



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Modelos Birnbaum-Saunders para Séries Temporais

Rubens Batista de Souza

Orientador: Prof. Dr. Helton Saulo Bezerra dos Santos

Brasília
2018

Modelos Birnbaum-Saunders para Séries Temporais

Rubens Batista de Souza

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Orientador: Helton Saulo Bezerra dos Santos

**Brasília
2018**

Modelos Birnbaum-Saunders para Séries Temporais

Rubens Batista de Souza

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Texto aprovado em: 14 de dezembro de 2018.

Banca examinadora

Prof. Dr. Helton Saulo Bezerra dos Santos (Orientador)

Universidade de Brasília - UnB

Prof. Dr. Roberto Vila Gabriel (Membro interno)

Universidade de Brasília - UnB

Prof. Dr. Mario Ernesto Piscoya Diaz (Membro externo)

Universidade Federal de Goiás - UFG

Prof. Dr. Raul Yukihiro Matsushita (Suplente)

Universidade de Brasília - UnB

Dedico este trabalho, com todo meu amor, a minha maravilhosa família: aos meus pais Adalto e Eleude e ao meu irmão Silas. Em especial, à minha amada mãe – Dona Eleude – pelo seu carinho, suas orações e amor incondicional.

Agradecimentos

Ao grandioso Deus – dono de todas as coisas, que por amor me permitiu a graça de chegar tão longe. A minha eterna gratidão.

Aos meus pais, Adalto e Eleude, e ao meu irmão, Silas, que de maneira incondicional me incentivaram e me apoiaram desde o início desta jornada. Obrigado por todo amor e carinho, que sempre tiveram comigo.

Ao professor Helton Saulo, pela excelente orientação, competência, dedicação, correções, paciência e disponibilidade em todas as etapas do desenvolvimento deste trabalho. Sou imensamente agradecido por toda ajuda, por todos os ensinamentos e, particularmente, pelo incentivo e exemplo pessoal de humildade profissional.

Aos professores Víctor Leiva e Robert G. Aykroyd, pelas valiosas contribuições a este trabalho e, sobretudo, por compartilharem alguns de seus conhecimentos e experiências.

Aos professores Mario Ernesto e Roberto Vila, por terem participado da banca examinadora, pelos questionamentos, correções e valiosas sugestões que contribuíram para o resultado final deste trabalho. Igualmente, agradeço ao professor Raul Matsushita, pelos proveitosos comentários, na ocasião do exame de qualificação.

Ao Programa de Pós-Graduação em Estatística – PPGEST/UnB – pelo ambiente propício para o aprendizado e pela oportunidade de fazer o curso com excelente qualidade. A todos os professores, pelos conhecimentos, ensinamentos e contribuições à minha formação acadêmica e profissional. Cada um de vocês contribuíram de forma particular para a realização deste feito.

A todos da Secretaria de Pós-Graduação em Estatística – PPGEST/UnB – pela eficiência, dedicação e carisma. Em especial, ao André, por todo o atendimento prestado.

Aos professores do Departamento de Matemática e Estatística – DME, da Universidade Federal de Rondônia – UNIR, por terem contribuído para minha formação. Em especial, aos professores Nerio e Roziane, por todo incentivo.

Aos meus familiares e amigos, que oraram a Deus, pelas minhas conquistas. Sou agradecido pelo amor e carinho de todos.

À Regina, pelo amor, carinho, amizade, companheirismo e por todos os momentos que já compartilhamos juntos.

Ao discípulo José Paulo, com quem compartilhei tantos momentos na graduação e agora no mestrado, verdadeiro amigo de todas as horas. Muito obrigado por toda ajuda durante essa jornada.

Aos colegas do curso de mestrado pela troca de experiência. Ao amigo Adolfo, pela sua humildade em ajudar e ensinar.

Aos meus amigos e amigas de trabalho, da escola EEEFM 28 de novembro, por todo o carinho e por se alegrarem com minhas conquistas.

Às minhas amigas Marivone e Rosiene por toda ajuda, cooperação e incentivo em todos os momentos que precisei. Sou imensamente agradecido por tudo.

Ao amigo Erildo que foi um grande companheiro no momento em que precisei de sua ajuda.

Às minhas amigas Cida, Raquel e Tânia pelas palavras de incentivo, por toda atenção e cuidado que sempre demonstraram ter comigo.

Às amigas Luana e Bruna pelo auxílio e hospitalidade nos primeiros dias que estive em Brasília.

Aos meus amigos e amigas, José Paulo, Isabel, Gesiane, Reinaldo, Leonardo, Kessys, Bruno e Regina pelos momentos de descontração compartilhados.

À SEDUC-RO, pela liberação da licença para cursar o Mestrado, proporcionando meu aperfeiçoamento acadêmico e profissional.

À todos aqueles que contribuíram de forma direta ou indireta para meu sucesso e crescimento pessoal e profissional.

Muitíssimo obrigado a todos!

Essencialmente, todos os modelos estão errados, mas alguns são úteis.

George E. P. Box (1919-2013)

Resumo

Este trabalho visa preencher uma lacuna existente na literatura da distribuição Birnbaum-Saunders pertinente à modelagem de dados com alguma estrutura de dependência. O principal objetivo é oferecer uma contribuição através da introdução de dois novos modelos, que serão apresentados em dois artigos. No primeiro artigo é proposto o modelo autorregressivo de médias móveis com distribuição Birnbaum-Saunders (BISARMA). A metodologia proposta inclui diversos aspectos relacionados à formulação do modelo, a estimação de seus parâmetros com base no método da máxima verossimilhança, análise dos resíduos e predição. No segundo artigo, é estudado um modelo que é especificado em termos de uma média condicional variável no tempo e pode ser considerado como uma extensão do modelo de regressão Birnbaum-Saunders reparametrizado (RBS). Nomeamos este modelo como RBSARMA. Os desempenhos de ambas metodologias são avaliadas através de simulações de Monte Carlo. Esta metodologia foi implementada usando o software estatístico R. Finalmente, as metodologias são aplicadas à análise de um conjunto de dados ambientais reais para ilustrar a flexibilidade e potencialidade desses modelos. Os resultados numéricos mostram o excelente desempenho de ambos os modelos, indicando que a distribuição Birnbaum-Saunders é uma boa alternativa de modelagem quando se trata de dados de séries temporais positivos e assimétricos e, portanto, pode ser uma adição valiosa ao conjunto de ferramentas estatísticas aplicadas. Entre os dois modelos, o modelo BISARMA apresentou melhor potencial.

Palavras-chave: Dados dependentes; Modelos ARMA; Distribuições Birnbaum-Saunders; Métodos de máxima verossimilhança; Seleção de modelos; Método de Monte Carlo; Análise residual; Software R.

Abstract

This work aims to fill a gap in the Birnbaum-Saunders distribution literature pertinent to modeling data with some dependency structure. The main objective is to offer a contribution through the introduction of two new models, which will be presented in two papers. In the first paper, we proposed an autoregressive moving-average model based on the Birnbaum-Saunders distribution (BISARMA). The proposed methodology includes several aspects related to formulation, estimation of its parameters based on the maximum likelihood method, residual analysis and prediction. In the second paper, it is studied a model that is specified in terms of a time-varying conditional mean and can be considered of as an extension of the reparameterized Birnbaum-Saunders regression model (RBS). We named this model as RBSARMA. The performance of both methodologies are evaluated through Monte Carlo simulations. All numerical evaluations were implemented using the statistical software R. Finally, we have applied the proposed methodologies to the analysis of a real environmental data set in order to illustrate their flexibility and potentiality. The numerical results show the excellent performance of both models, indicating that the BS distribution is a good modeling alternative when it comes to positive and asymmetric time series data. Therefore, this model can be a valuable addition to the set of applied statistical tools. Between the two models, the BISARMA model presented better results.

Keywords: Dependend data; ARMA models; Birnbaum-Saunders distributions; Maximum likelihood methods; Model selection; Monte Carlo method; Residual analysis; R software.

Lista de Tabelas

2.1	Estimativas de ML para α indicado, com base em simulações de Monte Carlo para o modelo BISARMA (1,1).	16
2.2	Estimativas ML para os diferentes valores de ϕ e θ , baseadas em simulações de Monte Carlo para o modelo BISARMA(1,1).	18
2.3	Estatísticas de avaliação de previsão para os diferentes valores de ϕ e θ , baseadas em simulações de Monte Carlo para o modelo BISARMA(1,1) e ARMA(1,1) (com as estatísticas entre parênteses).	20
2.4	Estatísticas de avaliação de previsão para os diferentes valores de α , baseadas em simulações de Monte Carlo para o modelo BISARMA(1,1) e ARMA(1,1) (com as estatísticas entre parênteses).	21
2.5	Estimativas e medidas de adequação de modelos para os dados da série de tempos de mortalidade no período de 10 anos (1970 - 1979) no Condado de Los Angeles.	25
3.1	Estimativas de máxima verossimilhança condicional para os diferentes valores de δ , baseadas em simulações de Monte Carlo para o modelo RBSARMA(1,1).	40
3.2	Estimativas de máxima verossimilhança condicional para os diferentes valores de ϕ e θ , baseadas em simulações de Monte Carlo para o modelo RBSARMA(1,1).	42
3.3	Estatísticas de avaliação de previsão para os diferentes valores de ϕ e θ , baseadas em simulações de Monte Carlo para o modelo RBSARMA(1,1) e ARMA(1,1) (com as estatísticas entre parênteses).	44
3.4	Estatísticas de avaliação de previsão para os diferentes valores de δ , baseadas em simulações de Monte Carlo para o modelo RBSARMA(1,1) e ARMA(1,1) (com as estatísticas entre parênteses).	45
3.5	Estatísticas descritivas para M_t , X_{1t} e X_{2t}	47
3.6	Estimativas e medidas de adequação de modelos para os dados da série de tempos de mortalidade no período de 10 anos (1979-1979) no Condado de Los Angeles.	50

Lista de Figuras

2.1	Funções densidade de probabilidade das distribuições: (a) $BS(\alpha, 1)$ e (b) $BS(0.1, \beta)$ para os valores indicados dos parâmetros de forma α e escala β , respectivamente.	6
2.2	Funções densidade da distribuição Log-BS: (a) $Log-BS(\alpha, 0)$ e (b) $Log-BS(1, \mu)$ para os valores indicados dos parâmetros de forma α e localização μ	9
2.3	Viés absoluto empírico (a), variância (b) e MSE (c) do estimador de α	17
2.4	Mortalidade cardiovascular (a), Temperatura (b) e Material Particulado (c) ao longo do período de 10 anos (1970 - 1979).	22
2.5	Dispersão para a relação entre mortalidade e particulado (a) e entre mortalidade e temperatura (b) durante o período de 10 anos (1970 - 1979) no condado de Los Angeles.	23
2.6	Representação gráfica das ACF (a) e PACF (b) dos resíduos do modelo de regressão para dados de séries temporais ao longo do período de 10 anos (1970 - 1979) no condado de Los Angeles.	24
2.7	Envelope de resíduos Cox-Snell generalizado (à esquerda), ACF (centro) e PACF (à direita) do modelo BISARMA(2,0) com dados de séries temporais ao longo do período de 10 anos (1970 - 1979) no Condado de Los Angeles.	26
2.8	Envelope de resíduos Cox-Snell generalizado (à esquerda), ACF (centro) e PACF (à direita) do modelo ARMA(2,0) com dados de séries temporais ao longo do período de 10 anos (1970 - 1979) no Condado de Los Angeles.	26
2.9	Série mortalidade cardiovascular no Condado de Los Angeles (cinza), ajustada conforme os modelos BISARMA(2,0) (preto) e ARMA(2,0) (preto - - -).	27
3.1	Funções densidade de probabilidade da distribuição RBS(μ, δ) para $\mu = 1$ fixo (a) e para $\delta = 50$ fixo (b).	34
3.2	Viés absoluto empírico (a), variância (b) e MSE (c) do estimador de δ	40
3.3	Mortalidade cardiovascular (a), Temperatura (b) e níveis de material particulado (c) ao longo do período de 10 anos (1970 - 1979).	46

3.4	Matriz de dispersão, histogramas e coeficientes de correlação das variáveis mortalidade (M_t), temperatura (X_{1t}) e níveis de material particulado (X_{2t}). As variáveis e seus histogramas são mostrados ao longo da diagonal da matriz. Abaixo da diagonal mostra os gráficos de dispersão entre duas variáveis, e o correspondente coeficiente de correlação linear é mostrado acima da diagonal.	47
3.5	Boxplots para as variáveis M_t (a), X_{1t} (b) e X_{2t} (c).	48
3.6	Representação gráfica das funções de (a) autocorrelação (ACF) e (b) autocorrelação parcial (PACF) dos resíduos do modelo de regressão para dados de séries temporais ao longo do período de 10 anos (1979-1979) no condado de Los Angeles.	49
3.7	Gráfico dos envelopes dos resíduos de Cox-Snell generalizado (a) e resíduos quantil randomizados (b), funções de autocorrelação (c) e autocorrelação parcial (d) para o modelo RB-SARMA(2,0).	51
3.8	Gráfico dos envelopes dos resíduos de Cox-Snell generalizado (a) e resíduos quantil randomizados (b), funções de autocorrelação (c) e autocorrelação parcial (d) para o modelo ARMA(2,0).	52
3.9	Série mortalidade cardiovascular no condado de Los Angeles (cinza), ajustadas conforme os modelos RBSARMA(2,0) (preto) e ARMA(2,0) (preto - - -)	53

Sumário

1	Introdução	1
2	BISARMA: Um modelo de séries temporais Birnbaum-Saunders e análise de dados para um processo ambiental	3
2.1	Introdução	4
2.2	Definições Básicas	5
2.2.1	Distribuição Birnbaum-Saunders	5
2.2.2	Distribuição log-Birnbaum-Saunders	7
2.3	Modelo BISARMA	9
2.3.1	Formulação	9
2.3.2	Estimação	11
2.3.3	Previsão	13
2.3.4	Análise residual	14
2.4	Simulações de Monte Carlo	14
2.4.1	Modelo BISARMA(1,1)	15
2.4.2	Medidas de desempenho e seleção de modelos	18
2.5	Aplicação	21
2.6	Considerações finais	27
3	RBSARMA: Um modelo de séries temporais Birnbaum-Saunders reparametrizado e sua aplicação no estudo de um processo de dados ambientais	28
3.1	Introdução	28
3.2	Definições Básicas	31
3.2.1	Distribuição Birnbaum-Saunders	31
3.2.2	Distribuição Birnbaum-Saunders reparametrizada	32
3.2.3	Modelo de regressão RBS	34
3.3	Modelo RBSARMA	35

3.3.1	Formulação	35
3.3.2	Estimação	36
3.3.3	Análise residual	37
3.4	Simulação de Monte Carlo	38
3.4.1	Modelo RBSARMA(1,1)	39
3.4.2	Medidas de desempenho e seleção de modelos	42
3.5	Aplicação	45
3.6	Considerações finais	53
4	Observações finais e trabalhos futuros	54
	Referências Bibliográficas	55
	Apêndice	60

Capítulo 1

Introdução

Este trabalho trata, de maneira independente, auto-contida, da apresentação de dois artigos pertinentes à modelagem de dados com alguma estrutura de dependência. O ponto comum desses artigos reside no uso de modelos Birnbaum-Saunders (BS) aplicados à análise de dados de um processo ambiental.

A distribuição Birnbaum-Saunders (BS) é uma família de modelos originados na lei de danos cumulativos relacionados à fadiga de materiais submetidos a ciclos de estresse e tensão. Esta distribuição é unimodal, com assimetria positiva e suporte definido nos números reais positivos, indexada por dois parâmetros que controlam a sua forma e escala. Desde as suas origens e numerosas aplicações nas ciências dos materiais, vêm recebendo interesse em diferentes campos do conhecimento; veja [Leiva \(2016\)](#) e [Balakrishnan e Kundu \(2018\)](#).

A atratividade da distribuição BS para a análise de dados deve-se, entre várias razões, as suas boas propriedades teóricas e a sua estreita relação com a distribuição normal. Na literatura, além de assumir um papel proeminente nas áreas de confiabilidade e análise de sobrevivência, destaca-se pelas suas diversas generalizações. Dentre as suas aplicações práticas, destacam-se as áreas de negócios, economia, finanças, engenharia, medicina, meio ambiente e agricultura. Para algumas de suas aplicações, veja [Leiva et al. \(2009, 2010, 2012, 2014a\)](#), [Bhatti \(2010\)](#), [Vilca et al. \(2010\)](#), [Paula et al. \(2012\)](#), [Marchant et al. \(2013\)](#), [Saulo et al. \(2013\)](#), [Garcia-Papani et al. \(2017\)](#), [Leiva e Saulo \(2017\)](#), [Leão et al. \(2017, 2018a,b\)](#) e [Saulo et al. \(2018a\)](#).

O objetivo principal deste trabalho é estudar duas novas classes de modelos de séries temporais baseadas na distribuição BS, que permite descrever dados positivos e assimétricos que possuem uma estrutura autorregressiva. O primeiro modelo é denominado de autorregressivo de médias móveis Birnbaum-Saunders (BISARMA), enquanto o segundo é denominado de autorregressivo de médias móveis Birnbaum-Saunders reparametrizado (RBSARMA). Para o modelo BISARMA, no Capítulo

2, apresentam-se: (i) a distribuição BS na sua versão original e sua versão logarítmica (log-BS), assim como algumas de suas propriedades; (ii) formulação do modelo BISARMA, estimativas com base no método de máxima verossimilhança, previsão e análise dos resíduos; (iii) realização de simulações de Monte Carlo para avaliar o desempenho da metodologia proposta; e (iv) uma aplicação do modelo proposto para dados de séries temporais ambientais reais. O segundo modelo permite modelar os dados com base na média. Para o modelo RBSARMA, discutido no Capítulo 3, apresentam-se: (i) a distribuição BS na sua versão original e a sua forma reparametrizada RBS, bem como algumas propriedades de ambas as distribuições; (ii) formulação do modelo RBSARMA, estimativas com base no método de máxima verossimilhança e análise dos resíduos ; (iii) realização da simulação de Monte Carlo para avaliar a metodologia; e (iv) uma aplicação do modelo aos mesmos dados considerados no Capítulo 2. No Capítulo 4, apresentam-se algumas conclusões dos estudos e sugestões para trabalhos futuros.

Capítulo 2

BISARMA: Um modelo de séries temporais Birnbaum-Saunders e análise de dados para um processo ambiental

Resumo: A distribuição Birnbaum-Saunders (BS) é um modelo que aparece frequentemente na literatura estatística e provou ser muito versátil e eficiente em uma ampla gama de aplicações. Contudo, apesar do crescente interesse no estudo desta distribuição e do desenvolvimento de muitos artigos relacionados, poucos desses artigos consideraram dados com alguma estrutura de dependência. Para preencher essa lacuna, pretende-se com este artigo introduzir uma nova classe de modelos de séries temporais baseada na distribuição BS, que permite descrever dados positivos e assimétricos que possuem uma estrutura autorregressiva. Nomeamos esses modelos de autorregressivos de média móvels Birnbaum-Saunders (BISARMA). A metodologia proposta inclui um estudo aprofundado de propriedades teóricas e questões práticas, tais como estimativa de parâmetros baseada no método de máxima verossimilhança, análise de resíduos e previsão. O desempenho da metodologia proposta é avaliado usando uma simulação de Monte Carlo. Finalmente, uma análise estatística com dados ambientais do mundo real é realizada utilizando a metodologia para mostrar seu potencial para aplicações. Os resultados numéricos mostram o excelente desempenho do modelo BISARMA, indicando que a distribuição BS é uma boa alternativa de modelagem quando se trata de dados de séries temporais que são positivos e assimétricos e, portanto, pode ser uma adição valiosa ao conjunto de ferramentas estatísticas aplicadas.

Palavras-chave: Modelos ARMA; Distribuição Birnbaum-Saunders; Dados dependentes ao longo do tempo; Métodos de máxima verossimilhança; Método de Monte Carlo; Seleção de modelos; Análise residual; Software R.

2.1 Introdução

A distribuição Birnbaum-Saunders (BS) é um modelo que surge constantemente na literatura estatística aplicada. Nas últimas décadas, tem se mostrado versátil e eficiente em diversos campos da ciência, sendo amplamente estudada, devido a justificativa teórica, suas boas propriedades e a sua estreita relação com a distribuição normal. [Mann et al. \(1974, p.155\)](#) demonstraram que a distribuição BS é unimodal. Além disso, ela possui assimetria positiva e suporte definido nos números reais positivos, indexada por dois parâmetros que controlam a sua forma e escala. O modelo BS é muitas vezes considerado como uma distribuição de vida devido às suas origens descrevendo a fadiga de materiais sujeitos a estresse. Assim, a distribuição BS assume um papel de destaque nas áreas de confiabilidade e análise de sobrevivência, sendo uma boa alternativa às distribuições clássicas. Para mais detalhes, sobre a distribuição BS veja [Birnbaum e Saunders \(1969a,b\)](#), [Johnson et al. \(1995, págs. 651-663\)](#), [Leiva \(2016\)](#) e [Balakrishnan e Kundu \(2018\)](#).

Embora a distribuição BS tenha suas origens em física e engenharia, também recebeu interesse em várias outras áreas e, particularmente, nas ciências da terra, devido à sua reformulação de processos ambientais formalizados por [Leiva et al. \(2015\)](#). Para algumas referências sobre o uso de modelos de BS em aplicações ambientais, veja [Leiva et al. \(2009, 2016\)](#), [Vilca et al. \(2010\)](#), [Ferreira et al. \(2012\)](#), [Marchant et al. \(2013, 2016a,b, 2018\)](#), [Saulo et al. \(2013\)](#), [Garcia-Papani et al. \(2017, 2018a,b\)](#) e [Lillo et al. \(2018\)](#).

O estudo da distribuição BS tem recebido crescente interesse e uma quantidade considerável de trabalho está disponível; veja as publicações recentes de [Leiva \(2016\)](#), [Balakrishnan e Kundu \(2018\)](#) e as referências nele, que resumem a maioria dos trabalhos publicados até a presente data. Contudo, pouco foi desenvolvido para a análise de dados com estrutura de dependência temporal, e particularmente, para os modelos de séries temporais baseados na distribuição BS. Alguns esforços nessa linha são atribuídos a [Bhatti \(2010\)](#), [Leiva et al. \(2014b\)](#) e [Saulo et al. \(2018a\)](#) que, motivados pelos trabalhos de [Engle e Russell \(1998\)](#), desenvolveram modelos autorregressivos baseados na distribuição BS. Outros estudos recentes relacionados aos modelos autorregressivos BS são atribuídos a [Fonseca e Cribari-Neto \(2018\)](#) e [Rahul et al. \(2018\)](#).

Devido ao exposto, o principal objetivo deste trabalho é propor uma nova classe de modelos baseados na distribuição BS para séries temporais, denominados como modelos autorregressivos de médias móveis Birnbaum-Saunders (BISARMA). O modelo proposto pode ser considerado análogo aos mo-

delos ARMAX; veja [Shumway e Stoffer \(2017\)](#). A abordagem utilizada é semelhante aos trabalhos de [Benjamin et al. \(2003\)](#), [Rocha e Cribari-Neto \(2009\)](#) e [Maior e Cysneiros \(2016\)](#), que estenderam o modelo de série temporal autorregressivo de média móvel (ARMA) para diferentes classes de distribuições, como o modelo ARMA generalizado (GARMA), que inclui a família exponencial, β -ARMA para distribuições beta e SYMARMA para distribuições simétricas.

Após essa introdução, esse artigo é organizado da seguinte maneira: A Seção 2.2 apresenta a distribuição BS, sua versão logarítmica (log-BS) e algumas de suas propriedades. A Seção 2.3 formula o modelo BISARMA, fornece estimativa, previsão e análise de resíduos com base no método de máxima verossimilhança (ML). A Seção 2.4 relata um estudo de simulação de Monte Carlo para avaliar o desempenho da metodologia proposta. A Seção 2.5 realiza uma aplicação do modelo BISARMA para dados de séries temporais ambientais reais, bem como uma análise residual para mostrar o seu potencial. Finalmente, a Seção 2.6 apresenta as considerações finais. Um apêndice com expressões matemáticas detalhadas também é apresentado.

2.2 Definições Básicas

2.2.1 Distribuição Birnbaum-Saunders

Se uma variável aleatória T segue uma distribuição BS, geralmente denotada por $T \sim \text{BS}(\alpha, \beta)$, esta distribuição pode ser representada pela função de distribuição acumulada (f.d.a.) dada por

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

em que $\Phi(\cdot)$ é a f.d.a. da normal padrão, α é o parâmetro de forma e β é um parâmetro de escala, assim como a mediana da distribuição.

A função densidade de probabilidade (f.d.p.) de T é obtida derivando (2.1) em relação a t , ou seja, $f_T(t; \alpha, \beta) = F'_T(t; \alpha, \beta)$. De modo que,

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

A Figura 2.1, mostra alguns exemplos de formas para a f.d.p. da distribuição $\text{BS}(\alpha, \beta)$. Note que na Figura 2.1 (a), à medida que o valor de α aumenta, o grau de assimetria aumenta, provocando o

aumento da variância e o achatamento da distribuição. Caso contrário, quanto mais α se aproxima de zero, mais a curva tende a ser simétrica em torno do valor fixo $\beta = 1$ (mediana da distribuição) e a sua variabilidade diminui. Na Figura 2.1 (b), fixando o valor $\alpha = 0.1$ e variando β , a forma do gráfico da f.d.p. é a mesma e somente a sua escala é alterada; isto é, a f.d.p. se desloca no mesmo sentido da variabilidade de β . Isto mostra o fato de que β é um parâmetro de escala e, além disso, também corresponde a mediana da distribuição.

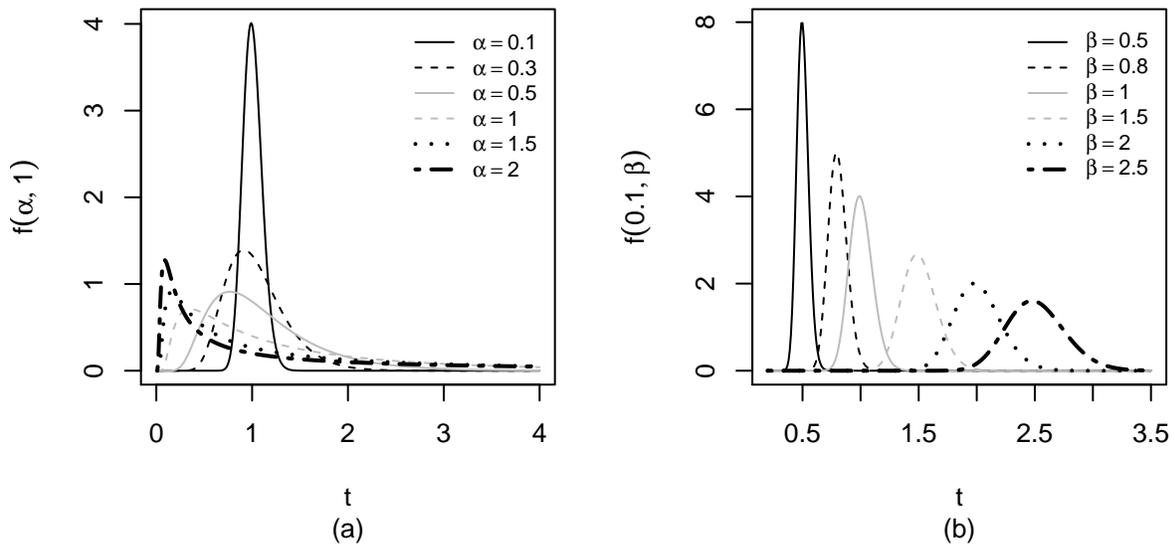


Figura 2.1: Funções densidade de probabilidade das distribuições: (a) $BS(\alpha, 1)$ e (b) $BS(0.1, \beta)$ para os valores indicados dos parâmetros de forma α e escala β , respectivamente.

Uma propriedade importante na construção dessa distribuição é que uma variável aleatória $T \sim BS(\alpha, \beta)$ pode ser gerada a partir de uma variável aleatória Z , que possui a distribuição normal padrão, mediante a seguinte relação

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

em que

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

Essa relação é extremamente útil e fornece um método para obter números pseudo-aleatórios da distribuição BS. Para mais detalhes veja [Leiva \(2016\)](#). Além disso, algumas propriedades da

distribuição BS são apresentadas a seguir. Se $T \sim \text{BS}(\alpha, \beta)$, então: (i) para qualquer constante real $k > 0$, tem-se $kT \sim \text{BS}(\alpha, k\beta)$, o que significa que a distribuição BS é fechada sob proporcionalidade; (ii) Se $1/T \sim \text{BS}(\alpha, \beta^{-1})$, implica que a distribuição BS também é fechada sob reciprocidade; (iii) O parâmetro β é um parâmetro de escala, em que, $T/\beta \sim \text{BS}(\alpha, 1)$. Além disso, β é a mediana da distribuição de T , que pode ser obtida diretamente fazendo $q = 0.5$ na sua função quantil que é dada por

$$t(q; \alpha, \beta) = F_T^{-1}(q; \alpha, \beta) = \beta \left[\frac{\alpha z(q)}{2} + \sqrt{\left(\frac{\alpha z(q)}{2}\right)^2 + 1} \right]^2, \quad 0 < q \leq 1,$$

em que z é a função quantil da normal padrão e F_T^{-1} é a inversa da função F_T expressa em (2.1); e (iv) α é um parâmetro de forma, tal que, quando $\alpha \rightarrow 0$, a distribuição BS tende para uma distribuição normal $N(\beta, \tau)$, onde $\tau \rightarrow 0$ quando $\alpha \rightarrow 0$. Além dessas propriedades, o r -ésimo momento da distribuição BS é dado por

$$E[T^r] = \beta^r \sum_{j=0}^r \binom{2r}{2j} \sum_{i=0}^j \binom{j}{i} \frac{(2(r-j+i))!}{2^{r-j+i}(r-j-i)!} \left(\frac{\alpha}{2}\right)^{2(r-j+i)}.$$

As expressões para a média, variância, coeficientes de assimetria e curtose, são dadas, respectivamente, por

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

$$CS[T] = 16\alpha^2 \frac{(11\alpha^2 + 6)}{(5\alpha^2 + 4)^3} \quad \text{e} \quad CK[T] = 3 + 6\alpha^2 \frac{(93\alpha^2 + 41)}{(5\alpha^2 + 4)^2}.$$

Conseqüentemente, conforme a propriedade da reciprocidade, temos que

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

2.2.2 Distribuição log-Birnbaum-Saunders

A distribuição log-BS, denotada por $\text{log-BS}(\alpha, \mu)$, é derivada como o logaritmo de uma variável aleatória $\text{BS}(\alpha, \beta)$. [Rieck e Nedelman \(1991\)](#) provaram que se $T \sim \text{BS}(\alpha, \beta)$, então $Y = \log(T) \sim \text{log-BS}(\alpha, \mu)$ com parâmetros de forma e localização, dados por $\alpha > 0$ e $\mu = \log(\beta) \in \mathbb{R}$, respectivamente.

Em uma formulação alternativa, uma variável aleatória Y segue uma distribuição log-BS, se e somente se

$$Z = \frac{2}{\alpha} \sinh \left(\frac{Y - \mu}{2} \right) \sim N(0, 1).$$

Então, a f.d.a. de Y é dada por

$$F_Y(y; \alpha, \mu) = \Phi \left[\frac{2}{\alpha} \sinh \left(\frac{y - \mu}{2} \right) \right], \quad y \in \mathbb{R}, \mu \in \mathbb{R}, \alpha > 0. \quad (2.2)$$

Derivando (2.2) com relação a y obtemos a f.d.p. dada por

$$f_Y(y; \alpha, \mu) = \frac{1}{\alpha\sqrt{2\pi}} \exp \left[-\frac{2}{\alpha^2} \sinh^2 \left(\frac{y - \mu}{2} \right) \right] \cosh \left(\frac{y - \mu}{2} \right), \quad y \in \mathbb{R}, \mu \in \mathbb{R}, \alpha > 0.$$

A média e a variância de Y podem ser obtidas usando a função geradora de momentos dada por

$$M_Y(s) = \exp(\mu s) \left[\frac{K_{[\sigma s + 1]/2}(\delta^{-2}) + K_{[\sigma s - 1]/2}(\delta^{-2})}{2K_{1/2}(\delta^{-2})} \right],$$

em que K_λ é a função Bessel modificada do terceiro tipo, dada por

$$K_\lambda(w) = \frac{w^\lambda}{2^{\lambda+1}} \int_0^\infty u^{-\lambda-1} \exp \{ -u - [w^2/4u] \} du.$$

Algumas propriedades da distribuição log-BS são apresentadas a seguir. Se $Y \sim \text{log-BS}(\alpha, \mu)$, então: (i) $T = \exp(Y) \sim \text{BS}(\alpha, \beta)$; (ii) $E[Y] = \mu = \log(\beta)$; (iii) não existe uma forma fechada para a variância de Y , porém com base em uma aproximação assintótica para a função geradora de momentos da distribuição log-BS, segue-se que, se $\alpha \rightarrow 0$, então a $\text{Var}[T] = \alpha^2 - \alpha^4/4$, por outro lado, se $\alpha \rightarrow \infty$, então $\text{Var}[T] = 4(\log^2 \sqrt{2}\alpha) + 2 - 2\log(\sqrt{2}\alpha)$; (iv) se $X = Y \pm k$, então $X \sim \text{log-BS}(\alpha, \mu \pm k)$; e (v) a distribuição log-BS é simétrica em torno de μ , unimodal para $\alpha \leq 2$ e bimodal para $\alpha > 2$ veja [Rieck e Nedelman \(1991\)](#); outra propriedade importante da distribuição BS é que sua versão logarítmica tem uma bimodalidade flexível. Note que: Se $T \sim \text{BS}(\alpha, \beta)$, então $Y = \log(T) \sim \text{log-BS}(\alpha, \mu)$, em que $\mu = \log(\beta)$. Para mais detalhes, veja [Desousa et al. \(2018\)](#) e [Leão et al. \(2017, 2018a,b\)](#).

Na Figura 2.2 (a) pode se observar algumas das características da distribuição log-BS descritas pelas propriedades acima. Observe que α modifica a forma da distribuição e, quando aumenta, a curtose da distribuição também aumenta, de modo que a curva da distribuição torna-se mais achatada. Em

particular para $\alpha \leq 2$ a distribuição é unimodal e exibe curtose menor que a distribuição normal. No entanto, se $\alpha > 2$, a distribuição começa a exibir bimodalidade, com modas amplamente separadas e a sua curtose é maior que a distribuição normal. A Figura 2.2 (b) mostra como μ modifica a localização.

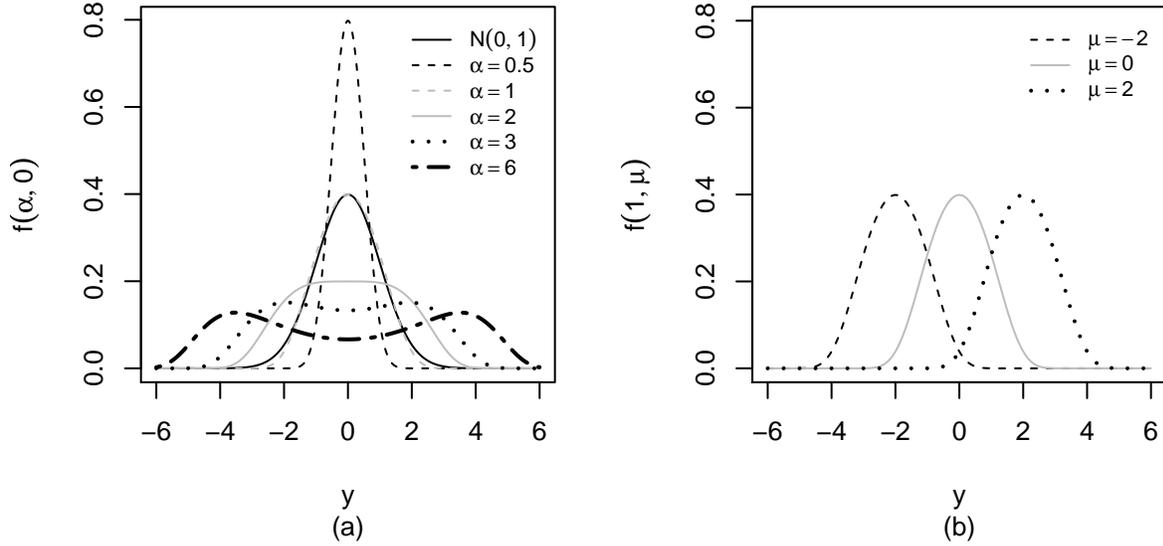


Figura 2.2: Funções densidade da distribuição Log-BS: (a) Log-BS($\alpha, 0$) e (b) Log-BS($1, \mu$) para os valores indicados dos parâmetros de forma α e localização μ .

2.3 Modelo BISARMA

2.3.1 Formulação

Sejam Y_1, \dots, Y_n variáveis aleatórias definidas em um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ e para cada $t = 1, \dots, n$, \mathcal{F}_t é a σ -álgebra gerada pela informação até o tempo t . Suponha que, para cada $t = 1, \dots, n$, a distribuição condicional de cada Y_t , dado o conjunto de informações passadas $\mathcal{F}_{t-1} = \{Y_{t-1}, \dots, Y_1, \mu_{t-1}, \dots, \mu_1\}$, segue uma distribuição log-BS, denotada por $Y_t | \mathcal{F}_{t-1} \sim \text{log-BS}(\alpha, \mu_t)$, com f.d.p. condicional dada por

$$f(y_t; \alpha, \mu_t | \mathcal{F}_{t-1}) = \frac{1}{\alpha \sqrt{2\pi}} \exp \left[-\frac{2}{\alpha^2} \sinh^2 \left(\frac{y_t - \mu_t}{2} \right) \right] \cosh \left(\frac{y_t - \mu_t}{2} \right), \quad y_t \in \mathbb{R}, \mu_t \in \mathbb{R}, \alpha > 0,$$

em que α e $\mu_t = E[Y_t | \mathcal{F}_{t-1}]$ são, respectivamente, o parâmetro de forma e a média condicional de Y_t .

A classe de modelos de regressão log-linear BS é definida por

$$Y_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, n, \quad (2.3)$$

em que $\mu_t = \mathbf{x}_t^\top \boldsymbol{\beta}$, e $\mathbf{x}_t^\top = [x_{t1}, \dots, x_{tk}]$ é um vetor contendo os valores das k variáveis explicativas, para $k < n$, n é o tamanho da amostra, $\boldsymbol{\beta} = [\beta_1, \dots, \beta_k]^\top$ é um vetor de parâmetros desconhecidos a serem estimados e ε_t é o termo do erro do modelo, que não são correlacionados, com $\varepsilon_t \sim \text{log-BS}(\alpha, 0)$. No entanto, no modelo BISARMA, o componente aleatório μ_t do modelo fornecido em (2.3) contém um elemento adicional dinâmico, τ_t , com estrutura ARMA, tal que agora

$$\mu_t = \mathbf{x}_t^\top \boldsymbol{\beta} + \tau_t. \quad (2.4)$$

O componente τ_t incluído em (2.4) é definido da seguinte forma. Considere $\psi_t = y_t - \mathbf{x}_t^\top \boldsymbol{\beta}$ como sendo um modelo ARMA(p, q), inicialmente, como função de ψ_t . Além disso, o componente ψ_t pode ser expresso por

$$\psi_t = \eta + \sum_{i=1}^p \phi_i \psi_{t-i} + \sum_{j=1}^q \theta_j u_{t-j} + u_t, \quad (2.5)$$

em que $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)^\top \in \mathbb{R}^p$, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)^\top \in \mathbb{R}^q$, e $p, q \in \mathbb{N}$ são os parâmetros autorregressivos e de médias móveis e suas ordens, respectivamente, e $\eta \in \mathbb{R}$ é uma constante. Os termos de médias móveis, u_t 's, são erros aleatórios não correlacionados (ruído branco), que podem ser, por exemplo, erros de medida na escala original (isto é, $u_t = y_t - \mu_t$). Tomando a esperança condicional com relação ao conjunto de informações passadas \mathcal{F}_{t-1} para ambos os lados da expressão (2.5), temos que

$$\begin{aligned} \mathbb{E}(\psi_t | \mathcal{F}_{t-1}) &= \eta + \sum_{i=1}^p \phi_i \mathbb{E}(\psi_{t-i} | \mathcal{F}_{t-1}) + \sum_{j=1}^q \theta_j \mathbb{E}(u_{t-j} | \mathcal{F}_{t-1}) + \mathbb{E}(u_t | \mathcal{F}_{t-1}), \\ &= \eta + \sum_{i=1}^p \phi_i \psi_{t-i} + \sum_{j=1}^q \theta_j u_{t-j} + \mathbb{E}(u_t | \mathcal{F}_{t-1}) \\ &= \eta + \sum_{i=1}^p \phi_i [y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}] + \sum_{j=1}^q \theta_j u_{t-j}. \end{aligned}$$

Dado que $\mathbb{E}(Y_t|\mathcal{F}_{t-1}) = \mu_t$ e $\mu_t = \mathbf{x}_t^\top \boldsymbol{\beta} + \tau_t$, obtemos

$$\mathbb{E}(\psi_t|\mathcal{F}_{t-1}) = \mathbb{E}(Y_t - \mathbf{x}_t^\top \boldsymbol{\beta}|\mathcal{F}_{t-1}) = \mathbb{E}(Y_t|\mathcal{F}_{t-1}) - \mathbf{x}_t^\top \boldsymbol{\beta} = \mu_t - \mathbf{x}_t^\top \boldsymbol{\beta} = \tau_t$$

e, desta forma,

$$\tau_t = \eta + \sum_{i=1}^p \phi_i [y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}] + \sum_{j=1}^q \theta_j u_{t-j}. \quad (2.6)$$

Portanto, a partir de (2.3), (2.4) e (2.6), propomos o modelo BISARMA expresso por

$$Y_t = \eta + \mathbf{x}_t^\top \boldsymbol{\beta} + \sum_{i=1}^p \phi_i [y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}] + \sum_{j=1}^q \theta_j u_{t-j} + \varepsilon_t, \quad t = 1, \dots, n, \quad (2.7)$$

em que $\varepsilon_t \sim \text{log-BS}(\alpha, 0)$, o que implica que

$$Y_t = \log(T_t) \sim \text{log-BS} \left(\alpha, \eta + \mathbf{x}_t^\top \boldsymbol{\beta} + \sum_{i=1}^p \phi_i (y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}) + \sum_{j=1}^q \theta_j u_{t-j} \right),$$

em que

$$T_t \sim \text{BS} \left(\alpha, \exp \left[\eta + \mathbf{x}_t^\top \boldsymbol{\beta} + \sum_{i=1}^p \phi_i (y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}) + \sum_{j=1}^q \theta_j u_{t-j} \right] \right).$$

2.3.2 Estimação

Seja Y_1, \dots, Y_n uma amostra do modelo BISARMA definido em (2.7), $\boldsymbol{\gamma} = (\alpha, \boldsymbol{\beta}, \eta, \boldsymbol{\phi}, \boldsymbol{\theta})$ é o vetor de parâmetros de interesse e $m = \max\{p, q\}$, com $n > m$. Para cada $t = m+1, \dots, n$, considere a função de log-verossimilhança $\ell_t(\alpha, \boldsymbol{\beta}, \eta, \boldsymbol{\phi}, \boldsymbol{\theta}) = \log(f(y_t; \alpha, \mu_t|\mathcal{F}_{t-1}))$ dado \mathcal{F}_{t-1} . Portanto, $\ell_t = \ell_t(\alpha, \boldsymbol{\beta}, \eta, \boldsymbol{\phi}, \boldsymbol{\theta})$ é expressa como

$$\ell_t \propto -\log(\alpha) + \log \left[\cosh \left(\frac{y_t - \mu_t}{2} \right) \right] - \frac{2}{\alpha^2} \sinh^2 \left(\frac{y_t - \mu_t}{2} \right).$$

Assim, a função de log-verossimilhança para $\boldsymbol{\gamma}$ do modelo BISARMA é dada por

$$\ell = \sum_{t=m+1}^n \ell_t(\alpha, \boldsymbol{\beta}, \eta, \boldsymbol{\phi}, \boldsymbol{\theta}). \quad (2.8)$$

Para realizar a estimação dos parâmetros no modelo BISARMA, calculamos o vetor escore, que é definido como $\boldsymbol{U}(\boldsymbol{\gamma})$, formado pelas derivadas parciais de primeira ordem, com relação a cada um

dos parâmetros, da função log-verossimilhança . Assim, a partir de (2.8), temos

$$\frac{\partial \ell_t}{\partial \alpha} = -\frac{1}{\alpha} + \frac{4}{\alpha^3} \left[\sinh \left(\frac{y_t - \mu_t}{2} \right) \right]^2$$

e conseqüentemente

$$\frac{\partial \ell}{\partial \alpha} = \frac{1}{\alpha} \sum_{t=m+1}^n \left[\frac{4}{\alpha^2} \sinh^2 \left(\frac{y_t - \mu_t}{2} \right) - 1 \right].$$

Observe que

$$\frac{\partial \ell_t}{\partial \mu_t} = \tanh \left(\frac{y_t - \mu_t}{2} \right) \left[\frac{2}{\alpha^2} \cosh^2 \left(\frac{y_t - \mu_t}{2} \right) - \frac{1}{2} \right].$$

Portanto,

$$\frac{\partial \ell_t}{\partial \beta_k} = \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \beta_k} = \left(x_{t,k} - \sum_{k=1}^p \phi_i x_{t-i,k} \right) \frac{\partial \ell_t}{\partial \mu_t},$$

$$\frac{\partial \ell_t}{\partial \eta} = \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta} = \frac{\partial \ell_t}{\partial \mu_t}.$$

Conseqüentemente, os escores associadas aos parâmetros β_k e η são

$$\frac{\partial \ell}{\partial \beta_k} = \sum_{t=m+1}^n \left(x_{t,k} - \sum_{k=1}^p \phi_i x_{t-i,k} \right) \tanh \left(\frac{y_t - \mu_t}{2} \right) \left[\frac{2}{\alpha^2} \cosh^2 \left(\frac{y_t - \mu_t}{2} \right) - \frac{1}{2} \right],$$

$$\frac{\partial \ell}{\partial \eta} = \sum_{t=m+1}^n \tanh \left(\frac{y_t - \mu_t}{2} \right) \left[\frac{2}{\alpha^2} \cosh^2 \left(\frac{y_t - \mu_t}{2} \right) - \frac{1}{2} \right].$$

Dado que

$$\frac{\partial \ell_t}{\partial \phi_i} = \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \phi_i} = (y_{t-i} - \mathbf{x}_{t-i} \boldsymbol{\beta}) \frac{\partial \ell_t}{\partial \mu_t} \quad \text{e} \quad \frac{\partial \ell_t}{\partial \theta_j} = \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \theta_j} = u_{t-j} \frac{\partial \ell_t}{\partial \mu_t},$$

obtêm-se

$$\frac{\partial \ell}{\partial \phi_i} = \sum_{t=m+1}^n (y_{t-i} - \mathbf{x}_{t-i} \boldsymbol{\beta}) \tanh \left(\frac{y_t - \mu_t}{2} \right) \left[\frac{2}{\alpha^2} \cosh^2 \left(\frac{y_t - \mu_t}{2} \right) - \frac{1}{2} \right],$$

$$\frac{\partial \ell}{\partial \theta_j} = \sum_{t=m+1}^n u_{t-j} \tanh \left(\frac{y_t - \mu_t}{2} \right) \left[\frac{2}{\alpha^2} \cosh^2 \left(\frac{y_t - \mu_t}{2} \right) - \frac{1}{2} \right].$$

O estimador ML, $\hat{\gamma}$, do vetor de parâmetros γ é obtido através das soluções do sistema de equações $\mathbf{U}_\eta(\gamma) = 0$, $\mathbf{U}_{\beta_k}(\gamma) = 0$ para $k = 1, \dots, w$, $\mathbf{U}_\alpha(\gamma) = 0$, $\mathbf{U}_{\phi_i}(\gamma) = 0$ para $i = 1, \dots, p$ e $\mathbf{U}_{\theta_j}(\gamma) = 0$ para $j = 1, \dots, q$ em que $\gamma = (\alpha, \beta, \eta, \phi, \theta)$, $\mathbf{U}_\eta(\gamma) = \partial\ell/\partial\eta$, $\mathbf{U}_{\beta_k}(\gamma) = \partial\ell/\partial\beta_k$, $\mathbf{U}_\alpha(\gamma) = \partial\ell/\partial\alpha$, $\mathbf{U}_{\phi_i}(\gamma) = \partial\ell/\partial\phi_i$ e $\mathbf{U}_{\theta_j}(\gamma) = \partial\ell/\partial\theta_j$. Exceto para $\mathbf{U}_\alpha = 0$, em que o estimador de ML é dado por

$$\hat{\alpha} = \sqrt{\frac{4}{n} \sum_{t=m+1}^n \left[\sinh\left(\frac{y_t - \mu_t}{2}\right) \right]^2},$$

as equações do vetor escore $\mathbf{U}(\gamma) = \partial\ell/\partial\gamma = \mathbf{0}$ não possuem soluções analíticas e, portanto, é necessário o uso de algum método de otimização não-linear para maximizar a função log-verossimilhança definida em (2.8). O algoritmo Broyden-Fletcher-Goldfarb-Shanno (BFGS), conhecido também como algoritmo quase-Newton BFGS, é uma boa opção para a resolução de sistemas de equações não lineares, pois, na maioria dos casos, consegue alcançar a solução mais rapidamente que outros métodos. O método de quase-Newton aproxima a matriz de segundas derivadas, reduzindo assim o número de operações por iteração. Para mais detalhes sobre métodos de maximização numérica e o método BFGS, veja [Nocedal e Wright \(1999\)](#), [Press et al. \(1992\)](#) e [Lange \(2001\)](#). O método BFGS está implementado no software R pelas funções `optim` e `optimx`; veja www.R-project.org e [R Core Team \(2016\)](#). Os resultados para a matriz de informações de Fisher são fornecidos no Apêndice deste trabalho.

2.3.3 Previsão

A previsão na origem t e horizonte h é denotada por \hat{y}_{t+h} . Defina o seguinte:

$$\hat{y}_{t+h} = \begin{cases} \hat{y}_t(h), & \text{se } h > 0; \\ y_{t+h}, & \text{se } h \leq 0; \end{cases} \quad \text{e} \quad \hat{\varepsilon}_{t+h} = \begin{cases} 0, & \text{se } h > 0; \\ \hat{\varepsilon}_{t+h}, & \text{se } h \leq 0. \end{cases}$$

O fato de $\hat{\varepsilon}_{t+h} = 0$ para $h > 0$ indica que a observação y_{t+h} é prevista corretamente.

As estimativas para μ_t com $t = m + 1, \dots, n$, denotada por $\hat{\mu}_t$, são obtidas a partir das estimativas ML de γ , $\hat{\gamma}$, como

$$\hat{\mu}_t = \hat{\eta} + \mathbf{x}_t^\top \hat{\beta} + \sum_{i=1}^p \hat{\phi}_i [y_{t-i} - \mathbf{x}_{t-i}^\top \hat{\beta}] + \sum_{j=1}^q \hat{\theta}_j \hat{u}_{t-j}.$$

Utilizando $\hat{\mu}_t$ podemos obter $\hat{\varepsilon}_t$, em que $t = m + 1, \dots, n$. Por exemplo, para $\varepsilon_t = y_t - \mu_t$, isto é,

erros de medida na escala original, temos que $\widehat{\varepsilon}_t = y_t - \widehat{\mu}_t$. Assim a estimativa y_{n+1} é dada por

$$\widehat{y}_{n+1} = \widehat{\eta} + \mathbf{x}_{n+1}^\top \widehat{\boldsymbol{\beta}} + \sum_{i=1}^p \widehat{\phi}_i [y_{n+1-i} - \mathbf{x}_{n+1-i}^\top \widehat{\boldsymbol{\beta}}] + \sum_{j=1}^q \widehat{\theta}_j \widehat{u}_{n+1-j}.$$

Para o tempo $n + 2$, obtemos

$$\widehat{y}_{n+2} = \widehat{\eta} + \mathbf{x}_{n+2}^\top \widehat{\boldsymbol{\beta}} + \sum_{i=1}^p \widehat{\phi}_i [y_{n+2-i} - \mathbf{x}_{n+2-i}^\top \widehat{\boldsymbol{\beta}}] + \sum_{j=1}^q \widehat{\theta}_j \widehat{u}_{n+2-j},$$

e assim por diante.

2.3.4 Análise residual

A análise de resíduos desempenha um papel fundamental na validação de qualquer modelo estatístico. Esta análise permite detectar a existência de possíveis observações discrepantes (*outliers*) e seu efeito na inferência ou no ajuste. Para avaliar o ajuste do modelo definido em (2.3), considerou-se o resíduo de Cox-Snell generalizado (GCS). No caso deste modelo, o resíduo GCS é definido como

$$r_t = -\log(\widehat{S}(y_t | \mathcal{F}_{t-1})), \quad (2.9)$$

em que $\widehat{S}(\cdot)$ é a função de sobrevivência estimada para o modelo avaliado no caso t , definida como

$$\widehat{S}(Y_t; \alpha, \boldsymbol{\beta}, \mathbf{x}) = \Phi \left[-\frac{2}{\widehat{\alpha}} \sinh \left(\frac{y_t - \widehat{\mu}_t}{2} \right) \right]. \quad (2.10)$$

Os resíduos do GCS seguem uma distribuição exponencial unitária, EXP (1), quando o modelo é especificado corretamente. Como os resíduos GCS têm uma distribuição de EXP (1), um gráfico de quantil teórico *versus* quantil empírico (QQ) de r_t com envelope, definido em (2.9), pode ser empregado para avaliar o ajuste do modelo.

2.4 Simulações de Monte Carlo

Nesta seção, simulações de Monte Carlo são utilizadas para avaliar o desempenho e o comportamento dos estimadores de ML dos parâmetros do modelo BISARMA. Todas as rotinas de simulação e estimativas são desenvolvidas utilizando o software R. Para o estudo, em todos os cenários, o número

de réplicas de Monte Carlo está definido em $N_r = 10.000$. As estimativas de ML são obtidas maximizando a função de log-verossimilhança definida em (2.8) utilizando-se o algoritmo BFGS. A partir das 10.000 réplicas, para cada tamanho de amostra (n) e cenários considerados, foram calculados a média dos valores estimados, o viés empírico, a variância e o erro quadrático médio (MSE), respectivamente dados por

$$\bar{\hat{\varphi}} = \frac{1}{N_r} \sum_{r=1}^{N_r} \hat{\varphi}_r, \quad \text{Viés}(\hat{\varphi}) = \bar{\hat{\varphi}} - \varphi,$$

$$\widehat{Var}(\hat{\varphi}) = \frac{1}{N_r} \sum_{r=1}^{N_r} (\hat{\varphi}_r - \bar{\hat{\varphi}})^2 \quad \text{e} \quad \widehat{MSE}(\hat{\varphi}) = \frac{1}{N_r} \sum_{r=1}^{N_r} (\hat{\varphi}_r - \varphi)^2,$$

em que $\hat{\varphi}_r$ é a estimativa obtida na r -ésima réplica, φ é o valor verdadeiro do parâmetro, e N_r é o número de réplicas de Monte Carlo. Com exceção da média, para todas as outras estatísticas, à medida que o valor diminui, o desempenho da estimativa melhora. O viés tem essa característica, mas quando analisado em termos do seu valor absoluto.

2.4.1 Modelo BISARMA(1,1)

A seguir, são apresentados os resultados do estudo de simulação realizado usando o modelo BISARMA (1,1). Avaliou-se o desempenho do estimador de ML para os parâmetros de forma (α), autorregressivos (ϕ) e de média móvel (θ). Especificamente, o estudo é realizado em dois casos. Para o primeiro caso, considerou-se os tamanhos de amostra $n \in \{50, 100, 200, 500\}$, com os seguintes valores para os parâmetros de interesse: $\alpha \in \{0.25, 0.5, 1.5, 2.5\}$; $\beta = 0.7$, $\eta = 1.0$, $\phi = 0.7$ e $\theta = 0.5$. Os tamanhos de amostra considerados permitem verificar se há melhorias na estimativa dos parâmetros do modelo à medida que tamanho da amostra aumenta. O conjunto de valores para α , é escolhido para obter diferentes formas. Para o segundo caso, utilizou-se tamanhos de amostra $n \in \{100, 200, 500\}$, com os valores de $\alpha = 0.5$, $\beta = 0.7$ e $\eta = 1.0$. Os parâmetros ϕ e θ são fixados em valores de 0.3, 0.5 e 0.7. Em ambos os casos, foram considerados 10.000 réplicas de Monte Carlo para cada tamanho de amostra. Os resultados das estatísticas calculadas, para o primeiro e segundo casos são expostos nas Tabelas 2.1 e 2.2.

Os resultados da Tabela 2.1 mostram que, em geral, o desempenho das estimativas de α está diretamente relacionado ao tamanho da amostra; isto é, à medida que n aumenta, a precisão das estimativas melhora, como esperado. Estas conclusões também são válidas para α , β , η , ϕ e θ utilizados no primeiro caso.

Tabela 2.1: Estimativas de ML para α indicado, com base em simulações de Monte Carlo para o modelo BISARMA (1,1).

n	α	$\hat{\alpha}$			
		Média	Viés	Variância	MSE
50	0.25	0.2353	-0.0147	0.0006	0.0008
	0.5	0.4697	-0.0303	0.0025	0.0034
	1.5	1.3972	-0.1028	0.0232	0.0337
	2.5	2.3268	-0.1732	0.0696	0.0995
100	0.25	0.2429	-0.0071	0.0003	0.0004
	0.5	0.4854	-0.0146	0.0013	0.0015
	1.5	1.4507	-0.0493	0.0116	0.0140
	2.5	2.4187	-0.0813	0.0337	0.0403
200	0.25	0.2465	-0.0035	0.0002	0.0002
	0.5	0.4927	-0.0073	0.0006	0.0007
	1.5	1.4753	-0.0247	0.0057	0.0063
	2.5	2.4599	-0.0401	0.0164	0.0180
500	0.25	0.2487	-0.0013	0.0001	0.0001
	0.5	0.4972	-0.0028	0.0003	0.0003
	1.5	1.4907	-0.0093	0.0023	0.0024
	2.5	2.4850	-0.0150	0.0064	0.0066

Observe que quando o tamanho da amostra aumenta de $n = 50$ para $n = 500$ o viés em valor absoluto do estimador de $\alpha = 0.5$, em média, diminui consideravelmente, passando de 0.0303 para 0.0028. Por outro lado, para um tamanho de amostra fixo, o viés absoluto dos estimadores aumenta à medida que α aumenta. Por exemplo, quando $n = 100$ e $\alpha = 0.25$ o viés absoluto de $\hat{\alpha}$ é 0.0071. No entanto, quando $\alpha = 2.5$ esse viés aumenta para 0.0813. Tal comportamento é semelhante para a variância e o MSE do estimador de α . Em todos os cenários considerados, o parâmetro α é, em média, subestimado, isto é, $\hat{\alpha}$ é menor que o valor verdadeiro do parâmetro. Observe que a variabilidade de $\hat{\alpha}$ diminui quando n aumenta, como esperado, com resultados semelhantes para o MSE. Isso indica que o estimador do parâmetro α do modelo BISARMA(1,1) obtido pelo estimador ML é acurado.

A Figura 2.3 mostra os resultados em forma gráfica para os valores indicados de n e α . Observe na Figura 2.3 (a) que, quando n aumenta, o viés do estimador é menor em valor absoluto. Também para um tamanho de amostra fixo, o viés em valor absoluto aumenta à medida que α aumenta. Este comportamento é semelhante para a variância e o MSE de $\hat{\alpha}$; veja Figuras 2.3 (b) e (c).

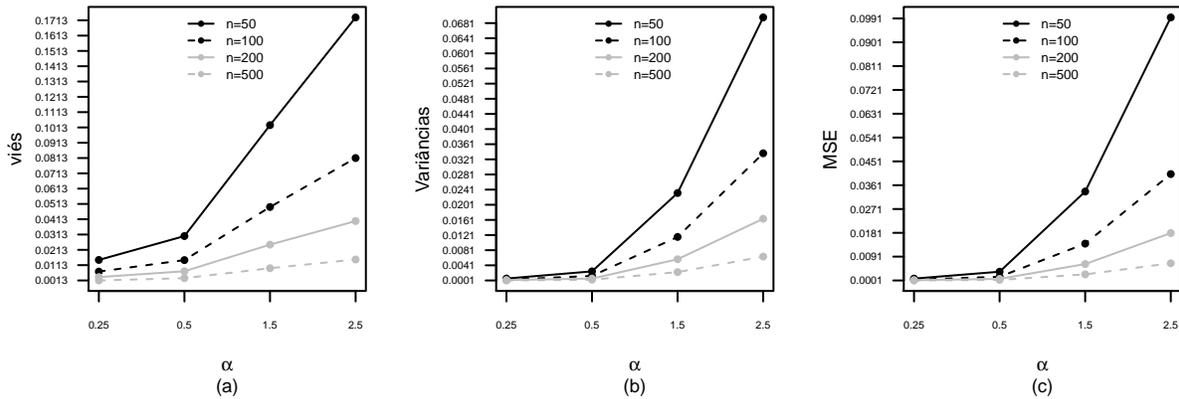


Figura 2.3: Viés absoluto empírico (a), variância (b) e MSE (c) do estimador de α .

A Tabela 2.2 apresenta algumas estatísticas resumo para as estimativas dos parâmetros ϕ e θ . Note que o estimador ML para ϕ e θ do modelo BISARMA(1,1) fornecem estimativas bastante precisas para o valor verdadeiro dos dois parâmetros. Por exemplo, para um tamanho de amostra $n = 500$, $\phi = 0.5$ e $\theta = 0.3$ as estimativas são bastante próximas do valor verdadeiro dos parâmetros, isto é, $\hat{\phi} = 0.4914$ e $\hat{\theta} = 0.3067$. Em média, o viés absoluto empírico de $\hat{\phi}$ e $\hat{\theta}$ é sempre menor que 0.0490. Os maiores valores dos MSE's empíricos são observados para $\phi = 0.3$, $\theta = 0.3$ e $n = 100$. Também pode ser visto que os valores obtidos pelo estimador de ML ϕ e θ se tornam mais precisos à medida que o tamanho da amostra aumenta. Considerando um tamanho de amostra fixo, nota-se uma ligeira redução da variância e do MSE de $\hat{\phi}$ ou $\hat{\theta}$ à medida que os parâmetros ϕ e/ou θ aumentam. Note também que ϕ é, em média, subestimado. isto é, $\hat{\phi}$ é menor que seu verdadeiro valor, em todos os cenários considerados, enquanto, θ é superestimado.

Tabela 2.2: Estimativas ML para os diferentes valores de ϕ e θ , baseadas em simulações de Monte Carlo para o modelo BISARMA(1,1).

n	ϕ	θ	$\hat{\phi}$				$\hat{\theta}$			
			Média	Viés	Variância	MSE	Média	Viés	Variância	MSE
100	0.3	0.3	0.2510	-0.0490	0.0364	0.0388	0.3435	0.0435	0.0377	0.0396
		0.5	0.2584	-0.0416	0.0243	0.0261	0.5383	0.0383	0.0236	0.0250
		0.7	0.2644	-0.0356	0.0185	0.0198	0.7311	0.0311	0.0162	0.0172
	0.5	0.3	0.4562	-0.0438	0.0188	0.0207	0.3367	0.0367	0.0233	0.0246
		0.5	0.4595	-0.0405	0.0149	0.0165	0.5354	0.0354	0.0178	0.0190
		0.7	0.4632	-0.0368	0.0124	0.0138	0.7306	0.0306	0.0140	0.0149
		0.3	0.6575	-0.0425	0.0097	0.0115	0.3334	0.0334	0.0166	0.0177
		0.5	0.6587	-0.0413	0.0086	0.0103	0.5343	0.0343	0.0145	0.0156
		0.7	0.6607	-0.0393	0.0076	0.0091	0.7322	0.0322	0.0126	0.0136
200	0.3	0.3	0.2768	-0.0232	0.0170	0.0175	0.3198	0.0198	0.0177	0.0181
		0.5	0.2789	-0.0211	0.0115	0.0119	0.5182	0.0182	0.0110	0.0113
		0.7	0.2828	-0.0172	0.0090	0.0093	0.7145	0.0145	0.0078	0.0080
	0.5	0.3	0.4794	-0.0206	0.0087	0.0091	0.3165	0.0165	0.0111	0.0113
		0.5	0.4800	-0.0200	0.0070	0.0074	0.5163	0.0163	0.0084	0.0087
		0.7	0.4823	-0.0177	0.0060	0.0063	0.7141	0.0141	0.0068	0.0070
		0.3	0.6799	-0.0201	0.0043	0.0047	0.3150	0.0150	0.0078	0.0081
		0.5	0.6800	-0.0200	0.0038	0.0042	0.5155	0.0155	0.0069	0.0071
		0.7	0.6812	-0.0188	0.0035	0.0039	0.7150	0.0150	0.0061	0.0063
500	0.3	0.3	0.2904	-0.0096	0.0066	0.0067	0.3080	0.0080	0.0069	0.0070
		0.5	0.2916	-0.0084	0.0045	0.0046	0.5073	0.0073	0.0043	0.0044
		0.7	0.2923	-0.0077	0.0035	0.0036	0.7062	0.0062	0.0031	0.0031
	0.5	0.3	0.4914	-0.0086	0.0034	0.0035	0.3067	0.0067	0.0043	0.0044
		0.5	0.4920	-0.0080	0.0027	0.0028	0.5066	0.0066	0.0033	0.0033
		0.7	0.4922	-0.0078	0.0023	0.0024	0.7060	0.0060	0.0026	0.0027
		0.3	0.6917	-0.0083	0.0016	0.0017	0.3061	0.0061	0.0030	0.0031
		0.5	0.6919	-0.0081	0.0015	0.0015	0.5064	0.0064	0.0026	0.0027
		0.7	0.6919	-0.0081	0.0014	0.0014	0.7061	0.0061	0.0023	0.0024

2.4.2 Medidas de desempenho e seleção de modelos

Uma etapa fundamental na modelagem estatística é a seleção de um modelo específico que melhor se ajusta aos dados, o que também é válido ao prever séries temporais. Diversas medidas de desempenho são utilizadas para estimar a precisão das previsões e comparar diferentes modelos. Além disso, a qualidade do ajuste dos modelos BISARMA e ARMA gaussiano pode ser comparada com essas medidas. Para avaliar e comparar a adequação do modelo BISARMA, gerou-se valores simulados do modelo BISARMA e comparou-se a capacidade preditiva desse modelo com o ARMA gaussiano

através da computação da raiz do erro quadrático médio (RMSE) dado por

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2},$$

em que n é o número de observações da série, y_t é o valor observado no instante t e \hat{y}_t é o valor previsto para y_t . Utilizou-se ainda, o critério de informação de Akaike (AIC) proposto por [Akaike \(2012\)](#) e o critério de informação Bayesiano (BIC) proposto por [Schwarz \(1978\)](#), os quais são definidos como

$$\text{AIC} = -2\ln(L) + 2k, \quad \text{BIC} = -2\ln(L) + 2k\ln(n),$$

em que L é o valor maximizado da função de verossimilhança para o modelo estimado, n é o número de observações da amostra e k é o número de parâmetros estimados. O modelo que fornece o menor valor de AIC e BIC é selecionado como o modelo de melhor ajuste.

A Tabela 2.3 apresenta alguns dos resultados dessas medidas para os tamanhos de amostra $n \in \{100, 200, 500\}$ do modelo BISARMA(1,1), com $\eta = 1.0$, $\beta = 0.7$, $\alpha = 0.5$ e $\phi, \theta \in \{0.3, 0.5, 0.7\}$. Para cada combinação de parâmetros são utilizadas 1000 replicações de Monte Carlo. Comparando os modelos BISARMA(1,1) e ARMA (1,1) gaussiano com base nas estatísticas descritas, na Tabela 2.3 verifica-se que os valores de AIC e/ou BIC destacam o fato de que o modelo BISARMA se ajusta melhor aos dados do que o modelo ARMA gaussiano. O modelo BISARMA também fornece valores menores de RMSE, indicando uma melhor capacidade de previsão e ajuste.

Tabela 2.3: Estatísticas de avaliação de previsão para os diferentes valores de ϕ e θ , baseadas em simulações de Monte Carlo para o modelo BISARMA(1,1) e ARMA(1,1) (com as estatísticas entre parênteses).

n	ϕ	θ	AIC	BIC	RMSE
100	0.3	0.3	141.8991 (142.7913)	154.9250 (155.8172)	0.4729 (0.4814)
		0.5	141.9712 (143.7678)	154.9971 (156.7937)	0.4731 (0.4845)
		0.7	141.9093 (145.7418)	154.9351 (158.7677)	0.4728 (0.4915)
	0.5	0.3	141.7895 (142.8292)	154.8154 (155.8551)	0.4726 (0.4822)
		0.5	141.8908 (143.9018)	154.9166 (156.9276)	0.4729 (0.4860)
		0.7	141.8989 (146.1938)	154.9248 (159.2196)	0.4727 (0.4945)
	0.7	0.3	141.6287 (142.8097)	154.6546 (155.8356)	0.4722 (0.4830)
		0.5	141.7572 (143.9976)	154.7831 (157.0235)	0.4726 (0.4876)
		0.7	141.8046 (146.6471)	154.8304 (159.6729)	0.4725 (0.4977)
200	0.3	0.3	280.6618 (281.7446)	297.1533 (298.2362)	0.4787 (0.4829)
		0.5	280.8021 (282.8071)	297.2937 (299.2987)	0.4789 (0.4842)
		0.7	280.8887 (285.1818)	297.3803 (301.6734)	0.4790 (0.4888)
	0.5	0.3	280.5644 (281.7629)	297.0559 (298.2545)	0.4786 (0.4833)
		0.5	280.7148 (282.9062)	297.2064 (299.3978)	0.4788 (0.4848)
		0.7	280.9012 (285.7727)	297.3928 (302.2643)	0.4790 (0.4905)
	0.7	0.3	280.4677 (281.7850)	296.9593 (298.2765)	0.4785 (0.4839)
		0.5	280.6136 (283.0111)	297.1051 (299.5027)	0.4787 (0.4856)
		0.7	280.8398 (286.4212)	297.3314 (302.9128)	0.4789 (0.4924)
500	0.3	0.3	699.6953 (701.3716)	720.7683 (722.4447)	0.4837 (0.4854)
		0.5	700.0944 (702.7698)	721.1674 (723.8428)	0.4839 (0.4863)
		0.7	699.3433 (704.6478)	720.4163 (725.7208)	0.4835 (0.4875)
	0.5	0.3	699.5842 (701.3973)	720.6573 (722.4703)	0.4836 (0.4855)
		0.5	699.9943 (702.8966)	721.0674 (723.9697)	0.4838 (0.4866)
		0.7	699.2964 (705.2882)	720.3694 (726.3612)	0.4835 (0.4882)
	0.7	0.3	699.4819 (701.4246)	720.5549 (722.4977)	0.4836 (0.4857)
		0.5	699.9119 (703.0665)	720.9849 (724.1395)	0.4838 (0.4869)
		0.7	699.1754 (706.0010)	720.2485 (727.0740)	0.4834 (0.4890)

Para verificar os efeitos do parâmetro de forma α sobre o desempenho do modelo. Considerou-se os valores de $\eta = 1.0$, $\beta = 0.7$ e $\phi = 0.6$ e $\theta = 0.2$ e $\alpha = \{0.5, 1.0, 1.5, 2.0, 2.5\}$, com 1000 réplicas de Monte Carlo. Os resultados são apresentados na Tabela 2.4. Neste caso, o modelo BISARMA também fornece valores menores de AIC, BIC e RMSE, indicando melhor capacidade de previsão e ajuste.

Tabela 2.4: Estatísticas de avaliação de previsão para os diferentes valores de α , baseadas em simulações de Monte Carlo para o modelo BISARMA(1,1) e ARMA(1,1) (com as estatísticas entre parênteses).

n	α	AIC	BIC	RMSE
100	0.5	141.9090 (142.8055)	154.9349 (155.8314)	0.4729 (0.4813)
	1.0	265.7652 (268.0712)	278.7910 (281.097)	0.8913 (0.8985)
	1.5	329.6598 (334.6509)	342.6856 (347.6767)	1.2499 (1.2553)
	2.0	369.6198 (378.0031)	382.6457 (391.0290)	1.5593 (1.5614)
	2.5	397.1729 (409.2500)	410.1987 (422.2758)	1.8297 (1.8275)
200	0.5	280.5448 (281.5542)	297.0364 (298.0457)	0.4786 (0.4831)
	1.0	-349.1334 (-347.1097)	-332.6419 (-330.6182)	0.0983 (0.1120)
	1.5	657.9656 (666.9540)	674.4572 (683.4456)	1.2625 (1.2651)
	2.0	738.2936 (753.9990)	754.7852 (770.4906)	1.5730 (1.5737)
	2.5	793.6655 (816.7358)	810.1571 (833.2274)	1.8436 (1.8419)
500	0.5	699.4671 (701.0986)	720.5401 (722.1716)	0.4836 (0.4852)
	1.0	1323.3460 (1331.902)	1344.4190 (1352.9750)	0.9101 (0.9113)
	1.5	1644.9330 (1666.830)	1666.0060 (1687.9030)	1.2736 (1.2743)
	2.0	1845.8550 (1884.830)	1866.9280 (1905.9030)	1.5853 (1.5852)
	2.5	1984.2180 (2041.935)	2005.2910 (2063.0080)	1.8563 (1.8552)

2.5 Aplicação

Nesta seção utilizou-se três conjuntos de dados reais para ilustrar o desempenho do modelo BISARMA no ajuste de séries temporais. Os dados são de frequência semanal e correspondem a três séries temporais com 508 observações. Referem-se aos possíveis efeitos da variação de temperatura e dos níveis de materiais particulado (partículas muito finas de sólidos ou líquidos suspensos no ar) sobre a mortalidade cardiovascular no condado de Los Angeles, no período compreendido entre 1970 e 1979; veja Figura 2.4. Os dados estão disponíveis no software R por meio do pacote `astsa`; veja [Shumway e Stoffer \(2017\)](#). As variáveis em estudo são mortalidade (M_t), temperatura (X_{1t}) e particulado (X_{2t}).

A Figura 2.4 (a) mostra um comportamento de tendência de queda ao longo da série M_t . Em todas as séries, a presença de picos sazonais, correspondentes às variações de inverno-verão, também é evidente.

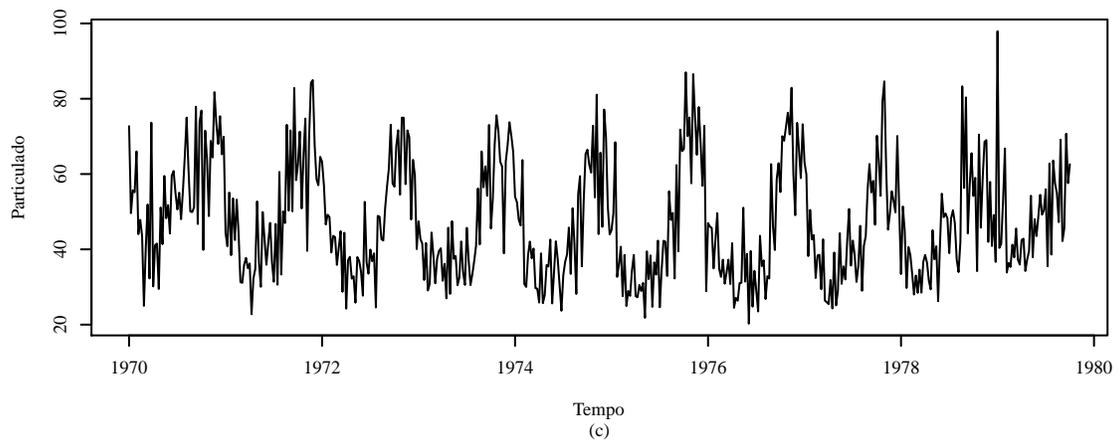
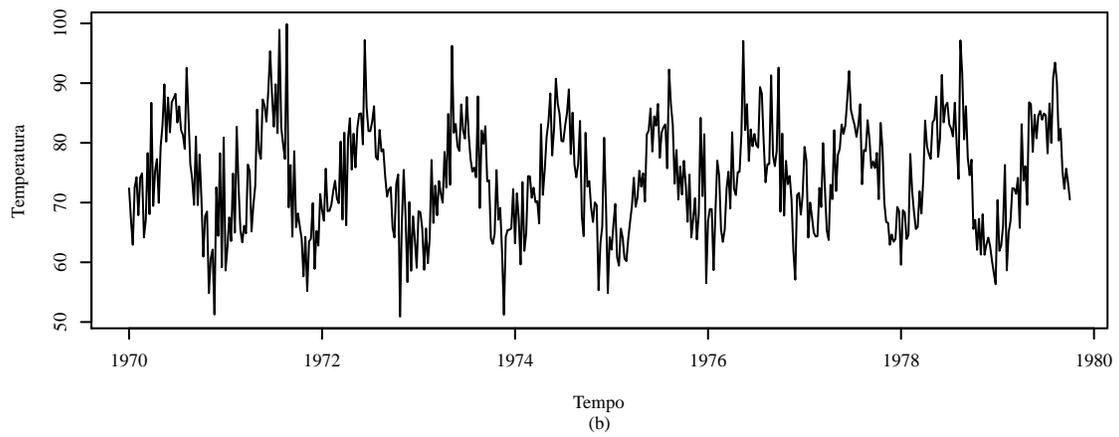
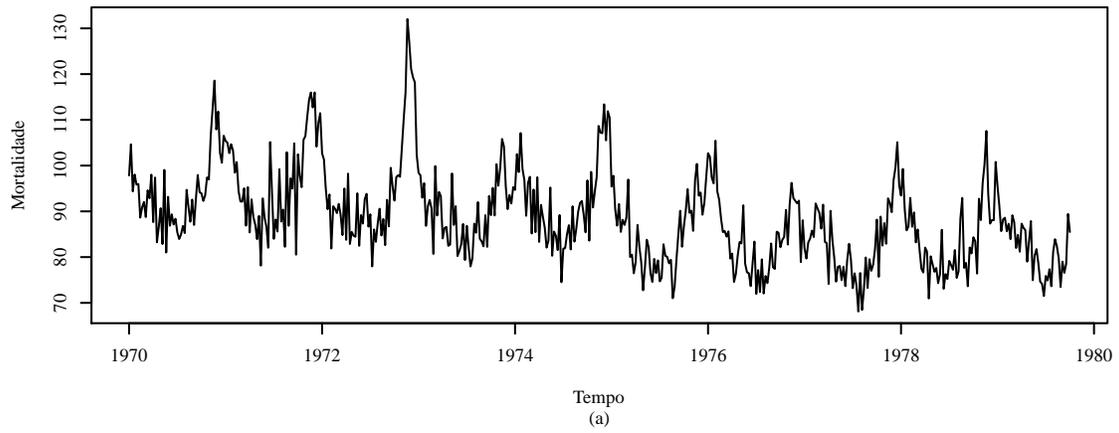


Figura 2.4: Mortalidade cardiovascular (a), Temperatura (b) e Material Particulado (c) ao longo do período de 10 anos (1970 - 1979).

A Figura 2.5 mostra os gráficos de dispersão para a relação entre mortalidade e particulado (a) e mortalidade e temperatura (b), respectivamente. A Figura 2.5 (a) indica uma possível relação linear entre a mortalidade e os níveis de materiais particulado. A Figura 2.5 (b) mostra a relação curvilínea entre a mortalidade e a temperatura, indicando que as flutuações da mortalidade estão fortemente associadas às variações da temperatura. Um estudo desse conjunto de dados pode ser visto em [Shumway e Stoffer \(2017\)](#).

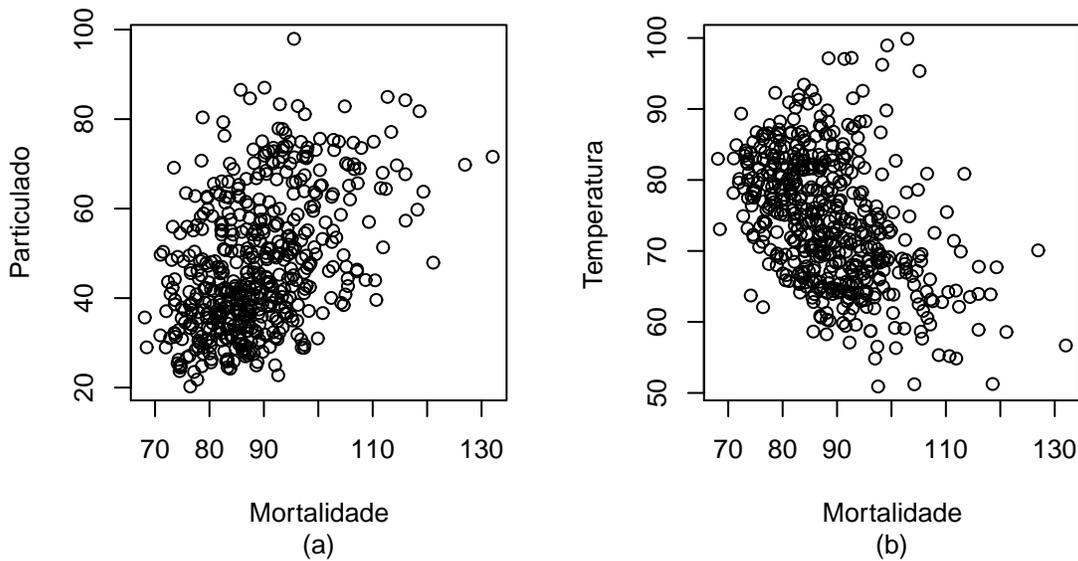


Figura 2.5: Dispersão para a relação entre mortalidade e particulado (a) e entre mortalidade e temperatura (b) durante o período de 10 anos (1970 - 1979) no condado de Los Angeles.

Baseado na análise da Figura 2.5, que mostra a relação entre as variáveis M_t , X_{1t} e X_{2t} e, considerando M_t como a variável resposta, essas relações podem ser modeladas ao longo do tempo t por meio dos valores observados de X_{1t} e X_{2t} , x_{1t} e x_{2t} como

$$M_t = \eta + \beta_1 \text{trend}_t + \beta_2(x_{1t} - \bar{x}_1) + \beta_3(x_{1t} - \bar{x}_1)^2 + \beta_4 x_{2t} + \varepsilon_t, \quad (2.11)$$

em que \bar{x}_1 é a temperatura média adicionada no modelo para evitar colinearidade, “trend_{*t*}” é uma tendência linear descendente em t observada na Figura 2.4 (a), e ε_t é um erro aleatório ou um processo de ruído consistindo de variáveis independentes que são identicamente distribuídas com média zero e variância σ_ε^2 ; veja [Shumway e Stoffer \(2017\)](#).

A Figura 2.6 mostra a função de autocorrelação (ACF) em (a) e a função de autocorrelação parcial

(PACF) em (b) dos resíduos do ajuste de mínimos quadrados de (2.11). Uma análise do gráfico da ACF e da PACF sugere a característica de um modelo estacionário AR(p) de ordem $p = 2$ para os resíduos.

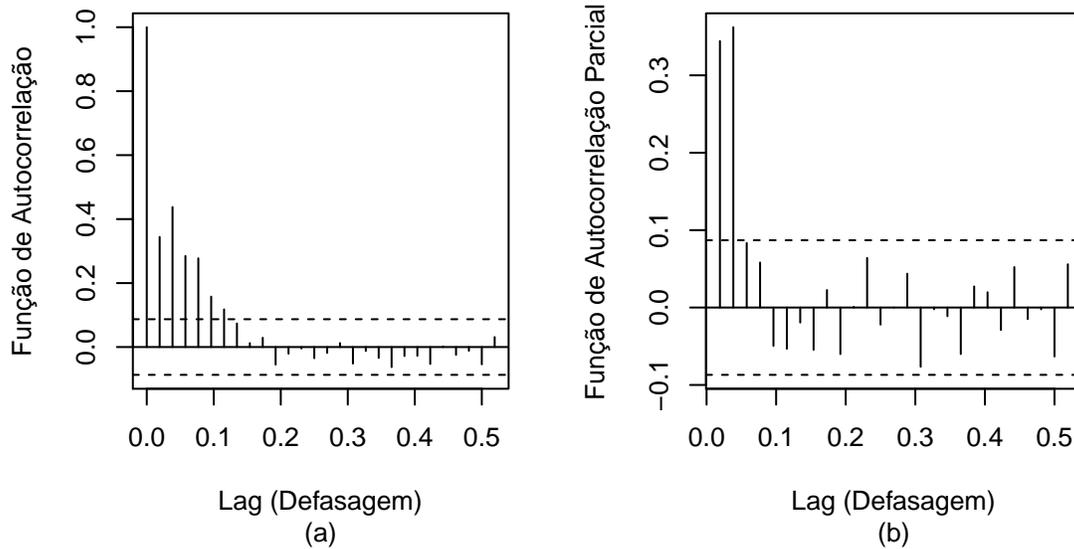


Figura 2.6: Representação gráfica das ACF (a) e PACF (b) dos resíduos do modelo de regressão para dados de séries temporais ao longo do período de 10 anos (1970 - 1979) no condado de Los Angeles.

Assim, o modelo do erro correlacionado definido em (2.11) é expresso da seguinte forma

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + u_t,$$

em que ε_t é um modelo AR(2) e u_t é um ruído branco. Os resultados para este modelo são obtidos usando a função `gammaFit` do pacote `gamlss.util`; veja <http://www.gamlss.org>.

Considere agora uma análise do modelo BISARMA definido em (2.7), com

$$Y_t = \log(M_t) = \eta + \beta_1 \text{trend}_t + \beta_2 (x_{1t} - \bar{x}_1) + \beta_3 (x_{1t} - \bar{x}_1)^2 + \beta_4 x_{2t} + \sum_{i=1}^2 \phi_i [y_{t-i} - (\eta + \beta_1 \text{trend}_{t-i} + \beta_2 (x_{1t-i} - \bar{x}_1) + \beta_3 (x_{1t-i} - \bar{x}_1)^2 + \beta_4 x_{2t-i})] + \varepsilon_t,$$

em que $\varepsilon_t \sim \log\text{-BS}(\alpha, 0)$, para o conjunto de dados associado às variáveis M_t , com $M_t = T_t$ conforme as expressões que se seguem abaixo de (2.7), X_{1t} e X_{2t} .

A Tabela 2.5 reporta os valores das estimativas de ML, RMSE e AIC / BIC. A partir desta tabela,

observe que o modelo BISARMA (2,0) fornece um melhor ajuste do que o modelo ARMA (2,0) baseado nos valores AIC / BIC. Além disso, o modelo BISARMA (2,0) apresenta menores valores de RMSE, indicando melhor capacidade de previsão.

Tabela 2.5: Estimativas e medidas de adequação de modelos para os dados da série de tempos de mortalidade no período de 10 anos (1970 - 1979) no Condado de Los Angeles.

Modelo	Parâmetros	Estimativa	AIC	BIC	RMSE
BISARMA(2,0)	ϕ_1	0.4057	-1487.9150	-1454.0710	0.0547
	ϕ_2	0.2779			
	η	38.0662			
	β_1	-0.0171			
	β_2	-0.0017			
	β_3	0.0002			
	β_4	0.0023			
	α	0.0548			
ARMA(2,0)	ϕ_1	0.3881	3100.1290	3133.9730	8.1362
	ϕ_2	0.4321			
	η	2831.4911			
	β_1	-1.3932			
	β_2	-0.0169			
	β_4	0.1554			

Os gráficos do quantil teórico *versus* quantil empírico (QQ) dos resíduos do GCS, com envelopes de simulação, indicam melhor concordância com a distribuição do EXP (1) no modelo BISARMA (2,0); veja a Figura 2.7 (à esquerda) e a Figura 2.8 (à esquerda). No entanto, nos gráficos ACF e PACF, observe que ambos os modelos produzem resíduos não autocorrelacionados; veja as Figuras 2.7 e 2.8 (ao centro e à direita).

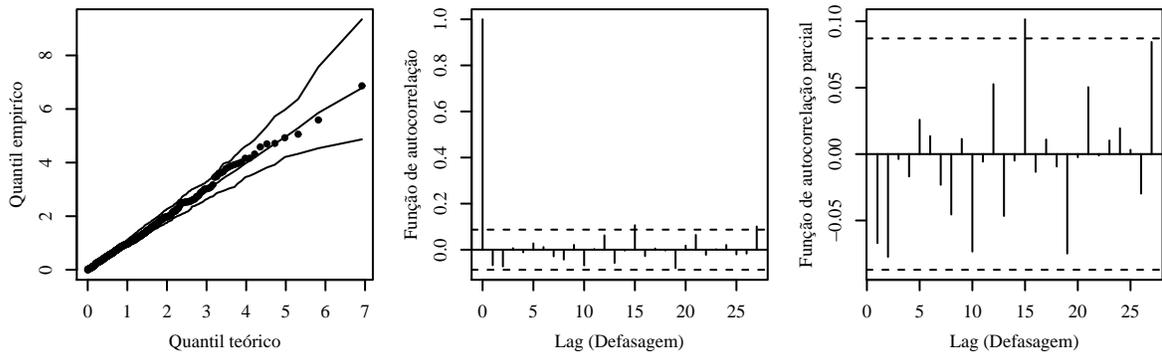


Figura 2.7: Envelope de resíduos Cox-Snell generalizado (à esquerda), ACF (centro) e PACF (à direita) do modelo BISARMA(2,0) com dados de séries temporais ao longo do período de 10 anos (1970 - 1979) no Condado de Los Angeles.

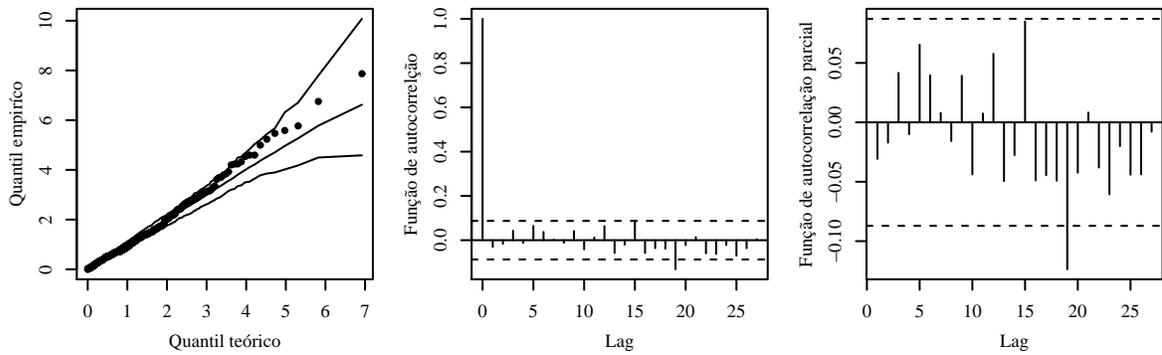


Figura 2.8: Envelope de resíduos Cox-Snell generalizado (à esquerda), ACF (centro) e PACF (à direita) do modelo ARMA(2,0) com dados de séries temporais ao longo do período de 10 anos (1970 - 1979) no Condado de Los Angeles.

As séries cronológicas ajustadas pelos modelos BISARMA (2,0) e ARMA (2,0) são apresentadas em conjunto com as séries temporais observadas na Figura 2.9.

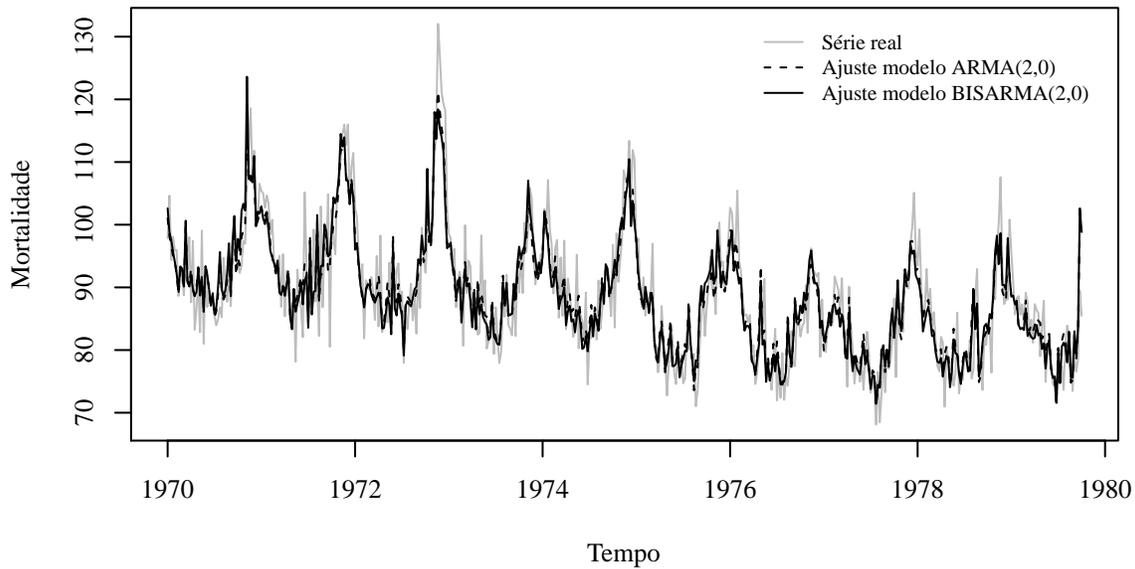


Figura 2.9: Série mortalidade cardiovascular no Condado de Los Angeles (cinza), ajustada conforme os modelos BISARMA(2,0) (preto) e ARMA(2,0) (preto - - -).

2.6 Considerações finais

Neste trabalho, foi proposto um novo modelo autorregressivo de média móvel baseado na distribuição Birnbaum-Saunders. O modelo provou ser uma boa alternativa para descrever dados nos quais a variável dependente é autocorrelacionada. Foi realizado uma avaliação numérica dos estimadores de máxima verossimilhança dos parâmetros do modelo através da simulações de Monte Carlo utilizando como medidas de desempenho a média empírica, viés, variância e erro quadrático médio dos estimadores. O estudo de simulação mostrou o bom desempenho dos estimadores de máxima verossimilhança. Uma análise residual para verificar a qualidade do ajuste também foi realizada. Finalmente, uma aplicação do novo modelo autorregressivo Birnbaum-Saunders foi realizada com dados do mundo real das ciências ambientais. A aplicação mostrou a superioridade do novo modelo sobre o clássico modelo ARMA gaussiano, fornecendo fortes evidências de que a distribuição Birnbaum-Saunders é uma boa alternativa de modelagem para lidar com dados temporais. Estes resultados sugerem que o modelo autorregressivo de médias móveis Birnbaum-Saunders (BISARMA) pode se tornar um novo padrão para a análise de rotina de dados de séries temporais positivos e assimétricos nas ciências ambientais e em outros campos do conhecimento.

Capítulo 3

RBSARMA: Um modelo de séries temporais Birnbaum-Saunders reparametrizado e sua aplicação no estudo de um processo de dados ambientais

Resumo: A classe de modelos Birnbaum-Saunders (BS), sobretudo os modelos de regressão BS, vêm recebendo considerável destaque na literatura estatística, sendo aplicados com bastante sucesso em diferentes áreas do conhecimento, tais como: negócios, finanças, seguros, medicina, meio ambiente, entre outras. Devido a essa importância, três diferentes modelos de regressão BS foram propostos em [Rieck e Nedelman \(1991\)](#), [Leiva et al. \(2014a\)](#) e [Balakrishnan e Zhu \(2015\)](#). Um ponto comum nesses modelos é que a presença de correlação não é considerada. Neste artigo, estudamos um modelo de série temporal baseado na distribuição BS que permite analisar dados em termos de uma média condicional variável no tempo. O modelo pode ser considerado como uma extensão do modelo de regressão proposto por [Leiva et al. \(2014a\)](#) e é baseado na distribuição BS reparametrizada (RBS). Particularmente, trata-se de um modelo autorregressivo de médias móveis (ARMA) com covariáveis e uma distribuição condicional RBS (RBSARMA). Além disso, um estudo de simulação de Monte Carlo é realizado para examinar o desempenho dos estimadores propostos. A metodologia é ilustrada com dados de séries temporais de um processo ambiental.

Palavras-chave: Modelo ARMA; Distribuição Birnbaum-Saunders; Dados dependentes; Métodos de máxima verossimilhança; Seleção de modelos; Método de Monte Carlo; Linguagem computacional R; Análise residual.

3.1 Introdução

A distribuição Birnbaum-Saunders (BS) é um modelo do tempo de vida que, nas últimas e recentes décadas, vem sendo bastante utilizado em diferentes campos das ciências. Essa distribuição é unimo-

dal, com assimetria positiva e suporte definido nos reais positivos, indexada por dois parâmetros α e β que modificam a sua forma e a escala, respectivamente. Proposta por [Birnbaum e Saunders \(1969a\)](#) a distribuição BS teve a sua origem baseada em um problema físico relacionado a um tipo específico de fadiga nos materiais sob estresse cíclico, onde o crescimento de uma rachadura causa a falha. Posteriormente, [Desmond \(1985\)](#) relaxou algumas suposições feitas por [Birnbaum e Saunders \(1969a\)](#) e reforçou a justificativa física desse modelo apresentando uma derivação mais geral. Para mais detalhes sobre a distribuição BS no que diz respeito às suas propriedades, veja [Johnson et al. \(1995, pags. 651-663\)](#) e [Leiva \(2016\)](#).

Desde a sua primeira utilização e numerosas aplicações na área de engenharia e confiabilidade de materiais, a família de distribuição BS tem sido considerada em diferentes campos do conhecimento, incluindo ciências ambientais; veja [Leiva et al. \(2009, 2016\)](#), [Vilca et al. \(2010\)](#), [Ferreira et al. \(2012\)](#), [Marchant et al. \(2013, 2018\)](#), [Saulo et al. \(2013\)](#), [Garcia-Papani et al. \(2017, 2018a,b\)](#) e [Lillo et al. \(2018\)](#). O amplo interesse nessa distribuição se deve aos seus argumentos teóricos, suas boas propriedades e sua relação próxima com a distribuição normal. Diversos estudos foram desenvolvidos em relação aos aspectos de estimação, inferência, generalizações, modelagem e diagnóstico da distribuição BS. Uma síntese sobre os principais estudos da distribuição BS pode ser consultada em [Balakrishnan e Kundu \(2018\)](#).

Na classe de modelos de regressão BS algumas formas de modelagem foram propostas. [Rieck e Nedelman \(1991\)](#) foram os pioneiros nesta linha. Eles introduziram o modelo log-linear para a distribuição BS e desenvolveram métodos de estimação. Posteriormente, outros trabalhos foram realizados sobre modelos de regressão BS, por exemplo, [Galea et al. \(2004\)](#) e [Xie e Wei \(2010\)](#) apresentaram vários métodos de diagnósticos para os modelos log-linear BS. [Leiva et al. \(2007\)](#) formularam o modelo de regressão log-BS com observações censuradas. [Díaz-García e Leiva \(2005\)](#) propuseram uma generalização da distribuição BS, que foi derivada com base em distribuições de contornos elípticos e foi denominada de distribuição BS generalizada (GBS). [Barros et al. \(2008\)](#) propuseram então, uma nova classe de modelos de regressão em que os erros seguem uma distribuição log-GBS baseada no modelo t-Student. [Barros et al. \(2009\)](#) desenvolveram um pacote no programa R para análise da distribuição GBS, considerando casos de dados com ou sem censura, possibilitando uma maior aplicação da distribuição BS em dados de análise de sobrevivência. Em todos esses modelos, a resposta original deve ser transformada em escala logarítmica, o que induz a um problema de interpretação dos resultados e, a uma redução do poder do estudo. Além disso, embora nessa escala

se esteja modelando a média $\mu = \log(\beta)$, na escala original está se modelando $\beta = \exp(\mu)$, que, no caso da distribuição BS, corresponde à mediana.

Neste contexto, [Santos-Neto et al. \(2012\)](#) propuseram várias reparametrizações da distribuição BS. Uma dessas reparametrizações indexa a distribuição BS por seus parâmetros de média e precisão, denominada de distribuição BS reparametrizada (RBS). Tal reparametrização permite modelar diretamente a média sem a necessidade de transformações de maneira similar aos modelos lineares generalizados (MLG), mas baseado na distribuição BS que não pertence a família exponencial. Entretanto, uma modelagem do tipo MLG baseada na distribuição RBS também pode ser executada; veja [Leiva et al. \(2014a\)](#). Nesses estudos, a resposta média está relacionada a um preditor linear por uma das várias funções possíveis de ligação, englobando todos os parâmetros a serem estimados. Diferentemente de todos os modelos de regressão existentes da BS, a abordagem proposta por [Leiva et al. \(2014a\)](#) permite que os dados sejam modelados em sua escala original com uma ampla flexibilidade.

Contudo, apesar do crescente interesse no estudo da distribuição BS e do desenvolvimento de uma quantidade considerável de trabalhos, pouco foi desenvolvido para a análise de dados com alguma estrutura de correlação serial. Além disso, este trabalho evidencia uma área relativamente pouco explorada no contexto de modelagem BS, que são os modelos de séries temporais baseados na distribuição BS. Alguns esforços nessa área são atribuídos a [Bhatti \(2010\)](#), [Leiva et al. \(2014b\)](#), [Saulo et al. \(2018a\)](#), [Fonseca e Cribari-Neto \(2018\)](#) e [Rahul et al. \(2018\)](#).

O principal objetivo deste trabalho é estudar um modelo de série temporal proposto por [Saulo et al. \(2018b\)](#) que é baseado na distribuição RBS. Esse modelo é especificado em termos de uma média condicional variando no tempo e pode ser considerado como uma extensão do modelo de regressão RBS proposto por [Leiva et al. \(2014a\)](#). A abordagem deste trabalho é semelhante aos trabalhos de [Benjamin et al. \(2003\)](#), [Rocha e Cribari-Neto \(2009\)](#) e [Maior e Cysneiros \(2016\)](#) que estenderam os modelos de séries temporais autorregressivos de médias móveis (ARMA) para diferentes classes de distribuições.

O restante desse trabalho está organizado da seguinte forma. A Seção 3.2 apresenta a distribuição BS na sua versão original e a sua forma reparametrizada RBS, bem como algumas propriedades de ambas as distribuições. Também, o modelo de regressão RBS proposto por [Leiva et al. \(2014a\)](#) será abordado, sendo apresentado sua forma e estimação de seus parâmetros. A Seção 3.3 formula o modelo RBSARMA, fornece estimativas com base no método de máxima verossimilhança e análise de resíduos. A Seção 3.4 fornece a realização da simulação de Monte Carlo para avaliar o desempenho

da metodologia proposta. A Seção 3.5 realiza uma aplicação do modelo RBSARMA para dados de séries temporais ambientais reais, para mostrar o seu potencial. Os resultados serão comparados com os ajustes produzidos por um modelo ARMA gaussiano. Finalmente, a Seção 3.6 discute algumas observações finais sobre este trabalho.

3.2 Definições Básicas

3.2.1 Distribuição Birnbaum-Saunders

Uma variável aleatória Y segue uma distribuição BS, usualmente denotada por $Y \sim \text{BS}(\alpha, \beta)$, se sua função de distribuição acumulada (f.d.a.) é expressa por

$$F(y; \alpha, \beta) = \Phi \left[\frac{1}{\alpha} \left(\sqrt{\frac{y}{\beta}} - \sqrt{\frac{\beta}{y}} \right) \right], \quad y > 0, \quad (3.1)$$

em que $\Phi(\cdot)$ é a f.d.a. da normal padrão, $\alpha > 0$ é o parâmetro de forma e $\beta > 0$ é tanto o parâmetro de escala quanto a mediana da distribuição.

A função densidade de probabilidade (f.d.p) de Y é obtida derivando (3.1) em relação a y , ou seja, $f(y; \alpha, \beta) = F'_Y(y; \alpha, \beta)$. De modo que,

$$f(y; \alpha, \beta) = \frac{y^{-3/2}(y + \beta)}{2\alpha\sqrt{2\pi\beta}} \exp \left\{ -\frac{1}{2\alpha^2} \left[\frac{y}{\beta} + \frac{\beta}{y} - 2 \right] \right\}, \quad y > 0.$$

Uma propriedade importante na construção desta distribuição é que uma variável aleatória $Y \sim \text{BS}(\alpha, \beta)$ pode ser construída a partir de uma variável aleatória Z com a distribuição normal padrão mediante a seguinte relação estocástica

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

em que

$$Z = \frac{1}{\alpha} \left[\sqrt{\frac{Y}{\beta}} - \sqrt{\frac{\beta}{Y}} \right] \sim \text{N}(0, 1).$$

Essa relação é extremamente útil e pode ser usada para obtenção de números pseudo-aleatórios provenientes da distribuição BS. Mais detalhes veja, por exemplo, [Leiva \(2016\)](#), [Díaz-García e Leiva](#)

(2005).

Outras propriedades da distribuição BS são apresentadas a seguir. Se $Y \sim \text{BS}(\alpha, \beta)$, então: (i) para qualquer constante real $k > 0$, têm-se $kY \sim \text{BS}(\alpha, k\beta)$, o que significa que a distribuição BS é fechada sob proporcionalidade; (ii) $Y^{-1} \sim \text{BS}(\alpha, \beta^{-1})$, implicando que a distribuição BS também é fechada sob reciprocidade; (iii) O parâmetro β é um parâmetro de escala, em que, $Y\beta^{-1} \sim \text{BS}(\alpha, 1)$. Além disso, β é a mediana da distribuição de Y , e pode ser obtida diretamente quando $q = 0,5$ de sua função quantílica dada por

$$y(q; \alpha, \beta) = F^{-1}(q; \alpha, \beta) = \beta \left[\frac{\alpha z(q)}{2} + \sqrt{\left(\frac{\alpha z(q)}{2}\right)^2 + 1} \right]^2, \quad 0 < q < 1,$$

em que $z(q)$ é a q -ésimo quantil da distribuição normal padrão e F_Y^{-1} é o inverso da função de distribuição de Y aplicada em q ; e (iv) O parâmetro α é um parâmetro de forma, tal que, quando $\alpha \rightarrow 0$, a distribuição BS tende para uma distribuição normal $N(\beta, \tau)$, onde $\tau \rightarrow 0$ quando $\alpha \rightarrow 0$. Além dessas propriedades, o r -ésimo momento da distribuição BS é dado por,

$$E(Y^r) = \beta^r \sum_{j=0}^r \binom{2r}{2j} \sum_{i=0}^j \binom{j}{i} \frac{(2(r-j+i))!}{2^{r-j+i}(r-j-i)!} \left(\frac{\alpha}{2}\right)^{2(r-j+i)}. \quad (3.2)$$

Utilizando a Equação (3.2), as expressões para a média, variância, coeficientes de assimetria e curtose de Y , são dadas respectivamente por

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

$$\alpha_3[Y] = 16\alpha^2 \frac{(11\alpha^2 + 6)}{(5\alpha^2 + 4)^3} \quad \text{e} \quad \alpha_4[Y] = 3 + 6\alpha^2 \frac{(93\alpha^2 + 41)}{(5\alpha^2 + 4)^2}.$$

Essas propriedades também são mantidas em outras parametrizações da distribuição BS como por exemplo, a distribuição Birnbaum-Saunders reparametrizada.

3.2.2 Distribuição Birnbaum-Saunders reparametrizada

A distribuição RBS foi proposta por Santos-Neto et al. (2012) como uma das várias formas de parametrizações da distribuição BS e foi introduzida utilizando uma reparametrização da distribuição BS em função de sua média. Tal distribuição, permite que diversas características de modelagem

de dados sejam consideradas; veja Santos-Neto et al. (2014) e Santos-Neto et al. (2016). Considere os parâmetros da distribuição BS definida em (3.1) como $\alpha = \sqrt{2/\delta}$ e $\beta = \mu\delta/(\delta + 1)$, os novos parâmetros da forma reparametrizada da distribuição BS são expressos da seguinte forma

$$\mu = \beta \left(1 + \frac{1}{2}\alpha^2\right) \quad \text{e} \quad \delta = \frac{2}{\alpha^2},$$

em que $\mu > 0$ é a média da distribuição e também um parâmetro de escala e $\delta > 0$ é um parâmetro de forma e precisão. Neste caso, usamos a notação $Y \sim \text{RBS}(\mu, \delta)$.

A f.d.a. da distribuição BS reparametrizada por μ e δ é dada por

$$F(y; \mu, \delta) = \Phi \left(\frac{\delta}{2} \right)^{\frac{1}{2}} \left[\left(\frac{(\delta + 1)y}{\mu\delta} \right)^{\frac{1}{2}} - \left(\frac{\mu\delta}{(\delta + 1)y} \right)^{\frac{1}{2}} \right], \quad y > 0, \quad (3.3)$$

em que $\Phi(\cdot)$ é a f.d.a da normal padrão.

A f.d.p. de Y é obtida derivando a expressão (3.3) em relação y . De modo que

$$f(y; \mu, \delta) = \frac{\exp\left\{\frac{\delta}{2}\right\} \sqrt{\delta + 1}}{4\sqrt{\pi\mu}} y^{-3/2} \left(y + \frac{\mu\delta}{(\delta + 1)} \right) \exp\left\{-\frac{\delta}{4} \left(\frac{(\delta + 1)y}{\mu\delta} + \frac{\mu\delta}{(\delta + 1)y} \right)\right\}, \quad y > 0. \quad (3.4)$$

Através das relações da distribuição BS em sua versão original, a distribuição RBS possui a seguinte relação estocástica com a distribuição normal

$$Y = \frac{\mu\delta}{\delta + 1} \left[\frac{Z}{\sqrt{2\delta}} + \sqrt{\left(\frac{Z}{\sqrt{2\delta}} \right)^2 + 1} \right]^2,$$

em que

$$Z = \left(\frac{\delta}{2} \right)^{\frac{1}{2}} \left[\left(\frac{(\delta + 1)Y}{\mu\delta} \right)^{\frac{1}{2}} - \left(\frac{\mu\delta}{(\delta + 1)Y} \right)^{\frac{1}{2}} \right] \sim N(0, 1).$$

A função quantílica para a distribuição RBS é dada por

$$y(q; \mu, \delta) = F^{-1}(q; \mu, \delta) = \frac{\mu\delta}{\delta + 1} \left[\frac{z(q)}{\sqrt{2\delta}} + \sqrt{\left(\frac{z(q)}{\sqrt{2\delta}} \right)^2 + 1} \right]^2, \quad 0 < q < 1,$$

em que $z(q)$ é o q -ésimo quantil da distribuição normal padrão e F_Y^{-1} é o inverso da função de distribuição Y aplicada em q .

As expressões para a média e a variância da distribuição RBS, são dadas, respectivamente, por

$E(Y) = \mu$ e $\text{Var}(Y) = \mu^2 \cdot [\text{CV}(Y)]^2$, em que $\text{CV}(Y) = \sqrt{2\delta + 5}/(\delta + 1) \in (0, \sqrt{5})$ é o coeficiente de variação de Y . Como mencionado, δ pode ser interpretado como um parâmetro de precisão, isto é, para valores fixos de μ , quando $\delta \rightarrow \infty$, a variância de Y tende a zero. Além disso, para valores fixos de μ se $\delta \rightarrow 0$, então $\text{Var}(Y) = 5\mu^2$. A mediana de Y é $\delta\mu/(\delta + 1)$ e representa uma proporção da média. Note que, para μ fixo $\delta\mu/(\delta + 1) \rightarrow \mu$ quando $\delta \rightarrow \infty$.

A Figura 3.1, apresenta algumas formas da distribuição de $Y \sim \text{RBS}(\mu, \delta)$. A partir da Figura 3.1 (a) observe que o parâmetro δ além de ser um parâmetro de precisão, também é um parâmetro de forma e, à medida que δ aumenta, a densidade da distribuição é mais concentrada em torno da média $\mu = 1$ e, portanto, a variabilidade diminui. Na Figura 3.1 (b), observe que o parâmetro μ mesmo sendo a média da distribuição também se comporta como um parâmetro de escala e, à medida que aumenta, provoca o aumento da variância e o achatamento na densidade da distribuição.

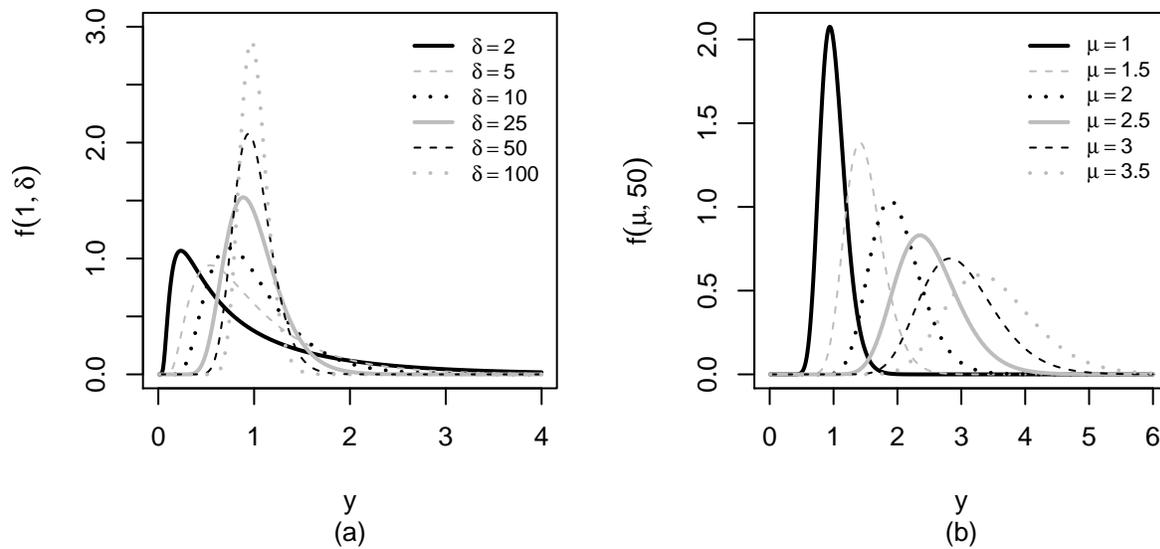


Figura 3.1: Funções densidade de probabilidade da distribuição $\text{RBS}(\mu, \delta)$ para $\mu = 1$ fixo (a) e para $\delta = 50$ fixo (b).

3.2.3 Modelo de regressão RBS

Baseados na distribuição RBS, [Leiva et al. \(2014a\)](#), propuseram uma nova abordagem ao modelo de regressão da distribuição BS. Nesta abordagem, a construção do modelo de regressão é semelhante aos MLG's, no qual modela-se diretamente a média sem que seja necessário a realização de uma transformação da variável dependente para uma escala logarítmica. Formalmente, considere

Y_1, \dots, Y_n uma amostra de variáveis aleatórias independentes, em que cada $Y_t \sim \text{RBS}(\mu_t, \delta)$, para $t = 1, \dots, n$ e suas respectivas observações $y = [y_1, \dots, y_n]^\top$. Então, um modelo de regressão baseado em (3.4) é definido pela seguinte componente sistemática

$$g(\mu_t) = \alpha_t = \mathbf{x}_t^\top \boldsymbol{\beta}, \quad t = 1, \dots, n, \quad (3.5)$$

em que $\mathbf{x}_t = (x_{t1}, \dots, x_{tp})^\top$ é um vetor de variáveis explicativas sobre p regressores conhecidos, para $t = 1, \dots, n$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ para $p < n$ é um vetor de coeficientes de regressão desconhecidos a serem estimados e α_t é o preditor linear.

Como nesse caso uma função da média é que possui relação linear com as variáveis explicativas e o vetor de parâmetros, tem-se que a média da variável resposta é dada por $\mu_t = g^{-1}(\mathbf{x}_t^\top \boldsymbol{\beta})$, em que $g(\cdot)$ é uma função de ligação $g : \mathbb{R} \rightarrow \mathbb{R}^+$ estritamente monótona, positiva e pelo menos duas vezes diferenciável.

O logaritmo da função verossimilhança do modelo de regressão RBS para o vetor de parâmetros $\boldsymbol{\gamma} = (\boldsymbol{\beta}^\top, \delta)^\top$ tem a seguinte forma $\ell(\boldsymbol{\gamma}) = \sum_{t=1}^n \ell_t(y_t; \mu_t, \delta)$, em que

$$\ell_t(y_t; \mu_t, \delta) = \frac{\delta}{2} - \frac{\log(16\pi)}{2} - \frac{1}{2} \log \left[\frac{(\delta + 1)y_t^3 \mu_t}{(\delta y_t + y_t + \delta \mu_t)^2} \right] - \frac{(\delta + 1)y_t}{4\mu_t} - \frac{\delta^2 \mu_t}{4(\delta + 1)y_t}. \quad (3.6)$$

As estimativas de máxima verossimilhança do vetor de parâmetros $\boldsymbol{\gamma}$ é dada através das soluções do sistema de equações $\mathbf{U}_{\beta_j}(\boldsymbol{\gamma}) = 0$ ($j = 1, \dots, k$) e $\mathbf{U}_\delta(\boldsymbol{\gamma}) = 0$ em que $\mathbf{U}_{\beta_j}(\boldsymbol{\gamma}) = \partial \ell(\boldsymbol{\gamma}) / \partial \beta_j$ e $\mathbf{U}_\delta(\boldsymbol{\gamma}) = \partial \ell(\boldsymbol{\gamma}) / \partial \delta$. Neste caso não é possível encontrar uma solução analítica. Sendo assim, as estimativas de máxima verossimilhança são obtidas numericamente com o uso de algum método iterativo para problemas de otimização não-linear. Os resultados teóricos para este modelo são obtidos usando a função `gamlss` do pacote RBS implementado no software R; veja [Stasinopoulos e Rigby \(2007\)](#), www.R-project.org e [R Core Team \(2016\)](#).

3.3 Modelo RBSARMA

3.3.1 Formulação

Sejam $\{Y_t\}$, para $t = 1, \dots, n$ variáveis aleatórias e assuma que a distribuição condicional de cada Y_t , dado o conjunto de informações passadas $\mathcal{F}_{t-1} = \{y_{t-1}, \dots, y_1, \mu_{t-1}, \dots, \mu_1\}$, seguem uma

distribuição RBS, denotada por $Y_t|\mathcal{F}_{t-1} \sim \text{RBS}(\mu_t, \delta)$. Então, a f.d.p. condicional de Y_t dado \mathcal{F}_{t-1} é dada por

$$f(y_t; \mu_t, \delta | \mathcal{F}_{t-1}) = \frac{\exp\left\{\frac{\delta}{2}\right\} \sqrt{\delta+1}}{4\sqrt{\pi\mu_t}} y_t^{-3/2} \left(y_t + \frac{\delta\mu_t}{(\delta+1)}\right) \exp\left\{-\frac{\delta}{4} \left(\frac{(\delta+1)y_t}{\delta\mu_t} + \frac{\delta\mu_t}{(\delta+1)y_t}\right)\right\}, \quad y_t > 0, \quad (3.7)$$

em que $\delta > 0$ e $\mu_t = E[Y_t|\mathcal{F}_{t-1}]$ são, respectivamente, o parâmetro de precisão e a média condicional de Y_t . Baseado no modelo de regressão RBS apresentado em (3.5) no modelo RBSARMA, é possível acomodar uma componente dinâmica adicional, τ_t , com uma estrutura ARMA dada por (Saulo et al., 2018b)

$$\tau_t = \eta + \sum_{i=1}^p \phi_i [g(y_{t-i}) - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}] + \sum_{j=1}^q \theta_j [g(y_{t-j}) - \alpha_{t-j}], \quad (3.8)$$

tal que agora

$$g(\mu_t) = \alpha_t = \mathbf{x}_t^\top \boldsymbol{\beta} + \tau_t, \quad t = 1, \dots, n, \quad (3.9)$$

em que $g(\cdot)$, \mathbf{x}_t e $\boldsymbol{\beta}$ são definidas como em (3.5), $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)^\top \in \mathbb{R}^p$, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)^\top \in \mathbb{R}^q$, e $p, q \in \mathbb{N}$ são os parâmetros autorregressivos e de médias móveis e suas ordens, respectivamente; e $\eta \in \mathbb{R}$ é uma constante.

Portanto,

$$\begin{aligned} g(\mu_t) = \alpha_t &= \mathbf{x}_t^\top \boldsymbol{\beta} + \eta + \sum_{i=1}^p \phi_i [g(y_{t-i}) - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}] + \sum_{j=1}^q \theta_j [g(y_{t-j}) - \alpha_{t-j}] \\ &= \eta + \left(\mathbf{x}_t^\top \boldsymbol{\beta} - \sum_{i=1}^p \phi_i \mathbf{x}_{t-i}^\top \boldsymbol{\beta} \right) + \sum_{i=1}^p \phi_i g(y_{t-i}) + \sum_{j=1}^q \theta_j g(y_{t-j}) - \alpha_{t-j}. \end{aligned} \quad (3.10)$$

O modelo RBSARMA é definido pela distribuição de Y_t dado o conjunto de informações passadas \mathcal{F}_{i-1} em (3.7) e pela componente dada em (3.10). Perceba que a formulação do modelo RBSARMA, com exceção da expressão (3.7), segue a mesma estrutura dos modelos GARMA propostos por Benjamin et al. (2003), o qual tem como objetivo estender os modelos ARMA ao contexto da família exponencial. Para o modelo RBSARMA, a função de ligação escolhida é a identidade.

3.3.2 Estimação

A estimação dos parâmetros do modelo RBSARMA é realizada considerando o método da máxima verossimilhança condicional para as primeiras m observações, em que $m = \max\{p, q\}$ e $n > m$.

A partir da expressão (3.7) tem-se que o logaritmo da função de verossimilhança para o vetor de parâmetros $\gamma = (\delta, \eta, \beta^\top, \phi^\top, \theta^\top)^\top$ condicional às m observações é dado por

$$\ell(\gamma) = \sum_{t=m+1}^n \ell_t(\delta, \beta, \eta, \phi, \theta), \quad (3.11)$$

em que $\ell_t(\delta, \beta, \eta, \phi, \theta) = \log f(y_t; \mu_t, \delta | \mathcal{F}_{t-1})$ é definido por

$$\ell_t = \left\{ \frac{\delta}{2} + \log \left(\frac{\log(16\pi)}{2} \right) - \frac{1}{2} \log \left(\frac{[\delta + 1] Y_t^3 \mu_t}{[(\delta + 1) Y_t + \delta \mu_t]^2} \right) - \frac{Y_t[\delta + 1]}{4\mu_t} - \frac{\delta^2 \mu_t}{4[\delta + 1] Y_t} \right\}.$$

A estimativa de máxima verossimilhança condicional do vetor de parâmetros γ pode ser obtida maximizando a função log-verossimilhança definida em (3.11) igualando o vetor escore $U(\gamma) = \partial \ell / \partial \gamma$ para zero. Isso é feito resolvendo um processo iterativo não-linear de otimização, em particular pode-se usar o algoritmo Broyden-Fletcher-Goldfarb-Shanno (BFGS) conhecido também como algoritmo quase-Newton BFGS. Para mais detalhes sobre métodos de maximização numérica e o método BFGS, veja [Nocedal e Wright \(1999\)](#), [Press et al. \(1992\)](#) e [Lange \(2001\)](#). O método BFGS está implementado no software R pelas funções `optim` e `optimx`; veja www.R-project.org e [R Core Team \(2016\)](#).

3.3.3 Análise residual

A análise de resíduos desempenha um papel fundamental na validação de qualquer modelo estatístico. Esta análise permite detectar a existência de possíveis observações discrepantes (*outliers*) e seus efeitos na inferência ou nos resultados do ajuste. Em particular, dois tipos de resíduos são propostos neste trabalho. O primeiro é um resíduo de Cox-Snell generalizado (GCS) dado por

$$r_t^{GCS} = -\log(\widehat{S}(y_t | \mathcal{F}_{t-1})), \quad (3.12)$$

em que $\widehat{S}(y_t | \mathcal{F}_{t-1})$ é a função de sobrevivência estimada para o modelo avaliado, definida como

$$\widehat{S}(y_t; \mu_t, \delta) = \Phi \left\{ - \left(\frac{\delta}{2} \right)^{\frac{1}{2}} \left[\left(\frac{(\delta + 1)y_t}{\mu_t \delta} \right)^{\frac{1}{2}} - \left(\frac{\mu_t \delta}{(\delta + 1)y_t} \right)^{\frac{1}{2}} \right] \right\}, \quad y_t, \mu_t, \delta > 0, \quad (3.13)$$

em que Φ é a f.d.a. da distribuição normal padrão.

O resíduo GCS segue uma distribuição exponencial unitária, EXP(1), quando o modelo é especi-

ficado corretamente, e um gráfico de quantil teórico *versus* quantil empírico (QQ) de r_t^{GCS} , definido em (3.12), pode ser empregado para avaliar o ajuste do modelo.

O quantil residual (QS) é o segundo tipo de resíduo proposto. O QS é dado por

$$r_t^{QS} = \Phi^{-1}(\widehat{S}(y_t|\mathcal{F}_{t-1})), \quad (3.14)$$

em que Φ^{-1} é a função inversa da f.d.a da distribuição normal padrão e $\widehat{S}(y_t|\mathcal{F}_{t-1})$ é a função de sobrevivência estimada, ajustada como em (3.13). O residual QS segue uma distribuição normal padrão quando o modelo é especificado corretamente. Da mesma forma, um gráfico QQ dos resíduos definidos como em (3.14), pode ser empregado para avaliar o ajuste do modelo.

3.4 Simulação de Monte Carlo

Nesta seção são apresentados os resultados das simulações realizadas no modelo RBSARMA(1,1), baseados em amostras de tamanho $n \in \{100, 200, 500\}$. Para estes estudos, dois casos foram considerados. No primeiro caso foi realizado as simulações com os valores $\delta \in \{8, 15, 25, 50\}$ e $\beta = 0.7$, $\eta = 1.0$, $\phi = 0.7$ e $\theta = 0.5$. Para o segundo caso, os parâmetros autorregressivo (ϕ) e de médias móveis (θ) tomam os valores de 0.3, 0.5 e 0.7 com $\delta = 8$, $\beta = 0.7$ e $\eta = 1.0$.

O objetivo destas simulações é avaliar o desempenho e o comportamento dos estimadores de máxima verossimilhança condicional (EMLC) para os parâmetros autorregressivos (ϕ), de médias móveis (θ) e de forma (e/ou precisão) (δ) do modelo RBSARMA(1,1). O estudo de simulação foi baseado em 1000 réplicas de Monte Carlo para cada tamanho (n) de amostra. Os tamanhos amostrais propostos têm por objetivo verificar se existe melhorias na estimação dos parâmetros de interesse conforme o aumento do tamanho da amostra. Os estimadores de máxima verossimilhança condicional foram obtidos através da maximização da expressão (3.11) utilizando o algoritmo de otimização não-linear quase-Newton BFGS. Os critérios utilizados para avaliar e comparar o desempenho dos EMLC dos parâmetros ϕ , θ e δ foram a média dos valores estimados, o viés, a variância e o erro quadrático

médio (MSE) dados, respectivamente, por

$$\begin{aligned} \bar{\hat{\varphi}} &= \frac{1}{N_r} \sum_{r=1}^{N_r} \hat{\varphi}_r, & \text{Viés}(\hat{\varphi}) &= \bar{\hat{\varphi}} - \varphi, \\ \widehat{Var}(\hat{\varphi}) &= \frac{1}{N_r} \sum_{r=1}^{N_r} (\hat{\varphi}_r - \bar{\hat{\varphi}})^2 & \text{e} & \quad \widehat{MSE}(\hat{\varphi}) = \frac{1}{N_r} \sum_{r=1}^{N_r} (\hat{\varphi}_r - \varphi)^2, \end{aligned} \tag{3.15}$$

em que $\hat{\varphi}_r$ é a estimativa obtida na r -ésima réplica para cada parâmetro, φ representa o valor verdadeiro do parâmetro e N_r é o número de réplicas de Monte Carlo.

Com exceção da média, para todas as outras estatísticas calculadas, quanto menor o seu valor, melhor o desempenho do método de estimação. Ressalta-se, que o viés tem essa característica quando analisado em termos do seu valor absoluto. Todas as rotinas de simulação e estimação foram desenvolvidas utilizando o software R.

3.4.1 Modelo RBSARMA(1,1)

Nas Tabelas 3.1 e 3.2 são apresentadas a média, o viés, a variância e o MSE das estimativas dos parâmetros de forma (e/ou precisão) (δ) e dos parâmetros autorregressivos (ϕ) e de médias móveis (θ), respectivamente, obtidas conforme a metodologia mencionada para as simulações de Monte Carlo.

Na Tabela 3.1 pode se observar as estimativas para o parâmetro δ , fixado de acordo com o primeiro caso. Note que o desempenho das estimativas de δ está relacionada com o tamanho da amostra. Por exemplo, quando aumenta o tamanho da amostra de $n = 100$ para $n = 500$ o viés em valor absoluto da estimativa de $\delta = 8$, em média, diminui consideravelmente, passando de 0.4705 para 0.0720. Consequentemente, a média da estimativa de δ tende a ficar mais próxima do verdadeiro parâmetro. Por outro lado, para um mesmo tamanho de amostra, o viés das estimativas em valor absoluto aumenta à medida que o valor de δ cresce. Por exemplo, quando $n = 200$ e $\delta = 8$ o viés absoluto da estiva de δ é de 0.2414, no entanto, quando $\delta = 50$ a medida do viés aumenta para 1.1722. Esse comportamento é semelhante, também, para a variância e o MSE das estimativas de δ .

Em todos os cenários considerados, o parâmetro δ foi, em média, superestimado, isto é, a estimativa $\hat{\delta}$ fornecida pelo EMLC para δ foi maior que o valor verdadeiro do parâmetro.

Tabela 3.1: Estimativas de máxima verossimilhança condicional para os diferentes valores de δ , baseadas em simulações de Monte Carlo para o modelo RBSARMA(1,1).

n	δ	$\hat{\delta}$			
		Média	Viés	Variância	MSE
100	8	8.4705	0.4705	1.5462	0.2334
	15	15.8429	0.8429	5.3862	0.7171
	25	26.3139	1.3139	14.8842	1.7304
	50	52.2676	2.2676	59.0348	5.1441
200	8	8.2414	0.2414	0.7216	0.0640
	15	15.4353	0.4353	2.5264	0.1926
	25	25.6816	0.6816	7.0042	0.4665
	50	51.1722	1.1722	27.9215	1.3750
500	8	8.0720	0.0720	0.2464	0.0072
	15	15.1332	0.1332	0.8625	0.0189
	25	25.2044	0.2044	2.3963	0.0425
	50	50.3276	0.3276	9.5864	0.1077

Na Figura 3.2 esses resultados são apresentados em forma de gráficos o que facilita a comparação das estatísticas calculadas em relação ao tamanho da amostra e aos valores verdadeiros de δ . Nota-se na Figura 3.2 (a) que à medida que $n \rightarrow \infty$ menores são os valores do viés das estimativas em valor absoluto. Também para um mesmo tamanho de amostra, o viés das estimativas em valor absoluto aumenta à medida que o valor de α cresce. Esse comportamento é semelhante, para a variância e o MSE das estimativas de δ (veja Figura 3.2 (b) e (c)).

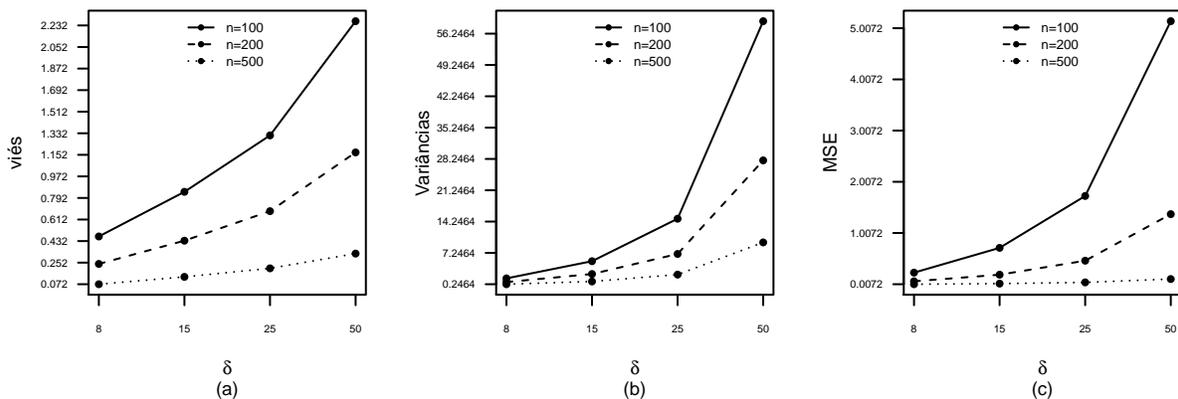


Figura 3.2: Viés absoluto empírico (a), variância (b) e MSE (c) do estimador de δ .

As análises dos gráficos da Figura 3.2 permitem concluir que, de uma maneira geral, o desempenho das estimativas de δ está diretamente relacionada com o tamanho da amostra; isto é, à medida que $n \rightarrow \infty$, menores são os valores das estatísticas de teste e, conseqüentemente, melhor é o desempenho das estimativas. Tal comportamento já era esperado, pois, quanto maior o tamanho da amostra, em princípio, mais informações se dispõe para se estimar os parâmetros.

Na Tabela 3.2 apresentam-se algumas estatísticas de resumo para as estimativas dos parâmetros ϕ e θ , fixados conforme as configurações descritas para o segundo caso. Note que as estimativas dos parâmetros ϕ e θ são bastante precisas à medida que $n \rightarrow \infty$. Isto faz com que os resultados obtidos para o MSE sejam equivalentes aos da variância estimada. Por exemplo, para um tamanho de amostra $n = 500$, $\phi = 0.5$ e $\theta = 0.3$ as estimativas são bastante próximas do valor verdadeiro dos parâmetros, isto é, $\hat{\phi} = 0.4922$ e $\hat{\theta} = 0.3034$. Em média, os vieses em valores absolutos estimados de ϕ ou θ foram sempre menores que 0.0336. Os maiores valores dos MSE estimados foram observados para $\phi = 0.3$ e $\theta = 0.3$ para um tamanho de amostra igual a 100. Considerando um tamanho de amostra fixo, nota-se um leve decaimento da variância e do MSE dos parâmetros ϕ e θ à medida que os parâmetros ϕ e/ou θ aumentam de valor.

Também foi observado que os valores estimados para ϕ e/ou θ estão, em média, subestimados. isto é, as estimativas $\hat{\phi}$ e/ou $\hat{\theta}$ foi menor que o parâmetro verdadeiro, em maioria dos cenários considerados.

Tabela 3.2: Estimativas de máxima verossimilhança condicional para os diferentes valores de ϕ e θ , baseadas em simulações de Monte Carlo para o modelo RBSARMA(1,1).

n	ϕ	θ	$\hat{\phi}$				$\hat{\theta}$			
			Média	Viés	Variância	MSE	Média	Viés	Variância	MSE
100	0.3	0.3	0.2957	-0.0043	0.0258	0.0258	0.2953	-0.0047	0.0285	0.0286
		0.5	0.3087	0.0087	0.0190	0.0191	0.4791	-0.0209	0.0207	0.0211
		0.7	0.3098	0.0098	0.0139	0.0140	0.6681	-0.0319	0.0142	0.0152
	0.5	0.3	0.4761	-0.0239	0.0149	0.0155	0.3114	0.0114	0.0188	0.0189
		0.5	0.4863	-0.0137	0.0127	0.0129	0.4946	-0.0054	0.0151	0.0152
		0.7	0.4855	-0.0145	0.0110	0.0112	0.6755	-0.0245	0.0122	0.0128
	0.7	0.3	0.6664	-0.0336	0.0084	0.0096	0.3171	0.0171	0.0136	0.0139
		0.5	0.6733	-0.0267	0.0080	0.0087	0.5005	0.0005	0.0119	0.0119
		0.7	0.6725	-0.0275	0.0074	0.0081	0.6717	-0.0283	0.0109	0.0117
200	0.3	0.3	0.2954	-0.0046	0.0143	0.0143	0.2980	-0.0020	0.0151	0.0151
		0.5	0.3004	0.0004	0.0083	0.0083	0.4913	-0.0087	0.0078	0.0079
		0.7	0.3046	0.0046	0.0066	0.0066	0.6818	-0.0182	0.0052	0.0055
	0.5	0.3	0.4853	-0.0147	0.0079	0.0082	0.3066	0.0066	0.0096	0.0096
		0.5	0.4897	-0.0103	0.0054	0.0055	0.4983	-0.0017	0.0060	0.0060
		0.7	0.4920	-0.0080	0.0048	0.0049	0.6852	-0.0148	0.0045	0.0048
	0.7	0.3	0.6811	-0.0189	0.0041	0.0045	0.3093	0.0093	0.0067	0.0068
		0.5	0.6841	-0.0159	0.0033	0.0036	0.5006	0.0006	0.0050	0.0050
		0.7	0.6868	-0.0132	0.0033	0.0034	0.6817	-0.0183	0.0044	0.0047
500	0.3	0.3	0.2958	-0.0042	0.0058	0.0058	0.3002	0.0002	0.0063	0.0063
		0.5	0.2978	-0.0022	0.0036	0.0036	0.4984	-0.0016	0.0032	0.0032
		0.7	0.3018	0.0018	0.0025	0.0025	0.6922	-0.0078	0.0017	0.0018
	0.5	0.3	0.4922	-0.0078	0.0030	0.0030	0.3034	0.0034	0.0040	0.0040
		0.5	0.4936	-0.0064	0.0023	0.0024	0.5011	0.0011	0.0024	0.0024
		0.7	0.4966	-0.0034	0.0018	0.0019	0.6936	-0.0064	0.0015	0.0016
	0.7	0.3	0.6916	-0.0084	0.0014	0.0015	0.3037	0.0037	0.0029	0.0029
		0.5	0.6927	-0.0073	0.0013	0.0014	0.5013	0.0013	0.0020	0.0020
		0.7	0.6965	-0.0035	0.0013	0.0013	0.6905	-0.0095	0.0015	0.0015

3.4.2 Medidas de desempenho e seleção de modelos

Diversas medidas de desempenho são utilizadas para estimar o nível de precisão das previsões e comparar diferentes modelos. Cada uma dessas medidas de desempenho é uma função dos valores reais e previstos da série temporal. Nesta seção, foram gerados valores simulados do modelo RBSARMA e foram estimados os modelos RBSARMA e ARMA gaussiano. Em seguida, o desempenho e o ajuste dos modelos foram comparados. Para avaliar e comparar a capacidade preditiva dos modelos utilizou-se o erro percentual absoluto médio (MAPE) dado por

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100, \quad (3.16)$$

em que n é o número de observações da série, y_t é o valor real no instante t e \hat{y}_t é o valor previsto de y_t . Entre os critérios de seleção para o melhor modelo utilizou-se o critério de informação de Akaike (AIC) e o critério de informação Bayesiano (BIC), dados, respectivamente, pelas seguintes por:

$$\text{AIC} = -2\ln(L) + 2k \quad (3.17)$$

e

$$\text{BIC} = -2\ln(L) + 2k\ln(n), \quad (3.18)$$

em que L é o valor maximizado da função de verossimilhança para o modelo estimado, n é o número de observações da amostra e k é o número de parâmetros estimados. 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 da amostra. Valores menores de AIC e/ou BIC indicam modelos com melhores ajustes; veja [Akaike \(2012\)](#) e [Schwarz \(1978\)](#).

A Tabela 3.3 apresenta alguns dos resultados dessas medidas para tamanhos de amostra $n \in \{100, 200, 500\}$ do modelo RBSARMA(1,1), com $\eta = 1.0$, $\beta = 0.7$, $\delta = 8$ e $\phi, \theta \in \{0.3, 0.5, 0.7\}$. Para cada combinação de parâmetros foram utilizadas 1000 replicações de Monte Carlo. Também ajustou-se o modelo ARMA(1,1) gaussiano com relação ao modelo RBSARMA(1,1).

Comparando os modelos RBSARMA(1,1) e ARMA (1,1) gaussiano com base nas estatísticas descritas, na Tabela 3.3 verifica-se que os valores de AIC e/ou BIC destacam o fato de que o modelo RBSARMA(1,1) ajusta-se aos dados melhor que o modelo ARMA (1,1) gaussiano. No que se refere ao desempenho das previsões dos modelos, também o modelo RBSARMA(1,1) fornece valores mais baixos do MAPE, indicando melhor capacidade de previsão.

Tabela 3.3: Estatísticas de avaliação de previsão para os diferentes valores de ϕ e θ , baseadas em simulações de Monte Carlo para o modelo RBSARMA(1,1) e ARMA(1,1) (com as estatísticas entre parênteses).

n	ϕ	θ	AIC	BIC	MAPE
100	0.3	0.3	365.1127 (424.5509)	378.1385 (437.5767)	47.5600 (50.5495)
		0.5	358.6198 (434.0864)	371.6457 (447.1122)	47.4919 (53.2613)
		0.7	353.1928 (450.4833)	366.2187 (463.5092)	47.9737 (59.0348)
	0.5	0.3	346.8663 (426.9802)	359.8921 (440.0060)	47.5695 (53.5615)
		0.5	337.8665 (441.0506)	350.8924 (454.0764)	47.5570 (58.4819)
		0.7	329.9848 (461.8964)	343.0103 (474.9222)	48.2969 (67.8709)
	0.7	0.3	304.6918 (427.5490)	317.7176 (440.5749)	47.6495 (62.3559)
		0.5	290.1231 (448.0035)	303.1489 (461.0293)	47.7313 (74.1163)
		0.7	276.5373 (474.0917)	289.5631 (487.1176)	48.8350 (94.6032)
200	0.3	0.3	730.0360 (850.0534)	746.5276 (866.5450)	48.0250 (50.4204)
		0.5	717.5611 (873.3260)	734.0527 (889.8176)	48.0832 (53.3298)
		0.7	708.0656 (915.5029)	724.5571 (931.9945)	48.5205 (59.2581)
	0.5	0.3	694.7785 (858.7896)	711.2701 (875.2812)	48.0156 (53.1801)
		0.5	677.0623 (893.5079)	693.5539 (909.9995)	48.1152 (58.4859)
		0.7	662.9814 (947.2945)	679.4730 (963.7861)	48.6620 (68.0547)
	0.7	0.3	612.9718 (873.1339)	629.4634 (889.6255)	48.0465 (61.6831)
		0.5	583.2384 (924.1664)	599.7300 (940.6580)	48.2035 (74.2198)
		0.7	558.8758 (995.1106)	575.3674 (1011.6021)	48.9833 (95.2990)
500	0.3	0.3	1830.5340 (2144.955)	1851.6070 (2166.0280)	48.4820 (50.6497)
		0.5	1798.2570 (2204.2160)	1819.3300 (2225.2890)	48.5240 (53.4874)
		0.7	1768.6360 (2307.8030)	1789.7090 (2328.8760)	48.6783 (59.2412)
	0.5	0.3	1742.7360 (2176.732)	1763.8090 (2197.8050)	48.4793 (53.3932)
		0.5	1697.1640 (2265.1910)	1718.2370 (2286.2640)	48.5363 (58.5531)
		0.7	1654.8520 (2400.0260)	1675.9250 (2421.099)	48.7389 (67.9263)
	0.7	0.3	1538.2870 (2238.3820)	1559.3600 (2259.4550)	48.4948 (62.0646)
		0.5	1462.2430 (2374.6440)	1483.3160 (2395.7170)	48.5859 (74.5613)
		0.7	1391.7030 (2559.9320)	1412.7760 (2581.00500)	48.9514 (96.0765)

A fim de verificar os efeitos do parâmetro δ no desempenho do modelo. Considerou-se os valores de $\eta = 1.0$, $\beta = 0.7$, $\phi = 0.7$, $\theta = 0.5$ e $\delta \in \{8, 15, 25, 50\}$. Foram realizadas 1000 réplicas de Monte Carlo. Os resultados estão expostos na Tabela 3.4. Neste caso, também o modelo RBSARMA(1,1) fornece valores mais baixos de AIC, BIC e MAPE, indicando melhor capacidade de previsão e ajuste.

Tabela 3.4: Estatísticas de avaliação de previsão para os diferentes valores de δ , baseadas em simulações de Monte Carlo para o modelo RBSARMA(1,1) e ARMA(1,1) (com as estatísticas entre parênteses).

n	δ	AIC	BIC	MAPE
100	8	290.1231 (448.0035)	303.1489 (461.02934)	47.7313 (74.1163)
	15	282.6548 (386.9450)	295.6806 (399.9709)	32.0136 (42.2746)
	25	258.3715 (335.5085)	271.3974 (348.5343)	23.8601 (29.8148)
	50	210.3474 (265.9603)	223.3732 (278.9862)	16.4679 (20.0521)
200	8	583.2384 (924.1664)	599.7300 (940.6580)	48.2035 (74.2198)
	15	567.9856 (790.4170)	584.4772 (806.9086)	32.2836 (42.1439)
	25	518.7238 (681.6172)	535.2154 (698.1088)	24.0106 (29.5744)
	50	421.3633 (537.5253)	437.8549 (554.0169)	16.4974 (19.7077)
500	8	1462.2430 (2374.6440)	1483.3160 (2395.7170)	48.5859 (74.5613)
	15	1425.7260 (2014.7160)	1446.7990 (2035.7890)	32.4906 (42.3578)
	25	1302.9140 (1731.6710)	1323.9870 (1752.7440)	24.1287 (29.6706)
	50	1058.8260 (1363.2620)	1079.8990 (1384.3350)	16.5303 (19.6674)

3.5 Aplicação

Esta seção trata da aplicação e da avaliação do desempenho do modelo RBSARMA no ajuste de um processo ambiental real, composto por três séries temporais. Os dados analisados estão disponíveis no software R por meio do pacote `astsa` e correspondem a 508 observações contínuas das médias semanais de mortalidade cardiovascular no condado de Los Angeles, a partir de 1970 até 1979, associadas aos efeitos da variação de temperatura e aos níveis de material particulado (partículas muito finas de sólidos ou líquidos suspensos no ar). As variáveis em estudo são mortalidade (M_t), temperatura (X_{1t}) e particulado (X_{2t}). Um estudo semelhante a este é realizado por [Shumway e Stoffer \(2017\)](#) que utilizaram o mesmo conjunto de dados para modelos de regressão no contexto de séries temporais.

O comportamento das variáveis M_t , X_{1t} e X_{2t} ao longo do tempo são apresentados na Figura 3.3. Pode constatar-se, que todas as séries são pautadas por uma visível sazonalidade que correspondem às possíveis variações de inverno-verão. Além disso, a Figura 3.3 (a) mostra uma tendência de queda ao longo da série de mortalidade

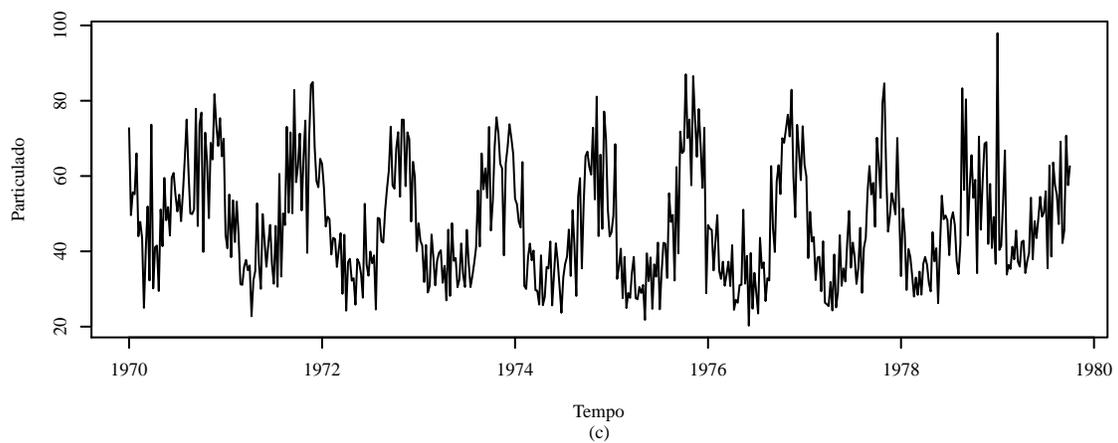
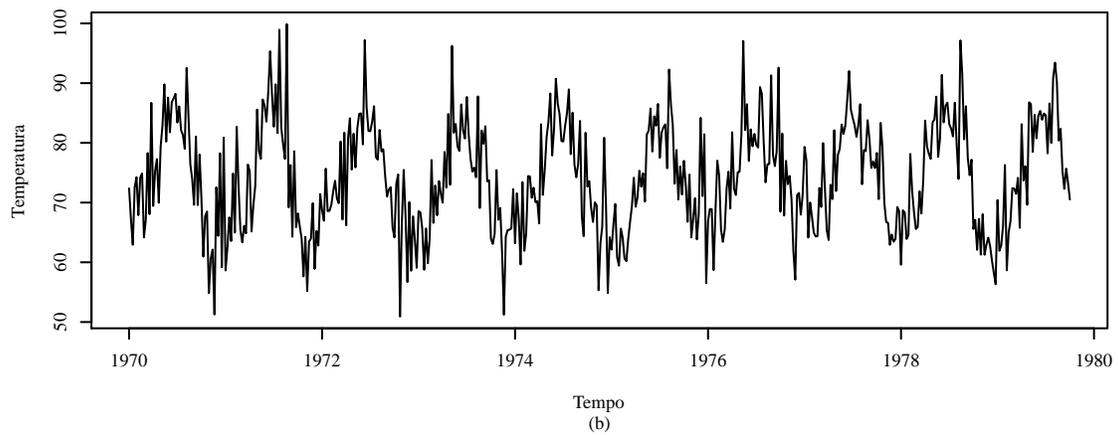
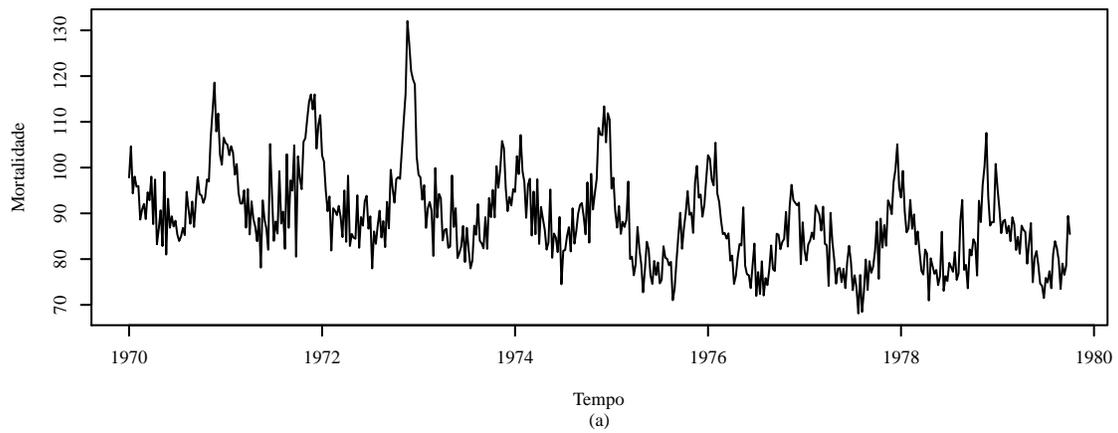


Figura 3.3: Mortalidade cardiovascular (a), Temperatura (b) e níveis de material particulado (c) ao longo do período de 10 anos (1970 - 1979).

A Tabela 3.5 fornece algumas medidas descritivas para cada uma das variáveis, que incluem: tamanho da amostra, valor mínimo, valor máximo, mediana, desvio-padrão (DP), coeficiente de variação (CV), coeficiente de assimetria (CA) e coeficiente de curtose (CC).

Tabela 3.5: Estatísticas descritivas para M_t , X_{1t} e X_{2t} .

n	Variáveis	Min.	Max.	Mediana	Média	DP	CV	CA	CC
508	M_t	68.11	132.04	87.33	88.699	9.999	11.273	0.804	0.981
	X_{1t}	50.91	99.88	74.055	74.260	9.014	12.138	0.095	-0.459
	X_{2t}	20.25	97.94	44.25	47.413	15.138	31.928	0.57	-0.474

A Figura 3.4 mostra uma matriz de dispersão e correlação entre as variáveis M_t , X_{1t} e X_{2t} . As variáveis e seus histogramas são mostrados ao longo da diagonal da matriz. Abaixo da diagonal é mostrado os diagramas de dispersão entre duas variáveis, e acima da diagonal é mostrado o correspondente coeficiente de correlação de Pearson (ρ).

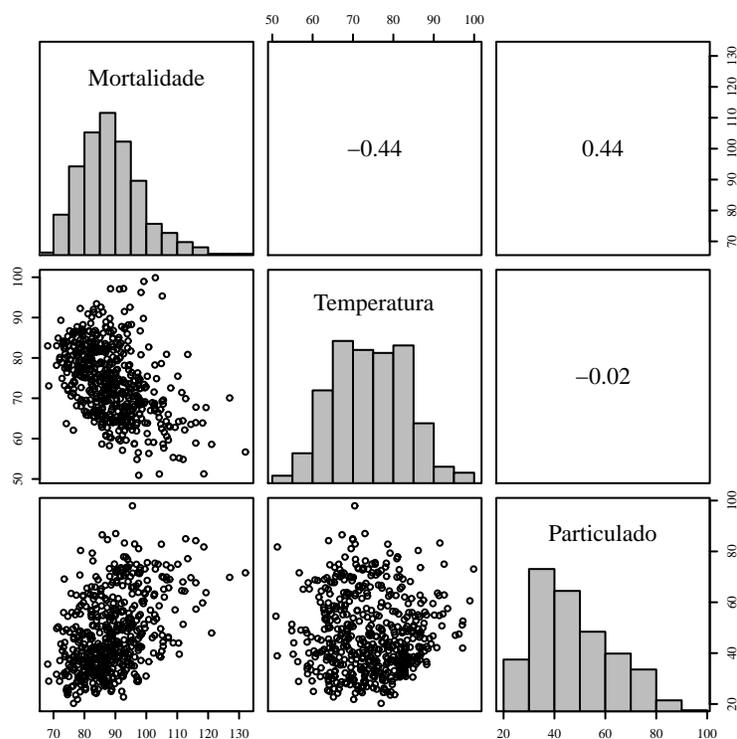


Figura 3.4: Matriz de dispersão, histogramas e coeficientes de correlação das variáveis mortalidade (M_t), temperatura (X_{1t}) e níveis de material particulado (X_{2t}). As variáveis e seus histogramas são mostrados ao longo da diagonal da matriz. Abaixo da diagonal mostra os gráficos de dispersão entre duas variáveis, e o correspondente coeficiente de correlação linear é mostrado acima da diagonal.

A análise da matriz e de seus diagramas, permitem identificar que as variáveis M_t e X_{2t} estão

moderadamente correlacionadas ($\rho = 0.44$). Além disso, também existe uma relação curvilínea entre as variáveis M_t e X_{1t} , indicando que as temperaturas mais altas e as temperaturas mais baixas estão moderadamente associadas ao aumento da mortalidade cardiovascular. No entanto, as variáveis X_{1t} e X_{2t} não possuem quaisquer correlação ($\rho = -0.02$). Os histogramas corroboram com os resultados dos CA's obtidos na Tabela 3.5 e, mostram que as variáveis M_t e X_{2t} possuem um comportamento assimétrico positivo. Também, esse comportamento é confirmado pelos Boxplots apresentados na Figura 3.5 (a) e (c) e além disso, nota-se ainda a presença de valores extremos ou observações discrepantes para as variáveis M_t e X_{2t} . Para a variável X_{1t} há indícios de assimetria fraca (quase simétrica).

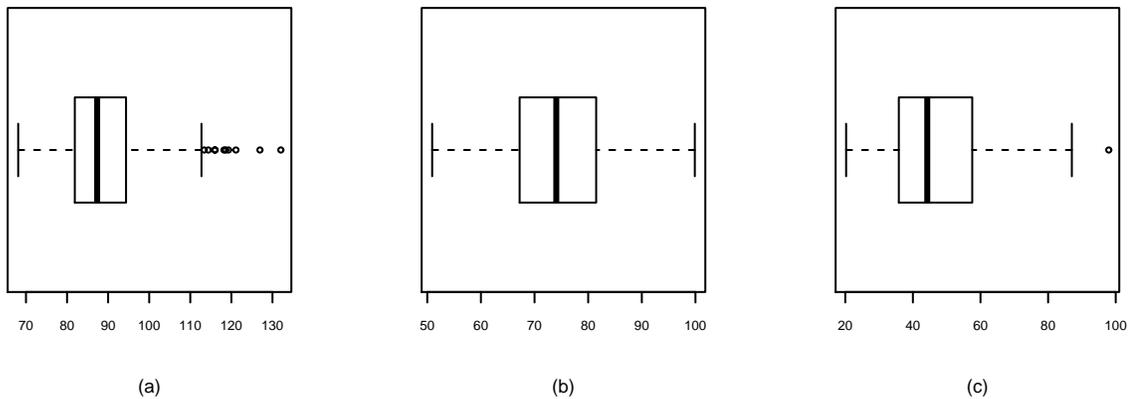


Figura 3.5: Boxplots para as variáveis M_t (a), X_{1t} (b) e X_{2t} (c).

Baseado na análise da Figura 3.4, que mostra a relação entre as variáveis M_t , X_{1t} e X_{2t} e, considerando M_t como a variável de resposta, essas relações podem ser modeladas ao longo do tempo t por meio dos valores observados de X_{1t} e X_{2t} , x_{1t} e x_{2t} como

$$M_t = \eta + \beta_1 \text{trend}_t + \beta_2(x_{1t} - \bar{x}_1) + \beta_3(x_{1t} - \bar{x}_1)^2 + \beta_4 x_{2t} + \varepsilon_t, \quad (3.19)$$

em que \bar{x}_1 é a temperatura média adicionada no modelo para evitar colinearidade, “trend_{*t*}” é uma tendência linear descendente em t observada na Figura 3.3 (a), e ε_t é um erro aleatório ou um processo de ruído consistindo de variáveis independentes e identicamente distribuídas com média zero e variância σ_ε^2 ; ver [Shumway e Stoffer \(2017\)](#).

A Figura 3.6 mostra os gráficos das funções de autocorrelação (ACF) (a) e de autocorrelação parcial (PACF) (b) dos resíduos do ajuste de mínimos quadrados da Equação (3.19). A análise do

gráfico da ACF e da PACF, sugere intuitivamente a característica de um modelo estacionário AR(p) de ordem $p = 2$ para os resíduos.

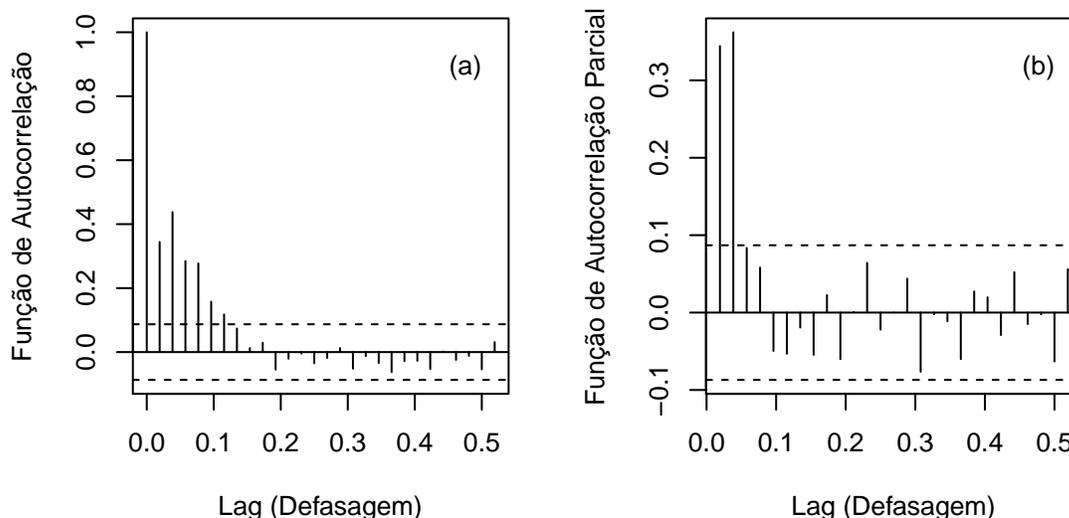


Figura 3.6: Representação gráfica das funções de (a) autocorrelação (ACF) e (b) autocorrelação parcial (PACF) dos resíduos do modelo de regressão para dados de séries temporais ao longo do período de 10 anos (1979-1979) no condado de Los Angeles.

Assim, o modelo do erro correlacionado definido em (3.19) é expresso da seguinte forma

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + u_t,$$

em que ε_t é um modelo AR(2) e u_t é um ruído branco. Os resultados para este modelo são obtidos usando a função `gammaFit` do pacote `gamlss.util`; veja <http://www.gamlss.org>.

Considere agora uma análise com o modelo RBSARMA definido por (3.7) e pela componente dada em (3.10). A Tabela 3.6 reporta os valores das estimativas de máxima verossimilhança, MAPE e AIC / BIC. A partir desta tabela, observe que o modelo RBSARMA (2,0) fornece um melhor ajuste do que o modelo ARMA (2,0) baseado nos valores AIC / BIC. Além disso, o modelo RBSARMA (2,0) apresenta menor valor de MAPE, indicando melhor capacidade de previsão.

Tabela 3.6: Estimativas e medidas de adequação de modelos para os dados da série de tempos de mortalidade no período de 10 anos (1979-1979) no Condado de Los Angeles.

Modelo	Parâmetros	Estimativa	AIC	BIC	MAPE
RBSARMA(2,0)	ϕ_1	0.3646			
	ϕ_2	0.4393			
	η	2842.8252			
	β_1	-1.3990	3078.4330	3112.2770	4.8128
	β_2	-0.0161			
	β_3	0.0154			
	β_4	0.1503			
	δ	623.5548			
ARMA(2,0)	ϕ_1	0.3881			
	ϕ_2	0.4321			
	η	2831.4911			
	β_1	-1.3932	3100.1290	3133.9730	4.8151
	β_2	-0.0169			
	β_3	0.0154			
	β_4	0.1554			

Os gráficos do quantil teórico *versus* quantil empírico (QQ) dos resíduos do GCS e QS, com envelopes de simulação, apresentados, respectivamente na Figura 3.7 (a) e (b), indicam melhor concordância com a distribuição do EXP (1) no modelo RBSARMA(2,0). No entanto, para uma mesma análise referente ao modelo ARMA(2,0), nota-se que o gráfico dos resíduos GCS e QS, com envelopes de simulação produzem alguns pontos de resíduos que situam-se distante da linha diagonal e fora do envelope; veja Figura 3.8 a,b. Nos gráficos ACF e PACF, observe que ambos os modelos produzem resíduos não autocorrelacionados; veja as Figuras 3.7 c,d e 3.8 c,d.

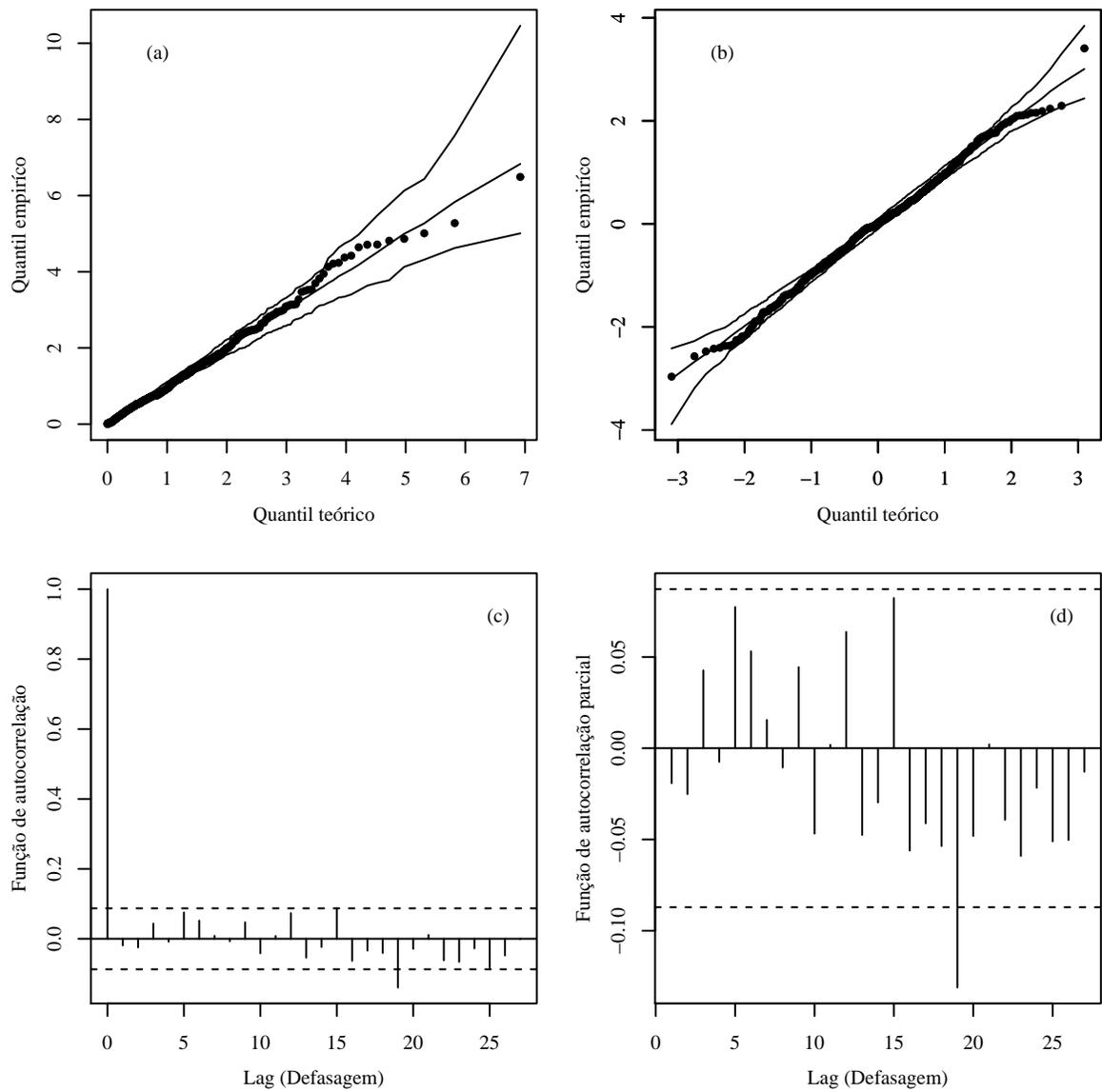


Figura 3.7: Gráfico dos envelopes dos resíduos de Cox-Snell generalizado (a) e resíduos quantil randomizados (b), funções de autocorrelação (c) e autocorrelação parcial (d) para o modelo RBSARMA(2,0).

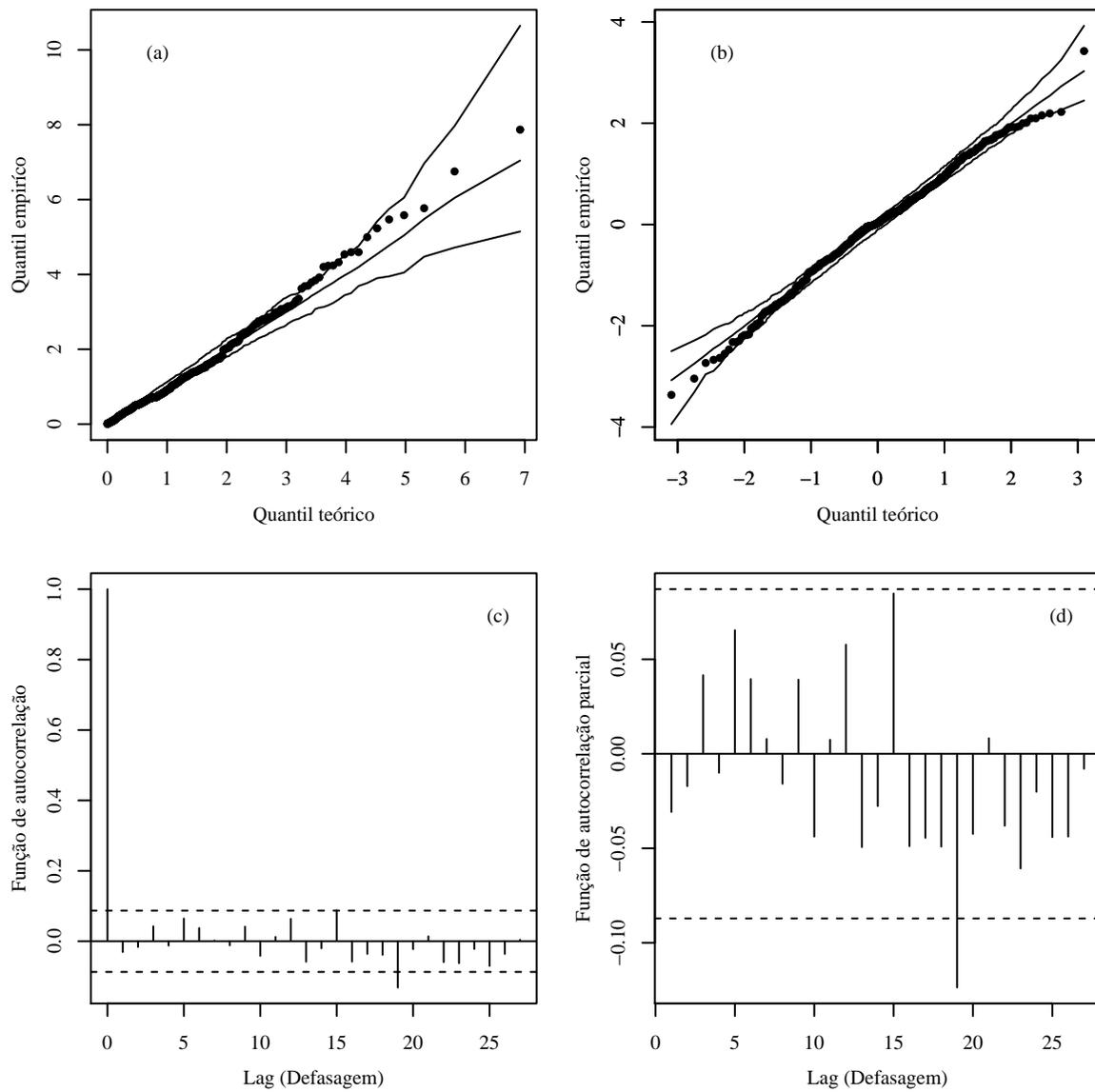


Figura 3.8: Gráfico dos envelopes dos resíduos de Cox-Snell generalizado (a) e resíduos quantil randomizados (b), funções de autocorrelação (c) e autocorrelação parcial (d) para o modelo ARMA(2,0).

As séries cronológicas ajustadas pelos modelos RBSARMA (2,0) e ARMA (2,0) são apresentadas em conjunto com as séries temporais observadas na Figura 3.9.

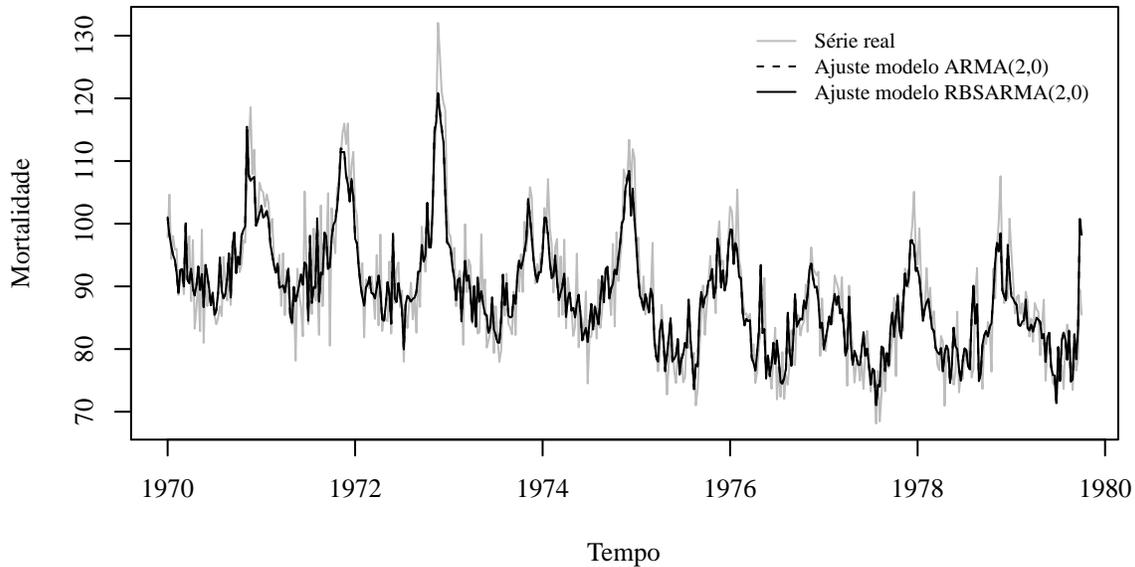


Figura 3.9: Série mortalidade cardiovascular no condado de Los Angeles (cinza), ajustadas conforme os modelos RBSARMA(2,0) (preto) e ARMA(2,0) (preto - - -) .

3.6 Considerações finais

Neste trabalho, foi estudado um novo modelo autorregressivo de média móvel baseado na distribuição Birnbaum-Saunders reparametrizada, denominado de RBSARMA. Estudos de simulação de Monte Carlo mostraram o bom desempenho dos estimadores de máxima verossimilhança. Na aplicação, o modelo RBSARMA(2,0) apresenta resultado superior ao clássico modelo ARMA(2,0) gaussiano, fornecendo fortes evidências de que a distribuição Birnbaum-Saunders é uma boa alternativa para lidar com dados temporais. Conseqüentemente, os resultados sugerem que o modelo RBSARMA pode se tornar uma valiosa ferramenta para a análise de dados de séries temporais positivos e assimétricos nas ciências ambientais e em outros campos do conhecimento.

Capítulo 4

Observações finais e trabalhos futuros

Apesar do crescente interesse no estudo da distribuição Birnbaum-Saunders e do desenvolvimento de muitos artigos de revisão, poucos trabalhos consideraram dados com uma estrutura de dependência. Para preencher essa lacuna, introduzimos duas novas classes de modelos de séries temporais baseada na distribuição Birnbaum-Saunders, que permite descrever dados positivos e assimétricos que possuem uma estrutura autorregressiva. Primeiro foi proposto o modelo autorregressivo de médias móveis Birnbaum-Saunders (BISARMA). Em seguida, foi estudado o modelo autorregressivo de médias móveis Birnbaum-Saunders reparametrizado (RBSARMA). Os resultados numéricos mostram o excelente desempenho de ambos os modelos, indicando que a distribuição Birnbaum-Saunders é uma boa alternativa de modelagem quando se trata de dados de séries temporais que são positivos e assimétricos e, portanto, pode ser uma adição valiosa ao conjunto de ferramentas estatísticas aplicadas. Em geral, o modelo BISARMA mostrou maior potencial em relação aos modelos RBSARMA e ARMA gaussiano.

Considerando o que foi exposto nesse trabalho, surgem perspectivas de desenvolvimento de algumas linhas de pesquisa. Por exemplo, modelos multivariados podem ser explorados, permitindo a inserção de uma estrutura de correlação. Ainda, metodologias para detecção de *outliers* podem ser implementadas. Finalmente, ferramentas de diagnóstico baseadas em influência local e global são de particular interesse.

Referências Bibliográficas

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19:716–723.
- Balakrishnan, N. e Zhu, X. (2015). Inference for the Birnbaum–Saunders Lifetime Regression Model with Applications. *Communications in Statistics-Simulation and Computation*, 44:2073–2100.
- Balakrishnan, N. e Kundu, D. (2018). Birnbaum-Saunders distribution: A review of models, analysis, and applications. *Applied Stochastic Models in Business and Industry*, page in press.
- Barros, M., Paula, G. A. e Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis*, 14:316–332.
- Barros, M., Paula, G. A. e Leiva, V. (2009). An R implementation for generalized Birnbaum-Saunders distributions. *Computational Statistics and Data Analysis*, 53:1511–1528.
- Benjamin, M. A., Rigby, R. A., e Stasinopoulos, D. M. (2003). Generalized autoregressive moving average models. *Journal of the American Statistical Association*, 98:214–223.
- Bhatti, C. (2010). The Birnbaum-Saunders autoregressive conditional duration model. *Mathematics and Computers in Simulation*, 80:2062–2078.
- Birnbaum, Z. W. e Saunders, S. C. (1969a). A new family of life distributions. *Journal of Applied Probability*, 6:319–327.
- Birnbaum, Z. W. e Saunders, S. C. (1969b). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability*, 6:328–347.
- Desmond, A. (1985). Stochastic models of failure in random environments. *Canadian Journal of Statistics*, 13:171-183.
- Desousa, M. F., Saulo, H., Leiva, V., e Scalco, P. (2018). On a tobit-Birnbaum-Saunders model with an application to antibody response to vaccine. *Journal of Applied Statistics*, 45:932–955.
- Díaz-García, J. A. e Leiva, V. (2005). A new family of life distributions based on elliptically contoured distributions. *Journal of Statistical Planning and Inference*, 128:445-457.
- Engle, R. e Russell, J. (1998). Autoregressive conditional duration: A new method for irregularly spaced transaction data. *Econometrica*, 66:1127–1162.
- Ferreira, M., Gomes, M. I., e Leiva, V. (2012). On an extreme value version of the Birnbaum-Saunders distribution. *REVSTAT Statistical Journal*, 10:181–210.

- Fonseca, R. V. e Cribari-Neto, F. (2018). Bimodal Birnbaum-Saunders generalized autoregressive score model. *Journal of Applied Statistics*, 45:2585–2606.
- Galea, M., Leiva, V. e Paula, G. A. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal of Applied Statistics*, 31:1049–1064.
- Garcia-Papani, F., Uribe-Opazo, M. A., Leiva, V. e Aykroyd, R. G. (2017). Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. *Stochastic Environmental Research and Risk Assessment*, 31:105–124.
- Garcia-Papani, F., Leiva, V., Ruggeri, F. e Uribe-Opazo, M. A. (2018a). Kriging with external drift in a Birnbaum-Saunders geostatistical model. *Stochastic Environmental Research and Risk Assessment*, 32:1517–1530.
- Garcia-Papani, F., Leiva, V., Uribe-Opazo, M. A. e Aykroyd, R. G. (2018b). Birnbaum-Saunders spatial regression models: Diagnostics and application to chemical data. *Chemometrics and Intelligent Laboratory Systems*, 177:114–128.
- Johnson, N. L., Kotz, S. e Balakrishnan, N. (1995). *Continuous Univariate Distributions*, volume 2. Wiley, New York, US.
- Lange, K. (2001). *Numerical Analysis for Statisticians*. Springer, New York, US.
- Leão, J., Leiva, V., Saulo, H. e Tomazella, V. (2017). Birnbaum-Saunders frailty regression models: Diagnostics and application to medical data. *Biometrical Journal*, 59:291–314.
- Leão, J., Leiva, V., Saulo, H., e Tomazella, V. (2018a). A survival model with Birnbaum-Saunders frailty for uncensored and censored cancer data. *Brazilian Journal of Probability and Statistics*, 32:707–729.
- Leão, J., Leiva, V., Saulo, H. e Tomazella, V. (2018b). Incorporation of frailties into a cure rate regression model and its diagnostics and application to melanoma data. *Statistics in Medicine*, page in press.
- Leiva, V., Barros, M., Paula, G. A. e Galea, M. (2007). Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Computational Statistics and Data Analysis*, 51:5694–5707.
- Leiva, V., Sanhueza, A. e Angulo, J. M. (2009). A length-biased version of the Birnbaum-Saunders distribution with application in water quality. *Stochastic Environmental Research and Risk Assessment*, 23:299–307.

- Leiva, V., Vilca, F., Balakrishnan, N. e Sanhueza, A. (2010). A skewed sinh-normal distribution and its properties and application to air pollution. *Communications in Statistics: Theory and Methods*, 39:426–443.
- Leiva, V., Ponce, M. G., Marchant, C. e Bustos, O. (2012). Fatigue Statistical Distributions Useful for Modeling Diameter and Mortality of Trees. *Revista Colombiana de Estadística*, 35:349–370.
- Leiva, V., Santos-Neto, M., Cysneiros, F. J. A. e Barros, M. (2014a). Birnbaum-Saunders statistical modelling: a new approach. *Statistical Modelling*, 14:21-48.
- Leiva, V., Saulo, H., Leão, J., e Marchant, C. (2014b). A family of autoregressive conditional duration models applied to financial data. *Computational Statistics and Data Analysis*, 79:175–191.
- Leiva, V., Marchant, C., Ruggeri, F. e Saulo, H. (2015). A criterion for environmental assessment using Birnbaum-Saunders attribute control charts. *Environmetrics*, 26:463–476.
- Leiva, V., Ferreira, M., Gomes, M. I. e Lillo, C. (2016). Extreme value Birnbaum-Saunders regression models applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, 30:1045–1058.
- Leiva, V. (2016). *The Birnbaum-Saunders Distribution*. Academic Press, New York, US.
- Leiva, V. e Saulo, H. (2017). Environmental applications based on Birnbaum–Saunders models. In: Adhikari, A., Adhikari, M. R. e Chaubey, Y. P. eds. *Mathematical and Statistical Applications in Life Sciences and Engineering*. New York, NY: Springer; 2017:283–304.
- Lillo, C., Leiva, V., Nicolis, O. e Aykroyd, R. (2018). L-moments of the Birnbaum-Saunders distribution and its extreme value version: Estimation, goodness of fit and application to earthquake data. *Journal of Applied Statistics*, 45:187–209.
- Mann, N. R., Schafer, R. E. e Singpurwalla, N. D. (1974). *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley and Sons, New York, US.
- Maior, V. Q. S. e Cysneiros, F. J. A. (2016). Symarma: A new dynamic model for temporal data on conditional symmetric distribution. *Statistical Papers*, 59:75–97.
- Marchant, C., Leiva, V., Cavieres, M. e Sanhueza, A. (2013). Air contaminant statistical distributions with application to PM10 in Santiago, Chile. *Reviews of Environmental Contamination and Toxicology*, 223:1–31.
- Marchant, C., Leiva, V. e Cysneiros, F. J. A. (2016a). A multivariate log-linear model for Birnbaum-Saunders distributions. *IEEE Transactions on Reliability*, 65:816–827.
- Marchant, C., Leiva, V., Cysneiros, F. J. A., e Vivanco, J. F. (2016b). Diagnostics in multivariate generalized Birnbaum-Saunders regression models. *Journal of Applied Statistics*, 43:2829–2849.

- Marchant, C., Leiva, V., Cysneiros, F. J. A., e Liu, S. (2018). Robust multivariate control charts based on Birnbaum-Saunders distributions. *Journal of Statistical Computation and Simulation*, 88:182–202.
- Nocedal, J. e Wright, S. (1999). *Numerical Optimization*. Springer, New York, US.
- Paula, G. A., Leiva, V., Barros, M. e Liu, S. (2012). Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry*, 28:16-34.
- Press, W. H., Teulosky, S. A., Vetterling, W. T., e Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing*. Prentice-Hall, London, UK.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rahul, T., Balakrishnan, N., e Balakrishna, N. (2018). Time series with Birnbaum-Saunders marginal distributions. *Applied Stochastic Models in Business and Industry*, 34:562–581.
- Rieck, J. R. e Nedelman, J. R. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics*, 3:51–60.
- Rocha, A. V. e Cribari-Neto, F. (2009). Beta autoregressive moving average models. *TEST*, 18:529–545.
- Santos-Neto, M., Cysneiros, F. J. A., Leiva, V. e Ahmed, S. E. (2012). On new parameterizations of the Birnbaum-Saunders distribution. *Pakistan Journal of Statistics*, 28:1–26.
- Santos-Neto, M., Cysneiros, F. J. A., Leiva, V. e Barros, M. (2014). On new parameterizations of the Birnbaum-Saunders distribution and its moments, estimation and application. *Revstat Statistical Journal*, 12:247–272.
- Santos-Neto, M., Cysneiros, F. J. A., Leiva, V. e Barros, M. (2016). Reparameterized Birnbaum-Saunders regression models with varying precision. *Electronic Journal of Statistics*, 2:2825–2855.
- Saulo, H., Leiva, V., Ziegelmann, F. A., e Marchant, C. (2013). A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, 27:1479–1491.
- Saulo, H., Leão, J., Leiva, V., e Aykroyd, R. G. (2018a). Birnbaum–Saunders autoregressive conditional duration models applied to high-frequency financial data. *Statistical Papers*, 1–25.
- Saulo, H., Leão, J. e Santos-Neto, M. (2018b). Discussion of “Birnbaum-Saunders distribution: A review of models, analysis, and applications” by N. Balakrishnan and Debasis Kundu. *Applied Stochastic Models in Business and Industry*, 1–4

- Shumway, R. H. e Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples*. Springer, New York, US.
- Schwarz, G. (1978). Estimation of the dimension of the model. *The Annals of Statistics*, 6:461–464.
- Stasinopoulos, D. e Rigby, R. (2007). Generalized additive models for location, scale and shape (GAMLSS). *Journal of Statistical Software*, 23:1-46.
- Vilca, F., Sanhueza, A., Leiva, V. e Christakos, G. (2010). An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile. *Stochastic Environmental Research and Risk Assessment*, 24:771–782.
- Xie, F. C. e Wei, B. C. (2007). Diagnostics analysis for log-Birnbaum-Saunders regression models. *Computational Statistics and Data Analysis*, 51:4692–4706.

Apêndice

Matriz informação observada de Fisher

Este apêndice contém expressões para derivadas parciais de segunda ordem da função log-verossimilhança do modelo BISARMA dadas em (2.8). As derivadas de primeira e segunda ordem da função de log-verossimilhança $\ell_t(\alpha, \beta, \eta, \phi, \theta)$ em relação a μ_t são expressas, respectivamente, por

$$\frac{\partial \ell_t}{\partial \mu_t} = \tanh\left(\frac{y_t - \mu_t}{2}\right) \left[\frac{2}{\alpha^2} \cosh^2\left(\frac{y_t - \mu_t}{2}\right) - \frac{1}{2} \right],$$

$$\frac{\partial^2 \ell_t}{\partial \mu_t^2} = \frac{\partial}{\partial \mu_t} \left(\frac{\partial \ell_t}{\partial \mu_t} \right) = \frac{1}{\alpha^2} \left\{ \frac{\alpha^2}{4} \operatorname{sech}^2\left(\frac{y_t - \mu_t}{2}\right) - 2 \operatorname{senh}^2\left(\frac{y_t - \mu_t}{2}\right) - 1 \right\}.$$

(a) Segunda derivada da função log-verossimilhança em relação a η

Note que

$$\begin{aligned} \mathbf{H}_{\eta\eta} &= \frac{\partial^2 \ell_t}{\partial \eta^2} = \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \\ &= \sum_{t=m+1}^n \frac{1}{\alpha^2} \left\{ \frac{\alpha^2}{4} \operatorname{sech}^2\left(\frac{y_t - \mu_t}{2}\right) - 2 \operatorname{senh}^2\left(\frac{y_t - \mu_t}{2}\right) - 1 \right\}. \end{aligned}$$

(b) Segunda derivada da função log-verossimilhança em relação a β

Para $\beta_{v,w}$ com $v, w = 1, \dots, k$, tem-se que

$$\begin{aligned} \mathbf{H}_{\beta\beta} &= \frac{\partial^2 \ell_t}{\partial \beta_v \partial \beta_w} = \sum_{t=m+1}^n \frac{\partial}{\partial \beta_v} \left(\frac{\partial \ell_t}{\partial \beta_w} \right) \\ &= \sum_{t=m+1}^n \frac{\partial}{\partial \beta_v} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \beta_w} \right) \\ &= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \beta_v} \frac{\partial \mu_t}{\partial \beta_w} + \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \beta_v \partial \beta_w} \\ &= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \left(x_{tv} - \sum_{i=1}^p \phi_i x_{(t-i)v} \right) \left(x_{tw} - \sum_{i=1}^p \phi_i x_{(t-i)w} \right), \end{aligned}$$

em que

$$\frac{\partial \mu_t}{\partial \beta_v} = \left(x_{tv} - \sum_{i=1}^p \phi_i x_{(t-i)v} \right), \quad \frac{\partial \mu_t}{\partial \beta_w} = \left(x_{tw} - \sum_{i=1}^p \phi_i x_{(t-i)w} \right)$$

e

$$\frac{\partial^2 \mu_t}{\partial \beta_v \partial \beta_w} = \frac{\partial}{\partial \beta_v} \left(\frac{\partial \mu_t}{\partial \beta_w} \right) = 0.$$

(c) Segunda derivada da função log-verossimilhança em relação a β_v e η

Para β_v com $v = 1, \dots, k$, tem-se que

$$\begin{aligned} \mathbf{H}_{\beta\eta} &= \frac{\partial^2 \ell_t}{\partial \beta_v \partial \eta} = \sum_{t=m+1}^n \frac{\partial}{\partial \eta} \left(\frac{\partial \ell_t}{\partial \beta_v} \right) \\ &= \sum_{t=m+1}^n \frac{\partial}{\partial \eta} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \beta_v} \right) \\ &= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \eta} \frac{\partial \mu_t}{\partial \beta_v} + \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \eta \partial \beta_v} \\ &= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \left(x_{tv} - \sum_{i=1}^p \phi_i x_{(t-i)v} \right), \end{aligned}$$

em que

$$\frac{\partial^2 \mu_t}{\partial \eta \partial \beta_v} = \frac{\partial}{\partial \eta} \left(\frac{\partial \mu_t}{\partial \beta_v} \right) = 0.$$

(d) Segunda derivada da função log-verossimilhança em relação a α

Note que

$$\frac{\partial \ell}{\partial \alpha} = \frac{1}{\alpha} \sum_{t=m+1}^n \left\{ \frac{4}{\alpha^2} \sinh^2 \left(\frac{y_t - \mu_t}{2} \right) - 1 \right\}.$$

Portanto,

$$\mathbf{H}_{\alpha\alpha} = \frac{\partial^2 \ell}{\partial \alpha^2} = -\frac{1}{\alpha^2} \sum_{t=m+1}^n \left\{ \frac{12}{\alpha^2} \sinh^2 \left(\frac{y_t - \mu_t}{2} \right) - 1 \right\}.$$

(e) Segunda derivada da função log-verossimilhança em relação a β_v e α

Para β_v com $v = 1, \dots, k$, tem-se que

$$\begin{aligned}
\mathbf{H}_{\alpha\beta_v} &= \frac{\partial^2 \ell}{\partial \alpha \partial \beta_v} = \sum_{t=m+1}^n \frac{\partial}{\partial \alpha} \left(\frac{\partial \ell_t}{\partial \beta_v} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \alpha} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \beta_v} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \alpha} \left\{ \tanh \left(\frac{y_t - \mu_t}{2} \right) \left[\frac{2}{\alpha^2} \cosh^2 \left(\frac{y_t - \mu_t}{2} \right) - \frac{1}{2} \right] \right\} \frac{\partial \mu_t}{\partial \beta_v} \\
&= \sum_{t=m+1}^n \left\{ -\frac{4}{\alpha^3} \cosh \left(\frac{y_t - \mu_t}{2} \right) \sinh \left(\frac{y_t - \mu_t}{2} \right) \right\} \frac{\partial \mu_t}{\partial \beta_v} \\
&= \sum_{t=m+1}^n -\frac{4}{\alpha^3} \cosh \left(\frac{y_t - \mu_t}{2} \right) \sinh \left(\frac{y_t - \mu_t}{2} \right) \left(x_{tv} - \sum_{i=1}^p \phi_i x_{(t-i)v} \right).
\end{aligned}$$

(f) Segunda derivada da função log-verossimilhança em relação a α e η

Note que

$$\frac{\partial \ell}{\partial \eta} = \sum_{t=m+1}^n \tanh \left(\frac{y_t - \mu_t}{2} \right) \left[\frac{2}{\alpha^2} \cosh^2 \left(\frac{y_t - \mu_t}{2} \right) - \frac{1}{2} \right];$$

Portanto,

$$\mathbf{H}_{\alpha\eta} = \frac{\partial^2 \ell}{\partial \alpha \partial \eta} = \frac{\partial}{\partial \alpha} \left(\frac{\partial \ell}{\partial \eta} \right) = \sum_{t=m+1}^n \frac{-4}{\alpha^3} \sinh \left(\frac{y_t - \mu_t}{2} \right) \cosh \left(\frac{y_t - \mu_t}{2} \right).$$

(g) Segunda derivada da função log-verossimilhança em relação a ϕ_i e ϕ_j

Para $\phi_i, \phi_j, i, j = 1, \dots, k$, tem-se que

$$\begin{aligned}
\mathbf{H}_{\phi_i\phi_j} &= \frac{\partial^2 \ell}{\partial \phi_i \partial \phi_j} = \sum_{t=m+1}^n \frac{\partial}{\partial \phi_i} \left(\frac{\partial \ell_t}{\partial \phi_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \phi_i} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \phi_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \phi_i} \frac{\partial \mu_t}{\partial \phi_j} + \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \phi_i \partial \phi_j} \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} (y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}) (y_{t-j} - \mathbf{x}_{t-j}^\top \boldsymbol{\beta}).
\end{aligned}$$

em que

$$\frac{\partial^2 \mu_t}{\partial \phi_i} = (y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}), \quad \frac{\partial^2 \mu_t}{\partial \phi_j} = (y_{t-j} - \mathbf{x}_{t-j}^\top \boldsymbol{\beta}) \quad \text{e} \quad \frac{\partial^2 \mu_t}{\partial \phi_i \partial \phi_j} = \frac{\partial}{\partial \phi_i} \left(\frac{\partial \mu_t}{\partial \phi_j} \right) = 0.$$

(h) Segunda derivada da função log-verossimilhança em relação a α e ϕ_i

Para $\phi_i, j = 1, \dots, k$, tem-se que

$$\begin{aligned} \mathbf{H}_{\alpha \phi_i} &= \frac{\partial^2 \ell}{\partial \alpha \partial \phi_i} = \sum_{t=m+1}^n \frac{\partial}{\partial \alpha} \left(\frac{\partial \ell_t}{\partial \phi_i} \right) \\ &= \sum_{t=m+1}^n \frac{\partial}{\partial \alpha} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \phi_i} \right) \\ &= \sum_{t=m+1}^n \frac{\partial}{\partial \alpha} \left\{ \tanh \left(\frac{y_t - \mu_t}{2} \right) \left[\frac{2}{\alpha^2} \cosh^2 \left(\frac{y_t - \mu_t}{2} \right) - \frac{1}{2} \right] \right\} \frac{\partial \mu_t}{\partial \phi_i} \\ &= \sum_{t=m+1}^n \left\{ -\frac{4}{\alpha^3} \cosh \left(\frac{y_t - \mu_t}{2} \right) \sinh \left(\frac{y_t - \mu_t}{2} \right) \right\} \frac{\partial \mu_t}{\partial \phi_i} \\ &= \sum_{t=m+1}^n -\frac{4}{\alpha^3} \cosh \left(\frac{y_t - \mu_t}{2} \right) \sinh \left(\frac{y_t - \mu_t}{2} \right) (y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}). \end{aligned}$$

(i) Segunda derivada da função log-verossimilhança em relação a β e ϕ_i

Para $\phi_i = 1, \dots, k$, tem-se que

$$\begin{aligned} \mathbf{H}_{\beta_v \phi_i} &= \frac{\partial^2 \ell}{\partial \beta_v \partial \phi_i} = \sum_{t=m+1}^n \frac{\partial}{\partial \beta_v} \left(\frac{\partial \ell_t}{\partial \phi_i} \right) \\ &= \sum_{t=m+1}^n \frac{\partial}{\partial \beta_v} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \phi_i} \right) \\ &= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \beta_v} \frac{\partial \mu_t}{\partial \phi_i} + \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \beta_v \partial \phi_i} \\ &= \sum_{t=m+1}^n \left\{ \frac{\partial^2 \ell_t}{\partial \mu_t^2} \left(x_{tv} - \sum_{i=1}^p \phi_i x_{(t-i)v} \right) (y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}) - \frac{\partial \ell_t}{\partial \mu_t} \mathbf{x}_{(t-i)v} \right\}, \end{aligned}$$

em que

$$\frac{\partial^2 \mu_t}{\partial \beta_v \partial \phi_i} = \frac{\partial}{\partial \beta_v} \left(\frac{\partial \mu_t}{\partial \phi_i} \right) = -x_{(t-i)v}.$$

(j) Segunda derivada da função log-verossimilhança em relação a ϕ_i e η Para ϕ_i com $i = 1, \dots, k$,

tem-se que

$$\begin{aligned}
\mathbf{H}_{\phi_i\eta} &= \frac{\partial^2 \ell_t}{\partial \phi_i \partial \eta} = \sum_{t=m+1}^n \frac{\partial}{\partial \eta} \left(\frac{\partial \ell_t}{\partial \phi_i} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \eta} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \phi_i} \right) \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \eta} \frac{\partial \mu_t}{\partial \phi_i} + \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \eta \partial \phi_i} \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} (y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}),
\end{aligned}$$

em que

$$\frac{\partial^2 \mu_t}{\partial \eta \partial \phi_i} = \frac{\partial}{\partial \eta} \left(\frac{\partial \mu_t}{\partial \phi_i} \right) = 0.$$

(k) Segunda derivada da função log-verossimilhança em relação a θ_i e θ_j

Para $\theta_i, \theta_j, i, j = 1, \dots, k$, tem-se que

$$\begin{aligned}
\mathbf{H}_{\theta_i\theta_j} &= \frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j} = \sum_{t=m+1}^n \frac{\partial}{\partial \theta_i} \left(\frac{\partial \ell_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \theta_i} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \theta_i} \frac{\partial \mu_t}{\partial \theta_j} + \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \theta_i \partial \theta_j} \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} u_{t-i} u_{t-j},
\end{aligned}$$

em que

$$\frac{\partial^2 \mu_t}{\partial \theta_i} = u_{t-i}, \quad \frac{\partial^2 \mu_t}{\partial \theta_j} = u_{t-j} \quad \text{e} \quad \frac{\partial^2 \mu_t}{\partial \theta_i \partial \theta_j} = \frac{\partial}{\partial \theta_i} \left(\frac{\partial \mu_t}{\partial \theta_j} \right) = 0.$$

(l) Segunda derivada da função log-verossimilhança em relação a ϕ_i e θ_j

Para $\phi_i, \theta_j, i, j = 1, \dots, k$, tem-se que

$$\begin{aligned}
\mathbf{H}_{\phi_i \theta_j} &= \frac{\partial^2 \ell}{\partial \phi_i \partial \theta_j} = \sum_{t=m+1}^n \frac{\partial}{\partial \phi_i} \left(\frac{\partial \ell_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \phi_i} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \phi_i} \frac{\partial \mu_t}{\partial \theta_j} + \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \phi_i \partial \theta_j} \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial^2 \mu_t} (y_{t-i} - \mathbf{x}_{t-i}^\top \boldsymbol{\beta}) u_{t-j},
\end{aligned}$$

em que

$$\frac{\partial^2 \mu_t}{\partial \phi_i \partial \theta_j} = \frac{\partial}{\partial \phi_i} \left(\frac{\partial \mu_t}{\partial \theta_j} \right) = 0.$$

(m) Segunda derivada da função log-verossimilhança em relação a α e θ_j

Para $\theta_j, j = 1, \dots, k$, tem-se que

$$\begin{aligned}
\mathbf{H}_{\alpha \theta_j} &= \frac{\partial^2 \ell}{\partial \alpha \partial \theta_j} = \sum_{t=m+1}^n \frac{\partial}{\partial \alpha} \left(\frac{\partial \ell_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \alpha} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \alpha} \left\{ \tanh \left(\frac{y_t - \mu_t}{2} \right) \left[\frac{2}{\alpha^2} \cosh^2 \left(\frac{y_t - \mu_t}{2} \right) - \frac{1}{2} \right] \right\} \frac{\partial \mu_t}{\partial \theta_j} \\
&= \sum_{t=m+1}^n \left\{ -\frac{4}{\alpha^3} \cosh \left(\frac{y_t - \mu_t}{2} \right) \sinh \left(\frac{y_t - \mu_t}{2} \right) \right\} \frac{\partial \mu_t}{\partial \theta_j} \\
&= \sum_{t=m+1}^n -\frac{4}{\alpha^3} \cosh \left(\frac{y_t - \mu_t}{2} \right) \sinh \left(\frac{y_t - \mu_t}{2} \right) u_{t-j}.
\end{aligned}$$

(n) Segunda derivada da função log-verossimilhança em relação a β e θ_j

Para $\theta_j = 1, \dots, k$, tem-se que

$$\begin{aligned}
\mathbf{H}_{\beta_v \theta_j} &= \frac{\partial^2 \ell}{\partial \beta_v \partial \theta_j} = \sum_{t=m+1}^n \frac{\partial}{\partial \beta_v} \left(\frac{\partial \ell_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \beta_v} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \beta_v} \frac{\partial \mu_t}{\partial \theta_j} + \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \beta_v \partial \theta_j} \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial^2 \mu_t} \left(x_{tv} - \sum_{i=1}^p \theta_j x_{(t-i)v} \right) u_{t-j}.
\end{aligned}$$

(o) Segunda derivada da função log-verossimilhança em relação a θ_j e η

Para θ_j com $i = 1, \dots, k$, tem-se que

$$\begin{aligned}
\mathbf{H}_{\theta_j \eta} &= \frac{\partial^2 \ell_t}{\partial \theta_j \partial \eta} = \sum_{t=m+1}^n \frac{\partial}{\partial \eta} \left(\frac{\partial \ell_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial}{\partial \eta} \left(\frac{\partial \ell_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \theta_j} \right) \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \eta} \frac{\partial \mu_t}{\partial \theta_j} + \frac{\partial \ell_t}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \eta \partial \theta_j} \\
&= \sum_{t=m+1}^n \frac{\partial^2 \ell_t}{\partial \mu_t^2} u_{t-j},
\end{aligned}$$

em que

$$\frac{\partial^2 \mu_t}{\partial \eta \partial \theta_j} = \frac{\partial}{\partial \eta} \left(\frac{\partial \mu_t}{\partial \theta_j} \right) = 0.$$