



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

**As Distribuições Beta Burr XII e
Beta Weibull Exponenciada:
uma abordagem Bayesiana**

por

Alex Felipe Rodrigues Lima

Orientadora: Prof. Dra. Cibele Queiroz da Silva

Dezembro de 2016

Alex Felipe Rodrigues Lima

As Distribuições Beta Burr XII e Beta Weibull Exponenciada: uma abordagem Bayesiana

Dissertação apresentada ao Departamento de
Estatística do Instituto de Ciências Exatas
da Universidade de Brasília como requisito
parcial à obtenção do título de Mestre em
Estatística.

Universidade de Brasília
Brasília, Dezembro de 2016

TERMO DE APROVAÇÃO

Alex Felipe Rodrigues Lima

**As Distribuições Beta Burr XII e
Beta Weibull Exponenciada:
uma abordagem Bayesiana**

Dissertação de Mestrado apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Data da defesa: 05 de Dezembro de 2016

Orientadora:

Profa. Dra. Cibele Queiroz da Silva
Departamento de Estatística, UnB

Comissão Examinadora:

Prof. Dr. Helton Saulo Bezerra dos Santos
Instituto de Matemática e Estatística, UFG

Prof. Dr. Eduardo Yoshio Nakano
Departamento de Estatística, UnB

Brasília, Dezembro de 2016

Ficha Catalográfica

LIMA, Alex Felipe Rodrigues

As distribuições Beta Burr XII e Beta Weibull Exponenciada:
uma abordagem Bayesiana, (UnB - IE, Mestre em Estatística, 2016).

Dissertação de Mestrado - Universidade de Brasília. Departamento de Estatística
- Instituto de Ciências Exatas.

1. Beta BurrXII 2. Beta Weibull Exponenciada 3. Abordagem Bayesiana

É concedida à Universidade de Brasília a permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta monografia pode ser reproduzida sem a autorização por escrito do autor.

Alex Felipe Rodrigues Lima

Agradecimentos

A minha esposa e minha filha, Tamara e Heloísa, a quem tenho grande amor, por sempre estarem sempre do meu lado, mesmo nos momentos mais difíceis, me apoiando, me concedendo todo carinho e, também, por toda paciência que tiveram comigo. Obrigado por tudo! Amo vocês.

A meus pais, Bertolino e Ester, meus irmãos, Guilherme e Letícia, meus sogros, Lucimar e Mônica, meus sobrinhos, Davi e Miguel, meu afilhado, Guidson, meus tios, João e Nalcisa por todo carinho, estímulo e por suportarem minha ausência várias vezes durante todo esse período de dedicação aos estudos. Obrigado por tudo! Amo vocês.

A minha orientadora, Cibele Queiroz da Silva, que sem ela, nada disto teria sido possível, pois soube sempre me ajudar a ultrapassar os obstáculos que surgiram, aliás não foram poucos. Sempre me incentivou, teve muita paciência, e foi, sem dúvida nenhuma, sempre muita dedicada a este trabalho. Obrigado por tudo e desculpa pelas minhas falhas.

A meus amigos e colegas da pós-graduação, Luciana, Marcílio, Vanessa, Felipe e Evelyn; do IMB, Sérgio, Murilo, Guilherme, Luiz Batista e Paulo; da UFG, Prof. Mário e Prof. Valdivino; de infância, Vinicius, Mauro, Dirceu, Rogério e Bruno pelos conselhos e motivação em todos os momentos.

Por fim, agradeço a todas as pessoas que de uma forma direta ou indireta contribuíram para o meu crescimento profissional e acadêmico.

Lista de Tabelas

4.1	Estimativas do WinBUGS com $a=16$, $b=8$, $k=6.0$, $s=1.0$, $c=5.0$	38
4.2	Estimativas marginais a posteriori obtida pelo <i>JAGS</i>	40
5.1	Estimativas marginais a posteriori (MCMC) dos parâmetros da Distribuição BBXII com tamanho amostral 300.	46
5.2	Estimativas marginais à posteriori da Distribuição BBXII com amostra 200.	52
5.3	Estimativas marginais a posteriori da Distribuição BBXII utilizando a abordagem da Seção 5.2.2.	56
5.4	EMV e estimativas bayesianas obtidos para a Distribuição BBXII com <i>prioris</i> Beta e Qui-Quadrado.	59
5.5	Estimativas marginais à posteriori da Distribuição BBXII utilizando a abordagem da Seção 5.3.	64
5.6	EMVs e estimativas bayesianas obtidas para a Distribuição BBXII.	67
6.1	Estimativas marginais à posteriori da Distribuição BEW utilizando a abordagem da Subseção 6.1.	72
6.2	EMVs e estimativas bayesianas obtidas para a Distribuição BEW.	75
7.1	EMV dos parâmetros dos modelos BBXII, BurrXII, Beta Weibull, Log-logística e Weibull para os dados de melanoma (Ibrahim et. al., 2001) e intervalos de confiança (95%) correspondentes.	78
7.2	AIC, BIC e CAIC das distribuições BBXII, BurrXII, Beta Weibull, Log-logística e Weibull para os dados melanoma.	79

7.3	Estimativas Bayesianas dos parâmetros dos modelos BBXII para os dados de melanoma, intervalos HPD (95%) correspondentes e taxas de aceitação.	83
8.1	Estimativas a posteriori (MCMC) da BBurrXII.	97
8.2	Estimativas marginais a posteriori da BBurrXII.	108

Lista de Figuras

1.1	Gráfico da densidade BBXII	11
1.2	Gráfico da densidade BEW	14
5.1	Traço, densidade e autocorrelação em amostras MCMC para os parâmetros a, b, k e c da distribuição BBXII.	47
5.2	Curva associada a densidade verdadeira e a curva estimada.	48
5.3	Traço, densidade e autocorrelação das amostras MCMC para os parâmetros a, b, k e c da distribuição BBXII.	53
5.4	(a) Histograma dos dados e a Curva estimada (b) Box-plot e o valor do verdadeiro parâmetro (linha horizontal).	54
5.5	Traço, densidade, autocorrelação e boxplots associados as amostras MCMC para os parâmetros μ , γ , k e c da distribuição BBXII sob a proposta da Seção 5.2.2.	57
5.6	Traço, Densidade, Autocorrelação e Boxplot's das amostras MCMC para os parâmetros μ , γ , k e c da distribuição BBXII sob a proposta da Seção 5.3.	65
6.1	Traço, Densidade, Autocorrelação e Boxplots das amostras MCMC para os parâmetros μ , γ , α , λ e c da distribuição BEW sob a proposta da Seção 6.1	74
7.1	Função empírica de sobrevivência para os dados de Melanoma (Ibrahim et al, 2001).	79
7.2	Análise gráfica dos resíduos de Cox-Snell dos modelos Weibull, BurrXII, Log-logística, Beta Weibull e BBXII ajustado aos dados de melanoma (Ibrahim et al, 2001).	81

7.3	Traço, Densidade, Autocorrelação e Boxplots das amostras MCMC para os parâmetros μ , γ , s , k e c da distribuição BBXII obtido via aproximação gaussiana dos parâmetros e vetor de candidatos - Metodologia proposta na Seção 5.3.	84
7.4	Traço, Densidade, Autocorrelação e Boxplots das amostras MCMC para os parâmetros μ , γ , s , k e c da distribuição BBXII obtido via aproximação gaussiana dos parâmetros e candidatos univariados- Metodologia proposta na Seção 6.1	85
7.5	Função empírica de sobrevivência ajustada aos dados de Melanoma (Ibrahim et al, 2001).	86
7.6	Análise gráfica dos resíduos de Cox-Snell do BBXII obtidos pelas abordagens das Seções 6.1 e 5.3 e ajustados aos dados de melanoma (Ibrahim et al, 2001).	86
8.1	Gráfico dos valores iniciais e dos estimados	98
8.2	Histograma dos dados e a curva estimada com os valores iniciais (curva pontilhada)	108
8.3	Histograma dos dados e a curva estimada com os parâmetros verdadeiros e estimados	109
8.4	Traço, Densidade, Autocorrelação e Boxplot's das amostras MCMC para os parâmetros a , b , k e c da distribuição BBXII	110

Lista de Abreviaturas e Siglas

AIC Critério de informação de *Akaike*

BBXII Beta BurrXII

Beta-G Beta Generalizada

BEW Beta Weibull Exponenciada

BIC Critério de informação de Bayesiano

CAIC Critério de informação de *Akaike* corrigido

EMV Estimativa de máxima verossimilhança

EW Distribuição Weibull Exponencializada

fda Função de densidade acumulada

fdp Função densidade de probabilidade

fgm Função geradora de momentos

HPD highest posterior density (intervalo de credibilidade de maior densidade a posteriori)

JAGS Just Another Gibbs Sampler

KM Kaplan-Meier

MCMC Monte Carlo via Cadeia de Markov

ME Modelo específico

TRV Teste de razão de verossimilhança

Lista de Códigos

1.1	Código para geração de dados da Distribuição BBurrXII	12
1.2	Código em R para geração de dados da Distribuição BEW	16
4.1	Modelo BBurrXII no JAGS e no WinBUGS	36
4.2	Comandos para inserção de dados no WinBUGS	37
4.3	Comandos para inserção de dados e obtenção dos resultados no JAGS	39
5.1	Algoritmo para gerar um candidato pela aproximação da Normal com restrição no intervalo (a, b).	46

Sumário

Agradecimentos

Resumo 4

Abstract 5

Introdução 6

1 As Distribuições BBXII e BEW 9

1.1 Distribuição BBXII 9

1.1.1 Geração de dados 12

1.2 Distribuição BEW 13

1.2.1 Geração de dados 16

2 A Abordagem Frequentista 17

2.1 Caso BBXII 17

2.2 Caso BEW 19

2.3 Verossimilhança com dados censurados
ao acaso 20

2.3.1 Caso BBXII 22

2.3.2 Caso BEW 23

3 Inferência Bayesiana 24

3.1 Inferência Bayesiana 24

3.2 Estimativa Pontual 26

3.3 Estimativa Intervalar 27

3.4 MCMC 28

3.4.1	Metropolis-Hastings	29
3.4.2	Amostrador de Gibbs	30
3.5	Avaliação do MCMC	31
3.5.1	Análise Gráfica	31
3.5.2	Método de Gelman e Rubin	31
3.5.3	Método de Geweke	32
3.5.4	Taxa de Aceitação	33
4	A Abordagem Bayesiana via JAGS e WinBUGS	34
4.1	Distribuição a Posteriori e Condiçoes Completas	35
4.2	JAGS e WinBUGS	36
4.2.1	Dados e Resultados no WinBUGS	37
4.2.2	Dados e Resultados no JAGS	39
5	Implementações Bayesianas da distribuição BBXII utilizando o R	41
5.1	Utilização de prioris Gama	41
5.1.1	Proposta de Função Geradora de Candidatos para o Parâmetro a	42
5.1.2	Proposta de Função Geradora de Candidatos para o Parâmetro b	43
5.1.3	Proposta de Função Geradora de Candidatos para o Parâmetro k	43
5.1.4	Proposta de Função Geradora de Candidatos para o Parâmetro s	44
5.1.5	Proposta de Função Geradora de Candidatos para o Parâmetro c	45
5.1.6	Resultados	46
5.2	Nova parametrização da Distribuição Beta	49
5.2.1	Distribuição a posteriori da BBXII com a Nova Parametrização da Beta e Função Geradora de Candidatos	50
5.2.1.1	Resultados	52
5.2.2	Prioris Beta e Qui-Quadrado	55
5.2.2.1	Resultados	55
5.2.2.2	Estudo Comparativo entre as abordagens frequentista e Bayesiana	56
5.3	Aproximações Gaussianas para os parâmetros da distribuição BBXII .	60
5.3.1	Função Geradora de Candidatos Conjunta	62

5.3.2	Resultados	64
5.3.2.1	Estudo Comparativo entre as abordagens frequentista e Bayesiana	64
6	Implementação Bayesiana da distribuição BEW utilizando o R	68
6.1	Aproximações Gaussianas para os parâmetros da distribuição BEW .	68
6.1.1	Função Geradora de Candidatos Univariado	71
6.1.2	Resultados	71
6.1.3	Estudo Comparativo entre as abordagens frequentista e Bayesiana	72
7	Aplicação a dados reais	76
7.1	Abordagem Clássica	77
7.1.1	Resultados	78
7.2	Abordagem Bayesiana	80
7.2.1	Resultados	82
8	Discussão	87
	Referências Bibliográficas	90
	Apêndice	93

Resumo

O interesse principal do trabalho está na proposta de uma abordagem Bayesiana adequada para a estimação dos parâmetros das distribuições Beta BurrXII (BBXII) e Beta Weibull Exponenciada (BEW). Essas distribuições pertencem à classe de distribuições Beta Generalizadas (Beta-G). Sendo a distribuição Beta uma parte integrante das distribuições BBXII e da BEW, constatou-se que a sua reparametrização, proposta por Ferrari e Cribari-Neto (2004), fornece vantagens computacionais para a convergência das estimativas Bayesianas. Foram propostas duas abordagens Bayesianas para a estimação dos parâmetros da distribuição BBXII e uma abordagem para a distribuição BEW. Para a BBXII, a primeira abordagem considera *prioris* e funções geradoras de candidatos Beta, para o parâmetro μ , e Qui-Quadrado para os demais parâmetros. A segunda abordagem considera transformações *logit* para μ e log para os demais parâmetros. As *prioris* e funções geradoras de candidatos adotadas foram Beta para μ e Gama para os demais parâmetros. Nessa abordagem, obtém-se um vetor de candidatos Gaussianos de acordo com uma adaptação da proposta Gaussiana de Gray(2001). Para a BEW, considerou-se as transformações, *prioris* e funções geradoras de candidatos equivalentes à segunda proposta da BBXII, diferindo na obtenção de candidatos, que somente puderam ser obtidos de forma univariada para cada parâmetro.

Palavras Chave: *Beta BurrXII, Beta Weibull Exponenciada e Abordagem Bayesiana.*

Abstract

The aim of this work is the development of an adequate Bayesian approach for the estimation of the parameters of the Beta BurrXII (BBXII) and Beta Exponentiated Weibull (BEW) distributions. Both of these distributions belong to the Beta Generalized Class (Beta-G). Since the Beta distribution is used in the construction of the BBXII and the BEW distribution, It was observed that the Beta distribution reparametrization proposed by Ferrari and Cribari-Neto (2004) provides computational advantages with respect to the convergence of the Bayesian estimates. Two Bayesian approaches were proposed to the estimation of the parameters of the BBXII distribution and one approach was proposed to the BEW one. In the case of the BBXII distribution, a first approach incorporates both beta priors and beta proposal distributions for the parameters μ and both Chi-square priors and Chi-square proposal distributions for the other parameters. In the second approach it was applied a set of transformations to the original parameters in order to make possible the use of Gaussian approximations. It was applied a *logit* transformation to μ and a log transformation to the others parameters. The proposal distributions were set up according to an adaptation of the Gaussian approaches proposed by Gray (2001). In the case of the BEW distribution it was also used Gaussian approximations to the original parameters, with the difference that for the proposals it only possible the use of univariate distributions.

key words: *Beta BurrXII, Beta Exponentiated Weibull, Bayesian approach.*

Introdução

A família de distribuições Beta-G (Eugene et al., 2002) tem motivado um extenso número de pesquisas no desenvolvimento de novas distribuições de probabilidade desde sua criação. Eugene et al. (2002) postularam que, para a função distribuição acumulada G de uma dada variável aleatória (em que a função de densidade de probabilidade e acumulada são denotadas por $g(x)$ e $G(x)$, respectivamente), a função distribuição acumulada beta-G é definida por

$$F(x) = \frac{1}{B(a, b)} \int_0^{G(x)} w^{(a-1)}(1-w)^{(b-1)} dw \quad (1)$$

e, conseqüentemente, sua função de densidade de probabilidade é dada por

$$f(x) = \frac{d}{dx}F(x) = \frac{g(x)}{B(a, b)}G(x)^{a-1}(1-G(x))^{b-1}, \quad (2)$$

em que $B(a, b)$ é a função Beta.

Tahir e Nadarajah (2015) fazem uma compilação dos artigos de desenvolvimento relativos a distribuição Beta-G. Esta relação inclui, entre outras, a distribuição Beta Burr XII (Paranaíba, et. al., 2011) e a distribuição Beta Weibull Exponenciada (Cordeiro et al., 2013), que foram descritas ao considerar-se, respectivamente, as funções distribuição G relativas à distribuição Burr XII (Burr, 1942) e à distribuição Weibull (Weibull, 1951), que são de interesse central nesta dissertação.

Apesar de haver uma extensa literatura sobre as distribuições na família Beta-G e sobre os procedimentos frequentistas, via máxima verossimilhança, adotados na estimação dos parâmetros dos modelos propostos, são escassos os trabalhos que utilizam o paradigma Bayesiano no processo inferencial. Uma das vantagens do uso da abordagem Bayesiana é a obtenção de intervalos de credibilidade para os parâmetros

do modelo beta-G em estudo sem a necessidade do uso de argumentos assintóticos.

Neste trabalho de dissertação objetiva-se o desenvolvimento de metodologia Bayesiana adequada para a estimação dos parâmetros dos modelos Beta BurrXII (BBXII) e Beta Weibull Exponeciada (BEW). A dissertação está organizada como a seguir: no Capítulo 1 apresenta-se as definições dos modelos BBXII e BEW assim como os procedimentos necessários para a simulação de dados a partir de tais distribuições.

No Capítulo 2 descreve-se os procedimentos frequentistas para a estimação dos parâmetros das distribuições BBXII e BEW e as formas funcionais das respectivas verossimilhanças para o caso de dados censurados com censuras ao acaso.

No Capítulo 3 são apresentados conceitos básicos de inferência Bayesiana e no Capítulo 4 algumas tentativas de abordagem de estimação Bayesiana das distribuições BBXII e BEW utilizando-se os softwares WinBUGS e JAGS. No entanto, tais iniciativas mostraram-se infrutíferas de modo que foram necessários a implementação vários códigos computacionais em R para viabilizar a exploração dos métodos que propusemos nesta dissertação.

No Capítulo 5 são apresentadas novas metodologias Bayesianas para a estimação dos parâmetros da distribuição BBXII. Constatou-se que o uso da forma reparametrizada da distribuição Beta, como descrito em Ferrari e Cribari-Neto (2004), traz vantagens computacionais e conduz à convergência das cadeias de Markov geradas via método Monte Carlo via Cadeia de Markov (MCMC).

Os modelos BBXII e BEW incorporam elementos da distribuição Beta e, inicialmente trabalhou-se com os parâmetros desta última distribuição em sua forma original. No entanto, em inúmeras estratégias que foram tentadas, ou não obteve-se a convergência das cadeias MCMC ou os resultados obtidos das estimativas dos parâmetros para dados simulados foram muito insatisfatórios.

Constatou-se (Capítulo 5) que o uso da parametrização da distribuição Beta, como descrita em Ferrari e Cribari-Neto (2004), traz vantagens computacionais e conduz à convergência das cadeias de Markov geradas via método MCMC, além de conduzir a valores estimados muito mais próximos aos valores reais dos parâmetros.

Na tentativa da obtenção de procedimentos mais robustos e precisos das distribuições BBXII e BEW foram propostos, nos Capítulos 5 e 6, aproximações Gaussianas da distribuição a posteriori conjunta de transformações dos parâmetros origi-

nais de modo a trabalhar-se com espaços paramétricos descritos nos reais. Obteve-se resultados satisfatórios com tal estratégia.

No Capítulo 7 descreve-se as análises Bayesianas para o modelo BBXII utilizando os dados de melanoma de Ibrahim et al. (2001). Acreditamos que os métodos desenvolvidos nesta dissertação conduzem a um avanço em relação a proposta Bayesiana de Paranaíba et. al. (2011) para a distribuição BBXII. Além disso, propusemos um método Bayesiano (que acreditamos seja o primeiro) para a distribuição BEW.

Capítulo 1

As Distribuições BBXII e BEW

Nesta seção serão apresentados as Distribuições BBXII (Paranaíba et. al. 2011) e BEW (Cordeiro et al. 2011) e algumas propriedades. Também será provado que a função de distribuição de probabilidade de ambas distribuições podem ser escritas como uma combinação linear da distribuição BurrXII e Weibull, respectivamente.

1.1 Distribuição BBXII

A distribuição BBXII foi proposta por Paranaíba et.al (2011). Ela é uma composição da distribuição Beta generalizada com a distribuição Burr XII. A primeira foi proposta por Eugene et. al. (2002), que definem uma classe de distribuições beta generalizadas (beta-G), que consiste numa generalização da distribuição Beta. Considere $G(x)$ denotando uma função de distribuição acumulada (fda) de uma variável aleatória contínua e $a > 0$ e $b > 0$ (parâmetros de forma) da distribuição beta. Tem-se que a fda da beta-G é dada por

$$F(x) = \frac{1}{B(a, b)} \int_0^{G(x)} w^{(a-1)}(1-w)^{(b-1)} dw, \quad (1.1)$$

em que $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ é a função *beta* e $\Gamma(\cdot)$ é a função *Gama*. Pelas propriedades de probabilidade tem-se que a função densidade de probabilidade fdp associada a $F(x)$ é dada por:

$$f(x) = \frac{d}{dx}F(x) = \frac{g(x)}{B(a, b)}G(x)^{a-1}(1-G(x))^{b-1}, \quad (1.2)$$

onde $g(x) = dG(x)/dx$ é função densidade da distribuição de probabilidade (fdp) a ser definida.

Dessa forma, a distribuição BBXII é obtida ao assumir $G(x)$ e $g(x)$ sendo a fda e fdp da distribuição BurrXII, respectivamente. A distribuição BurrXII foi proposta por Burr (1942), possui 3 parâmetros ($s > 0, k > 0$ e $c > 0$) e tem a seguinte forma:

$$G(x; s, k, c) = 1 - \left[1 + \left(\frac{x}{s}\right)^c\right]^{-k}, \quad x > 0, \quad (1.3)$$

e

$$g(x; s, k, c) = ck s^{-c} \left[1 + \left(\frac{x}{s}\right)^c\right]^{-k-1} x^{c-1}, \quad x > 0. \quad (1.4)$$

Diante disso, assumindo $G(x)$ a fda da BurrXII e substituindo na fda da beta-G, tem-se que a fda da distribuição BBXII é dada por ($x > 0$)

$$F(x) = I_{\{1 - [1 + (x/s)^c]^{-k}\}}(a, b) = \frac{1}{B(a, b)} \int_0^{G(x)=1 - [1 + (x/s)^c]^{-k}} w^{a-1} (1-w)^{b-1} dw \quad (1.5)$$

e sua fdp é

$$f(x) = \frac{ckx^{c-1}}{s^c B(a, b)} [1 + (x/s)^c]^{-(kb+1)} \{1 - [1 + (x/s)^c]^{-k}\}^{a-1}. \quad (1.6)$$

Dessa forma, a Função de Risco da distribuição BBXII é dada por ($x > 0$):

$$h(x) = \frac{ckx^{c-1} [1 + (x/s)^c]^{-(kb+1)} \{1 - [1 + (x/s)^c]^{-k}\}^{a-1}}{s^c B(a, b) [1 - I_{\{1 - [1 + (x/s)^c]^{-k}\}}(a, b)]}. \quad (1.7)$$

Na Figura 1.1 ilustra-se algumas formas possíveis da distribuição BBXII.

Uma característica da distribuição BBXII é que podemos obter uma série distribuições em casos particulares. A exemplo, quando $a = b = 1$, obtêm-se a distribuição BXII; quando $b = 1$, temos a distribuição BurrXII exponencializada; e quando temos $a = c = 1$ ou $a = b = c = 1$, temos a Beta-Pareto tipo II (BP-II) e Pareto tipo II (PII), respectivamente.

Outra característica dessa distribuição é que a fdp pode ser escrita como uma combinação linear de distribuições BurrXII. Considerando a expansão de $(1 - z)^{b-1}$

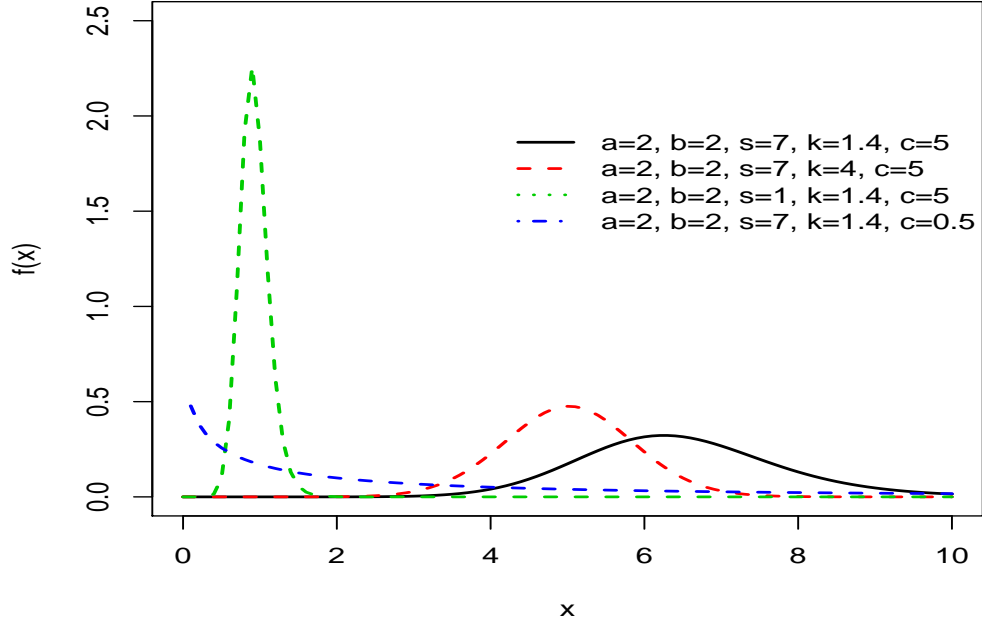


Figura 1.1: Gráfico da densidade BBXII

quando $b > 0$ e $|z| < 1$ (Nadarajah; Kotz, 2004),

$$(1 - z)^{b-1} = \sum_{i=0}^{\infty} \frac{(-1)^j \Gamma(b)}{\Gamma(b-j)j!} z^j. \quad (1.8)$$

Substituindo (1.8) em (1.5), temos que

$$F(x) = \frac{1}{B(a, b)} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{\Gamma(b-j)j!} \int_0^{1-(1+(x/s)^c)^{-k}} w^{a+j-1} dw. \quad (1.9)$$

Logo,

$$F(x) = \frac{1}{B(a, b)} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b) \{1 - (1 + (x/s)^c)^{-k}\}^{a+j}}{\Gamma(b-j)j!(a+j)}. \quad (1.10)$$

Aplicando a derivada expressa em (1.2), tem-se:

$$f(x) = \sum_{j=0}^{\infty} \frac{ckx^{c-1}(-1)^j \Gamma(b)(1 + (x/s)^c)^{-(k+1)}}{B(a, b)s^c \Gamma(b-j)j!} \{1 - [1 + (x/s)^c]^{-k}\}^{a+j-1}. \quad (1.11)$$

Usando a expansão (1.8) em $\{1 - [1 + (x/s)^c]^{-k}\}^{a+j-1}$, tem-se que

$$\begin{aligned}
f(x) &= \sum_{j=0}^{\infty} \frac{ckx^{c-1}(-1)^j\Gamma(b)(1 + (x/s)^c)^{-(k+1)}}{B(a, b)s^c\Gamma(b-j)j!} \sum_{r=0}^{\infty} \frac{(-1)^r\Gamma(a+j)}{\Gamma(a+j-r)j!} (1 + (x/s)^c)^{-kr} \\
&= \sum_{j,r=0}^{\infty} \frac{(-1)^{j+r}\Gamma(b)\Gamma(a+j)ckx^{c-1}}{B(a, b)\Gamma(b-j)\Gamma(a+j-r)j!r!s^c} [1 + (x/s)^c]^{-k(r+1)-1} \\
&= \sum_{j,r=0}^{\infty} w_r(a, b) \frac{ckx^{c-1}}{s^c} [1 + (x/s)^c]^{-k(r+1)-1} \\
&= \sum_{j,r=0}^{\infty} w_r(a, b)g(x; s, k(r+1), c), \tag{1.12}
\end{aligned}$$

em que $g(x; s, k(r+1), c)$ é a fdp da distribuição Burr XII com parâmetros s , c e $k(r+1)$ e w_r é dado por:

$$w_r = w_r(a, b) = \frac{(-1)^{j+r}\Gamma(b)\Gamma(a+j)}{B(a, b)\Gamma(b-j)\Gamma(a+j-r)(r+1)!}. \tag{1.13}$$

Por integração de (1.13), para a e b positivos reais não-inteiros, obtém-se a fda da distribuição BBXII, que pode ser escrita como combinação linear da funções de densidade da distribuição BurrXII,

$$F(x) = \sum_{j,r=0}^{\infty} w_{j,r}G(x; s, k(r+1), c) \tag{1.14}$$

em que w_r foi definido em (1.13). Do mesmo modo, a função geradora de momentos (*fgm*) da distribuição pode ser escrita em termos da *fgm* da distribuição BurrXII reparametrizada.

1.1.1 Geração de dados

Os dados da distribuição BBXII podem ser gerados pelo método da inversão. Se $X \sim BBXII(a, b, k, s, c)$ e $V \sim Beta(a, b)$, uma amostra de tamanho n pode ser gerada conforme o Código 1.1.

Código 1.1: Código para geração de dados da Distribuição BBurrXII

```

1 Ger.VA.BURRXII<-function(a,b,s,k,cc,n)
  {

```

```

3  set.seed(8982229)
   x<-c()
5  V<-rbeta(n, a, b)
   x<-s*(((1-V)^(-1/k)-1)^(1/cc))
7  return(x=x)
}

```

Na próxima Seção, apresenta-se a distribuição BEW e será provado que a função de distribuição pode ser escrita como uma combinação linear da distribuição Weibull.

1.2 Distribuição BEW

A distribuição BEW foi proposto por Cordeiro et al (2011). Ela é uma composição da distribuição Weibull com a distribuição Exponencializada (ou Exponenciada) e a Beta generalizada. A primeira é uma distribuição muito conhecida e tem sido amplamente utilizada em estudos nas áreas médicas, engenharia e confiabilidade.

A distribuição Exponencializada é obtida quando a distribuição acumulada ($G(x)$) é elevada a um expoente arbitrário ($\alpha > 0$). Então a nova distribuição acumulada é dada por $F(x) = G(x)^\alpha$ e pode ser chamada de distribuição G exponencializada (ou exponenciada). Logo, a fda dessa distribuição é $f(x) = \alpha G(x)^{\alpha-1}g(x)$.

Assumindo que $G(x)$ seja a fda da distribuição Weibull, tem-se a fda da distribuição Weibull Exponencializada (EW), que foi proposta por Mudholkar e Srivastava (1993) e é dada por

$$G_{\lambda,\alpha,c}(x) = \{1 - \exp[-(\lambda x)^c]\}^\alpha, \quad x > 0, \quad (1.15)$$

onde $\alpha > 0$ e $c > 0$ são parâmetros de forma e $\lambda > 0$ é parâmetros de escala. Nota-se que um caso particular da distribuição EW é a distribuição Exponencial, quando $\alpha = c = 1$.

A distribuição Beta-G foi proposta por Eugene et al (2002) e foi definida em (1.1). Assumindo $G(x)$ dada em (1.15) e substituindo em (1.1) obtém-se a distribuição Beta

Weibull Exponencializada (BEW) que possui a seguinte forma:

$$\begin{aligned} F(x) &= I_{1-(1-e^{-(\lambda x)^c})^\alpha}(a, b) \\ &= \frac{1}{B(a, b)} \int_0^{G_{\lambda, \alpha, c}(x) = \{1 - \exp[-(\lambda x)^c]\}^\alpha} w^{a-1} (1-w)^{b-1} dw, w > 0, \end{aligned} \quad (1.16)$$

e sua $f dp$ é dado por:

$$f(x) = \frac{\alpha c \lambda^c}{B(a, b)} x^{c-1} e^{-(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{\alpha a - 1} \left\{ 1 - (1 - e^{-(\lambda x)^c})^\alpha \right\}^{b-1}. \quad (1.17)$$

Logo, a Função de Risco da distribuição BEW é dada por ($x > 0$):

$$h(x) = \frac{\lambda c \lambda^c x^{c-1} e^{-(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{\alpha a - 1} \left\{ 1 - (1 - e^{-(\lambda x)^c})^\alpha \right\}^{b-1}}{B(a, b) I_{1-(1-e^{-(\lambda x)^c})^\alpha}(a, b)}, \quad (1.18)$$

Na Figura 1.2 ilustra-se algumas formas possíveis da distribuição BEW. Uma car-

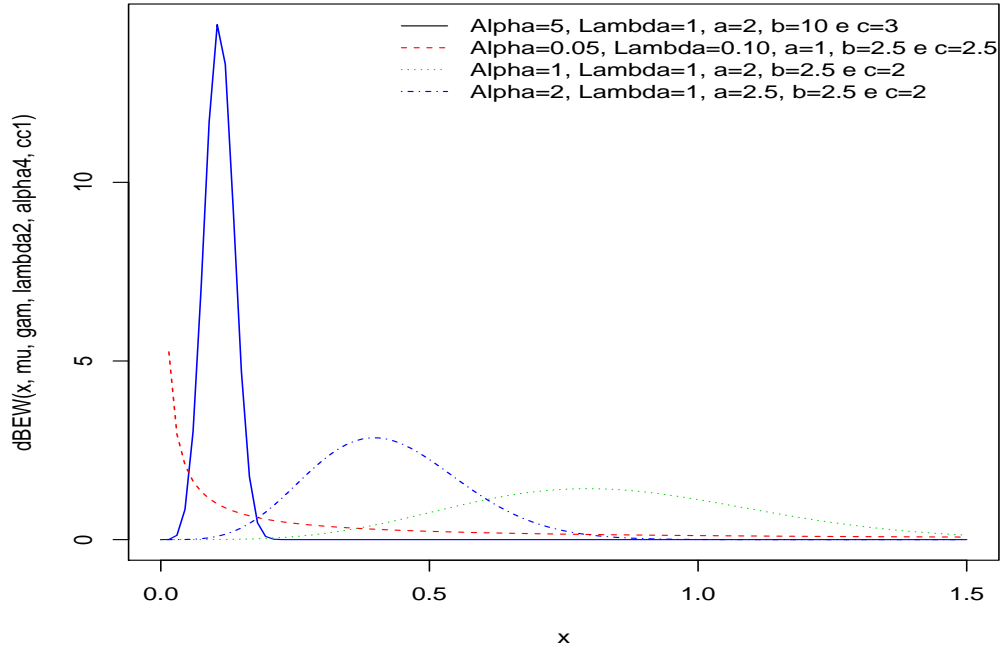


Figura 1.2: Gráfico da densidade BEW

acterística interessante dessa distribuição é a possibilidade de obtenção de várias distribuições como caso particular da BEW. Por exemplo, se $\alpha = a = b = 1$ tem-se a

distribuição Weibull, se $a = b = 1$ tem-se a distribuição Weibull Exponenciada e se $\alpha = b = c = 1$ obtém-se a distribuição Exponencial Exponenciada.

Outra característica interessante dessa distribuição é que sua fdp ou fda pode ser escrita pela soma ponderada de fdp ou fda da Weibull, respectivamente. Primeiramente, se $b > 0$ real e não inteiro, pode-se substituir $(1 - w)^{b-1}$ da integral pela série de potência e obter a seguinte integral:

$$\int_0^x w^{a-1}(1-w)^{b-1}dw = \sum_{j=0}^{\infty} \frac{(-1)^j \binom{b-1}{j}}{(a+j)} x^{a+j}, \quad (1.19)$$

quando o termo binomial $\binom{b-1}{j} = \Gamma(b)/\Gamma(b-j)!$ é definido para todo número real b . Da equação (1.16), tem-se que

$$F(x) = \frac{1}{B(a,b)} \sum_{j=0}^{\infty} \frac{(-1)^j \binom{b-1}{j}}{d(a+j)} \{1 - \exp[-(\lambda x)^c]\}^{\alpha(a+j)}. \quad (1.20)$$

Agora, da expansão binomial e de (1.15), tem-se que

$$F(x) = \sum_{j,k=0}^{\infty} \frac{(-1)^{j+k} \binom{b-1}{j} \binom{\alpha(a+j)}{k}}{(a+j)B(a,b)} [1 - G_{\lambda_k,1,c}(x)], \quad (1.21)$$

quando $G_{\lambda_k,1,c}(x)$ é uma fdp da Weibull com parâmetro de escala $\lambda_k = k^{1/c}\lambda$ e parâmetro de forma c . Diferenciando (1.21) tem-se uma expansão útil para a função de densidade que é dada por

$$f(x) = \sum_{k=1}^{\infty} w_{+,k} g_{\lambda_k,1,c}(x), \quad (1.22)$$

quando $g_{\lambda_k,1,c} = dG_{\lambda_k,1,c}(x)/dx$ denota a densidade da Weibull com parâmetro de escala λ_k , parâmetro de forma c e $w_{+,k} = \sum_{j=0}^{\infty} w_{j,k}$, onde $w_{j,k}$ são dados por:

$$w_{j,k} = \frac{(-1)^{j+k+1} \binom{b-1}{j} \binom{\alpha(a+j)}{k}}{(a+j)B(a,b)}. \quad (1.23)$$

1.2.1 Geração de dados

Os dados da distribuição BEW podem ser gerados pelo método da inversão. Se $X \sim BEW(a, b, \alpha, \lambda, c)$ e $V \sim Beta(a, b)$, uma amostra de tamanho n pode ser gerada conforme o Código 1.2.

Código 1.2: Código em R para geração de dados da Distribuição BEW

```
Ger.VA.BEW<-function(a,b,alpha,lambda,cc,n)
2 {
  set.seed(8982229)
4 x<-c()
  V<-rbeta(n,a,b)
6 x<-(1/lambda)*(-log(1-V^{1/alpha}))^(1/cc)
  return(x=x)
8 }
```

Capítulo 2

A Abordagem Frequentista

Nesta seção serão apresentadas as funções *score* das distribuições BBXII e BEW, além de um teste de razão de verossimilhança para avaliar a adequação das respectivas distribuições e qualquer outro de seus sub-modelos.

2.1 Caso BBXII

Seja x_1, x_2, \dots, x_n uma amostra aleatória de uma distribuição BBXII (vide (1.6)) com vetor de parâmetros $\theta=(a, b, s, k, c)^T$, então a correspondente função de verossimilhança é dada por,

$$L(\theta) = \prod_{i=1}^n \frac{ckx_i^{c-1}}{s^c B(a, b)} [1 + (x_i/s)^c]^{-(kb+1)} \left\{ 1 - [1 + (x_i/s)^c]^{-k} \right\}^{a-1}, \quad (2.1)$$

e, conseqüentemente, a função *log-verossimilhança* é:

$$l(\theta) = n \log(c) + n \log(k) + (c-1) \sum_{i=1}^n \log(x_i) - nc \log(s) - n \log(B(a, b)) - \quad (2.2) \\ (kb+1) \sum_{i=1}^n \log(1 + (x_i/s)^c) + (a-1) \sum_{i=1}^n \log\left(1 - [1 + (x_i/s)^c]^{-k}\right).$$

Fazendo $u_i = 1 - (1 + (x_i/s)^c)^{-k}$, as funções *score* dos parâmetros a, b, s, k e c são dadas por

$$\frac{\partial l(\theta)}{\partial a} = U_a(\theta) = -n[\psi(a) - \psi(a+b)] + \sum_{i=1}^n \log(u_i), \quad (2.3)$$

$$\frac{\partial l(\theta)}{\partial b} = U_b(\theta) = -n[\psi(b) - \psi(a+b)] + \sum_{i=1}^n \log(1 - u_i), \quad (2.4)$$

$$\frac{\partial l(\theta)}{\partial s} = U_s(\theta) = -\frac{rc}{s} + (a-1) \sum_{i=1}^n \left(\frac{x_i^c}{s^{c+1}} \right) \left[\frac{-kc(1-u_i)^{\frac{k+1}{k}}}{u_i} + \frac{c(kb+1)}{(1-u_i)^{-1/k}} \right], \quad (2.5)$$

$$\frac{\partial l(\theta)}{\partial k} = U_k(\theta) = \frac{r}{k} - \frac{1}{k} \sum_{i=1}^n \log(1 - u_i) \left[\frac{(a-1)(1-u_i)}{u_i} - b \right], \quad (2.6)$$

$$\frac{\partial l(\theta)}{\partial c} = U_c(\theta) \quad (2.7)$$

$$= \frac{r}{c} - r \log(s) + \sum_{i=1}^n \left(\frac{x_i}{s} \right)^c \log \left(\frac{x_i}{s} \right) \left[\frac{(a-1)k(1-u_i)^{\frac{k+1}{k}}}{u_i} - \frac{(kb+1)}{(1-u_i)^{-1/k}} \right], \quad (2.8)$$

em que $\psi(\cdot)$ é a função *digama*. Então o estimador de máxima verossimilhança $\hat{\theta}$ de θ é obtido resolvendo o sistema de equações não lineares tomado pelas equações $U_a(\theta) = 0$, $U_b(\theta) = 0$, $U_s(\theta) = 0$, $U_k(\theta) = 0$ e $U_c(\theta) = 0$.

Também é possível realizar um teste para verificar se a BBXII pode ser reduzida a algum submodelo. Para tanto, obtém-se a *log-verossimilhança* restrita ao submodelo a *log-verossimilhança* irrestrita maximizada e calcula-se a estatística de razão de *log-verossimilhança* (*TRV*). Um exemplo é o teste para avaliar as distribuições Burr XII exponencializada ($b = 1$) e a BBXII. Dessa forma as hipóteses são

$$\begin{cases} H_0 : b = 1, \\ H_1 : b \neq 1. \end{cases}$$

Com isso, a estatística de teste é dada por

$$w = 2 \left\{ l(\hat{a}, \hat{b}, \hat{s}, \hat{k}, \hat{c}) - l(\tilde{a}, 1, \tilde{s}, \tilde{k}, \tilde{c}) \right\}, \quad (2.9)$$

onde $\hat{a}, \hat{b}, \hat{s}, \hat{k}, \hat{c}$ são estimados sob H_1 e $\tilde{a}, 1, \tilde{s}, \tilde{k}, \tilde{c}$ são estimados sob H_0 . Sob essa hipótese $w \rightarrow \chi_m^2$, onde m é a diferença entre o número de parâmetros do espaço

paramétrico sob H_1 (θ_{H_1}) e do espaço paramétrico sob H_0 .

2.2 Caso BEW

Seja x_1, x_2, \dots, x_n uma amostra aleatória de uma distribuição BEW (vide (1.17)) com vetor de parâmetros $\theta = (a, b, \alpha, \lambda, c)^T$, então a correspondente função de máxima verossimilhança é dada por,

$$L(\theta) = \prod \left[\frac{\alpha c \lambda^c}{B(a, b)} x^{c-1} e^{-(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{\alpha a - 1} \left\{ 1 - (1 - e^{-(\lambda x)^c})^\alpha \right\}^{b-1} \right], \quad (2.10)$$

a correspondente função *log-verossimilhança* é dada por:

$$\begin{aligned} l(\theta) &= \log(\alpha) + \log(c) + c \log(\lambda) - \log[B(a, b)] + (c-1) \log(x) + \\ &\quad - (\lambda x)^c + (\alpha a - 1) \log \{1 - \exp[-(\lambda x)^c]\} \\ &\quad + (b-1) \log \{1 - [1 - \exp\{- (\lambda x)^c\}]^\alpha\}. \end{aligned} \quad (2.11)$$

Logo os componentes do vetor *score* $U = (\partial l / \partial a, \partial l / \partial b, \partial l / \partial \alpha, \partial l / \partial \lambda, \partial l / \partial c)^T$ são dados por

$$\begin{aligned} \frac{\partial l}{\partial \alpha} &= \frac{1}{\alpha} + \log \{1 - \exp[-(\lambda x)^c]\} - \\ &\quad - (b-1) \frac{\{1 - \exp[-(\lambda x)^c]\}^\alpha \log \{1 - \exp[-(\lambda x)^c]\}}{1 - \{1 - \exp[-(\lambda x)^c]\}^\alpha}, \end{aligned} \quad (2.12)$$

$$\begin{aligned} \frac{\partial l}{\partial \lambda} &= \frac{c}{\lambda} - c \lambda^{c-1} x^c + (\alpha a - 1) \frac{c x^c \lambda^{c-1} \exp[-(\lambda x)^c]}{1 - \exp[-(\lambda x)^c]} \\ &\quad - (b-1) \frac{\alpha \{1 - \exp[-(\lambda x)^c]\}^{\alpha-1} x^c c \lambda^{c-1} \exp[-(\lambda x)^c]}{1 - \{1 - \exp[-(\lambda x)^c]\}^\alpha}, \end{aligned} \quad (2.13)$$

$$\frac{\partial l}{\partial a} = \psi(a+b) - \psi(a) + \lambda \log \{1 - \exp[-(\lambda x)^c]\}, \quad (2.14)$$

$$\frac{\partial l}{\partial b} = \psi(a+b) - \psi(b) + \log \{1 - (1 - \exp[-(\lambda x)^c])^\alpha\}, \quad (2.15)$$

$$\begin{aligned} \frac{\partial l}{\partial c} &= \frac{1}{c} + \log(\lambda) + \log(x) - (\lambda x)^c \log(\lambda x) + (\alpha a - 1) \frac{(\lambda x)^c \log(\lambda x) \exp\{- (\lambda x)^c\}}{1 - \exp\{- (\lambda x)^c\}} \\ &\quad - (b-1) \alpha \frac{\{1 - \exp[-(\lambda x)^c]\}^{\alpha-1} (\lambda x)^c \log(\lambda x) \exp\{- (\lambda x)^c\}}{1 - \{1 - \exp[-(\lambda x)^c]\}^\alpha}, \end{aligned} \quad (2.16)$$

onde $\psi(\cdot)$ é a função *digama*. Então o estimador de máxima verossimilhança $\hat{\theta}$ de θ é obtido resolvendo o sistema de equações não-lineares $U_\alpha(\theta) = 0$, $U_\lambda(\theta) = 0$, $U_a(\theta) = 0$, $U_b(\theta) = 0$ e $U_c(\theta) = 0$.

Também é possível realizar um teste para verificar se a BEW pode ser reduzida a algum sub-modelo. Para tanto, obtém-se a *log-verossimilhança* restrita ao sub-modelo e a *log-verossimilhança* irrestrita e calcula-se a estatística de razão de *log-verossimilhança* (*TRV*). Um exemplo é o teste para avaliar se a BEW se reduz à distribuição Beta Weibull. Dessa forma, as hipóteses são

$$\begin{cases} H_0 : \alpha = 1, \\ H_1 : \alpha \neq 1. \end{cases}$$

Com isso, a estatística de teste é dada por

$$w = 2 \left\{ l(\hat{a}, \hat{b}, \hat{\lambda}, \hat{\alpha}, \hat{c}) - l(\tilde{a}, \tilde{b}, \tilde{\lambda}, 1, \tilde{c}) \right\}, \quad (2.17)$$

onde $\hat{a}, \hat{b}, \hat{\lambda}, \hat{\alpha}, \hat{c}$ são estimados sob H_1 e $\tilde{a}, 1, \tilde{\lambda}, \tilde{\alpha}, \tilde{c}$ são estimados sob H_0 .

2.3 Verossimilhança com dados censurados ao acaso

Seja t_i o tempo decorrido até a ocorrência de um dado evento em estudo para o i -ésimo indivíduo. Logo, a variável aleatória que designa o “tempo até a ocorrência do evento” será denotado por T_i . Considerando-se a variável indicadora

$$\delta_i = \begin{cases} 1, & \text{se } t_i \text{ é o “tempo” associado ao indivíduo } i, \\ 0, & \text{se o tempo está censurado,} \end{cases}$$

onde os tempos censurados serão considerados aleatórios, denotados por C_i . Dessa forma, t_i é, então, uma observação da variável aleatória $\tau_i = \min(T_i, C_i)$, em que

T_i, C_i são independentes. Considerando dados censurados, tem-se que

$$\begin{aligned} P(\tau_i = t, \delta_i = 0) &= P(C_i = t, T_i > t) \\ &= P(C_i = t) P(T_i > t), \end{aligned} \quad (2.18)$$

e para o caso em que os dados não são censurados

$$\begin{aligned} P(\tau_i = t, \delta_i = 0) &= P(C_i > t, T_i = t) \\ &= P(T_i = t) P(C_i > t), \end{aligned} \quad (2.19)$$

mas

$$\begin{aligned} P(\tau_i = t) &= P(\tau_i = t, \delta_i = 0) + P(\tau_i = t, \delta_i = 1) \\ &= P(C_i = t, T_i > t) + P(T_i = t, C_i > t) \\ &= P(C_i = t) P(T_i > t) + P(T_i = t) P(C_i > t). \end{aligned} \quad (2.20)$$

Logo,

$$P(\tau_i = t) = f_{C_i}(t)S_{T_i}(t) + f_{T_i}(t)S_{C_i}(t). \quad (2.21)$$

Considerando que o par (t_i, δ_i) é observado, então, para cada indivíduo, apenas uma das duas partes de (2.21) ocorre

$$\begin{cases} f_{C_i}(t)S_{T_i}(t), & \text{se } \delta_i = 0, \\ f_{T_i}(t)S_{C_i}(t), & \text{se } \delta_i = 1. \end{cases}$$

Logo, a verossimilhança associada a (τ_i, τ_n) é dada por

$$\begin{aligned} L(\theta) &= \prod_{i=1}^n P(\tau_i = t_i) \\ &= \prod_{i=1}^n [f_{T_i}(t_i)S_{C_i}(t_i)]^{\delta_i} [f_{C_i}(t_i)S_{T_i}(t_i)]^{1-\delta_i}. \end{aligned} \quad (2.22)$$

Logo,

$$L(\theta) = \prod_{i=1}^n [f_{C_i}(t_i)]^{1-\delta_i} [S_{C_i}(t_i)]^{\delta_i} \times \prod_{i=1}^n [f_{T_i}(t_i)]^{\delta_i} [S_{T_i}(t_i)]^{1-\delta_i}. \quad (2.23)$$

Sob o pressuposto de que a censura é não informativa, os parâmetros envolvidos no primeiro produtório não envolverá qualquer parâmetro que esteja relacionado no segundo produtório. Dessa forma, a verossimilhança dos dados observados é proporcional a

$$L(\theta) \propto \prod_{i=1}^n [f_{T_i}(t_i)]^{\delta_i} [S_{T_i}(t_i)]^{1-\delta_i}, \quad (2.24)$$

em que

$$S_{T_i}(t_i) = P(T_i > t) = 1 - P(T_i \leq t) = 1 - F_{T_i}(t). \quad (2.25)$$

2.3.1 Caso BBXII

A verossimilhança para dados censurados para a distribuição BBXII possui a forma dada em (2.24), onde

$$\begin{aligned} S_{T_i}(t_i) &= P(T_i > t) = 1 - P(T_i \leq t) = 1 - F_{T_i}(t) \\ &= 1 - I_{u_i}(a, b) = 1 - \frac{1}{B(a, b)} \int_0^{u_i} w^{a-1} (1-w)^{b-1} dw, \end{aligned} \quad (2.26)$$

em que $u_i = 1 - [1 + (\frac{x_i}{s})^c]^{-k}$. Seja $q_i = 1 - I_{u_i}(a, b)$. Substituindo (2.26) em (2.24), tem-se que

$$\begin{aligned} L(\theta) &\propto \prod_{i=1}^n [f_{T_i}(t_i)]^{\delta_i} [1 - I_{u_i}(a, b)]^{1-\delta_i} \\ &\propto \prod_{i=1}^n [f_{T_i}(t_i)]^{\delta_i} [q_i]^{1-\delta_i}. \end{aligned} \quad (2.27)$$

A *log-verossimilhança* (com $n = 1$) é dada por

$$l(\theta) = \delta \log(f_T(t)) + (1 - \delta) \log(q). \quad (2.28)$$

2.3.2 Caso BEW

A verossimilhança para os dados censurados para a distribuição BEW é da forma dada em (2.24), em que

$$\begin{aligned} S_{T_i}(t_i) &= P(T_i > t) = 1 - P(T_i \leq t) = 1 - F_{T_i}(t) \\ &= 1 - I_{r_i}(a, b) = 1 - \frac{1}{b(a, b)} \int_0^{r_i} w^{a-1} (1-w)^{b-1} dw, \end{aligned} \quad (2.29)$$

onde $r_i = [1 - \exp(-(\lambda x)^c)]^\alpha$. Seja $p_i = 1 - I_{r_i}(a, b)$. Substituindo (2.26) em (2.24), tem-se que

$$\begin{aligned} L(\theta) &\propto \prod_{i=1}^n [f_{T_i}(t_i)]^{\delta_i} [1 - I_{r_i}(a, b)]^{1-\delta_i} \\ &\propto \prod_{i=1}^n [f_{T_i}(t_i)]^{\delta_i} [p_i]^{1-\delta_i}. \end{aligned} \quad (2.30)$$

A *log-verossimilhança* (com $n = 1$) é dada por

$$l(\theta) = \delta \log(f_T(t)) + (1 - \delta) \log(p). \quad (2.31)$$

Capítulo 3

Inferência Bayesiana

Neste Capítulo, será apresentada uma introdução à teoria da Inferência Bayesiana. Na Seção 3.1, são descritos os aspectos teóricos da Inferência Bayesiana e seus objetivos. Na Seção 3.2 e 3.3 são apresentadas as estimativas bayesianas pontuais e intervalares. Na Seção 3.4, são apresentadas algumas técnicas para a obtenção da densidade marginal a posteriori pelo método de Monte Carlo via Cadeias de Markov (MCMC). São elas: os métodos de Metropolis-Hastings e Amostrador de Gibbs. Por fim, na Seção 3.5 são apresentadas algumas formas de avaliação das estimativas obtidas pelo MCMC.

3.1 Inferência Bayesiana

A Inferência Bayesiana é uma abordagem estatística em que toda e qualquer incerteza sobre um fenômeno em estudo é expresso em termos estocásticos. Inicia-se a descrição probabilística do problema com a formulação de um modelo estatístico adequado, estabelecendo-se as distribuições a priori dos parâmetros desconhecidos do modelo em questão. Isso tem por objetivo a descrição probabilística do conhecimento prévio do pesquisador acerca dos parâmetros do modelo.

Dessa forma, essa abordagem apresenta dois objetivos fundamentais: o primeiro é a realização de inferências sobre o parâmetro não observável θ ; o segundo é a possibilidade de, a partir de um conjunto de observações, realizar previsões sobre o conjunto de observações ainda não observados.

Diante disso, o primeiro objetivo é, ao mesmo tempo, coincidente e conflitante com a abordagem clássica. Pois, de um lado, a abordagem clássica é baseada em probabilidades associadas a diferentes amostras que poderiam ocorrer para algum valor fixo, mas desconhecido, do parâmetro θ . Por outro lado, o Bayesiano está baseada em probabilidades subjetivas, associadas a diferentes valores do parâmetro θ e condicionadas à amostra observada.

O'Hagan (1994) afirma que os Bayesianos podem emitir enunciados probabilísticos sobre os parâmetros, pois o consideram variáveis aleatórias, mas isso não se torna possível para os clássicos.

Nesse contexto, na abordagem Bayesiana, o valor de θ é desconhecido e aleatório. Portanto, θ possui uma distribuição de probabilidade associada. Assim, toda a informação disponível *a priori* em relação ao parâmetro θ é traduzida na forma de uma distribuição de probabilidade $\pi(\theta)$.

Vale ressaltar que $\pi(\theta)$ pode ser informativa ou não. Quando se tem algum conhecimento prévio sobre os parâmetros usa-se *prioris* informativas. Caso contrário, quando existe pouca ou nenhuma informação, usa-se as *prioris* não-informativas.

Considere $x=(x_1, x_2, \dots, x_n)^T$ um vetor com n observações, com vetor de parâmetros $\theta=(\theta_1, \theta_2, \dots, \theta_p)$. Assuma-se que a informação disponível sobre θ pode ser resumida probabilisticamente, de acordo com alguma função de densidade de probabilidade $f(x|\theta)$.

Dessa forma, combinando a informação *a priori* $\pi(\theta)$, que é o conhecimento prévio do pesquisador, e a verossimilhança (que é função dos dados) obteve-se a distribuição *a posteriori* $\pi(\theta|x)$, que é obtida pelo Teorema de Bayes:

$$\pi(\theta|x) = \frac{f_X(x|\theta)\pi(\theta)}{f_X(x)} \propto f_X(x|\theta)\pi(\theta), \quad (3.1)$$

onde

$$f_X(x) = \int f_X(X|\theta)\pi(\theta)d\theta \quad \text{e} \quad \int \pi(\theta|x)d\theta = 1. \quad (3.2)$$

Intuitivamente, observa-se que $\pi(\theta|x)$ incorpora toda e qualquer informação disponível sobre os parâmetros, além de ser usado como a regra de atualização para quantificar

o aumento de informação.

Em linhas gerais, a Inferência Bayesiana segue um esquema básico. Primeiramente, formula-se o modelo matemático (experimental), então de posse dos dados, a função de verossimilhança ($f(x|\theta)$), que mede toda informação sobre θ , é fornecida pelos dados. Define-se as distribuições *a priori* $\pi(\theta)$ sobre os parâmetros e, utilizando-se do Teorema de Bayes, obtêm-se a distribuição a posteriori.

Uma vez construída a posteriori, alguns interesses residem em:

- Avaliar o *efeito marginal* de um conjunto de parâmetros no modelo multivariado. Suponha que $\theta = (\theta_1, \dots, \theta_p)$ denote um modelo p -dimensional. Então, nosso interesse está em obter $\pi(\theta_i|x)$ onde $\theta_i \in \theta$, e

$$\begin{aligned}\pi(\theta_i|x) &= \int \pi(\theta_i, \theta_{-i}|x) d\theta_{-i} \\ &= \int \pi(\theta_i|\theta_{-i}, x) \pi(\theta_{-i}|x) d\theta_{-i}.\end{aligned}\tag{3.3}$$

- Previsões a posteriori de alguma quantidade de interesse são obtidas de forma natural. Denote por \tilde{x} a quantidade a ser predita, então a função preditiva a posteriori é dada por

$$\begin{aligned}\pi(\tilde{x}|x) &= \int \pi(\tilde{x}|\theta, x) \pi(\theta|x) d\theta \\ &= \int \pi(\tilde{x}|\theta) \pi(\theta|x) d\theta.\end{aligned}\tag{3.4}$$

Note que, \tilde{x} e x são condicionalmente independentes, dado θ . Na maioria dos casos, a posteriori obtida não possui distribuição conhecida e nem é trivial a sua estimativa por métodos não iterativos. O método de Monte Carlo via Cadeias de Markov representa uma boa alternativa, sendo a mais utilizada em problemas Bayesianos.

3.2 Estimativa Pontual

Sabe-se que a distribuição a posteriori de um parâmetro θ contém toda a informação probabilística do referido parâmetro. Dessa forma, se faz necessário resumir a informação contida no parâmetro através de valores numéricos, entre eles, as

medidas de tendência central.

Na abordagem bayesiana, adota-se o conceito de Função Perda, como auxílio, na escolha do estimador de θ , enquanto na abordagem clássica o problema é resolvido, por exemplo, encontrando-se o estimador de máxima verossimilhança da quantidade desconhecida (porém fixa) θ .

Segundo Ehlers (2007), a Função Perda, denotada por $L(\alpha, \theta)$, determina a perda sofrida ao se tornar a decisão α dado o real estado $\theta \in \Theta$. Esta perda é expressa como um número real e é definida por

$$E [L(\alpha, \theta|y)] = \int L(\alpha, \theta|y)h(\theta|y)\partial\theta, \quad (3.5)$$

onde α é escolhida de tal forma que a perda esperada a posteriori seja minimizada. α é chamado de estimador de Bayes e depende da função perda que é adotada.

Se a Função Perda é definida como $L(\alpha - \theta) = (\alpha - \theta)^2$ então o estimador de Bayes é a média a posteriori. Se adotar a Função Perda absoluta ($L(\alpha - \theta) = |\alpha - \theta|$) tem-se que o estimador de Bayes é a mediana a posteriori. Se a Função Perda denominada 0-1 é adotada, o estimador de Bayes será a moda a posteriori. Esse estimador de θ também é chamado de estimador de máxima verossimilhança generalizado e é o mais de fácil de ser obtido dentre os apresentados. No caso contínuo, deve-se obter a solução da seguinte equação

$$\frac{\partial \pi(\theta|x)}{\partial \theta} = 0. \quad (3.6)$$

3.3 Estimativa Intervalar

No entanto, a principal restrição da estimação pontual se deve ao resumo de toda a informação da distribuição a posteriori a um único número. Dessa forma, faz-se necessária associar alguma informação sobre o quão precisa é a especificação deste número, objetivo do intervalo de credibilidade, ou “intervalos de confiança” bayesianos.

Definição 3.3.1. *C é um intervalo de credibilidade de $100(1 - \alpha)\%$, ou nível de credibilidade (ou de confiança) $1 - \alpha$, para θ se $P(\theta \in C) \geq 1 - \alpha$.*

Pela definição acima é possível obter uma infinidade de intervalos, mas o interesse está naquele intervalo que possui o menor comprimento possível. Segundo Ehlers (2007), pode-se mostrar que os intervalos de comprimento mínimo são obtidos tomando-se os valores de θ com maior densidade a posteriori. Esses intervalos são denominados intervalos Highest Posterior Density (HPD).

Definição 3.3.2. *Intervalo HPD: Um intervalo de credibilidade C de $100(1 - \alpha)\%$ para θ é de máxima densidade a posteriori se $C = \{\theta \in \Theta : \pi(\theta|x) \geq k(\alpha)\}$ onde $k(\alpha)$ é a maior constante tal que $P(\theta \in C) \geq 1 - \alpha$.*

Por essa definição, todos os pontos dentro dessa intervalo terão maior densidade do que qualquer ponto fora dele. Além disso, no caso de distribuição bicaudal (exemplo: Normal, t de Student) é obtido de modo que as caudas tenham a mesma probabilidade. De forma equivalente a abordagem clássica, esses intervalos de credibilidades não são invariantes a transformações 1 a 1, a não ser para transformações lineares.

Se a distribuição em estudo for simétrica e unimodal, esse intervalo corresponde ao intervalo simétrico. Se for assimétrica e unimodal corresponde ao intervalo cuja a posteriori apresenta o mesmo valor nos limites. Em geral, nesse caso, o HPD não é obtido numericamente e, assim, deve-se empregar métodos numéricos para construí-los.

3.4 MCMC

O método de Monte Carlo via Cadeias de Markov (MCMC) tem a finalidade de obter uma amostra da distribuição a posteriori, com o objetivo de calcular as estimativas amostrais das características da distribuição em estudo. Esse é um método alternativo aos métodos não iterativos em problemas complexos. Esse método se constitui numa integração numérica baseada em simulações estocásticas.

Os métodos de Metropolis-Hastings (Metropolis et. al., 1953; Hastings, 1970) e amostrador de Gibbs (Geman e Geman, 1984; Gelfand e Smith, 1990) são tipos específicos do MCMC. Em linhas gerais, nesses métodos, simula-se um passeio aleatório no espaço de θ que convergirá para uma distribuição estacionária, que é a distribuição de interesse no problema. Diante disso, Kruschke(2014) propõe 3 objetivos principais

na geração de uma amostra da distribuição a posteriori pelo MCMC:

- Os valores da cadeia devem ser representativos da distribuição a posteriori. Eles não devem ser influenciados, de forma indevida, por valores iniciais arbitrários da cadeia e, ainda, devem varrer o espaço paramétrico do parâmetro em estudo;
- A cadeia deve ter tamanho suficiente para que as estimativas se tornem precisas e estáveis. Principalmente, as estimativas de tendência central (tais como a mediana ou moda), além dos limites do intervalo HPD 95%, não devem ser muito diferentes se a análise MCMC é executada novamente;
- A cadeia deve ser gerada de forma eficiente com tão poucos passos quanto possível, de modo que não exceda o poder de computação;

Diante disso, esse algoritmo de MCMC nos dá um método geral para construir uma cadeia de Markov com distribuição estacionária, dada por uma função de probabilidade arbitrária $f(x)$ (geralmente é a distribuição de interesse, ou alvo). Na maioria das vezes a complexidade de $f(x)$ faz com que a geração de dados a partir dela não seja trivial.

3.4.1 Metropolis-Hastings

O método de Metropolis-Hastings é a junção de dois trabalhos, o primeiro de Metropolis et al. (1953) e o segundo de Hastings (1970). Estes são considerados como os principais artigos para a caracterização do método, embora outros apresentaram contribuições relevantes. O primeiro artigo foi proposto para resolver um problema da área da físico-química e somente mais tarde foi verificado o grande impacto na área da Estatística.

Em termos gerais, esse método usa a mesma ideia dos métodos de rejeição, de forma que um valor é gerado (θ') de uma distribuição $q(.|\theta)$ e é aceito de acordo com uma dada probabilidade α , dada por

$$\alpha(\theta, \theta') = \min \left\{ 1, \frac{\pi(\theta')q(\theta|\theta')}{\pi(\theta)q(\theta'|\theta)} \right\}, \quad (3.7)$$

onde $\pi()$ é a distribuição de interesse. Esse método é de grande importância nos casos em que a distribuição a posteriori não é conhecida.

Dessa forma, o algoritmo de Metropolis-Hastings pode ser realizado conforme os seguintes passos:

- 1. Inicialize o contador de iterações $t = 0$ e especifique um valor inicial θ^0 ;
- 2. Gere um novo valor θ' da distribuição $q(\cdot|\theta)$;
- 3. Calcule a probabilidade de aceitação $\alpha(\theta, \theta')$ e gere $U \sim \text{Uniforme}(0, 1)$.
- 4. Se $u < \alpha$, então o novo valor deve ser aceito ($\theta^{t+1} = \theta'$), caso contrário rejeite e faça $\theta^{t+1} = \theta$;
- 5. Faça $t = t + 1$ e volte ao passo 2.

3.4.2 Amostrador de Gibbs

O Amostrador de Gibbs (Geman e Geman,1984; Gelfand e Smith,1990) é um caso especial do algoritmo de Metropolis-Hastings com probabilidade de aceitação igual a 1. Logo, é um esquema MCMC onde as transições são dadas por condicionais completas com formas totalmente conhecidas.

A ideia básica de Gibbs é a possibilidade de tornar um problema multivariado numa sequência de problemas univariados, para o quais itera-se e assim produza uma Cadeia de Markov. Isso nos mostra que a mudança de estado depende apenas do estado atual. Temos também que a distribuição de equilíbrio é a distribuição a posteriori desejada.

Assumindo que a distribuição conjunta de interesse é dada por $\pi(\theta) = \pi(\theta_1, \dots, \theta_d)'$ e que pode ser caracterizada pelas densidade condicionais completas $\pi_i(\theta_i) = \pi_i(\theta_i|\theta_{i-1})$, $i = 1, \dots, d$. Logo, o algoritmo de Gibbs pode ser descrito da seguinte maneira:

- 1. Inicialize o contador de iteração da cadeia e o valor inicial $\theta^0 = (\theta_1^0, \theta_2^0, \dots, \theta_d^0)$;
- 2. obtenha os novos valores $\theta^j = (\theta_1^j, \theta_2^j, \dots, \theta_d^j)$ dado $\theta^{(j-1)}$ através da geração

sucessiva de valores

$$\begin{aligned}\theta_1^j &= \pi(\theta_1|\theta_2^{j-1}, \dots, \theta_d^{j-1}) \\ \theta_2^j &= \pi(\theta_2|\theta_1^j, \theta_2^{j-1}, \dots, \theta_d^{j-1}) \\ &\vdots \\ \theta_d^j &= \pi(\theta_d|\theta_1^j, \theta_2^j, \dots, \theta_{d-1}^j); \end{aligned}$$

- 3. Fazer $j \leftarrow j + 1$ e retorne ao passo 2 até que a convergência seja alcançada;

Quando a convergência é alcançada, a amostra de θ converge para a distribuição estacionária. Ressalta-se que, mesmo num problema multidimensional, em alguns casos, todas as simulações podem ser univariadas, o que pode ser uma vantagem computacional.

3.5 Avaliação do MCMC

Nesta seção, objetiva-se apresentar alguns métodos para avaliar a convergência das estimativas obtidas pelo MCMC. Na literatura, existem várias estratégias. Aqui serão apresentados os métodos mais utilizados.

3.5.1 Análise Gráfica

Na análise gráfica, primeiramente, temos que supor a estratégia de cadeias múltiplas (com valores iniciais diferentes) no método de MCMC. Dessa forma, as observações serão geradas e a trajetória monitorada ao longo dos passos.

Com isso, a análise será realizada através da análise gráfica da sobreposição da trajetória das cadeias. Se as estimativas forem graficamente indistinguíveis, existe indicativo de convergência (Gelfand et. al., 1990).

3.5.2 Método de Gelman e Rubin

O segundo método apresentado se dá pela descrição numérica da convergência. Gelman e Rubin (1992) propuseram um critério baseado na análise de variância. Em

linhas gerais, nesse método utiliza-se duas ou mais cadeias paralelas, com diferentes pontos iniciais e realiza-se a comparação da variância dentro e entre as cadeias.

Suponha que existem m cadeias paralelas, de tamanho n cada. Seja ϕ_{ij} , com $i = 1, \dots, n$ e $j = 1, \dots, m$, onde i representa a observação e j a cadeia. Denotando V_D e V_E as variâncias *entre* e *dentro* das cadeias, respectivamente, então

$$V_E = \frac{n}{m-1} \sum_{j=1}^m (\bar{\phi}_{.j} - \bar{\phi})^2 \quad (3.8)$$

e

$$V_D = \frac{1}{n(m-1)} \sum_{i=1}^m \sum_{j=1}^n (\phi_{ij} - \bar{\phi}_{.j})^2, \quad (3.9)$$

onde $\bar{\phi}_{.j}$ é a média das observações da cadeia j , e $\bar{\phi}$ é a média dessas médias. Logo, a variância marginal a posteriori de ϕ ($V(\hat{\phi}|y)$) é estimada pela média ponderada de V_E e V_D

$$V(\hat{\phi}|y) = \frac{n-1}{n} V_D + \frac{1}{n} V_E. \quad (3.10)$$

Então, o monitoramento da convergência da cadeia é denotado "fator de redução potencial de escala" (R),

$$\hat{R} = \sqrt{\frac{V(\hat{\phi}|y)}{V_D}}. \quad (3.11)$$

de forma que a medida que n cresce, \hat{R} tende a 1. Se $\hat{R} \approx 1$, tem-se o indicativo de que o período de *burn-in* pode terminar e o processo iterativo pode continuar.

3.5.3 Método de Geweke

Esse método é baseado em aplicações de técnicas usuais de séries temporais para averiguar a convergência do método de MCMC (Geweke, 1992). Seja $g(\theta)$ a função do parâmetro de interesse. Sabe-se que o objetivo do MCMC é atingido quando simula-se valores de θ^t de uma determinada Cadeia de Markov.

O valor esperado da estimativa marginal a posteriori de $g(\theta)$ é dada pela média

ergódica dos $g(\theta^{(j)})$. Com isso, considerando que $g(\theta)$ é uma função real e sua trajetória g^1, g^2, \dots , obtida a partir de $g^t = g(\theta^t)$, define uma série temporal.

Assim, observa-se a série ao longo de n simulações, suficientemente grande, e calcula-se a média das n_α primeiras iterações $g_\alpha = \frac{1}{n_\alpha} \sum_{t=1}^{n_\alpha} g(\theta^t)$, bem como a média das últimas n_β iterações ($g_\beta = \frac{1}{n_\beta} \sum_{t=n-n_\beta+1}^n g(\theta^t)$).

Logo, se a cadeia é estacionária, então a média da primeira parte da cadeia (das n_α iterações) deve ser semelhante à média da segunda parte da cadeia das últimas n_β iterações. Então, se $n \rightarrow \infty$ temos que

$$\frac{(g_\alpha - g_\beta)}{\sqrt{(s_\alpha^2/n_\alpha) + (s_\beta^2/n_\beta)}} \longrightarrow N(0, 1), \quad (3.12)$$

admitindo n_α e n_β fixos e s_α^2 e s_β^2 são estimativas independentes das variâncias assintóticas de g_α e g_β . De acordo com o valor da estatística dada por (3.12) pode se verificar se ocorre ou não a convergência.

3.5.4 Taxa de Aceitação

Por final, pode-se verificar a taxa de aceitação do algoritmo MCMC (no caso do Metropolis-Hastings) para cada parâmetro. Gelman et. al. (2003) propõem que a taxa de aceitação do algoritmo deve ser próxima de 30%.

Capítulo 4

A Abordagem Bayesiana via JAGS e WinBUGS

A seguir, apresenta-se todas as tentativas desenvolvidas para aprimorar o processo de estimação dos parâmetros da distribuição BBXII utilizando a abordagem Bayesiana. Em cada seção, apresenta-se uma descrição da metodologia a ser exposta, bem como os resultados obtidos e uma breve discussão sobre as limitações e vantagens do método em estudo.

Na Seção 4.2 são apresentados os procedimentos utilizados para implementar a distribuição BBXII nos softwares *JAGS* e *WinBUGS*, bem como a estimação e os resultados.

No Apêndice são apresentadas algumas tentativas de processos de estimação bayesiano para os quais não logrou-se sucesso. No Apêndice A, é apresentada uma possibilidade de inclusão, ao algoritmo de MCMC, dos momentos da distribuição BBXII e os resultados das simulações. No Apêndice B, é apresentada uma abordagem baseada em McCormik (2012).

Na Apêndice C é apresentada uma forma alternativa de obtenção das estimativas bayesianas dos parâmetros relacionados à distribuição Beta(a , b) por uma aproximação da moda a posteriori pelo estimador de máxima verossimilhança da função de distribuição condicional completa bivariada do vetor (a, b) e os resultados obtidos.

4.1 Distribuição a Posteriori e Condicionais Completas

A distribuição a posteriori dos parâmetros associados ao modelo BBXII é dada por

$$\pi(a, b, s, k, c|X) \propto \left[\prod_{i=1}^n f(x_i|a, b, s, k, c) \right] \pi(a, b, s, k, c), \quad (4.1)$$

onde $f(x_i)$ é dado pela expressão (1.6).

Assuma que os parâmetros do modelo sejam independentes *a priori* de modo que

$$\pi(a, b, s, k, c) = \pi(a)\pi(b)\pi(s)\pi(k)\pi(c). \quad (4.2)$$

As distribuições condicionais completas a posteriori de cada um dos parâmetros são dadas por:

$$\begin{aligned} \pi(a|b, s, k, c, x) &\propto \left[\prod_{i=1}^n f(x_i|a, b, s, k, c) \right] \times \pi(a) \\ &\propto [B(a, b)]^{-n} \prod_{i=1}^n \left(1 - [1 + (x_i/s)^c]^{-k} \right)^a \times \pi(a), \end{aligned} \quad (4.3)$$

$$\begin{aligned} \pi(b|a, s, k, c, x) &\propto \left[\prod_{i=1}^n f(x_i|a, b, s, k, c) \right] \times \pi(b) \\ &\propto [B(a, b)]^{-n} \prod_{i=1}^n [1 + (x_i/s)^c]^{-(kb)} \times \pi(b), \end{aligned} \quad (4.4)$$

$$\begin{aligned} \pi(s|a, b, k, c, x) &\propto \left[\prod_{i=1}^n f(x_i|a, b, s, k, c) \right] \times \pi(s) \\ &\propto s^{-nc} \prod_{i=1}^n [1 + (x_i/s)^c]^{-(kb+1)} \left(1 - [1 + (x_i/s)^c]^{-k} \right)^{a-1} \times \pi(s), \end{aligned} \quad (4.5)$$

$$\begin{aligned}\pi(k|a, b, s, c, x) &\propto \left[\prod_{i=1}^n f(x_i|a, b, s, k, c) \right] \times \pi(k) \\ &\propto k^n \prod_{i=1}^n [1 + (x_i/s)^c]^{-(kb)} \left(1 - [1 + (x_i/s)^c]^{-k} \right)^{a-1} \times \pi(k),\end{aligned}\quad (4.6)$$

$$\begin{aligned}\pi(c|a, b, s, k, x) &\propto \left[\prod_{i=1}^n f(x_i|a, b, s, k, c) \right] \times \pi(c) \\ &\propto c^n s^{-nc} \prod_{i=1}^n x_i^c [1 + (x_i/s)^c]^{-(kb+1)} \left(1 - [1 + (x_i/s)^c]^{-k} \right)^{a-1} \times \pi(c).\end{aligned}\quad (4.7)$$

4.2 JAGS e WinBUGS

Uma primeira implementação do problema de estimação Bayesiana da distribuição BBXII foi realizada utilizando-se os programas JAGS e WinBUGS, considerando-se *prioris Gama* independentes para todos os parâmetros do modelo. No entanto, não foi possível trabalhar com tais *prioris* devido a problemas numéricos. Então aplicou-se a transformação *log* nos parâmetros, na tentativa de evitar tais problemas numéricos. Com os parâmetros expressos na escala *log*, trabalhou-se com *prioris* uniformes restritas ao intervalo $(-5, 5)$.

O primeiro desafio foi encontrar a forma correta de programar uma nova distribuição (não padrão) nos referidos programas. Isso porque a distribuição BBXII ainda não está implementada em ambos os *softwares*.

Depois de algum tempo de pesquisa, constatou-se que a forma correta para a implementação de um modelo não padrão, assim como a inserção de dados, e também para obtenção dos resultados, deve seguir os procedimentos descritos por Lunn et. al. (2013). Tanto no JAGS quanto no WinBUGS, utiliza-se o que se denomina “Bernoulli ones trick” ou “Poisson zeros trick” (e.g, Lunn et al., 2013).

Novas funções de verossimilhança podem ser descritas de maneira semelhante nos dois softwares, conforme o Código 4.1. A única diferença é que no JAGS deve ser inserida mais uma linha com o comando `#monitor#c, s, k, a, b`, que tem por objetivo informar quais são os parâmetros de interesse no MCMC.

Código 4.1: Modelo BBurrXII no JAGS e no WinBUGS

```

model{
2 for(i in 1:N){

```

```

zeros[i]<-0
4 phi[i]<- -log(y[i])
zeros[i] ~ dpois(phi[i])
6 y[i]<-exp(loggam(exp(r1)+exp(r2))-loggam(exp(r1))
-loggam(exp(r2))+r5+r3-exp(r5)*r4+
8 (exp(r5)-1)*log(x[i])-(exp(r3)*exp(r2)+1)*log(1+
pow(x[i]/exp(r4),exp(r5))) +
10 (exp(r1)-1)*log(1-pow(1+pow(x[i]/exp(r4),exp(r5)),
-exp(r3))))
12 }
r1~dunif(-3,3)
14 r2~dunif(-3,3)
r3~dunif(-5,0)
16 r4~dunif(-5,0)
r5~dunif(-5,5)
18 a<-exp(r1)
b<-exp(r2)
20 k<-exp(r3)
s<-exp(r4)
22 c<-exp(r5)
}

```

Como mencionado anteriormente, a forma de inserção do novo modelo é equivalente nos dois *softwares*, que diferem na maneira de inserir os dados e na obtenção dos resultados.

4.2.1 Dados e Resultados no WinBUGS

O Código 4.2 ilustra a sintaxe de inserção dos dados no WinBUGS, a título de exemplo. Para a obtenção dos resultados, os dados foram gerados no software *R*, de acordo com o Código 1.1, da Seção 1.1.1, com tamanho 300.

Código 4.2: Comandos para inserção de dados no WinBUGS

```

1 Data
list(N=10, x=c( 0.7997404 ,0.7836458 ,0.6757031 ,0.8283034
,0.7734923, 0.8068642 ,0.7907033,0.8234490 ,0.7720038 ,0.7909191)

```

Como já mencionamos, trabalhou-se com os parâmetros do modelo expressos na es-

cala logarítmica, para evitar problemas numéricos. Além disso, utilizou-se *prioris* uniformes independentes. Achcar (2013) obtém estimativas bayesianas dos parâmetros de uma nova distribuição utilizando “zero’s trick” no WinBUGS e considerando *prioris* independentes uniformes.

Dentre todas as tentativas, o resultado mais coerente está apresentado na Tabela 4.1. Isso porque foi verificado uma grande sensibilidade na escolha dos hiperparâmetros das *prioris* uniformes. Nota-se que as estimativas dos parâmetros apresentadas na Tabela 4.1 não são “próximas” dos verdadeiros valores em todos os casos. Ainda foram observados problemas de autocorrelação nas estimativas.

Tabela 4.1: Estimativas do WinBUGS com transformação *log*, *prioris* Uniformes e dados gerados com $a=16$, $b=8$, $k=6.0$, $s=1.0$, $c=5.0$

$n = 300$	Valor real	Média	Desvio Padrão	2,5%	50,0%	97,5%
a	16,0	1,702	0,864	0,584	1,502	4,013
b	8,0	10,280	5,102	2,259	9,965	19,460
c	6,0	23,760	8,017	12,780	22,300	44,510
k	5,0	0,524	0,250	0,099	0,522	0,965
s	1,0	0,835	0,024	0,800	0,831	0,889

Vale ressaltar que tentou-se implementar no *Winbugs* os cálculos de forma direta (sem o uso de transformações nos parâmetros), mas isso não foi possível na maioria dos casos. O programa indicava que os resultados não eram definidos, com a possibilidade de erros numéricos dos tipos *overflow* ou *underflow*. As razões da ocorrência de tais erros, conforme a mensagem no *software*, estão relacionados a seguir:

- Valores iniciais gerados a partir de uma distribuição a priori “vaga” podem ser numericamente extremos - especificar os valores iniciais adequados;
- Valores numericamente impossíveis, tais como registro de um número não-positivo;
- Dificuldades numéricas em amostragem. As soluções possíveis incluem: melhores valores iniciais; *prioris* mais informativas ou utilização de *prioris* uniformes, mas com a sua gama restrita a valores plausíveis; melhor parametrização para melhorar ortogonalidade; padronização de covariáveis a ter média 0 e desvio padrão 1.

Observa-se, de modo geral, que os resultados obtidos não foram satisfatórios, pois mesmo com valores iniciais semelhantes aos verdadeiros, as estimativas obtidas são totalmente diferentes dos verdadeiros, ainda com o aumento considerável do número de simulações. Isso pode estar relacionado com problemas numéricos. Além das *prioris Uniformes*, também foram investigadas *prioris Gama* para todos os parâmetros. No entanto, os resultados obtidos foram bastante insatisfatórios.

4.2.2 Dados e Resultados no JAGS

No JAGS, em continuidade aos Códigos 4.1 e obedecendo a forma de geração de dados da Seção 1.1.1, os dados simulados e os resultados podem ser obtidos através do Código 4.3. Vale ressaltar que, primeiramente, deve ser instalado o pacote *runjags* no R. Utilizou-se 35.000 simulações, uma cadeia e *burn-in* de tamanho 4.000, que são as mesmas especificações utilizadas no WinBUGS.

Código 4.3: Comandos para inserção de dados e obtenção dos resultados no JAGS

```
require(runjags)
2 a<-16
  b<-8
4 k<-5
  s<-1
6 cc<-6
  x<-Ger.VA.BURRXII(a,b,k,s,cc,n)
8 N<-length(x)
  const <- 10000
10 z<-rep(0,N)
  dados.jags<-list('x' = x,'N' = N,'const'=const,'z'=z)
12 model <- read.jagsfile(string)
  results <- run.jags(string, n.chains=1,data=dados.jags,
14 method="rjags",sample=35000)
  summary(results$mcmc)
```

Os resultados obtidos no JAGS são apresentados na Tabela 4.2. Verifica-se que as estimativas não são próximas dos valores verdadeiros dos parâmetros em todos os casos, exceto para o parâmetro s .

Diante das dificuldades, dos resultados insatisfatórios e de limitações dos *softwares*

Tabela 4.2: Estimativas do JAGS com Transformação *log, priors* Uniformes.

$n = 300$	Valor real	Média	Desvio Padrão	2,5%	50,0%	97,5%
a	16,0	2,700	0,017	2,653	2,705	2,717
b	8,0	2,463	0,210	1,935	2,516	2,710
c	6,0	25,396	19,800	3,857	18,219	69,431
k	5,0	0,038	0,033	0,007	0,026	0,127
s	1,0	0,994	0,005	0,978	0,995	0,999

WinBUGS e JAGS, no que diz respeito à falta de flexibilidade para implementação de modelagens Bayesianas mais complexas, que fogem a certos padrões no uso desses programas, optou-se pela implementação de programas utilizando o software R.

Capítulo 5

Implementações Bayesianas da distribuição BBXII utilizando o R

Nesta seção, apresenta-se alguns desenvolvimentos Bayesianos para a estimação dos parâmetros das distribuições BBXII. Utilizou-se o *software* R para o desenvolvimento de scripts a serem utilizados nas estimações frequentistas e Bayesianas da distribuição BBXII.

5.1 Utilização de prioris Gama

Considerando as expressões (4.1) a (4.7), nesta seção apresenta-se o desenvolvimento da proposta Bayesiana para a estimação dos parâmetros da distribuição BBXII, utilizando-se *prioris* Gama para os parâmetros a , b , k e c e distribuição Gama inversa para o parâmetro s , uma vez que todos esses parâmetros são não-negativos. Dessa forma, as *prioris* descritas nas equações (4.3) a (4.7) são tais que

$$a \sim \text{Gama}(a_1, b_1), \quad (5.1)$$

$$b \sim \text{Gama}(a_2, b_2), \quad (5.2)$$

$$s \sim \text{Gama} - \text{Inv}(a_3, b_3), \quad (5.3)$$

$$k \sim \text{Gama}(a_4, b_4), \quad (5.4)$$

$$c \sim \text{Gama}(a_5, b_5). \quad (5.5)$$

Tais *prioris*, quando substituídas nas equações (4.3) a (4.7) auxiliam na obtenção de propostas para a definição de funções geradoras de candidatos a serem utilizadas nos procedimentos MCMC, para cada um dos parâmetros acima.

Para tanto, a estratégia foi utilizar aproximações para as distribuições condicionais completas obtidas com as escolhas mencionadas de *prioris* que conduzam a distribuições conhecidas e fáceis de simular. A seguir, apresenta-se os referidos desenvolvimentos para cada parâmetro do modelo.

5.1.1 Proposta de Função Geradora de Candidatos para o Parâmetro a

Tome $G(x, s, k, c) = 1 - (1 + (x/s)^c)^{-k} = v$ como a função de distribuição acumulada de uma variável aleatória BBXII. Logo, a distribuição condicional completa de a é dada por:

$$\pi(a|b, s, k, c, x) \propto [B(a, b)]^{-n} \left[\prod_{i=1}^n v_i^a \right] a^{a_1-1} e^{-b_1 a}. \quad (5.6)$$

Tomando $h_i = \log(v_i)$, então,

$$\begin{aligned} \pi(a|b, s, k, c, x) &\propto [B(a, b)]^{-n} \left[\prod_{i=1}^n \exp(a h_i) \right] a^{a_1-1} e^{-b_1 a} \\ &\propto [B(a, b)]^{-n} \exp \left(a \sum_{i=1}^n h_i \right) a^{a_1-1} e^{-b_1 a} \\ &\propto [B(a, b)]^{-n} a^{a_1-1} \exp \left(- \left[b_1 - \sum_{i=1}^n h_i \right] a \right) \\ &\approx a^{a_1-1} \exp \left(- \left[b_1 - \sum_{i=1}^n h_i \right] a \right), \quad \text{se } b = 1. \end{aligned} \quad (5.7)$$

Como a distribuição Gama com parâmetros (α, β) é dada por

$$P(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad (5.8)$$

tentaremos a função geradora de candidatos de a dada por:

$$P(a) = {}^d Gama \left(a_1, b_1 - \sum_{i=1}^n h_i \right) = Gama \left(a_1, b_1 - \sum_{i=1}^n \log(v_i) \right). \quad (5.9)$$

5.1.2 Proposta de Função Geradora de Candidatos para o Parâmetro b

Agora a função de distribuição condicional completa de b é dada por:

$$\pi(b|a, s, k, c, x) \propto [B(a, b)]^{-n} \left[\prod_{i=1}^n (1 + (x_i/s)^c)^{-(kb)} \right] \times b^{a_2-1} e^{-b_2 b}. \quad (5.10)$$

Como $v = G(x, s, k, c) = 1 - (1 + (x/s)^c)^{-k}$, logo $1 - v = 1 - G(x, s, k, c) = (1 + (x/s)^c)^{-k}$.

Portanto,

$$\begin{aligned} \pi(b|a, s, k, c) &\propto [B(a, b)]^{-n} \left[\prod_{i=1}^n (1 - v_i)^b \right] x b^{a_2-1} e^{-b_2 b} \\ &\propto [B(a, b)]^{-n} \left[\prod_{i=1}^n \exp(b \log(1 - v_i)) \right] b^{a_2-1} e^{-b_2 b}. \end{aligned} \quad (5.11)$$

Então,

$$\begin{aligned} \pi(b|a, s, k, c, x) &\propto [B(a, b)]^{-n} b^{a_2-1} \exp \left\{ - \left[b_2 - \sum_{i=1}^n \log(1 - v_i) \right] b \right\} \\ &\approx b^{a_2-1} \exp \left\{ - \left[b_2 - \sum_{i=1}^n \log(1 - v_i) \right] b \right\}, \quad \text{se } a = 1. \end{aligned} \quad (5.12)$$

Portanto, tentaremos a função geradora de candidatos de b dada por:

$$P(b) = {}^d Gama \left(a_2, b_2 - \sum_{i=1}^n \log(1 - v_i) \right). \quad (5.13)$$

5.1.3 Proposta de Função Geradora de Candidatos para o Parâmetro k

Para o parâmetro k , temos que, de acordo com Paranaíba et. al. (2011), se $a = b = c = 1$, a distribuição BBXII reduz-se à distribuição Beta Pareto, que tem a

seguinte expressão:

$$f(x|s, k) = \frac{k}{s} [1 + (x/s)]^{-(k+1)}. \quad (5.14)$$

Considerando a distribuição condicional completa de k baseada na distribuição Beta Pareto e uma *priori* Gama para o parâmetros k , tem-se que:

$$P(k|s, x) \propto k^n \left[\prod_{i=1}^n (1 + (x_i/s))^{-k} \right] \times k^{a_4-1} e^{-b_4 k}. \quad (5.15)$$

Seja $z_i = 1 + (x_i/s)$. Então,

$$\begin{aligned} P(k|s, x) &\propto k^n \exp \left\{ -k \sum_{i=1}^n \log(z_i) \right\} x k^{a_4-1} e^{b_4 k} \\ &\propto k^{n+a_4-1} \exp \left\{ -(b_4 + \sum_{i=1}^n \log(z_i)) k \right\}, \end{aligned} \quad (5.16)$$

que é o *kernel* de uma distribuição Gama. Logo, tentaremos como função geradora de k

$$P(k) = {}^d \text{Gama} \left(n + a_4, \quad b_4 + \sum_{i=1}^n \log(z_i) \right). \quad (5.17)$$

5.1.4 Proposta de Função Geradora de Candidatos para o Parâmetro s

Segundo Paranaíba et. al. (2011), se $a = b = 1$ e $k \rightarrow \infty$, a distribuição BBXII reduz-se à distribuição Weibull:

$$f(x|c, s) = \frac{c}{s} \left(\frac{x}{s} \right)^{c-1} \exp \left(- \left[\frac{x}{s} \right]^c \right). \quad (5.18)$$

Considerando uma distribuição condicional completa baseada na distribuição Weibull e uma priori com distribuição Gama para s , temos que:

$$\begin{aligned}
P(s|c, x) &\propto s^{-n} \left[\prod_{i=1}^n \left(\prod_{i=1}^n (x_i/s)^{c-1} \right) \exp \left(- \left[\frac{x_i}{s} \right]^c \right) \right] \times s^{-a_3-1} e^{-b_3/s} \\
&\propto s^{-n} x s^{-n(c-1)} \exp \left(-s^{-c} \sum_{i=1}^n x_i \right) s^{-a_3-1} e^{-b_3/s} \\
&\propto s^{-nc-a_3-1} \exp \left(-s^{-c} \sum_{i=1}^n x_i - s^{-1} b_3 \right). \tag{5.19}
\end{aligned}$$

No caso de $c = 1$, temos que

$$P(s|x) \propto s^{-(n+a_3)-1} \exp \left(-s^{-1} \left(\sum_{i=1}^n x_i + b_3 \right) \right), \tag{5.20}$$

que é o kernel de uma distribuição Gama-Inversa. Logo,

$$S \sim \text{Gama-Inversa} \left(n + a_3, b_3 + \sum_{i=1}^n x_i \right). \tag{5.21}$$

5.1.5 Proposta de Função Geradora de Candidatos para o Parâmetro c

No caso do parâmetro c , não foi possível obter nenhuma forma funcional conhecida para ser utilizada como função geradora de candidatos. Nesse caso, utilizaremos a proposta de Hoff (2009, pg 214), que é baseada na construção de uma função geradora de candidatos utilizando-se a distribuição Normal. Em linhas gerais, a proposta de Hoff é dada a seguir.

Seja $X \sim N(\mu, \sigma^2)$. Para gerar dados sujeitos à restrição de $x \in [a, b]$, então faça:

a. $u \sim \text{Unif}(\Phi[(a^* - \mu)/\sigma], \Phi[(b^* - \mu)/\sigma])$

b. $x = \mu + \sigma \Phi^{-1}(u)$;

Seja θ um parâmetro não-negativo, isto é, $\theta \in (0, \infty)$. Como $\Phi(\infty) = 1$ e $a > 0$,

$a^* \leftarrow 0.000001$

$u \leftarrow \text{runif}(1, \text{pnorm}(a^* - \mu), 1)$

$x \leftarrow \mu + \text{qnorm}(u)$,

onde μ é o último valor aceito de x no MCMC. No R, a ideia do algoritmo acima para o parâmetro c da distribuição BBXII é dado pelo Código 5.1.

Código 5.1: Algoritmo para gerar um candidato pela aproximação da Normal com restrição no intervalo (a, b) .

```

1 a.star<-0.000001
  cc.old<-cc
3 u<-runif(1,pnorm(a.star-cc.old),1)
  cc.cand<-cc.old+qnorm(u)

```

5.1.6 Resultados

Implementando-se os desenvolvimentos anteriores num algoritmo MCMC, verificou-se que as estimativas dos parâmetros k e s foram ruins, comparativamente às obtidas para os demais parâmetros, que foram, de fato, inaceitáveis.

A proposta também conduz a forte autocorrelação entre os valores gerados nas cadeias de MCMC. Os resultados são apresentados na Tabela 5.1

Tabela 5.1: Estimativas marginais a posteriori (MCMC) dos parâmetros da Distribuição BBXII com tamanho amostral 300

	Valor Real	Média	Desvio Padrão	2,5%	50,0%	97,5%	Taxa de Aceitação(%)
a	16,000	568,547	80,494	449,261	562,269	707,222	73,859
b	8,000	287,621	38,290	218,282	291,223	339,991	74,549
s	1,000	0,730	0,017	0,697	0,729	0,767	12,353
k	5,000	1,461	0,050	1,3591	1,453	1,568	4,906
c	6,000	1,195	0,095	1,015	1,193	1,375	4,226

A Figura 5.1 apresenta o traço das estimativas dos parâmetros via MCMC, a aproximação da densidade marginal, além da autocorrelação. Verifica-se que as trajetórias dos gráficos relacionados ao traço não apresentaram a estacionariedade esperada.

Nota-se que as autocorrelações são muito elevadas, não obtendo-se atenuação ao longo do tempo, e as densidades marginais estimadas para alguns parâmetros não são unimodais, principalmente aos parâmetros relacionados à distribuição Beta.

A Figura 5.2 apresenta as curvas verdadeiras e estimadas segundo as estimativas marginais obtidas pelo MCMC. Apesar de todos os problemas no geral, observa-se uma boa aderência entre essa curva estimada e a curva de densidade verdadeira.

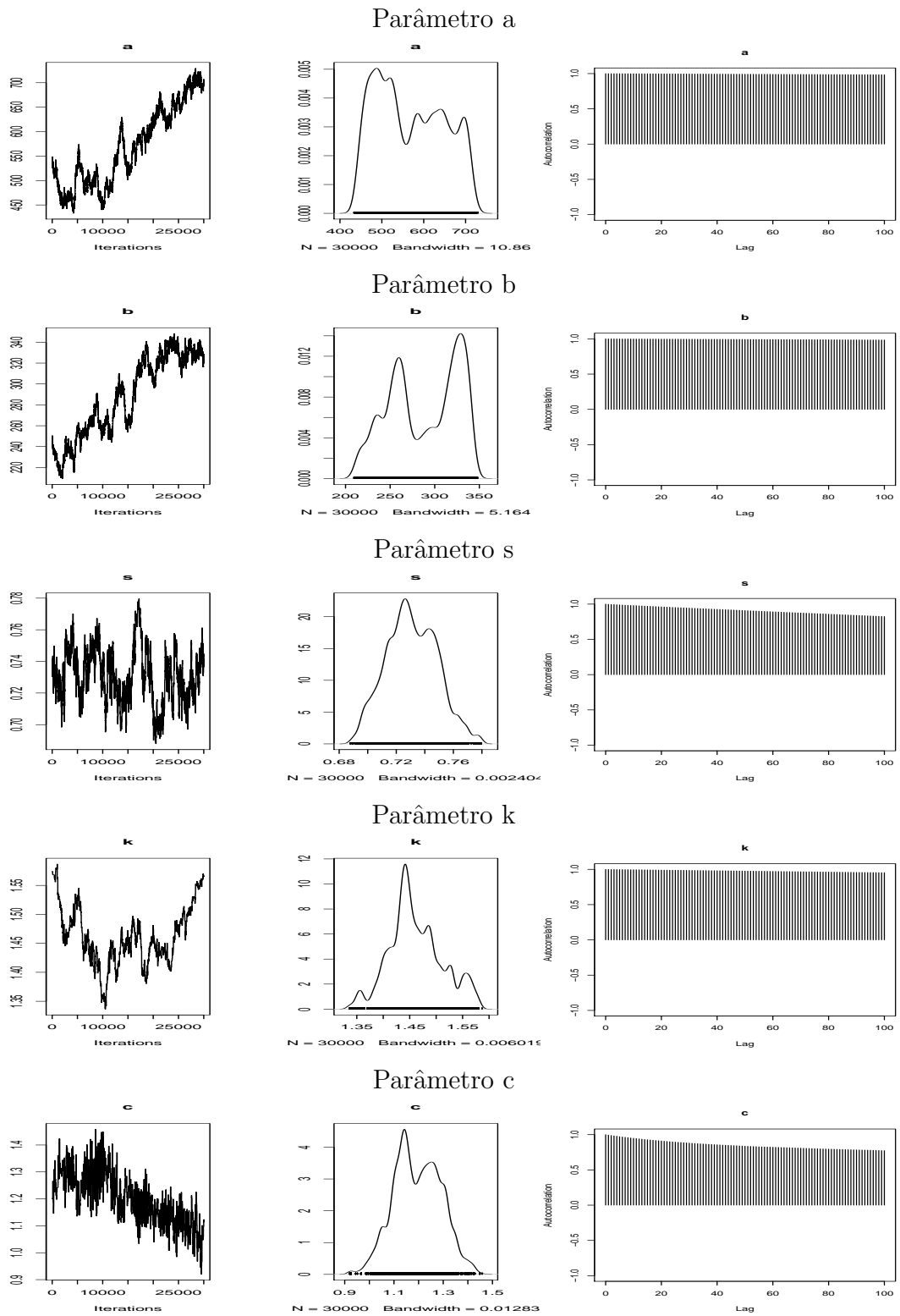


Figura 5.1: Traço, densidade e autocorrelação em amostras MCMC para os parâmetros a , b , k e c da distribuição BBXII.

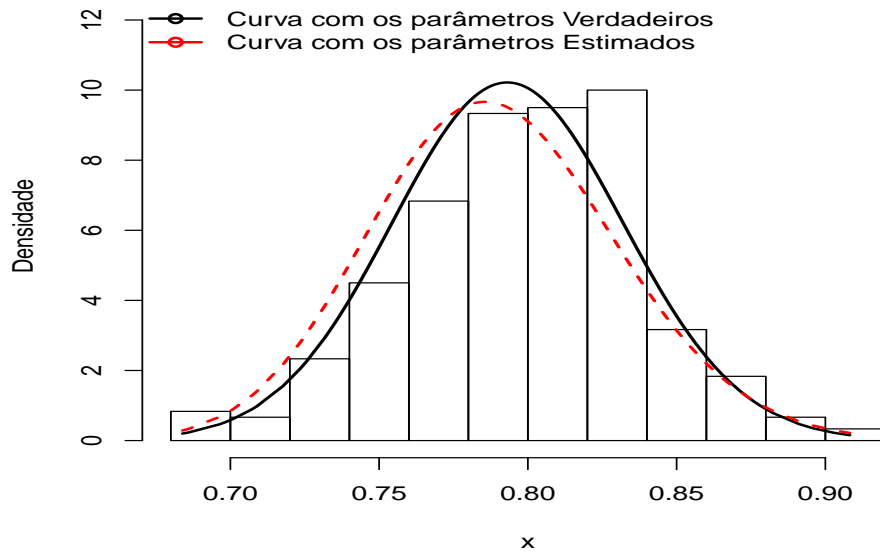


Figura 5.2: Curva associada a densidade verdadeira e a curva estimada.

Na próxima seção apresenta-se outros desenvolvimentos visando um tratamento Bayesiano mais satisfatório dos parâmetros, em especial dos parâmetros a e b associados à distribuição Beta. Outras abordagens que não tiveram sucesso são apresentadas no Apêndice. Em uma destas abordagens verificou-se uma dependência entre as estimativas dos parâmetros a e b . Numa tentativa de eliminar esse problema, propõe-se o uso de outra parametrização para a distribuição Beta, que é discutida na Seção 5.2.

5.2 Nova parametrização da Distribuição Beta

Nesta seção faremos uma mudança na parametrização da distribuição Beta associada a distribuição BBXII, conforme proposta de Ferrari e Cribari-Neto (2004).

Para uma variável aleatória $X \sim Beta(a, b)$, a fdp correspondente é dada por

$$f_{Beta}(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1}, \quad (5.22)$$

com $a > 0$, $b > 0$ e $B(a, b)$ descrevendo a função Beta. Sabe-se que a média e variância de X são dadas, respectivamente, por

$$E(X) = \frac{a}{a+b}, \quad (5.23)$$

$$Var(X) = \frac{ab}{(a+b)^2(a+b+1)}. \quad (5.24)$$

Ferrari e Cribari-Neto (2004) propõem outra parametrização da distribuição Beta, que consiste em tomar $\mu = a/(a+b)$ e $\gamma = a+b$. Dessa forma $a = \mu\gamma$ e $b = \gamma(1-\mu)$. Então,

$$E(X) = \frac{a}{a+b} = \mu \quad (5.25)$$

e

$$V(X) = \frac{ab}{(a+b)^2(a+b+1)} = \frac{\gamma(1-\mu)}{1+\gamma}. \quad (5.26)$$

Sendo assim, μ corresponde à média da variável resposta e γ é interpretado como um parâmetro de precisão, de modo que, dado μ fixo, quanto maior o valor de γ , menor será a variância de X .

Diante disso, espera-se que essa parametrização elimine a dependência nas estimativas dos parâmetros a e b que foram observadas em alguns estudos. Com essa nova parametrização, a densidade da variável aleatória Beta é dada por:

$$f_{Beta-G}(x) = \frac{1}{B(\mu\gamma, \gamma(1-\mu))} x^{\mu\gamma-1} (1-x)^{\gamma(1-\mu)-1}. \quad (5.27)$$

Nessa nova parametrização, a fdp da classe de distribuições Beta generalizada, descrita pela expressão (1.2) (Eugenne et. al., 2002) é expressa por:

$$f_{Beta-G}(x) = \frac{g(x)}{B(\mu\gamma, \gamma(1-\mu))} G(x)^{\mu\gamma-1} (1-G(x))^{\gamma(1-\mu)-1}, \quad (5.28)$$

em que $G(x)$ e $g(x)$ são descritos pelas expressões (1.3) e (1.4), respectivamente. Logo, a distribuição BBXII nessa nova parametrização, é dada por

$$f(x) = \frac{ckx^{c-1}}{s^c B(\mu\gamma, \gamma(1-\mu))} [1 + (x/s)^c]^{-(k\gamma(1-\mu)+1)} \{1 - [1 + (x/s)^c]^{-k}\}^{\mu\gamma-1}. \quad (5.29)$$

Na próxima Seção apresenta-se a distribuição a posteriori da BBXII sob a nova reparametrização e uma proposta de função geradora de candidatos.

5.2.1 Distribuição a posteriori da BBXII com a Nova Parametrização da Beta e Função Geradora de Candidatos

A distribuição a *posteriori* dos parâmetros associados ao modelo BBXII, sob a nova parametrização é dada por:

$$\pi(\mu, \gamma, s, k, c|X) \propto \left[\prod_{i=1}^n f(x_i|\mu, \gamma, s, k, c) \right] \pi(\mu, \gamma, s, k, c), \quad (5.30)$$

onde $f(x_i)$ é dada pela expressão (5.29). De forma análoga à Seção 4.1, por simplicidade, assumo que os parâmetros desse modelo sejam independentes *a priori*, de forma que

$$\pi(\mu, \gamma, s, k, c) = \pi(\mu)\pi(\gamma)\pi(s)\pi(k)\pi(c). \quad (5.31)$$

As distribuições condicionais completas *a posteriori* de cada um dos parâmetros nessa parametrização são dadas por:

$$\pi(\mu|\gamma, s, k, c, x) \propto \left[\prod_{i=1}^n f(x_i|\mu, \gamma, s, k, c) \right] \pi(\mu), \quad (5.32)$$

$$\pi(\gamma|\mu, s, k, c, x) \propto \left[\prod_{i=1}^n f(x_i|\mu, \gamma, s, k, c) \right] \pi(\gamma), \quad (5.33)$$

$$\pi(k|\mu, \gamma, s, c, x) \propto \left[\prod_{i=1}^n f(x_i|\mu, \gamma, k, c, x) \right] \pi(k), \quad (5.34)$$

$$\pi(c|\mu, \gamma, s, k, x) \propto \left[\prod_{i=1}^n f(x_i|\mu, \gamma, k, c, x) \right] \pi(c), \quad (5.35)$$

em que, a princípio, as *prioris* foram definidas da seguinte maneira, exceto para o parâmetro s^1 ,

$$\mu \sim \text{Beta}(1, 1), \quad (5.36)$$

$$\gamma \sim \text{Log-normal}(0, \sigma_\gamma), \quad (5.37)$$

$$\log(k) \sim N(0, \sigma_k) \quad (5.38)$$

$$e \ c \propto \frac{1}{c}. \quad (5.39)$$

Definiu-se a função geradora de candidatos dos parâmetros do modelo BBXII da seguinte forma

$$P(\mu) \sim \text{Beta}(\mu^{(t-1)}\gamma^{(t-1)}, \gamma^{(t-1)}(1 - \mu^{(t-1)})), \quad (5.40)$$

$$P(\gamma) \sim \text{Log-normal}(\log(\gamma^{(t-1)}), \sigma_{\gamma_1}), \quad (5.41)$$

$$P(\log(k)) \sim \text{Normal}(\log(k^{(t-1)}), \sigma_{k_1}), \quad (5.42)$$

$$P(\log(c)) \sim \text{Normal}(\log(c^{(t-1)}), \sigma_{c_1}), \quad (5.43)$$

onde as distribuições dadas em (5.38), (5.42) e (5.43) tratou-se do log do parâmetro, daí o modelo gaussiano.

¹Para evitar problemas numéricos, assume-se nesse caso $s = 1$

5.2.1.1 Resultados

Implementando-se os desenvolvimentos das seções anteriores num algoritmo MCMC e com os parâmetros k e c da distribuição BBXII na transformação “log” foi realizado o estudo de simulação com $\mu = 0.66$, obtida por (5.25), e $\gamma = 24$, pois $\gamma = a + b$, onde $a = 16$, $b = 8$, $k = 5$ e $c = 6$ são iguais às definidas nas simulações anteriores.

Vale lembrar que, para a realização da simulação, definiu-se os valores dos hiperparâmetros da seguinte maneira: $\sigma_\gamma = 0.08$, $\sigma_{\gamma_1} = 2$, $\sigma_k = 3$, $\sigma_{k_1} = \sigma_{c_1} = 0.055$ e $\sigma_c = 3$. Os resultados das estimativas dos parâmetros são apresentados na Tabela 5.2. Nota-se que já ocorre uma “aproximação” entre as estimativas e os valores verdadeiros em todos os parâmetros. Embora, a taxa de aceitação de alguns parâmetros ainda não estejam em níveis aceitáveis.

Tabela 5.2: Estimativas marginais à posteriori da Distribuição BBXII com utilizando a abordagem da Subseção 5.2.1

	Valor Real	Média	Desvio Padrão	2,5%	50,0%	97,5%	Taxa de Aceitação(%)
μ	0,666	0,677	0,005	0,666	0,6775	0,688	6,990
γ	24,000	21,178	1,246	18,774	21,145	23,712	61,031
k	5,000	5,730	0,314	5,140	5,718	6,391	34,436
c	6,000	6,405	0,252	5,910	6,399	6,902	26,648

A Figura 5.3 apresenta os gráficos do traço da estimativa marginal a *posteriori* do MCMC, a aproximação da densidade marginal, além da autocorrelação para todos os parâmetros da distribuição BBXII. Verifica-se que a densidade apresenta a forma de uma distribuição unimodal e as trajetórias dos gráficos relacionadas ao traço apresentaram a estacionariedade esperada.

Pela Figura 5.4 (a) observa-se que a curva estimada segundo as estimativas marginais obtidas pelo MCMC apresenta boa aderência com a curva de densidade verdadeira. Pela Figura 5.4 (b), tem-se que o verdadeiro valor de cada parâmetro (representado pela linha horizontal) se encontra dentro do intervalo interquartilico, em todos os casos. Isso revela que essa abordagem se mostra, a princípio, adequada na análise bayesiana.

No entanto, a *priori* e a função geradora de candidatos do parâmetro γ revelou-se instável computacionalmente. O algoritmo convergiu somente para valores de σ_γ (vide expressão (5.37)) tão pequenos quanto 0.08. Considerou-se, então, a proposta atual ainda insatisfatória, de modo que, outras propostas, mais flexíveis, precisam ser

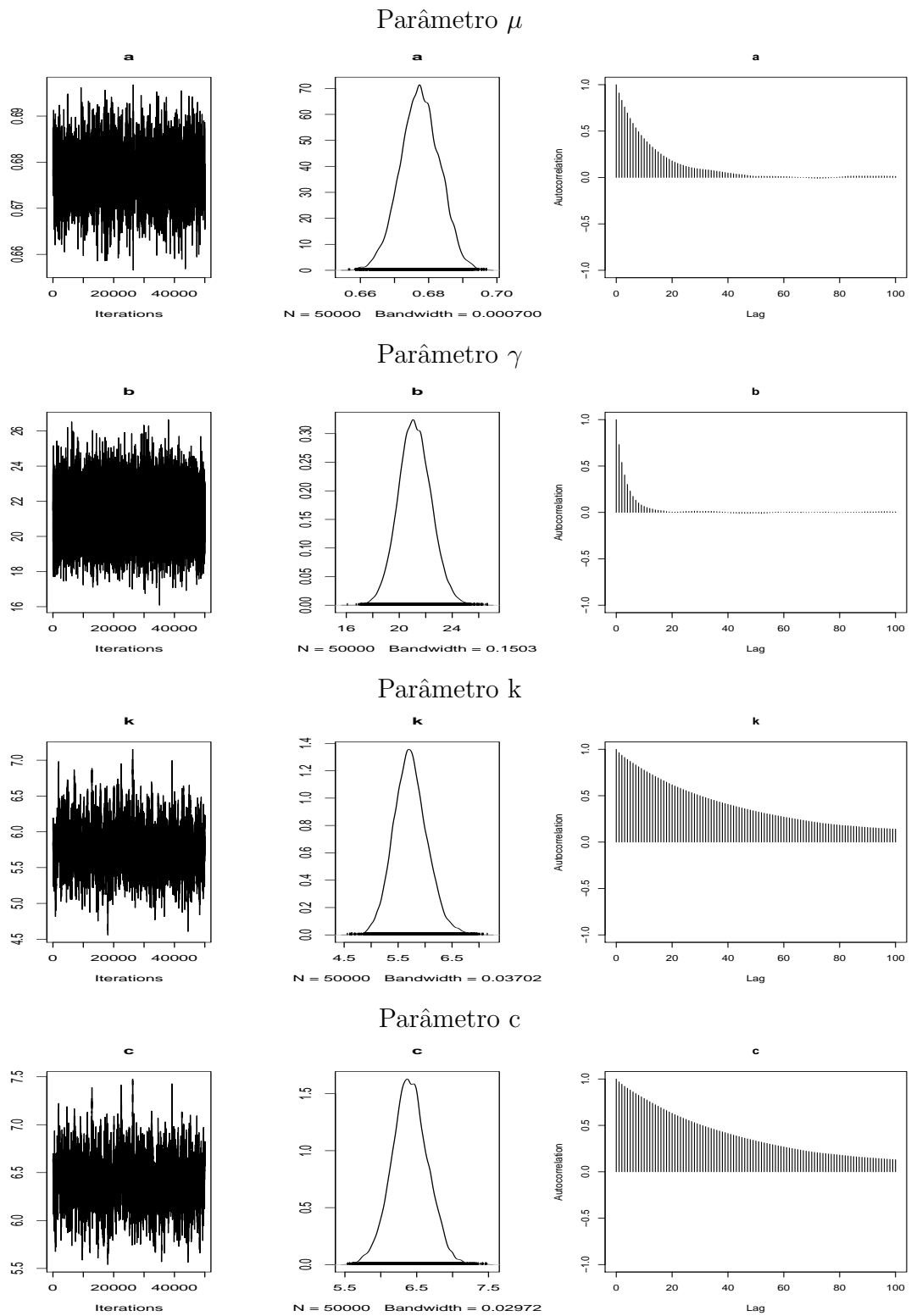
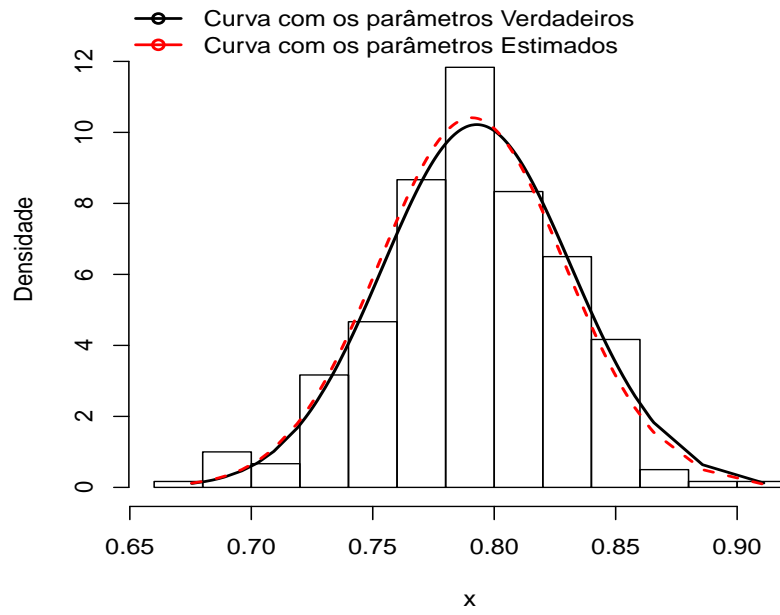


Figura 5.3: Traço, densidade e autocorrelação das amostras MCMC para os parâmetros a, b, k e c da distribuição BBXII.

(a)



(b)

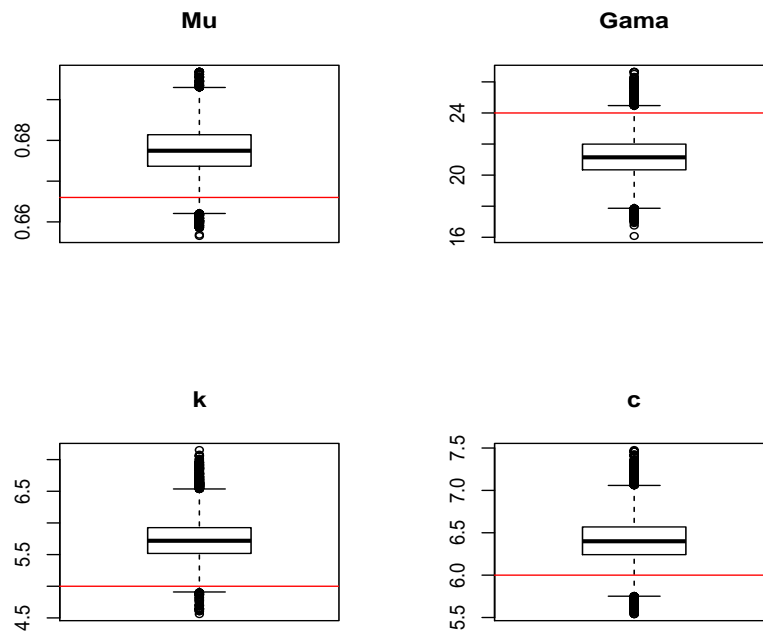


Figura 5.4: (a) Histograma dos dados e a Curva estimada (b) Box-plot e o valor do verdadeiro parâmetro (linha horizontal).

vislumbradas.

Dessa forma, observa-se a necessidade de novas propostas de *prioris* e de geradoras de candidatos que sejam mais flexíveis. Na próxima seção apresenta-se uma abordagem em que considera-se *prioris* Beta e Qui-quadrado.

5.2.2 Prioris Beta e Qui-Quadrado

Seguindo os procedimentos descritos nas equações (5.32) a (5.35), na tentativa de descrever *prioris* e funções geradoras de candidatos mais flexíveis, nessa seção as *prioris* dos parâmetros da distribuição BBXII foram definidas da seguinte maneira:

$$\mu \sim \text{Beta}(a_1, b_1), \quad (5.44)$$

$$\gamma \sim \text{Qui-quadrado}(a_2), \quad (5.45)$$

$$s \sim \text{Qui-quadrado}(a_3), \quad (5.46)$$

$$k \sim \text{Qui-quadrado}(a_4), \quad (5.47)$$

$$c \sim \text{Qui-quadrado}(a_5), \quad (5.48)$$

e as funções geradoras de candidatos dos parâmetros da seguinte forma

$$\mu \sim \text{Beta}(\mu^{(t-1)}\gamma^{(t-1)}, \gamma^{(t-1)}(1 - \mu^{(t-1)})), \quad (5.49)$$

$$\gamma \sim \text{Qui-quadrado}(\gamma^{(t-1)}), \quad (5.50)$$

$$k \sim \text{Qui-quadrado}(k^{(t-1)}), \quad (5.51)$$

$$c \sim \text{Qui-quadrado}(c^{c-1}). \quad (5.52)$$

5.2.2.1 Resultados

Para esta nova proposta considerou-se um algoritmo MCMC, com *burn-in* de 10 mil e 100 mil simulações e com amostra sistemáticas (*gap*) de tamanho 50. Os valores iniciais foram definidos como $\mu_0 = 0.6$, $\gamma_0 = 10$, $k_0 = 8$ e $c_0 = 8$. Os dados foram gerados utilizando uma amostra de tamanho $n = 200$ de uma distribuição BBXII com parâmetros $\mu = 0.4$, $\gamma = 18$, $k = 5$ e $c = 6$, com $s = 1$ fixo (vide Seção 1.1.1).

A Tabela 5.3 apresenta os resultados das estimativas dos parâmetros. Observa-se que, para todos os parâmetros o intervalo HPD contempla o verdadeiro valor, exceto

para o parâmetro μ , embora o valor estimado deste parâmetro esteja próximo do valor verdadeiro. No entanto, o intervalo HPD para μ aparentemente exibe amplitude estreita demais, o que nos motiva a buscar outras possibilidades de efetuar a estimação Bayesiana da distribuição BBXII. Tais procedimentos estão descritos na Seção 5.3. Em relação a taxa de aceitação, não foi possível obter taxas em níveis entre 25% e 50% para todos os parâmetros.

Tabela 5.3: Estimativas marginais a posteriori da Distribuição BBXII utilizando a abordagem da Seção 5.2.2.

Parâmetro	Valor Real	Mediana	Média	HPD (2, 5%)	HPD(97, 5%)	Tx. de Aceitação(%)
μ	0,400	0,344	0,345	0,328	0,364	8,487
γ	18,000	15,070	16,050	7,010	28,163	30,470
k	5,000	5,209	5,784	2,273	10,712	7,385
c	6,000	6,782	6,859	4,797	9,102	3,471

Na Figura 5.5 são apresentados os gráficos da trajetória da cadeia, da densidade empírica estimada, da autocorrelação e dos boxplots para cada parâmetro. Nota-se que, a densidade estimada possui unimodalidade e a trajetória possui aparente estacionariedade. A autocorrelação possui rápido decaimento para o parâmetro μ , enquanto para os outros parâmetros possui decaimento mais lento. Exceto pelo parâmetro μ , os boxplots revelam que o verdadeiro valor de cada parâmetro está inserido dentro do intervalo interquartilício.

5.2.2.2 Estudo Comparativo entre as abordagens frequentista e Bayesiana

Foi realizado um estudo comparativo entre as estimativas obtidas via máxima verossimilhança e as obtidas pelo método Bayesiano descrito na Seção 5.2.2. Para tanto, obteve-se o EMV, para o caso clássico, e a média e mediana a posteriori para o caso Bayesiano. Adicionalmente, obteve-se os seus respectivos vícios e a raiz quadrada do EQM. Considerou-se os tamanhos amostrais $n = 50, 100, 200, 300, 500$ e 1000 , de modo a possibilitar a análise de aspectos assintóticos dos estimadores.

Para obter comparabilidade entre as estimativas, tomou-se o cuidado em trabalhar com as mesmas amostras, em ambos métodos. Para isso, considerou-se somente as amostras em que o método frequentista foi convergente. Nesse estudo, as amostras foram geradas conforme a Seção 1.1.1. Os valores iniciais nesse caso foram: $\mu_0 = 0.6$, $\gamma_0 = 10$, $k_0 = 8$ e $c_0 = 8$.

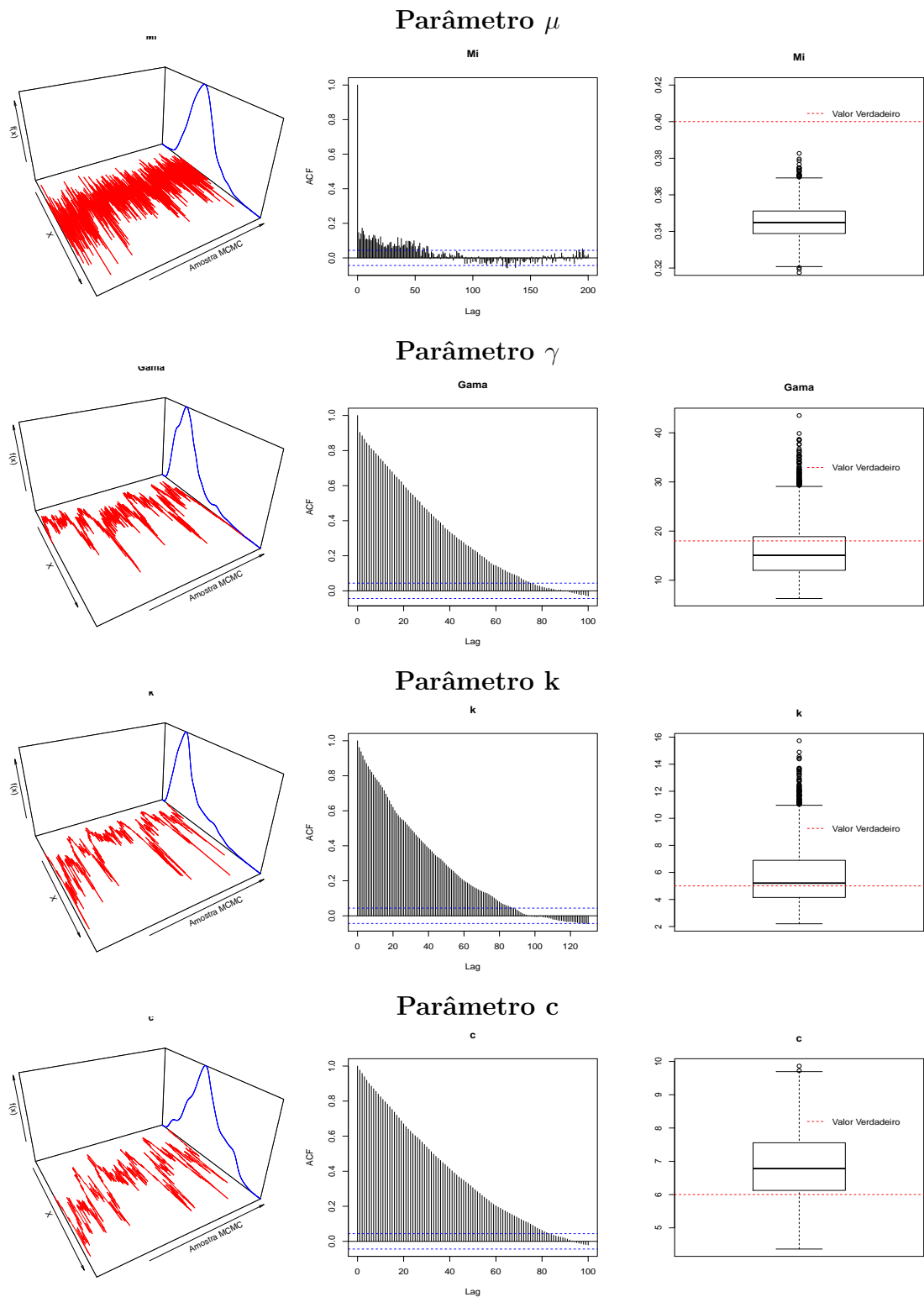


Figura 5.5: Traço, densidade, autocorrelação e boxplots associados as amostras MCMC para os parâmetros μ , γ , k e c da distribuição BBXII sob a proposta da Seção 5.2.2.

A Tabela 5.4 apresenta os resultados obtidos para 50 amostras da distribuição BBXII sob diferentes tamanhos amostrais. Verifica-se que, em linhas gerais, a medida que o tamanho amostral aumenta, o vício e a raiz do EQM diminuem, em todos os casos. Tais medidas (vício e EQM) apresentam valores menores quando obtidos pelo método Bayesiano, na maioria dos casos, exceto para o parâmetro c .

Uma outra alternativa de abordagem de estimação bayesiana dos parâmetros da distribuição BBXII será discutida na próxima seção. A abordagem consiste em proceder a uma transformação na distribuição a posteriori. Nesse caso, considera-se as seguintes transformações nos parâmetros: $\eta_1 = \log(\mu/(1 + \mu))$, $\eta_2 = \log(\gamma)$, $\eta_3 = \log(k)$ e $\eta_4 = \log(c)$.

Dessa forma, pode-se descrever *prioris* e funções geradoras de candidatos gaussianas, para cada parâmetro η_i , separadamente ou para o vetor $\eta = (\eta_1, \eta_2, \eta_3, \eta_4)^T$. Tal procedimento visa a obtenção de metodologia mais adequada na estimativa de μ .

Tabela 5.4: EMV e estimativas Bayesianas obtidos para a Distribuição BBXII com *prioris* Beta e Qui-Quadrado.

n	μ (0,4)			γ (18)			k (5)			c(6)		
	Bayesiano		EMV	Bayesiano		EMV	Bayesiano		EMV	Bayesiano		EMV
	Média	Mediana	Média	Mediana	Média	Média	Mediana	Média	Média	Mediana	Média	Mediana
50	0,5005	0,3399	0,3406	17,8631	11,8231	19,0598	6,1269	6,7201	8,4026	7,2199	7,2344	7,2344
100	0,5420	0,3368	0,3374	17,9448	13,2651	10,7397	6,3521	6,9211	6,903	7,3298	7,3434	7,3434
200	0,5281	0,3357	0,3362	17,9759	13,8482	9,9884	6,4025	6,945	6,8878	7,3389	7,3766	7,3766
300	0,4934	0,3351	0,3355	17,9862	14,6534	9,4776	6,3649	6,7976	7,0313	7,3183	7,3223	7,3223
500	0,4936	0,3353	0,3356	17,9933	15,1856	8,8991	6,0244	6,4231	6,8267	7,1621	7,1977	7,1977
1000	0,4036	0,3347	0,3349	17,9975	15,939	5,1532	5,8786	6,1185	6,0356	7,0662	7,0823	7,0823
Vício												
50	0,1005	-0,0601	-0,0594	-0,1369	-6,1769	14,0598	1,1269	1,7201	2,4026	1,2199	1,2344	1,2344
100	0,1420	-0,0632	-0,0626	-0,0552	-4,7349	5,7397	1,3521	1,9211	0,9030	1,3298	1,3434	1,3434
200	0,1281	-0,0643	-0,0638	-0,0241	-4,1518	4,9884	1,4025	1,945	0,8878	1,3389	1,3766	1,3766
300	0,0934	-0,0649	-0,0645	-0,0138	-3,3466	4,4776	1,3649	1,7976	1,0313	1,3183	1,3223	1,3223
500	0,0936	-0,0647	-0,0644	-0,0067	-2,8144	3,8991	1,0244	1,4231	0,8267	1,1621	1,1977	1,1977
1000	0,0036	-0,0653	-0,0651	-0,0025	-2,061	0,1532	0,8786	1,1185	0,0356	1,0662	1,0823	1,0823
Raiz do EQM												
50	0,3195	0,0617	0,0617	6,2786	6,2734	39,8699	1,8600	1,8796	8,0582	8,0582	1,273	1,273
100	0,2698	0,0638	0,0638	4,8433	4,8459	6,2904	2,0583	2,0751	6,2865	1,3784	1,3784	1,3784
200	0,2363	0,0645	0,0645	4,4126	4,4278	5,2380	2,2145	2,2259	6,2182	1,4438	1,4438	1,4438
300	0,1894	0,0649	0,0649	3,7193	3,6946	4,6165	2,1389	2,1246	6,1507	1,415	1,4002	1,4002
500	0,1593	0,0647	0,0647	3,5251	3,5742	4,0059	1,8307	1,8626	6,0885	1,2885	1,2955	1,2955
1000	0,0237	0,0653	0,0653	3,5803	3,6647	0,7015	1,7647	1,7570	0,2531	1,2606	1,2474	1,2474

5.3 Aproximações Gaussianas para os parâmetros da distribuição BBXII

Considere a nova parametrização da distribuição BBXII, proposta na seção 5.2 (vide expressão (5.29)) e o parâmetro de escala $s=1$ fixo. Seja $\Theta = (\mu, \gamma, k, c) = (\theta_1, \theta_2, \theta_3, \theta_4)$ e admita que os θ_i 's são independentes *a priori*, com as seguintes distribuições:

$$\mu \sim \text{Beta}(a_1, b_1), \quad (5.53)$$

$$\gamma \sim \text{Gama}(a_2, b_2), \quad (5.54)$$

$$k \sim \text{Gama}(a_3, b_3), \quad (5.55)$$

$$c \sim \text{Gama}(a_4, b_4), \quad (5.56)$$

ou seja,

$$P(\mu) = \frac{\Gamma(a_1 + b_1)}{\Gamma(a_1)\Gamma(b_1)} \mu^{a_1-1} (1 - \mu)^{b_1-1}, \quad (5.57)$$

$$P(\gamma) = \frac{b_2^{a_2-1}}{\Gamma(a_2)} \gamma^{a_2-1} e^{-b_2\gamma}, \quad (5.58)$$

$$P(k) = \frac{b_3^{a_3-1}}{\Gamma(a_3)} k^{a_3-1} e^{-b_3k}, \quad (5.59)$$

$$P(c) = \frac{b_4^{a_4-1}}{\Gamma(a_4)} c^{a_4-1} e^{-b_4c}. \quad (5.60)$$

Sejam as transformações:

$$\eta_1 = \log\left(\frac{\mu}{1 - \mu}\right), \quad \text{isto é,} \quad \mu = \frac{e^{\eta_1}}{1 + e^{\eta_1}}, \quad (5.61)$$

$$\eta_2 = \log(\gamma), \quad \text{isto é,} \quad \gamma = e^{\eta_2}, \quad (5.62)$$

$$\eta_3 = \log(k), \quad \text{isto é,} \quad k = e^{\eta_3}, \quad (5.63)$$

$$\eta_4 = \log(c), \quad \text{isto é,} \quad c = e^{\eta_4}. \quad (5.64)$$

A distribuição a posteriori de $\eta = (\eta_1, \eta_2, \eta_3, \eta_4)$ é dada por

$$\pi_\eta(\eta_1, \eta_2, \eta_3, \eta_4 | X) \propto \pi_\Theta(\theta_1, \theta_2, \theta_3, \theta_4 | X) \times |J|, \quad (5.65)$$

em que $\theta_1 = \exp(\eta_1)/(1 + \exp(\eta_1))$, $\theta_2 = \exp(\eta_2)$, $\theta_3 = \exp(\eta_3)$ e $\theta_4 = \exp(\eta_4)$ e J é o Jacobiano da transformação, que é dado por:

$$J = \begin{bmatrix} \frac{\partial \theta_1}{\partial \eta_1} & \frac{\partial \theta_1}{\partial \eta_2} & \frac{\partial \theta_1}{\partial \eta_3} & \frac{\partial \theta_1}{\partial \eta_4} \\ \frac{\partial \theta_2}{\partial \eta_1} & \frac{\partial \theta_2}{\partial \eta_2} & \frac{\partial \theta_2}{\partial \eta_3} & \frac{\partial \theta_2}{\partial \eta_4} \\ \frac{\partial \theta_3}{\partial \eta_1} & \frac{\partial \theta_3}{\partial \eta_2} & \frac{\partial \theta_3}{\partial \eta_3} & \frac{\partial \theta_3}{\partial \eta_4} \\ \frac{\partial \theta_4}{\partial \eta_1} & \frac{\partial \theta_4}{\partial \eta_2} & \frac{\partial \theta_4}{\partial \eta_3} & \frac{\partial \theta_4}{\partial \eta_4} \end{bmatrix} = \begin{bmatrix} \frac{e^{\eta_1}}{(1+e^{\eta_1})^2} & 0 & 0 & 0 \\ 0 & e^{\eta_2} & 0 & 0 \\ 0 & 0 & e^{\eta_3} & 0 \\ 0 & 0 & 0 & e^{\eta_4} \end{bmatrix}$$

$$|J| = \frac{\exp(\eta_1 + \eta_2 + \eta_3 + \eta_4)}{(1 + e^{\eta_1})^2}. \quad (5.66)$$

Admitindo-se a expressão (1.3), a parametrização dos parâmetros a e b proposta na Seção 5.2 e as transformações (5.61) a (5.64), e por brevidade, denotando-se

$$G(x) = G(x|1, \eta_3, \eta_4) = 1 - \left[1 + \left(\frac{x}{1} \right)^{\exp(\eta_4)} \right]^{-\exp(\eta_3)}, \quad (5.67)$$

tem-se que a distribuição a posteriori com respeito a η é dada por:

$$\begin{aligned} \pi(\eta|X) &\propto e^{\eta_3 + \eta_4} [1 + x^{e^{\eta_4}}]^{-(e^{\eta_3} + 1)} x^{e^{\eta_4} - 1} \times \left[\text{Beta} \left(\frac{\exp(\eta_1 + \eta_2)}{(1 + e^{\eta_1})}, \frac{e^{\eta_2}}{1 + e^{\eta_1}} \right) \right]^{-1} \times \\ &\times [G(x)]^{\left(\frac{\exp(\eta_1 + \eta_2)}{1 + \exp(\eta_1)} \right) - 1} \times [1 - G(x)]^{\left(\frac{e^{\eta_2}}{1 + e^{\eta_1}} \right) - 1} \times \\ &\times \left[\frac{e^{\eta_1}}{(1 + e^{\eta_1})} \right]^{a_1 - 1} \left[\frac{1}{(1 + e^{\eta_1})} \right]^{b_1 - 1} \times [e^{\eta_2}]^{a_2 - 1} e^{-b_2 e^{\eta_2}} \times \\ &\times [e^{\eta_3}]^{a_3 - 1} e^{-b_3 e^{\eta_3}} \times [e^{\eta_4}]^{a_4 - 1} e^{-b_4 e^{\eta_4}} \times \frac{\exp(\eta_1 + \eta_2 + \eta_3 + \eta_4)}{(1 + e^{\eta_1})^2}. \end{aligned} \quad (5.68)$$

Dessa forma, a posteriori de η (para $n = 1$) é dada por:

$$\begin{aligned} \pi(\eta|X) &\propto [G(x)]^{\left(\frac{\exp(\eta_1 + \eta_2)}{(1 + e^{\eta_1})} \right) - 1} \times [1 - G(x)]^{\left(\frac{e^{\eta_2}}{(1 + e^{\eta_1})} \right) - 1} \times \\ &\times \left[\text{Beta} \left(\frac{\exp(\eta_1 + \eta_2)}{(1 + e^{\eta_1})}, \frac{e^{\eta_2}}{1 + e^{\eta_1}} \right) \right]^{-1} \times x^{e^{\eta_4} - 1} \times [e^{\eta_1}]^{a_1} \times [e^{\eta_2}]^{a_2} \times \\ &\times [e^{\eta_3}]^{a_3 + 1} \times [e^{\eta_4}]^{a_4 + 1} \times e^{-b_2 [e^{\eta_2}]} \times e^{-b_3 [e^{\eta_3}]} \times e^{-b_4 [e^{\eta_4}]} \times \\ &\times (1 + e^{\eta_1})^{-(a_1 + b_1)} \times [1 + (x)^{e^{\eta_4}}]^{-(e^{\eta_3} + 1)}. \end{aligned} \quad (5.69)$$

A log-posteriori de η é dada por:

$$\begin{aligned}
\log(\pi(\eta|X)) &\propto \left[\frac{e^{\eta_1+\eta_2}}{1+e^{\eta_1}} - 1 \right] \log(G(x)) + \left[\frac{e^{\eta_2}}{1+e^{\eta_1}} - 1 \right] \log(1-G(x)) + \\
&+ (e^{\eta_4} - 1) \log(x) + (e^{\eta_3} + 1) \log(1 + (x)^{e^{\eta_4}}) - \\
&- \log\left(\text{Beta} \left[\frac{e^{\eta_1+\eta_2}}{1+e^{\eta_1}}, \frac{e^{\eta_2}}{1+e^{\eta_1}} \right] \right) + a_1\eta_1 + a_2\eta_2 + (a_3 + 1)\eta_3 + \\
&+ (a_4 + 1)\eta_4 - b_2 [e^{\eta_2}] - b_3 [e^{\eta_3}] - b_4 [e^{\eta_4}] - \\
&- (a_1 + b_1) \log(1 + e^{\eta_1}). \tag{5.70}
\end{aligned}$$

Na próxima Seção apresenta-se a forma de obtenção das distribuições geradoras de candidatos para cada um dos parâmetros. Nesse caso, será obtido um vetor de candidatos.

5.3.1 Função Geradora de Candidatos Conjunta

Gray (2001) propõe o uso de uma distribuição geradora de candidatos baseada em uma aproximação Gaussiana da distribuição a posteriori com respeito a η . Assuma que

$$\eta \sim N_4(\hat{\eta}, I^{-1}(\hat{\eta})), \tag{5.71}$$

descreve a distribuição geradora de candidatos η , onde $\hat{\eta}$ é a moda a posteriori de $\pi(\eta|X)$ e $I^{-1}(\hat{\eta})$ representa o inverso da matriz de informação observada, que é dada por menos o inverso da matriz hessiana de $\log(\pi(\eta|X))$ avaliada em $\hat{\eta}$, isto é,

$$I(\hat{\eta}) = - \frac{d^2}{d\eta d\eta^t} \log[\pi(\eta|X)] \Bigg|_{\eta=\hat{\eta}}. \tag{5.72}$$

Considere a decomposição de *Choleski* de $I(\hat{\eta})$ de modo que

$$I(\hat{\eta}) = B^T B. \tag{5.73}$$

Portanto,

$$I^{-1}(\hat{\eta}) = B^{-1}(B^T)^{-1} = B^{-1}(B^{-1})^T. \tag{5.74}$$

De acordo com Gray (2001) os vetores de candidatos de η podem ser obtidas a partir da distribuição descrita em (5.71) e utilizando a seguinte estratégia: considere um vetor candidato η descrito por

$$\eta^{(s)} = B^{-1}Z^{(s)} + \hat{\eta}, \quad (5.75)$$

onde $Z^{(s)} = (Z_1, Z_2, Z_3, Z_4)$ e $Z_i \sim N(0, 1)$ independentes. A distribuição geradora de candidatos é normal multivariada, em que

$$g(\eta) \propto \exp\left(-\frac{1}{2}(\eta - \hat{\eta})^T (I^{-1}(\hat{\eta})) (\eta - \hat{\eta})\right) \quad (5.76)$$

$$\propto \exp\left(-\frac{1}{2}(\eta - \hat{\eta})^T I(\hat{\eta}) (\eta - \hat{\eta})\right). \quad (5.77)$$

No processo de Metropolis-Hastings precisa-se calcular:

$$\alpha \propto (\eta^{(old)}, \eta^{(s)}) = \min\{1, r\}, \quad (5.78)$$

onde

$$r = \frac{\pi(\eta^s|X)}{\pi(\eta^{old}|X)} \frac{g(\eta^{old})}{g(\eta^s)}, \quad (5.79)$$

em que

$$r = \frac{\pi(\eta^{(s)}|X)/\pi(\hat{\eta}|X)}{\pi(\eta^{(old)}|X)/\pi(\hat{\eta}|X)} \times \frac{\exp\{-\frac{1}{2}(\eta^{(old)} - \hat{\eta})^T I(\hat{\eta}) (\eta^{(old)} - \hat{\eta})\}}{\exp\{-\frac{1}{2}(\eta^{(s)} - \hat{\eta})^T I(\hat{\eta}) (\eta^{(s)} - \hat{\eta})\}}. \quad (5.80)$$

Como $\eta^{(s)} - \hat{\eta} = (B^{-1}Z^{(s)} + \hat{\eta}) - \hat{\eta} = B^{-1}Z^s$, no denominador da razão à direita de (5.80), tem-se:

$$\begin{aligned} \exp\left\{-\frac{1}{2}(\eta^{(s)} - \hat{\eta})^T I(\hat{\eta}) (\eta^{(s)} - \hat{\eta})\right\} &= \exp\left\{-\frac{1}{2}(Z^s)^T + (B^{-1})^T [(B^{-1})^T]^{-1} B B^{-1} Z^{(s)}\right\} \\ &= \exp\left\{-\frac{1}{2}(Z^{(s)})^T Z^{(s)}\right\} \\ &= \exp\left\{-\frac{1}{2}\sum_{i=1}^p (Z_i^{(s)})^2\right\}. \end{aligned}$$

Então, de (5.80) tem-se:

$$r = \frac{\pi(\eta^{(s)}|X) / \pi(\hat{\eta}|X)}{\pi(\eta^{(old)}|X) / \pi(\hat{\eta}|X)} \times \frac{\exp\left\{-\frac{1}{2} \sum_{i=1}^p (Z_i^{(old)})^2\right\}}{\exp\left\{-\frac{1}{2} \sum_{i=1}^p (Z_i^{(s)})^2\right\}}. \quad (5.81)$$

5.3.2 Resultados

Implementando-se os desenvolvimentos da Seção 5.3.1 num algoritmo MCMC e considerando os dados gerados utilizando uma amostra de tamanho $n=200$ de uma distribuição BBXII com parâmetros $\mu = 0.4$, $\gamma = 18$, $k = 5$ e $c = 6$, com $s = 1$ fixo, adotou-se um cadeia de tamanho 100.000, *burn-in* de 10.000 amostras e amostra sistemática (*gap*) com período $k=100$.

A Tabela 5.5 apresenta os resultados das estimativas dos parâmetros. Observa-se que, em todos os casos, o valor verdadeiro de cada parâmetro está inserido no intervalo HPD. Nota-se que a taxa de aceitação do vetor de candidatos está em níveis aceitáveis. Dessa forma, aparentemente, a proposta desta seção é superior a seção 5.2.2 (com *prioris* Chi-quadrado).

Tabela 5.5: Estimativas marginais à posteriori da Distribuição BBXII utilizando a abordagem da Seção 5.3.

Parâmetro	Valor Real	Mediana	Média	HPD (2, 5%)	HPD(97, 5%)	Tx. de Aceitação(%)
μ	0,400	0,561	0,557	0,369	0,775	32,700
γ	18,000	16,607	17,943	8,298	30,901	32,700
k	5,000	6,531	7,216	2,653	15,004	32,700
c	6,000	5,424	5,467	4,243	6,914	32,700

A Figura 5.6 apresenta os gráficos da densidade empírica estimada, do traço, da autocorrelação e os boxplots para cada parâmetro da distribuição BBXII. Verifica-se que, em todos os casos, a densidade possui forma unimodal, a trajetória é aparentemente estacionária e observa-se rápido decaimento da autocorrelação.

5.3.2.1 Estudo Comparativo entre as abordagens frequentista e Bayesiana

Foi realizado um estudo comparativo entre os métodos frequentista e Bayesiano. Para ter comparabilidade entre os métodos, tomou-se o cuidado em utilizar as mesmas

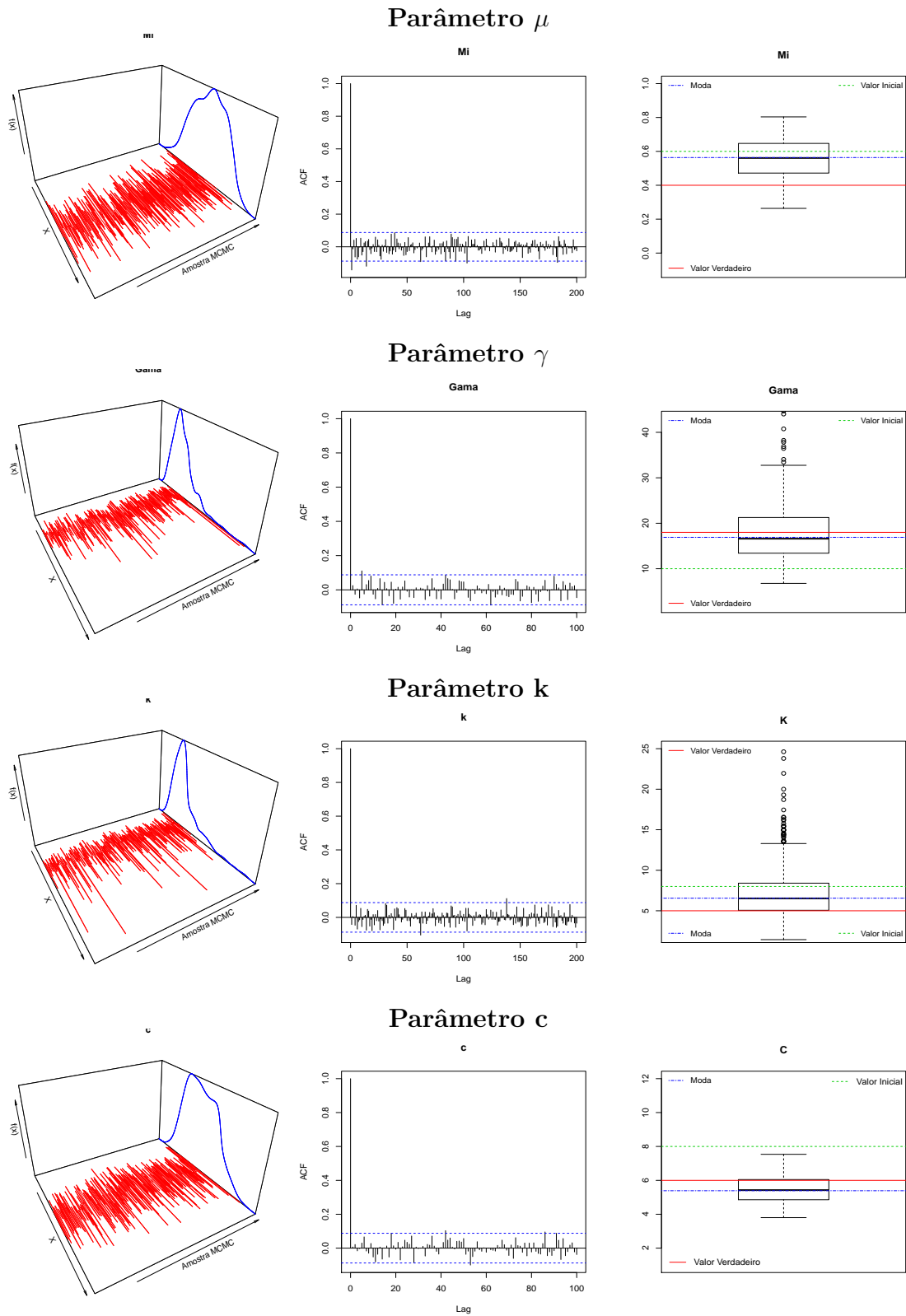


Figura 5.6: Traço, Densidade, Autocorrelação e Boxplot's das amostras MCMC para os parâmetros μ , γ , k e c da distribuição BBXII sob a proposta da Seção 5.3.

amostras, tanto para a obtenção das estimativas de máxima verossimilhança, quanto das Bayesianas.

Para isso, precisou-se primeiramente obter a moda a *posteriori* e a respectiva matriz hessiana sob a decomposição de *Choleski* (vide Seção 5.3.1). Logo depois, obteve-se os EMV's e as estimativas Bayesianas.

A Tabela 5.6 apresenta os resultados obtidos para 50 amostras da distribuição BBXII sob diferentes tamanhos amostrais. Nota-se que a raiz do EQM diminui com o aumento do tamanho amostral. Outro ponto que se pode observar é que o vício e o EQM diminuem mais rapidamente, a medida que n cresce para as estimativas via EMV, exceto para o parâmetro c . No entanto, para amostras pequenas ($n = 50$) o estimador Bayesiano apresenta, em geral, menor vício e EQM, conferindo uma superioridade importante do método Bayesiano neste caso, tão frequente na prática.

A lenta convergência observada para o verdadeiro valor pode estar relacionado com a escolha dos valores iniciais.

Tabela 5.6: EMVs e estimativas bayesianas obtidas para a Distribuição BBXII.

		μ (0.4)			γ (18)			k (5)			c(6)					
Média																
		Bayesiano			EMV			Bayesiano			EMV					
n	Média	Moda	Mediana	Média	Média	Moda	Mediana	Média	Média	Moda	Mediana	Média	Média	Moda	Mediana	Média
50	0,4746	0,4759	0,4759	0,4758	15,3490	14,0780	14,0830	14,1030	12,7430	12,2200	12,2150	12,2420	8,3200	7,4440	7,4420	7,4450
100	0,4836	0,4929	0,4930	0,4930	13,0960	13,8610	13,8570	13,8620	10,8970	11,5860	11,5870	11,5940	7,6960	7,2400	7,2390	7,2400
200	0,5010	0,6366	0,5441	0,5441	12,1160	14,6220	12,8980	12,9020	9,5880	9,7270	9,7280	9,7300	7,0100	6,5930	6,5940	6,5940
300	0,5191	0,5541	0,5542	0,5539	12,1740	12,8730	12,8830	12,8900	9,4480	9,4780	9,4780	9,4980	6,8010	6,4510	6,4490	6,4510
500	0,4905	0,5637	0,5637	0,5625	17,0200	17,0540	17,1600	17,2640	10,1920	9,7940	9,8010	10,1490	8,9350	6,4750	6,4770	6,4820
1000	0,4657	0,5553	0,5550	0,5548	17,5700	16,4270	16,3150	16,3930	9,0980	9,9630	9,9580	10,1260	8,2900	6,5730	6,5720	6,5750
Vício																
50	-0,1914	-0,1901	-0,1901	-0,1902	-8,6510	-9,9220	-9,9170	-9,8970	7,7430	7,2200	7,2150	7,2420	2,3200	1,4440	1,4420	1,4450
100	-0,1824	-0,1731	-0,1730	-0,1730	-10,9040	-10,1390	-10,1430	-10,1380	5,8970	6,5860	6,5870	6,5940	1,6960	1,2400	1,2390	1,2400
200	-0,1650	-0,0294	-0,1219	-0,1219	-11,8840	-9,3780	-11,1020	-11,0980	4,5880	4,7270	4,7280	4,7300	1,0100	0,5930	0,5940	0,5940
300	-0,1469	-0,1119	-0,1118	-0,1121	-11,8260	-11,1270	-11,1170	-11,1100	4,4480	4,4780	4,4780	4,4980	0,8010	0,4510	0,4490	0,4510
500	-0,1755	-0,1023	-0,1023	-0,1035	-6,9800	-6,9460	-6,8400	-6,7360	5,1920	4,7940	4,8010	5,1490	2,9350	0,4750	0,4770	0,4820
1000	-0,2003	-0,1107	-0,1110	-0,1112	-6,4300	-7,5730	-7,6850	-7,6070	4,0980	4,9630	4,9580	5,1260	2,2900	0,5730	0,5720	0,5750
Raiz do EQM																
50	0,2866	0,1535	0,1535	0,1535	7,4905	7,7894	7,7983	7,7983	9,7917	10,9943	10,9806	10,9806	4,5483	3,2152	3,2141	3,2141
100	0,2514	0,1709	0,1710	0,1710	5,9187	7,3744	7,3793	7,3793	6,5897	9,7137	9,7112	9,7112	3,2122	2,9290	2,9281	2,9281
200	0,2027	0,1995	0,1994	0,1994	6,3494	6,4647	6,4615	6,4615	4,7514	5,2535	5,2533	5,2533	1,6733	1,5693	1,5693	1,5693
300	0,2045	0,2047	0,2048	0,2048	6,0773	6,0329	6,0251	6,0251	4,6030	4,7367	4,7376	4,7376	1,5315	1,3985	1,3985	1,3985
500	0,1404	0,1881	0,1885	0,1885	6,0753	5,6476	5,6489	6,6489	3,6216	4,8359	4,8554	4,8554	1,2186	0,9430	0,9412	0,9412
1000	0,0996	0,1734	0,1731	0,1731	5,6922	6,8039	6,7999	6,7999	3,0880	5,1306	5,1351	5,1351	1,1352	0,9792	0,9804	0,9804

Capítulo 6

Implementação Bayesiana da distribuição BEW utilizando o R

Nesta seção, apresenta-se o desenvolvimento Bayesiano para a estimação dos parâmetros da distribuição BEW.

6.1 Aproximações Gaussianas para os parâmetros da distribuição BEW

Considere a distribuição BEW dada em (1.17) e a nova parametrização da distribuição Beta-G dada em (5.28). Logo a distribuição da BEW sob a nova parametrização é dada por

$$f_{BEW}(x) = \frac{\alpha c \lambda^c x^{c-1} e^{-(\lambda x)^c}}{B(\mu\gamma, \gamma(1-\mu))} (1 - e^{-(\lambda x)^c})^{\alpha\mu\gamma-1} \left\{ 1 - (1 - e^{-(\lambda x)^c})^\alpha \right\}^{\gamma(1-\mu)-1}. \quad (6.1)$$

Tome $\Theta = (\mu, \gamma, \alpha, \lambda, c) = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ e admita que os θ_i 's são independentes

a priori com as seguintes distribuições:

$$\mu \sim \text{Beta}(a_1, b_1), \quad (6.2)$$

$$\gamma \sim \text{Gama}(a_2, b_2), \quad (6.3)$$

$$\alpha \sim \text{Gama}(a_3, b_3), \quad (6.4)$$

$$\lambda \sim \text{Gama}(a_4, b_4), \quad (6.5)$$

$$c \sim \text{Gama}(a_5, b_5), \quad (6.6)$$

ou seja,

$$P(\mu) = \frac{\Gamma(a_1 + b_1)}{\Gamma(a_1)\Gamma(b_1)} \mu^{a_1-1} (1 - \mu)^{b_1-1}, \quad (6.7)$$

$$P(\gamma) = \frac{b_2^{a_2-1}}{\Gamma(a_2)} \gamma^{a_2-1} e^{-b_2 \gamma}, \quad (6.8)$$

$$P(\alpha) = \frac{b_3^{a_3-1}}{\Gamma(a_3)} \alpha^{a_3-1} e^{-b_3 \alpha}, \quad (6.9)$$

$$P(\lambda) = \frac{b_4^{a_4-1}}{\Gamma(a_4)} \lambda^{a_4-1} e^{-b_4 \lambda}, \quad (6.10)$$

$$P(c) = \frac{b_5^{a_5-1}}{\Gamma(a_5)} c^{a_5-1} e^{-b_5 c}. \quad (6.11)$$

Considere as transformações

$$\eta_1 = \log\left(\frac{\mu}{1 - \mu}\right) \quad \text{isto é,} \quad \mu = \frac{e^{\eta_1}}{1 + e^{\eta_1}}, \quad (6.12)$$

$$\eta_2 = \log(\gamma) \quad \text{isto é,} \quad \gamma = e^{\eta_2}, \quad (6.13)$$

$$\eta_3 = \log(\alpha) \quad \text{isto é,} \quad \alpha = e^{\eta_3}, \quad (6.14)$$

$$\eta_4 = \log(\lambda) \quad \text{isto é,} \quad \lambda = e^{\eta_4}, \quad (6.15)$$

$$\eta_5 = \log(c) \quad \text{isto é,} \quad c = e^{\eta_5}. \quad (6.16)$$

Seguindo os desenvolvimentos análogos ao da Seção 5.3 tem-se que a distribuição a posteriori de $\eta = (\eta_1, \eta_2, \eta_3, \eta_4, \eta_5)$ é dada por

$$\pi_\eta(\eta_1, \eta_2, \eta_3, \eta_4, \eta_5 | X) \propto \pi_\Theta(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5 | X) \times |J|, \quad (6.17)$$

em que $\theta_1 = \exp(\eta_1)/(1 + \exp(\eta_1))$, $\theta_2 = \exp(\eta_2)$, $\theta_3 = \exp(\eta_3)$, $\theta_4 = \exp(\eta_4)$ e

$\theta_5 = \exp(\eta_5)$ e J é o jacobiano da transformação que é dado por:

$$J = \begin{bmatrix} \frac{\partial \theta_1}{\partial \eta_1} & \frac{\partial \theta_1}{\partial \eta_2} & \frac{\partial \theta_1}{\partial \eta_3} & \frac{\partial \theta_1}{\partial \eta_4} & \frac{\partial \theta_1}{\partial \eta_5} \\ \frac{\partial \theta_2}{\partial \eta_1} & \frac{\partial \theta_2}{\partial \eta_2} & \frac{\partial \theta_2}{\partial \eta_3} & \frac{\partial \theta_2}{\partial \eta_4} & \frac{\partial \theta_2}{\partial \eta_5} \\ \frac{\partial \theta_3}{\partial \eta_1} & \frac{\partial \theta_3}{\partial \eta_2} & \frac{\partial \theta_3}{\partial \eta_3} & \frac{\partial \theta_3}{\partial \eta_4} & \frac{\partial \theta_3}{\partial \eta_5} \\ \frac{\partial \theta_4}{\partial \eta_1} & \frac{\partial \theta_4}{\partial \eta_2} & \frac{\partial \theta_4}{\partial \eta_3} & \frac{\partial \theta_4}{\partial \eta_4} & \frac{\partial \theta_4}{\partial \eta_5} \\ \frac{\partial \theta_5}{\partial \eta_1} & \frac{\partial \theta_5}{\partial \eta_2} & \frac{\partial \theta_5}{\partial \eta_3} & \frac{\partial \theta_5}{\partial \eta_4} & \frac{\partial \theta_5}{\partial \eta_5} \end{bmatrix} = \begin{bmatrix} \frac{e^{\eta_1}}{(1+e^{\eta_1})^2} & 0 & 0 & 0 & 0 \\ 0 & e^{\eta_2} & 0 & 0 & 0 \\ 0 & 0 & e^{\eta_3} & 0 & 0 \\ 0 & 0 & 0 & e^{\eta_4} & 0 \\ 0 & 0 & 0 & 0 & e^{\eta_5} \end{bmatrix}$$

$$|J| = \frac{\exp(\eta_1 + \eta_2 + \eta_3 + \eta_4 + \eta_5)}{(1 + e^{\eta_1})^2}. \quad (6.18)$$

Portanto, a distribuição a posteriori de η é dada por:

$$\begin{aligned} \pi(\eta|X) &= \frac{e^{\eta_3} e^{\eta_5} e^{\eta_4} e^{\eta_5}}{B\left(\frac{e^{\eta_1}}{1+e^{\eta_1}} e^{\eta_2}, e^{\eta_2} \left(1 - \frac{e^{\eta_1}}{1+e^{\eta_1}}\right)\right)} x^{e^{\eta_5}-1} e^{-(e^{\eta_4} x)^{e^{\eta_5}}} \left(1 - e^{-(e^{\eta_4} x)^{e^{\eta_5}}}\right)^{e^{\eta_3} \frac{e^{\eta_1}}{1+e^{\eta_1}} e^{\eta_2}-1} \times \\ &\times \left\{1 - \left(1 - e^{-(e^{\eta_4} x)^{e^{\eta_5}}}\right)^{e^{\eta_3}}\right\}^{e^{\eta_2} \left(1 - \frac{e^{\eta_1}}{1+e^{\eta_1}}\right)-1} \times \left[\frac{e^{\eta_1}}{(1+e^{\eta_1})}\right]^{a_1-1} \left[\frac{1}{(1+e^{\eta_1})}\right]^{b_1-1} \times \\ &\times [e^{\eta_2}]^{a_2-1} e^{-b_2 e^{\eta_2}} \times [e^{\eta_3}]^{a_3-1} e^{-b_3 e^{\eta_3}} \times [e^{\eta_4}]^{a_4-1} e^{-b_4 e^{\eta_4}} \times [e^{\eta_5}]^{a_5-1} e^{-b_5 e^{\eta_5}} \times \\ &\times \frac{\exp(\eta_1 + \eta_2 + \eta_3 + \eta_4 + \eta_5)}{(1 + e^{\eta_1})^2}. \end{aligned} \quad (6.19)$$

A log-posteriori de η é dada por

$$\begin{aligned} \log(\pi(\eta|X)) &= \eta_3 + \eta_5 + \eta_4 e^{\eta_5} - \log\left(B\left(\frac{e^{\eta_1}}{1+e^{\eta_1}} e^{\eta_2}, e^{\eta_2} \left(1 - \frac{e^{\eta_1}}{1+e^{\eta_1}}\right)\right)\right) + \\ &+ (\eta_5 - 1) \log(x) - (e^{\eta_4} x)^{e^{\eta_5}} \left(e^{\eta_3} \frac{e^{\eta_1}}{1+e^{\eta_1}} e^{\eta_2} - 1\right) \log\left[1 - e^{-(e^{\eta_4} x)^{e^{\eta_5}}}\right] + \\ &+ \left(e^{\eta_2} \left(1 - \frac{e^{\eta_1}}{1+e^{\eta_1}}\right) - 1\right) \log\left\{1 - \left(1 - e^{-(e^{\eta_4} x)^{e^{\eta_5}}}\right)^{e^{\eta_3}}\right\} + \\ &+ (a_1 - 1) \log\left(\frac{e^{\eta_1}}{1+e^{\eta_1}}\right) + (b_1 - 1) \log\left[\frac{1}{1+e^{\eta_1}}\right] + (a_2 - 1) \eta_2 - b_2 \eta_2 + \\ &+ (a_3 - 1) \eta_3 - b_3 e^{\eta_3} + (a_4 - 1) \eta_4 - b_4 e^{\eta_4} + (a_5 - 1) \eta_5 - b_5 e^{\eta_5} + \\ &+ \log\left[\frac{\exp(\eta_1 + \eta_2 + \eta_3 + \eta_4 + \eta_5)}{(1 + e^{\eta_1})^2}\right]. \end{aligned} \quad (6.20)$$

Infelizmente a implementação do processo MCMC para o vetor η não apresentou resultados satisfatórios, uma vez que não obteve-se a convergência das cadeias. Implementou-se, então, uma estratégia baseada em propostas individuais para cada

coordenada de η .

6.1.1 Função Geradora de Candidatos Univariado

Considere a proposta de uma função geradora de candidatos baseada na aproximação Gaussiana da distribuição a posteriori com respeito a η , dada por Gray (2001) na Seção 5.3.1. Dessa forma tem-se que

$$\eta \sim N_5(\hat{\eta}, I^{-1}(\hat{\eta})), \quad (6.21)$$

descreve a distribuição geradora de candidatos de η , onde $\hat{\eta}$ é a moda a posteriori de $\pi(\eta|X)$ e $I^{-1}(\hat{\eta})$ representa o inverso da matriz de informação observada, que é dada por menos o inverso da matriz hessiana de $\log(\pi(\eta|X))$ avaliada em $\hat{\eta}$, isto é,

$$I(\hat{\eta}) = - \left. \frac{\partial^2}{\partial \eta \partial \eta^T} \log[\pi(\eta|X)] \right|_{\eta=\hat{\eta}}. \quad (6.22)$$

Considere a matriz $D = \text{diag}(I^{-1}(\hat{\eta}))$. Então, teoricamente as amostras podem ser obtidas a partir da distribuição descrita em (6.21). No entanto, como mencionado, não logramos êxito com esse procedimento para a distribuição BEW. Considerou-se, então, a estratégia da geração de candidatos de η_i a partir da expressão (6.23)

$$\eta_i^{(s)} = D_{ii} Z_i^{(s)} + \hat{\eta}_i, \quad (6.23)$$

onde $Z_i^{(s)} \sim N(0, 1)$. No processo de Metropolis-Hastings a regra de aceitação (r) para o i -ésimo parâmetro é dada por

$$r_i = \frac{\pi(\eta_i^s|X)}{\pi(\eta_i^{old}|X)}. \quad (6.24)$$

6.1.2 Resultados

Implementando-se os desenvolvimentos da Seção 6.1.1 num algoritmo MCMC para a distribuição BEW e realizou-se um estudo para uma amostra gerada com $\mu = 0.66$, $\gamma = 24$, $\alpha = 3$, $\lambda = 5$ e $c = 6$ de tamanho 200. Adotou-se uma cadeia de tamanho 50.000, *burn-in* de 20.000 amostras e amostras sistemáticas com período

100.

Os valores iniciais foram escolhidos da seguinte maneira: primeiramente gerou-se 100 valores para cada parâmetro. Depois estimou-se o EMV para cada conjunto e avaliou se houve ou não a convergência no ponto estimado. Dentre todos os conjuntos que tiveram convergência, o ponto inicial escolhido foi a média desses valores iniciais.

A Tabela 6.1 apresenta os resultados das estimativas nesse desenvolvimento. Nota-se que, exceto pelo parâmetro c , os valores verdadeiros dos parâmetros estão inseridos no intervalo HPD. Em relação a taxa de aceitação, todos os parâmetros estão em níveis aceitáveis.

Tabela 6.1: Estimativas marginais à posteriori da Distribuição BEW utilizando a abordagem da Seção 6.1.

Parâmetro	Valor Real	Mediana	Média	HPD (2, 5%)	HPD(97, 5%)	Tx. de Aceitação(%)
μ	0,666	0,473	0,473	0,428	0,524	24,606
γ	24,000	18,670	18,670	14,191	23,584	45,518
α	3,000	2,382	2,394	2,099	2,780	24,860
λ	5,000	4,592	4,591	4,537	4,645	23,374
c	6,000	7,896	7,929	7,097	8,859	47,154

A Figura 6.1 apresenta os gráficos da densidade, do traço, da autocorrelação e os boxplot para cada parâmetro. Verifica-se que a densidade possui a unimodalidade esperada, a trajetória possui estacionariedade, a autocorrelação possui decaimento rápido. Pelos boxplots nota-se a aproximação dos valores gerados pela cadeia do verdadeiro valor em cada parâmetro.

6.1.3 Estudo Comparativo entre as abordagens frequentista e Bayesiana

Nesta seção será apresentado um estudo comparativo entre o método frequentista e o bayesiano utilizando a abordagem trabalhada nesse capítulo. Tomou-se o cuidado em utilizar as mesmas amostras nas estimações de ambos métodos. Primeiramente, para a seleção de pontos iniciais gerou-se 100 pontos aleatórios para cada parâmetro e obteve-se o EMV para cada conjunto de valores iniciais possíveis. Selecionou-se com ponto inicial a média dos pontos gerados em que as estimativas de máxima verossimilhança foram convergentes. Assumiu-se esses valores iniciais em todas as 50

amostras de tamanhos $n=50, 100, 200, 300, 500$ e 1000 .

Uma vez definido esses valores iniciais, selecionou-se apenas as amostras em que foi possível a obtenção de estimativas de moda a posteriori cuja matriz hessiana associada fosse positiva definida, de modo a permitir o cálculo da decomposição de *Choleski*. Feito isso, obteve-se as estimativas de máxima verossimilhança e as estimativas Bayesianas.

A Tabela 6.2 apresenta os resultados obtidos das 50 amostras. Para ambos os métodos, nota-se que existe uma lenta convergência para o verdadeiro valor do parâmetro. O parâmetro α , no caso Bayesiano, é sistematicamente estimado com maior vício e EQM do que o estimado via EMV. Para os outros parâmetros, as estimativas frequentistas e Bayesianas são muito próximas.

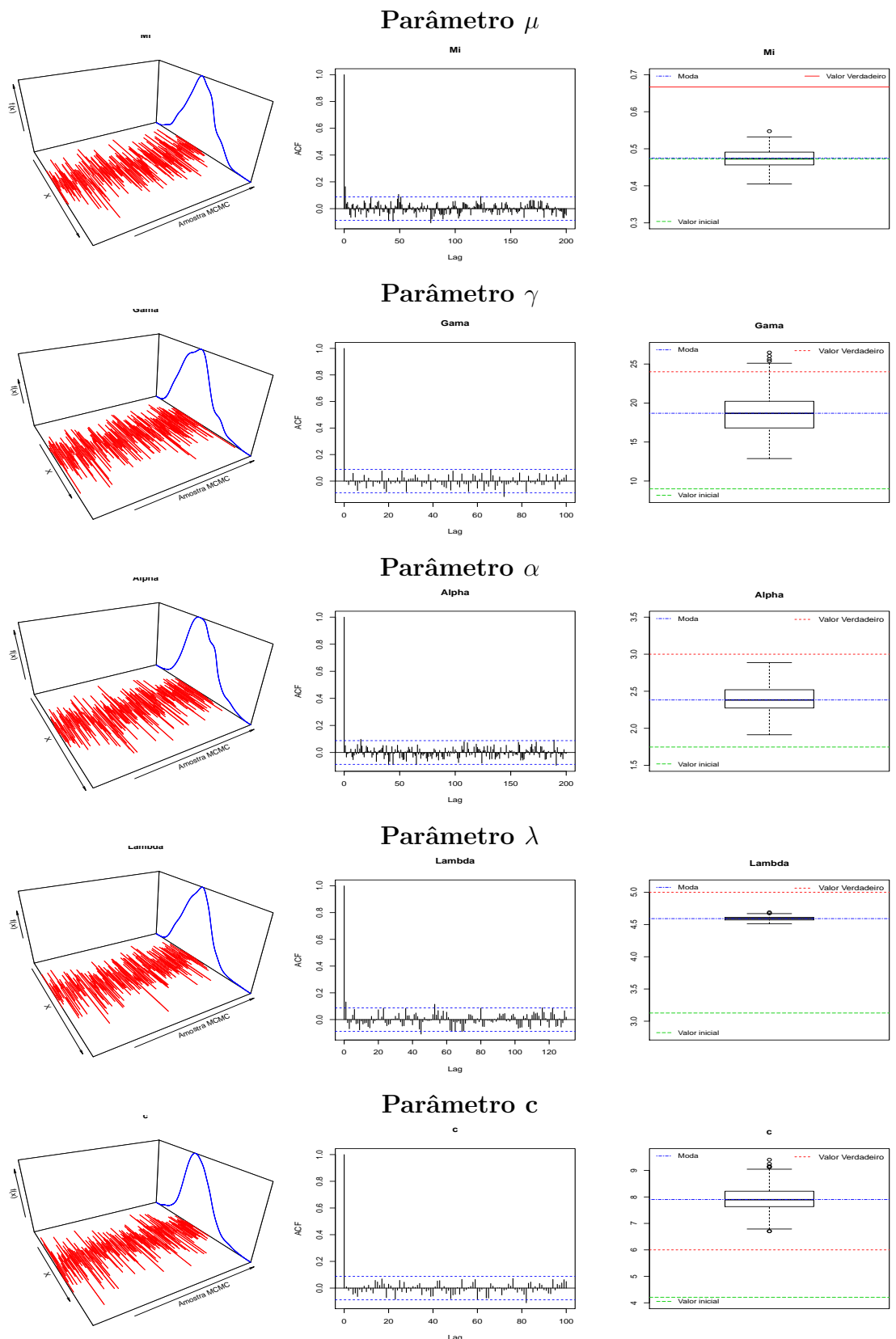


Figura 6.1: Traço, Densidade, Autocorrelação e Boxplots das amostras MCMC para os parâmetros μ , γ , α , λ e c da distribuição BEW sob a proposta da Seção 6.1

Capítulo 7

Aplicação a dados reais

Neste capítulo é apresentada uma aplicação a um conjunto de dados reais nos quais se aplicam o modelo BBXII e seus submodelos. O conjunto de dados refere-se a um estudo que teve por objetivo a prevenção da reincidência do melanoma cutâneo. Segundo Ibrahim; Chen e Sinha (2001), o período de coleta de dados dos pacientes foi de 1991 a 1995 e o acompanhamento foi até 1998. A variável resposta foi o tempo (em anos) até a morte do paciente. O tamanho original da amostra foi de 427 pacientes, dos quais 10 não apresentaram mudança no tamanho do tumor e forem retiradas da análise, restando 417 pacientes. Observou-se que 56% dos dados são censurados.

Dessa maneira ajustou-se a distribuição BBXII (definida em (1.6)) e seus submodelos. O objetivo é verificar qual distribuição proporciona melhor ajuste a esse conjunto de dados, sendo a distribuição BBXII dada por

$$f_{BBXII}(x) = \frac{ckx^{c-1}}{s^c B(a, b)} [1 + (x/s)^c]^{-(kb+1)} \{1 - [1 + (x/s)^c]^{-k}\}^{a-1}. \quad (7.1)$$

Para $a = b = 1$ a distribuição BBXII reduz-se à distribuição BurrXII (BXII) cuja densidade é dada por

$$f_{BXII}(x; s, k, c) = cks^{-c} \left[1 + \left(\frac{x}{s}\right)^c\right]^{-k-1} x^{c-1}. \quad (7.2)$$

Para $a = b = 1$, $s = m^{-1}$ e $k = 1$, a distribuição BBXII reduz-se à distribuição

log-logística (LL) cuja densidade é dada por

$$f_{LL}(x; m, c) = \frac{c}{m} \left(\frac{x}{m}\right)^{c-1} \left[1 + \left(\frac{x}{m}\right)^c\right]^{-2}. \quad (7.3)$$

No caso de $k \rightarrow \infty$, a distribuição BBXII converge para a distribuição Beta Weibull (Cordeiro, 2011):

$$f_{BW}(x; \mu, \gamma, \lambda, c) = \frac{c\lambda^c x^{c-1} \exp\{-\gamma(1-\mu)(\lambda x)^c\}}{B(\mu\gamma, \gamma(1-\mu))} [1 - \exp\{-(\lambda x)^c\}]^{\mu\gamma-1}. \quad (7.4)$$

Para $\mu\gamma = \gamma(1-\mu) = 1$, além de $k \rightarrow \infty$, a distribuição BBXII reduz-se a distribuição Weibull:

$$f_{Weibull}(x; \lambda, c) = c\lambda^c x^{c-1} \exp\{-(x\lambda)^c\}. \quad (7.5)$$

Nas próximas seções serão apresentados os resultados das estimativas de máxima verossimilhança para a distribuição BBXII e seus submodelos, além de seus erros padrões e os valores dos critérios de informação de *Akaike* (AIC), Bayesiano (BIC) e o *Akaike* corrigido (CAIC). Também serão apresentados os resultados obtidos pela abordagem Bayesiana, ou seja, as estimativas a posteriori da BBXII obtidas via MCMC, além dos intervalos de credibilidade considerando as abordagens das Seções 5.3 e 6.1.

Em ambas abordagens assumiu-se valores iniciais diferentes dos adotadas por Paranaíba (2012). Pois, verificou-se que os valores iniciais não resultam EMVs convergentes, nem matriz hessiana positiva definida, tanto na distribuição BBXII, quanto na BurrXII.

7.1 Abordagem Clássica

Para cada uma das distribuições apresentadas acima, de acordo com a correspondente função de verossimilhança é descrita pela expressão (2.24):

$$L(\theta|X) = \prod_{i=1}^n [f_{T_i}(t_i)]^{\delta_i} [S_{T_i}(t_i)]^{1-\delta_i}$$

Para a distribuição BBXII adotou-se, em ambos métodos (clássico e Bayesiano), os seguintes valores iniciais: $a_0 = 0.1, b_0 = 1.2, s_0 = 0.1, k_0 = 0.02$ e $c_0 = 5.5$, e conseqüentemente, $\mu = 0.076$ e $\gamma = 1.3$. Para a distribuição BurrXII foram os seguintes pontos iniciais: $s_0 = 0.87, k_0 = 0.14$ e $c_0 = 2.5$. Para a distribuição Beta Weibull foram os seguintes pontos iniciais: $a_0 = 1, b_0 = 0.5, \lambda_0 = 0.08$ e $c_0 = 5$, e conseqüentemente, $\mu_0 = 0.666$ e $\gamma_0 = 1.5$. Para a distribuição Log-logística foram os seguinte pontos iniciais: $c_0 = 2.5$ e $s_0 = 1$, conseqüentemente, $m_0 = 1/s_0 = 1/1 = 1$. Por fim, para a distribuição Weibull foram os seguintes pontos iniciais: $\lambda_0 = 0.1$ e $c_0 = 0.8$.

7.1.1 Resultados

A Tabela 7.1 apresenta as estimativas de máxima verossimilhança (e os correspondentes limites dos intervalos de confiança entre parênteses) obtidas para a distribuição BBXII e outros submodelos.

Tabela 7.1: EMV dos parâmetros dos modelos BBXII, BurrXII, Beta Weibull, Log-logística e Weibull para os dados de melanoma (Ibrahim et. al., 2001) e intervalos de confiança (95%) correspondentes.

Distribuição	μ	γ	s	k	c
BBXII	0,270 (0,268;0,272)	1,069 (1,061;1,076)	1,703 (1,702;1,704)	0,016 (0,015;0,017)	6,077 (6,068;6,086)
BurrXII			0,873 (0,872;0,874)	0,146 (0,145;0,147)	2,566 (2,565;2,567)
Beta Weibul	0,666 (0,662;0,669)	737,531 (630,383;884,678)	λ 2,896 (2,463;3,328)		0,035 (0,032; 0,037)
Log-logística	c 1,266 (1,264;1,267)	m 0,210 (0,207;0,212)			
Weibull	α 6,943 (6,942;6.945)	γ 1,050 (1,050;1,051)			

A Tabela 7.2 apresenta os critérios de informação de *Akaike* (AIC), o Bayesiano (BIC) e *Akaike* corrigido (CAIC) para os modelos ajustados para os dados de melanoma. Nota-se que os modelos BurrXII e BBXI possuem os menores valores de AIC, BIC e CAIC. Por outro lado, a distribuição Weibull apresenta os maiores valores.

A Figura 7.1 descreve uma análise gráfica comparando a função de sobrevivência

Tabela 7.2: AIC, BIC e CAIC das distribuições BBXII, BurrXII, Beta Weibull, Log-logística e Weibull para os dados melanoma.

Distribuição	AIC	BIC	CAIC
BBXII	1057,214	1077,380	1077,392
BurrXII	1058,348	1070,447	1070,455
Beta Weibul	1074,517	1082,583	1082,588
Log-logística	1086,319	1094,385	1094,389
Weibull	1102,108	1110,174	1110,179

empírica obtida a partir do estimador não paramétrico de Kaplan-Meier (Kaplan e Meier, 1958; Lawless, 1982) e a função de sobrevivência estimada para a distribuição BBXII e seus submodelos. Observa-se a boa adequabilidade do ajuste das distribuições BBXII e BurrXII.

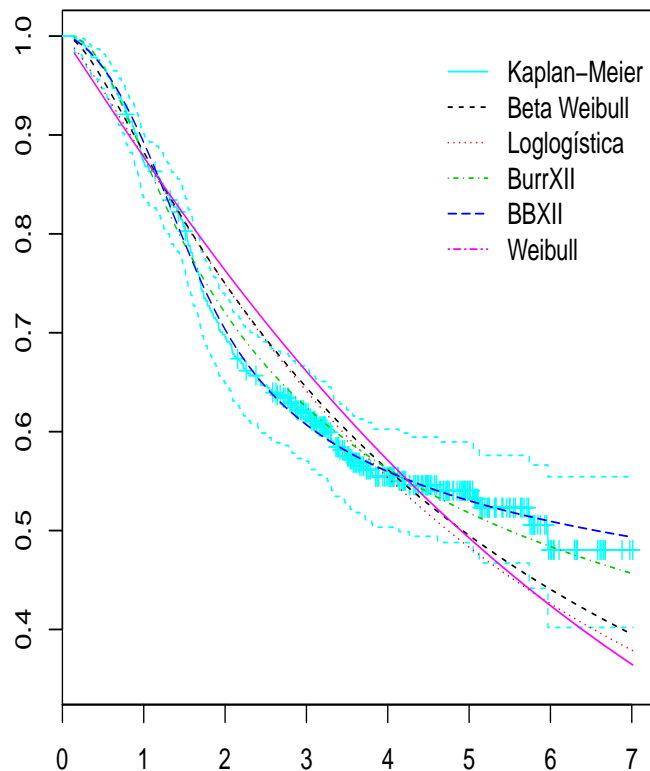


Figura 7.1: Função empírica de sobrevivência para os dados de Melanoma (Ibrahim et al, 2001).

A adequabilidade dos modelos também será avaliada através da análise de resíduos de Cox-Snell (Cox e Snell, 1968). Os resíduos de Cox-Snell são um tipo de resíduo padronizado utilizados em análises de confiabilidade. Os resíduos de Cox-Snell são iguais a menos o logaritmo natural da probabilidade de sobrevivência para cada observação, ou seja,

$$\hat{e}_i = -\log \left(\hat{S}(t_i|x_i) \right). \quad (7.6)$$

Nas análises gráficas dos resíduos constrói-se gráficos de resíduos Cox-Snell para o estimador de Kaplan-Meier ($\hat{S}(\hat{e}_i)_{KM}$) versus os resíduos de Cox-Snell para um modelo ajustado específico ($\hat{S}(\hat{e}_i)_{ME}$). Para modelos que proporcionam boas descrições dos dados, os pares $\left(\hat{S}(\hat{e}_i)_{KM}, \hat{S}(\hat{e}_i)_{ME} \right)$ estão próximos de uma reta que passa pela origem.

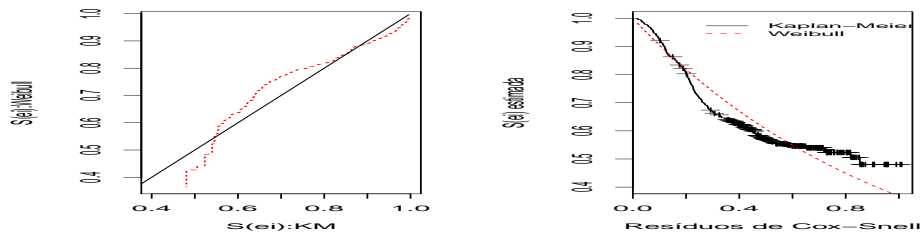
Pela Figura 7.2, comparando-de os gráficos de todos os modelos ajustados, confirma-se que o melhor ajustamento foram dos modelos BBXII e BurrXII, por meio da análise residual, o que indica a boa adequabilidade do ajuste.

7.2 Abordagem Bayesiana

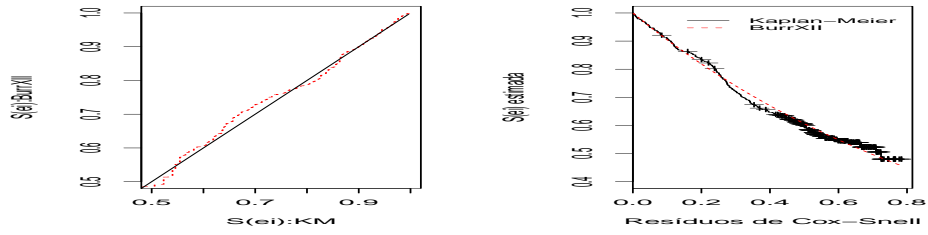
No caso Bayesiano, considerou-se *prioris* independentes com distribuição beta para o parâmetro μ e distribuição *gama* para os demais parâmetros. Trabalhou-se com a distribuição BBXII na escala “log” nos parâmetros e, conseqüentemente, a distribuição a posteriori é dada por

$$P(\eta_1, \eta_2, \eta_3, \eta_4, \eta_5 | T, C) \propto \left[\prod_{i=1}^n f(x|\eta) \right] \times \pi(\eta_1)\pi(\eta_2)\pi(\eta_3)\pi(\eta_4)\pi(\eta_5) \times |J|, \quad (7.7)$$

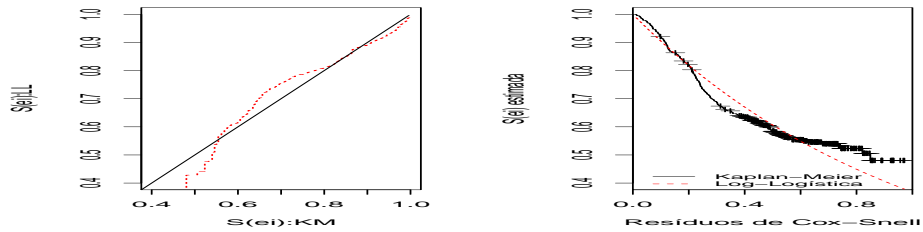
Distribuição Weibull



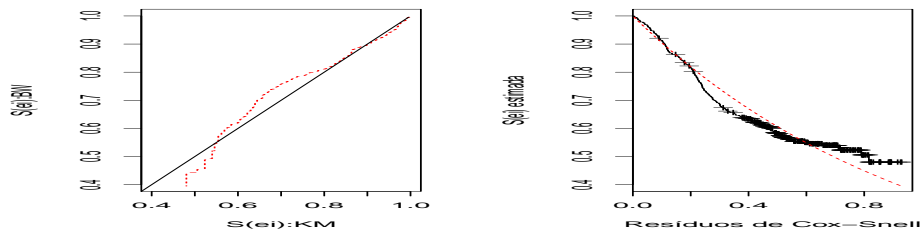
Distribuição BurrXII



Distribuição Log-logística



Distribuição Beta Weibull



Distribuição BBXII

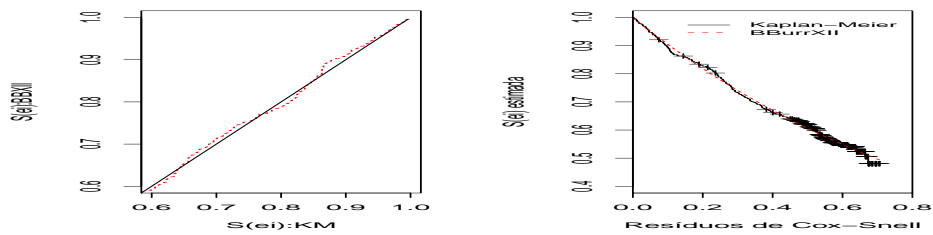


Figura 7.2: Análise gráfica dos resíduos de Cox-Snell dos modelos Weibull, BurrXII, Log-logística, Beta Weibull e BBXII ajustado aos dados de melanoma (Ibrahim et al, 2001).

onde T é o tempo, C é a censura, $|J| = \exp(\eta_1 + \eta_2 + \eta_3 + \eta_4 + \eta_5)/(1 + e^{\eta_1})^2$ e

$$\begin{aligned} \eta_1 &= \log\left(\frac{\mu}{1-\mu}\right), & \text{isto é,} & & \mu &= \frac{e^{\eta_1}}{1+e^{\eta_1}}, \\ \eta_2 &= \log(\gamma), & \text{isto é,} & & \gamma &= e^{\eta_2}, \\ \eta_3 &= \log(s), & \text{isto é,} & & s &= e^{\eta_3}, \\ \eta_4 &= \log(k), & \text{isto é,} & & k &= e^{\eta_4}, \\ \eta_5 &= \log(c), & \text{isto é,} & & c &= e^{\eta_5}. \end{aligned}$$

Na seção seguinte apresenta-se os resultados obtidos sob as abordagens trabalhadas nas seções 5.3 e 6.1 para a distribuição BBXII.

7.2.1 Resultados

A Tabela 7.3 apresenta os resultados das estimativas Bayesianas das seções 5.3 e 6.1 para a distribuição BBXII. Nota-se que as estimativas pontuais são similares em ambos os métodos, mas os intervalos HPD obtido pela abordagem da Seção 5.3 possui amplitude menor do que os obtidos com a abordagem da Seção 6.1. A taxa de aceitação do método proposto na Seção 5.3 é bastante alta (66%), ao passo que as taxas de aceitação obtidas para o método descrito na Seção 6.1 estão em níveis aceitáveis.

A Figura 7.3 apresenta os boxplots, a autocorrelação, o traço e a densidade estimada para todos os parâmetros obtidos pela abordagem Bayesiana da Seção 5.3. A trajetória se manteve estacionária ao longo dos passos da cadeia, a densidade possui aspecto unimodal e a autocorrelação decai rapidamente para todos os parâmetros da distribuição BBXII.

A Figura 7.4 apresenta os boxplots, a autocorrelação, o traço e a densidade estimada para todos os parâmetros obtido pela abordagem Bayesiana da Seção 6.1. Nota-se que a trajetória se manteve estacionária ao longo dos passos da cadeia, a densidade possui aspecto unimodal e a autocorrelação decai rapidamente em todos os parâmetros da distribuição BBXII.

A Figura 7.5 apresenta a curva empírica de sobrevivência obtida com as estimativas Bayesianas, sob as abordagens das seções 5.3 e 6.1, e de máxima verossimilhança.

Tabela 7.3: Estimativas Bayesianas dos parâmetros dos modelos BBXII para os dados de melanoma, intervalos HPD (95%) correspondentes e taxas de aceitação.

Estimativas	μ	γ	s	k	c
Aproximação gaussiana e vetor de candidatos (Seção 5.3)					
Moda a posteriori	0,440	0,622	1,715	0,052	6,382
Mediana a posteriori	0,441	0,622	1,715	0,052	6,380
Média a posteriori	0,441	0,623	1,715	0,052	6,381
HPD (2,5%)	0,400	0,557	1,704	0,037	6,243
HPD (97,5%)	0,477	0,687	1,727	0,064	7,521
Taxa de Aceitação (%)	66,842	66,842	66,842	66,842	66,842
		AIC	BIC	CAIC	
		1061,151	1089,382	1061,425	
Aproximação gaussiana e candidatos univariados (Seção 6.1)					
Moda a posteriori	0,440	0,622	1,715	0,052	6,382
Mediana a posteriori	0,442	0,613	1,716	0,051	6,403
Média a posteriori	0,442	0,624	1,720	0,053	6,425
HPD (2,5%)	0,387	0,463	1,488	0,027	5,188
HPD (97,5%)	0,508	0,803	1,992	0,084	7,887
Taxa de Aceitação (%)	42,276	35,904	48,670	43,706	50,068
		AIC	BIC	CAIC	
		1061,155	1089,387	1061,429	

Nota-se a sobreposição das curvas e a grande aderência com a curva de kaplan-Meier.

Para verificar a adequabilidade do modelo BBXII sob a abordagem Bayesiana calculou-se os resíduos Cox-Snell. Nota-se pela Figura 7.6 que, para ambas as abordagens, os resíduos possuem comportamento adequado e muito próximo.

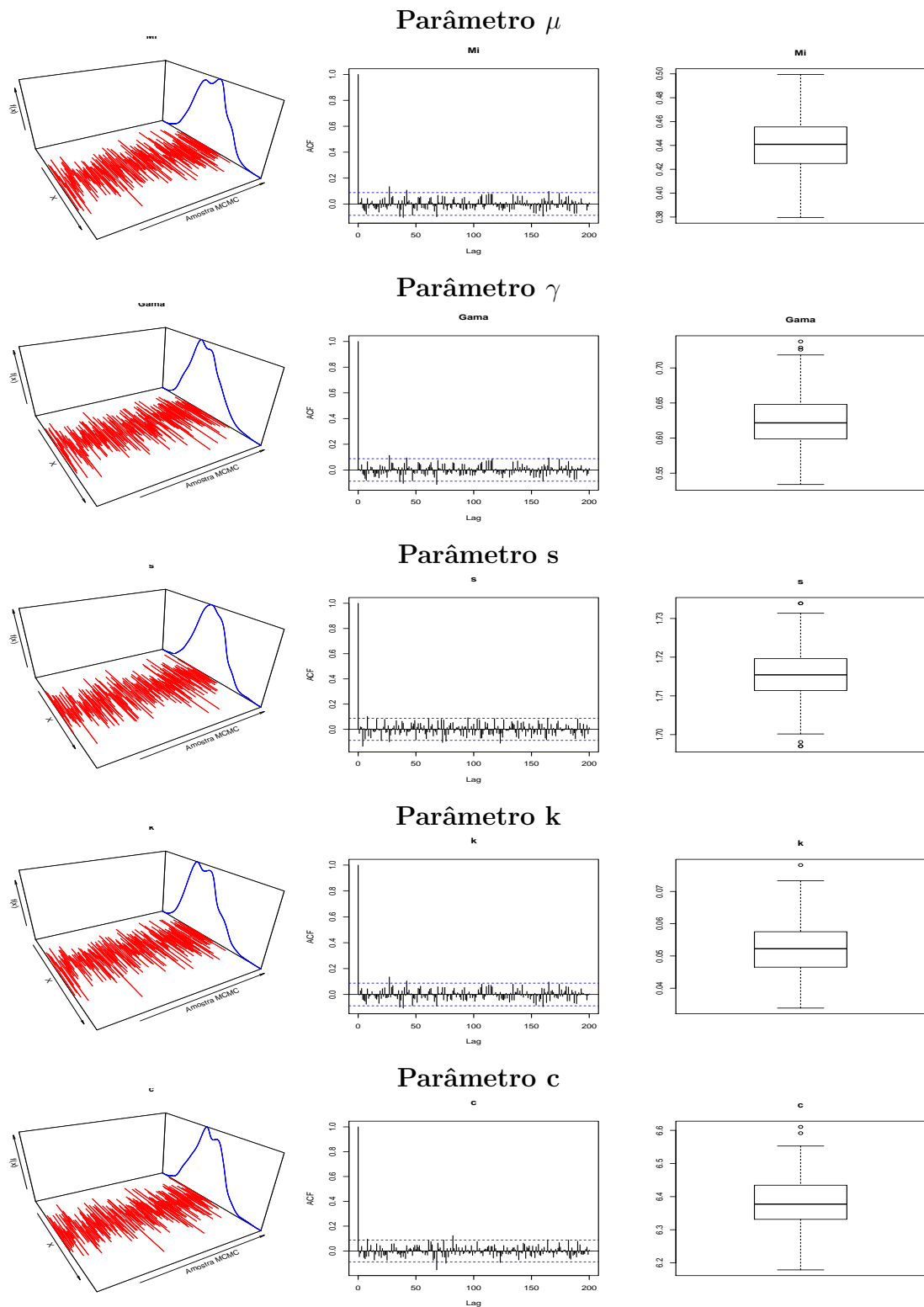


Figura 7.3: Traço, Densidade, Autocorrelação e Boxplots das amostras MCMC para os parâmetros μ , γ , s , k e c da distribuição BBXII obtido via aproximação gaussiana dos parâmetros e vetor de candidatos - Metodologia proposta na Seção 5.3.

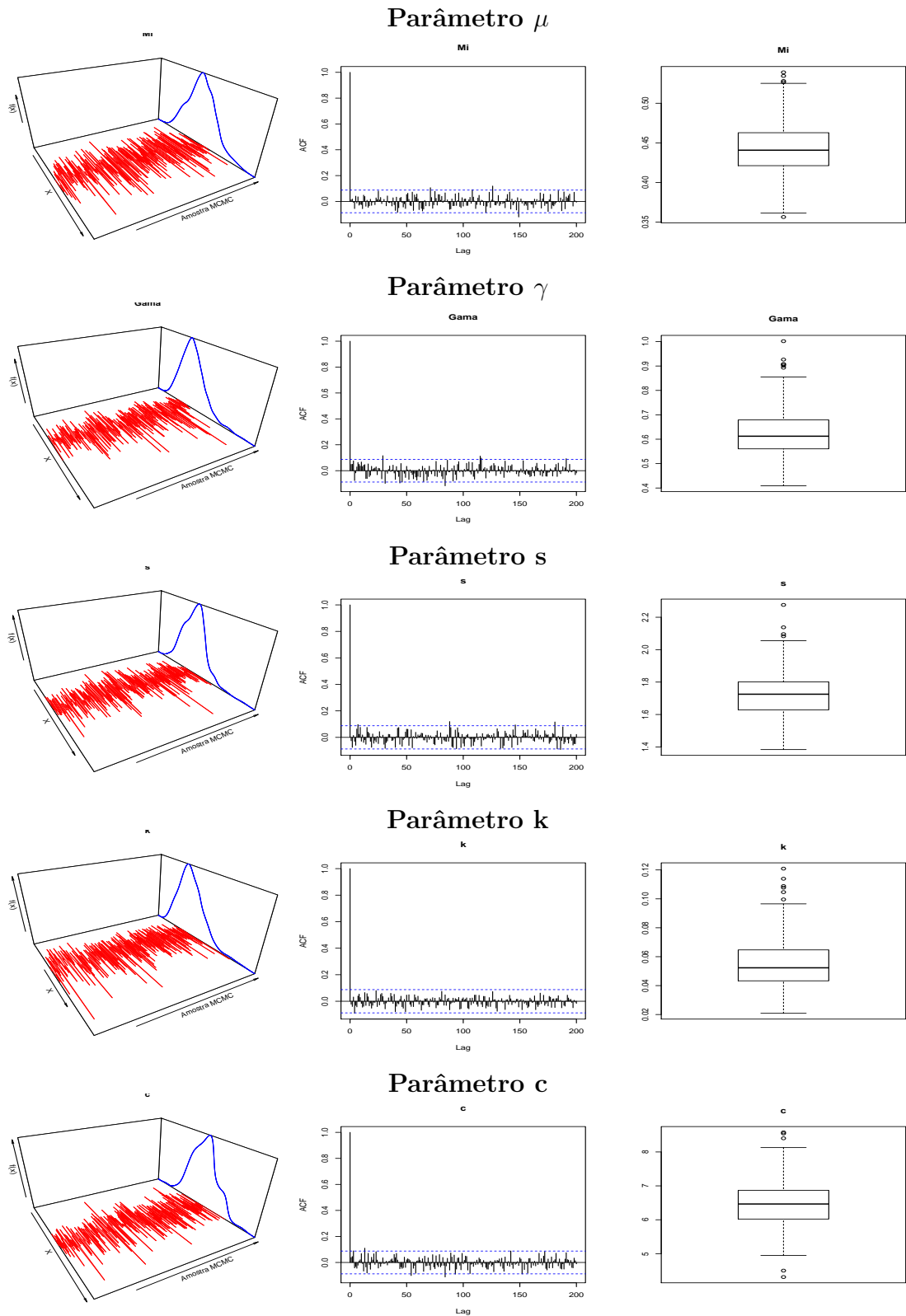


Figura 7.4: Traço, Densidade, Autocorrelação e Boxplots das amostras MCMC para os parâmetros μ , γ , s , k e c da distribuição BBXII obtido via aproximação gaussiana dos parâmetros e candidatos univariados- Metodologia proposta na Seção 6.1

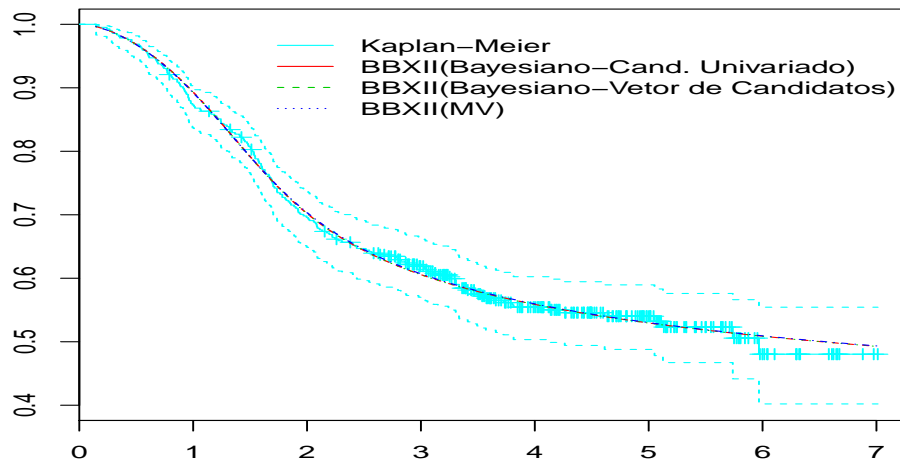
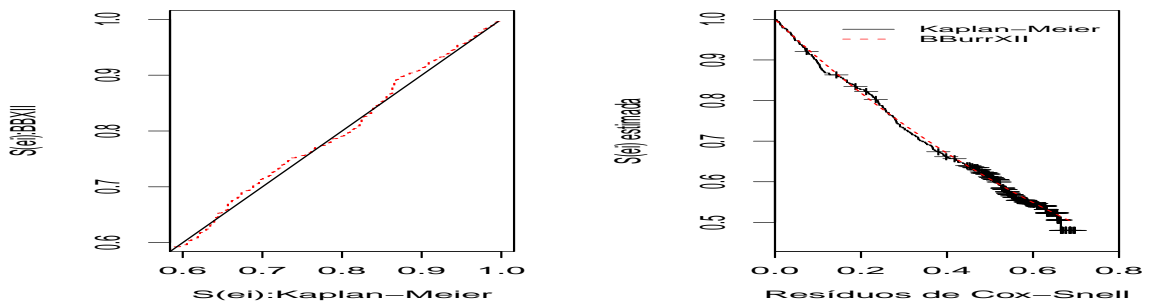


Figura 7.5: Função empírica de sobrevivência ajustada aos dados de Melanoma (Ibrahim et al, 2001).

Candidatos Univariados - Abordagem da Seção 6.1



Vetor de Candidatos - Abordagem da Seção 5.3

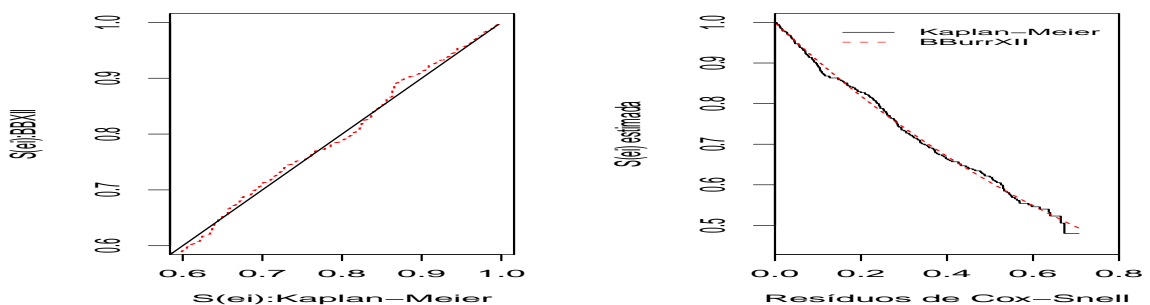


Figura 7.6: Análise gráfica dos resíduos de Cox-Snell do BBXII obtidos pelas abordagens das Seções 6.1 e 5.3 e ajustados aos dados de melanoma (Ibrahim et al, 2001).

Capítulo 8

Discussão

Neste trabalho de dissertação objetivou-se o desenvolvimento de métodos Bayesianos para a estimação dos parâmetros das distribuições BBXII e BEW. O desenvolvimento de metodologia adequada para tal propósito revelou-se bastante desafiador, devido a complexidade de tais modelos.

Tentativas frustradas do uso de alguns pacotes computacionais destinados à estimação Bayesiana, tais como o WinBUGS e JAGS (Capítulo 4) motivaram o desenvolvimento de nossas próprias rotinas em R. Tais pacotes não são dotados de flexibilidade suficiente para permitir a descrição de métodos Bayesianos que não seguem determinada construção padrão.

Uma vez que os modelos BBXII e BEW incorporam elementos da distribuição Beta, inicialmente trabalhou-se com os parâmetros desta última distribuição em sua forma original. No entanto, em nosso estudo (Capítulo 5, Seção 5.2) constatou-se que o uso da parametrização da distribuição Beta, como descrita em Ferrari e Cribari-Neto (2004), traz vantagens computacionais e conduz à convergência das cadeias de Markov geradas via método MCMC. No entanto, as distribuições geradoras de candidatos de alguns parâmetros ainda necessitavam de maiores desenvolvimentos de modo a robustificar o método.

De modo a obter estimativas mais precisas dos parâmetros e cobertura adequada nos intervalos de credibilidade, foram desenvolvidas métodos baseados em aproximações Gaussianas (Capítulo 5, Seção 5.3 e Capítulo 6) para a descrição de distribuições geradoras de candidatos dos parâmetros, agora expressos ou na escala log

ou logit e, portanto, definidos nos reais.

Alguns estudos de simulação foram empreendidos de modo a permitir a comparação entre as metodologias Bayesianas que foram propostas nesta dissertação e as estimativas obtidas através de estimação por máxima verossimilhança. Em linhas gerais os métodos Bayesianos e de máxima verossimilhança conduziram a estimativas pontuais muito próximas.

No entanto, as estimativas por máxima verossimilhança conduzem a intervalos de confiança estreitos em demasia, o que nos causa a suspeita de que os mesmos são irrealisticamente otimistas, implicando em pequena probabilidade de cobertura. Na prática isso significa que, frequentemente os intervalos de confiança construídos não incluirão o verdadeiro valor do parâmetro. Um exemplo deste comportamento pode ser observado nos intervalos de confiança (credibilidade) na Tabela 7.3.

Além dos aspectos acima, no cálculo dos intervalos de confiança frequentistas faz-se necessário a inversão da matriz de informação observada, que, muitas vezes, não é positiva definida, impossibilitando o cálculo do desvio-padrão das estimativas, por consequência, dos respectivos intervalos de confiança.

Por outro lado, na construção dos intervalos de credibilidade, para o método Bayesiano, utiliza-se simplesmente as amostras MCMC após a convergência da cadeia, pois não se utiliza nenhum argumento assintótico, tal como no caso frequentista.

Os desenvolvimentos da Seção 5.2.2, em que se utiliza *prioris* Qui-Quadrado para vários parâmetros, nos parece promissor, no sentido de que o método proporciona boa varredura do espaço paramétrico.

No Capítulo 7 descreveu-se nossas análises Bayesianas para o modelo BBXII utilizando os dados de melanoma de Ibrahim et al. (2001). Utilizando o mesmo conjunto de dados, Paranaíba et. al. (2011), procedeu a análises Bayesianas para o modelo BBXII utilizando *prioris* gama(0.01,0.01) para todos os parâmetros (com a e b na sua parametrização original).

Os autores utilizaram distribuições geradores de candidatos gaussianas para todos os parâmetros (que são não-negativos), tomando-se o valor absoluto dos valores gerados. No entanto, tais distribuições geradoras de candidatos gaussianas foram parametrizadas com médias fixadas nos valores estimados dos parâmetros que foram obtidos pelo método de máxima verossimilhança. Adicionalmente, para o

parâmetro de dispersão adotou-se um valor extremamente pequeno, conduzindo a pequena varredura do espaço paramétrico.

Acreditamos que os métodos desenvolvidos nesta dissertação conduzem a um avanço em relação a proposta Bayesiana de Paranaíba et. al.(2011) para a distribuição BBXII. Além disso, propusemos um método Bayesiano (que acreditamos seja o primeiro) para a distribuição BEW.

Os trabalhos futuros envolvem a execução de mais simulações e a escrita de um artigo para ser submetido a um periódico internacional.

Referências Bibliográficas

ACHCAR, J. A.; COELHO-BARROS, E. A.; CORDEIRO, G. M.. *Beta generalized distributions and related exponentiated models: a Bayesian approach*. Brazilian Journal of Probability and Statistics, 27(1), 1-19,2013.

BURR,I.W.. *Cumulative frequency functions*, Annals of Mathematical Statistics, 13, 215-232, 1942.

CORDEIRO, G.M.; GOMES, A.E.; DA SILVA, C.Q.; ORTEGA, E.M.M.. *The beta exponentiated Weibull distribution*. J Stat Comput Simul, 83: 114-138, 2013.

CORDEIRO, G.M.; SIMAS, A. B. ; STOSIC, B.. *Closed Form expressions for moments of the beta Weibull distribution*. Anais da Academia Brasileira de Ciências (Impresso), v. 83, p. 357-373, 2011.

COX, D. R.; SNELL, E. J. A general definition of residuals. Journal of the Royal Statistical Society B, London, v. 30, n. 2, p. 248-254, Mar. 1968.

DENWOOD, M. J..*runjags: An R package providing interface utilities, parallel computing methods and additional distributions for MCMC models in JAGS*. Journal of Statistical Software. Retrieved from <http://cran.r-project.org/web/packages/runjags/>.

EHLERS, R. S..*Inferência Bayesiana*. Disponível em:<http://conteudo.icmc.usp.br/pessoas/ehlers/bayes/> . Acesso em: Outubro de 2016.

EUGENE, N.; LEE, C.; FAMOYE, F..*Beta-normal distribution and its applications*. Commun. Statist. - Theory and Methods, 31, 497-512, 2002.

FERRARI, S.L.P.; CRIBARI-NETO, F.. *Beta regression for modelling rates and proportions*. Journal of Applied Statistics, 31(7), 799-815, 2004.

GEWEKE, J.; *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (with discussion)*. In Bayesian Statistics. (J.M. Bernardo,

- J. O. Berger, A. P. Dawid e A. F. M. Smith, eds), 169-193. Oxford University Press, Oxford, 1992.
- GELFAND, A.; SMITH, A.. *Sampling - based approaches to calculating marginal densities*. Journal of the American Statistical Association, (85) : 398 - 409, 1990.
- GELMAN, A.; RUBIN, D. B.. *Inference from iterative simulation using multiple sequences* *Statistical Science*, 7, 457-511, 1992.
- GEMAN, S., GEMAN, D.. *Stochastic relaxation, gibbs distribution and the bayesian restoration of images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, (6): 721-741, 1984.
- GRAY, R.. *BIO 248 cd Advanced Statistical Computing*. Disponível em: <https://www.ida.liu.se/732A26/info/report.pdf>. Acesso em: Agosto de 2016.
- HATINGS, W.. *Monte Carlo sampling methods using markov chains and their applications*. Biometrika, (57):97-109, 1970.
- HOFF, P. D.. *A first course in Bayesian statistical methods*. Springer Science & Business Media, 2009.
- LAWLESS, J. F.. *Statistical Models and Lifetime Data*. Ed. John Wiley & Sons, 1982.
- LUNN, D.; JACKSON, C.; BEST, N., THOMAS, A.; SPIEGELHALTER, D.. *The BUGS book: A practical introduction to Bayesian analysis*. Boca Raton, FL: CRC Press, 2013.
- KAPLAN, E. L.; MEIER, P.: *Nonparametric estimation from incomplete observations*. J. Amer. Statist. Assn. 53:457-481, 1958.
- KRUSCHKE, J.; *Doing Bayesian Data Analysis A Tutorial with R, JAGS, and Stan*, Academic Press, 2 Edição, 2014.
- McCORMIK, T. H.; RAFTERY, A. E.; MADIGAN, D.; BURD, R. S.. Dynamic logistic regression and dynamic model averaging for binary classification. *Biometrics*, 68(1), 23-30, 2012.
- METROPOLIS, N.; ROSENBLUTH, A., ROSENBLUTH, M., TELLER, A., TELLER, E.. *Equations of state calculations by fast computing machines*. Journal of Chemical Physics, (21): 1087-1092, 1953.

MUDHOLKAR,G.S.; SRIVASTAVA,D.K..*Exponentiated Weibull family for analyzing bathtub failure-real data*, IEEE Trans. Reliab. 42, pp. 299–302, 1993.

NADARAJAH, S,; KOTZ, S.. *The beta Gumbel distribution*. Mathematical Problems in Engineering, New York, v.10, p.323-332, 2004.

O'HAGAN, A.. *Bayesian Inference*. Kendall's Advanced Theory of Statistics, Vol. 2B. Arnold, London, 1994.

PARANAÍBA, P. F.. *Caracterização e extensões da distribuição Burr XII: propriedades e aplicações*. 2012. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2012. Disponível em: <http://www.teses.usp.br/teses/disponiveis/11/11134/tde-29102012-085146/> . Acesso em: 2016-06-07.

PARANAÍBA, P. F.; ORTEGA, E. M.; CORDEIRO, G. M.; PESCIM, R. R.. *The beta Burr XII distribution with application to lifetime data*. Computational Statistics & Data Analysis, 55(2), 1118-1136, 2011.

R Core Team.*R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>, 2013.

SPIEGELHALTER, D.J.; THOMAS, A.. *WinBUGS Version 1.2 User Manual*. MRC Biostatistics Unit, 1999.

UGRAY, Z.; LASDON, L.; PLUMMER, J.; GLOVERr, F.; KELLY, J.. *R.: Scatter Search and Local NLP Solvers: A Multistart Framework for Global Optimization*. INFORMS Journal on Computing, 19, No. 3, 328-340, 2007.

TAHIR, M.H.; NADARAJAH, S.. *Parameter induction in continuous univariate distributions: Well-established G families*. Anais da Academia Brasileira de Ciências (Annals of the Brazilian Academy of Sciences), 87(2): 539-568, 2015.

WEIBULL, W. . Wide applicability. Journal of applied mechanics, 103, 33, 1951.

ZIMMER, W. J.; KEATS, J. B.; WANG, F.K.. *The Burr XII distribution in reliability analysis*. Journal of Quality Technology, Milwaukee, v. 30, p. 389-394, 1998.

Apêndice A:

Elicitação de Funções Geradoras de Candidatos Poisson para os parâmetros a e b

Nessa seção será proposto uma outra forma para obtenção das estimativas bayesianas, em especial para os parâmetros a e b . Para tanto, serão utilizadas algumas propriedades da Distribuição BBXII, inclusive relacionadas à família exponencial.

Objetiva-se criar distribuições geradoras de candidatos para os parâmetros a e b , baseadas na Distribuição Poisson com parâmetros definidos através de um processo de elicitacão a ser detalhado a seguir.

A distribuição BBXII pode ser escrita através de

$$f(x) = g(x) \frac{1}{B(a, b)} G(x)^{a-1} [1 - G(x)]^{b-1} \quad (8.1)$$

$$= g(x; s, k, c) h(x; a, b, s, k, c),$$

$$G(x; s, k, c) = 1 - \left[1 + \left(\frac{x}{s} \right)^c \right]^{-k} = G(x) \quad (8.2)$$

$$h(x; a, b, s, k, c) = \frac{1}{B(a, b)} G(x)^{a-1} [1 - G(x)]^{b-1}, \quad (8.3)$$

e

$$g(x) = cks^{-k} \left[1 + \left(\frac{x}{s} \right)^c \right]^{-(k+1)} x^{c-1}. \quad (8.4)$$

Dessa forma, a distribuição a *posteriori* envolvendo os parâmetros do modelo pode

ser escrita como

$$\pi(a, b, s, k, c|x) \propto \prod_{i=1}^n g(x_i, s, k, c)h(x; a, b, s, k, c)\pi(a, b, s, k, c). \quad (8.5)$$

A distribuição condicional completa conjunta do vetor (a, b) é dada por:

$$\pi(a, b|s, k, c, x) \propto \left[\prod_{i=1}^n h(x_i, a, b, s, k, c) \right] \pi(a, b). \quad (8.6)$$

Logo, a verossimilhança tem a seguinte forma:

$$\begin{aligned} \prod_{i=1}^n h(x_i, a, b, s, k, c) &= \prod_{i=1}^n \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} G(x_i)^{a-1} (1-G(x_i))^{b-1} \\ &= \exp \{ aT_1(G(x_i)) + bT_2(G(x_i)) - \gamma(a, b) \} \end{aligned} \quad (8.7)$$

onde

$$\begin{aligned} \gamma(a, b) &= -n [\log(\Gamma(a+b)) - \log(\Gamma(a)) - \log(\Gamma(b))] \\ T_1(G(x)) &= \sum_{i=1}^n \log(G(x_i)) \quad e \quad T_2(G(x)) = \sum_{i=1}^n \log(1-G(x_i)). \end{aligned} \quad (8.8)$$

Pelas propriedades da família exponencial,

$$\begin{aligned} E(T_1) &= \frac{\partial \gamma}{\partial a} = n(\psi(a) - \psi(a+b)), & E(T_2) &= \frac{\partial \gamma}{\partial b} = n(\psi(b) - \psi(a+b)) \\ V(T_1) &= \frac{\partial^2 \gamma}{\partial a^2} = n(\psi'(a) - \psi'(a+b)), & V(T_2) &= \frac{\partial^2 \gamma}{\partial b^2} = n(\psi'(b) - \psi'(a+b)) \end{aligned}$$

e

$$cov(T_1, T_2) = \frac{\partial^2 \gamma}{\partial a \partial b} = -n\psi'(a, b),$$

onde

$$\psi(z) = \frac{\gamma'(z)}{\gamma(z)} \quad e \quad \psi'(z) = \frac{d\psi(z)}{dz},$$

que são, respectivamente, as funções *digama* e *trigama*. Seja,

$$\begin{aligned}\mu &= E \left[\sum_{i=1}^n \log \left(\frac{G(x_i)}{1 - G(x_i)} \right) \right] = \\ &= E \left[\sum_{i=1}^n \log (G(x_i)) - \sum_{i=1}^n \log (1 - G(x_i)) \right] \\ &= E(T_1) - E(T_2) = n(\psi(a) - \psi(b))\end{aligned}\quad (8.9)$$

e

$$\begin{aligned}\sigma^2 &= Var \left(\sum_{i=1}^n \log \left(\frac{G(x_i)}{1 - G(x_i)} \right) \right) = Var (T_1 - T_2) \\ &= Var(T_1) + Var(T_2) - 2Cov (T_1, T_2) = \\ &= n \left(\psi'(a) + \psi'(b) \right)\end{aligned}\quad (8.10)$$

Para grande valores de z , $\psi(z) \approx \log(z)$, enquanto $\psi'(z) \approx z^{-1}$. Com essas aproximações tem-se que:

$$\mu \approx n \log \left(\frac{a}{b} \right) \quad (8.11)$$

e

$$\sigma^2 \approx n \left(\frac{1}{a} + \frac{1}{b} \right). \quad (8.12)$$

Logo,

$$a = \frac{n}{\sigma^2} \left(1 + \exp \left(\frac{\mu}{n} \right) \right), \quad (8.13)$$

$$b = \frac{n}{\sigma^2} \left(1 + \exp \left(-\frac{\mu}{n} \right) \right). \quad (8.14)$$

Note que as fórmulas (8.13) e (8.14) indicam que a e b podem ser elicitadas quando os valores de μ e σ^2 são conhecidos. Portanto, faz-se necessário calcular estes valores.

Note que

$$\begin{aligned}\mu &= E \left[\sum_{i=1}^n \log \left(\frac{G(x_i)}{1 - G(x_i)} \right) \right] \\ &= \sum_{i=1}^n E \left[\log \left(\frac{G(x_i)}{1 - G(x_i)} \right) \right] = nE \left[\log \left(\frac{G(x_i)}{1 - G(x_i)} \right) \right],\end{aligned}\quad (8.15)$$

onde as $G(x_i)$'s são *i.i.d.*. Então

$$E \left[\log \left(\frac{G(x)}{1-G(x)} \right) \right] = \int \log \left(\frac{G(x)}{1-G(x)} \right) f(x) dx \quad (8.16)$$

onde $f(x)$ é a densidade da BBXII. O cálculo de (8.16) encontra-se no Apêndice D. Além disso,

$$\begin{aligned} \sigma^2 &= Var \left(\sum_{i=1}^n \log \left(\frac{G(x_i)}{1-G(x_i)} \right) \right) = \sum_{i=1}^n Var \left(\log \left(\frac{G(x_i)}{1-G(x_i)} \right) \right) \\ &= \sum_{i=1}^n \left\{ E \left[\log \left(\frac{G(x_i)}{1-G(x_i)} \right) \right] - \frac{\mu^2}{n} \right\} \\ &= nE \left\{ \left[\log \left(\frac{G(x_i)}{1-G(x_i)} \right) \right]^2 \right\} - \mu^2. \end{aligned} \quad (8.17)$$

Os cálculos de (8.17) também são encontrados no Apêndice D. Dessa forma, a condicional de (a,b) é dada por:

$$\pi(a, b|s, k, c, x) \propto \left[\prod_{i=1}^n h_i(x_i, a, b, s, k, c) \right] \pi(a, b). \quad (8.18)$$

Tomando $\pi(a, b) = \pi(a)\pi(b)$ e função geradora de candidatos $P(a^*)$ e $P(b^*)$, onde a^* e b^* são os valores obtidos na elicitacão de a e b descritos através das fórmulas (8.13) e (8.14), com μ e σ obtidos por (8.15) e (8.17), respectivamente, tem-se que

$$a^{*(i)} = a(s^{(i)}, c^{(i)}, k^{(i)}) \quad \text{e} \quad b^{*(i)} = b(s^{(i)}, c^{(i)}, k^{(i)}). \quad (8.19)$$

Uma observacão adicional a ser feita é que, nesta abordagem, os valores gerados pela função geradora de candidatos são obrigatoriamente inteiros.

Simulacão

De posse dos cálculos acima, foram realizadas as devidas adaptaçoes no algoritmo de MCMC na tentativa de melhoria dos resultados bayesianos obtidas nas seções anteriores. Diante disso, verificou-se problemas nesse algoritmo relacionado com os valores iniciais dos parâmetros de interesse. Como exemplo, se fossem considerados

os valores iniciais como os verdadeiros, o algoritmo sempre retornava um erro.

Para resolver esse problema, utilizou-se o comando *MultiStart* do pacote *BB* do software *R*, para obter pontos iniciais ótimos para a função de verossimilhança de BBXII. De posse desses pontos iniciais, foi resolvido esse problema, mas observou-se que no algoritmo MCMC os resultados estimados eram equivalentes aos pontos iniciais para alguns parâmetros, ou seja, os pontos iniciais determinavam as estimativas marginais a posteriori.

Num estudo de simulação foram obtidos 20 vetores de pontos iniciais “ótimos” e, assim, realizado o procedimento de MCMC. Na Figura 8.1, visualiza-se que, para os parâmetros a e b , as estimativas marginais a posteriori são iguais aos valores iniciais.

Outro ponto problemático foi que os dados foram gerados com $a=16$ e $b=8$ e, nesse procedimento, os parâmetros foram estimados próximos de 2, em todos os casos.

Também observou-se problemas relacionados a taxa de aceitação, conforme é apresentado na Tabela 8.1. Por outro lado, a autocorrelação ficou em níveis aceitáveis. Diante dessas constatações, resolveu-se tentar outras formas para resolução desses problemas.

Tabela 8.1: Estimativas a posteriori (MCMC) da BBurrXII

	Valor Real	Média	Desvio Padrão	2.5%	50%	97.5%
a	16.0	3.916	0.073	3.844	3.919	3.919
b	8.0	1.666	0.028	1.664	1.664	1.707
s	1.0	0.772	0.003	0.765	0.772	0.780
k	5.0	1.455	0.059	1.341	1.454	1.579
c	6.0	17.217	0.789	15.393	17.242	18.594

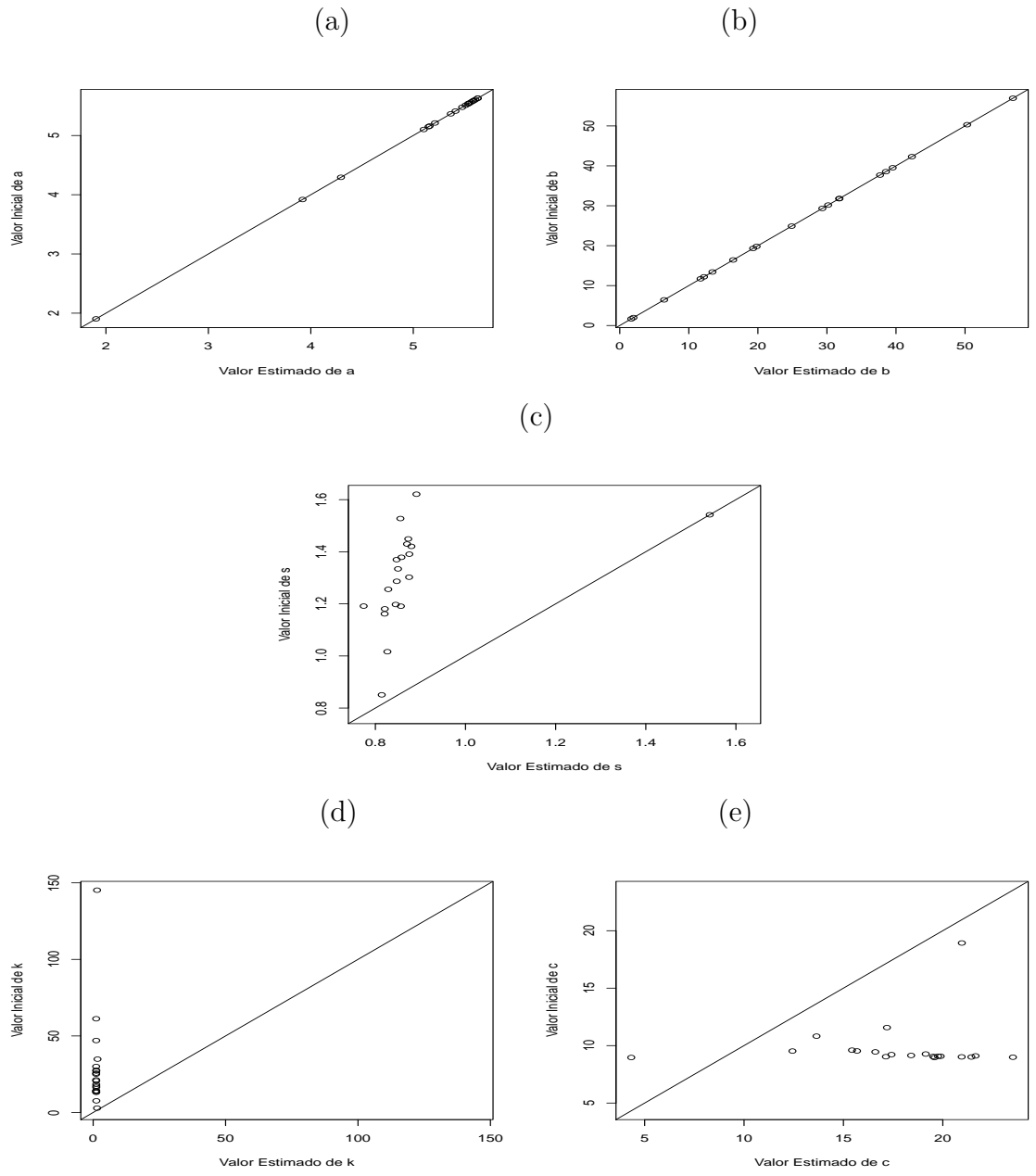


Figura 8.1: Gráfico dos valores iniciais e dos estimados

Apêndice B: Aproximações Gaussianas para a e b via Método de McCormick

Como os resultados anteriores não foram satisfatórios, a nova abordagem é baseada nos métodos de McCormick (2012) que propõe aproximações gaussianas para distribuições a posteriori e estimação de parâmetros de interesse utilizando variações do método de Newton-Raphson.

Como já descrito anteriormente, a distribuição BBXII pode ser escrita como

$$f(x) = \frac{g(x)}{G(x)} G(x)^{a-1} (1 - G(x))^{b-1} = f(x, a, b, s, k, c) \quad (8.20)$$

onde:

$$G(x) = G(x; s, k, c) = 1 - \left[1 + \left(\frac{x}{s}\right)^c\right]^{-k} \quad (8.21)$$

$$g(x) = g(x; s, k, c) = cks^{-c} \left[1 + \left(\frac{x}{s}\right)^c\right]^{-(k+1)}. \quad (8.22)$$

Seja

$$h(x) = \frac{1}{B(a, b)} G(x)^{a-1} (1 - G(x))^{b-1}, \quad (8.23)$$

logo, a função densidade da BBurrXII pode ser escrita da seguinte forma:

$$f(x; a, b, s, k, c) = g(x; s, k, c)h(x; a, b, s, k, c). \quad (8.24)$$

Considerando a perspectiva Bayesiana, a posteriori conjunta de $\theta = (a, b, s, k, c)$ é dada por:

$$\pi(a, b, s, k, c|x) \propto \left[\prod_{i=1}^n f(x_i; a, b, k, s, c) \right] \pi(a, b, s, k, c), \quad (8.25)$$

em que a $\pi(a, b, s, k, c)$ precisa ser especificada. No entanto, trabalharemos com transformações dos parâmetros a e b .

Considere as transformações $\eta_1 = \log(a)$ e $\eta_2 = \log(b)$. Então, com base em (8.24) teremos

$$\begin{aligned} \pi_\eta(x, \eta_1, \eta_2, s, k, c) &= f(x, e^{\eta_1}, e^{\eta_2}, s, k, c) \\ &= g(x; s, k, x)h(x, e^{\eta_1}, e^{\eta_2}, s, k, c). \end{aligned} \quad (8.26)$$

Com base em (8.26), a distribuição condicional completa a *posteriori* de $(\eta_1, \eta_2)^T$ é dada por

$$\pi(\eta_1, \eta_2|s, k, c, x) \propto \left[\prod_{i=1}^n h(x, e^{\eta_1}, e^{\eta_2}, s, k, c) \right] \pi(\eta_1, \eta_2|s, k, c, x). \quad (8.27)$$

Seja

$$(\eta_1, \eta_2|s, k, c, x) \sim N(\mu, R) \quad (8.28)$$

a ser especificada. Denote,

- $\eta = (\eta_1, \eta_2)^T$;
- $L(\eta) = \pi(\eta|s, k, c, x)$
- $l(\eta) = \log(L(\eta))$

Logo de (8.27) e de (8.23) e considerando uma amostra aleatória de tamanho n de uma distribuição BBurrXII, tem-se que:

$$L(\eta) = \prod_{i=1}^n \left[\frac{1}{B(e^{\eta_1}, e^{\eta_2})} (G(x))^{(e^{\eta_1}-1)} (1-G(x))^{(e^{\eta_2}-1)} \right] \times \\ \times \exp \left\{ -\frac{1}{2}(\eta - \mu)^T R^{-1}(\eta - \mu) \right\} \quad (8.29)$$

e

$$l(\eta) = -n \log(B(e^{\eta_1}, e^{\eta_2})) + (e^{\eta_1} - 1) \sum_{i=1}^n \log(G(x_i)) + \\ + (e^{\eta_2} - 1) \sum_{i=1}^n \log(1 - G(x_i)) - \frac{1}{2}(\eta - \mu)^T R^{-1}(\eta - \mu). \quad (8.30)$$

Mas,

$$(\eta - \mu)^T R^{-1}(\eta - \mu) = \eta^T R^{-1}\eta - 2\eta^T R^{-1}\mu + \mu^T R^{-1}\mu \quad (8.31)$$

e

$$\log(B(e^{\eta_1}, e^{\eta_2})) = \log(\Gamma(e^{\eta_1})) + \log(\Gamma(e^{\eta_2})) - \log(\Gamma(e^{\eta_1} + e^{\eta_2})). \quad (8.32)$$

Então

$$l(\eta) = -n \log(e^{\eta_1}) - n \log(e^{\eta_2}) + n \log(e^{\eta_1} + e^{\eta_2}) + e^{\eta_1} \sum_{i=1}^n \log(G(x_i)) + \\ + e^{\eta_2} \sum_{i=1}^n \log(1 - G(x_i)) - \frac{1}{2}(\eta^T R^{-1}\eta - 2\eta^T R^{-1}\mu). \quad (8.33)$$

Sejam as derivadas

$$\partial l(\eta) = \frac{\partial l(\eta)}{\partial \eta} \quad e \quad D^2 l(\eta) = \frac{\partial^2 l(\eta)}{\partial \eta \partial \eta^T}, \quad (8.34)$$

em que

$$\frac{dl(\eta)}{d\eta} = \left[\frac{dl(\eta)}{d\eta_1} \quad \frac{dl(\eta)}{d\eta_2} \right]^T, \quad \frac{d(\eta^T R^{-1}\eta)}{d\eta} = 2R^{-1}\eta \quad e \quad \frac{d(\eta^T R^{-1}\mu)}{d\eta} = R^{-1}\mu. \quad (8.35)$$

Também considere

$$\psi(\alpha) = \frac{d}{d\alpha} \log(\Gamma(\alpha)) \quad \text{e} \quad \psi'(\alpha) = \frac{d^2}{d\alpha^2} \log(\Gamma(\alpha)) \quad (8.36)$$

que denotam as funções digama e trigama, com $\psi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$. Logo,

$$\frac{dl(\eta)}{d\eta} = -\eta \begin{pmatrix} e^{\eta_1} [\psi(e^{\eta_1}) - \psi(e^{\eta_1} + e^{\eta_2})] \\ e^{\eta_2} [\psi(e^{\eta_1}) - \psi(e^{\eta_1} + e^{\eta_2})] \end{pmatrix} + \begin{pmatrix} e^{\eta_1} \sum_{i=1}^n \log(G(x_i)) \\ e^{\eta_2} \sum_{i=1}^n \log(1 - G(x_i)) \end{pmatrix} + R^{-1}(\mu - \eta).$$

Além disso,

$$\frac{dl^2l(\eta)}{d\eta d\eta^T} = \begin{pmatrix} \frac{dl^2l(\eta)}{d\eta_1^2} & \frac{dl^2l(\eta)}{d\eta_1 d\eta_2} \\ \frac{dl^2l(\eta)}{d\eta_1 d\eta_2} & \frac{dl^2l(\eta)}{d\eta_2^2} \end{pmatrix}$$

Considerando os termos de $\frac{dl(\eta)}{d\eta}$, tem-se:

$$\begin{aligned} a_{11} &= \frac{d}{d\eta_1} [e^{\eta_1} \psi(e^{\eta_1}) - e^{\eta_1} \psi(e^{\eta_1} + e^{\eta_2})] \\ &= e^{\eta_1} \psi(e^{\eta_1}) + e^{2\eta_1} \psi'(e^{\eta_1}) - e^{\eta_1} \psi(e^{\eta_1} + e^{\eta_2}) - e^{2\eta_1} \psi'(e^{\eta_1} + e^{\eta_2}) \\ &= e^{\eta_1} [\psi(e^{\eta_1}) - \psi(e^{\eta_1} + e^{\eta_2})] + e^{2\eta_1} [\psi'(e^{\eta_1}) - \psi'(e^{\eta_1} + e^{\eta_2})]. \end{aligned} \quad (8.37)$$

Analogamente,

$$\begin{aligned} a_{22} &= \frac{d}{d\eta_2} [e^{\eta_2} \psi(e^{\eta_2}) - e^{\eta_2} \psi(e^{\eta_1} + e^{\eta_2})] \\ &= e^{\eta_2} [\psi(e^{\eta_2}) - \psi(e^{\eta_1} + e^{\eta_2})] + e^{2\eta_2} [\psi'(e^{\eta_2}) - \psi'(e^{\eta_1} + e^{\eta_2})] \end{aligned}$$

e

$$a_{12} = a_{21} = \frac{d}{d\eta_2} [e^{\eta_1} \psi(e^{\eta_1}) - e^{\eta_1} \psi(e^{\eta_1} + e^{\eta_2})] \quad (8.38)$$

$$= -e^{\eta_1 + \eta_2} \psi'(e^{\eta_1} + e^{\eta_2}). \quad (8.39)$$

Então, temos que

$$D^2l(\eta) = -n \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} e^{\eta_1} \sum_{i=1}^n \log(G(x_i)) & 0 \\ 0 & e^{\eta_2} \sum_{i=1}^n \log(1 - G(x_i)) \end{pmatrix} - R^{-1}$$

em que os a 's são funções de η . Com isso a estimação de $\hat{\eta}^{(t)}$ é feita a partir do

método de *Newton – Raphson*:

$$\hat{\eta}_t = \hat{\eta}_{t-1} - [D^2 l(\hat{\eta}_{t-1})]^{-1} D l(\hat{\eta}_{t-1}). \quad (8.40)$$

Considerando $\mu = \hat{\eta}_{t-1}$ e $R_t = \hat{\Sigma}_{t-1}$, em que $\hat{\Sigma}_t = -[D^2(l(\hat{\eta}_{t-1}))]^{-1}$. Logo, no MCMC substitui-se a atualização de a e b por esse procedimento.

Simulação

Para simplificar um pouco nessa simulação foi assumido o valor de $s = 1$. Dessa forma, os cálculos acima foram adaptados para a estimação de todos os parâmetros, principalmente aos relacionados a distribuição Beta (a e b). Nas simulações realizadas foi observada a grande sensibilidade do método a valores iniciais, grande autocorrelação em todas as estimativas e a baixa taxa de aceitação em todos os parâmetros.

Uma limitação é que o método depende da especificação dos valores iniciais de R_t e, em algumas simulações, isso apresentou uma forte influência nos resultados. Dessa forma, optou-se por outra abordagem, no qual as estimativas não dependessem de R_t . Por isso não serão apresentados os resultados das simulações, conforme as seções anteriores.

Apêndice C:

Aproximações Gaussianas para a Distribuição Condicional Completa de η

Nesta seção, proporemos uma outra abordagem Gaussianas na construção de uma função geradora de candidatos do vetor $\eta = (\eta_1, \eta_2)^T = (\log(a), \log(b))^T$. A distribuição condicional completa a posteriori de $\eta = (\eta_1, \eta_2)^T$ é dada por

$$\pi(\eta_1, \eta_2 | s, k, c, x) \propto \left[\prod_{i=1}^n h(x_i, e^{\eta_1}, e^{\eta_2}, s, k, c) \right] \pi(\eta_1, \eta_2). \quad (8.41)$$

Sejam η_1, η_2 independentes a priori com $\eta_i \sim N(0, \sigma^2)$. Logo,

$$\pi(\eta) \propto \exp(-\eta\eta^T/2\sigma^2), \quad (8.42)$$

onde $\eta = (\eta_1, \eta_2)^T$. Logo, a condicional completa conjunta de η é dada por:

$$\begin{aligned} \pi(\eta_1, \eta_2 | s, k, c, x) &= \left\{ \prod_{i=1}^n \left[\frac{1}{B(\eta_1, \eta_2)} [G(x_i)]^{e^{\eta_1}-1} [1 - G(x_i)]^{e^{\eta_2}-1} \right] \right\} \exp \left\{ -\frac{1}{2\sigma^2} \eta^T \eta \right\} \\ &\propto L(\eta) \pi(\eta). \end{aligned} \quad (8.43)$$

Para obter uma função geradora de candidatos de η , vamos construir uma aproximação Normal para a distribuição a posteriori de η . Nessa aproximação pretendida faremos a média da distribuição igual à moda da distribuição condicional

completa de η ($\hat{\eta}$) e a matriz de covariância dada por menos o inverso da matriz hessiana com respeito a η . Logo,

$$(\eta^* | \eta^{(k)}) \sim N_2(\hat{\eta}, I^{-1}(\hat{\eta})), \quad (8.44)$$

representa a distribuição geradora de candidatos de η condicional ao valor atual de η (denotado por $\eta^{(k)}$) na cadeia MCMC, em que

- $\hat{\eta}$ é a moda a posteriori de η , e

-

$$I(\eta) = -\frac{d^2}{d\eta d\eta^T} \log(L(\eta)\pi(\eta)).$$

Seja

$$l(\eta) = \log(\pi(\eta|s, k, c, x)) \propto \log(L(\eta)) + \log(\pi(\eta)).$$

Então,

$$l(\eta) = -n \log(\Gamma(e^{\eta_1})) - n \log(\Gamma(e^{\eta_2})) + n \log(\Gamma(e^{\eta_1} + e^{\eta_2})) + e^{\eta_1} \sum_{i=1}^n \log(G(x_i)) + e^{\eta_2} \sum_{i=1}^n \log(1 - G(x_i)) - \frac{1}{2\sigma^2} (\eta_1^2 + \eta_2^2), \quad (8.45)$$

e a sua derivada é dada por

$$\begin{aligned} \frac{dl(\eta)}{d\eta} &= -n \begin{pmatrix} e^{\eta_1} [\psi(e^{\eta_1}) - \psi(e^{\eta_1} + e^{\eta_2})] + \\ e^{\eta_2} [\psi(e^{\eta_2}) - \psi(e^{\eta_1} + e^{\eta_2})] \end{pmatrix} \\ &+ \begin{pmatrix} e^{\eta_1} \sum_{i=1}^n \log(G(x_i)) \\ e^{\eta_2} \sum_{i=1}^n \log(1 - G(x_i)) \end{pmatrix} - \frac{1}{\sigma^2} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, \end{aligned} \quad (8.46)$$

com $\psi(\alpha) = \frac{d}{d\alpha} \ln(\Gamma(\alpha))$ e $\psi'(\alpha) = \frac{d^2}{d\alpha^2} \ln\Gamma(\alpha)$ e

$$\begin{aligned} \frac{\partial^2 l(\eta)}{\partial \eta \eta^T} &= \begin{bmatrix} \frac{\partial^2 l(\eta)}{\partial \eta_1^2} & \frac{\partial^2 l(\eta)}{\partial \eta_1 \partial \eta_2} \\ \frac{\partial^2 l(\eta)}{\partial \eta_1 \partial \eta_2} & \frac{\partial^2 l(\eta)}{\partial \eta_2^2} \end{bmatrix} = -n \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\ &+ \begin{bmatrix} e^{\eta_1} \sum_{i=1}^n \log(G(x_i)) & 0 \\ 0 & e^{\eta_2} \sum_{i=1}^n \log(1 - G(x_i)) \end{bmatrix} + \frac{1}{\sigma^2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned} \quad (8.47)$$

onde

$$a_{11} = e^{\eta_1} [\psi(e^{\eta_1}) - \psi(e^{\eta_1} + e^{\eta_2})] + e^{2\eta_1} [\psi'(e^{\eta_1}) - \psi'(e^{\eta_1} + e^{\eta_2})], \quad (8.48)$$

$$a_{22} = e^{\eta_2} [\psi(e^{\eta_2}) - \psi(e^{\eta_1} + e^{\eta_2})] + e^{2\eta_2} [\psi'(e^{\eta_2}) - \psi'(e^{\eta_1} + e^{\eta_2})] \quad (8.49)$$

e

$$a_{12} = a_{21} = -e^{\eta_1 + \eta_2} \psi(e^{\eta_1} + e^{\eta_2}). \quad (8.50)$$

Logo,

$$I(\eta) = -\frac{d^2 l(\eta)}{d\eta d\eta^T}. \quad (8.51)$$

Dessa forma, para obter uma estimativa da moda a *posteriori* de η , ($\hat{\eta}$), pode-se recorrer à função *nminb* ou *optim* do *R*.

- Seja $\hat{\eta}_0$ a estimativa de η obtida com a maximização de (8.45);
- Obtenha

$$I(\hat{\eta}_0) = -\left. \frac{d^2 l(\eta)}{d\eta d\eta^T} \right|_{\eta=\hat{\eta}_0}.$$

Amostrando a partir da Função Geradora de Candidatos Proposta

Para selecionar amostras a partir da distribuição geradora de candidatos

$$(\eta^* | \eta^{(k)}) \sim_2 (\hat{\eta}, I^{-1}(\hat{\eta})),$$

considere $\hat{\eta}_0$ uma estimativa de η . A decomposição de Choleski de $I^{-1}(\hat{\eta}_0)$ é dada por

$$I^{-1}(\hat{\eta}_0) = B^{-1}(B^{-1})^T. \quad (8.52)$$

Um vetor candidato é gerado calculando-se

$$\eta^* = B^{-1}Z^* + \hat{\eta}_0 \quad (8.53)$$

onde Z^* é um vetor de variáveis aleatórias $N(0, 1)$ independentes.

Resultados

Nesta simulação, os dados foram gerados com $s = 1$ fixo e conhecido, $a = 16$, $b = 8$, $k = 5$ e $c = 6$ numa amostra de tamanho 300. Primeiramente, foi verificado que os valores iniciais obtidos pelo comando *MultStart* do pacote *BB* do software *R* não eram razoáveis, pois não indicavam valores consistentes com os verdadeiros. Dessa forma, propôs-se uma alternativa para a obtenção dos valores iniciais que leva em consideração a média de uma amostra de valores iniciais, que conduzem a resultados convergentes em estimativas de máxima verossimilhança.

De posse da média desses pontos iniciais, foi estimada a curva de densidade com esses valores (Figura 8.2), com objetivo de verificar se há alguma coerência entre esta e a curva obtida com os parâmetros verdadeiros. Verificou-se que os valores são razoáveis, pois a moda da curva estimada com a média dos valores iniciais não é excessivamente distante da moda observada no conjunto de dados.

Verificou-se que outro ponto relevante é a especificação do parâmetro σ da distribuição *a priori* conjunta (vide equação 8.42), pois exerce influência na estimativa da moda *a posteriori*. Optou-se pela especificação de uma *priori* vaga, fixando-se $\sigma = 300$.

Os resultados das estimativas marginais *a posteriori* são apresentados na Tabela 8.2. Verifica-se que esses resultados apresentam severos vícios relativamente aos valores verdadeiros. Além disso, as taxas de aceitação são muito altas.

Adicionalmente, observa-se que as autocorrelações das estimativas são muito altas. Outro ponto importante é que as estimativas de a e b foram parecidas em todas as simulações realizadas, mesmo nos casos em que os pontos iniciais foram bem diferentes. Ou seja, observou-se uma espécie de comportamento “geminado” nas estimativas de a e b .

Vale ressaltar que, mesmo com as estimativas dos parâmetros a e b bem diferentes

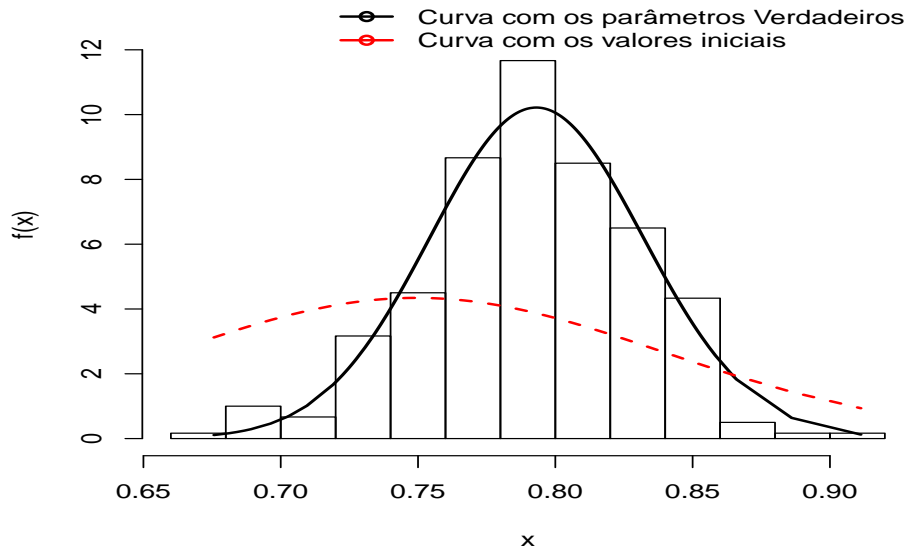


Figura 8.2: Histograma dos dados e a curva estimada com os valores iniciais (curva pontilhada)

Tabela 8.2: Estimativas marginais a posteriori da BBurrXII

	Valor Real	Média	Desvio Padrão	2.5%	50%	97.5%	Taxa de Aceitação(%)
a	16	5.072	1.872	2.819	4.625	9.804	82.169
b	8	5.074	1.870	2.814	4.616	9.797	82.272
k	5	8.975	3.188	4.268	8.437	15.531	70.802
c	6	10.561	1.656	7.429	10.577	13.364	56.318

dos verdadeiros, a curva ajustada para essas estimativas não é ruim, conforme Figura 8.3, pois quase está sobreposta a curva obtida com os verdadeiros valores.

A Figura 8.4 apresenta o traço, a densidade estimada e a autocorrelação obtida no algoritmo de MCMC. Verifica-se que no gráfico do traço das estimativas marginais não observa-se uma estabilização da cadeia construída. Ainda que se aumente o número de simulações MCMC, a estabilização do traço não ocorre.

Notoriamente, constata-se problemas de autocorrelação em todas as estimativas. Algumas densidades empíricas das estimativas marginais apresentam várias modas, indicando problemas de convergência nas cadeias MCMC.

No intuito de resolver o problema das estimativas “geminadas” de a e b , na próxima implementação utilizaremos uma nova parametrização para a distribuição beta.

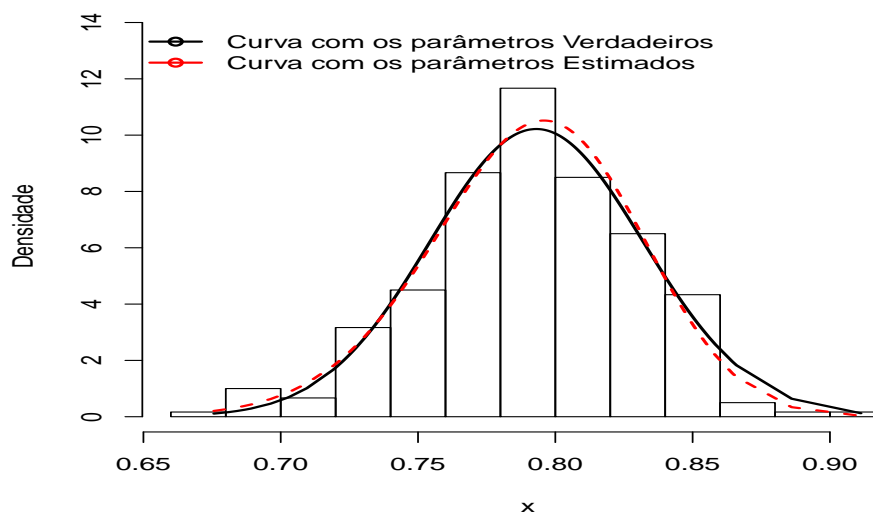


Figura 8.3: Histograma dos dados e a curva estimada com os parâmetros verdadeiros e estimados

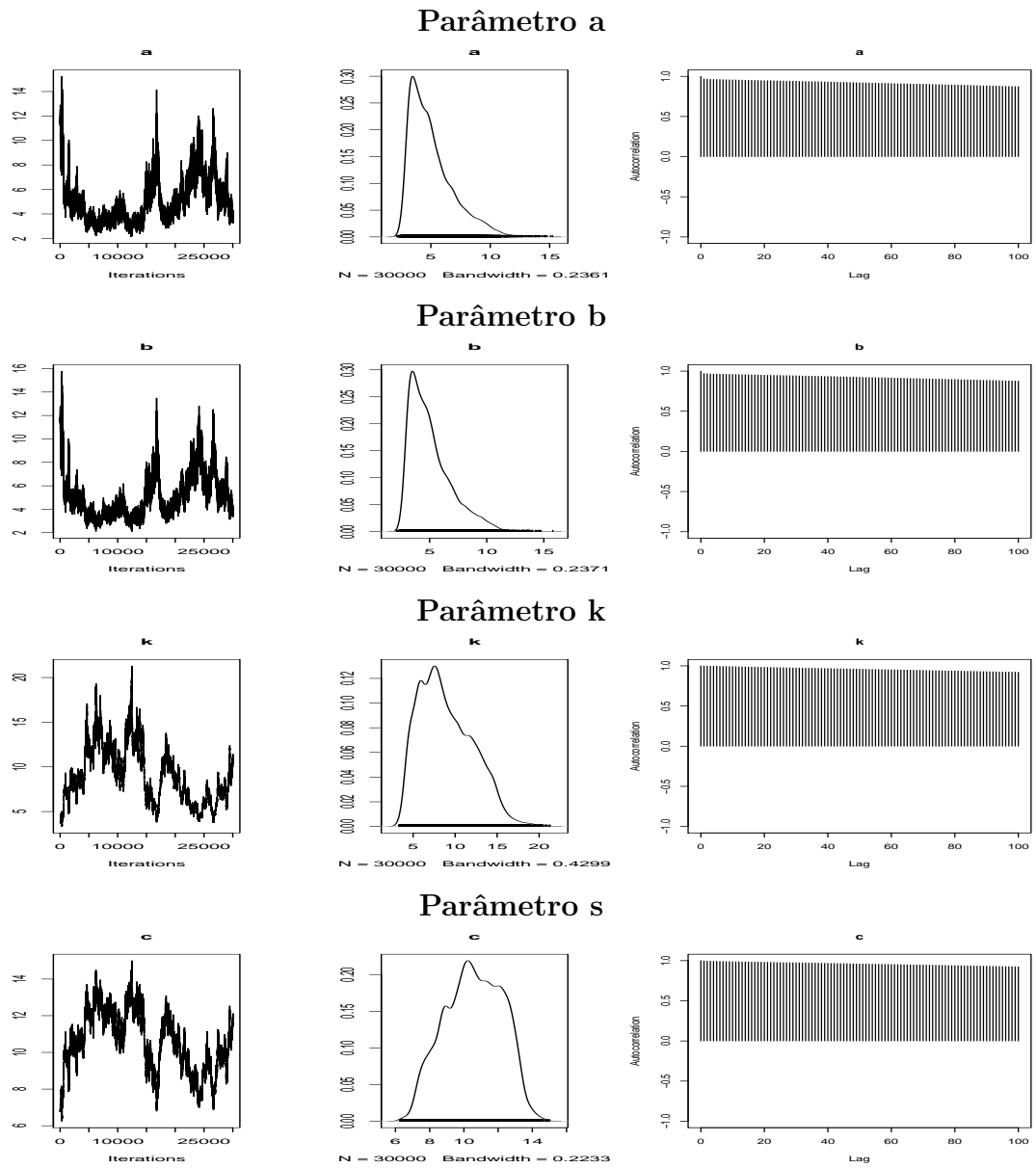


Figura 8.4: Traço, Densidade, Autocorrelação e Boxplot's das amostras MCMC para os parâmetros a , b , k e c da distribuição BBXII

Apêndice D

Cálculos dos valores esperados no Apêndice A

$$E\left(\log\left(\frac{G(x)}{1-G(x)}\right)\right) = \frac{1}{B(a,b)} \int_0^\infty \log\left(\frac{G(x)}{1-G(x)}\right) G(x)^{a-1} \times \\ \times (1-G(x))^{b-1} g(x) dx \quad (8.54)$$

fazendo $u = G(x)$ e $du = g(x)dx$

$$E\left(\log\left(\frac{G(x)}{1-G(x)}\right)\right) = \frac{1}{B(a,b)} \int_0^1 \log\left(\frac{u}{1-u}\right) u^{a-1} (1-u)^{b-1} du \\ = \frac{1}{B(a,b)} \left[\int_0^1 \log(u) u^{a-1} (1-u)^{b-1} du - \right. \\ \left. - \int_0^1 \log(1-u) u^{a-1} (1-u)^{b-1} du \right] \quad (8.55)$$

fazendo $-\log(u) = z$, $u = e^{-z}$ logo $du = -e^{-z}$ e $-\log(1-u) = z$, $u = 1 - e^{-z}$ e $du = e^{-z} dz$, então

$$\begin{aligned}
E \left(\log \left(\frac{G(x)}{1-G(x)} \right) \right) &= \\
&= \frac{1}{B(a,b)} \left(- \int_0^\infty z e^{-(a-1)z} (1-e^{-z})^{b-1} e^{-z} dz + \right. \\
&\quad \left. + \int_{i=0}^\infty z (1-e^{-z})^{a-1} e^{-(b-1)z} e^{-z} dz \right) \\
&= \frac{1}{B(a,b)} \left(- \sum_{j=0}^\infty (-1)^j C_j^{b-1} \int_0^\infty z e^{-(a+j)z} dz + \right. \\
&\quad \left. + \sum_{j=0}^\infty (-1)^j C_j^{a-1} \int_0^\infty z e^{-(b+j)z} dz \right) \\
&= \frac{1}{B(a,b)} \left(- \sum_{j=0}^\infty (-1)^j C_j^{b-1} \frac{1}{(a+j)^2} + \sum_{j=0}^\infty (-1)^j C_j^{a-1} \frac{1}{(b+j)^2} \right) \quad (8.56)
\end{aligned}$$

O Segundo momento é dado por:

$$\begin{aligned}
E \left(\left[\log \left(\frac{G(x)}{1-G(x)} \right) \right]^2 \right) &= \\
&= \frac{1}{B(a,b)} \int_0^\infty \left(\log \left(\frac{G(x)}{1-G(x)} \right) \right)^2 G(x)^{a-1} (1-G(x))^{b-1} g(x) dx \quad (8.57) \\
&= \frac{1}{B(a,b)} \left(\int_0^\infty [\log(G(x))]^2 G(x)^{a-1} (1-G(x))^{b-1} g(x) dx + \right. \\
&\quad \left. + \int_0^\infty [\log(1-G(x))]^2 G(x)^{a-1} (1-G(x))^{b-1} g(x) dx - \right. \\
&\quad \left. - 2 \int_{i=1}^\infty [\log(G(x))] [1-\log(x)] G(x)^{a-1} (1-G(x))^{b-1} g(x) dx \right) \quad (8.58)
\end{aligned}$$

Da parcela da integral dada em (8.58) será feito a substituição $z = -\log(G(x))$, $dz = (-1/G(x))g(x)dx$, $G(x) = e^z$ e $g(x)dx = -e^z dz$, da parcela de 8.58 será feito a seguinte substituição $-\log(1-G(x)) = z$, $G(x) = 1 - e^{-z}$ e $g(x)dx = e^{-z} dz$ e da parcela (8.58) será feito $G(x) = u$ e $g(x)dx = du$, logo

$$\begin{aligned}
E \left(\left(\log \left(\frac{G(x)}{1-G(x)} \right) \right)^2 \right) &= \\
&= \frac{1}{B(a,b)} \left(\int_0^\infty z^2 e^{-(a-1)z} (1-e^{-z})^{b-1} e^{-z} dz - \right. \\
&- \int_0^\infty z^2 (1-e^{-z})^{a-1} e^{-(b-1)z} e^{-z} dz - \\
&- \int_0^1 (\log(u)) [\log(1-u)] u^{a-1} (1-u)^{b-1} du \Big) \\
E \left(\left(\log \left(\frac{G(x)}{1-G(x)} \right) \right)^2 \right) &= \\
&= \frac{1}{B(a,b)} \left[\sum_{j=0}^\infty (-1)^j C_j^{b-1} \int_0^\infty z^2 e^{-(a+j)z} dz + \right. \\
&+ \sum_{j=0}^\infty (-1)^j C_j^{a-1} \int_0^\infty z^2 e^{-(b+j)z} dz + \\
&+ \left. 2 \sum_{k=1}^\infty \frac{1}{k} \int_0^1 \log(u) u^{a+j+k-1} du \right] \tag{8.59}
\end{aligned}$$

Assumindo que Z e W possuem distribuições exponenciais com parâmetros (a+j) e (b+j), respectivamente, então

$$\begin{aligned}
\int_0^\infty z^2 e^{-(a+j)z} dz &= \text{Var}(Z) + (E(z))^2 = \frac{2}{(a+j)^2}, \\
\int_0^\infty z^2 e^{-(b+j)z} dz &= \text{Var}(W) + (E(W))^2 = \frac{2}{(b+j)^2},
\end{aligned}$$

logo

$$\begin{aligned}
E \left(\left(\log \left(\frac{G(x)}{1-G(x)} \right) \right)^2 \right) &= \frac{1}{B(a,b)} \left[\sum_{j=0}^\infty (-1)^j C_j^{b-1} \frac{2}{(a+j)^3} + \sum_{j=0}^\infty (-1)^j C_j^{a-1} \frac{2}{(b+j)^3} - \right. \\
&\left. - 2 \sum_{k=1}^\infty \sum_{j=0}^\infty \frac{(-1)^j}{k} C_j^{b-1} \int_0^\infty z e^{-(a+j+k)z} dz \right],
\end{aligned}$$

$$E \left(\left(\log \left(\frac{G(x)}{1-G(x)} \right) \right)^2 \right) = \frac{1}{B(a,b)} \left[\sum_{j=0}^{\infty} (-1)^j C_j^{b-1} \frac{2}{(a+j)^3} + \sum_{j=0}^{\infty} (-1)^j C_j^{a-1} \frac{2}{(b+j)^3} - \right. \\ \left. - 2 \sum_{k=1}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^j}{k} C_j^{b-1} \frac{1}{(a+j+k)^2} \right].$$