

Thiago Rezende dos Santos

INFERÊNCIA, PREVISÃO E SUAVIZAÇÃO
EM MODELOS ESTRUTURAIS GAUSSIANOS
E NÃO-GAUSSIANOS

Belo Horizonte
Março / 2012

INFERÊNCIA, PREVISÃO E SUAVIZAÇÃO EM MODELOS ESTRUTURAIS GAUSSIANOS E NÃO-GAUSSIANOS

Autor: Thiago Rezende dos Santos

Orientadora: Glaura C. Franco

Tese de Doutorado apresentada
ao Departamento de Estatística do
Instituto de Ciências Exatas da UFMG
como parte dos requisitos necessários
para a obtenção do grau de Doutor em Ciências.

Doutorado em Estatística
Departamento de Estatística
Instituto de Ciências Exatas
Universidade Federal de Minas Gerais

Belo Horizonte
Março / 2012

Tese de Doutorado sob o título “*Inferência, Previsão e Suavização em Modelos Estruturais Gaussianos e Não-Gaussianos*”, defendida por Thiago Rezende dos Santos e aprovada em 02 de março de 2012, em Belo Horizonte, Minas Gerais, pela banca examinadora constituída pelos doutores:

Profa. Dra. Glaura C. Franco
EST - UFMG
Orientadora

Profa. Dra. Clélia M. C. Toloí
IME - USP

Prof. Dr. Helio dos Santos Migon
DME - UFRJ

Prof. Dr. Renato Martins Assunção
EST - UFMG

Profa. Dra. Rosângela Helena Loschi
EST - UFMG

RESUMO

Neste trabalho, uma das abordagens existentes para a modelagem de séries temporais, denominada modelos estruturais (ME), é considerada. O objetivo principal desta tese é propor novos procedimentos de inferência, previsão e suavização em ME sob os contextos gaussiano e não-gaussiano. Quando se desconhece o valor verdadeiro dos hiperparâmetros dos ME, uma prática comum consiste em substituí-lo pela sua respectiva EMV (Estimativa de Máxima Verossimilhança). Isso inevitavelmente conduz à subestimação do EQM (Erro quadrático Médio) de suavização e previsão. Na literatura, alguns procedimentos foram propostos para solucionar esse problema em ME, incorporando à incerteza associada à estimação dos hiperparâmetros, tais como os métodos de Hamilton (1986) e Ansley & Kohn (1986). Entretanto, os últimos lançam mão da amostragem da distribuição assintótica do EMV dos hiperparâmetros - a qual é problemática para tamanhos amostrais pequenos. Neste sentido, novos procedimentos, que são variações de métodos usado para este fim, são propostos, usando a técnica *bootstrap*, evitando amostragem da distribuição assintótica. Vale ressaltar que um procedimento inédito de previsão é apresentado. Nos casos de previsão e suavização, a hipótese de normalidade pode ser irrealística, contudo a robustez dos métodos a essa hipótese é averiguada. Estudos de simulação são conduzidos a fim que se possa avaliar o desempenho dessas propostas. Além disso, aplicações a séries reais são feitas de forma a exemplificar os métodos desenvolvidos. A última parte do trabalho destina-se aos ME não-gaussianos. Propõe-se uma nova família de ME não-gaussianos que tem a vantagem, sobre outros procedimentos existentes na literatura, de possuir função de verossimilhança exata. Além da caracterização dessa família de modelos, são apresentados novos modelos, que são casos particulares dessa família, como os modelos Borel-Tanner, Gama Generalizada, Weibull, Pareto, Beta e Rayleigh. São desenvolvidos também procedimentos de suavização e de previsão para essa família e para modelos que condicionalmente pertencem à mesma. Considerando os modelos Poisson e Gama pertencentes a essa família, estudos de simulação são realizados a fim de verificar o desempenho dos estimadores dos parâmetros sob as perspectivas clássica e Bayesiana. Várias aplicações a séries reais em diversas

áreas do conhecimento, tais como Meio-ambiente, Finanças e Confiabilidade de *software*, são realizadas, ilustrando bem a metodologia proposta e desenvolvida.

Palavras-chave: modelos dinâmicos, modelos espaço de estados; filtro de Kalman, *bootstrap*, MCMC.

ABSTRACT

In this work, one of the many procedures for modeling time series, called structural model (SM), is considered. The main objective of this thesis is to propose new procedures of inference, forecasting and smoothing in gaussian and non-gaussian SM. When the hyperparameters of SM are unknown, a common practice is to replace them by their respective MLE. This leads to underestimate of the forecasting and smoothing MSE. In the literature, some procedures were proposed for solving this trouble in SM, incorporating the uncertain related to the hyperparameters estimation, such as Hamilton (1986) and Ansley & Kohn (1986)'s procedures. Although, they use the sampling of the asymptotic distribution of the hyperparameters MLE, which is problematic for small sample sizes. Then, new methods - that are variations of the procedures cited above - are proposed using the bootstrap technique, avoiding the sampling of MLE asymptotic distribution. It is important to highlight that an unpublished new forecasting method is presented and discussed. In the smoothing and forecasting cases, the normality hypothesis can be unrealistic, however the methods are evaluated without the normality hypothesis in order to verify if they work well in this context. Simulation studies are conducted for assessing the performance of these proposals. In addition, applications to real series are made, so that the methods are illustrated. The last part of work is devoted to non-gaussian SM. We propose a new family of non-gaussian SM whose the likelihood function is exact, which is an advantage over other procedures in the literature. Besides the characterization of the family, particular cases of it are also shown, such as the Borel-Tanner, Generalized gamma, Weibull, Pareto, Rayleigh and Beta models. Smoothing and forecasting methods for this family and for models that conditionally belong to it are discussed. Considering the Poisson and Gamma models, simulation studies are performed to verify the performance of estimators for the parameters under the classical and Bayesian perspectives. Several applications to real series in areas of knowledge, such as Environment, Finance and Software Reliability, are analyzed, illustrating the methodology proposed and developed in this thesis.

Keywords: dynamic models, state space models, Kalman filter, bootstrap, MCMC.

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

MEE - Modelos de Espaço de Estados

MNL - Modelo de Nível Local

MTL - Modelo de Tendência Linear Local

MEB - Modelo Estrutural Básico

FK - Filtro de Kalman

MCMC - Markov chain Monte Carlo

n - Número de observações da série

MC - Número de replicações Monte Carlo

EMV - Estimador de máxima verossimilhança

EQM - Erro quadrático médio

B - Número de replicações *bootstrap*

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

FMEVE - Família de Modelos de Espaço de Estados de Verossimilhança Exata

*“E todos estes que aí estão atravancando o meu caminho Eles passarão...
Eu passarinho!”*
(Mário Quintana)

AGRADECIMENTOS

Eu gostaria de agradecer a Deus 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, Israel, Daniel pelo apóio e orações essenciais para que eu continuasse a trilhar esse difícil e tortuoso caminho da academia. Aos amigos da UFMG e/ou UFOP, Ricardo, Thaís, Maria Claudia, Fernandão, Fábio, Emerson, Flávio e Roger. De maneira alguma, eu poderia esquecer dos meus queridos orientadores, Profa. Glaura e Prof. Dani (orientador informal), pois este trabalho não seria realizado sem a imensa contribuição deles. À Profa. Glaura fica a minha incomensurável e eterna gratidão por me acompanhar, orientar desde da graduação e me ajudar nas minhas limitações. Ao Departamento de Estatística da UFMG, pela oportunidade de realizar este estudo. À UFOP e ao Departamento de Matemática da UFOP, por me apoiar no desenvolvimento deste trabalho. À CAPES, CNPq e FAPEMIG, pelo suporte financeiro.

SUMÁRIO

1. Introdução	1
2. Modelos Estruturais	4
2.1 Modelos Estruturais (ME)	5
2.1.1 Definição	5
2.1.2 A Forma de Espaço de Estados	7
2.1.3 Filtro de Kalman	9
2.2 Inferência Clássica em ME	10
2.2.1 Estimador de máxima verossimilhança	11
2.3 <i>Bootstrap</i>	12
2.3.1 <i>Bootstrap</i> não-paramétrico nos resíduos	13
2.3.2 <i>Bootstrap</i> paramétrico	14
3. Suavização em Modelos Estruturais Gaussianos	15
3.1 O suavizador do estado	18
3.2 Suavização: Incorporando a Incerteza dos Hiperparâmetros	19
3.2.1 Procedimento de Hamilton Bootstrap (HaB)	19
3.2.2 Procedimento de Ansley e Kohn <i>bootstrap</i> (AKB)	20
3.2.3 Procedimento de Pfeffermann e Tiller	22
3.3 Estudos de Simulação	23
3.4 Aplicação a uma série temporal real	30
3.4.1 Série Temporal da CHESF	30

3.4.2	Série IPCA	32
4.	<i>Previsão em Modelos Estruturais Gaussianos</i>	36
4.1	Previsão	37
4.2	Intervalos de Previsão para Valores Futuros	38
4.2.1	Procedimento de Hamilton com Amostragem Bootstrap	39
4.2.2	Procedimento de Ansley e Kohn	41
4.2.3	Procedimento de Pfeiffermann e Tiller (PT)	43
4.2.4	Um Novo Procedimento <i>Bootstrap</i> Paramétrico	44
4.3	Estudos de Simulação	46
4.4	Aplicações a séries temporais reais	56
4.4.1	Série do log da incidência de casos de Dengue em BH .	56
4.4.2	Série <i>Mortgages</i>	58
5.	<i>Modelos de Espaços de Estados Não-Gaussianos</i>	65
5.1	Modelos Lineares Dinâmicos Generalizados	67
5.2	Família de modelos de espaço de estados com verossimilhança exata	69
5.2.1	Definição	69
5.2.2	Procedimento de inferência	72
5.2.2.1	Inferência Clássica	72
5.2.2.2	Inferência Bayesiana	75
5.2.3	Previsão	76
5.2.4	Suavização	78
5.2.5	Adequação do modelo e ajuste	79
5.3	Estudos de simulação em ME Não-Gaussianos Poisson e Gama	81
5.3.1	Modelo de Poisson	81
5.3.2	Modelo Gama	84
5.4	Aplicações a séries temporais reais	88

5.4.1	Dados de doença respiratória	91
5.4.2	Série Temporal dos Retornos da Petrobrás	92
5.4.3	Dados de retornos da CEMIG	95
5.4.4	Dados de confiabilidade de software	96
6.	<i>Conclusão e Considerações Finais</i>	106
	<i>Apêndice</i>	109
	<i>Referências bibliográficas</i>	119

1. INTRODUÇÃO

A metodologia de *modelos estruturais* (ME) ou *modelos dinâmicos*, é uma das várias abordagens para a 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 C. 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). As variâncias ou as volatilidades dos termos de erros desses modelos são denominados hiperparâmetros e, quando desconhecidos, devem ser estimados. Existem basicamente dois procedimentos de inferência em ME, o primeiro utilizando a abordagem clássica (Harvey, 1989) e o segundo a abordagem Bayesiana (West & Harrison, 1997). A abordagem clássica usa o estimador de máxima verossimilhança (EMV), obtido através de métodos de otimização numérica. Já a abordagem Bayesiana utiliza os estimadores média, mediana e moda *a posteriori* os quais são aproximados através de métodos *Markov chain Monte Carlo* (MCMC). Um maior detalhamento desses métodos na estimação dos hiperparâmetros em ME podem ser encontrados em Santos (2009).

Além da inferência para os (hiper)parâmetros, em geral, os ME podem ser ajustados visando à previsão para os valores futuros da série e/ou a explicação da mesma através de seus componentes não-observáveis, para os quais podem ser obtidas estimativas suavizadas. Neste estudo, serão abordados alguns métodos de previsão dos valores futuros da série e suavização dos componentes não-observáveis detalhadamente sob o contexto de observações Gaussianas, usando o *bootstrap*. Os principais trabalhos que utilizam a técnica *bootstrap*, visando à previsão de valores futuros incluem Rodriguez & Ruiz (2010a) e Wall & Stoffer (2002). No caso da suavização, pode-se citar os artigos de Pfeiffermann & Tiller (2005) e Queenville & Singh (2000).

Stoffer & Wall (1991) propuseram um procedimento que possibilita aplicar a técnica de *bootstrap* em modelos estruturais. Esse método é valioso, já que, através dele, os intervalos de confiança *bootstrap* para os hiperparâmetros e intervalos previsão *bootstrap* podem ser construídos, além da correção de vícios.

Os ME mais difundidos são construídos baseados nas suposições de independência, homoscedasticidade e normalidade dos erros. Porém, em alguns casos, as hipóteses de homoscedasticidade e normalidade 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) podem ser usados, a fim de que se possa mensurar o impacto desse evento na série. Maiores detalhes podem ser vistos em Santos, Franco & Gamerman (2010).

Já para os casos nos quais a suposição de normalidade é violada, uma solução natural é aplicar 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, 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 enorme interesse devido à sua grande aplicabilidade em diversas áreas do conhecimento.

Assim, a grande contribuição deste trabalho é a proposta de uma classe (família) de modelos (distribuições) denominada Modelos de Espaço de Estados com Verossimilhança Exata (FMEVE), englobando vários modelos não-

gaussianos inclusive alguns que não pertencem à família exponencial. Essa classe permite o cálculo exato da função de verossimilhança, seguindo a idéia de Smith & Miller (1986). Para se fazer inferência sobre o vetor de estados latentes, um procedimento de suavização também é desenvolvido. Além disso, um procedimento de previsão é desenvolvido e justificado. Esta proposta abre leque muito grande de trabalhos futuros e extensões. Trabalhos de mestrado e doutorado estão sendo desenvolvidos a partir dessa família.

Neste estudo, são consideradas no processo de inferência a abordagem clássica, utilizando o estimador de máxima verossimilhança (EMV), e a abordagem Bayesiana, utilizando métodos MCMC para obter as estimativas Bayesianas (EB) aproximadas.

Em resumo, o foco principal desta tese está voltada para o desenvolvimento e/ou aprimoramento de métodos de inferência, previsão e suavização em modelos estruturais gaussianos e não-gaussianos. Neste sentido, a proposta de trabalho se divide em 3 objetivos específicos:

1. Desenvolvimento de procedimentos alternativos com o uso do *bootstrap*, que incorporem a incerteza associada à estimação dos hiperparâmetros, para a estimação do Erro Quadrático Médio (EQM) dos componentes não-observáveis (vetor de estados) na suavização;
2. Desenvolvimento de novos procedimentos para o cálculo de intervalos de previsão em ME Gaussianos, utilizando a técnica *bootstrap*;
3. A proposta de uma família mais geral, que incorpore modelos não-gaussianos.

O trabalho está organizado como segue. No Capítulo 2, um panorama geral sobre os modelos estruturais é feito. Os Capítulos 3 e 4 tratam de métodos de suavização e de previsão em modelos estruturais gaussianos, respectivamente. No Capítulo 5, são abordados os modelos estruturais não-gaussianos. Finalmente, no Capítulo 6, são discorridos as considerações finais e trabalhos futuros.

2. MODELOS ESTRUTURAIS

Os modelos estruturais decompõem a série temporal em termos de seus componentes não-observáveis, os quais têm uma interpretação direta, e são conhecidos também como variáveis latentes, similar aos modelos normais de efeitos mistos. Na inferência Bayesiana, esses modelos são denominados modelos dinâmicos. Essa metodologia é muito flexível, englobando por exemplo dados multivariados ou dados ausentes com relativa facilidade. Este capítulo discorre sobre aspectos gerais dos ME para um melhor entendimento dos mesmos, antes dos objetivos deste trabalho serem abordados.

Em análise de séries temporais, os dados obtidos de um determinado fenômeno possuem uma estrutura de dependência entre suas observações que não pode ser desprezada, ou seja, a ordem em que as observações foram obtidas não pode ser descartada. Uma série temporal pode ser definida informalmente como uma coleção de observações ordenadas no tempo de acordo com Brockwell & Davis (1996). 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. É importante ressaltar que nesta tese serão trabalhadas séries no domínio do tempo cujo intervalo de observação é igualmente espaçado, apesar dos mesmos incorporarem observações irregularmente espaçadas.

Os modelos estruturais são apenas uma das metodologias para a análise de séries temporais, porém existem outros procedimentos como os modelos ARIMA (Box & Jenkins, 1976), os modelos de longa dependência (ARFIMA) (Granger & Joyeux, 1980), os modelos não-lineares ARCH (Engle, 1982), entre outros. Para obter uma explicação mais aprofundada e detalhada 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 & Tolo (2004), este último em português.

Com o intuito de uma melhor explicação sobre a metodologia e o contexto na qual ela está inserida, primeiramente são apresentados os modelos

estruturais com a forma de espaço de estados, o filtro de Kalman e alguns modelos estruturais especiais. Os procedimentos de inferência também são mostrados, assim como o método computacional *bootstrap*.

2.1 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 Filtro de Kalman (FK) e sua variância, que são utilizados diretamente na construção da função de verossimilhança.

2.1.1 Definição

Uma série temporal pode ser decomposta em componentes não-observáveis. 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:

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

sendo $\epsilon_t \sim N[0, \sigma_\epsilon^2]$, independentes entre si.

Neste estudo, abordam-se os modelos de nível local (MNL), tendência linear local (MTL) e estrutural básico (MEB), os quais serão descritos bre-

vemente. Mais detalhes sobre esses modelos podem ser encontrados em West & Harrison (1997) e Harvey (1989).

Modelo 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:

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

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

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

Para que esse modelo incorpore uma estrutura autoregressiva, acrescenta-se um parâmetro $0 < \rho \leq 1$ tal que a equação do nível assuma a forma $\mu_t = \rho\mu_{t-1} + \eta_t$.

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:

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

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

Modelo 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 defasagens sazonais. Defasagens são múltiplos de um período, que se denotará por s , e que pode ser semanal, mensal, trimestral etc.

Há várias maneiras de se tratar a sazonalidade (Proietti, 2000, ver). Uma delas é modelar a sazonalidade por fatores e, uma outra, por componentes harmônicas ou trigonométricas. A última não é discutida neste estudo; en-

tretanto o leitor interessado pode consultar os livros 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:

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

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 $\omega_t \sim N(0, \sigma_\omega^2)$. A modelagem por fatores sazonais, atualmente, não é a mais preferida (Proietti, 2000, ver), porém é muito utilizada em estudos empíricos.

2.1.2 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):

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

$$\alpha_t = \mathbf{T}_t \alpha_{t-1} + \mathbf{c}_t + \mathbf{R}_t \eta_t, \quad \eta_t \sim N(0, \mathbf{Q}_t) \quad (2.4)$$

onde $t = 1, 2, \dots, n$, ϵ_t são ruídos não correlacionados, η_t é um vetor de ruídos serialmente não-correlacionados, cuja matriz de covariâncias é dada por \mathbf{Q}_t e α_t é o vetor de estados. Assume-se também que ϵ_t e η_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. Os componentes do erro (ruído) podem assumir outras distribuições, não necessariamente distribuições gaussianas. 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 $E(\epsilon_t \eta_j) = \mathbf{0}$ e $E(\eta_t \alpha_0) = \mathbf{0} \quad \forall j, t = 1, \dots, n$. Considera-se a seguinte notação $E(\alpha_0) = \mathbf{a}_0$ e $Cov(\alpha_0, \alpha_0) = \mathbf{P}_0$, onde α_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).

2.1.3 Filtro de Kalman

O vetor de estados (α_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 $t - 1$, $E(\alpha_0) = \mathbf{a}_0$ e $Cov(\alpha_0, \alpha_0) = \mathbf{P}_0$, supondo-se que as observações têm distribuição normal e utilizando as propriedades da normal multivariada, tem-se que:

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

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 ν_t o erro de previsão um passo à frente ou inovação

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

Assim, $E(\nu_t) = 0$ e $Var(\nu_t) = F_t$. Então, substituindo ν_t em (2.5-2.6) tem-se

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

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 z'_t , T_t e R_t da forma de espaço de estados são iguais a um,

$$z'_t = T_t = R_t = 1.$$

Além disso, $c_t = d_t = 0$ e as variâncias dos resíduos ϵ_t e η_t são $h_t = \sigma_\epsilon^2$ 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:

$$\begin{aligned} \nu_t &= y_t - a_{t|t-1}, \\ \mathbf{a}_{t+1|t} &= \mathbf{a}_{t|t-1} + K_t \cdot \nu_t, \\ F_t &= P_{t|t-1} + \sigma_\epsilon^2, \\ K_t &= P_{t|t-1} / F_t, \\ P_{t+1|t} &= P_{t|t-1} - K_t^2 F_t + \sigma_\eta^2. \end{aligned}$$

O FK tem outras características além da predição como, por exemplo, a suavização.

2.2 Inferência Clássica em ME

Seja $\psi = (\psi_1, \dots, \psi_p)'$ o vetor de hiperparâmetros. Por exemplo, os vetores paramétricos dos modelos MNL, MTL e MEB são $\psi = (\sigma_\eta^2, \sigma_\epsilon^2)'$, $\psi = (\sigma_\eta^2, \sigma_\xi^2, \sigma_\epsilon^2)'$ e $\psi = (\sigma_\eta^2, \sigma_\xi^2, \sigma_\omega^2, \sigma_\epsilon^2)'$, 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 $f(y_1, \dots, y_n; \psi)$ em relação a ψ .

2.2.1 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:

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

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

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

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

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

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

O vetor $\hat{\psi}$ 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.

Além da estimação pontual, os intervalos de confiança assintóticos (estimação intervalar) podem ser construídos para os hiperparâmetros, baseados nas propriedades assintóticas do EMV de ψ . Detalhes sobre esses intervalos

podem ser encontrados em Franco, Santos, Ribeiro & Cruz (2008), Santos (2009) e Harvey (1989).

Aspectos relacionados a inferência Bayesiana sobre os hiperparâmetros dos ME como estimadores Bayesianos pontuais e intervalares, distribuições *a priori*, bem como métodos MCMC (por exemplo o algoritmo de Metropolis-Hastings) podem ser consultados em Santos (2009) e Migon et al. (2005).

2.3 Bootstrap

O *bootstrap* (Efron, 1979) é um método computacional desenvolvido principalmente 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 (Efron & Tibshirani, 1993).

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 $\mathbf{T}^*(\mathbf{X}^{*b})$ 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 exemplo, o *bootstrap em bloco* (Efron & Tibshirani, 1993, p.99) - que permite a reamostragem diretamente nas observações da série, entretanto requer que a mesma seja estacionária. 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.

2.3.1 *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 η_t da equação da medida e transição, respectivamente. Para tanto, tem-se que executar o FK para obter os valores das inovações estimadas, $\hat{\nu}_t$, 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, $\hat{\nu}_t$, 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, as mesmas média e variância (Stoffer & Wall, 1991). Ou seja,

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

$t = 1, 2, \dots, n$, onde $\bar{\hat{\nu}}_t = \frac{\sum_{j=1}^n \hat{\nu}_j}{n}$. Agora amostras podem ser tomadas, com reposição, de $\hat{e}_t, t=1, 2, \dots, n$, para se 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

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

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

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

$$\mathbf{S}_t = \mathbf{A}_t \mathbf{S}_{t-1} + \mathbf{B}_t e_t, \quad (2.10)$$

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

Portanto, resolvendo a equação (2.10) com \hat{e}_t^* no lugar de e_t e utilizando os valores estimados F_t e \mathbf{K}_t , obtém-se a série *bootstrap* y_t^* , $t = 1, 2, \dots, n$.

2.3.2 *Bootstrap paramétrico*

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{e}_t^* e $\hat{\eta}_t^*$ são obtidos e a série temporal *bootstrap* y_t^* pode ser construída através das equações:

$$y_t^* = \mathbf{z}'_t \alpha_t^* + d_t + \hat{e}_t^*,$$

$$\alpha_t^* = \mathbf{T}_t \alpha_{t-1}^* + \mathbf{c}_t + \mathbf{R}_t \hat{\eta}_t^*,$$

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

Para uma discussão mais aprofundada sobre os intervalos de confiança *bootstrap*, sobre o número de replicações *bootstrap*, o leitor pode consultar Santos (2009).

3. SUAVIZAÇÃO EM MODELOS ESTRUTURAIS GAUSSIANOS

Além de encontrar as estimativas para os hiperparâmetros dos modelos estruturais, a obtenção de estimativas suavizadas dos componentes não-observáveis dos modelos estruturais constitui um tópico muito importante devido ao fato desses componentes terem uma interpretação direta e serem baseados em toda informação disponível. Há vários trabalhos aplicados em diversas áreas nos quais as estimativas desses componentes têm um grande impacto, como estudos sobre a inflação, taxas de fertilidade, demanda de energia elétrica, ver referências em Rodriguez & Ruiz (2010b), que trabalham com as estimativas de predição ao invés de estimativas suavizadas. As novidades deste capítulo são propostas de procedimentos e/ou variações que incorporem a incerteza associada à estimação de ψ através do *bootstrap*, bem como a utilização do algoritmo Suavizador do estado (ver Seção 4.2).

O filtro de Kalman (FK) é uma ferramenta muito empregada para estimar o vetor de estado, α_t , utilizando a informação disponível até o tempo t , Y_t (atualização), ou até o tempo $t - 1$ (previsão). Um outro procedimento, que também pode utilizar o FK com o mesmo objetivo, porém com a diferença de levar em conta todas as observações da amostra, Y_n , é conhecido como suavização. Na verdade, suavização, atualização e previsão são denominações para os casos em que o instante de tempo t é menor, igual ou maior que o número de observações disponíveis, respectivamente.

Há três algoritmos de suavização ligados diretamente com o FK em modelos lineares, a saber, método de ponto fixo, o de “lag” fixo e o do intervalo fixo. Uma exposição mais detalhada destes métodos pode ser encontrada em Harvey (1989). Neste trabalho, será abordado o método do suavizador do estado (de Jong, 1988, 1989). Ver Tsimikas & Ledolter (1997) para maiores detalhes sobre melhoramentos nestes algoritmos. Esses algoritmos são executados a fim de se calcular os estimadores suavizados do estado cuja forma é apresentada a seguir.

Em modelos gaussianos lineares, o estimador suavizado do vetor de es-

tados é denotado por $\mathbf{a}_{t|n}$. Então,

$$\mathbf{a}_{t|n} = E(\alpha_t | Y_n, \psi) = E_n(\alpha_t). \quad (3.1)$$

Assim como o estimador da filtragem $\mathbf{a}_{t|t-1, \psi}$, $\mathbf{a}_{t|n, \psi}$ é também o estimador de α_t de EQM mínimo, baseado em toda informação disponível. Se a suposição de gaussianidade (normalidade) é desconsiderada, ele é o BLUP (*Best Linear Unbiased Predictor*) de α_t (Harvey, 1989, ver).

A matriz de covariâncias de α_t condicional a todas as n observações é denotada por

$$P_{t|n}(\psi) = EQM_{t|n, \psi} = E_n \left[(\alpha_t - \mathbf{a}_{t|n, \psi}) (\alpha_t - \mathbf{a}_{t|n, \psi})' \right]. \quad (3.2)$$

Sendo $\mathbf{a}_{t|n, \psi}$ um estimador de α_t , $P_{t|n, \psi} = EQM_{t|n, \psi}$ é a sua matriz do EQM.

Caracterização do problema:

Quando os (hiper)parâmetros do modelo têm que ser estimados, $P_{t|n, \hat{\psi}}$ irá subestimar o valor verdadeiro da matriz do EQM na Equação (3.2) (Harvey, 1989), uma vez que a incerteza devida à estimação de ψ não é incorporada. Os valores $\mathbf{a}_{t|n, \hat{\psi}}$ e $P_{t|n, \hat{\psi}}$ são denominados estimadores naturais de α_t e do seu EQM verdadeiro, respectivamente. Esses estimadores podem ser obtidos usando qualquer algoritmo de suavização.

Devido a esse problema de subestimação do EQM verdadeiro, várias propostas e maneiras de corrigir essa subestimação foram feitas. Destacam-se os métodos de Hamilton (1986), Ansley & Kohn (1986) e Queenville & Singh (2000), os quais lançam mão das propriedades assintóticas do EMV de ψ .

Incorporar a incerteza devido à estimação dos hiperparâmetros usando a distribuição assintótica do EMV pode ser uma aproximação pobre principalmente para pequenas amostras. Por isso, os métodos *bootstrap* surgem como uma alternativa plausível para problemas desse porte. A abordagem *bootstrap* adotada por Pfeffermann & Tiller (2005) é adequada para o contexto de pequenas amostras, porém é muito demorada em termos de tempo computacional e implementação, assim como algumas propostas Bayesianas que usam métodos MCMC - os quais podem ser muito dispendiosos (Harvey, 2000; Pfeffermann & Tiller, 2005; Rodriguez & Ruiz, 2010b, ver).

A nossa meta, neste capítulo, está voltada para a incorporação da incerteza da estimação dos (hiper)parâmetros, de forma que seja corrigida a subestimação do valor verdadeiro do EQM de suavização do vetor de estados,

mesmo sob o contexto de pequenas amostras. As nossas propostas de procedimento - que são variações de propostas anteriores na literatura - levam em conta três hipóteses, nesta ordem de importância:

- a correção da subestimação do EQM verdadeiro;
- o uso do *bootstrap* para o problema de pequenas amostras;
- custo computacional pequeno.

É importante propor procedimentos considerando os itens mencionados acima. Embora haja várias sugestões e propostas para a construção de estimadores do EMQ de suavização do vetor de estados levando em conta a incerteza devida à estimação (Harvey, 1989; Pfeiffermann & Tiller, 2005; Rodriguez & Ruiz, 2010b), sob o conhecimento do autor, não há nenhum trabalho que utilize esses procedimentos aplicados diretamente na suavização, usando o algoritmo Suavizador do estado e o *bootstrap*. Apenas o procedimento de Pfeiffermann & Tiller (2005) utiliza a técnica *bootstrap* para incorporar a incerteza de $\hat{\psi}$ e não utiliza o suavizador do estado.

Portanto, os objetivos deste capítulo são:

- Implementar e comparar os procedimentos de Hamilton (Ha), Ansley e Kohn (AK), Pfeiffermann e Tiller (PT) que incorporam a incerteza devida à estimação dos hiperparâmetros através do *bootstrap*, usando o algoritmo Suavizador do estado.
- Propor variações dos procedimentos acima mencionados, levando em conta as hipóteses acima citadas;
- Estudar o comportamento desses métodos em amostras finitas;
- Aplicar esses procedimentos a dados reais.

Vale destacar que duas propostas inéditas de correção dos estimadores do EQM do vetor de estados são apresentadas. A primeira é feita baseada em uma proposta de Rodriguez & Ruiz (2010b) para a predição do vetor de estados um passo à frente com algumas modificações (Procedimento 1, HaB). A segunda, é uma extensão do método de Ansley & Kohn (1986) usando o método *bootstrap* para aproximar a matriz informação de Fisher de ψ (Procedimento 2, AKB). Além disso, essas propostas inéditas podem ser estendidas facilmente para atualização e previsão.

A seguir, primeiramente, é discorrido sobre o método suavizador do estado.

3.1 O suavizador do estado

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 três métodos clássicos para calcular os estimadores suavizados do estado $\mathbf{a}_{t|n,\psi} = E(\alpha_t|\mathbf{Y}_n, \psi)$ e sua respectiva matriz de covariâncias $\mathbf{P}_{t|n,\psi} = Var(\alpha_t|\mathbf{Y}_n, \psi)$, a saber, o método de ponto fixo, o de "lag" fixo e o do intervalo fixo. Uma exposição mais detalhada destes métodos pode ser encontrada em Harvey (1989). Aqui trabalha-se com um algoritmo de suavização melhorado, desenvolvido por de Jong (1988, 1989) e Ansley & Kohn (1986) e que segue imediatamente do método do intervalo fixo desenvolvido por Anderson & Moore (1979), mas com maior eficiência computacional. De acordo com a definição dada em Koopman (1992), esses estimadores são chamados de suavizadores do estado. Pode-se mostrar que $(\alpha_t|\mathbf{Y}_n) \sim N(\mathbf{a}_{t|n}, P_{t|n})$. Os momentos $\mathbf{a}_{t|n}$ e $P_{t|n}$ podem ser encontrados recursivamente através das seguintes equações:

$$\begin{aligned} \mathbf{r}_{t-1} &= \mathbf{z}_t F_t^{-1} \nu_t + \mathbf{L}'_t \mathbf{r}_t, \\ \mathbf{N}_{t-1} &= \mathbf{z}'_t F_t^{-1} \mathbf{z}_t + \mathbf{L}'_t \mathbf{N}_t \mathbf{L}_t, \\ \mathbf{a}_{t|n} &= \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{r}_{t-1}, \\ \mathbf{P}_{t|n} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{N}_{t-1} \mathbf{P}_{t|t-1}, \end{aligned} \quad (3.3)$$

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 ν_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}_n e \mathbf{N}_n podem ser fixados em zero. Analogamente ao que foi exposto na atualização, aqui também $\mathbf{a}_{t|n}$ é um vetor de estimativas dos componentes de estados.

Um intervalo de confiança de nível $1 - \kappa$ para α_t ($t = 1, \dots, n$) é dado por:

$$\left[\mathbf{a}_{t|n} \pm z_{\kappa/2} \sqrt{\mathbf{P}_{t|n}} \right], \quad (3.4)$$

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

Como já dito anteriormente, dado o valor de $\hat{\psi}$, denotam-se $\mathbf{a}_{t|n}$ e $\mathbf{P}_{t|n}$ como estimadores naturais (N) do estado α_t e de seu EQM, respectivamente.

3.2 Suavização: Incorporando a Incerteza dos Hiperparâmetros

Nesta seção, são discutidas formas de incorporar a incerteza devida à estimação dos hiperparâmetros nos estimadores do EQM suavizado do estado, dado pela Equação (3.2). Segundo Harvey (1989, pag. 151), quando os hiperparâmetros do modelo têm que ser estimados, $\mathbf{P}_{t|n}$ subestimar o valor verdadeiro da variância ou do EQM de $\mathbf{a}_{t|n}$.

3.2.1 Procedimento de Hamilton Bootstrap (HaB)

A seguir, é proposto um procedimento que leva em conta a incerteza associada à estimação de ψ na estimação do EQM de suavização do vetor de estados (o próprio procedimento de Hamilton inclui um termo respectivo à incerteza da estimação de ψ), usando o suavizador do estado com o procedimento de Hamilton, sendo que, ao invés de usar a distribuição assintótica do EMV de ψ , é usada a distribuição empírica *bootstrap* de $\hat{\psi}$.

Procedimento 1:

1. Obtenha a EMV de ψ , baseada na série original y_t ;
2. Gere as B séries *bootstrap* (y_t^*) (não-)parametricamente de posse de $\hat{\psi}$. Esse procedimento de geração está bem detalhado na Seção 2.3;
3. Obtenha a EMV $\hat{\psi}^{*(i)}$ baseada nas séries *bootstrap* $y_t^{*(i)}$, para $i = 1, \dots, B$;
4. De posse de $\hat{\psi}^{*(1)}, \dots, \hat{\psi}^{*(B)}$ e usando a série original y_t , faça $i=0$:
 - (a) Faça $i = i + 1$;
 - (b) Obtenha $\tilde{\alpha}_{t|n}^{(i)}$ de (3.3);
 - (c) Se $i < B$, volte para o passo (a); caso contrário, pare.
5. Calcule a estimativa suavizada do vetor de estados α_t como

$$\tilde{\alpha}_{t|n} \cong \frac{1}{B} \sum_{i=1}^B \tilde{\alpha}_{t|n}^{(i)}.$$

6. Calcule e estime o EQM de $\alpha|Y_n$ como

$$\begin{aligned}\widehat{EQM}_{t|n}^{HaB} &= E [(\alpha_t - \tilde{\alpha}_{t|n})^2 | Y_n] \\ &\cong \frac{1}{B} \sum_{i=1}^B \widehat{EQM}(\alpha_t | Y_n, \hat{\psi}^{*(i)}) + \frac{1}{B} \sum_{i=1}^B (\tilde{\alpha}_{t|n}^{(i)} - \tilde{\alpha}_{t|n})^2 \\ &= \frac{1}{B} \sum_{i=1}^B \mathbf{P}_{t|n, \hat{\psi}^{*(i)}} + \frac{1}{B} \sum_{i=1}^B (\tilde{\alpha}_{t|n}^{(i)} - \tilde{\alpha}_{t|n})^2.\end{aligned}$$

Um intervalo de confiança de nível $1 - \kappa$ para α_t ($t = 1, \dots, n$) é dado por:

$$\left[\tilde{\alpha}_{t|n} \pm |z_{\kappa/2}| \sqrt{\widehat{EQM}_{t|n}^{HaB}} \right], \quad (3.5)$$

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

De acordo com esse procedimento, de uma forma simples e rápida se comparado a outros procedimentos, consegue-se incorporar a incerteza devida à estimação de ψ , evitando os problemas ocasionados pela amostragem da distribuição assintótica do EMV de ψ .

Rodriguez & Ruiz (2010b) propuseram um procedimento para a predição do estado um passo à frente um pouco diferente do procedimento acima. Adaptando para o nosso contexto - que é suavização no Passo 6, no cálculo do EQM de $\hat{\alpha}_{t|n}$, considere a estimativa suavizada natural $\mathbf{a}_{t|n}$ de α_t ao invés de considerar a média das estimativas do estado $\tilde{\alpha}_{t|n}$.

3.2.2 Procedimento de Ansley e Kohn bootstrap (AKB)

A seguir, é proposto um procedimento que leva em conta a incerteza associada à estimação de ψ (sugerido por Ansley & Kohn (1986)) na estimação do EQM verdadeiro de suavização do vetor de estados, usando o algoritmo suavizador do estado e a distribuição empírica *bootstrap* de $\hat{\psi}$. Esse procedimento é uma variação do procedimento original de Ansley & Kohn (1986). Denota-se por $I(\psi)$ a matriz de informação de Fisher.

Procedimento 2: _____

1. Obtenha a EMV de ψ , baseada na série original y_t ;

2. Gere as B séries *bootstrap* (y_t^*) (não-)parametricamente de posse de $\hat{\psi}$;
3. Obtenha a EMV $\hat{\psi}^{*(i)}$ baseada nas séries *bootstrap* $y_t^{*(i)}$, para $i = 1, \dots, B$;
4. De posse de $\hat{\psi}^{*(1)}, \dots, \hat{\psi}^{*(B)}$ e usando a série original y_t , faça $i=0$:
 - (a) Faça $i = i + 1$;
 - (b) Obtenha $\tilde{\alpha}_{t|n}^{(i)}$ de (3.3);
 - (c) Se $i < B$, volte para o passo (a); caso contrário, pare.
5. Calcule a estimativa suavizada do vetor de estados α_t como

$$\tilde{\alpha}_{t|n} \cong \frac{1}{B} \sum_{i=1}^B \tilde{\alpha}_{t|n}^{(i)}.$$

6. Calcule a variância de $\hat{\alpha}_{t|n}$ como

$$\widehat{EQM}_{t|n}^{\text{AKB}} = \widehat{EQM}(\alpha_t | Y_n, \hat{\psi}) + \left[\frac{\partial \alpha_t | Y_n, \psi}{\partial \psi} \right]_{\psi=\hat{\psi}}' \left[I^{-1}(\hat{\psi}) \right] \left[\frac{\partial \alpha_t | Y_n, \psi}{\partial \psi} \right]_{\psi=\hat{\psi}},$$

Um intervalo de confiança de nível $1 - \kappa$ para α_t ($t = 1, \dots, n$) é dado por:

$$\left[\tilde{\alpha}_{t|n} \pm |z_{\kappa_0/2}| \sqrt{\widehat{EQM}_{t|n}^{\text{AKB}}} \right], \quad (3.6)$$

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

O procedimento de Ansley & Kohn (1986) incorpora a incerteza da estimação dos hiperparâmetros de uma maneira bem elegante, porém traz as dificuldades (Pfeffermann & Tiller, 2005, ver) do cálculo da matriz de informação de Fisher e da derivada do vetor de estados com respeito à ψ . Pode-se resolver o primeiro problema de duas maneiras:

1. Utilizar uma aproximação numérica rápida estudada e implementada por Franco, Santos, Ribeiro & Cruz (2008);
2. Usar o fato que $p(\hat{\psi}^* | Y_n) \cong p(\hat{\psi} | Y_n)$ (argumentado no próximo capítulo), quando $B \rightarrow \infty$, fazendo $I_n^{-1}(\hat{\psi}) \cong \widehat{Cov}(\hat{\psi}^*)$.

Usando a primeira maneira, o Procedimento 2 é denominado procedimento de Ansley e Kohn *Bootstrap 2* (AKB2); usando a segunda, procedimento de Ansley e Kohn *Bootstrap* (AKB). Com respeito ao segundo problema, o cálculo da derivada $\frac{\partial \alpha_t(\hat{\psi})}{\partial \psi} \Big|_{\psi=\hat{\psi}}$ pode ser feito numericamente.

3.2.3 Procedimento de Pfeiffermann e Tiller

Pfeiffermann & Tiller (2005) propuseram procedimentos que obtêm o EQMP (Erro Quadrático Médio de Previsão) de $a_{t|t-1}$ que incorporam a incerteza dos hiperparâmetros. Eles consideram a decomposição do EQMP, tomando as esperanças sobre todas as realizações de $\{y_1, \dots, y_n\}$ e $\{\alpha_1, \dots, \alpha_n\}$, ao invés das esperanças condicionais sobre todas as observações disponíveis. No caso da suavização, naturalmente, essas esperanças são tomadas sobre todas as observações. Uma vantagem deste procedimento é que Pfeiffermann & Tiller (2005) garantem que o estimador EQMP é de ordem $O(1/n^2)$ sob certas condições, ao contrário dos métodos de Hamilton (1986) e Ansley & Kohn (1986) os quais são de ordem $O(1/n)$. Por isso, espera-se que o método de Pfeiffermann & Tiller (2005) produza estimativas melhores do $\text{EQM}_{t|n, \psi}$. No Procedimento 3, é descrito o procedimento *bootstrap* de Pfeiffermann & Tiller (2005).

Procedimento 3:

1. Gere (não-)parametricamente um número grande de replicações *bootstrap* $y_t^{*(b)}$ ($t = 1, \dots, n$) do modelo na forma de espaço de estados dada em (2.3-2.4) com os hiperparâmetros $\psi = \hat{\psi}$ estimados com a série original;
2. Estime o vetor de hiperparâmetros de cada série *bootstrap* gerada usando o mesmo método utilizado para encontrar $\hat{\psi}$, levando as estimativas $\hat{\psi}^{*b}$ com $b = 1, \dots, B$;
3. Estime $\text{EQM}_{t|n}^{PFTi} = E \left[\alpha_{t|n}(\hat{\psi}) - \alpha_t \right]^2$ como

$$\widehat{\text{EQM}}_{t|n}^{PFTi} = \text{EQM}_{t|n}^{bs} + 2P_{t|n}(\hat{\psi}) - \bar{P}_{t|n}^{bs},$$

em que

$$\text{EQM}_{t|n}^{bs} = \frac{1}{B} \sum_{b=1}^B \left[\alpha_{t|n}^{(b)}(\hat{\psi}^{(b)}) - \alpha_{t|n}^{(b)}(\hat{\psi}) \right]^2$$

e

$$\bar{P}_{t|n}^{bs} = \frac{1}{B} \sum_{b=1}^B P_{t|n}(\hat{\psi}^{(b)}).$$

É importante observar que $\alpha_{t|n}^{(b)}(\hat{\psi})$ e $\alpha_{t|n}^{(b)}(\hat{\psi}^{(b)})$ são as estimativas suavizadas do estado com a série *bootstrap* e usando o algoritmo Suavizador do estado, variando a estimativa de ψ . $P_{t|n}(\hat{\psi}^{(b)})$ é o estimador natural do *EQM* de suavização com a série original, aplicado à estimativa $\hat{\psi}^{(b)}$ - a qual é obtida com a b-ésima série *bootstrap*.

O intervalo de confiança para o vetor de estados pode ser construído de forma similar ao Procedimento 1. Para maiores esclarecimentos, consulte Pfeiffermann & Tiller (2005) e Stoffer & Wall (1991).

3.3 Estudos de Simulação

Para averiguar se os métodos incorporam a incerteza devida à estimação dos hiperparâmetros (ψ) na suavização, geram-se séries temporais via Monte Carlo sob o modelo de nível local (MNL) para vários tamanhos de amostra e razão do sinal-ruído $q = 0, 50$.

Um estudo Monte Carlo é realizado, só para ilustrar como a não incorporação da incerteza de ψ afeta a construção de intervalos de confiança, por exemplo - e apenas o procedimento de HaB é usado para a incorporação da incerteza de ψ . Nas Figuras (3.1-3.2), note que as linhas tracejadas em vermelho, que são os limites de confiança em (3.5) e incorporam a incerteza associada à $\hat{\psi}$, são mais amplas que as linhas pontilhadas cujos limites são dados em (3.4), principalmente para tamanhos de amostra pequenos como $n = 50$ (ver Figura 3.1). À medida que n aumenta, as linhas tracejadas e pontilhadas se aproximam mais (ver Figura 3.2), pois a incerteza com respeito à $\hat{\psi}$ diminui, como já era esperado.

Para averiguar se os métodos incorporam a incerteza devida à estimação dos hiperparâmetros (ψ), realizou-se Monte Carlo de séries temporais sob o modelo de nível local (MNL) com erros gaussianos e não-gaussianos, no caso com distribuição Gama, para vários tamanhos de amostra $n = 40$, $n = 100$ e $n = 500$ e razão do sinal-ruído $q = 0, 50$, isto é, $\sigma_\eta^2 = 0, 50$ e $\sigma_\epsilon^2 = 1, 0$. Esse valor de q foi escolhido, seguindo o mesmo padrão de estudos anteriores (Franco, Santos, Ribeiro & Cruz, 2008). O MNL com erros Gaussianos está definido no Capítulo 2 e o modelo de nível local com erros não-gaussianos

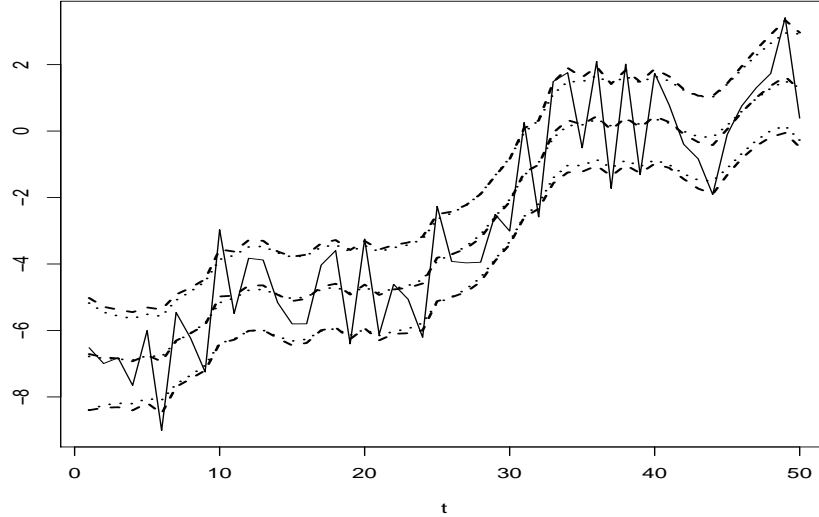


Fig. 3.1: As linhas contínua, tracejadas e pontilhadas representam a série temporal simulada y_t sob o MNL com $n = 50$, as estimativas suavizadas pontuais e intervalares usando o Procedimento de HaB e as estimativas suavizadas pontuais e intervalares usando o procedimento Natural (N), respectivamente. Assumiu-se um nível nominal de 95%.

Gama é dado por:

$$y_t = \mu_t + \sigma_\epsilon \epsilon_t, \quad (3.7)$$

$$\mu_t = \mu_{t-1} + \sigma_\eta \eta_t, \quad (3.8)$$

com $\epsilon_t = \nu_t - \frac{4}{3}$, $\nu_t \sim \text{Gama}(16/9, 3/4)$ e $\eta_t = \omega_t - \frac{5}{4}$, $\omega_t \sim \text{Gama}(25/16, 4/5)$ não-correlacionados para $t = 1, \dots, n$. Esses valores são escolhidos de forma que os erros continuem possuindo média zero e variância 1, porém a distribuição seja assimétrica, ver Figura 3.3. É importante salientar que as replicações Monte Carlo geradas com os erros Não-gaussianos são para a averiguação da robustez dos métodos sob esse contexto. Porém, o procedimento inferencial é feito considerando esse MNL de erros Não-gaussianos como um MNL de erros gaussianos.

Para comparar e avaliar os métodos propostos para a estimação do $EQM_{t|n,\psi}$ é preciso conhecê-lo. Para tanto, simulam-se 10.000 séries de tamanhos $n = 40$, $n = 100$ e $n = 500$, calcula-se o erro quadrático médio verdadeiro da

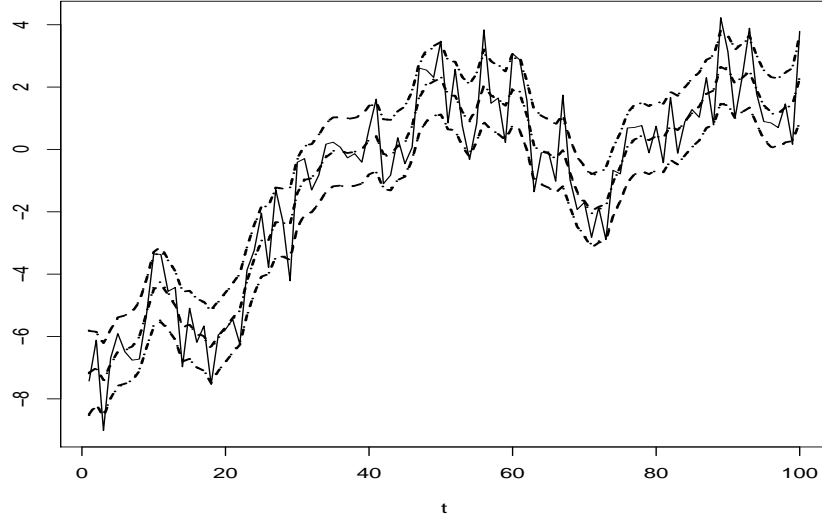


Fig. 3.2: As linhas contínua, tracejadas e pontilhadas representam a série temporal simulada y_t sob o MNL com $n = 100$, as estimativas suavizadas pontuais e intervalares usando o Procedimento de HaB e as estimativas suavizadas pontuais e intervalares o procedimento Natural (N), respectivamente. Assumiu-se um nível nominal de 95%.

suavização ($EQM_{t|n,\psi}$) para cada tempo t , como

$$EQM_{t|n,\psi} \cong \sum_{i=1}^{10.000} \frac{(a_{t|n,\hat{\psi}}^{(i)} - \alpha_t^{(i)})^2}{10.000}, \quad (3.9)$$

para $t = 1, \dots, n$. O erro Monte Carlo é bem pequeno (de ordem 10^{-2}). As medidas de avaliação abaixo, como vício relativo e a raiz quadrada do EQM relativo, dependem do valor de $EQM_{t|n,\psi}$.

Utiliza-se o vício relativo (Vício-Rel) e a raiz quadrada do erro quadrático médio relativo (REQM-Rel), seguindo o mesmo procedimento e sugestão de Pfeffermann & Tiller (2005) - que avaliaram o desempenho do *bootstrap* na estimação do vetor de estados. Denote

$$d_t^{(i)[Proc]} = \left[E\hat{Q}M_{t|n}^{(i)[Proc]} - EQM_{t|n,\psi} \right]$$

o erro na estimação do valor verdadeiro do $EQM_{t|n}$ no tempo t para i -ésima

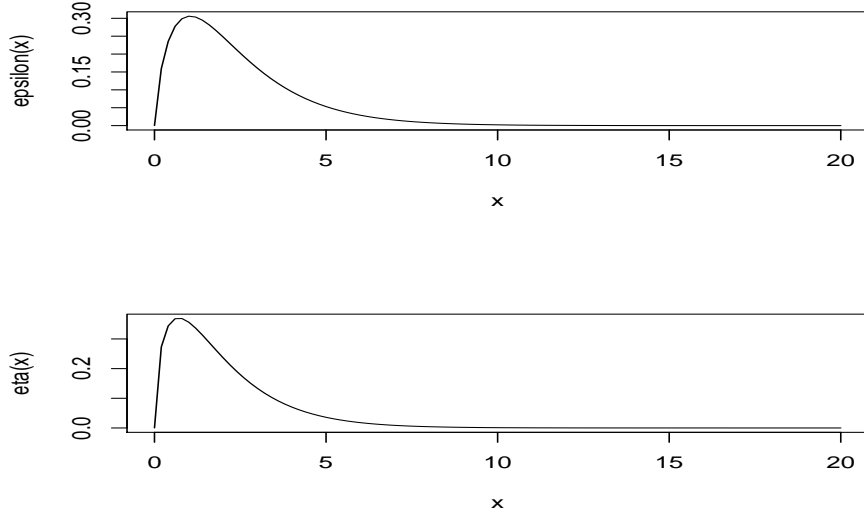


Fig. 3.3: Função densidade de probabilidade da distribuição Gama das equações das observações e de transição, respectivamente.

série pelo estimador $\widehat{EQM}_{t|n}^{(i)[Proc]}$ e sejam

$$\bar{d}_t^{[Proc]} = \sum_{i=1}^{1.000} \frac{d_t^{(i)[Proc]}}{1.000}$$

e

$$\bar{d}_t^{(2)[Proc]} = \sum_{i=1}^{1.000} \frac{d_t^{2(i)[Proc]}}{1.000},$$

em que Proc = N, HaB, AKB2, AKB ou PT.

O Vício-Relativo e o REQM-Relativo são definidos como

$$\text{Vício-Rel}^{[Proc]} = \frac{100}{n} \sum_{t=1}^n \frac{\bar{d}_t^{[Proc]}}{EQM_{t|n}}$$

e

$$\text{REQM-Rel}^{[Proc]} = \frac{100}{n} \sum_{t=1}^n \frac{\left(\bar{d}_t^{(2)[Proc]}\right)^{1/2}}{EQM_{t|n}}.$$

Os métodos a serem comparados são: padrão ou natural (N), o procedimento de Hamilton (HaB) (Procedimento 1), o procedimento de Ansley e Kohn (AKB) e (AKB2) (Procedimento 2) - usando uma aproximação

Tab. 3.1: Vício médio relativo percentual (%) e raiz quadrada do EQM relativo percentual (%) dos estimadores suavizados do $EQM_{t|n,\psi}$ para o modelo de nível local com erros normais e razão sinal-ruído $q = 0, 50$.

Método	n = 40		n = 100		n = 500	
	Vício-Rel	REQM-Rel	Vício-Rel	REQM-Rel	Vício-Rel	REQM-Rel
N	-23,587	21,257	-8,149	9,640	-1,119	3,520
AKB2	-14,048	20,108	-4,199	9,574	-0,316	3,562
HaBNP	-10,372	25,695	-6,874	10,297	-1,097	3,686
AKBNP	-8,669	24,343	-3,672	9,767	-0,295	3,566
PTBNP	1,056	20,613	1,098	9,508	0,531	3,579
HaBP	-8,280	24,502	-6,762	10,088	-1,666	3,513
AKBP	-9,749	21,966	-3,860	9,318	-0,878	3,622
PTBP	-1,236	19,061	0,598	9,305	-0,067	3,554

numérica para matriz informação de Fisher e o *bootstrap*, o procedimento de Pfefferman e Tiller (PT) (Procedimento 3). É aplicado tanto o *bootstrap* não-paramétrico quanto o paramétrico, os quais são denotados pelas siglas (NP) e (P) respectivamente. Os resultados mostrados a seguir para a comparação dos métodos são baseados em 1.000 simulações Monte Carlo e 2.000 replicações *bootstrap*. Todas as rotinas são implementadas no *software* Ox (Doornik, 2002).

A Tabela 3.1 mostra o vício médio relativo percentual e a raiz quadrada do $EQM_{t|n}$ relativo percentual dos estimadores suavizados do $EQM_{t|n}$ para o MNL com erros normais. A razão sinal ruído considerada é $q = 0, 50$. Também é utilizado tanto a técnica *bootstrap* paramétrico (P) quanto *bootstrap* não-paramétrico (NP). O método de Pfefferman e Tiller (PT) apresenta o menor valor do vício médio relativo em ambos métodos *bootstrap* e para todos os tamanhos de amostra. Em seguida, os métodos AKBP e HaBP apresentam os menores e similares valores de vício médio relativo para $n = 40$. Para $n = 100$ e $n = 500$, os métodos AKBP e AKBNP possuem os menores valores, excluindo o método PT. Vale ressaltar que o método de PT requer uma quantidade de operações bem maior que os outros métodos, tendo um custo computacional maior. Observe que o estimador natural (N), apesar de ser mais viciado que os outros, possui REQM-Rel similar ou inferior aos outros estimadores. Isso não é uma surpresa, já que a correção do vício pode aumentar a variância, ver Pfeffermann & Tiller (2005). Note que o Vício-Rel e REQM-Rel diminuem à medida que o tamanho da amostra aumenta.

Na Tabela 3.2 estão o vício médio relativo percentual e a raiz quadrada do EQM relativo percentual dos estimadores suavizados do $EQM_{t|n,\psi}$ para o

Tab. 3.2: Vício médio relativo percentual (%) e raiz quadrada EQM relativo percentual (%) dos estimadores suavizados do $EQM_{t|n,\psi}$ para o modelo de nível local com erros não-gaussianos e razão sinal-ruído $q = 0, 50$.

Método	n = 40		n = 100		n = 500	
	Vício-Rel	REQM-Rel	Vício-Rel	REQM-Rel	Vício-Rel	REQM-Rel
N	-24,952	27,769	-9,600	15,319	-1,578	6,336
AKB2	-13,759	28,486	-5,743	15,149	-0,613	6,064
HaBNP	-13,536	34,921	-10,102	16,990	-1,592	6,269
AKBNP	-10,934	31,445	-4,906	15,621	-0,547	6,084
PTBNP	2,570	31,125	1,147	15,012	0,519	5,987
HaBP	-12,446	31,584	-8,864	16,146	-1,571	6,433
AKBP	-12,902	28,769	-4,784	15,669	-0,625	6,394
PTBP	-2,541	27,665	0,240	15,093	0,347	6,329

MNL com erros Não-gaussianos. A razão sinal-ruído considerada é $q = 0, 50$. Também é utilizado tanto o método *bootstrap* paramétrico (BP) quanto *bootstrap* não-paramétrico (BNP). Os estimadores PT apresentam os menores valores do vício relativo, seguido do estimador AK. Em geral, os estimadores corrigidos têm REQM-Rel maior que o estimador natural (N). Os vícios relativos na Tabela 3.2 são similares aos da Tabela 3.1, evidenciando que os métodos não são sensíveis à suposição de não-gaussianidade. Novamente, o Vício-Rel e REQM-Rel diminuem à medida que o tamanho da amostra aumenta.

Apesar do procedimento PT apresentar um desempenho melhor que os demais métodos, os procedimentos inéditos propostos neste trabalho têm a vantagem da simplicidade e de necessitarem de menos cálculos que o método de PT. Isso pode trazer um ganho significativo quando o modelo estrutural tiver muitos componentes.

Para a comparação dos nossos resultados com os obtidos por Pfeiffermann & Tiller (2005), faz-se um estudo com $q = 0, 25$ com o tamanho amostral $n = 40$ e 100 . Observe que os estimadores propostos que utilizam a técnica *bootstrap* na Tabela 3.3 possuem vício relativo menor e a raiz quadrada do vício relativo menores que os resultados obtidos por Pfeiffermann & Tiller (2005). Enquanto o REQM-Rel está em torno de 35% em Pfeiffermann & Tiller (2005), em nossas propostas está em torno de 25%, para $n = 40$. O estimador de PT é o que tem o melhor desempenho, como já era esperado. Usando o método de suavização Suavizador do estado, o método de PT apresenta REQM-Rel aproximadamente de 19% contra 34% mostrado por Pfeiffermann & Tiller (2005). Para $n = 100$, os valores do vício relativo são equivalentes

Tab. 3.3: Vício médio relativo percentual (%) e raiz quadrada do EQM relativo percentual (%) dos estimadores suavizados do $EQM_{t|n,\psi}$ para o modelo de nível local com erros gaussianos e razão sinal-ruído $q = 0, 25$.

Método	n = 40		n = 100	
	Vício-Rel	REQM-Rel	Vício-Rel	REQM-Rel
N	-23,422	19,413	-8,677	10,315
AKB2	-13,441	18,189	-4,690	10,239
HaBNP	-8,224	24,204	-5,726	10,407
AKBNP	-4,460	28,131	-4,003	11,224
PTBNP	0,110	19,828	0,742	10,108
HaBP	-5,971	23,748	-5,421	10,122
AKBP	-4,659	25,913	-4,326	10,204
PTBP	-1,772	18,782	0,376	9,755

aos apresentados por Pfeffermann & Tiller (2005), porém, novamente, os valores da REQM-Rel em nossas propostas são menores. Enquanto REQM-Rel está em torno de 18% para as propostas deles, em nossas propostas está torno de 10%. É um ganho significativo, que pode ser explicado pelo o uso do algoritmo suavizador do estado. Isto é, essa diferença se deve ao uso de algoritmos de suavização diferentes.

3.4 Aplicação a uma série temporal real

3.4.1 Série Temporal da CHESF

O Modelo de Nível Local (MNL) foi ajustado à série temporal de eletricidade na região Nordeste do Brasil. Esses dados foram obtidos de um grande estudo sobre a quantidade de energia necessária para atender a demanda máxima no intervalo de 18h to 21h. A série é mensal do consumo de eletricidade da CHESF (Companhia Hidroelétrica do São Francisco), é mensal e observada no período de Maio 1991 a Dezembro 1996 ($n = 68$).

Os dados são mostrados na Figura 3.4 e a série não parece apresentar sazonalidade, portanto o MNL pode ser escolhido para modelar os dados.

Os hiperparâmetros foram estimados usando o *software* Ox (Doornik (2002)). As estimativas e os intervalos de confiança estão na Tabela 3.4. Observe que a razão sinal-ruído é menor que 1 (0.56). A análise dos resíduos foi feita e não revelou evidências de inadequação do ajuste do modelo, ver Figura 3.5. Como o limite inferior dos intervalos de confiança *bootstrap* para ψ está distante do valor zero, o MNL pode ser um modelo adequado para essa série.

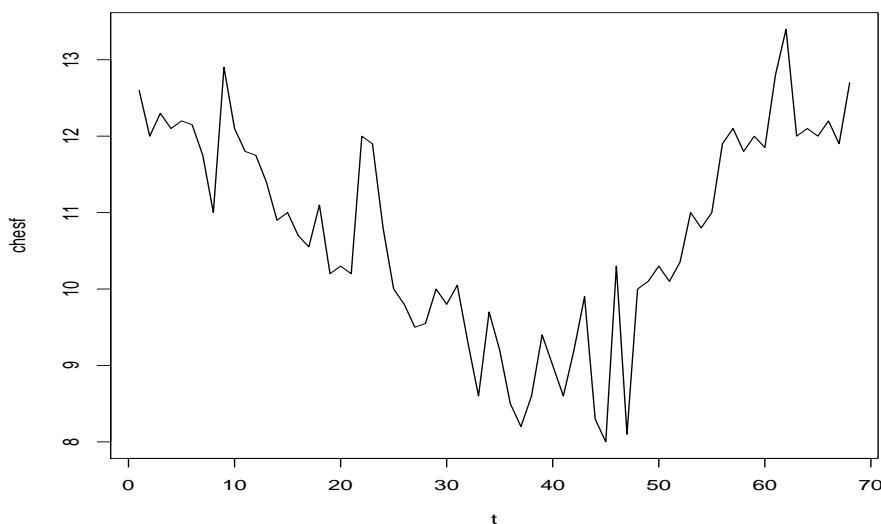


Fig. 3.4: Consumo de Energia elétrica da CHESF. Valores foram divididos por 100.

Tab. 3.4: Estimativas de máxima verossimilhança e intervalos de confiança *bootstrap* para os hiperparâmetros da série da CHESF.

ψ	MLE	Int. Conf. Boot. 95%
σ_η^2	0,14	[0,03; 0,31]
σ_ϵ^2	0,25	[0,11; 0,40]

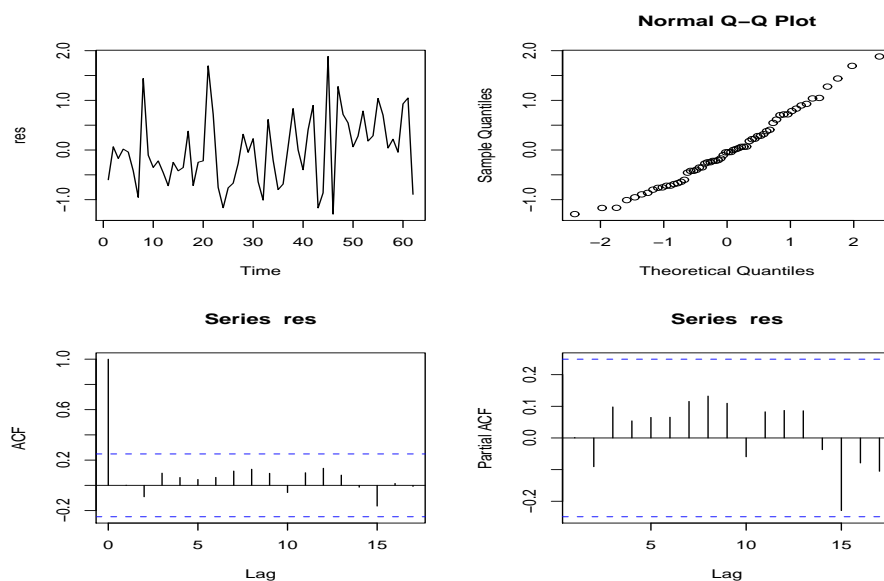


Fig. 3.5: Análise dos resíduos do ajuste do MNL à série da CHESF.

Na Figura 3.6, estão as estimativas do EQM de suavização do vetor de estados para vários métodos. A linha tracejada em negrito representa o método Natural e note que não possui flutuações induzidas pelas variações locais da série. Os métodos PTBNP e AKBNP têm um bom desempenho, pois os seus valores são maiores que os demais métodos para os vários instantes de tempo. Com exceção do estimador de PT, os outros estimadores são mais sensíveis a flutuações da série original. A Tabela 3.5 mostra o percentual médio do aumento do EQM frente ao estimador Natural. Observe que o método de PT tem o maior percentual, enquanto que o método AKB2NP (Ansley Kohn *Bootstrap* 2) possui o menor percentual. Os métodos AKBNP e HaBNP têm um desempenho parecido.

Escolheu-se apenas o método de PTBNP para determinar-se as estimativas suavizadas do estado nesta aplicação por revelar melhores resultados se comparado aos outros métodos ou estimadores. A Figura 3.7 mostra a série da CHESF juntamente com a estimativa suavizada do estado e os limites do

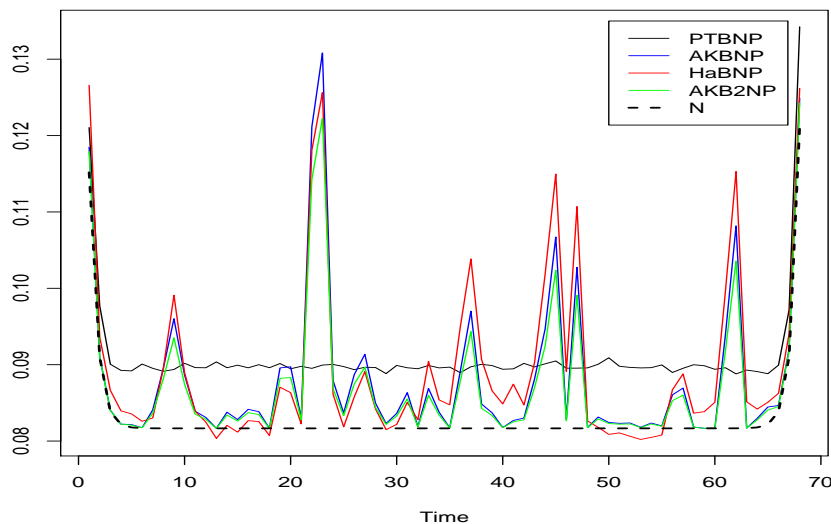


Fig. 3.6: Consumo de Energia elétrica da CHESF. Gráfico das estimativas do EQM de suavização do estado obtido por vários métodos. As linhas em preto, azul, vermelho, verde e tracejada em negrito indicam os métodos PTBNP, AKBNP, HaBNP, AKB2NP e N (Natural), respectivamente.

intervalo de confiança, com nível de 95%, obtidos pelo método de PTBNP. Observe que essas estimativas acompanham bem o comportamento da série.

3.4.2 Série IPCA

Os dados do Índice de Preço ao Consumidor Amplo (IPCA) 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 por 141 observações no período de julho/1997 a março/2009. Esta aplicação ilustra os procedimentos de suavização em uma série de tamanho maior.

Da Figura 3.8, a série IPCA parece apresentar um comportamento de um modelo de nível local (MNL), com uma intervenção em torno de outubro/2002. Esse salto foi devido a temores na economia após a eleição do presidente de esquerda Luís Inácio Lula da Silva. As estimativas do vetor de estado acompanham bem o comportamento da série inclusive nesse evento da eleição do presidente Luís Inácio Lula da Silva. O procedimento HaBNP,

Tab. 3.5: Porcentagem média de aumento do EQM de suavização sobre o método (estimador) Natural para a série temporal da CHESF.

Métodos	Porcentagem de aumento no EQM
PTBNP	9,56
HaBNP	7,71
AKBNP	6,38
AKB2NP	5,27

AKBNP e N foram escolhidos para esta série como ilustração das propostas de incorporação da incerteza da estimação dos hiperparâmetros nos procedimentos de suavização (os intervalos são mais amplos), entretanto não há muita diferença entre os procedimentos de HaBNP, AKBNP e o Natural quando o tamanho de amostra não é pequeno, como foi mostrado anteriormente nos estudos de simulação.

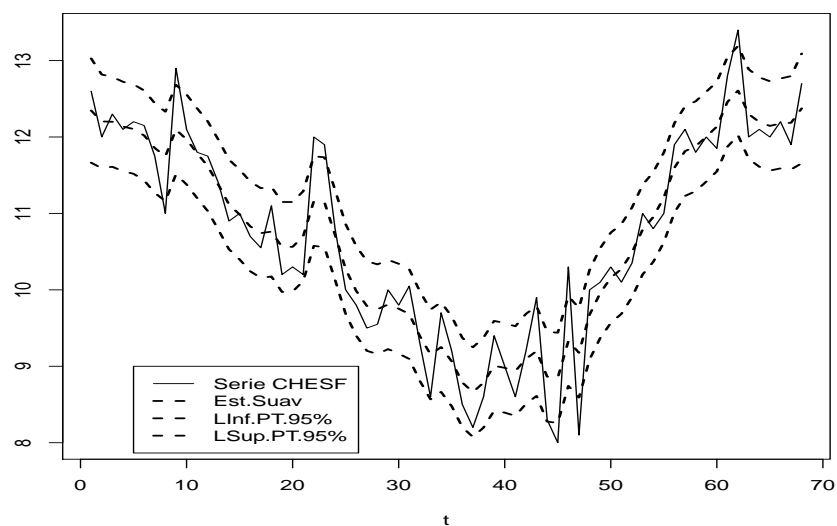


Fig. 3.7: A linha contínua em preto indica a série do Consumo de Energia elétrica da CHESF; a linha tracejada central a estimativa suavizada do estado; as outras linhas tracejadas os limites de confiança com nível de 95%, usando o método de PTBNP (Procedimento 4).

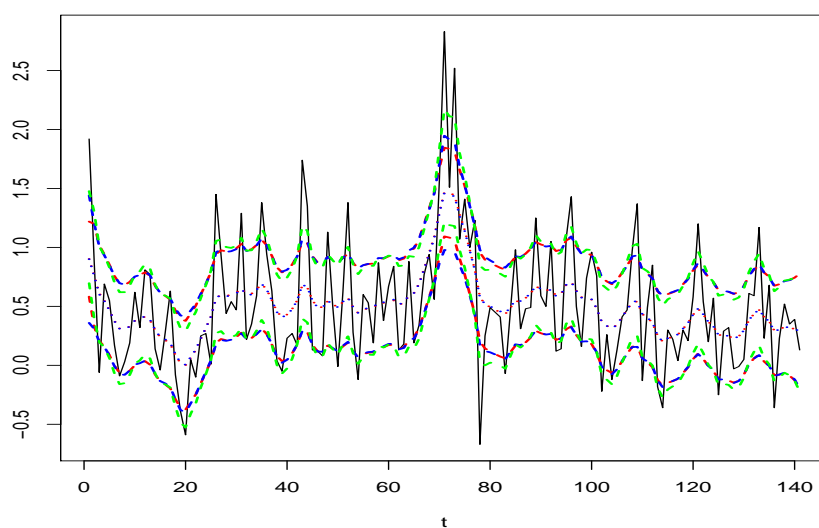


Fig. 3.8: As linhas contínua, tracejadas em azul, em verde e em vermelho representam a série temporal IPCA ($n=141$), as estimativas suavizadas intervalares usando os procedimentos HaBNP, AKBNP e N (Natural), respectivamente. As estimativas suavizadas pontuais são representadas pelas linhas pontilhadas. Assumiu-se um nível nominal de 95%.

4. PREVISÃO EM MODELOS ESTRUTURAIS GAUSSIANOS

Um dos objetivos principais de se modelar séries temporais é prever valores futuros. Portanto, o desenvolvimento das funções de previsão torna-se um importante tópico na análise de séries temporais. Há muitos trabalhos aplicados voltados para a previsão em séries temporais em várias áreas do conhecimento.

No contexto de modelos de espaço de estados, existem alguns trabalhos utilizando o *bootstrap* para a construção de intervalos de previsão para os valores futuros, entre os quais destacam-se Wall & Stoffer (2002) e Rodriguez & Ruiz (2010a). Esses artigos contêm propostas para a previsão de observações futuras utilizando a técnica *bootstrap*. Wall & Stoffer (2002) propõem um método complexo e custoso em termos de implementação computacional para a construção da distribuição empírica do erro de previsão das observações futuras. Já Rodriguez & Ruiz (2010a) apresentam um procedimento *bootstrap* não-paramétrico mais simples se comparado à proposta de Wall & Stoffer (2002), porém sem uma fundamentação teórica bem formulada e justificada. O funcionamento do método é justificado por resultados de simulação Monte Carlo em amostras finitas. Eles empregam uma abordagem diferente utilizando o estimador de quasi-verossimilhança de ψ , que possui algumas propriedades interessantes, como consistência (Harvey, 1989, ver).

Os procedimentos de previsão acima citados incorporam a incerteza de $\hat{\psi}$ usando a técnica *bootstrap*, mas há outras abordagens. Queenville & Singh (2000) propuseram uma correção para os métodos de Hamilton (1986) e Ansley & Kohn (1986) - que incorporam a incerteza de ψ nas estimativas do EQM do vetor de estados, através do estimador corrigido dos hiperparâmetros (ψ) - que também pode ser feito na previsão. Abordagem similar foi adotada por Shephard (1993) que escreveu o modelo de nível local (MNL) como um modelo normal de efeitos mistos a fim de usar a função de verossimilhança restrita e o EMVR (Estimador de Máxima Verossimilhança Restrita). Tsimikas & Ledolter (1994) apresentam formas alternativas de construir a função

de verossimilhança restrita, utilizando modelos de efeitos mistos.

Neste trabalho, a estimação dos parâmetros é feita utilizando-se a função de verossimilhança usual, e o método *bootstrap* é usado para incorporar a incerteza associada à estimação de ψ nos estimadores do EQM de previsão. Além do *bootstrap*, há outras formas de diminuir a incerteza da estimação de ψ nos estimadores do EQM de previsão, buscando estimadores corrigidos para ψ , por exemplo, o EMVR (Estimador de Máxima Verossimilhança Restrita). Em geral, para obter o EMVR de ψ é necessário escrever os modelos de espaço de estados ou modelos estruturais na forma de um modelo linear de efeito mistos, porém isso não é uma tarefa trivial. Um exemplo disso é que todos os trabalhos, sob o conhecimento do autor, fazem estudos de simulação e aplicações só com o modelo estrutural mais simples, o MNL. Uma outra forma é incorporar a incerteza devido à estimação dos hiperparâmetros amostrando a distribuição assintótica do EMV, que pode ser uma aproximação pobre principalmente para pequenas amostras. Por isso, os métodos *bootstrap* surgem como uma alternativa plausível para problemas desse porte.

O objetivo deste capítulo é propor intervalos de previsão que incorporem a incerteza devida à estimação dos (hiper)parâmetros usando a técnica *bootstrap* - que leva em conta a subestimação do EQM de previsão verdadeiro - de maneira mais simples e com custo computacional menor. Basicamente, são feitas propostas e extensões de alguns métodos, que consideram a incerteza dos (hiper)parâmetros para a previsão de valores futuros, baseadas em alguns métodos que foram propostos para estimação do EQMP (Erro Quadrático Médio de Previsão) do vetor de estados α_t , tais como os de Hamilton (1986), Ansley & Kohn (1986) e Pfeiffermann & Tiller (2005).

4.1 Previsão

A previsão de um valor futuro y_{n+k} baseada em toda informação disponível da amostra, \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

$$\begin{aligned} \boldsymbol{\alpha}_{n+k} &= \left(\prod_{i=1}^k \mathbf{T}_{n+i} \right) \boldsymbol{\alpha}_n + \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \mathbf{T}_{n+k-j+1} \right) \mathbf{c}_{n+i} \\ &+ \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \mathbf{T}'_{n+k-j+1} \right) \mathbf{R}'_{n+i} \boldsymbol{\eta}_{n+i}. \end{aligned} \quad (4.1)$$

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

$$\tilde{y}_{n+k|n,\psi} = \mathbf{z}'_{n+k} \left[\left(\prod_{i=1}^k \mathbf{T}_{n+i} \right) \mathbf{a}_n + \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \mathbf{T}_{n+k-j+1} \right) \mathbf{c}_{n+i} \right] + d_{n+k}, \quad (4.2)$$

e o EQM de $(\tilde{y}_{n+k} | \mathbf{Y}_n, \boldsymbol{\psi})$ é dado por

$$\begin{aligned} & \mathbf{z}'_{n+k} \left(\prod_{i=1}^k \mathbf{T}_{n+i} \right) \mathbf{P}_n \left(\prod_{i=1}^k \mathbf{T}'_{n+i} \right) \mathbf{z}_{n+k} \\ & + \mathbf{z}'_{n+k} \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \mathbf{T}_{n+k-j+1} \right) \mathbf{R}_{n+i} \mathbf{Q}_{n+i} \mathbf{R}'_{n+i} \\ & \left(\prod_{j=1}^{k-i} \mathbf{T}'_{n+k-j+1} \right) \mathbf{z}_{n+k} + h_{n+k}. \end{aligned} \quad (4.3)$$

Como exemplo, a previsão k passos à frente para o MNL é dada por

$$\tilde{y}_{n+k|n,\psi} = \mathbf{a}_n$$

e o seu EQM é igual a

$$\mathbf{P}_n + k\sigma_\eta^2 + \sigma_\varepsilon^2.$$

A forma da função previsão abordada acima pode ser denominada como natural ou padrão.

Deve-se observar que a obtenção de $\tilde{y}_{n+k|n,\psi}$ e seu EQM não encerram o processo de inferência, pois eles ainda podem depender de quantidades desconhecidas (como $\boldsymbol{\psi}$) - as quais devem ser estimadas.

A incerteza associada à estimação dessas quantidades desconhecidas conduz inevitavelmente à subestimação do EQM de previsão na abordagem clássica (Harvey, 1989). Por isso, neste estudo, métodos previsão são trabalhados sob essa perspectiva, já que, na abordagem Bayesiana, a incorporação da incerteza é feita de maneira natural (Santos, 2009, ver).

4.2 Intervalos de Previsão para Valores Futuros

Quando se conhece o valor verdadeiro de ψ , sob o contexto de ME gaussianos, o intervalo de previsão de confiança de $1 - \kappa$ $n + k$ passos à frente se torna

$$\left[\tilde{y}_{n+k|n,\psi} \pm |z_{\kappa/2}| \sqrt{\widehat{EQM}(y_{n+k|n,\psi})} \right],$$

onde $|z_{\kappa/2}|$ é o percentil de ordem $\kappa/2$ de uma distribuição Normal-padrão.

O problema acontece quando se desconhece o valor de ψ (muito comum na prática), logo deve-se estimá-lo, como dito anteriormente. A prática mais comum é substituir ψ por $\hat{\psi}$ (Brockwell & Davis, 1996, ver). Quando isto é feito, o intervalo de confiança para previsão, de nível $1 - \kappa$, passa a ser denominado como Natural (N) ou padrão e sua forma é dada por:

$$\left[\tilde{y}_{n+k|n,\hat{\psi}} \pm |z_{\kappa/2}| \sqrt{\widehat{EQM}(\tilde{y}_{n+k|n,\hat{\psi}})} \right], \quad (4.4)$$

onde $|z_{\kappa/2}|$ é o percentil de ordem $\kappa/2$ de uma distribuição Normal-padrão. Observe que o intervalo em (4.4) não incorpora a incerteza associada a ψ . Quando se desconhece ψ , deseja-se trabalhar com $\tilde{y}_{n+k|n}$ ao invés de $\tilde{y}_{n+k|n,\hat{\psi}}$, isto é, incorporar toda incerteza associada a $\hat{\psi}$. Um outro questionamento que pode ser feito quanto ao intervalo de confiança padrão definido em (4.4) é com relação à hipótese de normalidade dos valores futuros, que pode ser irrealística. Por exemplo, se as observações futuras assumem uma distribuição assimétrica (não-normal), a taxa de cobertura desse intervalo padrão pode ser afetada, segundo Rodriguez & Ruiz (2010a). De acordo com os nossos estudos de simulação - os quais serão mostrados mais adiante, sob o contexto de observações muito assimétricas, os intervalos não mantêm taxa de cobertura muito próxima do nível nominal assumido (principalmente em amostras pequenas).

4.2.1 Procedimento de Hamilton com Amostragem Bootstrap

Seguindo a mesma idéia de Hamilton (1986) - o qual incorpora a incerteza da estimação dos hiperparâmetros na estimação do vetor de estados (que também está descrita em Pfeiffermann & Tiller (2005)), pode-se eliminar o efeito da incerteza associada à estimação de ψ na previsão usando integração Monte Carlo. Abaixo, é calculada a distribuição preditiva das observações futuras.

$$\begin{aligned} p(y_{n+k}|Y_n) &= \int p(y_{n+k}, \psi|Y_n) d\psi \\ &= \int p(y_{n+k}|Y_n, \psi) p(\psi|Y_n) d\psi \\ &\cong \frac{1}{MC} \sum_{i=1}^{MC} p(y_{n+k}|Y_n, \psi^i), \end{aligned} \quad (4.5)$$

onde ψ^i são amostras obtidas de $p(\hat{\psi}|Y_n) \cong N(\hat{\psi}, I^{-1}(\hat{\psi}))$.

No caso da média - a qual é uma função da distribuição preditiva -, basta tomar a esperança em ambos lados da igualdade em (4.5) para obter:

$$\tilde{y}_{n+k|n} = \int \tilde{y}_{n+k|n,\psi} p(\psi|Y_n) d\psi \quad (4.6)$$

$$\cong \frac{1}{MC} \sum_{i=1}^{MC} \tilde{y}_{n+k|n,\psi^i}. \quad (4.7)$$

Agora, deseja-se avaliar a variância ou EQM de previsão $E[(y_{n+k} - \tilde{y}_{n+k|n})^2|Y_n]$. Segundo o resultado de Hamilton (1986) e algumas adaptações,

$$\begin{aligned} EQM_{n+k|n}^{Ha} &= E[(y_{n+k} - \tilde{y}_{n+k|n})^2|Y_n] \\ &= EQM_{p(\psi|Y_n)}(y_{n+k}|Y_n) + E_{p(\psi|Y_n)}[(y_{n+k|n,\psi} - \tilde{y}_{n+k|n})^2|Y_n]. \end{aligned} \quad (4.8)$$

Logo,

$$\begin{aligned} EQM_{n+k|n}^{Ha} &\cong \frac{1}{MC} \sum_{i=1}^{MC} \widehat{EQM}(\tilde{y}_{n+k|n}|Y_n, \psi^i) + \\ &\quad \frac{1}{MC} \sum_{i=1}^{MC} (\tilde{y}_{n+k|n,\psi^i} - \tilde{y}_{n+k|n})^2, \end{aligned} \quad (4.9)$$

em que $\tilde{y}_{n+k|n}$ é dado pela equação (4.6), $\tilde{y}_{n+k|n,\psi^i}$ em (4.2) e $E\widehat{QM}(\tilde{y}_{n+k|n}|Y_n, \psi^i)$ em (4.3) e ψ^i são amostras obtidas de $p(\hat{\psi}|Y_n) \cong N(\hat{\psi}, I^{-1}(\hat{\psi}))$.

Pfeffermann & Tiller (2005) argumentam que amostrar ψ^i de $N(\hat{\psi}, I^{-1}(\hat{\psi}))$ pode acarretar em vários problemas e dificuldades, tais como problemas de fronteira, a distribuição de $\hat{\psi}$ ser assimétrica em amostras pequenas e a dificuldade do cálculo da matriz de informação de Fisher. Para contornar esses problemas de maneira simples e eficiente, pode-se utilizar procedimentos *bootstrap* para obter amostras de $\psi^{*(i)}$. A seguir, mostra-se um procedimento, que é uma variação do procedimento de Hamilton discutido acima, utilizando amostragem *bootstrap*.

Procedimento 1: _____

1. Obtenha a EMV ($\hat{\psi}$) de ψ , baseada na série original y_t ;
2. De posse de $\hat{\psi}$, gere B séries *bootstrap* (y_t^*) (não-)parametricamente;

3. Obtenha a EMV $\hat{\psi}^{*(b)}$ baseada nas séries *bootstrap* $y_t^{*(b)}$, para $b = 1, \dots, B$;
4. De posse de $\hat{\psi}^{*(1)}, \dots, \hat{\psi}^{*(B)}$, calcule:
 - $\tilde{y}_{n+k|n}$ dado pela equação (4.6);
 - O EQM ou a variância da previsão dado pela Equação (4.9).
5. O intervalo de previsão de nível $(1 - \kappa) I^{HaB}$ é dado por:

$$\left[\tilde{y}_{n+k|n} \pm |z_{\kappa/2}| \sqrt{\widehat{EQM}_{n+k|n}^{HaB}} \right].$$

De acordo com esse procedimento, de uma forma simples e rápida, consegue-se incorporar a incerteza devida à estimação de ψ , evitando os problemas ocasionados pela amostragem da distribuição assintótica do EMV de ψ .

Validade Assintótica do Procedimento 1:

Se é assumida a forma de espaço de estados com erros gaussianos dada em (2.3-2.4), Stoffer & Wall (1991, pags. 1031-1032) asseguram que $p(\hat{\psi}^*|Y_n) \cong p(\hat{\psi}|Y_n)$, quando B (número de réplicas *bootstrap*) é suficientemente grande. Isto é, a distribuição do EMV gaussiano de ψ é aproximada pela distribuição do EMV *bootstrap* e são equivalentes quando o tamanho amostral é suficientemente grande. Portanto, pode-se substituir a distribuição $p(\hat{\psi}|Y_n)$ por $p(\hat{\psi}^*|Y_n)$ no procedimento acima. Ou seja, é válido amostrar da distribuição do EMV *bootstrap* de ψ ao invés da distribuição assintótica do EMV.

4.2.2 Procedimento de Ansley e Kohn

Ansley & Kohn (1986) propuseram um procedimento para a incorporação da incerteza da estimação de ψ , o qual é trazido para o contexto de previsão de acordo com uma sugestão de Harvey (1989, pag. 151), que consiste em aproximar $EQM(\tilde{y}_{n+k}|Y_n, \psi)$ por $EQM(\tilde{y}_{n+k}|Y_n, \hat{\psi})$ e expandir $\tilde{y}_{n+k}|Y_n, \hat{\psi}$ em torno de $\tilde{y}_{n+k}|Y_n, \psi$ para a aproximação do segundo termo. O estimador do EQM de previsão é

$$\widehat{EQM}_{n+k|n}^{AK} = \widehat{EQM}(\tilde{y}_{n+k}|Y_n, \hat{\psi}) + \left[\frac{\partial \tilde{y}_{n+k}|Y_n, \psi}{\partial \psi} \right]_{\psi=\hat{\psi}}' \left[I^{-1}(\hat{\psi}) \right] \left[\frac{\partial \tilde{y}_{n+k}|Y_n, \psi}{\partial \psi} \right]_{\psi=\hat{\psi}} \quad (4.10)$$

onde $\hat{\psi}$ é o EMV de ψ e $I^{-1}(\hat{\psi})$ é a correspondente matriz de informação de Fisher avaliada em $\hat{\psi}$.

O intervalo de confiança para previsão de nível $(1 - \kappa)$ I^{AK} é dado por:

$$\left[\tilde{y}_{n+k|n,\hat{\psi}} \pm |z_{\kappa/2}| \sqrt{\widehat{EQM}_{n+k|n}^{\text{AK}}} \right]. \quad (4.11)$$

O procedimento de Ansley & Kohn (1986) incorpora a incerteza da estimação dos hiperparâmetros de uma maneira bem elegante, porém traz a dificuldade do cálculo da matriz de informação de Fisher e da derivada da função de previsão com respeito à ψ . Pode-se resolver o primeiro problema de duas maneiras:

1. Utilizar uma aproximação numérica, como a apresentada em Franco, Santos, Ribeiro & Cruz (2008);
2. Usar uma aproximação para a matriz informação de Fisher $I_n^{-1}(\hat{\psi}) \cong \widehat{Cov}(\hat{\psi}^*)$. Essa aproximação é suportada pelo mesmo argumento da validade assintótica do Procedimento 1.

Para o segundo problema, o cálculo de $\frac{\partial \tilde{y}_{n+k|n,\psi}}{\partial \psi} \Big|_{\psi=\hat{\psi}}$ pode ser feito usando derivada numérica e é necessário no cálculo do EQM de previsão em (4.10).

Utilizando a distribuição empírica *bootstrap* do EMV, é proposto o seguinte **procedimento de Ansley e Kohn com amostragem Bootstrap** (AKB). Essa proposta é uma variação da proposta de Ansley e Kohn na qual a matriz de informação de Fisher é aproximada pela matriz de covariâncias da distribuição empírica *bootstrap* do EMV de ψ .

Procedimento 2: _____

1. Obtenha a EMV de ψ , baseada na série original y_t ;
2. Gere B séries *bootstrap* (y_t^*) (não-)parametricamente de posse de $\hat{\psi}$;
3. Obtenha a EMV $\hat{\psi}^{*(i)}$ baseada nas séries *bootstrap* $y_t^{*(i)}$, para $i = 1, \dots, B$;
4. De posse de $\hat{\psi}^{*(1)}, \dots, \hat{\psi}^{*(B)}$, calcule:

- $Cov(\hat{\psi}^*)$;
 - $\tilde{y}_{n+k|n}$ dado pela equação (4.6);
 - O EQM ou a variância estimada da previsão dado pela Equação (4.10).
5. O intervalo de confiança para previsão de nível $(1 - \kappa)$ I^{AKB} é dado por:

$$\left[\tilde{y}_{n+k|n} \pm |z_{\kappa/2}| \sqrt{\widehat{EQM}_{n|n+k}^{AK}} \right].$$

Validade Assintótica do Procedimento 2:

A matriz informação de Fisher $I_n^{-1}(\hat{\psi})$ é substituída pela matriz de covariâncias de $\hat{\psi}^*$, que é válido pelo mesmo argumento do Procedimento 1. A distribuição do EMV de ψ $p(\hat{\psi}|Y_n)$ pode ser aproximada por $p(\hat{\psi}^*|Y_n)$ quando $B \rightarrow \infty$ e são equivalentes quando o tamanho amostral é suficientemente grande.

4.2.3 Procedimento de Pfeffermann e Tiller (PT)

Pfeffermann & Tiller (2005) propuseram um procedimento que estima o EQMP (Erro quadrático Médio de Previsão) que incorpora a incerteza da estimação dos hiperparâmetros. Uma vantagem deste procedimento é que o estimador do EQMP é de ordem $O(1/n^2)$ sob certas condições. Por isso, espera-se que o método de Pfeffermann & Tiller (2005) produza estimativas melhores para o EQM de previsão dos valores futuros. No Procedimento 3 é descrito esse procedimento com o uso do *bootstrap* paramétrico.

Procedimento 3:

1. Gere parametricamente um número grande de replicações *bootstrap* $y_t^{*(b)}$ ($t = 1, \dots, n$) do modelo na forma de espaço de estados dada em (2.3-2.4) com os hiperparâmetros $\psi = \hat{\psi}$ estimados com a série original. Esse procedimento de geração está bem descrito na Seção 2.3.1;

2. Estime o vetor de hiperparâmetros de cada série *bootstrap* gerada usando o mesmo método utilizado para encontrar $\hat{\psi}$, levando às estimativas $\hat{\psi}^{*b}$ com $b = 1, \dots, B$;
3. Estime $\text{EQM}_{n+k|n}^{PFTi} = E \left[\tilde{y}_{n+k|n, \hat{\psi}} - y_{n+k} \right]^2$ como

$$\widehat{\text{EQM}}_{n+k|n}^{PFTi} = \widehat{\text{EQM}}_{n+k|n}^{bs} + 2\widehat{\text{EQM}}(\tilde{y}_{n+k} | \mathbf{Y}_n, \hat{\psi}) - \text{EQM}(\tilde{y}_{n+k}^{bs} | Y_n),$$

$$\text{em que } \widehat{\text{EQM}}_{n+k|n}^{bs}(\tilde{y}_{n+k|n}) = \frac{1}{B} \sum_{b=1}^B \left[y_{n+k}^{*(b)} | Y_n, \hat{\psi}^{*b} - y_{n+k}^{*(b)} | Y_n, \hat{\psi} \right]^2 \text{ e}$$

$$\text{EQM}(\tilde{y}_{n+k}^{bs} | Y_n) = \frac{1}{B} \sum_{b=1}^B \widehat{\text{EQM}}(y_{n+k} | \mathbf{Y}_n, \hat{\psi}^{*b}).$$

É importante observar que $y_{n+k}^{*(b)} | Y_n, \hat{\psi}$ e $y_{n+k}^{*(b)} | Y_n, \hat{\psi}^{*b}$ são as estimativas das observações futuras obtidas com a série *bootstrap*, variando a estimativa de ψ . $\widehat{\text{EQM}}(\tilde{y}_{n+k} | \mathbf{Y}_n, \hat{\psi}^{*b})$ é o estimador natural dado pela Equação (4.3) com a série original, aplicado à estimativa $\hat{\psi}^{*b}$ que é obtida com a b-ésima série *bootstrap*.

O intervalo para previsão de nível $(1 - \kappa)$ I^{PFTi} é dado por:

$$\left[\tilde{y}_{n+k|n} \pm |z_{\kappa/2}| \sqrt{\widehat{\text{EQM}}_{n+k|n}^{PFTi}} \right],$$

onde $\tilde{y}_{n+k|n}$ é dado pela Equação (4.6).

Note que o número de operações desse procedimento é maior que os descritos anteriormente. Para efetuar o procedimento *bootstrap* não-paramétrico, basta alterar os passos 1 e 2 do Procedimento 3, gerando as séries *bootstrap*, usando a forma de inovação e reamostrando os resíduos similarmente como foi descrito na Seção 2.3.1. Para maiores esclarecimentos, consulte Pfeiffermann & Tiller (2005) e Stoffer & Wall (1991).

4.2.4 Um Novo Procedimento Bootstrap Paramétrico

Nesta subseção, é proposto um novo procedimento *bootstrap* paramétrico - que será denominado pela sigla NPB e que incorpora a incerteza devida à

estimação dos hiperparâmetros para a construção dos intervalos de previsão. Apesar desse procedimento ser proposto para observações não necessariamente gaussianas, será assumida a hipótese de que as observações futuras sejam gaussianas. O procedimento funciona bem sob o contexto de modelos de espaço de estados gaussianos e é bem mais simples que o sofisticado procedimento proposto por Wall & Stoffer (2002) e o procedimento de Rodriguez & Ruiz (2010a), preservando boas propriedades assintóticas.

Procedimento 4:

1. Obtenha a EMV de ψ , baseada na série original y_t ;
2. Gere as B séries *bootstrap* (y_t^*) (não-)parametricamente de posse de $\hat{\psi}$;
3. Obtenha a EMV $\hat{\psi}^{*(b)}$ baseada nas séries *bootstrap* $y_t^{*(i)}$, para $b = 1, \dots, B$;
4. De posse de $\hat{\psi}^{*(1)}, \dots, \hat{\psi}^{*(B)}$, execute o FK com série original e obtenha $a_n^{(i)}$ e $\mathbf{P}_n^{(b)}$, para $b = 1, \dots, B$;
5. Gere a previsão *bootstrap* k passos à frente usando as seguintes equações:

- (a) Para $b = 1, \dots, B$, faça

$$\alpha_{n+k}^{*(b)} = \left(\prod_{i=1}^k \mathbf{T}_{n+i}^{*(b)} \right) \alpha_n^{*(b)} + \sum_{i=1}^k \left(\prod_{j=1}^{k-i} T_{n+k-j+1} \right) \mathbf{c}_{n+i} \\ + \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \mathbf{T}'_{n+k-j+1} \right) \mathbf{R}'_{n+i} \eta_{n+i}^{*(b)}, \text{ onde } \eta_{n+i}^{*(b)} \sim F(\mathbf{0}, \hat{\mathbf{Q}}_t^{*(b)}) \text{ e } \alpha_n^{*(b)}$$

é gerado de uma distribuição Normal com média $a_n^{(b)}$ e matriz de covariâncias $\mathbf{P}_n^{(b)} + (a_n^{(b)} - a_n) \times (a_n^{(b)} - a_n)'$.

- (b) $y_{n+k}^{*(b)} = \mathbf{z}_{n+k}' \alpha_{n+k}^{*(b)} + \epsilon_{n+k}^{*(b)}$, onde $\epsilon_{n+k}^{*(b)}$ é gerado de uma distribuição G com média zero e variância $\hat{h}_t^{*(b)}$.

6. O intervalo de confiança *bootstrap* para previsão de nível $(1 - \kappa)$ é dado por:

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

onde y_{n+k}^* é o vetor das previsões *bootstrap* ordenadas das B séries *bootstrap*.

As distribuições F e G não são gaussianas necessariamente, apesar de ser o caso mais recorrente e que será discutido neste estudo. O termo $(a_n^{(b)} - a_n) \times (a_n^{(b)} - a_n)'$ foi introduzido na matriz de covariâncias do item 5 (a) para incorporar a incerteza da estimação de ψ na geração de valores do vetor de estados.

A validade assintótica do Procedimento 4 é garantida pelo Teorema 4.1, considerando a forma de espaço de estados gaussianas em (2.3-2.4) e as distribuições F e G gaussianas.

Teorema 4.1: Supondo a forma de espaço de estados com erros Gaussianos dada em (2.3-2.4) e seguindo os passos do Procedimento 4 com F e G sendo distribuições gaussianas, então

$$y_{n+k}^* | \mathbf{Y}_n \xrightarrow{P} y_{n+k},$$

quando $l \rightarrow \infty$ (l e n são o tamanho amostral e o índice da última observação, respectivamente). Essa troca de índices, neste capítulo, é necessária, pois para estudar as propriedades assintóticas do método, é preciso considerar uma amostra suficientemente grande. Se é usado o índice n como tamanho amostral e fazendo $n \rightarrow \infty$, é inviável determinar o horizonte futuro (as observações futuras que estão além do conjunto de dados). A demonstração na íntegra do Teorema 4.1, encontra-se no Apêndice 1.

4.3 Estudos de Simulação

Nesta seção os resultados obtidos por experimentos Monte Carlo (MC), implementados no *software* Ox (Doornik, 2002), são mostrados. O desempenho dos métodos de previsão Natural (N) (Eq. 4.4), o Procedimento 1 (HaB), o Procedimento 2 de Ansley e Kohn *bootstrap* (AKB), o Procedimento 3 de Pfeiffermann e Tiller (PT) e o Procedimento 4 (NPB) serão comparados quanto à amplitude média e taxa de cobertura do intervalo de previsão. Não serão abordados os procedimentos nos quais é necessário amostrar a distribuição assintótica do EMV dos hiperparâmetros e sim os quais usam distribuição empírica *bootstrap* do estimador.

Os valores dos hiperparâmetros foram escolhidos de forma a obter as seguintes razões sinal-ruído: $q = 0, 1, q = 1$ e $q = 2$, onde $q = \sigma_\eta^2 / \sigma_\epsilon^2$ e $\sigma_\epsilon^2 = 1$. Neste caso, séries temporais de tamanho $n = 50, 100$ e 500 foram geradas com um *burn-in* igual a 100 para o modelo de nível local em dois casos:

-erros normais; erros da equação das observações $\text{Gama}(1/9, 1/3)$ centrada e re-escalada, de forma que os erros tenham média zero e variância 1. A Figura 4.1 apresenta a função densidade de probabilidade dessa distribuição - que é bem assimétrica. O ajuste do modelo é sempre realizado usando o FK e a função de verossimilhança gaussiana independente da distribuição dos erros. A previsão foi feita k passos à frente com $k = 1, 5$ e 15 . Foram fixados $MC = 1000$ (número replicações Monte Carlo), $B = 2000$ (número de reamostras *bootstrap*) e um nível nominal de confiança de 95%.

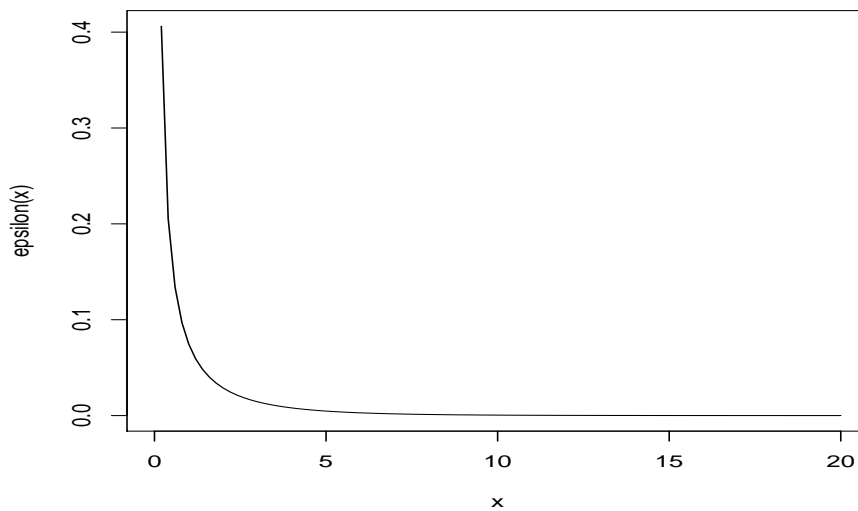


Fig. 4.1: Função densidade de probabilidade da distribuição Gama da equação das observações.

A Tabela 4.1 contém os resultados da simulação Monte Carlo com $q = 0, 1$ e erros gaussianos para diversos métodos, tamanhos de amostra e *lags* (defasagens). Em geral, os métodos AKB, PT, NPB e HaB possuem uma taxa ligeiramente mais próxima do nível nominal de 95%. Note que possuem uma amplitude maior que o método N, principalmente para pequenas amostras, por incorporar a incerteza inerente à estimação de ψ . Porém, essas diferenças desaparecem quando o tamanho amostral aumenta, como já era esperado, uma vez que a incerteza associada à estimação dos hiperparâmetros diminui. O NPB possui um melhor desempenho quando $k = 1$, mas a sua amplitude é um pouco maior se comparado aos outros intervalos. A cobertura nas caudas parece ser simétrica entre todos os métodos.

A Tabela 4.2 contém os resultados da simulação Monte Carlo com $q = 1, 0$

Tab. 4.1: Intervalos de confiança de nível de 95% para previsão k passos à frente com $q = 0, 1$ e erros gaussianos.

n	k	N	NPB	HaB	AKB	PT
		Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas
50	1	4,462	4,707	4,536	4,590	4,601
		0,913	0,933	0,919	0,920	0,922
		0,041/0,046	0,035/0,032	0,038/0,043	0,037/0,043	0,037/0,041
100	1	4,533	4,631	4,574	4,561	4,603
		0,955	0,953	0,954	0,954	0,956
		0,021/0,024	0,019/0,028	0,020/0,026	0,020/0,026	0,020/0,024
500	1	4,582	4,566	4,579	4,586	4,592
		0,949	0,952	0,955	0,954	0,955
		0,022/0,029	0,020/0,028	0,017/0,028	0,016/0,030	0,016/0,029
50	5	5,099	5,340	5,174	5,229	5,215
		0,938	0,942	0,939	0,941	0,941
		0,031/0,031	0,028/0,030	0,030/0,031	0,029/0,032	0,029/0,030
100	5	5,147	5,240	5,180	5,174	5,213
		0,936	0,944	0,938	0,939	0,941
		0,031/0,033	0,025/0,031	0,03/0,032	0,030/0,031	0,030/0,029
500	5	5,214	5,198	5,210	5,218	5,225
		0,950	0,949	0,951	0,951	0,951
		0,023/0,027	0,026/0,025	0,023/0,026	0,023/0,026	0,023/0,026
50	15	6,331	6,606	6,410	6,448	6,406
		0,918	0,926	0,923	0,924	0,919
		0,044/0,033	0,045/0,029	0,047/0,030	0,044/0,032	0,049/0,032
100	15	6,409	6,504	6,429	6,440	6,472
		0,926	0,930	0,928	0,929	0,930
		0,037/0,037	0,037/0,033	0,038/0,034	0,037/0,034	0,037/0,033
500	15	6,485	6,464	6,479	6,489	6,498
		0,946	0,938	0,944	0,946	0,947
		0,028/0,026	0,033/0,029	0,029/0,027	0,028/0,026	0,028/0,025

e erros gaussianos para diversos métodos, tamanhos de amostra e *lags*. Novamente, os intervalos obtidos pelos procedimentos HaB, AKB, PT e NPB apresentaram uma performance superior ao intervalo N principalmente para $q = 0, 10$ e $q = 1, 0$. Os métodos que incorporam a incerteza da estimação possuem uma amplitude maior especialmente para tamanhos de amostra pequenos (por exemplo, $n = 50$). Entretanto, todos os métodos se aproximam bem quando n é grande. A amplitude média dos intervalos é maior que no caso $q = 0, 10$.

A Tabela 4.3 contém os resultados da simulação Monte Carlo com $q = 2$ e erros gaussianos para diversos métodos, tamanhos de amostra e *lags*. Os mesmos comentários feitos para Tabela 4.2 podem ser feitos para esta tabela. O NPB teve um desempenho melhor se comparado a tabela anterior com $q = 1$. As amplitudes dos intervalos maiores que nos casos $q = 0, 1$ e $q = 1, 0$.

A única hipótese que pode ser irrealística é a de normalidade das observações futuras, porém serão feitos exercícios de simulação para averiguar a robustez dos métodos sob o contexto de observações Não-gaussianas. A Tabela 4.4 contém os resultados da simulação Monte Carlo com $q = 0, 1$ e erros Não-gaussianos para diversos métodos, tamanhos de amostra e *lags*. Os intervalos obtidos pelos procedimentos AKB, PT, HaB e NPB mostraram um desempenho superior que o método N, destacando o NPB principalmente para tamanho amostral pequeno. Os métodos que incorporam a incerteza da estimação possuem uma amplitude maior especialmente para tamanhos de amostra pequenos (por exemplo, $n = 50$). Entretanto, todos os métodos se aproximam bem quando n é grande. O NPB tem amplitude maiores que os demais intervalos. Em geral, a taxa de cobertura nas caudas não se manteve simétrica principalmente para $q = 0, 10$.

A Tabela 4.5 mostra os resultados da simulação Monte Carlo com $q = 1$ e erros não-gaussianos para diversos métodos. O intervalo NPB possui uma taxa cobertura mais próxima do nível nominal assumido e uma amplitude um pouco maior se comparado aos intervalos de AKB, PT, HaB e N para $n = 50$. Contudo, os métodos apresentam um desempenho melhor que o método N. Vale destacar que os métodos AK, PT, HaB e NPB tendem a se aproximar quanto a taxa de cobertura e amplitude à medida que o tamanho da amostra cresce. Note que a taxa de cobertura nas caudas, na maioria dos casos mas com o número de casos inferior ao da tabela anterior, é assimétrica mesmo considerando os erros não-Gaussianos e o mesmo não pode ser dito para a Tabela 4.6 com $q = 2$.

A Tabela 4.6 apresenta os resultados da simulação Monte Carlo com $q = 2$ e erros não-Gaussianos para vários intervalos, tamanhos amostrais e passos à

Tab. 4.2: Intervalos de confiança de nível de 95% para previsão k passos à frente com $q = 1, 0$ e erros gaussianos.

n	k	N	NPB	HaB	AKB	PT
		Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas
50	1	6,212	6,398	6,282	6,292	6,381
		0,924	0,933	0,927	0,929	0,932
		0,041/0,035	0,036/0,031	0,038/0,035	0,036/0,035	0,038/0,030
100	1	6,287	6,330	6,292	6,326	6,366
		0,952	0,952	0,952	0,955	0,956
		0,022/0,026	0,023/0,025	0,022/0,026	0,021/0,024	0,021/0,023
500	1	6,326	6,315	6,326	6,333	6,334
		0,949	0,946	0,949	0,949	0,949
		0,020/0,031	0,020/0,034	0,020/0,031	0,020/0,031	0,020/0,031
50	5	9,772	10,015	9,861	9,829	9,845
		0,925	0,933	0,927	0,927	0,929
		0,036/0,039	0,032/0,035	0,035/0,038	0,032/0,041	0,033/0,038
100	5	9,915	9,996	9,956	9,940	9,928
		0,945	0,938	0,944	0,944	0,944
		0,024/0,031	0,026/0,036	0,024/0,032	0,024/0,032	0,024/0,032
500	5	10,035	10,008	10,046	10,040	10,033
		0,949	0,943	0,948	0,948	0,950
		0,026/0,025	0,031/0,026	0,026/0,026	0,025/0,027	0,025/0,025
50	15	15,504	15,836	15,521	15,504	15,517
		0,911	0,913	0,910	0,911	0,913
		0,048/0,041	0,047/0,04	0,048/0,042	0,047/0,042	0,047/0,04
100	15	15,793	15,943	15,830	15,808	15,788
		0,928	0,927	0,925	0,925	0,925
		0,034/0,038	0,036/0,037	0,035/0,04	0,036/0,039	0,037/0,038
500	15	15,881	15,842	15,886	15,884	15,882
		0,955	0,952	0,955	0,955	0,955
		0,021/0,024	0,021/0,027	0,021/0,024	0,021/0,024	0,021/0,024

Tab. 4.3: Intervalos de confiança de nível de 95% para previsão k passos à frente com $q = 2, 0$ e erros gaussianos.

n	k	N	NPB	HaB	AKB	PT
		Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas
50	1	7,435	7,598	7,493	7,520	7,608
		0,926	0,933	0,926	0,929	0,930
		0,04/0,034	0,036/0,031	0,039/0,035	0,039/0,032	0,039/0,031
100	1	7,514	7,560	7,521	7,557	7,597
		0,948	0,950	0,950	0,950	0,951
		0,022/0,030	0,021/0,029	0,021/0,029	0,020/0,030	0,020/0,029
500	1	7,548	7,522	7,549	7,556	7,563
		0,943	0,945	0,943	0,943	0,943
		0,021/0,036	0,019/0,036	0,021/0,036	0,021/0,036	0,021/0,036
50	5	13,017	13,242	13,079	13,070	13,098
		0,934	0,935	0,937	0,939	0,938
		0,032/0,034	0,030/0,035	0,030/0,033	0,028/0,033	0,029/0,033
100	5	13,232	13,363	13,323	13,258	13,196
		0,939	0,939	0,937	0,938	0,940
		0,029/0,032	0,030/0,031	0,030/0,033	0,028/0,034	0,028/0,032
500	5	13,368	13,347	13,410	13,372	13,334
		0,943	0,942	0,943	0,943	0,943
		0,028/0,029	0,028/0,030	0,028/0,029	0,028/0,029	0,028/0,029
50	15	21,418	21,751	21,401	21,449	21,515
		0,915	0,915	0,914	0,914	0,915
		0,048/0,038	0,049/0,036	0,049/0,037	0,048/0,038	0,048/0,037
100	15	21,890	22,129	22,006	21,905	21,804
		0,927	0,929	0,929	0,928	0,926
		0,033/0,040	0,033/0,038	0,033/0,038	0,033/0,039	0,034/0,040
500	15	21,955	21,965	22,019	21,958	21,896
		0,951	0,949	0,952	0,951	0,950
		0,023/0,026	0,023/0,028	0,022/0,026	0,023/0,026	0,024/0,026

Tab. 4.4: Intervalos de confiança de nível de 95% para previsão k passos à frente com $q = 0, 1$ e erros não-gaussianos.

n	k	N	NPB	HaB	AKB	PT
		Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas
50	1	3,731	4,386	3,778	4,241	3,853
		0,879	0,929	0,884	0,897	0,883
		0,062/0,059	0,011/0,060	0,057/0,059	0,045/0,058	0,059/0,058
100	1	4,111	4,502	4,145	4,188	4,174
		0,932	0,968	0,935	0,938	0,936
		0,044/0,024	0,006/0,026	0,041/0,024	0,037/0,025	0,040/0,024
500	1	4,409	4,531	4,407	4,414	4,420
		0,955	0,956	0,955	0,954	0,954
		0,021/0,024	0,008/0,036	0,020/0,025	0,021/0,025	0,021/0,025
50	5	4,418	5,170	4,475	4,857	4,511
		0,863	0,941	0,864	0,882	0,871
		0,094/0,043	0,023/0,036	0,031/0,045	0,073/0,045	0,087/0,042
100	5	4,764	5,168	4,789	4,846	4,820
		0,924	0,959	0,925	0,926	0,926
		0,047/0,029	0,004/0,037	0,038/0,030	0,037/0,031	0,045/0,029
500	5	5,036	5,151	5,035	5,04	5,04
		0,931	0,952	0,932	0,932	0,932
		0,038/0,031	0,009/0,039	0,038/0,030	0,037/0,031	0,037/0,031
50	15	5,592	6,435	5,665	5,944	5,638
		0,836	0,924	0,837	0,847	0,835
		0,129/0,035	0,040/0,036	0,128/0,035	0,119/0,034	0,131/0,034
100	15	6,070	6,562	6,086	6,157	6,122
		0,900	0,940	0,899	0,900	0,897
		0,075/0,025	0,022/0,038	0,075/0,026	0,073/0,027	0,077/0,026
500	15	6,306	6,468	6,301	6,310	6,319
		0,922	0,960	0,922	0,921	0,921
		0,065/0,013	0,016/0,024	0,065/0,013	0,066/0,013	0,066/0,013

Tab. 4.5: Intervalos de confiança de nível de 95% para previsão k passos à frente com $q = 1$ e erros não-gaussianos.

n	k	N	NPB	HaB	AKB	PT
		Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas
50	1	5,483	6,155	5,527	5,571	5,627
		0,887	0,939	0,891	0,893	0,894
		0,057/0,056	0,025/0,036	0,053/0,056	0,049/0,058	0,052/0,054
100	1	5,867	6,216	5,873	5,906	5,940
		0,929	0,951	0,933	0,932	0,932
		0,044/0,027	0,025/0,024	0,040/0,027	0,039/0,029	0,041/0,027
500	1	6,088	6,267	6,087	6,095	6,102
		0,945	0,954	0,945	0,945	0,945
		0,028/0,027	0,018/0,028	0,028/0,027	0,028/0,027	0,028/0,027
50	5	8,954	9,974	8,939	9,015	9,096
		0,885	0,929	0,886	0,890	0,893
		0,064/0,051	0,031/0,040	0,053/0,051	0,050/0,051	0,061/0,046
100	5	9,458	9,994	9,461	9,478	9,499
		0,929	0,949	0,929	0,930	0,929
		0,034/0,037	0,013/0,038	0,033/0,038	0,032/0,038	0,034/0,037
500	5	9,763	10,011	9,767	9,767	9,767
		0,919	0,940	0,921	0,920	0,920
		0,041/0,040	0,025/0,035	0,039/0,040	0,041/0,039	0,041/0,039
50	15	14,166	15,655	14,083	14,202	14,325
		0,885	0,915	0,878	0,883	0,884
		0,074/0,041	0,043/0,042	0,079/0,043	0,076/0,041	0,076/0,040
100	15	15,105	15,981	15,107	15,12	15,133
		0,906	0,923	0,904	0,907	0,907
		0,050/0,044	0,036/0,041	0,051/0,045	0,048/0,041	0,049/0,044
500	15	15,440	15,917	15,447	15,443	15,439
		0,944	0,943	0,944	0,944	0,945
		0,032/0,024	0,025/0,032	0,032/0,024	0,032/0,024	0,032/0,023

frente. Os métodos AKB, PT e HaB que incorporam à incerteza da estimação dos hiperparâmetros possuem um desempenho ligeiramente melhor que o N. O método NPB tem um desempenho bem superior aos demais com respeito à taxa de cobertura. Diferentemente das tabelas anteriores, na maioria dos casos há simetria nas caudas.

Conclui-se que o fato dos erros serem não-gaussianos com uma assimetria muito forte interfere diretamente na construção dos intervalos de previsão, isto é, os métodos não são robustos com respeito à não-normalidade dos erros para a construção de intervalos. As taxas de cobertura dos intervalos sofrem grandes mudanças principalmente para tamanhos amostrais pequenos, com exceção do método NPB. Entretanto, estudos anteriores feitos pelo autor mostraram que os métodos são robustos a uma assimetria de leve a moderada.

Tab. 4.6: Intervalos de confiança de nível de 95% para previsão k passos à frente com $q = 2$ e erros não-gaussianos.

n	k	N	NPB	HaB	AKB	PT
		Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas	Amp. Cobertura Caudas
50	1	6,654	7,415	6,669	6,723	6,809
		0,898	0,942	0,899	0,902	0,903
		0,051/0,051	0,028/0,030	0,048/0,053	0,040/0,052	0,040/0,057
100	1	7,068	7,468	7,076	7,108	7,143
		0,928	0,948	0,931	0,931	0,931
		0,043/0,030	0,026/0,026	0,039/0,030	0,038/0,037	0,041/0,028
500	1	7,299	7,497	7,299	7,307	7,314
		0,945	0,950	0,945	0,946	0,946
		0,026/0,029	0,023/0,027	0,026/0,029	0,026/0,028	0,026/0,028
50	5	11,293	13,162	11,875	11,975	12,096
		0,892	0,923	0,892	0,896	0,899
		0,058/0,050	0,037/0,040	0,057/0,051	0,054/0,050	0,055/0,046
100	5	12,596	13,323	12,617	12,617	12,168
		0,933	0,946	0,935	0,934	0,935
		0,032/0,035	0,015/0,039	0,030/0,035	0,029/0,037	0,030/0,035
500	5	13,026	13,359	13,055	13,030	13,005
		0,932	0,944	0,932	0,932	0,932
		0,034/0,034	0,022/0,034	0,034/0,034	0,034/0,034	0,034/0,034
50	15	19,571	21,483	19,417	19,596	19,786
		0,885	0,921	0,882	0,889	0,889
		0,071/0,044	0,042/0,037	0,070/0,048	0,068/0,044	0,071/0,040
100	15	20,837	22,034	20,869	20,850	20,830
		0,910	0,921	0,908	0,909	0,911
		0,042/0,048	0,034/0,045	0,042/0,050	0,041/0,050	0,021/0,048
500	15	21,425	22,034	21,481	21,428	21,373
		0,941	0,944	0,942	0,941	0,941
		0,028/0,031	0,024/0,032	0,028/0,030	0,028/0,031	0,028/0,031

4.4 Aplicações a séries temporais reais

4.4.1 Série do log da incidência de casos de Dengue em BH

Para ilustrar o funcionamento dos métodos de previsão desenvolvidos nas seções anteriores em um modelo mais complexo que o MNL, os métodos Natural, HaB e o intervalo de previsão do modelo ARIMA (Box & Jenkins, 1976) serão comparados. Para tanto, serão considerados os dados do número de casos notificados de Dengue em BH do período de 01/2002 a 03/2008, totalizando 75 observações. Como a natureza desses dados é de contagem, será trabalhada a série transformada do logaritmo natural da incidência de casos de Dengue. A incidência em um instante de tempo é definida como o número de casos notificados dividido pela população sob risco neste instante de tempo. É comum, em estudos de previsão de casos de Dengue, a modelagem da série log-incidência (Luz et al., 2008, ver). Essas informações foram cedidas pela Secretaria Municipal de Saúde de BH. Serão construídas estimativas pontuais e intervalos de confiança assintóticos e *bootstrap* para as observações futuras.

O gráfico da série, os correlogramas e o gráfico de probabilidade Normal estão na Figura 4.2. Os dados não parecem seguir uma distribuição Normal de probabilidade e possuem correlações significativas. A série apresenta um comportamento sazonal que também pode ser verificado na Figura 4.3. Os meses de dezembro-março têm uma incidência maior de Dengue, já que se tratam de meses de verão os quais possuem temperaturas elevadas e chuvas (clima propício a proliferação do mosquito transmissor da Dengue).

Como a série possui componentes de nível e sazonalidade, o Modelo Estrutural Básico (MEB) - descrito no Capítulo 2 - com sazonalidade modelada por variáveis *dummy* será ajustado a essa série transformada. As EMV e o intervalo de confiança *bootstrap* para os hiperparâmetros do modelo estão na Tabela 4.7. Parece que os componentes de inclinação e sazonalidade não são estocásticos, porém manter-se-á o ajuste do MEB.

Os modelos ARIMA (Box & Jenkins, 1976) são ajustados à série, para fazer uma comparação sobre a capacidade preditiva desses modelos frente aos ME. Segundo os critérios *AIC* e log-verossimilhança, foi escolhido o modelo $ARIMA(1, 0, 0) \times (1, 0, 0)_{12}$ - que será utilizado para a comparação das previsões com os modelos estruturais.

A Tabela 4.8 contém as estimativas de máxima verossimilhança dos parâmetros do modelo ARIMA. Note que todos os parâmetros são signifi-

Tab. 4.7: Estimativas de máxima verossimilhança e intervalos de confiança *bootstrap* para os hiperparâmetros do MEB ajustado à série do logaritmo da incidência de casos de Dengue em BH.

ψ	Estimativa	LInf.95%	LSup.95%
σ_η^2	0,176	0,095	0,219
σ_ξ^2	$2,280 \times 10^{-13}$	$7,730 \times 10^{-10}$	0,001
σ_ω^2	$5,670 \times 10^{-09}$	$4,290 \times 10^{-20}$	0,002
σ_ϵ^2	$1,620 \times 10^{-04}$	$3,300 \times 10^{-04}$	0,035

Tab. 4.8: Estimativas de máxima verossimilhança e erros padrão para os parâmetros do modelo ARIMA incompleto, ajustado à série do logaritmo da incidência de casos de Dengue em BH.

	Coefficiente	Erro padrão
Intercepto	2,345	1,216
AR(1)	0,866	0,054
Sazonal AR(1)	0,802	0,063

cativos ao nível de 5% de significância.

O MEB apresentou o $AIC = 93,00$ e a log-verossimilhança = $-43,00$, enquanto que o melhor modelo ARIMA incompleto encontrado obteve o $AIC = 109,54$ e a log-verossimilhança = $-50,86$. Isso evidencia um melhor ajuste do MEB.

Analisando a Figura 4.4, percebe-se que o ajuste do MEB é adequado. Os resíduos não evidenciam nenhum padrão, não são autocorrelacionados e se aderem a uma distribuição Normal. Além da análise gráfica, foram feitos testes. O valor da estatística de teste de Ljung-Box e seu p-valor para os lags 10, 12, 24 e 36 foram $Q(10) = 11,44$ (p-valor = 0,32), $Q(12) = 13,57$ (p-valor = 0,33), $Q(24) = 32,66$ (p-valor = 0,12) e $Q(36) = 42,69$ (p-valor = 0,21), respectivamente. Não se rejeita a hipótese de que os resíduos são independentes. Já o p-valor do teste de normalidade foi maior que 0,10, ou seja, não se rejeita a hipótese que os resíduos são oriundos de uma distribuição Normal.

Analisando a Figura 4.5, não há evidência de violação das suposições do ajuste do modelo ARIMA. O valor da estatística de teste de Ljung-Box e seu p-valor para os lags 6, 12, 24 e 36 foram $Q(6) = 9,73$ (p-valor = 0,136), $Q(12) = 12,532$ (p-valor = 0,404), $Q(24) = 20,16$ (p-valor = 0,69) e $Q(36) = 21,43$ (p-valor = 0,97), respectivamente. Não se rejeita a hipótese

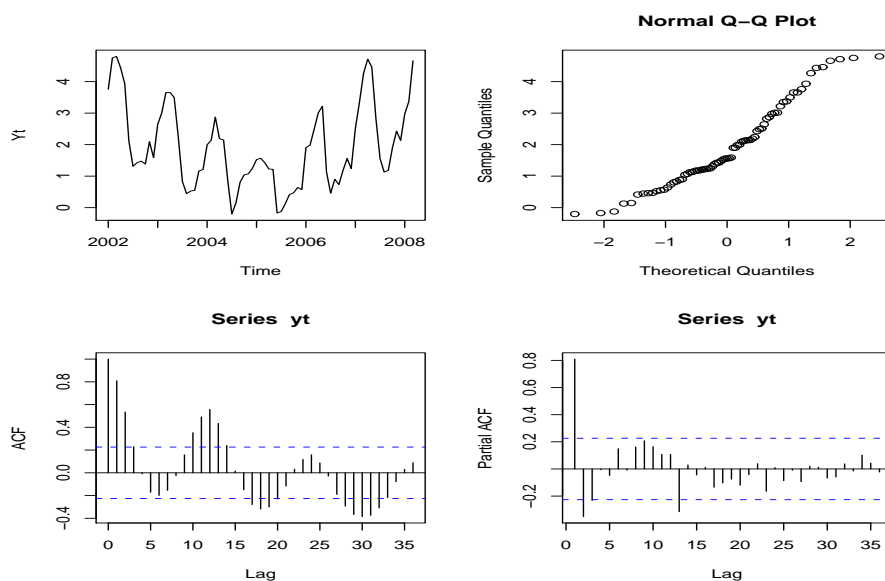


Fig. 4.2: No topo à direita: O gráfico temporal da série do logaritmo natural da incidência de Dengue por 100.000 habitantes; No topo à esquerda: O QQ-Plot Normal; Abaixo à esquerda: Função de autocorrelação amostral da série; Abaixo à direita: Função de autocorrelação parcial amostral da série.

de que os resíduos são independentes. Também não se rejeita a hipótese que os resíduos são provenientes de uma distribuição Normal (p -valor = 0,88).

Na Figura 4.6 estão as previsões e intervalos de confiança de 95% para as observações futuras, obtidos através dos modelos ajustados aos dados. A linha contínua após a linha vertical representa os valores reais. Note que os intervalos de confiança assintótico (natural) e *bootstrap* cobriram todos os valores reais futuros, sob o MEB. O intervalo *bootstrap* (I^{HaB}) é mais amplo que o natural, já que incorpora a incerteza da estimação dos hiperparâmetros pelo procedimento de Hamilton *bootstrap* (HaB). Optou-se por esse procedimento por incorporar a incerteza e ser mais simples. O mesmo não pode ser dito para o intervalo de confiança assintótico obtido através do ajuste do modelo ARIMA.

4.4.2 Série Mortgages

Mortgages é um pedido ou uma solicitação contra uma propriedade real feita por quem empresta de forma a se obter o ressarcimento do valor do

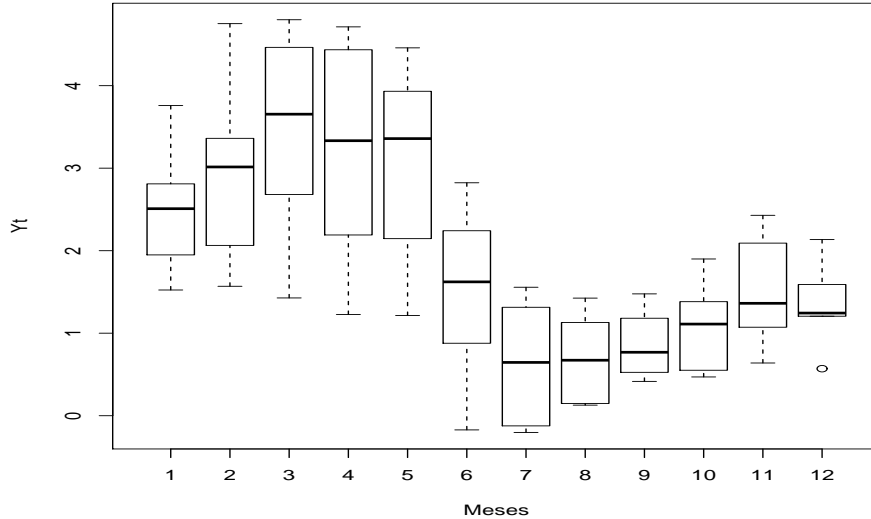


Fig. 4.3: Boxplots da série log incidência de casos de Dengue em BH segundo os meses.

empréstimo. Isto é, um empréstimo com uma propriedade (patrimônio) como garantia. *Mortgages* geralmente são executados em um período de 10 a 30 anos, durante o qual o empréstimo é pago. Esse índice é mensal dado pelo seguro do Tesouro dos EUA. A série temporal de *Mortgages* consiste de 63 observações mensais. As últimas 2 observações da série são omitidas do ajuste para a comparação dos intervalos de previsão para os valores futuros. O gráfico temporal da série está na Figura 4.7. Os dados não parecem seguir uma distribuição Normal de probabilidade e possuem correlações significativas.

Um MNL é ajustado à série e os resíduos não seguem uma distribuição normal (ver Figura 4.8). A Tabela 4.9 apresenta as estimativas pontual e intervalar para os hiperparâmetros de um MNL. Note que os intervalos de confiança *bootstrap* estão afastados do valor zero. Todos os intervalos de previsão (métodos), apresentados nas seções anteriores teoricamente, estão na Figura 4.9 e cobrem os 2 valores futuros. Note que os demais intervalos possuem uma amplitude maior que o intervalo (I^N), como foi também verificado nos estudos de simulação.

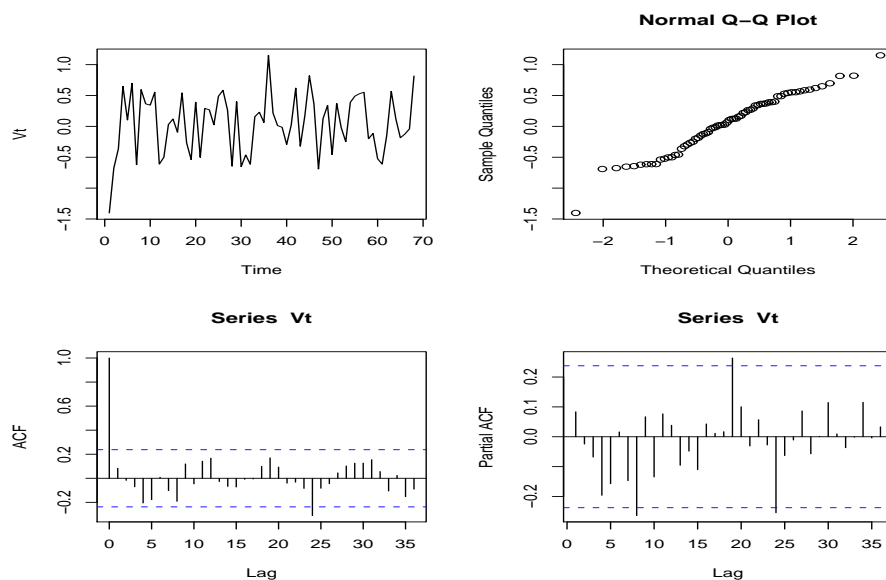


Fig. 4.4: Análise dos resíduos do MEB. No topo à direita: O gráfico temporal dos resíduos; No topo à esquerda: O QQ-Plot Normal dos resíduos; Abaixo à direita: Função de autocorrelação amostral dos resíduos; Abaixo à esquerda: Função de autocorrelação parcial amostral dos resíduos.

Tab. 4.9: Estimativas de máxima verossimilhança e intervalos de confiança *bootstrap* para os hiperparâmetros do MNL ajustado à série Mortgages.

ψ	Estimativa	LInf.95%	LSup.95%
σ_{η}^2	0,086	0,020	0,181
σ_{ϵ}^2	0,1023	0,034	0,183

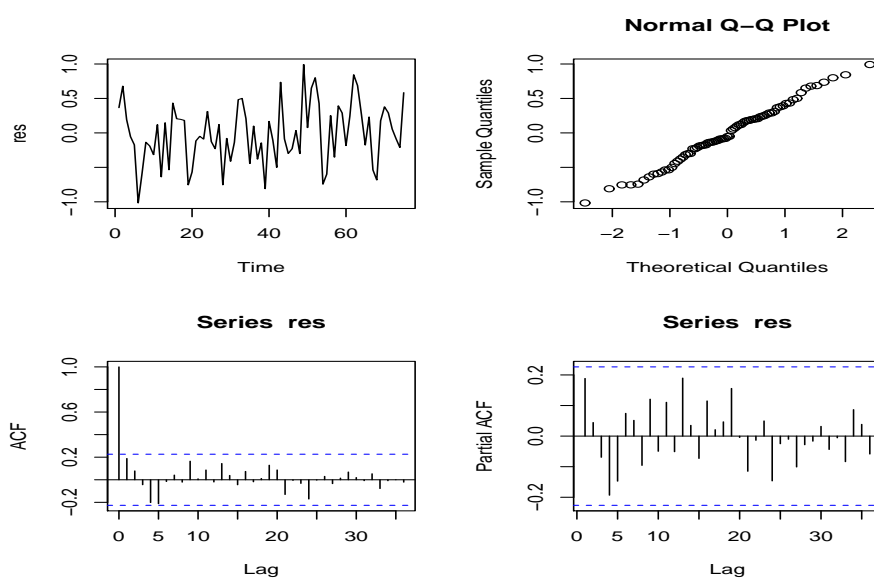


Fig. 4.5: Análise dos resíduos do modelo ARIMA. No topo à direita: O gráfico temporal dos resíduos; No topo à esquerda: O QQ-Plot Normal dos resíduos; Abaixo à direita: Função de autocorrelação amostral dos resíduos; Abaixo à esquerda: Função de autocorrelação parcial amostral dos resíduos.

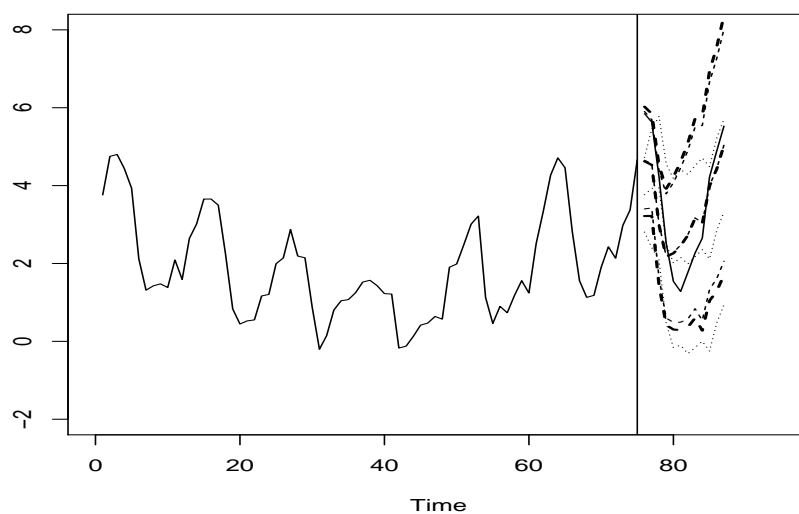


Fig. 4.6: Previsão para a série da log-incidência de dengue em BH. Linha contínua: a série do logaritmo da incidência de casos de Dengue em BH; Linha pontilhada: as estimativas pontual e intervalar do ajuste do modelo ARIMA; Linha tracejada: a estimativa pontual e o intervalo de previsão assintótico (I^N) do ajuste do MEB; Linha tracejada em negrito: a estimativa pontual e o intervalo de previsão *bootstrap* I^{HaB} . A linha vertical separa os dados usados na estimação do modelo das previsões. O nível de confiança adotado é de 95%.

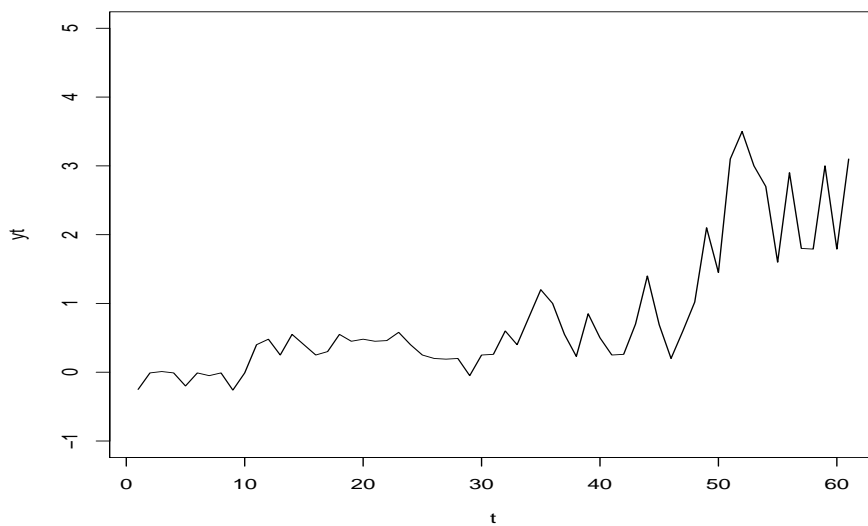


Fig. 4.7: A série temporal Mortgages.

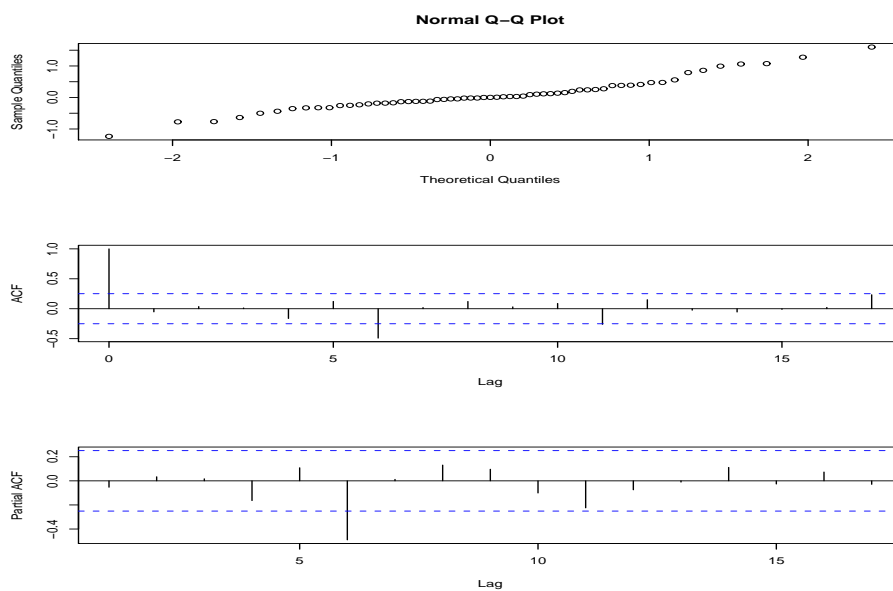


Fig. 4.8: Análise dos resíduos do ajuste de um MNL à série Mortgages.

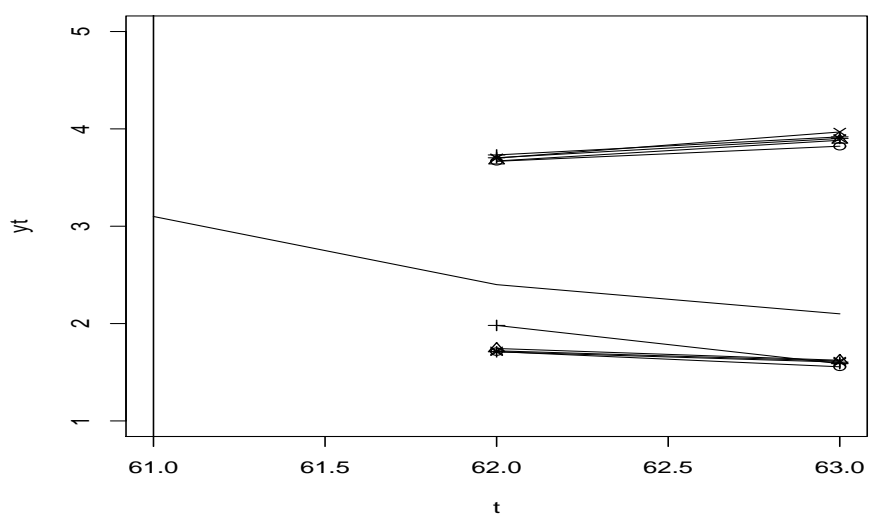


Fig. 4.9: As linhas contínuas com círculo, x, +, asterisco e triângulo indicam a série Mortgages, o intervalo de previsão assintótico (I^N), I^{Pfti} , I^{NPB} , I^{AKB} e I^{HaB} , respectivamente. A linha vertical separa os dados usados na estimação do modelo das previsões. O nível de confiança adotado é de 95%.

5. MODELOS DE ESPAÇOS DE ESTADOS NÃO-GAUSSIANOS

Este capítulo destina-se à introdução de uma família de distribuições e/ou modelos dinâmicos, bem como vários resultados inéditos de inferência, previsão e suavização em modelos de espaço de estados não-gaussianos.

Na literatura, há vários modelos que são construídos baseados nas suposições de normalidade, homoscedasticidade e independência dos erros, porém nem sempre é possível satisfazer essas suposições. Sob o contexto de séries temporais, a hipótese de independência dos erros não é satisfeita, e a suposição de normalidade tem sido central às principais abordagens existentes para tratamento de séries temporais.

A modelagem via espaço de estados, que é o tema abordado nesta tese, possui uma grande quantidade de trabalhos e resultados, baseados na suposição de normalidade. Neste trabalho, serão apresentadas algumas possibilidades para o tratamento de séries temporais que extrapolam essa restrição.

O ponto de partida para essa extensão é o artigo de Nelder & Wedderburn (1972), que propôs a família por eles denominada de *modelos lineares generalizados* (MLG), unificando vários modelos, então existentes de forma isolada, em uma classe. 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, uma família de distribuições é proposta e, assim como os MLG, também permite um leque de opções para a distribuição para a variável-resposta (Y). De uma maneira geral, define-se que uma variável aleatória Y possui uma distribuição na família não-gaussiana, se a sua distribuição é escrita na forma:

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

onde $y \in H(\varphi) \subset \mathfrak{R}$ e 0 , caso contrário. O termo $\mu = \lambda g(x' \beta)$ é conhecido função de ligação, onde $\lambda > 0$ é um parâmetro de escala, x' é um vetor de covariáveis e β os parâmetros associados às mesmas. φ é um vetor de parâmetros desconhecidos e estáticos. As funções $a(\cdot, \cdot) > 0$, $b(\cdot, \cdot)$, $c(\cdot, \cdot)$ e $H(\varphi)$ são tais que $p(y|\mu, \varphi) \geq 0$ e que a integral de Lebesgue-Stieljes $\int dP(y|\mu, \varphi) = 1$. Caso $b(y, \varphi) = b(y)$ ou $c(y, \varphi) = c(y)$ e $S(\varphi)$ seja um conjunto que não dependa de φ , a família não-Gaussiana de distribuições torna-se um caso especial da família exponencial de distribuições. $a(\cdot, \cdot) > 0$ e μ têm o mesmo sinal.

No contexto de séries temporais, a estrutura de correlação das observações não pode ser desprezada. Nesse sentido, 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. Dentre os quais pode-se citar os trabalhos de Gamerman & West (1987), Grunwald, Raftery & Guttorp (1993), Fahrmeir (1987), Fruhwirth-Schnatter (1994), Lindsey & Lambert (1995), Gamerman (1991); Gamerman (1998), Chiogna & Gaetan (2002), Hemming & Shaw (2002) e Godolphin & Triantafyllopoulos (2006). Trabalhos que tratam de séries temporais não-Gaussianas, não necessariamente os MLDG, incluem Smith (1979), Smith (1981), Cox (1981), Kaufmann (1987), Kitagawa (1987), Shephard & Pitt (1997), Smith & Miller (1986), Harvey & Fernandes (1989), Jørgensen et al. (1999) e Durbin & Koopman (2000), entre outros.

O problema com essa classe de modelos (MLDG) é que a tratabilidade analítica é facilmente perdida, mesmo para componentes muito simples. Assim, a verossimilhança preditiva, que é fundamental para o processo de inferência, pode apenas ser obtida de forma aproximada. Uma particularização desses modelos assume que apenas a tendência flutua e os efeitos das covariáveis são fixos ao longo do tempo. Para esses casos, é proposto uma classe bastante ampla de modelos que permite o cálculo exato da verossimilhança preditiva, a *Família não-Gaussiana de Modelos de Espaço de Estados com Verossimilhança Exata* (FMEVE).

Nos últimos dez anos, sob a abordagem Bayesiana, muitos pesquisadores têm trabalhado em procedimentos auxiliares a fim de se fazer inferência sobre os modelos de séries temporais, como os modelos de espaço de estados. Fruhwirth-Schnatter & Wagner, H. (2006) usaram procedimentos auxiliares de amostragem de mistura e Pitt & Walker (2005) utilizaram variáveis auxiliares para obtenção de séries estacionárias. Andrieu & Doucet (2002) and

Silva, et al. (2009) adotaram filtro de partículas no contexto de modelos de espaço de estados. Entretanto, é importante enfatizar que esses métodos são aproximações e o tempo computacional pode ser muito alto.

O objetivo principal deste capítulo é propor a família de modelos de espaço de estados com verossimilhança exata, onde se tem a tratabilidade analítica. Essa família é obtida a partir de uma generalização de um resultado de Smith & Miller (1986). Eles propuseram uma equação de evolução exata, possibilitando assim a integração analítica dos estados e a obtenção das distribuições preditivas que compõem a função de verossimilhança.

Assim, as contribuições deste capítulo são propor e caracterizar a FMEVE, apresentar casos particulares pertencentes a essa família, como os modelos Borel-Tanner, Gama Generalizada, Weibull, Rayleigh, Gaussiana Inversa, Laplace, Pareto e Beta, propor uma forma de suavização do componente de nível do modelo e um método de previsão.

A família de modelos de espaço de estados com verossimilhança exata será apresentada na Seção 5.2. Seus principais resultados teóricos serão fornecidos bem como será descrita a forma de fazer inferência. Novamente, a abordagem clássica, utilizando o estimador de máxima verossimilhança (EMV), e a abordagem Bayesiana, utilizando os estimadores Bayesianos (EB) que não possuem forma analítica, logo os métodos MCMC são usados para obter aproximações para os mesmos, são consideradas no processo de inferência. Além disso, os intervalos de confiança assintótico e de credibilidade são construídos para os parâmetros. A seguir, na Seção 5.3 serão apresentados casos particulares de interesse dentro dessa família. A Seção 5.4 trata da comparação dos pontos de vista clássico e Bayesiano através de exercícios de simulação. O capítulo é concluído com aplicações a dados reais.

5.1 Modelos Lineares Dinâmicos Generalizados

Essa classe de modelos foi introduzida por West, Harrison & Migon (1985) e consiste numa extensão direta do MEE (Modelo de Espaço de Estados) dado em (2.3)-(2.4) ao permitir especificações mais gerais para a distribuição das observações da série, pertencente à família exponencial de distribuições. A outra novidade aqui com relação à especificação do Capítulo 2 é a presença de uma função de ligação para estabelecer a conexão entre as equações das observações e dos estados. Mais especificamente, diz-se que uma série temporal y_t é descrita por um Modelo Linear Dinâmico Generalizado (MLDG)

se segue

$$p(y_t|\theta_t, \phi) = a(y_t, \phi) \exp\{\phi[\theta_t y_t - b(\theta_t)]\} \quad \text{onde } \mu_t = E(y_t|\theta_t), \quad (5.2)$$

$$g(\mu_t) = z_t' \alpha_t + d_t, \quad (5.3)$$

$$\alpha_t = \mathbf{T}_t \alpha_{t-1} + \mathbf{c}_t + \mathbf{R}_t \eta_t, \quad \eta_t \sim N(0, \mathbf{Q}_t), \quad (5.4)$$

onde o suporte de variação de y_t não depende de θ_t , η_t é um vetor de ruídos serialmente não-correlacionados, cuja matriz de covariâncias é dada por \mathbf{Q}_t e α_t é o vetor de estados, para $t = 1, 2, \dots, n$. Assume-se também que η_t sejam independentes entre si com média 0, para $t = 1, 2, \dots, n$. Novamente, as matrizes \mathbf{z}_t , \mathbf{T}_t e \mathbf{R}_t são conhecidas como matrizes do sistema e assumidas determinísticas e conhecidas e os termos d_t e \mathbf{c}_t são covariáveis que podem ser inseridas nas equações das observações e de transição, respectivamente. O modelo é completado com a especificação inicial $\alpha_0 \sim N(\mathbf{a}_0, \mathbf{P}_0)$.

Note que a equação de observação depende do parâmetro canônico θ_t e do parâmetro de dispersão ϕ , aqui assumido constante ao longo do tempo. Os hiperparâmetros do modelo (5.2)-(5.4) são dados pela dispersão ϕ e pelas componentes desconhecidas em \mathbf{Q}_t .

Resultados sobre conjugação dos parâmetros canônicos foram então utilizados por West, Harrison & Migon (1985) para propor formas aproximadas para um procedimento sequencial de inferência supondo conhecidos os hiperparâmetros ψ . Com isso, as distribuições preditivas podem ser obtidas e seu produto fornece a função de verossimilhança preditiva para os hiperparâmetros. Inferência tanto sob a ótica clássica quanto Bayesiana pode ser realizada. Vale ressaltar que essa verossimilhança só pode ser obtida a partir de uma aproximação no ciclo de inferência. Na seção seguinte, será apresentada uma classe relativamente ampla de modelos dinâmicos onde também não se assume normalidade dos dados, mas a função de verossimilhança preditiva pode ser obtida de forma exata.

A maioria dos trabalhos na área seguiu preferencialmente dois caminhos. No primeiro deles, optou-se por assumir hiperparâmetros conhecidos ou estimados por algum procedimento *ad-hoc*, como uso de fatores de desconto. Essa foi a abordagem original de West, Harrison & Migon (1985) mas também foi utilizada por outros autores como Singh & Roberts (1992) e Fahrmeir & Wagenpfeil (1997). O segundo caminho foi tratar todos os parâmetros de estados e hiperparâmetros conjuntamente e aproximar a distribuição *a posteriori* conjunta através de amostras obtidas via algum algoritmo de simulação. Existem inúmeras formas de fazer isso: Shephard & Pitt (1997), Gamerman (1998) e Knorr-Held (1999) usaram MCMC com diferentes estratégias de

formulação de propostas ao passo que Durbin & Koopman (2000) usaram simulação Monte Carlo via reamostragem por importância.

5.2 Família de modelos de espaço de estados com verossimilhança exata

Smith & Miller (1986), Harvey & Fernandes (1989) e Shephard (1994) 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 de modelos de espaço de estados de verossimilhança exata. Uma vantagem desses modelos frente aos MLDG é que a equação de evolução é exata e não aproximada. Por outro lado, não é uma tarefa trivial inserir outros componentes estocásticos no mesmo, por exemplo, os componentes de tendência e sazonalidade. Entretanto, o efeito desses componentes determinísticos podem ser captados no modelo através de covariáveis. Nesta seção, são abordados a definição da FMEVE, os procedimentos de inferência (clássico e Bayesiano), procedimentos de previsão e suavização, bem como a adequação do modelo.

5.2.1 Definição

De uma maneira geral, define-se que a série temporal $\{y_t\}$ possui uma distribuição na família não-gaussiana de modelos de espaço de estados com verossimilhança exata, se a sua distribuição é escrita na forma:

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = a(y_t, \boldsymbol{\varphi}) \mu_t^{b(y_t, \boldsymbol{\varphi})} \exp(-\mu_t c(y_t, \boldsymbol{\varphi})), \quad (5.5)$$

onde $y_t \in S(\boldsymbol{\varphi}) \subset \mathfrak{R}$ e $p(y_t|\mu_t, \boldsymbol{\varphi}) = 0$, caso contrário. μ_t são parâmetros latentes e $\boldsymbol{\varphi}$ um vetor de parâmetros estáticos. As funções $a(\cdot)$, $b(\cdot)$, $c(\cdot)$ e $S(\cdot)$ são tais que $p(y_t|\mu_t, \boldsymbol{\varphi}) \geq 0$ e a integral de Lebesgue-Stieljes $\int dP(y_t|\mu_t, \boldsymbol{\varphi}) = 1$ em que $F(\cdot)$ denota a função distribuição acumulada de y_t .

A FMEVE é definida da seguinte forma:

- A1 Se x_t é um vetor de covariáveis, a função de ligação g relaciona o preditor com o parâmetro μ_t através da relação $\mu_t = \lambda_t g(x_t, \boldsymbol{\beta})$, onde $\boldsymbol{\beta}$ contém coeficientes de regressão (um dos componentes de $\boldsymbol{\varphi}$) e λ_t é o parâmetro responsável pela descrição do nível dinâmico. Se o preditor for linear então, $g(x_t, \boldsymbol{\beta}) = g(x_t' \boldsymbol{\beta})$.

A2 O nível dinâmico λ_t evolui de acordo com a equação de evolução $\lambda_{t+1} = w^{-1}\lambda_t\varsigma_{t+1}$ onde $\varsigma_{t+1}|\mathbf{Y}_t \sim \text{Beta}(w\lambda_t, (1-w)\lambda_t)$, isto é,

$$w\frac{\lambda_{t+1}}{\lambda_t} | \lambda_t, \mathbf{Y}_t \sim \text{Beta}(w\lambda_t, (1-w)\lambda_t),$$

where $0 < w \leq 1$, $\mathbf{Y}_t = \{Y_0, y_1, \dots, y_t\}$, for $t = 1, 2, \dots$. Y_0 representa a informação anterior disponível.

A3 A dinâmica do nível λ_t é inicializada com a especificação inicial *a priori* $\lambda_0|Y_0 \sim \text{Gama}(a_0, b_0)$. Logo, usando a propriedade de escala da distribuição Gama, $\mu_0|Y_0 \sim \text{Gama}(a_0, b_0)$.

Uma das especificações mais comuns para a função de ligação g é a função logarítmica. É interessante destacar que, nesse caso, é obtida a equação de evolução $\ln(\lambda_t) = \ln(\lambda_{t-1}) + \varsigma_t^*$, onde $\varsigma_t^* = \ln(\varsigma_t/w) \in \mathfrak{R}$. Essa equação é similar à usual equação de evolução dada por um passeio aleatório, como no modelo de nível local. O parâmetro w varia entre 0 e 1 e também compõe φ . Como será visto a seguir, w cumpre a função de aumentar multiplicativamente a variância devido à passagem do tempo. Assim, ele desempenha um papel similar ao das variâncias do sistema e idêntico ao dos fatores de desconto, usados na abordagem Bayesiana para substituir essas variâncias.

Caso $b(y_t, \varphi) = b(y)$ ou $c(y_t, \varphi) = c(y)$ e $S(\varphi)$ seja um conjunto que não dependa de φ , a família não-gaussiana de modelos de espaço de estados com verossimilhança exata torna-se um caso especial da família exponencial de distribuições.

A Tabela 5.1 fornece a forma das funções a , b , c e S para algumas distribuições nessa família. Os casos mais comuns são Poisson e Exponencial que foram anteriormente destacados na literatura. Vários outros casos dessa família são introduzidos aqui e incluem distribuições discretas e contínuas. Algumas delas são bem conhecidas como a Normal e a Pareto, mas a família também inclui as distribuições de Borel-Tanner (Haight and Breuer, 1960) e Rayleigh, por exemplo. A figura que emerge é uma coleção que é capaz de representar uma variedade de características as quais estão presentes em aplicações de séries temporais e, logo, são de importância prática.

Para uma melhor caracterização da FMEVE, alguns resultados são apresentados no Teorema 5.1 tais como a distribuição *a posteriori* do nível λ_t e as distribuições preditivas. Esses resultados são fundamentais no processo inferencial da FMEVE.

Tab. 5.1: Casos especiais de MEE não-Gaussianos da FMEVE.

Modelos	φ	$a(y_t, \varphi)$	$b(y_t, \varphi)$	$c(y_t, \varphi)$	$S(\varphi)$
Poisson	$(w, \beta)'$	$(y_t!)^{-1}$	y_t	1	$\{0, 1, \dots\}$
Borel-Tanner	$(w, \rho, \beta)'$	$\frac{\rho}{(y_t - \rho)!} y_t^{y_t - \rho - 1}$	$y_t - \rho$	y_t	$\{\rho, \rho + 1, \dots\}$
Gamma	$(w, \chi, \beta)'$	$y_t^{\chi-1} / \Gamma(\chi)$	χ	y_t	$(0, \infty)$
Weibull	$(w, \nu, \beta)'$	$\nu(y_t)^{\nu-1}$	1	$(y_t)^\nu$	$(0, \infty)$
Pareto	$(w, \rho, \beta)'$	y_t^{-1}	1	$\ln y_t - \ln \rho$	(ρ, ∞)
Normal	$(w, \theta)'$	$(2\pi)^{-1/2}$	1/2	$(y_t - \theta)^2 / 2$	$(-\infty, \infty)$
Laplace	$(w, \theta)'$	$\frac{1}{\sqrt{2}}$	1	$\sqrt{2} y_t - \theta $	$(-\infty, \infty)$
Gaussiana	$(w, \theta)'$	$\frac{1}{\sqrt{2\pi y_t^3}}$	1/2	$\frac{(y_t - \theta)^2}{2y_t \theta^2}$	$(0, \infty)$
Inversa					
Rayleigh	$(w, \theta)'$	y_t	1	$\frac{1}{2}(y_t - \theta)^2$	$(0, \infty)$
Power	$(w, \nu, \kappa, \theta)'$	$\frac{\nu}{\kappa 2^{\frac{\nu+1}{\nu}} \Gamma(1/\nu)}$	1/ν	$\frac{(y_t - \theta)^\nu}{2\kappa^\nu}$	$(-\infty, \infty)$
Exponential (GED)					
Generalized Gamma	$(w, \nu, \chi)'$	$\nu y_t^{\chi-1} / \Gamma(\chi/\nu)$	1	y_t^ν	$(0, \infty)$

Teorema 5.1

Se o modelo está definido na forma descrita na Equação (5.5), os seguintes resultados podem ser obtidos:

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

$$a_{t|t-1} = w a_{t-1}, \quad (5.6)$$

$$b_{t|t-1} = w b_{t-1}, \quad (5.7)$$

e $0 < w \leq 1$.

2. $(\mu_t = \lambda_t g(x_t, \beta)) | \mathbf{Y}_{t-1}$, que é Gama($a_{t|t-1}^*$, $b_{t|t-1}^*$), onde

$$a_{t|t-1}^* = w a_{t-1}, \quad (5.8)$$

$$b_{t|t-1}^* = w b_{t-1} [g(x_t, \beta)]^{-1}. \quad (5.9)$$

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

$$a_t^* = a_{t|t-1}^* + b(y_t, \varphi), \quad (5.10)$$

$$b_t^* = b_{t|t-1}^* + c(y_t, \varphi). \quad (5.11)$$

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

$$a_t = a_{t|t-1} + b(y_t, \boldsymbol{\varphi}), \quad (5.12)$$

$$b_t = b_{t|t-1} + c(y_t, \boldsymbol{\varphi})g(x_t, \boldsymbol{\beta}). \quad (5.13)$$

5. A função de densidade preditiva um passo a frente é dada por

$$p(y_t|\mathbf{Y}_{t-1}, \boldsymbol{\varphi}) = \frac{\Gamma(b(y_t, \boldsymbol{\varphi}) + a_{t|t-1}^*)a(y_t, \boldsymbol{\varphi})(b_{t|t-1}^*)^{a_{t|t-1}^*}}{\Gamma(a_{t|t-1}^*)[c(y_t, \boldsymbol{\varphi}) + b_{t|t-1}^*]^{b(y_t, \boldsymbol{\varphi}) + a_{t|t-1}^*}}, \quad y_t \in S(\boldsymbol{\varphi}). \quad (5.14)$$

$\forall t \in N; t \leq n$ onde n é o tamanho da série temporal e $\Gamma(\cdot)$ é a função gama. A prova do Teorema 5.1 é encontrada no Apêndice 2. Fica fácil ver a partir de (5.6)- (5.7) que $Var(\lambda_t|\mathbf{Y}_{t-1}) = w^{-1}Var(\lambda_{t-1}|\mathbf{Y}_{t-1})$. Pensando-se em termos da precisão (inverso da variância) como medida de informação, tem-se que a passagem de $t-1 \rightarrow t$ implica que apenas $100w\%$ da informação é preservada. É exatamente esse o uso de fatores de desconto em West & Harrison (1997). Esses fatores medem a quantidade de informação (medida pela precisão do sistema) preservada com o passar do tempo.

5.2.2 Procedimento de inferência

Os parâmetros do modelo podem ser divididos em estados latentes λ_t e parâmetros fixos, usualmente denominados hiperparâmetros ($\boldsymbol{\varphi}$). A inferência *on-line* para os λ_t 's foi tratada na Seção 5.2.1 e a inferência suavizada será tratada na Seção 5.2.4. Nesta subseção, será abordado a inferência sobre os hiperparâmetros.

5.2.2.1 Inferência Clássica

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

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

onde τ é o instante da primeira observação diferente de zero; φ é composto por ω , β e por parâmetros específicos do modelo; $\mathbf{Y}_n = (y_1, \dots, y_n)'$.

Na Equação (5.15), a definição de τ , como apresentada acima, se deve ao seguinte fato: a distribuição *a priori* Gama, isto é, a distribuição inicial μ_t 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. 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. Entretanto, se $a_0 > 0$ e $b_0 > 0$, não é necessária a utilização de τ . Será assumido, daqui por diante, que $a_0 > 0$ e $b_0 > 0$.

Pelas propriedades assintóticas do EMV (Harvey, 1989), sob certas condições de regularidade, conclui-se que

$$I_n^{1/2}(\hat{\varphi})(\hat{\varphi} - \varphi) \rightarrow N_p[\mathbf{0}, I_p], \quad (5.16)$$

onde I_p é a matriz identidade de dimensão $p \times p$.

Se $\hat{\varphi}$ é obtido maximizando a função de log-verossimilhança marginal, que é obtida analiticamente como o produto das distribuições preditivas, o resultado acima está sujeito a algumas das seguintes condições (Harvey, 1989):

1. φ é um ponto interior do espaço paramétrico;
2. as derivadas da log-verossimilhança até a ordem 3 com respeito a φ existem e são contínuas na vizinhança do valor verdadeiro do parâmetro;
3. φ é identificável.

Condição (1): Quando $w = 1$ e/ou algum parâmetro do modelo definido na semi-reta positiva estiver na borda do espaço paramétrico, φ não será um ponto interior ao espaço paramétrico e a distribuição limite pode ser afetada.

Condição (2): As derivadas, para um instante de tempo t arbitrário com $t = 1, \dots, n$, existem e podem ser calculadas derivando a função de densidade preditiva em (5.14) com respeito a φ e a sua continuidade decorre da continuidade de $a_{t|t-1}^*$ e $b_{t|t-1}^*$. No caso em que um dos parâmetros do modelo dependa do suporte, as derivadas deste respectivo parâmetro podem não existir.

Condição (3): Como a função de verossimilhança é obtida analiticamente pelo produto das funções densidades preditivas em (5.14) e sendo φ um ponto

interior ao espaço paramétrico, dado dois pontos φ_1 e φ_2 arbitrários e diferentes, a família de densidade conjunta $\mathcal{F} = \{p(y_1, \dots, y_n; \varphi); \varphi \in \Phi\}$ produzirá modelos diferentes e, conseqüentemente, valores diferentes da função de verossimilhança.

O intervalo de confiança assintótico para φ é construído baseado em uma aproximação numérica para a matriz de informação de Fisher, utilizando $I(\varphi) \cong -H(\varphi)$ em que $-H(\varphi)$ é a matriz de informação observada, calculada através das derivadas segundas da função de log-verossimilhança com respeito aos parâmetros. Como o cálculo das derivadas não é uma tarefa trivial, utilizam-se derivadas numéricas.

Seja φ_i , $i = 1, \dots, p$, um parâmetro qualquer no vetor φ . Então, um intervalo de confiança assintótico de $100(1 - \kappa)\%$ para φ_i é dado por:

$$\hat{\varphi}_i \pm z_{\kappa/2} \sqrt{\widehat{Var}(\hat{\varphi}_i)},$$

onde $z_{\kappa/2}$ é o $\kappa/2$ percentil da distribuição Normal-padrão e $\widehat{Var}(\hat{\varphi}_i)$ é obtido dos elementos da diagonal da inversa da matriz de informação de Fisher.

A matriz de informação observada é assintoticamente equivalente à matriz de informação esperada - resultado já conhecido (Migon & Gamerman, 1999) e aparentemente corroborado para modelos espaço de estados pelas simulações conduzidas em Cavanaugh & Shumway (1996). Essa prática de aproximar a matriz de informação esperada pela matriz de informação observada é relativamente comum e sugerida em muitos textos como Cavanaugh & Shumway (1996) e Sallas & Harville (1988), principalmente para grandes amostras.

Em alguns problemas, o interesse principal está em calcular uma função do estimador do parâmetro. Nestes casos, utiliza-se o método Delta (Casella & Berger, 2002), o qual é definido logo abaixo. Seja $g(\cdot)$ uma função um-a-um cuja primeira derivada existe e é diferente de zero. Utilizando a propriedade assintótica dos EMV em (5.16), pelo método Delta, tem-se que

$$I_{ii}^{1/2}(\hat{\varphi}) \left[g'(\hat{\varphi}_i) \right]^{-1} [g(\hat{\varphi}_i) - g(\varphi_i)] \rightarrow N[0, 1], \quad (5.17)$$

para $i = 1, \dots, p$. $I_{ii}^{-1}(\varphi)$ é o i -ésimo elemento da diagonal da matriz de informação de Fisher $I^{-1}(\varphi)$.

Sob o enfoque clássico, intervalos de confiança assintóticos são construídos para os parâmetros, mas esses intervalos podem apresentar problemas de fronteira, isto é, os limites dos intervalos ultrapassam as bordas do espaço

paramétrico. Nesses casos, aplica-se o método Delta para corrigir esse problema da seguinte maneira:

1. O intervalo de confiança assintótico para $g(\varphi_i)$ é construído para $i = 1, \dots, p$, usando o método Delta;
2. A seguir, a transformação inversa $g^{-1}(\cdot)$ é aplicada nos limites do intervalo, obtendo-se, assim, o intervalo assintótico para φ .

5.2.2.2 Inferência Bayesiana

Já para se fazer inferência Bayesiana sobre os parâmetros do modelo, como a distribuição *a posteriori* dos parâmetros não é analiticamente tratável, é usado o método MCMC, algoritmo de Metropolis-Hastings (Gamerman & Lopes, 2006), descrito em detalhes em Santos (2009), - a fim de se obter uma amostra da distribuição *a posteriori* do vetor de parâmetros φ cuja forma é dada por:

$$\pi(\varphi | \mathbf{Y}_n) \propto L(\varphi; \mathbf{Y}_n)\pi(\varphi), \quad (5.18)$$

onde $L(\varphi; \mathbf{Y}_n)$ é a função de verossimilhança obtida em (5.15) e $\pi(\varphi)$ é a distribuição *a priori* para φ . Neste trabalho, uma distribuição *a priori* uniforme própria é usada, dada por $\pi(\varphi) = c$ para todos os valores possíveis de φ em um intervalo limitado pré-fixado e 0, caso contrário.

Intervalos de credibilidade percentílico 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 $100(1 - \kappa)\%$.

Inferência para as variáveis latentes pode ser feita com a saída do algoritmo de MCMC. Uma vez que uma amostra $\varphi^{(1)}, \dots, \varphi^{(M)}$ está disponível, amostras *a posteriori* $\lambda^{(1)}, \dots, \lambda^{(M)}$ das variáveis latentes são obtidas como se segue:

1. fixe $j = 1$;
2. amostre o hiperparâmetro $\varphi^{(j)}$ através do algoritmo de MCMC;

3. amostre o conjunto $\lambda^{(j)}$ de variáveis latentes de $p(\lambda|\varphi^{(j)}, \mathbf{Y}_n)$ usando Teorema 5.3;
4. aumente $j \rightarrow j + 1$ e retorne ao passo 1, até $j = M$.

Novamente, comentários similares são válidos para as distribuições preditivas. Note que

$$p(y_{t+h}|\mathbf{Y}_t) = \int p(y_{t+h}|\mathbf{Y}_t, \varphi) \pi(\varphi|\mathbf{Y}_n) d\varphi.$$

Portanto, as distribuições preditivas h passos à frente podem ser aproximadas por

$$\frac{1}{M} \sum_{j=1}^M p(y_{t+h}|\mathbf{Y}_t, \varphi^{(j)})$$

das quais resumos tais como médias, variâncias e intervalos de credibilidade podem ser obtidos.

Maiores informações e detalhes sobre os estimadores Bayesianos e algoritmo de Metropolis-Hastings podem ser vistos em Santos (2009), Gamerman & Lopes (2006), Gilks, Richardson & Spiegelhalter (1996) e Migon & Gamerman (1999).

5.2.3 Previsão

A previsão dos valores futuros de uma série constitui um tópico importante em séries temporais, devido ao grande interesse em extrapolar os resultados do ajuste do modelo e, assim, projetar os valores futuros. As previsões para as observações futuras da série podem ser obtidas através das distribuições preditivas.

O Teorema 5.2 fornece resultados para a distribuição preditiva das observações futuras. A prova é também apresentada no Apêndice 3. Tentativas para se obter essa distribuição assumindo a validade de **A2** apenas para $h = 0$ não são uma tarefa trivial, de acordo com Harvey (1989, p. 353). As extensões introduzidas aqui permitem um cálculo fácil e também provêm benefícios adicionais, apresentados na sequência. É importante enfatizar que o Teorema 2 provém resultados aproximados para as distribuições das variáveis latentes e das observações futuras.

Teorema 5.2 Considerando a seguinte equação de evolução, aproximadamente consistente,

$$w \frac{\lambda_{t+h+1}}{\lambda_{t+h}} \mid \lambda_{t+h}, \mathbf{Y}_t \sim \text{Beta}(w^{h+1}a_t, (1-w)w^h a_t), \text{ para } h = 0, 1, 2, \dots,$$

onde $0 < w \leq 1$, $\mathbf{Y}_t = \{Y_0, y_1, \dots, y_t\}$, para $t = 1, 2, \dots$ e Y_0 representa a informação anterior disponível.

Se $h > 0$, as distribuições dos estados latentes no tempo $t+h$ dado toda informação disponível até o instante t são dadas por

$$\lambda_{t+h} \mid \mathbf{Y}_t, \boldsymbol{\varphi} \sim \text{Gama}(a_{t+h|t}, b_{t+h|t}) \quad (5.19)$$

onde $a_{t+h|t} = w^h a_t$ e $b_{t+h|t} = w^h b_t$. Além disso, a função de densidade preditiva das observações h passos à frente é dada por

$$p(y_{t+h} \mid \mathbf{Y}_t, \boldsymbol{\varphi}) = \frac{\Gamma(b(y_{t+h}, \boldsymbol{\varphi}) + a_{t+h|t}) a(y_{t+h}, \boldsymbol{\varphi}) (b_{t+h|t})^{a_{t+h|t}}}{\Gamma(a_{t+h|t}) [c(y_{t+h}, \boldsymbol{\varphi}) + b_{t+h|t}]^{b(y_{t+h}, \boldsymbol{\varphi}) + a_{t+h|t}}}, \quad y_{t+h} \in S(\boldsymbol{\varphi}). \quad (5.20)$$

A distribuição em (5.19) descreve a incerteza associada com as previsões do nível dinâmico e a distribuição em (5.20) descreve a incerteza associada com as previsões dos valores futuros da série temporal. Medidas de resumo, como média, mediana e percentis podem ser facilmente extraídas. Caso seja necessário a inclusão de covariáveis no modelo, resultados para previsão podem ser obtidos de maneira análoga, substituindo λ_{t+h} por μ_{t+h} .

Algumas importantes consequências dos resultados são imediatos. O papel exercido pelo fator de desconto w e sua natureza multiplicativa podem ser melhor entendidos agora. A preservação das médias dos níveis dinâmicos é mantida mesmo para *lags* grandes, porque $E[\lambda_{t+h} \mid \mathbf{Y}_t] = E[\lambda_t \mid \mathbf{Y}_t]$, para $h = 1, 2, \dots$ A natureza multiplicativa do fator de desconto w é evidente de $\text{Var}[\lambda_{t+h} \mid \mathbf{Y}_t, \boldsymbol{\varphi}] = w^{-h} \text{Var}(\lambda_{t-1} \mid \mathbf{Y}_{t-1}, \boldsymbol{\varphi})$, para $h = 1, 2, \dots$

O Teorema 5.2 é também útil para lidar com dados ausentes em séries temporais de forma exata quando instantes de tempo são menores que o tamanho amostral e de forma aproximada quando são maiores que o tamanho amostral. Se y_s não está disponível para um tempo s , então $\mathbf{Y}_s = \mathbf{Y}_{s-1}$. Além disto, a inferência sequencial requer a evolução de λ_{s-1} diretamente de λ_{s+1} . A distribuição *on-line* no tempo $s-1$ é $\lambda_{s-1} \mid \mathbf{Y}_{s-1}, \boldsymbol{\varphi} \sim \text{Gama}(a_{s-1}, b_{s-1})$. A direta aplicação do Teorema 5.2 leva a $\lambda_{s+1} \mid \mathbf{Y}_{s-1}, \boldsymbol{\varphi} \sim \text{Gama}(w^2 a_{s-1}, w^2 b_{s-1})$ e a inferência sequencial pode ser processada para $s+1$ de maneira usual. As modificações óbvias podem ser aplicadas quando l consecutivas (ao invés de uma única) observações da série temporal não estão

disponíveis. Assim, se o conjunto $\{y_s, \dots, y_{s+l-1}\}$ de l consecutivos valores da série temporal são dados ausentes, então $\mathbf{Y}_{s+l-1} = \mathbf{Y}_{s-1}$ e $\lambda_{s+l} | \mathbf{Y}_{s+l-1}, \boldsymbol{\varphi} \sim \text{Gama}(w^{l+1}a_{s-1}, w^{l+1}b_{s-1})$.

Comentários similares são feitos para os tempos irregulares observados t_1, t_2, \dots . Então, a inferência sequencial deve ser baseada na evolução de t_{i-1} para t_i , para $i = 0, 1, 2, \dots$. O Teorema 5.2 assegura que se $\lambda_{t_{i-1}} | \mathbf{Y}_{t_{i-1}}, \boldsymbol{\varphi} \sim \text{Gama}(a_{i-1}, b_{i-1})$ então $\lambda_{t_i} | \mathbf{Y}_{t_{i-1}}, \boldsymbol{\varphi} \sim \text{Gama}(w^{t_i-t_{i-1}}a_{i-1}, w^{t_i-t_{i-1}}b_{i-1})$, para $i = 0, 1, 2, \dots$ e as distribuições *on-line* e preditivas podem ser obtidas pelo Teorema 5.1.

5.2.4 Suavização

O interesse principal em muitas situações é estimar o componente de nível $\lambda = (\lambda_1, \dots, \lambda_n)'$ baseado em toda a informação disponível (\mathbf{Y}_n) ao invés da sequência de distribuições *on-line* $\lambda_t | \mathbf{Y}_t, \forall t$. Técnicas de suavização devem ser usadas nesses casos. Harvey & Fernandes (1989) apresentam uma estimativa do componente de nível em uma aplicação a uma série real, obtida através de um algoritmo aproximado de suavização do intervalo fixo (Harvey, 1989). Eles reconhecem a natureza da aproximação do esquema pela própria denominação feita por eles como um procedimento de quase-suavização.

No contexto de modelos lineares normais, a distribuição conjunta de todos os parâmetros dado toda informação disponível é normal multivariada e sua expressão é dada em Migon et al. (2005). A precisão ou matriz de covariâncias está na forma bloco tridiagonal, como um reflexo da estrutura Markoviana do modelo. Isso abre a possibilidade para criar melhores esquemas baseados no fato da matriz ser esparsa. Isso é especialmente relevante quando o tamanho da série temporal é grande. Um desses esquemas é o algoritmo FFBS (Forward Filtering Backing Sampling) proposto por Fruhwirth-Schnatter (1994) e Carter & Kohn (1994). Esse algoritmo mostra como a distribuição conjunta normal multivariada dos estados pode ser dividida em partes menores, mas ainda em componentes normais.

A decomposição do FFBS é sempre possível em MEE, mas MEE normais e lineares são apenas uma situação conhecida que leva a uma solução tratável. O Teorema 5.3 abaixo fornece uma versão não-gaussiana do FFBS para modelos deste trabalho, mostrando que eles também levam a uma solução tratável.

A distribuição conjunta de $(\lambda|\mathbf{Y}_n, \boldsymbol{\varphi})$ pode sempre ser escrita como

$$\begin{aligned} p(\lambda|\boldsymbol{\varphi}, \mathbf{Y}_n) &= p(\lambda_n|\boldsymbol{\varphi}, \mathbf{Y}_n) \prod_{t=1}^{n-1} p(\lambda_t|\lambda_{t+1}, \dots, \lambda_n, \boldsymbol{\varphi}, \mathbf{Y}_n) p(\boldsymbol{\varphi}|\mathbf{Y}_n) \\ &= p(\lambda_n|\boldsymbol{\varphi}, \mathbf{Y}_n) \prod_{t=1}^{n-1} p(\lambda_t|\lambda_{t+1}, \boldsymbol{\varphi}, \mathbf{Y}_t) p(\boldsymbol{\varphi}|\mathbf{Y}_n), \end{aligned}$$

onde a última igualdade segue da estrutura Markoviana do modelo.

A equação anterior é a chave para a solução do problema de suavização, que é realizado via as distribuições obtidas no Teorema 5.3 abaixo.

Teorema 5.3 A distribuição de $(\lambda_t|\lambda_{t+1}, \boldsymbol{\varphi}, \mathbf{Y}_t)$, é dada por

$$\lambda_t - w\lambda_{t+1}|\lambda_{t+1}, \boldsymbol{\varphi}, \mathbf{Y}_t \sim \text{Gama}((1-w)a_t, b_t), \forall t \geq 0. \quad (5.21)$$

□

A prova do Teorema 5.3 pode ser encontrada no Apêndice 4.

Baseado no Teorema 5.3, uma amostra exata da distribuição conjunta de $(\lambda|\boldsymbol{\varphi}, \mathbf{Y}_n)$ pode ser obtida seguindo o algoritmo abaixo:

1. fixe $t = n$ e amostre $p(\lambda_n|\boldsymbol{\varphi}, \mathbf{Y}_n)$, usando o Teorema 5.1 com $t = n$;
2. fixe $t = t - 1$ e amostre $p(\lambda_t|\lambda_{t+1}, \boldsymbol{\varphi}, \mathbf{Y}_t)$, usando (5.21);
3. se $t > 1$, volte para o passo 2; caso contrário, a amostra de $(\lambda_1, \dots, \lambda_n|\boldsymbol{\varphi}, \mathbf{Y}_n)$ está completa.

O resultado em (5.21) permite a implementação do passo 2 do algoritmo acima e, assim, uma amostra exata da distribuição suavizada dos estados, uma vez que os hiperpâmetros são conhecidos. Esse resultado será crucial para a inferência do vetor de dos estados mesmo quando os hiperparâmetros não forem conhecidos.

5.2.5 Adequação do modelo e ajuste

A verificação da adequação do modelo é uma parte importante na modelagem. Após o ajuste do modelo aos dados, é necessário verificar a adequação

do modelo. Na literatura, há várias sugestões de métodos de diagnóstico. Abaixo, serão descritos alguns deles.

Harvey & Fernandes (1989) sugerem alguns métodos de diagnóstico, baseados nos resíduos padronizados, conhecidos também como resíduos Pearson. Esses resíduos são definidos por

$$\nu_t^p = \frac{y_t - E(y_t | \mathbf{Y}_{t-1}, \varphi)}{DP(y_t | \mathbf{Y}_{t-1}, \varphi)}, \quad (5.22)$$

onde $DP(y_t | \mathbf{Y}_{t-1}, \varphi)$ é o desvio-padrão da distribuição de $y_t | \mathbf{Y}_{t-1}, \varphi$.

Alguns métodos de diagnóstico:

1. Examinar o gráfico dos resíduos versus o tempo e versus uma estimativa do componente de nível.
2. Verificar se a variância amostral dos resíduos padronizados é próxima de 1. Um valor maior que 1 indica superdispersão relativa ao modelo que é ajustado aos dados, entre outros métodos sugeridos.

Para maiores detalhes sobre esses e outros métodos de diagnóstico, consulte Harvey & Fernandes (1989).

Outra alternativa é usar os resíduos *deviance* (McCulagh & Nelder, 1989) que podem ser expressos como:

$$\nu_t^d = \left\{ 2 \ln \left[\frac{p(y_t | y_t, \varphi)}{p(y_t | \hat{\gamma}_t, \varphi)} \right] \right\}^{1/2}, \quad (5.23)$$

onde $\hat{\gamma}_t = E(y_t | \mathbf{Y}_{t-1}, \varphi)$.

Por exemplo, considerando o modelo Poisson, os resíduos *deviance* são dados por

$$\begin{aligned} \nu_t^d &= \left\{ 2 \ln \left[\frac{p(y_t | y_t, \varphi)}{p(y_t | \hat{\gamma}_t, \varphi)} \right] \right\}^{1/2} \\ &= \left\{ 2 \ln \left[\frac{y_t^{y_t} \exp(-y_t) / y_t!}{\hat{\gamma}_t^{y_t} \exp(-\hat{\gamma}_t) / y_t!} \right] \right\}^{1/2} \\ &= \{ 2 [y_t \ln(y_t / \hat{\gamma}_t) - (y_t - \hat{\gamma}_t)] \}^{1/2}, \end{aligned}$$

onde $\hat{\gamma}_t = E(y_t | \mathbf{Y}_{t-1}, \varphi) = a_{t+1|t}^* / b_{t+1|t}^*$. Gráficos de diagnóstico que podem ser construídos incluem resíduos versus ordem e valores ajustados ou preditos.

Às vezes, é possível ter mais de um modelo para os dados, porém é necessário escolher um deles. Portanto, abaixo, são discutidos alguns critérios para a escolha de modelos.

Critério de ajuste

De acordo com Harvey (1989), os critérios AIC e BIC são adotados sendo

$$AIC = -2l(\hat{\varphi}) + 2p$$

e

$$BIC = -2l(\hat{\varphi}) + p \ln(n),$$

em que $l(\cdot)$ é a função de log-verossimilhança, p o número de parâmetros e n o número de observações.

O critério DIC (*Deviance Information Criterion*) (Spiegelhalter et al. , 2002, ver) é usado para comparar os modelos Bayesianos. Seja $\varphi^{(1)}, \dots, \varphi^{(M)}$ uma amostra de $p(\varphi|\mathbf{Y}_n)$ via MCMC. Logo,

$$DIC = \bar{D} + k,$$

onde $\bar{D} \approx \frac{-2}{M} \sum_{i=1}^M \ln L(\varphi^{(i)}; Y_n)$ e $k \approx \bar{D} - 2 \ln L(\varphi^*; Y_n)$ é o número efetivo de parâmetros. φ^* é a média *a posteriori* dos parâmetros.

5.3 Estudos de simulação em ME Não-Gaussianos Poisson e Gama

Nesta seção, simulações Monte Carlo são realizadas para dois modelos na FMEVE: Poisson e Gama. 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 a taxa de cobertura. Todos os resultados apresentados nesta seção foram obtidos através de programas desenvolvidos no *software* Ox (Doornik , 2002).

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 e EB-mediana - foram investigadas para séries temporais de tamanho $n = 100$ e $n = 500$, geradas sob o modelo Poisson com uma covariável $x_t = \cos(2\pi t/12)$,

para $t = 1, \dots, n$, e com parâmetros $w = 0,90$ e $\beta = 1$. Duas cadeias de 8000 amostras foram geradas das quais as 5000 primeiras foram excluídas. O número de replicações Monte Carlo foi fixado em 500. O nível dos intervalos foi fixado em 0,95. A condição inicial do estado assumida foi $\lambda_0|Y_0 \sim \text{Gama}(100, 1)$. São adotadas distribuições *a priori* Uniforme(0,1) e Uniforme(-1000,1000) próprias para w e β , respectivamente.

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 as linhas tracejada e pontilhada referentes à média suavizada dos ajustes clássico e Bayesiano, respectivamente, do modelo Poisson acompanham bem o comportamento da série (linha contínua).

Das Figuras 5.2 e 5.3, percebe-se que os valores do método de Gelman e Rubin, implementados no pacote Coda (Plummer *et al.*, 2005) no *software* R (R, 2005) para a averiguação da convergência das cadeias geradas por MCMC de w e β para cada Monte Carlo estão abaixo de 1,20 (valor de referência, ver Gelman (1996)), respectivamente. As Figuras 5.4 e 5.5 apresentam a função de autocorrelação das cadeias dos parâmetros dos 500 Monte Carlos sob o modelo Poisson. Observe que tanto a função de autocorrelação média quanto os limites do envelope de 95%, decaem exponencialmente - característica típica de distribuições estacionárias. Assim, têm-se indícios da convergência das cadeias (Gelman, 1996, ver).

Da Tabela 5.2, o EMV e os estimadores Bayesianos são comparados com respeito ao vício e ao EQM. Para β , todos os estimadores têm vício e EQM muito próximos. Já para w , o EMV apresenta vício menor que os estimadores EB-média e EB-mediana, enquanto que os valores do EQM são similares. É interessante observar que quando o tamanho da amostra aumenta, o vício e o EQM diminuem. Não se observou nenhum problema na estimação de nenhum parâmetro como viés alto e etc.

Na Tabela 5.3, os intervalos de confiança e de credibilidade percentílico são comparados através da taxa de cobertura e amplitude. Para β , a taxa de cobertura do intervalo de credibilidade está mais próxima do nível nominal de 95% que os intervalos de confiança. Ambos intervalos têm taxa de cobertura similar para w . O intervalo de confiança tem a amplitude ligeiramente menor que os intervalos de credibilidade em geral.

Tab. 5.2: EMV e EB para o modelo Poisson.

n	φ	EMV	EB-Mediana	EB-Média
		estimativa	estimativa	estimativa
		Vício (EQM)	Vício (EQM)	Vício (EQM)
100	$w = 0,90$	0,889	0,874	0,868
		-0,011	-0,026	-0,032
		(0,002)	(0,003)	(0,003)
	$\beta = 1,00$	0,958	0,960	0,961
		-0,042	-0,040	-0,039
		(0,011)	(0,011)	(0,011)
500	$w = 0,90$	0,901	0,898	0,897
		0,001	-0,002	-0,003
		(0,000)	(0,000)	(0,000)
	$\beta = 1,00$	0,993	0,993	0,993
		-0,007	-0,007	-0,007
		(0,002)	(0,002)	(0,002)

Tab. 5.3: Intervalos de confiança e de credibilidade para o modelo Poisson com nível nominal de 95%.

n	φ	Int. Conf.	Int. Cred. Perc.
		limites médios amplitude (cobertura)	limites médios amplitude (cobertura)
100	$w = 0,90$	[0,786; 0,991]	[0,749; 0,952]
		0,205	0,203
		(0,97)	(0,93)
	$\beta = 1,00$	[0,765; 1,150]	[0,762; 1,165]
		0,385	0,403
		(0,92)	(0,94)
500	$w = 0,90$	[0,865; 0,938]	[0,856; 0,930]
		0,073	0,074
		(0,94)	(0,96)
	$\beta = 1,00$	[0,914; 1,072]	[0,913; 1,075]
		0,158	0,162
		(0,95)	(0,96)

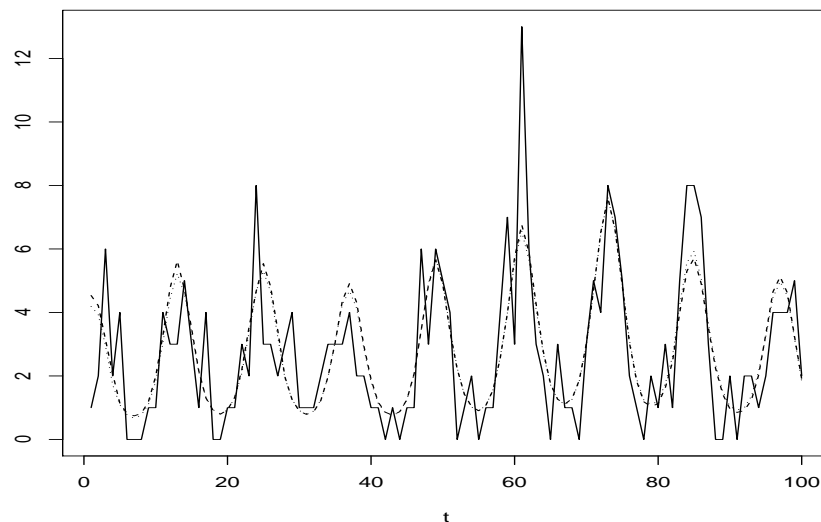


Fig. 5.1: A linha contínua representa a série temporal simulada sob o modelo Poisson, as linhas tracejada e pontilhada indicam a média suavizada dos ajustes clássico e Bayesiano, usando o método de suavização, respectivamente.

5.3.2 Modelo Gama

Da mesma forma que a simulação MC do modelo Poisson a qual foi mostrada na seção anterior, o desempenho do estimador EMV e dos estimadores Bayesianos - EB-Média e EB-Mediana - foram investigadas via experimentos de Monte Carlo para séries temporais de tamanho $n = 100$ e $n = 500$, geradas sob o modelo Gama com uma covariável $x_t = \cos(2\pi t/12)$, para $t = 1, \dots, n$, e parâmetros $\omega = 0, 90$, $\chi = 5, 00$ e $\beta = 0, 10$. Duas cadeias de 5000 amostras foram geradas das quais as 3000 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. A condição inicial do estado assumida foi $\lambda_0|Y_0 \sim Gama(100, 1)$. Distribuições *a priori* Uniforme(0,1), Uniforme(-1000,1000) e Uniforme(0,1000) próprias foram assumidas para os parâmetros w , β e χ , respectivamente.

A Figura 5.6 apresenta um exemplo de uma série simulada segundo esse modelo. As linhas tracejada e pontilhada que representam a média suavizada dos ajustes clássico e Bayesiano, respectivamente, têm um comportamento

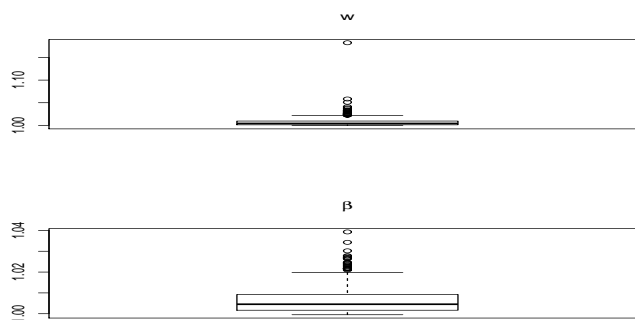


Fig. 5.2: Boxplots dos valores de Gelman e Rubin para diagnóstico de convergência das cadeias dos parâmetros w e β para os 500 Monte Carlo gerados sob o modelo Poisson com $n = 100$, respectivamente.



Fig. 5.3: Boxplots dos valores de Gelman e Rubin para diagnóstico de convergência das cadeias dos parâmetros w e β para os 500 Monte Carlo gerados sob o modelo Poisson com $n = 500$, respectivamente.

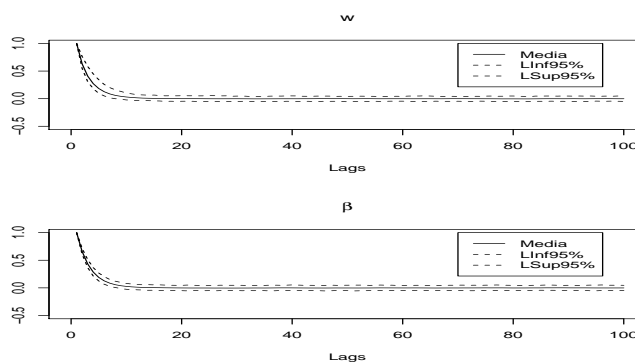


Fig. 5.4: As linhas contínua e pontilhada indicam a função de autocorrelação média e o envelope de 95% das cadeias dos parâmetros w e β para os 500 Monte Carlo gerados sob o modelo Poisson com $n = 100$, respectivamente.

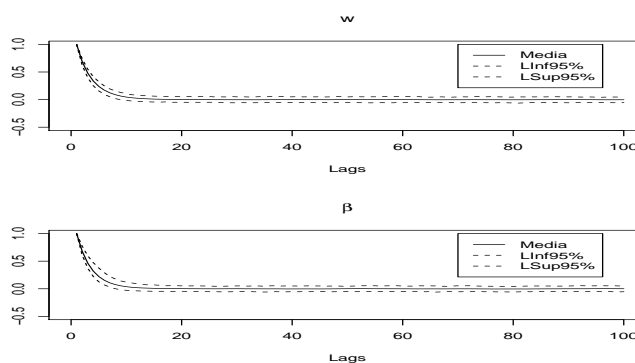


Fig. 5.5: As linhas contínua e pontilhada indicam a função de autocorrelação média e o envelope de 95% das cadeias dos parâmetros w e β para os 500 Monte Carlo gerados sob o modelo Poisson com $n = 500$, respectivamente.

Tab. 5.4: EMV e EB para o modelo Gama.

n		EMV Vício (EQM)	EB-Média Vício (EQM)	EB-Mediana Vício (EQM)
100	$\omega = 0,90$	0,916	0,875	0,885
		0,016	-0,025	-0,015
		(0,003)	(0,003)	(0,003)
	$\beta = 0,10$	0,087	0,085	0,085
		-0,013	-0,015	-0,015
		(0,003)	(0,003)	(0,003)
$\chi = 5,00$	5,264	5,487	5,430	
	0,264	0,487	0,430	
	(0,391)	(0,526)	(0,473)	
500	$\omega = 0,90$	0,901	0,895	0,896
		0,001	-0,005	-0,004
		(0,000)	(0,000)	(0,000)
	$\beta = 0,10$	0,097	0,097	0,097
		-0,003	-0,003	-0,003
		(0,001)	(0,001)	(0,001)
$\chi = 5,00$	5,033	5,075	5,065	
	0,033	0,075	0,065	
	(0,101)	(0,108)	(0,106)	

similar.

Das Figuras 5.7 e 5.8, observe que os valores do critério de Gelman e Rubin para as cadeias geradas por MCMC para w , β and χ são menores que 1,2 (valor de referência), para o modelo Gama com $n = 500$ e $n = 100$. As Figuras 5.9 e 5.10 apresentam a função de autocorrelação das cadeias dos parâmetros dos 500 Monte Carlo sob o modelo Gama. Tanto a função de autocorrelação média quanto os limites do envelope de 95%, decaem exponencialmente - característica típica de distribuições estacionárias. Assim, têm-se evidências da convergência das cadeias.

Da Tabela 5.4 - que mostra os resultados do estudo Monte Carlo dos estimadores pontuais clássico e Bayesianos -, note que os estimadores têm o mesmo EQM, exceto o parâmetro χ cujo EMV tem o EQM ligeiramente menor. O EMV e o EB-Mediana possuem o vício menor que o EB-Média para todos os parâmetros do modelo. Considerando, $n = 500$, todos os estimadores têm o EQM parecido. Não se observou nenhum problema na estimação de nenhum parâmetro como viés alto e etc.

Da Tabela 5.5, observe que os intervalos de credibilidade percentílicos possuem a taxa de cobertura mais próxima do nível nominal assumido de 0,95. Em geral, o intervalo de confiança tem a amplitude ligeiramente menor

Tab. 5.5: Intervalos de confiança e de credibilidade para o modelo Gama com nível nominal de 95%.

n		Int. Conf. limites médios amplitude (cobertura)	Int. Cred. Perc. limites médios amplitude (cobertura)
$n = 100$	$\omega = 0,90$	[0,547; 0,968] 0,421 (1,000)	[0,730; 0,965] 0,235 (0,970)
	$\beta = 0,10$	[-0,039; 0,213] 0,252 (0,970)	[-0,051; 0,222] 0,273 (0,970)
	$\chi = 5,00$	[3,755; 6,773] 3,018 (1,000)	[3,971; 7,328] 3,357 (0,970)
$n = 500$	$\omega = 0,90$	[0,855; 0,933] 0,078 (0,954)	[0,849; 0,930] 0,081 (0,948)
	$\beta = 0,10$	[0,041; 0,154] 0,113 (0,946)	[0,039; 0,155] 0,116 (0,954)
	$\chi = 5,00$	[4,395; 5,671] 1,276 (0,954)	[4,443; 5,754] 1,311 (0,946)

que o intervalo de credibilidade.

5.4 Aplicações a séries temporais reais

Nesta seção, os modelos Poisson, Normal, Gamma, Weibull, Laplace e Gaussiana Inversa são aplicados a séries temporais reais, usando as inferências clássica e Bayesiana. Em geral, distribuições *a priori* vagas e próprias como Uniforme e Gama são assumidas para os hiperparâmetros e para parâmetros latentes, respectivamente. Para a implementação da inferência Bayesiana, o algoritmo de Metropolis-Hastings é usado para a geração das cadeias. O pacote Coda do *software* R é utilizado para os métodos de diagnóstico, checando a convergência das cadeias através de métodos gráficos tais como o autocorrelograma e os gráficos da série e do traço.

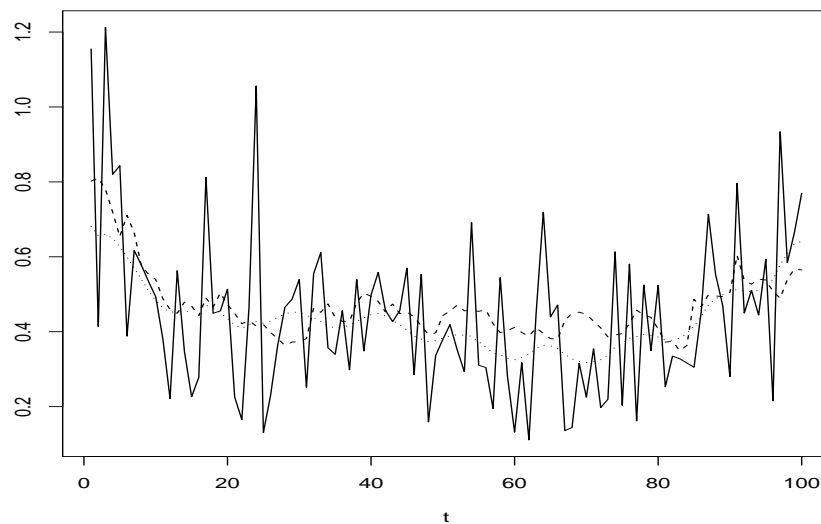


Fig. 5.6: A linha contínua representa a série temporal simulada sob o modelo Gama, as linhas tracejada e pontilhada indicam a média suavizada dos ajustes clássico e Bayesiano, usando o método de suavização, respectivamente.

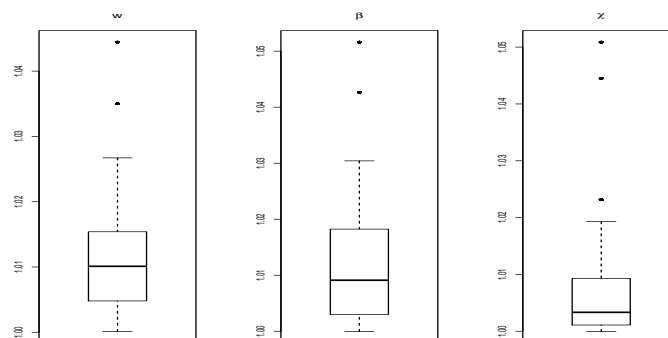


Fig. 5.7: Boxplots dos valores de Gelman e Rubin para diagnóstico de convergência das cadeias dos parâmetros ω , β e χ para os 500 Monte Carlo gerados sob o modelo Gama com $n = 100$, respectivamente.

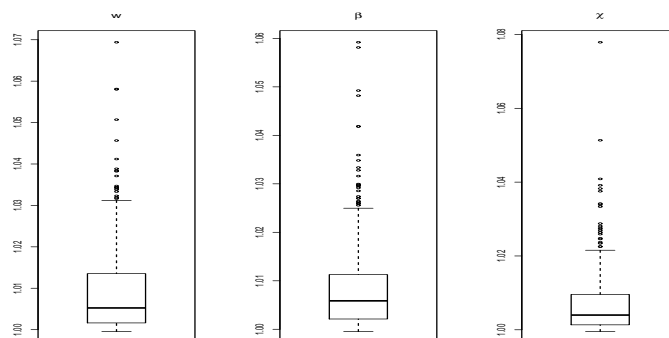


Fig. 5.8: Boxplots dos valores de Gelman e Rubin para diagnóstico de convergência das cadeias dos parâmetros ω , β e χ para os 500 Monte Carlo gerados sob o modelo Gama com $n = 500$, respectivamente.

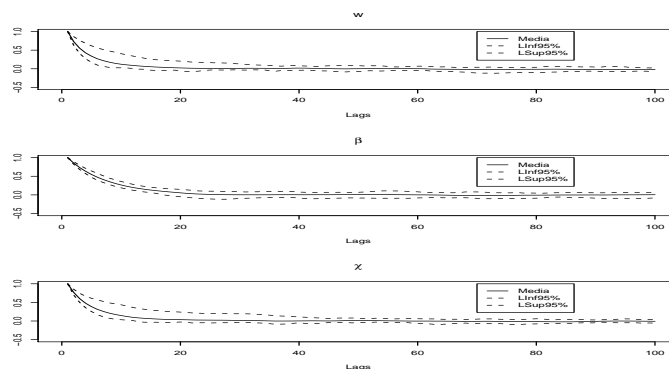


Fig. 5.9: As linhas contínua e pontilhada indicam a função de autocorrelação média e o envelope de 95% das cadeias dos parâmetros w e β para os 500 Monte Carlo gerados sob o modelo Gama com $n = 100$, respectivamente.

5.4.1 Dados de doença respiratória

O modelo Poisson é usado para os dados diários do número de pacientes com doença respiratória (DR) em São Paulo, Brasil, de 02/12/1999 a 31/12/2000 (396 observações). As últimas 10 observações são excluídas do ajuste com o objetivo de comparar as previsões, e então $n = 386$. As séries temporais do dióxido de enxofre SO_2 (x_{1t}) e monóxido de carbono CO (x_{2t}), $t = 1, \dots, n$, foram consideradas como covariáveis, já que a exposição à poluição pode aumentar o risco de doenças respiratórias (Ferris *et al.*, 1983; Alves, Garmerman and Ferreira, 2010). Essas séries foram escolhidas porque, normalmente, altas concentrações de CO produzidas por veículos encontradas em áreas urbanas e os combustíveis baseados em enxofre (por exemplo, óleos) em combustão produzem SO_2 . As séries de DR e poluentes SO_2 e CO foram obtidas, respectivamente, no site do ministério da Saúde brasileiro (<http://www.datasus.gov.br>) e Companhia Tecnológica e Ambiental de São Paulo (CETESB). Os dados são ajustados a dois modelos, os quais são comparados:

- O modelo Poisson (FMEVE) apresentado nesta tese cuja forma está na Tabela 5.1;
- O modelo Poisson com erros normais, segundo Durbin & Koopman (2001).

Os dados foram ajustados ao modelo de espaço de estados Poisson com erros normais (MEEPEN), ver Durbin & Koopman (2001),

$$\begin{aligned} y_t &\sim Po(\lambda_t), \\ \mu_t &= \ln(\lambda_t) = T_t + \beta_1 x_{1t} + \beta_2 x_{2t}, \\ T_t &= T_{t-1} + \omega_t, \omega_t \sim N(0, W), \\ T_0 &\sim N(0, 1000), \end{aligned}$$

usando o pacote `sspir` (Dethlefsen & Lundbye-Christensen, 2006) no *software* R. Ele usa a técnica de amostragem por importância para fazer inferências aproximadas. A previsão é feita de acordo com Durbin & Koopman (2001, p. 214-215) e as estimativas suavizadas dos componentes do estado são obtidos usando o *iterated extended Kalman filter* (Durbin & Koopman, 2001).

A Tabela 5.6 apresenta a EMV e as estimativas Bayesianas dos parâmetros w , β_1 e β_2 para o modelo Poisson, assim como intervalos de confiança e credibilidade para os parâmetros. As estimativas de w estão em torno de 0,78.

Os coeficientes de regressão da nossa abordagem e dos MEEPEN são similares, bem como os intervalos. Os coeficientes de regressão dos poluentes são positivos, como esperado. A significância é obtida para uma das covariáveis, considerando um nível de confiança de 95%. Essa característica é esperada devida à alta correlação entre os níveis dos poluentes. As análises foram repetidas excluindo a série *CO* não significativa, e os resultados permanecem basicamente os mesmos. Como há evidência na literatura de que ambos poluentes influenciam a série *DR*, essa covariável foi mantida no modelo.

Tab. 5.6: estimação pontual e intervalar para o modelo Poisson sobre as abordagens dos MEEPEN e FMEVE ajustados à série *DR*.

φ	FMEVE					MEEPEN	
	EMV	Int. Conf.	EB-Mediana	EB-Média	Int. Cred.	EMV	Int. Conf.
w	0,784	[0,735; 0,833]	0,778	0,775	[0,721; 0,821]	-	-
β_1	0,005	[0,001; 0,008]	0,005	0,005	[0,001; 0,008]	0,006	[0,002; 0,010]
β_2	0,017	[-0,003; 0,038]	0,017	0,017	[-0,003; 0,038]	0,010	[-0,014; 0,032]

Os valores do critério de Gelman e Rubin (Gelman, 1996) para as duas cadeias dos parâmetros w , β_1 e β_2 estão em torno de 1,00 e essas cadeias se sobrepõem, indicando convergência. O ajuste do modelo não parece apresentar nenhuma evidência de inadequação, como pode ser visto das análises dos resíduos de Pearson e deviance na Figura 5.11.

A Figura 5.12 mostra a série *DR* com suas estimativas suavizadas da média. As estimativas da média parecem seguir melhor o comportamento da série FMEVE. Medidas de resumo baseadas nos erros do ajuste e de previsão são comparados entre os dois modelos, com uma ligeira preferência pela FMEVE. Essa melhora é provavelmente devida a uma melhor adaptação às mudanças pelos erros log-beta. A figura também mostra as previsões para valores futuros, baseado em um cenário futuro para as covariáveis e nos resultados do Teorema 5.2. As previsões parecem seguir o comportamento dos valores estimados da média suavizada no final da série, e a amplitude do intervalo apresenta um crescimento com o horizonte de previsão, como esperado.

5.4.2 Série Temporal dos Retornos da Petrobrás

Os modelos Normal e Gama, descritos na Tabela 5.1, são ajustados aos dados diários dos retornos das ações da Petrobrás (Petro) no período de 03/01/1995 a 27/12/2000, totalizando 1498 observações. Os dados podem ser obtidos pelo site [http : //www.ime.usp.br/ ~ pam/ef.html](http://www.ime.usp.br/~pam/ef.html). Essa série também

foi analisada no livro de Morettin (2006) no qual vários modelos ARMA-GARCH foram utilizados e os resultados desses ajustes são aproveitados para se fazer comparação com os ajustes dos modelos deste trabalho.

A Figura 5.20 mostra o gráfico temporal do logaritmo dos retornos da Petrobrás. Morettin (2006) ajusta um modelo AR(1) à série, uma vez que a mesma apresenta uma estrutura de autocorrelação com respeito à média. Seguindo a mesma idéia, primeiro ajusta-se um modelo de nível local (MNL) com uma estrutura autoregressiva como foi abordado na Seção 2.1. Os resíduos do ajuste do MNL também são mostrados na Figura 5.20 e é usado para o ajuste dos modelos Qui-quadrado e Gama.

Logo após ao ajuste do MNL, assume-se que os resíduos do ajuste $e_t | \sigma_t^2 \sim N(0, \sigma_t^2)$. Utilizando a hipótese assumida acima, pode-se concluir que $(e_t^2 / \sigma_t^2) | \sigma_t^2 \sim \chi_{(1)}^2 \equiv \text{Gama}(1/2, 1/2)$. Portanto, $e_t^2 | \sigma_t^2 \sim \text{Gama}(1/2, (1/2) \times (1/\sigma_t^2))$, que toma a forma do modelo Gama de volatilidade estocástica com parâmetros $\chi = 1/2$ e $\lambda_t = \frac{1}{\sigma_t^2}$ (precisão). Esse modelo é denominado MNL + Qui-quadrado com um grau de liberdade ($\chi_{(1)}$).

Aos resíduos do ajuste do MNL é também ajustado o modelo Gama com o parâmetro de forma χ desconhecido de acordo com a Tabela 5.1, que é denominado MNL + Gama.

Já o modelo Normal com média desconhecida é ajustado considerando a média $y_{t-1} \phi$ com uma estrutura autoregressiva de ordem 1 e precisão $\lambda_t = \frac{1}{\sigma_t^2}$, que denomina-se modelo Normal nas tabelas e gráficos que serão mostrados abaixo. Observe que para tratar tanto a média quanto a variância da série é feito apenas o ajuste de um modelo, enquanto que nas outras propostas é realizada uma combinação de modelos, isto é, o tratamento é feito em duas etapas: um modelo para média e um modelo para a variância.

A Tabela 5.8 contém os valores da log-verossimilhança, AIC e BIC para alguns modelos ajustados e o DIC para os ajustes Bayesianos. Para se fazer uma comparação justa dos modelos na Tabela 5.8, o valor da log-verossimilhança do modelo Gama e Qui-quadrado são corrigidos pela inclusão do termo Jacobiano da transformação acima mencionada.

Observe que os modelos MNL + Gama e Normal têm os maiores valores da log-verossimilhança e os menores AIC e BIC se comparado aos modelos AR-GARCH, sendo que o modelo Normal possui o menor AIC e BIC. O modelo Gama conseguiu lidar melhor com as caudas pesadas, porém vale mencionar que os modelos Normais dinâmicos têm resultados melhores ainda que os modelos GARCH.

Tab. 5.7: Valores máximos da log-verossimilhança, AIC, BIC e DIC para os modelos ajustados aos retornos da Petrobrás.

Modelos	log-verossimilhança	AIC	BIC	DIC
AR(1)+ARCH(3)*	3152,52*	-6295,04	-6268,48	-
AR(1)+GARCH(1,1)*	3181,79*	-6355,58	-6334,33	-
AR(1)+EGARCH(1,1)*	3218,44*	-6426,88	-6400,32	-
MNL + Gama clássico	3222,05	-6434,10	-6407,54	-
MNL + Gama Bayesiano	-	-	-	-6439,90
MNL + $\chi_{(1)}$ clássico	3212,29	-6416,58	-6395,33	-
MNL + $\chi_{(1)}$ Bayesiano	-	-	-	-6422,50
Normal clássico	3210,80	-6417,60	-6406,98	-
Normal Bayesiano	-	-	-	-6417,50

Nota: * extraído da Tabela 5.1 de Morettin (2006), ajustes sob o enfoque clássico.

Os resultados da estimação dos modelos ajustados são apresentados a seguir. As EMV do MNL (enfoque clássico) são $\hat{\sigma}_\eta^2 = 0,001$, $\hat{\sigma}_\epsilon^2 = 0,000$ e $\hat{\rho} = 0,106$. O modelo Qui-quadrado é ajustado aos resíduos do ajuste do MNL, sendo que w é o único parâmetro a ser estimado. A EMV, o EB-Mediana e o EB-Média de w são 0,798, 0,798 e 0,798, respectivamente. Enquanto que os intervalos de confiança e de credibilidade de nível de 95% são $[0,769; 0,827]$ e $[0,765; 0,828]$, respectivamente. O modelo Gama é ajustado aos resíduos e w e χ são os únicos parâmetros a serem estimados. A EMV, o EB-Mediana e o EB-Média de w e χ são 0,844, 0,843 e 0,843; 0,414, 0,414 e 0,414, respectivamente. Enquanto que os intervalos de confiança e de credibilidade de nível de 95% para w e χ são $[0,825; 0,863]$ e $[0,825; 0,861]$; $[0,398; 0,430]$ e $[0,398; 0,431]$, respectivamente.

Já para o ajuste do modelo Normal com média desconhecida, a EMV, o EB-Mediana e o EB-Média de ϕ são 0,126, 0,126 e 0,126 e 0,799, 0,798 e 0,798 são para w , respectivamente. Os intervalos de confiança e de credibilidade de nível de 95% para o parâmetro ϕ são $[0,074; 0,178]$, $[0,071; 0,181]$ e $[0,770; 0,828]$, $[0,768; 0,828]$ para w , respectivamente.

O valor referente ao critério de convergência de Gelman e Rubin (Gelman, 1996) das duas cadeias para os parâmetros w e χ são iguais a 1,00. Observe também que, na Figura 5.14, as duas cadeias para cada parâmetro se sobrepõem, indicando a convergência das mesmas. O mesmo também pode ser observado para o ajuste Bayesiano dos Modelos Normal e Qui-quadrado. Não há evidências de inadequação dos modelos Qui-quadrado, Normal e Gama ajustados à série. Ver Figuras 5.17-5.19.

Na Figura 5.15, encontram-se os gráficos da estimativa da volatilidade

obtida pelos modelo Gama com o parâmetro de forma desconhecido, caso particular da FMEVE, sob os enfoques clássico e Bayesiano em vários instantes de tempo. Observe que ele possui vários valores discrepantes (pulsos) os quais são justificados por períodos de crise, conhecidos na literatura. Vale ressaltar que o ajuste conseguiu captar a queda da Nasdaq em Abril/2000 ao contrário dos modelos ajustados por Morettin (2006).

Por curiosidade, atualizou-se a série dos retornos da Petrobrás até meados de 2010, esperando que o ajuste do modelo MNL-Gama conseguisse captar a crise imobiliária que atingiu os EUA nos últimos anos. Não se trabalhou com a série atualizada desde o princípio para não se perder a comparabilidade com os modelos ajustados por Morettin (2006). Note que, na Figura 5.16, a estimativa da volatilidade evidencia a crise imobiliária nos EUA no segundo semestre de 2007 e o seu auge em 2008 com a falência do Lehman-Brothers.

5.4.3 Dados de retornos da CEMIG

O terceiro exemplo refere-se aos dados diários dos retornos do ativo da CEMIG no período de 03/01/2005 a 08/06/2011 (1590 observações). Aqui o retorno no tempo t é definido como $y_t = \frac{P_t}{P_{t-1}} - 1$, onde P_t é o preço diário de fechamento. A irregularidade dos dados devida aos feriados e finais de semana será ignorada. A Figura 5.20 apresenta o gráfico temporal dos retornos da CEMIG. Uma característica distintiva de séries financeiras é que elas normalmente apresentam variância condicional não-constante ou volatilidade. Para a análise da série da CEMIG, uma estrutura temporal de correlação é assumida para a variância e duas abordagens são consideradas. Na primeira, os modelos Normal e Laplace são considerados. Na segunda, a transformação quadrática y_t^2 é usada, para que os modelos Gama e Gaussiano Inverso (com o suporte positivo) listados na Tabela 5.1 possam ser ajustados à série. O modelo Gaussiano Inverso foi sugerido por Barndorff-Nielsen & Shephard (2001) para a modelagem de séries com volatilidade.

A Tabela 5.8 mostra a log-verossimilhança, AIC e BIC para os ajustes sob o enfoque clássico, e o critério DIC para os ajustes Bayesianos. Para uma comparação justa entre os modelos, a log-verossimilhança na segunda abordagem é corrigida pelo Jacobiano da transformação.

Os resultados na Tabela 5.8 mostram que o modelo Gaussiano Inverso pode ser descartado. Os modelos restantes são comparáveis com uma ligeira preferência pelo modelo Gama.

A Tabela 5.9 apresenta as estimativas pontuais e intervalares para os

Tab. 5.8: Valores da log-verossimilhança, AIC, BIC e DIC para os modelos ajustados aos retornos da CEMIG.

Modelos	log-verossimilhança	AIC ^(c)	BIC ^(c)	DIC ^(b)
Gama	3928,00	-4,94	-4,93	-7852,00
Normal	3926,00	-4,93	-4,93	-7850,00
Laplace	3863,74	-4,86	-4,85	-7725,80
Gaussiano Inverso	2629,25	-3,31	-3,30	-5256,25

Note: (c) classical fit. (b) Bayesian fit.

parâmetros do modelo Gama. A média e a mediana *a posteriori* são muito similares e apenas a primeira é mostrada na tabela. As estimativas pontuais e intervalares clássica e Bayesiana são muito próximas. A análise dos resíduos não mostra nenhuma inadequação do modelo (ver Figura 5.21).

Tab. 5.9: Estimação pontual e intervalar para o modelo Gama ajustado à série da CEMIG.

φ	EMV	Conf. Int.	EB	Int. Cred.
w	0,945	[0,929; 0,963]	0,945	[0,927; 0,961]
χ	0,468	[0,440; 0,490]	0,468	[0,439; 0,491]

Para ilustração, a Figura 5.22 mostra o gráfico da volatilidade obtida pelo Modelo Gama sob a abordagem Bayesiana. Os picos correspondem a períodos de crise conhecidos na literatura e são apontados pelo modelo.

5.4.4 Dados de confiabilidade de software

Essa aplicação lida com problemas de Confiabilidade de *software* (CS). Uma grande variedade de modelos é empregado na área de CS, tais como o modelo JM (Jelinski & Moranda, 1972), o modelo TGO (Goel & Okumoto, 1979), o modelo CS (Chen & Singpurwalla, 1994) e, recentemente, o modelo GJM (Chang & Liu, 2009). O modelo JM é um MEE Exponencial com uma equação de evolução Poisson, enquanto que o modelo CS é um MEE Gama. O modelo GJM é uma generalização do modelo JM para prever falhas de *software* com *debugging* imperfeito. O modelo TGO é baseado em um Processo de Poisson Não-Homogêneo (PPNH). Maiores detalhes sobre esses modelos podem ser encontrados em Chang & Liu (2009) e uma revisão geral em Chen & Singpurwalla (1994).

A distribuição Weibull é usada aqui para modelar o tempo entre 101 sucessivas falhas de computador do “System 40”, anteriormente analisado por Chen & Singpurwalla (1994) e Chang & Liu (2009). O parâmetro de desconto w foi adaptado para acomodar o tempo decorrido entre as observações e é então substituído pelo fator de desconto variando no tempo dado por $w_t = w^{z_t}$, onde $z_t = (y_t - y_{t-1})/\bar{y}$, para $t = 1, \dots, n$. Essa mudança é baseada nos comentários da Seção 5.2.3 sobre séries temporais irregulares e reflete mais adequadamente a perda de informação com a passagem do tempo. A mudança é suportada pelos dados quando seu ajuste é comparado contra o ajuste do modelo com o fator de desconto constante w .

O modelo Weibull é definido da seguinte forma:

- Equação da observação: $y_t | \mu_t, \varphi \sim \text{Weibull}(\mu_t, \nu)$;
- Função de ligação: $\mu_t = \lambda_t \exp(x_t, \beta)$, onde x_t é uma covariável;
- Equação de sistema: como em (A2) com w substituído por w_t , para $t = 1, \dots, n$;
- Condição inicial: $\lambda_0 | Y_0 \sim \text{Gama}(a_0, b_0)$.

A covariável considerada, x_t , é o número de falhas até o tempo t .

Os resultados são mostrados na Figura 5.23 e Tabela 5.10. As estimativas suavizadas da média seguem bem os valores observados da série. Os valores negativos para o parâmetro β indicam que a confiabilidade decresce à medida que o número de falhas aumenta, como esperado. As estimativas para o parâmetro de forma ν mostram evidências dos dados em favor de taxas de falhas decrescentes. O histograma da distribuição *a posteriori* dos parâmetros está na Figura 5.24 e corrobora a significância do parâmetro β e que $\nu < 1$.

Tab. 5.10: Estimação pontual e intervalar (nível de 90%) para o modelo Weibull ajustado aos dados do “System 40”.

φ	EMV	Int. Conf.	EB	Int. Cred.
w	0,888	[0,822; 0,954]	0,885	[0,811; 0,961]
β	-0,028	[-0,038; -0,018]	-0,027	[-0,038; -0,016]
ν	0,716	[0,621; 0,812]	0,720	[0,622; 0,822]

Para avaliar o desempenho preditivo do modelo Weibull versus os outros modelos, a Soma dos Erros Absolutos (SEA) é usada. Essa estatística é

definida como $SEA = \sum_{i=1}^n |y_t - \tilde{y}_t|$, onde \tilde{y}_t é a média suavizada do modelo Weibull, obtida via o procedimento de suavização da Seção 5.2.4.

Como pode ser visto da Tabela 5.11, o modelo Weibull usando a média suavizada tem um desempenho superior aos modelos CS, TGO e GJM. Então, pode-se concluir que a melhora dos valores do SEA é provavelmente devida aos procedimentos de inferência para os modelos não-gaussianos propostos neste trabalho.

Tab. 5.11: valores do SEA para os modelos Weibull, GJM, TGO e CS para os dados do “System 40”.

Modelos	SEA ($\times 10^7$)
Weibull	1,6397
GJM	1,8855
TGO	2,6133
CS	2,6328

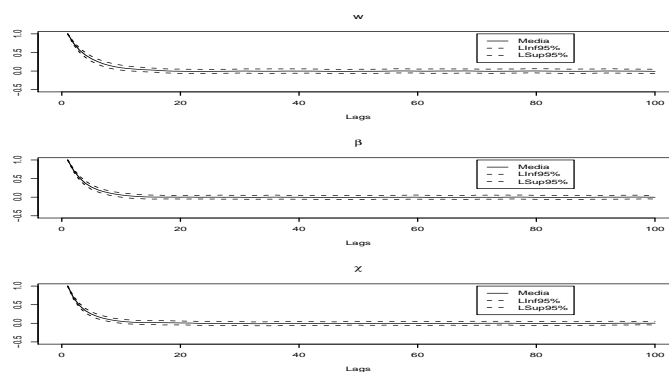


Fig. 5.10: As linhas contínua e pontilhada indicam a função de autocorrelação média e o envelope de 95% das cadeias dos parâmetros w e β para os 500 Monte Carlo gerados sob o modelo Gama com $n = 500$, respectivamente.

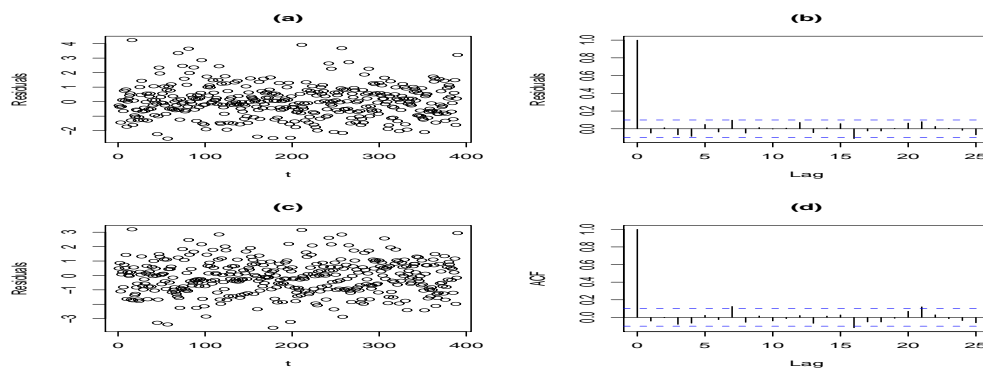


Fig. 5.11: Análise dos resíduos de Pearson e deviance para o modelo Poisson. (a): Gráfico temporal dos resíduos de Pearson; (b): autocorrelograma dos resíduos de Pearson; (c): Gráfico temporal dos resíduos deviance; (d): autocorrelograma dos resíduos deviance.

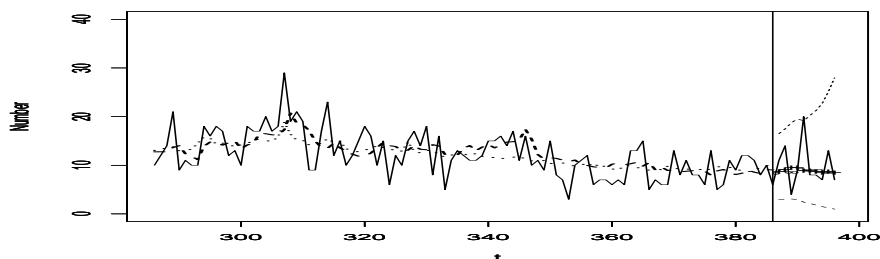


Fig. 5.12: as linhas contínua, tracejada e pontilhada representam, respectivamente, a série DR, a média suavizada da estimação Bayesiana obtida através da abordagem FMEVE e a média suavizada obtida através da abordagem MEEPEN. A linha vertical separa os dados ajustados do horizonte de previsão. a) média FMEVE: \circ ; b) média MEEPEN: $+$; FMEVE 95% interval limits: $- - -$.

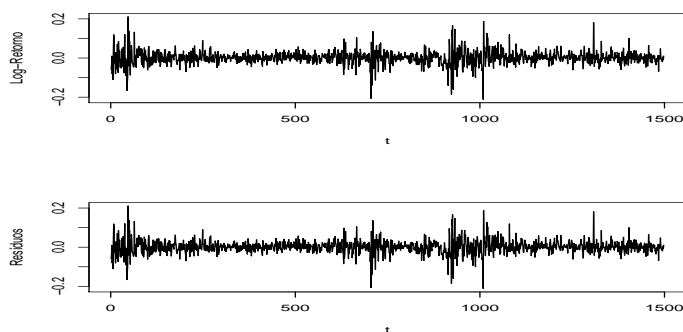


Fig. 5.13: Gráficos do log-retorno da Petrobrás e resíduos do ajuste do MNL, respectivamente.

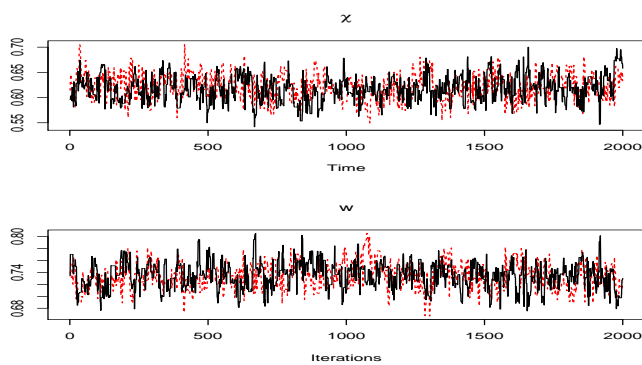


Fig. 5.14: Gráficos das duas cadeias geradas por MCMC dos parâmetros w e χ do ajuste Bayesiano do modelo Gama com χ desconhecido à série log-retorno Petrobrás, respectivamente.

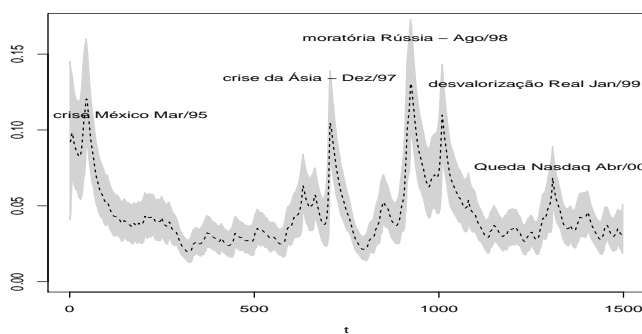


Fig. 5.15: As linhas contínua e pontilhada representam a estimativa suavizada da volatilidade obtida pelo ajuste do modelo Gama com χ desconhecido à série log-retorno Petrobrás sob os enfoques clássico e Bayesiano, respectivamente.

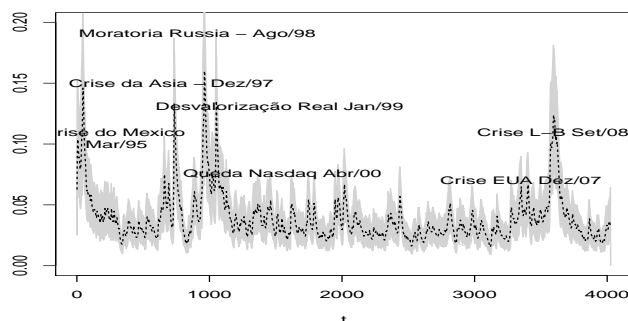


Fig. 5.16: Estimativa suavizada e limites de 95% da volatilidade obtida pelo ajuste do modelo Gama com o parâmetro de forma desconhecido à série log-retorno Petrobrás com observações até a metade de 2010.

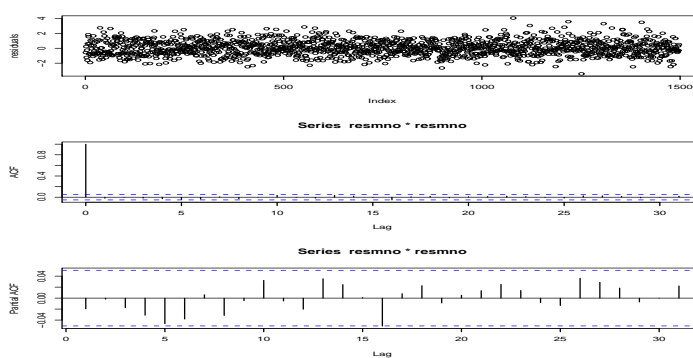


Fig. 5.17: Análise dos resíduos Deviance do modelo Normal com média desconhecida à série log-retorno Petrobrás.

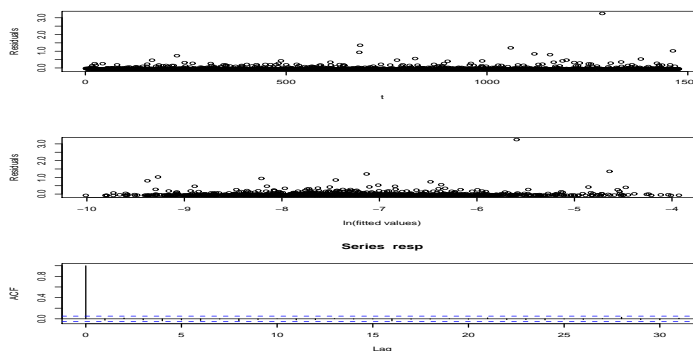


Fig. 5.18: Análise dos resíduos de Pearson do modelo Qui-quadrado à série log-retorno Petrobrás.

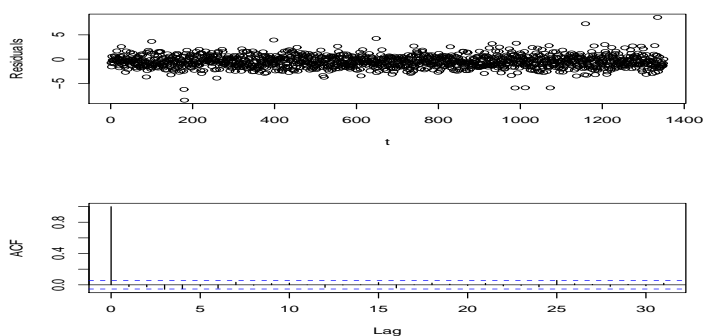


Fig. 5.19: Análise dos resíduos deviance do modelo Gama com χ desconhecido da série log-retorno Petrobrás.

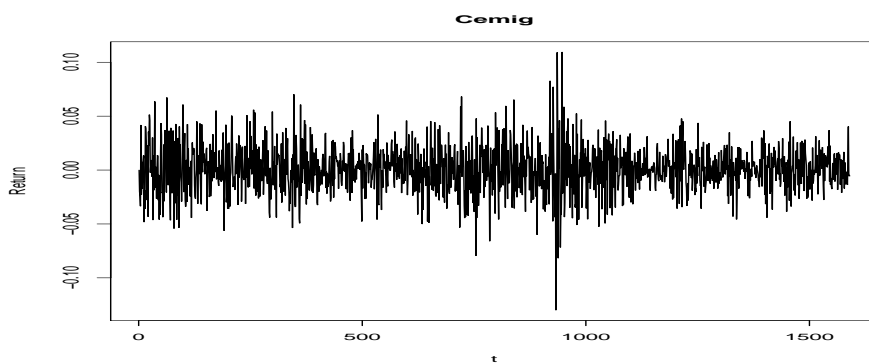


Fig. 5.20: Série de retornos da CEMIG.

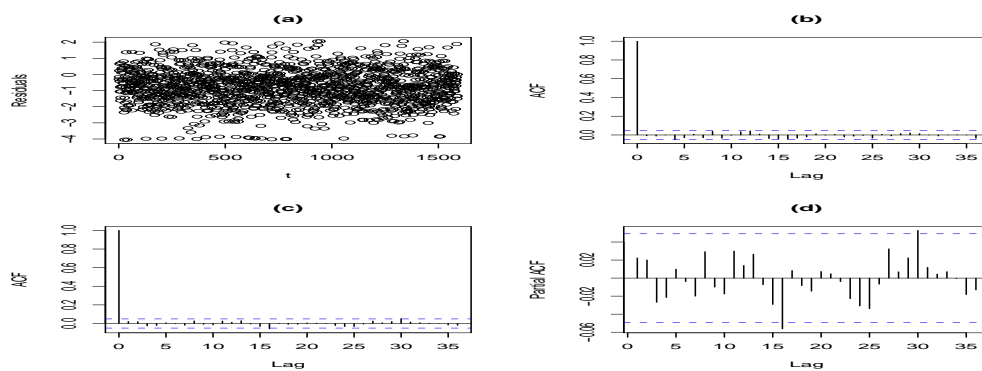


Fig. 5.21: Análise dos resíduos Deviance para o modelo Gama ajustado à série de retornos da CEMIG. (a): Gráfico temporal dos resíduos Deviance; (b): autocorrelograma dos resíduos Deviance; (c): autocorrelograma do quadrado dos resíduos; (d): autocorrelograma parcial do quadrado dos resíduos.

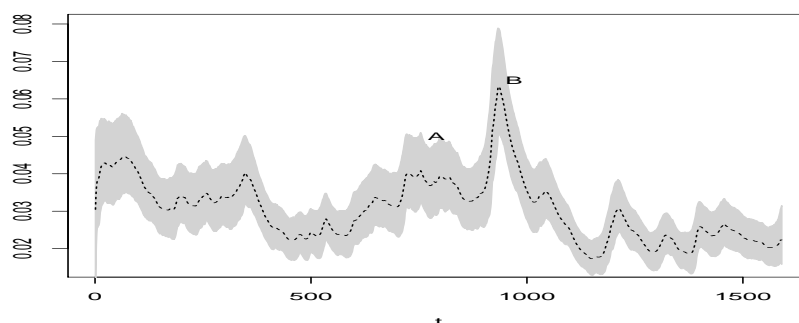


Fig. 5.22: Os dados CEMIG: A linha tracejada representa a estimativa suavizada da volatilidade estocástica, obtida pelo ajuste do modelo Gama sob a abordagem Bayesiana. As linhas hachuradas indicam os intervalos de credibilidade de nível de 95%. A representa a crise imobiliária nos EUA em Agosto de 2007, B representa a crise do L-B em Setembro de 2008.

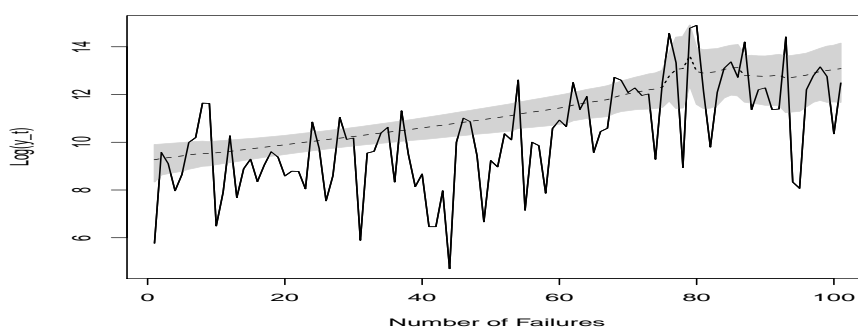


Fig. 5.23: Dados “System 40”: as linhas contínuas e tracejadas representam, respectivamente, a série e as estimativas suavizadas da média dos dados, obtido pelo ajuste do modelo Weibull sob a perspectiva Bayesiana. A área em cinza indicam os intervalos de credibilidade de nível de 95%.

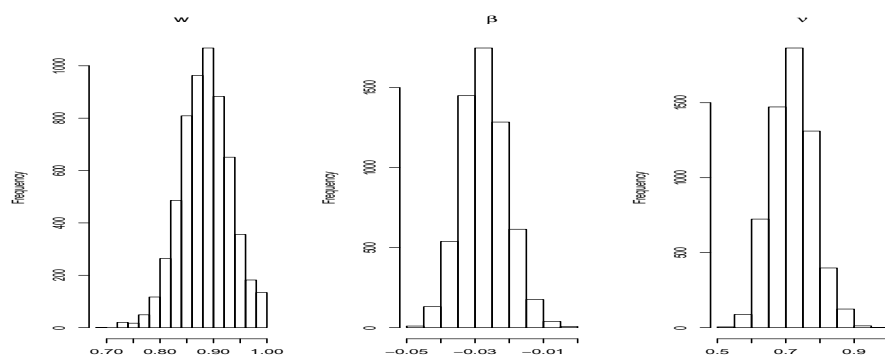


Fig. 5.24: Dados “System 40”: Histograma da distribuição *a posteriori* dos parâmetros do modelo Weibull.

6. CONCLUSÃO E CONSIDERAÇÕES FINAIS

Neste trabalho, abordaram-se os modelos estruturais Gaussianos sob o contexto de previsão dos valores futuros da série e suavização dos componentes de estados (latentes). Os modelos não-gaussianos foram explorados para uma classe de modelos denominada família de modelos de espaço de estados não-gaussianos de verossimilhança exata. Os objetivos desta tese foram atingidos nos três capítulos anteriores nos quais foram feitas várias propostas, justificadas teoricamente e/ou empiricamente para os MEE gaussianos e não-gaussianos. A seguir, é feita uma breve descrição de cada tópico abordado nesses capítulos, assim como idéias e propostas de trabalhos futuros.

Vários métodos de previsão em MEE gaussianos foram implementados e comparados, os quais levam em conta a incerteza associada à estimação dos hiperparâmetros. Além disso, foram feitas propostas, que são variações dos procedimentos de Hamilton (1986) e Ansley & Kohn (1986) e um novo procedimento *bootstrap* (NPB) de previsão foi desenvolvido. Os diversos métodos não foram robustos a hipótese de não-gaussianidade com forte assimetria da distribuição dos erros da equação das observações, mas foram com assimetria de leve a moderada. As aplicações foram satisfatórias.

Futuramente, será desenvolvido métodos de previsão que não consideram a hipótese de normalidade para os valores futuros, principalmente no NPB, já que foco desta tese na parte de previsão foi MEE gaussianos. Segundo Stoffer & Wall (2004), o procedimento *bootstrap* de Stoffer & Wall (1991) suporta também a hipótese de não-gaussianidade e, neste caso, o EMV dos hiperparâmetros passa a ser o estimador de quasi-verossimilhança gaussiana dos hiperparâmetros, o qual também é consistente (Harvey, 1989, ver).

Vários métodos de estimação do EQM de suavização do vetor de estados foram estudados e algumas variações dos métodos de Hamilton (1986) e Ansley & Kohn (1986) foram propostas. Em geral, os resultados mostraram um desempenho melhor do método de PT (Pfeffermann & Tiller, 2005), porém vale ressaltar que esse método possui um tempo computacional maior, devido a número maior de operações que devem ser realizadas. Os procedimentos de HaB e AKB apresentaram resultados razoáveis, considerando a simplicidade

de implementação dos mesmos e o tempo computacional menor. Os ajustes das séries reais ilustram bem o que se queria mostrar quanto à subestimação do EQM de suavização quando os hiperparâmetros devem ser estimados em um contexto de amostras pequenas.

Trabalhos futuros incluem o estudo desses procedimentos em MEE mais complexos e uma aplicação mais elaborada considerando a inflação e/ou a demanda de energia. É interessante permitir que alguns fatores sejam componentes estocásticos e dinâmicos, os quais poderão ser estimados e uma estimativa corrigida do seu EQM pode ser obtida pelos procedimentos abordados nesta tese.

Neste trabalho, também foi apresentada uma nova família de modelos dinâmicos não-gaussianos (FMEVE), bem como casos particulares da mesma e uma forma de se fazer suavização do componente de nível da série. Através de experimentos Monte Carlo, foi averiguado o comportamento dos estimadores pontuais e intervalares (clássico e Bayesiano) dos parâmetros dos modelos Poisson e Gama em amostras finitas. Os resultados mostraram que ambos estimadores possuem um EQM relativamente pequeno. Já os estimadores intervalares têm um comportamento parecido quanto à taxa de cobertura, entretanto o intervalo de confiança possui uma amplitude menor.

Vários novos modelos particulares pertencentes à família foram introduzidos e aplicados a séries reais. Além disso, três teoremas foram desenvolvidos, suportando as nossas propostas de filtragem, de previsão dos valores futuros, bem como de suavização do vetor de estados latentes. Vale destacar que um artigo já foi escrito e submetido a uma excelente revista com essa parte do trabalho de modelos não-Gaussianos.

Com respeito aos modelos de resposta não-gaussiana, vários trabalhos podem ser desenvolvidos. Um estudo explorando as propriedades da família seria muito relevante. Casos específicos de interesse da família não-gaussiana de modelos de espaço de estados com verossimilhança exata podem ser encontrados. Uma equação de evolução como uma estrutura autoregressiva poderia ser proposta, já que, como mostrado na Subseção 5.5.2, os modelos da FMEVE têm desempenho superior se comparado a outros, principalmente, na modelagem da volatilidade estocástica (estudos e pesquisas, nesta direção, estão sendo desenvolvidos).

O modelo exponencial por partes (Gamberman, 1994) é muito utilizado e de interesse principalmente em estudos de confiabilidade e se encaixa nos moldes dos modelos de resposta não-Gaussiana.

Testes de hipóteses, intervalos de confiança e correções de viés podem

ser explorados, utilizando a técnica *bootstrap* na FMEVE, a fim de que se possa fazer inferência sobre os parâmetros dos modelos em estudo. Franco *et al.* (1999) trabalharam com testes *bootstrap* para os hiperparâmetros em ME gaussianos. Um outro trabalho interessante que poderia se desenvolvido é a comparação do modelo de resposta não-gaussiana com os modelos lineares dinâmicos generalizados (West, Harrison & Migon, 1985; West & Harrison, 1997) seja via experimentos Monte Carlo seja utilizando séries temporais reais.

APÊNDICE

Apêndice 1

Demonstração do Teorema 4.1 (Um novo procedimento de previsão *Bootstrap*)

Antes, demonstram-se os seguintes lemas que serão utilizados na prova do resultado principal.

Lema A: Se o modelo está na forma de espaço de estados (equações 2.3-2.4) e utilizando o procedimento proposto por Stoffer & Wall (1991) - descrito no Capítulo 2, então $\hat{\psi}^{*(b)} \xrightarrow{P} \hat{\psi}$, com $b = 1, \dots, B$, quando $B \rightarrow \infty$ e $\hat{\psi} \xrightarrow{P} \psi$, $l \rightarrow \infty$.

Prova:

Segundo Stoffer & Wall (1991, pg. 1031), $\hat{\psi}^{*(b)} \xrightarrow{P} \hat{\psi}$, com $b = 1, \dots, B$, quando $B \rightarrow \infty$, ou seja, a distribuição do EMV $\hat{\psi}$ pode ser aproximada pela distribuição de $\hat{\psi}^{*(b)}$. Além disso, quando o tamanho amostral é suficientemente grande ($l \rightarrow \infty$), $\hat{\psi} \xrightarrow{P} \psi$ (propriedade assintótica dos EMV sob certas condições de regularidade) e as distribuições de $\hat{\psi}$ e $\hat{\psi}^{*(b)}$ são equivalentes. Esse artigo garante que $\hat{\psi}^{*(b)} \xrightarrow{P} \hat{\psi}$ quando $B \rightarrow \infty$ e $\hat{\psi} \xrightarrow{P} \psi$ quando $l \rightarrow \infty$.

□

Lema B: Se $\hat{\psi} \xrightarrow{P} \psi$, quando $l \rightarrow \infty$, então $\hat{\eta}_t \xrightarrow{P} \eta_t$ e $\hat{\epsilon}_t \xrightarrow{P} \epsilon_t$ em que $\eta_t \sim N(0, Q_t)$ e $\epsilon_t \sim N(0, \sigma_\epsilon^2)$.

Prova:

$\hat{\eta}_t = \hat{Q}_t^{1/2} Z_t$, onde $Z_t \sim N(0, 1)$

$Z_t \xrightarrow{P} Z$, onde $Z_t \sim N(0, 1)$ e $Z \sim N(0, 1)$.

$\hat{Q}_t^{1/2} \xrightarrow{P} Q_t^{1/2}$, $l \rightarrow \infty$.

Segundo um Teorema em Sen & Singer (1993),

$\hat{\eta}_t = \hat{Q}_t^{1/2} Z_t \xrightarrow{P} \eta_t = Q_t^{1/2} Z$.

Similarmente, pode-se provar que $\hat{\epsilon}_t \xrightarrow{P} \epsilon_t$, quando $l \rightarrow \infty$, onde $\epsilon_t \sim N(0, \sigma_\epsilon^2)$.

□

Observação: Se $\hat{\psi}^{*(b)} \xrightarrow{P} \hat{\psi}$ quando $B \rightarrow \infty$, pode-se mostrar que $\hat{\eta}_t^{*(b)} \xrightarrow{P} \hat{\eta}_t$ e $\hat{\epsilon}_t^{*(b)} \xrightarrow{P} \hat{\epsilon}_t$ em que $\hat{\eta}_t \sim N(0, \hat{Q}_t)$ e $\hat{\epsilon}_t \sim N(0, \hat{\sigma}_\epsilon^2)$.

Lema C: Se g é uma função contínua e $\hat{\psi}^{*(b)} \xrightarrow{P} \hat{\psi}$ quando $B \rightarrow \infty$ e $\hat{\psi} \xrightarrow{P} \psi$ quando $l \rightarrow \infty$, então

$$g(\hat{\psi}^{*(b)}) \xrightarrow{P} g(\hat{\psi})$$

e

$$g(\hat{\psi}) \xrightarrow{P} g(\psi).$$

Prova:

Funções contínuas preservam a convergência em probabilidade (Sen & Singer, 1993, ver), logo

$g(\hat{\psi}^{*(b)}) \xrightarrow{P} g(\hat{\psi})$, quando $B \rightarrow \infty$ e $g(\hat{\psi}) \xrightarrow{P} g(\psi)$, quando $l \rightarrow \infty$.

□

Observação: Usando o Lema C, $z'_{n+k}(\hat{\psi}^{*(b)}) \xrightarrow{P} z'_{n+k}(\hat{\psi})$ e $T_{n+k}(\hat{\psi}^{*(b)}) \xrightarrow{P} T_{n+k}(\hat{\psi})$, quando $B \rightarrow \infty$.

$z'_{n+k}(\hat{\psi}) \xrightarrow{P} z'_{n+k}(\psi)$ e $T_{n+k}(\hat{\psi}) \xrightarrow{P} T_{n+k}(\psi)$, quando $l \rightarrow \infty$, desde que z'_{n+k} e T_{n+k} sejam funções contínuas.

Prova do Teorema 4.1:

Da Equação (4.1), tem-se que

$$\begin{aligned} \alpha_{n+k}^{*(b)} = & \prod_{i=1}^k T_{n+i}(\hat{\psi}^{*(b)}) \hat{\alpha}_n^{*(b)} + \sum_{i=1}^k \prod_{j=1}^{k-i} T_{n+k-j+1}(\hat{\psi}^{*(b)}) c_{n+i} + \\ & \sum_{i=1}^k \left(\prod_{j=1}^{k-i} T_{n+k-j+1}(\hat{\psi}^{*(b)}) \right) R'_{n+i} \hat{\eta}_{n+i}^*, \end{aligned} \quad (6.1)$$

onde $\hat{\eta}_{n+i}^* \sim N(\mathbf{0}, \hat{Q}_t^*)$ e $\hat{\alpha}_n^{*(b)}$ é gerado de uma distribuição Normal com média $a_n(\hat{\psi}^{*(b)})$ e matriz de covariância $\mathbf{P}_n(\hat{\psi}^{*(b)}) + [a_n(\hat{\psi}^{*(b)}) - a_n(\hat{\psi})] \times [a_n(\hat{\psi}^{*(b)}) - a_n(\hat{\psi})]'$.

1. Observe que

$$\hat{\alpha}_n^{*(b)} \sim N \left\{ a_n(\hat{\psi}^{*(b)}), \mathbf{P}_n(\hat{\psi}^{*(b)}) + [a_n(\hat{\psi}^{*(b)}) - a_n(\hat{\psi})] \times [a_n(\hat{\psi}^{*(b)}) - a_n(\hat{\psi})]' \right\}$$

Note que

$$\begin{aligned} a_n(\hat{\psi}^{*(b)}) &\xrightarrow{P} a_n(\hat{\psi}), \\ \mathbf{P}_n(\hat{\psi}^{*(b)}) &\xrightarrow{P} \mathbf{P}_n(\hat{\psi}), \end{aligned}$$

quando $B \rightarrow \infty$, pelos Lemas A e C. a_n e \mathbf{P}_n são funções contínuas por se tratarem da média e da variância de uma distribuição normal. Segundo Watanabe (1985),

$$\begin{aligned} a_n(\hat{\psi}) &\xrightarrow{P} a_n(\psi), \\ \mathbf{P}_n(\hat{\psi}) &\xrightarrow{P} \mathbf{P}_n(\psi), \end{aligned}$$

quando $l \rightarrow \infty$.

O segundo termo da matriz de covariâncias

$$[a_n(\hat{\psi}^{*(b)}) - a_n(\hat{\psi})] \times [a_n(\hat{\psi}^{*(b)}) - a_n(\hat{\psi})]' \xrightarrow{P} \mathbf{0},$$

quando $B \rightarrow \infty$, uma vez que $a_n(\hat{\psi}^{*(b)}) \xrightarrow{P} a_n(\hat{\psi})$. Portanto, dada toda informação disponível até o instante n ($\mathbf{Y}_n = [y_{n-l+1}, \dots, y_n]$),

$\hat{\alpha}_n^{*(b)} \xrightarrow{P} N[a_n(\hat{\psi}), \mathbf{P}_n(\hat{\psi})]$, isto é,

$$\hat{\alpha}_n^{*(b)} \xrightarrow{P} \hat{\alpha}_n,$$

em que $\hat{\alpha}_n \sim N[a_n(\hat{\psi}), \mathbf{P}_n(\hat{\psi})]$, quando $B \rightarrow \infty$.

E

$$\hat{\alpha}_n \xrightarrow{P} \alpha_n,$$

em que $\alpha_n \sim N[a_n(\psi), \mathbf{P}_n(\psi)]$, quando $l \rightarrow \infty$.

Logo, novamente aplicando o Lema C e sabendo que $T_{n+i}(\hat{\psi}^{*(b)}) \xrightarrow{P} T_{n+i}(\hat{\psi})$ e $T_{n+i}(\hat{\psi}) \xrightarrow{P} T_{n+i}(\psi)$ (observação do Lema C), o primeiro termo da Equação (6.1) converge em probabilidade para $\prod_{i=1}^k T_{n+i}(\hat{\psi}) \hat{\alpha}_n$,

quando $B \rightarrow \infty$ e $\prod_{i=1}^k T_{n+i}(\hat{\psi}) \hat{\alpha}_n$ converge em probabilidade para $\prod_{i=1}^k T_{n+i}(\psi) \alpha_n$ quando $l \rightarrow \infty$.

2. c_{n+i} é uma constante e $T_{n+k-j+1}(\hat{\psi}^{*(b)}) \xrightarrow{P} T_{n+k-j+1}(\hat{\psi})$ e $T_{n+k-j+1}(\hat{\psi}) \xrightarrow{P} T_{n+k-j+1}(\psi)$, logo o segundo termo da Equação (6.1) converge em probabilidade para $\sum_{i=1}^k \left(\prod_{j=1}^{k-i} T_{n+k-j+1}(\hat{\psi}) \right) c_{n+i}$, quando $B \rightarrow \infty$ e $\sum_{i=1}^k \left(\prod_{j=1}^{k-i} T_{n+k-j+1}(\hat{\psi}) \right) c_{n+i}$

converge em probabilidade para $\sum_{i=1}^k \left(\prod_{j=1}^{k-i} T_{n+k-j+1}(\psi) \right) c_{n+i}$, quando $l \rightarrow \infty$.

3. Usando o Lema A, B e C, tem-se que $\hat{\eta}_{n+i}^* \xrightarrow{P} \hat{\eta}_{n+i}$ e $\hat{\eta}_{n+i} \xrightarrow{P} \eta_{n+i}$, $T_{n+k-j+1}(\hat{\psi}^{*(b)}) \xrightarrow{P} T_{n+k-j+1}(\hat{\psi})$ e $T_{n+k-j+1}(\hat{\psi}) \xrightarrow{P} T_{n+k-j+1}(\psi)$, $B \rightarrow \infty$ e $l \rightarrow \infty$, respectivamente, logo

$$\sum_{i=1}^k \left(\prod_{j=1}^{k-i} T_{n+k-j+1}(\hat{\psi}^{*(b)}) \right) R'_{n+i} \hat{\eta}_{n+i}^* \xrightarrow{P} \sum_{i=1}^k \left(\prod_{j=1}^{k-i} T_{n+k-j+1}(\hat{\psi}) \right) R'_{n+i} \hat{\eta}_{n+i}$$

e

$$\sum_{i=1}^k \left(\prod_{j=1}^{k-i} T_{n+k-j+1}(\hat{\psi}) \right) R'_{n+i} \hat{\eta}_{n+i} \xrightarrow{P} \sum_{i=1}^k \left(\prod_{j=1}^{k-i} T_{n+k-j+1}(\psi) \right) R'_{n+i} \eta_{n+i}$$

pelo Lema C.

Somando os três itens anteriores, pelo Lema C, tem-se que

$$\hat{\alpha}_{n+k}^{*(b)} \xrightarrow{P} \hat{\alpha}_{n+k}, \quad (6.2)$$

$$\hat{\alpha}_{n+k} \xrightarrow{P} \alpha_{n+k}, \quad (6.3)$$

quando $B \rightarrow \infty$ e $l \rightarrow \infty$, respectivamente.

Da Equação (2.3), tem-se que

$$y_{n+k}^* = z'_{n+k}(\hat{\psi}^{*(b)}) \hat{\alpha}_{n+k}^* + d_{n+k} + \hat{\epsilon}_{n+k}^*,$$

onde $\hat{\epsilon}_{n+k}^* \sim N(0, \hat{h}_{n+k}^*)$.

Segundo os resultados anteriores e os Lemas A-C:

1. $\hat{\alpha}_{n+k}^* \xrightarrow{P} \hat{\alpha}_{n+k}$ e $\hat{\alpha}_{n+k} \xrightarrow{P} \alpha_{n+k}$.
2. $z'_{n+k}(\hat{\psi}^{*(b)}) \xrightarrow{P} z'_{n+k}(\hat{\psi})$ e $z'_{n+k}(\hat{\psi}) \xrightarrow{P} z'_{n+k}(\psi)$.
3. $\hat{\epsilon}_{n+k}^* \xrightarrow{P} \hat{\epsilon}_{n+k}$, $\hat{\epsilon}_{n+k} \sim N(0, \hat{h}_{n+k})$ e $\hat{\epsilon}_{n+k} \xrightarrow{P} \epsilon_{n+k}$, $\epsilon_{n+k} \sim N(0, h_{n+k})$.

Portanto, fazendo o produto do item (2) com o item (1) e somando com o último, tem-se que

$$y_{n+k}^* | \mathbf{Y}_n \xrightarrow{P} z'_{n+k}(\hat{\psi}) \hat{\alpha}_{n+k} + d_{n+k} + \hat{\epsilon}_{n+k} = \hat{y}_{n+k},$$

onde $\hat{\epsilon}_{n+k} \sim N(0, \hat{h}_{n+k})$ e

$$\hat{y}_{n+k} | \mathbf{Y}_n \xrightarrow{P} z'_{n+k} \alpha_{n+k} + d_{n+k} + \epsilon_{n+k},$$

onde $\epsilon_{n+k} \sim N(0, h_{n+k})$.

Portanto, conclui-se que

$$y_{n+k}^* | \mathbf{Y}_n \xrightarrow{P} \hat{y}_{n+k}$$

e

$$\hat{y}_{n+k} | \mathbf{Y}_n \xrightarrow{P} y_{n+k}$$

onde $\mathbf{Y}_n = [y_{n-l+1}, \dots, y_n]$ é toda informação disponível até o instante n , quando $B \rightarrow \infty$ e $l \rightarrow \infty$, respectivamente.

□

Apêndice 2

Demonstração do Teorema 5.1

Assumindo a definição do modelo na Seção 5.2, podem-se obter as equações de atualização. Será feita a prova dos itens 1, 2, 3, 4 e 5 do Teorema 5.1.

OBS.: Para a notação não ficar muito carregada, omitir-se-à o vetor φ nos Lemas II e III e na prova dos Resultados (1-2).

A fim de provar o teorema, os seguintes lemas são necessários.

Lema II. A distribuição de $\lambda_t = w^{-1} \lambda_{t-1} \varsigma_t$.

Prova:

sabendo que $\varsigma_t \sim B(wa_{t-1}, (1-w)a_{t-1})$ pelo suposição **A2** com $h = 0$ e usando o método Jacobiano, tem-se que

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

□

Lema III.

$\lambda_t|\mathbf{Y}_{t-1} \sim \text{Gama}(a_{t|t-1}, b_{t|t-1})$.

Prova:

Para tanto, integra-se fora em λ_{t-1} e utiliza-se o Lema II. Logo,

$$\begin{aligned} p(\lambda_t|\mathbf{Y}_{t-1}, \varphi) &= \int p(\lambda_{t-1}|\mathbf{Y}_{t-1}, \varphi)p(\lambda_t|\lambda_{t-1}, \mathbf{Y}_{t-1}, \varphi)d\lambda_{t-1} \\ &= \int_{w\lambda_t}^{\infty} \left[\frac{\lambda_{t-1}^{a_{t-1}-1} \exp(-b_{t-1}\lambda_{t-1})}{\Gamma(a_{t-1})b_{t-1}^{-a_{t-1}}} \right] \left[\frac{w\lambda_{t-1}^{-1} \left(\frac{w\lambda_t}{\lambda_{t-1}}\right)^{wa_{t-1}-1} \left(1 - \frac{w\lambda_t}{\lambda_{t-1}}\right)^{(1-w)a_{t-1}-1}}{\frac{\Gamma(wa_{t-1})\Gamma((1-w)a_{t-1})}{\Gamma(a_{t-1})}} \right] d\lambda_{t-1} \end{aligned}$$

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

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

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

$$\begin{aligned}
p(\lambda_t | \mathbf{Y}_{t-1}, \varphi) &= c \int_0^{\infty} \exp[-b_{t-1}(z + w\lambda_t)] z^{(1-w)a_{t-1}-1} dz \\
&= c \exp(-wb_{t-1}\lambda_t) \int_0^{\infty} \exp(-b_{t-1}z) z^{(1-w)a_{t-1}-1} dz \\
&= \frac{w(w\lambda_t)^{wa_{t-1}-1}}{\Gamma(a_{t-1})b_{t-1}^{-a_{t-1}} B(wa_{t-1}, (1-w)a_{t-1})} \frac{\Gamma((1-w)a_{t-1})}{b_{t-1}^{(1-w)a_{t-1}}} \exp(-wb_{t-1}\lambda_t) \\
&= \frac{w^{wa_{t-1}-1+1} (\lambda_t)^{wa_{t-1}-1} \Gamma((1-w)a_{t-1})}{\Gamma(a_{t-1})b_{t-1}^{-wa_{t-1}} \Gamma(wa_{t-1}) \Gamma((1-w)a_{t-1}) [\Gamma(a_{t-1})]^{-1}} \exp(-wb_{t-1}\lambda_t) \\
&= \frac{\lambda_t^{wa_{t-1}-1} \exp(-wb_{t-1}\lambda_t)}{(wb_{t-1})^{-wa_{t-1}} \Gamma(wa_{t-1})}, \lambda_t > 0.
\end{aligned}$$

□

Prova do Teorema 5.1:

Se $t = 1$, $\lambda_0 | Y_0 \sim \text{Gama}(a_0, b_0)$ e $\mu_0 | Y_0 \sim \text{Gama}(a_0, b_0 [g(x_t, \beta)]^{-1})$ - que é verdade pela Suposição **A3**;

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

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 pela hipótese de indução e a última pelo Lema II abaixo.

Prova do Resultado (1) do Teorema 5.1:

Pelo Lema III, $\lambda_t | \mathbf{Y}_{t-1} \sim \text{Gama}(a_{t|t-1}, b_{t|t-1})$ onde $a_{t|t-1} = wa_{t-1}$ e $b_{t|t-1} = wb_{t-1}$.

Prova do Resultado (2) do Teorema 5.1:

Portanto, do Resultado (1) e fazendo uma transformação, $(\mu_t = \lambda_t g(x_t, \beta)) | \mathbf{Y}_{t-1} \sim \text{Gama}(a_{t|t-1}^*, b_{t|t-1}^*)$ onde $a_{t|t-1}^* = a_{t|t-1}$ e $b_{t|t-1}^* = b_{t|t-1} g(x_t, \beta)^{-1}$.

Prova do Resultado (3) do Teorema 5.1:

Pelo Teorema de Bayes,

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

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

Prova do Resultado (4) do Teorema 5.1:

Usando o Resultado (3) e fazendo uma transformação,

$(\lambda_t = \mu_t g(x_t, \boldsymbol{\beta})^{-1}) | \mathbf{Y}_t \sim \text{Gama}(a_t, b_t)$, em que $a_t = a_{t|t-1} + b(y_t, \boldsymbol{\varphi})$ e $b_t = b_{t|t-1} + c(y_t, \boldsymbol{\varphi}) g(x_t, \boldsymbol{\beta})$; $\forall t \in N$, $t \leq n$ onde n é o tamanho da série temporal. A hipótese indutiva é verificada.

Prova do Resultado (5) do Teorema 5.1:

$$\begin{aligned} p(y_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) &= \int_0^\infty p(y_t | \mu_t, \boldsymbol{\varphi}) p(\mu_t | \mathbf{Y}_{t-1}, \boldsymbol{\varphi}) d\mu_t \\ &= \int_0^\infty \left[a(y_t, \boldsymbol{\varphi}) \mu_t^{b(y_t, \boldsymbol{\varphi})} \exp(-\mu_t c(y_t, \boldsymbol{\varphi})) \right] \left[\frac{\mu_t^{a_{t|t-1}^* - 1} \exp(-b_{t|t-1}^* \mu_t)}{\Gamma(a_{t|t-1}^*) (b_{t|t-1}^*)^{-a_{t|t-1}^*}} \right] d\mu_t \\ &= \frac{a(y_t, \boldsymbol{\varphi})}{\Gamma(a_{t|t-1}^*) (b_{t|t-1}^*)^{-a_{t|t-1}^*}} \int_0^\infty \left[\mu_t^{b(y_t, \boldsymbol{\varphi}) + a_{t|t-1}^* - 1} \exp(-\mu_t (c(y_t, \boldsymbol{\varphi}) + b_{t|t-1}^*)) \right] d\mu_t \\ &= \frac{\Gamma(b(y_t, \boldsymbol{\varphi}) + a_{t|t-1}^*) a(y_t, \boldsymbol{\varphi}) (b_{t|t-1}^*)^{a_{t|t-1}^*}}{\Gamma(a_{t|t-1}^*) (c(y_t, \boldsymbol{\varphi}) + b_{t|t-1}^*)^{a_{t|t-1}^* + b(y_t, \boldsymbol{\varphi})}}; \quad \text{onde} \end{aligned}$$

$$a_{t|t-1}^* = w a_{t-1}, \quad b_{t|t-1}^* = w b_{t-1} g(x_t, \boldsymbol{\beta})^{-1} \text{ e } y_t \in H(\boldsymbol{\varphi}).$$

□

Apêndice 3

Prova do Teorema 5.2

Primeiro resultado: $\lambda_{t+h} | \mathbf{Y}_t \sim \text{Gama}(a_{t+h|t}, b_{t+h|t})$, onde $a_{t+h|t} = w^h a_t$ e $b_{t+h|t} = w^h b_t$.

Prova:

Se $h = 1$, pelo item 1 do Teorema 5.1, $\lambda_{t+1} | \mathbf{Y}_t \sim \text{Gama}(a_{t+1|t}, b_{t+1|t})$. Agora suponha que $\lambda_{t+h} | \mathbf{Y}_t \sim \text{Gama}(w^h a_t, w^h b_t)$ é válido. Para verificar a hipótese indutiva, é necessário demonstrar que a hipótese acima é também válida para $h + 1$.

Pela hipótese **A2**, a evolução para $t + h + 1$ dada por

$$w \frac{\lambda_{t+h+1}}{\lambda_{t+h}} | \lambda_{t+h}, \mathbf{Y}_t \sim \text{Beta} [w^{h+1} a_t, (1-w)w^h a_t].$$

A combinação da equação de evolução $t + h + 1$ com a hipótese indutiva acima leva a, depois de integrar fora em λ_{t+h} , $\lambda_{t+h+1} | \mathbf{Y}_t \sim \text{Gama}(w(w^h a_t), w(w^h b_t)) = \text{Gama}(w^{h+1} a_t, w^{h+1} b_t)$.

Portanto, a hipótese indutiva é verificada, com

$$\lambda_{t+h} | \mathbf{Y}_t \sim \text{Gama}(a_{t+h|t}, b_{t+h|t}),$$

com $a_{t+h|t} = w^h a_t$ e $b_{t+h|t} = w^h b_t$,

for all $h \in N$.

□

Segundo resultado:

$$p(y_{t+h} | \mathbf{Y}_t, \boldsymbol{\varphi}) = \frac{(b(y_{t+h}, \boldsymbol{\varphi}) + a_{t+h|t}) a(y_{t+h}, \boldsymbol{\varphi}) (b_{t+h|t})^{a_{t+h|t}}}{(a_{t+h|t}) (c(y_{t+h}, \boldsymbol{\varphi}) + b_{t+h|t})^{a_{t+h|t} + b(y_{t+h}, \boldsymbol{\varphi})}}, \text{ para } y_{t+h} \in S(\boldsymbol{\varphi}).$$

Prova:

Usando a equação das observações no tempo $t + h$ e o resultado acima,

a distribuição preditiva h passos à frente pode ser encontrado exatamente como no Resultado 5 do Teorema 5.1 com meras mudanças de notação de $(y_{t+1}, a_{t+1|t}, b_{t+1|t})$ por $(y_{t+h}, a_{t+h|t}, b_{t+h|t})$. Portanto, a função de densidade preditiva h passos à frente é dada por

$$p(y_{t+h}|\mathbf{Y}_t, \boldsymbol{\varphi}) = \frac{\Gamma(b(y_{t+h}, \boldsymbol{\varphi}) + a_{t+h|t}) a(y_{t+h}, \boldsymbol{\varphi}) (b_{t+h|t})^{a_{t+h|t}}}{\Gamma(a_{t+h|t}) (c(y_{t+h}, \boldsymbol{\varphi}) + b_{t+h|t})^{a_{t+h|t} + b(y_{t+h}, \boldsymbol{\varphi})}},$$

onde $a_{t+h|t} = w^h a_t$, $b_{t+h|t} = w^h b_t$ e $y_{t+h} \in S(\boldsymbol{\varphi})$.

□

Apêndice 4

Prova do Teorema 5.3

A requerida função de densidade é dada por

$$\begin{aligned}
p(\lambda_{t-1}|\lambda_t, \mathbf{Y}_{t-1}) &= \frac{p(\lambda_t|\lambda_{t-1}, \mathbf{Y}_{t-1})p(\lambda_{t-1}|\mathbf{Y}_{t-1})}{p(\lambda_t|\mathbf{Y}_{t-1})} \\
&= \frac{\Gamma(wa_{t-1})\Gamma((1-w)a_{t-1})}{\Gamma(a_{t-1})} \frac{w}{\lambda_{t-1}} \left(\frac{w\lambda_t}{\lambda_{t-1}}\right)^{wa_{t-1}-1} \left(1 - \frac{w\lambda_t}{\lambda_{t-1}}\right)^{(1-w)a_{t-1}-1} \times \\
&\quad \frac{b_{t-1}^{a_{t-1}}}{\Gamma(a_{t-1})} \lambda_{t-1}^{a_{t-1}-1} \exp(-\lambda_{t-1}b_{t-1}) \frac{\frac{\Gamma(wa_{t-1})}{(wb_{t-1})^{wa_{t-1}}}}{\lambda_t^{wa_{t-1}-1} \exp(-\lambda_t wb_{t-1})} \\
&\propto (\lambda_{t-1} - w\lambda_t)^{(1-w)a_{t-1}-1} \exp(-b_{t-1}(\lambda_{t-1} - w\lambda_t)).
\end{aligned}$$

Agora, defina $\eta_{t-1} = \lambda_{t-1} - w\lambda_t$ com a remoção da constante $w\lambda_t$ de λ_{t-1} . Então, $p(\eta_{t-1}|\lambda_t, \mathbf{Y}_{t-1}) \propto \eta_{t-1}^{(1-w)a_{t-1}-1} \exp(-b_{t-1}\eta_{t-1})$. É claro que $\eta_{t-1} = \lambda_{t-1} - w\lambda_t | \lambda_t, \mathbf{Y}_{t-1} \sim \text{Gama}((1-w)a_{t-1}, b_{t-1})$, completando a prova.

□

REFERÊNCIAS BIBLIOGRÁFICAS

- Alves, M. B., Gamerman, D. and Ferreira, M. A. R. (2010). Transfer functions in dynamic generalized linear models. *Statistical Modelling*, **10**, 3-40. 91
- Andrieu, C. & Doucet, A. (2002), Particle filtering for partially observed Gaussian state space models, *Journal of the Royal Statistical Society, Ser. B*, **64**, 827-836. 66
- Anderson, B. D. O. & Moore, J. B. (1979). *Optimal Filtering*, Englewood Cliffs, New Jersey, Prentice Hall. 18
- Ansley, C. F. & Kohn, R. (1986). Prediction mean squared error for state space models with estimated parameters. *Biometrika*, **73**, 467-73. iv, vi, 16, 17, 18, 20, 21, 22, 36, 37, 41, 42, 106
- Barndorff-Nielsen, O.E. & Shephard, N. (2001), Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion), *Journal of the Royal Statistical Society, Ser. B*, **63**, 167241. 95
- Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control*, 3rd ed. New Jersey: Prentice-Hall. 4
- Box, G. E. P. & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*, 2nd ed. San Francisco: Holden Day. 1, 4, 56
- Brockwell, P. J. & Davis, R. A. (1996). *Introduction to Time Series and Forecasting*, Springer text in Statistics. New York: Springer. 4, 39
- Carter C.K., Kohn R. (1994). On Gibbs sampling for state space models, *Biometrika*, 81: 541-533. 78
- Casella, G. & Berger, R. L. (2002). *Statistical Inference*, Duxbury, Pacific Grove. 11, 74

- Cavanaugh, J.E. and Shumway, R.H. (1996). On computing the expected Fisher information matrix for state-space model parameters. *Statistics and Probability Letters*, Vol. **26**, p. 347-355. 74
- Chang, Y. and Liu, C. (2009). A generalized JM model with applications to imperfect debugging in software reliability. *Applied Mathematical Modelling*, **33**, 3578-3588. 96, 97
- Chang, Y. and Leu, L. (1998). A state space model for software reliability. *Annals of the Institute of Statistical and Mathematics*, **50**, 789-799.
- Chen, Y. and Singpurwalla, N. D. (1994). A non-gaussian Kalman filter model for tracking software reliability. *Statistica Sinica*, **4**, 535-548. 96, 97
- Chiogna, M. and Gaetan, C. (2002). Dynamic generalized linear models with application to environmental epidemiology. *Applied Statistics*, **51**, 453-468. 66
- Cox, D. R. (1981). Statistical analysis of time-series: some recent developments. *Scandinavian Journal of Statistics*, **8**, 93-115. 66
- de Jong, P. (1988). A cross-validation filter for time series models. *Biometrika*, **75**, 594-600. 15, 18
- de Jong, P. (1989). Smoothing and interpolation with the state space model. *Journal of the American Statistical Association*, **84**, 1085-1088. 15, 18
- Dethlefsen, C. & Lundbye-Christensen, S. (2006). Formulating state space models in R with focus on longitudinal regression models. *J. Statist. Software* **16**, 1-15. 91
- Dobson, A. J. (2002). *An Introduction to Generalized Linear Models*. 2nd edition, Chapman and Hall, New York. 65
- Doornik, J. A. (2002). *Ox: An Object-Oriented Matrix Language*. 3. ed. London: Timberlake Consultants Press. 27, 30, 46, 81
- 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. 66, 69
- Durbin, J. & Koopman, S. J. (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press. 8, 91

- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, **7**, 1-26. 12
- Efron, B. & Tibshirani, R. (1993). *An Introduction to the Bootstrap*. London: Chapman and Hall. 12, 13
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U. K. inflation. *Econometrica*, **50**, 987-1008. 4
- Fahrmeir, L. & Wagenpfeil, S. (1997). Penalized likelihood estimation and iterative Kalman smoothing for non-gaussian dynamic regression models. *Computational Statistics and Data Analysis*, **24**, 295-320. 68
- Fahrmeir, L. (1987). Regression models for nonstationary categorical time series. *Journal of Time Series Analysis*, **8**, 147-160. 66
- Ferris, B. G., Dockery, D. W., Ware, J. H., Speizer, F.E. and Spiro, R. (1983). The Six-City Study: examples of problems in analysis of the data. *Environmental Health Perspectives*, **52**, 115-123. 91
- 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. 108
- 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. 12, 21, 23, 42
- Fruhworth-Schnatter, S. (1994). Applied state space modelling of non-Gaussian time series using integration-based Kalman filtering. *Statistical Computation*, **4**, 259-269. 66, 78
- Fruhworth-Schnatter, S. and Wagner, H. (2006). Auxiliary mixture sampling for parameterdriven models of time series of counts with applications to state space modelling, *Biometrika*, **93**, 827-841. 66
- Gamerman, D. & West, M. (1987). An application of dynamic survival models in unemployment studies. *Statistician*, **36**, 269-274. 66
- Gamerman, D. (1991). Dynamic Bayesian models for survival data. *Applied Statistics*, **40**, 63-79. 66

- Gamerman, D. (1994). Bayes estimation of the piece-wise exponential distribution. *IEEE Transactions on Reliability*, **43**, 128-131. 107
- Gamerman, D. (1998). Markov chain Monte Carlo for dynamic generalised linear models. *Biometrika*, **85**, 215-227. 66, 68
- Gamerman, D. & Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. London: Chapman and Hall. 75, 76
- 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. 82, 92, 94
- Gilks, W.R., Richardson, S. and Spiegelhalter, D. (1996). *Markov Chain Monte Carlo in practice*, Chapman and Hall, London. 76
- Godolphin, E. J. & Triantafyllopoulos, K. (2006). Decomposition of time series models in state-space form. *Computational Statistics and Data Analysis*, **50**, 2232-2246. 66
- Goel, A. L. and Okumoto, K. (2006). Time-dependent error detection rate model for software reliability and other performance measures. *IEEE Trans. on Reliability*, **28**, 206-211. 96
- 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. 4
- 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. 66
- Haight, F. A. and Breuer, M. A. (1960). The Borel-Tanner Distribution. *Biometrika*, **47**, 143-150. 70
- Hamilton, J. D. (1986). A standard error for the estimated state vector of a state space model. *Journal of Econometrics*, **33**, 387-97. iv, vi, 16, 22, 36, 37, 39, 40, 106
- Harrison, P. J. & Stevens, C. F. (1971). A Bayesian approach to short-term forecasting. *Operational Research Quarterly*, **22**, 341-362. 1
- Harrison, P. J. & Stevens, C. F. (1976). Bayesian forecasting. *Journal of the Royal Statistical Society: series B*, **38**, 205-247. 1

- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: University Press. 1, 5, 6, 7, 8, 10, 12, 15, 16, 17, 18, 19, 36, 38, 41, 73, 76, 78, 81, 106
- Harvey, A. C. (2000). Comments of "Time Series analysis of non-gaussian observations based on state space models from both classical and Bayesian perspectives". *Journal of the Royal Statistical Society, Ser. B*, **62**, 34-35. 16
- 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. 66, 69, 78, 80
- Hemming, K. & Shaw, J. E. H. (2002). A parametric dynamic survival model applied to breast cancer survival times. *Applied Statistics*, **51**, 421-435. 66
- Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted moving averages. ONR Research Memorandum 52, Carnegie Institute of Technology, Pittsburgh, Pennsylvania. 1
- Jelinski, Z. and Moranda, P. B. (1972), *in: W. freiberger (Ed.), Software Reliability Research*. In Statistical Computer Performance Evaluation, Academic Press, New York, 465-497. 96
- Jørgensen, B., Lundbye-Christensen, S., Song, P.X.K. and Sun, L. (1999). A state space models for multivariate longitudinal count data. *Biometrika*, **86**, 169-181. 66
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.*, **82**, 35-45. 9
- Kaufmann, H. (1987). Regression models for nonstationary categorical time series: asymptotic estimation theory. *The Annals of Statistics*, **15**, 79-98. 66
- Kitagawa, G. (1987). Non-Gaussian state-space modelling of nonstationary time series. *Journal of the American Statistical Association*, **82**, 1032-1063. 66
- Knorr-Held, L. (1999). Conditional prior proposal in dynamic models. *Scandinavian Journal of Statistics*, **26**, 129-144. 68
- Koopman, S. J. (1992). *Diagnostic checking and intra-daily effects in time series models*. Amsterdam: Thesis Publishers. 18

- Koopman, S. J., Shephard, N. & Doornik, J. A. (1999). Statistical algorithms for inference in state space using SsfPack 2.2 *Econometrics Journal*, **47**, 129-139.
- Lindsey, J. K. & Lambert, P. (1995). Dynamic generalized linear models and repeated measurements. *Journal of Statistical Planning and Inference*, **47**, 129-139. 66
- Luz, P. M., Mendes, V. M., Codeço, C. T., Struchiner, C.J. & Galvani, A.P. (2008). Time series analysis of Dengue Incidence in Rio de Janeiro, Brazil. *American Journal of Tropical Medicine and Hygiene*, **79**, 933-939. 56
- Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association*, **55**, 299-305. 1
- Migon, H. S. & Gamerman, D. (1999). *Statistical Inference: an Integrated Approach*. London: Arnold. 11, 74, 76
- Migon H.S, Gamerman D., Lopes H. F., Ferreira M. A. R. (2005). *Dynamic Models*. In: Editors: C.R. Rao, Dipak K. Dey, Elsevier B. V. (Eds.). *Handbook of Statistics*, 25, p. 553-588. 12, 78
- Morettin, P. A. (2006) *Um curso em Séries Temporais Financeiras*. 17º SINAPE, Caxambu-MG, Associação Brasileira de Estatística. 93, 95
- Morettin, P. A. & Toloí, C. M. (2004). *Análise de Séries Temporais*. São Paulo: Editora E. Blucher-Associação Brasileira de Estatística. 4
- McCullagh, P. & Nelder. J.A. (1989). *Generalized Linear Models*. Chapman and Hall: London. 80
- Nelder. J.A. & Wedderburn, R.W.M. (1972). Generalized linear models. *Journal of the Royal Statistical Society A*, 135, 370-384. 65
- Nervole, M. & Wage, S. (1964). On the optimality of adaptive forecasting. *Management Science*, **10**, 207-229. 1
- Pfeffermann, D. & Tiller, R. (2005). Bootstrap approximation to prediction MSE for state-space models with estimated parameters. *Journal of Time Series Analysis*, **26**, 893-916. 2, 16, 17, 21, 22, 23, 25, 27, 28, 29, 37, 39, 40, 43, 44, 106
- Pitt, M. K. and Walker, S. G. (2005). Constructing stationary time series models using auxiliary variables with applications, *Journal of the American Statistical Association*, **100**, 554-564. 66

- 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/>. 82
- Proietti, T. (2000). Comparing seasonal components for structural time series models, *International Journal of Forecasting*, **16**: 247-260. 6, 7
- Queenville, B. & Singh, A. C. (2000). Bayesian prediction mean squared error for state space models with estimated parameters. *Journal of Time Series Analysis*, **21**, 2, 219-236. 2, 16, 36
- R Development Core Team. (2005). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna. 82
- Rodriguez, A. & Ruiz, E. (2010a) Bootstrap Prediction Intervals in State Space Models. *Journal of the Time Series Analysis*, **30**, 167-178. 2, 36, 39, 45
- Rodriguez, A. & Ruiz, E. (2010b) Bootstrap prediction mean squared errors of unobserved states based on Kalman filter with estimated parameters. Technical report. Department of Statistics, Carlos III de Madrid University. 15, 16, 17, 20
- Sallas, W.M. and Harville, D.A. (1988). *Noninformative Priors and Restricted Maximum Likelihood Estimation in the Kalman Filter*. In: James C. Spall. *Bayesian Analysis of Time Series and Dynamic Models*. New York, Marcel Dekker Inc. 74
- Santos, T. R. & Franco, G. C. (2010). Inference for the hyperparameters of structural models under classical and Bayesian perspectives: a comparison study. *Communications in Statistics: Simulation and Computation*, **39**: 1671-1689.
- Santos, T. R. (2009). Inferência sobre os hiperparâmetros dos modelos estruturais sob as perspectivas clássica e Bayesiana. Dissertação de mestrado, DEST, UFMG, Belo Horizonte. 1, 12, 14, 38, 75, 76
- Santos, T.R., Franco, G.C. & Gamerman, D. (2010). Comparison of classical and Bayesian approaches for intervention analysis in structural models. *International Statistical Review*, **78**, 2, 218-239. 2
- Sen, P. K. & Singer, J. M. (1993). *Large Sample Methods in Statistics: An introduction with applications*. Chapman-Hall, London. 109, 110

- Shephard, N. (1993) Maximum likelihood estimation of regression models with stochastic trend components. *Journal of the American Statistical Association*, **88**, 590-595. 36
- Shephard, N. (1994) Local scale models: state space alternative to integrated GARCH processes. *Journal of Econometrics*, **60**, 181-202. 69
- Shephard, N. & Pitt, M. K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika*, **84**, 653-667. 66, 68
- Silva, R.S., Giordani, P., Konh, R. and Pitt, M.K.(2009). Particle filtering within adaptive Metropolis-Hastings sampling. Technical report. School of Economics, University of New South Wales, Australia. 67
- Singh, A. C. and Roberts, G. R. (1992). State space modelling of cross-classified time series of counts. *International Statistical Review*, **60**, 321-36. 68
- Smith, J. Q. (1979). A Generalization of the Bayesian Steady Forecasting Model. *Journal of the Royal Statistical Society, Ser. B*, **41**, 375-387. 66
- Smith, J. Q. (1981). The Multiparameter Steady Model. *Journal of the Royal Statistical Society, Ser. B*, **43**, 256-260. 66
- 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. 3, 66, 67, 69
- Spall, J. C. (1988). *Bayesian Analysis of Time Series and Dynamic Models*. New York, Marcel Dekker Inc.
- Spiegelhalter, D. J., Best, N. G., Carlin, B.P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion and rejoinder). *Journal of the Royal Statistical Society, Ser. B*, **64**, 583-639. 81
- 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. 2, 13, 23, 41, 44, 106, 109
- Stoffer, D. S. & Wall, K. D. (2004). *Resampling in State Space Models*, in *State Space and Unobserved Components Models: Theory and Applications*, A. C. Harvey, S.J. Koopman and N. Shephard , Eds., Cambridge University Press. 106

- Theil, H. & Wage, S.(1964). Some observations on adaptative forecasting. *Management Science*, **10**, 198-206. 1
- Tsimikas, J. & Ledolter, J. (1994). REML and best linear unbiased prediction in state space models. *Communications in Statistics: Theory and Methods*, **23** (8), 2253-2268. 36
- Tsimikas, J. & Ledolter, J. (1997). Mixed model representation of state space models: new smoothing results and their application to REML estimation. *Statistica Sinica*, **7**, 973-991. 15
- Wall, K. D. & Stoffer, D. S. (2002) A state space approach to bootstrapping conditional forecasts in ARMA models. *Jr. Time Series Analysis*, **23**, 733-751. 2, 36, 45
- Watanabe, N. (1985) Note on the Kalman filter with estimated parameters. *Jr. Time Series Analysis*, **6**, 269-278. 111
- Wei, W. W. S. (1990). *Time Series Analysis - Univariate and Multivariate Methods*. New York: Addison-Wesley. 4, 7
- 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. 2, 66, 67, 68, 108
- West, M. & Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. New York: Springer. 1, 2, 5, 6, 7, 8, 72, 108
- Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, **6**, 324-342. 1