

**Modelos de volatilidade estocástica
com distribuições de caudas pesadas**

Iván Robert Enríquez Guzmán

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
DOUTOR EM CIÊNCIAS

Programa: Estatística

Orientador: Prof. Dr. Pedro Alberto Morettin

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CNPq

São Paulo, novembro de 2010

Modelos de volatilidade estocástica com distribuições de caudas pesadas

Este exemplar corresponde à redação
final da tese devidamente corrigida
e defendida por Ivan Robert Enríquez Guzmán
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Prof. Dr. Pedro Alberto Morettin (orientador) - IME-USP.
- Prof. Dra. Clélia Maria de Castro Toloí - IME-USP.
- Prof. Dr. Pedro Luiz Valls Pereira - FGV-SP.
- Prof. Dr. Thelma Safadi - DEX-UFLA.
- Prof. Dr. Luiz Koodi Hotta - IMECC-UNICAMP.

Agradecimentos

Agradeço em primeiro lugar a Deus que dispensa justificativas. A minha família, que em forma silenciosa mais sempre presente, nunca dexaram de me incentivar, a minha mãe Matilde, ao meu pãe Samuel; que agora mais que nunca posso perceber o grande sacrificio que fizeram para eu chegar aqui.

Quero agradecer de forma especial, ao meu orientador, mestre e amigo, Professor Pedro Alberto Morettin. À profesora Dra. Lucia Barroso, por sua amizade e seu carinho. Aos diversos professores do departamento que contribuíram para minha formação.

A meu filho Pedro, que é minha razão e minha força de luta do día a día nesta vida, e minha esposa Patricia, pela amor que me brinda, a compreensão de minha ausência, quando enfatizei a tese.

Aos meus amigos Cristian Bayes, Pablo Laura, Christian Noriega, Cristhian Paz, pelo apoio e diversas formas de alento, em continuar nesta etapa de minha vida. Aos colegas do doutorado, Michel Helcias, Cristian Villalobos, Sandro, Pedro Fujita, Cibele, Alexandre Patriota e Rafael Bras que com seu companherismo fizeram sem dúvida mais agradável o dia a dia.

Finalmente Agradeço na CNPq pelo apoio financeiro para o desenvolvimento desta tese.

Resumo

A presente tese trata nos modelos de Volatilidade Estocástica usando a classe de mistura na escala de normais e a distribuição normal assimétrica. Os primeiros permitem que o erro do modelo possua distribuições de caudas mais pesadas que a distribuição normal. Esta tentativa já foi estudada anteriormente por diferentes pesquisadores, considerando o modelo com erro t-Student. Implementamos neste contexto, uma alternativa de modelar a característica assimétrica e leptocurtica que apresentam as séries financeiras. Intuitivamente, a idéia é incorporar erros que permitam considerar pontos mais extremos que permite a normal, por meio de distribuições que apresentam caudas mais pesadas, assim como a leve assimetria dos retornos financeiros. Em particular, utilizamos as distribuições Skew Normal (SN), e distribuições de mistura de normais na escala (MEN), dentro da família dos modelos elípticos. Iniciamos com o estudo dos modelos clássicos de volatilidade estocástica (VE), isto é com erro gaussiano, depois consideramos os modelos de volatilidade com mistura na escala de normais, incorporando o erro deste tipo só na série dos log retornos, e depois em ambas as séries de retornos e estados. Finalmente implementamos os modelos de volatilidade estocástica com erro normal assimétrico. Todos estes modelos foram estudados no contexto de considerar o efeito alavanca e sem considerar essa característica. Foram feitas aplicações com a série *S&P500*, encontrando que os modelos VE-Slash e VE-VG (volatilidade estocástica com erro de distribuição Slash e Variância Gama) foram os mais adequados. Incorporamos uma correlação entre os erros com a finalidade de avaliar o efeito de alavancagem, onde os modelos mais adequados resultaram VELE-Slash e VELE-VG (volatilidade estocástica com efeito alavanca e erro de distribuição Slash e Variância Gama). Comparando com a normal assimétrica a família de MEN sem efeito alavanca resulta ser mais adequada segundo o Fator de Bayes e o critério de seleção DIC, o que não acontece ao comparar com os modelos com efeito alavanca.

Palavras-chave: Volatilidade Estocástica , Misturas na Escala de Distribuições Normais , Markov Chain Monte Carlo , Modelos de Espaços de Estados Não Lineares.

Abstract

This thesis is focused on stochastic volatility models using the class of mixture of normals on the scale and the skew normal distribution. The first of them allows the models error to have distributions with tails heavier than the normal distribution. This attempt has already been studied previously by different researchers, considering the model with t-Student error. We implemented an alternative to model the asymmetric and leptokurtic characteristic which is featured by the financial séries. Intuitively, the idea is to incorporate errors which allow to consider the extreme points which allow distributions with heavier tails, such as the slight asymmetry of the financial returns. In particular, we use the Skew Normal distributions (SN), and scale mixture of normal distributions (SMN), within the family of elliptical models. We started with the study of classical stochastic volatility models (SV), ie with Gaussian error, then we consider volatility models with mixture in the scale of normals, incorporating the error of this type only in the log-return séries, and then in both the return and the state series. Finally, we implement the stochastic volatility models with skew normal error. All these models were studied in the context of considering the leverage effect and without considering this feature. Applications were made with the *S&P500* series, to find that the SV-Slash and SV-VG models were the most appropriate. We incorporate a correlation between the errors in order to evaluate the leverage effect, where the most suitable models resulted SV-SlashLE and SV-VGLE. Compared with the skew normal, SMNs family without leverage effect proves to be the most appropriate according to the Bayes factor and the selection criteria DIC.

Keywords: Stochastic Volatility, Scale Mixture of Normal Distributions, Markov Chain Monte Carlo, Non Linear State Space Models.

Sumário

Lista de Abreviaturas	xi
Lista de Símbolos	xiii
1 Introdução	1
1.1 Conceitos Importantes	3
1.1.1 Estimador Monte Carlo Ordinário	3
1.1.2 Intervalos de Credibilidade	4
1.1.3 Métodos de amostragem Gibbs em Inferência Bayesiana	4
1.1.4 O Algoritmo de Metropolis-Hastings (M-H)	5
1.1.5 Critério de Convergência de Geweke (CD)	6
1.1.6 Critérios de Comparação dos Modelos de VE	7
1.1.7 Fator de Bayes	8
2 O Modelo Básico de Volatilidade Estocástica	9
2.1 Métodos utilizados para estimar os modelos de VE	10
2.1.1 Single-move Sampler	11
2.1.2 Mixture Sampler	12
2.1.3 Multi-move Sampler	13
2.1.4 Produto de misturas gaussianas	15
2.1.5 Aplicação	15

2.2	Modelos de Volatilidade Estocástica com Efeito Alavanca	19
2.2.1	Aplicação	20
3	Modelo de Volatilidade Estocástica com Caudas Pesadas	25
3.1	Distribuições de contorno elíptico	25
3.2	Distribuições de misturas de normais na escala (MEN)	26
3.2.1	Definição	27
3.3	Modelos de volatilidade estocástica com erro na família MEN	29
3.3.1	O modelo VE-MEN1:	31
3.3.2	O modelo VE-MEN2:	44
4	Modelos de Volatilidade Estocástica Normais Assimétricos	59
4.1	Definições Prévias	59
4.2	Casos Particulares	60
4.3	Modelo de volatilidade estocástica normal assimétrico	63
4.4	Modelo de volatilidade estocástica com erro normal assimétrico e efeito alavanca	64
4.5	Estimação	68
4.6	Aplicação	71
5	Simulações	77
5.1	Introdução	77
5.2	Comparação de estatísticas de diagnóstico	81
5.3	Ranking de Modelos	88
6	Conclusões	89
6.1	Pesquisas Futuras	90
A	Distribuição normal assimétrica	93
A.1	Distribuição normal assimétrica padrão	93

<i>SUMÁRIO</i>	ix
A.1.1 Função de densidade e distribuição	93
A.1.2 Algumas Considerações	96
B Distribuição normal assimétrica multivariada	99
Referências Bibliográficas	103

Lista de Abreviaturas

- MEN Mistura na Escala de Normais (*Mixture Scala Normal*).
- VE Volatilidade Estocástica (*Stochastic Volatilidade*).
- MEN1 Mistura na Escala de Normais na equação da série.
- MEN2 Mistura na Escala de Normais para ambas equações, equação da série e de estados.

Lista de Símbolos

$\Phi(., .)$ Distribuição de Probabilidade Normal.

Capítulo 1

Introdução

Séries temporais financeiras apresentam características especiais, exigindo a introdução de modelos adequados. Um dessas características é precisamente a heteroscedasticidade da série (conhecida como volatilidade), em outras palavras permitir que a variância da série, condicional a informação passada, seja modelada. A volatilidade tem sido modelada de forma não estocástica e considerando aleatoriedade na equação, tornando-se um processo estocástico. Sob certas restrições próprias de cada modelo, surgem teorias tais como dos processos ARCH, [Engle \(1982\)](#), e sua generalização, os processos GARCH, [Bollerslev \(1986\)](#). De forma paralela os modelos de Volatilidade Estocástica (VE) para séries temporais têm sido expandidos, desde a primeira vez em que foi formulado por [Taylor \(1982\)](#). [Ghysels et al. \(1996\)](#) e [Shephard e Pitt \(1997\)](#) fornecem consideráveis contribuições para o modelo univariado de volatilidade estocástica. Os modelos GARCH são modelos de variância condicional que assumem funções determinísticas dependentes dos retornos passados e variâncias passadas, enquanto nos modelos VE a volatilidade é um processo estocástico.

Alguns estudos têm mostrado que existem evidências de que os erros dos modelos baseados nos retornos apresentam distribuições de caudas pesadas, ver [Jacquier et al. \(2004\)](#), [Kim et al. \(1998\)](#), [Shephard \(1996\)](#), [Omori et al. \(2007\)](#), [Jun \(2006\)](#). Além disso, é conhecido que eles apresentam curtoses muito altas, e ligeiras assimetrias nas suas distribuições. Na literatura existem modelos que medem de alguma forma a assimetria nos estados, como os modelos EGARCH e TGARCH, assim como algumas propostas dos modelos de mistura como o estudado por [Jun \(2005\)](#), que impõe erros com distribuição t-Student, que captura melhor a curtose dos dados. Outros tentam capturar a assimetria diretamente nos dados, pode-se citar o Skew-Garch, [Genton et al. \(2004\)](#) entre outros.

As distribuições de mistura na escala de normais (MEN, [Andrews e Mallows \(1974\)](#)),

nos proporcionam algumas distribuições com caudas pesadas que são freqüentemente mais robustas no processo de inferência, quando temos dados que provêm de uma distribuição assimétrica. Esse fato nos motivou a tentar modelar os modelos de VE de modo a permitir distribuições elípticas para os erros. Tendo em vista o exposto, [Abanto-Valle *et al.* \(2010\)](#) introduziram a classe de modelos de volatilidade com erros seguindo uma distribuição na mistura de escala de normais, de tal forma que permite estabelecer previsões mais robustas. No entanto, não é considerada a correlação entre os erros das equações de estados e retornos, pois consideram não significativa a correlação mencionada, além de ignorar a assimetria que apresentam os retornos (a distribuição do erro é elíptica).

Deve-se notar que nos modelos EGARCH o parâmetro de assimetria é imposto na equação da volatilidade, especificamente na função da curva de impacto da informação dada por,

$$g(\epsilon_t) = \theta\epsilon_t + \gamma(|\epsilon_t|) - E(|\epsilon_t|). \quad (1.1)$$

Para que retornos negativos tenham maior impacto na volatilidade, esperamos $\gamma < 0$. Essa característica que a volatilidade responde mais rapidamente a retornos negativos do que a positivos, é conhecido como efeito alavancagem. Esses modelos têm mais afinidade com os modelos VE do que com os modelos GARCH, pois a equação das volatilidades é modelada como nos modelos de VE, desde que possuem uma componente aleatória dentro da função da curva de impacto $g(\epsilon_t)$. Observamos que pela forma da equação de $g(\epsilon_t)$ tem-se indícios de que a função que controla a assimetria nos estados do modelo é uma combinação linear de um processo gaussiano e uma variável com distribuição normal-truncada. Note que uma forma de gerar uma normal assimétrica é mediante uma combinação linear côncava de uma normal e uma normal truncada, que são independentes, isto é

$$Y = \mu + \sigma \frac{\lambda}{\sqrt{1+\lambda^2}} |U| + \frac{\sigma}{\sqrt{1+\lambda^2}} Z, \quad (1.2a)$$

$$U, Z \sim N(0, 1). \quad (1.2b)$$

Na equação (1.2) U, Z são independentes o que não acontece com ϵ_t e $|\epsilon_t|$ na função $g(\epsilon_t)$.

Objetivos da tese.

Implementamos nesta tese o estudo da VE com ruídos não correlacionados e considerando a correlação entre a série observada de retornos e o log-retorno, modelos conhecidos como VE com efeito alavanca. Assim, expomos o modelo VE-N e VE-NLE (volatilidade estocástica com erro Normal e VE com erro Normal e efeito alavanca [Harvey *et al.* \(1994\)](#)), VE-T e

VE-TLE (volatilidade estocástica com erro t-Student sem e com efeito alavanca), generalizamos este último para os modelos de VE-MEN, o qual considera os anteriores mencionados, além de algumas outras distribuições com caudas pesadas, como a Slash, Variância Gama, Pearson IV, entre outras. Implementamos o modelo VE-Slash e VE-SlashLE, assim como VE-VG e VE-VGLE (volatilidade estocástica com erro Slash/Variância Gama com e sem efeito alavanca, respetivamente).

Nós aplicamos esses modelos à série de retornos da S&P500, obtendo como resultados (baseados na estatística DIC e Fator de Bayes), que o modelo com Slash e Variância Gama dentro da sua respectiva classe (com e sem efeito de alavanca) resultam serem os mais adequados.

Nesta tese também consideramos o estudo da assimetria nos retornos financeiros do ponto de vista da distribuição do ruído, além de considerar o efeito de alavanca. Neste contexto incorporamos a assimetria na série de retornos e permitimos que as log-volatilidades sejam um processo gaussiano.

1.1 Conceitos Importantes

Neste capítulo daremos alguns conceitos que serão utilizados neste trabalho.

1.1.1 Estimador Monte Carlo Ordinário

Consideremos o problema de resolver a integral da forma

$$\int g(\theta)h(\theta|X)d\theta = E[g(\theta)|x], \quad (1.3)$$

onde θ e X podem ser vetores, $g(\theta)$ alguma função integrável. Em muitas ocasiões teremos necessidade de realizar integrais como em (1.3), com a particularidade que seu desenvolvimento analítico não seja fácil de obter. É o caso de médias e covariâncias a posteriori de componentes de θ , da probabilidade a posteriori de marginais de θ em que $g(\theta)$ é, respetivamente, representada por θ , e $I_A(\theta)$ para $A \in \Theta$.

Assumindo que podemos simular uma amostra aleatória $\theta_1, \dots, \theta_n$ da densidade a posteriori $p(\theta|X)$, o método de estimador Monte Carlo aproxima o integral (1.3) pela média empírica

$$\hat{E}[g(\theta|X)] = \frac{1}{n} \sum_{i=1}^n g(\theta_i), \quad (1.4)$$

a qual, pela Lei Forte dos Grandes Números, converge quase certamente para $E(g(\theta)|X)$. Uma medida de precisão desta aproximação é dada pelo erro estimado de Monte Carlo,

$$\frac{1}{\sqrt{n(n-1)}} \left(\sum_{t=1}^n [g(\theta_t) - \frac{1}{n} \sum_{t=1}^n g(\theta_t)] \right)^{1/2}. \quad (1.5)$$

Estes métodos baseados em simulação estocástica são aproximados, já que a amostra simulada, da qual se obtém as quantidades de interesse, não fornece toda a distribuição. No entanto, e desde que o custo computacional cada vez é um fator a favor e não contra, pela potência crescente dos computadores atuais, o nível de precisão é determinado pelo número de simulações, que pode ser aumentado tanto quanto quisermos, permitindo resultados cada vez mais exatos.

1.1.2 Intervalos de Credibilidade

O intervalo de credibilidade central a $100\gamma\%$, $0 < \gamma < 1$, para θ é definido por $\mathcal{R}_c(\gamma) = (\theta_{\frac{\gamma}{2}}, \theta_{1-\frac{\gamma}{2}})$, onde os extremos definem os quantis de probabilidade a posteriori de θ . Uma aproximação Monte Carlo de $\mathcal{R}_c(\gamma)$ é obtida ordenando uma amostra aleatória e calculando os quantis empíricos, assim a estimativa Monte Carlo de \mathcal{R}_c é definida como

$$\hat{\mathcal{R}}_c = \left(\theta_{[\frac{\gamma}{2}]}, \theta_{[1-\frac{\gamma}{2}]} \right) \quad (1.6)$$

onde $[\cdot]$ denota a função parte inteira. Em muitos casos e sobretudo quando temos distribuições assimétricas para os dados, é preferível a implementação do intervalo HPD (Highest Posterior Density), muito utilizado em problemas de inferência Bayesiana.

1.1.3 Métodos de amostragem Gibbs em Inferência Bayesiana

Geman e Geman (1984) introduziram o método de amostragem Gibbs aplicados a problemas de resolução de imagem. O algoritmo é baseado no fato que, se a distribuição conjunta de $f(\theta|X)$ for positiva em $\Theta_1, \dots, \Theta_k$ sendo Θ o suporte da distribuição de θ_i , para $i = 1, \dots, k$, então ela é univocamente determinada pelas distribuições condicionais completas $f(\theta_i|X, \theta_{/i})$ onde $\theta_{/i}$ representa o vector θ sem a i -ésima componente.

O algoritmo é descrito da seguinte forma:

- Seja $\theta^0 = (\theta_1^0, \dots, \theta_k^0)$ um valor arbitrário inicial para o vector θ . Realizamos os seguintes passos:

- Amostramos $\theta_1^{(1)}$ de $p(\theta_1|x, \theta_1^{(0)}, \dots, \theta_k^{(0)})$,
 - Amostramos $\theta_2^{(1)}$ de $p(\theta_2|x, \theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_k^{(0)})$,
 - Amostramos $\theta_3^{(1)}$ de $p(\theta_3|x, \theta_1^{(1)}, \theta_2^{(1)}, \theta_4^{(0)}, \dots, \theta_k^{(0)})$,
 - \vdots
 - Amostramos $\theta_k^{(1)}$ de $p(\theta_k|x, \theta_1^{(1)}, \dots, \theta_{k-1}^{(1)})$,
- Completando uma iteração do algoritmo ao completar todos os valores de $\theta^1 = (\theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_k^{(1)})$.
 - Repete-se o esquema anterior fazendo $\theta^1 = \theta^0$, para obter um novo valor para θ , (θ^2).

Este ciclo é repetido t vezes, gerando desta forma uma amostra $\theta^0, \dots, \theta^t$ do vetor θ . Resultados teóricos nos permite concluir que, quando $t \rightarrow \infty$, $(\theta_1^t, \dots, \theta_k^t)$ tende em distribuição para o vetor aleatório cuja densidade de probabilidade conjunta é $p(\theta|X)$. Em particular, θ_i^t tende a distribuição marginal da posteriori de θ_i . O número de simulações até alcançar convergência é uma tanto arbitrário. Existem numerosos estudos sobre diversos problemas que acarretam realizar um amostrador de Gibbs, uma breve revisão pode-se encontrar em [Gilks *et al.* \(1996\)](#). O problema mais notório é sobre a convergência da cadeia gerada para o estado de equilíbrio. [Gelfand et al. \(1990\)](#) propõem métodos baseados em representações gráficas, monitorando as convergências das médias ergódicas. [Geweke \(1992\)](#) sugerem métodos baseados em séries temporais. [Gelman e Rubin\(1992 a,b\)](#) sugerem métodos baseados na análise de variância, entre outros. Na presente tese fazemos uso do método de convergência de [Geweke \(1992\)](#).

1.1.4 O Algoritmo de Metropolis-Hastings (M-H)

O método de M-H é implementado com base na metodologia da simulação Monte Carlo em Cadeias de Markov (MCMC). O objetivo principal é determinar a distribuição de probabilidade invariante ou estacionária, para a qual convergem as cadeias, a qual chamaremos de $\pi(x)$. As simulações das cadeias são geradas até alcançar a distribuição de equilíbrio, além de obter a condição de invertibilidade, isto é,

$$\pi(x)p(x, y) = \pi(y)p(y, x), \quad (1.7)$$

onde, $\pi(\cdot)$ é invariante com respeito a $p(x, y)$ ([Tierney \(1994\)](#)).

Como no método de aceitação-rejeição, suponha que temos uma densidade de onde se pode gerar valores candidatos para a distribuição do parâmetro de interesse, que chamaremos densidade proposta, denotada por $q(x, y)$. Esta densidade é interpretada como a densidade de transição de passar do estado x para o estado y , isto é, estando no estado x a probabilidade de passar para o estado y é $q(x, y)$. No entanto, a condição de equilíbrio nem sempre se cumpre, mas podemos estabelecer a igualdade considerando um novo fator,

$$\pi(x)p(x, y)\alpha(x, y) = \pi(y)p(y, x)\alpha(y, x), \quad (1.8)$$

$\alpha(x, y)$ é a probabilidade de movimento, definida como

$$\alpha(x, y) = \min \left[\frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}, 1 \right], \quad \text{se } \pi(x)q(x, y) > 0 \quad (1.9)$$

e 1, em caso contrário. Assim, o algoritmo proposto por [Hastings \(1970\)](#), é resumido nos seguintes passos:

- Inicializamos com um valor arbitrário x^0 ;
- Repetimos para $j = 1, 2, \dots, N$:
 - Geramos y da distribuição proposta $q(x^j, \cdot)$ e u da distribuição $U(0, 1)$,
 - Se $u < \alpha(x^j, y)$ então $x^{j+1} = y$, caso contrário $x^{j+1} = x^j$;
- Os valores obtidos da distribuição $\pi(x)$ são (x^0, \dots, x^N) .

A relação necessária entre $q(x, y)$ e $\pi(x, y)$ é que $q(x)$ seja uma densidade sobre o suporte de $\pi(\cdot)$ ou um subconjunto do dito suporte. Um caso particular desta metodologia é o algoritmo do amostrador de Gibbs (veja [Siddhartha e Greenberg \(1995\)](#)). Na presente tese implementamos o algoritmo de M-H, no cálculo de parâmetros do Modelo de VE, como veremos mais adiante.

1.1.5 Critério de Convergência de Geweke (CD)

[Geweke \(1992\)](#), sugere avaliar a convergência da MCMC, comparando os valores no início da sequência com aqueles no final da sequência. Seja $\theta^{(i)}$ a i -ésima simulação do parâmetro de interesse, $\bar{\theta}_A = \frac{\sum_1^{n_A} \theta^{(i)}}{n_A}$ e $\bar{\theta}_B = \frac{\sum_{M+1-n_B}^M \theta^{(i)}}{n_B}$, onde M é o total de simulações geradas na cadeia de Markov. [Geweke \(1992\)](#) usa esses valores e propõe a estatística de convergência,

chamada "convergence diagnostics"(CD) dado por

$$CD = \frac{\bar{\theta}_A - \bar{\theta}_B}{\sqrt{\frac{\hat{\sigma}_A^2}{n_A} + \frac{\hat{\sigma}_B^2}{n_B}}},$$

onde $\hat{\sigma}_A^2 = Var(\bar{\theta}_A)$ e $\hat{\sigma}_B^2 = Var(\bar{\theta}_B)$. Se a sequência é estacionária, a estatística converge em distribuição para uma Normal padrão. Os valores de n_A e n_B são assumidos como 0,010M e 0,05M, respectivamente.

1.1.6 Critérios de Comparação dos Modelos de VE

O critério de informação de Akaike (AIC), tem sido utilizado como critério de seleção em muitos estudos estatísticos. No entanto, ele requer a logverossimilhança e uma medida de penalidade que depende do número de parâmetros considerados no modelo. Logo, para o cálculo do AIC é preciso o número de parâmetros livres. Esse fato gera um problema quando nos encontramos com os modelos Bayesianos Hierárquicos, onde a especificação do número de parâmetros é arbitrária, porque eles incorporam no modelos variáveis latentes, em quantidades nada desprezíveis. É este o contexto em que os modelos de volatilidade estocástica se encontra, dado que no modelo, as log-volatilidades (h_t) são consideradas como estados, que são incorporadas no modelagem como variáveis latentes. Consequentemente, o AIC não pode ser implementado nos modelos de VE, [Berg et al. \(2004\)](#).

Com a finalidade de obter a generalização do critério AIC, [Spiegelhalter et al. \(2002\)](#) propõem o critério de informação de desvio, DIC, utilizado para modelos hierárquicos. Este contém dois componentes,

$$DIC = \bar{D} + p_D = -2E_{\theta|y_n}[\log L(y_n|\theta)] + p_D, \quad (1.10)$$

onde \bar{D} é o desvio médio do modelo, p_D é o termo de penalização pela a complexidade no modelo ($p_D = 2[\log L(y_n|\bar{\theta}) - E_{\theta|y_n}[\log L(y_n|\theta)]]$). De forma similar ao AIC, o modelo que tenha menor valor do DIC deve ser considerado como o modelo mais adequado para o conjunto de dados. Todos os modelos nesta tese foram implementados em WinBUGS (Windows version of Bayesian analysis Using Gibbs Sanpler, [Spiegelhalter et al. \(2002\)](#)).

1.1.7 Fator de Bayes

Outro critério que utilizaremos para comparar os diferentes modelos desenvolvidos é o Fator de Bayes (FB). Em geral, o FB para comparar os modelos M_i com M_j é dado por

$$FB = \frac{p(y|M_i)}{p(y|M_j)},$$

onde $p(y|M_i)$ é a densidade preditiva sob o modelo M_i , logo o fator de Bayes nos indicará a chance de um modelo com respeito a outro.

Capítulo 2

O Modelo Básico de Volatilidade Estocástica

Segundo [Shephard \(2005\)](#), a publicação mais referenciada e considerada a primeira publicação do modelo de volatilidade estocástica é [Taylor \(1982\)](#), além do livro [Taylor \(1986\)](#). Entre as variantes dos modelos de Volatilidade Estocástica (VE), [Taylor \(1982, 1986\)](#), formulou o modelo de VE em tempo discreto como

$$y_t = e^{\frac{h_t}{2}} \varepsilon_t, \quad (2.1a)$$

$$h_t = \alpha + \phi h_{t-1} + \sigma_\eta \eta_t, \quad (2.1b)$$

onde y_t e h_t são os log-retornos e os estados (log-volatilidades), respectivamente, no tempo $t = 1, \dots, n$. As inovações ε_t e η_t são supostas independentes e normalmente distribuídas, com média zero e variância um. Denotamos este modelo como VE-N.

O modelo proposto por Taylor tem um diferencial, pois está ligado à equação EDP de Black-Scholes para opções europeias dado por,

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0,$$

onde $V(S,t)$ é o preço de um derivativo como função do tempo e do preço da ação (S) (veja-se [Steele \(2001\)](#)). O modelo de Black-Scholes supõe que o preço da ação segue um movimento Browniano geométrico com tendência (μ) e volatilidade (σ) constantes, i.e, a equação diferencial estocástica que gera o EDP de Black-Scholes é

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (2.2)$$

$$d\beta_t = r \beta_t dt, \quad (2.3)$$

onde S_t é o preço da opção no instante t e W_t é um processo Browniano (veja [Steele \(2001\)](#)). Em (2.3), β_t denota a condição de contorno para o tempo t e r é a taxa de reversão à média de S_t .

2.1 Métodos utilizados para estimar os modelos de VE

[Taylor \(1994\)](#), [Ghysels et al. \(1996\)](#) e [Shephard \(1996\)](#), mostram como os modelos de volatilidade estocástica têm vantagens comparados com os modelos ARCH, mas eles são difíceis de implementar pela complexidade da função de verossimilhança. Com o desenvolvimento computacional, foram desenvolvidas algumas formas para solucionar o problema no contexto dos métodos MCMC (Markov Chain Monte Carlo) [Gilks et al. \(1996\)](#). Para modelos VE podemos mencionar os trabalhos de [Jacquier et al. \(1994\)](#), [Shephard e Pitt \(1997\)](#), e [Kim et al. \(1998\)](#) entre outros.

Fazendo a transformação seguinte,

$$\log(y_t^2) = h_t + \log \epsilon_t^2, \quad (2.4a)$$

[Harvey et al. \(1994\)](#), desenvolvem uma metodologia baseada no filtro de Kalman para estimar os parâmetros do modelo, maximizando a função de quase-verossimilhança. O método para estimar os parâmetros do modelo VE, assumindo erro normal, por quase-verossimilhança, não é muito adequado como analisa [Kim et al. \(1998\)](#).

A função de verossimilhança $f(y|\theta) = \int f(y|h, \theta) f(h|\theta) \partial h$ é difícil de implementar, mas do ponto de vista da teoria bayesiana, podemos obter a densidade a posteriori $\pi(\theta/y)$ pelo método MCMC. O método implica em estimar os parâmetros partindo da densidade completa $\pi(\theta, h|y)$, onde $h = (h_1, \dots, h_n)$ é o vetor de volatilidades, visto como uma variável latente. Assim podem ser estimados momentos da posteriori e densidades marginais.

[Jun \(2000\)](#), ilustra como os modelos de volatilidade estocástica podem ser implementados fazendo uso de BUGS (Bayesian analysis Using Gibbs Sampling). [Asai \(2005\)](#) avalia os prós e contras dos diferentes métodos em MCMC utilizados em VE. Neste artigo, [Asai \(2005\)](#), compara o "single-move sampler" e o "multi-move sampler", menciona que os métodos eficientes são o "multi-move sampler" e "mixture sampler". Além disso, o método "integration sampler", baseado em "mixture sampler", é mencionado e analisado como outro método eficiente, e propõe uma técnica alternativa baseada em "multi-move sampler", encontrando que esta é a melhor das técnicas para os dados analisados.

Descrevemos a seguir esses métodos como referências.

2.1.1 Single-move Sampler

O modelo básico de volatilidade estocástica é expresso em forma hierárquica como

$$y_t = e^{\frac{h_t}{2}} \epsilon_t, \quad (2.5a)$$

$$h_t = \alpha + \phi h_{t-1} + \sigma_\eta \eta_t, \quad (2.5b)$$

$$\epsilon_t \sim N(0, 1), \quad \eta_t \sim N(0, 1), \quad \eta_t \perp\!\!\!\perp \epsilon_t, \quad (2.5c)$$

onde o vetor de parâmetros será $\boldsymbol{\theta} = (\alpha, \phi, \sigma_\eta^2)$.

A verossimilhança $P(\mathbf{h}_{0:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T})$, não tem forma fechada, de modo que utilizamos o amostrador de Gibbs para simular os valores de interesse dos parâmetros. Seja $\boldsymbol{\theta}$ o vetor de parâmetros, $\mathbf{h}_{0:T} = (h_0, h_1, \dots, h_T)'$ o vetor de log volatilidades, e $\mathbf{y}_{1:T} = (y_1, \dots, y_T)'$ a informação avaliada até o tempo T . Consideramos $\mathbf{h}_{0:T}$ como parâmetros latentes. Usando o Teorema de Bayes, a densidade a posterior dos parâmetros e das variáveis latentes pode ser escrita como

$$p(\mathbf{h}_{0:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T}) \propto p(\mathbf{y}_{1:T} \mid \mathbf{h}_{0:T}) p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}), \quad (2.6)$$

onde

$$p(\mathbf{y}_{1:T} \mid \mathbf{h}_{0:T}) \propto \prod_{t=1}^T e^{-\frac{h_t + y_t^2 e^{-h_t}}{2}}, \quad (2.7)$$

$$p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta}) \propto e^{-\frac{1-\phi^2}{2\sigma_\eta^2} (h_0 - \frac{\alpha}{1-\phi})^2} \prod_{t=1}^T e^{-\frac{1}{2\sigma_\eta^2} (h_t - \alpha - \phi h_{t-1})^2}, \quad (2.8)$$

sendo $p(\boldsymbol{\theta})$ a distribuição a priori.

O algoritmo conhecido por "Single-move Sampler", sugerido por [Jacquier et al. \(1994\)](#) e [Shephard \(1993\)](#), é dado da forma seguinte:

Algoritmo 1

1. Inicializamos ($i = 0$) os valores de $\boldsymbol{\theta}^{(i)}$ e $h^{(i)}$;
2. Amostramos $\mathbf{h}_t^{(i+1)}$, de $\mathbf{h}_{1:T}^{(i+1)} / h_t^{(i+1)}, y_t, \boldsymbol{\theta}^{(i)}$, $t = 1, \dots, T$;
3. Amostramos $\sigma_\eta^{2(i+1)}$ de $p(\sigma_\eta^2 \mid \mathbf{y}_{1:T}, \mathbf{h}_{0:T}^{(i+1)}, \phi^{(i)}, \mu^{(i)})$;
4. Amostramos $\phi^{(i+1)}$ de $p(\phi \mid \mathbf{h}_{0:T}^{(i+1)}, \sigma_\eta^{2(i+1)}, \mu^{(i)})$;

5. Amostramos $\alpha^{(i+1)}$ de $p(\alpha \mid \mathbf{h}_{0:T}^{(i+1)}, \phi^{(i+1)}, \sigma_\eta^{2(i+1)})$;
6. Fazemos $i = i + 1$ e repetimos 2 até convergir.

Aqui denotamos por $h_{/t}$ o vetor de estados no qual o estado no tempo t não é considerado. O índice i indica a i -ésima iteração do termo no algoritmo. O ciclo de passos 2 a 5 é chamado "sweep". [Kim et al. \(1998\)](#), nota que o problema é a amostragem de $h_t \mid h_{/t}, y_t, \boldsymbol{\theta}$, pois em cada iteração o passo 2 tem que realizar T amostras, até completar todo o vetor \mathbf{h}_t , além de dar diferentes soluções alternativas, onde $f(h_t \mid h_{/t}, y_t, \boldsymbol{\theta})$ é limitada por uma distribuição normal de $f(h_t \mid h_{/t}, \boldsymbol{\theta}) = f(h_t \mid h_{t-1}, h_{t+1}, \boldsymbol{\theta}) = f(h_t \mid h_{t^*}, v^2)$. Assim, o método de aceitação- rejeição pode ser implementado para os h_t (veja [Kim et al. \(1998\)](#)).

2.1.2 Mixture Sampler

A metodologia do "mixture sampler" é usada por Jong e Shephard (1995) e [Kim et al. \(1998\)](#), considerando a transformação

$$y_t^* = h_t + \log(\epsilon_t^2), \quad (2.9)$$

onde $y_t^* = \log(y_t^2 + c)$, $\log(\epsilon_t^2) = z_t$, e

$$f(z_t) = \sum_{i=1}^k q_i f_N(z_t \mid m_i - 1.2704, v_t^2), \quad (2.10)$$

ou seja, temos uma aproximação por uma mistura de k funções de densidade normais (f_N) com pesos (probabilidades) q_i . Fuller em 1996 foi o primeiro a introduzir aquelas misturas (2.10) no contexto da volatilidade estocástica, procurando fazer mais robusta a estimativa dada pela função de quase-verosimilhança(QML) para y^2 . Hierarquicamente, o modelo pode ser escrito em função da variável indicadora $s_t = i$ para $i = 1, \dots, k$, onde $Pr(s_t = i) = q_i$, $\sum_1^k q_i = 1$, como

$$z_t \mid s_t = i \sim N(m_i - 1.2704, v_t^2), \quad (2.11a)$$

$$Pr(s_t = i) = q_i. \quad (2.11b)$$

[Kim et al. \(1998\)](#) propõe que o número adequado para k seja 7, que indica a quantidade de misturas normais. O algoritmo geral foi proposto simultânea e independentemente por [Harvey et al. \(1994\)](#) e Carter e Kohn (1994). [Omori et al. \(2007\)](#) estende o algoritmo para

uma composição de $k = 10$ normais e prova que aproximação para a log normal é muito mais adequada. Denotamos a densidade a posteriori de interesse como $\pi(s, h, \phi, \sigma_\eta^2, \mu)$ e $s = (s_1, \dots, s_n)$. Neste caso, s e h são obtidos independentemente em blocos, utilizando o seguinte algoritmo:

Algoritmo 2

1. Inicializamos ($i = 0$) os valores de $s, \phi, \sigma_\eta^2, \mu$ e h ;
2. Amostramos \mathbf{h}_t de $\mathbf{h}_{1:n}/y_t^*, s, \phi, \sigma_\eta^2, \mu$, $t = 1, \dots, n$;
3. Amostramos s de $p(s | y_t^*, h_t)$;
4. Amostramos σ_η^2 de $p(\sigma_\eta^2 | \mathbf{y}_{1:T}, \mathbf{h}_{0:T}, \phi, \mu)$;
4. Amostramos ϕ de $p(\phi | \mathbf{h}_{0:T}, \sigma_\eta^2, \mu)$;
5. Amostramos μ de $p(\mu | \mathbf{h}_{0:T}, \phi, \sigma_\eta^2)$;
6. Fazemos $t = t + 1$ e repetimos o passo 2 até que se obtenha convergência.

Kim *et al.* (1998) ressalta que neste método comparado com o anterior (single-move) é mais fácil de amostrar de forma eficiente pois estamos diante de uma distribuição gaussiana multivariada, dado que $h_t | s, \phi, \sigma_\eta^2, \mu$ é um processo gaussiano, e $y_t^* | s, \phi, \sigma_\eta^2, \mu$ é uma série temporal gaussiana de alta dimensão. Assim, pode-se utilizar o Filtro de Kalman, pois obtemos um modelo na forma de espaços de estados. A literatura chama a estes modelos de Parcialmente Gaussianos ou Condicionalmente Gaussianos.

2.1.3 Multi-move Sampler

Com a finalidade de reduzir o problema apresentado pelo método de "single-move", Shephard e Pitt (1997) propuseram o "multi-move sampler", para amostrar as variáveis latentes (variáveis de estado) da densidade a posteriori, dados os parâmetros, pois que o modelo VE pode ser considerado como modelo de espaço de estados não gaussiano e não linear. O método consiste em dividir em blocos representativos ou significativos, e amostrar todo o bloco em cada passo, a cada tempo t , baseados na variação do algoritmos de Metropolis Hastings proposto por Tierney (1994). O método permite uma rápida convergência dos parâmetros do modelo, no entanto apresenta pequeno viés nas estimativas dos mesmos ao

aumentar o número de blocos, como menciona [Watanabe e Omori \(2004\)](#), que fazem uma proposta de melhorar a técnica "multi-move". Além disso, [Watanabe e Omori \(2004\)](#) ressaltam que na amostragem do "multi-move", assume-se que a equação do estados é não linear, e que o vetor de observação e vetor de estados são condicionalmente independentes. Com estas suposições ficam de fora importantes modelos que consideram correlação entre ambas as equações, por exemplo os modelos com efeito alavanca. Precisamente este caso é tratado por [Omori et al. \(2007\)](#), que estende o método de amostragem por mistura de normais de [Kim et al. \(1998\)](#). Estes autores aproximam o modelo não gaussiano utilizando uma distribuição normal bivariada, condicionado ao sinal das variáveis dependentes observadas.

O "multimove sampler" se baseia na aproximação da verdadeira densidade de um bloco de erros (dos estados) dados os parâmetros e os outros blocos, associados por um processo gaussiano.

Seja a distribuição condicional dos distúrbios $\{\eta_s\}_{s=t}^{t+k}$, dados $\xi = (\alpha_{t-1}, \dots, \alpha_{t+k+1}, y, \theta)$, ou seja a log-verossimilhança,

$$\ln f(\{\eta_s\}_{s=t}^{t+k} | \xi) = \text{const} + \ln f(\{\eta_s\}_{s=t}^{t+k}) + \ln f(\{y_s\}_{s=t}^{t+k} | \{\alpha_s\}_{s=t}^{t+k}) \quad (2.12)$$

$$\ln f(\{\eta_s\}_{s=t}^{t+k} | \xi) = \text{const} - \frac{1}{2\sigma_\eta^2} \sum_{s=t}^{t+k} \eta_s^2 + \sum_{s=t}^{t+k} l(\alpha_s), \quad (2.13)$$

onde

$$l(\alpha) = \ln f(y_s | \alpha_s, \beta) = \text{const} - \frac{1}{2}\sigma_s - \frac{y_t^2}{2} \exp(-\alpha_s). \quad (2.14)$$

Aplicando a aproximação da expansão de Taylor para o logaritmo da verossimilhança $\ln f(\{\eta_s\}_{s=t}^{t+k} | \xi) = \ln(g)$ próximo de alguma estimativa preliminar $\{\hat{\eta}_s\}_{s=t}^{t+k}$, temos,

$$\ln f(\{\eta_s\}_{s=t}^{t+k} | \xi) = \text{const} - \frac{1}{2\sigma_\eta^2} \sum_{s=t}^{t+k} \eta_s^2 + \sum_{s=t}^{t+k} \left\{ l(\hat{\alpha}_s) + (\alpha_s - \hat{\alpha}_s)l'(\hat{\alpha}_s) + \frac{1}{2}(\alpha_s - \hat{\alpha}_s)l''(\hat{\alpha}_s) \right\}, \quad (2.15)$$

onde $\{\hat{\alpha}_s\}_{s=t}^{t+k}$ é a estimativa correspondente para $\{\hat{\eta}_s\}_{s=t}^{t+k}$ e l' e l'' são as primeira e segunda derivadas de $l(\alpha_s)$ com respeito a α_s .

Então aproxima-se a densidade normal de $f^* = \log f(\eta_s, \dots, \eta_{s+m-1} | \alpha_s, \alpha_s, y_s, \dots, y_{s+m})$ como

segue,

$$\begin{aligned}
& \log f(\eta_s, \dots, \eta_{s+m-1} \mid \alpha_s, \alpha_s, y_s, \dots, y_{s+m}) & (2.16) \\
\approx \text{const} - \frac{1}{2} \sum_{t=s}^{s+m-1} \eta'_t \eta_t + \hat{L} + \frac{\partial L}{\partial \eta'} \Big|_{\eta=\hat{\eta}} (\eta - \hat{\eta}) + \frac{1}{2} (\eta - \hat{\eta})' E \left(\frac{\partial L^2}{\partial \eta \partial \eta'} \right) \Big|_{\eta=\hat{\eta}} (\eta - \hat{\eta}) \\
& = \text{const} - \frac{1}{2} \sum_{t=s}^{s+m-1} \eta'_t \eta_t + \hat{L} + \hat{d}' (\alpha - \hat{\alpha}) + \frac{1}{2} (\alpha - \hat{\alpha})' \hat{Q} (\eta - \hat{\eta}) \\
& = \text{const} + \log f^*(\eta_s, \dots, \eta_{s+m-1} \mid \alpha_s, \alpha_s, y_s, \dots, y_{s+m})
\end{aligned}$$

Omori *et al.* (2007) obtiveram estimativas $\hat{\eta}$ da densidade posteriori de η .

2.1.4 Produto de misturas gaussianas

Outro método destacado é o produto de misturas gaussianas. Sejam $p_1(x), \dots, p_d(x)$ d misturas de densidades normais, ou seja,

$$p_i(x) = \sum_{l_i} w_i N(x; \mu_{l_i}, \lambda_i), \quad i = 1, \dots, d, \quad (2.17)$$

onde l_i denota o conjunto de índices das componentes densidades normais de p_i com média μ_{l_i} , variância λ_i e w_{l_i} são os pesos dos componentes. Por simplicidade assumimos que as misturas de normais são de igual quantidade de somandos n , e que as variâncias λ_i são as mesmas em cada mistura. O objetivo principal é amostrar eficientemente da densidade gerada pela mistura de densidades $p(x) \propto \prod_{i=1}^d p_i(x)$ de dimensão n^d .

Esta amostragem pode ser descomposta em dois passos, primeiro selecionamos um dos produtos de densidades de n^d componentes, depois amostramos para cada densidade correspondente.

2.1.5 Aplicação

Para observar o desempenho dos método MCMC nos modelos de volatilidade estocástica, realizamos um amostragem “single-move” usando os 2022 retornos diários obtidos da série S&P500 para o período de janeiro de 1980 até dezembro de 1987, utilizados por Jun (2000), os mesmos que serão utilizados em todos os exemplos durante esta tese. Foram consideradas como prioris as propostas por Kim *et al.* (1998), Jacquier *et al.* (2004), Omori *et al.* (2007) e Shephard (1996), entre outros. Assim temos $\mu \sim \mathcal{N}(0, 25)$, por motivos computacionais se considerou a transformação $\phi^* = \frac{\phi+1}{2}$ na qual $\phi^* \sim Be(20, 1.5)$ (distribuição Beta de parametros 20 e 1.5), $\alpha = \mu(1 - \phi)$ e finalmente $\sigma_\eta^2 \sim \mathcal{IG}(2.5, 0.025)$.

Resultados descritivos da série mostram que o valor da curtose é muito maior do que o valor correspondente à distribuição normal (> 3), isto é um indicativo de que supor normalidade não é muito adequado. O gráfico dos retornos da série *S&P500* é dado na Figura 2.1. Podemos observar a grande concentração de valores em torno da média. Na tabela 2.1 apresentamos alguns resultados de estatísticas descritivas correspondentes à série de retornos.

Tabela 2.1: Estatísticas para a série de retornos da S&P500

	média	s.d.	max	min	assimetria	curtose
Retornos	0.000	1.03	8.247	-7.35	-2.553	7.2629

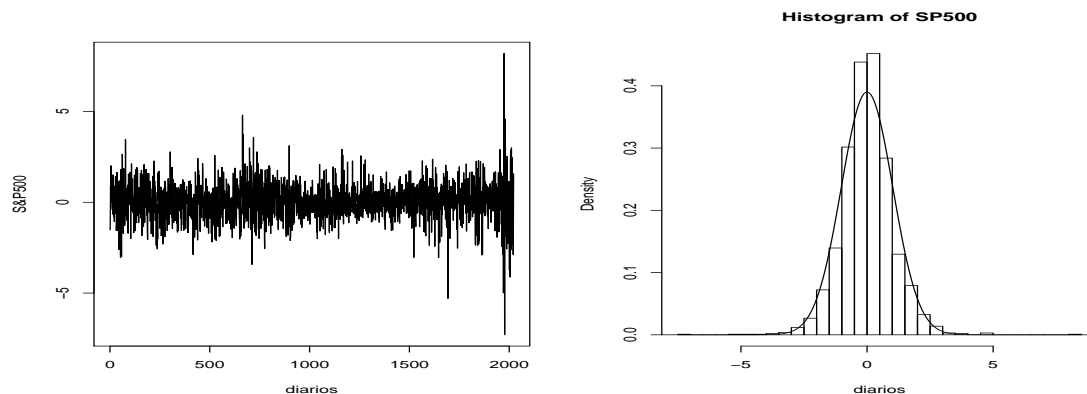


Figura 2.1: Gráfico correspondente aos retornos da série S&P500 para o período de janeiro 1980 até dezembro 1987. Esquerda: série. Direita: Histograma de retornos.

Implementamos a metodologia do amostrador de Gibbs, em particular a metodologia do “single move” e consideramos os modelos de volatilidade estocástica com distribuição normal e t-Student para o ruído. Neste caso foram realizadas 100.000 simulações, no entanto, as primeiras 5.000 foram consideradas no período de “burn-in”, depois foi selecionada uma amostra sistemática de 25 em 25 para todos os parâmetros em ambos os modelos. Para 38.000 simulações selecionadas de ambos modelos, verificamos que a função de autocorrelação decresce para zero. Baseados na amostra, calculamos a média a posteriori e algumas outras estatísticas como o intervalo de 95% de credibilidade, o desvio padrão ou erro Monte Carlo da média posteriori e a estatística de convergência CD de Geweke (1992) (veja a Tabela 2.2).

Apresentamos nas figuras 2.2 e 2.3 as simulações resultantes das cadeias de Markov para o modelo VE-N e VE-T; nelas pode-se notar as características mencionadas acima sobre o

Tabela 2.2: Estimativas para a série de retornos S&P500. Primeira linha: média posteriori. Segunda linha: intervalo de credibilidade de 95%. Terceira linha: Erro Monte Carlo da média posteriori. Quarta linha: estatística CD

Parameter	VE-N	VE-Tstudent
α	-0,00335	-0,00145
	(-0,0109,0,0027)	(-0,005924, 0,003165)
	$0,34 \times 10^{-4}$	$0,2335 \times 10^{-2}$
	-0,09435	0,4066
ϕ	0,9748	0,9838
	(0,9577,0,9904)	(0,9699, 0,9956)
	$3,04 \times 10^{-4}$	$0,6845 \times 10^{-2}$
	-0,886	0,86
σ^2	0,1492	0,934
	(0,1058,0,1798)	(0,0664,0,1228)
	$0,19 \times 10^{-1}$	$0,149 \times 10^{-2}$
	1,79	-0,85
ν	—	9,374
	—	(6,001,13,2)
	—	0,194
	—	1,055

comportamento das simulações. A convergência do parâmetro ν para o valor de 9,374 nos faz pensar que considerar o ruído com distribuição $t - Student$ é relevante.

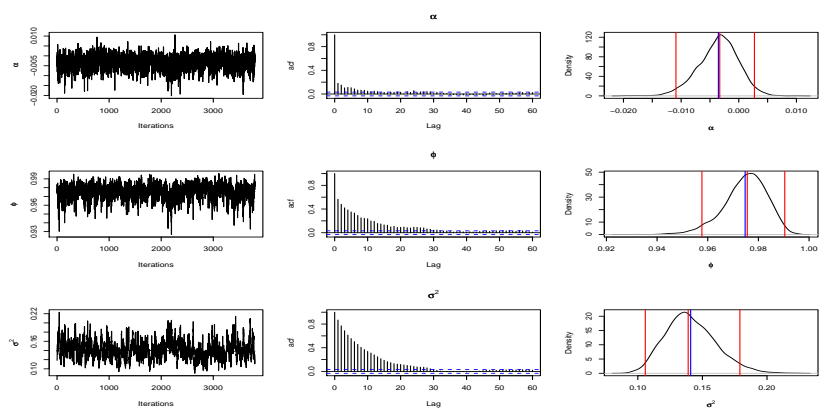


Figura 2.2: Comportamento das Simulações dos parâmetros α , ϕ e σ obtidas no método MCMC, para a série de retornos diários da S&P500 (modelo VE-N). Simulações (esquerda), função de autocorrelação das iterações (centro), densidade posteriori (direita), as linhas verticais indicam o intervalo de maior credibilidade de nível 5% e as linhas centrais a média e mediana a posteriori.

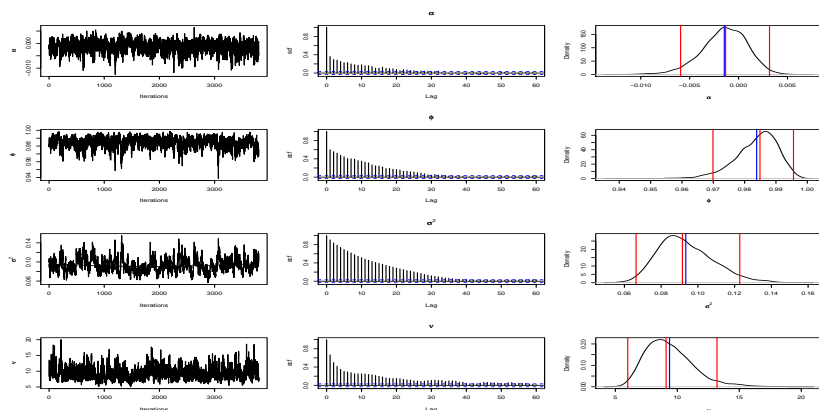


Figura 2.3: Comportamento das Simulações dos parâmetros α , ϕ , σ e ν obtidas ao implementar o amostrador de Gibbs para a série de retornos diários da S&P500 (modelo VET-Student). Simulações (esquerda), função de autocorrelação das iterações (centro), densidade posteriori (direita), as linhas verticais indicam o intervalo de maior credibilidade de nível 5% e as linhas centrais a média e mediana a posteriori.

2.2 Modelos de Volatilidade Estocástica com Efeito Alavanca

Existem evidências de que as séries dos retornos financeiros respondam assimetricamente a choques positivos e negativos, acarretando em uma correlação negativa entre a série e a volatilidade. Do ponto de vista dos modelos GARCH foram incorporadas ao modelo algumas alternativas para solucionar o problema, assim surgem os modelos EGARCH e TGARCH por exemplo. Entre os modelos de volatilidade estocástica VE, estão os Modelos com Leverage Effect, Modelos Fatoriais, Modelos VE elípticos ou de caudas pesadas, e os que levam em conta a correlação dos estados e a série variando no tempo.

A fim de generalizar a equação de Black-Scholes para preços de ativos financeiros, [Harvey e Shephard \(1996\)](#) propõem o modelo de VE univariado de tempo discreto com leverage effect, [Wiggins \(1987\)](#) e [Chesney et al. \(2004\)](#) entre outros, consideraram o modelo em tempo contínuo. O modelo de [Harvey e Shephard \(1996\)](#) pode ser escrito:

$$y_t = \sigma_t \epsilon_t, \quad t = 1, 2, \dots, n, \quad (2.18a)$$

$$\ln \sigma_{t+1}^2 = \alpha + \phi \ln \sigma_t^2 + \sigma_\eta \eta_t, \quad t = 1, 2, \dots, n-1, \quad (2.18b)$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right), \quad (2.18c)$$

onde $|\phi| < 1$, σ_t é considerada a volatilidade da resposta y_t . A correlação entre os erros é considerada significativa e negativa a fim de explicar o efeito de alavancagem. A queda dos retornos é seguida de um elevado valor de volatilidade, que é a definição de "efeito alavanca" ([Black \(1976\)](#) e [Nelson \(1991\)](#)).

[Jun \(2005\)](#) faz uma análise comparativa entre os modelos de volatilidade estocástica propostos até então, o modelo de aproximação do Euler, mais conhecido para tempos contínuos, e o modelo de VE com efeito alavanca em tempos discretos de [Jacquier et al. \(2004\)](#), e chega a conclusão de que tanto no nível de interpretação como no da conservação das propriedades do martingale, o modelo proposto em [Jacquier et al. \(2004\)](#) é inferior. Alguns autores tratam de simular este modelo mediante a metodologia da amostragem "single move" para os estados, que considera encontrar um h_t em cada tempo t , as outras variáveis de estado dadas, o que implica que a amostragem baseada no MCMC seja muito lenta, havendo a necessidade de iterar um grande número de vezes até a convergência dos parâmetros, devido a alta correlação que apresentam as variáveis latentes.

Omori e Watanabe (2008), propõem uma rápida e eficiente metodologia "mixture sampler" de estimação dos estados e mostra que a performance de estimação é bem melhor e mais eficiente que a amostragem single move. No entanto, o método é só aplicável a classe de modelos (2.18a) e (2.18b), desconsiderando características nos retornos como o efeito alavanca.

Os modelos de caudas pesadas são considerados por Kim *et al.* (1998), em que os modelos de VE podem ser estendidos para permitir que ε_t em (2.1a) assumam distribuições de caudas mais pesadas que a distribuição normal. Segundo Harvey *et al.* (1994) entre outros, pode-se obter valores da distribuição $t - Student$ padronizada como,

$$\varepsilon_t = \frac{\sqrt{\frac{v-2}{v}}\epsilon_t}{\sqrt{\frac{\chi_{1,v}^2}{v}}} \quad (2.19)$$

$$\epsilon_t \stackrel{iid}{\sim} N(0, 1), \chi_{1,v}^2 \stackrel{iid}{\sim} \chi_v^2, \quad (2.20)$$

onde χ_v^2 denota a v.a *qui*-quadrado e ϵ_t e χ_v^2 são independentes entre si. Kim *et al.* (1998) analisam a alternativa de estimar os parâmetros deste modelo utilizando o método de single move e o "offset mixture", incorporando a amostragem Gibbs da nova variável latente.

2.2.1 Aplicação

Como exemplo vamos considerar a série S&P500, e por motivos comparativos o período da análise foi de 5 janeiro 1980 até dezembro do 1987 (2022 observações), dados considerados em Jun (2005).

Para exemplificar o modelo de volatilidade estocástica com efeito alavanca (VE-LE) implementamos os modelos VELE-N e VELE-T. Deste modo foi desenvolvido o algoritmo do Gibbs em MCMC, e consideramos novamente como distribuições a priori, $\mu \sim \mathcal{N}(0, 25)$, onde $\alpha = \mu(1 - \phi)$, $\phi^* \sim \text{Be}(20; 1, 5)$, onde $\phi^* = \frac{\phi+1}{2}$ e $\sigma_\eta^2 \sim \mathcal{IG}(2, 5; 0, 025)$. Além disso, assumimos que a distribuição a priori para o parâmetro de correlação é $\rho \sim \mathcal{U}(-1, 1)$.

A metodologia utilizada para realizar a simulação dos parâmetros é o amostrador de Gibbs considerando um "single move" descrito na seção 3.3.1. Note que no caso de modelo V.E. com efeito alavanca o vetor de parâmetros θ é $(\alpha, \phi, \sigma, \rho)$, logo o algoritmo Gibbs fica:

Algoritmo

1. Inicializamos ($i = 0$) os valores de $\boldsymbol{\theta}^{(i)}$ e $h^{(i)}$;
2. Amostramos $\mathbf{h}_t^{(i+1)}$ de $\mathbf{h}_{1:T}^{(i+1)}/h_{/t}^{(i+1)}, \mathbf{y}_t, \boldsymbol{\theta}^{(i)}$, $t = 1, \dots, T$;
3. Amostramos $\sigma_\eta^{2(i+1)}$ de $p(\sigma_\eta^2 | \mathbf{y}_{1:T}, \mathbf{h}_{0:T}^{(i+1)}, \phi^{(i)}, \mu^{(i)})$;
4. Amostramos $\phi^{(i+1)}$ de $p(\phi | \mathbf{h}_{0:T}^{(i+1)}, \sigma_\eta^{2(i+1)}, \mu^{(i)})$;
5. Amostramos $\mu^{(i+1)}$ de $p(\mu | \mathbf{h}_{0:T}^{(i+1)}, \phi^{(i+1)}, \sigma_\eta^{2(i+1)})$;
5. Amostramos $\rho^{(i+1)}$ de $p(\rho | \mathbf{h}_{0:T}^{(i+1)}, \phi^{(i+1)}, \sigma_\eta^{2(i+1)})$;
6. Fazemos $i = i + 1$ e repetimos os passos de 2 a 6 até convergir.

Aqui denotamos como $h_{/t}$ o vetor de estados no qual o estado no tempo t não é considerado. O índice i indica a i -ésima iteração do termo no algoritmo.

Realizamos simulações para estimar os parâmetros do modelo, destacamos que devido à alta correlação apresentada pelos valores gerados nas cadeias dos parâmetros, tomamos uma amostra sistemática de 25 em 25, uma vez ordenados os dados, por isso, das 95000 simulações, nós consideramos 3800. Baseados na amostra calculamos a média a posteriori e seu respectivo desvio padrão, assim como o intervalo de credibilidade ao nível de 95% de confiança, resultados que apresentamos na Tabela 2.3. A estimação dos parâmetros $\alpha, \phi, \sigma, \rho$, resultaram coerentes com respeito aos valores encontrados na literatura (0,00, 0,98, 0,10, -0,3).

Neste exemplo para o modelo VELE-N (volatilidade estocástica com efeito alavanca e erro normal) temos que $\alpha, \phi, \sigma, \rho$ resultam -0,002263467 (dp=0,00360185), 0,9716 (dp=0,00981), 0,1513087 (dp=0,02124793) e -0,3210059 (d.p=0,08444381). Para o modelo VELE-t (volatilidade estocástica com erro t-Student e efeito alavanca) os resultados dos parâmetros $\alpha, \phi, \sigma, \rho, v$ e seus desvios padrão são 0,000578 (dp=0,0023), 0,98246 (dp=0,0017), 0,098 (dp=0,016), -0,287 (dp=0,08) e 9,047 (dp=0,181). A estatística de convergência de Geweke (1992) foi calculada, resultando que todos os parâmetros convergem ao 95% de confiança em ambos os modelos.

Nas figuras 2.4 e 2.5 observamos o desempenho das cadeias de Markov para cada parâmetro nos modelos VELE-N e VELE-T para os retornos da série S&P500. Pode-se notar que a função de autocorrelação decresce para zero.

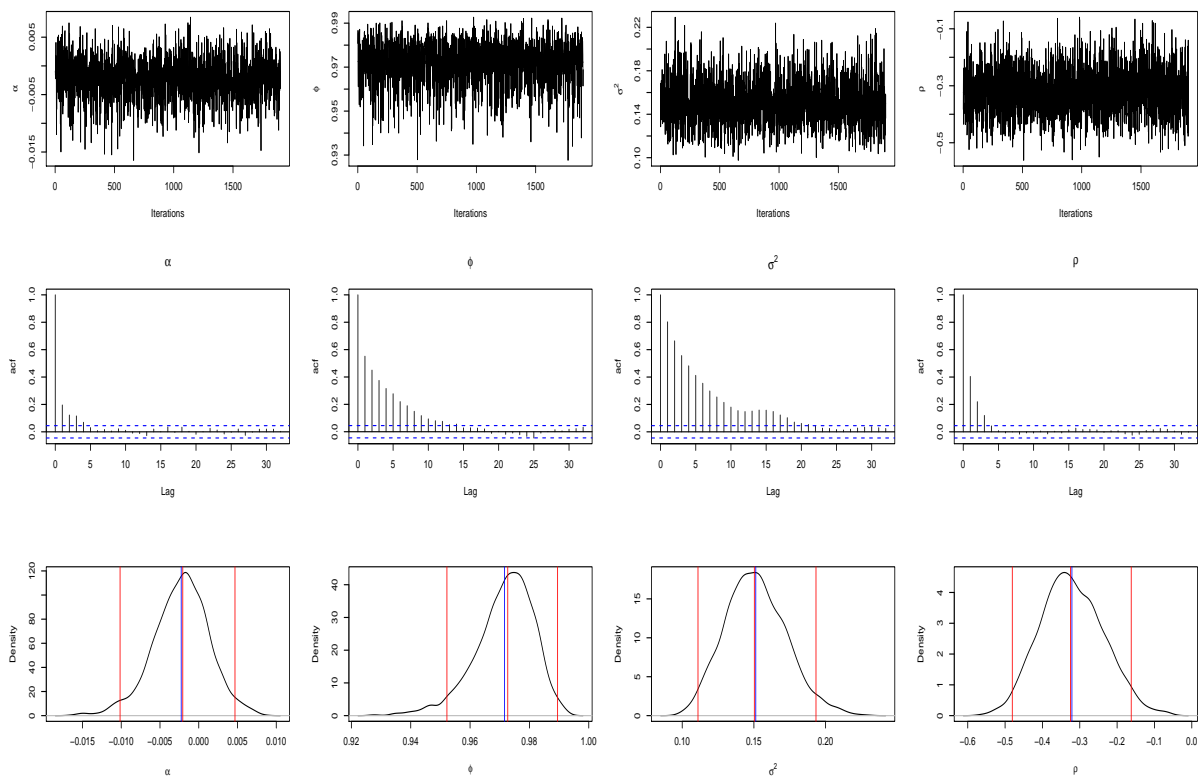


Figura 2.4: Série de iterações para α , ϕ , σ^2 e ρ dos resultados do MCMC para a série de retornos diários S&P500 (VE- Normal com Efeito Alavanca). Iterações amostradas (primeira linha), autocorrelações amostrais (segunda linha), densidade posteriori (terceira linha), As linhas indicam os percentiles 2,5% e 97,5% e as linhas do meio a média posteriori e mediana.

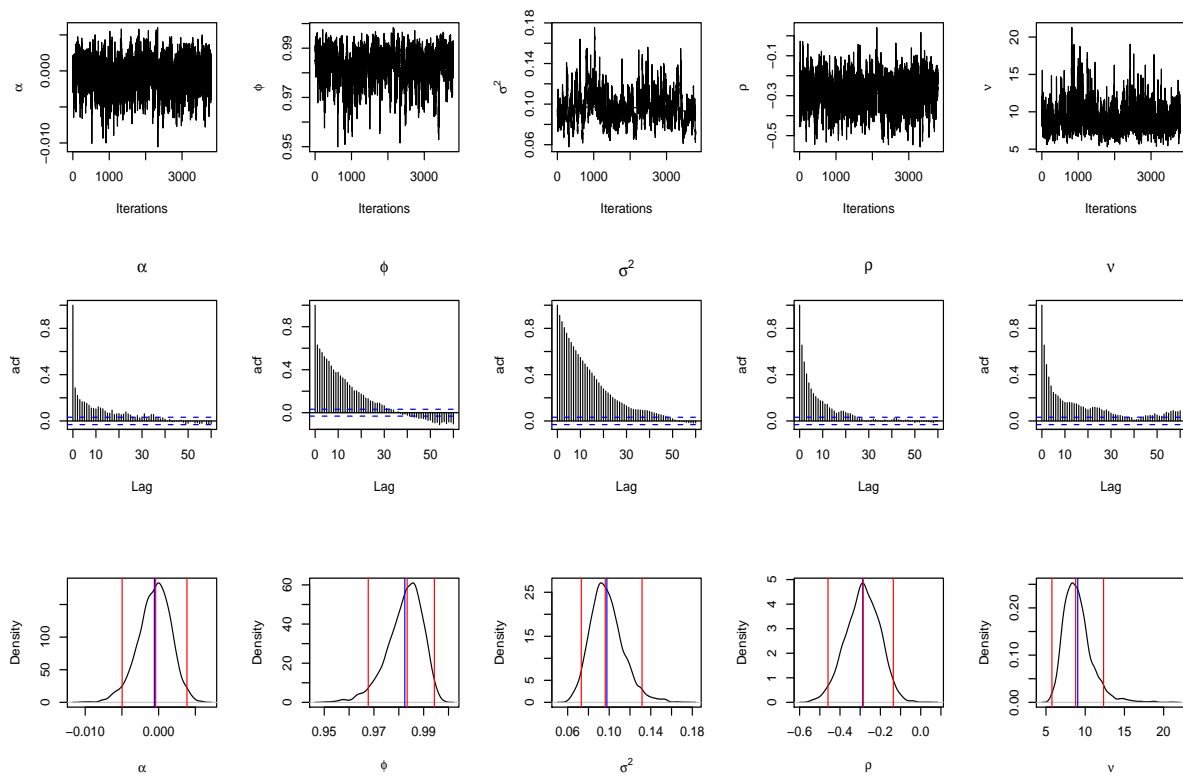


Figura 2.5: Série de iterações para α e ϕ dos resultados do MCMC para a série de retornos diários S&P500 (VE- T-student com Efeito Alavanca). Iterações amostradas (primeira linha), autocorrelações amostrais (segunda linha), densidade posteriori (terceira linha), As linhas indicam os percentiles 2,5% e 97,5% e as linhas do meio a média posteriori e mediana.

Tabela 2.3: Estimativas resultantes para a série de retornos S&P500. Primeira linha: média posteriori. Segunda linha: intervalo de credibilidade posterior de 95% entre parêntesis. Terceira linha: Erro Monte Carlo erro da média a posteriori. Quarta linha: Estatística de convergência CD

Parameter	VELE-N	VELE-T
α	-0,002263467	0,000578
	(-0,0102; 0,0047)	(-0,0050; 0,00386)
	$0,360185 \times 10^{-2}$	$0,223 \times 10^{-2}$
	-0,01231	0,4231
ϕ	0,9716	0,9824
	(0,9522; 0,9895)	(0,9677; 0,9943)
	$0,981 \times 10^{-2}$	$0,7128 \times 10^{-2}$
	-0,5233	0,9714
σ^2	0,1513087	0,0977
	(0,111; 0,1933)	(0,07031; 0,1316)
	$0,2125 \times 10^{-1}$	$1,58 \times 10^{-2}$
	0,9998	-1,848
ρ	-0,3210059	-0,2875
	(-0,4807; -0,1622)	(-0,4594; -0,136)
	$0,8444381 \times 10^{-1}$	$8,37 \times 10^{-2}$
	-0,9379	1,117
ν	—	9,05
	—	(5,757, 12,34)
	—	0,1836
	—	-1,356

Capítulo 3

Modelo de Volatilidade Estocástica com Caudas Pesadas

Várias generalizações de modelos que substituem a normal padrão por outras distribuições têm sido propostas. Dentre os modelos de VE que permitem a presença de distribuições para os distúrbios com caudas mais pesadas que a normal, pode-se citar, [Mandelbrot \(1963\)](#), [Fama \(1965\)](#), [Liesenfeld e Jung \(2000\)](#), [Chib *et al.* \(2002\)](#), [Jacquier *et al.* \(2004\)](#) e [Cathy *et al.* \(2008\)](#).

Neste contexto o modelo de volatilidade estocástica com erro t-Student (VE-t), tem sido o mais estudado ou ao menos mencionado como o modelo que considera caudas mais pesadas nos retornos. Estendemos o modelo VE-t para uma classe flexível de modelos de distribuições com misturas de normais na escala (MEN, [Andrews e Mallows \(1974\)](#), [Lange e Sinsheimer \(1993\)](#), [Fernandez e Steel \(2000\)](#), [Chow e Chan \(2008\)](#)), uma vez que esta classe considera como elementos próprios a distribuição normal (VE-N), a Student-t (VE-t), a slash (VE-S) e a distribuição "variance gamma" (VE-VG). Todas estas distribuições têm caudas mais pesadas do que a normal e podem ser usadas como modelos que proporcionam inferências mais robustas neste tipo de modelos. Denotamos estes modelos por VE-MEN.

3.1 Distribuições de contorno elíptico

Dizemos que o vetor aleatório m-dimensional Y tem distribuição elíptica multivariada com vetor de posição $\mu \in R^m$ e matriz de escala positiva definida Λ se sua função de densidade assume a forma

$$f(y) = \|\Lambda\|^{-1/2} g \left[(y - \mu)^T \Lambda^{-1} (y - \mu) \right], \quad (3.1)$$

em que $g : \Re \rightarrow [0; 1)$ satisfaz a condição $\int_0^\infty u^{m/2-1}g(u)du < \infty$. A função g é conhecida como função geradora de densidade. Quando um vetor aleatório tem densidade dada por (3.1) usamos a notação $Y \sim EC_m(\mu; \Lambda; g)$.

Um estudo completo das propriedades da classe de distribuições elípticas assim como alguns exemplos podem ser encontrados em Fang, Kotz, e Zhang (1990b), Fang, Kotz, e Ng (1990a) e Azzalini e Valle (1996). A seguir apresentamos uma propriedade que estabelece que a classe de distribuições de contornos elípticos é fechada sob transformações lineares.

Proposição 2.1 Seja $\mathbf{Y} \sim EC_m(\mu; \Lambda; g)$. Se $\mathbf{a} \in R^n$ e \mathbf{A} é uma matriz $n \times m$ com posto $(\mathbf{A}) = n \leq m$, então

$$X = \mathbf{A}Y + a \sim EC_n(\mathbf{A}\mu + a; \mathbf{A}\Lambda\mathbf{A}^T; g). \quad (3.2)$$

A partir de (3.2) podemos notar que qualquer distribuição marginal de $\mathbf{Y} \sim EC_m(\mu; \Lambda; g)$ também é elíptica.

3.2 Distribuições de misturas de normais na escala (MEN)

As distribuições de mistura na escala de normais desempenham um papel importante na modelagem estatística, elas são derivadas da mistura de uma variável aleatória com distribuição normal (Z) com uma variável aleatória não negativa (λ), como

$$Y = \mu + \kappa^{1/2}(\lambda)Z,$$

onde μ é o parâmetro de localização, λ é uma variável aleatória positiva com função de densidade $h(\lambda|\boldsymbol{\nu})$, independente de $Z \sim \mathcal{N}(0, \sigma^2)$, onde $\boldsymbol{\nu}$ é um parâmetro escalar ou parâmetro vetorial que indexa a distribuição de λ e $\kappa(\cdot)$ é a função de ponderação. Como em Lange e Sinsheimer (1993) e Chow e Chan (2008), neste trabalho enfocamos o caso em que $\kappa(\lambda) = 1/\lambda$. Isto é, dado λ , $Y|\lambda \sim \mathcal{N}(\mu, \lambda^{-1}\sigma^2)$ e a fdp de Y é dada por

$$f(y|\mu, \sigma^2, \boldsymbol{\nu}) = \int_0^\infty \mathcal{N}(y|\mu, \lambda^{-1}\sigma^2)h(\lambda|\boldsymbol{\nu})d\lambda. \quad (3.3)$$

A escolha da densidade de mistura $h(\cdot|\boldsymbol{\nu})$, uma classe de distribuições contínuas simétricas unimodais descritas pela densidade dada em (3.3), é feita de tal forma que se possa modelar séries com caudas pesadas.

3.2.1 Definição

Seja X variável aleatória contínua com parâmetro de localização μ e parâmetro de escala σ . A v.a X se diz ter distribuição de mistura de normais na escala, se sua função de densidade de probabilidade puder ser expressa como,

$$f_x(x | \mu, \sigma, \lambda) = \int_0^\infty N(x | \mu, \kappa(\lambda), \sigma^2) h(\lambda) d(\lambda), \quad (3.4)$$

onde $\mathcal{N}(x | \cdot, \cdot)$ é a densidade da distribuição normal, $\kappa(\lambda)$ é uma função positiva de λ , e $h(\cdot)$ é uma função de densidade definida em \mathfrak{R}^+ . Nos referimos a λ e $h(\cdot)$ como o parâmetro de mistura e densidade de mistura, respectivamente.

Alguns exemplos são dados a seguir.

- A distribuição *t-Student*, $Y \sim t(\mu, \sigma^2, \nu)$

Usar a distribuição t como uma alternativa à distribuição normal, têm sido frequentemente sugerido na literatura, por exemplo, [Little \(1988\)](#) e [Lange et al. \(1989\)](#) utilizam a distribuição para modelos robustos. A fdp da distribuição t- Student com parâmetro de localização μ , escala σ e graus de liberdade ν , pode ser expressa como

$$f(y|\mu, \sigma, \nu) = \int_0^\infty N\left(y|\mu, \frac{\sigma^2}{\lambda}\right) \mathcal{G}\left(\lambda\left|\frac{\nu}{2}, \frac{\nu}{2}\right.\right) d\lambda. \quad (3.5)$$

onde $\mathcal{G}(\cdot|a, b)$ é a função de densidade gama, dada por

$$\mathcal{G}(\lambda|a, b) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda), \quad \lambda, a, b > 0, \quad (3.6)$$

e $\Gamma(a)$ é a função gama com argumento $a > 0$. Isto é, se $Y \sim t_\nu(\mu, \sigma)$, temos a forma hierárquica:

$$Y|\mu, \sigma^2, \nu, \lambda \sim N\left(\mu, \frac{\sigma^2}{\lambda}\right), \quad \lambda|\nu \sim \mathcal{G}(\nu/2, \nu/2). \quad (3.7)$$

Neste caso $\kappa(\lambda) = \frac{1}{\lambda}$ e $\lambda \sim \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$. $h(\cdot)$ é $\mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$ com momento finito,

$$E[\lambda^{-m}] = \frac{(\frac{\nu}{2})^m \Gamma(\frac{\nu}{2} - m)}{\Gamma(\frac{\nu}{2})}, \quad m < \frac{\nu}{2}. \quad (3.8)$$

- A distribuição Slash, $Y \sim S(\mu, \sigma^2, \nu)$, $\nu > 0$.

É uma distribuição elíptica com caudas mais pesadas do que a distribuição normal, que a inclui quando $\nu \uparrow \infty$. Sua fdp é dada por

$$f(y|\mu, \sigma, \nu) = \nu \int_0^1 \lambda^{\nu-1} N\left(y|\mu, \frac{\sigma^2}{\lambda}\right) d\lambda, \quad (3.9)$$

onde λ é Beta ($\mathcal{B}e(\nu, 1)$), com densidade

$$h(\lambda|\nu) = \nu \lambda^{\nu-1} \mathbb{I}_{(0,1)}. \quad (3.10)$$

Expressando a distribuição Slash de uma forma hierárquica temos,

$$Y|\mu, \sigma^2, \nu, \lambda \sim N\left(\mu, \frac{\sigma^2}{\lambda}\right), \quad \lambda|\nu \sim \mathcal{B}e(\nu, 1). \quad (3.11)$$

Neste caso $\kappa(\lambda) = \frac{1}{\lambda}$ e $\lambda \sim \mathcal{B}e(\nu, 1)$, com $h(u, \nu) = \nu u^{\nu-1} \mathbb{I}_{(0,1)}$, com momentos dados por

$$E[\lambda^{-m}] = \frac{\nu}{\nu - m}, \quad m < \nu. \quad (3.12)$$

A distribuição Slash tem sido principalmente usada em estudos de simulação porque representa uma situação extrema, veja [Andrews et al. \(1972\)](#), [Gross \(1973\)](#), e [Morgenthaler e Tukey \(1991\)](#).

- A distribuição Variância Gama, $Y \sim VG(\mu, \sigma^2, \nu)$, $\nu > 0$.

A distribuição variância gama (VG) foi proposta por Madan e Seneta (1990) para modelos de retornos de mercado. Esta distribuição é controlada por um parâmetro de forma $\nu > 0$ e apresenta caudas mais pesadas que a distribuição normal. Pode-se provar que tem como densidade

$$f(y|\mu, \sigma, \nu) = \int_0^\infty N\left(y|\mu, \frac{\sigma^2}{\lambda}\right) \mathcal{IG}\left(\lambda\left|\frac{\nu}{2}, \frac{\nu}{2}\right.\right) d\lambda. \quad (3.13)$$

Isto é, a distribuição VG é equivalente a seguinte forma hierárquica:

$$Y|\mu, \sigma^2, \nu, \lambda \sim N\left(\mu, \frac{\sigma^2}{\lambda}\right), \quad \lambda|\nu \sim \mathcal{IG}\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \quad (3.14)$$

onde $\mathcal{IG}(a, b)$ é a distribuição gama inversa

$$\mathcal{IG}(\lambda|a, b) = \frac{b^a}{\Gamma(a)} \lambda^{-(a+1)} \exp\left(-\frac{b}{\lambda}\right).$$

Se $\nu = 2$, a distribuição VG é a distribuição de Laplace. Temos que,

$$E[\lambda^{-m}] = \frac{(\frac{\nu}{2})^m \Gamma(\frac{\nu}{2} + m)}{\Gamma(\frac{\nu}{2})}, \quad m < \frac{\nu}{2}. \quad (3.15)$$

A seguir enfocamos esta classe de distribuições no contexto da volatilidade estocástica.

3.3 Modelos de volatilidade estocástica com erro na família MEN

Consideremos o modelo

$$y_t = \exp^{\frac{h_t}{2}} \epsilon_t, \quad (3.16a)$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \tau\eta_t, \quad (3.16b)$$

onde

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} | \lambda_t \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \kappa(\lambda_t) \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right). \quad (3.17a)$$

Vamos considerar o valor de $\kappa(\lambda_t) = [E(\lambda_t^{-1})]^{-1} \frac{1}{\sqrt{\lambda}}$, e verificar as diferentes distribuições que possam ser a densidade de mistura e por isso os diferentes valores de $[E(\lambda_t^{-1})]^{-1}$. Esta função é que indexa no modelo uma variável aleatória (λ_t) de mistura, com um novo parâmetro (v). Considerando, por exemplo,

$$u_t = \frac{\epsilon_t}{\sqrt{\lambda_t}} \sqrt{\kappa(\lambda_t)}, \quad (3.18)$$

onde $\epsilon_t \sim N(0, 1)$ e $\lambda_t(v) \sim Ga(\frac{v}{2}, \frac{v}{2})$, vemos que u_t têm distribuição t-Student.

O valor do fator $\sqrt{\kappa_t(\lambda)}$ que é incorporado na transformação, a fim de que a variável aleatória u_t tenha a propriedade de ser uma v.a com $E(u_t) = 0$ e $Var(u_t) = 1$, é deduzido como segue. Como $\epsilon_t \sim N(0, 1)$, temos os resultados seguintes:

$$u_t | \lambda_t \sim N(0; \kappa(\lambda_t)^2 \frac{1}{\lambda_t}), \quad (3.19)$$

$$E[u_t] = E_{\lambda_t} E(u_t | \lambda_t) = 0, \quad (3.20)$$

e substituindo no valor de u_t ,

$$E[u_t^2] = E_{\lambda_t} \left\{ E\left(\frac{\epsilon_t^2 \kappa(\lambda_t)}{\lambda_t}\right) \right\} = 1, \quad (3.21)$$

logo

$$\kappa(\lambda_t) E[\lambda_t^{-1}] = 1,$$

do que decorre

$$\kappa(\lambda_t) = E[\lambda_t^{-1}]^{-1}.$$

Para o caso da t-Student,

$$E[\lambda_t^{-1}] = \frac{\frac{v}{2}(\Gamma(\frac{v}{2} - 1))}{\Gamma(\frac{v}{2})} = \frac{v}{v - 2}. \quad (3.22)$$

Logo, se, $\lambda_t \sim Ga(\frac{v}{2}, \frac{v}{2})$ e $\kappa(\lambda_t) = \frac{1}{\lambda_t}(\frac{v-2}{v})$, então,

$$u_t \sim t_{(v)}. \quad (3.23)$$

Similarmente, para a distribuição Slash e Variância Gama temos o seguinte.

Se $\lambda_t \sim Beta(\nu, 1)$, $E[\lambda_t^{-m}] = \frac{\nu}{\nu - m}$ e $\kappa(\lambda) = (\frac{1}{\lambda_t})^{\frac{\nu+1}{\nu}}$ então,

$$u_t \sim Slash_{(\nu)}. \quad (3.24)$$

Para Variância Gama,

$$\lambda_t \sim IGa\left(\frac{\nu}{2}, \frac{\nu}{2}\right), E[\lambda_t^{-m}] = \frac{\frac{\nu}{2}^{-1} \Gamma(\frac{\nu}{2} + m)}{\Gamma(\frac{\nu}{2})} \text{ e } \kappa(\lambda) = \left(\frac{1}{\lambda_t}\right)^{\frac{\nu}{2} \Gamma(\frac{\nu}{2})}, \text{ então,}$$

$$u_t = \frac{\epsilon_t}{\sqrt{\lambda_t}} \sqrt{\frac{\frac{\nu}{2} \Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2}) + 1}} \sim V.G_{(\nu)}. \quad (3.25)$$

3.3.1 O modelo VE-MEN1:

As distribuições de misturas na escala de distribuições normais podem ser consideradas para os distúrbios da série e dos estados, que por serem da família exponencial cumprem as propriedades de um processo Markoviano. Neste primeiro caso, consideramos o erro com distribuição na escala normal, só na equação da série, permanecendo a equação de estados como um processo gaussiano, vamos denotar este modelo por VE-MEN1. Isto é,

$$y_t = e^{\frac{h_t}{2}} \epsilon_t, \quad (3.26a)$$

$$h_t = \alpha + \phi h_{t-1} + \sigma_\eta \eta_t, \quad (3.26b)$$

onde y_t e h_t são respectivamente os retornos e log-volatilidade no tempo t , $|\phi| < 1$ e $h_0 \sim N(\mu, \frac{\sigma_\eta^2}{1-\phi^2})$. Aqui, $\mu = \frac{\alpha}{1-\phi}$, que representa a média do processo estocástico da log-volatilidade. As inovações ϵ_t são distribuídas independentemente com distribuição de escala normal e η_t são assumidos mutuamente independentes normalmente distribuídos com média zero e variância unitária, assim como ϵ_t e η_t independentes entre si. Escrito na forma hierárquica,

$$y_t = e^{\frac{h_t}{2}} \lambda_t^{-\frac{1}{2}} \epsilon_t, \quad (3.27a)$$

$$h_t = \alpha + \phi h_{t-1} + \sigma_\eta \eta_t, \quad (3.27b)$$

$$\lambda_t \sim p(\lambda_t), \quad \epsilon_t \sim N(0, 1), \quad \eta_t \sim N(0, 1), \quad (3.27c)$$

onde $p(\lambda_t)$ é a distribuição de mistura de parâmetro ν , o vetor de parâmetros será $\boldsymbol{\theta} = (\alpha, \phi, \sigma^2, \nu)$.

A verossimilhança $P(\mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T})$, não tem forma fechada, de modo que utilizaremos o algoritmo Gibbs Sampler para simular os valores de interesse dos parâmetros.

Amostragem do modelo e estimação via MCMC

Para estimar os parâmetros nos modelos VE-MEN1 (equações (3.27a)-(3.27c)) usamos a técnica MCMC. Seja $\boldsymbol{\theta}$ o vetor de parâmetros, $\mathbf{h}_{0:T} = (h_0, h_1, \dots, h_T)'$ o vetor de log volatilidades, $\boldsymbol{\lambda}_{1:T} = (\lambda_1, \dots, \lambda_T)'$ as variáveis de mistura e $\mathbf{y}_{1:T} = (y_1, \dots, y_T)'$ a informação avaliada até o tempo T . Consideramos $\mathbf{h}_{0:T}$ e $\boldsymbol{\lambda}_{1:T}$ como parâmetros latentes. Usando o Teorema de Bayes, a densidade a posteriori dos parâmetros e das variáveis latentes pode ser escrita

$$p(\mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T}) \propto p(\mathbf{y}_{1:T} \mid \mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T})p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta})p(\boldsymbol{\lambda}_{1:T} \mid \boldsymbol{\theta})p(\boldsymbol{\theta}), \quad (3.28)$$

onde

$$p(\mathbf{y}_{1:T} \mid \boldsymbol{\lambda}_{1:T}, \mathbf{h}_{0:T}) \propto \prod_{t=1}^T \lambda_t^{1/2} e^{-\frac{h_t + \lambda_t y_t^2 e^{-h_t}}{2}}, \quad (3.29)$$

$$p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta}) \propto e^{-\frac{1-\phi^2}{2\sigma_\eta^2}(h_0 - \frac{\alpha}{1-\phi})^2} \prod_{t=1}^T e^{-\frac{1}{2\sigma_\eta^2}(h_t - \alpha - \phi h_{t-1})^2}, \quad (3.30)$$

$$p(\boldsymbol{\lambda}_{1:T} \mid \boldsymbol{\theta}) = \prod_{t=1}^T p(\lambda_t), \quad (3.31)$$

sendo $p(\boldsymbol{\theta})$ a distribuição a priori. Para os parâmetros comuns da classe de modelos VE-MEN, as distribuições a prioris são dadas como: $\alpha \sim \mathcal{N}(\bar{\alpha}, \sigma_\alpha^2)$, $\phi \sim \mathcal{N}_{(-1,1)}(\bar{\phi}, \sigma_\phi^2)$, e $\sigma_\eta^2 \sim \mathcal{IG}(\frac{T_0}{2}, \frac{M_0}{2})$, onde $\mathcal{N}_{(a,b)}(\cdot, \cdot)$ denota a distribuição normal truncada no intervalo(a,b).

Como a densidade a posteriori $p(\mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$ não tem forma analítica fechada, usamos o algoritmo de Metrópolis-Hastings para obter valores dos parâmetros, primeiramente amostramos $\boldsymbol{\theta}$, depois as variáveis latentes $\boldsymbol{\lambda}_{1:T}$ e $\mathbf{h}_{0:T}$. O algoritmo é descrito a seguir:

Algoritmo 3.1

1. Para $i = 0$ inicializamos os valores dos parâmetros $\boldsymbol{\theta}^{(i)}$, os estados $\boldsymbol{\lambda}_{1:T}^{(i)}$ e $\mathbf{h}_{0:T}^{(i)}$;
2. Amostramos $\boldsymbol{\theta}^{(i+1)} \sim p(\boldsymbol{\theta} \mid \mathbf{h}_{0:T}^{(i)}, \boldsymbol{\lambda}_{1:T}^{(i)}, \mathbf{y}_{1:T})$;
3. Amostramos $\boldsymbol{\lambda}_{1:T}^{(i+1)} \sim p(\boldsymbol{\lambda}_{1:T} \mid \boldsymbol{\theta}^{(i+1)}, \mathbf{h}_{0:T}^{(i)}, \mathbf{y}_{1:T})$;
4. Amostramos $\mathbf{h}_{0:T}^{(i+1)} \sim p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta}^{(i+1)}, \boldsymbol{\lambda}_{1:T}^{(i+1)}, \mathbf{y}_{1:T})$;

5. Fazemos $i = i + 1$ e repetimos até convergir.

Como descrevemos no algoritmo 3.1, o Gibbs sampler requer para amostrar os parâmetros, as distribuições condicionais completas das variáveis latentes. Considerando a estrutura Markoviana dos estados desenvolvemos um "multi-move sampler" (Shephard e Pitt (1997), Watanabe e Omori (2004), Omori e Watanabe (2008)) para amostrar $\mathbf{h}_{0:T}$ por blocos. Detalhes das condicionais completas de $\boldsymbol{\theta}$ e das variáveis latentes $\boldsymbol{\lambda}_{1:T}$ são dados a seguir (para mais detalhes veja Abanto-Valle *et al.* (2010)).

Distribuição Condicional Completa de α, ϕ e σ_η^2

As distribuições a priori dos parâmetros mencionados são considerados na literatura como: $\alpha \sim N(\bar{\alpha}, \sigma_\alpha^2)$ (Kim *et al.* (1998), Jun (2006), Shephard (1996)), Ando (2007), $\phi \sim \mathcal{N}_{(-1,1)}(\bar{\phi}, \sigma_\phi^2)$ (Jun (2000), Kim *et al.* (1998)), Ando (2007) entre outros. Também é usual considerar a priori de ϕ como uma distribuição *Beta*(20, 1.5), e $\sigma_\eta^2 \sim \mathcal{IG}(\frac{T_0}{2}, \frac{M_0}{2})$. De acordo com (3.28) e (3.30), a distribuição condicional para α :

$$p(\alpha \mid \mathbf{h}_{0:T}, \phi, \sigma_\eta^2) \propto \exp\left\{-\frac{a_\alpha}{2} \left(\alpha - \frac{b_\alpha}{a_\alpha}\right)^2\right\}, \quad (3.32)$$

que é uma distribuição normal com média $\frac{b_\alpha}{a_\alpha}$ e variância $\frac{1}{a_\alpha}$, onde

- $a_\alpha = \frac{1}{\sigma_\alpha^2} + \frac{T}{\sigma_\eta^2} + \frac{1+\phi}{\sigma_\eta^2(1-\phi)}$ e
- $b_\alpha = \frac{\bar{\alpha}}{\sigma_\alpha^2} + \frac{(1+\phi)}{\sigma_\eta^2} h_0 + \frac{\sum_{t=1}^T (h_t - \phi h_{t-1})}{\sigma_\eta^2}$.

De forma similar, de (3.30) e (3.28), a distribuição condicional de ϕ é dada por,

$$p(\phi \mid \mathbf{h}_{0:T}, \alpha, \sigma_\eta^2, \nu) \propto Q(\phi) \exp\left\{-\frac{a_\phi}{2\sigma_\eta^2} \left(\phi - \frac{b_\phi}{a_\phi}\right)^2\right\} \mathbb{I}_{|\phi| < 1}, \quad (3.33)$$

onde

- $Q_\phi = \sqrt{1 - \phi^2} \exp\left\{-\frac{1}{2\sigma_\eta^2} [(1 - \phi^2)(h_0 - \frac{\alpha}{1-\phi})^2]\right\}$,
- $a_\phi = \sum_{t=1}^T h_{t-1}^2 + \frac{\sigma_\eta^2}{\sigma_\phi^2}$,
- $b_\phi = \sum_{t=1}^T h_{t-1}(h_t - \alpha) + \bar{\phi} \frac{\sigma_\eta^2}{\sigma_\phi^2}$

e $\mathbb{I}_{|\phi|<1}$ é uma variável indicadora.

Como $p(\phi \mid \mathbf{h}_{0:T}, \alpha, \sigma_\eta^2, \rho, \nu)$ em (3.33) não tem forma fechada, amostramos usando o algoritmo Metropolis-Hastings. Neste caso consideramos como a densidade proposta a distribuição normal truncada $\mathcal{N}_{(-1,1)}(\frac{b_\phi}{a_\phi}, \frac{\sigma_\eta^2}{a_\phi})$.

De (3.30) e (3.28), a posteriori condicional de σ_η^2 é $\mathcal{IG}(\frac{T_1}{2}, \frac{M_1}{2})$, onde

- $T_1 = T_0 + T + 1$ e
- $M_1 = M_0 + [(1 - \phi^2)(h_0 - \frac{\alpha}{1-\phi})^2] + \sum_{t=1}^T (h_t - \alpha - \phi h_{t-1})^2$.

Condicionais completas de λ_t e ν

• Caso VE-t

Como $\lambda_t \sim \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$, a condicional completa de λ_t é dada por

$$p(\lambda_t \mid y_t, h_t, \nu) \propto \lambda_t^{\frac{\nu+1}{2}-1} e^{-\frac{\lambda_t}{2}(y_t^2 e^{-h_t} + \nu)}, \quad (3.34)$$

que é a distribuição gama, $\mathcal{G}(\frac{\nu+1}{2}, \frac{y_t^2 e^{-h_t}}{2})$.

Assumimos a distribuição a priori de ν como $\mathcal{G}(a_\nu, b_\nu)\mathbb{I}_{2<\nu\leq 40}$. Então, a condicional completa de ν é

$$p(\nu \mid \boldsymbol{\lambda}_{1:T}) \propto \frac{\left[\frac{\nu}{2}\right]^{\frac{T\nu}{2}} \nu^{a_\nu-1} e^{-\frac{\nu}{2} \sum_{t=1}^T [(\lambda_t - \log \lambda_t) + 2b_\nu]}}{\Gamma(\frac{\nu}{2})} \mathbb{I}_{2<\nu\leq 40}. \quad (3.35)$$

Amostramos ν pelo algoritmo de aceitação/rejeição Metropolis-Hastings (Tierney (1994), Chib et al. (2002)). Seja ν^* a moda (ou aproximação da moda) de $p(\nu \mid \boldsymbol{\lambda}_{1:T})$, e seja $\ell(\nu) = \log p(\nu \mid \boldsymbol{\lambda}_{1:T})$. Como $\ell(\nu)$ é concava, usamos a densidade proposta $\mathcal{N}_{(2,40)}(\mu_\nu, \sigma_\nu^2)$, onde $\mu_\nu = \nu^* - \ell'(\nu^*)/\ell''(\nu^*)$ e $\sigma_\nu^2 = -1/\ell''(\nu^*)$. Aqui $\ell'(\nu^*)$ e $\ell''(\nu^*)$ são a primeira e segunda derivada de $\ell(\nu)$ calculada em $\nu = \nu^*$. Para provar a concavidade de $\ell(\nu)$, temos o resultado de Abramowitz e Stegun (1970), na qual $\log \Gamma(\nu)$ pode-se aproximar como

$$\log \Gamma(\nu) = \frac{\log(2\pi)}{2} + \frac{2\nu - 1}{2} \log(\nu) - \nu + \frac{\theta}{12\nu}, \quad 0 < \theta < 1. \quad (3.36)$$

Tomando a segunda derivada de $\ell(\nu)$ para (3.35) e usando (3.36), temos

$$\ell''(\nu) = -\frac{T\theta}{3\nu^3} - \frac{(T + 2a_\nu - 2)\nu}{2\nu^2} < 0.$$

• **Caso VE-S**

Sabemos que $\lambda_t \sim \mathcal{B}e(\nu, 1)$, logo temos que a condicional completa de λ_t é dada por

$$p(\lambda_t \mid y_t, h_t, \nu) \propto \lambda_t^{\nu+\frac{1}{2}-1} e^{-\frac{\lambda_t}{2} y_t^2 e^{-h_t}} \mathbb{I}_{0 < \lambda_t < 1}, \quad (3.37)$$

isto é $\lambda_t \sim \mathcal{G}_{(0 < \lambda_t < 1)}(\nu + \frac{1}{2}, \frac{1}{2} y_t^2 e^{-h_t})$, ou seja, o lado direito é a distribuição gama truncada. Assumindo a distribuição a priori $\nu \sim \mathcal{G}(a_\nu, b_\nu)$, a distribuição condicional completa de ν é dada por

$$p(\nu \mid \mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}) \propto \nu^{T+a_\nu-1} e^{-\nu(b_\nu - \sum_{t=1}^T \log \lambda_t)} \mathbb{I}_{\nu > 1}, \quad (3.38)$$

ou seja, é uma $\mathcal{G}_{\nu > 1}(T + a_\nu, b_\nu - \sum_{t=1}^T \log \lambda_t)$, ou seja, uma distribuição gama truncada. Simulamos desta distribuição gama truncada usando o algoritmo proposto por [Philippe \(1997\)](#).

• **Caso VE-VG**

Como $\lambda_t \sim \mathcal{IG}(\frac{\nu}{2}, \frac{\nu}{2})$, a condicional completa de λ_t é dada por

$$p(\lambda_t \mid y_t, h_t, \nu) \propto \lambda_t^{-\frac{\nu}{2} + \frac{1}{2} - 1} e^{-\frac{1}{2}(\lambda_t y_t^2 e^{-h_t} + \frac{\nu}{\lambda_t})}, \quad (3.39)$$

que é a distribuição gaussiana inversa generalizada, $\mathcal{GIG}(-\frac{\nu}{2} + \frac{1}{2}, y_t^2 e^{-h_t}, \nu)$.

Assumimos a distribuição a priori de ν como $\mathcal{G}(a_\nu, b_\nu) \mathbb{I}_{0 < \nu \leq 40}$. Então, a condicional completa de ν é

$$p(\nu \mid \mathbf{y}_{1:T}, \mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}) \propto \frac{\left[\frac{\nu}{2} \right]^{\frac{T\nu}{2}} \nu^{a_\nu-1} e^{-\frac{\nu}{2} \sum_{t=1}^T [(\frac{1}{\lambda_t} + \log \lambda_t) + 2b_\nu]}}{\Gamma(\frac{\nu}{2})} \mathbb{I}_{0 < \nu \leq 40}, \quad (3.40)$$

que é côncava. Então, amostramos ν pelo algoritmo de aceitação/rejeição Metropolis-Hastings como no caso do modelo VE-t com desidade proposta $\mathcal{N}_{(0,40)}(\mu_\nu, \sigma_\nu^2)$.

Aplicação

Baseados nos dados do exemplo anterior, vamos considerar agora a implementação dos modelos VE-MEN1. Uma vez mais para a estimação dos parâmetros consideramos o amostrador de Gibbs, com as prioris sugeridas no exemplo anterior, a elas se adiciona o parâmetro $\nu = \frac{1}{\nu^*}$, para o qual se considerou a distribuição a priori para ν^* de $U_{(0,0025;0,25)}$.

Realizamos igual número de simulações que nos outros casos (100000) e consideramos o

mesmo período “burn-in” de 5000 iterações. No entanto, e devido a alta correlação das amostras geradas dos parâmetros, escolhemos uma amostra sistemática de 30 em 30 para os parâmetros do modelo VEN, para o modelo VET, a escolha foi de 25 em 25. No caso do modelo VEMEN1-T, como no caso anterior foi de 25 em 25, da VEMEN1-Slash de 50 em 50 e para a VEMEN1-VG escolhemos de 50 em 50.

Os gráficos mostram o desempenho do "single move" para cada parâmetro, assim pode-se observar que as cadeias convergem lentamente, enquanto na maioria dos parâmetros as funções de autocorrelações decaem lentamente para zero. Isto acontece devido a alta correlação dos valores gerados na cadeia de Markov. Por esse motivo, no modelo onde o erro tem distribuição Variância Gama, realizamos 110.000 simulações, descartando as primeiras 60.000, destas escolhemos uma amostra sistemática com intervalo 50.

Finalmente, para garantir a convergência calculamos as estatísticas de CD, do critério de convergência de Geweke (1992), que indicaram convergência ao nível de 95% de credibilidade como se pode observar na Tabela 3.1.

Tabela 3.1: Estimativas dos parâmetros para a série de retornos S&P500. Primeira linha: Média a posteriori. Segunda linha: Intervalo de credibilidade a posteriori, nível de confiança de 95% entre parênteses. Terceira linha: O erro Monte Carlo da média a posteriori. Quarta linha: estatística de convergência de Geweke(CD).

Parameter	VE-N	VE-MENI-t	VE-MENI-Slash	VE-MENI-NContam	VE-MENI-VG
α	-0,00334 (-0,0109;0,0027) $0,35 \times 10^{-4}$ -0,09435	-0,000905 (-0,0051;0,0034) $0,215 \times 10^{-2}$ -0,8494	-0,008618 (-0,01781;-0,00089) $4,168 \times 10^{-3}$ 0,1059	0,07061 (0,02152; 0,1271) $2,89 \times 10^{-3}$ -0,2144	-0,00368 (-0,00916; 0,000928) $0,281 \times 10^{-2}$ 1,399
ϕ	0,9757 (0,9577;0,9904) $8,6 \times 10^{-3}$ -0,886	0,9869 (0,9745;0,996) $6,06 \times 10^{-3}$ -0,4321	0,9828 (0,9679;0,9956) $6,63 \times 10^{-3}$ -0,295	0,9735 (0,9512; 0,9902) $1,03 \times 10^{-2}$ 0,5433	0,9875 (0,9742; 0,9972) $0,5804248 \times 10^{-2}$ 1,219
σ^2	0,1393 (0,1058;0,1798) $1,9 \times 10^{-2}$ 0,18	0,0936 (0,0675;0,1279) $1,53 \times 10^{-2}$ -0,1849	0,1098 (0,0769;0,1473) $1,7 \times 10^{-2}$ 1,236	0,1446 (0,1029;0,1896) $2,48 \times 10^{-2}$ -0,8928	0,0915 (0,06387; 0,124) $1,5545480 \times 10^{-2}$ -0,7753
ν	— — — —	10,21 (6,72;16,61) 0,253 -1,127	5,485 (5,00;6,854) 0,66 0,1056	0,08396 (0,06543;0,1071) 0,01130 -1,682	10,18 (7,039; 13,23) 0,28 0,4332
γ	— — — —	— — — —	— — — —	0,4749 (0,03583;0,9421) 0,2843 0,5691	— — — —

Ao analisar os resultados (veja a Tabela 3.1), os valores dos parâmetros α , ϕ , e σ não se diferenciam muito nas estimativas entre os modelos, convergindo para valores razoáveis e coerentes com a literatura, por exemplo ϕ converge para um valor perto de um, indicando uma alta persistência da volatilidade; entretanto, o mesmo não acontece com a variável de mistura ν . Podemos observar que as suas estimativas diferem razoavelmente entre os modelos. No modelo $VE-VGMEN1$ a média a posteriori resulta 10,18, na $VE-TMEN1$ resulta 10,207 e no modelo $VE-SlashMEN1$ o resultado é 5,55. As figuras 3.1-3.5 mostram o desempenho das cadeias para os diferentes modelos e parâmetros.

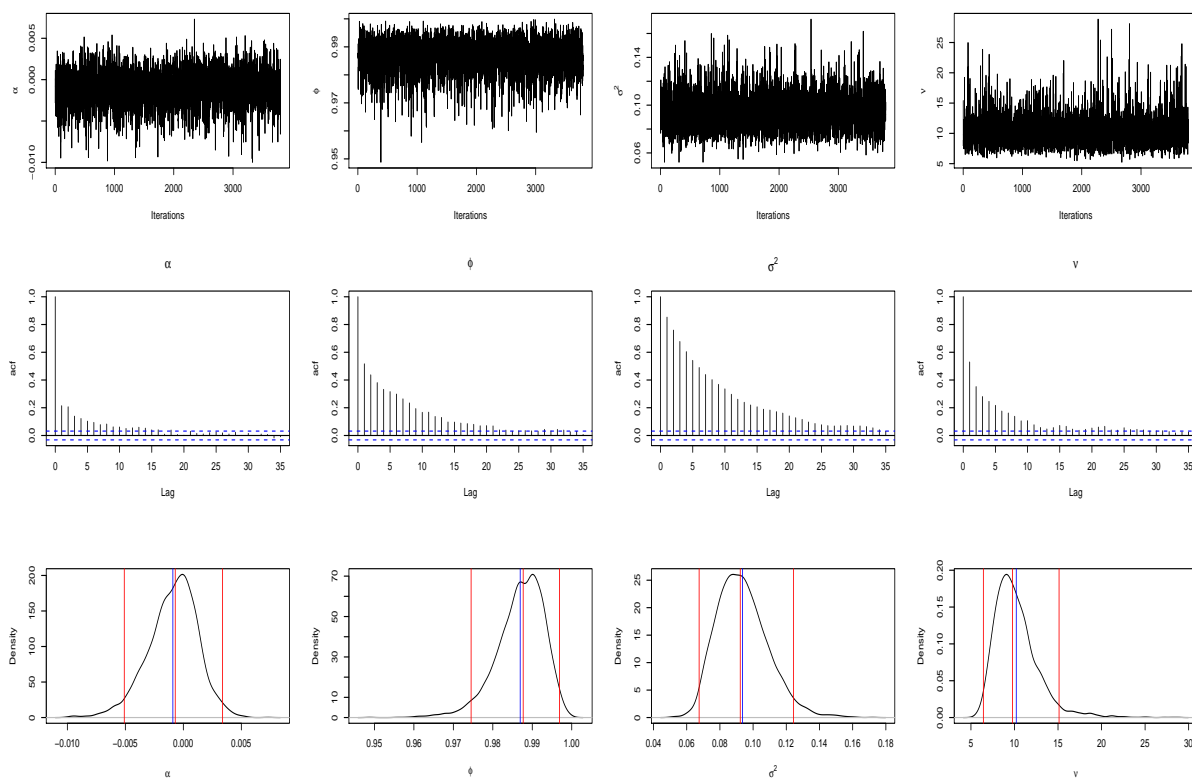


Figura 3.1: Gráfico da série de iterações dos parâmetros α , ϕ , σ e ν obtidas nas MCMC para a série de retornos diários da S&P500 (modelo VE-MEN1-t-Student). Simulações (Esquerda), função de autocorrelação das iterações(Centro), densidade posteriori (Direita), as linhas indicam o intervalo de maior credibilidade do nível 5% e as linhas centrais a média e mediana posteriori.

3.3. MODELOS DE VOLATILIDADE ESTOCÁSTICA COM ERRO NA FAMÍLIA MEN 39

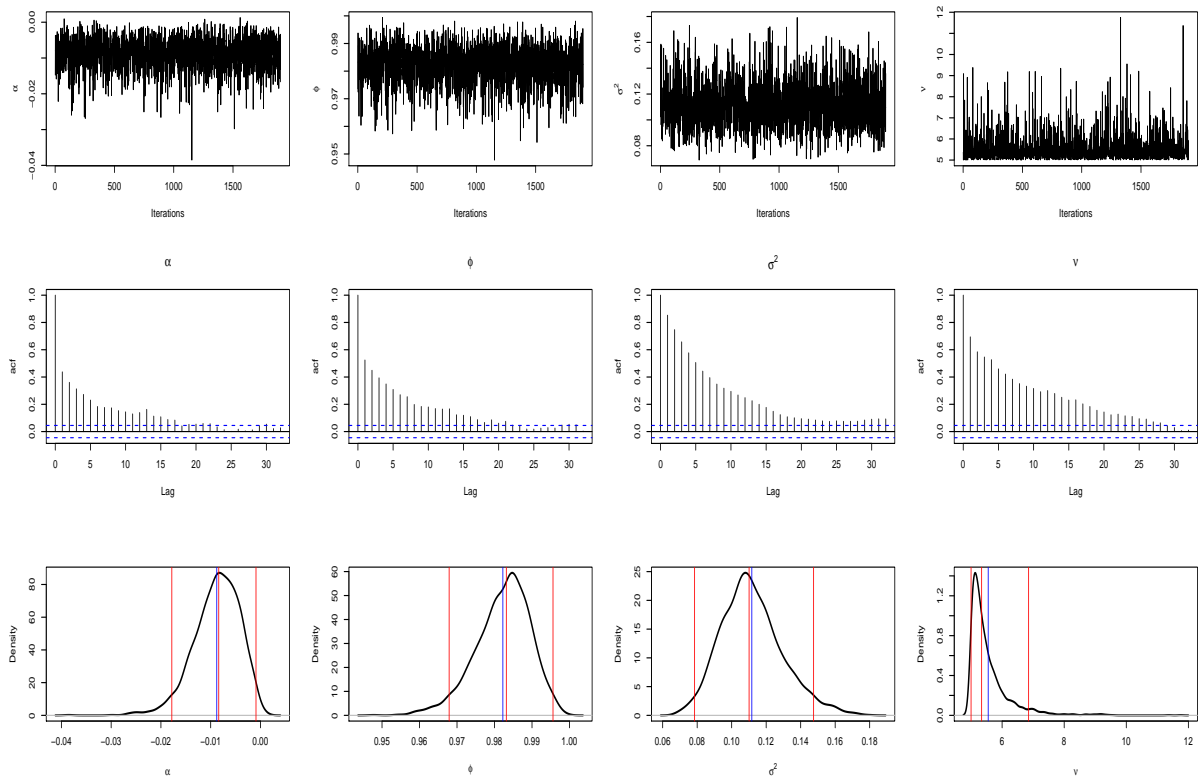


Figura 3.2: Gráfico da série de iterações dos parâmetros α , ϕ , σ e ν obtidas nas MCMC para a série de retornos diários da S&P500 (modelo VE-MEN1-Slash). Simulações (Esquerda), função de autocorrelação das iterações (Centro), densidade posteriori (Direita), as linhas indicam o intervalo de maior credibilidade do nível 5% e as linhas centrais a média e mediana posteriori.

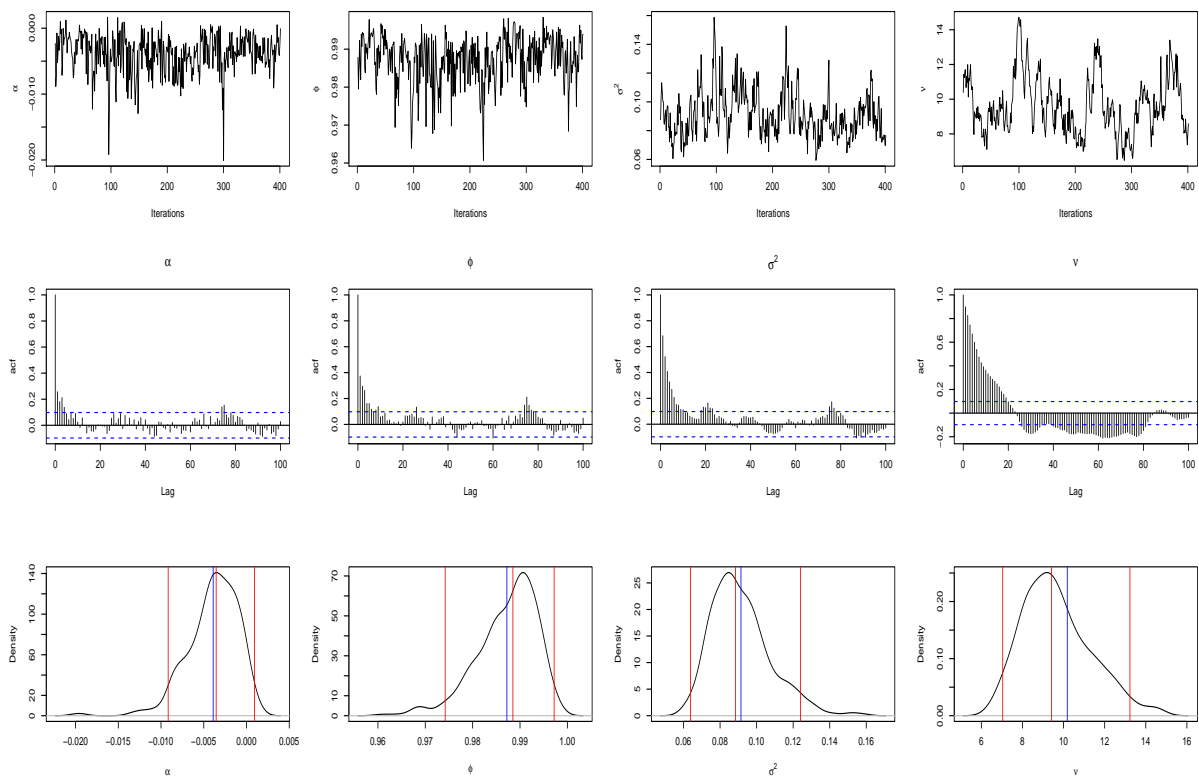


Figura 3.3: Gráfico da série iterações dos parâmetros α , ϕ , σ e ν obtidas nas MCMC para a série de retornos diários da S&P500 (modelo VE-MEN1-Variância Gamma). Simulações (Esquerda), função de autocorrelação das iterações (Centro), densidade posteriori (Direita), as linhas indicam o intervalo de maior credibilidade do nível 5% e as linhas centrais a média e mediana posteriori.

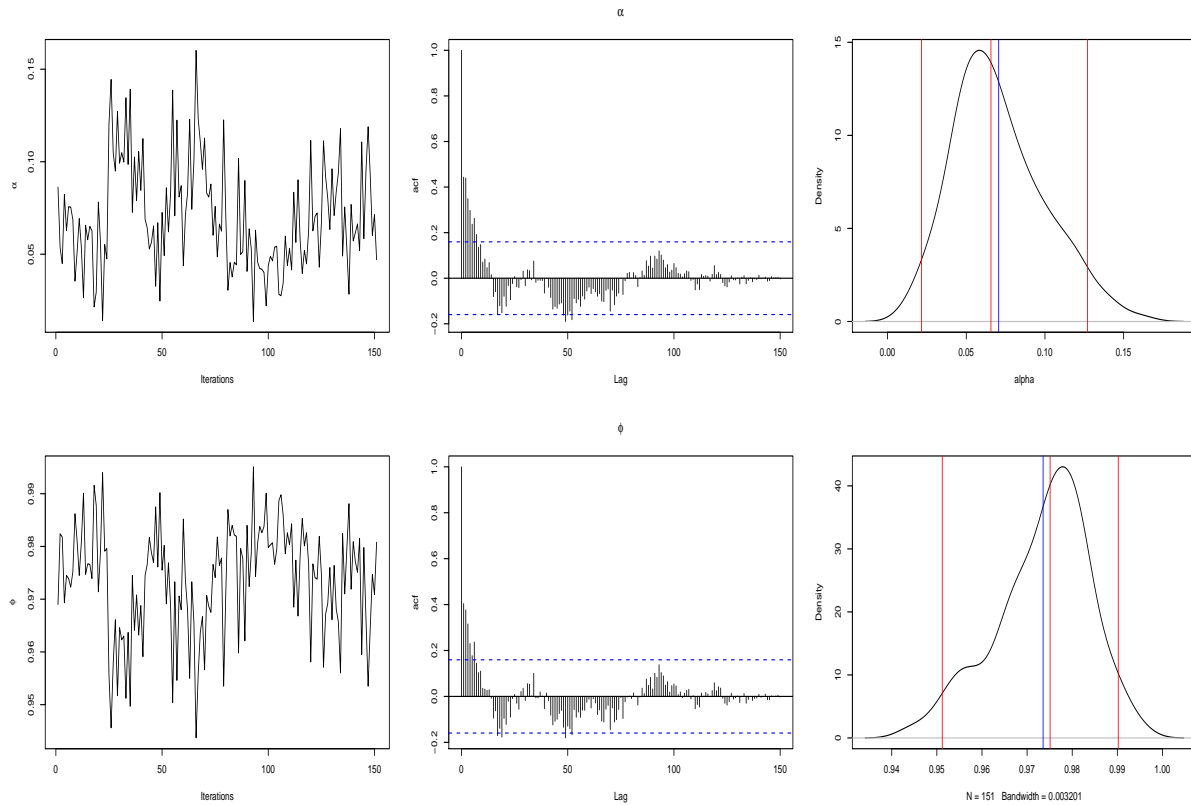


Figura 3.4: Gráfica da série de iterações dos parâmetros α e ϕ obtidas nas MCMC para a série de retornos diários da S&P500 (modelo VE-MEN1-NormalContaminada). Simulações (Esquerda), função de autocorrelação das iterações(Centro), densidade posteriori (Direita), as linhas indica o intervalo de maior credibilidade do nível 5% e as linhas centrais a média e mediana posteriori.

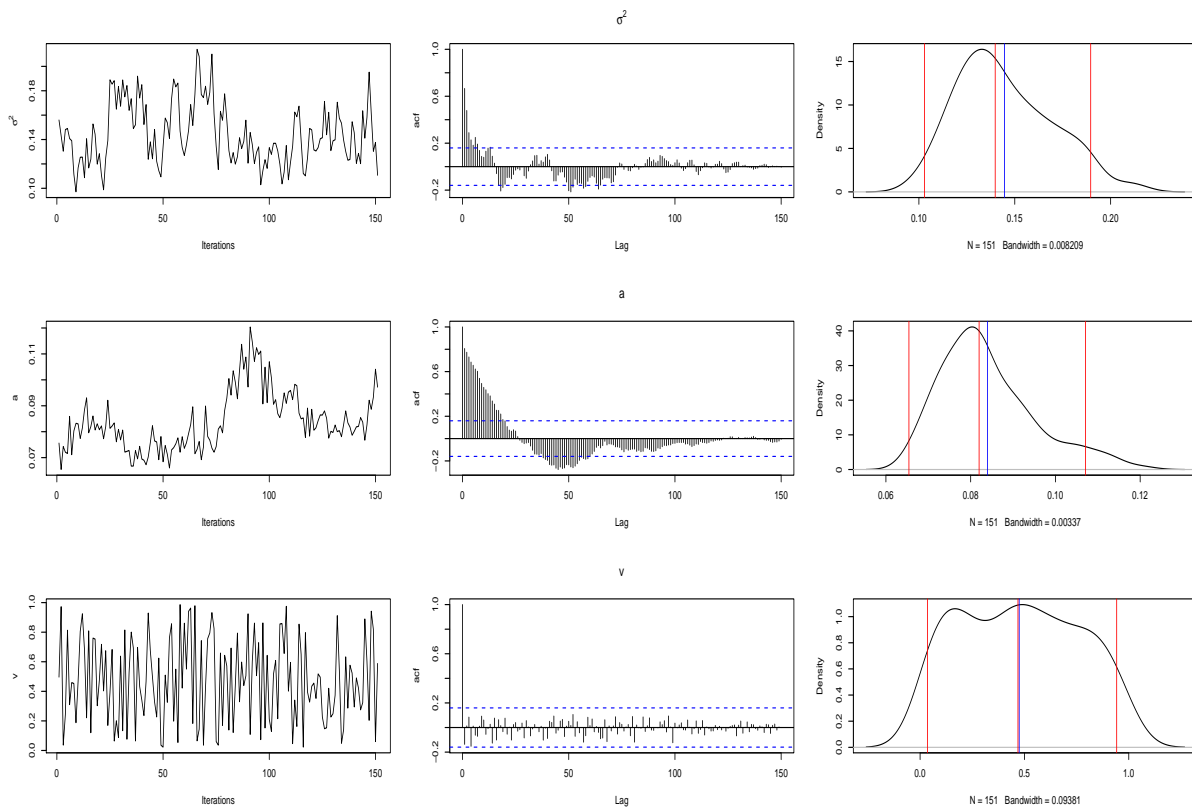


Figura 3.5: Gráfica da série iterações dos parâmetros σ , γ e ν obtidas nas MCMC para a série de retornos diários da S&P 500 (modelo VE-MEN1-Normal Contaminada). Simulações (Esquerda), função de autocorrelação das iterações (Centro), densidade posteriori (Direita), as linhas indica o intervalo de maior credibilidade do nível 5% e as linhas centrais a média e mediana posteriori.

Fazendo uso do critério DIC de seleção de modelos, comparamos os modelos VE-MEN1, resultados que se observam na Tabela 3.2. Neste caso, o modelo de mistura na escala de normal, com distribuição Variância Gamma, seguida da distribuição normal contaminada resulta ser o mais adequado para os retornos da S&P500. Note que os modelos com MEN resultam melhores que o modelo clássico com erro normal, devido à característica de caudas pesadas das série de retorno S&P500. Este fato é confirmado pelo valor do parâmetro ν dos diferentes modelos simulados (Tabela 3.1).

Tabela 3.2: Comparação entre os modelo VE-MEN1: VE-Normal com Efeito alavanca, t-Student com efeito alavanca, VE-Slash com efeito alavanca, VE-Variância Gama com efeito alavanca e VE-Normal Contaminada com efeito alavanca.

Modelo	DIC ($pD = \text{var}(\text{deviance})/2$)	Ordem
VE-N	5502.03 ($pD = 93.159$)	5
VE-TMEN1	5476.9 ($pD = 235.4$)	4
VE-SlashMEN1	5473.0 ($pD = 97.2$)	3
VE-VGMEN1	5397 ($pD = 157$)	1
VE-NCMEN1	5409.5 ($pD = -0.3$)	2

3.3.2 O modelo VE-MEN2:

Neste caso, consideramos o erro com distribuição na escala de normais, em ambas as equações, de estados e observação. Além disso, na tentativa de medir o efeito de alavancagem consideramos correlação entre ambos distúrbios (VE-MEN2). Isto é,

$$y_t = e^{\frac{h_t}{2}} \varepsilon_t, \quad (3.41a)$$

$$h_t = \alpha + \phi h_{t-1} + \sigma_\eta \eta_t, \quad (3.41b)$$

com

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim SMN \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \kappa(\lambda) \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}; v \right), \quad (3.42a)$$

onde y_t e h_t são respectivamente os retornos e log-volatilidade no tempo t . Observe que as inovações ε_t e η_t são distribuídas com misturas na escala normal e assumimos uma correlação entre elas. A distribuição condicionada delas, dada a variável de mistura λ_t , é normal bivariada com média zero e variância como indicada na equação (3.42a). Espera-se que a correlação seja negativa devido ao possível efeito de alavanca. Na forma hierárquica temos,

$$y_t = e^{\frac{h_t}{2}} \lambda_t^{-\frac{1}{2}} \varepsilon_t, \quad (3.43a)$$

$$h_t = \alpha + \phi h_{t-1} + \sigma_\eta \lambda_t^{-\frac{1}{2}} \eta_t', \quad (3.43b)$$

$$\lambda_t \sim p(\lambda_t), \quad \varepsilon_t \sim N(0, 1), \quad \eta_t' \sim N(0, 1). \quad (3.43c)$$

onde $\varepsilon_t = \lambda_t^{-\frac{1}{2}} \varepsilon_t$ e $\eta_t = \lambda_t^{-\frac{1}{2}} \eta_t'$, de modo que

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} | \lambda_t \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \kappa(\lambda) \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right). \quad (3.44a)$$

Analogamente ao modelo VE-MEN1 a verossimilhança $P(\mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}, \boldsymbol{\theta} | \mathbf{y}_{1:T})$, em que $\boldsymbol{\theta} = (\alpha, \phi, \sigma^2, v)$, não tem forma fechada. Utilizaremos o algoritmo Gibbs Sampler para simular os valores de interesse dos parâmetros. Realizamos uma aproximação para estimar os parâmetros nos modelos VE-MEN2, definido pelas equações anteriores mediante a técnica MCMC.

De forma similar que o caso VE-MEN1, usando o Teorema de Bayes, a densidade a posteriori

dos parâmetros e das variáveis latentes fica

$$p(\mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T}) \propto p(\mathbf{y}_{1:T} \mid \mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T})p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta})p(\boldsymbol{\lambda}_{1:T} \mid \boldsymbol{\theta})p(\boldsymbol{\theta}), \quad (3.45)$$

onde

$$\begin{aligned} p(\mathbf{y}_{1:T} \mid \boldsymbol{\lambda}_{1:T}, \mathbf{h}_{0:T}, \alpha, \sigma_\eta, \phi, \rho, \nu) &\propto \prod_{t=1}^T \frac{1}{(e^{h_t}(1-\rho^2)K(\lambda_t))^{1/2}} e^{-\frac{1}{2e^{h_t}(1-\rho^2)K(\lambda_t)} \left(y_t - \frac{e^{h_t/2}\rho}{\sigma_\eta} (h_{t+1} - \alpha - \phi h_t) \right)^2}, \\ p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta}) &\propto \frac{\sqrt{1-\phi^2}}{\sigma_\eta} \times e^{\frac{-(1-\phi^2)}{2\sigma_\eta^2} (h_0 - \frac{\alpha}{1-\phi})^2} \times \prod_{t=1}^T \frac{1}{\sigma_\eta} e^{\frac{-1}{2\sigma_\eta^2 K(\lambda_t)} (h_t - \alpha - \phi h_{t-1})^2}, \\ p(\boldsymbol{\lambda}_{1:T}) &\propto p(\boldsymbol{\lambda}_{1:T}/\nu), \\ p(\boldsymbol{\theta}) &\propto p(\alpha)p(\phi)p(\sigma_\eta^2)p(\rho)p(\nu). \end{aligned}$$

Como a densidade a posteriori $p(\mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T})$ não tem forma fechada, primeiramente amostramos $\boldsymbol{\theta}$, depois as variáveis latentes $\boldsymbol{\lambda}_{1:T}$ e $\mathbf{h}_{0:T}$ usando o amostrador de Gibbs. O algoritmo é descrito a seguir.

Algoritmo 3.2

1. Inicializamos ($i = 0$) os valores de $\boldsymbol{\theta}$ e h e λ ;
2. Amostramos \mathbf{h}_t de $\mathbf{h}_{1:n}/h_{/t}, y_t, \boldsymbol{\theta}$, $t = 1, \dots, n$;
2. Amostramos λ_t de $\lambda_{1:n}/\mathbf{h}_{0:T}, \mathbf{y}_t, \boldsymbol{\theta}$, $t = 1, \dots, n$;
3. Amostramos σ_η^2 de $p(\sigma_\eta^2 \mid \mathbf{y}_{1:T}, \mathbf{h}_{0:T}, \phi, \alpha, \rho, \nu)$;
4. Amostramos ϕ de $p(\phi \mid \mathbf{h}_{0:T}, \sigma_\eta^2, \alpha, \rho, \nu)$;
5. Amostramos α de $p(\alpha \mid \mathbf{h}_{0:T}, \phi, \sigma_\eta^2, \rho, \nu)$;
4. Amostramos ρ de $p(\rho \mid \mathbf{h}_{0:T}, \sigma_\eta^2, \alpha, \phi, \nu)$;
5. Amostramos ν de $p(\nu \mid \mathbf{h}_{0:T}, \phi, \sigma_\eta^2, \alpha, \nu)$;
6. Faz-se $t = t + 1$ e repete-se até que convirja.

Condicionais Completas dos parâmetros comuns α , ϕ , σ e ρ .

As distribuições das condicionais completas são as seguintes:

A distribuição condicional de α é uma $N(\mu^*, \sigma^*)$ onde,

- $\mu^* = \frac{a_\alpha}{b_\alpha}$ e $\sigma^* = \frac{1}{b_\alpha^2}$,
- $a_\alpha = \frac{\sum_{t=1}^T (h_t - \phi h_{t-1})}{\sigma_\eta^2 \kappa(\lambda)} + \frac{\bar{\alpha}}{\sigma_\alpha^2} + \frac{(1+\phi)}{\sigma_\eta^2} h_0 + \frac{\sum_{t=1}^T \rho e^{h_t}}{\kappa(\lambda) \sigma_\eta (1-\rho^2)} \left(\frac{\rho e^{h_t}}{\sigma_\eta} (h_t - \phi h_{t-1}) + y_t \right)$,
- $b_\alpha = \frac{1}{\sigma_\alpha^2} + \sum_{t=1}^T \frac{1}{\sigma_\eta^2 \kappa(\lambda)} + \frac{1+\phi}{\sigma_\eta^2 (1-\phi)} + \frac{\sum_{t=1}^T \rho^2}{\kappa(\lambda) \sigma_\eta^2 (1-\rho^2)}$.

A distribuição condicional de ϕ é dada por

$$p(\phi \mid \mathbf{h}_{0:T}, \alpha, \sigma_\eta^2, \rho, \nu) \propto Q(\phi) \exp\left\{-\frac{a_\phi}{2\sigma_\phi^2} \left(\phi - \frac{b_\phi}{a_\phi}\right)^2\right\} \mathbb{I}_{|\phi| < 1}, \quad (3.46)$$

onde

- $Q_\phi = \sqrt{1 - \phi^2} \exp\left\{-\frac{1}{2\sigma_\eta^2} [(1 - \phi^2)(h_0 - \frac{\alpha}{1-\phi})^2]\right\}$,
- $a_\phi = \sum_{t=1}^T \frac{h_{t-1}^2}{k(\lambda)} + \frac{\sigma_\eta^2}{\sigma_\phi^2} + \frac{\sum_{t=1}^T \rho^2 h_t^2}{(1-\rho^2)k(\lambda)}$,
- $b_\phi = \sum_{t=1}^T \frac{h_{t-1}(h_t - \alpha)}{k(\lambda)} + \bar{\phi} \frac{\sigma_\eta^2}{\sigma_\phi^2} + \frac{\sum_{t=1}^T \rho^2 h_t (h_{t+1} - y_t A_t^{-1} - \alpha)}{(1-\rho^2)k(\lambda)}$, $A_t = \frac{\rho e^{h_t}}{\sigma_\eta}$ e $\mathbb{I}_{|\phi| < 1}$ é uma variável indicadora.

Como $p(\phi \mid \mathbf{h}_{0:T}, \alpha, \sigma_\eta^2, \rho, \nu)$ não tem forma fechada, amostramos usando o algoritmo Metropolis-Hastings com a densidade truncada proposta $\mathcal{N}_{(-1,1)}\left(\frac{b_\phi}{a_\phi}, \frac{\sigma_\eta^2}{a_\phi}\right)$.

Para a posteriori condicional de σ_η^2 tendo como priori $P(\sigma_\eta^2) \sim \mathcal{IG}\left(\frac{T_0}{2}, \frac{M_0}{2}\right)$, a função condicional completa para σ_η^2 é

$$p(\sigma_\eta^2 \mid \mathbf{y}_{1:T}, \mathbf{h}_{0:T}, \phi, \alpha, \rho, \nu) \propto e^{-\frac{1}{2} [y_t - \frac{\rho e^{h_t/2} (h_{t+1} - \alpha - \phi h_t)}{\sigma_\eta}]^2} \times \frac{1}{(\sigma_\eta^2)^{1/2}} e^{\left[\frac{-1}{2\sigma_\eta} (1 - \phi^2) (h_0 - \frac{\alpha}{1-\phi})^2\right]} \frac{1}{(\sigma_\eta^2)^{T/2}} e^{\left[\frac{-1}{2\sigma_\eta} (h_t - \alpha - \phi h_{t-1})^2\right]} p(\sigma_\eta^2). \quad (3.47)$$

De onde,

$$\sigma_\eta^2 \sim Q(\sigma_\eta^2) \times (\sigma_\eta^2)^{-\left[\frac{T_0+T+1}{2}+1\right]} e^{-\frac{1}{2} \left(\frac{M_0 + \sum_{t=1}^T (h_t - \alpha - \phi h_{t-1})^2 + (1-\phi^2)(h_0 - \frac{\alpha}{1-\phi})^2 + B_t^2}{\sigma_\eta^2}\right)}, \quad (3.48)$$

onde $B_t = \rho e^{\frac{h_t}{2}}(h_{t+1} - \alpha - \phi h_t)$. Isto é, $\sigma_t^2 \sim \mathcal{IG}(\frac{T_1}{2}, \frac{M_1}{2})$, onde $T_1 = T_0 + T + 1$ e $M_1 = M_0 + [(1 - \phi^2)(h_0 - \frac{\alpha}{1-\phi})^2] + \sum_{t=1}^T (h_t - \alpha - \phi h_{t-1})^2 + B_t^2$.

Finalmente para completar as distribuições dos parâmetros comuns temos a distribuição a posteriori condicional de ρ^2 :

$$p(\rho^2 \mid \mathbf{h}_{0:T}, \sigma_\eta^2, \alpha, \phi, v) \propto \prod_{t=1}^T \frac{1}{\sqrt{k(\lambda_t) e^{h_t} (1 - \rho^2)}} e^{-\frac{1}{2} [\frac{1}{k(\lambda_t) e^{h_t} (1 - \rho^2)}] [y_t - \frac{\rho e^{h_t/2}}{\sigma_\eta} (h_{t+1} - \alpha - \phi h_t)]^2} \times p(\rho), \quad (3.49)$$

$$p(\rho^2 \mid \mathbf{h}_{0:T}, \sigma_\eta^2, \alpha, \phi, v) \propto \frac{1}{\sqrt{1 - \rho^2}} e^{-\frac{b_\rho^2}{2(1 - \rho^2)} [\rho - (\frac{a_\rho}{b_\rho})]^2}, \quad (3.50)$$

$$p(\rho^2 \mid \mathbf{h}_{0:T}, \sigma_\eta^2, \alpha, \phi, v) \propto Q(\rho) e^{-\frac{b_\rho^2}{2(1 - \rho^2)} (\rho - (\frac{a_\rho}{b_\rho}))^2}. \quad (3.51)$$

Aqui, $Q(\rho) = \frac{1}{\sqrt{1 - \rho^2}}$, $a_\rho = \sum_{t=1}^T A_1 C_1$, $b_\rho^2 = \sum_{t=1}^T (C_1^2 - h_t)$, $C_1 = \frac{(h_{t+1} - \alpha - \phi h_t)}{k(\lambda_t)^{1/2} \sigma_\eta}$ and $A_1 = \frac{y_t e^{-h_t/2}}{k(\lambda_t^{1/2})}$.

Tomando a distribuição normal truncada $\mathcal{N}_{(-1,1)}(\frac{a_\rho}{b_\rho}, \frac{1}{b_\rho^2})$ como a densidade geradora. Como nos casos anteriores, quando não temos forma fechada, poderemos utilizar o algoritmo de Metropolis- Hastings de Aceitação/Rejeição.

Distribuição Condicional Completa de λ_t e ν

As distribuições das condicionais completas para ν e λ_t são as seguintes:

• Caso VE-t-Student

Como $\lambda_t \sim \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$, e considerando $f(\nu) = \frac{\nu - 2}{\nu}$, a condicional completa de λ_t para $t = 1, \dots, T - 1$, é dada por

$$p(\lambda_t \mid y_t, h_t, \nu) \propto \frac{\lambda_t^{\frac{\nu}{2}}}{f(\nu) e^{h_t} (1 - \rho^2)} e^{-\frac{\lambda_t^{1/2}}{2 * f(\nu) * e^{h_t} * (1 - \rho^2)} (y_t - A_t)^2} e^{-\frac{\lambda_t}{2} \frac{(h_t - \alpha - \phi * h_{t-1})^2}{f(\nu) * \sigma^2}} * p(\lambda), \quad (3.52)$$

que é a distribuição gama, $\mathcal{G}(\frac{\nu+1}{2}, \frac{(y_t - A_t)^2 e^{-h_t}}{2f(\nu)(1 - \rho^2)} + \frac{(h_t - \alpha - \phi h_{t-1})^2}{2f(\nu)\sigma^2} + \frac{\nu}{2})$. Já para $t = T$ temos, $\mathcal{G}(\frac{\nu+2}{2}, \frac{y_T^2 e^{h_T + \nu}}{2(1 - \rho^2)})$, considerando $A_t = \frac{\rho e^{h_t}}{\sigma_\eta}$

Assumimos a distribuição a priori de ν como $\mathcal{G}(a_\nu, b_\nu) \mathbb{I}_{2 < \nu \leq 40}$. Então, a condicional completa de ν é

$$p(\nu \mid \boldsymbol{\lambda}_{1:T}) \propto \frac{\left[\frac{\nu}{2}\right]^{\frac{T\nu}{2}} \nu^{a_\nu - 1} e^{-\frac{\nu}{2} \sum_{t=1}^T [(\lambda_t - \log \lambda_t) + 2b_\nu]}}{\Gamma(\frac{\nu}{2})} e^{-\frac{\lambda_t^2}{\nu} \frac{(y_t - A_t)^2}{(1 - \rho^2)}} \mathbb{I}_{2 < \nu \leq 40}. \quad (3.53)$$

Amostramos ν pelo algoritmo de aceitação/rejeição Metropolis-Hastings como no caso VE-MEN1.

• **Caso VE-S**

Sabemos que $\lambda_t \sim \mathcal{B}e(\nu, 1)$, ($p(\lambda_t) = \nu\lambda_t^{\nu-1}$), temos que a condicional completa de λ_t é dada por

$$p(\lambda_t | y_t, h_t, \nu) \propto \frac{\lambda_t^{\frac{1}{2}}}{f(\nu)e^{h_t(1-\rho^2)}} e^{-\frac{\lambda_t}{2f(\nu)e^{h_t(1-\rho^2)}}(y_t-A_t)^2} \frac{\lambda_t^{1/2}}{\sigma} e^{-\frac{\lambda_t}{2} \frac{(h_t-\alpha-\phi h_{t-1})^2}{f(\nu)\sigma^2}} p(\lambda_t), \mathbb{I}_{0 < \lambda_t < 1}, \quad (3.54)$$

isto é $\lambda_t \sim \mathcal{G}(\frac{\nu+1}{2}, \frac{(y_t-A_t)^2 e^{-h_t}}{2f(\nu)(1-\rho^2)} + \frac{(h_t-\alpha-\phi h_{t-1})^2}{2f(\nu)\sigma^2})$, ou seja, o lado direito é a distribuição gama truncada.

Como a distribuição a priori de ν é $\mathcal{G}(a_\nu, b_\nu)$ e da equação (3.28), temos que a distribuição condicional completa de ν é dada por

$$p(\nu | \lambda_t, \mathbf{h}_{0:T}, \sigma^2, \alpha, \phi) \propto \nu^{a_\nu-1} e^{\nu(b_\nu - \frac{\lambda_t^2(y_t-A_t)^2}{(\nu-1)(1-\rho^2)})}, \quad (3.55)$$

ou seja, é uma $\mathcal{G}_{\nu > 1}(a_\nu, b_\nu - \frac{\lambda_t^2(y_t-A_t)^2}{(\nu-1)(1-\rho^2)})$, isto é, uma distribuição gama truncada. Simulamos desta distribuição gama truncada usando a metodologia do caso VE-MEN1.

• **Caso VE-VG**

Como $\lambda_t \sim \mathcal{IG}(\frac{\nu}{2}, \frac{\nu}{2})$, considerando $f(\nu) = \frac{\nu(\nu+2)}{4}$, a condicional completa de λ_t é dada por

$$p(\lambda_t | y_t, h_t, \nu) \propto \lambda_t^{-\frac{\nu}{2}} e^{-\frac{1}{2}(\frac{\lambda_t(y_t-A_t)^2}{f(\nu)e^{h_t(1-\rho^2)}} + \lambda_t \frac{(h_t-\alpha-\phi h_{t-1})^2}{f(\nu)\sigma^2}) - \frac{\nu}{2\lambda_t}}, \quad (3.56)$$

que é a distribuição gaussiana inversa generalizada, $\mathcal{GIG}(-\frac{\nu}{2}+1, \frac{-1}{2}(\frac{(y_t-A_t)^2}{f(\nu)e^{h_t(1-\rho^2)}} + \frac{(h_t-\alpha-\phi h_{t-1})^2}{f(\nu)\sigma^2}), \nu)$.

Assumimos a distribuição a priori de ν como $\mathcal{G}(a_\nu, b_\nu)\mathbb{I}_{0 < \nu \leq 40}$. Então, a condicional completa de ν é

$$p(\nu | \mathbf{y}_{1:T}, \mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}) \propto \frac{\left[\frac{\nu}{2} \right]^{\frac{T\nu}{2}} \nu^{a_\nu-1} e^{-\frac{\nu}{2} \sum_{t=1}^T [(\frac{1}{\lambda_t} + \log \lambda_t) + 2b_\nu] - \frac{\lambda_t}{2f(\nu)} \left(\frac{(y_t-A_t)^2}{e^{h_t(1-\rho^2)}} + \frac{(h_t-\alpha-\phi h_{t-1})^2}{\sigma^2} \right)}}{\Gamma(\frac{\nu}{2})^T} \mathbb{I}_{0 < \nu \leq 40},$$

que é côncava. Então, amostramos ν pelo algoritmo de aceitação/rejeição Metropolis-Hastings, similarmente aos casos anteriores.

Aplicação

Continuando o exemplo anterior, para os dados da série *S&P500*, neste caso vamos implementar os modelos VE-MEN2. Utilizamos as mesmas distribuições a priori, às quais adicionamos a distribuição Uniforme ($U_{[-1,-0.02]}$) sugerido para o parâmetro ρ do modelo. Considerando as condicionais completas mencionadas acima, nós implementamos o algoritmo de amostrador de Gibbs.

Para o modelo VELE-t-Student realizamos 100.000 simulações, e consideramos 5.000 simulações iniciais para o período de burn-in. No entanto, devido a alta correlação das estimativas dos parâmetros, se escolheu uma amostra sistemática de 25 em 25. No caso da VELE-VG e VELE-Slash, consideramos 10.000 observações iniciais para o período burn-in das 260.000 e 110.000 realizadas respectivamente, logo realizamos uma seleção sistemática de 100 em 100 para VELE-VG e para o modelo VELE-Slash a amostra foi de 50 em 50.

Os gráficos mostram o desempenho do “single move” para a amostra selecionada, assim pode-se observar que as cadeias dos parâmetros estão estáveis ao redor dos valores de suas médias. Para as amostras simuladas selecionadas verificou-se que as funções de autocorrelações decrescem lentamente a zero. Finalmente, calculamos a média da posteriori, o erro padrão Monte Carlo, o intervalo credibilidade do 95% e a estatística de convergência de Geweke (1992) (CD), indicando a convergência da cadeia (veja-se a Tabela 3.3).

Apresentamos nas figuras a seguir (3.6 - 3.8) o desenvolvimento das cadeias simuladas para os diferentes parâmetros.

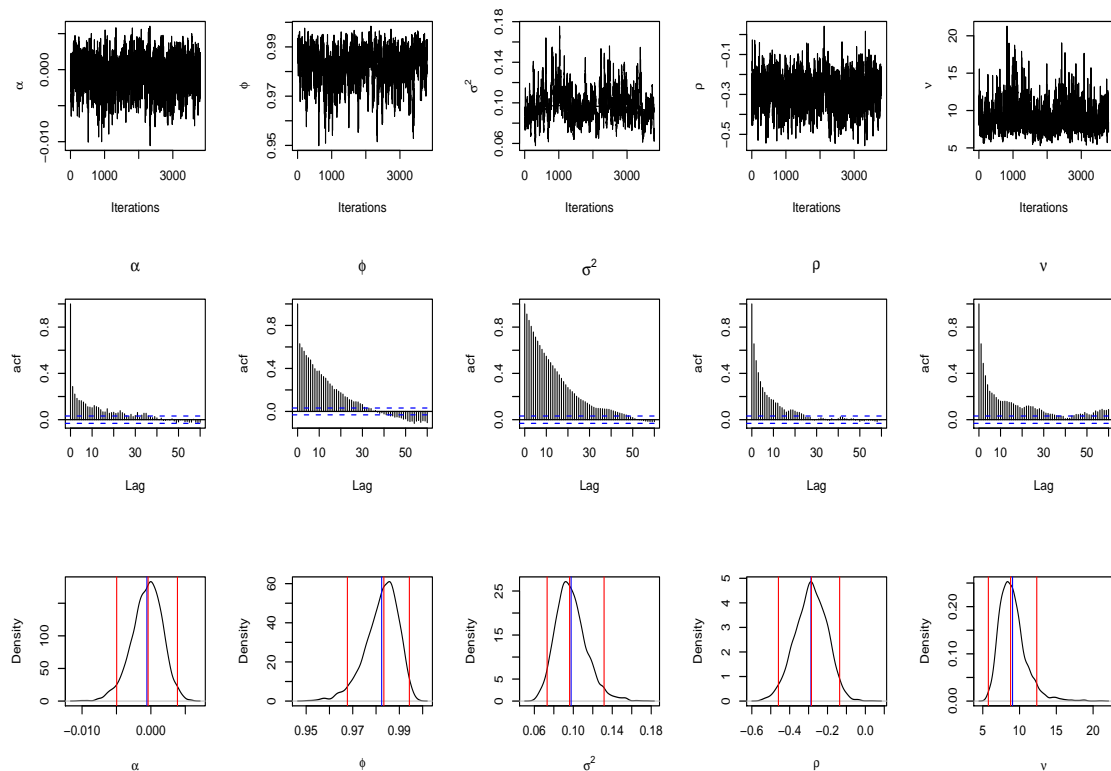


Figura 3.6: Comportamento das simulações dos parâmetros α , ϕ , σ^2 , ρ e ν obtidas do Amostrador de Gibbs para a série de retornos diários S&P500 (VE-T-Student com Efeito Alavanca). Iterações amostradas (Primeira linha), autocorrelações amostradas (Segunda linha), densidade posteriori (Terceira linha). As linhas indicam os percentil 2.5% e 97.5% e as linhas do meio a média posteriori e mediana.

3.3. MODELOS DE VOLATILIDADE ESTOCÁSTICA COM ERRO NA FAMÍLIA MEN 51

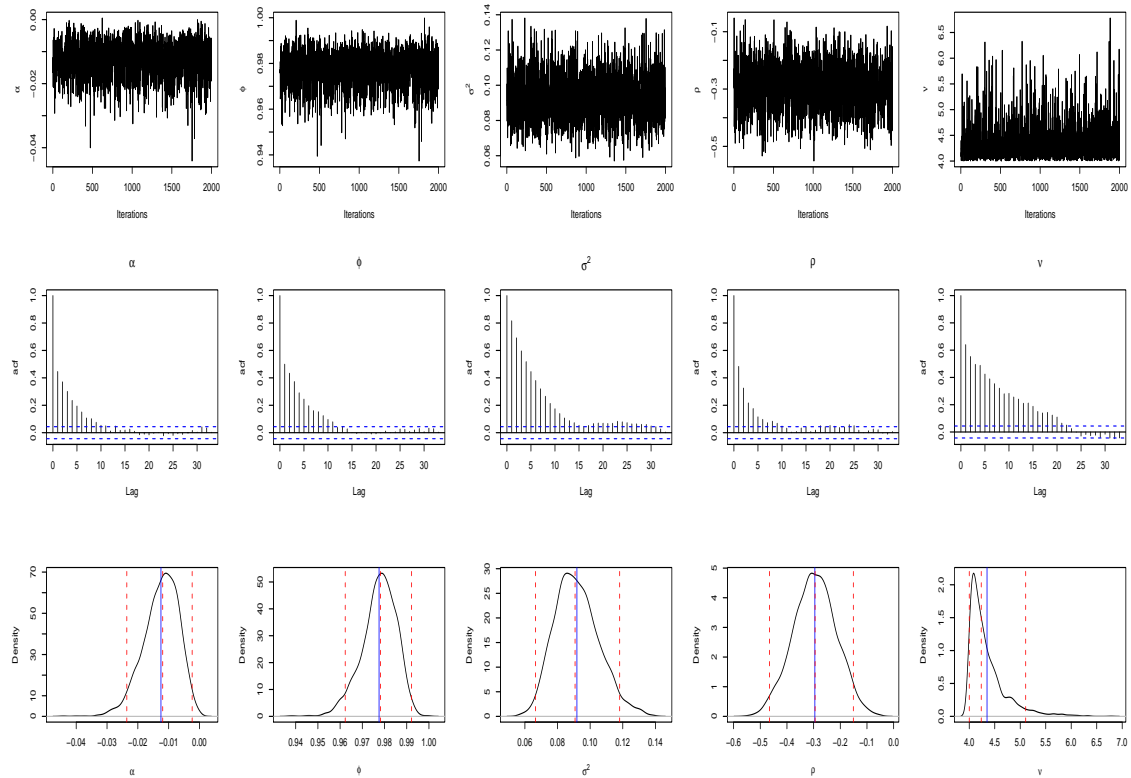


Figura 3.7: Comportamento das simulações dos parâmetros α , ϕ , σ^2 , ρ e ν obtidas do Amostrador de Gibbs para a série de retornos diários S&P500 (VE-Slash com Efeito Alavanca). Iterações amostradas (Primeira linha), autocorrelações amostrais (Segunda linha), densidade posteriori (Terceira linha). As linhas indicam os percentil 2.5% e 97.5% e as linhas do meio a média posteriori e mediana.

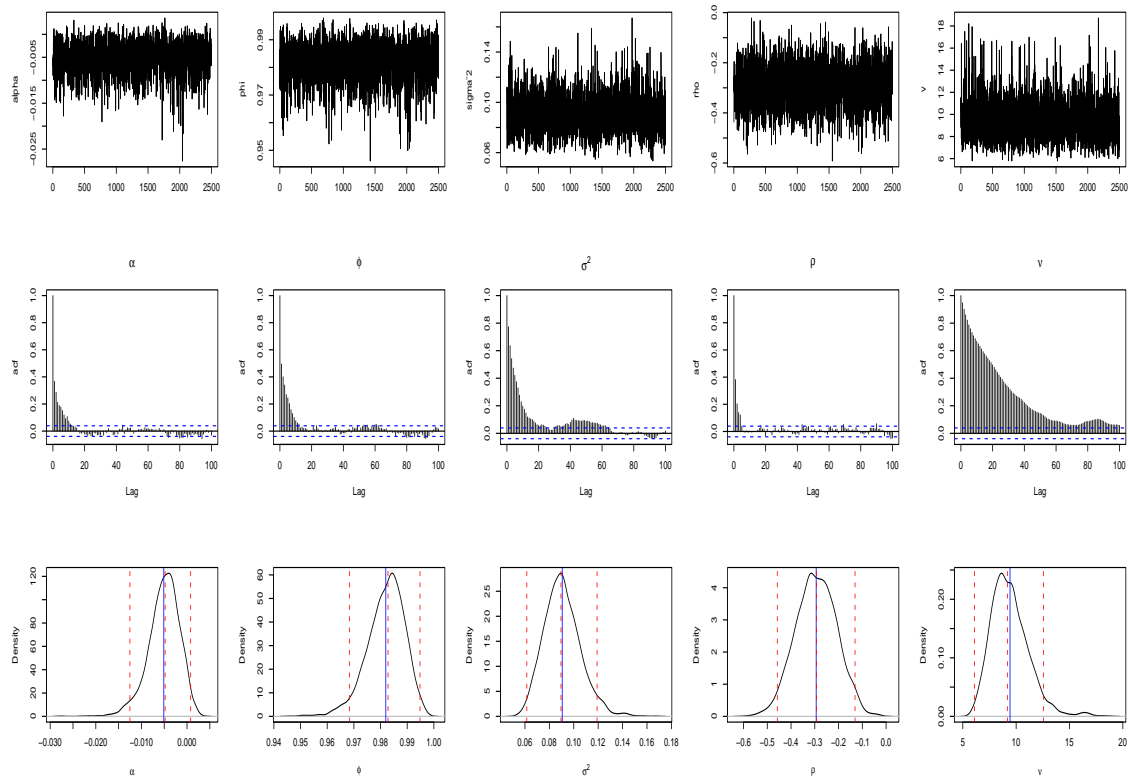


Figura 3.8: Comportamento das simulações dos parâmetros α , ϕ , σ^2 , ρ e ν obtidos do Amostrador de Gibbs para a série de retornos diários S&P500 (VE- Variância Gamma com Efeito Alavanca). Iterações amostradas (Primeira linha), autocorrelações amostrais (Segunda linha), densidade posteriori (Terceira linha). As linhas indicam os percentil 2.5% e 97.5% e as linhas do meio a média posteriori e mediana.

Tabela 3.3: Estimativas dos parâmetros dos modelos VE-MEN2 para a série de retornos S&P500. Primeira linha: média a posteriori. Segunda linha: intervalo de credibilidade de 95% entre parêntesis. Terceira linha: Erro Monte Carlo da média a posteriori. Quarta linha: Estatística de convergência de Geweke (CD)

Parameter	VE-t	VELE-t	VELE-Slash	VELE-VG
α	-0,001455	-0,000578	-0,012506	-0,005089
	(-0,005924,0,003165)	(-0,004974, 0,003859)	(-0,02358,-0,002382)	(-0,001252, 0,0008004)
	$0,234 \times 10^{-2}$	$0,22 \times 10^{-2}$	$0,5753 \times 10^{-2}$	$0,3461 \times 10^{-2}$
	0,4066	0,4231	-1,114	-0,3243
ϕ	0,9838	0,9824	0,9775	0,9819
	(0,9699,0,9956)	(0,9677,0,9943)	(0,9623,0,9922)	(0,9683, 0,9948)
	$0,6845 \times 10^{-2}$	$0,7128 \times 10^{-2}$	$0,7853 \times 10^{-2}$	$7,025 \times 10^{-3}$
	0,8695	0,9714	-0,6289	-1,175
σ^2	0,09342	0,0977	0,0919	0,09056
	(0,06635,0,1228)	(0,07031, 0,1316)	(0,06656,0,1181)	(0,06155, 0,119)
	$1,497 \times 10^{-2}$	$1,58 \times 10^{-2}$	$0,135 \times 10^{-1}$	$0,1497 \times 10^{-1}$
	-0,8543	-1,848	0,7516	1,010
ν	9,374	9,047	4,348	9,426
	(6,001;13,2)	(5,757;12,34)	(4;5,107)	(6,093; 12,57)
	1,94	1,836	0,369	$1,784 \times 10^{-1}$
	1,055	-1,356	-1,640	-1,366
ρ	—	-0,2875	-0,2645	-0,2954
	—	(-0,4594, ;-0,136)	(-0,4657;-0,1511)	(-0,4578;-0,1306)
	—	0,0837	0,0807	$0,85633 \times 10^{-1}$
	—	1,117	0,6036	1,502

Fazendo uso do critério DIC de seleção de modelos, comparamos os modelos VE-MEN2, resultados que se observam na Tabela 3.4. Neste caso, os modelo de mistura na escala normal Variância Gama e Slash resultam ser os mais adequados para os retornos da *S&P500*. Notamos que como no caso dos modelos de VE-MEN1, o modelo adequado é o modelo de misturas na escala de normais, o modelo Variância Gama com efeito alavanca.

Os valores das estimativas dos parâmetros dos diferentes modelos considerados mostram resultados coerentes, assim o parâmetro de persistência é próximo da unidade, o valor de variância do processo da volatilidade é perto de 0.10, para o parâmetro de mistura (ν), os valores resultantes são menores que 10, indicativos de uma diferença significativa com a distribuição normal. Com respeito ao parâmetro ρ , eles resultam ser negativos, confirmando o conceito do efeito alavanca, significativos, como se pode observar nos intervalos de credibilidade dos diferentes modelos que consideram esse parâmetro.

Pode-se concluir, com base nos resultados, que os modelos de volatilidade estocástica com distribuição do erro na mistura na escala resultam ser mais adequados para série financeiras, pois eles capturam pontos mais discrepantes, permitindo identificar esses valores e incorporando-os ao modelo, isto é, os modelos VE-MEN2 resultam ser mais robustos que os

Tabela 3.4: Comparação entre os modelo VE-MEN2: VE-Normal com Efeito alavanca, T-Student com efeito alavanca, VE-Slash com efeito alavanca, e VE-Variância Gamma com efeito alavanca.

Modelo	DIC (pD = var(deviance)/2)	Ordem
VELE-N	5503,15 (pD = 92,19)	4
VELE-T	5392,84 (pD = 386,5)	3
VELE-Slash	5390,10 (pD = 269,1)	2
VELE-VG	5382,8 (pD = 380,6)	1

modelos com erro gaussiano.

Uma forma de identificar esses valores discrepantes é ponderando os erros da série de estimativas dos retornos pelas estimativas da variável de mistura correspondente. As figuras 3.9-3.10 nos mostram como qualquer um dos modelos com MEN2 captura melhor o comportamento da volatilidade que o processo com erro gaussiano com e sem efeito alavanca. Os modelos que incorporam erros distribuídos de mistura na escala de normais são utilizados no sentido de obter modelos mais robustos, que permitam capturar pontos discrepantes dos demais dados. No âmbito dos modelos de volatilidade vai cumprir o mesmo papel, tentando identificar valores de logretornos discrepantes do resto. Nos modelos propostos neste capítulo, os logretornos, modelados através da equação 3.16a, incorporam uma componente, a variável de mistura, que altera o valor estimado da série (série suavizada), de forma tal que possa estimar aqueles dados discrepantes ou distanciados do comportamento padrão do conjunto de dados. Na figura 3.9 observamos os valores das variáveis de mistura para cada um dos modelos VELE, é notório que a Variância Gama apresenta valores menores que o resto dos modelos. Este fato nos faz concluir que a suavização de $e^{\frac{h_t}{2}}$ ponderado pela variável de mistura alcançará maior valor com pequenos valores de λ_t , como se nota nas figuras 3.11-3.10.

3.3. *MODELOS DE VOLATILIDADE ESTOCÁSTICA COM ERRO NA FAMÍLIA MEN 55*

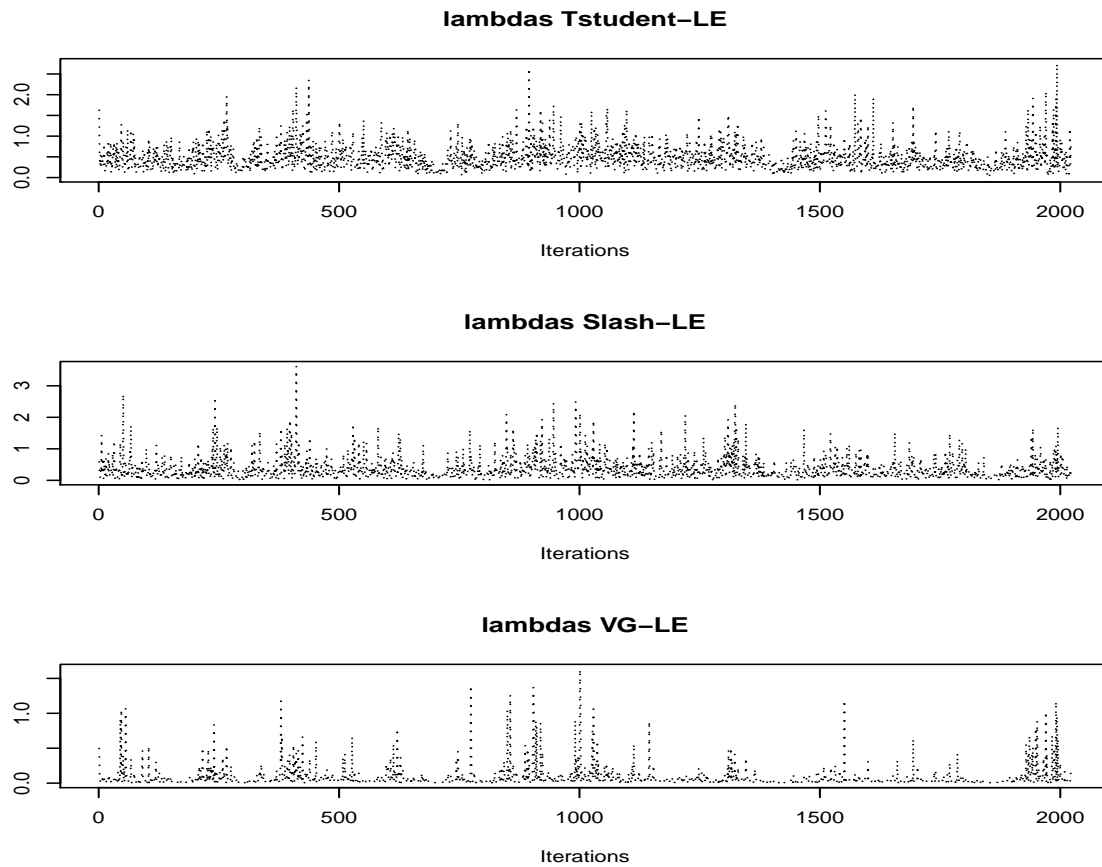


Figura 3.9: Variáveis de mistura dos modelos de volatilidade estocástica com efeito alavanca

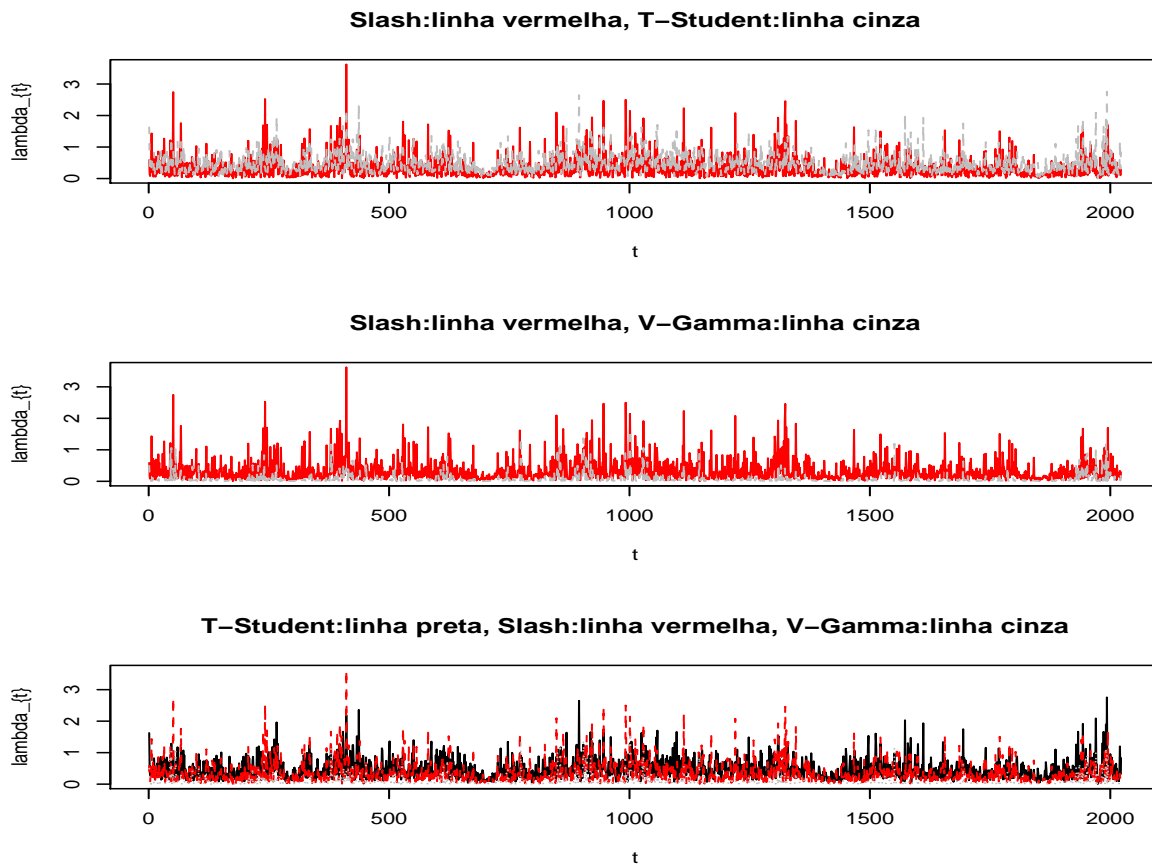


Figura 3.10: Comparação entre as diferentes variáveis de mistura dos modelos VELE. No gráfico superior o modelo VELE-Slash (linha) vs. o modelo VELE-T-Student(tracejada). No gráfico central o modelo VELE-Slash (linha) vs. o modelo VELE-VG(tracejada). No gráfico inferior os tres modelos, VELE-Slash (linha), VELE-VG (pontilhada)e VELE-Slash (tracejada)

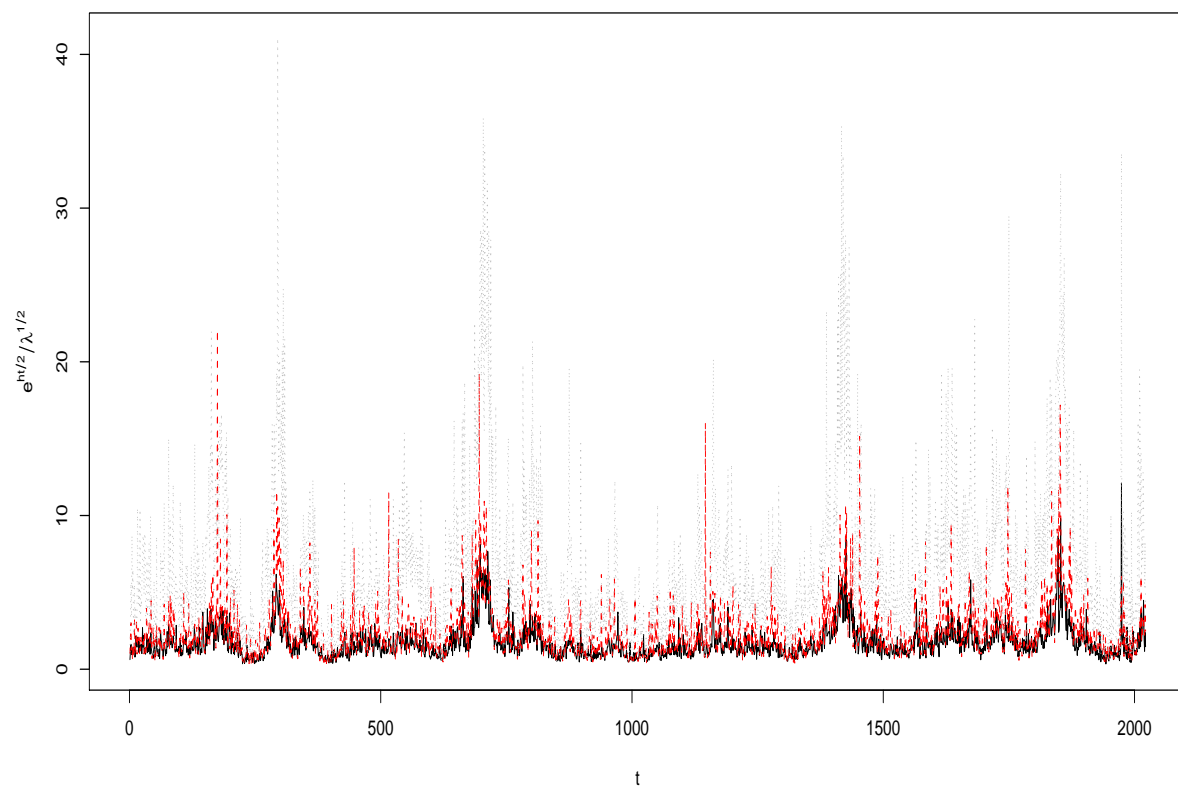


Figura 3.11: Estimativa dos log-retornos nos modelos de volatilidade estocástica com caudas pesadas e efeito alavanca. t-Student (linha), Slash (tracejada), VG (pontilhada).

Capítulo 4

Modelos de Volatilidade Estocástica Normais Assimétricos

Vários modelos foram desenvolvidos dentro do contexto da análise da assimetria nas séries de retornos que apresentam heteroscedasticidade, entre eles podemos mencionar a família de modelos que consideram a volatilidade condicional como função de volatilidades passadas e retornos passados, paralelamente surgem os modelos de volatilidade estocástica (SV), onde a volatilidade é considerada como um processo estocástico markoviano, na maioria das vezes gaussiano. No entanto, alguns modelos têm sugerido a necessidade de incorporar distribuições de caudas mais pesadas comparadas com a normal, pois este tipo de séries apresentam um coeficiente de curtose maior que o da normal. Da mesma forma, as séries de retornos apresentam distribuições levemente assimétricas, o que induz fazer o estudo do comportamento da assimetria no modelo e como ele afeta os resultados finais de previsão.

4.1 Definições Prévias

Definição 4.1.1 (*Azzalini (1985a)*) Uma variável aleatória Z tem distribuição normal assimétrica padrão se sua função de densidade de probabilidade é dada por

$$f_Z(z) = 2\phi(z)\Phi(\lambda z), \quad -\infty < z < +\infty, \quad (4.1)$$

onde $\phi(\cdot)$ e $\Phi(\cdot)$ são as funções de densidade e de distribuição de probabilidade de uma normal padrão, respectivamente. O parâmetro λ ($\lambda \in \Re$) caracteriza a forma da distribuição e também a forma da assimetria, pois valores negativos de λ indicam assimetria negativa e valores positivos assimetria positiva. Se $\lambda = 0$ a densidade acima coincide com a densidade da distribuição normal padrão e, portanto simétrica. Utilizaremos a notação $Z \sim SN(\lambda)$

para denotar que uma v.a tem distribuição normal assimétrica padrão. Uma notação mais completa é $Z \sim SN(0, 1, \lambda)$.

Definição 4.1 [Azzalini \(1985a\)](#), Se $Z \sim SN(\lambda)$ e $Y = \epsilon + \sigma Z$, onde $\epsilon \in \mathfrak{R}, \sigma \in \mathfrak{R}^+$, então temos que $Y \sim SN(\epsilon, \sigma^2, \lambda)$, e sua função de densidade é dada por

$$f_Y(y) = 2\phi\left(\frac{y - \epsilon}{\sigma}\right)\Phi\left(\lambda\frac{y - \epsilon}{\sigma}\right). \quad (4.2)$$

A reparametrização estocástica do modelo segundo [Azzalini \(1985a\)](#) é,

$$Y = \mu + \sigma \frac{\lambda}{\sqrt{1 + \lambda^2}} |X| + \frac{\sigma}{\sqrt{1 + \lambda^2}} Z, \quad (4.3a)$$

$$X \sim N(0, 1) \perp\!\!\!\perp Z \sim N(0, 1), \quad (4.3b)$$

Definimos a transformação $\delta = \frac{\lambda}{(1 + \lambda^2)^{1/2}} \in (-1, 1)$ o qual nos permite uma melhor interpretação da assimetria dos dados. Neste caso denotamos $SN(\mu, \sigma^2, \delta)$. Alguns detalhes podem ser encontrados no Apêndice A. A seguir consideramos algumas observações mais relevantes para nosso modelo de SV.

Como se pode observar nas figuras a seguir, quando o parâmetro de assimetria δ aumenta, aproximando-se de um, a cauda positiva da distribuição normal assimétrica é cada vez mais pesada do que a da distribuição normal, no entanto, a cauda esquerda é mais leve que a normal, sendo limitada pela normal que resulta quando $\delta = 0$, como era de se esperar. Similar interpretação pode-se dar quando analisamos o caso δ negativo. Este aspecto é importante no caso de empregar o modelo para uma análises de quantis, dependentes dos valores das caudas, como por exemplo o calculo do valor em risco.

4.2 Casos Particulares

Deve-se notar que se uma v.a. tem distribuição $SN(\lambda)$ ela é centrada em 0 mas sua esperança é diferente de 0, isto é, se $X \sim SN(\lambda)$ sua função de densidade é

$$f_X(x) = 2\phi(x)\Phi(\lambda x), \quad -\infty < x < +\infty. \quad (4.4)$$

Sua esperança e variância ficam $E[X] = \frac{\lambda}{(1 + \lambda^2)^{1/2}} \sqrt{\frac{2}{\pi}}$ e $Var(X) = \left(1 - \frac{2\lambda^2}{\pi(1 + \lambda^2)}\right)$, i.e, não necessariamente tem média 0 e variância 1. Nas séries de retornos, os erros geralmente são considerados já centrados em zero, para poder seguir permitindo essa característica

consideramos a seguinte distribuição,

$$y \sim SN \left(-\sqrt{\frac{2}{\pi}} \frac{\lambda}{(1 + \lambda^2)^{1/2}}, \sigma^2 \left(1 - \frac{2\lambda^2}{\pi(1 + \lambda^2)} \right)^{-1}, \lambda \right). \quad (4.5)$$

Então a v.a y tem $E[y] = 0$ e $Var(y) = 1$.

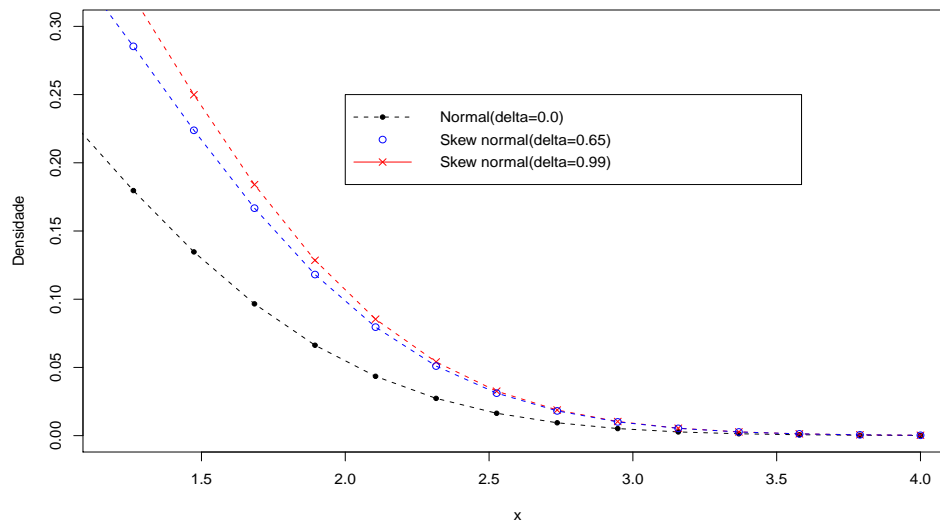


Figura 4.1: Comparação das caudas direitas das distribuições normais assimétricas ($\delta = 0.65$ e 0.99) e da normal padrão.

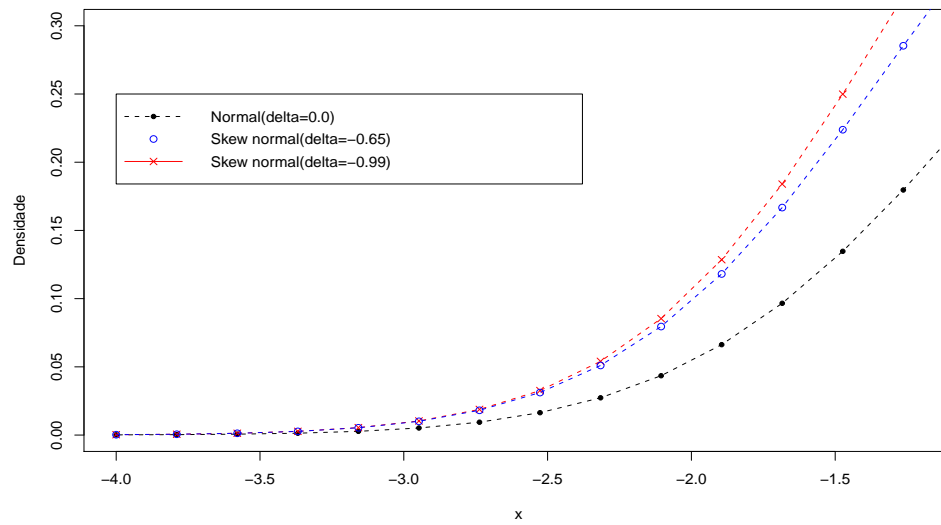


Figura 4.2: Comparação das caudas esquerdas das diferentes normais assimétricas e da normal padrão.

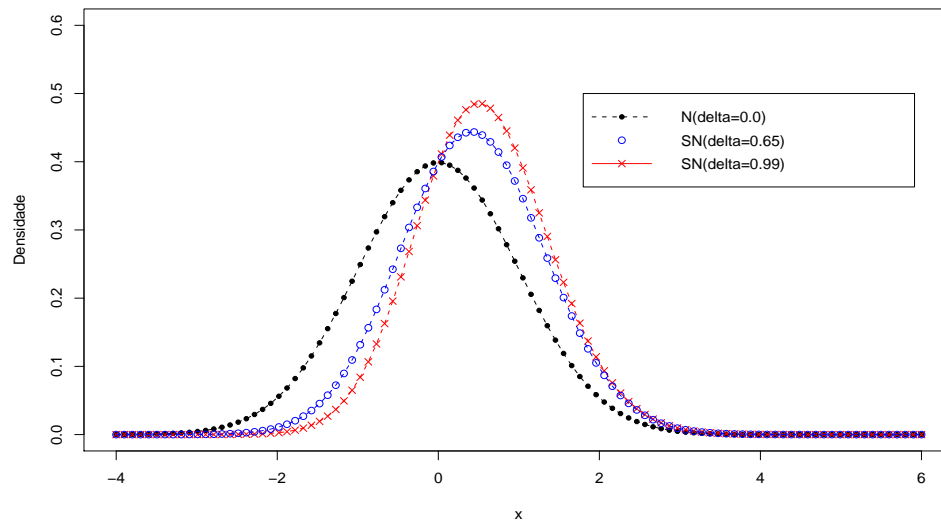


Figura 4.3: Gráfico de normais assimétricas (δ negativo) e da normal padrão.

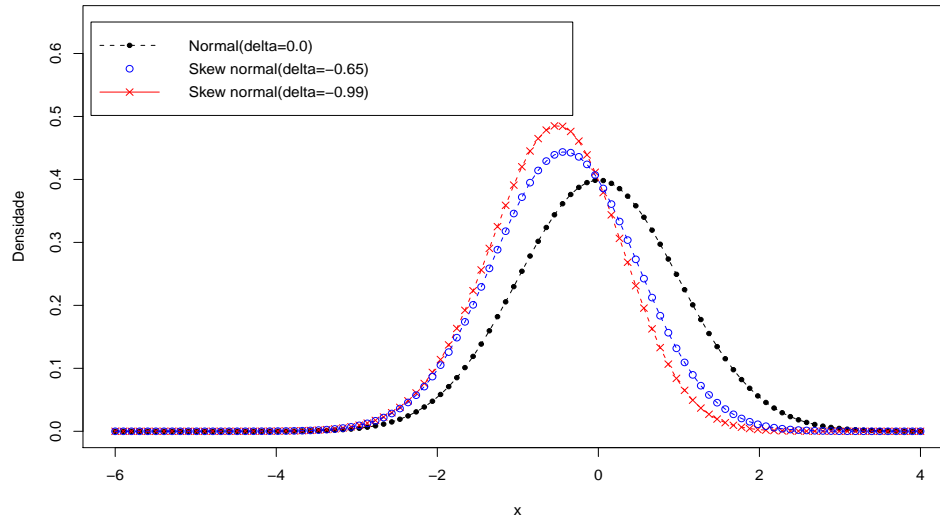


Figura 4.4: Gráfico de normais assimétricas (δ positivo) e da normal padrão.

4.3 Modelo de volatilidade estocástica normal assimétrico

Definimos o seguinte modelo, no qual não consideramos o efeito de alavancagem, para $t = 1, 2, \dots, T$:

$$Y_t = \exp^{\frac{h_t}{2}} \varepsilon_t, \quad (4.6)$$

$$h_t = \alpha + \phi h_{t-1} + \sigma_\eta \eta_t,$$

$$\varepsilon_t = \mu(\delta) + \eta(\delta)U_t + \tau(\delta)Z_t,$$

$$U_t \sim N(0, 1)I_{(0, \infty)},$$

$$Z_t, \eta_t \sim N(0, 1),$$

onde Y_t é a série de log-retornos, h_t é a log-volatilidade, $|\phi| < 1$, $h_0 \sim N(\mu, \frac{\sigma_\eta^2}{1-\phi^2})$, com $\mu = \frac{\alpha}{1-\phi}$, que representa a média do processo estocástico da log-volatilidade. A característica deste modelo é que há um parâmetro de assimetria na série de retornos (δ), pois consideremos um processo normal assimétrico; no entanto a equação dos estados segue sendo um processo gaussiano. Neste caso o vetor do parâmetros de nosso modelo é $\{\alpha, \phi, \sigma_\eta, \delta\}$, e consideramos como variáveis latentes, h_t, U_t e Z_t .

Do ponto de vista Bayesiano podemos usar um algoritmo baseado no amostrador do Gibbs

em cadeias de Markov. Nesse sentido, expressando o modelo na forma hierárquica, isto é, dados $h_t, \delta, \alpha, \phi, \sigma, U_t, Z_t$, para $t = 1, \dots, T$ a distribuição condicional de $Y_t | h_t, \delta, \alpha, \phi, \sigma, U_t, Z_t \sim \mathcal{N}(e^{\frac{h_t}{2}} [\mu(\delta) + \eta(\delta)U_t], \tau^2(\delta))$, temos

$$Y_t | h_t, \delta, \alpha, \phi, \sigma, U_t, Z_t = e^{\frac{h_t}{2}} \varepsilon_t \sim N(\exp^{\frac{h_t}{2}} [\mu(\delta) + \eta(\delta)U_t], \tau^2(\delta)), \quad (4.7)$$

$$\varepsilon_t \sim SN(\mu(\delta); \sigma^2(\delta), \lambda(\delta)),$$

$$h_t \sim N(\alpha + \phi h_{t-1}; \sigma_\eta^2),$$

$$U_t \sim N(0; 1) I_{(0, \infty)},$$

$$Z_t \sim N(0, 1),$$

$$\eta_t \sim N(0, 1).$$

Como desejamos que o processo Y_t seja tal que seu valor esperado seja zero e sua variância igual a um, podemos obter η e τ como segue, $E[Y] = \mu + \sigma \delta \sqrt{\frac{2}{\pi}} = 0$, $\text{Var}(Y) = \sigma^2(1 - \frac{2}{\pi}\delta^2) = 1$, em que a distribuição de $\varepsilon_t \sim SN[\mu(\delta), \tau^2(\delta), \eta(\delta)]$, sendo $\mu(\delta) = -\delta \sqrt{\frac{2}{\pi}(1 - \frac{2\delta^2}{\pi})^{-1/2}}$, $\tau^2(\delta) = (1 - \delta^2)^{-1} (1 - \frac{2\delta^2}{\pi})^{-1}$, $\eta(\delta) = \delta(1 - \frac{2\delta^2}{\pi})^{-1/2}$. Desta forma pode-se implementar um algoritmo de estimação baseado no amostrador de Gibbs. Especificações da normal assimétrica proposta por [Sahu et al. \(2003\)](#) estão nos Apêndices A e B.

4.4 Modelo de volatilidade estocástica com erro normal assimétrico e efeito alavanca

Nesta seção consideramos o modelo de SV com efeito alavanca, onde incorporamos um parâmetro de assimetria na série de retornos ao considerar que estes têm distribuição normal assimétrica. A respeito existem algumas pesquisas no contexto dos modelos GARCH ([Genton et al. \(2004\)](#)). Para o desenvolvimento do modelo precisamos fazer uso da normal assimétrica multivariada e algumas de suas propriedades (veja o Apêndice B ou [Azzalini e Valle \(1996\)](#), [Azzalini \(2005\)](#) e [Sahu et al. \(2003\)](#)). A diferença destas duas formas de expressar a distribuição normal assimétrica reside basicamente na forma paramétrica de suas representações. Neste contexto, ambas distribuições são iguais no caso univariado.

Definição 4.2 ([Azzalini e Valle \(1996\)](#)) Dizemos que Y tem distribuição normal assimétrica multivariada se

$$\mathbf{Y} = \boldsymbol{\mu} + \text{Diag}(\boldsymbol{\eta}) |\mathbf{U}| + \mathbf{Z},$$

4.4. MODELO DE VOLATILIDADE ESTOCÁSTICA COM ERRO NORMAL ASSIMÉTRICO E EFEITO

onde $U \sim N(0, I)$, $Z \sim N(0, \Sigma)$, $\mathbf{Y} \in \mathfrak{R}^n$, Σ matriz simétrica definida positiva, e escrevemos $\mathbf{Y} \sim SN(\boldsymbol{\mu}, \Sigma, \boldsymbol{\eta})$.

A literatura reconhece esta distribuição como normal assimétrica de Azzalini, ou fundamental. Nesta forma da distribuição SN, a estrutura de correlação e assimetria estão relacionadas, desde que a matriz de variâncias covariâncias depende de Σ como de $\boldsymbol{\eta}$.

Definição 4.3 (Sahu *et al.* (2003)) Seja $\boldsymbol{\varepsilon}$ e \mathbf{U} vetores aleatórios m -dimensionais, seja $\boldsymbol{\mu}$ um vetor m -dimensional e Σ uma matriz definida positiva. Suponha que

$$\begin{pmatrix} \boldsymbol{\varepsilon} \\ \mathbf{Z} \end{pmatrix} \sim El \left(\begin{pmatrix} \boldsymbol{\mu} \\ 0 \end{pmatrix}, \Omega = \begin{pmatrix} \Sigma & 0 \\ 0 & \mathbf{I} \end{pmatrix}, g^{(2m)} \right), \quad (4.9a)$$

isto é, $\begin{pmatrix} \boldsymbol{\varepsilon} \\ \mathbf{Z} \end{pmatrix}$, têm distribuição de contorno elíptico. Se consideramos a transformação

$$\mathbf{Y} = \text{Diag}(\boldsymbol{\delta}) |\mathbf{Z}| + \boldsymbol{\varepsilon},$$

onde $\text{Diag}(\boldsymbol{\delta})$ é uma matriz diagonal com elementos δ_i , a classe de distribuições elípticas assimétricas é desenvolvida considerando a variável aleatória $[\mathbf{Y}|\mathbf{Z} > 0]$, onde $\mathbf{Z} > 0$ implica que cada $z_i > 0$, para toda i -ésima componente de \mathbf{Z} . Notamos que se $\delta = 0$, nós estamos gerando valores de uma distribuição elíptica. Denotamos esta distribuição como $SN(\boldsymbol{\mu}, \Sigma, D)$ ou SN_{Sahu} .

Como a matriz D é assumida diagonal, a introdução da assimetria não afeta a estrutura de correlação. Isto não acontece com a distribuição proposta por Azzalini e Capitanio(1999). Este fato é importante para o modelo de volatilidade estocástica com erro normal assimétrico e efeito alavanca ao implementar a metodologia de estimação dos parâmetros.

Proposição 4.4 Se

$$\mathbf{Y} \sim SN(\boldsymbol{\mu}, \Sigma, \boldsymbol{\eta}),$$

com $Y \in \mathfrak{R}^n$ e $X = \mathbf{A}\mathbf{Y}$, então, $X \sim SN(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\Sigma\mathbf{A}^T, \mathbf{A}\boldsymbol{\eta})$ fundamental. Note que se $\mathbf{A}\boldsymbol{\eta}$ é uma matriz diagonal $n \times n$ então $X \sim SN_{(Sahu)}$

Considerando $\boldsymbol{\varepsilon}^*$ e $\boldsymbol{\eta}^*$ distribuídos como na equação (4.6), temos,

$$\begin{aligned} \boldsymbol{\varepsilon}^* &\sim SN(\boldsymbol{\mu}(\boldsymbol{\delta}); \tau^2(\boldsymbol{\delta}); \boldsymbol{\eta}(\boldsymbol{\delta})), \\ \boldsymbol{\eta}^* &\sim N(0, 1). \end{aligned}$$

Então, pode-se expressar as equações na forma matricial,

$$\begin{pmatrix} \eta^* \\ \varepsilon^* \end{pmatrix} = \begin{pmatrix} 0 \\ \mu(\delta) \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \eta(\delta) \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} + \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}, \quad (4.10a)$$

onde

$$z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim N \left(0; \begin{pmatrix} 1 & 0 \\ 0 & \tau^2(\delta) \end{pmatrix} \right) \quad (4.11)$$

e $\mu(\delta)$, $\tau^2(\delta)$ e $\eta(\delta)$ são as mesmas deduzidas que no caso univariado, isto é, os parâmetros de localização, escala e assimetria reparametrizados em função do parâmetro δ .

Pela Definição 4.1 temos,

$$\begin{pmatrix} \eta^* \\ \varepsilon^* \end{pmatrix} \sim SN \left(\begin{pmatrix} 0 \\ \mu(\delta) \end{pmatrix}; \begin{pmatrix} 1 & 0 \\ 0 & \tau^2(\delta) \end{pmatrix}; \begin{pmatrix} 0 \\ \eta(\delta) \end{pmatrix} \right), \quad (4.12)$$

onde o valor esperado e variância são:

$$\mathbf{E} \begin{bmatrix} \eta^* \\ \varepsilon^* \end{bmatrix} = 0, \quad \mathbf{Var} \begin{bmatrix} \eta^* \\ \varepsilon^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (4.13)$$

Usando o resultado anterior, consideramos a seguinte proposição.

Proposição 4.5 Seja $\begin{pmatrix} \eta_t^* \\ \varepsilon_t^* \end{pmatrix}$ a variável definida como em (4.12), para todo $t = 1, \dots, T$, e considere

$$\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} \eta_t^* \\ \varepsilon_t^* \end{pmatrix}. \quad (4.14a)$$

Então, temos:

$$\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} \sim SN_{Sahu} \left(\begin{pmatrix} 0 \\ \sqrt{1-\rho^2}\mu(\delta) \end{pmatrix}; \begin{pmatrix} 1 & \rho \\ \rho & \rho^2 + (1-\rho^2)\tau^2(\delta) \end{pmatrix}; \begin{pmatrix} 0 \\ \sqrt{1-\rho^2}\eta(\delta) \end{pmatrix} \right) \quad (4.15a)$$

e o valor esperado e variância ficam

$$\mathbf{E} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix} = 0, \quad \mathbf{Var} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \quad (4.16)$$

4.4. MODELO DE VOLATILIDADE ESTOCÁSTICA COM ERRO NORMAL ASSIMÉTRICO E EFEITO

Além disso, o parâmetro de assimetria é dado por

$$\lambda = \frac{\eta(\delta)}{\sqrt{\tau^2(\delta) + \frac{\rho^2}{1-\rho^2}}}, \quad (4.17)$$

onde $\delta \perp \rho \in [-1, 1]$. Logo, pode-se escrever

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim SN \left(\begin{pmatrix} \sqrt{1-\rho^2}\mu(\delta) \\ 0 \end{pmatrix}; \begin{pmatrix} \sqrt{1-\rho^2}\eta(\delta) \\ 0 \end{pmatrix}; \begin{pmatrix} \rho^2 + (1-\rho^2)\tau^2(\delta) & \rho \\ \rho & 1 \end{pmatrix} \right) \quad (4.18a)$$

e

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} | U_t \sim N \left(\begin{pmatrix} \sqrt{1-\rho^2}[\mu(\delta) + \eta\delta U_t] \\ 0 \end{pmatrix}; \begin{pmatrix} \rho^2 + (1-\rho^2)\tau^2(\delta) & \rho \\ \rho & 1 \end{pmatrix} \right). \quad (4.19)$$

Por isso, o modelo SV normal assimétrico com efeito alavanca pode ser deduzido como segue.

Proposição 4.6 Seja,

$$\begin{aligned} Y_t &= \exp^{h_t/2} \varepsilon_t, \\ h_{t+1} &= \alpha + \phi h_t + \sigma_\eta \eta_{t+1}, \end{aligned} \quad (4.20)$$

onde, $\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix}$ é um vetor aleatório definido como na proposição anterior. Então podemos escrever o modelo de volatilidade estocástica de efeito alavanca como

$$\begin{pmatrix} Y_t \\ h_t \end{pmatrix} | U_t, h_t, F_t = \begin{pmatrix} 0 \\ \alpha + \phi h_t \end{pmatrix} + \begin{bmatrix} \exp^{h_t/2} & 0 \\ 0 & \sigma_\eta \end{bmatrix} \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} | U_t, \quad (4.21)$$

ou seja,

$$\begin{pmatrix} Y_t \\ h_t \end{pmatrix} | U_t, h_t, F_t \sim N \left(\begin{pmatrix} \sqrt{1-\rho^2}[\mu(\delta) + \eta\delta U_t] \\ \alpha + \phi h_{t-1} \end{pmatrix}; \begin{pmatrix} \exp^{h_t} [\rho^2 + (1-\rho^2)\tau^2(\delta)] & \rho\sigma_\eta \exp^{h_t/2} \\ \rho\sigma_\eta \exp^{h_t/2} & \sigma_\eta^2 \end{pmatrix} \right).$$

Assim, o modelo expresso em forma hierárquica fica

$$\begin{aligned}
Y_t | h_{t+1} &\sim N \left(\exp^{h_t/2} \sqrt{1 - \rho^2} [\mu(\delta) + \eta\delta U_t] + \frac{\rho \exp^{h_t/2}}{\sigma_\eta} (h_{t+1} - \alpha - \phi h_t); \exp^{h_t} (1 - \rho^2) \tau^2(\delta) \right), \\
h_{t+1} &\sim N(\alpha + \phi h_t; \sigma^2(\delta)), \\
U_t &\sim N(0; 1) I_{(0, \infty)}, \\
Z_t &\sim N(0, 1), \\
\eta_t &\sim N(0, 1).
\end{aligned} \tag{4.22a}$$

$$\varepsilon_t \sim SN(\mu(\delta); \sigma^2(\delta), \lambda(\delta)), \tag{4.23}$$

$$h_t = \alpha + \phi h_{t-1} + \sigma_\eta \eta_t, \tag{4.24}$$

$$U_t \sim N(0; 1) I_{(0, \infty)}, \tag{4.25}$$

$$Z_t \sim N(0, 1), \tag{4.26}$$

$$\eta_t \sim N(0, 1), \tag{4.27}$$

isto é a distribuição de $\eta_t \sim SN[\mu(\delta), \tau^2(\delta), \eta(\delta)]$, onde $\mu(\delta) = -\delta \sqrt{\frac{2}{\pi}} (1 - \frac{2\delta^2}{\pi})^{-1/2}$, $\tau^2(\delta) = (1 - \delta^2)^{-1} (1 - \frac{2\delta^2}{\pi})^{-1}$, $\eta(\delta) = \delta (1 - \frac{2\delta^2}{\pi})^{-1/2}$. Finalmente, podemos desenvolver o algoritmo de estimação baseado no amostrador de Gibbs.

4.5 Estimação

Para estimar os parâmetros nos modelos SV-SN definidos pelas equações de (4.6) utilizamos a técnica MCMC. Seja $\boldsymbol{\theta}$ o vetor de parâmetros de SV-SN, sejam $\mathbf{h}_{0:T} = (h_0, h_1, \dots, h_T)'$ o vetor de log-volatilidades, $U_{1:T} = (U_1, \dots, U_T)'$ as variáveis latentes, $\mathbf{y}_{1:T} = (y_1, \dots, y_T)'$ a informação avaliada até o tempo t e consideremos $\mathbf{h}_{0:T}$ e $U_{1:T}$ como parâmetros latentes. Usando o Teorema de Bayes, a densidade a posteriori dos parâmetros e das variáveis latentes fica

$$p(\mathbf{h}_{0:T}, U_{1:T}, \boldsymbol{\theta} | \mathbf{y}_{1:T}) \propto p(\mathbf{y}_{1:T} | \mathbf{h}_{0:T}, U_{1:T}) p(\mathbf{h}_{0:T} | \boldsymbol{\theta}) p(U_{1:T} | \boldsymbol{\theta}) p(\boldsymbol{\theta}), \tag{4.28}$$

onde

$$p(\mathbf{y}_{1:T} | U_{1:T}, \mathbf{h}_{0:T}, \boldsymbol{\theta}) \propto \prod_{t=1}^T \tau(\delta) e^{-\frac{y_t - e^{h_t/2}(\mu(\delta) + \eta(\delta)U_t)^2}{2\tau(\delta)^2}}, \quad (4.29)$$

$$p(\mathbf{h}_{0:T} | \boldsymbol{\theta}) \propto e^{-\frac{1-\phi^2}{2\sigma_\eta^2}(h_0 - \frac{\alpha}{1-\phi})^2} \prod_{t=1}^T e^{-\frac{1}{2\sigma_\eta^2}(h_t - \alpha - \phi h_{t-1})^2}, \quad (4.30)$$

$$p(U_{1:T} | \boldsymbol{\theta}) = \prod_{t=1}^T e^{-\frac{(U_t)^2}{2}} I(0, \infty), \quad (4.31)$$

e $p(\boldsymbol{\theta})$ é a distribuição a priori. Para os parâmetros comuns da classe de modelos SV–SN, as distribuições a priori são dadas como: $\alpha \sim \mathcal{N}(\bar{\alpha}, \sigma_\alpha^2)$, $\phi \sim \mathcal{N}_{(-1,1)}(\bar{\phi}, \sigma_\phi^2)$ e $\sigma_\eta^2 \sim \mathcal{IG}(\frac{T_0}{2}, \frac{M_0}{2})$, onde $\mathcal{N}_{(a,b)}(\cdot, \cdot)$ denota a distribuição normal truncada no intervalo (a,b).

Como a densidade a posteriori $p(\mathbf{h}_{0:T}, U_{1:T}, \boldsymbol{\theta} | \mathbf{y}_{1:T})$ não tem forma fechada, primeiramente amostramos $\boldsymbol{\theta}$, depois as variáveis latentes $U_{1:T}$ e $\mathbf{h}_{0:T}$ usando o amostrador de Gibbs. O algoritmo é descrito a seguir:

Algoritmo 4.1

1. Para $(i = 0)$ inicializamos os valores dos parâmetros; $\boldsymbol{\theta}^{(i)}$, os estados $\mathbf{U}_{1:T}^{(i)}$ e $\mathbf{h}_{0:T}^{(i)}$;
2. Amostramos $\boldsymbol{\theta}^{(i+1)}$ de $p(\boldsymbol{\theta} | \mathbf{h}_{0:T}^{(i)}, \mathbf{U}_{1:T}^{(i)}, \mathbf{y}_{1:T})$;
3. Amostramos $\mathbf{U}_{1:T}^{(i+1)}$ de $p(\mathbf{U}_{1:T} | \boldsymbol{\theta}^{(i+1)}, \mathbf{h}_{0:T}^{(i)}, \mathbf{y}_{1:T})$;
4. Amostramos $\mathbf{h}_{0:T}^{(i+1)}$ de $p(\mathbf{h}_{0:T} | \boldsymbol{\theta}^{(i+1)}, \mathbf{U}_{1:T}^{(i+1)}, \mathbf{y}_{1:T})$;
5. Fazemos $i = i + 1$ e repetimos até que convirja.

Como descrito no algoritmo 4.1, o Gibbs sampler requer das variáveis latentes as condicionais completas para amostrar os parâmetros .

Detalhes das condicionais completas de $\boldsymbol{\theta}$ e das variáveis latentes $U_{1:T}$ são dadas a seguir. As distribuições a priori dos parâmetros mencionados são consideradas como acima. De acordo com (4.30), a distribuição condicional para α :

$$p(\alpha | \mathbf{h}_{0:T}, \phi, \sigma_\eta^2) \propto \exp\left\{-\frac{a_\alpha}{2}\left(\alpha - \frac{b_\alpha}{a_\alpha}\right)^2\right\}, \quad (4.32)$$

que é uma distribuição normal com média $\frac{b_\alpha}{a_\alpha}$ e variância $\frac{1}{a_\alpha}$, onde $a_\alpha = \frac{1}{\sigma_\alpha^2} + \frac{T}{\sigma_\eta^2} + \frac{1+\phi}{\sigma_\eta^2(1-\phi)}$ e $b_\alpha = \frac{\bar{\alpha}}{\sigma_\alpha^2} + \frac{(1+\phi)}{\sigma_\eta^2}h_0 + \frac{\sum_{t=1}^T(h_t - \phi h_{t-1})}{\sigma_\eta^2}$.

De forma similar, de (4.30), a distribuição condicional de ϕ é dada por

$$p(\phi \mid \mathbf{h}_{0:T}, \alpha, \sigma_\eta^2) \propto Q(\phi) \exp\left\{-\frac{a_\phi}{2\sigma_\eta^2}(\phi - \frac{b_\phi}{a_\phi})^2\right\} \mathbb{I}_{|\phi| < 1}, \quad (4.33)$$

onde $Q_\phi = \sqrt{1 - \phi^2} \exp\left\{-\frac{1}{2\sigma_\eta^2}[(1 - \phi^2)(h_0 - \frac{\alpha}{1-\phi})^2]\right\}$, $a_\phi = \sum_{t=1}^T h_{t-1}^2 + \frac{\sigma_\eta^2}{\phi^2}$, $b_\phi = \sum_{t=1}^T h_{t-1}(h_t - \alpha) + \bar{\phi} \frac{\sigma_\eta^2}{\phi^2}$ e $\mathbb{I}_{|\phi| < 1}$ é uma variável indicadora. Como $p(\phi \mid \mathbf{h}_{0:T}, \alpha, \sigma_\eta^2)$ em (4.33) não tem forma fechada, amostramos usando o algoritmo Metropolis-Hastings com a densidade truncada proposta $\mathcal{N}_{(-1,1)}(\frac{b_\phi}{a_\phi}, \frac{\sigma_\eta^2}{a_\phi})$.

De (4.30), a posteriori condicional de σ_η^2 é $\mathcal{IG}(\frac{T_1}{2}, \frac{M_1}{2})$, onde $T_1 = T_0 + T + 1$ e $M_1 = M_0 + [(1 - \phi^2)(h_0 - \frac{\alpha}{1-\phi})^2] + \sum_{t=1}^T (h_t - \alpha - \phi h_{t-1})^2$.

Finalmente para U_t temos,

$$p(U_t \mid \mathbf{h}_{0:T}, \phi, \sigma_\eta^2, \delta) \propto \exp\left\{-\frac{a_{U_t}}{2}(U_t - \frac{b_{U_t}}{a_{U_t}})^2\right\} I_{(0,\infty)}, \quad (4.34)$$

que é uma distribuição normal truncada com média $\frac{b_{U_t}}{a_{U_t}}$ e variância $\frac{1}{a_{U_t}}$, onde $a_{U_t} = \frac{e^{h_t} \eta(\delta) + \tau^2(\delta)}{\tau^2(\delta)}$ e $b_{U_t} = \mu(\delta) \eta(\delta) e^{h_t} - y_t e^{h_t/2} \eta(\delta)$.

Para o caso SV-SNLE, o algoritmo será similar, como indicado a seguir. Seja $\boldsymbol{\theta}$ o vetor de parâmetros do modelo SV-SNLE, sejam $\mathbf{h}_{0:T} = (h_0, h_1, \dots, h_T)'$ o vetor de log volatilidades, $U_{1:T} = (U_1, \dots, U_T)'$ variáveis latentes com distribuição normal truncada, $\mathbf{y}_{1:T} = (y_1, \dots, y_T)'$ a informação avaliada até o tempo t , consideremos $\mathbf{h}_{0:T}$ and $U_{1:T}$ como parâmetros latentes. Usando o Teorema de Bayes, a densidade a posteriori dos parâmetros e das variáveis latentes pode ser escrita como

$$p(\mathbf{h}_{0:T}, U_{1:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T}) \propto p(\mathbf{y}_{1:T} \mid \mathbf{h}_{0:T}, U_{1:T}) p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta}) p(U_{1:T} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}), \quad (4.35)$$

onde

$$p(\mathbf{y}_{1:T} | U_{1:T}, \mathbf{h}_{0:T}, \boldsymbol{\theta}) \propto \prod_{t=1}^T \exp^{-h_t/2} [\rho^2 + (1 - \rho^2)\tau^2(\delta)]^{-1/2} e^{-\frac{y_t - e^{h_t/2}(\sqrt{1-\rho^2})(\mu(\delta) + \eta(\delta)U_t)^2}{2\tau(\delta)^2}} \quad (4.36)$$

$$p(\mathbf{h}_{0:T} | \boldsymbol{\theta}) \propto e^{-\frac{1-\phi^2}{2\sigma_\eta^2}(h_0 - \frac{\alpha}{1-\phi})^2} \prod_{t=1}^T e^{-\frac{1}{2\sigma_\eta^2}(h_t - \alpha - \phi h_{t-1})^2}, \quad (4.37)$$

$$p(U_{1:T} | \boldsymbol{\theta}) = \prod_{t=1}^T e^{-\frac{(U_t)^2}{2}} I(0, \infty), \quad (4.38)$$

e $p(\boldsymbol{\theta})$ é a distribuição a priori. O algoritmo Gibbs Sampler é similar ao caso SV-SN. Detalhes das condicionais completas são mostrados no Apêndice B.

4.6 Aplicação

Utilizando a série de retornos S&P500 dos capítulos anteriores, fazemos uma aplicação da metodologia mencionada neste capítulo. Como nos casos anteriores, utilizamos as mesmas prioris da literatura, que temos utilizado nesta tese, isto é, considerando $\alpha = \mu(1 - \phi)$, então $\mu \sim \mathcal{N}(0; 0, 25)$, $\phi^* \sim \text{Be}(20, 1.5)$ com $\phi^* = \frac{\phi+1}{2}$ e $\sigma_\eta^2 \sim \mathcal{IG}(2.5, 0.025)$. Para o parâmetro de assimetria, $\delta \sim \mathcal{U}(-1, 1)$. Na estimação dos parâmetros consideramos o amostrador de Gibbs, com as prioris sugeridas. Realizamos igual número de simulações que nos outros casos (100000) e consideramos o mesmo período “burn-in” de 5000 iterações. No entanto, e devido a alta correlação das amostras geradas dos parâmetros escolhemos uma amostra sistemática de 320 em 320 para cada parâmetro. As Figuras 4.5-4.7 mostram o desempenho do “single move”, assim pode-se observar que as cadeias encontram-se estáveis ao redor dos valores de suas médias, enquanto as funções de autocorrelações decaem para zero. Finalmente, calculamos a média da posteriori, o erro padrão Monte Carlo, o intervalo de credibilidade de 95% e a estatística de convergência CD (veja a Tabela 5.3)

As estimativas dos parâmetros (α, ϕ, σ^2) são coerentes com resultados encontrados na literatura. A média a posteriori de ϕ é próxima da unidade, em ambos modelos, fato que é conhecido devido a alta persistência da volatilidade na série dos log-retornos corrigidos, entretanto para o parâmetro σ^2 a VE-SN se aproxima melhor do valor correspondente de outros modelos. Analisando o parâmetro de assimetria (δ) deve-se destacar que a série de retornos apresenta assimetria muito baixa, o que é corroborado pelos resultados. Entretanto, no parâmetro do efeito de alavancagem (ρ), o modelo VELE-SN aproxima de forma similar aos modelos VE-MEN2 o qual é acorde com valores dados na literatura.

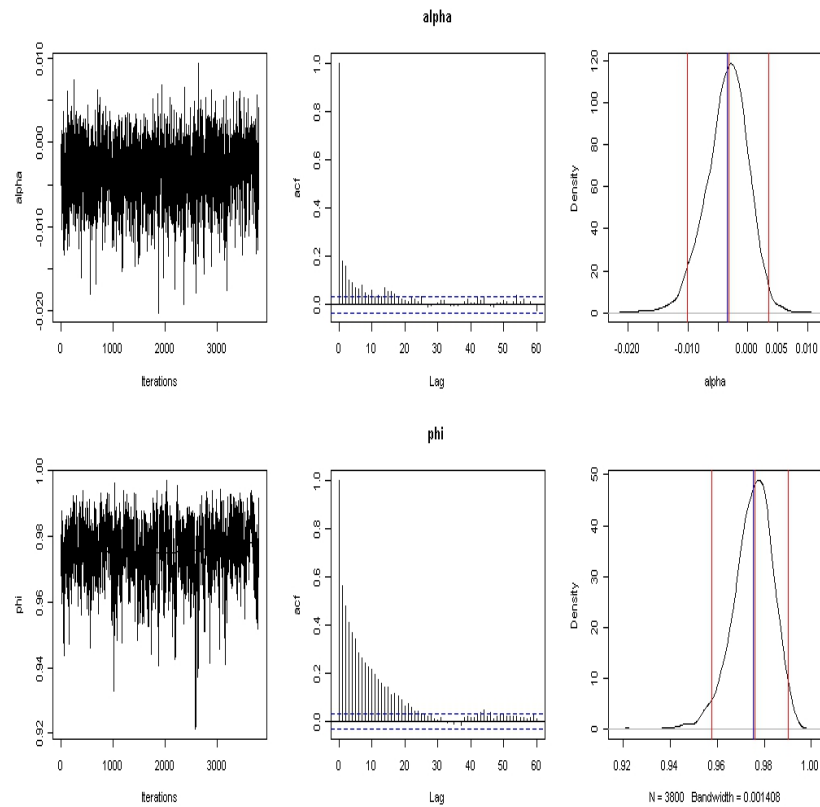


Figura 4.5: Comportamento das simulações dos parâmetros α e ϕ , no MCMC, para a série de retornos diários S&P500 (VE-Normal Assimétrica). Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2.5% e 97.5% e as linhas do meio a média a posteriori e mediana.

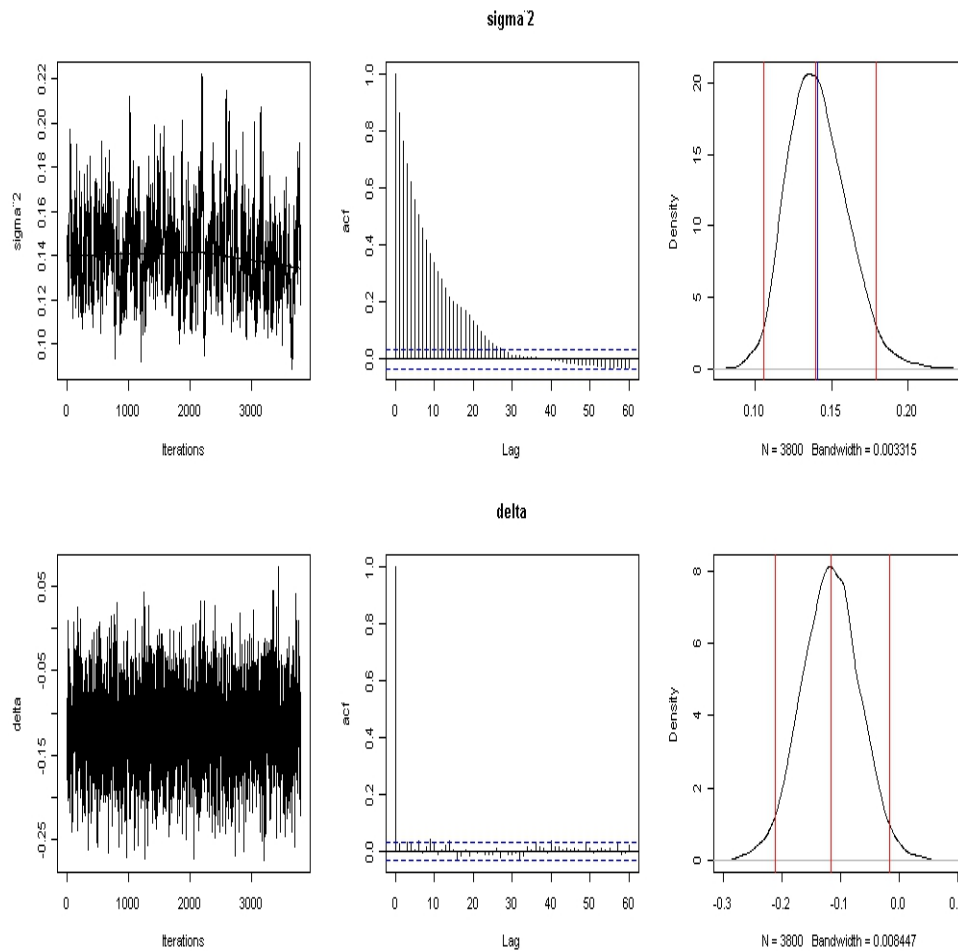


Figura 4.6: Comportamento das simulações dos parâmetros σ^2 e δ no MCMC, para a série de retornos diários S&P500 (VE-Normal Assimétrica). Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2.5% e 97.5% e as linhas do meio a média a posteriori e mediana.

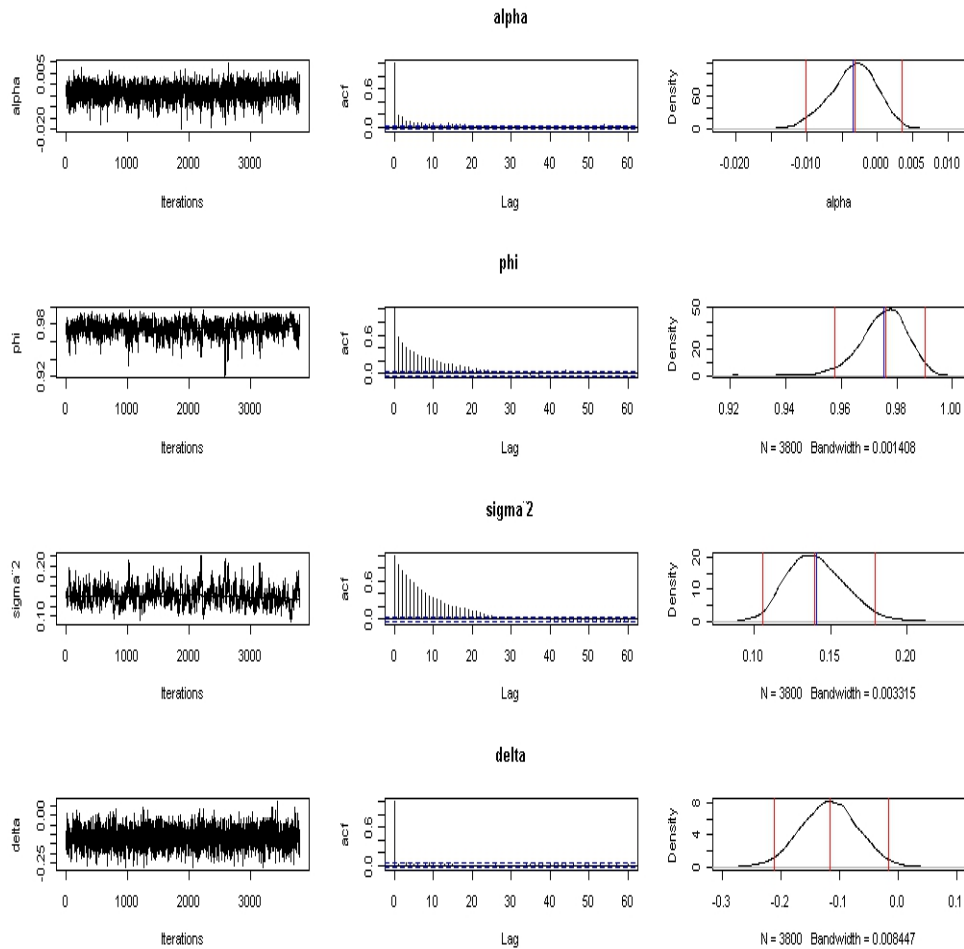


Figura 4.7: Comportamento das simulações dos parâmetros α , ϕ , σ^2 e δ , no MCMC, para a série de retornos diários S&P500 (VE-Normal Assimétrica). Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

Tabela 4.1: Estimativas dos parâmetros dos modelos VE-N e VE-SN para a série de retornos S&P500 return. Primeira linha: Média a posteriori. Segunda linha: Intervalo de credibilidade a posteriori, nível de confiança de 95% entre parênteses. Terceira linha: O erro Monte Carlo da média a posteriori. Quarta linha: estatística de convergência de Geweke(CD).

Parâmetro	VE-N	VE-SN	VELE-SN
α	-0,00335 (-0,0109,0,0027) $0,34 \times 10^{-4}$ -0,009435	-0,003354 (-0,01199, 0,003059) $0,3848 \times 10^{-2}$ -0,3416	-0,0019164 (-0,009609, 0,003696) $0,3564083 \times 10^{-2}$ 0,3397
ϕ	0,9748 (0,9577,0,9904) $3,04 \times 10^{-4}$ -0,886	0,9755 (0,9547, 0,9899) $0,90998 \times 10^{-1}$ -1,737	0,9729 (0,9557, 0,99) $0,09298353 \times 10^{-1}$ -0,3594
σ^2	0,1492 (0,1058,0,1798) $0,19 \times 10^{-1}$ 1,79	0,13902 (0,09834,0,1786) $0,2061 \times 10^{-1}$ -0,8931	0,1480 (0,1058,0,1833) $0,21023 \times 10^{-1}$ -0,58
δ	— — — —	-0,028 (-0,6279, 0,625) 0,38518 0,09968	-0,2241 (-0,8048, 0,4938) 0,40586 -1,452
ρ	— — — —	— — — —	-0,333 (-0,5075, -0,1369) 0,092986 -0,3292

Tabela 4.2: SP&500 return data set. DIC: deviance information criterion.

Modelo	DIC		pD
	Valor	Ordem	Valor
VE-N	6889,6	1	93,159
VE-SN	8000,6	3	58,0
VELE-SN	7830,3	2	135,5

Comparamos os modelos fazendo uso do critério DIC de seleção de modelos, resultados que se observam na Tabela 4.2. Neste caso o modelo Normal resulta ser o mais adequado para os retornos da *S&P500*. Devemos de ter presente que a comparação entre estes modelos é válida desde que estes pertencem à mesma família de modelos de VE, que comparados com os modelos de misturas na escala de normais, têm um único elemento comum, o modelo com distribuição Normal no ruído. Devido aos resultados e para efeitos onde se possa comparar os modelos, pode-se considerar uma família de modelos maior, os modelos de volatilidade Skew-elípticos,(ver Branco e Day (2001)), família que permitirá o estudo conjunto das caudas pesadas e a assimetria dos retornos em forma conjunta.

Capítulo 5

Simulações

5.1 Introdução

Nesta seção apresentamos resultados de o caso simulações realizadas para o modelo de volatilidade estocástica considerados anteriormente. Implementamos a metodologia do "Mixture Sampler", desenvolvida no Capítulo 1, para obter dados simulados de um processo de volatilidade estocástica com erro t-Student. Escolhemos a distribuição t-Student para o ruído do modelo uma vez que é uma das mais utilizada na literatura. Nós geramos 3000 valores com $\alpha=0,002$, $\phi= 0,97$, $\sigma= 0,10$, $\rho= -0,3$ e $\nu= 15$, valores utilizados por vários autores, veja por exemplo [Ando \(2006\)](#), [Ando \(2007\)](#), [Nakajima e Omori \(2009\)](#).

Implementamos os diferentes modelos, tanto da família VE-MEN1, como da VE-MEN2, como foi feito no capítulo anterior, para modelar a série de retornos S&P500. Neste caso empregando a série de 3000 dados gerada, com ruído t-Student e efeito alavanca, simulados via metologia "Mixture Sampler" utilizada por [Kim *et al.* \(1998\)](#).

Estimamos os parâmetros para os diferentes modelos de caudas pesadas, e os resultados das simulações resultam valores próximos aos verdadeiros valores dos parâmetros. Estimativas do parâmetro α convergem a 0, aproximando-se mais ao modelos de difusão, para o caso contínuo. Para o parâmetro de persistência de volatilidade (ϕ), o modelo que mais se aproxima do real valor é o modelo VE-MEN1-VG. Observando os resultados com respeito a volatilidade das log volatilidades (σ^2), as estimativas nos modelos são próximas a seu verdadeiro valor, destacando os modelos VE-t, VELE-Slash e VELE-t, no entanto o modelo Normal com e sem efeito alavanca super estima o valor verdadeiro.

No que se refere às estimativas do parâmetro da variável de mistura ($\nu = 15$), os modelos com erro t-Student com efeito alavanca (VELE-t) conservam o valor do parâmetro relati-

vamente próximo; já os demais modelos de mistura de escala de normais resultam em uma convergência do parâmetro próxima de 18, super-estimando de forma nada desprezível o valor verdadeiro, ou sub-estimando o valor, como podemos ver nos modelos Slash. Dentre os modelos com efeito alavanca, a convergência da correlação dos ruídos (ρ) resulta adequada só no modelo VELE-t. Resumindo o modelo VELE-t, recupera de forma adequada os valores dos parâmetros do modelo simulado, sendo este, como se esperava.

Nós comparamos os resultados com o critério DIC, pois os critérios AIC e BIC, consideram o número de parâmetros do modelo, e neste tipo de modelo essa quantidade não é fixa (veja-se Ando,2006). Devido ao fato que a implementação do ponto de vista clássico é pouco factível, pela forma da distribuição conjunta dos parâmetros, utilizamos os métodos de MCMC para as estimativas dos parâmetros (algoritmo Gibbs e Metropolis-Hastings). A implementação é feita de forma hierárquica, o que implica incorporar no modelo certo número de parâmetros ou tantos parâmetros quantos se disponham, para cada variável latente considerada. Na Tabela 5.6 pode-se observar a comparação dos diferentes modelos expostos nas simulações.

Entre os modelos da família VE-MEN1, Tabela 5.5, o melhor modelo segundo o DIC resulta ser o VG-MEN1 seguido do Slash-MEN1, e entre os modelos com efeito alavanca VE-MEN2, Tabela 5.4, o melhor modelo resultou o VELE-VG, seguido do modelo VELE-t, que corresponde aos dados. Esse resultado nos indica que o critério DIC não é um bom critério para os modelos de VE-MEN, ou com caudas pesadas.

Tabela 5.1: Estimativas dos parâmetros dos modelos da família VE-MEN2 para série simulada do modelo VELE-t. Primeira linha: média a posteriori. Segunda linha: intervalo de credibilidade posterior de 95% entre parêntesis. Terceira linha: Erro Monte Carlo da média a posteriori. Quarta linha: Estatística de convergência CD.

Parâmetro	VELE-N	VELE-t	VELE-S	VELE-VG
$\alpha=0,002$	-0,0047909 (-0,0186, 0,006436) $0,643 \times 10^{-2}$ 0,5047	-0,001049 (-0,001103, 0,001566) $0,6339 \times 10^{-2}$ -0,02737	-0,049536 (-0,1795, -0,01777) $0,4954 \times 10^{-1}$ -0,5747	0,0033388 (-0,06761, -0,003191) $0,19632 \times 10^{-1}$ 0,5308
$\phi=0,97$	0,8437 (0,7231, 0,9424) $0,6422 \times 10^{-1}$ -1,380	0,8133 (0,6083, 0,9339) $0,9158 \times 10^{-1}$ 0,4133	0,8304 (0,6941, 0,9486) $0,7872 \times 10^{-1}$ -0,8503	0,8069 (0,6504, 0,9543) $0,9453 \times 10^{-1}$ 0,2807
$\sigma^2=0,10$	0,15157 (0,07887, 0,242) $0,4483 \times 10^{-2}$ 1,611	0,09535 (0,05631, 0,1462) $2,53 \times 10^{-2}$ 0,8575	0,1012 (0,0515, 0,1579) $0,29086 \times 10^{-1}$ 1,906	0,08443 (0,05196, 0,1247) $0,203186 \times 10^{-1}$ -0,04172
$\nu=15$	— — — —	14,78 (7,49, 24,44) 0,2114 -0,643	4,643 (2,695, 7,459) 1,348 1,801	12,977 (8,359, 18,61) $3,021 \times 10^{-1}$ -0,961
$\rho=-0,3$	0,1634 (-0,04603, 0,3507) 0,10663 -0,4167	-0,3349 (-0,4594, -0,136) 4,6943 -1,403	0,1957 (-0,1229, 0,6217) 0,17449 0,7422	-0,4336 (-0,02865, -0,8674) $2,341 \times 10^{-1}$ 1,270

Tabela 5.2: Estimativas dos parâmetros dos modelos da família VE-MENI para série simulada do modelo VELE-t. Primeira linha: média a posteriori. Segunda linha: intervalo de credibilidade posterior de 95% entre parêntesis. Terceira linha: Erro Monte Carlo da média a posteriori. Quarta linha: Estatística de convergência CD

Parameter	VE-N	VE-t	VE-MENI-t	VE-MENI-Slash	VE-MENI-VG
$\alpha=0,002$	-0,004131	-0,0001882	-0,0002236	-0,06248	-0,0006989
	(-0,0176,0,005311)	(-0,01061,0,007935)	(-0,01035,0,0155)	(-0,1198,-0,0169)	(-0,03135,0,0268)
	$0,6296 \times 10^{-2}$	$0,448 \times 10^{-2}$	$0,55083 \times 10^{-2}$	$0,2956 \times 10^{-2}$	$0,5204554 \times 10^{-2}$
$\phi=0,97$	-1,019	-0,5202	0,2094	0,6453	0,498
	0,8518	0,8670	0,8579	0,8634	0,8719
	(0,7353,0,9476)	(0,7475,0,9532)	(0,683,0,9491)	(0,7504,0,9553)	(0,7599,0,9645)
$\sigma^2=0,10$	$0,62349 \times 10^{-1}$	$0,6544 \times 10^{-1}$	$0,7344 \times 10^{-1}$	$0,6216 \times 10^{-1}$	$5,731522 \times 10^{-3}$
	-1,3971	0,5708	0,1581	-1,57	-0,0649
	0,1501	0,10023	0,1048	0,11754	0,10519
$\nu^2=15$	(0,08281,0,2362)	(0,06207,0,1528)	(0,05618,0,1609)	(0,05435,0,1854)	(0,06282,0,1508)
	$0,3986 \times 10^{-1}$	$2,61 \times 10^{-2}$	$0,3192 \times 10^{-2}$	$0,3598 \times 10^{-1}$	$0,58539417 \times 10^{-1}$
	0,7946	0,08876	-0,2988	1,865	0,6484
$\nu=15$	—	18,725	18,11	4,958	18,58
	—	(10,91,27,59)	(9,855,27,31)	(3,002,8,966)	(11,09,28,445)
	—	5,07	4,49	0,19554	$3,21727647 \times 10^{-1}$
$\nu=15$	—	0,0927	1,031	1,618	-1,46

Tabela 5.3: Estimativas dos parâmetros dos modelos VE-N e VE-SN para simulada do modelo VELE-t. Primeira linha: Média a posteriori. Segunda linha: Intervalo de credibilidade a posteriori, nível de confiança de 95% entre parênteses. Terceira linha: O erro Monte Carlo da média a posteriori. Quarta linha: estatística de convergência de Geweke(CD).

Parâmetro	VE-N	VE-SN
α	-0,00335	-0,003943
	(-0,0109,0,0027)	(-0,01583; 0,006204)
	$0,34 \times 10^{-4}$	$0,5822967 \times 10^{-2}$
	-0,009435	0,1184
ϕ	0,9748	0,8473
	(0,9577,0,9904)	(0,7097; 0,9331)
	$3,04 \times 10^{-4}$	$0,05886 \times 10^{-1}$
	-0,886	1,140
σ^2	0,1492	0,15761
	(0,1058,0,1798)	(0,09222; 0,2338)
	$0,19 \times 10^{-1}$	$0,388121 \times 10^{-1}$
	1,79	-1,225
δ	—	-0,02971
	—	(-0,6066; 0,5414)
	—	0,338112
	—	-1,450

5.2 Comparação de estatísticas de diagnóstico

Fazendo uso do critério DIC de seleção de modelos, rodamos duas vezes os modelos sob as condições acima. Comparando os modelos VE-MEN2, observamos em uma primeira rodada dos modelos, o modelo de mistura na escala de normal Slash com efeito alavanca tem desempenho melhor nas simulações realizadas. Na segunda implementação dos modelos (Tabela 5.4), o melhor resultado foi do modelo Variância Gama com efeito alavanca. Além disso, no caso dos modelos de VE-MEN1, o modelo adequado resultou ser o modelo Slash sem efeito alavanca para a primeira corrida, já na segunda, como se mostra na Tabela 5.5, o melhor segundo o DIC, foi o Variância Gama.

No geral todos os modelos que considera o efeito alavanca apresentam melhores resultados que os modelos sem correlação. Observando as Tabelas 5.4 e 5.6 vemos que os modelos Variância Gama e Slash considerando o efeito alavanca deveriam ser escolhidos como os possíveis modelos adequados. No entanto, lembremos que os dados foram gerados do modelo de volatilidade estocástica com efeito alavanca e erro t-Student(VELE-t) Isso nos faz pensar que o critério DIC pode não ser adequado para a comparação dos modelos, pois não identifica o modelo do qual procedem os dados.

Como alternativa para a comparação dos modelos, calculamos o Fator de Bayes, sendo

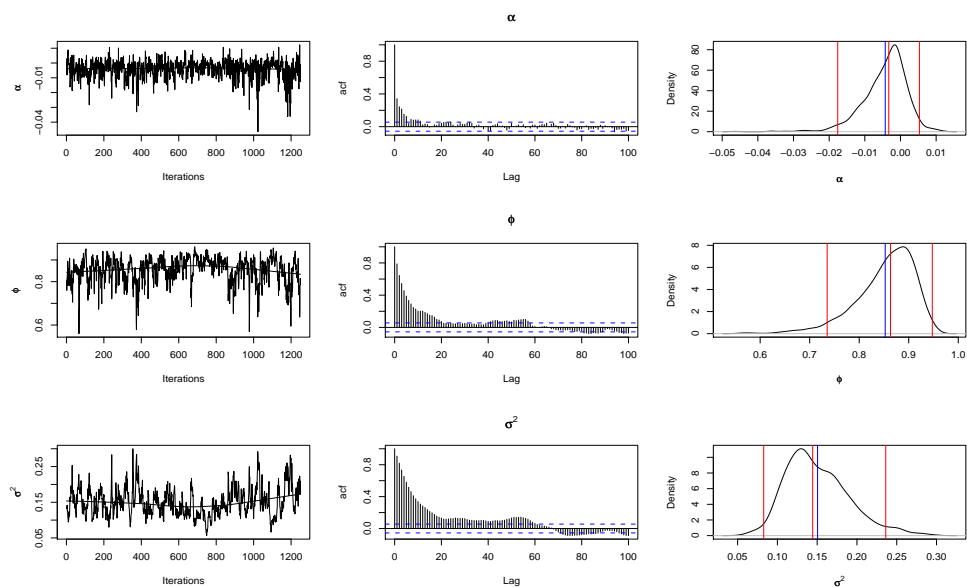


Figura 5.1: Comportamento das simulações dos parâmetros no modelo VE com erro normal. Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

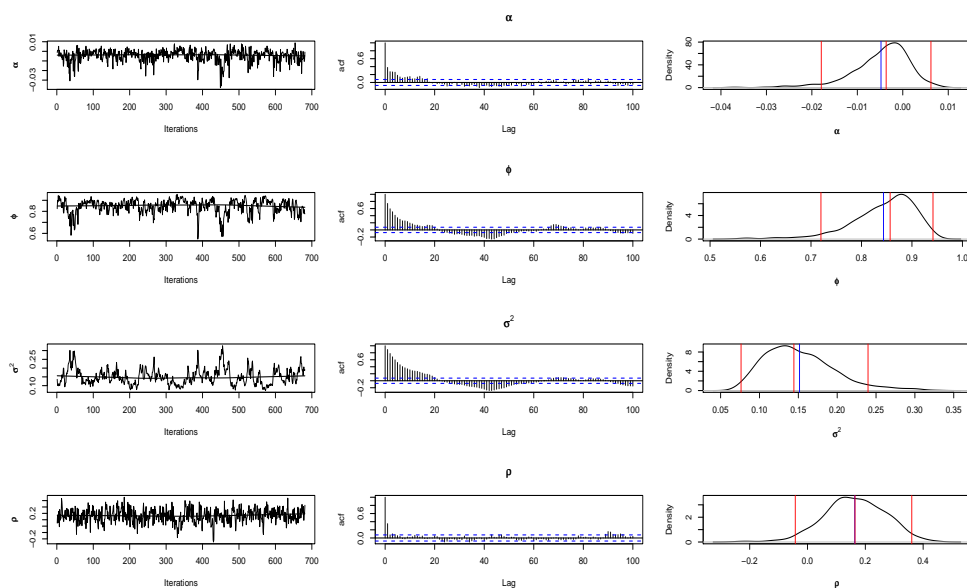


Figura 5.2: Comportamento das simulações dos parâmetros no modelo normal com efeito aleatória. Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

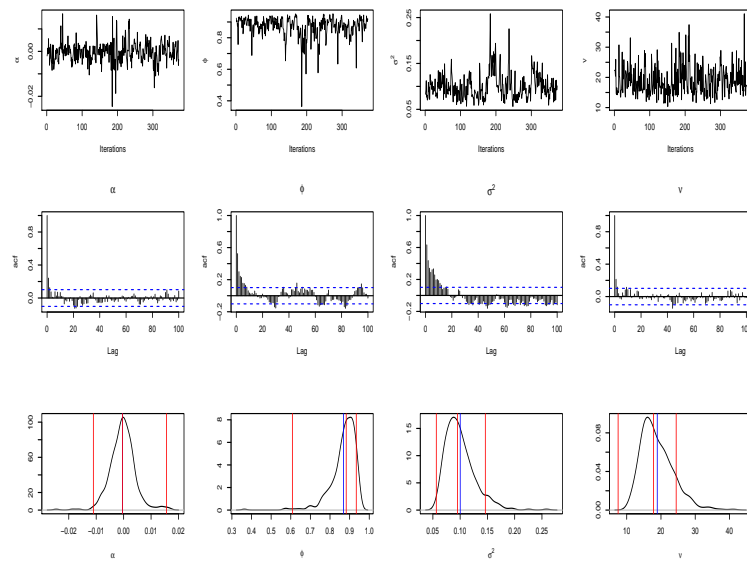


Figura 5.3: Comportamento das simulações dos parâmetros do modelo VE- t-Student. Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

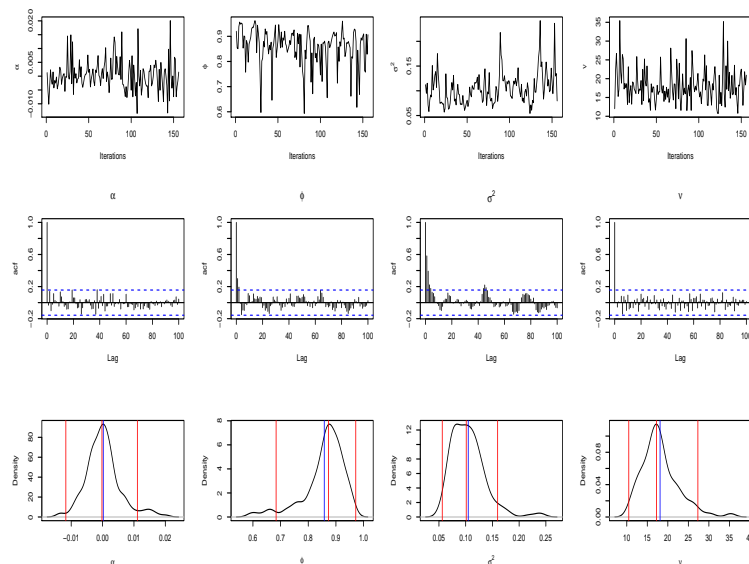


Figura 5.4: Comportamento das simulações dos parâmetros do modelo VE- t-Student-MEN1 na série. Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

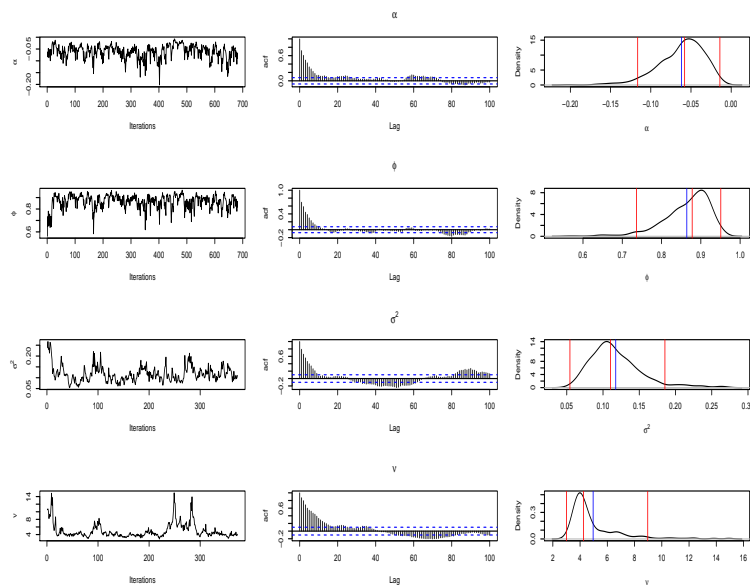


Figura 5.5: Comportamento das simulações dos parâmetros do modelo VE-Slash-MEN1. Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

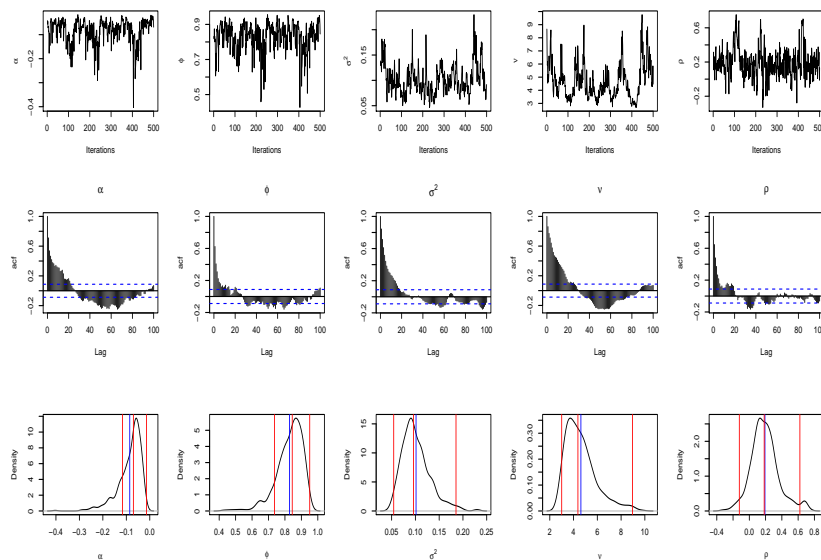


Figura 5.6: Comportamento das simulações dos parâmetros do modelo VE-Slash com efeito alavanca. Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

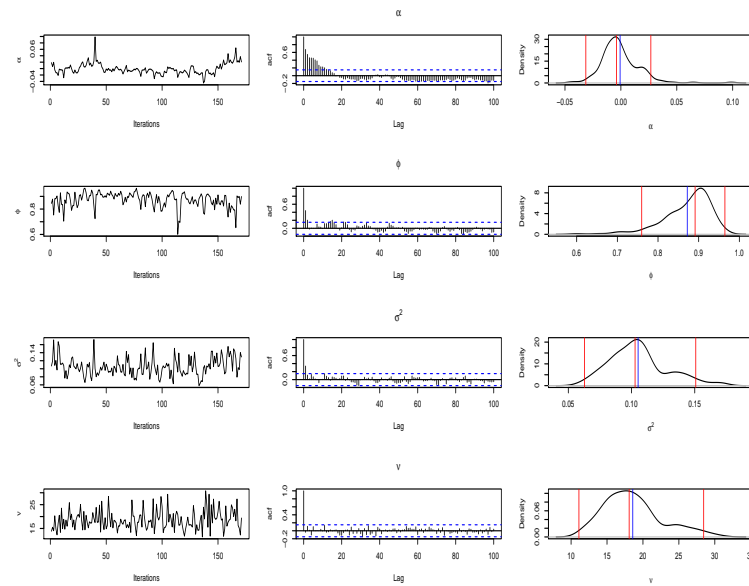


Figura 5.7: Comportamento das simulações dos parâmetros do modelo VE-VG-MEN1. Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

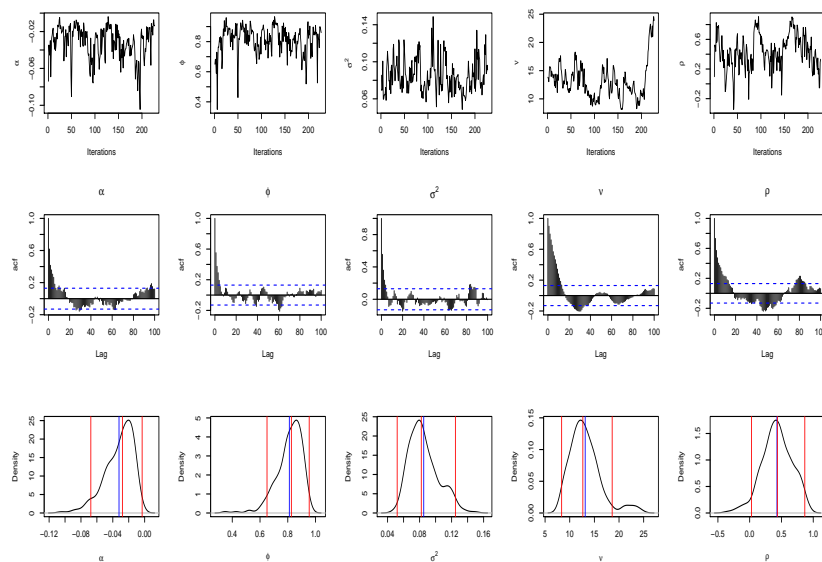


Figura 5.8: Comportamento das simulações dos parâmetros do modelo VE-VG com efeito alavanca. Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

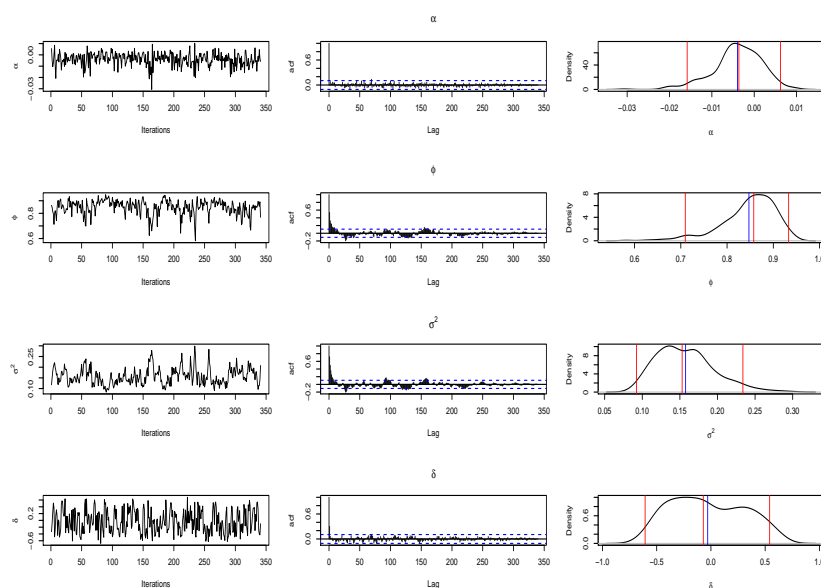


Figura 5.9: Comportamento das simulações dos parâmetros do modelo VE-Normal Assimétrica. Iterações amostradas (esquerda), autocorrelações amostrais (centro), densidade a posteriori (direita). As linhas indicam os percentis 2,5% e 97,5% e as linhas do meio a média a posteriori e mediana.

o melhor modelo o que tem maior chance de ocorrência. Isto é, dado o quociente entre a verossimilhança do modelo considerando os parâmetros do modelo A, com a verossimilhança do modelo considerando os parâmetros do modelo B, o fator de Bayes é,

$$FB = \frac{\text{verossimilhança}(\text{Modelo A})}{\text{verossimilhança}(\text{Modelo B})} \quad (5.1)$$

Na Tabela 5.7, temos os logaritmos dos fatores de Bayes, que mostram que o modelo VELE-t tem maior chance de ocorrência que os outros considerados.

Tabela 5.4: Comparação entre os modelo VE-MEN2: VE-Normal com Efeito alavanca, t-student com efeito alavanca, VE-Slash com efeito alavanca, e VE-Variância Gama com efeito alavanca.

Modelo	DIC (pD)	Ordem	SQE	Ordem
VELE-N	8469,1 (pD = 157,2)	4	3608,791	1
VELE-t	8005,6 (pD = 290,1)	2	3957,005	4
VELE-Slash	8281,7 (pD = 136,8)	3	5839,404	3
VELE-VG	7677,2 (pD = 400,8)	1	3655,092	2

Tabela 5.5: Comparação entre os modelo VE-MEN1: VE-Normal, t-student, VE-Slash, e VE-Variância Gama sem efeito alavanca.

Modelo	DIC (pD)	Ordem	SQE	Ordem
VE-N	8525,0 (pD = 98,40)	3	3739,097	4
VE-t	8525,4 (pD = 210,7)	4	12489,24	2
VE-tMEN1	8526,7 (pD = 214,0)	5	11817,56	1
VE-SlashMEN1	8502,7 (pD = 123,3)	2	5970,211	5
VE-VGMEN1	8490,9 (pD = 171,9)	1	3291,116	3

Tabela 5.6: Comparação entre os modelo VE-MEN2: VE-Normal com Efeito alavanca, t-student com efeito alavanca, VE-Slash com efeito alavanca, VE-Variância Gama com efeito alavanca, e o Normal.

Modelo	DIC (pD)	Ordem	SQE	Ordem
VE-N	8525,0 (pD = 98,40)	7	3739,097	4
VELE-N	8469,1 (pD = 157,2)	4	3608,791	2
VE-t	8525,4 (pD = 210,7)	8	12489,24	9
VELE-t	8005,6 (pD = 290,1)	2	9957,005	7
VE-TMEN1	8526,7 (pD = 214,0)	9	11817,56	8
VELE-Slash	8281,7 (pD = 136,8)	3	5839,404	5
VE-SlashMEN1	8502,7 (pD = 123,3)	6	5970,211	6
VELE-VG	7677,2 (pD = 400,8)	1	3655,092	3
VE-VGMEN1	8490,9 (pD = 171,9)	5	3291,116	1

Tabela 5.7: Logaritmos dos fatores de Bayes.

	N	LE-N	LE-S	S	VG	LE-VG	T	LE-t	SkewN
N	0,00	-441,08	-454,65	-63,51	-91,16	-499,08	-109,76	-536,99	-311,69
LE-N	441,08	0,00	-13,57	377,57	349,93	-58,00	331,32	-95,91	129,39
LE-S	454,65	13,57	0,00	391,14	363,50	-44,43	344,89	-82,34	142,97
S	63,51	-377,57	-391,14	0,00	-27,65	-435,57	-46,25	-473,48	-248,18
VG	91,16	-349,93	-363,50	27,65	0,00	-407,93	-18,61	-445,84	-220,53
LE-VG	499,08	58,00	44,43	435,57	407,93	0,00	389,32	-37,91	187,40
T	109,76	-331,32	-344,89	46,25	18,61	-389,32	0,00	-427,23	-201,92
LE-T	536,99	95,91	82,34	473,48	445,84	37,91	427,23	0,00	225,31
SkewN	311,69	-129,39	-142,97	248,18	220,53	-187,40	201,92	-225,31	0,00

5.3 Ranking de Modelos

Nós utilizamos o critério de informação do desvio (DIC), e o fator de Bayes (FB) como critérios de comparação entre modelos, na simulação e na aplicação. O melhor modelo será aquele que possuir menor DIC, e maior chance de ocorrência (FB). Na Tabela 5.4, o critério DIC nos indica o modelo Variância Gama e o modelo t-Student, ambos com efeito alavanca, como os mais adequados para a série de dados. Entretanto, na Tabela 5.5, o critério DIC indica o modelo Variância Gama e o modelo Slash, ambos sem efeito alavanca, como os mais adequados para a série de dados.

Com respeito ao Fator de Bayes, como um indicativo da chance de um modelo em respeito do outro, se a verossimilhança do modelo A, for maior ao do modelo B, o fator de Bayes definido acima, terá um valor maior que um, e logaritmo do FB positivo, caso contrário será negativo. Na Tabela 5.7 podemos ver que o modelo com parâmetro de assimetria se adequa melhor aos dados considerados que qualquer um dos modelos pertencentes na família VE-MEN1 (sem efeito alavanca). No entanto o modelo de VE-SN é inferior a qualquer um daqueles de que consideram efeito alavanca. Considerando todos os modelos, o modelo que melhor se adequa aos dados é o modelo com erro t-Student com efeito alavanca, que é o modelo que gerou os dados. Podemos dizer que o Fator de Bayes é um bom critério para comparar modelos de volatilidade estocástica.

Capítulo 6

Conclusões

Na presente tese, desenvolvemos a família de modelos de volatilidade estocástica com misturas na escala de normais, e normais assimétricas. Do ponto de vista empírico, os modelos desenvolvidos nos permitem capturar "outliers" nas séries de retornos, tornando-se então modelos mais robustos do que o modelo normal. Outra característica, não menos importante, é a leve assimetria que possuem as séries financeiras, e assim, desenvolvemos o modelo normal assimétrico com e sem efeito alavanca.

Os principais resultados obtidos foram

- Estendemos os modelos de volatilidade estocástica básico de Taylor, relaxando a suposição do erro ser um processo gaussiano, permitindo ter uma distribuição de mistura na escala de normais, que em geral possuem caudas mais pesadas do que a distribuição normal.
- Aplicações empíricas em dados reais e simulações nos mostram como os modelos de mistura na escala de normais são notoriamente melhores do que os modelos de volatilidade estocástica convencionais, como os de erro normal ou t-Student.
- Considerando a evidência empírica de que retornos reagem de forma assimétrica a choques positivos e negativos, efeito alavanca, consideramos os modelos de volatilidade estocástica com efeito alavanca. Neste sentido, consideramos a família de modelos VE-MEN2, os quais consideram uma mesma variável de mistura para as equações de log-retornos e log-volatilidades, além de considerar uma correlação entre ambas equações. Obtendo como resultados que, na maioria dos casos, estes modelos são mais adequados que os modelos propostos pela literatura (erro gaussiano e t-Student), assim como dos sem efeito alavanca.

- Para a série de retornos *SP500* o modelo que melhor ajusta os dados resultou ser o modelo VELE-VG, e entre os modelos sem considerar o efeito alavanca, resultam os modelos VEMEN1-NC e VEMEN1-Slash.
- O modelo de VE de erro com distribuição normal assimétrica apresenta melhor desempenho que a maioria dos modelos VEMEN1, pois acreditamos que o primeiro modelo considera assimetria observada na distribuição dos retornos. Entretanto, se consideramos os modelos com efeito alavanca, o modelo normal assimétrico resulta ser menos favorecido pelos critérios de seleção.
- Considerando os modelos VE-SkewNormal com efeito alavanca, os critérios de comparação o desfavorecem, concluindo-se que considerar assimetria propriamente dos retornos, distinta aos produzidos pelo efeito alavanca, só é relevante em séries financeiras onde a característica da assimetria na série seja relevante.
- Ao comparar os modelos de volatilidade estocástica de efeito alavanca com os modelos de erro normal assimétrico, os primeiros resultam ser melhores que estes últimos, o que poderia nos indicar que é preferível caracterizar as caudas pesadas do que a leve assimetria.
- Com respeito aos critérios de comparação, o DIC resulta ser não tão robusto ao comparar os diferentes modelos de VEMEN1 e VEMEN2. Já o fator de Bayes consegue identificar o melhor modelo do qual temos certeza que provém os dados (no caso das simulações).

Os resultados expostos acima nos fazem concluir que considerar modelos com efeito alavanca e de erro mistura na escala de normal são mais adequados para as séries de retornos financeiros, e que tentar considerar um parâmetro de assimetria próprio na série poderia dar-se só no caso que a série de log-retornos seja bastante assimétrica.

6.1 Pesquisas Futuras

Fica aberta a possibilidade de fazer algumas extensões aos modelos propostos na presente tese. De forma imediata e até natural resulta a implementação de modelos que considerem a característica de assimetria e de caudas pesadas, estamos nos referindo aos modelos que considerem misturas de escala de normais assimétrica elípticas nos seus erros. Outra extensão seria considerar diferentes variáveis de misturas para os erros da equação de log-retornos

e a equação de log-volatilidades e algumas combinações entre os modelos com erros normais assimétricos. Outra extensão natural, seguindo o desenvolvimento dos modelos de volatilidade em tempo contínuo, é considerar um processo de Lévy na série de log-retornos, o que permitirá incorporar ao modelo um parâmetro que controle os saltos dos retornos. Poder-se-ia considerar, também, os modelos de volatilidade estocástica com mudanças de regime com erros de mistura na escala de normais, assim como as extensões já mencionadas.

Evidentemente, no contexto da econometria, os modelos de volatilidade estocástica são aplicados para o cálculo do valor em risco, VaR. Os diferentes modelos mencionados devem ser avaliados neste contexto. E, do ponto de vista Bayesiano, deve-se realizar análise de influência dos diferentes modelos expostos.

No âmbito da implementação de novas metodologias, podemos desenvolver os modelos propostos nesta tese, com métodos mais eficientes, como o "multi-move" ou Filtro de Partículas, de Shephard e Pitts (1992), ou alternativamente algum método não paramétrico que possa aproximar a verossimilhança no lugar de realizar os métodos baseados em MCMC.

Apêndice A

Distribuição normal assimétrica

A.1 Distribuição normal assimétrica padrão

Lema A.1.1 *Se f_0 é uma função de densidade de probabilidade uni-dimensional ao redor do zero, e G é função de distribuição uni-dimensional onde G' existe e é uma densidade simétrica em torno de zero, então*

$$f_Z(z) = 2f_0(z)G(w(z)), \quad (-\infty < z < +\infty) \quad (\text{A.1})$$

é uma função de densidade para alguma função $w(\cdot)$. (Para a prova veja (Azzalini, 1985a), (Azzalini, 2005)). Uma representação estocástica associada ao lema anterior, é dada a seguir ((Azzalini, 1985b)).

Representação estocástica Se $X \sim G'$ e $Y \sim f_0$ são variáveis independentes, então podemos gerar números aleatórios da distribuição dada por (A.1) fazendo

$$Z = \begin{cases} Y, & \text{se } X < G(w(Y)) \\ -Y, & \text{c.c.} \end{cases} \quad (\text{A.2a})$$

A.1.1 Função de densidade e distribuição

A distribuição normal assimétrica resulta do caso particular de utilizar o Lema A.1 quando $f_0 = \phi$ e $G = \Phi$, a função de densidade e de distribuição da $N(0, 1)$ respectivamente, e $w(x) = \alpha x$ onde, $\alpha \in \Re$ ((Azzalini, 1985a)).

Definição A.1.1 *Uma variável aleatória Z tem distribuição normal assimétrica padrão se sua função de densidade de probabilidade é dada por*

$$f_Z(z) = 2\phi(z)\Phi(\lambda z), \quad -\infty < z < +\infty. \quad (\text{A.3})$$

O parâmetro λ ($\lambda \in \mathfrak{R}$) caracteriza a forma da distribuição e também a forma da assimetria, pois valores negativos de λ indicam assimetria negativa e valores positivos assimetria positiva. Se $\lambda = 0$ a densidade acima coincide com a densidade da distribuição normal padrão e portanto simétrica. Utilizaremos a notação $Z \sim SN(\lambda)$; para uma denotação mais geral da normal assimétrica padrão temos $Z \sim SN(0, 1, \lambda)$.

Proposição A.1.1.1 (*Azzalini, 1985a*). *Se $Z \sim SN(\lambda)$ e $Y = \epsilon + \sigma Z$, onde $\epsilon \in \mathfrak{R}$, $\lambda \in \mathfrak{R}$, $\sigma \in \mathfrak{R}^+$, então $Y \sim SN(\epsilon, \sigma^2, \lambda)$, e sua função de densidade é dada por*

$$f_Y(y) = 2\phi\left(\frac{y - \epsilon}{\sigma}\right)\Phi\left(\lambda\frac{y - \epsilon}{\sigma}\right). \quad (\text{A.4})$$

Propriedades

Denotamos por $HN(0,1)$ a distribuição "half-normal" que tem densidade dada por $2\phi(Z)I_{[0,\infty)}(Z)$.

Para a distribuição normal assimétrica temos:

1. Se $Z \sim SN(\lambda)$, então $|Z| \sim HN(0,1)$.
2. Quando $\lambda \rightarrow \infty$, a densidade (A.4) converge para uma $HN(0,1)$.
3. Se $Z \sim SN(\lambda)$, então $-Z \sim SN(-\lambda)$.
4. A densidade (A.4) é log-côncava.
5. $1 - F_Z(-z; \lambda) = F_Z(z; -\lambda)$.
6. $F_Z(z; 1) = \Phi(Z)^2$.
7. Se $Z \sim SN(\lambda)$, então $Z^2 \sim \chi_1^2$.
8. A função geradora de momentos da normal assimétrica padrão é dada por

$$M_Z(t) = 2 \exp \frac{t^2}{2} \Phi\left(\frac{\lambda}{\sqrt{1 + \lambda^2}} t\right) = 2 \exp \frac{t^2}{2} \Phi(\delta t). \quad (\text{A.5})$$

9. A função geradora de momentos da normal assimétrica $Y \sim SN(\epsilon, \sigma^2, \lambda)$ é dada por

$$M(t) = E \{ \exp tY \} = 2e^{(\epsilon t + \frac{\sigma^2 t^2}{2})} \Phi(\delta \sigma t), \quad (\text{A.6})$$

onde $\delta = \frac{\lambda}{(1+\lambda^2)^{1/2}} \in (-1, 1)$.

Lema A.1.2 Se $U \sim N(0, 1)$ e $a, b \in \Re$, então

$$E[\Phi(a + bU)] = \Phi \left\{ \frac{a}{(1 + b^2)^{1/2}} \right\}. \quad (\text{A.7})$$

Assim, da função geradora de momentos temos,

$$E[Y] = \epsilon + \sigma \mu_z, \quad (\text{A.8})$$

$$Var[Y] = \sigma^2 \left(1 - \frac{2\delta^2}{\pi} \right) = \sigma^2 (1 - \mu_z^2). \quad (\text{A.9})$$

Note que $\mu_z = \delta \sqrt{\frac{2}{\pi}}$.

Usando A.7, podemos demonstrar que,

$$\phi(z) \frac{\Phi(\tau \sqrt{1 + \alpha^2} + \alpha z)}{\Phi(\tau)} \quad (\text{A.10})$$

é uma função de densidade para $\tau \in \mathcal{R}$. A expressão dada em (A.10) é uma extensão da distribuição normal assimétrica, que se reduz a (A.4) quando $\tau = 0$ (Azzalini, 1985a).

Adicionalmente à forma de gerar uma normal assimétrica dado em (A.2a), que permite gerar uma distribuição de tipo (A.4) ou (A.10), existe o método da distribuição condicional (Birnbaum, 1950): dado um vetor (U_0, U_1) , distribuído como uma normal bivariada, com correlação ρ , a distribuição condicional $Z = U_1 | U_0 > \tau^1$ têm distribuição normal assimétrica tipo (A.10), conhecida na literatura como tipo fundamental, com $\tau = -\tau^1$ e $\alpha = \frac{\rho}{\sqrt{1-\rho^2}}$.

Outra forma de gerar é considerando o vetor (U_0, U_1) , e distribuição do $\min(U_0, U_1)$ e $\max(U_0, U_1)$. A distribuição do $\max(U_0, U_1)$ é $SN(\frac{1-\rho}{1+\rho})$, para o mínimo o parâmetro de forma muda somente de sinal.

Weistein(1964) propõe a distribuição de uma variável aleatória gerada pela soma de uma normal e uma normal truncada. Nelson (1964) nos proporciona a expressão computacional do problema tratado por Weistein(1964), que no fundo é a distribuição de (A.10). Já no contexto

bayesiano, O'Hagan e Leonard (1976) gera em forma hierárquica uma normal truncada com parâmetro $\Theta > 0$ da distribuição de $\Theta|\mu \sim N(\mu, \sigma^2)$ onde μ , tem distribuição normal, resultando a expressão

$$Z = \delta X_0 + \sqrt{1 - \delta} X_1 \quad (\text{A.11})$$

onde X_0 e X_1 são normais independentes $N(0, 1)$ e $\delta \in (-1, 1)$. A distribuição de Z é $SN(\frac{\delta}{\sqrt{1-\delta^2}})$ isto é, do tipo (A.4).

A.1.2 Algumas Considerações

Seja $Y \sim SN(\mu, \sigma, \delta)$ com função de densidade,

$$f_Y(y) = 2\phi\left(\frac{y - \mu}{\sigma}\right)\Phi\left(\lambda\frac{y - \mu}{\sigma}\right). \quad (\text{A.12})$$

A reparametrização estocástica do modelo segundo Azzalini (1985a) é,

$$Y = \mu + \sigma \frac{\lambda}{\sqrt{1 + \lambda^2}} |X| + \frac{\sigma}{\sqrt{1 + \lambda^2}} Z, \quad (\text{A.13a})$$

$$X \sim N(0, 1) \perp\!\!\!\perp Z \sim N(0, 1), \quad (\text{A.13b})$$

onde a esperança e variância são

$$E[Y] = \mu + \sigma \frac{\lambda}{\sqrt{1 + \lambda^2}} E[|X|] + \frac{\sigma}{\sqrt{1 + \lambda^2}} E[Z], \quad (\text{A.14a})$$

$$E[Y] = \mu + \eta \sqrt{\frac{2}{\pi}} = \mu + \sigma \delta \sqrt{\frac{2}{\pi}}, \quad (\text{A.14b})$$

$$Var[Y] = \left(\sigma \frac{\lambda}{\sqrt{1 + \lambda^2}}\right)^2 Var[|X|] + \left(\frac{\sigma}{\sqrt{1 + \lambda^2}}\right)^2 Var[Z], \quad (\text{A.14c})$$

$$Var[Y] = \left(\sigma \frac{\lambda}{\sqrt{1 + \lambda^2}}\right)^2 \left(1 - \frac{2}{\pi}\right) + \left(\frac{\sigma}{\sqrt{1 + \lambda^2}}\right)^2 = \sigma^2 \left(1 - \frac{2}{\pi} \delta^2\right). \quad (\text{A.14d})$$

Fazendo a reparametrização,

$$\eta = \sigma \frac{\lambda}{\sqrt{1 + \lambda^2}}, \quad \tau = \frac{\sigma}{\sqrt{1 + \lambda^2}}, \quad (\text{A.15})$$

então $Y \sim SN(\mu, \sigma, \delta)$ pode ser reparametrizada como $Y \sim SN(\mu, \eta, \tau)$ e sua representação estocástica é da forma

$$Y = \mu + \eta |X| + \tau Z, \quad (\text{A.16a})$$

$$X \perp\!\!\!\perp Z \sim N(0, 1), \quad (\text{A.16b})$$

$$E[Y] = \mu + \sqrt{\frac{2}{\pi}}\eta, \quad (\text{A.16c})$$

$$Var[Y] = \eta^2\left(1 - \frac{2}{\pi}\right) + \tau^2. \quad (\text{A.16d})$$

Desta forma a esperança e variância podem ser escritas

$$E[Y] = \mu + \eta\sqrt{\frac{2}{\pi}} = \mu + \sigma\delta\sqrt{\frac{2}{\pi}} = \mu + \sigma\frac{\lambda}{(1 + \lambda^2)^{1/2}}\sqrt{\frac{2}{\pi}}, \quad (\text{A.17})$$

$$Var[Y] = \tau^2 + \eta^2\left(1 - \frac{2}{\pi}\right) = \sigma^2\left(1 - \frac{2\delta^2}{\pi}\right) = \sigma^2\left(1 - \frac{2\lambda^2}{\pi(1 + \lambda^2)}\right), \quad (\text{A.18})$$

obtendo as seguintes relações:

$$\eta = \sigma\frac{\lambda}{\sqrt{1 + \lambda^2}}, \quad \tau = \frac{\sigma}{\sqrt{1 + \lambda^2}}, \quad (\text{A.19a})$$

$$\lambda = \frac{\eta}{\tau}, \quad \sigma^2 = \eta^2 + \tau^2. \quad (\text{A.19b})$$

Apêndice B

Distribuição normal assimétrica multivariada

Para introduzir a versão multivariada da normal assimétrica, é usual considerar a extensão do Lema A.1 para o caso d dimensional.

Lema B.1. *Seja f_0 uma fdp d -dimensional tal que $f_0(x) = f_0(-x)$ para $x \in \mathbf{R}^d$, G é uma função de distribuição unidimensional diferenciável tal que G é simétrica em torno do zero, e w é uma função de valor real tal que $w(-x) = -w(x)$ para todo $x \in \mathcal{R}^d$. Então*

$$f(z) = 2f_0(z)G(w(z)), \quad z \in \mathcal{R}^d, \quad (\text{B.1})$$

é uma função em \mathcal{R}^d .

Arellano e Valle (2002) obtiveram a forma geral de (B.1), expressaram a forma geral da distribuição de $Z = (U_1|U_0 > 0)$, que é dada por

$$f_z(z) = f_{U_1}(z) \frac{P(U_0|U_1 = z)}{P(U_0 > 0)}. \quad (\text{B.2})$$

O problema é a implementação computacional, que envolve integrais de alta dimensão.

Definição B.1 Considere o caso em que em (B.1) f_0 seja $\phi_d(x, \Omega)$, a função de densidade de uma variável normal d -dimensional, onde Ω é uma matriz definida positiva; e o caso $G = \Phi$ e w uma função linear, permitindo a presença de um parâmetro de localização ϵ . A função de densidade fica

$$f(y) = 2\phi(y - \epsilon; \Omega)\Phi(\alpha^T w^{-1}(y - \epsilon)), \quad y \in \mathcal{R}^d, \quad (\text{B.3})$$

onde $\alpha \in \mathcal{R}^d$ é o parâmetro de forma, e w é uma matriz diagonal formada pelos desvios padrões de Ω . Se a v.a. d -dimensional Y tem densidade dada por (B.3), dizemos que Y é distribuída como uma normal assimétrica multivariada, denotada por $Y \sim SN_d(\epsilon, \Omega, \alpha)$.

De forma análoga a (A.7), estendemos para uma SN multivariada, com densidade

$$f(y) = \phi(y - \epsilon; \Omega) \Phi(\alpha_0 + \alpha^T \omega^{-1}(y - \epsilon)) / \Phi(\tau), \quad (\text{B.4})$$

onde $\alpha_0 = \tau(1 + \alpha^T \Omega \alpha)^{1/2}$. Podemos escrever $Y \sim SN_d(\epsilon, \Omega, \alpha, \tau)$. Se $\tau = 0$, temos a definição dada em (B.3).

As integrais envolvidas em (B.2) podem ser resolvidas graças às propriedades da distribuição normal, e a densidade pode-se re-escrever como

$$f(x) = \phi_d(x - \psi; \Omega) \frac{\Phi_m(\tau + \gamma(x - \psi), \Gamma^*)}{\Phi_m(\tau, \Gamma)}. \quad (\text{B.5})$$

De forma similar ao caso univariado, existem métodos para obter amostras com distribuição normal assimétrica, como a distribuição condicional ou a soma de uma normal e uma normal truncada (veja [Azzalini \(2005\)](#)). A distribuição expressa por (B.5), é conhecida na literatura como fundamental, ao qual denotamos $SN_{Fundamental}$. Uma proposta alternativa é dada por [Sahu et al. \(2003\)](#), para o caso em que $m = 1$ na expressão (4.18a), a que denominamos normal assimétrica de Sahu, denotamos por SN_{Sahu} . Esse último caso é mais factível na implementação, num contexto bayesiano.

Consideramos aqui algumas das propriedades da normal multivariada assimétrica.

Proposição B.1 *Se Y_0 e Y_1 são v.a. $N(0, 1)$ independentes e $\delta \in (-1, 1)$, então*

$$\mathbf{Z} = \delta \mathbf{Y}_0 + (1 - \delta^2)^{1/2} \mathbf{Y}_1 \quad (\text{B.6})$$

é $SN(\lambda(\delta))$.

Definição B.2 [Azzalini \(1985a\)](#) *Considere a v.a. normal k -dimensional $Y = (Y_1, \dots, Y_k)^T$ com distribuições marginais padronizadas, independentes de $Y_0 \sim N(0, 1)$; então*

$$\begin{pmatrix} Y_0 \\ Y \end{pmatrix} \sim N_{k+1} \left(0, \begin{pmatrix} 1 & 0 \\ 0 & \Phi \end{pmatrix} \right), \quad (\text{B.7a})$$

onde Φ é $k \times k$ (matriz correlação). Se $(\delta_1, \dots, \delta_k)$ definidos em $(-1,1)$, definimos

$$\mathbf{z}_j = \delta_j \mathbf{Y}_0 + (1 - \delta_j^2)^{1/2} \mathbf{Y}_j, \quad (j = 1, \dots, k), \quad (\text{B.8})$$

isto é $z_j \sim SN(\lambda(\delta_j))$, por (B.6). A densidade da função de z é dada por

$$f_k(z) = 2\Phi_k(z, \mathbf{\Omega})\Phi(\alpha^T z), \quad (z \in \mathfrak{R}^k), \quad (\text{B.9})$$

isto é, a extensão de A.4 para a forma k -dimensional, [Azzalini e Valle \(1996\)](#), onde

$$\alpha^T = \frac{\lambda^T \mathbf{\Psi}^{-1} \mathbf{\Delta}^{-1}}{(1 + \lambda^T \mathbf{\Psi}^{-1} \lambda)^{1/2}}, \quad (\text{B.10})$$

$$\mathbf{\Delta} = \text{diag}((1 - \delta_1^2)^{1/2}, \dots, (1 - \delta_k^2)^{1/2}), \quad (\text{B.11})$$

$$\lambda = (\lambda(\delta_1), \dots, \lambda(\delta_k))^T, \quad (\text{B.12})$$

$$\mathbf{\Omega} = \mathbf{\Delta}(\mathbf{\Psi} + \lambda\lambda^T)\mathbf{\Delta}, \quad (\text{B.13})$$

e $\Phi_k(z, \mathbf{\Omega})$ denota a função de densidade da distribuição normal multivariada k -dimensional com marginais padronizadas e matriz de correlação $\mathbf{\Omega}$. Vamos nos referir a variável aleatória z , com função de densidade (B.9), como uma v.a normal assimétrica k -dimensional, com vetor de parâmetro de forma λ e parâmetro de dependência $\mathbf{\Psi}$. Denotamos como

$$Z \sim SN(\lambda, \mathbf{\Phi}). \quad (\text{B.14})$$

Proposição B.2 *Seja*

$$\mathbf{Y} = \boldsymbol{\mu} + \text{Diag}(\boldsymbol{\eta}) |\mathbf{U}| + \mathbf{Z} \quad (\text{B.15})$$

onde $U \sim N(0,1)$, $Z \sim N(0, \mathbf{\Sigma})$, $\mathbf{Y} \in \mathfrak{R}^n$, $\mathbf{\Sigma}$ matriz simétrica positiva definida. Então $\mathbf{Y} \sim SN(\boldsymbol{\mu}, \mathbf{\Sigma}, \boldsymbol{\eta})$.

Proposição B.3 *Se*

$$\mathbf{Y} \sim SN(\boldsymbol{\mu}, \mathbf{\Sigma}, \boldsymbol{\eta}), \quad (\text{B.16})$$

com $Y \in \mathfrak{R}^n$ e $X = A\mathbf{Y}$, então, $X \sim SN_{Fund}(A\boldsymbol{\mu}, A\mathbf{\Sigma}A^T, A\boldsymbol{\eta})$. Note que se $A\boldsymbol{\eta}$ é uma matriz diagonal $n \times n$ então $X \sim SN_{(Sahu)}$, (Sahu et al. (1993)).

Referências Bibliográficas

- Abanto-Valle, C. A., Bandyopadhyay, D., Lachos, V. H. e Enriquez, I. R. (2010). Robust bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions. *Computational Statistics and Data Analysis*, 54, 2884–2893. 2, 33
- Abramowitz, M. e Stegun, N. (1970). *Handbook of Mathematical Functions*. New York: Dover Publications, Inc. 34
- Ando, T. (2006). Bayesian inference for nonlinear and non-gaussian stochastic volatility model with leverage effect. *Journal of Japan Statistical Society*, 36, 173–197. 77
- Ando, T. (2007). Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models. *Biometrika*, 94, 443–458. 33, 77
- Andrews, D. F., Bickel, P. J., Hampel, F. R., Huber, P. J., Rogers, W. H. e Tukey, J. (1972). *Robust Estimates of Location: Survey and Advances*. Princeton, NJ: Princeton University Press. 28
- Andrews, D. F. e Mallows, S. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society, Series B*, 36, 99–102. 1, 25
- Asai, M. (2005). Comparison of MCMC methods for estimating stochastic volatility models. *Computational Econometrics*, 25, 281–301. 10
- Azzalini, A. (1985a). A class of distribution which includes the normal ones. *Scand. Journal Statist*, 12, 171–178. 59, 60, 93, 94, 96, 100
- Azzalini, A. (1985b). Further results on a class of distribution which includes the normal ones. *Biometrika*, 83, 715–726. 93
- Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *The Scandinavian Journal of Statistical*, 32, 159–188. 64, 93, 100

- Azzalini, A. e Valle, A. D. (1996). The multivariate skew-normal distribution. *Journal of the Royal Statistical Society, Series B*, 46, 253–280. 26, 64, 101
- Berg, A., Meyer, R. e Yu, J. (2004). Deviance Information Criterion for comparing stochastic volatility models. *Journal of Business and Economic Statistics*, 22, 107–120. 7
- Black, F. (1976). Studies of stock market volatility changes. proceeding of the Americans Statistics Association. *J. Business and Economic Statistics*, 3, 177–181. 19
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307–327. 1
- Branco, M. D. e Day, D. K. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, 79, 99–113. 75
- Cathy, W., Chen, S., Liu, F., Mike, K. e So, P. (2008). Heavy-tailed-distributed threshold stochastic volatility models in financial time series. *Australian and New Zealand Journal of Statistics*, 50, 29–51. 25
- Chesney, M., Kohn, R. e Kirby, C. (2004). Pricing European currency options: a comparison of the modified Black-Scholes model and a random variance model. *Journal of Financial and Quantitative Analysis*, 24, 267–284. 19
- Chib, S., Nardari, F. e Shepard, N. (2002). Markov Chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics*, 108, 281–316. 25, 34
- Chow, S. T. B. e Chan, J. S. K. (2008). Scale mixtures distributions in statistical modelling. *Australian and New Zealand Journal of Statistics*, 50, 135–146. 25, 26
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica*, 50, 987–1008. 1
- Fama, E. (1965). Portfolio analysis in a stable Paretian market. *Managament Science*, 11, 404–419. 25
- Fang, K. T., Kotz, S. e Ng, K. W. (1990a). Symmetric multivariate and related distributions. *Chapman and Hall, London/New York*,. 26
- Fang, K. T., Kotz, S. e Zhang, Y. T. (1990b). Generalized Multivariate Analysis. *Springer-Verlag, New York-Berlin*,. 26

- Fernandez, C. e Steel, M. F. J. (2000). Bayesian Regression analysis with scale mixtures of normals. *Econometric Theory*, 16, 80–101. 25
- Genton, M. G., Luca, G. D. e Leoporfido, N. (2004). A multivariate Skew-GARCH model. *A Journey Beyond Normality. Genton, M, Ed., Chapman and Hall Boca Raton,* 205–222. 1, 64
- Geweke, J. (1992). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*. Oxford, U.K.: Oxford University Press. 5, 6, 16, 21, 36, 49
- Ghysels, E., Harvey, A. C. e Renault, E. (1996). Stochastic volatility. *Edo. Rao C. R. and G.S. Manddala, Statistics models in finance, North-Holland, Amsterdam*, 3569–3649. 1, 10
- Gilks, W. R., Richardson, S. e Spiegelhalter, D. J. (1996). Markov Chain Monte Carlo in practice. *Chapman and Hall, London,*. 5, 10
- Gross, A. M. (1973). A Monte Carlo swindle for estimators of location. *Journal of the Royal Statistical Society, Series C. Applied Statistics*, 22, 347–353. 28
- Harvey, A. C., Ruiz, E. e Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, 61, 247–264. 2, 10, 12, 20
- Harvey, A. C. e Shephard, N. (1996). Multivariate stochastic variance models. *Journal of Business and Economic Statistics*, 14, 429–434. 19
- Hastings, W. (1970). Monte-carlo sampling methods using markov chains and their applications. *Biometrika*, 57, 97–109. 6
- Jacquier, E., Polson, N. G. e Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, 12, 371–418. 10, 11
- Jacquier, E., Polson, N. G. e Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122, 185–212. 1, 15, 19, 25
- Jun, Y. (2000). Bugs for bayesian analysis of stochastic volatility models. *Eonometrics Journal*, 3, 198–215. 10, 15, 33
- Jun, Y. (2005). On leverage in stochastic volatility model. *Journal of Econometrics*, 127, 165–178. 1, 19, 20

- Jun, Y. (2006). Multivariate stochastic volatility model. Bayesian estimation and model comparison. *Econometric Reviews*, 25, 361–364. 1, 33
- Kim, S., Shepard, N. e Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies*, 65, 361–393. 1, 10, 12, 13, 14, 15, 20, 33, 77
- Lange, K. L., Little, R. J. A. e Taylor, J. M. G. (1989). Robust statistical modeling using t distribution. *Journal of the American Statistical Association*, 84, 881–896. 27
- Lange, K. L. e Sinsheimer, J. S. (1993). Normal/independent distributions and their applications in robust regression. *J. Comput. Graph. Stat*, 2, 175–198. 25, 26
- Liesenfeld, R. e Jung, R. C. (2000). Stochastic volatility models: Conditional normality versus heavy-tailed distributions. *Journal of Applied Econometrics*, 15, 137–160. 25
- Little, R. J. A. (1988). Robust estimation of the mean and covariance matrix from data with missing values. *Applied Statistics*, 37, 23–38. 27
- Madan, D. B. e Seneta, E. (1990). The variance gamma (v.g) model for share market return. *Journal of Business*, 63, 511–524. 29
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36, 314–419. 25
- Morgenthaler, S. e Tukey, J. (1991). *Configural Polysampling: A Route to Practical Robustness*. New York: Wiley. 28
- Nakajima, J. e Omori, Y. (2009). Leverage, heavy-tails and correlated jumps in stochastic volatility models. *Computational Statistics and Data Analysis*, 53, 2535–2553. 77
- Nelson, D. B. (1991). Conditional heteroscedasticity in asset returns: a new approach. *Econometrica*, 59, 347–370. 19
- Nelson, L. S. (1964). The sum of the values from a normal and a truncated normal distribution. *Technometrics*, 6, 469–471. 95
- O’Hagan, A. e Leonard, T. (1976). Bayesian estimation subject to uncertainty about parameter constraints. *Biometrika*, 63, 201–202. 96

- Omori, Y., Chib, S. e Jouchi, N. (2007). Block sampler and posterior mode estimation for a nonlinear and non gaussian state-space model with correlated errors. *Journal of Economics*,. 12, 14, 15
- 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. 1
- Omori, Y. e Watanabe, T. (2008). Block sampler and posterior mode estimation for asymmetric stochastic volatility models. *Computational Statistics and Data Analysis*, 52, 2892–2910. 19, 33
- Philippe, A. (1997). Simulation of right and left truncated gamma distributions by mixtures. *Statistics and Computing*, 7, 173–181. 35
- Sahu, S., Day, D. e Branco, M. (2003). A new class of multivariate skew distributions with applications to bayesian regression models. *Canadian Journal of Statistics*, 31, 129–150. 64, 65, 100
- Shephard, N. (1993). Fitting non-linear time series models, with applications to stochastic variance model. *Journal of Applied Econometrics*, 8, 135–152. 11
- Shephard, N. (1996). *Statistical aspects of ARCH and stochastic volatility*. London:Chapman & Hall: D.R. Cox, O. E. Barndorff Nielson and D.V. Hinkley. 1, 10, 15, 33
- Shephard, N. (2005). *Stochastic Volatility*. Selected Readings: Oxford Press. 9
- Shephard, N. e Pitt, M. (1997). Likelihood analysis of non-Gaussian measurements time series. *Biometrika*, 84, 653–667. 1, 10, 13, 33
- Siddhartha, C. e Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49, 327–335. 6
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. e Linde, A. V. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, 64, 621–622. 7
- Steele, J. (2001). *Stochastic Calculus and Financial Applications*. New York: Springer-Verlag. 9, 10

- Taylor, S. (1982). *Financial returns modelled by the product of two stochastic processes—a study of the daily sugar prices 1961-75*. 1, 9
- Taylor, S. (1986). *Modeling Financial Time Series*. Chichester: Wiley. 9
- Taylor, S. (1994). Modeling stochastic volatility. *Mathematical Finance*, 4, 183–204. 10
- Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics*, 21, 1701–1762. 5, 13, 34
- Watanabe, T. e Omori, Y. (2004). A multi-move sampler for estimate non-Gaussian time series model: Comments on Shepard and Pitt (1997). *Bimetrika*, 91, 246–248. 14, 33
- Wiggins, J. B. (1987). Options values under stochastic volatility: theory and empirical estimate. *Journal of Financial Economics*, 19, 351–372. 19