KLEBER NAPOLEÃO NUNES DE OLIVEIRA BARROS

CLASSE DE DISTRIBUIÇÕES DE MARSHALL-OLKIN GENERALIZADA EXPONENCIADA

RECIFE - PE
DEZEMBRO - 2014

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

CLASSE DE DISTRIBUIÇÕES DE MARSHALL-OLKIN GENERALIZADA EXPONENCIADA

Tese apresentada à Universidade Federal Rural de Pernambuco, para obtenção do título de Doutor em Biometria e Estatística Aplicada.

Área de Concentração: Modelagem e estatística aplicada

Estudante: Kleber Napoleão Nunes de Oliveira Barros
Orientador: Prof. Dr. Kleber Régis Santoro

Recife, dezembro de 2014.

Ficha catalográfica

B277c Barros, Kleber Napoleão Nunes de Oliveira
Classe de distribuições de Marshall-Olkin generalizada exponenciada / Kleber Napoleão Nunes de Oliveira Barros. – Recife, 2014.
165 f. : il.

Orientador: Kleber Régis Santoro.
Tese (Doutorado em Biometria e Estatística Aplicada) – Universidade Federal Rural de Pernambuco, Departamento de Biometria e Estatística Aplicada, Recife, 2014.
Inclui referências e apêndice(s).

1. Classe de distribuição Marshall-Olkin 2. Novas distribuições 3. Análise de sobrevivência I. Santoro, Kleber Régis, orientador II. Título

CDD 574.018

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

CLASSE DE DISTRIBUIÇÃO DE MARSHALL-OLKIN GENERALIZADA EXPONENCIADA

Kleber Napoleão Nunes de Oliveira Barros

Tese julgada adequada para a obtenção do título de Doutor em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 19/12/2014 pela banca examinadora.

Orientador:

Prof. Dr. Kleber Régis Santoro
Universidade Federal Rural de Pernambuco

Banca Examinadora:

Prof. Dr. Tiago Alessandro E. Ferreira
Universidade Federal Rural de Pernambuco

Prof. Dr. Wilson Rosa de Oliveira
Universidade Federal Rural de Pernambuco

Prof. Dr. Francisco Cribari Neto
Universidade Federal de Pernambuco

Prof. Dr. Ricardo Alves de Olinda
Universidade Estadual da Paraíba

Recife, 19/12/2014.

Dedico a todos que de alguma forma tornaram a realização deste trabalho possível. Em especial aos meus pais Djalma e Célia, e avós Severino e Josefa.

Agradecimentos

Primeiramente, gostaria de agradecer ao meu orientador professor Doutor Kleber Régis Santoro por sua competência, conhecimento transmitido e atenção dedicada no desenvolvimento dessa tese.

Aos professores e funcionários do Programa de Pós-Graduação em Biometria e Estatística Aplicada e do Departamento de Estatística e Informática pela convivência agradável durante esse período. Muitas foram às pessoas que de alguma maneira contribuíram para a realização deste trabalho. A todas elas, meu sincero agradecimento.

A minha esposa Patrícia e ao meu filho Khalel pelo suporte familiar durante a realização desta pesquisa.

Aos colegas de curso, em especial a Cícero Brito, Gabriel R. de Melo por suas contribuições durante o curso e realização deste trabalho. Outros que não poderia deixar de citar são Luciano de Souza, Joseilme Gouveia, Paulo Duarte, Sílvio Fernando, Macio Albuquerque, Fábio Jaques, Erinaldo Leite, entre outros, que foram amigos e companheiros durante o curso.

Resumo

O presente trabalho generaliza a família de distribuições Marshall-Olkin pela adição de parâmetros, tornando-a uma nova classe mais flexível, criando-se a nova distribuição Marshall-Olkin Generalizada Exponenciada Weibull (MOGEW). Foi estudado o comportamento da função densidade de probabilidade MOGEW e sua respectiva função de risco com resultados promissores. Encontrou-se algumas quantidades tais como função geradora de momentos, função quantílica e mediana, além das curvas de Bonferroni e Lorenz, para a distribuição proposta. Obteve-se uma simulação e utilizou-se o método de reamostragem bootstrap para obter os erros padrão dos estimadores dos parâmetros do modelo. Para aplicação foram utilizados dados de magnitudes de abalos sísmicos próximos ao arquipélago de Fiji, dados de resistência de fibras de vidro ajustando o modelo proposto, submodelos e distribuições concorrentes. Também se obteve um modelo de regressão para dados censurados que foi aplicado a dados de um estudo sobre AIDS e um modelo Bayesiano para dados de quebra de fibras de carbono. Os resultados mostraram que a distribuição apresenta ajuste superior, em comparação às distribuições concorrentes, para os conjuntos de dados aplicados.

Palavras-chave: Classe de Distribuição Marshall-Olkin, Novas Distribuições, Análise de Sobrevivência.

Abstract

This work generalizes the family of Marshall-Olkin distributions by adding parameters, making it a new more flexible class, creating the new Generalized Exponentialized Marshall-Olkin Weibull distribution (GEMOW). Its probability density function and the associated risk function were studied with promising results. We found some quantities such as moments, moment generating function, quantile function and median, as well Bonferroni and Lorenz curves, for the proposed distribution. We drawed a simulation and we employed the bootstrap resampling procedure for the standard errors of the estimators of the model parameters. We applied the new distribution to magnitudes earthquakes dataset from Fiji archipelago, glass fiber resistance dataset to the proposed model, sub-models and competitors distributions. Also it was obtained a regression model for censored data that was applied to data from a study of AIDS, and a Bayesian model implemented for carbon fibre data. Comparing with the others distributions, the results demonstrate that GEMOW has superior fit to the applied dataset.

Palavras-chave: Generalized Exponentialized Marshall-Olkin Class, New Distributions, Survival Analysis.

Lista de Figuras

Lista de Tabelas

Sumário

CAPÍTULO 1

Introdução

A estatística paramétrica é um dos mais conhecidos e mais promissores ramos da inferência estatística (COX, 2006). Modelos probabilísticos e métodos de estimação têm sido produzidos de forma cada vez mais intensa. Nos últimos quinze anos a quantidade de novos modelos paramétricos vem aumentando com o maior acesso a ferramentas computacionais. Distribuições como normal, exponencial, gama, beta, de Laplace, normal-inversa, de Rayleigh, de Weibull, Gumbel e Fréchet entre outras, vêm sendo generalizadas pela adição de parâmetros que possibilitam uma maior flexibilidade. Em particular, para dados de sobrevivência (ou taxa de falha) a intensidade de criação de novos modelos tem sido grande, haja vista a grande aplicabilidade de distribuições que atendam tais características.

A distribuição de Weibull foi nomeada em homenagem a Waloddi Weibull (WEIBULL, 1951), que a descreveu em detalhes em 1951. No entanto, Fréchet (1927) e Rosin e Rammler (1933) foram os primeiros a identificá-la e aplicá-la, respectivamente. Notadamente, a distribuição de Weibull tem grande importância entre as distribuições contínuas positivas, devido a suas aplicações nas mais diversas áreas tais como análise de sobrevivência, confiabilidade, engenharia industrial, hidrologia, etc. Outra distribuição bastante conhecida, que também é utilizada em análise de sobrevivência, é a distribuição gama.

O trabalho está organizado da seguinte forma. No Capítulo 2 é relatada de forma breve a teoria da Análise de Sobrevivência, algumas quantidades e modelos probabilísticos positivos, faz-se um breve histórico das generalizações de distribuições que serviram de guias para este trabalho. No Capítulo 3 são descritas algumas quantidades do modelo proposto. Também serão estudadas as formas da função densidade de probabilidade e função de risco para esta generalização. No Capítulo 5 se obtém diversas quantidades para o modelo proposto compondo-o com a distribuição de Weibull, obtendo-se assim, a distribuição Marshall-Olkin Generalizada Exponenciada de Weibull (MOGEW). São obtidas diversas quantidades como a função geradora de momentos, função quantílica, mediana, curvas de Bonferroni, de Lorenz, log-verossimilhança. Por fim, no Capítulo 6, se aplica a distribuição generalizada proposta e seus submodelos para uma simulação, para dados de magnitudes de terremotos no arquipélago de Fiji, dados de resistência de fibras de vidro, dados de Análise de Sobrevivência de um ensaio clínico sobre HIV e dados de quebras de fibras de carbono com um modelo bayesiano, acompanhados de uma breve discussão dos resultados obtidos.

O objetivo geral deste trabalho é obter uma generalização da família de distribuições de Marshall-Olkin (MARSHALL; OLKIN, 1997) pela introdução de novos parâmetros, obtendo uma maior flexibilidade (i.e., uma maior quantidade de formatos possíveis aos quais uma curva possa se ajustar), sendo portanto capaz de modelar comportamentos diversos de dados positivos. Tem-se o objetivo específico de se obter diversas quantidades para esta nova generalização utilizando a distribuição de Weibull que, por fim, será aplicada a alguns conjuntos de dados simulados e reais.

CAPÍTULO 2

Revisão de Literatura

Em Análise de Sobrevivência (AS) a variável de interesse, T, o **tempo** até o acontecimento do evento, denominado tempo de falha, que está relacionada à variável auxiliar indicadora de **censura** (KALBFLEISCH; PRENTICE, 2011; LAWLESS, 2011). A censura acontece pela incorporação de informações parciais devido à perda ou retirada de um elemento do estudo (COLOSIMO; GIOLO, 2006). Entre as diversas áreas de interesse da AS, estão Medicina, Engenharia, Biologia e Economia (RODRIGUES *et al.*, 2008).

No que diz respeito as diversas abordagens em AS, Rodrigues *et al.* (2008) afirma que:

> *"Em geral, os modelos paramétricos não são triviais de serem utilizados em situações práticas devido às suposições exigidas na sua formulação. Por outro lado, eles são mais informativos do que os modelos não-paramétricos e permitem interpretar de forma objetiva o mecanismo biológico de interesse do pesquisador."*

Os dados utilizados em AS, na presença de covariáveis $\boldsymbol{x}$ (por exemplo: idade, sexo, tipo de tratamento, peso, etc...), para o i-ésimo paciente, $i = 1, 2, \ldots, n$, são coletados e representados pelo vetor $(t_i, \delta_i, \boldsymbol{x}_i)$, em que t_i é uma observação do tempo de falha (morte) ou censura (T_i) e δ_i indica se houve censura, isto é,

$$\delta_i = \begin{cases} 1, & \text{se } t_i \text{ é um tempo de falha,} \\ 0, & \text{se } t_i \text{ é uma censura.} \end{cases}$$

2.1 Função de Sobrevivência e de Risco

Seja $f(t)$ uma função densidade de probabilidade (fpd) de T, uma variável aleatória (v.a.) positiva, que representa o tempo de vida. A fdp $f(t)$ deve atender as seguintes condições

(i) $f(t) \geq 0, \ \forall t$,

(ii) $\int_0^\infty f(t)dt = 1$.

Define-se a função de distribuição acumulada (fda) da variável T por

$$F(t) = P(T \leq t) = \int_0^t f(t)dt.$$

Uma função de interesse em AS é a denominada função de Sobrevivência. Ela é definida como a probabilidade de uma observação (peça mecânica, dispositivo, pessoa, etc.) vir a falhar num tempo superior a t, sendo dada por

$$S(t) = P(T > t) = 1 - P(T \leq t). \tag{2.1}$$

A função de sobrevivência é o complemento da função de distribuição, isto é, $S(t) = 1 - F(t)$, e pode-se interpretá-la como a probabilidade de uma observação não vir a falhar num tempo inferior a t.

Outra função muito importante neste campo é a função de risco ou função taxa de falha, definida por

$$h(t) = \lim_{h \to 0} \frac{P(t \leq T \leq t + h | T \geq t)}{h} = \frac{f(t)}{S(t)}, \tag{2.2}$$

em que $f(t) = F'(t)$ é a função de densidade de probabilidade. Note também que a função de risco pode ser determinada por $h(t) = -d \log S(t)/dt$.

A função de risco é interpretada como a probabilidade de um indivíduo falhar no instante t, desde que a falha não tenha ocorrido antes de t. Representa, pois, o risco eminente do indivíduo falhar em t.

Na Figura 2.1 são apresentadas as formas básicas da função de risco. Para o gráfico $h_1(t)$ o risco é constante, isto é, com o passar do tempo a probabilidade de um indivíduo qualquer do estudo falhar não se altera, dado que não falhou até o tempo t. Em $h_2(t)$ temos riscos crescentes, também conhecido na literatura como forma de J, assim os indivíduos têm baixa probabilidade de falhar quando entram no estudo e esta probabilidade aumenta com o passar do tempo. A função $h_3(t)$ tem comportamento decrescente, também conhecido como forma de J invertido, típico de equipamentos eletrônicos, cuja probabilidade de falha diminui com o passar do tempo. Por sua vez, a função $h_4(t)$ tem concavidade positiva e é conhecida na literatura como forma de banheira ou de U (COLOSIMO; GIOLO, 2006). Essa característica é expressa por seres humanos e animais para muitas doenças. Já $h_5(t)$ é convexa e é referida como unimodal ou U invertido. Os três primeiros gráficos podem ser obtidos, por exemplo, com a distribuição bi-paramétrica de Weibull, enquanto os gráficos $h_4(t)$ e $h_5(t)$ requerem distribuições mais flexíveis, que possuem mais parâmetros.

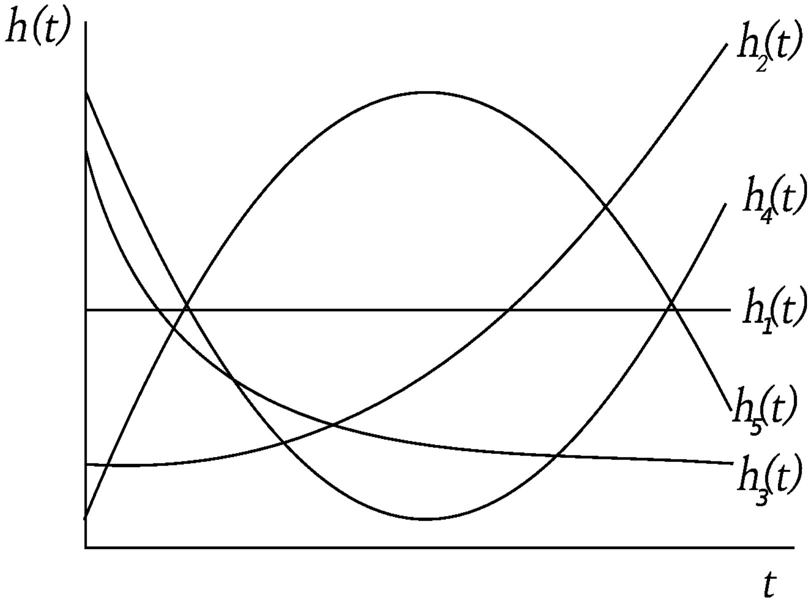


Figura 2.1: Comportamentos típicos de funções de risco.

A função de risco é mais informativa do que a função de sobrevivência para a mesma densidade, já que funções de sobrevivência semelhantes podem gerar funções de taxa de falha significativamente diferentes. Desta forma, a função de risco é utilizada como instrumento natural de estimação de modelos (COLOSIMO; GIOLO, 2006). De fato, Cox e Oakes (1984) listam algumas razões para se preferir a função de risco à outras medidas de tempo de vida:

(i) pode ser fisicamente esclarecedor considerar o 'risco' imediato associado a um indivíduo sabidamente estar vivo na idade t,

(ii) comparação de grupos de indivíduos são as vezes intensivamente feitas via risco,

(iii) modelos baseados em risco são frequentemente convenientes quando há censura ou há diversos tipos de falhas,

(iv) a comparação com uma distribuição exponencial é particularmente simples em termos do risco,

(v) o risco é a forma especial para o sistema de 'falha simples' da função de intensidade completa para processos pontuais mais elaborados, i.e., sistemas nos quais diversos eventos pontuais podem ocorrer para cada indivíduo. [1]

2.2 Gráfico TTT

O gráfico do tempo total de teste (gráfico TTT) é uma metodologia gráfica muito utilizada para seleção de modelos em AS (AARSET, 1987). O gráfico TTT é obtido 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 $r = 1, \ldots, n$ e $T_{i:n}$ é i-ésima a estatística de ordem da amostra.

O gráfico TTT, apresentado na Figura 2.2, tem alguns formatos típicos que indicam funções de risco diferentes. Quando o gráfico é uma reta diagonal (curva A), a função de risco é constante, como em $h_1(t)$; quando a curva é côncava (curva B) convexa (curva C), tem-se um comportamento monotônico crescente ($h_2(t)$) ou decrescente ($h_3(t)$), respectivamente; quando a concavidade muda de negativa para positiva (curva D), tem-se uma taxa de falha em forma de U, como em ($h_4(t)$); e finalmente quando a curva é côncava e depois convexa (curva E), tem-se um comportamento unimodal, como em $h_5(t)$.

2.3 Função de Verossimilhança

O método de Máxima Verossimilhança é usado para estimar os parâmetros que melhor expliquem a amostra observada. Para um modelo usual de inferência, deve-se maximizar a função de verossimilhança:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(t_i; \boldsymbol{\theta})$$

[1] "(i) it may be physically enlightening to consider the immediate 'risk' attaching to an individual known to be alive at age t, (ii) comparison of groups of individuals are sometimes intensively made via the hazard, (iii) hazard–based models are often convenient when there is censoring or there are several types of failure, (iv) comparison with an exponential distribution is particular simple in terms of the hazard, (v) the hazard is the special form for the 'single failure' system of the complete intensity function for more elaborate point processes, i.e., systems in which several point events can occur for each individual."

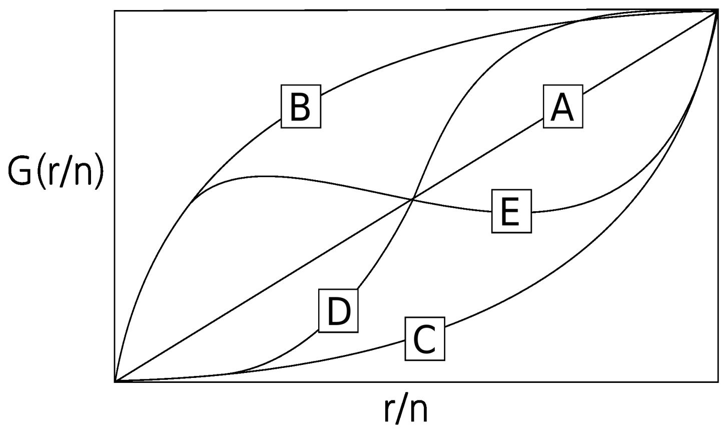


Figura 2.2: Gráfico TTT utilizado na estimação/validação de modelos em Análise de Sobrevivência.

em relação a $\boldsymbol{\theta}$ para se obter tal estimativa.

A logaritmo da função de verossimilhança (log-verossimilhança) é definida por

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{r} \log f(t_i; \boldsymbol{\theta}),$$

que é mais simples do ponto de vista computacional, quando comparada com a função de verossimilhança $L(\theta)$, uma vez que o logaritmo é uma função capaz de diminuir a complexidade de operações aritméticas.

2.4 Tipos de Censura

Há três tipos básicos de censura. Uma observação censurada é considerada do tipo I quando ocorre devido ao término de um período pré-estabelecido do estudo. Neste caso, cada indivíduo que não falhou ao fim do estudo será censurado. Seja n o número de indivíduos submetidos a um tratamento. Suponha que $d = \sum_{i=1}^{n} \delta_i < n,\ i = 1, 2, \ldots, n$ indivíduos não sobreviveram até o término do experimento e que $n - d$ indivíduos foram censurados, pois estavam vivos até o tempo final C do estudo. Assim, cada observação é representada por (t_i, δ_i) com

$$\delta_i = \begin{cases} 1, & \text{para } t_i \leq C, \\ 0, & \text{para } t_i > C. \end{cases}$$

Neste caso, ao finalizar o experimento se observa d falhas, cada uma com informação completa especificada pela fdp $f(t_i)$ e $n-d$ censuras, cuja informação parcial é obtida pela função de sobrevivência $S(t_i)$. Na censura do tipo II o estudo termina após um número d fixado previamente de falhas e uma quantidade não determinada de censuras. A censura aleatória ou não-informativa acontece quando cada indivíduo tem tempos de censura C_i e de falha T_i estatisticamente independentes. Para $i = 1, 2, \ldots, n$, o tempo observado será $t_i = \min\{T_i, C_i\}$. Denota-se $f_T(t)$ e $f_C(c)$, como sendo as densidades do tempo de falha e de censura, respectivamente, e $S_T(t)$ e $S_C(c)$, as funções de sobrevivência do tempo de falha e de censura, respectivamente. Pode-se mostrar-se que, de um modo geral, a função de verossimilhança para os três esquemas de censura é

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{r} f(t_i; \boldsymbol{\theta}) \prod_{i=r+1}^{n} S(t_i; \boldsymbol{\theta}). \tag{2.3}$$

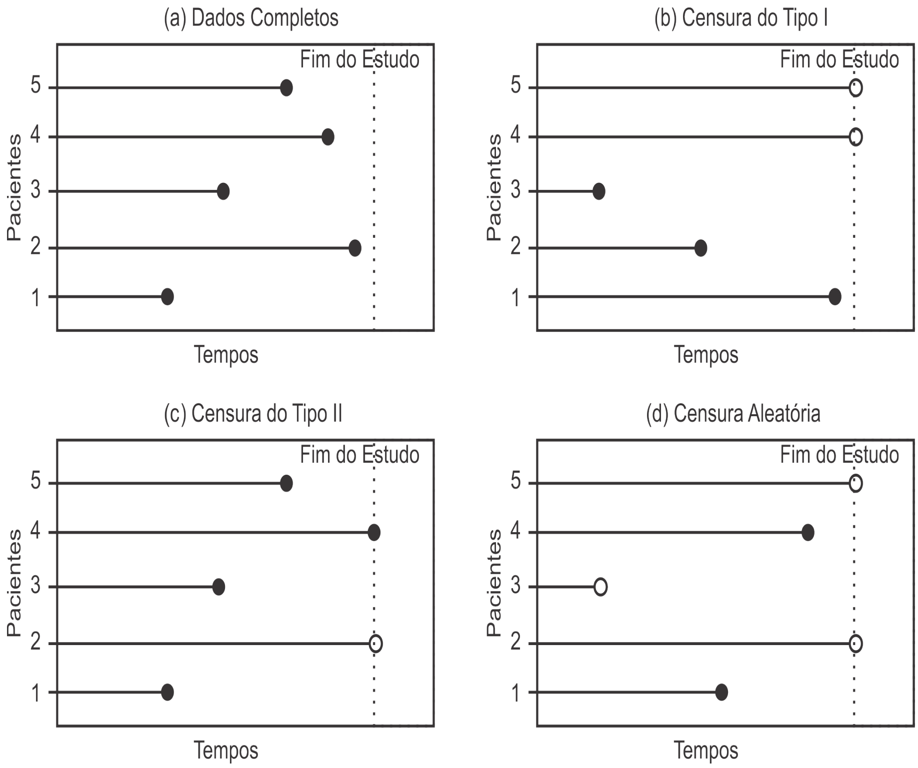
Na Figura 2.3 é possível se observar uma ilustração dos tipos de censura estudados aqui. Na Figura 2.3.a são apresentados dados em que todos os pacientes falharam e se diz que este é um estudo de dados completos. Em (b) observa-se a censura do tipo I, isto é, após um tempo pré-estabelecido, o estudo termina e se contabiliza os tempos de falhas e censuras existentes. Em (c) os dados se têm o esquema de censura do tipo II, em que o estudo termina após um número pré-estabelecido de falhas (quatro neste caso). Na Figura 2.3.d alguns pacientes são censurados antes do fim do estudo e outros ao fim do estudo.

2.5 Estimador de Kaplan-Meier

O estimador de Kaplan-Meier, proposto por Kaplan e Meier (1958), é uma adaptação da função de sobrevivência empírica para dados com censura. Considere $t_1 < t_2 < \cdots < t_k$, os k tempos distintos e ordenados de falha, d_j o número de falhas no instante t_j, $j = 1, \ldots, k$, e n_j o número de indivíduos em risco (indivíduos que não falharam e não foram censurados) antes do instante t_j. O estimador de Kaplan-Meier é definido por

$$\hat{S}(t) = \prod_{j:t_j<t} \left(\frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j<t} \left(1 - \frac{d_j}{n_j} \right). \tag{2.4}$$

O estimador de Kaplan-Meier faz uma estimativa da probabilidade condicional de sobreviver no instante t_j dado que sobreviveu até antes de t_j. Algumas Propriedades do estimador de Kaplan-Meier são expostas por (BRESLOW; CROWLEY, 1974) e (KAPLAN; MEIER, 1958). O estimador não-paramétrico de Kaplan-Meier é, amplamente utilizado no ajuste de funções de sobrevivência paramétricas, por conta das suas propriedades. Visualmente, quanto mais próximo uma curva paramétrica estiver do estimador



Fonte: Adaptado de Colosimo e Giolo (2006).

Figura 2.3: Esquema representativo dos tipos de censura.

de Kaplan-Meier, melhor será o ajuste.

2.6 Algumas Medidas Importantes

A seguir, se descreve de forma sucinta algumas medidas que serão utilizadas ao longo do texto.

2.6.1 Expansão Binomial

Muitas vezes é conveniente representar uma função em série. A expansão binomial será uma fórmula usada recorrentemente, quando se estiver tratando de expansões da função de distribuição, função de densidade de probabilidade, além dos momentos de uma distribuição. É definida por

$$(1-x)^q = \sum_{j=0}^{\infty} \binom{q}{j} (-1)^j x^j \qquad \text{para} \qquad -1 < x < 1 \quad \text{e} \quad q \in \mathbb{R}. \tag{2.5}$$

Particularmente, para q negativo se pode escrever

$$(1-x)^q = \sum_{j=0}^{\infty} \frac{\Gamma(q+j)}{\Gamma(q)\, j!} x^j \qquad \text{para} \quad -1 < x < 1 \quad \text{e} \quad q \in (-\infty,\, 0), \tag{2.6}$$

em que $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx$ é a função gama.

Nos casos especiais, quando $q = -1$ ou $q = -2$ se tem, respectivamente,

$$(1-x)^{-1} = \sum_{j=0}^{\infty} x^j \tag{2.7}$$

e

$$(1-x)^{-2} = \sum_{j=0}^{\infty} (j+1) x^j. \tag{2.8}$$

2.6.2 Momentos e Mediana

Os momentos são medidas importantes das distribuições. Diversas características de uma distribuição são determinadas pelos seus momentos.

Defininição 1. *O r-ésimo momento $r = 1, 2, \ldots,$ de uma v.a. X é definido por $E(X^r) = \mu_r' = \int x dF$, desde que $E(|X^r|) < \infty$, i.e., o momento absoluto de ordem r seja finito. O r-ésimo momento central é $\mu_r = E[(X - E(X))^k]$. Se a variável aleatória X é discreta $E(X^r) = \sum_i x_i^r\, p(x_i)$, se é contínua $E(X^r) = \int x^r\, f(x)\, dx$.*

O r-ésimo momento central μ_r pode ser expresso em termos do r-ésimo momento μ_r' pela seguinte fórmula recursiva

$$\begin{aligned}
\mu_r = E[(X-\mu)^r] &= \int_{-\infty}^{\infty} (x-\mu)^r\, dF(x) \\
&= \int_{-\infty}^{\infty} \sum_{j=0}^{\infty} \binom{r}{j} x^{r-j} (-\mu)^r\, dF(x) \\
&= \sum_{j=0}^{\infty} \binom{r}{j} (-1)^r \mu^r \int_{-\infty}^{\infty} x^{r-j}\, dF(x) \\
&= \sum_{j=0}^{r} \binom{r}{j} (-1)^j \mu^j \mu_{r-j}'
\end{aligned}$$

(PAPOULIS; PILLAI, 2002; BRITO, 2014).

O momento de 1ª ordem $\mu = E(X)$ recebe algumas denominações: média, esperança e valor esperado são as mais conhecidas. A média é uma medida de tendência central da distribuição. Uma interpretação física da esperança matemática é centro de gravidade de uma massa. Assim, a esperança desempenha um papel central também em Física e nas Engenharias. Essa definição coincide com o baricentro do Cálculo e Geometria.

Outra medida de tendência central é a mediana m, definida por

$$\int_{-\infty}^{m} f(x)\,dx = \int_{m}^{\infty} f(x)\,dx,$$

para variáveis contínuas. A mesma tem utilidade em Robustez Estatística, pois é menos sensível a extremos (isto é, à presença de observações extremas não modifica de modo significativo a curva da mediana) que a média (JURECKOVÁ; SEN, 1996).

A variância $Var(x) = \sigma^2 = E[(X - E(X))^2] = E(X^2) - E(X)^2$ é outra medida importante obtida a partir dos momentos. Sua raiz quadrada σ, o desvio-padrão, é a medida de dispersão (ou variabilidade) mais importante da variável X. A vantagem de σ é que tem a mesma unidade de medida de X e portanto a comparação é direta.

O coeficiente de assimetria é o terceiro momento central padronizado, definido por $\gamma_1 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$ e que, para distribuições unimodais, mede o quanto uma cauda (extremidade) da densidade da distribuição difere da outra. Para distribuições simétricas $\gamma_1 = 0$. Distribuições com $\gamma_1 > 0$ são conhecidas como distribuições assimétricas positivas ou à direita; distribuições com $\gamma_1 < 0$ são assimétricas negativas ou à esquerda. O coeficiente γ_1 define a relação entre μ e m. Se $\gamma_1 < 0$, então $\mu \leq m$. Entretanto, se $\gamma_1 > 0$, então $\mu \geq m$. De fato, Pearson definiu o coeficiente de assimetria não-paramétrico (KENDALL *et al.*, 1946) como $\frac{3(\bar{x}-m_x)}{\sigma}$ de onde se verifica a relação anterior, em que $\bar{x}$ e m_x são a média e a mediana amostral.

O coeficiente de curtose é o quarto momento central padronizado $\gamma_2 = E\left[\left(\frac{X-\mu}{s}\right)^4\right]$, sendo uma medida de achatamento da distribuição. A curtose da distribuição normal é $\gamma_2 = 3$, assim uma definição moderna do coeficiente de curtose é $\gamma_2 = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] - 3$, assim a curtose da distribuição fica sendo zero e a comparação é facilitada. Distribuições com $\gamma_2 = 0$ são ditas ser mesocúrticas. Quando $\gamma_2 < 0$ a distribuição é platicúrtica, o que significa que ela é mais achatada que a normal e tem caudas mais pesadas que a normal. Se por outro lado, $\gamma_2 > 0$, a distribuição é dita leptocúrtica, o que indica que ela é menos achatada que a distribuição normal e tem caudas mais leves que a normal (BALANDA; MACGILLIVRAY, 1988; CYSNEIROS *et al.*, 2005).

2.6.3 Função Geradora de Momentos e Função Característica

A função geradora de momentos (f.g.m.), como o nome diz, é útil para se encontrar os momentos de uma distribuição. A f.g.m. de uma variável aleatória X é definida por

$$M_X(t) = E(e^{t\,X}), \quad -t_0 < t < t_0.$$

A função geradora de momentos pode ser reescrita como $M_X(t) = \sum_{j=0}^{\infty} \frac{t^j\,E(X^j)}{j!}$ que é uma combinação infinita dos momentos da v.a. X. Qualquer momento r pode ser obtido se derivando $M_X(t)$, r vezes e fazendo $t = 0$, isto é,

$$E(X^r) = \frac{d^{(r)}M_X(t)}{d\,t^r}\,|_{t=0}\;.$$

Uma importante propriedade das funções geradoras de momentos conhecida como teorema da unicidade. Para todo valor de t,

$$M_X(t) = M_Y(t) \quad \Rightarrow \quad F_X(x) = F_Y(y).$$

Isto é, se duas distribuições X e Y têm a mesma f.g.m., então elas são idênticas em quase toda parte. A existência de momentos não implica a existência da f.g.m. (lognormal, por exemplo), assim a propriedade acima não é equivalente a dizer que se duas v.a. têm os mesmos momentos, então elas têm a mesma distribuição. Heyde (1963) citado por Durrett (2010) mostra o seguinte exemplo, seja a distribuição log-normal e sua versão perturbada,

$$f_0(x) = \frac{1}{x\sqrt{2\pi}}\exp -(\log(x))^2/2$$

e

$$f(x) = f_0(x)\left\{1 + \alpha\,\text{sen}[2\pi\;\log(x)]\right\}.$$

As funções $f_0(x)$ e $f(x)$ têm os mesmos momentos $E(X^r) = e^{r^2/2}$, porém claramente não são a mesma distribuição. No entanto, se a função geradora de momentos de uma distribuição tem raio de convergência[2] positivo, então a v.a. é determinada pelos seus momentos (BILLINGSLEY, 2008).

A função característica é uma alternativa à função geradora de momentos. Seja X

[2]Seja uma série de potências definida por

$$f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n,$$

em que a é uma constante complexa (o centro do disco de convergência), c_n é um coeficiente complexo dependente de n, e z é uma variável complexa.

O raio de convergência r é um real não negativo tal que a série converge se

$$|z-a| < r.$$

uma variável aleatória. A função característica de X é

$$\phi_X(t) = E(e^{\imath tX}), \quad t \in \mathbb{R},$$

em que $\imath = \sqrt{-1}$ é a unidade imaginária. O principal ganho da função característica em relação à função geradora de momentos é que ela sempre existe para qualquer t real (ROUSSAS, 1997; MAGALHÃES, 2011).

Todas as propriedades da função geradora de momentos, em particular o teorema da unicidade, podem ser estendidas para a função característica. Outra propriedade interessante das funções características, que se relaciona com a família de locação e escala, definida posteriormente, é que se as variáveis aleatórias X e Z se relacionam por $X = \sigma Z + \mu$, em que μ e σ são quantidades desconhecidas, então:

$$\phi_X(t) = e^{\imath t\mu}\phi_Z(\sigma t).$$

Outra propriedade das funções características é a conhecida fórmula da inversão, que descreve, através da função característica, como se determina a função de distribuição de X, pela fórmula

$$\tilde{F}(b) - \tilde{F}(a) = \lim_{c\to\infty} \frac{1}{2\pi} \int_{-c}^{c} \frac{e^{-\imath at} - e^{-\imath bt}}{\imath t} \phi_X(t)\, dt,$$

sendo $\tilde{F}(w) = \frac{F(w)+F(w^-)}{2}$, $\forall w \in \mathbb{R}$. Para a demonstração desse teorema, consulte Roussas (1997) ou Magalhães (2011).

2.6.4 Função Quantílica

A função quantílica Q é a inversa da função de distribuição acumulada F da variável aleatória X. Para uma função de distribuição estritamente monótona, Q retorna o valor x abaixo do qual se encontra $p\%$ da massa X. Sua definição formal é

$$Q(p) \;=\; \inf\{x \in \mathbb{R} : p \leq F(x)\}$$

em que $0 < p < 1$. Esta definição continua válida para distribuições discretas. Outras notações para a função quantílica são $F^{-1}(p)$ e x_p. Casos especiais da função quantílica são o primeiro quartil $Q(0,25)$ e o terceiro quartil $Q(0,75)$. Note que a mediana é o segundo quartil, isto é, $m = Q(0,5)$. O percentil é definido por

$$P(p) = 100\, Q(p)\%.$$

Gilchrist (2002) oferece diversas propriedades e aplicações da função quantílica em alguns contextos, como Hidrologia, Controle Estatístico de Qualidade Análise de Sobrevivência e Confiabilidade.

A esperança e a variância da v.a. X pode ser escrita em termos da função quantílica:

$$E(X) = \int_0^1 Q(u)\,du$$

e

$$Var(X) = \int_0^1 [Q(u) - E(X)]^2\,du.$$

Para outras propriedades, veja Parzen (2004).

2.6.5 Desvio Médio e Mediano

Outra medida importante é a dispersão de uma população, que pode ser mensurada pelo total de desvios da média e mediana. Se T tem distribuição cuja fdp é $f(t)$, então os desvios médios da média $\mu_1 = E(T)$ e da mediana m são dados por

$$\delta_1 = \int_0^\infty |t - \mu_1| f(t) dt \qquad \text{e} \qquad \delta_2 = \int_0^\infty |t - m| f(t) dt,$$

respectivamente.

2.6.6 Curvas de Bonferroni e Lorenz

Duas curvas que aparecem frequentemente em artigos explorando novas distribuições são as curvas de Bonferroni e Lorenz, que têm aplicações na área econômica (GASTWIRTH, 1972; KAKWANI, 1977) em estudos de renda e pobreza. O largamente utilizado índice (ou coeficiente) de Gini (CERIANI; VERME, 2012) é a razão entre a área entre a função identidade e a curva de Lorenz (Área A da Figura 2.4) e o total abaixo da linha de igualdade (soma das áreas A e B). Assim, o índice de Gini é $G = A/(A+B)$. Se uma população tem renda igualmente distribuída, então $G = 1$, de modo que a curva de Lorenz coincide com a função identidade.

As curvas são definidas por

$$B(p) = \frac{1}{p\mu_1} \int_0^q t f(t) dt \qquad \text{e} \qquad L(p) = \frac{1}{\mu_1} \int_0^q t f(t) dt,$$

respectivamente, em que $\mu_1 = E(T)$ e $q = F^{-1}(p)$.

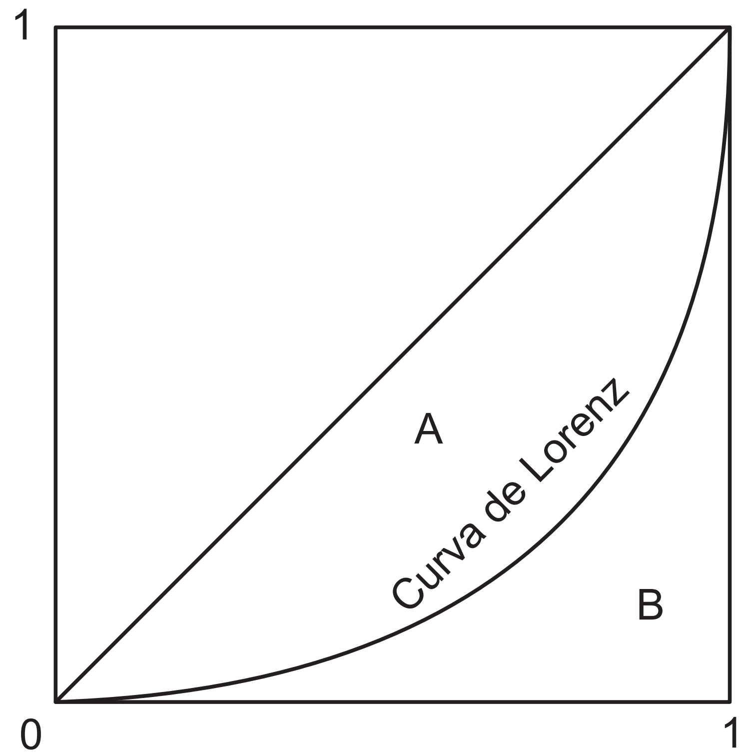


Figura 2.4: Diagrama da curva de Lorenz.

2.6.7 Momentos Probabilisticamente Ponderados

Uma quantidade, definida por Greenwood *et al.* (1979), conhecida como momentos probabilisticamente ponderados, que aparece com frequência, para obtenção de determinadas expressões, quando não há forma analítica fechada, é

$$\tau_{r,s} = E[X^r F^s(x)] = \int_{-\infty}^{\infty} x^r F^s(x) f(x) dx. \tag{2.9}$$

Esta integral, por sua vez, pode ser expressa em termos da função quantílica $Q(x) = F^{-1}(x)$ da distribuição $F(x)$:

$$\tau_{r,s} = \int_0^1 Q(u)^r u^s du.$$

Nadarajah *et al.* (2012) utilizam uma quantidade similar definida por

$$\rho_{r,s} = E[e^{rX} F^s(x)] = \int_{-\infty}^{\infty} \exp\{rx\} F^s(x) f(x) dx \tag{2.10}$$

que também pode ser expressa como função do quantil $Q(x) = F^{-1}(x)$ da distribuição

$F(x)$:

$$\rho_{r,s} = \int_0^1 \exp\{rQ(u)\}u^s du.$$

Série de potência elevada a um inteiro aparecem em diversos contextos. Por Gradshteyn e Ryzhik (2000) (pág. 17), tem-se a seguinte recorrência:

$$\left(\sum_{k=0}^{\infty} a_k x^k\right)^n = \sum_{k=0}^{\infty} c_k x^k, \tag{2.11}$$

com $c_0 = a_0^n, \quad c_m = \frac{1}{ma_0}\sum_{k=1}^{m}(kn - m + k)a_k c_{m-k}, \quad m \geq 1,\, n \in \mathbb{N}.$

2.7 Alguns Modelos Probabilísticos

Antes de se apresentar modelos probabilísticos que são frequentemente utilizados em AS, faz-se necessário definir alguns conceitos ligados aos tipos de parâmetros que um modelo probabilístico pode apresentar.

Defininição 2. *A variável aleatória X tem um modelo de locação se existem uma função f_X e um parâmetro μ tais que $f_X(x;\mu) = f(x-\mu)$. Alternativamente, seja uma variável aleatória Z. Pode-se criar um modelo de locação X se fazendo $X = Z + \mu$.*

O parâmetro μ é dito ser um *parâmetro de locação*. Por exemplo, a variável aleatória $X \sim U_c[\mu-1, \mu+1],\ \mu - 1 < x < \mu + 1$ é um modelo de locação. Os parâmetros de locação estão associados à média da distribuição e não alteram o formato da distribuição (CASELLA; BERGER, 2002).

Defininição 3. *A variável aleatória X tem um modelo de escala se existem uma função f_X e um parâmetro σ tais que $f_X(x;\sigma) = \frac{1}{\sigma}f(\frac{x}{\sigma})$. Alternativamente, para a v.a. Z, um modelo de escala X é obtido se fazendo $X = \sigma Z$.*

Neste caso σ é conhecido como *parâmetro de escala*. Como exemplo, pode-se citar a densidade $f_x(x;\sigma) = \frac{1}{\sigma}\exp\left(\frac{x}{\sigma} - \exp\{\frac{x}{\sigma}\}\right),\ -\infty < x < \infty$ que é uma fdp de um modelo de escala. Os parâmetros de escala estão associados à dispersão da distribuição.

Defininição 4. *A variável aleatória X tem um modelo de locação e escala se existem uma função f_X e os parâmetro μ e σ tais que $f_X(x;\mu,\sigma) = \frac{1}{\sigma}f(\frac{x-\mu}{\sigma})$. Alternativamente, seja a v.a. Z um modelo de locação e escala X é obtido se fazendo $X = \sigma Z + \mu$.*

Como exemplo, no modelo $X \sim N(\mu, \sigma^2)$, o parâmetro μ é de locação e σ é de escala. Um parâmetro que não é nem de locação e nem de escala é chamado de *parâmetro de forma*.

A seguir, na Tabela 2.1 são encontrados alguns dos principais modelos probabilísticos utilizados na literatura de modelos de AS.

Tabela 2.1: Algumas distribuições importantes e funções especiais relacionadas.

Distribuição	$f(x)$	$F(x)$	$h(x)$
Exponencial	$\lambda \exp(-\lambda x)$	$1-\exp(-\lambda x)$	$1/\lambda$
Weibull	$\frac{\kappa}{\lambda}\left(\frac{t}{\lambda}\right)^{\kappa-1}\exp\left\{-\left(\frac{t}{\lambda}\right)^{\kappa}\right\}$	$1-\exp\left\{-\left(\frac{t}{\lambda}\right)^{\kappa}\right\}$	$\frac{\kappa}{\lambda^{\kappa}}t^{\kappa-1}$
Gama	$\frac{t^{\kappa-1}}{\lambda^{\kappa}\Gamma(\kappa)}\exp\left\{-\frac{t}{\lambda}\right\}$	$\int_t^{\infty}\frac{t^{\kappa-1}}{\lambda^{\kappa}\Gamma(\kappa)}\exp\left\{-\frac{t}{\lambda}\right\}dt$	$\frac{t^{\kappa-1}}{\lambda^{\kappa}\Gamma(\kappa)\sum_{x=0}^{\kappa-1}\frac{(t)^x}{\lambda^x x!}}$
Log-normal	$\frac{1}{\sigma t\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{\log(t)-\mu}{\sigma}\right)^2\right\}$	$1-\Phi\left(\frac{-\log(t)+\mu}{\sigma}\right)$	$\frac{f(t)}{1-F(t)}$
Log-logística	$\frac{\kappa}{\lambda^{\kappa}}t^{\kappa-1}\left\{1+\left(\frac{t}{\lambda}\right)^{\kappa}\right\}^{-2}$	$\frac{\left(\frac{t}{\lambda}\right)^{\kappa}}{1+\left(\frac{t}{\lambda}\right)^{\kappa}}$	$\frac{\kappa(t/\lambda)^{\kappa-1}}{\lambda\{1+(t/\lambda)^{\kappa}\}}$
Birnbaum-Saunders	$\phi(a_t)\frac{da_t}{dt}$	$1-\Phi(-a_t)$	$\frac{\phi(a_t)\frac{da_t}{dt}}{\Phi(-a_t)}$
Fréchet	$\frac{\kappa}{\sigma}\left(\frac{x-\mu}{\sigma}\right)^{-1-\kappa}e^{-\left(\frac{x-\mu}{\sigma}\right)^{-\kappa}}$	$e^{-\left(\frac{x-\mu}{\sigma}\right)^{-\kappa}}$	$\frac{\kappa}{\sigma}\left(\frac{x-\mu}{\sigma}\right)^{-1-\kappa}$

em que $a_t = \frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}}-\sqrt{\frac{\beta}{t}}\right)$, $\phi(x)=\Phi'(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ e $\Gamma(a)=\int_0^{\infty}x^{a-1}e^{-x}\,dx$.

Para a distribuição Weibull, utilizando a transformação $Y=\log(T)$ obtém-se que a distribuição condicionada de $Y|X=x$ é

$$f_Y(y|x)=\frac{1}{\sigma}\exp\left\{\frac{y-\mu(x)}{\sigma}-\exp\left\{\frac{y-\mu(x)}{\sigma}\right\}\right\}$$

que é a densidade da distribuição de Gumbel com $\mu(x)=\log\{\lambda\}$ e $\sigma=1/\kappa$, também conhecida como distribuição do valor extremo. Assim se $T\sim Weibull(\lambda,\kappa)$, então $Y\sim Gumbel(\mu,\sigma)$. Aqui, μ age como parâmetro de locação e σ como parâmetro de escala. A função de sobrevivência condicionada em $Y|X=x$ é

$$S_Y(y|x)=\exp\left\{-\exp\left\{\frac{y-\mu(x)}{\sigma}\right\}\right\}.$$

Um modelo de regressão segue, como antes, da função (2.3) com parâmetros β_0, β_1 e σ.

Essa abordagem por transformação é bastante conveniente, pois em AS se lida com tempos de vida, que pode variar de forma não-linear de indivíduo para indivíduo (COLOSIMO; GIOLO, 2006). Explicando em termos populares, uma pessoa pode falecer logo após o início do estudo, enquanto outra pode passar anos e anos viva mesmo após o fim do estudo. Com as duas equações anteriores é possível se obter um modelo de regressão exponencial para dados de Análise de Sobrevivência maximizando a função de verossimilhança (2.3) nos parâmetros β_0 e β_1. A componente μ pode ser generalizada para a adição de diversas covariáveis por $\mu(\mathbf{x})=\beta_0+\beta_1x_1+\cdots+\beta_px_p$, $p<n$.

Rinne (2010) descreve uma interpretação física para a distribuição de Weibull. Se um sistema é composto de n componentes em série com distribuição comum uniforme, então a distribuição do primeiro a falhar (distribuição do mínimo) é exponencial. Se, no entanto, os componentes seguem a distribuição potência, a distribuição do mínimo é Weibull.

A distribuição gama aparece, no contexto das Ciências dos Materiais e Confiabilidade, quando uma ruptura ou falha de um material é causada por um processo repetido de choques, que acontecem de acordo com um processo homogêneo de Poisson (LEIVA *et al.*, 2009). Neste sentido λ pode ser entendido como a taxa de choques e κ como sendo o número de ciclos. Nesse contexto, ela foi utilizada por Brown e Flood (1947) para ajustar o tempo de vida de copos de vidro em uma cafeteira. Também foi utilizada para ajustar o tempo de vida de dispositivos eletrônicos (BIRNBAUM; SAUNDERS, 1958).

A distribuição log-normal foi introduzida por Aitchison e Brown (1957) e mais tarde utilizada como distribuição de tempos de vida por Nelson e Hahn (1972). Mann *et al.* (1974) justificam o uso da distribuição log-normal num processo de tempo de fadiga de materiais. A partir de então, ela têm sido aplicada na literatura para caracterizar fadiga de sistemas (STAHL; GEYER, 1984) e estruturas metálicas (ZHAO *et al.*, 1994), bem como tempos de reparo e manutenção de um sistema (O'CONNOR; KLEYNER, 2011). Também é aplicada em tempos de vida de pacientes com leucemia (ARMENIAN; LILIENFELD, 1974; COLOSIMO; GIOLO, 2006; LIMA *et al.*, 2010).

Bennett (1983) usa o modelo de log-logístico aplicando-os a dados de sobrevivência em pacientes com câncer. Leroy *et al.* (2003) investigam 4468 crianças em Franders, França, quanto ao tempo de aparecimento de dentes permanentes. Utilizando a distribuição log-logística eles foram capazes de discriminar que os dentes permanentes aparecem primeiro em meninas (SANTANA *et al.*, 2012)

A distribuição Birnbaum-Saunders proposta por Birnbaum e Saunders (1969a, 1969b) é baseada no argumento físico que danos sucessivos causados por choques produzem uma fadiga cumulativa nos materiais que acabam por se quebrarem. Esse argumento é conhecido como "Regra de Miner" (MINER, 1945). Eles derivaram o modelo baseando-se na suposição de que o tempo total até a ocorrência de um dano acumulado, produzido pelo crescimento de uma fissura dominante, ultrapasse um certo limiar e cause a falha no material.

A distribuição de Fréchet foi introduzida por Maurice Fréchet, um matemático francês (FRÉCHET, 1927). Utilizada em aplicações de financiamento, a distribuição de Fréchet tem sido de grande utilidade para a modelagem adequada dos retornos do mercado, que têm muitas vezes caudas pesadas (LONGIN, 1996; EMBRECHTS *et al.*, 1997). Em hidrologia, a distribuição de Fréchet tem sido utilizada para modelar a precipitação máximo anual (COLES *et al.*, 2001; CORNELL, 1968).

2.8 Classe de Distribuições Exponenciadas

Mudholkar *et al.* (1995) propõem a distribuição Weibull Exponenciada e Gupta e Kundu (1999) obtêm as relações para um caso particular, denominando-o de distribuição Exponencial Generalizada (EG) como alternativa às distribuições gama e de Weibull. Desde então, diversos trabalhos em AS têm sido produzidos com a EG: (RAQAB, 2002; RAQAB; AHSANULLAH, 2001; ZHENG, 2002; SARHAN, 2007; GUPTA; KUNDU, 2008; ACHCAR; BOLETA, 2009). Nadarajah (2011) revisa a distribuição EG acrescentando-se novos resultados.

O princípio por trás de tais distribuições é simples. Dada uma distribuição com função de distribuição acumulada (fda) contínua $G(t)$ conhecida como função de distribuição base, sua generalização ou exponenciação que pode ser escrita como

$$F(t) = G(t)^{\alpha}, \quad \alpha > 0, \tag{2.12}$$

consequentemente, a fdp é obtida diretamente por $f(t) = \alpha\, g(t)\, G(t)^{\alpha-1}$, em que $g(t) = d\,G(t)/d\,t$ é a função densidade de probabilidade da distribuição base. Quando

$$G(t) = \begin{cases} 1 - e^{-(t/\lambda)^{\gamma}}, & se\ t \geq 0; \\ 0, & se\ t < 0, \end{cases}$$

tem-se a distribuição Weibull exponenciada. Adicionalmente, quando $\gamma = 1$ se tem a distribuição exponencial exponenciada.

2.9 Classe de Distribuições Beta

As distribuições da família beta foram obtidas por Nadarajah e Kotz (2006), com a inserção da função de distribuição acumulada da distribuição exponencial na integral da distribuição Beta, formando-se a distribuição Beta Exponencial (BE). Seja uma distribuição com função de distribuição contínua $G(y)$. A distribuição beta-G será

$$F(t) = \frac{1}{B(\alpha,\beta)} \int_0^{G(t)} w^{\alpha-1}(1-w)^{\beta-1} dw,$$

com $\alpha > 0$ e $\beta > 0$, em que

$$B(\alpha,\beta) = \int_0^1 w^{\alpha-1}(1-w)^{\beta-1} dw.$$

A função densidade da classe de distribuições beta é dada por

$$f(t) = \frac{1}{B(\alpha, \beta)} g(t) G(t)^{\alpha-1} (1 - G(t))^{\beta-1}, \quad t > 0.$$

Esta classe de distribuições tem sido bastante desenvolvida nos últimos anos. Eugene *et al.* (2002) apresenta e discute algumas propriedades da distribuição beta normal, Ojo e Olapade (2003) propõem e demonstram alguns teoremas da distribuição beta logística e definem a distribuição beta log-logística, Cordeiro *et al.* (2011) explicitam diversas quantidades, tais como momentos, expansões e entropia de Rényi, para a distribuição beta Weibull introduzida por Famoye *et al.* (2005).

Recentemente, Pescim *et al.* (2010) apresentaram a beta half-normal generalizada, que tem como casos especiais a half-normal e a half-normal generalizada (COORAY; ANANDA, 2008). Foi mostrada a importância da nova distribuição através de um ajuste a um conjunto de dados, descrito no artigo original, em comparação com a half-normal generalizada, half-normal e normal.

Entretanto, um inconveniente das distribuições beta-G é que elas são funções da distribuição beta incompleta, dada por

$$B_p(\alpha, \beta) = \int_0^p w^{\alpha-1} (1 - w)^{\beta-1} dw,$$

que para $0 < p < 1$ não tem forma analítica fechada, sendo portanto necessário algum método numérico para obtenção dos resultados.

2.10 Classe de Distribuições Kumaraswamy

Para contornar o problema das distribuições beta, Jones (2008), Cordeiro *et al.* (2010), Cordeiro e Castro (2011) propõem a distribuição de Kumaraswamy (Kw), definindo-se as distribuições K-G por

$$F(t) = 1 - (1 - G(t)^\alpha)^\beta, \quad \alpha > 0, \beta > 0,$$

pois esta leva a uma função de distribuição analiticamente solúvel para uma dada distribuição primitiva G. A função densidade geral dessa família é dada por

$$f(t) = \alpha \, \beta \, g(t) \, G(t)^{\alpha-1} \{1 - G(t)^\alpha\}^{\beta-1}.$$

Evidentemente, quando $\alpha = \beta = 1$ se tem a densidade da distribuição de base. Os parâmetros adicionais α e β controlam a assimetria e as caudas da distribuição (NADARAJAH; ELJABRI, 2013). Eles também mostram uma interpretação física para esta classe de distribuições quando α e β são inteiros positivos. Seja um sistema composto por β subsistemas independentes cada um composto por α componentes independentes. Suponha que o sistema falha se qualquer subsistema falhar e que cada subsistema falha se um de seus componentes falhar. Sejam $T_{j1}, T_{j2}, \cdots, T_{j\alpha}$ os tempos de vida dos componentes no subsistema j, $j = 1, 2, \ldots, \beta$ independentes e identicamente distribuídos da distribuição de base G. Se T_j denota o tempo de vida do j-ésimo subsistema, $j = 1, \ldots, \beta$, e seja T o tempo de vida do sistema. Assim, a função de distribuição acumulada de T é

$$\begin{aligned} P(T \leq t) &= 1 - [P(T_j > t)]^{\beta} = 1 - [1 - P(T_j \leq t)]^{\beta} \\ &= 1 - [1 - P(T_{j1} \leq t, T_{j2} \leq t, \ldots, T_{j\alpha} \leq t)]^{\beta} \\ &= 1 - [1 - P(T_{j1} \leq t)^{\alpha}]^{\beta} = 1 - [1 - G(t)^{\alpha}]^{\beta}. \end{aligned}$$

Desta forma, segue que a função de distribuição da classe Kumaraswamy é o tempo de falha do próprio sistema.

Há vários trabalhos recentes com esta classe de distribuições. Por exemplo, Cordeiro *et al.* (2010) propuseram Kw normal, Kw gumbel, Kw gama entre outras. Pascoa *et al.* (2011) apresenta a Kw gama generalizada. Enquanto Saulo *et al.* (2012) estudam a Kw Birnbaum–Saunders capaz de modelar dados com caudas leves/pesadas. Santana *et al.* (2012) introduz um modelo de regressão baseado na distribuição Kw log-logística. Mais recentemente, Bourguignon *et al.* (2013) deriva a distribuição Kw Pareto que tem como sub-casos a Pareto e a Pareto exponenciada.

2.11 Classes de Distribuições Gama Generalizada

Foram desenvolvidas e discutidas várias formas de generalização de distribuições nos últimos anos. Foi introduzida em (ZOGRAFOS; BALAKRISHNAN, 2009) a família gama-generalizada de distribuições em que é definida a função de distribuição acumulada da gama-generalizada (para $x \in R$) por:

$$F(x) = \frac{1}{\Gamma(\delta)} \int_0^{-log[S(x)]} t^{\delta-1} e^{-t} dt, \tag{2.13}$$

em que $S(x) = 1 - F(x)$ é a função de sobrevivência e $\Gamma(\delta) = \int_0^{\infty} x^{\delta-1} e^{-x} dx$ é a função gama.

Ristić e Balakrishnan (2012) propuseram uma alternativa definida pela função de distribuição acumulada dada por:

$$F(x) = \frac{1}{\Gamma(\delta)} \int_0^{-log[F(x)]} t^{\delta-1} e^{-t} dt, \tag{2.14}$$

em que $F(x)$ é a função de distribuição acumulada. Com essa generalização introduziram o modelo Gama Exponencial Exponenciado (GEE).

Com a generalização (2.14) Pinho *et al.* (2012) introduziram o modelo Gamma Weibull exponenciado (GWE)e suas propriedades.

Ramos *et al.* (2013) utilizaram a generalização (2.13) e propuseram a distribuição Log-logística Zografos-Balakrishnan, que contém a distribuição log-logística como um modelo especial.

Nadarajah *et al.* (2015) forneceram um tratamento completo de propriedades matemáticas gerais de distribuições Zografos-Balakrishnan-G. Discutiram também a estimativa dos parâmetros por máxima verossimilhança e forneceram uma aplicação para um conjunto de dados reais.

2.12 Classe de Distribuições de Marshall-Olkin

Por outra linha, Marshall e Olkin (1997) usaram o seguinte método de adição de parâmetros para uma dada distribuição G:

$$F(t) = \frac{G(t)}{G(t) + \beta(1 - G(t))}, \quad \beta > 0. \tag{2.15}$$

Nanda e Das (2012) interpretaram β como sendo um parâmetro de deslocamento da função de risco, pois a função de risco de F é deslocada acima se $\beta > 1$ ou é deslocada abaixo $0 < \beta < 1$ da função de risco de G. Quando $\beta = 1$ a distribuição de Marshall-Olkin se reduz a distribuição de base e dessa forma a função de risco não se modifica.

Thomas e Jose (2004) desenvolveram várias características das distribuições Marshall-Olkin semi-pareto bivariada e Marshall-Olkin pareto bivariada, e introduziram um modelo de séries temporais de k-ésima ordem, sendo a distribuição Marshall-Olkin semi-pareto bivariada a distribuição marginal obtida.

Ghitany *et al.* (2007) introduziram o modelo Marshall-Olkin Lomax, a partir da distribuição Lomax, também conhecida como Pareto do segundo tipo. Os referidos autores mostraram que a mesma pode ser definida como uma mistura de distribuições exponenciais e obtiveram as funções de densidade, taxa de falha e log-verossimilhança.

Zhang e Xie (2007) utilizaram procedimentos de estimação baseados no gráfico de probabilidade Weibull para obter as estimativas dos parâmetros da distribuição Marshall-Olkin Weibull. Preda *et al.* (2011) utilizam a família de Marshall e Olkin para obter, de forma condicional, a distribuição exponencial-Poisson modificada.

Ghitany *et al.* (2005) utilizam de uma modificação da distribuição de Weibull, pela transformação de um parâmetro, para obter a distribuição Marshall-Olkin Weibull Extendida, com

$$G(t) = 1 - e^{\lambda t^{\delta}}, \quad \lambda > 0,\ \delta > 0$$

e fazem uma aplicação à dados censurados. Cordeiro e Lemonte (2013) também estudam esta distribuição, obtendo diversas propriedades.

Santos-Neto *et al.* (2014) juntam a classe de distribuições de Marshall-Olkin e a classe de distribuições estendida de Weibull e propõem a classe de distribuições de Marshall-Olkin estendida de Weibull e mostram que diversas distribuições são sub-casos dessa classe.

2.13 Critério de Informação de Akaike - AIC

Akaike (1974) utilizou a Informação de Kullback-Leibler para analisar se um dado modelo é adequado. Porém seu uso é limitado, pois depende da distribuição g (modelo verdadeiro), que é desconhecida. Demonstrou que o viés é dado assintoticamente por p, em que p é o número de parâmetros a serem estimados no modelo, definindo seu critério de informação como:

$$AIC = -2 \log L(\hat{\theta}) + 2(p),$$

em que $L(\theta)$ é a verossimilhança do modelo avaliada pelo estimador de máxima verossimilhança $\hat{\theta}$. Dentre todos os modelos considerados, deve-se preferir aquele que tem o menor AIC. Burnham e Anderson (2002) apontam uma regra de decisão para a discriminação de modelos. Eles consideram a medida

$$\Delta_{AIC} = \Delta_i = AIC_i - min\{AIC\}$$

em que AIC_i é o AIC do modelo i e min$\{AIC\}$ é o AIC do melhor modelo (o modelo melhor ajustado). Assim, conforme disposto na Tabela 2.2, eles definem uma regra de evidência a favor do modelo i (contra o modelo melhor ajustado):

A partir da tabela acima, diversos artigos (ARMSTRONG; EWEN, 2002; WINNIE-JR *et al.*, 2006; MAZEROLLE, 2006; HAZZAH *et al.*, 2009; MØLLER *et al.*, 2008) têm usado a regra prática $\Delta_i > 2$ para rejeitar o modelo i.

Tabela 2.2: Regras de evidência a favor do modelo i propostas Burnham e Anderson (2002).

Δ_i	Nível de Suporte Empírico do Modelo i
$0-2$	Substancial
$4-7$	Consideravelmente menor
> 10	Essencialmente nenhuma

Fonte: Burnham e Anderson (2002), página 70.

Alternativamente, Burnham e Anderson (2002) definem o AIC ponderado:

$$w_i = \frac{\exp(-\frac{1}{2}\Delta_i)}{\sum_{i=1}^{r}\exp(-\frac{1}{2}\Delta_i)}$$

em que r é o número de r modelos competitivos. A medida w_i pode ser interpretada como a probabilidade relativa de que o i-ésimo modelo minimize a informação de Kullback-Leibler. Os w_i's são equivalantes às probabilidade a posteriori da teoria Bayesiana (BURNHAM; ANDERSON, 2004; MAZEROLLE, 2006).

2.14 Critério de Informação de Akaike Corrigido - AICc

Sugiura (1978) propõe o AICc que uma correção para populações finitas do AIC:

$$AICc = -2\log L(\hat{\theta}) + \frac{2np}{n-p-1}.$$

Burnham e Anderson (2002) recomendam fortemente utilizar está versão corrigida quando n é pequeno. Uma vez que

$$\lim_{n\to\infty}\left[-2\log L(\hat{\theta}) + \frac{2np}{n-p-1}\right] = -2\log L(\hat{\theta}) + 2(p),$$

isto é, o AICc converge para o AIC quando n tende para infinito, não há perdas ao se utilizar AICc em vez do AIC para grandes amostras. Porém, para amostras pequenas, a penalidade do AICc é maior que a penalidade do AIC para modelos com maior quantidade de parâmetros. Assim, utilizando AICc se rejeita modelos com maior quantidade de parâmetros, com maior frequência.

2.15 Critério de Informação Bayesiano - BIC

O Critério de Informação Bayesiano (BIC), proposto por Schwarz (1978), é dado por

$$BIC = -2\log L(\hat{\theta}) + p\log n,$$

em que $L(\theta)$ é a verossimilhança do modelo escolhido, p é o número de parâmetros a serem estimados e n é o número de observações da amostra. Tal como o AIC, este critério é utilizado para discriminar o modelo melhor ajustado entre todos os modelos considerados aquele que tem o menor BIC.

Aitkin (1991) define o fator bayesiano a posteriori do modelo M_1 contra o modelo M_2

$$B_{12} = \frac{\int f(\boldsymbol{y}|\boldsymbol{\theta}_1)\pi(\boldsymbol{\theta}_1|\boldsymbol{y})d\boldsymbol{\theta}_1}{\int f(\boldsymbol{y}|\boldsymbol{\theta}_2)\pi(\boldsymbol{\theta}_2|\boldsymbol{y})d\boldsymbol{\theta}_2}$$

em que $\boldsymbol{\theta}_i$, $i = 1,2$ é o vetor paramétrico do modelo i. Seja BIC_i o BIC do modelo $M_i, i = 1,2$. Schwarz (1978) mostram que, para grandes amostras, $\Delta_{BIC} = BIC_2 - BIC_1$ é uma boa aproximação para $2\log B_{12}$. Kass e Raftery (1995) definem a Tabela 2.3 de evidências do M_1 contra o modelo M_2

Tabela 2.3: Regras de evidências propostas por Kass e Raftery (1995).

B_{12}	Δ_{BIC}	Evidência em favor de M_1
1 - 3	0 - 2	não vale a pena mencionar
3 - 12	2 - 5	substancial
12 - 100	5 - 10	forte
> 100	> 10	decisiva

Fonte: Adaptado de Kass e Raftery (1995).

Assim, pode-se outra vez utilizar uma regra prática: se $\Delta_{BIC} > 2$, tem-se evidências em favor do modelo que têm menor BIC.

2.16 Critério de Informação de Hannan-Quinn - HQIC

O critério de informação de Hannan-Quinn (HQIC) é um critério de classificação de modelos alternativo ao AIC e BIC. O HQIC foi originalmente proposto por Hannan e Quinn (1979) para determinar a ordem de um modelo auto-regressivo (CLAESKENS; HJORT, 2008). É definido por

$$HQIC = -2\log L(\hat{\theta}) + 2p\log\log(n).$$

Burnham e Anderson (2002) se limitam a dizer que este critério tem pouco uso prático. Claeskens e Hjort (2008) observam que HQIC, via lei do logaritmo iterado, apresenta consistência forte para a informação de Kullback–Leibler. Porém, eles mostram que HQIC e BIC não são assintoticamente eficientes, ao contrário do AIC. Eles acrescentam que o termo $\log\log(n)$ é pequeno mesmo para grandes amostras.

Assim como os critérios anteriores, o critério de informação de Hannan-Quinn deve ser mínimo para o melhor modelo.

2.17 Teste de Wald

A estatística do teste de Wald, devido à Abraham Wald (1902-1950) que teve grande contribuição na inferência estatística em meados do seculo XX (WASSERMAN, 2004), é obtida por uma razão, em que o numerador é a diferença da estimativa de máxima verossimilhança do parâmetro $(\widehat{\theta})$ e o próprio parâmetro (aqui $\theta = 0$), e o denominador é a respectiva estimativa do erro-padrão do estimador $\widehat{\theta}$. Sob a hipótese nula $\mathrm{H}_0 : \ \theta = 0$, tal quociente tem distribuição assintótica normal padrão, isto é

$$\frac{\widehat{\theta} - 0}{\widehat{EP}(\widehat{\theta})} \overset{a}{\sim} N(0, 1).$$

A estatística do teste de Wald fica então definida por

$$W = \frac{\widehat{\theta}}{\widehat{EP}(\widehat{\theta})}.$$

em que $\widehat{EP}(\widehat{\theta})$ é o erro-padrão de $\widehat{\theta}$ estimado pelo método de máxima verossimilhança.

Esta quantidade é comparada com o valor crítico da distribuição normal padrão, que ao nível de confiança de 95% é, aproximadamente, igual a 2 (sendo mais rigoroso: $1,96$). Assim, se a estimativa é duas vezes maior que o erro-padrão, em valores absolutos, a estimativa é dita significativa. Hauck e Donner (1977) ao examinar o teste de Wald descobriram anomalias em determinadas situações. O teste, por exemplo, não rejeitava a hipótese nula quando o coeficiente é significativo. Recomendaram, então, a utilização do teste da razão de verossimilhança, como alternativa, quando não se fosse rejeitada a hipótese nula no teste de Wald para uma determinada estimativa.

2.18 Teste da Razão de Verossimilhanças - TRV

O Teste da Razão de Verossimilhanças é um teste bastante geral, podendo ser utilizado para testar um vetor paramétrico (WASSERMAN, 2004). Suponha que se deseje testar:

$$H_0 : \theta \in \Theta_0 \; vs \; H_1 : \theta \notin \Theta_0.$$

A estatística do **Teste da Razão de Verossimilhanças** é dada, sob H_0, por

$$\Lambda = 2 \log \left(\frac{\sup_{\theta \in \Theta} L(\theta)}{\sup_{\theta \in \Theta_0} L(\theta)} \right) = 2\ell(\widehat{\theta}) - 2\ell(\theta_0) \overset{a}{\sim} \chi^2_{p-q}$$

em que $\ell(\widehat{\theta})$ é a função de log-verossimilhança obtida no seu ponto máximo, assim denominado estimador de máxima verossimilhança, sob o modelo proposto com p parâmetros e $\ell(\widehat{\theta_0})$ é a log-verossimilhança obtida no seu ponto máximo (estimador de máxima verossimilhança), sob o modelo reduzido (sub-modelo) com q parâmetros. Como Λ tem distribuição assintótica qui-quadrado, para se proceder o teste deve-se comparar a estatística Λ ao quantil desejado da distribuição χ^2. Suponha que o modelo completo tem três parâmetros, enquanto seu sub-modelo (modelo proposto sob H_0) tem um parâmetro. Se $\Lambda \geq \chi^2_2 = 5,99$ rejeita-se H_0 ao nível de confiança de 95%, isto é, o modelo proposto é melhor.

Engle (1984) mostrou que o teste da razão de verossimilhanças, o teste de Wald e o teste escore são assintoticamente equivalentes.

2.19 Teste de Aderência a uma Distribuição

Os testes de Cramer-Von Mises e Anderson-Darling são baseados na Função de Distribuição Empírica (FDE) dos dados, e apresentam vantagens sobre o teste de aderência qui-quadrado, incluindo maior poder e invariância em relação aos pontos médios dos intervalos escolhidos. Os testes Anderson-Darling e Cramer-von Mises pertencem à classe quadrática de estatísticas baseadas na FDE, pois trabalham com as diferenças quadráticas entre a distribuição empírica e a hipotética. As estatísticas de Cramer-von Mises (W^*) e Anderson-Darling (A^*) são descritas em detalhes em (CHEN; BALAKRISHNAN, 1995). Em geral, quanto menor for o valor das estatísticas W^* e A^*, melhor o ajuste para os dados.

2.19.1 Teste de Anderson-Darling

O teste de Anderson-Darling foi proposto por Anderson e Darling (1952) e é mais utilizado quando o tamanho da amostra não é maior que 25. Este teste baseia-se na função de distribuição empírica.

Seja F a função de distribuição (desconhecida) da população, as hipóteses a serem testadas são:

$H_0 : F(x) = F_0(x), -\infty < x < \infty$

$H_1 : F(x) \neq F_0(x),$

em que F_0 é a função de distribuição proposta, contínua e completamente especificada.

Considere $\delta_i = F(x_{(i)}; \theta)$ uma f.d.a., com $x_{(i)}$ em ordem crescente. Faça $y_{(i)} = \Phi^{-1}(\delta_i)$, em que Φ representa a função de distribuição normal padrão.

Seja $p_{(i)} = \Phi\left(\frac{[y_{(i)} - \bar{y}]}{s_y}\right)$, em que $\bar{y}$ é a média e s_y o desvio padrão dos $y_{(i)}$, respectivamente.

Seja

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}[(2i-1)log(p_{(i)}) + (2n+1-2i)log(1-p_{(i)})], \tag{2.16}$$

em que $p_{(i)} = \Phi(\frac{[y_{(i)} - \bar{y}]}{s_y})$ são percentis ordenados da distribuição normal padrão. A estatística de Anderson-Darling é dada por $A^* = A^2\left(1 + \frac{0{,}75}{n} + \frac{2{,}25}{n^2}\right)$.

2.19.2 Teste de Cramér - Von Mises

Este teste também se baseia na distribuição acumulada e foi proposto por Darling (1957). A expressão da estatística de teste de Cramér - Von Mises é calculada da seguinte forma

$$W^2 = \frac{1}{12n} + \sum_{i=1}^{n}\left(p_{(i)} - \frac{2i-1}{2n}\right)^2, \tag{2.17}$$

em que $p_{(i)}$ é definido como na seção anterior. A estatística de Cramér - Von Misses é dada por $W^* = W^2\left(1 + \frac{0{,}5}{n}\right)$.

As hipóteses a serem testadas são:

$H_0 : F(x) = F_0(x), -\infty < x < \infty$

$H_1 : F(x) \neq F_0(x).$

2.19.3 Teste de Kolmogorov-Smirnov (K-S)

O teste de Kolmogorov-Smirnov foi proposto por Kolmogorov (1933). Posteriormente, Smirnov (1948) forneceu uma tabela para o cálculo da distribuição empírica. Para uma variável aleatória X, o teste K-S baseia-se na análise do ajustamento entre a função de distribuição populacional admitida em H_0, F_0, e a função de distribuição empírica $\hat{F}_n$. Como antes, as hipóteses a serem testadas são:

$H_0 : F(x) = F_0(x), -\infty < x < \infty$

$H_1 : F(x) \neq F_0(x).$

No teste de Kolmogorov-Smirnov considera-se a estatística

$$D_n = \sup_{-\infty < x < \infty} |F_n(x) - F_0(x)|,$$

como uma medida da discrepância entre a função de distribuição da amostra F_n e a função de distribuição proposta F_0. Observe-se que D_n representa a distância vertical máxima entre as imagens da função de distribuição da amostra, $F_n(x)$, e da função de distribuição proposta $F_0(x)$, dando assim uma ideia do ajustamento, como aliás se pretendia. Ou seja, quanto menor o valor de D_n melhor o ajustamento.

2.20 *p*-valor

O p-valor, ou nível descritivo, é definido como a probabilidade de se obter um valor tão ou mais extremo (desfavorável) que o valor observado, se H_0 for verdadeiro. É, usualmente, uma das saídas fornecidas por softwares estatísticos. Na Tabela 2.4, a seguir, são apresentados vários níveis de evidência contra a hipótese nula para várias faixas do nível descritivo.

Tabela 2.4: Níveis de evidências contra H_0 para diversos intervalos do p-valor.

p-valor	evidência contra H_0
$< 0,01$	evidência determinante contra H_0
$0,01 - 0,05$	evidência forte contra H_0
$0,05 - 0,10$	evidência fraca contra H_0
$> 0,10$	pouca ou nenhuma evidência contra H_0

O p-valor é uma medida largamente difundida nas mais diversas áreas científicas que vão desde biologia (REN *et al.*, 2000; DRAGHICI *et al.*, 2007) até ciências sociais (WETZELS *et al.*, 2011; BABBIE, 2012). Tradicionalmente, o p-valor é comparado com nível de significância de 5% ou 1% (NUZZO, 2014).

Recentemente, o p-valor tem recebido diversas críticas por conta da incompatibilidade com o princípio da verossimilhança e da dependência do tamanho da amostra (CASSON, 2011). Pelo princípio da verossimilhança, toda informação da amostra pode ser resumida na função de verossimilhança. Assim, se dois experimentos conduzem a mesma função de verossiminhança, eles são essencialmente o mesmo experimento, isto é, não há dependência do desenho experimental. No entanto, o p-valor, contrariamente a este princípio, fornece conclusões diferentes para resultados essencialmente iguais. A dependência do p-valor em relação ao tamanho da amostra pode ser traduzida pelo clássico problema da moeda, em que para amostras pequenas, o p-valor não rejeita H_0 e para amostras grandes H_0 é rejeitado sempre. De fato, Sir. Ronald A. Fisher (o considerado pai da estatística moderna) afirmava que se deveria combinar o p-valor com outras medidas tais como resultados obtidos em experimentos anteriores (HUBBARD; LINDSAY, 2008). Desta forma, o p-valor têm sido ferrenhamente criticado em ciências médicas (GARDNER; ALTMAN, 1986; GOODMAN, 1999; HUBBARD, 2011).

2.21 Bootstrap

O método de reamostragem bootstrap é uma técnica robusta para se obter estimativas e erros padrão quaisquer (WASSERMAN, 2004). Foi inicialmente proposto por Efron (1979). Alguns livros importantes incluindo este tópico são Efron e Tibshirani (1994), Davison (1997), Godfrey (2009), Shao e Tu (1995).

No mundo real, tomamos uma amostra aleatória $X_1, \ldots, X_n$ de uma distribuição F e construímos uma função $\hat{\theta} = g(X_1, \ldots, X_n)$ chamada estatística da qual se realizam inferências sobre um parâmetro θ da população. O método bootstrap essencialmente repete esse procedimento utilizando a amostra aleatória como base para se construir uma distribuição empírica $\hat{F}_n$ (JAMES *et al.*, 2013). O esquema a seguir ilustra o método:

$$\begin{array}{rcccccc} \text{Mundo Real:} & F & \Longrightarrow & X_1, \ldots, X_n & \Longrightarrow & \hat{\theta} = g(X_1, \ldots, X_n) \\ \text{Mundo Bootstrap:} & \hat{F}_n & \Longrightarrow & X_1^*, \ldots, X_n^* & \Longrightarrow & \hat{\theta}^* = g(X_1^*, \ldots, X_n^*) \end{array}$$

Conforme o esquema acima, são simuladas n realizações $X_1^*, \ldots, X_n^*$ de $\hat{F}_n$ com reposição, assim X_i têm probabilidade $1/n$ de ser escolhido. O processo é repetido B vezes, para se computar o erro-padrão bootstrap (EP_{boot}) de $\hat{\theta}^*$. A seguir, é mostrado o algoritmo que resume todo o processo.

Algoritmo Bootstrap

(1) Simule $X_1^*, \ldots, X_n^* \sim \hat{F}_n$;

(2) Calcule $\hat{\theta}^* = g(X_1^*, \ldots, X_n^*)$;

(3) Repita os passos (1) e (2) B vezes e compute $\hat{\theta}_1^*, \hat{\theta}_2^*, \ldots, \hat{\theta}_B^*$;

(4) Calcule

$$EP_{boot}(\hat{\theta}^*) = \sqrt{\frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\theta}_b^* - \hat{\theta}_{(\cdot)}^*\right)^2}, \tag{2.18}$$

em que $\hat{\theta}_{(\cdot)}^* = \frac{1}{B} \sum_{r=1}^{B} \hat{\theta}_r^*$.

Wasserman (2004) discute três métodos para obtenção de intervalos de confiança:

IC Normal. O intervalo de confiança normal é derivado do intervalo de confiança frequentista

$$\hat{\theta}^* \pm z_{\alpha/2}\, EP_{boot}.$$

Ele alerta que este método não é acurado, a menos que a distribuição de $\hat{\theta}^*$ seja aproximadamente normal.

IC Pivotal. O intervalo de confiança pivotal se baseia na quantidade pivotal $R_n = \hat{\theta}_n - \theta$. Após algum desenvolvimento algébrico, se chega no seguinte intervalo

$$(2\hat{\theta}_n^* - \hat{\theta}_{1-\alpha/2}^*;\, 2\hat{\theta}_n^* - \hat{\theta}_{\alpha/2}^*)$$

IC Percentil. O intervalo de confiança percentil é definido por

$$(\hat{\theta}_{\alpha/2}^*;\, \hat{\theta}_{1-\alpha/2}^*).$$

Justificativas e detalhes para os intervalos pivotal e percentil podem ser encontradas em (WASSERMAN, 2004).

Uma correção para as estimativas bootstrap pode ser obtida definindo

$$\tilde{\theta} = \hat{\theta}^* - \widehat{\text{viés}}$$

e fazendo $\widehat{\text{viés}} = \hat{\theta}_{(\cdot)}^* - \hat{\theta}^*$, de tal forma que

$$\tilde{\theta} = 2\hat{\theta}^* - \hat{\theta}_{(\cdot)}^*. \tag{2.19}$$

Supondo independência entre $\hat{\theta}^*$ e $\hat{\theta}^*_{(\cdot)}$ o erro-padrão de $\tilde{\theta}$ fica

$$EP(\tilde{\theta}) = \sqrt{4Var(\hat{\theta}^*) + Var(\hat{\theta}^*_{(\cdot)})}. \tag{2.20}$$

em que $Var(\hat{\theta}^*)$ é o quadrado de $EP_{boot}(\hat{\theta}^*)$. Para se obter uma estimativa para $Var(\hat{\theta}^*_{(\cdot)})$ se repete o algoritmo de bootstrap uma quantidade arbitrária C, tomando-se ao final do processo a variância amostral de $\hat{\theta}^*_{(\cdot),1}, \ldots, \hat{\theta}^*_{(\cdot),C}$.

Efron (1981) adapta a técnica de reamostragem bootstrap para Análise de Sobrevivência. A única modificação no algoritmo é a substituição da reamostra $X_1^*, \ldots, X_n^*$ por $(T_1^*, D_1^*, X_1^*), \ldots, (T_n^*, D_n^*, X_n^*)$, em que T é o tempo de falha, D é o indicador de censura e X é um vetor de covariáveis. Ele utiliza o método para obter estimativas para a variância do estimador de Kaplan Meyer. Neste trabalho, utilizar-se-á a técnica bootstrap para obter melhores estimativas e respectivos erros-padrão de ajustes de densidades a conjuntos univariados de dados, bem como em Análise de Sobrevivência com o mesmo propósito.

2.22 Inferência Bayesiana

Na inferência bayesiana qualquer abordagem parte de uma simples regra de probabilidade condicional, conhecida como Teorema de Bayes. A fórmula do Teorema de Bayes encontra-se num antigo artigo póstumo (BAYES; PRICE, 1763) do reverendo Thomas Bayes que estudou um problema do que se chamava de "probabilidade inversa": dada uma amostra independente de falhas e sucessos $y_1, y_2, \ldots, y_n$, qual é a probabilidade de sucesso? A teoria da probabilidade resolvia o "problema direto": dado θ, o número de sucessos $s = \sum_{i=1}^n y_i$ tem a distribuição $s|\theta \sim Binomial(n, \theta)$, que é, $p(s|\theta) = \binom{n}{\theta}\theta^s(1-\theta)^{n-s}$. A quantidade desconhecida θ é especificada pela densidade a priori $\pi(\theta)$. A solução de Bayes, redescoberta independentemente por Laplace (LAPLACE, 1995) em 1774, pela seguinte fórmula:

$$\pi(\theta|s) = \frac{\pi(\theta)p(s|\theta)}{\int \pi(\theta)p(s|\theta)d\theta}.$$

Utilizando a posteriori, pode-se responder uma pergunta específica sobre θ, por exemplo

$$P(a < \theta < b|s) = \int_a^b \pi(\theta|s)d\theta.$$

Observe que tal resolução em nenhum momento utiliza artifícios tais como resultados assintóticos ou quantidades pivotais, tais como são necessários na metodologia clássica. Outro ponto contra a inferência convencional advém da necessidade de amostras relativamente grandes para se ter resultados relevantes, o que não é necessário na inferência bayesiana (PAULINO *et al.*, 2003).

No contexto bayesiano, a informação prévia que o pesquisador/especialista tem sobre um determinado fenômeno é incorporada pela especificação de uma distribuição a priori $p(\theta)$ para a quantidade de interesse θ (GAMERMAN; LOPES, 2006). Esta distribuição deve representar (probabilisticamente) o conhecimento que se tem sobre θ antes da realização do experimento. A distribuição dos dados observados $p(y|\theta)$ é conhecida como verossimilhança. Conhecendo-se os dados, a informação prévia é atualizada na densidade a posteriori via o teorema de Bayes. Na Figura 2.5 é apresentada uma situação típica da relação entre priori, verossimilhança e posteriori. Note que a posteriori é uma especie de ponderação entre a priori e a verossimilhança.

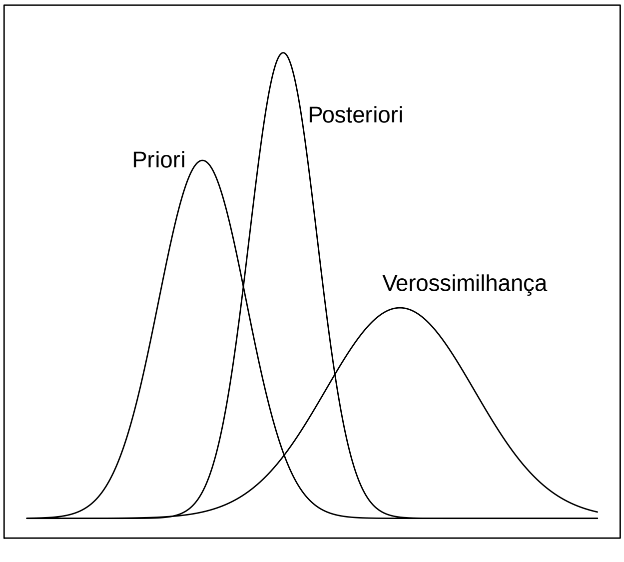


Figura 2.5: Gráfico típico das relações entre as densidades de Bayes.

Na sequência, este trabalho trata de técnicas de integração numéricas, conhecidos como métodos de integração Monte Carlo, que são muito utilizados em integrais que surgem nos cálculos da abordagem Bayesiana.

2.23 Integração de Monte Carlo

A distribuição a posteriori pode ser escrita em termos de esperanças de funções particulares do parâmetro θ, ou seja

$$E[g(\theta)|x] = \int g(\theta)\pi(\theta|x)d\theta. \tag{2.21}$$

Dessa forma, o problema geral da inferência bayesiana consiste em calcular tais valores esperados segundo a distribuição a posteriori de θ.

A inferência exata só é possível se estas integrais puderem ser calculadas analiticamente, caso contrário deve-se usar aproximações, caso em que se recorre a métodos numéricos.

2.23.1 Método de Monte Carlo Simples

A ideia é escrever a integral que se deseja calcular como um valor esperado. Seja θ um parâmetro de interesse e seja $x_1, \ldots, x_n$ uma amostra da distribuição $X_1, \ldots, X_N|\theta$. Uma simulação com S valores independentes θ da posteriori pode ser obtida por

$$\theta^{(1)}, \ldots, \theta^{(S)} \sim \pi(\theta|x_1, \ldots, x_n).$$

A distribuição empírica $\{\theta^{(1)}, \ldots, \theta^{(S)}\}$ é uma aproximação Monte Carlo de $\pi(\theta|x_1, \ldots, x_n)$, que é consistente quando S aumenta. Pela Lei dos Grandes Números (HOFF, 2009):

$$\frac{\sum_{s=1}^{S} g(\theta^{(S)})}{S} \to E[g(\theta)|x_1, \ldots, x_n].$$

Assim, pode-se utilizar um software que simule a distribuição $\pi(\theta|x_1, \ldots, x_n)$ e obter as quantidades

$$\begin{aligned}
\bar{\theta} = \frac{\sum_{s=1}^{S} \theta^{(S)}}{S} &\rightarrow E[\theta|x_1, \ldots, x_n] \quad \text{(esperança a posteriori);}\\
\frac{\sum_{s=1}^{S} (\theta^{(S)} - \bar{\theta})^2}{S} &\rightarrow Var[\theta|x_1, \ldots, x_n] \quad \text{(variância a posteriori);}\\
\{\theta^{(1)}, \ldots, \theta^{(S)}\} &\rightarrow p(\theta|x_1, \ldots, x_n) \quad \text{(distribuição a posteriori);}\\
\frac{\#(\theta^{(S)} \leq c)}{S} &\rightarrow P(\theta \leq c|x_1, \ldots, x_n) \quad \text{(probabilidade a posteriori);}\\
\{\theta^{(1)}, \ldots, \theta^{(S)}\} &\rightarrow \theta_\alpha \quad \text{(percentil a posteriori).}
\end{aligned}$$

2.23.2 Monte Carlo via Função de Importância

Utiliza-se a amostragem por importância quando é impossível ou muito custoso simular valores da distribuição a posteriori. Neste caso, recorre-se a função $q(\theta)$ que é de fácil amostragem, chamada de função de importância.

Seja $q(\theta)$ uma função de densidade definida no mesmo espaço variação de θ então a integral (2.21) pode ser aproximada por

$$\frac{1}{n} \sum_{i=1}^{n} \frac{g(\theta_i)p(\theta_i)}{q(\theta_i)} \rightarrow E\left[\frac{g(\theta)p(\theta)}{q(\theta)}\right]$$

e tem as mesmas propriedades do estimador de Monte Carlo simples.

2.24 Métodos de Reamostragem

Os métodos de reamostragem geram valores em duas etapas. Na primeira etapa gera-se valores de uma distribuição auxiliar conhecida. Na segunda etapa utiliza-se uma correção para que os valores sejam representativos (ao menos aproximadamente) da distribuição a posteriori. Conforme Smith e Gelfand (1992), toma-se a priori como distribuição auxiliar.

2.24.1 Método de Rejeição

Seja a densidade a priori $p(\theta)$ uma densidade auxiliar e uma constante A tal que $p(x|\theta) < A$. Tomando-se A como sendo o valor máximo da função de verossimilhança, ou seja $A = p(x|\hat{\theta})$ a probabilidade de aceitação de uma valor $\theta \sim p(\theta)$ é $p(x|\theta)/p(x|\hat{\theta})$.

O seguinte algoritmo gere valores da densidade a posteriori:

1. gerar um valor θ^* da distribuição a priori;

2. gerar $u \sim U(0,1)$;

3. aceitar θ^* como um valor da posteriori se $u < p(x|\theta^*)/p(x|\hat{\theta})$, caso contrário θ^* e retornar ao item 1.

Um problema desse método é a necessidade de se maximizar a função de verossimilhança o que pode não ser uma tarefa simples em modelos mais complexos.

2.24.2 Reamostragem Ponderada

Utiliza a mesma ideia de gerar valores de uma distribuição auxiliar porém sem a necessidade de maximização da verossimilhança. A desvantagem é que os valores obtidos são apenas aproximadamente distribuídos segundo a posteriori.

Este método toma uma segunda amostra (ou reamostra) de tamanho S da distribuição discreta em $\theta_1, \ldots, \theta_n$ com probabilidades $w_1, \ldots, w_S$. Tomando-se novamente a priori como densidade auxiliar, ou seja $q(\theta) = p(\theta)$ se calcula os pesos

$$w_i = \frac{\frac{p(x|\theta_i)\pi(\theta_i)}{q(\theta_i)}}{\sum_{j=1}^{n} \frac{p(x|\theta_i)\pi(\theta_i)}{q(\theta_i)}} = \frac{p(x|\theta_i)}{\sum_{j=1}^{n} p(x|\theta_j)}, \quad i = 1, \ldots, n$$

dessa forma, o algoritmo para geração de valores (aproximadamente) da posteriori fica

1. gerar valores $\theta_1, \ldots, \theta_n$ da distribuição a priori;

2. calcular os pesos $w_i, i = 1, \ldots, n$;

3. reamostrar valores com probabilidades $w_1, \ldots, w_n$.

Observe que esse método é um boostrap ponderado pelos w_i's.

2.25 Critério de Informação *Deviance* - DIC

O Critério de Informação *Deviance* (DIC) é uma generalização do AIC e do BIC para seleção de modelos Bayesianos (BEST *et al.*, 2003).

Seja a *deviance* $D(\theta) = -\log(p(x|\theta)) + C$. A constante C é anulada ao se comparar modelos e portanto pode ser desconsiderada. Seja a esperança $\bar{D} = E[D(\theta)]$ e o número efetivos de parâmetros $p_D = \bar{D} - D(\bar{\theta})$, em que $\bar{\theta}$ é a média de θ. Gelman *et al.* (2013) introduz a seguinte modificação $p_D = \frac{1}{2}\widehat{Var}(D(\theta))$. O DIC é, então

$$DIC = \bar{D} + p_D.$$

O melhor modelo bayesiano deve ter DIC mínimo dentre todos os modelos testados.

CAPÍTULO 3

Família de Distribuições de Marshall-Olkin Generalizada Exponenciada

Neste capítulo são apresentadas diversas propriedades do modelo proposto, que passará a ser denominado, a partir de então, como família de distribuições de Marshall-Olkin Generalizada Exponenciada, cuja função de distribuição é dada a seguir.

Defininição 5. *Seja $X \sim MOGE(\alpha, \beta, \gamma, \delta, \theta)$ uma variável aleatória da família de distribuições de Marshall-Olkin Generalizada Exponenciada (MOGE). A função de distribuição de X é dada por:*

$$F(x) = \left\{ \frac{G(x)^{\alpha}}{G(x)^{\alpha} + \beta[1 - G(x)^{\gamma}]^{\delta}} \right\}^{\theta}, \quad \alpha > 0,\ \beta > 0,\ \gamma > 0,\ \delta > 0,\ \theta > 0. \tag{3.1}$$

Teorema 1. *Seja $X \sim MOGE(\alpha, \beta, \gamma, \delta, \theta)$ uma variável aleatória da família de distribuições de Marshall-Olkin Generalizada Exponenciada (MOGE). A função de distribuição de X é dada por:*

$$\begin{aligned} f(x) = {} & \beta\, \theta\, g(x)\, G^{-1-\alpha}(x)[1 - G^{\gamma}(x)]^{\delta - 1} \times \\ & \times [\alpha + (\gamma\delta - \alpha) G^{\gamma}(x)] \left\{ \frac{G^{\alpha}(x)}{G^{\alpha}(x) + \beta[1 - G^{\gamma}(x)]^{\delta}} \right\}^{\theta + 1}. \end{aligned} \tag{3.2}$$

Demonstração: Note que a função de distribuição pode ser escrita como

$$F(x) = \left\{1 + \frac{\beta[1 - G(x)^\gamma]^\delta}{G(x)^\alpha}\right\}^{-\theta}.$$

Fazendo

$$u = 1 + \frac{\beta[1 - G(x)^\gamma]^\delta}{G(x)^\alpha},$$

então

$$\frac{du}{dx} = -\beta\, g(x)\, G^{-1-\alpha}(x)[1 - G^\gamma(x)]^{\delta-1}[\alpha + (\gamma\delta - \alpha)G^\gamma(x)].$$

A regra da cadeia completa a demonstração.

(c.q.d)

Utilizando-se a equação (2.1), a função de sobrevivência da classe de distribuições de Marshall-Olkin generalizada exponenciada fica

$$S(x) = 1 - \left\{\frac{G(x)^\alpha}{G(x)^\alpha + \beta\,[1 - G(x)^\gamma]^\delta}\right\}^\theta. \tag{3.3}$$

Substituindo (3.2) e (3.3) em (2.2), obtém-se a seguinte função de risco da família MOGE:

$$h(x) = \frac{\beta\,\theta\, g(x)\, G(x)^{-1-\alpha}\,[1 - G(x)^\gamma]^{\delta-1}\,[\alpha + (-\alpha + \gamma\delta)G(x)^\gamma]\left\{\frac{G(x)^\alpha}{G(x)^\alpha + \beta[1-G(x)^\gamma]^\delta}\right\}^{1+\theta}}{1 - \left\{\frac{G(x)^\alpha}{G(x)^\alpha + \beta[1-G(x)^\gamma]^\delta}\right\}^\theta}. \tag{3.4}$$

Três limites notáveis existem para esta classe são

$$\lim_{\beta\to 0^+} h(x) = \frac{[(\alpha - \gamma\delta)G(x)^\gamma - \alpha]\, g(x)}{G(x)\,(-1 + G(x)^\gamma)},$$

$$\lim_{\delta\to 0^+} h(x) = \frac{\alpha\beta\theta G(x)^{-1-\alpha}\left(\frac{G(x)^\alpha}{\beta + G(x)^\alpha}\right)^{1+\theta} g(x)}{\left(\frac{G(x)^\alpha}{\beta + G(x)^\alpha}\right)^\theta - 1}$$

e

$$\lim_{\theta\to 0^+} h(x) = -\frac{\beta\,(1 - G(x)^\gamma)^{-1+\delta}\,(\alpha + (-\alpha + \gamma\delta)G(x)^\gamma)\, g(x)}{G(x)\left(G(x)^\alpha + \beta\,(1 - G(x)^\gamma)^\delta\right)\log\left(\frac{G(x)^\alpha}{G(x)^\alpha + \beta(1-G(x)^\gamma)^\delta}\right)}.$$

Não há limite quando um dos parâmetros tende a infinito, portanto não há valor assintótico nestes casos.

Embora as equações (3.1) e (3.2) sejam formas analíticas exatas, a fim de obter certas propriedades, é conveniente escrevê-las como séries infinitas.

3.1 Expansões da Função de Distribuição e de Densidade

Teorema 2. *Seja $X \sim MOGE(\alpha, \beta, \gamma, \delta, \theta)$. A função de distribuição (3.1) pode ser escrita como uma combinação linear da distribuição de base $G(x)$.*

Prova

Para $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$, tem-se que

$$\begin{aligned}
F(x) &= \left\{ \frac{G^{\alpha}(x)}{G^{\alpha}(x) + \beta[1 - G^{\gamma}(x)]^{\delta}} \right\}^{\theta} \\
&= \frac{G^{\alpha\theta}(x)}{\beta^{\theta}[1 - G^{\gamma}(x)]^{\delta\theta}} \left\{ 1 + \frac{G^{\alpha}(x)}{\beta[1 - G^{\gamma}(x)]^{\delta}} \right\}^{-\theta} \\
&= \sum_{j=0}^{\infty} \frac{\Gamma(\theta + j)(-1)^{j}}{\Gamma(\theta)\, j!\, \beta^{j+\theta}} G^{\alpha(\theta+j)}(x)[1 - G^{\gamma}(x)]^{-\delta(\theta+j)} \\
&= \sum_{j,k=0}^{\infty} \frac{\Gamma(\theta + j)\Gamma(\delta(\theta + j) + k)(-1)^{j}}{\Gamma(\theta)\Gamma(\delta(\theta + j))\, j!\, k!\, \beta^{\theta+j}} G^{\alpha(\theta+j)+\gamma k}(x) \\
&= \sum_{j,k=0}^{\infty} \nu_{j,k}\, G^{\alpha(\theta+j)+\gamma k}(x), \qquad (3.5)
\end{aligned}$$

em que $\nu_{j,k} = \nu_{j,k}(\beta, \delta, \theta) = \frac{\Gamma(\theta+j)\Gamma(\delta(\theta+j)+k)(-1)^{j}}{\Gamma(\theta)\Gamma(\delta(\theta+j))\, j!\, k!\, \beta^{\theta+j}}$.

Para $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$, tem-se que

$$
\begin{aligned}
F(x) &= \left\{ \frac{G^{\alpha}(x)}{G^{\alpha}(x) + \beta[1 - G^{\gamma}(x)]^{\delta}} \right\}^{\theta} \\
&= \left\{ 1 + \frac{\beta[1 - G^{\gamma}(x)]^{\delta}}{G^{\alpha}(x)} \right\}^{-\theta} \\
&= \sum_{j=0}^{\infty} \frac{\Gamma(\theta + j)(-1)^{j} \beta^{j}}{\Gamma(\theta)\, j!} G^{-\alpha\, j}(x)[1 - G^{\gamma}(x)]^{\delta j} \\
&= \sum_{j,k=0}^{\infty} \binom{\delta j}{k} (-1)^{j+k} \frac{\Gamma(\theta + j)\, \beta^{j}}{\Gamma(\theta)\, j!} G^{\gamma k - \alpha j}(x) \\
&= \sum_{j,k=0}^{\infty} \omega_{j,k}\, G^{\gamma k - \alpha j}(x),
\end{aligned} \tag{3.6}
$$

com $\omega_{j,k} = \omega_{j,k}(\beta, \delta, \theta) = \binom{\delta j}{k} (-1)^{j+k} \frac{\Gamma(\theta+j)\, \beta^{j}}{\Gamma(\theta)\, j!}$.

Assim, tanto para $\beta > \frac{[1-G^{\gamma}(x)]^{\delta}}{G^{\alpha}(x)}$ como para $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$, $F(x)$ pode ser escrita como somas de produtos entre funções dos parâmetros e $G(x)$ elevada a uma certa potência.

(c.q.d)

Corolário 1. *Seja $X \sim MOGE(\alpha, \beta, \gamma, \delta, \theta)$ com função de distribuição (3.1) e forma expandida (3.5) ou (3.6). A respectiva função de densidade $f(x)$ tem forma expandida expressa por uma combinação linear de produtos da densidade $g(x)$ e de potências da função de distribuição $G(x)$.*

Prova

Para $\beta > \frac{[1-G^{\gamma}(x)]^{\delta}}{G^{\alpha}(x)}$, derivando-se a função (3.5) em relação a x, tem-se

$$
f(x) = \frac{dF(x)}{dx} = \sum_{j,k=0}^{\infty} v_{j,k}\, g(x)\, G^{\alpha(\theta+j)+\gamma k - 1}(x), \tag{3.7}
$$

em que $v_{j,k} = v_{j,k}(\alpha, \beta, \gamma, \delta, \theta) = [\alpha(\theta + j) + \gamma k]\, \nu_{j,k} = \frac{\Gamma(\theta+j)\Gamma(\delta(\theta+j)+k)(-1)^{j}[\alpha(\theta+j)+\gamma k]}{\Gamma(\theta)\Gamma(\delta(\theta+j))\, j!\, k!\, \beta^{\theta+j}}$.

Se $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$, se utiliza a função (3.6) e se deriva

$$
f(x) = \frac{dF(x)}{dx} = \sum_{j,k=0}^{\infty} w_{j,k}\, g(x)\, G^{\gamma k - \alpha j - 1}(x), \tag{3.8}
$$

em que $w_{j,k} = w_{j,k}(\alpha, \beta, \gamma, \delta, \theta) = (\gamma k - \alpha j)\, \omega_{j,k} = \binom{\delta j}{k} (-1)^{j+k} \frac{\Gamma(\theta+j)\, \beta^{j} (\gamma k - \alpha j)}{\Gamma(\theta)\, j!}$.

Os resultados dessa seção são importantes porque permitem obter diversas propriedades da variável X expressas a partir de funções da distribuição de base.

3.2 Combinação Linear de Exponencializadas

Considere a família exponenciada da seguinte forma

$$H_r(x) = G^r(x) \qquad \text{com} \qquad h_r(x) = r\, g(x)\, G(^{r-1}(x)), \quad r > 0.$$

Para $\beta > \frac{[1-G^\gamma(x)]^\delta}{G^\alpha(x)}$ e utilizando-se a função de distribuição e a densidade do modelo proposto na forma expandida (3.5) e (3.7), obtém-se, respectivamente,

$$\begin{aligned} F(x) &= \sum_{j,k=0}^{\infty} \nu_{j,k} G^{r_{j,k}}(x) \\ &= \sum_{j,k}^{\infty} \nu_{j,k} H_{r_{j,k}}(x) \end{aligned}$$

e

$$\begin{aligned} f(x) &= \sum_{j,k=0}^{\infty} \nu_{j,k} r_{j,k} g(x) G^{r_{j,k}-1}(x) \\ &= \sum_{j,k}^{\infty} \nu_{j,k} h_{r_{j,k}}(x). \end{aligned}$$

Como antes, $\nu_{j,k} = \nu_{j,k}(\beta, \delta, \theta) = \frac{\Gamma(\theta+j)\Gamma(\delta(\theta+j)+k)(-1)^j}{\Gamma(\theta)\Gamma(\delta(\theta+j))\, j!\, k!\, \beta^{\theta+j}}$ e $r_{j,k} = r_{j,k}(\alpha, \gamma, \theta) = \alpha(\theta + j) + \gamma k$.

Se $\beta < \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$ se utiliza as expansões (3.6) e (3.8) para obter

$$\begin{aligned} F(x) &= \sum_{j,k=0}^{\infty} \omega_{j,k} G^{r_{j,k}}(x) \\ &= \sum_{j,k}^{\infty} \nu_{j,k} H_{r_{j,k}}(x) \end{aligned}$$

e

$$\begin{aligned} f(x) &= \sum_{j,k=0}^{\infty} \omega_{j,k} r_{j,k} g(x) G^{r_{j,k}-1}(x) \\ &= \sum_{j,k}^{\infty} \omega_{j,k} h_{r_{j,k}}(x) \end{aligned}$$

com $\omega_{j,k} = \omega_{j,k}(\beta, \delta, \theta) = \binom{\delta j}{k}(-1)^{j+k} \frac{\Gamma(\theta+j)\, \beta^j}{\Gamma(\theta)\, j!}$ e $r_{j,k} = r_{j,k}(\alpha, \gamma) = \gamma k - \alpha j$.

Desta forma, as funções de distribuição e de densidade da família de distribuições MOGE podem ser escritas como combinações lineares da família de distribuição exponenciada.

3.3 Momento de Ordem r

Utilizando-se a expressão (2.9) é possível se obter os momentos da família de distribuições $MOGE$.

Teorema 3. *Seja a densidade da família de distribuição $X \sim MOGE(\alpha, \beta, \gamma, \delta, \theta)$. O $r-$ésimo momento é uma combinação linear de momentos probabilisticamente ponderados (2.9).*

Prova

Seja a densidade na forma expandida (3.7) para $\beta > \frac{[1-G^{\gamma}(x)]^{\delta}}{\Gamma^{\alpha}(x)}$. Então o $s-$ésimo momento é

$$\begin{aligned} E(X^r) &= \int_{-\infty}^{\infty} x^r f(x)\, dx \\ &= \sum_{j,k=0}^{\infty} v_{j,k} \int_{-\infty}^{\infty} x^r G^{\alpha(\theta+j)+\gamma k-1}(x) g(x) dx \\ &= \sum_{j,k=0}^{\infty} v_{j,k}\, \tau_{r,\alpha(\theta+j)+\gamma k-1}. \end{aligned} \tag{3.9}$$

Brito (2014) define os momentos probabilísticos ponderados generalizados, para incluir integrais de Lesbegue:

$$\tau_{r,s,u} = E[X^r f(X)^s F(X)^u] = \int_{-\infty}^{\infty} x^r f(x)^s F(x)^u dF(x). \tag{3.10}$$

Em sua notação a equação (3.9) se torna

$$E(X^r) = \sum_{j,k=0}^{\infty} v_{j,k}\, \tau_{r,0,\alpha(\theta+j)+\gamma k-1}. \tag{3.11}$$

Por outro lado, se $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$, utiliza-se a função densidade (3.8) e se encontra

$$
\begin{aligned}
E(X^r) &= \int_{-\infty}^{\infty} x^r f(x)\, dx \\
&= \sum_{j,k=0}^{\infty} w_{j,k}\, \tau_{r,\gamma k-\alpha j-1}, \qquad (3.12)
\end{aligned}
$$

(c.q.d.)

em que $\tau_{r,s}$ dado pela fórmula (2.9) é o momento probabilisticamente ponderado da distribuição de base. Utilizando a equação (3.10), o r-ésimo momento (3.12) se torna

$$
E(X^r) = \sum_{j,k=0}^{\infty} w_{j,k}\, \tau_{r,0,\gamma k-\alpha j-1}. \qquad (3.13)
$$

3.4 Função Geradora de Momentos

Para a família de distribuições $MOGE$ segue que $\mu'_r = E(X^r)$ é dado pela expressão (3.9) se $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$ ou é dado por (3.12) para $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$. No entanto, da forma em que está escrita a função de distribuição perde uma de suas principais utilidades, uma vez que desta forma ela requer a equação dos momentos, que geralmente se tem interesse em obter da própria fgm.

A seguir se propõe duas formas alternativas para a fgm da família de distribuições de Marshall-Olkin Generalizada Exponenciada.

Teorema 4. *Seja $X \sim MOGE(\alpha, \beta, \gamma, \delta, \theta)$, sua função geradora de momentos pode ser escrita como combinação linear entre momentos probabilisticamente ponderados.*

Prova

$$
\begin{aligned}
M_X(t) &= E(e^{tX}) \\
&= \sum_{j,k=0}^{\infty} v_{j,k} \int_{-\infty}^{\infty} e^{tx} G^{\alpha(\theta+j)+\gamma k-1} g(x) dx.
\end{aligned}
$$

Utilizando-se a expansão $\exp\{u\} = \sum_{r=0}^{\infty} u^r/r!$:

$$M_X(t) = \sum_{j,k,r=0}^{\infty} v_{j,k,r} \int_{--\infty}^{\infty} x^r G^{\alpha(\theta+j)+\gamma k-1} g(x)dx$$
$$= \sum_{j,k,r=0}^{\infty} v_{j,k,r}\, \tau_{r,\alpha(\theta+j)+\gamma k-1},$$

em que

$$v_{j,k,r} = \frac{\Gamma(\theta+j)\Gamma(\delta(\theta+j)+k)(-1)^j[\alpha(\theta+j)+\gamma k]t^r}{\Gamma(\theta)\Gamma(\delta(\theta+j))\,\beta^{\theta+j}\,j!\,k!\,r!}$$

para $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$. Caso contrário, para $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$:

$$M_X(t) = \sum_{j,k,r=0}^{\infty} w_{j,k,r}\, \tau_{r,\gamma k-\alpha j-1},$$

em que

$$w_{j,k,r} = \binom{\delta j}{k} \frac{(-1)^{j+k}\Gamma(\theta+j)\beta^j(\gamma k-\alpha j)t^r}{\Gamma(\theta)j!r!}.$$

(c.q.d.)

Teorema 5. *Seja $X \sim MOGE(\alpha,\beta,\gamma,\delta,\theta)$, sua função geradora de momentos pode ser escrita como combinação linear de expressões da forma (2.10).*

Prova

$$M_X(t) = \sum_{j,k=0}^{\infty} v_{j,k} \int_{-\infty}^{\infty} e^{tx} G^{\gamma k+\alpha(j+\theta)-1} g(x)dx$$
$$= \sum_{j,k=0}^{\infty} v_{j,k}\, \rho_{t,\gamma k+\alpha(j+\theta)-1}.$$

A equação acima é válida para $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$. Se, por outro lado, $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$ se tem que

$$M_X(t) = \sum_{j,k=0}^{\infty} w_{j,k} \int_{-\infty}^{\infty} e^{tx} G^{\gamma k-\alpha j-1} g(x)dx$$
$$= \sum_{j,k=0}^{\infty} w_{j,k}\, \rho_{t,\gamma k-\alpha j-1}.$$

(c.q.d.)

3.5 Expansão para a Função Característica para a Classe Marshall-Olkin Generalizada Exponenciada

A seguir são desenvolvidos os cálculos da expansão para a função característica para a classe Marshall-Olkin generalizada exponenciada:

Para $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$, tem-se que:

$$\begin{aligned}
\phi_X(t) &= \int_{-\infty}^{+\infty} e^{\imath tx} \sum_{j,k=0}^{\infty} w_{j,k} g(x) G^{\gamma k-\alpha j-1}(x) dx \\
&= \sum_{j,k=0}^{\infty} w_{j,k} \int_{-\infty}^{+\infty} e^{\imath tx} g(x) G^{\gamma k-\alpha j-1}(x) dx.
\end{aligned}$$

Dado que

$$e^{\imath tx} = \sum_{m=0}^{\infty} \frac{\imath^m t^m x^m}{m!},$$

segue que:

$$\begin{aligned}
\phi_X(t) &= \sum_{j,k=0}^{\infty} w_{j,k} \int_{-\infty}^{+\infty} \sum_{m=0}^{\infty} \frac{\imath^m t^m x^m}{m!} g(x) G^{\gamma k-\alpha j-1}(x) dx \\
&= \sum_{j,k,m=0}^{\infty} w_{j,k} \frac{\imath^m t^m}{m!} \int_{-\infty}^{+\infty} x^m g(x) G^{\gamma k-\alpha j-1}(x) dx.
\end{aligned}$$

Portanto,

$$\phi_X(t) = \sum_{j,k,m=0}^{\infty} \frac{w_{j,k} \imath^m t^m}{m!} \tau_{m,0,\gamma k-\alpha j-1}.$$

Para $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$, tem-se que:

$$\begin{aligned}
\phi_X(t) &= \int_{-\infty}^{+\infty} e^{\imath tx} \sum_{j,k=0}^{\infty} v_{j,k} g(x) G^{\alpha(\theta+j)+\gamma k-1}(x) dx \\
&= \sum_{j,k=0}^{\infty} v_{j,k} \int_{-\infty}^{+\infty} e^{\imath tx} g(x) G^{\alpha(\theta+j)+\gamma k-1}(x) dx.
\end{aligned}$$

Como

$$e^{itx} = \sum_{m=0}^{\infty} \frac{\imath^m t^m x^m}{m!},$$

segue que

$$\phi_X(t) = \sum_{j,k=0}^{\infty} v_{j,k} \int_{-\infty}^{+\infty} \sum_{m=0}^{\infty} \frac{\imath^m t^m x^m}{m!} g(x) G^{\alpha(\theta+j)+\gamma k-1}(x) dx$$
$$= \sum_{j,k=0}^{\infty} v_{j,k} \frac{\imath^m t^m}{m!} \int_{-\infty}^{+\infty} x^m g(x) G^{\alpha(\theta+j)+\gamma k-1}(x) dx.$$

Portanto,

$$\phi_X(t) = \sum_{j,k,m=0}^{\infty} \frac{v_{j,k} \imath^m t^m}{m!} \tau_{m,0,\alpha(\theta+j)+\gamma k-1}.$$

3.6 Expansão para os Momentos Centrais de Ordem r para a Classe Marshall-Olkin Generalizada Exponenciada

A seguir, se desenvolvem os cálculos da expansão para os momentos centrais de ordem r para a classe Marshall-Olkin generalizada exponenciada. Utilizando-se a equação recursiva dos momentos, para $\beta < \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$, tem-se que

$$\mu_r = \sum_{i=0}^{r} \binom{r}{i} (-1)^i \mu^i \mu'_{r-i}.$$

Utilizando-se (3.13) se obtém

$$\mu_{r-i} = \sum_{j,k=0}^{\infty} w_{j,k}\, \tau_{r-i,0,\gamma k-\alpha j-1}.$$

Assim,

$$\mu_r = \sum_{i=0}^{r} \binom{r}{i} (-1)^i \mu^i \sum_{j,k=0}^{\infty} w_{j,k}\, \tau_{r-i,0,\gamma k-\alpha j-1}.$$

Portanto,

$$\mu_r = \sum_{j,k=0,i=0}^{\infty,r} w_{j,k} \binom{r}{i} (-1)^i \mu^i \tau_{r-i,0,\gamma k-\alpha j-1}.$$

Em particular, tem-se que a expansão da variância para a classe Marshall-Olkin Generalizada Exponenciada é dada por

$$\sigma^2 = \mu_2 = \sum_{j,k=0,i=0}^{\infty,2} w_{j,k} \binom{2}{i} (-1)^i \mu^i \tau_{2-i,0,\gamma k-\alpha j-1}.$$

Para $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$, utilizando (3.11) se tem que

$$\mu_i = \sum_{i=0}^{r} \binom{r}{i} (-1)^i \mu^i \mu'_{r-i}.$$

Como

$$\mu_{r-i} = \sum_{j,k=0}^{\infty} v_{j,k}\, \tau_{r-i,0,\alpha(\theta+j)+\gamma k-1},$$

obtém-se que

$$\mu_r = \sum_{i=0}^{r} \binom{r}{i} (-1)^i \mu^i \sum_{j,k=0}^{\infty} v_{j,k}\, \tau_{r-i,0,\alpha(\theta+j)+\gamma k-1}.$$

Portanto,

$$\mu_r = \sum_{j,k=0,i=0}^{\infty,r} v_{j,k} \binom{r}{i} (-1)^i \mu^i \tau_{r-i,0,\alpha(\theta+j)+\gamma k-1}.$$

Em particular, tem-se que a expansão da variância para a classe Marshall-Olkin Generalizada Exponenciada é dada por

$$\sigma^2 = \mu_2 = \sum_{j,k=0,i=0}^{\infty,2} v_{j,k} \binom{2}{i} (-1)^i \mu^i \tau_{2-i,0,\alpha(\theta+j)+\gamma k-1}.$$

3.7 Expansão para o Coeficiente Geral para a Classe Marshall-Olkin Generalizada Exponenciada

A seguir se desenvolvem os cálculos da expansão para o coeficiente geral para a classe Marshall-Olkin generalizada exponenciada. Sabe-se que

$$C_g(r) = \frac{E[(X-\mu)^r]}{\sqrt{\{E[(X-\mu)^2]\}}^r} = \frac{E[(X-\mu)^r]}{\sigma^r} = \frac{\mu_r}{\sigma^r},$$

então se obtém, para $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$, que

$$C_g(r) = \frac{\sum_{j,k=0,i=0}^{\infty,r} w_{j,k} \binom{r}{i} (-1)^i \mu^i \tau_{r-i,0,\gamma k-\alpha j-1}}{\left(\sum_{j,k=0,i=0}^{\infty,2} w_{j,k} \binom{2}{i} (-1)^i \mu^i \tau_{2-i,0,\gamma k-\alpha j-1}\right)^{r/2}}.$$

Em particular, como $\gamma_1 = C_g(3)$ se tem que a expansão para o coeficiente geral para a classe Marshall-Olkin generalizada exponenciada é dada por

$$\gamma_1 = \frac{\sum_{j,k=0,i=0}^{\infty,3} w_{j,k} \binom{3}{i} (-1)^i \mu^i \tau_{3-i,0,\gamma k-\alpha j-1}}{\left(\sum_{j,k=0,i=0}^{\infty,2} w_{j,k} \binom{2}{i} (-1)^i \mu^i \tau_{2-i,0,\gamma k-\alpha j-1}\right)^{3/2}}.$$

Similarmente, como $\gamma_2 = C_g(4)$, tem-se que a expansão para o coeficiente de curtose para a classe Marshall-Olkin generalizada exponenciada:

$$\gamma_2 = \frac{\sum_{j,k=0,i=0}^{\infty,4} w_{j,k} \binom{4}{i} (-1)^i \mu^i \tau_{4-i,0,\gamma k-\alpha j-1}}{\left(\sum_{j,k=0,i=0}^{\infty,2} w_{j,k} \binom{2}{i} (-1)^i \mu^i \tau_{2-i,0,\gamma k-\alpha j-1}\right)^{2}}.$$

Para $\beta > \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$, tem-se

$$C_g(r) = \frac{\sum_{j,k=0,i=0}^{\infty,r} v_{j,k} \binom{r}{i} (-1)^i \mu^i \tau_{r-i,0,\alpha(\theta+j)+\gamma k-1}}{\left(\sum_{j,k=0,i=0}^{\infty,2} v_{j,k} \binom{2}{i} (-1)^i \mu^i \tau_{2-i,0,\alpha(\theta+j)+\gamma k-1}\right)^{r/2}}.$$

Em particular, como $\gamma_1 = C_g(3)$ se tem que a expansão para o coeficiente geral para a classe Marshall-Olkin generalizada exponenciada é dada por:

$$\gamma_1 = \frac{\sum_{j,k=0,i=0}^{\infty,3} v_{j,k} \binom{3}{i} (-1)^i \mu^i \tau_{3-i,0,\alpha(\theta+j)+\gamma k-1}}{\left(\sum_{j,k=0,i=0}^{\infty,2} v_{j,k} \binom{2}{i} (-1)^i \mu^i \tau_{2-i,0,\alpha(\theta+j)+\gamma k-1}\right)^{3/2}}.$$

Similarmente, como $\gamma_2 = C_g(4)$ se tem que a expansão para o coeficiente de curtose para a classe Marshall-Olkin generalizada exponenciada:

$$\gamma_2 = \frac{\sum_{j,k=0,i=0}^{\infty,4} v_{j,k} \binom{4}{i} (-1)^i \mu^i \tau_{4-i,0,\alpha(\theta+j)+\gamma k-1}}{\left(\sum_{j,k=0,i=0}^{\infty,2} v_{j,k} \binom{2}{i} (-1)^i \mu^i \tau_{2-i,0,\alpha(\theta+j)+\gamma k-1}\right)^{2}}.$$

3.8 Função Quantílica

Para a família de distribuições de Marshall-Olkin generalizada exponenciada se utiliza o seguinte conjunto de restrições: $\gamma = \alpha$ e $\delta = 1$, a fim de obter uma solução analítica para função quantílica. Assim, se

$$\left\{\frac{G^\alpha(x)}{G^\alpha(x) + \beta[1 - G^\gamma(x)]^\delta}\right\}^\theta = u$$

sujeita as restrições acima, então

$$Q(p) = F^{-1}(p) = G^{-1}\left[\left(\frac{\beta p^{1/\theta}}{1+\beta p^{1/\theta} - p^{1/\theta}}\right)^{1/\alpha}\right], \tag{3.14}$$

em que $G^{-1}(\cdot)$ é a inversa da função de distribuição de base.

3.9 Estatísticas de Ordem

Estatísticas de ordem são quantidades que frequentemente aparecem em aplicações estatísticas. Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória, isto é, independente e identicamente distribuída. Após ordenados do menor para o maior os elementos desta amostra são escritos na notação

$$X_{(1)} = \min\{X_1, \ldots, X_n\}, \ldots, X_{(k)}, \ldots, X_{(n)} = \max\{X_1, \ldots, X_n\}.$$

O $r-$ésimo elemento $X_{(r)}$ é denominado estatística de ordem da amostra. A densidade e a função de distribuição da estatística de ordem são denotados por $f_{r:n}(x)$ e $F_{r:n}(x)$.

Teorema 6. *Seja uma amostra aleatória $X_1, \ldots, X_n \sim MOGE(\alpha, \beta, \gamma, \delta, \theta)$. A densidade da estatística de ordem $f_{r:n}$ pode ser escrita como a mistura de densidades e funções de distribuições - elevadas a um expoente - da distribuição de base.*

Prova

Para $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$

$$\begin{aligned}
f_{r:n}(x) &= \{B(r, n-r+1)\}^{-1} f(x)\, F^{r-1}(x)[1-F(x)]^{n-r} \\
&= \{B(r, n-r+1)\}^{-1} f(x) \sum_{l=0}^{\infty} \binom{n-r}{l} (-1)^l F^{l+r-1}(x) \\
&= \{B(r, n-r+1)\}^{-1} \sum_{j,k=0}^{\infty} v_{j,k}\, g(x)\, G^{\alpha(\theta+j)+\gamma k-1}(x) \sum_{l=0}^{\infty} \binom{n-r}{l} (-1)^l F^{l+r-1}(x).
\end{aligned}$$

Pode-se utilizar a equação (3.5) e separar o duplo somatório, conforme a expressão recursiva (2.11), no fator

$$F^{l+r-1}(x) = G^{\alpha\theta(l+r-1)}(x) \underbrace{\left\{\sum_{j=0}^{\infty} \nu_j\, G^{\alpha j}(x)\right\}^{l+r-1}}_{(i)} \underbrace{\left\{\sum_{k=0}^{\infty} \nu_{(j),k}\, G^{\gamma k}(x)\right\}^{l+r-1}}_{(ii)},$$

em que

$$\nu_j = \frac{\Gamma(\theta+j)(-1)^j}{\Gamma(\theta)\Gamma(\delta(\theta+j))j!\beta^{\theta+j}} \quad \text{e} \quad \nu_{(j),k} = \frac{\Gamma(\delta(\theta+j)+k)}{k!}.$$

Nessa notação ν_j e $\nu_{(j),k}$ foram separados, de tal forma que ν_j não depende de k.

(i)

$$\left\{\sum_{j=0}^{\infty} \nu_j\, G^{\alpha j}(x)\right\}^{l+r-1} = \sum_{j=0}^{\infty} c_j G^{\alpha j}(x),$$

com $\quad c_0 = \nu_0^{l+r-1} \quad$ e $\quad c_m = \frac{1}{m\,\nu_0}\sum_{j=1}^{m}[j(l+r-1)-m+j]\nu_j c_{m-j}$.

(ii)

$$\left\{\sum_{k=0}^{\infty} \nu_{(j),k}\, G^{\gamma k}(x)\right\}^{l+r-1} = \sum_{k=0}^{\infty} d_k G^{\gamma k}(x),$$

com $\quad d_0 = \nu_{(j),0}^{l+r-1} \quad$ e $\quad d_p = \frac{1}{p\,\nu_{(j),0}}\sum_{k=1}^{p}[k(l+r-1)-p+k]\nu_{(j),k} d_{p-k}$.

Portanto,

$$f_{r:n}(x) = \sum_{j,k,l=0}^{\infty} \xi_{j,k,l}\,\cdot\, g(x)\, G^{\alpha[2j+\theta(l+r)]+2\gamma k-1},$$

em que

$$\xi_{j,k,l} = \binom{n-r}{l}\frac{(-1)^l v_{j,k}\, c_j\, d_k}{B(r,n+r-1)}.$$

Para $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$

$$f_{r:n}(x) = \{B(r,n-r+1)\}^{-1}\sum_{j,k=0}^{\infty} w_{j,k}\, g(x)\, G^{\gamma k-\alpha j-1}(x)\sum_{l=0}^{\infty}\binom{n-r}{l}(-1)^l F^{l+r-1}(x).$$

De forma semelhante ao desenvolvimento anterior, obtém-se

$$F^{l+r-1}(x) = \underbrace{\left\{\sum_{j=0}^{\infty} \omega_j\, G^{-\alpha j}(x)\right\}^{l+r-1}}_{(i)} \underbrace{\left\{\sum_{k=0}^{\infty} \omega_{(j),k}\, G^{\gamma k}(x)\right\}^{l+r-1}}_{(ii)},$$

em que

$$\omega_j = \frac{\Gamma(\theta+j)(-1)^j\beta^j}{\Gamma(\theta)j!} \quad \text{e} \quad \omega_{(j),k} = \binom{\delta j}{k}(-1)^k.$$

(i)

$$\left\{\sum_{j=0}^{\infty} \omega_j\, G^{-\alpha j}(x)\right\}^{l+r-1} = \sum_{j=0}^{\infty} c_j^{*} G^{-\alpha j}(x),$$

com $\quad c_0^* = 1 \quad$ e $\quad c_m^* = \frac{1}{m}\sum_{j=1}^{m}[j(l+r-1)-m+j]\omega_j c_{m-j}^*.$

(ii)

$$\left\{\sum_{k=0}^{\infty}\omega_{(j),k}\, G^{\gamma k}(x)\right\}^{l+r-1} = \sum_{k=0}^{\infty} d_k^* G^{\gamma k}(x),$$

com $\quad d_0^* = 1 \quad$ e $\quad d_p^* = \frac{1}{p}\sum_{k=1}^{p}[k(l+r-1)-p+k]\omega_{(j),k} d_{p-k}^*.$

De forma que,

$$f_{r:n}(x) = \sum_{j,k,l=0}^{\infty} \zeta_{j,k,l} \cdot g(x)\, G^{2(\gamma k-\alpha j)-1}(x),$$

em que

$$\xi_{j,k,l} = \binom{n-r}{l}\frac{(-1)^l w_{j,k}\, c_j^*\, d_k^*}{B(r,n+r-1)}.$$

(c.q.d.)

Em particular, a função de densidade do mínimo é, para $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$,

$$f_{(1)}(x) = \sum_{j,k,l=0}^{\infty} \xi_{j,k,l} \cdot g(x)\, G^{\alpha[2j+\theta(l+1)]+2\gamma k-1},$$

em que

$$\xi_{j,k,l} = \binom{n-1}{l}\frac{(-1)^l v_{j,k}\, c_j\, d_k}{B(1,n)},$$

para $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$,

$$f_{(1)}(x) = \sum_{j,k,l=0}^{\infty} \zeta_{j,k,l} \cdot g(x)\, G^{2(\gamma k-\alpha j)-1}(x),$$

com

$$\xi_{j,k,l} = \binom{n-1}{l}\frac{(-1)^l w_{j,k}\, c_j^*\, d_k^*}{B(1,n)}.$$

A função de densidade do máximo é, para $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$,

$$f_{(n)}(x) = \sum_{j,k,l=0}^{\infty} \xi_{j,k,l} \cdot g(x)\, G^{\alpha[2j+\theta(l+n)]+2\gamma k-1},$$

em que

$$\xi_{j,k,l} = \frac{(-1)^l v_{j,k}\, c_j\, d_k}{B(n,2n-1)}$$

e para $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$

$$f_{(n)}(x) = \sum_{j,k,l=0}^{\infty} \zeta_{j,k,l} \cdot g(x)\, G^{2(\gamma k-\alpha j)-1}(x),$$

em que

$$\xi_{j,k,l} = \frac{(-1)^l w_{j,k}\, c_j^*\, d_k^*}{B(n, 2n-1)}.$$

Também é possível se realizar desenvolvimento semelhante para a função de distribuição da estatística de ordem:

$$F_{r:n}(x) = \sum_{h=r}^{\infty} \binom{n}{h} F^h(x)[1 - F(x)]^{n-h} \tag{3.15}$$

$$= \sum_{h=r}^{\infty} \sum_{i=0}^{\infty} \binom{n}{h} \binom{n-h}{i} (-1)^i F(x)^{h+i}. \tag{3.16}$$

Supondo-se que $\beta > \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$ e aplicando-se a equação recursiva (2.11) em (3.5) para separar o duplo somatório se tem

$$F^{h+i}(x) = G^{\alpha\theta(h+i)}(x) \underbrace{\left\{ \sum_{j=0}^{\infty} \nu_j\, G^{\alpha j}(x) \right\}^{h+i}}_{(i)} \underbrace{\left\{ \sum_{k=0}^{\infty} \nu_{(j),k}\, G^{\gamma k}(x) \right\}^{h+i}}_{(ii)}$$

sendo

$$\nu_j = \frac{\Gamma(\theta + j)(-1)^j}{\Gamma(\theta)\Gamma(\delta(\theta + j))j!\beta^{\theta+j}} \quad \text{e} \quad \nu_{(j),k} = \frac{\Gamma(\delta(\theta + j) + k)}{k!}.$$

(i)

$$\left\{ \sum_{j=0}^{\infty} \nu_j\, G^{\alpha j}(x) \right\}^{h+i} = \sum_{j=0}^{\infty} c_j G^{\alpha j},$$

$$c_0 = \nu_0^{h+i}, \quad c_m = \frac{1}{m\nu_0} \sum_{j=1}^{\infty} [j(h + i) - m + j] \nu_j c_{m-j}.$$

(ii)

$$\left\{ \sum_{k=0}^{\infty} \nu_{(j),k}\, G^{\gamma k}(x) \right\}^{h+i} = \sum_{k=0}^{\infty} d_k\, G^{\gamma k}(x),$$

$$d_0 = \nu_{(j),0}^{h+i}, \quad d_p = \frac{1}{p\, \nu_{(j),0}} \sum_{k=1}^{\infty} [k(h + i) - p + k] \nu_{(j),k} d_{p-k}.$$

Assim,

$$F_{r:n} = \sum_{h=r}^{\infty} \sum_{i,j,k=0}^{\infty} \binom{n}{h} \binom{n-h}{i} (-1)^i\, c_j\, d_k\, G^{\alpha\theta(h+i)+\alpha j+\gamma k}(x).$$

Se, entretanto $\beta < \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$,

$$F^{h+i}(x) = \underbrace{\left\{\sum_{j=0}^{\infty} \omega_j \, G^{-\alpha j}(x)\right\}^{h+i}}_{(i)} \underbrace{\left\{\sum_{k=0}^{\infty} \omega_{(j),k} \, G^{\gamma k}(x)\right\}^{h+i}}_{(ii)},$$

sendo

$$\omega_j = \frac{\Gamma(\theta + j)(-1)^j \beta^j}{\Gamma(\theta) j!} \quad \text{e} \quad \omega_{(j),k} = \binom{\delta j}{k}(-1)^k,$$

tem-se

(i)

$$\left\{\sum_{j=0}^{\infty} \omega_j \, G^{-\alpha j}(x)\right\}^{h+i} = \sum_{j=0}^{\infty} c_j^* G^{-\alpha j},$$

$$c_0^* = 1, \quad c_m^* = \frac{1}{m} \sum_{j=1}^{\infty} [j(h+i) - m + j] \omega_j c_{m-j}^*;$$

(ii)

$$\left\{\sum_{k=0}^{\infty} \omega_{(j),k} \, G^{\gamma k}(x)\right\}^{h+i} = \sum_{k=0}^{\infty} d_k^* G^{\gamma k}(x)$$

$$d_0^* = 1, \quad d_p = \frac{1}{p} \sum_{k=1}^{\infty} [k(h+i) - p + k] \omega_{(j),k} d_{p-k}^*.$$

Assim,

$$F_{r:n} = \sum_{h=r}^{\infty} \sum_{i,j,k=0}^{\infty} \binom{n}{h} \binom{n-h}{i} (-1)^i \, c_j^* \, d_k^* \, G^{\gamma k - \alpha j}(x).$$

Restringindo os parâmetros $\gamma = \alpha$ e $\delta = 1$ os momentos da estatística de ordem são facilmente obtidos por

$$\begin{aligned}
\mu_{r:n}^{(k)} &= \{B(r, n-r+1)\}^{-1} \int_{-\infty}^{\infty} x^k F^{r-1}(x)[1 - F(x)]^{n-r} f(x) dx \\
&= \{B(r, n-r+1)\}^{-1} \sum_{j=0}^{\infty} \binom{n-r}{j} (-1)^j \int_{-\infty}^{\infty} x^k F^{j+r-1}(x) f(x) dx \\
&= \{B(r, n-r+1)\}^{-1} \sum_{j=0}^{\infty} \binom{n-r}{j} (-1)^j \tau_{k,0,j+r-1}
\end{aligned}$$

3.10 Entropia de Rényi

Seja uma variável aleatória X, entropia é uma medida de incerteza, no sentido que se o valor da entropia for maior, menor será a informação e maior será a incerteza sobre a variável X. Considere a entropia de Rényi definida por

$$\mathcal{I}_R(\eta) = \frac{1}{1-\eta} \log \left\{ \int_{-\infty}^{\infty} f^\eta(x) dx \right\}.$$

Utilizando-se a equação (3.2)

$$f^\eta(x) = \beta^\eta \, \theta^\eta \, g^\eta(x) \, G^{-\eta(1+\alpha)}(x)[1 - G^\gamma(x)]^{\eta(\delta-1)} \times$$
$$\times \, [\alpha + (\gamma\delta - \alpha)G^\gamma(x)]^\eta \left\{ \frac{G^\alpha(x)}{G^\alpha(x) + \beta[1 - G^\gamma(x)]^\delta} \right\}^{\eta(\theta+1)}.$$

Agora expandindo-se os termos necessários de $f^\eta(x)$ por (2.5):

$$[1 - G^\gamma(x)]^{\eta(\delta-1)} \left\{ \frac{G^\alpha(x)}{G^\alpha(x) + \beta[1 - G^\gamma(x)]^\delta} \right\}^{\eta(\theta+1)} =$$
$$= \begin{cases} \displaystyle\sum_{j,k=0}^{\infty} \nu_{j,k}(\eta(\theta+1)) G^{\alpha[j+\eta(\theta+1)]+\gamma k}, & \beta > \dfrac{G^\alpha(x)}{[1 - G^\gamma(x)]^\delta} \\ \displaystyle\sum_{j,k=0}^{\infty} \omega_{j,k}(\eta(\theta+1)) G^{\alpha[j+\eta(\theta+1)]+\gamma k}, & \beta < \dfrac{G^\alpha(x)}{[1 - G^\gamma(x)]^\delta}, \end{cases}$$

em que $\nu_{j,k}(\theta) = \nu_{j,k}(\beta, \delta, \theta)$ e $\omega_{j,k}(\theta) = \omega_{j,k}(\beta, \delta, \theta)$, e

$$[\alpha + (\gamma\delta - \alpha)G^\gamma(x)]^\eta = \begin{cases} \displaystyle\sum_{l=0}^{\infty} \binom{n}{l} \frac{(\gamma\delta - \alpha)^l G^{\gamma l}(x)}{\alpha^{l-\eta}}, & \left| \dfrac{(\gamma\delta - \alpha)}{\alpha} G^\gamma(x) \right| < 1 \\ \displaystyle\sum_{l=0}^{\infty} \binom{n}{l} \alpha^l (\gamma\delta - \alpha)^{\eta-l} G^{\gamma(\eta-l)}(x), & \left| \dfrac{(\gamma\delta - \alpha)}{\alpha} G^\gamma(x) \right| > 1. \end{cases}$$

Entretanto, as expansões acima geram quatro sub-casos que se passa a descrever. Para $\left| \frac{(\gamma\delta-\alpha)}{\alpha} G^\gamma(x) \right| < 1$ e $\beta > \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$ se tem

$$f^\eta(x) = \sum_{i,j,k,l=0}^{\infty} \zeta_{i,j,k,l}^{1,1} \, g^\eta(x) \, G^{a_{1,1}}(x),$$

com

$$\zeta_{i,j,k,l}^{1,1} = \beta^\eta \, \theta^\eta \binom{\eta(\delta-1)}{i} \binom{\eta}{l} (-1)^i \frac{(\gamma\delta - \alpha)^l}{\alpha^{l-\eta}} \nu_{j,k}(\eta(\theta+1))$$

e

$$a_{1,1} = \alpha[j + \eta(2+\theta)] + \gamma(i + k + l) - \eta.$$

Se $\left| \frac{(\gamma\delta-\alpha)}{\alpha} G^\gamma(x) \right| < 1$ e $\beta < \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$ se tem

$$f^{\eta}(x) = \sum_{i,j,k,l=0}^{\infty} \zeta_{i,j,k,l}^{1,2}\, g^{\eta}(x)\, G^{a_{1,2}}(x),$$

com

$$\zeta_{i,j,k,l}^{1,2} = \beta^{\eta}\, \theta^{\eta} \binom{\eta(\delta-1)}{i}\binom{\eta}{l}(-1)^{i}\alpha^{l}(\gamma\delta-\alpha)^{\eta-l}\nu_{j,k}(\eta(\theta+1))$$

e

$$a_{1,2} = \alpha[j+\eta(2+\theta)] + \gamma(i+k-l+\eta) - \eta.$$

Para $\left|\frac{(\gamma\delta-\alpha)}{\alpha}G^{\gamma}(x)\right| > 1$ e $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$ se tem que

$$f^{\eta}(x) = \sum_{i,j,k,l=0}^{\infty} \zeta_{i,j,k,l}^{2,1}\, g^{\eta}(x)\, G^{a_{2,1}}(x),$$

com

$$\zeta_{i,j,k,l}^{2,1} = \beta^{\eta}\, \theta^{\eta} \binom{\eta(\delta-1)}{i}\binom{\eta}{l}(-1)^{i}\frac{(\gamma\delta-\alpha)^{l}}{\alpha^{l-\eta}}\omega_{j,k}(\eta(\theta+1))$$

e

$$a_{2,1} = \alpha[\eta-j] + \gamma(i+k+l) - \eta.$$

Se $\left|\frac{(\gamma\delta-\alpha)}{\alpha}G^{\gamma}(x)\right| > 1$ e $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$ se tem

$$f^{\eta}(x) = \sum_{i,j,k,l=0}^{\infty} \zeta_{i,j,k,l}^{2,2}\, g^{\eta}(x)\, G^{a_{2,2}}(x),$$

com

$$\zeta_{i,j,k,l}^{2,2} = \beta^{\eta}\, \theta^{\eta} \binom{\eta(\delta-1)}{i}\binom{\eta}{l}(-1)^{i}\alpha^{l}(\gamma\delta-\alpha)^{\eta-l}\omega_{j,k}(\eta(\theta+1))$$

e

$$a_{2,2} = \alpha[\eta-j] + \gamma(i+k-l+\eta) - \eta.$$

Então, por exemplo, para $\left|\frac{(\gamma\delta-\alpha)}{\alpha}G^{\gamma}(x)\right| > 1$ e $\beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$ a entropia de Rényi é

$$\mathcal{I}_R(\eta) = \frac{1}{1-\eta}\log\left\{\sum_{i,j,k,l=0}^{\infty} \zeta_{i,j,k,l}^{1,1}\int_{-\infty}^{\infty} g^{\eta}(x)\, G^{a_{1,1}}(x)dx\right\}, \qquad (3.17)$$

isto é, a entropia de Rényi pode ser escrita em termos da função de distribuição e de densidade da distribuição de base. A integral $\int_{-\infty}^{\infty} g^{\eta}(x)\, G^{a}(x)dx$ é solúvel quando $g(x)$ e $G(x)$ são indicadas.

3.11 Mediana

Teorema 7. *A mediana m da família de distribuições $MOGE(\alpha, \beta, \gamma, \delta, \theta)$ é a solução da equação*

$$\sum_{j,k=0}^{\infty} v_{j,k} \frac{G^{\alpha(\theta+j)+\gamma k}(m)}{\alpha(\theta+j)+\gamma k} = \frac{1}{2}, \quad \textit{para} \quad \beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}},$$

ou

$$\sum_{j,k=0}^{\infty} w_{j,k} G^{\gamma k-\alpha j}(m) = \frac{1}{2}, \quad \textit{para} \quad \beta < \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}.$$

Prova

A mediana m é a solução da equação $\int_{-\infty}^{m} f(x)\,dx = 1/2$. Primeiramente, tem-se que para $F(x)$ e $f(x)$ sendo uma função de distribuição e de densidade da mesma variável aleatória, vale a seguinte integral

$$\int_{-\infty}^{q} f(x)F^{r}(x)dx = \int_{0}^{F(q)} F^{r}(x)dF(x) = \frac{F^{r+1}(q)}{r+1}.$$

Assim, para $\beta > \frac{G^{\alpha}}{[1-G^{\gamma}]^{\delta}}$, utilizando-se a equação (3.7) a mediana é solução da equação

$$\sum_{j,k=0}^{\infty} v_{j,k} \int_{-\infty}^{m} g(x)G^{\alpha(\theta+j)+\gamma k-1}(x)dx = \frac{1}{2}$$

$$\sum_{j,k=0}^{\infty} v_{j,k} \frac{G^{\alpha(\theta+j)+\gamma k}(m)}{\alpha(\theta+j)+\gamma k} = \frac{1}{2}.$$

Aplicando-se o mesmo raciocínio para $\beta < \frac{G^{\alpha}}{[1-G^{\gamma}]^{\delta}}$, o que implica que a densidade da distribuição proposta é da forma (3.8), conclui-se que

$$\sum_{j,k=0}^{\infty} w_{j,k} \frac{G^{\gamma k-\alpha j}(m)}{\gamma k-\alpha j} = \frac{1}{2}.$$

(c.q.d)

3.12 Desvio Médio e Desvio Mediano

Duas quantidades que medem a dispersão da variável X são o desvio em relação à média e o desvio em relação à mediana, definidos por

$$\delta_1(X) = \int_{-\infty}^{\infty} |x-\mu| f(x)\,dx \quad \text{e} \quad \delta_2(x) = \int_{-\infty}^{\infty} |x-m| f(x)\,dx,$$

respectivamente, em que μ, calculado por (3.9) ou (3.12) com $r = 1$, é a média de X e m é a mediana de X. Para distribuições de suporte positivo, as medidas $\delta_1(X)$ e $\delta_2(X)$ são expressas em Cordeiro e Lemonte (2013) por

$$\delta_1(X) = 2\,\mu\,F(\mu) - 2\,J(\mu) \quad \text{e} \quad \delta_2(X) = \mu - 2\,J(m), \tag{3.18}$$

em que $F(q)$ é obtido de (3.1) e $J(q) = \int_{-\infty}^{q} x\,f(x)\,dx$. Considerando-se o caso em que $\beta > \frac{G^\alpha}{[1-G^\gamma]^\delta}$, pode-se escrever

$$\begin{aligned} J(q) &= \int_{-\infty}^{q} x \sum_{j,k=0}^{\infty} v_{j,k}\, g(x)\, G^{\alpha(\theta+j)+\gamma k-1}(x)\, dx \\ &= \sum_{j,k=0}^{\infty} v_{j,k}\, \tau^{(q)}_{1,\,\alpha(\theta+j)+\gamma k-1}, \end{aligned} \tag{3.19}$$

em que

$$\tau_{s,\,r}(q) = \int_{-\infty}^{q} x^s G^r(x) g(x) dx = \int_{0}^{G(q)} Q(x)^s u^r du$$

e $Q(u) = G^{-1}(x)$, sendo $G(x)$ a função de distribuição de base. Se $\beta > \frac{G^\alpha}{[1-G^\gamma]^\delta}$ a equação (3.19) é substituída por

$$J(q) = \sum_{j,k=0}^{\infty} w_{j,k}\, \tau_{1,\,\gamma k-\alpha j-1}(q). \tag{3.20}$$

3.13 Estimação por Máxima Verossimilhança

Nesta seção se considera o vetor paramétrico $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \delta, \theta, \mu)'$, com $(\alpha, \beta, \gamma, \delta, \theta)'$ sendo o conjunto de parâmetros da classe de distribuições de Marshall-Olkin generalizada exponenciada e μ é um parâmetro da distribuição de base, cujas funções de distribuição e de densidade são representadas por $G(x; \mu)$ e $g(x; \mu) = dG(x; \mu)/d\,x$, respectivamente. Utilizando-se a f.d.p. (3.2), a função de log-verossimilhança fica

$$\begin{aligned} l(\boldsymbol{\theta}) = n\log(\beta) + n\log(\theta) &+ \sum_{j=0}^{\infty} \log\left(g\left(x_j; \mu\right)\right) - (\alpha+1)\sum_{j=0}^{\infty} \log\left(G(x_j; \mu)\right) \\ &+ (\delta-1)\sum_{j=0}^{\infty} \log\left(1 - G(x_j; \mu)^\gamma\right) + \sum_{j=0}^{\infty} \log\left(\alpha + (\gamma\delta - \alpha)G(x_j; \mu)^\gamma\right) \\ &+ (\theta+1)\sum_{j=0}^{\infty} \log\left(G(x_j; \mu)^\alpha + \beta\left(1 - G(x_j; \mu)^\gamma\right)^\delta\right). \end{aligned}$$

Sejam as seguintes derivadas:

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha}=\sum_{j=0}^{\infty}\frac{1-G(x_j;\mu)^{\gamma}}{\alpha+(-\alpha+\gamma\delta)G(x_j;\mu)^{\gamma}}+(1+\theta)\sum_{j=0}^{\infty}\frac{G(x_j;\mu)^{\alpha}\log(G(x_j;\mu))}{G(x_j;\mu)^{\alpha}+\beta\left(1-G(x_j;\mu)^{\gamma}\right)^{\delta}}$$
$$-\sum_{j=0}^{\infty}\log(G(x_j;\mu))$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta}=\frac{n}{\beta}+(1+\theta)\sum_{j=0}^{\infty}\frac{\left(1-G(x_j;\mu)^{\gamma}\right)^{\delta}}{G(x_j;\mu)^{\alpha}+\beta\left(1-G(x_j;\mu)^{\gamma}\right)^{\delta}}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \gamma}=-(\delta-1)\sum_{j=0}^{\infty}\frac{G(x_j;\mu)^{\gamma}\log(G(x_j;\mu))}{1-G(x_j;\mu)^{\gamma}}$$
$$+\sum_{j=0}^{\infty}\frac{\delta G(x_j;\mu)^{\gamma}+(-\alpha+\gamma\delta)G(x_j;\mu)^{\gamma}\log(G(x_j;\mu))}{\alpha+(-\alpha+\gamma\delta)G(x_j;\mu)^{\gamma}}$$
$$-(1+\theta)\sum_{j=0}^{\infty}\frac{\beta\delta G(x_j;\mu)^{\gamma}\left(1-G(x_j;\mu)^{\gamma}\right)^{-1+\delta}\log(G(x_j;\mu))}{G(x_j;\mu)^{\alpha}+\beta\left(1-G(x_j;\mu)^{\gamma}\right)^{\delta}}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \delta}=\sum_{j=0}^{\infty}\frac{\gamma G(x_j;\mu)^{\gamma}}{\alpha+(-\alpha+\gamma\delta)G(x_j;\mu)^{\gamma}}+\sum_{j=0}^{\infty}\log\left(1-G(x_j;\mu)^{\gamma}\right)$$
$$+(1+\theta)\sum_{j=0}^{\infty}\frac{\beta\left(1-G(x_j;\mu)^{\gamma}\right)^{\delta}\log\left(1-G(x_j;\mu)^{\gamma}\right)}{G(x_j;\mu)^{\alpha}+\beta\left(1-G(x_j;\mu)^{\gamma}\right)^{\delta}}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \theta}=\frac{n}{\theta}+\sum_{j=0}^{\infty}\log\left(G(x_j;\mu)^{\alpha}+\beta\left(1-G(x_j;\mu)^{\gamma}\right)^{\delta}\right)$$

$$
\begin{aligned}
\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} &= \sum_{j=0}^{\infty} \frac{\partial \log(G(x_j;\mu))}{\partial \mu} - (1+\alpha)\sum_{j=0}^{\infty} \frac{\partial \log(G(x_j;\mu))}{\partial \mu} \\
&+ (-1+\delta) - \sum_{j=0}^{\infty} \frac{\gamma G(x_j;\mu)^{-1+\gamma}}{1-G(x_j;\mu)^{\gamma}} \frac{\partial G(x_j;\mu)}{\partial \mu} \\
&+ (1+\theta)\sum_{j=0}^{\infty} \frac{\alpha G(x_j;\mu)^{-1+\alpha} - \beta\gamma\delta G(x_j;\mu)^{-1+\gamma}\left(1-G(x_j;\mu)^{\gamma}\right)^{-1+\delta}}{G(x_j;\mu)^{\alpha} + \beta\left(1-G(x_j;\mu)^{\gamma}\right)^{\delta}} \frac{\partial G(x_j;\mu)}{\partial \mu} \\
&+ \sum_{j=0}^{\infty} \frac{\gamma(-\alpha+\gamma\delta)G(x_j;\mu)^{-1+\gamma}}{\alpha+(-\alpha+\gamma\delta)G(x_j;\mu)^{\gamma}} \frac{\partial G(x_j;\mu)}{\partial \mu},
\end{aligned}
$$

que formam o vetor escore $U = \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha}, \frac{\partial l(\boldsymbol{\theta})}{\partial \beta}, \frac{\partial l(\boldsymbol{\theta})}{\partial \gamma}, \frac{\partial l(\boldsymbol{\theta})}{\partial \delta}, \frac{\partial l(\boldsymbol{\theta})}{\partial \theta}, \frac{\partial l(\boldsymbol{\theta})}{\partial \mu}\right)'$. Resolvendo o sistema $U = 0$ se encontra o estimador de máxima verossimilhança (EMV) $\widehat{\boldsymbol{\theta}} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\delta}, \widehat{\theta}, \widehat{\mu})'$ de $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \delta, \theta, \mu)'$. Uma vez que não é possível a solução analítica dessas equações, necessita-se de um método numérico-iterativo, tal como o método BFGS (BROYDEN, 1970; FLETCHER, 1970; GOLDFARB, 1970; SHANNO, 1970) ou método do gradiente conjugado (HESTENES; STIEFEL, 1952), ambos implementáveis no software estatístico R (R Core Team, 2014).

Pode-se utilizar a aproximação normal do estimador de máxima verossimilhança de $\boldsymbol{\theta}$ para testar hipóteses e construir intervalos de confiança aproximados para os parâmetros α, β, γ, δ, θ e μ. Sob a suposição de que a diferença entre $\widehat{\boldsymbol{\theta}}$ e $\boldsymbol{\theta}$ tem distribuição normal assintótica, tem-se que $\boldsymbol{\sqrt{n}(\hat{\theta} - \theta) \overset{a}{\sim} N_p(0, n^{-1}J_n(\theta))}$, em que p é o número de parâmetros em $\boldsymbol{\theta}$ ($p = 6$, neste exemplo) e $\boldsymbol{J_n(\theta)}$ é a matriz de informação observada de Fisher

$$
\boldsymbol{J_n(\theta)} = -\begin{bmatrix}
\boldsymbol{U_{\alpha,\alpha}} & \boldsymbol{U_{\alpha,\beta}} & \boldsymbol{U_{\alpha,\gamma}} & \boldsymbol{U_{\alpha,\delta}} & \boldsymbol{U_{\alpha,\theta}} & \boldsymbol{U_{\alpha,\mu}} \\
\cdot & \boldsymbol{U_{\beta,\beta}} & \boldsymbol{U_{\beta,\gamma}} & \boldsymbol{U_{\beta,\delta}} & \boldsymbol{U_{\beta,\theta}} & \boldsymbol{U_{\beta,\mu}} \\
\cdot & \cdot & \boldsymbol{U_{\gamma,\gamma}} & \boldsymbol{U_{\gamma,\delta}} & \boldsymbol{U_{\gamma,\theta}} & \boldsymbol{U_{\gamma,\mu}} \\
\cdot & \cdot & \cdot & \boldsymbol{U_{\delta,\delta}} & \boldsymbol{U_{\delta,\theta}} & \boldsymbol{U_{\delta,\mu}} \\
\cdot & \cdot & \cdot & \cdot & \boldsymbol{U_{\theta,\theta}} & \boldsymbol{U_{\theta,\mu}} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \boldsymbol{U_{\mu,\mu}}
\end{bmatrix}.
$$

Veja o **Apêndice A** para a expressão de cada elemento da matriz de informação observada. Conforme Seção 2.21, pode-se utilizar uma estratégia bootstrap para obter erros-padrão estimados para $\widehat{\boldsymbol{\theta}}$. Para se testar um conjunto de hipóteses para modelos hierárquicos (modelo versus sub-modelo) utiliza-se o teste da razão de verossimilhanças discutido anteriormente.

CAPÍTULO 4

A Distribuição Marshall-Olkin Generalizada Exponenciada de Weibull

Seja $T \sim MOGE(\alpha, \beta, \gamma, \delta, \theta)$ com função de distribuição dada em (3.1) e respectiva função de densidade (3.2). Seja ainda a distribuição de Weibull com função de distribuição $G(t) = 1 - e^{-(t/\lambda)^k}$. É possível se obter a função de distribuição da distribuição de Marshall-Olkin generalizada exponenciada de Weibull $T \sim MOGE(\alpha, \beta, \gamma, \delta, \theta, \lambda, k)$:

$$F(t) = \left\{ \frac{\left(1 - e^{-(t/\lambda)^k}\right)^{\alpha}}{\left(1 - e^{-(t/\lambda)^k}\right)^{\alpha} + \beta \left[1 - \left(1 - e^{-(t/\lambda)^k}\right)^{\gamma}\right]^{\delta}} \right\}^{\theta} \tag{4.1}$$

sendo $\lambda > 0$ o parâmetro de escala, $\theta > 0$ o parâmetro de forma, $\alpha > 0, \beta > 0, \gamma > 0, \delta > 0$ e $\theta > 0$ os parâmetros adicionais, referentes à nova classe. Se $\beta = \delta = \theta = 1$ e $\alpha = \gamma$, tem-se a distribuição de Weibull Exponenciada (WE) e se $\alpha = \gamma = \delta = \theta = 1$, tem-se a distribuição de Marshall-Olkin Weibull (MOW). Note que a distribuição MOW é a mesma tratada por Cordeiro e Lemonte (2013) fazendo $\gamma = 1/\lambda^k$, isto é, a distribuição de Marshall-Olkin Weibull Estendida é a distribuição de Marshall-Olkin Weibull com uma reparametrização diferente, sendo portanto um sub-modelo da nossa distribuição. Particularmente, se todos os parâmetros adicionais são unitários, tem-se a distribuição de Weibull, isto é: $F(t) = G(t)$. Utilizando-se a expansão (3.5), para $\beta > \frac{G^{\alpha}(x)}{[1-G^{\gamma}(x)]^{\delta}}$ se tem que

$$F(t) = \sum_{j,k=0}^{\infty} \nu_{j,k} \, \left[1 - e^{-(t/\lambda)^k}\right]^{\alpha(\theta+j)+\gamma k},$$

em que $\nu_{j,k} = \nu_{j,k}(\beta, \delta, \theta) = \frac{\Gamma(\theta+j)\Gamma(\delta(\theta+j)+k)(-1)^j}{\Gamma(\theta)\Gamma(\delta(\theta+j))\, j!\, k!\, \beta^{\theta+j}}$. Para $\beta < \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$ se substitui $\nu_{j,k}$ por $\omega_{j,k} = \omega_{j,k}(\beta, \delta, \theta) = \binom{\delta j}{k}(-1)^{j+k}\frac{\Gamma(\theta+j)\,\beta^j}{\Gamma(\theta)\, j!}$ e $\alpha(\theta + j) + \gamma k$ por $\gamma k - \alpha j$ na equação acima, isto é,

$$F(t) = \sum_{j,k=0}^{\infty} \omega_{j,k} \, \left[1 - e^{-(t/\lambda)^k}\right]^{\gamma k-\alpha j}.$$

A densidade é obtida de (3.7) fazendo a função de distribuição de base se $G(t) = 1 - e^{-(t/\lambda)^k}$ e a densidade da distribuição de base ser $g(t) = k/\lambda(t/\lambda)^{k-1}e^{-(t/\lambda)^k}$:

$$f(t) = \sum_{j,k=0}^{\infty} v_{j,k} \, \frac{k}{\lambda} \left(\frac{t}{\lambda}\right)^{k-1} e^{-(t/\lambda)^k} \, \left[1 - e^{-(t/\lambda)^k}\right]^{\alpha(\theta+j)+\gamma k-1}, \tag{4.2}$$

em que

$$v_{j,k} = v_{j,k}(\alpha, \beta, \gamma, \delta, \theta) = [\alpha(\theta+j)+\gamma k]\, \nu_{j,k} = \frac{\Gamma(\theta + j)\Gamma(\delta(\theta + j) + k)(-1)^j[\alpha(\theta + j) + \gamma k]}{\Gamma(\theta)\Gamma(\delta(\theta + j))\, j!\, k!\, \beta^{\theta+j}},$$

como antes. Se $\beta < \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$:

$$f(x) = \sum_{j,k=0}^{\infty} w_{j,k} \, \frac{k}{\lambda} \left(\frac{t}{\lambda}\right)^{k-1} e^{-(t/\lambda)^k} \, \left[1 - e^{-(t/\lambda)^k}\right]^{\gamma k-\alpha j-1}(x), \tag{4.3}$$

em que $w_{j,k} = w_{j,k}(\alpha, \beta, \gamma, \delta, \theta) = (\gamma k - \alpha j)\, \omega_{j,k} = \binom{\delta j}{k}(-1)^{j+k}\frac{\Gamma(\theta+j)\,\beta^j(\gamma k-\alpha j)}{\Gamma(\theta)\, j!}$.

Ainda é possível se expandir, com o auxílio da fórmula (2.5), os termos entre colchetes de (4.2) e (4.3) obtendo-se

$$f(t) = \begin{cases} \sum_{j,k,l=0}^{\infty} v^*_{j,k,l}\, f_{\lambda,k,(1+l)}(t) & \text{se } \beta > \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}; \\ \sum_{j,k,l=0}^{\infty} w^*_{j,k,l}\, f_{\lambda,k,(1+l)}(t) & \text{se } \beta < \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}, \end{cases} \tag{4.4}$$

em que

$$v^*_{j,k,l} = \binom{\alpha(\theta + j) + \gamma k - 1}{l}(-1)^l v_{j,k},$$

$$w^*_{j,k,l} = \binom{\gamma k - \alpha j}{l}(-1)^l w_{j,k}$$

e $f_{\lambda,k,\vartheta}(t)$ é a densidade de uma nova distribuição não identificada na literatura consultada, dada por

$$f_{\lambda,k,\vartheta}(t)\frac{k\left(\frac{t}{\lambda}\right)^{k-1}}{\lambda\Gamma\left(1+\frac{1}{\vartheta}\right)}\exp\left\{-\left(\frac{t}{\lambda}\right)^{\vartheta k}\right\}, \tag{4.5}$$

cujo r-ésimo momento é

$$E(T^r) = \frac{\lambda^r\Gamma\left(\frac{k+r}{k\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)}.$$

Essa distribuição tem densidade e propriedades muito similares às da distribuição gama generalizada, embora não sejam a mesma distribuição. O principal ganho em se utilizar as equações (4.4) é que quantidades, tais como momentos da nova distribuição podem ser calculados diretamente das quantidades de (4.5).

Na Figura 4.1 é representada a densidade $MOGEW$ se variando um ou dois parâmetros e fixando os demais. É possível se observar que, mesmo se variando apenas um ou dois parâmetros, os formatos obtidos são bastante variados, incluindo forma leptocúrtica (curva vermelha tracejada), simétrica (preta cheia) assimétricas positivas (curvas cheias verde e azul) e bimodal (preta tracejada), o que fornece um indício da flexibilidade do modelo.

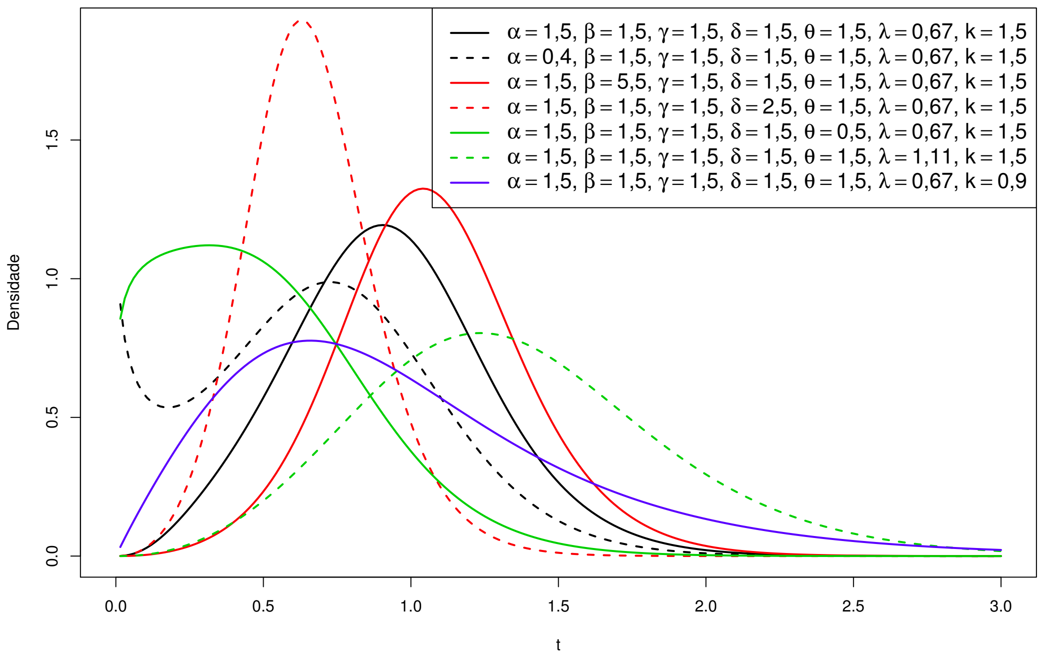


Figura 4.1: Gráfico para a função de densidade da distribuição MOGEW para diversos valores de $\alpha, \beta, \gamma, \delta, \theta, \lambda$ e k.

A função de sobrevivência, por outro lado, é obtida pela equação (3.3)

$$S(x) = 1 - \left\{ \frac{G(x)^{\alpha}}{G(x)^{\alpha} + \beta\,[1 - G(x)^{\gamma}]^{\delta}} \right\}^{\theta},$$

fazendo-se $G(t) = 1 - e^{-(t/\lambda)^k}$.

Assim, a função de risco é, segundo a equação (3.4), dada por

$$h(x) = \frac{\beta\,\theta\,g(x)\,G(x)^{-1-\alpha}\,[1 - G(x)^{\gamma}]^{\delta-1}\,[\alpha + (-\alpha + \gamma\delta)G(x)^{\gamma}]\left\{\frac{G(x)^{\alpha}}{G(x)^{\alpha}+\beta[1-G(x)^{\gamma}]^{\delta}}\right\}^{1+\theta}}{1 - \left\{\frac{G(x)^{\alpha}}{G(x)^{\alpha}+\beta[1-G(x)^{\gamma}]^{\delta}}\right\}^{\theta}}, \tag{4.6}$$

fazendo-se $G(t) = 1 - e^{-(t/\lambda)^k}$.

A Figura 4.2 expressa diversos formatos da função de risco dessa distribuição. É possível se notar comportamentos constantes, crescentes e decrescentes, como também os formatos de U (curva preta tracejada) e de U invertido (curvas cheias verde e vermelha), bastante requeridos em Análise de Sobrevivência (COLOSIMO; GIOLO, 2006). Também é possível se notar os dois comportamentos na mesma função (isto é, risco decrescente-crescente-decrescente: curva verde tracejada). Isso, outra vez, evidencia a flexibilidade do modelo.

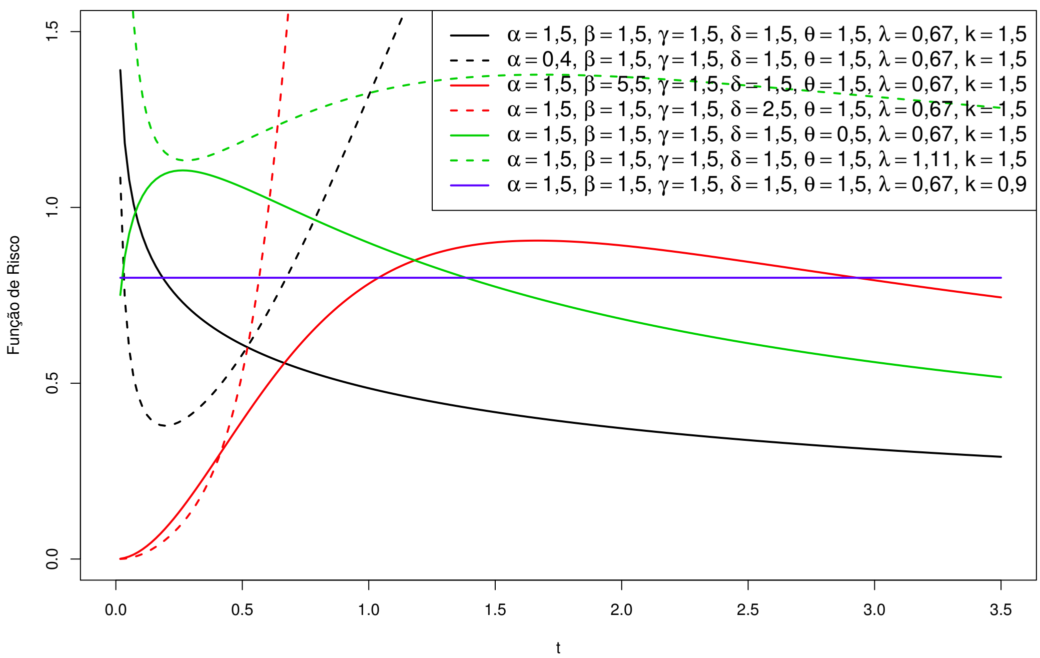


Figura 4.2: Gráfico para a função de risco da distribuição MOGEW para diversos valores de $\alpha, \beta, \gamma, \delta, \theta, \lambda$ e k.

O $r-$ésimo momento é obtido por uma das equações (3.9) e (3.12), com $G(t) = 1 - e^{-(t/\lambda)^k}$ e $g(t) = k/\lambda(t/\lambda)^{k-1}e^{-(t/\lambda)^k}$. Ou ainda, utilizando (4.4), o r-ésimo momento fica

$$E(T^r) = \begin{cases} \displaystyle\sum_{j,k,l=0}^{\infty} v^*_{j,k,l} \frac{\lambda^r \Gamma\left(\frac{k+r}{k(l+1)}\right)}{\Gamma\left(\frac{1}{l+1}\right)} & \text{se } \beta > \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}; \\ \displaystyle\sum_{j,k,l=0}^{\infty} w^*_{j,k,l} \frac{\lambda^r \Gamma\left(\frac{k+r}{k(l+1)}\right)}{\Gamma\left(\frac{1}{l+1}\right)} & \text{se } \beta < \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}. \end{cases} \tag{4.7}$$

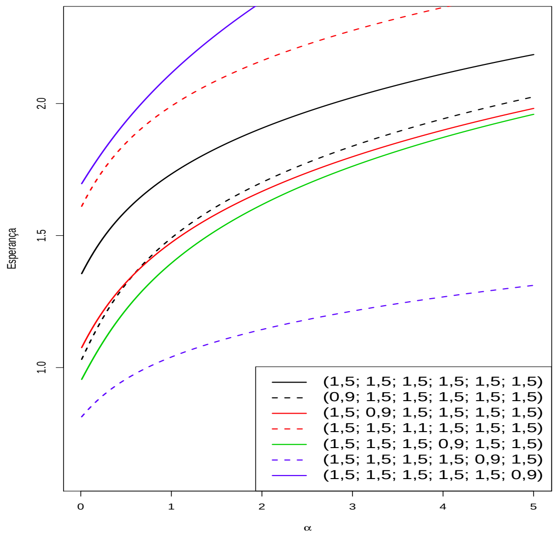
Estas equações permitem encontrar os momentos centrais, particularmente, média, variância, coeficiente de assimetria e de curtose. Assim, por exemplo para $\beta > \frac{G^\alpha(x)}{[1-G^\gamma(x)]^\delta}$, a esperança e a variância de T são, respectivamente,

$$E(T) = \sum_{j,k,l=0}^{\infty} v^*_{j,k,l} \frac{\lambda^r \Gamma\left(\frac{k+r}{k(l+1)}\right)}{\Gamma\left(\frac{1}{l+1}\right)}.$$

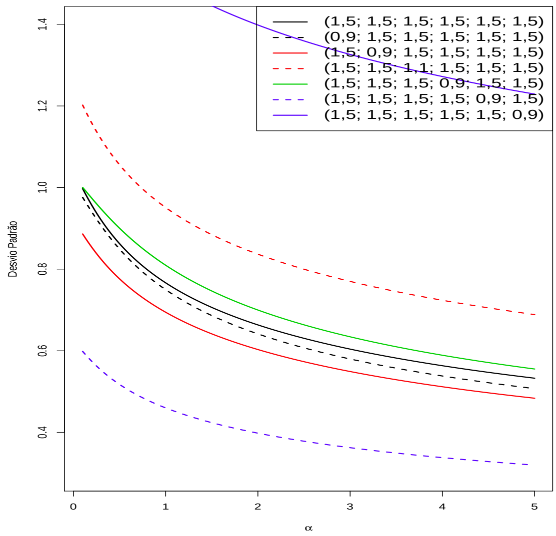
e

$$Var(T) = \sum_{j,k,l=0}^{\infty} v^*_{j,k,l} \frac{\lambda^2 \left(\Gamma\left(\frac{1}{l+1}\right)\Gamma\left(\frac{2+k}{k(l+1)}\right)\right) - \Gamma\left(\frac{1+k}{k(l+1)}\right)^2}{\Gamma\left(\frac{1}{l+1}\right)^2}.$$

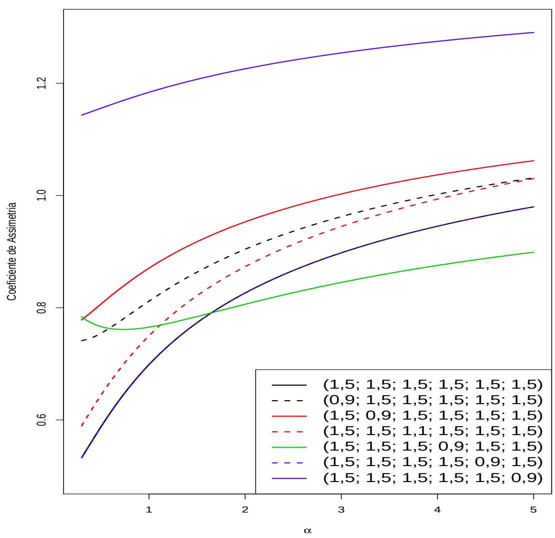
Na Figura 4.3 são mostradas estas medidas para a distribuição MOGEW para o vetor $(\beta, \gamma, \delta, \theta, \lambda, k)'$ em função de α. No gráfico (a) se observa que a esperança cresce quando α aumenta. Se tomar $\beta = 1,5; \gamma = 1,5; \delta = 1,5; \theta = 1,5; \lambda = 1,5; k = 1,5$ (linha preta cheia) se observa que quando se diminui o valor de β (preta tracejada), γ (vermelha cheia), θ (verde cheia) ou λ (azul tracejada) a média da distribuição MOGEW diminui; ao passo que quando se diminui δ (vermelha tracejada) ou k (azul cheia) a média aumenta. Em (b) se observa que o desvio-padrão diminui quando se aumenta α. Diminuir δ, k ou θ aumenta a variância, enquanto diminuir β, γ ou λ diminuem a variância. No gráfico (c) se observa que ao se aumentar α, de modo geral, que o coeficiente de assimetria aumenta, exceto para valores de θ (verde cheia) num intervalo entre zero e um, quando a assimetria diminui. Observa-se que λ não tem influência expressiva na assimetria, pois as curvas preta cheia e azul tracejada estão praticamente justapostas. A assimetria, em geral diminui quando θ diminui. Para os outros parâmetros ela aumenta. Em (d) se observa que o coeficiente de curtose aumenta quando α aumenta. Para θ menor, a densidade tende a ser mais achatada, enquanto para os demais tende a ser mais alongada que a densidade de referencia, exceto para λ que não influencia o coeficiente de curtose na mesma proporção que os demais parâmetros, a exemplo do que ocorre com o coeficiente de assimetria.



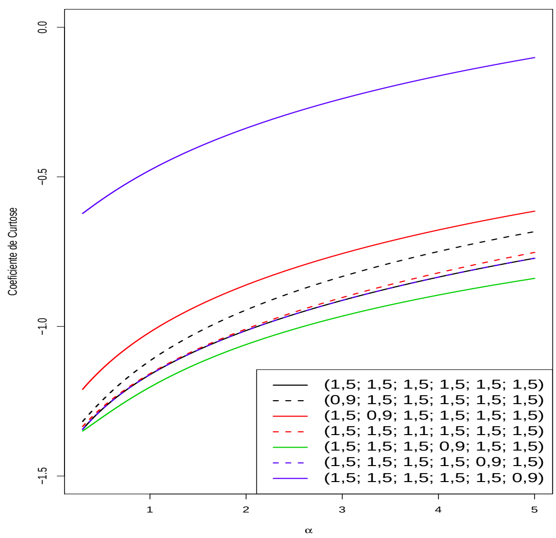
(a)



(b)



(c)



(d)

Figura 4.3: Gráficos das principais funções do r-ésimo momento para a distribuição MO-GEW em função do parâmetro α. Em (a) a esperança, em (b) o desvio-padrão, em (c) o coeficiente de assimetria, e em (d) o coeficiente de curtose para alguns valores de $(\beta, \gamma, \delta, \theta, \lambda, k)'$.

4.1 Função Quantílica

Invertendo a f.d. (4.1), para $\gamma = 1$ (restrição necessária para se obter uma solução para a inversa) se obtém a função quantílica da distribuição Marshall-Olkin Generalizada Exponenciada de Weibull

$$t = Q(p) = \frac{1}{\lambda}\left\{-\log\left[1-\left(\frac{p^{1/\theta}\beta}{1-p^{1/\theta}+p^{1/\theta}\beta}\right)^{1/\alpha}\right]\right\}^{1/\kappa}. \tag{4.8}$$

Com a função quantílica (4.8) é possível simular números pseudo-aleatórios da distribuição $T \sim MOGEW(\alpha,\beta,\gamma,\delta,\theta,\lambda,k)$ a partir de $p \sim U(0,1)$, isto é, p simulado de uma distribuição uniforme no intervalo zero-um. Na Figura 4.4 foi gerada uma amostra de tamanho 10000 da distribuição T, pela função (4.8) com parâmetros fixados em $\alpha = 3$, $\beta = 5$, $\gamma = \alpha$, $\delta = 1$, $\theta = 2$, $\lambda = 2$ e $\kappa = 1$. Também foi ajustada, em azul, a fdp (3.2) verdadeira, isto é, com os mesmos valores dos parâmetros.

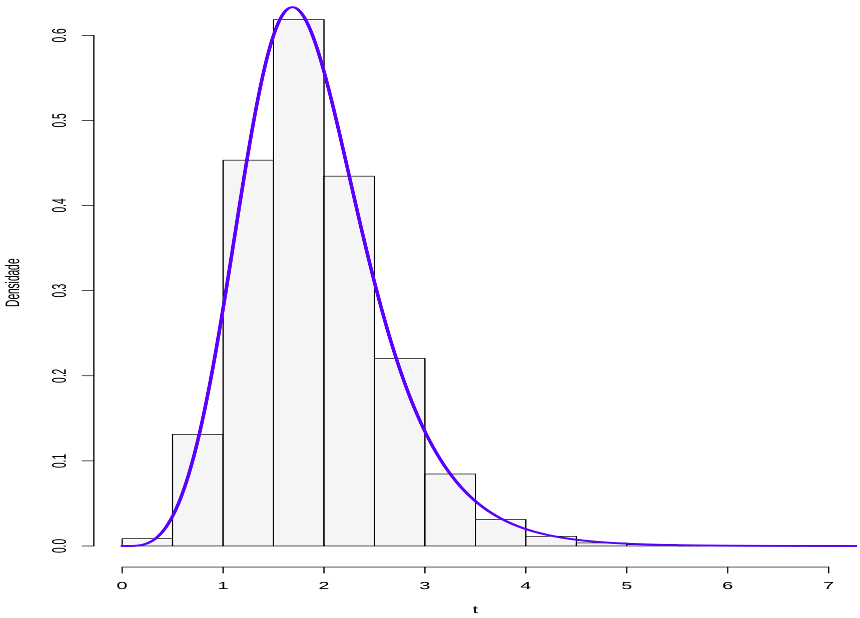


Figura 4.4: Histograma de 10000 números pseudo-aleatórios gerados pela função (4.8) para os valores $\alpha = 3$, $\beta = 5$, $\gamma = \alpha$, $\delta = 1$, $\theta = 2$, $\lambda = 2$ e $\kappa = 1$ com respectiva densidade verdadeira.

Nas duas seções seguintes, afim de se obter resultados analíticos e numéricos para as expressões, os cálculos serão obtidos para $\gamma = \alpha$ e $\delta = \theta = 1$.

4.2 Desvio Médio e Desvio Mediano

Sejam $\gamma = \alpha$ e $\delta = \theta = 1$. As densidade de $T \sim MOGEW(\alpha,\beta,\alpha,1,1,\lambda,\kappa)$ para $\beta \in (0,1)$ pode ser representada por

$$f(t) = \sum_{j,k,l,m=0}^{\infty} w_{j,k,l,m} \cdot g_{\lambda_{l,m},\theta}(t), \tag{4.9}$$

em que $g_{\lambda_{l,m},\theta}(t) = \kappa/\lambda(t/\lambda)^{\kappa-1}e^{-(t/\lambda)^{\kappa}}$ representa a fdp da distribuição de Weibull com parâmetros $\lambda_{l,m} = \frac{\lambda}{(1+l+m)^{1/\kappa}}$ e κ, e

$$w_{j,k,l,m} = \frac{(-1)^{k+l+m}(j+1)}{1+l+m}\binom{j}{k}\binom{\alpha\, k}{l}\binom{\alpha-1}{m}\alpha\beta\bar{\beta}^{j}$$

para $\bar{\beta} = 1-\beta$. Se pelo contrário $\beta > 1$, substitui-se $w_{j,k,l,m}$ por

$$v_{j,k,l,m} = \frac{(-1)^{j+k+l+m}(j+1)}{1+l+m}\binom{j}{k}\binom{\alpha\, k}{l}\binom{\alpha-1}{m}\alpha\sum_{s=0}^{\infty}\binom{s}{j}(1-1/\beta)^{s}.$$

Os desvios médio e mediano

$$\delta_1(T) = \int_0^{\infty}|t-\mu|f(t)\,dt \quad \text{e} \quad \delta_2(T) = \int_0^{\infty}|t-m|f(t)\,dt,$$

são expressos em Cordeiro e Lemonte (2013) por

$$\delta_1(T) = 2\,\mu\, G(\mu) - 2\,J(\mu) \quad \text{e} \quad \delta_2(T) = \mu - 2\,J(m), \tag{4.10}$$

em que $G(q)$ é obtido de (4.1) e $J(q) = \int_0^q t\,f(t)\,dt$. Considerando o caso $\beta \in (0,1)$, pode-se escrever

$$\begin{aligned} J(q) &= \sum_{j,k,l,m}^{\infty} w_{j,k,l,m}\int_0^q t g_{\lambda_{l,m},\kappa}\,d\,t \\ &= \sum_{j,k,l,m}^{\infty} w_{j,k,l,m}\frac{\lambda_{l,m}}{\kappa}\left[\Gamma(1/\kappa) - \kappa\,\Gamma\left(1+1/\kappa,\left(q\lambda_{l,m}^{-1}\right)^{\kappa}\right)\right], \end{aligned} \tag{4.11}$$

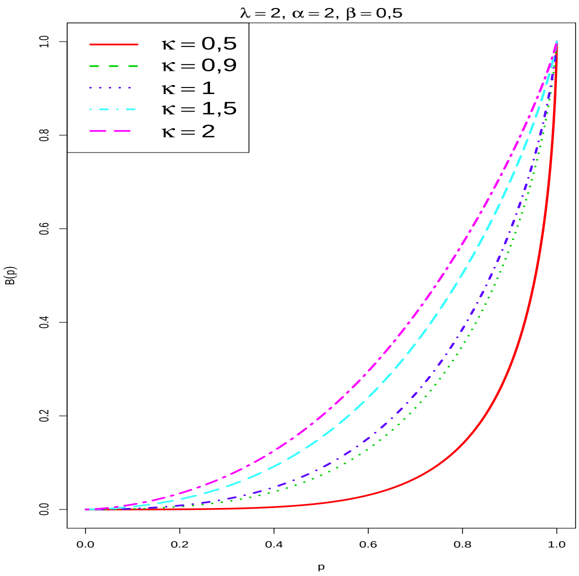
em que $\Gamma(a,b) = \int_b^{\infty} t^{u-1}e^{-t}\,dt$ para α, $\beta > 0$ é a função gama incompleta complementar. Se $\beta > 1$, substituí-se $w_{j,k,l,m}$ por $v_{j,k,l,m}$ em (4.11).

4.3 Curvas de Bonferroni e Lorenz

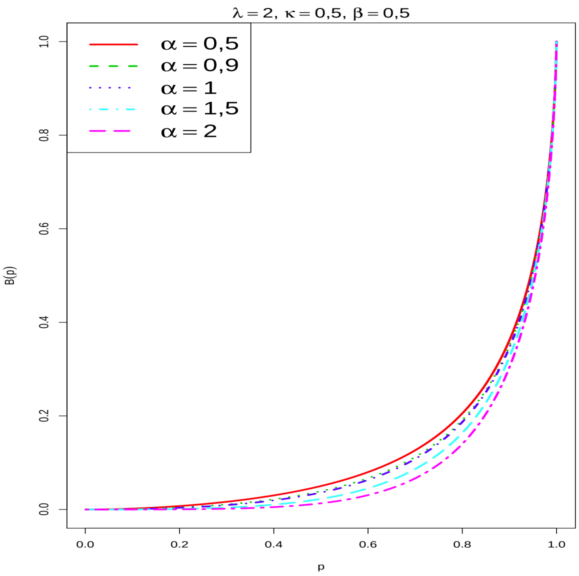
A equação (4.11) é utilizada em economia para a construção das curvas de Bonferroni e Lorenz. Elas são dadas por

$$B(p) = \frac{J(q)}{p\,\mu} \quad \text{e} \quad L(p) = \frac{J(q)}{\mu},$$

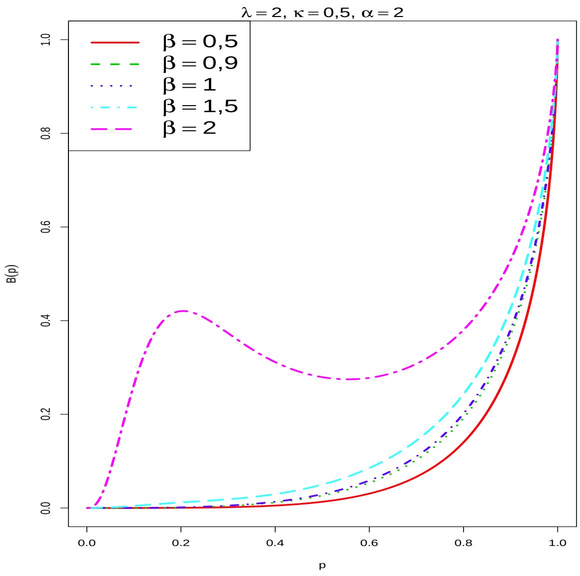
respectivamente, em que $q = Q(p)$ vem da equação (4.8), $J(q)$ é expressa por (4.11) para uma medida de probabilidade p. Seja $T \sim MOGEW(\alpha,\beta,\alpha,1,1,\lambda,\kappa)$ a distribuição de renda numa determinada população, a área entre a reta $L(p) = p$ e a curva de Lorenz pode ser interpretada como uma medida de desigualdade da renda X (NADARAJAH, 2011). Na Figura 4.5 são apresentadas a função da curva de Bonferroni para (a) $\lambda = 2$, $\alpha = 2$ e $\beta = 0,5$ e diferentes valores de κ, (b) $\lambda = 2$, $\kappa = 0,5$ e $\beta = 0,5$ com vários valores de α e (c) $\lambda = 2$, $\kappa = 0,5$ e $\alpha = 2$ variando-se β.



(a)



(b)

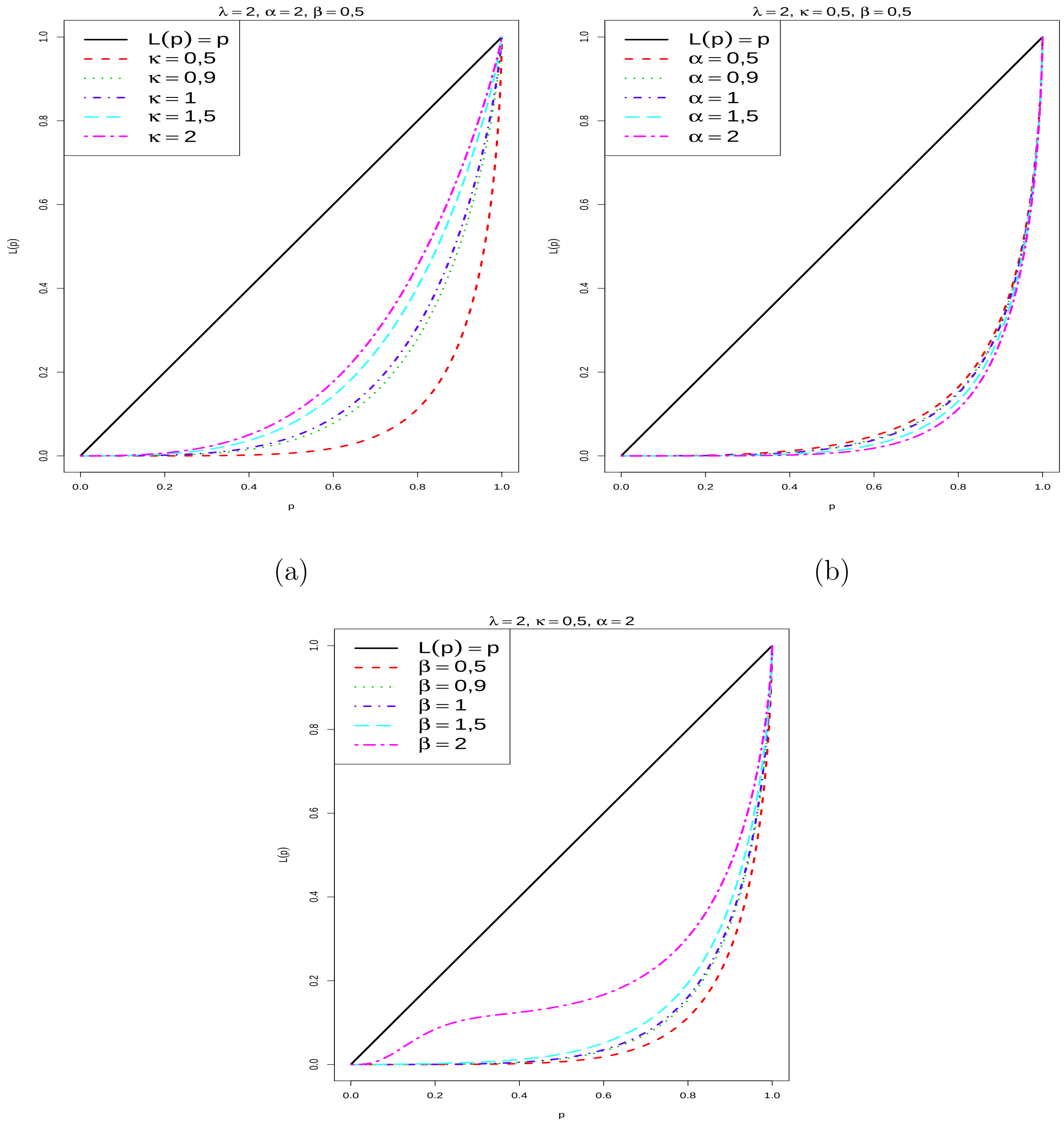


(c)

Figura 4.5: Gráfico $B(p) \times p$ para diversos valores (a) do parâmetro κ; (b) de α; e (c) de β.

Na Figura 4.6 se encontra o gráfico da função $L(p) \times p$ para (a) diferentes valores de κ, (b) de α e (c) de β, respectivamente.

Pode-se visualizar que a curva de Lorenz é menos sensível a mudanças em α que em



(c)

Figura 4.6: Curvas de Lorenz para diversos valores (a) do parâmetro κ; (b) de α; e (c) de β.

β e κ, isto é, modificações na quantidade α não modificam acentuadamente a curva de $L(p)$. Ao contrário, o mesmo não se verifica quando se modifica β ou κ. Analisando-se sob o enfoque da aplicação, suponha que a distribuição de renda de uma determinada população seja modelada por uma variável aleatória $T \sim MOGEW(\alpha, \beta, \alpha, 1, 1, \lambda, \kappa)$. Pelos gráficos é possível se notar que quanto maior κ ou maior β, menor a desigualdade de renda na respectiva população (i.e., menor a área entre a reta e a curva correspondente). O contrário ocorre entre o parâmetro α e a desigualdade de renda.

CAPÍTULO 5

Resultados e Discussões

Neste capítulo são apresentados resultados de simulação e diversas aplicações a dados reais utilizando a distribuição Marshall-Olkin Generalizada Exponenciada de Weibull. Outras distribuições são também ajustadas com o propósito de se verificar qual dentre as distribuições testadas é a mais adequada para cada situação.

5.1 Simulação

Seja $\boldsymbol{\xi} = (\alpha, \beta, \gamma, \delta, \theta, \lambda, \kappa)'$ o vetor paramétrico da distribuição $MOGEW$. Sem perda de generalidade os parâmetros foram fixados em $\alpha = 1,5,\ \beta = 1,2,\ \gamma = 1,5,\ \delta = 1,0,$ $\theta = 0,3,\ \lambda = 1,7$ e $\kappa = 2,5$. Para efeito de comparação entre métodos de estimação foi realizada uma simulação Monte Carlo de tamanho $m = 1000$ para amostras de tamanhos $n = 10, 20, 50$ e 100 para as estimativas de máxima verossimilhança $\hat{\boldsymbol{\xi}}$ (EMV) via algoritmo de estimação $BFGS$, bootstrap $\boldsymbol{\xi}^{*}$ (B) e boostrap corrigido $\tilde{\boldsymbol{\xi}} = 2\boldsymbol{\xi}^{*} - \hat{\boldsymbol{\xi}}^{*}_{(\cdot)}$ (BS). Para as estimativas bootstrap se utilizou uma reamostragem de $B = 500$.

Na Tabela 5.1 se encontram as estimativas dos vieses relativos , via Monte Carlo, para os métodos de máxima verossimilhança, bootstrap e bootstrap corrigido. Nota-se que para todos os tamanhos de amostra, as estimativas de máxima verossimilhança e bootstrap são semelhantes, no sentido de que o viés é aproximadamente o mesmo para cada parâmetro. Diferentemente, as estimativas corrigidas têm vieses sistematicamente superiores. Nota-se ainda que para tamanhos de amostra maiores os vieses tendem a diminuir. Pode-se

Tabela 5.1: Estimativas dos vieses relativos, considerando os métodos de máxima verossimilhança, bootstrap e bootstrap corrigido ($\alpha = 1,5; \beta = 1,2; \gamma = 1,5; \delta = 1,0; \theta = 0,3; \lambda = 1,7; \kappa = 2,5$).

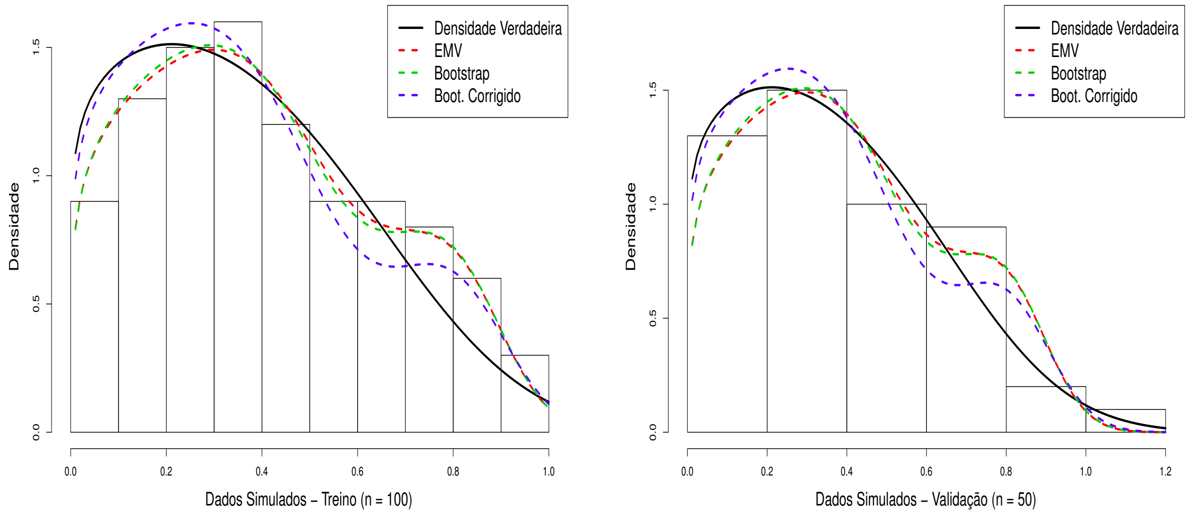
n	Método	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\kappa}$
10	EMV	-0,2789	0,7423	5,9129	2,1581	9,9686	0,7862	1,5201
	B	-0,2568	0,9440	6,0370	2,5355	10,3314	0,7375	1,5322
	BC	-0,5135	1,8880	12,0739	5,0711	20,6629	1,4751	3,0645
20	EMV	-0,2148	0,7328	4,0218	0,8944	6,2399	0,7173	1,0785
	B	-0,2335	0,6059	4,3624	1,0733	7,1114	0,7006	1,1013
	BC	-0,4669	1,2118	8,7248	2,1466	14,2228	1,4011	2,2027
50	EMV	-0,0197	0,7191	1,7603	0,1649	1,9708	0,3140	0,7974
	B	-0,0490	0,6617	1,6676	0,1722	1,9894	0,3090	0,8440
	BC	-0,0980	1,3234	3,3351	0,3444	3,9789	0,6180	1,6879
100	EMV	0,0554	0,5337	0,8589	0,0079	1,1868	0,1850	0,5381
	B	0,0816	0,5512	0,8516	-0,0380	1,2401	0,1795	0,5460
	BC	0,1632	1,1024	1,7032	-0,0761	2,4801	0,3591	1,0920

concluir dessa tabela que o método de máxima verossiminhança e boostrap produzem estimativas semelhantes.

Para verificar se um dos métodos é melhor, sob algum critério que outro, se testou a adequação da distribuição proposta com 150 valores simulados da distribuição $MOGEW$ com $\alpha = 1,5$, $\beta = 1,2$, $\gamma = 1,5$, $\delta = 1,0$, $\theta = 0,3$, $\lambda = 1,7$ e $\kappa = 2,5$, como antes. Os dados foram então subdivididos em dois conjuntos, um de **treino** ou ajuste com 100 valores e um segundo de **validação** com 50 valores. Foram ajustadas, aos dados de treino, a densidade com os parâmetros verdadeiros $\boldsymbol{\xi}$ acima, sua estimativa de máxima verossimilhança (EMV) $\hat{\boldsymbol{\xi}}$ para os dados de treino. Também se repetiu a simulação de máxima verossimilhança utilizando-se reamostragem bootstrap $\boldsymbol{\xi}^*$ e bootstrap corrigido $\tilde{\boldsymbol{\xi}}$. Os códigos feitos no software estatístico R utilizados neste exemplo se encontram no **APÊNDICE B**. Na Figura 5.1.a se pode observar o histograma e os respectivos ajustes para os dados de treino. Já na Figura 5.1.b estão dispostos o histograma e as respectivas densidades ajustadas na fase de treino, aplicados agora para os dados de validação.

Visualmente, tanto as densidades estimadas pelas estimativas por máxima verossimilhança, via bootstrap e bootstrap corrigido parecem condizer tanto com a densidade dos parâmetros, como com os histogramas de treino e de validação. A densidade da EMV e bootstrap se comportam de forma bastante semelhante aos histogramas, enquanto a densidade do método bootstrap corrigido capta melhor a densidade verdadeira (que em aplicações à dados reais não se conhece).

Na Tabela 5.2 se encontram os valores dos parâmetros verdadeiros, dos ajustes de



(a) (b)

Figura 5.1: Histograma para 100 valores gerados de uma distribuição MOGEW com $\alpha = 1,5$, $\beta = 1,2$, $\gamma = 1,5$, $\delta = 1,0$, $\theta = 0,3$, $\lambda = 1,7$ e $\kappa = 2,5$ utilizados para o treino (a) e 50 valores utilizados para validação (b).

máxima verossimilhança, bootstrap e bootstrap corrigido aplicados no conjunto de dados de treino. Cada valor entre parênteses representa o erro-padrão da estimativa imediatamente acima. Os erros-padrão do método bootstrap foram obtidos por (2.18), enquanto que os erros-padrão de (2.20) com $B = 1000$ e $C = 5$. Observe que, ao nível de significância de 5%, aplicando-se o teste de Wald para $H_0 : \alpha = 0$ contra $H_1 : \alpha \neq 0$, não se descarta H_0. Aplicando-se as mesmas hipóteses para β, γ, δ, θ e λ, as mesmas conclusões são obtidas, o que compromete todo o modelo. Situação semelhante ocorre com as estimativas de β e δ do método bootstrap corrigido. As estimativas do método bootstrap são significativas, ao nível de 5% pelo teste de Wald.

Tabela 5.2: Estimativas dos parâmetros para os dados de treino.

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\kappa}$
Valores Reais	1,500	1,200	1,500	1,000	0,300	1,700	2,500
EMV	1,0870	1,4847	2,6082	0,5147	0,2521	1,6668	4,3841
	(0,8182)	(3,7262)	(6,3534)	(1,3417)	(0,2046)	(0,9330)	(0,3152)
Bootstrap	2,333	1,066	4,268	1,532	0,199	1,614	2,593
	(0,132)	(0,448)	(0,455)	(0,223)	(0,043)	(0,082)	(0,231)
Boot. Corrigido	1,109	1,190	2,302	0,520	0,230	1,617	4,588
	(0,267)	(0,969)	(0,865)	(0,542)	(0,076)	(0,174)	(0,419)

Na Tabela 5.3 a seguir são apresentados os critérios AIC, AICC, BIC e HQIC para os diferentes métodos de estimativas empregadas, bem como para os parâmetros verdadeiros.

As medidas da segunda parte da tabela foram obtidas simplesmente substituindo-se os dados de treino pelos de validação, porém continuando com as estimativas da tabela acima. Observa-se que as medidas obtidas pelas estimativas de máxima verossimilhança, bootstrap e boostrap corrigido são similares para os dois conjuntos de dados, e portanto podem ser utilizadas sem distinção. Em comparação com o ajuste dos valores verdadeiros, pelos critérios de AIC e BIC é possível se notar que as estimativas têm melhor ajuste aos dados de treino e de validação, enquanto que AICc e HQIC indicam que os valores verdadeiros seriam melhores. É claro que, tal tipo de comparação só é possível porque numa simulação se dispõe dos valores verdadeiros, o que é impraticável para dados reais. Se observa ainda que há uma coerência entre as estimativas, no sentido de que o que ocorreu nos dados de treino se repetiu nos dados de validação.

Tabela 5.3: Critérios de informação das estimativas de máxima verossimilhança, bootstrap e bootstrap corrigido para os dados simulados.

	Treino				Validação			
	AIC	AICc	BIC	HQIC	AIC	AICc	BIC	HQIC
Parâmetros	3,620	8,867	4,838	10,084	21,857	27,103	11,001	16,247
EMV	-1,098	14,694	0,119	15,911	17,138	32,930	6,282	22,074
Bootstrap	-1,096	14,713	0,122	15,931	17,140	32,950	6,285	22,094
Boot Corrigido	-0,956	13,624	0,261	14,841	17,280	31,860	6,424	21,004

Na Tabela 5.4 estão as estatísticas dos testes de aderência de Anderson-Darling, Cramér Von Mises e p-valor associado à estatística de Kolmogorov-Smirnov. Dentre as estimativas, o método bootstrap apresenta menor estatística de Anderson-Darling tanto nos dados de treino, quanto nos dados de validação. No que se refere a estatística de Cramér - Von Mises, a estimativa de máxima verossimilhança é melhor para os dados de treino, porém para os dados de validação, mais uma vez o método bootstrap é melhor. Pelo teste de Kolmogorov-Smirnov, não se rejeita que nenhuma das densidades se ajuste aos dados de treino ou de validação.

Do exposto se observa que o método bootstrap é compatível com as estimativas de máxima verossimilhança e podem ser adotados quando as mesmas gerarem estimativas não significantes.

5.2 Aplicação 2: Dados de Fibras de Vidro

O primeiro conjunto de dados reais estudado foi obtido de Smith e Naylor (1987) e representa a resistência de $1,5$ cm fibras de vidro, mensuradas no Laboratório Nacional de Física, Inglaterra. A unidade de medida é omitida no trabalho original. Na Tabela 5.5 a seguir estão dispostos os 63 resultados do experimento.

Tabela 5.4: Estatísticas de teste para aderência das estimativas de máxima verossimilhança, bootstrap e bootstrap corrigido e dados para os dados simulados.

	Treino		
	Cramér-Von Mises	Anderson-Darling	K-S (p-valor)
Parâmetros	0,0254	0,2403	0,0691 (0,7257)
EMV	0,0212	0,1946	0,042 (0,9946)
Bootstrap	0,0211	0,1911	0,0435 (0,9915)
Boot Corrigido	0,0201	0,1867	0,0485 (0,9726)
	Validação		
	Cramér-Von Mises	Anderson-Darling	K-S (p-valor)
Parâmetros	4,3082	19,8489	0,0878 (0,8037)
EMV	4,2186	19,4282	0,0681 (0,9624)
Bootstrap	4,2431	19,5214	0.0715 (0,9444)
Boot Corrigido	4,2321	19,5055	0,0594 (0.9901)

Tabela 5.5: Dados relativos à resistência de fibras de vidro fornecidos por Smith e Naylor (1987).

0,55	0,93	1,25	1,36	1,49	1,52	1,58	1,61	1,64	1,68	1,73	1,81	2,00	0,74	1,04
1,27	1,39	1,49	1,53	1,59	1,61	1,66	1,68	1,76	1,82	2,01	0,77	1,11	1,28	1,42
1,50	1,54	1,60	1,62	1,66	1,69	1,76	1,84	2,24	0,81	1,13	1,29	1,48	1,50	1,55
1,61	1,62	1,66	1,70	1,77	1,84	0,84	1,24	1,30	1,48	1,51	1,55	1,61	1,63	1,67
1,7	1,78	1,89												

Barreto-Souza *et al.* (2010) obtêm um bom ajuste para este conjunto de dados com a distribuição beta generalizada exponencial (BGE), cuja densidade é

$$f(x) = \frac{\kappa\lambda}{B(a,b)} e^{-\lambda x} \left(1 - e^{-\lambda x}\right)^{a\kappa - 1} \{1 - \left(1 - e^{-\lambda x}\right)^{\kappa}\}^{b-1}.$$

Barreto-Souza *et al.* (2011) utilizam os mesmos dados, desta vez com a distribuição beta Fréchet (BF), cuja densidade é

$$f(x) = \frac{\kappa\lambda^{\kappa}}{B(a,b)} e^{-a\left(\frac{\lambda}{x}\right)^{\kappa}} \{1 - e^{-\left(\frac{\lambda}{x}\right)^{\kappa}}\}^{b-1}.$$

Na Tabela 5.6 se encontram as estimativas de cada parâmetro, com os respectivos erros-padrão, utilizando-se o método BFGS com bootstrap. Traço significa que o modelo não têm o parâmetro correspondente. Pelo teste de Wald, ao nível de significância de 5%, conclui-se que a estimativa $\hat{\beta}$ do modelo beta Fréchet é não significativo. Se rejeita e hipótese de que os demais parâmetros dos três modelos sejam iguais a zero.

Na Figura 5.2 são apresentadas as densidades de probabilidade das distribuições Marshall-Olkin generalizada exponenciada de Weibull, a beta exponencial generalizada

Tabela 5.6: Estimativas dos parâmetros (erros-padrão, entre parênteses) para as distribuições MOGEW, BGE e beta Fréchet para o conjunto de dados relativos à resistência de fibras de vidro.

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\kappa}$
MOGEW	6,6528	6,9192	52,9672	1,5442	0,8656	1,6079	1,7309
	(0,9866)	(1,7478)	(2,3529)	(0,1568)	(0,1353)	(0,0629)	(0,0663)
BGE	0,3489	88,2031	-	-	-	0,9924	26,6632
	(0,0309)	(18,6182)				(0,0484)	(2,8512)
BFréchet	0,3535	61,9248	-	-	-	1,5729	4,7493
	(0,0347)	(44,7839)				(0,0912)	(0,6524)

e a beta Fréchet ajustados ao conjunto de dados de resistência de fibras de vidro. Pode-se notar, pelo gráfico, que as densidades MOGEW, BGE, BF parecem se ajustar bem aos dados, sendo que a MOGEW e a BGE parecem ter melhores ajustes que a beta Fréchet.

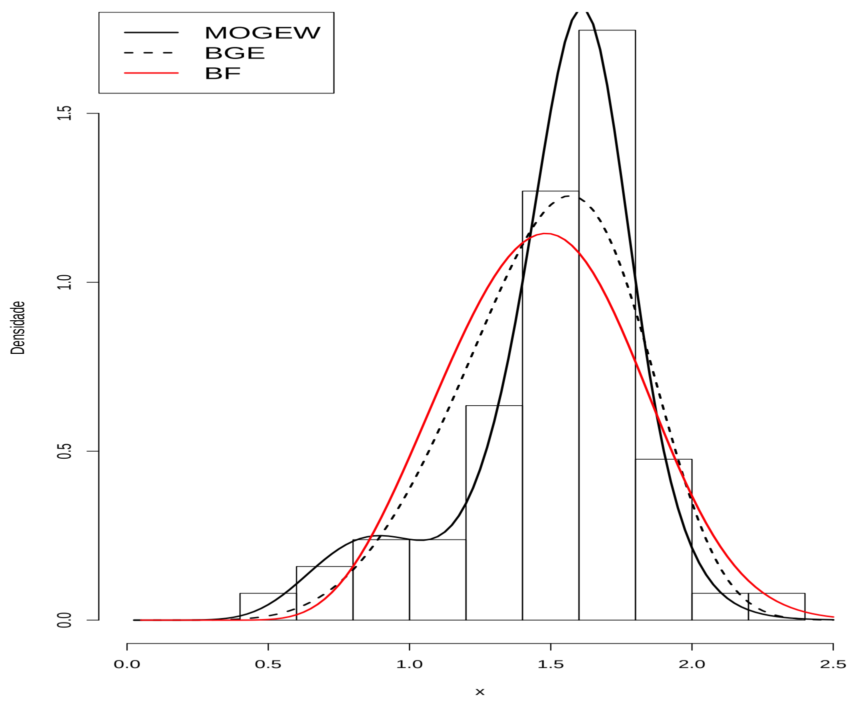


Figura 5.2: Histograma dos dados de fibras de vidro e o confronto do modelo proposto e modelos de artigos que utilizam os mesmos dados.

Para confirmar as conclusões baseadas no gráfico, na Tabela 5.7 são mostradas os valores obtidos pelos critérios AIC, AICc, BIC e HQIC das distribuições de Marshall Olkin generalizada exponencial de Weibull, beta exponencial generalizada e beta Fréchet para o conjunto de dados de fibras de vidro. Para todos os critérios, o modelo MOGEW é indicado como melhor. Note ainda que, usando a regra prática, não há diferença entre os modelos MOGEW e BGE segundo o BIC. No entanto, para os demais a diferença é considerável em favor do modelo proposto.

Tabela 5.7: Critérios de informação AIC, AICc, BIC e HQIC das distribuições MOGEW, BGE e BF para o conjunto de dados relativos à resistência de fibras de vidro.

Modelo	AIC	AICc	BIC	HQIC
MOGEW	32,6009	34,6372	47,6028	38,5012
BGE	39,2439	39,9336	47,8164	42,6155
B-Fréchet	51,8917	52,5814	60,4643	55,2634

Analisando-se a Tabela 5.8 se nota que, para os dados de resistência de fibras de vidro, os menores valores das Estatísticas de Anderson-Darling, Cramér-Von Mises e Kolmogorov-Smirnov são obtidos para o modelo MOGEW. Note ainda que, ao nível de significância de 5%, o único modelo que não se descarta a hipótese de que se ajuste aos dados é o Marshall Olkin generalizado de Weibull.

Tabela 5.8: Estatísticas de teste para aderência das estimativas das distribuições MOGEW, BGE e BF para o conjunto de dados de resistência de fibras de vidro.

	Anderson-Darling	Cramér-Von Mises	K-S (p-valor)
MOGEW	0,2448	0,042	0,0818 (0,7628)
BGE	1,4117	0,2573	0,1771 (0,0336)
BF	2,3485	0,4310	0,2187 (0,0040)

Comparando-se os resultados com os obtidos pelos dois artigos, é possível notar que a complementação da maximização com a técnica bootstrap, não só permitiu calcular o erro-padrão de cada estimativa dos modelos dos artigos, como também melhorou a estimativa de máxima verossimilhança para os mesmos. Ainda assim, pelos testes e critérios calculados há fortes indícios em favor da distribuição de Marshall Olkin generalizada exponenciada de Weibull. O código dessa aplicação se encontra no **APÊNDICE C**.

5.3 Aplicação 3: Dados de Terremotos em Fiji

Um sismo, popularmente conhecido como terremoto, é uma vibração abrupta e momentânea na superfície da Terra, resultante de movimentações das placas tectônicas, atividades vulcânicas ou de migrações gasosas no interior planetário, liberando bruscamente grande quantidade de energia em forma de ondas de choque, conhecidas como ondas sísmicas, que se propagam a partir de um epicentro. Dados de terremotos têm as características de que, para cada grande tremor, seguem-se centenas ou milhares de tremores de menores intensidades associados ao mesmo. Além disso, grandes tremores são cada vez mais raros conforme se aumenta a intensidade, por isso são frequentemente modelados por distribuições de valores extremos como a distribuição de Gumbel ou Fréchet.

A escala de magnitude de momentos (MMS) é usada pelos sismólogos para medir a magnitude dos terremotos em termos de energia liberada e substitui a escala Richter (HANKS; KANAMORI, 1979). Conforme sua predecessora, ela também se apresenta na escala logarítmica e é usada para estimar as magnitudes de todos os terremotos modernos (United States Geological Survey, 2013). É definida pela seguinte fórmula

$$M_w = \frac{2}{3}\log_{10} M_o - 10,7,$$

em que M_o é o momento sísmico medido em dina·centímetros ($10^{-7} N \cdot m$) (HANKS; KANAMORI, 1979).

Os dados são referentes a 1.000 abalos sísmicos mensurados na escala de momento com magnitudes variando entre $4,0$ e $6,4$. Os dados são utilizados por (WASSERMAN, 2004) e podem ser encontradas no seu website. Na Tabela 5.9, a seguir são apresentados, para o conjunto de dados das magnitudes de terremotos no arquipélago de Fiji, as seguintes estatísticas descritivas: média, desvio padrão, coeficiente de variação, mínimo, mediana e máximo amostral.

Tabela 5.9: Estatísticas descritivas para os dados de magnitude de terremotos no arquipélago de Fiji.

$\bar{x}$	s	$CV\%$	Min	Med	Max
$4,620$	$0,4028$	$8,717$	$4,000$	$4,600$	$6,400$

A região do Pacífico onde fica o arquipélago de Fiji na Oceania é conhecida como anel (ou círculo) de fogo (Figura 5.3), por concentrar 90% dos grandes terremotos (United States Geological Survey, 2013).

Na Figura 5.4 se ajustou o modelo proposto, seus submodelos e distribuições concorrentes ao conjunto de dados descritos acima utilizando-se o algoritmo BFGS com método de reamostragem bootstrap, descrito anteriormente e implementado no software R cujo código encontra-se no **APÊNDICE D**. Foram ajustadas as densidade MOGEW (o modelo proposto) com parametrização completa, a densidade MOGEW quando $\alpha = \gamma = \delta = \theta = 1$ (o submodelo Marshall-Olkin Weibull - MOW), a densidade MOWE quando $\alpha = \beta = \gamma = \delta = 1$ (o submodelo de Weibull Exponenciado - EW), além das densidades gama, de Weibull, log-normal, Birnbaum-Saunders e log-logística (frequentemente utilizadas para dados positivos) ao conjunto de dados já citado e a distribuição de Fréchet utilizada em dados de extremos como terremotos e inundações. É possível notar que, exceto as distribuições gama e de Weibull, todas as distribuições parecem se ajustar

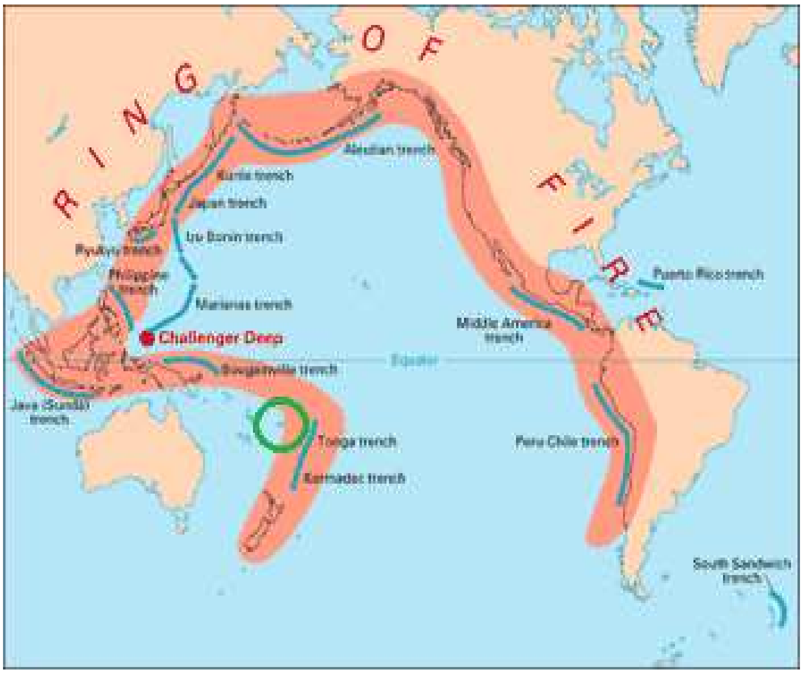


Figura 5.3: Anel de fogo do pacífico com o arquipélago de Fiji em destaque (círculo verde). Fonte: adaptada de United States Geological Survey (2013)

bem aos dados.

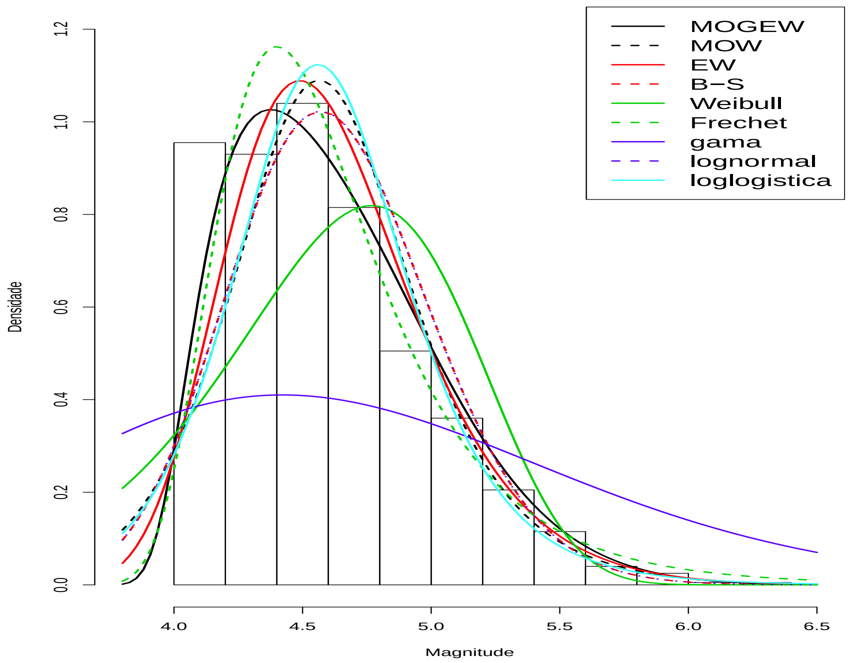


Figura 5.4: Histograma dos dados e o confronto do modelo proposto com seus submodelos e outras densidade positivas.

Na Tabela 5.10 são apresentadas as estimativas para cada parâmetro de cada distribuição ajustada aos dados de magnitudes de terremotos próximos ao arquipélago de Fiji. O valor 1 na estimativa do parâmetro significa que aquela distribuição é um caso particular da distribuição Marshall-Olkin generalizada exponenciada de Weibull quando o dado parâmetro vale um. Já um traço, como antes, significa que aquela distribuição não tem o parâmetro em questão e portanto, não é um sub-modelo da MOGEW. Os valores dos correspondentes erros-padrão (calculados com o método de reamostragem bootstrap) estão, como antes, abaixo das estimativas, entre parênteses.

Tabela 5.10: Estimativas dos parâmetros (erros-padrão, entre parênteses) para as distribuições MOGEW, MOW, EW, Birnbaum-Saunders, gama, Weibull, Fréchet, log-normal e log-logística para o conjunto de dados de magnitudes de terremotos próximos a Fiji.

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\kappa}$
MOGEW	37,6752	1,5308	0,0555	0,4431	29,9137	0,3580	3,8650
	(12,1149)	(0,3920)	(0,0188)	(0,1733)	(4,7048)	(0,0372)	(0,1409)
MOW	1	0,0067	1	1	1	0,1700	20,1532
		(0,0007)				(0,0010)	(0,1017)
EW	1	1	1	1	131,7438	0,4072	2,6621
					(32,1227)	(0,0156)	(0,0884)
B-S	0,08513	4,6041	-	-	-	-	-
	$(6 \cdot 10^{-5})$	(0,0004)					
Weibull	1	1	1	1	1	0,2078	10,6728
						$(2 \cdot 10^{-5})$	(0,0086)
Frétchet	-	-	-	-	-	4,4213	13,9360
						(0,0003)	(0,0090)
gama	-	-	-	-	-	21,6814	4,6978
						(6.3800)	(1.3813)
log-normal	-	-	-	-	-	1,5268	0,08508
						$(8 \cdot 10^{-5})$	$(5 \cdot 10^{-5})$
log-logística	-	-	-	-	-	4,5823	20,5126
						(0,0004)	(0,01572)

Pode-se notar que todas as estimativas do modelo MOGEW são diferentes de zero, ao nível de 5% de confiança pelo teste de Wald. O mesmo ocorre com os demais sobmodelos: Marshall-Olkin Weibull, Weibull Exponenciada e Weibull; e distribuições concorrentes: Birnbaum-Saunders, Fréchet, gama, log-normal e log-logística.

Os critérios AIC, AICc, BIC e HQIC são apresentados, para todas distribuições ajustadas, na Tabela 5.11. Pode-se notar que os critérios AIC, AICc e HQIC da distribuição Marshall-Olkin Generalizada Exponenciada de Weibull são os menores, enquanto a distribuição de Weibull tem o menor BIC. Note ainda que, utilizando-se a regra prática, $\Delta_{AIC} > 2$, não há evidência favorável para os modelos concorrentes, portanto o modelo melhor ajustado, segundo o AIC é o MOGEW. Note também que, a diferença entre os BIC's dos modelos MOGEW e de Weibull é menor que dois, portanto não há evidência que suporte um ou o outro, de forma que o BIC é inconclusivo. Assim, três critérios apontam que o melhor modelo é o proposto.

Tabela 5.11: Critérios de informação AIC, AICc, BIC e HQIC das distribuições MOGEW, MOW, EW, Birnbaum-Saunders, gama, Weibull, Fréchet, log-normal e log-logística para o conjunto de dados de magnitudes de terremotos próximos a Fiji.

	AIC	AICc	BIC	HQIC
MOGEW	889,3104	889,4233	923,6647	902,3674
MOW	990,7853	990,8094	1.005,5086	996,3812
EW	917,7937	917,8178	932,5170	923,3896
B-S	1.270,4781	1.270,4901	1.280,2936	1.274,2087
Weibull	912,6033	912,6153	922,4188	916,3339
Frétchet	1.985,9404	1.985,9525	1.995,7559	1.989,6710
gama	966,1242	966,1363	975,9397	969,8548
log-normal	966,1264	966,1385	975,9420	969,8570
log-logística	987,4369	987,4490	997,2524	991,1675

Na sequência, foi obtida a Tabela 5.12 na qual se encontram os testes de aderência das distribuições ajustadas aos dados de magnitude de terremotos em Fiji. As menores estatísticas de Anderson-Darling, Cramér-Von Mises e Kolmogorov-Smirnov são obtidas para o modelo MOGEW. Note que todos os p-valores da estatística de Kolmogorov-Smirnov indicam que se deveria rejeitar a hipótese de que os dados seguem as distribuições ajustadas. Isso contrasta com a observação gráfica, onde não há motivo para tal rejeição em pelo menos sete das nove distribuições ajustadas. Portanto, se tem uma indicação de que os p-valores estão sendo influenciados pelo tamanho da amostra $n = 1000$ (dependência do tamanho da amostra), devendo se desconsiderar o p-valor nesta aplicação. Assim, se escolhe o modelo MOGEW como o melhor ajustado.

Tabela 5.12: Estatísticas de teste para aderência das estimativas das distribuições MOGEW, MOW, EW, Birnbaum-Saunders, gama, Weibull, Fréchet, log-normal e log-logística para o conjunto de dados de magnitudes de terremotos próximos a Fiji.

	Cramér-Von Mises	Anderson-Darling	K-S (p-valor)
MOGEW	0,6082	3,7178	0,0620 (0,0009)
MOW	1,2732	8,1128	0,0729 ($4 \cdot 10^{-5}$)
EW	0,6941	4,4015	0,0626 (0,0008)
B-S	1,0082	6,3476	0,8085 ($< 2 \cdot 10^{-16}$)
Weibull	4,4235	6,3636	0,2258 ($< 2 \cdot 10^{-16}$)
Frétchet	0,8793	5,601	0,0784 ($9 \cdot 10^{-6}$)
gama	1,4017	8,7441	1,0000 ($< 2 \cdot 10^{-16}$)
log-normal	334,676	1993,458	0,5205 ($2 \cdot 10^{-16}$)
log-logística	1,2523	8,0751	0,0728 ($5 \cdot 10^{-5}$)

Em seguida, realizou-se o Teste da Razão de Verossimilhanças (TRV) para confrontar o modelo proposto contra seus sub-modelos: Marshall-Olkin Weibull, Exponenciado de

Weibull e Weibull. Os resultados estão na Tabela 5.13. Ao nível de confiança de 95% para os valores críticos da distribuição qui-quadrado a quatro e a cinco graus de liberdade, observa-se que os parâmetros da distribuição MOGEW são significativos para os dados da aplicação, portanto, ao nível especificado, se rejeita as hipóteses nulas (de que os dados seguem um dos submodelos). Concluindo-se em favor da distribuição de Marshall-Olkin generalizada exponenciada de Weibull.

Tabela 5.13: Teste da Razão de Verossimilhanças para o modelo proposto e os seus submodelos para o conjunto de dados de magnitude de terremotos em Fiji.

	log-verossimilhança	$2\ell(\hat{\theta}) - 2\ell(\theta_0)$	χ^2_{p-q}
MOGEW	-438,1150	–	–
MOW	-492,3927	108,5554	$\chi^2_{4;0,95} = 9,4877$
EW	-455,8968	35,5636	$\chi^2_{4;0,95} = 9,4877$
Weibull	-454,3016	32,3732	$\chi^2_{5;0,95} = 11,0705$

Pelos resultados dos critérios de adequação e teses de hipóteses se conclui que a distribuição Marshall-Olkin generalizada exponenciada de Weibull é o modelo que melhor representa os dados de magnitude de terremotos no arquipélago de Fiji.

5.4 Aplicação 4: Ensaio Clínico de AIDS (Análise de Sobrevivência)

Nesta seção se estuda dados de um ensaio clínico fornecido pelo Grupo de Ensaios Clínicos de AIDS (ACTG 320). Os dados vêm de um experimento duplo-cego, com controle placebo e comparado com três tipos de drogas administrados em pacientes infectados com HIV (HAMMER *et al.*, 1997). Os pacientes foram considerados aptos para o estudo se eles não tivessem mais que 200 células CD4 por milímetro cúbico e ao menos três semanas de uso anterior de terapia com Zidovudina. A randomização foi estratificada por contagem de células CD4 no instante da admissão no ensaio. O interesse era o tempo até a constatação de AIDS ou a morte. Os dados são estudados por (HOSMER-JR.; LEMESHOW, 1999) e podem ser encontrados em (University of Massachusetts Amherst, 2014) para $n = 576$ pacientes.

Os dados foram analisados segundo os modelos MOGEW, MOW, EW e Weibull. Considerou-se:

t_i : tempo, em dias, até o i-ésimo paciente ser considerado aidético ou falecer;

δ_i : indicador de censura da i-ésima observação, $i = 1, ..., 576$;

x_i : placebo ($x_i = 0$) ou tratamento ($x_i = 1$);

e a equação:

$$y_i = \log t_i = \beta_0 + \beta_1 x_i + \sigma z_i, \quad i = 1, ..., 576.$$

sendo erros z_i distribuídos segundo a distribuição Marshall-Olkin generalizada exponenciada do extremo valor, cuja densidade é 3.2 com

$$g(z) = exp\{z - exp(z)\} \quad \text{e} \quad G(z) = 1 - exp\{-exp(z)\}.$$

Na Figura 5.5 se encontra o gráfico TTT, construído para indicar o modelo apropriado. A curva produzida no gráfico apresenta formato côncavo (B), indicando que a função de risco dos dados é crescente, isto é, com o passar do tempo os indivíduos ficam cada vez mais propensos à falha (morte ou diagnóstico).

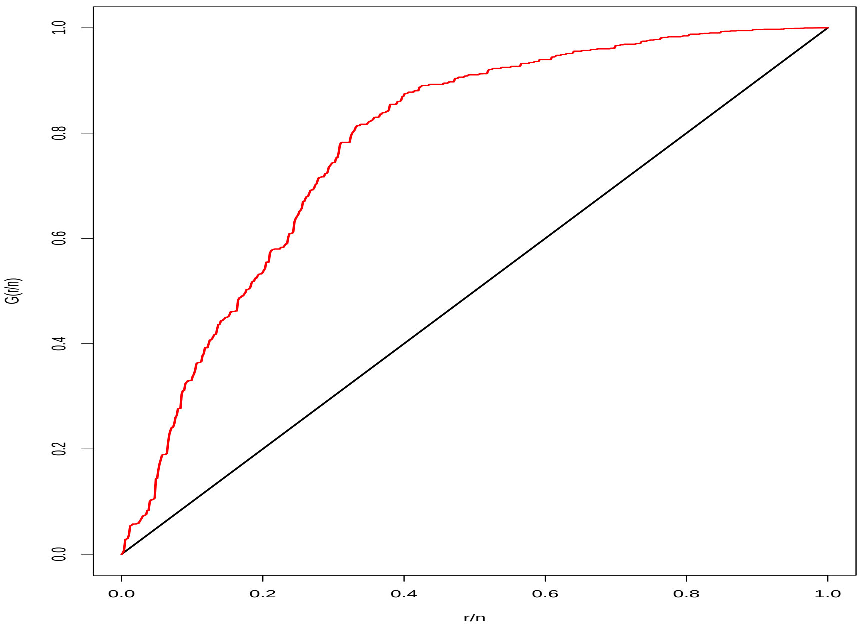


Figura 5.5: Gráfico TTT para os dados do ensaio clínico sobre AIDS (ACTG 320).

As estimativas dos parâmetros e respectivos erros-padrão se encontram na Tabela 5.14 para os modelos MOGEW, MOW, EW e Weibull ajustados os dados do estudo ACTG 320 com o método de maximização BFGS e reamostragem bootstrap com $B = 1000$. Para esta aplicação se fez $\beta = \theta = 1$, pois este submodelo apresentou maior log-verossimilhança que o modelo completo. Não obstante, este submodelo continua sendo uma distribuição ainda não estudada. Assim, sem perda de generalidade o mesmo será designado como MOGEW.

Pelo teste de Wald, ao nível de 5%, o parâmetro β_1 é não significativo. Assim, segundo o modelos Weibull exponenciado, não há influência do tratamento sobre a expectativa de vida dos pacientes. O restante dos modelos teve todos os parâmetros significativos, e em particular, apontam que há diferenças entre tratamento e placebo. Hosmer-Jr. e Lemeshow (1999) também encontram significância para a variável tratamento utilizando um modelo de riscos proporcionais.

Tabela 5.14: Estimativas dos parâmetros (erros-padrão, entre parênteses) para as distribuições MOGEW, MOW, EW e Weibull, para o conjunto de dados de AIDS.

Modelo	b_0	b_1	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$
MOGEW	2,6906	0,0224	0,2316	0,2841	1	1,8464	30,5475
	(0,0387)	(0,0064)	(0,0368)	(0,0513)		(0,4002)	(5,4185)
MOW	1,7737	0,0645	0,5569	1	4,7875	1	1
	(0,1256)	(0,0277)	(0,0675)		(1,3834)		
EW	2,4243	0,0330	0,1819	0,3142	1	1	1
	(0,0142)	(0,0210)	(0,0135)	(0,0348)			
Weibull	2,1936	0,0790	0,4322	1	1	1	1
	(0,0223)	(0,0235)	(0,0264)				

Pela Tabela 5.15 se observa que todos os critérios de informação são mínimos para o modelo de regressão MOGEW. Se percebe ainda que a diferença entre o modelo proposto e seus submodelos, para todos os critérios, supera duzentos pontos, indicando extrema evidência em favor do mesmo.

Tabela 5.15: Critérios de informação AIC, AICc, BIC e HQIC das distribuições MOGEW, MOW, EW e Weibull para o conjunto de dados de AIDS.

Modelo	AIC	AICc	BIC	HQIC
MOGEW	727,3586	727,5062	726,4551	753,4953
MOW	958,9316	959,0016	958,3293	976,356
EW	928,4101	928,4801	927,8078	945,8345
Weibull	1057,131	1057,173	1056,679	1070,199

O teste da razão de verossimilhanças é feito na Tabela 5.16 para o modelo proposto e seus sub-casos. Se observa que a hipótese de nulidade é rejeitada para cada uma das distribuições concorrentes, ao nível de significância de 5%. Conclui-se que, dentre os modelos testados, aquele com melhor ajuste aos dados clínicos de paciente com HIV é o modelo MOGEW.

Assim, a função de sobrevivência, condicionada o valor de x_i, do modelo escolhido, na escala logarítmica do tempo, é

$$\hat{S}(y|x) = 1 - \frac{\left\{1 - \exp\left[-\exp\left(\frac{y-\hat{\mu}_i}{0{,}2316}\right)\right]\right\}^{0{,}2841}}{\left\{1 - \exp\left[-\exp\left(\frac{y-\hat{\mu}_i}{0{,}2316}\right)\right]\right\}^{0{,}2841} + \left[1 - \left\{1 - \exp\left[-\exp\left(\frac{y-\hat{\mu}_i}{0{,}2316}\right)\right]\right\}^{1{,}8464}\right]^{30{,}5475}}.$$

Tabela 5.16: Teste da Razão de Verossimilhanças para o modelo proposto e os seus submodelos para o conjunto de dados do ensaio clínico de AIDS (ACTG 320).

	log-verossimilhança	$2\ell(\hat{\theta}) - 2\ell(\theta_0)$	χ^2_{p-q}
MOGEW	-360,0695	–	–
MOW	-478,0468	235,9546	$\chi^2_{2;0,95} = 5,9915$
EW	-461,9694	203,7998	$\chi^2_{2;0,95} = 5,9915$
Weibull	-525,8555	331,5721	$\chi^2_{3;0,95} = 7,8147$

Na escala linear do tempo, a função de sobrevivência, condicionada à presença ($x_i = 1$) ou ausência de tratamento ($x_i = 0$) é

$$S(t|x_i) = 1 - \frac{\left\{1 - \exp\left[-\left(\frac{t}{\exp(\hat{\mu})}\right)^{4,3178}\right]\right\}^{0,2841}}{\left\{1 - \exp\left[-\left(\frac{t}{\exp(\hat{\mu})}\right)^{4,3178}\right]\right\}^{0,2841} + \left[1 - \left\{1 - \exp\left[-\left(\frac{t}{\exp(\hat{\mu})}\right)^{4,3178}\right]\right\}^{1,8464}\right]^{30,5475}},$$

em que $\hat{\mu}_i = 2,6906 + 0,0224x_i$.

A função de sobrevivência empírica de Kaplan-Meier e as funções de sobrevivência ajustadas das distribuições Marshall-Olkin generalizada exponenciada de Weibull, Marshall-Olkin de Weibull, exponenciada de Weibull e de Weibull se encontram na Figura 5.6.

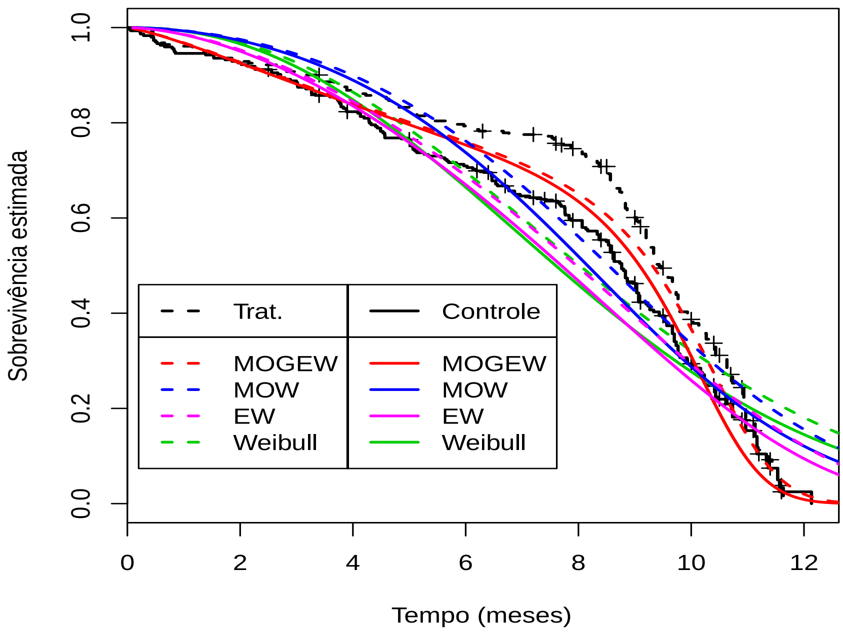


Figura 5.6: Gráfico da função de sobrevivência empírica de Kaplan-Meier e funções de sobrevivência ajustadas para os dados do ensaio clínico sobre AIDS (ACTG 320).

Tomando a função de sobrevivência empírica como base, se observa que há uma tendência dos modelos MOW, EW e Weibull de sobrestimar os dados entre os meses zero e quatro, de subestimar entre os meses quatro e dez e de sobrestimar (outra vez) após isso. Em contrapartida, o modelo proposto segue o comportamento da curva de Kaplan-Meier.

Uma ressalva, no entanto deve ser feita, pois o modelo parece não discriminar tão bem as diferenças entre tratamento e controle no intervalo de quatro a nove meses - nesse intervalo as curvas preditas estão próximas, enquanto as curvas de Kaplan-Meier estão afastadas. Porém, o mesmo aconteceu com as demais funções de sobrevivência ajustadas que não foram capazes de modelar, de forma razoável, o comportamento dos dados no período citado.

Conforme indicado pelo gráfico TTT, a função de risco, representada na Figura 5.7 do modelo proposto é crescente. Assim, segundo o modelo o risco dos pacientes aumentam a medida que o tempo passa. Esse risco de falha é cinco vezes maior para um paciente que está à dez meses no estudo de que para aquele que foi engajado à dois meses. Em comparação, o risco de falha para pacientes que estão no tratamento à um ano é 30 vezes maior que o de novatos.

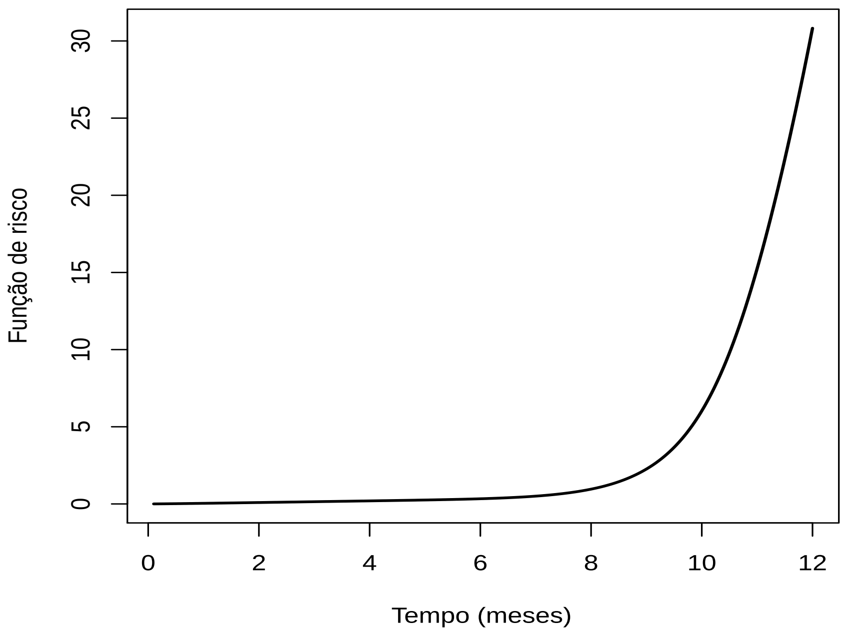


Figura 5.7: Gráfico da função de função de risco do modelo proposto para os dados do ensaio clínico sobre AIDS (ACTG 320).

Do exposto, verifica-se que o melhor ajuste aos dados do estudo clínico sobre AIDS ocorre para o modelo de regressão para dados censurados de Marshall-Olkin generalizado exponenciado de Weibull, cujos erros provêm de uma distribuição de Marshall-Olkin generalizada exponenciada do extremo valor. No **APÊNDICE E** se encontram os códigos dessa seção.

5.5 Aplicação 5: Dados de Fibras de Carbono (Inferência Bayesiana)

Nichols e Padgett (2006) fornecem um conjunto de dados de 100 observações de fibras de carbono submetidos a um estresse até quebrarem. Os dados são utilizados como exemplo para ajustar a distribuição Marshall-Olkin exponencial (LUNN *et al.*, 2009). Os dados são mostrados na Tabela 5.17 a seguir

Tabela 5.17: Dados de quebras de fibras de carbono (NICHOLS; PADGETT, 2006)

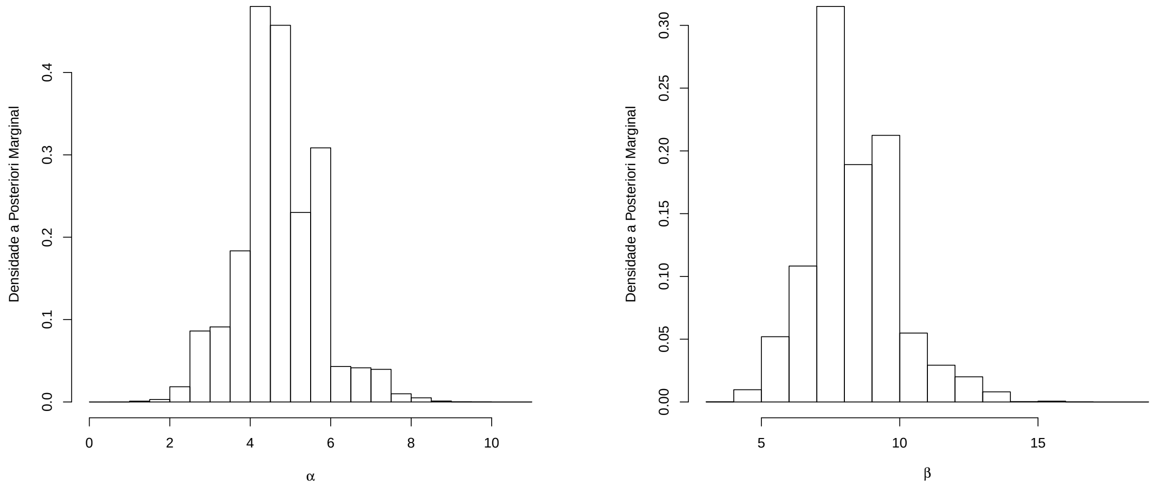
3,70	2,74	2,73	2,50	3,60	3,11	3,27	2,87	1,47	3,11
4,42	2,41	3,19	3,22	1,69	3,28	3,09	1,87	3,15	4,90
3,75	2,43	2,95	2,97	3,39	2,96	2,53	2,67	2,93	3,22
3,39	2,81	4,20	3,33	2,55	3,31	3,31	2,85	2,56	3,56
3,15	2,35	2,55	2,59	2,38	2,81	2,77	2,17	2,83	1,92
1,41	3,68	2,97	1,36	0,98	2,76	4,91	3,68	1,84	1,59
3,19	1,57	0,81	5,56	1,73	1,59	2,00	1,22	1,12	1,71
2,17	1,17	5,08	2,48	1,18	3,51	2,17	1,69	1,25	4,38
1,84	0,39	3,68	2,48	0,85	1,61	2,79	4,70	2,03	1,80
1,57	1,08	2,03	1,61	2,12	1,89	2,88	2,82	2,05	3,65

Os dados foram utilizados para ajustar um modelo de inferência bayesiana utilizando o processo de amostragem ponderada para se obter as estimativas da densidade a posteriori da distribuição de Marshall-Olkin generalizada exponenciada expoencial (MOGEE):

$$\left.\begin{array}{rcl} x_i & \sim & \text{dmogeweib}(\alpha, \beta, 1, 1, 1, 1, 1) \end{array}\right\} 1 \le i \le 100$$
$$\begin{array}{rcl} \alpha & \sim & \text{dgamma}(0{,}001; 0{,}001) \\ \beta & \sim & \text{dgamma}(0{,}001; 0{,}001) \end{array}$$

em que $\alpha \sim \text{dgamma}(0{,}001; 0{,}001)$ e $\beta \sim \text{dgamma}(0{,}001; 0{,}001)$ são as densidades a priori não informativa. Um segundo modelo é ajustado, trocando-se $x_i \sim \text{dmogeweib}(\alpha, \beta, 1, 1, 1, 1, 1)$ por $x_i \sim \text{dmogeweib}(1, \beta, 1, 1, 1, 1, 1)$, que é a distribuição de Marshall-Olkin exponencial com $\lambda = 1$. Foi simulada uma amostra de 500.000 valores para cada parâmetro, dos quais os 10.000 primeiros foram descartados. Foi tomada uma em cada 20 das observações da série resultante para se produzir a amostra final para cada parâmetro. Os resultados a posteriori estão na Tabela 5.18.

Na Figura 5.8 a seguir estão os histogramas das densidades a posteriori marginais para os parâmetros α e β da distribuição MOGEE.



(a) (b)

Figura 5.8: Histograma para 24.500 valores gerados de uma distribuição MOGEW com $\gamma = 1,0$, $\delta = 1,0$, $\theta = 1,0$, $\lambda = 1,0$ e $\kappa = 1,0$ para α (a) e para β (b).

Tabela 5.18: Estimativas a posteriori para os dados de fibras de carbono.

	parâmetro	média	erro-padrão	$Q_{2,5\%}$	mediana	$Q_{97,5\%}$
MOE	α	12,1777	1,9524	8,7841	11,9951	16,4400
$DIC = 112,4742$						
MOGEE	α	4,7337	1,0079	2,7899	4,7484	7,0002
	β	8,2724	1,6060	5,4278	8,1713	12,1301
$DIC = 50,09263$						

O código encontra no **APÊNDICE F**. Pelo DIC, o modelo melhor ajustado é o MOGEE.

De um modo geral, o modelo proposto teve bons ajustes para diversos conjuntos de dados positivos. Entre os complementos a serem estudados posteriormente se pode citar: a forma da nova distribuição através de suas derivadas e limites; encontrar as extensões das distribuições gama, de Fréchet, log-logística e log-normal a partir da nova classe de distribuições; aplicar o modelo proposto a outros conjuntos de dados de Análise de Sobrevivência e obter medidas de resíduos do modelo de regressão; calcular as medidas de entropias de Rènyi e Shannon; e um estudo mais aprofundado em inferência bayesiana.

CAPÍTULO 6

Conclusão

Neste trabalho, foi proposta uma classe de distribuições denominada família Marshall-Olkin Generalizada Exponenciada a partir de duas classes de distribuições conhecidas e bem estudadas em Estatística. Particularmente, foi estudada uma distribuição, denominada distribuição de Marshall-Olkin Generalizada Exponenciada de Weibull.

Em seguida, obteve-se alguns resultados analíticos e discutiu-se os gráficos das funções de distribuição, densidade de probabilidade e função de risco para diversos valores de α, β, γ, δ, θ, λ e κ, mostrando-se que a distribuição é flexível o bastante para modelar dados de Análise de Sobrevivência. Obteve-se algumas das principais quantidades e comparou-se o ajuste de algumas distribuições e sub-modelos da distribuição em simulação e dados reais. Foi obtido também, um modelo de regressão para dados censurados. Por fim, foi estudado um modelo Bayesiano para dados de quebra de fibras de carbono submetidos a estresse. Os resultados foram favoráveis a distribuição proposta.

Conclui-se que a distribuição de Marshall-Olkin Generalizada Exponenciada de Weibull compete com distribuições frequentemente utilizadas na literatura e assim requer um estudo mais minucioso. Este trabalho não é exaustivo e entre os complementos a serem estudados posteriormente, podendo-se citar: encontrar as extensões das distribuições gama, Fréchet, log-logística e log-normal a partir da nova classe de distribuições; aplicar o modelo proposto a outros conjuntos de dados de Análise de Sobrevivência, particularmente com outros tipos de censura; calcular as entropias de Rènyi e Shannon para o modelo completo.

Referências Bibliográficas

AARSET, M. V. How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, v. 36, n. 1, p. 106–108, 1987.

ACHCAR, J. A.; BOLETA, J. Distribuiçao exponencial generalizada: uso de métodos bayesianos. *Rev. Bras. Biom*, v. 27, n. 4, p. 644–658, 2009.

AITCHISON, J.; BROWN, J. A. The lognormal distribution with special reference to its uses in economics. 1957.

AITKIN, M. Posterior Bayes factors. *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 111–142, 1991.

AKAIKE, H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, v. 19, n. 6, p. 716–723, 1974.

ANDERSON, T. W.; DARLING, D. A. Asymptotic theory of certain"goodness of fit"criteria based on stochastic processes. *The Annals of Mathematical Statistics*, p. 193–212, 1952.

ARMENIAN, H. K.; LILIENFELD, A. M. The distribution of incubation periods of neoplastic diseases. *American Journal of Epidemiology*, v. 99, n. 2, p. 92–100, 1974.

ARMSTRONG, D. P.; EWEN, J. G. Dynamics and viability of a new zealand robin population reintroduced to regenerating fragmented habitat. *Conservation Biology*, v. 16, n. 4, p. 1074–1085, 2002.

BABBIE, E. *The practice of social research.* Belmont, California: Cengage Learning, 2012.

BALANDA, K. P.; MACGILLIVRAY, H. Kurtosis: a critical review. *The American Statistician*, v. 42, n. 2, p. 111–119, 1988.

BARRETO-SOUZA, W.; CORDEIRO, G. M.; SIMAS, A. B. Some results for beta fréchet distribution. *Communications in Statistics—Theory and Methods*, v. 40, n. 5, p. 798–811, 2011.

BARRETO-SOUZA, W.; SANTOS, A. H.; CORDEIRO, G. M. The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*, v. 80, n. 2, p. 159–172, 2010.

BAYES, M.; PRICE, M. An essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, frs communicated by mr. price, in a letter to john canton, amfrs. *Philosophical Transactions (1683-1775)*, p. 370–418, 1763.

BENNETT, S. Log-logistic regression models for survival data. *Applied Statistics*, p. 165–171, 1983.

BEST, N.; CARLIN, B.; LINDE, A. Van der *et al.* Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B*, v. 64, n. 4, p. 583–616, 2003.

BILLINGSLEY, P. *Probability and measure*. New York, USA: John Wiley & Sons, 2008.

BIRNBAUM, Z.; SAUNDERS, S. C. A statistical model for life-length of materials. *Journal of the American Statistical Association*, v. 53, n. 281, p. 151–160, 1958.

BIRNBAUM, Z.; SAUNDERS, S. C. A new family of life distributions. *Journal of Applied Probability*, p. 319–327, 1969a.

BIRNBAUM, Z.; SAUNDERS, S. C. Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability*, p. 328–347, 1969b.

BOURGUIGNON, M.; SILVA, R. B.; ZEA, L. M.; CORDEIRO, G. M. The kumaraswamy pareto distribution. *Journal of Statistical Theory and Applications*, v. 12, n. 2, p. 129–144, 2013.

BRESLOW, N.; CROWLEY, J. A large sample study of the life table and product limit estimates under random censorship. *The Annals of Statistics*, v. 2, n. 3, p. 437–453, 1974.

BRITO, C. C. R. de. *Método Gerador de Distribuições e Classes de Distribuições Probabilísticas*. Tese (Doutorado) — Universidade Federal Rural de Pernambuco - UFRPE, 2014.

BROWN, G. W.; FLOOD, M. M. Tumbler mortality. *Journal of the American Statistical Association*, v. 42, n. 240, p. 562–574, 1947.

BROYDEN, C. G. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics*, v. 6, n. 1, p. 76–90, 1970.

BURNHAM, K. P.; ANDERSON, D. R. *Model selection and multi-model inference: a practical information-theoretic approach*. New York, USA: Springer, 2002.

BURNHAM, K. P.; ANDERSON, D. R. Multimodel inference understanding aic and bic in model selection. *Sociological methods & research*, v. 33, n. 2, p. 261–304, 2004.

CASELLA, G.; BERGER, R. L. *Statistical inference*. Crawfordville, Florida: Duxbury Pacific Grove, CA, 2002.

CASSON, R. J. The pesty p value. *Clinical & experimental ophthalmology*, v. 39, n. 9, p. 849–850, 2011.

CERIANI, L.; VERME, P. The origins of the Gini index: extracts from variabilità e mutabilità (1912) by Corrado Gini. *The Journal of Economic Inequality*, v. 10, n. 3, p. 421–443, 2012.

CHEN, G.; BALAKRISHNAN, N. A general purpose approximate goodness-of-fit test. *Journal of Quality Technology*, v. 27, n. 2, p. 154–161, 1995.

CLAESKENS, G.; HJORT, N. L. *Model selection and model averaging*. London, GB: Cambridge University Press Cambridge, 2008.

COLES, S.; BAWA, J.; TRENNER, L.; DORAZIO, P. *An introduction to statistical modeling of extreme values*. Gateshead, GB: Springer, 2001.

COLOSIMO, E. A.; GIOLO, S. R. *Análise de Sobrevivência Aplicada*. São Paulo, SP: Associação Brasileira de Estatística., 2006.

COORAY, K.; ANANDA, M. M. A generalization of the half-normal distribution with applications to lifetime data. *Communications in Statistics—Theory and Methods*, v. 37, n. 9, p. 1323–1337, 2008.

CORDEIRO, G. M.; CASTRO, M. A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, v. 81, n. 7, p. 883–898, 2011.

CORDEIRO, G. M.; LEMONTE, A. J. On the Marshall-Olkin extended Weibull distribution. *Statistical Papers*, v. 54, n. 2, p. 333–353, 2013.

CORDEIRO, G. M.; ORTEGA, E. M.; NADARAJAH, S. The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*, v. 347, n. 8, p. 1399–1429, 2010.

CORDEIRO, G. M.; SIMAS, A. B.; STOSIC, B. D. Closed form expressions for moments of the beta Weibull distribution. *Annals of the Brazilian Academy of Sciences*, v. 83, p. 357 – 373, 2011.

CORNELL, C. A. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, v. 58, n. 5, p. 1583–1606, 1968.

COX, D. R. *Principles of statistical inference*. Cambridge: Cambridge University Press, 2006.

COX, D. R.; OAKES, D. *Analysis of survival data*. [S.l.]: CRC Press, 1984.

CYSNEIROS, F. J. A.; PAULA, G. A.; GALEA, M. Modelos simétricos aplicados. *ABE: São Paulo-IX Escola de Modelos de Regressão, 100p*, 2005.

DARLING, D. A. The kolmogorov-smirnov, cramer-von mises tests. *The Annals of Mathematical Statistics*, p. 823–838, 1957.

DAVISON, A. C. *Bootstrap methods and their application*. USA: Cambridge university press, 1997.

DRAGHICI, S.; KHATRI, P.; TARCA, A. L.; AMIN, K.; DONE, A.; VOICHITA, C.; GEORGESCU, C.; ROMERO, R. A systems biology approach for pathway level analysis. *Genome research*, v. 17, n. 10, p. 1537–1545, 2007.

DURRETT, R. *Probability: theory and examples*. New York, USA: Cambridge university press, 2010.

EFRON, B. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, p. 1–26, 1979.

EFRON, B. Censored data and the bootstrap. *Journal of the American Statistical Association*, USA, v. 76, n. 374, p. 312–319, 1981.

EFRON, B.; TIBSHIRANI, R. J. *An introduction to the bootstrap*. [S.l.]: CRC press, 1994.

EMBRECHTS, P.; KLÜPPELBERG, C.; MIKOSCH, T. *Modelling extremal events: for insurance and finance*. [S.l.]: Springer, 1997.

ENGLE, R. F. Wald, likelihood ratio, and lagrange multiplier tests in econometrics. *Handbook of Econometrics*, v. 2, p. 775–826, 1984.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal distribution and its applications. *Commun Stat-Theor M.*, v. 31, p. 497 – 512, 2002.

FAMOYE, F.; LEE, C.; OLUMOLADE, O. The beta-weibull distribution. *Journal of Statistical Theory and Applications*, v. 4, n. 2, p. 121–136, 2005.

FLETCHER, R. A new approach to variable metric algorithms. *The computer journal*, v. 13, n. 3, p. 317–322, 1970.

FRÉCHET, M. Sur la loi de probabilité de l'écart maximum. *Ann. soc. plon. math*, v. 6, p. 93–116, 1927.

GAMERMAN, D.; LOPES, H. F. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. [S.l.]: CRC Press, 2006.

GARDNER, M. J.; ALTMAN, D. G. Confidence intervals rather than p values: estimation rather than hypothesis testing. *British medical journal (Clinical research ed.)*, v. 292, n. 6522, p. 746, 1986.

GASTWIRTH, J. L. The estimation of the Lorenz curve and Gini index. *The Review of Economics and Statistics*, p. 306–316, 1972.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; DUNSON, D. B.; VEHTARI, A.; RUBIN, D. B. *Bayesian data analysis*. [S.l.]: CRC press, 2013.

GHITANY, M.; AL-AWADHI, F.; ALKHALFAN, L. Marshall-Olkin extended lomax distribution and its application to censored data. *Communications in Statistics—Theory and Methods*, v. 36, n. 10, p. 1855–1866, 2007.

GHITANY, M.; AL-HUSSAINI, E.; AL-JARALLAH, R. Marshall-Olkin extended Weibull distribution and its application to censored data. *Journal of Applied Statistics*, v. 32, n. 10, p. 1025–1034, 2005.

GILCHRIST, W. *Statistical modelling with quantile functions*. USA: CRC Press, 2002.

GODFREY, L. *Bootstrap tests for regression models*. [S.l.]: Palgrave Macmillan, 2009.

GOLDFARB, D. A family of variable-metric methods derived by variational means. *Mathematics of Computation*, v. 24, n. 109, p. 23–26, 1970.

GOODMAN, S. N. Toward evidence-based medical statistics. 1: The p value fallacy. *Annals of internal medicine*, v. 130, n. 12, p. 995–1004, 1999.

GRADSHTEYN, I. S.; RYZHIK, I. M. Table of integrals, series and products (corrected and enlarged edition prepared by a. jeffrey and d. zwillinger). *Academic Press, New York*, 2000.

GREENWOOD, J. A.; LANDWEHR, J. M.; MATALAS, N. C.; WALLIS, J. R. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, v. 15, n. 5, p. 1049–1054, 1979.

GUPTA, R. D.; KUNDU, D. Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, v. 41, n. 2, p. 173–188, 1999.

GUPTA, R. D.; KUNDU, D. Generalized exponential distribution: Bayesian inference. *Comput. Stat. Data Anal.*, v. 52, n. 4, p. 1873–1883, 2008.

HAMMER, S. M.; SQUIRES, K. E.; HUGHES, M. D.; GRIMES, J. M.; DEMETER, L. M.; CURRIER, J. S.; ERON JR, J. J.; FEINBERG, J. E.; BALFOUR JR, H. H.; DEYTON, L. R. *et al.* A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and cd4 cell counts of 200 per cubic millimeter or less. *New England Journal of Medicine*, v. 337, n. 11, p. 725–733, 1997.

HANKS, T. C.; KANAMORI, H. A moment magnitude scale. *Journal of Geophysical Research: Solid Earth (1978–2012)*, v. 84, n. B5, p. 2348–2350, 1979.

HANNAN, E. J.; QUINN, B. G. The determination of the order of an autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 190–195, 1979.

HAUCK, W. W. J.; DONNER, A. Wald's test as applied to hypotheses in logit analysis. *J. Am. Stat. Assoc.*, v. 72, p. 851–853, 1977.

HAZZAH, L.; BORGERHOFF MULDER, M.; FRANK, L. Lions and warriors: social factors underlying declining african lion populations and the effect of incentive-based management in kenya. *Biological Conservation*, v. 142, n. 11, p. 2428–2437, 2009.

HESTENES, M. R.; STIEFEL, E. *Methods of conjugate gradients for solving linear systems*. [S.l.]: National Bureau of Standards Washington, DC, 1952.

HEYDE, C. C. On a property of the lognormal distribution. *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 392–393, 1963.

HOFF, P. D. *A first course in Bayesian statistical methods*. New York, USA: Springer, 2009.

HOSMER-JR., D. W.; LEMESHOW, S. *Applied survival analysis: regression modeling of time to event data*. [S.l.]: John Wiley and Sons, 1999.

HUBBARD, R. The widespread misinterpretation of p-values as error probabilities. *Journal of Applied Statistics*, v. 38, n. 11, p. 2617–2626, 2011.

HUBBARD, R.; LINDSAY, R. M. Why p values are not a useful measure of evidence in statistical significance testing. *Theory & Psychology*, v. 18, n. 1, p. 69–88, 2008.

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. *An introduction to statistical learning*. USA: Springer, 2013.

JONES, M. C. Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, v. 6, p. 70 – 81, 2008.

JURECKOVÁ, J.; SEN, P. K. *Robust statistical procedures: asymptotics and interrelations*. [S.l.]: John Wiley & Sons, 1996.

KAKWANI, N. C. Applications of lorenz curves in economic analysis. *Econometrica: Journal of the Econometric Society*, p. 719–727, 1977.

KALBFLEISCH, J. D.; PRENTICE, R. L. *The statistical analysis of failure time data*. USA: John Wiley & Sons, 2011.

KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, v. 53, n. 282, p. 457–481, 1958.

KASS, R. E.; RAFTERY, A. E. Bayes factors. *Journal of the American Statistical Association*, v. 90, n. 430, p. 773–795, 1995.

KENDALL, M. G. *et al.* The advanced theory of statistics. *The advanced theory of statistics.*, n. 2nd Ed, 1946.

KOLMOGOROV, A. N. *Sulla determinazione empirica di una legge di distribuzione*. [S.l.: s.n.], 1933.

LAPLACE, P. S. *Pierre Simon Laplace Philosophical Essay on Probabilities*. USA: Springer, 1995.

LAWLESS, J. F. *Statistical models and methods for lifetime data*. USA: John Wiley & Sons, 2011.

LEIVA, V.; BARROS, M.; PAULA, G. Generalized birnbaum-saunders models using r. *Brazilian Statistical Association, São Paulo, Brazil (in English)*, 2009.

LEROY, R.; BOGAERTS, K.; LESAFFRE, E.; DECLERCK, D. The emergence of permanent teeth in flemish children. *Community Dentistry and Oral Epidemiology*, v. 31, n. 1, p. 30–39, 2003.

LIMA, M. de; GIRALT, S.; THALL, P. F.; PADUA SILVA, L. de; JONES, R. B.; KOMANDURI, K.; BRAUN, T. M.; NGUYEN, H. Q.; CHAMPLIN, R.;

GARCIA-MANERO, G. Maintenance therapy with low-dose azacitidine after allogeneic hematopoietic stem cell transplantation for recurrent acute myelogenous leukemia or myelodysplastic syndrome. *Cancer*, v. 116, n. 23, p. 5420–5431, 2010.

LONGIN, F. M. The asymptotic distribution of extreme stock market returns. *Journal of Business*, p. 383–408, 1996.

LUNN, D.; SPIEGELHALTER, D.; THOMAS, A.; BEST, N. The bugs project: Evolution, critique and future directions. *Statistics in Medicine*, v. 28, n. 25, p. 3049–3067, 2009.

MAGALHÃES, M. N. *Probabilidade e variáveis aleatórias*. São Paulo, SP: Edusp, 2011.

MANN, N. R.; SINGPURWALLA, N. D.; SCHAFER, R. E. Methods for statistical analysis of reliability and life data. 1974.

MARSHALL, A. W.; OLKIN, I. A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. *Biometrika*, v. 84, n. 3, p. 641–652, 1997.

MAZEROLLE, M. J. Improving data analysis in herpetology: using Akaike's information criterion (AIC) to assess the strength of biological hypotheses. *Amphibia Reptilia*, v. 27, n. 2, p. 169–180, 2006.

MINER, M. A. Cumulative damage in fatigue. *Journal of Applied Mechanics*, v. 12, n. 3, p. 159–164, 1945.

MØLLER, A.; MOUSSEAU, T.; LYNNN, C.; OSTERMILLER, S.; RUDOLFSEN, G. Impaired swimming behaviour and morphology of sperm from barn swallows hirundo rustica in Chernobyl. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, v. 650, n. 2, p. 210–216, 2008.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; M., F. The exponentiated Weibull family. *Technometrics*, v. 37, p. 436–45, 1995.

NADARAJAH, S. The exponentiated exponential distribution: a survey. *Advances in Statistical Analysis*, v. 95, n. 3, p. 219–251, 2011.

NADARAJAH, S.; CORDEIRO, G. M.; ORTEGA, E. M. General results for the kumaraswamy-g distribution. *Journal of Statistical Computation and Simulation*, v. 82, n. 7, p. 951–979, 2012.

NADARAJAH, S.; CORDEIRO, G. M.; ORTEGA, E. M. The zografos–balakrishnan-g family of distributions: Mathematical properties and applications. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 44, n. 1, p. 186–215, 2015.

NADARAJAH, S.; ELJABRI, S. The kumaraswamy gp distribution. *Journal of Data Science*, v. 11, n. 4, p. 739–766, 2013.

NADARAJAH, S.; KOTZ, S. The beta exponential distribution. *Reliability Engineering and System Safety*, 2006.

NANDA, A. K.; DAS, S. Stochastic orders of the marshall–olkin extended distribution. *Statistics & Probability Letters*, v. 82, n. 2, p. 295–302, 2012.

NELSON, W.; HAHN, G. J. Linear estimation of a regression relationship from censored data part i—simple methods and their application. *Technometrics*, v. 14, n. 2, p. 247–269, 1972.

NICHOLS, M. D.; PADGETT, W. A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International*, v. 22, n. 2, p. 141–151, 2006.

NUZZO, R. Statistical errors. *Nature*, v. 506, n. 13, p. 150–152, 2014.

O'CONNOR, P.; KLEYNER, A. *Practical reliability engineering*. New Delhi, India: John Wiley & Sons, 2011.

OJO, M. O.; OLAPADE, A. On the generalized logistic and log-logistic distributions. *Kragujevac Journal of Mathematics*, v. 25, p. 65–73, 2003.

PAPOULIS, A.; PILLAI, S. U. *Probability, random variables, and stochastic processes*. [S.l.]: Tata McGraw-Hill Education, 2002.

PARZEN, E. Quantile probability and statistical data modeling. *Statistical Science*, v. 19, n. 4, p. 652–662, 2004.

PASCOA, M. A. de; ORTEGA, E. M.; CORDEIRO, G. M. The kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*, v. 8, n. 5, p. 411–433, 2011.

PAULINO, C. D. M.; TURKMAN, M. A. A.; MURTEIRA, B. *Estatística Bayesiana*. [S.l.: s.n.], 2003.

PESCIM, R. R.; DEMÉTRIO, C. G.; CORDEIRO, G. M.; ORTEGA, E. M.; URBANO, M. R. The beta generalized half-normal distribution. *Computational Statistics & Data Analysis*, v. 54, n. 4, p. 945–957, 2010.

PINHO, L.; CORDEIRO, G.; NOBRE, J. The gamma-exponentiated weibull distribution. *Journal of Statistical Theory and Applications*, v. 11, n. 4, p. 379–395, 2012.

PREDA, V.; PANAITESCU, E.; CIUMARA, R. The modified exponential-poisson distribution. *Proceedings of the Romanian Academy*, v. 12, n. 1, p. 22–29, 2011.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2014. Disponível em: <http://www.R-project.org/>.

RAMOS, M. W. A.; CORDEIRO, G. M.; MARINHO, P. R. D.; DIAS, C. R. B.; HAMEDANI, G. G. The zografos-balakrishnan log-logistic distribution: Properties and applications. *Journal of Statistical Theory and Applications*, v. 12,n. 3, p. 225–244, 2013.

RAQAB, M. Z. Inferences for generalized exponential distribution based on record statistics. *J. Stat. Plann. Infer.*, v. 104, p. 339–350, 2002.

RAQAB, M. Z.; AHSANULLAH, M. Estimation of the location and scale parameters of generalized exponential distribution based on order statistics. *J. Stat. Comput. Simul.*, v. 69, p. 109–124, 2001.

REN, B.; ROBERT, F.; WYRICK, J. J.; APARICIO, O.; JENNINGS, E. G.; SIMON, I.; ZEITLINGER, J.; SCHREIBER, J.; HANNETT, N.; KANIN, E. *et al.* Genome-wide location and function of dna binding proteins. *Science*, v. 290, n. 5500, p. 2306–2309, 2000.

RINNE, H. *The Weibull distribution: a handbook*. USA: CRC Press, 2010.

RISTIĆ, M. M.; BALAKRISHNAN, N. The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, v. 82, n. 8, p. 1191–1206, 2012.

RODRIGUES, J.; CANCHO, V.; DE CASTRO, M. Teoria unificada de análise de sobrevivência. *Associação Brasileira de Estatística, São Paulo, SP*, 2008.

ROSIN, P.; RAMMLER, E. The laws governing the fineness of powdered coal. *J. Inst. Fuel*, v. 7, n. 31, p. 29–36, 1933.

ROUSSAS, G. G. *A course in mathematical statistics*. USA: Academic Press, 1997.

SANTANA, T. V. F. de; ORTEGA, E. M.; CORDEIRO, G. M.; SILVA, G. O. The kumaraswamy-log-logistic distribution. *Journal of Statistical Theory and Applications*, v. 11, n. 3, p. 265–291, 2012.

SANTOS-NETO, M.; BOURGUIGNON, M.; ZEA, L. M.; NASCIMENTO, A. D.; CORDEIRO, G. M. The Marshall-Olkin extended Weibull family of distributions. *Journal of Statistical Distributions and Applications*, v. 1, n. 1, p. 9, 2014.

SARHAN, A. M. Analysis of incomplete, censored data in competing risks models with generalized exponential distributions. *IEEE Trans. Reliab.*, v. 56, p. 132–138, 2007.

SAULO, H.; LEÃO, J.; BOURGUIGNON, M. The kumaraswamy birnbaum–saunders distribution. *Journal of Statistical Theory and Practice*, v. 6, n. 4, p. 745–759, 2012.

SCHWARZ, G. Estimating the dimension of a model. *The annals of statistics*, v. 6, n. 2, p. 461–464, 1978.

SHANNO, D. F. Conditioning of quasi-newton methods for function minimization. *Mathematics of computation*, v. 24, n. 111, p. 647–656, 1970.

SHAO, J.; TU, D. *The jackknife and bootstrap*. USA: Springer, 1995.

SMIRNOV, N. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, p. 279–281, 1948.

SMITH, A. F.; GELFAND, A. E. Bayesian statistics without tears: a sampling–resampling perspective. *The American Statistician*, v. 46, n. 2, p. 84–88, 1992.

SMITH, R. L.; NAYLOR, J. A comparison of maximum likelihood and bayesian estimators for the three-parameter Weibull distribution. *Applied Statistics*, p. 358–369, 1987.

STAHL, B.; GEYER, J. F. Fatigue reliability of parallel member systems. *Journal of Structural Engineering*, v. 110, n. 10, p. 2307–2323, 1984.

SUGIURA, N. Further analysts of the data by Akaike's information criterion and the finite corrections: Further analysts of the data by Akaike's. *Communications in Statistics-Theory and Methods*, v. 7, n. 1, p. 13–26, 1978.

THOMAS, A.; JOSE, K. Bivariate semi-Pareto minification processes. *Metrika*, v. 59, n. 3, p. 305–313, 2004.

United States Geological Survey. *Historic Earthquakes and Earthquake Statistics: where do earthquakes occur?* junho de 2013. Disponível em: <earthquakes.usgs.gov>.

University of Massachusetts Amherst. *Index of Survival Analysis Datasets*. novembro de 2014. Disponível em: <http://www.umass.edu/statdata/statdata/stat-survival.html>.

WASSERMAN, L. *All of Statistics: a concise course in statistics inference*. 2. ed. Heidelberg: Springer, 2004.

WEIBULL, W. A statistical distribution function of wide applicability. *Journal of applied mechanics*, v. 18, n. 3, p. 293–297, 1951.

WETZELS, R.; MATZKE, D.; LEE, M. D.; ROUDER, J. N.; IVERSON, G. J.; WAGENMAKERS, E.-J. Statistical evidence in experimental psychology an empirical comparison using 855 t tests. *Perspectives on Psychological Science*, v. 6, n. 3, p. 291–298, 2011.

WINNIE-JR, J.; CHRISTIANSON, D.; CREEL, S.; MAXWELL, B. Elk decision-making rules are simplified in the presence of wolves. *Behavioral Ecology and Sociobiology*, v. 61, n. 2, p. 277–289, 2006.

ZHANG, T.; XIE, M. Failure data analysis with extended weibull distribution. *Communications in Statistics—Simulation and Computation*, v. 36, n. 3, p. 579–592, 2007.

ZHAO, Z.; HALDAR, A.; BREEN JR, F. L. Fatigue-reliability evaluation of steel bridges. *Journal of Structural Engineering*, v. 120, n. 5, p. 1608–1623, 1994.

ZHENG, G. Fisher information matrix in type-ii censored data from exponentiated exponential family. *Biometrical J.*, v. 44, p. 353–357, 2002.

ZOGRAFOS, K.; BALAKRISHNAN, N. On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, v. 6, n. 4, p. 344–362, 2009.

APÊNDICE

APÊNDICE A

Matriz Observada de Fisher

$$U_{\alpha,\alpha} = -\sum_{j=0}^{\infty} \frac{\left(1-G(\mu)^{\gamma}\right)^2}{\left(\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\right)^2} + (1+\theta)\sum_{j=0}^{\infty}\left(\frac{G(\mu)^{\alpha}\log^2(G(\mu))}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}} - \frac{G(\mu)^{2\alpha}\log^2(G(\mu))}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^2}\right)$$

$$U_{\alpha,\beta} = -(1+\theta)\sum_{j=0}^{\infty}\frac{G(\mu)^{\alpha}\left(1-G(\mu)^{\gamma}\right)^{\delta}\log(G(\mu))}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^2}$$

$$U_{\alpha,\gamma} = (1+\theta)\sum_{j=0}^{\infty}\frac{\beta\delta G(\mu)^{\alpha+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}\log^2(G(\mu))}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^2} - \sum_{j=0}^{\infty}\left(\frac{G(\mu)^{\gamma}\log(G(\mu))}{\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}} + \frac{\left(1-G(\mu)^{\gamma}\right)\left(\delta G(\mu)^{\gamma}+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\log(G(\mu))\right)}{\left(\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\right)^2}\right)$$

$$U_{\alpha,\delta} = -\sum_{j=0}^{\infty} \frac{\gamma G(\mu)^{\gamma}\left(1 - G(\mu)^{\gamma}\right)}{\left(\alpha + (-\alpha + \gamma\delta)G(\mu)^{\gamma}\right)^{2}} + (1+\theta)\sum_{j=0}^{\infty} -\frac{\beta G(\mu)^{\alpha}\left(1 - G(\mu)^{\gamma}\right)^{\delta}\log(G(\mu))\log\left(1 - G(\mu)^{\gamma}\right)}{\left(G(\mu)^{\alpha} + \beta\left(1 - G(\mu)^{\gamma}\right)^{\delta}\right)^{2}}$$

$$U_{\alpha,\theta} = \sum_{j=0}^{\infty} \frac{G(\mu)^{\alpha}\log(G(\mu))}{G(\mu)^{\alpha} + \beta\left(1 - G(\mu)^{\gamma}\right)^{\delta}}$$

$$\begin{aligned}
U_{\alpha,\mu} = &-\sum_{j=0}^{\infty} \frac{G'(\mu)}{G(\mu)} + \sum_{j=0}^{\infty}\left(-\frac{\gamma(-\alpha+\gamma\delta)G(\mu)^{-1+\gamma}\left(1 - G(\mu)^{\gamma}\right)G'(\mu)}{\left(\alpha + (-\alpha+\gamma\delta)G(\mu)^{\gamma}\right)^{2}} - \frac{\gamma G(\mu)^{-1+\gamma}G'(\mu)}{\alpha + (-\alpha+\gamma\delta)G(\mu)^{\gamma}}\right) \\
&+ (1+\theta)\sum_{j=0}^{\infty}\left(\frac{G(\mu)^{-1+\alpha}G'(\mu)}{G(\mu)^{\alpha} + \beta\left(1 - G(\mu)^{\gamma}\right)^{\delta}} + \frac{\alpha G(\mu)^{-1+\alpha}\log(G(\mu))G'(\mu)}{G(\mu)^{\alpha} + \beta\left(1 - G(\mu)^{\gamma}\right)^{\delta}}\right. \\
&\left. - \frac{G(\mu)^{\alpha}\log(G(\mu))\left(\alpha G(\mu)^{-1+\alpha}G'(\mu) - \beta\gamma\delta G(\mu)^{-1+\gamma}\left(1 - G(\mu)^{\gamma}\right)^{-1+\delta}G'(\mu)\right)}{\left(G(\mu)^{\alpha} + \beta\left(1 - G(\mu)^{\gamma}\right)^{\delta}\right)^{2}}\right)
\end{aligned}$$

$$U_{\beta,\beta} = -\frac{n}{\beta^{2}} + (1+\theta)\sum_{j=0}^{\infty} -\frac{\left(1 - G(\mu)^{\gamma}\right)^{2\delta}}{\left(G(\mu)^{\alpha} + \beta\left(1 - G(\mu)^{\gamma}\right)^{\delta}\right)^{2}}$$

$$U_{\beta,\gamma} = (1+\theta)\sum_{j=0}^{\infty}\left(\frac{\beta\delta G(\mu)^{\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+2\delta}\log(G(\mu))}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^{2}} - \frac{\delta G(\mu)^{\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}\log(G(\mu))}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}\right)$$

$$U_{\beta,\delta} = (1+\theta)\sum_{j=0}^{\infty}\left(\frac{\left(1-G(\mu)^{\gamma}\right)^{\delta}\log\left(1-G(\mu)^{\gamma}\right)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}} - \frac{\beta\left(1-G(\mu)^{\gamma}\right)^{2\delta}\log\left(1-G(\mu)^{\gamma}\right)}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^{2}}\right)$$

$$U_{\beta,\theta} = \sum_{j=0}^{\infty}\frac{\left(1-G(\mu)^{\gamma}\right)^{\delta}}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}$$

$$U_{\beta,\mu} = -(1+\theta)\sum_{j=0}^{\infty}\left(\frac{\gamma\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}G'(\mu)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}} + \frac{\left(1-G(\mu)^{\gamma}\right)^{\delta}\left(\alpha G(\mu)^{-1+\alpha}G'(\mu)-\beta\gamma\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}G'(\mu)\right)}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^{2}}\right)$$

$$
\begin{aligned}
U_{\gamma,\gamma} = {} & (\delta-1)\sum_{j=0}^{\infty}\left(-\frac{G(\mu)^{2\gamma}\log^2(G(\mu))}{(1-G(\mu)^{\gamma})^2}-\frac{G(\mu)^{\gamma}\log^2(G(\mu))}{1-G(\mu)^{\gamma}}\right) \\
& +(1+\theta)\sum_{j=0}^{\infty}\left(-\frac{\beta^2\delta^2G(\mu)^{2\gamma}\left(1-G(\mu)^{\gamma}\right)^{-2+2\delta}\log^2(G(\mu))}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^2}\right. \\
& \left.+\frac{\beta(-1+\delta)\delta G(\mu)^{2\gamma}\left(1-G(\mu)^{\gamma}\right)^{-2+\delta}\log^2(G(\mu))}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}-\frac{\beta\delta G(\mu)^{\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}\log^2(G(\mu))}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}\right) \\
& +\sum_{j=0}^{\infty}\left(-\frac{(\delta G(\mu)^{\gamma}+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\log(G(\mu)))^2}{(\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma})^2}+\frac{2\delta G(\mu)^{\gamma}\log(G(\mu))+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\log^2(G(\mu))}{\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}}\right)
\end{aligned}
$$

$$
\begin{aligned}
U_{\gamma,\delta} = {} & \sum_{j=0}^{\infty}-\frac{G(\mu)^{\gamma}\log(G(\mu))}{1-G(\mu)^{\gamma}}+\sum_{j=0}^{\infty}\left(\frac{G(\mu)^{\gamma}+\gamma G(\mu)^{\gamma}\log(G(\mu))}{\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}}-\frac{\gamma G(\mu)^{\gamma}\left(\delta G(\mu)^{\gamma}+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\log(G(\mu))\right)}{(\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma})^2}\right) \\
& +(1+\theta)\sum_{j=0}^{\infty}\left(-\frac{\beta G(\mu)^{\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}\log(G(\mu))}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}+\frac{\beta^2\delta G(\mu)^{\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+2\delta}\log(G(\mu))\log\left(1-G(\mu)^{\gamma}\right)}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^2}\right. \\
& \left.-\frac{\beta\delta G(\mu)^{\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}\log(G(\mu))\log\left(1-G(\mu)^{\gamma}\right)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}\right)
\end{aligned}
$$

$$
U_{\gamma,\theta}=\sum_{j=0}^{\infty}-\frac{\beta\delta G(\mu)^{\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}\log(G(\mu))}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}
$$

$$
\begin{aligned}
U_{\gamma,\mu} = {} & (-1+\delta)\sum_{j=0}^{\infty}\left(-\frac{G(\mu)^{-1+\gamma}G'(\mu)}{1-G(\mu)^{\gamma}} - \frac{\gamma G(\mu)^{-1+2\gamma}\log(G(\mu))G'(\mu)}{\left(1-G(\mu)^{\gamma}\right)^{2}} - \frac{\gamma G(\mu)^{-1+\gamma}\log(G(\mu))G'(\mu)}{1-G(\mu)^{\gamma}}\right) \\
& + (1+\theta)\sum_{j=0}^{\infty}\left(-\frac{\beta\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}G'(\mu)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}} + \frac{\beta\gamma(-1+\delta)\delta G(\mu)^{-1+2\gamma}\left(1-G(\mu)^{\gamma}\right)^{-2+\delta}\log(G(\mu))G'(\mu)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}\right. \\
& - \frac{\beta\gamma\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}\log(G(\mu))G'(\mu)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}} \\
& \left. + \frac{\beta\delta G(\mu)^{\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}\log(G(\mu))\left(\alpha G(\mu)^{-1+\alpha}G'(\mu)-\beta\gamma\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}G'(\mu)\right)}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^{2}}\right) \\
& + \sum_{j=0}^{\infty}\left(-\frac{\gamma(-\alpha+\gamma\delta)G(\mu)^{-1+\gamma}\left(\delta G(\mu)^{\gamma}+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\log(G(\mu))\right)G'(\mu)}{\left(\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\right)^{2}}\right. \\
& \left. + \frac{\gamma\delta G(\mu)^{-1+\gamma}G'(\mu)+(-\alpha+\gamma\delta)G(\mu)^{-1+\gamma}G'(\mu)+\gamma(-\alpha+\gamma\delta)G(\mu)^{-1+\gamma}\log(G(\mu))G'(\mu)}{\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}}\right)
\end{aligned}
$$

$$
U_{\delta,\delta} = \sum_{j=0}^{\infty} -\frac{\gamma^{2}G(\mu)^{2\gamma}}{\left(\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\right)^{2}} + (1+\theta)\sum_{j=0}^{\infty}\left(-\frac{\beta^{2}\left(1-G(\mu)^{\gamma}\right)^{2\delta}\log^{2}\left(1-G(\mu)^{\gamma}\right)}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^{2}} + \frac{\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\log^{2}\left(1-G(\mu)^{\gamma}\right)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}\right)
$$

$$
U_{\delta,\theta} = \sum_{j=0}^{\infty}\frac{\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\log\left(1-G(\mu)^{\gamma}\right)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}
$$

$$U_{\delta,\mu} = \sum_{j=0}^{\infty} -\frac{\gamma G(\mu)^{-1+\gamma} G'(\mu)}{1-G(\mu)^{\gamma}} + \sum_{j=0}^{\infty} \left(-\frac{\gamma^2(-\alpha+\gamma\delta)G(\mu)^{-1+2\gamma}G'(\mu)}{(\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma})^2} + \frac{\gamma^2 G(\mu)^{-1+\gamma}G'(\mu)}{\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}} \right)$$
$$+ (1+\theta)\sum_{j=0}^{\infty} \left(-\frac{\beta\gamma G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta} G'(\mu)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}} - \frac{\beta\gamma\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}\log\left(1-G(\mu)^{\gamma}\right)G'(\mu)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}} \right.$$
$$\left. - \frac{\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\log\left(1-G(\mu)^{\gamma}\right)\left(\alpha G(\mu)^{-1+\alpha}G'(\mu) - \beta\gamma\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}G'(\mu)\right)}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^2} \right)$$

$$U_{\theta,\theta} = -\frac{n}{\theta^2}$$

$$U_{\theta,\mu} = \sum_{j=0}^{\infty} \frac{\alpha G(\mu)^{-1+\alpha}G'(\mu) - \beta\gamma\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}G'(\mu)}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}$$

$$
\begin{aligned}
U_{\mu,\mu} = & \sum_{j=0}^{\infty}\left(-\frac{g'(\mu)^2}{g(\mu)^2}+\frac{g''(\mu)}{g(\mu)}\right)-(1+\alpha)\sum_{j=0}^{\infty}\left(-\frac{G'(\mu)^2}{G(\mu)^2}+\frac{G''(\mu)}{G(\mu)}\right)\\
& +(-1+\delta)\sum_{j=0}^{\infty}\left(-\frac{\gamma^2 G(\mu)^{-2+2\gamma}G'(\mu)^2}{\left(1-G(\mu)^{\gamma}\right)^2}-\frac{(-1+\gamma)\gamma G(\mu)^{-2+\gamma}G'(\mu)^2}{1-G(\mu)^{\gamma}}-\frac{\gamma G(\mu)^{-1+\gamma}G''(\mu)}{1-G(\mu)^{\gamma}}\right)\\
& +\sum_{j=0}^{\infty}\left(-\frac{\gamma^2(-\alpha+\gamma\delta)^2 G(\mu)^{-2+2\gamma}G'(\mu)^2}{\left(\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}\right)^2}+\frac{(-1+\gamma)\gamma(-\alpha+\gamma\delta)G(\mu)^{-2+\gamma}G'(\mu)^2}{\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}}+\frac{\gamma(-\alpha+\gamma\delta)G(\mu)^{-1+\gamma}G''(\mu)}{\alpha+(-\alpha+\gamma\delta)G(\mu)^{\gamma}}\right)\\
& (1+\theta)\sum_{j=0}^{\infty}\Bigg(-\frac{\left(\alpha G(\mu)^{-1+\alpha}G'(\mu)-\beta\gamma\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}G'(\mu)\right)^2}{\left(G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}\right)^2}\\
& +\frac{1}{G(\mu)^{\alpha}+\beta\left(1-G(\mu)^{\gamma}\right)^{\delta}}\Big[(-1+\alpha)\alpha G(\mu)^{-2+\alpha}G'(\mu)^2+\beta\gamma^2(-1+\delta)\delta G(\mu)^{-2+2\gamma}\left(1-G(\mu)^{\gamma}\right)^{-2+\delta}G'(\mu)^2\\
& -\beta(-1+\gamma)\gamma\delta G(\mu)^{-2+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}G'(\mu)^2+\alpha G(\mu)^{-1+\alpha}G''(\mu)-\beta\gamma\delta G(\mu)^{-1+\gamma}\left(1-G(\mu)^{\gamma}\right)^{-1+\delta}G''(\mu)\Big]\Bigg)
\end{aligned}
$$

APÊNDICE B

Código R Utilizado na Simulação

Algoritmo R utilizado na seção 5.1.

Arquivo fonte com densidades

```
#------------ Kumaraswamy Beta -------------------------------------#

# Kumaraswamy Beta - Probability density function.
pdf_kwbeta <- function(par,x){
   a = par[1]
   beta = par[2]
   b = par[3]
   alpha = par[4]
   (a*b*x^(alpha-1)*(1-x)^(beta-1)*(pbeta(x,alpha,beta))^(a-1)*
   (1-pbeta(x,alpha,beta)^a)^(b-1))/beta(alpha,beta)
}
# Kumaraswamy Beta - Cumulative distribution function.
cdf_kwbeta <- function(par,x){
   a = par[1]
   beta = par[2]
   b = par[3]
   alpha = par[4]
```

```
   1 - (1 - pbeta(x,alpha,beta)^a)^b
}


#------------ Exponentiated Weibull --------------------------------#
# Exponentiated Weibull - Probability density function.
pdf_expweibull <- function(par,x){
   a = par[1]
   c = par[2]
   beta = par[3]
   a * beta * c * exp(-(beta*x)^c) * (beta*x)^(c-1) * (1 - exp(-(beta*x)^c))^(a-1)
}
# Exponentiated Weibull - Cumulative distribution function.
cdf_expweibull <- function(par,x){
   a = par[1]
   c = par[2]
   beta = par[3]
   (1 - exp(-(beta*x)^c))^a
}


#----------- Kumaraswamy Weibull Poisson -----------------------------#
# Kumaraswamy Weibull Poisson - Probability density function.
pdf_kwweibullpoisson <- function(par,x){
   a = par[1]
   b = par[2]
   c = par[3]
   lambda = par[4]
   beta = par[5]
   (a*b*c*lambda*(beta^c)*(x^(c-1))*((1-exp(-(x*beta)^c))^(a-1)) *
   ((1-(1-exp(-(beta*x)^c))^a)^(b-1)) *
   exp(-lambda*(1-(1-(1-exp(-(beta*x)^c))^a)^b)
   - (beta*x)^c))/(1-exp(-lambda))
}
# Kumaraswamy Weibull Poisson - Cumulative distribution function.
cdf_kwweibullpoisson <- function(par,x){
   a = par[1]
   b = par[2]
   c = par[3]
```

```
   lambda = par[4]
   beta = par[5]
   (1 - exp(lambda*(-(1-(1-(1-exp(-(x*beta)^c))^a)^b))))/(1-exp(-lambda))
}


#----------- BGE -----------------------------#
pdf_bge <- function(par,x){
   a = par[1]
   b = par[2]
   alpha = par[3]
   lamb = par[4]
   g = alpha*lamb * exp(-lamb*x)*(1-exp(-lamb*x))^(alpha-1)
   G = (1 - exp(-lamb*x))^alpha
   1/beta(a,b)*g*G^(a-1)*(1-G)^(b-1)
}
cdf_bge <- function(par,x){
   a = par[1]
   b = par[2]
   alpha = par[3]
   lamb = par[4]
   G = (1 - exp(-lamb*x))^alpha
   pbeta(G,a,b)
}



#----------- BFrechet -----------------------------#
pdf_bfret <- function(par,x){
   a = par[1]
   b = par[2]
   k = par[3]
   s = par[4]
   g = k*s^k*(x)^(-k-1)*exp(-(s/x)^(k))
   G = exp(-(s/x)^k)
   1/beta(a,b)*g*G^(a-1)*(1-G)^(b-1)
}
cdf_bfret <- function(par,x){
   a = par[1]
```

```
   b = par[2]
   k = par[3]
   s = par[4]
   G = exp(-(s/x)^k)
   pbeta(G,a,b)
}


#----------- Weibull Geometrica -----------------------------#
pdf_wgeo <- function(par,x){
   a = par[1]
   b = par[2]
   p = par[3]
   while(0>p | p>1) p = runif(1)
   g = a*b^a*(1-p)*x^(a-1)*exp(-(b*x)^a)/(1 - p*exp(-(b*x)^a))^2
   G = (1 - exp(-(b*x)^a) )/(1 - p*exp(-(b*x)^a) )
   1/beta(a,b)*g*G^(a-1)*(1-G)^(b-1)
}
cdf_wgeo <- function(par,x){
   a = par[1]
   b = par[2]
   p = par[3]
   while(0>p | p>1) p = runif(1)
   G = (1 - exp(-(b*x)^a) )/(1 - p*exp(-(b*x)^a) )
   pbeta(G,a,b)
}



#------------ MOGEW -------------------#
# MOGEW - Probability density function.
pdf_mogeweib <- function(par,x){
   a = par[1]
   b = par[2]
   gm= par[3]
   d = par[4]
   th = par[5]
   lamb = par[6]
   k =par[7]
```

```
   G = 1-exp(-(lamb*x)^k)
   g = lamb*k*(lamb*x)^(k-1)*exp(-(lamb*x)^k)
   b*th*g*G^(-a-1)*(1-G^gm)^(d-1)*(a+(d*gm-a)*G^gm)*
   (G^a/(G^a+b*(1-G^gm)^d))^(th+1)
}
# MOGEW - Cumulative distribution function.
cdf_mogeweib <- function(par,x){
   a = par[1]
   b = par[2]
   gm= par[3]
   d = par[4]
   th = par[5]
   lamb = par[6]
   k =par[7]
   G = 1-exp(-(lamb*x)^k)
   g = lamb*k*(lamb*x)^(k-1)*exp(-(lamb*x)^k)
   (G^a/(G^a+b*(1-G^gm)^d) )^th
}



#------------ Marshall-Olkin Weibull --------------------------------#
# Marshall-Olkin Weibull - Probability density function.
pdf_moweib <- function(par,x){
   a = 1
   b = par[1]
   gm= 1
   d = 1
   th = 1
   lamb = par[2]
   k =par[3]
   G = 1-exp(-(lamb*x)^k)
   g = lamb*k*(lamb*x)^(k-1)*exp(-(lamb*x)^k)
   b*th*g*G^(-a-1)*(1-G^gm)^(d-1)*(a+(d*gm-a)*G^gm)*
   (G^a/(G^a+b*(1-G^gm)^d))^(th+1)
}
# Marshall-Olkin Weibull - Cumulative distribution function.
cdf_moweib <- function(par,x){
   a = 1
```

```
   b = par[1]
   gm= 1
   d = 1
   th = 1
   lamb = par[2]
   k =par[3]
   G = 1-exp(-(lamb*x)^k)
   g = lamb*k*(lamb*x)^(k-1)*exp(-(lamb*x)^k)
   (G^a/(G^a+b*(1-G^gm)^d) )^th
}



#------------ Exponential Weibull --------------------------------#
# Marshall-Olkin Weibull - Probability density function.
pdf_eweib <- function(par,x){
   a = par[1]
   b = 1
   gm= par[1]
   d = 1
   th = 1
   lamb = par[2]
   k =par[3]
   G = 1-exp(-(lamb*x)^k)
   g = lamb*k*(lamb*x)^(k-1)*exp(-(lamb*x)^k)
   b*th*g*G^(-a-1)*(1-G^gm)^(d-1)*(a+(d*gm-a)*G^gm)*(G^a/(G^a+b*(1-G^gm)^d))^(th+1)
}
# Marshall-Olkin Weibull - Cumulative distribution function.
cdf_eweib <- function(par,x){
   a = par[1]
   b = 1
   gm= par[1]
   d = 1
   th = 1
   lamb = par[2]
   k =par[3]
   G = 1-exp(-(lamb*x)^k)
   g = lamb*k*(lamb*x)^(k-1)*exp(-(lamb*x)^k)
   (G^a/(G^a+b*(1-G^gm)^d) )^th
```

```
}
```

```
#------------ Marshall-Olkin Exponentialized Exponencial ------------------#
pdf_mogeexp <- function(par,x){
   a = par[1]
   b = par[2]
   gm= par[3]
   d = par[4]
   th = par[5]
   lamb = par[6]
   k = 1
   G = 1-exp(-(lamb*x)^k)
   g = lamb*k*(lamb*x)^(k-1)*exp(-(lamb*x)^k)
   b*th*g*G^(-a-1)*(1-G^gm)^(d-1)*(a+(d*gm-a)*G^gm)*
   (G^a/(G^a+b*(1-G^gm)^d))^(th+1)
}
cdf_mogeexp <- function(par,x){
   a = par[1]
   b = par[2]
   gm= par[3]
   d = par[4]
   th = par[5]
   lamb = par[6]
   k = 1
   G = 1-exp(-(lamb*x)^k)
   #g = lamb*k*(lamb*x)^(k-1)*exp(-(lamb*x)^k)
   (G^a/(G^a+b*(1-G^gm)^d) )^th
}
```

```
#---------------------------- Weibull --------------------------------#
# Weibull - Probability density function.
pdf_weib <- function(par,x){
   lamb = par[1]
   th = par[2]
```

```
   lamb*th*(lamb*x)^(th-1)*exp(-(lamb*x)^th)
}
# Weibull - Cumulative distribution function.
cdf_weib <- function(par,x){
   lamb = par[1]
   th = par[2]
   1 - exp(-(lamb*x)^th)
}




#----------- Marshall-Olkin Exponentialized Fretchet -------------#
# M-O Exponentialized Fretchet - Probability density function.
pdf_moefret <- function(par,x){
   a = par[1]
   b = par[2]
   gm= par[3]
   d = par[4]
   lamb = par[5]
   th = par[6]
   m = 0
   z=(x-m)/lamb
   g = th/lamb*z^(-1-th)*exp(-z^(-th))
   G =    exp(-z^(-th))
   b*g*G^(a-1)*(1-G^gm)^(d-1)*(a+(d*gm-a)*
   G^gm)/(G^a+b*(1-G^gm)^d)^2
}
# M-O Exponentialized Fretchet - Cumulative distribution function.
cdf_moefret <- function(par,x){
   a = par[1]
   b = par[2]
   gm= par[3]
   d = par[4]
   lamb = par[5]
   th = par[6]
   m = 0
   z=(x-m)/lamb
   #g = th/lamb*z^(-1-th)*exp(-z^(-th))
   G =    exp(-z^(-th))
```

```
   G^a/(G^a+b*(1-G^gm)^d)
}


#---------------------------- Fretchet -------------------------------#
# Fretchet - Probability density function.
pdf_fret <- function(par,x){
   m = 0
   lamb = par[1]
   th = par[2]
   z=(x-m)/lamb
   th/lamb*z^(-1-th)*exp(-z^(-th))
}
# Fretchet - Cumulative distribution function.
cdf_fret <- function(par,x){
   m = 0
   lamb = par[1]
   th = par[2]
   z=(x-m)/lamb
   exp(-z^(-th))
}


#---------------------------- Gumbel -------------------------------#
# Gumbel - Probability density function.
pdf_gumb <- function(par,x){
   mu = par[1]
   sig = par[2]
   z=(x-mu)/sig
   1/sig*exp(-(z+exp(-z)))
}
# Gumbel - Cumulative distribution function.
cdf_gumb <- function(par,x){
   mu = par[1]
   sig = par[2]
   z=(x-mu)/sig
   exp(-exp(-z))
}
```

```
#---------------------------- Gama --------------------------------#
# Gama - Probability density function.
pdf_gama <- function(par,x){
   a = par[1]
   b = par[2]
   x^(a-1)*b^a*exp(-b*x)/gamma(a)
}
# Gama - Cumulative distribution function.
cdf_gama <- function(par,x){
   a = par[1]
   b = par[2]
   pgamma(x,a,b)
   #integrand <- function(x) {x^(a-1)*b^a*exp(-b*x)/gamma(a)}
   #integrate(integrand, lower = 0, upper = Inf)$value
}


#---------------------------- Birnbaum Saunderes----------------------#
# Birnbaum Saunderes - Probability density function.
pdf_bs = function(par, x){
   a = par[1]
   b = par[2]
   at=1/a*(sqrt(x/b)-sqrt(b/x))
   jacobian = (sqrt(b/x) + sqrt(x/b))/(2*a*x)
   dnorm(at)*jacobian
}
# Birnbaum Saunderes - Cumulative distribution function.
cdf_bs = function(par, x){
   a = par[1]
   b = par[2]
   at=1/a*(sqrt(x/b)-sqrt(b/x))
   1 - pnorm(-at)
}


#---------------------------- Birnbaum Saunderes----------------------#
# Birnbaum Saunderes - Probability density function.
pdf_lnorm = function(par, x){
```

```
   mu = par[1]
   sig = par[2]
   z = (-log(x) + mu)/sig
   1/(sqrt(2*pi)*x*sig)*exp(-.5*z^2)
}
cdf_lnorm = function(par, x){
   mu = par[1]
   sig = par[2]
   z = (-log(x) + mu)/sig
   1 - pnorm(-z)
}


#----------------------------- Log Logistica----------------------#
# Birnbaum Saunderes - Probability density function.
pdf_llogis = function(par, x){
   lamb = par[1]
   k = par[2]
   k/lamb^k*x^(k-1)*(1+(x/lamb)^k)^(-2)
}
cdf_llogis = function(par, x){
   lamb = par[1]
   k = par[2]
   1 - 1/(1+(x/lamb)^k)
}
```

Arquivo fonte com função bootstrap

```
library(boot)

mle.boot = function(data, pdfm, cdfm, par, meth, R){
bs <- function(data, indices) {
  n = length(data)-1
  fit <- goodness.fit(pdf=pdfm, cdf=cdfm,
         starts = par, data =  sample(data,n,rep=F),
         method=meth, domain = c(0, Inf),mle=NULL)
  return( fit$mle)
}
```

```
boot(data=data, statistic=bs,
    R=R,  parallel = 'multicore')
}
```

Arquivo principal

```
setwd('C:/Users/Kleber/Dropbox/Tese/Tese1/Comandos/Estimação')
library(AdequacyModel)
source('pdfs.txt')
source('bootstrapping.txt')

setwd('C:/Users/Kleber/Dropbox/Tese/Tese1/Graficos')


a = 1.5; b = 1.2; gm = a; d = 1;
th = .3; lamb = 1.7; k = 2.5
par = c(a, b, gm, d, th, lamb, k)
u = runif(150, 0, 1)

v = ( b*u^(1/th) / ( 1 + b*u^(1/th) - u^(1/th) ) )^(1/a)
x = (- (1/lamb^k) * log(1 - v))^(1/k)
x.test = x[101:150]; x = x[1:100]

fit <- goodness.fit(pdf=pdf_mogeweib, cdf=cdf_mogeweib,
         starts = par, data =  x,
         method='B', domain=c(0,Inf),mle=NULL)

fitb = mle.boot(x, pdf_mogeweib, cdf_mogeweib,
par, 'B', 1000)

p=0
for(i in 1:length(fitb$t0)) p[i]= mean(fitb$t[,i])

fitc = 2*fitb$t0 - p

#p=matrix(0,5,7)
```

```
#for(i in 1:nrow(p)){
#   fitb = mle.boot(x, pdf_mogeweib, cdf_mogeweib,
#   par, 'B', 1000)
#   for(j in 1:ncol(p)) p[i,j]= mean(fitb$t[,j])
#}
#round(apply(p, 2, var),4)
var.p = c(0.0002, 0.0027, 0.0009, 0.0010, 0.0000, 0.0002, 0.0002)


#png('histtrain.png')
postscript('histtrain.eps')
hist(x, freq=F, main= '',
xlab= 'Dados Simulados - Treino (n = 100)', ylab='Densidade', cex.lab = 1.5)
curve(pdf_mogeweib(par,x),add=T,lwd=3)
curve(pdf_mogeweib(fit$mle,x),add=T,col=2,lwd=3, lty=2)
curve(pdf_mogeweib(fitb$t0,x),add=T,col=3,lwd=3, lty=2)
curve(pdf_mogeweib(fitc,x),add=T,col=4,lwd=3, lty=2)
legend('topright',c('Parâmetro', 'EMV', 'Bootstrap',
'Boot. Corrigido'), col=1:4, lwd=3, lty = c(1,2,2,2), cex = 1.5)
dev.off()

#png('histvalid.png')
postscript('histvalid.eps')
hist(x.test, freq=F, main= '',
xlab= 'Dados Simulados - Validação (n = 50)',
ylab='Densidade',ylim=c(0,1.8), cex.lab = 1.5)
curve(pdf_mogeweib(par,x),add=T,lwd=3, lty = c(1,2,2,2))
curve(pdf_mogeweib(fit$mle,x),add=T,col=2,lwd=3, lty=2)
curve(pdf_mogeweib(fitb$t0,x),add=T,col=3,lwd=3, lty=2)
curve(pdf_mogeweib(fitc,x),add=T,col=4,lwd=3, lty=2)
legend('topright',c('Parâmetro', 'EMV', 'Bootstrap',
'Boot. Corrigido'), col=1:4, lwd=3, lty = c(1,2,2,2), cex = 1.5)
dev.off()

lt1 = sum(log(pdf_mogeweib(par, x)))
lt2 = sum(log(pdf_mogeweib(fit$mle, x)))
lt3 = sum(log(pdf_mogeweib(fitb$t0, x)))
lt4 = sum(log(pdf_mogeweib(p, x)))
```

```
lv1 = sum(log(pdf_mogeweib(par, x.test)))
lv2 = sum(log(pdf_mogeweib(fit$mle, x.test)))
lv3 = sum(log(pdf_mogeweib(fitb$t0, x.test)))
lv4 = sum(log(pdf_mogeweib(fitc, x.test)))


AICt1 = - 2*lt1 + 2*7
AICt2 = - 2*lt2 + 2*7
AICt3 = - 2*lt3 + 2*7
AICt4 = - 2*lt4 + 2*7

AICv1 = - 2*lv1 + 2*7
AICv2 = - 2*lv2 + 2*7
AICv3 = - 2*lv3 + 2*7
AICv4 = - 2*lv4 + 2*7

AICCt1 = - 2*lt1 + 2*7*100/(100-7-1)
AICCt2 = - 2*lt2 + 2*7*100/(100-7-1)
AICCt3 = - 2*lt3 + 2*7*100/(100-7-1)
AICCt4 = - 2*lt4 + 2*7*100/(100-7-1)

AICCv1 = - 2*lv1 + 2*7*100/(100-7-1)
AICCv2 = - 2*lv2 + 2*7*100/(100-7-1)
AICCv3 = - 2*lv3 + 2*7*100/(100-7-1)
AICCv4 = - 2*lv4 + 2*7*100/(100-7-1)

BICt1 = - 2*lt1 + 7*log(100)
BICt2 = - 2*lt2 + 7*log(100)
BICt3 = - 2*lt3 + 7*log(100)
BICt4 = - 2*lt4 + 7*log(100)

BICv1 = - 2*lv1 + 7*log(100)
BICv2 = - 2*lv2 + 7*log(100)
BICv3 = - 2*lv3 + 7*log(100)
BICv4 = - 2*lv4 + 7*log(100)

HQICt1 = - 2*lt1 + 2*7*log(log(100))
HQICt2 = - 2*lt2 + 2*7*log(log(100))
HQICt3 = - 2*lt3 + 2*7*log(log(100))
```

```
HQICt4 = - 2*lt4 + 2*7*log(log(100))

HQICv1 = - 2*lv1 + 2*7*log(log(100))
HQICv2 = - 2*lv2 + 2*7*log(log(100))
HQICv3 = - 2*lv3 + 2*7*log(log(100))
HQICv4 = - 2*lv4 + 2*7*log(log(100))

round(AICt1,3);round(AICt2,3);round(AICt3,3);round(AICt4,3)
round(AICv1,3);round(AICv2,3);round(AICv3,3);round(AICv4,3)

round(AICCt1,3);round(AICCt2,3);round(AICCt3,3);round(AICCt4,3)
round(AICCv1,3);round(AICCv2,3);round(AICCv3,3);round(AICCv4,3)

round(BICt1,3);round(BICt2,3);round(BICt3,3);round(BICt4,3);
round(BICv1,3);round(BICv2,3);round(BICv3,3);round(BICv4,3);

round(HQICt1,3);round(HQICt2,3);round(HQICt3,3);round(HQICt4,3);
round(HQICv1,3);round(HQICv2,3);round(HQICv3,3);round(HQICv4,3);

round(
matrix(
c(AICt1, AICt2, AICt3, AICt4,
  AICv1, AICv2, AICv3, AICv4,
  AICCt1, AICCt2, AICCt3, AICCt4,
  AICCv1, AICCv2, AICCv3, AICCv4,
  BICt1, BICt2, BICt3, BICt4,
  BICv1, BICv2, BICv3, BICv4,
  HQICt1, HQICt2, HQICt3, HQICt4,
  HQICv1, HQICv2, HQICv3, HQICv4),
 4, 8),3)


round(par,3)

round(fit$mle,3)
round(fit$Erro,3)

round(results$t0,3)
round(sd(fitb$t[,1]),3);round(sd(fitb$t[,2]),3);
```

```
round(sd(fitb$t[,3]),3);round(sd(fitb$t[,4]),3)
round(sd(fitb$t[,5]),3);round(sd(fitb$t[,6]),3);round(sd(fitb$t[,7]),3)

round(fitc,3)
round(sqrt(4*var(fitb$t[,1])+var.p[1]),3)
round(sqrt(4*var(fitb$t[,2])+var.p[2]),3)
round(sqrt(4*var(fitb$t[,3])+var.p[3]),3)
round(sqrt(4*var(fitb$t[,4])+var.p[4]),3)
round(sqrt(4*var(fitb$t[,5])+var.p[5]),3)
round(sqrt(4*var(fitb$t[,6])+var.p[6]),3)
round(sqrt(4*var(fitb$t[,7])+var.p[7]),3)


cdf= function(par,x) cdf_mogeweib(par, x)
data = x

parameters = par
parameters = fit$mle
parameters = results$t0
parameters = fitc
data_orderdenados = sort(data)
v = cdf(as.vector(parameters), data_orderdenados)
n = length(data)
y = qnorm(v)
u = pnorm((y - mean(y))/sqrt(var(y)))
W_temp <- vector()
A_temp <- vector()
for (i in 1:n) {
  W_temp[i] = (u[i] - (2 * i - 1)/(2 * n))^2
  A_temp[i] = (2 * i - 1) * log(u[i]) + (2 * n + 1 -
                                         2 * i) * log(1 - u[i])
}
A_2 = -n - mean(A_temp)
W_2 = sum(W_temp) + 1/(12 * n)
W_star = W_2 * (1 + 0.5/n)
A_star = A_2 * (1 + 0.75/n + 2.25/n^2)
p = length(parameters)

ks.testg = function(...) tryCatch(ks.test(...), warning = function(war) NA)
```

```
KS = ks.testg(x = jitter(data,.1), y = "cdf", par = as.vector(parameters))
round(W_star,4)
round(A_star,4)
KS
```

APÊNDICE C

Código R Utilizado na Aplicação das Fibras de Vidro

Algoritmo R utilizado na seção 5.2

```
setwd('C:/Users/Kleber/Dropbox/Tese/Tese1/Comandos/Estimação')
library(AdequacyModel)
source('pdfs.txt')
source('bootstrapping.txt')

setwd('C:/Users/Kleber/Dropbox/Tese/Tese1/Comandos/Estimação/dados')
data = x = source('GlassFibres.txt')$value

par = c(14.19,8.1,46.52,5.57,.71,2.33,1)
fitmogew = mle.boot(x, pdf_mogeweib, cdf_mogeweib,
par, 'B', 1000)
llmogew = sum(log(pdf_mogeweib(fitmogew$t0, x)))


par = c(1,1,1,1)
fitbge = mle.boot(x, pdf_bge, cdf_bge,
par, 'B', 1000)
llbge = sum(log(pdf_bge(fitbge$t0, x)))

par = c(.1,.1,.1,.1)
```

```
fitbfret = mle.boot(x, pdf_bfret, cdf_bfret,
par, 'B', 1000)
llbfret = sum(log(pdf_bfret(fitbfret$t0, x)))

setwd('C:\\Users\\Kleber\\Dropbox\\Tese\\Tese1\\Graficos')
postscript('aplicacaofibrasdevidro.eps',horizontal=FALSE)
hist(x,freq=F,main='', ylab = 'Densidade',ylim=c(0,1.73),
xlim=c(0,2.5))
curve(pdf_mogeweib(fitmogew$t0, x), add=T, lty=1, lwd = 2)
curve(pdf_bge(fitbge$t0, x),add=T,    col=1, lty=2, lwd = 2)
curve(pdf_bfret(fitbfret$t0, x),add=T,col=2, lty=1, lwd = 2)
legend('topleft',c('MOGEW', 'BGE', 'BF'),
  col=c(1,1,2),lty=c(1:2),lwd=2, cex=1.5)
dev.off()


AICmogew  = - 2*llmogew + 2*7
AICbge    = - 2*llbge + 2*4
AICbfret   = - 2*llbfret + 2*4



AICCmogew = - 2*llmogew + 2*7*63/(63-7-1)
AICCbge   = - 2*llbge + 2*4*63/(63-4-1)
AICCbfret   = - 2*llbfret + 2*4*63/(63-4-1)

BICmogew  = - 2*llmogew + 7*log(63)
BICbge    = - 2*llbge + 4*log(63)
BICbfret    = - 2*llbfret + 4*log(63)

HQICmogew  = - 2*llmogew + 2*7*log(log(63))
HQICbge    = - 2*llbge + 2*4*log(log(63))
HQICbfret    = - 2*llbfret + 2*4*log(log(63))


c(AICmogew, AICbge, AICbfret)
c(AICCmogew, AICCbge, AICCbfret)
c(BICmogew, BICbge, BICbfret)
c(HQICmogew, HQICbge, HQICbfret)
```

```
round(fitmogew$t0,4);round(fitbge$t0,4);round(fitbfret$t0,4)

round(apply(fitmogew$t,2,sd),4)
round(apply(fitbge$t,2,sd),4)
round(apply(fitbfret$t,2,sd),4)



parameters = fitmogew$t0
cdf= function(par,x) cdf_mogeweib(par, x)
parameters = fitbge$t0
cdf= function(par,x) cdf_bge(par, x)
parameters = fitbfret$t0
cdf= function(par,x) cdf_bfret(par, x)

data_orderdenados = sort(data)
v = cdf(as.vector(parameters), data_orderdenados)
n = length(data)
y = qnorm(v)
u = pnorm((y - mean(y))/sqrt(var(y)))
W_temp <- vector()
A_temp <- vector()
for (i in 1:n) {
  W_temp[i] = (u[i] - (2 * i - 1)/(2 * n))^2
  A_temp[i] = (2 * i - 1) * log(u[i]) + (2 * n + 1 -
                                          2 * i) * log(1 - u[i])
}
A_2 = -n - mean(A_temp)
W_2 = sum(W_temp) + 1/(12 * n)
W_star = W_2 * (1 + 0.5/n)
A_star = A_2 * (1 + 0.75/n + 2.25/n^2)
p = length(parameters)

ks.testg = function(...) tryCatch(ks.test(...), warning = function(war) NA)
KS = ks.testg(x = jitter(data,.1), y = "cdf", par = as.vector(parameters))
round(A_star,4)
round(W_star,4)
KS
```

APÊNDICE D

Código R Utilizado na Aplicação dos Terremotos em Fiji

Algoritmo R utilizado na seção 5.3

```
setwd('C:/Users/Kleber/Dropbox/Tese/Tese1/Comandos/Estimação')
library(AdequacyModel)
source('pdfs.txt')
source('bootstrapping.txt')

setwd('C:/Users/Kleber/Dropbox/Tese/Tese1/Comandos/Estimação/dados')
data = x = source('TerremotoFiji.txt')$value
n = length(x)

par = c(.1,.1,.1,.1,.1,.1,.1)
fitmogew = mle.boot(x, pdf_mogeweib, cdf_mogeweib,
par, 'B', 2000)
llmogew = sum(log(pdf_mogeweib(fitmogew$t0, x)))

par = c(.1,.1,.1)
fitmow = mle.boot(x, pdf_moweib, cdf_moweib,
par, 'B', 1000)
llmow = sum(log(pdf_moweib(fitmow$t0, x)))

par = c(.1,.1,.1)
```

```
fitew = mle.boot(x, pdf_eweib, cdf_eweib,
par, 'B', 1000)
llew = sum(log(pdf_eweib(fitew$t0, x)))


par = c(2,1)
fitweib = mle.boot(x, pdf_weib, cdf_weib,
par, 'B', 1000)
llweib = sum(log(pdf_weib(fitweib$t0, x)))

par = c(1,1)
fitfret = mle.boot(x, pdf_fret, cdf_fret,
par, 'B', 1000)
llfret = sum(log(pdf_fret(fitfret$t0, x)))

par = c(.1,.1)
fitbs = mle.boot(x, pdf_bs, cdf_bs,
par, 'B', 1000)
llbs = sum(log(pdf_bs(fitbs$t0, x)))

par = c(2,1)
fitgama = mle.boot(x, pdf_gama, cdf_gama,
par, 'C', 1000)
llgama = sum(log(pdf_gama(fitgama$t0, x)))

par = c(1,1)
fitlnorm = mle.boot(x, pdf_lnorm, cdf_lnorm,
par, 'B', 1000)
lllnorm = sum(log(pdf_lnorm(fitlnorm$t0, x)))


par = c(.5,.5)
fitllogis = mle.boot(x, pdf_llogis, cdf_llogis,
par, 'B', 1000)
llllogis = sum(log(pdf_llogis(fitllogis$t0, x)))

#par = c(1,1,1,1)
#fitbge = mle.boot(x, pdf_bge, cdf_bge,
#par, 'B', 1000)
```

```
#llbge = sum(log(pdf_bge(fitbge$t0, x)))

setwd('C:\\Users\\Kleber\\Dropbox\\Tese\\Tese1\\Graficos')
postscript('aplicacaoterremoto.eps',horizontal=FALSE)
hist(x,freq=F,main='', ylab = 'Densidade',ylim=c(0,1.2),
xlim=c(3.8,6.5), xlab = 'Magnitude')
curve(pdf_mogeweib(fitmogew$t0, x), add=T, lty=1, lwd = 2)
curve(pdf_moweib(fitmow$t0, x),     add=T, lty=2, lwd = 2)
curve(pdf_eweib(fitew$t0, x), add=T,  col=2, lty=1, lwd = 2)
#curve(pdf_bge(fitbge$t0, x),add=T,    col=2, lty=2, lwd = 2)
curve(pdf_bs(fitbs$t0, x),add=T,      col=2, lty=2, lwd = 2)
curve(pdf_weib(fitweib$t0, x),add=T,  col=3, lty=1, lwd = 2)
curve(pdf_fret(fitfret$t0, x),add=T,  col=3, lty=2, lwd = 2)
curve(pdf_gama(fitgama$t0, x),add=T,  col=4, lty=1, lwd = 2)
curve(pdf_lnorm(fitlnorm$t0, x),add=T,col=4, lty=3, lwd = 2)
curve(pdf_llogis(fitllogis$t0, x),add=T,col=5, lty=1, lwd = 2)
legend('topright',c('MOGEW', 'MOW', 'EW', 'B-S', 'Weibull',
'Frechet', 'gama','lognormal', 'loglogistica'),
  col=c(1,1,2,2,3,3,4,4,5,5),lty=c(1:2),lwd=2, cex=1.5)
dev.off()



AICmogew  = - 2*llmogew + 2*7
AICmow    = - 2*llmow + 2*3
AICew     = - 2*llew + 2*3
#AICbge    = - 2*llbge + 2*4
AICweib   = - 2*llweib + 2*2
AICfret   = - 2*llfret + 2*2
AICbs     = - 2*llbs + 2*2
AICgama   = - 2*llgama + 2*2
AIClnorm  = - 2*lllnorm + 2*2
AICllogis = - 2*llllogis + 2*2

AICCmogew = - 2*llmogew + 2*7*n/(n-7-1)
AICCmow   = - 2*llmow + 2*3*n/(n-3-1)
AICCew    = - 2*llew + 2*3*n/(n-3-1)
#AICCbge   = - 2*llbge + 2*4*n/(n-4-1)
AICCweib  = - 2*llweib + 2*2*n/(n-2-1)
```

```
AICCfret  = - 2*llfret + 2*2*n/(n-2-1)
AICCbs    = - 2*llbs + 2*2*n/(n-2-1)
AICCgama  = - 2*llgama + 2*2*n/(n-2-1)
AICClnorm = - 2*lllnorm + 2*2*n/(n-2-1)
AICCllogis= - 2*llllogis + 2*2*n/(n-2-1)

BICmogew  = - 2*llmogew + 7*log(n)
BICmow    = - 2*llmow + 3*log(n)
BICew     = - 2*llew + 3*log(n)
#BICbge    = - 2*llbge + 4*log(n)
BICweib   = - 2*llweib + 2*log(n)
BICfret   = - 2*llfret + 2*log(n)
BICbs     = - 2*llbs + 2*log(n)
BICgama   = - 2*llgama + 2*log(n)
BIClnorm  = - 2*lllnorm + 2*log(n)
BICllogis = - 2*llllogis + 2*log(n)

HQICmogew  = - 2*llmogew + 2*7*log(log(n))
HQICmow    = - 2*llmow + 2*3*log(log(n))
HQICew     = - 2*llew + 2*3*log(log(n))
#HQICbge    = - 2*llbge + 2*4*log(log(n))
HQICweib   = - 2*llweib + 2*2*log(log(n))
HQICfret   = - 2*llfret + 2*2*log(log(n))
HQICbs     = - 2*llbs + 2*2*log(log(n))
HQICgama   = - 2*llgama + 2*2*log(log(n))
HQIClnorm  = - 2*lllnorm + 2*2*log(log(n))
HQICllogis = - 2*llllogis + 2*2*log(log(n))

fitmogew$t0;fitmow$t0;fitew$t0;fitbs$t0;fitweib$t0;
fitfret$t0;fitgama$t0;fitlnorm$t0;fitllogis$t0

apply(fitmogew$t,2,sd);apply(fitmow$t,2,sd)
apply(fitew$t,2,sd);apply(fitbs$t,2,sd)
apply(fitweib$t,2,sd);apply(fitfret$t,2,sd)
apply(fitgama$t,2,sd);apply(fitlnorm$t,2,sd)
apply(fitllogis$t,2,sd)

c(AICmogew, AICmow, AICew, AICweib,
AICfret, AICgama, AICbs, AIClnorm, AICllogis)
```

```
c(AICCmogew, AICCmow, AICCew, AICCweib,
AICCfret, AICCgama, AICCbs, AICClnorm, AICCllogis)

c(BICmogew, BICmow, BICew, BICweib,
BICfret, BICgama, BICbs, BIClnorm, BICllogis)

c(HQICmogew, HQICmow, HQICew, HQICweib,
HQICfret, HQICgama, HQICbs, HQIClnorm, HQICllogis)
```

```
data = x

parameters = fitmogew$t0
cdf= function(par,x) cdf_mogeweib(par, x)
parameters = fitmow$t0
cdf= function(par,x) cdf_moweib(par, x)
parameters = fitew$t0
cdf= function(par,x) cdf_eweib(par, x)
parameters = fitweib$t0
cdf= function(par,x) cdf_weib(par, x)
parameters = fitfret$t0
cdf= function(par,x) cdf_fret(par, x)
parameters = fitgama$t0
cdf= function(par,x) cdf_gama(par, x)
parameters = fitlnorm$t0
cdf= function(par,x) cdf_lnorm(par, x)
parameters = fitllogis$t0
cdf= function(par,x) cdf_llogis(par, x)
parameters = fitbs$t0
cdf= function(par,x) cdf_bs(par, x)
data_orderdenados = sort(data)
v = cdf(as.vector(parameters), data_orderdenados)
n = length(data)
y = qnorm(v)
u = pnorm((y - mean(y))/sqrt(var(y)))
W_temp <- vector()
```

```
A_temp <- vector()
for (i in 1:n) {
  W_temp[i] = (u[i] - (2 * i - 1)/(2 * n))^2
  A_temp[i] = (2 * i - 1) * log(u[i]) + (2 * n + 1 -
                                          2 * i) * log(1 - u[i])
}
A_2 = -n - mean(A_temp)
W_2 = sum(W_temp) + 1/(12 * n)
W_star = W_2 * (1 + 0.5/n)
A_star = A_2 * (1 + 0.75/n + 2.25/n^2)
p = length(parameters)

ks.testg = function(...) tryCatch(ks.test(...), warning = function(war) NA)
KS = ks.testg(x = jitter(data,.1), y = "cdf", par = as.vector(parameters))
round(W_star,4)
round(A_star,4)
KS
```

APÊNDICE E

Código R Utilizado na Aplicação de Análise de Sobrevivência

Algoritmo R utilizado na seção 5.4

Código fonte das densidades transformadas

```
dweib = function (t, x, b0 = 1, b1 = 1, theta = 1)
{
   y = log(t)
   sig = 1/theta
   z <- (y - b0 - b1*x)/sig
   density <- exp(z - exp(z)) * (1/sig)
   return(density)
}

pweib = function (t, x = 1, b0 = 1, b1 = 1, theta = 1)
{
   y = log(t)
   sig = 1/theta
   z <- (y - b0 - b1*x)/sig
   cdf <- 1 - exp(-exp(z))
   return(cdf)
}
```

```
dlnorm = function (t, x, b0 = 1, b1 = 1, theta = 1)
{
   y = log(t)
   sig = 1/theta
   z <- (y - b0 - b1*x)/sig
   density <- dnorm(z)
   return(density)
}

plnorm = function (t, x = 1, b0 = 1, b1 = 1, theta = 1)
{
   y = log(t)
   sig = 1/theta
   z <- (y - b0 - b1*x)/sig
   cdf <- pnorm(z)
   return(cdf)
}
```

```
dexpweib = function (t, x, b0 = 1, b1 = 1, theta = 1, a = 1)
{
   y = log(t)
   sig = 1/theta
   z = (y - b0 - b1*x)/sig
   g = exp(z - exp(z)) * (1/sig)
   G = 1 - exp(-exp(z))
```

```
    density <- a * g * G^(a-1)
    return(density)
}

pexpweib = function (t, x = 1, b0 = 1, b1 = 1, theta = 1, a = 1)
{
    y = log(t)
    sig = 1/theta
    z <- (y - b0 - b1*x)/sig
    G = 1 - exp(-exp(z))
    cdf <- G^a
    return(cdf)
}


dmoweib = function (t, x, b0 = 1, b1 = 1, theta = 1, b = 1)
{
    y = log(t)
    sig = 1/theta
    z = (y - b0 - b1*x)/sig
    g = exp(z - exp(z)) * (1/sig)
    G = 1 - exp(-exp(z))
    density <- b * g /(G + b*(1-G))^2
    return(density)
}

pmoweib = function (t, x = 1, b0 = 1, b1 = 1, theta = 1, b = 1)
{
    y = log(t)
    sig = 1/theta
    z <- (y - b0 - b1*x)/sig
    G = 1 - exp(-exp(z))
    cdf <- G/(G + b*(1-G))
    return(cdf)
}
```

```
dmoeweib = function (t, x, b0, b1, theta, a, b, c, d, q)
{
   y = log(t)
   sig = 1/theta
   z = (y - b0 - b1*x)/sig
   g = exp(z - exp(z)) * (1/sig)
   G = 1 - exp(-exp(z))
   density <- b * q *  g * G^(-a-1) * (1-G^c)^(d-1) *
   (a + (d*c-a)*G^c)* (G^a/(G^a + b*(1-G^c)^d))^(q+1)
   return(density)
}

pmoeweib = function (t, x, b0, b1, theta, a, b, c, d, q)
{
   y = log(t)
   sig = 1/theta
   z <- (y - b0 - b1*x)/sig
   G = 1 - exp(-exp(z))
   cdf <- (G^a/(G^a + b*(1-G^c)^d))^q
   return(cdf)
}
```

Código principal

```
setwd('C:/Users/Kleber/Desktop/Estimação - Novo')
source('pdfs.txt')
pdf = dmoeweib
cdf = pmoeweib
setwd('C:/Users/Kleber/Desktop/Estimação - Novo/DadosCensurados/Amherst')
data = read.table('actg320ncc.txt',head=T)

x = data$tx
cens = 0
t = data$time/30
for(i in 1:length(x)) cens[i] = ifelse(100*t[i]-floor(100*t[i])==0, 0, 1)
```

```
#cens = c
dados = data.frame(t,cens,x)
n = length(x)

b0.b = b1.b = sig.b = a.b = b.b = c.b = d.b = q.b =0
for(i in 1:5){
aux = sample(1:n, rep=T)
dados.b = dados[aux,]

logLik <- function(par, data)
{
b0 = par[1]; b1 = par[2]; theta = par[3]; a = 1; b = 1
c = 1; d = 1; q = 1
with(data,
   sum(
       (cens)*log( pdf(t, x, b0, b1, theta, a, b, c, d, q)) +
           (1-cens)* log( 1 - cdf(t, x, b0, b1, theta, a, b, c, d, q) )
      )
   )
}



result <- optim(par = c(0, 0, 1, 1, 1, 1, 1, 1), logLik,  data = dados.b,
hessian = TRUE, method='CG',control=list(fnscale=-1))
#sqrt(diag(solve(-result$hessian)))
b0.b[i] = result$par[1]
b1.b[i] = result$par[2]
sig.b[i] = 1/result$par[3]
a.b[i] = result$par[4]
b.b[i] = result$par[5]
c.b[i] = result$par[6]
d.b[i] = result$par[7]
q.b[i] = result$par[8]
}

fit0 = NULL
fit0$b0 = mean(b0.b); fit0$sb0 = sd(b0.b)
fit0$b1 = mean(b1.b); fit0$sb1 = sd(b1.b)
fit0$sig = mean(sig.b); fit0$ssig = sd(sig.b)
```

```
fit0$a = mean(a.b); fit0$sa = sd(a.b)
fit0$b = mean(b.b); fit0$sb = sd(b.b)
fit0$c = mean(c.b); fit0$sc = sd(c.b)
fit0$d = mean(d.b); fit0$sd = sd(d.b)
fit0$q = mean(q.b); fit0$sq = sd(q.b)

attach(fit0)
ll0 = logLik(c(b0, b1, 1/sig, a, b, c, d, q), dados)


library(survival)
ekm = survfit(Surv(t,cens)~strata(x))
plot(ekm, conf.int=F, mark.time=T, xlab="Tempo (meses)",
ylab="Sobrevivência estimada", lty= c(1,2), lwd=2)
tt = seq(.1,20, l=1000)
x = 0
mu = b0 + b1*x
z = (tt/exp(mu))^(1/sig)
G = 1 - exp(-z)
S = 1 - (G^a/(G^a + b*(1-G^c)^d))^q
lines(tt, S, lty=1, col=3, lwd=2)


x = 1
mu = b0 + b1*x
z = (tt/exp(mu))^(1/sig)
G = 1 - exp(-z)
S = 1 - (G^a/(G^a + b*(1-G^c)^d))^q
lines(tt, S, lty=2, col=3, lwd=2)

detach(fit0)

##########################################################################


b0.b = b1.b = sig.b = a.b = b.b = c.b = d.b = q.b =0
for(i in 1:5){
aux = sample(1:n, rep=T)
dados.b = dados[aux,]
```

```
logLik <- function(par, data)
{
b0 = par[1]; b1 = par[2]; theta = par[3]; a = par[4]; b = 1
c = par[6]; d = par[7]; q = 1
with(data,
   sum(
       (cens)*log( pdf(t, x, b0, b1, theta, a, b, c, d, q)) +
           (1-cens)* log( 1 - cdf(t, x, b0, b1, theta, a, b, c, d, q) )
      )
   )
}

result <- optim(par = c(0, 0, 1,  1, 1, 1, 1, 1), logLik,  data = dados.b,
hessian = TRUE, method='BFGS',control=list(fnscale=-1))
#sqrt(diag(solve(-result$hessian)))
b0.b[i] = result$par[1]
b1.b[i] = result$par[2]
sig.b[i] = 1/result$par[3]
a.b[i] = result$par[4]
b.b[i] = result$par[5]
c.b[i] = result$par[6]
d.b[i] = result$par[7]
q.b[i] = result$par[8]
}

fit1 = NULL
fit1$b0 = mean(b0.b);   fit1$sb0 = sd(b0.b)
fit1$b1 = mean(b1.b);   fit1$sb1 = sd(b1.b)
fit1$sig = mean(sig.b); fit1$ssig = sd(sig.b)
fit1$a = mean(a.b);     fit1$sa = sd(a.b)
fit1$b = mean(b.b);     fit1$sb = sd(b.b)
fit1$c = mean(c.b);     fit1$sc = sd(c.b)
fit1$d = mean(d.b);     fit1$sd = sd(d.b)
fit1$q = mean(q.b);     fit1$sq = sd(q.b)

attach(fit1)

ll1 = logLik(c(b0, b1, 1/sig, a, b, c, d, q), dados)
```

```
library(survival)
x = 0
mu = b0 + b1*x
z = (tt/exp(mu))^(1/sig)
G = 1 - exp(-z)
S = 1 - (G^a/(G^a + b*(1-G^c)^d))^q
lines(tt, S, lty=1, col=2, lwd=2)

x = 1
mu = b0 + b1*x
z = (tt/exp(mu))^(1/sig)
G = 1 - exp(-z)
S = 1 - (G^a/(G^a + b*(1-G^c)^d))^q
lines(tt, S, lty=2, col=2, lwd=2)

detach(fit1)
```

```
#######################################
setwd('C:/Users/Kleber/Desktop/Estimação - Novo')
source('pdfs.txt')
pdf = dmoweib
cdf = pmoweib

setwd('C:/Users/Kleber/Desktop/Estimação - Novo/DadosCensurados/Amherst')
data = read.table('actg320ncc.txt',head=T)

x = data$tx
cens = 0
t = data$time/30
for(i in 1:length(x)) cens[i] = ifelse(100*t[i]-floor(100*t[i])==0, 0, 1)
```

```
#cens = c
dados = data.frame(t,cens,x)
n = length(x)

b0.b = b1.b = sig.b = b.b =0
for(i in 1:5){
aux = sample(1:n, rep=T)
dados.b = dados[aux,]

logLik <- function(par, data)
{
b0 = par[1]; b1 = par[2]; theta = par[3]; b = par[4]
with(data,
   sum(
       (cens)*log( pdf(t, x, b0, b1, theta, b)) +
           (1-cens)* log( 1 - cdf(t, x, b0, b1, theta, b) )
      )
   )
}


result <- optim(par = c(-0.67, 1.4, 0.54,  1.5), logLik,  data = dados.b,
hessian = TRUE, method='CG',control=list(fnscale=-1))
b0.b[i] = result$par[1]
b1.b[i] = result$par[2]
sig.b[i] = 1/result$par[3]
b.b[i] = result$par[4]
}

fit2 = NULL
fit2$b0 = mean(b0.b);   fit2$sb0 = sd(b0.b)
fit2$b1 = mean(b1.b);   fit2$sb1 = sd(b1.b)
fit2$sig = mean(sig.b); fit2$ssig = sd(sig.b)
fit2$b = mean(b.b);     fit2$sb = sd(b.b)

attach(fit2)

ll2 = logLik(c(b0, b1, 1/sig, b), dados)
```

```
library(survival)
x = 0
mu = b0 + b1*x
z = (tt/exp(mu))^(1/sig)
G = 1 - exp(-z)
S = 1 - (G/(G + b*(1-G)))
lines(tt, S, lty=1, col=4, lwd=2)

x = 1
mu = b0 + b1*x
z = (tt/exp(mu))^(1/sig)
G = 1 - exp(-z)
S = 1 - (G/(G + b*(1-G)))
lines(tt, S, lty=2, col=4, lwd=2)

detach(fit2)



###############################################################
setwd('C:/Users/Kleber/Desktop/Estimação - Novo')
source('pdfs.txt')
pdf = dexpweib
cdf = pexpweib
setwd('C:/Users/Kleber/Desktop/Estimação - Novo/DadosCensurados/Amherst')
data = read.table('actg320ncc.txt',head=T)

x = data$tx
cens = 0
t = data$time/30
for(i in 1:length(x)) cens[i] = ifelse(100*t[i]-floor(100*t[i])==0, 0, 1)
#cens = c
dados = data.frame(t,cens,x)
n = length(x)

b0.b = b1.b = sig.b = a.b  =0
for(i in 1:5){
```

```
aux = sample(1:n, rep=T)
dados.b = dados[aux,]

logLik <- function(par, data)
{
b0 = par[1]; b1 = par[2]; theta = par[3]; a = par[4]
with(data,
   sum(
       (cens)*log( pdf(t, x, b0, b1, theta, a)) +
           (1-cens)* log( 1 - cdf(t, x, b0, b1, theta, a) )
      )
   )
}


result <- optim(par = c(2.7, 0.06, 1, 1), logLik,  data = dados.b,
hessian = TRUE, method='CG',control=list(fnscale=-1))
#sqrt(diag(solve(-result$hessian)))
b0.b[i] = result$par[1]
b1.b[i] = result$par[2]
sig.b[i] = 1/result$par[3]
a.b[i] = result$par[4]
}

fit3 = NULL
fit3$b0 = mean(b0.b);   fit3$sb0 = sd(b0.b)
fit3$b1 = mean(b1.b);   fit3$sb1 = sd(b1.b)
fit3$sig = mean(sig.b); fit3$ssig = sd(sig.b)
fit3$a = mean(a.b);     fit3$sa = sd(a.b)


attach(fit3)
ll3 = logLik(c(b0, b1, 1/sig, a), dados)


library(survival)
#ekm = survfit(Surv(t,cens)~strata(x))
#plot(ekm, conf.int=F, mark.time=T, xlab="Tempo",
#ylab="Sobrevivência estimada", lty= c(1,2))
```

```
#tt = seq(.1,1100, l=1000)
x = 0
mu = b0 + b1*x
z = (tt/exp(mu))^(1/sig)
G = 1 - exp(-z)
S = 1 - (G)^a
lines(tt, S, lty=1, col=6, lwd=2)


x = 1
mu = b0 + b1*x
z = (tt/exp(mu))^(1/sig)
G = 1 - exp(-z)
S = 1 - (G)^a
lines(tt, S, lty=2, col=6, lwd=2)

detach(fit3)


legend(0.2,.46, c('Trat.        '),col= 1,lty=2, lwd=2)
legend(0.2,.35, c('MOGEW', 'MOW', 'EW', 'Weibull'),
col= c(2,4,6,3),lty=2, lwd=2)

legend(3.9,.46, c('Controle '),col= 1,lty=1, lwd=2)
legend(3.9,.35, c('MOGEW', 'MOW', 'EW', 'Weibull'),
col= c(2,4,6,3),lty=1, lwd=2)




-2*ll0 + 2*3;-2*ll1 + 2*6;-2*ll2 + 2*4; -2*ll3 + 2*4

-2*ll0 + 2*n*3/(n-3-1);-2*ll1 + 2*n*6/(n-6-1);
-2*ll2 + 2*n*4/(n-4-1);-2*ll3 + 2*n*4/(n-4-1);

-2*ll0 + 3*log(log(n));-2*ll1 + 6*log(log(n));
-2*ll2 + 4*log(log(n));-2*ll3 + 4*log(log(n))

-2*ll0 + 3*log(n);-2*ll1 + 6*log(n);-2*ll2 + 4*log(n);-2*ll3 + 4*log(n)
```

APÊNDICE F

Código R Utilizado na Aplicação dos dados de Fibras de Carbono

Algoritmo R utilizado na seção 5.5

```
setwd('C:/Users/Kleber/Dropbox/Tese/Tese1/Comandos/Estimação')
source('pdfs.txt')

n = 250000
alpha = rgamma(n,0.01,0.01)
beta = rgamma(n,0.1,0.1)

x=c(3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11,
4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.90,
3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.53, 2.67, 2.93, 3.22,
3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 2.56, 3.56,
3.15, 2.35, 2.55, 2.59, 2.38, 2.81, 2.77, 2.17, 2.83, 1.92,
1.41, 3.68, 2.97, 1.36, 0.98, 2.76, 4.91, 3.68, 1.84, 1.59,
3.19, 1.57, 0.81, 5.56, 1.73, 1.59, 2.00, 1.22, 1.12, 1.71,
2.17, 1.17, 5.08, 2.48, 1.18, 3.51, 2.17, 1.69, 1.25, 4.38,
1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.80,
1.57, 1.08, 2.03, 1.61, 2.12, 1.89, 2.88, 2.82, 2.05, 3.65)

w = 0
```

```
for(i in 1:n){
w[i] = prod(pdf_mogeweib(c(alpha[i],beta[i],1,1,1,1,1),x))
if(i%%1000==0) print(i)
}

w = w/sum(w)
#w

alpha.p = sample(alpha,rep=T, prob=w)
hist(alpha.p,breaks=20,prob=T,main='',xlab=expression(alpha),
 ylab = 'Densidade a Posteriori Marginal')
beta.p = sample(beta,rep=T, prob=w)
hist(beta.p,breaks=20,prob=T,main='',xlab=expression(beta),
 ylab = 'Densidade a Posteriori Marginal')
#plot(beta.p,type='l')



D = -2*log(pdf_mogeweib(c(alpha[i],beta[i],1,1,1,1,1),x))
D.bar = mean(D)


pD = D.bar + 2*log(pdf_mogeweib(c(mean(alpha),mean(beta),1,1,1,1,1),x))
pD = 0.5*var(D)


DIC = D.bar + pD
mean(DIC)

mean(alpha.p)
sd(alpha.p)
quantile(alpha.p,c(.025, 0.5, 0.975))

mean(beta.p)
sd(beta.p)
quantile(beta.p,c(.025, 0.5, 0.975))
```