UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Faculdade de Ciências e Tecnologia
Câmpus de Presidente Prudente

Família Distribuição Gama Exponenciada

Guilherme Aparecido Santos Aguilar

Orientador: Prof. Dr. Fernando Antônio Moala

Programa: Matemática Aplicada e Computacional

Presidente Prudente, Março de 2017

UNIVERSIDADE ESTADUAL PAULISTA

Faculdade de Ciências e Tecnologia de Presidente Prudente

Programa de Pós-Graduação em Matemática Aplicada e Computacional

Família Distribuição Gama Exponenciada

Guilherme Aparecido Santos Aguilar

Orientador: Prof. Dr. Fernando Antônio Moala

Dissertação apresentada ao Programa de Pós-Graduação em Matemática Aplicada e Computacional da Faculdade de Ciências e Tecnologia da UNESP para obtenção do título de Mestre em Matemática Aplicada e Computacional.

Presidente Prudente, Março de 2017

FICHA CATALOGRÁFICA

Santos Aguilar, Guilherme Aparecido.
S233f Família distribuição Gama Exponenciada / Guilherme Aparecido Santos Aguilar. - Presidente Prudente : [s.n], 2017
98 f.

Orientador: Fernando Antônio Moala
Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências e Tecnologia
Inclui bibliografia

1. Gama Exponenciada. 2. Teoria de distribuições. 3. Estimadores. I. Moala, Fernando Antônio. II. Universidade Estadual Paulista. Faculdade de Ciências e Tecnologia. III. Título

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Presidente Prudente

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Familia Distribuição Gama Exponenciada

AUTOR: GUILHERME APARECIDO SANTOS AGUILAR
ORIENTADOR: FERNANDO ANTONIO MOALA

Aprovado como parte das exigências para obtenção do Título de Mestre em MATEMÁTICA APLICADA E COMPUTACIONAL, pela Comissão Examinadora:

Prof. Dr. FERNANDO ANTONIO MOALA
Departamento de Estatística / Faculdade de Ciencias e Tecnologia de Presidente Prudente

Prof. Dr. SERGIO MINORU OIKAWA
Departamento de Estatística / Faculdade de Ciencias e Tecnologia de Presidente Prudente

Prof. Dr. CARLOS APARECIDO DOS SANTOS
Departamento de Estatística / Universidade Estadual de Maringá

Presidente Prudente, 06 de março de 2017

À minha esposa, Amanda,
dedico!

Agradecimentos

A realização desta dissertação seria impossível, claro, sem a orientação do professor Professor Moala, por isso agradeço grandemente toda sua ajuda durante este longo processo de aprendizado, pois teve muita paciência comigo e foi peça importantíssima não só na minha evolução em relação à programação, que era bem precária, mas também foi um grande incentivador em relação a seguir a carreira acadêmica. Agradeço também ao Fernando Pacanelli pela ajuda com o *cluster*, sendo assim possível rodar programas pesados em um espaço de tempo bem menor.

Agradeço a CAPES pela bolsa, sem este apoio financeiro seria totalmente inviável a realização de todo o mestrado. Porém, com este auxílio, foi possível me dedicar totalmente e exclusivamente ao mestrado.

À minha família, também agradeço por permitirem que eu estudasse por muito tempo e me deixando livre de várias tarefas da casa que eram divididas entre todos que lá moravam. E por entenderem que eu ficar o dia na frente do computador significa que estou trabalhando e não sem fazer nada. À minha esposa pelo apoio, por estar sempre ao meu lado, me motivar muito e me ajudar gigantescamente com o inglês.

Agradeço ainda mais a Deus que sempre atendeu minhas preces, me dando força de vontade e motivação para trabalhar arduamente neste período de muito estresse que foi o mestrado.

"É o momento crítico que revela o homem.
Portanto, quando a crise te atingir, lembra-te que Deus,
como um treinador de lutadores, deu-te um antagonista áspero e rijo.
Com que fim?, perguntarás. Para que te consagres vitorioso nos Grandes Jogos."
Anacrúsio

Resumo

Devido aos inúmeros campos para aplicações na Análise de Sobrevivência, diferentes funções de risco são necessárias para modelar os mais diversos casos em estudo. Portanto, ao criar novas distribuições pode-se obter diferentes funções de risco com suas diferentes curvas, que podem ser utilizadas para diversos tipos de dados.

Serão apresentadas três novas distribuições de probabilidade, criadas a partir de três diferentes métodos, sendo a Gama Exponenciada Estendida de Marshall Olkin, Gama Exponenciada Poisson Truncada no Zero e também a Gama Exponenciada Bivariada. Para as distribuições de probabilidade univariadas foram obtidos resultados probabilísticos, tais como o n-ésimo momento; r-ésimo momento de vida média residual; r-ésimo momento de vida média residual invertido; ordenação estocástica; entropias; desvios médios; curvas de Bonferroni e de Lorenz; assimetria, curtose e seus gráficos; estatísticas de ordem e parâmetro $stress - strength$. Em relação a distribuição Gama Exponenciada Bivariada foi encontrada sua função acumulada; função densidade; função marginal; função condicional e seu n-ésimo momento.

Para as novas distribuições univariadas encontradas, também foram feitas simulações para diferentes valores de parâmetros com o intuito de verificar qual o melhor método de estimação, para cada parâmetro de cada distribuição. Os métodos utilizados foram: estimador de máxima verossimilhança, Mínimos Quadrados, Mínimos Quadrados Ponderados, Cramér-von-Mises, Anderson Darling, Anderson Darling -RT (cauda à direita), Anderson Darling - LT (cauda à esquerda), Anderson Darling - 2LT (cauda à esquerda de segunda ordem), Kolmogorov e também foi utilizado o método Bayesiano com priori Gama. Por último foram também realizadas aplicações em um banco de dados, uma para cada distribuição univariada, onde foi comparado o ajuste das novas distribuições propostas com outras já conhecidas na literatura.

Palavras-Chave: *Gama Exponenciada, Teoria de distribuições, Estimadores.*

Abstract

Due to the many fields for applications in Survival Analysis, different hazard functions are needed to modelling the various case studies. Therefore, creating new distributions can obtains different hazard functions with different graphics, which can be used for various types of data.

There will be presented three new probability distributions, created from three different methods, the Marshall Olkin Extendet Exponentiated Gamma, Poisson Zero Truncated Exponentiated Gamma and the Bivariate Exponentiated Gamma. For such univariate probability distributions it will be obtained some probabilistics results, like n-th time, r-th moment of residual life, r-th moment of residual life inverted, stochastic ordering, entropies, mean deviation, Bonferroni and Lorenz curve, skewness, kurtosis, order statistics and stress-strength parameter. Regarding the Bivariate Gamma Exponentiated was found your acumulated and density function; marginal function; conditional function and it's n-th moment.

For the new univariate distributions found, were also made simulations for different parameter values in order to find the best estimation method for each parameter. The methods used were: maximum likelihood, ordinary least-squares, weighted least-squares, Cramér-von-Mises, Anderson Darling, Anderson Darling - RT (right-tail), Anderson Darling - LT (left-tail), Anderson Darling - 2LT (left-tail second order), Kolmogorov and bayesian estimator with the prior Gamma. Some techniques to compare the estimators were used. Finally, applications were also performed, one for each univariate distribution, where the adjustment of some proposed distributions in relation to the database was tested.

Keywords: *Exponentiated Gamma, Theory of distributions, Estimators.*

Lista de Figuras

Lista de Tabelas

Lista de Siglas

GE: Gama Exponenciada

GEEMO: Gama Exponenciada Estendida de Marshall Olkin.

GEPTZ: Gama Exponenciada Poisson Truncada no Zero.

GEBV: Gama Exponenciada Bivariada.

Sumário

Capítulo

1

Introdução

1.1 Considerações Iniciais

A Análise de Sobrevivência/Confiabilidade é uma área da estatística que se dedica ao estudo de dados de tempo de vida de um indivíduo, componente ou item. Nos últimos anos, este foi um dos ramos da estatística que mais se desenvolveu, obtendo inúmeras aplicações principalmente em medicina (COLOSIMO; GIOLO, 2006). Uma evidência quantitativa desse sucesso é o número de aplicações em diversos campos, como medicina, biologia, saúde pública, epidemiologia, engenharia, economia, estudos demográficos, entre outros. Diante disso, diversas distribuições de probabilidade têm sido propostas, a fim de se obter diferentes funções de risco, as quais podem ser utilizadas na modelagem dos mais distintos fenômenos e circunstâncias.

Os dados em Análise de Sobrevivência referem-se ao tempo até a ocorrência de um evento de interesse. Esse tempo é denominado tempo de falha e apresenta algumas características especiais. A primeira é que a variável resposta que não é negativa, geralmente, apresenta distribuição assimétrica positiva, não sendo, portanto, adequado assumir que tenha distribuição normal. A segunda característica é a presença de dados censurados, isto é, para alguns elementos em estudo não se conhece o tempo de falha exato. Sabe-se que o tempo de falha ocorreu depois ou antes do valor registrado. É comum a ocorrência de dados censurados, uma vez que nem sempre é possível verificar a ocorrência do evento de interesse para todas as observações em teste (JUNIOR, 2013).

Na Análise de Sobrevivência, as distribuições clássicas associadas à variável resposta são as distribuições exponencial, Weibull, Log-Normal, Log-Logística, e Gama Generalizada (LAWLESS, 2011), entre outras. Nos últimos anos tem crescido muito a generalização e a modificação de algumas distribuições utilizadas em Análise de Sobrevivência. Existem diferentes formas de se modificar uma distribuição de probabilidade, como método desenvolvido por Marshall e Olkin (1997), as distribuições exponenciadas apresentadas, inicialmente por Mudholkar et al. (1995), as distribuições estendidas discutidas por Barros (2008), as distribuições betas que receberam maior atenção após o trabalho de (EUGENE et al., 2002), o método de generalização Kumaraswamy apresentado por Cordeiro e Castro (2011), que é baseada na distribuição Kumaraswamy (KUMARASWAMY, 1980).

Usualmente, dados censurados acontecem porque nem sempre é possível que o evento de interesse ocorra para todas as observações em teste. Existem diversos tipos de censuras. Além disso, é comum que a variável resposta (tempo até a ocorrência do evento de interesse) esteja relacionada com covariáveis que explicam a sua variabilidade. Para estudar o efeito dessas covariáveis sobre a variável resposta deve-se utilizar um modelo de regressão apropriado para dados censurados (SILVA, 2008). Os modelos de regressão paramétricos assumem que os dados seguem uma distribuição de probabilidade, (PASCOA, 2012).

A Análise de Sobrevivência é caracterizada pelo fato de que a variável resposta é composta de dois fatores: o tempo até a ocorrência de um evento de interesse e as censuras (COLOSIMO; GIOLO, 2006). O evento em estudo é denominado falha e o tempo até a ocorrência da falha é chamado de tempo de falha, enquanto que a censura é o registro parcial do tempo de falha, devido a perda ou retirada de um elemento do estudo. Por exemplo, no experimento conduzido por Elliott et al. (2000), 50 gatos domésticos foram observados. Os tempos de falha foram considerados como o tempo (em dias) a partir do diagnóstico de insuficiência renal até a sua morte. Por outro lado, se os animais não morreram até o final do estudo ou se morreram por outras causas, os tempos são considerados tempos censurados.

A censura é dita ser do tipo I quando ocorre devido ao término do estudo após um período de tempo pré determinado, do tipo II quando ocorre devido ao término do estudo após um número de falhas fixado previamente, ou pode ser aleatória, sendo a mais comum em situações práticas.

As censuras previamente citadas são conhecidas como censura à direita, pois a falha ocorre sempre à direita do tempo registrado. Existem ainda outros mecanismos de censura, como, por exemplo, as censuras à esquerda, em que o tempo registrado é maior que o tempo de falha e censura intervalar, na qual não se sabe o tempo exato de falha, sendo que a única informação disponível é que o tempo de falha ocorreu em um certo intervalo de tempo.

Dessa forma, para uma definição de censura aleatória, considere T uma variável aleatória não negativa que representa o tempo de falha de um elemento e C uma variável aleatória, independente de T, que representa o tempo de censura associado a esse elemento. Assim, os dados observados são representados por $t = min(T, C)$ e δ o indicador de censura, dado por:

$$\delta = \begin{cases} 1, & T \leq C \\ 0, & T > C, \end{cases}$$

em que $\delta = 0$ indica censura e $\delta = 1$ indica falha. A distribuição de probabilidade da variável aleatória T pode ser especificada por meio da função densidade de probabilidade, função de sobrevivência ou função risco, sendo as três formas equivalentes.

A função densidade de probabilidade, $f(t)$, é definida por:

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t)}{\Delta t} = \frac{\partial F(t)}{\partial t},$$

em que $F(t) = P[T \leq t]$ é a função de distribuição acumulada (fda) de T. No contexto da Análise de Sobrevivência, a função $f(t)$ pode ser interpretada como a probabilidade de um indivíduo sofrer o acontecimento de interesse no intervalo de tempo Δt e possui duas propriedades:

$$\begin{cases} f(t) \geq 0, \\ \int_0^\infty f(t)dt = 1. \end{cases}$$

A função de sobrevivência, $S(t)$ é definida como sendo a probabilidade de ocorrência do acontecimento de interesse após o instante t. A função de sobrevivência é definida da seguinte forma:

$$S(t) = P(T > t) = \int_t^\infty f(x)dx,$$

em que $\lim_{t\to 0} S(t) = 1$ e $\lim_{t\to\infty} S(t) = 0$.

A função risco, $h(t)$, representa o risco instantâneo no instante t condicional à sobrevivência até o tempo t e é definida por:

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

A função risco é definida como a probabilidade de um indivíduo sofrer o evento em um intervalo de tempo entre t e $t + \Delta t$, dado que ele sobreviveu até o tempo t. Uma forma empírica de determinar o comportamento da função risco se dá por meio da construção do gráfico do tempo total em teste (curva TTT), proposto por Aarset (1987). A curva TTT é obtida construindo um gráfico a partir da quantidade

$$G\left(\frac{r}{n}\right) = \frac{\sum_{i=1}^{r} T_{i:n} + (n-r)T_{r:n}}{\sum_{i=1}^{n} T_{i:n}} \quad por \quad \frac{r}{n}$$

em que n é o tamanho da amostra, $r = 1, \ldots, n$ e $T_{i:n}, i = 1, \ldots, n$ são estatísticas de ordem da amostra (MUDHOLKAR et al., 1996).

Este gráfico apresenta uma reta diagonal se o risco for constante (reta A), uma curva convexa se a função risco é decrescente (curva B) e côncava se o risco é crescente (curva C), uma curvatura primeiramente convexa e depois côncava (curva D) se o risco é em forma de "U", e no caso reverso (curva E) é unimodal. A Figura 1.1 ilustra as diversas formas que a curva TTT pode apresentar.

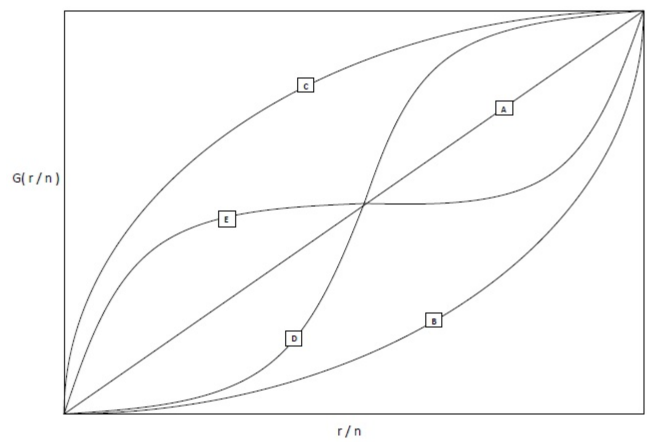


Figura 1.1: Ilustração de algumas curvas TTT.
Fonte: Pescim (2009)

Em Análise de Sobrevivência e confiabilidade a distribuição Exponencial é bastante utilizada para análise de dados e muitas vezes preferível para situações em que a taxa de falha é constante. Em caso de taxa de falha monótona as distribuições Weibull e Gama são mais frequentemente utilizadas (SINGH et al., 2015). A distribuição Gama tem uma limitação importante, que é o fato de sua função de sobrevivência e a função risco não terem formas fechadas. Em termos matemáticos, o modelo Exponencial é o modelo mais simples utilizado para descrever dados do tempo de falha, pois ele possui apenas um parâmetro e função de risco constante. Outras distribuições como a Exponencial Exponenciada e a Gama incorporam o comportamento constante para a função de risco como um caso particular (RISTIĆ; NADARAJAH, 2014). A distribuição Exponencial Generalizada, ou Exponencial Exponencializada, foi introduzida por Gupta e Kundu (1999) como alternativa para o modelo Gama ou Weibull em várias situações, principalmente na presença de censura ou dados agrupados.

Gupta et al. (1998) propuseram o uso da Distribuição Gama Exponenciada como uma alternativa, Shawky e Bakoban (2008) discutiram a aplicabilidade da distribuição Gama Exponenciadas e utilizou estimadores Bayesianos e não-Bayesianos. Além disso, Shawky e Bakoban (2009) considerou as estatísticas de ordem da distribuição Gama Exponenciadas e discutiu as várias propriedades matemáticas com base em estatísticas de ordem. Singh et al. (2011), Ghanizadeh et al. (2011), Khan e Kumar (2011), também propuseram o procedimento de estimação para a Gama Exponenciada sob dados completos e censurados.

No Capítulo 2 é apresentado o método que se usa para obter distribuições Exponenciadas e algumas de suas vantagens. Também é apresentada a distribuição base utilizada nos novos modelos propostos nesta dissertação, a distribuição Gama Exponenciada e algumas aplicações que foram realizadas com ela. No Capítulo 3 é construída uma nova distribuição, a Gama Exponenciada Estendida de Marshall Olkin, em seguida algumas de suas propriedades probabilísticas são encontradas, foi realizado um estudo de simulação comparando dez diferentes métodos de estimação e por fim foi feito uma aplicação num banco de dados, verificando o ajustamento da nova distribuição criada, em relação a outras já conhecidas. No Capítulo 4, como no Capítulo 3, também foi construída uma nova distribuição, a Gama Exponenciada Poisson Truncada no Zero, onde foram encontradas algumas de suas propriedades probabilísticas e realizado um estudo de simulação com dez diferentes estimadores, e por fim foi feito também uma aplicação num banco de dados, onde o ajustamento da nova distribuição foi comparado com demais funções já conhecidas. No Capítulo 5 foi construída a distribuição Gama Exponenciada Bivariada, onde foi encontrada sua função acumulada, função densidade, função marginal, função condicional, n-ésimo momento e o gráfico de sua densidade. Por último, no Capítulo 6, foi feito uma conclusão do trabalho realizado nesta dissertação.

Capítulo

2

Gama Exponenciada

2.1 Distribuições Exponenciadas

A classe de distribuições Exponenciadas é obtida elevando-se ao expoente θ, com $\theta > 0$, a função de distribuição acumulada de uma distribuição base dada por

$$F(x) = [G(x)]^{\theta} \tag{2.1}$$

onde $G(x)$ é a função acumulada da distribuição base e $F(x)$ é a nova distribuição Exponenciada. Consequentemente, a função densidade de probabilidade da nova distribuição é dada por $f(x) = \theta g(x) G(x)^{\theta-1}$.

O ganho dessa classe de distribuições é que o novo parâmetro θ flexibiliza a distribuição base. Dentro dessa classe de distribuições, cita-se a Weibull Exponenciada, apresentada por Mudholkar et al. (1995), que é a generalização da distribuição Weibull com parâmetros $\alpha > 0$ e $\gamma > 0$. A distribuição Weibull Exponenciada apresenta diferentes formas da função risco, como, por exemplo, formas crescente, decrescente, de banheira e unimodal, obtidas de quatro regiões do espaço paramétrico.

Gupta e Kundu (1999) observaram que esta família possui algumas características interessantes que são similares àquelas das distribuições Gama e Weibull. Dessa forma, este modelo pode ser usado como alternativa aos modelos Gama e Weibull em diversas situações. Algumas das propriedades da distribuição Exponencial Exponenciada, discutida por Gupta et al. (1998), foram estudadas por Gupta e Kundu (2001).

A distribuição Weibull Exponenciada é indicada para modelar dados em que o risco é não monótono e pode ser utilizada para testar a qualidade do ajuste da distribuição Weibull como sub modelo. Cinco conjuntos de dados sobre falha de motor de ônibus e dados de experimento clínico de câncer de cabeça e pescoço são analisados pelos autores, utilizando a distribuição Weibull Exponenciada. Gupta e Kundu (2001) propuseram a distribuição Exponencial Generalizada (que é a modificação da distribuição Exponencial com parâmetros de escala $\lambda > 0$ e locação $\alpha > 0$). Os autores também estudaram diferentes propriedades desta distribuição, que é um caso particular da distribuição Weibull Exponenciada proposta por Mudholkar et al. (1995), quando o parâmetro de locação é igual a zero.

Carrasco et al. (2008) definiram e estudaram a distribuição Weibull modificada generalizada que modela formas da função risco monótonas e não monótonas, e tem como sub modelos as distribuições Weibull, valor extremo, Weibull Exponenciada, Rayleigh generalizada e Weibull modificada, entre outras. Os autores também apresentaram duas representações dos momentos em série infinita e a função densidade de probabilidade das estatísticas de ordem foi obtida. A estimação dos parâmetros foi obtida pelo método de máxima verossimilhança e a matriz de informação observada foi apresentada. Finalmente, Alkasasbeh e Raqab (2009) estudaram a distribuição logística generalizada, definida por Balakrishnan e Leung (1988) e consideraram a estimação de máxima verossimilhança dos parâmetros, assim como outros cinco procedimentos de estimação, e compararam o desempenho destes procedimentos por meio de simulação numérica.

A distribuição Exponencial Exponenciada tem dois parâmetros (escala e forma) similares à distribuição Weibull e Gama. Portanto, essa distribuição também pode ser usada como uma possível alternativa às distribuições Weibull e Gama. Dois conjuntos de dados reais de tempo de falha foram ajustados ao modelo proposto. O primeiro conjunto de dados refere-se à resistência de rolamentos de esfera, citado em Lawless (2011), e o segundo conjunto refere-se ao tempo de falha de sistema de condicionamento de ar de um avião, citado em Linhart e Zucchini (1986), e comparados ao ajuste do modelo com os modelos Weibull e Gama. O modelo Exponencial Exponenciado teve melhor ajuste em um dos conjuntos de dados e o modelo Weibull se ajustou melhor no outro conjunto de dados. Por outro lado, Nadarajah e Kotz (2006b) discutiram algumas distribuições Exponenciada e suas extensões, e introduziram a distribuição Gama Exponenciada.

2.2 Distribuição Gama Exponenciada

A distribuição Gama Exponenciada foi proposta por Gupta et al. (1998), onde o modelo é obtido através do método $F(x) = [G(x)]^{\theta}$, sendo $G(x)$ a distribuição base, no caso em estudo, a distribuição Gama e θ (parâmetro forma) é um número real e positivo. Tal distribuição possui a flexibilidade suficiente para modelar taxas de falha monótonas e não monótonas (SHAWKY; BAKOBAN, 2011).

Shawky e Bakoban (2008) discutiram a distribuição Gama Exponenciada como um modelo importante entre os modelos de tempo de vida e seus estimadores Bayesianos e não Bayesianos do parâmetro de forma. Funções de confiabilidade e taxa de falha no caso de amostras completas e censuradas tipo II. Também as estatísticas de ordem da distribuição Gama Exponenciada e inferência associada foram discutidas por Shawky e Bakoban (2009).

Ghanizadeh et al. (2011), trabalharam com a estimação dos parâmetros da Gama Exponenciada com a presença de k *outliers*. Estes estimadores foram comparados empiricamente utilizando simulação Monte Carlo. Singh et al. (2011) propuseram estimadores de Bayes do parâmetro da distribuição Gama Exponenciada e associou a função sobrevivência com a função de perda da Entropia Geral para uma amostra censurada. Os estimadores propostos foram comparados com os correspondentes estimadores de Bayes, obtidos sobre a função de perda do erro ao quadrado e os estimadores de máxima verossimilhança através de suas simulações de riscos.

Khan e Kumar (2011) estabeleceram expressões explícitas e algumas relações de recorrência para momentos individuais e seus produtos de baixas estatísticas de ordem generalizadas da distribuição Gama Exponenciada. Feroze e Aslam (2012) introduziram diferentes prioris na análise bayesiana da distribuição Gama Exponenciada com amostras de censura do Tipo II. Recentemente, Nasiri et al. (2013) discutiram a estimativa clássica e Bayesiana dos parâmetros da distribuição Gama Exponenciada Generalizada.

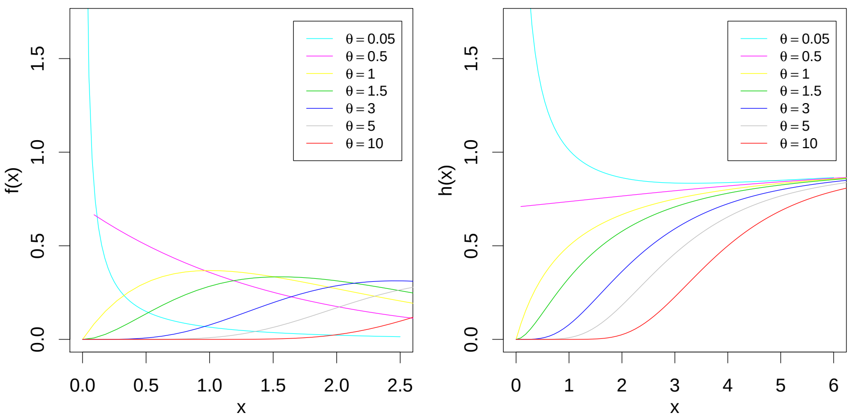
Muitos autores trabalharam com a generalização de alguma função já conhecida. Aryal e Tsokos (2009) apresentaram uma nova generalização da distribuição Weibull chamada de distribuição Weibull Transmutada. Mahmoud e Alam (2010) introduziram uma nova generalização da distribuição Exponencial Linear, a distribuição Exponencial Linear Generalizada. Estas distribuições são importantes uma vez que contêm sub modelos como casos especiais de algumas distribuições já conhecidas. Ele também fornece mais flexibilidade para analisar complexos conjuntos de dados reais. Vale ressaltar que recentemente, Lee e Tsai (2017) escreveram uma nota sobre erros ortográficos em algumas de suas equações matemáticas e uma prova incorreta do Lema 1.

As funções, densidade, acumulada e sobrevivência da Gama Exponenciada, possuem a seguinte forma, respectivamente

$$\begin{aligned} f_Y(y;\theta) &= \theta y e^{-y}(1-e^{-y}(y+1))^{\theta-1}, \\ F_Y(y;\theta) &= [1-e^{-y}(y+1)]^{\theta}, \\ S_Y(y;\theta) &= 1-[1-e^{-y}(y+1)]^{\theta}, \quad 0<y, \quad 0<\theta \end{aligned} \tag{2.2}$$

onde θ é o parâmetro forma. Vale ressaltar que quando $\theta = 1$ a distribuição Gama Exponenciada resulta numa Gama com parâmetro *forma* $\alpha = 2$ e parâmetro escala $\beta = 1$.

Os gráficos das funções densidade e risco da distribuição Gama Exponenciada são dados pela Figura (2.1), onde é possível notar que mesmo com apenas um parâmetro, a função risco da Gama Exponenciada tem uma maior flexibilidade que a função risco da distribuição Exponencial, por exemplo, que também possui apenas um parâmetro.



(a) Função Densidade (b) Função Risco

Figura 2.1: Gráfico da função densidade e risco da distribuição Gama Exponenciada.

Capítulo

3

Gama Exponenciada Estendida de Marshall Olkin

3.1 Composição entre distribuições

Utilizando o método de composição entre as distribuições Gama Exponenciada e a Geométrica, foi obtida uma nova distribuição denominada Gama Exponenciada Geométrica. Para melhor compreensão do processo de composição realizado, considere uma amostra aleatória $Y_1, Y_2, \cdots, Y_Z$ de uma distribuição Gama Exponenciada com função densidade de probabilidade (fdp) dada por (2.2). Além disso, seja Z uma variável aleatória com distribuição Geométrica, tal que Z represente o número de ensaios de Bernoulli realizados até a obtenção do primeiro fracasso, cuja função massa de probabilidade (fmp) é dada por:

$$P_Z(z;\lambda) = (1-\lambda)\lambda^{z-1}, \quad z = 1, 2, \cdots, \quad 0 \leq \lambda \leq 1. \tag{3.1}$$

Vale esclarecer que apesar das expressões (2.2) e (3.1) não serem as mesmas encontradas comumente na literatura, estas foram utilizadas em alguns trabalhos importantes (SHAWKY; BAKOBAN, 2008; ZAKERZADEH; MAHMOUDI, 2012) respectivamente, para a composição de novas distribuições utilizando o mesmo procedimento de composição.

Considerando a teoria de riscos competitivos, em que se observa o menor tempo de vida entre todos os fatores de falha, e a teoria de riscos complementares, na qual se observa o maior tempo entre os mesmos fatores, supondo que Y e Z são variáveis independentes, pode-se definir uma nova variável aleatória X como sendo o mínimo ou o máximo entre $Y_1, Y_2, \cdots, Y_Z$. A distribuição de X será dado o nome de distribuição Gama Exponeciada-Geométrica.

A seguir, esta distribuição será apresentada considerando os casos de riscos competitivos (mínimo) e riscos complementares (máximo).

3.2 Riscos Competitivos

Realizando a composição utilizando a teoria dos riscos competitivos, onde é considerado o mínimo do conjunto de variáveis aleatórias Y_i, com $i = 1, \ldots, Z$. Logo, considere $X = min(Y_1, Y_2, \cdots, Y_Z)$, supondo que Y e Z são variáveis aleatórias independentes, pode-se escrever a função densidade de probabilidade de X dado $Z = z$ como sendo:

$$f(x|Z = z, \theta) = z[1 - F_Y(x|\theta)]^{z-1} f_Y(x|\theta).$$

Daí segue que a fdp marginal de X é obtida da seguinte forma:

$$
\begin{aligned}
f(x|\theta,\lambda) &= \sum_{z=1}^{\infty} f(x|Z=z,\theta)P(Z=z) \\
&= \sum_{z=1}^{\infty} z[1-F_Y(x|\theta)]^{z-1} f_Y(x|\theta)(1-\lambda)\lambda^{z-1} \\
&= \frac{(1-\lambda)f_Y(x|\theta)}{\lambda(1-F_Y(x|\theta))} \sum_{z=1}^{\infty} z[\lambda(1-F_Y(x|\theta))]^z. \qquad (3.2)
\end{aligned}
$$

Observando a última soma obtida na expressão (3.2), temos que $|\lambda(1-F_Y(x|\theta))| < 1$, logo podemos encará-la como uma soma infinita de séries geométricas de razão $\lambda(1-F_Y(x|\theta))$, a qual converge para:

$$\sum_{z=1}^{\infty} z[\lambda(1-F_Y(x|\theta))]^z = \frac{(1-\lambda)f_Y(x|\theta)}{[1-\lambda(1-F_Y(x|\theta)]^2}. \qquad (3.3)$$

Sendo assim, a fdp marginal de $X = \min(Y_1, Y_2, ..., Y_Z)$ é dada por:

$$f(x|\theta,\lambda) = \frac{(1-\lambda)\theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1}}{\left[1-\lambda(1-(1-e^{-x}(x+1))^{\theta})\right]^2}, \quad 0<\theta, \quad 0 \le \lambda \le 1, \quad 0<x. \qquad (3.4)$$

A distribuição de $X = \min(Y_1, Y_2, ..., Y_Z)$, apresentada acima na expressão (3.4) é dita distribuição Gama Exponenciada-Geométrica (GEG_{min}).

3.3 Riscos Complementares

Para o caso dos riscos complementares, onde é considerado o máximo do conjunto de variáveis aleatórias Y_i, onde $i = 1, \ldots, Z$. Então, seja $X = \max(Y_1, Y_2, ..., Y_Z)$, supondo Y e Z variáveis aleatórias independentes, pode-se escrever a função densidade de probabilidade de X dado $Z = z$ como sendo:

$$f(x|Z=z,\theta) = z[F_Y(x|\theta)]^{z-1} f_Y(x|\theta).$$

Procedendo similarmente ao que foi feito no caso de riscos competitivos, tem-se que a fdp marginal de $X = \max(Y_1, Y_2, ..., Y_Z)$ é dada por:

$$f(x|\theta,\lambda) = \frac{(1-\lambda)\theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1}}{\left[1-\lambda(1-e^{-x}(x+1))^{\theta}\right]^2}, \quad 0<\theta, \quad 0 \le \lambda \le 1, \quad 0<x. \qquad (3.5)$$

A distribuição de $X = \max(Y_1, Y_2, ..., Y_Z)$, apresentada acima na expressão (3.5), é dita distribuição Gama Exponenciada-Geométrica (GEG_{max}).

3.4 Caso geral (Marshall Olkin)

Marshall e Olkin (1997) apresentaram um método para obter generalizações de distribuições de probabilidade assumindo um novo parâmetro, $\alpha > 0$, em uma família de distribuições (distribuições bases), dada pela equação:

$$S*(y) = \frac{\alpha S(y)}{1-\bar{\alpha}S(y)} = \frac{\alpha S(y)}{F(y)+\alpha S(y)}, \quad \alpha > 0, \qquad (3.6)$$

em que $F(y)$ e $S(y)$ são a função de distribuição acumulada e a função de sobrevivência da distribuição base respectivamente, e $S*(y)$ é a função de sobrevivência da nova família de distribuição, agora com o parâmetro adicional α, obtida através da equação (3.6). Se observa que se $\alpha = 1$ então $S*(y) = S(y)$.

Se a distribuição base tem função densidade de probabilidade $f(y)$ e função risco $h(y)$, então a nova função densidade de probabilidade, correspondente a $f*(y)$, é obtida por:

$$f*(y) = \frac{\alpha f(y)}{[1-(1-\alpha)S(y)]^2} = \frac{\alpha f(y)}{[F(y)+\alpha S(y)]^2} \tag{3.7}$$

e a função risco é dada por:

$$h*(y) = \frac{h(y)}{1-(1-\alpha)S(y)} = \frac{h(y)}{F(y)+\alpha S(y)}. \tag{3.8}$$

Marshall e Olkin (1997) modificaram a distribuição Exponencial baseados na equação (3.7) atribuindo o nome de distribuição Exponencial estendida com dois parâmetros, que pode servir como concorrente de distribuições já conhecidas de dois parâmetros como as distribuições Weibull, Gama e lognormal. A mesma modificação foi realizada com a distribuição Weibull, a qual é denominada de Weibull com três parâmetros. Marshall e Olkin também desenvolveram as versões bivariadas destas distribuições.

A distribuição Weibull com três parâmetros, obtida por (MARSHALL; OLKIN, 1997) foi estudada por Zhang e Xie (2007), que destacam uma das propriedades importantes dessa distribuição que é a sua função risco que modela forma crescente, decrescente ou inicialmente crescente, e então decrescente, e eventualmente crescente e tem como caso particular a distribuição Weibull para $\alpha = 1$. A caracterização do modelo é estudada, se baseando no gráfico de probabilidade Weibull (WPP). Os autores apresentaram um procedimento de estimação de parâmetros com base no WPP e adicionalmente desenvolveram o método de estimação por máxima verossimilhança.

Utilizando o método apresentado por Marshall e Olkin (1997), Thomas e Jose (2004) introduziram a distribuição Marshall-Olkin semi-pareto bivariada e Marshall- Olkin pareto bivariada e estudaram várias características dessas distribuições. Ghitany et al. (2007) investigaram as propriedades da inclusão de um novo parâmetro pelo método de Marshall e Olkin (1997), baseados no modelo lomax, também conhecido como distribuição Pareto de segundo tipo. Os autores mostraram que a distribuição proposta pode ser expressa como uma distribuição de mistura do modelo Exponencial. Condições simples e suficientes para o comportamento de forma das funções densidade de probabilidade e risco foram fornecidas. O método de máxima verossimilhança foi utilizado para estimar os parâmetros da distribuição.

Rao et al. (2009) desenvolveram um plano de teste de confiabilidade para não rejeição ou rejeição de muitos produtos submetidos a inspeção, em que o tempo de vida segue a distribuição Exponencial com dois parâmetros de Marshall-Olkin (PASCOA, 2012).

Seguindo Santo (2014), considerando agora que M tenha distribuição geométrica com função de probabilidade dada por $P(M = m) = (1-\lambda)^{m-1}\lambda, \quad m = 1, 2, ...$ e $0 < \lambda < 1$. As funções de sobrevivência da distribuição obtida pelo processo de composição de distribuição com estruturas latentes de ativação mínimo e máximo podem ser reescritas, respectivamente, na forma:

$$S_{\min}(y|\theta, \lambda) = \frac{\lambda S(y|\theta)}{1-\bar{\lambda} S(y|\theta)} \tag{3.9}$$

$$S_{\max}(y|\theta,\lambda) = \frac{S(y|\theta)}{1-\bar{\lambda}F(y|\theta)}, \tag{3.10}$$

onde $\bar{\lambda} = 1-\lambda$.

Ambas as funções de sobrevivência podem se igualar se for feito uma reparametrização, fazendo $\alpha = \frac{1}{\lambda} \geq 1, (0 < \lambda < 1)$. Marshall e Olkin (1997) chegaram nesse resultado quando propuseram estender distribuições através do acréscimo de um novo parâmetro de forma nas distribuições de probabilidade já existentes, a nova distribuição é conhecida como distribuição Estendida Marshall-Olkin. Este resultado foi possível utilizando o conceito de estabilidade extrema da distribuição geométrica.

Para entender a propriedade de geométrica-extrema estável deve ser lembrado que as distribuições de valores extremos limitam distribuições de extremos e, como tal, são, por vezes, aproximações úteis. Na prática, uma variável aleatória de interesse pode ser o extremo de apenas um número finito, possivelmente aleatório, o número N de variáveis aleatórias. Quando N tem uma distribuição geométrica, a variável aleatória tem uma propriedade de estabilidade particularmente agradável, não muito diferente das distribuições de valores extremos.

Suponha que N seja independente de $X_i's \sim Geo(p)$, logo, $P(N=n) = (1-p)^{n-1}p$, $(n = 1, 2, \dots)$, e seja:

$$U = \min(X_1, \dots, X_n), \quad V = \max(X_1, \dots, X_n).$$

Definição 1 *Se $F \in \Im$ implica que a distribuição $U(V)$ está em $\Im$, então $\Im$ é dito geométrica-mínimo estável(geométrica-máximo estável). Se $\Im$ é tanto geométrico-mínimo quanto geométrico-máximo estável, então $\Im$ é dito ser geométrico-extrema estável.*

Proposição 1 *A família paramétrica de distribuições na forma da função (3.6) é geométrica -extrema estável.*

A prova desta preposição e mais detalhes sobre a relação entre a distribuição geométrica e a extensão de Marshall Olkin podem ser encontrados em Marshall e Olkin (1997).

Desta forma conclui-se que o modelo obtido pelo processo de composição utilizando a distribuição Geométrica é um caso particular da extensão de Marshall Olkin. A distribuição estendida de Marshall Olkin é obtida da seguinte maneira

$$\begin{aligned} f(x|\theta,\alpha) &= \frac{\alpha f(x|\theta)}{[1-(1-\alpha)S(x|\theta)]^2} \\ &= \frac{\alpha\theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1}}{[1-(1-\alpha)(1-(1-e^{-x}(x+1))^{\theta})]^2}, \quad 0 < x, 0 < \alpha, 0 < \theta. \end{aligned} \tag{3.11}$$

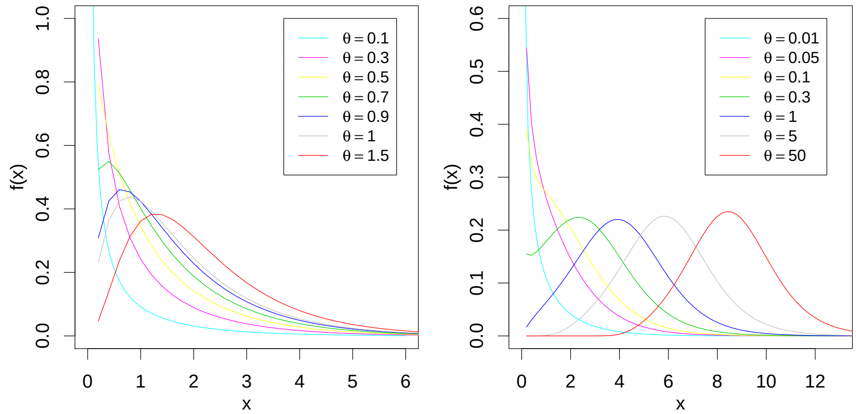
A distribuição que será utilizada a seguir para fazer os estudos probabilísticos será a Gama Exponenciada Estendida de Marshall Olkin (GEEMO).

A Figura (3.1) esboça o comportamento da função densidade quando se varia o parâmetro θ.

A Figura (3.2) esboça o comportamento da função densidade quando se varia o parâmetro α.

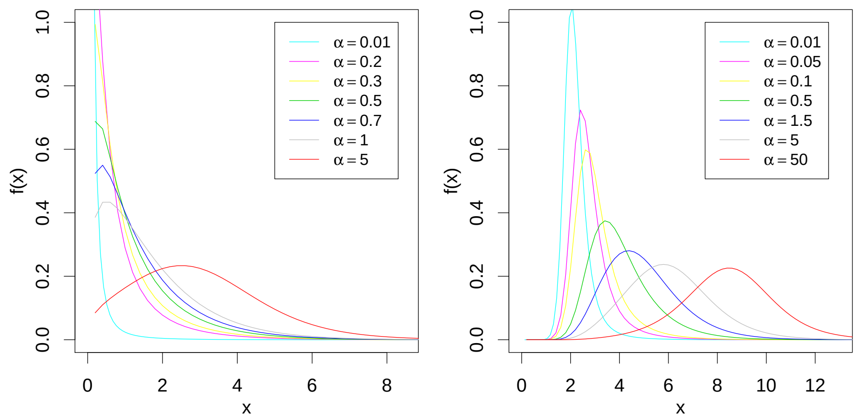
A função risco da distribuição GEEMO é dada por:

$$\begin{aligned} h(x|\theta,\alpha) &= \frac{h(x|\theta)}{1-\bar{\alpha}S(x|\theta)} \\ &= \frac{\theta\, x\mathrm{e}^{-x}\left(1-\mathrm{e}^{-x}\left(x+1\right)\right)^{\theta-1}}{\left(1-\left(1-\mathrm{e}^{-x}\left(x+1\right)\right)^{\theta}\right)\left(1-\bar{\alpha}\left(1-\left(1-\mathrm{e}^{-x}\left(x+1\right)\right)^{\theta}\right)\right)} \end{aligned} \tag{3.12}$$



(a) $\alpha = 1.5$ (b) $\alpha = 0.05$

Figura 3.1: Gráfico da função densidade para diferentes valores de θ.

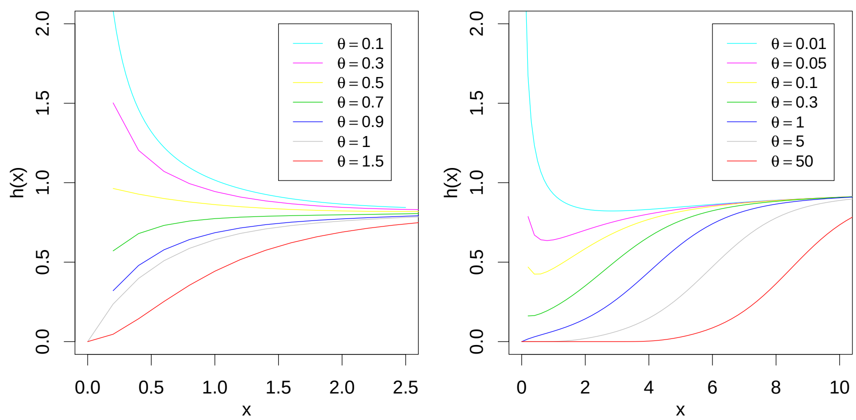


(a) $\theta = 0.5$ (b) $\theta = 5$

Figura 3.2: Gráfico da função densidade para diferentes valores de α.

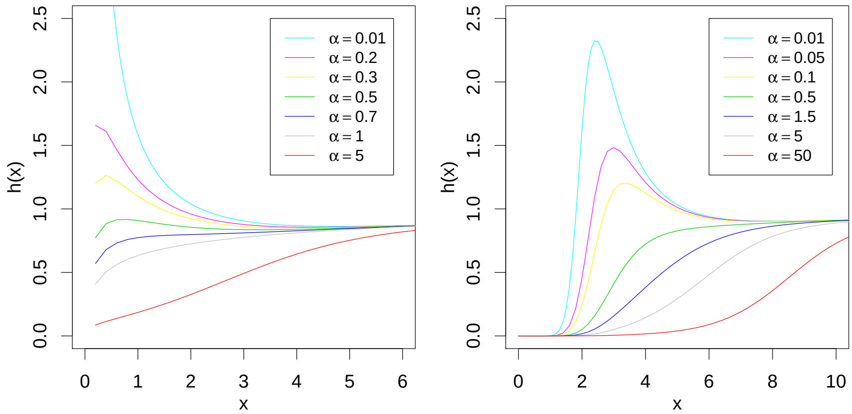
onde $\bar{\alpha} = (1 - \alpha)$.

É possível notar a partir da função (3.12), que $h(x|\theta, \alpha)$ é crescente em x para $1 \leq \alpha$ e decrescente em x para $0 < \alpha \leq 1$ como pode ser visto na Figura 3.3.



(a) $\alpha = 0.7$ (b) $\alpha = 10$

Figura 3.3: Gráfico da função risco para diferentes valores de θ.



(a) $\theta = 0.7$ (b) $\theta = 10$

Figura 3.4: Gráfico da função risco para diferentes valores de α.

A função distribuição e a função de sobrevivência da GEEMO são dadas por:

$$F(x|\theta,\alpha) = \frac{F(x|\theta)}{1-\bar{\alpha}S(x|\theta)} = \frac{(1-e^{-x}(x+1))^{\theta}}{1-\bar{\alpha}(1-(1-e^{-x}(x+1))^{\theta})} \tag{3.13}$$

$$S(y|\theta,\alpha) = \frac{\alpha S(x|\theta)}{1-\bar{\alpha}S(x|\theta)} = \frac{\alpha(1-(1-e^{-x}(x+1))^{\theta})}{1-\bar{\alpha}(1-(1-e^{-x}(x+1))^{\theta})} \tag{3.14}$$

onde $\bar{\alpha} = (1-\alpha)$.

A função sobrevivência foi plotada, como pode ser observado na Figura 3.5.

Seguindo Marshall e Olkin (1997), algumas propriedades da função risco e função de sobrevivência da distribuição GEEMO são obtidas em (3.15), mostrando os limites de tais funções em relação ao novo parâmetro adicionado:

$$\begin{aligned}
&\frac{\theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1}}{(1-(1-e^{-x}(x+1))^{\theta})} \leq h(x|\theta,\alpha) \leq \frac{\theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1}}{(1-(1-e^{-x}(x+1))^{\theta})\alpha} \quad (0<x, 0<\alpha\leq 1)\\
&\frac{\theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1}}{(1-(1-e^{-x}(x+1))^{\theta})\alpha} \leq h(x|\theta,\alpha) \leq \frac{\theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1}}{(1-(1-e^{-x}(x+1))^{\theta})} \quad (0<x, 1\leq\alpha)\\
&(1-(1-e^{-y}(y+1))^{\theta})^{1/\alpha} \leq S(x|\theta,\alpha) \leq 1-(1-e^{-y}(y+1))^{\theta} \quad (0<x, 0<\alpha\leq 1)\\
&1-(1-e^{-y}(y+1))^{\theta} \leq S(x|\theta,\alpha) \leq (1-(1-e^{-y}(y+1))^{\theta})^{1/\alpha} \quad (0<x, 1\leq\alpha)
\end{aligned} \tag{3.15}$$

3.5 Propriedades Estatísticas e Matemáticas da distribuição GEEMO

A vantagem em escrever a densidade utilizando expansões é que, desta forma ela é uma combinação linear de densidades da distribuição base. Esta combinação linear facilita encontrar as propriedades da distribuição obtida, basta ter as mesmas propriedades da distribuição base.

A função Quantil $x_q = Q(q) = F^{-1}(q)$, com $0 < q < 1$ é obtida da função (3.13), no entanto não existe forma fechada para esta distribuição, então, utilizando expansões em séries a fdp pode ser escrita de uma forma que se torna possível a obtenção de alguns resultados importantes. Esta nova forma de escrever a densidade está no teorema a seguir.

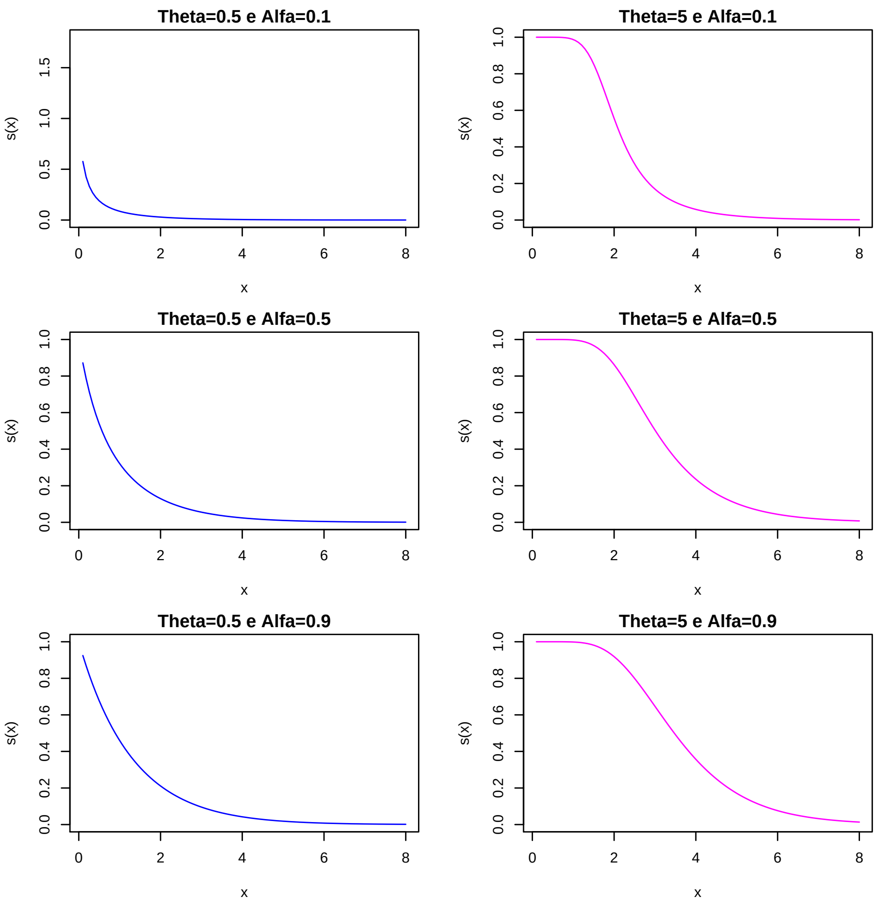


Figura 3.5: Comportamento da função de sobrevivência da GEEMO.

Teorema 1 *Seja X uma variável aleatória da distribuição GEEMO com parâmetros α e θ, então a função densidade (3.11) pode ser escrita na seguinte forma:*

$$f(x|\theta,\alpha)=\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\left(\begin{array}{c}\theta-1+\theta j\\ i\end{array}\right)\left(\begin{array}{c}i\\ k\end{array}\right)(-1)^{i+j}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^{j}\theta x^{1+k}e^{-x(1+i)}.$$

Demonstração. Baseado na expansão em série de Taylor, para $|z|<1$,

$$(1-p+pz)^{-a}=(1-p)^{-a}\sum_{k=0}^{\infty}\left(\begin{array}{c}-a\\ k\end{array}\right)\frac{(pz)^{k}}{(1-p)^{k}}$$

se obtêm

$$f(x|\theta,\alpha)=\alpha^{-2}\sum_{j=0}^{\infty}\left(\begin{array}{c}-2\\ j\end{array}\right)\left(\frac{1-\alpha}{\alpha}\right)^{j}\alpha\theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1+\theta j}.$$

Em seguida, utilizando expansão em série, com $|z|<1$,

$$(1-z)^{k-1}=\sum_{m=0}^{\infty}(-1)^{m}\left(\begin{array}{c}k-1\\ m\end{array}\right)z^{m}$$

a função é escrita como

$$f(x|\theta,\alpha) = \alpha^{-2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\left(\begin{array}{c}-2\\ j\end{array}\right)\left(\frac{1-\alpha}{\alpha}\right)^{j}\left(\begin{array}{c}\theta-1+\theta j\\ i\end{array}\right)(-1)^{i}\alpha\theta x e^{-x-xi}(x+1)^{i}$$

logo, utilizando a expansão binomial e também a transformação

$$\left(\begin{array}{c}-2\\ j\end{array}\right) = (-1)^{j}\left(\begin{array}{c}2+j-1\\ j\end{array}\right) = (-1)^{j}(j+1)$$

então a função pode ser reescrita da seguinte forma

$$f(x|\theta,\alpha) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\left(\begin{array}{c}\theta-1+\theta j\\ i\end{array}\right)\left(\begin{array}{c}i\\ k\end{array}\right)(-1)^{i+j}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^{j}\theta x^{1+k}e^{-x(1+i)}.$$

Esta forma de escrever a função densidade facilita a obtenção de suas integrais, para cálculos como o n-ésimo momento, a função geradora de momentos entre outros. ■

3.5.1 N-ésimo momento

Em Matemática, um momento é uma medida quantitativa específica, em forma de conjunto ou de pontos. Se os pontos representam a densidade de probabilidade, então o momento zero é a probabilidade total, ou seja, um.

Em Estatística, a expressão genérica de esperança, o n-ésimo momento ou momento de ordem n de uma variável aleatória X é dado por $E(x^n)$. Os momentos são muito importantes em Estatística para caracterizar distribuições de probabilidade. Por exemplo, a distribuição normal é caracterizada apenas pelo primeiro e pelo segundo momentos. Os primeiro, segundo, terceiro e quarto momentos caracterizam a tendência central, dispersão, assimetria e curtose, respectivamente, de uma distribuição de probabilidade. Mais conteúdo sobre os momentos pode ser encontrado em Casella e Berger (2002).

Os momentos mais importantes são os quatro primeiros, que são muito utilizados para caracterizar funções densidades. Para se obter o n-ésimo momento contínuo é feito

$$E(x^n) = \int_{0}^{\infty} x^n f(x)dx$$

agora, utilizando o Teorema 1 para calcular o n-ésimo momento

$$E(x^n) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\left(\begin{array}{c}\theta-1+\theta j\\ i\end{array}\right)\left(\begin{array}{c}i\\ k\end{array}\right)(-1)^{i+j}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^{j}\theta \times (1+i)^{-2-k-n}\Gamma(2+k+n). \quad (3.16)$$

3.5.2 Função Geradora de Momentos

Em teoria da probabilidade e Estatística, a função geradora de momentos (fgm) de uma variável aleatória é uma especificação alternativa de sua distribuição de probabilidade. Assim, ela fornece a base de uma forma alternativa para os resultados analíticos em comparação com resultados diretamente obtidos com funções densidade de probabilidade. Existem resultados particularmente simples para as funções de geração de momento de distribuições definidas pelas somas ponderadas de variáveis aleatórias.

Seja X uma variável aleatória. A função geradora de momentos da variável X é definida como sendo a função:

$$M_X(t) = E(e^{tX})$$

desde que $E(e^{tX})$ exista em algum intervalo do tipo $(-h, h)$ para algum número real $h > 0$.

A fgm possui esse nome pois, a partir dela, podemos encontrar todos os momentos da variável aleatória X (quando estes existem). A função só precisa estar definida em uma vizinhança do ponto zero, pois os momentos serão obtidos através de sucessivas diferenciações aplicadas em zero (CURTISS, 1942).

Por meio da expansão em série da fgm, esta também pode ser escrita da seguinte maneira:

$$\begin{aligned} M_X(t) = E(e^{tX}) = E\left(\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right) &= \sum_{n=0}^{\infty} \frac{t^n E(X^n)}{n!} \\ &= \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \sum_{k=0}^{i} \binom{\theta - 1 + \theta j}{i} \binom{i}{k} (-1)^{i+j} (j+1) \frac{t^n}{\alpha n!} \left(\frac{1-\alpha}{\alpha}\right)^j \theta \\ &\times (1+i)^{-2-k-n} \Gamma(2+k+n). \end{aligned} \tag{3.17}$$

3.5.3 Assimetria e Curtose

A tarefa fundamental em muitas análises Estatísticas é caracterizar a localização e variabilidade de um conjunto de dados. A caracterização adicional dos dados inclui assimetria e curtose. Medidas de assimetria e curtose são muitas vezes utilizados para descrever características da forma de uma distribuição. Elas também têm sido utilizadas em testes de normalidade e em estudos de robustez em teoria de procedimentos normais, como por exemplo, em Wilcox (1990).

Assimetria

A assimetria é uma medida de simetria, ou mais precisamente, a falta de simetria. Uma distribuição, ou conjunto de dados, é simétrico se os lados, tanto o direito quanto o esquerdo do ponto central forem iguais. O valor da assimetria pode ser positiva, negativa, ou ainda indefinida. A fórmula abaixo da assimetria é conhecida como o coeficiente de assimetria *Fisher-Pearson*.

$$\gamma = E\left(\frac{(x-\mu)^3}{\sigma^3}\right) = \frac{E(X^3) - 3E(X)E(X^2) + 2E^3(X)}{Var^{3/2}(X)}. \tag{3.18}$$

Curtose

A curtose é uma medida de dispersão que caracteriza o pico ou "achatamento"da curva da função de distribuição de probabilidade. Conjuntos de dados com alta curtose tendem a ter caudas pesadas, ou *outliers*. Conjuntos de dados com baixa curtose tendem a ter caudas leves, ou a falta de *outliers*. A medida padrão da curtose, baseia-se numa versão em escala reduzida do quarto momento dos dados ou população (PEARSON, 1905). É normalmente definida como:

$$\kappa = \frac{\mu^4}{\sigma^4} = \frac{E(X^4) - 4E(X)E(X^3) + 6E(X^2)(E(X))^2 - 3(E(X))^4}{(Var(X))^2}. \tag{3.19}$$

A tabela 3.1 mostra que a variância da função GEEMO é uma função crescente enquanto a assimetria e a curtose são funções decrescentes em relação aos parâmetros.

Tabela 3.1: Tabelas Assimetria e Curtose

(a) θ=0.7

α	Assimetria	Curtose	Variância
0.01	9.10088	138.487	0.07938
0.1	3.72683	24.26932	0.49945
0.5	2.05672	9.22583	1.33449
1	1.58127	6.63649	1.84144
1.5	1.34838	5.63901	2.15788
2	1.20003	5.09721	2.38338
3	1.01189	4.51639	2.69327
5	0.80563	4.01996	3.05493
10	0.57318	3.64666	3.46661

(b) α=0.5

θ	Assimetria	Curtose	Variância
0.01	16.26309	367.58970	0.04168
0.1	5.08144	38.82985	0.37220
0.5	2.34967	10.96594	1.15529
1	1.82026	8.00707	1.49790
1.5	1.62669	7.13877	1.63981
2	1.52789	6.74362	1.71128
3	1.42992	6.38790	1.77566
5	1.35587	6.14920	1.81032
10	1.30954	6.02561	1.80681

A Figura (3.6) mostra o ocorrido na tabela, onde conforme os valores dos parâmetros vão diminuindo a assimetria e curtose vão aumentando.

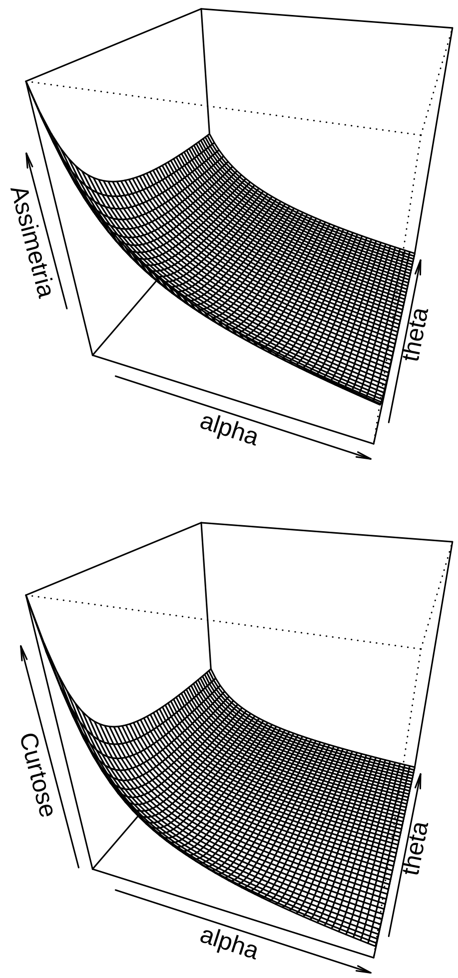


Figura 3.6: Gráfico da Assimetria e Curtose.

3.5.4 R-ésimo momento da Vida Média Residual

Dado que um componente sobrevive até ao tempo $t > 0$, a vida residual é o período além de t até o momento da falha e é definido pela esperança condicional de $X|X > t$. Portanto o r-ésimo momento da vida residual média é dado pela seguinte fórmula geral:

$$\mu_r(t) = E\left((X-t)^r|X>t\right) = \frac{\int\limits_t^\infty (x-t)^r f(x)dx}{S(t)}. \tag{3.20}$$

Utilizando o resultado do teorema 1 é obtido:

$$\mu_r(t) = \frac{1}{S(t)}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\sum_{m=0}^{r}\binom{r}{m}\binom{\theta-1+\theta j}{i}\binom{i}{k}(-1)^{i+j+m}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^j$$
$$\times\,\theta t^m\int\limits_t^\infty x^{1+k+r-m}e^{-x(1+i)}dx. \tag{3.21}$$

Portanto é obtido o r-ésimo momento da vida média residual:

$$\mu_r(t) = \frac{1}{S(t)}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\sum_{m=0}^{r}\binom{r}{m}\binom{\theta-1+\theta j}{i}\binom{i}{k}(-1)^{i+j+m}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^j$$
$$\times\,\theta t^m(1+i)^{-2-k-r+m}\Gamma(2+k+r-m,t(i+1)),\quad r\geq 1, \tag{3.22}$$

onde $\Gamma(s;t) = \int\limits_t^\infty x^{s-1}e^{-x}dx$ é a função Gama incompleta superior.

A vida média residual é um caso particular do r-ésimo momento da vida média residual, isso ocorre quando $r = 1$, também conhecida como *MRL* do inglês, *Mean Residual Life*, é a vida remanescente (X-t) esperada para a unidade, dado que no tempo t a unidade estava operante. Teorias e aplicações que utilizam *MRL* se estendem por vasto campos. Os métodos diferem consideravelmente de uma aplicação para outra. Testes acelerados, conjunto de modelagem difusa, misturas, avaliação de seguro da esperança de vida humana, manutenção e substituição de pontes, a substituição de componentes de segurança significativas em usinas de energia, e avaliação de sinais de degradação em sistemas são apenas alguns exemplos de aplicações de função *MRL*, (STEELE et al., 2011).

$$\mu(t) = E(X-t|X>t) = \frac{\int\limits_t^\infty S(x)dx}{S(t)}. \tag{3.23}$$

Utilizando a função (3.22) e igualando $r = 1$ é obtido:

$$\mu(t) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\frac{1}{S(t)}\binom{\theta-1+\theta j}{i}\binom{i}{k}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^j\theta$$
$$\times\left((-1)^{i+j+1}(i+1)^{-2-k}t\Gamma(2+k,t(i+1)) + (-1)^{i+j}(i+1)^{-3-k}\Gamma(3+k,t(i+1))\right).$$

É possível observar também que, igualando $t = 0$, obtém-se a média:

$$\mu(0) = E(X) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\binom{\theta-1+\theta j}{i}\binom{i}{k}(-1)^{i+j}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^j\theta$$
$$\times\,(1+i)^{-3-k}\Gamma(3+k). \tag{3.24}$$

De acordo com Colosimo e Giolo (2006) a *MRL* é definida condicional a um certo tempo de vida x. Ou seja, para indivíduos com idade x esta quantidade mede o tempo médio restante de vida e é, então, a área sob a curva de sobrevivência à direita do tempo x dividida por $S(x)$, onde a função obtida é (3.23). Em termos de confiabilidade, é conhecido que a função de vida média residual e a razão de dois momentos consecutivos de vida residual determinam a distribuição exclusivamente (GUPTA; GUPTA, 1983).

3.5.5 R-ésimo momento da Vida Média Residual Invertido

O tempo decorrido desde o fracasso de um item na condição de que essa falha tenha ocorrido em $[0, t]$ é conhecido na literatura como a vida média residual invertida. Também é referido pelo tempo médio de inatividade. O r-ésimo momento residual invertido é dado por:

$$m_r(t) = E\left((t-X)^r | X \leq t\right) = \frac{\int_0^t (t-x)^r f(x)dx}{F(t)}.$$

Recorrendo a conceitos similares aos utilizados em 3.22 é obtido a seguinte função do r-ésimo momento da vida média residual invertido:

$$\begin{aligned} m_r(t) &= \frac{1}{F(t)} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \sum_{k=0}^{i} \sum_{m=0}^{r} \binom{r}{m} \binom{\theta - 1 + \theta j}{i} \binom{i}{k} (-1)^{i+j-m+r} (j+1) \frac{1}{\alpha} \theta t^m \\ &\times \left(\frac{1-\alpha}{\alpha}\right)^j (i+1)^{-2-k-r+m} \gamma(2+k+r-m, t(i+1)), \quad r \geq 1, \end{aligned} \tag{3.25}$$

onde $\gamma(s;t) = \int_0^t x^{s-1} e^{-x} dx$ é a função Gama incompleta inferior.

Em uma situação real, onde os sistemas muitas vezes não são monitorados continuamente, pode existir o interesse em obter inferência sobre o histórico do sistema, por exemplo, quando os componentes individuais falharam. Suponha agora que um componente com vida útil X, que tenha falhado no momento ou antes do tempo $t, t \geq 0$. Considere a variável aleatória condicional $t - X | X \leq t$. Esta variável aleatória condicional mostra, de fato, o tempo decorrido desde a falha do componente dado que seu tempo de vida é menor ou igual a t. Esta variável aleatória pode também ser chamada de tempo de inatividade (ou tempo até a falha), para mais detalhes pode-se ver Nanda et al. (2003), Kundu e Nanda (2010). Assim, o tempo de vida médio passado *MPL*, (*mean past lifetime*) é um caso particular para o r-ésimo momento da vida média residual invertido com $r = 1$, e é dado pela seguinte expressão:

$$\begin{aligned} m(t) &= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \sum_{k=0}^{i} \frac{1}{F(t)} \binom{\theta - 1 + \theta j}{i} \binom{i}{k} (j+1) \frac{1}{\alpha} \left(\frac{1-\alpha}{\alpha}\right)^j \theta \\ &\times \left((-1)^{i+j+1} (i+1)^{-3-k} \gamma(3+k, t(i+1)) + (-1)^{i+j} t (i+1)^{-2-k} \gamma(2+k, t(i+1))\right). \end{aligned} \tag{3.26}$$

3.5.6 Ordenação Estocástica

A forma mais simples de confrontar duas variáveis aleatórias é comparando os respectivos valores esperados. No entanto, esta comparação é pouco informativa já que se baseia apenas em dois números. Para além disso, na maior parte das situações dispõe-se de informação muito mais detalhada acerca do comportamento das variáveis aleatórias, como é o caso das suas funções de distribuição, transformadas de Laplace, funções geradoras de momentos, funções taxa de falha e outros funcionais. A comparação destas características das variáveis aleatórias resulta no estabelecimento de diferentes relações de ordem estocástica entre essas mesmas variáveis aleatórias, de longe mais informativas que a mera comparação dos seus valores esperados (MORAIS; PACHECO, 1997).

O método mais simples e popular de comparar as magnitudes de duas variáveis aleatórias é através de suas médias e medianas. Pode acontecer que, em alguns casos, a média de X é maior do que a de Y, enquanto a mediana de X é menor do que a média de Y. No entanto, esta confusão não irá surgir se as variáveis aleatórias são ordenadas estocasticamente. Da mesma forma, o mesmo pode ocorrer se for comparado a variabilidade de X com a de Y com base apenas nas medidas numéricas como o desvio padrão, e assim por diante. Além disso, estas características de distribuições pode não existir, em alguns casos. Estes métodos são muito mais informativos do que os baseados apenas em algumas características numéricas das distribuições. As comparações de variáveis aleatórias com base em tais funções normalmente estabelecem ordens parciais entre elas e são chamadas de ordens estocásticas (KOCHAR, 2012).

Ordenação das distribuições, particularmente entre distribuições que modelam tempo de vida, desempenham um papel importante na literatura Estatística. Foram consideradas seis ordens estocásticas diferentes, a usual, a taxa de risco, a taxa de risco invertida, a média, a vida média residual, e a ordem da razão de verossimilhança para duas variáveis aleatórias GEEMO independentes.

Se X e Y são variáveis aleatórias independentes com funções distribuições acumuladas (fda) F_X e F_Y respectivamente, então X é dito ser menor que Y em:

- ordem estocástica ($X \leq_{st} Y$) se $F_X(x) \geq F_Y(x)$ $\forall x$;
- ordem da taxa de risco ($X \leq_{hr} Y$) se $h_X(x) \geq h_Y(x)$ $\forall x$;
- ordem da taxa de risco reversa ($X \leq_{rh} Y$) se $P(t - X > x | X \leq t) \geq P(t - Y > y | Y \leq t), \forall x \geq 0, \quad \forall t$;
- ordem da vida média residual ($X \leq_{mrl} Y$) se $\mu_X(x) \leq \mu_Y(x)$ $\forall x$;
- ordem da razão da verossimilhança ($X \leq_{lr} Y$) se $\frac{f_X(x)}{f_Y(x)}$ decresce em x.

Seguindo Kochar (2012), tem-se a seguinte cadeia de implicações entre as ordens estocásticas:

$$\begin{array}{ccccc} X \leq_{lr} Y & \Rightarrow & X \leq_{hr} Y & \Rightarrow & X \leq_{mrl} Y \\ \Downarrow & & & & \Downarrow \\ X \leq_{rh} Y & \Rightarrow & X \leq_{st} Y & \Rightarrow & E[X] \leq E[Y]. \end{array} \tag{3.27}$$

Teorema 2 *Seja* $X \sim GEEMO(\alpha_1, \theta_1)$ *e* $Y \sim GEEMO(\alpha_2, \theta_2)$*. Se* $\theta_1 = \theta_2 = \theta$ *e* $\alpha_2 \geq \alpha_1$*, então* $(X \leq_{st} Y)$*,* $(X \leq_{rh} Y)$*,* $(X \leq_{hr} Y)$*,* $(X \leq_{mrl} Y)$*,* $(X \leq_{lr} Y)$ *e* $E[X] \leq E[Y]$*.*

Demonstração. Dada a a razão da verossimilhança

$$\frac{f_X(x)}{f_Y(x)}$$

logo,

$$\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} \leq 0$$

para $\theta_1 = \theta_2 = \theta$ e $\alpha_2 > \alpha_1$, implicando que $(X \leq_{st} Y)$, então $(X \leq_{hr} Y)$, $(X \leq_{rh} Y)$, $(X \leq_{mrl} Y)$, $(X \leq_{lr} Y)$ e $E[X] \leq E[Y]$. (Detalhes das expressões deste teorema estão no apêndice A). ■

3.5.7 Entropia

A entropia de uma variável aleatória mede a variação da incerteza. Um grande valor de entropia indica a maior incerteza nos dados. Algumas medidas de entropia populares são entropia de Rényi (RÉNYI, 1961) e entropia de Shannon (SHANNON, 1951).

Entropia de Rényi

A entropia de Rényi é uma medida de variação de incerteza que tem sido usada em aplicações e caracterizações de muitas distribuições de probabilidade. Uma expressão para esta medida é deduzida. Para a função de densidade $f(x)$, a entropia de Rényi é definida por

$$H_I(\delta) = \frac{1}{1-\delta} \ln \left(\int_0^\infty f^\delta(x) dx \right), \quad \delta > 0, \quad \delta \neq 1 \tag{3.28}$$

Utilizando as mesmas expansões em série do Teorema 1 foi obtido a seguinte expressão:

$$H_I(\delta) = \frac{1}{1-\delta} \ln(\sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \sum_{k=0}^{i} \binom{i}{k} \binom{\delta(\theta-1)+\theta j}{i} \binom{-2\delta}{j} (-1)^i \theta^\delta$$
$$\times \left(\frac{1-\alpha}{\alpha}\right)^j \frac{1}{\alpha^\delta} (\delta+i)^{-1-k-\delta} \Gamma(\delta+k+i)), \quad \delta > 0, \quad \delta \neq 1. \tag{3.29}$$

Entropia de Shannon

O conceito de entropia de Shannon refere-se à incerteza de uma distribuição de probabilidade e ela destina-se a quantificar essa incerteza. A entropia de Shannon chama atenção para o fato de que a entropia H não é função da variável aleatória X, mas sim da distribuição de probabilidade dessa variável. Em outras palavras, não depende dos valores que X assume, mas das suas probabilidades (ARTUSO, 2012). Se X é uma variável aleatória contínua não-negativa com pdf $f(x)$, então a entropia de Shannon é definido como:

$$H(f) = E[-\ln f(x)] = -\int_0^\infty f(x) \ln(f(x)) dx. \tag{3.30}$$

De acordo com Al-Zahrani et al. (2015), a entropia de Shannon é:

$$\lim_{\delta \to 1} H_\delta(x) = H(f)$$

Como neste limite chega-se à indeterminação $\frac{0}{0}$, é usado a Regra de l'Hôpital, o resultado deste limite foi encontrado com o auxílio do *software Maple*, e a expressão obtida em (3.30) se encontra no apêndice A.

3.5.8 Desvios Médios

A quantidade de dispersão em uma população pode ser medida pela totalidade dos valores absolutos dos desvios em relação à média (no caso de uma distribuição simétrica) ou em relação à mediana (no caso de uma distribuição assimétrica) (PASCOA, 2012). Se T é uma variável aleatória com distribuição GEEMO com média $\mu = E(T)$ e mediana m, então o desvio médio em relação à média e o desvio médio em relação à mediana são definidos, respectivamente por

$$\vartheta_1 = \int_0^\infty |t-\mu| f(t)dt \qquad e \qquad \vartheta_2 = \int_0^\infty |t-m| f(t)dt. \tag{3.31}$$

Os desvios médios podem ser simplificados como (mais detalhes sobre a simplificação podem ser encontrados em Nadarajah e Kotz (2006a))

$$\vartheta_1 = 2\mu F(\mu) - 2I(\mu) \qquad e \qquad \vartheta_2 = \mu + 2mF(m) - m - 2I(m), \tag{3.32}$$

onde $F(\mu)$ e $F(m)$ são obtidos por meio da expressão 3.13 e a função $I(m)$ é obtida por:

$$I(m) = \int_0^m t f(t)dt.$$

Considerando $f(t)$ como a fdp definida em 3.11, tem-se:

$$\begin{aligned} I(m) = {} & \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i} \binom{\theta-1+\theta j}{i}\binom{i}{k}(-1)^{i+j}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^{j}\theta \\ & \times (i+1)^{-3-k}\gamma(3+k, m(1+i)). \end{aligned} \tag{3.33}$$

O resultado é análogo para $I(\mu)$.

3.5.9 Curva de Bonferroni e de Lorenz

Curvas de Bonferroni e Lorenz foram propostas por Bonferroni (1930). São importantes aplicações dos desvios médios, estas curvas têm aplicações não só em economia para estudar renda e pobreza, mas também em outros campos, como confiabilidade, demografia, seguro e medicina. A curva de Bonferroni pode ser expressa por

$$B(p) = \frac{1}{p\mu}\int_0^q x f(x)dx, \tag{3.34}$$

aplicando na distribuição GEEMO

$$\begin{aligned} B(p) = {} & \frac{1}{p\mu}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i} \binom{\theta-1+\theta j}{i}\binom{i}{k}(-1)^{i+j}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^{j}\theta \\ & \times (i+1)^{-3-k}\gamma(3+k, q(1+i)), \end{aligned} \tag{3.35}$$

onde $\mu = E(X)$ e $q = F^{-1}(p)$.

A curva de Lorenz é uma representação gráfica da função de distribuição cumulativa empírica da distribuição de probabilidade de riqueza. Em tal uso, muitos economistas consideram que é uma medida de desigualdade social para representar a desigualdade da distribuição da riqueza. As curvas de Lorenz são calculadas por:

$$L(p) = \frac{1}{\mu}\int_0^q xf(x)dx, \tag{3.36}$$

aplicando na distribuição GEEMO:

$$L(p) = \frac{1}{\mu}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\binom{\theta-1+\theta j}{i}\binom{i}{k}(-1)^{i+j}(j+1)\frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)^{j}\theta$$
$$\times\,(i+1)^{-3-k}\gamma(3+k, q(1+i)). \tag{3.37}$$

3.5.10 Estatística de Ordem

As Estatísticas de Ordem, assim como os momentos amostrais, desempenham um importante papel na Inferência Estatística. Momentos de Estatísticas de ordem desempenham um papel importante nos testes de controle de qualidade e confiabilidade, para prever o fracasso de itens futuros com base nos tempos de algumas falhas iniciais.

Sabe-se que $X_{(1)} \leq \cdots \leq X_{(n)}$ denota a ordem Estatística da amostra aleatória $X_1 \leq \cdots \leq X_n$ da população contínua com fda $F(x)$ e fdp $f(x)$ então a fdp $X_{(j)}$ será dada por:

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!}f(x)(F(x))^{j-1}(1-F(x))^{n-j},$$

para $j = 1, \ldots, n$. Então a fdp da j-ésima Estatística de ordem da distribuição GEEMO é dada por

$$f_{X_{(j)}}(x) = \sum_{i,k,m=0}^{\infty}\sum_{h=0}^{m}(-1)^{i+m}\binom{n-j}{i}\binom{i+j+1}{k}\binom{\theta(i+j+k)-1}{m}\binom{m}{h}$$
$$\times\,\frac{n!}{(j-1)!(n-j)!}\times\frac{\theta}{\alpha^{i+j}}\left(\frac{1-\alpha}{\alpha}\right)^{k}x^{1+h}e^{-x(1+m)}. \tag{3.38}$$

O t-ésimo momento de $X_{j:n}$ pode ser expresso por

$$E[X_{j:n}^{t}] = \sum_{i,k,m=0}^{\infty}\sum_{h=0}^{m}(-1)^{i+m}\binom{n-j}{i}\binom{i+j+1}{k}\binom{\theta(i+j+k)-1}{m}\binom{m}{h}$$
$$\times\,\frac{n!}{(j-1)!(n-j)!}\times\frac{\theta}{\alpha^{i+j}}\left(\frac{1-\alpha}{\alpha}\right)^{k}(1+m)^{-2-h-t}\Gamma(2+h+t). \tag{3.39}$$

Uma aplicação dos primeiros momentos de Estatísticas de ordem pode ser considerado no cálculo dos L-momentos que são na verdade as combinações lineares das Estatísticas de ordem esperadas.

3.5.11 Confiabilidade

Trabalhos que envolvem modelos de resistência-tensão possuem o interesse na estimação da confiabilidade denotada por $R = P(X_2 < X_1)$, em que X_1 e X_2 são variáveis aleatórias independentes que pertencem à mesma família univariada de distribuições. A forma algébrica para confiabilidade tem sido elaborada para a maioria das distribuições padrões conhecidas. Considerando X_1 e X_2 variáveis aleatórias independentes da distribuição GEEMO com parâmetros (α_1, θ_1) e (α_2, θ_2) respectivamente, então

$$\begin{aligned} R = P(X_2 < X_1) &= \int_0^{\infty} f_1(x) F_2(x) dx \\ &= \int_0^{\infty} \frac{\alpha_1 \theta_1 x e^{-x} (1 - e^{-x}(x+1))^{\theta_1 - 1}}{\left[1 - (1-\alpha_1)(1 - (1 - e^{-x}(x+1))^{\theta_1})\right]^2} \times \frac{(1 - e^{-x}(x+1))^{\theta_2}}{1 - \bar{\alpha}_2 (1 - (1 - e^{-x}(x+1))^{\theta_2})} dx. \end{aligned}$$

Na modelagem, R é a medida de confiabilidade do componente quando ele é submetido a uma tensão aleatória X_2 e com força X_1. Pode ser mencionado que R é de maior interesse do que apenas a confiabilidade, uma vez que fornece uma medida geral da diferença entre duas populações e tem aplicações em muitas áreas. Por exemplo, se X_1 é a resposta para um grupo controle, e X_2 refere-se a um grupo de tratamento, R será uma medida do efeito do tratamento (ASGHARZADEH et al., 2011).

Portanto, fazendo $\theta_1 = \theta_2 = \theta$, foi obtido:

$$R = \frac{\alpha_1 \left(\alpha_2 \ln(\alpha_2) - \alpha_2 \ln(\alpha_1) + \alpha_1 - \alpha_2\right)}{\left(-\alpha_2 + \alpha_1\right)^2}. \tag{3.40}$$

Para $(\alpha_1, \theta_1) \neq (\alpha_2, \theta_2)$ a função encontrada está no Apêndice A.

3.6 Estimadores

Foram utilizados dez diferentes métodos para se obter as estimativas dos parâmetros α e θ, sendo eles: Método de Máxima Verossimilhança, Mínimos Quadrados, Mínimos Quadrados Ponderados, Cramér-von-Mises, Anderson Darling, Anderson Darling - RT, Anderson Darling - LT, Anderson Darling - 2LT, Kolmogorov e também foi utilizado o método de estimação Bayesiana. Tais métodos serão apresentados a seguir.

Método de Máxima Verossimilhança

Seja X_i, $i = 1, 2, \ldots, n$, uma amostra aleatória com $X_i \sim GEEMO(\alpha, \theta)$ a função de verossimilhança é dada por

$$L(x|\theta, \alpha) = \prod_{i=1}^{n} f(x|\theta, \alpha).$$

Por propriedades matemáticas da função de verossimilhança, maximizar a verossimilhança é o mesmo que maximizar o seu logaritmo. Fazendo $L(\alpha, \theta|x) = l(\alpha, \theta|x)$, tem-se que

$$\begin{aligned} l(\alpha, \theta|x) &= n \ln(\alpha) + n \ln(\theta) + \sum_{i=1}^{n} \ln(x_i) - \sum_{i=1}^{n} x_i \\ &\quad + (\theta - 1) \sum_{i=1}^{n} \ln(1 - e^{-x}(x+1)) - 2 \sum_{i=1}^{n} \ln((1-\alpha)(1 - (1 - e^{-x}(x+1))^{\theta})) \end{aligned} \tag{3.41}$$

Derivando (3.41) em relação aos parâmetros α e θ, é obtido o vetor escore que é composto por

$$U_\alpha(\alpha,\theta|x) = \frac{\partial}{\partial\alpha} l\left(\alpha,\theta|x\right) = \frac{n}{\alpha} - 2\sum_{i=1}^{n} \frac{1-\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta}{1-\left(1-\alpha\right)\left(1-\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta\right)} \quad (3.42)$$

$$\begin{aligned} U_\theta(\alpha,\theta|x) &= \frac{\partial}{\partial\theta} l\left(\alpha,\theta|x\right) = \frac{n}{\theta} + \sum_{i=1}^{n} \ln\left(1-e^{-x_i}\left(x_i+1\right)\right) \\ &- 2\sum_{i=1}^{n} \frac{\left(1-\alpha\right)\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta \ln\left(1-e^{-x_i}\left(x_i+1\right)\right)}{1-\left(1-\alpha\right)\left(1-\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta\right)} \end{aligned} \quad (3.43)$$

Os estimadores de Máxima Verossimilhança (EMV) $\hat{\alpha}_{EMV}$ e $\hat{\theta}_{EMV}$ para os parâmetros α e θ são obtidos por meio das equações não lineares $U_\alpha(\alpha,\theta|x) = 0$ e $U_\theta(\alpha,\theta|x) = 0$.

Dado o vetor de parâmetros desconhecidos $\Theta = (\alpha, \theta)$, para amostras grandes é conhecido que a distribuição assintótica do EMV Θ é

$$(\hat{\Theta} - \Theta) \rightarrow N_2\left(0, I^{-1}(\Theta)\right).$$

Onde a matriz de informação de Fisher observada, $I(\Theta)$, é dada por

$$I(\Theta) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} \quad (3.44)$$

Implicando que a inversa da matriz de informação de Fisher observada, $I^{-1}(\Theta)$, é dada por

$$I^{-1}(\Theta) = \begin{pmatrix} \mathrm{var}(\hat{\alpha}_{EMV}) & cov(\hat{\alpha}_{EMV}, \hat{\theta}_{EMV}) \\ cov(\hat{\theta}_{EMV}, \hat{\alpha}_{EMV}) & \mathrm{var}(\hat{\theta}_{EMV}) \end{pmatrix} \quad (3.45)$$

Sendo assim, tem-se

- $$\frac{\partial^2 l(\alpha,\theta|x)}{\partial\alpha^2} = -\frac{n}{\alpha^2} - 2\sum_{i=1}^{n}\left(-\frac{\left(1-\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta\right)^2}{\left(1-\left(1-\alpha\right)\left(1-\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta\right)\right)^2}\right) \quad (3.46)$$

- $$\begin{aligned} \frac{\partial^2 l(\alpha,\theta|x)}{\partial\theta^2} &= -\frac{n}{\theta^2} \\ &- 2\sum_{i=1}^{n}\left(-\frac{\left(\alpha-1\right)\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta\left(\ln\left(1-e^{-x_i}\left(x_i+1\right)\right)\right)^2\alpha}{\left(-\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta - \alpha + \alpha\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta\right)^2}\right) \end{aligned} \quad (3.47)$$

- $$\frac{\partial^2 l(\alpha,\theta|x)}{\partial\theta\partial\alpha} = -2\sum_{i=1}^{n}\left(-\frac{\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta \ln\left(1-e^{-x_i}\left(x_i+1\right)\right)}{\left(-\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta - \alpha + \alpha\left(1-e^{-x_i}\left(x_i+1\right)\right)^\theta\right)^2}\right) \quad (3.48)$$

Portanto, a abordagem acima é usado para obter os $100\%(1-\tau)$ intervalos de confiança aproximados dos parâmetros $\Theta = (\alpha, \theta)$ como nas seguintes formas

$$\hat{\alpha}_{EMV} \pm z_{\frac{\tau}{2}} \sqrt{\mathrm{var}(\hat{\alpha}_{MLE})} \quad ; \quad \hat{\theta}_{EMV} \pm z_{\frac{\tau}{2}} \sqrt{\mathrm{var}(\hat{\theta}_{MLE})}$$

Onde, tem-se o $z_{\frac{\tau}{2}}$-ésimo percentil da distribuição normal padrão.

Método dos Mínimos Quadrados

Sejam $x_{1:n} < x_{2:n} < \cdots < x_{n:n}$ uma amostra aleatória de tamanho n com função de distribuição $F(x)$, então

$$E[F(x_{i:n}|\alpha,\theta)] = \tfrac{i}{n+1} \qquad \text{e} \qquad V[F(x_{i:n}|\alpha,\theta)] = \tfrac{i(n-i+1)}{(n+1)^2(n+2)}$$

Os Estimadores de Mínimos Quadrados (EMQ), para a GEEMO, $\hat{\alpha}_{EMQ}$ e $\hat{\theta}_{EMQ}$ para os parâmetros α e θ são obtidos através da minimização da função:

$$M(\alpha,\theta) = \sum_{i=1}^{n} \left[F(x_{i:n}|\alpha,\theta) - \frac{i}{n+1} \right]^2. \tag{3.49}$$

Os parâmetros são obtidos por meio das equações $\frac{\partial M}{\partial \alpha} = 0$ e $\frac{\partial M}{\partial \theta} = 0$, que são obtidas através de soluções numéricas, por não possuir forma analítica fechada. Os resultados das derivadas parciais são dadas por:

$$\begin{aligned}
\frac{\partial M}{\partial \alpha} &= \sum_{i=1}^{n} 2 \frac{\left(1-e^{-x_i}(x_i+1)\right)^{\theta} i \left(1-\left(1-e^{-x_i}(x_i+1)\right)^{\theta}\right)}{\left(-\alpha\left(1-e^{-x_i}(x_i+1)\right)^{\theta} + \left(1-e^{-x_i}(x_i+1)\right)^{\theta}+\alpha\right)^2 (n+1)} \\
&\quad - 2\, \frac{\left(\left(1-e^{-1x_i}(x_i+1)\right)^{\theta}\right)^2 \left(1-\left(1-e^{-x_i}(x_i+1)\right)^{\theta}\right)}{\left(-\alpha\left(1-e^{-x_i}(x_i+1)\right)^{\theta} + \left(1-e^{-x_i}(x_i+1)\right)^{\theta}+\alpha\right)^3}, \qquad (3.50) \\
\frac{\partial M}{\partial \theta} &= \sum_{i=1}^{n} \Big(\Big(\left(-2\,i\,\alpha^2 + (2i-2\,n-2)\,\alpha\right) \ln\left(1-e^{-x_i}(x_i+1)\right) \left(\left(1-e^{-x_i}(x_i+1)\right)^{\theta}\right)^2 \\
&\quad + 2\,\ln\left(1-e^{-x_i}(x_i+1)\right)\alpha^2 i\, \left(1-e^{-x_i}(x_i+1)\right)^{\theta}\Big) \\
&\quad \div \left(\alpha\left(1-e^{-x_i}(x_i+1)\right)^{\theta} - \left(1-e^{-x_i}(x_i+1)\right)^{\theta}-\alpha\right)^3 (n+1)\Big). \qquad (3.51)
\end{aligned}$$

Método dos Mínimos Quadrados Ponderados

A linearidade, homocedasticidade, não tendenciosidade, variância mínima e consistência são algumas das características apresentadas pelo estimador de mínimos quadrados, porém, na prática, nem sempre essas suposições são válidas. Por isso, foram desenvolvidos métodos alternativos como o estimador de mínimos quadrados ponderados que corrige a falta de homocedasticidade, que são as variâncias homogêneas.

Os Estimadores de Mínimos Quadrados Ponderados (EMQP), $\hat{\alpha}_{EMQP}$ e $\hat{\theta}_{EMQP}$ para os parâmetros α e θ são obtidos através da minimização da seguinte função

$$W(\alpha,\theta) = \sum_{i=1}^{n} \left[\frac{(n+1)^2(n+2)}{i(n-i+1)} F(x_{i:n}|\alpha,\theta) - \frac{i}{n+1} \right]^2. \tag{3.52}$$

O que diferencia da equação 3.49 é que neste método os estimadores são ponderados pela precisão da distribuição teórica. As estimativas para esses parâmetros são obtidas através de soluções numéricas, por não possuírem forma analítica. As derivadas parciais são dadas por

$$\frac{\partial W(\alpha,\theta|x)}{\partial\alpha} = \sum_{i=1}^{n}\left[\frac{(n+1)^2(n+2)}{i(n-i+1)} \times \frac{F(x_{i:n}|\theta)}{1-\bar{\alpha}S(x_{i:n}|\theta)} - \frac{i}{n+1}\right]\eta_1(\alpha,\theta|x_{i:n}) \tag{3.53}$$

$$\frac{\partial W(\alpha,\theta|x)}{\partial\theta} = \sum_{i=1}^{n}\left[\frac{(n+1)^2(n+2)}{i(n-i+1)} \times \frac{F(x_{i:n}|\theta)}{1-\bar{\alpha}S(x_{i:n}|\theta)} - \frac{i}{n+1}\right]\eta_2(\alpha,\theta|x_{i:n}), \tag{3.54}$$

com

$$\eta_1(\alpha,\theta|x_{i:n}) = -\frac{S(x_{i:n}|\theta)F(x_{i:n}|\theta)}{\left[1-\bar{\alpha}S(x_{i:n}|\theta)\right]^2},$$

$$\eta_2(\alpha,\theta|x_{i:n}) = \frac{(1-\bar{\alpha}S(x_{i:n}|\theta))\left[\frac{\partial}{\partial\theta}F(x_{i:n}|\theta)\right] + \bar{\alpha}F(x_{i:n}|\theta)\left[\frac{\partial}{\partial\theta}S(x_{i:n}|\theta)\right]}{\left[1-\bar{\alpha}S(x_{i:n}|\theta)\right]^2}.$$

Método de Cramér-von-Mises

As estimativas de Cramér-von-Mises (ECM) $\hat{\alpha}_{ECM}$ e $\hat{\theta}_{ECM}$ para os parâmetros α e θ são obtidos através da minimização da distância mínima através da Estatística *goodness-of-fit*, essa Estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$C(\alpha,\theta) = \frac{1}{12n} + \sum_{i=1}^{n}\left[F(x_{i:n}|\alpha,\theta) - \frac{2i-1}{2n}\right]^2. \tag{3.55}$$

Os parâmetros são obtidos por meio das equações $\frac{\partial C}{\partial\alpha} = 0$ e $\frac{\partial C}{\partial\theta} = 0$, que são obtidas através de soluções numéricas, por não possuir forma analítica. As derivadas parciais são dadas por

$$\frac{\partial C(\alpha,\theta|x)}{\partial\alpha} = \sum_{i=1}^{n}\left[\frac{F(x_{i:n}|\theta)}{1-\bar{\alpha}S(x_{i:n}|\theta)} - \frac{2i-1}{2n}\right]\eta_1(\alpha,\theta|x_{i:n}), \tag{3.56}$$

$$\frac{\partial C(\alpha,\theta|x)}{\partial\theta} = \sum_{i=1}^{n}\left[\frac{F(x_{i:n}|\theta)}{1-\bar{\alpha}S(x_{i:n}|\theta)} - \frac{2i-1}{2n}\right]\eta_2(\alpha,\theta|x_{i:n}). \tag{3.57}$$

Método de Anderson Darling

As estimativas de Anderson Darling (EAD) $\hat{\alpha}_{EAD}$ e $\hat{\theta}_{EAD}$ para os parâmetros α e θ são obtidos por meio da minimização da distância mínima através da Estatística *goodness-of-fit*, essa Estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$A(\alpha,\theta) = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\log\left(F(x_{i:n}|\alpha,\theta)\right) + \log\left(\bar{F}(x_{n+1-i:n}|\alpha,\theta)\right)\right\}. \tag{3.58}$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial A}{\partial\alpha} = 0$ e $\frac{\partial A}{\partial\theta} = 0$, essas equações são dadas por

$$\frac{\partial A(\alpha,\theta|x)}{\partial\alpha} = \sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)} - \frac{\eta_1(\alpha,\theta|x_{n+1-i:n})}{S(x_{n+1-i:n}|\alpha,\theta)}\right], \tag{3.59}$$

$$\frac{\partial A(\alpha,\theta|x)}{\partial\theta} = \sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)} - \frac{\eta_2(\alpha,\theta|x_{n+1-i:n})}{S(x_{n+1-i:n}|\alpha,\theta)}\right]. \tag{3.60}$$

Método de Anderson Darling - RT (right-tail)

As estimativas de Anderson Darling - RT (EADD) $\hat{\alpha}_{EADD}$ e $\hat{\theta}_{EADD}$ para os parâmetros α e θ são obtidos através da minimização da distância mínima por meio da Estatística *goodness-of-fit*, essa Estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$R(\alpha,\theta) = \frac{n}{2} - 2\sum_{i=1}^{n} F(x_{i:n}|\alpha,\theta) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\log\left(\bar{F}(x_{n+1-i:n}|\alpha,\theta)\right). \quad (3.61)$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial R}{\partial \alpha} = 0$ e $\frac{\partial R}{\partial \theta} = 0$, essas equações são dadas por

$$\frac{\partial R(\alpha,\theta|x)}{\partial \alpha} = -2\sum_{i=1}^{n}\eta_1(\alpha,\theta|x_{i:n}) + \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1(\alpha,\theta|x_{n+1-i:n})}{S(x_{n+1-i:n}|\alpha,\theta)}\right], \quad (3.62)$$

$$\frac{\partial R(\alpha,\theta|x)}{\partial \theta} = -2\sum_{i=1}^{n}\eta_2(\alpha,\theta|x_{i:n}) + \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2(\alpha,\theta|x_{n+1-i:n})}{S(x_{n+1-i:n}|\alpha,\theta)}\right]. \quad (3.63)$$

Método de Anderson Darling - LT (left-tail)

As estimativas de Anderson Darling - LT (EADL) $\hat{\alpha}_{EADL}$ e $\hat{\theta}_{EADL}$ para os parâmetros α e θ são obtidos através da minimização da distância mínima por meio da Estatística *goodness-of-fit*, essa Estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$LT = -\frac{3n}{2} + 2\sum_{i=1}^{n} F(x_{i:n}|\alpha,\theta) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)ln(F(x_{i:n}|\alpha,\theta)). \quad (3.64)$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial LT}{\partial \alpha} = 0$ e $\frac{\partial LT}{\partial \theta} = 0$, essas equações são dadas por

$$\frac{\partial LT(\alpha,\theta|x)}{\partial \alpha} = 2\sum_{i=1}^{n}\eta_1(\alpha,\theta|x_{i:n}) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)}\right], \quad (3.65)$$

$$\frac{\partial LT(\alpha,\theta|x)}{\partial \theta} = 2\sum_{i=1}^{n}\eta_2(\alpha,\theta|x_{i:n}) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)}\right]. \quad (3.66)$$

Método de Anderson Darling - 2LT (left-tail de segunda ordem)

As estimativas de Anderson Darling - 2LT (EAD2L) $\hat{\alpha}_{EAD2L}$ e $\hat{\theta}_{EAD2L}$ para os parâmetros α e θ são obtidos através da minimização da distância mínima por meio da Estatística *goodness-of-fit*, essa Estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$LT = 2\sum_{i=1}^{n} ln(F(x_{i:n}|\alpha,\theta)) + \frac{1}{n}\sum_{i=1}^{n}\frac{2i-1}{F(x_{i:n}|\alpha,\theta)}. \quad (3.67)$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial LT}{\partial \alpha} = 0$ e $\frac{\partial LT}{\partial \theta} = 0$, essas equações são dadas por

$$\frac{\partial LT(\alpha,\theta|x)}{\partial \alpha} = 2\sum_{i=1}^{n}\frac{\eta_1(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)} - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1(\alpha,\theta|x_{i:n})}{(F(x_{i:n}|\alpha,\theta))^2}\right], \tag{3.68}$$

$$\frac{\partial LT(\alpha,\theta|x)}{\partial \theta} = 2\sum_{i=1}^{n}\frac{\eta_2(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)} - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2(\alpha,\theta|x_{i:n})}{(F(x_{i:n}|\alpha,\theta))^2}\right]. \tag{3.69}$$

Método de Kolmogorov

As estimativas de Kolmogorov (EK) $\hat{\alpha}_{EK}$ e $\hat{\theta}_{EK}$ para os parâmetros α e θ são obtidos através da minimização da distância mínima por meio da Estatística *goodness-of-fit*, essa Estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$K(\alpha,\theta) = \max_{1\leq i\leq n}\left\{\frac{i}{n} - F(x_{i:n}|\alpha,\theta), F(x_{i:n}|\alpha,\theta) - \frac{i-1}{n}\right\}. \tag{3.70}$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial K}{\partial \alpha} = 0$ e $\frac{\partial K}{\partial \theta} = 0$.

Método Bayesiano

A Inferência Bayesiana se diferencia da clássica, pois permite que a opinião do especialista seja levada em consideração, ou seja, a estimativa dos parâmetros não depende apenas dos dados, mas também da informação a priori que é obtida pelo especialista do assunto tratado. Os parâmetros da distribuição em estudo são tidos como variáveis aleatórias na inferência Bayesiana, diferenciando-se da clássica, onde o mesmo é tido como um valor fixo desconhecido. No método Bayesiano os parâmetros assumem uma distribuição (distribuição à priori) e o mesmo é atualizado (distribuição a posteriori) pelo Teorema de Bayes, combinando com a informação trazida dos dados pela função de Verossimilhança.

Para os parâmetros desconhecidos da $GEEMO \sim (\alpha,\theta)$ será assumido que α e θ tenham como distribuições a priori independentes a Gama. Com função densidade probabilidade dadas por

$$g(\alpha) \propto \alpha^{a-1}e^{\alpha b}, \quad \alpha > 0,$$
$$g(\theta) \propto \theta^{c-1}e^{\theta d}, \quad \theta > 0.$$

Os hiperparâmetros a, b, c e d são conhecidos e não negativos.

Teorema de Bayes

O teorema de Bayes, utilizado para atualizar probabilidades, é um corolário do teorema da probabilidade total que permite calcular a seguinte probabilidade:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}. \tag{3.71}$$

Para realizar a inferência, pode-se reescrever 3.71 da seguinte forma:

$$p(\theta|x) = \frac{p(\theta)p(x|\theta)}{p(x)} = \frac{p(\theta)p(x|\theta)}{\int p(\theta)p(x|\theta)d\theta},$$

sendo que

- $p(\theta)$ é o a distribuição a priori do(s) parâmetro(s);
- $p(x|\theta)$ representa a informação proveniente dos dados;
- $p(x)$ é a distribuição marginal de x , havendo a necessidade de integrar $p(\theta)p(x|\theta)$ em relação a θ , cobrindo todo seu espaço paramétrico;
- $p(\theta|x)$ é a distribuição a posteriori de θ. É uma combinação da informação entre a priori e a informação do especialista.

Para obtenção dos estimadores foi utilizado o algoritmo de Metropolis-Hasting, que é um processo iterativo utilizado para contornar o problema da obtenção da posteriori de forma analítica, gerando possíveis valores da mesma, proposto inicialmente por Metropolis Nicholas e Rosenbluth (1953) e generalizado por Hastings (1970), onde ambos buscavam a obtenção de coeficientes de determinadas integrais inatingíveis analiticamente.

Utilizando uma distribuição de probabilidade auxiliar $q(\theta, \cdot)$, são gerados possíveis valores da distribuição a posteriori, onde o valor gerado é aceito mediante a satisfação da condição abaixo

$$\alpha(\theta^{(t)}, \theta^{(*)}) = \min\left[1; \frac{p(\theta^{(*)})q(\theta^{(t)}|\theta^{(*)})}{p(\theta^{(t)})q(\theta^{(*)}|\theta^{(t)})}\right] \geq u(0,1).$$

Então, são realizados os seguintes passos:

1 Proponha a distribuição auxiliar que gere possíveis valores de θ a posteriori;

2 Inicie o contador $t = 1$;

3 Escolha um valor inicial para $\theta^{(t)}$;

4 Gere um valor para $\theta^{(t)}$ utilizando $q(\theta, \cdot)$;

5 Teste a condição $\alpha(\theta^{(t)}, \theta^{(*)}) \geq u(0,1)$;

6 Se a condição for satisfeita: $\theta^{(t+1)} = \theta^{(*)}, t = t + 1$ e volte ao passo 4;

7 Se a condição não for satisfeita: $\theta^{(t+1)} = \theta^{(t)}, t = t + 1$ e volte ao passo 4;

8 Repita o procedimento até alcançar convergência.

O algoritmo foi escrito em linguagem de programação *R*, para cada uma das 1000 amostras foram realizadas 20000 iterações, porém as 5000 primeiras foram descartadas para o aquecimento inicial (*burn-in* inicial), e o resultado utilizado foi a moda e a mediana destes 15000 valores gerados, para α e θ respectivamente.

3.7 Simulação

Foram geradas $B = 1000$ amostras da distribuição GEEMO para diferentes valores dos parâmetros, para quatro diferentes tamanhos amostrais, sendo eles $n = 20$, 60, 100 e 200.

As amostras foram obtidas por meio da transformação inversa gerando-se um vetor u de tamanho 200000 tal que $u \sim U(0,1)$, então são encontradas as raízes para $F(x) = u \rightarrow x = F^{-1}(u)$, estas são obtidas pelo método numérico de Newton-Raphson. Tais raízes são armazenadas numa matriz X de 200 linhas (número máximo de tamanho da amostra) e 1000 colunas (número de amostras). Sendo assim, na primeira coluna da matriz X, as

20 primeiras linhas representam a primeira amostra de tamanho $n = 20$, as 60 primeiras linhas, ou seja, a primeira amostra adicionando-se as 40 próximas linhas da primeira coluna representa a primeira amostra de tamanho $n = 60$, considerando as 100 primeiras linhas da mesma coluna obtêm-se a primeira amostra de tamanho $n = 100$ e, finalmente, considerando-se todas as linhas da primeira coluna, tem-se a primeira amostra de tamanho $n = 200$. Repetindo esse processo nas 1000 colunas, obtemos as 1000 amostras de cada um dos diferentes tamanhos amostrais considerados no estudo.

3.8 Comparador de Estimadores

Para comparar as diferentes estimativas para cada um dos métodos apresentados anteriormente, foram calculados o vício ($BIAS$), o erro padrão médio ($RMSE$), a média da diferença absoluta entre a distribuição teórica e a estimada (D_{abs}) e a diferença absoluta máxima (D_{max}) entre as funções de distribuições verdadeira e estimada sendo que suas expressões são dadas por

$$
\begin{aligned}
BIAS(\hat{\alpha}) &= \tfrac{1}{B}\sum_{i=1}^{B}(\hat{\alpha}_i - \alpha), \qquad BIAS(\hat{\theta}) = \tfrac{1}{B}\sum_{i=1}^{B}(\hat{\theta}_i - \theta),\\
RMSE(\hat{\alpha}) &= \sqrt{\tfrac{1}{B}\sum_{i=1}^{B}(\hat{\alpha}_i - \alpha)^2}, \quad RMSE(\hat{\theta}) = \sqrt{\tfrac{1}{B}\sum_{i=1}^{B}(\hat{\theta}_i - \theta)^2},\\
D_{abs} &= \tfrac{1}{Bn}\sum_{i=1}^{B}\sum_{j=1}^{n}\left|F(x_{ij}|\alpha,\theta) - F(t_{ij}|\hat{\alpha},\hat{\theta})\right|,\\
D_{\max} &= \tfrac{1}{B}\sum_{i=1}^{B}\max_{j}\left|F(x_{ij}|\alpha,\theta) - F(x_{ij}|\hat{\alpha},\hat{\theta})\right|.
\end{aligned}
\tag{3.72}
$$

3.9 Resultados das Simulações

As comparações dos métodos de estimação foram realizadas neste tópico, com o intuito de verificar onde se encaixam os melhores estimadores para diferentes valores de parâmetros. Nas Tabelas de 3.2 até 3.8 estão os valores dos parâmetros obtidos pelos diferentes estimadores, os valores entre os parênteses representam os *ranks* dos estimadores.

É possível observar que para os estimadores de Mínimo Quadrado (Ponderado) e o Anderson Darling com segunda cauda para θ são viciados negativamente. Para o restante dos estimadores e incluindo todos os valores estimados tanto para θ quanto para α são viciados positivamente, porém vale a propriedade de não tendenciosidade, pois quando n aumenta o vício tende a zero.

Com relação ao erro padrão médio, para todos os métodos de estimação se observa a propriedade de consistência, ou seja, quando n é grande a variância tende a zero.

O estimador de Máxima Verossimilhança possui o menor vício para α e o maior vício para este parâmetro ficou para o estimador Anderson Darling Calda a Direita. Já para o parâmetro θ os estimadores de Mínimo Quadrado, Mínimo Quadrado Ponderado e Anderson Darling tiveram o menor vício enquanto que o estimador de Máxima Verossimilhança ficou com os maiores valores.

Em relação ao erro padrão médio, para α o de menor vício foi o estimador bayesiano, para valores maiores dos parâmetros, e o de Máxima Verossimilhança para parâmetros com valores mais baixos. O pior ficou com Anderson Darling Calda a Direita. Para θ, o menor valor no erro padrão médio foi o bayesiano e para pequenas amostras os estimadores Anderson Darling e Anderson Darling com cauda a Esquerda tiveram valores próximos, e para amostras maiores o estimador de Máxima Verossimilhança apresentou resultados

Tabela 3.2: Resultado da simulação para os parâmetros α=0.5 e θ=0.8

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.0705(1)	0.1261(10)	0.547(1)	0.3594(4)	0.0576(2)	0.0975(3)
MQ		1.5758(9)	-0.0272(2)	17.4137(8)	0.3857(6)	0.0598(6)	0.1001(6)
MQP		1.3019(8)	-0.0012(1)	25.0857(9)	0.3602(5)	0.0585(5)	0.0977(5)
CM		0.5846(6)	0.1163(8)	8.7437(6)	0.441(8)	0.0598(7)	0.1033(7)
KS		0.2557(5)	0.1251(9)	1.1527(5)	0.4418(9)	0.0608(8)	0.1052(8)
AD		0.9622(7)	0.0405(3)	13.28(7)	0.3494(2)	0.058(3)	0.0969(2)
ADR		3.1530(10)	0.0968(6)	30.7816(10)	0.5169(10)	0.0621(9)	0.1081(10)
ADL		0.1941(4)	0.0691(5)	1.0381(4)	0.3542(3)	0.0581(4)	0.0976(4)
AD2L		0.1870(3)	0.0999(7)	0.7722(3)	0.4249(7)	0.0622(10)	0.1063(9)
BAYES		0.1556(2)	0.0585(4)	0.5647(2)	0.3236(1)	0.0574(1)	0.0951(1)
MV	60	0.0238(1)	0.0409(10)	0.2397(1)	0.1847(2)	0.0338(2)	0.057(2)
MQ		0.1146(9)	-0.0072(1)	0.374(8)	0.2186(6)	0.0352(7)	0.0599(6)
MQP		0.0759(7)	0.0098(2)	0.3067(5)	0.1996(5)	0.0343(4)	0.0582(5)
CM		0.0567(4)	0.0371(9)	0.321(6)	0.228(8)	0.0352(6)	0.0605(7)
KS		0.0720(6)	0.0347(8)	0.3671(7)	0.2337(9)	0.0361(8)	0.0619(8)
AD		0.0681(5)	0.0143(5)	0.2997(4)	0.1924(4)	0.0342(3)	0.0577(3)
ADR		0.1253(10)	0.0273(7)	0.5745(10)	0.262(10)	0.0364(9)	0.0636(9)
ADL		0.0474(2)	0.0241(6)	0.2679(3)	0.1899(3)	0.0345(5)	0.058(4)
AD2L		0.1089(8)	0.0117(4)	0.4105(9)	0.2257(7)	0.0387(10)	0.0659(10)
BAYES		0.0531(3)	0.0113(3)	0.2598(2)	0.1725(1)	0.0328(1)	0.0544(1)
MV	100	0.0130(1)	0.0213(10)	0.1777(1)	0.1324(2)	0.0257(2)	0.0432(2)
MQ		0.0659(8)	-0.0099(5)	0.2522(8)	0.1586(6)	0.0267(6)	0.0455(6)
MQP		0.0413(7)	0.0025(1)	0.211(4)	0.1424(5)	0.026(4)	0.0439(4)
CM		0.0341(4)	0.0162(8)	0.2305(6)	0.1618(7)	0.0268(7)	0.0458(7)
KS		0.0392(5)	0.0182(9)	0.2516(7)	0.172(9)	0.0275(8)	0.0473(8)
AD		0.0402(6)	0.0039(2)	0.2121(5)	0.1404(4)	0.0259(3)	0.0438(3)
ADR		0.0686(9)	0.0104(6)	0.3343(10)	0.1942(10)	0.0278(9)	0.0485(9)
ADL		0.0272(2)	0.0105(7)	0.1939(3)	0.1369(3)	0.0262(5)	0.044(5)
AD2L		0.0834(10)	-0.0062(3)	0.3124(9)	0.1702(8)	0.0303(10)	0.0515(10)
BAYES		0.0310(3)	0.0068(4)	0.1855(2)	0.1310(1)	0.0247(1)	0.0413(1)
MV	200	0.0056(1)	0.0096(9)	0.1185(1)	0.0905(1)	0.0176(1)	0.0294(1)
MQ		0.0329(9)	-0.0082(8)	0.1496(7)	0.1122(6)	0.0184(7)	0.0314(6)
MQP		0.0189(5)	0.0003(2)	0.1305(4)	0.0996(5)	0.0178(4)	0.0301(4)
CM		0.0180(4)	0.0047(5)	0.1427(6)	0.1129(7)	0.0184(6)	0.0314(7)
KS		0.0192(7)	0.0065(7)	0.1509(8)	0.1203(9)	0.0191(8)	0.0327(8)
AD		0.0191(6)	0.0003(1)	0.1308(5)	0.0985(4)	0.0178(3)	0.03(3)
ADR		0.0307(8)	0.0026(3)	0.174(9)	0.1345(10)	0.0192(9)	0.0333(9)
ADL		0.0125(2)	0.004(4)	0.1256(2)	0.095(3)	0.018(5)	0.0302(5)
AD2L		0.0585(10)	-0.0143(10)	0.2184(10)	0.1187(8)	0.022(10)	0.0371(10)
BAYES		0.0142(3)	0.0059(6)	0.1275(3)	0.0936(2)	0.0178(2)	0.0299(2)

Tabela 3.3: Resultado da simulação para os parâmetros α=0.5 e θ=1.5

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.0704(1)	0.2366(9)	0.5466(1)	0.6738(4)	0.0576(1)	0.0975(2)
MQ		1.6419(9)	-0.0511(2)	19.8484(9)	0.7231(6)	0.0598(6)	0.1001(6)
MQP		0.9884(8)	-0.0023(1)	14.4253(8)	0.6754(5)	0.0585(4)	0.0977(5)
CM		0.7981(6)	0.2181(8)	13.229(7)	0.827(8)	0.0598(7)	0.1033(7)
KS		0.2557(5)	0.237(10)	1.1858(5)	0.8315(9)	0.0608(8)	0.1053(8)
AD		0.8309(7)	0.0759(3)	10.9628(6)	0.6551(2)	0.058(2)	0.0969(1)
ADR		3.2545(10)	0.1816(6)	33.1077(10)	0.9692(10)	0.0621(9)	0.1081(10)
ADL		0.1941(3)	0.1296(5)	1.0369(4)	0.6642(3)	0.0581(3)	0.0976(3)
AD2L		0.187(2)	0.1874(7)	0.7719(3)	0.7963(7)	0.0622(10)	0.1063(9)
BAYES		0.1965(4)	0.0990(4)	0.6123(2)	0.6099(1)	0.0589(5)	0.0976(4)
MV	60	0.0238(1)	0.0767(10)	0.2397(2)	0.3464(2)	0.0338(2)	0.057(2)
MQ		0.1147(9)	-0.0135(1)	0.374(8)	0.4098(6)	0.0352(7)	0.0599(6)
MQP		0.076(6)	0.0185(2)	0.3067(5)	0.3743(5)	0.0343(4)	0.0582(5)
CM		0.0568(3)	0.0695(9)	0.321(6)	0.4275(8)	0.0352(6)	0.0605(7)
KS		0.0721(5)	0.0649(8)	0.367(7)	0.4381(9)	0.0361(8)	0.0619(8)
AD		0.0682(4)	0.0268(4)	0.2997(4)	0.3608(4)	0.0342(3)	0.0577(3)
ADR		0.1253(10)	0.0513(6)	0.5744(10)	0.4912(10)	0.0364(9)	0.0636(9)
ADL		0.0475(2)	0.0452(5)	0.2679(3)	0.3561(3)	0.0345(5)	0.058(4)
AD2L		0.1089(8)	0.0219(3)	0.4106(9)	0.4232(7)	0.0387(10)	0.0659(10)
BAYES		0.0942(7)	-0.0626(7)	0.2250(1)	0.2663(1)	0.0317(1)	0.0518(1)
MV	100	0.013(1)	0.0399(9)	0.1777(2)	0.2482(2)	0.0257(2)	0.0432(2)
MQ		0.066(8)	-0.0186(4)	0.2522(8)	0.2974(6)	0.0267(6)	0.0455(6)
MQP		0.0413(6)	0.0048(1)	0.211(4)	0.2671(5)	0.026(4)	0.0439(4)
CM		0.0342(3)	0.0305(7)	0.2305(6)	0.3034(7)	0.0268(7)	0.0458(7)
KS		0.0394(4)	0.0334(8)	0.2514(7)	0.3214(9)	0.0275(8)	0.0472(8)
AD		0.0403(5)	0.0073(2)	0.2121(5)	0.2632(4)	0.0259(3)	0.0438(3)
ADR		0.0687(9)	0.0196(5)	0.3342(10)	0.3642(10)	0.0278(9)	0.0485(9)
ADL		0.0272(2)	0.0196(6)	0.1939(3)	0.2567(3)	0.0262(5)	0.044(5)
AD2L		0.0835(10)	-0.0116(3)	0.3123(9)	0.3192(8)	0.0303(10)	0.0515(10)
BAYES		0.0630(7)	-0.0480(10)	0.1765(1)	0.2141(1)	0.0247(1)	0.0408(1)
MV	200	0.0056(1)	0.0181(9)	0.1184(1)	0.1697(2)	0.0176(2)	0.0294(2)
MQ		0.0329(9)	-0.0154(8)	0.1496(7)	0.2103(6)	0.0184(7)	0.0314(6)
MQP		0.0189(5)	0.0006(2)	0.1305(4)	0.1867(5)	0.0178(4)	0.0301(4)
CM		0.018(3)	0.0089(6)	0.1427(6)	0.2117(7)	0.0184(6)	0.0314(7)
KS		0.0192(7)	0.0122(7)	0.1509(8)	0.2256(9)	0.0191(8)	0.0327(8)
AD		0.0192(6)	0.0005(1)	0.1308(5)	0.1847(4)	0.0178(3)	0.03(3)
ADR		0.0307(8)	0.0048(4)	0.1739(9)	0.2523(10)	0.0192(9)	0.0333(9)
ADL		0.0126(2)	0.0074(5)	0.1255(3)	0.1782(3)	0.018(5)	0.0302(5)
AD2L		0.0586(10)	-0.0269(10)	0.2184(10)	0.2225(8)	0.022(10)	0.0371(10)
BAYES		0.0182(4)	0.0044(3)	0.1187(2)	0.1684(1)	0.0171(1)	0.0286(1)

Tabela 3.4: Resultado da simulação para os parâmetros α=0.5 e θ=5

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.0705(1)	0.7886(9)	0.5472(1)	2.246(4)	0.0576(1)	0.0975(3)
MQ		1.2337(8)	-0.1702(2)	12.2539(7)	2.4101(6)	0.0598(6)	0.1001(6)
MQP		1.5106(9)	-0.0078(1)	27.5602(9)	2.2513(5)	0.0585(5)	0.0977(5)
CM		0.7937(6)	0.7273(8)	13.014(8)	2.7557(8)	0.0598(7)	0.1033(7)
KS		0.2555(5)	0.7902(10)	1.1647(5)	2.7722(9)	0.0608(8)	0.1054(8)
AD		0.8155(7)	0.2535(4)	11.1846(6)	2.1836(2)	0.058(2)	0.0969(2)
ADR		3.4999(10)	0.6066(6)	38.3057(10)	3.2284(10)	0.0621(9)	0.1081(10)
ADL		0.1941(3)	0.432(5)	1.0376(4)	2.2138(3)	0.0581(3)	0.0976(4)
AD2L		0.187(2)	0.6248(7)	0.7719(3)	2.6559(7)	0.0622(10)	0.1063(9)
BAYES		0.2376(4)	-0.2328(3)	0.6478(2)	1.9231(1)	0.0584(4)	0.0961(1)
MV	60	0.0238(1)	0.2554(10)	0.2397(1)	1.1546(2)	0.0338(2)	0.057(2)
MQ		0.1147(9)	-0.045(1)	0.374(8)	1.3661(6)	0.0352(7)	0.0599(6)
MQP		0.076(7)	0.0615(2)	0.3067(5)	1.2475(5)	0.0343(4)	0.0582(5)
CM		0.0568(3)	0.2317(9)	0.321(6)	1.4249(8)	0.0352(6)	0.0605(7)
KS		0.0721(6)	0.216(8)	0.3671(7)	1.4599(9)	0.0361(8)	0.0619(8)
AD		0.0682(4)	0.0892(5)	0.2997(4)	1.2025(4)	0.0342(3)	0.0577(3)
ADR		0.1253(10)	0.1708(7)	0.5744(10)	1.6373(10)	0.0364(9)	0.0636(9)
ADL		0.0475(2)	0.1506(6)	0.2679(2)	1.1871(3)	0.0345(5)	0.058(4)
AD2L		0.109(8)	0.073(3)	0.4106(9)	1.4107(7)	0.0387(10)	0.0659(10)
BAYES		0.0716(5)	-0.0887(4)	0.2706(3)	1.0708(1)	0.0335(1)	0.0555(1)
MV	100	0.013(1)	0.1329(10)	0.1777(1)	0.8272(2)	0.0257(2)	0.0432(2)
MQ		0.066(8)	-0.0621(5)	0.2522(8)	0.9913(6)	0.0267(6)	0.0455(6)
MQP		0.0413(6)	0.0159(1)	0.211(4)	0.8903(5)	0.026(4)	0.0439(4)
CM		0.0342(3)	0.1015(8)	0.2305(6)	1.0113(7)	0.0268(7)	0.0458(7)
KS		0.0391(4)	0.1143(9)	0.2516(7)	1.0756(9)	0.0275(8)	0.0473(8)
AD		0.0403(5)	0.0242(2)	0.2121(5)	0.8773(4)	0.0259(3)	0.0438(3)
ADR		0.0687(9)	0.0651(6)	0.3342(10)	1.214(10)	0.0278(9)	0.0485(9)
ADL		0.0272(2)	0.0654(7)	0.1939(3)	0.8556(3)	0.0262(5)	0.044(5)
AD2L		0.0835(10)	-0.0386(3)	0.3123(9)	1.0638(8)	0.0303(10)	0.0515(10)
BAYES		0.0422(7)	-0.0580(4)	0.1904(2)	0.8162(1)	0.0251(1)	0.0418(1)
MV	200	0.0057(1)	0.0602(9)	0.1185(1)	0.5654(1)	0.0176(1)	0.0294(1)
MQ		0.0329(9)	-0.0514(8)	0.1496(7)	0.7009(6)	0.0184(7)	0.0314(6)
MQP		0.0189(4)	0.0019(2)	0.1305(4)	0.6222(5)	0.0178(3)	0.0301(4)
CM		0.018(3)	0.0296(6)	0.1427(6)	0.7056(7)	0.0184(6)	0.0314(7)
KS		0.0192(6)	0.0405(7)	0.1509(8)	0.7521(9)	0.0191(8)	0.0327(8)
AD		0.0192(5)	0.0018(1)	0.1308(5)	0.6157(4)	0.0178(2)	0.03(3)
ADR		0.0307(8)	0.0161(4)	0.174(9)	0.8409(10)	0.0192(9)	0.0333(9)
ADL		0.0126(2)	0.0248(5)	0.1255(2)	0.5939(3)	0.018(5)	0.0302(5)
AD2L		0.0586(10)	-0.0896(10)	0.2184(10)	0.7418(8)	0.022(10)	0.0371(10)
BAYES		0.0196(7)	-0.0148(3)	0.1291(3)	0.5768(2)	0.0179(4)	0.0299(2)

Tabela 3.5: Resultado da simulação para os parâmetros α=1.5 e θ=0.8

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.2008(1)	0.2446(8)	1.8891(2)	0.5854(4)	0.0568(2)	0.0951(4)
MQ		19.8476(9)	-0.0054(1)	83.2247(9)	0.644(7)	0.059(6)	0.099(6)
MQP		11.4262(8)	0.0271(2)	80.6739(8)	0.5988(5)	0.0577(5)	0.0959(5)
CM		9.7177(7)	0.2362(7)	75.328(7)	0.7792(9)	0.0595(7)	0.1029(8)
KS		1.1258(4)	0.2463(10)	3.8142(4)	0.7489(8)	0.0604(9)	0.1043(9)
AD		8.7759(6)	0.0874(4)	72.7305(6)	0.5651(3)	0.0573(4)	0.0948(2)
ADR		24.2016(10)	0.2449(9)	102.0726(10)	0.9779(10)	0.0621(10)	0.1093(10)
ADL		3.5506(5)	0.1255(5)	44.1342(5)	0.5573(2)	0.0572(3)	0.0948(3)
AD2L		0.7994(3)	0.1755(6)	3.7174(3)	0.6237(6)	0.0601(8)	0.1013(7)
BAYES		0.2910(2)	0.0721(3)	1.3272(1)	0.4628(1)	0.0523(1)	0.0851(1)
MV	60	0.0708(1)	0.0784(10)	0.7947(1)	0.28(2)	0.0333(2)	0.055(2)
MQ		0.8946(9)	0.0004(1)	5.2236(9)	0.3625(7)	0.035(6)	0.0593(6)
MQP		0.3635(6)	0.0275(3)	1.3324(5)	0.3194(6)	0.034(4)	0.0569(5)
CM		0.4416(7)	0.0748(9)	2.6908(8)	0.3848(8)	0.035(7)	0.0601(7)
KS		0.4911(8)	0.0699(8)	1.886(7)	0.3982(9)	0.0362(8)	0.0621(9)
AD		0.3472(5)	0.0306(4)	1.4944(6)	0.2998(4)	0.0337(3)	0.0562(4)
ADR		2.6405(10)	0.068(7)	21.2351(10)	0.4605(10)	0.0365(9)	0.0641(10)
ADL		0.2144(3)	0.043(6)	1.0168(3)	0.2866(3)	0.034(5)	0.0562(3)
AD2L		0.3238(4)	0.0314(5)	1.2519(4)	0.3062(5)	0.0373(10)	0.0618(8)
BAYES		0.1856(2)	0.0167(2)	0.8547(2)	0.2513(1)	0.0314(1)	0.0514(1)
MV	100	0.0354(1)	0.0424(10)	0.579(1)	0.1952(2)	0.0253(2)	0.0417(2)
MQ		0.3625(9)	-0.0098(2)	1.3434(9)	0.2579(7)	0.0266(6)	0.045(6)
MQP		0.182(4)	0.0104(3)	0.8064(4)	0.2218(6)	0.0257(4)	0.0429(5)
CM		0.2086(6)	0.0338(8)	1.1042(7)	0.2659(8)	0.0266(7)	0.0454(7)
KS		0.2849(8)	0.0339(9)	1.2704(8)	0.29(9)	0.0276(8)	0.0474(8)
AD		0.1833(5)	0.0108(4)	0.8322(5)	0.2149(4)	0.0256(3)	0.0427(4)
ADR		1.0208(10)	0.0288(7)	12.7823(10)	0.3323(10)	0.0278(9)	0.0488(10)
ADL		0.1168(3)	0.0197(6)	0.6879(3)	0.2032(3)	0.0258(5)	0.0427(3)
AD2L		0.2357(7)	0.002(1)	0.9088(6)	0.2216(5)	0.0289(10)	0.0478(9)
BAYES		0.1134(2)	0.0111(5)	0.6235(2)	0.1891(1)	0.0239(1)	0.0394(1)
MV	200	0.014(1)	0.0202(10)	0.3899(1)	0.1305(1)	0.0172(1)	0.0282(1)
MQ		0.1566(8)	-0.0097(6)	0.6062(7)	0.1807(7)	0.0183(6)	0.031(6)
MQP		0.0782(4)	0.0044(2)	0.4663(4)	0.1524(6)	0.0176(4)	0.0293(4)
CM		0.0946(6)	0.0118(7)	0.5638(6)	0.183(8)	0.0183(7)	0.0311(7)
KS		0.1207(7)	0.0139(9)	0.6449(9)	0.2031(9)	0.0192(8)	0.0329(8)
AD		0.0822(5)	0.0032(1)	0.4711(5)	0.1496(4)	0.0176(3)	0.0292(3)
ADR		0.185(10)	0.0095(5)	0.8148(10)	0.2276(10)	0.0192(9)	0.0334(9)
ADL		0.0517(2)	0.0084(3)	0.4274(3)	0.1401(3)	0.0177(5)	0.0293(5)
AD2L		0.1627(9)	-0.0127(8)	0.6307(8)	0.1522(5)	0.021(10)	0.0345(10)
BAYES		0.0522(3)	0.0093(4)	0.4170(2)	0.1336(2)	0.0174(2)	0.0287(2)

Tabela 3.6: Resultado da simulação para os parâmetros α=1.5 e θ=1.5

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.2006(1)	0.4588(9)	1.8878(2)	1.0978(4)	0.0568(2)	0.0951(4)
MQ		22.7754(9)	-0.0102(1)	96.1041(9)	1.2075(7)	0.059(6)	0.0989(6)
MQP		12.3541(8)	0.0509(2)	90.0936(8)	1.1231(5)	0.0577(5)	0.0959(5)
CM		10.34(7)	0.4428(7)	74.8016(7)	1.461(9)	0.0595(7)	0.1029(8)
KS		1.1381(4)	0.4571(8)	3.8393(4)	1.3891(8)	0.0604(9)	0.1042(9)
AD		8.797(6)	0.1637(3)	67.9383(6)	1.0596(3)	0.0573(4)	0.0948(3)
ADR		25.9719(10)	0.459(10)	117.8518(10)	1.8337(10)	0.0621(10)	0.1093(10)
ADL		3.795(5)	0.2354(4)	49.7045(5)	1.0448(2)	0.0572(3)	0.0948(2)
AD2L		0.8012(3)	0.329(6)	3.7381(3)	1.1694(6)	0.0601(8)	0.1013(7)
BAYES		0.3341(2)	0.2433(5)	1.3569(1)	0.8656(1)	0.0559(1)	0.0903(1)
MV	60	0.0708(1)	0.1471(10)	0.7947(1)	0.525(2)	0.0333(2)	0.055(2)
MQ		1.0408(9)	0.0008(1)	9.4165(9)	0.6797(7)	0.035(6)	0.0593(6)
MQP		0.3635(6)	0.0516(3)	1.3323(5)	0.5989(6)	0.034(4)	0.0569(5)
CM		0.4697(7)	0.1402(9)	3.4054(8)	0.7215(8)	0.035(7)	0.0601(7)
KS		0.4984(8)	0.1328(8)	1.9199(7)	0.7509(9)	0.0361(8)	0.0621(9)
AD		0.3471(5)	0.0573(4)	1.4945(6)	0.5622(4)	0.0337(3)	0.0562(4)
ADR		3.028(10)	0.1276(7)	25.5184(10)	0.8634(10)	0.0365(9)	0.0641(10)
ADL		0.2144(3)	0.0807(6)	1.0167(3)	0.5374(3)	0.034(5)	0.0562(3)
AD2L		0.3238(4)	0.0587(5)	1.2515(4)	0.574(5)	0.0373(10)	0.0618(8)
BAYES		0.1915(2)	0.0473(2)	0.8518(2)	0.4675(1)	0.0321(1)	0.0522(1)
MV	100	0.0356(1)	0.0796(10)	0.5795(1)	0.3662(2)	0.0253(2)	0.0417(2)
MQ		0.3625(9)	-0.0184(2)	1.3436(9)	0.4835(7)	0.0266(6)	0.045(6)
MQP		0.1815(4)	0.0196(3)	0.8059(4)	0.4156(6)	0.0257(4)	0.0429(5)
CM		0.2086(6)	0.0634(8)	1.1044(7)	0.4986(8)	0.0266(7)	0.0454(7)
KS		0.2854(8)	0.0638(9)	1.2727(8)	0.5441(9)	0.0276(8)	0.0474(8)
AD		0.1833(5)	0.0202(4)	0.8321(5)	0.4029(4)	0.0256(3)	0.0427(4)
ADR		0.8657(10)	0.0541(7)	8.4605(10)	0.6231(10)	0.0278(9)	0.0488(10)
ADL		0.1168(2)	0.037(6)	0.6879(3)	0.3809(3)	0.0258(5)	0.0427(3)
AD2L		0.2355(7)	0.0039(1)	0.9087(6)	0.4156(5)	0.0289(10)	0.0478(9)
BAYES		0.1227(3)	0.0244(5)	0.6273(2)	0.3545(1)	0.0241(1)	0.0396(1)
MV	200	0.014(1)	0.0379(10)	0.3898(1)	0.2446(1)	0.0172(1)	0.0282(1)
MQ		0.1566(8)	-0.0182(6)	0.6062(7)	0.3388(7)	0.0183(6)	0.031(6)
MQP		0.0782(4)	0.0083(2)	0.4663(4)	0.2858(6)	0.0176(4)	0.0293(4)
CM		0.0947(6)	0.0221(7)	0.5638(6)	0.3429(8)	0.0183(7)	0.0311(7)
KS		0.1204(7)	0.0262(9)	0.6443(9)	0.3808(9)	0.0192(8)	0.0329(8)
AD		0.0822(5)	0.006(1)	0.4711(5)	0.2806(4)	0.0176(3)	0.0292(3)
ADR		0.185(10)	0.0179(5)	0.8148(10)	0.4267(10)	0.0192(9)	0.0334(9)
ADL		0.0516(2)	0.0158(3)	0.4274(3)	0.2626(3)	0.0177(5)	0.0293(5)
AD2L		0.1626(9)	-0.0237(8)	0.6306(8)	0.2854(5)	0.021(10)	0.0345(10)
BAYES		0.0584(3)	0.0170(4)	0.4205(2)	0.2504(2)	0.0175(2)	0.0287(2)

Tabela 3.7: Resultado da simulação para os parâmetros α=2 e θ=1.5

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.2662(1)	0.5636(8)	2.6704(2)	1.2898(4)	0.0565(1)	0.0944(3)
MQ		37.4401(9)	0.0278(1)	129.6031(8)	1.4181(7)	0.0586(5)	0.0982(5)
MQP		25.7858(8)	0.0887(2)	149.61(9)	1.3267(5)	0.0575(4)	0.0952(4)
CM		18.2977(7)	0.5521(7)	101.8159(6)	1.7443(9)	0.0593(6)	0.1025(7)
KS		1.4975(3)	0.567(9)	4.7768(3)	1.6296(8)	0.0599(8)	0.1029(8)
AD		17.184(6)	0.21(3)	107.5207(7)	1.2354(3)	0.057(3)	0.094(2)
ADR		39.3189(10)	0.6077(10)	168.0948(10)	2.2416(10)	0.062(9)	0.1091(9)
ADL		8.3735(5)	0.2869(5)	86.1526(5)	1.2155(2)	0.0569(2)	0.094(1)
AD2L		1.7689(4)	0.3971(6)	12.092(4)	1.3453(6)	0.0597(7)	0.1001(6)
BAYES		-0.6444(2)	0.2867(4)	1.1438(1)	1.0079(1)	0.0896(10)	0.1400(10)
MV	60	0.0968(1)	0.1799(10)	1.116(2)	0.6(2)	0.0331(1)	0.0544(1)
MQ		2.5983(9)	0.0113(1)	17.3432(9)	0.7993(7)	0.0349(5)	0.0591(5)
MQP		0.6225(5)	0.0691(2)	2.3326(5)	0.6977(6)	0.0338(3)	0.0565(4)
CM		1.2985(8)	0.1757(8)	10.4285(8)	0.8546(8)	0.035(6)	0.0599(6)
KS		0.7681(7)	0.1702(7)	2.654(6)	0.8758(9)	0.036(7)	0.0618(8)
AD		0.6803(6)	0.0719(3)	3.894(7)	0.6471(5)	0.0336(2)	0.0557(3)
ADR		5.9195(10)	0.1696(6)	40.5557(10)	1.0311(10)	0.0365(8)	0.0641(10)
ADL		0.3532(2)	0.0974(5)	1.5779(3)	0.6149(3)	0.0338(4)	0.0557(2)
AD2L		0.4536(3)	0.0766(4)	1.748(4)	0.6389(4)	0.037(9)	0.0608(7)
BAYES		-0.4915(4)	0.1785(9)	0.6250(1)	0.3838(1)	0.0394(10)	0.0624(9)
MV	100	0.0475(1)	0.098(9)	0.8092(2)	0.4141(2)	0.0251(1)	0.0413(1)
MQ		0.7414(9)	-0.0164(2)	4.0347(9)	0.5651(7)	0.0266(5)	0.0448(5)
MQP		0.292(3)	0.0275(4)	1.2286(4)	0.4793(6)	0.0256(3)	0.0426(4)
CM		0.4212(7)	0.0801(7)	2.4695(8)	0.5852(8)	0.0266(6)	0.0452(6)
KS		0.4787(8)	0.0819(8)	1.8974(7)	0.6356(9)	0.0275(7)	0.0473(9)
AD		0.3(4)	0.0267(3)	1.3006(6)	0.4608(5)	0.0255(2)	0.0423(3)
ADR		2.1496(10)	0.0733(6)	20.2339(10)	0.737(10)	0.0278(8)	0.0488(10)
ADL		0.186(2)	0.0454(5)	1.0119(3)	0.4329(3)	0.0257(4)	0.0423(2)
AD2L		0.3219(5)	0.0117(1)	1.2407(5)	0.455(4)	0.0286(9)	0.0469(8)
BAYES		-0.3649(6)	0.1300(10)	0.5274(1)	0.3101(1)	0.0293(10)	0.0466(7)
MV	200	0.0183(1)	0.047(10)	0.5455(2)	0.2739(2)	0.0171(1)	0.0278(1)
MQ		0.2597(9)	-0.0183(4)	0.9499(8)	0.3945(7)	0.0183(5)	0.0309(6)
MQP		0.1216(4)	0.0126(2)	0.6785(4)	0.3271(6)	0.0175(3)	0.029(4)
CM		0.1623(6)	0.0293(8)	0.8694(7)	0.4003(8)	0.0183(6)	0.031(7)
KS		0.2183(7)	0.0337(9)	1.033(9)	0.4492(9)	0.0192(9)	0.0329(8)
AD		0.1293(5)	0.0089(1)	0.6898(5)	0.3199(5)	0.0175(2)	0.0289(2)
ADR		0.3444(10)	0.0254(7)	1.4647(10)	0.5035(10)	0.0192(8)	0.0334(9)
ADL		0.0808(3)	0.0199(5)	0.6121(3)	0.2975(3)	0.0177(4)	0.029(3)
AD2L		0.2192(8)	-0.0214(6)	0.8541(6)	0.3103(4)	0.0207(10)	0.0338(10)
BAYES		-0.0636(2)	0.0154(3)	0.4968(1)	0.2674(1)	0.0189(7)	0.0304(5)

Tabela 3.8: Resultado da simulação para os parâmetros α=5 e θ=5

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.662(1)	3.897(8)	8.5017(3)	7.9157(5)	0.0558(2)	0.0944(4)
MQ		137.7753(10)	1.2904(2)	313.2926(7)	8.6794(8)	0.0571(5)	0.0982(6)
MQP		135.215(9)	1.3454(3)	394.8529(9)	8.2513(7)	0.0562(4)	0.0952(3)
CM		76.2086(5)	4.3965(9)	266.2909(4)	11.6518(9)	0.0586(8)	0.1025(8)
KS		2.5287(3)	2.9715(7)	7.9618(2)	7.9591(6)	0.0573(6)	0.1029(5)
AD		86.1423(7)	1.7243(4)	295.6509(6)	7.2827(3)	0.0558(1)	0.094(1)
ADR		102.9408(8)	5.3604(10)	277.8809(5)	15.7864(10)	0.0614(9)	0.1091(9)
ADL		77.7715(6)	2.0429(5)	328.7321(8)	7.1767(2)	0.0559(3)	0.094(2)
AD2L		62.9742(4)	2.6679(6)	403.064(10)	7.8382(4)	0.0585(7)	0.1001(7)
BAYES		-2.1239(2)	0.9880(1)	2.8359(1)	3.2176(1)	0.1112(10)	0.1665(10)
MV	60	0.2842(1)	1.2237(7)	3.5503(2)	3.3063(3)	0.0326(1)	0.0544(1)
MQ		29.4129(9)	0.4024(2)	92.1025(9)	4.7756(8)	0.0344(5)	0.0591(5)
MQP		10.7627(7)	0.6108(5)	54.174(6)	4.1314(6)	0.0334(3)	0.0565(4)
CM		17.4751(8)	1.3699(9)	68.2378(8)	5.2922(9)	0.0347(6)	0.0599(8)
KS		1.8329(3)	1.2628(8)	5.7987(3)	4.5075(7)	0.0348(7)	0.0618(6)
AD		10.7433(6)	0.5384(3)	57.1575(7)	3.6014(5)	0.0331(2)	0.0557(2)
ADR		32.5566(10)	1.6004(10)	97.8538(10)	6.5201(10)	0.036(8)	0.0641(9)
ADL		5.6667(5)	0.6448(6)	40.6521(5)	3.3949(4)	0.0334(4)	0.0557(3)
AD2L		1.8595(4)	0.5978(4)	8.0566(4)	3.2766(2)	0.0363(9)	0.0608(7)
BAYES		-0.9673(2)	0.3620(1)	2.1741(1)	2.2877(1)	0.0567(10)	0.0865(10)
MV	100	0.135(1)	0.6755(8)	2.5368(2)	2.1738(2)	0.0247(1)	0.0413(1)
MQ		12.8024(9)	0.0604(1)	55.8416(9)	3.299(7)	0.0263(5)	0.0448(5)
MQP		2.3525(6)	0.2747(5)	13.3374(6)	2.7212(6)	0.0253(3)	0.0426(4)
CM		7.8347(8)	0.6364(7)	39.6913(8)	3.4952(9)	0.0264(6)	0.0452(7)
KS		1.6123(5)	0.6844(9)	5.2205(4)	3.3265(8)	0.0266(7)	0.0473(8)
AD		3.5319(7)	0.2226(4)	21.4179(7)	2.5065(5)	0.0252(2)	0.0423(3)
ADR		16.9984(10)	0.7487(10)	59.0158(10)	4.4635(10)	0.0275(8)	0.0488(9)
ADL		1.2626(4)	0.3114(6)	5.5096(5)	2.3242(4)	0.0254(4)	0.0423(2)
AD2L		1.0491(3)	0.1939(2)	3.8795(3)	2.1777(3)	0.0278(9)	0.0469(6)
BAYES		-0.6395(2)	0.1673(3)	1.8248(1)	1.7832(1)	0.0407(10)	0.0621(10)
MV	200	0.0503(1)	0.3265(10)	1.713(2)	1.3721(2)	0.0168(1)	0.0278(1)
MQ		2.8013(9)	-0.0353(2)	13.1723(9)	2.3005(7)	0.0181(5)	0.0309(5)
MQP		0.6517(5)	0.1395(5)	2.7648(5)	1.7942(6)	0.0173(3)	0.029(4)
CM		1.88(8)	0.2518(7)	9.5622(8)	2.3598(8)	0.0181(6)	0.031(6)
KS		1.1194(7)	0.2917(9)	3.7878(7)	2.4316(9)	0.0188(7)	0.0329(7)
AD		0.7488(6)	0.0911(4)	3.1147(6)	1.7195(5)	0.0172(2)	0.0289(2)
ADR		6.0125(10)	0.2742(8)	25.9292(10)	3.0263(10)	0.019(8)	0.0334(9)
ADL		0.4469(3)	0.1449(6)	2.3202(3)	1.5725(4)	0.0174(4)	0.029(3)
AD2L		0.6464(4)	-0.0172(1)	2.4708(4)	1.4348(3)	0.0201(9)	0.0338(8)
BAYES		-0.3214(2)	0.0884(3)	1.4849(1)	1.3293(1)	0.0253(10)	0.0389(10)

próximos. Já o que teve os piores resultados foi o Anderson Darling Calda a Direita. A diferença máxima e diferença absoluta tiveram resultados similares, onde para $n = 20$ o melhor estimador foi o Anderson Darling, e para n grande o bayesiano foi mais eficiente.

3.10 Aplicação

Foi realizada uma aplicação para testar o ajustamento da nova distribuição GEEMO num certo banco de dados, onde também foi comparado o ajustamento desta distribuição com outras já existentes. O banco de dados utilizado foi obtido a partir de Box e Cox (1964), onde é apresentado o tempo de vida de 48 animais expostos a certos tipos de venenos e em seguida a certos tipos de tratamentos. É possível notar, observando a Figura (3.7), que o risco é crescente.

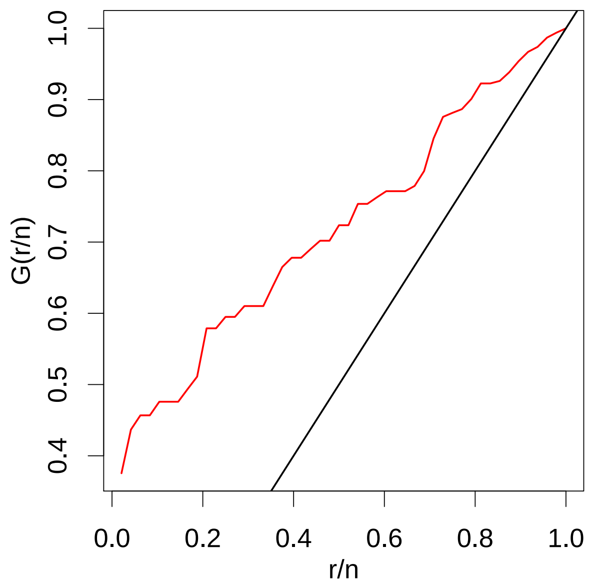


Figura 3.7: Gráfico TTT Plot.

As distribuições utilizadas para fins comparativos e seus respectivos parâmetros estimados pelo método da verossimilhança foram a GEEMO com $\alpha = 0.0036$ e $\theta = 2.0509$, Gama Exponencial (GE) com $\theta = 0.3681$, Nadarajah Haghighi Exponencial (NHE) com $\alpha = 0.0090$ e $\beta = 153.0002$, Weibull com $\alpha = 2.0606$ e $\beta = 0.5445$, Exponencial Power (EP) com $\lambda = 0.7603$ e $\theta = 1.3798$ e também a distribuição Gama com $\beta = 0.1112$ e $\alpha = 4.3079$.

Em problemas que envolvem incerteza, utilizar critérios que auxiliem na comparação e seleção de modelos paramétricos é de fundamental importância numa análise de dados. Alguns critérios comuns na literatura podem ser utilizados para seleção de modelos. Esses critérios levam em consideração a complexidade do modelo. São critérios que, essencialmente, penalizam a função verossimilhança pelo número de parâmetros do modelo e, eventualmente, o tamanho da amostra. Essa penalização é feita subtraindo-se do valor da função verossimilhança uma determinada quantidade que expressa o quão complexo é o modelo (JUNIOR, 2013).

Akaike (1974) propôs utilizar a informação de Kullback-Leibler para a seleção de modelos. O autor estabeleceu uma relação entre a função de verossimilhança maximizada e a informação de Kullback-Leibler desenvolvendo, então, o chamado critério de informação

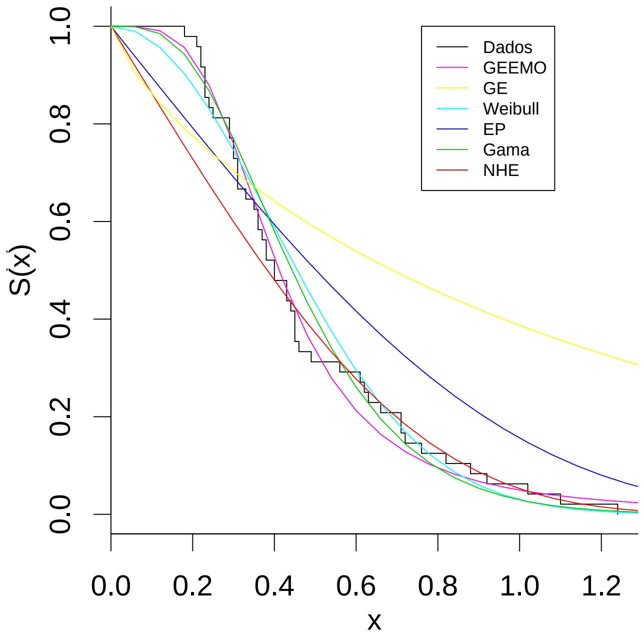


Figura 3.8: Gráfico dos dados e das distribuições ajustadas.

de Akaike (AIC). Liang e Zou (2008) propuseram uma versão melhorada (CAIC), em que é considerada uma ponderação entre o número de parâmetros e o tamanho amostral.

O critério de informaçãao bayesiano (BIC), também chamado de critério de Schwarz et al. (1978), é um critério de avaliação de modelos definido em termos da probabilidade a posteriori, sendo assim chamado porque Schwarz forneceu um argumento bayesiano para prová-lo. Segundo esses critérios o melhor modelo será aquele que apresentar menor valor para AIC, BIC e CAIC

$$\begin{aligned} AIC &= -2l(\theta) + 2d \\ BIC &= -2l(\theta) + \log(n) \\ CAIC &= AIC + \frac{2(d+2)(d+3)}{n-d-3} \end{aligned} \tag{3.73}$$

Tabela 3.9: Critérios de seleção do modelo

Modelo	AIC	BIC	CAIC
GEEMO	-9.709	-5.967	-8.779
NHE	12.762	16.504	13.692
Gama	-8.340	-4.5980	-7.409
GE	61.362	63.234	61.908
EP	5.672	9.414	6.602
Weibull	-2.506	1.236	-1.575

É possível notar na tabela (3.9) que a distribuição GEEMO teve um bom ajuste no conjunto de dados proposto. Seu resultado foi melhor, obviamente que seu caso particular, a GE, foi melhor também que as outras distribuições de dois parâmetros, tanto em relação a distribuições mais conhecidas na literatura, como em relação a outras distribuições propostas mais recentemente. Desta forma é possível notar a flexibilidade do modelo proposto e este passa a ser uma boa alternativa para realizar ajustes.

Capítulo

4

Gama Exponenciada Poisson Truncada no Zero

Como no capítulo 3, foi feito a composição entre as funções utilizando a teoria dos riscos competitivos e riscos complementares para a criação de novas distribuições. Agora também será proposto uma nova distribuição, onde a distribuição utilizada foi a Poisson Truncada no Zero.

O truncamento de uma distribuição é definido por

$$P(Y=y|y>0)=\frac{P(Y=y)}{1-P(Y=0)}, \quad y=1,2,\ldots$$

A distribuição de Poisson tem um único parâmetro desconhecido. O parâmetro pode ser interpretado como a média aritmética de ocorrências por intervalo de tempo ou espaço, o qual caracteriza o processo que gera a distribuição de Poisson. Por exemplo, ao se coletar os dados da produtividade científica não é possível observar a produtividade dos autores, diga-se, os improdutivos (ALVARADO, 2008). Isto é, existem situações em que algumas frequências não são verificadas, como no caso dos autores com zero colaborações num determinado período. Nessas situações, tem-se uma distribuição de Poisson truncada no zero, que é definida por:

$$P(M=m)=\frac{\lambda^m e^{-\lambda}}{m!(1-e^{-\lambda})}.$$

A partir da distribuição Gama Exponenciada foi obtido uma nova, a distribuição Gama Exponenciada Poisson Truncada no Zero, pelo método da composição. Considere uma amostra aleatória $Y_1, Y_2, \cdots, Y_M$ de uma distribuição Gama Exponenciada com função densidade de probabilidade (fdp) dada por (2.2).

A seguir, esta distribuição será apresentada considerando os casos de riscos competitivos (mínimo) e riscos complementares (máximo).

4.1 Composição entre distribuições

4.2 Riscos Competitivos

Seja $X=min(Y_1,Y_2,\cdots,Y_M)$, supondo que Y e M são variáveis aleatórias independentes e além disso, M é o tamanho da amostra, pode-se escrever a função densidade de probabilidade de X dado $M=m$ como sendo:

$$f(x|M=m,\theta)=m[S_Y(x|\theta)]^{m-1}f_Y(x|\theta)$$

Daí segue que a fdp marginal de X é obtida da seguinte forma

$$\begin{aligned} f(x|\theta,\lambda) &= \sum_{m=1}^{\infty} f(x|M=m,\theta)P(M=m) \\ &= \sum_{m=1}^{\infty} m[S(x|\theta)]^{m-1} f_Y(x|\theta) \frac{\lambda^m e^{-\lambda}}{m!(1-e^{-\lambda})} \\ &= \frac{\lambda e^{-\lambda F(y|\theta)} f(y|\theta)}{(1-e^{-\lambda})}. \end{aligned} \quad (4.1)$$

Sendo assim, a fdp marginal de $X = \min(Y_1, Y_2, ..., Y_M)$ é dada por

$$f(x|\theta,\lambda) = \frac{\lambda e^{-\lambda(1-e^{-x}(x+1))^{\theta}} \theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1}}{(1-e^{-\lambda})}, \quad 0<\theta, \quad 0<\lambda, \quad 0<x \quad (4.2)$$

A distribuição de $X = \min(Y_1, Y_2, ..., Y_M)$, apresentada acima na expressão (4.2) é dita distribuição Gama Exponenciada Poisson Truncada no Zero ($\text{GEPTZ}_{\text{min}}$).

A função acumulada, de sobrevivência e a de risco são dadas respectivamente por

$$\begin{aligned} F(y|\theta,\lambda) &= \frac{1-e^{-\lambda F(y|\theta)}}{(1-e^{-\lambda})}, \\ S(y|\theta,\lambda) &= \frac{e^{-\lambda F(y|\theta)} - e^{-\lambda}}{(1-e^{-\lambda})}, \\ h(y|\theta,\lambda) &= \frac{\lambda e^{-\lambda F(y|\theta)} f(y|\theta)}{e^{-\lambda F(y|\theta)} - e^{-\lambda}}. \end{aligned} \quad (4.3)$$

4.3 Riscos Complementares

Seja $X = \max(Y_1, Y_2, ..., Y_M)$, onde Y e M são variáveis aleatórias independentes, pode-se escrever a função densidade de probabilidade de X dado $M = m$ como sendo

$$f(x|M=m,\theta) = m[F_Y(x|\theta)]^{m-1} f_Y(x|\theta).$$

Procedendo similarmente ao que foi feito no caso de riscos competitivos, tem-se

$$\begin{aligned} f(y|\theta,\lambda) &= \sum_{m=1}^{\infty} m[F(y|\theta)]^{m-1} f(y|\theta) \frac{\lambda^m e^{-\lambda}}{m!(1-e^{-\lambda})} \\ &= \frac{\lambda e^{-\lambda S(y|\theta)} f(y|\theta)}{1-e^{-\lambda}}, \end{aligned} \quad (4.4)$$

logo, a fdp marginal de $X = \max(Y_1, Y_2, ..., Y_M)$ é dada por

$$f(x|\theta,\lambda) = \frac{\lambda e^{-\lambda(1-(1-e^{-x}(x+1))^{\theta})} \theta x e^{-x}(1-e^{-x}(x+1))^{\theta-1}}{1-e^{-\lambda}}, \quad 0<\theta, \quad 0<\lambda, \quad 0<x. \quad (4.5)$$

A distribuição de $X = \max(Y_1, Y_2, ..., Y_M)$, apresentada acima na expressão 4.5, é dita distribuição Gama Exponenciada Poisson Tuncada no Zero ($\text{GEPTZ}_{\text{max}}$).

A função acumulada, de sobrevivência e a de risco são dadas respectivamente por

$$\begin{aligned} F(y|\theta,\lambda) &= \frac{e^{-\lambda S(y|\theta)} - e^{-\lambda}}{1-e^{-\lambda}}, \\ S(y|\theta,\lambda) &= \frac{1-e^{-\lambda S(y|\theta)}}{1-e^{-\lambda}}, \\ h(y|\theta,\lambda) &= \frac{\lambda e^{-\lambda S(y|\theta)} f(y|\theta)}{1-e^{-\lambda F(y|\theta)}}. \end{aligned} \quad (4.6)$$

4.4 Caso geral

Seguindo Santo (2014), substituindo λ por $-\lambda$ na função de sobrevivência dos riscos competitivos tem-se a mesma expressão de sobrevivência obtida no caso dos riscos complementares. Logo, é possível obter um caso geral para o modelo Gama Exponenciada Poisson Truncada no Zero (GEPTZ).

A nova distribuição tem a seguinte densidade de probabilidade

$$f(x|\theta,\alpha) = \frac{\alpha e^{-\alpha F(x|\theta)} f(x|\theta)}{1-e^{-\alpha}}, \tag{4.7}$$

com $\alpha = \lambda > 0$ se $X = \min(Y_1, Y_2, ..., Y_M)$ e $\alpha = -\lambda < 0$ se $X = \max(Y_1, Y_2, ..., Y_M)$.

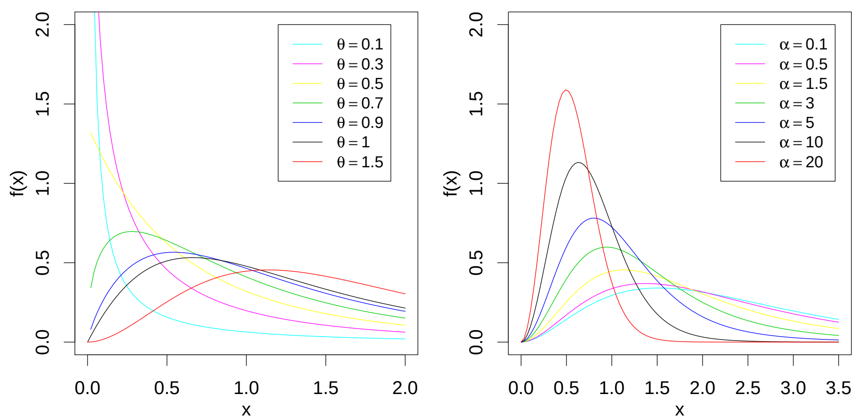
Portanto são obtidas as seguintes funções

$$\begin{aligned} F(x|\theta,\alpha) &= \frac{1-e^{-\alpha F(x|\theta)}}{1-e^{-\alpha}}, \\ S(x|\theta,\alpha) &= \frac{e^{-\alpha F(x|\theta)}-e^{-\alpha}}{1-e^{-\alpha}}, \\ h(x|\theta,\alpha) &= \frac{\alpha e^{-\alpha F(x|\theta)} f(x|\theta)}{e^{-\alpha F(x|\theta)}-e^{-\alpha}}. \end{aligned} \tag{4.8}$$

Logo, a função densidade no caso geral é dada por

$$f(x|\theta,\alpha) = \frac{\alpha e^{-\alpha(1-e^{-x}(x+1))^{\theta}} \theta x e^{-x} (1-e^{-x}(x+1))^{\theta-1}}{1-e^{-\alpha}}. \tag{4.9}$$

E a Figura 4.1 esboça o comportamento da função densidade quando se varia o valor de θ e de α.



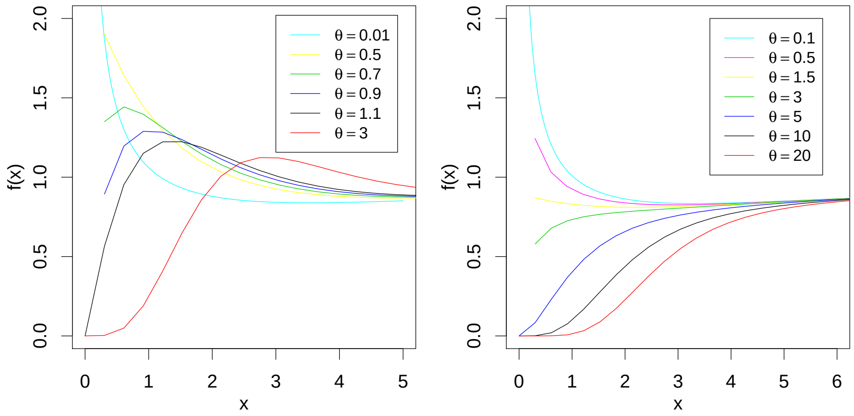
(a) $\alpha = 1.5$ (b) $\theta = 1.5$

Figura 4.1: Gráfico da função densidade para diferentes valores de θ e α.

A função risco da distribuição Gama Exponencial Poisson Truncada no Zero é dada por

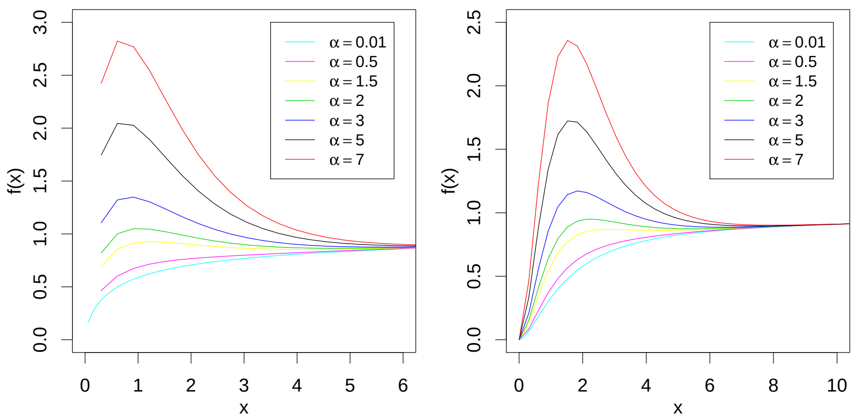
$$h(x|\theta,\alpha) = \frac{\alpha e^{-\alpha(1-(1-e^{-x}(x+1))^{\theta})} \theta x e^{-x} (1-e^{-x}(x+1))^{\theta-1}}{e^{-\alpha(1-(1-e^{-x}(x+1))^{\theta})}-e^{-\alpha}}. \tag{4.10}$$

A Figura 4.2 esboça o comportamento da função risco quando se varia o valor de α, enquanto a Figura 4.3 mostra o comportamento da função risco quando o valor de α é fixado e os valores de θ são variados. É possível observar que o novo modelo possui bastante flexibilidade, mesmo possuindo apenas dois parâmetros.



(a) $\alpha = 3$ (b) $\theta = 0.5$

Figura 4.2: Gráfico da função risco para diferentes valores de θ.



(a) $\theta = 0.8$ (b) $\alpha = 1.5$

Figura 4.3: Gráfico da função risco para diferentes valores de α.

4.5 Propriedades Estatísticas e Matemáticas da distribuição GEPTZ

Teorema 3 *Seja X uma variável aleatória da distribuição GEPTZ com parâmetros α e θ, então a função densidade (4.9) pode ser escrita na seguinte forma*

$$f(x|\theta,\alpha) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!}\left(\begin{array}{c}\theta(i+1)-1\\ j\end{array}\right)\left(\begin{array}{c}j\\ k\end{array}\right)(-1)^j e^{-x(j+1)}x^{1+k}.$$

Demonstração. Baseado na expansão em série de Taylor,

$$e^x = \sum_{i=0}^{\infty}\frac{x^i}{i!}$$

se obtém

$$f(x|\theta,\alpha) = \sum_{i=0}^{\infty}\frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!}xe^{-x}(1-e^{-x}(x+1))^{\theta i+\theta-1}.$$

Em seguida, utilizando expansão em série, com $|z| < 1$,

$$(1-z)^{k-1} = \sum_{m=0}^{\infty} (-1)^m \binom{k-1}{m} z^m$$

a função é escrita como

$$f(x|\theta,\alpha) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \binom{\theta(i+1)-1}{j} (-1)^j x e^{-x-xj}(x+1)^j$$

logo, utilizando a expansão binomial, a função pode ser reescrita da seguinte forma

$$f(x|\theta,\alpha) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \binom{\theta(i+1)-1}{j} \binom{j}{k} (-1)^j e^{-x(j+1)} x^{1+k}.$$

Esta forma de escrever a função densidade facilita a obtenção de suas integrais, para cálculos como o n-ésimo momento, a função geradora de momentos entre outros. ■

4.5.1 N-ésimo momento

Os momentos são muito importantes em Estatística para caracterizar distribuições de probabilidade. Para o cálculo do N-ésimo momento da distribuição GEPTZ será utilizado o resultado do Teorema 3, pois através dele é possível resolver a seguinte integral:

$$E(X^n) = \int_0^{\infty} x^n f(x|\theta,\alpha)dx,$$

então pode-se escrever essa função da seguinte forma:

$$E(X^n) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \binom{\theta(i+1)-1}{j} \binom{j}{k} (-1)^j \int_0^{\infty} e^{-x(j+1)} x^{1+k+n} dx,$$

e assim é obtido

$$E(X^n) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \binom{\theta(i+1)-1}{j} \binom{j}{k} (-1)^j \times (j+1)^{-2-k-n}\Gamma(2+k+n). \tag{4.11}$$

4.5.2 Função Geradora de Momentos

A função geradora de momentos (fgm) de uma variável aleatória é uma especificação alternativa de sua distribuição de probabilidade. Assim, ela fornece a base de uma forma alternativa para os resultados analíticos em comparação com resultados diretamente obtidos com funções densidade de probabilidade.

Seja X uma variável aleatória. A função geradora de momentos da variável X é definida como sendo a função

$$M_X(t) = E(e^{tX}),$$

desde que $E(e^{tX})$ exista em algum intervalo do tipo $(-h,h)$ para algum número real $h > 0$.

A fgm possui esse nome pois, a partir dela, podemos encontrar todos os momentos da variável aleatória X (quando estes existem). A função só precisa estar definida em uma vizinhança do ponto zero, pois os momentos serão obtidos através de sucessivas diferenciações aplicadas em zero (CURTISS, 1942). Utilizando o resultado da função do N-ésimo momento se chega na fgm da seguinte forma

$$M_X(t) = E(e^{tX}) = E\left(\sum_{r=0}^{\infty} \frac{t^r X^r}{r!}\right) = \sum_{r=0}^{\infty} \frac{t^r E(X^r)}{r!},$$

logo a fgm é dada por

$$\mathrm{M}_X(t) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{r=0}^{\infty}\sum_{k=0}^{j} \frac{\alpha(-\alpha)^i \theta t^r}{(1-e^{-\alpha})i!r!} \binom{\theta(i+1)-1}{j} \binom{j}{k} (-1)^j \times (j+1)^{-2-k-r}\Gamma(2+k+r). \tag{4.12}$$

4.5.3 Assimetria e Curtose

Medidas de assimetria e curtose são muitas vezes utilizadas para descrever características da forma de uma distribuição.

Na Tabela 4.1 é possível observar que a assimetria e a curtose, quando se tem um valor fixo para o parâmetro θ e varia o α, aumentam até um certo valor para α, depois começa a decrescer, o inverso ocorre para a variância.

Tabela 4.1: Tabelas Assimetria e Curtose

(a) θ=10

α	Assimetria	Curtose	Variância
0.01	1.060	5.017	2.228
0.1	1.076	5.072	2.200
0.5	1.148	5.352	2.065
1	1.240	5.773	1.878
1.5	1.329	6.262	1.683
2	1.410	6.795	1.489
3	1.517	7.844	1.137
5	1.410	8.610	0.661
10	0.512	4.443	0.290

(b) α=10

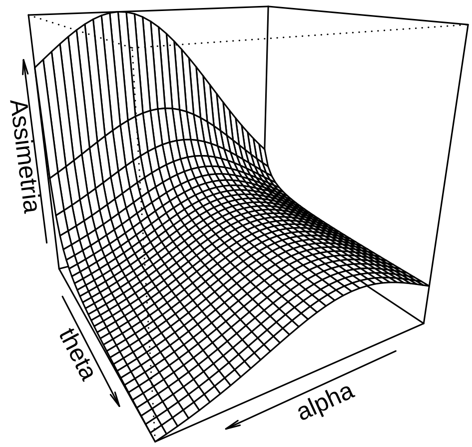
θ	Assimetria	Curtose	Variância
0.01	514.933	371287.900	0.000
0.1	82.694	12405.040	0.001
0.5	3.822	43.495	0.034
1	1.732	11.091	0.106
1.5	1.229	7.354	0.158
2	1.001	6.130	0.193
3	0.786	5.232	0.234
5	0.624	4.716	0.268
10	0.512	4.443	0.290

Na Figura 4.4 tem-se que, quando θ é menor que 1 o valor da assimetria e da curtose, principalmente, aumentam bastante. Também é possível observar que tanto a curtose quanto a assimetria têm picos, ou seja, elas vão aumentando até um certo valor dos parâmetros e depois diminuem.

4.5.4 R-ésimo momento da Vida Média Residual

O r-ésimo momento da vida residual média é dado pela seguinte fórmula

$$\mu_r(t) = E\left((X-t)^r | X > t\right) = \frac{\int_t^{\infty} (x-t)^r f(x)dx}{S(t)}. \tag{4.13}$$



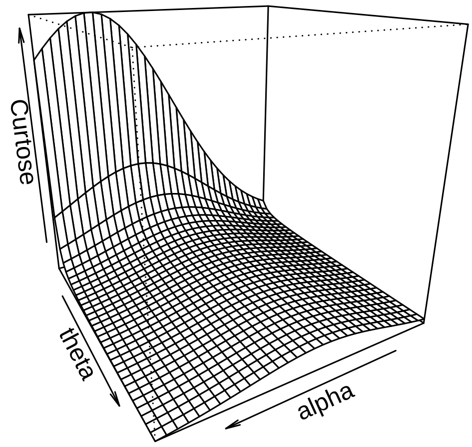


Figura 4.4: Gráfico da Assimetria e Curtose.

Substituindo a função densidade, $f(x)$, da fórmula do r-ésimo momento da vida residual média pelo resultado obtido no teorema 3, o $\mu_r(t)$ pode ser escrito da seguinte forma

$$\mu_r(t) = \frac{1}{S(t)} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{j} \frac{\alpha(-\alpha)^i \theta}{(1-e^{-\alpha}) i!} \binom{\theta(i+1)-1}{j} \binom{j}{k} (-1)^j \times \int_t^{\infty} e^{-x(j+1)} x^{1+k} (x-t)^r dx. \tag{4.14}$$

Agora, utilizando a série binomial apresentada em Zakerzadeh e Mahmoudi (2012), chega-se a expressão

$$\mu_r(t) = \frac{1}{S(t)} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{j} \sum_{m=0}^{r} \frac{\alpha(-\alpha)^i \theta}{(1-e^{-\alpha}) i!} \binom{\theta(i+1)-1}{j} \binom{j}{k} \binom{r}{m} (-1)^{j+m} t^m \times \int_t^{\infty} e^{-x(j+1)} x^{1+k+r-m} dx. \tag{4.15}$$

Deste modo chega-se no resultado final

$$\mu_r(t) = \frac{1}{S(t)} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j}\sum_{m=0}^{r} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \left(\begin{array}{c}\theta(i+1)-1\\ j\end{array}\right)\left(\begin{array}{c}j\\ k\end{array}\right)\left(\begin{array}{c}r\\ m\end{array}\right)(-1)^{j+m}t^m$$
$$\times (i+1)^{-2-k-r+m}\Gamma(2+k+r-m, t(j+1)), \quad r \geq 1, \tag{4.16}$$

onde $\Gamma(s;t) = \int\limits_t^{\infty} x^{s-1}e^{-x}dx$ é a função Gama incompleta superior.

A vida média residual é um caso particular do r-ésimo momento da vida média residual, isso ocorre quando $r = 1$, também conhecida como *MRL*, (*Mean Residual Life*), é a vida remanescente $(X - t)$ esperada para a unidade, dado que no tempo t a unidade estava operante.

$$\mu(t) = E(X - t|X > t) = \frac{\int\limits_t^{\infty} S(x)dx}{S(t)}.$$

Utilizando a função 4.16 e fazendo $r = 1$ é obtido

$$\mu(t) = \frac{1}{S(t)} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \left(\begin{array}{c}\theta(i+1)-1\\ j\end{array}\right)\left(\begin{array}{c}j\\ k\end{array}\right)$$
$$\times \left((i+1)^{-2-k}(-1)^{j+1}\Gamma(2+k, t(j+1))t + (i+1)^{-3-k}(-1)^j\Gamma(3+k, t(j+1))\right). \tag{4.17}$$

Por meio do *MRL* é possível obter a média, igualando $t = 0$

$$\mu(0) = E(X) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \left(\begin{array}{c}\theta(i+1)-1\\ j\end{array}\right)\left(\begin{array}{c}j\\ k\end{array}\right)$$
$$\times \left((i+1)^{-3-k}(-1)^j\Gamma(3+k)\right). \tag{4.18}$$

4.5.5 R-ésimo momento da Vida Média Residual Invertido

O tempo decorrido desde o fracasso de um item na condição de que essa falha tenha ocorrido em $[0, t]$ é conhecido na literatura como a vida média residual revertida. Também é referido pelo tempo médio de inatividade. O r-ésimo momento da vida média residual invertido é dada por

$$m_r(t) = E\left((t - X)^r|X \leq t\right) = \frac{\int\limits_0^{t} (t-x)^r f(x)dx}{F(t)}.$$

Recorrendo a conceitos similares aos utilizados em 4.16 é obtido a seguinte função do r-ésimo momento da vida média residual invertido

$$m_r(t) = \frac{1}{F(t)} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j}\sum_{m=0}^{r} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \left(\begin{array}{c}\theta(i+1)-1\\ j\end{array}\right)\left(\begin{array}{c}j\\ k\end{array}\right)\left(\begin{array}{c}r\\ m\end{array}\right)(-1)^{j+r-m}t^m$$
$$\times (i+1)^{-2-k-r+m}\gamma(2+k+r-m, t(j+1)), \quad r \geq 1, \tag{4.19}$$

onde $\gamma(s;t) = \int\limits_0^{t} x^{s-1}e^{-x}dx$ é a função Gama incompleta inferior.

Mais detalhes podem ser vistos em Nanda et al. (2003), Kundu e Nanda (2010). Assim, o tempo de vida médio passado *MPL*, (*mean past lifetime*) é um caso particular com r-ésimo momento da vida média residual invertido com $r = 1$, é dado pela seguinte expressão

$$m(t) = \frac{1}{F(t)} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{j} \frac{\alpha(-\alpha)^i \theta}{(1-e^{-\alpha})i!} \begin{pmatrix} \theta(i+1)-1 \\ j \end{pmatrix} \begin{pmatrix} j \\ k \end{pmatrix}$$
$$\times \left((i+1)^{-2-k}(-1)^j \gamma(2+k, t(j+1))t + (i+1)^{-3-k}(-1)^{j+1}\gamma(3+k, t(j+1)) \right). \tag{4.20}$$

4.5.6 Ordenação Estocástica

Ordenação das distribuições, particularmente entre distribuições que modelam tempo de vida desempenham um papel importante na literatura estatística. Foram considerados seis ordens estocásticas diferentes, a usual, a taxa de risco, a taxa de risco invertida, a média, a vida média residual e a ordem da razão de verossimilhança, para duas variáveis aleatórias GEPTZ independentes.

Se X e Y são variáveis aleatórias independentes com funções distribuições acumuladas (fda) F_X e F_Y respectivamente, então X é dito ser menor que Y em:

- ordem estocástica $(X \leq_{st} Y)$ se $F_X(x) \geq F_Y(x)$ $\forall x$;
- ordem da taxa de risco $(X \leq_{hr} Y)$ se $h_X(x) \geq h_Y(x)$ $\forall x$;
- ordem da taxa de risco reversa $(X \leq_{rh} Y)$ se $P(t - X > x | X \leq t) \geq P(t - Y > y | Y \leq t), \forall x \geq 0, \quad \forall t$;
- ordem da vida média residual $(X \leq_{mrl} Y)$ se $\mu_X(x) \leq \mu_Y(x)$ $\forall x$;
- ordem da razão da verossimilhança $(X \leq_{lr} Y)$ se $\frac{f_X(x)}{f_Y(x)}$ decresce em x.

Seguindo Kochar (2012), tem-se a seguinte cadeia de implicações entre as ordens estocásticas:

$$\begin{array}{ccccc} X \leq_{lr} Y & \Rightarrow & X \leq_{hr} Y & \Rightarrow & X \leq_{mrl} Y \\ \Downarrow & & & & \Downarrow \\ X \leq_{rh} Y & \Rightarrow & X \leq_{st} Y & \Rightarrow & E[X] \leq E[Y] \end{array} \tag{4.21}$$

Teorema 4 *Seja* $X \sim GEPTZ(\alpha_1, \theta_1)$ *e* $Y \sim GEPTZ(\alpha_2, \theta_2)$. *Se* $\theta_1 = \theta_2 = \theta$ *e* $\alpha_2 \leq \alpha_1$, *então* $(X \leq_{st} Y)$, $(X \leq_{hr} Y)$, $(X \leq_{rh} Y)$, $(X \leq_{mrl} Y)$, $(X \leq_{lr} Y)$ *e* $E[X] \leq E[Y]$.

Demonstração. Dada a a razão da verossimilhança

$$\frac{f_X(x)}{f_Y(x)}$$

logo,

$$\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} \leq 0,$$

para $\theta_1 = \theta_2 = \theta$ e $\alpha_2 > \alpha_1$, implicando que $(X \leq_{st} Y)$, então $(X \leq_{hr} Y)$, $(X \leq_{rh} Y)$, $(X \leq_{mrl} Y)$, $(X \leq_{lr} Y)$ e $E[X] \leq E[Y]$ (detalhes das expressões deste teorema estão no Apêndice A). ■

4.5.7 Entropia

Como apresentado no capítulo anterior, a entropia de uma variável aleatória mede a variação da incerteza. Um grande valor de entropia indica a maior incerteza nos dados. Algumas medidas de entropia populares são entropia de Rényi (RÉNYI, 1961) e entropia de Shannon (SHANNON, 1951).

Entropia de Rényi

A entropia de Rényi é uma medida de variação de incerteza que tem sido usada em aplicações e caracterizações de muitas distribuições de probabilidade. Nessa seção, deduzimos uma expressão para esta medida. Para a função de densidade $f(x)$, a entropia de Rényi é definida por

$$H_I(\delta) = \frac{1}{1-\delta} \ln \left(\int_0^\infty f^\delta(x)dx \right), \quad \delta > 0, \quad \delta \neq 1. \tag{4.22}$$

Utilizando as mesmas expansões em série do Teorema 1 foi obtido a seguinte expressão

$$\begin{aligned} H_I(\delta) = {} & \frac{1}{(1-e^{-\alpha})^\delta} \alpha^\delta \theta^\delta \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{j} (-\alpha\delta)^i \binom{\theta i + \delta(\theta - 1)}{j} \binom{j}{k} (-1)^j \\ & \times (j+\delta)^{-1-k-\delta} \Gamma(1+k+\delta). \end{aligned} \tag{4.23}$$

Entropia de Shannon

A entropia de Shannon, chama atenção para o fato de que a entropia H não é função da variável aleatória X, mas sim da distribuição de probabilidade dessa variável. Em outras palavras, não dependente dos valores que X assume, mas das suas probabilidades (ARTUSO, 2012). Se X é uma variável aleatória contínua não-negativa com pdf $f(x)$, a entropia de Shannon é definida como

$$H(f) = E[-\ln f(x)] = -\int_0^\infty f(x) \ln(f(x))dx. \tag{4.24}$$

De acordo com Al-Zahrani et al. (2015), a entropia de Shannon é

$$\underset{\delta \to 1}{lim} H_\delta(x) = H(f).$$

O resultado da função acima se encontra no Apêndice A.

4.5.8 Desvios Médios

Se T é uma variável aleatória com distribuição GEPTZ com média $\mu = E(T)$ e mediana m, então o desvio médio em relação à média e o desvio médio em relação à mediana são definidos, respectivamente por

$$\vartheta_1 = \int_0^\infty |t-\mu| f(t)dt \qquad e \qquad \vartheta_2 = \int_0^\infty |t-m| f(t)dt. \tag{4.25}$$

Os desvios médios podem ser simplificados e ficam da seguinte maneira

$$\vartheta_1 = 2\mu F(\mu) - 2I(\mu) \qquad e \qquad \vartheta_2 = \mu + 2mF(m) - m - 2I(m). \tag{4.26}$$

onde $F(\mu)$ e $F(m)$ são obtidos por meio da expressão (4.8) e a função $I(m)$ é obtida por

$$I(m) = \int_0^m t f(t) dt.$$

Considerando $f(t)$ como a fdp definida no Teorema 4, tem-se

$$\begin{aligned} I(m) = & \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!}\begin{pmatrix}\theta(i+1)-1\\ j\end{pmatrix}\begin{pmatrix} j\\ k\end{pmatrix} \\ & \times\left((i+1)^{-3-k}(-1)^j\gamma(3+k, m(1+j))\right). \end{aligned} \tag{4.27}$$

O resultado é análogo para $I(\mu)$.

4.5.9 Curva de Bonferroni e de Lorenz

Curvas de Bonferroni e Lorenz foram propostas por Bonferroni (1930). São importantes aplicações dos desvios médios, estas curvas têm aplicações não só em economia para estudar renda e pobreza, mas também em outros campos, como confiabilidade, demografia, seguro e medicina. A curva de Bonferroni pode ser expressa por

$$B(p) = \frac{1}{p\mu}\int_0^q x f(x) dx. \tag{4.28}$$

Desta forma, obtém-se o seguinte resultado

$$\begin{aligned} B(p) = & \frac{1}{p\mu}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \frac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!}\begin{pmatrix}\theta(i+1)-1\\ j\end{pmatrix}\begin{pmatrix} j\\ k\end{pmatrix} \\ & \times\left((i+1)^{-3-k}(-1)^j\gamma(3+k, m(1+j))\right), \end{aligned} \tag{4.29}$$

onde $\mu = E(X)$ e $q = F^{-1}(p)$.

A curva de Lorenz é uma representação gráfica da função de distribuição cumulativa empírica da distribuição de probabilidade de riqueza. Em tal uso, muitos economistas consideram que é uma medida de desigualdade social para representar a desigualdade da distribuição da riqueza. As curvas de Lorenz são calculadas por:

$$L(p) = \frac{1}{\mu}\int_0^q x f(x) dx. \tag{4.30}$$

A obtenção da expressão da curva de Lorenz é análoga à de Bonferroni.

4.5.10 Estatística de Ordem

Sabe-se que $X_{(1)} \leq \cdots \leq X_{(n)}$ denota a ordem estatística da amostra aleatória $X_1 \leq \cdots \leq X_n$ da população contínua com fda $F(x)$ e fdp $f(x)$, então a pdf $X_{(j)}$ será dada por

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f(x)(F(x))^{j-1}(1-F(x))^{n-j},$$

para $j = 1, \ldots, n$. Então a fdp da j-ésima estatística de ordem da distribuição GEPTZ é dada por

$$f_{X_{(j)}}(x) = \sum_{i,k,l,m=0}^{\infty} \sum_{h=0}^{m} (-1)^{i+m+k+h} \frac{n!}{(j-1)!(n-j)!} \binom{n-j}{i} \binom{i+j-1}{k} \times \binom{\theta(1+l)-1}{m} \binom{m}{h} \frac{\theta\alpha}{(1-e^{-\alpha})^{i+j} l!} (-\alpha(1+k))^l x^{1+h} e^{-x(1+m)}. \quad (4.31)$$

O t-ésimo momento de $X_{j:n}$ pode ser expresso por

$$E[X_{j:n}^t] = \sum_{i,k,l,m=0}^{\infty} \sum_{h=0}^{m} (-1)^{i+m+k+h} \frac{n!}{(j-1)!(n-j)!} \binom{n-j}{i} \binom{i+j-1}{k} \times \binom{\theta(1+l)-1}{m} \binom{m}{h} \frac{\theta\alpha}{(1-e^{-\alpha})^{i+j} l!} (-\alpha(1+k))^l (1+m)^{-2-h-t} \Gamma(2+h+t). \quad (4.32)$$

4.5.11 Confiabilidade

Considerando X_1 e X_2 variáveis aleatórias independentes da distribuição GEPTZ com parâmetros (α_1, θ_1) e (α_2, θ_2) respectivamente, então, a forma algébrica da confiabilidade é dada por

$$R = P(X_2 < X_1) = \int_0^{\infty} f_1(x) F_2(x) dx.$$

Fazendo $\theta_1 = \theta_2 = \theta$, obtém-se a seguinte expressão

$$R = \frac{(\mathrm{e}^{\alpha_1}\alpha_2 - \alpha_1 - \alpha_2)\,\mathrm{e}^{\alpha_2} + \alpha_1}{((\alpha_1+\alpha_2)\,\mathrm{e}^{\alpha_1} - \alpha_1 - \alpha_2)\,\mathrm{e}^{\alpha_2} + (-\alpha_1 - \alpha_2)\,\mathrm{e}^{\alpha_1} + \alpha_1 + \alpha_2}. \quad (4.33)$$

4.6 Estimadores

Como no Capítulo anterior, foram utilizados dez diferentes métodos para se obter as estimativas dos parâmetros α e θ, sendo eles: Método de Máxima Verossimilhança, Mínimos Quadrados, Mínimos Quadrados Ponderados, Cramér-von-Mises, Anderson Darling, Anderson Darling - RT,Anderson Darling - LT, Anderson Darling - 2LT, Kolmogorov e também foi utilizado o método de estimação Bayesiana. Tais métodos serão apresentados a seguir.

Método de Máxima Verossimilhança

Seja X_i, $i = 1, 2, \ldots, n$, uma amostra aleatória com $X_i \sim GEPTZ(\alpha, \theta)$ a função de verossimilhança é dada por

$$L(x|\theta, \alpha) = \prod_{i=1}^{n} f(x|\theta, \alpha).$$

Por propriedades matemáticas da função de verossimilhança, maximizar a verossimilhança é o mesmo que maximizar o seu logaritmo. Fazendo $L(\alpha, \theta|x) = l(\alpha, \theta|x)$, tem-se que

$$l(\alpha, \theta|x) = -\alpha \sum_{i=1}^{n} \left(1 - \mathrm{e}^{-x_i}(x_i + 1)\right)^{\theta} + (\theta - 1) \sum_{i=1}^{n} \ln\left(1 - \mathrm{e}^{-x_i}(x_i + 1)\right) + n \ln(\alpha)$$
$$+ n \ln(\theta) + \sum_{i=1}^{n} \ln(x_i) - n \ln\left(1 - \mathrm{e}^{-\alpha}\right) - \sum_{i=1}^{n} x_i. \quad (4.34)$$

Derivando 4.34 em relação aos parâmetros α e θ, é obtido o vetor escore que é composto por

$$U_\alpha(\alpha, \theta|x) = \frac{\partial}{\partial \alpha} l(\alpha, \theta|x) = -\sum_{i=1}^{n} \left(1 - \mathrm{e}^{-x_i}(x_i + 1)\right)^{\theta} + \frac{n}{\alpha} - \frac{n\mathrm{e}^{-\alpha}}{1 - \mathrm{e}^{-\alpha}}, \quad (4.35)$$

$$U_\theta(\alpha, \theta|x) = \frac{\partial}{\partial \theta} l(\alpha, \theta|x) =$$
$$\frac{-\sum_{i=1}^{n} \alpha(1 + (-x_i - 1)\mathrm{e}^{-x_i})^{\theta} \ln(1 + (-x_i - 1)\mathrm{e}^{-x_i})\theta + \sum_{i=1}^{n} \ln(1 + (-x_i - 1)\mathrm{e}^{-x_i})\theta + n}{\theta}. \quad (4.36)$$

Os estimadores de Máxima Verossimilhança (EMV) $\hat{\alpha}_{EMV}$ e $\hat{\theta}_{EMV}$ para os parâmetros α e θ são obtidos por meio das equações não lineares $U_\alpha(\alpha, \theta|x) = 0$ e $U_\theta(\alpha, \theta|x) = 0$.

Dado o vetor de parâmetros desconhecidos $\Theta = (\alpha, \theta)$, para amostras grandes é conhecido que a distribuição assintótica do EMV Θ é

$$(\hat{\Theta} - \Theta) \rightarrow N_2\left(0, I^{-1}(\Theta)\right).$$

A inversa da matriz de informação de Fisher observada, $I^{-1}(\Theta)$, é dada por

$$I^{-1}(\Theta) = \begin{pmatrix} \mathrm{var}(\hat{\alpha}_{EMV}) & cov(\hat{\alpha}_{EMV}, \hat{\theta}_{EMV}) \\ cov(\hat{\theta}_{EMV}, \hat{\alpha}_{EMV}) & \mathrm{var}(\hat{\theta}_{EMV}) \end{pmatrix} \quad (4.37)$$

Sendo assim, têm-se:

- $$\frac{\partial^2 l(\alpha, \theta|x)}{\partial \alpha^2} = -\frac{n}{\alpha^2} + \frac{n\mathrm{e}^{-\alpha}}{1 - \mathrm{e}^{-\alpha}} + \frac{n(\mathrm{e}^{-\alpha})^2}{(1 - \mathrm{e}^{-\alpha})^2}; \quad (4.38)$$

- $$\frac{\partial^2 l(\alpha, \theta|x)}{\partial \theta^2} = \frac{-\sum_{i=1}^{n} \alpha(1 + (-x_i - 1)\mathrm{e}^{-x_i})^{\theta}(\ln(1 + (-x_i - 1)\mathrm{e}^{-x_i}))^2\theta^2 - n}{\theta^2}; \quad (4.39)$$

- $$\frac{\partial^2 l(\alpha, \theta|x)}{\partial \theta \partial \alpha} = -\sum_{i=1}^{n} \left(1 + (-x_i - 1)\mathrm{e}^{-x_i}\right)^{\theta} \ln\left(1 + (-x_i - 1)\mathrm{e}^{-x_i}\right). \quad (4.40)$$

Portanto, a abordagem acima é usada para obter os $100\%(1-\tau)$ intervalos de confiança aproximados dos parâmetros $\Theta = (\alpha, \theta)$, como nas seguintes formas

$$\hat{\alpha}_{EMV} \pm z_{\frac{\tau}{2}} \sqrt{\mathrm{var}(\hat{\alpha}_{MLE})} \quad ; \quad \hat{\theta}_{EMV} \pm z_{\frac{\tau}{2}} \sqrt{\mathrm{var}(\hat{\theta}_{MLE})}$$

Onde, têm-se o $z_{\frac{\tau}{2}}$-ésimo percentil da distribuição normal padrão.

Método dos Mínimos Quadrados

Sejam $x_{1:n} < x_{2:n} < \cdots < x_{n:n}$ uma amostra aleatória de tamanho n com função de distribuição $F(x)$, então

$$E[F(x_{i:n}|\alpha,\theta)] = \tfrac{i}{n+1} \qquad \text{e} \qquad V[F(x_{i:n}|\alpha,\theta)] = \tfrac{i(n-i+1)}{(n+1)^2(n+2)}.$$

As estimativas de Mínimos Quadrados (EMQ) para a GEEMO, $\hat{\alpha}_{EMQ}$ e $\hat{\theta}_{EMQ}$ para os parâmetros α e θ são obtidos através da minimização da função

$$M(\alpha,\theta) = \sum_{i=1}^{n}\left[F(x_{i:n}|\alpha,\theta) - \frac{i}{n+1}\right]^2. \tag{4.41}$$

Os parâmetros são obtidos por meio das equações $\frac{\partial M}{\partial \alpha} = 0$ e $\frac{\partial M}{\partial \theta} = 0$, que são calculadas através de soluções numéricas, por não possuir forma analítica. Os resultados das derivadas parciais são dados por

$$\frac{\partial M}{\partial \alpha} = \sum_{i=1}^{n}\left(\frac{\alpha\theta x_i e^{-\alpha\left(1-\mathrm{e}^{-x_i}(x_i+1)\right)^{\theta}} e^{-x_i}(1-\mathrm{e}^{-x_i}(x_i+1))^{\theta-1}}{1-e^{-\alpha}} - \frac{i}{n+1}\right)\eta_1(\alpha,\theta|x_{i:n}) = 0, \tag{4.42}$$

$$\frac{\partial M}{\partial \theta} = \sum_{i=1}^{n}\left(\frac{\alpha\theta x_i e^{-\alpha\left(1-\mathrm{e}^{-x_i}(x_i+1)\right)^{\theta}} e^{-x_i}(1-\mathrm{e}^{-x_i}(x_i+1))^{\theta-1}}{1-e^{-\alpha}} - \frac{i}{n+1}\right)\eta_2(\alpha,\theta|x_{i:n}) = 0. \tag{4.43}$$

onde

$$\eta_1(\alpha,\theta|x_{i:n}) = -\frac{S(x_{i:n}|\theta)F(x_{i:n}|\theta)}{[1-\bar{\alpha}S(x_{i:n}|\theta)]^2},$$

$$\eta_2(\alpha,\theta|x_{i:n}) = \frac{(1-\bar{\alpha}S(x_{i:n}|\theta))\left[\frac{\partial}{\partial\theta}F(x_{i:n}|\theta)\right] + \bar{\alpha}F(x_{i:n}|\theta)\left[\frac{\partial}{\partial\theta}S(x_{i:n}|\theta)\right]}{[1-\bar{\alpha}S(x_{i:n}|\theta)]^2}.$$

Método dos Mínimos Quadrados Ponderados

A linearidade, homocedasticidade, não tendenciosidade, variância mínima e consistência são algumas das características apresentadas pelo estimador de mínimos quadrados, porém, na prática, nem sempre essas suposições são válidas. Por isso, foram desenvolvidos métodos alternativos como o estimador de mínimos quadrados ponderados que corrige a falta de homocedasticidade, que são as variâncias iguais.

Os Estimadores de Mínimos Quadrados Ponderados (EMQP), $\hat{\alpha}_{EMQP}$ e $\hat{\theta}_{EMQP}$ para os parâmetros α e θ são obtidos através da minimização da seguinte função

$$W(\alpha,\theta) = \sum_{i=1}^{n}\left[\frac{(n+1)^2(n+2)}{i(n-i+1)}F(x_{i:n}|\alpha,\theta) - \frac{i}{n+1}\right]^2. \tag{4.44}$$

O que diferencia da equação 3.49 é que neste método os estimadores são ponderados pela precisão da distribuição teórica. As estimativas para esses parâmetros são obtidas através de soluções numéricas, por não possuírem forma analítica. As derivadas parciais são dadas por

$$\frac{\partial W(\alpha,\theta|x)}{\partial \alpha} = \sum_{i=1}^{n}\left[\frac{(n+1)^2(n+2)}{i(n-i+1)} \times \frac{F(x_{i:n}|\theta)}{1-\bar{\alpha}S(x_{i:n}|\theta)} - \frac{i}{n+1}\right]\eta_1(\alpha,\theta|x_{i:n}), \tag{4.45}$$

$$\frac{\partial W(\alpha,\theta|x)}{\partial \theta} = \sum_{i=1}^{n}\left[\frac{(n+1)^2(n+2)}{i(n-i+1)} \times \frac{F(x_{i:n}|\theta)}{1-\bar{\alpha}S(x_{i:n}|\theta)} - \frac{i}{n+1}\right]\eta_2(\alpha,\theta|x_{i:n}). \tag{4.46}$$

com

$$\eta_1(\alpha,\theta|x_{i:n}) = -\frac{S(x_{i:n}|\theta)F(x_{i:n}|\theta)}{\left[1-\bar{\alpha}S(x_{i:n}|\theta)\right]^2},$$
$$\eta_2(\alpha,\theta|x_{i:n}) = \frac{\left(1-\bar{\alpha}S(x_{i:n}|\theta)\right)\left[\frac{\partial}{\partial\theta}F(x_{i:n}|\theta)\right] + \bar{\alpha}F(x_{i:n}|\theta)\left[\frac{\partial}{\partial\theta}S(x_{i:n}|\theta)\right]}{\left[1-\bar{\alpha}S(x_{i:n}|\theta)\right]^2}.$$

Método de Cramér-von-Mises

Os Estimadores de Cramér-von-Mises (ECM) $\hat{\alpha}_{ECM}$ e $\hat{\theta}_{ECM}$ para os parâmetros α e θ são obtidos através da minimização da distância mínima por meio da estatística *goodness-of-fit*, essa estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$C(\alpha,\theta) = \frac{1}{12n} + \sum_{i=1}^{n}\left[F(x_{i:n}|\alpha,\theta) - \frac{2i-1}{2n}\right]^2. \tag{4.47}$$

Os parâmetros são obtidos através das equações $\frac{\partial C}{\partial \alpha} = 0$ e $\frac{\partial C}{\partial \theta} = 0$, que são obtidas por meio de soluções numéricas, por não possuírem forma analítica. As derivadas parciais são dadas por

$$\frac{\partial C(\alpha,\theta|x)}{\partial\alpha} = \sum_{i=1}^{n}\left[\frac{F(x_{i:n}|\theta)}{1-\bar{\alpha}S(x_{i:n}|\theta)} - \frac{2i-1}{2n}\right]\eta_1(\alpha,\theta|x_{i:n}), \tag{4.48}$$
$$\frac{\partial C(\alpha,\theta|x)}{\partial\theta} = \sum_{i=1}^{n}\left[\frac{F(x_{i:n}|\theta)}{1-\bar{\alpha}S(x_{i:n}|\theta)} - \frac{2i-1}{2n}\right]\eta_2(\alpha,\theta|x_{i:n}). \tag{4.49}$$

Método de Anderson Darling

As estimativas de Anderson Darling (EAD) $\hat{\alpha}_{EAD}$ e $\hat{\theta}_{EAD}$ para os parâmetros α e θ são obtidos através da minimização da distância mínima, por meio da estatística *goodness-of-fit*, que se baseia na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$A(\alpha,\theta) = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\log\left(F(x_{i:n}|\alpha,\theta)\right) + \log\left(\bar{F}(x_{n+1-i:n}|\alpha,\theta)\right)\right\}. \tag{4.50}$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial A}{\partial\alpha} = 0$ e $\frac{\partial A}{\partial\theta} = 0$, essas equações são dadas por

$$\frac{\partial A(\alpha,\theta|x)}{\partial\alpha} = \sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)} - \frac{\eta_1(\alpha,\theta|x_{n+1-i:n})}{S(x_{n+1-i:n}|\alpha,\theta)}\right], \tag{4.51}$$
$$\frac{\partial A(\alpha,\theta|x)}{\partial\theta} = \sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)} - \frac{\eta_2(\alpha,\theta|x_{n+1-i:n})}{S(x_{n+1-i:n}|\alpha,\theta)}\right]. \tag{4.52}$$

Método de Anderson Darling - RT (right-tail)

As estimativas de Anderson Darling - RT (EADD) $\hat{\alpha}_{EADD}$ e $\hat{\theta}_{EADD}$ para os parâmetros α e θ são obtidos, através da minimização da distância mínima por meio da estatística *goodness-of-fit*, essa estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$R(\alpha,\theta)=\frac{n}{2}-2\sum_{i=1}^{n}F(x_{i:n}|\alpha,\theta)-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\log\left(\bar{F}(x_{n+1-i:n}|\alpha,\theta)\right). \tag{4.53}$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial R}{\partial\alpha}=0$ e $\frac{\partial R}{\partial\theta}=0$, essas equações são dadas por

$$\frac{\partial R(\alpha,\theta|x)}{\partial\alpha}=-2\sum_{i=1}^{n}\eta_1(\alpha,\theta|x_{i:n})+\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1(\alpha,\theta|x_{n+1-i:n})}{S(x_{n+1-i:n}|\alpha,\theta)}\right], \tag{4.54}$$

$$\frac{\partial R(\alpha,\theta|x)}{\partial\theta}=-2\sum_{i=1}^{n}\eta_2(\alpha,\theta|x_{i:n})+\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2(\alpha,\theta|x_{n+1-i:n})}{S(x_{n+1-i:n}|\alpha,\theta)}\right]. \tag{4.55}$$

Método de Anderson Darling - LT (left-tail)

As estimativas de Anderson Darling - LT (EADL) $\hat{\alpha}_{EADL}$ e $\hat{\theta}_{EADL}$ para os parâmetros α e θ são obtidos através da minimização da distância mínima por meio da estatística *goodness-of-fit*, essa estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$L=-\frac{3n}{2}+2\sum_{i=1}^{n}F(x_{i:n}|\alpha,\theta)-\frac{1}{n}\sum_{i=1}^{n}(2i-1)ln(F(x_{i:n}|\alpha,\theta)). \tag{4.56}$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial L}{\partial\alpha}=0$ e $\frac{\partial L}{\partial\theta}=0$, essas equações são dadas por

$$\frac{\partial L(\alpha,\theta|x)}{\partial\alpha}=2\sum_{i=1}^{n}\eta_1(\alpha,\theta|x_{i:n})-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)}\right], \tag{4.57}$$

$$\frac{\partial L(\alpha,\theta|x)}{\partial\theta}=2\sum_{i=1}^{n}\eta_2(\alpha,\theta|x_{i:n})-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)}\right]. \tag{4.58}$$

Método de Anderson Darling - 2LT (left-tail de segunda ordem)

As estimativas de Anderson Darling - 2LT (EAD2L) $\hat{\alpha}_{EAD2L}$ e $\hat{\theta}_{EAD2L}$ para os parâmetros α e θ são obtidos através da minimização da distância mínima por meio da estatística *goodness-of-fit*, essa estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$LT=2\sum_{i=1}^{n}ln(F(x_{i:n}|\alpha,\theta))+\frac{1}{n}\sum_{i=1}^{n}\frac{2i-1}{F(x_{i:n}|\alpha,\theta)}. \tag{4.59}$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial LT}{\partial\alpha}=0$ e $\frac{\partial LT}{\partial\theta}=0$, essas equações são dadas por

$$\frac{\partial LT(\alpha,\theta|x)}{\partial\alpha}=2\sum_{i=1}^{n}\frac{\eta_1(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)}-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1(\alpha,\theta|x_{i:n})}{(F(x_{i:n}|\alpha,\theta))^2}\right], \tag{4.60}$$

$$\frac{\partial LT(\alpha,\theta|x)}{\partial\theta}=2\sum_{i=1}^{n}\frac{\eta_2(\alpha,\theta|x_{i:n})}{F(x_{i:n}|\alpha,\theta)}-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2(\alpha,\theta|x_{i:n})}{(F(x_{i:n}|\alpha,\theta))^2}\right]. \tag{4.61}$$

Método de Kolmogorov

As estimativas de Kolmogorov (EK) $\hat{\alpha}_{EK}$ e $\hat{\theta}_{EK}$ para os parâmetros α e θ são obtidos por meio da minimização da distância mínima através da estatística *goodness-of-fit*, essa estatística baseia-se na diferença entre a estimativa da função distribuição acumulada e a distribuição empírica

$$K(\alpha,\theta) = \max_{1\leq i\leq n}\left\{\frac{i}{n} - F(x_{i:n}|\alpha,\theta), F(x_{i:n}|\alpha,\theta) - \frac{i-1}{n}\right\}. \quad (4.62)$$

As estimativas dos parâmetros são obtidas por meio da solução das equações não lineares oriundas de $\frac{\partial K}{\partial \alpha} = 0$ e $\frac{\partial K}{\partial \theta} = 0$.

Método Bayesiano

A inferência Bayesiana se diferencia da clássica, pois permite que a opinião do especialista seja levada em consideração, ou seja, a estimativa dos parâmetros não dependem apenas dos dados, mas também da informação a priori que é obtida pelo especialista do assunto tratado. Os parâmetros da distribuição em estudo são tidos como variáveis aleatórias na inferência Bayesiana, diferenciando-se da clássica, onde o mesmo é tido como um valor fixo desconhecido. No método Bayesiano os parâmetros assumem uma distribuição (distribuição à priori) e o mesmo é atualizado (distribuição a posteriori) pelo Teorema de Bayes, combinando com a informação trazida dos dados pela função de Verossimilhança.

Para os parâmetros desconhecidos da $GEPTZ \sim (\alpha,\theta)$ será assumido que α e θ tenham como distribuições a priori independentes a Gama. Com função densidade probabilidade dadas por

$$g(\alpha) \propto \alpha^{a-1}e^{\alpha b}, \quad \alpha > 0$$
$$g(\theta) \propto \theta^{c-1}e^{\theta d}, \quad \theta > 0.$$

Os hiperparâmetros a, b, c e d são conhecidos e não negativos.

Para obtenção dos estimadores foi utilizado o algoritmo de Metropolis-Hasting, que é um processo iterativo utilizado para contornar o problema da obtenção da posteriori de forma analítica, gerando possíveis valores da mesma, proposto inicialmente por Metropolis Nicholas e Rosenbluth (1953) e generalizado por Hastings (1970), onde ambos buscavam a obtenção de coeficientes de determinadas integrais inatingíveis analiticamente.

Utilizando uma distribuição de probabilidade auxiliar, $q(\theta,\cdot)$, são gerados possíveis valores da distribuição a posteriori. Onde o valor gerado é aceito mediante a satisfação da condição abaixo

$$\alpha(\theta^{(t)},\theta^{(*)}) = \min\left[1; \frac{p(\theta^{(*)})q(\theta^{(t)}|\theta^{(*)})}{p(\theta^{(t)})q(\theta^{(*)}|\theta^{(t)})}\right] \geq u(0,1).$$

Então, são realizados os seguintes passos:

1 Proponha a distribuição auxiliar que gere possíveis valores de θ a posteriori;

2 Inicie o contador $t = 1$;

3 Escolha um valor inicial para $\theta^{(t)}$;

4 Gere um valor para $\theta^{(t)}$ utilizando $q(\theta,\cdot)$;

5 Teste a condição $\alpha(\theta^{(t)},\theta^{(*)}) \geq u(0,1)$;

6 Se a condição for satisfeita: $\theta^{(t+1)} = \theta^{(*)}, t = t+1$ e volte ao passo 4;

7 Se a condição não for satisfeita: $\theta^{(t+1)} = \theta^{(t)}, t = t + 1$ e volte ao passo 4;

8 Repita o procedimento até alcançar convergência.

O algoritmo foi escrito em linguagem de programação R, para cada uma das 1000 amostras foram realizadas 20000 iterações, porém as 5000 primeiras foram descartadas por causa do aquecimento inicial (*burn-in* inicial), e o resultado utilizado foi a moda e a mediana destes 15000 valores gerados, para α e θ respectivamente.

4.7 Simulação

Analogamente ao que foi feito para o modelo anterior, foram geradas $B = 1000$ amostras da distribuição GEPTZ para diferentes valores dos parâmetros, para quatro diferentes tamanhos amostrais, sendo eles $n = 20$, 60, 100 e 200.

As amostras foram obtidas por meio da transformação inversa gerando-se um vetor u de tamanho 200000 tal que $u \sim U(0,1)$, então são encontradas as raízes para $F(x) = u \rightarrow x = F^{-1}(u)$, estas são obtidas pelo método numérico de Newton-Raphson. Tais raízes são armazenadas numa matriz X de 200 linhas (número máximo de tamanho da amostra) e 1000 colunas (número de amostras). Sendo assim, na primeira coluna da matriz X, as 20 primeiras linhas representam a primeira amostra de tamanho $n = 20$, as 60 primeiras linhas, ou seja, a primeira amostra adicionando-se as 40 próximas linhas da primeira coluna representa a primeira amostra de tamanho $n = 60$, considerando as 100 primeiras linhas da mesma coluna obtêm-se a primeira amostra de tamanho $n = 100$ e, finalmente, considerando-se todas as linhas da primeira coluna, tem-se a primeira amostra de tamanho $n = 200$. Repetindo esse processo nas 1000 colunas, obtemos as 1000 amostras de cada um dos diferentes tamanhos amostrais considerados no estudo.

4.8 Comparador de Estimadores

Para comparar as diferentes estimativas para cada um dos métodos apresentados anteriormente, foram calculados o vício ($BIAS$), o erro padrão médio ($RMSE$), a média da diferença absoluta entre a distribuição teórica e a estimada (D_{abs}) e a diferença absoluta máxima entre as funções de distribuições verdadeira e estimada, sendo que suas expressões são dadas por:

$$\begin{aligned} BIAS(\hat{\alpha}) &= \tfrac{1}{B}\sum_{i=1}^{B}(\hat{\alpha}_i - \alpha) \qquad BIAS(\hat{\theta}) = \tfrac{1}{B}\sum_{i=1}^{B}(\hat{\theta}_i - \theta) \\ RMSE(\hat{\alpha}) &= \sqrt{\tfrac{1}{B}\sum_{i=1}^{B}(\hat{\alpha}_i - \alpha)^2} \quad RMSE(\hat{\theta}) = \sqrt{\tfrac{1}{B}\sum_{i=1}^{B}(\hat{\theta}_i - \theta)^2} \\ D_{abs} &= \tfrac{1}{Bn}\sum_{i=1}^{B}\sum_{j=1}^{n}\left|F(x_{ij}|\alpha,\theta) - F(t_{ij}|\hat{\alpha},\hat{\theta})\right| \\ D_{\max} &= \tfrac{1}{B}\sum_{i=1}^{B}\max_{j}\left|F(x_{ij}|\alpha,\theta) - F(x_{ij}|\hat{\alpha},\hat{\theta})\right| \end{aligned} \tag{4.63}$$

4.9 Resultados das Simulações

As tabelas 4.2, 4.3 e 4.4 possuem os resultados dos comparadores de estimadores, onde são comparados os dez diferentes estimadores para alguns valores de parâmetros utilizados na simulação.

Tabela 4.2: Resultado da simulação para os parâmetros α=0.1 e θ=0.9

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.2493(7)	0.1522(10)	1.6406(2)	0.4479(3)	0.0555(3)	0.0929(4)
MQ		-0.5386(10)	-0.0414(3)	1.9608(7)	0.4673(6)	0.0563(6)	0.0939(6)
MQP		-0.4902(9)	-0.0263(1)	1.9694(8)	0.4563(5)	0.0553(2)	0.0919(2)
CM		0.0879(2)	0.1305(7)	1.9473(5)	0.5425(8)	0.058(8)	0.0999(8)
KS		0.0879(3)	0.1305(8)	1.9473(6)	0.5425(9)	0.058(9)	0.0999(9)
AD		-0.2545(8)	0.0297(2)	1.8808(3)	0.4382(2)	0.0556(5)	0.0922(3)
ADR		-0.0276(1)	0.1306(9)	1.9282(4)	0.5813(10)	0.0608(10)	0.1047(10)
ADL		-0.1301(4)	0.0514(4)	1.9834(9)	0.4546(4)	0.0555(4)	0.0931(5)
AD2L		0.2417(5)	0.0835(5)	4.8774(10)	0.5204(7)	0.0578(7)	0.0996(7)
BAYES		0.2424(6)	0.0974(6)	0.4551(1)	0.2646(1)	0.0459(1)	0.0701(1)
MV	60	-0.0222(1)	0.0397(10)	1.2075(6)	0.2563(2)	0.0332(2)	0.0552(2)
MQ		-0.3083(10)	-0.0362(9)	1.2745(8)	0.309(7)	0.0343(6)	0.058(6)
MQP		-0.2006(9)	-0.0103(4)	1.2078(7)	0.2845(5)	0.0336(4)	0.0565(5)
CM		-0.0645(2)	0.0286(6)	1.1679(4)	0.3109(8)	0.0346(7)	0.0593(7)
KS		-0.0645(3)	0.0286(7)	1.1679(5)	0.3109(9)	0.0346(8)	0.0593(8)
AD		-0.1521(6)	-0.0005(1)	1.133(2)	0.2685(4)	0.0335(3)	0.0561(3)
ADR		-0.1634(7)	0.0191(5)	1.3006(9)	0.3477(10)	0.0357(9)	0.062(10)
ADL		-0.0919(4)	0.0101(3)	1.1364(3)	0.2638(3)	0.0338(5)	0.0564(4)
AD2L		-0.1918(8)	-0.0091(2)	1.524(10)	0.2975(6)	0.0367(10)	0.0618(9)
BAYES		0.1064(5)	0.0345(8)	0.2645(1)	0.138(1)	0.0263(1)	0.0399(1)
MV	100	0.0248(1)	0.0226(8)	0.7324(2)	0.1774(2)	0.0253(1)	0.042(1)
MQ		-0.1983(9)	-0.0317(10)	0.9553(8)	0.2311(8)	0.0263(5)	0.0445(5)
MQP		-0.0961(7)	-0.009(7)	0.8239(5)	0.2016(5)	0.0256(2)	0.0429(3)
CM		-0.0619(3)	0.0064(4)	0.9109(6)	0.2297(6)	0.0264(6)	0.045(6)
KS		-0.0619(4)	0.0064(5)	0.9109(7)	0.2297(7)	0.0264(7)	0.045(7)
AD		-0.0884(6)	-0.0063(3)	0.8169(4)	0.1966(4)	0.0256(3)	0.0428(2)
ADR		-0.1339(8)	-0.0008(1)	1.0378(9)	0.27(10)	0.0275(8)	0.0477(8)
ADL		-0.0393(2)	0.0021(2)	0.7837(3)	0.1879(3)	0.0258(4)	0.043(4)
AD2L		-0.2154(10)	-0.0246(9)	1.3507(10)	0.2332(9)	0.0289(9)	0.0484(9)
BAYES		-0.0681(5)	-0.0085(6)	0.1256(1)	0.0896(1)	0.1723(10)	0.3986(10)
MV	200	0.0102(1)	0.0098(7)	0.5272(3)	0.1239(2)	0.0174(1)	0.0287(1)
MQ		-0.1093(9)	-0.0215(9)	0.6292(8)	0.1642(9)	0.0182(5)	0.031(5)
MQP		-0.042(3)	-0.0052(4)	0.5315(5)	0.1391(5)	0.0177(2)	0.0295(2)
CM		-0.0428(4)	-0.0023(2)	0.607(6)	0.1623(7)	0.0183(6)	0.0311(6)
KS		-0.0428(5)	-0.0023(3)	0.607(7)	0.1623(8)	0.0183(7)	0.0311(7)
AD		-0.0438(6)	-0.0054(5)	0.5299(4)	0.1373(4)	0.0177(3)	0.0295(3)
ADR		-0.0899(7)	-0.0083(6)	0.7276(9)	0.1953(10)	0.019(8)	0.033(8)
ADL		-0.0156(2)	-0.0002(1)	0.5072(2)	0.1299(3)	0.0179(4)	0.0297(4)
AD2L		-0.1512(10)	-0.0263(10)	0.8802(10)	0.1613(6)	0.0212(9)	0.0352(9)
BAYES		-0.0934(8)	-0.016(8)	0.116(1)	0.0632(1)	0.2146(10)	0.5028(10)

Tabela 4.3: Resultado da simulação para os parâmetros α=3 e θ=3

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.4652(5)	0.2149(9)	2.2084(5)	0.7957(2)	0.0581(2)	0.0956(2)
MQ		-0.0809(2)	-0.0918(3)	2.3469(6)	0.867(5)	0.0604(5)	0.0987(4)
MQP		-0.008(1)	-0.0354(1)	2.069(3)	0.8088(3)	0.0587(3)	0.0961(3)
CM		0.6421(7)	0.2136(7)	3.2218(8)	0.9424(8)	0.0604(6)	0.1015(7)
KS		0.6421(8)	0.2136(8)	3.2218(9)	0.9424(9)	0.0604(7)	0.1015(8)
AD		0.1483(3)	0.0512(2)	2.0094(2)	0.7661(1)	0.058(1)	0.0946(1)
ADR		0.2383(4)	0.1127(4)	2.1485(4)	0.9112(7)	0.0603(4)	0.1003(5)
ADL		0.5756(6)	0.1419(5)	2.962(7)	0.8806(6)	0.0604(8)	0.101(6)
AD2L		7.4094(10)	0.1983(6)	168.8932(10)	1.223(10)	0.0681(10)	0.1189(10)
BAYES		-0.9889(9)	-0.4529(10)	1.9995(1)	0.8647(4)	0.0668(9)	0.1086(9)
MV	60	0.1142(5)	0.0693(9)	0.8908(2)	0.4365(1)	0.0339(1)	0.0562(1)
MQ		-0.0541(4)	-0.0284(4)	0.9835(5)	0.4857(5)	0.0356(6)	0.059(4)
MQP		0.0025(1)	0.0065(1)	0.9131(3)	0.4516(3)	0.0345(3)	0.0572(3)
CM		0.139(7)	0.0672(7)	1.0322(6)	0.4983(6)	0.0355(4)	0.0595(6)
KS		0.139(8)	0.0672(8)	1.0322(7)	0.4983(7)	0.0355(5)	0.0595(7)
AD		0.0219(2)	0.0181(3)	0.8868(1)	0.4406(2)	0.0343(2)	0.0568(2)
ADR		0.0427(3)	0.0337(5)	0.9637(4)	0.5065(9)	0.0356(7)	0.0594(5)
ADL		0.1302(6)	0.0493(6)	1.1029(8)	0.4821(4)	0.0357(8)	0.0598(8)
AD2L		0.1502(9)	-0.0126(2)	1.9795(10)	0.6532(10)	0.0431(10)	0.0747(10)
BAYES		-0.3512(10)	-0.1624(10)	1.1076(9)	0.5062(8)	0.037(9)	0.0612(9)
MV	100	0.0618(6)	0.0314(7)	0.658(1)	0.3172(1)	0.0257(1)	0.0426(1)
MQ		-0.0485(5)	-0.0318(8)	0.7155(4)	0.3563(5)	0.0271(4)	0.0449(4)
MQP		-0.0038(2)	-0.0054(2)	0.668(3)	0.3279(3)	0.0261(2)	0.0432(3)
CM		0.0638(8)	0.0246(5)	0.7313(6)	0.3593(6)	0.0271(5)	0.0451(5)
KS		0.0638(9)	0.0246(6)	0.7313(7)	0.3593(7)	0.0271(6)	0.0451(6)
AD		0.0035(1)	-0.0015(1)	0.6617(2)	0.3245(2)	0.0261(3)	0.0431(2)
ADR		0.015(4)	0.0067(3)	0.7259(5)	0.3798(8)	0.0272(7)	0.0454(8)
ADL		0.0635(7)	0.0163(4)	0.777(8)	0.3455(4)	0.0272(8)	0.0453(7)
AD2L		-0.008(3)	-0.05(9)	1.3829(10)	0.5002(10)	0.0341(10)	0.0595(10)
BAYES		-0.2052(10)	-0.0942(10)	0.8797(9)	0.403(9)	0.0297(9)	0.0518(9)
MV	200	0.0286(7)	0.0125(7)	0.448(1)	0.2196(1)	0.0176(1)	0.029(1)
MQ		-0.0356(8)	-0.0236(9)	0.4968(6)	0.2545(6)	0.0187(7)	0.0311(5)
MQP		-0.0049(3)	-0.0061(6)	0.4577(2)	0.2314(4)	0.0179(2)	0.0297(2)
CM		0.0195(4)	0.0043(3)	0.5009(7)	0.2548(7)	0.0186(5)	0.0312(6)
KS		0.0195(5)	0.0043(4)	0.5009(8)	0.2548(8)	0.0186(6)	0.0312(7)
AD		-0.0028(1)	-0.0053(5)	0.4581(3)	0.2302(3)	0.0179(3)	0.0297(3)
ADR		0.0028(2)	-0.0009(1)	0.496(5)	0.2662(9)	0.0188(9)	0.0314(9)
ADL		0.0244(6)	0.003(2)	0.5178(9)	0.2427(5)	0.0187(8)	0.0313(8)
AD2L		-0.0993(10)	-0.0614(10)	0.8875(10)	0.35(10)	0.0246(10)	0.0426(10)
BAYES		-0.0474(9)	-0.0204(8)	0.4851(4)	0.2243(2)	0.0181(4)	0.03(4)

Tabela 4.4: Resultado da simulação para os parâmetros α=1.1 e θ=3

	n	$BIAS\ \alpha$	$BIAS\ \theta$	$RMSE\ \alpha$	$RMSE\ \theta$	$Dabs$	$Dmax$
MV	20	0.2324(7)	0.3171(9)	1.6544(2)	1.1296(2)	0.0571(1)	0.0955(3)
MQ		-0.4654(9)	-0.2048(5)	1.9872(7)	1.2642(5)	0.0584(5)	0.0972(4)
MQP		-0.3463(8)	-0.1177(2)	1.8506(4)	1.1931(4)	0.0573(3)	0.0953(2)
CM		0.1547(5)	0.2464(7)	1.9918(8)	1.3461(6)	0.0592(6)	0.1013(6)
KS		0.1547(6)	0.2464(8)	1.9918(9)	1.3461(7)	0.0592(7)	0.1013(7)
AD		-0.1414(3)	0.0306(1)	1.7638(3)	1.129(1)	0.0572(2)	0.0948(1)
ADR		-0.0713(2)	0.1592(4)	1.9438(5)	1.4135(8)	0.0607(8)	0.1037(8)
ADL		0.0465(1)	0.1215(3)	1.9677(6)	1.188(3)	0.0581(4)	0.0975(5)
AD2L		0.9308(10)	0.2109(6)	15.9401(10)	1.4377(9)	0.0625(9)	0.1084(9)
BAYES		-0.1429(4)	0.6457(10)	0.3619(1)	2.722(10)	0.0774(10)	0.1261(10)
MV	60	0.0538(6)	0.0993(9)	0.8687(3)	0.6204(1)	0.0338(1)	0.0565(1)
MQ		-0.1632(10)	-0.0766(8)	1.008(8)	0.7382(5)	0.0352(5)	0.0593(5)
MQP		-0.0712(8)	-0.0087(1)	0.8928(4)	0.6668(4)	0.0343(3)	0.0577(3)
CM		0.0222(2)	0.0716(6)	0.9619(6)	0.7391(6)	0.0352(6)	0.06(6)
KS		0.0222(3)	0.0716(7)	0.9619(7)	0.7391(7)	0.0352(7)	0.06(7)
AD		-0.0464(5)	0.0125(2)	0.8625(2)	0.6426(2)	0.0341(2)	0.0573(2)
ADR		-0.0631(7)	0.029(4)	1.0382(9)	0.8134(9)	0.036(8)	0.0618(9)
ADL		0.017(1)	0.0466(5)	0.9039(5)	0.6484(3)	0.0348(4)	0.0585(4)
AD2L		-0.0854(9)	-0.0265(3)	1.4702(10)	0.8078(8)	0.0397(10)	0.0677(10)
BAYES		-0.0235(4)	0.4716(10)	0.1861(1)	1.8126(10)	0.0364(9)	0.0604(8)
MV	100	0.0373(6)	0.0488(7)	0.6066(2)	0.4433(1)	0.0257(1)	0.0428(1)
MQ		-0.0983(9)	-0.0626(9)	0.7198(8)	0.535(7)	0.0268(5)	0.0452(6)
MQP		-0.0381(7)	-0.0148(4)	0.6497(4)	0.4775(4)	0.026(2)	0.0436(3)
CM		0.0063(1)	0.0236(5)	0.7036(6)	0.5331(5)	0.0269(6)	0.0455(7)
KS		0.0063(2)	0.0236(6)	0.7036(7)	0.5331(6)	0.0269(7)	0.0455(8)
AD		-0.0325(5)	-0.0088(2)	0.6453(3)	0.4712(3)	0.026(3)	0.0435(2)
ADR		-0.0507(8)	-0.0053(1)	0.7993(9)	0.6207(9)	0.0276(9)	0.0474(9)
ADL		0.01(3)	0.0139(3)	0.6598(5)	0.4643(2)	0.0265(4)	0.0444(4)
AD2L		-0.0994(10)	-0.0615(8)	1.0653(10)	0.6111(8)	0.0313(10)	0.0535(10)
BAYES		-0.0127(4)	0.2568(10)	0.1374(1)	1.2511(10)	0.0269(8)	0.0446(5)
MV	200	0.0175(5)	0.0202(7)	0.4165(2)	0.3065(1)	0.0176(1)	0.0293(1)
MQ		-0.0558(9)	-0.041(8)	0.4774(8)	0.3733(7)	0.0185(6)	0.0313(6)
MQP		-0.019(7)	-0.0107(6)	0.4369(4)	0.3328(4)	0.0179(2)	0.03(3)
CM		-0.0053(2)	0.0013(1)	0.4722(6)	0.3718(5)	0.0185(7)	0.0313(7)
KS		-0.0053(3)	0.0013(2)	0.4722(7)	0.3718(6)	0.0185(8)	0.0313(8)
AD		-0.0186(6)	-0.0099(4)	0.4364(3)	0.3299(3)	0.0179(3)	0.0299(2)
ADR		-0.0283(8)	-0.01(5)	0.5295(9)	0.4295(9)	0.0191(9)	0.0327(9)
ADL		0.0033(1)	0.0021(3)	0.4438(5)	0.3239(2)	0.0183(5)	0.0305(5)
AD2L		-0.1066(10)	-0.0722(9)	0.701(10)	0.4259(8)	0.0228(10)	0.0386(10)
BAYES		-0.0064(4)	0.1251(10)	0.0981(1)	0.8426(10)	0.018(4)	0.0302(4)

Os estimadores possuem a propriedade de não tendenciosidade, pois o vício tende a zero conforme o n aumenta. Quanto ao erro padrão médio, é possível observar a propriedade de consistência, pois quando o valor de n aumenta a variância tende a zero.

Na primeira tabela, onde os valores dos parâmetro são menores, os menores vícios foram dos estimadores de Máxima Verossimilhança, Anderson Darling e Anderson Darling Calda a Direita, com valores próximos um do outro. Nesta mesma tabela, em relação ao erro padrão médio, o estimador Bayesiano teve os melhores resultados, enquanto que os piores foram o estimador de Anderson Darling Calda a Direita para o parâmetro θ e o Anderson Darling Calda a Esquerda de Segunda Ordem para o parâmetro α. Em relação a distância máxima e distância absoluta os resultados foram próximos.

Na tabela do meio, onde ambos os parâmetros tem valor igual a 3, o pior estimador foi o de Anderson Darling Calda a Esquerda de Segunda Ordem. O de menor vício foi o Anderson Darling. O estimador com o menor erro padrão médio também foi o estimador de Anderson Darling, no geral, mas quando n é grande o estimador de Máxima Verossimilhança obteve melhores resultados.

Por último, na terceira tabela, onde $\alpha = 1.1$ e $\theta = 3$, o estimador Bayesiano foi o mais eficiente para o parâmetro α enquanto que o estimador de Máxima Verossimilhança foi o mais eficiente para o parâmetro θ. Tanto para a distância máxima e distância absoluta os resultados foram próximos.

4.10 Aplicação

Para realizar a aplicação foram utilizados dados de Smith e Naylor (1987) que representam a resistência de fibras de vidro de $1,5cm$, medidos no *National Physical Laboratory*, Inglaterra.

É possível notar, observando o TTT Plot na Figura (4.5), que o risco é crescente.

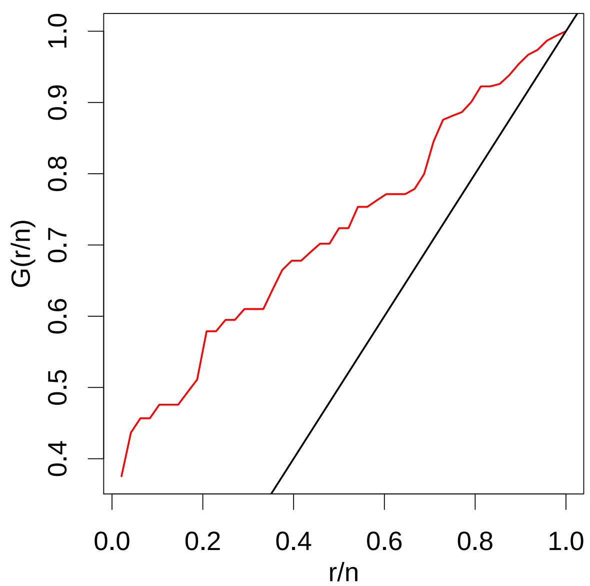


Figura 4.5: Gráfico TTT Plot.

Foram utilizadas 4 diferentes distribuições para comparar qual obteve o melhor ajuste perante os dados, entre elas, distribuições com 1, 2 e 3 parâmetros, distribuições muito conhecidas na literatura, outras de origens mais recentes e claro a nova distribuição em estudo, a GEPTZ.

As distribuições e seus respectivos parâmetros estimados pelo método da verossimilhança foram a GEPTZ com $\alpha = 43.7023$ e $\theta = 5.1264$, Gama Exponencial (GE) com $\theta = 1.1552$, Exponencial Generalizada (EG) com $\alpha = 31.3406$, $\lambda = 2.6116$ e por último a distribuição LogNormal (LN) com $\mu = 0.3810$ e $\sigma = 0.2578$.

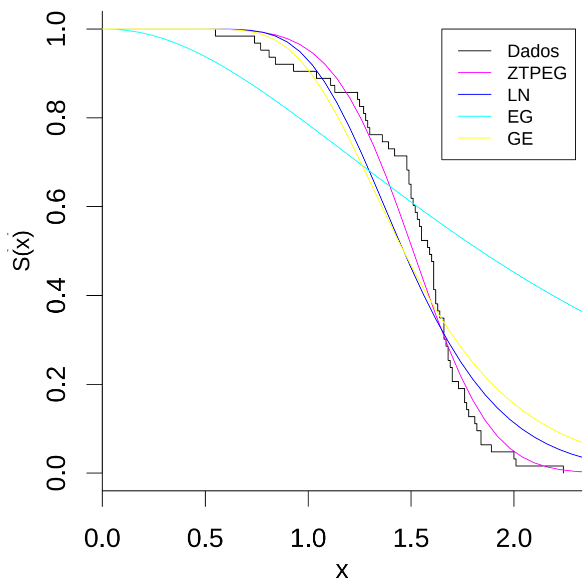


Figura 4.6: Gráfico dos dados de e as distribuições ajustadas.

Serão utilizados alguns critérios comuns na literatura para seleção de modelos. Esses critérios levam em consideração a complexidade do modelo. São critérios que, essencialmente, penalizam a função verossimilhança pelo número de parâmetros do modelo e, eventualmente, o tamanho da amostra. Essa penalização é feita subtraindo-se do valor da função verossimilhança uma determinada quantidade que expressa o quão complexo é o modelo (JUNIOR, 2013).

Akaike (1974) propôs utilizar a informação de Kullback-Leibler para a seleção de modelos. O autor estabeleceu uma relação entre a função de verossimilhança maximizada e a informação de Kullback-Leibler desenvolvendo, então, o chamado critério de informação de Akaike (AIC). Liang e Zou (2008) propuseram uma versão melhorada (CAIC), em que é considerada uma ponderação entre o número de parâmetros e o tamanho amostral.

O critério de informaçãao bayesiano (BIC), também chamado de critério de Schwarz et al. (1978), é um critério de avaliação de modelos definido em termos da probabilidade a posteriori, sendo assim chamado porque Schwarz forneceu um argumento bayesiano para prová-lo.

O melhor modelo será aquele que apresentar menor valor para AIC, BIC e CAIC

$$\begin{aligned} AIC &= -2l(\theta) + 2d \\ BIC &= -2l(\theta) + \log(n) \\ CAIC &= AIC + \frac{2(d+2)(d+3)}{n-d-3} \end{aligned} \tag{4.64}$$

Com base nos resultados apresentados na Tabela (4.5) é possível notar que a distribuição em estudo, a GEPTZ, obteve bom desempenho, pois obteve resultados muito melhores do que a GE, que é um caso particular da GEPTZ e tem apenas um parâmetro. Ela obteve resultados melhores também que outras distribuições de dois parâmetros, a

Tabela 4.5: Critérios de seleção do modelo

Modelo	AIC	BIC	CAIC
GEPTZ	39.9872	44.2735	40.6769
GE	142.5933	144.7365	143.0001
EG	66.7669	71.0532	67.4566
LN	60.0098	64.2961	61.0624

LN e a EG, ou seja, a nova distribuição GEPTZ se mostra uma boa alternativa para se trabalhar com ela.

Capítulo 5

Gama Exponenciada Bivariada

A distribuição Gama Exponenciada, introduzida por Gupta et al. (1998), foi utilizada como base para a criação de uma nova distribuição, a Gama Exponenciada Bivariada. A construção desta distribuição se torna útil pelo fato dela ser utilizada em casos onde os dados são bivariados, por exemplo, se é realizado um estudo onde é medido algum fator relacionado ao órgão ouvido, se torna necessário uma distribuição bivariada, pois uma fdp univariada não retrataria bem o caso em questão, pois o estudo seria apenas de um lado. Muitos outros casos também requerem o uso de distribuições bivariadas, onde é necessário o estudo utilizando duas variáveis aleatórias numa mesma distribuição. Esta nova distribuição bivariada de três parâmetros foi obtida utilizando o método proposto por Arnold (1967).

5.1 Construção do modelo

Para a criação da nova distribuição bivariada, deve-se supor que

$$U_1 \sim GE(\theta_1),$$

$$U_2 \sim GE(\theta_2),$$

$$U_3 \sim GE(\theta_3).$$

ou seja, U_1, U_2 e U_3 possuem distribuição Gama Exponenciada com seus respectivos parâmetros e são mutuamente independentes. Em seguida, foi definido

$$X = \max\{U_1, U_3\},$$

$$Y = \max\{U_2, U_3\}.$$

Então é dito que o vetor bivariado (X, Y) possue distribuição Gama Exponenciada Bivariada com os parâmetros θ_1, θ_2 e θ_3. Esta distribuição será denotada da seguinte maneira

$$(X,Y) \sim GEBV(\theta_1, \theta_2, \theta_3). \tag{5.1}$$

Para entender melhor como funciona na prática o uso de uma distribuição bivariada, supõe-se que em um sistema exista dois componentes. Cada componente está sujeito a tensão individual independente, digamos U_1 e U_2, respectivamente. O sistema tem uma tensão geral U_3 que foi transmitida a ambos os componentes igualmente, independentemente das suas tensões individuais. Portanto, a tensão observada nas duas componentes são $X = \max\{U_1, U_3\}$ e $Y = \max\{U_2, U_3\}$, respectivamente, (KUNDU; GUPTA, 2009).

Teorema 5 *Se $(X,Y) \sim GEBV(\theta_1, \theta_2, \theta_3)$, então a distribuição acumulada conjunta (X,Y) para $x > 0, y > 0$ é*

$$F_{X,Y}(x,y) = (1 - e^{-x}(x+1))^{\theta_1}(1 - e^{-y}(y+1))^{\theta_2}(1 - e^{-z}(z+1))^{\theta_3}, \tag{5.2}$$

onde $z = \min\{x, y\}$.

Demonstração. Seja $F_{GE}(x;\theta_i)$, $i = 1,2,3$, a distribuição acumulada da Gama Exponenciada, então

$$F_{X,Y}(x,y) = F_{GE}(x;\theta_1) \times F_{GE}(y;\theta_2) \times F_{GE}(z;\theta_3).$$

■

Corolário 1 *A distribuição GEBV também pode ser escrita da seguinte forma*

$$F_{X,Y}(x,y) = \begin{cases} F_{GE}(x;\theta_1+\theta_3) \times F_{GE}(y;\theta_2), & x < y \\ F_{GE}(x;\theta_1) \times F_{GE}(y;\theta_2+\theta_3), & y < x \\ F_{GE}(x;\theta_1+\theta_2+\theta_3), & x = y. \end{cases} \tag{5.3}$$

Teorema 6 *Se $(X,Y) \sim GEBV(\theta_1, \theta_2, \theta_3)$, então a função densidade de probabilidade da GEBV para $x > 0$, e $y > 0$, é*

$$\begin{cases} f_1(x,y), & 0 < x < y < \infty \\ f_2(x,y), & 0 < y < x < \infty \\ f_0(x), & 0 < x = y < \infty \end{cases} \tag{5.4}$$

onde

$$\begin{aligned} f_1(x,y) &= f_{GE}(x;\theta_1+\theta_3) f_{GE}(y;\theta_2) = \\ &= (\theta_1+\theta_3)\theta_2 e^{-y-x} xy (1 - e^{-y}(y+1))^{\theta_2-1}(1 - e^{-x}(x+1))^{\theta_1+\theta_3-1}, \end{aligned}$$

$$\begin{aligned} f_2(x,y) &= f_{GE}(x;\theta_1) f_{GE}(y;\theta_2+\theta_3) = \\ &= \theta_1(\theta_2+\theta_3) e^{-y-x} xy (1 - e^{-y}(y+1))^{\theta_2+\theta_3-1}(1 - e^{-x}(x+1))^{\theta_1-1}, \end{aligned}$$

e

$$\begin{aligned} f_0(x) &= \frac{\theta_3}{\theta_1+\theta_2+\theta_3} f_{GE}(x;\theta_1+\theta_2+\theta_3) = \\ &= \theta_3 e^{-x} x (1 - e^{-x}(x+1))^{\theta_1+\theta_2+\theta_3-1}. \end{aligned}$$

Demonstração. As funções f_1 e f_2 são obtidas facilmente fazendo $\frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$ para $x < y$ e $x > y$, respectivamente. No entanto, para calcular f_0 foi usados os seguintes fatos

$$\int_0^\infty \int_0^y f_1(x,y)dxdy + \int_0^\infty \int_0^x f_2(x,y)dydx + \int_0^\infty f_0(x)dx = 1,$$

$$\int_0^\infty \int_0^y f_1(x,y)dxdy = \int_0^\infty \theta_2 y e^{-y}(1-e^{-y}(y+1))^{\theta_1+\theta_2+\theta_3-1}dy = \frac{\theta_2}{\theta_1+\theta_2+\theta_3}$$

e

$$\int_0^\infty \int_0^x f_2(x,y)dydx = \int_0^\infty \theta_1 x e^{-x}(1-e^{-x}(x+1))^{\theta_1+\theta_2+\theta_3-1}dx = \frac{\theta_1}{\theta_1+\theta_2+\theta_3},$$

logo, note que

$$\int_0^\infty \theta_3 x e^{-x}(1-e^{-x}(x+1))^{\theta_1+\theta_2+\theta_3-1}dx = \frac{\theta_3}{\theta_1+\theta_2+\theta_3}.$$

■

Corolário 2 *A função densidade de probabilidade conjunta de X e Y dada em 5.4, pode ser expressa da seguinte forma, para $z = \min(x,y)$ e para $f_1(.,.)$ e $f_2(.,.)$*

$$f(x,y) = \frac{\theta_1+\theta_2}{\theta_1+\theta_2+\theta_3} f_a(x,y) + \frac{\theta_3}{\theta_1+\theta_2+\theta_3} f_s(x) \tag{5.5}$$

onde

$$f_a(x,y) = \begin{cases} \frac{\theta_1+\theta_2+\theta_3}{\theta_1+\theta_2} f_{GE}(x;\theta_1+\theta_3) f_{GE}(y;\theta_2), x < y \\ \frac{\theta_1+\theta_2+\theta_3}{\theta_1+\theta_2} f_{GE}(x;\theta_1) f_{GE}(y;\theta_2+\theta_3), x > y \end{cases}$$

e

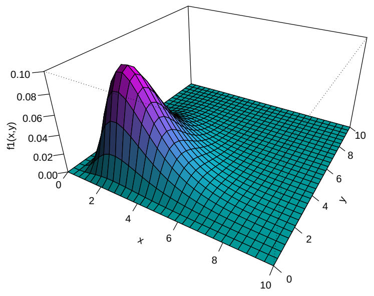
$$f_s(x) = f(x;\theta_1+\theta_2+\theta_3),$$

sendo que $f_a(x,y)$ é a parte absolutamente contínua e $f_s(x)$ é a parte singular.

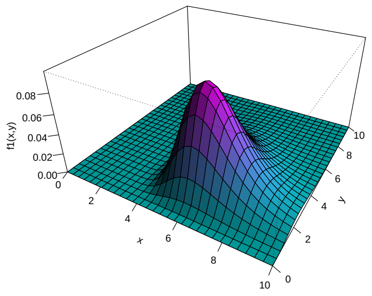
As Figuras 5.1 e 5.2 mostram o comportamento da função densidade, nos dois casos, quando $y > x$ e quando $y < x$. É possível observar, que o valor das densidades são baixos, ou seja, e distribuição deve ter mais facilidades para modelar dados com valores pequenos, que estejam próximos entre 0 e 1.

5.2 Função densidade de probabilidade marginal e condicional

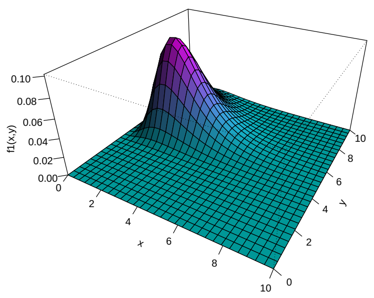
Foram encontradas as funções densidades marginais X_i e a função densidade condicional $X_i|X_j, (i \neq j)$.



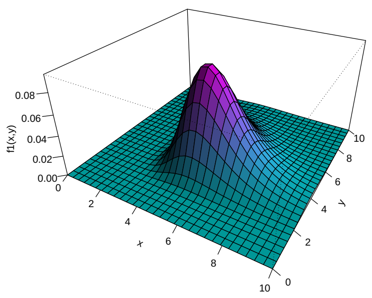
(a) $\theta_1 = 0.1$, $\theta_2 = 2.15$ e $\theta_3 = 2.5$



(b) $\theta_1 = 0.5$, $\theta_2 = 3$ e $\theta_3 = 50$



(c) $\theta_1 = 1.5$, $\theta_2 = 50$ e $\theta_3 = 0.5$



(d) $\theta_1 = 15$, $\theta_2 = 10$ e $\theta_3 = 10$

Figura 5.1: Gráfico da função densidade bivariada $(x < y)$.

5.2.1 Função densidade de probabilidade marginal

Teorema 7 *A função densidade de probabilidade marginal de* $X_i (i = 1, 2)$ *é dada por*

$$f_{X_i}(x_i) = (\theta_i + \theta_3) e^{-x_i} x_i (1 - e^{-x_i}(x_i + 1))^{\theta_i + \theta_3 - 1}. \tag{5.6}$$

Demonstração. Para encontrar a função densidade de probabilidade marginal, primeiro foi encontrado a conjunta da função acumulada de X_i, da seguinte forma

$$F_{X_i}(x_i) = P(\min(U_i, U_3) < x_i) = P(U_i < x_i, U_3 < x_i),$$

porém, como U_i e U_3 são independentes, então

$$P(U_i < x_i, U_3 < x_i) = P(U_i < x_i) P(U_3 < x_i),$$

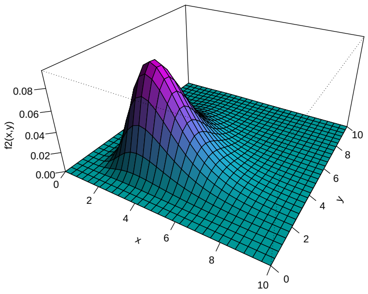
logo, tem-se

$$F_{X_i}(x_i) = (1 - e^{-x_i}(x_i + 1))^{\theta_i + \theta_3},$$

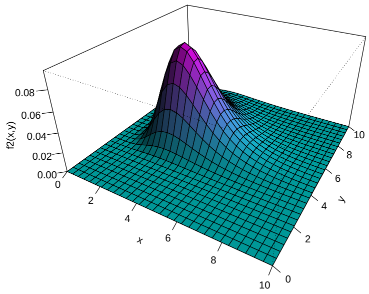
sendo assim, é possível chegar na função densidade de probabilidade marginal da seguinte forma

$$\frac{\partial}{\partial x_i} F_{X_i}(x_i) = f_{X_i}(x_i).$$

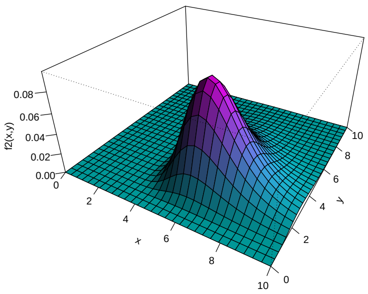
■



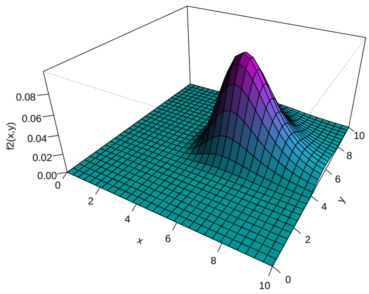
(a) $\theta_1 = 5$, $\theta_2 = 1.7$ e $\theta_3 = 2.5$



(b) $\theta_1 = 5$, $\theta_2 = 10$ e $\theta_3 = 15$



(c) $\theta_1 = 50$, $\theta_2 = 1.7$ e $\theta_3 = 2.5$



(d) $\theta_1 = 50$, $\theta_2 = 15$ e $\theta_3 = 25$

Figura 5.2: Gráfico da função densidade bivariada ($x > y$).

5.2.2 Função densidade de probabilidade condicional

Depois de obter a fdp marginal de $X_i, i = 1, 2$, pode-se obter a fdp condicional no presente Teorema (8)

Teorema 8 *A pdf condicional de X_i dado $X_j = x_j$ é denotada por*

$$f_{X_i|X_j}(x_i, x_j) = \begin{cases} f^{(1)}_{X_i|X_j}(x_i|x_j), x_i < x_j, \\ f^{(2)}_{X_i|X_j}(x_i|x_j), x_i > x_j, \\ f^{(3)}_{X_i|X_j}(x_i|x_j), x_i = x_j, \end{cases}$$

onde tem-se

$$f^{(1)}_{X_i|X_j}(x_i|x_j) = \frac{(\theta_i + \theta_3)\theta_j e^{-x_i} x_i (1 - e^{-x_j}(x_j + 1))^{\theta_3} (1 - e^{-x_i}(x_i + 1))^{\theta_i - 1}}{(\theta_j + \theta_3)},$$

$$f^{(2)}_{X_i|X_j}(x_i|x_j) = \theta_i e^{-x_i} x_i (1 - e^{-x_i}(x_i + 1))^{\theta_i - 1},$$

e

$$f^{(3)}_{X_i|X_j}(x_i|x_j) = \frac{\theta_3 (1 - e^{-x_i}(x_i + 1))^{\theta_i}}{(\theta_j + \theta_3)}.$$

Demonstração. O teorema segue prontamente substituindo a fdp conjunta de (X_1, X_2), presente na função (5.4), e a fdp marginal de $X_i (i = 1, 2)$, presente em (5.6), na seguinte relação

$$f_{X_i|X_j}(x_i|x_j) = \frac{f_{X_i,X_j}(x_i, x_j)}{f_{X_j}(x_j)}.$$

■

5.2.3 N-ésimo momento da GEBV

Com base nos resultados obtidos com a construção da fdp marginal de X_i será possível obter o seu n-ésimo.

Teorema 9 *O n-ésimo momento da variável aleatória X_i é dado por*

$$E(X_i^n) = (\theta_i + \theta_3) \sum_{k,m=0}^{\infty} \left(\begin{array}{c} \theta_i + \theta_3 - 1 \\ k \end{array} \right) \left(\begin{array}{c} k \\ m \end{array} \right) (-1)^k (1+k)^{-2-n-m} \Gamma(2+n+m).$$

Demonstração. Na expressão dada por

$$E(X_i^n) = \int_0^{\infty} x_i f(x_i) dx_i,$$

foram utilizados a expansão binomial e a seguinte série infinita

$$(1-x)^a = \sum_{k=0}^{\infty} \left(\begin{array}{c} a \\ k \end{array} \right) (-1)^k x^k$$

para chegar no resultado final, presente no teorema. ■

O n-ésimo momento possui grande importância na Estatística, pois através dele é possível obter a média e a variância, além de que é possível também obter assimetria, curtose entre outras medidas.

5.3 Primeiras impressões da GEBV

A distribuição Gama Exponenciada Bivariada, possue apenas três parâmetros, o que para uma distribuição bivariada é considerado um número pequeno, mesmo assim, ela se mostrou bem flexível, como é possível observar nos gráficos (5.1) e (5.2). Uma outra vantagem é que a GEBV possue funções densidades de probabilidades marginais conhecidas, que é a GE com parâmetros (θ_i e θ_3).

Capítulo 6

Considerações finais

Nesta dissertação foram construídas três novas distribuições a partir da pouco utilizada Gama Exponenciada de Gupta et al. (1998), que possui apenas um parâmetro, e feito um tratamento matemático para cada uma delas. As distribuições criadas foram a distribuição de probabilidade Gama Exponenciada Estendida de Marshall Olkin (GEEMO), Gama Exponenciada Poisson Truncada no Zero (GEPTZ) e a Gama Exponenciada Bivariada. Quanto as distribuições univariadas, pode ser observado que as distribuições são versáteis para modelar a função de risco nas formas crescente, decrescente e unimodal. Também foram calculadas várias expressões, tais como n-ésimo momento; r-ésimo momento de vida média residual; r-ésimo momento de vida média residual invertido; ordenação estocástica; entropias; desvios médios; curvas de Bonferroni e de Lorenz; assimetria, curtose e seus gráficos; estatísticas de ordem e parâmetro $stress-strength$. A vantagem de se utilizar as duas primeiras distribuições apresentadas é que ambas possuem apenas dois parâmetros e sua função risco apresenta mais de uma forma.

Em relação à distribuição Gama Exponenciada Bivariada foi encontrada sua função acumulada; função densidade; função marginal; função condicional e seu n-ésimo momento.

Aplicações em banco de dados também foram realizadas para as distribuições univariadas, onde foi possível observar que as novas distribuições propostas podem ser utilizadas mais efetivamente para proporcionar um melhor ajuste do que outros modelos já conhecidos na literatura.

O objetivo de obter novas distribuições a partir da Gama Exponenciada foi alcançado. Podendo ressaltar ainda que existem muitos estudos a serem explorados, principalmente em relação à Gama Exponenciada Bivariada, que foi introduzida nessa dissertação, e desenvolvida sua parte probabilística.

Referências

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

AKAIKE, H. A new look at the statistical model identification. *IEEE transactions on automatic control*, Ieee, v. 19, n. 6, p. 716–723, 1974.

AL-ZAHRANI, B.; FATTAH, A. A.; NADARAJAH, S.; AHMED, A.-H. N. The exponentiated transmuted weibull geometric distribution with application in survival analysis. *Communications in Statistics-Simulation and Computation*, Taylor & Francis, n. just-accepted, p. 00–00, 2015.

ALKASASBEH, M. R.; RAQAB, M. Z. Estimation of the generalized logistic distribution parameters: Comparative study. *Statistical Methodology*, Elsevier, v. 6, n. 3, p. 262–279, 2009.

ALVARADO, R. U. Aplicação da distribuição poisson zero truncada á produtividade de autores. *Perspectivas em ciência da informação*, v. 9, n. 1, 2008.

ARNOLD, B. C. A note on multivariate distributions with specified marginals. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 62, n. 320, p. 1460–1461, 1967.

ARTUSO, A. R. Entropias de shannon e rényi aplicadas ao reconhecimento de padrões. *Revista CIATEC-UPF*, v. 3, n. 2, p. 56–72, 2012.

ARYAL, G. R.; TSOKOS, C. P. On the transmuted extreme value distribution with application. *Nonlinear Analysis: Theory, Methods & Applications*, Elsevier, v. 71, n. 12, p. e1401–e1407, 2009.

ASGHARZADEH, A.; VALIOLLAHI, R.; RAQAB, M. Z. Stress-strength reliability of weibull distribution based on progressively censored samples. *SORT: statistics and operations research transactions*, Institut d'Estadística de Catalunya, v. 35, n. 2, p. 103–124, 2011.

BALAKRISHNAN, N.; LEUNG, M. Order statistics from the type i generalized logistic distribution. *Communications in Statistics-Simulation and Computation*, Taylor & Francis, v. 17, n. 1, p. 25–50, 1988.

BARROS, A. A. d. A. *Distribuiçoes exponencializadas e estendidas: uma abordagem clássica e bayesiana.* Tese (Doutorado) — UFRPE/BIOMETRIA E ESTATISTICA APLICADA, 2008.

BONFERRONI, C. *Elementi di statistica generale.* [S.l.]: Firenze, SEEBER, 1930.

BOX, G. E.; COX, D. R. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 211–252, 1964.

CARRASCO, J. M.; ORTEGA, E. M.; CORDEIRO, G. M. A generalized modified weibull distribution for lifetime modeling. *Computational Statistics & Data Analysis*, Elsevier, v. 53, n. 2, p. 450–462, 2008.

CASELLA, G.; BERGER, R. L. *Statistical inference.* [S.l.]: Duxbury Pacific Grove, CA, 2002. v. 2.

COLOSIMO, E. A.; GIOLO, S. R. Análise de sobrevivência aplicada. In: *ABE-Projeto Fisher*. [S.l.]: Edgard Blücher, 2006.

CORDEIRO, G. M.; CASTRO, M. de. A new family of generalized distributions. *Journal of statistical computation and simulation*, Taylor & Francis, v. 81, n. 7, p. 883–898, 2011.

CURTISS, J. H. A note on the theory of moment generating functions. *The Annals of Mathematical Statistics*, JSTOR, v. 13, n. 4, p. 430–433, 1942.

ELLIOTT, J.; RAWLINGS, J.; MARKWELL, P.; BARBER, P. Survival of cats with naturally occurring chronic renal failure: effect of dietary management. *Journal of Small Animal Practice*, Wiley Online Library, v. 41, n. 6, p. 235–242, 2000.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal distribution and its applications. *Communications in Statistics-Theory and methods*, Taylor & Francis, v. 31, n. 4, p. 497–512, 2002.

FEROZE, N.; ASLAM, M. Bayesian analysis of exponentiated gamma distribution under type ii censored samples. *Scientific Journal of Pure and Applied Sciences*, v. 1, n. 1, p. 30–39, 2012.

GHANIZADEH, A.; PAZIRA, H.; LOTFI, R. Classical estimations of the exponentiated gamma distribution parameters with presence of k outliers. *Aust J Basic Appl Sci*, v. 5, n. 3, p. 571–579, 2011.

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

GUPTA, P. L.; GUPTA, R. C. On the moments of residual life in reliability and some characterization results. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 12, n. 4, p. 449–461, 1983.

GUPTA, R. C.; GUPTA, P. L.; GUPTA, R. D. Modeling failure time data by lehman alternatives. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 27, n. 4, p. 887–904, 1998.

GUPTA, R. D.; KUNDU, D. Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, Wiley Online Library, v. 41, n. 2, p. 173–188, 1999.

GUPTA, R. D.; KUNDU, D. Exponentiated exponential family: an alternative to gamma and weibull distributions. *Biometrical journal*, Berlin, Akademie-Verlag, 1977-, v. 43, n. 1, p. 117–130, 2001.

HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, Biometrika Trust, v. 57, n. 1, p. 97–109, 1970.

JUNIOR, A. C. R. B. *Distribuiçoes das classes Kumaraswamy generalizada e exponenciada: propriedades e aplicaçoes.* Tese (Doutorado) — Escola Superior de Agricultura Luiz de Queiroz, 2013.

KHAN, R.; KUMAR, D. Lower generalized order statistics from exponentiated gamma distribution and its characterization. In: *ProbStat Forum*. [S.l.: s.n.], 2011. v. 4, p. 25–38.

KOCHAR, S. Stochastic comparisons of order statistics and spacings: a review. *ISRN Probability and Statistics*, Hindawi Publishing Corporation, v. 2012, 2012.

KUMARASWAMY, P. A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, Elsevier, v. 46, n. 1, p. 79–88, 1980.

KUNDU, C.; NANDA, A. K. Some reliability properties of the inactivity time. *Communications in Statistics - Theory and Methods*, Taylor & Francis, v. 39, n. 5, p. 899–911, 2010.

KUNDU, D.; GUPTA, R. D. Bivariate generalized exponential distribution. *Journal of Multivariate Analysis*, Elsevier, v. 100, n. 4, p. 581–593, 2009.

LAWLESS, J. F. *Statistical models and methods for lifetime data.* [S.l.]: John Wiley & Sons, 2011. v. 362.

LEE, C.-S.; TSAI, H.-J. A note on the generalized linear exponential distribution. *Statistics & Probability Letters*, Elsevier, 2017.

LIANG, H.; ZOU, G. Improved aic selection strategy for survival analysis. *Computational statistics & data analysis*, Elsevier, v. 52, n. 5, p. 2538–2548, 2008.

LINHART, H.; ZUCCHINI, W. Finite sample selection criteria for multinomial models. *Statistische Hefte*, Springer, v. 27, n. 1, p. 173–178, 1986.

MAHMOUD, M.; ALAM, F. M. A. The generalized linear exponential distribution. *Statistics & probability letters*, Elsevier, v. 80, n. 11, p. 1005–1014, 2010.

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*, Biometrika Trust, v. 84, n. 3, p. 641–652, 1997.

METROPOLIS NICHOLAS E ROSENBLUTH, A. W. e. R. M. N. e. T. A. H. e. T. E. Equation of state calculations by fast computing machines. *The journal of chemical physics*, AIP Publishing, v. 21, n. 6, p. 1087–1092, 1953.

MORAIS, M. J. C.; PACHECO, A. Ordenação estocástica: Um pouco de história e aplicações. 1997.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; FREIMER, M. The exponentiated weibull family: a reanalysis of the bus-motor-failure data. *Technometrics*, Taylor & Francis, v. 37, n. 4, p. 436–445, 1995.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; KOLLIA, G. D. A generalization of the weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 91, n. 436, p. 1575–1583, 1996.

NADARAJAH, S.; KOTZ, S. The beta exponential distribution. *Reliability engineering & system safety*, Elsevier, v. 91, n. 6, p. 689–697, 2006.

NADARAJAH, S.; KOTZ, S. The exponentiated type distributions. *Acta Applicandae Mathematica*, Springer, v. 92, n. 2, p. 97–111, 2006.

NANDA, A. K.; SINGH, H.; MISRA, N.; PAUL, P. Reliability properties of reversed residual lifetime. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 32, n. 10, p. 2031–2042, 2003.

NASIRI, P.; LOTFI, R.; VEISIPOUR, H. Classical and bayesian estimation of parameters on the generalized exponentiated gamma distribution. *Scientific Research and Essays*, Academic Journals, v. 8, n. 8, p. 309–314, 2013.

PASCOA, M. A. R. de. *Extensões da distribuição gama generalizada: propriedades e aplicações*. Tese (Doutorado) — Escola Superior de Agricultura Luiz de Queiroz, 2012.

PEARSON, K. "das fehlergesetz und seine verallgemeinerungen durch fechner und pearson."a rejoinder. *Biometrika*, JSTOR, v. 4, n. 1/2, p. 169–212, 1905.

PESCIM, R. R. *A distribuição beta generalizada semi-normal*. Tese (Doutorado) — Escola Superior de Agricultura Luiz de Queiroz, 2009.

RAO, G. S.; GHITANY, M.; KANTAM, R. Reliability test plans for marshall-olkin extended exponential distribution. *Applied mathematical sciences*, v. 3, n. 55, p. 2745–2755, 2009.

RÉNYI, A. On measures of entropy and information. 1961.

RISTIĆ, M. M.; NADARAJAH, S. A new lifetime distribution. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 84, n. 1, p. 135–150, 2014.

SANTO, A. P. J. d. E. Extensões da distribuição de lindley para análise de dados de sobrevivência. Universidade Estadual Paulista (UNESP), 2014.

SCHWARZ, G. et al. Estimating the dimension of a model. *The annals of statistics*, Institute of Mathematical Statistics, v. 6, n. 2, p. 461–464, 1978.

SHANNON, C. E. Prediction and entropy of printed english. *Bell system technical journal*, Wiley Online Library, v. 30, n. 1, p. 50–64, 1951.

SHAWKY, A.; BAKOBAN, R. Bayesian and non-bayesian estimations on the exponentiated gamma distribution. *Applied Mathematical Sciences*, v. 2, n. 51, p. 2521–2530, 2008.

SHAWKY, A.; BAKOBAN, R. Order statistics from exponentiated gamma distribution and associated inference. *Int. J. Contemp. Math. Sciences*, v. 4, n. 2, p. 71–91, 2009.

SHAWKY, I. A.; BAKOBAN, R. A. Certain characterizations of the exponentiated gamma distribution. *Journal of Statistics Science*, Journal of Statistics Science, v. 3, n. 2, 2011.

SILVA, G. O. *Modelos de regressao quando a funçao de taxa de falha nao é monótona e o modelo probabilıstico beta Weibull modificada*. Tese (Doutorado) — Escola Superior de Agricultura Luiz de Queiroz, 2008.

SINGH, S. K.; SINGH, U.; KUMAR, D. Bayesian estimation of the exponentiated gamma parameter and reliability function under asymmetric loss function. *REVSTAT–Stat J*, v. 9, n. 3, p. 247–260, 2011.

SINGH, U.; SINGH, S. K.; YADAV, A. S. Bayesian estimation for exponentiated gamma distribution under progressive type-ii censoring using different approximation techniques. *Journal of Data Science*, v. 13, n. 3, 2015.

SMITH, R. L.; NAYLOR, J. A comparison of maximum likelihood and bayesian estimators for the three-parameter weibull distribution. *Applied Statistics*, JSTOR, p. 358–369, 1987.

STEELE, J. C.; GUESS, F. M.; YOUNG, T. M.; EDWARDS, D. J. Mean residual life. In: *International Encyclopedia of Statistical Science*. [S.l.]: Springer, 2011. p. 791–792.

THOMAS, A.; JOSE, K. Bivariate semi-pareto minification processes. *Metrika*, Springer, v. 59, n. 3, p. 305–313, 2004.

WILCOX, R. R. Comparing the means of two independent groups. *Biometrical Journal*, Wiley Online Library, v. 32, n. 7, p. 771–780, 1990.

ZAKERZADEH, H.; MAHMOUDI, E. A new two parameter lifetime distribution: model and properties. *arXiv preprint arXiv:1204.4248*, 2012.

ZHANG, T.; XIE, M. Failure data analysis with extended weibull distribution. *Communications in Statistics Simulation and Computation®*, Taylor & Francis, v. 36, n. 3, p. 579–592, 2007.

Apêndice

A
Equações

No teorema 2 a expressão $\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)}$ é dada por:

$$
\begin{aligned}
&\left(-2\left(1.0 - \mathrm{e}^{-x}x - 1.0\mathrm{e}^{-x}\right)^{\theta}\theta\mathrm{e}^{-x}x\left(\alpha_1 - \alpha_2\right)\right) \\
&\div \left(\left(\left(1.0 - \mathrm{e}^{-x}x - 1.0\mathrm{e}^{-x}\right)^{\theta}\alpha_2 - \alpha_2 - \left(1.0 - \mathrm{e}^{-x}x - 1.0\mathrm{e}^{-x}\right)^{\theta}\right)\right. \\
&\left(\left(1.0 - \mathrm{e}^{-x}x - 1.0\mathrm{e}^{-x}\right)^{\theta}\alpha_1 - \alpha_1 - \left(1.0 - \mathrm{e}^{-x}x - 1.0\mathrm{e}^{-x}\right)^{\theta}\right) \\
&\left.\left(-1.0 + \mathrm{e}^{-x}x + 1.0\mathrm{e}^{-x}\right)\right)
\end{aligned}
$$

A função 3.30, que representa a entropia de Shannon da distribuição GEEMO é dada por:

$$
\begin{aligned}
&H(f) = -\int_0^{\infty} f(x)\ln(f(x))dx = \\
&-\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\binom{\theta - 1 + \theta j}{i}\binom{i}{k}(-1)^{i+j}(j+1)\tfrac{1}{\alpha}\left(\tfrac{1-\alpha}{\alpha}\right)^{j}\theta \\
&\times(\ln\left(\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{i}\binom{\theta - 1 + \theta j}{i}\binom{i}{k}(-1)^{i+j}(j+1)\tfrac{1}{\alpha}\left(\tfrac{1-\alpha}{\alpha}\right)^{j}\theta\right)(1+i)^{-2-k}\Gamma(2+k)) \\
&-(1+i)^{-3-k}\Gamma\left(3+k\right)\left(1+i\right) + \tfrac{\left(\left(-k^2-k\right)\ln(1+i)+\left(k^2+k\right)\Psi(k)+2k+1\right)\Gamma(k)(1+k)}{(1+i)^k(1+i)^2})
\end{aligned}
\tag{A.1}
$$

Onde $\Psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}$.

O Parâmetro Stress Strength para $(\alpha_1, \theta_1) \neq (\alpha_2, \theta_2)$ da distribuição GEEMO é dado por:

$$
\begin{aligned}
R &= \int_0^{\infty} \frac{\alpha_1\theta_1 x e^{-x}(1 - e^{-x}(x+1))^{\theta_1 - 1}}{\left[1 - (1-\alpha_1)(1 - (1 - e^{-x}(x+1))^{\theta_1})\right]^2} \times \frac{(1 - e^{-x}(x+1))^{\theta_2}}{1 - \bar{\alpha}_2(1 - (1 - e^{-x}(x+1))^{\theta_2})}dx \\
&= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{i=0}^{\infty}\sum_{h=0}^{i}\binom{-1}{j}\binom{i}{h}\binom{-2}{k}\binom{\theta_1(1+j) + \theta_2(1+k) - 1}{i}(-1)^i \\
&\times \frac{1}{\alpha_1\alpha_2}\left(\frac{1-\alpha_2}{\alpha_2}\right)^j\left(\frac{1-\alpha_1}{\alpha_1}\right)^k \alpha_1\theta_1(1+i)^{-2-h}\Gamma(2+h)
\end{aligned}
\tag{A.2}
$$

A função 3.30, que representa a entropia de Shannon da distribuição GEPTZ é dada por:

$$\begin{aligned}
&H(f) = -\int_0^\infty f(x)\ln(f(x))dx = \\
&-\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \tfrac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \left(\begin{array}{c} \theta(i+1)-1 \\ j \end{array}\right) \left(\begin{array}{c} j \\ k \end{array}\right) (-1)^j \\
&\times((1+j)^{-2-k}\Gamma(2+k)\ln\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{j} \tfrac{\alpha(-\alpha)^i\theta}{(1-e^{-\alpha})i!} \left(\begin{array}{c} \theta(i+1)-1 \\ j \end{array}\right) \left(\begin{array}{c} j \\ k \end{array}\right) (-1)^j\right) \\
&+\tfrac{\left(\left(-k^2-k\right)\ln(1+j)+\left(k^2+k\right)\Psi(k)+2k+1\right)(1+k)\Gamma(k)}{(1+j)^k(1+j)^2} - (1+j)^{-3-k}\Gamma\left(3+k\right)\left(1+j\right))
\end{aligned} \tag{A.3}$$

Onde $\Psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}$.

No teorema 4 a expressão $\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)}$ é dada por:

$$\frac{(1-e^{-x}(x+1))^{\theta}\theta e^{-x}x(\alpha_1-\alpha_2)}{-1+e^{-x}(x+1)}.$$