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

Modelo linear beta Weibull generalizado: propriedades,
estimação, diagnóstico e aplicações

Tiago Viana Flor de Santana

Tese apresentada para obtenção do título de Doutor em
Ciências. Área de concentração: Estatística e Experimen-
tação Agronômica

Piracicaba
2016

Tiago Viana Flor de Santana
Matemático

**Modelo linear beta Weibull generalizado: propriedades,
estimação, diagnóstico e aplicações**
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Prof^o Dr^o **EDWIN MOISÉS MARCOS ORTEGA**

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

Piracicaba
2016

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

Santana, Tiago Viana Flor de

Modelo linear beta Weibull generalizado: propriedades, estimação, diagnóstico e aplicações / Tiago Viana Flor de Santana. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2016.
112 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz".

1. Modelo linear generalizado 2. Modelo linear Weibull exponenciada generalizado
3. Modelo linear Beta Weibull generalizado 4. Influência local 5. Análise de resíduo
6. Função quantílica I. Título

CDD 519.532
S232m

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

DEDICO

Ao meu pai,

Levi Flor de Santana,

pelas orações, incentivo e esforço que fez para que eu pudesse continuar em frente.

AGRADEÇO

A **Deus**, pela força, sustento e por mais esta benção que me tem concedido.

Ao meu pai **Levi Flor de Santana**, a minha mãe **Valnice Aparecida dos Santos de Santana**, a minha irmã **Talita Flor de Santana** e a minha esposa **Ione Viana Flor de Santana** pelas orações, incentivo, amor e apoio incondicional.

Ao meu filho **Tiago Viana Flor de Santana Filho** pelos sorrisos encorajadores.

Ao meu orientador, Prof. Dr. Edwin Moises Marcos Ortega, pela orientação, confiança e exemplo de profissional.

Ao Prof. Dr. Adriano Kamimura Suzuki do Instituto de Ciências Matemáticas e de Computação da USP de São Carlos pela disponibilidade e contribuições na parte Bayesiana.

Agradeço especialmente aos meus amigos Ana Júlia Righetto, Tiago Gentil Ramires, Lucas Santana da Cunha, Rodrigo Pescim, Altemir da Silva Braga, Natália da Silva Martins, Alessandra dos Santos, Iábita Fabiana Souza, Kuang Hongyu, Elisângela Oliveira, José Nilton Cruz e Viviane Paulenas pelo compartilhamento de conhecimento, dicas e companherismo.

Aos professores e funcionários do Departamento de Ciências Exatas da ESALQ/USP, em especial as secretárias Solange de Assis Paes Sabadin, Luciane Brajão e Mayara Segatto, por estarem sempre prontas a ajudar e resolver nossos “*problemas*”.

A Coordenadoria para o Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro.

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

*“O temor do SENHOR é o princípio da ciência; os loucos
desprezam a sabedoria e a instrução”*

Provérbios 1:7

*“Pois, que adianta ao homem ganhar o mundo inteiro e perder a
salvação?”*

Marcos 8:36

SUMÁRIO

RESUMO	11
ABSTRACT	13
1 INTRODUÇÃO	15
Referências	18
2 MODELO LINEAR WEIBULL EXPONENCIADO GENERALIZADO: ESTI- MAÇÃO, DIAGNÓSTICO E ANÁLISE DE RESÍDUOS	19
Resumo	19
Abstract	19
2.1 Introdução	20
2.2 A distribuição Weibull exponenciada	21
2.2.1 Modelo linear Weibull exponenciado generalizado	24
2.2.2 Modelo de regressão	29
2.2.3 Estimação, testes e simulação	31
2.2.3.1 Estimação de máxima verossimilhança	31
2.2.3.2 Estudo de Simulação: Máxima verossimilhança	33
2.2.3.3 Estimação Bayesiana	34
2.2.3.4 Distribuições <i>a priori</i> e <i>a posteriori</i>	34
2.2.3.5 Critério de comparação de modelo	37
2.2.3.6 Estudo de simulação: Análise Bayesiana	38
2.2.4 Análise de sensibilidade	39
2.2.4.1 Influência local	40
2.2.5 Diagnóstico de casos influentes Bayesiano	43
2.2.5.1 Influência de observações <i>outlying</i>	45
2.2.6 Análise de resíduos	46
2.2.7 Aplicação	50
2.2.7.1 Valores pagos por danos com acidentes pessoais	51
2.2.7.2 Estimação de máxima verossimilhança	53
2.2.7.3 Análise Bayesiana	56
2.3 Conclusões	61
Referências	62

3 MODELO LINEAR BETA WEIBULL GENERALIZADO: ESTIMAÇÃO, DI- AGNÓSTICO E ANÁLISE DE RESÍDUOS	69
Resumo	69
Abstract	69
3.1 Introdução	70
3.2 A distribuição beta Weibull	72
3.2.1 Modelo linear beta Weibull generalizado	78
3.2.1.1 Modelo de regressão	83
3.2.1.2 Estimação de máxima verossimilhança	86
3.2.2 Estudo de Simulação	89
3.2.3 Análise de sensibilidade	91
3.2.3.1 Influência local	92
3.2.4 Análise de resíduos	95
3.2.5 Aplicação	97
3.3 Tempo de falha de filmes PET	99
3.4 Conclusões	102
Referências	103
4 PESQUISAS FUTURAS	107
5 APÊNDICE	109

RESUMO

Modelo linear beta Weibull generalizado: propriedades, estimação, diagnóstico e aplicações

Neste trabalho dois novos modelos estatísticos de regressão são propostos, com estrutura muito semelhante aos Modelos Lineares Generalizados (MLG) porém, admitindo as distribuições Weibull exponenciada (WE) e beta Weibull (BW) para o componente aleatório as quais não pertencem a família exponencial como é requerido em MLG. Os novos modelos trazem uma nova abordagem para as distribuições admitidas em modelos de regressão e estende o MLG para além da família exponencial. Os modelos, nomeados por Modelo Linear Weibull Exponenciada Generalizado (MLWEG) e Modelo Linear Beta Weibull Generalizado (MLBWG), possuem como caso particular o modelo Exponencial, pertencente a família de MLG, além de outros modelos que os MLGs não contemplam como, por exemplo: Weibull, WE, Exponencial Exponenciado (EE) entre outros. Além da função taxa de falha (ftf) constante da distribuição Exponencial, os novos modelos ajustam também formas monótonas e não monótonas da ftf. Quando se admite função de ligação logarítmica obtém-se o mesmo modelo de locação e escala, muito utilizado em análise de sobrevivência, sem a necessidade de transformação da variável resposta simplificando a modelagem e permitindo maior compreensão da influência das covariáveis na resposta. Método de estudo de observações influentes foi construído baseado na metodologia de influência local sobre três esquemas de perturbações: perturbação da verossimilhança, da variável resposta e das covariáveis e a análise de resíduo foi proposta a partir da função quantílica. Por fim, dois conjuntos de dados reais foram utilizados para ilustrar a aplicabilidade dos modelos propostos e seus resultados discutidos.

Palavras-chave: Modelo Linear Generalizado; Modelo Linear Weibull exponenciado Generalizado; Modelo Linear Beta Weibull Generalizado; Influência Local; Análise de Resíduo; Função quantílica

ABSTRACT

Generalized beta Weibull linear model: properties, estimation, diagnostics and applications

In this work two new statistical regression models are proposed, with very similar structure to Generalized Linear Models (GLM) but, assuming the exponentiated Weibull (EW) and beta Weibull (BW) distributions for the random component which do not belong to the exponential family as required in GLM. The new models bring a new approach to the distribution accepted in regression models and extend the GLM beyond of the exponential family. The models, named by Generalized Exponentiated Weibull Linear Model (GEWLM) and Generalized Beta Weibull Linear Model (GBWLM) have as a particular case the Exponential model, belonging to the family of GLM, and other models that GLMs do not include, for example : Weibull, EW, Exponentiated Exponential (EE) among others. Besides the failure rate function (frf) constant of Exponential distribution, the new models also model monotonous and not monotonous forms of frf. When it accepts logarithmic link function obtains the same location and scale model, widely used in the analysis of survival without the need to transform the response variable simplifying the modeling and allowing greater understanding of the influence of covariates on the response. Study of influential observations method was built based on the methodology of the local influence on three perturbations schemes: perturbation of the likelihood of the response variable and the covariates and residual analysis was proposed from the quantile function. Finally, two sets of real data are used to illustrate the applicability of the models proposed and results discussed.

Keywords: Generalized Linear Model; Generalized Exponentiated Weibull Linear Model; Generalized Beta Weibull Linear Model; Local Influence; Residual analysis; Quantile Function

1 INTRODUÇÃO

A distribuição Weibull foi proposta por Waloddi Weibull em 1939 e desde então vem sendo amplamente utilizada em estudos biomédicos e industriais (Colosimo e Giolo, 2006), bem como na análise de dados de tempo de vida e para a resolução de uma variedade de problemas em diversas áreas (Prudente e Cordeiro, 2010). Segundo Collet (2003), ela é tão importante para análise paramétrica em dados de sobrevivência quanto a distribuição normal é para modelos lineares. Sua grande popularidade em aplicações práticas deve-se a capacidade de acomodar uma grande variedade de formas, todas com uma propriedade básica: a sua função taxa de falha (ftf) é monótona, ou seja, apresenta formas crescente, decrescente e constante (Colosimo e Giolo, 2006).

Entretanto, essa mesma característica limita sua aplicação em situações que apresentam ftf não monótona, tais como ftf unimodal e em forma de U (ou de banheira). Formas unimodais e de banheira são frequentes em estudos de confiabilidade, ciências biológicas e na área de engenharias conforme afirmam Carrasco, Ortega e Cordeiro (2008) e Mudholkar e Srivastava (1995). Devido a restrição da distribuição Weibull na análise de dados com ftf não monótona Mudholkar e Srivastava propuseram a distribuição Weibull exponenciada (WE), que ajusta-se a conjunto de dados cuja ftf é monótona e não monótona e tem como submodelos a distribuição Weibull, exponencial, Rayleigh, Burr tipo X, exponencial exponenciada (EE), entre outras distribuições (Mudholkar e Srivastava, 1993 e 1995). Posteriormente Famoye, Lee e Olumolade (2005) propuseram uma distribuição mais geral que a WE e denominaram beta Weibull (BW). A BW também ajusta-se a dados com ftf monótona e não monótona, tem como submodelo a WE e se mostra mais adequada, em alguns casos, que a WE conforme apresentado no trabalho de Famoye, Lee e Olumolade (2005).

Em 2010, Prudente e Cordeiro propuseram uma nova classe de modelos de regressão seguindo os mesmos conceitos dos modelos lineares generalizados (MLG) porém, admitindo a distribuição Weibull na análise dos dados. O modelo linear Weibull generalizado (MLWG), como foi nomeado pelos autores, permite um novo uso da distribuição Weibull em uma estrutura muito semelhante aos modelos lineares generalizados. O modelo proposto foi definido como uma síntese de três partes estruturais importantes em MLG: o componente aleatório para a variável resposta seguindo distribuição Weibull parametrizada de forma conveniente, um componente sistemático especificado por uma

função linear que modela parte relevante da média da distribuição, chamada de preditor linear do modelo, e uma função de ligação, ligando o componente aleatório e o sistemático. Os MLWGs são muitos efetivos, segundo os autores, na modelagem de uma parte da resposta média em termos de uma equação de regressão para dados positivos contínuos e na definição do resíduo. E muitas outras medidas de diagnóstico seguem de perto a teoria dos MLGs. Adicionalmente Vanegas, Rondón e Cordeiro (2013), desenvolveram ferramentas de diagnóstico e validação para essa classe de modelos de regressão e realizaram análise de influência local e global para os esquemas de perturbação da verossimilhança, da resposta e das covariáveis.

Devido a importância notória das distribuições Weibull, beta Weibull e Weibull exponenciada na análise de dados de tempo de vida e baseado nos trabalhos de Prudente e Cordeiro (2010) e Vanegas, Rondón e Cordeiro (2013), propõem-se neste trabalho uma classe de modelos de regressão baseado na estrutura dos MLGs e na distribuição beta Weibull e Weibull exponenciada. De forma a ser um modelo competitivo ao já conhecido MLG, em particular os modelos gama e gaussiano inverso, e tendo como caso particular o modelo linear Weibull generalizado proposto por Prudente e Cordeiro (2010).

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 faz-se necessário o uso de medidas de influência baseadas em pequenas perturbações nos dados ou no modelo (Cook, 1986), chamada de influência local. O método de influência local tem se constituído em uma ferramenta bastante popular na análise da influência conjunta das observações nos resultados de um ajuste de regressão. Por outro lado, a validação desse ajuste deve passar também pela análise de resíduos em que se procura detectar pontos mal ajustados (aberrantes), bem como verificar se há indícios de afastamentos sérios das suposições feitas para o erro do modelo.

Com esse objetivo, propõem-se ainda o estudo do resíduo desenvolvido a partir da função quantílica, denominado resíduo quantil e a construção das bandas de confiança obtidas a partir da geração de envelopes (Atkinson, 1985) por meio da simulação de Monte Carlo como uma medida adicional para verificar a qualidade de ajuste do modelo

de regressão.

Nos dois próximos capítulos os modelos lineares generalizados Weibull exponenciado e beta Weibull serão desenvolvidos. No capítulo 2 é apresentado o modelo linear Weibull exponenciado generalizado (MLWEG) que possui três parâmetros, $(\lambda, \phi$ e $\delta)$, para o componente aleatório em que λ é um parâmetro de escala que também influencia na média e na variância da distribuição e conecta o componente aleatório, descrito pela distribuição Weibull exponenciada, ao componente sistemático, indicado pela equação $\eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$ em que $\mathbf{x}_i$ é o vetor de covariáveis associado a i -ésima observação e $\boldsymbol{\beta}$ é um vetor de parâmetros desconhecido, por meio de uma função de ligação $g(\lambda_i) = \eta_i$.

O modelo linear Beta Weibull generalizado (MLBWG) é apresentado e discutido no capítulo 3, onde um estudo sobre a influência dos quatros parâmetros da distribuição beta Weibull no comportamento da função densidade de probabilidade e função taxa de falha é realizado. As expressões para o logaritmo da função de verossimilhança, função score e matriz de informação de Fisher esperada e observada são apresentadas. A metodologia de influência local sobre três esquemas de perturbação: da variável resposta, da verossimilhança e das covariáveis é desenvolvida. E por fim, para verificar se existe afastamento nas suposições feitas ao modelo é proposto um resíduo desenvolvido a partir da função quantílica e realizado um estudo por meio da simulação de Monte Carlo para construção de envelopes simulados.

- ATKINSON, A. C. **Plots, transformations and regressions**. Oxford: Statistical Sciences Series Oxford, 1985. 296 p.
- CARRASCO, J. M. F.; ORTEGA, E. M. M.; CORDEIRO, G. M. A generalized modified Weibull distribution for lifetime modeling. **Computational Statistics and Data Analysis**, Oveido, n. 53, p. 450-462, 2008.
- COLLET, A. **Modelling Survival Data in Medical Research**. London: Chapman and Hall, 2003. 410 p.
- COLOSIMO, E. A.; GIOLO, S. R. **Análise de Sobrevivência aplicada**. São Paulo: Edgard Blücher, 2006. 392 p.
- COOK, R. D. Assessment of local influence (with discussion). **Journal of the Royal Statistical Society**, Minneapolis, v. 48, p. 133-169, 1986.
- FAMOYE, F.; LEE, C., OLUMOLADE, O. The Beta-Weibull Distribution. **Journal of Statistical Theory and Applications**, Lawrenceville, v. 4, p. 121-136, 2005.
- MUDHOLKAR, G.S.; SRIVASTAVA, D.K. Exponentiated Weibull family for analyzing bathtub failure-ratedata. **IEEE Transactions on Reliability**, Dallas, v. 42, p. 299-302, 1993.
- MUDHOLKAR, G.S.; SRIVASTAVA, D.K.; FREIMER, M. The exponentiated Weibull family; a reanalysis of the bus motor failure data. **Technometrics**, Toronto, v. 37, p. 436-445, 1995.
- PRUDENTE, A. A.; CORDEIRO, G. M. Generalized weibull linear models. **Communications in Statistics Theory and Methods**, Philadelphia, v. 39, p. 3739-3755, 2010.
- VANEGAS, L. H.; RONDÓN, L. M.; CORDEIRO, G. M., Diagnostic tools in generalized Weibull linear regression models. **Journal of Statistical Computation and Simulation**, Toronto, v. 83, p. 1-24, 2013.

2 MODELO LINEAR WEIBULL EXPONENCIADO GENERALIZADO: ESTIMAÇÃO, DIAGNÓSTICO E ANÁLISE DE RESÍDUOS

Resumo

Um novo modelo de regressão baseado na distribuição Weibull exponenciada (WE) (com um parâmetro de escala λ e dois parâmetros de forma ϕ e δ , denotada por $WE(\lambda, \phi, \delta)$) e na estrutura dos Modelos Lineares Generalizados (MLG) denominado Modelo Linear Weibull exponenciado Generalizado (MLWEG) é proposto. O MLWEG, da mesma forma que os MLGs, é composto por três partes estruturais importantes: o componente aleatório, caracterizado pela distribuição da variável resposta, o componente sistemático que incluem as variáveis explanatórias no modelo por meio de uma estrutura linear, chamada de preditor linear e uma função de ligação que liga a parte sistemática do modelo a parte aleatória. O MLWEG tem como submodelo o modelo linear Weibull generalizado (MLWG) proposto por Prudente e Cordeiro (2010), quando fixado o parâmetro $\delta = 1$. Além de ser mais flexível do que o MLWG, ele é capaz de modelar dados com funções taxa de falha monótonas e não monótonas. As expressões para o logaritmo da função de verossimilhança, o vetor escore para β , ϕ e δ foram desenvolvidos matricialmente e as matrizes de informação observada e esperada são apresentadas. O método de máxima verossimilhança e o procedimento Bayesiano foram adotados para estimar os parâmetros do modelo. Para detectar observações influentes no modelo proposto são usadas medidas de influência baseadas na influência local e Bayesiana para diagnosticar casos influentes. Também, mostra-se que as estimativas do MLWEG são robustas na presença de *outliers* nos dados. Adicionalmente, para verificar se existe afastamento das suposições feitas ao modelo, detectar pontos atípicos e como uma medida para verificar a qualidade de ajuste do modelo de regressão é proposto um resíduo desenvolvido a partir da função quantílica e realizado um estudo por meio da simulação de Monte Carlo para a construção das bandas de confiança construídas a partir da geração de envelopes (Atkinson, 1985). Finalmente, os resultados são aplicados a um conjunto de dados reais na área de seguro.

Palavras-chave: Modelo de regressão Weibull exponenciado; Modelos lineares generalizados; Influência local e global; Análise de resíduos

Abstract

A new regression model based on the exponentiated Weibull distribution and the structure of generalized linear models (GLMs), called the generalized exponentiated Weibull linear model (GEWLM), is proposed. The GEWLM, like the GLMs, is composed of three important structural parts: the random component, characterized by the distribution of the response variable; the systematic component, which includes the explanatory variables in the model by means of a linear structure; and a link function, which connects the systematic and random parts of the model. Explicit expressions for the logarithm of the likelihood function, score vector and observed and expected information matrices are presented. The method of maximum likelihood and a Bayesian procedure are adopted for estimating the model parameters. To detect influential observations in the new model, we use diagnostic measures based on the local

influence and Bayesian case influence diagnostics. Also, we show that the estimates of the GEWLM are robust to deal with the presence of outliers in data. Additionally, to check whether the model supports its assumptions, to detect atypical points and to verify the goodness of fit of the regression model, we propose a residuals based on the quantile function, and perform a Monte Carlo simulation study to construct the confidence bands from the generated envelopes. We apply the model to a dataset from the insurance area.

Keywords: Generalized exponentiated Weibull linear model; Generalized linear model; Local influence; Residual

2.1 Introdução

Desde que foi proposta, a distribuição WE têm sido bastante aplicada nas mais diversas áreas do conhecimento, tanto práticas quanto teóricas. Pesquisadores na área de ciência da computação, estudos clínicos, hidrologia, silvicultura, estatística, atuária, confiabilidade, ecologia e engenharia mecânica têm utilizado com êxito essa distribuição. Como exemplo de aplicação da WE, pode-se citar os trabalhos de Muldholkar e Hutson (1996) que modelaram dados de inundação do rio Floyd no estado americano de Iowa, Ahmad et al. (2006) envolvendo testes de vida acelerados, Surles e D'Ambrosio (2004) em aplicações com compostos de fibra de carbono, Wang e Rennolls (2005) que modelaram o diâmetro de árvores em plantações de pinheiros chineses, Zhang et al. (2005) no estudo de falha de sistemas de firmware, Chaubey (2007) na distribuição de dados de seguro com excesso de perda, Barghout (2009) que propôs um novo modelo de estatística de ordem para predição de confiabilidade para o tempo de detecção de falha, entre inúmeros outros autores. O artigo de Nadarajah, Cordeiro e Ortega (2013) apresenta uma extensa revisão da distribuição WE, detalha-se seu desenvolvimento, os resultados mais recentes e as principais aplicações dessa distribuição.

Devido a importância notória das distribuições Weibull e Weibull exponenciada na análise de dados de tempo de vida e baseado nos trabalhos de Prudente e Cordeiro (2010) e Vanegas, Rondón e Cordeiro (2013), propõem-se neste trabalho uma classe de modelos de regressão baseado na estrutura dos modelos lineares generalizados e na distribuição Weibull exponenciada. De forma a ser um modelo competitivo ao já conhecido modelo linear generalizado, em particular os modelos Gama e Gaussiano inverso, e tendo como caso particular o modelo linear Weibull generalizado proposto por Prudente e Cordeiro (2010).

Esse artigo está organizado da seguinte forma. Na seção 2.2 uma breve re-

visão da distribuição WE é realizada, destacando o ganho na adição do parâmetro δ na distribuição Weibull, o qual gera grande flexibilidade ao modelo WE e permite modelar funções taxa de falha com formas de U e unimodais. Na seção 2.2.1 o MLWEG é definido e suas propriedades apresentadas, o modelo de regressão é discutido na subseção 2.2.2 e o logaritmo da função de verossimilhança é apresentado na subseção 2.2.3.1, juntamente com as matrizes de informação observada e esperada. Um estudo de simulação é realizado na seção 2.2.3.2 utilizando a metodologia de Monte Carlo, para avaliar as estimativas de máxima verossimilhança do modelo MLWEG, considerando as funções de ligação identidade, recíproca e logística. Nas seções 2.2.3.3, 2.2.3.4, 2.2.3.5, e 2.2.3.6 a abordagem Bayesiana é adotada para o estudo de estimação dos parâmetros do modelo, discussão das distribuições *a priori* e *posteriori*, construção de critérios de comparação de modelos e estudo de simulação. A análise de sensibilidade é desenvolvida na seção 2.2.4 e subseção 2.2.4.1, nas quais destacam-se as expressões para o logaritmo da função de verossimilhança perturbada e a matriz Δ para os esquemas de perturbações da verossimilhança, da variável resposta e das covariáveis. Semelhantemente a metodologia Bayesiana é utilizada para analisar casos influentes nas seções 2.2.5 e 2.2.5.1. O estudo dos resíduos é discutido na seção 2.2.6, na qual o resíduo quantil é definido para MLWEG assim como o método de construção de envelope simulado, utilizado para avaliar o ajuste do modelo. Na seção 2.2.7 um conjunto de dados reais na área de seguro é utilizados para exemplificar o ajuste do MLWEG em contraste ao MLWG utilizando a abordagem clássica e Bayesiana e por fim na seção 2.3 as conclusões são discutidas.

2.2 A distribuição Weibull exponenciada

A distribuição WE foi proposta por (Mudholkar e Srivastava, 1993 e 1995). A distribuição WE é uma generalização da distribuição Weibull com dois parâmetros e sua construção é bastante simples. Se uma variável aleatória T , não negativa, tem distribuição Weibull com função distribuição acumulada

$$G(t; \alpha, \phi) = 1 - \exp \left[- \left(\frac{t}{\alpha} \right)^\phi \right] \quad \text{com } t > 0 \text{ e } \alpha, \phi > 0, \quad (1)$$

em que ϕ é o parâmetro de forma e α o parâmetro de escala, a distribuição acumulada da WE ($F(t)$) é obtida elevando $G(t)$ a um parâmetro $\delta > 0$ tal que $F(t) = G(t)^\delta$, obtendo-se

$$F(t; \alpha, \phi, \delta) = \left\{ 1 - \exp \left[- \left(\frac{t}{\alpha} \right)^\phi \right] \right\}^\delta \quad \text{com } t > 0 \text{ e } \alpha, \phi, \delta > 0, \quad (2)$$

em que α é o parâmetro de escala e ϕ e δ são os parâmetros de forma. A função densidade de probabilidade (fdp) da WE é obtida derivando-se (2) em relação a t , cuja expressão é

$$f(t) = \delta \phi t^{-1} \left(\frac{t}{\alpha} \right)^\phi \exp \left\{ - \left(\frac{t}{\alpha} \right)^\phi \right\} \left[1 - \exp \left\{ - \left(\frac{t}{\alpha} \right)^\phi \right\} \right]^{\delta-1}. \quad (3)$$

A adição do parâmetro δ trouxe uma grande flexibilidade à distribuição, permitindo a modelagem de uma grande variedade de formas para a fdp de acordo com os valores atribuídos para os parâmetros ϕ e δ . Se $0 < \phi\delta < 1$, então a fdp em (3) é monótona decrescente e ilimitada na sua calda esquerda, se $\phi\delta = 1$, então ela é monótona decrescente com a ordenada da cauda esquerda igual a $1/\alpha$ e se $\phi\delta > 1$, a fdp é unimodal com uma pequena cauda à esquerda. Esses três casos são ilustrados na Figura 1, fixando alguns valores para ϕ e δ e fazendo $\alpha = 2$. Nota-se ainda, as seguintes relações:

$$f(t) \sim \phi\delta\alpha^{-\phi\delta}t^{\phi\delta-1} \quad \text{quando } t \rightarrow 0$$

e

$$f(t) \sim \phi\delta\alpha^{-\phi}t^{\phi-1} \exp \left[- \left(\frac{t}{\alpha} \right)^\phi \right] \quad \text{quando } t \rightarrow \infty$$

Destaca-se aqui que a notação $a(t) \sim b(t)$, quando $t \rightarrow 0$, significa que $a(t)/b(t) \rightarrow 1$ e da mesma forma quando $t \rightarrow \infty$ significa que $a(t)/b(t) \rightarrow 1$. Portanto, as caudas à esquerda da fdp da WE comportam-se polinomialmente, enquanto que as caudas à direita se comportam exponencialmente. Consequentemente, se $t \rightarrow 0$ então $f(t) \rightarrow 0$ quando $\phi\delta > 1$, $f(t) = 1/\alpha$ quando $\phi\delta = 1$ e $f(t) \rightarrow \infty$ quando $0 < \phi\delta < 1$.

Outra característica importante da WE é que ela possui vários submodelos como casos particulares, por exemplo: para $\phi = 1$ e $\delta = 1$, obtém-se a distribuição exponencial como caso particular, fixando $\phi = 1$ obtém-se então a distribuição exponencial exponenciada (EE), para $\delta = 1$ têm-se a Weibull, se $\phi = 2$ a distribuição Burr tipo X é obtida como caso particular, para $\phi = 2$ e $\delta = 1$ em (3) resulta na fdp da Rayleigh entre

Figura 1 – (a) Formas da função densidade de probabilidade da WE, para $\phi\delta > 1$, $0 < \phi\delta < 1$, $\phi\delta = 1$ e $\alpha = 2$, (b) Formas da função taxa de falha da WE para alguns valores atribuídos para ϕ e δ

outros casos particulares.

O parâmetro δ não somente trouxe maior flexibilidade para a fdp, mas tornou possível o ajuste da distribuição WE a conjuntos de dados cuja a função taxa de falha têm forma de U, ou de banheira como também é conhecida, além de modelar todas as formas que a Weibull já modela.

A função taxa de falha de uma variável aleatória $T > 0$, é definida por $h(t) = f(t)/[1 - F(t)]$ em que f e F são as funções de densidade e acumulada, respectivamente. O gráfico da função taxa de falha apresenta algumas formas básicas conforme pode ser visualizado na Figura 2, com diferentes interpretações práticas. O gráfico pode ser crescente ou decrescente conforme indicado por $h_1(t)$ e $h_2(t)$ respectivamente, constante representado por $h_3(t)$, ou alternar entre intervalos decrescentes, constantes e crescentes, como por exemplo $h_4(t)$, que é uma forma da função taxa de falha conhecida por forma de U ou forma de banheira e $h_5(t)$ conhecida por forma unimodal, que inicia com um intervalo crescente até atingir um máximo e então decresce.

As formas da função taxa de falha para esta distribuição são obtidas de quatro regiões do espaço paramétrico determinadas pelas equações $\phi = 1$ e $\delta\phi = 1$ (Figura 2), tais que:

- i. Se $\phi > 1$ e $0 < \phi\delta < 1$ então obtém-se forma de U;
- ii. Se $0 < \phi < 1$ e $\phi\delta > 1$ então obtém-se forma unimodal;

Figura 2 – (a) Formas gerais para a função taxa de falha, (b) As quatro regiões do espaço paramétrico da WE determinadas pelas equações $\phi = 1$ e $\phi\delta = 1$ com as respectivas descrições das formas da função taxa de falha

- iii. Se $0 < \phi \leq 1$ e $0 < \phi\delta \leq 1$ então obtém-se forma decrescente;
- iv. Se $\phi \geq 1$ e $\phi\delta \geq 1$ então obtém-se forma crescente;
- v. Se $\phi = 1$ e $\delta = 1$ então obtém-se forma constante.

A Figura 1 ilustra as formas da função taxa de falha para alguns valores fixos de ϕ e δ , satisfazendo os 4 casos acima.

2.2.1 Modelo linear Weibull exponenciado generalizado

Prudente e Cordeiro (2010) publicaram um artigo intitulado *Generalized Weibull Linear Models* (Modelos lineares Weibull generalizado) onde propuseram uma nova classe de modelo de regressão com estrutura análoga aos modelos lineares generalizados, com base na distribuição Weibull, que não pertence a família exponencial.

O MLWG, como foi nomeado pelos autores, permite um novo uso da distribuição Weibull em uma estrutura muito semelhante aos modelos lineares generalizados. O modelo proposto foi definido como uma síntese de três componente estruturais importantes em MLG: o componente aleatório para a variável resposta seguindo distribuição Weibull parametrizada de forma conveniente, um componente sistemático especificado pelo preditor linear do modelo e por uma função adequada ligando uma parte relevante da média da distribuição e o componente sistemático.

Os MLWGs são muitos efetivos, como colocam os autores, na modelagem de uma parte da resposta média em termos de uma equação de regressão para dados positivos contínuos e a definição do resíduo e muitas outras medidas de diagnóstico podem seguir de perto a teoria dos MLG.

Sejam $Y_1, Y_2, \dots, Y_n$, n variáveis aleatórias independentes, com cada Y_i seguindo distribuição Weibull com parâmetro de forma ϕ e parâmetro de escala α_i o qual varia de acordo com as observações, com função distribuição acumulada (fda) e densidade dadas, respectivamente, por

$$F(y_i; \alpha_i, \phi) = 1 - \exp \left[- \left(\frac{y_i}{\alpha_i} \right)^\phi \right] \quad (4)$$

e

$$f(y_i; \alpha_i, \phi) = \phi y_i^{-1} \left(\frac{y_i}{\alpha_i} \right)^\phi \exp \left\{ - \left(\frac{y_i}{\alpha_i} \right)^\phi \right\} \quad \text{com } \phi, \alpha_i, y_i > 0, \quad (5)$$

com média e variância

$$E(Y_i) = \alpha_i \Gamma_1(\phi) \quad \text{e} \quad Var(Y_i) = \alpha_i^2 \{ \Gamma_2(\phi) - \Gamma_1^2(\phi) \}$$

em que $\Gamma_j(\phi) = \Gamma \left(1 + \frac{j}{\phi} \right)$ para $j \geq 1$, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ é a função Gama e $\Gamma'(x) = \partial \Gamma(x) / \partial x$.

Reparametrizando a equação (4) de acordo com a relação $\lambda_i = \alpha_i \exp \{ \phi^{-1} \Gamma'(2) \}$, proposta por Cox e Reid (1987), obtém-se ortogonalidade entre os parâmetros ϕ e λ_i , ou seja, $E(\partial^2 l(\boldsymbol{\theta}) / \partial \phi \partial \lambda_i) = E(\partial^2 l(\boldsymbol{\theta}) / \partial \lambda_i \partial \phi) = 0$. A vantagem em se ter parâmetros ortogonais consiste na independência assintótica das estimativas de máxima verossimilhança para ϕ e λ_i e na simplificação do desenvolvimento algébrico e computacional, uma vez que a matriz de informação esperada é diagonal. Dessa forma, as expressões reparametrizadas para a fda, fdp, esperança e variância são dadas, respectivamente, por

$$F(y_i; \lambda_i, \phi) = 1 - \exp \left[- \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \right], \quad (6)$$

$$f(y_i; \lambda_i, \phi) = \phi y_i^{-1} \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \exp \left\{ - \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \right\} \quad \text{com } \phi, \lambda_i, y_i > 0, \quad (7)$$

$$E(Y_i) = \lambda_i c_1(\phi) \quad \text{e} \quad Var(Y_i) = \lambda_i^2 c_2(\phi)$$

em que $c_1(\phi) = \exp\{-\Gamma'(2)/\phi\}\Gamma_1(\phi)$ e $c_2(\phi) = \exp\{-2\Gamma'(2)/\phi\}\{\Gamma_2(\phi) - \Gamma_1^2(\phi)\}$.

O MLWG é definido pela equação (7), a fdp da Weibull, como o componente aleatório e por uma função g aplicada em λ_i tal que $g(\lambda_i) = \eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$, como o componente sistemático, em que $\mathbf{x}_i$ é o vetor de covariáveis associado a i -ésima observação e $\boldsymbol{\beta}$ é um vetor de parâmetros desconhecido.

Elevando-se a equação (6) a potência $\delta > 0$ obtém-se a distribuição WE com um parâmetro de escala λ_i e dois parâmetros de forma ϕ e δ . Daqui em diante, a distribuição Weibull exponenciada reparametrizada será denotada simplesmente por WE e indicaremos que uma variável aleatória Y têm distribuição WE por $Y \sim WE(\lambda_i, \phi, \delta)$, cuja função distribuição acumulada e função densidade são definidas respectivamente por

$$F(y_i; \lambda_i, \phi, \delta) = \left\{ 1 - \exp \left[- \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \right] \right\}^\delta \quad (8)$$

e

$$f(y_i; \lambda_i, \phi, \delta) = \delta \phi y_i^{-1} \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \exp \left\{ - \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \right\} \left\{ 1 - \exp \left[- \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \right] \right\}^{\delta-1} \quad (9)$$

Note que se $\delta = 1$ e $\phi = 1$ obtém-se a distribuição exponencial, se $\phi = 1$ obtém-se a distribuição EE e se Y_i segue distribuição WE então a variável aleatória $A_i = (Y_i/\lambda_i)^\phi e^{\Gamma'(2)}$ segue distribuição EE com parâmetro de escala igual a um e parâmetro de forma δ . O teorema a seguir apresenta o resultado descrito de forma mais geral.

Teorema: Se $Y \sim WE(\lambda, \phi, \delta)$ e $Z = (Y/\lambda)^\phi e^{\Gamma'(2)}$ então $Z \sim EE(1, \delta)$.

Prova: Seja t um ponto no suporte da WE, então

$$\begin{aligned} F_Z(t) &= P(Z < t) = P\left(\frac{Y^\phi}{\lambda^\phi} e^{\Gamma'(2)} < t\right) = P\left(Y^\phi < t \frac{\lambda^\phi}{e^{\Gamma'(2)}}\right) = P\left(Y < t^{1/\phi} \frac{\lambda}{e^{\Gamma'(2)/\phi}}\right) = \\ &= F_Y\left(t^{1/\phi} \frac{\lambda}{e^{\Gamma'(2)/\phi}}\right) = \left\{ 1 - \exp \left[- \left(t^{1/\phi} \frac{\lambda}{e^{\Gamma'(2)/\phi}} \right)^\phi e^{\Gamma'(2)} \right] \right\}^\delta = [1 - \exp(-t)]^\delta. \end{aligned}$$

■

A função geradora de momentos da distribuição WE é obtida (sem restrições) segundo a expressão, apresentada por Choudhury (2005) o qual demonstrou a

convergência da série para $\phi > 0$, $\delta > 0$ e k inteiro positivo, cuja expressão é dada por

$$\mu'_k = \lambda_i^k \delta \exp \left\{ -\frac{k\Gamma'(2)}{\phi} \right\} \Gamma_k(\phi) \left[1 + \sum_{i=1}^{\infty} a_i (i+1)^{-(1+\frac{k}{\phi})} \right]$$

em que

$$a_i = (-1)^i \binom{\delta-1}{i} = (-1)^i \frac{\delta^1(\delta^1-1)(\delta^1-2)\dots(\delta^1-i+1)}{i!}, \quad \delta^1 = \delta - 1.$$

Portanto, os dois primeiros momentos ordinários da WE são

$$\mu'_1 = E(Y_i) = \lambda_i \delta \exp \left\{ -\frac{\Gamma'(2)}{\phi} \right\} \Gamma_1(\phi) \left[1 + \sum_{i=1}^{\infty} a_i (i+1)^{-(1+\frac{1}{\phi})} \right]$$

e

$$\mu'_2 = E(Y_i^2) = \lambda_i^2 \delta \exp \left\{ -\frac{2\Gamma'(2)}{\phi} \right\} \Gamma_2(\phi) \left[1 + \sum_{i=1}^{\infty} a_i (i+1)^{-(1+\frac{2}{\phi})} \right]$$

e logo, a variância e a esperança são expressas por:

$$E(Y_i) = \lambda_i c_1(\phi, \delta) \quad \text{e} \quad Var(Y_i) = \lambda_i^2 [h_2(\phi, \delta) - h_1^2(\phi, \delta)] = \lambda_i^2 c_2(\phi, \delta)$$

em que

$$c_1(\phi, \delta) = h_1(\phi, \delta) = \delta \exp \left\{ -\frac{\Gamma'(2)}{\phi} \right\} \Gamma_1(\phi) \left[1 + \sum_{i=1}^{\infty} a_i (i+1)^{-(1+\frac{1}{\phi})} \right]$$

e

$$h_2(\phi, \delta) = \delta \exp \left\{ -\frac{2\Gamma'(2)}{\phi} \right\} \Gamma_2(\phi) \left[1 + \sum_{i=1}^{\infty} a_i (i+1)^{-(1+\frac{2}{\phi})} \right]$$

Observa-se que o MLWG proposto por Prudente e Cordeiro (2010) é um caso particular quando se têm $\delta = 1$, além de similaridades entre o trabalho apresentado por Prudente e Cordeiro (2010), pois da mesma forma como discutem os autores o parâmetro λ_i é um tipo de parâmetro de locação o qual também afeta a variância e a função de variância da distribuição WE e em termos da nova parametrização é uma função quadrática da média.

Os gráficos da função densidade para alguns valores selecionados de δ e λ_i são apresentados nas Figuras 3 e 4.

O MLWEG fica então definido pela fdp em (9), especificando o componente aleatório para a variável resposta y_i e por uma função g conhecida, duas vezes diferenciável

Figura 3 – Densidades da WE

e bijetora que liga o parâmetro λ_i ao preditor η_i , denominada, portanto, de função de ligação, tal que:

$$g(\lambda_i) = \eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$$

em que $\mathbf{x}_i$ é o vetor $1 \times p$ de covariáveis relacionadas a i -ésima variável resposta y_i e $\boldsymbol{\beta}^T = (\beta_1, \dots, \beta_p)$ é o vetor de parâmetros da regressão a ser estimado. O parâmetro λ_i pode ser explicitado pela inversa da função de ligação, ou seja: $\lambda_i = g^{-1}(\eta_i) = g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta})$ e portanto $\mu_i = E(Y_i) = g^{-1}(\eta_i) c_1(\phi, \delta)$ em que se percebe que parte da resposta média para a i -ésima variável é explicada por λ_i , além de influenciar na variância do modelo, pois a função de variância da distribuição WE é uma função quadrática da média $Var(Y_i) = \lambda_i^2 h_2(\phi, \delta) - \mu_i^2$.

O comportamento de $c_1(\phi, \delta)$ e $c_2(\phi, \delta)$ é apresentado na Figura (5), para valores de δ igual 0.2, 1, 2 e 9.

Figura 4 – Densidades da WE

Observa-se a partir da Figura (5), que conforme os valores para ϕ aproximam-se de zero a esperança e variância aumentam e conforme ϕ aumenta a esperança e variância diminuem. Em particular se $\delta = 1$, quando $\phi \rightarrow \infty$, $c_1(\phi, 1) \rightarrow 1$ e $c_2(\phi, 1) \rightarrow 0$.

2.2.2 Modelo de regressão

Sabe-se que na prática existem muitas situações em que a variável resposta é influenciada por uma ou mais variáveis regressoras. Essas variáveis regressoras também são chamadas de variáveis explicativas ou covariáveis que podem ser: tratamentos, características intrínsecas das unidades amostrais, variáveis exógenas, interações, variáveis dependentes do tempo e etc.

Seja $\mathbf{x}_i^T = (x_{i1}, x_{i2}, \dots, x_{ip})$ um conjunto de p covariáveis associadas a i -

Figura 5 – Gráficos das funções $c_1(\phi, \delta)$ e $c_2(\phi, \delta)$

ésima variável resposta y_i para $i = 1, \dots, n$. Se $Y_1, \dots, Y_n$ é uma sequência de n variáveis aleatórias independentes, tais que $Y_i \sim WE(\lambda_i, \phi, \delta)$ e $g(\lambda_i) = \eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$ é a função de ligação, então a equação que melhor descreve o modelo de regressão é

$$Y_i = \lambda_i \varepsilon_i = g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta}) \varepsilon_i$$

com $\varepsilon_i \sim WE(1, \phi, \delta)$ e média e variância dadas, respectivamente, por

$$\begin{aligned} E(Y_i) &= \lambda_i E(\varepsilon_i) = \lambda_i c_1(\phi, \delta) = g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta}) c_1(\phi, \delta) \\ \text{Var}(Y_i) &= \lambda_i^2 \text{Var}(\varepsilon_i) = \lambda_i^2 c_2(\phi, \delta) = [g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta})]^2 c_2(\phi, \delta) \end{aligned}$$

O componente sistemático $\eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$ entra no modelo de regressão na forma de uma estrutura linear para as variáveis explicativas e a ligação entre o componente sistemático e o aleatório é realizada pela função de ligação g tal que $g^{-1} > 0$. A escolha da função de ligação depende muito do experimento e da relação entre a variável resposta e as covariáveis. Algumas possíveis escolhas para a função de ligação são, por exemplo, a identidade $g(\lambda_i) = \lambda_i$, a recíproca $g(\lambda_i) = 1/\lambda_i$ e a logarítmica $g(\lambda_i) = \log(\lambda_i)$. Dessa forma, o modelo de regressão é composto por uma estrutura linear nos parâmetros em uma função não-linear com erro multiplicativo.

2.2.3 Estimação, testes e simulação

2.2.3.1 Estimação de máxima verossimilhança

Considere agora uma amostra aleatória $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$ de n observações independentes, em que $\mathbf{x}_i^T = (x_{i1}, x_{i2}, \dots, x_{ip})$ com $i = 1, \dots, n$ é o vetor de covariáveis associado ao i -ésimo indivíduo e y_i a variável resposta seguindo distribuição $WE(\lambda_i, \phi, \delta)$ com $\lambda_i = g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta})$. O logaritmo da função de verossimilhança para os parâmetros do modelo $\boldsymbol{\theta}^T = (\phi, \delta, \boldsymbol{\beta}^T)$ é dado por

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \{ \log(\delta) - \log(y_i/\phi) + \log(z_i) - z_i + (\delta - 1) \log[1 - \exp(-z_i)] \} , \quad (10)$$

em que $z_i = (y_i/\lambda_i)^\phi e^{\Gamma'(2)}$.

O estimador de máxima verossimilhança $\hat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$ pode ser obtido pela maximização do logaritmo da função de verossimilhança dado em (10), o que implica resolver o sistema de equações não lineares $\partial \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} = \mathbf{0}$ que não possui, em geral, solução analítica, sendo necessário o uso de métodos computacionais iterativos. Portanto, para obtenção das estimativas dos parâmetros do modelo o procedimento NLMixed do *software* estatístico SAS foi utilizado.

Os vetores escore para $\boldsymbol{\beta}$, ϕ e δ podem ser escritos, respectivamente, por

$$U_{\boldsymbol{\beta}} = \frac{\partial \ell}{\partial \boldsymbol{\beta}} = \phi \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{s} , \quad U_{\phi} = \frac{\partial \ell}{\partial \phi} = \frac{n}{\phi} + \boldsymbol{\tau}^T \mathbf{1} \quad \text{e} \quad U_{\delta} = \frac{\partial \ell}{\partial \delta} = \boldsymbol{\kappa}^T \mathbf{1}$$

em que $\mathbf{W} = \text{diag}[(\lambda'_i/\lambda_i)^2]$ é uma matriz diagonal de ordem n . $\lambda'_i = \partial \lambda_i / \partial \eta_i$, $\lambda''_i = \partial^2 \lambda_i / \partial \eta_i^2$ são as derivadas da inversa da função de ligação, $\mathbf{1}$ é um vetor $n \times 1$ de uns e os vetores $\mathbf{s} = (s(z_1), s(z_2), \dots, s(z_n))^T$, $\boldsymbol{\tau} = (\tau_1, \tau_2, \dots, \tau_n)^T$ e $\boldsymbol{\kappa} = (\kappa_1, \kappa_2, \dots, \kappa_n)^T$ têm componentes expressos, respectivamente, por:

$$s(z_i) = (z_i - 1) - (\delta - 1)l_i , \quad \tau_i = -s(z_i) \log(y_i/\lambda_i) , \quad \kappa_i = \delta^{-1} + \log[1 - \exp(-z_i)]$$

e

$$z_i = \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} , \quad l_i = z_i \frac{\exp(-z_i)}{1 - \exp(-z_i)} .$$

Sob condições de regularidade, a distribuição assintótica de $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ é a distribuição Normal multivariada $N_{p+2}(0, \mathbf{I}_{\theta\theta}^{-1})$, em que $\mathbf{I}_{\theta\theta}$ é a matriz de informação

esperada, apresentada em (11). A matriz de covariância assintótica $\mathbf{I}_{\theta\theta}^{-1}$ pode ser aproximada pela inversa da matriz $(p+2) \times (p+2)$ de informação observada $\ddot{\mathbf{L}}_{\theta\theta}$ e então a inferência assintótica para o vetor de parâmetros $\boldsymbol{\theta}$ pode ser baseada na aproximação normal $N_{p+2}(0, \ddot{\mathbf{L}}_{\theta\theta}^{-1})$ para $\hat{\boldsymbol{\theta}}$. A matriz de informação observada é

$$\ddot{\mathbf{L}}_{\theta\theta} = \begin{bmatrix} \ddot{\mathbf{L}}_{\beta\beta} & \ddot{\mathbf{L}}_{\beta\phi} & \ddot{\mathbf{L}}_{\beta\delta} \\ \ddot{\mathbf{L}}_{\phi\beta} & \ddot{\mathbf{L}}_{\phi\phi} & \ddot{\mathbf{L}}_{\phi\delta} \\ \ddot{\mathbf{L}}_{\delta\beta} & \ddot{\mathbf{L}}_{\delta\phi} & \ddot{\mathbf{L}}_{\delta\delta} \end{bmatrix} = \begin{bmatrix} -\phi^2 \mathbf{X}^T \mathbf{W} [\mathbf{d}] \mathbf{X} & \mathbf{X}^T \mathbf{W}^{1/2} \underline{\mathbf{U}}_{\phi} & -\phi \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{L} \\ - & -\frac{n}{\phi^2} - \frac{1}{\phi^2} \mathbf{M}^T \mathbf{1} & \frac{1}{\phi} \mathbf{B}^T \mathbf{1} \\ - & - & -\frac{n}{\delta^2} \end{bmatrix},$$

em que

$$\mathbf{B} = (b_1, b_2, \dots, b_n)^T, \quad \mathbf{d} = \text{diag}[d_1, d_2, \dots, d_n], \quad \underline{\mathbf{U}}_{\phi} = (\underline{u}_1^{\phi}, \underline{u}_2^{\phi}, \dots, \underline{u}_n^{\phi})^T,$$

$$\mathbf{M} = (m_1, m_2, \dots, m_n)^T, \quad \mathbf{L} = (l_1, l_2, \dots, l_n)^T,$$

$$b_i = c_i l_i, \quad c_i = \phi \log \left(\frac{y_i}{\lambda_i} \right) = \log \left(\frac{z_i}{e^{\Gamma(2)}} \right), \quad f_i = \frac{\lambda_i'' \lambda_i}{(\lambda_i')^2},$$

$$d_i = \frac{1}{\phi} \{ \phi z_i - s(z_i)(f_i - 1) - \phi(\delta - 1)l_i[1 - \exp(z_i)l_i] \},$$

$$m_i = n_i c_i^2, \quad n_i = z_i - (\delta - 1)l_i[1 - \exp(z_i)l_i],$$

$$\underline{u}_i^{\phi} = s(z_i) + n_i c_i.$$

A matriz de informação esperada é

$$\mathbf{I}_{\theta\theta} = \begin{bmatrix} \mathbf{I}_{\beta\beta} & \mathbf{I}_{\beta\phi} & \mathbf{I}_{\beta\delta} \\ \mathbf{I}_{\phi\beta} & \mathbf{I}_{\phi\phi} & \mathbf{I}_{\phi\delta} \\ \mathbf{I}_{\delta\beta} & \mathbf{I}_{\delta\phi} & \mathbf{I}_{\delta\delta} \end{bmatrix} = \begin{bmatrix} \phi^2 \mathbf{X}^T \mathbf{W} \mathbf{V} \mathbf{X} & \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{O} & \phi \delta t_1 \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{1} \\ - & -\frac{n}{\phi^2} - \frac{1}{\phi^2} \mathbf{G}^T \mathbf{1} & -\frac{n\delta}{\phi^2} t_2 \\ - & - & \frac{n}{\delta^2} \end{bmatrix}, \quad (11)$$

em que

$$\mathbf{V} = \text{diag}(v_1, v_2, \dots, v_n)^T, \quad \mathbf{O} = (o_1, o_2, \dots, o_n)^T, \quad \mathbf{G} = (g_1, g_2, \dots, g_n)^T$$

$$T(d, b, a, e) = \int_0^{\infty} x^{d-1} e^{-bx} (1 - e^{-x})^{a-1} [\log(x)]^e dx$$

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt, \quad \psi(x) = \frac{\partial \log(\Gamma(x))}{\partial x}, \quad \psi'(x) = \frac{\partial^2 \log(\Gamma(x))}{\partial x^2},$$

$$t_1 = T(2, 2, \delta - 1, 0) \quad , \quad t_2 = T(3, 2, \delta - 1, 1) - \psi(2)T(3, 2, \delta - 1, 0) \quad ,$$

$$v_i = \frac{\delta - 1}{B(\delta, 1)} T(3, 2, \delta - 2, 0) + 1 \quad ,$$

$$o_i = 1 + \frac{1}{B(\delta, 1)} \left\{ T(2, 1, \delta, 1) + (\delta - 1) [T(3, 0, \delta - 1, 1) - T(2, 2, \delta - 1, 1) - \Gamma'(2)T(3, 2, \delta - 2, 0)] \right\} \quad ,$$

$$g_i = \psi^2(2) [\psi(\delta + 1) - \psi(1)] + \delta \{ T(2, 1, \delta, 2) - 2\psi(2)T(2, 1, \delta, 1) - (\delta - 1) [T(2, 2, \delta - 1, 2) - T(3, 2, \delta - 2, 2) - 2\psi(2) [T(2, 2, \delta - 1, 1) - T(3, 2, \delta - 2, 1)] + \psi^2(2) [T(2, 2, \delta - 1, 0) - T(3, 2, \delta - 2, 0)] \} \quad .$$

2.2.3.2 Estudo de Simulação: Máxima verossimilhança

Realizou-se um estudo de simulação, utilizando a metodologia de Monte Carlo, para avaliar as estimativas de máxima verossimilhança do modelo MLWEG, considerando o componente sistemático $\eta_i = \beta_0 + \beta_1 x_i, i = 1, \dots, n$, e funções de ligação recíproca ($g(\lambda_i) = 1/\lambda_i$), identidade ($g(\lambda_i) = \lambda_i$) e logarítmica ($g(\lambda_i) = \log(\lambda_i)$).

Os valores para a variável resposta foram simulados apartir da inversa da fda da WE, equação (8), tal que $Y_i = F^{-1}(U_i)$ com $U_i \sim U(0, 1)$ e os valores para a covariável X foram gerados pela equação $X_i = \exp(T_i)$ em que $T_i \sim N(0, 1)$. De forma que X apresenta forte assimetria à direita.

Portanto, para cada combinação $n, \beta_0, \beta_1, \phi, \delta$ e $g(\lambda_i)$, 1000 amostras aleatórias foram geradas e para cada uma dessas amostras as estimativas de máxima verossimilhança dos parâmetros foram obtidas. Dessas 1000 estimativas de cada parâmetro, a média, o viés e erro quadrático médio (EQM) foram calculados e listados nas Tabelas 1 e 2.

Para esse estudo, adotou-se quatro tamanhos amostrais ($n = 75, 150, 300$ e 750), as funções de ligação recíproca, identidade e logarítmica e os valores verdadeiros para os parâmetros foram fixados em $\beta_0 = 1, 5, \beta_1 = 2, 0, \phi = 0, 5$ e $\delta = 0, 5$ e $1, 5$.

Dos valores apresentados nas Tabelas 1 e 2, observa-se que, em todos os casos, a medida que o tamanho amostral aumenta o EQM e o viés diminuem e as estima-

tivas se aproximam dos verdadeiros valores. Nota-se ainda que o viés e o EQM são bem menores para os parâmetros estimados β_0 e β_1 quando considerado a função de ligação logarítmica, além da convergência para os valores verdadeiros dos parâmetros ser mais rápida. Por fim, para valores maiores de δ observa-se aumento no viés e no EQM para as estimativas desse parâmetro.

2.2.3.3 Estimação Bayesiana

Nesta seção os procedimentos inferenciais sob um enfoque Bayesiano é apresentado. Considerando uma amostra aleatória $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$ de n observações independentes, em que $\mathbf{x}_i^T = (x_{i1}, x_{i2}, \dots, x_{ip})$ com $i = 1, \dots, n$ é o vetor de covariáveis associado ao i -ésimo indivíduo e y_i a variável resposta seguindo distribuição $WE(\lambda_i, \phi, \delta)$ com $\lambda_i = g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta})$. O logaritmo da função de verossimilhança para os parâmetros do modelo $\boldsymbol{\theta}^T = (\phi, \delta, \boldsymbol{\beta}^T)$ é dado por

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \{ \log(\delta) - \log(y_i/\phi) + \log(z_i) - z_i + (\delta - 1) \log[1 - \exp(-z_i)] \} , \quad (12)$$

em que $z_i = (y_i/\lambda_i)^\phi e^{\Gamma'(2)}$.

2.2.3.4 Distribuições *a priori* e *a posteriori*

Neste trabalho, considerou-se uma distribuição *a priori* conjunta própria para os parâmetros do modelo para garantir que a distribuição *a posteriori* conjunta seja própria (Ibrahim et al., 2001). Seja β , δ e ϕ mutuamente independentes *a priori*. Assim, a densidade *a priori* conjunta de $\boldsymbol{\theta}$ é dada por

$$\pi(\boldsymbol{\theta}) = \prod_{i=1}^p \pi(\beta_i) \pi(\delta) \pi(\phi), \quad (13)$$

em que $\beta_j \sim N(0, \sigma_{\beta_j}^2)$, $j = 1, \dots, p$, $\log(\delta) \sim N(0, \sigma_\delta^2)$ e $\log(\phi) \sim N(0, \sigma_\phi^2)$. Todos os hiper-parâmetros foram especificados para expressar *prioris* não informativas.

Combinando a função de verossimilhança (12) e a distribuição *a priori* em (13), a distribuição *a posteriori* para $\boldsymbol{\theta}$ é obtida como

$$\pi(\boldsymbol{\theta}|\mathcal{D}) \propto L(\boldsymbol{\theta}; \mathcal{D}) \prod_{i=1}^p \pi(\beta_i) \pi(\delta) \pi(\phi).$$

Tabela 1 – Estimativas média, viés e EQM baseados em 1000 simulações do MLWEG com $\beta_0 = 1, 5$, $\beta_1 = 2, 0$, $\phi = 0, 5$ e $\delta = 0, 5$ e $1, 5$

$g(\lambda_i)$	n	θ	$\delta = 0.5$			θ	$\delta = 1.5$		
			Média	Viés	EQM		Média	Viés	EQM
$1/\lambda_i$	75	β_0	3,7136	2,2136	566,7043	β_0	3,0633	1,5633	52,5531
		β_1	5,2468	3,2468	1231,4334	β_1	4,1954	2,1954	158,4900
		ϕ	0,6288	0,1288	0,1077	ϕ	0,5806	0,0806	0,0636
		δ	0,5435	0,0435	0,1920	δ	1,8357	0,3357	2,7380
$1/\lambda_i$	150	β_0	1,7279	0,2279	2,3447	β_0	2,118	0,6180	7,1914
		β_1	2,4023	0,4023	3,6811	β_1	2,7272	0,7272	11,2634
		ϕ	0,5584	0,0584	0,0304	ϕ	0,5254	0,0254	0,0167
		δ	0,4999	-0,0001	0,0341	δ	1,6823	0,1823	0,7946
$1/\lambda_i$	300	β_0	1,6549	0,1549	0,9182	β_0	1,7592	0,2592	1,0130
		β_1	2,2022	0,2022	1,6008	β_1	2,2627	0,2627	1,2709
		ϕ	0,5250	0,0250	0,0102	ϕ	0,5073	0,0073	0,0060
		δ	0,5004	0,0004	0,0164	δ	1,5956	0,0956	0,2450
$1/\lambda_i$	750	β_0	1,5809	0,0809	0,3340	β_0	1,5818	0,0818	0,2399
		β_1	2,0757	0,0757	0,3906	β_1	2,0824	0,0824	0,3840
		ϕ	0,5065	0,0065	0,0031	ϕ	0,5044	0,0044	0,0025
		δ	0,5031	0,0031	0,0060	δ	1,5283	0,0283	0,0834
λ_i	75	β_0	2,0274	0,5274	5,5805	β_0	1,9304	0,4304	3,4099
		β_1	2,3816	0,3816	5,4610	β_1	2,4132	0,4132	4,5988
		ϕ	0,6596	0,1596	0,1335	ϕ	0,5913	0,0913	0,0817
		δ	0,4954	-0,0046	0,0844	δ	1,7533	0,2533	1,5490
λ_i	150	β_0	1,6272	0,1272	2,0534	β_0	1,5747	0,0747	1,1083
		β_1	2,3445	0,3445	2,9381	β_1	2,1675	0,1675	1,8551
		ϕ	0,5681	0,0681	0,0367	ϕ	0,5225	0,0225	0,0174
		δ	0,4995	-0,0005	0,0406	δ	1,7024	0,2024	0,7972
λ_i	300	β_0	1,6210	0,1210	0,9375	β_0	1,6188	0,1188	0,5708
		β_1	2,0686	0,0686	1,1237	β_1	2,0979	0,0979	0,7743
		ϕ	0,5236	0,0236	0,0102	ϕ	0,5173	0,0173	0,0072
		δ	0,5038	0,0038	0,0175	δ	1,5510	0,0510	0,2683
λ_i	750	β_0	1,5326	0,0326	0,2753	β_0	1,5656	0,0656	0,2185
		β_1	2,0224	0,0224	0,4127	β_1	2,0414	0,0414	0,3016
		ϕ	0,5067	0,0067	0,0031	ϕ	0,5085	0,0085	0,0027
		δ	0,5032	0,0032	0,0058	δ	1,5082	0,0082	0,0789

Tabela 2 – Estimativas média, viés e EQM baseados em 1000 simulações do MLWEG com $\beta_0 = 1,5$, $\beta_1 = 2,0$, $\phi = 0,5$ e $\delta = 0,5$ e $1,5$

$g(\lambda_i)$	n	θ	Média	Viés	EQM	θ	Média	Viés	EQM
$\log(\lambda_i)$	75	β_0	1,4416	-0,0584	0,7315	β_0	1,3060	-0,1940	1,1192
		β_1	1,9573	-0,0427	0,0455	β_1	1,9868	-0,0132	0,0138
		ϕ	0,6329	0,1329	0,1173	ϕ	0,5668	0,0668	0,0677
		δ	0,5394	0,0394	0,1877	δ	2,0828	0,5828	5,6226
$\log(\lambda_i)$	150	β_0	1,4700	-0,0300	0,3306	β_0	1,4362	-0,0638	0,4242
		β_1	1,9743	-0,0257	0,0168	β_1	1,9912	-0,0088	0,0054
		ϕ	0,5507	0,0507	0,0319	ϕ	0,5295	0,0295	0,0185
		δ	0,5211	0,0211	0,0423	δ	1,6864	0,1864	0,8804
$\log(\lambda_i)$	300	β_0	1,5056	0,0056	0,1509	β_0	1,4908	-0,0092	0,1684
		β_1	1,9842	-0,0158	0,0072	β_1	1,9940	-0,0060	0,0024
		ϕ	0,5237	0,0237	0,0104	ϕ	0,5168	0,0168	0,0074
		δ	0,5070	0,0070	0,0179	δ	1,5545	0,0545	0,2607
$\log(\lambda_i)$	750	β_0	1,4984	-0,0016	0,0556	β_0	1,4940	-0,0060	0,0618
		β_1	1,9938	-0,0062	0,0027	β_1	1,9980	-0,0020	0,0009
		ϕ	0,5095	0,0095	0,0032	ϕ	0,5058	0,0058	0,0025
		δ	0,5019	0,0019	0,0060	δ	1,5211	0,0211	0,0841

A distribuição *a posteriori* conjunta não é tratável analiticamente, assim métodos Monte Carlo em Cadeias de Markov (MCMC) podem ser utilizados para obter amostras, a partir de quais características da distribuição *a posteriori* de interesse podem ser inferidas (Gamerman e Lopes, 2006). O algoritmo foi implementado no *software* (R Core Team, 2011) e monitorou-se a convergência utilizando tanto pelo método proposto por Geweke (1992) quanto por métodos gráficos.

Para iniciar a implementação do algoritmo, transformou-se as variáveis $\xi = (\log(\delta), \log(\phi), \beta^T)^T$ para trabalhar com todos os parâmetros na reta. Assim, o espaço paramétrico é transformado em $\mathcal{R}^{p+2}$. Como em Suzuki et. al (2016), a matriz de transição utilizada é uma distribuição Normal multivariada, cujo vetor de médias corresponde ao ajuste clássico do logaritmo da distribuição *a posteriori* e, utiliza-se como matriz de covariância a decomposição de Choleski na negativa da inversa da matriz hessiana.

Os passos para a implementação do algoritmo Metropolis-Hastings são:

- (1) Iniciar com o vetor $\xi_{(j)}$ e o indicador de iterações $j = 0$;

- (2) Gerar um ponto ξ' segundo o núcleo de transição $Q(\xi', \xi_j) = N_{p+2}(\xi_j, \tilde{\Sigma})$, em que $\tilde{\Sigma}$ é a matriz de covariância de ξ , a qual é mesma em qualquer estágio;
- (3) Atualizar $\xi_{(j)}$ para $\xi_{(j+1)} = \xi'$ com probabilidade $p_j = \min\{1, \pi(\xi'|\mathcal{D})/\pi(\xi_{(j)}|\mathcal{D})\}$, ou manter $\theta_{(j)}$ com probabilidade $1 - p_j$;
- (4) Repetir os passos (2) e (3) incrementando o contador até o processo atingir a distribuição estacionária.

2.2.3.5 Critério de comparação de modelo

Na literatura, encontra-se metodologias que podem ser aplicadas para comparação de modelos competitivos para um conjunto de dados e seleção do modelo que melhor se ajusta aos dados. Uma das abordagens mais utilizadas é derivada das ordenadas da densidade preditiva condicional (CPO - *conditional predictive ordinate*). Para uma discussão detalhada sobre a estatística CPO e suas aplicações em seleção de modelos, veja Gelfand (1992).

Seja $\mathcal{D}$ e $\mathcal{D}^{(-i)}$ a representação dos dados completos e dos dados com a i -ésima observação excluída, respectivamente. A densidade *a posteriori* de θ dado $\mathcal{D}^{(-i)}$ é denotada por $\pi(\theta|\mathcal{D}^{(-i)})$, $i = 1, \dots, n$. Para a i -ésima observação, CPO_i pode ser escrita como

$$CPO_i = \int_{\theta \in \Theta} f(y_i|\theta) \pi(\theta|\mathcal{D}^{(-i)}) d\theta = \left\{ \int_{\theta} \frac{\pi(\theta|\mathcal{D})}{f(y_i|\theta)} d\theta \right\}^{-1}. \quad (14)$$

A CPO_i não tem forma fechada para o modelo proposto. Entretanto, uma estimativa de Monte Carlo de CPO_i pode ser obtida usando uma simples amostra MCMC da distribuição *a posteriori* $\pi(\theta|\mathcal{D})$.

Seja $\theta^{(1)}, \dots, \theta^{(Q)}$ uma amostra de tamanho Q de $\pi(\theta|\mathcal{D})$ após o *burn-in*. Uma aproximação Monte Carlo da CPO_i (Ibrahim et al., 2001) é dada por

$$\widehat{CPO}_i = \left\{ \frac{1}{Q} \sum_{q=1}^Q \frac{1}{f(y_i|\theta^{(q)})} \right\}^{-1}.$$

Para comparações de modelo usa-se o log pseudo verossimilhança marginal (LPML- *log*

pseudo marginal likelihood) definido por $LPML = \sum_{i=1}^n \log(\widehat{CPO}_i)$. Quanto maior o valor de LPML, melhor o ajuste do modelo.

O critério desvio-informação (DIC) proposto por Spiegelhalter et al. (2002), o critério informação de Akaike esperado (EAIC) - Brooks (2002) e o critério de informação Bayesiano (EBIC) - Carlin e Louis (2001) também podem ser utilizados. Eles são baseados na média *a posteriori* dos desvios, os quais podem ser aproximados por $\bar{d} = \sum_{q=1}^Q d(\boldsymbol{\theta}_q)/Q$, em que $d(\boldsymbol{\theta}) = -2 \sum_{i=1}^n \log [f(y_i|\boldsymbol{\theta})]$.

O critério DIC pode ser estimado usando as saídas MCMC por $\widehat{DIC} = \bar{d} + \hat{\rho}_d = 2\bar{d} - \hat{d}$, em que ρ_D é o número efetivo de parâmetros definido como $E\{d(\boldsymbol{\theta})\} - d\{E(\boldsymbol{\theta})\}$, em que $d\{E(\boldsymbol{\theta})\}$ é o desvio avaliado na média *a posteriori* e pode ser estimado como

$$\hat{D} = d \left(\frac{1}{Q} \sum_{q=1}^Q \boldsymbol{\beta}^{(q)}, \frac{1}{Q} \sum_{q=1}^Q \delta^{(q)}, \frac{1}{Q} \sum_{q=1}^Q \phi^{(q)} \right).$$

Da mesma forma, os critérios EAIC e EBIC podem ser estimados pelas médias de $\widehat{EAIC} = \bar{d} + 2\#(\boldsymbol{\theta})$ e $\widehat{EBIC} = \bar{d} + \#(\boldsymbol{\theta}) \log(n)$, em que $\#(\boldsymbol{\theta})$ é o número de parâmetros do modelo.

2.2.3.6 Estudo de simulação: Análise Bayesiana

Um estudo de simulação foi conduzido para avaliar as estimativas dos parâmetros para o modelo proposto por meio da análise Bayesiana. Considere o componente sistemático $\eta_i = \beta_0 + \beta_1 x_i, i = 1, \dots, n$ e a função de ligação logarítmica ($g(\lambda_i) = \log(\lambda_i)$).

Os valores para a variável resposta são simulados da inversa da fda da distribuição WE, tal que $Y_i = F^{-1}(U_i)$ com $U_i \sim U(0, 1)$. E, os valores das covariáveis x são gerados segundo $x_i = \exp(t_i)$, em que $t_i \sim N(0, 1)$, assim x é assimétrico fortemente a direita.

Foram geradas 1.000 réplicas com amostras de tamanho $n = 150$ e 300 assumindo a função de ligação logarítmica e os verdadeiros valores para os parâmetros fixados em $\beta_0 = 2, 0, \beta_1 = 1, 0, \phi = 0, 5$ e $\delta = 1, 5$. Os seguintes valores $\sigma_{\beta_0}^2 = \sigma_{\beta_1}^2 = \sigma_{\delta}^2 = \sigma_{\phi}^2 = 100$ para as distribuições *a priori* dadas em (13) foram considerados. Observe que assumi-se distribuições *a priori* próprias porém pouco informativas. Em todo o trabalho, foi considerado uma cadeia de tamanho 200.000 para cada parâmetro, após desconsiderar as primeiras 40.000 amostras (*burn-in*) e, para evitar problemas de autocorrelação, considerou-se

um espaçamento de tamanho 20, obtendo uma amostra efetiva de tamanho 10.000 sobre a qual a inferência *a posteriori* é baseada.

Adicionalmente, a média, o viés, o erro quadrático médio (EQM) e a probabilidade de cobertura de 95% foram calculados e apresentados na Tabela 3.

Tabela 3 – Média, viés, erro quadrático médio (EQM) e probabilidade de cobertura (PC) de 95% para os parâmetros estimados para ajuste do modelo MLWEG

n	Parâmetro	Média	Viés	EQM	PC
150	δ	2,0092	0,5092	3,0407	0,916
	ϕ	0,7096	0,2096	0,6888	0,920
	β_0	1,8663	-0,1337	1,3791	0,904
	β_1	1,0058	0,0058	0,0125	0,912
300	δ	1,7323	0,2323	0,7928	0,936
	ϕ	0,5059	0,0059	0,0137	0,940
	β_0	1,8778	-0,1222	0,3982	0,934
	β_1	1,0022	0,0022	0,0047	0,956

2.2.4 Análise de sensibilidade

Uma 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) 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 têm 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). Entretanto, 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 aplicado influência local em modelos de análise de regressão. Recentemente têm-se vários trabalhos publicados, a saber: Silva *et al.* (2008) estuda influência local e modelos de regressão log-Burr XII com dados censurados, Ortega *et al.* (2009) aplica influência local em modelos de regressão log-gama generalizado com fração de cura, Russo *et al.* (2012) aplica influência local em modelo elípticos autorregressivos não-lineares para dados correlacionados, Hashimoto *et al.* (2013) estuda influência local em modelos de regressão log-gama generalizados com censura intervalar, Paula (2013) estuda influência local em modelo linear generalizado duplos, Ortega *et al.* (2013) estuda influência local para o modelo de regressão log-beta Weibull, com aplicação para dados recorrentes de câncer de próstata, dentre outros.

2.2.4.1 Influência local

Quando se está ajustando um modelo a um conjunto de dados, é desejável que as estimativas obtidas a partir do modelo proposto sejam resistentes a pequenas perturbações no modelo ou nas observações. Por isso é importante que se faça um estudo sobre a robustez dos resultados obtidos, em termos dos vários aspectos que envolvem a formulação do modelo, e as estimativas dos seus parâmetros, ou seja, fazer uma análise de diagnóstico, que consiste por exemplo, de métodos para avaliar o grau de sensibilidade das inferências a pequenas perturbações nos dados ou mesmo no modelo proposto.

Enquanto a análise de resíduos é feita para pesquisar problemas com o modelo ajustado, tais como presença de observações aberrantes e afastamento sérios das suposições para a distribuição do erro, uma análise de influência é feita assumindo o modelo como correto, e estuda-se a robustez das conclusões à perturbações nos dados ou no modelo. Uma observação é dita influente se ela produzir alterações desproporcionais nos resultados da análise, quando a mesma for excluída, ou submetida a uma pequena perturbação.

Numa análise de diagnóstico, assumi-se que o modelo postulado é correto, e deseja-se, por exemplo, comparar as estimativas obtidas por meio desse modelo com as respectivas estimativas decorrentes de uma pequena perturbação no mesmo.

O interesse nessa seção é estudar o comportamento do afastamento da verossimilhança $LD(\boldsymbol{\omega}) = 2[\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})]$ em torno de um vetor de não perturbação, $\boldsymbol{\omega}_0$, tal que $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}_0}$ e, portanto, $LD(\boldsymbol{\omega}_0) = 0$. Segundo Cook (1986), esse estudo pode ser feito usando a curvatura normal que, na direção do vetor $\mathbf{d}$, é expressa como $C_d(\hat{\boldsymbol{\theta}}) = 2|\mathbf{d}^T \Delta^T (-\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}})^{-1} \Delta \mathbf{d}|$ em que $\Delta = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T$ e $\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}} = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$, tanto Δ quanto $\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}$ são avaliados em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega}_0$. Portanto, a curvatura máxima, $C_{d_{max}}$, é obtida na direção $\mathbf{d}_{max}$, em que $C_{d_{max}}$ é o maior autovalor de $\ddot{\mathbf{T}} = \Delta^T (-\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}})^{-1} \Delta$ e $\mathbf{d}_{max}$ é o autovetor correspondente. Alternativamente, pode-se examinar a influência local total $C_i = C_{d_i}(\boldsymbol{\theta})$, em que, $\mathbf{d}_i$ é um vetor $n \times 1$ de zeros com um na i -ésima posição. A metodologia de influência local é talvez a mais popular ferramenta moderna para diagnóstico de influência. Ela têm sido largamente utilizada, nos recentes anos, no diagnóstico de regressão.

Como $C_d(\boldsymbol{\theta})$ não é invariante sobre mudança uniforme de escala, Poon e Poon (1999) propuseram a curvatura normal conformal $B_d(\boldsymbol{\theta}) = C_d(\boldsymbol{\theta}) / \text{tr}\{2\ddot{\mathbf{T}}\}$, a qual é invariante sobre reparametrização de ω e $0 \leq B_d(\boldsymbol{\theta}) \leq 1$ para qualquer direção unitária $\mathbf{d}$. Dessa forma, para determinar se a i -ésima observação é possivelmente influente, Poon e Poon (1999), analisaram se a quantidade $M(0)_i = B_{d_i}$ associada a i -ésima observação, em que $\mathbf{d}_i$ é um vetor de dimensão $n \times 1$ de zeros com 1 na i -ésima posição, é maior que $\overline{M(0)} + c^* SM(0)$, em que $\overline{M(0)} = 1/n$, $SM(0)$ é o erro padrão amostral de $\{M(0)_k, k = 1, \dots, n\}$ e c^* é uma constante selecionada. Dependendo da aplicação real, c^* pode assumir qualquer valor. A seguir a expressão para a matriz Δ é obtida sobre três diferentes esquemas de perturbação.

• Perturbação da verossimilhança

A perturbação da verossimilhança é realizada atribuindo pesos a cada observação no logaritmo da função de verossimilhança. Assim o logaritmo da função de verossimilhança perturbado, neste caso, é dado por

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \{ \log(\delta) - \log(y_i/\phi) + \log(z_i) - z_i + (\delta - 1) \log[1 - \exp(-z_i)] \} ,$$

em que $z_i = (y_i/\lambda_i)^\phi e^{\Gamma'(2)}$ e $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$ é o vetor de não perturbação. Portanto, para o esquema de perturbação de casos, têm-se que

$$\Delta = \begin{bmatrix} \hat{\phi} \mathbf{X}^T \hat{\mathbf{W}}^{1/2} \underline{\mathbf{s}} \\ (u_1^\phi, \dots, u_n^\phi) \\ (\kappa_1, \dots, \kappa_n) \end{bmatrix},$$

em que $\underline{\mathbf{s}} = \text{diag}\{s(\hat{z}_1), \dots, s(\hat{z}_n)\}$, $u_i^\phi = \hat{\phi}^{-1}[1 - s(\hat{z}_i) \log(\hat{z}_i/e^{\Gamma'(2)})]$ e $\kappa_i = \delta^{-1} + \log[1 - \exp(-z_i)]$.

• Perturbação da variável resposta

A perturbação da variável resposta é realizada de forma aditiva e expressa por $y_{i\omega} = y_i + \omega_i s_y$ em que s_y é o desvio padrão amostral de Y . Consequentemente $\omega_0 = (0, \dots, 0)^T$ e o logaritmo da função de verossimilhança se reduz a

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \{ \log(\delta) - \log(y_{i\omega}/\phi) + \log(z_{i\omega}) - z_{i\omega} + (\delta - 1) \log[1 - \exp(-z_{i\omega})] \},$$

em que $z_{i\omega} = (y_{i\omega}/\hat{\lambda}_i)^{\hat{\phi}} e^{\Gamma'(2)}$. Neste caso, pode-se demonstrar que

$$\Delta = \hat{\phi} s_y \begin{bmatrix} \hat{\phi} \mathbf{X}^T \hat{\mathbf{W}}^{1/2} \underline{\mathbf{s}} \\ (\underline{u}_1^\phi, \dots, \underline{u}_n^\phi) \\ (\underline{l}_1, \dots, \underline{l}_n) \end{bmatrix},$$

em que $\underline{\mathbf{s}} = \text{diag}\{z_1/y_{1\omega}, \dots, z_n/y_{n\omega}\}$, $\underline{u}_i^\phi = -(\underline{u}_i^\phi/\hat{\phi})/y_{i\omega}$ e $\underline{l}_j = l_i/y_{i\omega}$.

• Perturbação na variável explanatória

Semelhante ao caso de perturbação da variável resposta, a perturbação da j -ésima variável (contínua) é realizada de forma aditiva e expressa por $x_{ij\omega} = x_{ij} + \omega_i s_{x_j}$, em que s_{x_j} é o desvio padrão amostral da j -ésima variável explanatória. Portanto, $\omega_0 = (0, \dots, 0)^T$ e o logaritmo da função de verossimilhança pode ser expresso por

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \left\{ \log(\delta) - \log\left(\frac{y_i}{\phi}\right) + \log(z_{i\omega}) - z_{i\omega} + (\delta - 1) \log[1 - \exp(-z_{i\omega})] \right\}$$

em que $z_{i\omega} = (y_i/\hat{\lambda}_{i\omega})^{\hat{\phi}} e^{\Gamma'(2)}$ e $\hat{\lambda}_{i\omega} = g^{-1}(\mathbf{x}_i^T \hat{\boldsymbol{\beta}} + s_{x_j} \hat{\beta}_j \omega_i)$. Assim, para a perturbação da variável explanatória, obtém-se

$$\Delta_{p+2 \times n} = (\Delta_{p \times n}^{\beta}, \Delta_{1 \times n}^{\phi}, \Delta_{1 \times n}^{\delta})^T ,$$

em que

$$\begin{aligned} \Delta_{ij}^{\beta} &= \begin{cases} \hat{\phi} \hat{\beta}_t s_{x_j} x_{ij} (\lambda'_i / \lambda_i)^2 \left[(f_i - 1) s(z_i) - \hat{\phi} n_i \right] , & j \neq t , \\ \hat{\phi} \hat{\beta}_t s_{x_j} x_{ij} (\lambda'_i / \lambda_i)^2 \left[(f_i - 1) s(z_i) - \hat{\phi} n_i \right] + \hat{\phi} s_{x_j} (\lambda'_i / \lambda_i) s(z_i) , & j = t . \end{cases} \\ \Delta_i^{\phi} &= \hat{\phi} \hat{\beta}_t s_{x_j} (\lambda'_i / \lambda_i) \left[\hat{\phi}^{-1} s(z_i) + n_i c_i \right] . \\ \Delta_i^{\delta} &= -\hat{\phi} \hat{\beta}_t s_{x_j} (\lambda'_i / \lambda_i) l_i . \end{aligned}$$

2.2.5 Diagnóstico de casos influentes Bayesiano

É sabido que modelos de regressão são sensíveis aos pressupostos do modelo, geralmente uma análise de sensibilidade é fortemente aconselhável. Cook (1986) usa esta ideia para motivar a avaliação da análise de influência. Ele sugere que maior confiança deve ser dada a um modelo relativamente estável sobre pequenas modificações. Os esquemas de perturbação mais conhecido são baseados na deleção de casos (Cook e Weisberg, 1982) cujos efeitos são estudados removendo completamente caso da análise. Este raciocínio será a base para a metodologia de influência global bayesiana para determinação de observações influentes na análise.

Neste trabalho, utiliza-se a medida de diagnóstico de influência de deleção de casos Bayesiano para a distribuição *a posteriori* conjunta baseada na divergência ψ ((Peng e Dey, 1995) e (Weiss, 1996)).

Seja $D_{\psi}(P, P_{(-i)})$ a ψ -divergência entre P e $P_{(-i)}$, em que P representa a distribuição *a posteriori* de $\boldsymbol{\theta}$ para os dados completos, e $P_{(-i)}$ representa a distribuição de $\boldsymbol{\theta}$ excluindo o i -ésimo caso. Especificamente,

$$D_{\psi}(P, P_{(-i)}) = \int \psi \left(\frac{\pi(\boldsymbol{\theta} | \mathcal{D}^{(-i)})}{\pi(\boldsymbol{\theta} | \mathcal{D})} \right) \pi(\boldsymbol{\theta} | \mathcal{D}) d\boldsymbol{\theta}. \quad (15)$$

em que ψ é uma função convexa com $\psi(1) = 0$. Várias escolhas a respeito de ψ são apre-

sentadas por Dey e Birmiwal (1994). Por exemplo, $\psi(z) = -\log(z)$ define a divergência Kullback-Leibler (K-L), $\psi(z) = 0,5|z - 1|$ define a norma L_1 (ou distância variacional) e $\psi(z) = (z - 1)\log(z)$ resulta a distância J (ou a versão simétrica da divergência K-L).

Seja $\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(Q)}$ uma amostra de tamanho Q de $\pi(\boldsymbol{\theta}|\mathcal{D})$. A $D_\psi(P, P_{(-i)})$ pode ser calculada numericamente por

$$\widehat{D}_\psi(P, P_{(-i)}) = \frac{1}{Q} \sum_{q=1}^Q \psi \left(\frac{\widehat{CPO}_i}{f(y_i|\boldsymbol{\theta}^{(q)})} \right), \quad (16)$$

em que $\widehat{CPO}_i = \left\{ \frac{1}{Q} \sum_{q=1}^Q \frac{1}{f(y_i|\boldsymbol{\theta}^{(q)})} \right\}^{-1}$ é a aproximação numérica da estatística preditiva condicional ordinal da i -ésima observação (Ibrahim et al., 2001).

Note que $D_\psi(P, P_{(-i)})$ pode ser interpretada como a divergência ψ do efeito de deleção do i -ésimo caso dos dados completos sobre a distribuição *a posteriori* conjunta de $\boldsymbol{\theta}$. Como pontuado por Peng e Dey (1995) e Weiss (1996), pode ser difícil para um profissional julgar o ponto de corte da medida de divergência, assim como para determinar se um pequeno subconjunto de observações é influente ou não. Assim, utilizamos a proposta de Peng e Dey (1995) e Weiss (1996) que considera uma moeda viesada, a qual tem probabilidade de sucesso p . Então a divergência ψ entre a moeda viesada e a não viesada é

$$D_\psi(f_0, f_1) = \int \psi \left(\frac{f_0(x)}{f_1(x)} \right) f_1(x) dx, \quad (17)$$

em que $f_0(x) = p^x(1-p)^{1-x}$ e $f_1(x) = 0,5$, $x = 0, 1$. Se $D_\psi(f_0, f_1) = d_\psi(p)$, pode ser facilmente checado que d_ψ satisfaz a seguinte equação

$$d_\psi(p) = \frac{\psi(2p) + \psi(2(1-p))}{2}. \quad (18)$$

Não é difícil ver que para as medidas de divergência consideradas, d_ψ aumenta à medida que p se afasta de 0,50. Além disso, $d_\psi(p)$ é simétrica em torno de $p = 0,50$ e d_ψ atinge seu mínimo em $p = 0,50$. Neste ponto, $d_\psi(0,50) = 0$ e $f_0 = f_1$. Portanto, se $p > 0,90$ (ou $p \leq 0,10$) é considerado como uma moeda muito viciada, então $d_{K-L}(0,90) = 0,51$, $d_J(0,90) = 0,88$ e $d_{L_1}(0,90) = 0,40$.

Assim, se utiliza-se a distância J , pode-se considerar que uma observação é influente quando $d_J > 0,88$. Da mesma forma, usando a divergência Kullback-Leibler e a norma L_1 , uma observação na qual $d_{K-L} > 0,51$ ou $d_{L_1} > 0,40$ pode ser considerada

influyente.

2.2.5.1 Influência de observações *outlying*

Um dos principais objetivos desse estudo é mostrar a necessidade de modelos robustos para lidar com a presença de outliers nos dados.

Para examinar o desempenho da medida de divergência ψ , geramos uma amostra do modelo MLWEG de tamanho 300 considerando a função de ligação logarítmica com parâmetros fixos $\beta_0 = 1,0$, $\beta_1 = 0,5$, $\phi = 2,0$, $\delta = 1,5$. Nestes dados simulados, y_i varia de 0,4535 a 744,8722 com mediana 4,1141, média 13,4078 e desvio padrão 56,6212.

Selecionamos os casos 25, 150 e 225 para perturbação. Para criar observações artificialmente influentes no conjunto de dados, escolhemos um, dois ou três casos selecionados. Para cada caso, perturbamos a variável resposta da seguinte forma: $\tilde{y}_i = y_i + 5S_y$, $i = 25, 150$ e 225 , em que S_y é o desvio padrão de y_i .

Realizamos os mesmos procedimentos descritos na Subseção 2.2.3.6 para a implementação do algoritmo MCMC, assim como a verificação da convergência das cadeias.

A Tabela 4 mostra que as inferências *a posteriori* sobre os parâmetros são sensíveis à perturbação dos casos selecionados, exceto o parâmetro β_1 . Nesta tabela, o conjunto de dados A denota os dados originais simulados sem perturbação e os dados B a H denotam os conjuntos de dados com casos perturbados.

Tabela 4 – Média e desvio padrão (DP) para os parâmetros do MLWEG para cada conjunto de dados simulados

Conjunto de dados	Identificação dos casos perturbados	Parâmetros			
		δ	ϕ	β_0	β_1
		Média (DP)	Média (DP)	Média (DP)	Média (DP)
A	Nenhum	1,633 (0,52)	1,853 (0,29)	0,989 (0,11)	0,490 (0,01)
B	25	27,496 (5,49)	0,490 (0,02)	-1,045 (0,20)	0,479 (0,02)
C	150	20,635 (4,86)	0,560 (0,03)	-0,699 (0,22)	0,489 (0,02)
D	225	26,634 (5,39)	0,499 (0,02)	-1,002 (0,21)	0,481 (0,02)
E	{25, 150}	30,943 (5,71)	0,450 (0,02)	-1,266 (0,20)	0,481 (0,02)
F	{25, 225}	33,288 (5,75)	0,431 (0,02)	-1,388 (0,20)	0,476 (0,02)
G	{150, 225}	30,541 (5,74)	0,454 (0,02)	-1,243 (0,21)	0,483 (0,02)
H	{25, 150, 22}	35,142 (6,22)	0,407 (0,02)	-1,541 (0,22)	0,479 (0,02)

A Tabela 5 mostra as estimativas dos critérios DIC, EAIC, EBIC e LPML

obtidas ao ajustar os diferentes conjuntos de dados perturbados. Observamos que, como esperado, os dados simulados sem perturbação (dados A) teve o melhor ajuste.

Tabela 5 – Critério Bayesiano

Conjunto de dados	Critérios			
	EAIC	EBIC	DIC	LPML
A	1376,4892	1391,3043	1371,7413	-686,1304
B	1502,2160	1517,0310	1497,8440	-755,9290
C	1477,5376	1492,3528	1473,2525	-742,6665
D	1497,5984	1512,4136	1493,3490	-753,0548
E	1542,6485	1557,4637	1538,3057	-772,7122
F	1555,7940	1570,6092	1551,7796	-780,6969
G	1539,2131	1554,0282	1535,1550	-772,1113
H	1592,2849	1607,1000	1589,8817	-796,7907

Para cada conjuntos de dado simulados, considerou-se as amostras da distribuição *a posteriori* dos parâmetros do modelo WE para calcular as três medidas de divergência (d_{KL} , d_J , d_{L_1}) descrita na Subseção 2.2.5. Os resultados na Tabela 6 mostram, antes da perturbação (dados A), os casos selecionados não são influentes segundo todas as medidas de divergência. Entretanto, após a perturbação (dados B-H), as medidas aumentaram, indicando que os casos perturbados são influentes. Dessa forma, claramente pode-se ver que todas as medidas de divergência detectam os pontos influentes.

As Figuras 6, 7, 8 e 9 mostram os gráficos de índices das três medidas divergência para os dados A, B, F e H, respectivamente. Claramente, observamos que todas as medidas identificam os casos influentes.

2.2.6 Análise de resíduos

Quando se procura ajustar um modelo a um conjunto de dados, a validação desse ajuste passa pela análise de uma estatística especial, que têm finalidade de medir a qualidade do modelo ajustado. Uma vez escolhido o modelo e realizado o ajuste, a análise de resíduos é uma maneira eficiente de verificar a qualidade do ajuste do modelo. Os resíduos também servem para outros propósitos, como por exemplo, detectar a presença de pontos aberrantes, identificar a relevância de um fator adicional omitido no modelo, assim como verificar se há indícios de afastamentos sérios da distribuição assumida para o erro.

Tabela 6 – Medidas de ψ -divergência para os dados simulados com ajuste do modelo WE

Conjunto de dados	Identificação dos casos	Medidas de divergência		
		d_{K-L}	d_J	d_{L_1}
A	25	0,0050	0,0100	0,0401
	150	0,0028	0,0057	0,0300
	225	0,0155	0,0312	0,0702
B	25	7,6356	13,4340	0,9523
C	150	6,5771	11,5979	0,9426
D	225	7,0510	12,0335	0,9352
E	25	3,2524	5,8476	0,7939
	150	1,1009	2,1237	0,5379
F	25	3,0605	6,0520	0,8067
	225	2,7802	5,5341	0,7852
G	150	1,4175	2,8145	0,6120
	225	4,0247	7,1761	0,8459
H	25	1,5909	3,0451	0,6263
	150	0,6040	1,1972	0,4204
	225	1,5199	2,9159	0,6156

Figura 6 – Gráficos de índices das medidas de divergência para o caso A

Por outro lado, sabe-se que os resíduos são usados para identificar discrepâncias entre o modelo ajustado e o conjunto de dados, logo é conveniente procurar uma definição para o resíduo que leve em conta a contribuição de cada observação sobre a medida usada de qualidade de ajuste.

Em síntese, os resíduos medem a qualidade do ajuste do modelo para cada observação e permitem estudar se as diferenças entre os valores observados e ajustados são devido ao acaso ou a um comportamento sistemático que pode ser modelado. Para

Figura 7 – Gráficos de índices das medidas de divergência para o caso B

Figura 8 – Gráficos de índices das medidas de divergência para o caso F

definir o resíduo, pode-se usar a metodologia proposta por Dunn e Smyth (1996) para obter o resíduo quantil. Dessa forma, o resíduo quantil para o modelo WE é definido por

$$t_i = \Phi^{-1}\{[1 - \exp(-z_i)]^\delta\},$$

em que $\Phi(\cdot)$ é a distribuição acumulada da Normal padrão.

Atkinson (1985), sugere que para poder interpretar melhor o gráfico normal de probabilidades dos resíduos, este deve ser auxiliado pela construção de envelopes, que são bandas de confiança simuladas que cobrem os resíduos. Esses envelopes são tais que, se o modelo estiver bem ajustado, a maioria dos pontos devem estar dentro dessas bandas e de forma aleatória.

Figura 9 – Gráficos de índices das medidas de divergência para o caso H

A construção das bandas de confiança segue os seguintes passos

- i. Ajustar o modelo proposto e calcular os resíduos t_i ;
- ii. Simular k amostras da variável resposta utilizando o modelo ajustado;
- iii. Fazer o ajuste do modelo para cada amostra, e calcular os resíduos t_{ij} , $j = 1, 2, \dots, k$ e $i = 1, 2, \dots, n$;
- iv. Colocar cada grupo de n resíduos em ordem crescente, obtendo $t_{(i)j}$ com $j = 1, 2, \dots, k$ e $i = 1, 2, \dots, n$;
- v. Para cada i obter a média, mínimo e máximo de $t_{(i)j}$, isto é,

$$t_{(i)M} = \sum_{j=1}^k \frac{t_{(i)j}}{k}, \quad t_{(i)I} = \min\{t_{(i)j} : 1 \leq j \leq k\} \quad \text{e} \quad t_{(i)S} = \max\{t_{(i)j} : 1 \leq j \leq k\};$$

- vi. Incluir as médias, mínimos e máximos conjuntamente com os valores t_i contra os percentis esperados da distribuição normal padrão.

O mínimo e o máximo dos t_i formam o envelope. Se o modelo estimado é correto, espera-se que os valores observados estejam dentro das bandas e de forma aleatória. Usando 19 simulações, quantidade sugerida por Atkinson (1985), existe 1 chance em 20 ou 5% de que o maior valor absoluto dos resíduos do conjunto de dados originais esteja fora do envelope simulado, quando o modelo ajustado for o correto. Esse resultado pode ser usado também para detectar a presença de pontos aberrantes.

2.2.7 Aplicação

Para ilustrar a aplicação do MLWEG foi utilizado um conjunto de dados reais na área de seguro.

Os valores pagos por indenizações pelas seguradoras em geral concentram-se em uma faixa de valores mais baixos e em poucos casos as indenizações assumem valores mais altos caracterizando uma distribuição com assimetria à direita dos dados. Várias são as distribuições com assimetria à direita e suporte positivo que podem ser usadas no ajuste desses dados. Dessa forma a aplicação da função taxa de falha é uma importante ferramenta na seleção de modelos, pois muitas vezes distribuições com comportamento da função taxa de falha totalmente diferentes possuem densidades semelhantes.

Na linguagem da análise de sobrevivência, a função taxa de falha representa o risco de uma observação falhar dado que sobreviveu até o tempo t . Na área de seguro a função taxa de falha pode ser interpretada como o risco por unidade monetária da seguradora de uma apólice ter de pagar ao segurado logo após o pagamento de uma quantia específica de dinheiro (Marshall e Olkin, 2007).

É possível extrair informações qualitativas sobre a função taxa de falha a partir das observações disponíveis por meio de análise gráfica. Aarset (1987) propôs um gráfico denominado tempo total em teste (gráfico TTT), obtido plotando os pontos $(r/n, G(r/n))$ obtidos a partir da expressão

$$G(r/n) = \frac{(n-r)T_{r:n} + \sum_{i=1}^r T_{i:n}}{\sum_{i=1}^r T_{i:n}},$$

em que $r = 1, \dots, n$ e $T_{i:n}$, $i = 1, \dots, n$ são as estatísticas de ordem das observações.

As diferentes formas que o gráfico TTT pode assumir são apresentadas na Figura 10. Quando a curva, obtida do gráfico TTT, é uma reta diagonal (curva A), a função taxa de falha é constante; quando a curva é convexa (curva B) ou côncava (curva C), a função taxa de falha é monotonicamente decrescente ou crescente, respectivamente; quando a curva é convexa e então côncava (curva D), a função taxa de falha têm forma de U, e no caso reverso (curva E) é unimodal.

Os modelos WE e Weibull são modelos encaixados, e portanto, o teste da razão de verossimilhança é útil para realizar comparação da adequabilidade entre esses modelos. A estatística do teste é obtida resolvendo a expressão $w = 2[\ell(\hat{\theta}) - \ell(\tilde{\theta})]$ sob a

Figura 10 – Gráfico TTT-plot

hipótese nula $H_0 : \delta = 1$ contra $H_1 : \delta \neq 1$, em que $\tilde{\theta}$ e $\hat{\theta}$ são as estimativas de máxima verossimilhança sob H_0 e H_1 , respectivamente e $\ell(\hat{\theta})$ e $\ell(\tilde{\theta})$ são os valores do logaritmo da função de verossimilhança maximizado do MLWEG e do MLWG, respectivamente. A estatística w tem distribuição aproximada Qui-quadrado com um grau de liberdade, correspondente a diferença entre os parâmetros dos dois modelos.

Para o cálculo das medidas de influência local, utilizadas para detectar possíveis pontos influentes, foi utilizado a linguagem de programação Ox versão 6.20 (Doornik, 2007) e as estimativas dos parâmetros dos modelos foram obtidas com o auxílio do *software* estatístico SAS - Statistical Analysis System (Institute, 2004).

2.2.7.1 Valores pagos por danos com acidentes pessoais

Para essa análise considerou-se os últimos 18 meses de estudo (Janeiro de 1998 a Junho de 1999), segurados com representação legal totalizando 542 observações e a covariável tempo operacional. Inicialmente realizou-se a análise exploratória dos dados, cujos resultados são apresentados na Tabela 7. O histograma para esse conjunto de dados é apresentado na Figura 11a. Dos resultados apresentados na Tabela 7 e da Figura 11 observa-se assimetria à direita dos dados indicando que o ajuste por uma distribuição assimétrica é o mais adequado. O gráfico TTT para esse conjunto de dados é apresentado na Figura 11b, na qual observa-se uma curva inicialmente côncava e em seguida convexa, entorno da linha diagonal de referência, indicando que a forma geral da função taxa de falha é unimodal e, portanto, o modelo linear WE generalizado é indicado para ajustar-se

aos dados.

É interessante ressaltar que se o gráfico TTT apresentasse uma curva inicialmente convexa e depois côncava, a função taxa de falha teria a forma de U e logo a distribuição Weibull, novamente, não seria indicada para o ajuste dos dados, uma vez que não modela forma de U, ao contrário da distribuição WE.

Tabela 7 – Estatísticas descritivas do conjunto de dados valores pagos por danos com acidentes pessoais

Mediana	Média	Desvio padrão	Coef. de variação	Assimetria	Curtose	Amplitude	Min	Máx	n
6765,24	8996,41	8790,80	97,71 %	5,59	54,32	116481	109	116570	542

(a)

(b)

Figura 11 – Histograma (a) e gráfico TTT (b) para os dados de valores pagos por danos com acidentes pessoais

As estatísticas AIC (Critério de Informação Akaike), CAIC (Critério Akaike de Informação Consistente) e BIC (Critério de Informação Bayesiano) foram calculadas (sem considerar a covariável tempo operacional) para os modelos Weibull e WE obtendo-se os valores AIC = 10885, CAIC = 10885 e BIC = 10893 para o modelo Weibull e AIC = 10871, CAIC = 10871 e BIC = 10883 para o modelo WE. Posteriormente, foi calculado a estatística da razão de verossimilhança para testar as hipóteses $H_0 : \delta = 1$ contra $H_1 : \delta \neq 1$ e comparar os modelos WE e Weibull, obtendo $w = 2[-5432, 5 - (-5440, 5)] = 8$ (p-valor = 0,0047). Analisando as estatísticas AIC, CAIC e BIC, observa-se que o MLWEG resultou nos menores valores para essas estatísticas e pelo teste de hipótese H_0 não deve ser aceita. Portanto, o modelo mais indicado para ajustar-se aos dados é o modelo WE.

2.2.7.2 Estimação de máxima verossimilhança

O modelo de regressão, considerando a distribuição WE, é dado por

$$y_i = g^{-1}(\beta_0 + \beta_1 x_i) \epsilon_i, \quad i = 1, \dots, 542$$

em que y_i é o valor pago da i -ésima observação, x_i é o tempo operacional associado a i -ésima observação, ϵ_i é o erro aleatório seguindo distribuição $WE(1, \phi, \delta)$ e $g^{-1}(\cdot)$ é a função de ligação inversa.

A Figura 12, apresenta o gráfico de dispersão da covariável tempo operacional, contra a variável resposta (valor do seguro) Figura 12a e o logaritmo neperiano da variável resposta Figura 12b, na qual observa-se que o logaritmo do valor pago cresce linearmente com o tempo operacional e a variabilidade é estabilizada. Portanto, adotou-se a função de ligação logarítmica $\log(\lambda_i) = \eta_i$ para ajuste do MLWEG. Logo, o modelo de regressão adotado é dado por

$$y_i = e^{\beta_0 + \beta_1 x_i} \epsilon_i.$$

Figura 12 – Gráfico de dispersão da covariável tempo operacional (x_i) contra a resposta valor do seguro (y_i) Figura (a) e gráfico de dispersão de x_i contra $\log(y_i)$, Figura (b)

As estatísticas AIC, CAIC e BIC, as estimativas dos parâmetros, o erro padrão (EP) e o nível descritivo (p-valor), para os modelos de regressão WE e Weibull são apresentadas na Tabela 8, na qual se observa que o modelo WE obteve novamente

os menores valores para as estatísticas AIC, CAIC e BIC em comparação com o modelo Weibull. Os parâmetros da regressão foram significativos em ambos os modelos e portanto devem permanecer no estudo.

Tabela 8 – EMV dos parâmetros para os modelos de regressão linear generalizado Weibull e Weibull exponenciado

Modelo	δ	ϕ	β_0	β_1
WE	6,6226	0,6859	8,0552	0,01769
EP	1,6332	0,06234	0,2047	0,003589
p-valor	-	-	$< 0,0001$	$< 0,0001$
AIC = 10738	CAIC = 10739	BIC = 10736		
Modelo	-	ϕ	β_0	β_1
Weibull	-	1,4022	9,1651	0,02726
EP	-	0,04124	0,05271	0,003632
p-valor	-	-	$< 0,0001$	$< 0,0001$
AIC = 10832	CAIC = 10832	BIC = 10845		

Os gráficos da Figura 13 apresentam as possíveis observações influentes no modelo de regressão WE. Para os gráficos de influência local (M(0) contra índices das obsevações), foi considerado os esquemas de perturbação da verossimilhança (Figura 13a), da variável resposta (Figura 13b) e da covariável tempo operacional (Figura 13c). As observações 10 e 28 aparecem nos três esquemas de perturbação como possíveis pontos influentes e portanto, serão retiradas do conjunto de dados individualmente e em conjunto e reestimados os parâmetros do modelo para estudar a influência dessas observações nas estimativas dos parâmetros. Essas observações são os menores valores pagos por danos com acidentes pessoais pelas seguradoras Australianas dentro da amostra estudada ($y_{10} = 253,26$ e $y_{28} = 109$), com tempo operacional $x_{10} = 0,4$ e $x_{28} = 1,1$.

Na Tabela 9 é apresentado a mudança relativa (em porcentagens) de cada estimativa dos parâmetros, após a retirada das observações 10, 28 e (10 e 28), definida por $RC_{\theta_j} = [|\hat{\theta}_j - \hat{\theta}_{j(i)}|/\hat{\theta}_j] \times 100$, em que $\hat{\theta}_{j(i)}$ é a estimativa de máxima verossimilhança retirada a i -ésima observação. Observa-se que a significância dos parâmetros da regressão β_0 e β_1 não se alterou, evidenciando a robustez do modelo. Dessa forma não existe mudança na inferência após a remoção das possíveis observações influentes (observações 10, 28 e [10 e 28]).

Figura 13 – Gráfico de $M(0)$ contra índices para os parâmetros do MLGWE sobre os esquemas de perturbações: (a) da verossimilhança, (b) da resposta e (c) da covariável

Tabela 9 – Mudanças relativas [RC em %], parâmetros estimados e o p-valor em parênteses quando as observações 10, 28 e (10 e 28) são eliminadas, respectivamente

Exclusão	$\hat{\delta}$	$\hat{\phi}$	$\hat{\beta}_0$	$\hat{\beta}_1$
Nenhuma	[-]	[-]	[-]	[-]
	1,6332	0,06234	0,2047	0,003589
	(-)	(-)	(< 0,0001)	(< 0,0001)
10	[17,9597]	[4,4175]	[1,4376]	[6,1617]
	2,0763	0,06172	0,2263	0,003575
	(-)	(-)	(< 0,0001)	(< 0,0001)
28	[37,9445]	[8,7768]	[3,0142]	[9,1577]
	2,6763	0,06255	0,2574	0,003573
	(-)	(-)	(< 0,0001)	(< 0,0001)
10 e 28	[84,2947]	[15,5708]	[5,8583]	[18,2024]
	4,1634	0,06292	0,3140	0,003579
	(-)	(-)	(< 0,0001)	(< 0,0001)

2.2.7.3 Análise Bayesiana

Realizou-se uma análise de sensibilidade quanto a escolhas das distribuições *a priori*. Diferentes valores para os componentes de variância das distribuições *a priori* foram consideradas: *Priori 1*: $\sigma_{\beta_0}^2 = \sigma_{\beta_1}^2 = \sigma_{\delta}^2 = \sigma_{\phi}^2 = 50$; *Priori 2*: $\sigma_{\beta_0}^2 = \sigma_{\beta_1}^2 = \sigma_{\delta}^2 = \sigma_{\phi}^2 = 100$; *Priori 3*: $\sigma_{\beta_0}^2 = \sigma_{\beta_1}^2 = \sigma_{\delta}^2 = \sigma_{\phi}^2 = 1000$ e *Priori 4*: $\sigma_{\beta_0}^2 = \sigma_{\beta_1}^2 = \sigma_{\delta}^2 = \sigma_{\phi}^2 = 10000$.

Na Tabela 10 os resumos *a posteriori* dos parâmetros dos modelos MLWEG e MLWG são apresentados. Os resultados obtidos não apresentam notável diferença quanto a escolha da *priori*.

De modo geral, comparando os modelos MLWEG e MLWG, uma vez que na Tabela 10 a estimativa de δ seja diferente de 1, $1 \notin (4.1627, 10.3994)$, temos uma indicação favorável para o modelo MLWEG. Também, com base em todos os critérios bayesianos apresentados na Tabela 11, há evidências positivas em favor do modelo MLWEG.

Também foram obtidas todas medidas divergência ψ descritas na Subseção 2.2.5. A Tabela 12 apresenta os resultados para as quatro observações que apresentaram os maiores valores, sendo que os casos 28 e 539 são possivelmente observações influentes. A observação 28 é o menor valor pago para indenização por danos com acidentes pessoais pelas seguradoras Australianas na amostra estudada ($y_{28} = 109$), com tempo operacional $x_{28} = 1,1$ enquanto que a observação 539 é o maior valor ($y_{539} = 116586,72$, $x_{539} = 31,4$).

Os resultados apresentados a seguir foram obtidos assumindo a *Priori 2*. As Figuras 14, 15, 16 e 17 mostram a aproximação marginal da densidade *a posteriori* para os parâmetros do modelo e o gráfico dos traços considerando as 10.000 observações

Tabela 10 – Dados reais. Resumos *a posteriori* dos parâmetros dos modelos MLWEG e MLWG

	Parâmetro	MLWEG			MLWG		
		Média	DP	IC (95%)	Média	DP	IC (95%)
<i>Priori 1</i>	δ	6,633	1,605	(4,162, 10,399)	-	-	-
	ϕ	0,693	0,060	(0,581, 0,815)	1,398	0,041	(1,318, 1,479)
	β_{01}	8,066	0,195	(7,655, 8,414)	9,166	0,053	(9,063, 9,273)
	β_{11}	0,017	0,003	(0,010, 0,024)	0,027	0,003	(0,020, 0,034)
<i>Priori 2</i>	δ	6,809	1,716	(4,171, 10,838)	-	-	-
	ϕ	0,688	0,062	(0,571, 0,815)	1,398	0,041	(1,318, 1,480)
	β_{01}	8,046	0,205	(7,608, 8,414)	9,166	0,053	(9,064, 9,273)
	β_{11}	0,017	0,003	(0,010, 0,024)	0,027	0,003	(0,020, 0,034)
<i>Priori 3</i>	δ	6,966	1,835	(4,248, 11,395)	-	-	-
	ϕ	0,683	0,062	(0,563, 0,810)	1,398	0,040	(1,318, 1,478)
	β_{01}	8,029	0,213	(7,567, 8,404)	9,167	0,052	(9,066, 9,272)
	β_{11}	0,017	0,003	(0,010, 0,024)	0,027	0,003	(0,020, 0,034)
<i>Priori 4</i>	δ	7,031	1,874	(4,248, 11,479)	-	-	-
	ϕ	0,681	0,063	(0,561, 0,809)	1,398	0,040	(1,317, 1,480)
	β_{01}	8,022	0,216	(7,556, 8,399)	9,167	0,053	(9,064, 9,275)
	β_{11}	0,017	0,003	(0,010, 0,024)	0,027	0,003	(0,020, 0,034)

Tabela 11 – Critérios Bayesianos

	Modelo	Critérios			
		EAIC	EBIC	DIC	LPML
<i>Priori 1</i>	GEWLM	10742,440	10759,620	10738,190	-5370,800
	GWLM	10835,162	10848,048	10832,194	-5418,818
<i>Priori 2</i>	GEWLM	10742,486	10759,667	10738,244	-5370,903
	GWLM	10835,130	10848,016	10832,132	-5418,637
<i>Priori 3</i>	GEWLM	10742,531	10759,712	10738,262	-5371,059
	GWLM	10835,110	10847,990	10832,080	-5418,720
<i>Priori 4</i>	GEWLM	10742,539	10759,721	10738,252	-5370,985
	GWLM	10835,137	10848,023	10832,141	-5418,749

geradas. A Figura 18 apresenta o gráfico de índices das três medidas de divergência.

A Figura 19 apresenta os gráficos de índices das medidas divergência para dados de pagamento de seguro por acidente pessoal retirando as observações 28 e 359. Para todas as medidas de divergência, podemos observar que após excluir as duas observações, não foram detectados mais pontos influentes.

A Tabela 13 apresenta a mudança relativa (RC) de cada estimativa dos parâmetros após remover as observações 28, 539, e (28 e 539), e o intervalo de credibilidade de 95% para os parâmetros. A RC (em percentagem) de cada parâmetro estimado é de-

Tabela 12 – Medidas de divergência

Identificação dos casos		Medidas de divergência		
		d_{K-L}	d_J	d_{L_1}
<i>Priori 1</i>	10	0,2577	0,5386	0,2842
	28	0,7266	1,5899	0,4687
	369	0,2566	0,5543	0,2824
	539	0,4666	0,9917	0,3830
<i>Priori 2</i>	10	0,2688	0,5583	0,2918
	28	0,7482	1,5887	0,4759
	369	0,2567	0,5462	0,2835
	539	0,4785	1,0395	0,3847
<i>Priori 3</i>	10	0,2929	0,6127	0,3031
	28	0,8646	1,9346	0,5083
	369	0,2523	0,5327	0,2816
	539	0,4668	1,0065	0,3809
<i>Priori 4</i>	10	0,2821	0,5820	0,2981
	28	0,7976	1,6793	0,4882
	369	0,2548	0,5432	0,2838
	539	0,4658	1,0026	0,3796

Figura 14 – Gráfico dos traços e aproximação da densidade marginal para o parâmetro δ

finido por $RC_{\theta_j} = |(\hat{\theta}_j - \hat{\theta}_{j(I)})/\hat{\theta}_j| \times 100\%$, em que $\hat{\theta}_{j(I)}$ é a notação da média *a posteriori* de θ_j , com $j = 1, \dots, 4$, após o conjunto I de observações ser removido. Podemos observar que uma pequena RC para o parâmetro β_0 e que a significância dos parâmetros da regressão β_0 e β_1 não mudam, indicando a robustez do modelo. Adicionalmente, segundo todos os critérios Bayesianos de comparação de modelos, o melhor ajuste é obtido quando a observação 28 é retirada, em comparação com o ajuste quando retira-se a observação 539.

Figura 15 – Gráfico dos traços e aproximação da densidade marginal para o parâmetro ϕ

Figura 16 – Gráfico dos traços e aproximação da densidade marginal para o parâmetro β_0

Tabela 13 – RC (em %) e o correspondente intervalo de credibilidade (IC) com 95% para ajuste do modelo WE aos dados reais retirando os casos influentes

Parâmetros	Observações retiradas					
	28		539		{28; 539}	
	RC	IC (95%)	RC	IC (95%)	RC	IC (95%)
δ	33,6288	(5,2704; 15,3334)	11,9081	(3,7309; 9,5729)	16,4957	(4,6816; 13,0944)
ϕ	7,5214	(0,5248; 0,7590)	6,3285	(0,6074; 0,8661)	1,6903	(0,5573; 0,8077)
β_0	2,6203	(7,3133; 8,2615)	1,5869	(7,7516; 8,5185)	0,8646	(7,4998; 8,3779)
β_1	8,5665	(0,0092; 0,0233)	9,2832	(0,0093; 0,0232)	16,1857	(0,0081; 0,0218)
EAIC	10708,190		10701,520		10667,745	
EBIC	10725,370		10718,693		10684,911	
DIC	10703,800		10697,318		10663,368	
LPML	-5353,220		-5350,178		-5332,747

Figura 17 – Gráfico dos traços e aproximação da densidade marginal para o parâmetro β_1

Figura 18 – Medidas de divergência para o conjunto de dados reais completo

Assim, a Tabela 8 e a Tabela 10 apresentam as estimativas dos parâmetros dos modelos, erro padrão e significância dos parâmetros para as EMVs e Bayesiana, respectivamente. Examinando os valores nessas tabelas, pode-se concluir que as estimativas pelos dois métodos são muito similar. E verifica-se que MLWEG parece ser o mais apropriado para ajustar-se ao conjunto de dados do que o MLWG.

Pode-se notar também que os resultados da análise de diagnóstico em ambas as abordagens foram similares.

A análise de resíduo também é realizada. A Figura 20b apresenta o gráfico normal de probabilidade para os resíduos quantil t_i com envelope simulado, o qual foi usado para verificar a qualidade do ajuste do modelo de regressão. O gráfico normal de

Figura 19 – Medidas de divergência para o conjunto de dados reais retirando as observações 28 e 539

probabilidade foi construído utilizando o conjunto completo de dados, pois os possíveis pontos influentes (10 e 28) não modificaram a significância das estimativas. Observa-se que todos os pontos estão dentro do envelope, o que indica o bom ajuste do modelo WE.

O modelo final, portanto, é dado por

$$\hat{y}_i = e^{8,0552+0,01769x_i} = 3150,13 \times 1,02^{x_i}$$

e dessa forma têm-se que aumentando em 2% o valor pago por danos com acidentes pessoais o tempo operacional aumenta em uma unidade ou seja em 1%.

2.3 Conclusões

O MLWEG traz uma nova abordagem da distribuição WE em modelos de regressão. Ele é eficaz para modelar uma parte da resposta média em dados positivos assimétricos incluindo as variáveis explanatórias por meio de uma função linear nos parâmetros denominada preditor linear ($\boldsymbol{\eta} = \mathbf{X}^T \boldsymbol{\beta}$) em um modelo não linear utilizando uma função de ligação g , tal que $g(\lambda) = \boldsymbol{\eta}$, que liga a parte sistemática do modelo a parte aleatória (Da mesma forma que o modelos lineares generalizados). A estimação dos parâmetros é abordada por dois diferentes métodos: por máxima verossimilhança e sob uma abordagem Bayesiana. Também, usa-se a influência local e Bayesiana para diagnosticar casos influentes para estudar a sensibilidade do modelo MLWEG. Esse modelo é mais flexível que o MLWG proposto por Prudente e Cordeiro (2010), modela formas monótonas

Figura 20 – Resíduo quantil (t_i) contra o índice das observações (a), gráfico normal de probabilidade para o resíduo quantil com envelope simulado (b)

e não monótonas da função taxa de falha e têm o MLWG como submodelo.

Na aplicação apresentada nesse trabalho o MLWEG se mostrou mais adequado ao ajuste que o MLWG, obtendo menores valores de AIC, BIC e CAIC. O modelo é robusto a pequenas perturbações nos dados e no modelo e todos os valores para o resíduo proposto permaneceram dentro das bandas de confiança do envelope simulado construído, indicando a qualidade do ajuste.

Referências

AARSET, M. V. How to identify bathtub rate. **Transactions on Realibility**, New York, v. 36, p. 106-108, 1987.

AHMAD, N.; ISLAM, A.; SALAM A. Analysis of optimal accelerated life test plans for periodic inspection: the case of exponentiated Weibull failure model. **International Journal of Quality & Reliability Management**, Bradford, v. 23, p. 1019-1046, 2006.

ATKINSON, A. C. **Plots, transformations and regressions**. Oxford: Statistical Sciences Series Oxford, 1985. 296 p.

BARGHOUT, M. An exponentiated Weibull software reliability model. **Advances and Applications in Statistics**. Allahabad, v. 13, p. 111-130, 2009.

BROOKS, S. P.; CATCHPOLE, E. A.; MORGAN, B. J. T.; HARRIS, M. Bayesian methods for analysing ringing data. **Journal of Applied Statistics**, Leeds, v. 29, p. 187-206, 2002.

CARLIN, B. P.; LOUIS, T. A. **Bayes and empirical bayes methods for data analysis**. 2nd ed., Boca Raton: Chapman and Hall, 2001. 274 p.

CHAUBEY, Y. P. Modeling of distribution for excess-of-loss insurance data. In: INTERNATIONAL CONGRESS ON INSURANCE: MATHEMATICS AND ECONOMICS, 11., 2007, Greece. **11th International Congress on insurance: Mathematics and Economics**. Greece: University of Piraeus, 2007. p. 10-12.

CHATTERJEE, S.; HADI, A. S. **Sensitiv Analysis in Linear Regression**. New York: John Wiley, 1988. 315 p. CHOUDHURY, A. A Simple Derivarion of Moments of the Exponentiated Weibull Distribution. **Metrika**, Heidelberg, v. 62, p. 17-22, 2005.

CHRISTENSEN, R.; PEARSON, L. M.; JOHNSON, W. Case-deletion diagnostics for mixed models. **Technometrics**, Gainesville, v. 34, p. 38-45, 1992.

COOK, R. D. Detection of influential observations in linear regression. **Technometrics**, Gainesville, v. 19, p. 15-18, 1977.

COOK, R. D. Assessment of local influence (with discussion). **Journal of the Royal Statistical Society, Series B**, Malden, v. 48, p. 133-169, 1986.

COOK, R.D.; WEISBERG, S. **Residuals and Influence in Regression**. New York: Chapman and Hall, 1982. 240 p.

COOK, R. D.; PEÑA, D.; WEISBERG, S. The likelihood displacement: A unifying principle for influence, **Communications in Statistics - Theory and Methods**, Philadelphia, v. 17, p. 623-640, 1988.

COX, D. R.; REID, N. Parameter orthogonality and aproximate conditional inference. **Journal of the Royal Statistical Society, Series B**, Malden, v. 49, p. 1-39, 1987.

DAVISON, A. C.; TSAI, C. L. Regression model diagnostics. **International Statistical Review**, Malden, v. 60, p. 337-353, 1992.

DEY, D.; BIRMIWAL, L. Robust Bayesian analysis using divergence measures. **Statistics & Probability Letters**, Amsterdam, v. 20, n. 4, p. 287-294, 1994.

DOORNIK, J. **Ox**: Object oriented matrix programming using Ox. London: Timberlake Consultants Ltd, 2007. 461 p.

DUNN, K. P.; SMYTH, G. K. Randomized quantile residuals. **Journal of Computational and Graphical Statistics**. Melbourne, v. 5, p. 236-244, 1996.

GAMERMAN, D.; LOPES, H.F. **Markov Chain Monte Carlo**: Stochastic Simulation for Bayesian Inference. 2.ed. Boca Raton: Chapman and Hall, 2006. 245 p.

GELFAND, A.E.; DEY, D.K.; CHANG, H. Model determination using predictive distributions with implementation via sampling-based methods. **Technical Report**, Palo Alto, v. 4, p. 147-167, 1992.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. **Bayesian Statistics**, London, p. 169-188, 1992.

HASHIMOTO, E. M.; ORTEGA, E. M. M.; CANCHO, V. G.; CORDEIRO, G. M. On estimation and diagnostics analysis in log-generalized gamma regression model for interval-censored data. **Statistics: A Journal of Theoretical and Applied Statistics**, Dortmund, v. 47, p. 379-398, 2013.

IBRAHIM, J. G.; CHEN, M. H.; SINHA, D. Bayesian Survival Analysis. **International Journal of Epidemiology**, New York, v. 31, n. 2, p. 479, 2001.

INSTITUTE, S. **SAS/STAT User's Guide Version 9**. Cary: SAS Institute, 2004. 4995 p.

JONG, P.; HELLER, G. Z. **Generalized linear models for insurance data**. Cambridge: Cambridge University Press, 2008. 208 p.

MARSHALL, A. W.; OLKIN, I. **Life Distributions. Structure of Nonparametric, Semiparametric and Parametric Families**, New York: Springer, 2007. 788 p.

MUDHOLKAR, G. S.; HUTSON, A. D. The exponentiated Weibull family: some properties and a flood data application. **Communications in Statistics Theory and Methods** 25, Philadelphia, v. 25, p. 3059-3083, 1996.

MUDHOLKAR, G.S.; SRIVASTAVA, D.K. Exponentiated Weibull family for analyzing bathtub failure-ratedata. **IEEE Transactions on Reliability**, Dallas, v. 42, p. 299-302, 1993.

MUDHOLKAR, G.S.; SRIVASTAVA, D.K.; FREIMER, M. The exponentiated Weibull family; a reanalysis of the bus motor failure data. **Technometrics**, Toronto, v. 37, p. 436-445, 1995.

NADARAJAH, S.; CORDEIRO, G. M.; ORTEGA, E. M. M. The exponentiated Weibull distribution: a survey, **Statistical Papers** Heidelberg, v. 54, p. 839-877, 2013.

ORTEGA, E. M. M.; CANCHO, V. G.; LACHOS, V. H. A Generalized Log-gamma Mixture Model for Cure Rate: Estimation and sensitivity analysis. **Sankhy: The Indian Journal of Statistics, Series B**, New York, v. 71, p. 1-29, 2009.

ORTEGA, E. M. M.; CORDEIRO, G. M.; KATTAN, M. W. The log-beta Weibull regression model with application to predict recurrence of prostate cancer, **Statistical Papers (1988)**, Heidelberg, v. 54, p. 113-132, 2013.

PAULA, G. A. On diagnostics in double generalized linear models. **Computational Statistics and Data Analysis**, Oviedo, v. 68, p. 44-51, 2013.

PENG, F.; DEY, D. Bayesian analysis of outlier problems using divergence measures. **The Canadian Journal of Statistics - La Revue Canadienne de Statistique**, Malden, v. 23, p. 199-213, 1995.

POON, W.; POON, Y. S. Conformal normal curvature and assessment of local influence. **Journal of the Royal Statistical Society, Series B**, Malden, v. 61, p. 51-61, 1999.

PRUDENTE, A. A.; CORDEIRO, G. M. Generalized weibull linear models.

Communications in Statistics Theory and Methods, Philadelphia, v. 39, p. 3739-3755, 2010.

R Development Core Team. **R**: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing, 2011. 3452 p.

RUSSO, C. M.; CYSNEIROS, F. J. A.; AOKI, R.; PAULA, G. A. Influence diagnostics in heteroscedastic and/or autoregressive nonlinear elliptical models for correlated data. **Journal of Applied Statistics**, Leeds, v. 39, p. 1049-1067, 2012.

SILVA, G. O.; ORTEGA, E. M. M.; CANCHO, V. G.; BARRETO, M. L. Log-Burr XII regression models with censored data. **Computational Statistics and Data Analysis**, Amsterdam, v. 52, p. 3820-3842, 2008.

SPIEGELHALTER, D.J.; BEST, N.G.; CARLIN, B.P.; VAN DER LINDE, A. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society Series B**, Malden, v. 64, p. 583-639, 2002.

SURLES, J. G.; D'AMBROSIO, D. M. A Burr type X chain-of-links model for carbon fibers. **Journal of Composite Materials**, Los Angeles, v. 38, p. 1337-1343, 2004.

SUZUKI, A. K.; CANCHO, V. G.; LOUZADA, F. The Poisson-Inverse-Gaussian regression model with cure rate: A bayesian approach and its case influence diagnostics. **Statistical Papers**, Heidelberg, v. 57, p. 133-159, 2016.

VANEGAS, L. H.; RONDÓN, L. M.; CORDEIRO, G. M. Diagnostic tolls in generalized Weibull linear regression models. **Journal of Statistical Computation and Simulation**, Blacksburg, v. 83, n. 12, p. 2315-2338, 2013.

WANG, M.; RENNOLLS, K. Tree diameter distribution modelling: introducing the logit-logistic distribution. **Canadian Journal of Forest Research**, Ottawa, v. 35, p. 1305-1313, 2005.

WEISS, R. An approach to Bayesian sensitivity analysis. **Journal of the Royal Statistical Society Series B**, Malden, v. 58, n. 4, p. 739-750, 1996.

XIE, F. C.; WEI, B. C. Diagnostics analysis for log-Birnbaum-Saunders regression models. **Computational Statistics & Data Analysis**, Oviedo, v. 51, p. 4692-4706, 2007.

ZHANG, T.; XIE, M.; TANG, L. C.; NG, S. H. Reliability and modeling of systems integrated with firmware and hardware. **International Journal of Reliability Quality and Safety Engineering**, Singapore, v. 12, p. 227-239, 2005.

3 MODELO LINEAR BETA WEIBULL GENERALIZADO: ESTIMAÇÃO, DIAGNÓSTICO E ANÁLISE DE RESÍDUOS

Resumo

Neste artigo é proposto um novo modelo de regressão baseado nos conceitos dos modelos lineares generalizados (MLG) e assumindo distribuição beta Weibull (BW) reparametrizada, que não pertence a família exponencial, como componente aleatório. Semelhante aos MLGs o modelo proposto, denominado modelo linear beta Weibull generalizado (MLBWG), é composto por três componentes estruturais importantes: o componente aleatório definido pela distribuição BW, o componente sistemático que inclui as covariáveis no modelo por meio de uma estrutura linear e o componente de ligação que conecta a parte sistemática à aleatória através de uma função conhecida. Um estudo sobre o comportamento e flexibilidade da BW é realizado e as expressões para o logaritmo da função de verossimilhança, função score e matriz de informação de Fisher esperada e observada são apresentadas. A estimação de máxima verossimilhança dos parâmetros admitindo três funções de ligação distintas é discutida utilizando dados simulados. A metodologia de influência local sobre três esquemas de perturbação: da variável resposta, da verossimilhança e das covariáveis é desenvolvida para o MLBWG. Adicionalmente, para verificar as suposições do modelo, detectar observações atípicas e para verificar a qualidade do ajuste do modelo de regressão foi proposto os resíduos baseado na função quantílica e realizado estudo de simulação de Monte Carlo para construção de envelopes simulados. Por fim, um conjunto de dados real, na área de engenharia elétrica, foi utilizado para ilustrar a aplicabilidade do MLBWG e seu ajuste comparado aos modelos Weibull e Weibull exponenciado.

Palavras-chave: Modelo Linear Generalizado; Modelo Linear Beta Weibull Generalizado; Influência Local; Análise de Resíduo

Abstract

In this paper we propose a new regression model based on the concepts of generalized linear models (GLM) and assuming beta Weibull distribution (BW) reparameterized, which does not belong to the exponential family, as random component. Similar to MLGs the proposed model, called generalized beta Weibull linear model (GBWLM) is composed of three major structural components: a random component defined by BW distribution, the systematic component that includes the covariates in the model by means of a linear structure, and connection member that connects the systematic part to the random through a known function. A study of the behavior and flexibility of BW is performed and the expressions for the logarithm of the likelihood function, function score and Fisher information matrix expected and observed are presented. The maximum likelihood estimation of the parameters assuming three distinct binding functions is discussed using simulated data. The methodology of local influence on three perturbation schemes: the response variable, the likelihood and covariates is developed for GBWML. Additionally, to check the model assumptions, detect atypical observations and to check the quality of the regression model adjustment was proposed waste based on the quantile function and performed Monte Carlo simulation study for

construction of simulated envelopes. Finally, a set of real data in the area of electrical engineering, was used to illustrate the applicability of GBWML and its adjustment compared to Weibull and exponentiated Weibull models.

Keywords: Generalized Beta Weibull linear model; Generalized linear model; Local influence; Residual

3.1 Introdução

A distribuição beta Weibull (BW) foi proposta por Famoye, Lee e Olumolade (2005) como alternativa a distribuição Weibull e Weibull exponenciada para ajustar-se a dados cuja função taxa de falha é não monótona, tendo como casos particulares a Weibull, exponencial, exponencial exponenciada (EE), Weibull exponenciada (WE), entre outras distribuições. Apesar de recente, a distribuição BW têm sido aplicada nas mais diversas áreas do conhecimento, pode-se citar por exemplo os trabalhos de Lee, Famoye e Olumolade (2007) que reanalisaram o conjunto de dados de falha de motores de ônibus publicado por Davis (1952), considerando a distribuição BW e comparando o ajuste com os ajustes das distribuições exponencial, Weibull e WE, obtendo resultados mais adequado quando a distribuição BW é admitida. Os autores também observaram que a frequência esperada do modelo BW é mais próxima das frequências observadas do que das demais distribuições. Além, de que somente a BW identificou corretamente que a função taxa de falha (ftf) tem forma de U, o que implicou no melhor ajuste em comparação com os demais modelos. Lee, Famoye e Olumolade também analisaram os dados de Efron's (1988) sobre estudo clínico de câncer de cabeça e garganta, obtendo ajuste muito similar ao da WE e melhores que o modelo Weibull.

Ortega, Cordeiro e Kattan (2013), propuseram um modelo de regressão considerando o logaritmo de um variável aleatória seguindo distribuição BW denominado Log beta Weibull (LBW), o qual pode ser aplicado a conjunto de dados censurados e não censurados e possui vários modelos conhecidos na literatura como caso particular. O modelo LBW foi utilizado para prever recorrência de câncer de próstata em pacientes com câncer de próstata localizado clinicamente e tratados por prostatectomia radical aberta, o qual mostrou melhor ajuste que os modelos Log Weibull (LW) e Risco proporcional de Cox segundo os critérios de AIC, BIC e deviance global.

Cordeiro, Nadarajah e Ortega (2013) analisaram três conjunto de dados admitindo o modelo BW e comparando a qualidade do ajuste com seus submodelos (WE,

EE e Weibull). No primeiro caso os autores ajustaram a BW aos dados apresentados por Aarset (1987) sobre o tempo de vida de 50 dispositivos industriais colocados em teste no tempo zero e compararam o ajuste obtido, obtendo os menores valores para as estatísticas AIC, BIC e CAIC para o modelo BW. Adicionalmente, as análises gráficas do ajuste da função de sobrevivência esperada sobre a empírica e do ajuste da função densidade de probabilidade (fdp) sobre o histograma indicaram a BW como o modelo mais adequado para ajustar-se aos dados. Para o segundo conjunto de dados sobre o tempo de sobrevivência de Golden shiner (*Notemigonus crysolucas*) obtido do experimento de campo conduzido no Lago Saint Pierre em Quebec no ano de 2005, (Laplante-Albert. 2008) o modelo BW novamente foi o que apresentou melhor ajuste em comparação aos modelos analisados. Por fim, o último conjunto de dados considerou o modelo de regressão Log beta Weibull (LBW) para estudar os dados de Efron (1988) sobre um estudo clínico de câncer de garganta e cabeça divididos em dois grupos. O grupo A tinha 51 pacientes tratados com terapia de radiação e o grupo B com 45 tratados com radiação mais quimioterapia. O ajuste do modelo de regressão LBW foi comparado ao ajuste do modelo LW. A estatística da razão de verossimilhança (LR) foi altamente significativa ($< 0,0001$) dando indicações favoráveis ao modelo LBW.

Ortega et. al. (2015) propuseram o modelo de regressão beta Weibull em Série de Potência (Power Series Beta Weibull - PSBW) com fração de cura para estimar a mortalidade por carcinoma mamária contra o tempo e a probabilidade de cura em mulheres que foram tratadas com mastectomia e linfonodectomia axilar, em que um número desconhecido de causas concorrentes, que podem influenciar o tempo de sobrevivência, é assumido seguir distribuição Série de Potência (PS) e o tempo das células carcinoma mamiárias restantes, depois da mastectomia para a metástase, é assumido seguir distribuição BW. Concluindo que a distribuição BW é mais flexível que outros modelos existentes para estimar os efeitos das variáveis explanatórias na função de sobrevivência e na proporção dos curados e a composição do modelo produz estimativas mais precisas dos efeitos das variáveis explanatórias sobre os tempos de recorrência.

A proposta deste trabalho é desenvolver uma modelo de regressão estatístico baseado na estrutura dos MLGs, porém tendo como componente aleatório a BW que não pertence a família exponencial. Estendendo a abrangência dos MLGs à família de distribuições BW e seus submodelos.

Este artigo está organizado em 8 capítulos, sendo que nos dois primeiros capítulos uma revisão sobre a distribuição beta Weibull é realizada apresentando os artigos e aplicações mais recentes, além de um estudo sobre o comportamento das funções densidade de probabilidade (fdp) e taxa de falha da BW. O modelo linear beta Weibull generalizado (MLBWG) é definido no capítulo 3 apresentando as expressões para o logaritmo da verossimilhança, função escore e matriz de informação de Fisher esperada e observada. Um estudo de simulação, utilizando a metodologia de Monte Carlo para avaliar as estimativas de máxima verossimilhança, a análise de influência local, para o modelo proposto, considerando os esquemas de perturbação da verossimilhança, da variável resposta e das covariáveis e a definição do resíduo quantil para análise de resíduo juntamente com a construção do envelope simulado são desenvolvidos nos capítulos 4, 5 e 6. Por fim, um conjunto real de dados na área de engenharia elétrica, sobre tempo de falha de filmes de tereftalato de polietileno em transformadores isolados com gás SF₆ é utilizado para ilustrar a aplicação do MLBWG.

3.2 A distribuição beta Weibull

A distribuição beta Weibull foi proposta por Famoye, Lee e Olumolade (2005) e baseou-se na generalização da distribuição Normal proposta por Eugene, Lee e Famoye (2002) denominada beta Normal. Qualquer distribuição contínua, pode ser generalizada segundo a proposta de Eugene, Lee e Famoye, após algumas operações matemáticas. O teorema a seguir apresenta o método para generalizar uma distribuição contínua qualquer $\mathcal{G}$ adicionando os parâmetros δ e b .

Teorema: Se T é uma variável aleatória (v.a.) com distribuição $\mathcal{G}$ e função distribuição acumulada (fda) $G(t)$, então $F(t)$ é a fda resultante da generalização da distribuição $\mathcal{G}$, denominada beta $\mathcal{G}$. Obtida da expressão:

$$F(t) = I_{G(t)}(\delta, b) = \frac{1}{B(\delta, b)} \int_0^{G(t)} w^{\delta-1} (1-w)^{b-1} dw \quad (19)$$

em que $B(\delta, b) = \Gamma(\delta + b)/[\Gamma(\delta)\Gamma(b)]$ é a função beta, $\Gamma(\alpha) = \int_0^\infty w^{\alpha-1} \exp(-w) dw$ é a função gama, $I_y(\delta, b) = B(\delta, b)^{-1} \int_0^y w^{\delta-1} (1-w)^{b-1} dw$ é a função razão beta incompleta e $\delta > 0$ e $b > 0$ são dois parâmetros de forma adicionais.

Prova: Seja a fda da distribuição beta

$$H(x) = \frac{1}{B(\delta, b)} \int_0^x \omega^{\delta-1} (1-\omega)^{b-1} d\omega, \quad 0 \leq x \leq 1$$

e seja $G(t)$ uma fda contínua e domínio indicado por $Dom(G)$. Composto as funções G e H tal que $F(t) = H \circ G(t) = H(G(t))$ tem-se que

1. $\lim_{t \rightarrow \infty} F(t) = \lim_{t \rightarrow \infty} H(G(t)) = H(1) = 1$, pois $\lim_{t \rightarrow \infty} G(t) = 1$ e $H(1) = 1$;
2. $\lim_{t \rightarrow -\infty} F(t) = \lim_{t \rightarrow -\infty} H(G(t)) = H(0) = 0$, pois $\lim_{t \rightarrow -\infty} G(t) = 0$ e $H(0) = 0$;
3. Sejam $t_1, t_2 \in Dom(G)$ tal que $t_1 < t_2$ segue que $F(t_1) = H(G(t_1)) < H(G(t_2)) = F(t_2)$ uma vez que $G(t_1) < G(t_2) \Rightarrow H(G(t_1)) < H(G(t_2))$;
4. $\forall a \in Dom(G) \lim_{t \rightarrow a^+} F(t) = \lim_{t \rightarrow a^+} H(G(t)) = H(G(a)) = F(a)$, pois G é contínua a direita. ■

Em particular, se $T > 0$ segue distribuição Weibull com parâmetros α e ϕ ($T \sim \text{Weibull}(\alpha, \phi)$) e fda

$$G(t; \alpha, \phi) = 1 - \exp \left[- \left(\frac{t}{\alpha} \right)^\phi \right] \quad \text{com } t > 0 \text{ e } \alpha, \phi > 0, \quad (20)$$

em que ϕ é o parâmetro de forma e α o parâmetro de escala, a distribuição acumulada da BW ($F(t)$) é obtida da equação (19) tal que

$$F(t) = I_{G(t; \alpha, \phi)}(\delta, b) = \frac{1}{B(\delta, b)} \int_0^{1 - \exp[-(t/\alpha)^\phi]} w^{\delta-1} (1-w)^{b-1} dw, \quad t > 0. \quad (21)$$

A função densidade de probabilidade (fdp) da BW é expressa por

$$f(t) = \frac{\phi}{B(\delta, b)} t^{-1} \left(\frac{t}{\alpha} \right)^\phi \exp \left[-b \left(\frac{t}{\alpha} \right)^\phi \right] \left\{ 1 - \exp \left[- \left(\frac{t}{\alpha} \right)^\phi \right] \right\}^{\delta-1} \quad (22)$$

com $t, \alpha, \phi, \delta, b > 0$ em que α é o parâmetro de escala e ϕ e δ são os parâmetros de forma e b um parâmetro adicional.

A inclusão dos parâmetros de forma $\delta > 0$ e $b > 0$ trouxe grande flexibilidade a distribuição permitindo a fdp modelar uma grande variedade de formas. O comportamento da fdp da BW pode ser estudado analisando a mudança de sinal da função:

$$T(z) = -bz + (1 - 1/\phi) + (\delta - 1)I(z) \quad (23)$$

em que $I(z) = ze^{-z}/(1 - e^{-z})$ e $z = (t/\alpha)^\phi$.

Dessa forma, se I_0 , I_1 e I_2 são subconjuntos do $Dom(f)$, apresentada em (22), tal que $I_0 \cup I_1 \cup I_2 = Dom(f)$ e $I_0 \cap I_1 \cap I_2 = \emptyset$. Tem-se que:

1. Se $T(z) > 0$ em $z \in I_0$ então a fdp é crescente em $z \in I_0$;
2. Se $T(z) = 0$ em $z \in I_1$ então a fdp tem um ponto de máximo ou mínimo em $z \in I_1$;
3. Se $T(z) < 0$ em $z \in I_2$ então a fdp é decrescente em $z \in I_2$.

Da equação (23) e das considerações acima segue o seguinte teorema.

Teorema: Seja T uma v.a. com distribuição $BW(\alpha, \phi, \delta, b)$ então se:

- i. $\phi \leq 1$ e $\delta \leq 1$ então a fdp é decrescente;
- ii. $\phi = 1$ e $\delta > 1$ então a fdp é unimodal;
- iii. $\phi > 1$ e $\delta = 1$ então a fdp é unimodal.

Prova:

- i. Se $\phi \leq 1$ e $\delta \leq 1$ então (23) se resume a $T(z) = -bz + (1 - 1/\phi)$ em que $T(z) < 0 \forall z \in Dom(f)$.
- ii. Se $\phi = 1$ e $\delta > 1$ então (23) se resume a $T(z) = -bz + (\delta - 1)I(z)$. Como $0 < I(z) < 1$ e $I(z) \rightarrow 1$ quando $z \rightarrow 0$ então $T(z) > 0 \forall z \in (0, z_0 = \log[\delta/b - 1/b + 1])$, $T(z_0) = 0$ e $T(z) < 0 \forall z \in (z_0, \infty)$;
- iii. Se $\phi > 1$ e $\delta = 1$ então (23) se resume a $T(z) = -bz + (1 - 1/\phi)$ em que $T(z) > 0 \forall z \in (0, z_0 = 1/b - b^{-1}/\phi)$, $T(z_0) = 0$ e $T(z) < 0 \forall z \in (z_0, \infty)$.

■

Figura 21 – Formas da fdp da BW, fixando os parâmetros $\alpha = 1$, ϕ e δ e variando b . Os parâmetros foram fixados em (a) $\phi = 0.3$ e $\delta = 0.5$, (b) $\delta = 2$ e (c) $\phi = 3$

A Figura 21 apresenta algumas formas da fdp da BW para valores fixos de α , ϕ e δ e variando o parâmetro b . As Figuras 21a, 21b e 21c apresentam as formas da fdp para $(\phi \leq 1, \delta \leq 1)$, $(\phi = 1, \delta > 1)$ e $(\phi > 1, \delta = 1)$. Em todos os casos o parâmetro α foi fixando em $\alpha = 1$ (por se tratar de um parâmetro de escala e não influenciar na forma da densidade). Nas três figuras observa-se que o parâmetro b não influencia na forma da curva (decrecente ou unimodal) porém, contribui suavizando a curva (formas decrecente) ou deslocando o ponto de máximo da curva (formas unimodais).

Fixando $b = 1$, e restringindo os parâmetros ϕ e δ tais que $\phi\delta < 1$, $\phi\delta = 1$ e $\phi\delta > 1$ a fdp assume forma decrecente e ilimitada na sua calda esquerda, decrecente com ordenada da cauda esquerda igual a $1/\alpha$ e unimodal com uma pequena cauda à

esquerda, respectivamente. A Figura 22a ilustra os três casos para valores escolhidos de ϕ e δ . A Figura 22b apresenta formas da fdp da BW para valores dos parâmetros não descritos acima, exemplificando a grande flexibilidade da distribuição BW. Destaca-se a quase bimodalidade obtida na Figura 22b quando $\alpha = 48$, $\phi = 6$, $\delta = 0.10$ e $b = 0.1$.

Figura 22 – Formas da fdp da BW. (a): Forma decrescente se $b = 1$ e $\phi\delta \leq 1$ e unimodal se $b = 1$ e $\phi\delta > 1$. (b) Formas decrescente, unimodal e quase bimodalidade para alguns pontos selecionados do espaço paramétrico $\theta = \{(\alpha, \phi, \delta, b) : b \neq 1, \phi > 1 \text{ e } \delta > 1\}$

Outra característica importante da BW é que ela possui vários submodelos, conhecidos na literatura, como casos particulares, por exemplo: para $\phi = 1$, $\phi = b = 1$ e $\phi = \delta = b = 1$ obtém-se as distribuições beta exponencial (BE), exponencial exponenciada (EE) e exponencial, respectivamente. Fixando os parâmetros em $b = 1$, $\delta = b = 1$ e $\delta = 1$ têm-se as distribuições Weibull exponenciada (WE), Weibull e Weibull com parâmetros $\lambda b^{-1/\phi}$ e ϕ . Se $\phi = 2$ e $b = 1$ a distribuição Burr tipo X é obtida como caso particular. A substituição dos parâmetros ϕ , δ e b pelas constantes 2, 1 e 1 em (22), respectivamente, resulta na fdp da Rayleigh.

Os parâmetros δ e b não somente trouxeram maior flexibilidade para a fdp, mas tornou possível o ajuste da distribuição BW a conjuntos de dados cuja a ftf têm forma de U, além de modelar todas as formas que a Weibull já modela.

A ftf de uma variável aleatória $T > 0$, é definida por $h(t) = f(t)/[1 - F(t)]$ em que f e F são as funções de densidade e acumulada, respectivamente. O gráfico da ftf apresenta algumas formas básicas conforme pode ser visualizado na Figura 23a, com diferentes interpretações práticas. O gráfico pode ser crescente ou decrescente conforme

indicado por $h_1(t)$ e $h_2(t)$ respectivamente, constante representado por $h_3(t)$, ou alternar entre intervalos decrescentes, constantes e crescentes, como por exemplo $h_4(t)$, que é uma forma da ftf conhecida por forma de U ou forma de banheira e $h_5(t)$ conhecida por forma unimodal, que inicia com um intervalo crescente até atingir um máximo e então decresce.

Figura 23 – (a) Formas gerais para a função taxa de falha, (b) As quatro regiões do espaço paramétrico da WE determinadas pelas equações $\phi = 1$ e $\phi\delta = 1$ com as respectivas descrições das formas da função taxa de falha

Glaser (1980) apresenta um método para estudar as formas da ftf analisando a variação do sinal da expressão $\eta'(t) = [f'(t)/f(t)]^2 - f''(t)/f(t)$ e estudando as raízes de $g(t) = [1 - F(t)]/f(t)$ em que $f(t)$ é a fdp de uma distribuição contínua, duas vezes diferenciável, com suporte em $(0, \infty)$ e $f'(t)$ e $f''(t)$ são as derivadas primeira e segunda de f em relação a t respectivamente. Os resultados obtidos por Glaser são apresentados a seguir:

- a. Se $\eta'(t) > 0$ para $t > 0$ então a ftf é crescente;
- b. Se $\eta'(t) < 0$ para $t > 0$ então a ftf é decrescente;
- c. Suponha que exista $t_0 > 0$ tal que $\eta'(t) < 0$ para $t \in (0, t_0)$, $\eta'(t_0) = 0$ e $\eta'(t) > 0$ para $t > t_0$;
 - i. Se existir $y_0 > 0$ tal que $g'(y_0) = 0$, então a ftf tem forma de U;
 - ii. Se não existir $y_0 > 0$ tal que $g'(y_0) = 0$, então a ftf tem forma crescente;

- d. Suponha que exista $t_0 > 0$ tal que $\eta'(t) > 0$ para $t \in (0, t_0)$, $\eta'(t_0) = 0$ e $\eta'(t) < 0$ para $t > t_0$;
- i. Se existir $y_0 > 0$ tal que $g'(y_0) = 0$, então a ftf tem forma unimodal;
- ii. Se não existir $y_0 > 0$ tal que $g'(y_0) = 0$, então a ftf tem forma decrescente.

A expressão de $\eta'(t)$ para a distribuição BW é da forma:

$$\eta'(t) = t^{-2} \{ (1 + \phi)\phi bz + (1 - \phi) - \phi(\delta - 1)J(z) \} \quad (24)$$

em que $J(z) = I(z)[1 + z - z^2 - I(z)]$, $I(z) = ze^{-z}/(1 - e^{-z})$ e $z = (t/\alpha)^\phi$.

Fixando $b = 1$ em (24) obtém-se a expressão de $\eta'(t)$ para a distribuição WE, cujas formas da ftf já são conhecidas e obtidas de quatro regiões do espaço paramétrico determinadas pelas equações $\phi = 1$ e $\delta\phi = 1$, veja Figura 23b. Fixando ϕ e b o termo $(1 + \phi)\phi bz + (1 - \phi)$ em (24) se torna uma função linear crescente em z e portanto a introdução do parâmetro b em (24) não altera o sentido do sinal de $\eta'(t)$, uma vez que $J(z)$ é uma função limitada que tende a zero quando $t \rightarrow \infty$. No entanto, b funciona como coeficiente angular mudando a inclinação de $(1 + \phi)\phi bz + (1 - \phi)$ e portanto o ponto t_0 tal que $\eta'(t_0) = 0$. Assim sendo o parâmetro b serve como um parâmetro de escala para a ftf da distribuição BW, como é ilustrado na Figura 24.

3.2.1 Modelo linear beta Weibull generalizado

Prudente e Cordeiro (2010) publicaram um artigo intitulado *Generalized Weibull Linear Models* (Modelos lineares Weibull generalizado), onde propuseram uma nova classe de modelo de regressão com estrutura análoga aos modelos lineares generalizados, com base na distribuição Weibull, que não pertence a família exponencial.

O MLWG, como foi nomeado pelos autores, permite um novo uso da distribuição Weibull em uma estrutura muito semelhante aos modelos lineares generalizados. O modelo proposto foi definido como uma síntese de três componente estruturais importantes em MLG: o componente aleatório para a variável resposta seguindo distribuição Weibull parametrizada de forma conveniente, um componente sistemático especificado pelo preditor linear do modelo e por uma função adequada ligando uma parte relevante da média da distribuição e o componente sistemático.

Os MLWGs são muito efetivos, conforme afirmam os autores, na modelagem

Figura 24 – Estudo da influência do parâmetro b na função taxa de falha da BW, fixando $\alpha = 1$, ϕ , δ e variando b

de uma parte da resposta média em termos de uma equação de regressão para dados positivos contínuos e a definição do resíduo e muitas outras medidas de diagnóstico podem seguir de perto a teoria dos MLG.

Sejam $Y_1, Y_2, \dots, Y_n$, n v.a.'s independentes, com cada Y_i seguindo distribui-

ção Weibull com parâmetro de forma ϕ e parâmetro de escala α_i o qual varia de acordo com as observações, com fda e fdp dadas, respectivamente, por

$$F(y_i; \alpha_i, \phi) = 1 - \exp \left[- \left(\frac{y_i}{\alpha_i} \right)^\phi \right] \quad (25)$$

e

$$f(y_i; \alpha_i, \phi) = \phi y_i^{-1} \left(\frac{y_i}{\alpha_i} \right)^\phi \exp \left\{ - \left(\frac{y_i}{\alpha_i} \right)^\phi \right\} \quad \text{com } \phi, \alpha_i, y_i > 0, \quad (26)$$

com média e variância

$$E(Y_i) = \alpha_i \Gamma_1(\phi) \quad \text{e} \quad Var(Y_i) = \alpha_i^2 \{ \Gamma_2(\phi) - \Gamma_1^2(\phi) \}$$

em que $\Gamma_j(\phi) = \Gamma \left(1 + \frac{j}{\phi} \right)$ para $j \geq 1$, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ é a função Gama e $\Gamma'(x) = \partial \Gamma(x) / \partial x$.

Reparametrizando a equação (25) de acordo com a relação $\lambda_i = \alpha_i \exp \{ \phi^{-1} \Gamma'(2) \}$, proposta por Cox e Reid (1987), obtém-se ortogonalidade entre os parâmetros ϕ e λ_i , ou seja, $E(\partial^2 \ell(\boldsymbol{\theta}) / \partial \phi \partial \lambda_i) = E(\partial^2 \ell(\boldsymbol{\theta}) / \partial \lambda_i \partial \phi) = 0$. A vantagem em se ter parâmetros ortogonais consiste na independência assintótica das estimativas de máxima verossimilhança para ϕ e λ_i e na simplificação do desenvolvimento algébrico, uma vez que a matriz de informação esperada é diagonal. Dessa forma, as expressões reparametrizadas para a fda, fdp, esperança e variância são dadas por

$$F(y_i; \lambda_i, \phi) = 1 - \exp \left[- \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \right], \quad (27)$$

$$f(y_i; \lambda_i, \phi) = \phi y_i^{-1} \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \exp \left\{ - \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \right\} \quad \text{com } \phi, \lambda_i, y_i > 0, \quad (28)$$

$$E(Y_i) = \lambda_i c_1(\phi) \quad \text{e} \quad Var(Y_i) = \lambda_i^2 c_2(\phi)$$

em que $c_1(\phi) = \exp \{ -\Gamma'(2)/\phi \} \Gamma_1(\phi)$ e $c_2(\phi) = \exp \{ -2\Gamma'(2)/\phi \} \{ \Gamma_2(\phi) - \Gamma_1^2(\phi) \}$.

O MLWG é definido pela equação (28) (a fdp da Weibull reparametrizada) representando o componente aleatório, por uma função g aplicada em λ_i tal que $g(\lambda_i) = \eta_i$ correspondendo a função de ligação e $\eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$ o componente sistemático, em que $\mathbf{x}_i$ é

o vetor de covariáveis associado a i -ésima observação e β é um vetor de parâmetros desconhecido.

Substituindo $G(t; \alpha, \phi)$ em (21) pela equação (27) obtém-se a distribuição BW com um parâmetro de escala λ_i e três parâmetros de forma ϕ , δ e b . Daqui em diante, a distribuição beta Weibull reparametrizada será denotada simplesmente por BW e indicaremos que uma v.a. Y segue distribuição BW por $Y \sim BW(\lambda_i, \phi, \delta, b)$, cuja fdp é definida por

$$f(y_i) = \frac{\phi}{B(\delta, b)} y_i^{-1} \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \exp \left[-b \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \right] \left\{ 1 - \exp \left[- \left(\frac{y_i}{\lambda_i} \right)^\phi e^{\Gamma'(2)} \right] \right\}^{\delta-1} \quad (29)$$

Note que se $\phi = 1$, $\delta = 1$ e $b = 1$ obtém-se a distribuição exponencial, se $\phi = 1$ e $b = 1$ obtém-se a distribuição EE, se $\phi = 1$ a distribuição BE e se Y_i segue distribuição BW então a variável aleatória $A_i = (Y_i/\lambda_i)^\phi e^{\Gamma'(2)}$ segue distribuição BE com parâmetro de escala igual a um e parâmetros de forma δ e b . O teorema a seguir apresenta o resultado descrito de forma mais geral.

Teorema: Se $Y \sim BW(\lambda, \phi, \delta, b)$, com reparametrização

$$f(y; \lambda, \phi, \delta, b) = \frac{\phi}{B(\delta, b)} y^{-1} \left(\frac{y}{\lambda} \right)^\phi k \exp \left[-b \left(\frac{y}{\lambda} \right)^\phi k \right] \left\{ 1 - \exp \left[- \left(\frac{y}{\lambda} \right)^\phi k \right] \right\}^{\delta-1},$$

em que $k > 0$ e $Z = (Y/\lambda)^\phi k$ então $Z \sim BE(1, \delta, b)$.

Prova: Seja t um ponto no suporte da BW, então

$$\begin{aligned} F_Z(t) &= P(Z < t) = P \left(k \left(\frac{Y}{\lambda} \right)^\phi < t \right) = P \left(Y < \lambda \left(\frac{t}{k} \right)^{1/\phi} \right) = \\ &= F_Y \left(\lambda \left(\frac{t}{k} \right)^{1/\phi} \right) = \frac{1}{B(\delta, b)} \int_0^{1-\exp(-t)} w^{\delta-1} (w-1)^{b-1}. \end{aligned}$$

■

A função geradora de momentos da BW é obtida segundo a expressão ¹

$$\mu'_k = \frac{\lambda_i^k}{B(\delta, b)} \exp \left\{ -k \frac{\Gamma'(2)}{\phi} \right\} \Gamma_k(\phi) \left[b^{-(1+\frac{k}{\phi})} + \sum_{i=1}^{\infty} a_i (i+b)^{-(1+\frac{k}{\phi})} \right] \quad (30)$$

¹A Prova da convergência da série $\sum_{j=1}^{\infty} a_j (j+b)^{-(1+k/\phi)}$ é analoga a prova feita por Choudhury, (2005) para $b = 1$.

$$a_i = (-1)^i \binom{\delta - 1}{i} = (-1)^i \frac{\delta^1(\delta^1 - 1)(\delta^1 - 2) \dots (\delta^1 - i + 1)}{i!}, \quad \delta^1 = \delta - 1.$$

Portanto, os dois primeiros momentos ordinários da BW são

$$\mu'_1 = E(Y_i) = \frac{\lambda_i}{B(\delta, b)} \exp \left\{ -\frac{\Gamma'(2)}{\phi} \right\} \Gamma_1(\phi) \left[b^{-(1+\frac{1}{\phi})} + \sum_{i=1}^{\infty} a_i(i+b)^{-(1+\frac{1}{\phi})} \right]$$

e

$$\mu'_2 = E(Y_i^2) = \frac{\lambda_i^2}{B(\delta, b)} \exp \left\{ -2\frac{\Gamma'(2)}{\phi} \right\} \Gamma_2(\phi) \left[b^{-(1+\frac{2}{\phi})} + \sum_{i=1}^{\infty} a_i(i+b)^{-(1+\frac{2}{\phi})} \right]$$

e logo, a variância e a esperança são expressas por:

$$E(Y_i) = \lambda_i c_1(\phi, \delta, b) \quad \text{e} \quad Var(Y_i) = \lambda_i^2 [h_2(\phi, \delta, b) - h_1^2(\phi, \delta, b)] = \lambda_i^2 c_2(\phi, \delta, b)$$

em que

$$c_1(\phi, \delta, b) = h_1(\phi, \delta, b) = \frac{e^{-\frac{\Gamma'(2)}{\phi}}}{B(\delta, b)} \Gamma_1(\phi) \left[b^{-(1+\frac{1}{\phi})} + \sum_{i=1}^{\infty} a_i(i+b)^{-(1+\frac{1}{\phi})} \right]$$

e

$$h_2(\phi, \delta, b) = \frac{e^{-2\frac{\Gamma'(2)}{\phi}}}{B(\delta, b)} \Gamma_2(\phi) \left[b^{-(1+\frac{2}{\phi})} + \sum_{i=1}^{\infty} a_i(i+b)^{-(1+\frac{2}{\phi})} \right]$$

Observa-se que o MLWG, proposto por Prudente e Cordeiro (2010), é um caso particular quando se têm $\delta = 1$ e $b = 1$, além de similaridades entre o trabalho apresentado por Prudente e Cordeiro (2010), pois da mesma forma como discutem os autores o parâmetro λ_i é um tipo de parâmetro de locação o qual também afeta a variância e a função de variância da distribuição BW e em termos da nova parametrização é uma função quadrática da média.

O MLBWG fica então definido pela fdp em (22), especificando o componente aleatório para a variável resposta y_i e por uma função g conhecida, duas vezes diferenciável e bijetora que liga o parâmetro λ_i ao preditor η_i , denominada, portanto, de função de ligação, tal que:

$$g(\lambda_i) = \eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$$

em que $\mathbf{x}_i$ é o vetor $1 \times p$ de covariáveis relacionadas a i -ésima variável resposta y_i e

$\beta^T = (\beta_1, \dots, \beta_p)$ é o vetor de parâmetros da regressão a ser estimado. O parâmetro λ_i pode ser explicitado pela inversa da função de ligação, ou seja: $\lambda_i = g^{-1}(\eta_i) = g^{-1}(\mathbf{x}_i^T \beta)$ e portanto $\mu_i = E(Y_i) = g^{-1}(\eta_i) c_1(\phi, \delta, b)$ em que se percebe que parte da resposta média para a i -ésima variável é explicada por λ_i , além de influenciar na variância do modelo, pois a função de variância da distribuição BW é uma função quadrática da média $Var(Y_i) = \lambda_i^2 h_2(\phi, \delta, b) - \mu_i^2$.

O comportamento de $c_1(\phi, \delta, b)$ e $c_2(\phi, \delta, b)$ é apresentado, respectivamente, nas Figuras 25 e 26 para $b \in \{0.1, 0.5, 1, 2.5, 5\}$ e variando os parâmetros ϕ e δ .

Visualizando a Figura 25 nota-se que a medida que os valores de ϕ e δ se aproximam da origem (ponto $(0, 0)$ do plano $\phi\delta$) $c_1(\phi, \delta, b)$ cresce, atingindo valores maiores a medida que b aproxima-se de 0 (Figura 25 para os casos de $b = 0.1$ e $b = 0.5$) e decresce aproximando de zero a medida que b aumenta (Figura 25 para os casos de $b = 5$ e $b = 10$). Entretanto, se ϕ e δ aumentam c_1 aproxima-se de 1 independentemente do valor de b . Esse comportamento influencia diretamente na modelagem da resposta média, ora atribuindo peso maior à parte aleatória sobre a média ora a parte sistemática, dependendo dos valores atribuídos (estimados) para ϕ , δ e b .

A variação de $c_2(\phi, \delta, b)$ pode ser observada na Figura 26 onde percebe-se que c_2 assume valores maiores para ϕ próximo de zero e conforme distancia-se da origem c_2 diminui tendendo a zero. Os maiores valores para c_2 são obtidos a medida que b aproxima-se de 0. A variabilidade é portanto influenciada pelo comportamento de $c_2(\phi, \delta, b)$, dependendo dos valores atribuídos (estimados) para ϕ , δ e b .

3.2.1.1 Modelo de regressão

Sabe-se que na prática existem muitas situações em que a variável resposta é influenciada por uma ou mais variáveis regressoras. Essas variáveis regressoras também são chamadas de variáveis explicativas ou covariáveis que podem ser: tratamentos, características intrínsecas das unidades amostrais, variáveis exógenas, interações, variáveis dependentes do tempo e etc.

Seja $\mathbf{x}_i^T = (x_{i1}, x_{i2}, \dots, x_{ip})$ um conjunto de p covariáveis associadas a i -ésima variável resposta y_i para $i = 1, \dots, n$. Se $Y_1, \dots, Y_n$ é uma sequência de n variáveis aleatórias independentes, tais que $Y_i \sim BW(\lambda_i, \phi, \delta, b)$ e $g(\lambda_i) = \eta_i = \mathbf{x}_i^T \beta$ é a função de

$b = 0.1$ $b = 0.5$ $b = 1.0$ $b = 2.5$ $b = 5.0$ $b = 10$

Figura 25 – Gráficos das funções $c_1(\phi, \delta, b)$

ligação, então a equação que melhor descreve o modelo de regressão é

$$Y_i = \lambda_i \varepsilon_i = g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta}) \varepsilon_i$$

com $\varepsilon_i \sim BW(1, \phi, \delta, b)$ e média e variância dadas, respectivamente, por

Figura 26 – Gráficos das funções $c_2(\phi, \delta, b)$

$$\begin{aligned} E(Y_i) &= \lambda_i E(\varepsilon_i) = \lambda_i c_1(\phi, \delta, b) = g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta}) c_1(\phi, \delta, b) \\ \text{Var}(Y_i) &= \lambda_i^2 \text{Var}(\varepsilon_i) = \lambda_i^2 c_2(\phi, \delta, b) = [g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta})]^2 c_2(\phi, \delta, b) \end{aligned}$$

O componente sistemático $\eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$ entra no modelo de regressão na forma de uma estrutura linear para as variáveis explicativas e a ligação entre o componente

sistemático e o aleatório é realizada pela função de ligação g tal que $g^{-1} > 0$. A escolha da função de ligação depende muito do experimento e da relação entre a variável resposta e as covariáveis. Algumas possíveis escolhas para a função de ligação são, por exemplo, a identidade $g(\lambda_i) = \lambda_i$, a recíproca $g(\lambda_i) = 1/\lambda_i$ e a logarítmica $g(\lambda_i) = \log(\lambda_i)$. Dessa forma, o modelo de regressão é composto por uma estrutura linear nos parâmetros em uma função não-linear com erro multiplicativo.

3.2.1.2 Estimação de máxima verossimilhança

Considere agora uma amostra aleatória $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$ de n observações independentes, em que $\mathbf{x}_i^T = (x_{i1}, x_{i2}, \dots, x_{ip})$ com $i = 1, \dots, n$ é o vetor de covariáveis associado ao i -ésimo indivíduo e y_i a variável resposta seguindo distribuição $BW(\lambda_i, \phi, \delta, b)$ com $\lambda_i = g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta})$. O logaritmo da função de verossimilhança para os parâmetros do modelo $\boldsymbol{\theta}^T = (\boldsymbol{\beta}^T, \phi, \delta, b)$ é dado por

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \{-\log[B(\delta, b)] + \log(\phi) - \log(y_i) + \log(z_i) - bz_i + (\delta - 1) \log[1 - \exp(-z_i)]\} , \quad (31)$$

em que $z_i = (y_i/\lambda_i)^\phi e^{\Gamma'(2)}$.

O estimador de máxima verossimilhança $\hat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$ pode ser obtido pela maximização do logaritmo da função de verossimilhança dado em (31), o que implica resolver o sistema de equações não lineares $\partial \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} = \mathbf{0}$ que não possui, em geral, solução analítica, sendo necessário o uso de métodos computacionais iterativos. Portanto, para obtenção das estimativas dos parâmetros do modelo o procedimento NLMixed do *software* estatístico SAS foi utilizado.

Os vetores escore para $\boldsymbol{\beta}$, ϕ , δ e b podem ser escritos, respectivamente, por

$$U_{\boldsymbol{\beta}} = \frac{\partial \ell}{\partial \boldsymbol{\beta}} = \phi \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{s} \quad , \quad U_{\phi} = \frac{\partial \ell}{\partial \phi} = \frac{n}{\phi} + \boldsymbol{\tau}^T \mathbf{1} \quad , \quad U_{\delta} = \frac{\partial \ell}{\partial \delta} = \boldsymbol{\kappa}^T \mathbf{1}$$

$$\text{e} \quad U_b = \frac{\partial \ell}{\partial b} = \boldsymbol{\vartheta}^T \mathbf{1}$$

em que $\mathbf{W} = \text{diag}[(\lambda'_i/\lambda_i)^2]$ é uma matriz diagonal de ordem n . $\lambda'_i = \partial \lambda_i / \partial \eta_i$, $\lambda''_i = \partial^2 \lambda_i / \partial \eta_i^2$ são as derivadas da inversa da função de ligação, $\mathbf{1}$ é um vetor $n \times 1$ de uns e os vetores $\mathbf{s} = (s(z_1), s(z_2), \dots, s(z_n))^T$, $\boldsymbol{\tau} = (\tau_1, \tau_2, \dots, \tau_n)^T$, $\boldsymbol{\kappa} = (\kappa_1, \kappa_2, \dots, \kappa_n)^T$ e

$\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \dots, \vartheta_n)^T$ têm componentes expressos, respectivamente, por:

$$s(z_i) = (bz_i - 1) - (\delta - 1)l_i \quad , \quad \tau_i = -s(z_i) \log(y_i/\lambda_i) \quad ,$$

$$\kappa_i = \Psi_{\delta,b,1}^0 + \log[1 - \exp(-z_i)] \quad , \quad \vartheta_i = \Psi_{\delta,b,1}^0 - z_i$$

e

$$z_i = \left(\frac{y_i}{\lambda_i}\right)^\phi e^{\Gamma'(2)} \quad , \quad l_i = z_i \frac{\exp(-z_i)}{1 - \exp(-z_i)} \quad , \quad \Psi_{l,m,n}^k = \Psi^k(l+m) - n\Psi^k(l) \quad \text{e}$$

$$\Psi^k(x) = \frac{\partial^{k+1}}{\partial x^{k+1}} \log[\Gamma(x)].$$

Sob condições de regularidade, a distribuição assintótica de $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ é a distribuição Normal multivariada $N_{p+3}(0, \mathbf{I}_{\theta\theta}^{-1})$ em que $\mathbf{I}_{\theta\theta}$ é a matriz de informação esperada, apresentada em (32). A matriz de covariância assintótica $\mathbf{I}_{\theta\theta}^{-1}$ pode ser aproximada pela inversa da matriz $(p+3) \times (p+3)$ de informação observada $\ddot{\mathbf{L}}_{\theta\theta}$ e então a inferência assintótica para o vetor de parâmetros $\boldsymbol{\theta}$ pode ser baseada na aproximação normal $N_{p+3}(0, \ddot{\mathbf{L}}_{\theta\theta}^{-1})$ para $\hat{\boldsymbol{\theta}}$. A matriz de informação observada é

$$\ddot{\mathbf{L}}_{\theta\theta} = \begin{bmatrix} \ddot{\mathbf{L}}_{\beta\beta} & \ddot{\mathbf{L}}_{\beta\phi} & \ddot{\mathbf{L}}_{\beta\delta} & \ddot{\mathbf{L}}_{\beta b} \\ \ddot{\mathbf{L}}_{\phi\beta} & \ddot{\mathbf{L}}_{\phi\phi} & \ddot{\mathbf{L}}_{\phi\delta} & \ddot{\mathbf{L}}_{\phi b} \\ \ddot{\mathbf{L}}_{\delta\beta} & \ddot{\mathbf{L}}_{\delta\phi} & \ddot{\mathbf{L}}_{\delta\delta} & \ddot{\mathbf{L}}_{\delta b} \\ \ddot{\mathbf{L}}_{b\beta} & \ddot{\mathbf{L}}_{b\phi} & \ddot{\mathbf{L}}_{b\delta} & \ddot{\mathbf{L}}_{bb} \end{bmatrix} =$$

$$= \begin{bmatrix} -\phi^2 \mathbf{X}^T \mathbf{W} [\mathbf{d}] \mathbf{X} & \mathbf{X}^T \mathbf{W}^{1/2} \underline{\mathbf{U}}_\phi & -\phi \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{L} & \phi \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{Z} \\ - & -\frac{n}{\phi^2} - \frac{1}{\phi^2} \mathbf{M}^T \mathbf{1} & \frac{1}{\phi} \mathbf{B}^T \mathbf{1} & -\frac{1}{\phi} \mathbf{P}^T \mathbf{1} \\ - & - & n\Psi_{\delta,b,1}^1 & n\Psi_{\delta,b,0}^1 \\ - & - & - & n\Psi_{b,\delta,1}^1 \end{bmatrix} \quad ,$$

em que

$$\mathbf{B} = (b_1, \dots, b_n)^T \quad , \quad \mathbf{d} = \text{diag}[d_1, \dots, d_n] \quad , \quad \underline{\mathbf{U}}_\phi = (\underline{u}_1^\phi, \dots, \underline{u}_n^\phi)^T \quad , \quad \mathbf{M} = (m_1, \dots, m_n)^T \quad ,$$

$$\mathbf{L} = (l_1, \dots, l_n)^T \quad , \quad \mathbf{P} = (p_1, \dots, p_n)^T \quad , \quad \mathbf{Z} = (z_1, \dots, z_n)^T \quad ,$$

$$b_i = c_i l_i \quad , \quad c_i = \phi \log \left(\frac{y_i}{\lambda_i} \right) = \log \left(\frac{z_i}{e^{\Gamma(2)}} \right) \quad , \quad f_i = \frac{\lambda_i'' \lambda_i}{(\lambda_i')^2} \quad , \quad p_i = z_i c_i \quad ,$$

$$d_i = \frac{1}{\phi} \{ \phi b z_i - s(z_i)(f_i - 1) - \phi(\delta - 1)l_i[1 - \exp(z_i)l_i] \} \quad ,$$

$$m_i = n_i c_i^2 \quad , \quad n_i = b z_i - (\delta - 1)l_i[1 - \exp(z_i)l_i] \quad , \quad \underline{u}_i^\phi = s(z_i) + n_i c_i \quad .$$

A matriz de informação esperada é

$$\mathbf{I}_{\theta\theta} = \begin{bmatrix} \mathbf{I}_{\beta\beta} & \mathbf{I}_{\beta\phi} & \mathbf{I}_{\beta\delta} & \mathbf{I}_{\beta b} \\ \mathbf{I}_{\phi\beta} & I_{\phi\phi} & I_{\phi\delta} & I_{\phi b} \\ \mathbf{I}_{\delta\beta} & I_{\delta\phi} & I_{\delta\delta} & I_{\delta b} \\ \mathbf{I}_{b\beta} & I_{b\phi} & I_{b\delta} & I_{bb} \end{bmatrix} \quad (32)$$

$$= \begin{bmatrix} \phi^2 \mathbf{X}^T \mathbf{W} \mathbf{V} \mathbf{X} & \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{O} & \phi t_1 \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{1} & \phi \Psi_{b,\delta,1}^0 \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{1} \\ - & -\frac{n}{\phi^2} - \frac{1}{\phi^2} \mathbf{G}^T \mathbf{1} & -\frac{n}{\phi} t_2 & \frac{n}{\phi} t_3 \\ - & - & n \Psi_{\delta,b,1}^1 & n \Psi_{\delta,b,0}^1 \\ - & - & - & n \Psi_{b,\delta,1}^1 \end{bmatrix} \quad , \quad (33)$$

em que

$$\mathbf{V} = \text{diag}(v_1, v_2, \dots, v_n)^T, \quad \mathbf{O} = (o_1, o_2, \dots, o_n)^T, \quad \mathbf{G} = (g_1, g_2, \dots, g_n)^T$$

$$T(d, b, a, e) = \int_0^\infty x^{d-1} e^{-bx} (1 - e^{-x})^{a-1} [\log(x)]^e dx$$

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad , \quad \psi(x) = \frac{\partial \log(\Gamma(x))}{\partial x} \quad , \quad \psi'(x) = \frac{\partial^2 \log(\Gamma(x))}{\partial x^2} \quad ,$$

$$t_1 = \frac{1}{B(\delta, b)} T(2, b+1, \delta-1, 0) \quad , \quad t_2 = T(2, b+1, \delta-1, 1) - \psi(2) T(2, b+1, \delta-1, 0) \quad ,$$

$$t_3 = \frac{T(2, b, \delta, 1)}{B(\delta, b)} - \psi(2) \Psi_{b,\delta,1}^0 \quad , \quad v_i = \frac{\delta-1}{B(\delta, b)} T(3, b+1, \delta-2, 0) + 1 \quad ,$$

$$o_i = -\psi(2) + \frac{1}{B(\delta, b)} \{ b T(2, b, \delta, 1) + (\delta-1) [T(3, b-1, \delta-1, 1) - T(2, b+1, \delta-1, 1) - \\ - \psi(2) T(3, b+1, \delta-2, 0)] \} \quad ,$$

$$g_i = \frac{b - (\delta-1)}{B(\delta, b)} \{ T(2, b, \delta, 2) - 2\psi(2) T(2, b, \delta, 1) + B(\delta, b) [\psi(2)]^2 \} + \frac{\delta-1}{B(\delta, b)} \{ T(3, b+1, \delta-2, 2) -$$

$$-2\psi(2)T(3, b+1, \delta-2, 1) + [\psi(2)]^2 T(3, b+1, \delta-2, 0)\}$$

3.2.2 Estudo de Simulação

Realizou-se um estudo de simulação, utilizando a metodologia de Monte Carlo, para avaliar as estimativas de máxima verossimilhança do modelo MLBWG, considerando o componente sistemático $\eta_i = \beta_0 + \beta_1 x_i, i = 1, \dots, n$, e funções de ligação recíproca ($g(\lambda_i) = 1/\lambda_i$), identidade ($g(\lambda_i) = \lambda_i$) e logarítmica ($g(\lambda_i) = \log(\lambda_i)$).

Os valores para a variável resposta foram simulados a partir da inversa da fda da BW, equações (19) e (27), tal que $Y_i = \lambda_i e^{\Gamma'(2)/\phi} [-\log(1 - w_i)]^{1/\phi}$ com $W_i \sim \text{Beta}(\delta, b)$ (Cordeiro, Nadarajah e Ortega, 2013) e os valores para a covariável X foram gerados pela equação $X_i = \exp(T_i)$ em que $T_i \sim N(0, 1)$. De forma que X apresenta forte assimetria à direita.

Portanto, para cada combinação $n, \beta_0, \beta_1, \phi, \delta, b$ e $g(\lambda_i)$, 1000 amostras aleatórias foram geradas e para cada uma dessas amostras as estimativas de máxima verossimilhança dos parâmetros foram obtidas. Dessas 1000 estimativas de cada parâmetro, a média, o viés e erro quadrático médio (EQM) foram calculados e listados na Tabela 14. As amostras foram geradas no *software* estatístico *R* (R Core Team, 2011) e os parâmetros foram estimados utilizando a função *optim* do *R*.

Para esse estudo, adotou-se quatro tamanhos amostrais ($n = 75, 150, 300$ e 750), as funções de ligação recíproca, identidade e logarítmica e os valores verdadeiros para os parâmetros foram fixados em $\beta_0 = 1,5; \beta_1 = 2,0; \phi = 0,5; \delta = 0,5$ e $b = 1,0$ e $1,5$.

Dos valores apresentados na Tabela 14, observa-se que, em todos os casos, a medida que o tamanho amostral aumenta o EQM e o viés diminuem e as estimativas se aproximam dos verdadeiros valores. Nota-se ainda que o viés e o EQM são bem menores para os parâmetros estimados β_0 e β_1 quando considerado a função de ligação logarítmica, além da convergência para os valores verdadeiros dos parâmetros ser mais rápida.

Tabela 14 – Estimativas média, viés e EQM baseados em 1000 simulações do MLBWG
com $\beta_0 = 1,5$, $\beta_1 = 2,0$, $\phi = 0,5$ e $\delta = 0,5$, $b = 1,0$ e $1,5$

(continua)

$g(\lambda_i)$	n	θ	$b = 1,0$			θ	$b = 1,5$		
			Média	Viés	EQM		Média	Viés	EQM
$1/\lambda_i$	75	β_0	1,5716	0,0716	1,9516	β_0	1,7386	0,2386	3,3973
		β_1	1,8785	-0,1215	1,5285	β_1	1,9369	-0,0631	2,2438
		ϕ	0,5865	0,0865	0,0680	ϕ	0,6100	0,1100	0,0926
		δ	0,5633	0,0633	0,1398	δ	0,5438	0,0438	0,1283
		b	1,1631	0,1631	0,2652	b	1,7542	0,2542	0,5425
$1/\lambda_i$	150	β_0	1,5925	0,0925	2,1974	β_0	1,6335	0,1335	1,3611
		β_1	2,0334	0,0334	1,3811	β_1	1,9636	-0,0364	1,2615
		ϕ	0,5440	0,0440	0,0225	ϕ	0,5618	0,0618	0,0315
		δ	0,5156	0,0156	0,0504	δ	0,4999	-0,0001	0,0378
		b	1,0489	0,0489	0,0906	b	1,6258	0,1258	0,2662
$1/\lambda_i$	300	β_0	1,5346	0,0346	0,5504	β_0	1,6105	0,1105	0,7993
		β_1	1,9950	-0,0050	0,5612	β_1	2,0419	0,0419	0,6513
		ϕ	0,5257	0,0257	0,0082	ϕ	0,5242	0,0242	0,0100
		δ	0,4936	-0,0064	0,0139	δ	0,5044	0,0044	0,0174
		b	1,0148	0,0148	0,0421	b	1,5437	0,0437	0,0843
$1/\lambda_i$	750	β_0	1,5716	0,0716	0,2583	β_0	1,5713	0,0713	0,9051
		β_1	2,0663	0,0663	0,2984	β_1	2,0681	0,0681	1,0038
		ϕ	0,5074	0,0074	0,0027	ϕ	0,5250	0,0250	0,0098
		δ	0,5020	0,0020	0,0053	δ	0,4982	-0,0018	0,0155
		b	0,9988	-0,0012	0,0157	b	1,5457	0,0457	0,0874
λ_i	75	β_0	1,5780	0,0780	2,8003	β_0	1,5995	0,0995	3,3486
		β_1	2,1275	0,1275	3,3990	β_1	2,0302	0,0302	2,9980
		ϕ	0,6171	0,1171	0,0946	ϕ	0,6170	0,1170	0,0941
		δ	0,5417	0,0417	0,1462	δ	0,5425	0,0425	0,1383
		b	1,0047	0,0047	0,1502	b	1,5505	0,0505	0,2522
λ_i	150	β_0	1,6064	0,1064	1,4808	β_0	1,5963	0,0963	1,6967
		β_1	2,0612	0,0612	1,6197	β_1	2,0361	0,0361	1,4970
		ϕ	0,5494	0,0494	0,0246	ϕ	0,5543	0,0543	0,0308
		δ	0,5044	0,0044	0,0358	δ	0,5147	0,0147	0,0481
		b	1,0034	0,0034	0,0532	b	1,5300	0,0300	0,1266
λ_i	300	β_0	1,6110	0,1110	0,8186	β_0	1,5731	0,0731	0,8386
		β_1	2,0883	0,0883	0,9167	β_1	2,0602	0,0602	1,0048
		ϕ	0,5213	0,0213	0,0082	ϕ	0,5205	0,0205	0,0105
		δ	0,5023	0,0023	0,0163	δ	0,5124	0,0124	0,0194
		b	1,0144	0,0144	0,0308	b	1,5308	0,0308	0,0657

Tabela 14 – Estimativas média, viés e EQM baseados em 1000 simulações do MLBWG com $\beta_0 = 1, 5$, $\beta_1 = 2, 0$, $\phi = 0, 5$ e $\delta = 0, 5$, $b = 1, 0$ e $1, 5$

(conclusão)

$g(\lambda_i)$	n	θ	$b = 1, 0$			θ	$b = 1, 5$		
			Média	Viés	EQM		Média	Viés	EQM
λ_i	750	β_0	1,5814	0,0814	0,3409	β_0	1,5348	0,0348	0,3484
		β_1	2,0768	0,0768	0,4260	β_1	2,0879	0,0879	0,3911
		ϕ	0,5067	0,0067	0,0028	ϕ	0,5083	0,0083	0,0033
		δ	0,5045	0,0045	0,0056	δ	0,5022	0,0022	0,0062
		b	1,0174	0,0174	0,0156	b	1,5260	0,0260	0,0309
$\log(\lambda_i)$	75	β_0	1,5661	0,0661	0,4755	β_0	1,4762	-0,0238	0,4603
		β_1	1,9670	-0,0330	0,0405	β_1	1,9637	-0,0363	0,0455
		ϕ	0,6116	0,1116	0,0861	ϕ	0,6313	0,1313	0,1101
		δ	0,5402	0,0402	0,1540	δ	0,5439	0,0439	0,1868
		b	1,0870	0,0870	0,3249	b	1,6071	0,1071	0,3358
$\log(\lambda_i)$	150	β_0	1,5900	0,0900	0,2110	β_0	1,5314	0,0314	0,2267
		β_1	1,9799	-0,0201	0,0175	β_1	1,9827	-0,0173	0,0155
		ϕ	0,5514	0,0514	0,0248	ϕ	0,5560	0,0560	0,0283
		δ	0,5015	0,0015	0,0345	δ	0,5027	0,0027	0,0400
		b	1,0492	0,0492	0,0540	b	1,5651	0,0651	0,1179
$\log(\lambda_i)$	300	β_0	1,6191	0,1191	0,1511	β_0	1,5731	0,0731	0,1538
		β_1	1,9847	-0,0153	0,0076	β_1	1,9928	-0,0072	0,0072
		ϕ	0,5224	0,0224	0,0091	ϕ	0,5268	0,0268	0,0112
		δ	0,5032	0,0032	0,0156	δ	0,5021	0,0021	0,0188
		b	1,0677	0,0677	0,0504	b	1,5793	0,0793	0,0879
$\log(\lambda_i)$	750	β_0	1,6036	0,1036	0,0991	β_0	1,5829	0,0829	0,0849
		β_1	1,9979	-0,0021	0,0026	β_1	1,9959	-0,0041	0,0024
		ϕ	0,5110	0,0110	0,0032	ϕ	0,5085	0,0085	0,0030
		δ	0,4990	-0,0010	0,0057	δ	0,4996	-0,0004	0,0059
		b	1,0553	0,0553	0,0263	b	1,5698	0,0698	0,0480

3.2.3 Análise de sensibilidade

Uma 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) 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. Uma desvantagem da deleção de casos é que ela não permite dizer em que aspecto específico do modelo o caso têm 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). Entretanto, 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 aplicado influência local em modelos de análise de regressão. Recentemente têm-se vários trabalhos publicados, a saber: Silva *et al.* (2008) estuda influência local e modelos de regressão log-Burr XII com dados censurados, Ortega *et al.* (2009) aplica influência local em modelos de regressão log-gama generalizado com fração de cura, Russo *et al.* (2012) aplica influência local em modelo elípticos autorregressivos não-lineares para dados correlacionados, Hashimoto *et al.* (2013) estuda influência local em modelos de regressão log-gama generalizados com censura intervalar, Paula (2013) estuda influência local em modelo linear generalizado duplos, Ortega *et al.* (2013) estuda influência local para o modelo de regressão log-beta Weibull, com aplicação para dados recorrentes de câncer de próstata, dentre outros.

3.2.3.1 Influência local

Quando se está ajustando um modelo a um conjunto de dados, é desejável que as estimativas obtidas a partir do modelo proposto sejam resistentes a pequenas perturbações no modelo ou nas observações. Por isso é importante que se faça um estudo sobre a robustez dos resultados obtidos, em termos dos vários aspectos que envolvem a formulação do modelo, e as estimativas dos seus parâmetros, ou seja, fazer uma análise de diagnóstico, que consiste por exemplo, de métodos para avaliar o grau de sensibilidade das inferências a pequenas perturbações nos dados ou mesmo no modelo proposto.

Enquanto a análise de resíduos é feita para pesquisar problemas com o modelo ajustado, tais como presença de observações aberrantes e afastamento sérios das suposições para a distribuição do erro, uma análise de influência é feita assumindo o modelo como correto, e estuda-se a robustez das conclusões à perturbações nos dados ou no modelo. Uma observação é dita influente se ela produzir alterações desproporcionais nos resultados da análise, quando a mesma for excluída, ou submetida a uma pequena perturbação.

O interesse nessa seção é estudar o comportamento do afastamento da verossimilhança $LD(\boldsymbol{\omega}) = 2[\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})]$ em torno de um vetor de não perturbação, $\boldsymbol{\omega}_0$, tal que $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}_0}$ e, portanto, $LD(\boldsymbol{\omega}_0) = 0$. Segundo Cook (1986), esse estudo pode ser feito usando a curvatura normal que, na direção do vetor $\mathbf{d}$, é expressa como $C_d(\hat{\boldsymbol{\theta}}) = 2|\mathbf{d}^T \Delta^T (-\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}})^{-1} \Delta \mathbf{d}|$ em que $\Delta = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T$ e $\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}} = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$, tanto Δ quanto $\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}$ são avaliados em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega}_0$. Portanto, a curvatura máxima, $C_{d_{max}}$, é obtida na direção $\mathbf{d}_{max}$, em que $C_{d_{max}}$ é o maior autovalor de $\ddot{\mathbf{T}} = \Delta^T (-\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}})^{-1} \Delta$ e $\mathbf{d}_{max}$ é o autovetor correspondente. Alternativamente, pode-se examinar a influência local total $C_i = C_{d_i}(\boldsymbol{\theta})$, em que, $\mathbf{d}_i$ é um vetor $n \times 1$ de zeros com um na i -ésima posição. A metodologia de influência local é talvez a mais popular ferramenta moderna para diagnóstico de influência. Ela têm sido largamente utilizada, nos recentes anos, no diagnóstico de regressão.

Como $C_d(\boldsymbol{\theta})$ não é invariante sobre mudança uniforme de escala, Poon e Poon (1999) propuseram a curvatura normal conformal $B_d(\boldsymbol{\theta}) = C_d(\boldsymbol{\theta}) / \text{tr}\{2\ddot{\mathbf{T}}\}$, a qual é invariante sobre reparametrização de $\boldsymbol{\omega}$ e $0 \leq B_d(\boldsymbol{\theta}) \leq 1$ para qualquer direção unitária $\mathbf{d}$. Dessa forma, para determinar se a i -ésima observação é possivelmente influente, Poon e Poon (1999), analisaram se a quantidade $M(0)_i = B_{d_i}$ associada a i -ésima observação, em que $\mathbf{d}_i$ é um vetor de dimensão $n \times 1$ de zeros com 1 na i -ésima posição, é maior que $\overline{M(0)} + c^* SM(0)$, em que $\overline{M(0)} = 1/n$, $SM(0)$ é o erro padrão amostral de $\{M(0)_k, k = 1, \dots, n\}$ e c^* é uma constante selecionada. Dependendo da aplicação real, c^* pode assumir qualquer valor. A seguir a expressão para a matriz Δ é obtida sobre três diferentes esquemas de perturbação.

• Perturbação da verossimilhança

A perturbação da verossimilhança é realizada atribuindo pesos a cada observação no logaritmo da função de verossimilhança tal que $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \ell(\lambda_i, \phi; y_i)$,

em que $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$ é o vetor de não perturbação. Assim o logaritmo da função de verossimilhança perturbado neste caso é dado por

$$\begin{aligned} \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = & \sum_{i=1}^n \omega_i \{ -\log[B(\delta, b)] + \log(\phi) - \log(y_i) + \log(z_i) - bz_i + \\ & + (\delta - 1) \log[1 - \exp(-z_i)] \} , \end{aligned} \quad (34)$$

em que $z_i = (y_i/\lambda_i)^\phi e^{\Gamma(2)}$. Portanto, para o esquema de perturbação de casos, têm-se que

$$\Delta = \begin{bmatrix} \hat{\phi} \mathbf{X}^T \hat{\mathbf{W}}^{1/2} \underline{\mathbf{s}} \\ (u_1^\phi, \dots, u_n^\phi) \\ (\kappa_1, \dots, \kappa_n) \\ (\vartheta_1, \dots, \vartheta_n) \end{bmatrix} ,$$

em que $\underline{\mathbf{s}} = \text{diag}\{s(\hat{z}_1), \dots, s(\hat{z}_n)\}$, $u_i^\phi = \hat{\phi}^{-1}[1 - s(\hat{z}_i) \log(\hat{z}_i/e^{\Gamma(2)})]$, $\kappa_i = \Psi_{b,1}^0 + \log[1 - \exp(-\hat{z}_i)]$ e $\vartheta_i = \Psi_{\delta,b,1}^0 - \hat{z}_i$.

• Perturbação da variável resposta

A perturbação da variável resposta é realizada de forma aditiva e expressa por $y_{i\omega} = y_i + \omega_i s_y$ em que s_y é o desvio padrão amostral de Y . Consequentemente $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$ e o logaritmo da função de verossimilhança se reduz a

$$\begin{aligned} \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = & \sum_{i=1}^n \{ -\log[B(\delta, b)] + \log(\phi) - \log(y_{i\omega}) + \log(z_{i\omega}) - bz_{i\omega} \\ & + (\delta - 1) \log[1 - \exp(-z_{i\omega})] \} , \end{aligned} \quad (35)$$

em que $z_{i\omega} = (y_{i\omega}/\hat{\lambda}_i)^\phi e^{\Gamma(2)}$. Neste caso, pode-se demonstrar que

$$\Delta = \hat{\phi} s_y \begin{bmatrix} \hat{\phi} \mathbf{X}^T \hat{\mathbf{W}}^{1/2} \underline{\mathbf{s}} \\ (-\underline{u}_1^\phi/\hat{\phi} y_1, \dots, -\underline{u}_n^\phi/\hat{\phi} y_n) \\ (l_1/y_1, \dots, l_n/y_n) \\ (z_1/y_1, \dots, z_n/y_n) \end{bmatrix} ,$$

em que $\underline{s} = \text{diag}\{n_1/y_1, \dots, n_n/y_n\}$, $\underline{u}_i^\phi = s(z_i) + n_i \log(z_i/e^{\Gamma'(2)})$, $n_i = bz_i - (\delta - 1)l_i(1 - e^{z_i}l_i)$, $s(z_i) = (bz_i - 1) - (\delta - 1)l_i$ and $l_i = z_i e^{-z_i}/(1 - e^{-z_i})$.

• Perturbação na variável explanatória

Semelhante ao caso de perturbação da variável resposta, a perturbação da j -ésima variável (contínua) é realizada de forma aditiva e expressa por $x_{ij\omega} = x_{ij} + \omega_i s_{x_j}$, em que s_{x_j} é o desvio padrão amostral da j -ésima variável explanatória. Portanto, $\omega_0 = (0, \dots, 0)^T$ e o logaritmo da função de verossimilhança pode ser expresso por

$$\ell(\theta) = \sum_{i=1}^n \{ -\log[B(\delta, b)] + \log(\phi) - \log(y_i) + \log(z_{i\omega}) - bz_{i\omega} + (\delta - 1) \log[1 - \exp(-z_{i\omega})] \} , \quad (36)$$

em que $z_{i\omega} = (y_i/\hat{\lambda}_{i\omega})^{\hat{\phi}} e^{\Gamma'(2)}$ e $\hat{\lambda}_{i\omega} = g^{-1}(\mathbf{x}_i^T \hat{\beta} + s_{x_j} \hat{\beta}_j \omega_i)$. Assim, para a perturbação da variável explanatória, obtém-se

$$\Delta_{p+3 \times n} = (\Delta_{p \times n}^\beta, \Delta_{1 \times n}^\phi, \Delta_{1 \times n}^\delta, \Delta_{1 \times n}^b)^T ,$$

em que

$$\begin{aligned} \Delta_{ij}^\beta &= \begin{cases} \hat{\phi} \hat{\beta}_t s_{x_j} x_{ij} (\lambda'_i / \lambda_i)^2 \left[(f_i - 1) s(z_i) - \hat{\phi} n_i \right] , & j \neq t , \\ \hat{\phi} \hat{\beta}_t s_{x_j} x_{ij} (\lambda'_i / \lambda_i)^2 \left[(f_i - 1) s(z_i) - \hat{\phi} n_i \right] + \hat{\phi} s_{x_j} (\lambda'_i / \lambda_i) s(z_i) , & j = t , \end{cases} \\ \Delta_i^\phi &= \hat{\phi} \hat{\beta}_t s_{x_j} (\lambda'_i / \lambda_i) \left[\hat{\phi}^{-1} s(z_i) + n_i c_i \right] , \\ \Delta_i^\delta &= -\hat{\phi} \hat{\beta}_t s_{x_j} (\lambda'_i / \lambda_i) l_i \quad \text{and} \\ \Delta_i^b &= \hat{\phi} \hat{\beta}_t s_{x_j} z_i . \end{aligned}$$

3.2.4 Análise de resíduos

Quando se procura ajustar um modelo a um conjunto de dados, a validação desse ajuste passa pela análise de uma estatística especial, que têm finalidade de medir a qualidade do modelo ajustado. Uma vez escolhido o modelo e realizado o ajuste, a análise de resíduos é uma maneira eficiente de verificar a qualidade do ajuste do modelo. Os resíduos também servem para outros propósitos, como por exemplo, detectar a presença de pontos aberrantes, identificar a relevância de um fator adicional omitido no modelo,

assim como verificar se há indícios de afastamentos sérios da distribuição assumida para o erro.

Por outro lado, sabe-se que os resíduos são usados para identificar discrepâncias entre o modelo ajustado e o conjunto de dados, logo é conveniente procurar uma definição para o resíduo que leve em conta a contribuição de cada observação sobre a medida usada de qualidade de ajuste.

Em síntese, os resíduos medem a qualidade do ajuste do modelo para cada observação e permitem estudar se as diferenças entre os valores observados e ajustados são devido ao acaso ou a um comportamento sistemático que pode ser modelado. Para definir os resíduos, pode-se usar a metodologia proposta Dunn e Smyth (1996) para obter os resíduos quantil.

O resíduo quantil para o modelo BW é definido por

$$t_i = \Phi^{-1}\{I_{1-e^{-z_i}}(\delta, b)\},$$

em que $I_{1-e^{-z_i}}(\delta, b) = B(\delta, b)^{-1} \int_0^{1-e^{-z_i}} \omega^{\delta-1} (\omega-1)^{b-1} d\omega$, $\Phi(\cdot)$ é a fda da distribuição normal padrão, e $z_i = (y_i/\lambda_i)^\phi e^{\Gamma'(2)}$.

Note que t_i segue distribuição normal padrão, exceto pela variabilidade amostral em $\lambda_i = g^{-1}(\mathbf{x}_i^T \boldsymbol{\beta})$, ϕ , δ e b . Portanto, a distribuição de t_i convergirá para a normal padrão se $\boldsymbol{\beta}$, ϕ , δ e b são consistentemente estimados.

Atkinson (1985) sugere, para poder interpretar melhor o gráfico normal de probabilidades dos resíduos, que este deve ser auxiliado pela construção de envelopes, que são bandas de confiança simuladas que cobrem os resíduos. Esses envelopes são tais que, se o modelo estiver bem ajustado, a maioria dos pontos devem estar dentro dessas bandas e de forma aleatória.

A construção das bandas de confiança segue os seguintes passos

- i. Ajustar o modelo proposto e calcular os resíduos t_i ;
- ii. Simular k amostras da variável resposta utilizando o modelo ajustado;
- iii. Fazer o ajuste do modelo para cada amostra, e calcular os resíduos t_{ij} , $j = 1, 2, \dots, k$ e $i = 1, 2, \dots, n$;
- iv. Colocar cada grupo de n resíduos em ordem crescente, obtendo $t_{(i)j}$ com $j =$

$1, 2, \dots, k$ e $i = 1, 2, \dots, n$;

v. Para cada i obter a média, mínimo e máximo de $t_{(i)j}$, isto é,

$$t_{(i)M} = \sum_{j=1}^k \frac{t_{(i)j}}{k}, \quad t_{(i)I} = \min\{t_{(i)j} : 1 \leq j \leq k\} \quad \text{e} \quad t_{(i)S} = \max\{t_{(i)j} : 1 \leq j \leq k\};$$

vi. Incluir as médias, mínimos e máximos conjuntamente com os valores t_i contra os percentis esperados da distribuição normal padrão.

O mínimo e o máximo dos t_i formam o envelope. Se o modelo estimado é correto, espera-se que os valores observados estejam dentro das bandas e de forma aleatória. Usando 19 simulações, quantidade sugerida por Atkinson (1985), existe 1 chance em 20 ou 5% de que o maior valor absoluto dos resíduos do conjunto de dados originais esteja fora do envelope simulado, quando o modelo ajustado for o correto. Esse resultado pode ser usado também para detectar a presença de pontos aberrantes.

3.2.5 Aplicação

Para ilustrar a aplicação do MLBWG foi utilizado um conjunto real de dados na área de engenharia elétrica sobre o tempo de falha de filmes de tereftalato de polietileno (PET) em transformadores isolados com gás SF₆. Os dados foram publicados por Hirose (1993) o qual utilizou dados de teste de vida acelerado para estimar o tempo de vida médio de filmes de PET submetidos a stress e estimar o stress limiar abaixo do qual uma falha é improvável de ocorrer. O objetivo do estudo é relacionar o tempo até o filme PET (usado como isolamento elétrico) falhar com o nível de voltagem submetido.

Os tempos de falhas de filmes PET tendem a ser maiores quando submetidos a níveis de voltagens menores e diminuem conforme a voltagem aumenta caracterizando uma distribuição com assimetria à direita dos dados. Várias são as distribuições com assimetria à direita e suporte positivo que podem ser usadas na modelagem desses dados. Desta forma, a aplicação da ftf é uma importante ferramenta na seleção de modelos, pois muitas vezes distribuições com comportamento da ftf totalmente diferentes possuem densidades semelhantes.

É possível extrair informações qualitativas sobre a ftf a partir das observações disponíveis por meio de análise gráfica. Aarset (1987) propôs um gráfico denominado

tempo total em teste (gráfico TTT), obtido plotando os pontos $(r/n, G(r/n))$ obtidos a partir da expressão

$$G(r/n) = \frac{(n-r)T_{r:n} + \sum_{i=1}^r T_{i:n}}{\sum_{i=1}^n T_{i:n}}$$

em que $r = 1, \dots, n$ e $T_{i:n}$, $i = 1, \dots, n$ são as estatísticas de ordem das observações.

As diferentes formas que o gráfico TTT pode assumir são apresentadas na Figura 27. Quando a curva, obtida do gráfico TTT, é uma reta diagonal (curva A), a ftf é constante; quando a curva é convexa (curva B) ou côncava (curva C), a ftf é monotonicamente decrescente ou crescente, respectivamente; quando a curva é convexa e então côncava (curva D), a ftf têm forma de U, e no caso reverso (curva E) é unimodal.

Figura 27 – Gráfico TTT

Os modelos BW, WE e Weibull são modelos encaixados, e portanto, o teste da razão de verossimilhança é útil para realizar comparação da adequabilidade entre esses modelos. A estatística do teste é obtida resolvendo a expressão $w = 2[\ell(\hat{\theta}) - \ell(\tilde{\theta})]$ sob a hipótese nula H_0 contra H_1 , em que $\tilde{\theta}$ e $\hat{\theta}$ são as estimativas de máxima verossimilhança sob H_0 e H_1 , respectivamente e $\ell(\hat{\theta})$ e $\ell(\tilde{\theta})$ são os valores do logaritmo da função de verossimilhança maximizado. As hipóteses a serem testadas são $H_0 : b = 1$ contra $H_1 : b \neq 1$ (BW contra WE) e $H_0 : (\delta, b)^T = (1, 1)^T$ contra $H_1 : (\delta, b)^T \neq (1, 1)^T$ (BW contra Weibull). A estatística w têm distribuição aproximada Qui-quadrado com grau de liberdade correspondente a diferença entre os parâmetros dos modelos testados.

Para o cálculo das medidas de influência local, utilizadas para detectar possíveis pontos influentes, foi utilizado a linguagem de programação Ox versão 6.20 (Doornik, 2007) e as estimativas dos parâmetros dos modelos foram obtidas com o auxílio do *software* estatístico SAS - Statistical Analysis System (Institute, 2004).

3.3 Tempo de falha de filmes PET

Para essa análise considerou-se os tempos em horas até os filmes de PET falharem e quatro níveis de voltagem (5, 7, 10 e 15) medidos em quilovolt (kV), totalizando 41 observações e a covariável nível de voltagem (três observações do conjunto de dados original foram consideradas censuras e portanto excluídas do conjunto). Inicialmente realizou-se a análise exploratória dos dados, cujos resultados são apresentados na Tabela 15 e o histograma na Figura 28a. Dos resultados apresentados na Tabela 15 e da Figura 28 observa-se assimetria à direita dos dados indicando que o ajuste por uma distribuição assimétrica é o mais adequado. O gráfico TTT é apresentado na Figura 28b, no qual observa-se uma curva inicialmente convexa e em seguida côncava, entorno da linha diagonal de referência, indicando que a ftf apresenta forma de U e portanto, o modelo linear BW generalizado é indicado para ajustar-se aos dados assim como seu submodelo WE. Para esse conjunto de dados o modelo MLWG não é indicado pois a ftf da Weibull não modela forma de U.

Tabela 15 – Estatísticas descritivas do conjunto de dados tempo de falha de filmes de PET

Mediana	Média	Desvio padrão	Coefficiente de variação	Assimetria	Curtose	Min	Máx	n
87,75	1510,83	3261,06	215,85 %	1,77	4,18	2,4	9104	41

Figura 28 – Histograma (a) e gráfico TTT (b) para os dados de tempo de falha de filmes PET

O modelo de regressão, considerando a distribuição BW, é dado por

$$y_i = g^{-1}(\beta_0 + \beta_1 x_i) \epsilon_i, \quad i = 1, \dots, 41$$

em que y_i é o tempo até a falha da i -ésima observação, x_i é o nível de voltagem associado a i -ésima observação, ϵ_i é o erro aleatório seguindo distribuição BW($1, \phi, \delta, b$) e $g^{-1}(\cdot)$ é a função de ligação inversa.

Em análise de tempo de falha é comum a transformação dos dados para a escala logarítmica tal que $h_i = \log(y_i)$ e o modelo de regressão é construído com base nos dados transformados, conhecido como modelo de locação escala. No MLBWG adotando-se a função de ligação logarítmica ($\log(\lambda_i) = \eta_i$) o mesmo modelo regressão pode ser obtido sem a necessidade de transformação nos dados. Logo, o modelo de regressão adotado é dado por

$$y_i = e^{\beta_0 + \beta_1 x_i} \epsilon_i.$$

As estatísticas AIC, CAIC e BIC, as estimativas dos parâmetros, o erro padrão (EP) e o nível descritivo (p-valor), para os modelos de regressão BW, WE e Weibull são apresentadas na Tabela 16, na qual observa-se que o modelo BW obteve os menores valores para as estatísticas AIC, CAIC e BIC em comparação aos modelos WE e Weibull. Os parâmetros da regressão foram significativos para os modelos BW, WE e Weibull.

O teste da razão de verossimilhança, cujos resultados são apresentados na Tabela 17, indicou o modelo BW como o mais adequado para ajustar-se aos dados em comparação aos modelos WE ($\omega = 5,4$ e p-valor = 0,02013) e Weibull ($\omega = 19$ e p-valor < 0,0001) ao nível de 5% de significância.

Os gráficos da Figura 29 apresentam as possíveis observações influentes no modelo de regressão BW. Nos gráficos são apresentados os esquemas de influência local (M(0) contra índices das observações), considerando os esquemas de perturbação da verossimilhança (Figura 29a), da variável resposta (Figura 29b) e da covariável nível de voltagem (Figura 29c). Visualmente não observa-se pontos influentes nos três esquemas de perturbação.

A Figura 30b apresenta o gráfico normal de probabilidade para os resíduos quantil t_i com envelope simulado, o qual foi usado para verificar a qualidade do ajuste

Tabela 16 – EMV dos parâmetros para os modelos de regressão linear generalizado Weibull e beta Weibull

Modelo	b	δ	ϕ	β_0	β_1
BW	0,2007	100,3800	0,4115	5,3860	-0,5743
EP	0,1216	25,1758	0,1082	0,9793	0,06369
p-valor	-	-	-	< 0,0001	< 0,0001
AIC = 486,4	CAIC = 488,2	BIC = 495,0			
Modelo	-	δ	ϕ	β_0	β_1
WE	-	100,04	0,2062	3,6874	-0,5671
EP	-	28,6288	0,02575	0,9948	0,05503
p-valor	-	-	-	0,0006	< 0,0001
AIC = 489,8	CAIC = 490,9	BIC = 496,7			
Modelo	-	-	ϕ	β_0	β_1
Weibull	-	-	0,8252	11,7694	-0,6881
EP	-	-	0,0998	0,4276	0,04184
p-valor	-	-	-	< 0,0001	< 0,0001
AIC = 501,4	CAIC = 502,1	BIC = 506,6			

Tabela 17 – Teste da razão de verossimilhança

Modelo	Hipóteses			w	p-valor
BW contra WE	$H_0 : b = 1$	contra	$H_0 : b \neq 1$	5,4	0,02013
BW contra Weibull	$H_0 : (\delta, b)^T = (1, 1)^T$	contra	$H_1 : (\delta, b)^T \neq (1, 1)^T$	19	< 0,0001

Figura 29 – Gráfico de $M(0)$ contra índices para os parâmetros do MLBWG sobre os esquemas de perturbações: (a) da verossimilhança, (b) da resposta e (c) da covariável

do modelo de regressão. O gráfico normal de probabilidade foi construído utilizando o conjunto completo de dados. Observa-se que todos os pontos estão dentro do envelope, o que indica o bom ajuste do modelo BW.

Figura 30 – Resíduo quantil (t_i) contra o índice das observações (a), gráfico normal de probabilidade para o resíduo quantil com envelope simulado (b)

O modelo final, portanto, é dado por

$$\hat{y}_i = e^{5,3860 - 0,5743x_i} = 211,8757 \times 0,5631^{x_i}$$

e dessa forma têm-se que aumentando o nível de voltagem em 1 kV o tempo de falha dos filmes de PET diminuem em 43,7%.

3.4 Conclusões

O MLBWG traz uma nova abordagem da distribuição BW em modelos de regressão. Ele é eficaz para modelar uma parte da resposta média em dados positivos assimétricos incluindo as variáveis explanatórias por meio de uma função linear nos parâmetros denominada preditor linear ($\boldsymbol{\eta} = \mathbf{X}^T \boldsymbol{\beta}$) em um modelo não linear utilizando uma função de ligação g , tal que $g(\lambda) = \boldsymbol{\eta}$, que liga a parte sistemática do modelo a parte aleatória (Da mesma forma que os modelos lineares generalizados). Esse modelo é mais flexível que o MLWG proposto por Prudente e Cordeiro (2010), modela formas monótonas e não monótonas da função taxa de falha e tem o MLWG como submodelo.

Na aplicação apresentadas nesse trabalho o MLBWG se mostrou mais adequado ao ajuste que o MLWG, obtendo menores valores de AIC, BIC e CAIC. O modelo é robusto a pequenas perturbações nos dados e no modelo e todos os valores para o resíduo proposto permaneceram dentro das bandas de confiança do envelope simulado construído, indicando a qualidade do ajuste.

Referências

- AARSET, M. V. How to identify bathtub rate. **Transactions on Realibility**, New York, v. 36, p. 106-108, 1987.
- ATKINSON, A. C. **Plots, transformations and regressions**. Oxford: Statistical Sciences Series Oxford, 1985. 296 p.
- CHATTERJEE, S.; HADI, A. S., **Sensitiv Analysis in Linear Regression**. New York: John Wiley, New York, 1988. 315 p.
- CHOUDHURY, A. A. Simple Derivarion of Moments of the Exponentiated Weibull Distribution. **Metrika**, Heidelberg, v. 62, p. 17-22, 2005.
- CHRISTENSEN, R.; PEARSON, L. M.; JOHNSON, W. Case-deletion diagnostics for mixed models. **Technometrics**, Gainesville, v. 34, p. 38-45, 1992.
- COOK, R. D. Detection of influential observations in linear regression. **Technometrics**, Gainesville, v. 19, p. 15-18, 1977.
- COOK, R. D. Assessment of local influence (with discussion). **Journal of the Royal Statistical Society, Series B**, Malden, v. 48, p. 133-169, 1986.
- COOK, R. D.; PEÑA, D.; WEISBERG, S. The likelihood displacement: A unifying principle for influence, **Communications in Statistics - Theory and Methods**, Philadelphia, v. 17, p. 623-640, 1988.
- CORDEIRO, G. M.; NADARAJAH, S.; ORTEGA, E. M. M. General results for the beta Weibull distribution. **Journal of Statistical Computation and Simulation**, New York, v. 83, n. 6, p. 1082-1114, 2013.
- COX, D. R.; REID, N. Parameter orthogonality and aproximate conditional inference. **Journal of the Royal Statistical Society, Series B**, Malden, v. 49, p. 1-39, 1987.
- DAVIS, D. J. An Analysis of Some Failure Data. **Journal of the American Statistical Association**, New York, v. 47, n. 258, p. 113-150, 1952.

DAVISON, A. C.; TSAI, C. L. Regression model diagnostics. **International Statistical Review**, Malden, v. 60, p. 337-353, 1992.

DOORNIK, J. **Ox**: Object oriented matrix programming using Ox. London: Timberlake Consultants Ltd, 2007. 461 p. DUNN, K. P.; SMYTH, G. K. Randomized quantile residuals. **Journal of Computational and Graphical Statistics**. Melbourne, v. 5, p. 236-244, 1996.

EFRON'S, B. Logistic regression, survival analysis, and the Kaplan-Meier curve. **Journal of the American Statistical Association**, New York, v. 83, n. 402, p. 414-425, 1988.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-Normal distribution and its applications. **Communications in Statistics - Theory and Methods**, Hamilton, v. 31, n. 4, p. 497-512, 2002.

FAMOYE, F.; LEE, C.; OLUMOLADE, O. The Beta-Weibull Distribution. **Journal of Statistical Theory and Applications**, Lawrenceville, v. 4, p. 121-136, 2005.

GLASER, R. E. Bathtub and related failure rate characterizations. **Journal of the American Statistical Association**, New York, v. 75, p. 667-672, 1980.

HASHIMOTO, E. M.; ORTEGA, E. M. M.; CANCHO, V. G.; CORDEIRO, G. M. On estimation and diagnostics analysis in log-generalized gamma regression model for interval-censored data. **Statistics: A Journal of Theoretical and Applied Statistics**, Dortmund, v. 47, p. 379-398, 2013.

HIROSE, H. Estimation of threshold stress in accelerated life-testing. **IEEE Transactions on Reliability** Dallas, v. 42, p. 650-657, 1993.

INSTITUTE, S. **SAS/STAT User's Guide Version 9**. Cary: SAS Institute, 2004. 4995 p.

LAPLANTE-ALBERT, K. A. **Habitat-dependent mortality risk in lacustrine fish**. Québec: Science Université du Québec à Trois-Rivières, 2008. p. 123.

- LEE, C.; FAMOYE, F.; OLUMOLADE, O. Beta-Weibull Distribution: Some Properties and Applications to Censored Data. **Journal of Modern Applied Statistical Methods**, Detroit, v. 6, n. 1, p. 173-186, 2007. ORTEGA, E. M. M.; CANCHO, V. G.; LACHOS, V. H. A Generalized Log-gamma Mixture Model for Cure Rate: Estimation and sensitivity analysis. **Sankhy: The Indian Journal of Statistics, Series B**, New York, v. 71, p. 1-29, 2009.
- ORTEGA, E. M. M.; CORDEIRO, G. M.; KATTAN, M. W. The log-beta Weibull regression model with application to predict recurrence of prostate cancer, **Statistical Papers (1988)**, Heidelberg, v. 54, p. 113-132, 2013.
- ORTEGA, E. M. M.; CORDEIRO, G. M.; CAMPELO, A. K.; KATTAN, M. W.; GANCHO, V. G. A power series beta Weibull regression model for predicting breast carcinoma, **Statistics in Medicine**, Malden, v. 34, p. 1366-1388, 2015.
- PAULA, G. A. On diagnostics in double generalized linear models. **Computational Statistics and Data Analysis**, Oviedo, v. 68, p. 44-51, 2013.
- POON, W.; POON, Y. S. Conformal normal curvature and assessment of local influence. **Journal of the Royal Statistical Society, Series B**, Malden, v. 61, p. 51-61, 1999.
- PRUDENTE, A. A.; CORDEIRO, G. M. Generalized weibull linear models. **Communications in Statistics Theory and Methods**, Philadelphia, v. 39, p. 3739-3755, 2010.
- R Development Core Team. **R: A Language and Environment for Statistical Computing**. Vienna: R Foundation for Statistical Computing, 2011. 3452 p.
- RUSSO, C. M.; CYSNEIROS, F. J. A.; AOKI, R.; PAULA, G. A. Influence diagnostics in heteroscedastic and/or autoregressive nonlinear elliptical models for correlated data. **Journal of Applied Statistics**, Leeds, v. 39, p. 1049-1067, 2012.
- SILVA, G. O.; ORTEGA, E. M. M.; CANCHO, V. G.; BARRETO, M. L. Log-Burr XII regression models with censored data. **Computational Statistics and Data Analysis**, Amsterdam, v. 52, p. 3820-3842, 2008.

XIE, F. C.; WEI, B. C. Diagnostics analysis for log-Birnbaum-Saunders regression models. **Computational Statistics & Data Analysis**, Oviedo, v. 51, p. 4692-4706, 2007.

4 PESQUISAS FUTURAS

Dando continuidade a esta pesquisa, pretende-se desenvolver a parte Bayesiana dos MLBWG.

Também pode ocorrer que em algumas aplicações a suposição de independência dos tempos de sobrevivência não é válida. Assim, uma pesquisa futura é introduzir o MLBWG com efeito aleatório para o qual pode-se desenvolver dois métodos de estimação: frequentista e bayesiano.

No MLBWG com efeito aleatório, pode-se estudar a parte de diagnóstico baseadas na função Q-afastamento para estudar medidas de influência local, conforme apresentado no artigo ZHU, H. T., LEE, S. Y., Local influence for incomplete data models. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Malden, v. 63, p. 111 - 126, 2001.

E por fim, desenvolver uma biblioteca com os modelos propostos.

5 APÊNDICE

A.1 - Linhas de comando do software R para estudo de simulação do MLWEG

```
rm(list=ls(all=TRUE))
log.vero<-function(par,bd){
  b0    <-par[1]
  b1    <-par[2]
  phi   <-par[3]
  delta <-par[4]
  y     <-bd[,1]
  x     <-bd[,2]
  #lambda<-exp(b0+b1*x)
  #lambda<-b0+b1*x
  lambda<- 1/(b0+b1*x)

  -sum( log(delta*phi*(1/y)*((y/lambda)^phi)*exp(digamma(2))*
  exp(-((y/lambda)^phi)*exp(digamma(2))))*((1-exp(-((y/lambda)^phi)
  *exp(digamma(2))))^(delta-1))) )
}

# Quantile distribution
qua<-function(u,x,b0,b1,phi,delta){
  #lambda<-exp(b0+b1*x)
  #lambda<-b0+b1*x
  lambda<- 1/(b0+b1*x)
  lambda*((1/exp(digamma(2)))*(-1)*log(1-u^(1/delta)))^(1/phi)
}

eqm<-function(n,b01,b11,phi1,delta1,casa.dec){
  b0<-c(); b1<-c(); phi<-c(); delta<-c();

  for (i in 1:1000){
    u    <-runif(n = n,min = 0,max = 1)
    N    <-rnorm(n = n,mean=0,sd=1)
    x    <-exp(N)
    y    <-qua(u,x,b01,b11,phi1,delta1)
    data<-cbind(y,x)

    hat<-optim(c(b01,b11,phi1,delta1), log.vero, bd=data,
    control=list(maxit = 300))
    b0    <-c(b0,hat$par[1])
    b1    <-c(b1,hat$par[2])
    phi   <-c(phi,hat$par[3])
    delta<-c(delta,hat$par[4])
  }
  means      <-c(mean(b0), mean(b1), mean(phi), mean(delta))
}
```

```

110
vars      <-c(var(b0), var(b1), var(phi), var(delta))
vies      <-means-c(b01, b11, phi1, delta1)
eqms      <-vars+vies^2
result    <-new.env()
result$means<-round(means,casa.dec)
#result$vars<-round(vars,casa.dec)
result$vies <-round(vies,casa.dec)
result$eqms <-round(eqms,casa.dec)
return(as.list(result))
}

```

A.2 - Linhas de comando do software SAS para estimação dos parâmetros do MLWEG

```

proc nlmixed corr data=lista1;

parms phi=0.6859 theta=1 b0=1 b1=0.5;
bounds phi>0;
bounds theta>0;

lambda = exp(b0 + b1*x1);

e=exp(1 - CONSTANT("EULER"));
a = (t/lambda)**(phi);
l = log(theta) + log(phi) - log(t) + log(a*e) - e*a + (theta-1)
*log(1-exp(-a*e));

model t ~ general(l);
run;

```

A.3 - Linhas de comando do software Ox para análise de influência do MLWEG

```

#include<oxstd.h>
#include<oxprob.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,xx1;

```

```

decl uns=ones(n,1);
decl vero=zeros(1,n);
decl phi=fabs(vP[0][0]); //phi
decl delta=fabs(vP[1][0]); //delta
decl beta0=vP[2][0]; //beta 0
decl beta1=vP[3][0]; //beta 1
decl w11=vP[4:][0]; /*peso*/

for(cont=0;cont<n;++cont)
{
y=g_mY[cont][0];
xx1=g_mY[cont][1];

decl xbeta=beta0+beta1*xx1;
decl lambda = exp(xbeta);
decl z= ((y/lambda)^phi)*(exp(polygamma(2,0)));

vero[0][cont]=w11[cont]*( log(delta) - log(y/phi) + log(z) - z
+ (delta - 1)*log( 1-exp(-z)) );
}
adFunc[0]=double(vero*uns);
return 1;
}

main()
{
g_mY = loadmat("seguro-ox.txt");
print(" Dados ",g_mY);
decl nc=rows(g_mY);
print("Imprimir numero dwe linhas \n", nc);
decl dfunc;
decl vP=0.6859|6.6226|8.0552|0.01769|ones(542,1);

log_vero(vP,&dfunc,0,1);
println("vero=" ,dfunc);

MaxControl(500,20);
decl mhess=0.00001*unit(544);
Num1Derivative(log_vero,vP,&mhess);
//print("gradiente=",mhess');
Num2Derivative(log_vero,vP,&mhess);
print ("segunda derivada=",mhess);
decl QQ, delta,FF, HHH, HHH1;
QQ=mhess[0:3][0:3];
delta=mhess[0:3][4:];
//print(mhess[1:4][1:4]);
print("matriz de informacao observada numerica",QQ);
print("matriz das derivadas cruzadas numerica",delta);

```



```
FF=delta'*(((1)*QQ)^(-1))*delta;
print("matriz FF",FF);
HHH=diagonal(FF);
HHH1=HHH';
print ("VECTOR DIAGONAL DE FF  TOTAL",HHH1);

decl aFFval, aFFvec;
eigensym(FF, &aFFval, &aFFvec);

print ("autovalores da matriz FF numerico", aFFval);
print ("autovetores da matriz FF numerico", aFFvec);
print("LMAX para TEMPO ", aFFvec[0:][:0]);
}
```