

Thiago Rezende dos Santos

***INFERÊNCIA SOBRE OS HIPERPARÂMETROS
DOS MODELOS ESTRUTURAIS SOB AS
PERSPECTIVAS CLÁSSICA E BAYESIANA***

Belo Horizonte - MG

Março / 2009

Thiago Rezende dos Santos

***INFERÊNCIA SOBRE OS HIPERPARÂMETROS
DOS MODELOS ESTRUTURAIS SOB AS
PERSPECTIVAS CLÁSSICA E BAYESIANA***

Dissertação de mestrado apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da UFMG, como parte dos requisitos necessários para obtenção do grau de Mestre em Ciências.

Orientadora:
Profa. Dra. Glaura C. Franco

Co-orientador:
Prof. Dr. Dani Gamerman

DEPARTAMENTO DE ESTATÍSTICA
INSTITUTO DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DE MINAS GERAIS

Belo Horizonte - MG

Março / 2009

Dissertação de Mestrado sob o título “*Inferência sobre os hiperparâmetros dos modelos estruturais sob as perspectivas Clássica e Bayesiana*”, defendida por Thiago Rezende dos Santos e aprovada em 09 de março de 2009, em Belo Horizonte, Minas Gerais, pela banca examinadora constituída pelos doutores:

Profa. Dra. Glaura C. Franco
Departamento de Estatística - UFMG
Orientadora

Prof. Dr. Dani Gamerman
Departamento de Métodos Estatísticos - UFRJ
Co-orientador

Prof. Dr. Pedro A. Morettin
IME - USP

Profa. Dra. Edna Afonso Reis
Departamento de Estatística - UFMG

Dedico esta dissertação ao Deus-vivo (Jesus), aos meus familiares, aos amigos visíveis e invisíveis e a todos que souberam aprender e sonhar.

AGRADECIMENTOS

Eu gostaria de agradecer, primeiramente, a Deus (sumo Deo) pela vida, aos meus amados, pais Hélio e Alzira, irmãos, Hugo e Daniela, Pr. Clébio, tios Alda e Roberto, primos, Krisna e Lucas, amigos, Alexandre, Francisco, Karlla pelo apóio, orações e amor essenciais para que eu continuasse a trilhar esse difícil e tortuoso caminho da academia. Aos amigos da UFMG, Gabriel, João Vitor, Markus, Fernandão, Fábio, Emerson, Fabiana, Nayara, Naila, Vinícius, Flávio e Roger, pelos impagáveis momentos de alegria e descontração e por me ajudarem "a estabilizar a variância" nos momentos difíceis. De maneira alguma, eu poderia esquecer dos meus queridos orientadores, Profa. Glaura e Prof. Dani, pois este trabalho não seria realizado sem a imensa contribuição deles. À Profa. Glaura, fica a minha incomensurável e eterna gratidão por me acompanhar e orientar desde da graduação e me ajudar nas minhas limitações. Ao Departamento de Estatística da UFMG, pela oportunidade de realizar este trabalho. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo suporte financeiro tão necessário, para que eu pudesse me dedicar exclusivamente a este estudo.

RESUMO

Esta dissertação dedica-se à discussão da metodologia dos modelos estruturais, também conhecidos como modelos dinâmicos, para a modelagem de séries temporais. Neste trabalho, o foco principal está voltado para a comparação dos estimadores clássicos e Bayesianos para se fazer inferências sobre os parâmetros do modelo. Para tanto, utilizam-se as técnicas computacionais *bootstrap*, enfoque clássico, e Markov chain Monte Carlo (MCMC), no enfoque bayesiano. Através de experimentos Monte Carlo, avaliam-se o vício e o erro quadrático médio dos estimadores de máxima verossimilhança e Bayesianos. Avaliando os estimadores pontuais, percebe-se que o estimador de máxima verossimilhança (EMV) e a moda *a posteriori* apresentam um melhor desempenho, em geral. Verifica-se também que as estimativas *bootstrap* acompanham bem o comportamento das estimativas de máxima verossimilhança nas replicações Monte Carlo dos modelos. Discute-se também a construção de intervalos de confiança assintóticos, *bootstrap* e intervalos de credibilidade para os parâmetros, comparando-os quanto à porcentagem de cobertura e à amplitude. Analisando os estimadores intervalares, nota-se que os intervalos de credibilidade possuem resultados melhores que os intervalos de confiança clássicos principalmente os assintóticos que, em alguns casos, apresentam limites que ultrapassam o limiar do espaço paramétrico, o que ocorre geralmente para séries de tamanho pequeno. Isso justifica também a construção de outros intervalos para os parâmetros do modelo. Este trabalho também apresenta uma extensão da metodologia para modelar séries que possuem quebras estruturais, utilizando funções de transferência. A parte final destina-se à apresentação dos modelos de resposta não-gaussiana e a introdução de uma nova família (mais ampla) de modelos de nível local não-gaussianos, assim como casos particulares dessa família. Como ilustração das metodologias apresentadas, são feitas aplicações a algumas séries reais. O ajuste dos modelos a essas séries foram satisfatórios, exemplificando bem as metodologias apresentadas neste estudo.

Palavras-chave: modelos estruturais, modelos dinâmicos, função de transferência, filtro de Kalman, *bootstrap*, MCMC, modelos de resposta não-gaussiana.

ABSTRACT

This dissertation is dedicated to the discussion of the methodology of structural models, also known as dynamic models, for modeling time series. In this work, the main objective is the comparison of classical and Bayesian estimators in order to make inferences about the parameters of the model. For this purpose, the computational techniques bootstrap, in the classical approach, and Markov chain Monte Carlo (MCMC), in the Bayesian approach, are used. Through Monte Carlo simulations, the bias and the mean square error of the maximum likelihood and Bayesian estimators are compared. Evaluating the point estimators, the maximum likelihood and the posterior mode estimators present in general better performance. It can be also checked that the bootstrap estimates mimics well the behavior of the maximum likelihood estimates in the Monte Carlo replications of models. Besides, asymptotic and bootstrap confidence intervals and credibility intervals for the hyperparameters are built, and they are compared with respect to the coverage rate and width. The credibility intervals show a better performance than the classical confidence intervals especially the asymptotic ones that, in some cases, present limits that go beyond the parametric space, mainly for small time series. This work also presents an extension of the methodology to model series that possesses structural breaks, using transfer functions. In the final part, a proposal of a new family of non-gaussian models is presented. As illustration of the methodologies, applications to some real time series are performed.

Keywords: structural models, dynamic models, transfer function, Kalman filter, bootstrap, MCMC, non-gaussian models.

SUMÁRIO

Lista de Figuras

Lista de Tabelas

1	Introdução	15
1.1	Motivação	17
1.2	Objetivos	17
1.3	Organização deste trabalho	18
2	Metodologia	20
2.1	Séries temporais	20
2.2	Modelos Estruturais (ME)	22
2.2.1	Definição	22
2.2.2	Alguns Modelos Estruturais Específicos	23
2.2.3	A Forma de Espaço de Estados	25
2.2.4	Filtro de Kalman	26
2.2.5	Suavização	28
2.3	Inferência Clássica	29
2.3.1	Estimador de máxima verossimilhança	29
2.3.2	Intervalos de confiança assintóticos	30
2.4	Inferência Bayesiana	31
2.4.1	Estimador de Bayes	32
2.4.2	Intervalo de credibilidade	32

2.5	<i>Bootstrap</i>	33
2.5.1	<i>Bootstrap</i> não-paramétrico nos resíduos	34
2.5.2	<i>Bootstrap</i> paramétrico nos resíduos	35
2.5.3	Intervalos de Confiança <i>Bootstrap</i>	35
2.5.4	Comparações entre os intervalos de confiança <i>bootstrap</i>	37
2.6	Métodos de Simulação MCMC	37
2.6.1	Algoritmo de Metropolis-Hastings	38
2.6.2	Algoritmo de Metropolis-Hastings híbrido	39
2.7	Previsão	41
3	Comparação de métodos clássicos e Bayesianos para inferência sobre os hiperparâmetros em modelos estruturais	43
3.1	Comparando os procedimentos de estimação	43
3.2	Intervalos de confiança e de credibilidade	45
3.3	Aplicação à série da precipitação de SO_4 em Nova York	48
4	Função de transferência	52
4.1	Análise de intervenção em modelos estruturais	53
4.1.1	Modelo 1	55
4.1.2	Modelo 2	56
4.1.3	Estimação	57
4.2	Resultados de Simulação	58
4.2.1	Modelo 1	59
4.2.2	Modelo 2	63
4.3	Aplicação à séries reais	65
4.3.1	Série IPCA	65
4.3.2	Série IBOVESPA	68

5	Modelos estruturais não-gaussianos	70
5.1	Família Gama de modelos dinâmicos	71
5.1.1	Definição	72
5.1.2	Procedimento de inferência	73
5.2	Casos particulares da família Gama de modelos dinâmicos	74
5.2.1	Modelo Poisson	75
5.2.2	Modelo Gama	76
5.2.3	Modelo Weibull	77
5.2.4	Modelo Pareto	78
5.2.5	Modelo Beta	79
5.2.6	Modelo Normal com média conhecida	80
5.3	Aplicações a modelos estruturais de Poisson	81
5.3.1	Modelo de Poisson	81
5.3.2	Aplicação à série temporal histórica do confronto entre Brasil e Argentina	82
6	Conclusões e trabalhos futuros	86
	Apêndice	90
	Referências bibliográficas	93

LISTA DE FIGURAS

3.1	Logaritmo Precipitação de SO4 em Nova York.	50
4.1	Efeito da intervenção das funções pulso e degrau.	54
4.2	Exemplos das funções pulso e degrau para o Modelo 1, com $n = 100$, $\beta = 4$ e $\rho = 0, 0.50$ e 0.99 . A linha vertical indica o instante da intervenção, $T = 50$, a linha contínua indica a série temporal, y_t , e linha tracejada indicada a resposta média α_t	56
4.3	Exemplos da função degrau para o Modelo 2, com $n = 100$ e $\rho = 0.00, 0.50$ e 0.99 . A linha vertical indica o instante da intervenção, $T = 50$, a linha contínua indica a série, y_t , e a linha tracejada indica a resposta média α_t	57
4.4	Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função pulso e $\rho = 0.99$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.	60
4.5	Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função pulso e $\rho = 0.50$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.	60
4.6	Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função pulso e $\rho = 0.00$. As linhas horizontal e vertical indicam o vício e a raiz quadrada do EQM, respectivamente.	61
4.7	Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função degrau e $\rho = 0.99$. As linhas horizontal e vertical indicam o vício e a raiz quadrada do EQM, respectivamente.	61
4.8	Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função degrau e $\rho = 0.50$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.	62

4.9	Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função degrau e $\rho = 0$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.	62
4.10	Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 2 com função degrau e $\rho = 0.99$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.	64
4.11	Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 2 com função degrau e $\rho = 0.50$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.	64
4.12	Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 2 com função degrau e $\rho = 0$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.	67
4.13	A linha contínua indica a série do IPCA, y_t , e as linhas pontilhada e tracejada indicam as estimativas da resposta média do Modelo 1, segundo as abordagens Bayesiana e clássica, respectivamente.	67
4.14	A linha contínua indica a série do retorno do IBOVESPA, y_t , e as linhas pontilhada e tracejada indicam as respostas médias do Modelo 2 Clássico e do Modelo 2 Bayesiano, respectivamente.	68
5.1	Série temporal simulada sob o modelo Poisson. A linha contínua representa a série temporal e a linha tracejada a média da distribuição preditiva do ajuste clássico.	83
5.2	Série temporal do número de gols marcados pelo Brasil no confronto com a Argentina. Os pontos representam a série temporal e as linhas tracejada e pontilhada representam as médias da distribuição preditiva dos ajustes clássico e Bayesiano, respectivamente.	84

LISTA DE TABELAS

3.1	Resultados da estimação no MNL segundo os métodos de Máxima verossimilhança, <i>bootstrap</i> e estimação Bayesiana.	44
3.2	Resultados da estimação no MTL segundo os métodos de Máxima verossimilhança, <i>bootstrap</i> e estimação Bayesiana.	45
3.3	Resultados da estimação no MEB segundo os métodos de Máxima verossimilhança, <i>bootstrap</i> e estimação Bayesiana.	46
3.4	Intervalos de confiança e de credibilidade de 95% para o MNL.	47
3.5	Intervalos de confiança e de credibilidade de 95% para o MTL.	48
3.6	Intervalos de confiança e de credibilidade de 95% para o MEB.	49
3.7	Estimativas de Máxima verossimilhança, <i>bootstrap</i> e estimação Bayesiana para os hiperparâmetros da série de SO_4	51
3.8	Intervalos de confiança e de credibilidade para os hiperparâmetros da série de SO_4	51
4.1	Intervalos de confiança e de credibilidade para o Modelo 1 com função pulso.	63
4.2	Intervalos de confiança e de credibilidade para o Modelo 1 com função degrau.	65
4.3	Intervalos de confiança e de credibilidade para o Modelo 2.	66
4.4	AIC e a previsão um passo à frente para a série IPCA.	66
4.5	Ajuste do Modelo 1 à série IPCA.	68
4.6	AIC e a previsão um passo à frente para a série IBOVESPA.	69
4.7	Ajuste do Modelo 2 à série IBOVESPA.	69
5.1	EMV e EB para o modelo Poisson.	82
5.2	Intervalos de confiança e de credibilidade para o modelo Poisson.	82

5.3	Ajuste do modelo Poisson à série do número de gols marcados pelo Brasil no confronto com a Argentina.	84
5.4	Distribuição de probabilidade preditiva do número de gols do Brasil na próxima partida contra a Argentina.	85

LISTA DE ABREVIATURAS E SÍMBOLOS

y_t - Série temporal univariada y_1, y_2, \dots

μ_t - Componente de tendência

γ_t - Componente sazonal

δ_t - Componente cíclico

ε_t - Componente aleatório ou erro

Ψ - Vetor de hiperparâmetros

ϕ - Vetor composto por hiperparâmetros e parâmetros

ME - Modelos Estruturais

MNL - Modelo de Nível Local

MTL - Modelo de Tendência Linear Local

MEB - Modelo Estrutural Básico

NID - Independente e normalmente distribuído

i.i.d. - Independente e identicamente distribuído

FK - Filtro de Kalman

MCMC - Markov chain Monte Carlo

N_i - Número de iterações do Algoritmo Metroplis-Hasting

n - Número de observações da série

T - Ponto ou instante da intervenção

MC - Número de replicações Monte Carlo

EMV - Estimador de máxima verossimilhança

EB - Estimador de Bayes

EQM - Erro quadrático médio

B - Número de replicações *bootstrap*

* - Símbolo indicativo dos resultados obtidos utilizando-se o *bootstrap*

1 INTRODUÇÃO

A metodologia de *modelos estruturais* (ME) ou *modelos dinâmicos*, é uma das várias abordagens para modelagem de séries temporais. Esses modelos decompõem uma série temporal em termos de seus componentes não-observáveis, como nível, tendência e sazonalidade, os quais têm uma interpretação direta (Harvey, 1989).

A idéia dessa decomposição da série temporal surgiu nos trabalhos de Holt (1957) e Winters (1960), que desenvolveram as técnicas de alisamento exponencial. Aproveitando essa idéia, surgiram alguns trabalhos formalizando a metodologia de modelos estruturais, dentre os quais citam-se os de Muth (1960), Theil & Wage (1964) e Nervole & Wage (1964). Entretanto, mediante a dificuldade computacional da época e o aparecimento dos modelos de Box & Jenkins (1976), essa metodologia foi um pouco esquecida. Apenas os autores Harrison & Stevens (1971, 1976) trabalharam com essa metodologia sob a perspectiva Bayesiana. Com a criação do *software STAMP (Structural Time Series Analyser, Modeller and Predictor)* no final da década de 1980, e com a importante colaboração de Andrew Harvey, os modelos estruturais tornaram-se mais populares.

Levando em conta essa decomposição da série, alguns modelos estruturais específicos podem ser obtidos, tais como o modelo de nível local (MNL), o modelo de tendência linear local (MTL) e o modelo estrutural básico (MEB), que são construídos baseados nas suposições de independência, homoscedasticidade e normalidade dos erros. Porém, em alguns casos, essas suposições não são atendidas devido a algum evento e/ou fator externo (que influencia os valores da série em questão) ou pelo fato da própria série temporal não ter resposta gaussiana.

Para os casos nos quais as suposições do modelo são violadas em virtude de algum evento externo, os modelos de *função de transferência* (FT) são introduzidos, a fim de que se possa mensurar o impacto desse evento na série. Estes foram introduzidos por Box & Jenkins (1976) no contexto dos modelos ARIMA. Tendo em vista os modelos de FT de Box & Jenkins (1976), a introdução da FT no modelo estrutural ou no modelo dinâmico é quase natural (Harvey, 1989; West & Harrison, 1997), uma vez que os modelos ARIMA podem ser escritos na forma de

modelos estruturais.

Já para os casos nos quais as suposições de normalidade e/ou homoscedasticidade são violadas, uma solução natural é tentar uma transformação na série temporal. Porém, nem sempre uma transformação é suficiente para que a suposição de normalidade seja satisfeita. Logo, outra possibilidade é construir um modelo estrutural com resposta não-gaussiana (Smith & Miller, 1986), o que inclusive traz um ganho de interpretação do modelo (West, Harrison & Migon, 1985; West & Harrison, 1997). Uma estrutura mais geral, denominada *modelos lineares dinâmicos generalizados* (MLDG), foi proposta por West, Harrison & Migon (1985), despertando um grande interesse devido à sua grande aplicabilidade dos mesmos em diversas áreas do conhecimento.

Em todos os modelos citados anteriormente, é necessário fazer inferências sobre os parâmetros e hiperparâmetros. Neste estudo, são consideradas no processo de inferência a abordagem clássica, utilizando o estimador de máxima verossimilhança (EMV), e a abordagem Bayesiana, utilizando métodos *Markov chain Monte Carlo* (MCMC) para obter as estimativas Bayesianas (EB).

Na perspectiva clássica, a estimação dos hiperparâmetros é feita maximizando-se a função de verossimilhança que não tem uma "forma simples" por ser uma função não-linear do vetor hiperparâmetros. Por isso, utilizam-se métodos numéricos para maximizá-la.

Sob a perspectiva Bayesiana, a estimação dos hiperparâmetros pode não ser "tão" trivial, pois, na maioria dos casos, a distribuição *a posteriori* dos hiperparâmetros não é analiticamente tratável. Mediante essa dificuldade, os métodos MCMC provocaram uma revolução na inferência Bayesiana, porque são capazes de gerar uma amostra da distribuição *a posteriori* de um parâmetro de interesse (Gelman & Lopes, 2006; Gilks, Richardson & Spiegelhalter, 1996).

Como há interesse também em se obter estimativas intervalares para os hiperparâmetros, intervalos de confiança assintóticos e intervalos de credibilidade podem ser construídos.

O *bootstrap* (Efron, 1979) é um método de reamostragem de dados muito utilizado na inferência estatística quando se objetiva a construção de intervalos de confiança, realização de testes de hipóteses, estimação de vícios, previsão, seleção de modelos. Stoffer & Wall (1991) propuseram um procedimento que possibilita aplicar a técnica de *bootstrap* em modelos estruturais. Portanto, neste trabalho, esse é o procedimento usado para construir intervalos de confiança *bootstrap* para os hiperparâmetros.

1.1 MOTIVAÇÃO

A motivação principal desta dissertação está voltada para a comparação dos estimadores clássicos e Bayesianos e a aplicação dos métodos computacionais *bootstrap* e MCMC em modelos estruturais, para se fazer inferências sobre os hiperparâmetros, bem como estudar alguns modelos estruturais de função de transferência e modelos de resposta não-gaussiana. Existem estudos que aplicam o *bootstrap* em ME sob a abordagem clássica e também há alguns trabalhos que aplicam os métodos MCMC em ME sob a abordagem Bayesiana, mas não se tem conhecimento de trabalhos comparando as duas abordagens. Além disso, nos modelos de FT, não há estudos no enfoque clássico, de conhecimento dos autores, que consideram o parâmetro de intervenção variando no tempo, segundo por alguma lei estocástica.

1.2 OBJETIVOS

Este trabalho foi realizado com o propósito de atingir os seguintes objetivos:

1. Avaliar o vício e o Erro Quadrático Médio (EQM) dos estimadores de máxima verossimilhança e dos estimadores de Bayes, com distribuições *a priori* uniforme e de Jeffreys (não-conjugadas) em modelos estruturais específicos, tais como MNL, MTL e MEB;
2. Construir intervalos de confiança assintóticos e *bootstrap* (percentílico, com correção de vício e com correção de vício e aceleração) e intervalos de credibilidade para os hiperparâmetros. Comparar os intervalos quanto à porcentagem de cobertura e à amplitude em modelos estruturais específicos tais como MNL, MTL e MEB;
3. Apresentar os modelos estruturais de FT com os parâmetros das variáveis de intervenção tanto fixos quanto variando no tempo. Comparar os estimadores pontuais e intervalares dos parâmetros, utilizando as técnicas clássicas e Bayesianas. Além disso, comparar os intervalos de confiança *bootstrap* e de credibilidade para a previsão um passo à frente quanto à taxa de cobertura;
4. Propor os modelos estruturais de resposta não-gaussiana e a família gama de modelos dinâmicos. Apresentar alguns estudos de simulação comparando os estimadores de máxima verossimilhança e os estimadores Bayesianos nestes modelos;
5. Aplicar as metodologias discutidas a séries reais.

1.3 ORGANIZAÇÃO DESTE TRABALHO

Este trabalho já gerou uma publicação, dois artigos submetidos e um último artigo em fase de preparação. Os assuntos abordados em cada artigo são descritos abaixo.

- No primeiro artigo (Franco, Santos, Ribeiro & Cruz, 2008), intitulado "Confidence intervals in structural models", publicado na revista *Communications in Statistics*, foram implementados os intervalos de confiança assintótico e *bootstrap* percentílico não-paramétrico, os quais foram comparados quanto à amplitude e à porcentagem de cobertura. O EQM e o vício dos estimadores de máxima verossimilhança e *bootstrap* também foram avaliados. Além disso, uma aplicação da metodologia à série real do IPCA de Belo Horizonte foi discutida. Esse artigo resultou da monografia de conclusão de curso de graduação deste autor, que serviu de base para esta dissertação. Todos os programas computacionais foram desenvolvidos no *software* Ox (Doornik, 1999).
- O segundo artigo (Santos & Franco, 2008), sob o título "Inference for the hyperparameters of structural models under classical and Bayesian perspectives", submetido à revista *Computational Statistics*, é uma extensão do primeiro artigo, incluindo a abordagem Bayesiana e os intervalos de confiança *bootstrap* BC e BCa para os hiperparâmetros. Ressalta-se que as estimativas Bayesianas são obtidas usando distribuições *a priori* de referência para os hiperparâmetros. Adicionalmente, uma aplicação da metodologia à série real da precipitação de sulfato em Nova York é feita. Os resultados deste artigo serão mostrados no Capítulo 3.
- No terceiro artigo (Santos, Franco & Gamerman, 2009), intitulado "Comparison of Classical and Bayesian approaches for intervention analysis in structural models", submetido à revista *International Statistical Review*, são abordados os modelos estruturais de função de transferência. Os métodos computacionais MCMC e *bootstrap* são comparados quanto a sua eficiência para se fazer inferências sobre os parâmetros dos modelos. Além disso, aplicações da metodologia a duas séries reais são feitas. Os resultados deste artigo serão apresentados no Capítulo 4.
- No quarto artigo, que está em fase de elaboração, os ME de resposta não-gaussiana são abordados, assim como uma família gama de modelos dinâmicos. São percorridos alguns casos particulares dessa família e alguns resultados de simulação são obtidos para o modelo Poisson, comparando os métodos clássicos e Bayesianos para se fazer inferência sobre os parâmetros do modelo. Os resultados deste artigo serão mostrados no Capítulo 5.

Assim, para melhor compreensão da teoria subjacente a todos os trabalhos e para a exposição dos resultados obtidos em cada artigo, este texto foi elaborado da seguinte forma. No Capítulo 2, a metodologia dos modelos estruturais, bem como os métodos computacionais utilizados, são discutidos. No Capítulo 3, apresenta-se a comparação de métodos clássicos e Bayesianos em ME para se fazer inferência sobre os parâmetros. No Capítulo 4, a metodologia de função de transferência e alguns estudos de simulação são mostrados. No Capítulo 5, a família gama de modelos dinâmicos e alguns resultados de simulação são apresentados. Finalmente, no Capítulo 6 são apresentadas as conclusões e as propostas de trabalhos futuros.

2 METODOLOGIA

Os modelos estruturais (como são conhecidos na inferência clássica) ou modelos dinâmicos (como são conhecidos na inferência Bayesiana) são utilizados para modelar séries temporais a fim de explicá-los e/ou fazer previsões para as mesmas. A flexibilidade dessa modelagem, incluindo a capacidade de lidar com dados multivariados e processos não-estacionários, confere-lhes uma significativa vantagem frente às demais metodologias de análise de séries temporais (Spall, 1988). Este capítulo é destinado à parte metodológica desta dissertação, isto é, à formalização dos métodos bem como suas descrições. Para um melhor entendimento da metodologia, primeiramente a definição de séries temporais é introduzida. A seguir, são apresentados os modelos estruturais com a forma de espaço de estados, o filtro de Kalman e alguns modelos estruturais específicos. Os procedimentos de inferência clássica e Bayesiana também são mostrados, assim como os métodos computacionais *bootstrap* e MCMC.

2.1 SÉRIES TEMPORAIS

Segundo Brockwell & Davis (1996), uma série temporal pode ser definida informalmente como uma coleção de observações ordenadas no tempo. O intervalo de observação é considerado igualmente espaçado neste trabalho. Existem vários exemplos de séries temporais em diversas áreas como as séries de índices de preços ou de mercado de ações (na Economia), do índice de precipitação atmosférica diária (na Meteorologia), de batimentos cardíacos (na Medicina) etc.

Um aspecto relevante e interessante é que, em séries temporais, a amostra obtida de um determinado fenômeno possui uma estrutura de dependência entre suas observações, ou seja, a ordem em que as observações foram obtidas não pode ser descartada. Existe uma quantidade muito grande de dados com essa característica nas mais diversas áreas do conhecimento como Demografia, Economia, Engenharia, Biologia, Geofísica, Meio-Ambiente, entre outras.

Em cada instante de tempo a série temporal pode assumir diversos valores, porém, na

prática, apenas um único valor é observado. Isto é, uma série temporal é apenas uma das várias *trajetórias* que poderiam ser observadas ao longo do tempo. O conjunto de todas as trajetórias possíveis é denominado um *processo estocástico*.

Formalmente, um *processo estocástico* é uma família $y = \{y(t), t \in N\}$, tal que, para cada $t \in N$ (conjunto dos números inteiros), $y(t)$ é uma variável aleatória (Morettin & Tolo, 2004). Desta forma, um processo estocástico é uma coleção de variáveis aleatórias, definidas em um mesmo espaço de probabilidade (Ω, A, P) , formado pela terna espaço amostral (Ω) , a σ -álgebra (A) contida nas partes de Ω e a probabilidade (P) definida em A . O processo estocástico estará especificado quando a distribuição conjunta das variáveis aleatórias for definida.

A classe de processos estocásticos denominada *processos estacionários* - que podem ser fortemente ou fracamente estacionários - é de grande interesse na modelagem de séries temporais, pois vários modelos a utilizam como uma suposição simplificadora.

Definição 1. Um processo estocástico $y = \{y(t), t \in N\}$ é denominado *fracamente estacionário* se, e somente se,

$$E(y(t)) = \mu(t) = \mu, \text{ constante para todo } t \in N;$$

$$E(y^2(t)) < \infty \text{ para todo } t \in N;$$

$\gamma(t_1, t_2) = Cov(y(t_1), y(t_2))$ é uma função apenas de $|t_1 - t_2|$. Mais detalhes podem ser vistos em Morettin & Tolo (2004).

Fixado um $\omega \in \Omega$, uma função de t é obtida, isto é, uma *série temporal* (uma *trajetória do processo* ou uma *realização*) e a notação y_t é uma simplificação de $y(\omega, t)$. Uma série temporal não precisa necessariamente ser univariada e unidimensional, ela pode ser *multivariada* (um vetor aleatório \mathbf{y}_t) e/ou *multidimensional* (\mathbf{t} é um vetor).

Entre os principais objetivos de estudar séries, podem-se citar:

1. Análise - descrever a série, verificando suas principais características e possíveis relações com outras séries;
2. Geração - Investigar o mecanismo pelo qual as observações foram geradas;
3. Previsão - fazer previsões para os valores futuros, baseadas em toda informação disponível.

Harvey (1989) fornece duas razões para a modelagem de séries temporais que são prover uma descrição da série em termos de seus componentes de interesse e projetar valores futuros.

A análise e modelagem de séries temporais podem ser feitas em duas vertentes: no domínio do tempo e no domínio da frequência. Várias metodologias foram desenvolvidas no domínio do tempo, utilizando modelos paramétricos e visando a modelagem de séries temporais. Pode-se citar, dentre outras, os modelos ARIMA (Box & Jenkins, 1976), os modelos de longa dependência (ARFIMA) (Granger & Joyeux, 1980), os modelos não-lineares ARCH (Engle, 1982) e os modelos estruturais - que são o tema desta dissertação e serão apresentados na próxima seção.

Para obter uma melhor explanação sobre a formulação, definição e conceitos fundamentais em séries temporais, o leitor é fortemente indicado a ler os textos de Wei (1990), Box, Jenkins & Reinsel (1994), Brockwell & Davis (1996) e Morettin & Toloi (2004), este último em português.

2.2 MODELOS ESTRUTURAIS (ME)

Nesta seção, primeiramente, a definição e a decomposição dos ME em termos dos seus componentes são apresentadas. Para facilitar a estimação e reduzir o número de equações desses modelos, a seguir eles são escritos na forma de espaço de estados. Uma vez que os ME estão escritos nessa forma, eles podem ser aplicados ao algoritmo recursivo do filtro de Kalman, obtendo, dessa forma, o erro de previsão do FK e sua variância, que são utilizados diretamente na função de verossimilhança. De posse desta, pode-se fazer inferência clássica e Bayesiana (definindo a distribuição *a priori*) sobre os parâmetros do modelo.

2.2.1 DEFINIÇÃO

Uma série temporal pode ser decomposta em componentes não-observáveis. Um dos procedimentos disponíveis para realizar essa modelagem é conhecido como modelos estruturais. Os componentes não-observáveis que ocorrem com maior frequência em uma série temporal (Harvey, 1989; West & Harrison, 1997) são:

1. Tendência (μ_t): refere-se à direção geral segundo a qual o gráfico da série se desenvolve em um longo intervalo de tempo.
2. Componente Sazonal (γ_t): refere-se a padrões semelhantes que uma série temporal parece obedecer durante os mesmos meses de anos sucessivos, resultantes de eventos periódicos que ocorrem anualmente.
3. Componente cíclica (δ_t): refere-se às oscilações em longo prazo ou aos desvios em torno da reta ou da curva de tendência.

4. Componente aleatória ou erro (ε_t): referem-se aos deslocamentos esporádicos das séries temporais, provocados por eventos casuais.

Portanto, através dos modelos estruturais, pode-se escrever uma série temporal univariada $y_t, t = 1, 2, \dots, n$, da seguinte forma:

$$y_t = \mu_t + \gamma_t + \delta_t + \varepsilon_t,$$

sendo $\varepsilon_t \sim [0, \sigma_\varepsilon^2]$, independentes entre si.

2.2.2 ALGUNS MODELOS ESTRUTURAIS ESPECÍFICOS

Neste estudo, abordam-se os modelos de nível local (MNL), tendência linear local (MTL), de tendência polinomial (MTP) e estrutural básico (MEB). O modelo estrutural com o componente de ciclo foge ao escopo deste estudo. Mais detalhes sobre esses modelos podem ser encontrados em West & Harrison (1997) e Harvey (1989).

MODELOS DE NÍVEL LOCAL (MNL)

O modelo de nível local ou modelo linear dinâmico (MLD) de primeira ordem é o mais simples, porque possui apenas o componente de nível (μ_t) e do erro (ε_t). É também denominado como um passeio aleatório acrescido de um erro. A série se movimenta ao longo do tempo sem uma trajetória fixa, pois o nível atual é igual ao nível no período anterior mais um ruído branco. A forma do modelo é dada por:

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad (2.1)$$

$$\mu_t = \mu_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2), \quad (2.2)$$

com ε_t e η_t não-correlacionados e $t = 1, \dots, n$.

MODELO DE TENDÊNCIA LINEAR LOCAL (MTL)

Este modelo também é conhecido como *modelo linear dinâmico de segunda ordem*. A tendência em uma série pode ser observada quando existe um aumento ou decréscimo gradual

do nível das observações ao longo do tempo. Sabendo-se que $t = 1, \dots, n$, tem-se:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim N(0, \sigma_\varepsilon^2) \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t, & \eta_t &\sim N(0, \sigma_\eta^2) \\ \beta_t &= \beta_{t-1} + \xi_t, & \xi_t &\sim N(0, \sigma_\xi^2), \end{aligned}$$

onde ε_t , η_t e ξ_t são erros aleatórios mutuamente não-correlacionados.

MODELO DE TENDÊNCIA POLINOMIAL (MTP)

O modelo de tendência polinomial é um modelo de tendência mais geral que os outros dois modelos apresentados anteriormente. O *modelo de tendência polinomial de ordem k* pode ser definido como se segue:

$$\begin{aligned} y_t &= \alpha_{t,1} + \varepsilon_t, & \varepsilon_t &\sim N(0, \sigma_\varepsilon^2) \\ \alpha_{t,j} &= \alpha_{t-1,j} + \alpha_{t,j+1} + \eta_{t,j}, & (j &= 1, \dots, k-1), \\ \alpha_{t,k} &= \alpha_{t-1,k} + \eta_{t,k}, \end{aligned}$$

onde $\eta_t \sim N(\mathbf{0}, \mathbf{Q}_t)$ e \mathbf{Q}_t é uma matriz de covariâncias diagonal para $t = 1, \dots, n$.

Se $k = 1$ e $\alpha_t = \mu_t$, o MNL é obtido. O MTL também é um caso particular desse modelo quando $k = 2$ e $\alpha_t = (\mu_t, \beta_t)'$. Já o *modelo linear dinâmico de tendências quadráticas* pode ser obtido fazendo $k = 3$ e $\alpha_t = (\mu_t, \beta_t, \gamma_t)'$.

MODELO ESTRUTURAL BÁSICO (MEB)

O componente sazonal aparece quando a série apresenta algum tipo de periodicidade que ocorre regularmente de ano para ano. Portanto, as séries sazonais são caracterizadas por apresentarem correlações altas em *lags* sazonais. *Lags* são múltiplos de um período, que se denotará por s , e que pode ser semanal, mensal, trimestral etc.

Há duas maneiras de se "tratar" a sazonalidade. A primeira é modelar a sazonalidade por fatores e, a segunda, por componentes harmônicas ou trigonométricas. A segunda não é discutida neste estudo; entretanto o leitor interessado pode consultar os trabalhos de Harvey (1989) e West & Harrison (1997).

O Modelo Estrutural Básico é o modelo de tendência linear local acrescido do componente sazonal (γ_t) e está definido na forma:

$$\begin{aligned}
y_t &= \mu_t + \gamma_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2), \\
\mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2), \\
\beta_t &= \beta_{t-1} + \xi_t, \quad \xi_t \sim N(0, \sigma_\xi^2), \\
\gamma_t &= -\gamma_{t-1} - \dots - \gamma_{t-s+1} + \omega_t, \quad \omega_t \sim N(0, \sigma_\omega^2),
\end{aligned}$$

assumindo que $t = 1, \dots, n$ e que s refere-se ao número de períodos sazonais. ε_t , η_t , ξ_t e ω_t são distúrbios tipo ruído branco mutuamente não-correlacionados.

Para modelar um componente sazonal por fatores é necessário a restrição de que a soma dos componentes sazonais seja zero, isto é, $\sum_{j=0}^{s-1} \gamma_{t-j} = 0$. Obtém-se uma modelagem estocástica para o componente sazonal fazendo $\sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t$ sendo $\omega_t \sim N(0, \sigma_\omega^2)$.

2.2.3 A FORMA DE ESPAÇO DE ESTADOS

A representação de espaço de estados de um sistema é um conceito fundamental na teoria do controle moderna (Wei, 1990). Modelos de espaço de estados foram originalmente desenvolvidos por engenheiros de controle, particularmente para aplicação em sistemas de navegação, como, por exemplo, no monitoramento da posição de um foguete espacial. Somente mais tarde verificou-se que esses modelos eram também valiosos na análise de muitos outros tipos de séries temporais. De fato, a modelagem em espaço de estados é muito flexível, aplicável em uma grande variedade de problemas.

O objetivo de se escrever os modelos estruturais na forma de espaço de estados é reduzir o número de equações para apenas duas, facilitando, assim, o processo de estimação de tais modelos. Os parâmetros a serem estimados podem ser as variâncias dos erros associados a cada componente não-observável, denominados hiperparâmetros, ou coeficientes relacionados a variáveis e a componentes não-observáveis incluídas no modelo.

Como dito anteriormente, na forma de espaço de estados, é possível representar os modelos estruturais através de duas equações - a equação das observações (ou da medida) e a equação do estado (ou de transição), dadas respectivamente por (Harvey, 1989; West & Harrison, 1997):

$$y_t = \mathbf{z}_t' \boldsymbol{\alpha}_t + d_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, h_t) \quad (2.3)$$

$$\boldsymbol{\alpha}_t = \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \mathbf{R}_t \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(0, \mathbf{Q}_t) \quad (2.4)$$

onde $t = 1, 2, \dots, n$, ε_t são ruídos não correlacionados, $\boldsymbol{\eta}_t$ é um vetor de ruídos serialmente não-correlacionados, cuja matriz de covariâncias é dada por \mathbf{Q}_t e $\boldsymbol{\alpha}_t$ é o vetor de estados. Assume-se

também que ε_t e $\boldsymbol{\eta}_t$ sejam independentes entre si para $t = 1, 2, \dots, n$. As matrizes \mathbf{z}'_t , \mathbf{T}_t e \mathbf{R}_t são conhecidas como matrizes do sistema e assumidas determinísticas e conhecidas. Os termos d_t e \mathbf{c}_t são covariáveis que podem ser inseridas nas equações das observações e de transição, respectivamente. O modelo de espaço de estados é dito invariante no tempo quando as matrizes de sistema não mudam ao longo do tempo. O modelo tem as seguintes suposições $E(\varepsilon_t \boldsymbol{\eta}_j) = \mathbf{0}$ e $E(\boldsymbol{\eta}_t \boldsymbol{\alpha}_0) = \mathbf{0} \forall j, t = 1, \dots, n$. Considera-se a seguinte notação $E(\boldsymbol{\alpha}_0) = \mathbf{a}_0$ e $Cov(\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_0) = \mathbf{P}_0$, onde $\boldsymbol{\alpha}_0$ é vetor de estados inicial.

Uma das grandes vantagens dos modelos de espaço de estados decorre de sua natureza Markoviana, que permite que os cálculos sejam feitos recursivamente, de forma particularmente conveniente à computação. Segundo Durbin & Koopman (2001), essa natureza recursiva dos modelos e as técnicas computacionais usadas para analisá-los conferem aos modelos de espaço de estados grande flexibilidade, pois permitem que alterações na estrutura do sistema sejam feitas com relativa facilidade.

Outro ponto a favor da modelagem em espaço de estados é que ela permite trabalhar com observações faltantes e agregação temporal (Harvey, 1989, p.10). Além disso, modelos podem ser reformulados para tempo contínuo, habilitando-os, por exemplo, a tratar situações com observações irregularmente espaçadas no tempo.

Aqui é apresentada a forma de espaço de estados para séries temporais univariadas, mas esse modelo contempla também séries multivariadas. O interessado neste tópico pode consultar os livros de Harvey (1989) e West & Harrison (1997).

2.2.4 FILTRO DE KALMAN

O vetor de estados ($\boldsymbol{\alpha}_t$) é não-observável; logo, é necessário um algoritmo recursivo para estimá-lo. O Filtro de Kalman (FK) (Kalman, 1960) é um algoritmo recursivo que determina a estimativa do vetor de estados no tempo t dada toda a informação disponível até o instante $t - 1$. Outros algoritmos recursivos poderiam ser escolhidos, mas o FK é preferido devido a algumas propriedades interessantes, que serão vistas a seguir.

Sabendo-se que \mathbf{Y}_{t-1} é o vetor das observações até o instante $t - 1$, $E(\boldsymbol{\alpha}_0) = \mathbf{a}_0$ e $Cov(\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_0) = \mathbf{P}_0$, supondo-se que as observações têm distribuição normal e utilizando as propriedades da normal multivariada, tem-se que:

1. $(\boldsymbol{\alpha}_t | \mathbf{Y}_{t-1}) \sim N(\mathbf{a}_{t|t-1}, \mathbf{P}_{t|t-1})$, sendo $\mathbf{a}_{t|t-1} = E(\boldsymbol{\alpha}_t | \mathbf{Y}_{t-1}) = \mathbf{T}_t \mathbf{a}_{t-1} + \mathbf{c}_t$,
2. $\mathbf{P}_{t|t-1} = Var(\boldsymbol{\alpha}_t | \mathbf{Y}_{t-1}) = \mathbf{T}_t \mathbf{P}_{t-1} \mathbf{T}'_t + \mathbf{R}_t \mathbf{Q}_t \mathbf{R}'_t$,

3. $(y_t | \mathbf{Y}_{t-1}) \sim N(\tilde{y}_{t|t-1}, F_t)$, sendo $\tilde{y}_{t|t-1} = E(y_t | \mathbf{Y}_{t-1}) = \mathbf{z}'_t \mathbf{a}_{t|t-1} + d_t$,
4. $F_t = \text{Var}(y_t | \mathbf{Y}_{t-1}) = \mathbf{z}'_t \mathbf{P}_{t|t-1} \mathbf{z}_t + h_t$.

Através de recursividade, encontram-se os importantes resultados:

$$\mathbf{a}_t = \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{z}_t F_t^{-1} (y_t - \tilde{y}_{t|t-1}), \quad (2.5)$$

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{z}_t F_t^{-1} \mathbf{z}'_t \mathbf{P}_{t|t-1}. \quad (2.6)$$

As equações (2.5-2.6) são conhecidas como equações de atualização, pois através delas é possível atualizar os estimadores \mathbf{a}_t e \mathbf{P}_t do vetor de estado no instante t . Para facilitar o manuseio das fórmulas, denota-se por v_t o erro de previsão um passo à frente ou inovação

$$v_t = y_t - \tilde{y}_{t|t-1} \Rightarrow v_t = y_t - \mathbf{z}'_t \mathbf{a}_{t|t-1} - d_t.$$

Assim, $E(v_t) = 0$ e $\text{Var}(v_t) = F_t$. Então, substituindo v_t em (2.5-2.6) tem-se

$$\mathbf{a}_t = \mathbf{a}_{t|t-1} + \mathbf{T}_{t+1}^{-1} \mathbf{K}_t v_t$$

e

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{T}_{t+1}^{-1} \mathbf{K}_t \mathbf{z}'_t \mathbf{P}_{t|t-1},$$

onde

$$\mathbf{K}_t = \mathbf{T}_{t+1} \mathbf{P}_{t|t-1} \mathbf{z}_t F_t^{-1}$$

é conhecida como matriz de ganho de Kalman. Para simplificar as equações do FK, pode-se escrever $\mathbf{a}_{t+1|t}$ e $\mathbf{P}_{t+1|t}$ diretamente de $\mathbf{a}_{t|t-1}$ e $\mathbf{P}_{t|t-1}$. Logo, tem-se que

$$\mathbf{a}_{t+1|t} = \mathbf{T}_{t+1} \mathbf{a}_{t|t-1} + \mathbf{c}_{t+1} \quad (2.7)$$

$$\mathbf{P}_{t+1|t} = \mathbf{T}_{t+1} \mathbf{P}_{t|t-1} \mathbf{T}'_{t+1} - \mathbf{K}_t F_t \mathbf{K}'_t + \mathbf{R}_{t+1} \mathbf{Q}_{t+1} \mathbf{R}'_{t+1} \quad (2.8)$$

Para inicializar o filtro de Kalman toma-se $\mathbf{a}_0 = \mathbf{0}$ ou $\mathbf{a}_0 = E(y_t)$ e \mathbf{P}_0 como sendo um valor suficientemente grande (Harvey, 1989).

Como exemplo de aplicação do FK, será considerado o modelo de nível local, descrito nas equações (2.1-2.2). Neste caso, as matrizes de sistema \mathbf{z}'_t , T_t e R_t da forma de espaço de estados são iguais a um,

$$\mathbf{z}'_t = T_t = R_t = 1.$$

Além disso, $c_t = d_t = 0$ e as variâncias dos resíduos ε_t e η_t são $h_t = \sigma_\varepsilon^2$ e $Q_t = \sigma_\eta^2$, respectivamente. O estado do sistema, α_t , é referido como μ_t . Tomando-se as equações em (2.1-2.2)

na forma de espaço de estados do MNL e os valores encontrados das matrizes de sistema, substituindo-os nas fórmulas do filtro de Kalman e atendendo às suposições, encontram-se, para o MNL, equações mais simples e fáceis de serem usadas, porém com a mesma finalidade:

$$\begin{aligned}v_t &= y_t - a_{t|t-1}, \\a_{t+1|t} &= a_{t|t-1} + K_t \cdot v_t \\F_t &= P_{t|t-1} + \sigma_\varepsilon^2, \\K_t &= P_{t|t-1} / F_t, \\P_{t+1|t} &= P_{t|t-1} - K_t^2 F_t + \sigma_\eta^2.\end{aligned}$$

O FK tem outras características além de predição como, por exemplo, a suavização - que será definida a seguir.

2.2.5 SUAVIZAÇÃO

O FK pode ser utilizado também para estimar o vetor de estado, α_t , baseado em toda informação da amostra obtida, \mathbf{Y}_n . Existem vários métodos, que não são apresentados aqui, para calcular os estimadores suavizados do estado $\mathbf{a}_{t|n} = E(\alpha_t | \mathbf{Y}_n)$ e $\mathbf{P}_{t|n} = Var(\alpha_t | \mathbf{Y}_n)$. De acordo com a definição dada em Koopman (1992), esses estimadores são chamados de suavizadores do estado. Para encontrar esses estimadores, as seguintes equações são utilizadas:

$$\begin{aligned}\mathbf{r}_{t-1} &= \mathbf{z}_t F_t^{-1} + \mathbf{L}_t' \mathbf{r}_t, \\N_{t-1} &= \mathbf{z}_t F_t^{-1} \mathbf{z}_t' + \mathbf{L}_t' N_t \mathbf{L}_t, \\\mathbf{a}_{t|n} &= \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{r}_{t-1}, \\\mathbf{P}_{t|n} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} N_{t-1} \mathbf{P}_{t|t-1},\end{aligned}$$

em que $\mathbf{L}_t = \mathbf{T}_t - \mathbf{K}_t \mathbf{z}_t'$. São necessárias as quantidades $\mathbf{a}_{t|t-1}$, $\mathbf{P}_{t|t-1}$, \mathbf{K}_t , F_t e v_t obtidas através do FK para a execução do algoritmo, o qual é iniciado no instante $t = n$ e o suavizador do estado é obtido na ordem inversa até $t = 1$. Os valores iniciais \mathbf{r}_T e N_T podem ser fixados em zero. Uma exposição mais cuidadosa sobre esse tópico está em (Harvey, 1989, p.149).

2.3 INFERÊNCIA CLÁSSICA

Seja $\boldsymbol{\psi} = (\psi_1, \dots, \psi_p)'$ o vetor de hiperparâmetros. Por exemplo, os vetores paramétricos dos modelos MNL, MTL, MTP e MEB são $\boldsymbol{\psi} = (\sigma_\eta^2, \sigma_\varepsilon^2)'$, $\boldsymbol{\psi} = (\sigma_\eta^2, \sigma_\xi^2, \sigma_\varepsilon^2)'$, $\boldsymbol{\psi} = (\text{diag}(\mathbf{Q}_t), \sigma_\varepsilon^2)'$ e $\boldsymbol{\psi} = (\sigma_\eta^2, \sigma_\xi^2, \sigma_\omega^2, \sigma_\varepsilon^2)'$, respectivamente. Por se tratar de variâncias, o espaço paramétrico dos vetores é o \mathfrak{R}_+^p . O método de estimação por máxima verossimilhança será usado para estimar os hiperparâmetros de um modelo através da maximização da função densidade conjunta $f(y_1, \dots, y_n; \boldsymbol{\psi})$ em relação a $\boldsymbol{\psi}$. As propriedades assintóticas destes estimadores serão usadas para a construção do intervalo de confiança para $\boldsymbol{\psi}$, descritos na Seção 2.3.2.

2.3.1 ESTIMADOR DE MÁXIMA VEROSSIMILHANÇA

A função de verossimilhança pode ser calculada através das quantidades obtidas pelo filtro de Kalman, supondo que $(y_t | \mathbf{Y}_{t-1}) \sim N(\tilde{y}_{t|t-1}, F_t)$, onde $\mathbf{Y}_{t-1} = (y_1, \dots, y_{t-1})$. Neste caso, a função de densidade preditiva é dada por:

$$p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\psi}) = (2\pi)^{-1/2} |F_t|^{-1/2} \exp \left\{ -\frac{1}{2} (y_t - \tilde{y}_{t|t-1})' F_t^{-1} (y_t - \tilde{y}_{t|t-1}) \right\}.$$

Substituindo $v_t = y_t - \tilde{y}_{t|t-1}$ na equação anterior tem-se que:

$$p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\psi}) = (2\pi)^{-1/2} |F_t|^{-1/2} \exp \left\{ -\frac{1}{2} v_t' F_t^{-1} v_t \right\}.$$

Para encontrar o estimador de máxima verossimilhança de $\boldsymbol{\psi}$, deve-se maximizar a função de verossimilhança, que neste caso é o produto das distribuições preditivas $L(\boldsymbol{\psi}; \mathbf{Y}_n) = p(y_1, \dots, y_n | \boldsymbol{\psi}) = \prod_{t=1}^n p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\psi})$.

Aplicando-se o logaritmo natural em $L(\boldsymbol{\psi}; \mathbf{Y}_n)$ para simplificar os cálculos, a função de log-verossimilhança é obtida:

$$\ln L(\boldsymbol{\psi}; \mathbf{Y}_n) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^n \ln |F_t| - \frac{1}{2} \sum_{t=1}^n v_t' F_t^{-1} v_t. \quad (2.9)$$

O vetor $\hat{\boldsymbol{\psi}}$ de hiperparâmetros estimado é obtido maximizando a função de log-verossimilhança (2.9). Como esta é uma função não-linear do vetor de hiperparâmetros, a estimação deve ser realizada numericamente através de algoritmos de otimização. Há várias maneiras e métodos para solucionar problemas de otimização. Um dos métodos de maximização mais utilizados é o conhecido BFGS (em referência aos seus autores, Broyden-Fletcher-Goldfarb-Shanno), que será utilizado neste trabalho.

A vantagem de se considerar o método de máxima verossimilhança para obter estimadores dos hiperparâmetros é que se pode valer das propriedades ótimas que esses estimadores gozam sob certas condições de regularidade (Migon & Gamerman, 1999; Casella & Berger, 2002). Por exemplo, os EMV são assintoticamente consistentes e não-viesados. A normalidade assintótica é outra propriedade importante dos EMV - que será utilizada na próxima subseção para construir intervalos de confiança assintóticos para os hiperparâmetros.

2.3.2 INTERVALOS DE CONFIANÇA ASSINTÓTICOS

Sob certas condições de regularidade, o estimador de máxima verossimilhança, $\hat{\boldsymbol{\psi}}$, é assintoticamente normal com média $\boldsymbol{\psi}$ e matriz de covariância assintótica $\mathbf{aVar}(\hat{\boldsymbol{\psi}}) = n^{-1}\mathbf{IA}^{-1}(\boldsymbol{\psi})$, na qual $\mathbf{IA}(\boldsymbol{\psi}) = \lim_{n \rightarrow \infty} n^{-1}I(\boldsymbol{\psi})$ e $I(\boldsymbol{\psi})$ é a matriz de informação de Fisher. Além disso, em amostras finitas, $\mathbf{aVar}(\hat{\boldsymbol{\psi}})$ pode ser considerada como o inverso da matriz de informação, ver Casella & Berger (2002) e Migon & Gamerman (1999). Usando a distribuição assintótica dos estimadores de máxima verossimilhança, intervalos de confiança assintóticos podem ser construídos para os hiperparâmetros (Harvey, 1989).

Seja ψ_i , $i = 1, \dots, p$, um hiperparâmetro qualquer no vetor $\boldsymbol{\psi}$. Então, um intervalo de confiança assintótico para ψ_i de confiança $100.(1 - \kappa)\%$ é dado por:

$$\hat{\psi}_i \pm z_{\kappa/2} \sqrt{\mathbf{aVar}(\hat{\psi}_i)};$$

em que $z_{\kappa/2}$ é o $\kappa/2$ percentil de uma distribuição Normal-padrão e $\mathbf{aVar}(\hat{\psi}_i)$ é obtido dos elementos da diagonal da matriz $\mathbf{aVar}(\hat{\boldsymbol{\psi}})$. Para que as condições de regularidade sejam satisfeitas, em modelos mais comportados como MNL, MTL e MEB, basta que os hiperparâmetros sejam pontos interiores ao espaço paramétrico (ver Harvey (1989)).

O cálculo das derivadas do erro de previsão e de sua variância com respeito aos hiperparâmetros da matriz de informação não é uma tarefa trivial. Harvey (1989) propõe uma forma de aproximação numérica no cálculo das derivadas de F_t e v_t , que é extremamente útil no cálculo da matriz informação de Fisher cuja forma, em uma expressão simplificada, é dada por:

$$\mathbf{I}_{ij}(\boldsymbol{\psi}) = \frac{1}{2} \sum_t \left[\text{tr} \left[F_t^{-1} \frac{\partial F_t}{\partial \psi_i} F_t^{-1} \frac{\partial F_t}{\partial \psi_j} \right] \right] + \text{E} \left[\sum_t \left(\frac{\partial v_t}{\partial \psi_i} \right)' F_t^{-1} \frac{\partial v_t}{\partial \psi_j} \right], \quad i, j = 1, \dots, p \quad (2.10)$$

onde $t = 1, \dots, n$. As derivadas de F_t e v_t podem ser calculadas pelo seguinte procedimento:

1. Faça $i = 1$ e acrescente uma pequena quantidade, δ_i , ao hiperparâmetro ψ_i ;
2. Executa-se o Filtro de Kalman com o novo valor do hiperparâmetro (ψ_i) e mantém-se

constante os valores dos outros hiperparâmetros;

3. Um novo vetor do erro de previsão e sua variância são obtidos, $v_t^{(i)}$ e $F_t^{(i)}$;
4. As expressões $\delta_i^{-1}[v_t^{(i)} - v_t]$ e $\delta_i^{-1}[F_t^{(i)} - F_t]$ são, então, aproximações numéricas para as requeridas derivadas;
5. Faça $i = i + 1$ e volte ao passo 1, até $i = p$.

O primeiro artigo citado na Seção 1.2 (Franco, Santos, Ribeiro & Cruz, 2008) apresenta um estudo Monte Carlo para a determinação da quantidade δ_j . Foi mostrado que, para os ME analisados nesta dissertação, $\delta_j = 0,0001$ é uma boa escolha.

2.4 INFERÊNCIA BAYESIANA

A inferência Bayesiana sobre os parâmetros de modelos estruturais tem atraído um grande interesse recentemente. Admitindo que uma distribuição de probabilidade pode descrever as incertezas coerentemente, tal procedimento de inferência consiste em descrever a incerteza inicial de um pesquisador sobre um parâmetro através da distribuição *a priori* $\pi(\cdot)$ e depois combinar a informação proveniente dos dados, resultando na distribuição *a posteriori*. Se uma distribuição *a priori* $\pi(\boldsymbol{\psi})$ é especificada para $\boldsymbol{\psi}$, então a distribuição *a posteriori* de $\boldsymbol{\psi}$ é dada por:

$$\pi(\boldsymbol{\psi}|\mathbf{Y}_n) = \frac{L(\boldsymbol{\psi}; \mathbf{Y}_n)\pi(\boldsymbol{\psi})}{\int L(\boldsymbol{\lambda}; \mathbf{Y}_n)\pi(\boldsymbol{\lambda})d\boldsymbol{\lambda}}.$$

DISTRIBUIÇÕES A PRIORI

Na maioria das vezes, não é uma tarefa trivial especificar $\pi(\boldsymbol{\psi})$. Uma estratégia comum é a adoção de distribuições *a priori* vagas (com valores de variância grandes). No entanto, esta não é a estratégia mais adequada, uma vez que distribuições *a priori* vagas podem ser informativas, segundo Migon & Gamerman (1999). Na ausência de informação suficiente *a priori*, tem sido uma prática comum a especificação de distribuições *a priori* não-informativas apropriadas, obtidas através de algum método formal. A mais comum dessas distribuições *a priori* é a de Jeffreys (1961).

Neste trabalho serão utilizadas as seguintes distribuições *a priori* para os hiperparâmetros:

Uniforme: $\pi(\boldsymbol{\psi}) \propto c$, para todos os valores possíveis de $\boldsymbol{\psi}$, e 0, caso contrário, onde $c \in \mathfrak{R}_+$.

Jeffreys: $\pi(\boldsymbol{\psi}) = \det(\mathbf{I}(\boldsymbol{\psi}))^{1/2}$, onde $\mathbf{I}(\boldsymbol{\psi})$ é matriz de informação de Fisher dada em (2.10). Para o cálculo de $\mathbf{I}(\boldsymbol{\psi})$ utiliza-se o mesmo procedimento numérico proposto por Harvey (1989) para aproximação da matriz informação, que foi usado para construir o intervalo de confiança assintótico na Subseção 2.3.1.

2.4.1 ESTIMADOR DE BAYES

A distribuição *a posteriori* é toda a informação disponível sobre $\boldsymbol{\psi}$ após os dados serem observados. Apesar disto, muitas vezes é preciso resumi-la em um único número. Quando este número é obtido minimizando uma função de perda previamente escolhida, torna-se o estimador de Bayes. Seja $G(\boldsymbol{\psi}, \boldsymbol{\delta})$ a função de perda segundo a qual se quer ser penalizado, onde $\boldsymbol{\psi}$ é o valor real e $\boldsymbol{\delta}$ são possíveis estimativas de $\boldsymbol{\psi}$. O estimador de Bayes para $\boldsymbol{\psi}$ é o valor $\boldsymbol{\delta}$ que torna mínima a perda esperada *a posteriori*, isto é: $\hat{\boldsymbol{\psi}}_{EB} = \min_{\boldsymbol{\delta}} E(G(\boldsymbol{\psi}, \boldsymbol{\delta}) | \mathbf{Y}_n) = \min_{\boldsymbol{\delta}} \int G(\boldsymbol{\psi}, \boldsymbol{\delta}) \cdot \pi(\boldsymbol{\psi} | \mathbf{Y}_n) d\boldsymbol{\psi}$. Se a função de perda quadrática $G(\boldsymbol{\psi}, \boldsymbol{\delta}) = (\boldsymbol{\psi} - \boldsymbol{\delta})^2$ é considerada, o estimador de Bayes é a média *a posteriori*. Se, por outro lado, escolhem-se as funções de perda 0-1 e absoluta, a moda e mediana *a posteriori* são o estimador de Bayes $\hat{\boldsymbol{\psi}}_{EB}$, respectivamente. Deve-se ressaltar que diferentes funções de perda geram diferentes estimadores de Bayes e deve-se frisar que a escolha da função de perda é completamente subjetiva, ver Migon & Gamerman (1999).

2.4.2 INTERVALO DE CREDIBILIDADE

Seja ψ_i , $i = 1, \dots, p$, um hiperparâmetro qualquer no vetor $\boldsymbol{\psi}$. Fixado $0 < \kappa < 1$, qualquer intervalo $(t_1, t_2)'$ satisfazendo

$$\int_{t_1}^{t_2} \pi(\psi_i | \mathbf{Y}_n) d\psi_i = 1 - \kappa$$

é um intervalo de credibilidade para ψ_i de nível $100 \cdot (1 - \kappa)\%$. Esses intervalos são aproximados pelas correspondentes estatísticas de ordem da amostra gerada via MCMC de ψ_i , tomando os percentis simétricos. Para maiores detalhes ver Migon & Gamerman (1999).

A seguir, discutir-se-á um pouco sobre as técnicas computacionais *bootstrap* e MCMC. Essas técnicas serão extremamente úteis para se fazer inferências sobre os hiperparâmetros dos modelos descritos anteriormente.

2.5 BOOTSTRAP

O *bootstrap* (Efron, 1979) é um método computacional desenvolvido para inferência estatística. A idéia central desse método consiste em aproximar a distribuição dos dados (F) pela distribuição empírica (\hat{F}), baseada nas observações obtidas de F . Um resultado importante é que a distribuição estimada de F converge, em probabilidade, para a distribuição dos dados.

Existem basicamente duas maneiras de se realizar o *bootstrap*: não-paramétrica e paramétrica. Na primeira, a reamostragem é feita, com reposição, na própria amostra e cada observação tem a chance de $1/n$ de ser reamostrada, ou seja, a distribuição estimada empírica \hat{F}_{npar} é definida como uma distribuição discreta. Na segunda, \hat{F}_{par} é uma estimativa de F (conhecida e derivada de algum modelo paramétrico), isto é, as observações são reamostradas da distribuição geradora dos dados cujos parâmetros desconhecidos são substituídos por suas respectivas estimativas. O *bootstrap* paramétrico é útil em problemas onde se tem algum conhecimento sobre a forma da distribuição da população.

Suponha que a estatística de interesse é dada por $T(\mathbf{X})$, onde \mathbf{X} é uma amostra aleatória da distribuição dos dados. Usando o procedimento *bootstrap*, as B amostras *bootstrap* são obtidas $\mathbf{X}^{*1}, \mathbf{X}^{*2}, \dots, \mathbf{X}^{*B}$ e, conseqüentemente, as réplicas *bootstrap* da estatística $T^*(\mathbf{X}^{*b})$ para $b = 1, 2, \dots, B$. Pode-se usar $T^*(\mathbf{X}^*)$ como uma aproximação verdadeira da estatística $T(\mathbf{X})$ e, portanto, fazer inferências sobre o parâmetro de interesse. Esse procedimento é denominado como princípio *plug-in* por Efron & Tibshirani (1993, p.31). Segundo os mesmos, a função $\psi = T_F(\mathbf{X})$ da distribuição de F é estimada pela mesma distribuição empírica \hat{F} , $\hat{\psi} = T_{\hat{F}}(\mathbf{X})$, logo $\hat{\psi}$ é a estimativa do parâmetro ψ .

Efron & Tibshirani (1993) discorrem sobre duas vantagens do método *bootstrap* sobre os métodos tradicionais. A primeira, quando o modo não-paramétrico é usado, o analista ou pesquisador não tem que fazer suposições paramétricas sobre a população latente. Segunda, quando o modo paramétrico é utilizado, ele obtém respostas mais acuradas que as fórmulas tradicionais.

Sabe-se que para fazer inferências usando o *bootstrap* é necessário que a suposição de independência das observações seja válida. Levando em consideração que as observações em séries temporais são correlacionadas, uma das possibilidades é aplicar o *bootstrap* nos resíduos do modelo estrutural ajustado, utilizando uma proposta de Stoffer & Wall (1991). Os autores fornecem uma justificativa assintótica para a validade do procedimento *bootstrap* aplicado a modelos de espaço de estados.

Existem outras alternativas para realizar o *bootstrap* em séries temporais, como por ex-

emplo, o *bootstrap em bloco* (Efron & Tibshirani, 1993, p.99) - que permite a reamostragem diretamente nas observações da série. Essa técnica consiste em dividir a série em blocos de mesmo tamanho e reamostrá-los com reposição, obtendo, assim, a nova série *bootstrap*. Desta forma, a estrutura de correlação das observações em um mesmo bloco é preservada.

2.5.1 BOOTSTRAP NÃO-PARAMÉTRICO NOS RESÍDUOS

Seja o modelo na forma de espaço de estados dado em (2.3-2.4) com $d_t = 0$ e $\mathbf{c}_t = \mathbf{0}$. Inicialmente, deve-se estimar os hiperparâmetros do modelo, que, neste estudo, são as variâncias desconhecidas dos erros ε_t e $\boldsymbol{\eta}_t$ da equação da medida e transição, respectivamente. Para tanto, tem-se que executar o FK para obter os valores das inovações estimadas, \hat{v}_t , e sua variância \hat{F}_t . Deve-se observar que essas quantidades são funções dos hiperparâmetros desconhecidos.

Após a estimação dos hiperparâmetros, por máxima verossimilhança, têm-se as inovações estimadas, \hat{v}_t , que são reamostradas para a construção da série *bootstrap*. Antes, porém, é necessário centrar e padronizar as inovações para garantir que todos os resíduos possuam, pelo menos, os dois primeiros momentos iguais (Stoffer & Wall, 1991). Logo,

$$\hat{e}_t = \frac{\hat{v}_t - \bar{\hat{v}}_t}{\sqrt{\hat{F}_t}},$$

$t = 1, 2, \dots, n$, onde $\bar{\hat{v}}_t = \frac{\sum_{j=1}^n \hat{v}_j}{n}$. Agora amostras podem ser tomadas, com reposição, de \hat{e}_t , $t=1, 2, \dots, n$ para obter as inovações *bootstrap*, \hat{e}_t^* .

A série *bootstrap*, y_t^* , é construída, recursivamente, utilizando-se as inovações *bootstrap* obtidas acima e os valores \hat{F}_t e $\hat{\mathbf{K}}_t$ calculados pelo FK. Inicialmente, as equações do filtro de Kalman para $\mathbf{a}_{t+1|t}$ e y_t são escritas em função das inovações. Portanto, tem-se que

$$\mathbf{a}_{t+1|t} = \mathbf{T}_t \mathbf{a}_{t|t-1} + \mathbf{K}_t v_t$$

$$y_t = \mathbf{z}_t' \mathbf{a}_{t|t-1} + v_t.$$

A seguir, o vetor $\mathbf{S}_t = \begin{bmatrix} \mathbf{a}_{t+1|t} \\ y_t \end{bmatrix}$ é definido como

$$\mathbf{S}_t = \mathbf{A}_t \mathbf{S}_{t-1} + \mathbf{B}_t e_t, \quad (2.11)$$

$$t = 1, 2, \dots, n, \text{ onde } \mathbf{A}_t = \begin{bmatrix} \mathbf{T}_t & 0 \\ \mathbf{z}_t & 0 \end{bmatrix} \text{ e } \mathbf{B}_t = \begin{bmatrix} \mathbf{K}_t \sqrt{\hat{F}_t} \\ \sqrt{\hat{F}_t} \end{bmatrix}.$$

Portanto, resolvendo a equação (2.11) com \hat{e}_t^* no lugar de e_t e utilizando os valores estima-

dos F_t e K_t , obtém-se a série *bootstrap* y_t^* , $t = 1, 2, \dots, n$.

2.5.2 BOOTSTRAP PARAMÉTRICO NOS RESÍDUOS

Diferentemente do *bootstrap* não-paramétrico, no qual os resíduos do modelo ajustado são reamostrados com reposição, o *bootstrap* paramétrico utiliza apenas as estimativas dos parâmetros da série original. O procedimento é realizado da seguinte forma: primeiro, os parâmetros do modelo são estimados baseados na série original. A seguir, as variâncias dos erros nas equações das observações e da transição em (2.3-2.4) são substituídas pelas suas respectivas estimativas. Realizando reamostragens, os termos de erro *bootstrap* $\hat{\varepsilon}_t^*$ e $\hat{\eta}_t^*$ são obtidos e a série temporal *bootstrap* y_t^* pode ser construída através das equações:

$$y_t^* = \mathbf{z}_t' \boldsymbol{\alpha}_t^* + d_t + \hat{\varepsilon}_t^*,$$

$$\boldsymbol{\alpha}_t^* = \mathbf{T}_t \boldsymbol{\alpha}_{t-1}^* + \mathbf{c}_t + \mathbf{R}_t \hat{\eta}_t^*,$$

$t = 1, 2, \dots, n$.

2.5.3 INTERVALOS DE CONFIANÇA BOOTSTRAP

Um dos principais objetivos da teoria *bootstrap* é produzir bons intervalos de confiança automaticamente. "Bons" significa que os intervalos *bootstrap* devem estar estreitamente em acordo com os intervalos de confiança exatos naquelas situações especiais onde a teoria Estatística apresenta uma resposta exata, e devem fornecer probabilidades de cobertura acuradas em todas as situações, ver Efron & Tibshirani (1993).

• INTERVALO DE CONFIANÇA BOOTSTRAP PERCENTÍLICO

Neste estudo, utiliza-se um dos métodos que envolvem a construção de intervalos de confiança *bootstrap*, o *bootstrap* percentílico, para os hiperparâmetros em modelos estruturais. Esse método foi proposto por Efron & Tibshirani (1993).

Na prática, depois de estimar os valores dos hiperparâmetros para cada uma das B séries *bootstrap*, ordenam-se esses valores e toma-se o percentil $100 \cdot (\kappa/2)^0$ como o limite inferior do intervalo e o percentil $100 \cdot (1 - \kappa/2)^0$ como o limite superior do intervalo. Pode-se definir o intervalo $100 \cdot (1 - \kappa)\%$ como

$$\left[\hat{\psi}_i^{*(\kappa/2)}, \hat{\psi}_i^{*(1-\kappa/2)} \right],$$

onde $\hat{\psi}_i^*$, $i = 1, \dots, p$ é o vetor das estimativas ordenadas das B séries *bootstrap* de um hiperparâmetro qualquer em ψ .

• INTERVALO DE CONFIANÇA *BOOTSTRAP* BC

Esse método também utiliza os percentis da distribuição *bootstrap*, mas não exatamente o $(\kappa/2)^0$ e o $(1 - \kappa/2)^0$. Ao invés disso, ele corrige esses valores para possíveis vícios na estimação de ψ , através de uma quantidade \hat{z}_0 que mede o vício mediano de $\hat{\psi}$.

O intervalo BC (*Bias-corrected*) 100.(1 - κ)% é definido por

$$\left[\hat{\psi}_i^{*(\kappa_1)}; \hat{\psi}_i^{*(\kappa_2)} \right],$$

onde $\kappa_1 = \Phi(2\hat{z}_0 + z^{(\kappa/2)})$ e $\kappa_2 = \Phi(2\hat{z}_0 + z^{(1-\kappa/2)})$. A função $\Phi(\cdot)$ é a função de distribuição acumulada de uma Normal-padrão e $z^{(k/2)}$ seu 100.($\kappa/2$)^o percentil.

O valor de \hat{z}_0 é calculado usando a proporção de amostras *bootstrap* que são menores que os hiperparâmetros estimados usando a série original:

$$z_0 = \Phi^{-1} \left(\frac{\#\hat{\psi}_i^*(b) < \hat{\psi}_i}{B} \right).$$

• INTERVALO DE CONFIANÇA *BOOTSTRAP* BCA

No intervalo BCa (*Bias-corrected and accelerated*), além do cálculo da quantidade \hat{z}_0 , também é necessário calcular a aceleração dada por

$$\hat{a} = \frac{\sum_{i=1}^n (\hat{\psi}_i^{(\cdot)} - \hat{\psi}_i^{(j)})^3}{\left[\sum_{i=1}^n (\hat{\psi}_i^{(\cdot)} - \hat{\psi}_i^{(j)})^2 \right]^{3/2}},$$

onde $\hat{\psi}_i^{(j)}$ é uma estimativa do hiperparâmetro com a j-ésima observação deletada e $\hat{\psi}_i^{(\cdot)} = \sum_{j=1}^n \hat{\psi}_i^{(j)} / n$. Então, da mesma forma que se calcula o intervalo BC, calcula-se o intervalo BCa de nível 100.(1 - κ)%

$$\left[\hat{\psi}_i^{*(\kappa_1)}; \hat{\psi}_i^{*(\kappa_2)} \right]$$

mas com $\kappa_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\kappa/2)}}{1 - \hat{a}(\hat{z}_0 + z^{(\kappa/2)})}\right)$ e $\kappa_2 = \Phi\left(z_0 + \frac{z_0 + z^{(1-\kappa/2)}}{1 - \hat{a}(\hat{z}_0 + z^{(1-\kappa/2)})}\right)$.

2.5.4 COMPARAÇÕES ENTRE OS INTERVALOS DE CONFIANÇA *BOOTSTRAP*

Uma forma de comparar os intervalos descritos na Seção 2.5.3 consiste em avaliar se a probabilidade de cobertura dos mesmos converge para o nível nominal $(1 - \kappa)$ quando $n \rightarrow \infty$ (consistência) e quão rápida é essa convergência (precisão).

Definição 2. Um intervalo de confiança I_n para ψ com nível nominal $1 - \kappa$ é consistente se $P(\psi \in I_n) \rightarrow 1 - \kappa$ quando $n \rightarrow \infty$.

Os procedimentos *bootstrap* percentílico, BC e BCa são consistentes, segundo Shao & Tu (1995). Para intervalos bilaterais, os procedimentos *bootstrap* percentílico, BC e BCa são precisos de segunda ordem e também invariantes a transformações (isto é, se os pontos limites do intervalo são transformados corretamente quando se muda de ψ para uma função de ψ). Diferentemente da noção de acurácia, a definição de correção refere-se a quão próximo o ponto limite do intervalo candidato está do ponto limite do intervalo exato ou ideal. Em situações onde os limites podem ser definidos, a normal-padrão e a t-Student produzem apenas pontos limites de primeira ordem corretos, enquanto que alguns métodos *bootstrap* produzem pontos limites de confiança corretos de segunda ordem (Efron & Tibshirani, 1993, p.322).

O número de replicações *bootstrap* pode variar de um problema para outro. Por exemplo, na estimação do erro padrão, $B = 50$ é frequentemente suficiente (Efron & Tibshirani, 1993, p.52). Já na construção intervalos de confiança *bootstrap*, Efron & Tibshirani (1993) sugerem um valor de $B \geq 500$, dependendo da complexidade do problema e também do custo computacional. Hall (1986) aconselha tomar a probabilidade de cobertura nominal como um múltiplo de $(B + 1)^{-1}$. Para intervalos de confiança, na literatura, os valores $B = 500, 1000$ e 2000 são os mais recorrentes. Davidson & Hinkley (1997) propõem utilizar $B \geq 100$, para obter uma boa aproximação, se o objetivo for calcular o nível de significância em testes de hipóteses. Às vezes, a melhora nas estimativas de um determinado parâmetro é muito pequena, passando de um determinado valor de B para um valor maior. Então, o número de réplicas *bootstrap* deve ser escolhida com cautela.

2.6 MÉTODOS DE SIMULAÇÃO MCMC

Os métodos Markov chain Monte Carlo (MCMC) possibilitam a simulação de distribuições de maneira indireta, usando os algoritmos de Amostrador de Gibbs e Metropolis-Hastings. O MCMC consiste, essencialmente, em integração Monte Carlo usando cadeias de Markov. De

acordo com Gilks, Richardson & Spiegelhalter (1996), Bayesianos necessitam integrar fora distribuições de probabilidade de alta-dimensão com o objetivo de se fazer inferência sobre os parâmetros do modelo ou fazer previsões. Já os frequentistas precisam integrar fora a distribuição das observações, dado os valores dos parâmetros.

A idéia central dos métodos MCMC é construir uma cadeia de Markov da qual seja fácil gerar uma amostra e que tenha distribuição de equilíbrio igual à distribuição de interesse, onde cada estado pode ser atingido a partir de um outro qualquer em um número finito de iterações (ver Gamerman & Lopes (2006) e Gilks, Richardson & Spiegelhalter (1996)). Após um número suficientemente grande de iterações, a cadeia converge para a distribuição de interesse.

Mesmo após um grande número de iterações, a cadeia pode ainda apresentar altas autocorrelações entre seus valores, algo que não é desejado. Porém, o teorema ergódico (Gamerman & Lopes, 2006) garante que para cadeias de Markov satisfazendo determinadas condições e com $E_{\pi}(t(\boldsymbol{\psi})) < \infty$

$$\bar{t} = Ni^{-1} \sum_{k=1}^{Ni} t(\boldsymbol{\psi}^{(k)}) \longrightarrow E_{\pi}(t(\boldsymbol{\psi})), \quad (2.12)$$

quando $Ni \longrightarrow \infty$ com probabilidade 1, onde Ni é o número de valores da cadeia e $t(\boldsymbol{\psi})$ é alguma função de interesse de $\boldsymbol{\psi}$, por exemplo, no caso da média *a posteriori*, $t(\boldsymbol{\psi}) = \boldsymbol{\psi}$. Ele garante que a média dos valores da cadeia provê um estimador consistente dos parâmetros da distribuição limite $\pi(\cdot)$, apesar da sua dependência.

Esses métodos são muito utilizados na estatística Bayesiana para obter amostras de uma densidade *a posteriori* $\pi(\boldsymbol{\psi}|\mathbf{Y}_n)$, cuja geração direta é custosa ou complicada.

2.6.1 ALGORITMO DE METROPOLIS-HASTINGS

Esse algoritmo foi proposto por Metropolis *et al.* (1953), generalizado por Hastings (1970) e será utilizado neste trabalho para obter as médias, medianas e quantis das diferentes distribuições de interesse. Dada uma distribuição alvo $\pi(\cdot)$, o algoritmo de Metropolis-Hastings (M-H) e suas variantes fornecem meios para gerar cadeias de Markov ergódicas que tenham $\pi(\cdot)$ como distribuição de equilíbrio. Esse algoritmo é utilizado para simular uma distribuição *a posteriori*, utilizando uma densidade de proposição ou de transição de um estado x para um y , $q(y|x)$, que é uma distribuição condicional que governa as transições do estado, tendo as seguintes propriedades:

- $\int q(y|x)dy = 1$;
- $q(y|x)$ pode ser avaliada para todo x e y ;

- para cada x é possível gerar realizações que tenham distribuição $q(\cdot|x)$.

Seja $\boldsymbol{\psi} = (\psi_i, \dots, \psi_p)'$ o conjunto de parâmetros de interesse. O algoritmo de Metropolis-Hastings é inicializado a partir de um ponto arbitrário $\boldsymbol{\psi}^{(0)}$. Ele passa de um ponto $\boldsymbol{\psi}^{(j-1)}$ a outro $\boldsymbol{\psi}^{(j)}$ da cadeia de acordo com as seguintes etapas:

1. Faça $j = 1$;

2. Simule

$$\boldsymbol{\psi}^{(j)} \sim q(\cdot|\boldsymbol{\psi}^{(j-1)})$$

3. Avalie a razão

$$R = \frac{\pi(\boldsymbol{\psi}^{(j)})q(\boldsymbol{\psi}^{(j-1)}|\boldsymbol{\psi}^{(j)})}{\pi(\boldsymbol{\psi}^{(j-1)})q(\boldsymbol{\psi}^{(j)}|\boldsymbol{\psi}^{(j-1)})};$$

4. o próximo valor da cadeia será $\boldsymbol{\psi}^{(j)}$ com probabilidade igual $\min(1, R)$, ficando no ponto $\boldsymbol{\psi}^{(j-1)}$ com probabilidade complementar;

5. Faça $j = j + 1$ e retorne ao passo 2 até a convergência ser alcançada.

Às vezes, é muito custoso ou não é possível amostrar $\boldsymbol{\psi} = (\psi_i, \dots, \psi_p)'$ conjuntamente. Uma solução, nestes casos, é utilizar uma variação do algoritmo de Metropolis conhecida como M-H híbrido ou M-H com uma variável por vez.

2.6.2 ALGORITMO DE METROPOLIS-HASTINGS HÍBRIDO

Neste estudo, adota-se uma versão híbrida do algoritmo de M-H (Gamerman & Lopes, 2006). Nessa versão, os p hiperparâmetros de $\boldsymbol{\psi}$ são atualizados separadamente, com diferentes densidades de proposição. Os seguintes passos compõem, então, o algoritmo:

1. Inicie $\boldsymbol{\psi}_0$ e faça o contador das iterações $j = 1$;

2. Faça o contador das componentes $i = 1$;

3. Gere $\psi_i^{(j)} \sim q_i(\cdot|\boldsymbol{\psi}_i^{(j-1)})$;

4. Calcule a razão de Hastings, $R = \left[\pi_i(\psi_i^{(j)})q_i(\boldsymbol{\psi}_i^{(j-1)}|\boldsymbol{\psi}_i^{(j)}) \right] / \left[\pi_i(\boldsymbol{\psi}_i^{(j-1)})q_i(\psi_i^{(j)}|\boldsymbol{\psi}_i^{(j-1)}) \right]$,

onde $\pi_i(\psi_i)$ é a condicional completa de ψ_i . Nota-se que $\frac{\pi_i(\psi_i^{(j)})}{\pi_i(\boldsymbol{\psi}_i^{(j-1)})} = \frac{\pi(\boldsymbol{\Psi}_i^{(j)})}{\pi(\boldsymbol{\Psi}_i^{(j-1)})}$, em que

$$\boldsymbol{\Psi}_i^{(j)} = (\boldsymbol{\psi}_1^{(j)}, \dots, \boldsymbol{\psi}_{i-1}^{(j)}, \psi_i^{(j)}, \boldsymbol{\psi}_{i+1}^{(j-1)}, \dots, \boldsymbol{\psi}_p^{(j-1)})';$$

5. Faça a probabilidade de aceitação $\tau_i(\boldsymbol{\psi}_i^{(j-1)}, \boldsymbol{\psi}_i^{(j)}) = \min\{1, R\}$;

6. O próximo valor da cadeia será $\psi_i^{(j)}$ com probabilidade igual $\tau_i(\psi_i^{(j-1)}, \psi_i^{(j)})$, ficando no ponto $\psi_i^{(j-1)}$ com probabilidade complementar;
7. Mude o contador $i = i + 1$ e retorne ao passo 2 até $i = p$. Quando $i = p$, vá para o passo 8.
8. Mude o contador $j = j + 1$ e retorne ao passo 2 até a convergência ser alcançada.

A probabilidade de aceitação $\tau_i(\psi_i^{(j-1)}, \psi_i^{(j)})$ é definida de forma a garantir a reversibilidade da cadeia e sua convergência para a distribuição de equilíbrio $\pi(\cdot)$.

Há várias possibilidades de densidades de proposição na literatura. Neste trabalho, usa-se a densidade de proposição sendo um passeio aleatório $q_i(\psi_i^{(j)} | \psi_i^{(j-1)}) = \psi_i^{(j)} = \psi_i^{(j-1)} + \varepsilon$, onde $\varepsilon \sim N(0, \sigma^2)$ (Chib & Greenberg, 1995). Densidades de proposição que geram passos curtos conduzirão a altas taxas de aceitação, porém com baixa mistura, posto que a cadeia caminhará lentamente. Por outro lado, densidades que geram passos muito longos tenderão a ser rejeitadas, mantendo a cadeia estagnada e levando a baixas taxas de aceitação. Por isso, o valor de σ^2 é definido tal que as taxas de aceitação dos hiperparâmetros estejam entre 20% e 50% (Chib & Greenberg, 1995).

Em alguns casos, quando o espaço paramétrico de um parâmetro, por exemplo, é restrito a semi-reta positiva, valores negativos podem ser obtidos para o mesmo. Neste caso, faz-se necessário restringir essa densidade a um intervalo $[a, b]$, mas, para garantir que essa densidade preserve aquelas propriedades que definiram-se anteriormente, tem-se que fazer uma alteração na razão R que passa a ser

$$R = \left[\pi_i(\psi_i^{(j)}) q_i(\psi_i^{(j-1)} | \psi_i^{(j)}) \left(\Phi\left(\frac{a - \psi_i^{(j-1)}}{\sigma}\right) - \Phi\left(\frac{b - \psi_i^{(j-1)}}{\sigma}\right) \right) \right] / \left[\pi_i(\psi_i^{(j-1)}) q_i(\psi_i^{(j)} | \psi_i^{(j-1)}) \left(\Phi\left(\frac{a - \psi_i^{(j)}}{\sigma}\right) - \Phi\left(\frac{b - \psi_i^{(j)}}{\sigma}\right) \right) \right],$$

onde $\Phi(\cdot)$ denota a função de distribuição da Normal-padrão.

Outro aspecto importante é a convergência do algoritmo que pode ser verificada através do método de Gelman (1996), no qual múltiplas cadeias com diferentes valores iniciais são usadas. Esse método está disponível no R (2005) no pacote Coda (Plummer *et al.*, 2005). O traço da cadeia gerada é observado, procurando movimentos qualitativamente similares, os quais são um forte indício que a convergência da cadeia para distribuição estacionária foi atingida. Além disso, a função de autocorrelação das cadeias geradas também é analisada, pois uma distribuição estacionária não apresenta autocorrelações entre as observações. A questão do número de cadeias que se deve utilizar é um tanto quanto controversa na literatura, pois alguns

autores sugerem muitas cadeias pequenas, outros pesquisadores sugerem muitas cadeias longas ou, até mesmo, uma cadeia muita longa, ver Gilks, Richardson & Spiegelhalter (1996). Os valores iniciais devem ser escolhidos com cuidado, para que o *burn-in* da cadeia fique muito longo. Mas, se a cadeia é irredutível, os valores iniciais não afetarão a distribuição estacionária.

2.7 PREVISÃO

Como foi dito na Seção 2.1, um dos objetivos de modelar séries temporais é prever valores futuros. Portanto, o desenvolvimento das funções de previsão torna-se um importante tópico na análise de séries temporais. A previsão de um valor futuro y_{n+k} baseada em toda informação \mathbf{Y}_n pode ser obtida pela combinação da equação (2.3) no tempo $n+k$ com a equação do estado k passos à frente dada por

$$\boldsymbol{\alpha}_{n+k} = \left(\prod_{i=1}^k \mathbf{T}_{n+i} \right) \boldsymbol{\alpha}_n + \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \mathbf{T}_{n+k-j+1} \right) \mathbf{c}_{n+i} + \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \mathbf{T}'_{n+k-j+1} \right) \mathbf{R}'_{n+i} \boldsymbol{\eta}_{n+i}. \quad (2.13)$$

Para obter a previsão k passos à frente, define-se $\tilde{y}_{n+k|n}(\boldsymbol{\psi}) = E(y_{n+k} | \mathbf{Y}_n, \boldsymbol{\psi})$. Utilizando a forma de espaço de estados em (2.3-2.4), $\tilde{y}_{n+k|n}(\boldsymbol{\psi})$ é dada por

$$\mathbf{z}'_{n+k} \left(\left(\prod_{i=1}^k \mathbf{T}_{n+i} \right) \mathbf{a}_n + \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \mathbf{T}_{n+k-j+1} \right) \mathbf{c}_{n+i} \right) + d_{n+k}. \quad (2.14)$$

A variância de $(y_{n+k} | \mathbf{Y}_n, \boldsymbol{\psi})$ é dada por

$$\begin{aligned} & \mathbf{z}'_{n+k} \left(\prod_{i=1}^k \mathbf{T}'_{n+i} \right) \mathbf{P}_n \left(\prod_{i=1}^k \mathbf{T}_{n+i} \right) \mathbf{z}_{n+k} + \mathbf{z}'_{n+k} \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \mathbf{T}'_{n+k-j+1} \right) \mathbf{R}'_{n+i} \mathbf{Q}_{n+i} \mathbf{R}_{n+i} \\ & \left(\prod_{j=1}^{k-i} \mathbf{T}_{n+k-j+1} \right) \mathbf{z}_{n+k} + h_{n+k} \end{aligned} \quad (2.15)$$

Como exemplo, a previsão k passos à frente para o MNL é dada por $\tilde{y}_{n+k|n}(\boldsymbol{\psi}) = \mathbf{a}_n$ e a sua variância é igual a $\mathbf{P}_n + k\sigma_\eta^2 + \sigma_\varepsilon^2$.

ABORDAGEM CLÁSSICA

Neste caso, a função de previsão é dada por $\tilde{y}_{n+k|n}(\hat{\boldsymbol{\psi}})$, ou seja, ela é obtida substituindo-se o vetor de parâmetros $\boldsymbol{\psi}$ pelo seu estimador de máxima verossimilhança $\hat{\boldsymbol{\psi}}$ (Brockwell & Davis, 1996). Da mesma forma, a variância da função de previsão é obtida substituindo $\boldsymbol{\psi}$ por $\hat{\boldsymbol{\psi}}$ em (2.15).

Por exemplo, a função de previsão do MNL é dada por $\tilde{y}_{n+k|n}(\hat{\boldsymbol{\psi}}) = a_n$ e a sua variância por $\text{Var}(y_{n+k}|\mathbf{Y}_n, \hat{\boldsymbol{\psi}}) = P_n + k\hat{\sigma}_\eta^2 + \hat{\sigma}_\varepsilon^2$.

Um intervalo de confiança *bootstrap* percentílico de nível $(1 - \kappa)100\%$ para y_{n+k} é dado por

$$\left[\tilde{y}_{n+k}^{*(\kappa/2)}; \tilde{y}_{n+k}^{*(1-\kappa/2)} \right],$$

onde $\tilde{y}_{n+k}^{*(\kappa)}$ é o $B.(\kappa)$ -ésimo valor ordenado da replicação *bootstrap* para a previsão. Neste trabalho, a previsão *bootstrap* é calculada baseada no trabalho de Thombs & Schuncary (1990).

ABORDAGEM BAYESIANA

A função de previsão é dada pela média da distribuição preditiva de $y_{n+k}|\mathbf{Y}_n$ e é denotada por $\tilde{y}_{n+k}^{(b)}$. É obtida resolvendo a seguinte integral

$$\tilde{y}_{n+k}^{(b)} = \int \tilde{y}_{n+k}(\boldsymbol{\psi})\pi(\boldsymbol{\psi}|\mathbf{Y}_n)d\boldsymbol{\psi}$$

e um intervalo de credibilidade de $(1 - \kappa)100\%$ para y_{n+k} é dado por

$$\int_{li}^{ls} p(y_{n+k}|\mathbf{Y}_n)d\mathbf{Y}_n = 1 - \kappa.$$

Os limites li e ls podem ser obtidos aproximadamente por MCMC e os passos do algoritmo são descritos abaixo, uma vez que uma amostra de $\boldsymbol{\psi}^{(1)}, \dots, \boldsymbol{\psi}^{(m)}$ é obtida de $\pi(\boldsymbol{\psi}|\mathbf{Y}_n)$ para cada j com $j = 1, \dots, m$.

1. $\boldsymbol{\alpha}_n^{(j)}$ é gerado da distribuição $p(\boldsymbol{\alpha}_n|\boldsymbol{\psi}^{(j)}, \mathbf{Y}_n)$, obtido através do filtro de Kalman;
2. $\boldsymbol{\alpha}_n^{(j)}$ é gerado da distribuição de $p(\boldsymbol{\alpha}_n|\boldsymbol{\alpha}_n^{(j)}, \boldsymbol{\psi}^{(j)}, \mathbf{Y}_n)$, obtido pela equação (2.13);
3. Então, $y_{n+k}^{(j)} = \mathbf{z}_{n+k}'\boldsymbol{\alpha}_{n+k}^{(j)} + \varepsilon_{n+k}^{(j)}$, onde $\varepsilon_{n+k}^{(j)}$ é gerado de uma distribuição Normal com média zero e variância $\sigma_\varepsilon^{2(j)}$.

Finalmente, os valores de $y_{n+k}^{(1)}, \dots, y_{n+k}^{(m)}$ são ordenados e os percentis $\kappa/2$ e $1 - \kappa/2$ são tomados como os limites inferior e superior do intervalo, respectivamente.

3 COMPARAÇÃO DE MÉTODOS CLÁSSICOS E BAYESIANOS PARA INFERÊNCIA SOBRE OS HIPERPARÂMETROS EM MODELOS ESTRUTURAIS

Nesta seção os resultados obtidos dos experimentos Monte Carlo (MC), implementados na linguagem Ox (Doornik, 1999), são apresentados. Os desempenhos dos estimadores de máxima verossimilhança (EMV), *bootstrap* (BOOT) e dos estimadores de Bayes (EB) foram avaliados para o MNL, MTL e MEB. Neste trabalho, são mostrados os resultados para os valores dos hiperparâmetros fixados em $\sigma_{\eta}^2 = 0.50$, $\sigma_{\varepsilon}^2 = 1.00$, $\sigma_{\xi}^2 = 0.10$ e $\sigma_{\omega}^2 = 0.03$. Outras combinações de valores foram utilizadas, mas como os resultados foram similares aos apresentados aqui, não são mostradas. Séries de tamanho $n = 60$ e 200 foram geradas com um *burn-in* igual a 100 . Para os EB, 10.300 amostras foram geradas das quais as 300 primeiras foram excluídas. O número de MC e *bootstrap* foram fixados em 500 . O nível e a probabilidade dos intervalos de confiança e de credibilidade, respectivamente, foram fixados em 0.95 . As estimativas dos parâmetros são comparadas através do vício e do erro quadrático médio (EQM). Todas as tabelas deste capítulo foram retiradas de Santos & Franco (2008).

3.1 COMPARANDO OS PROCEDIMENTOS DE ESTIMAÇÃO

As Tabelas 3.1 a 3.3 mostram, respectivamente, os resultados da estimação dos hiperparâmetros nos modelos MNL, MTL e MEB.

Da Tabela 3.1, pode ser visto que os estimadores Bayesianos apresentam, para o componente de nível (σ_{η}^2), um vício maior que o EMV. As estimativas Bayesianas tendem a superestimar o valor real deste hiperparâmetro, mesmo para séries de tamanho $n = 200$. Entretanto,

Tabela 3.1: Resultados da estimação no MNL segundo os métodos de Máxima verossimilhança, *bootstrap* e estimação Bayesiana.

n	ψ	EMV	BOOT	EB-Mediana	EB-Media	EB-Mediana	EB-Media
		Priori Uniforme				Priori Jeffreys	
60	$\sigma_\eta^2 = 0.5$	0.483	0.517	0.654	0.725	0.513	0.580
		(-0.017)	(0.017)	(0.154)	(0.225)	(0.013)	(0.080)
		0.063	0.072	0.093	0.124	0.057	0.066
	$\sigma_\varepsilon^2 = 1.0$	1.017	1.002	1.020	1.055	1.055	1.088
		(0.017)	(0.002)	(0.020)	(0.055)	(0.055)	(0.087)
		0.081	0.086	0.084	0.089	0.101	0.108
200	$\sigma_\eta^2 = 0.5$	0.500	0.503	0.540	0.560	0.510	{0.529}
		(0.000)	(0.003)	(0.040)	(0.060)	(0.010)	(0.029)
		0.017	0.017	0.026	0.026	0.016	0.017
	$\sigma_\varepsilon^2 = 1.0$	1.000	1.001	1.020	1.004	1.020	1.013
		(0.000)	(0.001)	(0.020)	(0.004)	(0.020)	(0.013)
		0.016	0.017	0.024	0.024	0.025	0.025

Obs.: Os números entre parêntesis são o vício e, em negrito, são o erro quadrático médio.

o EQM tem a mesma magnitude para todos os estimadores. Com respeito ao componente de erro, σ_ε^2 , não há muita diferença entre os métodos. Também nota-se que o *bootstrap* imita o comportamento do EMV, o que permite a essa técnica ser usada como ferramenta para construir intervalos de confiança neste caso. Comparando os estimadores Bayesianos, a distribuição *a priori* de Jeffreys parece levar a estimativas mais próximas do valor real dos hiperparâmetros do que a distribuição *a priori* uniforme. Como esperado, o EQM decresce quando o tamanho da amostra aumenta.

Os resultados para o MTL são apresentados na Tabela 3.2. As mesmas conclusões obtidas para os hiperparâmetros σ_η^2 e σ_ε^2 no modelo MNL com respeito ao vício e ao EQM podem ser conseguidas aqui. Quanto ao componente de tendência, σ_ξ^2 , parece não existir muita diferença entre os procedimentos, sendo todas estimativas muito próximas do valor real do hiperparâmetro.

Simulações para o MEB são apresentadas na Tabela 3.3. Parece que a introdução do componente sazonal no modelo traz algumas mudanças nos métodos de estimação com respeito às conclusões retiradas para os outros dois modelos. Por exemplo, o vício ainda é maior para os estimadores Bayesianos no caso do componente de nível, σ_η^2 , mas agora o EB com a distribuição *a priori* uniforme apresenta o menor EQM. Para a tendência, σ_ξ^2 , o EB com a distribuição *a priori* uniforme possui maior vício que os outros procedimentos, entretanto o EQM é muito similar para todos eles. O componente de erro, σ_ε^2 , tem um menor vício que para o EB com distribuição *a priori* de Jeffreys, mas o menor EQM para os EB com distribuição *a priori* uniforme, entretanto os resultados são muito parecidos para tamanho de amostra $n = 200$. Com

Tabela 3.2: Resultados da estimação no MTL segundo os métodos de Máxima verossimilhança, *bootstrap* e estimação Bayesiana.

n	ψ	MLE	BOOT	EB-Mediana	EB-Media	EB-Mediana	EB-Media
		Priori Uniforme				Priori Jeffreys	
60	$\sigma_\eta^2 = 0.5$	0.562	0.670	0.681	0.758	0.688	0.740
		(0.062)	(0.170)	(0.181)	(0.258)	(0.188)	(0.240)
		0.364	0.287	0.160	0.201	0.199	0.215
	$\sigma_\xi^2 = 0.1$	0.097	0.100	0.199	0.233	0.070	0.076
		(-0.003)	(0.000)	(0.099)	(0.133)	(-0.030)	(-0.024)
		0.005	0.004	0.236	0.252	0.004	0.004
	$\sigma_\varepsilon^2 = 1.0$	0.963	0.912	0.943	0.952	0.907	0.916
		(-0.037)	(-0.088)	(-0.057)	(-0.048)	(-0.093)	(-0.084)
		0.153	0.132	0.098	0.095	0.085	0.082
200	$\sigma_\eta^2 = 0.5$	0.522	0.530	0.553	0.554	0.575	0.573
		(0.022)	(0.030)	(0.053)	(0.054)	(0.075)	(0.073)
		0.120	0.112	0.091	0.090	0.098	0.097
	$\sigma_\xi^2 = 0.1$	0.099	0.099	0.107	0.092	0.110	0.097
		(-0.001)	(-0.001)	(0.007)	(-0.008)	(0.010)	(-0.003)
		0.001	0.001	0.001	0.001	0.001	0.001
	$\sigma_\varepsilon^2 = 1.0$	0.982	0.979	0.991	0.955	0.993	0.956
		(-0.018)	(-0.022)	(-0.009)	(-0.045)	(-0.007)	(-0.044)
		0.045	0.044	0.038	0.038	0.043	0.042

Obs.: Os números entre parêntesis são o vício e, em negrito, são o erro quadrático médio.

respeito ao componente sazonal, o EMV é muito similar ao EB com distribuição *a priori* de Jeffreys, apresentando vício e EQM menores. Novamente, o *bootstrap* mostrou um desempenho próximo dos resultados do EMV. Comparando os estimadores Bayesianos, a mediana apresenta menores valores do vício e do EQM, independente da distribuição *a priori* utilizada, exceto para σ_ε^2 , onde a média parece ter um desempenho melhor.

3.2 INTERVALOS DE CONFIANÇA E DE CREDIBILIDADE

É bem conhecido que a probabilidade de credibilidade e o nível de confiança têm diferentes significados e interpretações. Citando Casella e Berger (2002), "A primeira reflete a crença subjetiva do pesquisador, enquanto a última reflete a incerteza no procedimento amostral". Assim, neste trabalho, uma tentativa de comparar os resultados obtidos usando esses dois procedimentos será feita. A porcentagem de vezes que os intervalos construídos contêm o valor verdadeiro dos hiperparâmetros é obtida baseada em um número grande de simulações Monte Carlo. Os resultados, junto com a amplitude dos intervalos, são mostrados nas Tabelas 3.4 a 3.6, para o MNL, MTL e MEB, respectivamente.

Da Tabela 3.4 pode ser visto que, para o MNL, todos intervalos apresentam, em geral,

Tabela 3.3: Resultados da estimação no MEB segundo os métodos de Máxima verossimilhança, *bootstrap* e estimação Bayesiana.

n	ψ	MLE	BOOT	EB-Mediana	EB-Media	EB-Mediana	EB-Media
				Priori Uniforme		Priori Jeffreys	
60	$\sigma_\eta^2 = 0.5$	0.550	0.507	0.614	0.663	0.616	0.677
		(0.050)	(0.007)	(0.114)	(0.163)	(0.116)	(0.177)
		0.223	0.143	0.110	0.124	0.172	0.188
	$\sigma_\xi^2 = 0.03$	0.026	0.036	0.049	0.056	0.023	0.038
		(-0.004)	(0.006)	(0.019)	(0.026)	(-0.007)	(0.008)
		0.001	0.002	0.001	0.002	0.001	0.001
	$\sigma_\omega^2 = 0.1$	0.105	0.097	0.152	0.166	0.103	0.124
		(0.005)	(0.004)	(0.052)	(0.066)	(0.003)	(0.024)
		0.007	0.006	0.011	0.013	0.007	0.008
	$\sigma_\varepsilon^2 = 1.0$	0.973	1.109	0.914	0.932	0.991	1.009
		(-0.027)	(0.109)	(-0.086)	(-0.068)	(0.009)	(0.009)
		0.197	0.218	0.140	0.132	0.175	0.167
200	$\sigma_\eta^2 = 0.5$	0.523	0.500	0.539	0.551	0.521	0.536
		(0.023)	(0.000)	(0.039)	(0.051)	(0.021)	(0.036)
		0.072	0.080	0.055	0.058	0.075	0.073
	$\sigma_\xi^2 = 0.03$	0.029	0.030	0.034	0.036	0.029	0.033
		(-0.001)	(0.000)	(0.004)	(0.006)	(0.001)	(0.003)
		0.000	0.000	0.000	0.000	0.000	0.000
	$\sigma_\omega^2 = 0.1$	0.100	0.100	0.112	0.115	0.096	0.102
		(0.000)	(0.000)	(0.012)	(0.015)	(0.004)	(0.002)
		0.001	0.001	0.002	0.002	0.001	0.001
	$\sigma_\varepsilon^2 = 1.0$	0.996	1.053	0.989	0.994	1.004	1.008
		(-0.004)	(0.053)	(-0.011)	(0.006)	(0.004)	(0.008)
		0.054	0.071	0.057	0.054	0.058	0.057

Obs.: Os números entre parêntesis são o vício e, em negrito, são o erro quadrático médio.

Tabela 3.4: Intervalos de confiança e de credibilidade de 95% para o MNL.

n	ψ	Assint.	Cred.Unif	Cred.Jeff	Perc	BC	BCa
60	$\sigma_{\eta}^2 = 0.5$	[0.03; 0.94]	[0.25; 1.61]	[0.19; 1.36]	[0.11; 1.13]	[0.16; 1.26]	[0.16; 1.26]
		(0.91)	(1.36)	(1.17)	(1.02)	(1.10)	(1.10)
		0.88	0.95	0.98	0.90	0.92	0.92
	$\sigma_{\varepsilon}^2 = 1.0$	[0.45; 1.58]	[0.22; 1.83]	[0.51; 1.85]	[0.44; 1.62]	[0.47; 1.65]	[0.43; 1.61]
(1.13)		(1.61)	(1.34)	(1.18)	(1.18)	(1.18)	
		0.95	0.97	0.94	0.94	0.94	0.95
200	$\sigma_{\eta}^2 = 0.5$	[0.24; 0.76]	[0.32; 0.92]	[0.30; 0.87]	[0.27; 0.79]	[0.28; 0.81]	[0.28; 0.82]
		(0.52)	(0.60)	(0.57)	(0.52)	(0.53)	(0.54)
		0.93	0.94	0.97	0.94	0.94	0.94
	$\sigma_{\varepsilon}^2 = 1.0$	[0.69; 1.31]	[0.72; 1.38]	[0.71; 1.36]	[0.70; 1.33]	[0.71; 1.34]	[0.70; 1.33]
(0.62)		(0.66)	(0.65)	(0.63)	(0.63)	(0.63)	
		0.97	0.96	0.95	0.93	0.97	0.97

Obs.: Os números em colchetes são os limites médios dos intervalos, em parêntesis a amplitude e, em negrito, a taxa de cobertura dos intervalos.

taxas de cobertura próximas do nível nominal assumido de 0.95, a única exceção sendo o intervalo assintótico para $n = 60$ e $\sigma_{\eta}^2 = 0.5$. No entanto, os intervalos de confiança *bootstrap* e assintótico devem ser preferidos, porque possuem menores amplitudes.

Na Tabela 3.5 estão os resultados para o MTL. Para o componente de nível, σ_{η}^2 , os intervalos de credibilidade, em geral, fornecem a melhor combinação de taxa de cobertura mais próxima do nível nominal de 0.95 e menores amplitudes. Para o componente de tendência, σ_{ξ}^2 , os intervalos *bootstrap*, especialmente o BC e o BCa, mostram um melhor desempenho. Finalmente, para o componente do erro, os intervalos *bootstrap* para σ_{ε}^2 têm a taxa de cobertura próxima de 0.95, entretanto a amplitude desses intervalos é ligeiramente maior do que dos intervalos de credibilidade. Pode ser notado também que, muitas vezes, os intervalos assintóticos apresentam valores negativos para o limite inferior dos intervalos, e isso não é desejável, uma vez que os hiperparâmetros são variâncias.

Para o MEB, mostrado na Tabela 3.6, os intervalos de credibilidade apresentam, em geral, a melhor combinação de menor amplitude para os intervalos e taxa de cobertura mais próxima de 0.95. Em alguns casos os intervalos BC e BCa estão mais próximos do ponto 0.95, com amplitude aceitável, como por exemplo em $n = 200$ com σ_{ω}^2 e σ_{ε}^2 . Novamente, para o limite inferior do intervalo assintótico, todos os limites são negativos, exceto para σ_{ε}^2 .

Tabela 3.5: Intervalos de confiança e de credibilidade de 95% para o MTL.

n	ψ	Assint.	Cred.Unif	Cred.Jeff	Perc	BC	BCa
60	$\sigma_{\eta}^2 = 0.5$	[-0.67; 1.79]	[0.07;1.89]	[0.14;1.65]	[0.01;2.15]	[0.01;2.62]	[0.01;2.55]
		(2.46)	(1.82)	(1.51)	(2.15)	(2.61)	(2.54)
		0.97	0.99	0.93	0.99	0.99	0.99
	$\sigma_{\xi}^2 = 0.1$	[-0.04;0.23]	[0.04; 0.84]	[0.02;0.17]	[0.01; 0.25]	[0.01;0.33]	[0.01;0.32]
		(0.27)	(0.80)	(0.15)	(0.24)	(0.32)	(0.31)
		0.88	0.89	0.71	0.86	0.97	0.95
	$\sigma_{\varepsilon}^2 = 1.0$	[0.22;1.71]	[0.41;1.52]	[0.44;1.43]	[0.19;1.64]	[0.15;1.64]	[0.16;1.65]
		(1.49)	(1.11)	(0.99)	(1.45)	(1.49)	(1.49)
		0.92	0.93	0.90	0.94	0.96	0.96
200	$\sigma_{\eta}^2 = 0.5$	[-0.16; 1.20]	[0.14;1.14]	[0.16;1.09]	[0.05;1.30]	[0.05;1.46]	[0.05;1.43]
		(1.36)	(1.00)	(0.93)	(1.25)	(1.41)	(1.38)
		0.97	0.93	0.90	0.98	0.99	0.99
	$\sigma_{\xi}^2 = 0.1$	[0.02;0.18]	[0.06; 0.18]	[0.05; 0.15]	[0.04; 0.19]	[0.04;0.19]	[0.04;0.19]
		(0.16)	(0.13)	(0.10)	(0.15)	(0.15)	(0.15)
		0.90	0.90	0.87	0.94	0.95	0.95
	$\sigma_{\varepsilon}^2 = 1.0$	[0.57;1.40]	[0.68;1.31]	[0.66;1.26]	[0.55;1.39]	[0.54;1.38]	[0.54;1.40]
		(0.83)	(0.63)	(0.60)	(0.84)	(0.84)	(0.86)
		0.96	0.92	0.88	0.96	0.95	0.95

Obs.: Os números em colchetes são os limites médios dos intervalos, em parêntesis a amplitude e em negrito a taxa de cobertura dos intervalos.

3.3 APLICAÇÃO À SÉRIE DA PRECIPITAÇÃO DE SO_4 EM NOVA YORK

Nesta seção, as metodologias descritas nas seções anteriores são aplicadas a uma série temporal real da precipitação do sulfato (SO_4) em Nova York, EUA. Os dados são coletados pelo *National Atmospheric Deposition Program*¹ com o objetivo de monitoramento geográfico e temporal das precipitações químicas. Os dados são a média (mg/L) da precipitação de SO_4 de vários lugares na cidade de Nova York. A série de SO_4 é composta de 271 observações mensais no período de Janeiro de 1985 a Julho de 2007. Uma transformação logarítmica foi feita na série, para estabilizar a variância da mesma.

Pode ser visto na Figura 3.1 que os componentes de nível e sazonalidade estão presentes na série, entretanto pode existir alguma dúvida quanto ao componente de inclinação, que parece apresentar um comportamento descendente suave. Então, o MEB será ajustado e os intervalos de confiança e de credibilidade serão construídos para os hiperparâmetros, a fim de checar se o MEB é adequado neste caso.

¹O programa Nacional Atmosférico é um programa cooperativo de suporte à pesquisa das estações de experimentos dos estados agrícolas, agências do Estado, Federal e organizações não-governamentais de pesquisa nos EUA. Os dados estão disponíveis em <http://nadp.sws.uiuc.edu/>.

Tabela 3.6: Intervalos de confiança e de credibilidade de 95% para o MEB.

n	ψ	Assint.	Cred.Unif	Cred.Jeff	Perc	BC	BCa
60	$\sigma_{\eta}^2 = 0.50$	[-0.34;1.43] (1.77) 0.96	[0.10;1.51] (1.41) 0.97	[0.13;1.58] (1.45) 0.92	[0.00;1.71] (1.71) 0.96	[0.00;2.25] (2.25) 0.99	[0.00; 2.19] (2.19) 0.99
	$\sigma_{\xi}^2 = 0.03$	[-0.02;0.08] (0.10) 0.75	[0.01;0.15] (0.14) 0.91	[0.01;0.21] (0.20) 0.96	[0.00;0.14] (0.14) 0.91	[0.00;0.40] (0.40) 0.96	[0.00;0.38] (0.38) 0.96
	$\sigma_{\omega}^2 = 0.10$	[-0.05;0.27] (0.32) 0.82	[0.04;0.38] (0.34) 0.88	[0.03;0.36] (0.33) 0.92	[0.01;0.28] (0.27) 0.86	[0.01;0.43] (0.43) 0.99	[0.00; 0.44] (0.44) 0.99
	$\sigma_{\varepsilon}^2 = 1.00$	[0.12;1.78] (1.66) 0.90	[0.31;1.64] (1.33) 0.91	[0.40;1.71] (1.31) 0.88	[0.14; 2.11] (1.97) 0.99	[0.11;2.11] (2.00) 0.98	[0.11; 2.09] (1.98) 0.98
200	$\sigma_{\eta}^2 = 0.50$	[-0.17;1.20] (1.37) 0.93	[0.13;1.22] (1.09) 0.95	[0.15;1.28] (1.13) 0.89	[0.02;1.45] (1.43) 0.96	[0.04;1.77] (1.73) 0.99	[0.04; 1.75] (1.71) 0.99
	$\sigma_{\xi}^2 = 0.03$	[-0.01;0.07] (0.08) 0.82	[0.01;0.09] (0.08) 0.91	[0.01;0.12] (0.11) 0.96	[0.00;0.08] (0.08) 0.89	[0.00;0.10] (0.09) 0.98	[0.00;0.10] (0.10) 0.98
	$\sigma_{\omega}^2 = 0.10$	[-0.01;0.21] (0.22) 0.87	[0.05;0.26] (0.21) 0.87	[0.03;0.25] (0.22) 0.89	[0.01;0.23] (0.22) 0.89	[0.02;0.26] (0.24) 0.95	[0.02; 0.27] (0.25) 0.95
	$\sigma_{\varepsilon}^2 = 1.00$	[0.32;1.60] (1.28) 0.91	[0.46;1.51] (1.05) 0.92	[0.53;1.61] (1.08) 0.88	[0.28; 1.82] (1.54) 0.96	[0.24;1.81] (1.57) 0.96	[0.23; 1.77] (1.54) 0.96

Obs.: Os números em colchetes são os limites médios dos intervalos, em parêntesis, a amplitude e, em negrito, a taxa de cobertura dos intervalos.

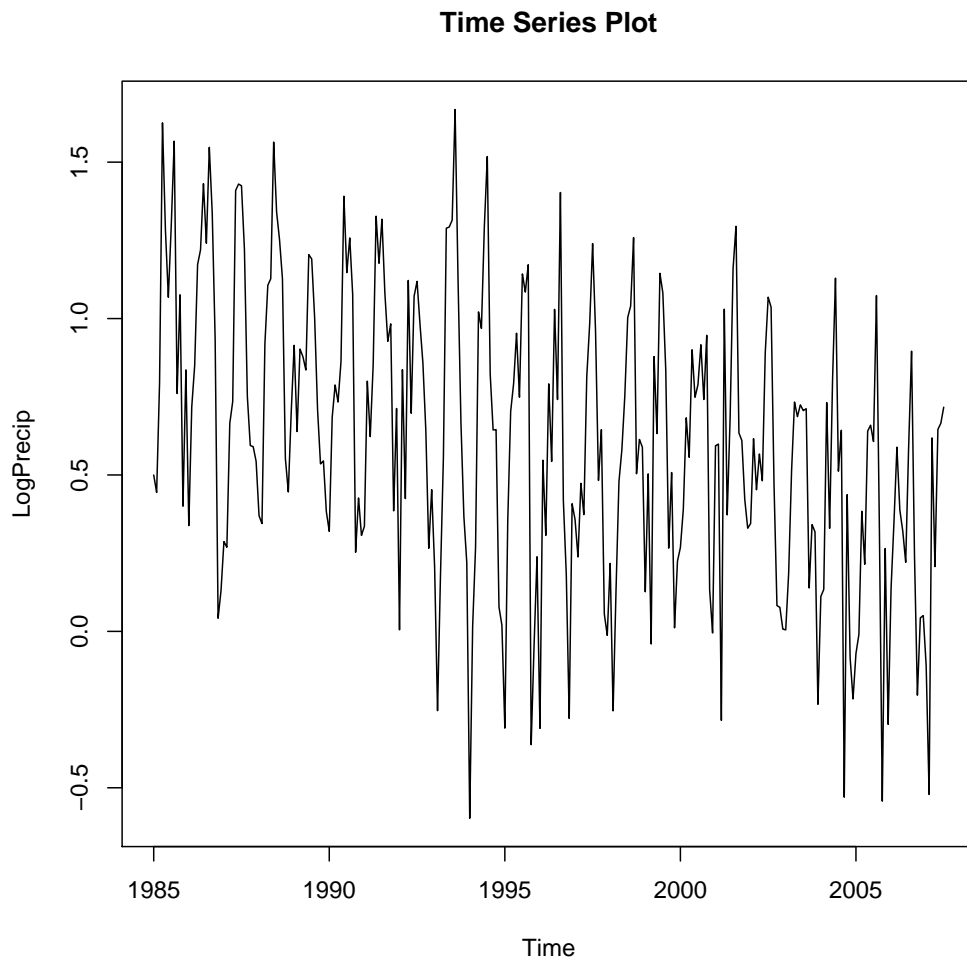


Figura 3.1: Logaritmo Precipitação de SO4 em Nova York.

A Tabela 3.7 mostra o ajuste do MEB para a série. Pode ser observado que σ_{ξ}^2 , a variância relacionada ao componente de inclinação, foi estimada como zero (considerando até sexta casa decimal) em todos os métodos utilizados. Os intervalos de confiança e de credibilidade, mostrados na Tabela 3.8, confirmam esse fato, já que todos os intervalos têm os limites inferior e superior iguais a zero, exceto o intervalo assintótico que inclui o zero. Como esse hiperparâmetro está na fronteira do espaço paramétrico, a distribuição assintótica de $\hat{\psi}$ é modificada para levar em conta a fronteira, mas a distribuição univariada dos hiperparâmetros não é afetada. Então, o intervalo de confiança assintótico para σ_{ξ}^2 neste caso não é confiável, justificando o uso dos intervalos *bootstrap* e de credibilidade. Os outros hiperparâmetros não estão próximos de zero. Note que os limites dos intervalos *bootstrap* percentílico, BC e BCa para os hiperparâmetros estão próximos. A mesma observação pode ser feita para os intervalos de credibilidade entre si. A análise dos resíduos não mostrou nenhuma violação das suposições do modelo.

Tabela 3.7: Estimativas de Máxima verossimilhança, *bootstrap* e estimação Bayesiana para os hiperparâmetros da série de SO_4 .

ψ	EMV	BOOT	EB-Mediana		EB-Media	
			Priori Uniforme	Priori Jeffreys	Priori Uniforme	Priori Jeffreys
σ_η^2	0.13	0.13	0.24	0.26	0.12	0.14
σ_ξ^2	0.00	0.00	0.00	0.00	0.00	0.00
σ_ω^2	0.24	0.28	0.52	0.56	0.41	0.47
σ_ε^2	68.20	67.96	67.52	67.54	67.45	67.68

Obs.: Os valores desta tabela estão multiplicados por 1000.

Tabela 3.8: Intervalos de confiança e de credibilidade para os hiperparâmetros da série de SO_4 .

Intervalos	σ_η^2	σ_ξ^2	σ_ω^2	σ_ε^2
Assint.	[-0.41;0.67]	[-0.05;0.05]	[-0.56;1.04]	[54.03;82.37]
Cred.Unif	[0.02;0.66]	[0.00;0.00]	[0.04;1.34]	[58.63;76.64]
Cred.Jeff	[0.01;0.44]	[0.00;0.00]	[0.02;1.36]	[58.96;76.80]
Perc	[0.00;0.63]	[0.00;0.00]	[0.00;1.13]	[53.45; 81.42]
BC	[0.00;0.78]	[0.00;0.00]	[0.00;1.22]	[52.27;80.89]
BCa	[0.00;0.77]	[0.00;0.00]	[0.00;1.22]	[52.35;80.93]

Obs.: Os valores desta tabela estão multiplicados por 1000.

4 FUNÇÃO DE TRANSFERÊNCIA

Séries temporais são frequentemente afetadas por eventos externos, conhecidos como intervenções, que podem mudar a estrutura da série (tais como quebras, mudanças políticas, etc.). Há evidências na literatura que as primeiras propostas de análise de intervenção surgiram nas Ciências Sociais com o trabalho de Campbell & Stanley (1968). Embora o termo intervenção tenha sido introduzido pela primeira vez por Glass (1972), foi apenas em 1975 que Box e Tiao desenvolveram a teoria de análise de intervenção para estudar mudanças estruturais em séries temporais (Box & Tiao, 1975).

Ao mesmo tempo, Box & Jenkins (1976) introduziram os modelos de função de transferência (FT) no contexto dos processos Autoregressivos Integrados de Média Móveis (ARIMA). Os modelos de função de transferência foram criados para mensurar a relação entre uma série resposta e uma ou mais séries explicativas. Por exemplo, no caso de uma série resposta y_t e uma série x_t explicativa, a função de transferência relaciona as variáveis através de um filtro linear da forma

$$y_t = \vartheta(B)x_t + \varepsilon_t$$

onde $\vartheta(B) = \sum_{j=-\infty}^{\infty} \vartheta_j B^j$ e B é o operador de retardo, $B^k y_t = y_{t-k}$.

Se a série explicativa x_t é uma função determinística que descreve as mudanças estruturais em y_t , a FT pode ser usada para modelar quebras estruturais seguindo a mesma idéia de Box & Tiao (1975).

Devido ao fato que os modelos ARIMA podem ser escritos na forma de espaço de estados, a introdução da idéia das funções de transferência nos modelos estruturais é quase natural (Harvey, 1989; West & Harrison, 1997). A flexibilidade da representação de espaço de estados permite a inserção de covariáveis nas equações das observações e de transição. Alguns trabalhos sobre esse assunto incluem Penzer (2007), de Jong & Penzer (1998), Harvey & Koopman (1992), Salvador & Gargallo (2004), Ravines, Schmidt & Migon (2008) e Alves, Gamerman & Ferreira (2009).

Neste trabalho, dois tipos de procedimentos de intervenção são descritos: as funções pulso e degrau. O parâmetro do modelo associado com a intervenção assume duas formas: na primeira, ele é fixo no tempo e, na segunda, ele varia no tempo. Inferências sobre os parâmetros da FT e variâncias dos erros nas equações da observação e transição são feitas usando as abordagens clássica e Bayesiana. Trabalhos utilizando inferência clássica para modelar os parâmetros associados à intervenção variando no tempo não são de conhecimento dos autores. Intervalos de confiança sob o paradigma clássico são construídos usando o *bootstrap* sob a abordagem paramétrica. Na abordagem Bayesiana, os métodos de Markov Chain Monte Carlo (MCMC) são utilizados para obter as estimativas pontuais e os intervalos de credibilidade.

4.1 ANÁLISE DE INTERVENÇÃO EM MODELOS ESTRUTURAIS

Se uma mudança de nível é observada na série, seu efeito pode ser modelado inserindo um componente E_t na equação do estado. Esse componente E_t é escrito como o filtro da função de transferência de Box, Jenkins & Reinsel (1994),

$$E_t = \frac{\beta(B)B^b}{\rho(B)}x_t = \vartheta(B)x_t = \sum_{j=0}^{\infty} \vartheta_j B^j x_t \quad (4.1)$$

onde x_t é uma variável exógena, $\beta(B) = \beta_0 - \beta_1 B - \dots - \beta_s B^s$, $\rho(B) = \rho_0 - \rho_1 B - \dots - \rho_r B^r$ e b é o parâmetro de atraso. Os coeficientes ϑ_j no modelo de função de transferência, denotado aqui por FT(r, b, s), são denominados função-resposta impulso. Mais detalhes podem ser vistos em Wei (1990).

Neste capítulo, a variável exógena x_t em (4.1) representará a intervenção. Na prática, há muitas possibilidades para a ocorrência de intervenção. Por exemplo, o impacto de um evento externo pode ser sentido b períodos após a intervenção, com um efeito apenas no momento da ocorrência. Neste caso, o modelo de função de transferência é um FT(0,0, b):

$$E_t = \beta_0 B^b x_t = \beta_0 x_{t-b}.$$

Ou pode acontecer que o impacto é sentido no momento da intervenção, mas a resposta é gradual. Então, um modelo apropriado é um FT(1,0,0),

$$E_t = \frac{\beta_0 B^0}{\rho_0 - \rho_1 B} x_t \Rightarrow (\rho_0 - \rho_1 B)E_t = \beta_0 x_t \Rightarrow E_t = \rho E_{t-1} + \beta x_t, \quad (4.2)$$

onde $\rho = \rho_1/\rho_0$ e $\beta = \beta_0/\rho_0$.

Box & Tiao (1975) definem dois tipos comuns de variáveis de intervenção, função pulso e degrau, que podem ser representadas pelas seguintes variáveis *dummy*:

(1) Função degrau: Se a intervenção ocorre em algum tempo T fixado e o efeito persiste após a intervenção:

$$S_t^T = \begin{cases} 0, & t < T \\ 1, & t \geq T \end{cases} ;$$

(2) Função pulso: Se a intervenção ocorre em algum tempo T fixado e tem efeito apenas neste instante:

$$I_t^T = \begin{cases} 1, & t = T \\ 0, & t \neq T \end{cases}$$

A Figura 4.1 apresenta o comportamento do bloco estrutural E_t para séries de tamanho $n = 100$ com uma mudança de nível no tempo $T = 50$, sob o Modelo FT(1,0,0), para funções pulso e degrau. Para a função pulso nota-se que, quando aumenta-se o valor de ρ , a série tem um retorno muito lento para o seu nível médio anterior à intervenção. Para a função degrau, se ρ aumenta, a série leva mais tempo para atingir um novo nível médio.

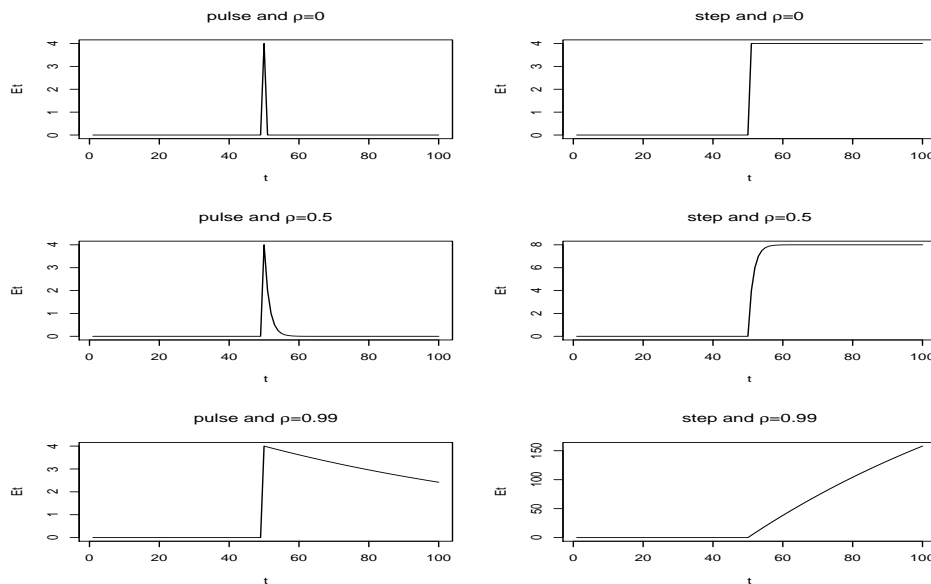


Figura 4.1: Efeito da intervenção das funções pulso e degrau.

Neste trabalho, FT(1,0,0) será considerada com dois tipos de função-resposta impulso. Na primeira (Modelo 1) o fator de ganho β na equação (4.2) é fixado no tempo, enquanto que o Modelo 2 permite β variar no tempo, ou seja, β_t . Esse último modelo é conhecido como um modelo de função de transferência com fator de ganho dinâmico, ver Alves, Gamerman &

Ferreira (2009).

O modelo estrutural mais simples - o modelo de nível local (MNL), descrito no Capítulo 2 - é considerado com um componente de intervenção. As estruturas dos Modelos 1 e 2 sob a abordagem de nível local são descritas abaixo, com exemplos de séries geradas desses modelos.

4.1.1 MODELO 1

O MNL com um componente de intervenção, assumindo o coeficiente β fixo no tempo tem a seguinte estrutura:

$$\begin{cases} y_t = \mu_t + E_t + \varepsilon_t; & \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \\ \mu_t = \mu_{t-1} + \eta_t; & \eta_t \sim N(0, \sigma_\eta^2) \\ E_t = \rho E_{t-1} + \beta x_t \end{cases} \quad (4.3)$$

onde $t = 1, 2, \dots, n$, ε_t e η_t são independentes e $0 \leq \rho \leq 1$. ρ é definido entre 0 e 1 para que o comportamento da função-resposta impulso tenha um decaimento geométrico. Por essa razão, ρ é normalmente chamado de persistência ou memória do efeito. A resposta média da série é dada por $\alpha_t = \mu_t + E_t$.

O modelo em (4.3) pode ser escrito na forma de espaço de estados com matrizes dadas por $\mathbf{Z}'_t = [1 \quad 1]$, $\mathbf{R}_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{T}_t = \begin{bmatrix} 1 & 0 \\ 0 & \rho \end{bmatrix}$, $\boldsymbol{\eta}_t = \begin{bmatrix} \eta_t \\ 0 \end{bmatrix}$, $h_t = [\sigma_\varepsilon^2]$, $\mathbf{Q}_t = \begin{bmatrix} \sigma_\eta^2 & 0 \\ 0 & 0 \end{bmatrix}$, $d_t = 0$, $\mathbf{c}_t = \begin{bmatrix} 0 \\ \beta x_t \end{bmatrix}$ e $\boldsymbol{\alpha}_t = \begin{bmatrix} \mu_t \\ E_t \end{bmatrix}$.

A Figura 4.2 mostra algumas séries simuladas sob o Modelo 1. Como esperado, o comportamento da série é muito similar ao comportamento do bloco E_t mostrado na Figura 4.1. Para a função pulso, pode-se notar que pequenos valores de ρ causam na série um efeito similar à presença de um *outlier* no ponto da intervenção. Para grandes valores de ρ , a série apresenta um salto no ponto da intervenção, mas com um retorno gradual ao valor médio. Para a função degrau, quando $\rho = 0.00$ ou 0.50 , há um salto repentino no tempo da intervenção, com uma mudança de nível, e o restante da série permanece no novo nível. Quando $\rho = 0.99$, há uma gradual mudança de nível, e a série leva mais tempo para atingir um novo nível.

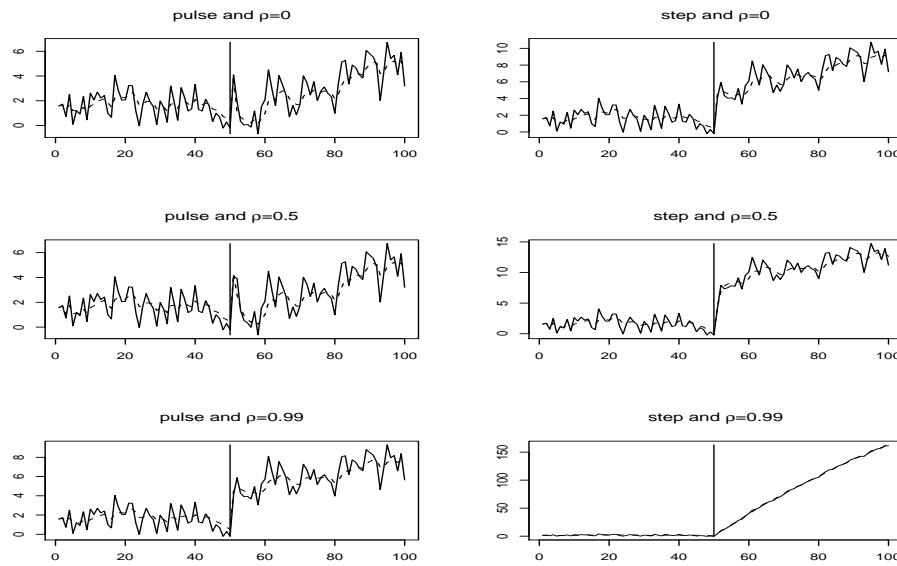


Figura 4.2: Exemplos das funções pulso e degrau para o Modelo 1, com $n = 100$, $\beta = 4$ e $\rho = 0, 0.50$ e 0.99 . A linha vertical indica o instante da intervenção, $T = 50$, a linha contínua indica a série temporal, y_t , e linha tracejada indicada a resposta média α_t .

4.1.2 MODELO 2

O modelo de nível local com um fator de ganho dinâmico (Alves, Gamerman & Ferreira, 2009) sendo dado por

$$\begin{cases} y_t = \mu_t + E_t + \varepsilon_t; & \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \\ \mu_t = \mu_{t-1} + \eta_t; & \eta_t \sim N(0, \sigma_\eta^2) \\ E_t = \rho E_{t-1} + \beta_{t-1} x_t \\ \beta_t = \beta_{t-1} + \xi_t; & \xi_t \sim N(0, \sigma_\xi^2) \end{cases};$$

onde $t = 1, 2, \dots, n$, ε_t , η_t e ξ_t são conjuntamente independentes.

Ao contrário do Modelo 1, no Modelo 2 o parâmetro β_t é estocástico e é obtido dinamicamente através do tempo por alguma lei estocástica (aqui um passeio aleatório).

O Modelo 2 pode ser escrito na forma de espaço de estados com matrizes dadas por

$$\mathbf{Z}'_t = [1 \ 1 \ 0], \mathbf{R}_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{T}_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho & x_t \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{\eta}_t = \begin{bmatrix} \eta_t \\ 0 \\ \xi_t \end{bmatrix}, h_t = [\sigma_\varepsilon^2], \mathbf{Q}_t = \begin{bmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_\xi^2 \end{bmatrix}, d_t = 0, \mathbf{c}_t = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ e } \boldsymbol{\alpha}_t = \begin{bmatrix} \mu_t \\ E_t \\ \beta_t \end{bmatrix}.$$

A Figura 4.3 apresenta algumas séries simuladas sob o Modelo 2. O comportamento é muito similar ao do Modelo 1, mas neste caso, o β dinâmico causa mais ruído na série. Se uma função pulso é usada, o Modelo 2 não é adequado. Por essa razão, apenas a função degrau será usada no Modelo 2.

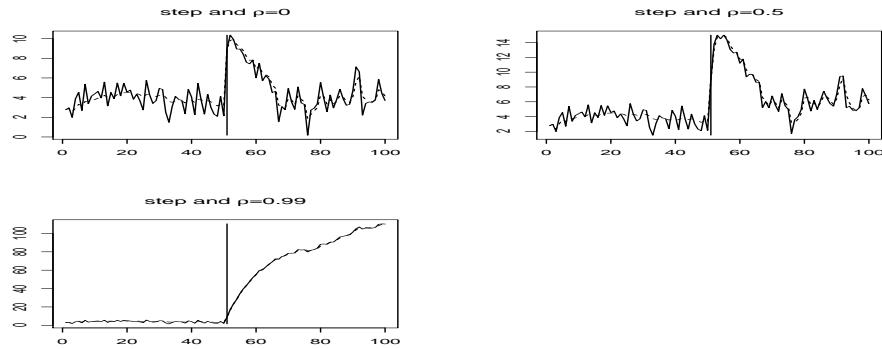


Figura 4.3: Exemplos da função degrau para o Modelo 2, com $n = 100$ e $\rho = 0.00, 0.50$ e 0.99 . A linha vertical indica o instante da intervenção, $T = 50$, a linha contínua indica a série, y_t , e a linha tracejada indica a resposta média α_t .

4.1.3 ESTIMAÇÃO

A estimação dos parâmetros dos Modelos 1 e 2 pode ser feita usando métodos clássicos e/ou métodos Bayesianos. Em ambos casos, o algoritmo do filtro de Kalman (Kalman, 1960) será utilizado para estimar os componentes do vetor de espaço de estados, α_t , dadas as observações $\mathbf{Y}_t = \{y_1, \dots, y_t\}$. Na Seção 2.2.4 foi visto que o estimador linear de α_t e sua matriz de variância são dados respectivamente por

$$\mathbf{a}_t = E(\alpha_t | \mathbf{Y}_t) \quad \text{e} \quad \mathbf{P}_t = \text{Var}(\alpha_t | \mathbf{Y}_t).$$

O filtro de Kalman pode ser usado para calcular o erro de previsão, v_t , e sua variância

$$v_t = y_t - E(y_t | \mathbf{Y}_{t-1}) = y_t - \tilde{y}_{t|t-1} \quad \text{e} \quad F_t = \mathbf{Z}_t \mathbf{P}_{t|t-1} \mathbf{Z}_t' + h_t,$$

onde $\mathbf{P}_{t|t-1} = \text{Var}(\alpha_t | \mathbf{Y}_{t-1})$.

Assumindo que os erros ε_t e η_t são normalmente distribuídos e o vetor paramétrico é dado por $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_p)'$ - que inclui os hiperparâmetros (ou seja, as variâncias dos termos de erros) e também parâmetros associados com variáveis exógenas ou componentes não-observáveis tais como o ρ e o β , a função de verossimilhança $L(\boldsymbol{\varphi}; \mathbf{Y}_n)$ pode ser calculada depois de integrar fora as variáveis do vetor de estado. Para uma série univariada de tamanho n , o logaritmo da

função de verossimilhança (como definido em 2.9) é dado por

$$\ln L(\boldsymbol{\varphi}; \mathbf{Y}_n) = \ln \prod_{t=1}^n p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^n \ln |F_t| - \frac{1}{2} \sum_{t=1}^n \mathbf{v}_t' F_t^{-1} \mathbf{v}_t \quad (4.4)$$

onde $p = 4$ se $\boldsymbol{\varphi} = (\sigma_\varepsilon^2, \sigma_\eta^2, \rho, \beta)'$ para o Modelo 1 ou $\boldsymbol{\varphi} = (\sigma_\varepsilon^2, \sigma_\eta^2, \sigma_\xi^2, \rho)'$ para o Modelo 2.

Para obter estimadores para $\boldsymbol{\varphi}$ usando inferência clássica, deve-se maximizar o logaritmo da função de verossimilhança, dado na equação (4.4), com respeito a $\boldsymbol{\varphi}$. Como visto na Seção 2.3.1, a função de verossimilhança é uma função não-linear de $\boldsymbol{\varphi}$, portanto o algoritmo de otimização BFGS será empregado (ver Franco, Santos, Ribeiro & Cruz (2008) para maiores detalhes). Inferências para os parâmetros sob a abordagem clássica serão feitas usando o *bootstrap* paramétrico (Efron, 1979), descrito na Seção 2.5.2.

Neste capítulo, a distribuição *a priori* uniforme, dada por $\pi(\boldsymbol{\varphi}) \propto c$, onde $c \in \mathfrak{R}$, é usada para permitir uma comparação justa com a abordagem clássica. Como, neste caso, a distribuição *a posteriori* não tem uma forma fechada, métodos Markov chain Monte Carlo (MCMC) são usado para obter estimativas pontuais e intervalares para os parâmetros, ver Gamerman & Lopes (2006) para detalhes. Neste capítulo, uma versão híbrida do algoritmo de Metropolis-Hastings é adotada na qual os componentes de $\boldsymbol{\varphi}$ são atualizados separadamente, com diferentes densidades de proposição. Reis, Salazar & Gamerman (2006) obtêm evidências empíricas substanciais em favor deste esquema de amostragem frente a outros esquemas.

4.2 RESULTADOS DE SIMULAÇÃO

Os procedimentos descritos na Seção 4.1 são agora investigados através de experimentos Monte Carlo (MC), implementados na linguagem Ox (Doornik, 1999). Séries com intervenções pulso e degrau para o modelo de nível local foram simuladas, baseadas no Modelo 1 e Modelo 2 (só intervenção do tipo degrau). Os desempenhos do estimador de máxima verossimilhança (EMV), *bootstrap* (BOOT) e estimadores de Bayes - média (Média), mediana (Med) e moda (Moda) - foram avaliadas para séries de tamanho $n = 100$ e $\rho = 0, 0.50$ and 0.99 . Em todos os casos, $\sigma_\varepsilon^2 = 1.00$, $\sigma_\eta^2 = 0.10$, $\sigma_\xi^2 = 0.50$ (para o Modelo 2) e $\beta = 4$ (para o Modelo 1). Para os estimadores de Bayes, duas cadeias com 2000 amostras foram geradas das quais as 1000 primeiras foram excluídas. O número de MC e *bootstrap* foram fixados em 500. O nível e a probabilidade dos intervalos de confiança e de credibilidade, respectivamente, foram fixados em 0.95.

4.2.1 MODELO 1

As Figuras 4.4 a 4.9 apresentam o vício e o EQM das 500 replicações MC para o Modelo 1 com funções pulso e degrau.

Para a função pulso (Figuras 4.4, 4.5 e 4.6), a primeira conclusão que pode ser obtida é que ρ é sempre subestimado, exceto no caso $\rho = 0$. Se ρ é grande ($\rho = 0.99$), o melhor estimador é a Moda, com vício e EQM pequenos, mas para $\rho = 0.50$ os melhores estimadores são a Mediana e Média e para $\rho = 0$ o melhor é o EMV. Os outros parâmetros apresentam desempenhos muito satisfatórios para todos os métodos. Os melhores estimadores para σ_η^2 são a Média e a Moda, independente do valor de ρ . Para σ_ε^2 , todos os procedimentos apresentam o mesmo comportamento. No caso do parâmetro β , se ρ é grande, todos os procedimentos mostram aproximadamente o mesmo desempenho, mas a Média e Mediana são ligeiramente melhor para $\rho = 0.50$ and 0.00 . Com respeito ao *bootstrap*, pode-se ver que ele imita bem o comportamento do EMV, o que permite a essa técnica ser usada na abordagem clássica para construir os intervalos de confiança para os parâmetros.

Para a função degrau (Figuras 4.7, 4.8 e 4.9), ρ é subestimado nos casos que $\rho = 0.99$ e $\rho = 0.50$ pelos estimadores Bayesianos. Todos os procedimentos mostram um excelente desempenho para estimar ρ quando esse parâmetro assume um valor grande, mas para $\rho = 0.50$ e 0.00 , o EMV e a Moda *a posteriori* são um pouco melhores. Os desempenhos dos métodos para estimar os outros parâmetros com a função degrau são muito similares à função pulso, exceto no caso $\rho = 0.99$. Parece que, neste caso, o EMV não é capaz de estimar σ_η^2 , σ_ε^2 e β muito precisamente, apresentando vício e EQM maiores que os estimadores Bayesianos. Novamente, o comportamento do *bootstrap* é muito similar ao do EMV.

As Tabelas 4.1 e 4.2 apresentam os intervalos de confiança *bootstrap* e de credibilidade para o Modelo 1 com funções pulso e degrau, respectivamente. Nessas tabelas, além dos estimadores, a previsão um passo à frente é também incluída. Quando $\rho = 0.99$, para todos os parâmetros considerados, ambos intervalos têm taxas de cobertura bem próximas do nível nominal 0.95, exceto o intervalo de confiança *bootstrap* para ρ . Por outro lado, o desempenho do intervalo de credibilidade para função degrau é bem melhor que o intervalo *bootstrap*, exceto para previsão um passo à frente, onde ambos estão longe do nível de 0.95 assumido. Se $\rho = 0.50$ ou 0.00 , para funções pulso e degrau, ambos intervalos mostram aproximadamente o mesmo comportamento para todos parâmetros, com taxa de cobertura próxima de 0.95. As únicas exceções são o intervalo *bootstrap* para σ_η^2 , o qual é ligeiramente pior do que o intervalo de credibilidade, e os intervalos para ρ na função degrau, onde ambos métodos apresentam taxas de cobertura abaixo de 0.90. Deve-se ressaltar que os intervalos para $\rho = 0$ não são con-

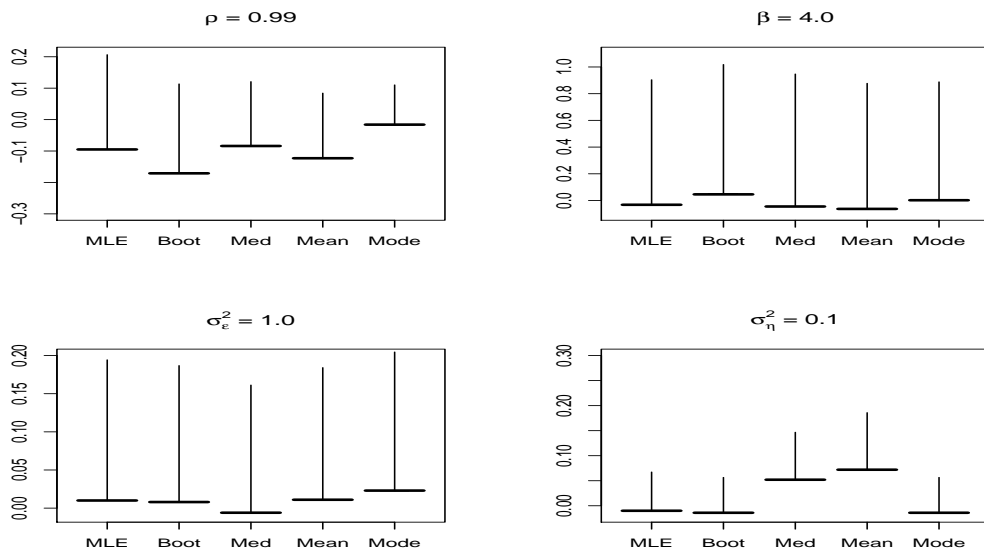


Figura 4.4: Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função pulso e $\rho = 0.99$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.

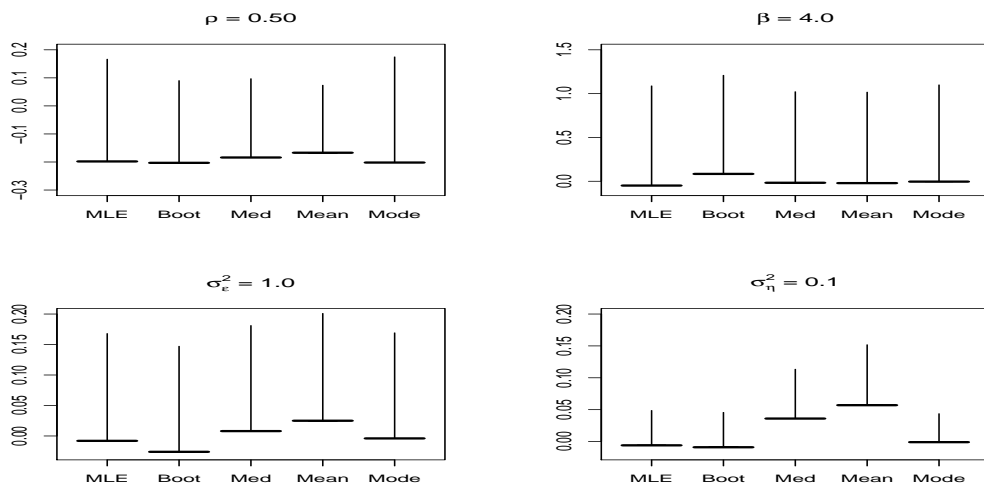


Figura 4.5: Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função pulso e $\rho = 0.50$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.

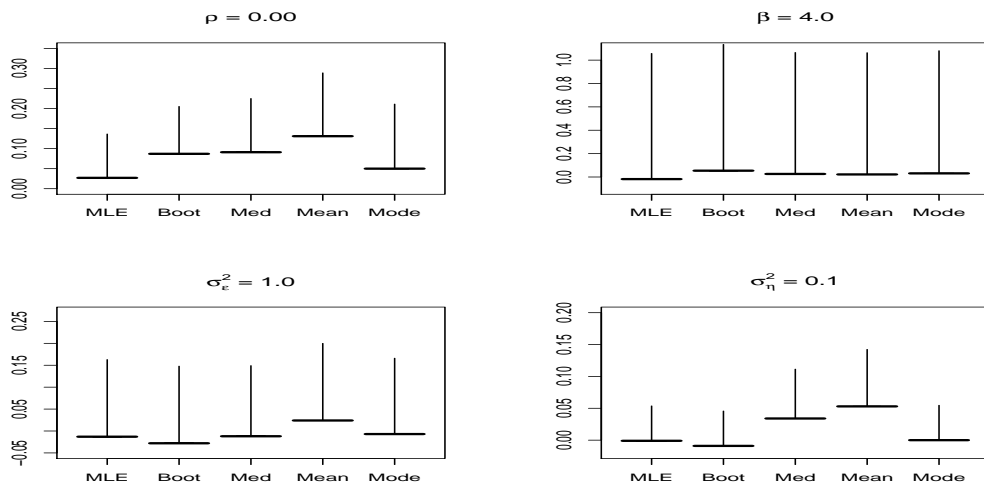


Figura 4.6: Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função pulso e $\rho = 0.00$. As linhas horizontal e vertical indicam o vício e a raiz quadrada do EQM, respectivamente.

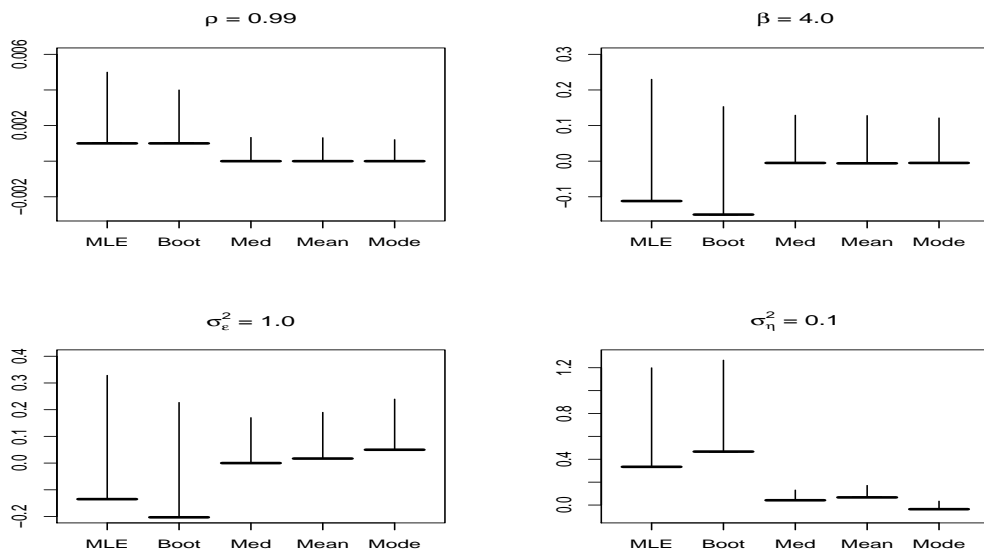


Figura 4.7: Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função degrau e $\rho = 0.99$. As linhas horizontal e vertical indicam o vício e a raiz quadrada do EQM, respectivamente.

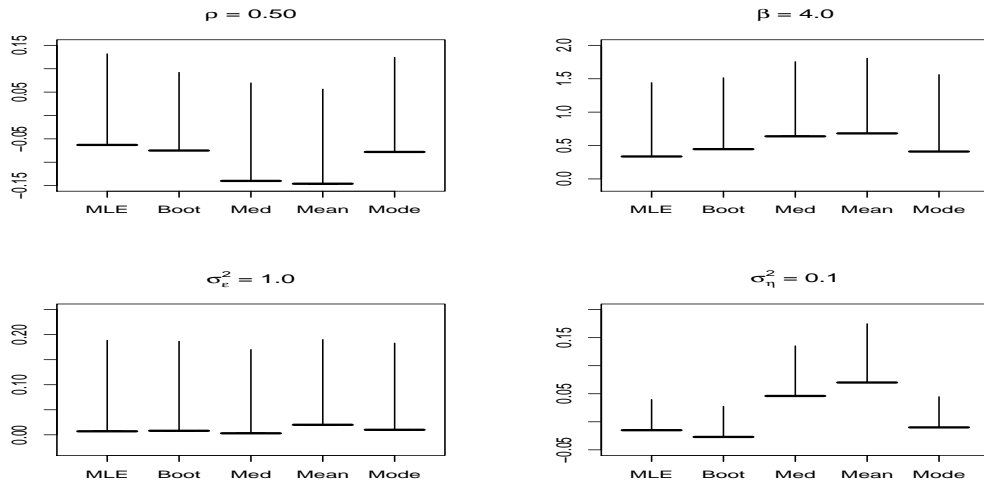


Figura 4.8: Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função degrau e $\rho = 0.50$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.

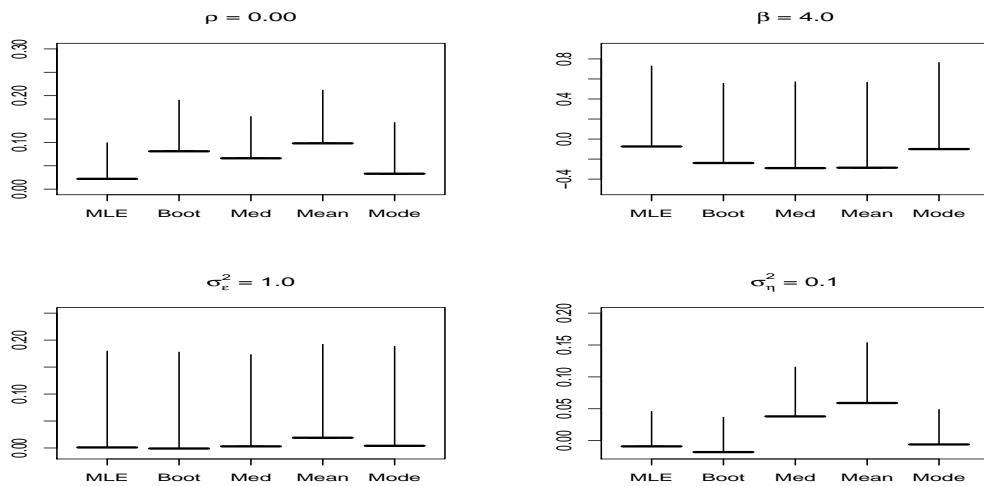


Figura 4.9: Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 1 com função degrau e $\rho = 0$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.

struídos, pois os intervalos não cobrem o valor verdadeiro deste parâmetro, estando no limite do espaço paramétrico.

Tabela 4.1: Intervalos de confiança e de credibilidade para o Modelo 1 com função pulso.

	valor verdadeiro	Intervalo de Confiança			Intervalo de Credibilidade		
		limites médios	amplitude	cobertura	limites médios	amplitude	cobertura
ρ	0.99	[0.177; 0.983]	0.806	0.87	[0.536; 0.994]	0.458	0.95
β	4.00	[2.167; 5.681]	3.514	0.91	[1.921; 5.845]	3.924	0.97
σ_ε^2	1.00	[0.687; 1.371]	0.684	0.94	[0.683; 1.427]	0.744	0.95
σ_η^2	0.10	[0.007; 0.256]	0.249	0.93	[0.043; 0.449]	0.406	0.93
y_{n+1}		[0.221; 4.590]	4.369	0.93	[0.013; 4.275]	4.262	0.93
ρ	0.50	[0.011; 0.757]	0.746	0.91	[0.040; 0.724]	0.684	0.89
β	4.00	[1.985; 6.144]	4.159	0.92	[1.724; 6.197]	4.473	0.96
σ_ε^2	1.00	[0.668; 1.328]	0.660	0.93	[0.700; 1.445]	0.745	0.95
σ_η^2	0.10	[0.011; 0.213]	0.202	0.87	[0.046; 0.390]	0.344	0.94
y_{n+1}		[-2.179; 1.960]	4.139	0.92	[-2.427; 2.024]	4.451	0.95
ρ	0.00	-	-	-	-	-	-
β	4.00	[1.944; 6.148]	4.204	0.94	[1.802; 6.231]	4.429	0.96
σ_ε^2	1.00	[0.663; 1.326]	0.663	0.93	[0.702; 1.441]	0.739	0.95
σ_η^2	0.10	[0.015; 0.221]	0.206	0.90	[0.046; 0.377]	0.331	0.94
y_{n+1}		[-2.149; 1.944]	4.093	0.92	[-2.424; 2.025]	4.449	0.95

Obs.: Não há um único valor verdadeiro para y_{n+1} , já que cada série tem o seu próprio valor y_{n+1} .

4.2.2 MODELO 2

As Figuras 4.10, 4.11 e 4.12 apresentam o vício e o EQM sobre as 500 replicações MC para o Modelo 2. Como foi explicado na Seção 4.1.2, neste caso o modelo não é capaz de estimar a influência da função pulso. Portanto, apenas os resultados para função degrau são apresentados abaixo.

Pode-se notar que, em geral, o EMV e Moda *a posteriori* são ligeiramente melhores que os outros procedimentos para estimar todos parâmetros, com vício e EQM menores. Mais uma vez, ρ é subestimado, exceto para $\rho = 0$.

Na Tabela 4.3 estão os resultados para os intervalos *bootstrap* e de credibilidade com função degrau. Como no Modelo 1 com função degrau, o desempenho do intervalo de credibilidade é novamente melhor do que a do intervalo *bootstrap*, no caso $\rho = 0.99$. Se $\rho = 0.50$ ou 0, o intervalo de credibilidade mostra, em geral, taxas de cobertura mais próximas de 0.95, exceto para σ_ε^2 , onde o intervalo *bootstrap* é melhor. Novamente, os intervalos para $\rho = 0$ não são construídos.

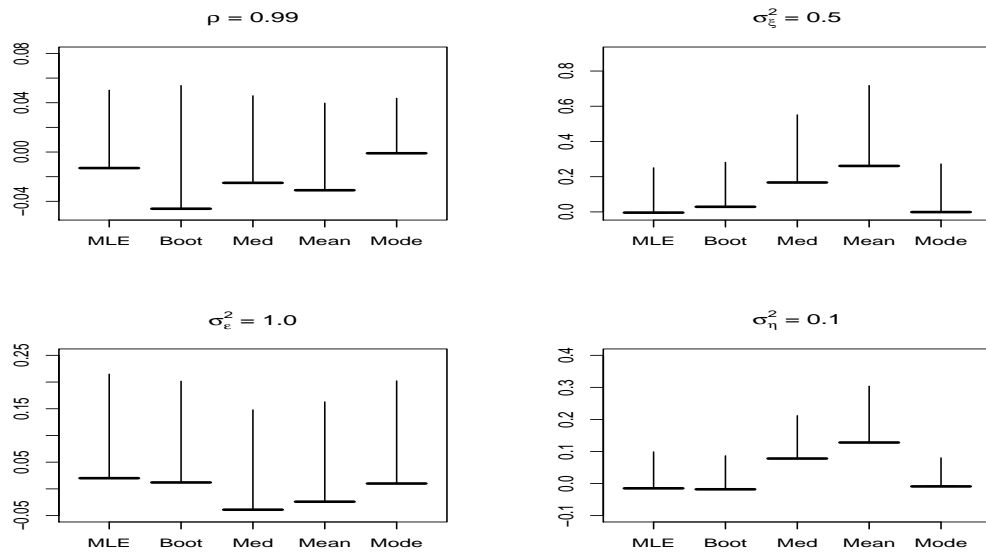


Figura 4.10: Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 2 com função degrau e $\rho = 0.99$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.

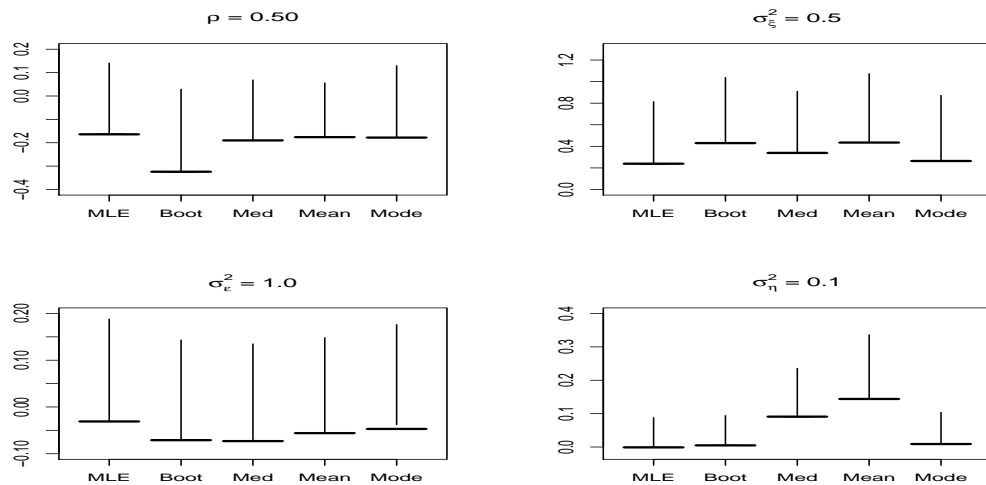


Figura 4.11: Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 2 com função degrau e $\rho = 0.50$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.

Tabela 4.2: Intervalos de confiança e de credibilidade para o Modelo 1 com função degrau.

	valor verdadeiro	Intervalo de Confiança			Intervalo de Credibilidade		
		limites médios	amplitude	cobertura	limites médios	amplitude	cobertura
ρ	0.99	[0.988; 0.998]	0.010	0.75	[0.987; 0.992]	0.005	0.88
β	4.00	[3.165; 4.169]	1.004	0.78	[3.726; 4.257]	0.531	0.90
σ_{ε}^2	1.00	[0.027; 1.255]	1.228	0.79	[0.688; 1.434]	0.746	0.94
σ_{η}^2	0.10	[0.201; 2.238]	2.037	0.81	[0.035; 0.443]	0.407	0.95
y_{n+1}		[155.64; 160.35]	4.71	0.51	[155.45; 159.83]	4.38	0.38
ρ	0.50	[0.139; 0.655]	0.516	0.89	[0.090; 0.601]	0.511	0.81
β	4.00	[2.839; 6.330]	3.491	0.92	[2.964; 6.634]	3.670	0.91
σ_{ε}^2	1.00	[0.695; 1.365]	0.670	0.94	[0.692; 1.441]	0.749	0.96
σ_{η}^2	0.10	[0.005; 0.187]	0.182	0.80	[0.046; 0.436]	0.390	0.94
y_{n+1}		[5.761; 10.052]	4.291	0.93	[5.536; 9.819]	4.283	0.94
ρ	0.00	-	-	-	-	-	-
β	4.00	[2.135; 5.329]	3.194	0.94	[0.729; 5.465]	4.736	0.95
σ_{ε}^2	1.00	[0.686; 1.360]	0.674	0.93	[0.697; 1.434]	0.737	0.95
σ_{η}^2	0.10	[0.007; 0.201]	0.194	0.84	[0.046; 0.395]	0.349	0.93
y_{n+1}		[1.792; 5.977]	4.185	0.92	[1.543; 5.816]	4.273	0.94

Obs.: Não há um único valor verdadeiro para y_{n+1} , já que cada série tem o seu próprio valor y_{n+1} .

4.3 APLICAÇÃO À SÉRIES REAIS

Nesta seção, a metodologia anteriormente descrita é aplicada a duas séries temporais reais. A primeira é o Índice de Preço ao Consumidor Amplo (IPCA) na cidade de Belo Horizonte do período de julho de 1997 a junho de 2008. Essa série parece ter uma intervenção pulso próximo de Outubro de 2002. A segunda série é o retorno da BOVESPA (mercado de ações de São Paulo), no período de janeiro de 1991 a agosto de 2008. Para essa série, a intervenção tem a forma de função degrau, apresentando um salto em torno da metade de 1994.

4.3.1 SÉRIE IPCA

Os dados do Índice de Preço ao Consumidor Amplo são coletados pelo IPEAD - *Fundação Instituto de Pesquisas Econômicas, Administrativas e Contábeis de Minas Gerais*. Esse índice mede a evolução das rendas em famílias que gastam de 1 a 40 salários mínimos por mês. A série IPCA é composta de 132 observações mensais do período de julho, 1997 a junho, 2008.

Da Figura 4.13, a série IPCA parece apresentar um comportamento de um Modelo de Nível Local (MNL), com uma intervenção em torno de outubro, 2002. É importante ressaltar que, no ano de 2002, iniciou o primeiro mandato do presidente Luís Inácio Lula da Silva. Portanto, o MNL foi ajustado com e sem o componente de intervenção, para avaliar a relevância do mesmo. Como foi visto na Seção 4.1.2, o Modelo 2 não é apropriado neste caso, quando a intervenção

Tabela 4.3: Intervalos de confiança e de credibilidade para o Modelo 2.

	valor verdadeiro	Intervalos de Confiança			Intervalos de Credibilidade		
		limites médios	amplitude	cobertura	limites médios	amplitude	cobertura
ρ	0.99	[0.671; 0.998]	0.327	0.98	[0.887; 0.997]	0.110	0.96
σ_{ε}^2	0.50	[0.128; 1.243]	1.115	0.99	[0.255; 1.815]	1.125	0.93
σ_{ε}^2	1.00	[0.656; 1.407]	0.751	0.96	[0.604; 1.436]	0.832	0.97
σ_{η}^2	0.10	[0.001; 0.280]	0.279	0.74	[0.032; 0.722]	0.690	0.95
y_{n+1}		[-27.19; -22.71]	4.48	0.27	[-29.49; -20.53]	6.956	0.81
ρ	0.50	[0.001; 0.608]	0.607	0.67	[0.103; 0.613]	0.510	0.81
σ_{ε}^2	0.50	[0.169; 2.058]	1.889	0.93	[0.240; 2.174]	1.934	0.90
σ_{ε}^2	1.00	[0.562; 1.343]	0.781	0.90	[0.559; 1.420]	0.861	0.95
σ_{η}^2	0.10	[0.003; 0.309]	0.306	0.85	[0.041; 0.762]	0.721	0.92
y_{n+1}		[-3.434; 0.855]	4.289	0.77	[-3.258; 2.660]	5.920	0.84
ρ	0.00	-	-	-	-	-	-
σ_{ε}^2	0.50	[0.052; 1.141]	1.089	0.92	[0.102; 1.526]	1.424	0.97
σ_{ε}^2	1.00	[0.636; 1.417]	0.781	0.93	[0.623; 1.471]	0.848	0.97
σ_{η}^2	0.10	[0.003; 0.270]	0.267	0.85	[0.037; 0.588]	0.551	0.93
y_{n+1}		[-2.759; 1.368]	4.127	0.79	[-2.662; 3.329]	5.991	0.87

Obs.: Não há um único valor verdadeiro para y_{n+1} , já que cada série tem o seu próprio valor y_{n+1} .

é da forma pulso.

Da Tabela 4.4, é verificado que o Modelo 1, em ambas estimações clássica e Bayesiana, apresentam menor AIC do que o MNL, confirmando, então, que a intervenção pulso deve ser incluída no modelo. A previsão um passo à frente foi calculada para julho de 2008. O valor verdadeiro é 0.35, e pode ser visto que a predição Bayesiana é mais próxima do valor verdadeiro que a clássica, com menor amplitude do intervalo. Estimativas para os parâmetros são mostradas na Tabela 4.5. Parece que o termo de nível μ_t é constante no tempo, já que o intervalo de credibilidade para σ_{η}^2 é bem concentrado em torno de zero. O valor mais provável de ρ é 0,80. A estimativa da moda *a posteriori* de β é 1.968, que corresponde à magnitude do salto da intervenção, compatível com uma inspeção visual.

Tabela 4.4: AIC e a previsão um passo à frente para a série IPCA.

Modelos	AIC	Previsão (valor verdadeiro=0.350)	Int. Previsão
MNL Clássico	212.650	0.320	[-0.508; 1.456]
MNL Bayesiano	212.600	0.330	[-0.302; 0.962]
Modelo 1 Clássico	198.050	0.419	[-0.407; 1.442]
Modelo 1 Bayesiano	198.101	0.349	[-0.443; 1.430]

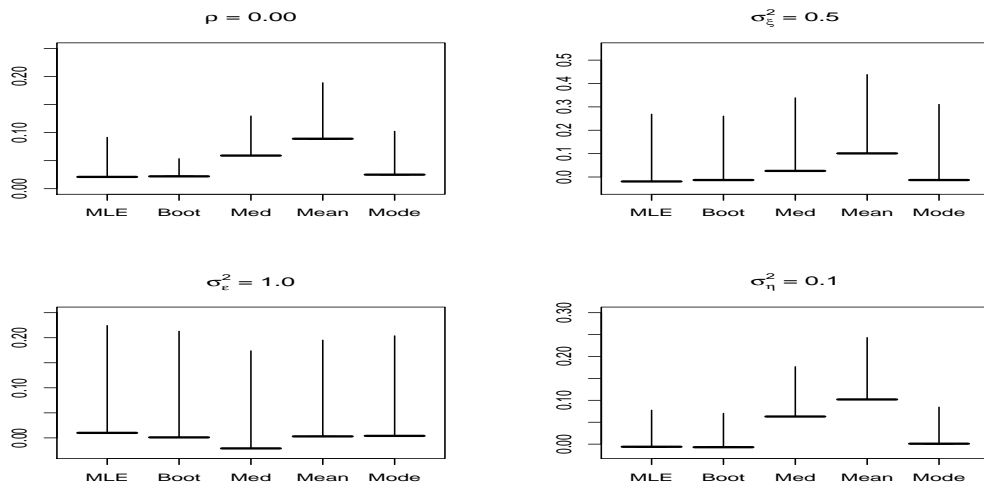


Figura 4.12: Máxima verossimilhança, bootstrap e estimação Bayesiana para o Modelo 2 com função degrau e $\rho = 0$. As linhas horizontal e vertical (o comprimento da linha) indicam o vício e a raiz quadrada do EQM, respectivamente.

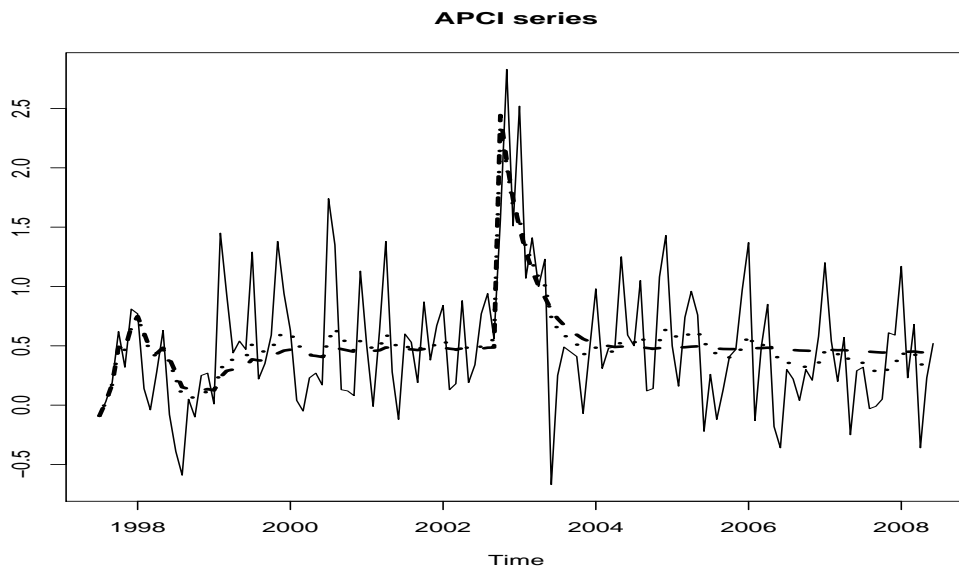


Figura 4.13: A linha contínua indica a série do IPCA, y_t , e as linhas pontilhada e tracejada indicam as estimativas da resposta média do Modelo 1, segundo as abordagens Bayesiana e clássica, respectivamente.

Tabela 4.5: Ajuste do Modelo 1 à série IPCA.

	Inferência Clássica		Inferência Bayesiana	
	EMV	Intervalo	Moda	Intervalo
ρ	0.797	[0.030; 0.883]	0.801	[0.558; 0.884]
β	1.975	[1.411; 2.888]	1.968	[0.851; 2.685]
σ_ε^2	0.225	[0.167; 0.277]	0.222	[0.168; 0.298]
σ_η^2	5.88×10^{-12}	$[2.13 \times 10^{-18}; 1.63 \times 10^{-3}]$	3.52×10^{-5}	$[2.88 \times 10^{-5}; 2.52 \times 10^{-2}]$

4.3.2 SÉRIE IBOVESPA

A série IBOVESPA é uma série mensal do retorno do mercado de ações de São Paulo, no período de janeiro de 1991 a agosto de 2008. Essa série é composta de 212 observações mensais. Da Figura 4.14, parece que a IBOVESPA segue um MNL com um componente de intervenção da forma degrau. A série possui valores acima de zero até a metade de 1994 e próximos de zero após esse período. Uma razão para tal comportamento é a introdução do Plano Real que mudou a moeda no Brasil, implementado em julho de 1994. A fim de comparar os procedimentos, o MNL e os Modelos 1 e 2 com uma intervenção da forma degrau foram ajustados à série IBOVESPA.

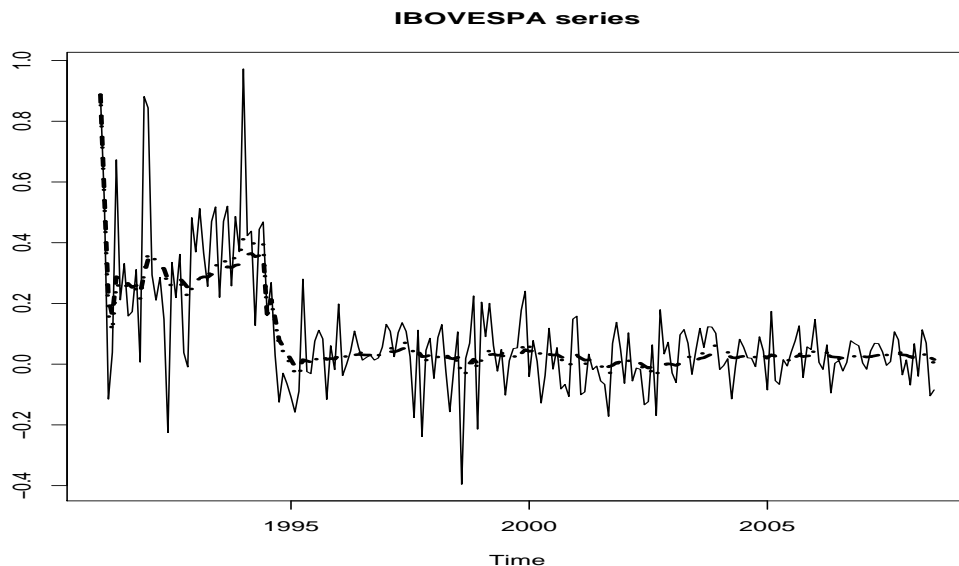


Figura 4.14: A linha contínua indica a série do retorno do IBOVESPA, y_t , e as linhas pontilhada e tracejada indicam as respostas médias do Modelo 2 Clássico e do Modelo 2 Bayesiano, respectivamente.

A Tabela 4.6 mostra o AIC e a previsão um passo à frente usando os três modelos mencionados acima. Pode-se ver que o Modelo 2, em ambas abordagens clássica e Bayesiana, apresenta um menor AIC do que o MNL e o Modelo 1, mas o procedimento Bayesiano fornece uma previsão mais próxima do valor real -0.064 que os outros procedimentos clássicos. As

Tabela 4.6: AIC e a previsão um passo à frente para a série IBOVESPA.

Modelos	AIC	previsão (valor verdadeiro=-0.064)	Intervalo previsão
MNL Clássico	-174.389	-0.020	[-0.291; 0.229]
MNL Bayesiano	-174.042	-0.009	[-0.314; 0.294]
Modelo 1 Clássico	-212.488	0.020	[-0.255; 0.296]
Modelo 1 Bayesiano	-212.050	0.003	[-0.229; 0.348]
Modelo 2 Clássico	-255.019	0.010	[-0.291; 0.306]
Modelo 2 Bayesiano	-255.009	-0.016	[-0.195; 1.001]

Tabela 4.7: Ajuste do Modelo 2 à série IBOVESPA.

	Inferência Clássica		Inferência Bayesiana	
	EMV	Intervalo	Moda	Intervalo
ρ	0	[0; 0.022]	0.002	[0.001; 0.184]
σ_{ξ}^2	0.079	[0.042; 0.123]	0.090	[0.046; 0.164]
σ_{ε}^2	0.010	[0.010; 0.012]	0.010	[0.008; 0.013]
σ_{η}^2	2.30×10^{-13}	$[3.69 \times 10^{-21}; 3.23 \times 10^{-5}]$	1.086×10^{-6}	$[1.815 \times 10^{-7}; 2.427 \times 10^{-4}]$

estimativas dos parâmetros são mostrados na Tabela 4.7. Mais uma vez, parece que o termo de nível μ_t é constante no tempo, como pode ser verificado pelo intervalo de credibilidade para σ_{η}^2 . O valor de ρ é muito pequeno, que significa que o modelo não guarda uma memória dos instantes anteriores.

5 **MODELOS ESTRUTURAIS NÃO-GAUSSIANOS**

Na literatura, há vários modelos que são construídos baseados nas suposições de normalidade, homocedasticidade e independência dos erros, porém nem sempre é possível satisfazer essas suposições.

Uma proposta inovadora foi apresentada por Nelder & Wedderburn (1972), que propuseram *os modelos lineares generalizados* (MLGs). A idéia básica desses modelos consiste em abrir o leque de opções para a distribuição da variável-resposta, permitindo que a mesma pertença à família exponencial de distribuições, o que inclusive traz um ganho na questão de interpretação do modelo. A função de ligação dos dados faz o papel de relacionar a média dos dados ao preditor linear, segundo Nelder & Wedderburn (1972) e Dobson (2002).

Neste trabalho, a *família gama* (FG) de distribuições é proposta. De uma maneira geral, define-se que uma variável aleatória Y possui uma distribuição na FG, se a sua distribuição é escrita na forma:

$$p(y|\mu, \boldsymbol{\varphi}) = a(y, \boldsymbol{\varphi})\mu^{b(y, \boldsymbol{\varphi})} \exp(-\mu c(y, \boldsymbol{\varphi})), \quad (5.1)$$

onde $y \in H(\boldsymbol{\varphi}) \subset \mathfrak{R}$ e 0 , caso contrário. O termo $\mu = g(x'\boldsymbol{\beta})$ é conhecido função de ligação, onde x' é um vetor de covariáveis e $\boldsymbol{\beta}$ os parâmetros associados às mesmas. As funções $a(\cdot, \cdot)$, $b(\cdot, \cdot)$, $c(\cdot, \cdot)$ e $H(\cdot)$ são tais que $p(y|\mu, \boldsymbol{\varphi}) \geq 0$ e que a integral de Lebesgue $\int p(y|\mu, \boldsymbol{\varphi}) dy = 1$. Caso $b(y, \boldsymbol{\varphi}) = b(\boldsymbol{\varphi})$ ou $c(y, \boldsymbol{\varphi}) = c(\boldsymbol{\varphi})$ e $H(\boldsymbol{\varphi})$ seja uma função constante (que não dependa de $\boldsymbol{\varphi}$), a família gama de distribuições torna-se um caso especial da família exponencial de distribuições.

No contexto de séries temporais, a estrutura de correlação das observações não pode ser desprezada. Portanto, várias metodologias foram propostas, levando em conta essa estrutura.

Há vários trabalhos na literatura propondo modelos que levam em consideração a correlação das observações e o fato da série temporal ser não-gaussiana. Trabalhos abordando esse tópico incluem Cox (1981), Kaufmann (1987), Kitagawa (1987), Shephard & Pitt (1997), Smith &

Miller (1986), Harvey & Fernandes (1989) e Durbin & Koopman (2000), entre outros. Uma estrutura mais geral, denominada por *Modelos Lineares Dinâmicos Generalizados* (MLDG), foi proposta por West, Harrison & Migon (1985), despertando um imenso interesse devido à grande aplicabilidade dos mesmos em diversas áreas do conhecimento. Prova disso é o grande número de trabalhos publicados sobre esses modelos. Citam-se ainda os trabalhos de Gamerman & West (1987), Grunwald, Raftery & Guttorp (1993), Fahrmeir (1987), Fruhwirth-Schnatter (1994), Lindsey & Lambert (1995), Gamerman (1991, 1998), Chiogna & Gaetan (2002), Hemming & Shaw (2002) e Godolphin & Triantafyllopoulos (2006).

O objetivo principal deste capítulo é propor uma *Família Gama de Modelos Dinâmicos* (FGMD), generalizando um resultado apresentado por Smith & Miller (1986) - que propuseram um modelo dinâmico para série temporal exponencial e para séries que sejam transformações um-a-um da mesma, considerando uma equação de evolução exata - para qualquer série temporal com distribuição pertencente a essa família.

Nos modelos pertencentes à FGMD, faz-se necessário fazer inferências sobre os parâmetros. Neste estudo, a abordagem clássica, utilizando o estimador de máxima verossimilhança (EMV), e a abordagem Bayesiana, utilizando métodos MCMC para obter os estimadores Bayesianos (EB), são consideradas no processo de inferência. Além disso, os intervalos de confiança assintótico e de credibilidade são construídos para os parâmetros. Assim, um segundo objetivo é comparar a eficiência dos métodos clássicos e Bayesianos para se fazer inferência sobre os parâmetros de um caso particular da FGMD, a saber, série temporal com distribuição Poisson.

Também, como ilustração da metodologia apresentada, modelos na FGMD são ajustados à series temporais reais.

5.1 FAMÍLIA GAMA DE MODELOS DINÂMICOS

Smith & Miller (1986) e Harvey & Fernandes (1989) apresentaram casos particulares de modelos dinâmicos de resposta não-gaussiana. Neste trabalho, a partir desses casos é feita uma generalização, isto é, uma família mais ampla - que os engloba - é introduzida, denominada *família gama de modelos dinâmicos*. Uma vantagem desses modelos frente aos MLDG é que a equação de evolução é exata e não aproximada. Por outro lado, uma desvantagem é que é complicado inserir outros componentes estocásticos no mesmo, como por exemplo, os componentes de tendência e sazonalidade. Entretanto, o efeito desses componentes determinísticos podem ser captados no modelo através de covariáveis.

5.1.1 DEFINIÇÃO

De uma maneira geral, define-se que a série temporal $\{y_t\}$ possui uma distribuição na família gama de modelos dinâmicos, se a sua distribuição é escrita na forma:

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = a(y_t, \boldsymbol{\varphi})\mu_t^{b(y_t, \boldsymbol{\varphi})} \exp(-\mu_t c(y_t, \boldsymbol{\varphi})), \quad (5.2)$$

se $y_t \in H(\boldsymbol{\varphi}) \subset \mathfrak{R}$ e $p(y_t|\mu_t, \boldsymbol{\varphi}) = 0$, caso contrário.

As funções $a(\cdot)$, $b(\cdot)$, $c(\cdot)$ e $H(\cdot)$ são tais que $p(y_t|\mu_t, \boldsymbol{\varphi}) \geq 0$ e $\int p(y_t|\mu_t, \boldsymbol{\varphi}) dy_t = 1$.

O modelo é definido da seguinte forma:

1. a função de ligação é $\mu_t = \lambda_t g(x_t' \boldsymbol{\beta})$, onde $g(\cdot)$ é uma função do preditor linear, x_t' é um vetor de covariáveis, $\boldsymbol{\beta}$ é seu coeficiente de regressão (um dos componentes de $\boldsymbol{\varphi}$) e λ_t é dado pela equação de evolução logo abaixo;
2. A equação de evolução é dada por: $\lambda_t = w^{-1} \lambda_{t-1} \zeta_t$, onde $\zeta_t \sim \text{Beta}(w a_{t-1}, (1-w) a_{t-1})$;
3. $\lambda_0 | Y_0 \sim \text{Gama}(a_0, b_0)$, logo, usando a propriedade de escala da distribuição Gama, $\mu_0 | Y_0 \sim \text{Gama}(a_0, b_0 [g(x_0' \boldsymbol{\beta})]^{-1})$.

É importante destacar que se o logaritmo da equação de evolução é tomado, é obtida a seguinte equação $\ln(\lambda_t) = \ln(\lambda_{t-1}) + \ln(\frac{\zeta_t}{w})$, onde $\ln(\frac{\zeta_t}{w}) \in \mathfrak{R}$. Essa equação é similar à usual equação de evolução dada por um passeio aleatório. w varia entre 0 e 1 e também compõe $\boldsymbol{\varphi}$. Em geral, deseja-se modelar a variância ou a média dessa distribuição, que será função de alguns parâmetros invariantes no tempo e do parâmetro de escala μ_t , isto é, $E(y_t|\mu_t, \boldsymbol{\varphi}) = f(\mu_t, \boldsymbol{\varphi})$.

Teorema 1. Se o modelo está definido na forma descrita na Equação (5.2), os seguintes resultados podem ser obtidos:

1. a distribuição *a priori* $\lambda_t | \mathbf{Y}_{t-1}$ segue uma distribuição $\text{Gama}(a_{t|t-1}, b_{t|t-1})$ tais que

$$a_{t|t-1} = w a_{t-1}, \quad (5.3)$$

$$b_{t|t-1} = w b_{t-1}, \quad (5.4)$$

e $0 < w \leq 1$.

2. $\mu_t = \lambda_t g(x_t' \boldsymbol{\beta}) | \mathbf{Y}_{t-1}$, que é $\text{Gama}(a_{t|t-1}^*, b_{t|t-1}^*)$, onde

$$a_{t|t-1}^* = w a_{t-1}, \quad (5.5)$$

$$b_{t|t-1}^* = w b_{t-1} [g(x_t' \boldsymbol{\beta})]^{-1}. \quad (5.6)$$

3. A distribuição *a posteriori* de $\mu_t | \mathbf{Y}_t$ é Gama (a_t^*, b_t^*), onde

$$a_t^* = a_{t|t-1}^* + b(y_t, \boldsymbol{\varphi}), \quad (5.7)$$

$$b_t^* = b_{t|t-1}^* + c(y_t, \boldsymbol{\varphi}). \quad (5.8)$$

4. $\lambda_t = \mu_t [g(x_t' \boldsymbol{\beta})]^{-1} | \mathbf{Y}_t$ tem também distribuição Gama(a_t, b_t) (atualização) em que

$$a_t = a_{t|t-1} + b(y_t, \boldsymbol{\varphi}), \quad (5.9)$$

$$b_t = b_{t|t-1} + c(y_t, \boldsymbol{\varphi}) g(x_t' \boldsymbol{\beta}). \quad (5.10)$$

5. A função de densidade preditiva é dada por

$$p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = \frac{\Gamma(b(y_t, \boldsymbol{\varphi}) + a_{t|t-1}^*) a(y_t, \boldsymbol{\varphi}) (b_{t|t-1}^*)^{a_{t|t-1}^*}}{\Gamma(a_{t|t-1}^*) [c(y_t, \boldsymbol{\varphi}) + b_{t|t-1}^*]^{b(y_t, \boldsymbol{\varphi}) + a_{t|t-1}^*}}, \quad y_t \in H(\boldsymbol{\varphi}). \quad (5.11)$$

$\forall t \in N; t \leq n$ onde n é o tamanho da série temporal.

As respectivas provas dos resultados do Teorema 1 são encontradas no Apêndice.

Exemplo. Considerando a função de ligação logarítmica, tem-se $g(x_t' \boldsymbol{\beta}) = \exp(x_t' \boldsymbol{\beta})$ e os resultados do Teorema 1 podem ser obtidos para esse caso.

5.1.2 PROCEDIMENTO DE INFERÊNCIA

Uma das formas de se fazer inferência clássica sobre o vetor de parâmetros $\boldsymbol{\varphi}$ é através da função de log-verossimilhança cuja forma é dada por

$$\begin{aligned} \ln L(\boldsymbol{\varphi}; \mathbf{Y}_n) &= \ln \prod_{t=\tau+1}^n p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = \sum_{t=\tau+1}^n \ln \Gamma(a_{t|t-1}^* + b(y_t, \boldsymbol{\varphi})) - \ln \Gamma(a_{t|t-1}^*) + \\ & a_{t|t-1}^* \ln b_{t|t-1}^* + \ln(a(y_t, \boldsymbol{\varphi})) - (b(y_t, \boldsymbol{\varphi}) + a_{t|t-1}^*) \ln(c(y_t, \boldsymbol{\varphi}) + b_{t|t-1}^*), \end{aligned} \quad (5.12)$$

onde τ é o instante da primeira observação diferente de zero, $\boldsymbol{\varphi}$ é composto por ω , $\boldsymbol{\beta}$ e por parâmetros específicos do modelo e $\mathbf{Y}_n = (y_1, \dots, y_n)'$. Na Equação (5.12), a definição de τ , como apresentada acima, se deve ao seguinte fato: a distribuição *a priori* Gama, isto é, a distribuição de μ_t inicial tende a se tornar não-informativa quando $a_0, b_0 \rightarrow 0$, entretanto ela é imprópria quando $a_0 = b_0 = 0$. Nota-se que se $a_0, b_0 \rightarrow 0$ e $y_1 = 0$, a distribuição *a posteriori* $p(\mu_t | \mathbf{Y}_t)$ pode ser imprópria, logo as funções de densidades preditivas podem não estar

definidas. Observa-se que

$$\begin{aligned} p(\mu_1|Y_1) &\propto p(Y_1|\mu_1)p(\mu_1|Y_0) = \frac{\exp(-\mu_1 c(y_1, \boldsymbol{\varphi}))}{[a(y_1, \boldsymbol{\varphi})\mu_1^{b(y_1, \boldsymbol{\varphi})}]^{-1}} \left(\frac{\exp\{-wb_0 \exp(-x'_t \boldsymbol{\beta})\mu_1\}\mu_1^{wa_0}}{\Gamma(a_0)(wb_0 \exp(-x'_t \boldsymbol{\beta}))^{-wa_0}} \right) \infty \\ &\propto \exp\left\{-\mu_1(c(y_1, \boldsymbol{\varphi}) + wb_0 \exp(-x'_t \boldsymbol{\beta}))\right\} \mu_1^{(b(y_1, \boldsymbol{\varphi}) + wa_0)} \end{aligned}$$

se $a_0, b_0 \rightarrow 0$ e $0 < w \leq 1$, então $p(\mu_1|Y_1) \sim Gama((b(y_1, \boldsymbol{\varphi}), c(y_1, \boldsymbol{\varphi}))$. Portanto, se $y_1 = 0$ a distribuição *a posteriori* pode ser imprópria. Logo, as funções de densidade preditivas não estão definidas e, conseqüentemente, a função de log-verossimilhança também pode não estar definida até a primeira observação $y_t \neq 0$. Uma distribuição própria para μ_t pode ser obtida no tempo $t = \tau$, em que τ é o índice da primeira observação diferente de zero.

O intervalo de confiança assintótico para $\boldsymbol{\varphi}$ é construído baseado em uma aproximação numérica para a matriz de informação de Fisher obtida do próprio método de maximização numérica BFGS, utilizado para maximizar a função de log-verossimilhança com respeito a $\boldsymbol{\varphi}$.

Já para se fazer inferência Bayesiana sobre os parâmetros do modelo, é usado o método MCMC, algoritmo de Metropolis-Hastings (Gamerman & Lopes, 2006) - a fim de se obter uma amostra da distribuição *a posteriori* do vetor de parâmetros $\boldsymbol{\varphi}$ cuja forma é dada por:

$$\pi(\boldsymbol{\varphi}|\mathbf{Y}_n) \propto L(\boldsymbol{\varphi}; \mathbf{Y}_n)\pi(\boldsymbol{\varphi}), \quad (5.13)$$

onde $L(\boldsymbol{\varphi}; \mathbf{Y}_n)$ é a função de verossimilhança obtida em (5.12) e $\pi(\boldsymbol{\varphi})$ é a distribuição *a priori* para $\boldsymbol{\varphi}$. Neste trabalho, uma distribuição *a priori* uniforme, dada por $\pi(\boldsymbol{\varphi}) \propto c$ para todos os valores possíveis de $\boldsymbol{\varphi}$ e 0, caso contrário, é usada.

Intervalos de credibilidade para φ_i , $i = 1, \dots, p$ são construídos como se segue. Dado um valor $0 < \kappa < 1$, todo intervalo $(t_1, t_2)'$ satisfazendo

$$\int_{t_1}^{t_2} \pi(\varphi_i | \mathbf{Y}_n) d\varphi_i = 1 - \kappa$$

é um intervalo de credibilidade para φ_i com nível $100(1 - \kappa)\%$.

5.2 CASOS PARTICULARES DA FAMÍLIA GAMA DE MODELOS DINÂMICOS

Nas próximas subseções, serão discutidos alguns casos particulares da família Gama de modelos dinâmicos tais como os modelos Poisson, Gama, Weibull, Beta e Normal com média conhecida.

5.2.1 MODELO POISSON

Suponha que uma observação no tempo t é retirada de uma distribuição de Poisson com média μ_t ,

$$p(y_t | \mu_t, \boldsymbol{\varphi}) = \mu_t^{y_t} \exp(-\mu_t) / y_t!, \quad (5.14)$$

onde $y_t = 0, 1, \dots$, $\mu_t = \lambda_t \exp(x_t' \boldsymbol{\beta})$, função de ligação logarítmica. Esse modelo pertence à família Gama de modelos dinâmicos em que $a(y_t, \boldsymbol{\varphi}) = (y_t!)^{-1}$, $b(y_t, \boldsymbol{\varphi}) = y_t$ e $c(y_t, \boldsymbol{\varphi}) = 1$. Logo, $\boldsymbol{\varphi} = (w, \boldsymbol{\beta})'$

A distribuição *a priori* é a mesma do Teorema 1. Com as as funções $b(\cdot, \cdot)$ e $c(\cdot, \cdot)$ identificadas, utilizando o Teorema 1, a distribuição *a posteriori* de $\mu_t | \mathbf{Y}_t$ é dada pela distribuição Gama com parâmetros

$$a_t^* = a_{t|t-1}^* + y_t,$$

$$b_t^* = b_{t|t-1}^* + 1.$$

Logo, segue-se que $\lambda_t = \mu_t [g(x_t' \boldsymbol{\beta})]^{-1} | \mathbf{Y}_t$ tem também distribuição Gama com parâmetros (equações de atualização)

$$a_t = w a_{t-1} + y_t,$$

$$b_t = w b_{t-1} + g(x_t' \boldsymbol{\beta}).$$

Substituindo as funções $a(\cdot, \cdot)$, $b(\cdot, \cdot)$, $c(\cdot, \cdot)$ e usando o Teorema 1, obtém-se a distribuição preditiva, que é Binomial negativa, dada por

$$p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = \binom{a_{t|t-1}^* + y_t + 1}{y_t} (b_{t|t-1}^*)^{a_{t|t-1}^*} (1 + b_{t|t-1}^*)^{-(a_{t|t-1}^* + y_t)},$$

em que $y_t = 0, 1, 2, \dots$

e

$$\binom{a_{t|t-1}^* + y_t + 1}{y_t} = \frac{\Gamma(a_{t|t-1}^* + y_t)}{\Gamma(y_t + 1) \Gamma(a_{t|t-1}^*)}.$$

A função de log-verossimilhança é obtida cuja forma é dada por

$$\begin{aligned} \ln L(\boldsymbol{\varphi}; \mathbf{Y}_n) = & \sum_{t=\tau+1}^n \ln \Gamma(a_{t|t-1}^* + y_t) - \ln y_t! - \ln \Gamma(a_{t|t-1}^*) + \\ & a_{t|t-1}^* \ln b_{t|t-1}^* - (a_{t|t-1}^* + y_t) \ln(1 + b_{t|t-1}^*), \end{aligned} \quad (5.15)$$

onde τ é o instante da primeira observação diferente de zero e $\boldsymbol{\varphi} = (\omega, \boldsymbol{\beta})'$.

Então, segue das propriedades da distribuição binomial negativa, que a média e variância da distribuição preditiva de $y_{n+1}|Y_n, \boldsymbol{\varphi}$ são, respectivamente,

$$y_{n+1} = E(y_{n+1}|Y_n, \boldsymbol{\varphi}) = a_{n+1|n}^*/b_{n+1|n}^*$$

e

$$\text{var}(y_{n+1}|Y_n, \boldsymbol{\varphi}) = a_{n+1|n}^*(1 + b_{n+1|n}^*)/(b_{n+1|n}^*)^2.$$

5.2.2 MODELO GAMA

Suponha que a série temporal $\{y_t\}$ é gerada de uma distribuição Gama com o parâmetro de forma χ desconhecido e o parâmetro de escala $\chi\mu_t$, logo:

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = \frac{y_t^{\chi-1} \exp(-\mu_t \chi y_t)}{\Gamma(\chi) (\mu_t \chi)^{-\chi}}, \quad (5.16)$$

onde $y_t > 0$, $\mu_t = \lambda_t \exp(x_t' \boldsymbol{\beta})$ e $y_t > 0; \forall t \leq n$. O valor esperado de $(y_t|\mu_t, \boldsymbol{\varphi})$ é $1/\mu_t$. Se $\chi = 1$, $(y_t|\mu_t, \boldsymbol{\varphi})$ tem distribuição exponencial com média $1/\mu_t$.

O modelo Gama pode ser escrito na forma da família Gama de modelos dinâmicos em que $a(y_t, \boldsymbol{\varphi}) = \frac{y_t^{\chi-1} \chi^\chi}{\Gamma(\chi)}$, $b(y_t, \boldsymbol{\varphi}) = \chi$ e $c(y_t, \boldsymbol{\varphi}) = \chi y_t$.

Pelo Teorema 1, dado a t -ésima observação e as funções $b(\cdot, \cdot)$ e $c(\cdot, \cdot)$ devidamente discriminadas, a distribuição *a posteriori* de $\mu_t|Y_t$ é Gama com parâmetros

$$a_t^* = a_{t|t-1}^* + \chi$$

$$b_t^* = b_{t|t-1}^* + \chi y_t.$$

Logo, tem-se que $\lambda_t = \mu_t [g(x_t' \boldsymbol{\beta})]^{-1} | Y_t \sim \text{Gama}(a_t, b_t)$ e as equações de atualização são dadas por:

$$a_t = a_{t|t-1} + \chi$$

$$b_t = b_{t|t-1} + \chi y_t g(x_t' \beta).$$

Substituindo as funções $a(\cdot, \cdot)$, $b(\cdot, \cdot)$ e $c(\cdot, \cdot)$ em (5.11), a função de densidade preditiva $y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi} \sim \text{Gama-gama}(a_{t|t-1}^*, b_{t|t-1}^* / \chi, \chi)$ cuja forma é dada por:

$$p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = \frac{\Gamma(\chi + a_{t|t-1}^*) y_t^{\chi-1}}{\Gamma(a_{t|t-1}^*) \Gamma(\chi) (b_{t|t-1}^* / \chi)^{-a_{t|t-1}^*} (y_t + b_{t|t-1}^* / \chi)^{\chi + a_{t|t-1}^*}},$$

se $y_t > 0$ e 0 , caso contrário.

A função de verossimilhança é o produto das funções de densidade preditivas dada por:

$$\ln L(\boldsymbol{\varphi}; \mathbf{Y}_n) = \ln \prod_{t=\tau+1}^n p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = \sum_{t=\tau+1}^n \ln p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = \ln \Gamma(\chi + a_{t|t-1}^*) - \ln \Gamma(\chi) \Gamma(a_{t|t-1}^*) + a_{t|t-1}^* \ln(b_{t|t-1}^* / \chi) + \ln y_t^{\chi-1} - (\chi + a_{t|t-1}^*) \ln(y_t + b_{t|t-1}^* / \chi),$$

onde τ é definido como em (5.15) e $\boldsymbol{\varphi} = (\omega, \beta, \chi)'$.

Das propriedades da distribuição Gama-gama, a média e a variância de y_{n+1} condicionada a informação até o tempo n são

$$y_{n+1|n} = \mathbb{E}(y_{n+1} | \mathbf{Y}_n, \boldsymbol{\varphi}) = \frac{b_{t|t-1}^*}{a_{t|t-1}^* - 1}$$

e

$$\text{var}(y_{n+1} | \mathbf{Y}_n, \boldsymbol{\varphi}) = \frac{(b_{t|t-1}^*)^2 [\chi^2 + \chi(a_{t|t-1}^* - 1)]}{(a_{t|t-1}^* - 1)^2 (a_{t|t-1}^* - 2)}.$$

5.2.3 MODELO WEIBULL

Se as observações no tempo t são geradas de uma distribuição Weibull triparamétrica (ver Ross (2002)) e os parâmetros $v_t = v$ e $\rho_t = \rho$ são invariantes no tempo e desconhecidos, então:

$$p(y_t | \boldsymbol{\mu}_t, \boldsymbol{\varphi}) = v \boldsymbol{\mu}_t (y_t - \rho)^{v-1} \exp[-\boldsymbol{\mu}_t (y_t - \rho)^v], \quad (5.17)$$

onde $y_t > \rho$, $\boldsymbol{\mu}_t, v > 0$ e $\boldsymbol{\mu}_t = \lambda_t g(x_t' \beta)$.

O modelo Weibull pode ser escrito na forma da família gama de modelos dinâmicos onde $a(y_t, \boldsymbol{\varphi}) = v(y_t - \rho)^{v-1}$, $b(y_t, \boldsymbol{\varphi}) = 1$ e $c(y_t, \boldsymbol{\varphi}) = (y_t - \rho)^v$.

Pelo Teorema 1, a distribuição *a posteriori* de $\mu_t|Y_t$ é Gama com parâmetros

$$\begin{aligned} a_t^* &= a_{t|t-1}^* + 1, \\ b_t^* &= b_{t|t-1}^* + (y_t - \rho)^v. \end{aligned}$$

Então, segundo o Teorema 1, tem-se que $\lambda_t = \mu_t[g(x_t'\beta)]^{-1}|Y_t \sim \text{Gama}(a_t, b_t)$ e as equações de atualização são dadas por:

$$\begin{aligned} a_t &= a_{t|t-1} + 1, \\ b_t &= b_{t|t-1} + (y_t - \rho)^v g(x_t'\beta). \end{aligned}$$

Em dados de análise de sobrevivência, pode-se assumir $a_t = a_{t|t-1} + \delta_t$, onde δ_t é um indicador de censura à direita.

Conhecendo as funções $a(\cdot, \cdot)$, $b(\cdot, \cdot)$ e $c(\cdot, \cdot)$ e utilizando o Teorema 1, função de densidade preditiva $y_t|\mathbf{Y}_{t-1}$, $\boldsymbol{\varphi}$ é dada por:

$$\frac{\Gamma(1 + a_{t|t-1}^*) v (y_t - \rho)^{v-1}}{\Gamma(a_{t|t-1}^*) (b_{t|t-1}^*)^{-a_{t|t-1}^*} [(y_t - \rho)^v + b_{t|t-1}^*]^{1+a_{t|t-1}^*}};$$

onde $y_t > \rho$, $a_{t|t-1}^*$ e $b_{t|t-1}^*$ são dadas pelo Teorema 1.

A função de verossimilhança, que é o produto das funções de densidade preditivas, é dada por:

$$\begin{aligned} \ln L(\boldsymbol{\varphi}; \mathbf{Y}_n) &= \ln \prod_{t=\tau+1}^n p(y_t|\mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = \sum_{t=\tau+1}^n \ln \Gamma(1 + a_{t|t-1}^*) + \ln v (y_t + \kappa)^{v-1} \\ &\quad - \ln \Gamma(a_{t|t-1}^*) + a_{t|t-1}^* \ln b_{t|t-1}^* - (1 + a_{t|t-1}^*) \ln [(y_t - \kappa)^v + b_{t|t-1}^*], \end{aligned}$$

em que $\boldsymbol{\varphi} = (\omega, \beta, v, \rho)'$ e τ é definido como em (5.15).

5.2.4 MODELO PARETO

A distribuição de Pareto é uma distribuição que possui muitas aplicações em problemas econômicos, sociais e geofísicos. Se as observações no tempo t são geradas de uma distribuição Pareto com parâmetros $\rho > 0$, desconhecido e invariante no tempo, e $\mu_t > 0$, então:

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = \mu_t \rho^{\mu_t} y_t^{-\mu_t-1}, \quad (5.18)$$

onde $y_t > \rho$ e $\mu_t = \lambda_t g(x_t'\beta)$.

O modelo Pareto pode ser escrito na forma da família gama de modelos dinâmicos em que

$$a(y_t, \boldsymbol{\varphi}) = y_t^{-1}, \quad b(y_t, \boldsymbol{\varphi}) = 1, \quad c(y_t, \boldsymbol{\varphi}) = \ln y_t - \ln \rho.$$

As distribuições *a priori* são as mesmas da família gama de modelos dinâmicos. Quando a t -ésima observação é obtida e com as funções $b(\cdot, \cdot)$ e $c(\cdot, \cdot)$ a distribuição *a posteriori* de $\mu_t | Y_t$, pelo Teorema 1, é Gama com parâmetros

$$\begin{aligned} a_t^* &= a_{t|t-1}^* + 1, \\ b_t^* &= b_{t|t-1}^* - \ln \rho + \ln y_t. \end{aligned}$$

Logo, fazendo a transformação inversa, tem-se que $\lambda_t = \mu_t [g(x_t' \boldsymbol{\beta})]^{-1} | Y_{t-1} \sim \text{Gama}(a_t, b_t)$ onde as equações de atualização são dadas por:

$$\begin{aligned} a_t &= a_{t|t-1} + 1, \\ b_t &= b_{t|t-1} + (-\ln \rho + \ln y_t) g(x_t' \boldsymbol{\beta}). \end{aligned}$$

Pelo Teorema 1, de posse das funções $a(\cdot, \cdot)$, $b(\cdot, \cdot)$ e $c(\cdot, \cdot)$, a função de densidade preditiva $y_t | Y_{t-1}, \boldsymbol{\varphi}$ tem a seguinte forma:

$$\frac{a_{t|t-1}^* y_t^{-1} [b_{t|t-1}^*]^{a_{t|t-1}^*}}{[b_{t|t-1}^* - \ln \rho + \ln y_t]^{a_{t|t-1}^* + 1}},$$

onde $y_t > \rho$.

Com as funções de densidade preditivas, é possível determinar a função de verossimilhança, que é dada por:

$$\begin{aligned} \ln L(\boldsymbol{\varphi}; \mathbf{Y}_n) &= \ln \prod_{t=\tau+1}^n p(y_t | Y_{t-1}, \boldsymbol{\varphi}) = \sum_{t=\tau+1}^n \ln p(y_t | Y_{t-1}, \boldsymbol{\varphi}) = \\ &= \sum_{t=\tau+1}^n \ln(a_{t|t-1}^* y_t^{-1} (b_{t|t-1}^*)^{a_{t|t-1}^*}) - (a_{t|t-1}^* + 1) \ln[b_{t|t-1}^* - \ln \rho + \ln y_t], \end{aligned}$$

onde $\boldsymbol{\varphi} = (\omega, \boldsymbol{\beta}, \rho)'$.

5.2.5 MODELO BETA

Quando um parâmetro da distribuição Beta é igual a 1, pode-se escrever esse modelo na forma de FGMD. Esse modelo é muito útil para modelar séries temporais que são proporções e probabilidades. Suponha que a série temporal $\{y_t\}$ é gerada de uma distribuição Beta com parâmetros μ_t e 1 cuja função de densidade é dada por

$$p(y_t | \mu_t, \boldsymbol{\varphi}) = \mu_t \mathbf{y}_t^{\mu_t - 1}, \quad (5.19)$$

onde $0 < y_t < 1$.

Esse modelo também pertence à FGMD no qual $a(y_t, \boldsymbol{\varphi}) = y_t^{-1}$, $b(y_t, \boldsymbol{\varphi}) = 1$ e $c(y_t, \boldsymbol{\varphi}) = -\ln(y_t)$. As equações de atualização e a função de densidade preditiva podem ser encontradas similarmente aos modelos anteriores, usando o Teorema 1. Neste caso, $\boldsymbol{\varphi} = (\omega, \beta)'$.

5.2.6 MODELO NORMAL COM MÉDIA CONHECIDA

Essa metodologia pode ser aplicada também em modelos gaussianos. Se as observações no tempo t são geradas de uma distribuição Normal com média zero e parâmetro de precisão $\mu_t > 0$, então:

$$p(y_t | \mu_t, \boldsymbol{\varphi}) = \frac{\mu_t^{1/2}}{\sqrt{2\pi}} \exp\left(\frac{-\mu_t y_t^2}{2}\right), \quad (5.20)$$

onde $-\infty < y_t < \infty$ e $\mu_t = \lambda_t g(x_t' \beta)$.

O modelo Normal pode ser escrito na forma da família Gama de modelos dinâmicos em que $a(y_t, \boldsymbol{\varphi}) = (2\pi)^{-1/2}$, $b(y_t, \boldsymbol{\varphi}) = 1/2$ e $c(y_t, \boldsymbol{\varphi}) = y_t^2/2$.

A distribuição *a priori* de $\lambda_t | \mathbf{Y}_{t-1}$ é dada pelo item 2 do Teorema 1. Segundo o Teorema 1, a distribuição *a posteriori* de $\mu_t | \mathbf{Y}_t$ também pode ser obtida, sendo Gama com parâmetros

$$\begin{aligned} a_t^* &= a_{t|t-1}^* + 1/2, \\ b_t^* &= b_{t|t-1}^* + y_t^2/2. \end{aligned}$$

Logo, usando novamente a propriedade de escala da distribuição Gama, tem-se que $\lambda_t = \mu_t [g(x_t' \beta)]^{-1} | \mathbf{Y}_t \sim \text{Gama}(a_t, b_t)$ onde as equações de atualização são dadas por:

$$\begin{aligned} a_t &= a_{t|t-1} + 1/2, \\ b_t &= b_{t|t-1} + (y_t^2/2)g(x_t' \beta). \end{aligned}$$

A função de densidade preditiva $y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}$, pelo Teorema 1, tem a seguinte forma:

$$\frac{\Gamma(a_{t|t-1}^* + 1/2) (2\pi)^{-1/2} (b_{t|t-1}^*)^{a_{t|t-1}^*}}{\Gamma(a_{t|t-1}^*) [y_t^2/2 + b_{t|t-1}^*]^{a_{t|t-1}^* + 1/2}},$$

onde $-\infty < y_t < \infty$.

Com as funções de densidade preditivas, é possível determinar a função de verossimilhança,

que é dada por:

$$\ln L(\boldsymbol{\varphi}; \mathbf{Y}_n) = \ln \prod_{t=\tau+1}^n p(y_t | Y_{t-1}, \boldsymbol{\varphi}) = \sum_{t=\tau+1}^n \ln p(y_t | Y_{t-1}, \boldsymbol{\varphi}) = \sum_{t=\tau+1}^n \ln \Gamma(a_{t|t-1}^* + 1/2) (2\pi)^{-1/2} (b_{t|t-1}^*)^{a_{t|t-1}^*} - \ln \Gamma(a_{t|t-1}^*) [y_t^2/2 + b_{t|t-1}^*]^{a_{t|t-1}^* + 1/2},$$

em que $\boldsymbol{\varphi} = (\omega, \beta)'$.

Da mesma forma que o modelo Normal com média conhecida e evolução na variância foi construído, os modelos Log-normal e Normal Inversa podem ser desenvolvidos. O modelo Valor extremo também pode ser construído, uma vez que a distribuição do Valor extremo é uma transformação um-a-um de uma distribuição Exponencial (Smith & Miller, 1986).

5.3 APLICAÇÕES A MODELOS ESTRUTURAIS DE POISSON

Neste capítulo, simulações Monte Carlo são realizadas para o modelo de Poisson. O EMV e os estimadores Bayesianos são comparados quanto ao vício e ao EQM, assim como os intervalos de credibilidade e de confiança são comparados com respeito à amplitude e taxa de cobertura. Adicionalmente, para ilustração da metodologia apresentada no capítulo anterior, um modelo de Poisson é ajustado à série temporal do número de gols marcados pelo Brasil em confrontos com a Argentina.

5.3.1 MODELO DE POISSON

Através de simulações Monte Carlo, os desempenhos do estimador de máxima verossimilhança (EMV) e dos estimadores Bayesianos - EB-média, EB-mediana e EB-moda - foram investigadas para séries temporais de tamanho $n = 100$, geradas sob o modelo Poisson com uma covariável $x_t = \cos(2\pi t/12)$ e parâmetros $w = 0.90$ e $\beta = 1$. Duas cadeias de 2000 amostras foram geradas das quais as 1000 primeiras foram excluídas. O número de replicações Monte Carlo foi fixado em 500. O nível de confiança e a probabilidade de credibilidade dos intervalos de confiança e de credibilidade, respectivamente, foram fixados em 0,95.

Na Figura 5.1, está uma série simulada sob o modelo Poisson com os mesmos valores dos parâmetros descritos anteriormente. Observa-se que a série oscila em torno de um nível médio igual a 3, aproximadamente. Nota-se também que a linha tracejada referente à média da distribuição preditiva do modelo Poisson acompanha bem o comportamento da série (linha contínua).

Na Tabela 5.1, o EMV e os estimadores Bayesianos são comparados quanto ao vício e ao EQM. Para w , todos os estimadores possuem valores de vício e EQM muito próximos. Já para β , o EMV e o EB-moda apresentam valores do EQM menores que os outros métodos.

Na Tabela 5.2, os intervalos de confiança e de credibilidade são comparados através da taxa de cobertura e amplitude. A taxa de cobertura do intervalo de credibilidade está mais próxima do nível nominal 0.95 assumido que o intervalo assintótico para w . Para β ambos têm taxas de cobertura próximas. A amplitude do intervalo de credibilidade é ligeiramente menor que a do intervalo de confiança. Nota-se que o limite superior do intervalo de confiança do parâmetro ω ultrapassou o limite do espaço paramétrico. Isso pode ter ocorrido por se tratar de um intervalo de confiança assintótico.

Tabela 5.1: EMV e EB para o modelo Poisson.

	EMV	Mediana	Media	Moda
	estimativa	estimativa	estimativa	estimativa
	Vício	Vício	Vício	Vício
φ	(EQM)	(EQM)	(EQM)	(EQM)
$w = 0.90$	0.912	0.889	0.883	0.910
	0.012	-0.011	-0.017	0.010
	(0.003)	(0.003)	(0.003)	(0.004)
$\beta = 1.00$	0.990	1.038	1.039	1.005
	-0.010	0.038	0.039	0.005
	(0.068)	(0.148)	(0.150)	(0.017)

Tabela 5.2: Intervalos de confiança e de credibilidade para o modelo Poisson.

	Int. Cred.	Int. Assint.
	limites	limites
	amplitude	amplitude
φ	(cobertura)	(cobertura)
$w = 0.90$	[0.771; 0.962]	[0.816; 1.013]
	0.191	0.197
	(0.962)	(0.922)
$\beta = 1.00$	[0.787; 1.308]	[0.812; 1.207]
	0.521	0.395
	(0.956)	(0.944)

5.3.2 APLICAÇÃO À SÉRIE TEMPORAL HISTÓRICA DO CONFRONTO ENTRE BRASIL E ARGENTINA

Nesta seção, como ilustração da metodologia dos modelos de resposta não-gaussiana apresentada, o modelo Poisson é ajustado à série do número de gols marcados pelo Brasil contra a

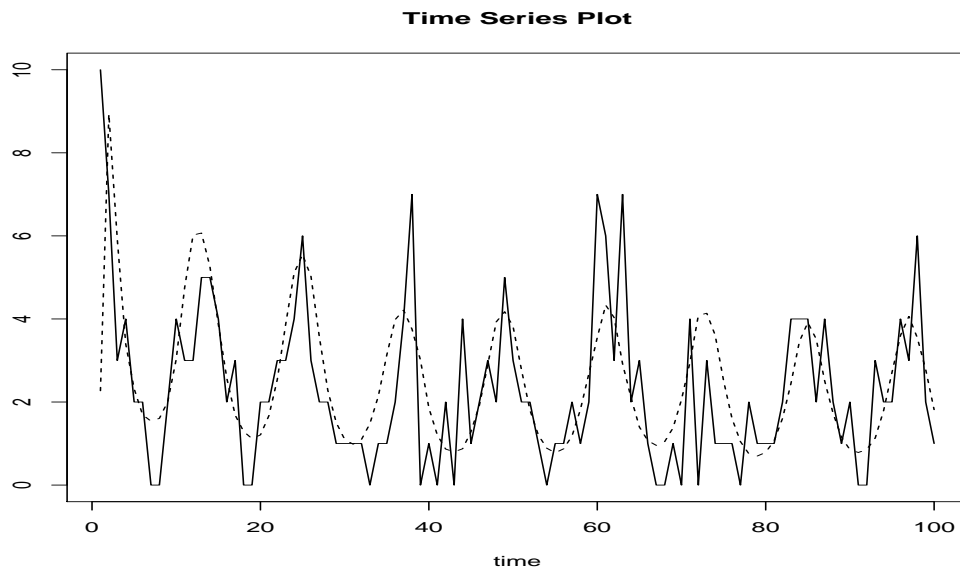


Figura 5.1: Série temporal simulada sob o modelo Poisson. A linha contínua representa a série temporal e a linha tracejada a média da distribuição preditiva do ajuste clássico.

Argentina. No site da FIFA, foram obtidos os placares¹ de 93 jogos entre Brasil e Argentina no período de 1914 a 2008. Como em alguns anos não houve jogos, tem-se dados faltantes que terão que ser tratados. Também houve anos em que ocorreram mais de um jogo no mesmo ano. Para contornar esse problema, alterações nas equações de evolução e atualização do algoritmo recursivo foram realizadas. Da Figura 5.2, percebe-se que a média de gols marcados pelo Brasil está em torno de 2, salvo alguns anos nos quais o número de gols foi superior a 2. Como os dados referem-se a jogos no Brasil e também fora do Brasil, o modelo Poisson é ajustado com uma covariável x_t que recebe 1 se o jogo foi realizado no Brasil e 0 fora. Tem-se interesse em saber se realmente faz diferença jogar contra Argentina em casa.

Na Tabela 5.3, o EMV, os estimadores pontuais Bayesianos, intervalos de confiança e de credibilidade para φ são apresentados do ajuste do modelo Poisson à série real. Nota-se que os valores do EMV e do EB-Moda ficaram bem próximos tanto para o parâmetro w quanto para β . Os limites do intervalo de credibilidade para β são similares aos do intervalo de confiança assintótico, porém isso não acontece para w . Percebe-se que o limite superior do intervalo assintótico está fora do espaço paramétrico de w . Isso pode acontecer por se tratar de um intervalo aproximado. Observa-se que $\hat{\beta} = 0,485$ (maior que zero), indicando que jogar em casa confere ao Brasil uma certa vantagem, como esperado. Tomando $\exp(\hat{\beta})$, obtém-se 1,624, isto é, o número esperado de gols marcados pelo Brasil aumenta 62,40%, quando o Brasil joga em casa contra Argentina. Esse valor é significativo ao nível de 0,95 de confiança e de

¹Os dados estão disponíveis em <http://es.fifa.com>

Tabela 5.3: Ajuste do modelo Poisson à série do número de gols marcados pelo Brasil no confronto com a Argentina.

Métodos	w	β
EMV	1,000	0,485
Mediana	0,956	0,461
Media	0,949	0,475
Moda	0,999	0,506
Int. Cred.	[0,856; 0,997]	[0,125; 0,817]
Int. Assint.	[0,721; 1,190]	[0,153; 0,812]

credibilidade, pois o valor zero não está contido tanto no intervalo de confiança quanto no intervalo de credibilidade. Os resíduos de Pearson padronizados foram avaliados e a variância amostral dos mesmos foi igual a 1,042. Através da análise dos gráficos dos resíduos, não observou-se nenhuma violação das suposições do modelo.

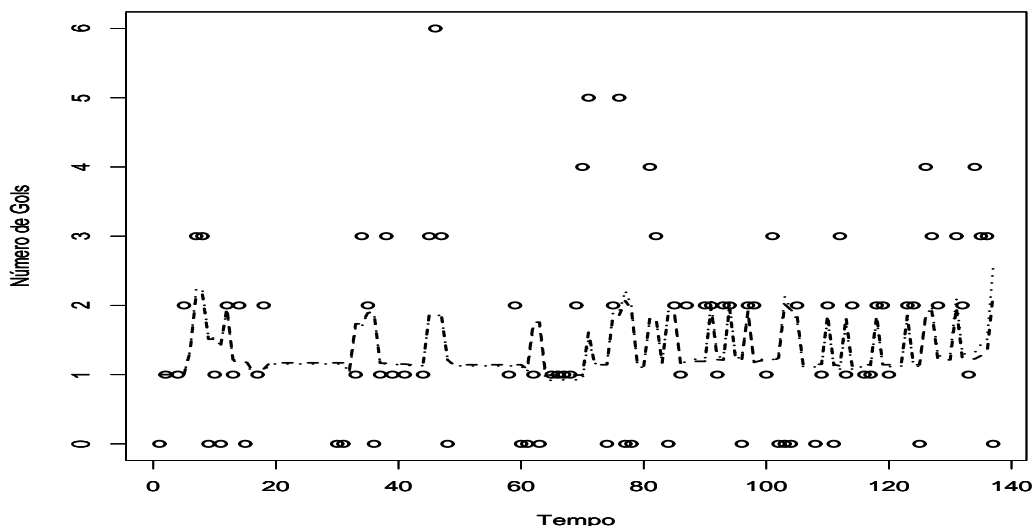


Figura 5.2: Série temporal do número de gols marcados pelo Brasil no confronto com a Argentina. Os pontos representam a série temporal e as linhas tracejada e pontilhada representam as médias da distribuição preditiva dos ajustes clássico e Bayesiano, respectivamente.

A Tabela 5.4 apresenta as probabilidades do Brasil marcar 0,1,2,3,4 ou mais gols contra Argentina no próximo jogo em casa ou fora de casa. Nota-se que a probabilidade do Brasil fazer nenhum gol jogando fora de casa é maior do que jogando em casa. O ajuste do modelo pode ser melhorado, pois há inúmeros fatores que influenciaram essa série temporal histórica que poderiam ser discutidos e incluídos no modelo, porém, este estudo não entrará nesse mérito. O intuito desta aplicação é mostrar a potencialidade e aplicabilidade dos modelos de resposta não-gaussiana nos mais diversos tipos de problemas.

Tabela 5.4: Distribuição de probabilidade preditiva do número de gols do Brasil na próxima partida contra a Argentina.

Número de gols	Lugar	
	Em casa	Fora de casa
0	0,125	0,277
1	0,257	0,353
2	0,267	0,227
3	0,187	0,099
4	0,099	0,033
>4	0,065	0,011

6 CONCLUSÕES E TRABALHOS FUTUROS

A tônica deste trabalho foi a comparação de métodos clássicos e Bayesianos na estimação dos hiperparâmetros e parâmetros em modelos estruturais específicos, bem como em modelos de função de transferência e em modelos estruturais com resposta não-gaussiana. Do ponto de vista clássico, o método de máxima verossimilhança foi utilizado no processo de inferência. Por outro lado, sob o enfoque Bayesiano, a estimação foi feita através de estimadores de Bayes, utilizando-se métodos MCMC devido ao fato da distribuição *a posteriori* não ser analiticamente tratável. Usaram-se distribuições *a priori* não-informativas (uniforme e de Jeffreys) e foram escolhidas a média, a mediana e a moda *a posteriori* como estimadores Bayesianos.

Intervalos de confiança assintóticos e *bootstrap* e intervalos de credibilidade para os hiperparâmetros em modelos estruturais específicos (MNL, MTL e MEB) foram construídos e comparados com respeito à amplitude e à porcentagem de cobertura. Os intervalos de confiança *bootstrap* considerados foram o percentílico, o BC (Bias Corrected) e o BCa (Bias Corrected and accelerated).

Os resultados obtidos ao longo deste trabalho foram transformados em artigos que estão submetidos, em fase de preparação ou já publicados.

No primeiro artigo que foi a base deste trabalho, publicado no periódico *Communications in Statistics*, foi feita uma comparação entre os intervalos de confiança assintóticos e *bootstrap*. Os resultados mostraram que a porcentagem de cobertura dos intervalos *bootstrap* ficou mais próxima do nível nominal. Foi ressaltado que os intervalos assintóticos possuíam problemas de fronteira, o que reforçou o uso de intervalos *bootstrap*. Além das simulações, foi feita também uma aplicação à série IPCA (Índice de Preço ao Consumidor Amplo) de Belo Horizonte. Um modelo de tendência linear local foi ajustado à mesma e intervalos de confiança assintótico e *bootstrap* foram construídos. Concluiu-se que o modelo de nível local é o mais indicado para descrever o comportamento da série.

No segundo artigo, submetido à revista *Computational Statistics*, foi realizada uma comparação entre os intervalos de confiança assintóticos e *bootstrap* e os intervalos de credibilidade e uma comparação entre o EMV e os estimadores Bayesianos. Através de simulações Monte Carlo, o vício e o EQM dos estimadores foram calculados. Os resultados mostraram que, em geral, o EMV é menos viesado, embora, em alguns casos, os estimadores Bayesianos possuem um menor EQM. Foi aplicado o *bootstrap* não-paramétrico nos resíduos do modelo e foi verificado que as estimativas *bootstrap* acompanham bem o comportamento das EMV nas replicações Monte Carlo. Os intervalos de credibilidade mostraram, em geral, um melhor desempenho com uma amplitude menor e uma porcentagem de cobertura mais próxima do nível nominal assumido. Além disso, a metodologia foi aplicada à série da precipitação de SO_4 (sulfato) e concluiu-se que a série segue um modelo estrutural básico, sem o componente estocástico de tendência. Nenhuma das suposições do modelo estrutural gaussiano foi violada.

No terceiro artigo, submetido à revista *International Statistical Review*, foi abordado o tópico de análise de intervenção em modelos estruturais, sob as perspectivas clássica e Bayesiana. Intervenções foram consideradas para o nível médio das séries em duas formas canônicas: funções pulso e degrau. Essas funções tratam de quebras nas formas abruptas e graduais, respectivamente. Através de simulações Monte Carlo, percebeu-se que o EMV e a moda *a posteriori* apresentaram um melhor desempenho com menores vício e EQM, em geral. Os intervalos de credibilidade obtiveram resultados melhores que os intervalos de confiança (I.C.) com função degrau e resultados similares aos dos I.C. com função pulso, em geral. Essa metodologia foi aplicada a duas séries econômicas reais de índices brasileiros e os resultados mostraram se satisfatórios para os modelos de intervenção considerados.

O quarto artigo trata de modelos estruturais com resposta não-gaussiana e está em fase de elaboração. Uma família gama de modelos dinâmicos foi apresentada bem como alguns estudos de simulação para um caso particular dessa família, o modelo Poisson. Comparações entre os estimadores pontuais Bayesianos e clássicos foram realizadas, assim como entre os intervalos de confiança e de credibilidade para os parâmetros do modelo. Os resultados mostraram que as taxas de cobertura dos intervalos de credibilidade e confiança ficaram bem próximos do nível nominal. O EMV e o EB-Moda apresentam valores do EQM ligeiramente menores para todos os parâmetros do modelo considerados se comparados aos outros estimadores. A título de ilustração da metodologia desenvolvida, o modelo Poisson foi ajustado à série temporal do número de gols marcados pelo Brasil contra a Argentina em confrontos de 1914 a 2008 e os resultados foram satisfatórios.

Deve-se ressaltar que as técnicas *bootstrap* e MCMC demandam um tempo computacional

muito elevado. Para se ter uma idéia, em um modelo simples como MNL, a execução do método MCMC leva em torno de meio minuto para uma replicação Monte Carlo de uma série temporal de tamanho $n = 100$, enquanto que o *bootstrap* gasta aproximadamente um minuto. Já as estimativas de máxima verossimilhança e do intervalo de confiança assintótico são obtidas instantaneamente.

Alguns tópicos podem ser desenvolvidos futuramente. Outros modelos, além dos modelos de nível local, de tendência linear local e estrutural básico, que incluem outros componentes podem ser estudados como, por exemplo, o componente de ciclo.

Há um número de outras questões relacionadas aos modelos de função de transferência que não foram abordadas aqui. Este trabalho foi focado na estimação dos hiperparâmetros e predição de valores. A distribuição suavizada para os parâmetros do vetor de espaço de estados pode ser obtida facilmente, mas não foi abordada. Outros componentes não-observáveis podem ser incluídos facilmente como o de sazonalidade. Os resultados podem ser estendidos aos modelos de função de transferência na família exponencial, ver Alves, Gamerman & Ferreira (2009), mas a verossimilhança exata não é obtida analiticamente. Uma discussão mais completa sobre previsão com a incorporação da incerteza sobre os hiperparâmetros poderia ser feita. Um outro tópico que poderia ser melhor explorado e discutido é a aplicação do *bootstrap* não-paramétrico em modelos estruturais de FT bem como a construção de uma aproximação melhor para variância assintótica do vetor paramétrico desses modelos que incluem parâmetros e hiperparâmetros.

Com respeito aos modelos de resposta não-gaussiana, vários trabalhos podem ser desenvolvidos. Casos específicos de interesse da família gama de modelos dinâmicos podem ser encontrados. Um tópico interessante seria trabalhar com suavização nesses modelos - que ainda está em aberto. O modelo exponencial por partes é muito utilizado e de interesse principalmente em estudos de confiabilidade e se encaixa nos moldes dos modelos de resposta não-gaussiana, embora a dependência das observações não é considerada no modelo exponencial por partes, mas equivalentemente aplicável a modelos que levem em conta a autocorrelação das observações. Um outro trabalho interessante que poderia ser desenvolvido é a comparação do modelo de resposta não-gaussiana com os modelos lineares dinâmicos generalizados (West & Harrison, 1997) seja via experimentos Monte Carlo seja utilizando séries reais.

Em todos os artigos que compõem esta dissertação, os intervalos de confiança e de credibilidade foram abordados. Embora eles sejam também um tipo de teste de hipóteses, seria importante construir testes de hipóteses específicos para os hiperparâmetros e, conseqüentemente, verificar a presença ou não de determinados componentes estocásticos não-observáveis.

Testes assintóticos, *bootstrap* (Franco *et al.*, 1999) e Bayesianos poderiam ser construídos e comparados quanto ao nível de significância e poder para vários modelos estruturais, seguindo a mesma linha que foi adotada neste trabalho.

Alguma atenção foi dada à parte de previsão nos modelos de função transferência, porém esse tópico deve ser melhor explorado em todos os modelos citados anteriormente, pois um dos objetivos de se modelar séries temporais é fazer previsões para valores futuros da mesma. As previsões clássica e Bayesiana podem ser comparadas em um trabalho específico tratando somente desse assunto.

APÊNDICE

Demonstração do Teorema 1.

Assumindo a definição do modelo da Seção 5.1.1, podem-se obter as equações de atualização:

Por indução, supõe-se que $\lambda_{t-1}|\mathbf{Y}_{t-1} \sim Gama(a_{t-1}, b_{t-1})$ é válido para t e, consequentemente, $\mu_{t-1}|\mathbf{Y}_{t-1} \sim Gama(a_{t-1}^*, b_{t-1}^*)$:

- Se $t = 1$, $\lambda_0|Y_0 \sim Gama(a_0, b_0)$ e $\mu_0|Y_0 \sim Gama(a_0, b_0)$ - que é verdade pela Suposição 4 do modelo;
- Agora, basta provar que essa suposição vale para $t + 1$. As distribuições de $\lambda_{t-1}|\mathbf{Y}_{t-1}$ e $\lambda_t|\lambda_{t-1}, \mathbf{Y}_{t-1}$ são conhecidas. A primeira por suposição e a última pela demonstração feita na Prova 1a abaixo.

Integrando fora em λ_{t-1} , conclue-se que $\lambda_t|\mathbf{Y}_{t-1} \sim Gama(a_{t|t-1} = w a_{t-1}, b_{t|t-1} = w b_{t-1})$ (ver demonstração na Prova 1a abaixo). Portanto,

$$\mu_t = \lambda_t g(x_t' \beta) | \mathbf{Y}_{t-1} \sim Gama(a_{t|t-1}^* = a_{t|t-1}, b_{t|t-1}^* = b_{t|t-1} g(x_t' \beta)^{-1});$$

Pelo Teorema de Bayes,

$$p(\mu_t | \mathbf{Y}_t) \propto p(y_t | \mu_t) p(\mu_t | \mathbf{Y}_{t-1}) \propto \mu_t^{(a_{t|t-1}^* + b(y_t, \boldsymbol{\varphi})) - 1} \exp[-\mu_t (b_{t|t-1}^* + c(y_t, \boldsymbol{\varphi}))].$$

Então, segue-se que $\mu_t | \mathbf{Y}_t \sim Gama(a_t^* = a_{t|t-1}^* + b(y_t, \boldsymbol{\varphi}), b_t^* = b_{t|t-1}^* + c(y_t, \boldsymbol{\varphi}))$. Logo,

$\lambda_t = \mu_t g(x_t' \beta)^{-1} | \mathbf{Y}_t \sim Gama(a_t = a_{t|t-1} + b(y_t, \boldsymbol{\varphi}), b_t = b_{t|t-1} + c(y_t, \boldsymbol{\varphi}) g(x_t' \beta))$, $\forall t \in N$; $t \leq n$ onde n é o tamanho da série temporal. A hipótese indutiva é verificada.

Prova 1a:

Deseja-se encontrar a distribuição de $\lambda_t = w^{-1} \lambda_{t-1} \zeta_t$, onde $\zeta_t \sim B(w a_{t-1}, (1-w) a_{t-1})$.

Então, usando o Jacobiano, tem-se que

$$p_{\lambda_t | \lambda_{t-1}, \mathbf{Y}_{t-1}}(\lambda_t^*) = \begin{cases} \frac{w}{\lambda_{t-1}} \left(\frac{w \lambda_t^*}{\lambda_{t-1}} \right)^{w a_{t-1} - 1} \left(1 - \frac{w \lambda_t^*}{\lambda_{t-1}} \right)^{(1-w) a_{t-1} - 1}; & \text{se } 0 < \lambda_t^* < w^{-1} \lambda_{t-1} \\ 0; & \text{c.c.} \end{cases} .$$

Deseja-se encontrar a distribuição de $\lambda_t | \mathbf{Y}_{t-1}$. Para tanto, integra-se fora em λ_{t-1} e utiliza-se o resultado obtido na Prova 1a.

$$p(\lambda_t | \mathbf{Y}_{t-1}) = \int p(\lambda_{t-1} | \mathbf{Y}_{t-1}) p(\lambda_t | \lambda_{t-1}, \mathbf{Y}_{t-1}) d\lambda_{t-1} =$$

$$\int_{w\lambda_t}^{\infty} \left[\frac{\lambda_{t-1}^{a_{t-1}-1} \exp(-b_{t-1}\lambda_{t-1})}{\Gamma(a_{t-1}) b_{t-1}^{-a_{t-1}}} \right] \left[\frac{w\lambda_{t-1}^{-1} \left(\frac{w\lambda_t}{\lambda_{t-1}}\right)^{wa_{t-1}-1} \left(1 - \frac{w\lambda_t}{\lambda_{t-1}}\right)^{(1-w)a_{t-1}-1}}{B(wa_{t-1}, (1-w)a_{t-1})} \right] d\lambda_{t-1}$$

Seja $c = \frac{w(w\lambda_t)^{wa_{t-1}-1}}{\Gamma(a_{t-1}) b_{t-1}^{-a_{t-1}} B(wa_{t-1}, (1-w)a_{t-1})}$, logo,

$$p(\lambda_t | \mathbf{Y}_{t-1}) = c \int_{w\lambda_t}^{\infty} \left[\lambda_{t-1}^{a_{t-1}-1-wa_{t-1}+1-1} \exp(-b_{t-1}\lambda_{t-1}) \right] \left[\left(1 - \frac{w\lambda_t}{\lambda_{t-1}}\right)^{(1-w)a_{t-1}-1} \right] d\lambda_{t-1} =$$

$$c \int_{w\lambda_t}^{\infty} \left[\lambda_{t-1}^{(1-w)a_{t-1}-1} \exp(-b_{t-1}\lambda_{t-1}) \right] \left[\left(1 - \frac{w\lambda_t}{\lambda_{t-1}}\right)^{(1-w)a_{t-1}-1} \right] d\lambda_{t-1} =$$

$$c \int_{w\lambda_t}^{\infty} \exp(-b_{t-1}\lambda_{t-1}) (\lambda_{t-1} - w\lambda_t)^{(1-w)a_{t-1}-1} d\lambda_{t-1}$$

Considere $z = \lambda_{t-1} - w\lambda_t$, então

$$p(\lambda_t | \mathbf{Y}_{t-1}) = c \int_0^{\infty} \exp[-b_{t-1}(z + w\lambda_t)] z^{(1-w)a_{t-1}-1} dz =$$

$$c \exp(-wb_{t-1}\lambda_t) \int_0^{\infty} \exp(-b_{t-1}z) z^{(1-w)a_{t-1}-1} dz =$$

$$\frac{w(w\lambda_t)^{wa_{t-1}-1}}{\Gamma(a_{t-1}) b_{t-1}^{-a_{t-1}} B(wa_{t-1}, (1-w)a_{t-1})} \frac{\Gamma((1-w)a_{t-1})}{b_{t-1}^{(1-w)a_{t-1}}} \exp(-wb_{t-1}\lambda_t) =$$

$$\frac{w^{wa_{t-1}-1+1} (\lambda_t)^{wa_{t-1}-1} \Gamma((1-w)a_{t-1})}{\Gamma(a_{t-1}) b_{t-1}^{-wa_{t-1}} \Gamma(wa_{t-1}) \Gamma((1-w)a_{t-1}) [\Gamma(a_{t-1})]^{-1}} \exp(-wb_{t-1}\lambda_t) = \frac{\lambda_t^{wa_{t-1}-1} \exp(-wb_{t-1}\lambda_t)}{(wb_{t-1})^{-wa_{t-1}} \Gamma(wa_{t-1})};$$

$$\lambda_t > 0.$$

Prova 1b:

Demonstração da distribuição preditiva da FGMD:

$$p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = \int_0^{\infty} p(y_t | \mu_t, \boldsymbol{\varphi}) p(\mu_t | \mathbf{Y}_{t-1}) d\mu_t = \int_0^{\infty} \left[a(y_t, \boldsymbol{\varphi}) \mu_t^{b(y_t, \boldsymbol{\varphi})} \exp(-\mu_t c(y_t, \boldsymbol{\varphi})) \right]$$

$$\left[\frac{\mu_t^{a_{i|t-1}^*} \exp(-b_{i|t-1}^* \mu_t)}{\Gamma(a_{i|t-1}^*) (b_{i|t-1}^*)^{-a_{i|t-1}^*}} \right] d\mu_t = \frac{a(y_t, \boldsymbol{\varphi})}{\Gamma(a_{i|t-1}^*) (b_{i|t-1}^*)^{-a_{i|t-1}^*}}$$

$$\int_0^\infty \left[\mu_t^{b(y_t, \boldsymbol{\varphi}) + a_{i|t-1}^* - 1} \exp\left(-\mu_t (c(y_t, \boldsymbol{\varphi}) + b_{i|t-1}^*)\right) \right] d\mu_t = \frac{\Gamma\left(b(y_t, \boldsymbol{\varphi}) + a_{i|t-1}^*\right) a(y_t, \boldsymbol{\varphi}) (b_{i|t-1}^*)^{a_{i|t-1}^*}}{\Gamma(a_{i|t-1}^*) \left(c(y_t, \boldsymbol{\varphi}) + b_{i|t-1}^*\right)^{a_{i|t-1}^* + b(y_t, \boldsymbol{\varphi})}};$$

onde $a_{i|t-1}^* = wa_{t-1}$, $b_{i|t-1}^* = wb_{t-1}g(x_t'\boldsymbol{\beta})^{-1}$ e $y_t \in H(\boldsymbol{\varphi})$.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alves, M., Gamerman, D. & Ferreira, M. A. R. (2009). Transfer functions in dynamic generalized linear models. To appear in *Statistical Models*.
- Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control*, 3rd ed. New Jersey: Prentice-Hall.
- Box, G.E.P. & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of American Statistical Association*, **70**, 70-79.
- Box, G. E. P. & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*, 2nd ed. San Francisco: Holden Day.
- Brockwell, P. J. & Davis, R. A. (1996). *Introduction to Time Series and Forecasting*, Springer text in Statistics. New York: Springer.
- Campbell, D. T. & Stanley, J. C. (1968). The Connecticut crackdown on speeding: Time series data in quasi-experimental analysis. *Law and Society Review*, **3**, 33-53.
- Casella, G. & Berger, R. L. (2002). *Statistical Inference*, Duxbury, Pacific Grove.
- Chib, S. & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *American Statistician*, **49**, 327-335.
- Chiogna, M. and Gaetan, C. (2002). Dynamic generalized linear models with application to environmental epidemiology. *Applied Statistics*, **51**, 453-468.
- Cox, D. R. (1981). Statistical analysis of time-series: some recent developments. *Scandinavian Journal of Statistics*, **8**, 93-115.
- Davidson, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their applications*. Wiesbaden: Courses notes.
- de Jong, P. & Penzer, J. (1998). Diagnosing shocks in time series. *Journal of the American Statistical Association*, **93**, 796-806.

- Dobson, A. J. (2002). *An Introduction to Generalized Linear Models*. 2nd edition, Chapman and Hall, New York.
- Doornik, J. A. (1999). *Ox: An Object-Oriented Matrix Language*, 3rd ed. London: Timberlake Consultants Press.
- Durbin, J. and Koopman, S. J. (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives (with discussion). *Journal of the Royal Statistical Society Series B*, **62**, 3-56.
- Durbin, J. & Koopman, S. J. (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, **7**, 1-26.
- Efron, B. & Tibshirani, R. (1993). *An Introduction to the Bootstrap*. London: Chapman and Hall.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U. K. inflation. *Econometrica*, **50**, 987-1008.
- Fahrmeir, L. (1987). Regression models for nonstationary categorical time series. *Journal of Time Series Analysis*, **8**, 147-160.
- Franco, G. C., Koopman, S. J. & Souza, R. C. (1999). Bootstrap tests when parameters of nonstationary time series lie on the boundary of the parameter space. *Brazilian Journal of the Probability and Statistics*, **13**, 41-54.
- Franco, G. C., Santos, T. R., Ribeiro, J. A. & Cruz, F. R. B. (2008). Confidence intervals for hyperparameters in structural models. *Communications in Statistics: Simulation and Computation*, **37**(3), 486-497.
- Fruhworth-Schnatter, S. (1994). Applied state space modelling of non-Gaussian time series using integration-based Kalman filtering. *Statistical Computation*, **4**, 259-269.
- Gamerman, D. & West, M. (1987). An application of dynamic survival models in unemployment studies. *Statistician*, **36**, 269-274.
- Gamerman, D. (1991). Dynamic Bayesian models for survival data. *Applied Statistics*, **40**, 63-79.

- Gamerman, D. (1998). Markov chain Monte Carlo for dynamic generalised linear models. *Biometrika*, **85**, 215-227.
- Gamerman, D. & Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. London: Chapman and Hall.
- Glass, G.V. (1972). Estimating the effects of intervention into a nonstationary time series. *American Educational Research Journal*, **9**(3), 463-477.
- Gelman, A. (1996). *Inference and monitoring convergence*, in *Markov Chain Monte Carlo in Practice*, W.R. Gilks et al., eds., Chapman and Hall, London, pp. 131-140.
- Gilks, W.R., Richardson, S. and Spiegelhalter, D. (1996). *Markov Chain Monte Carlo in practice*, Chapman and Hall, London.
- Gomes, C. H. A. (2006). *Bayesian Estimation with Jeffreys Priors for State Space Models*. Unpublished M.Sc. Thesis. Universidade Federal do Rio de Janeiro (in Portuguese).
- Godolphin, E. J. & Triantafyllopoulos, K. (2006). Decomposition of time series models in state-space form. *Computational Statistics and Data Analysis*, **50**, 2232-2246.
- Granger, C. M. G. & Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, **1**, 15-29.
- Grunwald, G. K., Raftery, A. E. & Guttorp, P. (1993). Time series of Continuous proportions. *Journal of the Royal Statistical Society series B*, **55**, 1, pp. 103-116.
- Hall, P. (1986). On the number of bootstrap simulations required to construct a confidence interval. *The Annals of Statistics*, **14**, 125-129.
- Harrison, P. J. & Stevens, C. F. (1971). A Bayesian approach to short-term forecasting. *Operational Research Quarterly*, **22**, 341-362.
- Harrison, P. J. & Stevens, C. F. (1976). Bayesian forecasting. *Journal of the Royal Statistical Society: series B*, **38**, 205-247.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: University Press.
- Harvey, A. C. & Fernandes, C. (1989). Time Series Models for Count or Qualitative Observations. *Journal of Business and Economic Statistics*, **7**, 4, pp. 407-417.

- Harvey, A. C. & Koopman, S. J. (1992). Diagnostic checking of unobserved components time series models. *Journal of Business and Economic Statistics*, **10**, 377-389.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*, **57**, pp. 97-109.
- Hemming, K. & Shaw, J. E. H. (2002). A parametric dynamic survival model applied to breast cancer survival times. *Applied Statistics*, **51**, 421-435.
- Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted moving averages. ONR Research Memorandum 52, Carnegie Institute of Technology, Pittsburgh, Pennsylvania.
- Jeffreys, H. (1961). *Theory of Probability*. Oxford: Clarendon Press.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.*, **82**, 35-45.
- Kaufmann, H. (1987). Regression models for nonstationary categorical time series: asymptotic estimation theory. *The Annals of Statistics*, **15**, 79-98.
- Kitagawa, G. (1987). Non-Gaussian state-space modelling of nonstationary time series. *Journal of the American Statistical Association*, **82**, 1032-1063.
- Koopman, S. J. (1992). *Diagnostic checking and intra-daily effects in time series models*. Amsterdam: Thesis Publishers.
- Lindsey, J. K. & Lambert, P. (1995). Dynamic generalized linear models and repeated measurements. *Journal of Statistical Planning and Inference*, **47**, 129-139.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953). Equations of state calculations by fast computing machine. *Journal of Chemistry and Physics*, **21**, pp. 1087-1091.
- Migon, H. S. & Gamerman, D. (1999). *Statistical Inference: an Integrated Approach*. London: Arnold.
- Morettin, P. A. & Toloi, C. M. (2004). *Análise de Séries Temporais*. São Paulo: Editora E. Blucher-Associação Brasileira de Estatística.

- Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association*, **55**, 299-305.
- Nelder, J.A. & Wedderburn, R.W.M. (1972). Generalized linear models. *Journal of the Royal Statistical Society A*, 135, 370-384.
- Nervole, M. & Wage, S. (1964). On the optimality of adaptive forecasting. *Management Science*, **10**, 207-229.
- Penzer, J. (2007). State space models for time series with patches of unusual observations. *Journal of Time Series Analysis*, **28**, 629-645.
- Plummer, M., Best, N., Cowles, K. & Vines, K. (2005). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, **6**, 7-11. Software available at <http://www-fis.iarc.fr/coda/>.
- Ravines, R. E. R., Schmidt, A. M. & Migon, H. S. (2008). Revisiting distributed lag models through a Bayesian perspective. *Applied Stochastic Models in Business Industry*, **22**, 193-210.
- Reis, E. A., Salazar, E. & Gamerman, D. (2006). Comparison of sampling schemes for dynamic linear models. *International Statistical Review*, **74**, 203-214.
- R Development Core Team. (2005). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna.
- Ross, S. (2002). *A First Course in Probability*, 6th ed., Prentice Hall.
- Salvador, M. & Gargallo, P. (2004). Automatic monitoring and intervention in multivariate dynamic linear models. *Computational Statistics and Data Analysis*, **47**, 401-431.
- Santos, T. R. & Franco, G. C. (2008). Inference for hyperparameters in structural models under Bayesian and classical perspectives (submitted). *Computational Statistics (Zeitschrift)*.
- Santos, T.R., Franco, G.C. & Gamerman, D. (2009). Comparison of classical and Bayesian approaches for intervention analysis in structural models. Relatório Técnico do LES - UFRJ.
- Shao, J. & Tu, D. (1995). *The jackknife and bootstrap*. New York: Springer-Verlag.
- Shephard, N. & Pitt, M. K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika*, **84**, 653-667.
- Smith, R. L., & Miller, J. E. (1986). A Non-Gaussian State Space Model and Application to Prediction of Records. *Journal of the Royal Statistical Society, Ser. B*, **48**, 79-88.

Spall, J. C. (1988). *Bayesian Analysis of Time Series and Dynamic Models*. New York, Marcel Dekker Inc.

Stoffer, D. S. & Wall, K. D. (1991). Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter. *Journal of the American Statistical Association*, **86**, 1024-1033.

Theil, H. & Wage, S.(1964). Some observations on adaptative forecasting. *Management Science*, **10**, 198-206.

Thombs, L. A. & Schucany, W. R. (1990). Bootstrap prediction intervals for autoregression. *Journal of the American Statistical Association*, **85**, 486-492.

Wei, W. W. S. (1990). *Time Series Analysis - Univariate and Multivariate Methods*. New York: Addison-Wesley.

West, M., Harrison, P. J. & Migon, H. S.(1985). Dynamic Generalized Linear Models and Bayesian Forecasting (with discussion). *Journal of the American Statistical Association*, **81**, 741-750.

West, M. & Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. New York: Springer.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, **6**, 324-342.