



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Projeto para Qualificação em Mestrado

Um Estudo sobre Modelos para Volatilidade
Estocástica

por

André Silva de Queiroz

Orientadora: Prof.^a Cibele Queiroz da Silva, Ph.D.

Junho de 2015

Sumário

Introdução	2
1 Modelo Linear Dinâmico	3
1.1 Definição do Modelo	3
1.2 Breve Resumo do Teorema de Bayes	5
1.3 Equações de Atualização	6
1.4 Previsões	7
2 Modelo de Volatilidade Estocástica	8
2.1 Definição do Modelo Canônico	9
2.2 Um Modelo de Espaço-Estado	9
2.2.1 Equação das Observações	9
2.2.2 Equação do Sistema	11
2.2.3 Modelo Completo	11
2.3 Definição do Modelo Bayesiano	12
2.4 Estudo de Simulação	13
2.4.1 Amostrador de Gibbs via JAGS	14
2.5 Um Exemplo com Dados Reais	20
2.5.1 Modelo I	20
2.5.2 Modelo II	21
2.6 Estimação dos parâmetros	21
Considerações Preliminares	23
A Códigos-Fonte	27

Introdução

Modelos de dados temporais são de extrema importância para a sociedade. Seja para entender um processo natural ou um fenômeno econômico, sua aplicação adequada é fundamental para gerar boas estimativas do comportamento do sistema. Afinal, predição de valores futuros é um dos principais objetivos de um modelo desse tipo, Brockwell e Davis (2002). Eventualmente, o conjunto de dados de interesse pode apresentar uma variância distinta ao longo do tempo. Os modelos ARCH e GARCH são a solução tradicional para esse tipo de situação, porém tratam a variância de forma determinística. Em um panorama mais geral, a variância ou volatilidade dos dados pode ser considerada de natureza estocástica, assim surgem os modelos de volatilidade estocástica, que são o objeto principal de estudo nesse trabalho.

A inferência Bayesiana tem se mostrado uma excelente ferramenta para análise de dados. Como cita Pearl (2009), ela se conecta perfeitamente com a maneira humana de raciocínio diante da incerteza. Pensar em problemas sob essa perspectiva num contexto de séries temporais é possível através dos modelos dinâmicos.

Nesse início de trabalho será introduzida uma pequena apresentação dos modelos lineares dinâmicos com seus conceitos básicos, como em West e Harrison (1997). Então o modelo de volatilidade estocástica será definido no capítulo seguinte, e como ele pode ser visto também conforme um modelo de espaço-estado, assim como os modelos lineares dinâmicos. Em sequência, um conjunto de dados simulados será utilizado para testar um algoritmo de estimação dos seus parâmetros baseado no amostrador de Gibbs. Também será apresentada a replicação dos resultados de Achcar et al. (2011), que foi feita como um exercício inicial para entendimento do problema. Essa etapa se encerra com algumas considerações sobre a direção a ser tomada no desenvolvimento da futura dissertação.

Capítulo 1

Modelo Linear Dinâmico

O uso de modelos lineares dinâmicos (MLD) cresceu bastante nas últimas décadas. Sua aplicabilidade tem se estendido nos campos da biologia, genética, geofísica, economia entre outros, (Petrís et al., 2009). O desenvolvimento computacional das técnicas Bayesianas de estimação dos parâmetros é um dos fatores que impulsionou esse crescimento recente.

Eles constituem uma família muito importante, pois, além de trazerem os modelos tradicionais de séries de tempo para a abordagem Bayesiana, expandem a gama de possibilidades de aplicações através da sua estrutura flexível e do próprio paradigma Bayesiano em si. São, ainda, um caso particular linear e gaussiano dos modelos mais gerais de espaço-estado.

1.1 Definição do Modelo

O modelo linear dinâmico mais geral é definido em relação a um vetor de observações \mathbf{Y}_t , porém, para efeitos de simplificação, será considerado o caso univariado Y_t como em West e Harrison (1997). Portanto, sejam Y_t uma variável observada de interesse, $\boldsymbol{\theta}_t$ um vetor latente de variáveis que representam o sistema gerador de Y_t e D_t toda informação disponível a respeito do sistema. Desse modo, para todo tempo t , o modelo linear dinâmico é:

$$\text{Equação das Observações:} \quad Y_t = \mathbf{F}_t' \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim \mathcal{N}(0, V_t), \quad (1.1)$$

$$\text{Equação do Sistema:} \quad \boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t). \quad (1.2)$$

A quádrupla $\{\mathbf{F}_t, \mathbf{G}_t, V_t, \mathbf{W}_t\}$ é composta pelas matrizes que caracterizam o modelo. \mathbf{F}_t é a matriz de *design*, ou o vetor de regressão, no caso univariado, \mathbf{G}_t é a matriz de evolução do sistema, V_t é a variância observacional e \mathbf{W}_t é a matriz de variância da evolução do sistema. Se esses elementos forem invariantes no tempo, o modelo $\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}$ é chamado de constante, e engloba essencialmente todos os modelos lineares tradicionais

de séries temporais.

Os termos ν_t e ω_t são os erros associados às equações (1.1) e (1.2). Por definição, eles são independentes, no tempo e entre si, e apresentam distribuição normal com média zero. Modelos ainda mais gerais podem ser definidos com ν_t e ω_t autocorrelacionados e correlacionados entre si, West e Harrison (1997), entretanto, destaca que tais modelos podem sempre ser reescritos em termos mais simples satisfazendo as condições de independência. Eles são denominados erro observacional e erro do sistema, respectivamente.

A informação inicial sobre θ_t é representada por:

$$\text{Informação Inicial:} \quad (\theta_0|D_0) \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0), \quad (1.3)$$

onde \mathbf{m}_0 e \mathbf{C}_0 são o vetor de médias e a matriz de variâncias e covariâncias da distribuição proposta, respectivamente.

A Figura 1.1 mostra um exemplo de 100 observações simuladas a partir de um modelo linear dinâmico constante definido pela quádrupla:

$$\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\} = \left\{ \begin{bmatrix} 1 \\ \pi \end{bmatrix}, \begin{bmatrix} 0,15 & -0,08 \\ -0,08 & 0,15 \end{bmatrix}, 1, \begin{bmatrix} 1 & 0,3 \\ 0,3 & 1 \end{bmatrix} \right\}. \quad (1.4)$$

A série partiu da distribuição inicial:

$$(\theta_0|D_0) \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}\right). \quad (1.5)$$

Os valores dos parâmetros nesse exemplo foram selecionados arbitrariamente, apenas

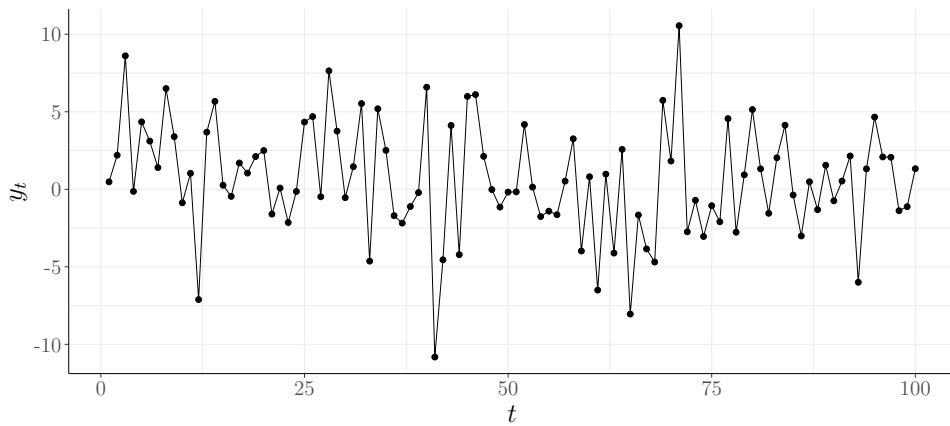


Figura 1.1: Exemplo simulado de 100 observações geradas a partir de um modelo polinomial de primeira ordem.

tomando-se o cuidado para que não fossem muito grandes de modo a gerar observações numa escala elevada.

Dentre as finalidades de um modelo de dados temporais está a previsão de observa-

ções futuras. Sob esse aspecto, o objetivo a ser alcançado agora é definir as distribuições de probabilidade de $(Y_t|D_{t-1})$ e $(\boldsymbol{\theta}_t|D_t)$. A primeira é a distribuição da previsão de Y no tempo t dado todo o conhecimento anterior a esse tempo. A outra é a distribuição do vetor $\boldsymbol{\theta}$ no tempo t dado todo conhecimento disponível até então. Antes da definição dessas distribuições, deve ser apresentado um breve resumo por trás da ideia do teorema de Bayes que é fundamental no desenvolvimento dos próximos passos.

1.2 Breve Resumo do Teorema de Bayes

Sejam dois eventos distintos A e B , a probabilidade de ocorrência conjunta deles é dada pela regra do produto:

$$P(A,B) = P(A|B)P(B). \quad (1.6)$$

Naturalmente, a ordem dos eventos pode ser mudada, isto é, $P(B,A) = P(B|A)P(A)$. A partir dessas duas equações surge a relação:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \quad (1.7)$$

O resultado em (1.7) tem sua origem nas ideias do Reverendo Thomas Bayes no século XVIII, porém Pierre Simon de Laplace é o autor dessa equação conhecida como teorema de Bayes, Sivia e Skilling (2006). No contexto dos modelos lineares dinâmicos,

$$f(\boldsymbol{\theta}_t|y_t, D_{t-1}) = \frac{g(y_t|\boldsymbol{\theta}_t, D_{t-1})\pi(\boldsymbol{\theta}_t|D_{t-1})}{h(y_t|D_{t-1})}. \quad (1.8)$$

O processo latente definido por $\boldsymbol{\theta}_t$ é estimado através da realização y_t de Y e do conjunto de informações relevantes, D_{t-1} . Porém, a incerteza (ou certeza) inicial existente sobre $\boldsymbol{\theta}_t$ deve ser expressa através de uma distribuição de probabilidade adequada (Jaynes, 2003), no caso, $\pi(\boldsymbol{\theta}_t|D_{t-1})$. Ela é chamada de distribuição *a priori* de $(\boldsymbol{\theta}_t|D_{t-1})$.

Conforme surjam novas realizações dos dados, o conhecimento sobre o processo latente deve ser atualizado. Isso é feito pela função $g(y_t|\boldsymbol{\theta}_t, D_{t-1})$, denominada de verossimilhança.

A função $h(y_t|D_{t-1})$ é chamada de distribuição preditiva *a priori* de y_t . Ela é muito útil quando é desejável fazer inferências a respeito de uma observação ainda desconhecida, (Gelman et al., 2014). No caso, quando for o interesse estimar a , até então desconhecida, observação y_t de Y .

Por fim, $f(\boldsymbol{\theta}_t|y_t, D_{t-1})$ é a distribuição resultante da interação das três anteriores. Ela permite fazer inferências sobre o processo latente e recebe o nome de distribuição *a posteriori* de $(\boldsymbol{\theta}_t|D_t)$, uma vez que o conhecimento de y_t e D_{t-1} dá origem a D_t .

Recursivamente, a distribuição *a posteriori* de $(\boldsymbol{\theta}_t|D_t)$ pode se tornar a distribuição *a priori* de $(\boldsymbol{\theta}_{t+1}|D_t)$. Assim o teorema de Bayes flui naturalmente no contexto dos modelos lineares dinâmicos. Como pode ser visto através das equações de atualização.

1.3 Equações de Atualização

Uma das motivações ao se propor um modelo de dados temporais é fazer previsões futuras. O modelo linear dinâmico oferece um conjunto de equações que são atualizadas ao longo do tempo, e que permitem estimar essas informações. O raciocínio a seguir segue o proposto em West e Harrison (1997). Apenas para simplificar a álgebra futura, será considerado o modelo linear dinâmico constante.

Seja, para algum \mathbf{m}_t e \mathbf{C}_t , a *a posteriori* de $\boldsymbol{\theta}_t$:

$$(\boldsymbol{\theta}_t|D_t) \sim \mathcal{N}(\mathbf{m}_t, \mathbf{C}_t). \quad (1.9)$$

A equação (1.2) permite, então, calcular a distribuição *a priori* de $\boldsymbol{\theta}_{t+1}$ dada a informação em t que é:

$$\begin{aligned} \boldsymbol{\theta}_{t+1} &= \mathbf{G}\boldsymbol{\theta}_t + \boldsymbol{\omega}_{t+1}, \\ (\boldsymbol{\theta}_{t+1}|D_t) &\sim \mathbf{G}\mathcal{N}_{(\boldsymbol{\theta}_t|D_t)}(\mathbf{m}_t, \mathbf{C}_t) + \mathcal{N}_{\boldsymbol{\omega}_{t+1}}(\mathbf{0}, \mathbf{W}), \\ (\boldsymbol{\theta}_{t+1}|D_t) &\sim \mathcal{N}(\mathbf{a}_{t+1}, \mathbf{R}_{t+1}), \end{aligned} \quad (1.10)$$

onde $\mathbf{a}_{t+1} = \mathbf{G}\mathbf{m}_t$ e $\mathbf{R}_{t+1} = \mathbf{G}\mathbf{C}_t\mathbf{G}' + \mathbf{W}$.

Semelhantemente, a equação (1.1) é utilizada para calcular a distribuição da previsão da observação de Y_{t+1} dada a informação em t , ou a primeira previsão:

$$\begin{aligned} Y_{t+1} &= \mathbf{F}'\boldsymbol{\theta}_{t+1} + \nu_{t+1}, \\ (Y_{t+1}|D_t) &\sim \mathbf{F}'\mathcal{N}_{(\boldsymbol{\theta}_{t+1}|D_t)}(\mathbf{a}_{t+1}, \mathbf{R}_{t+1}) + \mathcal{N}_{\nu_{t+1}}(0, V), \\ (Y_{t+1}|D_t) &\sim \mathcal{N}(f_{t+1}, Q_{t+1}), \end{aligned} \quad (1.11)$$

onde $f_{t+1} = \mathbf{F}'\mathbf{a}_{t+1}$ e $Q_{t+1} = \mathbf{F}'\mathbf{R}_{t+1}\mathbf{F} + V$.

A distribuição *a posteriori* de $\boldsymbol{\theta}_{t+1}$ dada a informação em $t + 1$, que será *a priori* na próxima iteração, será:

$$(\boldsymbol{\theta}_{t+1}|D_{t+1}) \sim \mathcal{N}(\mathbf{m}_{t+1}, \mathbf{C}_{t+1}), \quad (1.12)$$

onde, $\mathbf{m}_{t+1} = \mathbf{a}_{t+1} + \mathbf{A}_{t+1}e_{t+1}$ e $\mathbf{C}_{t+1} = \mathbf{R}_{t+1} - \mathbf{A}_{t+1}Q_{t+1}\mathbf{A}_{t+1}'$, com $\mathbf{A}_{t+1} = \mathbf{R}_{t+1}\mathbf{F}Q_{t+1}^{-1}$ e $e_{t+1} = Y_{t+1} - f_{t+1}$.

1.4 Previsões

A previsão para o primeiro passo a frente é dada pela distribuição (1.11) descrita anteriormente. Para definir o k -ésimo valor predito, é necessário antes encontrar a distribuição de $\boldsymbol{\theta}_{t+k}$. Isso é feito a partir da distribuição *a priori* (1.10) aplicada na equação do sistema (1.2) sucessivamente. Assim, para D_t que é ainda a última informação disponível e para $k = 2$:

$$\begin{aligned}\boldsymbol{\theta}_{t+2} &= \mathbf{G}\boldsymbol{\theta}_{t+1} + \boldsymbol{\omega}_{t+2}, \\ (\boldsymbol{\theta}_{t+2}|D_t) &\sim \mathbf{GN}_{(\boldsymbol{\theta}_{t+1}|D_t)}(\mathbf{a}_{t+1}, \mathbf{R}_{t+1}) + \mathcal{N}_{\boldsymbol{\omega}_{t+2}}(0, \mathbf{W}), \\ (\boldsymbol{\theta}_{t+2}|D_t) &\sim \mathcal{N}(\mathbf{a}_{t+2}, \mathbf{R}_{t+2}),\end{aligned}\tag{1.13}$$

onde $\mathbf{a}_{t+2} = \mathbf{G}\mathbf{a}_{t+1}$ e $\mathbf{R}_{t+2} = \mathbf{G}\mathbf{R}_{t+1}\mathbf{G}' + \mathbf{W}$.

Para um valor de $k \geq 2$, Pole et al. (1994) mostra que:

$$(\boldsymbol{\theta}_{t+k}|D_t) \sim \mathcal{N}(\mathbf{a}_{t+k}, \mathbf{R}_{t+k}),\tag{1.14}$$

onde $\mathbf{a}_{t+k} = \mathbf{G}^{k-1}\mathbf{a}_{t+1}$ e $\mathbf{R}_{t+k} = \mathbf{G}^{k-1}\mathbf{R}_{t+1}(\mathbf{G}^{k-1})' + \sum_{j=2}^k \mathbf{G}^{k-j}\mathbf{W}(\mathbf{G}^{k-j})'$. E, com isso, a distribuição da k -ésima predição dada a informação D_t será:

$$(Y_{t+k}|D_t) \sim \mathcal{N}(f_{t+k}, Q_{t+k}),\tag{1.15}$$

onde $f_{t+k} = \mathbf{F}'\mathbf{a}_{t+k}$ e $Q_{t+k} = \mathbf{F}'\mathbf{R}_{t+k}\mathbf{F} + V$.

Capítulo 2

Modelo de Volatilidade Estocástica

O modelo de volatilidade estocástica (MVE) surgiu como uma alternativa aos tradicionais modelos ARCH e GARCH, (Engle (1982); Bollerslev (1986)). Nesse novo modelo, a variabilidade dos dados deixa de ter um padrão determinístico e passa também a ser interpretada como uma variação aleatória. Entretanto, Bos (2012) cita que há uma grande desvantagem dos modelos de volatilidade estocástica em relação aos modelos ARCH e GARCH em termos de aplicações. Isso ocorre pois os modelos tradicionais possuem muitas variações, porém basicamente apenas uma maneira de se estimar os parâmetros, e que está presente em grande parte dos *softwares* com essa finalidade. Já os modelos de volatilidade estocástica apresentam poucas variações, porém diversos autores sugerem maneiras distintas de se estimar os parâmetros, e quase nenhuma delas está disponível facilmente através de algum pacote computacional. A Tabela 2.1 foi extraída do texto supracitado e apresenta de forma sintetizada as principais referências sobre os diversos métodos de estimação já publicados.

Tabela 2.1: Métodos de estimação dos parâmetros do MVE e principais referências, (Bos, 2012).

Método	Referência	Paradigma
Quasi-Maximum Likelihood (QML)	Harvey et al. (1994)	Clássico
Gaussian Mixture Sampling (GMS)	Kim et al. (1998)	Bayesiano
Simulated Method of Moments (SMM)	Gallant e Tauchen (1996)	Clássico
Importance Sampling (IS)	Durbin e Koopman (1997)	Clássico
Efficient Importance Sampling (EIS)	Richard e Zhang (2007)	Clássico
Improved Importance Sampling (IIS)	Nguyen (2007)	Clássico
Single Site Sampler (SSS)	Carter e Kohn (1994)	Bayesiano
MultiMove Sampler (MMS)	Shephard e Pitt (1997)	Bayesiano

2.1 Definição do Modelo Canônico

Essa modelagem foi inicialmente proposta por Taylor (1982). Ele sugere que o logaritmo do quadrado da volatilidade ao longo do tempo deve ser modelado por um processo latente auto-regressivo de primeira ordem. Kim et al. (1998) apresenta o modelo canônico para dados igualmente espaçados.

O modelo é basicamente composto por uma equação que descreve a variável observada Y_t em função de uma variável latente h_t , cuja função é captar a variabilidade dos dados no decorrer do tempo. Essa variável latente é modelada a partir de um modelo auto-regressivo de primeira ordem, AR(1), o que define a segunda parte do modelo de volatilidade estocástica. Assim, para $t = 1, \dots, N$:

$$\text{Modelo Canônico : } \begin{cases} Y_t &= e^{\frac{h_t}{2}} \delta_t, \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\ h_0 &\sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1-\phi^2}\right). \end{cases} \quad (2.1)$$

Os termos δ_t e η_t são ruídos gaussianos, independentes no tempo e entre si, com distribuições:

$$\delta_t \sim \mathcal{N}(0, 1) \quad \text{e} \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \quad (2.2)$$

Observe que, $Y_t \sim \mathcal{N}(0, e^{h_t})$, ou como requerido pelo modelo, a variância das observações é totalmente descrita no tempo t em função da variável latente h_t . O conjunto de parâmetros do modelo a serem estimados é definido pelo vetor:

$$\boldsymbol{\psi} = (\mu, \phi, \sigma_\eta^2). \quad (2.3)$$

2.2 Um Modelo de Espaço-Estado

2.2.1 Equação das Observações

Os modelos de volatilidade estocástica podem ser reescritos na notação de espaço-estado proposta por West e Harrison (1997). Uma das vantagens disso é poder utilizar o aparato ferramental e metodológico já implementado computacionalmente (Petrís et al. (2009); Petris (2010)). Isso é possível a partir da extensão da proposta em Zivot e Yollin (2012). Pela definição em (2.1):

$$\begin{aligned} Y_t &= e^{\frac{h_t}{2}} \delta_t, \\ \ln Y_t^2 &= h_t + \ln \delta_t^2, \end{aligned} \quad (2.4)$$

O valor esperado da distribuição de $\ln \delta_t^2$ pode ser separado no lado direito da equação (2.4) de modo que $\varepsilon_t \sim (0, \text{Var}(\ln \delta_t^2))$ seja um ruído em torno de zero. Assim,

$$\ln Y_t^2 = h_t + E(\ln \delta_t^2) + \varepsilon_t. \quad (2.5)$$

Apesar dessa transformação linearizar o problema, ela eliminou o ruído gaussiano δ_t . Durbin e Koopman (2012) afirmam que isso não impede a utilização das técnicas associadas aos modelos gaussianos, uma vez que se assuma a distribuição de ε_t aproximadamente normal com média zero. Entretanto, Omori et al. (2007) sugere uma mistura de $r = 10$ normais que se aproxima da distribuição do erro na equação (2.4), $\ln \delta_t^2$, que também será utilizada nesse estudo:

$$f(\ln \delta_t^2) = \xi_t = \sum_{j=1}^{10} w_j \mathcal{N}(m_j, s_j^2). \quad (2.6)$$

O peso, a média e a variância, w_j , m_j e s_j^2 , respectivamente, das 10 distribuições sugeridas por Omori et al. (2007) são detalhados na Tabela 2.2.

Tabela 2.2: Parâmetros da distribuição de ξ_t em (2.6), (Omori et al., 2007).

j	1	2	3	4	5	6	7	8	9	10
w_j	0,006	0,048	0,131	0,207	0,227	0,188	0,120	0,056	0,016	0,001
m_j	1,927	1,347	0,735	0,023	-0,852	-1,973	-3,468	-5,552	-8,684	-14,650
s_j^2	0,113	0,178	0,268	0,406	0,627	0,986	1,575	2,545	4,166	7,333

A Figura 2.1 mostra a comparação entre as distribuições de $\ln \delta_t^2$, que é o verdadeiro erro do modelo, de ε_t , que corresponde à aproximação ingênua, e de ξ_t , que é a mistura de 10 normais sugerida.

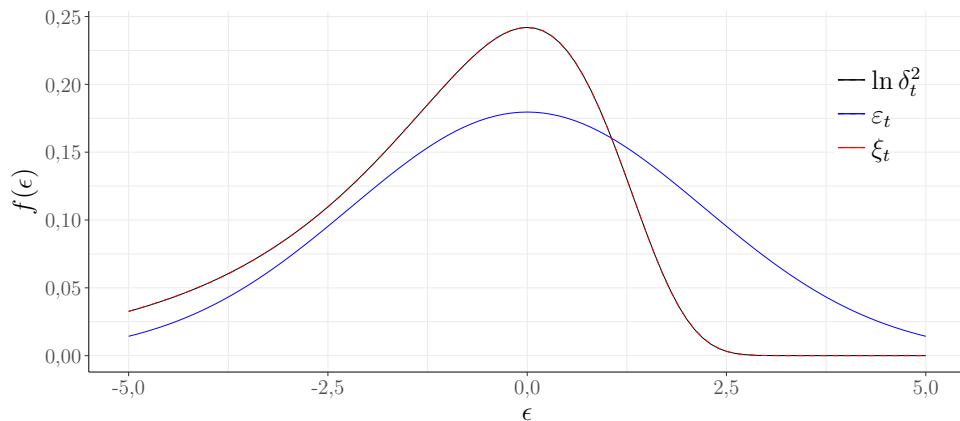


Figura 2.1: Comparação das distribuições de $\ln \delta_t^2$, ε_t e ξ_t .

2.2.2 Equação do Sistema

A equação que define a evolução da variável latente, h_t , em (2.1) pode ser reescrita:

$$\begin{aligned} h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\ h_t &= \mu - \phi\mu + \phi h_{t-1} + \eta_t, \\ h_t &= (1 - \phi)\mu + \phi h_{t-1} + \eta_t. \end{aligned} \tag{2.7}$$

Como essa transformação é um mero rearranjo dos termos, a estrutura gaussiana dos erros em η_t não foi perdida, como no caso da equação das observações.

2.2.3 Modelo Completo

Utilizando a notação de espaço-estado para os modelos lineares dinâmicos em West e Harrison (1997), descrita no capítulo 1, a equação das observações (2.4) será:

$$\begin{aligned} Y_t &= \mathbf{F}_t' \boldsymbol{\theta}_t + \nu_t, \\ \ln Y_t^2 &= h_t + E(\ln \delta_t^2) + \varepsilon_t, \\ \ln Y_t^2 &= \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_t \\ \mu \\ E(\ln \delta_t^2) \end{bmatrix} + \varepsilon_t. \end{aligned} \tag{2.8}$$

Já a equação do sistema (2.7) será:

$$\begin{aligned} \boldsymbol{\theta}_t &= \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \\ h_t &= \phi h_{t-1} + (1 - \phi)\mu + \eta_t, \\ \begin{bmatrix} h_t \\ \mu \\ E(\ln \delta_t^2) \end{bmatrix} &= \begin{bmatrix} \phi & 1 - \phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_{t-1} \\ \mu \\ E(\ln \delta_t^2) \end{bmatrix} + \begin{bmatrix} \eta_t \\ 0 \\ 0 \end{bmatrix}. \end{aligned} \tag{2.9}$$

Com isso, o modelo de volatilidade estocástica é definido como um modelo de espaço-estado constante através das matrizes:

$$\boldsymbol{\theta}_t = \begin{bmatrix} h_t \\ \mu \\ E(\ln \delta_t^2) \end{bmatrix}, \quad \mathbf{F}_t = \mathbf{F} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{e} \quad \mathbf{G}_t = \mathbf{G} = \begin{bmatrix} \phi & 1 - \phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.10}$$

Os termos que representam os erros nas equações das observações e do sistema são:

$$\nu_t = \varepsilon_t \quad \text{e} \quad \boldsymbol{\omega}_t = \begin{bmatrix} \eta_t \\ 0 \\ 0 \end{bmatrix}. \quad (2.11)$$

Consequentemente, as variâncias de ν_t e $\boldsymbol{\omega}_t$ são definidas, respectivamente, por:

$$V_t = V = \text{Var}(\varepsilon_t) \quad \text{e} \quad \mathbf{W}_t = \mathbf{W} = \begin{bmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (2.12)$$

A variância V é obtida diretamente através da distribuição de ε_t . A matriz \mathbf{W} , por sua vez, é determinada a partir da matriz $\boldsymbol{\omega}_t$ em (2.11) e da distribuição de η_t em (2.2).

Há de se destacar que, apesar de h_t ser a variável que determina a evolução do sistema pela definição do próprio modelo em (2.1), o vetor $\boldsymbol{\theta}_t$, que capta essa característica, só pôde ser escrito em função de h_t , do parâmetro a ser estimado μ e da constante $E(\ln \delta_t^2)$.

A distribuição *a priori* de $\boldsymbol{\theta}_t$ é definida por $\boldsymbol{\theta}_0 \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0)$, onde:

$$\mathbf{m}_0 = \begin{bmatrix} \mu \\ \mu \\ E(\ln \delta_t^2) \end{bmatrix} \quad \text{e} \quad \mathbf{C}_0 = \begin{bmatrix} \frac{\sigma_\eta^2}{1-\phi^2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (2.13)$$

2.3 Definição do Modelo Bayesiano

A estimação Bayesiana dos parâmetros do modelo de volatilidade estocástica consiste em determinar a distribuição *a posteriori* de (2.3). Apesar de $\boldsymbol{\psi}$ ser um vetor com três elementos, no processo de estimação ele pode ser dividido em blocos. Assim, como Petris et al. (2009) diz, não é necessário obter a distribuição *a posteriori* conjunta de μ , ϕ e σ_η^2 . Portanto, foram tomadas três distribuições, uma para cada elemento de $\boldsymbol{\psi}$, independentes entre si.

Como o paradigma Bayesiano sugere, deve-se definir as distribuições *a priori* dos parâmetros. Foram assumidas as distribuições *a priori* conforme Kastner e Frühwirth-Schnatter (2014) e Kim et al. (1998).

O nível μ da volatilidade tem seu suporte em \mathbb{R} , e será atribuída uma distribuição *a priori* com densidade gaussiana para o seu valor:

$$\mu \sim \mathcal{N}(a_\mu, B_\mu). \quad (2.14)$$

O parâmetro ϕ determina a persistência da volatilidade, e $|\phi| < 1$. A distribuição beta é bastante flexível e a escolha mais comum quando se deseja modelar alguma variável

cujo valor esteja no intervalo $(0, 1)$. Com a finalidade de aproveitar essa vantagem da distribuição beta, seja uma nova variável ϕ_0 cuja distribuição *a priori* é $\mathcal{B}(a_\phi, b_\phi)$. A persistência então será $\phi = 2\phi_0 - 1$ e sua distribuição *a priori* será:

$$\pi(\phi) = \frac{\Gamma(a_\phi + b_\phi)}{2\Gamma(a_\phi)\Gamma(b_\phi)} \left(\frac{1 + \phi}{2}\right)^{a_\phi - 1} \left(\frac{1 - \phi}{2}\right)^{b_\phi - 1}. \quad (2.15)$$

O terceiro e último parâmetro do modelo é σ_η^2 , a variância da volatilidade, e seus possíveis valores estão em \mathbb{R}^+ . Sua distribuição *a priori* será:

$$\sigma_\eta^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2B_\sigma}\right). \quad (2.16)$$

2.4 Estudo de Simulação

A fim de avaliar alguns métodos de estimação a seguir, foram criadas populações com dados simulados, conforme a definição em (2.1). As populações se caracterizam pelo produto cruzado dos possíveis valores predeterminados para μ , ϕ e σ_η^2 a seguir:

- $\mu \in \{-5, 4\}$,
- $\phi \in \{0, 0; 0, 65; 0, 99\}$,
- $\sigma_\eta^2 \in \{0, 1^2; 0, 3^2; 0, 5^2\}$.

No caso de μ , a opção de se trabalhar com um único valor fixado é devido ao conhecimento prévio de que tal parâmetro não apresenta maiores dificuldades técnicas em sua estimação, Kastner (2015).

Foram gerados 64 conjuntos de dados para cada um dos nove possíveis valores distintos de $\psi = (\mu, \phi, \sigma_\eta^2)$. Todas as populações tinham o mesmo tamanho de $n = 4 \times 365,25 = 1.461$ elementos, o que corresponde a quatro anos de observações diárias. A geração dos dados bem como os cálculos computacionais de estimação foram todos implementados em R (Ihaka e Gentleman, 1996). O Código A.1 no apêndice A apresenta a função `sv_data` que é responsável por gerar os dados simulados do modelo de volatilidade estocástica descrito em 2.1. Essa função possui quatro parâmetros: o tamanho da amostra desejada `n`, o valor `mu` do nível μ , o valor `phi` da persistência ϕ e o valor `sigma` do desvio padrão σ_η . O resultado da função é um objeto do tipo `data.frame`, que na verdade é um conjunto de dados com os valores simulados `y` da variável observada Y_t e `h` da variável latente h_t .

A Figura 2.2 mostra um exemplo de um conjunto de dados simulados com a persistência extremamente alta, $\phi = 0,99$, e a volatilidade a maior possível do arranjo proposto, $\sigma_\eta^2 = 0,5^2$.

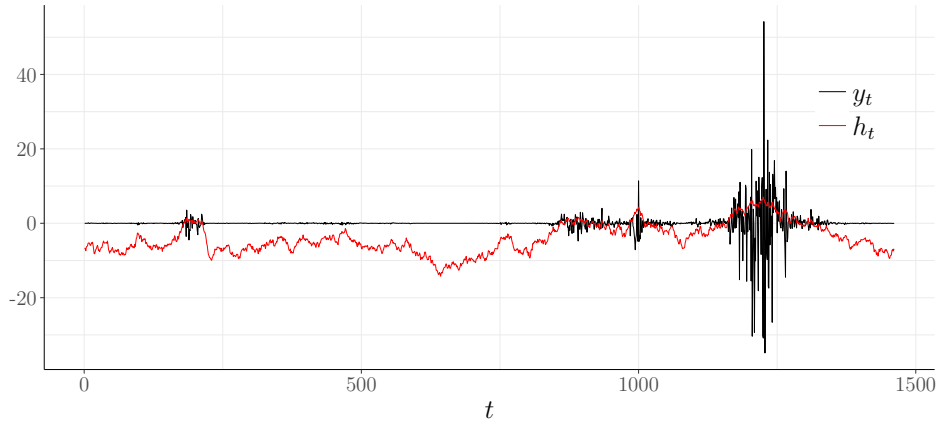


Figura 2.2: Conjunto de dados simulados a partir do modelo (2.1), $\mu = -5,4$, $\phi = 0,99$ e $\sigma_\eta^2 = 0,5^2$.

Devido a complexidade do modelo, formas fechadas para as distribuições *a posteriori* dos parâmetros não são obtidas trivialmente. Portanto, num primeiro momento, serão utilizados algoritmos baseados em métodos de Monte Carlo em cadeias de Markov (MCMC) para a obtenção de observações das distribuições *a posteriori* de μ , ϕ e σ_η^2 .

2.4.1 Amostrador de Gibbs via JAGS

O método de estimação dos parâmetros inicialmente testado será baseado no amostrador de Gibbs através de sua implementação no *software* JAGS, Plummer (2003). JAGS é um acrônimo para a expressão em inglês **J**ust **A**nother **G**ibbs **S**ampler, e sua escolha foi motivada por dois fatores.

O primeiro motivo é em relação ao desenvolvimento e manutenção do programa. Enquanto a solução mais tradicional, WinBUGS (Lunn et al., 2000), foi descontinuada em 2007, o desenvolvimento do JAGS permanece ativo. Assim, a princípio, é um programa menos suscetível a *bugs* ou outros problemas. O segundo fator é devido a ótima integração com o R, através do pacote **R2jags** (Su e Yajima, 2015).

O JAGS, apesar de integrado ao R, possui uma linguagem própria, descrita em Plummer (2015). As variáveis do modelo estudado devem ser declaradas segundo a sua natureza (determinística ou estocástica). Sua sintaxe é muito parecida com R. Contudo, deve-se tomar cuidado quanto à implementação das distribuições. A normal, por exemplo, foi programada a partir de sua precisão ($\tau = \sigma^{-2}$).

O código fonte do modelo canônico de volatilidade estocástica (2.1) em JAGS está no Código A.2 do apêndice A. As distribuições de Y_t e h_t , que são a base dessa

implementação são:

$$Y_t \sim \mathcal{N}(0, e^{h_t}) \quad \text{e} \quad (2.17)$$

$$(h_t|h_{t-1}) \sim \mathcal{N}(\mu + \phi(h_{t-1} - \mu), \sigma_\eta^2). \quad (2.18)$$

Em cada população simulada, os parâmetros μ , ϕ e σ_η^2 foram estimados com base no seguinte procedimento. O modelo foi calibrado através de um *burn-in* igual a 5.000. Então, mais 10.000 iterações foram executadas, e uma a cada 10 observações foi salva. Resultando, assim, numa amostra *a posteriori* de tamanho 1.000 para cada um dos três parâmetros em cada população.

As distribuições *a priori* dos parâmetros (2.14), (2.15) e (2.16) foram definidas como:

$$\mu \sim \mathcal{N}(0, 100), \quad \phi_0 \sim \mathcal{B}(1, 1) \quad \text{e} \quad \sigma_\eta^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right). \quad (2.19)$$

Os parâmetros da *priori* de μ , em geral, não influenciam no seu processo de estimação. Como cita Kastner (2015), a verossimilhança carrega informação suficiente para estimar μ com sucesso.

Para ϕ foi adotada inicialmente uma *priori* não informativa, contrário ao recomendado por Kim et al. (1998), que alega que uma *priori* vaga pode gerar problemas de estimação quando os dados forem próximos da não-estacionariedade.

O parâmetro da distribuição de σ_η^2 foi escolhido de maneira que uma massa considerável de probabilidade estivesse concentrada no intervalo (0, 1).

Após o procedimento computacional, os resultados foram compilados e podem ser observados nas figuras e tabelas que seguem.

A Figura 2.3 mostra as distribuições das médias estimadas de μ , $\bar{\mu}$, para os 64 conjuntos de dados simulados a partir dos nove possíveis valores de ψ . Levando em consideração que o valor real de μ é -5,4, a estimação do parâmetro, quando $\phi = 0,99$, foi a pior de todas. O valor não parece viesado, porém apresentou a maior dispersão entre os demais. A Tabela 2.3 mostra a média, o mínimo, a mediana, o máximo e o desvio padrão dos valores estimados $\bar{\mu}$, respectivamente.

A Figura 2.4 mostra que quando a persistência ϕ é elevada, o valor real desse parâmetro é razoavelmente estimado. Valores intermediários de ϕ são estimados com maior precisão a medida que o valor da volatilidade, σ_η^2 , aumenta. Quando a persistência é inexistente, $\phi = 0$, os valores estimados do parâmetro são, em média, não viesados, no entanto, apresentam alta dispersão. A Tabela 2.4 mostra novamente a média, o mínimo, a mediana, o máximo e o desvio padrão dos valores estimados $\bar{\phi}$, respectivamente.

A Figura 2.5 mostra que a volatilidade, σ_η^2 , somente é estimada de maneira aceitável quando o valor de ϕ é próximo de um, pois a distribuição dos valores estimados de σ_η torna-se leptocúrtica com moda em torno de σ_η real. A Tabela 2.5 mostra as estatísticas de $\bar{\sigma}_\eta$ que comprovam essa afirmação.

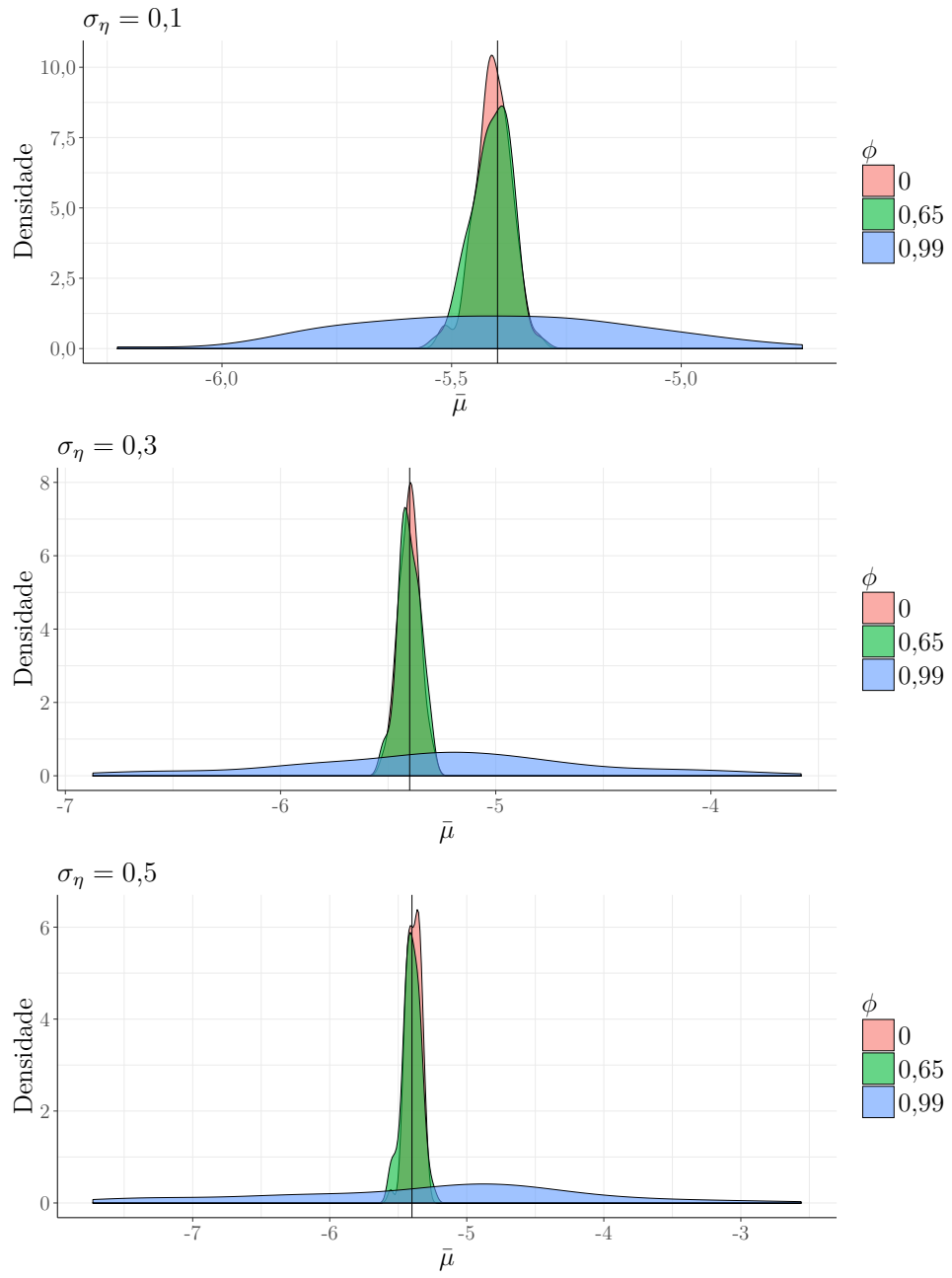


Figura 2.3: Valores de $\bar{\mu}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.

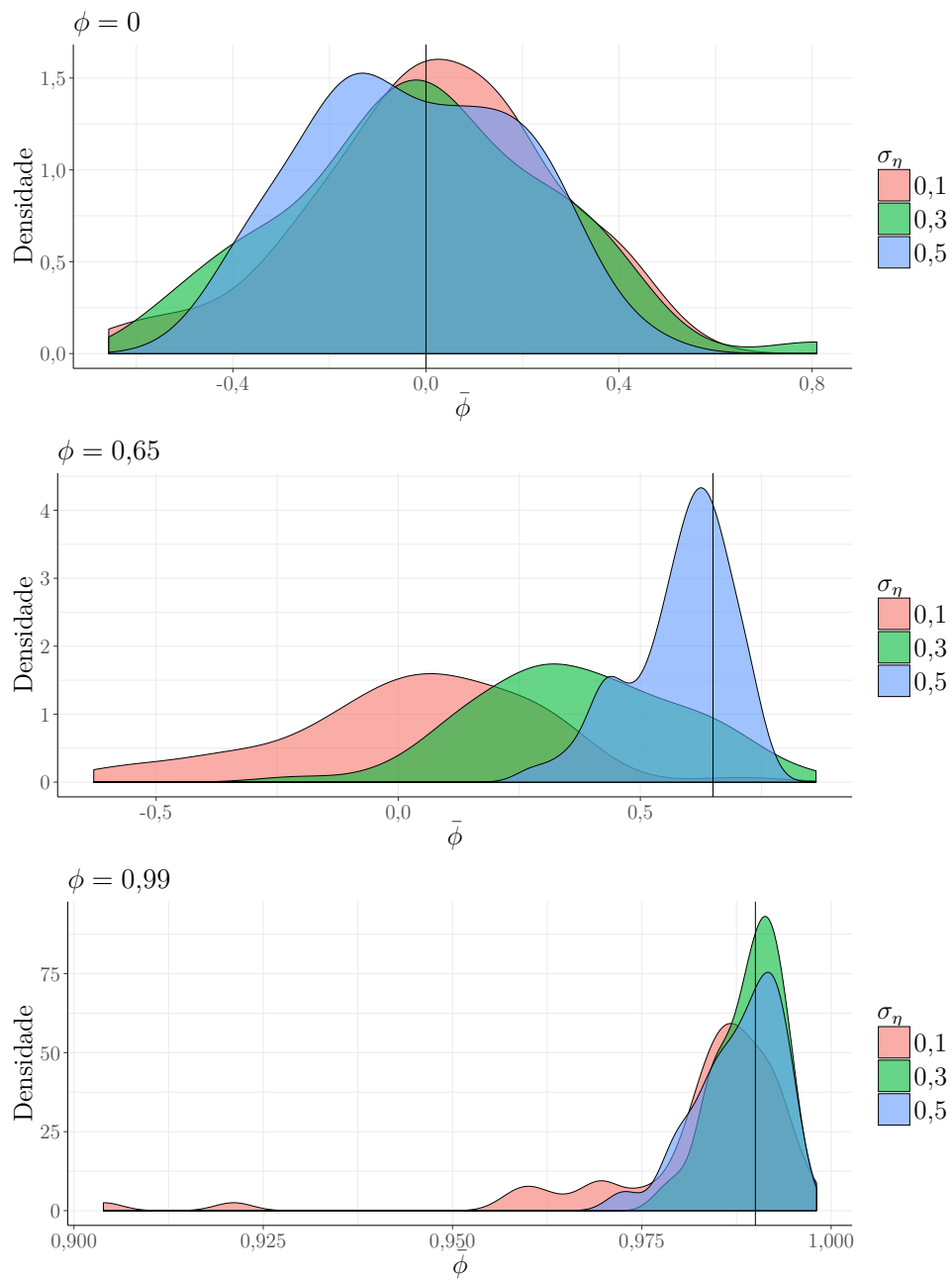


Figura 2.4: Valores de $\bar{\phi}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.

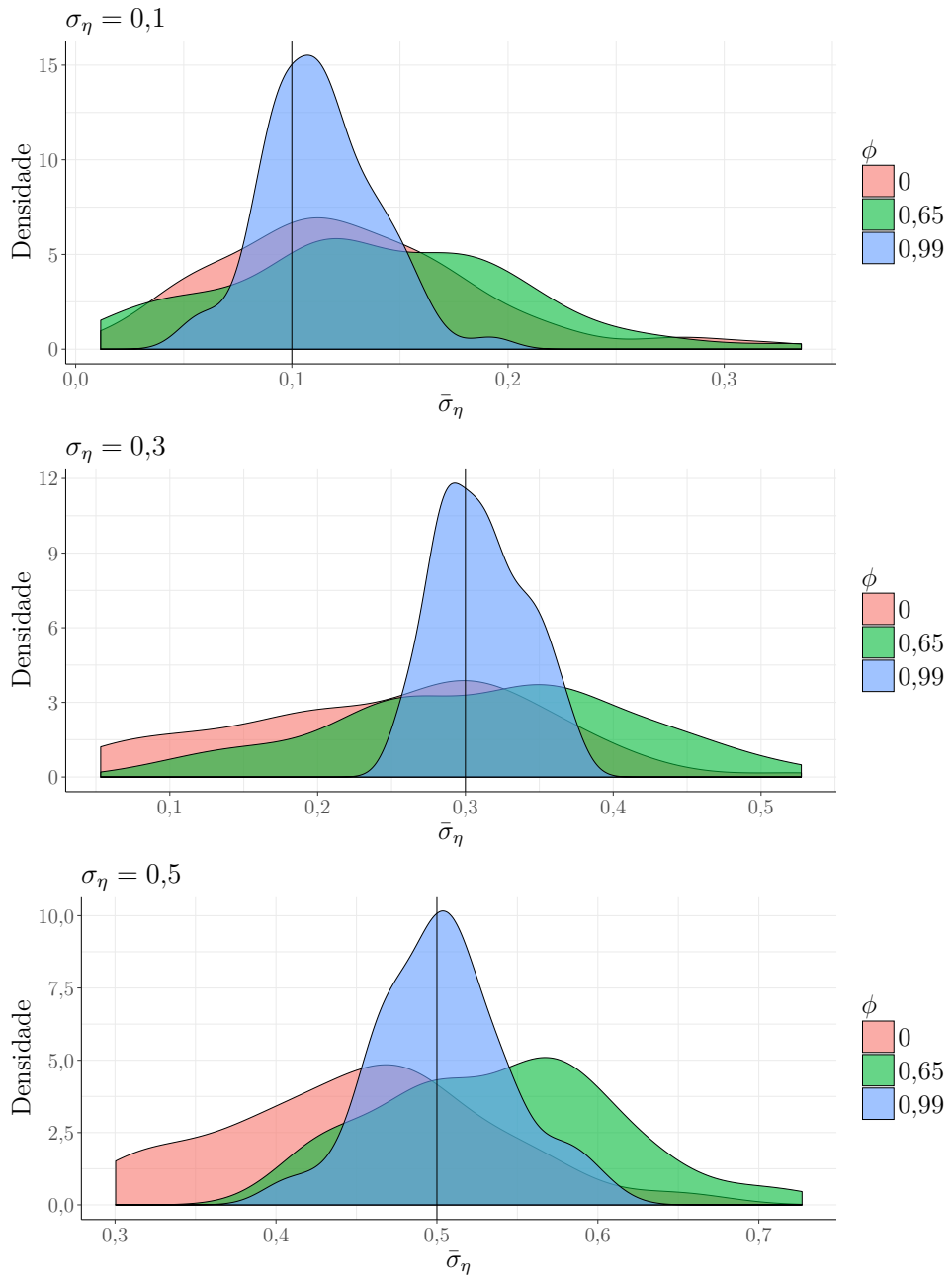


Figura 2.5: Valores de $\bar{\sigma}_\eta$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.

Tabela 2.3: Principais estatísticas dos valores estimados $\bar{\mu}$ através do amostrador de Gibbs para o conjunto de populações de dados simulados.

ϕ_{real}	$\sigma_{\eta_{\text{real}}}$	média($\bar{\mu}$)	mín($\bar{\mu}$)	Md($\bar{\mu}$)	máx($\bar{\mu}$)	$\sigma_{\bar{\mu}}$
0	0,1	-5,410	-5,521	-5,409	-5,312	0,039
0	0,3	-5,403	-5,527	-5,401	-5,284	0,049
0	0,5	-5,388	-5,555	-5,386	-5,293	0,053
0,65	0,1	-5,414	-5,533	-5,411	-5,308	0,043
0,65	0,3	-5,402	-5,529	-5,408	-5,292	0,052
0,65	0,5	-5,402	-5,558	-5,401	-5,242	0,065
0,99	0,1	-5,411	-6,228	-5,413	-4,736	0,297
0,99	0,3	-5,240	-6,872	-5,241	-3,581	0,699
0,99	0,5	-5,277	-7,728	-4,995	-2,560	1,142

Tabela 2.4: Principais estatísticas dos valores estimados $\bar{\phi}$ através do amostrador de Gibbs para o conjunto de populações de dados simulados.

ϕ_{real}	$\sigma_{\eta_{\text{real}}}$	média($\bar{\phi}$)	mín($\bar{\phi}$)	Md($\bar{\phi}$)	máx($\bar{\phi}$)	$\sigma_{\bar{\phi}}$
0	0,1	0,012	-0,657	0,017	0,503	0,248
0	0,3	-0,006	-0,547	-0,013	0,810	0,271
0	0,5	-0,027	-0,465	-0,032	0,448	0,218
0,65	0,1	0,025	-0,629	0,054	0,708	0,259
0,65	0,3	0,373	-0,209	0,364	0,863	0,213
0,65	0,5	0,582	0,277	0,602	0,750	0,105
0,99	0,1	0,982	0,904	0,986	0,998	0,016
0,99	0,3	0,989	0,977	0,990	0,997	0,004
0,99	0,5	0,988	0,972	0,989	0,995	0,005

Tabela 2.5: Principais estatísticas dos valores estimados $\bar{\sigma}_{\eta}$ através do amostrador de Gibbs para o conjunto de populações de dados simulados.

ϕ_{real}	$\sigma_{\eta_{\text{real}}}$	média($\bar{\sigma}_{\eta}$)	mín($\bar{\sigma}_{\eta}$)	Md($\bar{\sigma}_{\eta}$)	máx($\bar{\sigma}_{\eta}$)	$\sigma_{\bar{\sigma}_{\eta}}$
0	0,1	0,126	0,012	0,118	0,323	0,061
0	0,3	0,247	0,053	0,271	0,527	0,104
0	0,5	0,451	0,300	0,454	0,658	0,082
0,65	0,1	0,137	0,013	0,128	0,336	0,067
0,65	0,3	0,314	0,097	0,333	0,527	0,100
0,65	0,5	0,539	0,412	0,546	0,727	0,072
0,99	0,1	0,113	0,055	0,112	0,192	0,026
0,99	0,3	0,310	0,258	0,308	0,372	0,030
0,99	0,5	0,504	0,404	0,501	0,604	0,041

2.5 Um Exemplo com Dados Reais

O amostrador de Gibbs via JAGS, aplicado anteriormente em dados simulados, foi utilizado para replicar o estudo em Achcar et al. (2011). No referido artigo, os autores propõem dois modelos de volatilidade estocástica levemente diferente do modelo canônico em 2.1 para estudar a concentração de ozônio na Cidade do México. A base de dados corresponde a 16 anos (1º de janeiro 1990 a 31 de dezembro 2005) de medições da média semanal do nível diário máximo de ozônio em cinco regiões da Cidade do México.

Em ambos os casos, a variável aleatória de interesse, X_t , é a concentração de ozônio no ar da Cidade do México em partes por milhão (ppm) aferida no tempo $t = 1, \dots, N$ em $j = 1, \dots, K$ regiões. Entretanto a variável, Y_t , a ser modelada é a série de log-retornos de X_t , e é definida, na região j , por:

$$Y_{j,t} = \ln \left(\frac{X_{j,t}}{X_{j,t-1}} \right), \quad t = 1, \dots, N. \quad (2.20)$$

2.5.1 Modelo I

Os autores definem o primeiro modelo, no tempo $t = 1, \dots, N$ e na região $j = 1, \dots, K$, como:

$$\text{Modelo I: } \begin{cases} Y_{j,t} \sim \mathcal{N}(0, e^{h_{j,t}}) \\ h_{j,t} = \begin{cases} \mu_j + \eta_{j,t}, & t = 1, \\ \mu_j + \phi_j(h_{j,t-1} - \mu_j) + \eta_{j,t}, & t = 2, \dots, N. \end{cases} \end{cases} \quad (2.21)$$

onde $h_{j,t}$ é a variável latente que contabiliza a volatilidade do sistema no tempo t e na região j , e um modelo AR(1) é utilizado para representar a evolução dessa variável. Os parâmetros μ_j e ϕ_j representam o nível e a persistência dos log-retornos, respectivamente. O termo $\eta_{j,t}$ é um ruído aleatório com distribuição $\mathcal{N}(0, \sigma_{\eta_j}^2)$, e, em $t = 1, \dots, N$, são independentes entre si.

Essa definição difere um pouco do modelo canônico enunciado em Kim et al. (1998) e sintetizada em (2.1). A parte auto-regressiva que modela a volatilidade nessa nova definição não leva em consideração a distribuição *a priori* de $h_{j,0}$ como em (2.1). Nesse caso, é como se *a priori* da volatilidade fosse no tempo $t = 1$ e definida por:

$$h_{j,1} \sim \mathcal{N}(\mu_j, \sigma_{\eta_j}^2). \quad (2.22)$$

Os valores de μ_j , ϕ_j e $\sigma_{\eta_j}^2$ devem ser estimados. Para tanto, foram assumidas as mesmas distribuições *a priori* dos parâmetros definidas em (2.14), (2.15) e (2.16).

O Código A.3 no apêndice A mostra a tradução em JAGS do modelo I, (2.21).

2.5.2 Modelo II

O segundo modelo se difere do primeiro, (2.21), apenas na variância de $Y_{j,t}$. Um termo com a finalidade de captar um efeito comum à volatilidade de todas as regiões é adicionado. Assim, o modelo II, no tempo $t = 1, \dots, N$ e na região $j = 1, \dots, K$, é:

$$\text{Modelo II : } \begin{cases} Y_{j,t} \sim \mathcal{N}(0, e^{h_{j,t} + \omega_t}). \\ h_{j,t} = \begin{cases} \mu_j + \eta_{j,t}, & t = 1, \\ \mu_j + \phi_j(h_{j,t-1} - \mu_j) + \eta_{j,t}, & t = 2, \dots, N, \end{cases} \\ \omega_t \sim \mathcal{N}(0, \sigma_\omega^2). \end{cases} \quad (2.23)$$

Como η_t em (2.1), a variável ω_t tem distribuição $\mathcal{N}(0, \sigma_\omega^2)$ independente no tempo, e sua variância σ_ω^2 deve ser estimada. Sua distribuição *a priori* é definida de forma semelhante àquela de σ_η^2 em (2.16):

$$\sigma_\omega^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2B_{\sigma_\omega}}\right). \quad (2.24)$$

Os demais parâmetros e variáveis em (2.23) são idênticos àqueles já definidos no modelo I na seção 2.5.1.

O código fonte do modelo II definido em (2.23) está no Código A.4 no apêndice A.

2.6 Estimação dos parâmetros

A estimação dos parâmetros foi feita da mesma maneira que aquela para dados simulados descrita na seção 2.4.1. Foi utilizado o amostrador de Gibbs via JAGS e o procedimento foi idêntico ao já descrito na página 15. Em suma, foi feito um *burn-in* de 5.000 iterações, seguido de mais 10.000. Dessas uma a cada 10 amostras dos parâmetros foi salva, resultando assim, numa amostra final de 1.000 observações de cada um dos parâmetros do modelo.

Os valores estimados dos parâmetros foram obtidos através da média *a posteriori* e correspondentes estimativa de desvio-padrão.

A Tabela 2.6 sumariza a comparação entre os valores estimados dos parâmetros em Achcar et al. (2011) com aqueles obtidos através da replicação do modelo I, (2.21), e do modelo canônico, (2.1). Já a Tabela 2.7 faz o mesmo, porém tendo por base o modelo II, (2.23).

Na primeira linha de cada parâmetro nas tabelas 2.6 e 2.7 são apresentados os valores estimados obtidos na implementação via WinBUGS em Achcar et al. (2011), que até onde se percebe, foi feita de maneira equivocada. Talvez devido a uma pequena desatenção, pois foi utilizada uma parametrização da variância diferente daquela descrita no manual do *software* utilizado (Lunn et al. (2000) e Plummer (2015)). Nas referidas

tabelas essa parametrização indevida é representada por σ^2 .

Na segunda linha de cada parâmetro nas tabelas 2.6 e 2.7 são mostrados os valores estimados dos parâmetros para o modelo I e II, respectivamente, obtidos com a parametrização das variâncias que está de acordo com os manuais do WinBUGS e do JAGS, representada por $1/\sigma^2$.

Na segunda linha de cada parâmetro nas tabelas 2.6 e 2.7 são apresentados os valores estimados dos parâmetros pelo modelo canônico.

Tabela 2.6: Comparação entre valores estimados dos parâmetros do modelo I em Achcar et al. (2011) e do modelo canônico para a região NW.

Região	Parâmetro	Modelo	Parame- trização	Valor estimado	Desvio padrão	95% Int. de Credibilidade
NW	μ	Modelo I	σ^2	-2,817	0,075	(-2,972; -2,670)
		Modelo I	$1/\sigma^2$	-2,796	0,080	(-2,950; -2,635)
		M. Canônico	$1/\sigma^2$	-2,789	0,078	(-2,947; -2,631)
	ϕ	Modelo I	σ^2	0,690	0,086	(0,498; 0,830)
		Modelo I	$1/\sigma^2$	0,791	0,076	(0,628; 0,912)
		M. Canônico	$1/\sigma^2$	0,775	0,093	(0,546; 0,903)
	σ_η	Modelo I	σ^2	5,413	1,420	(3,092; 8,610)
		Modelo I	$1/\sigma^2$	0,336	0,077	(0,203; 0,500)
		M. Canônico	$1/\sigma^2$	0,327	0,075	(0,214; 0,485)

Tabela 2.7: Comparação entre valores estimados dos parâmetros do modelo II em Achcar et al. (2011) e do modelo canônico para a região NW.

Região	Parâmetro	Modelo	Parame- trização	Valor estimado	Desvio padrão	95% Int. de Credibilidade
NW	μ	Modelo II	σ^2	-3,008	0,072	(-3,148; -2,854)
		Modelo II	$1/\sigma^2$	-2,993	0,045	(-3,063; -2,902)
		M. Canônico	$1/\sigma^2$	-3,000	0,055	(-3,094; -2,871)
	ϕ	Modelo II	σ^2	0,300	0,184	(0,021; 0,700)
		Modelo II	$1/\sigma^2$	0,009	0,504	(-0,755; 0,892)
		M. Canônico	$1/\sigma^2$	-0,254	0,544	(-0,915; 0,833)
	σ_η	Modelo II	σ^2	8,755	2,172	(5,027; 13,490)
		Modelo II	$1/\sigma^2$	0,036	0,025	(0,010; 0,086)
		M. Canônico	$1/\sigma^2$	0,055	0,027	(0,020; 0,126)
-	σ_ω	Modelo II	σ^2	1,279	0,106	(1,091; 1,512)
		Modelo II	$1/\sigma^2$	0,912	0,036	(0,843; 0,983)
		M. Canônico	$1/\sigma^2$	0,916	0,034	(0,848; 0,983)

Considerações Preliminares

Essa primeira etapa do trabalho foi concluída com sucesso, uma vez que o objetivo principal era entender o modelo de volatilidade estocástica junto com suas principais dificuldades. Em suma, o processo de estimação dos parâmetros do modelo é bastante complicado, uma vez que mesmo em situações com dados simulados os resultados podem ser inconclusivos. Isso ocorre principalmente quando os valores de ϕ ou σ_η^2 são próximos de zero. O que faz sentido, uma vez que esses parâmetros são dominantes na modelagem da variável latente h_t , e se seus valores forem nulos deve-se inclusive questionar a aplicabilidade do MVE.

Para a dissertação, será estudado o algoritmo *forward filtering-backward sampling* proposto por Carter e Kohn (1994), Frühwirth-Schnatter e Kaufmann (1994) e Shephard (1994). Ele é baseado no amostrador de Gibbs e será o elo de ligação entre o modelo de volatilidade estocástica e a estimação dos seus parâmetros como uma modelo de espaço-estado.

Um estudo sobre a convergência dos parâmetros em algoritmos baseados em Monte Carlo em cadeias de Markov, bem como a própria estrutura desses algoritmos será detalhado futuramente.

Por fim, será feita uma análise de valores preditos pelo modelo de volatilidade estocástico em conjuntos simulados, bem como em dados reais.

Referências Bibliográficas

- Achcar, J. A., Rodrigues, E. R., e Tzintzun, G. (2011). Using stochastic volatility models to analyse weekly ozone averages in Mexico City. *Environmental and Ecological Statistics*, 18(2):271–290.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327.
- Bos, C. S. (2012). *Handbook of Volatility Models and their Applications*, chapter Relating Stochastic Volatility Estimation Methods, pages 147–174. John Wiley & Sons.
- Brockwell, P. J. e Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer Texts in Statistics. Springer, New York, 2nd edition.
- Carter, C. K. e Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81:541–553.
- Durbin, J. e Koopman, S. J. (1997). Monte Carlo maximum likelihood estimation for non-gaussian state-space models. *Biometrika*, 89:603–615.
- Durbin, J. e Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, 2nd edition.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007.
- Frühwirth-Schnatter, S. e Kaufmann, S. (1994). Model-based clustering of multiple time series. *Journal of Business and Economic Statistics*, 26:78–89.
- Gallant, A. R. e Tauchen, G. (1996). Which moments to match? *Econometric Theory*, 12:657–681.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., e Rubin, D. B. (2014). *Bayesian Data Analysis*. Chapman & Hall/CRC, Boca Raton, 3rd edition.
- Harvey, A., Ruiz, E., e Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, 61:247–264.

- Ihaka, R. e Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphics Statistics*, 5(3):299–314.
- Jaynes, E. T. (2003). *Probability Theory – The Logic of Science*. Cambridge University Press, Cambridge, 1st edition.
- Kastner, G. (2015). Dealing with stochastic volatility in time series using the R package stochvol. *Journal of Statistical Software*, (forthcoming):to appear.
- Kastner, G. e Frühwirth-Schnatter, S. (2014). Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics and Data Analysis*, 76:408–423.
- Kim, S., Shepard, N., e Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, 65(3):361–393.
- Lunn, D. J., Thomas, A., Best, N., e Spiegelhalter, D. (2000). WinBUGS - A Bayesian modelling framework: Concepts, structure and extensibility. *Statistics and Computing*, 10(4):325–337.
- Nguyen, T. M. (2007). A new efficient algorithm for the analysis with non-linear and non-gaussian state space models. Master’s thesis, VU University Amsterdam.
- Omori, Y., Chib, S., Shephard, N., e Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140:425–449.
- Pearl, J. (2009). *Causality: Models, Reasoning and Inference*. Cambridge University Press, Cambridge, 2nd edition.
- Petris, G. (2010). An R package for dynamic linear models. *Journal of Statistical Software*, 36(12):1–16.
- Petris, G., Petrone, S., e Campagnoli, P. (2009). *Dynamic Linear Models with R*. Springer, New York, 1st edition.
- Plummer, M. (2003). JAGS: A program for analysis of bayesian graphical models using Gibbs sampling. In *Distributed Statistical Computing*, Vienna University of Technology.
- Plummer, M. (2015). *JAGS Version 4.0.0. user manual*.
- Pole, A., West, M., e Harrison, J. (1994). *Applied Bayesian Forecasting and Time Series Analysis*. Chapman & Hall/CRC, Boca Raton, 1st edition.
- Richard, J. F. e Zhang, W. (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics*, 141:1385–1411.

- Shephard, N. (1994). Partial non-gaussian state space models. *Biometrika*, 81:115–131.
- Shephard, N. e Pitt, M. K. (1997). Likelihood analysis of non-gaussian measurement time series. *Biometrika*, 84:653–667.
- Sivia, D. e Skilling, J. (2006). *Data Analysis: A Bayesian Tutorial*. Oxford University Press, Oxford, 2nd edition.
- Su, Y. e Yajima, M. (2015). *R2jags: Using R to Run JAGS*. R package version 0.5-6.
- Taylor, S. (1982). Financial returns modelled by the product of two stochastic processes – a study of daily sugar prices 1961-75. *Time Series Analysis: Theory and Practice*, 1:203–226.
- West, M. e Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer, New York, 2nd edition.
- Zivot, E. e Yollin, G. (2012). Time series forecasting with state space models. In *R/Finance Workshop, University of Chicago*.

Apêndice A

Códigos-Fonte

Código A.1: Função Responsável por Gerar os Dados Simulados.

```
1 sv_data <- function(n, mu, phi, sigma) {  
2   h0 <- rnorm(1, mu, sigma / sqrt(1 - phi ** 2))  
3   h <- rep(mu + phi * (h0 - mu) + sigma * rnorm(1), n)  
4   y <- rep(exp(h[1] / 2) * rnorm(1), n)  
5   for(i in 2:n){  
6     h[i] <- mu + phi * (h[i - 1] - mu) + sigma * rnorm(1)  
7     y[i] <- exp(h[i] / 2) * rnorm(1)  
8   }  
9   return(data.frame(y = y, h = h))  
10 }
```

Código A.2: Modelo Canônico de Volatilidade Estocástica em JAGS definido em (2.1).

```
1 model {  
2   y[1] ~ dnorm(0, exp(-h[1]))  
3   h[1] ~ dnorm(mu + phi * (h0 - mu), 1 / sigma2)  
4   for (t in 2:N) {  
5     y[t] ~ dnorm(0, exp(-h[t]))  
6     h[t] ~ dnorm(mu + phi * (h[t - 1] - mu), 1 / sigma2)  
7   }  
8   h0 ~ dnorm(mu, (1 - phi^2) / sigma2)  
9   mu ~ dnorm(0, 1 / 100)  
10  phi0 ~ dbeta(1.0, 1.0)  
11  phi <- 2 * phi0 - 1  
12  sigma2 ~ dgamma(0.5, 0.5)  
13 }
```

Código A.3: Modelo I de Volatilidade Estocástica em JAGS definido em (2.21).

```
1 model {  
2   for (j in 1:K) {  
3     y[j, 1] ~ dnorm(0.0, exp(-h[j, 1]))  
4     h[j, 1] ~ dnorm(mu[j], 1 / sigma[j]^2)  
5     for (t in 2:N) {  
6       y[j, t] ~ dnorm(0.0, exp(-h[j, t]))  
7       h[j, t] ~ dnorm(mu[j] + phi[j] * (h[j, t - 1] - mu[j]), 1 / sigma[j]^2)  
8     }  
9     mu[j] ~ dnorm(0.0, 0.01)  
10    phi0[j] ~ dbeta(1.0, 1.0)  
11    phi[j] <- 2 * phi0[j] - 1  
12    sigma2[j] ~ dgamma(0.5, 0.5)  
13    sigma[j] <- sqrt(sigma2[j])  
14  }  
15 }
```

Código A.4: Modelo II de Volatilidade Estocástica em JAGS definido em (2.23).

```
1 model {
2   for (j in 1:K) {
3     y[j, 1] ~ dnorm(0.0, exp(-h[j, 1] - w[1]))
4     h[j, 1] ~ dnorm(mu[j], 1 / sigma[j]^2)
5   }
6   w[1] ~ dnorm(0.0, 1 / sigmaomega^2)
7   for (t in 2:N) {
8     for (j in 1:K) {
9       y[j, t] ~ dnorm(0.0, exp(-h[j, t] - w[t]))
10      h[j, t] ~ dnorm(mu[j] + phi[j] * (h[j, t - 1] - mu[j]), 1 / sigma[j]^2)
11    }
12    w[t] ~ dnorm(0.0, 1 / sigmaomega^2)
13  }
14  for (j in 1:K) {
15    mu[j] ~ dnorm(0.0, 0.01)
16    phi0[j] ~ dbeta(1.0, 1.0)
17    phi[j] <- 2 * phi0[j] - 1
18    sigma2[j] ~ dgamma(0.5, 0.5)
19    sigma[j] <- sqrt(sigma2[j])
20  }
21  sigma2omega ~ dgamma(0.5, 0.5)
22  sigmaomega <- sqrt(sigma2omega)
23 }
```