for(cont=0;cont<n;++cont) {
//t=log(g_mY[cont][0]);
t=g_mY[cont][0]; xx0=g_mY[cont][2]; xx1=g_mY[cont][3];
xx2=g_mY[cont][4]; xx3=g_mY[cont][5]; xx4=g_mY[cont][6];
xx5=g_mY[cont][7]; xx6=g_mY[cont][8]; xx7=g_mY[cont][9];

decl xbeta=xx0*a3+xx1*a4+xx2*a5+xx3*a6+xx4*a7+xx5*a8+xx6*a9+xx7*a10;
decl z=(t-(xbeta))/a2;

if(g_mY[cont][1]==1) vero[0][cont]=log(a1+0.0000000001)
-log(a2+0.0000000001)+xbeta+z+exp(z)+a1*exp(xbeta)*(1-exp(exp(z)));
if(g_mY[cont][1]==0) vero[0][cont]= a1*exp(xbeta)*(1-exp(exp(z))); }

adFunc[0]=double(vero*uns);
return 1; }

main() {
g_mY = loadmat("c:\TESE\New_Weibull\especie22.txt") ;

decl nc=rows(g_mY); decl dfunc; decl avScore;

decl
vP=<0.5;3;2.675;1.071;0.07151;-0.088;0.02455;0.009911;0.04693;-0.0231>;

decl LD=zeros(nc,1); log_vero(vP,&dfunc,&avScore,0);

println("vero=" ,dfunc);

MaxControl(-1,20); decl mhess=0.00001*unit(4); decl ir, var, var1;
ir=MaxBFGS(log_vero,&vP,&dfunc,&mhess,1);
Num1Derivative(log_vero,vP,&avScore);
Num2Derivative(log_vero,vP,&mhess);

var=-1/mhess; var1=invertsym((-1)*mhess);

```

```

print(" os valores da mhess  \n",mhess);

print(" os valores das variâncias  \n",var);

print(" os valores das variâncias 1  \n",var1);

print(" os valores dos parametros  \n", vP);

print(" os valores score  \n", avScore);

decl deriv1=zeros(nc,10); decl CD2=zeros(nc,1);

//Cálculo das estimativas para jackknife usando fórmula do artigo

decl t,xx0, xx1, xx2, xx3, xx4,xx5, xx6, xx7; decl j; decl i; decl
theta=zeros(nc,10); decl veronova=ones(nc,1); decl
PseudoValues=ones(nc,10);

for (j=0; j<nc ; ++j) {

decl mat1=dropr(g_mY,j); //elimina linha j dos dados

decl nc1=rows(mat1); decl theta1=zeros(nc1,10);

decl uns1=ones(nc1,1); decl uns2=ones(1,nc); decl vero1= ones(nc,1);

for (i=0; i<nc1 ; ++i) {

decl zz=(mat1[i] [0]-(vP[2] [0].*mat1[i] [2]+
                vP[3] [0].*mat1[i] [3]+
                vP[4] [0].*mat1[i] [4]+
                vP[5] [0].*mat1[i] [5]+
                vP[6] [0].*mat1[i] [6]+
                vP[7] [0].*mat1[i] [7]+
                vP[8] [0].*mat1[i] [8]+
                vP[9] [0].*mat1[i] [9]));

decl xbeta1=(vP[2] [0].*mat1[i] [2]+
             vP[3] [0].*mat1[i] [3]+
             vP[4] [0].*mat1[i] [4]+
             vP[5] [0].*mat1[i] [5]+
             vP[6] [0].*mat1[i] [6]+
             vP[7] [0].*mat1[i] [7]+

```

```

vP[8][0].*mat1[i][8]+
vP[9][0].*mat1[i][9]);

//lambda
if(mat1[i][1]==1)
theta1[i][0]=(1/vP[0][0])+(exp(xbeta1)*(1-exp(exp(zz/vP[1][0]))));
if(mat1[i][1]==0)
theta1[i][0]=(exp(xbeta1)*(1-exp(exp(zz/vP[1][0]))));

//sigma
if(mat1[i][1]==1)
theta1[i][1]=-(1/vP[1][0])-(1/vP[1][0])*(zz/vP[1][0])-(1/vP[1][0])*
(zz/vP[1][0])*exp(zz/vP[1][0])+(vP[0][0]/vP[1][0])*exp(xbeta1)*
(zz/vP[1][0])*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
if(mat1[i][1]==0)
theta1[i][1]=(vP[0][0]/vP[1][0])*exp(xbeta1)*(zz/vP[1][0])*
exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));

//beta0
if(mat1[i][1]==1) theta1[i][2]=mat1[i][2]-(mat1[i][2]/vP[1][0])-
(mat1[i][2]/vP[1][0])*exp(zz/vP[1][0])+vP[0][0]*mat1[i][2]*
exp(xbeta1)-vP[0][0]*mat1[i][2]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][2]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
if(mat1[i][1]==0) theta1[i][2]=vP[0][0]*mat1[i][2]*
exp(xbeta1)-vP[0][0]*mat1[i][2]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][2]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));

//beta1
if(mat1[i][1]==1) theta1[i][3]=mat1[i][3]-(mat1[i][3]/vP[1][0])-
(mat1[i][3]/vP[1][0])*exp(zz/vP[1][0])+vP[0][0]*mat1[i][3]*
exp(xbeta1)-vP[0][0]*mat1[i][3]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][3]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
if(mat1[i][1]==0) theta1[i][3]=vP[0][0]*mat1[i][3]*
exp(xbeta1)-vP[0][0]*mat1[i][3]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][3]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));

//beta2
if(mat1[i][1]==1) theta1[i][4]=mat1[i][4]-(mat1[i][4]/vP[1][0])-
(mat1[i][4]/vP[1][0])*exp(zz/vP[1][0])+vP[0][0]*mat1[i][4]*
exp(xbeta1)-vP[0][0]*mat1[i][4]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+

```

```

vP[0][0]*(mat1[i][4]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
if(mat1[i][1]==0) theta1[i][4]=vP[0][0]*mat1[i][4]*
exp(xbeta1)-vP[0][0]*mat1[i][4]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][4]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
//beta3
if(mat1[i][1]==1) theta1[i][5]=mat1[i][5]-(mat1[i][5]/vP[1][0])-
(mat1[i][5]/vP[1][0])*exp(zz/vP[1][0])+vP[0][0]*mat1[i][5]*
exp(xbeta1)-vP[0][0]*mat1[i][5]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][5]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
if(mat1[i][1]==0) theta1[i][5]=vP[0][0]*mat1[i][5]*
exp(xbeta1)-vP[0][0]*mat1[i][5]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][5]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
//beta4
if(mat1[i][1]==1) theta1[i][6]=mat1[i][6]-(mat1[i][6]/vP[1][0])-
(mat1[i][6]/vP[1][0])*exp(zz/vP[1][0])+vP[0][0]*mat1[i][6]*
exp(xbeta1)-vP[0][0]*mat1[i][6]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][6]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
if(mat1[i][1]==0) theta1[i][6]=vP[0][0]*mat1[i][6]*
exp(xbeta1)-vP[0][0]*mat1[i][6]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][6]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
//beta5
if(mat1[i][1]==1) theta1[i][7]=mat1[i][7]-(mat1[i][7]/vP[1][0])-
(mat1[i][7]/vP[1][0])*exp(zz/vP[1][0])+vP[0][0]*mat1[i][7]*
exp(xbeta1)-vP[0][0]*mat1[i][7]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][7]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
if(mat1[i][1]==0) theta1[i][7]=vP[0][0]*mat1[i][7]*
exp(xbeta1)-vP[0][0]*mat1[i][7]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][7]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
//beta6
if(mat1[i][1]==1) theta1[i][8]=mat1[i][8]-(mat1[i][8]/vP[1][0])-
(mat1[i][8]/vP[1][0])*exp(zz/vP[1][0])+vP[0][0]*mat1[i][8]*
exp(xbeta1)-vP[0][0]*mat1[i][8]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][8]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
if(mat1[i][1]==0) theta1[i][8]=vP[0][0]*mat1[i][8]*

```



```

exp(xbeta1)-vP[0][0]*mat1[i][8]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][8]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
//beta7
if(mat1[i][1]==1) theta1[i][9]=mat1[i][9]-(mat1[i][9]/vP[1][0])-
(mat1[i][9]/vP[1][0])*exp(zz/vP[1][0])+vP[0][0]*mat1[i][9]*
exp(xbeta1)-vP[0][0]*mat1[i][9]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][9]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0]));
if(mat1[i][1]==0) theta1[i][9]=vP[0][0]*mat1[i][9]*
exp(xbeta1)-vP[0][0]*mat1[i][9]*exp(xbeta1)*exp(exp(zz/vP[1][0]))+
vP[0][0]*(mat1[i][9]/vP[1][0])*
exp(xbeta1)*exp(zz/vP[1][0])*exp(exp(zz/vP[1][0])); }

deriv1[j][0]=(theta1[][0])'*uns1; deriv1[j][1]=(theta1[][1])'*uns1;
deriv1[j][2]=(theta1[][2])'*uns1; deriv1[j][3]=(theta1[][3])'*uns1;
deriv1[j][4]=(theta1[][4])'*uns1; deriv1[j][5]=(theta1[][5])'*uns1;
deriv1[j][6]=(theta1[][6])'*uns1; deriv1[j][7]=(theta1[][7])'*uns1;
deriv1[j][8]=(theta1[][8])'*uns1; deriv1[j][9]=(theta1[][9])'*uns1;
decl parte1=(var1*(deriv1[j][]))';
//estimativas dos parâmetros sem observação j
theta[j][0]=vP[0][0]+parte1[0][0];
theta[j][1]=vP[1][0]+parte1[1][0];
theta[j][2]=vP[2][0]+parte1[2][0];
theta[j][3]=vP[3][0]+parte1[3][0];
theta[j][4]=vP[4][0]+parte1[4][0];
theta[j][5]=vP[5][0]+parte1[5][0];
theta[j][6]=vP[6][0]+parte1[6][0];
theta[j][7]=vP[7][0]+parte1[7][0];
theta[j][8]=vP[8][0]+parte1[8][0];
theta[j][9]=vP[9][0]+parte1[9][0];
//distancia de cook
CD2[j][0]=2.*(((theta[j][])'-vP)')*(-mhess)*((theta[j][])'-vP);
//jackknife
PseudoValues[j][0]=nc.*fabs(theta[j][0])-(nc-1).*fabs(vP[0][0]);
PseudoValues[j][1]=nc.*fabs(theta[j][1])-(nc-1).*fabs(vP[1][0]);
PseudoValues[j][2]=nc.*theta[j][2]-(nc-1).*vP[2][0];
PseudoValues[j][3]=nc.*theta[j][3]-(nc-1).*vP[3][0];
PseudoValues[j][4]=nc.*theta[j][4]-(nc-1).*vP[4][0];
PseudoValues[j][5]=nc.*theta[j][5]-(nc-1).*vP[5][0];
PseudoValues[j][6]=nc.*theta[j][6]-(nc-1).*vP[6][0];
PseudoValues[j][7]=nc.*theta[j][7]-(nc-1).*vP[7][0];
PseudoValues[j][8]=nc.*theta[j][8]-(nc-1).*vP[8][0];

```

```

PseudoValues[j][9]=nc.*theta[j][9]-(nc-1).*vP[9][0];

decl m; for(m=0;m<nc;++m) { t=g_mY[m][0]; xx0=g_mY[m][2];
xx1=g_mY[m][3]; xx2=g_mY[m][4]; xx3=g_mY[m][5]; xx4=g_mY[m][6];
xx5=g_mY[m][7]; xx6=g_mY[m][8]; xx7=g_mY[m][9]; decl
xbeta=xx0*theta[j][2]+xx1*theta[j][3]+xx2*theta[j][4]+xx3*theta[j][5]
+xx4*theta[j][6]+xx5*theta[j][7]+xx6*theta[j][8]+xx7*theta[j][9];
decl z=(t-(xbeta))/theta[j][1];

if(g_mY[m][1]==1) vero1[m][0]=log(theta[j][0]+0.00000000001)
-log(theta[j][1]+0.00000000001)+xbeta+z+exp(z)+theta[j][0]
*exp(xbeta)*(1-exp(exp(z)));
if(g_mY[m][1]==0) vero1[m][0]=
theta[j][0]*exp(xbeta)*(1-exp(exp(z))); }

veronova[j][0]=(uns2*vero1); decl veroori=-204.778459528;
LD[j][0]=2*(veroori-veronova[j][0]); }

print(" os valores das estimativas sem uma observação \n", theta);
print(" log vero sem uma observação \n", veronova); print("
distancia de cook CD2 \n", CD2); print(" afastamento de vero LD \n",
LD);

//Calculando estimativas de jackknife

decl m; decl mediaPV = zeros (1,10); decl desvioPV = zeros (1,10);
for (m=0; m < 10; ++m) {
// calculando média
mediaPV[0][m]=meanc(PseudoValues[][m]);
desvioPV[0][m]=sqrt(varc(PseudoValues[][m])); }
print(" média dospseudo valores \n", mediaPV); // estimativas de
jackknife print(" desvio padrão dos pseudo valores \n", desvioPV); }

```

Apêndice H - Programa na linguagem de programação matricial OX para obter medidas de influência local para o modelo de regressão log-Weibull estendida ajustado aos dados de peixes da espécie *Golden shiner*

```

/**PROGRAMA para Influencia Local - perturbacao de casos**\
    /**Modelo Log nova Weibull Estendida**//
    /**DADOS REAIS - Peixes da espécie Golden shiner **//

```

```

#include<oxstd.h> #include<oxdraw.h> #include<oxfloat.h>
#include<maximize.h> #include<simula.h> #pragma link("maximize.oxo")
static decl g_mY; static decl g_mX;

log_vero(const vP,const adFunc, const avScore,const amHessian) {
decl n=rows(g_mY); decl cont,y,xx0, xx1, xx2, xx3, xx4,xx5, xx6,
xx7; decl uns=ones(n,1); decl vero=zeros(1,n);

decl a1=fabs(vP[0][0]); //lambda
decl a2=fabs(vP[1][0]); //sigma
decl a3=vP[2][0]; //beta 0 decl a4=vP[3][0]; //beta 1

decl a5=vP[4][0]; //beta 2 decl a6=vP[5][0]; //beta 3

decl a7=vP[6][0]; //beta 4 decl a8=vP[7][0]; //beta 5

decl a9=vP[8][0]; //beta 6 decl a10=vP[9][0]; //beta 7

decl w11=vP[10:][0]; /*peso*/

for(cont=0;cont<n;++cont) { y=g_mY[cont][0]; xx0=g_mY[cont][2];
xx1=g_mY[cont][3]; xx2=g_mY[cont][4]; xx3=g_mY[cont][5];
xx4=g_mY[cont][6]; xx5=g_mY[cont][7]; xx6=g_mY[cont][8];
xx7=g_mY[cont][9];
decl xbeta=xx0*a3+xx1*a4+xx2*a5+xx3*a6+xx4*a7+xx5*a8+xx6*a9+xx7*a10;
decl z=(y-(xbeta))/a2;

if(g_mY[cont][1]==1) vero[0][cont]=w11[cont]*(log(a1+0.00000000001)
-log(a2+0.00000000001)+xbeta+z+exp(z)+a1*exp(xbeta)*(1-exp(exp(z)))));
if(g_mY[cont][1]==0) vero[0][cont]=w11[cont]*
a1*exp(xbeta)*(1-exp(exp(z))); } adFunc[0]=double(vero*uns); return
1; }

main() {
g_mY = loadmat("c:\TESE\New_Weibull\especie22.txt") ;
decl nc=rows(g_mY); decl dfunc; decl
vP=0.018947|1.5969|7.3999|-3.3241|-0.23705|0.23085|
-0.091456|-0.11203|-0.18403|-0.094425|ones(nc,1);
//ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA
log_vero(vP,&dfunc,0,0); println("vero=" ,dfunc);
MaxControl(500,20); decl mhess=0.00001*unit(80);
Num1Derivative(log_vero,vP,&mhess);
Num2Derivative(log_vero,vP,&mhess);
decl QQ, delta,FF, HHH, HHH1;

```

```

    QQ=mhess[0:9][0:9];
    delta=mhess[0:9][10:];
FF=delta*(invertsym((-1)*QQ))*delta; HHH=diagonal(FF); HHH1=HHH';
decl aFFval, aFFvec; eigensym(FF, &aFFval, &aFFvec);
    print(" primeiro autovetor da matriz FF ",aFFvec[][0]);
    print("VALORES DA MEDIDA DE influencia local total",HHH1);
}

```

Apêndice I - Programa no *software* R para resíduo tipo martingale com envelope simulado para o modelo de regressão log- Weibull estendida ajustado aos dados da espécie de peixe *Golden Shiner*

```

residuos.dat<-scan("residuos.dat",
what=list(min=0,media=0,max=0,deviance=0))
dados<-attach(residuos.dat)
faixa<-range(deviance,min,max)par(pty="s")
qqnorm(deviance,main="",xlab="Percentis da N(0,1)",ylab="Resíduo
tipo Martingale", ylim=faixa)
par(new=T)
qqnorm(min,main="",axes=F,xlab="",
ylab="",type="l",ylim=faixa,lty=1)
par(new=T)
qqnorm(max,main="",axes=F,xlab="",
ylab="",type="l",ylim=faixa,lty=1)
par(new=T)
qqnorm(media,main="",axes=F,xlab="",
ylab="",type="l",ylim=faixa,lty=1)

```

Apêndice J - Programa na linguagem de programação matricial OX para simular dados com efeito aleatório e observações censuradas para o modelo de regressão log-Burr XII com efeito aleatório gama e com o algoritmo EMMC para estimar

```

/**DADOS SIMULADOS **//

#include<oxstd.h> #include<oxfloat.h> #include <oxprob.h>
#include<maximize.h> #include<simula.h> #include<oxdraw.h>
#include<packages/gnudraw/gnudraw.h> #import <maxsq>

```

```

#pragma link("maximize.oxo") static decl g_mY; static decl g_mY1;
static decl g_mX;

static decl g_mG1; static decl g_mG2; static decl g_mG3; static decl
g_mG4; static decl s_mWe1; //grupo1 static decl s_mWe2;//grupo2
static decl s_mWe3; //grupo3 static decl s_mWe4; //grupo4

static decl s_mWe11; //grupo1 static decl s_mWe21; //grupo2
static decl s_mWe31; //grupo3 static decl s_mWe41; //grupo4

static decl s_mC; //dados do log tempo imputado

static decl s_mC1; //dados da censura

static decl s_mW; //efeito aleatorio

//grupo 1
verossi1(w1,const vP, const adparte1) {
g_mG1=selectifr(g_mY1,g_mY1[][2].==1); //selecionar grupos

decl n1=rows(g_mG1[][3]); decl vero1=ones(1,n1);

decl cont,t,xx0, xx1; decl a1=fabs(vP[0][0]); //k

decl a2=fabs(vP[1][0]); //sigma decl a3=vP[2][0]; //beta 0

decl a4=vP[3][0]; //beta 1

for(cont=0;cont<n1;++cont) { t=log(g_mG1[cont][0]);
xx0=g_mG1[cont][3]; //inter
xx1=g_mG1[cont][4]; //x1

decl xbeta1=xx0*a3+xx1*a4; decl z=(t-w1-xbeta1)/a2;
vero1[0][cont]=(a1/a2)*((1+exp(z))^-a1-1))*exp(z); }
adparte1[0]=prodr(vero1); return 1; }

//grupo 2
verossi2(w2, const vP,const adparte2) {
g_mG2=selectifr(g_mY1,g_mY1[][2].==2); //selecionar grupos

decl n2=rows(g_mG2[][3]); decl vero2=ones(1,n2); decl cont,t,xx0,
xx1; decl a1=fabs(vP[0][0]); //k decl a2=fabs(vP[1][0]); //sigma
decl a3=vP[2][0]; //beta 0 decl a4=vP[3][0]; //beta 1

```

```

for( cont=0;cont<n2;++cont) { t=log(g_mG2[cont][0]);
xx0=g_mG2[cont][3]; xx1=g_mG2[cont][4];

decl xbeta2=xx0*a3+xx1*a4; decl z=(t-w2-xbeta2)/a2;
vero2[0][cont]=(a1/a2)*((1+exp(z))(-a1-1))*exp(z); }
adparte2[0]=prodr(vero2); return 1; }

//grupo 3
verossi3(w3, const vP,const adparte3) {
g_mG3=selectifr(g_mY1,g_mY1[][2].==3); //seleccionar grupos

decl n3=rows(g_mG3[][3]); decl vero3=ones(1,n3);

decl cont,t,xx0, xx1; decl a1=fabs(vP[0][0]); //k

decl a2=fabs(vP[1][0]); //sigma decl a3=vP[2][0]; //beta 0

decl a4=vP[3][0]; //beta 1

for( cont=0;cont<n3;++cont) { t=log(g_mG3[cont][0]);
xx0=g_mG3[cont][3]; xx1=g_mG3[cont][4];

decl xbeta3=xx0*a3+xx1*a4; decl z=(t-w3-xbeta3)/a2;
vero3[0][cont]=(a1/a2)*((1+exp(z))(-a1-1))*exp(z); }
adparte3[0]=prodr(vero3); return 1; }

//grupo 4
verossi4(w4, const vP,const adparte4) {
g_mG4=selectifr(g_mY1,g_mY1[][2].==4); //seleccionar grupos

decl n4=rows(g_mG4[][3]); decl vero4=ones(1,n4); decl cont,t,xx0,
xx1;

decl a1=fabs(vP[0][0]); //k decl a2=fabs(vP[1][0]); //sigma

decl a3=vP[2][0]; //beta 0 decl a4=vP[3][0]; //beta 1

for(cont=0;cont<n4;++cont) { t=log(g_mG4[cont][0]);
xx0=g_mG4[cont][3]; xx1=g_mG4[cont][4];

decl xbeta4=xx0*a3+xx1*a4; decl z=(t-w4-xbeta4)/a2;
vero4[0][cont]=(a1/a2)*((1+exp(z))(-a1-1))*exp(z); }
adparte4[0]=prodr(vero4); return 1; }

```

```

log_vero(const vP,const adFunc, const avScore,const amHessian) {
decl m_p=1001; decl n=rows(g_mY); decl uns=ones(n,1);

decl vero=zeros(1,n); decl uns1=ones(m_p,1);

decl vero1=zeros(1,m_p); decl cont,t,xx0, xx1;

decl a1=fabs(vP[0][0]); //k decl a2=fabs(vP[1][0]); //sigma

decl a3=vP[2][0]; //beta0 decl a4=vP[3][0]; //beta 1 decl p;

for(p=0;p<m_p;++p) { for(cont=0;cont<n;++cont) {

t=log(s_mC[cont][0]);
//t=g_mY[cont][2];
xx0=g_mY[cont][3]; xx1=g_mY[cont][4];

decl xbeta=xx0*a3+xx1*a4;

//grupo 1
if(g_mY[cont][2]==1 )
vero[0][cont]=log(a1+0.0000001)-log(a2+0.00000001)-(a1+1)*
(log(1+exp((t-xbeta-s_mWe1[p][0])/a2)))+
(t-xbeta-s_mWe1[p][0])/a2;

//grupo 2
if(g_mY[cont][2]==2 )
vero[0][cont]=log(a1+0.0000001)-log(a2+0.00000001)-(a1+1)*
(log(1+exp((t-xbeta-s_mWe2[p][0])/a2)))+
(t-xbeta-s_mWe2[p][0])/a2;

//grupo 3
if(g_mY[cont][2]==3)
vero[0][cont]=log(a1+0.0000001)-log(a2+0.00000001)-(a1+1)*
(log(1+exp((t-xbeta-s_mWe3[p][0])/a2)))+
(t-xbeta-s_mWe3[p][0])/a2;

//grupo 4
if(g_mY[cont][2]==4 )
vero[0][cont]=log(a1+0.0000001)-log(a2+0.00000001)-(a1+1)*
(log(1+exp((t-xbeta-s_mWe4[p][0])/a2)))+
(t-xbeta-s_mWe4[p][0])/a2;
}
}
}

```

```

} vero1[0][p]=double(vero*uns); }

adFunc[0]=(double(vero1*uns1))/m_p; return 1; }

log_vero1(const vP1,const adFunc1, const avScore1,const amHessian1)
{ decl m_p=1001; decl efal=ones(1,m_p); decl uns=ones(m_p,1);
  decl a6=fabs(vP1); //xi

  decl p;

  for(p=0;p<m_p;++p) {

    //ef aleat. gama
    efal[0][p]=((1/a6)-1)*
    (log(s_mWe1[p][0])+log(s_mWe2[p][0])+log(s_mWe3[p][0])+log(s_mWe4[p][0]))
    -(1/a6)*(s_mWe1[p][0]+s_mWe2[p][0]+s_mWe3[p][0]+s_mWe4[p][0])-
    (4/a6)*log(a6)-4*log(gammafact(1/a6)));

  }

  adFunc1[0]=double(efal*uns)/m_p; return 1; }

main() {
  decl time = timer();

  decl nc=120; g_mY = ones(nc,5) ;

  decl unif; unif=ranu(nc,1); g_mY[][4]=ranu(nc,1); //Covariável beta1

  decl grupo1=constant(1,30,1); decl grupo2=constant(2,30,1); decl
  grupo3=constant(3,30,1); decl grupo4=constant(4,30,1);

  g_mY[][2]=grupo1|grupo2|grupo3|grupo4;

  //gerando T
  decl beta0=1.5; decl beta1=1.5; decl
  mu=beta0.*g_mY[][3]+beta1.*g_mY[][4]; decl
  alpha=exp(beta0.*g_mY[][3]+beta1.*g_mY[][4]); decl k=0.25; decl
  sigma=0.5;

  decl valorZ = log( ((1-unif).^(-1/k))-1 ) ; decl teta=3;

```



```

decl w=rangamma(4,1,1/teta,1/ teta); // efeito aleatório
print(" os valores simulados do efeito aleatório\n", w);

decl v; for (v=0; v<120; v++)
  {   if  (g_mY[v][2] == 1 )
  {g_mY[v][0]=w[0][0]+mu[v][0]+sigma*valorZ[v][0];}

  if  (g_mY[v][2] == 2 )
  {g_mY[v][0]=w[1][0]+mu[v][0]+sigma*valorZ[v][0];}

  if  (g_mY[v][2] == 3 )
  {g_mY[v][0]=w[2][0]+mu[v][0]+sigma*valorZ[v][0];}

  if  (g_mY[v][2] == 4 )
  {g_mY[v][0]=w[3][0]+mu[v][0]+sigma*valorZ[v][0];} }
g_mY[] [0]=exp(g_mY[] [0]);

decl m; decl scensur; scensur=0; s_mC1=100*ranu(nc,1);

for (m=0; m<30; m++)
  { if  (s_mC1[m][0]-g_mY[m][0] < 0 && scensur < 3)
  { g_mY[m][0]= s_mC1[m][0];
  g_mY[m][1]=0;
  scensur=scensur+1; }
  }

  for (m=30; m<60; m++)
  { if  (s_mC1[m][0]-g_mY[m][0] < 0 && scensur < 6)
  { g_mY[m][0]= s_mC1[m][0];
  g_mY[m][1]=0;
  scensur=scensur+1; }
  }

  for (m=60; m<90; m++)
  { if  (s_mC1[m][0]-g_mY[m][0] < 0 && scensur < 9)
  { g_mY[m][0]= s_mC1[m][0];
  g_mY[m][1]=0;
  scensur=scensur+1; }
  }

  for (m=90; m<120; m++)
  { if  (s_mC1[m][0]-g_mY[m][0] < 0 && scensur < 12)
  { g_mY[m][0]= s_mC1[m][0];

```

```

    g_mY[m][1]=0;
    scensur=scensur+1; }
}

print(" Dados simulados \n", g_mY);

decl m_p=1001;
// Gerando efeito aleatório ... (Metropolis)
decl dparte1,dparte2,dparte3,dparte4,dfunc,dfunc1;

decl vP=<0.25;0.5;1.5;1.5>;

decl vLo=<0;0;-.Inf;-.Inf>; decl vHi=<+.Inf;+.Inf;+.Inf;+.Inf>;

decl vP1=<3>; decl vLo1=<0>; decl vHi1=<+.Inf>; decl inter=1;

decl L=20000; //número de iterações
decl estimativas=ones(5,L);

decl verossimi=ones(1,L); decl criterio1=ones(1,L),
criterio2=ones(1,L), taxa=ones(4,1);

estimativas[0:3][0]=vP; estimativas[4][0]=vP1;

s_mC=ones(nc,1);
s_mWe1=zeros(L,1),s_mWe2=zeros(L,1),
s_mWe3=zeros(L,1),s_mWe4=zeros(L,1);
s_mWe1=zeros(1001,1),s_mWe2=zeros(1001,1),
s_mWe3=zeros(1001,1),s_mWe4=zeros(1001,1);

do //inter {

s_mC[][0]=g_mY[][0];

// gerando dados censurados para w_{p}
decl cont,t,xx0,xx1,xx2,unif,gy,gt,Ft,gerado;

for(cont=0;cont<nc;++cont) { t=(g_mY[cont][0]);
  xx0=g_mY[cont][3];
//inter xx1=g_mY[cont][4];    //x1

decl xbeta=xx0*estimativas[2][inter-1]+xx1*estimativas[3][inter-1];

```

```

decl
efal1=(rangamma(1,1,1/estimativas[4][inter-1],1/estimativas[4][inter-1]));
decl
efal2=(rangamma(1,1,1/estimativas[4][inter-1],1/estimativas[4][inter-1]));
decl
efal3=(rangamma(1,1,1/estimativas[4][inter-1],1/estimativas[4][inter-1]));
decl
efal4=(rangamma(1,1,1/estimativas[4][inter-1],1/estimativas[4][inter-1]));

if(g_mY[cont][2]==1 && g_mY[cont][1]==0)
  {unif=ranu(1,1);
  gt= ( ((1-unif)^(-1/estimativas[0][inter-1]))*
  (1+ (t/exp(xbeta+efal1))^(1/estimativas[1][inter-1]))-1)
  ^ (estimativas[1][inter-1]);
  gy=gt*exp(xbeta+efal1);
  s_mC[cont][0]=(gy);}

if(g_mY[cont][2]==2 && g_mY[cont][1]==0) {unif=ranu(1,1);
gt= ( ((1-unif)^(-1/estimativas[0][inter-1]))*
(1+ (t/exp(xbeta+efal2))^(1/estimativas[1][inter-1]))-1)
^ (estimativas[1][inter-1]);
gy=gt*exp(xbeta+efal2);
s_mC[cont][0]=(gy);}

if(g_mY[cont][2]==3 && g_mY[cont][1]==0)
  {unif=ranu(1,1);
  gt= ( ((1-unif)^(-1/estimativas[0][inter-1]))*
  (1+ (t/exp(xbeta+efal3))^(1/estimativas[1][inter-1]))-1)
  ^ (estimativas[1][inter-1]);
  gy=gt*exp(xbeta+efal3);
  s_mC[cont][0]=(gy);}

if(g_mY[cont][2]==4 && g_mY[cont][1]==0)
  {unif=ranu(1,1);
  gt= ( ((1-unif)^(-1/estimativas[0][inter-1]))*
  (1+ (t/exp(xbeta+efal4))^(1/estimativas[1][inter-1]))-1)
  ^ (estimativas[1][inter-1]);
  gy=gt*exp(xbeta+efal4);
  s_mC[cont][0]=(gy);}
} //for

g_mY1=s_mC[][0]~g_mY[][1]~g_mY[][2]~g_mY[][3]~g_mY[][4];

```

```

decl burnin = L*0.1; // iteracoes de burnin decl i; decl p_acep;
decl j1=0,j2=0,j3=0,j4=0; //para calcular taxa de aceitacao
i=1;

do { decl w1=(rangamma(1,1,1/vP1, 1/vP1));

decl w2=(rangamma(1,1,1/vP1, 1/vP1));

decl w3=(rangamma(1,1,1/vP1,1/vP1));

decl w4=(rangamma(1,1,1/vP1, 1/vP1));

verossi1(w1,vP,&dparte1); decl num1=dparte1;
verossi1(s_mWe11[i-1][0],vP,&dparte1); decl den1=dparte1;
verossi2(w2,vP,&dparte2); decl num2=dparte2;
verossi2(s_mWe21[i-1][0],vP,&dparte2); decl den2=dparte2;
verossi3(w3,vP,&dparte3); decl num3=dparte3;
verossi3(s_mWe31[i-1][0],vP,&dparte3); decl den3=dparte3;
verossi4(w4,vP,&dparte4); decl num4=dparte4;
verossi4(s_mWe41[i-1][0],vP,&dparte4); decl den4=dparte4;

// probabilidade de aceitacao do Metropolis
p_acep = min(1, (num1/den1) ); if ( p_acep > ranu(1,1) )
    {s_mWe11[i][0] = w1;
    ++j1; }
else {s_mWe11[i][0] = s_mWe11[i-1][0];}

p_acep = min(1, (num2/den2)); if ( p_acep > ranu(1,1) )
    {s_mWe21[i][0] = w2;
    ++j2; }
else {s_mWe21[i][0] = s_mWe21[i-1][0];}

p_acep = min(1, (num3/den3) ); if ( p_acep > ranu(1,1) )
    {s_mWe31[i][0] = w3;
    ++j3; }
else {s_mWe31[i][0] = s_mWe31[i-1][0];}

p_acep = min(1, (num4/den4) ); if ( p_acep > ranu(1,1) )
    {s_mWe41[i][0] = w4;
    ++j4; }
else {s_mWe41[i][0] = s_mWe41[i-1][0];}

++i;} while (i<L);

```

```

taxa[0][0]=j1/L; taxa[1][0]=j2/L; taxa[2][0]=j3/L; taxa[3][0]=j4/L;

s_mWe1=s_mWe11[range(burnin-1,L,18)];
savemat("w1G_4G_TESTE_TRUNC.xls",s_mWe1 , {"w1"});
savemat("w11G_4G_TESTE_TRUNC.xls",s_mWe11 , {"w11"});
s_mWe2=s_mWe21[range(burnin-1,L,18)];
savemat("w2G_4G_TESTE_TRUNC.xls",s_mWe2, {"w2"});
savemat("w21G_4G_TESTE_TRUNC.xls",s_mWe21 , {"w21"});
s_mWe3=s_mWe31[range(burnin-1,L,18)];
savemat("w3G_4G_TESTE_TRUNC.xls", s_mWe3, {"w3"});
savemat("w31G_4G_TESTE_TRUNC.xls",s_mWe31 , {"w31"});
s_mWe4=s_mWe41[range(burnin-1,L,18)];
savemat("w4G_4G_TESTE_TRUNC.xls",s_mWe4 , {"w4"});
savemat("w41G_4G_TESTE_TRUNC.xls", s_mWe41, {"w41"});

println("taxa de aceitacao = ", taxa);

log_vero(estimativas[0:3][inter-1],&dfunc,!0,0); println("vero="
,dfunc);

log_vero1(estimativas[4][inter-1],&dfunc1,0,0); println("vero1="
,dfunc1);

verossimi[0][0]=dfunc+dfunc1; println("vero total="
,verossimi[0][0]);

MaxControl(-1,20); decl mhess=0.00001*unit(4); decl
mhess1=0.00001*unit(1); decl ir,ir1, var, var1,vscore1,vscore;

ir=MaxSQP(log_vero, &vP,&dfunc,&mhess,1, 0,0,vLo, vHi);
ir1=MaxSQP(log_vero1, &vP1,&dfunc1,&mhess1,1, 0,0,vLo1, vHi1);

if (ir != MAX_CONV) break; if (ir1 != MAX_CONV) break;

Num1Derivative(log_vero,vP,&vscore);
Num1Derivative(log_vero1,vP1,&vscore1);

print(" os valores da mhess \n",mhess); print(" os valores da mhess
1 \n",mhess1); println("vero=" ,dfunc); println("vero1=" ,dfunc1);
verossimi[][inter]=dfunc+dfunc1; println("vero total="
,verossimi[][inter]);

print("função escore \n", vscore); print("função escore 1\n",

```

```

vscore1);

vP[0][0]=fabs(vP[0][0]); vP[1][0]=fabs(vP[1][0]);
vP1[0][0]=fabs(vP1[0][0]);

estimativas[0:3][inter]=vP; estimativas[4][inter]=vP1;
println("estimativas",vP); println("estimativas",vP1);

//criterio de convergência

decl
maximo=(fabs(estimativas[][inter]-estimativas[][inter-1])./(
(fabs(estimativas[][inter-1])+0.001));
println ("maximo",maximo); criterio1[0][inter]=maxc(maximo); println
("criterio",criterio1[0][inter]);

s_mWe11[0][0]=s_mWe1[m_p-1][0]; s_mWe21[0][0]=s_mWe2[m_p-1][0];
s_mWe31[0][0]=s_mWe3[m_p-1][0]; s_mWe41[0][0]=s_mWe4[m_p-1][0];

++inter;

} while (inter<=100);

decl vscore2,dfunc2; decl mhess2=0.00001*unit(4);
println("estimativas",estimativas[][0:inter-1]); println ("criterio
1 = ",criterio1[0][1:inter-1]);

//Gráficos

DrawMatrix(0, verossimi[0][0:inter-1],{"log-vero"},1,1,2);
SaveDrawWindow("veroG_4G_TESTE_TRUNC.ps");
ShowDrawWindow();
CloseDrawWindow();

DrawMatrix(0, criterio1[0][1:inter-1],{"Criterio"},1,1,2);
SaveDrawWindow("criterio1G_4G_TESTE_TRUNC.ps");
ShowDrawWindow();
CloseDrawWindow();

DrawMatrix(0, estimativas[0:1][1:inter-1],{"k","sigma"},1,1,2);
SaveDrawWindow("est1G_4G_TESTE_TRUNC.ps");
ShowDrawWindow();
CloseDrawWindow();

```

```

DrawMatrix(0, estimativas[2][1:inter-1],{"beta0"},1,1,2);
  SaveDrawWindow("est2G_4G_TESTE_TRUNC.ps");
  ShowDrawWindow();
  CloseDrawWindow();

DrawMatrix(0, estimativas[3:][1:inter-1],{"beta1"},1,1,2);
  SaveDrawWindow("est3G_4G_TESTE_TRUNC.ps");
  ShowDrawWindow();
  CloseDrawWindow();

  DrawMatrix(0, s_mWe21[][0]',{"w2"},1000,18,0);
  SaveDrawWindow("grafW2G_4G_TESTE.ps");
  DrawCorrelogram(1, s_mWe21[][0]',{"ACF w2"},10);
  SaveDrawWindow("correlW2G_4G_TESTE_TRUNC.ps");
  CloseDrawWindow();

// Tempo duracao:
  println("\n tempo gasto: ", timespan(time), "\n");
}

```

Apêndice K - Programa no SAS usando proc NLMIXED para estimar os parâmetros do modelo de regressão log-Burr XII com efeito aleatório gama - dados simulados

```

data simulado;
input tempo  delta  grupo inter x1 ;
logtempo=log(tempo);
datalines;
      65.696      0.00000      1.0000      1.0000      0.44151
      33.387      0.00000      1.0000      1.0000      0.33195
      11.525      0.00000      1.0000      1.0000      0.10067
...
      355.53      1.0000      4.0000      1.0000      0.15445
      157.33      1.0000      4.0000      1.0000      0.23864
      41.542      1.0000      4.0000      1.0000      0.11473
run;

proc print data=simulado; run;

proc nlmixed data=simulado methods=GAUSS NOAD fd qpoints=10;

```

```

parms b0=1.5 b1=1.5 sigma=0.5 k=0.25 teta=3; bounds sigma, k>0;

bounds teta>0; p_i=CDF('NORMAL',u); if (p_i>0.999999) then
p_i=0.999999 ;

g_i2=quantile('GAMMA',p_i,1/teta);

g_i=teta*g_i2; eta=b0*1+b1*x1 +(g_i); z=(logtempo-eta)/sigma;
logp=delta*(log(k/sigma)+z-(k+1)*log(1+exp(z)))-(1-delta)*(k)*log(1+exp(z));
model logtempo ~general(logp); random u ~ normal(0,1) subject=grupo;
run;

```

Apêndice L - Programa na linguagem de programação matricial Ox para simular dados da distribuição beta Weibull modificada

```

//Programa para simular amostra da distribuição Beta Weibull modificada//
#include <oxstd.h> #include <oxprob.h> #import <solvenle>
#include<packages/gnudraw/gnudraw.h> #include <oxdraw.h> static decl
V;
// amostra gerada com dist. beta
amostra_BMW(const avF, const vX) { decl gamma, lambda, alpha;
lambda=1.5; gamma=3; alpha=0.5;
//Equacao nao-linear de beta
avF[0]=(vX[0]^gamma)*exp(lambda*vX[0])+(1/alpha)*log(1-V); return
1;}

main() {
// Gerando vetor aleatorio da distribuicao Beta(a,b)
// Gerador de numeros aleatorios
ranseed({1111,1111}); // semente do gerador decl a,b,n; a=1.5;b=0.5;
n=10000; //tamanho da amostra gerada decl cont; decl
tempo=ones(n,1); for(cont=0;cont<n;++cont) { V = ranbeta(1,1,a,b);
decl x=<1>; SolveNLE(amostra_BMW, &x); tempo[cont][0]=x; }
println("amostras geradas \n", tempo);
savemat("amostra.txt",tempo,{"amostra"}) ; }

```


Apêndice M - Elementos da matriz de informação observada $L(\boldsymbol{\theta})$ para os parâmetros $(a, b, \alpha, \gamma, \lambda)$ da distribuição beta Weibull modificada

$$L_{aa}(\boldsymbol{\theta}) = -n \left\{ \frac{\Gamma''(a)}{\Gamma(a)} - \psi^2(a) - \Gamma(a+b) \left[\psi'(a+b) + \psi^2(a+b) \right]^2 + \psi^2(a+b) \right\},$$

$$L_{bb}(\boldsymbol{\theta}) = -n \left\{ \frac{\Gamma''(b)}{\Gamma(b)} - \psi^2(b) - \Gamma(a+b) \left[\psi'(a+b) + \psi^2(a+b) \right]^2 + \psi^2(a+b) \right\},$$

$$L_{\alpha\alpha}(\boldsymbol{\theta}) = - \sum_{i=1}^n \left[\frac{x_i^{2\gamma} \exp(2\lambda x_i)}{v_i^2} \right] - (a-1) \sum_{i=1}^n \left\{ \frac{\exp(-v_i) x_i^{2\gamma} \exp(2\lambda x_i)}{[1 - \exp(-v_i)]^2} \right\},$$

$$L_{\gamma\gamma}(\boldsymbol{\theta}) = - \sum_{i=1}^n (\gamma + \lambda x_i)^{-2} + (a-1) \sum_{i=1}^n y_i [\log(x_i)]^2 - b \sum_{i=1}^n v_i [\log(x_i)]^2,$$

$$L_{\lambda\lambda}(\boldsymbol{\theta}) = - \sum_{i=1}^n \left(\frac{x_i}{\gamma + \lambda x_i} \right)^2 + (a-1) \sum_{i=1}^n y_i x_i^2 - b \sum_{i=1}^n v_i x_i^2,$$

$$L_{\alpha\gamma}(\boldsymbol{\theta}) = (a-1) \sum_{i=1}^n y_i v_i^{-1} x_i^\gamma \exp(\lambda x_i) \log(x_i) - b \sum_{i=1}^n x_i^\gamma \exp(\lambda x_i) \log(x_i),$$

$$L_{\alpha\lambda}(\boldsymbol{\theta}) = (a-1) \sum_{i=1}^n y_i v_i^{-1} \exp(\lambda x_i) x_i^{\gamma+1} - b \sum_{i=1}^n x_i^{\gamma+1} \exp(\lambda x_i),$$

$$L_{\gamma\lambda}(\boldsymbol{\theta}) = (a-1) \sum_{i=1}^n y_i x_i \log(x_i) - b \sum_{i=1}^n v_i x_i \log(x_i), \quad L_{a\lambda}(\boldsymbol{\theta}) = \sum_{i=1}^n \frac{\exp(-v_i) v_i x_i}{[1 - \exp(-v_i)]},$$

$$L_{a\alpha}(\boldsymbol{\theta}) = \sum_{i=1}^n \frac{\exp(-v_i) x_i^\gamma \exp(\lambda x_i)}{[1 - \exp(-v_i)]}, \quad L_{a\gamma}(\boldsymbol{\theta}) = \sum_{i=1}^n \frac{\exp(-v_i) v_i \log(x_i)}{[1 - \exp(-v_i)]}, \quad L_{ab}(\boldsymbol{\theta}) = -n\psi'(a+b),$$

$$L_{b\alpha}(\boldsymbol{\theta}) = - \sum_{i=1}^n x_i^\gamma \exp(\lambda x_i), \quad L_{b\gamma}(\boldsymbol{\theta}) = - \sum_{i=1}^n v_i \log(x_i), \quad L_{b\lambda}(\boldsymbol{\theta}) = - \sum_{i=1}^n v_i x_i,$$

em que $y_i = \frac{\exp(-v_i)v_i}{1-\exp(-v_i)} \left[1 - \frac{v_i}{1-\exp(-v_i)} \right]$ é outra observação transformada, $\Gamma''(\cdot)$ é a derivada de segunda ordem da função gama e $\psi'(\cdot)$ é a derivada de primeira ordem da função digama.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)