

UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE CIÊNCIAS EXATAS
POS-GRADUAÇÃO EM MATEMÁTICA

**Aplicação em modelos de variação autorregressiva condicional baseada na distribuição
Birnbaum-Saunders**

ERICO JANDER DA SILVA LOPES

MANAUS-AM

2019

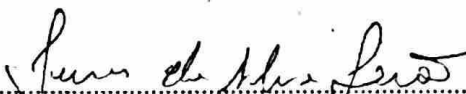
Erico Jander da Silva Lopes

APLICAÇÃO EM MODELOS DE VARIAÇÃO AUTORREGRESSIVA CONDICIONAL
BASEADA NA DISTRIBUIÇÃO BIRNBAUM-SAUNDERS

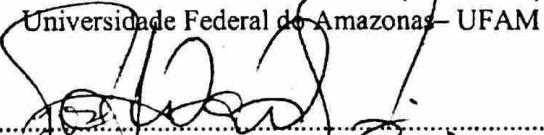
Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Amazonas, como requisito final para obtenção do título de Mestre em Matemática, área de concentração em Estatística.

Aprovado por:

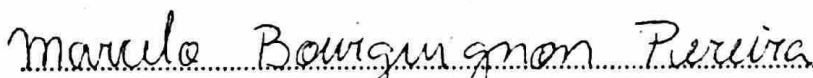
BANCA EXAMINADORA



Prof. Dr. Jeremias da Silva Leão (Orientador)
Universidade Federal do Amazonas - UFAM



Prof. Dr. José Mir Justino da Costa (Membro)
Universidade Federal do Amazonas - UFAM



Prof. Dr. Marcelo Bourguignon (Membro Externo)
Universidade Federal do Rio Grande do Norte - UFRN

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

L864a	<p>Lopes, Erico Jander da</p> <p>Aplicação em modelos de variação autorregressiva condicional baseada na distribuição Birnbaum-Saunders / Erico Jander da Lopes. 2017 36 f.: il.; 31 cm.</p> <p>Orientador: Jeremias da Silva leão Dissertação (Mestrado em Matemática - Estatística) - Universidade Federal do Amazonas.</p> <p>1. Distribuição Birnbaum-Saunders,. 2. Modelos CARR,. 3. Análise de resíduos.. 4. Análise de diagnóstico.. I. leão, Jeremias da Silva II. Universidade Federal do Amazonas III. Título</p>
-------	---



UFAM

APLICAÇÃO EM MODELOS DE VARIAÇÃO AUTORREGRESSIVA CONDICIONAL
BASEADA NA DISTRIBUIÇÃO BIRNBAUM-SAUNDERS

Erico Jander da Silva Lopes

Dissertação de Mestrado apresentada ao
Programa de Pós-graduação em Matemática,
da Universidade Federal do Amazonas, como
parte dos requisitos necessários à obtenção do
título de Mestre em Matemática

Orientador: Jeremias da Silva Leão

Manaus

Abril de 2019

*Esta dissertação é dedicada aos
meus pais e as minhas mães.*

Agradecimentos

Primeiramente a Deus, por iluminar os caminhos nos momentos difíceis.

A minha companheira que está comigo em todos os momentos de alegrias e dificuldades.

Ao professor Jeremias da Silva Leão (o cara) por sua paciência em compartilhar seus conhecimentos.

A minha mãe, Maria de Jesus e a minha irmã pelo incentivo.

Ao meu "irmão" de mestrado Tiago de Melo.

Aos amigos de caminhada, Guilherme, Antônio, Alice, Renata, Milena, Alex, Jhonata, Ciço, Vinícius, Natan e o professor Nelson.

Aos professores e colegas do Departamento de Estatística pelos ensinamentos.

Ao Departamento de Estatística.

Ao Instituto de Ciências Exatas.

A Universidade federal do Amazonas.

A todos que de uma forma ou de outra contribuíram para o desenvolvimento deste trabalho.

E principalmente aos meus acolhedores, São Jorge, Nossa Senhora dos Navegantes e São Sebastião.

Resumo da Dissertação apresentada ao Programa de Pós-Graduação em Matemática, da Universidade Federal do Amazonas, como parte dos requisitos necessários para a obtenção do grau de Mestre em Matemática. (M.Sc.)

APLICAÇÃO EM MODELOS DE VARIAÇÃO AUTORREGRESSIVA CONDICIONAL
BASEADA NA DISTRIBUIÇÃO BIRNBAUM-SAUNDERS

Erico Jander da Silva Lopes

Orientador: Jeremias da Silva Leão

Área de Concentração: Estatística

O modelo de variação autorregressivo condicional (CARR), proposto por Chou (2005) se mostrou eficiente para estimar a volatilidade do preço de ativos. Entretanto, a estimativa requer uma densidade do erro adequada, onde se usa comumente a distribuição de Weibull. Xie & Wu (2017) propôs um modelo baseado na distribuição gamma (GCARR), com resultados satisfatórios em relação a redução de problema de inlier e outlier. Neste trabalho, propomos o modelo de variação autorregressiva condicional baseado na distribuição Birnbaum-Saunders (BSCARR). Implementamos uma abordagem baseada no método da máxima verossimilhança para obter as estimativas dos parâmetros e derivamos medidas para análise de resíduos e diagnóstico. Em seguida fizemos um estudo via simulações e Monte Carlo com o objetivo de avaliar o desempenho dos estimadores de máxima verossimilhança do modelo proposto. Por fim, ilustramos a metodologia proposta usando um conjunto de dados reais.

Palavras-chave: Distribuição Birnbaum-Saunders, Modelos CARR, Análise de resíduos e diagnóstico.

Abstract of Dissertation presented to Postgraduate in Mathematics, of the Federal University of Amazonas, as a partial fulfillment of the requirements for the degree of Master of Mathematics. (M.Sc.)

SUBTITULO

Erico Jander da Silva Lopes

Advisor: Jeremias da Silva Leão

Research lines: Statistics

The conditional autoregressive variation (CARR) model proposed by Chou (2005) proved to be efficient in estimating asset price volatility. However, the estimation requires an adequate error density, where the Weibull distribution is commonly used. Xie & Wu (2017) proposed a model based on gamma distribution (GCARR), with satisfactory results in inlier and outlier problem reduction. In this work, we propose the conditional autoregressive variation model based on the Birnbaum-Saunders distribution (BSCARR). We implemented an approach based on the maximum likelihood method to obtain the parameter estimates and derive measurements for residue analysis and diagnosis. We then carried out a simulation and Monte Carlo study with the objective of evaluating the performance of the maximum likelihood estimators of the proposed model. Finally, we illustrate the proposed methodology using a set of real data.

Keyword: Birnbaum-Saunders Distribution, CARR Models, Waste Analysis and Diagnosis.

Sumário

Lista de Figuras	vii
Lista de Tabelas	viii
1 Introdução	1
1.1 Contextualização	1
1.2 Distribuição Birnbaum-Saunders	3
1.3 Estrutura da dissertação	6
2 Modelo autorregressivo de variação condicional (CARR)	7
2.1 Introdução	7
2.2 Modelo de variação autorregressiva condicional BS	9
2.3 Diagnóstico de influência e análise de resíduos	10
2.4 Influência global	11
2.5 Influência local	11
2.5.1 Perturbação de casos	13
2.5.2 Perturbação na resposta	13
2.6 Perturbação renovada	14
3 Avaliação numérica de dados	15
3.1 Estudo de Simulação	15
3.2 Análise de dados financeiros	17
3.2.1 Estimção e validação do modelo	20
3.2.2 Análise de diagnóstico	20

4	Considerações Finais	33
	Referências Bibliográficas	34

Lista de Figuras

1	2
2	3
1.1	CDF e PDF de $T \sim BS(\kappa, \sigma)$ para diferentes valores de κ com $\sigma=1.0$	6
3.1	Gráficos dos retornos e amplitudes diárias dos dados da BVSP, HSI e TSEC (linha), no período de 01/07/2007 a 31/07/2017.	17
3.2	Histograma, gráfico TTT e os boxplots dos dados da BVSP do Brasil(BVSP), HSI de HongKong(HSI) e TSEC Taiwan (TSEC)	19
3.3	QQplot do resíduo RQ para os dados da BVSP, HSI e TSEC, respectivamente, segundo o ajuste dos modelos WCARR, GCARR e BSCARR.	22
3.4	Distância de Cook Generalizada para o modelo BSCARR para os dados da BVSP do Brasil(BVSP), HSI de HongKong(HSI) e TSEC Taiwan (TSEC) respectivamente.	23
3.5	Índice C_i para o vetor de parâmetros sob perturbação de casos no modelo BSCARR com os dados BVSP, HSI e TSEC	24
3.6	Índice C_i para o parâmetro sob perturbação na Variável Resposta no modelo BSCARR com os dados BVSP, HSI e TSEC.	25
3.7	Índice C_i para o parâmetro sob perturbação na resposta dos dados no modelo BSCARR com os dados BVSP, HSI e TSEC	26
3.8	Gráficos ACF e PACF da transformada integral de probabilidade z_i para os modelos BSCARR para os dados da BVSP, HSI e TSEC	32

Lista de Tabelas

3.1	Estatísticas da análise de simulação para a estimativa do modelo BSCARR (1, 1).	16
3.2	Estatísticas descritivas da amplitude dos índices.	18
3.3	Estimativa do modelo CARR (1,1) com termo de perturbação com distribuição Birnbaum-Saunders, Weibull e Gama.	21
3.4	RCs (%) nas estimativas de ML e SEs correspondentes para o parâmetro indicado e os casos retirados, e respectivos valores de p	28
3.5	Previsão j passos a frente e erro quadrático médio (MSE) dos modelos WCARR, GCARR e BSCARR para os índices.	30
3.6	Resultados do teste LB obtidas para os modelos WCARR, GCARR e BSCARR para os índices BVSP, HSI e TSEC.	31

Capítulo 1

Introdução

1.1 Contextualização

Entender a natureza da dinâmica da volatilidade dos preços de ativos especulativos vem sendo bastante estudada na literatura, uma vez que está intimamente relacionado com o funcionamento e a instabilidade dos mercados financeiros, e vinculados diretamente com o funcionamento e as flutuações da economia real. Nas últimas três décadas foram propostos uma variedade de modelos para descrever a dinâmica de volatilidade, em que tradicionalmente esses modelos são estimados a partir do preço de fechamento.

No que se refere à modelagem da volatilidade podemos destacar a família de modelos ARCH/GARCH, Engle (1982), Bollerslev (1986) e Nelson (1991). Um levantamento completo, com uma revisão crítica sobre a metodologia ARCH pode ser vista em Bollerslev *et al.* (1992). Um modelo competitivo de volatilidade para o ARCH é o modelo de volatilidade estocástica (doravante SV) de Taylor (2008) e Heston (1993). Engle & Russell (1998) propuseram o modelo de duração autorregressiva condicional (ACD) para modelar dados de transações financeiras, em que a duração é, geralmente, definida como o intervalo de tempo entre eventos consecutivos, por exemplo, o intervalo de tempo entre duas transações de uma ação na bolsa de valores ou a diferença entre os tempos de chegada de dois clientes em uma estação de serviço. Como esta duração é necessariamente não negativa, o modelo ACD pode, também, ser usado para modelar séries temporais de observações positivas, como por exemplo o logarítimo do preço diário de um ativo.

Como vimos, anteriormente, a modelagem da volatilidade pode ser feita, não somente através do preço de fechamento dos mercados. Dessa forma, Chou (2005) propôs o modelo autorregressivo condicional baseado na amplitude (CARR), podemos usar, também, a terminologia de modelo autorregressivo de variação condicional. Vale salientar que apesar desta “classe” de modelos parecerem recentes, o uso da variação de preços em finanças não é um novo conceito, Mandelbrot (1971) e outros o empregam para testar a existência de dependência nos preços dos ativos. Chou (2005) mostra que estimativas de volatilidade para o modelo CARR são mais eficientes quando comparadas às obtidas por um modelo GARCH padrão na previsão de volatilidade para os logaritmos dos retornos semanais de uma amostra do índice S&P500. Entretanto, a eficiência na estimativa requer uma densidade do erro adequada para a metodologia CARR. É comumente utilizado no modelo CARR a distribuição Weibull, sendo reconhecida como WCARR. Parkinson (1980) argumentou que a amplitude de um preço de ativo durante um dia de negociação pode ser usado para medir a sua volatilidade e também demonstrou existir uma “superioridade” no uso do intervalo como um estimador de volatilidade em comparação com os métodos padrões. Dentre os trabalhos recentes com a abordagem CARR podemos destacar:

Chan *et al.* (2012) propuseram o modelo de variação autorregressiva condicional baseado no processo geométrico. Segundo os autores as teorias de valores extremos indicam que o intervalo é um estimador eficiente da volatilidade local em séries temporais financeiras, dessa forma uma estrutura de processo geométrico (GP) que incorpora a condição. Chiang *et al.* (2016) desenvolveram procedimentos para detecção de “outliers” no modelo logarítmico da variação autorregressiva condicional baseado na distribuição lognormal (Log-CARR). Xie & Wu (2017) introduziu o modelo CARR baseado na distribuição gamma, que denotaremos por GCARR. Os autores mostram que o GCARR têm resultados satisfatórios em relação a redução do problema de inlier, como também diminuiu os problema com outlier em comparação com o modelo WCARR. Entretanto o modelo GCARR apresenta evidências empíricas que ainda existe algum desvio na distribuição da gama e que o modelo reduz apenas parcialmente os problemas de inlier e outlier.

Diante disso, uma boa alternativa à distribuição gama é a distribuição Birnbaum-Saunders (BS), proposta por Birnbaum & Saunders (1969). Este modelo tem sido amplamente abordado na literatura devido às suas propriedades matemáticas, bem como sua estreita relação com a

distribuição normal. A BS é uma distribuição assimétrica, contínua e unimodal. É, também, conhecida como modelo de tempo de fadiga. A BS tem sido extensivamente aplicada para modelar tempos de falha em engenharia, diversas aplicações em ciências ambientais, biológicas e financeiras também foram consideradas, ver por exemplo Desmond (1985), Kotz *et al.* (2010), Saulo *et al.* (2013), e Leiva *et al.* (2014a,b, 2015, 2017).

Em face disso, podemos destacar alguns dos objetivos deste estudo, que são: (i) apresentar o desenvolvimento dos processos de estimação baseado no princípio da máxima verossimilhança, bem como um estudo de simulação para avaliar propriedades dos estimadores; (ii) introduzir métodos de análise de resíduos e desenvolver medidas de influência e diagnóstico para o modelo proposto; e (iii) aplicar a metodologia proposta em aplicações à dados reais para ilustrar as potencialidades do modelo.

1.2 Distribuição Birnbaum-Saunders

Motivados pela falha de materiais e em problemas nos novos aviões comerciais, Birnbaum & Saunders (1969) derivaram um nova família de distribuições de vida que modela o tempo de vida de materiais e equipamentos sujeitos a cargas dinâmicas. Considere que um material é submetido a um padrão cíclico de tensão e força, onde a “carga” é uma função que representa o resultado da tensão lançada no material no tempo t . Um ciclo contém m oscilações, onde cada aplicação da i -ésima oscilação em um ciclo resulta em uma extensão aleatória da fissura X_i . O processo de fadiga baseou-se nas seguintes suposições:

1. Uma amostra de material é submetida a cargas cíclicas de tensão e força, que produzem uma fissura ou desgaste;
2. A falha ocorre quando o tamanho da fissura na amostra de material excede certo nível de resistência, denotado por w ;
3. A sequência de cargas impostas no material é a mesma de um ciclo para outro;
4. A extensão incremental da fissura X_i resultante da aplicação da i -ésima oscilação de carga é uma variável aleatória com uma distribuição que só depende da fissura atual causada pela tensão neste ciclo;

5. A extensão total da fissura durante o $(j + 1)$ -ésimo ciclo é

$$Y_{j+1} = Y_{jm+1} + \dots + Y_{jm+m}, \quad j = 0, 1, \dots$$

A extensão da fissura após o j -ésimo ciclo é dada por

$$Y_j = \sum_{i=1}^m X_i,$$

em que Y_j é uma variável aleatória com distribuição de média μ e variância ζ^2 , $\forall j = 1, 2, \dots$

A extensão total da rachadura após z ciclos é dada por

$$W_z = \sum_{j=1}^z Y_j,$$

com função de distribuição

$$H_z(w) = P(W_z \leq w)$$

para $z = 1, 2, 3, \dots$

Seja N o numero de ciclos até a falha, onde a falha ocorre quando o comprimento da fissura excede um dado comprimento crítico w . A função de distribuição de N é dada por

$$P(N \leq z) = P\left(\sum_{j=1}^z Y_j > w\right) = 1 - H_z(w).$$

Supondo que os Y_j 's são variáveis aleatórias independentes e identicamente distribuídas, podemos aproximar N pelo Teorema central do Limite, isto é,

$$\begin{aligned} P(N \leq z) &= 1 - P\left(\sum_{j=1}^z \frac{Y_j - \mu}{\zeta\sqrt{z}} \leq \frac{w - Z\mu}{\zeta\sqrt{z}}\right). \\ &= 1 - P\left(\sum_{j=1}^z \frac{Y_j - \mu}{\zeta\sqrt{z}} \leq \frac{w}{\zeta\sqrt{z}} - \frac{\mu\sqrt{z}}{\zeta}\right). \\ &\cong 1 - \Phi\left(\frac{w}{\zeta\sqrt{z}} - \frac{\mu\sqrt{z}}{\zeta}\right). \end{aligned}$$

em que $\Phi(\cdot)$ representa a função distribuição acumulada (CDF) da normal padrão. Birnbaum & Saunders (1969) definiram que se substituirmos z por uma variável real não-negativa t , a

variável aleatória T é a extensão da variável aleatória N . Dessa forma T é considera o tempo até a falha. A CDF de T pode ser escrita como

$$F(t; \kappa, \varsigma) = P(T \leq t) = \Phi \left[\frac{1}{\kappa} \left(\sqrt{\frac{t}{\varsigma}} - \sqrt{\frac{\varsigma}{t}} \right) \right], t > 0, \kappa > 0, \varsigma > 0, \quad (1.1)$$

em que κ é o parâmetro de forma e σ o parâmetro de escala, que é dado por

$$\kappa = \frac{\varsigma}{\sqrt{\mu w}} \quad e \quad \sigma = \frac{w}{\mu}$$

A função densidade de probabilidade (PDF) da BS obtida a partir de (1.1) é dada da por

$$f(t; \kappa, \sigma) = \frac{1}{2\kappa\sigma\sqrt{2\pi}} \left[\left(\frac{\kappa}{t} \right)^{\frac{1}{2}} + \left(\frac{\kappa}{t} \right)^{\frac{3}{2}} \right] \exp \left\{ -\frac{1}{2\kappa^2} \left(\frac{t}{\kappa} + \frac{\kappa}{t} - 2 \right) \right\},$$

$t > 0$, e $\kappa, \sigma > 0$. Considerando a seguinte transformação

$$T = \sigma \left\{ \frac{\kappa Z}{2} \left[\left(\frac{\kappa Z}{2} \right)^2 + 1 \right]^{\frac{1}{2}} \right\}^2,$$

em que Z segue uma distribuição normal padrão, temos que o valor esperado e a variância da BS são dados por

$$\mathbb{E}[T] = \sigma \left(1 + \frac{\kappa^2}{2} \right) \quad e \quad \text{Var}[T] = (\kappa\sigma)^2 \left(1 + \frac{5\kappa^2}{4} \right)$$

A BS satisfaz as propriedades recíproca, ou seja, se $T \sim BS(\kappa, \sigma)$ então $1/T \sim BS(\kappa, \sigma^{-1})$; e a propriedade de multiplicação por escalar, onde $\forall b > 0$, $bT \sim BS(\kappa, b\sigma)$. Consequentemente, temos que

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

Como podemos observar na Figura 1.1, as formas de PDF da BS varia de uma alta distorção com cauda longa (valores alpha grandes) a quase simétricas e de cauda curta à medida que o alpha diminui.

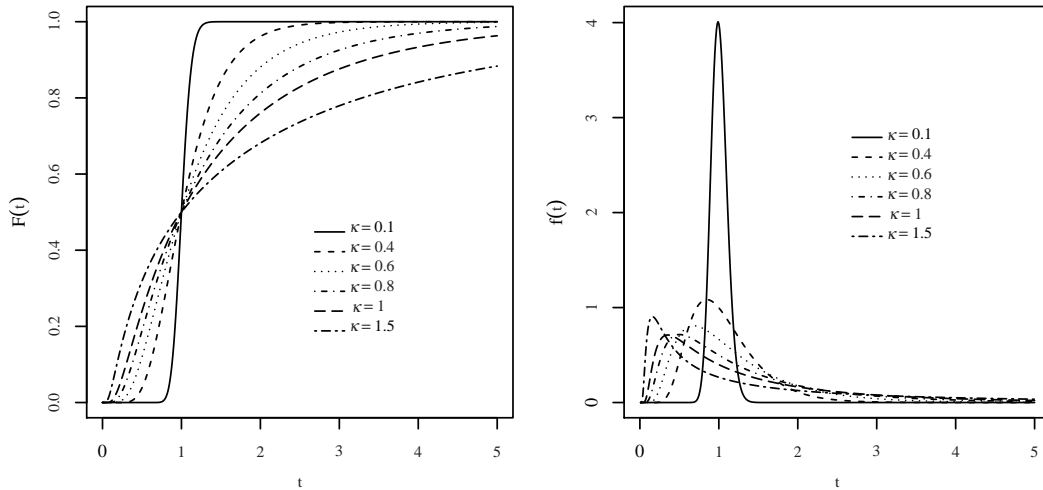


Figura 1.1: CDF e PDF de $T \sim BS(\kappa, \sigma)$ para diferentes valores de κ com $\sigma=1.0$.

1.3 Estrutura da dissertação

A presente dissertação é constituída por quatro capítulos. No primeiro capítulo, definimos o objetivo do trabalho e descrevemos algumas propriedades sobre a distribuição BS. Em seguida, no segundo capítulo, apresentamos o modelo proposto bem como o processo de estimação por máxima verossimilhança dos parâmetros, introduzimos as medidas de análise de resíduos e influência e diagnóstico para a metodologia apresentada. No terceiro capítulo, apresentamos os resultados e análises do estudo de simulações para modelo BSCARR, como também a aplicação a dados reais. Por fim, no quarto capítulo, discutimos as principais características observadas no estudo do modelo BSCARR e as perspectivas de trabalhos futuros.

Capítulo 2

Modelo autorregressivo de variação condicional (CARR)

2.1 Introdução

Como apresentado em Chou (2005), temos que P_t é logaritmo do preço de um ativo, em que o objetivo é analisar a amplitude (variação) medida em intervalos discretos (por exemplo, diariamente, semanalmente). Dessa forma, a variação observada (amplitude) é dada por

$$R_t = \text{Max}\{P_\tau\} - \text{Min}\{P_\tau\}, \quad (2.1)$$

em que $\tau = t - 1, t - 1 + \frac{1}{n}, t - 1 + \frac{2}{n}, \dots, t, n$ é o número de intervalos utilizados na medição do preço dentro de cada intervalo medido pelo intervalo. Considerando que os preços estejam na escala logarítmica, podemos definir $r_t = P_t - P_{t-1}$ como o retorno obtido no período $(t - 1, t)$. De acordo com Parkinson (1980) e Lo (1991) a amplitude (variação) pode ser usada como uma medida de variabilidade. Com isso, Chou (2005) definiu o modelo CARR como

$$\begin{aligned} R_t &= \lambda_t \varepsilon_t, \\ \lambda_t &= \omega + \sum_{j=1}^p \alpha_j \lambda_{t-j} + \sum_{j=1}^q \beta_j R_{t-j}, \end{aligned} \quad (2.2)$$

em que λ_t é a média condicional do intervalo com base em todas as informações até o tempo t . Supõe-se que o termo de aleatório (erro) ε_t assume uma distribuição com função de densi-

dade $f(\cdot)$ com média unitária. Os coeficientes ω , α_i e β_j na equação média condicional, são todos positivos, para garantir o intervalo seja positivo. Os parâmetros ω , α_i e β_j , caracterizam a incerteza inerente ao intervalo, o efeito de impacto a curto prazo e o efeito de impacto no intervalo a longo prazo (ou a volatilidade do retorno), respectivamente. Para que o processo seja estacionário as raízes do polinômio característico devem estar fora do círculo unitário $\sum_{j=1}^q \alpha_j + \sum_{j=1}^p \beta_j < 1$. Para motivar a especificação do modelo que será proposto, vamos revisar a especificação do modelo de variação autorregressiva condicional baseado na distribuição Weibull (WCARR). Assim sendo, temos que a PDF de uma variável aleatória (RV) Weibull é dada por

$$f(r) = \frac{\theta}{\sigma} \left(\frac{r}{\sigma}\right)^{\theta-1} \exp\left[-\left(\frac{r}{\sigma}\right)^\theta\right], \quad (2.3)$$

em que $r > 0$, $\theta > 0$ e $\sigma > 0$, os parâmetros θ e σ representam a forma e a escala da distribuição, respectivamente. Com isso temos que $\mathbb{E}(R) = \sigma\Gamma(1 + \theta^{-1})$. Considerando a transformação de variável dada por $\varepsilon_t = R_t/\sigma\Gamma(1 + \theta^{-1})$ em (2.3), obtemos a distribuição Weibull com média unitária

$$f_\varepsilon(\varepsilon_t) = \frac{\theta}{\Gamma(1 + \theta^{-1})^{-1}} \left(\frac{\varepsilon_t}{\Gamma(1 + \theta^{-1})^{-1}}\right)^{\theta-1} \exp\left[-\left(\frac{\varepsilon_t}{\Gamma(1 + \theta^{-1})^{-1}}\right)^\theta\right], \quad (2.4)$$

em que $\varepsilon_t > 0$ e $\theta > 0$. Devemos, ainda, aplicar uma segunda transformação de variável a fim de obter a distribuição de R_t parametrizada em termos da média condicional λ_t . Dessa forma, considerando que $\varepsilon_t = R_t/\lambda_t$, temos a PDF condicional será dada por

$$f_{R_t|\lambda_t}(R_t) = \frac{\theta}{\lambda_t/\Gamma(1 + \theta^{-1})} \left(\frac{R_t}{\lambda_t/\Gamma(1 + \theta^{-1})}\right)^{\theta-1} \exp\left[-\left(\frac{R_t}{\lambda_t/\Gamma(1 + \theta^{-1})}\right)^\theta\right]. \quad (2.5)$$

A estruturação do modelo CARR é concluída especificando a forma dinâmica para a média condicional λ_t . Xie & Wu (2017) propôs o modelo GCARR, o qual os passos usados para obtenção da dinâmica para a média condicional é semelhante ao caso do modelo WCARR. A função de verossimilhança do modelo CARR é derivada da suposição $R_t = \lambda_t \varepsilon_t$. Devemos notar que a função de verossimilhança do modelo (2.5) é obtida diretamente sem essa suposição. Assim, reparametrizando (2.3) e usando $\lambda_t = \sigma\Gamma(1 + \theta^{-1})$ obtemos $f_{R_t|\lambda_t}(R_t)$. Dessa forma, o modelo CARR formulado diretamente da distribuição de R_t parametrizada em termos da média condicional λ_t ou o modelo CARR formulado a partir da suposição $R_t = \lambda_t \varepsilon_t$ são equivalentes

pelo princípio de verossimilhança, com isso ambos produzem as mesmas estimativas para os parâmetros.

2.2 Modelo de variação autorregressiva condicional BS

O modelo de variação autorregressiva condicional BS, que denotaremos por BSCARR, é motivado devido ao fato de se trabalhar diretamente com a distribuição de R_t , condicionado a um parâmetro natural que varia no tempo. Com isso, os modelos ECARR, WCARR e GCARR podem ser obtidos considerando um parâmetro de escala variando no tempo. No caso do modelo ECARR o parâmetro de escala é a média. Considerando o WCARR e o GCARR o parâmetro de escala e a média são equivalentes por uma constante e, no caso do modelo BSCARR, o parâmetro da escala é a mediana.

Com isso teremos que o modelo BSCARR é um modelo de processo de ponto dinâmico, contudo não é um modelo verdadeiro de CARR no sentido tradicional. A diferença entre o modelo BSCARR e os modelos CARR usuais é o fato que o modelo BSCARR é especificado em termos da amplitude (variação) mediana condicional, o parâmetro de escala natural da distribuição BS(κ, σ). A PDF condicional de R_t dado σ_t é dada por

$$f_{R|\sigma}(R_t|\sigma_t) = \frac{1}{2\sqrt{2\pi\kappa\sigma_t}} \left[\left(\frac{\sigma_t}{R_t} \right)^{1/2} + \left(\frac{\sigma_t}{R_t} \right)^{3/2} \right] \exp \left(-\frac{1}{2\kappa^2} \left[\frac{R_t}{\sigma_t} + \frac{\sigma_t}{R_t} - 2 \right] \right). \quad (2.6)$$

Com isso a função de log verossimilhança do modelo é dada por

$$\begin{aligned} \log f_{R|\sigma}(R_t|\sigma_t) = & -\log(2\sqrt{2\pi}) - \log(\kappa) - \log(\sigma_t) - \frac{1}{2\kappa^2} \left[\frac{R_t}{\sigma_t} + \frac{\sigma_t}{R_t} - 2 \right] \\ & + \log \left[\left(\frac{\sigma_t}{R_t} \right)^{1/2} + \left(\frac{\sigma_t}{R_t} \right)^{3/2} \right], \end{aligned} \quad (2.7)$$

e a dinâmica na mediana condicional variando no tempo é definida por

$$\log(\sigma_t) = \omega + \sum_{j=1}^q \alpha_j \log(\sigma_{t-j}) + \sum_{j=1}^p \beta_j \left(\frac{R_{t-j}}{\sigma_{t-j}} \right).$$

A dinâmica GARCH para $\ln(\sigma_t)$ assegura que $\sigma_t > 0$ sem exigir restrições no proce-

dimento de estimativa, como também facilita a inclusão de covariáveis de microestrutura de mercado, ver Bhatti (2009). As estimativas do vetor parâmetros do modelo BSCARR(p, q), $\xi = [\omega, \alpha_1, \dots, \alpha_{p1}, \beta_1, \dots, \beta_{q1}, \kappa]^\top$ foram obtidos através do método da máxima verossimilhança (ML). Tendo em vista que os estimadores dos parâmetros não são obtidos de forma analítica, devemos utilizar procedimento de otimização não linear, como por exemplo o método de quasi-Newton de Broyden-Fletcher-Goldfarb-Shanno (BFGS) para obtenção dos estimadores. Esse método é implementado no software R disponível em www.r-project.org, através do uso das funções `optim` e `maxBFGS`, detalhes sobre esses métodos podem ser encontrados em Mittelhammer *et al.* (2000).

A inferência para ξ no modelo BSCARR, foi encontrada baseando-se na distribuição assintótica de ML do estimador $\hat{\xi}$, em que o estimador é consistente e possui distribuição assintótica multivariada normal com média ξ e matriz de covariância $\Sigma_{\hat{\xi}}$, que pode ser obtida através da esperança da matrix de Informação de Fisher $\mathcal{J}(\xi)$. Assim, temos que

$$\sqrt{n}[\hat{\xi} - \xi] \xrightarrow{D} N_{p1+q1+2}(0, \Sigma_{\hat{\xi}} = \mathcal{J}(\xi)^{-1}), \quad (2.8)$$

como $n \rightarrow \infty$, em que \xrightarrow{D} indica convergência em distribuição e $\mathcal{J}(\xi) = \lim_{n \rightarrow \infty} [1/n] \mathcal{J}(\xi)$, observando que $\mathcal{J}(\xi)^{-1}$ é um estimador consistente da matriz de covariância assintótica de $\hat{\xi}$. Podemos assim aproximar a esperança da matriz de informação Fisher por sua versão obtida através da matriz de Hessian $\mathcal{J}''(\xi)$, que contém as segundas derivadas de $\ell(\xi)$ (função de verossimilhança do modelo). Além disso, os elementos diagonais de sua matriz inversa podem ser usados para aproximar os erros padrão correspondentes (SEs), ver Cox & Hinkley (1978).

2.3 Diagnóstico de influência e análise de resíduos

A Estatística dispõe de diversas metodologias que auxiliam na análise dos dados. Dentre estas, temos ferramentas que ajudam a verificar presença de observações discrepantes e quais os efeitos que são causados nas estimativas dos parâmetros dos modelos. Nesse sentido, podemos destacar algumas propostas de análise de observações atípicas (*outliers*) sob o contexto de séries temporais: (a) *outliers* aditivos (OA) ; (b) *outliers* inovadores (OI); (c) *outliers* de mudança de nível; e (d) *outliers* de alteração transitória. Os tipos (a) e (b) são os mais

comuns em dados de séries temporais. Os *outliers* aditivo afetam um único ponto, enquanto que o outlier inovador afeta não somente uma observação em particular, mas também os casos subsequentes. De acordo com Zevallos *et al.* (2012) há uma relação de similaridade entre as estatísticas para determinar pontos influentes e testes para detecção de outlier. Dessa forma os autores propuseram procedimentos de detecção de outliers aditivos em modelos Log-ACD discutidos por Bauwens & Giot (2000).

Neste trabalho, além de considerarmos procedimentos para detecção de OA e OI, usamos como forma de analisar a influência global de casos individuais a distância de Cook generalizada, proposta por Cook (1987).

2.4 Influência global

A influência global é uma abordagem utilizada para avaliar o efeito da exclusão (omissão) de observações no conjunto de dados. Com isso, uma quantidade com subscrito “(i)” é calculada com a exclusão da observação i , em que $\ell_{(i)}$ é a função de log-verossimilhança definida (2.7), mas avaliada em $\hat{\xi}_{(i)} = [\hat{\omega}_{(i)}, \hat{\alpha}_{1(i)}, \dots, \hat{\alpha}_{p_1(i)}, \hat{\beta}_{1(i)}, \dots, \hat{\beta}_{q_1(i)}, \hat{\kappa}_{(i)}]^\top$. Uma das primeiras medidas de influência global proposta é definida em termos da norma padronizada de $\hat{\xi}_{(i)} - \hat{\xi}$, conhecida como a distância de Cook generalizada (GCD), que é dada por

$$\text{GCD}_{(i)}(\xi) = [\hat{\xi}_{(i)} - \hat{\xi}]^\top [-\ddot{\ell}(\hat{\xi})][\hat{\xi}_{(i)} - \hat{\xi}] \quad i = 1, \dots, n, \quad (2.9)$$

em que $\ddot{\ell}(\hat{\xi})$ é a matriz Hessiana de $\ell(\xi)$ avaliada em $\hat{\xi}$. De forma alternativa, pode-se computar $\text{GCD}_i(\omega)$, $\text{GCD}_i(\alpha_m)$, $\text{GCD}_i(\beta_s)$ e $\text{GCD}_i(\kappa)$, com $m = 1, \dots, p_1$ e $s = 1, \dots, q_1$, cujos valores revelam o impacto do caso i nas estimativas de ω , α_m , β_s e κ . Uma outra forma de comparação entre $\hat{\xi}_{(i)}$ e $\hat{\xi}$ é a distância entre as verossimilhança (LD), que é descrita como $\text{LD}_i(\xi) = 2[\ell(\hat{\xi}) - \ell(\hat{\xi}_{(i)})]$, para $i = 1, \dots, n$.

2.5 Influência local

A ideia base da influência local é utilizar o conceito de curvatura da superfície na avaliação do comportamento local da função de deslocamento de verossimilhança. Seja $\varpi =$

$[\varpi_{(1)}, \dots, \varpi_{(n)}]^\top$ um vetor de perturbação, e $\ell(\boldsymbol{\xi}_\varpi)$ a função de log verossimilhança do modelo perturbado por ϖ , o deslocamento da verossimilhança (LL) é dado por $LL_i(\boldsymbol{\xi}) = 2[\ell(\widehat{\boldsymbol{\xi}}) - \ell(\widehat{\boldsymbol{\xi}}_\varpi)]$, para $i = 1, \dots, n$, em que $\widehat{\boldsymbol{\xi}}_\varpi$ corresponde ao estimador de MV de $\boldsymbol{\xi}$ para o modelo com perturbação. O comportamento de LL_i em torno de ϖ_0 , ou seja, o vetor sem perturbação foi estudado por Cook (1987), que demonstrou a curvatura da normal para $\boldsymbol{\xi}$ na direção \mathbf{d} , com $\|\mathbf{d}\| = 1$, é expressa como $C_{(d)}(\boldsymbol{\xi}) = 2|\mathbf{d}^\top \nabla^\top \ddot{\ell}(\widehat{\boldsymbol{\xi}})^{-1} \nabla \mathbf{d}|$, em que ∇ é uma matriz $[1 + p_1 + q_1] \times n$ de perturbações com elementos

$$\nabla_{ji} = \left. \frac{\partial^2 \ell(\boldsymbol{\xi}_\varpi)}{\partial \xi_j \partial \varpi_{(i)}} \right|_{\boldsymbol{\xi}=\widehat{\boldsymbol{\xi}}, \varpi=\varpi_0}, \quad j = 1, \dots, p_1 + q_1, \quad i = 1, \dots, n. \quad (2.10)$$

Assim, temos que a avaliação de observações potencialmente influentes através dos métodos de influência local é geralmente baseado em gráficos de índices. Consideramos o gráfico de índice do autovetor \mathbf{d}_{\max} que corresponde ao maior autovetor

$$\ddot{\mathbf{F}} = -\nabla^\top \ddot{\ell}(\widehat{\boldsymbol{\xi}})^{-1} \nabla, \quad (2.11)$$

em que $C_{\mathbf{d}_{\max}}(\boldsymbol{\xi})$ indica quais os casos que são potencialmente influentes em $\widehat{\boldsymbol{\xi}}$. No entanto, podemos estar interessado apenas $\widehat{\boldsymbol{\omega}}, \widehat{\boldsymbol{\alpha}} = (\widehat{\alpha}_{1(i)}, \dots, \widehat{\alpha}_{p_1(i)})^\top, \widehat{\boldsymbol{\beta}} = (\widehat{\beta}_{1(i)}, \dots, \widehat{\beta}_{q_1(i)})^\top$ ou $\widehat{\boldsymbol{\kappa}}$ tal que a curvatura da normal na direção \mathbf{d} é $C_{(d)}(\widehat{\boldsymbol{v}}) = 2|\mathbf{d}^\top \nabla^\top [\ddot{\ell}(\widehat{\boldsymbol{\xi}})^{-1} - \ddot{\ell}(\widehat{\boldsymbol{v}})] \nabla \mathbf{d}|$, com $\mathbf{v} = \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}$, e

$$\ddot{\ell}(\widehat{\boldsymbol{\omega}}) = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & \ddot{\ell}(\widehat{\boldsymbol{\alpha}})^{-1} & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \ddot{\ell}(\widehat{\boldsymbol{\beta}})^{-1} & \mathbf{0} \\ \cdot & \cdot & \cdot & \ddot{\ell}(\widehat{\boldsymbol{\kappa}})^{-1} \end{bmatrix}, \quad \ddot{\ell}(\widehat{\boldsymbol{\alpha}}) = \begin{bmatrix} \ddot{\ell}(\widehat{\boldsymbol{\omega}})^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \ddot{\ell}(\widehat{\boldsymbol{\beta}})^{-1} & \mathbf{0} \\ \cdot & \cdot & \cdot & \ddot{\ell}(\widehat{\boldsymbol{\kappa}})^{-1} \end{bmatrix},$$

$$\ddot{\ell}(\widehat{\boldsymbol{\beta}}) = \begin{bmatrix} \ddot{\ell}(\widehat{\boldsymbol{\omega}})^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & \ddot{\ell}(\widehat{\boldsymbol{\alpha}})^{-1} & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \cdot & \ddot{\ell}(\widehat{\boldsymbol{\kappa}})^{-1} \end{bmatrix}, \quad \ddot{\ell}(\widehat{\boldsymbol{\kappa}}) = \begin{bmatrix} \ddot{\ell}(\widehat{\boldsymbol{\omega}})^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & \ddot{\ell}(\widehat{\boldsymbol{\alpha}})^{-1} & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \ddot{\ell}(\widehat{\boldsymbol{\beta}})^{-1} & \mathbf{0} \\ \cdot & \cdot & \cdot & \mathbf{0} \end{bmatrix}.$$

Podemos, também, considerar a direção de $\mathbf{d} = \mathbf{e}_{in}$, com \mathbf{e}_{in} sendo um vetor $n \times 1$ de zeros com um na posição i , ou seja, a base canônica de \mathbb{R}^n , com $\{\mathbf{e}_{in}, 1 \leq i \leq n\}$. Com isso, a curvatura

normal é dada por $C_i(\boldsymbol{\xi}) = 2\mathbf{e}_{in}|\ddot{\mathbf{F}}|\mathbf{e}_{in} = 2|\ddot{F}_{ii}|$, para $i = 1, \dots, n$, em que \ddot{F}_{ii} é i -ésimo elemento da diagonal $\ddot{\mathbf{F}}$. Portanto, se $C_i(\hat{\boldsymbol{\xi}}) > 2\bar{C}(\hat{\boldsymbol{\xi}})$, onde $\bar{C}(\hat{\boldsymbol{\xi}}) = \sum_{i=1}^n C_i(\hat{\boldsymbol{\xi}})/n$, então a observação i é considerada como potencialmente influente. Este procedimento é denominado o método de influência local total, ver Lesaffre & Verbeke (1998).

2.5.1 Perturbação de casos

Neste tipo de perturbação, tem como objetivo observar se a ponderação nas observações influencia na estimativa de ML do vetor de parâmetros $\boldsymbol{\xi}$. Consideremos o vetor de pesos (ponderações) $\boldsymbol{\varpi} = [\varpi_{(1)}, \dots, \varpi_{(n)}]^\top$, em que a função log verossimilhança será dada por $\ell(\boldsymbol{\xi}, \boldsymbol{\varpi}) = \sum_{t=1}^n \varpi_t \ell_t(\boldsymbol{\xi})$, com $0 \leq \varpi_i \leq 1$, para $t = 1, \dots, n$ e $\boldsymbol{\varpi}_0 = [1, \dots, 1]^\top$, em que $\ell_t(\boldsymbol{\xi})$ é dada por

$$\begin{aligned} \ell_t(\boldsymbol{\xi}) = & -\log(2\sqrt{2\pi}) - \log(\kappa) - \frac{3}{2}\log(r_t) + \log(r_t + \sigma_t) \\ & - \frac{1}{2}\log(\sigma_t) - \frac{1}{2\kappa^2} \left(\frac{r_t}{\sigma_t} + \frac{\sigma_t}{r_t} - 2 \right), \quad r_t > 0. \end{aligned} \quad (2.12)$$

2.5.2 Perturbação na resposta

Assumindo-se que há uma perturbação aditiva para cada caso do modelo descrito em (2.2) teremos que

$$R_t(\boldsymbol{\omega}_t) = R_t + \boldsymbol{\omega}_t s(R_t) = \sigma_t(\boldsymbol{\omega}_t) \varepsilon_t, \quad \varepsilon_t \sim BS(\kappa, 1), \quad (2.13)$$

com

$$\log(\sigma_t(\boldsymbol{\omega}_t)) = \omega + \sum_{j=1}^q \alpha_j \log(\sigma_{t-j}(\boldsymbol{\omega}_t)) + \sum_{j=1}^p \beta_j \left(\frac{R_{t-j}(\boldsymbol{\omega}_{t-j})}{\sigma_{t-j}(\boldsymbol{\omega}_{t-j})} \right), \quad (2.14)$$

em que $s(R_t)$ é um fator de escala e $\boldsymbol{\omega}_t \in \mathbb{R}$, para $i = 1, \dots, n$. Note que $R_t(\boldsymbol{\omega}_t) \sim BS(\kappa, \sigma(\boldsymbol{\omega}_t))$ e sua correspondente função do log-verossimilhança é dada por $\ell(\boldsymbol{\xi}, \boldsymbol{\varpi}) = \sum_{t=1}^n \ell_{\boldsymbol{\omega}_t}(\boldsymbol{\xi})$, com

$$\begin{aligned} \ell_{\boldsymbol{\omega}_t}(\boldsymbol{\xi}) = & -\log(2\sqrt{2\pi}) - \log(\kappa) - \frac{3}{2}\log[r_t(\boldsymbol{\omega}_t)] + \log[r_t(\boldsymbol{\omega}_t) + \sigma_t(\boldsymbol{\omega}_t)] \\ & - \frac{1}{2}\log[\sigma_t(\boldsymbol{\omega}_t)] - \frac{1}{2\kappa^2} \left(\frac{r_t(\boldsymbol{\omega}_t)}{\sigma_t(\boldsymbol{\omega}_t)} + \frac{\sigma_t(\boldsymbol{\omega}_t)}{r_t(\boldsymbol{\omega}_t)} - 2 \right), \end{aligned} \quad (2.15)$$

com $r_t(\varpi_t) > 0$ e $\varpi_0 = [0, \dots, 0]^\top$.

2.6 Perturbação renovada

Nesse caso, temos que a função log-verossimilhança dada em (2.7), considera a suposição que os ε_t 's são variáveis aleatórias independentes e identicamente distribuídas com distribuição BS. No entanto, o ajuste do modelo pode ser fortemente afetado pela existência de observações influentes. Assumindo um vetor de perturbação ϖ é introduzido em (2.2)

$$Z_t = R_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim \text{BS}(\kappa, 1/\varpi_t), \quad (2.16)$$

com

$$\log(\sigma_t) = \omega + \sum_{j=1}^p \alpha_j \log(\sigma_{t-j}) + \sum_{j=1}^q \beta_j \left(\frac{R_{t-j}}{\sigma_{t-j}} \right).$$

A função log-verossimilhança correspondente é dada por $\ell(\xi_{\varpi}) = \sum_{t=1}^n \ell_{\varpi_t}(\xi)$, é dada por

$$\begin{aligned} \ell_{\varpi_t}(\xi) = & -\log(2\sqrt{2\pi}) - \log(\kappa) - \frac{3}{2} \log(r_t) + \log(r_t + \sigma_t \varpi_t^{-1}) \\ & - \frac{1}{2} \log(\sigma_t / \varpi_t) - \frac{1}{2\kappa^2} \left(\frac{r_t \varpi_t}{\sigma_t} + \frac{\sigma_t}{\varpi_t r_t} - 2 \right), \end{aligned} \quad (2.17)$$

para $r_t > 0$ e $\varpi_0 = [1, \dots, 1]^\top$.

Capítulo 3

Avaliação numérica de dados

3.1 Estudo de Simulação

Foi realizado um estudo de simulação de Monte Carlo (MC) implementado no software R e disponível em www.r-project.org para avaliar o desempenho do estimador de máxima verossimilhança para o modelo BSCARR($p = 1, q = 1$). Consideramos como tamanhos de amostras $n \in \{250, 500, 1000, 2000\}$, os verdadeiros valores para o vetor de parâmetros é $[\omega, \alpha, \beta, \kappa] = [0.3, 0.7, 0.3, 1.1]$, com 5000 réplicas de MC para cada tamanho de amostra. Foram obtidas as seguintes estatísticas da amostra: média empírica, coeficientes de assimetria e curtose, viés, erro quadrático médio (EQM). Através dos resultados apresentados na Tabela 3.1 podemos concluir que os estimadores apresentam um comportamento esperado, ou seja, considerando as estatísticas obtidas nota-se que os valores destas diminuem conforme aumenta o tamanho da amostral.

Tabela 3.1: Estatísticas da análise de simulação para a estimativa do modelo BSCARR (1,1).

n		250	500	1,000	2,000
ω	Valor Real	0.3000	0.3000	0.3000	0.3000
	Média	0.3108	0.3062	0.3039	0.3020
	Assimetria	0.6708	0.3839	0.3478	0.1817
	Curtose	4.2165	3.4455	3.7494	3.2913
	Viés($\times 10^{-4}$)	0.0108	0.0062	0.0039	0.0020
	EQM($\times 10^{-4}$)	0.0297	0.0134	0.0063	0.0032
α	Valor Real	0.7000	0.700	0.7000	0.7000
	Média	0.6708	0.6943	0.6968	0.6985
	Assimetria	-0.6188	-0.3761	-0.5315	-0.0710
	Curtose	4.2758	3.8217	6.1960	3.2823
	Viés($\times 10^{-4}$)	-0.0108	0.5789	-0.0032	-0.0015
	EQM($\times 10^{-4}$)	0.0046	0.0020	0.0010	0.0005
β	Valor Real	0.3000	0.3000	0.3000	0.3000
	Média	0.3081	0.3062	0.3019	0.3008
	Assimetria	0.0074	0.241	-0.5127	-2.2001
	Curtose	3.4106	3.3324	9.4656	37.5334
	Viés($\times 10^{-4}$)	0.0081	0.2454	0.0019	0.0008
	EQM($\times 10^{-4}$)	0.0014	0.0006	0.0003	0.0002
κ	Valor Real	1.1000	1.1000	1.1000	1.1000
	Média	1.1095	1.1052	1.1028	1.1015
	Assimetria	0.3369	0.2239	0.8279	0.6690
	Curtose	3.8519	3.3643	12.2063	10.4207
	Viés($\times 10^{-4}$)	0.0095	0.0211	0.0028	0.0015
	EQM($\times 10^{-4}$)	0.0032	0.0014	0.0007	0.0003

3.2 Análise de dados financeiros

Os estudos foram realizados para três conjuntos de índices de ativos observados diariamente no período de 01/07/2007 a 31/07/2017. Todos os conjuntos de dados foram obtidos do sítio `www.finance.yahoo.com`, em que para cada dia observado, estão disponíveis quatro informações sobre os preços, os preços máximos, mínimos, de abertura e de fechamento.

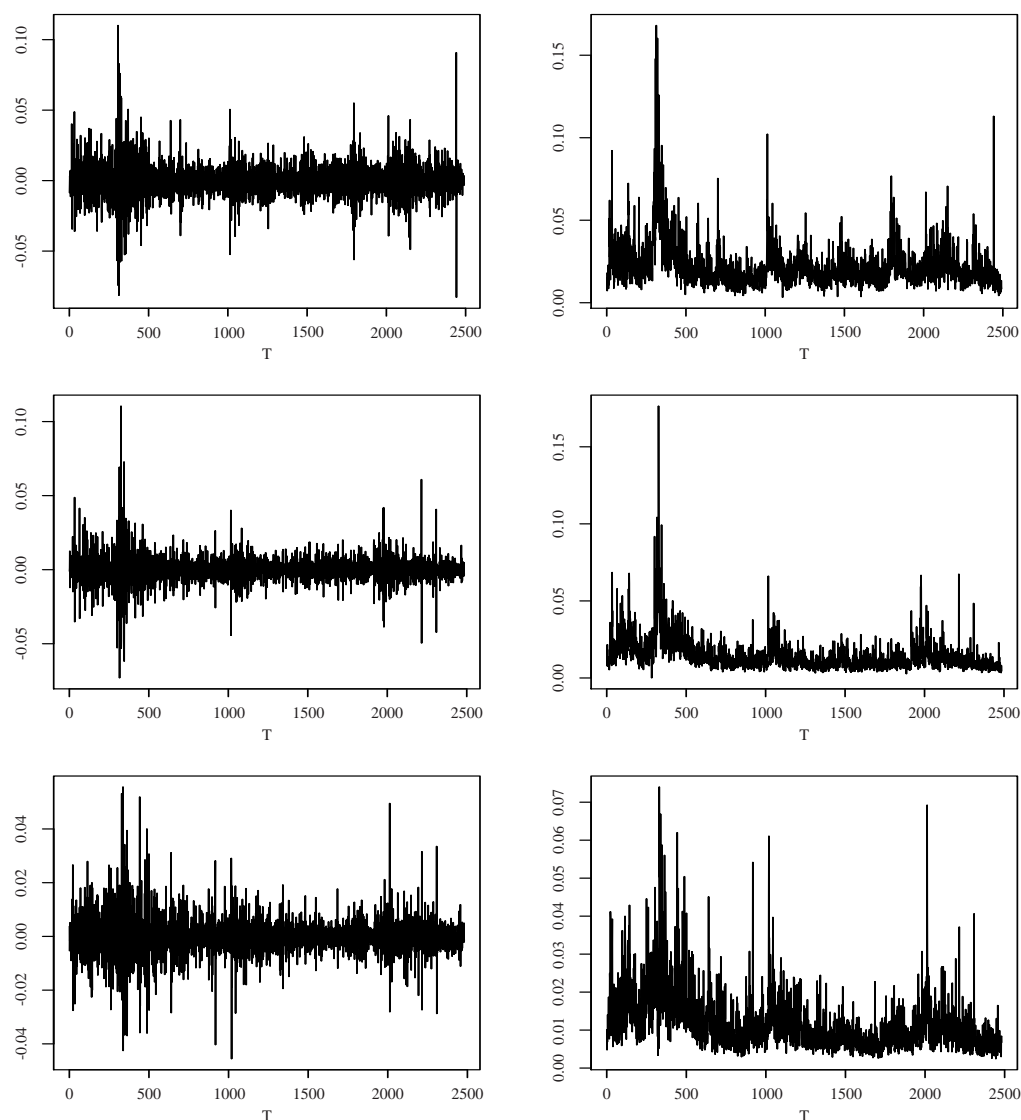


Figura 3.1: Gráficos dos retornos e amplitudes diários dos dados da BVSP, HSI e TSEC (linha), no período de 01/07/2007 a 31/07/2017.

Foram coletados os índices BVSP do Brasil, HSI de Hong Kong e TSEC de Taiwan, em que a amplitude foi construída a partir dos preços máximo e mínimo. Algumas medidas descritivas para a amplitude (variação) dos preços, foram calculadas para os dados considerados na

análise, como a média, desvio padrão, coeficiente de variação, máximo, mínimo, coeficientes de assimetria e curtose, além da estatística Ljung-Box (LB) $Q(\gamma)$ para correlação de ordem γ e p -valores para o teste de normalidade Bera-Jarque (JB) proposto em Jarque & Bera (1987).

Tabela 3.2: Estatísticas descritivas da amplitude dos índices.

	BVSP	HSI	TSEC
Média	0.0221	0.0143	0.0117
Assimetria	3.3501	4.1973	2.4041
Curtose	19.6983	36.5934	9.2602
Desvio Padrão	0.0180	0.0105	0.0077
CV	63.0246	73.3039	66.0675
Mínimo	0.0032	0.0001	0.0024
Máximo	0.1681	0.1765	0.0740
JB (p -valor)	<0.0001	<0.0001	<0.0001
Q(12)	1640.80	1375.00	1138.70

A Tabela 3.2 apresenta valores elevados para a curtose e assimetria, indicando um forte desvio da distribuição normal. Isto pode ser observado através da Figura 3.2 (histograma), que para todas as séries analisadas apresentou uma assimetria positiva dos dados. Através dos resultados dispostos na Tabela 3.2, observamos os valores do teste JB, o qual indica que a suposição de normalidade parece não ser adequada. As estatísticas Ljung-Box Q apresentam valores que giram em torno de 1138.70 (TSEC) à 1640 (BVSP) o que indica uma forte persistência na volatilidade nos dados.

A função de risco de falha (HR) de uma variável aleatória (RV) R é definida por $h(r) = f(r)/[1 - F(r)]$, em que $f(\cdot)$ e $F(\cdot)$ são a PDF e CDF da RV R . Uma forma de caracterizar a HR é através da função de tempo total do teste (TTT). Com isso podemos detectar a forma da HR por meio da curva TTT fornecida pelos dados e assim escolher uma distribuição adequada, ver Aarset (1987). A função TTT é definida por $W(u) = H^{-1}(u)/H^{-1}(1)$, para $0 \leq u \leq 1$, em que $H^{-1}(u) = \int_0^{F^{-1}(u)} [1 - F(y)] dy$, em que F^{-1} é a função inversa de CDF de R . Uma aproximação para W pode ser observada através de um gráfico de pontos $[k/n, W_n(k/n)]$, $r_{(i)}$ sendo a i -ésima estatística de ordem observada, com

$$W_n(k/n) = \frac{\sum_{i=1}^k r_{(i)} + [n - k]r_k}{\sum_{i=1}^k r_{(i)}} \quad i = 1, \dots, n.$$

De acordo com a Figura 3.2 observamos que os TTT-plots sugerem HR com compor-

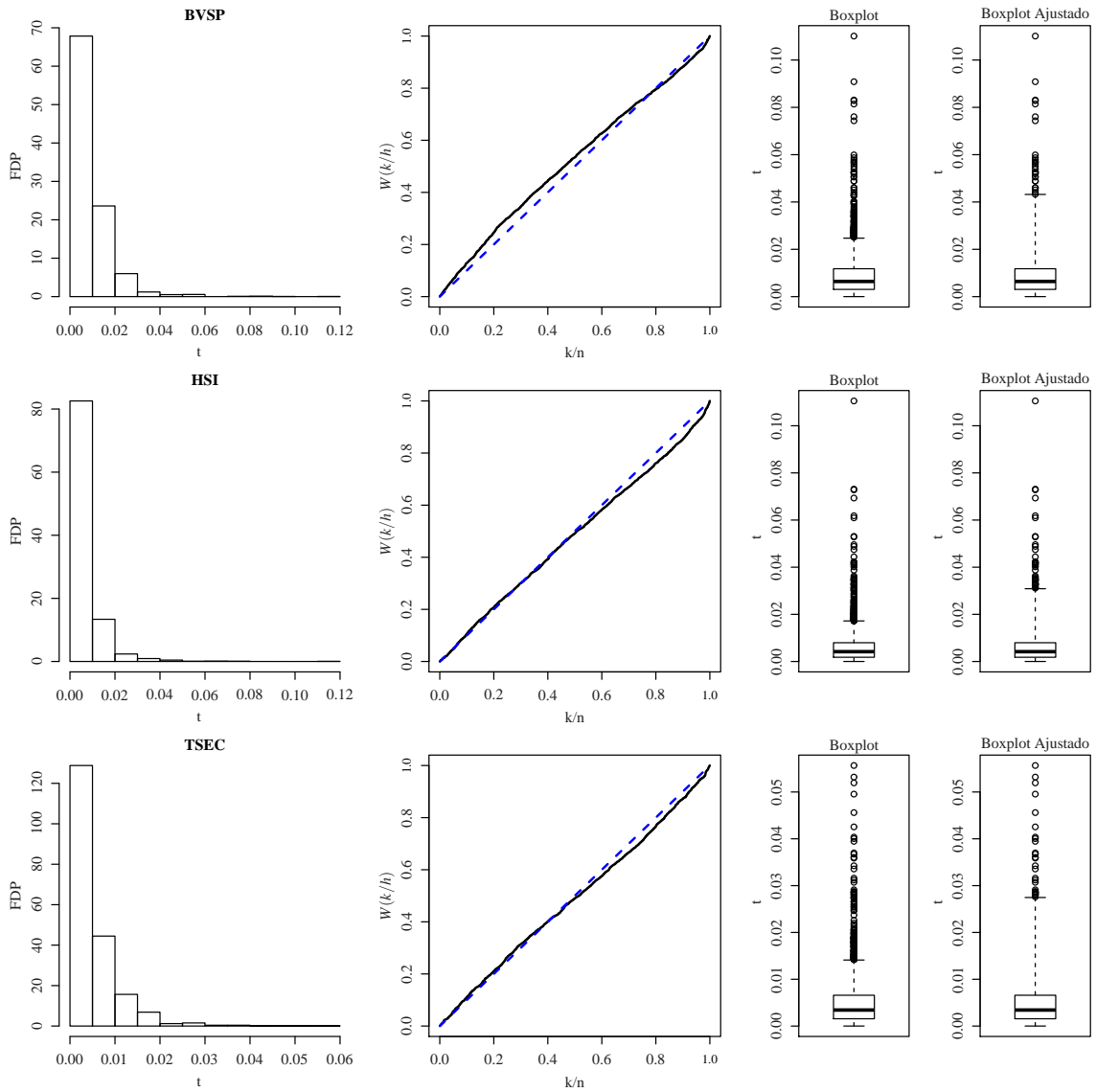


Figura 3.2: Histograma, gráfico TTT e os boxplots dos dados da BVSP do Brasil(BVSP), HSI de HongKong(HSI) e TSEC Taiwan (TSEC)

tamento unimodal. Para confirmar essa detecção, fizemos um pequeno estudo de simulação da distribuição BS. Primeiro, geramos dados de BS com uma configuração similar àquela encontrada em séries financeiras baseadas na amplitude. Com os dados gerados, plotamos a curva teórica HR da BS e o TTT-plot. O gráfico de HR da BS apresentou uma forma unimodal, enquanto o gráfico TTT-plot empírico foi muito semelhante ao encontrado nos conjuntos de dados BVSP, HSI e TSEC. A Figura 3.2 apresenta os boxplots usuais e o ajustado, sendo este último indicado nos casos em que os dados seguem uma distribuição assimétrica, ver Hubert & Van der Vaeken (2008). A partir dessa figura, notamos que alguns potenciais *outliers* identificados pelo boxplot usual não são outliers quando o boxplot ajustado é considerado.

3.2.1 Estimação e validação do modelo

A Tabela 3.3 fornece as estimativas de ML, SEs e p -valores do teste para os parâmetros do modelo BSCARR(1,1). Consideramos que uma estrutura dinâmica simples é satisfatória para o modelo CARR e também consistente com a literatura GARCH, em que uma especificação GARCH (1,1) é suficiente para uma grande classe de retornos de ativos especulativos, ver Chou (2005) e Bollerslev *et al.* (1992). Além disso, apresentamos os critérios de informação de Akaike (AIC) e Bayesiana (BIC). Para comparação, os resultados do modelo WCARR e GCARR são fornecidos. Podemos notar que o modelo CARR baseado na BS fornecem melhores ajustes em comparação com os outros modelos baseados nos valores dos AIC's e BIC's.

A Figura 3.3 exhibe os QQplots do resíduo quantílico (RQ) para os modelos BSCARR, WCARR e GCARR. A partir desta figura, observe que os resíduos do RQ mostram uma boa concordância com a distribuição $N(0,1)$ no modelo BSCARR em todas os conjuntos de dados analisados. Podemos notar que os modelos WCARR e GCARR apresentam comportamento semelhantes no que se refere a análise de resíduos. É importante destacar que tal análise só corrobora os resultados apresentados na Tabela 3.3.

3.2.2 Análise de diagnóstico

Influência global

A Figura 3.4 apresenta o gráfico da GCD discutida no capítulo anterior. Observamos na referida figura, que as estatísticas $GCD_i(\xi)$ apresentam diversas observações como poten-

Tabela 3.3: Estimativa do modelo CARR (1,1) com termo de perturbação com distribuição Birnbaum-Saunders, Weibull e Gama.

	BSCARR(1,1)	BVSP	HSI	TSEC
ω	0.00120	0.00231	0.00048	
p -valor	[0.00025]	[0.00038]	[0.00025]	
α	0.25091	0.20753	0.14910	
p -valor	[0.02893]	[0.02131]	[0.04378]	
β	0.57257	0.28355	0.72120	
p -valor	[0.05179]	[0.09084]	[0.10205]	
κ	1.56159	1.89518	1.68913	
p -valor	[0.02250]	[0.02792]	[0.02451]	
AIC	18856.86	20616.22	21517.38	
BIC	18862.45	20621.80	21522.96	
WCARR(1,1)				
ω	0.00156	0.00194	0.00032	
p -valor	[0.00034]	[0.00031]	[0.00012]	
α	0.25018	0.31725	0.16268	
p -valor	[0.03181]	[0.02942]	[0.03183]	
β	0.57148	0.36291	0.77773	
p -valor	[0.06551]	[0.06952]	[0.05351]	
θ	1.11632	1.04496	1.08543	
p -valor	[0.01813]	[0.01649]	[0.01673]	
AIC	18892.80	20716.98	21565.32	
BIC	18898.37	20722.56	21570.90	
GCARR(1,1)				
ω	0.00155	0.00197	0.00034	
p -valor	[0.00035]	[0.00031]	[0.00013]	
α	0.24908	0.31785	0.16759	
p -valor	[0.03247]	[0.02981]	[0.03365]	
β	0.57503	0.36183	0.76737	
p -valor	[0.06693]	[0.07000]	[0.05665]	
ϕ	1.05404	1.04412	1.17678	
p -valor	[0.09826]	[0.09591]	[0.11635]	
AIC	18894.80	20718.98	21567.34	
BIC	18901.78	20725.96	21574.31	

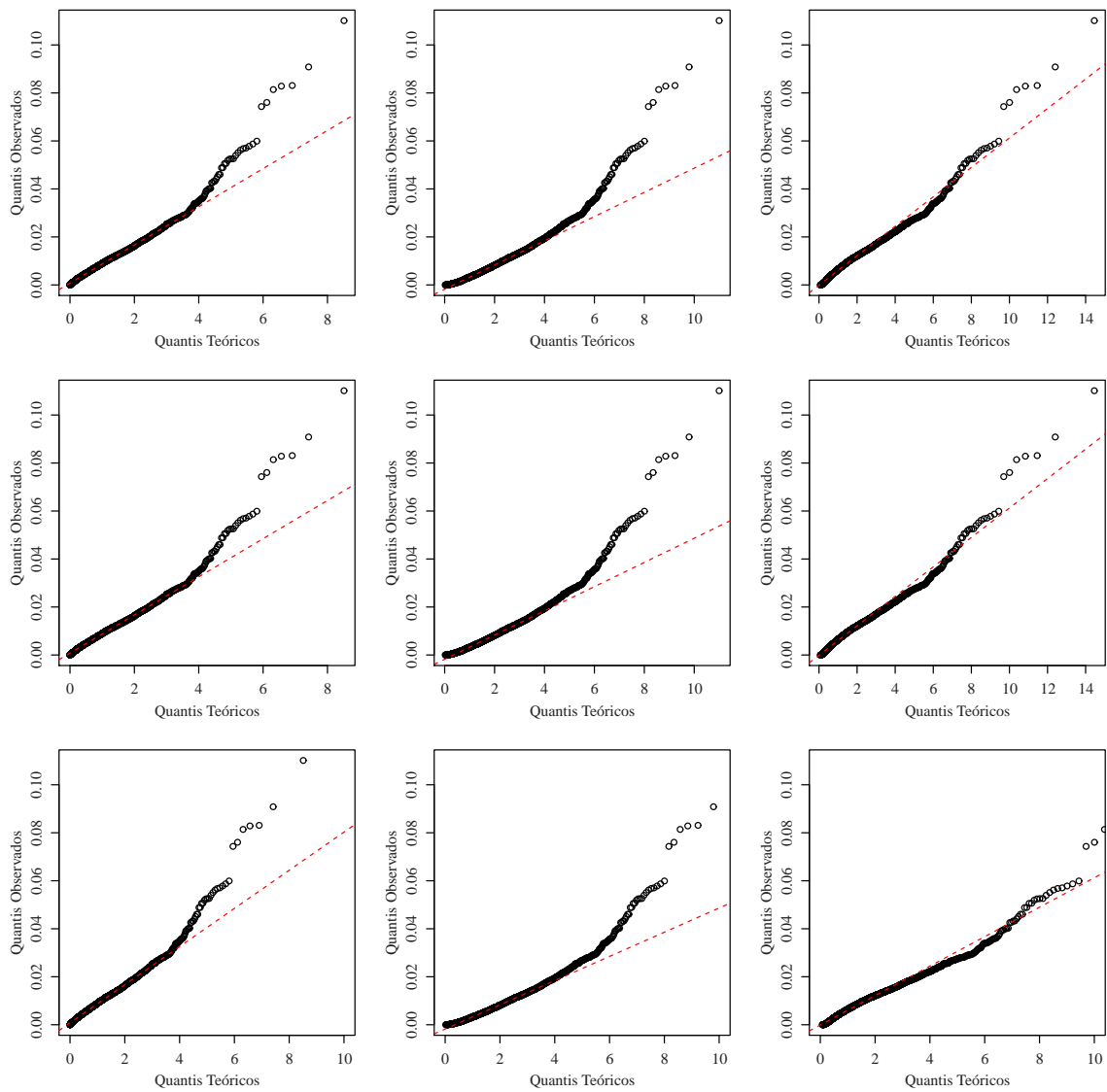


Figura 3.3: QQplot do resíduo RQ para os dados da BVSP, HSI e TSEC, respectivamente, segundo o ajuste dos modelos WCARR, GCARR e BSCARR.

cialmente influentes. Considerando o índice BVSP podemos verificar que os pontos #1028 e #1375 são indicados como possivelmente influentes. Para a segunda série estudada que é a HSI os pontos #1313, #1314 e #1853 aparecem como possíveis *outliers*. Por fim, o índice TSEC e de acordo com a GCD os casos #344, #475, #676 e #2133 estão entre as observações potencialmente influentes.

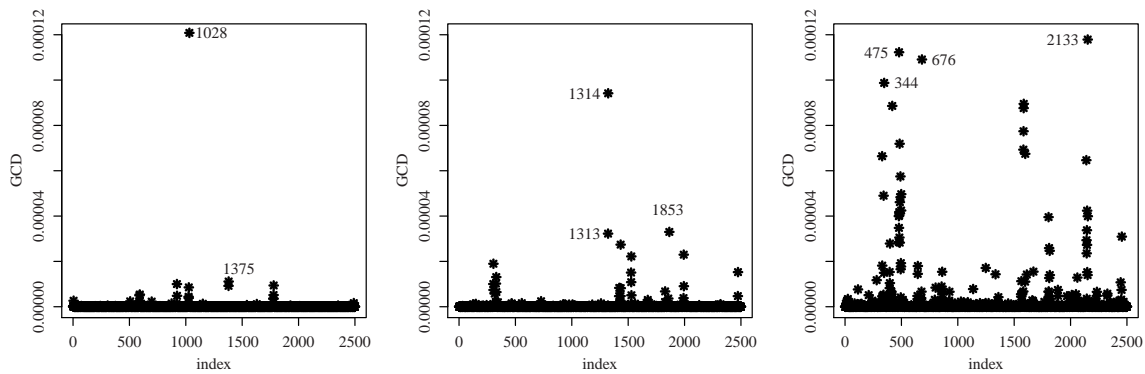


Figura 3.4: Distância de Cook Generalizada para o modelo BSCARR para os dados da BVSP do Brasil(BVSP), HSI de HongKong(HSI) e TSEC Taiwan (TSEC) respectivamente.

Influência local

Perturbação de casos

De acordo com o que mencionamos no capítulo anterior, temos que a perturbação do peso dos casos, avalia a contribuição do peso para as observações e sua influência na estimativa ML do vetor de parâmetros ξ . Os gráficos de índice C_i sob este esquema são apresentados na Figura 3.5, em que se destacaram como observações potencialmente influentes, os casos #1028, #1375 e #1773 para os índices da BVSP, para o HSI foram verificadas as observações #328, #1314, #1853 e #1773, e entre o índice TSEC foram destacadas as #397, #482, #643 e #856.

Perturbação na resposta

A Figura 3.6 indica os casos que foram considerados potencialmente influentes para este tipo de perturbação. Deste modo, destacam-se as observações #591, #1028 e #1777 para o índice BVSP, os casos #1314, #1853 e #1984 para o HSI e, por fim, as observações #397, #482, #856 e #1575 considerando o índice TSEC.

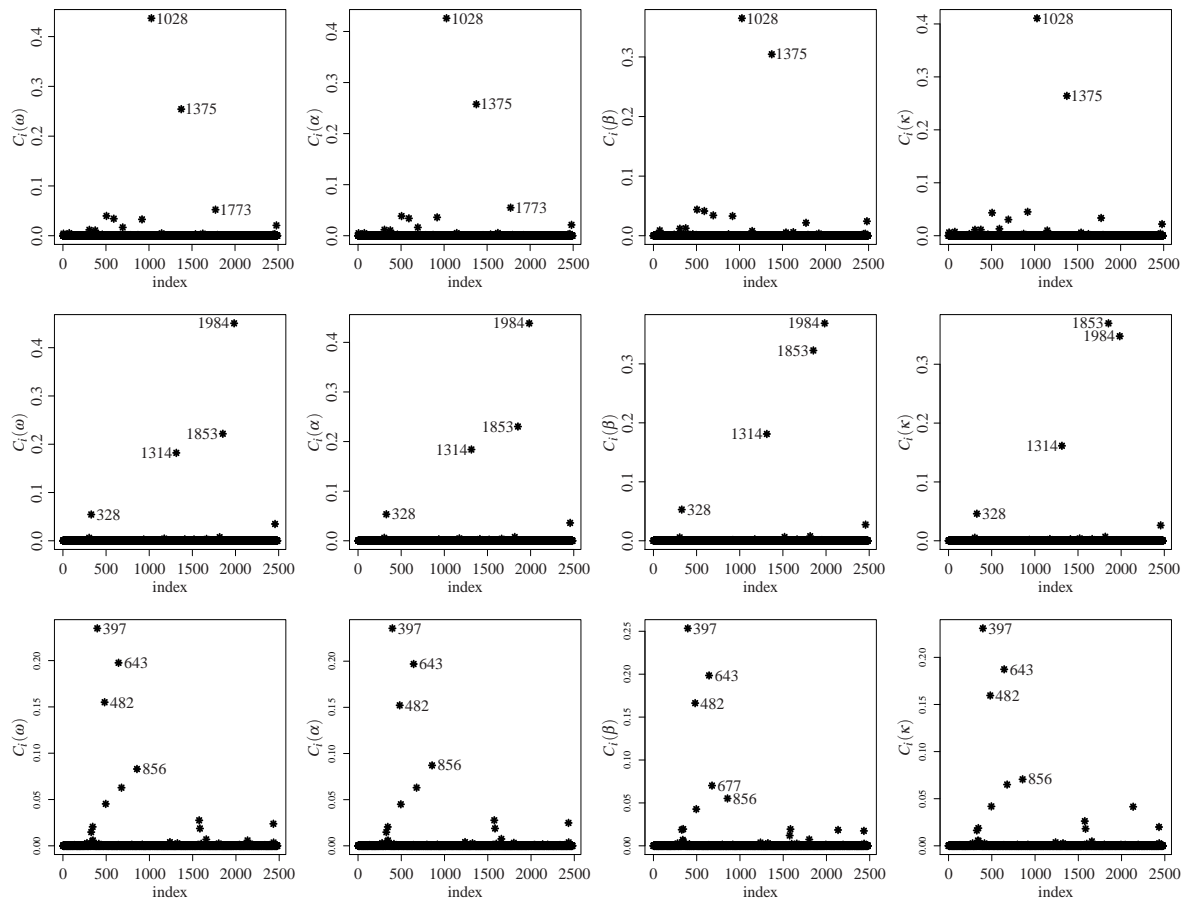


Figura 3.5: Índice C_i para o vetor de parâmetros sob perturbação de casos no modelo BSCARR com os dados BVSP, HSI e TSEC

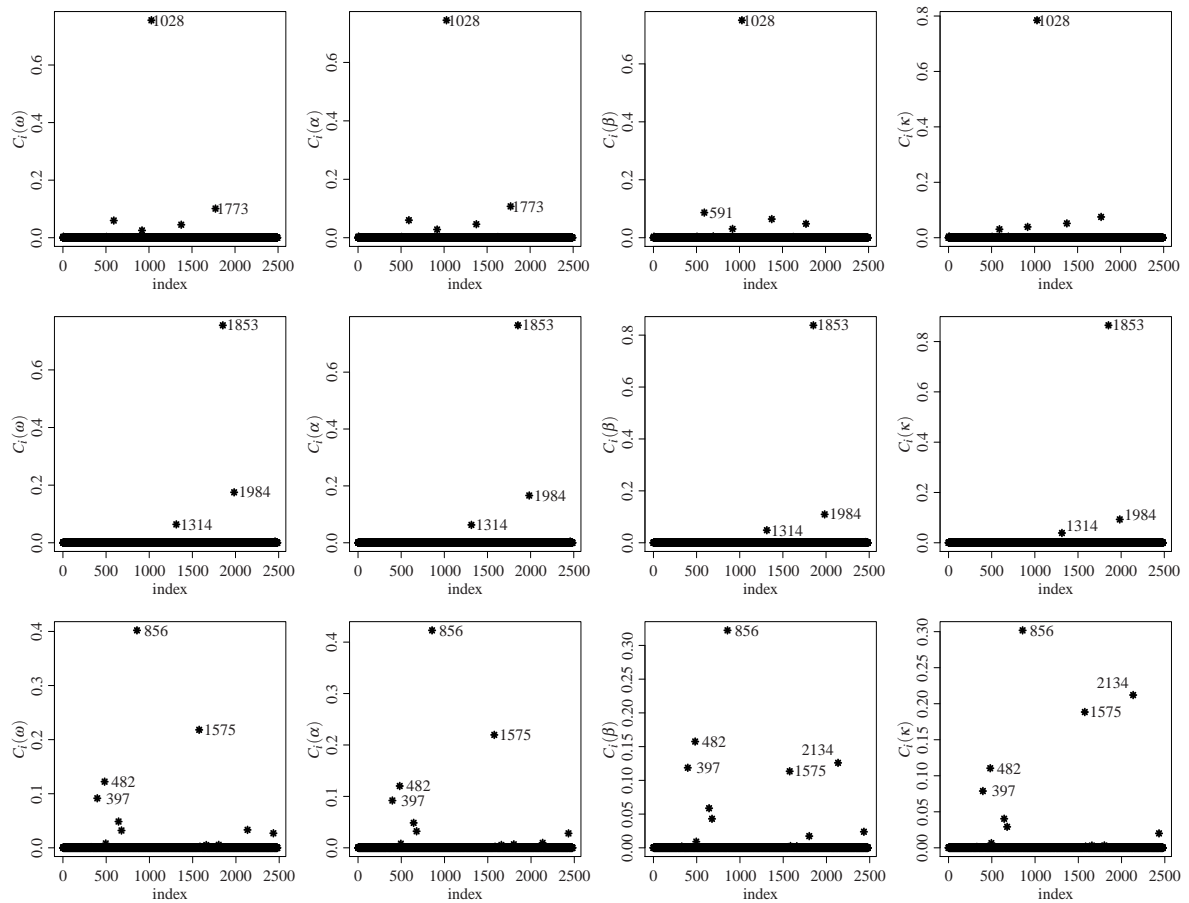


Figura 3.6: Índice C_i para o parâmetro sob perturbação na Variável Resposta no modelo BSCARR com os dados BVSP, HSI e TSEC.

Perturbação renovada

A Figura 3.7 apresenta os gráficos de índices C_i sob perturbação renovada. Com isso, de maneira similar ao caso de perturbação na resposta, também foram detectados como possíveis observações influentes os pontos #1028 e #1777, adicionando-se a observação #1375 para o índice BVSP. Comportamento análogo é observado ao dados da série HSI, que além dos valores #1314, #1853 e #1984, ainda agregou as observações #328 e #2458. Já para o índice de preços TSCE destacamos quatro possíveis pontos potencialmente influentes que são: #397, #482 e #856 e, neste caso, inclui-se o ponto #643.

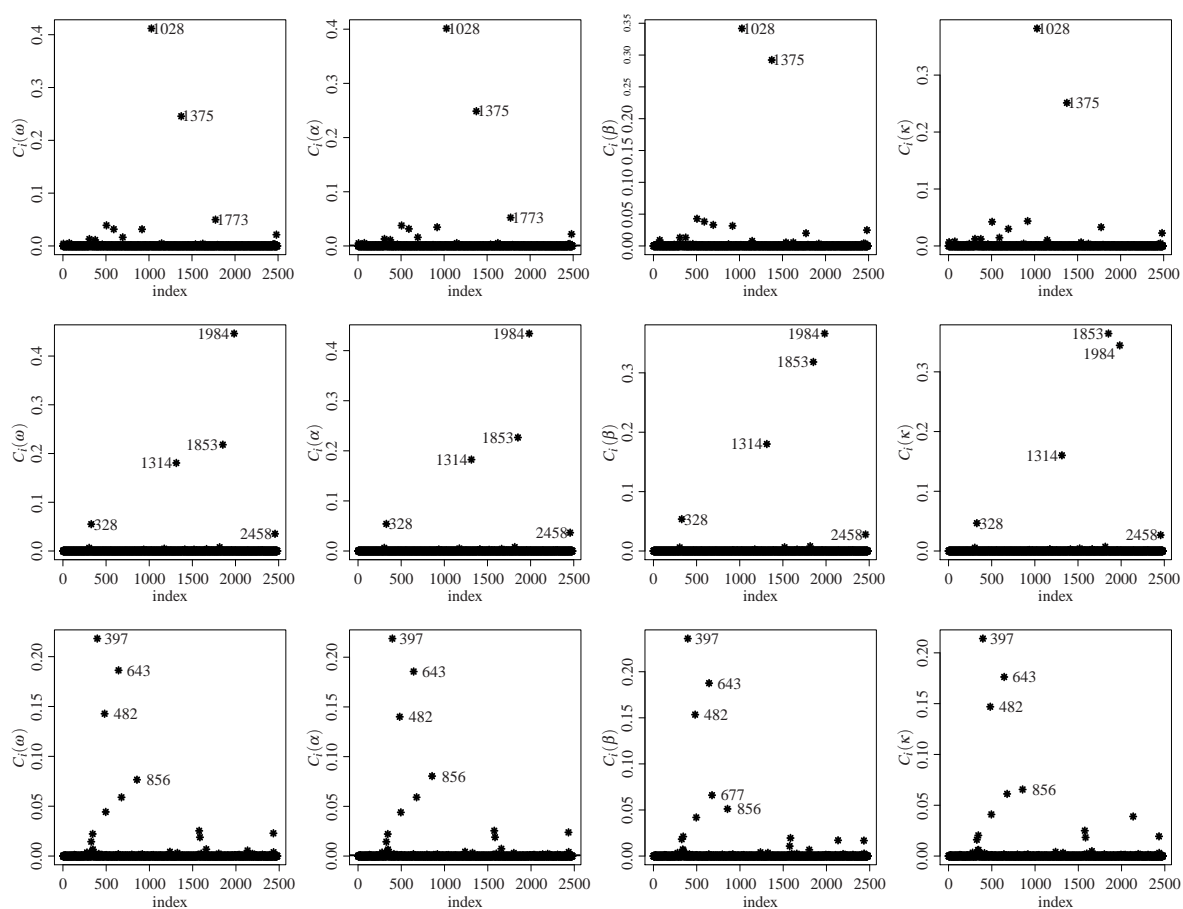


Figura 3.7: Índice C_i para o parâmetro sob perturbação na resposta dos dados no modelo BSCARR com os dados BVSP, HSI e TSEC

Mudança Relativa

Uma vez que se detecta os casos discrepantes é fundamental avaliar o impacto dos mesmos nas estimativas dos parâmetros do modelo. Para avaliar este impacto calculamos a

mudança relativa (RC), que é obtida através da remoção dos possíveis casos influentes, em que os parâmetros do modelo são reestimados sem o ponto potencialmente influente. A RC é expressa por

$$RC(\xi_{j(i)}) = \left| \frac{\widehat{\xi}_j - \widehat{\xi}_{j(i)}}{\widehat{\xi}_j} \right| \times 100\%, \quad RC(SE(\theta_{j(i)})) = \left| \frac{\widehat{SE}(\widehat{\xi}_j) - \widehat{SE}(\widehat{\xi}_{j(i)})}{\widehat{SE}(\widehat{\xi}_j)} \right| \times 100\%,$$

em que $\widehat{\xi}_{j(i)}$ e $\widehat{SE}(\widehat{\xi}_{j(i)})$ são as estimativas de ML de ξ_j e seu correspondente SE, respectivamente, depois de eliminar o caso i , para $j = 1, \dots, 4$ e $i = 1, \dots, n$, com $\xi_1 = \omega$, $\xi_2 = \alpha$, $\xi_3 = \beta$ e $\xi_4 = \kappa$. A Tabela 3.4 apresenta os RC's das estimativas dos parâmetros, seus correspondentes SE's estimados e os p -valores para o teste t após a retirada da i -ésimo observação. De acordo com os resultados podemos observar para o índice BVSP que o maior valor do RC está relacionado a retirada do ponto #1828. Já para o índice HSI podemos notar um valor elevado RC's nos parâmetros α e β após a remoção das observações #1313, #1314 e #1853. Isto se repete com maior intensidade nos casos #397, #482, #643 e #856 para os dados TSEC.

É importante observar que não há uma mudança na significância dos parâmetros, mesmo com a remoção e reestimação das observações indicadas como potencialmente influentes, para todos os dados analisados. Isso pode ser visto considerando os p -valores significativos para o teste t ao nível de 5% apresentados na Tabela 3.4 quando comparados com os resultados apresentados na Tabela 3.3. Este resultado pode ser interpretado como uma robustez do modelo BSCARR para observações atípicas, ou seja, as medidas usadas na análise de diagnóstico identificam casos potencialmente influentes, mas não alteram a inferência do modelo proposto.

Desempenho de Previsão

Nesta subseção apresentamos as duas maneiras para avaliar o desempenho da capacidade preditiva do modelo proposto: (i) técnica de previsão usual, um método bastante usado para os modelos GARCH e ACD, ver Tsay (2009); (ii) método da densidade preditiva proposto Diebold *et al.* (1998). Ambos os métodos podem ser adaptados para o modelo CARR para qualquer distribuição que seja assumida na modelagem. A seguir apresentamos alguns detalhes metodológicos de cada uma das abordagens considerando o modelo BSCARR

Tabela 3.4: RCs (%) nas estimativas de ML e SEs correspondentes para o parâmetro indicado e os casos retirados, e respectivos valores de p .

Retirada de caso(s)	modelo BSCARR			
	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\kappa}$
BVSP				
{#1028}				
RC($\theta_{j(i)}$)	60.0838	0.0299	31.0275	2.7498
RC(SE($\theta_{j(i)}$))	(38.1371)	(0.0196)	(26.9066)	(3.1306)
p -valor	[<0.001]	[<0.001]	[<0.001]	[<0.001]
{#1375}				
RC($\theta_{j(i)}$)	0.1725	10.4496	9.1290	0.9283
RC(SE($\theta_{j(i)}$))	(5.5110)	(3.3542)	(0.9618)	(0.9854)
p -valor	[<0.001]	[<0.001]	[<0.001]	[<0.001]
{#1028,#1375}				
RC($\theta_{j(i)}$)	46.7437	2.1619	23.6784	3.5477
RC(SE($\theta_{j(i)}$))	(21.0477)	(6.1991)	(8.8216)	(3.9197)
p -valor	[<0.001]	[<0.001]	[<0.001]	[<0.001]
HSI				
{#1313}				
RC($\theta_{j(i)}$)	4.1129	17.0976	19.6044	0.6553
RC(SE($\theta_{j(i)}$))	(1.1432)	(4.7453)	(13.4101)	(0.6958)
p -valor	[<0.001]	[<0.001]	[0.004]	[<0.001]
{#1314}				
RC($\theta_{j(i)}$)	1.5619	15.4597	32.2408	2.2634
RC(SE($\theta_{j(i)}$))	(1.0396)	(2.2561)	(16.7072)	(2.6417)
p -valor	[<0.001]	[<0.001]	[0.005]	[<0.001]
{#1853}				
RC($\theta_{j(i)}$)	1.7339	22.8680	1.9157	7.5798
RC(SE($\theta_{j(i)}$))	(3.8020)	(2.3022)	(4.0195)	(8.7351)
p -valor	[<0.001]	[<0.001]	[<0.001]	[<0.001]
{#1984}				
RC($\theta_{j(i)}$)	9.3451	3.2830	18.1811	0.0340
RC(SE($\theta_{j(i)}$))	(1.0317)	(3.2986)	(14.5979)	(0.0477)
p -valor	[<0.001]	[<0.001]	[0.003]	[<0.001]
{#1313,#1314,#1853,#1984}				
RC($\theta_{j(i)}$)	3.3988	33.7931	8.9351	9.7635
RC(SE($\theta_{j(i)}$))	(38.8619)	(0.8938)	(4.5629)	(11.1099)
p -valor	[<0.001]	[<0.001]	[<0.001]	[<0.001]
TSEC				
{#397}				
RC($\theta_{j(i)}$)	64.4500	142.0373	297.0022	418.1710
RC(SE($\theta_{j(i)}$))	(0.8300)	(12.0066)	(10.9500)	(1.3000)
p -valor	[<0.002]	[<0.001]	[<0.001]	[<0.001]
{#482}				
RC($\theta_{j(i)}$)	43.0000	196.9530	130.3041	758.8030
RC(SE($\theta_{j(i)}$))	(60.6800)	(41.4100)	(55.2000)	(2.3700)
p -valor	[<0.004]	[<0.001]	[<0.001]	[<0.001]
{#643}				
RC($\theta_{j(i)}$)	13.6500	59.3300	355.1620	318.6980
RC(SE($\theta_{j(i)}$))	(40.9300)	(32.2800)	(39.7900)	(0.9800)
p -valor	[<0.005]	[<0.001]	[<0.001]	[<0.001]
{#856}				
RC($\theta_{j(i)}$)	73.0100	518.8510	223.2164	578.7290
RC(SE($\theta_{j(i)}$))	(12.0070)	(21.4700)	(6.9600)	(1.8200)
p -valor	[<0.003]	[<0.001]	[<0.001]	[<0.001]
{#397,#482,#643,#856}				
RC($\theta_{j(i)}$)	15.4700	644.7370	140.0110	212.5570
RC(SE($\theta_{j(i)}$))	(45.6700)	(28.3800)	(43.0800)	(6.500)
p -valor	[<0.003]	[<0.001]	[<0.001]	[<0.001]

Técnica de previsão usual: Consideremos a previsão um passo a frente para o BSCARR dada por $r_{h+1} = \sigma_{h+1} \varepsilon_{h+1}$, para uma previsão com a origem em $i = h$, em que $\sigma_{h+1} = \exp(\omega + \alpha \log(\sigma_h) + \beta r_h / \sigma_h)$. Seja $r_h(1)$ a previsão a um passo a frente com origem h , em que $r_h(1) = F_{BS}^{-1}(0.5 | \Omega_h) = \sigma_{h+1}$. Assim, podemos obter a previsão j -passos a frente ($j > 1$) usando $r_{h+j} = \sigma_{h+j} \varepsilon_{h+j}$, em que $r_{h+j} = R_h(j) = \mathbb{E}[\sigma_{h+j} \varepsilon_{h+j}]$, assim, para $j = 1$

$$\begin{aligned} \mathbb{E}[\sigma_{h+j} \varepsilon_{h+j}] &= \mathbb{E} \left[\exp \left(\omega + \alpha \log(\sigma_{h+j-1}) + \beta \left(\frac{r_{h+j-1}}{\lambda_{h+j-1}} \right) \right) \right] \\ &= \mathbb{E} \left[\exp(\omega + \alpha \log(\sigma_{h+j-1})) \exp \left(\beta \frac{R_{h+j-1}}{\sigma_{h+j-1}} \right) \right] \end{aligned}$$

em que $\beta \frac{R_{h+j-1}}{\sigma_{h+j-1}} = \beta \varepsilon_{h+j-1}$. Fazendo $\mathbb{E}[\exp(g(\varepsilon_{h+j-1}))] = \Xi_1$, com $\exp(g(\varepsilon_{h+j-1})) = \beta \varepsilon_{h+j-1}$, temos que $R_h(1) = \Xi_1 \mathbb{E}[\exp(\omega + \alpha \log(\sigma_{h+j-1}))]$. Para $m = 1, \dots, j-1$, $r_h(j)$ pode ser escrito como

$$r_h(j) = \mathbb{E}[\sigma_{h+j} \varepsilon_{h+j}] = \Xi_1 \Xi_2 \dots \Xi_{j-1} \left(\exp \left(\frac{\omega[1-\alpha^{j-1}]}{1-\alpha} + \alpha^{j-1} \log(\sigma_{h+1}) \right) \right), \quad (3.1)$$

onde $\Xi_m = \mathbb{E}[\exp(\alpha^{m-1} g(\varepsilon_{h+j-m}))]$.

Os resultados apresentados na Tabela 3.5 são referentes ao Erro Quadrático Médio (MSE) das previsões dos modelos para 6 passos a frente. A proximidade nos valores apresentados indicam um bom desempenho dos modelos BSCARR em relação aos outros modelos, o que confirma uma boa capacidade preditiva dos três modelos estudados. Através Tabela 3.5 é possível observar que à medida em que aumentamos o número de passos a frente, o EMQ aumenta para todos os modelos estudados, o que expressa um comportamento já esperado. Contudo é possível observar, também, que dentre os modelos analisados para todas as séries referente à amplitude dos índices de preços abordados, o modelo BSCARR é o que apresentou, de maneira geral, os menores valores para o EQM em quase todos os passos de previsão.

Tabela 3.5: Previsão j passos a frente e erro quadrático médio (MSE) dos modelos WCARR, GCARR e BSCARR para os índices.

Dados	Modelo	j					
		1	2	3	4	5	6
BVSP	WCARR	0.00367	0.00453	0.00511	0.00548	0.00573	0.00588
	GCARR	0.00367	0.00458	0.00517	0.00554	0.00578	0.00593
	BSCARR	0.00291	0.00403	0.00453	0.00475	0.00485	0.00489
HSI	WCARR	0.00367	0.00458	0.00517	0.00554	0.00578	0.00593
	GCARR	0.00367	0.00458	0.00517	0.00554	0.00578	0.00593
	BSCARR	0.00367	0.00453	0.00511	0.00548	0.00573	0.00588
TSEC	WCARR	0.00259	0.00287	0.00337	0.00359	0.00378	0.00378
	GCARR	0.00259	0.00290	0.00318	0.00343	0.00366	0.00386
	BSCARR	0.00227	0.00268	0.00301	0.00328	0.00349	0.00365

Técnica da densidade preditiva: Esse método é usado para comparar a capacidade preditiva dos modelos BSCARR, WCARR e GCARR. Esta metodologia consiste em verificar se uma sequência de densidades preditivas um passo a frente gerados por um modelo CARR, $\{f_i(r_i|\Omega_{i-1})\}$, e uma sequência de PDF's que define o processo de geração dos dados $\{p_i(r_i|\Omega_{i-1})\}$ são tais que

$$\{f_i(r_i|\Omega_{i-1})\} = \{p_i(r_i|\Omega_{i-1})\}, \quad (3.2)$$

ou não. Devido ao fato que $\{p_i(r_i|\Omega_{i-1})\}$ nunca é observado temos que usar o resultado obtido da integral transformada dada por

$$z_t = \int_{-\infty}^{r_t} f_t(u) du. \quad (3.3)$$

Sob a hipótese dada por (3.2) a sequência $\{z_t\}$ e $\{r_t\}$, com respeito a $\{f_i(r_i|\Omega_{i-1})\}$ são variáveis aleatórias independentes e identicamente distribuídas com distribuição $U(0,1)$, para mais detalhes sobre esse fato ver Diebold *et al.* (1998) e Bauwens *et al.* (2004). Como forma de avaliar a independência foi utilizada a análise gráfica da função de autocorrelação (ACF) e função de autocorrelação parcial (PACF), além disso, foi calculado a estatística LB para ratificar a independência dos $\{z_t^2\}$.

A Tabela 3.6 apresenta os p -valores do teste LB para $\{z_t\}$ e $\{z_t^2\}$ considerando os *lags* 4 e 16, em que $\{z_t^2\}$ é usado para testar uma possível correlação não linear. A sequência $\{z_i\}$ é calculada fora da amostra, ou seja, as estimativas dos parâmetros são calculadas na primeira parte da amostra, em seguida, $\{z_t\}$ é obtida com base na outra parte. Os resultados indicam

Tabela 3.6: Resultados do teste LB obtidas para os modelos WCARR, GCARR e BSCARR para os índices BVSP, HSI e TSEC.

Dados	Modelo	p -valor			
		$Q(4)$	$Q^2(4)$	$Q(12)$	$Q^2(12)$
BVSP	WCARR	0.00524	0.00038	0.07593	0.01445
	GCARR	0.00489	0.00018	0.09614	0.01293
	BSCARR	0.17130	0.02687	0.55750	0.24350
HSI	WCARR	0.05554	0.00429	0.42940	0.06468
	GCARR	0.05494	0.00455	0.39510	0.05293
	BSCARR	0.51350	0.10100	0.88910	0.53420
TSEC	WCARR	0.06425	0.01755	0.23330	0.05394
	GCARR	0.06474	0.0189	0.22890	0.05704
	BSCARR	0.09306	0.02645	0.24060	0.10580

que a independência da sequência $\{z_t\}$ pode ser observada para o modelo BSCARR por meio dos gráficos ACF e PACF apresentados na Figura 3.8. Conforme os resultados apresentados na Tabela 3.6, os p -valores do teste LB, confirmam a independência na maioria dos $\{z_t\}$ no modelo BSCARR, o que não acontece para os modelos WCARR e GCARR na maioria dos lags dos índices estudados. A partir desses resultados, podemos concluir que os modelos BSCARR têm um desempenho de previsão melhor que os modelos GCARR e WCARR para os índices BVSP, HSI e TSEC.

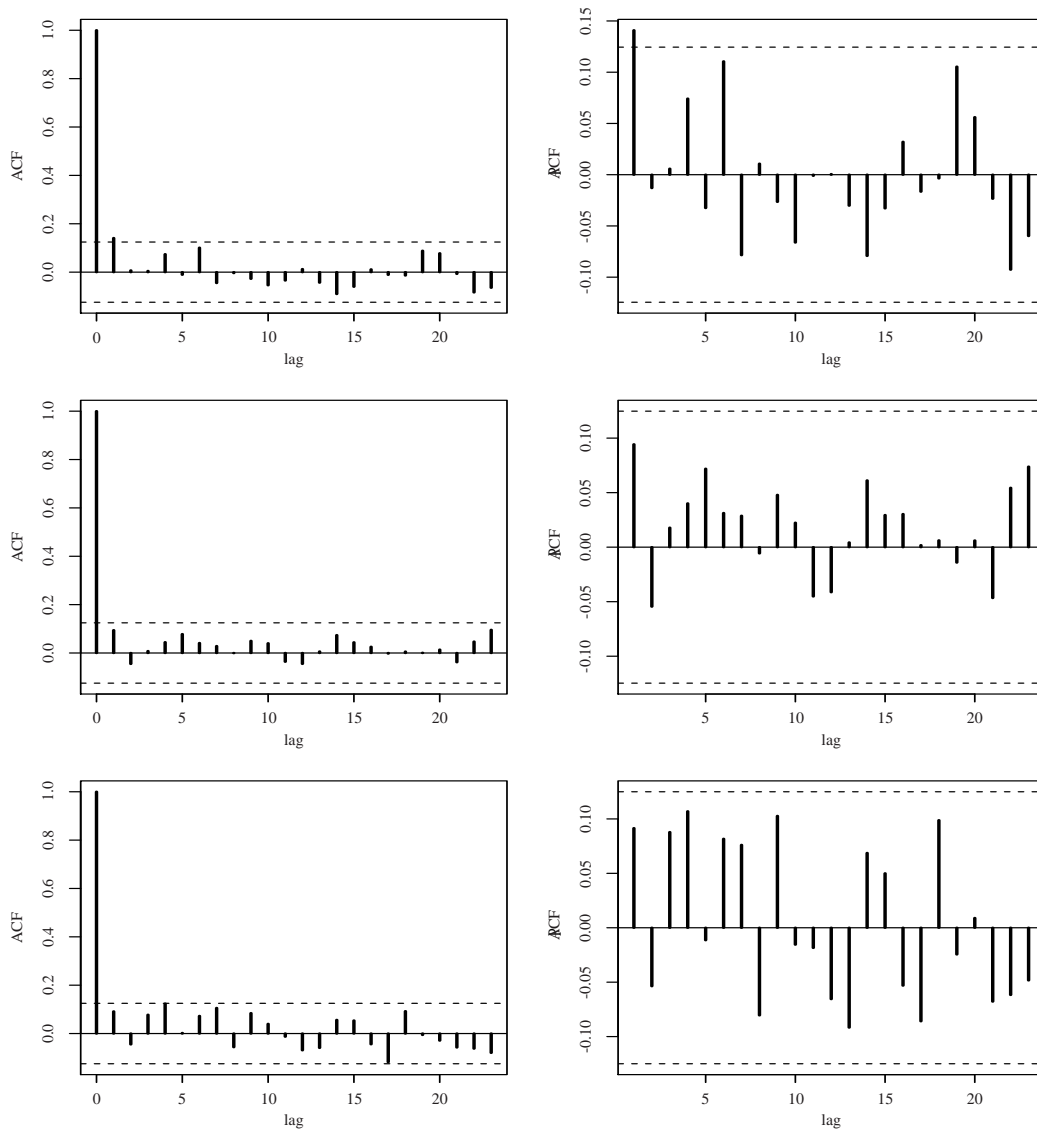


Figura 3.8: Gráficos ACF e PACF da transformada integral de probabilidade z_i para os modelos BS-CARR para os dados da BVSP, HSI e TSEC

Capítulo 4

Considerações Finais

O modelo autorregressivo de variação condicional (CARR) fornece uma estrutura simples e eficiente para modelar a dinâmica da volatilidade baseada na amplitude. Neste trabalho, apresentamos uma nova abordagem para modelo CARR, utilizando a distribuição Birnbaum-Saunders para o termo aleatório do modelo. A metodologia proposta considerou aspectos inferenciais para parâmetros do modelo e sua verificação com base em análise residual e técnicas de diagnóstico e previsão. Realizamos um estudo de simulação de Monte Carlo para avaliar o comportamento dos estimadores dos parâmetros. Os resultados obtidos a partir deste estudo de simulação mostraram bom desempenho da proposta e corroboraram empiricamente algumas propriedades esperadas. Além disso, comparamos os modelos de CARR já existentes na literatura como por exemplo os modelos CARR baseados nas distribuições Weibull e gamma. Como parte de uma pesquisa futura temos: (i) considerar outros procedimentos de estimação, por exemplo o método de momentos modificados; (ii) propor uma abordagem bayesiana via Markov Chain Monte Carlo (MCMC) para a estimação dos parâmetros do modelo; e (iii) podemos ainda avaliar o modelo BSCARR baseado no processo geométrico, como abordado por Chan *et al.* (2012).

Referências Bibliográficas

- Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, **36**(1), 106–108.
- Bauwens, L. & Giot, P. (2000). The logarithmic acid model: an application to the bid-ask quote process of three nyse stocks. *Annales d'Economie et de Statistique*, pages 117–149.
- Bauwens, L., Giot, P., Grammig, J. & Veredas, D. (2004). A comparison of financial duration models via density forecasts. *International Journal of Forecasting*, **20**(4), 589–609.
- Bhatti, C. R. (2009). Intraday trade and quote dynamics: A cox regression analysis. *Mathematics and Computers in Simulation*, **79**(7), 2240–2249.
- Birnbaum, Z. W. & Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied probability*, **6**(2), 319–327.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, **31**(3), 307–327.
- Bollerslev, T., Chou, R. Y. & Kroner, K. F. (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of econometrics*, **52**(1-2), 5–59.
- Chan, J. S., Lam, C., Yu, P., Choy, S. B. & Chen, C. W. (2012). A bayesian conditional autoregressive geometric process model for range data. *Computational Statistics & Data Analysis*, **56**(11), 3006–3019.
- Chiang, M.-H., Chou, R. Y. & Wang, L.-M. (2016). Outlier detection in the lognormal logarithmic conditional autoregressive range model. *Oxford Bulletin of Economics and Statistics*, **78**(1), 126–144.
- Chou, R. Y.-T. (2005). Forecasting financial volatilities with extreme values: the conditional autoregressive range (carr) model. *Journal of Money, Credit, and Banking*, **37**(3), 561–582.

- Cook, R. D. (1987). Influence assessment. *Journal of Applied Statistics*, **14**(2), 117–131.
- Cox, D. R. & Hinkley, D. V. (1978). Problems and solutions in theoretical statistics.
- Desmond, A. (1985). Stochastic models of failure in random environments. *Canadian Journal of Statistics*, **13**(3), 171–183.
- Diebold, F. X., Gunther, T. A. & Tay, A. S. (1998). Evaluating density forecasts with applications to financial risk management. v inter# national economic review, vol. 39, no. 4. In *Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance*, page 863.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.
- Engle, R. F. & Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, pages 1127–1162.
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, **6**(2), 327–343.
- Hubert, M. & Van der Veeken, S. (2008). Outlier detection for skewed data. *Journal of Chemometrics: A Journal of the Chemometrics Society*, **22**(3-4), 235–246.
- Jarque, C. M. & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review/Revue Internationale de Statistique*, pages 163–172.
- Kotz, S., Leiva, V. & Sanhueza, A. (2010). Two new mixture models related to the inverse gaussian distribution. *Methodology and Computing in Applied Probability*, **12**(1), 199–212.
- Leiva, V., Santos-Neto, M., Cysneiros, F. J. A. & Barros, M. (2014a). Birnbaum–saunders statistical modelling: a new approach. *Statistical Modelling*, **14**(1), 21–48.
- Leiva, V., Saulo, H., Leão, J. & Marchant, C. (2014b). A family of autoregressive conditional duration models applied to financial data. *Computational Statistics & Data Analysis*, **79**, 175–191.
- Leiva, V., Marchant, C., Ruggeri, F. & Saulo, H. (2015). A criterion for environmental assessment using birnbaum–saunders attribute control charts. *Environmetrics*, **26**(7), 463–476.

- Leiva, V., Ruggeri, F., Saulo, H. & Vivanco, J. F. (2017). A methodology based on the birnbaum–saunders distribution for reliability analysis applied to nano-materials. *Reliability Engineering & System Safety*, **157**, 192–201.
- Lesaffre, E. & Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, pages 570–582.
- Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica*, *59*, pages 1279–1313.
- Mandelbrot, B. B. (1971). When can price be arbitrated efficiently? a limit to the validity of the random walk and martingale models. *The Review of Economics and Statistics*, pages 225–236.
- Mittelhammer, R. C., Judge, G. G. & Miller, D. J. (2000). *Econometric Foundations Pack with CD-ROM*, volume 1. Cambridge University Press.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370.
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of business*, pages 61–65.
- Saulo, H., Leiva, V., Ziegelmann, F. A. & Marchant, C. (2013). A nonparametric method for estimating asymmetric densities based on skewed birnbaum–saunders distributions applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, **27**(6), 1479–1491.
- Taylor, S. J. (2008). *Modelling financial time series*. world scientific.
- Tsay, R. S. (2009). Autoregressive conditional duration models. In *Palgrave Handbook of Econometrics*, pages 1004–1024. Springer.
- Xie, H. & Wu, X. (2017). A conditional autoregressive range model with gamma distribution for financial volatility modelling. *Economic Modelling*, **64**, 349–356.
- Zevallos, M., Santos, B. & Hotta, L. K. (2012). A note on influence diagnostics in ar (1) time series models. *Journal of Statistical Planning and Inference*, **142**(11), 2999–3007.