

UNIVERSIDADE FEDERAL DO AMAZONAS  
INSTITUTO DE CIÊNCIAS EXATAS  
PÓS-GRADUAÇÃO EM MATEMÁTICA

**INFERÊNCIA ESTATÍSTICA PARA O PROCESSO GEOMÉTRICO BASEADO NA  
DISTRIBUIÇÃO BIRNBAUM-SAUNDERS REPARAMETRIZADA**

NAOR LIMA DE SOUZA

MANAUS-AM

2020

Naor Lima de Souza<sup>1</sup>

INFERÊNCIA ESTATÍSTICA PARA O PROCESSO GEOMÉTRICO BASEADO NA  
DISTRIBUIÇÃO BIRNBAUM-SAUNDERS REPARAMETRIZADA

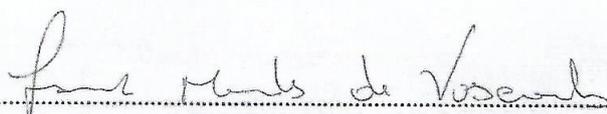
Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Amazonas, como requisito final para obtenção do título de Mestre em Matemática, área de concentração em Estatística.

Aprovado por:

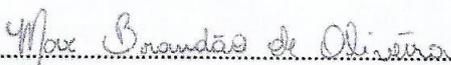
BANCA EXAMINADORA



Prof. Dr. Jeremias Da Silva Leão (Orientador)  
Universidade Federal do Amazonas – UFAM



Prof. Dr. Josimar Mendes de Vasconcelos (Membro Externo)  
Universidade Federal Rural de Pernambuco – UFRPE



Prof. Dr. Max Brandão de Oliveira (Membro Externo)  
Universidade Federal do Piauí - UFPI

---

<sup>1</sup> Foi bolsista durante todo o mestrado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.

### Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

S729i Souza, Naor Lima de  
Inferência Estatística para o Processo Geométrico baseado na  
distribuição Birnbaum-Saunders Reparametrizada / Naor Lima de  
Souza . 2020  
47 f.: il. color; 31 cm.

Orientador: Jeremias da Silva Leão  
Dissertação (Mestrado em Matemática - Estatística) -  
Universidade Federal do Amazonas.

1. Distribuição Birnbaum-Saunders Reparametrizada. 2.  
Inferência Estatística. 3. Modelo GP-RBS. 4. Processo Geométrico.  
I. Leão, Jeremias da Silva. II. Universidade Federal do Amazonas  
III. Título



**UFAM**

INFERÊNCIA ESTATÍSTICA PARA O PROCESSO GEOMÉTRICO BASEADO NA  
DISTRIBUIÇÃO BIRNBAUM–SAUNDERS REPARAMETRIZADA

Naor Lima de Souza

Dissertação de Mestrado apresentada ao  
Programa de Pós-graduação em Matemática,  
da Universidade Federal do Amazonas, como  
parte dos requisitos necessários à obtenção do  
título de Mestre em Matemática

Orientador: Jeremias da Silva Leão

Manaus

Fevereiro de 2021

# Agradecimentos

A Deus, por ser o meu guia nessa jornada, por me conceder sabedoria e por ter proporcionado a mim o gosto pela Matemática e Estatística, a vontade de ajudar, contribuir e lutar por uma educação melhor para este país.

A minha família, principal suporte para eu ter chegado até aqui. Desde meu pai, José Paz, o qual é o meu espelho, a minha mãe, Maria Trindade, meu irmão, Jessé Lima, e minha irmã, Mirian Lima, pelas mensagens e pelo exemplo que são na minha vida. Não tenho dúvidas que este apoio foi o alicerce fundamental para a minha formação como pessoa.

Ao Professor Doutor Jeremias da Silva Leão, pelo imenso empenho e dedicação para a realização deste trabalho. Pelas suas orientações que foram preponderantes para chegarmos a esse resultado.

A banca examinadora deste trabalho: o Prof. Dr. Max Brandão de Oliveira, pelas brilhantes considerações que guiaram a confecção final deste trabalho, e também o Prof. Dr. Josimar Mendes Vasconcelos cujas observações foram igualmente imprescindíveis

A todos os professores do Programa de Pós-Graduação em Matemática (PPGM) pelas aulas e ensinamentos transmitidos.

Aos amigos de mestrado, Antônio Ferreira (O coronel), Josemar Alves (O 12) e a Paola Martins pela parceria nesses anos de formação, pela união e pelas múltiplas conversas.

Ao Instituto de Ciências Exatas.

À Universidade Federal do Amazonas.

À CAPES pelo suporte financeiro.

E, finalmente, a todos que, com boa intenção, colaboraram diretamente ou indiretamente para a realização e finalização deste trabalho.

Resumo da Dissertação apresentada ao Programa de Pós-Graduação em Matemática, da Universidade Federal do Amazonas, como parte dos requisitos necessários para a obtenção do grau de Mestre em Matemática. (M.Sc.)

## INFERÊNCIA ESTATÍSTICA PARA O PROCESSO GEOMÉTRICO BASEADO NA DISTRIBUIÇÃO BIRNBAUM–SAUNDERS REPARAMETRIZADA

Naor Lima de Souza

Orientador: Jeremias da Silva Leão

Área de Concentração: Estatística

Neste trabalho, investigamos o problema da inferência estatística para o processo geométrico (GP) quando a distribuição do tempo da primeira ocorrência for Birnbaum–Saunders Reparametrizada (RBS). Para isto, propomos um modelo intitulado GP-RBS capaz de atuar nesse problema. São obtidos os estimadores de máxima verossimilhança (EMV), estimadores via método dos momentos (EMM) e estimadores de momentos modificados (MM) para os parâmetros do GP. Além disso, discutimos algumas propriedades assintóticas importantes desses estimadores, como distribuição assintótica e consistência. Em seguida, realizamos alguns estudos de simulações de Monte Carlo com diferentes valores de parâmetros para comparar os desempenhos dos resultados obtidos e, por fim, apresentamos as aplicações em quatro conjuntos de dados reais.

**Palavras-chave:** Distribuição Birnbaum-Saunders reparametrizada, inferência estatística; modelo GP-RBS; processo geométrico.

Abstract of Dissertation presented to Postgraduate in Mathematics, of the Federal University of Amazonas, as a partial fulfillment of the requirements for the degree of Master of Mathematics. (M.Sc.)

STATISTICAL INFERENCE FOR GEOMETRIC PROCESSES WITH  
REPARAMETERIZED BIRNBAUM–SAUNDERS DISTRIBUTION

Naor Lima de Souza

Advisor: Jeremias da Silva Leão

Research lines: Statistics

In this monography, we investigate the problem of statistical inference for the geometric process (GP) when the distribution of the first occurrence time is Reparameterized Birnbaum-Saunders (RBS). We propose a model entitled GP-RBS capable of acting in this problem. Maximum likelihood estimators (EMV), estimators by the moment method (MME) and modified moment estimators (MM) are obtained. In addition, we discuss some important asymptotic properties of these estimators, such as asymptotic distribution and consistency. Next, we conduct some studies of Monte Carlo simulations with different parameter values to compare the performance of the results obtained and, finally, we present the applications in four real data sets.

**Keyword:** Geometric process; GP-RBS model; reparameterized Birnbaum–Saunders distribution; statistical inference.

# Sumário

<b>Lista de Figuras</b>	<b>vi</b>
<b>Lista de Tabelas</b>	<b>viii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Contextualização . . . . .	1
1.2 Estrutura da dissertação . . . . .	3
<b>2 Notações e definições preliminares</b>	<b>4</b>
2.1 Processo de Poisson, renovação e o processo geométrico . . . . .	4
2.2 A distribuição BS . . . . .	6
2.3 A distribuição RBS . . . . .	8
2.3.1 Estimação clássica para a distribuição RBS . . . . .	13
<b>3 O modelo GP-RBS</b>	<b>16</b>
3.1 Estimador de máxima verosimilhança . . . . .	18
3.2 Estimação por momentos modificados . . . . .	19
3.3 Estimação pelo método dos momentos . . . . .	19
<b>4 Aplicação numérica, resultados e discussões</b>	<b>21</b>
4.1 Estudo de simulação de Monte Carlo . . . . .	22
4.2 Aplicações em dados reais . . . . .	32
4.2.1 Dados 1 - aircraft . . . . .	32
4.2.2 Dados 2 - Coal Mining . . . . .	34
4.2.3 Dados 3 - Carnero . . . . .	35

4.2.4	Dados 4 - jiwani . . . . .	36
<b>5</b>	<b>Conclusões e trabalhos futuro</b>	<b>38</b>
<b>A</b>	<b>Matriz Informação de Fisher</b>	<b>40</b>
	<b>Referências Bibliográficas</b>	<b>43</b>

# Lista de Figuras

2.1	Gráficos da PDF $BS(\alpha, \beta)$ : (a) para diferentes valores de $\alpha$ com $\beta = 1.0$ ; (b) para $\alpha = 0.1$ e diferentes valores de $\beta$ . . . . .	7
2.2	Gráficos da PDF da distribuição $RBS(\varphi, \kappa)$ : (a) para diferentes valores de $\varphi$ com $\kappa = 100.0$ ; (b) para $\varphi = 1.0$ e diferentes valores de $\kappa$ . . . . .	10
2.3	Comparação entre a média e a mediana da RBS. . . . .	10
4.1	Distribuição empírica de $\widehat{\varpi}$ , para $\varpi = 0.90$ , com $\varphi = 2.00$ e $IC(\varpi; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	26
4.2	Distribuição empírica de $\widehat{\varphi}$ , para $\varphi = 2.00$ , com $\varpi = 0.90$ e $IC(\varphi; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	27
4.3	Distribuição empírica de $\widehat{\kappa}$ , para $\kappa = 0.50$ , em que $\varpi = 0.90$ e $\varphi = 2.00$ , com $IC(\kappa; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	27
4.4	Distribuição empírica de $\widehat{\varpi}$ , para $\varpi = 1.10$ , com $\varphi = 2.00$ e $IC(\varpi; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	28
4.5	Distribuição empírica de $\widehat{\varphi}$ , para $\varphi = 2.00$ , com $\varpi = 1.10$ e $IC(\varphi; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	28
4.6	Distribuição empírica de $\widehat{\kappa}$ , para $\kappa = 0.50$ , em que $\varphi = 2.00$ e $\varpi = 1.10$ , com $IC(\kappa; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	29
4.7	Distribuição empírica de $\widehat{\varpi}$ , para $\varpi = 0.90$ , com $\varphi = 1.50$ e $IC(\varpi; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	29
4.8	Distribuição empírica de $\widehat{\varphi}$ , para $\varphi = 1.50$ , com $\varpi = 0.90$ e $IC(\varphi; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	30
4.9	Distribuição empírica de $\widehat{\varpi}$ , para $\varpi = 1.10$ , com $\varphi = 1.50$ e $IC(\varpi; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	30

4.10	Distribuição empírica de $\hat{\varphi}$ , para $\varphi = 1.50$ , com $\varpi = 1.10$ e $IC(\varphi; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	31
4.11	Distribuição empírica de $\hat{\kappa}$ , para $\kappa = 0.50$ , em que $\varphi = 1.50$ e $\varpi = 0.90$ , com $IC(\kappa; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	31
4.12	Distribuição empírica de $\hat{\kappa}$ , para $\kappa = 0.50$ , em que $\varphi = 1.50$ e $\varpi = 1.10$ , com $IC(\kappa; [1-\alpha] \times 100\%)$ para cada $n$ . . . . .	32

# Lista de Tabelas

4.1	Medidas descritivas para os estimadores EMV e EMM dos parâmetros $\varpi$ , $\varphi$ e $\kappa$ , quando $\varphi = 1.0$ . . . . .	23
4.2	Medidas descritivas para os estimadores EMV e EMM dos parâmetros $\varpi$ , $\varphi$ e $\kappa$ , quando $\varphi = 1.5$ . . . . .	24
4.3	Medidas descritivas para os estimadores EMV e EMM dos parâmetros $\varpi$ , $\varphi$ e $\kappa$ , quando $\varphi = 2.0$ . . . . .	25
4.4	Estatísticas descritivas para os Dados 1. . . . .	33
4.5	Estimativas dos parâmetros $\varpi$ , $\varphi$ e $\kappa$ para os Dados 1. . . . .	33
4.6	Estimativas da log-verossimilhança, AIC, BIC e BF segundo cada modelo para os Dados 1. . . . .	33
4.7	Estatísticas descritivas para os Dados 2. . . . .	34
4.8	Estimativas dos parâmetros $\varpi$ , $\varphi$ e $\kappa$ para os Dados 2. . . . .	34
4.9	Estimativas da log-verossimilhança para, AIC, BIC e BF para cada modelo considerando os Dados 2. . . . .	35
4.10	Estatísticas descritivas para os Dados 3. . . . .	35
4.11	Estimativas dos parâmetros $\varpi$ , $\varphi$ e $\kappa$ para os Dados 3. . . . .	36
4.12	Estatísticas descritivas para os Dados 4. . . . .	36
4.13	Estimativas dos parâmetros $\varpi$ , $\varphi$ e $\kappa$ para os Dados 4. . . . .	37

# Capítulo 1

## Introdução

### 1.1 Contextualização

Na tentativa de melhorar um sistema ou modelar uma deterioração, quando se refere a conjunto de dados com tempos de ocorrência de eventos sucessivos, com alguma tendência, em diferentes áreas, como ciências, engenharia, medicina e ciências sociais, uma abordagem mais direta é a introdução de um modelo de processo monótono; ver por exemplo Lam (2007), Lam (1988a), Kara et al. (2015) e Bicer et al. (2018). Nesse ramo, o processo geométrico (GP) se apresenta como um instrumento importante para modelar esses sistemas em deterioração, já que possibilita, diferentemente do processo de renovação (RP), que atua quando os dados em intervalos de tempos sucessivos são independentes e identicamente distribuídos (IID)(sem tendência), uma abordagem simples, flexível e não necessita que os dados sejam sem tendência ou distribuídos identicamente. Isto é extremamente vantajoso pois os dados, geralmente, contêm uma tendência monótona devido ao efeito de envelhecimento e ao desgaste acumulado, como sugere Chan et al. (2014).

No GP a suposição sobre a distribuição de  $X_1$ , que é a variável do tempo de primeira passagem, tem um significado especial devido ao fato de a distribuição de  $X_1$  e as outras variáveis aleatórias  $X_2, \dots, X_n$  serem da mesma família de distribuições com um conjunto diferente de parâmetros, ou seja,  $X_i$ 's ( $i = 1, 2, \dots, n$ ) são distribuídos independentemente, mas não de forma idêntica, como destacam Kara et al. (2015), Chan et al. (2014), Lam e Chan (1998) e Bicer (2018).

Assim, já que a distribuição da variável aleatória  $X_1$  determina a distribuição das outras variáveis, a seleção da distribuição de  $X_1$  com base nos dados observados é muito importante para a inferência estatística, como ressaltado em Bicer e Bicer (2017).

Como uma generalização de um processo de renovação, o GP foi introduzido pela primeira vez e aplicado a um problema de substituição por Lam (1988a). As propriedades teóricas do GP foram estudadas por Lam (1988a), Lam (1988b) e Braun et al. (2005). Atualmente, estudos relacionados ao problema de estimação de parâmetros sobre a distribuição do primeiro tempo de ocorrência têm impulsionado o estudo sobre GP. Lam e Chan (1998) propuseram o processo geométrico baseado na distribuição lognormal. Chan et al. (2014) consideraram a distribuição gama e, em seguida, Aydogdu et al. (2010) trabalharam com distribuição Weibull. O GP, considerando a distribuição gaussiana inversa, foi estudado por Kara et al. (2015). Recentemente, Bicer (2018) usou a distribuição Power Lindley e Bicer et al. (2018) trabalharam com a distribuição Rayleigh. É interessante notar que, em quase todos esses estudos, os métodos de estimação abordados foram o método dos momentos (MM) e o de máxima verossimilhança (MV), entre outros. A partir das análises apontadas em cada um desses trabalhos tem-se a indicação de que o estimador de máxima verossimilhança (EMV) apresentou melhores propriedades quando comparado com os outros estimadores da Inferência Clássica.

Em nosso estudo, vamos considerar que a distribuição do primeiro momento de ocorrência é Birnbaum–Saunders Reparametrizada (RBS), apresentada em Santos-Neto et al. (2012) e Santos-Neto et al. (2014). A distribuição RBS surgiu como uma nova versão da distribuição Birnbaum-Saunders (BS). O modelo de probabilidade BS é comumente conhecido como distribuição de tempo de fadiga e foi proposto por Birnbaum e Saunders (1969) em um ambiente de confiabilidade motivado por problemas de vibração em aeronaves comerciais que causavam fadiga nos materiais. Essa distribuição é assimétrica à direita, contínua e unimodal, com dois parâmetros que modificam sua forma e escala. A distribuição RBS nos permite fazer uma analogia com a parametrização empregada para a distribuição normal ou gaussiana, mas em uma estrutura assimétrica, que é útil para a modelagem de dados de vida útil e confiabilidade. Como destacado em Santos-Neto et al. (2016), uma das vantagens do uso da RBS é que ela não requer a transformação logarítmica da variável resposta, por exemplo no contexto de modelos de regressão quando usa-se a BS usual.

Dessa forma, temos que a motivação deste trabalho dar-se-á devido a dois fatores gerais:

(1) o GP é um processo estocástico simples e tem sido aplicado em problemas de substituição industrial. Em alguns problemas fornece soluções tratáveis para no contexto de confiabilidade, ver Lam (1988a) e Lam et al. (2003). Chan et al. (2012) introduziram uma análise de séries temporais financeiras considerando esses modelos; (2) a distribuição BS e suas versões reparametrizadas têm sido amplamente estudadas devido as suas propriedades teóricas e aplicações em diversas áreas.

Em face disso, o objetivo deste trabalho é apresentar o processo geométrico baseado na distribuição RBS. Nesse sentido serão apresentados: (1) o desenvolvimento dos processos iterativos de estimação baseados nos métodos da máxima verossimilhança, momentos e momentos modificados; (2) estudos de simulação de Monte Carlo (MC) serão condizidos para avaliar a performance dos estimadores propostos; e (3) por fim, será apresentado a aplicação a quatro conjuntos de dados.

## **1.2 Estrutura da dissertação**

O presente trabalho está organizado da seguinte forma: o Capítulo 2 introduz as notações, definições preliminares sobre a construção do GP. Adicionado a isso, apresentamos brevemente conceitos e propriedades sobre as distribuições BS e RBS. Em seguida, no Capítulo 3, introduzimos o modelo GP-RBS e apresentamos o desenvolvimento inferencial para a obtenção dos estimadores dos parâmetros do modelo, baseado no método da máxima verossimilhança, método dos momentos e momentos modificados. A avaliação numérica, isto é, um estudo de simulação de MC para avaliação dos estimadores dos parâmetros do modelo proposto, como também a aplicação a quatro conjuntos de dados são apresentados no Capítulo 4. Uma discussão geral sobre o modelo estudado e as perspectivas para trabalhos futuros estão dispostas no Capítulo 5.

# Capítulo 2

## Notações e definições preliminares

### 2.1 Processo de Poisson, renovação e o processo geométrico

Os três processos de contagem lidam com a análise estatística de um conjunto de dados com tempos de ocorrências de eventos sucessivos, cada um com suas particularidades, e são reduções do processo estocástico, conforme veremos a seguir. Isto porque, um processo estocástico  $\{X(t), t \in \tau\}$  é uma família de variáveis aleatórias, de modo que cada  $t \in \tau$  é uma variável aleatória, em que  $\tau$  é chamado o conjunto dos índices, por definição. Dito isto, as seguintes definições evidenciam tal fato.

**Definição 2.1.** *Um processo estocástico  $\{N(t), t \geq 0\}$  é dito ser um processo de contagem se  $N(t)$  é o número total de eventos ocorridos pelo tempo  $t$ . Propriedades do processo de contagem, são tais que*

1.  $N(t) \geq 0$ ;
2.  $N(t)$  é inteiro;
3. se  $s < t$ , então  $N(s) < N(t)$ ;
4. se  $s < t$ , então  $N(t) - N(s)$  representa o número de eventos ocorridos em  $(s, t]$ .

Com isso, o processo de Poisson é definido como segue:

**Definição 2.2.** *O processo de contagem  $\{N(t), t \geq 0\}$  é chamado Processo de Poisson, com taxa  $\lambda > 0$ , se*

1.  $N(0) = 0$ ;
2. O processo de contagem tem incrementos independentes;
3. Para todos  $s, t \geq 0$ ,

$$P\{N(t+s) - N(s) = n\} = \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, \dots$$

, ver mais detalhes em Lam (2007).

Lam (2007) mostrou que, em um processo de Poisson com taxa  $\lambda$ , os tempos de  $X_n, n = 1, 2, \dots$ , são variáveis aleatórias IID com uma distribuição exponencial, que denotaremos por  $\text{Exp}(\lambda)$ . Dessa forma, considerando que  $\{X_n, n = 1, 2, \dots\}$  é uma sequência variáveis aleatórias IID não-negativas, cada uma com distribuição  $F$ , com  $F(0) = P\{X_n = 0\} < 1$ , denotando

$$\mu = E[X_n] = \int_0^{\infty} x dF(x), \quad 0 < \mu \leq \infty,$$

em que se interpretarmos  $X_n$  como o tempo de inter-chegada entre o  $(n-1)$  e o  $n$ -ésimo evento (ou renovação), então podemos definir o tempo de chegada do  $n$ -ésimo evento por

$$S_n = \sum_{i=1}^n X_i,$$

com  $S_0 = 0$ . Logo, o número de eventos ocorrido pelo tempo  $t$  é então dado por

$$N(t) = \sup\{n : S_n \leq t\}.$$

Com isso  $\{N(t), t \geq 0\}$  é um processo de contagem, o que implica na seguinte definição a seguir.

**Definição 2.3.** O processo de contagem  $\{N(t), t \geq 0\}$  é chamado processo de renovação, com taxa  $\lambda > 0$ , como descrito em Lam (2007).

Quando a distribuição em questão  $F$  for exponencial, o processo de renovação (PR) torna-se um processo de Poisson. Ou melhor, o processo de renovação é uma generalização do processo de Poisson; ver Lam (2007). Por outro lado, a definição 2.4 deixa claro a posição

hierárquica dos processos citados, pondo o GP como o caso mais geral entre esses processos de contagem e, também, um tanto simples de manuseá-lo, como veremos no Capítulo 3.

**Definição 2.4.** *Seja  $X_1, X_2, \dots$  um conjunto de variáveis aleatórias não-negativas. Suponha que exista um número real positivo  $\varpi$  tal que*

$$Y_t = \varpi^{t-1} X_t, \quad t = 1, 2, \dots,$$

*forma um RP com média  $E[Y_t] = \mu$  e variância  $\text{Var}[Y_t] = \sigma^2$ . Então o processo  $\{X_t, t = 1, 2, \dots\}$  é chamado GP com razão  $\varpi$ .*

Seja  $X_t$  um GP, então temos que o valor esperado e a variância são  $E[X_t] = \mu/\varpi^{t-1}$  e  $\text{Var}[X_t] = \sigma^2/\varpi^{2(t-1)}$ , respectivamente. O parâmetro  $\varpi$  determina o comportamento do GP: (i) se  $0 < \varpi \leq 1$ , temos que o GP é estocasticamente não decrescente; (ii) já quando  $\varpi \geq 1$ , o GP é estocasticamente não crescente; e (iii) se  $\varpi = 1$ , o GP torna-se um RP.

## 2.2 A distribuição BS

Birnbaum e Saunders (1969) propuseram a distribuição BS com o objetivo de entender como modelar o tempo de vida de materiais e equipamentos submetidos a cargas dinâmicas, principalmente em problemas de fadiga de materiais e em aviões comerciais, à época. Atualmente, ela é usada extensivamente em aplicações de confiabilidade para modelar tempos de falha, em áreas como engenharia, ciências ambientais, biológicas e financeira, apresentadas em Leiva (2016). Uma variável aleatória  $T$ , com distribuição BS, é considerada como o tempo até a falha e tem a função distribuição acumulada (CDF) expressa como

$$F(t; \alpha, \beta) = P(T \leq t) = \Phi \left[ \frac{1}{\alpha} \left( \sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right], \quad t > 0, \alpha > 0, \beta > 0, \quad (2.1)$$

em que  $\Phi(\cdot)$  representa a CDF distribuição normal padrão, em que  $\alpha$  é o parâmetro de forma e  $\beta$  o parâmetro de escala, que são dados por

$$\alpha = \frac{\sigma}{\sqrt{\mu w}} \quad \text{e} \quad \beta = \frac{w}{\mu},$$

em que os parâmetros  $\mu$  e  $\sigma$  representam, respectivamente, a média e o desvio padrão de uma variável aleatória  $U_j$  que modela extensão da fissura após o  $j$ -ésimo ciclo e  $w$  denota a falha que ocorre quando o tamanho da fissura na amostra de material excede certo nível de resistência, para uma amostra de material submetida a cargas cíclicas de tensão e força, que produzem uma fissura ou desgaste, para detalhes sobre  $U_j$ ; ver Birnbaum e Saunders (1969).

A função densidade de probabilidade (PDF) da distribuição BS, obtida a partir de (2.1), é dada da por

$$f(t; \alpha, \beta) = \frac{1}{2\alpha\beta\sqrt{2\pi}} \left[ \left(\frac{\alpha}{t}\right)^{\frac{1}{2}} + \left(\frac{\alpha}{t}\right)^{\frac{3}{2}} \right] \exp \left\{ -\frac{1}{2\beta^2} \left( \frac{t}{\alpha} + \frac{\alpha}{t} - 2 \right) \right\}, \quad (2.2)$$

com  $t > 0$ , e  $\alpha, \beta > 0$ . Na Figura 2.1 é perceptível que, fixando  $\beta$  e variando  $\alpha$ , há uma alteração na assimetria da distribuição. Agora, fixando  $\alpha$  e variando  $\beta$ , há uma mudança no parâmetro de localização da distribuição, visto que  $\beta$  é a mediana do modelo, conseqüentemente há uma modificação no valor esperado, ou seja, a média da distribuição.

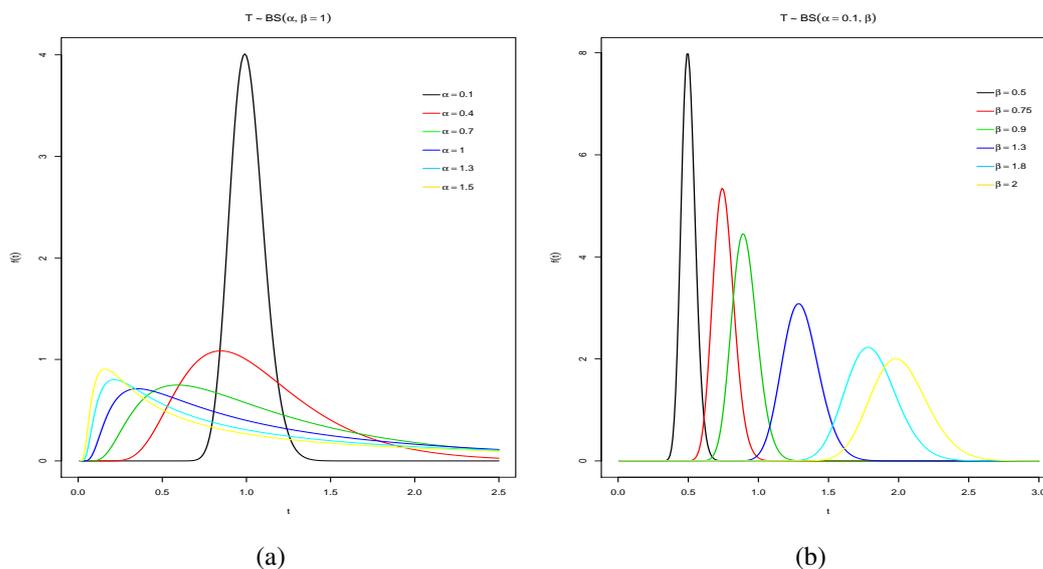


Figura 2.1: Gráficos da PDF  $BS(\alpha, \beta)$ : (a) para diferentes valores de  $\alpha$  com  $\beta = 1.0$ ; (b) para  $\alpha = 0.1$  e diferentes valores de  $\beta$ .

Considerando a seguinte transformação

$$T = \beta \left\{ \frac{\alpha Z}{2} \left[ \left( \frac{\alpha Z}{2} \right)^2 + 1 \right]^{\frac{1}{2}} \right\},$$

em que  $Z$  segue uma distribuição normal padrão, temos que o valor esperado e a variância da BS são dados por

$$E[T] = \beta \left(1 + \frac{\alpha^2}{2}\right) \quad \text{e} \quad \text{Var}[T] = (\alpha\beta)^2 \left(1 + \frac{5\alpha^2}{4}\right).$$

A distribuição BS é unimodal, com assimetria à direita e satisfaz as propriedades recíproca e multiplicação por escalar, isto é, se  $T \sim BS(\alpha, \beta)$ , então  $1/T \sim BS(\alpha, \beta^{-1})$  e também  $\forall b > 0, bT \sim BS(\alpha, b\beta)$ , conseqüentemente, temos que

$$E[T^{-1}] = \beta^{-1} \left(1 + \frac{\alpha^2}{2}\right) \quad \text{e} \quad \text{Var}[T^{-1}] = \left(\frac{\alpha}{\beta}\right)^2 \left(1 + \frac{5\alpha^2}{4}\right).$$

Vale ressaltar que, na literatura atual, para essa parametrização da BS, há diversos trabalhos científicos. Tomando como exemplo, Leiva et al. (2007) e Lemonte e Cordeiro (2009) consideraram o modelo de regressão não linear baseado na BS. Rieck e Nedelman (1991) consideraram o modelo log-linear para a distribuição BS. Também há outras extensões para a distribuição BS: como descreve Leiva (2016), vários autores estenderam e generalizaram esta distribuição. A primeira extensão da distribuição Birnbaum-Saunders se deve a Volodin e Dzhungurova (2000); em seguida, Diaz-Garcia e Leiva (2005) apresentaram a distribuição generalizada Birnbaum-Saunders. Gomez et al. (2009) estenderam a distribuição BS a partir do modelo slash-elíptico. Ahmed et al. (2010) analisaram uma versão truncada da distribuição Birnbaum-Saunders; Kotz et al. (2010) estudaram modelos de mistura relacionados com a distribuição Birnbaum-Saunders; Vilca et al. (2010) e Castillo et al. (2009) desenvolveram a distribuição epsilon-skew BS; Balakrishnan et al. (2011) consideraram distribuições de mistura BS. Cordeiro e Lemonte (2011) definiram a distribuição  $\beta$ -Birnbaum-Saunders. Leiva et al. (2012) e Santos-Neto et al. (2014) reparametrizaram a distribuição BS obtendo propriedades interessantes; Saulo et al. (2012) apresentaram a distribuição Kumaraswamy BS; Fierro et al. (2013) geraram a distribuição Birnbaum-Saunders de um processo Poisson não-homogógeno; e Bourguignon et al. (2015) derivaram a distribuição transmutada Birnbaum-Saunders.

## 2.3 A distribuição RBS

Santos-Neto et al. (2012) propuseram onze novas parametrizações do modelo BS com diferentes justificativas. Dentre estas, temos uma que é indexada por dois parâmetros relacionados à média e precisão da distribuição dos dados, respectivamente. Assim sendo, temos que os parâmetros da RBS em termos da BS original são definidos como  $\varphi = \beta(1 + \alpha^2/2)$  e  $\kappa = 2/\alpha^2$ , em que  $\alpha > 0$  e  $\beta > 0$  são os parâmetros originais da BS; ver por exemplo Birnbaum e Saunders (1969),  $\varphi > 0$  é um parâmetro de escala, nesse caso a média da distribuição, enquanto que  $\kappa > 0$  é um parâmetro de forma e precisão. A notação  $Y \sim \text{RBS}(\varphi, \kappa)$  é usada quando a variável aleatória  $Y$  segue essa distribuição. A média e a variância de  $Y$  são dadas por  $E[Y] = \varphi$  e  $\text{Var}[Y] = \varphi^2/\phi$ , respectivamente, em que  $\phi = (\kappa + 1)^2/(2\kappa + 5)$ . Como mencionado,  $\kappa$  pode ser interpretado como um parâmetro de precisão porque, para valores fixos de  $\varphi$ , quando  $\kappa$  tende ao infinito, a variação de  $Y$  tende a zero.

Além do que, para  $\varphi$  fixo, quando  $\kappa$  se aproximar de zero, então  $\text{Var}[Y]$  tende a  $5\varphi^2$ . Se  $Y \sim \text{RBS}(\varphi, \kappa)$ , então a CDF e PDF são expressas por

$$F_Y(y; \varphi, \kappa) = \Phi \left( \sqrt{\frac{\kappa}{2}} \left( \sqrt{\frac{(\kappa + 1)y}{\kappa\varphi}} - \sqrt{\frac{\kappa\varphi}{y(\kappa + 1)}} \right) \right), \quad y > 0, \quad (2.3)$$

e

$$f_Y(y; \varphi, \kappa) = \frac{\exp(\kappa/2)\sqrt{\kappa + 1}}{4y^{3/2}\sqrt{\pi\varphi}} \left[ y + \frac{\kappa\varphi}{\kappa + 1} \right] \times \exp \left( -\frac{\kappa}{4} \left[ \frac{y(\kappa + 1)}{\kappa\varphi} + \frac{\kappa\varphi}{y(\kappa + 1)} \right] \right). \quad (2.4)$$

A Figura 2.2 mostra o comportamento gráfico da distribuição  $\text{RBS}(\varphi, \kappa)$  para diferentes valores de  $\varphi$  e  $\kappa$ . Note que, para  $\kappa = 100.0$  com  $\varphi$  variando, percebe-se que  $\varphi$  controla a escala e a média da PDF; enquanto que para  $\varphi = 1.0$ ,  $\kappa$  regula a forma da PDF. Ainda com a mudança de  $\varphi$ , mantendo  $\kappa$  fixo, há uma alteração na curtose da variável aleatória e com o aumento de  $\kappa$ , há um aumento da assimetria.

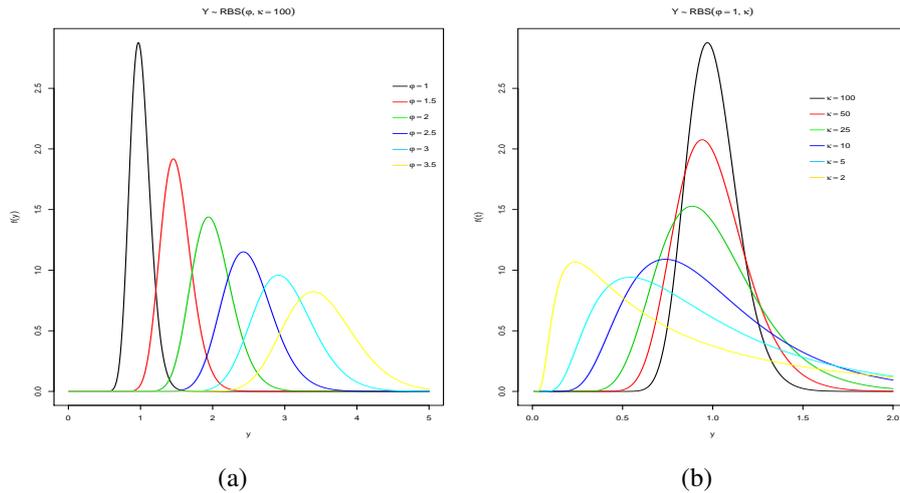


Figura 2.2: Gráficos da PDF da distribuição RBS( $\varphi, \kappa$ ): (a) para diferentes valores de  $\varphi$  com  $\kappa = 100.0$ ; (b) para  $\varphi = 1.0$  e diferentes valores de  $\kappa$ .

A Figura 2.3 retrata uma comparação entre a média e a mediana dos modelos RBS e BS, respectivamente. Podemos notar que à medida que o parâmetro de escala  $\kappa$  cresce, este se aproxima da média, ou seja, temos que dependendo do valor do parâmetro  $\kappa$  os resultados obtidos a partir das duas parametrizações serão bem próximos. Apesar disso, é importante destacar que, no contexto de modelos regressão, a parametrização pela média tem a vantagem em termos interpretativos, visto que não há necessidade de transformação na variável resposta; detalhes podem ser encontrados em Santos-Neto et al. (2012) e Santos-Neto et al. (2014).

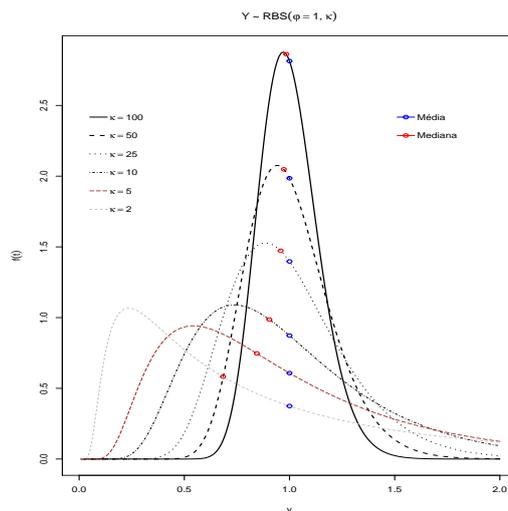


Figura 2.3: Comparação entre a média e a mediana da RBS.

Considerando a equação (2.2) e a parametrização proposta em Santos-Neto et al. (2014) temos que a RBS assume a representação abaixo

$$Z = \sqrt{\frac{\kappa}{2}} \left[ \sqrt{\frac{(\kappa+1)Y}{\varphi\kappa}} - \sqrt{\frac{\varphi\kappa}{(\kappa+1)Y}} \right]. \quad (2.5)$$

A partir desta podemos obter

$$V = Z^2 = \frac{\kappa}{2} \left[ \frac{(\kappa+1)Y}{\varphi\kappa} + \frac{\varphi\kappa}{(\kappa+1)Y} - 2 \right], \quad (2.6)$$

em que  $V$  segue uma distribuição qui-quadrado com 1 grau de liberdade,  $\chi_{(1)}^2$ , tal que  $E(V) = 1$  e  $\text{Var}(V) = 2$ , ou ainda

$$Y^2 - \frac{\varphi\kappa}{(\kappa+1)} \left( 2 + 2\frac{V}{\kappa} \right) Y + \left[ \frac{\varphi\kappa}{(\kappa+1)} \right]^2; \quad (2.7)$$

ver detalhes em Leiva (2016). Através do gerador de números pseudo-aleatórios para a distribuição BS apresentado em Chang e Tang (1994) podemos obter números pseudo-aleatórios para a distribuição RBS. Para isso, note que a equação (2.7) possui duas soluções: as raízes  $y_1$  e  $y_2$ , em que  $y_1 y_2 = [\varphi\kappa/(\kappa+1)]^2$ , conseqüentemente temos que  $y_2$  pode ser expresso em função de  $y_1$  (vice-versa). Nesse caso teremos que considerar um método de geração de números pseudo-aleatórios para transformações que possuem múltiplas raízes, isto foi abordado em Michael et al. (1976). Segundo Leiva (2016), este método pode ser adaptado para a distribuição BS porque consiste em duas variáveis aleatórias contínuas, digamos  $V$  e  $Y$ , definida por  $V = g(Y)$ , ou seja,  $V$  é uma função de  $Y$  através de  $g$ . Conseqüentemente, podemos adaptá-lo à distribuição RBS. O referido método usa uma relação do tipo  $V = g(Y)$ , que tem  $n$  raízes denotadas por  $y_1, y_2, \dots, y_n$ ; um dos quais corresponde a um valor  $y_0$  com probabilidade

$$p_i(y_0) = \left[ 1 + \sum_{j=i \neq 1}^n \frac{|g'(y_i)| f_Y(y_j)}{|g'(y_j)| f_Y(y_i)} \right]^{-1}, \quad (2.8)$$

sendo  $g'(\cdot)$  a derivada de  $g(\cdot)$  dada em 2.8,  $|x|$  é o valor absoluto de  $x$  e  $f_Y(\cdot)$  a função de

densidade de probabilidade de  $Y$ . Para a distribuição RBS, temos  $n = 2$  (onde  $n$  é o número de parâmetros) e a probabilidade  $p_1(y_0) = 1 - p_2(y_0)$  dada em (2.8), a partir da qual devemos escolher a menor raiz, digamos  $y_1$ , sendo igual a  $1/2$ . Isso é obtido a partir da (2.8), em que considerando  $y_1 y_2 = [\varphi \kappa (\kappa + 1)]^2$  e  $y_2 = [\varphi \kappa (\kappa + 1)]^2 / y_1$  temos que

$$\frac{g'(y_1)}{g'(y_2)} = \frac{(\kappa + 1)/\varphi \kappa - \varphi \kappa / (\kappa + 1) y_1^2}{(\kappa + 1)/\varphi \kappa - \varphi \kappa / (\kappa + 1) y_2^2} = [\varphi \kappa / (\kappa + 1) y_1]^2. \quad (2.9)$$

e

$$\begin{aligned} \frac{f_Y(y_j)}{f_Y(y_i)} &= \frac{y_1^{3/2} [y_2 + \varphi \kappa / (\kappa + 1)]}{y_2^{3/2} [y_1 + \varphi \kappa / (\kappa + 1)]} \\ &\quad \times \exp \left\{ -\frac{\kappa}{4} \left[ \left( \sqrt{y_2 (\kappa + 1) / \varphi \kappa} - \sqrt{\varphi \kappa / y_2 (\kappa + 1)} \right)^2 \right. \right. \\ &\quad \left. \left. + \left( \sqrt{y_1 (\kappa + 1) / \varphi \kappa} - \sqrt{\varphi \kappa / y_1 (\kappa + 1)} \right)^2 \right] \right\} \\ &= [y_1 (\kappa + 1) / \varphi \kappa]^2. \end{aligned} \quad (2.10)$$

Dessa forma, das equações (2.9) (2.10), por meio de (2.8), é comprovado que  $p_1(y_0) = 1 - p_2(y_0) = 1/2$ ; ver Leiva (2016). Assim, uma variável aleatória com distribuição RBS pode ser gerada por

$$y_1 I_{[-\infty, 0.5]}(u) + y_2 I_{[0.5, \infty]}, \quad (2.11)$$

em que  $I_A(\cdot)$  é a função indicadora do conjunto  $A$ . Logo, a expressão dada em (2.11), de acordo com Michael et al. (1976), garante que a probabilidade de escolher uma das duas raízes  $y_1$  ou  $y_2$ , dado que  $V = v$ , é igual a  $1/2$ . Assim, o Algoritmo 1, disposto abaixo, descreve os passos para gerar números aleatórios para a distribuição RBS:

---

**Algoritmo 1** – Gerador de números aleatórios da distribuição RBS

---

- 1: Gerar um número aleatório  $u$  de uma variável aleatória  $U \sim U(0, 1)$
  - 2: Gerar um número aleatório  $z$  de uma variável aleatória  $Z \sim N(0, 1)$
  - 3: Definir valores para  $\varphi$  e  $\kappa$  de  $Y \sim \text{RBS}(\varphi, \kappa)$
  - 4: Calcular um número aleatório  $y = y_1$  ou  $y = y_2$  de  $Y \sim \text{RBS}(\varphi, \kappa)$ , por meio dos seguintes critérios:
    - i: Se  $u \leq 0.5$ , então
$$y_1 = \frac{\left(\frac{\varphi\kappa}{(\kappa+1)}\left(2+2\frac{V}{\kappa}\right)\right) - \sqrt{\left(\frac{\varphi\kappa}{(\kappa+1)}\left(2+2\frac{V}{\kappa}\right)\right)^2 - 4\frac{\varphi\kappa}{(\kappa+1)}}}{2};$$
    - ii: Caso contrário, isto é, se  $u > 0.5$ , faça
$$y_2 = \frac{\left(\frac{\varphi\kappa}{(\kappa+1)}\left(2+2\frac{V}{\kappa}\right)\right) + \sqrt{\left(\frac{\varphi\kappa}{(\kappa+1)}\left(2+2\frac{V}{\kappa}\right)\right)^2 - 4\frac{\varphi\kappa}{(\kappa+1)}}}{2};$$
  - 4: Repetir as etapas de 1 a 4 até que a quantidade necessária de números seja obtida
- 

Por fim, vale ressaltar que, o diferencial entre as distribuições BS e RBS está, por exemplo, nas estimativas de momentos: para a parametrização original, não há uma forma fechada; porém, isto é possível com a reparametrização de Santos-Neto et al. (2012). Além disso, essa reparametrização permite que a variável resposta possa ser modelada em sua escala original; ver por exemplo Leiva et al. (2014), o que não é possível com as parametrizações propostas até agora. Santos-Neto et al. (2014) conduziram um estudo de simulação para avaliar propriedades dos estimadores dos parâmetros do modelo, em que obtiveram boas propriedades. Nesse ensaio, os autores concluíram que os estimadores de máxima verossimilhança e de momentos modificados apresentam propriedades estatísticas semelhantes e melhores às dos outros estimadores considerados: método de momentos e método de momentos generalizado. A seguir, apresentamos brevemente, como foram obtidos estimadores através do método de máxima verossimilhança e momentos para o modelo RBS, como descrito em Santos-Neto et al. (2014).

### 2.3.1 Estimação clássica para a distribuição RBS

A seguir, apresentamos as estimativas de máxima verossimilhança, de momentos e algumas inferências para a reparametrização considerada, que auxiliarão no tratamento do modelo do Capítulo 3. Denotaremos o vetor de parâmetros por  $\theta = (\varphi, \kappa)^\top$ .

## Estimação por máxima verossimilhança

Seja  $\mathbf{Y} = [Y_1, Y_2, \dots, Y_n]^\top$  uma amostra aleatória de tamanho  $n$  de  $Y \sim \text{RBS}(\varphi, \kappa)$ . Então, a função log-verossimilhança para  $\boldsymbol{\theta}$  é

$$\ell(\boldsymbol{\theta}) = \sum_{j=1}^n \ell_j(\boldsymbol{\theta}), \quad (2.12)$$

em que  $\ell_j(\boldsymbol{\theta})$  é o logaritmo da PDF dada em (2.3) substituindo  $y$  por  $y_j$ . Para obter o EMV para  $\boldsymbol{\theta}$ , igualamos as funções escore a zero, em que neste caso o vetor escore para  $\boldsymbol{\theta}$  é dado por  $U(\boldsymbol{\theta}) = [U_\varphi, Y_\kappa]^\top$ , em que

$$U_\varphi = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \varphi} = \sum_{j=1}^n \left[ \frac{\kappa}{\kappa y_j + y_j + \varphi \kappa} + \frac{y_j(\kappa + 1)}{4\varphi^2} - \frac{\kappa^2}{4y_j(\kappa + 1)} - \frac{1}{2\varphi} \right]. \quad (2.13)$$

e

$$U_\kappa = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \kappa} = \sum_{j=1}^n \left[ \frac{y_j + \varphi}{\kappa y_j + y_j + \varphi \kappa} - \frac{y_j}{4\varphi} - \frac{\kappa \kappa + 2\varphi}{4y_j(\kappa + 1)^2} + \frac{\kappa}{2(\kappa + 1)} \right]. \quad (2.14)$$

Para maximizar a função  $\ell(\boldsymbol{\theta})$ , dada em (2.12), devemos utilizar um método numérico iterativo como, por exemplo, de Newton-Raphson, visto que para o modelo RBS não é possível obter estimador de máxima verossimilhança de forma fechada, assim como no caso da BS original. No entanto, Santos-Neto et al. (2014) lembram que  $\hat{\boldsymbol{\theta}} = (\hat{\varphi}, \hat{\kappa})^\top$  é o EMV para  $\boldsymbol{\theta}$  e isto segue aproximadamente a distribuição

$$N\left(\hat{\boldsymbol{\theta}}, E(\hat{\boldsymbol{\theta}})^{-1}\right), \quad (2.15)$$

na qual  $E(\hat{\boldsymbol{\theta}}) = [E_{\theta_j \theta_k}]$  corresponde a informação esperada de Fisher.

## Estimação pelo método dos momentos

Consideremos, novamente,  $\mathbf{Y} = [Y_1, Y_2, \dots, Y_n]^\top$  uma amostra aleatória de tamanho  $n$  de  $Y \sim \text{RBS}(\varphi, \kappa)$ . Segundo Santos-Neto et al. (2014),  $\tilde{\varphi}$  e  $\tilde{\kappa}$  são os estimadores de pelo método dos momentos de  $\varphi$  e  $\kappa$ , dados por

$$\tilde{\varphi} = \bar{Y} \quad \text{e} \quad \tilde{\kappa} = \frac{\bar{Y}^2 - S^2 + \sqrt{(\bar{Y}^4 + 3\bar{Y}^2 S^2)}}{S^2}, \quad (2.16)$$

em que  $\bar{Y} = \frac{1}{n} \sum_{j=1}^n Y_j$  e  $S^2 = \frac{1}{n} \sum_{j=1}^n [Y_j - \bar{Y}]^2$ . Ainda segundo esses autores,  $\tilde{\varphi}$  e  $\tilde{\kappa}$  têm uma distribuição conjunta assintótica normal bivariada normal, isto é,

$$\begin{pmatrix} \tilde{\varphi} \\ \tilde{\kappa} \end{pmatrix} \sim AN \left( \begin{pmatrix} \varphi \\ \kappa \end{pmatrix}, \begin{pmatrix} \frac{\varphi^2(2\kappa+5)}{(\kappa+1)^2} & -\frac{\varphi(2\kappa^2+8\kappa-3)}{(\kappa+1)(\kappa+4)} \\ -\frac{\varphi(2\kappa^2+8\kappa-3)}{(\kappa+1)(\kappa+4)} & \frac{2\kappa^4+28\kappa^3+122\kappa^2+126\kappa+57}{(\kappa+4)^2} \end{pmatrix} \right).$$

# Capítulo 3

## O modelo GP-RBS

Um dos grandes motivos em usar o GP para a modelagem em questão, além da sua extensa aplicabilidade, se dá pelo fato de que a inferência estatística pode ser facilmente realizada, como discutido na Subseção 1.1. Esse importante aspecto, somado ao fato da distribuição RBS possuir a propriedade de multiplicação por um escalar (ver Santos-Neto et al. (2012) e Santos-Neto et al. (2014)), nos proporciona a seguinte PDF para o processo geométrico com a distribuição Birnbaum–Saunders reparametrizada (GP-RBS)

$$f_X(x; \varpi, \varphi, \kappa) = \frac{\exp(\kappa/2)\sqrt{\kappa+1}\sqrt{\varpi^{i-1}}}{4x^{3/2}\sqrt{\pi\varphi}} \left[ x + \frac{\kappa\varphi}{(\kappa+1)\varpi^{i-1}} \right] \times \exp\left(-\frac{\kappa}{4} \left[ \frac{x(\kappa+1)\varpi^{i-1}}{\kappa\varphi} + \frac{\kappa\varphi\varpi^{i-1}}{x(\kappa+1)} \right]\right), \quad (3.1)$$

para  $x > 0$  e  $\varpi, \varphi, \kappa > 0$ . Com isso, percebe-se que a distribuição GP-RBS pode ser vista como a RBS( $\frac{\varphi}{\varpi^{i-1}}, \kappa$ ), com  $\varpi$  sendo o parâmetro razão do GP. Logo, números aleatórios da distribuição do modelo acima podem ser obtidos baseando-se no gerador descrito no Algoritmo 1 indicado na Subseção 2.3.

Assumindo agora que  $\{X_1, X_2, X_3, \dots, X_n\}$  é um conjunto de dados proveniente de um GP com parâmetro razão  $\varpi > 0$  e  $X_1 \sim \text{RBS}(\varphi, \kappa)$ , então, pelas propriedades da distribuição

RBS e pela definição do GP (2.4), a função de verossimilhança para  $X_i, i = 1, 2, \dots, n$  é

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \frac{\exp(\kappa/2) \sqrt{\kappa+1} \sqrt{\varpi^{i-1}}}{4x_i^{3/2} \sqrt{\pi\varphi}} \left[ x_i + \frac{\kappa\varphi}{(\kappa+1)\varpi^{i-1}} \right] \\ \times \exp \left( -\frac{\kappa}{4} \left[ \frac{x_i(\kappa+1)\varpi^{i-1}}{\kappa\varphi} + \frac{\kappa\varphi\varpi^{i-1}}{x_i(\kappa+1)} \right] \right), \quad (3.2)$$

em que  $\boldsymbol{\theta} = (\varpi, \varphi, \kappa)^\top$  é o vetor de parâmetros. Consequentemente, temos a função de log-verossimilhança dada por

$$\ell(\boldsymbol{\theta}) = \frac{n}{2} \left\{ \kappa + \log(\kappa+1) - 4\log(2) - \log(\pi) - \log(\varphi) + \left( \frac{n-1}{2} \right) \log(\varpi) \right\} \\ - \sum_{i=1}^n \left[ \frac{3}{2} \log(x_i) - \log \left[ x_i + \frac{\kappa\varphi}{\varpi^{i-1}(\kappa+1)} \right] \right] \\ - \frac{\kappa}{4} \sum_{i=1}^n \left[ \frac{x_i(\kappa+1)\varpi^{i-1}}{\kappa\varphi} + \frac{\kappa\varphi}{x_i(\kappa+1)\varpi^{i-1}} \right]. \quad (3.3)$$

Então, as primeiras derivadas da função log-verossimilhança dadas na equação (3.3), em relação aos três parâmetros, são obtidas para estabelecer o vetor escore  $U(\boldsymbol{\theta}) = \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ .

Os elementos do vetor escore são expressados como

$$\frac{\partial\ell(\boldsymbol{\theta})}{\partial\varpi} = \frac{n(n-1)}{4} \left( \frac{1}{\varpi} \right) - \sum_{i=1}^n \frac{(i-1)\kappa\varphi}{\kappa\varphi\varpi + x_i(\kappa+1)\varpi^i} \\ - \frac{\kappa}{4} \sum_{i=1}^n (i-1) \left[ \frac{x_i(\kappa+1)\varpi^{i-2}}{\kappa\varphi} - \frac{\kappa\varphi}{x_i(\kappa+1)\varpi^i} \right] = 0, \\ \frac{\partial\ell(\boldsymbol{\theta})}{\partial\kappa} = \frac{n(\kappa+2)}{2(\kappa+1)} + \sum_{i=1}^n \frac{\varphi\varpi}{(\kappa+1)[\kappa\varphi\varpi + x_i(\kappa+1)\varpi^i]} \\ + \frac{\kappa}{4} \sum_{i=1}^n \frac{[(\varphi\kappa\varpi)^2(\kappa+2) - (x_i(\kappa+1)\varpi^i)^2]}{x_i\varphi\kappa\varpi^{i+1}[(\kappa+1)]^2} = 0, \quad (3.4) \\ \frac{\partial\ell(\boldsymbol{\theta})}{\partial\varphi} = -\frac{n}{2} \frac{1}{\varphi} - \frac{\kappa}{4} \sum_{i=1}^n \left( \frac{\kappa}{x_i\varpi^{i-1}(\kappa+1)} - \frac{x_i(\kappa+1)\varpi^{i-1}}{\kappa\varphi^2} \right) \\ + \sum_{i=1}^n \frac{\kappa\varpi}{\kappa\varphi\varpi + x_i(\kappa+1)\varpi^i} = 0.$$

Com base nisso, nas próximas seções iremos apresentar, para este modelo, como obtemos os estimadores do ponto de vista clássico via EMV e MM.

### 3.1 Estimador de máxima verosimilhança

Não é possível obter analiticamente os EMV para  $\theta$ , pelas equações em (3.4). No entanto, tais equações podem ser resolvidas simultaneamente usando um método numérico como o de Newton-Raphson para estimar esses parâmetros; ver por exemplo Gentle (2009) e Monahan (2011), da forma

$$\theta_{n+1} = \theta_n - H^{-1}(\theta_n)\nabla(\theta_n), \quad (3.5)$$

em que  $\nabla(\theta) = (\frac{\partial \ell(\theta)}{\partial \varpi}, \frac{\partial \ell(\theta)}{\partial \kappa}, \frac{\partial \ell(\theta)}{\partial \varphi})^\top$  é o vetor gradiente para  $\theta$ ,  $n$  é o número de iterações, e  $H^{-1}(\theta)$  é a matriz Hessiana inversa de  $H(\theta)$ , com  $H(\theta) \in \mathbb{R}^{3 \times 3}$ . Temos que os elementos da matriz  $H(\theta)$ , denotados por  $h_{ij}$ , ( $i, j = 1, 2, 3$ ), são as segundas derivadas de (3.4), exibidas no Apêndice A, com relação aos parâmetros  $\varpi$ ,  $\kappa$  e  $\varphi$ , de modo que

$$H^{-1}(\theta) = \frac{1}{\text{Det}(H(\theta))} \begin{bmatrix} h_{22}h_{33} - h_{23}h_{32} & -h_{12}h_{33} - h_{13}h_{32} & h_{12}h_{23} - h_{13}h_{22} \\ -h_{21}h_{33} - h_{31}h_{23} & h_{11}h_{33} - h_{13}h_{31} & -h_{11}h_{23} - h_{21}h_{13} \\ h_{21}h_{32} - h_{22}h_{31} & -h_{11}h_{32} - h_{12}h_{31} & h_{11}h_{22} - h_{12}h_{21} \end{bmatrix},$$

com  $\text{Det}(H(\theta)) = h_{11}h_{22}h_{33} - h_{11}h_{23}h_{32} - h_{12}h_{21}h_{33} + h_{12}h_{31}h_{23} + h_{21}h_{13}h_{32} + h_{13}h_{22}h_{31}$ . Os elementos  $h_{ij}$ , ( $i, j = 1, 2, 3$ ) da matriz  $H(\theta)$  são apresentados no Apêndice A. Dessa forma, os EMV's para  $\varpi$ ,  $\kappa$  e  $\varphi$ , denotados por  $\hat{\varpi}$ ,  $\hat{\kappa}$  e  $\hat{\varphi}$  são obtidos como os respectivos elementos de  $\theta_n$ , ao tomarmos  $\theta_0$  como estimativa inicial, via (3.5). Ademais, segundo Cox e Barndorff-Nielsen (1994), a distribuição conjunta desses estimadores é assintoticamente normal, com vetor de médias  $(\varpi, \kappa, \varphi)$  e matriz de covariâncias  $I^{-1}(\varpi, \kappa, \varphi)$ , em que  $I$  é a matriz informação de Fisher, descrita abaixo:

$$I = - \begin{bmatrix} E(\frac{\partial^2 \ln L(\theta)}{\partial \varpi^2}) & E(\frac{\partial^2 \ln L(\theta)}{\partial \varpi \partial \kappa}) & E(\frac{\partial^2 \ln L(\theta)}{\partial \varpi \partial \varphi}) \\ E(\frac{\partial^2 \ln L(\theta)}{\partial \kappa \partial \varpi}) & E(\frac{\partial^2 \ln L(\theta)}{\partial \kappa^2}) & E(\frac{\partial^2 \ln L(\theta)}{\partial \kappa \partial \varphi}) \\ E(\frac{\partial^2 \ln L(\theta)}{\partial \varphi \partial \varpi}) & E(\frac{\partial^2 \ln L(\theta)}{\partial \varphi \partial \kappa}) & E(\frac{\partial^2 \ln L(\theta)}{\partial \varphi^2}) \end{bmatrix}.$$

Como vemos no Apêndice A, não é possível obter uma forma explícita da matriz de informação de Fisher  $I$ , neste caso. Mas, os elementos de  $I$  são imediatamente escritos a partir de elementos da matriz hessiana  $H(\theta)$ . Embora uma forma explícita da matriz  $I$  não possa ser

encontrada, como estimador, então a matriz de informação observada pode ser usada em vez de matriz  $I$ , como ressalta Cox e Barndorff-Nielsen (1994).

## 3.2 Estimação por momentos modificados

Como visto na Definição 2.4, o GP é um processo estocástico monótono. Assim, alguns problemas de divergência podem surgir na estimativa do parâmetro razão  $\varpi$  (Bicer (2018)). No entanto, é proposto um método não paramétrico amplamente utilizado na inferência estatística, para superar tal problema e estimar o parâmetro  $\varpi$  no GP. Esse estimador não paramétrico é representado por  $\hat{\varpi}_{NP}$ ; (ver Lam e Chan (1998) e Lam et al. (2004)) e é dado por:

$$\hat{\varpi}_{NP} = \exp \left( \frac{6}{(n-1)n(n+1)} \sum_{i=1}^n (n-2i+1) \ln X_i \right), \quad (3.6)$$

o qual é não-viesado e segue uma distribuição assintoticamente normal, segundo Lam et al. (2004). Além disso, podemos representar o processo apresentado na Definição 2.4 ao substituir  $\varpi^{(i-1)}$  por (3.6), e encontrar o estimador por momentos modificados,

$$\hat{Y}_i = \hat{\varpi}_{NP}^{(i-1)} X_i, i = 1, 2, \dots, n, \quad (3.7)$$

que possibilitará estimar os parâmetros  $\kappa$  e  $\varphi$  por outros métodos mais adiante.

## 3.3 Estimação pelo método dos momentos

Consideremos, agora,  $X_1, X_2, \dots, X_n$  uma amostra aleatória de um GP com razão  $\varpi > 0$ ,  $X_1 \sim \text{RBS}(\varphi, \kappa)$  e o estimador obtido em (3.6). Seja, ainda,  $m_1$  a média amostral e  $S^2$  a variância amostral para a amostra  $X_1, X_2, \dots, X_n$ , tais que

$$m_1 = \frac{1}{n} \sum_{i=1}^n \hat{\varpi}_{NP}^{(i-1)} X_i, i = 1, 2, \dots, n \quad (3.8)$$

e

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (\hat{Y}_i - m_1)^2, i = 1, 2, \dots, n. \quad (3.9)$$

com isso, pelas propriedades da distribuição RBS, pondo  $\phi = (\kappa + 1)^2 / (2\kappa + 5)$ , temos que

$$E[X] = \phi \quad \text{e} \quad E[X^2] = \frac{\phi^2}{\phi} + \phi^2.$$

Dessa forma, os estimadores obtidos via método dos momentos, para os parâmetros  $\phi$  ( $\hat{\phi}_{MM}$ ) e  $\kappa$  ( $\hat{\kappa}_{MM}$ ), podem ser obtidos pela solução do seguinte sistema não-linear:

$$\phi - m_1 = 0 \quad \text{e} \quad \phi^2(2\kappa + 5) / (\kappa + 1)^2 + \left[ \frac{(\kappa + 1)^2}{(2\kappa + 5)} \right]^2 - m_2 = 0. \quad (3.10)$$

Através das equações em (2.16) temos que  $\tilde{\phi}_{MM}$  e  $\tilde{\kappa}_{MM}$  são dados por

$$\tilde{\phi}_{MM} = m_1 \quad \text{e} \quad \tilde{\kappa}_{MM} = \frac{m_1^2 - S^2 + \sqrt{(m_1^4 + 3m_1^2 S^2)}}{S^2}, \quad (3.11)$$

são os estimadores obtidos via método dos momentos, em função do estimador obtido em (3.2).

## Capítulo 4

# Aplicação numérica, resultados e discussões

Neste capítulo, apresentaremos um estudo de simulação de Monte Carlo para avaliar a performance dos estimadores dos parâmetros do modelo, em que nesse caso consideramos o estimador de máxima verossimilhança e o obtido via método dos momentos. Em seguida ilustramos o modelo proposto através da aplicação em quatro conjuntos de dados reais, a saber: `aircraft`, `coal mining`, `carnero` e `jiwani`, os quais podem ser vistos na Seção 4.2. Esses dados foram analisados em outros trabalhos, como por exemplo, Proschan (1963), Andrews e Herzberg (1985), Linhart e Zucchini (1986) e Bakouch et al. (2019). O primeiro conjunto de dados pode ser encontrado no pacote `boot` do software R, ver R-Team (2019).

Comparamos o modelo GP-RBS com os modelos Power Lindley, log-lornal, gama e gaussiana inversa, propostos por Bicer (2018), Lam e Chan (1998), Chan et al. (2014) e Kara et al. (2015), respectivamente. Como medida de avaliação de “bondade de ajuste” usamos o critério de informação de Akaike (AIC), critério Bayesiano de Schwarz (BIC) e o fator de Bayes (BF), propostos em Akaike (1974), Schwarz et al. (1978) e Kass e Raftery (1995), respectivamente. Nós usamos o BF para avaliar a magnitude da diferença entre dois valores de BIC; veja Kass e Raftery (1995), Leiva et al. (2015) e Ferreira et al. (2012).

## 4.1 Estudo de simulação de Monte Carlo

Conforme a literatura vigente e aos modelos que iremos comparar, consideramos os seguintes cenários para as simulações: fixamos o parâmetro de forma,  $\kappa$ , no valor de 0.5, enquanto que o parâmetro de escala,  $\varphi$ , variou no conjunto  $\{1.0, 1.5, 2.0\}$ . Enquanto isso, para o parâmetro de razão,  $\varpi$ , definimos sua variação no conjunto  $\{0.90, 0.95, 1.05, 1.10\}$ . Consideramos diferentes tamanhos amostrais, isto é  $n = \{30, 50, 100, 150\}$ , baseados em 3000 replicas de Monte Carlo. Para cada parâmetro e tamanho amostral distinto, obtivemos medidas descritivas tais como: a média, variância, viés e erro quadrático médio (EQM). Esses resultados estão organizados nas Tabelas 4.1, 4.2 e 4.3.

Através dos resultados apresentados nas Tabela 4.1 a 4.3, temos que os valores do viés e EQM diminuem à medida que  $n$  aumenta para todos os estimadores em cada um dos métodos de estimação avaliados. Isso ocorre de acordo com o esperado, indicando que os estimadores dos parâmetros do modelo são assintoticamente consistentes. Em relação ao EMM e EMV dos parâmetros  $\varpi$  e  $\varphi$ , observamos, em geral, que o EMV apresenta valores do EQM quando comparado os observados a partir do EMM. Nessa mesma linha, notamos para o parâmetro  $\kappa$  que o EQM dos EMV's são, claramente, menores em comparação aos obtidos via método dos momentos.

Tabela 4.1: Medidas descritivas para os estimadores EMV e EMM dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$ , quando  $\varphi = 1.0$ .

$\varpi$	$n$	Estimador	$\varpi$			$\varphi$			$\kappa$		
			Média	Viés	EQM	Média	Viés	EQM	Média	Viés	EQM
0.90	30	EMV	0.9002	0.0002	0.0007	1.0651	0.0651	0.3356	0.5890	0.0890	0.0385
		EMM	0.9004	0.0004	0.0009	1.1044	0.1044	0.4572	0.7530	0.2530	0.2724
	50	EMV	0.9001	0.0002	0.0001	1.0385	0.0385	0.1824	0.5509	0.0509	0.0171
		EMM	0.9002	0.0002	0.0002	1.0635	0.0635	0.2496	0.6535	0.1535	0.1505
	100	EMV	0.9000	<0.0001	<0.0001	1.0188	0.0188	0.0842	0.5238	0.0238	0.0066
		EMM	0.9001	0.0001	<0.0001	1.0347	0.0347	0.1176	0.5841	0.0841	0.0742
	150	EMV	0.9000	<0.0001	<0.0001	1.0123	0.0123	0.0523	0.5154	0.0154	0.0038
		EMM	0.9001	0.0001	<0.0001	1.0255	0.0255	0.0731	0.5623	0.0623	0.0489
0.95	30	EMV	0.9508	0.0008	0.0008	1.0693	0.0693	0.3494	0.5872	0.0872	0.0381
		EMM	0.9513	0.0013	0.0010	1.1111	0.1111	0.4890	0.7432	0.2432	0.2634
	50	EMV	0.9502	0.0002	0.0002	1.0397	0.0397	0.1750	0.5525	0.0525	0.0166
		EMM	0.9502	0.0002	0.0002	1.0630	0.0630	0.2515	0.6552	0.1552	0.1494
	100	EMV	0.9499	-0.0001	<0.0001	1.0082	0.0082	0.0814	0.5259	0.0259	0.0067
		EMM	0.9499	-0.0001	<0.0001	1.0212	0.0212	0.1072	0.5824	0.0824	0.0730
	150	EMV	0.9500	<0.0001	<0.0001	1.0157	0.0157	0.0563	0.5146	0.0146	0.0037
		EMM	0.9500	<0.0001	<0.0001	1.0270	0.0270	0.0750	0.5525	0.0525	0.0481
1.05	30	EMV	1.0494	-0.0006	0.0009	1.0536	0.0536	0.3418	0.5836	0.0836	0.0349
		EMM	1.0494	-0.0006	0.0012	1.0827	0.0827	0.4266	0.7447	0.2447	0.2748
	50	EMV	1.0504	0.0004	0.0002	1.0388	0.0388	0.1718	0.5462	0.0462	0.0158
		EMM	1.0504	0.0004	0.0003	1.0621	0.0621	0.2409	0.6536	0.1536	0.1434
	100	EMV	1.0500	<0.0001	<0.0001	1.0206	0.0206	0.0826	0.5246	0.0246	0.0064
		EMM	1.0500	<0.0001	<0.0001	1.0303	0.0303	0.1114	0.5857	0.0857	0.0726
	150	EMV	1.0500	<0.0001	<0.0001	1.0146	0.0146	0.0451	0.5132	0.0132	0.0041
		EMM	1.0499	-0.0001	<0.0001	1.0152	0.0152	0.0696	0.5482	0.0482	0.0476
1.10	30	EMV	1.0998	-0.0002	0.0011	1.0574	0.0574	0.3362	0.5878	0.0878	0.0388
		EMM	1.0997	-0.0003	0.0014	1.0852	0.0852	0.4436	0.7570	0.2570	0.2794
	50	EMV	1.1001	0.0001	0.0002	1.0290	0.0290	0.1813	0.5539	0.0539	0.0177
		EMM	1.1003	0.0003	0.0003	1.0554	0.0554	0.2563	0.6593	0.1593	0.1560
	100	EMV	1.0999	-0.0001	<0.0001	1.0118	0.0118	0.0655	0.5202	0.0202	0.0060
		EMM	1.0999	-0.0001	<0.0001	1.0293	0.0293	0.1162	0.5758	0.0758	0.0721
	150	EMV	1.0994	-0.0006	<0.0001	1.0053	0.0053	0.0002	0.4621	-0.0379	0.0020
		EMM	1.1001	0.0001	<0.0001	1.0270	0.0270	0.0765	0.5473	0.0473	0.0465

Tabela 4.2: Medidas descritivas para os estimadores EMV e EMM dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$ , quando  $\varphi = 1.5$ .

$\varpi$	$n$	Estimador	$\varpi$			$\varphi$			$\kappa$		
			Média	Viés	EQM	Média	Viés	EQM	Média	Viés	EQM
0.90	30	EMV	0.9009	0.0009	0.0007	1.5984	0.0984	0.7177	0.5890	0.0890	0.0369
		EMM	0.9016	0.0016	0.0009	1.6682	0.1682	0.0009	0.7604	0.2604	0.2828
	50	EMV	0.9005	0.0005	0.0001	1.5684	0.0684	0.4205	0.5524	0.0524	0.0165
		EMM	0.9004	0.0004	0.0002	1.5925	0.0925	0.0002	0.6541	0.1541	0.1530
	100	EMV	0.9000	<0.0001	<0.0001	1.5278	0.0278	0.1962	0.5250	0.0250	0.0069
		EMM	0.8999	-0.0001	<0.0001	1.5381	0.0381	<0.0001	0.5871	0.0871	0.0739
	150	EMV	0.9000	<0.0001	<0.0001	1.5201	0.0201	0.1208	0.5140	0.0140	0.0040
		EMM	0.8999	-0.0001	<0.0001	1.5240	0.0240	<0.0001	0.5591	0.0591	0.0518
0.95	30	EMV	0.9497	-0.0003	0.0008	1.5880	0.0880	0.7478	0.5820	0.0820	0.0346
		EMM	0.9502	0.0002	0.0010	1.6500	0.1500	1.0265	0.7483	0.2483	0.2766
	50	EMV	0.9502	0.0002	0.0002	1.5777	0.0777	0.4550	0.5470	0.0470	0.0158
		EMM	0.9501	0.0001	0.0002	1.6048	0.1048	0.5976	0.6527	0.1527	0.1489
	100	EMV	0.9500	<0.0001	<0.0001	1.5327	0.0327	0.1870	0.5217	0.0217	0.0063
		EMM	0.9500	<0.0001	<0.0001	1.5450	0.0450	0.2499	0.5792	0.0792	0.0717
	150	EMV	0.9500	<0.0001	<0.0001	1.5125	0.0125	0.1237	0.5157	0.0157	0.0038
		EMM	0.9500	<0.0001	<0.0001	1.5297	0.0297	0.1751	0.5547	0.0547	0.0486
1.05	30	EMV	1.0504	0.0004	0.0010	1.5955	0.0955	0.7502	0.5918	0.0918	0.0405
		EMM	1.0503	0.0003	0.0013	1.6411	0.1411	1.0633	0.7602	0.2602	0.2749
	50	EMV	1.0499	-0.0001	0.0002	1.5394	0.0394	0.3914	0.5529	0.0529	0.0170
		EMM	1.0501	0.0001	0.0003	1.5836	0.0836	0.5865	0.6534	0.1534	0.1560
	100	EMV	1.0501	0.0001	<0.0001	1.5387	0.0387	0.1985	0.5278	0.0278	0.0069
		EMM	1.0501	0.0001	<0.0001	1.5504	0.0504	0.2586	0.5777	0.0777	0.0716
	150	EMV	1.0501	0.0001	<0.0001	1.5277	0.0277	0.1127	0.5142	0.0142	0.0039
		EMM	1.0500	<0.0001	<0.0001	1.5342	0.0342	0.1595	0.5530	0.0530	0.0498
1.10	30	EMV	1.1000	<0.0001	0.0011	1.5976	0.0976	0.7896	0.5932	0.0932	0.0396
		EMM	1.1002	0.0002	0.0014	1.6497	0.1497	1.1159	0.7451	0.2451	0.2718
	50	EMV	1.1000	<0.0001	0.0002	1.5448	0.0448	0.3925	0.5512	0.0512	0.0173
		EMM	1.1003	0.0003	0.0003	1.5914	0.0914	0.5631	0.6546	0.1546	0.1505
	100	EMV	1.1000	<0.0001	<0.0001	1.5289	0.0289	0.1937	0.5244	0.0244	0.0063
		EMM	1.1000	<0.0001	<0.0001	1.5512	0.0512	0.2653	0.5700	0.0700	0.0748
	150	EMV	1.1000	<0.0001	<0.0001	1.5238	0.0238	0.1220	0.5150	0.0150	0.0040
		EMM	1.1000	<0.0001	<0.0001	1.5364	0.0364	0.1654	0.5501	0.0501	0.0482

Tabela 4.3: Medidas descritivas para os estimadores EMV e EMM dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$ , quando  $\varphi = 2.0$ .

$\varpi$	$n$	Estimador	$\varpi$			$\varphi$			$\kappa$		
			Média	Viés	EQM	Média	Viés	EQM	Média	Viés	EQM
0.90	30	EMV	0.9002	0.0002	0.0007	2.1250	0.1250	1.3215	0.5865	0.0865	0.0375
		EMM	0.9006	0.0006	0.0009	2.2165	0.2165	2.0266	0.7610	0.2610	0.2841
	50	EMV	0.9002	0.0002	0.0001	2.0616	0.0616	0.6671	0.5499	0.0499	0.0163
		EMM	0.9002	0.0002	0.0002	2.1114	0.1114	0.9547	0.6564	0.1564	0.1543
	100	EMV	0.8998	-0.0002	<0.0001	2.0228	0.0228	0.3371	0.5260	0.0260	0.0072
		EMM	0.8999	-0.0001	<0.0001	2.0541	0.0541	0.4733	0.5855	0.0855	0.0719
	150	EMV	0.9000	<0.0001	<0.0001	2.0245	0.0245	0.2233	0.5136	0.0136	0.0038
		EMM	0.9000	<0.0001	<0.0001	2.0405	0.0405	0.3030	0.5558	0.0558	0.0475
0.95	30	EMV	0.9492	-0.0008	0.0008	2.0977	0.0977	1.4648	0.5916	0.0916	0.0393
		EMM	0.9490	-0.0010	0.0010	2.1626	0.1626	1.9888	0.7522	0.2522	0.2759
	50	EMV	0.9501	0.0001	0.0002	2.0777	0.0777	0.7076	0.5496	0.0496	0.0169
		EMM	0.9500	<0.0001	0.0002	2.1226	0.1226	1.0118	0.6569	0.1569	0.1554
	100	EMV	0.9500	<0.0001	<0.0001	2.0511	0.0511	0.3506	0.5221	0.0221	0.0062
		EMM	0.9500	<0.0001	<0.0001	2.0706	0.0706	0.4726	0.5729	0.0729	0.0705
	150	EMV	0.9500	<0.0001	<0.0001	2.0283	0.0283	0.2114	0.5144	0.0144	0.0039
		EMM	0.9500	<0.0001	<0.0001	2.0409	0.0409	0.2927	0.5569	0.0569	0.0484
1.05	30	EMV	1.0501	0.0001	0.0010	2.0944	0.0944	1.2392	0.5909	0.0909	0.0387
		EMM	1.0497	-0.0003	0.0012	2.1497	0.1497	1.7376	0.7424	0.2424	0.2790
	50	EMV	1.0499	-0.0001	0.0002	2.0492	0.0492	0.7040	0.5523	0.0523	0.0165
		EMM	1.0500	-0.0003	0.0003	2.0963	0.0963	0.9469	0.6534	0.1534	0.1450
	100	EMV	1.0500	<0.0001	<0.0001	2.0447	0.0492	0.3401	0.5237	0.0237	0.0069
		EMM	1.0501	0.0001	<0.0001	2.0736	0.0736	0.4540	0.5798	0.0798	0.0690
	150	EMV	1.0500	<0.0001	<0.0001	2.0096	0.0096	0.1987	0.5171	0.0171	0.0039
		EMM	1.0500	<0.0001	<0.0001	2.0345	0.0345	0.2905	0.5510	0.0510	0.0500
1.10	30	EMV	1.1012	0.0012	0.0010	2.1572	0.1572	1.3554	0.5948	0.0948	0.0403
		EMM	1.1018	0.0018	0.0013	2.2216	0.2216	1.7619	0.7607	0.2607	0.2768
	50	EMV	1.0999	-0.0001	0.0002	2.0599	0.0599	0.7419	0.5525	0.0525	0.0168
		EMM	1.0996	-0.0004	0.0003	2.0880	0.0880	0.9632	0.6596	0.1596	0.1542
	100	EMV	1.0996	-0.0004	<0.0001	2.0023	0.0023	0.3059	0.5236	0.0236	0.0065
		EMM	1.0996	-0.0004	<0.0001	2.0246	0.0246	0.4228	0.5767	0.0767	0.0754
	150	EMV	1.1000	<0.0001	<0.0001	2.0211	0.0211	0.2227	0.5157	0.0157	0.0039
		EMM	1.1000	<0.0001	<0.0001	2.0366	0.0366	0.3005	0.5500	0.0500	0.0501

Por outro lado, para o EMV do parâmetro de forma  $\kappa$ , na Tabela 4.1, o seu EQM, tanto para  $n$  pequeno quanto para o maior  $n$  considerado, é bem próximo nestes tamanhos amostrais, ou seja, a convergência assintótica do seu estimador para o seu verdadeiro valor paramétrico parece não ser afetada. Para este mesmo parâmetro, o estimador obtido via MM tem um viés maior, mesmo quando  $n$  é relativamente grande ou moderado. Esse comportamento é observado nos demais cenários considerados, que estão dispostos nas Tabelas 4.2 e 4.3. Assim, podemos afirmar que os estimadores obtidos via método dos momentos tendem a convergir para o verdadeiro valor do parâmetro, contudo a sua respectiva convergência é mais lenta quando comparada ao método de máxima verossimilhança.

Nesta perspectiva, foram obtidas as distribuições empíricas dos estimadores dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$ , que apontam ilustrativamente para esta direção; ver Figuras 4.1 a 4.12. As figuras também apresentam os intervalos de confiança (IC), de 90%, 95% e 99%. Vale ressaltar que, em cada figura, estão na esquerda a estimação via MV; na direita, a estimação via MM.

De acordo com as Figuras 4.1 a 4.12 há uma indicação de que a distribuição empírica dos estimadores de momentos e máxima verossimilhança, respectivamente, parecem ser simétricas em torno do verdadeiro valor do parâmetro, à medida que  $n$  cresce, como esperado. Mas a variabilidade do estimador obtido através dos momentos é um pouco maior do que os MV, quando consideramos os parâmetros  $\varpi$  e  $\varphi$ .

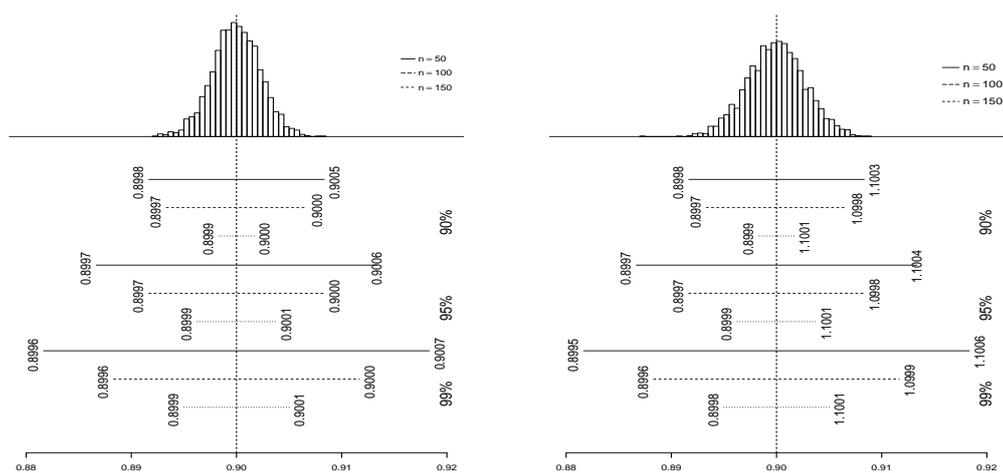


Figura 4.1: Distribuição empírica de  $\hat{\varpi}$ , para  $\varpi = 0.90$ , com  $\varphi = 2.00$  e  $IC(\varpi; [1-\alpha] \times 100\%)$  para cada  $n$ .

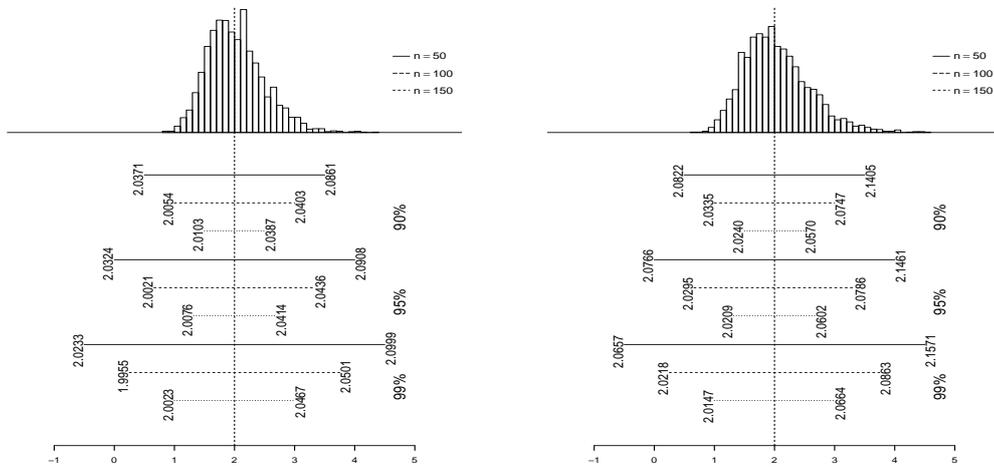


Figura 4.2: Distribuição empírica de  $\hat{\varphi}$ , para  $\varphi = 2.00$ , com  $\varpi = 0.90$  e  $IC(\varphi; [1-\alpha] \times 100\%)$  para cada  $n$ .

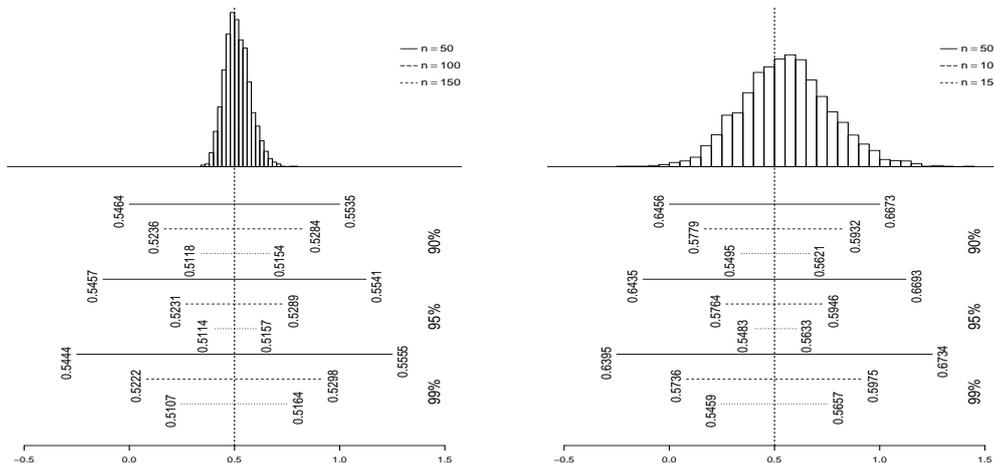


Figura 4.3: Distribuição empírica de  $\hat{\kappa}$ , para  $\kappa = 0.50$ , em que  $\varpi = 0.90$  e  $\varphi = 2.00$ , com  $IC(\kappa; [1-\alpha] \times 100\%)$  para cada  $n$ .

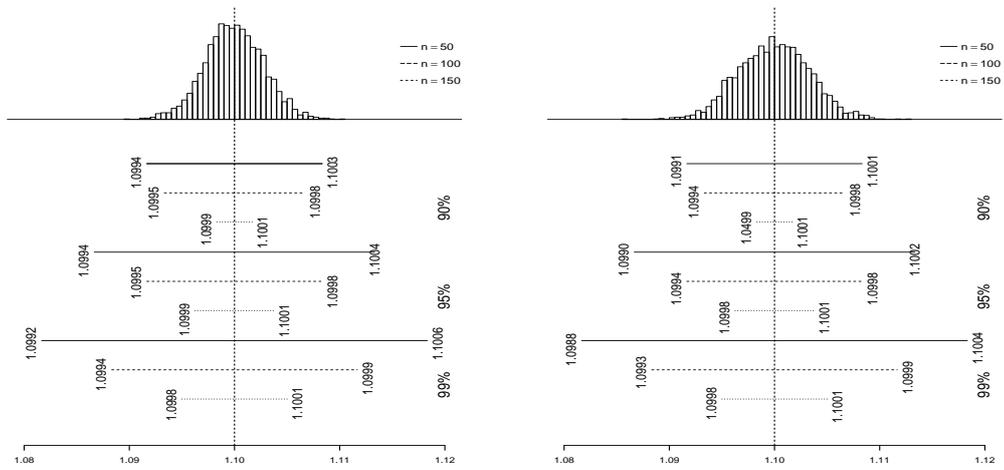


Figura 4.4: Distribuição empírica de  $\hat{\omega}$ , para  $\omega = 1.10$ , com  $\phi = 2.00$  e  $IC(\omega; [1-\alpha] \times 100\%)$  para cada  $n$ .

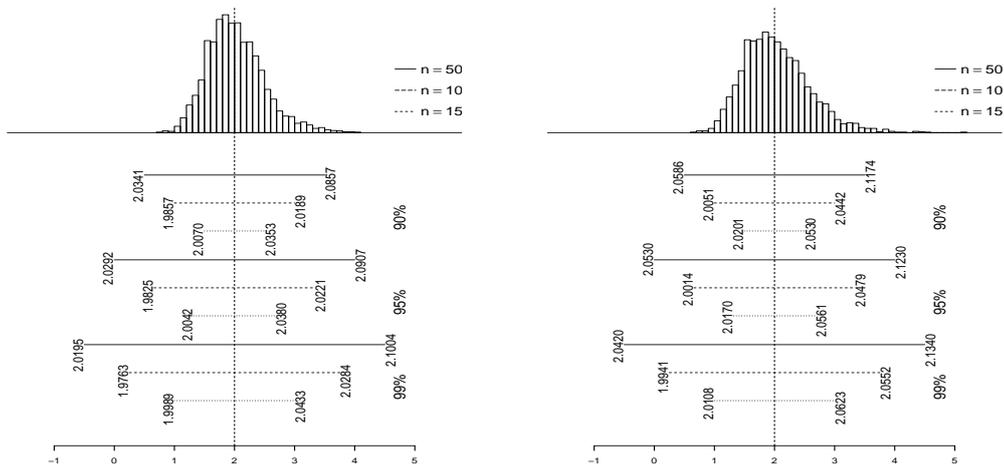


Figura 4.5: Distribuição empírica de  $\hat{\phi}$ , para  $\phi = 2.00$ , com  $\omega = 1.10$  e  $IC(\phi; [1-\alpha] \times 100\%)$  para cada  $n$ .

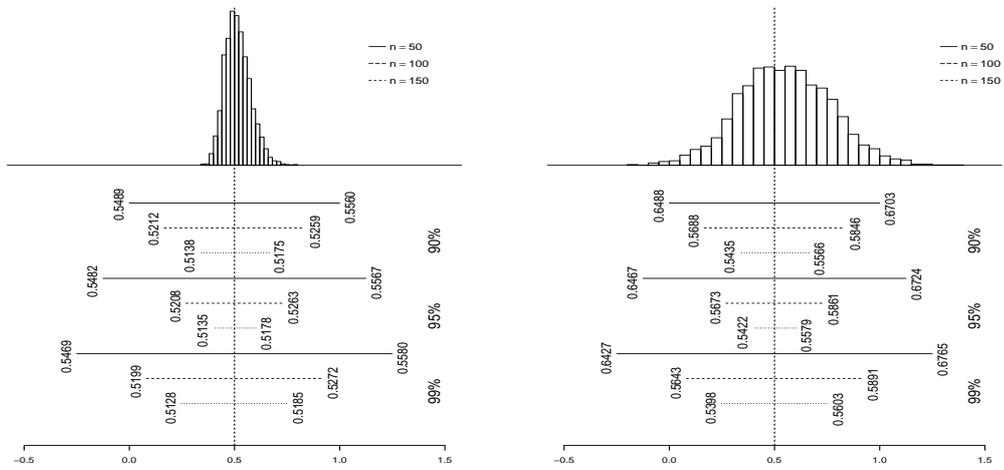


Figura 4.6: Distribuição empírica de  $\hat{\kappa}$ , para  $\kappa = 0.50$ , em que  $\varphi = 2.00$  e  $\varpi = 1.10$ , com  $IC(\kappa; [1-\alpha] \times 100\%)$  para cada  $n$ .

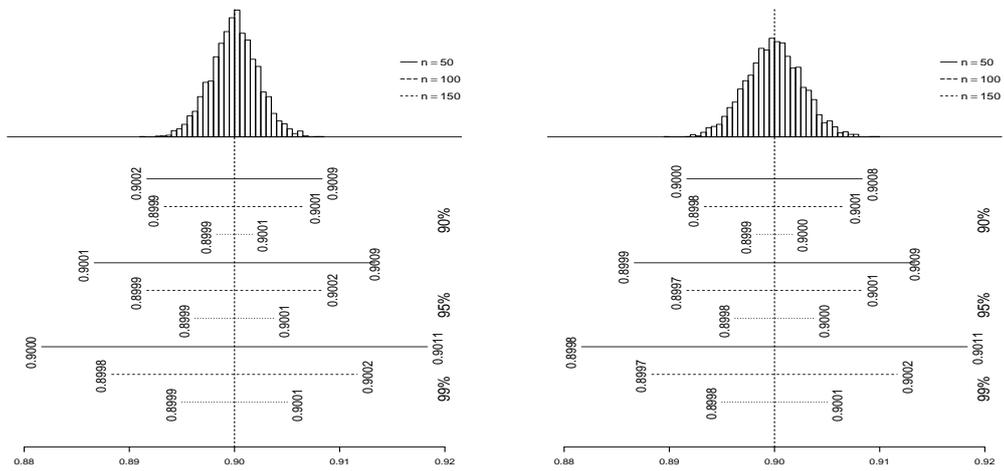


Figura 4.7: Distribuição empírica de  $\hat{\varpi}$ , para  $\varpi = 0.90$ , com  $\varphi = 1.50$  e  $IC(\varpi; [1-\alpha] \times 100\%)$  para cada  $n$ .

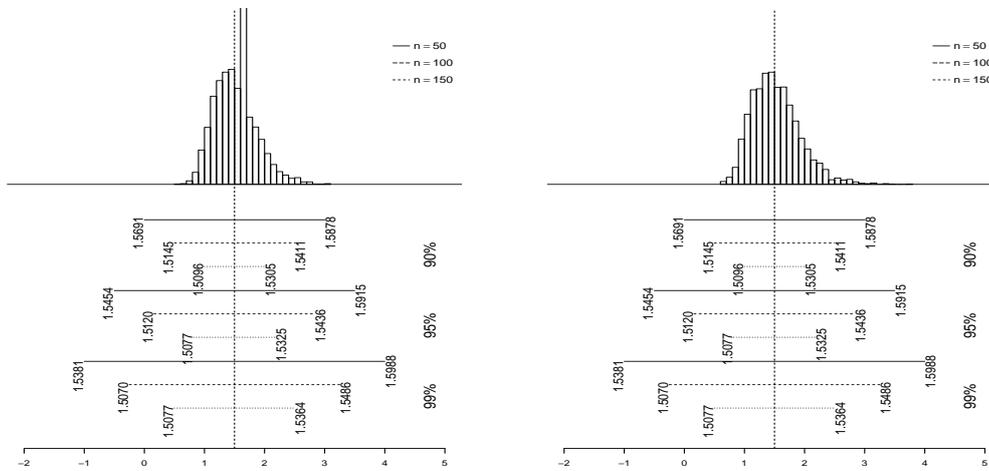


Figura 4.8: Distribuição empírica de  $\hat{\varphi}$ , para  $\varphi = 1.50$ , com  $\varpi = 0.90$  e  $IC(\varphi; [1-\alpha] \times 100\%)$  para cada  $n$ .

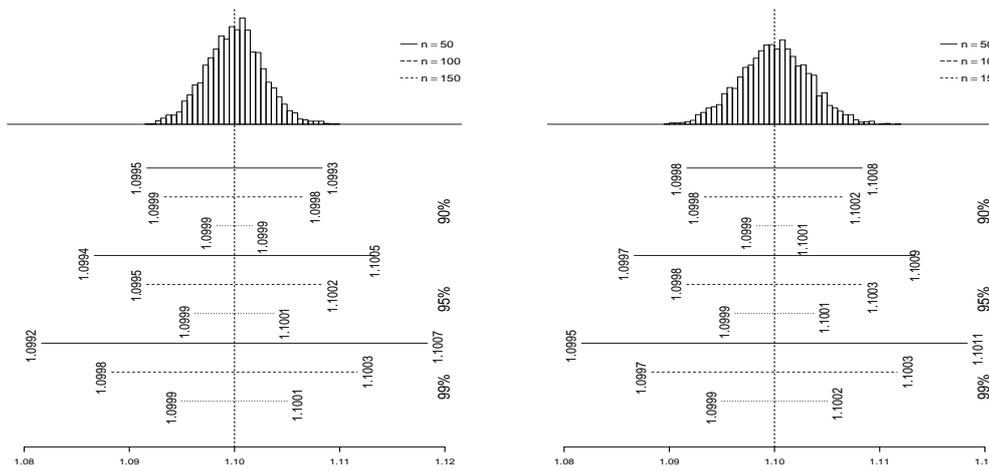


Figura 4.9: Distribuição empírica de  $\hat{\vartheta}$ , para  $\vartheta = 1.10$ , com  $\varpi = 1.50$  e  $IC(\vartheta; [1-\alpha] \times 100\%)$  para cada  $n$ .

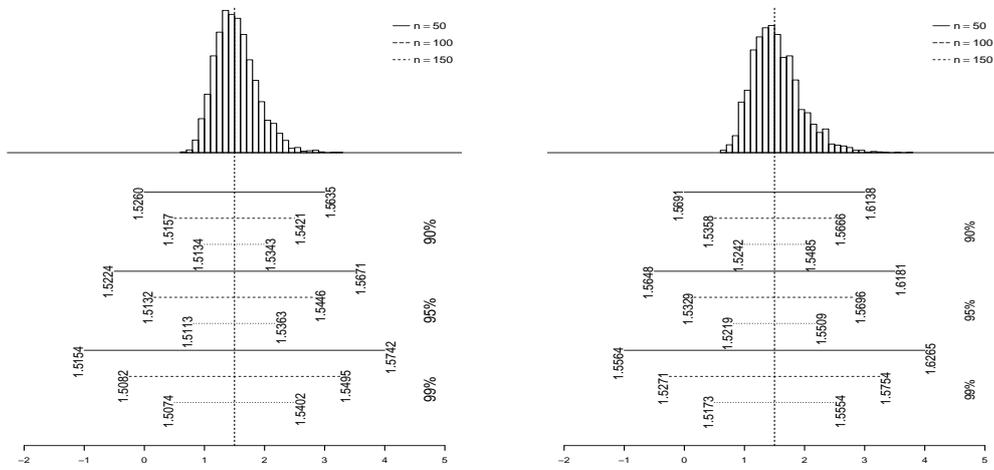


Figura 4.10: Distribuição empírica de  $\hat{\varphi}$ , para  $\varphi = 1.50$ , com  $\varpi = 1.10$  e  $IC(\varphi; [1-\alpha] \times 100\%)$  para cada  $n$ .

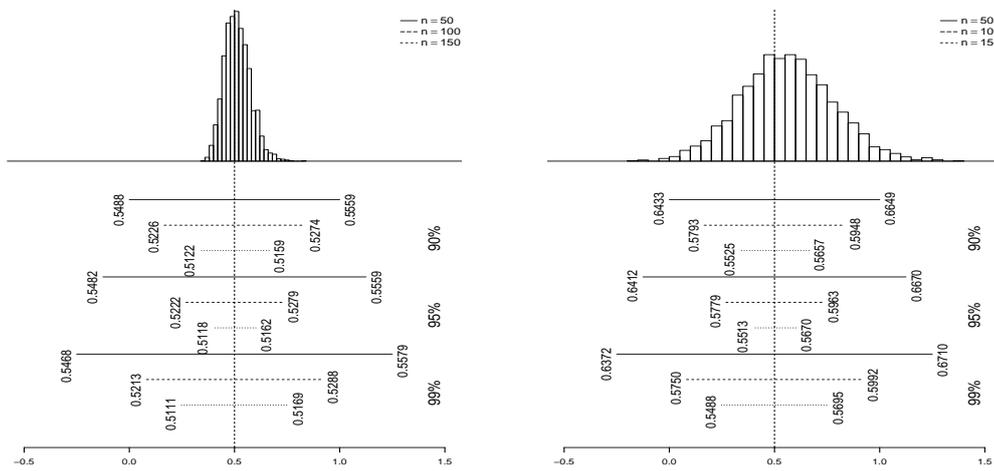


Figura 4.11: Distribuição empírica de  $\hat{\kappa}$ , para  $\kappa = 0.50$ , em que  $\varphi = 1.50$  e  $\varpi = 0.90$ , com  $IC(\kappa; [1-\alpha] \times 100\%)$  para cada  $n$ .

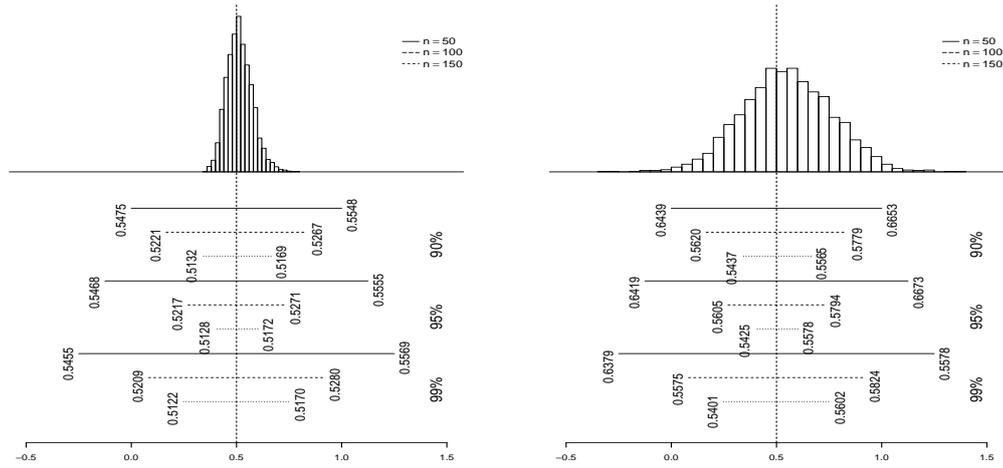


Figura 4.12: Distribuição empírica de  $\hat{\kappa}$ , para  $\kappa = 0.50$ , em que  $\varphi = 1.50$  e  $\varpi = 1.10$ , com  $IC(\kappa; [1-\alpha] \times 100\%)$  para cada  $n$ .

## 4.2 Aplicações em dados reais

Em busca de exemplificar, validar e comparar nosso modelo proposto, utilizaremos quatro bancos de dados reais. O primeiro é o `aircraft`: estes dados fornecem 30 observações que tratam dos tempos de falha do sistema de ar condicionado do modelo de aeronave Boeing 720, de número 7912, estudado originalmente por Proschan (1963). O segundo, `coal mining`, refere-se aos dados de desastres da mineração de carvão da Grã-Bretanha; esses dados já foram analisados em alguns trabalhos, dos quais destacamos Bicer e Bicer (2017), Chan et al. (2014), Kara et al. (2015) e Bicer (2018). O terceiro e o quarto conjuntos de dados são `carnero` e `jiwani`, respectivamente, os quais foram analisados por Linhart e Zucchini (1986) e Bakouch et al. (2019).

### 4.2.1 Dados 1 - aircraft

Os sucessivos tempos de falha, em horas, nos dados da aeronave, `aircraft`, são dados por: 23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95. A Tabela 4.4 apresenta medidas descritivas para estes dados, a saber: média (M), mediana (Med), variância (Var), desvio padrão (DP), coeficiente de variação (CV), assimetria e curtose. De acordo com a Tabela 4.4, notamos uma assimetria positiva e a indicação de caudas leves devido à estimativa do valor da curtose, tais fatos confirmam que o

modelo GP-RBS pode ser usado para o ajuste dos dados.

Tabela 4.4: Estatísticas descritivas para os Dados 1.

<b>M</b>	<b>Med</b>	<b>Var</b>	<b>DP</b>	<b>CV</b>	<b>Assimetria</b>	<b>Curtose</b>
59.60	22.00	5167.42	71.88	1.21	1.61	1.64

Na Tabela 4.5 apresentamos as estimativas dos parâmetros do modelo proposto, obtidos via MV e MM. Os resultados obtidos para o modelo GP-RBS são comparados com outros modelos baseados no processo geométrico que foram propostos na literatura, como por exemplo: Power Lindley, lognormal, gama e gaussiana inversa. Como medidas de “bondade de ajuste” foram obtidos o AIC, BIC e BF. De acordo com os resultados dispostos na Tabela 4.6, que são a log-verossimilhança ( $\ell$ ), AIC, BIC e BF para todos os modelos usados, temos que a decisão sobre o melhor ajuste é tomada de acordo com a interpretação do BF apresentada na Tabela 6 em Leiva et al. (2015). Dessa forma, a Tabela 4.6 indica que o modelo GP-RBS fornece o melhor ajuste quando comparado com os demais modelos em termos do AIC, BIC e BF.

Tabela 4.5: Estimativas dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$  para os Dados 1.

<b>Método</b>	$\hat{\varpi}$	$\hat{\varphi}$	$\hat{\kappa}$
MV	1.06	123.50	0.95
MM	1.05	112.78	1.66

Tabela 4.6: Estimativas da log-verossimilhança, AIC, BIC e BF segundo cada modelo para os Dados 1.

	<b>Modelo</b>				
	<b>Power Lindley</b>	<b>lognormal</b>	<b>gama</b>	<b>gaussiana inversa</b>	<b>GP-RBS</b>
$-\ell$	150.21	150.98	275.54	150.83	85.03
AIC	306.42	307.99	557.08	307.65	176.06
BIC	310.63	312.17	561.28	311.86	180.26
BF	130.36	131.90	381.02	131.60	—

## 4.2.2 Dados 2 - Coal Mining

A próxima aplicação, pertencente a Andrews e Herzberg (1985), é um exemplo sobre os intervalos em dias entre desastres sucessivos na mineração de carvão da Grã-Bretanha, no período de 1851 a 1962. Estes dados incluem 191 intervalos sucessivos e possuem as seguintes estatísticas descritivas, expostas na Tabela 4.7. De maneira semelhante ao exemplo anterior, temos indicação que o modelo GP-RBS pode ser utilizado na modelagem, porém nesse caso temos uma assimetria e curtose menor do que as observadas para os Dados 1, como também uma variabilidade menor que pode ser vista através das estimativas do DV e CV.

Tabela 4.7: Estatísticas descritivas para os Dados 2.

<b>M</b>	<b>Me</b>	<b>Var</b>	<b>DP</b>	<b>CV</b>	<b>Assimetria</b>	<b>Curtose</b>
1889.04	1880.06	859.12	29.31	0.02	0.77	- 0.58

As Tabelas 4.8 e 4.9, exibidas abaixo, apresentam as estimativas dos parâmetros do modelo GP-RBS( $\varpi$ ,  $\varphi$  e  $\kappa$ ). De acordo com as medidas usadas como indicação do modelo mais adequado, observamos através das estimativas dispostas na Tabela 4.9 que o modelo proposto apresenta valores extremamente significativos. Com isso, temos novamente a indicação que o modelo GP-RBS é mais indicado quando comparamos aos modelos propostos por Kara et al. (2015), Chan et al. (2014), Lam e Chan (1998) e Bicer (2018).

Tabela 4.8: Estimativas dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$  para os Dados 2.

<b>Método</b>	$\hat{\varpi}$	$\hat{\varphi}$	$\hat{\kappa}$
MV	0.99	1841.13	101880.04
MM	0.99	1841.47	101887.30

Tabela 4.9: Estimativas da log-verossimilhança para, AIC, BIC e BF para cada modelo considerando os Dados 2.

	<b>Modelo</b>				
	<b>power lindley</b>	<b>lognormal</b>	<b>gamma</b>	<b>gaussiana inversa</b>	<b>GP-RBS</b>
$-\ell$	11785.00	11928.00	11787.00	12528.00	2014.45
AIC	2365.00	2426.00	2366.00	2517.00	2086.99
BIC	23585.76	23871.76	23589.75	25071.76	4209.26
BF	19541.10	19827.10	19545.10	21027.10	—

### 4.2.3 Dados 3 - Carnero

O conjunto de dados Carnero contém valores máximos, de cada ano, de fluxo. As informações detalhadas estão disponíveis no site de pesquisa Geológica (USGS) dos EUA, em <https://earthexplorer.usgs.gov/>. Os valores destes dados são: 144, 179, 105, 280, 81, 35, 320, 248, 159, 570, 278, 315, 327, 182, 186, 102, 31, 350, 435, 520, 715, 1600, 660, 173, 239, 667, 44, 82, 70, 68, 69, 42, 16, 450, 333, 114, 121, 175, 299, 102, 93, 287, 64, 36, 438, 63, 146, 48, 37, 214, 25, 161, 104, 115, 32, 109, 128, 30. A Tabela 4.10 apresenta algumas estatísticas descritivas para estes dados. De acordo com os resultados dispostos nesta tabela, podemos notar que há a indicação de uma distribuição bastante assimétrica, visto o valor estimado da assimetria, como também uma distribuição com caudas pesadas por conta da estimativa do valor da curtose. Dessa forma, temos que o modelo GP-RBS poderá ser usado na modelagem deste banco de dados.

Tabela 4.10: Estatísticas descritivas para os Dados 3.

<b>M</b>	<b>Me</b>	<b>Var</b>	<b>DP</b>	<b>CV</b>	<b>Assimetria</b>	<b>Curtose</b>
224.41	145.00	64328.42	253.63	1.13	3.04	12.68

Tabela 4.11: Estimativas dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$  o para os Dados 3.

Método	$\hat{\varpi}$	$\hat{\varphi}$	$\hat{\kappa}$
MV	1.0205	378.7394	1.9567
MM	1.0207	380.9086	1.8100

A Tabela 4.11 apresenta os resultados das estimativas dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$  para os dados *Carnero*. Além disso, foram obtidos:  $AIC = 460.03$ ,  $-\ell = 227.01$ . Ademais, o BIC encontrado foi de 466.21. Vale ressaltar que os resultados obtidos, neste caso, foram favoráveis ao nosso modelo, pois, como comparativo, em Bakouch et al. (2019) , os autores obtiveram  $AIC = 742.40$  e  $-\ell = 369.20$ . Com isso, em termos de medidas de bondade de ajuste o modelo proposto apresenta um melhor desempenho. Isto é, é o mais indicado em comparação com os demais modelos propostos na literatura.

#### 4.2.4 Dados 4 - *jiwani*

Os dados *jiwani* descreve a precipitação, em *mm* por ano, máxima do Paquistão entre 1981 e 2010. Os valores são: 21.7, 172.9, 69.5, 96.5, 12.6, 265.5, 154, 28, 142.8, 14.2, 74.8, 32.5, 25, 28.5, 113.8, 25.7, 116.3, 28, 16.9, 6, 9, 17.6, 47.3, 55, 129, 72, 92, 28, 113, 194. Logo abaixo estão as estatísticas descritivas destes dados. Através dos resultados dispostos na Tabela 4.12 temos, de forma semelhante aos casos anteriores que o modelo GP-RBS também poderá ser usado no ajuste dos dados.

Tabela 4.12: Estatísticas descritivas para os Dados 4.

M	Me	Var	DP	CV	Assimetria	Curtose
73.4033	51.15	4209.33	64.88	0.88	1.09	0.53

Tabela 4.13: Estimativas dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$  para os Dados 4.

<b>Método</b>	$\hat{\varpi}$	$\hat{\varphi}$	$\hat{\kappa}$
MV	1.01	80.97	1.77
MM	1.00	76.86	2.67

Na Tabela 4.13, disponibilizamos as estimativas dos parâmetros  $\varpi$ ,  $\varphi$  e  $\kappa$  para o conjunto de dados *jiwani*. Obtivemos  $AIC = 206.35$ ,  $-\ell = 100.17$  e  $BIC = 210.5516$ . Novamente, os resultados indicam um melhor desempenho do modelo GP-RBS, uma vez que em Bakouch et al. (2019), os autores obtiveram  $AIC = 319.04$  e  $-\ell = 157.52$ .

# Capítulo 5

## Conclusões e trabalhos futuro

Em busca de um modelo atual, elementar e útil para conjunto de dados com tempos de ocorrência de eventos sucessivos, discutimos neste trabalho, o problema de estimativa de parâmetros para o processo geométrico. Especificamente, propusemos o processo geométrico quando a distribuição do primeiro tempo de ocorrência é Birnbaum–Saunders reparametrizada, resultando no modelo intitulado GP-RBS.

Como mostrado nos capítulos anteriores, o problema de estimação dos parâmetros foi discutido a partir de dois métodos, a saber: estimador de máxima verossimilhanças e o estimador de momentos. Vimos que o estimador de máxima verossimilhanças tem distribuição assintoticamente normal. Como a forma analítica do estimador obtido via método de momentos não pode ser escrita, não podemos dizer o mesmo sobre as propriedades assintóticas deste estimador, porém podemos avaliar numericamente este estimador.

Através do estudo de simulação analisado para o modelo GP-RBS, pode-se concluir que todos os estimadores são assintoticamente não viesados e consistentes, já que tanto o viés e os valores de erro quadrático médio de todos os estimadores diminuem à medida que o tamanho da amostra aumenta.

Quanto aos conjuntos de dados utilizados para demonstrar o desempenho de nosso modelo, os resultados obtidos para o GP-RBS são comparados com outros modelos propostos na literatura, como por exemplo: Power Lindley, lognormal, gamma e gaussiana inversa. Como medidas de “bondade de ajuste” consideramos os seguintes: Critério de Informação de Akaike (AIC), Critério Bayesiano de Schwarz (BIC) e fator bayesiano (BF). Assim sendo, a partir dos resultados obtidos em todas as aplicações consideradas temos a indicação que o GP-RBS apre-

senta o “melhor ajuste” quando comparado com os demais modelos em termos do AIC, BIC e BF. Isso implica dizer que o GP com distribuição RBS é uma alternativa bastante razoável para um GP com distribuições já conhecidas como gama, lognormal, gaussiana inversa ou Power Lindley na modelagem dos sucessivos tempos de inter-chegada.

Por fim, como propostas de trabalhos futuros, pelos resultados alcançados, muitas possibilidades surgiram, como as listadas a seguir: 1 - comparar os métodos de estimações com outros métodos alternativos ( método de log-cumulantes, percentis e função característica, por exemplo); 2 - construir um método de diagnóstico para avaliar o ajuste do modelo, tal qual pelo método de log-cumulantes e/ou distância quadrática; 3 - construir um pacote em R para essa metodologia empregada; 4 - propor uma abordagem Bayesiana via Markov Chain Monte Carlo (MCMC) para a estimação dos parâmetros do modelo; e 5 - podemos ainda empregar o modelo GP-RBS em modelos de variação autorregressiva condicional , como abordado por Chan et al. (2012).

# Apêndice A

## Matriz Informação de Fisher

Para o cálculo da matriz  $I$ , inicialmente, consideramos a matriz Hessiana  $H(\varpi, \varphi, \kappa) \in \mathbb{R}^{3 \times 3}$ , cujos elementos  $h_{ij}$ , ( $i, j = 1, 2, 3$ ) são as segundas derivadas da equação (3.4). Isto é:

$$h_{11} = \frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varpi^2} = -\frac{n(n-1)}{4\varpi^2} - \sum_{i=1}^n \frac{(i-1)\kappa\varphi(\kappa\varphi + ix_i(\kappa+1)\varpi^{i-1})}{(\kappa\varphi\varpi + x_i(\kappa+1)\varpi^i)^2} - \frac{\kappa}{4} \left( \sum_{i=1}^n (i-1)(i-2) \frac{x_i(\kappa+1)\varpi^{i-3}}{\varphi} + \sum_{i=1}^n i(i-1) \frac{\varphi\kappa}{x_i(\kappa+1)\varpi^{i+1}} \right),$$

$$h_{22} = \frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varphi^2} = \frac{n}{2\varphi^2} + \frac{1}{2} \sum_{i=1}^n \frac{x_i(\kappa+1)\varpi^{i-1}}{\varphi^3} - \sum_{i=1}^n \frac{(\kappa\varpi)^2}{(\varpi\varphi\kappa + x_i(\kappa+1)\varpi^i)^2},$$

$$h_{33} = \frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \kappa^2} = -\frac{n}{(\kappa+1)^2} - \sum_{i=1}^n \frac{\varphi\varpi(\varpi\varphi\kappa + x_i(\kappa+1)\varpi^{i-1}) + (\kappa+1)(\varphi\varpi + x_i\varpi^i)}{\alpha} + \sum_{i=1}^n \frac{\varphi(\kappa^3 + 3\kappa^2 + 4\kappa)}{4(\kappa+1)^3(x_i\varpi^{i-1})},$$

$$h_{12} = \frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varpi \partial \varphi} = -\sum_{i=1}^n (i-1)\kappa \frac{x_i(\kappa+1)}{(\varpi\varphi\kappa + x_i(\kappa+1)\varpi^i)^2} + \frac{\kappa}{4} \left( \sum_{i=1}^n (i-1) \frac{\kappa}{x_i(\kappa+1)\varpi^i} + \sum_{i=1}^n (i-1) \frac{x_i(\kappa+1)\varpi^{(i-2)}}{\varphi^2\kappa} \right),$$

$$h_{13} = \frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varpi \partial \kappa} = - \sum_{i=1}^n (i-1) \left[ \frac{\varphi(\varpi \varphi \kappa + x_i(\kappa+1)\varpi^i) + \kappa(\varphi \kappa + x_i \varpi^i)}{(\varpi \varphi \kappa + x_i(\kappa+1)\varpi^i)^2} \right] \\ - \frac{1}{4} \sum_{i=1}^n (i-1) \left[ \frac{x_i \varpi^{i-2}}{\varphi} - \frac{\varphi(\kappa^2 + 2\kappa)}{x_i \varpi^i (\kappa+1)^2} \right],$$

$$h_{23} = \frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varphi \partial \kappa} = \frac{1}{2} \sum_{i=1}^n \frac{x_i \varpi^{i-1}}{\varphi^3} \\ - \sum_{i=1}^n \left[ \frac{2\varpi^2 \kappa (\varpi \varphi \kappa + x_i(\kappa+1)\varphi^i)^2 - \varpi^2 \kappa^2 2(\varpi \varphi \kappa + x_i(\kappa+1)\varphi^i)(\varpi \varphi + x_i \varphi^i)}{(\varpi \varphi \kappa + x_i(\kappa+1)\varpi^i)^4} \right],$$

$$h_{21} = \frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varphi \partial \varpi} = - \sum_{i=1}^n (i-1) \kappa \frac{x_i(\kappa+1)}{(\varpi \varphi \kappa + x_i(\kappa+1)\varpi^i)^2} \\ + \frac{\kappa}{4} \left( \sum_{i=1}^n (i-1) \frac{\kappa}{x_i(\kappa+1)\varpi^i} + \sum_{i=1}^n (i-1) \frac{x_i(\kappa+1)\varpi^{(i-2)}}{\varphi^2 \kappa} \right),$$

$$h_{31} = \frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \kappa \partial \varpi} = - \sum_{i=1}^n (i-1) \left[ \frac{\varphi(\varpi \varphi \kappa + x_i(\kappa+1)\varpi^i) + \kappa(\varphi \kappa + x_i \varpi^i)}{(\varpi \varphi \kappa + x_i(\kappa+1)\varpi^i)^2} \right] \\ - \frac{1}{4} \sum_{i=1}^n (i-1) \left[ \frac{x_i \varpi^{i-2}}{\varphi} - \frac{\varphi(\kappa^2 + 2\kappa)}{x_i \varpi^i (\kappa+1)^2} \right],$$

$$h_{32} = \frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \kappa \partial \varphi} = \frac{1}{2} \sum_{i=1}^n \frac{x_i \varpi^{i-1}}{\varphi^3} \\ - \sum_{i=1}^n \left[ \frac{2\varpi^2 \kappa (\varpi \varphi \kappa + x_i(\kappa+1)\varphi^i)^2 - \varpi^2 \kappa^2 2(\varpi \varphi \kappa + x_i(\kappa+1)\varphi^i)(\varpi \varphi + x_i \varphi^i)}{(\varpi \varphi \kappa + x_i(\kappa+1)\varpi^i)^4} \right].$$

Com isso, a matriz de Informação de Fisher

$$I = - \begin{bmatrix} E\left(\frac{\partial^2 \ln L(\alpha, \theta, r)}{\partial r^2}\right) & E\left(\frac{\partial^2 \ln L(\alpha, \theta, r)}{\partial r \partial \alpha}\right) & E\left(\frac{\partial^2 \ln L(\alpha, \theta, r)}{\partial r \partial \theta}\right) \\ E\left(\frac{\partial^2 \ln L(\alpha, \theta, r)}{\partial \alpha \partial r}\right) & E\left(\frac{\partial^2 \ln L(\alpha, \theta, r)}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 \ln L(\alpha, \theta, r)}{\partial \alpha \partial \theta}\right) \\ E\left(\frac{\partial^2 \ln L(\alpha, \theta, r)}{\partial \theta \partial r}\right) & E\left(\frac{\partial^2 \ln L(\alpha, \theta, r)}{\partial \theta \partial \alpha}\right) & E\left(\frac{\partial^2 \ln L(\alpha, \theta, r)}{\partial \theta^2}\right) \end{bmatrix}$$

possui os seguintes elementos:

$$-E\left(\frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varpi^2}\right) = -E[h_{11}] := E_{11},$$

$$-E\left(\frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varphi^2}\right) = -E[h_{22}] := E_{22},$$

$$-E\left(\frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \kappa^2}\right) = -E[h_{33}] := E_{33},$$

$$-E\left(\frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varpi \partial \varphi}\right) = -E[h_{12}] := E_{12},$$

Concomitantemente,

$$-E\left(\frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \varphi \partial \varpi}\right) = -E[h_{21}] := E_{21},$$

$$-E\left(\frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \kappa \partial \varphi}\right) = -E[h_{31}] := E_{31},$$

$$-E\left(\frac{\partial^2 \ln L(\varpi, \varphi, \kappa)}{\partial \kappa \partial \varphi}\right) = -E[h_{32}] := E_{32}.$$

Ou seja,

$$I = - \begin{bmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{bmatrix}$$

Vale ressaltar que, como mencionado anteriormente, não há uma forma fechada da matriz informações de Fisher  $I$ ; ela é obtida maneira numérica, como descrito no Capítulo 3. No entanto, como estimador, a matriz informação observada pode ser usada em vez da matriz  $I$ , já a matriz de informações observada dos estimadores é o valor negativo da matriz  $H(\theta)$  obtida na última iteração. E, além disso, os elementos da matriz informação de Fisher em (3.1) são imediatamente escritos a partir de elementos da matriz hessiana  $H(\theta)$ .

# Referências Bibliográficas

AHMED, S.; CASTRO-KURISS, C.; FLORES, E.; LEIVA, V.; SANHUEZA, A. A truncated version of the birnbaum-saunders distribution with an application in financial risk. **Pakistan Journal of Statistics**, v. 26, p. 293–311, 01 2010.

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

ANDREWS, D.; HERZBERG, A. **Data**. [S.l.]: Springer: New York, 1985.

AYDOGDU, H.; SENOGLU, B.; KARA, M. Parameter estimation in geometric process with weibull distribution. **Applied Mathematics and Computation**, Elsevier, v. 217, p. 2657–2665, 2010.

BAKOUCH, H.; SABOOR, A.; CHESNEAU, C.; SAEED, A. A new lifetime model by mixing gamma and geometric distributions with fitting hydrologic data. 01 2019.

BALAKRISHNAN, N.; GUPTA, R.; KUNDU, D.; LEIVA, V.; SANHUEZA, A. On some mixture models based on the birnbaum-saunders distribution and associated inference. **Journal of Statistical Planning and Inference**, v. 141, n. 7, p. 2175 – 2190, 2011. ISSN 0378-3758. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S037837581000563X>>.

BICER, C. Statistical inference for geometric process with the power lindley distribution. **Entropy**, v. 20, p. 723, 09 2018.

BICER, C.; BICER, H. Discrimination between gamma and lognormal distributions for geometric process data. Ankara, p. 2830–2836, 2017.

BICER, C.; BICER, H.; KARA, M.; AYDOGDU, H. Statistical inference for geometric process with the rayleigh distribution. **Communications Faculty Of Science University of Ankara Series A1Mathematics and Statistics**, v. 68, p. 149–160, 04 2018.

BIRNBAUM, Z.; SAUNDERS, S. A new family of life distributions. **Journal of applied probability**, JSTOR, p. 319–327, 1969.

- BOURGUIGNON, M.; LEAO, J.; LEIVA, V.; SANTOS-NETO, M. The transmuted birnbaum-saunders distribution. **Working Paper**, 2015.
- BRAUN, W.; LI, W.; ZHAO, Y. Properties of the geometric and related processes. **Nav Res Log**, v. 52, p. 607–616, 10 2005.
- CASTILLO, N.; GOMEZ, G.; BOLFARINE, H. Epsilon birnbaum-saunders distribution family: Properties and inference. **Statistical Papers**, v. 52, p. 871–883, 11 2009.
- CHAN, J.; LAM, C.; YU, P.; CHOY, S.; CHEN, C. A bayesian conditional autoregressive geometric process model for range data. **Computational Statistics & Data Analysis**, Elsevier, v. 56, n. 11, p. 3006–3019, 2012.
- CHAN, J.; LAM, Y.; LEUNG, D. Statistical inference for geometric processes with gamma distributions. **Computational Statistics & Data Analysis**, Elsevier, v. 47, p. 565–581, 2014.
- CHANG, D.; TANG, L. Percentile bounds and tolerance limits for the birnbaum-saunders distribution. **Communications in Statistics - Theory and Methods**, Taylor & Francis, v. 23, n. 10, p. 2853–2863, 1994. Disponível em: <<https://doi.org/10.1080/03610929408831420>>.
- CORDEIRO, G.; LEMONTE, A. The  $\beta$ -birnbaum-saunders distribution: An improved distribution for fatigue life modeling. **Computational Statistics & Data Analysis**, v. 55, n. 3, p. 1445 – 1461, 2011. ISSN 0167-9473. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167947310003816>>.
- COX, D.; BARNDORFF-NIELSEN, O. **Inference and Asymptotics**. [S.l.]: Chapman and Hall/CRC:Boca Raton, FL, USA, 1994.
- DIAZ-GARCIA, J.; LEIVA, V. A new family of life distributions based on the elliptically contoured distributions. **Journal of Statistical Planning and Inference**, v. 128, n. 2, p. 445 – 457, 2005. ISSN 0378-3758. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378375804000072>>.
- FERREIRA, M.; GOMES, I.; LEIVA, V. On an extreme value version of the birnbaum-saunders distribution. Instituto Nacional de Estatística (INE), 2012.
- FIERRO, R.; LEIVA, V.; RUGGERI, F.; SANHUEZA, A. On a birnbaum-saunders distribution arising from a non-homogeneous poisson process. **Statistics & Probability Letters**, v. 83, n. 4, p. 1233 – 1239, 2013. ISSN 0167-7152. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167715212004774>>.
- GENTLE, J. **Computational Statistics**. [S.l.]: Springer-Verlag New York, 2009.

GOMEZ, H.; OLIVARES-PACHECO, J.; BOLFARINE, H. An extension of the generalized birnbaum-saunders distribution. **Statistics & Probability Letters**, v. 79, n. 3, p. 331 – 338, 2009. ISSN 0167–7152. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S016771520800415X>>.

KARA, M.; AYDOGDU, H.; TURKSEN, O. Statistical inference for geometric process with the inverse gaussian distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 85, n. 16, p. 3206–3215, 2015.

KASS, R. E.; RAFTERY, A. Bayes factors. **Journal of the american statistical association**, Taylor & Francis, v. 90, n. 430, p. 773–795, 1995.

KOTZ, S.; LEIVA, V.; SANHUEZA, A. Two new mixture models related to the inverse gaussian distribution. **Methodol Comput Appl Probab** 12, p. 199–212, 2010.

LAM, Y. Geometric processes and replacement problem. **Acta Mathematicae Applicatae Sinica**, Springer, v. 4, p. 366–377, 1988.

\_\_\_\_\_. A note on the optimal replacement problem. **Advances in Applied Probability**, Applied Probability Trust, v. 20, n. 2, p. 479–482, 1988. ISSN 00018678. Disponível em: <<http://www.jstor.org/stable/1427402>>.

\_\_\_\_\_. **Geometric processes and Its Applications**. [S.l.]: World Scientific:Singapore, 2007.

LAM, Y.; CHAN, S. Statistical inference for geometric processes with lognormal distributions. **Computational Statistics & Data Analysis**, Elsevier, v. 27, p. 99–112, 1998.

LAM, Y.; ZHENG, Y.; ZHANG, Y. Some limit theorems in geometric processes. **Acta Mathematicae Applicatae Sinica, English Series, English Series**, Springer, v. 27, p. 405–416, 2003.

LAM, Y.; ZHU, L.; CHAN, J.; LIU, Q. Analysis of data from a series of events by a geometric process model. **Acta Mathematicae Applicatae Sinica, English Series**, v. 20, n. 2, p. 263–282, Jun 2004. ISSN 1618-3932.

LEIVA, V. **the Birnbaum-Saunders Distribution**. [S.l.]: Academic Press, 2016.

LEIVA, V.; BARROS, M.; PAULA, G.; GALEA, M. Influence diagnostics in log-birnbaum-saunders regression models with censored data. **Computational Statistics & Data Analysis**, v. 51, n. 12, p. 5694 – 5707, 2007. ISSN 0167-9473. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167947306003495>>.

LEIVA, V.; PONCE, M.; MARCHANT, C.; BUSTOS-LETELIER, O. Fatigue statistical distributions useful for modeling diameter and mortality of trees. **Revista Colombiana de Estadística**, v. 35, p. 349–367, 12 2012.

LEIVA, V.; SANTOS-NETO, M.; CYSNEIROS, F.; BARROS, M. Birnbaum-saunders statistical modelling: A new approach. **Statistical Modelling**, v. 14, p. 21–48, 02 2014.

LEIVA, V.; TEJO, M.; GUIRAUD, P.; SCHMACHTENBERG, O.; ORIO, P.; MARMOLEJO-RAMOS, F. Modeling neural activity with cumulative damage distributions. **Biological Cybernetics**, v. 109, n. 4-5, p. 421–433, 2015.

LEMONTE, A.; CORDEIRO, G. Birnbaum-saunders nonlinear regression models. **Computational Statistics & Data Analysis**, v. 53, n. 12, p. 4441 – 4452, 2009. ISSN 0167-9473. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S016794730900245X>>.

LINHART, H.; ZUCCHINI, W. **Model Selection**. [S.l.]: Wiley, New York, 1986.

MICHAEL, J.; SCHUCANY, W.; HAAS, R. Generating random variates using transformations with multiple roots. **The American Statistician**, [American Statistical Association, Taylor & Francis, Ltd.], v. 30, n. 2, p. 88–90, 1976. ISSN 00031305. Disponível em: <<http://www.jstor.org/stable/2683801>>.

MONAHAN, J. **Numerical Methods of Statistics**. [S.l.]: Cambridge University Press, 2011.

PROSCHAN, F. Theoretical explanation of observed decreasing failure rate. **Technometrics**, Taylor & Francis, v. 5, n. 3, p. 375–383, 1963. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/00401706.1963.10490105>>.

R-TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2019. Disponível em: <[www.R-project.org](http://www.R-project.org)>.

RIECK, J.; NEDELMAN, J. A log-linear model for the birnbaum-saunders distribution. **Technometrics**, [Taylor & Francis, Ltd., American Statistical Association, American Society for Quality], v. 33, n. 1, p. 51–60, 1991. ISSN 00401706. Disponível em: <<http://www.jstor.org/stable/1269007>>.

SANTOS-NETO, M.; CYSNEIROS, F.; LEIVA, V.; AHMED, S. On new parameterizations of the birnbaum-saunders distribution. **Pakistan Journal of Statistics**, v. 28, n. 1, 2012.

SANTOS-NETO, M.; CYSNEIROS, F.; LEIVA, V.; BARROS, M. On a reparameterized birnbaum-saunders distribution and its moments, estimation and applications. **REVSTAT-Statistical Journal**, v. 12, n. 3, p. 247–272, 2014.

SANTOS-NETO, M.; CYSNEIROS, F.; LEIVA, V.; BARROS, M. et al. Reparameterized birnbaum-saunders regression models with varying precision. **Electronic Journal of Statistics**, The Institute of Mathematical Statistics and the Bernoulli Society, v. 10, n. 2, p. 2825–2855, 2016.

SAULO, H.; LEAO, J.; BOURGUIGNON, M. The kumaraswamy birnbaum-saunders distribution. **Journal of Statistical Theory and Practice**, v. 6, 12 2012.

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.

VILCA, F.; SANHUEZA, A.; LEIVA, V. An extended birnbaum-saunders model and its application. **Stoch Environ Res Risk**, p. 771–782, 2010.

VOLODIN, I.; DZHUNGUROVA, O. On limit distributions emerging in the generalized birnbaum-saunders model. **J Math Sci** **99**, p. 1348–1366, 2000.