



Universidade Federal da Paraíba
Centro de Informática
Programa de Pós-Graduação em Modelagem Matemática e Computacional

DISTRIBUIÇÃO BETA PRIME INFLACIONADA EM ZERO

Gedeao do Nascimento Corpes

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional, UFPB, da Universidade Federal da Paraíba, como parte dos requisitos necessários à obtenção do título de Mestre em Modelagem Matemática e Computacional.

Orientadores: Tarciana Liberal Pereira
Tatiene Correia de Souza

João Pessoa
Junho de 2021

Ata da Sessão Pública de Defesa de Dissertação de Mestrado de GEDEÃO DO NASCIMENTO CORPES, candidato ao título de Mestre em Matemática Computacional, na Área de Modelagem Matemática e Computacional, realizada no 14 de julho de 2021.

1 Aos quatorze dias do mês de julho do ano de 2021, às quatorze horas, por meio de
2 videoconferência, reuniram-se os membros da Banca Examinadora constituída para julgar
3 o Trabalho Final do Sr. **GEDEÃO DO NASCIMENTO CORPES**, vinculado a
4 Universidade Federal da Paraíba sob a matrícula nº **20181021027**, candidato ao grau de
5 Mestre em “*Modelagem Matemática e Computacional*”, na linha de pesquisa “*Modelagem*
6 *Probabilística*” do Programa de Pós-Graduação em Modelagem Matemática e
7 Computacional. A comissão examinadora foi composta pelos professores Tarciana Liberal
8 Pereira de Araújo, Orientadora e Presidente da Banca; Tatiene Correia de Souza,
9 Coorientadora; Cláudio Javier Tablada, Examinador Interno ao Programa; e Rodrigo
10 Bernardo da Silva, Examinador Externo ao Programa. Dando início aos trabalhos, a
11 Presidente da Banca cumprimentou os presentes, comunicou a finalidade da reunião e
12 passou a palavra ao candidato para que o mesmo fizesse, oralmente, a exposição do
13 trabalho de dissertação intitulado “*Distribuição Beta Prime Inflacionada em Zero*”.
14 Concluída a exposição, o candidato foi arguido pela Banca Examinadora que emitiu o
15 seguinte parecer: “*aprovado*”. Do ocorrido, eu, Gean Paulo Pereira Maurício de Barros,
16 Secretário do Programa de Pós-Graduação em Modelagem Matemática e Computacional
17 (PPGMMC), lavrei a presente ata que vai assinada por mim e pelos membros da Banca
18 Examinadora.

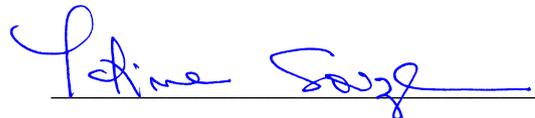
João Pessoa, 14 de julho de 2021.

Gean Paulo Pereira Maurício de Barros
Secretário do PPGMMC
SIAPE 2326476

Prof. ^a Dr.^a Tarciana Liberal de Araújo
Orientadora (PPGMMC)



Prof. ^a Dr.^a Tatiene Correia de Souza
Coorientadora (PPGMMC)



Prof. Dr. Cláudio Javier Tablada
Examinador Interno ao Programa (PPGMMC)



Prof. Dr. Rodrigo Bernardo da Silva
Examinador Externo ao Programa (UFPB)



Catálogo na publicação
Seção de Catalogação e Classificação

C822d Corpes, Gedeão do Nascimento.
Distribuição beta prime inflacionada em zero / Gedeão
do Nascimento Corpes. - João Pessoa, 2021.
64 f. : il.

Orientação: Tarciana Liberal Pereira.
Coorientação: Tatiene Correia Souza.
Dissertação (Mestrado) - UFPB/CI.

1. Matemática computacional. 2. Distribuição
inflacionada. 3. Distribuição beta prime. 4. Simulação
de Monte Carlo. I. Pereira, Tarciana Liberal. II.
Souza, Tatiene Correia. III. Título.

UFPB/BC

CDU 519.6(043)

DISTRIBUIÇÃO BETA PRIME INFLACIONADA EM ZERO

Gedeao do Nascimento Corpes

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL (PPGMMC) DO CENTRO DE INFORMÁTICA DA UNIVERSIDADE FEDERAL DA PARAÍBA E AOS MEMBROS EXTERNOS A ESTE PROGRAMA COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM MODELAGEM MATEMÁTICA E COMPUTACIONAL.

Examinada por:

Prof. Tarciana Liberal Pereira, D.Sc.

Prof. Tatiene Correia de Souza, D.Sc.

Prof. Claudio Javier Tablada, D.Sc.

Prof. Rodrigo Bernardo da Silva, D.Sc.

JOÃO PESSOA, PB – BRASIL

JUNHO DE 2021

*Dedico este trabalho ao meu avô
Arlindo (in memoriam).*

Agradecimentos

Gostaria de agradecer, primeiramente a Deus por ter me dado saúde e forças para encarar mais essa jornada em minha vida com êxito, apesar de todos os obstáculos que surgiram.

Aos meus pais, Francisco Corpes e Gersonita Corpes, aos meus irmãos Gerfson e Fabrícia, por estarem sempre ao meu lado, me apoiando em todas minhas decisões.

As minhas orientadoras, Tarciana e Tatiene, as quais me estenderam as mãos em um dos momentos mais difíceis da minha vida e me ajudaram a continuar nessa jornada. Sou eternamente grato por tudo que vocês fizeram por mim.

Ao meu amigo Carlos Edilon que juntos, enfrentamos diversos obstáculos em João Pessoa, ele no seu programa de mestrado e eu, no PPMMC. Nunca desistimos um do outro, sempre nos demos motivações para continuar.

Aos meus amigos no Pará, que mesmo de longe me deram apoio para continuar, em especial aos meus amigos Francisco Dênis, Antônia Adyelen, Ludimila Farias, Aline Lucimar, Luan Moraes, Fabricio Brasil, Mamede Simas, vocês são especiais na minha vida, obrigado por tudo.

Aos meus amigos, Jefferson e Oscar, que foram meus companheiros de estudo no PPGMMC.

Aos professores do Programa em Pós-graduação em Modelagem Matemática e Computacional da UFPB.

A Prefeitura Municipal de Parauapebas que me proporcionou apoio via licença estudo.

Aos membros da banca examinadora, pelas sugestões e correções.

Resumo da Dissertação apresentada ao PPGMMC/CI/UFPB como parte dos requisitos necessários para a obtenção do grau de Mestre em Modelagem Matemática e Computacional (M.Sc.)

DISTRIBUIÇÃO BETA PRIME INFLACIONADA EM ZERO

Gedeao do Nascimento Corpes

Junho/2021

Orientadores: Tarciana Liberal Pereira
Tatiene Correia de Souza

Programa: Modelagem Matemática e Computacional

As distribuições de probabilidade, sejam elas discretas ou contínuas, podem encontrar barreiras quando falamos de dados que contenham "zeros em excessos" [LORD *et al.* \(2005\)](#). Para que essas distribuições possam mensurar tais dados, criam-se distribuições mistas chamadas de distribuições inflacionadas. A forma como esse inflacionamento vai ocorrer dependerá do conjunto suporte da distribuição. Nesta pesquisa, propomos a construção da distribuição beta prime inflacionada em zero (BPIZ) a partir da reparametrização apresentada em [BOURGUIGNON *et al.* \(2018\)](#). Determinamos também estimadores de máxima verossimilhança e intervalos de confiança para o modelo BPIZ. Avaliamos numericamente os estimadores e os intervalos de confiança, onde verificamos sua eficiência assintótica. Por fim, realizamos uma aplicação para verificar, também, sua aplicabilidade em relação à dados reais.

Palavras chave: Distribuições inflacionadas; distribuição Beta Prime; Estimadores de Máxima Verossimilhança; Intervalo de Confiança; Simulação de Monte Carlo.

Abstract of Dissertation presented to PPGMMC/CI/UFPB as a partial fulfillment of the requirements for the degree of Master of Mathematical and Computational Modeling (M.Sc.)

DISTRIBUIÇÃO BETA PRIME INFLACIONADA EM ZERO

Gedeao do Nascimento Corpes

June/2021

Advisors: Tarciana Liberal Pereira
Tatiene Correia de Souza

Program: Computational Mathematical Modelling

The probability distributions whether discrete or continuous can encounter barriers when we talk about data that contains "zeros in excess" [LORD *et al.* \(2005\)](#). In order for these distributions to measure such data, mixed distributions called inflated distributions are created. The way in which this inflation will occur will depend of the distribution support set . In this research, we propose the construction of the zero inflated beta prime distribution (BPIZ) from the reparametrization presented in [BOURGUIGNON *et al.* \(2018\)](#). We also determined maximum likelihood estimators and confidence intervals for the model BPIZ. We evaluate numerically estimators and confidence intervals where we check its asymptotic efficiency. Finally, we performed an application to check also its applicability in relation to real data.

Keywords: Inflated Distribution; Beta Prime distribution; Maximum Likelihood Estimators; Confidence Interval; Monte Carlo Simulation.

Sumário

Lista de Figuras	viii
Lista de Tabelas	ix
1 Introdução	1
2 Conceitos Básicos	4
2.1 Inferência Estatística	4
2.1.1 Estimação Pontual	4
2.1.2 Estimação Intervalar	7
2.2 Distribuição Beta	9
2.3 Distribuição Beta Prime	10
3 Distribuição Beta Prime Inflacionada	12
3.1 Valor Esperado e Variância da Distribuição BPIZ	14
3.2 Estimação por Máxima Verossimilhança da Distribuição BPIZ	15
3.3 Matriz de Informação de Fisher BPIZ	19
3.4 Intervalos de Confiança BPIZ	22
4 Avaliação Numérica	24
4.1 Estimação Pontual	24
4.2 Estimação Intervalar	30
5 Aplicação	35
5.1 Aplicação da Distribuição BPIZ	36
6 Conclusões	40
Referências Bibliográficas	41
A Algumas Taxas de Cobertura	44
B Algoritmos	47

Lista de Figuras

3.1	Densidades da distribuição BPIZ, com parâmetro de mistura $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$, média $\mu = 0.9$ e precisão $\phi = 35$	13
3.2	Densidades da distribuição BPIZ, com parâmetro de mistura $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$, média $\mu = 10$ e precisão $\phi = 35$	13
3.3	Densidades da distribuição BPIZ, com parâmetro de mistura $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$, média $\mu = 0.9$ e precisão $\phi = 70$	14
3.4	Densidades da distribuição BPIZ, com parâmetro de mistura $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$, média $\mu = 10$ e precisão $\phi = 70$	14
5.1	Histograma e densidades estimadas das 2982 observações da concentração de ER de pacientes diagnosticados com câncer de mama primário.	37
5.2	Funções de distribuição acumuladas para as 2982 observações da concentração de ER de pacientes diagnosticados com câncer de mama primário.	38
5.3	Gráfico Quantil-Quantil (Q-Q Plot) para as 2982 observações da concentração de ER de pacientes diagnosticados com câncer de mama primário.	39

Lista de Tabelas

4.1	Cenários 1, 2 e 3, contendo os resultados da avaliação numérica da distribuição BPIZ para a Média, EQM e VR dos estimadores. Considerando $\mu = 0.9$ e $\phi = 35$, com $\lambda = (0.15, 0.25, 0.35)$	26
4.2	Cenários 4, 5 e 6, contendo os resultados da avaliação numérica da distribuição BPIZ para a Média, EQM e VR dos estimadores. Considerando $\mu = 10$ e $\phi = 35$, com $\lambda = (0.15, 0.25, 0.35)$	27
4.3	Cenários 7, 8 e 9, contendo os resultados da avaliação numérica da distribuição BPIZ para a Média, EQM e VR dos estimadores. $\mu = 0.9$ e $\phi = 70$, com $\lambda = (0.15, 0.25, 0.35)$	28
4.4	Cenários 10, 11 e 12, contendo os resultados da avaliação numérica da distribuição BPIZ para a Média, EQM e VR dos estimadores. $\mu = 10$ e $\phi = 70$, com $\lambda = (0.15, 0.25, 0.35)$	29
4.5	Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança dos parâmetros da distribuição beta prime inflacionada com $\mu = 0.9$ e $\phi = 35$, para $\lambda = (0.15, 0.25, 0.35)$	31
4.6	Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança dos parâmetros da distribuição beta prime inflacionada com $\mu = 10$ e $\phi = 35$, para $\lambda = (0.15, 0.25, 0.35)$	32
4.7	Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança dos parâmetros da distribuição beta prime inflacionada com $\mu = 0.9$ e $\phi = 70$, para $\lambda = (0.15, 0.25, 0.35)$	33
4.8	Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança dos parâmetros da distribuição beta prime inflacionada com $\mu = 10$ e $\phi = 70$, para $\lambda = (0.15, 0.25, 0.35)$	34
5.1	Medidas descritivas da quantidade de proteína ER presente na corrente sanguínea dos pacientes avaliados.	36

A.1	Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança, à nível de confiança de 90%, dos parâmetros da distribuição beta prime inflacionada com $\mu = 0.9$ e $\phi = 35$, para $\lambda = (0.15, 0.25, 0.35)$	44
A.2	Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança, à nível de confiança de 90%, dos parâmetros da distribuição beta prime inflacionada com $\mu = 10$ e $\phi = 35$, para $\lambda = (0.15, 0.25, 0.35)$	45
A.3	Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança, à nível de confiança de 90%, dos parâmetros da distribuição beta prime inflacionada com $\mu = 0.9$ e $\phi = 70$, para $\lambda = (0.15, 0.25, 0.35)$	45
A.4	Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança, à nível de confiança de 90%, dos parâmetros da distribuição beta prime inflacionada com $\mu = 10$ e $\phi = 70$, para $\lambda = (0.15, 0.25, 0.35)$	46

Capítulo 1

Introdução

O processo de estudo desta pesquisa está ligado diretamente ao inflacionamento da distribuição beta prime, que será apresentada no capítulo 2. Mas antes precisamos compreender o porquê e como surgem as distribuições inflacionadas. Na literatura existem vários autores que propuseram modelos de distribuições inflacionadas, dos quais podemos citar: [LORD *et al.* \(2005\)](#), [Ospina e Ferrari \(2012\)](#), [TONG *et al.* \(2013\)](#), [LEIVA *et al.* \(2016\)](#), [Rodrigues \(2016\)](#), [TOMAZELLA *et al.* \(2019\)](#). Essas pesquisas mostram quando devemos inflacionar uma distribuição. Os autores nos dizem que quando os dados coletados apresentam zeros em excessos, sob certas circunstâncias (razoavelmente comuns), tais como baixa exposição e/ou seleção inadequada, cria-se a necessidade de encontrar uma distribuição que seja capaz de abranger os zeros em excesso. Quando isso acontece, recorreremos à distribuição que será denominada mista, em que são ‘misturadas’ duas distribuições, uma chamada distribuição degenerada em zero e outra que pode ser uma distribuição discreta ou contínua.

A primeira distribuição inflacionada em zero foi apresentada de forma explícita por [Johnson e Kotz \(1969\)](#) que propuseram um modelo de Poisson modificado ou conhecido como Poisson com zeros adicionados. [Ospina e Ferrari \(2010\)](#) inflacionaram a distribuição beta em zero e/ou um, para modelar dados nos intervalos $[0, 1)$, $(0, 1]$ e $[0, 1]$, adicionalmente propuseram modelos de regressão beta inflacionados. [TONG *et al.* \(2013\)](#) propuseram um modelo de distribuição gama com zeros ajustados para analisar a perda de empréstimos hipotecários em casos de inadimplência. [LEIVA *et al.* \(2016\)](#) propuseram uma metodologia de logística de estoque, modelados por meio da distribuição Birnbaum-Saunders reparametrizada com zeros ajustados, para permitir a mensuração de dados que contenham zeros. [NOBRE *et al.* \(2017\)](#) utilizaram a distribuição gama inflacionada para avaliar o comportamento de pessoas que praticam atividades físicas e mensurar quantas horas por dia elas gastam treinando. O estudo mostrou que há situações em que pessoas não praticam nenhuma atividade física, implicando que os dados podem conter zeros. [TOMAZELLA *et al.* \(2019\)](#)

propuseram uma distribuição mista com um ponto de concentração no zero para a distribuição Birnbaum-Saunders reparametrizada. [Cribari-Neto e Santos \(2019\)](#), propuseram distribuições Kumaraswamy inflacionadas em zeros e/ou uns, considerando o intervalo unitário padrão, fechado em zero ou em um, ou ainda, em zero e em um como uma alternativa em relação ao uso da distribuição beta inflacionada.

A distribuição beta prime (BP), conhecida na literatura como distribuição beta do segundo tipo ou mesmo distribuição beta invertida, é uma distribuição de probabilidade absolutamente contínua, com suporte nos reais positivos. De acordo com [BOURGUIGNON *et al.* \(2018\)](#), existem poucos trabalhos em relação a distribuição BP. Dos trabalhos existentes na literatura, podemos citar por exemplo, [Abernathy e Smith \(1993\)](#) onde os autores utilizaram a expansão em séries para a distribuição beta invertida com o intuito de encontrar percentis da distribuição $-F$. [McDonald e Xu \(1995\)](#) mostraram a distribuição BP como consequência da generalização da distribuição beta do segundo tipo. [BEKKER *et al.* \(2009\)](#) utilizaram uma generalização da distribuição BP e mostraram que ela é útil em estudos de distribuições preditivas, para analisar o comportamento de ações no mercado financeiro. [TULUPYEV *et al.* \(2013\)](#) trouxeram a distribuição beta prime, nos parâmetros usuais, em modelo de regressão a fim de analisar dados de uma pesquisa realizada dentro do grupo de risco de pessoas suscetíveis a contrair doenças infectocontagiosas. [BOURGUIGNON *et al.* \(2018\)](#) realizaram um estudo sobre a distribuição BP e definiram uma nova parametrização em função da média e da precisão para facilitar o uso em modelos de regressão. Ainda segundo os autores, a distribuição BP pode ser utilizada em estudos de análise de sobrevivência para mensurar dados sobre o tempo de ocorrência de eventos de interesse.

A distribuição BP mostra-se como uma alternativa em relação a algumas distribuições contínuas cujo suporte são os reais positivos ([BOURGUIGNON *et al.*, 2018](#)), como por exemplo as distribuições Birnbaum-Saunders(BS), gama (GA) e Weibull (WB). Porém a distribuição está definida para variáveis aleatórias que assumem valores maiores que zero. Mas o que acontece quando uma variável aleatória que segue uma distribuição beta prime apresenta zeros? [TULUPYEV *et al.* \(2013\)](#) mostraram a necessidade de acomodar valores iguais a zero no uso da distribuição beta prime. Os autores identificaram, que quando os dados continham valores iguais a zero, esses dados não poderiam ser modelados pela distribuição BP e, como recurso, optaram por substituir os valores iguais a zero por 0.5 a fim de que esses dados fossem acomodados por tal distribuição. Segundo [Sellers e Raim \(2016\)](#), numerosos conjuntos de dados contêm zeros em excesso, limitando sua capacidade de serem descritos por meio de um modelo de distribuição absolutamente contínuo. Isso reforça a importância e a necessidade de propor distribuições inflacionadas, capazes de abranger tais dados. No início deste texto mencionamos que, quando uma distribuição de

probabilidade não mensura dados que contenham zeros, é necessário recorrer a um modelo de distribuição que possa abrangê-los, neste caso, um modelo de distribuição misto que conterá uma distribuição degenerada concentrando os zeros e uma outra distribuição discreta ou contínua. Sabe-se que a forma padrão da distribuição beta prime não é capaz de abranger dados que contenham zeros. Considerando esse fato, esta pesquisa propõe a distribuição beta prime inflacionada em zero, seguindo a reparametrização proposta por [BOURGUIGNON *et al.* \(2018\)](#).

Para a obtenção dos resultados desta dissertação recorreremos à linguagem de programação R, um ambiente de software livre para computação e gráficos estatísticos, a qual está disponível para download de forma gratuita na plataforma: <https://www.r-project.org/>. Indicamos, para um estudo mais aprofundado em relação a linguagem R, o manual do R desenvolvido por ([R Development Core Team, 2021](#)).

Este trabalho está organizado em 6 capítulos. No capítulo 2, apresentamos alguns conceitos de inferência estatística, estimação pontual e estimação intervalar, além de tópicos relacionados a distribuição beta. Ainda neste capítulo apresentamos a distribuição BP e algumas definições. No Capítulo 3, apresentamos a proposta da distribuição beta prime inflacionada em zero. A partir desse inflacionamento, são obtidos os estimadores de máxima verossimilhança, a matriz de informação de Fisher e os intervalos de confiança para os parâmetros. No Capítulo 4, apresentamos e comentamos os resultados numéricos para avaliar os estimadores de máxima verossimilhança e intervalos de confiança, apresentados no capítulo 3. Já no capítulo 5, apresentamos aplicações da distribuição beta prime inflacionada à dados reais, com resultados e discussões. As considerações finais seguem no capítulo 6.

Capítulo 2

Conceitos Básicos

Neste capítulo construiremos a base desta pesquisa, a partir de conceitos essenciais de Inferência Estatística, norteados pela estimação pontual e intervalar. Apresentaremos de forma breve, a distribuição beta e por fim, apresentaremos a distribuição beta prime e os principais resultados ligados a ela, bem como valor esperado e variância, além da reparametrização apresentada por [BOURGUIGNON *et al.* \(2018\)](#).

2.1 Inferência Estatística

Para que possamos desenvolver esta pesquisa, precisaremos recorrer a conceitos ligados à inferência estatística, um importante ramo da estatística que objetiva fazer uma análise populacional a partir das informações contidas em uma amostra. De acordo com [Wasserman \(2013\)](#), muitos problemas inferenciais podem ser identificados como sendo um dos três tipos: estimação pontual, intervalos de confiança ou teste de hipóteses. A seguir tomaremos um breve nota relacionada a estimação pontual e estimação intervalar.

2.1.1 Estimação Pontual

Uma sequência Y_1, \dots, Y_n de n variáveis aleatórias (**v.a**), independentes e identicamente distribuídas (*i.i.d.*) com função densidade de probabilidade (*fdp*), $f(y; \theta)$, é chamada de amostra aleatória de tamanho n da população Y , em que θ é um parâmetro desconhecido. Assim, podemos escrever a seguinte expressão:

$$f(y_1, \dots, y_n; \theta) = \prod_{i=1}^n f(y_i; \theta) = f(y_1; \theta) \times \dots \times f(y_n; \theta), \quad (2.1)$$

em que $f(y_1, \dots, y_n; \theta)$ é a densidade conjunta da amostra Y_1, \dots, Y_n . A expressão (2.1) pode ser escrita da seguinte forma:

$$\mathbf{L}(\theta; \mathbf{y}) = \prod_{i=1}^n f(y_i; \theta). \quad (2.2)$$

A expressão (2.2) é denominada de função de verossimilhança de θ relacionada a amostra observada $\mathbf{y} = (y_1, \dots, y_n)$. Para que ela seja tratada de forma mais prática, aplica-se sobre a expressão (2.2) o logaritmo natural, gerando a equação a seguir:

$$\ell(\theta; \mathbf{y}) = \sum_{i=1}^n \log [f(y_i; \theta)]. \quad (2.3)$$

A expressão (2.3) denominada de função de log-verossimilhança, é uma expressão mais simples e de derivação mais prática. A seguir é apresentada a função escore, correspondente a primeira derivada da função (2.3), sendo dada por:

$$U(\theta) = \frac{\partial \ell(\theta; \mathbf{y})}{\partial \theta}. \quad (2.4)$$

O valor esperado da função escore (2.4), sob certas condições de regularidade é igual a zero (Bolfarine e Sandoval, 2001), (Bartoszynski e Niewiadomska-Bugaj, 2007). Segundo Bolfarine e Sandoval (2001, p. 6) um dos grandes problemas da estatística está em encontrar um bom estimador para o parâmetro desconhecido θ . Uma das formas de identificar um bom estimador é utilizar o erro quadrático médio (EQM), que avalia o desempenho de um estimador. O EQM é definido a seguir.

Definição 2.1. *O EQM de um estimador $\hat{\theta}$ do parâmetro θ é dado por:*

$$\text{EQM}[\hat{\theta}] = \text{E}[(\hat{\theta} - \theta)^2]. \quad (2.5)$$

E ainda, o EQM pode ser descrito da seguinte forma:

$$\text{EQM}[\hat{\theta}] = \text{Var}[\hat{\theta}] + B^2(\hat{\theta}),$$

em que $B(\hat{\theta}) = \text{E}[\hat{\theta} - \theta]$ é denominado de viés do estimador de θ .

O uso do viés nos leva a verificar se o estimador é tendencioso, ou seja, se é ou não, não viesado. Para Wasserman (2013, p. 90) muitos estimadores são tendenciosos, por isso, é importante estudar a convergência do estimador, e verificar se ele converge para o valor verdadeiro do parâmetro, a medida que o tamanho amostral aumenta.

Definição 2.2. *Um estimador $\hat{\theta}$ do parâmetro θ é consistente, se $\hat{\theta} \xrightarrow{P} \theta$, em que \xrightarrow{P} denota convergência em probabilidade.*

Ou seja,

$$\lim_{n \rightarrow \infty} P \left(\left| \hat{\theta} - \theta \right| < \epsilon \right) = 1, \quad \forall \epsilon > 0,$$

ou ainda,

$$\lim_{n \rightarrow \infty} P \left(\left| \hat{\theta} - \theta \right| \geq \epsilon \right) = 0, \quad \forall \epsilon > 0.$$

Em que ϵ é um número positivo muito pequeno. Então podemos dizer que um estimador $\hat{\theta}$ do parâmetro θ é consistente se converge para o verdadeiro valor do parâmetro, conforme o tamanho da amostra n aumenta.

Existem vários métodos que nos permitem encontrar um estimador de θ , dentre os quais podemos citar, por exemplo: método dos momentos; método de máxima verossimilhança; método Bayes; método dos mínimos quadrados e etc. Aqui iremos dar nota sobre o método de máxima verossimilhança que será utilizado nesta dissertação.

O método de máxima verossimilhança é recorrido para estimar os parâmetros de um determinado modelo a partir de estimativas que tornam máximo o valor da função de verossimilhança (2.2). Este método foi usado pela primeira vez por Sir Ronald Fisher, em 1922, (Fisher (1922b, p. 310) e Fisher (1922a, p.600)) para encontrar um estimador de um parâmetro desconhecido. No entanto, o método se originou nos trabalhos de Gauss e Bernoulli (Gauss (1809)). Hald et al. (1999) apresentaram um pouco da história acerca da construção do método de máxima verossimilhança. De acordo com Casella e Berger (2010), é considerada a técnica mais popular para derivar estimadores. Então, podemos definir que,

Definição 2.3. *O estimador de máxima verossimilhança (EMV) de θ é o valor de $\theta \in \Theta$ (espaço paramétrico) que maximiza a função de verossimilhança $L(\theta, \mathbf{y})$.*

Então, para obtermos o EMV, igualamos a zero a função escore, $U(\theta) = 0$, uma condição necessária, porém não suficiente, de acordo com Bolfarine e Sandoval (2001, p.36). Para que o EMV seja máximo, é necessário verificar que as segundas derivadas sejam negativas para encontrar o ponto que maximiza a função. Sahoo (2013, p. 421) define alguns passos que resumem o método de estimação por máxima verossimilhança, a saber:

- 1) Obtenha uma amostra aleatória y_1, y_2, \dots, y_n da distribuição de uma população Y com função de densidade de probabilidade $f(y; \theta)$;
- 2) Defina a função de verossimilhança para a amostra y_1, y_2, \dots, y_n ;
- 3) Encontre a expressão para θ que maximize $L(\theta)$, alternativamente, isso pode ser feito diretamente maximizando $\ell(\theta)$;

- 4) Substitua θ por $\hat{\theta}$ para obter uma expressão para o estimador de máxima verossimilhança para θ ;
- 5) Encontre o valor que maximiza a função de verossimilhança.

Existem situações em que não é possível maximizar a função de verossimilhança de forma analítica, então recorreremos a recursos computacionais para maximizá-la numericamente.

A partir da segunda derivada da função log-verossimilhança é possível obter a informação de Fisher (IF), uma forma bem recorrente para determinar a variância do EMV. De acordo com [Wasserman \(2013, p. 128\)](#), a distribuição de $\hat{\theta}$ é aproximadamente Normal e sua variância pode ser calculada analiticamente. A IF é menos o valor esperado das segundas derivadas da função de log-verossimilhança (2.3).

Definição 2.4. *Seja Y uma observação de uma população com fdp $f(y; \theta)$, sendo $f(y; \theta)$ contínua, diferenciável duas vezes tal que seu suporte não depende de θ . Então a IF(θ) de uma única observação Y é dada por:*

$$\text{IF}(\theta) = -\text{E} \left[\frac{\partial^2 \log[f(y_i; \theta)]}{\partial \theta \partial \theta} \right]. \quad (2.6)$$

Seja Y_1, \dots, Y_n uma amostra aleatória da **v.a** Y com fdp $f(y; \theta)$ e IF(θ), a informação total de Fisher (IF $_n$ (θ)) de θ correspondente à amostra observada é a soma da IF das n observações da amostra. Então,

$$\text{IF}_n(\theta) = -\text{E} \left[\sum_{i=1}^n \frac{\partial^2 \log[f(y_i; \theta)]}{\partial \theta \partial \theta} \right] = n\text{IF}(\theta). \quad (2.7)$$

Definição 2.5. *Se Y é uma observação feita sob um vetor paramétrico $\theta = (\theta_1, \dots, \theta_n)$ com função de probabilidade $f(y; \theta)$, então a matriz informação de Fisher em relação a θ a partir de Y é dada por:*

$$\text{IF}(\theta) = (\text{IF}_{i,j}(\theta))_{i,j},$$

em que,

$$\text{IF}_{i,j}(\theta) = -\text{E} \left[\frac{\partial^2 \ell(\theta; \mathbf{y})}{\partial \theta_i \partial \theta_j} \right].$$

2.1.2 Estimação Intervalar

De acordo com [Morettin \(2010, p.223\)](#) a estimação pontual de um parâmetro não possui uma medida do possível erro cometido na estimação. Então para que

possamos expressar a precisão da estimação, se faz necessário estabelecer limites que, com certa probabilidade, incluam o valor verdadeiro do parâmetro. Esses intervalos são conhecidos como intervalos de confiança (IC) e fornecem um método para adicionar mais informações a um estimador $\hat{\theta}$ quando precisamos estimar um parâmetro desconhecido θ (DeGroot, 2012).

Em muitas situações acabamos por fazer uso de distribuições assintóticas, já que nem sempre é fácil trabalhar com a verdadeira distribuição, isso geralmente acontece em grande amostras. Para essas distribuições, são construídos IC assintóticos. De acordo com Mood (1950), os intervalos de confiança assintóticos com base nos estimadores de máxima verossimilhança serão mais curtos, em média, do que os intervalos determinados por qualquer outro estimador.

Sejam Y_1, \dots, Y_n uma amostra aleatória grande de Y com *fdp* $f(y; \theta)$, com o estimador de θ sendo $\hat{\theta}$. Temos:

$$Q(\mathbf{y}; \theta) = \sqrt{n} \frac{\hat{\theta} - \theta}{\sqrt{\text{IF}(\theta)^{-1}}} \stackrel{A}{\approx} N(0, 1), \quad (2.8)$$

em que $Q(\mathbf{y}; \theta)$ é a quantidade pivotal (com distribuição aproximadamente normal $N(0, 1)$), $\stackrel{A}{\approx}$ indica que a distribuição é assintoticamente normal e $\text{IF}(\theta)^{-1}$ é a inversa da matriz de informação de Fisher. A partir da expressão (2.8) podemos construir intervalos de confiança aproximados para θ com coeficiente de confiança $1 - \alpha$. Assim temos a seguinte relação:

$$P \left\{ -z_{\frac{\alpha}{2}} < \sqrt{n} \frac{\hat{\theta} - \theta}{\sqrt{\text{IF}(\hat{\theta})^{-1}}} < z_{\frac{\alpha}{2}} \right\} = 1 - \alpha, \quad (2.9)$$

em que $z_{\frac{\alpha}{2}}$ é o valor tabelado da distribuição Normal padrão, ao nível de confiança $1 - \frac{\alpha}{2}$. Essa desigualdade é equivalente à seguinte expressão:

$$\hat{\theta} - z_{\frac{\alpha}{2}} \frac{\sqrt{\text{IF}(\hat{\theta})^{-1}}}{\sqrt{n}} < \theta < \hat{\theta} + z_{\frac{\alpha}{2}} \frac{\sqrt{\text{IF}(\hat{\theta})^{-1}}}{\sqrt{n}}, \quad (2.10)$$

então o IC a partir da quantidade pivotal apresentada em (2.8), é dado por:

$$\left(\hat{\theta} - z_{\frac{\alpha}{2}} \frac{\sqrt{\text{IF}(\hat{\theta})^{-1}}}{\sqrt{n}}; \hat{\theta} + z_{\frac{\alpha}{2}} \frac{\sqrt{\text{IF}(\hat{\theta})^{-1}}}{\sqrt{n}} \right). \quad (2.11)$$

A expressão (2.11) é chamada de IC assintótico de θ com coeficiente de confiança de aproximadamente $1 - \alpha$ ou ainda, com taxa de cobertura $(1 - \alpha) * 100\%$. Na prática, é comum utilizarmos taxas de cobertura de 90%, 95% e 99%.

2.2 Distribuição Beta

Nesta seção apresentamos a distribuição de probabilidade beta como suporte para o desenvolvimento desta pesquisa, haja vista que a distribuição beta prime é uma das generalizações da família de distribuição beta.

A distribuição beta é uma das distribuições de probabilidade mais conhecidas da estatística, e de acordo com [Sahoo \(2013, p. 163\)](#), ela tem muitas aplicações em Estatística Clássica e Bayesiana. Seu estudo principal está ligado a modelagem do comportamento de variáveis aleatórias positivas, limitadas ao intervalo unitário padrão $(0, 1)$, a exemplo, proporções e porcentagens.

Definição 2.6. *Seja Y uma variável aleatória que segue uma distribuição beta $Y \sim \text{Beta}(\alpha, \beta)$. Sua fdp é dada por:*

$$f(y) = \begin{cases} \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1} & \text{se } 0 < y < 1, \\ 0, & \text{caso contrário.} \end{cases} \quad (2.12)$$

com $y \in (0, 1)$ e α, β valores reais positivos, $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$ a função beta e

$\Gamma(\cdot)$ a função gama definida por: $\Gamma(y) = \int_0^{\infty} e^{-r} r^{y-1} dr$.

O valor esperado e a variância da distribuição beta podem ser obtidos a partir da função geradora de momentos, dada por:

$$E(Y^k) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) + \Gamma(\beta)} \int_0^1 y^{\alpha+k-1} (1-y)^{\beta-1} dy = \frac{\Gamma(\alpha + \beta)\Gamma(\alpha + k)}{\Gamma(\alpha + \beta + 1)\Gamma(\alpha)}. \quad (2.13)$$

A seguir são apresentados o primeiro e o segundo momento da função (2.13), que são uteis para obter o valor esperado e a variância da distribuição beta.

O valor esperado é dado por:

$$E(Y) = \frac{\Gamma(\alpha + \beta)\Gamma(\alpha + 1)}{\Gamma(\alpha + \beta + 1)\Gamma(\alpha)} = \frac{\alpha}{\alpha + \beta}.$$

e a variância é dada por:

$$\text{Var}(Y) = E(Y^2) - E^2(Y) = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}.$$

De acordo com [McDonald e Xu \(1995\)](#), a distribuição beta prime é uma das generalizações da distribuição beta generalizada do segundo tipo.

2.3 Distribuição Beta Prime

Definição 2.7. *Seja Y uma v.a que segue uma distribuição beta prime (BP), $Y \sim BP(y; \alpha, \beta)$. A fdp é dada por:*

$$f(y; \alpha, \beta) = \frac{y^{\alpha-1}(1+y)^{-(\alpha+\beta)}}{B(\alpha, \beta)}, \quad y > 0, \quad (2.14)$$

em que $\alpha > 0$ e $\beta > 0$ são parâmetros de forma e $B(\alpha, \beta)$ é a função beta.

Definição 2.8. *A fda da distribuição BP é dada por:*

$$F(y; \alpha, \beta) = \mathbb{I}_{\frac{y}{1+y}}(\alpha, \beta). \quad (2.15)$$

em que $\mathbb{I}_{\frac{y}{1+y}}(\alpha, \beta) = B(y; \alpha, \beta) = \int_0^y t^{\alpha-1}(1+t)^{-(\alpha+\beta)} dt$ é a função beta incompleta regularizada.

Para determinar o valor esperado e a variância de uma v.a Y que segue a distribuição beta prime, usamos o r -ésimo momento que é dado por:

$$\mathbb{E}[Y^r] = \prod_{i=1}^r \frac{\alpha + i - 1}{\beta - i}, \quad (2.16)$$

com $r \in \mathbb{N}$.

Seguindo (2.16), a esperança e a variância de uma distribuição BP são dadas, respectivamente, por:

$$\mathbb{E}[Y] = \frac{\alpha}{\beta - 1}, \quad (2.17)$$

e

$$\text{Var}[Y] = \frac{\alpha(\alpha + \beta - 1)}{(\beta - 1)^2(\beta - 2)}. \quad (2.18)$$

BOURGUIGNON *et al.* (2018) apresentam uma nova parametrização da distribuição BP em função da média μ e da precisão ϕ para propor um novo modelo de regressão, em que consideraram $\mu = \alpha/(\beta - 1)$ e $\phi = \beta - 2$, com $\alpha = \mu(1 + \phi)$ e $\beta = 2 + \phi$. Com isso, a distribuição beta prime reparametrizada é dada por:

$$f(y; \mu, \phi) = \frac{y^{\mu(\phi+1)-1}(1+y)^{-[\mu(\phi+1)+\phi+2]}}{B(\mu(1+\phi), \phi+2)}, \quad y > 0 \quad (2.19)$$

em que $\mu > 0$ e $\phi > 0$. Para indicar que uma $\mathbf{v.a}$ segue uma distribuição beta prime, usaremos a notação $Y \sim BP(\mu, \phi)$. O valor esperado e a variância, a partir de (2.19), são dados, respectivamente, por:

$$E[Y] = \frac{\mu(1+\phi)}{2+\phi-1} = \mu, \quad (2.20)$$

e

$$\text{Var}[Y] = \frac{\mu(1+\phi)[\mu(1+\phi) + (\phi+2-1)]}{(\phi+2-1)^2(\phi+2-2)} = \frac{\mu(1+\mu)}{\phi}. \quad (2.21)$$

A distribuição BP não pode ser utilizada quando os dados apresentem zeros, ou seja, quando a $\mathbf{v.a}$ Y assume o valor zero. Um exemplo para tal situação poderá ser identificada em [TULUPYEV *et al.* \(2013\)](#), em que é possível verificar que foi proposta a substituição dos valores iguais a zero por 0.5 para que fosse possível utilizar a distribuição BP.

Capítulo 3

Distribuição Beta Prime Inflacionada

Neste capítulo propomos a distribuição beta prime inflacionada em zero (BPIZ) para que a distribuição possa acomodar valores iguais a zero. Aqui serão apresentadas a função densidade de probabilidade e a função de distribuição acumulada, assim como o valor esperado e a variância. Posteriormente, serão construídos os estimadores e intervalos de confiança para a distribuição BPIZ.

Definição 3.1. *A distribuição BPIZ, tem função de distribuição acumulada definida por:*

$$\text{BPIZ}(y; \lambda, \mu, \phi) = \lambda + (1 - \lambda)BP(y; \mu, \phi), \quad (3.1)$$

em que $BP(y; \mu, \phi)$ é a fda da distribuição BP e $\lambda = P(Y = 0)$ é o parâmetro de mistura, com $0 < \lambda < 1$.

Definição 3.2. *A fdp da distribuição BPIZ será dada por:*

$$\text{bpiz}(y; \lambda, \mu, \phi) = \{\lambda^{\mathbb{I}_{\{0\}}(y)}(1 - \lambda)^{1 - \mathbb{I}_{\{0\}}(y)}\} \{bp(y; \mu, \phi)^{1 - \mathbb{I}_{\{0\}}(y)}\}, \quad (3.2)$$

em que $BP(y; \mu, \phi)$ é a fdp da distribuição BP, $\mathbb{I}_{\{0\}}(y)$ é a função indicadora avaliada no ponto zero (assume 1 se $Y = 0$ e 0 se $Y \neq 0$) e $0 < \lambda < 1$, $\mu, \phi > 0$ são os parâmetros da distribuição beta prime inflacionada, tal que μ é a média condicional e ϕ é a precisão.

A seguir apresentaremos algumas representações gráficas da densidade da distribuição BPIZ para os valores dos parâmetros usados na simulação. As Figuras (3.1) e (3.2), representam densidades com média $\mu = 0.9$ e $\mu = 10$, respectivamente, e precisão $\phi = 35$ em ambas figuras. O parâmetro de mistura λ é dado por: $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$, em ambas as figuras. Já as Figuras (3.3) e (3.4), representam densidades com média $\mu = 0.9$ e $\mu = 10$, respectivamente, e precisão $\phi = 70$ em ambas figuras. O parâmetro de mistura λ é dado por: $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$,

em ambas as figuras. Nas figuras, perceba que existem pontos de concentração de zeros, representados nos gráficos pela linha vertical. Veja ainda que o pico da curva é mais acentuado quando a média é pequena e a concentração de zeros é baixa, porém quando a média é grande, a curva tende a ser mais baixa e ainda, como podemos ver também, a linha vertical de concentração de zeros é mais alta que o pico. Note que os gráficos apresentam comportamento assimétricos, haja vista que estamos avaliando um função mista, (com uma parte discreta) a fim de considerar os zeros presentes.

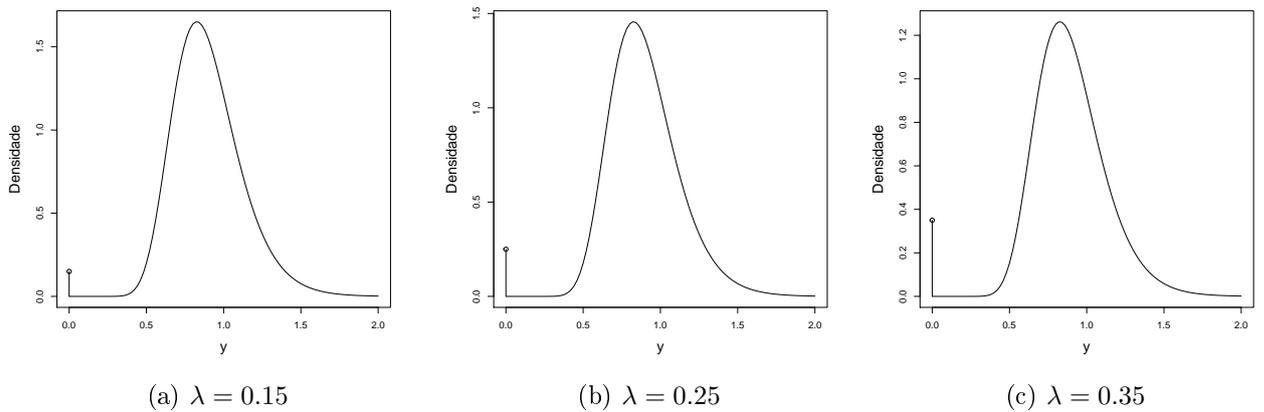


Figura 3.1: Densidades da distribuição BPIZ, com parâmetro de mistura $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$, média $\mu = 0.9$ e precisão $\phi = 35$.

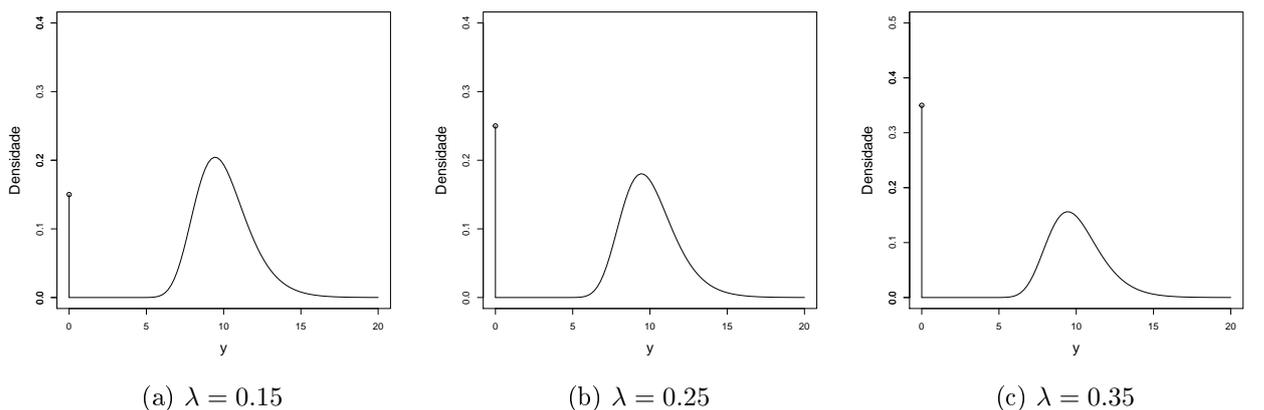


Figura 3.2: Densidades da distribuição BPIZ, com parâmetro de mistura $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$, média $\mu = 10$ e precisão $\phi = 35$.

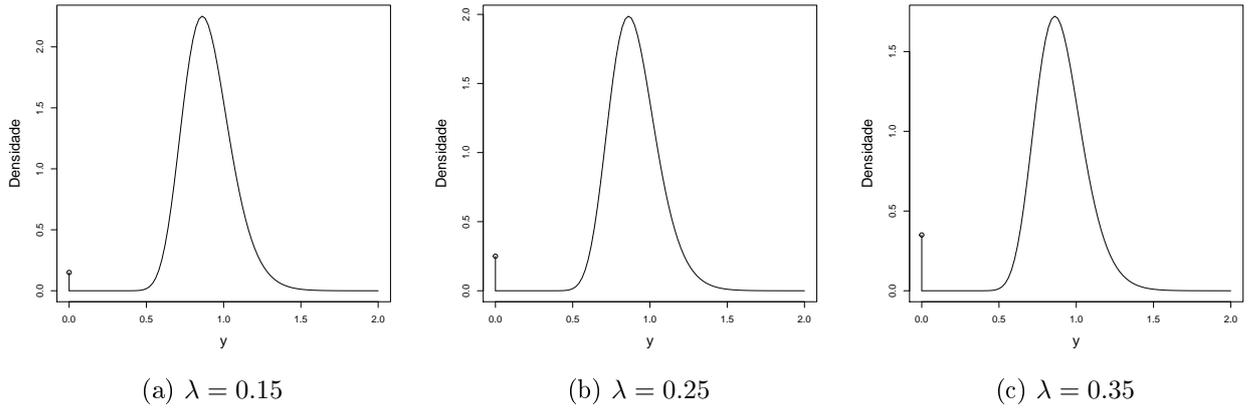


Figura 3.3: Densidades da distribuição BPIZ, com parâmetro de mistura $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$, média $\mu = 0.9$ e precisão $\phi = 70$.

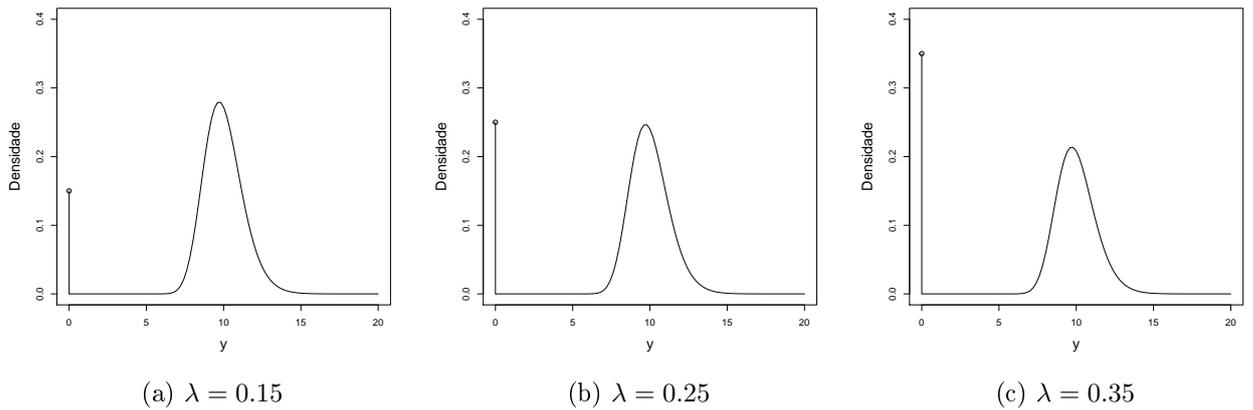


Figura 3.4: Densidades da distribuição BPIZ, com parâmetro de mistura $\lambda = 0.15$; $\lambda = 0.25$ e $\lambda = 0.35$, média $\mu = 10$ e precisão $\phi = 70$.

3.1 Valor Esperado e Variância da Distribuição BPIZ

O valor esperado e a variância de uma variável aleatória $Y \sim \text{BPIZ}(y; \lambda, \mu, \phi)$ podem ser obtidos a partir da seguinte relação de igualdade:

$$E[Y^r] = E[E(Y^r | \mathbb{I}_{\{0\}}(y))]$$

e

$$\text{Var}(Y) = E[\text{Var}(Y | \mathbb{I}_{\{0\}}(y))] + \text{Var}[E(Y | \mathbb{I}_{\{0\}}(y))],$$

em que:

$$E(Y^r | \mathbb{I}_{\{0\}}(y)) = \begin{cases} 0, & \text{com probabilidade } \lambda, \\ \mu_r, & \text{com probabilidade } 1 - \lambda, \end{cases}$$

e

$$\text{Var}(Y | \mathbb{I}_{\{0\}}(y)) = \begin{cases} 0, & \text{com probabilidade } \lambda, \\ \frac{\mu(1+\mu)}{\phi}, & \text{com probabilidade } 1 - \lambda. \end{cases} \quad (3.3)$$

O r -ésimo momento e a variância da distribuição BPIZ são dados por:

$$\begin{aligned} E(Y^r) &= (1 - \lambda)\mu_r \\ \text{Var}(Y) &= (1 - \lambda)\frac{\mu(1 + \mu)}{\phi} + \lambda(1 - \lambda)\mu^2, \end{aligned} \quad (3.4)$$

em que μ_r é o r -ésimo momento em torno da distribuição $BP(\mu, \phi)$ dada em (2.19), o valor esperado de uma variável aleatória que segue uma distribuição BPIZ é dado por:

$$E[Y] = (1 - \lambda)\mu,$$

sendo μ a média condicional de Y dado $Y > 0$ e ϕ é considerado parâmetro de precisão da distribuição beta prime.

3.2 Estimação por Máxima Verossimilhança da Distribuição BPIZ

Nesta seção construiremos estimadores de máxima verossimilhança a partir da distribuição BPIZ. Para isso, precisaremos construir mecanismos que nos possibilitem obter tais estimadores. Assim seguiremos a seguinte ordem: função de verossimilhança; função log-verossimilhança e função score. Após a obtenção desses estimadores é construída a matriz de informação de Fisher.

Dada a expressão (3.2) que representa a função densidade de probabilidade de uma distribuição beta prime inflacionada em zero, seja y_1, y_2, \dots, y_n uma amostra aleatória, de tamanho n , *i.i.d* com $\theta = (\lambda, \mu, \phi)$. Podemos construir a função de verossimilhança da seguinte forma,

$$L(\theta; y) = \prod_{i=1}^n \text{bpiz}(y_i, \lambda, \mu, \phi).$$

Podemos expressar $L(\theta; y)$ como:

$$L(\theta; y) = L_1(\lambda; y) \times L_2(\mu, \phi; y)$$

em que,

$$\begin{aligned} L_1(\lambda; y) &= \prod_{i=1}^n \lambda^{\mathbb{I}_{\{0\}}(y_i)} (1 - \lambda)^{1 - \mathbb{I}_{\{0\}}(y_i)} \\ &= \lambda^{\sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i)} (1 - \lambda)^{n - \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i)} \end{aligned}$$

e,

$$L_2(\mu, \phi; y) = \prod_{i=1}^n f(y_i; \mu, \phi)^{1 - \mathbb{I}_{\{0\}}(y_i)}.$$

em que $f(\cdot)$ é a função densidade de uma $v.a$ que segue uma distribuição BP. Note que $L(\theta; y)$ foi fatorada em dois termos, onde o primeiro depende apenas de λ e o segundo depende apenas de μ e ϕ .

A função log-verossimilhança da distribuição BPIZ, é dada da seguinte forma:

$$\begin{aligned} \ell(\theta; y) &= \log L(\theta; y) = \log L_1(\lambda; y) + \log L_2(\mu, \phi; y) \\ \ell(\theta; y) &= \ell_1(\lambda; y) + \ell_2(\mu, \phi; y), \end{aligned}$$

em que

$$\ell_1(\lambda; y) = \log(\lambda) \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) + \log(1 - \lambda) \left[n - \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) \right]$$

e

$$\begin{aligned} \ell_2(\mu, \phi; y) &= [\mu(\phi + 1) - 1] \sum_{i=1}^n \log(y_i)[1 - \mathbb{I}_{\{0\}}(y_i)] \\ &\quad - [\mu(\phi + 1) + \phi + 2] \sum_{i=1}^n \log(1 + y_i)[1 - \mathbb{I}_{\{0\}}(y_i)] + \{-\log[\Gamma(\mu(\phi + 1))]\} \\ &\quad - \log[\Gamma(\phi + 2)] + \log[\Gamma(\mu(\phi + 1) + \phi + 2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)]. \end{aligned}$$

A função escore é determinada a partir da primeira derivada em relação a função de log-verossimilhança, com respeito aos parâmetros desconhecidos, com $U(\theta) = [U_\lambda(\lambda), U_\mu(\mu, \phi), U_\phi(\mu, \phi)]$, sendo

$$U_\lambda(\lambda) = \frac{\partial \ell_1(\lambda; y)}{\partial \lambda} = \frac{1}{\lambda} \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) - \frac{1}{1 - \lambda} \left[n - \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) \right].$$

Após calcularmos a derivada de $\ell_1(\lambda; y)$ em relação a λ , vamos determinar a derivada de $\ell_2(\mu, \phi; y)$ em relação a μ , tal que $U_\mu(\mu, \phi) = \frac{\partial \ell_2(\mu, \phi)}{\partial \mu}$, logo temos que:

$$\begin{aligned} U_\mu(\mu, \phi) &= (\phi + 1) \left[\sum_{i=1}^n \log(y_i)[1 - \mathbb{I}_{\{0\}}(y_i)] \right] - (\phi + 1) \left[\sum_{i=1}^n \log(1 + y_i)[1 - \mathbb{I}_{\{0\}}(y_i)] \right] \\ &\quad + [-(\phi + 1)\Psi(\mu(\phi + 1)) + (\phi + 1)\Psi(\mu(\phi + 1) + \phi + 2)] \left[\sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right], \end{aligned}$$

em que, $\Psi(\cdot)$ é a função digama, definida por $\Psi(h) = \frac{\partial \log(\Gamma(h))}{\partial h}$, isto é, $\Psi(\cdot)$ a primeira derivada em relação a função gama. Então, $U_\mu(\mu, \phi)$ pode ser organizada da seguinte forma:

$$\begin{aligned} U_\mu(\mu, \phi) &= (\phi + 1) \left[\sum_{i=1}^n \log(y_i)[1 - \mathbb{I}_{\{0\}}(y_i)] - \sum_{i=1}^n \log(1 + y_i)[1 - \mathbb{I}_{\{0\}}(y_i)] \right] \\ &\quad - (\phi + 1) [\Psi(\mu(\phi + 1)) - \Psi(\mu(\phi + 1) + \phi + 2)] \left[\sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right]. \end{aligned}$$

Daí segue que,

$$\begin{aligned} U_\mu(\mu, \phi) &= (\phi + 1) \left\{ \left[\sum_{i=1}^n \log\left(\frac{y_i}{1 + y_i}\right) [1 - \mathbb{I}_{\{0\}}(y_i)] \right] \right. \\ &\quad \left. - [\Psi(\mu(\phi + 1)) - \Psi(\mu(\phi + 1) + \phi + 2)] \left[\sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right] \right\}. \end{aligned}$$

Assim,

$$U_\mu(\mu, \phi) = (\phi + 1) \left\{ \sum_{i=1}^n \left[\log \left(\frac{y_i}{1 + y_i} \right) - [\Psi(\mu(\phi + 1)) - \Psi(\mu(\phi + 1) + \phi + 2)] \right] [1 - \mathbb{I}_{\{0\}}(y_i)] \right\}.$$

Assim, escrevemos

$$U_\mu(\mu, \phi) = (\phi + 1) \sum_{i=1}^n (y_i^* - \mu^*) [1 - \mathbb{I}_{\{0\}}(y_i)],$$

em que $y_i^* = \log \left(\frac{y_i}{1 + y_i} \right)$ e $\mu^* = [\Psi(\mu(\phi + 1)) - \Psi(\mu(\phi + 1) + \phi + 2)]$.

A derivada de $\ell_2(\mu, \phi; y)$ em relação a ϕ , tal que $U_\phi(\mu, \phi) = \frac{\partial \ell_2(\mu, \phi)}{\partial \phi}$, é dada por:

$$U_\phi(\mu, \phi) = \mu \left[\sum_{i=1}^n \log(y_i) [1 - \mathbb{I}_{\{0\}}(y_i)] \right] - (\mu + 1) \left[\sum_{i=1}^n \log(1 + y_i) [1 - \mathbb{I}_{\{0\}}(y_i)] \right] +$$

$$+ [-\mu \Psi(\mu(\phi + 1)) - \Psi(\phi + 2) + (\mu + 1) \Psi(\mu(\phi + 1) + \phi + 2)] \left[\sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right].$$

Assim, $U_\phi(\mu, \phi)$ pode ser organizada da seguinte forma:

$$U_\phi(\mu, \phi) = \left[\sum_{i=1}^n \mu \log(y_i) [1 - \mathbb{I}_{\{0\}}(y_i)] - \sum_{i=1}^n (\mu + 1) \log(1 + y_i) [1 - \mathbb{I}_{\{0\}}(y_i)] \right] +$$

$$\{-\mu \Psi(\mu(\phi + 1)) - \Psi(\phi + 2) + (\mu + 1) \Psi(\mu(\phi + 1) + \phi + 2)\} \left[\sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right].$$

$$U_\phi(\mu, \phi) = \sum_{i=1}^n \log \left(\frac{y_i^\mu}{(1 + y_i)^{(\mu+1)}} \right) [1 - \mathbb{I}_{\{0\}}(y_i)]$$

$$- \mu [\Psi(\mu(\phi + 1))] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)]$$

$$+ [-\Psi(\phi + 2) + (\mu + 1) \Psi(\mu(\phi + 1) + \phi + 2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)].$$

Note que $\mu^* = [\Psi(\mu(\phi + 1)) - \Psi(\mu(\phi + 1) + \phi + 2)]$, logo temos que,

$$U_\phi(\mu, \phi) = \sum_{i=1}^n \log \left(\frac{y_i^\mu}{(1 + y_i)^{(\mu+1)}} \right) [1 - \mathbb{I}_{\{0\}}(y_i)] - \mu \mu^* \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)]$$

$$+ [-\Psi(\phi + 2) + \Psi(\mu(\phi + 1) + \phi + 2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)].$$

Agora tomando $\gamma = -\Psi(\phi + 2) + \Psi(\mu(\phi + 1) + \phi + 2)$ e fazendo alguns ajustes, temos a seguinte expressão:

$$U_\phi(\mu, \phi) = \sum_{i=1}^n \log \left(\frac{y_i^\mu}{(1 + y_i)^{(\mu+1)}} \right) [1 - \mathbb{I}_{\{0\}}(y_i)] - \mu \left(\mu^* - \frac{\gamma}{\mu} \right) \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)].$$

Assim segue que,

$$U_\phi(\mu, \phi) = \sum_{i=1}^n (y_i^\dagger - \mu^\dagger) [1 - \mathbb{I}_{\{0\}}(y_i)],$$

em que $y_i^\dagger = \log \left(\frac{y_i^\mu}{(1 + y_i)^{(\mu+1)}} \right)$ e $\mu^\dagger = \mu \left(\mu^* - \frac{\gamma}{\mu} \right)$.

O estimador de máxima verossimilhança, $\hat{\lambda}$, é obtido a partir da equação $U_\lambda(\lambda) = 0$. Assim segue que,

$$\begin{aligned} \frac{1}{\hat{\lambda}} \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) - \frac{1}{1 - \hat{\lambda}} \left[n - \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) \right] &= 0 \\ \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) \left[\frac{1}{\hat{\lambda}} + \frac{1}{1 - \hat{\lambda}} \right] &= \frac{n}{1 - \hat{\lambda}} \\ \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) &= \frac{\frac{n}{1 - \hat{\lambda}}}{\frac{1}{\hat{\lambda}} + \frac{1}{1 - \hat{\lambda}}} \\ \hat{\lambda} &= \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i). \end{aligned}$$

Como $\hat{\mu}$ e $\hat{\phi}$ não apresentam forma fechada, precisaremos recorrer a um algoritmo iterativo de otimização não-linear, para que seja possível maximizar numericamente a função de log-verossimilhança.

3.3 Matriz de Informação de Fisher BPIZ

Nesta seção obteremos a matriz de informação de Fisher (IF), relacionada à distribuição BPIZ. Ela será construída a partir dos valores esperados das derivadas de segunda ordem em torno da função escore. A IF poderá ser utilizada na obtenção de intervalos de confiança relacionados a distribuição BPIZ.

Denotaremos por: $M_{\lambda,\lambda}, M_{\lambda,\mu}, M_{\lambda,\phi}, M_{\mu,\lambda}, M_{\phi,\lambda}, M_{\mu,\mu}, M_{\mu,\phi}, M_{\phi,\mu}, M_{\phi,\phi}$, as derivadas segundas em relação a $\theta = (\lambda, \mu, \phi)$. A partir daí temos que,

$$\begin{aligned}
M_{\lambda\lambda} &= \frac{\partial^2 l_1(\lambda; y)}{\partial \lambda \partial \lambda} \\
&= -\frac{1}{\lambda^2} \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) - \frac{1}{(1-\lambda)^2} \left[n - \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) \right],
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l_1(\lambda; y)}{\partial \lambda \partial \mu} &= \frac{\partial^2 l_1(\lambda; y)}{\partial \lambda \partial \phi} = \frac{\partial^2 l_2(\mu, \phi; y)}{\partial \mu \partial \lambda} = \frac{\partial^2 l_2(\mu, \phi; y)}{\partial \phi \partial \lambda} = 0, \\
\Rightarrow M_{\lambda, \mu} &= M_{\lambda, \phi} = M_{\mu, \lambda} = M_{\phi, \lambda} = 0,
\end{aligned}$$

$$\begin{aligned}
M_{\mu, \mu} &= \frac{\partial^2 l_2(\mu, \phi; y)}{\partial \mu \partial \mu} \\
&= [-(\phi+1)^2 \Psi'(\mu(\phi+1)) + (\phi+1)^2 \Psi'(\mu(\phi+1) + \phi+2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \\
&= -(\phi+1)^2 [\Psi'(\mu(\phi+1)) - \Psi'(\mu(\phi+1) + \phi+2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)],
\end{aligned}$$

em que $\Psi'(\cdot)$ é a primeira derivada da função digama $\Psi(\cdot)$, denominada função trigama.

$$\begin{aligned}
M_{\phi, \phi} &= \frac{\partial^2 l_2(\mu, \phi; y)}{\partial \phi \partial \phi} \\
&= [-\mu^2 \Psi'(\mu(\phi+1)) - \Psi'(\phi+2) + (\mu+1)^2 \Psi'(\mu(\phi+1) + \phi+2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)],
\end{aligned}$$

$$\begin{aligned}
M_{\mu, \phi} &= \frac{\partial^2 l_2(\mu, \phi; y)}{\partial \mu \partial \phi} \\
&= \sum_{i=1}^n \log(y_i) [1 - \mathbb{I}_{\{0\}}(y_i)] - \sum_{i=1}^n \log(1+y_i) [1 - \mathbb{I}_{\{0\}}(y_i)] \\
&\quad + [-\Psi(\mu(\phi+1)) - (\phi+1)\mu \Psi'(\mu(\phi+1)) + \Psi(\mu(\phi+1) + \phi+2) \\
&\quad + (\phi+1)(\mu) \Psi'(\mu(\phi+1) + \phi+2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)].
\end{aligned}$$

Note que $M_{\mu, \phi} = M_{\phi, \mu}$.

$$\begin{aligned}
M_{\mu, \phi} &= M_{\phi, \mu} \\
&= \sum_{i=1}^n \log\left(\frac{y_i}{1+y_i}\right) [1 - \mathbb{I}_{\{0\}}(y_i)] \\
&\quad - [\Psi(\mu(\phi+1)) - \Psi(\mu(\phi+1) + \phi+2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] + \\
&\quad + [\mu(\phi+1) (\Psi'(\mu(\phi+1) + \phi+2) - \Psi'(\mu(\phi+1)))] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)].
\end{aligned}$$

Note que $\log\left(\frac{y_i}{1+y_i}\right) = y_i^*$ e $[\Psi(\mu(\phi+1)) - \Psi(\mu(\phi+1) + \phi + 2)] = \mu^*$. Então podemos substituir esses valores, obtendo a seguinte expressão:

$$M_{\mu,\phi} = M_{\phi,\mu} = \sum_{i=1}^n (y_i^* - \mu^*) [1 - \mathbb{I}_{\{0\}}(y_i)] \\ + \mu(\phi+1) [\Psi'(\mu(\phi+1) + \phi + 2) - \Psi'(\mu(\phi+1))] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)].$$

Aplicando o negativo do valor esperado, temos que:

$$-\mathbb{E}[M_{\lambda,\lambda}] = -\mathbb{E} \left\{ -\frac{1}{\lambda^2} \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) - \frac{1}{(1-\lambda)^2} \left[n - \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) \right] \right\}, \\ = \mathbb{E} \left[\frac{1}{\lambda^2} \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) \right] + \mathbb{E} \left[\frac{1}{(1-\lambda)^2} \left(n - \sum_{i=1}^n \mathbb{I}_{\{0\}}(y_i) \right) \right], \\ = \frac{n\lambda}{\lambda^2} + \frac{n}{(1-\lambda)^2} - \frac{\lambda n}{(1-\lambda)^2}, \\ = \frac{n - n\lambda + n\lambda}{\lambda(1-\lambda)}, \\ = \frac{n}{\lambda(1-\lambda)},$$

$$-\mathbb{E}[M_{\mu,\mu}] = -\mathbb{E} \left\{ -(\phi+1)^2 [\Psi'(\mu(\phi+1)) - \Psi'(\mu(\phi+1) + \phi + 2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right\} \\ = (\phi+1)^2 [\Psi'(\mu(\phi+1)) - \Psi'(\mu(\phi+1) + \phi + 2)] \mathbb{E} \left\{ \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right\} \\ = n(1-\lambda) \{ +(\phi+1)^2 [\Psi'(\mu(\phi+1)) - \Psi'(\mu(\phi+1) + \phi + 2)] \},$$

$$-\mathbb{E}[M_{\phi,\phi}] = -\mathbb{E} \{ [-\mu^2 \Psi'(\mu(\phi+1)) - \Psi'(\phi+2) + \\ + (\mu+1)^2 \Psi'(\mu(\phi+1) + \phi + 2)] \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \}, \\ = [-\mu^2 \Psi'(\mu(\phi+1)) - \Psi'(\phi+2) + (\mu+1)^2 \Psi'(\mu(\phi+1) + \phi + 2)] \times \\ \times \left(-\mathbb{E} \left[\sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right] \right), \\ = n(1-\lambda) [\mu^2 \Psi'(\mu(\phi+1)) + \Psi'(\phi+2) - (\mu+1)^2 \Psi'(\mu(\phi+1) + \phi + 2)],$$

Para o valor esperado de $M_{\mu,\phi}$ tomaremos, $(\phi+1)\Psi'(\mu(\phi+1) + \phi + 2) = k_1$ e

$\mu(\phi + 1) (\Psi'(\mu(\phi + 1) + \phi + 2) - \Psi'(\mu(\phi + 1))) = k_2$. Assim temos que,

$$\begin{aligned} -E[M_{\mu,\phi}] &= -E \left\{ \sum_{i=1}^n (y_i^* - \mu^*) [1 - \mathbb{I}_{\{0\}}(y_i)] + (k_1 + k_2) \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right\}, \\ &= -E \left[\sum_{i=1}^n (y_i^* - \mu^*) [1 - \mathbb{I}_{\{0\}}(y_i)] \right] - E \left[(k_1 + k_2) \sum_{i=1}^n [1 - \mathbb{I}_{\{0\}}(y_i)] \right]. \end{aligned}$$

Uma vez que o valor esperado da primeira derivada da função log-verossimilhança é igual a zero, tem que $E[y_i^*] = \mu^*$. Logo, isso implica que $\sum_{i=1}^n E[(y_i^* - \mu^*) [1 - \mathbb{I}_{\{0\}}(y_i)]] = 0$. Assim podemos calcular a esperança de $E[M_{\mu,\phi}]$ por:

$$-E[M_{\mu,\phi}] = -n[1 - \lambda](k_1 + k_2).$$

Daí temos que,

$$\begin{aligned} -E[M_{\mu,\phi}] &= -n[1 - \lambda] \{ (\phi + 1) \Psi'(\mu(\phi + 1) + \phi + 2) + \\ &\quad + \mu(\phi + 1) [\Psi'(\mu(\phi + 1) + \phi + 2) - \Psi'(\mu(\phi + 1))] \}, \end{aligned}$$

em que , $-E[M_{\mu,\phi}] = -E[M_{\phi,\mu}]$.

Após os cálculos acima, a matriz de informação de Fisher é dada por:

$$\text{IF}[\theta] = \begin{bmatrix} a_{\lambda\lambda} & 0 & 0 \\ 0 & a_{\mu\mu} & a_{\mu\phi} \\ 0 & a_{\phi\mu} & a_{\phi\phi} \end{bmatrix},$$

em que, $a_{\lambda\lambda} = -E[M_{\lambda,\lambda}]$, $a_{\mu\mu} = -E[M_{\mu,\mu}]$, $a_{\mu\phi} = -E[M_{\mu,\phi}]$, $a_{\phi\mu} = -E[M_{\phi,\mu}]$, e $a_{\phi\phi} = -E[M_{\phi,\phi}]$.

A obtenção da Matriz de IF nos permite obter importantes resultados de estimação intervalar, os quais serão apresentados na próxima seção e terão uma forte contribuição para a obtenção dos resultados desta pesquisa.

3.4 Intervalos de Confiança BPIZ

A partir da construção da matriz de informação de Fisher é possível obter resultados importantes da Inferência Estatística, como avaliação da probabilidade de cobertura, em estimação intervalar. A respeito dessa informação, o Teorema Central do Limite (TCL) nos diz que a média amostral de uma dada população

com média μ e variância σ^2 , desde que certas condições sejam satisfeitas, é assintoticamente normal, ou seja, $Z = \frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} \stackrel{A}{\sim} N(0, 1)$. Com base nisso podemos dizer que a distribuição assintótica dos estimadores de máxima verossimilhança da distribuição BPIZ, em grandes amostras, é normal, isto é $\hat{\theta} \stackrel{A}{\sim} N_3(\theta, \text{IF}[\theta]^{-1})$, em que $\hat{\theta} = (\hat{\lambda}, \hat{\mu}, \hat{\phi})$ é o EMV de θ , $\stackrel{A}{\sim}$ indica que a distribuição é assintoticamente normal, N_3 indica que a distribuição é normal trivariada e $\text{IF}[\theta]^{-1}$ é a inversa da matriz da informação de Fisher.

Assim, os ICs assintóticos para os parâmetros da distribuição BPIZ são descritos a seguir:

$$\left[\hat{\theta} - z_{1-\frac{\alpha}{2}} \sqrt{\text{IF}[\theta]^{-1}}; \hat{\theta} + z_{1-\frac{\alpha}{2}} \sqrt{\text{IF}[\theta]^{-1}} \right] \quad (3.5)$$

Os ICs assintóticos para λ , μ e ϕ , obtidos em (3.5), são dados, respectivamente, por: $\hat{\lambda} \pm z_{(1-\frac{\alpha}{2})} ep(\hat{\lambda})$, $\hat{\mu} \pm z_{(1-\frac{\alpha}{2})} ep(\hat{\mu})$ e $\hat{\phi} \pm z_{(1-\frac{\alpha}{2})} ep(\hat{\phi})$, em que ep é o erro padrão do estimador dado, onde: $ep(\hat{\lambda}) = \sqrt{a_{\lambda\lambda}}$; $ep(\hat{\mu}) = \sqrt{a_{\mu\mu}}$ e $ep(\hat{\phi}) = \sqrt{a_{\phi\phi}}$. Isso significa que o erro padrão relacionado a cada parâmetro é calculado a partir da raiz quadrada da variância estimada para cada estimador de máxima verossimilhança. Ainda, $z_{(1-\frac{\alpha}{2})}$ é o quantil $1 - \frac{\alpha}{2}$ da distribuição Normal padrão, e $1 - \alpha$ é o nível de confiança sugerido.

Neste capítulo propomos a distribuição beta prime inflacionada em zero, onde foi possível construir estimadores de máxima verossimilhança e intervalos de confiança. No capítulo seguinte desta pesquisa, será apresentada a avaliação numérica com base na distribuição BPIZ, a partir de estimação pontual e intervalos de confiança.

Capítulo 4

Avaliação Numérica

Neste capítulo apresentaremos os resultados obtidos a partir das simulações realizadas para analisar o desempenho da distribuição BPIZ, considerando estimação pontual e estimação intervalar. A seção 4.1 apresenta os resultados relacionados a estimação pontual, onde avaliamos numericamente os estimadores de máxima verossimilhança dos parâmetros da distribuição BPIZ. As medidas avaliadas foram: média, erro quadrático médio (EQM) e o viés relativo (VR), sendo o VR definido por $\frac{E[\hat{\theta} - \theta]}{\theta} \times 100$. A seção 4.2 traz os resultados da estimação intervalar, na qual foram avaliadas a probabilidade de cobertura e de não cobertura, tanto a esquerda quanto a direita, dos intervalos de confiança assintóticos, com nível de confiança de 90%, 95% e 99%.

Os resultados que serão apresentados neste capítulo são oriundos de simulações realizadas em linguagem de programação R, a partir do programa RStudio. Os algoritmos utilizados para a avaliação dos estimadores são descritos no apêndice deste trabalho.

4.1 Estimação Pontual

Nesta seção apresentamos os cenários avaliados a partir dos EMV da distribuição BPIZ. Para a média μ , consideramos dois valores, à saber, 0.9 e 10, já para o parâmetro de mistura λ atribuímos três valores distintos, a saber: 0.15, 0.25 e 0.35, lembrando que λ é a probabilidade da variável aleatória Y assumir o valor zero. Para ϕ , foram considerados dois valores de precisão: 35 e 70, respectivamente, (baixa e alta precisão). Para cada cenário, consideramos três tamanhos amostrais, a saber: 50, 100 e 200. A seguir apresentaremos as tabelas contendo os resultados das simulações de MC para a distribuição BPIZ, organizados em cenários da seguinte

forma:

- 1) A Tabela (4.1) contém os resultados da avaliação da distribuição BPIZ para a média, EQM e VR, com $\mu = 0.9$ e $\phi = 35$;
- 2) A Tabela (4.2) contém os resultados da avaliação da distribuição BPIZ para a média, EQM e VR, com $\mu = 10$ e $\phi = 35$;
- 3) A Tabela (4.3) contém os resultados da avaliação da distribuição BPIZ para a média, EQM e VR, com $\mu = 0.9$ e $\phi = 70$;
- 4) A Tabela (4.4) contém os resultados da avaliação da distribuição BPIZ para a média, EQM e VR, com $\mu = 10$ e $\phi = 70$.

As Tabelas 4.1, 4.2, 4.3 e 4.4 estão divididas em três cenários cada, a saber: o primeiro cenário contém resultados para $\lambda = 0.15$; o segundo cenário contém resultados para $\lambda = 0.25$ e o terceiro cenário contém resultados $\lambda = 0.35$. Nos cenários apresentados, é possível verificar que, conforme o tamanho da amostra n aumenta, as médias das estimativas de máxima verossimilhança convergem para os verdadeiros valores dos parâmetros, observa-se também o decaimento do EQM e do VR, conforme aumento do tamanho da amostra. Dessa forma, os resultados mostram um comportamento esperado. Nota-se ainda que, conforme o valor de λ aumenta, ou seja, conforme uma maior concentração de zeros, o EQM e o VR apresentam um decaimento um pouco mais lento, como pode ser observado nas tabelas. Vale ressaltar também que o VR é menor quando a precisão é alta.

Os resultados aqui apresentados mostram que os estimadores de máxima verossimilhança para a distribuição beta prime inflacionada em zero, apresentam resultados satisfatórios, isto é, apresentam um comportamento assintótico esperado. Assim conforme exposto, nos incentiva o uso da distribuição BPIZ em outras situações, como por exemplo, intervalos de confiança (IC) e testes de hipóteses.

Tabela 4.1: Cenários 1, 2 e 3, contendo os resultados da avaliação numérica da distribuição BPIZ para a Média, EQM e VR dos estimadores. Considerando $\mu = 0.9$ e $\phi = 35$, com $\lambda = (0.15, 0.25, 0.35)$.

Cenário	λ	Medida	Estimador	n		
				50	100	200
1	0.15	MÉDIA	$\hat{\lambda}$	0.1490	0.1497	0.1504
			$\hat{\mu}$	0.8998	0.9007	0.8998
			$\hat{\phi}$	37.9889	36.3508	35.6368
		EQM	$\hat{\lambda}$	0.0026	0.0013	0.0006
			$\hat{\mu}$	0.0011	0.0006	0.0003
			$\hat{\phi}$	92.0586	37.9058	17.3347
		VR	$\hat{\lambda}$	-0.6747	-0.2093	0.2720
			$\hat{\mu}$	-0.0193	0.0756	-0.0224
			$\hat{\phi}$	8.5398	3.8594	1.8195
2	0.25	MÉDIA	$\hat{\lambda}$	0.2481	0.2494	0.2505
			$\hat{\mu}$	0.8993	0.9007	0.8998
			$\hat{\phi}$	38.1890	36.5714	35.7314
		EQM	$\hat{\lambda}$	0.0037	0.0019	0.0009
			$\hat{\mu}$	0.0013	0.0006	0.0003
			$\hat{\phi}$	104.7085	43.5816	20.3472
		VR	$\hat{\lambda}$	-0.7408	-0.2280	0.1832
			$\hat{\mu}$	-0.0792	0.0808	-0.0241
			$\hat{\phi}$	9.1115	4.4897	2.0898
3	0.35	MÉDIA	$\hat{\lambda}$	0.3490	0.3486	0.3508
			$\hat{\mu}$	0.8995	0.9003	0.8999
			$\hat{\phi}$	38.8490	36.6697	35.8858
		EQM	$\hat{\lambda}$	0.0045	0.0023	0.0011
			$\hat{\mu}$	0.0015	0.0007	0.0004
			$\hat{\phi}$	131.4332	51.5252	23.2303
		VR	$\hat{\lambda}$	-0.2709	-0.3891	0.2297
			$\hat{\mu}$	-0.0604	0.0356	-0.0056
			$\hat{\phi}$	10.9970	4.7705	2.5309

Tabela 4.2: Cenários 4, 5 e 6, contendo os resultados da avaliação numérica da distribuição BPIZ para a Média, EQM e VR dos estimadores. Considerando $\mu = 10$ e $\phi = 35$, com $\lambda = (0.15, 0.25, 0.35)$.

Cenário	λ	Medida	Estimador	n		
				50	100	200
4	0.15	MÉDIA	$\hat{\lambda}$	0.1490	0.1497	0.1504
			$\hat{\mu}$	9.9986	10.0054	9.9983
			$\hat{\phi}$	37.9843	36.3498	35.6355
		EQM	$\hat{\lambda}$	0.0026	0.0013	0.0006
			$\hat{\mu}$	0.0725	0.0367	0.0185
			$\hat{\phi}$	91.8151	37.8194	17.2818
		VR	$\hat{\lambda}$	-0.6747	-0.2093	0.2720
			$\hat{\mu}$	-0.0136	0.0545	-0.0165
			$\hat{\phi}$	8.5266	3.8564	1.8157
5	0.25	MÉDIA	$\hat{\lambda}$	0.2481	0.2494	0.2505
			$\hat{\mu}$	9.9943	10.0058	9.9982
			$\hat{\phi}$	38.1830	36.5701	35.7300
		EQM	$\hat{\lambda}$	0.0037	0.0019	0.0009
			$\hat{\mu}$	0.0820	0.0415	0.0215
			$\hat{\phi}$	104.4170	43.4783	20.2924
		VR	$\hat{\lambda}$	-0.7408	-0.2280	0.1832
			$\hat{\mu}$	-0.0574	0.0584	-0.0178
			$\hat{\phi}$	9.0940	4.4861	2.0856
6	0.35	MÉDIA	$\hat{\lambda}$	0.3491	0.3486	0.3508
			$\hat{\mu}$	9.9956	10.0025	9.9996
			$\hat{\phi}$	38.8422	36.6674	35.8844
		EQM	$\hat{\lambda}$	0.0045	0.0023	0.0011
			$\hat{\mu}$	0.0944	0.0476	0.0246
			$\hat{\phi}$	131.0696	51.3961	23.1666
		VR	$\hat{\lambda}$	-0.2709	-0.3891	0.2297
			$\hat{\mu}$	-0.0437	0.0251	-0.0043
			$\hat{\phi}$	10.9777	4.7641	2.5268

Tabela 4.3: Cenários 7, 8 e 9, contendo os resultados da avaliação numérica da distribuição BPIZ para a Média, EQM e VR dos estimadores. $\mu = 0.9$ e $\phi = 70$, com $\lambda = (0.15, 0.25, 0.35)$.

Cenário	λ	Medida	Estimador	n		
				50	100	200
7	0.15	MÉDIA	$\hat{\lambda}$	0.1490	0.1497	0.1504
			$\hat{\mu}$	0.8999	0.9005	0.8998
			$\hat{\phi}$	75.8324	72.6452	71.2401
		EQM	$\hat{\lambda}$	0.0026	0.0013	0.0006
			$\hat{\mu}$	0.0006	0.0003	0.0001
			$\hat{\phi}$	349.9995	144.3088	65.7646
		VR	$\hat{\lambda}$	-0.6747	-0.2093	0.2720
			$\hat{\mu}$	-0.0129	0.0527	-0.0174
			$\hat{\phi}$	8.3320	3.7789	1.7716
8	0.25	MÉDIA	$\hat{\lambda}$	0.2481	0.2494	0.2505
			$\hat{\mu}$	0.8995	0.9005	0.8998
			$\hat{\phi}$	76.2137	73.0765	71.4251
		EQM	$\hat{\lambda}$	0.0037	0.0019	0.0009
			$\hat{\mu}$	0.0006	0.0003	0.0002
			$\hat{\phi}$	397.8773	165.8507	77.3068
		VR	$\hat{\lambda}$	-0.7408	-0.2280	0.1832
			$\hat{\mu}$	-0.0578	0.0571	-0.0190
			$\hat{\phi}$	8.8767	4.3950	2.0358
9	0.35	MÉDIA	$\hat{\lambda}$	0.3491	0.3486	0.3508
			$\hat{\mu}$	0.8996	0.9002	0.8999
			$\hat{\phi}$	77.5046	73.2622	71.7285
		EQM	$\hat{\lambda}$	0.0045	0.0023	0.0011
			$\hat{\mu}$	0.0007	0.0004	0.0002
			$\hat{\phi}$	499.5394	195.9792	88.2533
		VR	$\hat{\lambda}$	-0.2709	-0.3891	0.2297
			$\hat{\mu}$	-0.0434	0.0223	-0.0050
			$\hat{\phi}$	10.7208	4.6603	2.4692

Tabela 4.4: Cenários 10, 11 e 12, contendo os resultados da avaliação numérica da distribuição BPIZ para a Média, EQM e VR dos estimadores. $\mu = 10$ e $\phi = 70$, com $\lambda = (0.15, 0.25, 0.35)$.

Cenário	λ	Medida	Estimador	n		
				50	100	200
10	0.15	MÉDIA	$\hat{\lambda}$	0.1490	0.1497	0.1504
			$\hat{\mu}$	9.9991	10.0038	9.9987
			$\hat{\phi}$	75.8279	72.6447	71.2387
		EQM	$\hat{\lambda}$	0.0026	0.0013	0.0006
			$\hat{\mu}$	0.0363	0.0184	0.0092
			$\hat{\phi}$	349.5472	144.1614	65.6604
		VR	$\hat{\lambda}$	-0.6747	-0.2093	0.2720
			$\hat{\mu}$	-0.0091	0.0380	-0.0128
			$\hat{\phi}$	8.3256	3.7781	1.7696
11	0.25	MÉDIA	$\hat{\lambda}$	0.2482	0.2494	0.2505
			$\hat{\mu}$	9.9958	10.0041	9.9986
			$\hat{\phi}$	76.2072	73.0757	71.4235
		EQM	$\hat{\lambda}$	0.0037	0.0019	0.0009
			$\hat{\mu}$	0.0410	0.0208	0.0108
			$\hat{\phi}$	397.3217	165.6685	77.2032
		VR	$\hat{\lambda}$	-0.7408	-0.2280	0.1832
			$\hat{\mu}$	-0.0418	0.0412	-0.0139
			$\hat{\phi}$	8.8674	4.3938	2.0336
12	0.35	MÉDIA	$\hat{\lambda}$	0.3490	0.3486	0.3508
			$\hat{\mu}$	9.9969	10.0016	9.9996
			$\hat{\phi}$	77.4978	73.2602	71.7271
		EQM	$\hat{\lambda}$	0.0045	0.0023	0.0011
			$\hat{\mu}$	0.0473	0.0238	0.0123
			$\hat{\phi}$	498.8618	195.7448	88.1341
		VR	$\hat{\lambda}$	-0.2709	-0.3891	0.2297
			$\hat{\mu}$	-0.0314	0.0158	-0.0037
			$\hat{\phi}$	10.7112	4.6575	2.4672

4.2 Estimação Intervalar

Nesta seção apresentamos os resultados das simulações de MC da distribuição BPIZ para a avaliação da taxa(%) de cobertura e não cobertura dos IC assintóticos relacionados à distribuição. Para tal avaliação, consideramos níveis de confiança de 95% e de 99%, com média $\mu = 0.9$ e $\mu = 10$, precisão $\phi = 35$ e $\phi = 70$. Para o parâmetro de mistura λ consideramos os valores 0.15, 0.25 e 0.35. Consideramos também três tamanhos amostrais, a saber: 50, 100 e 200.

- 1) A Tabela (4.5) contém os resultados da avaliação da distribuição BPIZ para a taxa (%) de cobertura e não cobertura (à esquerda; à direita) dos IC, com $\mu = 0.9$ e $\phi = 35$;
- 2) A Tabela (4.6) contém os resultados da avaliação da distribuição BPIZ para a taxa (%) de cobertura e não cobertura (à esquerda; à direita) dos IC, com $\mu = 10$ e $\phi = 35$;
- 3) A Tabela (4.7) contém os resultados da avaliação da distribuição BPIZ para a taxa (%) de cobertura e não cobertura (à esquerda; à direita) dos IC, com $\mu = 0.9$ e $\phi = 70$;
- 4) A Tabela (4.5) contém os resultados da avaliação da distribuição BPIZ para a taxa (%) de cobertura e não cobertura (à esquerda; à direita) dos IC, com $\mu = 10$ e $\phi = 70$.

Nas Tabelas 4.5, 4.6, 4.7 e 4.8, para cada medida de $\lambda = (0.15, 0.25, 0.35)$, teremos dois níveis de confiança, a saber: 95% e 99%. As tabelas apresentam os resultados da simulação realizada para verificar se os intervalos de confiança contém os verdadeiros valores dos parâmetros relacionados. É possível verificar que a medida em que n aumenta, as probabilidades de cobertura convergem para os níveis de confiança adotados. Por exemplo, na Tabela 4.5, para um nível de confiança de 95%, $\lambda = 0.25$ e $n = 50$, as taxas de cobertura para λ , μ e ϕ são, respectivamente, 93.80, 94.24 e 95.82. Quando $n = 200$, as taxas de cobertura para λ , μ e ϕ são, respectivamente, 94.40, 94.36 e 94.86, logo é possível perceber, que de fato, existe a convergência dos valores para o nível de confiança adotado. As taxas de não cobertura, a medida que o tamanho amostral n aumenta, apresentam-se mais simétricas (como era esperado) já que os intervalos de confiança assintóticos são baseados na distribuição normal.

No apêndice desta pesquisa apresentaremos mais alguns resultados para as taxas(%) de cobertura e não cobertura dos IC assintóticos da distribuição BPIZ. Por

fim, com base nas simulações aqui apresentadas, no próximo capítulo abordaremos uma aplicação à dados reais da distribuição BPIZ, no intuito de mostrar sua aplicabilidade e também, tornar a distribuição BPIZ uma alternativa a outras distribuições que possuam o mesmo suporte.

Tabela 4.5: Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança dos parâmetros da distribuição beta prime inflacionada com $\mu = 0.9$ e $\phi = 35$, para $\lambda = (0.15, 0.25, 0.35)$.

λ	$1 - \delta$	$\hat{\theta}$	n		
			50	100	200
0.15	95%	$\hat{\lambda}$	93.86 (1.44; 4.70)	93.58 (1.02; 5.40)	94.60 (1.68; 3.72)
		$\hat{\mu}$	93.68 (2.20; 4.12)	94.56 (2.24; 3.20)	95.00 (2.02; 2.98)
		$\hat{\phi}$	95.64 (1.68; 2.68)	95.30 (2.00; 2.70)	95.30 (2.16; 2.54)
	99%	$\hat{\lambda}$	95.14 (0.16; 4.70)	97.00 (0.14; 2.86)	98.54 (0.24; 1.22)
		$\hat{\mu}$	98.34 (0.40; 1.26)	98.94 (0.32; 0.74)	98.70 (0.32; 0.98)
		$\hat{\phi}$	98.92 (0.16; 0.92)	99.12 (0.18; 0.70)	99.24 (0.22; 0.54)
0.25	95%	$\hat{\lambda}$	93.80 (1.32; 4.88)	94.34 (1.84; 3.82)	94.40 (2.04; 3.56)
		$\hat{\mu}$	94.24 (1.66; 4.10)	94.70 (2.00; 3.30)	94.36 (2.28; 3.36)
		$\hat{\phi}$	95.82 (1.36; 2.82)	95.64 (1.86; 2.50)	94.86 (2.22; 2.92)
	99%	$\hat{\lambda}$	97.68 (0.28; 2.04)	98.62 (0.28; 1.10)	98.78 (0.30; 0.92)
		$\hat{\mu}$	98.28 (0.34; 1.38)	98.78 (0.34; 0.88)	98.64 (0.36; 1.00)
		$\hat{\phi}$	99.18 (0.04; 0.78)	99.08 (0.16; 0.76)	98.90 (0.36; 0.74)
0.35	95%	$\hat{\lambda}$	94.96 (1.88; 3.16)	93.76 (2.42; 3.82)	95.22 (2.14; 2.64)
		$\hat{\mu}$	94.04 (1.96; 4.00)	94.56 (2.00; 3.44)	94.34 (2.08; 3.58)
		$\hat{\phi}$	96.04 (1.28; 2.68)	95.62 (1.76; 2.62)	95.44 (2.06; 2.50)
	99%	$\hat{\lambda}$	97.92 (0.50; 1.58)	98.82 (0.52; 0.66)	99.14 (0.30; 0.56)
		$\hat{\mu}$	98.26 (0.44; 1.30)	98.84 (0.30; 0.86)	98.94 (0.24; 0.82)
		$\hat{\phi}$	99.16 (0.02; 0.82)	99.10 (0.12; 0.78)	99.08 (0.20; 0.72)

Tabela 4.6: Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança dos parâmetros da distribuição beta prime inflacionada com $\mu = 10$ e $\phi = 35$, para $\lambda = (0.15, 0.25, 0.35)$.

λ	$1 - \delta$	$\hat{\theta}$	n		
			50	100	200
0.15	95%	$\hat{\lambda}$	93.86 (1.44; 4.70)	93.58 (1.02; 5.40)	94.60 (1.68; 3.72)
		$\hat{\mu}$	93.70 (2.28; 4.02)	94.62 (2.28; 3.10)	95.02 (2.04; 2.94)
		$\hat{\phi}$	95.64 (1.70; 2.66)	95.34 (1.96; 2.70)	95.30 (2.14; 2.56)
	99%	$\hat{\lambda}$	95.14 (0.16; 4.70)	97.00 (0.14; 2.86)	98.54 (0.24; 1.22)
		$\hat{\mu}$	98.36 (0.44; 1.20)	98.96 (0.32; 0.72)	98.70 (0.34; 0.96)
		$\hat{\phi}$	98.90 (0.16; 0.94)	99.14 (0.16; 0.70)	99.28 (0.20; 0.52)
0.25	95%	$\hat{\lambda}$	93.80 (1.32; 4.88)	94.34 (1.84; 3.82)	94.40 (2.04; 3.56)
		$\hat{\mu}$	94.14 (1.84; 4.02)	94.74 (2.04; 3.22)	94.38 (2.28; 3.34)
		$\hat{\phi}$	95.82 (1.36; 2.82)	95.66 (1.86; 2.48)	94.86 (2.22; 2.92)
	99%	$\hat{\lambda}$	97.68 (0.28; 2.04)	98.62 (0.28; 1.10)	98.78 (0.30; 0.92)
		$\hat{\mu}$	98.30 (0.34; 1.36)	98.78 (0.34; 0.88)	98.62 (0.38; 1.00)
		$\hat{\phi}$	99.18 (0.04; 0.78)	99.08 (0.16; 0.76)	98.90 (0.36; 0.74)
0.35	95%	$\hat{\lambda}$	94.96 (1.88; 3.16)	93.76 (2.42; 3.82)	95.22 (2.14; 2.64)
		$\hat{\mu}$	94.06 (2.02; 3.92)	94.52 (2.04; 3.44)	94.32 (2.16; 3.52)
		$\hat{\phi}$	96.06 (1.26; 2.68)	95.68 (1.76; 2.56)	95.40 (2.08; 2.52)
	99%	$\hat{\lambda}$	97.92 (0.50; 1.58)	98.82 (0.52; 0.66)	99.14 (0.30; 0.56)
		$\hat{\mu}$	98.26 (0.48; 1.26)	98.82 (0.34; 0.84)	98.94 (0.24; 0.82)
		$\hat{\phi}$	99.14 (0.02; 0.84)	99.10 (0.12; 0.78)	99.10 (0.18; 0.72)

Tabela 4.7: Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança dos parâmetros da distribuição beta prime inflacionada com $\mu = 0.9$ e $\phi = 70$, para $\lambda = (0.15, 0.25, 0.35)$.

λ	$1 - \delta$	$\hat{\theta}$	n		
			50	100	200
0.15	95%	$\hat{\lambda}$	93.86 (1.44; 4.70)	93.58 (1.02; 5.40)	94.60 (1.68; 3.72)
		$\hat{\mu}$	93.86 (2.50; 3.64)	94.60 (2.46; 2.94)	95.16 (2.06; 2.78)
		$\hat{\phi}$	95.64 (1.68; 2.68)	95.38 (1.94; 2.68)	95.30 (2.1; 2.60)
	99%	$\hat{\lambda}$	95.14 (0.16; 4.70)	97.00 (0.14; 2.86)	98.54 (0.24; 1.22)
		$\hat{\mu}$	98.36 (0.58; 1.06)	98.92 (0.42; 0.66)	98.74 (0.38; 0.88)
		$\hat{\phi}$	98.92 (0.16; 0.92)	99.12 (0.16; 0.72)	99.28 (0.2; 0.52)
0.25	95%	$\hat{\lambda}$	93.80 (1.32; 4.88)	94.34 (1.84; 3.82)	94.40 (2.04; 3.56)
		$\hat{\mu}$	94.20 (1.94; 3.86)	94.66 (2.22; 3.12)	94.46 (2.46; 3.08)
		$\hat{\phi}$	95.92 (1.28; 2.80)	95.62 (1.88; 2.50)	94.80 (2.22; 2.98)
	99%	$\hat{\lambda}$	97.68 (0.28; 2.04)	98.62 (0.28; 1.10)	98.78 (0.30; 0.92)
		$\hat{\mu}$	98.42 (0.38; 1.20)	98.86 (0.42; 0.72)	98.70 (0.42; 0.88)
		$\hat{\phi}$	99.16 (0.04; 0.80)	99.08 (0.16; 0.76)	98.90 (0.38; 0.72)
0.35	95%	$\hat{\lambda}$	94.96 (1.88; 3.16)	93.76 (2.42; 3.82)	95.22 (2.14; 2.64)
		$\hat{\mu}$	94.16 (2.22; 3.62)	94.50 (2.24; 3.26)	94.24 (2.32; 3.44)
		$\hat{\phi}$	96.02 (1.26; 2.72)	95.64 (1.80; 2.56)	95.42 (2.02; 2.56)
	99%	$\hat{\lambda}$	97.92 (0.50; 1.58)	98.82 (0.52; 0.66)	99.14 (0.30; 0.56)
		$\hat{\mu}$	98.26 (0.54; 1.20)	98.84 (0.42; 0.74)	98.98 (0.24; 0.78)
		$\hat{\phi}$	99.12 (0.02; 0.86)	99.10 (0.12; 0.78)	99.08 (0.20; 0.72)

Tabela 4.8: Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança dos parâmetros da distribuição beta prime inflacionada com $\mu = 10$ e $\phi = 70$, para $\lambda = (0.15, 0.25, 0.35)$.

λ	$1 - \delta$	$\hat{\theta}$	n		
			50	100	200
0.15	95%	$\hat{\lambda}$	93.86 (1.44; 4.70)	93.58 (1.02; 5.40)	94.60 (1.68; 3.72)
		$\hat{\mu}$	93.88 (2.54; 3.58)	94.52 (2.56; 2.92)	95.14 (2.08; 2.78)
		$\hat{\phi}$	95.60 (1.70; 2.70)	(95.42) (1.94; 2.64)	95.30 (2.10; 2.60)
	99%	$\hat{\lambda}$	95.14 (0.16; 4.70)	97.00 (0.14; 2.86)	98.54 (0.24; 1.22)
		$\hat{\mu}$	98.36 (0.58; 1.06)	98.90 (0.42; 0.68)	98.74 (0.38; 0.88)
		$\hat{\phi}$	98.92 (0.16; 0.92)	(99.10) (0.16; 0.74)	99.28 (0.20; 0.52)
0.25	95%	$\hat{\lambda}$	93.80 (1.32; 4.88)	94.34 (1.84; 3.82)	94.40 (2.04; 3.56)
		$\hat{\mu}$	94.26 (1.92; 3.82)	94.66 (2.24; 3.10)	94.48 (2.46; 3.06)
		$\hat{\phi}$	95.96 (1.26; 2.78)	95.64 (1.88; 2.48)	94.80 (2.24; 2.96)
	99%	$\hat{\lambda}$	97.68 (0.28; 2.04)	98.62 (0.28; 1.10)	98.78 (0.30; 0.92)
		$\hat{\mu}$	98.38 (0.42; 1.20)	98.86 (0.42; 0.72)	98.74 (0.42; 0.84)
		$\hat{\phi}$	99.18 (0.04; 0.78)	99.08 (0.16; 0.76)	98.90 (0.38; 0.72)
0.35	95%	$\hat{\lambda}$	94.96 (1.88; 3.16)	93.76 (2.42; 3.82)	95.22 (2.14; 2.64)
		$\hat{\mu}$	94.14 (2.30; 3.56)	94.52 (2.26; 3.22)	94.28 (2.32; 3.40)
		$\hat{\phi}$	96.02 (1.26; 2.72)	95.64 (1.80; 2.56)	95.40 (2.02; 2.58)
	99%	$\hat{\lambda}$	97.92 (0.50; 1.58)	98.82 (0.52; 0.66)	99.14 (0.30; 0.56)
		$\hat{\mu}$	98.24 (0.56; 1.20)	98.84 (0.42; 0.74)	98.96 (0.28; 0.76)
		$\hat{\phi}$	99.10 (0.02; 0.88)	99.10 (0.12; 0.78)	99.10 (0.20; 0.70)

Capítulo 5

Aplicação

Nos capítulos anteriores mostramos a importância de inflacionarmos a distribuição BP, isto é, criarmos um modelo capaz de mensurar variáveis aleatórias que assumam o valor zero. Vimos também sua eficiência no capítulo 4, quando avaliamos numericamente a distribuição BPIZ. Agora, aplicaremos o modelo BPIZ a dados reais a fim de mostrar sua importância como alternativa a distribuições que não possam mensurar dados que contenham zeros, ou mesmo outras distribuições inflacionadas cujo suporte seja os reais positivos. Além disso, será feita uma comparação com a distribuição Birnbaum-Saunders Reparametrizada com Zeros Ajustados (ZARBS) que apresenta suporte nos reais positivos \mathbb{R}^* (LEIVA *et al.*, 2016).

Uma variável aleatória que segue uma distribuição ZARBS, com parâmetros p , μ e ϕ , tem densidade dada por:

$$\text{zarbs}(y; p, \mu, \phi) = \begin{cases} p, & \text{se } y = 0, \\ (1 - p)f(y; \mu, \phi), & \text{se, } y > 0, \end{cases}$$

em $0 < p < 1$ é o parâmetro de mistura e $f(y; \mu, \phi)$ é a *fdp* da distribuição Birnbaum-Saunders Reparametrizada, com μ (média) e ϕ (parâmetro de precisão) (SANTOS-NETO *et al.*, 2012).

Estimaremos os parâmetros da distribuição BPIZ ajustando as funções presentes no pacote `santosneto/BPmodel` do R, apresentado por BOURGUIGNON *et al.* (2018). Para estimar os parâmetros da distribuição ZARBS, utilizaremos as funções presentes na biblioteca do R, `santosneto/RBS` (SANTOS-NETO *et al.*, 2016).

5.1 Aplicação da Distribuição BPIZ

O banco de dados utilizado para a aplicação da distribuição BPIZ foi apresentado por FOEKENS *et al.* (2000) e contém informações acerca do câncer de mama primário. Esse banco contém 2982 observações de pacientes diagnosticados com câncer de mama primário e está registrado no Banco de tumores de Rotterdam (Royston e Altman, 2013). A aplicabilidade do conjunto de dados foi feita sobre as informações apresentadas para os *Receptores de Estrogênio* (ER), que são proteínas capazes de se ligar a certas substâncias sanguíneas. O nível de concentração (*fmol/l*) dessa proteína pode indicar a presença de células cancerígenas. Porém, pacientes com concentração zero dessa proteína podem ter câncer de mama e não serem diagnosticados. A seguir descrevemos os dados e apresentamos os resultados da aplicação da distribuição BPIZ.

Nesse conjunto de dados é possível verificar que cerca de 13% das observações são de pacientes que apresentaram concentração zero da proteína ER e foram diagnosticados com câncer de mama. Pelo fato de conter zeros, uma distribuição comum não poderia mensurar todos esses dados, por isso a importância das distribuições inflacionadas. Aqui aplicamos o modelo objeto de nosso estudo, a distribuição BPIZ. A seguir, na Tabela (5.1) apresentamos um quadro resumido das medidas descritivas da quantidade de proteína ER presente na corrente sanguínea dos pacientes avaliados. É possível observar que o valor mínimo da variável de interesse é zero e que 75% das observações estão abaixo de 0.2028.

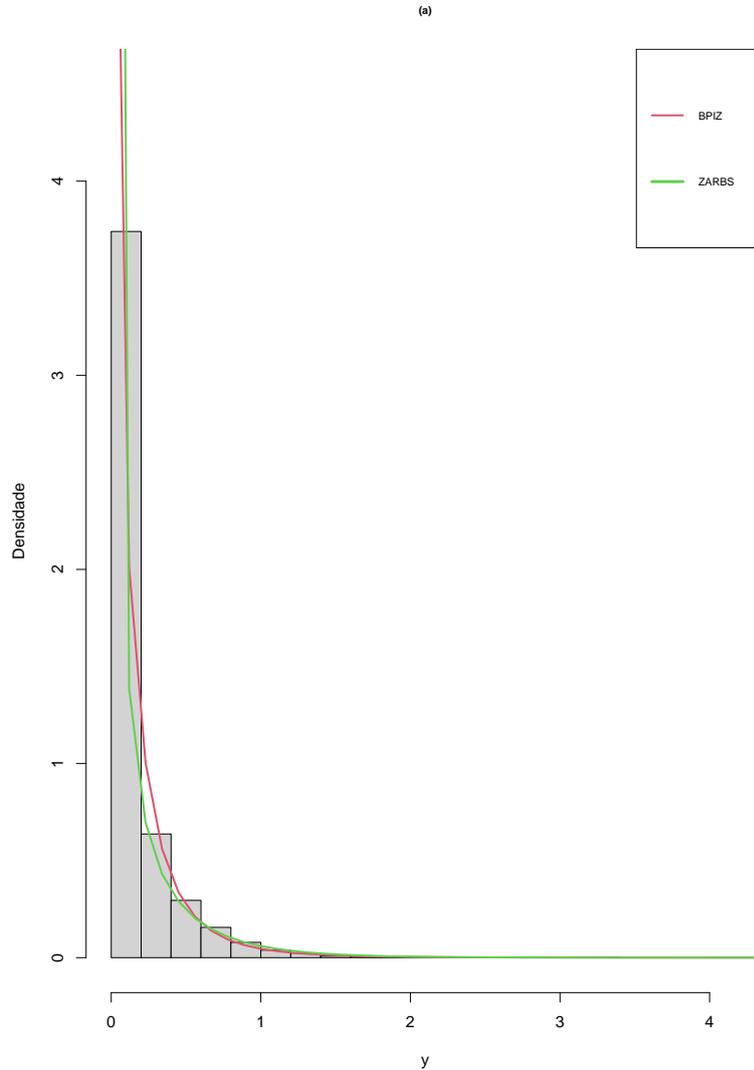
Tabela 5.1: Medidas descritivas da quantidade de proteína ER presente na corrente sanguínea dos pacientes avaliados.

Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0.000	0.0110	0.0610	0.1666	0.2028	3.2750

As estimativas de máxima verossimilhança dos parâmetros da distribuição BPIZ são: $\hat{\lambda} = 0.1317$; $\hat{\mu} = 0.1909$ e $\hat{\phi} = 2.7705$. Para a distribuição ZARBS, encontramos estimativas para $\hat{\lambda} = 0.1317$; $\hat{\mu} = 0.1781$ e $\hat{\phi} = 0.4584$.

Na Figura (5.1) apresentamos o histograma da distribuição dos dados em conjunto com as distribuições BPIZ e ZARBS. A imagem abaixo indica que a densidade possui forma de “J” invertido, indicado que há uma concentração maior de valores abaixo de 1. É possível verificar que dados com esse comportamento são modelados facilmente pela distribuição BPIZ.

Figura 5.1: Histograma e densidades estimadas das 2982 observações da concentração de ER de pacientes diagnosticados com câncer de mama primário.



Na Figura (5.2) apresentamos a *fda* empírica e a *fda* das distribuições BPIZ e ZARBS. No gráfico é possível observar que a distribuição BPIZ se ajusta melhor aos dados. Na Figura (5.3) apresentamos os gráficos quantil-quantil (Q-Q plot) para as distribuições BPIZ (5.3a) e ZARBS (5.3b). Nos referidos gráficos Q-Q plot ratificamos que a distribuição BPIZ se ajusta melhor aos dados, haja vista que os pontos se ajustam melhor a diagonal apresentada em (5.3a), indicando uma melhor adequação da distribuição BPIZ aos dados.

Figura 5.2: Funções de distribuição acumuladas para as 2982 observações da concentração de ER de pacientes diagnosticados com câncer de mama primário.

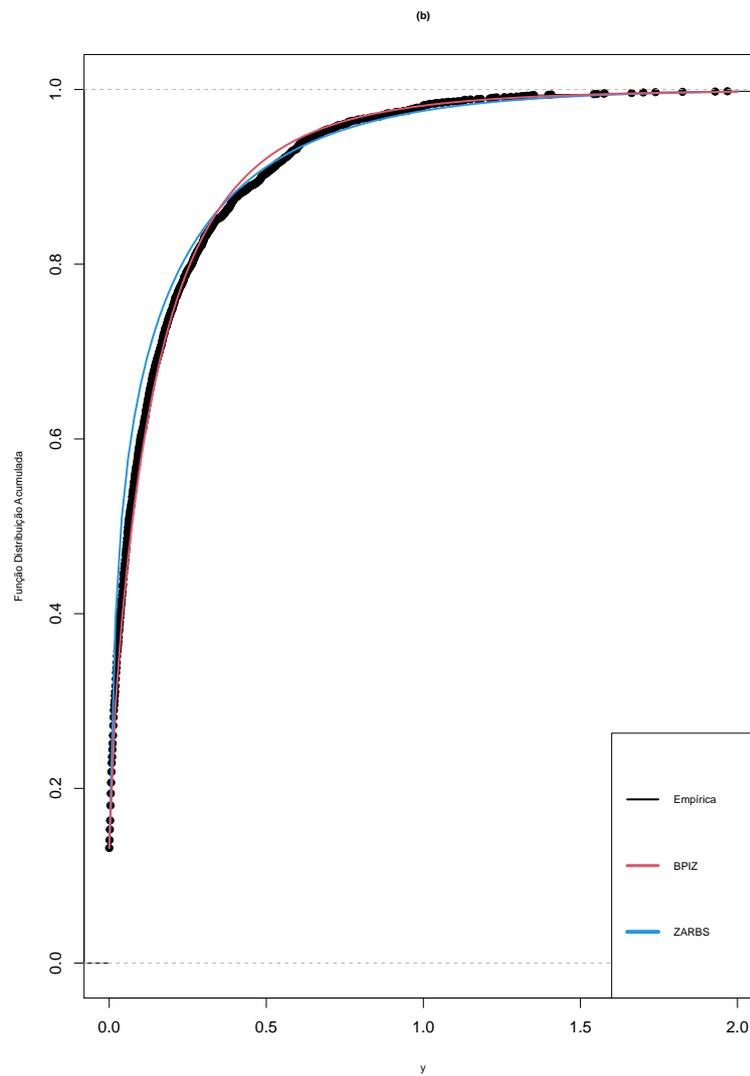
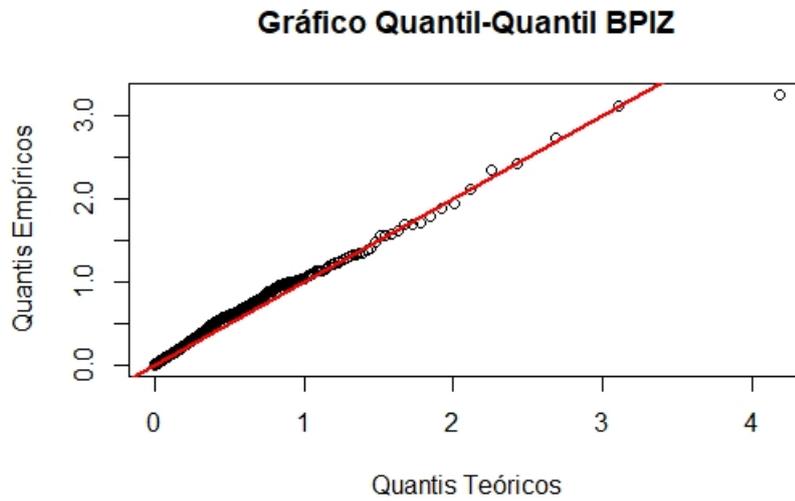
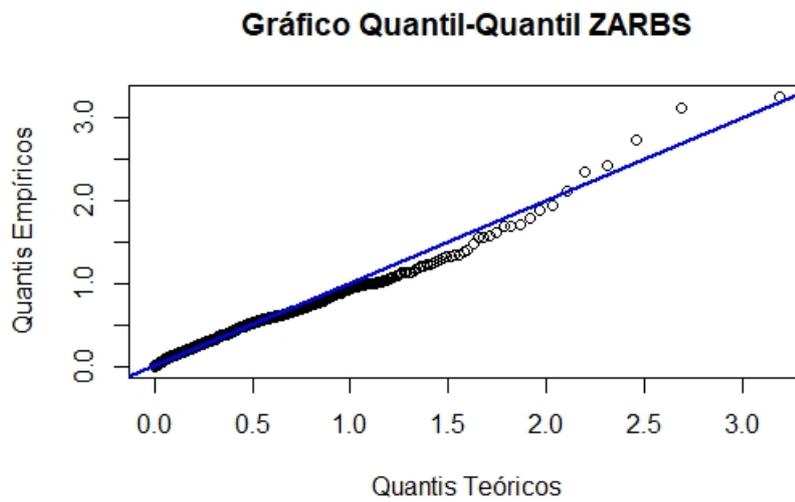


Figura 5.3: Gráfico Quantil-Quantil (Q-Q Plot) para as 2982 observações da concentração de ER de pacientes diagnosticados com câncer de mama primário.

(a) Q-Q Plot da Distribuição BPIZ.



(b) Q-Q Plot da Distribuição ZARBS.



Capítulo 6

Conclusões

Algumas distribuições de probabilidade, sejam elas discretas ou contínuas, não são capazes de contemplar dados que contenham zeros. A alternativa mais prudente é construir um modelo de distribuição misto capaz de considerar essas informações. Mostramos aqui que esse modelo misto é composto por uma parte discreta e uma parte contínua, os zeros serão acomodados por essa parte discreta e os demais dados serão de competência da parte contínua. Aqui construímos um modelo misto para a distribuição beta prime, nosso principal objeto de estudo.

A nossa pesquisa mostrou que é possível sim inflacionar a distribuição BP e evitar problemas na hora de mensurar dados que contenham zeros. Desenvolvemos um alternativa, por exemplo para a pesquisa apresentada por [TULUPYEV *et al.* \(2013\)](#), a qual apresenta a utilização da distribuição BP, porém com a modificação dos zeros por um outro número real positivo para que a distribuição fosse capaz de aceitar os valores apresentados. A essa alternativa denominamos distribuição beta prime inflacionada em zero (BPIZ). Foi necessário verificar a eficiência da distribuição proposta, então construímos EMV e IC para o nosso modelo e realizamos avaliação numérica, onde constatamos a eficiência da distribuição, como mostrado no capítulo 4.

Construído todo o embasamento da distribuição BPIZ foi necessário mostrar sua aplicabilidade à dados reais e conseguimos verificar que de fato é possível utilizar o modelo aqui por nós proposto, mostramos ainda que ele é eficiente em comparação distribuição Birnbaum-Saunders reparametrizada com zeros ajustados. Para concluir esta pesquisa, podemos afirmar que a distribuição BPIZ é uma alternativa viável para mensurar variáveis aleatórias que assumam o valor zero e tenham suporte nos reais positivos.

Referências Bibliográficas

- ABERNATHY, R. W., SMITH, R. P., 1993, “Applying series expansion to the inverse Beta distribution to find percentiles of the F-distribution”, *ACM Transactions on Mathematical Software (TOMS)*, v. 19, n. 4, pp. 474–480.
- BARTOSZYNSKI, R., NIEWIADOMSKA-BUGAJ, M., 2007, *Probability and statistical inference*, v. 549. John Wiley & Sons.
- BEKKER, A., ROUX, J., PHAM-GIA, T., 2009, “The type I distribution of the ratio of independent “Weibullized” generalized beta-prime variables”, *Statistical Papers*, v. 50, n. 2, pp. 323–338.
- BOLFARINE, H., SANDOVAL, M. C., 2001, *Introdução à inferência estatística*, v. 2. SBM.
- BOURGUIGNON, M., SANTOS-NETO, M., CASTRO, M., 2018, “A New Regression Model For Positive Data”, *arXiv e-prints*.
- CASELLA, G., BERGER, R. L., 2010, *Inferência estatística. 2a Edição*. São Paulo: Editora Cengage Learning. 588p.
- CRIBARI-NETO, F., SANTOS, J., 2019, “Inflated Kumaraswamy distributions”, *Anais da Academia Brasileira de Ciências*, v. 91, n. 2.
- DEGROOT, MORRIS H., S. M. J., 2012, *Probability and Statistics*. Pearson Education, Inc.
- FISHER, R. A., 1922a, “The goodness of fit of regression formulae, and the distribution of regression coefficients”, *Journal of the Royal Statistical Society*, v. 85, n. 4, pp. 597–612.
- FISHER, R. A., 1922b, “On the mathematical foundations of theoretical statistics”, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, v. 222, n. 594-604, pp. 309–368.

- FOEKENS, J. A., PETERS, H. A., LOOK, M. P., et al., 2000, “The urokinase system of plasminogen activation and prognosis in 2780 breast cancer patients”, *Cancer research*, v. 60, n. 3, pp. 636–643.
- GAUSS, C. F., 1809, *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, v. 7. Perthes et Besser.
- HALD, A., OTHERS, 1999, “On the history of maximum likelihood in relation to inverse probability and least squares”, *Statistical Science*, v. 14, n. 2, pp. 214–222.
- JOHNSON, N. L., KOTZ, S., 1969. “Distributions in statistics–Discrete distributions”. .
- LEIVA, V., SANTOS-NETO, M., CYSNEIROS, F. J. A., et al., 2016, “A methodology for stochastic inventory models based on a zero-adjusted Birnbaum-Saunders distribution”, *Applied Stochastic Models in Business and Industry*, v. 32, n. 1, pp. 74–89.
- LORD, D., WASHINGTON, S. P., IVAN, J. N., 2005, “Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory”, *Accident Analysis & Prevention*, v. 37, n. 1, pp. 35–46.
- MCDONALD, J. B., XU, Y. J., 1995, “A generalization of the beta distribution with applications”, *Journal of Econometrics*, v. 66, n. 1-2, pp. 133–152.
- MOOD, A. M., 1950, “Introduction to the Theory of Statistics.” .
- MORETTIN, L. G., 2010, *Estatística básica: probabilidade e inferência: volume único*. Pearson Prentice Hall.
- NOBRE, A. A., CARVALHO, M. S., GRIEP, R. H., et al., 2017, “Multinomial model and zero-inflated gamma model to study time spent on leisure time physical activity: an example of ELSA-Brasil”, *Revista de saude publica*, v. 51, pp. 76.
- OSPINA, R., FERRARI, S. L., 2010, “Inflated beta distributions”, *Statistical papers*, v. 51, n. 1, pp. 111.
- OSPINA, R., FERRARI, S. L., 2012, “A general class of zero-or-one inflated beta regression models”, *Computational Statistics & Data Analysis*, v. 56, n. 6, pp. 1609–1623.

- R DEVELOPMENT CORE TEAM, 2021, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<http://www.R-project.org>>. ISBN 3-900051-07-0.
- RODRIGUES, J. S., 2016, “Análise de diagnóstico em modelos de regressão ZAGA e ZAIG”, .
- ROYSTON, P., ALTMAN, D. G., 2013, “External validation of a Cox prognostic model: principles and methods”, *BMC medical research methodology*, v. 13, n. 1, pp. 1–15.
- SAHOO, P., 2013, “Probability and mathematical statistics”, *University of Louisville*.
- SANTOS-NETO, M., CYSNEIROS, F. J. A., LEIVA, V., et al., 2012, “On new parameterizations of the Birnbaum-Saunders distribution”, *Pakistan Journal of Statistics*, v. 28, n. 1, pp. 1–26.
- SANTOS-NETO, M., CYSNEIROS, F. J. A., LEIVA, V., et al., 2016, “Reparameterized Birnbaum-Saunders regression models with varying precision”, *Electronic Journal of Statistics*, v. 10, n. 2, pp. 2825–2855.
- SELLERS, K. F., RAIM, A., 2016, “A flexible zero-inflated model to address data dispersion”, *Computational Statistics & Data Analysis*, v. 99, pp. 68–80.
- TOMAZELLA, V., PEREIRA, G. H., NOBRE, J. S., et al., 2019, “Zero-adjusted reparameterized Birnbaum–Saunders regression model”, *Statistics & Probability Letters*, v. 149, pp. 142–145.
- TONG, E. N., MUES, C., THOMAS, L., 2013, “A zero-adjusted gamma model for mortgage loan loss given default”, *International Journal of Forecasting*, v. 29, n. 4, pp. 548–562.
- TULUPYEV, A., SUVOROVA, A., SOUSA, J., et al., 2013, “Beta prime regression with application to risky behavior frequency screening”, *Statistics in medicine*, v. 32, n. 23, pp. 4044–4056.
- WASSERMAN, L., 2013, *All of statistics: a concise course in statistical inference*. Springer Science & Business Media.

Apêndice A

Algumas Taxas de Cobertura

No capítulo 4 apresentamos simulações de Monte Carlo para avaliar os EMVs da distribuição BPIZ, a fim de mostrar sua eficiência. Mostramos que o estimadores possuem um bom comportamento tanto para a estimação pontual quanto para estimação por intervalos. Aqui apresentaremos algumas taxas (%) de cobertura como complemento às apresentadas na seção 4.2 do referido capítulo.

Tabela A.1: Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança, à nível de confiança de 90%, dos parâmetros da distribuição beta prime inflacionada com $\mu = 0.9$ e $\phi = 35$, para $\lambda = (0.15, 0.25, 0.35)$.

λ	$1 - \delta$	$\hat{\theta}$	n		
			50	100	200
0.15	90%	$\hat{\lambda}$	84.92 (3.14; 11.94)	86.76 (3.54; 9.70)	90.40 (3.46; 6.14)
		$\hat{\mu}$	89.16 (4.30; 6.54)	89.98 (4.40; 5.62)	89.72 (4.42; 5.86)
		$\hat{\phi}$	91.70 (3.98; 4.32)	90.06 (4.80; 5.14)	90.06 (4.86; 5.08)
		$\hat{\lambda}$	87.62 (2.44; 9.94)	88.88 (4.62; 6.50)	90.70 (4.24; 5.06)
		$\hat{\mu}$	89.34 (3.60; 7.06)	89.20 (4.76; 6.04)	88.82 (4.84; 6.34)
		$\hat{\phi}$	91.68 (3.92; 4.40)	90.86 (4.38; 4.76)	90.40 (4.64; 4.96)
0.25	90%	$\hat{\lambda}$	89.76 (3.82; 6.42)	90.26 (3.84; 5.90)	90.28 (4.52; 5.20)
		$\hat{\mu}$	89.20 (4.16; 6.64)	89.66 (4.36; 5.98)	88.44 (5.00; 6.56)
		$\hat{\phi}$	91.42 (3.92; 4.66)	90.92 (4.52; 4.56)	90.78 (4.76; 4.46)

Tabela A.2: Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança, à nível de confiança de 90%, dos parâmetros da distribuição beta prime inflacionada com $\mu = 10$ e $\phi = 35$, para $\lambda = (0.15, 0.25, 0.35)$.

λ	$1 - \delta$	$\hat{\theta}$	n		
			50	100	200
0.15	90%	$\hat{\lambda}$	84.92 (3.14; 11.94)	86.76 (3.54; 9.70)	90.40 (3.46; 6.14)
		$\hat{\mu}$	89.12 (4.36; 6.52)	89.98 (4.44; 5.58)	89.66 (4.52; 5.82)
		$\hat{\phi}$	91.68 (3.98; 4.34)	90.04 (4.80; 5.16)	90.06 (4.90; 5.04)
0.25	90%	$\hat{\lambda}$	87.62 (2.44; 9.94)	88.88 (4.62; 6.50)	90.70 (4.24; 5.06)
		$\hat{\mu}$	89.26 (3.70; 7.04)	89.28 (4.80; 5.92)	89.00 (4.80; 6.20)
		$\hat{\phi}$	91.64 (3.94; 4.42)	90.82 (4.42; 4.76)	90.30 (4.70; 5.00)
0.35	90%	$\hat{\lambda}$	89.76 (3.82; 6.42)	90.26 (3.84; 5.90)	90.28 (4.52; 5.20)
		$\hat{\mu}$	89.22 (4.26; 6.52)	89.66 (4.42; 5.92)	88.42 (5.08; 6.50)
		$\hat{\phi}$	91.38 (3.98; 4.64)	90.94 (4.54; 4.52)	90.74 (4.80; 4.46)

Tabela A.3: Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança, à nível de confiança de 90%, dos parâmetros da distribuição beta prime inflacionada com $\mu = 0.9$ e $\phi = 70$, para $\lambda = (0.15, 0.25, 0.35)$.

λ	$1 - \delta$	$\hat{\theta}$	n		
			50	100	200
0.15	90%	$\hat{\lambda}$	84.92 (3.14; 11.94)	86.76 (3.54; 9.70)	90.40 (3.46; 6.14)
		$\hat{\mu}$	89.10 (4.52; 6.38)	89.64 (4.84; 5.52)	89.78 (4.64; 5.58)
		$\hat{\phi}$	91.60 (4.02; 4.38)	90.28 (4.68; 5.04)	90.10 (4.9; 5.00)
0.25	90%	$\hat{\lambda}$	87.62 (2.44; 9.94)	88.88 (4.62; 6.50)	90.70 (4.24; 5.06)
		$\hat{\mu}$	89.26 (4.02; 6.72)	89.42 (4.86; 5.72)	89.24 (4.78; 5.98)
		$\hat{\phi}$	91.60 (3.94; 4.46)	90.88 (4.40; 4.72)	90.22 (4.72; 5.06)
0.35	90%	$\hat{\lambda}$	89.76 (3.82; 6.42)	90.26 (3.84; 5.90)	90.28 (4.52; 5.20)
		$\hat{\mu}$	89.18 (4.54; 6.28)	89.68 (4.56; 5.76)	88.48 (5.24; 6.28)
		$\hat{\phi}$	91.32 (3.96; 4.72)	90.90 (4.54; 4.56)	90.60 (4.96; 4.44)

Tabela A.4: Taxas de cobertura (%) e de não cobertura (à esquerda; à direita) dos intervalos de confiança, à nível de confiança de 90%, dos parâmetros da distribuição beta prime inflacionada com $\mu = 10$ e $\phi = 70$, para $\lambda = (0.15, 0.25, 0.35)$.

λ	$1 - \delta$	$\hat{\theta}$	n		
			50	100	200
0.15	90%	$\hat{\lambda}$	84.92 (3.14; 11.94)	86.76 (3.54; 9.70)	90.40 (3.46; 6.14)
		$\hat{\mu}$	89.08 (4.54; 6.38)	89.62 (4.90; 5.48)	89.74 (4.72; 5.54)
		$\hat{\phi}$	91.60 (4.00; 4.40)	(90.22) (4.72; 5.06)	90.06 (4.92; 5.02)
		$\hat{\lambda}$	87.62 (2.44; 9.94)	88.88 (4.62; 6.50)	90.70 (4.24; 5.06)
		$\hat{\mu}$	89.26 (4.08; 6.66)	89.36 (4.94; 5.70)	89.26 (4.78; 5.96)
		$\hat{\phi}$	91.66 (3.92; 4.42)	90.80 (4.42; 4.78)	90.26 (4.68; 5.06)
0.35	90%	$\hat{\lambda}$	89.76 (3.82; 6.42)	90.26 (3.84; 5.90)	90.28 (4.52; 5.20)
		$\hat{\mu}$	89.18 (4.60; 6.22)	89.70 (4.54; 5.76)	88.50 (5.28; 6.22)
		$\hat{\phi}$	91.30 (3.96; 4.74)	90.90 (4.54; 4.56)	90.62 (4.96; 4.42)

Apêndice B

Algoritmos

Aqui são apresentados os algoritmos utilizados para realizar as simulações do capítulo 4.

Script da Simulação Monte Carlo da distribuição BPIZ

```
rm(list = ls())
library(gamlss.dist)
library(gamlss.inf)
library(LaplacesDemon)
library(Formula)
library(ggplot2)
library(extraDistr)
library(gamlss)
library(rootSolve)
library(VGAM)
library(devtools)
library(ggplot2)
Sys.setenv("R_REMOTES_NO_ERRORS_FROM_WARNINGS"=T)
devtools::install_github('santosneto/BPmodel')
library(BPmodel)
library(gamlss)
library(RBS)
source("gamlss_BP.R")
```

R=5000

```

set.seed(1000)
vn=c(50,100, 200)
mu=0.9
sigma=70
xi0=0.15
n=100 # se o for para o n der problema

cont_zero=0

results = c()
results1 = c()

#valores tabelados da distribuicao normal

vt10 = qnorm(0.95)
vt10

vt5 = qnorm(0.975)
vt5

vt1 = qnorm(0.995)
vt1

pa=c(xi0,mu,sigma)
for(n in vn){
coef_results<-c()
ep_results<-c()
results<-c()
nc=0
ni=0
i=0
gen.Zadj(family="BP")
while(i<R)
{
yt<-rBPZadj(n,mu=mu,sigma=sigma,xi0=xi0)
cont_zero=ifelse(yt==0, 1,0)
n<-length(yt)
n0<-sum(yt==0) #numero de zeros
lambdahat<- n0/n #estimador de lambda

```

```

yt_0<-yt[-which(yt==0)]
fit<-gamlss(yt_0~1,family =
BP(mu.link = 'identity',sigma.link = 'identity'),
method = CG(), trace=F)

coeff<- c(lambdahat,fit$mu.coefficients,
fit$sigma.coefficients)
coef_results<-rbind(coef_results, coeff)

#Matriz InformaÃ§Ã£o de Fisher

muhat<-fit$mu.coefficients
phihat<-fit$sigma.coefficients

K11 =n/(lambdahat*(1-lambdahat))
Kmm = n*(1-lambdahat)*((phihat+1)^2)*(trigamma(muhat*(phihat+1))
-trigamma(muhat*(phihat+1)+phihat+2))

Kpp= n*(1-lambdahat)*((muhat^2)*trigamma(muhat*(phihat+1))
+trigamma(phihat+2)-((muhat+1)^2)
*trigamma(muhat*(phihat+1)+phihat+2))

m1 = (phihat+1)*trigamma(muhat*(phihat+1)+phihat+2)
m2 = muhat*(phihat+1)*(trigamma(muhat*(phihat+1)
+phihat+2)-trigamma(muhat*(phihat+1)))
Kmp = -n*(1-lambdahat)*(m1+m2)

I<- matrix(c(K11,0,0,0,Kmm,Kmp,0,Kmp,Kpp), 3,3)

#vcov<- solve(I)
#vcov <- try(solve(I))
vcov <- chol2inv(chol(I))

ep_results<-rbind(ep_results,sqrt(diag(vcov)))

```

```

i = i+1
}
m<-colMeans(coef_results,na.rm=T)
sd<-apply(coef_results, 2, sd)
bias<-(m-pa)
var_est = apply(coef_results,2, var)
eqm<- var_est+(bias)^2
rb<- 100*(bias/pa)
cont_zero=sum(cont_zero)/R
M<-rbind(pa,m,rb)
row.names(M)<-c("Parâmetro", "Média", "VR")
colnames(M)<-c("lambda", "mu", "phi")
print(c("n", n), quote = F)
print(c("nc",nc-ni), quote = F)
print(c("ni", ni), quote = F)
print("#-----")
print("#-----")
print("avaliacao pontual")

print(round(M,4))
#print(ep_results)

# Intervalos de confiança

IC_inf_1 = coef_results - vt1*ep_results
IC_inf_5 = coef_results - vt5*ep_results
IC_inf_10 = coef_results - vt10*ep_results
IC_sup_1 = coef_results + vt1*ep_results
IC_sup_5 = coef_results + vt5*ep_results
IC_sup_10 = coef_results + vt10*ep_results

Coblambda_1 = (ifelse(((pa[1]>IC_inf_1[,1])&
(pa[1]<IC_sup_1[,1])),1, 0))
TbSlambda_1 = (ifelse((pa[1]>IC_sup_1[,1]),1,0))
TbIlambda_1 = (ifelse((pa[1]<IC_inf_1[,1]),1,0))

Coblambda_5 = (ifelse(((pa[1]>IC_inf_5[,1])&
(pa[1]<IC_sup_5[,1])),1, 0))

```

TbSlambda_5 = (ifelse((pa[1]>IC_sup_5[,1]),1,0))
TbIlambda_5 = (ifelse((pa[1]<IC_inf_5[,1]),1,0))

Coblamba_10 = (ifelse(((pa[1]>IC_inf_10[,1])&
(pa[1]<IC_sup_10[,1])),1, 0))
TbSlamba_10 = (ifelse((pa[1]>IC_sup_10[,1]),1,0))
TbIlamba_10 = (ifelse((pa[1]<IC_inf_10[,1]),1,0))

Cobmu_1 = (ifelse(((pa[2]>IC_inf_1[,2])&
(pa[2]<IC_sup_1[,2])),1, 0))
TbSmu_1 = (ifelse((pa[2]>IC_sup_1[,2]),1,0))
TbImu_1 = (ifelse((pa[2]<IC_inf_1[,2]),1,0))

Cobmu_5 = (ifelse(((pa[2]>IC_inf_5[,2])&
(pa[2]<IC_sup_5[,2])),1, 0))
TbSmu_5 = (ifelse((pa[2]>IC_sup_5[,2]),1,0))
TbImu_5 = (ifelse((pa[2]<IC_inf_5[,2]),1,0))

Cobmu_10 = (ifelse(((pa[2]>IC_inf_10[,2])&
(pa[2]<IC_sup_10[,2])),1, 0))
TbSmu_10 = (ifelse((pa[2]>IC_sup_10[,2]),1,0))
TbImu_10 = (ifelse((pa[2]<IC_inf_10[,2]),1,0))

Cobphi_1 = (ifelse(((pa[3]>IC_inf_1[,3])&
(pa[3]<IC_sup_1[,3])),1, 0))
TbSphi_1 = (ifelse((pa[3]>IC_sup_1[,3]),1,0))
TbIphi_1 = (ifelse((pa[3]<IC_inf_1[,3]),1,0))

Cobphi_5 = (ifelse(((pa[3]>IC_inf_5[,3])&
(pa[3]<IC_sup_5[,3])),1, 0))
TbSphi_5 = (ifelse((pa[3]>IC_sup_5[,3]),1,0))
TbIphi_5 = (ifelse((pa[3]<IC_inf_5[,3]),1,0))

Cobphi_10 = (ifelse(((pa[3]>IC_inf_10[,3])&

```

(pa[3]<IC_sup_10[,3])),1, 0))
TbSphi_10 = (ifelse((pa[3]>IC_sup_10[,3]),1,0))
TbIphi_10 = (ifelse((pa[3]<IC_inf_10[,3]),1,0))

## Taxa de não cobertura

tnc_S_lambda_1=(sum(TbSlambda_1)/R)*100
tnc_I_lambda_1=(sum(TbIlambda_1)/R)*100
leng_lambda_1 = ((sum(TbSlambda_1)/R)-(sum(TbIlambda_1)/R))*100

tnc_S_lambda_5=(sum(TbSlambda_5)/R)*100
tnc_I_lambda_5=(sum(TbIlambda_5)/R)*100
leng_lambda_5 = ((sum(TbSlambda_5)/R)-(sum(TbIlambda_5)/R))*100

tnc_S_lambda_10=(sum(TbSlambda_10)/R)*100
tnc_I_lambda_10=(sum(TbIlambda_10)/R)*100
leng_lambda_10 = ((sum(TbSlambda_10)/R)-(sum(TbIlambda_10)/R))*100

tnc_S_mu_1=(sum(TbSmu_1)/R)*100
tnc_I_mu_1=(sum(TbImu_1)/R)*100
leng_mu_1 = ((sum(TbSmu_1)/R)-(sum(TbImu_1)/R))*100

tnc_S_mu_5=(sum(TbSmu_5)/R)*100
tnc_I_mu_5=(sum(TbImu_5)/R)*100
leng_mu_5 = ((sum(TbSmu_5)/R)-(sum(TbImu_5)/R))*100

tnc_S_mu_10=(sum(TbSmu_10)/R)*100
tnc_I_mu_10=(sum(TbImu_10)/R)*100
leng_mu_10 = ((sum(TbSmu_10)/R)-(sum(TbImu_10)/R))*100

tnc_S_phi_1=(sum(TbSphi_1)/R)*100
tnc_I_phi_1=(sum(TbIphi_1)/R)*100
leng_phi_1 = ((sum(TbSphi_1)/R)-(sum(TbIphi_1)/R))*100

tnc_S_phi_5=(sum(TbSphi_5)/R)*100

```

```

tnc_I_phi_5=(sum(TbIphi_5)/R)*100
leng_phi_5 = ((sum(TbSphi_5)/R)-(sum(TbIphi_5)/R))*100

tnc_S_phi_10=(sum(TbSphi_10)/R)*100
tnc_I_phi_10=(sum(TbIphi_10)/R)*100
leng_phi_10 = ((sum(TbSphi_10)/R)-(sum(TbIphi_10)/R))*100

#Taxas de cobertura - 99%
tlambda_1 = (sum(Coblamba_1)/R)*100
tmu_1 = (sum(Cobmu_1)/R)*100
tphi_1 =(sum(Cobphi_1)/R)*100

#Taxas de cobertura - 95%
tlambda_5 = (sum(Coblamba_5)/R)*100
tmu_5 = (sum(Cobmu_5)/R)*100
tphi_5 =(sum(Cobphi_5)/R)*100

#Taxas de cobertura - 90%
tlambda_10 = (sum(Coblamba_10)/R)*100
tmu_10 = (sum(Cobmu_10)/R)*100
tphi_10 =(sum(Cobphi_10)/R)*100

lambda_90 = c( tlambda_10, tnc_I_lambda_10,tnc_S_lambda_10 )
lambda_95 = c( tlambda_5, tnc_I_lambda_5,tnc_S_lambda_5 )
lambda_99 = c( tlambda_1, tnc_I_lambda_1,tnc_S_lambda_1 )
lambda_90
lambda_95
lambda_99

mu_90 = c( tmu_10, tnc_I_mu_10,tnc_S_mu_10 )
mu_95 = c( tmu_5, tnc_I_mu_5,tnc_S_mu_5 )
mu_99 = c( tmu_1, tnc_I_mu_1,tnc_S_mu_1 )
mu_90
mu_95
mu_99

```

```

phi_90 = c( tphi_10, tnc_I_phi_10,tnc_S_phi_10 )
phi_95 = c( tphi_5, tnc_I_phi_5,tnc_S_phi_5 )
phi_99 = c( tphi_1, tnc_I_phi_1,tnc_S_phi_1 )

phi_90
phi_95
phi_99

result_90=c(lambda_90, mu_90, phi_90)
result_95=c(lambda_95, mu_95, phi_95)
result_99=c(lambda_99, mu_99, phi_99)

result_IC = rbind(result_90,result_95,result_99)
result_IC = round(result_IC,3)

colnames(result_IC)<-c("TC_l", "TNC_l_I","TNC_l_S", "TC_m", "TNC_m_I","TNC_m_S",
"TC_p", "TNC_p_I","TNC_p_S")

rownames(result_IC)<-c("90%", "95%","99%")

print("avaliacaoIC ")
print(round(result_IC,4))

print ("#####")

print ("#-----")
print ("#-----")
}

```