



NATHALIA LIMA CHAVES

MODELOS DE REGRESSÃO BIRNBAUM-SAUNDERS
BASEADOS NA DISTRIBUIÇÃO NORMAL
ASSIMÉTRICA CENTRADA

CAMPINAS
2015



UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Matemática, Estatística
e Computação Científica

NATHALIA LIMA CHAVES

**MODELOS DE REGRESSÃO BIRNBAUM-SAUNDERS
BASEADOS NA DISTRIBUIÇÃO NORMAL
ASSIMÉTRICA CENTRADA**

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em estatística.

Orientador: Caio Lucidius Naberezny Azevedo

Coorientador: Filidor Edilson Vilca Labra

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELA ALUNA NATHALIA LIMA CHAVES, E ORIENTADA PELO PROF. DR. CAIO LUCIDIUS NABEREZNY AZEVEDO.

Assinatura do Orientador

A handwritten signature in blue ink, written over a horizontal line. The signature is highly stylized and cursive.

Assinatura do Coorientador

A handwritten signature in blue ink, written over a horizontal line. The signature is highly stylized and cursive.

CAMPINAS
2015

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

C398m Chaves, Nathalia Lima, 1989-
Modelos de regressão Birnbaum-Saunders baseados na distribuição normal
assimétrica centrada / Nathalia Lima Chaves. – Campinas, SP : [s.n.], 2015.

Orientador: Caio Lucidius Naberezny Azevedo.
Coorientador: Filidor Edilson Vilca Labra.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de
Matemática, Estatística e Computação Científica.

1. Distribuição normal assimétrica. 2. Modelos lineares generalizados. 3.
Inferência bayesiana. 4. Inferência estatística. I. Azevedo, Caio Lucidius
Naberezny, 1979-. II. Vilca Labra, Filidor Edilson, 1964-. III. Universidade
Estadual de Campinas. Instituto de Matemática, Estatística e Computação
Científica. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Birnbaum-Saunders regression models based on skew-normal
centered distribution

Palavras-chave em inglês:

Skew-normal distribution
Generalized linear models
Bayesian inference
Statistical inference

Área de concentração: Estatística

Titulação: Mestra em Estatística

Banca examinadora:

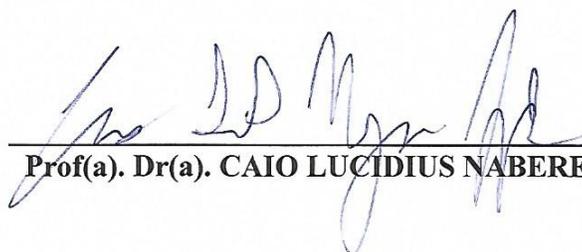
Caio Lucidius Naberezny Azevedo [Orientador]
Hildete Prisco Pinheiro
Gilberto Alvarenga Paula

Data de defesa: 27-02-2015

Programa de Pós-Graduação: Estatística

Dissertação de Mestrado defendida em 27 de fevereiro de 2015 e aprovada

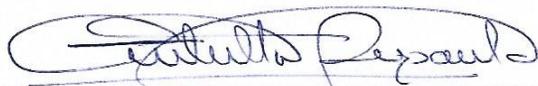
Pela Banca Examinadora composta pelos Profs. Drs.



Prof(a). Dr(a). CAIO LUCIDIUS NABEREZNY AZEVEDO



Prof(a). Dr(a). HILDETE PRISCO PINHEIRO



Prof(a). Dr(a). GILBERTO ALVARENGA PAULA

Abstract

The class of Birnbaum-Saunders (BS) models was developed from problems that arose in the field of material reliability. These problems generally are related to the study of material fatigue. However, in the last years, this class of models has been applied in areas outside that context, such as in health sciences, environmental, forestry, demographic, actuarial, financial, among others, due to its great versatility. In this work, we developed the skew-normal Birnbaum-Saunders distribution under the centered parameterization (BSNAC), which also represents an extension of the usual BS distribution and presents several advantages over the BS distribution based on the skew-normal distribution under the usual parameterization. We also developed a log-Birnbaum-Saunders linear regression model. We present several properties of both BSNAC distribution and the related regression model. We develop estimation procedures under the frequentist and Bayesian approaches, as well as diagnostic tools for the proposed models, contemplating residual analysis and measures of influence. We conducted simulation studies considering different scenarios, in order to compare the frequentist and Bayesian estimates and evaluate the performance of diagnostic measures. The methodology proposed here is illustrated with data sets from both simulation studies and real data sets.

Keywords: Birnbaum-Saunders distribution, skew-normal centered distribution, Generalized linear regression model, Bayesian inference, frequentist inference, measures of model validation.

Resumo

A classe de modelos Birnbaum-Saunders (BS) foi desenvolvida a partir de problemas que surgiram na área de confiabilidade de materiais. Tais problemas, em geral, são ligados ao estudo de fadiga de materiais. No entanto, nos últimos tempos, essa classe de modelos tem sido aplicada em áreas fora do referido contexto como, por exemplo, em ciências da saúde, ambiental, florestal, demográficas, atuariais, financeira, entre outras, devido à sua grande versatilidade. Neste trabalho desenvolvemos a distribuição Birnbaum-Saunders (BS) baseada na normal assimétrica padrão sob a parametrização centrada (BSNAC) que, além de representar uma extensão da distribuição BS usual, apresenta diversas vantagens em relação à distribuição BS baseada na distribuição normal assimétrica sob a parametrização usual. Desenvolvemos também um modelo de regressão linear log-Birnbaum-Saunders. Apresentamos, tanto para a distribuição BSNAC quanto para o respectivo modelo de regressão, diversas propriedades. Desenvolvemos procedimentos de estimação sob os enfoques frequentista e bayesiano, bem como ferramentas de diagnóstico para os modelos propostos, contemplando análise residual e medidas de influência. Realizamos estudos de simulação,

considerando diferentes cenários, com o intuito de comparar as estimativas frequentistas e bayesianas, bem como avaliar o desempenho das medidas de diagnóstico. A metodologia aqui proposta foi ilustrada tanto com dados provenientes de estudos de simulação, quanto com conjuntos de dados reais.

Palavras-chave: Distribuição Birnbaum-Saunders, Distribuição normal assimétrica centrada, Modelo de regressão linear generalizado, inferência bayesiana, inferência frequentista, medidas de validação de modelos.

Sumário

Dedicatória	xi
Agradecimentos	xiii
1 Introdução e Motivação	1
1.1 Introdução	1
1.2 Motivação	5
2 A Distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada	7
2.1 A distribuição Normal assimétrica sob a parametrização centrada	7
2.2 Definição e propriedades	10
2.3 Momentos	14
2.4 Métodos de estimação	16
2.4.1 Estimação por Máxima Verossimilhança via algoritmo ECM	19
2.4.2 Inferência bayesiana	25
2.5 Comparação entre a distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada e a Birnbaum-Saunders Normal assimétrica	29
3 Modelo de Regressão Log-Birnbaum-Saunders Normal assimétrica sob a parametrização centrada	35
3.1 A distribuição Senh-Normal	35
3.2 A distribuição SHNAC	36
3.3 O modelo de regressão log-Birnbaum-Saunders Normal assimétrica sob a parametrização centrada	41
3.4 Métodos de estimação	41
3.4.1 Estimação via algoritmo EM com maximização condicional	42
3.4.2 Inferência bayesiana	47
3.5 Análise de Diagnóstico e comparação de modelos	50
3.5.1 Análise Residual	50
3.5.2 Estatísticas para comparação de modelos	51
3.5.3 Medidas de Influência	52

4	Estudo de simulação	55
4.1	Estudo de simulação I	56
4.1.1	Estudo de Convergência	56
4.1.2	Comparação entre as estimativas frequentistas e bayesianas	64
4.2	Estudo de simulação II	75
4.2.1	Estudo de convergência	75
4.2.2	Comparação entre as estimativas frequentistas e bayesianas	83
4.2.3	Análise residual	95
4.2.4	Análise de influência	104
4.2.5	Seleção de modelos	107
5	Aplicação	109
5.1	Aplicação I	109
5.2	Aplicação II	112
5.2.1	Fadiga Biaxial (“do eixo duplo”)	112
5.2.2	Liga dos campeões da UEFA	120
6	Considerações finais	129
	Referências Bibliográficas	131
A	Construção dos envelopes simulados	137
A.1	Resíduos obtidos através das estimativas frequentistas	137
A.2	Resíduos obtidos através das estimativas bayesianas	137
B	Conjunto de dados Fadiga Biaxial	139
C	Liga dos campeões da UEFA	141

*A Deus pela dádiva da vida,
Aos meus pais,
À minha família.*

Agradecimentos

Meus sinceros agradecimentos:

A Deus por me conceder, acima de tudo, saúde, disposição, coragem para vencer todos os desafios e obstáculos encontrados pelo caminho e por iluminar constantemente as minhas decisões.

À minha mãe (por quem sinto um amor incondicional), pelo amor, carinho, pela educação, pela paciência e doçura, pela inteligência e delicadeza. Obrigada por estar tão presente na minha vida e tornar possível a realização desse sonho.

Ao meu pai (*In Memoriam*), pela educação, pela total dedicação à família e pelo exemplo de honestidade. Obrigada por estar tão presente na minha vida através dos teus valiosos ensinamentos e por caminhar ao meu lado, de uma forma ou de outra.

À minha irmã, pelo amor, carinho, por todos os bons momentos vivenciados, pelas conversas e brincadeiras. À toda minha família, pelo exemplo de união e pelo amor dedicado.

À Roberta, pelo carinho, pela compreensão e parceria, pela inteligência e sensatez compartilhadas, pelos inúmeros conselhos e aprendizados, pelos momentos bons e pelos momentos difíceis em que esteve sempre ao meu lado, por acreditar em mim e me dá forças pra seguir em frente, por ajudar a realizar muitos dos meus planos e por dividir comigo os meus sonhos. Obrigada pela tua contribuição nesse trabalho, compartilhando comigo os teus conhecimentos.

Ao meu orientador Caio Azevedo pela confiança que em mim foi depositada, pela paciência e compreensão, pelas dúvidas esclarecidas prontamente e pelo constante incentivo. E ao meu coorientador Filidor Vilca Labra, por compartilhar conosco a sua experiência e pela contribuição nesse trabalho.

Aos professores do Departamento de Estatística e Matemática Aplicada (DEMA) da Universidade Federal do Ceará. Em especial, aos professores: João Maurício e Juvêncio, por serem verdadeiros mestres e pela imensa contribuição na minha formação profissional.

A todos, que direta ou indiretamente, contribuíram com essa conquista.

À CAPES e à UNICAMP pelo apoio financeiro, fundamental para o desenvolvimento desse trabalho.

Lista de Ilustrações

1.1	Função de distribuição acumulada da Birnbaum-Saunders para diferentes valores de α e $\beta = 1$	3
1.2	Duas vezes a LVPR de γ do modelo BSNA (painel à esquerda) e do BSNAC (painel à direita).	5
1.3	Envelopes simulados (a) log-BSNAC e (b) log-BS.	6
2.1	Densidades da distribuição Normal assimétrica sob a parametrização centrada para diferentes valores de γ , mantendo fixos $\mu = 10$ e $\sigma = 5$	9
2.2	Densidades da BS para diferentes valores de α , fixando $\beta = 1$	11
2.3	Densidades da BSNAC para diferentes valores de γ , considerando $\beta = 1$, (a)–(b) $\alpha = 0, 2$ e (c)–(d) $\alpha = 0, 8$	12
2.4	Densidades da BSNAC para diferentes valores de α , considerando $\beta = 1$, (a) $\gamma = 0, 9$ e (b) $\gamma = -0, 9$	13
2.5	Densidades da BSNAC considerando (a) $\beta = 2$, (b) $\beta = 3$, (c) $\beta = 4$, (d) $\beta = 5$; os mesmos valores de α da Figura 2.4 e $\gamma = 0, 9$	14
2.6	Duas vezes a LVPR de γ	30
2.7	Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 20$	31
2.8	Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 100$	31
2.9	Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 500$	31
2.10	Duas vezes a LVPR de γ	33
2.11	Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 20$	33
2.12	Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 100$	34
2.13	Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 500$	34
3.1	Densidades da SHNAC para diferentes valores de γ , considerando $\beta = 1$ ($\rho = 0$), (a)–(b) $\alpha = 0, 2$ e (c)–(d) $\alpha = 0, 8$	37
3.2	Densidades da SHNAC para diferentes valores de α , considerando $\beta = 1$ ($\rho = 0$), (a) $\gamma = 0, 9$ e (b) $\gamma = -0, 9$	38
4.1	Gráficos de trajetórias e de autocorrelação dos parâmetros do modelo probabilístico BSNAC ($n=10$).	58
4.2	Estatística de Gelman-Rubin para o modelo probabilístico BSNAC. ($n=10$)	59
4.3	Gráficos de autocorrelação para o modelo probabilístico BSNAC. ($n=10$)	60

4.4	Gráficos de trajetórias e de autocorrelação dos parâmetros do modelo probabilístico BSNAC (n=200).	61
4.5	Estatística de Gelman-Rubin para o modelo probabilístico BSNAC. (n=200)	62
4.6	Gráficos de autocorrelação para o modelo probabilístico BSNAC. (n=200)	63
4.7	Gráficos de trajetórias e de autocorrelação dos parâmetros do modelo de regressão log-BSNAC (n=10).	77
4.8	Estatística de Gelman-Rubin para o modelo de regressão log-BSNAC. (n=10)	78
4.9	Gráficos de autocorrelação para o modelo de regressão log-BSNAC. (n=10)	79
4.10	Gráficos de trajetórias e de autocorrelação dos parâmetros do modelo de regressão log-BSNAC (n=200).	81
4.11	Estatística de Gelman-Rubin para o modelo de regressão log-BSNAC. (n=200)	82
4.12	Gráficos de autocorrelação para o modelo de regressão log-BSNAC. (n=200)	83
4.13	Análise de resíduos, obtidos através das estimativas frequentistas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BSNAC.	96
4.14	Análise de resíduos, obtidos através das estimativas bayesianas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BSNAC.	97
4.15	Análise de resíduos, obtidos através das estimativas frequentistas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS.	98
4.16	Análise de resíduos, obtidos através das estimativas bayesianas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS.	99
4.17	Análise de resíduos, obtidos através das estimativas frequentistas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS-t.	100
4.18	Análise de resíduos, obtidos através das estimativas bayesianas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS-t.	101
4.19	Análise de resíduos, obtidos através das estimativas frequentistas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS-Skew-t.	102
4.20	Análise de resíduos, obtidos através das estimativas bayesianas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS-Skew-t.	103
4.21	A divergência de K-L obtida a partir do ajuste do modelo log-BSNAC e de observações simuladas dos modelos: (a) log-BSNAC, (b) log-BS, (c) log-BS-t, (d) log-BS-t assimétrica.	105
4.22	Calibração da divergência de K-L para os quatro conjunto de dados simulados a partir dos modelos: (a) log-BSNAC, (b) log-BS, (c) log-BS-t, (d) log-BS-t assimétrica.	106
5.1	Boxplot do tempo gasto (em horas) em alimentação.	110
5.2	Boxplot do tempo gasto (em horas) em alimentação com as curvas ajustadas dos modelos BSNAC e BS.	111
5.3	Gráficos de dispersão e retas ajustadas considerando os modelos - (a) modelo dado em (5.2.1) e (b) modelo dado em (5.2.3).	113
5.4	Análise de resíduos, calculados a partir das estimativas frequentistas, do modelo log-BSNAC.	115
5.5	Análise de resíduos, calculados a partir das estimativas bayesianas, do modelo log-BSNAC.	116

5.6	Análise de resíduos, calculados a partir das estimativas frequentistas, do modelo log-BS.	117
5.7	Análise de resíduos, calculados a partir das estimativas bayesianas, do modelo log-BS.	118
5.8	Divergência de K-L para os modelos - (a) log-BSNAC e (b) log-BS.	119
5.9	Calibração da divergência de K-L para os modelos - (a) log-BSNAC e (b) log-BS.	119
5.10	Análise de resíduos, calculados a partir das estimativas frequentistas, do modelo log-BSNAC.	123
5.11	Análise de resíduos, calculados a partir das estimativas bayesianas, do modelo log-BSNAC.	124
5.12	Análise de resíduos, calculados a partir das estimativas frequentistas, do modelo log-BS.	125
5.13	Análise de resíduos, calculados a partir das estimativas bayesianas, do modelo log-BS.	126
5.14	Divergência de K-L para os modelos - (a) log-BSNAC e (b) log-BS.	127
5.15	Calibração da divergência de K-L para os modelos - (a) log-BSNAC e (b) log-BS.	127

Lista de Tabelas

2.1	Distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada: média e erros-padrão das N estimativas de $\hat{\theta}$	25
2.2	Prioris associadas aos parâmetros do modelo probabilístico.	27
3.1	Prioris associadas aos parâmetros do modelo de regressão.	48
4.1	Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = -0,67$	65
4.2	Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = -0,45$	66
4.3	Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = 0$	67
4.4	Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = 0,45$	68
4.5	Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = 0,67$	69
4.6	Resultados do estudo de simulação - $\alpha = 1,5$, $\beta = 1$, $\gamma = -0,67$	70
4.7	Resultados do estudo de simulação - $\alpha = 1,5$, $\beta = 1$, $\gamma = -0,45$	71
4.8	Resultados do estudo de simulação - $\alpha = 1,5$, $\beta = 1$, $\gamma = 0$	72
4.9	Resultados do estudo de simulação - $\alpha = 1,5$, $\beta = 1$, $\gamma = 0,45$	73
4.10	Resultados do estudo de simulação - $\alpha = 1,5$, $\beta = 1$, $\gamma = 0,67$	74
4.11	Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1,2)^T$ e $\gamma = -0,67$	85
4.12	Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1,2)^T$ e $\gamma = -0,45$	86
4.13	Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1,2)^T$ e $\gamma = 0$	87
4.14	Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1,2)^T$ e $\gamma = 0,45$	88
4.15	Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1,2)^T$ e $\gamma = 0,67$	89
4.16	Resultados do estudo de simulação - $\alpha = 1,5$, $\varphi = (1,2)^T$ e $\gamma = -0,67$	90
4.17	Resultados do estudo de simulação - $\alpha = 1,5$, $\varphi = (1,2)^T$ e $\gamma = -0,45$	91
4.18	Resultados do estudo de simulação - $\alpha = 1,5$, $\varphi = (1,2)^T$ e $\gamma = 0$	92
4.19	Resultados do estudo de simulação - $\alpha = 1,5$, $\varphi = (1,2)^T$ e $\gamma = 0,45$	93
4.20	Resultados do estudo de simulação - $\alpha = 1,5$, $\varphi = (1,2)^T$ e $\gamma = 0,67$	94
4.21	Cenários referentes à análise de resíduos.	95
4.22	Médias dos critérios - Modelo verdadeiro: log-BSNAC.	107
4.23	Médias dos critérios - Modelo verdadeiro: log-BS.	107
5.1	Estatísticas descritivas do tempo gasto (em horas) em alimentação.	110
5.2	Estimativas dos parâmetros e seus respectivos erros-padrão, intervalos de confiança e de credibilidade para os modelos BSNAC e BS.	110
5.3	Crítérios para seleção de modelos.	111

5.4	Estatísticas descritivas para o número de ciclos, N , até que ocorra a falha e para o trabalho por ciclo, X	112
5.5	Estimativas dos parâmetros e seus respectivos erros-padrão, intervalos de confiança e de credibilidade para os modelos log-BSNAC e log-BS.	114
5.6	Critérios para seleção de modelos.	114
5.7	Impacto percentual nas estimativas, após a retirada das observações potencialmente influentes.	120
5.8	Estatísticas descritivas para T_1 e T_2	120
5.9	Estimativas dos parâmetros e seus respectivos erros-padrão, intervalos de confiança e de credibilidade para os modelos log-BSNAC e log-BS.	121
5.10	Critérios para seleção de modelos.	121
5.11	Impacto percentual nas estimativas, após a retirada das observações potencialmente influentes.	128
B.1	Dados de fadiga Biaxial (Brown e Miller, 1978).	139
C.1	Liga dos campeões da UEFA	141

Capítulo 1

Introdução e Motivação

1.1 Introdução

Em 1969, motivados por problemas de vibração encontrados em aviões comerciais e por problemas de falhas de materiais, Birnbaum e Saunders (1969a) propuseram uma nova família de distribuições para modelar o tempo de vida de materiais e equipamentos sujeitos a cargas dinâmicas, sobre as quais falaremos a seguir. Esta distribuição foi derivada a partir de um modelo que mostra que as falhas acontecem devido ao desenvolvimento e ao crescimento de uma rachadura dominante. A seguir, apresentaremos uma breve descrição de como este modelo foi derivado.

Considere um material que é sujeito a um padrão cíclico de tensão e força. Aqui, está implícito que, em qualquer momento, “carga” é uma função que representa o resultado da tensão lançada no material neste tempo. A falha do material ocorre devido ao desenvolvimento e ao crescimento de uma rachadura dominante dentro do material. Um ciclo é definido como m oscilações e cada aplicação da i -ésima oscilação em um ciclo resulta em uma extensão aleatória, X_i , da rachadura. A distribuição desta variável aleatória depende somente da rachadura atual causada pela oscilação de carga neste ciclo, ou seja as variáveis aleatórias X_i são independentes. A extensão da rachadura devido ao j -ésimo ciclo é dada por

$$Y_j = \sum_{i=1}^m X_i,$$

em que Y_j é uma variável aleatória com média μ e variância σ^2 , $\forall j = 1, 2, 3, \dots$, desde que as variáveis aleatórias X_i tenham médias e variâncias finitas. A extensão total da rachadura após z ciclos é dada por

$$W_z = \sum_{j=1}^z Y_j,$$

com função distribuição

$$H_z(w) = P(W_z \leq w), z = 1, 2, \dots$$

Seja C o número de ciclos até a falha, onde tal falha ocorre quando o comprimento da rachadura dominante excede um dado comprimento crítico ω . A função de distribuição da variável C é

$$\mathbb{P}(C \leq z) = \mathbb{P}\left(\sum_{j=1}^z Y_j > \omega\right) = 1 - H_z(\omega).$$

Assumindo que os Y_j 's são variáveis aleatórias independentes e identicamente distribuídas, a função de distribuição de C pode ser aproximada usando o Teorema Central do Limite, isto é,

$$\begin{aligned} \mathbb{P}(C \leq z) &= 1 - \mathbb{P}\left(\sum_{j=1}^z \frac{Y_j - \mu}{\sigma\sqrt{z}} \leq \frac{\omega - z\mu}{\sigma\sqrt{z}}\right) \\ &= 1 - \mathbb{P}\left(\sum_{j=1}^z \frac{Y_j - \mu}{\sigma\sqrt{z}} \leq \frac{\omega}{\sigma\sqrt{z}} - \frac{\mu\sqrt{z}}{\sigma}\right) \\ &\cong 1 - \Phi\left(\frac{\omega}{\sigma\sqrt{z}} - \frac{\mu\sqrt{z}}{\sigma}\right) \\ &= \Phi\left(\frac{\mu\sqrt{z}}{\sigma} - \frac{\omega}{\sigma\sqrt{z}}\right), \end{aligned}$$

em que, $\Phi(\cdot)$ representa a função distribuição acumulada da Normal padrão.

Segundo Birnbaum e Saunders (1969a), se z é substituído por uma variável aleatória não-negativa t , a variável aleatória T é a extensão contínua da variável aleatória discreta C . Então, T pode ser considerada como o tempo até a falha. Portanto, a função de distribuição pode agora ser escrita como

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

em que

$$\alpha = \frac{\sigma}{\sqrt{\omega\mu}} > 0 \quad \text{e} \quad \beta = \frac{\omega}{\mu} > 0.$$

Dizemos que T segue uma distribuição Birnbaum-Saunders (BS), com parâmetros α e β , e denotaremos esta distribuição por $T \sim \text{BS}(\alpha, \beta)$.

Uma propriedade importante na construção desta distribuição é que

$$Z = \frac{1}{\alpha}\left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}}\right) \sim \text{N}(0, 1). \quad (1.1.1)$$

Assim, a distribuição Birnbaum-Saunders está estreitamente relacionada com a distribuição Normal através da relação

$$T = \beta \left[\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1} \right]^2, \quad (1.1.2)$$

em que $Z \sim N(0, 1)$.

A distribuição (1.1.1) possui algumas propriedades interessantes apresentadas em Birnbaum e Saunders (1969a). O parâmetro α é um parâmetro de forma e à medida que α tende a zero, a distribuição Birnbaum-Saunders tende para a distribuição Normal com média β e variância τ , em que $\tau \rightarrow 0$, quando $\alpha \rightarrow 0$. O parâmetro β é um parâmetro de escala, isto é, $T/\beta \sim BS(\alpha, 1)$. Além disso, β é a mediana da distribuição, pois $F_T(\beta) = \Phi(0) = 1/2$.

Na Figura (1.1), apresentamos os gráficos da função de distribuição acumulada, dada em (1.1.1), para diferentes valores de α , fixando $\beta = 1$. Observamos que para diferentes escolhas dos parâmetros, as suas respectivas funções de distribuição acumulada são diferentes, isto é, a distribuição Birnbaum-Saunders é identificável, na medida que temos (α_1, β_1) e (α_2, β_2) , com $\alpha_1 \neq \alpha_2$ ou $\beta_1 \neq \beta_2$, as probabilidades correspondentes a $F(\cdot; \alpha_1, \beta_1)$ são diferentes das probabilidades correspondentes à $F(\cdot; \alpha_2, \beta_2)$.

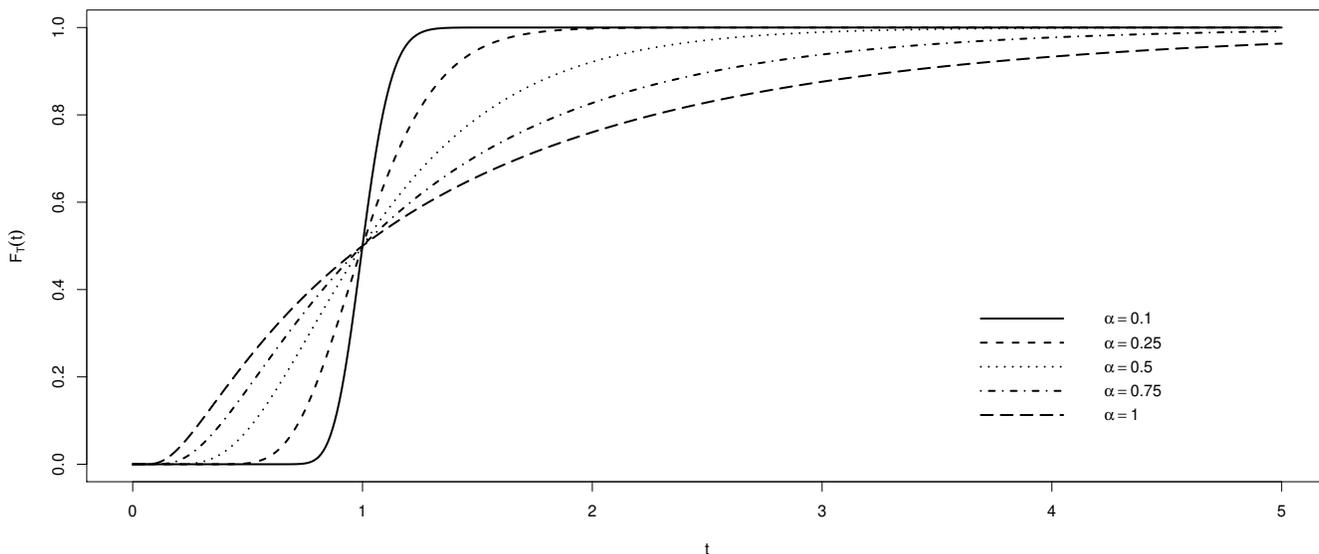


Figura 1.1: Função de distribuição acumulada da Birnbaum-Saunders para diferentes valores de α e $\beta = 1$.

A distribuição Birnbaum-Saunders vem sendo amplamente usada em diversas áreas do conhecimento. Consequentemente, há vários estudos que tem como base esta distribuição. A seguir, listamos alguns trabalhos relevantes. Birnbaum e Saunders (1969b) obtiveram originalmente os estimadores de máxima verossimilhança para α e β . Mann et al. (1974) mostraram que a distribuição (1.1.1) é unimodal. Engelhardt et al. (1981) propuseram intervalos de confiança e testes de hipóteses para α considerando β como um parâmetro de perturbação desconhecido e, da mesma forma, para β considerando α como um parâmetro de perturbação desconhecido. Desmond (1985) derivou (1.1.1) tendo por base um modelo biológico. O autor também estendeu a justificativa

física para o uso desta distribuição, relaxando algumas suposições feitas em Birnbaum e Saunders (1969a). Desmond (1986) investigou a relação entre a distribuição (1.1.1) e a distribuição gaussiana inversa. Rieck e Nedelman (1991) formularam e desenvolveram um modelo log-linear para a distribuição $BS(\alpha, \beta)$ e mostraram que este modelo pode ser usado para testes de vida acelerada ou para comparar vidas medianas de algumas populações. Achar (1993) desenvolveu procedimentos de estimação bayesiana para os parâmetros de (1.1.1) usando aproximações para as distribuições à posteriori marginais de α e β . Lu e Chang (1997) utilizaram métodos bootstrap para construir intervalos de predição para observações futuras de uma distribuição $BS(\alpha; \beta)$. Dupuis e Mills (1998) utilizaram métodos robustos de estimação para estimar os parâmetros da distribuição $BS(\alpha; \beta)$ na presença de outliers nos dados. Eles mostraram que o procedimento robusto é uma técnica alternativa poderosa quando os dados incluem observações discrepantes. Rieck (1999) derivou uma função geradora de momentos para a distribuição Senh-Normal(seno hiperbólico Normal) que pode ser usada para se obter momentos de ordem inteira ou fracionária da distribuição $BS(\alpha; \beta)$.

Algumas generalizações e aplicações da distribuição Birnbaum-Saunders foram propostas com o intuito de obter modelos mais robustos e mais flexíveis para serem utilizados em problemas reais. Por exemplo, Owen e Padgett (1999) desenvolveram a distribuição com três parâmetros. Barros et al. (2007) consideraram a distribuição t-Student padrão, $t(0, 1, \nu)$, ao invés da distribuição Normal padrão, aqui o parâmetro ν permite controlar por exemplo, a curtose da distribuição, e este novo modelo representa uma extensão robusta da distribuição proposta por Birnbaum e Saunders (1969a). Baseada nesta extensão, Leiva et al. (2007a), apresentaram um estudo de inferência e diagnóstico (baseado na metodologia de Cook, 1986) que representa uma extensão de alguns resultados desenvolvidos por Galea et al. (2004). Díaz-García e Leiva (2005) generalizaram a distribuição Birnbaum-Saunders baseado em uma classe de distribuições de contornos elípticos que inclui distribuições como Cauchy, Laplace, logística, Normal e t-Student com ν graus de liberdade.

Foram propostas ainda extensões da distribuição BS baseada na família de distribuições assimétricas, especificamente sob a classe das distribuições skew-elíptica. Alguns resultados podem ser encontrados em Vilca-Labra e Leiva (2006), e Leiva et al. (2007b, 2008) e Santana (2009). Nesses trabalhos foram obtidos resultados teóricos, estendendo as propriedades das distribuições BS e log-BS. Além das extensões já mencionadas, Leiva et al. (2009) apresentam uma revisão da distribuição BS.

Os principais objetivos deste trabalho consistem em: desenvolver e apresentar a distribuição Birnbaum-Saunders baseada na distribuição Normal assimétrica centrada, bem como um modelo de regressão linear log-Birnbaum-Saunders, destacando suas respectivas vantagens e propriedades. Desenvolver procedimentos de estimação sob os enfoques frequentista e bayesiano, bem como ferramentas de diagnóstico para os modelos propostos, contemplando análise residual e medidas de influência. Realizar estudos de simulação, considerando diferentes cenários, com o intuito de comparar as estimativas frequentistas e bayesianas, bem como avaliar o desempenho das medidas de diagnóstico. E por fim, ilustrar a metodologia proposta usando tanto dados provenientes de estudos de simulação, quanto conjuntos de dados reais.

1.2 Motivação

Os modelos Birnbaum-Saunders Normal assimétrica sob a parametrização centrada (BSNAC) e log-Birnbaum-Saunders Normal assimétrica sob a parametrização centrada (log-BSNAC) são definidos em função da distribuição Normal assimétrica sob a parametrização centrada (NAC), como uma proposta alternativa ao modelo original proposto por Birnbaum-Saunders (1969a) e ao modelo proposto por Santana (2009), que baseia-se na distribuição Normal assimétrica sob a parametrização usual.

Inicialmente, ilustraremos uma das principais vantagens da parametrização centrada frente à parametrização usual. A parametrização centrada além de remover a singularidade da matriz informação de Fisher esperada, que ocorre quando o parâmetro de assimetria $\gamma \rightarrow 0$, contorna o problema referente à existência de um ponto de inflexão na verossimilhança perfilada de γ . Detalhes a respeito da distribuição NAC serão apresentados na Seção 2.1.

A distribuição Birnbaum-Saunders Normal assimétrica (BSNA) (Santana, 2009) herda do modelo Normal assimétrico problemas na estimação por máxima verossimilhança e a utilização da parametrização centrada é uma forma de contornar tal dificuldade. Na Figura 1.2 é possível visualizar a log-verossimilhança perfilada relativa (LVPR) de γ para os modelos BSNA (painel à esquerda) e BSNAC (painel à direita). Os gráficos foram construídos considerando uma amostra aleatória das distribuições BSNA e BSNAC, respectivamente, fixando em ambos os casos, $\alpha = 0,5$, $\beta = 1$ e $\gamma = 0,67$.

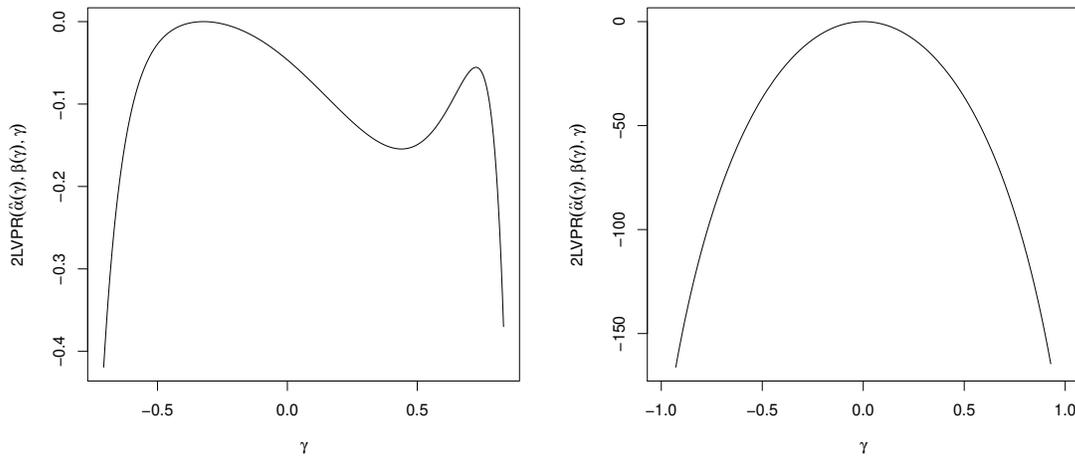


Figura 1.2: Duas vezes a LVPR de γ do modelo BSNA (painel à esquerda) e do BSNAC (painel à direita).

Nota-se uma forma não quadrática da log-verossimilhança do modelo BSNA gerada pela existência de um ponto de inflexão quando γ está bem próximo de zero, dificultando assim a obtenção dos estimadores de máxima verossimilhança. Já a log-verossimilhança do modelo BSNAC apresenta um comportamento mais regular e mais próximo de uma função côncava. Além disso, não

há ponto de inflexão em $\gamma = 0$. Uma comparação mais detalhada entre os modelos em questão será apresentada no Capítulo 2.

Para ilustrar a vantagem do modelo de regressão log-BSNAC sobre o modelo original log-BS, quando desejamos modelar dados assimétricos positivos ou dados assimétricos negativos, vamos considerar brevemente o conjunto de dados analisado por Meintanis (2007), que diz respeito à partidas de futebol da Liga dos Campeões da UEFA (Union of European Football Associations). As variáveis envolvidas nesse estudo, apresentam um comportamento fortemente assimétrico positivo. Ajustamos nesse caso, os modelos log-BS e log-BSNAC e obtivemos os envelopes simulados, conforme descrito no Apêndice A.1, Figura 1.3, para ambos os ajustes.

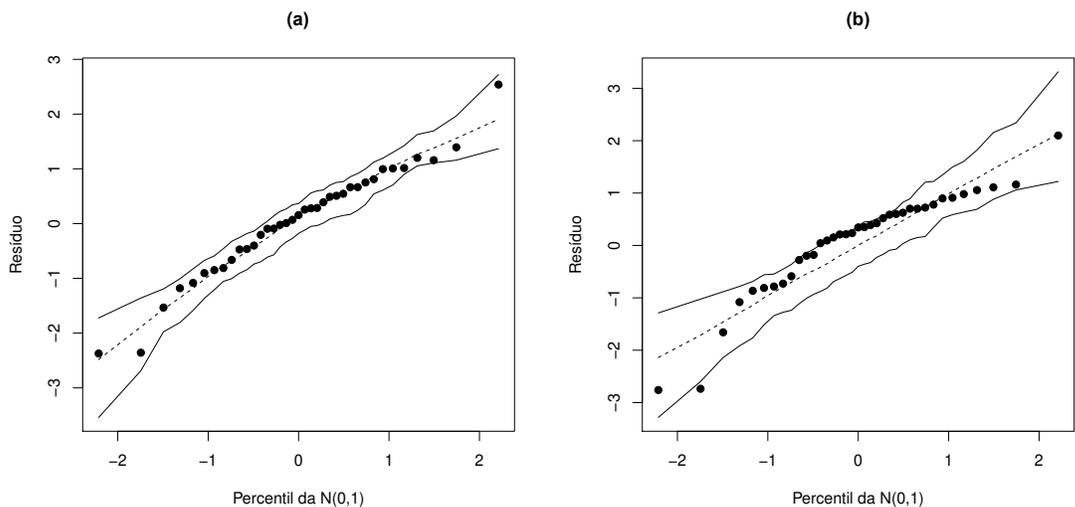


Figura 1.3: Envelopes simulados (a) log-BSNAC e (b) log-BS.

Na Figura 1.3 é possível notar que o modelo log-BSNAC ajusta-se melhor aos dados do que o modelo log-BS, visto que estamos diante de um conjunto de observações com comportamento assimétrico. No Capítulo 5, exploraremos com mais detalhes o conjunto de dados aqui mencionado, apresentando uma análise residual completa, análise de influência e critérios para comparação de modelos.

Capítulo 2

A Distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada

Neste capítulo será apresentado um novo modelo probabilístico para modelar dados positivos simétricos, assimétricos positivos e assimétricos negativos, incluindo tempos de falha. O modelo em questão é definido em função da distribuição Normal assimétrica padrão sob a parametrização centrada (NAC) como uma proposta alternativa ao modelo original proposto por Birnbaum-Saunders (1969a) e ao modelo proposto por Santana (2009), destacando suas propriedades e principais vantagens, e desenvolvendo procedimentos de estimação sob os enfoques frequentista e bayesiano.

2.1 A distribuição Normal assimétrica sob a parametrização centrada

A distribuição Normal assimétrica (NA) pertence a uma subclasse de distribuições pertencente à classe das distribuições elípticas, veja Azzalini (1985), Branco e Valle (2004) e Pewsey (2000), para mais detalhes. Esta distribuição tem sido utilizada na modelagem de dados assimétricos em várias áreas, incluindo a área psicométrica, como em Bazán et.al (2006) e Azevedo et.al (2012). Visando facilitar aspectos inferenciais, algumas representações estocásticas foram propostas na literatura. Duas representações bem conhecidas são as de Henze (1986) e Sahu et.al (2003). Azevedo et.al (2012) mostram que essas densidades são exatamente as mesmas sob a parametrização centrada. Dessa forma, os desenvolvimentos aqui apresentados se aplicam à qualquer uma dessas representações. Seguiremos utilizando a de Henze.

Uma variável aleatória Y é dita ter distribuição Normal assimétrica com parâmetro de localização $\alpha \in \mathbb{R}$, parâmetro de escala $\beta \in \mathbb{R}^+$ e parâmetro de assimetria $\lambda \in \mathbb{R}$ (notação: $NA(\alpha, \beta, \lambda)$), se sua f.d.p é dada por:

$$f(y; \alpha, \beta, \lambda) = 2\beta^{-1} \phi\left(\frac{y - \alpha}{\beta}\right) \Phi\left(\lambda \frac{y - \alpha}{\beta}\right) \mathbf{1}_{(-\infty, \infty)}(y) \quad , \quad (2.1.1)$$

em que ϕ e Φ denotam, respectivamente, as funções densidade e de distribuição acumulada da Normal padrão. A esperança e variância de Y , respectivamente, são dadas por:

$$\mathbb{E}(Y) = \alpha + \beta\delta r \quad \text{e} \quad \text{Var}(Y) = \beta^2(1 - r^2\delta^2),$$

em que $r = \sqrt{\frac{2}{\pi}}$ e $\delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}$, que assume valores no intervalo $(-1, 1)$ e pode ser usado como uma parametrização alternativa com interpretação similar a de λ .

Uma representação estocástica de Y , que facilita o desenvolvimento e a implementação computacional de algoritmos de estimação dos parâmetros, é dada por

$$Y \stackrel{d}{=} \alpha + \beta[\delta |X_0| + (1 - \delta^2)^{1/2} X_1], \quad (2.1.2)$$

em que, X_0 e X_1 são $\overset{iid}{\sim} N(0, 1)$.

Segundo Azzalini (1985), a parametrização centrada é definida a partir da seguinte identidade:

$$Y_c = \alpha + \beta Z = \mu + \sigma Z_0, \quad (2.1.3)$$

em que $Z \sim \text{NA}(0, 1, \lambda)$, $Z_0 = \sigma_z^{-1}(Z - \mu_z)$ e

$$\mu_z = \mathbb{E}(Z) = r\delta \quad \text{e} \quad \sigma_z^2 = \text{Var}(Z) = 1 - \mu_z^2. \quad (2.1.4)$$

A parametrização centrada é composta pelos parâmetros μ , σ e γ , cujas expressões são as seguintes:

$$\begin{aligned} \mu &= \mathbb{E}(Y_c) = \alpha + \beta\mu_z, \\ \sigma &= \sqrt{\text{Var}(Y_c)} = \beta\sqrt{1 - \mu_z^2}, \\ \gamma &= \frac{\mathbb{E}[(Y_c - \mathbb{E}(Y_c))^3]}{[\text{Var}(Y_c)]^{3/2}} = r\delta^3 \left[\frac{4}{\pi} - 1 \right] [1 - \mu_z^2]^{-3/2}. \end{aligned} \quad (2.1.5)$$

Neste caso, μ e σ representam o valor esperado e o desvio padrão da variável aleatória Y_c . O parâmetro γ corresponde ao coeficiente de assimetria de Pearson, que assume valores no intervalo $(-0,99527; 0,99527)$. Para denotar a Normal assimétrica sob a parametrização centrada será adotada a notação $Y_c \sim \text{NAC}(\mu, \sigma, \gamma)$. Pelo método do jacobiano, a partir da equação (2.1.3), pode-se obter a densidade de Y_c , que é dada pela seguinte expressão:

$$\begin{aligned} f_{Y_c}(y) &= 2\frac{\sigma_z}{\sigma} \phi \left[\frac{\sigma_z}{\sigma} \left(y - \mu + \frac{\sigma}{\sigma_z} \mu_z \right) \right] \Phi \left(\lambda \left[\frac{\sigma_z}{\sigma} \left(y - \mu + \frac{\sigma}{\sigma_z} \right) \right] \right) \\ &= 2\omega^{-1} \phi(\omega^{-1}(y - \xi)) \Phi[\lambda(\omega^{-1}(y - \xi))], \end{aligned} \quad (2.1.6)$$

em que,

$$\begin{aligned}
 \xi &= \mu - \sigma\gamma^{1/3}s, \\
 \omega &= \sigma\sqrt{1 + \gamma^{2/3}s^2}, \\
 \lambda &= \frac{\gamma^{1/3}s}{\sqrt{r^2 + s^2\gamma^{2/3}(r^2 - 1)}} \\
 s &= \left(\frac{2}{4 - \pi}\right)^{1/3}
 \end{aligned} \tag{2.1.7}$$

Na Figura 2.1 observa-se a densidade dada em (2.1.6) com $\mu = 10$, $\sigma = 5$ e diferentes valores do coeficiente de assimetria, γ . É possível perceber que valores positivos e negativos de γ fazem com que a densidade apresente comportamentos assimétricos positivos e negativos, respectivamente. A simetria da distribuição pode ser observada quando $\gamma = 0$. A parametrização centrada preserva a assimetria da densidade dada em (2.1.1).

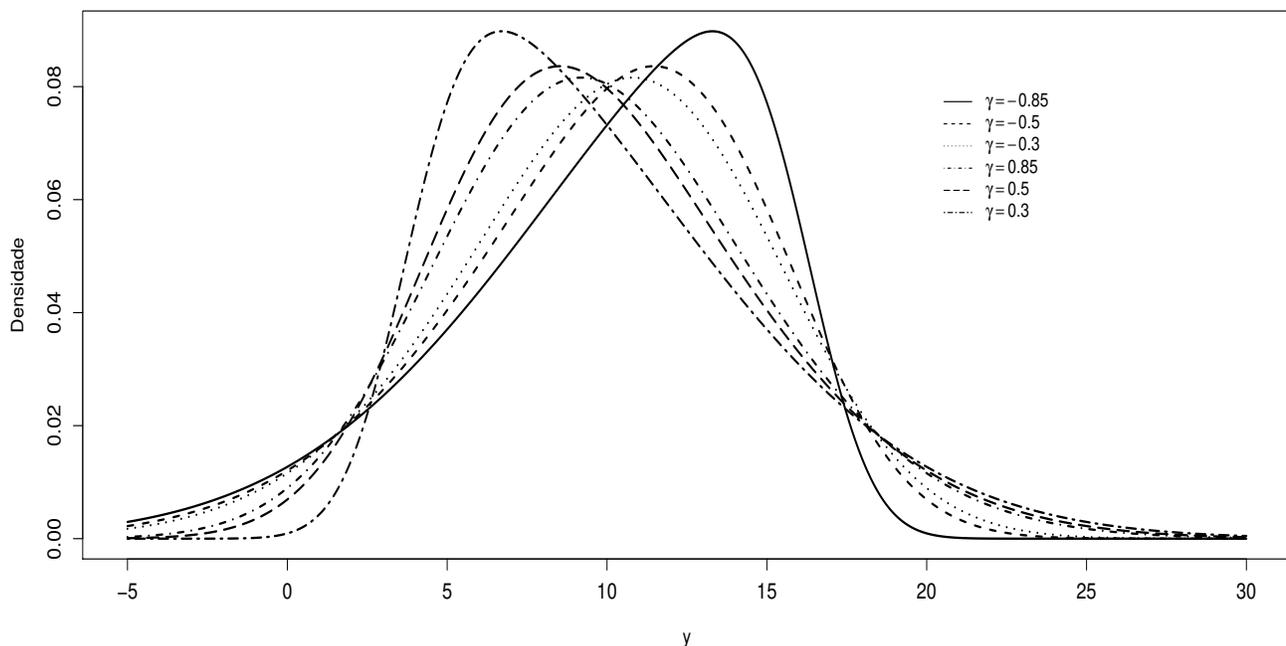


Figura 2.1: Densidades da distribuição Normal assimétrica sob a parametrização centrada para diferentes valores de γ , mantendo fixos $\mu = 10$ e $\sigma = 5$.

A representação estocástica de Y_c é obtida de forma direta a partir das equações (2.1.2) e (2.1.3), e é expressa da seguinte forma:

$$Y_c = \frac{\sigma}{\sigma_z} \left\{ \alpha + \beta \left[\delta |X_0| + (1 - \delta^2) X_1 \right] - \mu_z \right\} + \mu, \tag{2.1.8}$$

em que X_0 e X_1 são $\overset{iid}{\sim} N(0, 1)$.

2.2 Definição e propriedades

Inicialmente, apresentaremos algumas definições e propriedades referentes a distribuição Birnbaum-Saunders. Uma variável aleatória T é dita ter distribuição Birnbaum-Saunders se sua densidade é dada pela seguinte expressão

$$f_T(t; \alpha, \beta) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2\alpha^2} \left(\frac{t}{\beta} - \frac{\beta}{t} - 2 \right) \right\} \frac{t^{-3/2}(t + \beta)}{2\alpha\sqrt{\beta}}, t > 0, \quad (2.2.1)$$

em que, $\alpha, \beta > 0$. Denotaremos-na por $T \sim BS(\alpha, \beta)$. Fazendo,

$$a_t(\alpha, \beta) = \frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right] \quad \text{e} \quad A_t(\alpha, \beta) = \frac{d}{dt} a_t(\alpha, \beta) = \frac{t^{-3/2}(t + \beta)}{2\alpha\sqrt{\beta}}, \quad (2.2.2)$$

a densidade (2.2.1) pode ser escrita de forma simplificada como

$$f_T(t; \alpha, \beta) = \phi(a_t(\alpha, \beta))A_t(\alpha, \beta), \quad (2.2.3)$$

em que $\phi(\cdot)$ é a densidade da Normal padrão. Além disso, de (1.1.1), vamos ter que a acumulada $F_T(t)$ pode ser escrita como

$$F_T(t; \alpha, \beta) = \Phi(a_t(\alpha, \beta)), t > 0, \quad (2.2.4)$$

em que $\Phi(\cdot)$ representa a acumulada da Normal padrão.

A esperança, variância, coeficientes de assimetria (μ_3) e curtose (μ_4) da distribuição BS são dados, respectivamente, por

$$\mathbb{E}(T) = \beta \left(1 + \frac{\alpha^2}{2} \right), \quad \text{Var}(T) = (\alpha\beta)^2 \left(1 + \frac{5}{4}\alpha^2 \right),$$

$$\mu_3 = \frac{16\alpha^2(11\alpha^2 + 6)}{(5\alpha^2 + 4)^3} \quad \text{e} \quad \mu_4 = 3 + \frac{6\alpha^2(93\alpha^2 + 41)}{5 + \alpha^2 + 4}.$$

Além disso, se $T \sim BS(\alpha, \beta)$, então $aT \sim BS(\alpha, a\beta)$, $\forall a > 0$ e $T^{-1}BS(\alpha, \beta^{-1})$. Para mais detalhes, ver Leiva et al (2009) e Saunders (1974).

Na Figura apresentamos o gráfico da densidade (2.2.3) para alguns valores de α considerando $\beta = 1$. Observe que à medida que o valor de α diminui, a curva tende a ficar mais simétrica em torno de β , que é a mediana da distribuição. Da mesma forma, note que a variância também decresce com o aumento do valor de α .

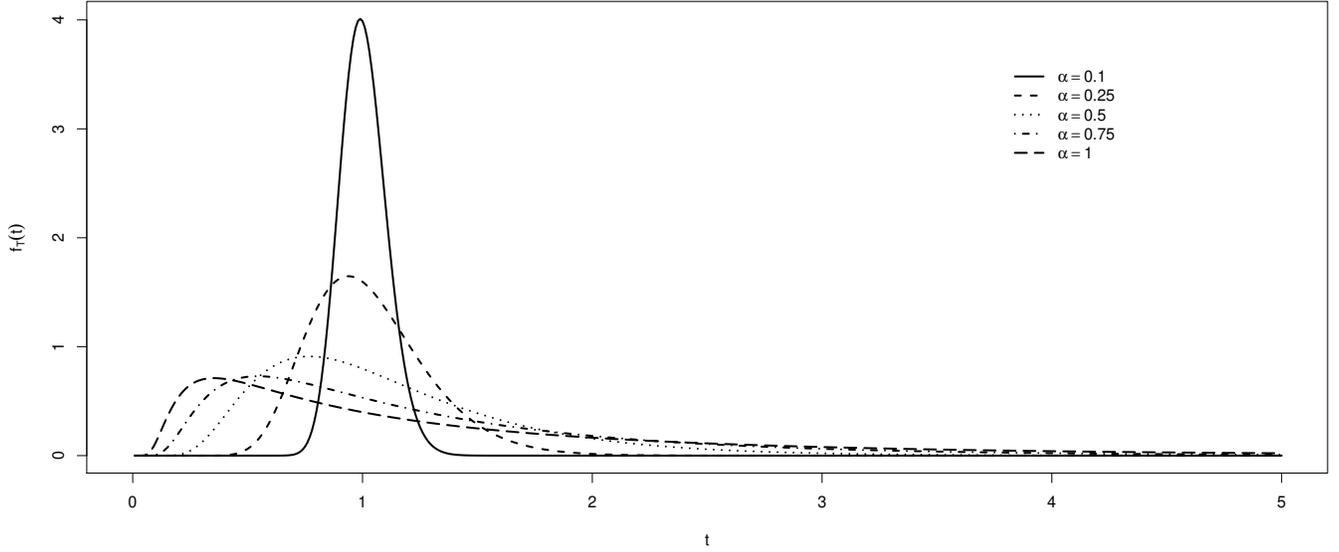


Figura 2.2: Densidades da BS para diferentes valores de α , fixando $\beta = 1$.

Definiremos, agora, a distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada. Assume-se agora, que a variável geradora da Birnbaum-Saunders ao invés de ser Normal padrão, como em (1.1.1), tem distribuição Normal assimétrica padrão sob a parametrização centrada (2.1.6), definida da seguinte forma

$$Y_c = \frac{1}{\alpha} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim \text{NAC}(0, 1, \gamma). \quad (2.2.5)$$

Assim, a distribuição da variável aleatória T é definida em termos de Y_c ,

$$T = \beta \left[\frac{\alpha Y_c}{2} + \sqrt{\left(\frac{\alpha Y_c}{2} \right)^2 + 1} \right]^2. \quad (2.2.6)$$

A variável aleatória T segue uma distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada (BSNAC) que será denotada por $T \sim \text{BSNAC}(\alpha, \beta, \gamma)$. Assim, formalmente tem-se a seguinte definição,

Definição 2.1 *Uma variável aleatória T é dita ter distribuição BSNAC com parâmetro de forma $\alpha > 0$, parâmetro de escala $\beta > 0$ e parâmetro de assimetria $\gamma \in (-0,99527; 0,99527)$, se sua f.d.p pode ser escrita como*

$$f_T(t) = 2\sigma_z \phi \left[\sigma_z \left(a_t(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \Phi \left\{ \lambda \left[\sigma_z \left(a_t(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \right\} A_t(\alpha, \beta) \mathbf{1}_{(0, \infty)}(t), \quad (2.2.7)$$

em que

$$a_t(\alpha, \beta) = \frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right)$$

$$A_t(\alpha, \beta) = \frac{d}{dt} a_t(\alpha, \beta) = \frac{t^{-3/2}(t + \beta)}{2\alpha\beta^{1/2}}, \quad (2.2.8)$$

sendo μ_z e σ_z são tais como em (2.1.4) e λ tal como em (2.1.7). Quando $\mu_z = 0$, $\sigma_z = 1$ e $\lambda = 0$, obtém-se a distribuição BS.

Nas Figuras 2.3 - 2.5, observa-se os gráficos da densidade da distribuição BSNAC dada pela equação (2.2.7) para diferentes valores de α , β e γ .

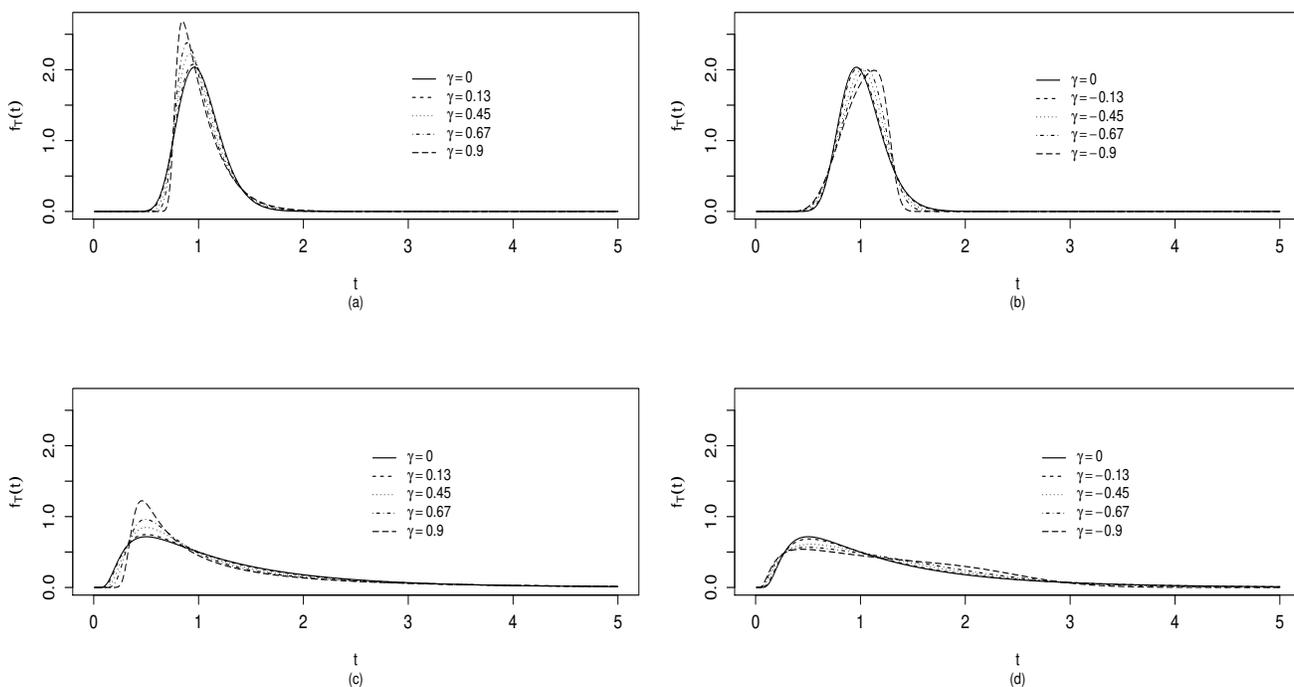


Figura 2.3: Densidades da BSNAC para diferentes valores de γ , considerando $\beta = 1$, (a)–(b) $\alpha = 0, 2$ e (c)–(d) $\alpha = 0, 8$.

Na Figura 2.3 apresentamos os gráficos da densidade (2.2.7) para diferentes valores de γ , considerando $\beta = 1$ e $\alpha = 0, 2$ (painéis (a) e (b)) e $\alpha = 0, 8$ (painéis (c) e (d)).

Os gráficos em questão sugerem que para valores positivos de γ e $\alpha = 0, 2$ (painel (a)), a densidade da BSNAC tende a se concentrar em torno de $\beta = 1$. Já para valores negativos de γ , mantendo $\alpha = 0, 2$ (painel (b)), percebe-se que a densidade permanece simétrica em torno de β , só que nesse caso, com uma variabilidade maior. Com o aumento do valor de α para 0, 8 (painéis (c) e (d)) nota-se uma mudança na forma da densidade.

Na Figura 2.4 apresentamos os gráficos da densidade (2.2.7) para diferentes valores de α , considerando $\beta = 1$ e $\gamma = 0,9$ (painel (a)) e $\gamma = -0,9$ (painel (b)). É possível perceber que a medida que o valor de α decresce, há uma diminuição da variabilidade, fazendo com que a densidade da BSNAC fique mais simétrica em torno de β .

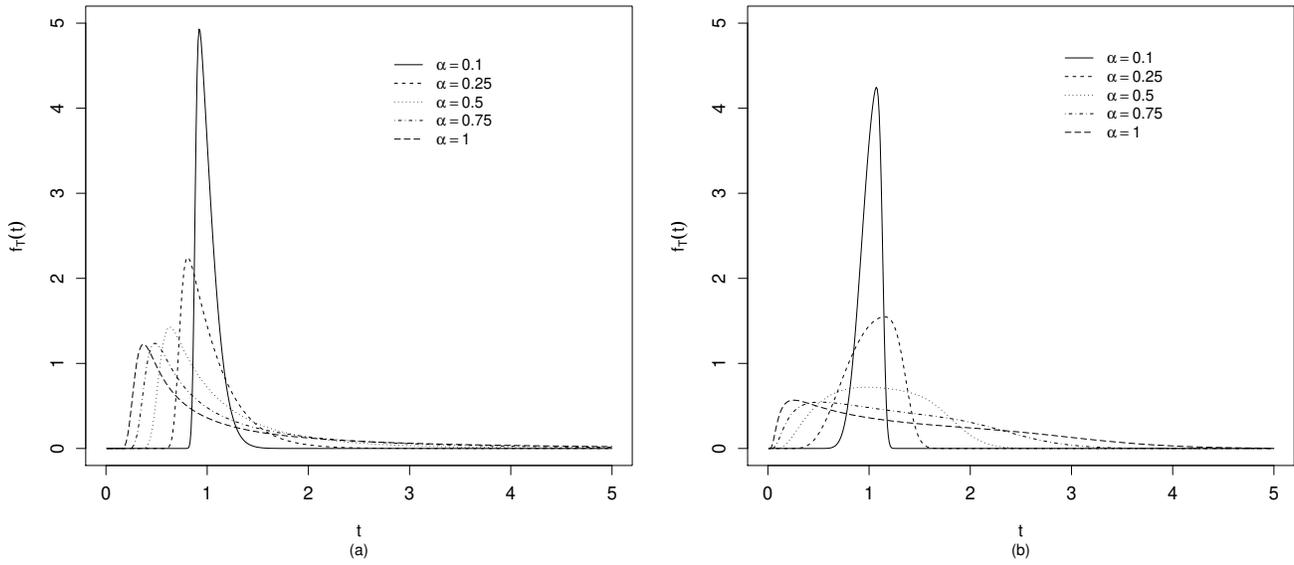


Figura 2.4: Densidades da BSNAC para diferentes valores de α , considerando $\beta = 1$, (a) $\gamma = 0,9$ e (b) $\gamma = -0,9$.

Na Figura 2.5 apresentamos os gráficos da densidade (2.2.7) para os diferentes valores de α , os mesmos usados na Figura 2.4, fixando $\gamma = 0,9$, e diferentes valores de β , $\beta = 2$ (painel (a)), $\beta = 3$ (painel (b)), $\beta = 4$ (painel (c)) e $\beta = 5$ (painel (d)). É possível perceber que quanto maior o valor de β e menor o valor fixado para α , a densidade da BSNAC apresenta um comportamento assimétrico negativo.

Em síntese, em termos dos três parâmetros da distribuição BSNAC, é possível observar a simetria da distribuição em torno de β , quando fixamos $\gamma = 0$ e pequenos valores para α . A assimetria positiva é observada à medida que α aumenta, β diminui e/ou γ assume valores positivos. Já a assimetria negativa é observada à medida que α diminui, β aumenta e/ou γ assume valores negativos.

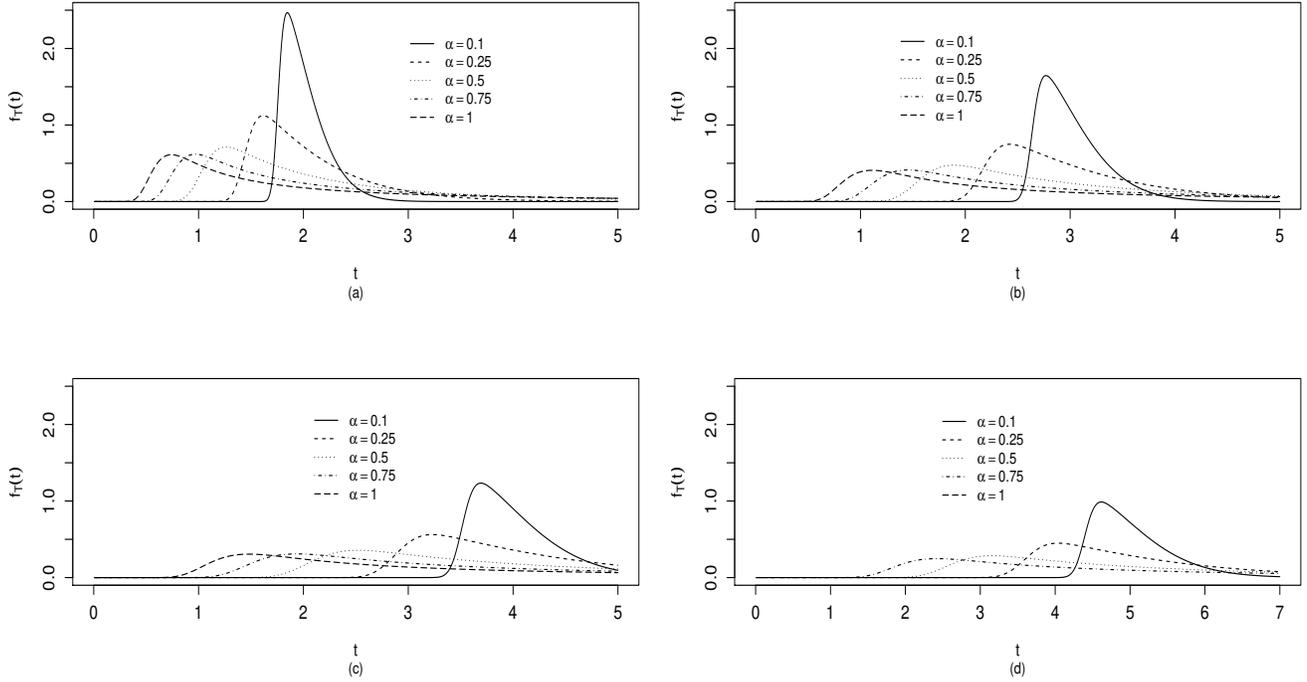


Figura 2.5: Densidades da BSNAC considerando (a) $\beta = 2$, (b) $\beta = 3$, (c) $\beta = 4$, (d) $\beta = 5$; os mesmos valores de α da Figura 2.4 e $\gamma = 0, 9$

2.3 Momentos

A fim de determinar os momentos da distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada faremos uso do Teorema Binomial (Coolidge, 1949). Os momentos de k -ésima ordem da variável aleatória T estão diretamente relacionados com os momentos de uma função da variável Y_c .

Teorema 2.1 *Seja $T \sim BSNAC(\alpha, \beta, \gamma)$ e Y_c como dado em (2.2.5). Então, $\mathbb{E}(T^n)$ existe e depende dos momentos*

$$\mathbb{E} \left\{ \left(\frac{\alpha Y_c}{2} \right)^{k+l} \left[\left(\frac{\alpha Y_c}{2} \right)^2 + 1 \right]^{\frac{k-l}{2}} \right\}, \quad (2.3.1)$$

com $k = 0, \dots, n$ e $l = 0, \dots, k$.

Demonstração:

De (2.2.6), tem-se que

$$\begin{aligned}\mathbb{E}\left(\frac{T}{\beta}\right) &= \mathbb{E}\left\{\left[\left(\frac{\alpha Y_c}{2} + \sqrt{\left(\frac{\alpha Y_c}{2}\right)^2 + 1}\right)^2\right]^n\right\} \\ &= \mathbb{E}\left\{\left[1 + \left(\frac{\alpha^2}{2} Y_c^2 + \alpha Y_c \sqrt{\left(\frac{\alpha Y_c}{2}\right)^2 + 1}\right)\right]^n\right\}.\end{aligned}$$

Usando o Teorema Binomial $(a + b)^m = \sum_{k=0}^m \binom{m}{k} a^{m-k} b^k$, tem-se que

$$\mathbb{E}\left(\frac{T}{\beta}\right)^n = \sum_{k=0}^n \binom{n}{k} \mathbb{E}\left\{\left[\frac{\alpha^2}{2} Y_c^2 + \alpha Y_c \sqrt{\left(\frac{\alpha Y_c}{2}\right)^2 + 1}\right]^k\right\}.$$

Expandindo o Binômio de Newton, obtém-se

$$\mathbb{E}\left(\frac{T}{\beta}\right)^n = \sum_{k=0}^n \binom{n}{k} \sum_{l=0}^k \binom{k}{l} 2^k \mathbb{E}\left\{\left(\frac{\alpha Y_c}{2}\right)^{k+l} \left[\left(\frac{\alpha Y_c}{2}\right)^2 + 1\right]^{\frac{k-l}{2}}\right\}. \quad (2.3.2)$$

Portanto, $\mathbb{E}\left(\frac{T}{\beta}\right)^n$ existe e depende dos momentos $\mathbb{E}\left\{\left(\frac{\alpha Y_c}{2}\right)^{k+l} \left[\left(\frac{\alpha Y_c}{2}\right)^2 + 1\right]^{\frac{k-l}{2}}\right\}$, $k = 0, \dots, n$ e $l = 0, \dots, k$, que também existe e são finitos.

De fato, seja $U_{kl} = \left(\frac{\alpha Y_c}{2}\right)^{k+l} \left[\left(\frac{\alpha Y_c}{2}\right)^2 + 1\right]^{\frac{k-l}{2}}$. Então, usando a Desigualdade de Markov, tem-se que

$$\begin{aligned}\mathbb{E}(|U_{kl}|) &\leq \mathbb{E}^{1/2} \left[\left(\frac{\alpha Y_c}{2}\right)^{2(k+l)}\right] \mathbb{E}^{1/2} \left\{\left[\left(\frac{\alpha Y_c}{2}\right)^2 + 1\right]^{k-l}\right\} \\ &< \infty,\end{aligned}$$

pois os momentos de Y_c são finitos, visto que $Z^2 \sim \mathcal{X}_1^2$ e Y_c é tal como definida em (2.1.3), e consequentemente $\mathbb{E}\left[\left(\frac{\alpha Y_c}{2}\right)^{2(k+l)}\right]$ e $\mathbb{E}\left\{\left[\left(\frac{\alpha Y_c}{2}\right)^2 + 1\right]^{k-l}\right\}$ são finitos.

Corolário 2.1 *Seja $T \sim \text{BSNAC}(\alpha, \beta, \gamma)$ e Y_c como dado em (2.2.5). Então,*

$$\mathbb{E}(T^n) = \beta^n \sum_{k=0}^n \binom{n}{k} \sum_{l=0}^k \binom{k}{l} 2^k \mathbb{E}\left\{\left(\frac{\alpha Y_c}{2}\right)^{k+l} \left[\left(\frac{\alpha Y_c}{2}\right)^2 + 1\right]^{\frac{k-l}{2}}\right\}. \quad (2.3.3)$$

A demonstração segue diretamente da equação (2.3.2) do Teorema 2.1.

Corolário 2.2 *Seja $T \sim \text{BSNAC}(\alpha, \beta, \gamma)$ e $U_k = \mathbb{E}(Y_c^k \sqrt{\alpha^2 Y_c^2 + 4})$. Então*

$$i) \mathbb{E}(T) = \frac{\beta}{2}(2 + \alpha^2 + 2\alpha U_1);$$

$$ii) \mathbb{E}(T^2) = \frac{\beta}{2} \left\{ 2\beta + 4\alpha^2\beta + \frac{\alpha^4\beta}{\sigma_z^4} [3 - \mu_z^2(6 - 4\delta^2 + 3\mu_z^2)] + 4\alpha\beta U_1 + 2\alpha^3\beta U_3 \right\};$$

$$iii) \text{Var}(T) = \frac{\beta}{2} \left\{ -8\alpha\beta U_1 + 2\alpha^2\beta(1 - U_1^2) + 2\alpha^3\beta(U_3 - U_1) - \alpha^4\beta \left[\frac{1}{\sigma_z^4} (3 - 6\mu_z^2 + 4\mu_z^2\delta^2 - 3\mu_z^4) - \frac{1}{2} \right] \right\},$$

em que $U_j = \mathbb{E} \left\{ Y_c^j \left[\left(\frac{\alpha Y_c}{2} \right)^2 + 1 \right]^{1/2} \right\}$, μ_z e σ_z são tais como em (2.1.4).

Para obter os valores de U_j , as integrais envolvidas nos cálculos desses momentos devem ser resolvidas através de métodos de integração numérica.

2.4 Métodos de estimação

Estudaremos inicialmente a obtenção dos estimadores de máxima verossimilhança, via algoritmo ECM (*Expectation conditional maximization*; Meng e Rubin, 1993). Em seguida, apresentaremos os estimadores obtidos sob a abordagem bayesiana, via algoritmo MCMC (Método de Monte Carlo via Cadeia de Markov; Gelman et al., 1997).

Apresentaremos a seguir, alguns resultados que serão de suma importância para obtenção dos estimadores de máxima verossimilhança. Em alguns momentos ao longo do trabalho mencionaremos e utilizaremos o parâmetro de assimetria δ , visto que facilita a obtenção dos resultados tanto analiticamente, quanto computacionalmente. Além disso, avaliar a magnitude da assimetria através de δ torna-se mais fácil. Vale ressaltar que o parâmetro de assimetria de interesse é o γ , cuja relação com δ é dada por (2.1.5).

Proposição 2.1 *Seja Y_c como em (2.2.5) e $T \sim BSNAC(\alpha, \beta, \delta)$. A representação estocástica de Y_c é dada por $Y_c = \frac{1}{\sigma_z} (\delta H + \sqrt{1 - \delta^2} X_1 - \mu_z)$, em que $H = |X_0|$, assim como pode ser visto em (2.1.8). Então,*

$$T|H = h \sim BSEC(\alpha_\delta, \beta, \sigma = 2, \nu_h),$$

em que $\alpha_\delta = \alpha \sqrt{\frac{1 - \delta^2}{1 - r^2 \delta^2}}$ e $\nu_h = -\frac{\delta(h-r)}{\sqrt{1 - \delta^2}}$. Aqui, usamos a sigla BSEC para representar a distribuição Birnbaum-Saunders Estendida sob a parametrização centrada. Para mais detalhes a respeito da distribuição Birnbaum-Saunders Estendida (BSE), ver Leiva et.al (2008).

A densidade condicional de T dada $H = h$ é dada por

$$f_{T|H}(t|h) = \phi(\nu_h + a_t(\alpha_\delta, \beta)) A_t(\alpha_\delta, \beta), \quad (2.4.1)$$

em que $a_t(\alpha_\delta, \beta)$ e $A_t(\alpha_\delta, \beta)$ são como em (2.2.8).

Demonstração:

Como $Y_c \sim \text{NAC}(0, 1, \gamma)$, usando a representação estocástica dada em (2.1.8) podemos expressar Y_c como

$$Y_c = \frac{1}{\sigma_z} [\delta H + \sqrt{1 - \delta^2} X_1 - \mu_z] = \frac{1}{\alpha} \left[\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right].$$

Assim a distribuição condicional de Y_c dado $H = h$ é dada por

$$Y_c | (H = h) = \frac{1}{\alpha} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \Big| (H = h) \sim \text{N}(\mu_h, \sigma^2),$$

em que $\mu_h = \frac{\delta(h - r)}{\sqrt{1 - r^2\delta^2}}$ e $\sigma^2 = \frac{1 - \delta^2}{1 - r^2\delta^2}$. Logo, condicionado à $H = h$ a distribuição de $W = (Y_c - \mu_h)/\sigma$ é dada por

$$W = -\frac{\mu_h}{\sigma} + \frac{1}{\sigma\alpha} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim \text{N}(0, 1).$$

em que, a variável aleatória T , nesse caso, pode ser escrita como:

$$T = \beta \left[\frac{\alpha}{2} (\sigma W + \mu_h) + \sqrt{\left[\frac{\alpha}{2} (\sigma W + \mu_h) \right]^2 + 1} \right].$$

Assim do resultado acima, conclui-se a prova.

Lema 2.1 *Sejam $Y \sim N(\mu_y, \sigma_y^2)$ e $X \sim N(\mu_x, \sigma_x^2)$. Então,*

$$\phi(y|\mu_y + ax, \sigma_y^2) \phi(x|\mu_x, \sigma_x^2) = \phi(y|\mu_y + a\mu_x, \sigma_y^2 + a^2\sigma_x^2) \phi(x|\mu_x + \Lambda \frac{a}{\sigma_y^2} (y - \mu_y - a\mu_x), \Lambda),$$

em que $\Lambda = \frac{\sigma_x^2 \sigma_y^2}{\sigma_y^2 + a^2 \sigma_x^2}$.

Lema 2.2 *Seja $X \sim N(\eta, \tau^2)$. Então, para qualquer constante real*

$$\begin{aligned} \mathbb{E}(X|X > a) &= \eta + \frac{\phi\left(\frac{a-\eta}{\tau}\right)}{1 - \Phi\left(\frac{a-\eta}{\tau}\right)} \tau. \\ \mathbb{E}(X^2|X > a) &= \eta^2 + \tau^2 + \frac{\phi\left(\frac{a-\eta}{\tau}\right)}{1 - \Phi\left(\frac{a-\eta}{\tau}\right)} (\eta + a) \tau. \end{aligned}$$

Proposição 2.2 *Seja Y_c como em (2.2.5) e $T \sim \text{BSNAC}(\alpha, \beta, \gamma)$. Então*

$$i) f_{H|T}(h|t) = \frac{\phi \left\{ h \left| \delta \sqrt{1 - r^2 \delta^2} \left(a_t(\alpha, \beta) + \frac{r\delta}{\sqrt{1 - r^2 \delta^2}} \right); 1 - \delta^2 \right. \right\}}{\Phi \left\{ \lambda \left[\sigma_z \left(a_t(\alpha, \beta) + \frac{r\delta}{\sqrt{1 - r^2 \delta^2}} \right) \right] \right\}} \mathbf{1}_{(0, \infty)}(h).$$

$$ii) \mathbb{E}(H|T = t) = \eta + W_{\Phi} \left(\frac{\eta}{\tau} \right) \tau,$$

$$iii) \mathbb{E}(H^2|T = t) = \eta^2 + \tau^2 + W_{\Phi} \left(\frac{\eta}{\tau} \right) (\eta\tau),$$

$$\text{em que, } \eta = \delta \sqrt{1 - r^2 \delta^2} \left(a_t(\alpha, \beta) + \frac{r\delta}{\sqrt{1 - r^2 \delta^2}} \right), \tau = \sqrt{1 - \delta^2} \text{ e } W_{\Phi} \left(\frac{\eta}{\tau} \right) = \frac{\phi \left(\frac{\eta}{\tau} \right)}{\Phi \left(\frac{\eta}{\tau} \right)}.$$

Demonstração:

Como $f_H(h) = 2\phi(h|0, 1) \mathbf{1}_{(0, \infty)}(h)$ e

$$\phi(\nu_h + a_t(\alpha_{\delta, \beta})) = \frac{\sqrt{1 - \delta^2}}{\sqrt{1 - r^2 \delta^2}} \phi \left(a_t(\alpha, \beta) \left| \frac{\delta(h - r)}{\sqrt{1 - r^2 \delta^2}}; \frac{1 - \delta^2}{1 - r^2 \delta^2} \right. \right).$$

Assim, usando o Lema 2.1, tem-se que

$$\begin{aligned} \phi \left(a_t(\alpha, \beta) \left| \frac{\delta(h - r)}{\sqrt{1 - r^2 \delta^2}}, \frac{1 - \delta^2}{1 - r^2 \delta^2} \right. \right) \phi(h|0, 1) &= \phi \left(a_t(\alpha, \beta) \left| -\frac{r\delta}{\sqrt{1 - r^2 \delta^2}}; \frac{1}{1 - r^2 \delta^2} \right. \right) \times \\ &\times \phi \left\{ h \left| \delta \sqrt{1 - r^2 \delta^2} \left(a_t(\alpha, \beta) + \frac{r\delta}{\sqrt{1 - r^2 \delta^2}} \right); 1 - \delta^2 \right. \right\}. \end{aligned}$$

Logo, a prova de i) segue diretamente da relação $f_{H|T}(h|t) = f_{T|H}(t|h) f_H(h) / f_T(t)$.

ii)–iii) Para $k = 1, 2$, tem-se que

$$\begin{aligned}
\mathbb{E}[H^k|T] &= \int_0^\infty \frac{h^k}{\Phi\left\{\lambda\left[\sigma_z\left(a_t(\alpha, \beta) + \frac{r\delta}{\sqrt{1-r^2\delta^2}}\right)\right]\right\}} \times \\
&\times \phi\left\{h|\delta\sqrt{1-r^2\delta^2}\left(a_t(\alpha, \beta) + \frac{r\delta}{\sqrt{1-r^2\delta^2}}\right); 1-\delta^2\right\} dh \\
&= \frac{1}{\Phi\left\{\lambda\left[\sigma_z\left(a_t(\alpha, \beta) + \frac{r\delta}{\sqrt{1-r^2\delta^2}}\right)\right]\right\}} \int_0^\infty h^k \times \\
&\times \phi\left\{h|\delta\sqrt{1-r^2\delta^2}\left(a_t(\alpha, \beta) + \frac{r\delta}{\sqrt{1-r^2\delta^2}}\right); 1-\delta^2\right\} dh \\
&= \mathbb{E}(X^k|X > 0).
\end{aligned}$$

Assim, usando diretamente as propriedades da distribuição Half-Normal descritas no Lema 2.1, conclui-se a demonstração.

2.4.1 Estimação por Máxima Verossimilhança via algoritmo ECM

Suponha n observações independentes, T_1, \dots, T_n , com $T_i \sim \text{BSNAC}(\alpha, \beta, \gamma)$, $i = 1, \dots, n$, em que α, β e γ são desconhecidos. A log-verossimilhança original de $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^\top$ é dada por

$$l(\boldsymbol{\theta}) = \sum_{i=1}^n l_i(\boldsymbol{\theta}), \quad (2.4.2)$$

em que,

$$\begin{aligned}
l_i(\boldsymbol{\theta}) &= \log(2) + \log(\sigma_z) + \log\left\{\phi\left[\sigma_z\left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z}\right)\right]\right\} \\
&+ \log\left\{\Phi\left[\lambda\left(\sigma_z\left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z}\right)\right)\right]\right\} + \log[A_{t_i}(\alpha, \beta)], \quad (2.4.3)
\end{aligned}$$

além disso, $a_{t_i}(\alpha, \beta)$ e $A_{t_i}(\alpha, \beta)$ são como em (2.2.8); μ_z e σ_z são como em (2.1.4).

A obtenção dos estimadores de máxima verossimilhança pode ser feita através da maximização de (2.4.2) mas, optamos por usar o algoritmo EM com maximização condicional (ECM) devido às simplicidades numéricas obtidas através dele. Nesse caso, devemos trabalhar com a verossimilhança completa, como apresentaremos a seguir.

Estimação via algoritmo EM com maximização condicional

Para efeito de estimação vamos considerar uma reparametrização em lugar de estimar $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^\top$, vamos estimar $\boldsymbol{\theta}$ dado por $\boldsymbol{\theta} = (\alpha, \beta, \delta)^\top$, em que $\delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}$ e λ é definido em função de γ tal como em (2.1.7). Como mencionado anteriormente, optamos por trabalhar com o parâmetro δ , visto que facilita a obtenção dos resultados tanto analiticamente, quanto computacionalmente. Além disso, avaliar a magnitude da assimetria através de δ torna-se mais fácil.

Utilizando a Proposição 2.1, a distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada tem a seguinte representação hierárquica

$$\begin{aligned} T_i | H_i = h_i &\stackrel{ind}{\sim} \text{BSEC}(\alpha_\delta, \beta, \sigma = 2, \nu_{h_i}) \\ H_i &\stackrel{ind}{\sim} \text{HN}(0, 1); i = 1, \dots, n, \end{aligned} \quad (2.4.4)$$

em que, $\alpha_\delta = \alpha \sqrt{\frac{1 - \delta^2}{1 - r^2 \delta^2}}$ e $\nu_{h_i} = -\frac{\delta(h_i - r)}{\sqrt{1 - \delta^2}}$.

Sejam $\mathbf{t} = (t_1, \dots, t_n)^\top$ e $\mathbf{h} = (h_1, \dots, h_n)^\top$, em que \mathbf{h} serão os dados faltantes, uma vez que eles são variáveis latentes, na linguagem da família de algoritmos EM. Então, de acordo com a representação hierárquica definida acima, segue que a log-verossimilhança completa associada à $\mathbf{t}^* = (\mathbf{t}^\top, \mathbf{h}^\top)^\top$ é dada por

$$\begin{aligned} l(\boldsymbol{\theta} | \mathbf{t}^*) &= \sum_{i=1}^n \log f_{T|H}(t_i | h_i) + \sum_{i=1}^n f_H(h_i) \\ &= c - \frac{\delta^2}{2(1 - \delta^2)} \sum_{i=1}^n h_i^2 + \frac{r\delta^2}{(1 - \delta^2)} \sum_{i=1}^n h_i - \frac{nr^2\delta^2}{2(1 - \delta^2)} + \frac{\delta\sqrt{1 - r^2\delta^2}}{1 - \delta^2} \sum_{i=1}^n h_i a_{t_i}(\alpha, \beta) - \\ &\quad - \frac{r\delta\sqrt{1 - r^2\delta^2}}{1 - \delta^2} \sum_{i=1}^n a_{t_i}(\alpha, \beta) - \frac{1 - r^2\delta^2}{2(1 - \delta^2)} \sum_{i=1}^n a_{t_i}^2(\alpha, \beta) + \frac{n}{2} \log(1 - r^2\delta^2) + \sum_{i=1}^n \log(t_i + \beta) - \\ &\quad - \frac{n}{2} \log(1 - \delta^2) - n \log(\alpha) - \frac{n}{2} \log \beta. \end{aligned} \quad (2.4.5)$$

Fazendo $\hat{h}_i = \mathbb{E}[H_i | \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, t_i]$ e $\hat{h}_i^2 = \mathbb{E}[H_i^2 | \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, t_i]$ aplicando as propriedades de esperança condicional de H dado $T = t$ descritas na Proposição 2.2 e usando os momentos da distribuição Half-Normal (Lachos, 2004), tem-se que

$$\hat{h}_i = \eta + W_\Phi\left(\frac{\eta}{\tau}\right) \tau, \quad (2.4.6)$$

e

$$\hat{h}_i^2 = \eta^2 + \tau^2 + W_\Phi\left(\frac{\eta}{\tau}\right) (\eta\tau), \quad (2.4.7)$$

em que, $\eta = \delta\sqrt{1 - r^2\delta^2} \left(a_{t_i}(\alpha, \beta) + \frac{r\delta}{\sqrt{1 - r^2\delta^2}} \right)$, $\tau = \sqrt{1 - \delta^2}$ e $W_\Phi\left(\frac{\eta}{\tau}\right) = \frac{\phi\left(\frac{\eta}{\tau}\right)}{\Phi\left(\frac{\eta}{\tau}\right)}$.

Seja $\boldsymbol{\theta}^{(r)} = (\alpha^{(r)}, \beta^{(r)}, \delta^{(r)})^\top$ a estimativa de $\boldsymbol{\theta}$ na k -ésima iteração. A esperança com respeito a t , condicionada à h , da log-verossimilhança completa, tem a forma

$$\begin{aligned}
Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) &= \mathbb{E}[l(\boldsymbol{\theta}|\mathbf{t}^*)|\mathbf{t}, \hat{\boldsymbol{\theta}}^{(r)}] \\
&= c - \frac{\delta^{(r)2}}{2(1-\delta^{(r)2})} \sum_{i=1}^n \hat{h}_i^2 + \frac{r\delta^{(r)2}}{(1-\delta^{(r)2})} \sum_{i=1}^n \hat{h}_i - \frac{nr^2\delta^{(r)2}}{2(\delta^{(r)2})} + \frac{\delta^{(r)}\sqrt{1-r^2\delta^{(r)2}}}{\alpha^{(r)}(\delta^{(r)2})} \sum_{i=1}^n \hat{h}_i b_{t_i} \\
&\quad - \frac{r\delta^{(r)}\sqrt{1-r^2\delta^{(r)2}}}{\alpha^{(r)}(1-\delta^{(r)2})} \sum_{i=1}^n b_{t_i} - \frac{1-r^2\delta^{(r)2}}{2\alpha^{(r)2}(1-\delta^{(r)2})} \sum_{i=1}^n b_{t_i}^2 + \frac{n}{2} \log(1-\delta^{(r)2}) + \sum_{i=1}^n \log(t_i + \beta^{(r)}) \\
&\quad - \frac{n}{2} \log(1-\delta^{(r)2}) - n \log(\alpha^{(r)}) - \frac{n}{2} \log(\beta^{(r)}), \tag{2.4.8}
\end{aligned}$$

em que, $b_{t_i} = b_{t_i}(\beta^{(r)}) = \sqrt{\frac{t_i}{\beta^{(r)}}} - \sqrt{\frac{\beta^{(r)}}{t_i}}$.

O algoritmo ECM é implementado da seguinte forma:

Passo E: Dado $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^{(r)}$, calcula-se $\hat{h}_i^{(r)}$ e $\hat{h}_i^2^{(r)}$, para $i = 1, \dots, n$ usando as equações (2.4.6) e (2.4.7),

Passo 1-CM: Fixar $\hat{\beta}^{(r)}$ e $\hat{\delta}^{(r)}$ e atualizar $\hat{\alpha}^{(r)}$ através da raiz positiva da seguinte equação do segundo grau

$$\hat{\alpha}^2 + \hat{\alpha} \frac{1}{(1-\hat{\delta}^2)} \left[\hat{\delta} \sqrt{1-r^2\hat{\delta}^2} \frac{1}{n} \sum_{i=1}^n \hat{h}_i b_{t_i} - r\hat{\delta} \sqrt{1-r^2\hat{\delta}^2} \frac{1}{n} \sum_{i=1}^n b_{t_i} \right] - \frac{(1-r^2\hat{\delta}^2)}{(1-\hat{\delta}^2)} \frac{1}{n} \sum_{i=1}^n \hat{h}_i b_{t_i}^2 = 0 \tag{2.4.9}$$

Ou seja,

$$\hat{\alpha} = \frac{-b + \sqrt{b^2 - 4ac}}{2a},$$

em que

$$\begin{aligned}
a &= 1, \\
b &= \frac{1}{(1-\hat{\delta}^2)} \left[\hat{\delta} \sqrt{1-r^2\hat{\delta}^2} \frac{1}{n} \sum_{i=1}^n \hat{h}_i b_{t_i} - r\hat{\delta} \sqrt{1-r^2\hat{\delta}^2} \frac{1}{n} \sum_{i=1}^n b_{t_i} \right], \\
c &= -\frac{(1-r^2\hat{\delta}^2)}{(1-\hat{\delta}^2)} \frac{1}{n} \sum_{i=1}^n \hat{h}_i b_{t_i}^2.
\end{aligned}$$

Passo 2-CM: Fixar $\hat{\alpha}^{(r+1)}$ e atualizar $\hat{\beta}^{(r)}$ e $\hat{\delta}^{(r)}$ usando

$$\begin{aligned}
\hat{\beta}^{(r+1)} &= \operatorname{argmax} Q(\hat{\alpha}^{(r+1)}, \beta, \hat{\delta}^{(r)}) \\
\hat{\delta}^{(r+1)} &= \operatorname{argmax} Q(\hat{\alpha}^{(r+1)}, \hat{\beta}^{(r+1)}, \delta) \tag{2.4.10}
\end{aligned}$$

A atualização de $\hat{\beta}^{(r+1)}$ e $\hat{\delta}^{(r+1)}$ pode ser feita através de algoritmos de otimização numérica. Para a obtenção dos resultados apresentados nesse trabalho, utilizamos a função de otimização

`optim`, disponível no software R (R Development Core Team, 2008), e dentro dela selecionamos o método L-BFGS-B (Byrd et al., 1995), para maximizar a função $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ com respeito a cada um dos parâmetros.

Para inicializar o algoritmo podemos usar como estimativas iniciais $\boldsymbol{\theta} = \boldsymbol{\theta}^{(0)}$ para α e β pode-se usar $\hat{\alpha}^{(0)} = [2(s/v) - 1]^{1/2}$ e $\hat{\beta}^{(0)} = (sv)^{1/2}$, em que $s = \frac{1}{n} \sum_{i=1}^n \hat{t}_i$ e $v = \left[\frac{1}{n} \sum_{i=1}^n 1/\hat{t}_i \right]^{-1}$. Após calcular $\hat{\alpha}^{(0)}$ e $\hat{\beta}^{(0)}$, obtem-se $z_i = \frac{1}{\hat{\alpha}^{(0)}} \left[\sqrt{t_i/\hat{\beta}^{(0)}} - \sqrt{\hat{\beta}^{(0)}/t_i} \right]$, observações que tem distribuição Normal assimétrica padrão sob a parametrização centrada. Assim, uma estimativa inicial para $\hat{\delta}^{(0)}$ pode ser obtida através da maximização da verossimilhança da NAC com respeito à δ . A verossimilhança da NAC pode ser definida da seguinte forma:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \left\{ 2\sigma_z \phi \left[\sigma_z \left(y_i + \frac{\mu_z}{\sigma_z} \right) \right] \Phi \left(\lambda \left[\sigma_z \left(y_i + \frac{\mu_z}{\sigma_z} \right) \right] \right) \right\}, \quad (2.4.11)$$

em que μ_z e σ_z são tais como em (2.1.4). Para mais detalhes, ver Vilca et.al (2011).

Por fim, podemos obter o parâmetro de interesse γ através da relação definida em (2.1.5). Como critério de convergência pode-se usar

$$\left\| \boldsymbol{\theta}^{(r)} - \boldsymbol{\theta}^{(r-1)} \right\| < \varepsilon.$$

Assim como sugerido por Vilca et.al (2011), a fim de garantir que as verdadeiras estimativas de máxima verossimilhança estão sendo obtidas, recomenda-se que o algoritmo ECM seja implementado utilizando diferentes estimativas iniciais e seja verificado se para todas elas o algoritmo está retornando estimativas similares.

Matriz de informação de Fisher observada

Assim como pode ser visto no Corolário 2.2, o cálculo dos momentos da variável aleatória T requer a utilização de métodos de integração numérica, dificultando assim a obtenção da matriz de informação de Fisher esperada. Portanto, optamos por trabalhar com a matriz de informação de Fisher observada.

Nesse caso, é importante ressaltar que sob certas condições de regularidade, para um tamanho de amostra suficientemente grande, o estimador de máxima verossimilhança, $\hat{\boldsymbol{\theta}}$, converge em distribuição para uma Normal p-variada com vetor de médias dado por $\boldsymbol{\theta}$ e matriz de variância-covariância dada pela inversa da matriz de informação de Fisher esperada, $I_E^{-1}(\boldsymbol{\theta})$, ou seja

$$\hat{\boldsymbol{\theta}} \xrightarrow{\mathcal{D}} N_p(\boldsymbol{\theta}, I_E^{-1}(\boldsymbol{\theta})). \quad (2.4.12)$$

Além disso, a matriz de informação de Fisher esperada converge em probabilidade para a matriz de informação de Fisher observada.

Antes de definirmos a matriz de informação de Fisher observada, consideremos T_1, \dots, T_n uma amostra aleatória da distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada, ou seja, $T_i \sim \text{BSNAC}(\alpha, \beta, \gamma)$, com $i = 1, \dots, n$. Assim, a log-verossimilhança original é da forma

$$l(\boldsymbol{\theta}) = \sum_{i=1}^n l_i(\boldsymbol{\theta}), \quad (2.4.13)$$

em que $l_i(\boldsymbol{\theta})$ é tal como (2.4.3). Logo, a matriz de informação de Fisher observada é dada por:

$$I(\boldsymbol{\theta}) = \begin{bmatrix} I_{\alpha\alpha} & I_{\beta\alpha} & I_{\delta\alpha} \\ I_{\alpha\beta} & I_{\beta\beta} & I_{\delta\beta} \\ I_{\alpha\delta} & I_{\beta\delta} & I_{\delta\delta} \end{bmatrix},$$

em que:

$$\begin{aligned} I_{\alpha\alpha} &= \frac{n}{\alpha^2} + \frac{6n\sigma_z^2}{\alpha^4} - \frac{3\sigma_z^2}{\alpha^4\beta} \sum_{i=1}^n t_i - \frac{3\sigma_z^2\beta}{\alpha^4} \sum_{i=1}^n \frac{1}{t_i} - \frac{2\mu_z\sigma_z}{\alpha^3} \sum_{i=1}^n \left(\sqrt{\frac{t_i}{\beta}} - \sqrt{\frac{\beta}{t_i}} \right) + \frac{2\lambda\sigma_z}{\alpha^3} \sum_{i=1}^n \left(\sqrt{\frac{t_i}{\beta}} - \sqrt{\frac{\beta}{t_i}} \right) \\ &\quad \times W_{\Phi}(\lambda\zeta) + \frac{\lambda^2\sigma_z^2}{\alpha^4} \sum_{i=1}^n \left(\frac{t_i}{\beta} - \frac{\beta}{t_i} - 2 \right) W_{\Phi}^{(1)}(\lambda\zeta), \\ I_{\alpha\beta} &= -\frac{\sigma_z^2}{\alpha^3\beta^2} \sum_{i=1}^n t_i - \frac{\sigma_z^2}{\alpha^3} \sum_{i=1}^n \frac{1}{t_i} - \frac{\mu_z\sigma_z}{2\alpha^2} \sum_{i=1}^n \left(\frac{\sqrt{t_i}}{\beta^{3/2}} + \frac{1}{\sqrt{t_i}\sqrt{\beta}} \right) + \frac{\lambda\sigma_z}{2\alpha^2} \sum_{i=1}^n \left(\frac{\sqrt{t_i}}{\beta^{3/2}} + \frac{1}{\sqrt{t_i}\sqrt{\beta}} \right) W_{\Phi}(\lambda\zeta) \\ &\quad + \frac{\lambda^2\sigma_z^2}{2\alpha^3} \sum_{i=1}^n \left(\frac{1}{\beta^2} - \frac{1}{t_i} \right) W_{\Phi}^{(1)}(\lambda\zeta), \\ I_{\alpha\delta} &= -\frac{2r^2\delta}{\alpha^3} \sum_{i=1}^n \left(\frac{t_i}{\beta} - \frac{\beta}{t_i} - 2 \right) + \frac{r(1-2\mu_z^2)}{\alpha^2\sigma_z} \sum_{i=1}^n \left(\sqrt{\frac{t_i}{\beta}} - \sqrt{\frac{\beta}{t_i}} \right) - \frac{r^2\delta^4 - 2\mu_z^2 + 1}{\alpha^2(1-\delta^2)^{3/2}\sigma_z} \sum_{i=1}^n \left(\frac{t_i}{\beta} - \frac{\beta}{t_i} - 2 \right) \\ &\quad \times W_{\Phi}(\lambda\zeta) + \frac{r\delta^2(\delta^2-2)\sigma_z}{\alpha^2(1-\delta^2)^2} \sum_{i=1}^n \left(\sqrt{\frac{t_i}{\beta}} - \sqrt{\frac{\beta}{t_i}} \right) W_{\Phi}^{(1)}(\lambda\zeta) - \frac{\delta(r^2\delta^4 - 2\mu_z^2 + 1)}{\alpha^3(1-\delta^2)^2} \sum_{i=1}^n \left(\frac{t_i}{\beta} - \frac{\beta}{t_i} - 2 \right) \\ &\quad \times W_{\Phi}^{(1)}(\lambda\zeta) \\ I_{\beta\beta} &= \frac{n}{2\beta^2} - \sum_{i=1}^n \frac{1}{(t_i + \beta)^2} - \frac{\sigma_z^2}{\alpha^2\beta^3} \sum_{i=1}^n t_i - \frac{\mu_z\sigma_z}{4\alpha} \sum_{i=1}^n \left(\frac{3\sqrt{t_i}}{\beta^{3/2}} - \frac{1}{\sqrt{t_i}\beta^{3/2}} \right) + \frac{\lambda\sigma_z}{4\alpha} \sum_{i=1}^n \left(\frac{3\sqrt{t_i}}{\beta^{3/2}} - \frac{1}{\sqrt{t_i}\beta^{3/2}} \right) \\ &\quad \times W_{\Phi}(\lambda\zeta) + \frac{\lambda^2\sigma_z^2}{4\alpha^2} \left(\frac{t_i}{\beta^3} + \frac{2}{\beta^2} + \frac{1}{\beta t_i} \right) W_{\Phi}^{(1)}(\lambda\zeta), \\ I_{\beta\delta} &= -\frac{r^2\delta}{\alpha^2} \sum_{i=1}^n \left(\frac{t_i}{\beta^2} - \frac{1}{t_i} \right) + \frac{r(1-2\mu_z^2)}{2\alpha\sigma_z} \sum_{i=1}^n \left(\frac{\sqrt{t_i}}{\beta^{3/2}} + \frac{1}{\sqrt{t_i}\sqrt{\beta}} \right) - \frac{r^2\delta^4 - 2\mu_z^2 + 1}{2\alpha(1-\delta^2)^{3/2}\sigma_z} \sum_{i=1}^n \left(\frac{\sqrt{t_i}}{\beta^{3/2}} + \frac{1}{\sqrt{t_i}\sqrt{\beta}} \right) \\ &\quad \times W_{\Phi}(\lambda\zeta) + \frac{r\delta^2(\delta^2-2)\sigma_z}{2\alpha(1-\delta^2)^2} \sum_{i=1}^n \left(\frac{\sqrt{t_i}}{\beta^{3/2}} + \frac{1}{\sqrt{t_i}\sqrt{\beta}} \right) W_{\Phi}^{(1)}(\lambda\zeta) - \frac{\delta(r^2\delta^4 - 2\mu_z^2 + 1)}{2\alpha^2(1-\delta^2)^2} \sum_{i=1}^n \left(\frac{t_i}{\beta^2} - \frac{1}{t_i} \right) \\ &\quad \times W_{\Phi}^{(1)}(\lambda\zeta), \end{aligned}$$

$$\begin{aligned}
I_{\delta\delta} &= -\frac{2nr^2}{\sigma_z^4} + \frac{r^2}{\alpha^2} \sum_{i=1}^n \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2 \right) - \frac{r^3\delta(2\mu_z^2 - 3)}{\alpha\sigma_z^3} \sum_{i=1}^n \left(\frac{\sqrt{t_i}}{\sqrt{\beta}} - \frac{\sqrt{\beta}}{\sqrt{t_i}} \right) + \frac{r(\delta^2 + 2)}{(1 - \delta^2)^{5/2}} \sum_{i=1}^n W_{\Phi}(\lambda\zeta) \\
&+ \left[\frac{\delta(-3r^2 + 2r^4\delta^2 - 2\mu_z^2 + \mu_z^4 - r^2\delta^4 + 3)}{\alpha(1 - \delta^2)^{5/2}\sigma_z^3} \right] \sum_{i=1}^n \left(\frac{\sqrt{t_i}}{\sqrt{\beta}} - \frac{\sqrt{\beta}}{\sqrt{t_i}} \right) W_{\Phi}(\lambda\zeta) \\
&+ \sum_{i=1}^n \left[\frac{1 + \mu_z^2(\delta^2 - 2)}{\alpha(1 - \delta^2)^{3/2}\sigma_z} \left(\frac{\sqrt{t_i}}{\sqrt{\beta}} - \frac{\sqrt{\beta}}{\sqrt{t_i}} \right) - \frac{\mu_z(\delta^2 - 2)}{(1 - \delta^2)^{3/2}} \right]^2 W_{\Phi}^{(1)}(\lambda\zeta),
\end{aligned}$$

em que $\zeta = \sigma_z \left[a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right]$; μ_z e σ_z são tais como em (2.1.4), $a_{t_i}(\alpha, \beta)$ é como em (2.2.8), $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ e λ é tal como em (2.1.7); $W_{\Phi}(\lambda\zeta) = \frac{\phi(\lambda\zeta)}{\Phi(\lambda\zeta)}$; $W_{\Phi}^{(1)}(\lambda\zeta) = -W_{\Phi}(\lambda\zeta)(\lambda\zeta + W_{\Phi}(\lambda\zeta))$ é derivada de $W_{\Phi}(\lambda\zeta)$ com relação à $\lambda\zeta$.

Utilizamos a matriz de informação de Fisher observada para obter os erros-padrão da seguinte forma

$$EP(\hat{\theta}) = \left[\text{Diag}(I^{-1}(\hat{\theta})) \right]^{1/2}.$$

Breve estudo de simulação

Um breve estudo de simulação foi conduzido com o intuito de ilustrar o fato de que o modelo BSNAC tem seus parâmetros bem recuperados e além disso, mostrar que a matriz de informação de Fisher observada provê boa aproximação para os erros-padrão, pois estes se aproximam daqueles obtidos pelas réplicas. Dessa forma, podemos comparar os modelos BSNAC e BSNA de forma mais apropriada.

Foram geradas $N = 1000$ réplicas de tamanho $n = 500$, para as quais foram considerados $\theta^{\top} = (\alpha, \beta, \gamma) = (0, 5; 1; 0, 67)$, valores estes que resultam em um comportamento fortemente assimétrico positivo da distribuição BSNAC.

Na Tabela 2.1 é possível observar a média das N estimativas, $\bar{\theta} = \sum_{j=1}^N \hat{\theta}_j / N$, bem como a média dos erros-padrão (EP), obtidos através da matriz de informação de Fisher observada, com mencionado anteriormente e também o erro padrão empírico das N estimativas que é calculado da seguinte forma

$$EP_{\text{emp}} = \sum_{j=1}^N (\hat{\theta}_j - \bar{\theta})^2 / N.$$

De acordo com a Tabela 2.1, percebe-se que o algoritmo ECM aqui desenvolvido comporta-se bem, pois as médias das estimativas dos parâmetros em questão aproximam-se dos verdadeiros valores previamente fixados e os erros-padrão, tanto o empírico como o teórico, são muito similares, indicando assim, que a utilização da Informação de Fisher observada como estimador da variância assintótica é apropriada.

Tabela 2.1: Distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada: média e erros-padrão das N estimativas de $\hat{\theta}$.

	$\hat{\theta}$	EP($\hat{\theta}$)	EP _{emp}
$\hat{\alpha}$	0,495	0,019	0,021
$\hat{\beta}$	1,003	0,032	0,028
$\hat{\gamma}$	0,667	0,015	0,012

2.4.2 Inferência bayesiana

Fundamentos teóricos sobre inferência bayesiana podem ser estudados em profundidade pela leitura de Berger (1985), Lee (1989), Press (1989), O'Hagan (1994), Bernardo e Smith (1994), Gamerman (1997), Gelman et al. (1997), entre outros. De modo geral, segundo esses autores, a inferência bayesiana se baseia na noção de probabilidade subjetiva, ou seja, no grau de credibilidade em determinada hipótese sobre quantidades de interesse não observáveis. Assim, essa probabilidade subjetiva pode variar entre indivíduos e/ou em função dos acontecimentos, estabelecendo-se uma probabilidade a priori sobre a hipótese. A atualização das informações com dados e evidências sobre a hipótese é realizada via teorema de Bayes, modificando e refinando a probabilidade subjetiva continuamente.

Segundo Gelman et al. (1997), para estabelecer probabilidades sobre parâmetros de interesse θ da população, sendo conhecidos os dados observados y , deve-se estabelecer um modelo de distribuição de probabilidades conjunta para θ e Y . A densidade conjunta é o produto das duas densidades que são frequentemente referenciadas como distribuição a priori de θ , $f(\theta)$, e distribuição amostral dos dados, $f(y|\theta)$. Tem-se assim que:

$$f(\theta, y) = f(\theta)f(y|\theta).$$

Utilizando-se o Teorema de Bayes obtém-se a partir dos valores conhecidos dos dados de y a densidade a posteriori $f(\theta|y)$:

$$f(\theta|y) = \frac{f(\theta, y)}{f(y)} = \frac{f(\theta)f(y|\theta)}{f(y)}, \quad (2.4.14)$$

onde $f(y) = \int f(\theta)f(y|\theta)d\theta$ é a distribuição marginal de y . Convenciona-se que a distribuição a priori seja denotada por $\pi(\theta)$ e a distribuição a posteriori por $\pi(\theta|y)$. A distribuição $f(y|\theta)$ é associada à verossimilhança $L(\theta|y)$, ou de forma simplificada $L(\theta)$. O fator $f(y)$, que não depende de θ , e com y fixado, pode ser considerado uma constante. Assim, a (2.4.14), pode ser reescrita como:

$$\pi(\theta|y) = cL(\theta)\pi(\theta),$$

em que, $c^{-1} = f(y) = \int \pi(\theta)L(\theta)d\theta$.

Método de Monte Carlo via Cadeia de Markov (MCMC)

Em muitos casos, não é possível obter as posteriores marginais de interesse e, portanto, métodos numéricos devem ser utilizados para se obter aproximações delas. Entre os principais métodos de simulação podem ser citados os métodos de integração de Monte Carlo e os de reamostragem - rejeição ou aceitação e rejeição (Newman, 1951 e Ripley, 1987), rejeição adaptativa (Gilks e Wild, 1992), reamostragem ponderada ou amostragem e reamostragem por importância (Rubin, 1988 e Smith e Gelfand, 1992).

Entre outros, Gilks et al. (1996) e Gelman et al. (1997) estudaram o método MCMC. De forma geral, trata-se de uma integração de Monte Carlo utilizando cadeias de Markov. Considere uma sequência de variáveis aleatórias X e uma função de interesse da população $f(X)$, a integração de Monte Carlo estima o valor esperado médio de $f(X)$ utilizando amostras geradas $X_t, t = 1, \dots, n$ da distribuição e aproximando $\mathbb{E}[f(X)]$:

$$\mathbb{E}[f(X)] \approx \frac{1}{n} \sum_{t=1}^n f(X_t).$$

O método de simulação MCMC gera amostras construindo uma cadeia de Markov constituída ao longo tempo até que a distribuição estacionária seja atingida. A maneira usual de constituir cadeias para realizar inferências sob o enfoque bayesiano é por meio de algoritmos como o amostrador de Gibbs (Geman e Geman, 1984) e o algoritmo de Metropolis-Hastings (Hastings, 1970).

A implementação da análise bayesiana utilizando o método de simulação via MCMC, foi realizada, no presente trabalho, com o auxílio do pacote `R2openBugs` (Sturtz et. al, 2005) do software R .

Distribuições à priori, densidade à posteriori conjunta e condicionais completas

Realizamos a implementação da análise bayesiana via MCMC, utilizando tanto a verossimilhança original do modelo BSNAC, quanto a estrutura hierárquica, da seguinte forma:

Verossimilhança original:

Seja n observações independentes, T_1, \dots, T_n , com $T_i \sim \text{BSNAC}(\alpha, \beta, \gamma), i = 1, \dots, n$, em que α, β e γ são desconhecidos. A verossimilhança de $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^\top$, dada a amostra observada $\mathbf{t} = (t_1, \dots, t_n)^\top$, é definida da seguinte forma:

$$L(\boldsymbol{\theta}|\mathbf{t}) = \prod_{i=1}^n 2\sigma_z \phi \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \Phi \left\{ \lambda \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \right\} A_{t_i}(\alpha, \beta), \quad (2.4.15)$$

em que $a_{t_i}(\alpha, \beta)$ e $A_{t_i}(\alpha, \beta)$ são tais como (2.2.8); μ_z e σ_z são tais como em (2.1.4) e λ tal como em (2.1.7).

Para uma completa definição do modelo, precisamos especificar as distribuições à priori para o vetor de parâmetros desconhecidos $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^\top$.

Para a escolha das prioris apresentadas na Tabela 2.2, levamos em consideração o espaço paramétrico em que cada um dos parâmetros estão definidos. Para o parâmetro de forma α , adotamos uma priori Gama com massa concentrada no intervalo $(0, 3)$, visto que tais valores são mais comumente observados na literatura. Para o parâmetro de escala, β , adotamos uma priori Gama pouco informativa e por fim, para o parâmetro de assimetria, γ , adotamos uma priori Uniforme contemplando todo o espaço paramétrico.

Tabela 2.2: Prioris associadas aos parâmetros do modelo probabilístico.

Parâmetros	Prioris
α	Gama ($r_\alpha; \lambda_\alpha$)
β	Gama ($r_\beta; \lambda_\beta$)
γ	U($a; b$)

Vamos assumir que os elementos do vetor $\boldsymbol{\theta}$ são independentes, de modo que a distribuição à priori conjunta de todas as quantidades desconhecidas tem densidade dada por:

$$\begin{aligned} \pi(\alpha, \beta, \gamma) &= \pi(\alpha)\pi(\beta)\pi(\gamma) \\ &\propto \alpha^{(r_\alpha-1)}\beta^{(r_\beta-1)} \exp\{-(\alpha\lambda_\alpha + \beta\lambda_\beta)\}. \end{aligned} \quad (2.4.16)$$

Combinando a verossimilhança (2.4.15) e as prioris adotadas em (2.4.16), a densidade à posteriori conjunta de todas as quantidades não observáveis é dada por

$$\begin{aligned} \pi(\boldsymbol{\theta}|\mathbf{t}) &\propto \prod_{i=1}^n \sigma_z \phi \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \Phi \left\{ \lambda \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \right\} A_{t_i}(\alpha, \beta) \\ &\times \left\{ \alpha^{(r_\alpha-1)} \beta^{(r_\beta-1)} \exp\{-(\alpha\lambda_\alpha + \beta\lambda_\beta)\} \right\}. \end{aligned} \quad (2.4.17)$$

Por fim, consideremos o modelo completo definido em (2.4.17). As condicionais completas de α , β e γ são dadas por

$$\begin{aligned} \pi(\alpha|\beta, \gamma, \mathbf{t}) &\propto \prod_{i=1}^n \left\{ \sigma_z \phi \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \Phi \left\{ \lambda \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \right\} A_{t_i}(\alpha, \beta) \right\} \\ &\times \left\{ \alpha^{(r_\alpha-1)} \exp\{-\alpha\lambda_\alpha\} \right\}, \\ \pi(\beta|\alpha, \gamma, \mathbf{t}) &\propto \prod_{i=1}^n \left\{ \sigma_z \phi \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \Phi \left\{ \lambda \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \right\} A_{t_i}(\alpha, \beta) \right\} \\ &\times \left\{ \beta^{(r_\beta-1)} \exp\{-\beta\lambda_\beta\} \right\}, \\ \pi(\gamma|\alpha, \beta, \mathbf{t}) &\propto \prod_{i=1}^n \sigma_z \phi \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \Phi \left\{ \lambda \left[\sigma_z \left(a_{t_i}(\alpha, \beta) + \frac{\mu_z}{\sigma_z} \right) \right] \right\} A_{t_i}(\alpha, \beta). \end{aligned} \quad (2.4.18)$$

Estrutura hierárquica:

Assim como mostrado anteriormente, a distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada tem a seguinte representação hierárquica

$$\begin{aligned} T_i | H_i = h_i &\stackrel{ind}{\sim} \text{BSEC}(\alpha_\delta, \beta, \sigma = 2, \nu_{h_i}) \\ H_i &\stackrel{ind}{\sim} \text{HN}(0, 1); i = 1, \dots, n, \end{aligned} \quad (2.4.19)$$

em que, $\alpha_\delta = \alpha \sqrt{\frac{1-\delta^2}{1-r^2\delta^2}}$ e $\nu_{h_i} = -\frac{\delta(h_i-r)}{\sqrt{1-\delta^2}}$. Consideremos agora, $\mathbf{t} = (t_1, \dots, t_n)^\top$ e $\mathbf{h} = (h_1, \dots, h_n)^\top$ denotando os dados observados e não observados, respectivamente. Assim, de acordo com a representação hierárquica definida em (2.4.19), segue que a verossimilhança associada aos dados completos $\mathbf{t}^* = (\mathbf{t}^\top, \mathbf{h}^\top)^\top$ é dada por

$$\begin{aligned} L(\boldsymbol{\theta} | \mathbf{t}^*) &= \prod_{i=1}^n f_{T|H}(t_i | h_i) f_H(h_i) \\ &= \prod_{i=1}^n \sqrt{2/\pi} \phi[\nu_{h_i} + a_{t_i}(\alpha, \beta)] A_{t_i}(\alpha, \beta) \exp\left\{-\frac{h_i^2}{2}\right\}. \end{aligned} \quad (2.4.20)$$

Combinando a verossimilhança completa em (2.4.20) e as prioris adotadas em (2.4.16), a densidade à posteriori conjunta de todas as quantidades não observáveis é dada por

$$\begin{aligned} \pi(\boldsymbol{\theta}, \mathbf{h} | \mathbf{t}) &\propto \prod_{i=1}^n \left\{ \phi[\nu_{h_i} + a_{t_i}(\alpha, \beta)] A_{t_i}(\alpha, \beta) \exp\left\{-\frac{h_i^2}{2}\right\} \right\} \times \\ &\quad \left\{ \alpha^{(r_\alpha-1)} \beta^{(r_\beta-1)} \exp\{-\alpha\lambda_\alpha + \beta\lambda_\beta\} \right\} \end{aligned} \quad (2.4.21)$$

Por fim, consideremos o modelo completo descrito em (2.4.21). As condicionais completas de α , β e $\boldsymbol{\gamma}$ e \mathbf{h} são dadas por:

$$\begin{aligned} \pi(\mathbf{h} | \alpha, \beta, \boldsymbol{\gamma}, \mathbf{t}) &\propto \prod_{i=1}^n \phi[\nu_{h_i} + a_{t_i}(\alpha, \beta)] A_{t_i}(\alpha, \beta) \exp\left\{-\frac{h_i^2}{2}\right\}, \\ \pi(\alpha | \beta, \boldsymbol{\gamma}, \mathbf{t}^*) &\propto \prod_{i=1}^n \left\{ \phi[\nu_{h_i} + a_{t_i}(\alpha, \beta)] A_{t_i}(\alpha, \beta) \exp\left\{-\frac{h_i^2}{2}\right\} \right\} \left\{ \alpha^{(r_\alpha-1)} \exp\{-\alpha\lambda_\alpha\} \right\}, \\ \pi(\beta | \alpha, \boldsymbol{\gamma}, \mathbf{t}^*) &\propto \prod_{i=1}^n \left\{ \phi[\nu_{h_i} + a_{t_i}(\alpha, \beta)] A_{t_i}(\alpha, \beta) \exp\left\{-\frac{h_i^2}{2}\right\} \right\} \left\{ \beta^{(r_\beta-1)} \exp\{-\beta\lambda_\beta\} \right\}, \\ \pi(\boldsymbol{\gamma} | \alpha, \beta, \mathbf{t}^*) &\propto \prod_{i=1}^n \phi[\nu_{h_i} + a_{t_i}(\alpha, \beta)] A_{t_i}(\alpha, \beta) \exp\left\{-\frac{h_i^2}{2}\right\}. \end{aligned} \quad (2.4.22)$$

As distribuições condicionais completas (2.4.18) e (2.4.22) podem ser facilmente simuladas/geradas no software `OpenBugs` (Spiegelhalter et. al, 2005) através da representação hierárquica e das distribuições a priori. Como é possível observar, as condicionais completas apresentadas não tem forma fechada, logo não é possível simular diretamente delas e nesse caso, se faz necessária a utilização de algoritmos tais como: *Metropolis-Hastings*, *slice sampling* e *rejeição adaptativa*.

Todas esses algoritmos estão implementados no `OpenBugs` e as condicionais completas podem ser facilmente implementadas no software `R`, por exemplo.

No Capítulo 4, apresentaremos um estudo de simulação completo, sob os enfoques frequentista e bayesiano, a fim de avaliar o modelo probabilístico aqui proposto com respeito a recuperação dos parâmetros.

2.5 Comparação entre a distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada e a Birnbaum-Saunders Normal assimétrica

Desde o artigo pioneiro de Azzalini (1985), tem-se conhecimento da grande dificuldade na estimação por máxima verossimilhança dos parâmetros do modelo Normal assimétrico. Tais dificuldades estão relacionadas com a não convergência dos estimadores para particulares resultados amostrais. Por exemplo, no caso uniparamétrico $NA(0, 1, \gamma)$, dado uma amostra aleatória dessa distribuição, pode-se mostrar (Branco et.al, 2004) que: se todas as observações x'_i s são positivas (negativas) então a verossimilhança será monotonamente crescente (decrecente) e o estimador de máxima verossimilhança de γ será infinito. No caso geral (três parâmetros) observamos que, há sempre um ponto de inflexão em $\gamma = 0$ na verossimilhança perfilada desse parâmetro. Correspondentemente, a matriz de informação de Fisher esperada é singular, quando $\gamma \rightarrow 0$.

Ilustraremos a seguir, que a distribuição Birnbaum-Saunders Normal assimétrica (Santana, 2009) herda do modelo Normal assimétrico esses problemas na estimação por máxima verossimilhança e que a utilização da parametrização centrada é uma forma de contornar tal dificuldade.

Para uma amostra aleatória T_1, T_2, \dots, T_n de uma $BSNA(\alpha, \beta, \gamma)$ a log-verossimilhança, definida em Santana(2009), é dada por

$$l(\boldsymbol{\theta}) = n \log(2) + \sum_{i=1}^n \log [A_{t_i}(\alpha, \beta)] + \sum_{i=1}^n \log \{ \phi [a_{t_i}(\alpha, \beta)] \} + \sum_{i=1}^n \log \{ \Phi [\lambda a_{t_i}(\alpha, \beta)] \}, \quad (2.5.1)$$

em que $a_{t_i}(\alpha, \beta)$ e $A_{t_i}(\alpha, \beta)$ são tais como em (2.2.8) e λ é como definido em (2.1.7).

Uma vez que a log-verossimilhança depende de três parâmetros α , β e γ consideraremos a correspondente log-verossimilhança perfilada. Por exemplo, a log-verossimilhança perfilada de γ é dada por $l(\hat{\alpha}(\gamma), \hat{\beta}(\gamma), \gamma)$, em que $\hat{\alpha}(\gamma)$ e $\hat{\beta}(\gamma)$ denotam os valores de α e β que maximizam a $l(\boldsymbol{\theta})$ para um dado valor de γ .

Seguindo Arellano-Valle and Azzalini (2008), para facilitar as interpretações, iremos considerar a log-verossimilhança perfilada relativa (LVPR), obtida pela subtração $l(\hat{\alpha}(\gamma), \hat{\beta}(\gamma), \gamma) - l(\hat{\alpha}(\gamma), \hat{\beta}(\gamma), \hat{\gamma})$, assim o máximo da nova função será igual a zero.

A Figura 2.6 mostra o gráfico de duas vezes a LVPR de γ , considerando uma amostra aleatória de tamanho 200 de uma $BSNA(0, 5; 1; 0, 67)$, gerada com o auxílio do pacote `sn` do software `R`. É possível observar uma forma não quadrática da log-verossimilhança gerada pela existência de um ponto de inflexão quando γ está bem próximo de zero, dificultando assim a obtenção dos estimadores de máxima verossimilhança, e introduzindo um comportamento assimétrico da distribuição amostral mesmo quando $n \rightarrow \infty$. Um outro problema que está relacionado à existência desse

ponto de inflexão é a singularidade da matriz de informação de Fisher esperada quando γ tende a zero. Este fato afeta não somente as propriedades assintóticas dos estimadores de máxima verossimilhança como também impede que avaliemos a existência de assimetria usando testes baseados nesta matriz, como por exemplo, o teste de Wald.

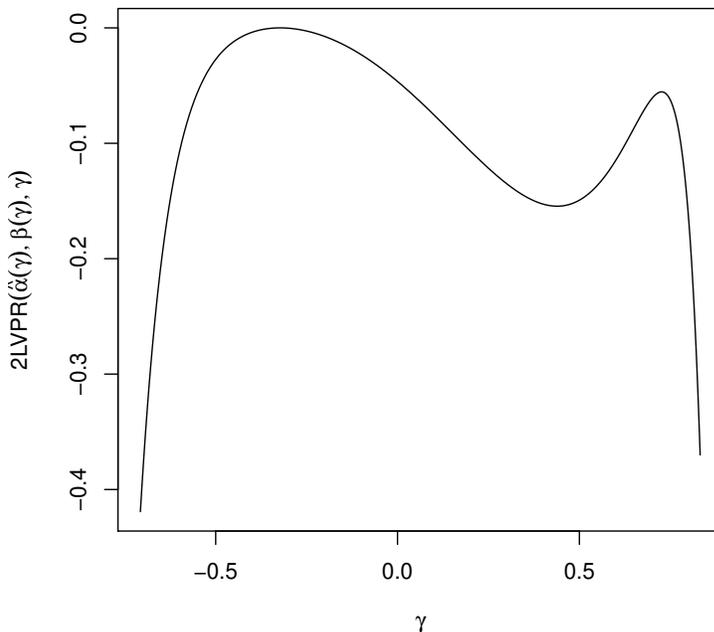


Figura 2.6: Duas vezes a LVPR de γ

Para estudar melhor as nuances desse problema, conduzimos um pequeno estudo de simulação. Neste estudo foram geradas 1000 réplicas de tamanho 20, 100, 500 e da BSNA(0, 5; 1; 0, 67), e para cada uma delas foram obtidas as estimativas de máxima verossimilhança $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$, via algoritmo ECM apresentado em Vilca et.al (2011).

As Figuras 2.7 - 2.9 mostram as correspondentes distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para os diferentes tamanhos amostrais em forma de histograma. A forma dessas distribuições amostrais sugerem, basicamente, um comportamento assimétrico que permanece mesmo quando $n \rightarrow \infty$.

Passando agora para o modelo aqui proposto, Birnbaum-Saunders Normal assimétrica sob a parametrização centrada, tem-se que a log-verossimilhança é dada em (2.4.2).

A Figura 2.10 mostra o gráfico de duas vezes a LVPR de γ , considerando uma amostra aleatória de tamanho 200 de uma BSNAC(0, 5; 1; 0, 67).

Diferentemente do que foi visto na Figura 2.6 o novo gráfico exibe um comportamento mais regular e mais próximo de uma função convexa. Além disso, não há ponto de inflexão em $\gamma = 0$.

De forma similar ao que foi feito anteriormente, conduzimos um pequeno estudo de simulação, agora usando a BSNAC(0, 5; 1; 0, 67). As Figuras 2.11 - 2.13 mostram as correspondentes distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para os diferentes tamanhos amostrais em forma de histograma.

Os gráficos sugerem que as distribuições amostrais dos três parâmetros utilizando a parametrização centrada são mais próximas da distribuição Normal, quando $n \rightarrow \infty$, do que as de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ utilizando a parametrização usual. De fato, a reparametrização parece remover a singularidade da matriz de informação de Fisher.

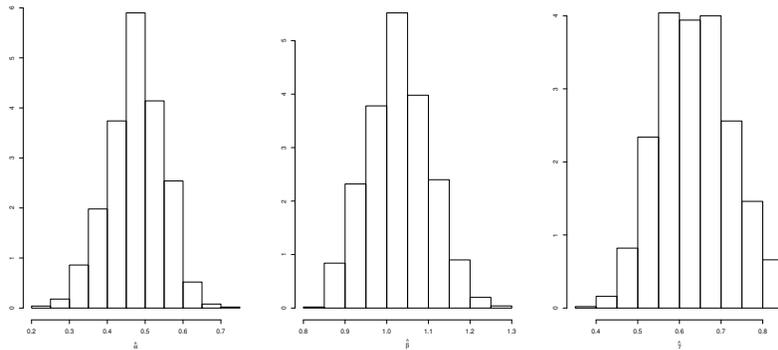


Figura 2.7: Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 20$.

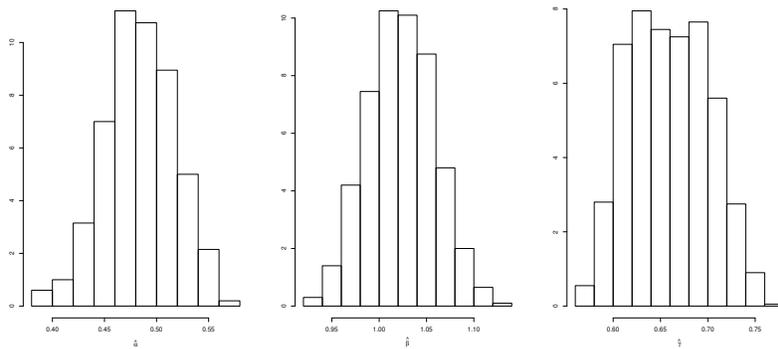


Figura 2.8: Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 100$.

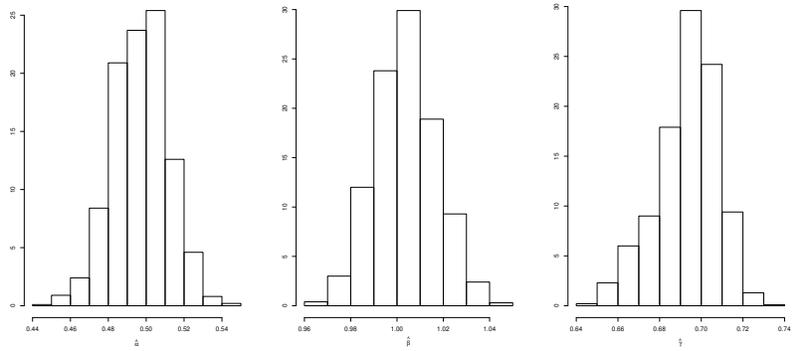


Figura 2.9: Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 500$.

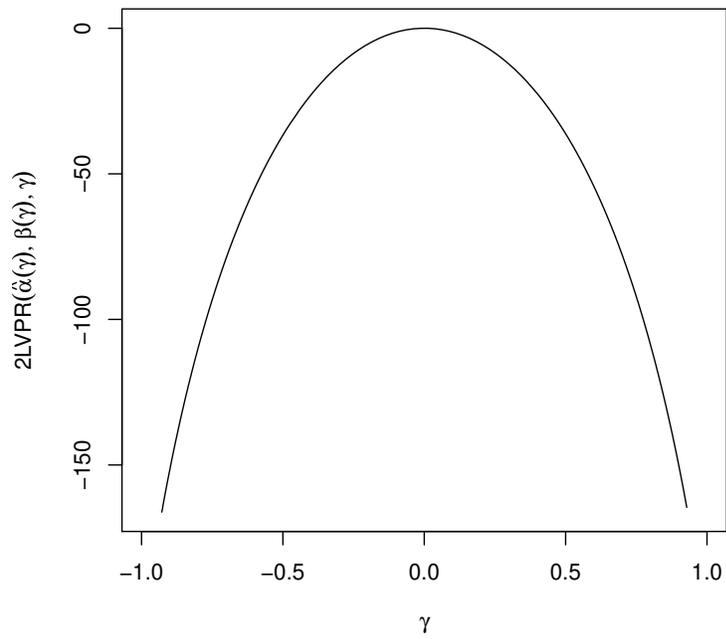


Figura 2.10: Duas vezes a LVPR de γ .

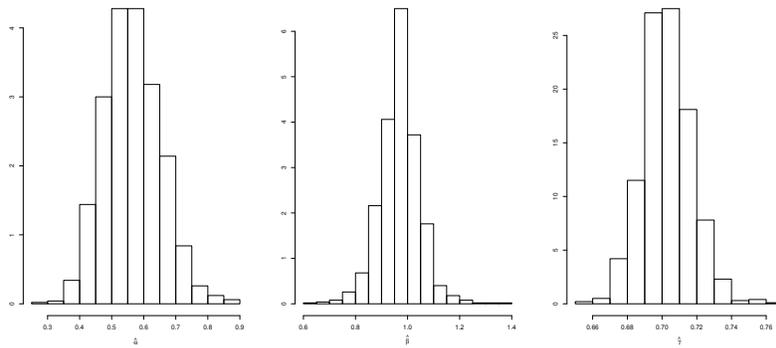


Figura 2.11: Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 20$.

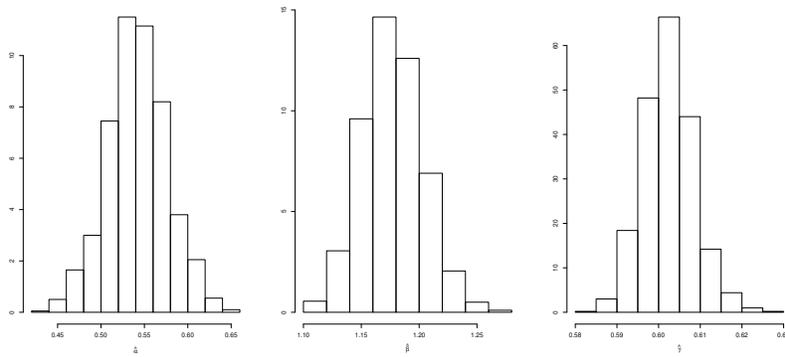


Figura 2.12: Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 100$.

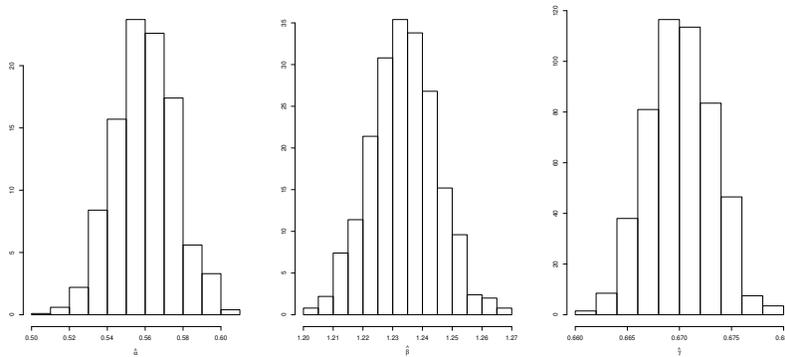


Figura 2.13: Distribuições amostrais de $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\gamma}$ para $n = 500$.

Capítulo 3

Modelo de Regressão Log-Birnbaum-Saunders Normal assimétrica sob a parametrização centrada

Neste capítulo iremos propor o modelo de regressão log-BSNAC. Especificamente vamos estudar propriedades da distribuição de $V = \log(T)$, quando $T \sim \text{BSNAC}(\alpha, \beta, \gamma)$, a qual será chamada de distribuição Seno Hiperbólico (Senh) Normal assimétrica sob a parametrização centrada (SHNAC), ou log-BSNAC. Destacaremos as principais propriedades do modelo de regressão log-BSNAC, bem como desenvolveremos procedimentos de estimação e análise de diagnóstico sob os enfoques frequentista e bayesiano. Além disso, apresentaremos ainda a análise de influência sob o ponto de vista bayesiano.

3.1 A distribuição Senh-Normal

Inicialmente, iremos apresentar a distribuição Seno Hiperbólico Normal, ou simplesmente Senh-Normal (SHN), que foi obtida através do log da distribuição BS. Rieck e Nedelman (1991) mostraram que essa distribuição pode ser usada em modelos para testes de vida acelerada ou para comparar vidas medianas de algumas populações. A distribuição SHN pode ser definida através de uma transformação da distribuição Normal baseada na seguinte relação

$$V = \rho + 2\text{arcsenh}\left(\frac{\alpha}{2}Z\right), \quad (3.1.1)$$

em que $Z \sim N(0, 1)$. Pode-se provar que sua densidade é dada por

$$\begin{aligned} f_V(v) &= \left(\frac{2}{\alpha\sigma\sqrt{2\pi}}\right) \exp\left[-2\alpha^{-2}\text{senh}^2\left(\frac{v-\rho}{\sigma}\right)\right] \cosh\left(\frac{v-\rho}{\sigma}\right), \\ &= \phi(b_v)b'_v, v \in \mathbb{R}. \end{aligned} \quad (3.1.2)$$

em que $\phi(\cdot)$ denota a densidade da Normal padrão. Além disso,

$$b_v = \frac{2}{\alpha} \sinh\left(\frac{v - \rho}{\sigma}\right) \quad \text{e} \quad b'_v = \frac{db_v}{dv} = \frac{2}{\alpha\sigma} \cosh\left(\frac{v - \rho}{\sigma}\right). \quad (3.1.3)$$

A distribuição de V é denotada por $V \sim \text{SHN}(\alpha, \rho, \sigma)$. A seguir, apresentaremos duas das principais propriedades da distribuição SHN.

- $Z = \frac{2}{\alpha} \sinh\left(\frac{V - \rho}{\sigma}\right) \sim \text{N}(0, 1)$;
- $W = a + bV \sim (\alpha, a + b\rho, b\sigma)$, em que $a \in \mathbb{R}$ e $b \in \mathbb{R} - \{0\}$. Em particular, se $V \sim \text{SHN}(\alpha, \rho, \sigma)$, então $[V - \rho] / \sigma \sim \text{SHN}(\alpha, 0, 1)$. Isto é, com esta transformação obtemos a distribuição SHN padrão.

A f.d.a de V é dada por

$$F_V(v) = \Phi\left[\frac{2}{\alpha} \sinh\left(\frac{v - \rho}{\sigma}\right)\right], \quad v \in \mathbb{R}, \quad (3.1.4)$$

em que Φ denota a acumulada da Normal padrão. Agora, se $T \sim \text{BS}(\alpha, \beta)$, então a distribuição de $V = \log(T)$ permite o estudo de modelos de regressão linear log-Birnbaum-Saunders, e neste caso, $V \sim \text{SHN}(\alpha, \rho, \sigma = 2)$.

A média e a variância da distribuição $\text{SHN}(\alpha, \rho, \sigma = 2)$ podem ser obtidas usando a função geradora de momentos apresentada em Gradshteyn e Randzhik (2000).

Rieck e Nedelman (1991) provaram que se $T \sim \text{BS}(\alpha, \beta)$, então $V = \log(T) \sim \text{SHN}(\alpha, \rho, \sigma = 2)$, em que $\rho = \log(\beta)$. Por isso, a distribuição Senh-Normal é denominada distribuição log-Birnbaum-Saunders. Para mais detalhes, ver também Leiva et al. (2009).

3.2 A distribuição SHNAC

Definamos agora, uma generalização da distribuição Senh-Normal (SHN), desenvolvida por Rieck e Nedelman (1991), que é baseada na distribuição de $V = \log(T)$, com $T \sim \text{BSNAC}(\alpha, \beta, \gamma)$. A variável aleatória V pode ser definida da seguinte forma

$$V = \rho + 2 \operatorname{arcsenh}\left(\frac{\alpha}{2} Y_c\right), \quad (3.2.1)$$

em que, $Y_c \sim \text{NAC}(0, 1, \gamma)$, $\alpha > 0$ é o parâmetro de forma e $\rho > 0$ é o parâmetro de localização. Ou seja, no lugar de assumir normalidade (caso Senh-Normal), vamos assumir que como variável geradora $Y_c \sim \text{NAC}(0, 1, \gamma)$. A distribuição de V será denotada por $V \sim \text{SHNAC}(\alpha, \rho, \sigma = 2, \gamma)$.

Definição 3.1 Dizemos que uma variável aleatória V tem distribuição SHNAC com parâmetro de forma $\alpha > 0$, parâmetro de localização $\rho \in \mathbb{R}$, parâmetro de escala $\sigma = 2$ e parâmetro de assimetria $\gamma \in (-0,99527; 0,99527)$, se sua densidade for dada por

$$f_V(v) = 2\sigma_z \phi\left[\sigma_z \left(b_v(\alpha, \rho) + \frac{\mu_z}{\sigma_z}\right)\right] b'_v(\alpha, \rho) \Phi\left\{\lambda \left[\sigma_z \left(b_v(\alpha, \rho) + \frac{\mu_z}{\sigma_z}\right)\right]\right\} \mathbf{1}_{(-\infty, \infty)}(v), \quad (3.2.2)$$

em que μ_z e σ_z são tais como em (2.1.4), λ tal como em (2.1.7), $\rho = \log(\beta)$ e

$$b_v(\alpha, \rho) = \frac{2}{\alpha} \sinh\left(\frac{v - \rho}{2}\right) \quad \text{e} \quad b'_v(\alpha, \rho) = \frac{1}{\alpha} \cosh\left(\frac{v - \rho}{2}\right). \quad (3.2.3)$$

Nas Figuras 3.1 e 3.2, observa-se os gráficos da densidade (3.2.2) para diferentes valores de α e γ .

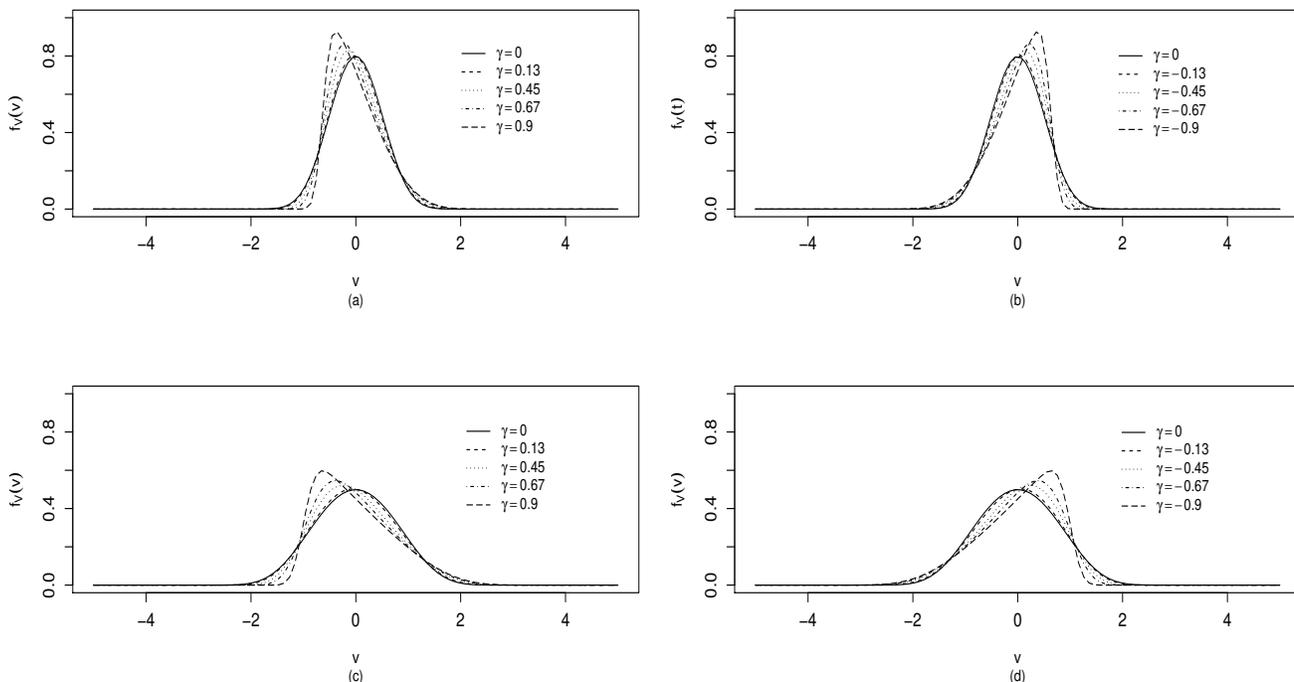


Figura 3.1: Densidades da SHNAC para diferentes valores de γ , considerando $\beta = 1$ ($\rho = 0$), (a)–(b) $\alpha = 0,2$ e (c)–(d) $\alpha = 0,8$.

Na Figura 3.1 apresentamos os gráficos da densidade (3.2.2) para diferentes valores de γ , considerando $\beta = 1$ ($\rho = 0$) e $\alpha = 0,2$ (painéis (a) e (b)) e $\alpha = 0,8$ (painéis (c) e (d)). É possível observar que valores negativos de γ induzem assimetria negativa (ou à esquerda) e valores positivos de γ induzem assimetria positiva (ou à direita). Se $\gamma = 0$ observamos um comportamento simétrico, além disso, à medida que aumentamos o valor de α é possível notar um aumento da variabilidade.

Na Figura 3.2 apresentamos os gráficos da densidade (3.2.2) para diferentes valores de α , considerando $\beta = 1$ ($\rho = 0$) e $\gamma = 0,9$ (painel (a)) e $\gamma = -0,9$ (painel (b)). É possível perceber que à medida que o valor de α decresce, há uma diminuição da variância, fazendo com que a densidade da SHNAC fique mais simétrica em torno de ρ .

Em síntese, em termos dos três parâmetros da distribuição SHNAC, é possível observar a simetria da distribuição em torno de ρ , quando fixamos $\gamma = 0$ e pequenos valores para α . A

assimetria positiva é observada à medida que α aumenta, ρ diminui e/ou γ assume valores positivos. Já a assimetria negativa é observada à medida que α diminui, ρ aumenta e/ou γ assume valores negativos. Lembrando que $\rho = \log(\beta)$.

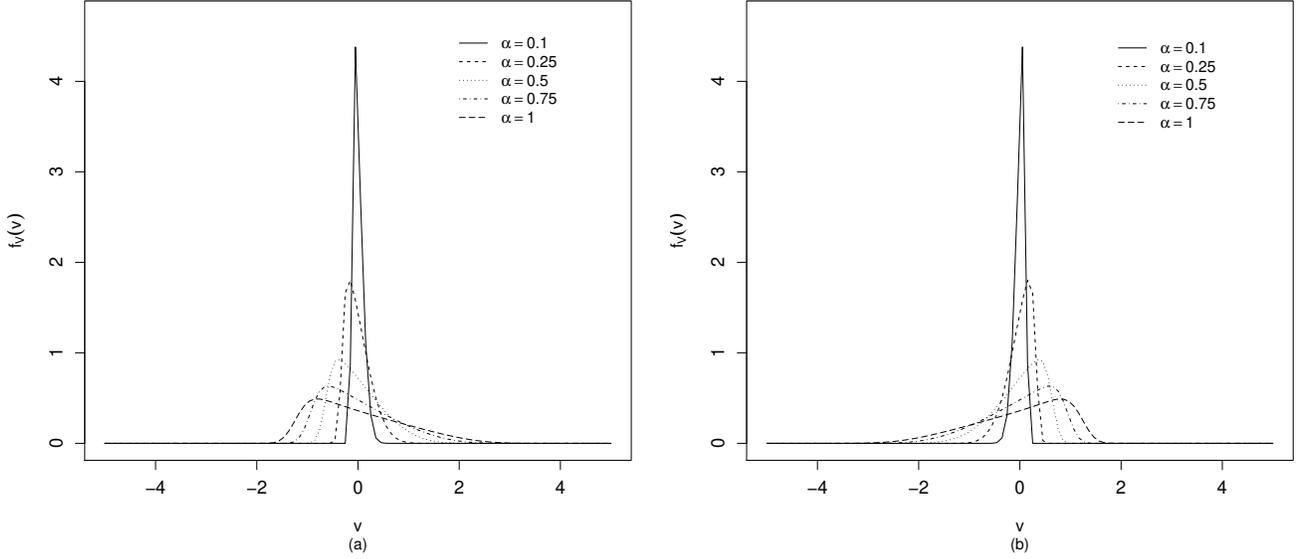


Figura 3.2: Densidades da SHNAC para diferentes valores de α , considerando $\beta = 1$ ($\rho = 0$), (a) $\gamma = 0,9$ e (b) $\gamma = -0,9$.

Apresentaremos a seguir, alguns resultados envolvendo a distribuição SHNAC que serão de suma importância para o desenvolvimento dos algoritmos de estimação do modelo de regressão.

De (3.2.1) temos que se, $V \sim \text{SHNAC}(\alpha, \rho, \sigma = 2, \gamma)$, então

$$Y_c = \frac{2}{\alpha} \sinh\left(\frac{v - \rho}{2}\right) \sim \text{NAC}(0, 1, \gamma), \quad (3.2.4)$$

em que $\rho = \log \beta$ e γ é tal como definido em (2.1.5).

Proposição 3.1 *Seja $V \sim \text{SHNAC}(\alpha, \rho, \sigma = 2, \gamma)$, a acumulada de SHNAC é dada por*

$$F_V(v) = \Phi_\gamma(b_v(\alpha, \rho)),$$

em que, $\Phi_\gamma(\cdot)$ é a acumulada de uma variável aleatória $\text{NAC}(0, 1, \gamma)$, que pode ser definida da seguinte forma

$$\Phi_\gamma(y_c) = \Phi(\sigma_z y_c + \mu_z) - 2T(\sigma_z y_c + \mu_z, a), \quad (3.2.5)$$

em que $T(w, a) = \frac{1}{2\pi} \int_0^a \frac{\exp[-w^2(1+x^2)/2]}{1+x^2} dx$. Essa função foi estudada por Owen (1956). Além disso, $\Phi(\cdot)$ representa a acumulada da Normal padrão, μ_z e σ_z são tais como em (2.1.4)

Proposição 3.2 *Seja Y_c como em (2.1.6) e $V \sim SHNAC(\alpha, \rho, \sigma = 2, \gamma)$. A variável aleatória Y_c tem representação estocástica dada como em (2.1.8). Então a função de densidade condicional de V dado $H = h$ é dada por*

$$f_{V|H}(v|h) = \phi(\nu_h + b_v(\alpha_\delta, \rho))b'_v(\alpha_\delta, \rho), \quad (3.2.6)$$

em que, $b_v(\alpha_\delta, \rho)$ e $b'_v(\alpha_\delta, \rho)$ são tais como (3.2.3), trocando apenas α por α_δ , $\alpha_\delta = \alpha\sqrt{\frac{1-\delta^2}{1-r^2\delta^2}}$ e $\nu_h = -\frac{\delta(h-r)}{\sqrt{1-\delta^2}}$, em que $\delta = \frac{1}{\sqrt{1+\lambda^2}}$ e λ é tal como em (2.1.7).

Demonstração:

Como $Y_c \sim NAC(0, 1, \delta)$, então usando a representação estocástica dada em (2.1.8) podemos expressar Y_c como

$$Y_c = \frac{1}{\sigma_z} \left[\delta H + \sqrt{1-\delta^2} X_1 - \mu_z \right] = \frac{2}{\alpha} \sinh \left(\frac{v-\rho}{2} \right).$$

Assim a distribuição condicional de Y_c dado $H = h$ é dada por

$$Y_c | (H = h) = \frac{2}{\alpha} \sinh \left(\frac{v-\rho}{2} \right) | (H = h) \sim N(\mu_h, \sigma_h^2),$$

em que $\mu_h = \frac{\delta(h-r)}{\sqrt{1-r^2\delta^2}}$ e $\sigma_h^2 = \frac{1-\delta^2}{1-r^2\delta^2}$. Logo, dado $H = h$ a distribuição de $W^* = (Y_c - \mu_h)/\sigma_h$ é dada por

$$W^* = -\frac{\mu_h}{\sigma_h} + \frac{2}{\alpha\sigma_h} \sinh \left(\frac{v-\rho}{2} \right) \sim N(0, 1).$$

Assim do resultado acima, conclui-se a prova.

Proposição 3.3 *Seja Y_c como em (2.1.6) e V como em (3.2.1), então*

$$i) f_{H|V}(h|v) = \frac{\phi \left\{ h \left| \delta \sqrt{1-r^2\delta^2} \left(b_v(\alpha, \rho) + \frac{r\delta}{\sqrt{1-r^2\delta^2}} \right); 1-\delta^2 \right\}}{\Phi \left\{ \lambda \left[\sigma_z \left(b_v(\alpha, \rho) + \frac{r\delta}{\sqrt{1-r^2\delta^2}} \right) \right] \right\}} \mathbf{1}_{(0, \infty)}(h).$$

$$ii) \mathbb{E}(H|V = v) = \eta^* + W_\Phi \left(\frac{\eta^*}{\tau} \right) \tau,$$

$$iii) \mathbb{E}(H^2|V = v) = \eta^{*2} + \tau^2 + W_\Phi \left(\frac{\eta^*}{\tau} \right) (\eta^* \tau),$$

$$\text{em que, } \eta^* = \delta \sqrt{1-r^2\delta^2} \left(b_v(\alpha, \rho) + \frac{r\delta}{\sqrt{1-r^2\delta^2}} \right), \tau = \sqrt{1-\delta^2} \text{ e } W_\Phi \left(\frac{\eta^*}{\tau} \right) = \frac{\phi \left(\frac{\eta^*}{\tau} \right)}{\Phi \left(\frac{\eta^*}{\tau} \right)}.$$

Demonstração: Como $f_H(h) = 2\phi(h|0, 1) \mathbf{1}_{(0, \infty)}(h)$ e

$$\phi(\nu_h + b_v(\alpha_{\delta, \rho})) = \frac{\sqrt{1 - \delta^2}}{\sqrt{1 - r^2\delta^2}} \phi\left(b_v(\alpha, \rho) \left| \frac{\delta(h - r)}{\sqrt{1 - r^2\delta^2}}; \frac{1 - \delta^2}{1 - r^2\delta^2} \right.\right).$$

Assim, usando o Lema 2.1, tem-se que

$$\begin{aligned} \phi\left(b_v(\alpha, \rho) \left| \frac{\delta(h - r)}{\sqrt{1 - r^2\delta^2}}, \frac{1 - \delta^2}{1 - r^2\delta^2} \right.\right) \phi(h|0, 1) &= \phi\left(b_v(\alpha, \rho) \left| -\frac{r\delta}{\sqrt{1 - r^2\delta^2}}; \frac{1}{1 - r^2\delta^2} \right.\right) \times \\ &\times \phi\left\{h \left| \delta\sqrt{1 - r^2\delta^2} \left(b_v(\alpha, \rho) + \frac{r\delta}{\sqrt{1 - r^2\delta^2}}\right); 1 - \delta^2 \right.\right\}. \end{aligned}$$

Logo, a prova de i) segue diretamente da relação $f_{H|V}(h|v) = f_{V|H}(v|h)f_H(h)/f_V(v)$.

ii)–iii) Para $k = 1, 2$, tem-se que

$$\begin{aligned} \mathbb{E}[H^k|V] &= \int_0^\infty \frac{h^k}{\Phi\left\{\lambda\left[\sigma_z\left(b_v(\alpha, \rho) + \frac{r\delta}{\sqrt{1 - r^2\delta^2}}\right)\right]\right\}} \times \\ &\times \phi\left\{h \left| \delta\sqrt{1 - r^2\delta^2} \left(b_v(\alpha, \rho) + \frac{r\delta}{\sqrt{1 - r^2\delta^2}}\right); 1 - \delta^2 \right.\right\} dh \\ &= \frac{1}{\Phi\left\{\lambda\left[\sigma_z\left(b_v(\alpha, \rho) + \frac{r\delta}{\sqrt{1 - r^2\delta^2}}\right)\right]\right\}} \int_0^\infty h^k \times \\ &\times \phi\left\{h \left| \delta\sqrt{1 - r^2\delta^2} \left(b_v(\alpha, \rho) + \frac{r\delta}{\sqrt{1 - r^2\delta^2}}\right); 1 - \delta^2 \right.\right\} dh \\ &= \mathbb{E}(X^k|X > 0). \end{aligned}$$

Assim, usando diretamente as propriedades da distribuição Half-Normal descritas no Lema 2.1, conclui-se a demonstração.

Proposição 3.4 *Seja $V \sim SHNAC(\alpha, \rho, \sigma = 2, \gamma)$ e $W = \mu + V$, $\mu \in R$, então*

$$W \sim SHNAC(\alpha, \mu + \rho, \sigma = 2, \gamma). \quad (3.2.7)$$

Demonstração: Usando o método do Jacobiano e a densidade de V dada em (3.2.2), temos que a densidade de W é dada por

$$f_W(w) = 2\sigma_z \phi \left\{ \sigma_z \left[b_w + \frac{\mu_z}{\sigma_z} \right] \right\} \Phi \left\{ \lambda \left[\sigma_z \left(b_w + \frac{\mu_z}{\sigma_z} \right) \right] \right\} b'_w \mathbf{1}_{(-\infty, \infty)}(w) ,$$

em que $b_w = \frac{2}{\alpha} \sinh \left(\frac{w - (\mu + \rho)}{2} \right)$ e $b'_w = \frac{1}{\alpha} \cosh \left(\frac{w - (\mu + \rho)}{2} \right)$.

Portanto, $W \sim \text{SHNAC}(\alpha, \mu + \rho, \sigma = 2, \gamma), \mu \in \mathbb{R}$

O resultado da Proposição 3.4 acima permite introduzir o modelo de regressão linear que discutiremos na seção seguinte.

3.3 O modelo de regressão log-Birnbaum-Saunders Normal assimétrica sob a parametrização centrada

Suponha que se tenha um conjunto de n observações independentes, denotadas por V_1, \dots, V_n , em que $V_i \sim \text{SHNAC}(\alpha, \mu_i, \sigma = 2, \gamma), i = 1, \dots, n$. Para definir o modelo log-Birnbaum-Saunders Normal assimétrica sob a parametrização centrada (log-BSNAC) procedemos como segue: associado com a observação i , considere \mathbf{x}_i um vetor $p \times 1$ de variáveis explicativas conhecidas associadas com a i -ésima resposta observada v_i , através do qual especifica-se o preditor linear $\mu_i = \mathbf{x}_i^\top \boldsymbol{\varphi}$, em que $\boldsymbol{\varphi}$ é um vetor p -dimensional de coeficientes de regressão desconhecidos. Assim, usando a Proposição 3.4, tem-se o modelo

$$V_i = \mathbf{x}_i^\top \boldsymbol{\varphi} + \varepsilon_i, \varepsilon_i \sim \text{SHNAC}(\alpha, 0, \sigma = 2, \gamma), i = 1, \dots, n. \quad (3.3.1)$$

Neste caso, $\varepsilon_i = \log T_i$, em que T_i tem distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada, $\text{BSNAC}(\alpha, 1, \gamma)$. Assim a log-verossimilhança para $\boldsymbol{\theta} = (\alpha, \boldsymbol{\varphi}^\top, \gamma)^\top$ pode ser escrita como

$$l(\boldsymbol{\theta}) = \sum_{i=1}^n \log \left\{ 2\sigma_z \phi \left\{ \sigma_z \left[b_{v_i} + \frac{\mu_z}{\sigma_z} \right] \right\} \Phi \left\{ \lambda \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] \right\} b'_{v_i} \right\}, \quad (3.3.2)$$

em que b_{v_i} e b'_{v_i} são tais como em (3.2.3).

A obtenção dos estimadores de máxima verossimilhança pode ser feita através da maximização de (3.3.2) mas, optamos por usar o algoritmo EM com maximização condicional (ECM) devido às simplicidades numéricas obtidas através dele. Nesse caso, devemos trabalhar com a verossimilhança completa, como apresentaremos a seguir.

3.4 Métodos de estimação

Nesta seção, apresentaremos as ferramentas necessárias para estimar os parâmetros do modelo de regressão log-BSNAC sob os pontos de vista frequentista e bayesiano.

3.4.1 Estimação via algoritmo EM com maximização condicional

Para efeito de estimação vamos considerar uma reparametrização, ou seja, em lugar de estimar $\boldsymbol{\theta} = (\alpha, \boldsymbol{\varphi}^\top, \gamma)^\top$, vamos estimar $\boldsymbol{\theta}$ dado por $\boldsymbol{\theta} = (\alpha, \boldsymbol{\varphi}^\top, \delta)^\top$, em que $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ e λ é definido em função de γ tal como em (2.1.7). Optamos por trabalhar com o parâmetro δ , visto que facilita a obtenção dos resultados tanto analiticamente, quanto computacionalmente. Além disso, avaliar a magnitude da assimetria através de δ torna-se mais fácil.

Para implementar o algoritmo ECM, primeiro note que, utilizando a representação estocástica da distribuição NAC dada em (2.1.8) e a Proposição 3.2, o modelo de regressão acima (4.2.1) pode ser escrito como

$$\begin{aligned} V_i | H_i = h_i &\stackrel{ind}{\sim} \text{SHNC}(\alpha_\delta, \mu_i, \sigma = 2, \nu_{h_i}) \\ H_i &\stackrel{ind}{\sim} \text{HN}(0, 1); i = 1, \dots, n, \end{aligned} \quad (3.4.1)$$

em que, $\alpha_\delta = \alpha \sqrt{\frac{1-\delta^2}{1-r^2\delta^2}}$ e $\nu_{h_i} = -\frac{\delta(h_i-r)}{\sqrt{1-\delta^2}}$.

A distribuição conjunta de v_i e h_i é dada por

$$f(v_i, h_i) = 2\phi(b_{v_i}^*) b_{v_i}^{*'} \phi(h_i) \mathbf{1}_{(0, \infty)}(h_i), \quad (3.4.2)$$

em que

$$b_{v_i}^* = \nu_{h_i} + \frac{2}{\alpha_\delta} \sinh\left(\frac{v_i - \mu_i}{2}\right) \quad \text{e} \quad b_{v_i}^{*'} = \frac{1}{\alpha_\delta} \cosh\left(\frac{v_i - \mu_i}{2}\right). \quad (3.4.3)$$

Seja $\mathbf{v} = (v_1, \dots, v_n)^\top$ e $\mathbf{h} = (h_1, \dots, h_n)^\top$. Então, tratando \mathbf{h} como dados faltantes, segue que a log-verossimilhança completa associada a $\mathbf{v}^* = (\mathbf{v}^\top, \mathbf{h}^\top)^\top$ é dada por

$$\begin{aligned} l(\boldsymbol{\theta} | \mathbf{v}^*) &= \sum_{i=1}^n \log f(v_i, h_i) + f_{H_i}(h_i) \\ &= c - n \log(\alpha) - \frac{n}{2} \log(1 - \delta^2) + \frac{n}{2} \log(1 - r^2\delta^2) + \sum_{i=1}^n \log \left[\cosh\left(\frac{v_i - \mu_i}{2}\right) \right] \\ &\quad - \frac{(1 - r^2\delta^2)}{2(1 - \delta^2)} \sum_{i=1}^n \left(b_{v_i} - \frac{\delta(h_i - r)}{\sqrt{1 - r^2\delta^2}} \right)^2, \end{aligned} \quad (3.4.4)$$

em que $b_{v_i} = \frac{2}{\alpha} \sinh\left(\frac{v_i - \mu_i}{2}\right)$, $r = \sqrt{\frac{2}{\pi}}$, $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ e λ é tal como em (2.1.5).

Fazendo $\hat{h}_i = \mathbb{E}[H_i | \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, v_i]$ e $\hat{h}_i^2 = \mathbb{E}[H_i^2 | \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, v_i]$ aplicando as propriedades de esperança condicional de H dado $V = v$ que foram obtidas na Proposição 3.3 e usando os momentos da distribuição Half-Normal (Lachos, 2004), tem-se que

$$\hat{h}_i = \eta^* + W_\Phi\left(\frac{\eta^*}{\tau}\right) \tau, \quad (3.4.5)$$

e

$$\hat{h}_i^2 = \eta^{*2} + \tau^2 + W_{\Phi} \left(\frac{\eta^*}{\tau} \right) (\eta^* \tau), \quad (3.4.6)$$

em que, $\eta^* = \delta \sqrt{1 - r^2 \delta^2} \left(b_{v_i} + \frac{r\delta}{\sqrt{1 - r^2 \delta^2}} \right)$, $\tau = \sqrt{1 - \delta^2}$ e $W_{\Phi} \left(\frac{\eta^*}{\tau} \right) = \frac{\phi \left(\frac{\eta^*}{\tau} \right)}{\Phi \left(\frac{\eta^*}{\tau} \right)}$.

Seja $\boldsymbol{\theta}^{(r)} = (\alpha^{(r)}, \boldsymbol{\varphi}^{(r)\top}, \delta^{(r)})^\top$ a estimativa de $\boldsymbol{\theta}$ na k -ésima iteração. A esperança da log-verossimilhança completa, que denotamos por Q , será dada por

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) &= \mathbb{E}[l(\boldsymbol{\theta}|\mathbf{v}^*)|\mathbf{v}, \hat{\boldsymbol{\theta}}^{(r)}] \\ &= c - \frac{\delta^{(r)2}}{2(1 - \delta^{(r)2})} \sum_{i=1}^n \hat{h}_i^{2(r)} + \frac{r\delta^{(r)2}}{1 - \delta^{(r)2}} \sum_{i=1}^n \hat{h}_i^{(r)} - \frac{nr^2\delta^{(r)2}}{2(1 - \delta^{(r)2})} + \frac{\delta^{(r)}\sqrt{1 - r^2\delta^{(r)2}}}{1 - \delta^{(r)2}} \sum_{i=1}^n b_{v_i} \hat{h}_i^{(r)} \\ &\quad - \frac{r\delta\sqrt{1 - r^2\delta^{(r)2}}}{1 - \delta^{(r)2}} \sum_{i=1}^n b_{v_i} - \frac{1 - r^2\delta^{(r)2}}{2(1 - \delta^{(r)2})} \sum_{i=1}^n b_{v_i}^2 + \frac{n}{2} \log(1 - r^2\delta^{(r)2}) - \frac{n}{2} \log(1 - \delta^{(r)2}) \\ &\quad - n \log(\alpha^{(r)}) + \sum_{i=1}^n \log \left[\cosh \left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2} \right) \right], \end{aligned}$$

em que $b_{v_i} = \frac{2}{\alpha^{(r)}} \sinh \left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2} \right)$.

A etapa M do algoritmo ECM é implementada maximizando a log-verossimilhança completa dada em (3.4.7). De forma análoga ao que foi apresentado no Capítulo 2, o algoritmo ECM é implementado da seguinte forma:

Passo E: Dado $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^{(r)}$, calcula-se $\hat{h}_i^{(r)}$ e $\hat{h}_i^{2(r)}$, para $i = 1, \dots, n$ usando as equações (3.4.5) e (3.4.6),

Passo 1-CM: Fixar $\hat{\boldsymbol{\varphi}}^{(r)}$ e $\hat{\delta}^{(r)}$ e atualizar $\hat{\alpha}^{(r)}$ através da raiz positiva da seguinte equação do segundo grau

$$\begin{aligned} \hat{\alpha}^2 &+ \hat{\alpha} \frac{2\delta\sqrt{1 - r^2\delta^2}}{n(1 - \delta^2)} \left[\sum_{i=1}^n \sinh \left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2} \right) \hat{h}_i^{(r)} - r \sum_{i=1}^n \sinh \left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2} \right) \right] - \\ &- \frac{4(1 - r^2\delta^2)}{n(1 - \delta^2)} \sum_{i=1}^n \sinh^2 \left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2} \right) = 0 \end{aligned} \quad (3.4.7)$$

Ou seja,

$$\hat{\alpha} = \frac{-b + \sqrt{b^2 - 4ac}}{2a},$$

em que

$$\begin{aligned}
a &= 1, \\
b &= \frac{2\delta\sqrt{1-r^2\delta^2}}{n(1-\delta^2)} \left[\sum_{i=1}^n \sinh\left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2}\right) \hat{h}_i^{(r)} - r \sum_{i=1}^n \sinh\left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2}\right) \right], \\
c &= -\frac{4(1-r^2\delta^2)}{n(1-\delta^2)} \sum_{i=1}^n \sinh^2\left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2}\right).
\end{aligned}$$

Passo 2-CM: Fixar $\hat{\alpha}^{(r+1)}$ e atualizar $\hat{\boldsymbol{\varphi}}^{(r)}$ e $\hat{\delta}^{(r)}$ usando

$$\begin{aligned}
\hat{\boldsymbol{\varphi}}^{(r+1)} &= \operatorname{argmax}_{\boldsymbol{\varphi}} Q\left(\hat{\alpha}^{(r+1)}, \boldsymbol{\varphi}, \hat{\delta}^{(r)}\right) \\
\hat{\delta}^{(r+1)} &= \operatorname{argmax}_{\delta} Q\left(\hat{\alpha}^{(r+1)}, \hat{\boldsymbol{\varphi}}^{(r+1)}, \delta\right)
\end{aligned} \tag{3.4.8}$$

A atualização de $\hat{\boldsymbol{\varphi}}^{(r+1)}$ e $\hat{\delta}^{(r+1)}$ pode ser feita através de algoritmos de otimização numérica. Para a obtenção dos resultados apresentados nesse trabalho de modo semelhante ao modelo probabilístico, utilizamos a função de otimização `optim`, disponível no software `R`, e dentro dela selecionamos o método L-BFGS-B, para maximizar $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ com respeito a cada um dos parâmetros.

Para inicializar o algoritmo podemos usar como estimativas iniciais $\boldsymbol{\theta} = \boldsymbol{\theta}^{(0)}$, para $\boldsymbol{\varphi}$, $\hat{\boldsymbol{\varphi}}^{(0)}$, usamos as estimativas de mínimos quadrados ordinários do modelo (4.2.1), $\hat{\boldsymbol{\varphi}}^{(0)} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{v}$, em que \mathbf{X} é a matriz de planejamento de dimensão $n \times p$ e $\mathbf{v} = (v_1, \dots, v_n)^\top$. Como estimativa inicial para α , $\hat{\alpha}^{(0)}$, pode-se usar $\hat{\alpha}^{(0)} = \left[\frac{4}{n} \sum_{i=1}^n \sinh^2\left(\frac{v_i - \mathbf{x}_i^\top \hat{\boldsymbol{\varphi}}^{(0)}}{2}\right) \right]^{1/2}$. Para mais detalhes, ver Lemonte e Cordeiro (2010).

Após calcular $\hat{\alpha}^{(0)}$ e $\hat{\boldsymbol{\varphi}}^{(0)}$, obtem-se $z_i = \left(\frac{2}{\hat{\alpha}^{(0)}}\right) \sinh\left(\frac{v_i - \mathbf{x}_i^\top \hat{\boldsymbol{\varphi}}^{(0)}}{2}\right)$; $i = 1, \dots, n$, observações que tem distribuição NAC. Assim, uma estimativa inicial para $\hat{\delta}^{(0)}$ pode ser obtida através da maximização da verossimilhança da NAC, (2.4.11), com respeito à δ .

Por fim, podemos recuperar o parâmetro de interesse γ através da relação definida em (2.1.5). Como critério de convergência pode-se usar

$$\left\| \boldsymbol{\theta}^{(r)} - \boldsymbol{\theta}^{(r-1)} \right\| < \varepsilon.$$

Recomenda-se que o algoritmo ECM seja implementado utilizando um conjunto de estimativas iniciais diferentes e seja verificado se para todas elas o algoritmo está retornando estimativas similares.

Matriz de informação de Fisher observada

Assim como mencionado no Capítulo 2, optamos por trabalhar com a matriz de informação de Fisher observada. Nesse caso, é importante ressaltar que sob certas condições de regularidade, para um tamanho de amostra suficientemente grande, o estimador de máxima verossimilhança, $\hat{\boldsymbol{\theta}}$, converge em distribuição para uma Normal p-variada com vetor de médias dado por $\boldsymbol{\theta}$ e matriz de variância-covariância dada pela inversa da matriz de informação de Fisher esperada, $I_E^{-1}(\boldsymbol{\theta})$, ou seja

$$\hat{\boldsymbol{\theta}} \xrightarrow{\mathcal{D}} N_p(\boldsymbol{\theta}, I_E^{-1}(\boldsymbol{\theta})). \quad (3.4.9)$$

Além disso, a matriz de informação de Fisher esperada converge em probabilidade para a matriz de informação de Fisher observada.

Antes de definirmos a matriz de informação de Fisher observada, consideremos V_1, \dots, V_n observações independentes, em que $V_i \sim \text{SHNAC}(\alpha, \mu_i, \sigma = 2, \gamma), i = 1, \dots, n$. A log-verossimilhança é da forma

$$l(\boldsymbol{\theta}) = \sum_{i=1}^n l_i(\boldsymbol{\theta}), \quad (3.4.10)$$

em que

$$l_i(\boldsymbol{\theta}) = \log(2) + \log(\sigma_z) + \log \left\{ \phi \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] \right\} + \log(b'_{v_i}) + \log \left\{ \Phi \left[\lambda \left(\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right) \right] \right\},$$

além disso, $b_{v_i} = \frac{2}{\alpha} \sinh \left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2} \right)$, $b'_{v_i} = \frac{1}{\alpha} \cosh \left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}^{(r)}}{2} \right)$, μ_z e σ_z são como em (2.1.4) e λ é tal como em (2.1.5). A matriz de informação de Fisher observada é dada por

$$I(\boldsymbol{\theta}) = \begin{bmatrix} I_{\alpha\alpha} & I_{\boldsymbol{\varphi}\alpha} & I_{\delta\alpha} \\ I_{\alpha\boldsymbol{\varphi}} & I_{\boldsymbol{\varphi}\boldsymbol{\varphi}} & I_{\delta\boldsymbol{\varphi}} \\ I_{\alpha\delta} & I_{\boldsymbol{\varphi}\delta} & I_{\delta\delta} \end{bmatrix},$$

em que,

$$\begin{aligned} I_{\alpha\alpha} &= \frac{n}{\alpha^2} - \frac{4\mu_z\sigma_z}{\alpha^3} \sum_{i=1}^n \sinh \left(\frac{v_i - \mu_i}{2} \right) - \frac{12\sigma_z^2}{\alpha^4} \sum_{i=1}^n \sinh^2 \left(\frac{v_i - \mu_i}{2} \right) + \frac{4\lambda\sigma_z}{\alpha^3} \sum_{i=1}^n W_\Phi(\lambda\zeta) \sinh \left(\frac{v_i - \mu_i}{2} \right) \\ &+ \frac{4\lambda^2\sigma_z^2}{\alpha^4} \sum_{i=1}^n W_\Phi^{(1)}(\lambda\zeta) \sinh^2 \left(\frac{v_i - \mu_i}{2} \right), \\ I_{\alpha\boldsymbol{\varphi}} &= -\frac{2\sigma_z^2}{\alpha^3} \mathbf{X}^\top \mathbf{a} + \frac{\lambda\sigma_z}{\alpha^2} \mathbf{X}^\top \text{Diag}(\mathbf{w}_1) \mathbf{f} + \frac{2\lambda^2\sigma_z^2}{\alpha^3} \mathbf{X}^\top \text{Diag}(\mathbf{w}_2) \text{Diag}(\mathbf{f}) \mathbf{a} - \frac{\mu_z\sigma_z}{\alpha^2} \mathbf{X}^\top \mathbf{f}, \end{aligned}$$

$$\begin{aligned}
I_{\alpha\delta} &= \frac{2r(1-2\mu_z^2)}{\alpha^2\sigma_z} \sum_{i=1}^n \sinh\left(\frac{v_i - \mu_i}{2}\right) - \frac{8r^2\delta}{\alpha^3} \sum_{i=1}^n \sinh^2\left(\frac{v_i - \mu_i}{2}\right) - \frac{2}{\alpha^2(1-\delta^2)^{3/2}} \sum_{i=1}^n W_{\Phi}(\lambda_{\zeta}) \times \\
&\times \sinh\left(\frac{v_i - \mu_i}{2}\right) - \frac{4\delta(1-2\mu_z^2+r^2\delta^4)}{\alpha^3\sigma_z(1-\delta^2)^2} \sum_{i=1}^n W_{\Phi}^{(1)}(\lambda_{\zeta}) \sinh^2\left(\frac{v_i - \mu_i}{2}\right) - \frac{2r\delta^2(2-\delta^2)}{\alpha^2(1-\delta^2)^2} \times \\
&\times \sum_{i=1}^n W_{\Phi}^{(1)}(\lambda_{\zeta}) \sinh\left(\frac{v_i - \mu_i}{2}\right), \\
I_{\varphi\varphi} &= \mathbf{X}^{\top} \mathbf{V} \mathbf{X} + \frac{\lambda\sigma_z}{2\alpha} \mathbf{X}^{\top} \text{Diag}(\mathbf{w}_1) \text{Diag}(\mathbf{b}) \mathbf{X} + \left(\frac{\lambda\sigma_z}{\alpha}\right)^2 \mathbf{X}^{\top} \mathbf{f} \mathbf{f}^{\top} \text{Diag}(\mathbf{w}_2) \mathbf{X} - \frac{\mu_z\sigma_z}{2\alpha} \mathbf{X}^{\top} \text{Diag}(\mathbf{b}) \mathbf{X}, \\
I_{\varphi\delta} &= -\frac{(1-2\mu_z^2+r^2\delta^4)}{\alpha\sigma_z(1-\delta^2)^{3/2}} \mathbf{X}^{\top} \text{Diag}(\mathbf{w}_1) \mathbf{f} - \frac{\delta(1-2\sigma_z+r^2\delta^4)}{\alpha^2(1-\delta^2)^2} \mathbf{X}^{\top} \text{Diag}(\mathbf{w}_2) \mathbf{a} - \frac{\sigma_z r \delta^2 (2-\delta^2)}{(1-\delta^2)^2} \times \\
&\times \mathbf{X}^{\top} \text{Diag}(\mathbf{w}_2) \mathbf{f} - \frac{2r^2\delta}{\alpha^2} \mathbf{X}^{\top} \mathbf{a} + \frac{r(1-2\mu_z^2)}{\alpha\sigma_z} \mathbf{X}^{\top} \mathbf{f}, \\
I_{\delta\delta} &= -\frac{nr^2(r^4\delta^2 - \mu_z^2 + 2)}{(1-\mu_z^2)^2} + \frac{2r^2 \left[2(1-\mu_z^2)^{3/2} + 2\mu_z(3-2\mu_z^2) \right]}{\alpha(1-\mu_z^2)^{3/2}} \sum_{i=1}^n \sinh\left(\frac{v_i - \mu_i}{2}\right) \\
&+ \frac{r \left[(2+\delta^2)\sqrt{1-\delta^2} + r\delta^2(2-\delta^2)^2 \right]}{(1-\delta^2)^3} \sum_{i=1}^n W_{\Phi}(\lambda_{\zeta}) + \frac{4\mu_z(2-\delta^2)(1-2\mu_z^2+r^2\delta^4)}{\alpha\sigma_z(1-\delta^2)^3} \times \\
&\times \sum_{i=1}^n W_{\Phi}^{(1)}(\lambda_{\zeta}) \sinh\left(\frac{v_i - \mu_i}{2}\right) + \frac{2\delta(-3r^2 - 2\mu_z^2 + 2r^4\delta^2 + \mu_z^4 - r^2\delta^4 + 3)}{\alpha\sigma_z^3(1-\delta^2)^{5/2}} \sum_{i=1}^n W_{\Phi}^{(1)}(\lambda_{\zeta}) \\
&+ \frac{4(1-2\mu_z^2+\mu_z^4)}{\alpha^2\sigma_z(1-\delta^2)^3} \sum_{i=1}^n W_{\Phi}^{(1)}(\lambda_{\zeta}) \sinh^2\left(\frac{v_i - \mu_i}{2}\right),
\end{aligned}$$

em que

$$\begin{aligned}
\mathbf{X} &= \begin{bmatrix} x_{11} & \dots & x_{1p} \\ x_{21} & \dots & x_{2p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{np} \end{bmatrix}, \\
\mathbf{a} &= \left(\sinh(v_1 - x_1^{\top} \boldsymbol{\varphi}), \dots, \sinh(v_n - x_n^{\top} \boldsymbol{\varphi}) \right)^{\top}, \\
\mathbf{b} &= \left(\sinh\left(\frac{v_1 - x_1^{\top} \boldsymbol{\varphi}}{2}\right), \dots, \sinh\left(\frac{v_n - x_n^{\top} \boldsymbol{\varphi}}{2}\right) \right)^{\top}, \\
\mathbf{c} &= \left(\text{sech}^2\left(\frac{v_1 - x_1^{\top} \boldsymbol{\varphi}}{2}\right), \dots, \text{sech}^2\left(\frac{v_n - x_n^{\top} \boldsymbol{\varphi}}{2}\right) \right)^{\top},
\end{aligned}$$

$$\begin{aligned}
\mathbf{d} &= \left(\cosh(v_1 - x_1^\top \boldsymbol{\varphi}), \dots, \cosh(v_n - x_n^\top \boldsymbol{\varphi}) \right)^\top \\
\mathbf{f} &= \left(\cosh\left(\frac{v_1 - x_1^\top \boldsymbol{\varphi}}{2}\right), \dots, \cosh\left(\frac{v_n - x_n^\top \boldsymbol{\varphi}}{2}\right) \right)^\top, \\
\mathbf{V} &= \frac{1}{4} \text{Diag}(\mathbf{c}) - \frac{1}{\alpha^2} \text{Diag}(\mathbf{d}),
\end{aligned}$$

além disso, o vetor $\mathbf{w}_1 = (w_{1_1}, \dots, w_{1_n})^\top$ tem como i -ésima componente $w_{1_i} = W_\Phi(\lambda \varsigma_i)$ e o vetor $\mathbf{w}_2 = (w_{2_1}, \dots, w_{2_n})^\top$ tem como i -ésima componente $w_{2_i} = W_\Phi^{(1)}(\lambda \varsigma_i)$; $\varsigma = \sigma_z \left[b_{v_i} + \frac{\mu_z}{\sigma_z} \right]$; μ_z e σ_z são tais como em (2.1.4); $W_\Phi^{(1)}(\lambda \varsigma) = -W_\Phi(\lambda \varsigma)(\lambda \varsigma + W_\Phi(\lambda \varsigma))$ é a derivada de $W_\Phi(\lambda \varsigma) = \frac{\phi(\lambda \varsigma)}{\Phi(\lambda \varsigma)}$ com relação à $\lambda \varsigma$; $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ e λ é tal como em (2.1.7). Para a obtenção de alguns dos resultados acima, usamos duas propriedades que envolvem funções trigonométricas, são elas:

- $2\sinh(x)\cosh(x) = \sinh\left(\frac{x}{2}\right),$
- $\frac{d \sinh(x)}{dx \cosh(x)} = \text{sech}^2(x).$

Utilizamos a matriz de informação de Fisher observada para obter os erros-padrão da seguinte forma

$$EP(\hat{\boldsymbol{\theta}}) = \left[\text{Diag}(I^{-1}(\hat{\boldsymbol{\theta}})) \right]^{1/2}.$$

3.4.2 Inferência bayesiana

Nesta seção, apresentaremos as ferramentas necessárias para estimar os parâmetros do modelo de regressão log-BSNAC sob o ponto de vista bayesiano.

Distribuições à priori, densidade à posteriori conjunta e condicionais completas

Realizamos a implementação da análise bayesiana via MCMC, utilizando tanto a verossimilhança original do modelo log-BSNAC, quanto a estrutura hierárquica, da seguinte forma:

Verossimilhança original:

Suponha que se tenha um conjunto de n observações independentes, denotadas por V_1, \dots, V_n , em que $V_i \sim \text{SHNAC}(\alpha, \mu_i, \sigma = 2, \gamma), i = 1, \dots, n$. A verossimilhança de $\boldsymbol{\theta} = (\alpha, \boldsymbol{\varphi}^\top, \gamma)^\top$, dada a amostra observada $\mathbf{v} = (v_1, \dots, v_n)^\top$, é definida da seguinte forma:

$$L(\boldsymbol{\theta}|\mathbf{v}) = \prod_{i=1}^n 2\sigma_z \phi \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] b'_{v_i} \Phi \left\{ \lambda \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] \right\}, \quad (3.4.11)$$

em que $b_{v_i} = \frac{2}{\alpha} \sinh\left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}}{2}\right)$, $b'_{v_i} = \frac{1}{\alpha} \cosh\left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}}{2}\right)$; μ_z e σ_z são como em (2.1.4) e λ é tal como em (2.1.5).

Para uma completa definição do modelo, precisamos especificar as distribuições à priori para o vetor de parâmetros desconhecidos $\boldsymbol{\theta} = (\alpha, \boldsymbol{\varphi}^\top, \gamma)^\top$.

Para a escolha das prioris apresentadas na Tabela 3.1, levamos em consideração o espaço paramétrico em que cada um dos parâmetros estão definidos. Para o parâmetro de forma α , adotamos uma priori Gama com massa concentrada no intervalo $(0, 3)$, visto que tais valores são mais comumente observados na literatura. Para os parâmetros de regressão, φ_k , adotamos uma priori Normal pouco informativa e por fim, para o parâmetro de assimetria, γ , adotamos uma priori Uniforme contemplando todo o espaço paramétrico.

Tabela 3.1: Prioris associadas aos parâmetros do modelo de regressão.

Parâmetros	Prioris
α	Gama ($r_\alpha; \lambda_\alpha$)
$\boldsymbol{\varphi}$	$N_p(\boldsymbol{\mu}_\varphi; \boldsymbol{\Psi})$
γ	$U(a; b)$

Vamos assumir que os elementos do vetor $\boldsymbol{\theta}$ são independentes, de modo que a distribuição à priori conjunta de todas as quantidades desconhecidas tem densidade dada por:

$$\begin{aligned} \pi(\alpha, \boldsymbol{\varphi}, \gamma) &= \pi(\alpha)\pi(\boldsymbol{\varphi})\pi(\gamma) \\ &\propto \alpha^{(r_\alpha-1)} \exp\{-\alpha\lambda_\alpha\} \exp\left\{-\frac{1}{2}(\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi)^\top \boldsymbol{\Psi}^{-1}(\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi)\right\}. \end{aligned} \quad (3.4.12)$$

Combinando a verossimilhança em (3.4.11) e as prioris adotadas em (3.4.12), a densidade à posteriori conjunta de todas as quantidades não observáveis é dada por

$$\begin{aligned} \pi(\boldsymbol{\theta}|\mathbf{v}) &\propto \prod_{i=1}^n \left\{ \sigma_z \phi \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] b'_{v_i} \Phi \left\{ \lambda \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] \right\} \right\}, \\ &\times \alpha^{(r_\alpha-1)} \exp\{-\alpha\lambda_\alpha\} \exp\left\{-\frac{1}{2}(\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi)^\top \boldsymbol{\Psi}^{-1}(\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi)\right\}. \end{aligned} \quad (3.4.13)$$

Por fim, consideremos o modelo completo descrito em (3.4.13). As condicionais completas de α , $\boldsymbol{\varphi}$ e γ são dadas por:

$$\begin{aligned}
\pi(\alpha|\boldsymbol{\varphi}, \gamma, \mathbf{v}) &\propto \prod_{i=1}^n \left\{ \sigma_z \phi \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] b'_{v_i} \Phi \left\{ \lambda \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] \right\} \right\} \\
&\times \left\{ \alpha^{(r\alpha-1)} \exp \{-\alpha \lambda_\alpha\} \right\}, \\
\pi(\boldsymbol{\varphi}|\alpha, \gamma, \mathbf{v}) &\propto \prod_{i=1}^n \left\{ \sigma_z \phi \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] b'_{v_i} \Phi \left\{ \lambda \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] \right\} \right\} \\
&\times \left\{ \exp \left\{ -\frac{1}{2} (\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi)^\top \boldsymbol{\Psi}^{-1} (\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi) \right\} \right\}, \\
\pi(\gamma|\alpha, \boldsymbol{\varphi}, \mathbf{v}) &\propto \prod_{i=1}^n \left\{ \sigma_z \phi \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] b'_{v_i} \Phi \left\{ \lambda \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] \right\} \right\}. \tag{3.4.14}
\end{aligned}$$

Estrutura hierárquica:

Assim como mostrado anteriormente, a distribuição SHNAC tem a seguinte representação hierárquica:

$$\begin{aligned}
V_i|H_i = h_i &\stackrel{ind}{\sim} \text{SHNC}(\alpha_\delta, \mu_i, \sigma = 2, \nu_{h_i}) \\
H_i &\stackrel{ind}{\sim} \text{HN}(0, 1); i = 1, \dots, n, \tag{3.4.15}
\end{aligned}$$

em que, $\alpha_\delta = \alpha \sqrt{\frac{1-\delta^2}{1-r^2\delta^2}}$ e $\nu_{h_i} = -\frac{\delta(h_i-r)}{\sqrt{1-\delta^2}}$.

Consideremos agora, $\mathbf{v} = (v_1, \dots, v_n)^\top$ e $\mathbf{h} = (h_1, \dots, h_n)^\top$ denotando os dados observados e não observados, respectivamente. Assim, de acordo com a representação hierárquica definida em (3.4.15), segue que a verossimilhança associada aos dados completos $\mathbf{v}^* = (\mathbf{v}^\top, \mathbf{h}^\top)^\top$ é dada por

$$\begin{aligned}
L(\boldsymbol{\theta}|\mathbf{v}^*) &= \prod_{i=1}^n f_{V|H}(v_i|h_i) f_H(h_i) \\
&= \prod_{i=1}^n 2\sqrt{2/\pi} \phi(b_{v_i}^*) b_{v_i}^{*'} \exp \left\{ -\frac{h_i^2}{2} \right\}, \tag{3.4.16}
\end{aligned}$$

em que $b_{v_i}^*$ e $b_{v_i}^{*'}$ são tais como em (3.4.3).

Combinando a verossimilhança completa em (3.4.16) e as prioris adotadas em (3.4.12), a densidade à posteriori conjunta de todas as quantidades não observáveis é dada por

$$\begin{aligned}
\pi(\boldsymbol{\theta}, \mathbf{h}|\mathbf{v}) &\propto \prod_{i=1}^n \left\{ \phi(b_{v_i}^*) b_{v_i}^{*'} \exp \left\{ -\frac{h_i^2}{2} \right\} \right\} \times \\
&\times \left\{ \alpha^{(r\alpha-1)} \exp \{-\alpha \lambda_\alpha\} \exp \left\{ -\frac{1}{2} (\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi)^\top \boldsymbol{\Psi}^{-1} (\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi) \right\} \right\}. \tag{3.4.17}
\end{aligned}$$

Por fim, consideremos o modelo completo descrito em (3.4.17). As condicionais completas de

α , $\boldsymbol{\varphi}$, γ e \mathbf{h} são dadas por:

$$\begin{aligned}
\pi(\mathbf{h}|\alpha, \boldsymbol{\varphi}, \gamma, \mathbf{v}) &\propto \prod_{i=1}^n \phi(b_{v_i}^*) b_{v_i}^{*'} \exp\left\{-\frac{h_i^2}{2}\right\}, \\
\pi(\alpha|\boldsymbol{\varphi}, \gamma, \mathbf{v}^*) &\propto \prod_{i=1}^n \left\{ \phi(b_{v_i}^*) b_{v_i}^{*'} \exp\left\{-\frac{h_i^2}{2}\right\} \right\} \left\{ \alpha^{(\tau_\alpha-1)} \exp\{-\alpha\lambda_\alpha\} \right\}, \\
\pi(\boldsymbol{\varphi}|\alpha, \gamma, \mathbf{v}^*) &\propto \prod_{i=1}^n \left\{ \phi(b_{v_i}^*) b_{v_i}^{*'} \exp\left\{-\frac{h_i^2}{2}\right\} \right\} \left\{ \exp\left\{-\frac{1}{2}(\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi)^\top \boldsymbol{\Psi}^{-1}(\boldsymbol{\varphi} - \boldsymbol{\mu}_\varphi)\right\} \right\}, \\
\pi(\gamma|\alpha, \boldsymbol{\varphi}, \mathbf{v}^*) &\propto \prod_{i=1}^n \phi(b_{v_i}^*) b_{v_i}^{*'} \exp\left\{-\frac{h_i^2}{2}\right\}. \tag{3.4.18}
\end{aligned}$$

Assim como no modelo probabilístico, é possível observar que as condicionais completas apresentadas em (3.4.14) e (3.4.18) não tem forma fechada, logo não é possível simular diretamente delas e nesse caso, se faz necessária a utilização de algoritmos tais como: *Metropolis-Hastings*, *slice sampling* e *rejeição adaptativa*.

No Capítulo 4, apresentaremos um estudo de simulação completo, sob os enfoques frequentista e bayesiano, a fim de avaliar o modelo regressão aqui proposto com respeito a recuperação dos parâmetros.

3.5 Análise de Diagnóstico e comparação de modelos

De acordo com Paula (2004) uma etapa importante na análise de um ajuste de um modelo de regressão (ou probabilístico) é a verificação de possíveis afastamentos das suposições feitas para o modelo, bem como a identificação de possíveis observações discrepantes (“*outliers*”) que possam causar algum tipo de interferência desproporcional ou inferencial nos resultados do ajuste.

Neste capítulo iremos apresentar a análise de resíduos do modelo log-BSNAC sob os enfoques frequentista e bayesiano, bem como estatísticas para comparação de modelos. Além disso, iremos desenvolver a análise de influência sob o ponto de vista bayesiano.

3.5.1 Análise Residual

A análise de resíduo é uma ferramenta importante na validação do modelo em estudo. É possível, através da análise de resíduos, verificar se existem observações discrepantes, assim como o afastamento das suposições feitas para o modelo.

Seguindo a metodologia proposta por Dunn e Smyth (1996), faremos uma análise residual para o modelo de regressão log-BSNAC usando os chamados resíduos quantis.

Seja $V \sim \text{SHNAC}(\alpha, \mathbf{x}_i^\top \boldsymbol{\varphi}, \sigma = 2, \gamma)$ uma variável aleatória contínua cuja distribuição acumulada é definida como:

$$F_V(v) = \Phi_\gamma(b_v(\alpha, \mathbf{x}_i^\top \boldsymbol{\varphi})),$$

em que, $\Phi_\gamma(\cdot)$ é a acumulada de uma variável aleatória NAC(0, 1, γ) e $b_v(\alpha, \mathbf{x}^\top \boldsymbol{\varphi})$ é tal como (3.4.3). Assim, podemos definir os resíduos quantis da seguinte forma

$$r_{q,i} = \Phi^{-1} \left\{ \Phi_\gamma \left[b_{v_i} \left(\hat{\alpha}, \mathbf{x}_i^\top \hat{\boldsymbol{\varphi}} \right) \right] \right\}, \quad (3.5.1)$$

em que $\widehat{(\cdot)}$ representa as estimativas obtidas pelos métodos frequentista e bayesiano. Portanto, sendo $\hat{\alpha}$ e $\hat{\boldsymbol{\varphi}}$ estimadores consistentes de α e $\boldsymbol{\varphi}$, respectivamente, temos que $r_{q,i}$ converge em distribuição para a Normal padrão.

A metodologia utilizada para a construção dos envelopes simulados que são apresentados ao longo do trabalho, está descrita no Apêndice A.

3.5.2 Estatísticas para comparação de modelos

Segundo Farias e Lemonte (2011), existe uma grande variedade de metodologias desenvolvidas para comparar modelos de regressão concorrentes, ajustados para um determinado conjunto de dados, e selecionar dentre esses modelos aquele que apresenta o melhor ajuste. Sob o ponto de vista frequentista e bayesiano consideramos algumas estatísticas para a comparação de modelos que podem ser calculadas a partir dos resultados obtidos via ECM e MCMC, respectivamente.

O critério de informação de Akaike (AIC) proposto por Akaike (1974) e o critério de informação bayesiano (BIC) proposto por Schwarz (1978) são medidas de comparação de modelos utilizadas tanto no contexto frequentista quanto no bayesiano. O AIC baseia-se na verossimilhança penalizada pelo número de parâmetros do modelo, enquanto que o BIC, além do número de parâmetros, pondera o tamanho amostral. Valores menores de AIC e BIC indicam modelos com melhores ajustes.

Seja $\boldsymbol{\theta} = (\alpha, \boldsymbol{\varphi}^\top, \gamma)^\top$ o vetor de parâmetros de interesse e $L(\mathbf{v}; \boldsymbol{\theta})$ a verossimilhança definida da seguinte maneira

$$L(\mathbf{v}; \boldsymbol{\theta}) = \prod_{i=1}^n 2\sigma_z \phi \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] b'_{v_i} \Phi \left\{ \lambda \left[\sigma_z \left(b_{v_i} + \frac{\mu_z}{\sigma_z} \right) \right] \right\}, \quad (3.5.2)$$

em que $b_{v_i} = \frac{2}{\alpha} \sinh \left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}}{2} \right)$, $b'_{v_i} = \frac{1}{\alpha} \cosh \left(\frac{v_i - \mathbf{x}_i^\top \boldsymbol{\varphi}}{2} \right)$; μ_z e σ_z são como em (2.1.4) e λ é tal como em (2.1.5).

No contexto frequentista, os critérios de seleção são definidos da seguinte forma:

$$AIC = -2 \log [L(\mathbf{v}; \boldsymbol{\theta})] + 2k \quad \text{e} \quad BIC = -2 \log [L(\mathbf{v}; \boldsymbol{\theta})] + k \log(n), \quad (3.5.3)$$

em que k é o número total de parâmetros do modelo e n é o número de observações.

Segundo Spiegelhalter et al. (2014), os critérios de seleção, no contexto bayesiano, são obtidos considerando a densidade a posteriori dos parâmetros do modelo. Definamos inicialmente $D(\boldsymbol{\theta}) = -2l(\boldsymbol{\theta}, \mathbf{v})$, em que $l(\boldsymbol{\theta}, \mathbf{v})$ é o logaritmo da verossimilhança dada em (3.4.11). Seja ainda $\boldsymbol{\theta}^{(m)}$, $m = 1, \dots, M$, o m -ésimo valor da amostra MCMC simulada válida, ou seja, a amostra MCMC obtida após o descarte do *burn-in* e do espaçamento entre os valores. Além disso,

seja $\widehat{\boldsymbol{\theta}}$ o vetor da esperança a posteriori de cada parâmetro considerando a amostra válida, e $\overline{D(\boldsymbol{\theta})} = \frac{1}{M} \sum_{i=1}^M D(\boldsymbol{\theta}^{(m)})$. Denotemos o desvio por $D(\bar{\boldsymbol{\theta}}) = -2l(\bar{\boldsymbol{\theta}}|\mathbf{v})$.

Os critérios EAIC e EBIC podem ser definidos da seguinte forma

$$EAIC = D(\bar{\boldsymbol{\theta}}) + 2k \quad \text{e} \quad EBIC = D(\bar{\boldsymbol{\theta}}) + k \log(n). \quad (3.5.4)$$

Definamos ainda o critério de informação do desvio (DIC) como

$$DIC = D(\bar{\boldsymbol{\theta}}) + 2p_D,$$

em que $p_D = \overline{D(\boldsymbol{\theta})} - D(\bar{\boldsymbol{\theta}})$.

O CPO (*conditional predictive ordinate*) associado à i -ésima observação é dado por

$$\widehat{CPO}_i = \left\{ \frac{1}{M} \sum_{i=1}^M \frac{1}{L(\boldsymbol{\theta}^{(m)}|\mathbf{v}_i)} \right\}^{-1}. \quad (3.5.5)$$

Para uma discussão mais detalhada a respeito da estatística CPO e sua aplicação na comparação de modelos, ver Gelfand et al. (1992). Como uma estatística resumo do CPO para todas as observações, usamos o LPML (*logarithm pseudo-marginal likelihood*), que é calculado da seguinte forma

$$LPML = \sum_{i=1}^N \ln(\widehat{CPO}_i).$$

Para mais detalhes a respeito do LPML, consultar Ibrahim et al. (2001). Como critérios para comparação de modelos utilizando as estatísticas aqui descritas, considera-se que: os menores valores do DIC, BIC e desvio, indicam o melhor ajuste. Por outro lado, os maiores valores do CPO, e por consequência do LPML, apontam o melhor modelo.

3.5.3 Medidas de Influência

Segundo Farias (2012), o diagnóstico de influência é uma das ferramentas mais utilizadas para a verificação da qualidade do ajuste do modelo proposto. Tais métodos visam a identificação de observações ou conjunto de observações que apresentam influência desproporcional nos resultados inferenciais e, por isto, merecem maior atenção. No contexto bayesiano, uma atenção especial tem sido dada para o desenvolvimento de técnicas de diagnóstico de influência utilizando a divergência de Kullback-Leibler (K-L). Cho et al. (2009) desenvolveram um diagnóstico de influência, para as distribuições conjuntas e marginais a posteriori, baseado na divergência de K-L, e apresentaram uma expressão simplificada para calcular tal medida de influência utilizando os resultados obtidos através da posteriori conjunta.

De acordo com Cho et al. (2009), seja $K(P, P_{(-i)})$ a divergência de K-L entre $P, P_{(-i)}$, em que P representa a distribuição a posteriori de $\boldsymbol{\theta}$ para os dados completos e $P_{(-i)}$ denota a distribuição a posteriori de $\boldsymbol{\theta}$ para os dados a menos da i -ésima observação. Assim, tem-se que

$$K(P, P_{(-i)}) = \int \pi(\boldsymbol{\theta}|\mathbf{v}) \ln \left\{ \frac{\pi(\boldsymbol{\theta}|\mathbf{v})}{\pi(\boldsymbol{\theta}|\mathbf{v}_{(-i)})} \right\} d\boldsymbol{\theta}, \quad (3.5.6)$$

em que $\mathbf{v}_{(-i)}$ corresponde a \mathbf{v} excluindo a i -ésima observação. Além disso, ainda segundo Cho et al. (2009), podemos estimar a divergência K-L, através dos resultados obtidos via método bayesiano, da seguinte forma

$$K(P, P_{(-i)}) = -\ln(\widehat{CPO}_i) + \sum_{i=1}^n \ln(L(\boldsymbol{\theta}^{(m)}|\mathbf{v}_i)), \quad (3.5.7)$$

em que \widehat{CPO}_i é dado pela equação ((3.5.5)).

Cho et al. (2009) também apresentam uma forma de calibrar a divergência $K(P, P_{(-i)})$. Esta calibração é obtida resolvendo para p_i a equação a seguir

$$K(P, P_{(-i)}) = K[Ber(1/2), Ber(p_i)] = \frac{1}{2} \log[4p_i(1 - p_i)], \quad (3.5.8)$$

em que $Ber(p)$ representa a distribuição Bernoulli com probabilidade de sucesso p . Note que considerar a expressão (3.5.8) como verdadeira, implica assumir que as distribuições a posteriori $\pi(\boldsymbol{\theta}|\mathbf{v})$ e $\pi(\boldsymbol{\theta}|\mathbf{v}_{(-i)})$ entre si estão tão próximas quanto as distribuições $Ber(1/2)$ e $Ber(p)$, o estão. Portanto, uma observação que induz a uma diferença grande entre essas distribuições à posteriori, irá influenciar na probabilidade de sucesso p_i , fazendo com que p_i seja igual a 1, quando a diferença entre as distribuições à posteriori for máxima. A calibração é dada por

$$p_i = \frac{1}{2} \left[1 + \sqrt{1 - \exp\{-2K(P, P_{(-i)})\}} \right].$$

Essa expressão implica em $0,5 \leq p_i \leq 1$. Cho et al. (2009) sugerem considerar a i -ésima observação como influente quando p_i for muito maior que 0,5.

Após a identificação das observações influentes, é de interesse avaliar o impacto que a retirada dessas observações exerce nas estimativas dos parâmetros em cada um dos modelos em questão. O impacto percentual pode ser calculado a partir da seguinte expressão

$$RC = \frac{|\hat{\theta}_{j(-i)} - \hat{\theta}_j|}{\hat{\theta}_j} \times 100, i = 1, \dots, n, \quad (3.5.9)$$

em que $\hat{\theta}_{j(-i)}$ denota a estimativa bayesiana após a retirada da i -ésima observação.

Nos Capítulos 4 e 5, ilustraremos a metodologia aqui proposta usando tanto os dados provenientes de estudos de simulação, quanto conjuntos de dados reais, respectivamente.

Capítulo 4

Estudo de simulação

São apresentados, neste capítulo, estudos de simulação para ilustrar a metodologia proposta nos Capítulos 2 e 3.

Inicialmente, apresentaremos o estudo de convergência do algoritmo MCMC desenvolvido para o modelo probabilístico BSNAC, em seguida, ilustraremos o comportamento dos algoritmos ECM e MCMC. É de interesse comparar as estimativas obtidas através dos dois métodos, bem como avaliar o impacto que diferentes tamanhos amostrais exercem na precisão das estimativas dos parâmetros. Os estudos de simulação referentes ao modelo probabilístico serão apresentados na Seção 4.1.

Além disso, na Seção 4.2, apresentaremos o estudo de convergência do algoritmo MCMC desenvolvido para o modelo log-BSNAC, e em seguida iremos comparar as estimativas dos parâmetros obtidas sob os pontos de vista frequentista e bayesiano. Ainda para o modelo de regressão log-BSNAC, apresentaremos estudos de simulação para ilustrar a análise de resíduos, a análise de influência e a metodologia apresentada a respeito dos critérios de comparação de modelos.

Para os estudos em questão consideramos 100 réplicas de Monte Carlo. Na abordagem bayesiana utilizou-se a esperança à posteriori como estimador, sendo que os algoritmos MCMC foram usados para amostrar das posteriores, enquanto que as estimativas de máxima verossimilhança foram obtidas utilizando-se os algoritmos ECM propostos nos Capítulos 2 e 3. A partir dessas 100 réplicas, obtivemos as seguinte quantidades: média, variância, vício, raiz quadrada do erro quadrático médio (REQM) e o valor absoluto do vício relativo (da sigla em inglês AVRBAbsolute Value of Relative Bias). Considere que $\hat{\theta}_r$ é a estimativa obtida na i -ésima réplica para cada parâmetro e θ representa o verdadeiro valor do parâmetro. Então, define-se cada uma das estatísticas como:

- **Média**-Média das R=100 réplicas:

$$\bar{\hat{\theta}}_R = \sum_{r=1}^R \frac{\hat{\theta}_r}{R}$$

- **Variância**- Variância das R=100 réplicas:

$$\widehat{Var}_{\theta} = Var(\bar{\hat{\theta}}_R) = \sum_{r=1}^R \frac{(\hat{\theta}_r - \bar{\hat{\theta}}_R)^2}{R - 1}.$$

- **Vício**-Diferença entre a média e o verdadeiro valor do parâmetro:

$$\widehat{\text{Vício}}_{\theta} = \text{Vício}(\bar{\theta}_R - \theta).$$

- **REQM**-Raiz quadrada da soma do quadrado do vício com a variância:

$$\widehat{\text{REQM}}_{\theta} = \text{REQM}(\bar{\theta}_R) = \sqrt{\widehat{\text{Vício}}_{\theta}^2 + \widehat{\text{Var}}_{\theta}}.$$

- **AVBR**-Divisão entre o valor absoluto do vício e o valor absoluto do verdadeiro valor do parâmetro:

$$\widehat{\text{AVBR}}_{\theta} = \frac{|\widehat{\text{Vício}}_{\theta}|}{|\theta|}.$$

Vale ressaltar que quando o verdadeiro valor do parâmetro for igual a zero, o AVBR não será calculado.

4.1 Estudo de simulação I

Nesta seção são apresentados resultados de estudos de simulação, referentes ao modelo probabilístico BSNAC, para verificar a convergência do algoritmo MCMC e ilustrar o comportamento dos algoritmos ECM e MCMC.

Para comparar as estimativas frequentistas e bayesianas, foram considerados um total de 30 diferentes cenários, resultantes da combinação de três diferentes tamanhos amostrais, $n = 10$, $n = 50$ e $n = 200$, com os seguintes valores fixados para os parâmetros de interesse: $\alpha \in (0, 5; 1, 5)$, $\beta = 1$ e $\gamma \in (-0, 67; -0, 45; 0; 0, 45; 0, 67)$. Os tamanhos amostrais aqui propostos tem por objetivo verificar se há melhoria na estimação dos parâmetros de interesse consoante o aumento do tamanho amostral. Os valores fixados para α e β , foram escolhidos com intuito de obter modelos com diferentes formas e pequena variabilidade. Os valores fixados para γ foram escolhidos com o intuito de obter modelos com assimetria forte, moderada e nula.

4.1.1 Estudo de Convergência

Nesta seção avaliaremos o algoritmo MCMC proposto para o modelo probabilístico no Capítulo 2 em termos da convergência: avaliando a magnitude das autocorrelações e, também, a sensibilidade com relação à escolha de valores iniciais e tamanhos amostrais. Utilizamos o conjunto de prioris descrito anteriormente, mais especificamente, para o parâmetro de forma $\alpha > 0$ adotamos uma priori Gama(1; 0,8) com massa concentrada no intervalo (0, 3), visto que tais valores são mais comumente observados na literatura. Para o parâmetro de escala, $\beta > 0$, adotamos uma priori Gama (0,001; 0,001) pouco informativa. Por fim, para o parâmetro de assimetria,

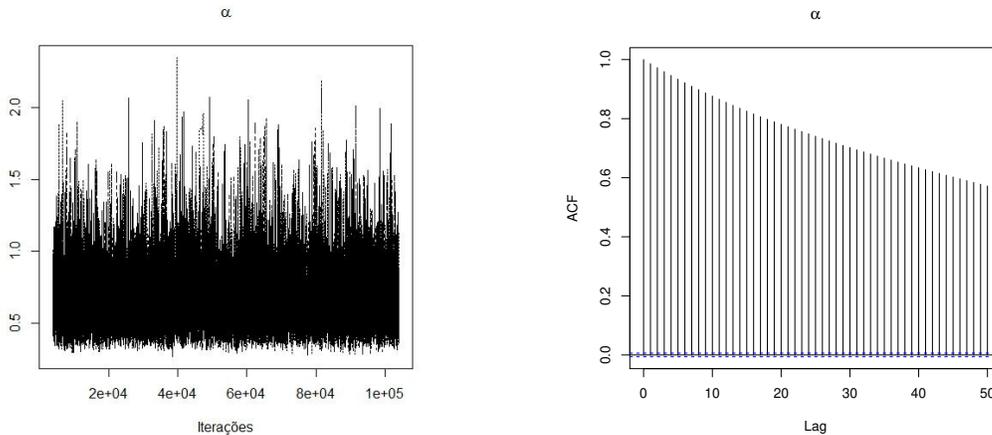
$\gamma \in (-0,99527; 0,99527)$, adotamos uma priori Uniforme contemplando todo o espaço paramétrico. Para ambos os tamanhos amostrais considerados nesse estudo, $n=10$ e $n=200$, fixamos $\alpha = 0,5$, $\beta = 1$, $\gamma = 0,67$.

O estudo consiste em iniciar paralelamente 3 cadeias. Em cada uma das cadeias o número de iterações foi igual a 104.000. Para auxílio da obtenção das estatísticas e gráficos de monitoramento da convergência das cadeias foi utilizado o pacote *coda* (Plummer et. al., 2006) do software R.

Para verificar convergência do algoritmo, utilizamos alguns métodos comumente considerados na literatura. Um deles é a inspeção visual através de gráficos das autocorrelações, e também da observação da trajetória de diferentes cadeias. Utilizamos também os critérios de Geweke, proposto em Geweke (1991) e o critério de Gelman e Rubin, sugerido por Gelman e Rubin (1992). O critério de Geweke sugere a verificação da convergência através de testes de médias, em intervalos diferentes, após o período de burn-in do algoritmo. Se a convergência foi alcançada, os comportamentos nos diferentes intervalos devem ser parecidos. Já o critério de Gelman e Rubin, avalia a convergência utilizando cadeias paralelas. O teste consiste em comparar a variância entre e dentro das cadeias. Gamerman e Lopes (2006) descrevem esses critérios com maiores detalhes.

As Figuras 4.1 - 4.3 dizem respeito ao estudo de convergência considerando um tamanho de amostra $n=10$. Na Figura 4.1 temos o gráfico de trajetórias das 3 cadeias completas (com as 104.000 iterações), assim como as autocorrelações obtidas. É possível observar uma sobreposição das 3 cadeias, e após o período de aquecimento de 4.000 primeiras iterações, há indícios de que diferentes valores iniciais não alteram significativamente os resultados. É possível observar ainda a presença de autocorrelação.

Considerando a estatística de Geweke, em todas as 3 cadeias dos parâmetros α , β e γ , a grande maioria dos valores ficaram dentro dos limites considerados aceitáveis, o que indica convergência das cadeias. Por simplicidade tais gráficos não serão apresentados.



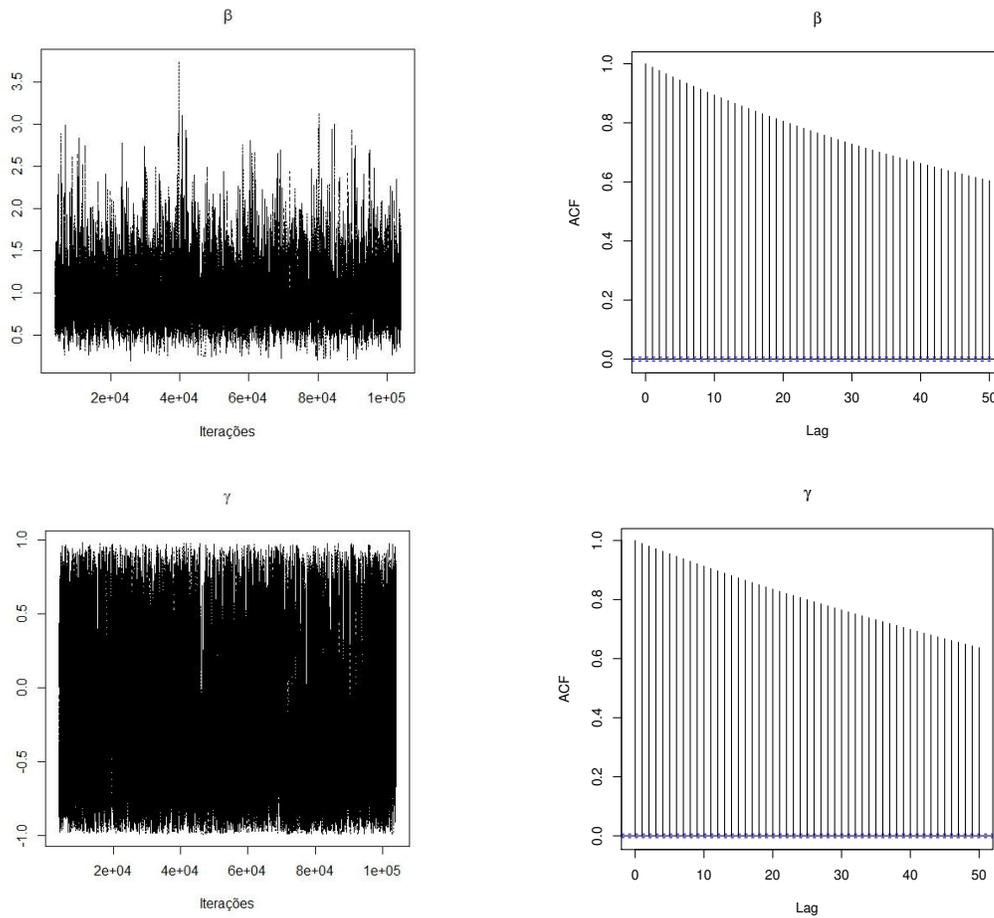


Figura 4.1: Gráficos de trajetórias e de autocorrelação dos parâmetros do modelo probabilístico BSNAC ($n=10$).

Na Figura 4.2 observa-se os gráficos referentes a estatística de Gelman-Rubin. Quanto mais perto do valor unitário, significa que as 3 cadeias convergiram conjuntamente. Há indícios de que de fato as cadeias convergem.

Na Figura 4.3 temos os correlogramas obtidas da amostra final, isto é, com as 1.000 estimativas selecionadas da cadeia de Markov, após a retirada das 4.000 primeiras iterações (burn-in) e do espaçamento igual a 100. Nota-se que os valores na sua grande maioria estão todos abaixo de 0,20, indicando assim que as amostras são aproximadamente não correlacionadas.

Nas Figuras 4.4 - 4.6 é possível visualizar os resultados do estudo de convergência considerando um tamanho de amostra $n=200$. Na Figura 4.4 temos o gráfico de trajetórias das 3 cadeias completas (com as 104.000 iterações), assim como as autocorrelações obtidas. É possível observar uma sobreposição das 3 cadeias.

Considerando a estatística de Geweke, em todas as 3 cadeias dos parâmetros α , β e γ , a grande maioria dos valores ficaram dentro dos limites considerados aceitáveis, o que indica convergência das cadeias.

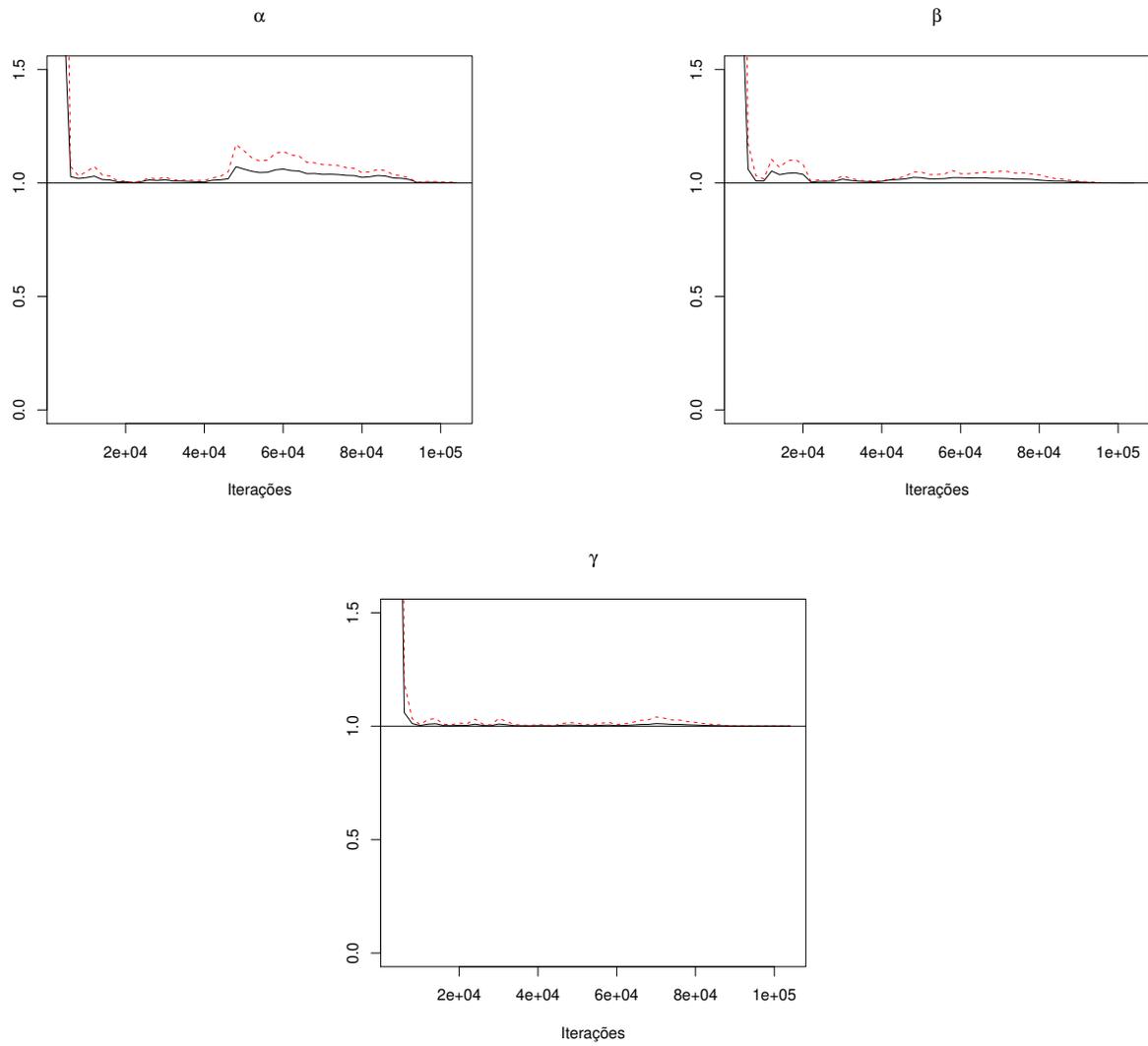


Figura 4.2: Estatística de Gelman-Rubin para o modelo probabilístico BSNAC. (n=10)

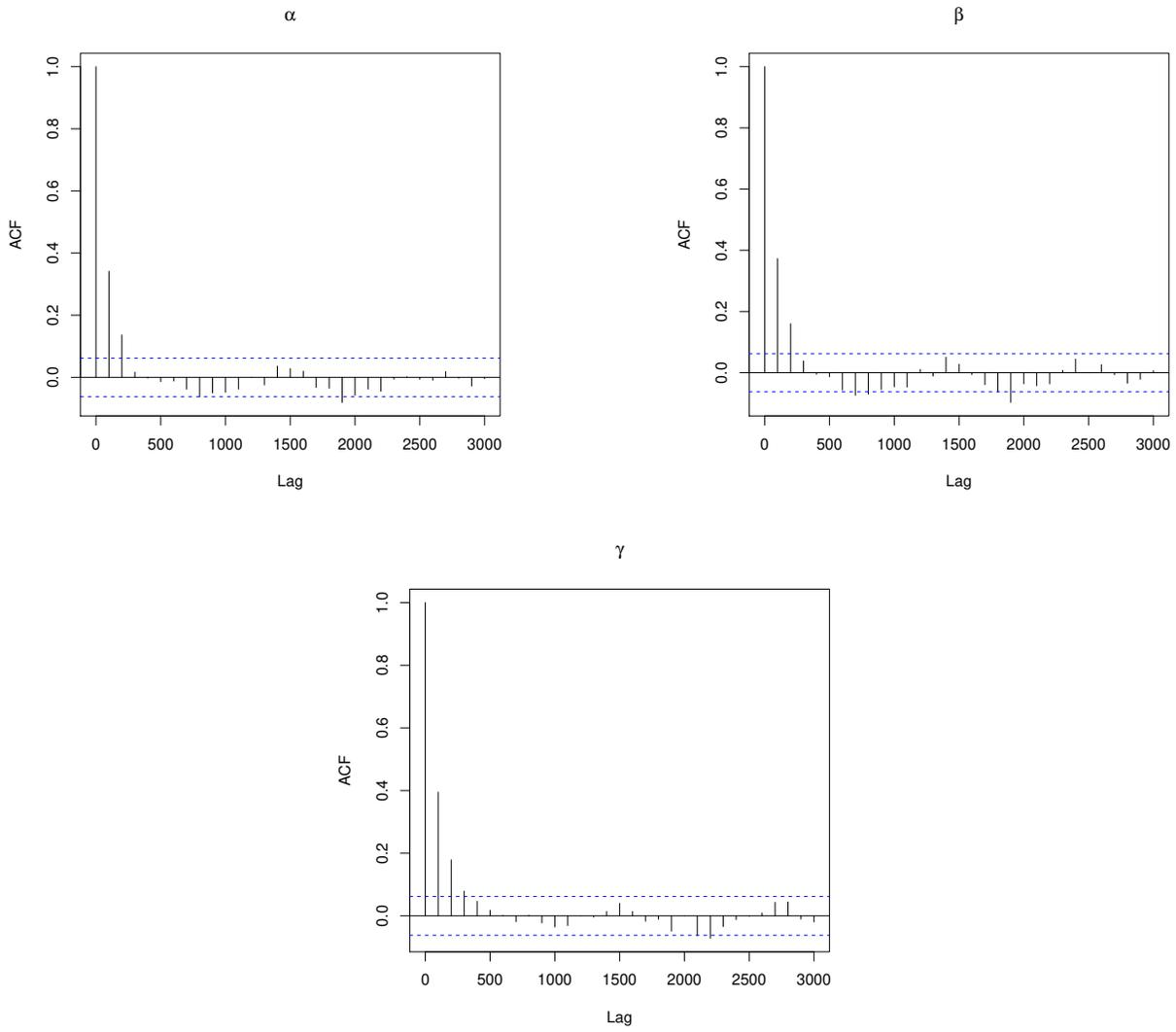


Figura 4.3: Gráficos de autocorrelação para o modelo probabilístico BSNAC. ($n=10$)

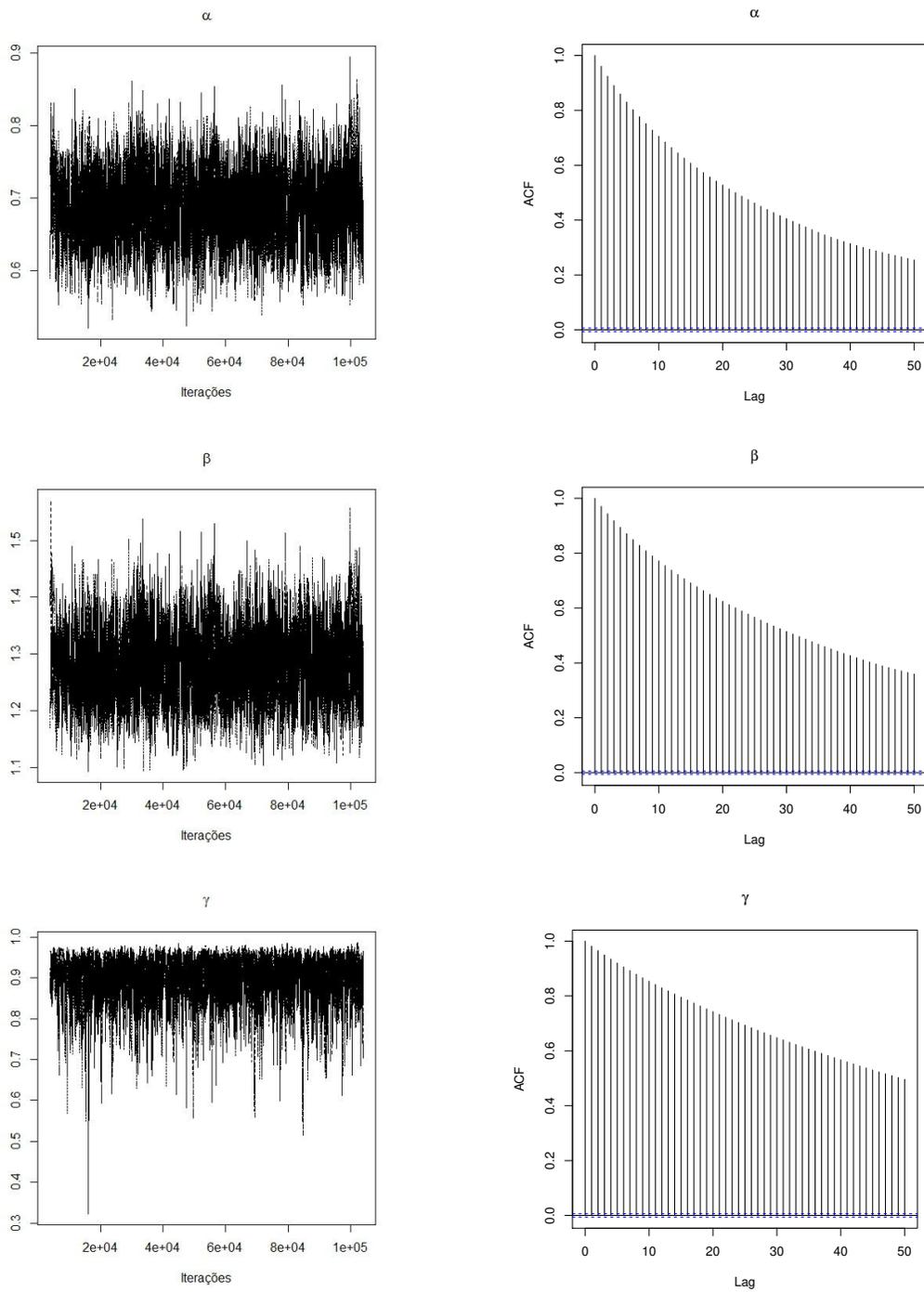


Figura 4.4: Gráficos de trajetórias e de autocorrelação dos parâmetros do modelo probabilístico BSNAC ($n=200$).

Na Figura 4.5 observa-se os gráficos referentes a estatística de Gelman-Rubin. Há indícios de que de fato as cadeias convergem e que tal convergência

Na Figura 4.6 temos os correlogramas obtidas da amostra final, isto é, com as 1.000 estimativas selecionadas da cadeia de Markov, após a retirada das 4000 primeiras iterações (burn-in) e do espaçamento igual a 100. Nota-se que os valores estão na sua grande maioria abaixo de 0,20, indicando assim que as amostras são aproximadamente não correlacionadas.

Com o aumento da amostra é possível notar que os valores apresentados na Figura 4.6 são ainda menores dos que aqueles mostrados na Figura 4.3. Além disso, consoante ao aumento do tamanho da amostra, é possível perceber que a convergência das cadeias é melhor atingida.

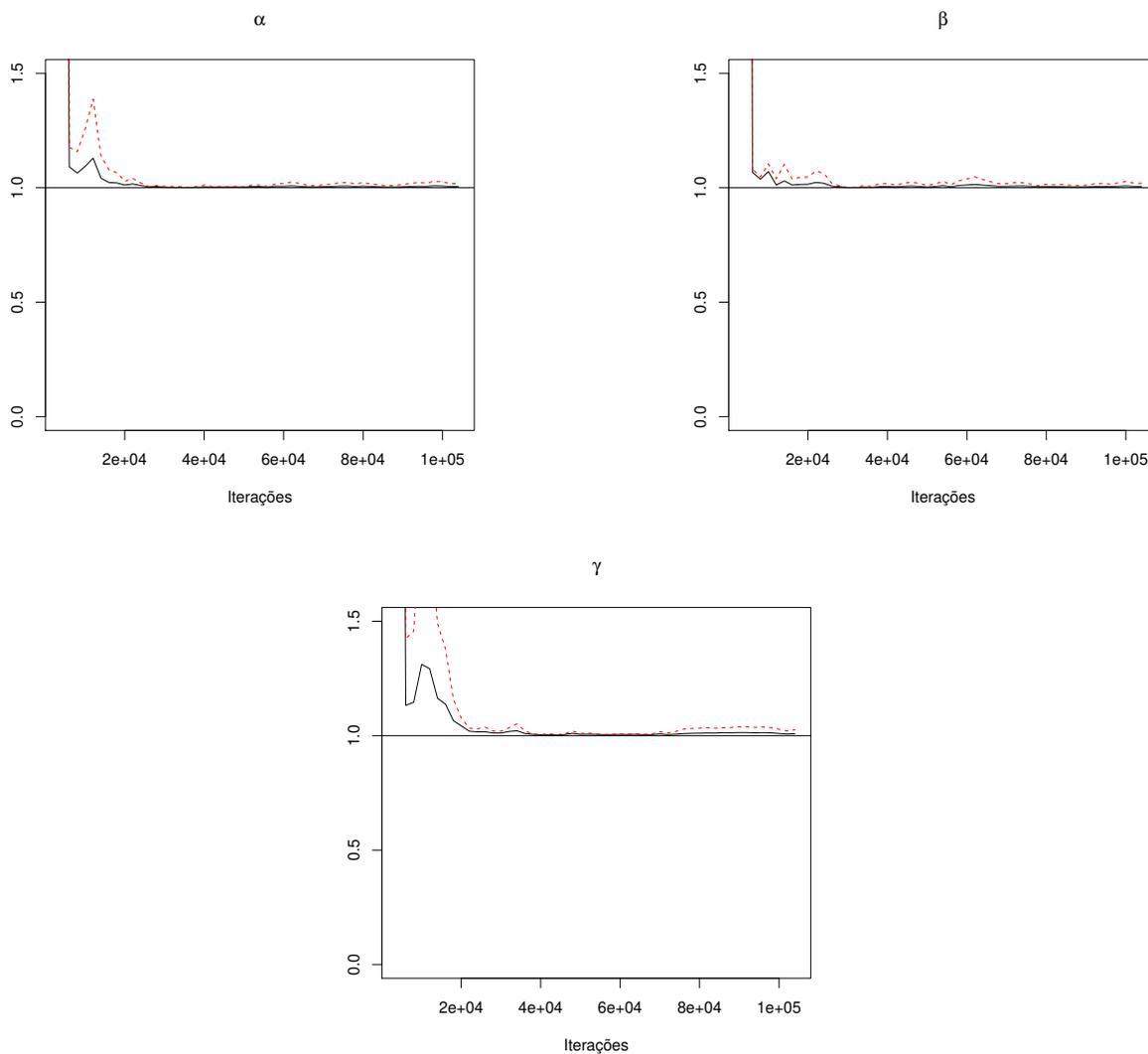


Figura 4.5: Estatística de Gelman-Rubin para o modelo probabilístico BSNAC. ($n=200$)

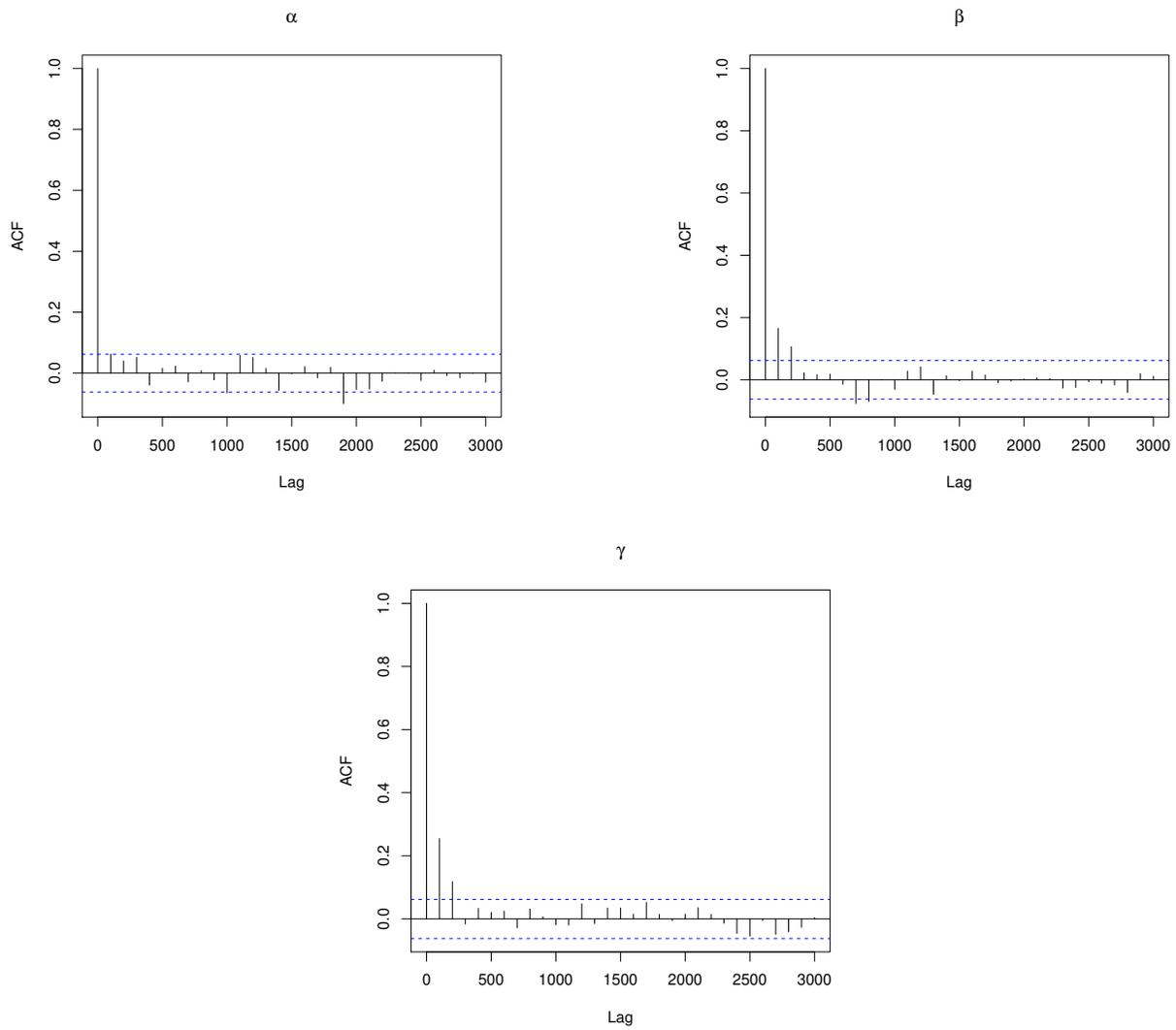


Figura 4.6: Gráficos de autocorrelação para o modelo probabilístico BSNAC. (n=200)

4.1.2 Comparação entre as estimativas frequentistas e bayesianas

Nesta seção, apresentaremos os resultados obtidos através das metodologias frequentista e bayesiana, desenvolvidas no Capítulo 2, considerando os cenários já descritos anteriormente. O algoritmo MCMC utilizado para a obtenção dos resultados bayesianos, foi implementado usando tanto a verossimilhança completa, quanto a verossimilhança original.

As Tabelas 4.1 - 4.5 dizem respeito aos resultados obtidos quando fixamos $\alpha = 0,5$, já as Tabelas 4.6 - 4.10 referem-se aos resultados obtidos quando fixamos $\alpha = 1,5$. Para ambos os conjuntos de tabelas, fixamos $\beta = 1$ e variamos o γ de acordo com os valores apresentados no início da Seção 4.1. A fim de facilitar a compreensão, utilizaremos ($< 0,001$) para representar valores positivos e ($> -0,001$) para denotar valores negativos, quando estiverem bem próximos de zero. Além disso, nomearemos como “Bayesiano 1” os resultados bayesianos obtidos a partir da verossimilhança completa, e como “Bayesiano 2” os que foram obtidos a partir da verossimilhança original.

De uma forma geral, analisando as tabelas em questão, nota-se que à medida que aumentamos o tamanho da amostra, as estimativas obtidas tanto pelo método frequentista, quanto pelo método bayesiano tendem a se aproximar dos verdadeiros valores dos parâmetros. É possível notar ainda, que as estimativas obtidas para α tanto pelo método frequentista, quanto pelo método “Bayesiano 2”, aproximam-se dos verdadeiros valores do parâmetro. Em particular, o método “Bayesiano 2” fornece estimativas com valores menores para o vício, REQM e AVBR. As estimativas obtidas via método “Bayesiano 1” ficaram em desvantagem frente as estimativas obtidas através dos outros dois métodos, provavelmente pelo fato de estarmos usando a estrutura hierárquica do modelo probabilístico.

Em relação à β , se observarmos as Tabelas 4.1 - 4.5, é possível notar que as estimativas obtidas pelos métodos frequentista e “Bayesiano 2”, à medida que aumentamos o tamanho da amostra, são próximas do verdadeiro valor. Por outro lado, se observarmos as Tabelas 4.6 - 4.10, é possível perceber que o método “Bayesiano 2” fornece estimativas melhores do que os métodos frequentista e “Bayesiano 1”.

Por fim, para estimar γ , é possível notar que o método bayesiano, tanto usando a verossimilhança completa, quanto a verossimilhança original, na maior parte dos cenários, apresenta estimativas com menores vício, REQM e AVBR, e mais próximas dos verdadeiros valores fixados para γ .

Tabela 4.1: Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = -0,67$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,700	< 0,001	0,200	0,200	0,400
		Bayesiano 2	0,578	0,001	0,078	0,082	0,156
		M.V.	0,520	0,071	0,020	0,267	0,040
	50	Bayesiano 1	0,634	0,001	0,134	0,136	0,268
		Bayesiano 2	0,511	< 0,001	0,011	0,015	0,021
		M.V.	0,498	0,001	-0,002	0,033	0,004
	200	Bayesiano 1	0,649	< 0,001	0,149	0,150	0,297
		Bayesiano 2	0,502	< 0,001	0,002	0,005	0,004
		M.V.	0,490	< 0,001	-0,010	0,012	0,019
β	10	Bayesiano 1	0,805	< 0,001	-0,195	0,196	0,195
		Bayesiano 2	1,004	< 0,001	0,004	0,021	0,004
		M.V.	1,105	0,214	0,105	0,474	0,105
	50	Bayesiano 1	0,856	0,005	-0,144	0,161	0,144
		Bayesiano 2	0,997	< 0,001	-0,003	0,009	0,003
		M.V.	1,039	0,018	0,039	0,140	0,039
	200	Bayesiano 1	0,998	0,010	-0,003	0,101	0,001
		Bayesiano 2	0,999	< 0,001	-0,001	0,005	0,001
		M.V.	0,997	< 0,001	-0,003	0,004	0,003
γ	10	Bayesiano 1	-0,139	0,065	0,531	0,589	0,793
		Bayesiano 2	-0,182	0,054	0,488	0,540	0,728
		M.V.	-0,603	0,028	0,067	0,179	0,100
	50	Bayesiano 1	-0,518	0,068	0,152	0,302	0,226
		Bayesiano 2	-0,505	0,049	0,165	0,276	0,247
		M.V.	-0,569	0,012	0,101	0,148	0,150
	200	Bayesiano 1	-0,713	0,013	-0,043	0,122	0,064
		Bayesiano 2	-0,601	0,015	0,069	0,141	0,103
		M.V.	-0,523	0,002	0,147	0,153	0,220

Tabela 4.2: Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = -0,45$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,703	< 0,001	0,203	0,203	0,405
		Bayesiano 2	0,583	< 0,001	0,083	0,085	0,166
		M.V.	0,528	0,040	0,028	0,203	0,055
	50	Bayesiano 1	0,622	0,001	0,122	0,124	0,243
		Bayesiano 2	0,515	< 0,001	0,015	0,017	0,030
		M.V.	0,582	0,197	0,082	0,451	0,165
	200	Bayesiano 1	0,612	0,001	0,112	0,116	0,225
		Bayesiano 2	0,504	< 0,001	0,004	0,005	0,008
		M.V.	0,509	0,001	0,009	0,037	0,018
β	10	Bayesiano 1	0,841	0,010	0,020	0,163	0,159
		Bayesiano 2	1,012	< 0,001	0,012	0,023	0,012
		M.V.	1,143	0,798	0,143	0,905	0,143
	50	Bayesiano 1	0,901	0,009	-0,099	0,137	0,099
		Bayesiano 2	0,999	< 0,001	-0,001	0,006	0,001
		M.V.	1,189	1,125	0,189	1,077	0,189
	200	Bayesiano 1	1,020	0,001	-0,159	0,101	0,020
		Bayesiano 2	1,000	<0,001	> -0,001	0,003	<0,001
		M.V.	0,997	<0,001	-0,003	0,007	0,003
γ	10	Bayesiano 1	-0,078	0,063	0,372	0,449	0,827
		Bayesiano 2	-0,085	0,054	0,365	0,433	0,811
		M.V.	-0,609	0,034	-0,159	0,244	0,353
	50	Bayesiano 1	-0,353	0,094	0,097	0,322	0,216
		Bayesiano 2	-0,356	0,069	0,094	0,278	0,209
		M.V.	-0,555	0,017	-0,105	0,168	0,234
	200	Bayesiano 1	-0,505	0,027	-0,055	0,173	0,123
		Bayesiano 2	-0,394	0,024	0,056	0,165	0,124
		M.V.	-0,568	0,029	-0,118	0,208	0,262

Tabela 4.3: Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = 0$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,703	< 0,001	0,203	0,204	0,406
		Bayesiano 2	0,581	< 0,001	0,081	0,084	0,163
		M.V.	0,474	< 0,001	-0,026	0,026	0,051
	50	Bayesiano 1	0,609	< 0,001	0,109	0,111	0,218
		Bayesiano 2	0,519	< 0,001	0,019	0,020	0,039
		M.V.	0,495	< 0,001	-0,005	0,006	0,010
	200	Bayesiano 1	0,564	< 0,001	0,064	0,066	0,129
		Bayesiano 2	0,506	< 0,001	0,006	0,006	0,012
		M.V.	0,499	< 0,001	-0,001	0,002	0,002
β	10	Bayesiano 1	1,041	0,010	0,041	0,108	0,041
		Bayesiano 2	1,020	< 0,001	0,020	0,028	0,020
		M.V.	0,999	< 0,001	-0,001	0,008	0,001
	50	Bayesiano 1	1,029	0,013	0,029	0,116	0,029
		Bayesiano 2	1,003	< 0,001	0,003	0,005	0,003
		M.V.	1,000	< 0,001	<0,001	0,005	< 0,001
	200	Bayesiano 1	1,005	0,007	0,005	0,087	0,005
		Bayesiano 2	1,001	< 0,001	0,001	0,002	0,001
		M.V.	1,000	< 0,001	< 0,001	0,002	< 0,001
γ	10	Bayesiano 1	-0,026	0,057	-0,026	0,240	-
		Bayesiano 2	0,017	0,066	0,017	0,257	-
		M.V.	> -0,001	<0,001	>-0,001	>-0,001	-
	50	Bayesiano 1	0,034	0,101	0,034	0,319	-
		Bayesiano 2	-0,004	0,091	-0,004	0,302	-
		M.V.	-0,009	0,012	-0,009	0,108	-
	200	Bayesiano 1	-0,012	0,044	-0,012	0,209	-
		Bayesiano 2	-0,009	0,033	-0,009	0,183	-
		M.V.	0,004	0,011	0,004	0,104	-

Tabela 4.4: Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = 0,45$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,704	< 0,001	0,204	0,205	0,408
		Bayesiano 2	0,580	< 0,001	0,080	0,082	0,159
		M.V.	0,487	0,001	-0,013	0,038	0,025
	50	Bayesiano 1	0,618	<0,001	0,118	0,120	0,237
		Bayesiano 2	0,516	<0,001	0,016	0,017	0,031
		M.V.	0,503	0,001	0,003	0,025	0,005
	200	Bayesiano 1	0,612	0,001	0,112	0,114	0,223
		Bayesiano 2	0,504	<0,001	0,004	0,005	0,008
		M.V.	0,505	0,001	0,005	0,026	0,009
β	10	Bayesiano 1	1,194	0,008	0,194	0,200	0,194
		Bayesiano 2	1,023	<0,001	0,023	0,030	0,023
		M.V.	0,986	0,013	-0,014	0,115	0,014
	50	Bayesiano 1	1,140	0,009	0,140	0,168	0,140
		Bayesiano 2	1,006	<0,001	0,006	0,009	0,006
		M.V.	0,988	0,006	-0,012	0,077	0,012
	200	Bayesiano 1	1,104	0,002	0,104	0,136	0,104
		Bayesiano 2	1,001	<0,001	0,001	0,003	0,001
		M.V.	1,001	0,001	0,001	0,038	0,001
γ	10	Bayesiano 1	0,127	0,048	-0,323	0,390	0,717
		Bayesiano 2	0,069	0,067	-0,381	0,460	0,847
		M.V.	0,554	0,024	0,104	0,187	0,232
	50	Bayesiano 1	0,342	0,074	-0,108	0,293	0,240
		Bayesiano 2	0,354	0,069	-0,096	0,279	0,213
		M.V.	0,545	0,015	0,095	0,156	0,211
	200	Bayesiano 1	0,501	0,027	0,051	0,172	0,113
		Bayesiano 2	0,405	0,022	-0,045	0,156	0,101
		M.V.	0,527	0,019	0,077	0,157	0,170

Tabela 4.5: Resultados do estudo de simulação - $\alpha = 0,5$, $\beta = 1$, $\gamma = 0,67$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,701	<0,001	0,201	0,201	0,402
		Bayesiano 2	0,574	0,001	0,074	0,078	0,148
		M.V.	0,491	0,002	-0,009	0,043	0,018
	50	Bayesiano 1	0,631	0,001	0,131	0,133	0,261
		Bayesiano 2	0,511	<0,001	0,011	0,015	0,022
		M.V.	0,499	0,003	-0,001	0,053	0,002
	200	Bayesiano 1	0,650	0,001	0,150	0,151	0,299
		Bayesiano 2	0,503	<0,001	0,003	0,005	0,005
		M.V.	0,490	<0,001	-0,010	0,014	0,020
β	10	Bayesiano 1	1,246	0,009	0,246	0,248	0,246
		Bayesiano 2	1,031	<0,001	0,031	0,038	0,031
		M.V.	0,996	0,012	-0,004	0,109	0,004
	50	Bayesiano 1	1,191	0,004	0,191	0,202	0,191
		Bayesiano 2	1,008	<0,001	0,008	0,012	0,008
		M.V.	0,991	0,005	-0,009	0,070	0,009
	200	Bayesiano 1	1,105	0,001	0,105	0,142	0,105
		Bayesiano 2	1,002	<0,001	0,002	0,005	0,002
		M.V.	1,003	<0,001	0,003	0,005	0,003
γ	10	Bayesiano 1	0,131	0,058	-0,539	0,590	0,805
		Bayesiano 2	0,167	0,065	-0,503	0,564	0,751
		M.V.	0,592	0,027	-0,078	0,183	0,116
	50	Bayesiano 1	0,498	0,042	-0,172	0,268	0,257
		Bayesiano 2	0,491	0,046	-0,179	0,279	0,268
		M.V.	0,543	0,008	-0,127	0,154	0,189
	200	Bayesiano 1	0,719	0,016	0,049	0,135	0,073
		Bayesiano 2	0,613	0,014	-0,057	0,130	0,085
		M.V.	0,527	0,002	-0,143	0,149	0,213

Tabela 4.6: Resultados do estudo de simulação - $\alpha = 1, 5, \beta = 1, \gamma = -0, 67$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	1,945	0,005	0,445	0,445	0,296
		Bayesiano 2	1,631	0,001	0,131	0,136	0,087
		M.V.	1,681	0,182	0,181	0,458	0,038
	50	Bayesiano 1	1,823	0,004	0,323	0,329	0,215
		Bayesiano 2	1,522	<0,001	0,022	0,029	0,015
		M.V.	1,702	0,086	0,202	0,356	0,135
	200	Bayesiano 1	1,823	0,001	0,323	0,330	0,215
		Bayesiano 2	1,505	<0,001	0,005	0,011	0,003
		M.V.	1,556	0,177	0,056	0,430	0,121
β	10	Bayesiano 1	0,641	0,125	0,034	0,355	0,359
		Bayesiano 2	1,026	0,012	0,026	0,114	0,026
		M.V.	1,121	2,344	0,121	1,536	0,121
	50	Bayesiano 1	0,690	0,008	-0,310	0,323	0,310
		Bayesiano 2	0,985	0,001	-0,015	0,037	0,015
		M.V.	1,281	0,946	0,281	1,012	0,281
	200	Bayesiano 1	1,034	0,002	-0,359	0,362	0,034
		Bayesiano 2	0,996	<0,001	-0,004	0,022	0,004
		M.V.	1,458	0,596	0,458	0,898	0,458
γ	10	Bayesiano 1	-0,158	0,088	0,512	0,591	0,764
		Bayesiano 2	-0,166	0,064	0,504	0,564	0,752
		M.V.	-0,509	0,398	0,161	0,651	0,240
	50	Bayesiano 1	-0,440	0,025	0,230	0,279	0,343
		Bayesiano 2	-0,485	0,051	0,185	0,292	0,276
		M.V.	-0,815	0,053	-0,145	0,272	0,217
	200	Bayesiano 1	-0,508	0,020	0,162	0,215	0,242
		Bayesiano 2	-0,596	0,015	0,074	0,144	0,111
		M.V.	-0,817	0,058	-0,147	0,283	0,220

Tabela 4.7: Resultados do estudo de simulação - $\alpha = 1, 5$, $\beta = 1$, $\gamma = -0, 45$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	1,944	0,005	0,444	0,444	0,296
		Bayesiano 2	1,639	0,001	0,139	0,143	0,093
		M.V.	1,564	0,214	0,139	0,467	0,043
	50	Bayesiano 1	1,785	0,004	0,285	0,291	0,190
		Bayesiano 2	1,532	<0,001	0,032	0,036	0,022
		M.V.	1,722	0,249	0,222	0,547	0,148
	200	Bayesiano 1	1,735	0,001	0,235	0,245	0,157
		Bayesiano 2	1,508	<0,001	0,008	0,010	0,005
		M.V.	1,639	0,047	0,064	0,257	0,093
β	10	Bayesiano 1	0,706	0,126	0,114	0,373	0,114
		Bayesiano 2	1,052	0,013	0,052	0,124	0,052
		M.V.	1,471	12,511	0,471	3,568	0,471
	50	Bayesiano 1	0,791	0,038	-0,209	0,285	0,209
		Bayesiano 2	0,997	0,001	-0,003	0,024	0,003
		M.V.	1,554	2,874	0,554	1,784	0,554
	200	Bayesiano 1	1,114	0,006	-0,294	0,303	0,294
		Bayesiano 2	0,998	<0,001	-0,002	0,013	0,002
		M.V.	1,393	0,547	0,393	0,837	0,393
γ	10	Bayesiano 1	-0,089	0,084	0,361	0,463	0,802
		Bayesiano 2	-0,120	0,061	0,330	0,412	0,732
		M.V.	-0,389	0,378	0,061	0,618	0,136
	50	Bayesiano 1	-0,301	0,051	0,149	0,271	0,331
		Bayesiano 2	-0,313	0,068	0,137	0,294	0,304
		M.V.	-0,700	0,158	-0,250	0,470	0,556
	200	Bayesiano 1	-0,337	0,019	0,113	0,177	0,250
		Bayesiano 2	-0,390	0,024	0,060	0,166	0,134
		M.V.	-0,742	0,029	-0,292	0,337	0,648

Tabela 4.8: Resultados do estudo de simulação - $\alpha = 1,5$, $\beta = 1$, $\gamma = 0$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	1,945	0,001	0,445	0,446	0,297
		Bayesiano 2	1,647	0,001	0,147	0,151	0,098
		M.V.	1,420	<0,001	-0,080	0,080	0,053
	50	Bayesiano 1	1,740	0,002	0,240	0,245	0,160
		Bayesiano 2	1,541	<0,001	0,041	0,042	0,028
		M.V.	1,485	< 0,001	-0,015	0,016	0,010
	200	Bayesiano 1	1,624	0,001	0,124	0,128	0,083
		Bayesiano 2	1,513	<0,001	0,013	0,013	0,008
		M.V.	1,497	<0,001	-0,003	0,006	0,002
β	10	Bayesiano 1	1,237	0,170	0,237	0,475	0,237
		Bayesiano 2	1,094	0,011	0,094	0,140	0,094
		M.V.	1,011	0,015	0,011	0,122	0,011
	50	Bayesiano 1	1,060	0,060	0,060	0,252	0,060
		Bayesiano 2	1,018	<0,001	0,018	0,026	0,018
		M.V.	0,998	0,003	-0,002	0,050	0,002
	200	Bayesiano 1	1,027	0,033	0,027	0,183	0,027
		Bayesiano 2	1,005	<0,001	0,005	0,008	0,005
		M.V.	1,000	0,001	>-0,001	0,029	<0,001
γ	10	Bayesiano 1	> -0,001	0,092	> -0,001	0,303	-
		Bayesiano 2	-0,010	0,055	-0,010	0,234	-
		M.V.	> -0,001	<0,001	> -0,001	<0,001	-
	50	Bayesiano 1	-0,017	0,059	-0,017	0,243	-
		Bayesiano 2	0,007	0,070	0,007	0,265	-
		M.V.	-0,003	0,016	-0,003	0,128	-
	200	Bayesiano 1	-0,005	0,016	-0,005	0,128	-
		Bayesiano 2	-0,008	0,033	-0,008	0,182	-
		M.V.	0,002	0,010	0,002	0,101	-

Tabela 4.9: Resultados do estudo de simulação - $\alpha = 1, 5$, $\beta = 1$, $\gamma = 0, 45$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	1,947	0,001	0,447	0,448	0,298
		Bayesiano 2	1,640	0,001	0,140	0,145	0,093
		M.V.	1,555	0,028	0,055	0,175	0,036
	50	Bayesiano 1	1,782	0,005	0,282	0,291	0,188
		Bayesiano 2	1,533	<0,001	0,033	0,037	0,022
		M.V.	1,602	0,019	0,102	0,172	0,068
	200	Bayesiano 1	1,732	0,004	0,232	0,240	0,154
		Bayesiano 2	1,502	<0,001	0,026	0,029	0,015
		M.V.	1,573	0,006	0,073	0,107	0,049
β	10	Bayesiano 1	1,443	0,180	0,443	0,552	0,354
		Bayesiano 2	1,141	0,015	0,141	0,185	0,141
		M.V.	1,028	0,045	0,028	0,213	0,028
	50	Bayesiano 1	1,377	0,073	0,377	0,464	0,377
		Bayesiano 2	1,032	0,001	0,032	0,043	0,032
		M.V.	0,964	0,037	-0,036	0,197	0,036
	200	Bayesiano 1	1,354	0,015	0,354	0,460	0,443
		Bayesiano 2	1,004	<0,001	0,026	0,032	0,018
		M.V.	0,922	0,015	-0,078	0,146	0,078
γ	10	Bayesiano 1	0,082	0,082	-0,368	0,467	0,819
		Bayesiano 2	0,081	0,063	-0,369	0,446	0,821
		M.V.	0,728	0,527	-0,084	0,731	0,187
	50	Bayesiano 1	0,277	0,063	-0,173	0,305	0,385
		Bayesiano 2	0,309	0,073	-0,141	0,305	0,314
		M.V.	0,592	0,196	0,142	0,465	0,316
	200	Bayesiano 1	0,328	0,016	-0,122	0,175	0,271
		Bayesiano 2	0,372	0,063	-0,043	0,201	0,114
		M.V.	0,366	0,032	0,278	0,331	0,619

Tabela 4.10: Resultados do estudo de simulação - $\alpha = 1, 5$, $\beta = 1$, $\gamma = 0, 67$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	1,945	0,001	0,445	0,446	0,297
		Bayesiano 2	1,636	0,002	0,136	0,142	0,091
		M.V.	1,555	0,025	0,055	0,167	0,037
	50	Bayesiano 1	1,827	0,005	0,327	0,335	0,218
		Bayesiano 2	1,522	0,001	0,022	0,033	0,014
		M.V.	1,636	0,022	0,136	0,202	0,090
	200	Bayesiano 1	1,828	0,005	0,328	0,336	0,219
		Bayesiano 2	1,505	<0,001	0,005	0,011	0,003
		M.V.	1,629	0,018	0,129	0,185	0,086
β	10	Bayesiano 1	1,463	0,182	0,591	0,630	0,463
		Bayesiano 2	1,175	0,017	0,175	0,217	0,175
		M.V.	1,065	0,039	0,065	0,208	0,065
	50	Bayesiano 1	1,534	0,057	0,534	0,585	0,534
		Bayesiano 2	1,047	0,002	0,047	0,063	0,047
		M.V.	1,044	0,023	0,044	0,157	0,044
	200	Bayesiano 1	1,591	0,011	0,463	0,600	0,591
		Bayesiano 2	1,010	<0,001	0,010	0,023	0,010
		M.V.	1,012	0,029	0,012	0,172	0,012
γ	10	Bayesiano 1	0,158	0,077	-0,512	0,582	0,764
		Bayesiano 2	0,143	0,056	-0,527	0,578	0,787
		M.V.	0,534	0,401	-0,136	0,648	0,203
	50	Bayesiano 1	0,426	0,050	-0,244	0,331	0,365
		Bayesiano 2	0,469	0,051	-0,201	0,302	0,299
		M.V.	0,692	0,148	0,022	0,386	0,033
	200	Bayesiano 1	0,518	0,020	-0,152	0,207	0,227
		Bayesiano 2	0,608	0,014	-0,062	0,133	0,092
		M.V.	0,768	0,091	0,098	0,316	0,146

4.2 Estudo de simulação II

Nesta seção, são apresentados resultados de estudos de simulação para verificar a convergência do algoritmo MCMC e ilustrar o comportamento dos algoritmos ECM e MCMC desenvolvidos para o modelo de regressão log-BSNAC ,

$$V_i = \varphi_0 + \varphi_1 x_i + \varepsilon_i, \varepsilon_i \sim \text{SHNAC}(\alpha, 0, \sigma = 2, \gamma), i = 1, \dots, n. \quad (4.2.1)$$

em que, $\varepsilon_i = \log T_i$, e T_i tem distribuição Birnbaum-Saunders Normal assimétrica sob a parametrização centrada, $\text{BSNAC}(\alpha, 1, \gamma)$.

Para comparar as estimativas frequentistas com as bayesianas, foram considerados um total de 30 diferentes cenários, resultantes da combinação de três diferentes tamanhos amostrais, $n = 10$, $n = 50$ e $n = 200$, com os seguintes valores fixados para os parâmetros de interesse: $\alpha \in (0, 5; 1, 5)$, $\varphi = (1, 2)^\top$ e $\gamma \in (-0, 67; -0, 45; 0; 0, 45; 0, 67)$, além disso as observações foram simuladas na presença de uma única variável regressora, gerada da distribuição $U(0,1)$.

Os tamanhos amostrais aqui propostos tem por objetivo verificar se há melhoria na estimação dos parâmetros de interesse consoante o aumento do tamanho amostral. Assim como no modelo probabilístico, os valores fixados para α e β , foram escolhidos com intuito de obter modelos com diferentes formas e pequena variabilidade. Os valores fixados para γ foram escolhidos com o intuito de obter modelos com assimetria forte, moderada e nula.

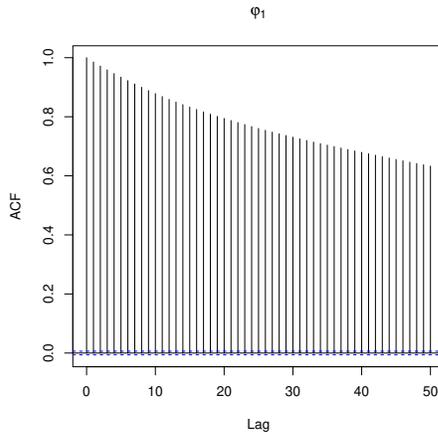
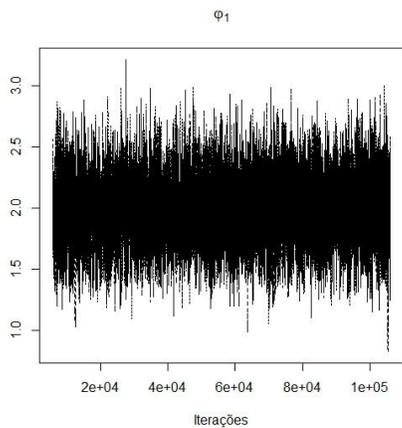
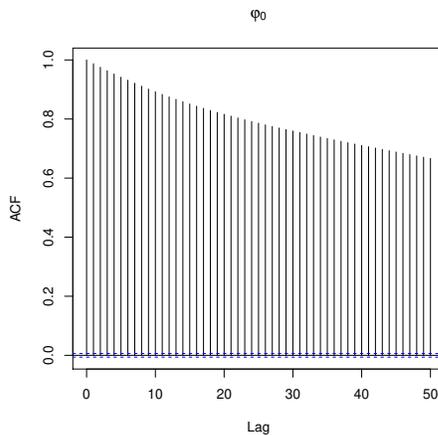
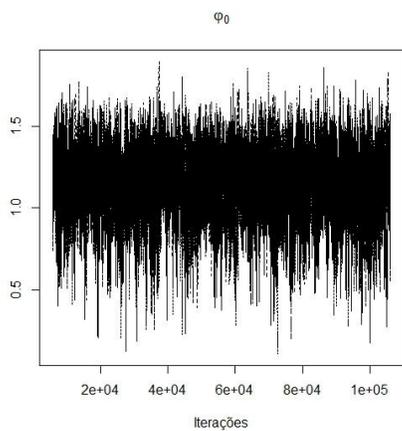
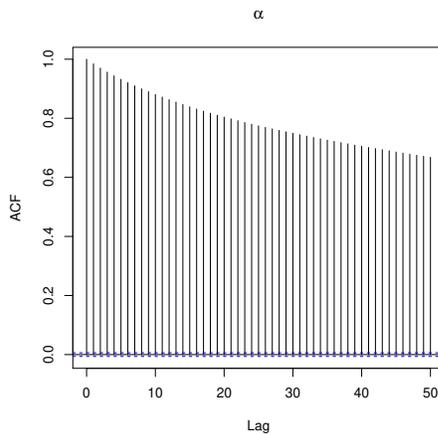
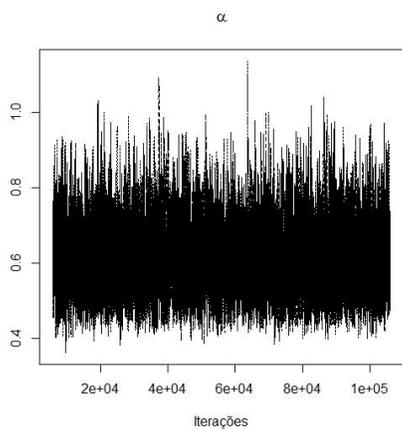
4.2.1 Estudo de convergência

Nesta seção avaliaremos o algoritmo MCMC proposto para o modelo de regressão log-BSNAC em termos da convergência: avaliando a magnitude das autocorrelações e, também, a sensibilidade com relação à escolha de valores iniciais e tamanhos amostrais. Utilizamos o conjunto de prioris descrito na Seção 3.4.2, mais especificamente, para o parâmetro de forma $\alpha > 0$ adotamos uma priori Gama(1; 0,8), o que implica que sua massa está concentrada no intervalo (0, 3), visto que tais valores são mais comumente observados na literatura. Para os parâmetros de regressão, φ_k ; $k = 1, 2$, adotamos uma priori $N(0; 1.000)$ pouco informativa. Por fim, para o parâmetro de assimetria, adotamos uma priori Uniforme contemplando todo o espaço paramétrico. Para ambos os tamanhos amostrais considerados nesse estudo, $n=10$ e $n=200$, fixamos $\alpha = 0, 5$, $\varphi = (1, 2)^\top$, $\gamma = 0, 67$.

O estudo consiste em iniciar paralelamente 3 cadeias. Em cada uma das cadeias o número de iterações foi igual a 104.000. Assim como mencionado na Seção 4.1.1, para verificar convergência do algoritmo, utilizamos alguns métodos comumente considerados na literatura. As Figuras 4.7 - 4.9 dizem respeito ao estudo de convergência considerando um tamanho de amostra $n=10$.

Na Figura 4.7 temos o gráfico de trajetórias das 3 cadeias completas (com as 104.000 iterações), assim como as autocorrelações obtidas. É possível observar uma sobreposição das 3 cadeias, e após o período de aquecimento de 4.000 primeiras iterações, há indícios de que diferentes valores iniciais não alteram, significativamente, os resultados. É possível perceber também a presença de autocorrelação.

Considerando a estatística de Geweke, em todas as 3 cadeias dos parâmetros α , β e γ , a grande maioria dos valores ficaram dentro dos limites considerados aceitáveis, o que indica convergência das cadeias. Por simplicidade tais gráficos não serão apresentados.



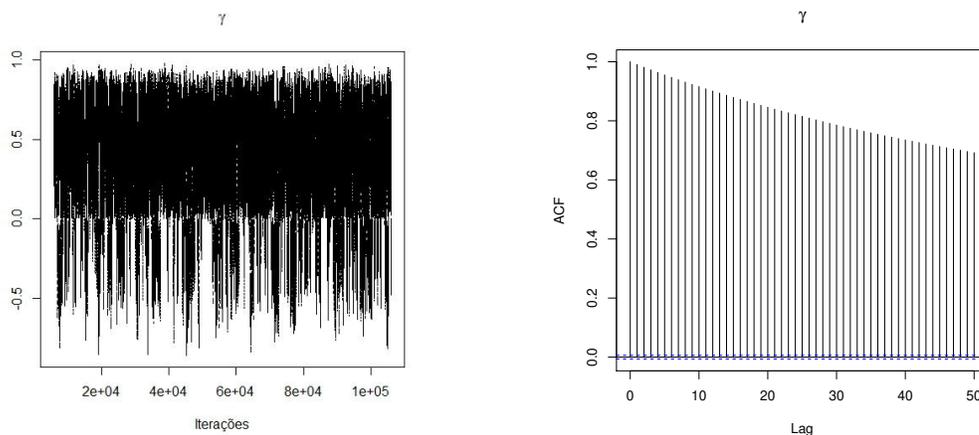


Figura 4.7: Gráficos de trajetórias e de autocorrelação dos parâmetros do modelo de regressão log-BSNAC ($n=10$).

Na Figura 4.8 observa-se os gráficos referentes a estatística de Gelman-Rubin. Há indícios de que, de fato, as cadeias convergem simultaneamente, visto que encontram-se bem próximas do valor unitário.

Na Figura 4.9 temos os correlogramas obtidas utilizando a amostra final, isto é, com as 1.000 estimativas selecionadas da cadeia de Markov, após a retirada das 4.000 primeiras iterações (burn-in) e do espaçamento igual a 100. Nota-se que os valores na sua grande maioria estão todos abaixo de 0,20, indicando assim que as amostras são aproximadamente não correlacionadas.

Nas Figuras 4.10 - 4.12 é possível visualizar os resultados do estudo de convergência considerando um tamanho de amostra $n=200$. Na Figura 4.10 apresentamos o gráfico de trajetórias das 3 cadeias completas (com as 104.000 iterações), assim como as autocorrelações obtidas. É possível observar uma sobreposição das 3 cadeias, e após o período de aquecimento de 4.000 primeiras iterações, há indícios de que diferentes valores iniciais não alteram, significativamente, os resultados.

Considerando a estatística de Geweke, em todas as 3 cadeias dos parâmetros α , φ e γ , a grande maioria dos valores ficaram dentro dos limites considerados aceitáveis, o que indica convergência das cadeias.

Na Figura 4.11 observa-se os gráficos referentes a estatística de Gelman-Rubin. Há indícios de que de fato as cadeias convergem.

Na Figura 4.12 temos os correlogramas obtidas da amostra final, isto é, com as 1.000 estimativas selecionadas da cadeia de Markov, após a retirada das 4000 primeiras iterações (burn-in) e do espaçamento igual a 100. Nota-se que os valores estão na sua grande maioria abaixo de 0,20, indicando assim que as amostras são aproximadamente não correlacionadas. Com o aumento da amostra é possível notar que os valores apresentados na Figura 4.12 são ainda menores dos que aqueles mostrados na Figura 4.9. Além disso consoante ao aumento da amostra, nota-se que a convergência das cadeias é melhor atingida.

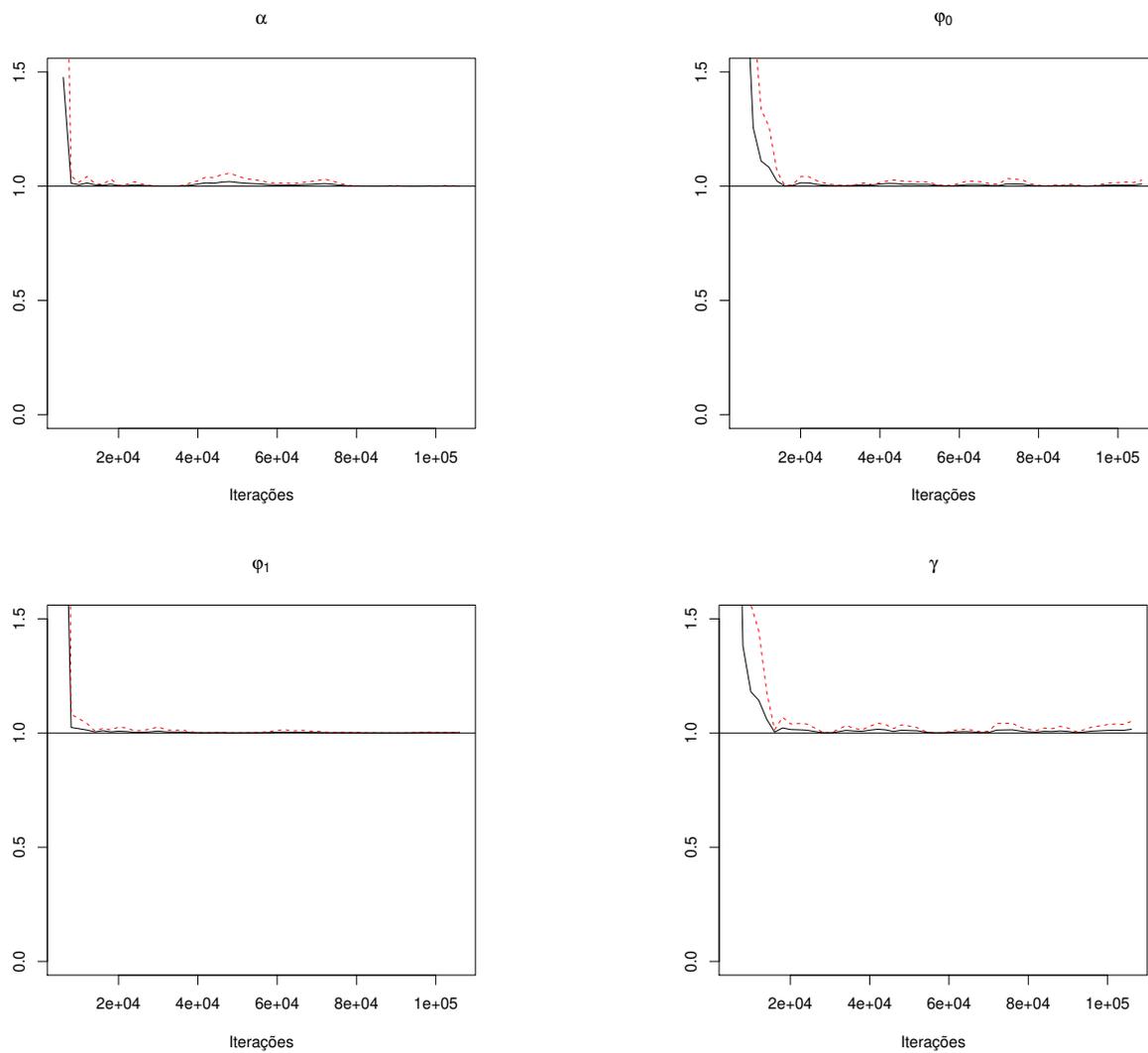


Figura 4.8: Estatística de Gelman-Rubin para o modelo de regressão log-BSNAC. ($n=10$)

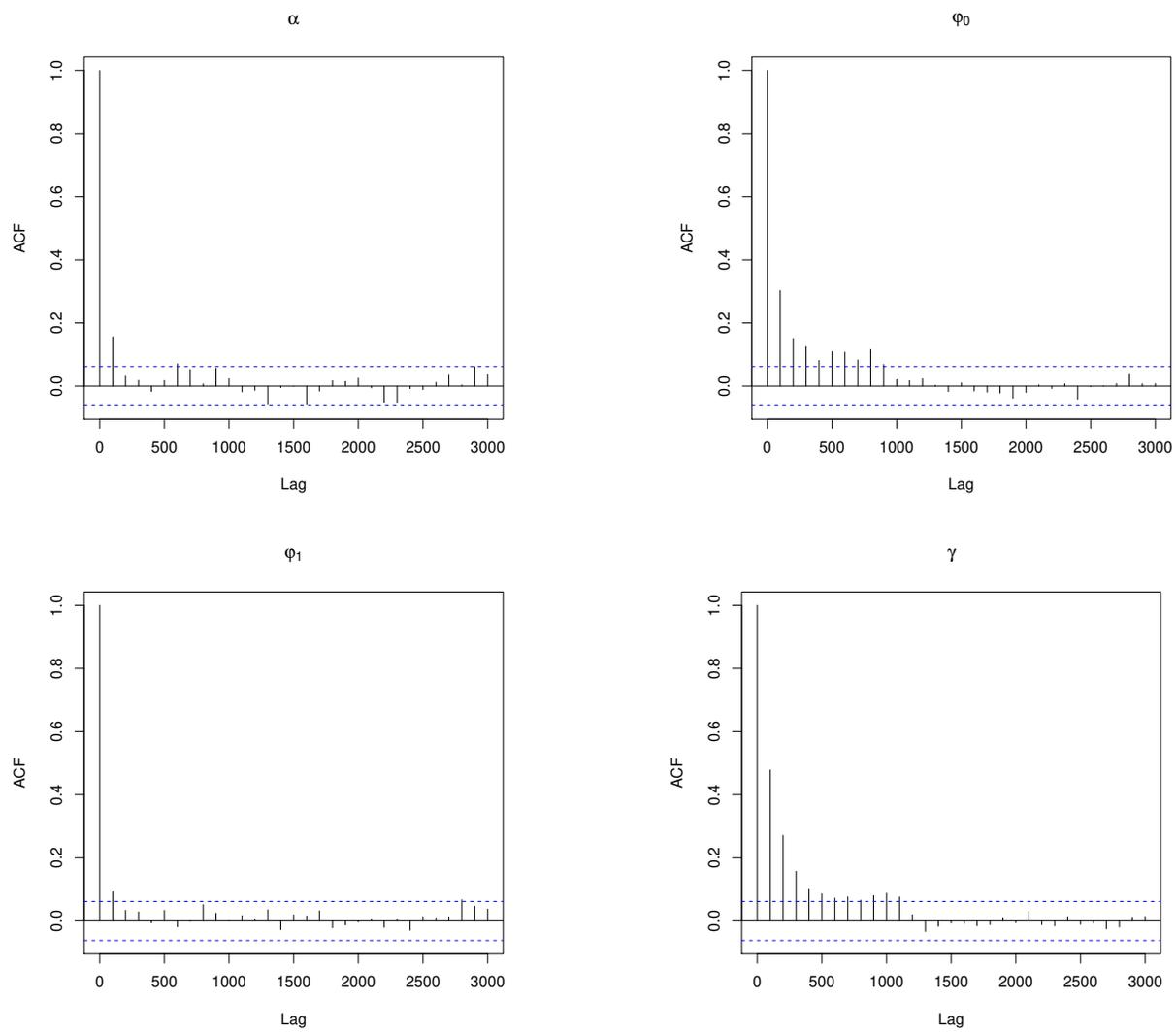
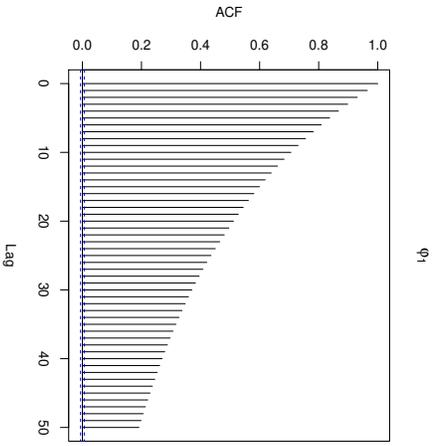
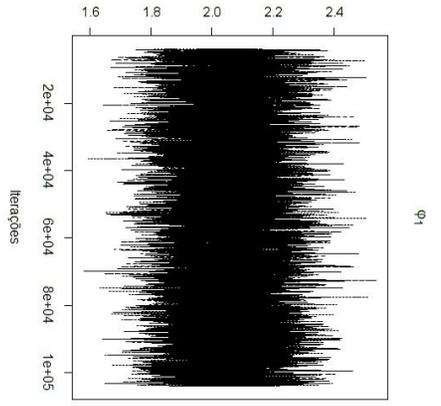
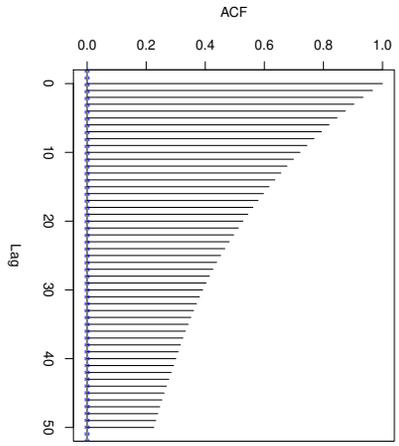
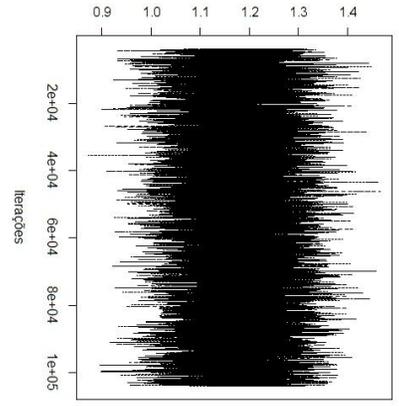
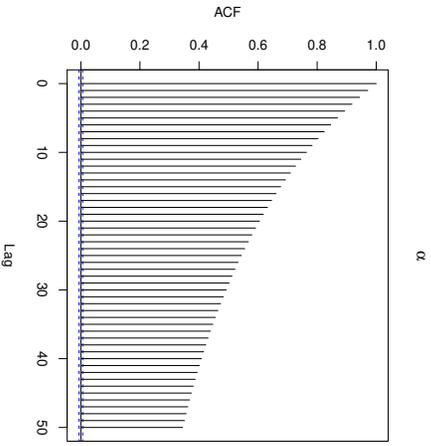
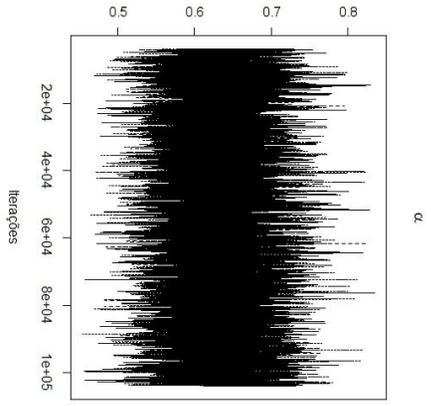


Figura 4.9: Gráficos de autocorrelação para o modelo de regressão log-BSNAC. ($n=10$)



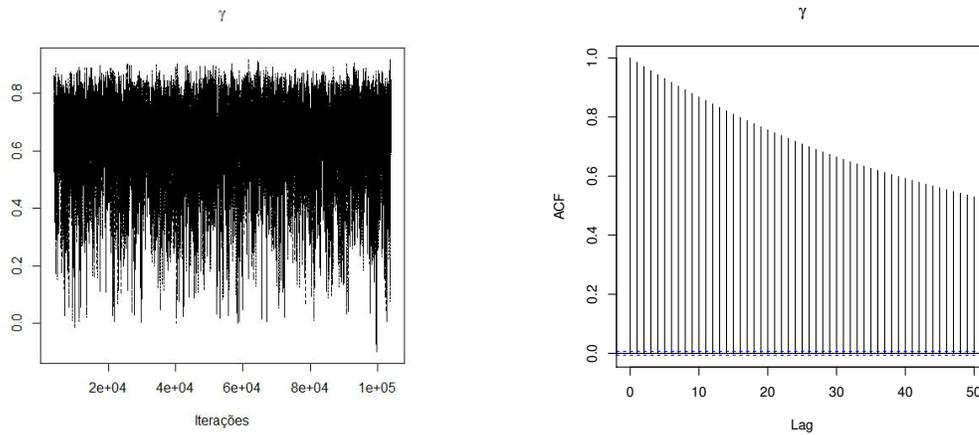
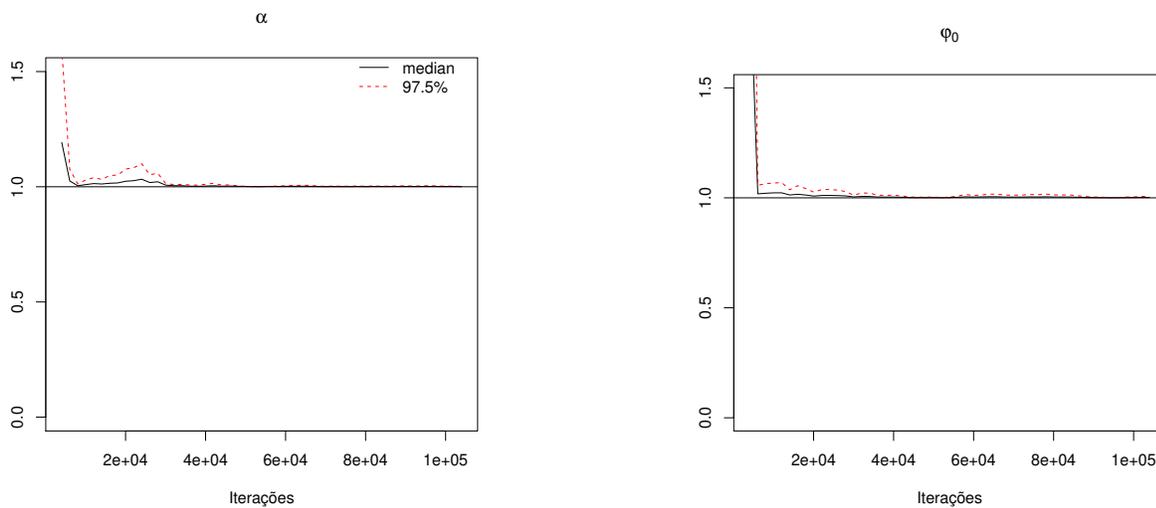


Figura 4.10: Gráficos de trajetórias e de autocorrelação dos parâmetros do modelo de regressão log-BSNAC ($n=200$).



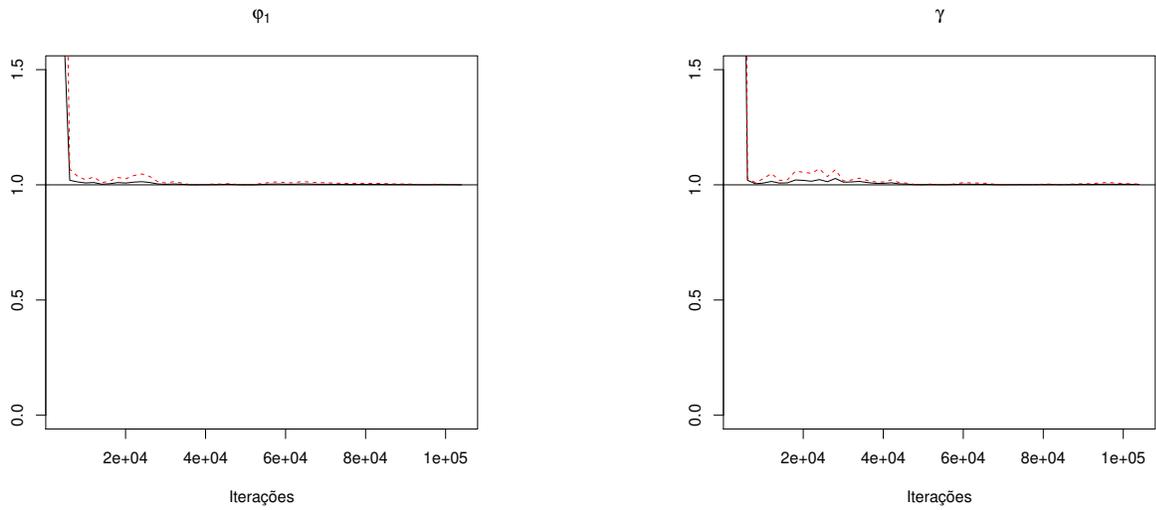
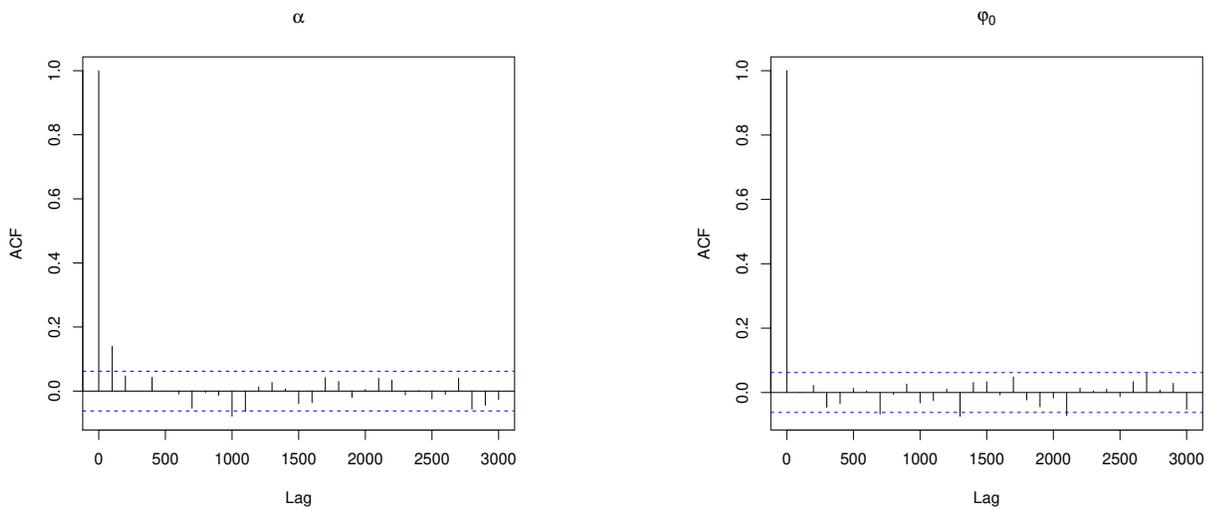


Figura 4.11: Estatística de Gelman-Rubin para o modelo de regressão log-BSNAC. ($n=200$)



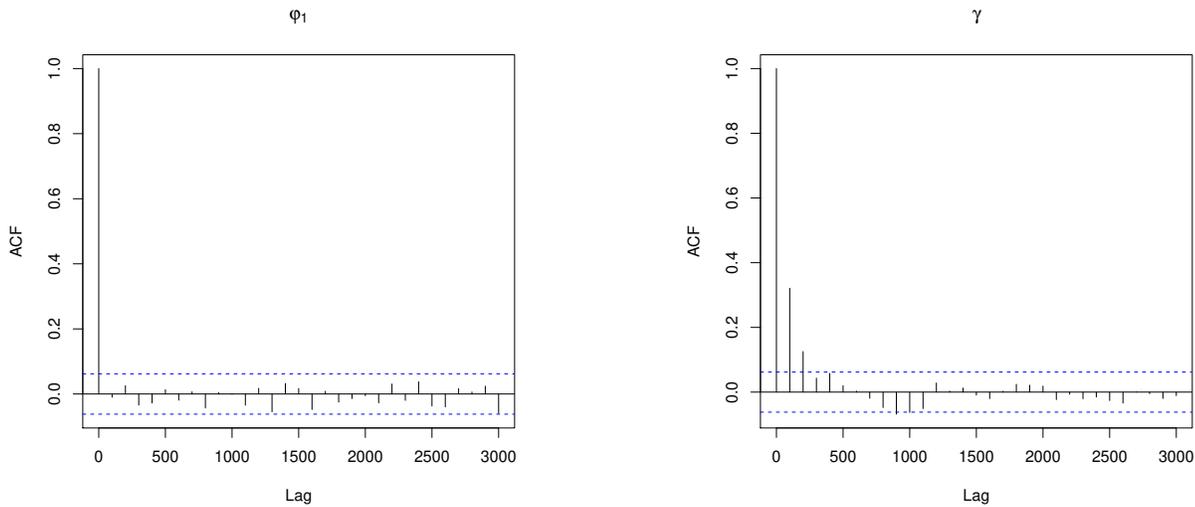


Figura 4.12: Gráficos de autocorrelação para o modelo de regressão log-BSNAC. ($n=200$)

4.2.2 Comparação entre as estimativas frequentistas e bayesianas

Nesta seção, apresentaremos os resultados obtidos através das metodologias frequentista e bayesiana, desenvolvidas no Capítulo 3, considerando os cenários já descritos anteriormente.

As Tabelas 4.11 - 4.15 dizem respeito aos resultados obtidos quando fixamos $\alpha = 0,5$, já as Tabelas 4.16 - 4.20 referem-se aos resultados obtidos quando fixamos $\alpha = 1,5$. Para ambos os conjuntos de tabelas, fixamos $\varphi = (1, 2)^\top$ e variamos o γ de acordo com os valores apresentados no início da Seção 4.2. A fim de facilitar a compreensão, utilizaremos ($< 0,001$) para representar valores positivos e ($> -0,001$) para denotar valores negativos, quando estiverem bem próximos de zero. Além disso, nomearemos como “Bayesiano 1” os resultados bayesianos obtidos a partir da verossimilhança completa, e como “Bayesiano 2” os que foram obtidos a partir da verossimilhança original.

De uma forma geral, analisando as tabelas em questão, nota-se que à medida que aumentamos o tamanho da amostra, as estimativas obtidas tanto pelo método frequentista, quanto pelo método bayesiano tendem a se aproximar dos verdadeiros valores dos parâmetros. É possível notar ainda, que as estimativas obtidas para α tanto pelo método frequentista, quanto pelo método “Bayesiano 2”, aproximam-se dos verdadeiros valores do parâmetro. Em particular, o método “Bayesiano 2” fornece estimativas com valores menores de REQm e AVBR. As estimativas obtidas via método “Bayesiano 1” ficaram em desvantagem frente as estimativas obtidas através dos outros dois métodos, provavelmente pelo fato de estarmos usando a estrutura hierárquica do modelo de regressão.

Para estimar φ_0 , é possível notar que os métodos frequentista e “Bayesiano 2” recuperam de forma satisfatória o verdadeiro valor do parâmetro. Além disso, as estimativas obtidas via método “Bayesiano 2” apresentam menores valores de REQm e AVBR.

Em relação à φ_1 , à medida que aumentamos o tamanho da amostra, nota-se que os métodos são igualmente eficientes e se aproximam do verdadeiro valor do parâmetro. No entanto, na grande maioria dos cenários apresentados, o método bayesiano fornece estimativas com menores variância, vício, REQM e AVBR.

Por fim, para estimar γ , é possível notar que o método bayesiano, tanto usando a verossimilhança completa, quanto a verossimilhança original, na maior parte dos cenários, apresenta estimativas com menores vício, REQM e AVBR, e mais próximas dos verdadeiros valores fixados para γ .

Tabela 4.11: Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1, 2)^\top$ e $\gamma = -0,67$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,838	0,016	0,338	0,362	0,677
		Bayesiano 2	0,607	0,003	0,107	0,118	0,214
		M.V.	0,447	0,002	-0,053	0,068	0,105
	50	Bayesiano 1	0,641	0,001	0,141	0,144	0,282
		Bayesiano 2	0,513	<0,001	0,013	0,020	0,026
		M.V.	0,504	<0,001	0,004	0,021	0,008
	200	Bayesiano 1	0,651	0,001	0,151	0,153	0,302
		Bayesiano 2	0,503	<0,001	0,003	0,005	0,005
		M.V.	0,515	<0,001	0,015	0,025	0,030
φ_0	10	Bayesiano 1	0,692	0,190	-0,308	0,533	0,308
		Bayesiano 2	1,026	0,071	0,026	0,267	0,026
		M.V.	0,966	0,080	-0,034	0,285	0,034
	50	Bayesiano 1	0,808	0,021	-0,192	0,241	0,192
		Bayesiano 2	1,001	0,012	0,001	0,108	0,001
		M.V.	0,998	0,013	-0,002	0,114	0,002
	200	Bayesiano 1	0,769	0,004	-0,231	0,239	0,231
		Bayesiano 2	0,993	0,003	-0,007	0,052	0,007
		M.V.	0,996	0,003	-0,004	0,054	0,004
φ_1	10	Bayesiano 1	2,049	0,625	0,049	0,792	0,025
		Bayesiano 2	1,902	0,302	-0,098	0,558	0,049
		M.V.	2,038	0,328	0,038	0,574	0,019
	50	Bayesiano 1	2,015	0,058	0,015	0,241	0,007
		Bayesiano 2	1,982	0,046	-0,018	0,215	0,009
		M.V.	2,013	0,053	0,013	0,230	0,006
	200	Bayesiano 1	2,023	0,010	0,023	0,101	0,011
		Bayesiano 2	2,009	0,011	0,009	0,104	0,004
		M.V.	2,019	0,011	0,019	0,104	0,009
γ	10	Bayesiano 1	-0,327	0,050	0,343	0,410	0,511
		Bayesiano 2	-0,147	0,051	0,523	0,571	0,781
		M.V.	-0,575	0,377	0,095	0,622	0,142
	50	Bayesiano 1	-0,529	0,054	0,141	0,272	0,211
		Bayesiano 2	-0,504	0,061	0,166	0,297	0,248
		M.V.	-0,718	0,080	-0,048	0,287	0,072
	200	Bayesiano 1	-0,713	0,014	-0,043	0,127	0,064
		Bayesiano 2	-0,607	0,015	0,063	0,137	0,094
		M.V.	-0,778	0,017	-0,108	0,169	0,161

Tabela 4.12: Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1, 2)^\top$ e $\gamma = -0,45$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,809	0,024	0,309	0,346	0,618
		Bayesiano 2	0,613	0,003	0,113	0,125	0,225
		M.V.	0,439	0,002	-0,061	0,078	0,121
	50	Bayesiano 1	0,626	0,001	0,126	0,129	0,252
		Bayesiano 2	0,519	<0,001	0,019	0,021	0,039
		M.V.	0,507	0,001	0,007	0,025	0,013
	200	Bayesiano 1	0,616	0,001	0,116	0,119	0,232
		Bayesiano 2	0,504	<0,001	0,004	0,005	0,009
		M.V.	0,519	0,001	0,019	0,031	0,038
φ_0	10	Bayesiano 1	0,758	0,254	-0,242	0,559	0,242
		Bayesiano 2	0,979	0,080	-0,021	0,284	0,021
		M.V.	0,978	0,107	-0,022	0,328	0,022
	50	Bayesiano 1	0,892	0,027	-0,108	0,196	0,108
		Bayesiano 2	0,987	0,010	-0,013	0,099	0,013
		M.V.	1,032	0,016	0,032	0,130	0,032
	200	Bayesiano 1	0,808	0,005	-0,192	0,204	0,192
		Bayesiano 2	0,999	0,003	-0,001	0,058	0,001
		M.V.	0,998	0,004	-0,002	0,062	0,002
φ_1	10	Bayesiano 1	1,989	0,865	-0,011	0,930	0,006
		Bayesiano 2	2,035	0,374	0,035	0,613	0,018
		M.V.	2,029	0,424	0,029	0,651	0,014
	50	Bayesiano 1	1,959	0,069	-0,041	0,266	0,020
		Bayesiano 2	2,014	0,038	0,014	0,197	0,007
		M.V.	1,952	0,064	-0,048	0,257	0,024
	200	Bayesiano 1	2,018	0,011	0,018	0,109	0,009
		Bayesiano 2	1,999	0,013	-0,001	0,116	0,001
		M.V.	2,017	0,015	0,017	0,123	0,008
γ	10	Bayesiano 1	-0,312	0,031	0,138	0,225	0,307
		Bayesiano 2	-0,036	0,058	0,414	0,479	0,920
		M.V.	-0,581	0,328	-0,131	0,588	0,292
	50	Bayesiano 1	-0,355	0,084	0,095	0,305	0,211
		Bayesiano 2	-0,342	0,075	0,108	0,295	0,239
		M.V.	-0,531	0,185	-0,081	0,438	0,180
	200	Bayesiano 1	-0,521	0,024	-0,071	0,169	0,157
		Bayesiano 2	-0,421	0,026	0,029	0,163	0,065
		M.V.	-0,627	0,046	-0,177	0,278	0,393

Tabela 4.13: Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1, 2)^T$ e $\gamma = 0$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,819	0,010	0,319	0,335	0,638
		Bayesiano 2	0,612	0,002	0,112	0,123	0,225
		M.V.	0,445	0,001	-0,055	0,064	0,110
	50	Bayesiano 1	0,615	0,001	0,115	0,117	0,229
		Bayesiano 2	0,521	<0,001	0,021	0,023	0,041
		M.V.	0,488	<0,001	-0,012	0,016	0,024
	200	Bayesiano 1	0,567	<0,001	0,067	0,069	0,133
		Bayesiano 2	0,506	<0,001	0,006	0,007	0,013
		M.V.	0,498	<0,001	-0,002	0,004	0,005
φ_0	10	Bayesiano 1	0,819	0,244	-0,181	0,526	0,181
		Bayesiano 2	0,966	0,088	-0,034	0,299	0,034
		M.V.	1,017	0,082	0,017	0,286	0,017
	50	Bayesiano	0,940	0,032	-0,060	0,187	0,060
		Bayesiano 2	1,000	0,016	> -0,001	0,127	<0,001
		M.V.	0,998	0,018	-0,002	0,136	0,002
	200	Bayesiano 1	0,985	0,013	-0,015	0,114	0,015
		Bayesiano 2	0,992	0,003	-0,008	0,052	0,008
		M.V.	1,001	0,004	0,001	0,061	0,001
φ_1	10	Bayesiano 1	1,974	0,720	-0,026	0,849	0,013
		Bayesiano 2	2,057	0,386	0,057	0,624	0,029
		M.V.	1,983	0,315	-0,017	0,562	0,008
	50	Bayesiano	2,008	0,083	0,008	0,289	0,004
		Bayesiano 2	2,003	0,067	0,003	0,259	0,001
		M.V.	2,002	0,072	0,002	0,268	0,001
	200	Bayesiano 1	1,998	0,015	-0,002	0,124	0,001
		Bayesiano 2	2,016	0,011	0,016	0,106	0,008
		M.V.	1,996	0,015	-0,004	0,122	0,002
γ	10	Bayesiano 1	-0,220	0,050	-0,220	0,314	-
		Bayesiano 2	-0,007	0,038	-0,007	0,195	-
		M.V.	-0,016	0,011	-0,016	0,108	-
	50	Bayesiano	-0,132	0,093	-0,132	0,333	-
		Bayesiano 2	-0,002	0,098	-0,002	0,314	-
		M.V.	-0,039	0,026	-0,039	0,166	-
	200	Bayesiano 1	-0,033	0,051	-0,033	0,228	-
		Bayesiano 2	0,044	0,032	0,044	0,183	-
		M.V.	0,006	0,012	0,006	0,109	-

Tabela 4.14: Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1, 2)^T$ e $\gamma = 0,45$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,807	0,027	0,307	0,348	0,614
		Bayesiano 2	0,619	0,002	0,119	0,129	0,238
		M.V.	0,439	0,003	-0,061	0,081	0,123
	50	Bayesiano 1	0,612	<0,001	0,112	0,114	0,224
		Bayesiano 2	0,518	<0,001	0,018	0,021	0,036
		M.V.	0,506	0,001	0,006	0,028	0,012
	200	Bayesiano 1	0,612	0,001	0,112	0,115	0,223
		Bayesiano 2	0,504	<0,001	0,004	0,005	0,009
		M.V.	0,513	<0,001	0,013	0,025	0,026
φ_0	10	Bayesiano 1	0,941	0,336	-0,059	0,583	0,059
		Bayesiano 2	1,040	0,087	0,040	0,298	0,040
		M.V.	1,051	0,128	0,051	0,361	0,051
	50	Bayesiano 1	1,085	0,020	0,085	0,165	0,085
		Bayesiano 2	1,008	0,015	0,008	0,121	0,008
		M.V.	0,992	0,016	-0,008	0,128	0,008
	200	Bayesiano 1	1,175	0,006	0,175	0,190	0,175
		Bayesiano 2	0,999	0,004	-0,001	0,060	0,001
		M.V.	1,001	0,003	0,001	0,056	0,001
φ_1	10	Bayesiano 1	1,788	0,885	-0,212	0,964	0,106
		Bayesiano 2	1,912	0,359	-0,088	0,606	0,044
		M.V.	1,907	0,426	-0,093	0,659	0,047
	50	Bayesiano 1	2,003	0,071	0,003	0,267	0,001
		Bayesiano 2	1,992	0,057	-0,008	0,239	0,004
		M.V.	2,001	0,067	0,001	0,258	0,001
	200	Bayesiano 1	1,994	0,012	-0,006	0,109	0,003
		Bayesiano 2	2,004	0,014	0,004	0,120	0,002
		M.V.	1,992	0,012	-0,008	0,112	0,004
γ	10	Bayesiano 1	-0,146	0,051	-0,596	0,637	1,325
		Bayesiano 2	0,042	0,040	-0,408	0,455	0,907
		M.V.	0,432	0,461	-0,018	0,679	0,040
	50	Bayesiano 1	0,277	0,067	-0,173	0,311	0,385
		Bayesiano 2	0,322	0,076	-0,128	0,304	0,285
		M.V.	0,525	0,156	0,075	0,401	0,166
	200	Bayesiano 1	0,499	0,028	0,049	0,176	0,109
		Bayesiano 2	0,414	0,026	-0,036	0,164	0,081
		M.V.	0,614	0,033	0,164	0,245	0,364

Tabela 4.15: Resultados do estudo de simulação - $\alpha = 0,5$, $\varphi = (1, 2)^T$ e $\gamma = 0,67$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	0,804	0,010	0,304	0,321	0,608
		Bayesiano 2	0,607	0,003	0,107	0,119	0,214
		M.V.	0,452	0,002	-0,048	0,066	0,096
	50	Bayesiano 1	0,631	0,001	0,131	0,133	0,262
		Bayesiano 2	0,512	<0,001	0,012	0,017	0,024
		M.V.	0,503	<0,001	0,003	0,020	0,007
	200	Bayesiano 1	0,648	0,001	0,148	0,150	0,296
		Bayesiano 2	0,502	<0,001	0,002	0,005	0,004
		M.V.	0,516	0,001	0,016	0,027	0,032
φ_0	10	Bayesiano 1	0,798	0,245	-0,202	0,534	0,202
		Bayesiano 2	1,009	0,104	0,009	0,323	0,009
		M.V.	0,988	0,107	-0,012	0,327	0,012
	50	Bayesiano 1	1,175	0,015	0,175	0,213	0,175
		Bayesiano 2	1,000	0,012	<0,001	0,112	<0,001
		M.V.	1,006	0,014	0,006	0,118	0,006
	200	Bayesiano 1	1,218	0,003	0,218	0,225	0,218
		Bayesiano 2	1,005	0,003	0,005	0,056	0,005
		M.V.	0,995	0,003	-0,005	0,053	0,005
φ_1	10	Bayesiano 1	2,047	0,754	0,047	0,870	0,023
		Bayesiano 2	2,002	0,361	0,002	0,601	0,001
		M.V.	2,028	0,358	0,028	0,599	0,014
	50	Bayesiano 1	1,960	0,051	-0,040	0,230	0,020
		Bayesiano 2	2,016	0,050	0,016	0,224	0,008
		M.V.	1,971	0,054	-0,029	0,234	0,015
	200	Bayesiano 1	1,994	0,009	-0,006	0,095	0,003
		Bayesiano 2	1,994	0,012	-0,006	0,111	0,003
		M.V.	1,997	0,011	-0,003	0,104	0,002
γ	10	Bayesiano 1	-0,141	0,053	-0,811	0,843	1,210
		Bayesiano 2	0,107	0,045	-0,563	0,601	0,840
		M.V.	0,576	0,348	-0,094	0,598	0,141
	50	Bayesiano 1	0,501	0,056	-0,169	0,290	0,252
		Bayesiano 2	0,499	0,060	-0,171	0,299	0,255
		M.V.	0,756	0,058	0,086	0,256	0,128
	200	Bayesiano 1	0,706	0,015	0,036	0,129	0,054
		Bayesiano 2	0,620	0,013	-0,050	0,126	0,074
		M.V.	0,782	0,019	0,112	0,178	0,167

Tabela 4.16: Resultados do estudo de simulação - $\alpha = 1, 5$, $\varphi = (1, 2)^\top$ e $\gamma = -0, 67$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	2,139	0,094	0,639	0,708	0,426
		Bayesiano 2	1,944	0,014	0,444	0,459	0,296
		M.V.	1,347	0,008	-0,153	0,177	0,102
	50	Bayesiano 1	1,963	0,025	0,463	0,489	0,309
		Bayesiano 2	1,536	0,001	0,036	0,050	0,024
		M.V.	1,481	0,003	-0,019	0,062	0,012
	200	Bayesiano 1	1,849	0,007	0,349	0,359	0,233
		Bayesiano 2	1,508	<0,001	0,008	0,013	0,005
		M.V.	1,511	0,001	0,011	0,038	0,007
φ_0	10	Bayesiano 1	0,690	0,233	-0,310	0,573	0,310
		Bayesiano 2	0,999	0,317	-0,001	0,563	0,001
		M.V.	0,955	0,281	-0,045	0,532	0,045
	50	Bayesiano 1	0,396	0,118	-0,604	0,695	0,604
		Bayesiano 2	0,944	0,090	-0,056	0,306	0,056
		M.V.	1,033	0,074	0,033	0,273	0,033
	200	Bayesiano 1	0,476	0,050	-0,524	0,570	0,524
		Bayesiano 2	0,981	0,014	-0,019	0,122	0,019
		M.V.	1,019	0,021	0,019	0,147	0,019
φ_1	10	Bayesiano 1	1,684	0,889	-0,316	0,994	0,158
		Bayesiano 2	1,864	1,257	-0,136	1,130	0,068
		M.V.	2,123	1,161	0,123	1,085	0,061
	50	Bayesiano 1	2,135	0,413	0,135	0,657	0,068
		Bayesiano 2	2,048	0,356	0,048	0,599	0,024
		M.V.	1,965	0,272	-0,035	0,523	0,017
	200	Bayesiano 1	2,079	0,123	0,079	0,360	0,039
		Bayesiano 2	2,013	0,055	0,013	0,235	0,007
		M.V.	2,006	0,080	0,006	0,283	0,003
γ	10	Bayesiano 1	-0,277	0,059	0,393	0,461	0,586
		Bayesiano 2	-0,141	0,037	0,529	0,563	0,790
		M.V.	-0,509	0,357	0,161	0,619	0,241
	50	Bayesiano 1	-0,501	0,035	0,169	0,252	0,252
		Bayesiano 2	-0,477	0,063	0,193	0,316	0,287
		M.V.	-0,666	0,092	0,004	0,303	0,007
	200	Bayesiano 1	-0,508	0,017	0,162	0,208	0,241
		Bayesiano 2	-0,600	0,015	0,070	0,140	0,104
		M.V.	-0,689	0,025	-0,019	0,158	0,028

Tabela 4.17: Resultados do estudo de simulação - $\alpha = 1, 5$, $\varphi = (1, 2)^\top$ e $\gamma = -0, 45$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	2,111	0,096	0,611	0,685	0,407
		Bayesiano 2	1,947	0,009	0,447	0,457	0,298
		M.V.	1,333	0,009	-0,167	0,191	0,111
	50	Bayesiano 1	1,843	0,012	0,343	0,360	0,229
		Bayesiano 2	1,550	0,001	0,050	0,058	0,033
		M.V.	1,481	0,002	-0,019	0,051	0,012
	200	Bayesiano 1	1,758	0,004	0,258	0,265	0,172
		Bayesiano 2	1,512	<0,001	0,012	0,015	0,008
		M.V.	1,517	0,001	0,017	0,038	0,012
φ_0	10	Bayesiano 1	0,700	0,275	-0,300	0,604	0,300
		Bayesiano 2	1,029	0,279	0,029	0,529	0,029
		M.V.	1,027	0,304	0,027	0,552	0,027
	50	Bayesiano 1	0,596	0,289	-0,404	0,672	0,404
		Bayesiano 2	0,976	0,072	-0,024	0,270	0,024
		M.V.	1,015	0,098	0,015	0,313	0,015
	200	Bayesiano 1	0,574	0,045	-0,426	0,476	0,426
		Bayesiano 2	1,000	0,018	<0,001	0,135	<0,001
		M.V.	1,029	0,025	0,029	0,162	0,029
φ_1	10	Bayesiano 1	1,689	1,037	-0,311	1,065	0,155
		Bayesiano 2	1,903	1,182	-0,097	1,091	0,048
		M.V.	1,938	1,205	-0,062	1,100	0,031
	50	Bayesiano 1	2,114	0,726	0,114	0,859	0,057
		Bayesiano 2	2,017	0,287	0,017	0,536	0,008
		M.V.	2,002	0,377	0,002	0,614	0,001
	200	Bayesiano 1	2,062	0,135	0,062	0,373	0,031
		Bayesiano 2	1,986	0,072	-0,014	0,268	0,007
		M.V.	1,995	0,092	-0,005	0,304	0,002
γ	10	Bayesiano 1	-0,248	0,059	0,202	0,317	0,449
		Bayesiano 2	-0,062	0,037	0,388	0,432	0,861
		M.V.	-0,469	0,382	-0,019	0,618	0,042
	50	Bayesiano 1	-0,276	0,057	0,174	0,295	0,387
		Bayesiano 2	-0,345	0,080	0,105	0,302	0,234
		M.V.	-0,347	0,127	0,103	0,371	0,228
	200	Bayesiano 1	-0,350	0,011	0,100	0,144	0,221
		Bayesiano 2	-0,416	0,026	0,034	0,164	0,076
		M.V.	-0,564	0,035	-0,114	0,220	0,254

Tabela 4.18: Resultados do estudo de simulação - $\alpha = 1, 5$, $\varphi = (1, 2)^T$ e $\gamma = 0$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	2,020	0,073	0,520	0,586	0,347
		Bayesiano 2	1,956	0,008	0,456	0,465	0,304
		M.V.	1,352	0,004	-0,148	0,162	0,099
	50	Bayesiano 1	1,792	0,009	0,292	0,308	0,195
		Bayesiano 2	1,561	0,001	0,061	0,066	0,040
		M.V.	1,461	0,001	-0,039	0,055	0,026
	200	Bayesiano 1	1,633	0,001	0,133	0,137	0,089
		Bayesiano 2	1,516	<0,001	0,016	0,017	0,011
		M.V.	1,494	<0,001	-0,006	0,012	0,004
φ_0	10	Bayesiano 1	0,777	0,295	-0,223	0,587	0,223
		Bayesiano 2	1,009	0,298	0,009	0,546	0,009
		M.V.	0,938	0,317	-0,062	0,566	0,062
	50	Bayesiano 1	0,727	0,358	-0,273	0,658	0,273
		Bayesiano 2	1,032	0,090	0,032	0,301	0,032
		M.V.	0,952	0,108	-0,048	0,332	0,048
	200	Bayesiano 1	0,986	0,074	-0,014	0,272	0,014
		Bayesiano 2	0,984	0,016	-0,016	0,129	0,016
		M.V.	0,999	0,023	-0,001	0,150	0,001
φ_1	10	Bayesiano 1	1,914	1,514	-0,086	1,234	0,043
		Bayesiano 2	1,987	1,345	-0,013	1,160	0,006
		M.V.	2,176	1,420	0,176	1,205	0,088
	50	Bayesiano 1	2,198	0,825	0,198	0,930	0,099
		Bayesiano 2	1,935	0,332	-0,065	0,580	0,032
		M.V.	2,095	0,421	0,095	0,656	0,047
	200	Bayesiano 1	2,030	0,124	0,030	0,353	0,015
		Bayesiano 2	2,034	0,065	0,034	0,258	0,017
		M.V.	2,011	0,089	0,011	0,298	0,005
γ	10	Bayesiano 1	-0,095	0,078	-0,095	0,295	-
		Bayesiano 2	-0,016	0,045	-0,016	0,213	-
		M.V.	0,012	0,029	0,012	0,170	-
	50	Bayesiano 1	-0,124	0,065	-0,124	0,284	-
		Bayesiano 2	0,014	0,085	0,014	0,292	-
		M.V.	-0,002	0,016	-0,002	0,126	-
	200	Bayesiano 1	0,006	0,019	0,006	0,139	-
		Bayesiano 2	0,043	0,031	0,043	0,181	-
		M.V.	0,017	0,010	0,017	0,101	-

Tabela 4.19: Resultados do estudo de simulação - $\alpha = 1, 5$, $\varphi = (1, 2)^T$ e $\gamma = 0, 45$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	2,038	0,088	0,538	0,614	0,359
		Bayesiano 2	1,954	0,010	0,454	0,465	0,303
		M.V.	1,353	0,007	-0,147	0,169	0,098
	50	Bayesiano 1	1,797	0,006	0,297	0,306	0,198
		Bayesiano 2	1,554	0,001	0,054	0,060	0,036
		M.V.	1,489	0,002	-0,011	0,046	0,007
	200	Bayesiano 1	1,726	0,004	0,226	0,234	0,151
		Bayesiano 2	1,512	<0,001	0,012	0,014	0,008
		M.V.	1,528	0,003	0,028	0,062	0,018
φ_0	10	Bayesiano 1	0,939	0,303	-0,061	0,554	0,061
		Bayesiano 2	0,991	0,314	-0,009	0,560	0,009
		M.V.	1,018	0,277	0,018	0,527	0,018
	50	Bayesiano 1	1,264	0,264	0,264	0,577	0,264
		Bayesiano 2	1,034	0,085	0,034	0,294	0,034
		M.V.	1,036	0,073	0,036	0,272	0,036
	200	Bayesiano 1	1,355	0,030	0,355	0,395	0,355
		Bayesiano 2	0,998	0,018	-0,002	0,134	0,002
		M.V.	0,964	0,027	-0,036	0,167	0,036
φ_1	10	Bayesiano 1	1,807	1,205	-0,193	1,114	0,096
		Bayesiano 2	2,093	1,326	0,093	1,155	0,047
		M.V.	1,987	1,127	-0,013	1,062	0,007
	50	Bayesiano 1	1,789	0,667	-0,211	0,843	0,105
		Bayesiano 2	1,972	0,338	-0,028	0,582	0,014
		M.V.	1,891	0,288	-0,109	0,548	0,055
	200	Bayesiano 1	1,983	0,071	-0,017	0,266	0,009
		Bayesiano 2	2,019	0,072	0,019	0,268	0,009
		M.V.	1,993	0,087	-0,007	0,296	0,003
γ	10	Bayesiano 1	-0,017	0,078	-0,467	0,544	1,037
		Bayesiano 2	0,060	0,036	-0,390	0,434	0,867
		M.V.	0,533	0,317	0,083	0,569	0,184
	50	Bayesiano 1	0,202	0,060	-0,248	0,348	0,550
		Bayesiano 2	0,272	0,056	-0,178	0,296	0,395
		M.V.	0,520	0,095	0,070	0,316	0,156
	200	Bayesiano 1	0,322	0,016	-0,128	0,181	0,285
		Bayesiano 2	0,409	0,025	-0,041	0,165	0,092
		M.V.	0,599	0,046	0,149	0,261	0,331

Tabela 4.20: Resultados do estudo de simulação - $\alpha = 1, 5$, $\varphi = (1, 2)^T$ e $\gamma = 0, 67$.

Parâmetro	n	Método	Média	Variância	Vício	REQM	AVRB
α	10	Bayesiano 1	1,989	0,040	0,489	0,528	0,326
		Bayesiano 2	1,945	0,011	0,445	0,458	0,297
		M.V.	1,333	0,009	-0,167	0,192	0,111
	50	Bayesiano 1	1,819	0,007	0,319	0,330	0,213
		Bayesiano 2	1,537	0,001	0,037	0,050	0,025
		M.V.	1,470	0,002	-0,030	0,051	0,020
	200	Bayesiano 1	1,819	0,005	0,319	0,328	0,213
		Bayesiano 2	1,507	<0,001	0,007	0,013	0,005
		M.V.	1,516	0,003	0,016	0,053	0,011
φ_0	10	Bayesiano 1	0,979	0,326	-0,021	0,571	0,021
		Bayesiano 2	1,106	0,313	0,106	0,570	0,106
		M.V.	0,994	0,409	-0,006	0,639	0,006
	50	Bayesiano 1	1,375	0,292	0,375	0,658	0,375
		Bayesiano 2	1,022	0,061	0,022	0,248	0,022
		M.V.	1,009	0,091	0,009	0,303	0,009
	200	Bayesiano 1	1,447	0,023	0,447	0,473	0,447
		Bayesiano 2	1,014	0,017	0,014	0,130	0,014
		M.V.	0,971	0,022	-0,029	0,152	0,029
φ_1	10	Bayesiano 1	1,735	1,272	-0,265	1,159	0,133
		Bayesiano 2	1,909	1,320	-0,091	1,152	0,045
		M.V.	2,038	1,658	0,038	1,288	0,019
	50	Bayesiano 1	1,871	0,703	-0,129	0,848	0,065
		Bayesiano 2	2,040	0,240	0,040	0,492	0,020
		M.V.	1,964	0,339	-0,036	0,584	0,018
	200	Bayesiano 1	1,996	0,078	-0,004	0,279	0,002
		Bayesiano 2	1,996	0,066	-0,004	0,256	0,002
		M.V.	2,005	0,079	0,005	0,280	0,003
γ	10	Bayesiano 1	-0,031	0,076	-0,701	0,753	1,047
		Bayesiano 2	0,079	0,043	-0,591	0,626	0,882
		M.V.	0,513	0,406	-0,157	0,656	0,234
	50	Bayesiano 1	0,379	0,072	-0,291	0,396	0,435
		Bayesiano 2	0,510	0,041	-0,160	0,258	0,239
		M.V.	0,660	0,107	-0,010	0,327	0,015
	200	Bayesiano 1	0,507	0,021	-0,163	0,217	0,243
		Bayesiano 2	0,615	0,013	-0,055	0,128	0,083
		M.V.	0,720	0,028	0,050	0,174	0,075

4.2.3 Análise residual

O principal objetivo da análise residual, que aqui será apresentada, consiste em avaliar o comportamento dos resíduos, quando as observações são geradas a partir de quatro diferentes modelos e o modelo log-BSNAC é usado para fazer inferência. Mais especificamente, utilizamos os modelos log-BSNAC, log-BS, log-BS-t-Student e log-BS-t-Student assimétrica para simular os diferentes conjuntos de observações. Para simular observações desses modelos, basta considerar que a variável (3.2.4) tem distribuição Normal padrão, t-Student padrão e t-Student assimétrica padrão, respectivamente. No nosso caso, para todos os modelos, consideramos $\alpha = 0,5$, $\varphi = (1, 2)^\top$. Além disso para os modelos log-BSNAC e log-BS-t assimétrico consideramos $\gamma = 0,67$, para os modelos log-BS-t e log-BS-t assimétrico fixamos $\nu = 4$, em que ν representa os graus de liberdade da distribuição t-Student. Consideramos nesse estudo um tamanho de amostra $n=200$.

Para facilitar a análise dos resultados aqui apresentados, iremos sintetizar as informações referentes à cada um dos cenários na Tabela 4.21.

Tabela 4.21: Cenários referentes à análise de resíduos.

Cenário	Modelo	Gráficos
C1	log-BSNAC	Figuras 4.13 e 4.14
C2	log-BS	Figuras 4.15 e 4.16
C3	log-BS-t	Figuras 4.17 e 4.18
C4	log-BS-t assimétrico	Figuras 4.19 e 4.20

No cenário C1, observando as Figuras 4.13 (a) e (b), é possível perceber que os resíduos encontram-se aleatoriamente dispersos, evidenciando assim a inexistência de autocorrelação e o mesmo pode-se concluir das Figuras 4.14 (a) e (b). As Figuras 4.13 (c) e 4.14 (c) nos dão indícios do comportamento simétrico dos resíduos. Por fim, analisando os gráficos de envelopes simulados, Figuras 4.13 (d) e 4.14 (d) nota-se que os resíduos encontram-se dentro das bandas de confiança e que parece não haver tendência. As conclusões referentes à análise residual do cenário C2 são semelhantes às do cenário C1. Isto é esperado, haja vista que o modelo log-BS utilizado para gerar o conjunto de observações é um caso particular do modelo log-BSNAC.

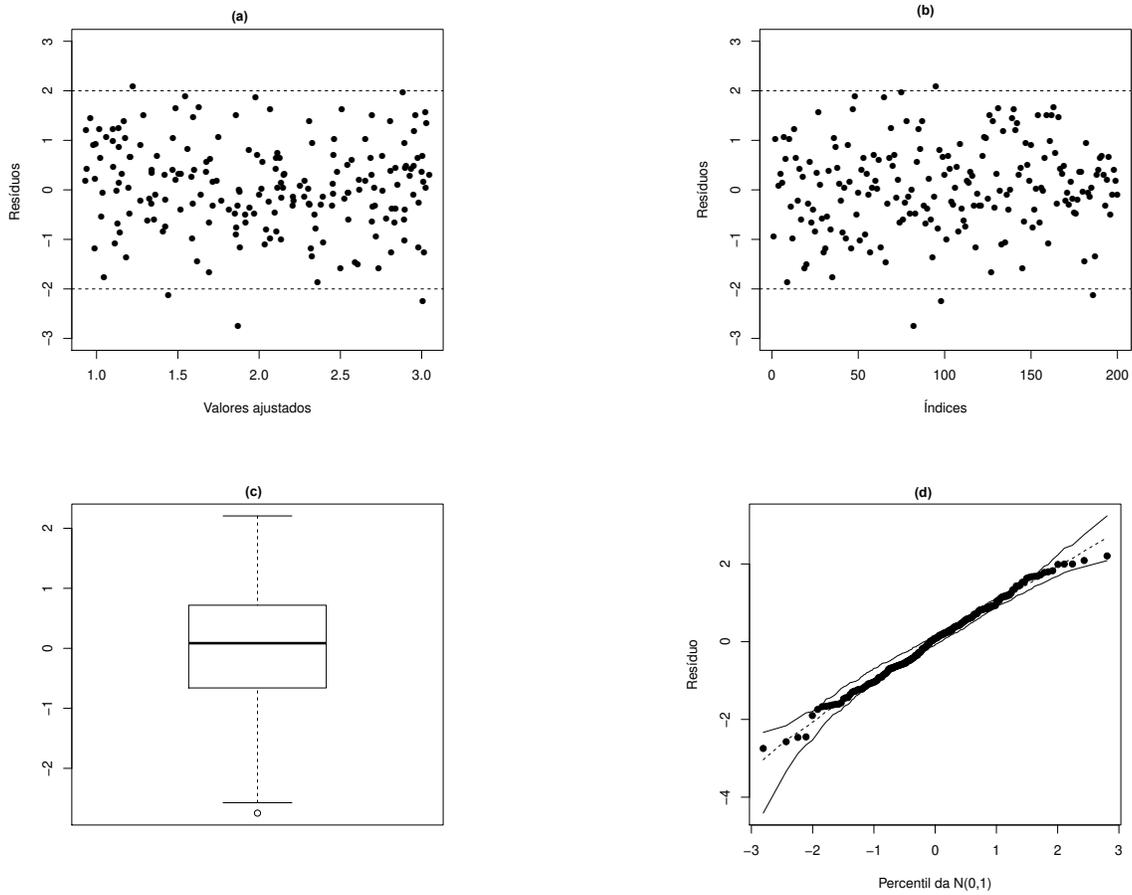


Figura 4.13: Análise de resíduos, obtidos através das estimativas frequentistas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BSNAC.

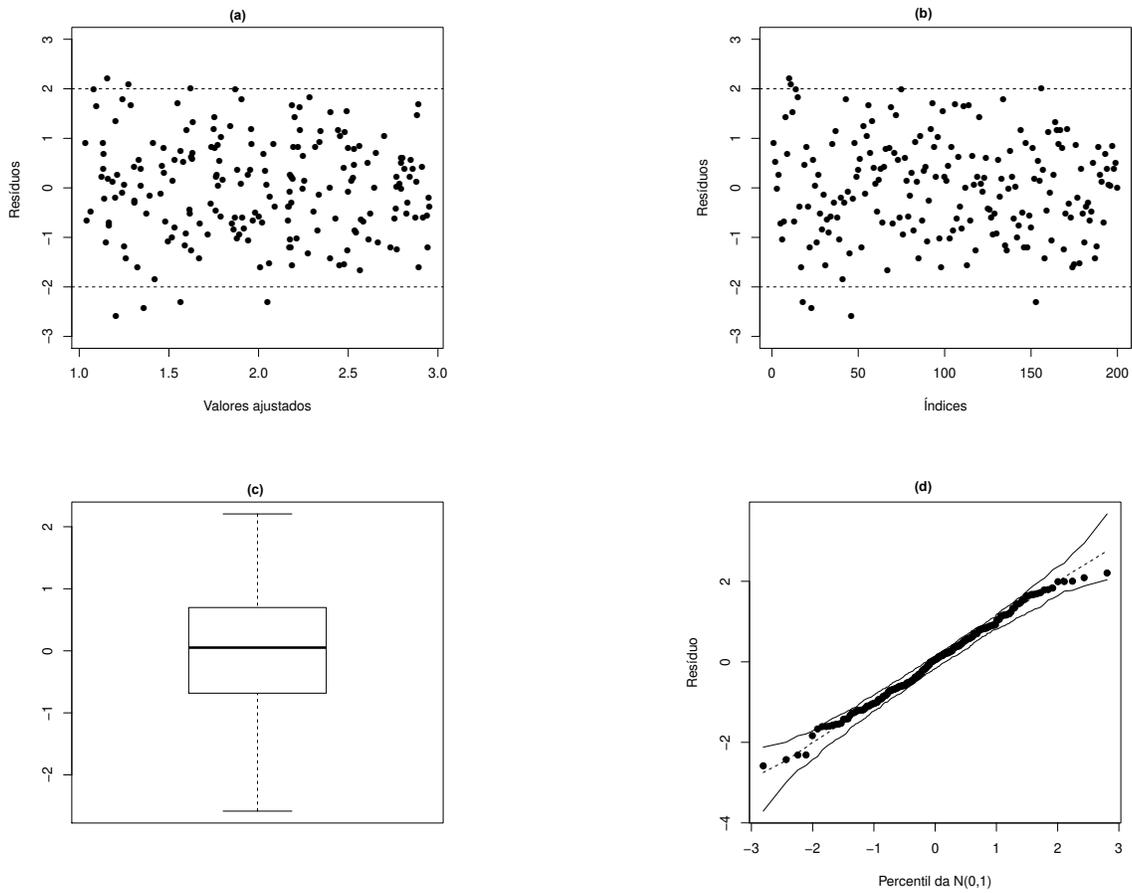


Figura 4.14: Análise de resíduos, obtidos através das estimativas bayesianas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BSNAC.

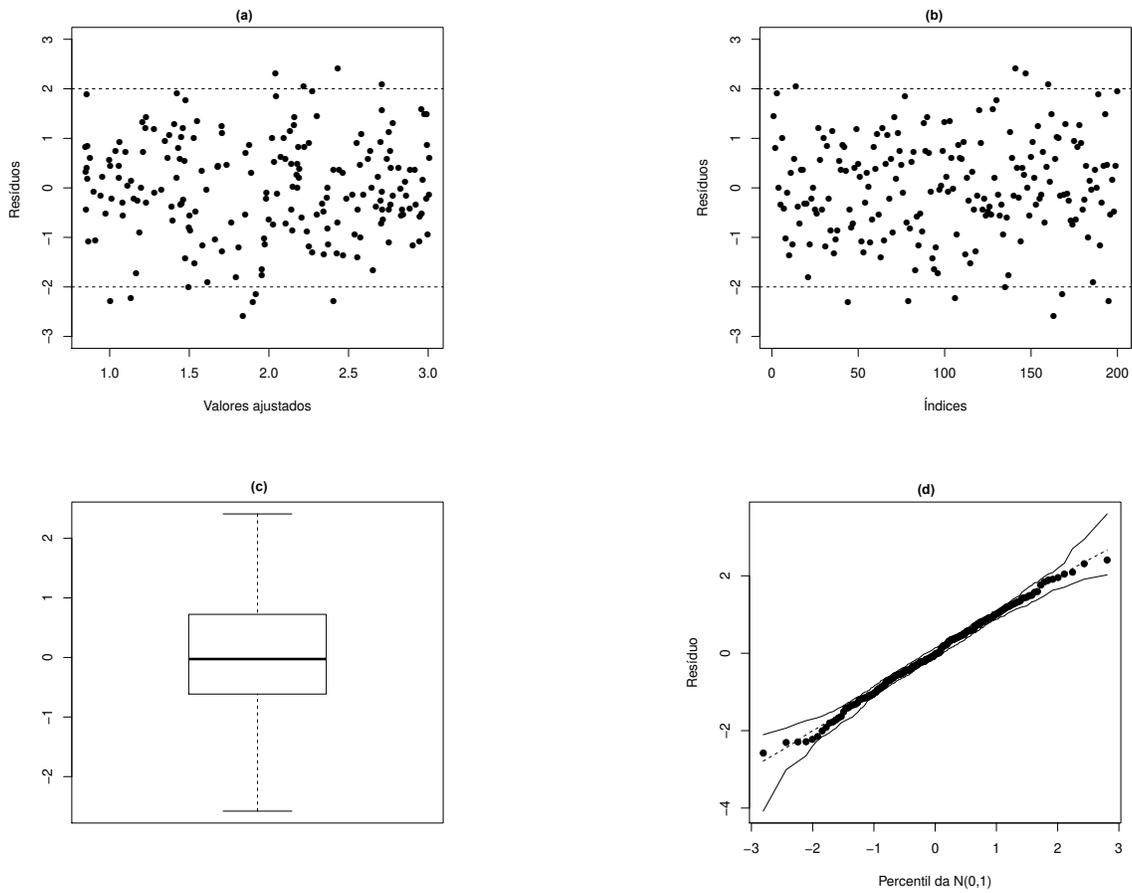


Figura 4.15: Análise de resíduos, obtidos através das estimativas frequentistas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS.

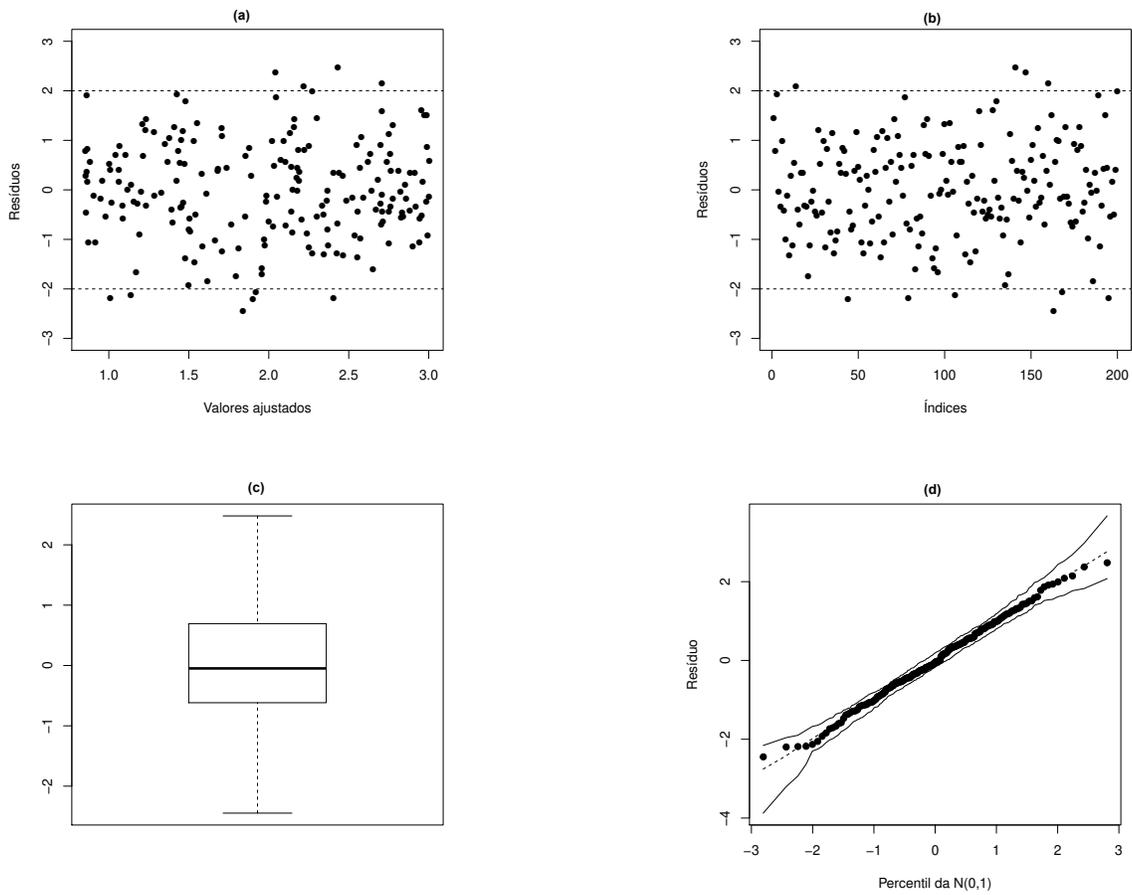


Figura 4.16: Análise de resíduos, obtidos através das estimativas bayesianas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS.

No cenário C3, observando as Figuras 4.17 (a) e (b), é possível perceber que os resíduos não estão aleatoriamente dispersos, visto que parece haver uma tendência. Além disso os resíduos tendem a se concentrar no intervalo $(-1,1)$. É possível notar ainda, que os resíduos apresentam uma leve tendência crescente e que a variabilidade parece aumentar consoante o aumento dos índices e valores ajustados, o mesmo pode-se concluir das Figuras 4.18 (a) e (b).

Nas Figuras 4.17 (c) e 4.18 (c), observa-se a presença de muitas observações discrepantes. Por fim, observando os gráficos referentes aos envelopes simulados, 4.17 (d) e 4.18 (d), nota-se que os resíduos concentram-se acima e abaixo da linha da média indicando assim, um mau ajuste do modelo às observações provenientes de uma distribuição com cauda pesada.

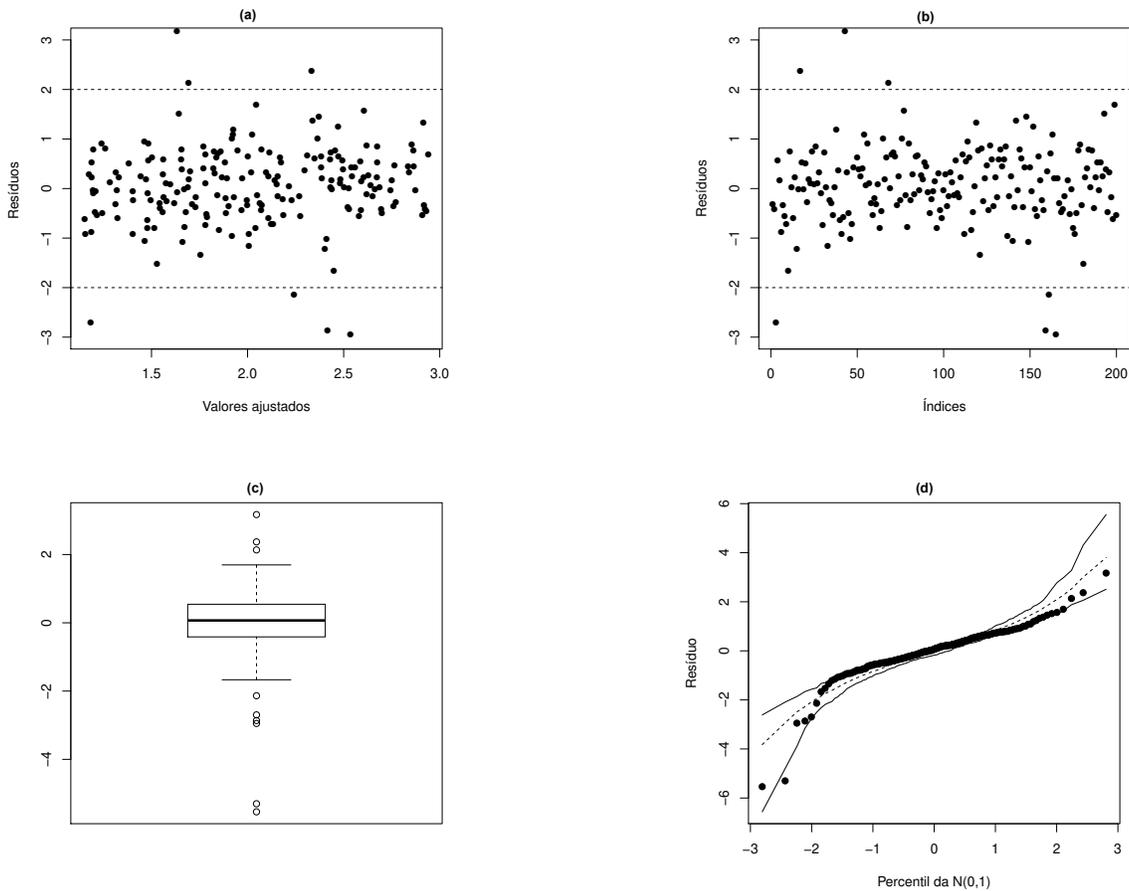


Figura 4.17: Análise de resíduos, obtidos através das estimativas frequentistas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS-t.

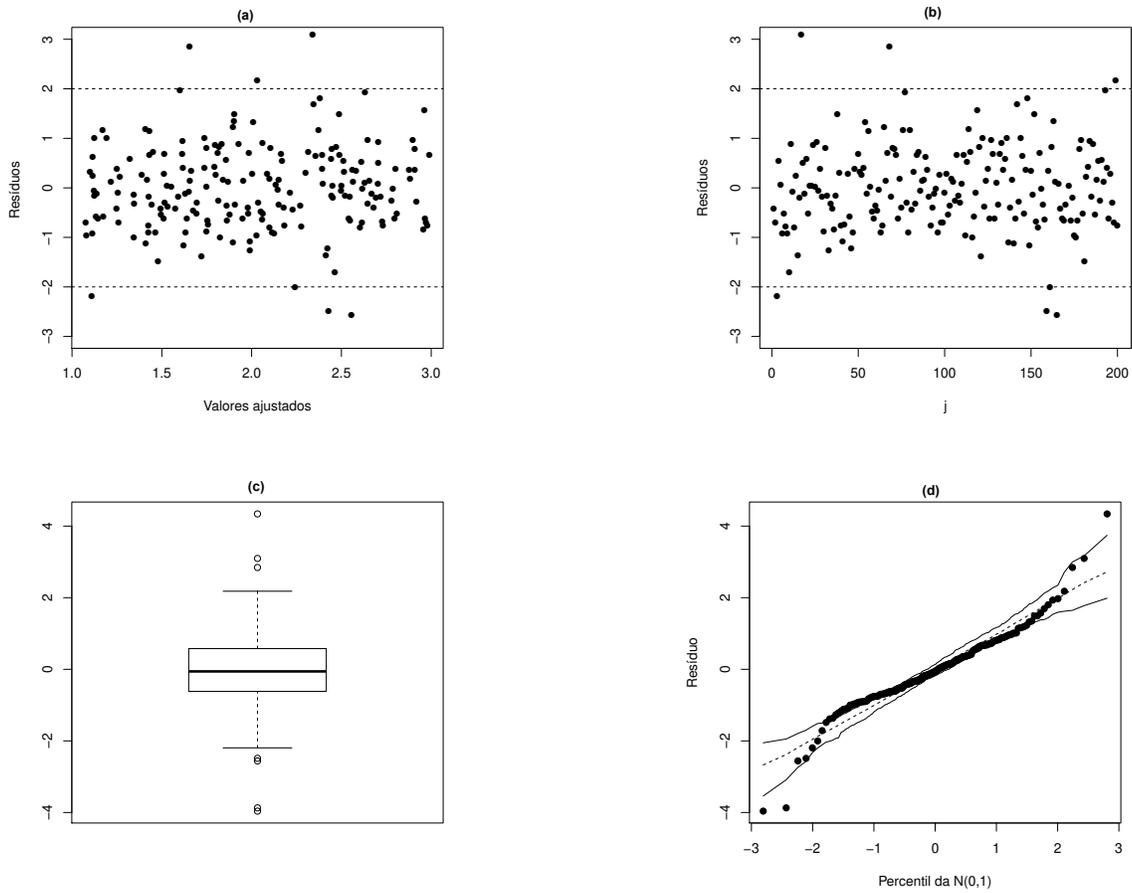


Figura 4.18: Análise de resíduos, obtidos através das estimativas bayesianas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS-t.

No cenário C4, observando as Figuras 4.19 (a) e (b), é possível perceber que os resíduos não estão aleatoriamente dispersos, visto que a variabilidade parece diminuir consoante o aumento dos índices e valores ajustados, o mesmo pode-se concluir das Figuras 4.20 (a) e (b).

Nas Figuras 4.19 (c) e 4.20 (c), observa-se a presença de observações discrepantes e além disso, nota-se ainda um comportamento assimétrico positivo dos resíduos.

Observando os gráficos referentes aos envelopes simulados, 4.19 (d) e 4.20 (d), nota-se alguns pontos fora das bandas de confiança superior. Além disso a disposição dos resíduos nos dá ideia de uma concavidade voltada para cima, evidenciando novamente um comportamento assimétrico positivo. Os gráficos do cenário C4 nos dão indícios de um mau ajuste do modelo às observações provenientes de uma distribuição assimétrica com cauda pesada.

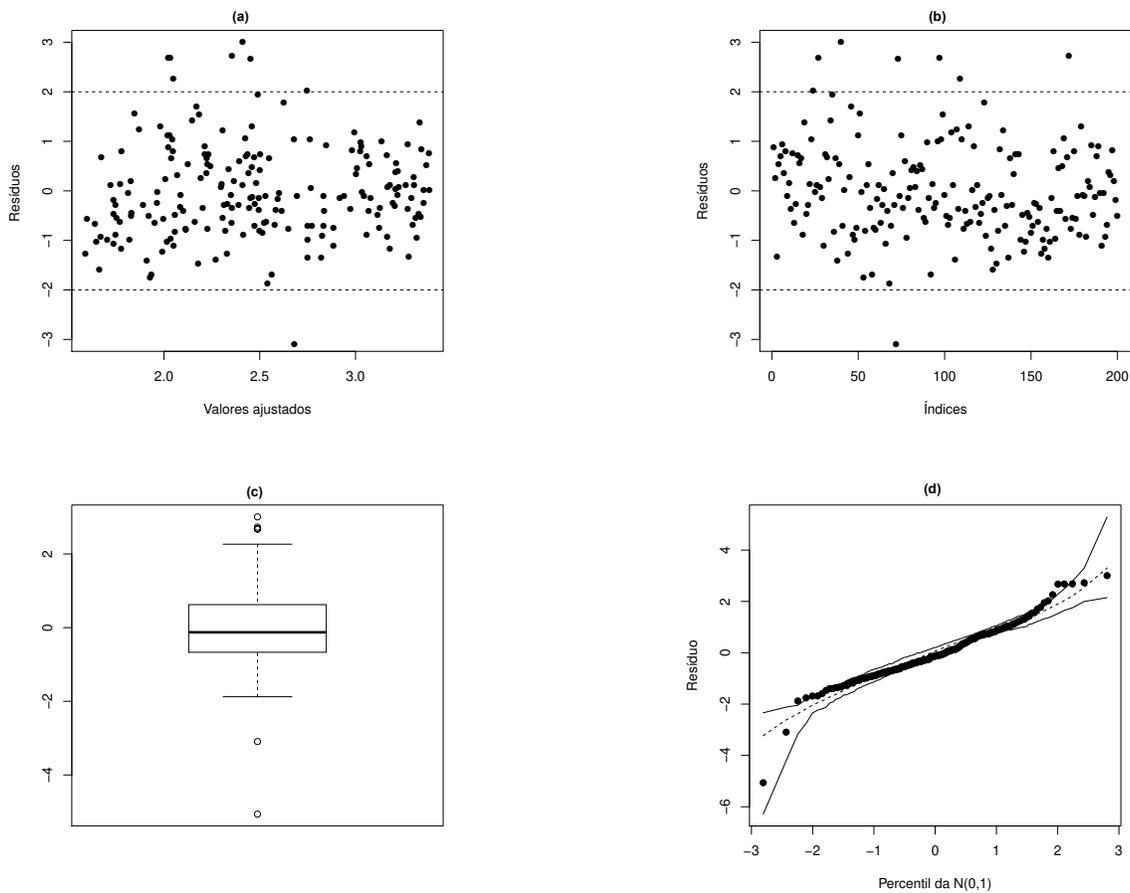


Figura 4.19: Análise de resíduos, obtidos através das estimativas frequentistas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS-Skew-t.

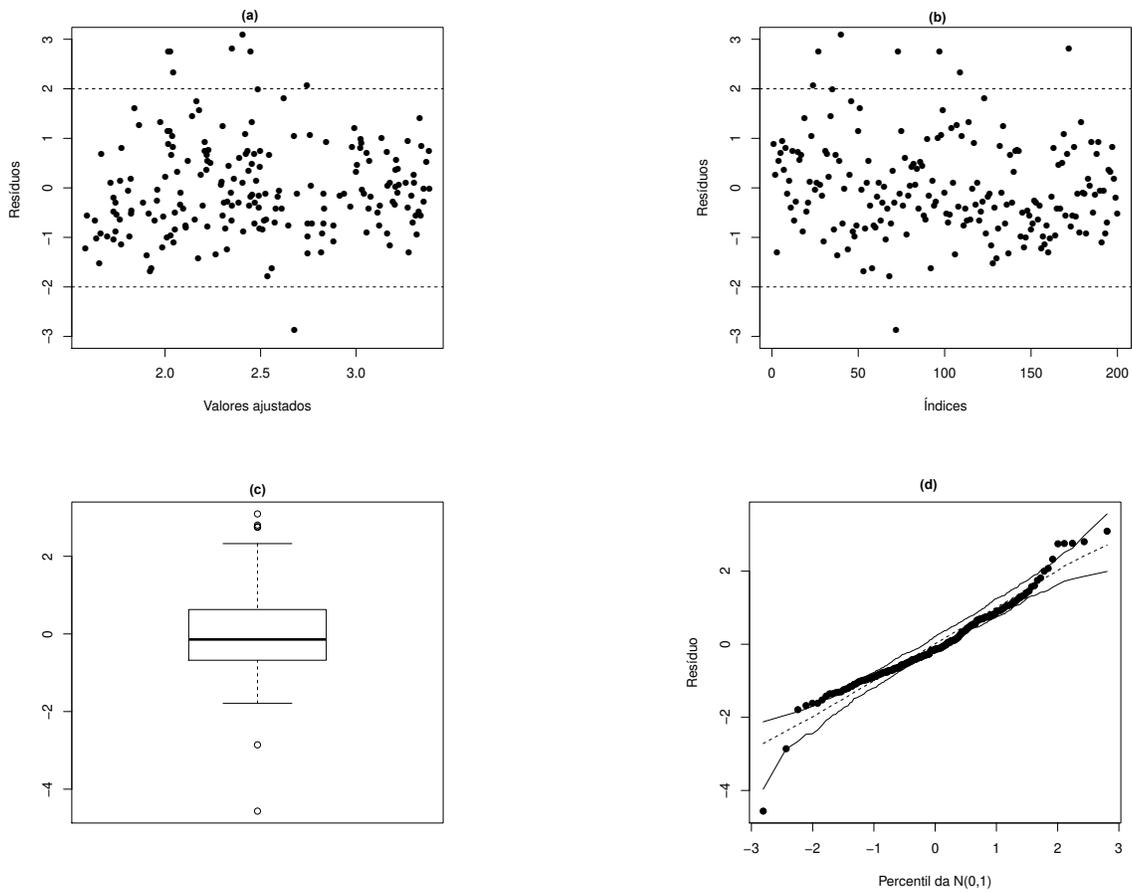


Figura 4.20: Análise de resíduos, obtidos através das estimativas bayesianas, do modelo log-BSNAC cuja as observações foram simuladas do modelo log-BS-Skew-t.

4.2.4 Análise de influência

Para ilustrar o uso de técnicas de diagnóstico de influência utilizando a divergência de Kullback-Leibler no contexto bayesiano, conduzimos um pequeno estudo de simulação e aplicamos a metodologia apresentada no Capítulo 3.

O estudo em questão consiste, inicialmente, em ajustar o modelo de regressão aqui proposto, log-BSNAC, à quatro diferentes conjunto de dados e em seguida, obter os gráficos das divergências de K-L e suas respectivas calibrações para cada uma das situações. Aqui, consideraremos como pontos influentes aqueles que apresentarem valores da divergência K-L significativamente maiores que os valores apresentados pelas demais observações do conjunto de dados e valores da calibração maiores do que 0,8.

Os conjuntos de dados foram obtidos a partir dos modelos log-BSNAC, log-BS, log-BS-t e log-BS-t assimétrica, assim como descrito na seção referente à análise residual. No nosso caso, para todos os modelos, consideramos $\alpha = 0,5$, $\boldsymbol{\varphi} = (1, 2)^\top$. Além disso para os modelos log-BSNAC e log-BS-t assimétrico consideramos $\gamma = 0,67$, para os modelos log-BS-t e log-BS-t assimétrico fixamos $\nu = 4$, em que ν representa os graus de liberdade da distribuição t-Student. Consideramos nesse estudo um tamanho de amostra $n=200$.

Analisando as Figuras 4.21 e 4.22 é possível notar que, quando ajustamos o modelo log-BSNAC aos conjuntos de dados simulados a partir dos modelos log-BSNAC e log-BS, identificamos apenas uma observação potencialmente influente, provavelmente devido à flutuações amostrais. Já quando simulamos do modelos log-BS-t e log-BS-t assimétrica, identificamos pelo menos três observações influentes. Os valores da divergência de K-L e da calibração, são significativamente menores para os casos que em simulamos dos dois primeiros modelos citados, indicando assim, que nesses casos, o modelo log-BSNAC parece ser menos influente à observações discrepantes.

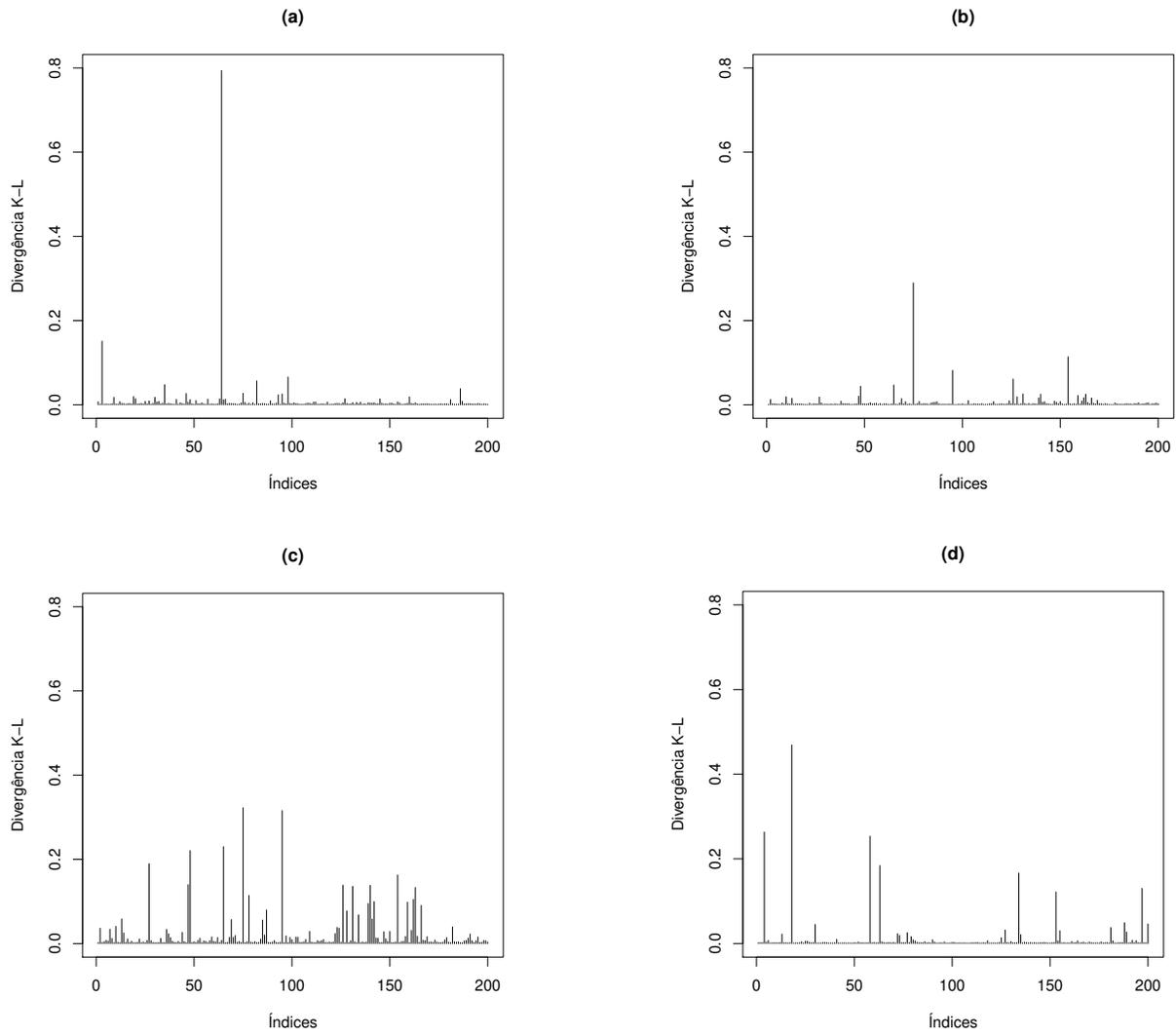


Figura 4.21: A divergência de K-L obtida a partir do ajuste do modelo log-BSNAC e de observações simuladas dos modelos: (a) log-BSNAC, (b) log-BS, (c) log-BS-t, (d) log-BS-t assimétrica.

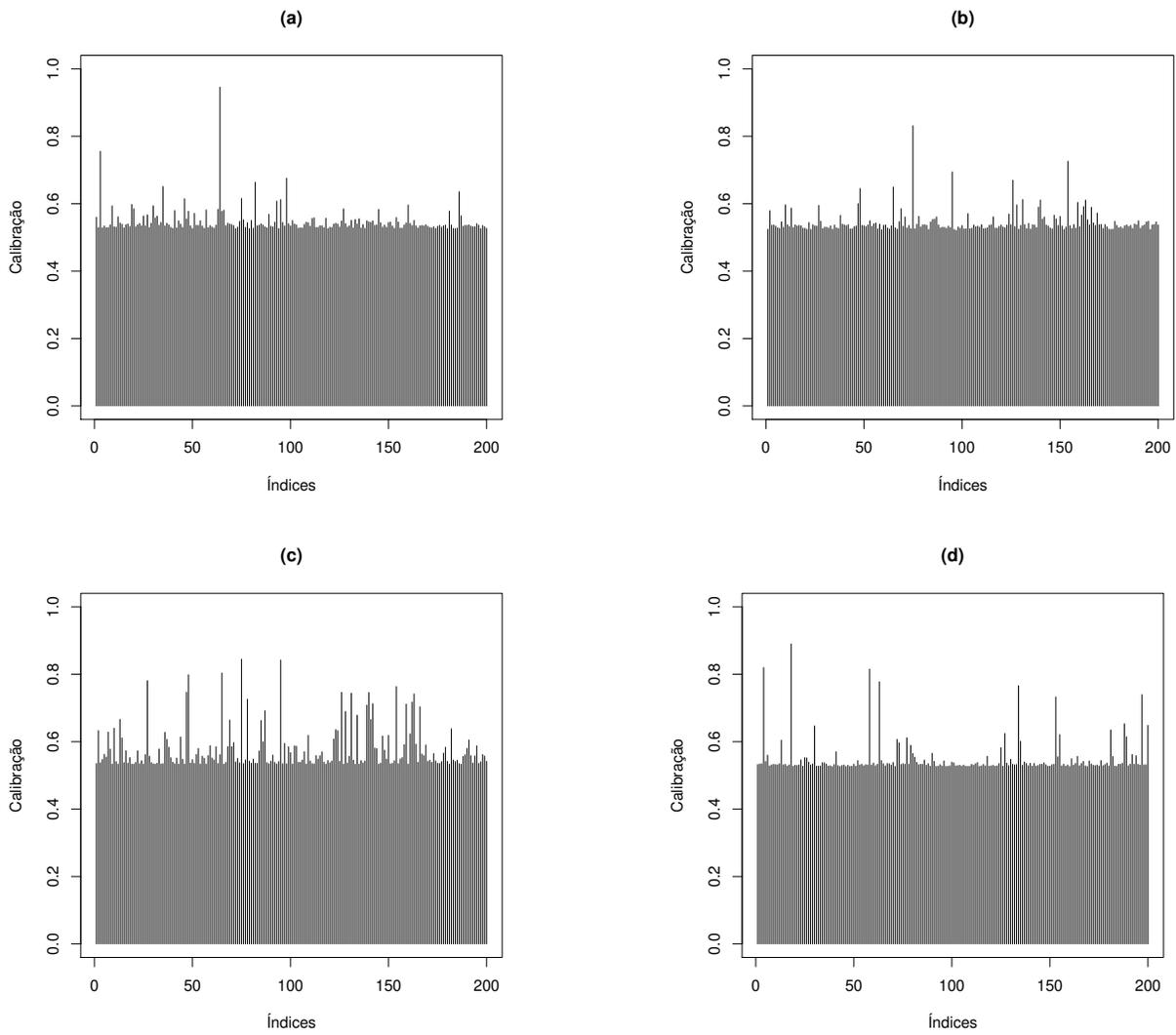


Figura 4.22: Calibração da divergência de K-L para os quatro conjunto de dados simulados a partir dos modelos: (a) log-BSNAC, (b) log-BS, (c) log-BS-t, (d) log-BS-t assimétrica.

4.2.5 Seleção de modelos

A fim de verificar a performance das estatísticas de comparação de modelos, apresentadas no Capítulo 3, na escolha do modelo correto, realizamos um estudo de simulação considerando dois diferentes cenários.

No primeiro cenário, simulamos observações do modelo de regressão log-BSNAC, fixando $\alpha = 0,5$, $\varphi = (1, 2)^\top$ e $\gamma = 0,67$, e ajustamos tanto o modelo aqui proposto, log-BSNAC, quanto o modelo de regressão log-BS. Já no segundo cenário, simulamos observações do modelo log-BS, fixando $\alpha = 0,5$, $\varphi = (1, 2)^\top$, e ajustamos novamente ambos os modelos. Para os dois cenários considerados, calculamos 100 vezes cada um dos critérios de seleção de modelos AIC e BIC, sob o ponto de vista frequentista, e os critérios bayesianos EAIC, EBIC, DIC e LPML, e verificamos em quantas delas o modelo verdadeiro foi escolhido.

No primeiro cenário, para um tamanho de amostra $n = 50$, os critérios AIC, BIC, EAIC, EBIC, DIC e LMPL, escolheram o modelo log-BSNAC como preferível 97, 95, 99, 97, 100 e 99 vezes, respectivamente. Já para um tamanho de amostra $n = 200$, os critérios AIC e BIC escolheram o log-BSNAC 98 vezes e os critérios EAIC, EBIC, DIC e LPML em todas as 100 vezes. Na Tabela 4.22, apresentamos a média calculada para cada um dos critérios.

Tabela 4.22: Médias dos critérios - Modelo verdadeiro: log-BSNAC.

Modelo ajustado	n	AIC	BIC	EAIC	EBIC	DIC	LPML
log-BSNAC	50	69,481	77,129	80,825	88,473	215,656	-38,630
	200	270,105	283,299	297,969	311,162	867,065	-146,664
log-BS	50	73,389	79,125	85,424	91,160	235,421	-41,268
	200	282,964	292,859	329,219	339,114	967,708	-162,727

No segundo cenário, para um tamanho de amostra $n = 50$, os critérios AIC, BIC, EAIC, EBIC, DIC e LMPL, escolheram o modelo log-BS como preferível 97, 99, 95, 96, 100 e 91 vezes, respectivamente. Já para um tamanho de amostra $n = 200$, os critérios AIC e BIC escolheram o log-BSNAC 100 vezes e os critérios EAIC, EBIC, DIC e LPML escolheram 97, 99, 100, 95 vezes, respectivamente. Na Tabela 4.23, apresentamos a média calculada para cada um dos critérios.

Tabela 4.23: Médias dos critérios - Modelo verdadeiro: log-BS.

Modelo ajustado	n	AIC	BIC	EAIC	EBIC	DIC	LPML
log-BSNAC	50	74,930	82,578	84,529	92,177	225,907	-40,521
	200	287,320	300,514	303,617	316,810	877,804	-152,385
log-BS	50	73,945	79,681	85,513	91,249	234,210	-41,534
	200	284,139	294,034	303,751	313,646	883,921	-152,953

Podemos concluir assim, que o modelo log-BSNAC ajusta-se melhor, do que o modelo log-BS original, a conjuntos de observações que apresentam assimetria, segundo as estatísticas de seleção de modelos aqui utilizadas.

Capítulo 5

Aplicação

Neste capítulo, ilustraremos as metodologias aqui propostas com conjuntos de dados reais encontrados na literatura. Mais especificamente, na Seção 5.1, apresentaremos uma aplicação para o modelo probabilístico BSNAC, e na Seção 5.2, apresentaremos duas aplicações para o modelo de regressão log-BSNAC.

5.1 Aplicação I

Para ilustrar brevemente a metodologia proposta no Capítulo 2, referente ao modelo probabilístico BSNAC, iremos considerar o conjunto de dados obtido originalmente por Volle (1985) e analisado por Lemonte et al. (2012), que diz respeito ao tempo gasto (em horas) em duas categorias de atividades durante 100 dias no ano 1976. Dentre as duas categorias de atividades estudadas, optamos por modelar, via modelos BS e BSNAC, a quantidade de tempo gasto em alimentação pelos 28 indivíduos envolvidos no estudo.

Na Tabela 5.1 e na Figura 5.1 apresentamos as estatísticas descritivas e o histograma do tempo gasto (em horas), respectivamente. Atavés do gráfico do histograma, é possível notar que as observações se concentram na classe que diz respeito aos tempos de 0 à 200 horas, indicando uma assimetria à direita (positiva).

O nosso objetivo consiste em verificar qual dos dois modelos melhor se ajusta ao conjunto de dados em estudo. Na Tabela 5.2, apresentamos as estimativas, sob as abordagens frequentista e bayesiana, bem como seus respectivos erros-padrão, intervalos de confiança e de credibilidade. As estimativas aqui apresentadas foram obtidas através dos algoritmos ECM e MCMC apresentados no Capítulo 2. O algoritmo MCMC foi implementado através do pacote `R2openBugs` do software R, considerando o conjunto de prioris apresentados, mais precisamente $\alpha \sim \text{Gama}(1; 0, 8)$, $\beta \sim \text{Gama}(0, 001; 0, 001)$ e $\gamma \sim U(-0, 99527; 0, 99527)$.

Tabela 5.1: Estatísticas descritivas do tempo gasto (em horas) em alimentação.

Média	Mediana	Mínimo	Máximo	DP	Assimetria	Curtose
294,286	130	100	960	330,529	1,317	-0.229

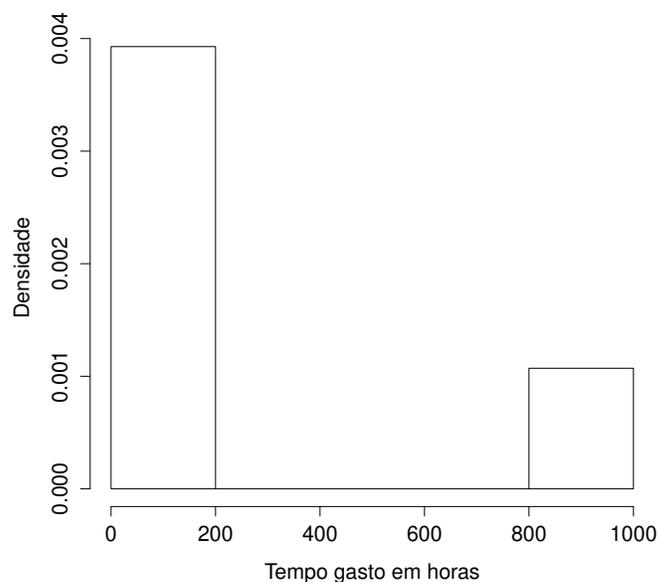


Figura 5.1: Boxplot do tempo gasto (em horas) em alimentação.

Tabela 5.2: Estimativas dos parâmetros e seus respectivos erros-padrão, intervalos de confiança e de credibilidade para os modelos BSNAC e BS.

Abordagem	Parâmetros	BSNAC			BS		
		Estimativas	EP	IC _{95%}	Estimativas	EP	IC _{95%}
Frequentista	α	0,692	0,102	[0,492; 0,892]	0,899	0,126	[0,6252; 1,146]
	β	209,569	26,127	[158,387; 260,805]	214,255	33,543	[148,511; 279,999]
	γ	0,919	0,032	[0,829; 0,982]	-	-	-
Bayesiana	α	0,773	0,115	[0,595; 1,036]	0,942	0,137	[0,721; 1,250]
	β	235,269	32,547	[185,482; 314,312]	216,966	35,616	[154,697; 294,925]
	γ	0,949	0,063	[0,781; 0,995]	-	-	-

As estimativas obtidas para α e γ , quando utilizamos o modelo BSNAC para modelar o tempo (em horas) em alimentação, resultam em um comportamento fortemente assimétrico positivo da distribuição BSNAC, visto que tal comportamento pode ser notado quando fixamos um valor para α grande e/ou um valor para γ grande e positivo. Muito embora a distribuição BS possa apresentar um comportamento assimétrico consoante aos valores dos parâmetros, a distribuição

BSNAC possui um parâmetro a mais que controla a assimetria da distribuição, tornando o modelo aqui proposto mais adequado para modelar observações com essa característica.

Na Tabela 5.3 e na Figura 5.2, apresentamos, respectivamente, os critérios para seleção de modelos e o histograma com as curvas ajustadas para ambos os modelos. Observando a Tabela 5.3 é possível perceber que todos os critérios, obtidos via abordagens frequentista e bayesiana, apontam o modelo probabilístico BSNAC como preferível para modelar o tempo gasto (em horas) em alimentação no estudo em questão. Além disso, é possível notar na Figura 5.2 que o modelo BSNAC acomoda melhor a assimetria positiva das observações do que o modelo BS.

Tabela 5.3: Critérios para seleção de modelos.

Critérios	BSNAC	BS
AIC	349,615	366,485
BIC	353,612	369,149
EAIC	349,087	368,603
EBIC	353,083	371,268
DIC	1028,252	1091,813
LPML	-172,341	-183,194

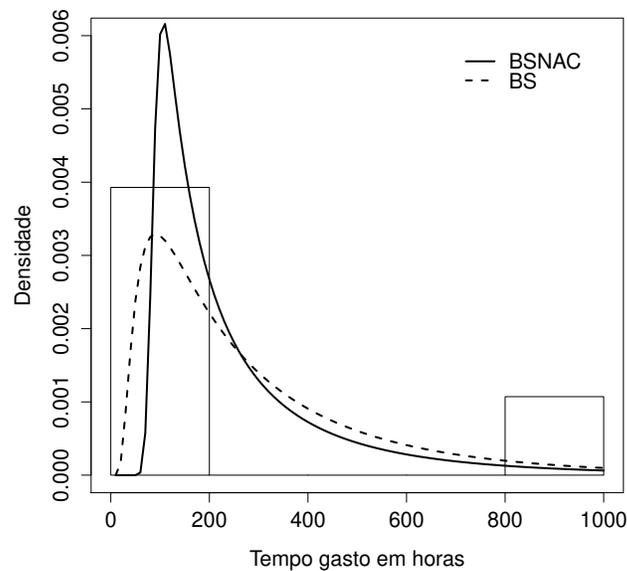


Figura 5.2: Boxplot do tempo gasto (em horas) em alimentação com as curvas ajustadas dos modelos BSNAC e BS.

5.2 Aplicação II

Nesta seção, ilustraremos a metodologia desenvolvida no Capítulos 3 para o modelo de regressão log-BSNAC utilizando dois conjuntos de dados reais.

5.2.1 Fadiga Biaxial (“do eixo duplo”)

Para esta aplicação vamos considerar o conjunto de dados analisado por Rieck and Nedelman (1991), Xie e Wei (2007), Santana et al. (2011) e Farias and Lemonte (2011), que diz respeito ao tempo de vida de um pedaço de metal sujeito à compressões cíclicas até falhar. A variável resposta N representa o número de ciclos até que ocorra a falha e a variável explicativa X é o trabalho por ciclo medido em mJ/m^3 . Os dados são apresentados na Tabela B.1. As estatísticas descritivas referentes a cada uma das variáveis são apresentadas na Tabela 5.4.

Tabela 5.4: Estatísticas descritivas para o número de ciclos, N , até que ocorra a falha e para o trabalho por ciclo, X .

	Mínimo	Média	Máximo	Q1	Mediana	Q3
N	125	943,652	5046	242,753	566	1086,257
X	11,524	40,293	100,511	24	33,276	60,125

Sob normalidade, esse conjunto de dados foi analisado por Rieck e Nedelman (1991) considerando o seguinte modelo de regressão

$$V_i = \varphi_0 + \varphi_1 \log(x_i) + \varepsilon_i, \quad (5.2.1)$$

em que $V_i = \log(N_i)$ e $\varepsilon_i \sim SHN(\alpha, 0, \sigma = 2)$, $i = 1, \dots, 46$.

Assim como proposto em Santana (2009), vamos considerar o seguinte modelo de regressão

$$\begin{aligned} V_i &= \varphi_0 + \varphi_1 \exp(-\vartheta_0/x_i) + \varepsilon_i \\ &= \varphi_0 + \varphi_1 g(x_i) + \varepsilon_i, i = 1, \dots, 46, \end{aligned}$$

em que ϑ_0 é uma quantidade conhecida. A quantidade ϑ_0 é obtida através da minimização da equação de mínimos quadrados dada por

$$MQ_k(\boldsymbol{\theta}) = \sum_{i=1}^n [v_{ik} - \hat{\varphi}_{0k} - \hat{\varphi}_{1k} \exp(-\vartheta_k/x_i)]^2, k = 1, \dots, m, \quad (5.2.2)$$

em que ϑ_k correspondem a m valores previamente gerados para ϑ . Ou seja, o valor de ϑ_0 escolhido será o $\min(MQ_k(\boldsymbol{\theta}))$, $k = 1, \dots, m$. Os valores $\hat{\varphi}_0$ e $\hat{\varphi}_1$ foram obtidos usando as estimativas de mínimos quadrados. Nesse caso, ainda segundo Santana (2009), o valor obtido para ϑ_0 foi 20,076.

Portanto, o modelo utilizado nessa aplicação será

$$V_i = \varphi_0 + \varphi_1 \exp(-20,076/x_i) + \varepsilon_i, i = 1, \dots, 46, \quad (5.2.3)$$

em que $\varepsilon_i \sim \text{SHN}(\alpha, 0, \sigma = 2)$.

Na Figura 5.3 (a)-(b) apresentamos o gráfico de dispersão e a reta ajustada do modelo (5.2.1) e o gráfico de dispersão e a reta ajustada do modelo (5.2.3), respectivamente. Note que o modelo (5.2.3) também parece apropriado para esse conjunto de dados.

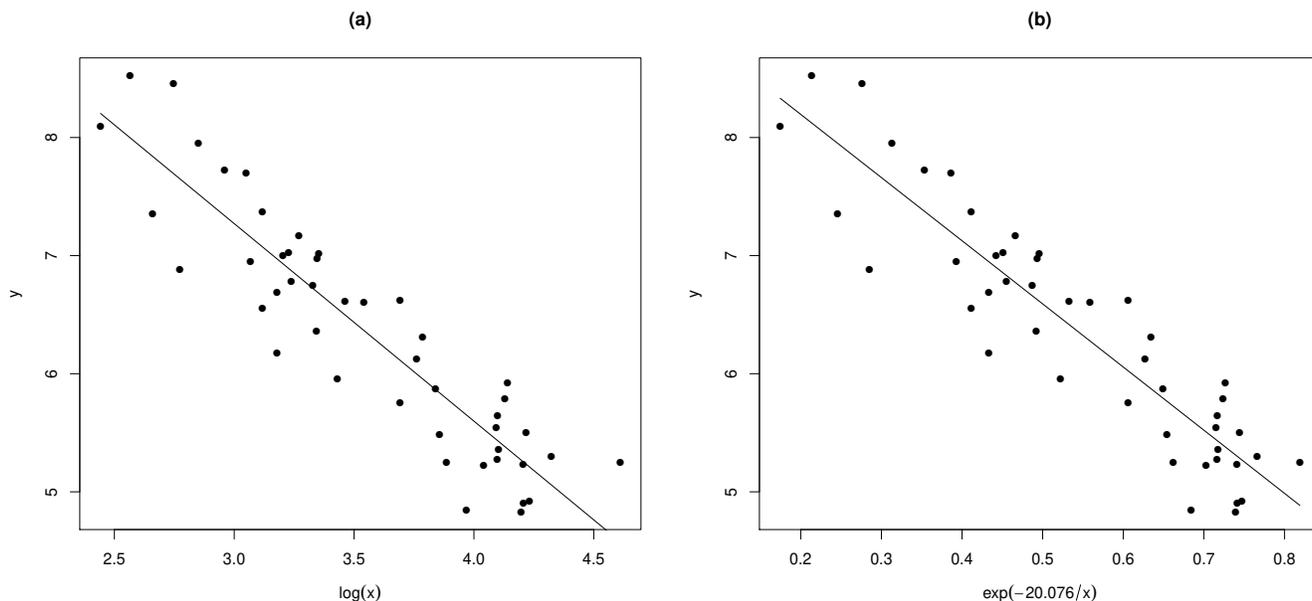


Figura 5.3: Gráficos de dispersão e retas ajustadas considerando os modelos - (a) modelo dado em (5.2.1) e (b) modelo dado em (5.2.3).

Vamos supor agora que em (5.2.3), $\varepsilon_i \sim \text{SHNAC}(\alpha, 0, \sigma = 2, \gamma)$. Nesse caso, temos o seguinte modelo,

$$V_i = \varphi_0 + \varphi_1 \exp(-20,076/x_i) + \varepsilon_i, i = 1, \dots, 46, \quad (5.2.4)$$

em que $\varepsilon_i \sim \text{SHNAC}(\alpha, 0, \sigma = 2, \gamma)$. O nosso objetivo consiste em verificar qual dos dois modelos melhor se ajusta ao conjunto de dados em estudo. Na Tabela 5.5, apresentamos as estimativas, sob as abordagens frequentista e bayesiana, bem como seus respectivos erros-padrão, intervalos de confiança e de credibilidade. As estimativas aqui apresentadas foram obtidas através dos algoritmos ECM e MCMC apresentados no Capítulo 3. O algoritmo MCMC foi implementado através do pacote `R2OpenBugs` do software R, considerando o conjunto de prioris apresentados, mais especificamente, $\alpha \sim \text{Gama}(1; 0, 8)$; $\varphi_k = N(0, 1000)$, $k = 1, 2$ e $\gamma \sim U(-0, 99527; 0, 99527)$.

Com o intuito de respeitar o espaço paramétrico do parâmetro de assimetria, truncamos o intervalo de confiança à esquerda do valor $-0, 99527$. Observando os intervalos de confiança e de credibilidade do parâmetro γ , pode-se observar que o parâmetro de assimetria é significativo.

Em termos dos parâmetros de regressão, se observarmos, por exemplo, os resultados frequentistas referentes ao modelo log-BSNAC, podemos concluir que a média do logaritmo do número de

ciclos é igual 9,408, quando observa-se o valor da função do trabalho por ciclo, $g_1(x_i)$, igual a zero. Além disso, há um decréscimo de 5,586 na média do logaritmo do número de ciclos, à medida que o valor da função do trabalho por ciclo, $g_1(x_i)$, sofre um aumento de uma unidade.

Tabela 5.5: Estimativas dos parâmetros e seus respectivos erros-padrão, intervalos de confiança e de credibilidade para os modelos log-BSNAC e log-BS.

Abordagem	Parâmetros	log-BSNAC			log-BS		
		Estimativas	EP	IC _{95%}	Estimativas	EP	IC _{95%}
Frequentista	α	0,428	0,036	[0,357; 0,498]	0,400	0,042	[0,318; 0,482]
	φ_0	9,408	0,167	[9,081; 9,735]	9,254	0,201	[8,856; 9,644]
	φ_1	-5,586	0,293	[-6,154; -5,001]	-5,335	0,351	[-6,023; -4,647]
	γ	-0,949	0,121	[-0,995; -0,712]	-	-	-
Bayesiana	α	0,424	0,050	[0,339; 0,541]	0,422	0,047	[0,345; 0,526]
	φ_0	9,344	0,177	[8,999; 9,684]	9,239	0,206	[8,825; 9,657]
	φ_1	-5,493	0,307	[-6,086; -4,888]	-5,304	0,365	[-6,050; -4,559]
	γ	-0,602	0,303	[-0,953; -0,119]	-	-	-

Na Tabela 5.5, apresentamos os critérios para seleção de modelos, obtidos através da metodologia apresentada na Seção 3.6.2. É possível perceber que todos os critérios, exceto o EBIC, apontam o modelo log-BSNAC como o que melhor se ajusta a esse conjunto de dados. Muito embora o EBIC indique o log-BS como preferível nesse caso, vale ressaltar que a diferença desse critério entre os modelos em questão é muito pequena.

Tabela 5.6: Critérios para seleção de modelos.

Critérios	log-BSNAC	log-BS
AIC	47,715	50,574
BIC	55,029	56,060
EAIC	52,195	53,645
EBIC	59,509	59,131
DIC	130,001	140,125
LPML	-23,857	-25,118

Nas Figuras 5.4 e 5.5, apresentamos a análise de resíduos do modelo log-BSNAC considerando os resíduos calculados a partir das estimativas frequentistas e bayesianas, respectivamente. Já nas Figuras 5.6 e 5.7, apresentamos a análise de resíduos do modelo log-BS considerando os resíduos calculados a partir das estimativas frequentistas e bayesianas, respectivamente.

Ao observarmos as Figuras 5.4 (a)-(b) é possível notar que os pontos parecem estar aleatoriamente dispersos e parece não haver tendências, o mesmo pode-se concluir das Figuras 5.5 (a)-(b). Já as Figuras 5.6 (a)-(b) nos dão indícios do aumento da variabilidade dos resíduos consoante ao aumento dos valores ajustados, e da diminuição da variabilidade consoante ao aumento dos índices, respectivamente. As Figuras 5.7 (a)-(b) nos dão indícios de uma tendência ora crescente, ora decrescente dos resíduos do modelo log-BS.

As Figuras 5.4 (c) 5.5 (c) nos dão evidências da simetria dos resíduos do modelo log-BSNAC, enquanto que as Figuras 5.6 (c) e 5.7 (c), nos dão indícios de uma leve assimetria à direita dos resíduos do modelo log-BS. Por fim, nas Figuras 5.6 (d) e 5.7 (d) é possível perceber que os resíduos tendem a sair das bandas de confiança, enquanto que nas Figuras 5.4 (d) e 5.5 (d) nota-se que os resíduos encontram-se todos dentro das bandas de confiança e parece não haver tendências.

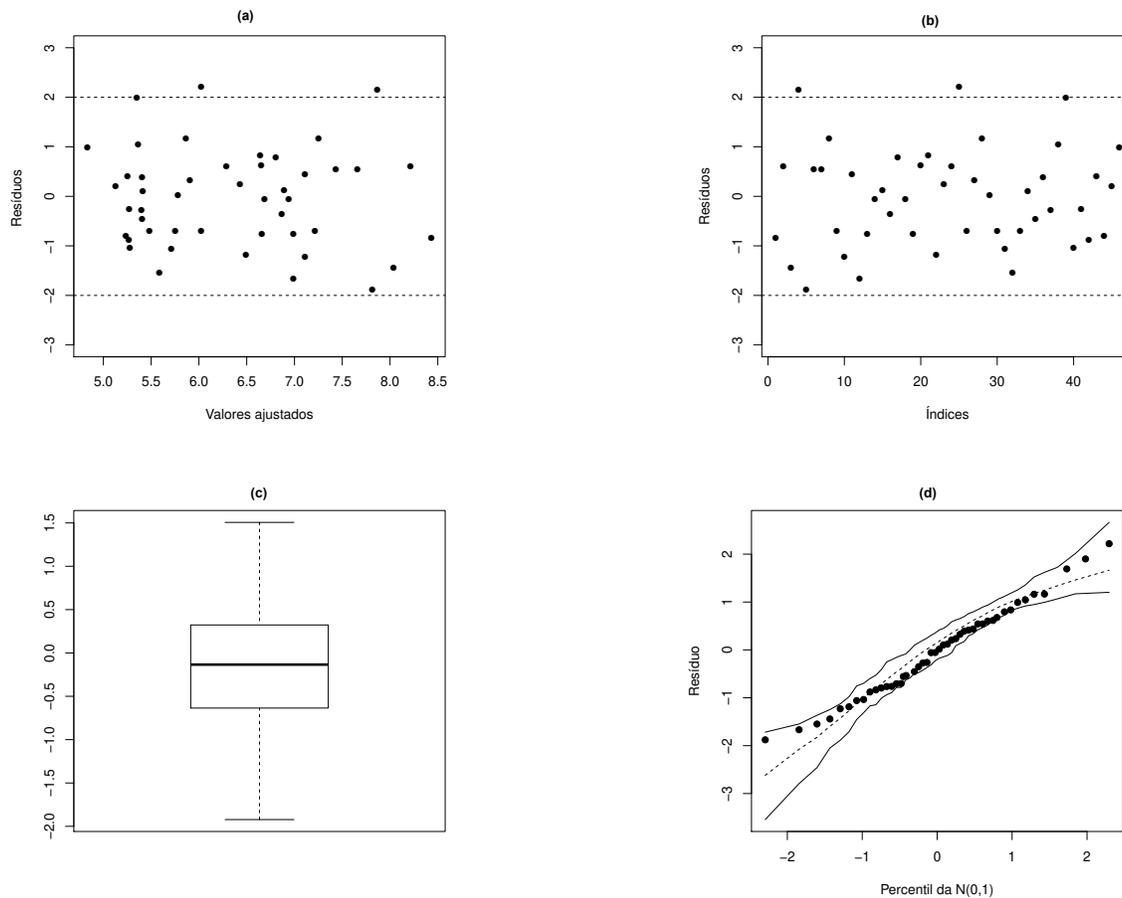


Figura 5.4: Análise de resíduos, calculados a partir das estimativas frequentistas, do modelo log-BSNAC.

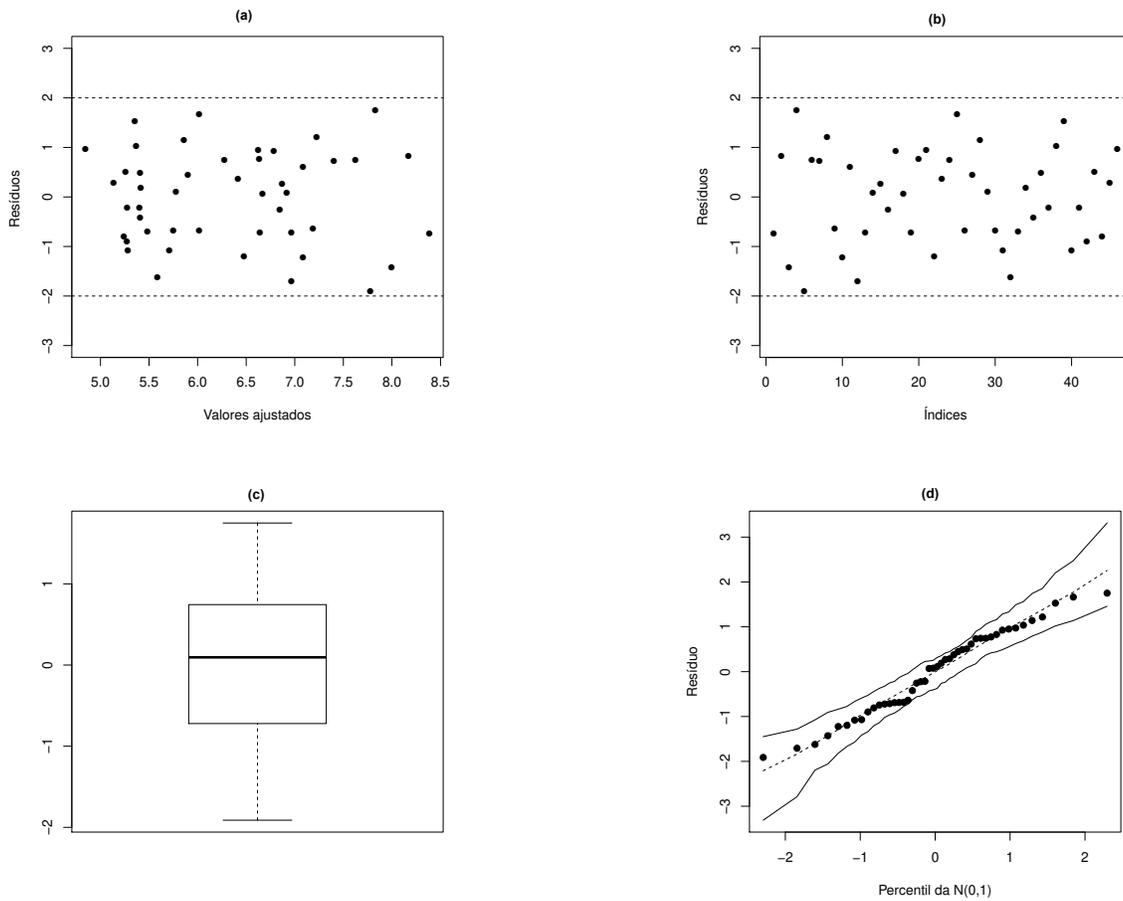


Figura 5.5: Análise de resíduos, calculados a partir das estimativas bayesianas, do modelo log-BSNAC.

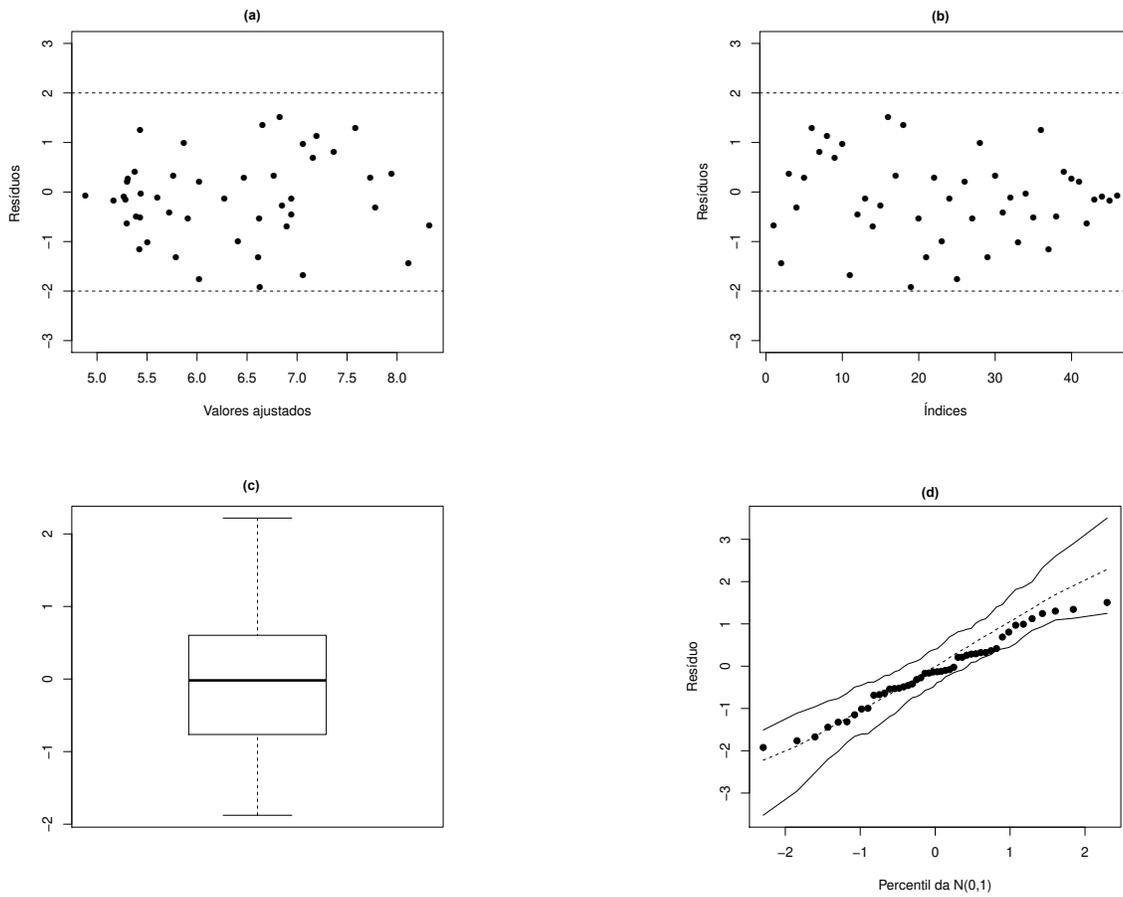


Figura 5.6: Análise de resíduos, calculados a partir das estimativas frequentistas, do modelo log-BS.

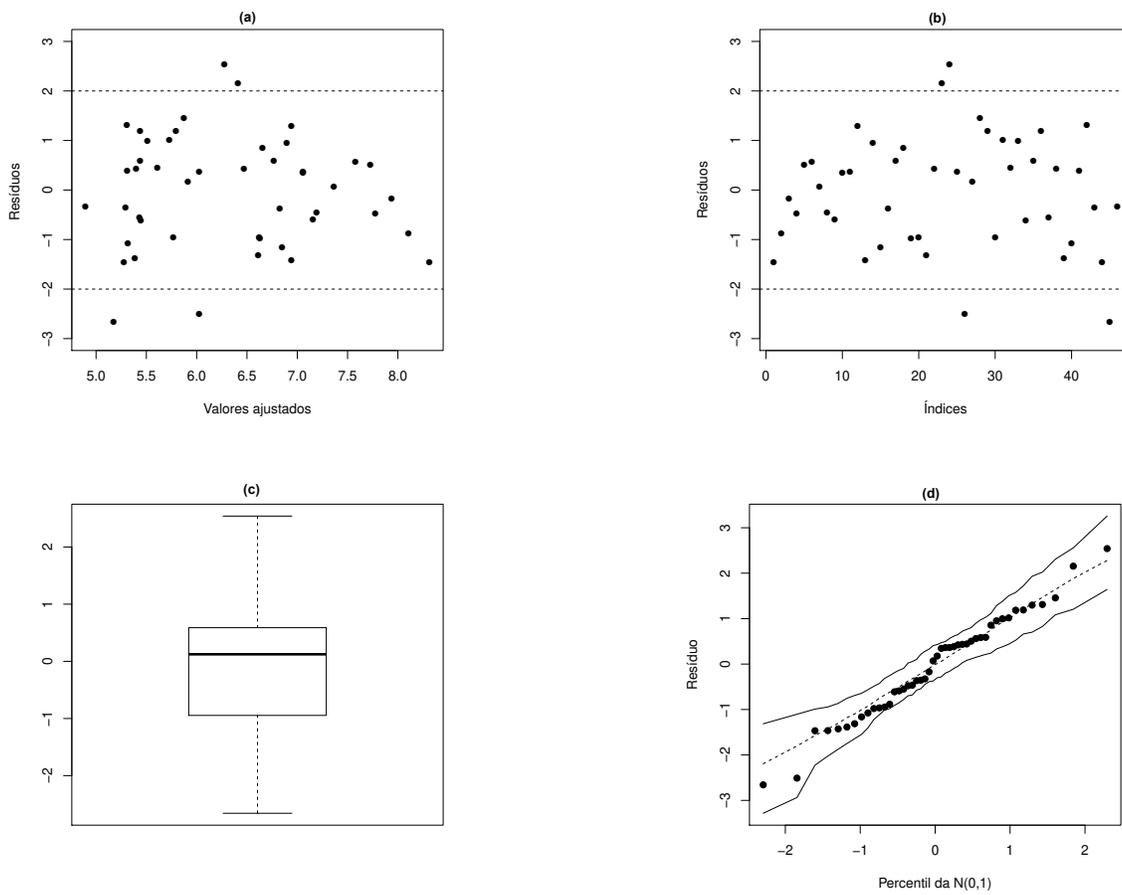


Figura 5.7: Análise de resíduos, calculados a partir das estimativas bayesianas, do modelo log-BS.

Com o intuito de identificar observações influentes, fizemos uma análise de influência utilizando a divergência de Kullback-Leibler assim como descrita na Seção 3.6.3. As Figuras 5.8 e 5.9 dizem respeito as divergências de K-L e suas respectivas calibrações para os modelos log-BSNAC e log-BS, nesse caso, iremos considerar como observação influente aquela que apresentar uma calibração maior ou igual a 0,8. Em sendo assim, a análise de influência detectou a observação 4 como potencialmente influente para o modelo log-BSNAC e a observação 5 para o modelo log-BS.

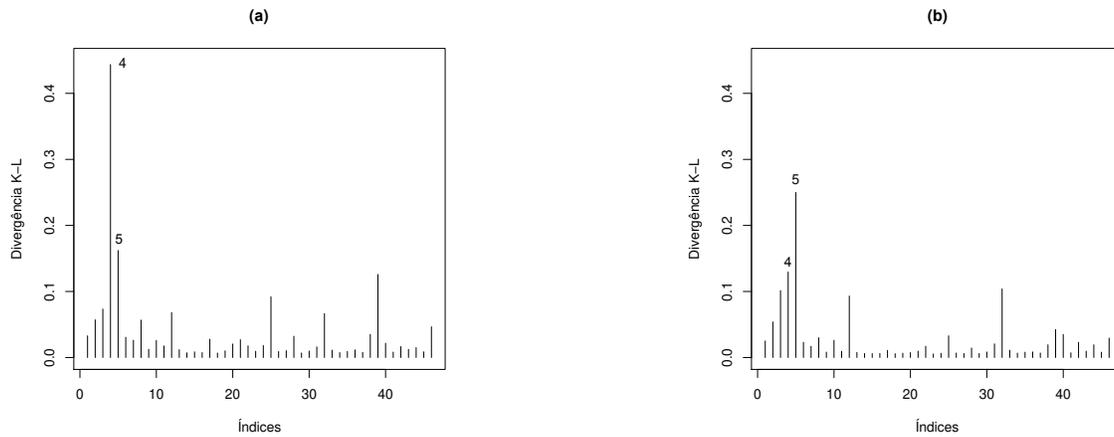


Figura 5.8: Divergência de K-L para os modelos - (a) log-BSNAC e (b) log-BS.

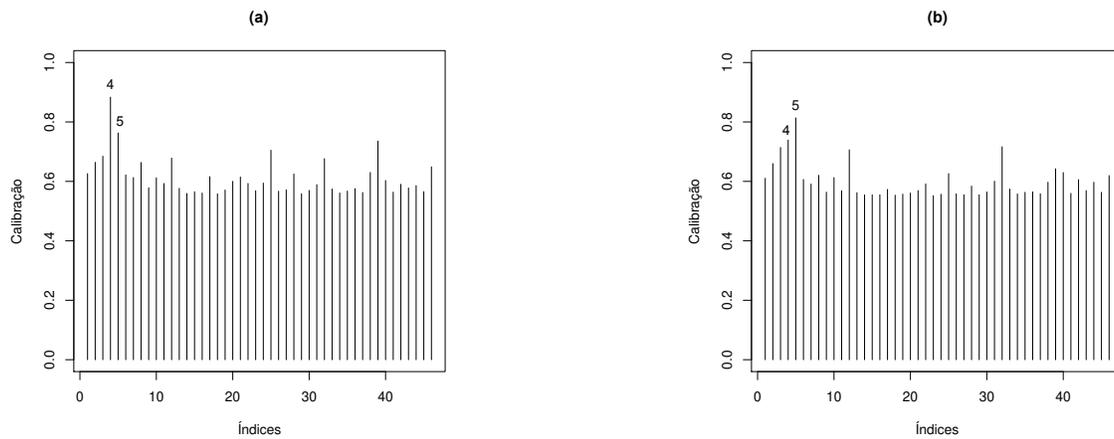


Figura 5.9: Calibração da divergência de K-L para os modelos - (a) log-BSNAC e (b) log-BS.

Após a identificação das observações influentes, é de interesse avaliar o impacto que a retirada dessas observações exerce nas estimativas dos parâmetros em cada um dos modelos em questão. Sendo assim, retiramos a observação 4 e ajustamos o modelo log-BSNAC, retiramos a observação 5 e ajustamos o modelo log-BS, para ambas as situações calculamos o impacto percentual, RC , tal como em (3.5.9), correspondente à cada uma das estimativas.

Na Tabela 5.7 apresentamos as novas estimativas para ambos os modelos, bem como o impacto percentual após a retirada das observações potencialmente influentes.

Tabela 5.7: Impacto percentual nas estimativas, após a retirada das observações potencialmente influentes.

	Obs.	α	φ_0	φ_1	γ
log-BSNAC	4	0,415 (- 2,12%)	9,224 (- 1,28%)	-5,298 (+ 3,55%)	-0,639 (- 6,15%)
log-BS	5	0,401 (+ 4,97%)	9,375 (+ 1,47%)	-5,523 (- 4,15%)	-

Os impactos percentuais que a retirada da observação 5 exerce nas estimativas dos parâmetros do modelo log-BS, é maior do que os impactos percentuais que a retirada da observação 4 exerce nas estimativas dos parâmetros do modelo log-BSNAC. Ou seja, o modelo de regressão log-BSNAC parece ser menos sensível à presença de observações discrepantes do que o modelo log-BS. Além disso, os gráficos referentes à análise de resíduos fornecem evidências, juntamente com os critérios para seleção de modelos, de que o modelo (5.2.4) proporciona um melhor ajuste aos dados.

5.2.2 Liga dos campeões da UEFA

Para essa aplicação iremos considerar o conjunto de dados analisado por Meintanis (2007), que diz respeito à partidas de futebol da Liga dos Campeões da UEFA (*Union of European Football Associations*) que encontra-se apresentado na Tabela C.1. Considere partidas de futebol em que (i) houve pelo menos um gol marcado pela equipe da casa, e (ii) houve pelo menos um gol marcado por qualquer equipe a partir de cobrança de pênalti, chute de falta, ou qualquer outro lance direto. Seja T_1 o tempo em minutos do primeiro gol marcado por qualquer equipe e T_2 representa o tempo em minutos do primeiro gol, de qualquer tipo, marcado pela equipe da casa. As estatísticas descritivas referentes a cada uma das variáveis são apresentadas na Tabela 5.8.

Tabela 5.8: Estatísticas descritivas para T_1 e T_2 .

	Mínimo	Média	Máximo	Q1	Mediana	Q3
T_1	2	40,892	82	25	41	54
T_2	2	32,865	85	14	28	48

Para ilustrar a metodologia apresentada no Capítulo 3, ajustamos o modelo log-BSNAC, definido da seguinte forma

$$V_i = \varphi_0 + \varphi_1 x_i + \varepsilon_i, i = 1, \dots, 37, \quad (5.2.5)$$

em que, pela construção do modelo, $V_i = \log(T_{2i})$, $x_i = \log(T_{1i})$ e $\varepsilon_i \sim \text{SHNAC}(\alpha, 0, \sigma = 2, \gamma)$. Ajustamos ainda o modelo log-BS, que pode ser definido como em (5.2.5), considerando que $\varepsilon_i \sim \text{SHN}(\alpha, 0, \sigma = 2)$. O nosso objetivo consiste em verificar qual dos dois modelos fornece um melhor ajuste ao conjunto de dados em questão.

Na Tabela 5.9, apresentamos as estimativas, sob as abordagens frequentista e bayesiana, bem como seus respectivos erros-padrão, intervalos de confiança e de credibilidade. As estimativas aqui apresentadas foram obtidas através dos algoritmos ECM e MCMC apresentados no Capítulo 3. O algoritmo MCMC foi implementado através do pacote `R2openBugs` do software `R`, considerando o conjunto de prioris apresentados, mais especificamente, $\alpha \sim \text{Gama}(1; 0, 8)$; $\varphi_k = N(0, 1000)$, $k = 1, 2$ e $\gamma \sim U(-0, 99527; 0, 99527)$.

Em termos dos parâmetros de regressão, se observarmos, por exemplo, os resultados frequentistas referentes ao modelo log-BSNAC, podemos concluir que a média do logaritmo do tempo em minutos do primeiro gol, de qualquer tipo, marcado pela equipe da casa, é igual a 0,871, quando observa-se o logaritmo do tempo em minutos do primeiro gol marcado por qualquer equipe, igual a zero. Além disso, há um acréscimo de 0,485 na média do tempo em minutos do primeiro gol, de qualquer tipo, marcado pela equipe da casa, quando o logaritmo do tempo em minutos do primeiro gol marcado por qualquer equipe sofre um aumento de uma unidade.

Tabela 5.9: Estimativas dos parâmetros e seus respectivos erros-padrão, intervalos de confiança e de credibilidade para os modelos log-BSNAC e log-BS.

Abordagem	Parâmetros	log-BSNAC			log-BS		
		Estimativas	EP	IC _{95%}	Estimativas	EP	IC _{95%}
Frequentista	α	0,871	0,136	[0,604; 1,137]	0,899	0,116	[0,672; 1,126]
	φ_0	1,384	0,963	[-0,503; 3,27]	1,059	0,702	[-0,317; 2,435]
	φ_1	0,485	0,245	[0,005; 0,965]	0,568	0,195	[0,186; 0,950]
	γ	-0,715	0,242	[-1,189; -0,241]	-	-	-
Bayesiana	α	0,943	0,124	[0,739; 1,239]	0,969	0,121	[0,766; 1,244]
	φ_0	1,305	0,789	[-0,270; 2,735]	0,982	0,778	[-0,561; 2,430]
	φ_1	0,501	0,218	[0,098; 0,932]	0,589	0,215	[0,168; 1,010]
	γ	-0,611	0,219	[-0,936; -0,116]	-	-	-

Na Tabela 5.5, apresentamos os critérios para seleção de modelos, obtidos através da metodologia apresentada na Seção 3.6.2. É possível perceber que todos os critérios apontam o modelo log-BSNAC como o que melhor se ajusta a esse conjunto de dados.

Tabela 5.10: Critérios para seleção de modelos.

Critérios	log-BSNAC	log-BS
AIC	92,865	97,468
BIC	99,308	102,300
EAIC	97,053	100,692
EBIC	103,496	105,524
DIC	263,933	281,259
LPML	-47,880	-49,268

Nas Figuras 5.10 e 5.11, apresentamos a análise de resíduos do modelo log-BSNAC considerando os resíduos calculados a partir das estimativas frequentistas e bayesianas, respectivamente. Já nas

Figuras 5.12 e 5.13, apresentamos a análise de resíduos do modelo log-BS considerando os resíduos calculados a partir das estimativas frequentistas e bayesianas, respectivamente.

Observando a Figura 5.10 (c) é possível perceber um comportamento simétrico dos resíduos e a presença de duas observações discrepantes, o mesmo pode-se concluir da 5.11 (c). Nas Figuras 5.10 (d) e 5.11 (d), nota-se que os resíduos encontram-se dentro das bandas de confiança e não apresentam tendências.

Já nas Figuras 5.12 (c) e 5.13 (c) é possível observar uma leve assimetria negativa dos resíduos. Observando as Figuras 5.12 (d) e 5.13 (d), nota-se uma espécie de concavidade voltada para baixo, evidenciando assim um comportamento assimétrico negativo. Portanto, a análise residual e os critérios de seleção de modelos obtidos para cada um dos ajustes, nos dão indícios de que o modelo log-BSNAC ajusta-se melhor aos dados do que o modelo log-BS.

Com o intuito de identificar observações influentes, fizemos uma análise de influência utilizando a divergência de Kullback-Leibler assim como descrita na Seção 3.6.3. As Figuras 5.14 (a)-(b) e 5.15 (a)-(b) dizem respeito as divergências de K-L e suas respectivas calibrações para os modelos log-BSNAC e log-BS, respectivamente. A análise de influência detectou as observações 14 (Olympiacos-Lyon), 15 (Internazionale-Porto), 34 (Arsenal-Rosenborg) e 37 (Bremen-Panathinaikos) como potencialmente influentes para os modelos log-BSNAC e log-BS.

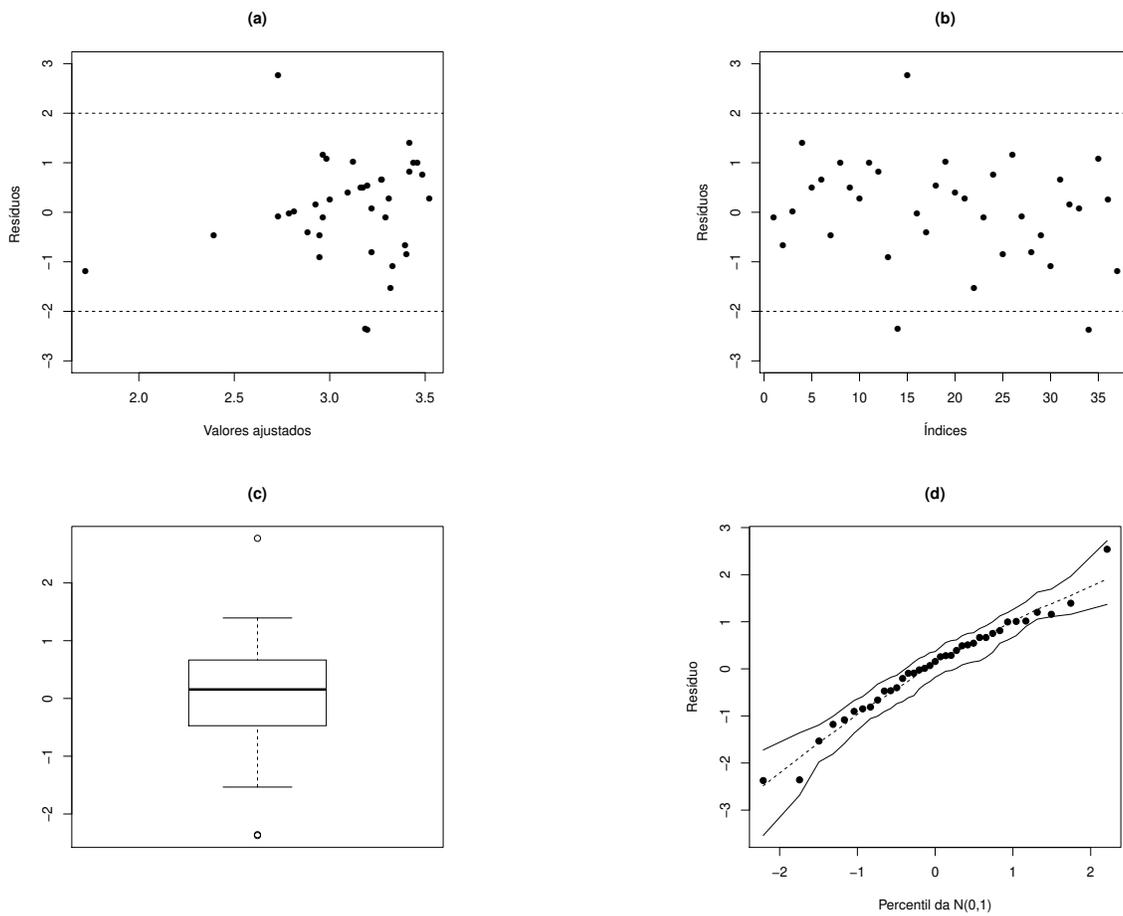


Figura 5.10: Análise de resíduos, calculados a partir das estimativas frequentistas, do modelo log-BSNAC.

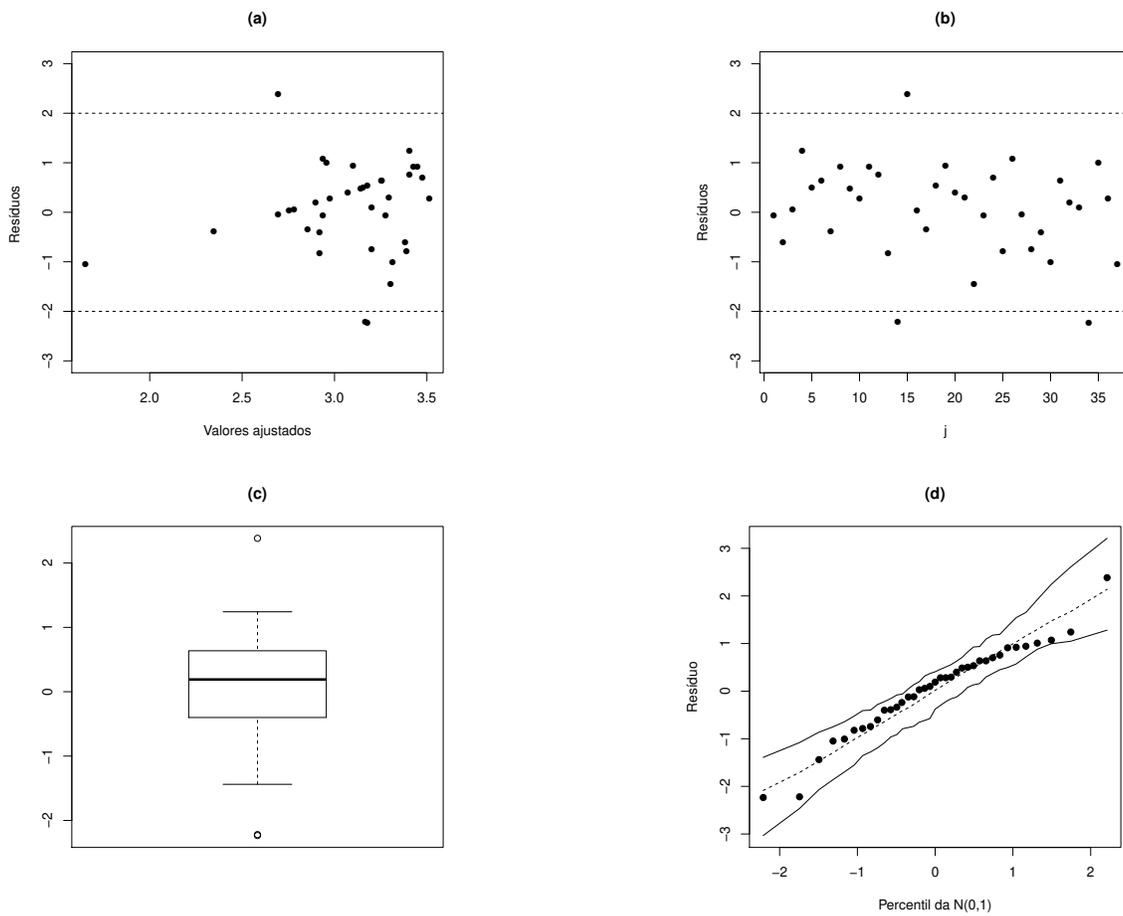


Figura 5.11: Análise de resíduos, calculados a partir das estimativas bayesianas, do modelo log-BSNAC.

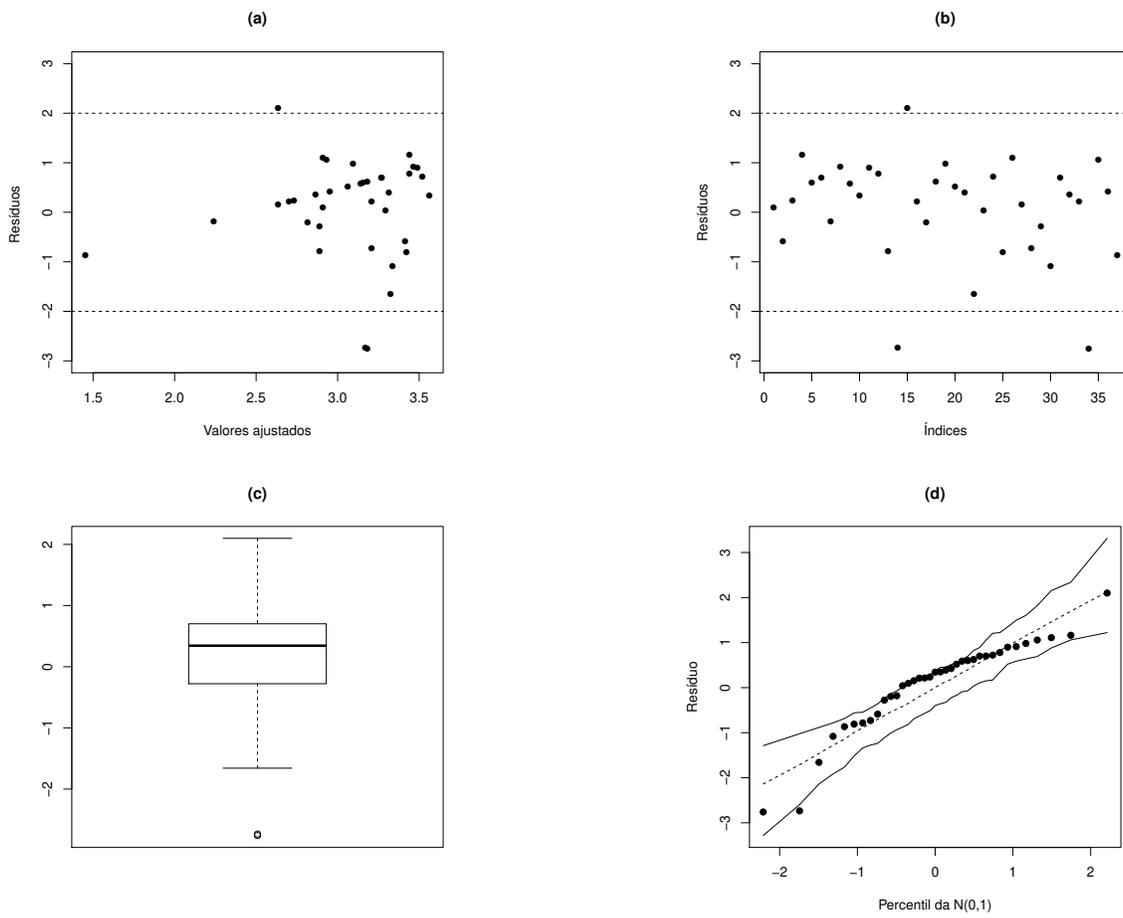


Figura 5.12: Análise de resíduos, calculados a partir das estimativas frequentistas, do modelo log-BS.

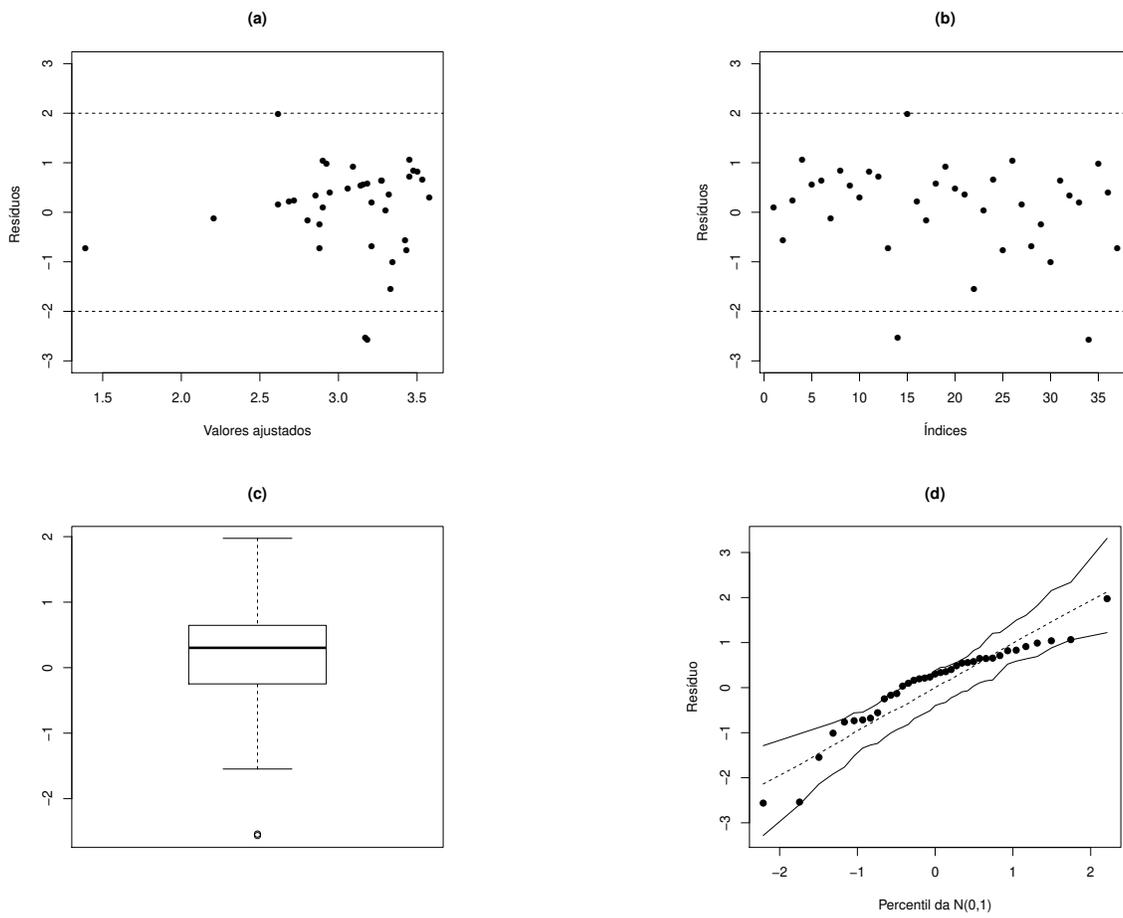


Figura 5.13: Análise de resíduos, calculados a partir das estimativas bayesianas, do modelo log-BS.

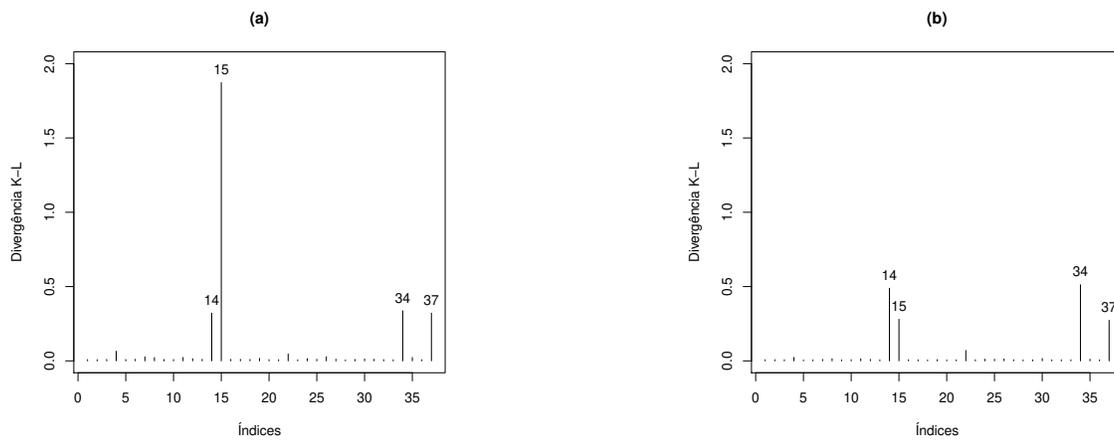


Figura 5.14: Divergência de K-L para os modelos - (a) log-BSNAC e (b) log-BS.

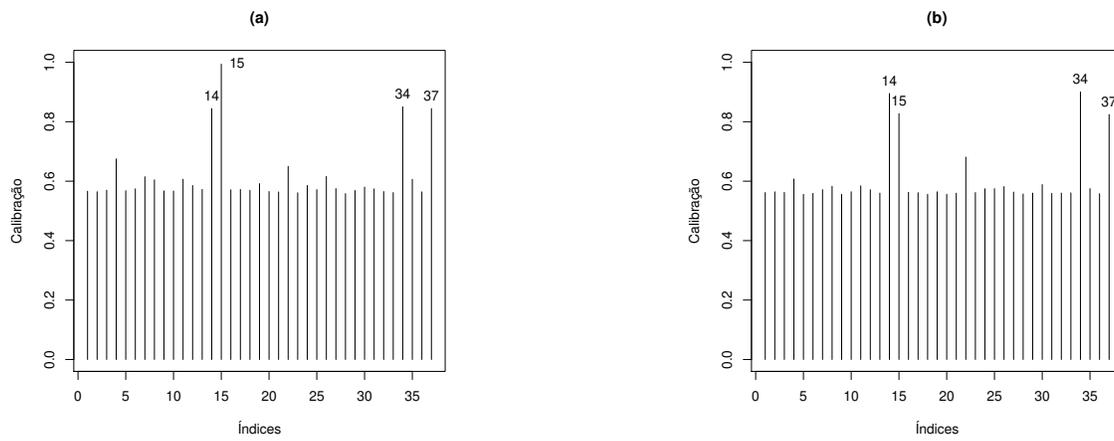


Figura 5.15: Calibração da divergência de K-L para os modelos - (a) log-BSNAC e (b) log-BS.

Tabela 5.11: Impacto percentual nas estimativas, após a retirada das observações potencialmente influentes.

	Obs.	α	φ_0	φ_1	γ
log-BSNAC	14	0,847 (-10,52%)	1,179 (-9,65%)	0,561 (+11,98%)	-0,534 (+12,60%)
	15	0,831 (-11,87%)	0,349 (-73,26%)	0,749 (+49,50%)	-0,829 (-35,68%)
	34	0,849 (-9,97%)	1,211 (-7,20%)	0,552 (+10,18%)	-0,528 (+13,58%)
	37	0,929 (-1,48%)	1,963 (+50,42%)	0,327 (-34,73%)	-0,642 (-5,07%)
	14, 15, 34 e 37	0,629 (-33,29%)	0,752 (-42,37%)	0,698 (+39,32%)	-0,718 (-17,51%)
log-BS	14	0,867 (-10,18%)	0,951 (-3,15%)	0,623 (+5,77%)	-
	15	0,912 (-5,88%)	0,553 (-43,68%)	0,694 (+17,8%)	-
	34	0,806 (-18,82%)	0,582 (-40,73%)	0,709 (+20,37%)	-
	37	0,963 (-0,62%)	1,694 (+72,50%)	0,397 (-32,59%)	-
	14, 15, 34 e 37	0,637 (-34,26%)	0,465 (-52,65%)	0,339 (-43,97%)	-

Após a identificação das observações influentes, avaliamos o impacto percentual, calculado a partir de (3.5.9), que a retirada dessas observações exerce nas estimativas dos parâmetros em cada um dos modelos em questão. Observando a Tabela 5.11 é possível notar que os impactos percentuais nas estimativas do modelo log-BSNAC, quando retiramos as observações 14 e 15, são maiores do que os impactos nas estimativas do modelo log-BS. Já a retirada da observação 34 impacta mais nas estimativas do modelo log-BS, do que nas estimativas do modelo log-BSNAC. A retirada da observação 37 causa um impacto percentual de 72,50% na estimativa de φ_0 do modelo log-BS, enquanto que no modelo log-BSNAC esse impacto é de 50,42%. Por fim, a retirada das quatro observações simultaneamente causa impactos percentuais maiores nas estimativas do modelo log-BS.

Capítulo 6

Considerações finais

Nesse trabalho, desenvolvemos métodos de estimação, sob as abordagens frequentista e bayesiana, para os modelos Birnbaum-Saunders Normal assimétrica sob a parametrização centrada (BSNAC) e log-Birnbaum-Saunders Normal assimétrica sob a parametrização centrada (log-BSNAC). Os modelos em questão são definidos em função da distribuição Normal assimétrica sob a parametrização centrada (NAC), como uma proposta alternativa aos modelos propostos Birnbaum-Saunders (1969a) e Santana (2009). A parametrização centrada além de remover a singularidade da matriz informação de Fisher esperada, que o ocorre quando o parâmetro de assimetria $\gamma \rightarrow 0$, contorna o problema referente à existência de um ponto de inflexão na verossimilhança perfilada de γ .

Para a estimação dos parâmetros dos modelos, apresentamos os algoritmos ECM e MCMC, que foram implementados com o auxílio do software R. Desenvolvemos ainda, técnicas de diagnóstico para o modelo de regressão log-BSNAC, mais especificamente, análise residual utilizando os resíduos obtidos a partir das estimativas frequentistas e bayesianas, e análise de influência apenas sob o ponto de vista bayesiano considerando a divergência de Kullback-Leibler. Além disso, apresentamos as estatísticas frequentistas e bayesianas para a comparação de modelos.

Ilustramos a metodologia desenvolvida para os modelos BSNAC e log-BSNAC tanto a partir de conjuntos de dados provenientes de estudos simulação, quanto a partir de dados reais. Assim, pudemos concluir que os modelos aqui propostos são preferíveis aos modelos originais, BS e log-BS, quando desejamos modelar dados assimétricos positivos ou assimétricos negativos.

Como perspectivas futuras, é do nosso interesse,

- Desenvolver a análise de influência para o modelo log-BSNAC sob o ponto de vista frequentista,
- Utilizar os modelos BSNAC e log-BSNAC para modelar dados censurados,
- Propor um modelo de regressão log-BSNAC não-linear,
- Desenvolver um modelo probabilístico Birnbaum-Saunders t-assimétrica, bem como um modelo de regressão log-BS-t-assimétrica para melhor acomodar observações presentes nas caudas.

Referências Bibliográficas

- [1] ACHCAR, J. Inferences for the birnbaum-saunders fatigue life model using bayesian methods. *Computational statistics & data analysis* 15, 4 (1993), 367–380.
- [2] AKAIKE, H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* 19, 6 (1974), 716–723.
- [3] ARELLANO-VALLE, R. B., AND AZZALINI, A. The centred parametrization for the multivariate skew-normal distribution. *Journal of Multivariate Analysis* 99, 7 (2008), 1362–1382.
- [4] AZEVEDO, C. L. N., B. H., AND ANDRADE, D. F. *A note on identification and metric issues for skew IRT model*. Relatório técnico, IMECC, Unicamp, 2012.
- [5] AZEVEDO, C., FOX, J., AND ANDRADE, D. Gibbs sampling for bayesian estimation and model fit assessment for a longitudinal irt model. *Computational Statistics & Data Analysis* 56 (2012), 4399–4412.
- [6] AZZALINI, A. A class of distributions which includes the normal ones. *Scandinavian journal of statistics* (1985), 171–178.
- [7] BARROS, M., PAULA, G., AND LEIVA, V. *Modelo de regressão t-Student log-Birnbaum-Saunders*. X Escola de Modelos de Regressão, Salvador. Programa e Resumos. São Paulo: ABE–Associação Brasileira de Estatística, 2007.
- [8] BAZÁN, J. L., BRANCO, M. D., AND BOLFARINE, H. A skew item response model. *Bayesian analysis* 1, 4 (2006), 861–892.
- [9] BERGER, J. O. *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media, 1985.
- [10] BERNARDO, J., AND SMITH, A. Bayesian theory. *Chichester: Wiley* (1994).
- [11] BIRNBAUM, Z., AND SAUNDERS, S. Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability* (1969b), 328–347.
- [12] BIRNBAUM, Z., AND SAUNDERS, S. C. A new family of life distributions. *Journal of applied probability* (1969a), 319–327.

- [13] BRANCO, M., AND ARELLANO-VALLE, R. Distribuições elípticas assimétricas. *São Paulo: Associação Brasileira de Estatística, 16^o Sinape-26 a 30* (2004).
- [14] BROWN, M., AND MILLER, K. J. *Biaxial fatigue data*. Department of Mechanical Engineering, University of Sheffield, 1978.
- [15] BYRD, R. H., LU, P., NOCEDAL, J., AND ZHU, C. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16, 5 (1995), 1190–1208.
- [16] CHO, H., IBRAHIM, J. G., SINHA, D., AND ZHU, H. Bayesian case influence diagnostics for survival models. *Biometrics* 65, 1 (2009), 116–124.
- [17] COOK, R. D. Assessment of local influence. *Journal of the Royal Statistical Society. Series B (Methodological)* (1986), 133–169.
- [18] COOLIDGE, J. L. The story of the binomial theorem. *American Mathematical Monthly* (1949), 147–157.
- [19] DESMOND, A. Stochastic models of failure in random environments. *Canadian Journal of Statistics* 13, 3 (1985), 171–183.
- [20] DESMOND, A. On the relationship between two fatigue-life models. *Reliability, IEEE Transactions on* 35, 2 (1986), 167–169.
- [21] DIAZ-GARCIA, J., AND LEIVA, V. A new family of life distributions based on the elliptically contoured distributions. *Journal of Statistical Planning and Inference* 128, 2 (2005), 445–457.
- [22] DUNN, P. K., AND SMYTH, G. K. Randomized quantile residuals. *Journal of Computational and Graphical Statistics* 5, 3 (1996), 236–244.
- [23] DUPUIS, D. J., AND MILLS, J. E. Robust estimation of the birnbaum-saunders distribution. *Reliability, IEEE Transactions on* 47, 1 (1998), 88–95.
- [24] ENGELHARDT, M., BAIN, L. J., AND WRIGHT, F. Inferences on the parameters of the birnbaum-saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics* 23, 3 (1981), 251–256.
- [25] FARIAS, R. B., AND LEMONTE, A. J. Bayesian inference for the birnbaum-saunders nonlinear regression model. *Statistical Methods & Applications* 20, 4 (2011), 423–438.
- [26] FARIAS, R. B. A. *Modelos multivariados binários com funções de ligação assimétricas*. Tese de Doutorado, Departamento de Estatística, IME-USP. São Paulo, 2012.
- [27] GALEA, M., LEIVA, V., AND PAULA, G. Influence diagnostics in log-birnbaum-saunders regression models. *Journal of Applied Statistics* 31, 9 (2004), 1049–1064.
- [28] GAMERMAN, D. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. CRC Press, 1997.

- [29] GELFAND, A. E., SMITH, A. F., AND LEE, T.-M. Bayesian analysis of constrained parameter and truncated data problems using gibbs sampling. *Journal of the American Statistical Association* 87, 418 (1992), 523–532.
- [30] GELMAN, A., CARLIN, J. B., STERN, H. S., AND RUBIN, D. B. *Bayesian data analysis*. Taylor & Francis, 1997.
- [31] GEMAN, S., AND GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 6 (1984), 721–741.
- [32] GEWEKE, J. Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints and the evaluation of constraint probabilities. In *Computing science and statistics: Proceedings of the 23rd symposium on the interface* (1991), Citeseer, pp. 571–578.
- [33] GILKS, W. R., RICHARDSON, S., AND SPIEGELHALTER, D. J. Introducing markov chain monte carlo. *Markov chain Monte Carlo in practice 1* (1996), 19.
- [34] GILKS, W. R., AND WILD, P. Adaptive rejection sampling for gibbs sampling. *Applied Statistics* (1992), 337–348.
- [35] GRADSHTEJN, I. S., AND RYZHIK, I. M. *Table of integrals, series and products*. Academic Press, 2000.
- [36] HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. *Biometrika* 57, 1 (1970), 97–109.
- [37] HENZE, N. A probabilistic representation of the 'skew-normal' distribution. *Scandinavian journal of statistics* (1986), 271–275.
- [38] IBRAHIM, J. G., CHEN, M.-H., AND SINHA, D. Criterion-based methods for bayesian model assessment. *Statistica Sinica* 11, 2 (2001), 419–444.
- [39] LACHOS, V. *Modelos lineares mistos assimétricos*. Tese de Doutorado, Departamento de Estatística, IME–USP. São Paulo, 2004.
- [40] LEE, P. M. *Bayesian statistics: an introduction*. John Wiley & Sons, 2012.
- [41] LEIVA, V. ; MICHELLI BARROS ; PAULA, G. *Generalized Birnbaum-Saunders Models using R*. XI Escola de Modelos de Regressão, Recife, 2009.
- [42] LEIVA, V., BARROS, M., PAULA, G. A., AND GALEA, M. Influence diagnostics in log-birnbaum–saunders regression models with censored data. *Computational Statistics & Data Analysis* 51, 12 (2007a), 5694–5707.

- [43] LEIVA, V., BARROS, M., PAULA, G. A., AND SANHUEZA, A. Generalized birnbaum-saunders distributions applied to air pollutant concentration. *Environmetrics* 19, 3 (2007b), 235–249.
- [44] LEIVA, V., VILCA, F., BALAKRISHNAN, N., AND SANHUEZA, A. A skewed sinh-normal distribution and its properties and application to air pollution. *Communications in Statistics—Theory and Methods* 39, 3 (2008), 426–443.
- [45] LEMONTE, A. J., AND CORDEIRO, G. M. Asymptotic skewness in birnbaum–saunders nonlinear regression models. *Statistics & probability letters* 80, 9 (2010), 892–898.
- [46] LEMONTE, A. J., MARTÍNEZ-FLOREZ, G., AND MORENO-ARENAS, G. Skewed multivariate birnbaum-saunders distributions. *arXiv preprint arXiv:1204.6286* (2012).
- [47] LU, M., AND CHANG, D. Bootstrap prediction intervals for the birnbaum-saunders distribution. *Microelectronics Reliability* 37, 8 (1997), 1213–1216.
- [48] MANN, N. R., SINGPURWALLA, N. D., AND SCHAFER, R. E. *Methods for statistical analysis of reliability and life data*. Wiley, New York, 1974.
- [49] MEINTANIS, S. G. Test of fit for marshall–olkin distributions with applications. *Journal of Statistical Planning and inference* 137, 12 (2007), 3954–3963.
- [50] MENG, X., AND RUBIN, D. Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika* 80, 2 (1993), 267–278.
- [51] NEWMAN, J. Various techniques used in connection with random digits. *NBS Appl. Math. Series* 36 (1951), 12.
- [52] OWEN, D. B. Tables for computing bivariate normal probabilities. *The Annals of Mathematical Statistics* (1956), 1075–1090.
- [53] OWEN, W., AND PADGETT, W. Accelerated test models for system strength based on birnbaum-saunders distributions. *Lifetime Data Analysis* 5, 2 (1999), 133–147.
- [54] O’HAGAN, A., AND FORSTER, J. Kendall’s advanced theory of statistic 2b: Bayesian inference. *London: Edward Arnold* (1994).
- [55] PAULA, G. A. *Modelos de regressão: com apoio computacional*. IME-USP São Paulo, 2004.
- [56] PEWSEY, A. Problems of inference for azzalini’s skewnormal distribution. *Journal of Applied Statistics* 27, 7 (2000), 859–870.
- [57] PLUMMER, M., BEST, N., COWLES, K., AND VINES, K. Coda: Convergence diagnosis and output analysis for mcmc. *R news* 6, 1 (2006), 7–11.
- [58] PRESS, S. J., AND PRESS, J. S. *Bayesian statistics: principles, models, and applications*. Wiley New York, 1989.

- [59] RIECK, J. R. A moment-generating function with application to the birnbaum-saunders distribution. *Communications in Statistics-Theory and Methods* 28, 9 (1999), 2213–2222.
- [60] RIECK, J. R., AND NEDELMAN, J. R. A log-linear model for the birnbaum-saunders distribution. *Technometrics* 33, 1 (1991), 51–60.
- [61] RIPLEY, B. D. *Stochastic simulation*. John Wiley & Sons, 1997.
- [62] RUBIN, D. B., ET AL. Using the sir algorithm to simulate posterior distributions. *Bayesian statistics* 3, 1 (1988), 395–402.
- [63] SAHU, S. K., DEY, D. K., AND BRANCO, M. D. A new class of multivariate skew distributions with applications to bayesian regression models. *Canadian Journal of Statistics* 31, 2 (2003), 129–150.
- [64] SANTANA, L. R. *Estimação e Diagnóstico em Modelos Birnbaum-Saunders Skew-Normal*. Tese de Mestrado, IMECC–UNICAMP. Campinas, 2009.
- [65] SAUNDERS, S. C. A family of random variables closed under reciprocation. *Journal of the American Statistical Association* 69, 346 (1974), 533–539.
- [66] SCHWARZ, G., ET AL. Estimating the dimension of a model. *The annals of statistics* 6, 2 (1978), 461–464.
- [67] SMITH, A. F., AND GELFAND, A. E. Bayesian statistics without tears: a sampling–resampling perspective. *The American Statistician* 46, 2 (1992), 84–88.
- [68] SPIEGELHALTER, D., THOMAS, A., BEST, N., AND LUNN, D. Openbugs version 2.10, user manual. *MRC Biostatistics Unit, Cambridge, United Kingdom* (2005).
- [69] SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P., AND LINDE, A. The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 76, 3 (2014), 485–493.
- [70] STURTZ, S., LIGGES, U., AND GELMAN, A. E. R2winbugs: a package for running winbugs from r. *Journal of Statistical software* 12, 3 (2005), 1–16.
- [71] TEAM, R. C. R: A language and environment for statistical computing. r foundation for statistical computing, 2008.
- [72] VILCA, F., SANTANA, L., LEIVA, V., AND BALAKRISHNAN, N. Estimation of extreme percentiles in birnbaum–saunders distributions. *Computational Statistics & Data Analysis* 55, 4 (2011), 1665–1678.
- [73] VILCA-LABRA, F., AND LEIVA-SÁNCHEZ, V. A new fatigue life model based on the family of skew-elliptical distributions. *Communications in Statistics—Theory and Methods* 35, 2 (2006), 229–244.

- [74] VOLLE, M. *Analyse des Donn'ees*, 3 ed. Economica, Paris, 1985.
- [75] XIE, F., AND WEI, B. Diagnostics analysis for log-birnbaum–saunders regression models. *Computational Statistics & Data Analysis* 51, 9 (2007), 4692–4706.

Apêndice A

Construção dos envelopes simulados

A.1 Resíduos obtidos através das estimativas frequentistas

Os gráficos de envelopes simulados, para os resíduos obtidos a partir das estimativas frequentistas, foram construídos da seguinte forma:

1. Ajusta-se o modelo log-BSNAC e gera-se uma amostra simulada de n observações independentes usando o modelo ajustado;
2. Ajusta-se o modelo à amostra n observações simulada no passo anterior e calcula-se os resíduos;
3. Repete-se os passos (1) e (2) k vezes;
4. Considera-se os n conjunto de observações formados pelas k estatísticas de ordem, para cada um desses n conjuntos calcula-se os valores médio, mínimo e máximo;
5. A partir dos valores mínimos e máximos das k estatísticas obtém-se o envelope simulado.

A.2 Resíduos obtidos através das estimativas bayesianas

Os gráficos de envelopes simulados para, os resíduos obtidos a partir das estimativas bayesianas, foram construídos da seguinte forma:

1. Gera-se n observações $N(0, 1)$ as quais são armazenadas em $\mathbf{e} = (e_1, \dots, e_n)^\top$.
2. A partir das observações geradas no passo 1, ordenamos essas observações e as consideramos como os resíduos, a saber: $e_i^* = e_i$.
3. Repete-se os passos (1)-(2), m vezes. Logo, teremos $e_{ij}^*, i = 1, \dots, n$ e $j = 1, \dots, m$.

4. Ao final, teremos uma matriz com os resíduos, ou seja, e_{ij}^* , $i = 1, \dots, n$, (tamanho da amostra), $j = 1, \dots, m$ (réplica).

$$\mathbf{E}_1 = \begin{bmatrix} e_{11}^* & e_{12}^* & \cdots & e_{1m}^* \\ e_{21}^* & e_{22}^* & \cdots & e_{2m}^* \\ \vdots & \vdots & \ddots & \vdots \\ e_{n1}^* & e_{n2}^* & \cdots & e_{nm}^* \end{bmatrix}$$

5. Dentro de cada amostra, ordena-se, de modo crescente, os resíduos, obtendo-se $e_{(i)j}^*$ (estatísticas de ordem):

$$\mathbf{E}_2 = \begin{bmatrix} e_{(1)1}^* & e_{(1)2}^* & \cdots & e_{(1)m}^* \\ e_{(2)1}^* & e_{(2)2}^* & \cdots & e_{(2)m}^* \\ \vdots & \vdots & \ddots & \vdots \\ e_{(n)1}^* & e_{(n)2}^* & \cdots & e_{(n)m}^* \end{bmatrix}$$

6. Obtém-se os limites $e_{(i)I}^* = \min_{1 \leq j \leq m} e_{(i)j}^*$ e $e_{(i)S}^* = \max_{1 \leq j \leq m} e_{(i)j}^*$, $j = 1, \dots, m$.

7. Na prática, considera-se $e_{(i)I}^* = \frac{e_{(i)(2)}^* + e_{(i)(3)}^*}{2}$ e $e_{(i)S}^* = \frac{e_{(i)(m-2)}^* + e_{(i)(m-1)}^*}{2}$ (refinamento das estimativas do mínimo e máximo), em que $e_{(i)(r)}^*$ é a r -ésima estatística de ordem dentro de cada linha, $i = 1, 2, \dots, n$. Além disso, consideramos como linha de referência $e_{(i)}^* = \frac{1}{m} \sum_{j=1}^m e_{(i)j}^*$, $i = 1, 2, \dots, n$.

Apêndice B

Conjunto de dados Fatiga Biaxial

Apresentamos aqui o conjunto de dados que diz respeito ao tempo de vida de um pedaço de metal sujeito a compressões cíclicas até falhar.

Tabela B.1: Dados de fadiga Biaxial (Brown e Miller, 1978).

Trabalho	Ciclos	Trabalho	Ciclos	Trabalho	Ciclos	Trabalho	Ciclos
11,5	3280	24,0	804	40,1	750	60,3	283
13,0	5046	24,6	1093	40,1	316	60,5	212
14,3	1563	25,2	1125	43,0	456	62,1	327
15,6	4707	25,5	884	44,1	552	62,8	373
16,0	977	26,3	1300	46,5	355	66,5	125
17,3	2834	27,9	852	47,3	242	67,0	187
19,3	2266	28,3	580	48,7	190	67,1	135
21,1	2208	28,4	1066	52,9	127	67,9	245
21,5	1040	28,6	1114	56,6	185	68,8	137
22,6	700	30,9	386	59,9	255	75,4	200
22,6	1583	31,9	745	60,2	195	100,5	190
24,0	482	34,5	736				

Apêndice C

Liga dos campeões da UEFA

Apresentamos aqui o conjunto de dados que diz respeito à partidas de futebol da Liga dos campeões da UEFA.

Tabela C.1: Liga dos campeões da UEFA

2005-2006	T_1	T_2	2004-2005	T_1	T_2
Lyon-Real Madrid	26	20	Internazionale-Bremen	34	34
Milan-Fenerbahce	62	18	Real Madrid-Roma	53	39
Chelsea-Anderlecht	19	19	Man. United-Fenerbahce	54	28
Club Brugge-Juventus	66	85	Bayern-Ajax	51	28
Fenerbahce-PSV	40	40	Moscow-PSV	76	64
Internazionale-Rangers	49	49	Barcelona-Shakhtar	64	15
Panathinaikos-Bremen	8	8	Leverkusen-Roma	26	48
Ajax-Arsenal	69	71	Arsenal-Panathinaikos	16	16
Man. United-Benfica	39	39	Dynamo Kiev-Real Madrid	44	13
Real Madrid-Rosenborg	82	48	Man. United-Sparta	25	14
Villareal-Benfica	72	72	Bayern-M. Tel Aviv	55	11
Juventus-Bayern	66	62	Bremen-Internazionale	49	49
Club Brugge-Rapid	25	9	Anderlecht-Valencia	24	24
Olympiacos-Lyon	41	3	Panathinaikos-PSV	44	30
Internacionale-Porto	16	75	Arsenal-Rosenborg	42	3
Schalke-PSV	18	18	Liverpool-Olympiacos	27	27
Barcelona-Bremen	22	14	M.Tel Aviv-Juventus	28	28
Milan-Schalke	42	42	Bremen-Panathinaikos	2	2
Rapid-Juventus	36	52			

