

Andressa da Silva Contarato

**Modelagem com o uso de séries temporais
utilizando distribuições hiperbólicas
generalizadas T-Student assimétrica no
mercado financeiro brasileiro**

Niterói - RJ, Brasil

8 de Julho de 2015

Andressa da Silva Contarato

**Modelagem com o uso de séries
temporais utilizando distribuições
hiperbólicas generalizadas T-Student
assimétrica no mercado financeiro
brasileiro**

Trabalho de Conclusão de Curso

Monografia apresentada para obtenção do grau de Bacharel em
Estatística pela Universidade Federal Fluminense.

Orientador: Prof. Dr. Marco Aurélio Sanfins

Coorientador: Prof. Dr. Valentin Sisko

Andressa da Silva Contarato

**Modelagem com o uso de séries temporais
utilizando distribuições hiperbólicas
generalizadas T-Student assimétrica no
mercado financeiro brasileiro**

Monografia de Projeto Final de Graduação sob o título “*Modelagem com o uso de séries temporais utilizando distribuições hiperbólicas generalizadas T-Student assimétrica no mercado financeiro brasileiro*”, defendida por Andressa da Silva Contarato e aprovada em 8 de Julho de 2015, na cidade de Niterói, no Estado do Rio de Janeiro, pela banca examinadora constituída pelos professores:

Prof. Dr. Marco Aurélio Sanfins
Departamento de Estatística – UFF

Prof. Dr. Valentin Sisko
Departamento de Estatística – UFF

Profa. Dra. Márcia de Carvalho
Departamento de Estatística – UFF

Niterói, 8 de Julho de 2015

Contarato, Andressa da Silva

Modelagem com o uso de séries temporais utilizando distribuições hiperbólicas generalizadas t-Student assimétrica no mercado financeiro brasileiro / Andressa da Silva Contarato; Marco Aurélio dos Santos Sanfins, orientador; Valentin Sisko, coorientador. Niterói, 2015.
67 f. : il.

Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Federal Fluminense, Instituto de Matemática e Estatística, Niterói, 2015.

1. Distribuição hiperbólica generalizada assimétrica t-Student (GHsT). 2. Séries temporais. 3. Previsão. 4. Value at risk (VaR). I. Sanfins, Marco Aurélio dos Santos, orientador. II. Sisko, Valentin, coorientador. III. Universidade Federal Fluminense. Instituto de Matemática e Estatística. IV. Título.

CDD -

Resumo

O escopo desta monografia foi discutir o uso da distribuição hiperbólica generalizada assimétrica T-student (GHsT) como modelo para os log retornos de três séries de dados do mercado financeiro Brasileiro. As séries em questão são o Índice Bovespa e os dois ativos das duas maiores empresas pertencentes ao Índice Bovespa: *PETRA4* (Petrobrás) e *VALE5* (Vale), para o período de 02 de Janeiro de 2004 a 12 de Junho de 2015. Os dados são formados pelos valores de fechamento diário. Utilizando o *software R* foram feitas análises de séries temporais para um melhor ajustamento dos modelos aos dados. A modelagem dos dados se deu por modelos do tipo *ARMA* e *GARCH*. Após este procedimento, foi feito o cálculo da volatilidade e da previsão dos modelos. Os resultados obtidos permitiram concluir que o melhor ajustamento foi feito pela hiperbólica generalizada assimétrica T-student comparando com a distribuição gaussiana. Ao final destas análises foi feito o cálculo do *Value at risk* como forma de prever a perda máxima.

Palavras-chaves: Distribuição hiperbólica generalizada assimétrica T-student (GHsT), Séries Temporais, Previsão, *Value at risk* (VaR).

Abstract

The scope of this paper was to discuss the use of asymmetric generalized hyperbolic distribution T-student (GHsT) as a model for the log returns of three Brazilian financial market data series. The series in question are the Bovespa Index and the two assets of the two largest companies belonging to the Bovespa Index: PETR4 and VALE5. The data is formed by the daily closing values. Using the R software time series analyzes were made for a better adjustment model to the data. The modeling of the data was by models of type ARMA and GARCH. After this procedure, it was made the calculation of volatility and the prediction models. The results showed that the best adjustment was made by asymmetric generalized hyperbolic T-student compared to the Gaussian distribution. At the end of these analyzes was made the calculation of Value at risk as a way of predicting the maximum loss.

Keywords: Generalized hyperbolic skew T-student distribution, GHsT, Time Series Forecasting Volatility, *Value at risk* (VaR).

“Alguns homens vêem as coisas como são, e dizem ‘Por quê?’ Eu sonho com as coisas que nunca foram e digo ‘Por que não?’”

(George Bernard Shaw)

Agradecimentos

A parte mais difícil de escrever este trabalho foi esta parte de agradecimentos. Não que seja dificultoso dizer quem me ajudou e incentivou, mas a forma de dizer e além disso ... expressar em tão poucas linhas meus sinceros sentimentos que tenho por cada um que cito aqui. Bom, vamos começar com Deus e Maria, mais que uma obrigação, é um prazer ser cristã. Agradeço as pessoas que me transformaram na pessoa que sou hoje: Marli, João e Vítor Ricardo. Mais do que minha família eles são a minha vida e com certeza não sou nada sem eles. Seguindo, a toda a minha família, mas em especial minhas avós (que por vezes me faziam estudar) e minha tia Glória (a pessoa mais paciente e amiga). E aos meus padrinhos que são meus exemplos de vida.

Agora, com relação a minha formação acadêmica, queria agradecer as minhas “tias” da escola: Patrícia (ensino fundamental), Hiraíma, Márcia, Bárbara, Patrícia (ensino médio). E aos meus professores de faculdade: Adrian, Márcia, Kelly, Jéssica, Jhony, Patrícia, Ludmilla, Mariana, Núbia, Francisco.

Mas em especial queria agradecer muito aos meus dois professores que tanto acreditaram e perseveraram comigo: Marco Aurélio e Valentin. Marco Aurélio foi um professor fundamental em minha vida, lembro que desde aquele minicurso em que a Daiane e eu ficamos vislumbradas com o mercado financeiro, senti que já havia escolhido meu orientador. Então, professor Marco Aurélio, obrigada por ser um ótimo professor e um amigo que eu sempre pude confiar mesmo tendo muita timidez. E, Valentin ... a pessoa mais inteligente, legal, surpreendente e compreensiva. Obrigada por me aturar, ajudar, trabalhar comigo e discutir filmes rs ...

Queria agora fazer um parágrafo falando sobre meus amigos, todos que diretamente e indiretamente contribuíram com minha formação. A começar por Jéssica Gonçalves, a pessoa que mais me apoiou nesta trajetória; Daiane Grugel e Sérgio Felipe, amigos e irmãos que sempre terei comigo; Larissa Eletério, Júlio César, Camila Domingos que sempre me aguentaram. E, meus amigos do mestrado: Daniela Albuquerque e Marcos Lima. Por fim, mas não menos importante quero agradecer muito a uma pessoa que me apoiou bastante nesta reta final da graduação Pedro Ivo Rodrigues.

Sumário

Lista de Figuras

Lista de Tabelas

1	Introdução	p. 14
1.1	Objetivos	p. 15
1.2	Revisão bibliográfica	p. 16
1.3	Organização	p. 17
2	Métodos	p. 19
2.1	Introdução a Séries Temporais	p. 19
2.2	<i>Processo Estocástico</i>	p. 20
2.3	Estacionariedade	p. 22
2.4	Retornos	p. 26
2.5	Tipos de modelos	p. 28
2.5.1	Modelos Auto-regressivos	p. 31
2.5.2	Médias Móveis	p. 32
2.5.3	Modelos ARMA	p. 33
2.5.3.1	Função de autocorrelação parcial	p. 34
2.5.4	Modelos não-estacionários	p. 35
2.5.4.1	ARIMA	p. 36
2.5.4.2	Identificação dos modelos	p. 36
2.5.5	Modelos não lineares	p. 39

2.5.5.1	Modelos ARCH	p. 39
2.5.5.2	Modelos GARCH	p. 41
2.5.6	Estimação dos modelos	p. 42
2.5.6.1	Estimação dos modelos ARIMA	p. 42
2.5.6.2	Estimação dos modelos GARCH	p. 44
2.5.6.3	Estimação dos modelos ARMA - GARCH	p. 45
2.5.7	Diagnóstico dos Modelos	p. 45
2.5.7.1	Testes de adequação do modelo - Teste de Box-Pierce .	p. 46
2.5.7.2	Testes de adequação do modelo - Teste de BROCK, DE- CHERT e SCHEINKMAN (BDS)	p. 46
2.5.7.3	Testes de adequação do modelo - Teste de Jarque-Bera	p. 47
2.6	Modelos de previsão	p. 48
2.6.1	Previsão para modelos lineares	p. 48
2.6.2	Previsão para modelos não lineares	p. 49
2.6.2.1	Previsões de volatilidade	p. 49
2.7	Tipos de distribuições	p. 50
2.7.1	Distribuição Normal	p. 50
2.7.2	Distribuições hiperbólicas Generalizadas Assimétricas	p. 51
2.7.2.1	Distribuição Hiperbólica Generalizada Assimétrica T- student (GHsT)	p. 52
2.7.3	<i>Value at risk</i>	p. 54
2.7.4	Teste de <i>Kupiec</i>	p. 56
3	Resultados	p. 58
3.1	Séries utilizadas na modelagem	p. 58
3.2	Resultados preliminares	p. 59
3.3	Modelagem e diagnóstico	p. 59

3.4	Previsão	p. 68
4	Conclusão	p. 71
5	Trabalhos futuros	p. 73
	Referências	p. 74
6	Anexo A - Cálculo dos parâmetros da GHsT	p. 76
7	Anexo B - Função de Bessel	p. 79

Lista de Figuras

1	Gráfico das séries PETR4, VALE5 e Índice Bovespa, considerando os valores em reais.	p. 20
2	Gráficos das séries Petrobrás (PETR4), Vale (VALE5) e Índice Bovespa (IBOV11) da primeira diferença do logaritmo neperiano.	p. 26
3	Autocorrelações e autocorrelações parciais amostrais para os dados de retorno das séries PETR4, VALE5 e Índice Bovespa.	p. 38
4	Gráfico das séries PETR4, VALE5 e Índice Bovespa, considerando os valores em reais.	p. 59
5	Gráficos das séries Petrobrás (PETR4), Vale (VALE5) e Índice Bovespa (IBOV11) da primeira diferença do logaritmo neperiano.	p. 60
6	Autocorrelações e autocorrelações parciais amostrais para os dados de retorno das séries PETR4, VALE5 e Índice Bovespa.	p. 61
7	Correlograma dos resíduos dos modelos para as séries de dados dos retornos das PETR4, VALE5 e Índice Bovespa usando o primeiro método de modelagem.	p. 63
8	Correlograma dos resíduos ao quadrado dos modelos para as séries de dados dos retornos das séries PETR4, VALE5 e Índice Bovespa considerando o primeiro método de modelagem.	p. 65
9	Gráfico dos valores de VaR Loss e o VaR Gain para o Índice Bovespa considerando a GHsT.	p. 70
10	Gráfico dos valores de VaR Loss e o VaR Gain para o Índice Bovespa considerando a Normal.	p. 70
11	Gráfico da função de Bessel modificada $K_\nu(z)$ de valor real e positiva, com $z > 0$ e $\nu > -1$	p. 79

Lista de Tabelas

1	Tabela de identificação dos modelos de séries temporais.	p. 35
2	PETR4 - Tabela dos valores dos coeficientes estimados para os dois métodos de estimação dos modelos, considerando os resíduos seguindo uma distribuição Normal.	p. 61
3	VALE5 - Tabela dos valores dos coeficientes estimados para os dois métodos de estimação dos modelos, considerando os resíduos seguindo uma distribuição Normal.	p. 61
4	Índice Bovespa - Tabela dos valores dos coeficientes estimados pelos dois métodos de estimação dos modelos, considerando os resíduos seguindo uma distribuição Normal.	p. 62
5	Ajuste do modelo, com os dois tipos de modelagem aplicadas nas três séries de dados analisadas, considerando os resíduos seguindo uma distribuição Normal.	p. 62
6	Primeiro método - Tabela dos p-valores do teste de Box-Pierce para os modelos estimados dos dados de retorno das séries PETR4, VALE5 e Índice Bovespa.	p. 63
7	Segundo método - Tabela dos p-valores do teste de Box-Pierce para os modelos estimados dos dados dos retornos para as séries PETR4, VALE5 e Índice Bovespa.	p. 64
8	Ajuste de modelos não lineares, com os dois tipos de métodos aplicados nas três séries de dados analisadas, considerando os resíduos seguindo distribuição Normal.	p. 64
9	Segundo método - Tabela dos valores dos coeficientes estimados dos modelos não lineares dos dados de retornos das séries PETR4, VALE5 e Índice Bovespa.	p. 65

10	Primeiro método - Tabela dos p-valores do teste de auto-correlação dos modelos não lineares propostos para os dados de retorno das séries PETR4, VALE5 e Índice Bovespa.	p. 66
11	Segundo método - Tabela dos p-valores do teste de auto-correlação dos modelos não lineares propostos dos dados de retorno das séries PETR4, VALE5 e Índice Bovespa.	p. 66
12	Teste de <i>Jarque-Bera</i> para o primeiro método dos dados dos retornos das séries analisadas.	p. 68
13	Teste de <i>Jarque-Bera</i> para o segundo método dos dados dos retornos das séries analisadas.	p. 68
14	Coefficientes da GHsT para os dados de retornos das séries PETR4, VALE5 e Índice Bovespa, para um dia e considerando o segundo método de modelagem.	p. 68
15	Previsão usando a GHsT para os dados de retornos das séries PETR4, VALE5 e Índice Bovespa, para um dia e considerando o segundo método de modelagem.	p. 68
16	P-valores do Teste de Kupiec considerando os modelos com os resíduos seguindo distribuição Normal e seguindo distribuição GHsT com o <i>VaR gain</i>	p. 69
17	P-valores do Teste de Kupiec considerando os modelos com os resíduos seguindo distribuição Normal e seguindo distribuição GHsT com o <i>VaR loss</i>	p. 69

1 Introdução

Cada vez mais diversas ciências vêm se interagindo com o sistema financeiro. E a estatística, entre elas, vem ganhando relevância principalmente na área de risco. De modo geral, a aplicação de métodos estatísticos tem uma mistura de ciência, tecnologia e lógica. O mercado financeiro tem o objetivo de monitorar o fluxo de recursos entre os agentes econômicos, estudá-lo é ter o conhecimento de como a alocação de recursos na economia está sendo distribuída. O uso da estatística neste tipo de problemática faz com que se aprimore o monitoramento de altas e baixas num sistema financeiro como abordado na Comissão de Valores Imobiliários (2004) [1].

Um dos diversos ramos de investimentos, designado como ação pode ser denominada como sendo a menor parcela de um capital social tanto de companhias quanto de sociedades. Ao se investir em alguma ação, três fatores são importantes: a rentabilidade, a liquidez e o risco. A rentabilidade é a renda do ativo naquele período de tempo já a liquidez é a facilidade de comprar e vender certo ativo e o grau de risco é uma incerteza ao se investir em um ativo financeiro. A volta de uma aplicação (seja positiva ou negativa) é chamada de retorno. Então, note que o controle (ou pelo menos uma maior sensibilidade) do ativo em um tempo futuro, como também abordado na Comissão de Valores Imobiliários (2004) [1].

Não é uma tarefa fácil prever o comportamento de ativos financeiros para horizontes futuros. É sabido que os dados de cotações de ações são ordenadas pelo tempo, diante disso várias técnicas de séries temporais podem ser empregadas. Mesmo com tantas facilidades já existentes na modelagem de retornos, devido a característica de ser livre de escala, não é fácil encontrar uma distribuição que melhor se adequa aos dados destes retornos. Uma distribuição que está sendo usada em dados financeiros é a distribuição hiperbólica (GH). Estudada pela primeira vez por Prause, ver Prause (1999) [2] e Teixeira (2006) [3], o uso desta possibilita comparar distribuições clássicas (como a normal) e optar pelo uso das distribuições hiperbólicas por suas propriedades como abordado em Silva (1999) [4].

Restringindo ainda mais, uma classe desta família de distribuições hiperbólicas é chamada de distribuição hiperbólica assimétrica, usada pela primeira vez por Hansen (1994), ver Aas et al. (2006) [5], que foi o primeiro a propor o uso das distribuições hiperbólicas assimétricas na modelagem de retornos. Ao passo que nos últimos anos vêm sendo usadas com as mesmas finalidades que as demais distribuições, contudo há um maior destaque no uso destas distribuições em dados financeiros pelo fato da assimetria verificada nos dados empiricamente.

O diferencial deste trabalho está na percepção do comportamento de um ativo no mercado financeiro. Este estudo se faz muito importante para que os *players* (pessoas que investem no mercado financeiro baseados nos diversos estudos sobre este) tenham um maior conhecimento dos dados das diversas séries de preços de ativos e saibam quando e como investir de uma forma mais segura, correndo menos riscos. Riscos sempre haverão neste mercado porque a probabilidade de se acertar uma observação futura é zero, por se tratar de dados contínuos. Entretanto, baseados nas análises, os investidores poderão ter um breve conhecimento se o preço irá cair ou não no próximo dia.

1.1 **Objetivos**

Uma série temporal pode apresentar as seguintes características: tendência, sazonalidade e ciclo. Sempre antes de se fazer análises é importante o tratamento do banco de dados antes da modelagem. É sabido que a estrutura da maioria das séries financeiras consistem em possuir tendência. Além da tendência, no momento que se faz a modelagem, a determinação da distribuição de probabilidade que os resíduos seguirão determinará a capacidade de precisão do modelo.

Com esse intuito, estudar diversos modelos e conhecer a distribuição de probabilidade dos erros faz com que possam ser calculadas previsões e controlar a volatilidade. Assim, o objetivo geral consiste em modelar séries temporais utilizando uma distribuição ligada a família de distribuições hiperbólicas, chamada de distribuição hiperbólica generalizada assimétrica T-student (GHsT), nos erros afim de determinar modelos com melhor previsão de preço e de volatilidade. Para que perguntas como: “O que acontecerá com preço de um tipo de ação?” e “Qual o seu risco de investir numa determinada ação?” possam ser sanadas.

Para isso, serão analisadas três séries de dados muito comuns no mercado financeiro: *Índice Bovespa*, *PETR4* e *VALE5*. Assim, serão aplicadas modelagens nos dados com o

uso de técnicas de séries temporais afim de descobrir qual o melhor modelo a ser usado em cada ativo. A comparação também será na utilização de distribuição *Normal* e *distribuição hiperbólica generalizada assimétrica T-student (GHsT)* para os erros destes ativos afim de comparar qual distribuição tem uma melhor adequação da série de retornos.

1.2 Revisão bibliográfica

As distribuições hiperbólicas foram propostas para o estudo de movimento de areia que pode ser provocada pelos ventos ou até mesmo pelas águas. A interseção entre a física do movimento da areia e os processos estocásticos são mais evidentes quando vemos o comportamento dos grãos, estes se comportam de forma hiperbólica. Os primeiros a introduzir a distribuição hiperbólica foram Barndorff-Nielsen em 1977, ver Silva (1999) [4]. E, um dos primeiros a propor esta ideia no mercado financeiro foram Eberlein et al. (1995) [6]. Futuramente, diversos outros trabalhos foram desenvolvidos com a aplicação desta família de distribuições no mercado financeiro.

No artigo de Silva (1999) [4], o autor faz uso de ativos financeiros brasileiros e nota a eficiência da modelagem da distribuição hiperbólica em comparação com a gaussiana. Fazendo uso de diversos testes para a verificação da normalidade dos dados, como também, usufruindo de ferramentas gráficas para uma melhor escolha do modelo. Fazendo o teste da χ^2 , rejeitou-se a suposição de normalidade nos resíduos além do teste de Jarque-Bera e de estatísticas descritivas e assim fez o uso da distribuição hiperbólica.

Em seu artigo, Fajardo et al. (2004) [8] usa diversas séries financeiras do mercado brasileiro fazendo o uso da distribuição hiperbólica assimétrica em comparação com a Normal inversa gaussiana. Eles desenvolveram no MatLab um algoritmo para estimação dos modelos, fizeram análises gráficas com os dados de retorno de cada série juntamente com as densidades das duas distribuições a serem comparadas. Usando o modelo GARCH para a modelagem com as distribuições Normal Inversa Gaussiana (NIG) e GHsT, o autor chega a conclusão de que quando se compara a qualidade de ajuste da distribuição hiperbólica generalizada assimétrica (GHs) com a NIG, a melhor escolha se faz no uso da distribuição hiperbólica. Também, ao calcular o VaR, a conclusão é a mesma. O interesse se fez em observar que a utilização de subclasses da distribuição hiperbólica generalizada tem o processo de estimação de parâmetros mais eficiente.

No artigo de Eberlein et al. (1995) [6], os autores utilizaram a distribuição hiperbólica para os dados vindos de *German stocks* e *American indexes*. Antes de fazer o uso da distri-

buição GHsT, os autores fizeram vários gráficos mostrando o comportamento dos retornos destas séries juntamente com a densidade de uma GHsT. Assim, usaram as distribuições hiperbólicas para os retornos nos dados financeiros e notaram que esta distribuição era realmente uma excelente candidata na modelagem como também providenciava um modelo mais realístico devido as análises gráficas, onde se construiu um gráfico com as densidades da Normal e da GHsT juntamente com os dados de retorno de cada série. Logo, a GHsT apresentou valores melhores no cálculo da opção de preços.

No artigo de Teixeira (2006) [3] o autor aborda o uso da distribuição hiperbólica assimétrica T-student como sendo a melhor escolha na modelagem dos retornos de dados financeiros do mercado português. Fazendo o uso de análises de comparação como o Qq plot e os gráficos dos log-retornos com as distribuições a serem usadas. Ao final, o autor chega a conclusão de que o uso da distribuição hiperbólica generalizada é a mais adequada para os resíduos do modelo. Um artigo análogo a este foi de Kim et al. (2007) [9] que faz uso de ferramentas gráficas e critérios de decisões ao modelar a série de retornos de ativos americanos, tendo como eleita a distribuição hiperbólica assimétrica T-student.

São também encontrados estudos mais aprofundados que utilizam as distribuições normal, inversa gama e hiperbólicas em geral comparando com a distribuição hiperbólica assimétrica T-student, como no artigo de Aas et al. (2006) [5]. Os autores fizeram análises gráficas como o Qq plot seguido de estimação dos parâmetros e fazendo uso do modelo GARCH(1,1), chegaram a conclusão da importância da aplicação da distribuição hiperbólica generalizada assimétrica T-student (GHsT) no uso de diversos ativos financeiros. E, além disso descreve suas propriedades como singulares no que diz respeito a caudas mais pesadas, como também, a propriedade de decaimento distinto das caudas.

1.3 Organização

Este trabalho está organizado em Introdução, Metodologia, Resultados e Conclusão. No que diz respeito a metodologia, a seção 2.1 apresenta noções de Séries temporais, a seção 2.2 aborda uma breve introdução a Processos estocásticos e a relação entre este e séries temporais, na seção 2.3 noções sobre estacionariedade e a aplicação desta nos dados da série original, na seção 2.4 é apresentado a definição de retorno, na seção 2.5 é discutido os tipo de modelos lineares e não lineares com suas respectivas características e método de identificação, na subseção 2.5.7 foi apresentado os testes com quais foram analisados a adequação dos modelos e na seção 2.8 foi apresentado as duas distribuições com as quais

foram adequados os resíduos dos modelos em questão e nesta mesma seção, foi abordado os conceitos do VaR como também do teste de Kupiec.

2 Métodos

2.1 Introdução a Séries Temporais

O monitoramento das ações num mercado financeiro pode ser diário, semanal, mensal, anual, *etc*, saber qual será o comportamento de um ativo amanhã ou após um determinado ciclo é de grande relevância para investidores. Uma definição informal de Séries Temporais é um conjunto de observações z_t , cada uma observação sendo gravada em um momento t específico, onde $t \in \mathbb{R}$. Em séries temporais é feito o uso de diferentes metodologias para analisar o que se quer. Técnicas são desenvolvidas ao longo do tempo, tais como as descritas em Steffen (2010) [10] e abordadas abaixo.

- Na década de 30 e 40 uso de métodos compostos relacionados a tendência, sazonalidade e aleatoriedade, isto é, o modelo a ser usado era: $Z_t = f(T_t, S_t, A_t)$ onde T_t é a tendência da série, S_t é a sazonalidade e A_t é o ciclo da série;
- Em 1960, técnicas de amortecimento exponencial;
- Já nos anos 70, houve a criação do Método Box e Jenkins, que será abordado mais adiante na determinação de modelos;
- E, em 1980, foi feito o uso de técnicas de inferência Bayesiana para uso em séries temporais.
- E ainda na década de 80, foi introduzido por Engle (1982) [11] e Bollerslev (1986) [12] os modelos não lineares.

Assim é fácil perceber que o uso de técnicas de séries temporais se dá pela facilidade de monitoramento e conhecimento dos dados de determinada pesquisa ao longo do tempo. Para a possibilidade de aplicação da técnica de séries temporais, são exigidas certas restrições para os dados como: a ordem temporal das observações, a série ser livre

de tendências, livre de sazonalidade, *etc*, a fim de permitir fazer modelagens e usar previsões nos dados. No entanto, há alguns problemas que acontecem tais como: observações perdidas e dados discrepantes.

Exemplo 2.1.1 *Aqui é apresentado o gráfico das três séries que foram estudadas. Pode-se ver tendências nos três gráficos, assim como, picos de alta e baixa devido aos acontecimentos históricos envolvidos nestas séries de dados como queda do PIB em 2004 e a crise de 2008.*

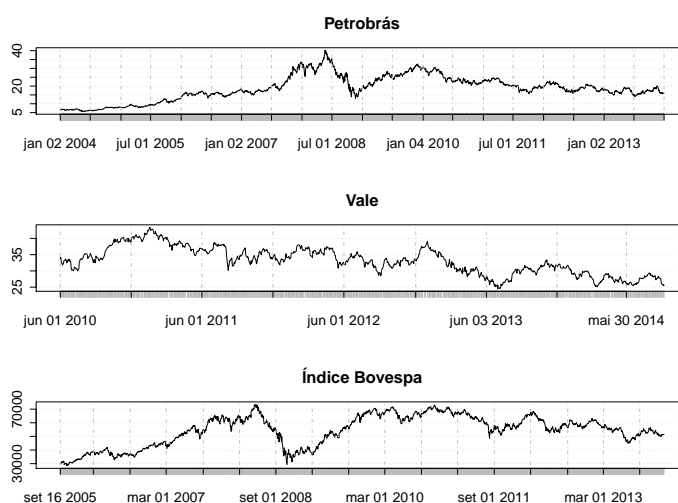


Figura 1: Gráfico das séries PETR4, VALE5 e Índice Bovespa, considerando os valores em reais.

2.2 Processo Estocástico

A ação de avaliar se subirá ou cairá e quanto irá cair ou subir um determinado ativo é parte crucial para o investimento. Um processo estocástico é muito importante na parte de previsão de ativos. No tocante à definição de séries temporais, os processos estocásticos podem ser definidos como um conjunto de observações ordenadas pelo tempo. Logo, um processo estocástico é uma família de variáveis aleatórias (*v.a.*'s) definidas no mesmo espaço de probabilidade, ou seja (Ω, \mathbb{A}, P) , onde Ω representa o espaço amostral, \mathbb{A} é definido como σ -álgebra de eventos e P é a função de probabilidade. Sabendo que \mathbb{T} é definido como o conjunto dos inteiros \mathbb{Z} ou dos reais \mathbb{R} . Assim, para $t \in \mathbb{T}$, Z_t é uma variável aleatória real. Com $\{Z_t\}$ definida sobre o espaço amostral Ω . Os resultados apresentados abaixo são extraídos do livro de Morettin et al. (2006) [13].

Definição 2.2.1 *Um processo estocástico é uma família $Z = \{Z_t, t \in \mathbb{T}\}$, onde, para cada $t \in \mathbb{T}$, Z_t é uma variável aleatória.*

Pode-se ver que $\{Z_t\}$ é uma função de dois argumentos, $Z_t(\omega)$, com $t \in \mathbb{T}$ e $\omega \in \Omega$. Uma observação a ser feita a respeito da distribuição de probabilidades $Z_t(\omega)$ é que esta fdp (função de densidade de probabilidade) no instante t_1 pode ser diferente da fdp no instante t_2 , para dois instantes de tempo t_1 e t_2 quaisquer. Mesmo assim, é mais usual a situação em que a fdp $Z(t, \omega)$ é a mesma, $\forall t \in \mathbb{T}$. Contrariamente, fixando um $\omega \in \Omega$, obteremos uma função de t que pode ser denominada como uma *realização* ou *trajetória* do processo, ou ainda, uma *série temporal*.

Note que, o termo série temporal é usado tanto para significar dados (i.e., uma realização de um processo) quanto para significar o próprio processo.

Observe que se o conjunto \mathbb{T} for finito ou, ao menos, enumerável como $\mathbb{T} = 1, 2, 3, \dots$ ou $\mathbb{T} = \mathbb{Z}$, o processo é chamado de *parâmetro discreto*. Por um outro lado, se \mathbb{T} for um intervalo de \mathbb{R} , obteremos um processo com *parâmetro contínuo*.

Sejam t_1, t_2, \dots, t_n elementos quaisquer de \mathbb{T} e consideremos

$$\begin{aligned} F(z_1, z_2, \dots, z_n; t_1, t_2, \dots, t_n) \\ = P(Z_{t_1} \leq z_1, Z_{t_2} \leq z_2, \dots, Z_{t_n} \leq z_n). \end{aligned} \quad (2.1)$$

Então o processo estocástico $Z = \{Z_t, t \in \mathbb{T}\}$ estará especificado se conhecermos as *distribuições finito-dimensionais* (2.1), para todo $n \geq 1$. Isto significa que, para $n = 1$, nós conhecemos as distribuições unidimensionais da v.a. $Z_{t_1}, t_1 \in \mathbb{T}$. Já para $n = 2$, também conhecemos as distribuições bidimensionais da v.a. $(Z_{t_1}, Z_{t_2}), t_1, t_2 \in \mathbb{T}$ e assim por diante. As funções de distribuição (2.1) devem obedecer as seguintes condições:

- Condição de simetria: para qualquer permutação j_1, j_2, \dots, j_n dos índices $1, 2, \dots, n$, temos: $F(z_{j_1}, z_{j_2}, \dots, z_{j_n}; t_{j_1}, t_{j_2}, \dots, t_{j_n}) = F(z_1, z_2, \dots, z_n; t_1, t_2, \dots, t_n)$;
- Condição de compatibilidade. Consideremos $m < n$, temos a seguinte relação:

$$\begin{aligned} F(z_1, z_2, \dots, z_m, +\infty, \dots, +\infty; t_1, t_2, \dots, t_m, t_{m+1}, \dots, t_n) \\ = F(z_1, z_2, \dots, z_m; t_1, t_2, \dots, t_m). \end{aligned} \quad (2.2)$$

Membro esquerdo de (2.2) deve ser entendido como:

$$\lim F(z_1, z_2, \dots, z_m; t_1, t_2, \dots, t_n),$$

para $z_{m+1} \rightarrow +\infty, z_{m+2} \rightarrow +\infty, \dots, z_n \rightarrow +\infty$.

2.3 Estacionariedade

Para uma melhor utilização dos processos de séries temporais, certas suposições tem que ser impostas afim de simplificar o modo de uso destes. Os tipos de suposições são muitos, como por exemplo: processos estacionários, processos gaussianos, processos markovianos, *etc.* Uma série não estacionária faz com que as autocorrelações decaiam para zero muito lentamente quando é considerado *lags* altas, para mais informações ver Ehlers (2009) [14]. Há duas formas de estacionariedade que são denominadas como: estrita (ou forte) e fraca (ou ampla) abordadas no livro de Brockwell et al. (2001) [15].

Definição 2.3.1 *Um processo estocástico $Z = \{Z_t, t \in \mathbb{Z}\}$ diz-se estritamente estacionária se e somente se:*

- *É definida pela condição que (Z_1, \dots, Z_n) ;*
- *$(Z_{1+h}, \dots, Z_{n+h})$ tem a mesma distribuição conjunta para todos os inteiros h e $n > 0$.*

Para uma melhor escolha de um modelo de séries temporais, restrições são feitas a momentos de primeira e segunda ordens. Surge, então, uma outra

Definição 2.3.2 *Dado Z_t uma série temporal com $E(Z_t^2) < \infty$ e $t \in \mathbb{Z}$, então:*

- *A função média de Z_t é $\mu_Z(t) = E(Z_t)$;*
- *A função covariância de Z_t é*

$$\gamma_Z(r, s) = \text{Cov}(Z(r), Z(s)) = E((Z(r) - \mu_Z(r))(Z(s) - \mu_Z(s))),$$
para todo r e s inteiros.

Definição 2.3.3 *Um processo estocástico $Z = \{Z_t, t \in \mathbb{Z}\}$ diz-se fracamente estacionário se e somente se:*

- *$\mu_Z(t)$ é independente de t ;*
- *$\gamma_Z(t+h, t)$ é independente de t , para todo $h \in \mathbb{Z}$.*

Onde, μ_Z e γ_Z são a função média e a função de covariância definidas abaixo.

Quando usamos o termo função covariância para um processo estacionário Z , vamos usar a função γ_Z de uma variável definida por

$$\gamma_Z(h) := \gamma_Z(h, 0) = \gamma_Z(t + h, t).$$

Neste trabalho a função $\gamma_Z(\cdot)$ é chamada de função de autocovariância. O valor $\gamma_Z(h)$ é o valor para lag h .

Aqui, seguindo o livro de Brockwell et al. (2001) [15], função γ_Z de duas variáveis é chamada de função de covariância (Definição 2.3.2) e a função γ_Z de uma variável é chamada de função de autocovariância. Mas, em alguns outros livros (por exemplo, Morettin et al. (2006) [13]) a função γ_Z de duas variáveis também é chamada função de autocovariância.

Definição 2.3.4 *Um processo estocástico $Z = \{Z_t, t \in \mathbb{Z}\}$ diz-se gaussiano se, para qualquer conjunto, t_1, t_2, \dots, t_n pertencentes a \mathbb{Z} , as variáveis aleatórias $Z(t_1), Z(t_2), \dots, Z(t_n)$ tem distribuição normal n -variada.*

Este será também determinado com o cálculo de médias e covariâncias. Então, para termos práticos se faz necessário o conhecimento de tais medidas.

Para um processo estacionário, média e variância mantem-se constantes ao longo do tempo, assim como a função de autocovariância se faz importante na hora de determinação de estacionariedade na série. A função de autocovariância, quando um processo $\{Z_t, t \in \mathbb{Z}\}$, será estacionário real discreto dado a seguinte

Proposição 2.3.5 *Consideremos $\{Z_t, t \in \mathbb{Z}\}$ um processo estacionário real discreto, cuja média é representada por zero e por conseguinte $\gamma_\tau = E(Z_t Z_{t+\tau})$. A função de autocovariância γ_τ satisfaz as propriedades abaixo:*

- $\gamma_0 > 0$;
- $\gamma_{-\tau} = \gamma_\tau$;
- $|\gamma_\tau| \leq \gamma_0$;
- γ_τ é não negativa definida em termos de $\sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma_{\tau_j} \gamma_{\tau_k}$. Onde a e τ são definidos para quaisquer números pertencentes ao conjunto dos inteiros.

Além da função de autocovariância, pode-se tirar também a função de autocorrelação, dada pela seguinte expressão:

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0}, \tau \in \mathbb{Z},$$

onde $\rho_0 = 1$ e quando se tem γ_0 representa a variância. Dado uma amostra a função de autocovariância amostral e autocorrelação amostral é dada abaixo, respectivamente:

$$\hat{\gamma}(h) := n^{-1} \sum_{t=1}^{n-|h|} (Z_{t+|h|} - \bar{x})(Z_t - \bar{x}), -n < h < n$$

e

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}.$$

Onde, \bar{x} é a média amostral definida como $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. Abaixo segue definições e proposições para uma melhor descrição de um processo estocástico definido na reta.

Definição 2.3.6 *Seja $\{Z_t, t \in \mathbb{Z}\}$ um processo de segunda ordem. $\{Z_t\}$ é contínuo em média quadrática no ponto t_0 , se e somente se:*

$$\lim_{t \rightarrow t_0} E(|Z_t - Z_{t_0}|^2) = 0.$$

Proposição 2.3.7 *Continuidade de γ_τ para $\tau = 0$ implica em continuidade de γ_τ para todo τ .*

Proposição 2.3.8 *Se γ_τ for contínua, então $\{Z_t\}$ é contínuo em média quadrática.*

Vale a pena observar que quando $|\tau| \rightarrow \infty$ a função de autocovariância deste processo estocástico tenderá a 0.

Teorema 2.3.9 *Uma função de real valor definida nos inteiros é a função de autocovariância da série temporal estacionária se e somente se esta é não negativa definida.*

Aqui será tratado modelos que foram apropriados para os chamados processos não estacionários homogêneos (nível muda conforme o tempo), tais processos podem ser transformados em estacionários, como por exemplo, através de sucessivas diferenças. Assim, se um processo estocástico for estacionário de segunda ordem ele será estritamente estacionário e se for gaussiano, determinar-se-á pela média e covariância como versado no livro de Brockwell et al. (2001) [15].

Definição 2.3.10 Um processo aleatório $a_t : t \in \mathbb{Z}$ é dito ser um ruído branco com variância σ^2 , onde $\sigma^2 > 0$, se e somente se:

- $E(a_t^2) < \infty, \forall t \in \mathbb{Z}$;
- $E(a_t) = \mu_a$, para $t \in \mathbb{Z}$;
- $Var(a_t) = \sigma^2, \forall t \in \mathbb{Z}$;
- $Cov(a_t, a_s) = 0, \forall t, s \in \mathbb{Z}, t \neq s$.

Note que, se a_t é uma sequência de variáveis aleatórias não correlacionadas, com média 0 e variância σ^2 , então a_t é claramente estacionária com a mesma função de covariância.

Notação: Pode-se dizer que $a_t \sim WN(0, \sigma^2)$, onde WN significa *White Noise*, tal notação foi retirada do livro de Brockwell et al. (2001) [15].

Afirmações com respeito a covariância são validadas, pois as variáveis aleatórias são quadraticamente integráveis. Se tomarmos uma noção de ruído branco mais restrita, teremos a definição de ruído independente e identicamente distribuído (*i.i.d.*).

É interessante ressaltar que, um processo aleatório $\{a_t : t \in \mathbb{Z}\}$ é dito ser um ruído *i.i.d.* se as *v.a.*'s a_t são independentes. Esta sequência, definida acima, é chamada também de *processo puramente aleatório*. Onde,

- As variáveis aleatórias $a_t, t \in \mathbb{Z}$ são independentes;
- $a_t \sim IID(0, \sigma_a^2)$.

Exemplo 2.3.11 Consideremos o método de estimação do modelo ARMA e GARCH separadamente com os resíduos normais, segue que o gráfico (5) da primeira diferença juntamente com a aplicação do logaritmo neperiano já torna as três séries estacionárias.

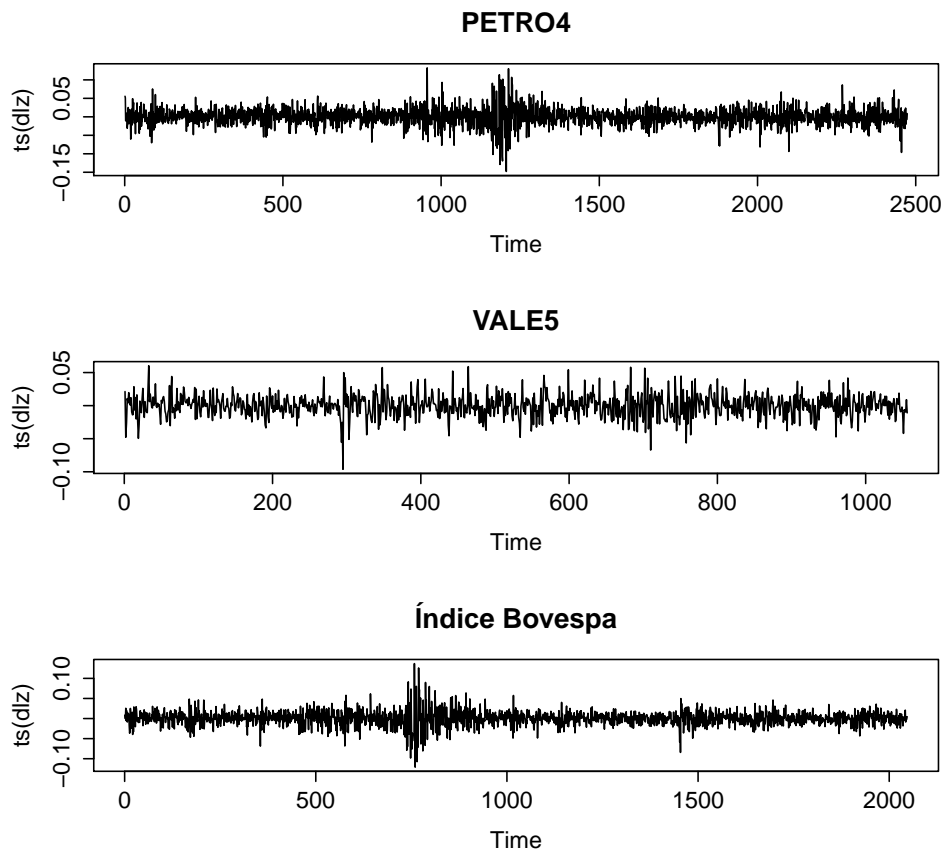


Figura 2: Gráficos das séries Petrobrás (PETR4), Vale (VALE5) e Índice Bovespa (IBOV11) da primeira diferença do logaritmo neperiano.

2.4 Retornos

Fundos de investimentos e títulos de dívidas públicas são os chamados ativos financeiros. Estes podem ser da forma de dinheiro ou representar algum tipo de posse por um bem. O dinheiro que uma pessoa recebe após um investimento num tipo de ativo é chamado de retorno. Em séries temporais é comum estudar a distribuição dos retornos. Porém, uma observação a ser feita é a de que em geral, a frequência de ocorrência de valores extremos nas séries financeiras geralmente apresentam um desvio quase universal da normalidade, sendo ordens de grandeza maior do que a prevista. Segue que, todos os resultados são retirados do livro de Morettin et al. (2006) [13].

Um dos principais motivos no estudo de finanças se faz pelo interesse de conhecer o mercado através dos ativos financeiros. Os dados dos ativos são muito complicados de se analisar e modelar, assim uma outra alternativa se faz através de análises e modelagem nos dados dos retornos destes ativos por serem livres de escalas. Para modelar tais dados, é comum o uso de diversas classes de modelos como: ARMA, ARCH e GARCH. Supondo

que não há dividendos pagos naqueles períodos, são apresentadas as fórmulas de *retorno líquido*, *log-retorno* e *retorno* respectivamente. Note que, a terceira equação (dos retornos) é tomado o logaritmo dos preços e em sequência a primeira diferença.

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}},$$

$$r_t = \ln \frac{P_t}{P_{t-1}} = \ln(1 + R_t) = p_t - p_{t-1}$$

e

$$r_t = \Delta \ln P_t.$$

Quando considera k no retorno de uma série, as fórmulas mudam de acordo com que segue abaixo:

$$R_t(k) = \frac{P_t - P_{t-k}}{P_{t-k}}.$$

Consideremos 1 período:

$$1 + R_t(k) = (1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k+1}) = \frac{P_t}{P_{t-1}} \frac{P_{t-1}}{P_{t-2}} \dots \frac{P_{t-k+1}}{P_{t-k}} = \frac{P_t}{P_{t-k}}.$$

Agora, consideremos que há pagamentos de dividendos no período, a fórmula mudará para:

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1.$$

Ou,

$$r_t = \log(1 + R_t) = \log(P_t + D_t) - \log(P_{t-1}).$$

Motivados pelo estudo de séries econômicas e financeiras, são apresentadas certas características comuns entre as demais séries. Tais como:

- Tendências;
- Sazonalidade;
- Pontos influentes;
- Heterocedasticidade condicional;
- Não linearidade.

De todas estas características, a mais difícil de analisar é a linearidade (tanto em finanças quanto em economia corresponde a uma mesma influência quando tem um aumento ou uma diminuição na série). Por outro lado, a grande maioria destes dados são livres de tendência (crescimento ou decrescimento no gráfico da série) e sazonalidade (comportamento que se repete de acordo com o ciclo). Visto que neste estudo foi usado os retornos dessas séries, a seguir são apresentadas certas características do retorno:

- Não-autocorrelacionados;
- Quando se eleva os valores ao quadrado, são auto-correlacionados;
- A distribuição (incondicional) desses retornos financeiros tem caudas mais pesadas que a distribuição Normal;
- Pode ter algumas séries não lineares.

2.5 Tipos de modelos

Por definição, um *Modelo de Séries temporais* para os dados observados z_t , onde $t \in \mathbb{Z}$, é uma especificação das distribuições conjuntas (ou, eventualmente, apenas a média e covariâncias) de uma sequência de variáveis aleatórias Z_t dos quais z_t é tida como uma realização, note que o termo *Série Temporal* deve ser utilizado tanto para os dados quanto para o processo que é uma realização. Um modelo probabilístico completo de séries temporais para a sequência de variáveis aleatórias Z_1, Z_2, \dots , deve especificar todas as distribuições conjuntas dos vetores aleatórios $(Z_1, Z_2, \dots, Z_n)'$, com $n = 1, 2, 3, \dots$ ou equivalentemente todas as probabilidades $P(Z_1 \leq z_1, Z_2 \leq z_2, \dots, Z_n \leq z_n)$ com $n = 1, 2, 3, \dots$ e $-\infty < z_1, z_2, \dots, z_n < \infty$, para mais informações ver o livro de Brockwell et al. (2001) [15].

Definição 2.5.1 *Um modelo de decomposição clássico é dado por:*

$$Z_t = \mu_t + s_t + a_t,$$

onde, μ_t é a tendência da série, s_t é a sazonalidade e a_t é o ruído branco. É notório que $E(a_t) = 0$, $s_{t+d} = s_t$ e $\sum_{j=1}^d s_j = 0$, com $t \in \mathbb{N}$.

O objetivo da determinação de um modelo de séries temporais é descrever os dados observados pelo tempo. Este modelo deve ter a característica de possuir uma parcela que

aborda a autocorrelação entre defasagens no passado e outra identificando a contribuição dos erros do passado afetando valores presentes. Ao analisar, por exemplo, um ativo financeiro a ideia é verificar o quanto os valores passados influenciam valores futuros e o quanto os erros do passado podem refletir nos valores futuros.

Para a construção de um modelo de uma série, as etapas consistem em usar procedimentos para modelar os dados com a classe de modelos lineares, logo após com a classe de modelos não lineares e ao final calcular a previsão da volatilidade do ativo e o VaR. Criado por Box e Jenkins no ano de 1970, os modelos lineares foram feitos para capturar a autocorrelação e criar modelos com uma parte mostrando o quanto o passado afeta o presente e em outra parte o quanto os erros afetam o hoje.

Um problema dos modelos lineares se dá quando há um aumento gradativo da variância. Então, Engle no ano de 1982 e Bollerslev no ano de 1986 criaram os chamados modelos não lineares afim de controlar a variância e verificar o comportamento dos resíduos e dos resíduos ao quadrado ao longo do tempo, livro de Morettin et al. (2006) [13]. Uma diferença entre estes dois tipos de modelos pode ser a quantidade de parâmetros, nos modelos paramétricos há um número finito e pequeno de parâmetros o que não acontece com os modelos não paramétricos.

Devido a esta razão, a utilização de modelos paramétricos torna mais fácil a identificação dos parâmetros que determinam o modelo. Nesta seção será abordado este tema além de introduzir uma noção de como discernir na busca do modelo que melhor representa uma série temporal. Usando a metodologia de Box e Jenkins, abaixo é dado o seguinte algoritmo de identificação de um modelo:

- Uma classe geral de modelos é considerada para a análise;
- Há identificação de um modelo, com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
- A seguir vem a fase de estimação, na qual os parâmetros do modelo identificado são estimados;
- Finalmente, há a verificação ou diagnóstico do modelo ajustado, através de uma análise de resíduos, para se saber se este é adequado para os fins em vista (previsão, por exemplo).

O algoritmo é feito repetidamente, do passo 2 em diante, até o modelo ser adequado. Usualmente é feito a identificação de vários tipos de modelos para a determinação dos

parâmetros seguindo da melhor escolha. Se o escopo da modelagem é a previsão, o modo com o qual será determinado o modelo se dá através do modelo que possuir o menor erro quadrático médio.

Uma desvantagem deste método será durante a escolha de um modelo, pois nesta etapa, a opção varia conforme as análises e sensibilidade. Assim, pode-se escolher modelos que são bons mas que ainda possam existir outros melhores para a modelagem das séries em questão. Em geral, os modelos são parcimoniosos, uma consequência é a precisão.

Seguindo, neste trabalho os modelos abordados serão modelos lineares estacionários, casos particulares de um modelo de filtro linear (procedimento que tem como escopo analisar séries quando estas possuem tendência). Estes tipos de modelos tem por suposição a construção através de um filtro linear, com a entrada sendo um ruído branco a_t . Deste procedimento, é dado o modelo abaixo.

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \psi(B)a_t. \quad (2.3)$$

Onde, $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$ é chamada de *função de transferência* do filtro, o parâmetro μ determina o nível da série em questão e B é o operador translação para o passado, ou seja, $BZ_t = Z_{t-1}$.

Definição 2.5.2 $\{Z_t\}$ na equação (2.3) é especificado como um processo linear discreto, se:

- $E(a_t) = 0, \forall t;$
- $Var(a_t) = \sigma_a^2, \forall t;$
- $E(a_t a_s) = 0, s \neq t.$

Lembremos que a média do ruído branco será zero, assim como a variância será a constante σ_a^2 e sua covariância será zero, nos dois primeiros casos para todo t e no último para qualquer s diferente de j .

Admitindo a seguinte transformação $\tilde{Z}_t = Z_t - \mu$ implica em $\tilde{Z}_t = \psi(B)a_t$. Se a sequência de pesos $\{\psi_j, j \neq 1\}$ for finita ou infinita e convergente este filtro será um *Filtro Estável* e a série $\{Z_t\}$ será classificada como estacionária, caso contrário a série será não-estacionária. O fato é que tendo uma série estacionária, o parâmetro μ tem significado como um ponto de referência para o nível da série o que não acontece com uma série

que não seja estacionária. E sua média será dada por $E(Z_t) = \mu + E\left(a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j}\right)$.

Segue-se que a função de autocovariância é dada por:

$$\gamma_j = \sigma_a^2 \sum_{i=0}^{\infty} \psi_j \psi_{j+i}. \quad (2.4)$$

Para que (2.4) exista, a condição é $\sum_{i=0}^{\infty} \psi_j^2 < \infty$. Tem-se que, a média e a variância são constantes e que a covariância só depende de j , pode admitir que a série $\{Z_t\}$ é estacionária. É possível ainda reescrever Z_t assim: $\tilde{Z}_t = \pi_1 \tilde{Z}_{t-1} + \pi_2 \tilde{Z}_{t-2} + \dots + a_t$. Segue-se que: $\left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) \tilde{Z}_t = a_t$.

Ou ainda, $\pi(B) \tilde{Z}_t = a_t$, onde $\pi(B)$ é o operador.

Definição 2.5.3 *Uma série temporal $\{Z_t\}$ é um processo linear se tem a seguinte representação:*

$$Z_t = \sum_{j=-\infty}^{\infty} \phi_j a_{t-j},$$

para todo t onde $a_t \sim WN(0, \sigma^2)$ e ϕ_t é uma sequência de constantes com $\sum_{j=-\infty}^{\infty} |\phi_j| < \infty$.

Proposição 2.5.4 *Um processo linear será estacionário se a série $\psi(B)$ convergir para $|B| \leq 1$; será invertível se $\pi(B)$ convergir para $|B| \leq 1$.*

Tais resultados desta seção foram extraídos do livro de Morettin et al. (2006) [13].

2.5.1 Modelos Auto-regressivos

Nesta seção será introduzido um dos modelos lineares chamado de auto-regressivo (AR) cujo parâmetro é p . Assim, como os demais modelos, os modelos auto-regressivos de ordem p tem que admitir certas propriedades como estacionariedade e também invertibilidade como discutido no livro de Brockwell et al. (2001) [15].

Definição 2.5.5 Z_t é um processo autoregressivo de ordem p , se:

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t. \quad (2.5)$$

Onde $\{a_t\} \sim WN(0, \sigma^2)$ e ϕ_1, \dots, ϕ_p são constantes e $\tilde{Z}_t = Z_t - \mu$.

Aqui, foram renomeados os pesos de π_j para ϕ_j , agora, definindo o operador auto-regressivo estacionário de ordem p é obtido:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p.$$

Reescrevendo a fórmula dada acima:

$$\phi(B)\tilde{Z}_t = a_t.$$

Ao tratar de invertibilidade e de estacionariedade de uma série, pode fazer uma expansão do operador linear $\pi(B) = \phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$. E, consideremos que seja finito não haverá restrições sobre os parâmetros do modelo sugerido que considere a invertibilidade da série Z_t . Por conseguinte, tomando $G_i^{-1}, i = 1, 2, \dots, p$ como raízes da equação característica $\phi(B) = (1 - G_1 B) \dots (1 - G_p B)$, é possível obter a fórmula abaixo:

$$\psi(B) = \phi^{-1}(B) = \sum_{i=1}^p \frac{A_i}{1 - G_i B}. \quad (2.6)$$

Modelos auto-regressivos são sempre invertíveis, porém, nem sempre estacionários. Para que na equação (2.6) a série seja estacionária, deve-se ter a convergência desta (caso em que $|B| \leq 1$ e também $|G_i| \leq 1, i = 1, 2, \dots, p$).

2.5.2 Médias Móveis

Considere agora que o processo seja linear e além disso que $\psi_j = 0, j > q$, para o modelo de médias móveis denotado por $MA(q)$ (de *moving average*), os resultados são retirados do livro de Morettin et al. (2006) [13] e Brockwell et al. (2001) [15].

Definição 2.5.6 Z_t é um processo de médias móveis de ordem q se:

$$\tilde{Z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}. \quad (2.7)$$

Onde, $\{a_t\} \sim WN(0, \sigma^2)$, $\theta_1, \dots, \theta_n$ são constantes e $\tilde{Z}_t = Z_t - \mu$.

E tendo $\tilde{Z}_t = Z_t - \mu$, de (2.7) tem-se:

$$\tilde{Z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t = \theta(B) a_t,$$

Em que, $\theta(B)$ é o operador de médias móveis cuja a ordem será q .

É comprovado que sempre um modelo de médias móveis será estacionário, então, é considerado $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$. Já em termos de invertibilidade, fazendo uma analogia ao raciocínio usado nos modelos auto-regressivos, para verificação de invertibilidade para um modelo MA(q) as raízes da equação característica denotada por $\theta(B) = 0$ estiverem fora do círculo unitário, então o modelo de médias móveis será invertível. Observe que um modelo de médias móveis de ordem q é análogo a um modelo auto-regressivo de ordem infinita.

Usando um raciocínio análogo ao que foi usado no modelo auto-regressivo, é obtido respectivamente a expressão para a autocorrelação do modelo de médias móveis.

2.5.3 Modelos ARMA

Os modelos anteriores, auto-regressivos e de médias móveis, são um caso particular de um modelo ARMA (modelos auto-regressivos e de médias móveis). Nesta seção, será apresentado o modelo ARMA, onde os resultados são extraídos do livro de Morettin et al. (2006) [13] e Brockwell et al. (2001) [15].

Definição 2.5.7 Z_t é um processo ARMA de ordem p e q se Z_t é estacionário e se para todo t :

$$\begin{aligned} Z_t - \phi_1 \tilde{Z}_{t-1} - \phi_2 \tilde{Z}_{t-2} - \dots - \phi_p \tilde{Z}_{t-p} = \\ a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}. \end{aligned} \quad (2.8)$$

Onde, $\{a_t\} \sim WN(0, \sigma^2)$ e os polinômios $(1 - \phi_1 z - \dots - \phi_p z^p)$ e $(1 - \theta_1 z - \dots - \theta_q z^q)$ não têm fatores comuns.

Sabendo que $\phi(B)$ e $\theta(B)$ são os operadores dos modelos auto-regressivos e de médias móveis, respectivamente, então é possível fazer uso da seguinte relação:

$$\phi(B) \tilde{Z}_t = \theta(B) a_t.$$

Este modelo é muito utilizado em diversas áreas como Economia, Ciências Físicas e Geofísicas. No caso em que as raízes de $\phi(B) = 0$ (ou seja, os valores de ϕ estiverem fora do círculo unitário) o modelo é estacionário e tendo as raízes $\theta(B) = 0$ (isto é, valores de θ fora do círculo unitário) o modelo será invertível. Fazendo um processo análogo ao que foi visto nos modelos auto-regressivos. Para chegar à fórmula de autocorrelação, basta multiplicar ambos os lados por \tilde{Z}_t e em seguida aplicar a esperança. O cálculo segue

abaixo.

$$\gamma_j = E(\tilde{Z}_t \tilde{Z}_{t-j}) = E((\phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}) \tilde{Z}_{t-j}).$$

Ou ainda, a equação apresentada acima pode estar expressa como a seguir:

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p} - \theta_1 \gamma_{za}(j-1) - \theta_2 \gamma_{za}(j-2) - \dots - \theta_q \gamma_{za}(j-q).$$

Onde, $\gamma_{za}(j) = E(a_t \tilde{Z}_{t-j})$ é a covariância cruzada entre o ruído e a $(t-j)$ -ésima defasagem. Deste modo, como as defasagens da série só dependem de choques a_t que aconteceram até o instante $t-j$, a $\gamma_{za}(j)$ será 0 para valores de j maiores que 0 e diferente de zero para valores de j menores que 0. Então, é obtido o seguinte resultado

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}, \quad j > 0$$

. Dividindo esta equação por γ_0 é encontrado a função de autocorrelação para modelos ARMA:

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}. \quad (2.9)$$

Na equação (2.9) se deduz que as *lags* (denominação das autocorrelações) sofrerão influência dos parâmetros das médias móveis do modelo ARMA, no entanto para $j > q$ estas se comportam como um modelo auto-regressivo.

Definição 2.5.8 Um ARMA(p, q) de um processo Z_t é invertível se existem constantes π_j tal que $\sum_{j=0}^{\infty} |\pi_j| < \infty$ e,

$$Z_t = \sum_{j=0}^{\infty} \pi_j Z_{t-j}, \quad \text{para todo } t. \quad (2.10)$$

2.5.3.1 Função de autocorrelação parcial

O procedimento de identificação de um modelo AR(p), MA(q) ou ARMA(p, q) consiste na observação dos gráficos da função de autocorrelação e função de autocorrelação parcial dos dados em questão. Cada modelo tem um padrão específico nestes dois gráficos, a seguir será apresentado uma tabela com o resumo do comportamento do correlograma em relação aos modelos especificados no livro de Morettin et al. (2006) [13].

Para a identificação dos modelos ARMA(p, q), não existe um padrão certo no corre-

Tabela 1: Tabela de identificação dos modelos de séries temporais.

Modelo	FAC	FACP
AR(p)	Decaimento exponencial e/ou senoide amortecida, infinita em extensão.	FACP $\phi_{kk} \neq 0$, para $k > p$.
MA(q)	FAC infinita, corte após a autocorrelação q .	Decaimento exponencial e/ou senoide amortecida, infinita em extensão.

lograma. Então um dos possíveis procedimentos será usar o critério de Akaike.

2.5.4 Modelos não-estacionários

Séries econômicas e financeiras são, em sua maioria, não-estacionárias. Devido a isso esta seção vem abordar modelos que se desenvolvem ao longo do tempo não tendo uma média e variância constantes. Note que muitas destas séries, denominadas como *séries não-estacionárias homogêneas* ou *portadoras de raízes unitárias*, depois de diferenciadas se tornam estacionárias. Ao se considerar um modelo ARMA como em (2.8), a condição para a série ser estacionária está em verificar se $\phi < 1$. Se $\phi = 1$ é obtido um processo não estacionário e para $\phi > 1$, o processo explode à medida que t aumenta. Porém, aqui será tratado séries que não explodem, todavia podem ser não-estacionárias e ainda possuir tendência determinística (quando possui intercepto) como no modelo a seguir do livro de Morettin et al. (2006) [13]:

$$X_t = \beta_0 + \beta_1 t + \varepsilon_t \quad (2.11)$$

Em que, $a_t \sim WN(0, \sigma^2)$.

- $E(X_t) = \mu_t = \beta_0 + \beta_1 t$;
- Mesmo tomando a diferença, a série $X_t - X_{t-1} = \beta_1 + a_t - a_{t-1}$ é ainda não-estacionária e também não-invertível, consideremos $\phi = \theta = 1$;
- Modelos de médias móveis são sempre estacionários, mas nem sempre invertíveis;
- Extraíndo a tendência de (2.11) o modelo torna estacionário.

2.5.4.1 ARIMA

No livro de Morettin et al. (2006) [13], para que os modelos não estacionários sejam estacionários é considerado $W_t = \Delta^d Z_t$ para a d -ésima diferença da série Z_t , assim é possível ter W_t como um modelo $ARMA(p, q)$, isto é:

$$\phi(B)W_t = \theta(B)a_t. \quad (2.12)$$

Tendo em vista que W_t é uma diferença da série original, chega a conclusão de que Z_t é uma integral de W_t , reescrevendo (2.12) como:

$$\phi(B)\Delta Z_t = \theta(B)a_t. \quad (2.13)$$

É dito que a série Z_t segue um modelo $ARIMA(p, d, q)$, onde p e q são as ordens de $\phi(B)$ e $\theta(B)$ respectivamente. Note que todas as raízes estão fora do círculo unitário. Então, reescrevendo a equação (2.12) e considerando $\varphi(B)$ como o operador auto-regressivo, tem a seguinte fórmula:

$$\varphi(B) = \phi(B)\Delta^d Z_t = \phi(B)(1 - B)^d.$$

Este modelo há de supor que a d -ésima diferença pode ser representada por um modelo ARMA (que seja estacionário e invertível). Em sua maioria, é necessário fazer somente até a segunda diferença, $d = 1$ como apresentado no livro de Morettin et al. (2006) [13].

2.5.4.2 Identificação dos modelos

O modelo de Box & Jenkins tem sua parte mais crítica na hora de identificar os parâmetros do modelo a ser usado. Pode-se ter bons modelos para uma mesma série temporal dependendo de como foi prescrito critérios de determinação destes. Neste tópico irá ser abordado o tema e técnicas para uma escolha adequada desses modelos. As escolhas terão como base as funções de autocorrelação e autocorrelação parcial estimadas representadas abaixo, note que os resultados foram retirados do livro de Morettin et al. (2006) [13]:

$$r_j = \frac{c_j}{c_0}, j = 0, 1, \dots, N - 1,$$

onde c_j é a estimativa da *facv* (função de autocovariância) γ_j ,

$$c_j = \frac{1}{N} \sum_{t=1}^{N-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})].$$

Sendo \bar{Z} a média amostral da variável Z e $r_j = r_{-j}$. A variância da *fac* (função de autocorrelação) estimada será:

$$\text{Var}(r_j) \cong \frac{1}{N} \sum_{v=-\infty}^{\infty} [\rho_v^2 + \rho_{v+j}\rho_{v-j} - 4\rho_j\rho_v\rho_{v-j} + 2\rho_v^2\rho_j^2]. \quad (2.14)$$

Se as autocorrelações forem nulas a fórmula (2.14) reduzirá para,

$$\text{Var}(r_j) \cong \frac{1}{N} [1 + 2 \sum_{v=1}^q \rho_v^2], j > q.$$

Como não é conhecido ρ_j , mas calculado uma estimativa então a equação (2.15) terá a seguinte forma:

$$\hat{\sigma}^2 \cong \frac{1}{N} [1 + 2 \sum_{v=1}^q r_v^2], j > q.$$

Para N suficientemente grande e consideremos $\rho = 0$, é tido a hipótese de normalidade do estimador r_j com média zero e variância dada por (2.15), podendo construir o seguinte intervalo de confiança:

$$r_j \pm t_\gamma \hat{\sigma}(r_j),$$

onde, t_γ é a estatística de teste *T - student* com $N - 1$ graus de liberdade. A ideia do teste será em analisar os valores calculados, caso estejam dentro do intervalo de confiança é porque há autocorrelação nos dados. O procedimento de identificação consiste em se obter os valores para os parâmetros p, d, q do modelo geral *ARIMA*, o processo consta de três etapas:

- Verificar a necessidade ou não de uma transformação na série original;
- Tomar d diferenças necessárias para a retirada de tendência na série (uma das hipóteses para a construção de um modelo *ARIMA*), para tal, temos o teste de **Dickey Fuller** para raízes unitárias;
- Através das contas *fac* e *facv* ou da análise gráfica do correlograma, faz o uso de

determinação do modelo AR de ordem p ou MA de ordem q que se tratará. Para a identificação de modelos ARMA, a alternativa será em comparar o Critério de Akaike.

Para o segundo passo, encontra-se um grande problema em relação ao número de vezes para a diferenciação numa série, McLeod em 1983 aborda este assunto com elaborados comentários.

- Um número excessivo de diferenças resulta em um valor negativo da autocorrelação de ordem 1 e da série diferenciada, neste caso $\rho_1 = 0,5$;
- Quando a série é completamente diferenciada a variância da série transformada diminui, por outro lado, excesso de diferenças aumentará essa variância. Assim, o monitoramento da variância é bastante útil para escolher o valor apropriado de d .

Exemplo 2.5.9 Consideremos o método de modelagem ARMA e GARCH separadamente e os resíduos seguindo uma distribuição normal, para as séries de dados que serão modeladas aqui, um dos primeiros procedimentos foi analisar o comportamento das observações através do correlograma para identificar qual modelo linear poderia ser usado.

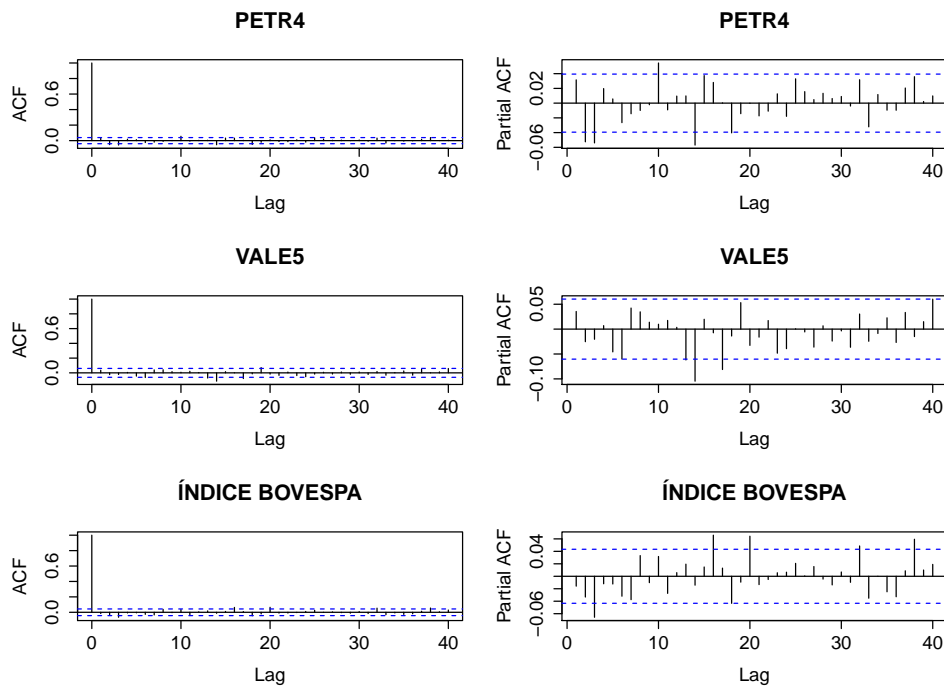


Figura 3: Autocorrelações e autocorrelações parciais amostrais para os dados de retorno das séries PETR4, VALE5 e Índice Bovespa.

2.5.5 Modelos não lineares

Mesmo após o tratamento de uma série, a adequação de um modelo do tipo ARMA(p, q), AR(p) ou MA(q) e seu diagnóstico, é possível haver o caso em que os resíduos sejam correlacionados quando elevados ao quadrado. Note que os testes na etapa de diagnóstico verificam somente se os resíduos são correlacionados ou não, estes não testam o caso em que os resíduos são independentes. Então, para alguns casos quando aplica o termo quadrado nos resíduos eles se apresentam correlacionados. Para resolver este problema foi feito modelos não lineares, introduzidos por Engle(1982) [11], como ARCH (*autoregressive conditional heterocedasticity*) e GARCH (*Generalized ARCH*) criados por Bollerslev (1986) [12]. Estes modelos também resolvem o caso de variância condicional, muito comum em séries financeiras, como discutido no livro de Morettin et al. (2006) [13].

2.5.5.1 Modelos ARCH

O objetivo dos modelos auto-regressivos com heterocedasticidade condicional, quando foram criados, era estimar variabilidade da inflação. O retorno X_t é não-correlacionado mas a volatilidade (que é a variância condicional) depende de retornos passados. A ideia de construção deste modelo está em considerar o retorno (ou os resíduos dos retornos, caso em que se usa os modelos lineares primeiro) como X_t é a volatilidade dependente de retornos (ou dos resíduos dos retornos) passados. Aqui, os resultados são retirados do livro de Morettin et al. (2006) [13]. Assim, o modelo ARCH de ordem r é dado pela seguinte

Definição 2.5.10 *Um modelo ARCH (r) é definido por:*

$$\begin{aligned} X_t &= \sqrt{h_t} \varepsilon_t, \\ h_t &= \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2, \end{aligned} \tag{2.15}$$

onde ε_t é uma sequência de v.a.'s i.i.d com média zero e variância constante e igual a 1. Na prática é comum supor que os resíduos têm distribuição Normal (0, 1) ou T-student com v graus de liberdade. Mas, aqui será considerado o seguinte caso:

- ε_t seguindo uma distribuição Normal;
- ε_t seguindo uma distribuição GHsT.

Para um modelo igual a expressão (2.16) em que $r = 1$,

$$\begin{aligned} X_t &= \sqrt{h_t} \varepsilon_t, \\ h_t &= \alpha_0 + \alpha_1 X_{t-1}^2. \end{aligned} \tag{2.16}$$

Tem-se as seguintes propriedades:

- $E(X_t) = E(X_t|F_{t-1}) = 0$;
- $Var(X_t) = \alpha_0 + \alpha_1 E(X_{t-1}^2) = \frac{\alpha_0}{1-\alpha_1}$. Só se é válida se X_t for definido como um processo estacionário de segunda ordem ($Var(X_t) > 0$ e $0 \leq \alpha_1 \leq 1$);
- $Cov(X_{t+k}, X_t) = 0$, para $k \neq 0$ e $\gamma_X(k) = 0$.

Nas séries financeiras os retornos apresentam geralmente caudas longas, por isso se faz tão importante o uso do modelo em questão. Então na hora do cálculo da curtose (importante como uma medida, tendo que ser maior do que 3) é preciso calcular o momento de quarta ordem de X_t . Para isso é suposto que X_t siga o modelo dado pela equação (2.16) e também é suposto que os ε_t sejam normais, para uma facilidade de cálculo. Assim, o momento de quarta ordem pode ser escrito como $\mu_4 = E(X_t^4)$.

$$\mu_4 = \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}.$$

Assim, pode-se calcular a curtose da seguinte maneira:

$$\begin{aligned} K &= \frac{E(X_t^4)}{[E(X_t^2)]^2}, \\ K &= \frac{\mu_4}{[Var(X_t)]^2} > 3. \end{aligned} \tag{2.17}$$

Vale a pena observar que quanto mais restrição ao modelo, mais restrição aos parâmetros. Para a identificação de um modelo ARCH de parâmetro r são feitos os seguintes procedimentos:

- Na primeira etapa é feito o ajuste de modelos ARMA nos dados para que seja possível a remoção de correlação serial na série, caso exista;
- Seguindo, é considerado que o ruído branco segue uma distribuição ARCH, ou seja, $a_t \sim ARCH(r)$.

Então, dois testes poderão ser usados para identificar a heterocedasticidade condicional:

- Teste de Box-Pierce (ou Ljung-Box) para X_t^2 ;
- Teste de multiplicadores de Lagrange.

Aqui, será usado o teste de Ljung-Box (teste de Box-Pierce) para modelos não lineares.

2.5.5.2 Modelos GARCH

A forma geral do modelo ARCH é o modelo GARCH (*Generalized ARCH*), introduzida por Bollerslev (1986). Este modelo tem uma relação com o modelo ARMA(p, q), enquanto que o modelo ARCH tem uma relação com o AR(p). Da mesma forma que os modelos AR e ARMA, o modelo GARCH pode ser estabelecido para se ter menos parâmetros em comparação com um modelo ARCH. Assim como o modelo ARCH, o uso do modelo GARCH é no sentido de descrever a volatilidade, porém com menos parâmetros. Resultados são retirados do livro de Morettin et al. (2006) [13].

Definição 2.5.11 *Um modelo da forma GARCH(r, s) é definido por:*

$$X_t = \sqrt{h_t} \varepsilon_t, \quad (2.18)$$

$$h_t = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}. \quad (2.19)$$

Onde ε_t é i.i.d. $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, \sum_{i=1}^q (\alpha_i + \beta_i) < 1, q = \max(r, s)$. Consideremos o retorno (ou os resíduos dos retornos, caso em que se usa os modelos lineares primeiro) como X_t . Como nos modelos ARCH em que é suposto normalidade dos resíduos, é encontrado em (2.20) um modelo para X_t^2 mas em geral v_t não é um processo *i.i.d.*. Como no modelo ARCH será tratado os resíduos como:

- ε_t seguindo uma distribuição Normal;
- ε_t seguindo uma distribuição GHsT.

Tomando $v_t = X_t^2 - h_t$, e substituindo em (2.19):

$$X_t^2 = \alpha_0 + \sum_{i=1}^q (\alpha_i + \beta_i) X_{t-i}^2 + v_t - \sum_{j=1}^s \beta_j v_{t-j} \quad (2.20)$$

Assim,

$$E(X_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^q (\alpha_i + \beta_i)}$$

Assim, o modelo em (2.20) será um ARMA(q,s) para a variável X_t^2 . Contudo, os resíduos representado por v_t não serão um processo *i.i.d.*, mas sim, uma diferença marginal. Com esperança dada por $E(v_t) = 0, \forall t \in T$ e $E(v_t|F_{t-1}) = 0, \forall t$. O modelo GARCH apresenta as mesmas vantagens e desvantagens de um modelo ARCH. Sua curtose é calculada como em (2.17). A identificação, numa série, deste modelo, é realmente difícil e assim se faz preferir o uso de ordens baixas para uma melhor identificação. A escolha dos modelos faz por meio de critérios como AIC, BIC, curtose, função perda e log-verossimilhança (2.26).

2.5.6 Estimação dos modelos

2.5.6.1 Estimação dos modelos ARIMA

Para a escolha de um modelo, o conhecimento dos valores dos parâmetros se faz necessário, porém saber quais valores assumirão tais parâmetros nem sempre é possível. Assim, uma alternativa será o uso de estimação afim de obter valores para estes parâmetros baseados nos dados anteriores. Na seção 2.5.4.2 foi abordado a noção de como escolher um modelo, daí a próxima etapa será usar este modelo provisório para a estimação dos parâmetros.

Será necessário um processo iterativo para a estimação não linear de mínimos quadrados, usar-se-ão as estimativas encontradas na fase de identificação como valores iniciais. Consideremos um modelo ARIMA e seus parâmetros $p + q + 1$ ao vetor $\varepsilon = (\phi, \theta, \sigma_a^2)$, onde ϕ e θ são vetores de ordem $p \times 1$, em caso de $d > 0$ teremos μ incluído no vetor ε e os parâmetros passarão para $p + q + 2$. Considere também $\eta = (\phi, \theta)$. Para estimar o vetor ε será usado o procedimento de Máxima Verossimilhança exata.

Função de máxima verossimilhança exata

Consideremos N suficientemente grande, aqui a forma de estimação exata foi desenvolvida por Newbold (1974). No livro de Morettin et al. (2006) [13] é citado outras referências adicionais deste método: Ansley (1979), Nicholls e Hall (1979), Ljung e Box (1979), Jenkins e Reinsel (1994).

Consideremos a série estacionária:

$$\hat{W}_t = \phi \hat{W}_{t-1} + a_t, \quad |\phi| < 1, \quad (2.21)$$

em que $\hat{W}_t = W_t - \mu$ e $a_t \sim N(0, \sigma_a^2)$ independentes. Pode-se escrever (2.21) como segue abaixo:

$$\hat{W}_t = \sum_{j=0}^{\infty} \phi^j a_{t-j},$$

onde supõe que W_t seja uma normal com média zero e variância dada por $\sigma_a^2(1 - \phi^2)^{-1}$. Daí, usando as duas equações anteriores, é possível reescrever \hat{W}_t como:

$$\hat{W}_t = \sum_{j=0}^{\infty} \phi^j a_{t-j}, \quad (2.22)$$

$$\hat{W}_t = \phi \hat{W}_{t-1} + a_t, t = 2, \dots, n. \quad (2.23)$$

Em que $(v_1, a_1, \dots, a_n) \sim N(0, \Sigma)$ e $\Sigma = \text{diag}((1 - \phi^2)^{-1}\sigma_a^2; \sigma_a^2, \dots, \sigma_a^2)$. Segue de (2.23) que a função de densidade conjunta será (v_1, a_2, \dots, a_n) :

$$f(v_1, a_2, \dots, a_n) = \left[\frac{(1 - \phi^2)}{2\pi\sigma_a^2} \right]^{\frac{1}{2}} \exp \left\{ -\frac{\hat{W}_1^2(1 - \phi^2)}{2\sigma_a^2} \right\} \left[\frac{1}{2\pi\sigma_a^2} \right]^{\frac{(n-1)}{2}} \exp \left[-\frac{1}{2\sigma_a^2} \sum_{t=2}^n (\hat{W}_t - \phi \hat{W}_{t-1})^2 \right]. \quad (2.24)$$

O logaritmo da função de verossimilhança exata do vetor $(\hat{W}_1, \dots, \hat{W}_n)$ será:

$$\ell(\mathbf{W}, \boldsymbol{\xi}) = -\frac{n}{2} \ln 2\pi + \frac{1}{2} \ln(1 - \phi^2) - \frac{n}{2} \ln \sigma_a^2 - \frac{S(\boldsymbol{\eta})}{2\sigma_a^2}. \quad (2.25)$$

Em que,

$$S(\boldsymbol{\eta}) = (W_1 - \mu)^2(1 - \phi^2) + \sum_{t=2}^n [(W_t - \mu) - \phi(W_{t-1} - \mu)]^2.$$

Para obter os estimadores é preciso derivar (2.25) e igualar a zero como expresso no livro de Morettin et al. (2006) [13]. Note que, para a estimação através da máxima verossimilhança exata só foi possível usar pois foi estimado o modelo $ARMA(p, q)$ separadamente do modelo $GARCH(r, s)$.

Exemplo 2.5.12 Usando o método de estimação dos parâmetros separadamente, ou seja, primeiro modela os modelos lineares, seguindo a modelagem por modelos não lineares e

depois o ajuste do erro para as demais análises e consideremos Z_t como a série original da Petrobrás com 2474 observações. A primeira etapa foi o cálculo da primeira diferença seguido da aplicação do logaritmo neperiano, ou seja:

$$W_t = \ln(Z_t - Z_{t-1}) = \ln(1 - B)Z_t = \ln \Delta Z_t,$$

onde, W_t é a série de retornos estacionária. O modelo escolhido foi um $AR(p = 18)$, assim:

$$\tilde{W}_t = \hat{\phi}_2 \tilde{W}_{t-2} + \hat{\phi}_3 \tilde{W}_{t-3} + \hat{\phi}_{10} \tilde{W}_{t-10} + \hat{\phi}_{14} \tilde{W}_{t-14} + \hat{\phi}_{18} \tilde{W}_{t-18} + a_t,$$

em que $\tilde{W}_t = W_t - \mu$ e $a_t \sim WN(0, \tilde{\sigma}_a^2)$. Dados os valores dos coeficientes, o modelo pode ser reescrito da seguinte maneira:

$$\tilde{W}_t = -0,0475\tilde{W}_{t-2} - 0,0536\tilde{W}_{t-3} + 0,0558\tilde{W}_{t-10} - 0,0531\tilde{W}_{t-14} - 0,0447\tilde{W}_{t-18},$$

note que neste modelo foi desconsiderada a constante.

2.5.6.2 Estimação dos modelos GARCH

A estimação de um processo ARCH de ordem r é feito através de estimadores de máxima verossimilhança condicional. Para isso, é preciso supor que os ε_t seguem uma distribuição normal. As funções são dadas abaixo:

$$L(\boldsymbol{\alpha}|x_1, x_2, \dots, x_N) = f(x_N|F_{N-1})f(x_{N-1}|F_{N-2}) \dots f(x_{r+1}|F_r)f(x_1, \dots, x_r|\boldsymbol{\alpha})$$

Para N suficientemente grande, $f(x_1, \dots, x_r|\boldsymbol{\alpha})$ pode ser desprezado.

Vale ressaltar que os estimadores do modelo são obtidos através da técnica de máxima verossimilhança condicional às r primeiras observações como em (2.26), mas considerando N suficientemente grande se aproxima da técnica de máxima verossimilhança exata.

$$\ell(\boldsymbol{\alpha}, \boldsymbol{\beta}|x_1, \dots, x_r) \propto -\frac{1}{2} \sum_{t=r+1}^N \ln(h_t) - \frac{1}{2} \sum_{t=r+1}^N \frac{x_t^2}{h_t} \quad (2.26)$$

Exemplo 2.5.13 Consideremos os dados da Petrobrás e o método de estimação separadamente de ARMA e GARCH, como visto em (2.5.12), a próxima etapa é modelar os erros. Assim, os resultados seguem abaixo:

$$X_t = \sqrt{h_t} \varepsilon_t,$$

$$h_t = 4,1544 \times 10^{-5} + 3,4523 \times 10^{-2} X_{t-1}^2 + 1,7767 \times 10^{-1} X_{t-2}^2 \\ + 4,4906 \times 10^{-1} h_{t-1} + 1,1139 \times 10^{-6} h_{t-2} + 2,6472 \times 10^{-1} h_{t-3}.$$

Note que $\varepsilon_t \sim N(0, 1)$.

2.5.6.3 Estimação dos modelos ARMA - GARCH

Para a estimação de um modelo onde se obtêm os coeficientes do ARMA e do GARCH juntos, a ideia será considerar o modelo ARMA dado por (2.8) como se segue

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} - \theta_1 X_{t-1} - \theta_2 X_{t-2} - \dots - \theta_q X_{t-q}$$

e o modelo GARCH considerado como

$$X_t = \sqrt{h_t} \varepsilon_t \\ h_t = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}.$$

Este processo de estimação é feito pela máxima verossimilhança. Em que o próximo procedimento está em calcular os coeficientes da distribuição GHsT.

2.5.7 Diagnóstico dos Modelos

A próxima etapa é de analisar realmente quando o modelo é adequado aos dados através de testes como: autocorrelação residual, box-pierce, autocorrelação cruzada, peridiograma acumulado, *etc.* Todavia, em termos práticos só serão usados os dois primeiros testes citados anteriormente. Todos estes testes verificam se a variável Z_t tem ou não uma certa dependência com as perturbações a_t que aconteceram na série em um tempo $t - 1$. Tais testes são de grande importância por que além de mostrar um modelo bem representativo para os dados, eles também permitem ver a insuficiência no modelo podendo resultar num acréscimo de mais parâmetros.

Este procedimento de acrescentar um parâmetro a mais no modelo tem um nome específico de *superajustamento*, na qual é estimado um modelo com números extras de parâmetros e se verifica quão representativo foi em comparação com um mais restrito de parâmetros, através da variância residual por exemplo. Este tipo de procedimento é de notória percepção que quando é sabido *a priori* em que direção está caminhando na decisão de um modelo de séries temporais como no livro de Morettin et al. (2006) [13].

2.5.7.1 Testes de adequação do modelo - Teste de Box-Pierce

Box e Pierce (1970) sugeriram um teste para identificação de autocorrelação nos resíduos, que após foi revisada por Ljung e Box (1978) conforme citado no livro de Moret-tin et al. (2006) [13]. Este teste permite saber se há quebras específicas no comportamento do ruído branco e também se há valores altos. Um teste semelhante ao teste de Box-Pierce é o teste de McLeod e Li (1983), ver Brockwell et al. (2001) [15].

Hipóteses a serem testadas:

$$\begin{cases} H_0 : \text{Não há autocorrelação na série.} \\ H_1 : \text{Há autocorrelação na série.} \end{cases}$$

Se o modelo for adequado a estatística de teste será - sob H_0 :

$$Q(K) = n(n+2) \sum_{j=1}^K \frac{\hat{r}_j^2}{(n-j)} \sim \chi_{(K-p-q)}^2.$$

É de notória percepção que a hipótese dos resíduos não serem um ruído branco é aceita quando se tem valores muito altos para a estatística de teste. Atualmente, este teste vem sendo criticado por ser uma medida de correlação e não de dependência. O teste de Box-pierce também serve na determinação de correlação quando se eleva os valores ao quadrado, assim:

Hipóteses a serem testadas:

$$\begin{cases} H_0 : \text{Não há autocorrelação nos valores ao quadrado da série.} \\ H_1 : \text{Há autocorrelação nos valores ao quadrado da série.} \end{cases}$$

A estatística de teste será a mesma.

2.5.7.2 Testes de adequação do modelo - Teste de BROCK, DECHERT e SCHEINKMAN (BDS)

Este teste identifica a dependência entre as observações capturadas pela integral de correlação, mais informações ver Bitencourt (2007) [16]. O procedimento do teste é aplicar o teste BDS na série original, com as seguintes

Hipóteses a serem testadas:

$$\begin{cases} H_0 : \text{Série } i.i.d. \text{ não há chances de modelagem.} \\ H_1 : \text{A série precisa ser feita modelagem.} \end{cases}$$

Se a hipótese nula for rejeitada, o próximo passo será usar a modelagem com modelos lineares e em seguida verificar se é preciso aplicar modelos não lineares. Assim,

Hipóteses a serem testadas:

$$\begin{cases} H_0 : \text{Série foi corretamente modelada.} \\ H_1 : \text{O modelo não seria adequado somente modelando de forma linear.} \end{cases}$$

A estatística de teste BDS, para estes dois casos seria - sob H_0 :

$$BDS(\varepsilon, m) = \frac{\sqrt{T}[C_m(\varepsilon) - (C_1(\varepsilon))^m]}{V_m^{\frac{1}{2}}} \sim N(0, 1). \quad (2.27)$$

Onde, V_m é a variância. A fórmula da variância é dada por:

$$V_m(\varepsilon) = 4 \left[K^m + 2 \sum_{j=1}^{m-1} K^{m-j} C^{2j} + (m-1)^2 C^{2m} - m^2 K C^{2m-2} \right],$$

sendo C o parâmetro estimado pela integral de correlação, cuja dimensão será 1. E o parâmetro K é a probabilidade calculada em cima das observações e $T_m = T - (m - 1)$. É comum usar $0,5 \leq \varepsilon \leq 2$ vezes o desvio padrão da série.

Vale ressaltar que essa aproximação é adequada desde que $\frac{T}{m} > 200$. Fora deste limite ou quando é usado os resíduos de um modelo não linear, é preciso o uso da técnica *debootstrap* para o cálculo dos valores da estatística de teste apresentada acima.

2.5.7.3 Testes de adequação do modelo - Teste de Jarque-Bera

Muito usado no mercado financeiro, o teste de Jarque-Bera verifica se os valores observados de uma determinada variável segue uma distribuição gaussiana ou não. Hipóteses a serem testadas:

$$\begin{cases} H_0 : \text{Os dados observados seguem uma distribuição normal.} \\ H_1 : \text{Os dados observados não seguem uma distribuição normal.} \end{cases}$$

Onde, a estatística de teste (sob H_0) é dada por:

$$JB = n \left(\frac{\alpha_3^2}{6} + \frac{(\alpha_4 - 3)^2}{24} \right),$$

em que, $\alpha_3 = \frac{1}{ns^3} \sum_{i=1}^n (y_i - \bar{y})^3$, $\alpha_4 = \frac{1}{ns^4} \sum_{i=1}^n (y_i - \bar{y})^4$ e $s^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$.

2.6 Modelos de previsão

A tecnologia vem se aprimorando e com isso é possível saber o que acontecerá num determinado instante de tempo seja considerando minuto, hora dia, semana ... o que se torna de grande importância para o rumo da economia. Baseados nisso, esta seção vem abordar os modelos de previsão dos modelos lineares e não lineares.

2.6.1 Previsão para modelos lineares

Consideremos um modelo Z_t sendo um modelo ARIMA (p, d, q) e utilizando a forma de equações de diferenças como segue abaixo segundo o livro e Morettin et al. (2006) [13]:

$$\begin{aligned} Z_{t+h} &= \varphi_1 Z_{t+h-1} + \varphi_2 Z_{t+h-2} + \dots + \varphi_{d+p} Z_{t+h-p-d} \\ &\quad - \theta_1 a_{t+h-1} - \theta_2 a_{t+h-2} - \dots - \theta_q a_{t+h-q} + a_{t+h}. \end{aligned}$$

Logo, uma maneira de expressar as previsões por EQM mínimo é

$$\begin{aligned} \hat{Z}_t(h) &= \varphi_1 [Z_{t+h-1}] + \varphi_2 [Z_{t+h-2}] + \dots + \varphi_{d+p} [Z_{t+h-p-d}] - \\ &\quad - \dots - \theta_1 [a_{t+h-1}] - \theta_2 [a_{t+h-2}] - \dots - \theta_q [a_{t+h-q}] + [a_{t+h}]. \end{aligned}$$

Para determinar o erro, é usual a fórmula abaixo:

$$e_t(h) = Z_{t+h} - \hat{Z}_t(h) = \psi_0 a_{t+h} + \psi_1 a_{t+h-1} + \dots + \psi_{h-1} a_{t+1} + \sum_{j=0}^{\infty} (\psi_{h+j} - \psi_{h+j}^*) a_{t-j}.$$

Portanto,

$$E[e_t(h)]^2 = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{h-1}^2) \sum_{j=0}^{\infty} (\psi_{h+j} - \psi_{h+j}^*)^2 \sigma^2. \quad (2.28)$$

Exemplo 2.6.1 Consideremos o modelo ARIMA(18,1,0) construído para a série de dados PETR4 ($Z_1 \dots Z_{2474}$) e o horizonte de tempo $h=1$, a partir do primeiro método de

modelagem tem:

$$(1 - \phi_2 B^2 - \phi_3 B^3 - \phi_{10} B^{10} - \phi_{14} B^{14} - \phi_{18} B^{18})(1 - B)Z_t = a_t$$

Então:

$$(1 - \phi_2 B^2 - \phi_3 B^3 - \phi_{10} B^{10} - \phi_{14} B^{14} - \phi_{18} B^{18})(1 - B)Z_{t+1} = a_{t+1}$$

Ou seja,

$$Z_{t+1} = \varphi_2 Z_{t-1} + \varphi_3 Z_{t-2} + \varphi_{10} Z_{t-9} + \varphi_{14} Z_{t-13} + \varphi_{18} Z_{t-17} + a_{t+1}.$$

2.6.2 Previsão para modelos não lineares

Em finanças o risco e o retorno se fazem importante para o conhecimento da série usada. O risco é denominado como uma medida de volatilidade (ou variabilidade ou incerteza) dos retornos. E, como já foi abordado, o retorno é o valor que se ganha de um bem ou um investimento. Assim, é comum análises de sensibilidade e de volatilidade para determinar o impacto de valores de uma variável numa determinada empresa. Então, determinando o modelo a próxima etapa é a de se fazer uso de previsões, mais especificamente previsões para a volatilidade. Dentro deste conceito temos dois tipos de previsão para a volatilidade como descrita no livro de Ehlers (2009) [14]:

- Previsões de Volatilidade: quando esta é determinística;
- Volatilidade Estocástica: a variância condicional varia estocasticamente ao longo do tempo

Para este trabalho, será usada a previsão de volatilidade.

2.6.2.1 Previsões de volatilidade

Consideremos uma série original W_t , o retorno (ou os resíduos dos retornos) como X_t e supondo que esta segue um processo ARCH de parâmetro r , as previsões de volatilidade

k passos a frente são obtidas, segundo Ehlers (2009) [14], como:

$$\begin{aligned}\hat{h}_t(k) &= E(h_{t+k}|I_t) \\ &= \alpha_0 + \sum_{i=1}^r E(X_{t+k-i}^2|I_t) \\ &= \alpha_0 + \sum_{i=1}^{k-1} h(k-i) + \sum_{i=k}^r \alpha_i X_{t+k-i}^2,\end{aligned}$$

onde $k \geq 1$.

Agora, consideremos uma série original W_t , o retorno (ou os resíduos dos retornos) como X_t e supondo que esta segue um processo GARCH de parâmetros r e s , as previsões de volatilidade k passos a frente são obtidas como segundo Ehlers (2009) [14]:

$$\begin{aligned}\hat{h}_t(k) &= E(h_{t+k}|I_t) \\ &= \alpha_0 + \sum_{i=1}^r E(X_{t+k-i}^2|I_t) + \sum_{j=1}^s \beta_j \\ &= \alpha_0 + \sum_{i=1}^{k-1} h(k-i) + \sum_{i=k}^r \alpha_i X_{t+k-i}^2 + \sum_{j=1}^{k-1} \hat{h}_t(k-j) + \sum_{j=k}^s h_{t+k-j},\end{aligned}$$

onde $k \geq 1$.

Exemplo 2.6.2 Consideremos X_t como a série dos retornos dos preços da série da Petrobrás e tomando o primeiro método, o próximo passo foi calcular manualmente a previsão para um dia a frente com a série devidamente modelada. Então, no primeiro método foi tido um modelo AR(17) e para os modelos não lineares foi escolhido um modelo GARCH(2,3). Consideremos a previsão estocástica e o horizonte de tempo $k = 1$, a previsão se dará por:

$$h_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2 + \hat{\alpha}_2 X_{t-2}^2 + \hat{\beta}_1 h_{t-1} + \hat{\beta}_2 h_{t-2} + \hat{\beta}_3 h_{t-3}, \quad (2.29)$$

assim, $X_{t+1} = \sqrt{h_{t+1}}\varepsilon_{t+1}$.

2.7 Tipos de distribuições

2.7.1 Distribuição Normal

Descoberta no século XVIII por Karl F. Gauss (1777 – 1855) segundo Jorion (2003) [17], as distribuições normais são importantes pelas características de representarem com

boa aproximação distribuições de frequência observadas de fenômenos naturais e físicos, quando considerado n suficientemente grande a distribuição normal serve de boa aproximação de probabilidades binomiais, é utilizada na amostragem por ter muitas distribuições de médias e proporções tendendo a ser distribuídas normalmente como discutido em Stevenson et al. (1981) [18].

A distribuição normal apresenta certas características singulares como:

- O gráfico em forma de sino;
- A curva de um gráfico de uma distribuição normal se prolonga indefinidamente em qualquer uma das direções;
- Os valores possíveis, teoricamente, vão de $-\infty$ à ∞ ;
- Como a distribuição Normal é especificada pela sua média e desvio padrão, então, para cada combinação desses parâmetros é observado uma distribuição normal diferente.

Na área de mercado financeiro, a distribuição Normal é muito usada para poder modelar os dados de séries financeiras. Assim, uma estratégia a ser feita é supor que os resíduos seguem uma distribuição Normal para poder ter uma ideia de como se comporta o modelo. Sua densidade é dada por:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right\}, x \in (-\infty, \infty),$$

em que a média é dada por μ e variância igual a σ^2 .

2.7.2 Distribuições hiperbólicas Generalizadas Assimétricas

Como já foi visto, séries de dados financeiros tem uma forma característica diferente das demais séries a serem estudadas. E, certas distribuições convencionais não são adequadas para a modelagem dos dados financeiros (mais especificamente dos retornos) devido ao fato que estas séries possuem caudas mais pesadas e assimétricas. Assim, uma família de distribuições assimétricas foram sendo desenvolvidas ao longo do tempo afim de proporcionar um melhor entendimento dos dados em questão.

Conforme o artigo de Silva (1999) [4], o autor explica que distribuições hiperbólicas são de suma importância por possuírem caudas mais pesadas e terem um bom ajuste dos dados

quando se trata de retornos de ativos financeiros. Ainda, além de se ter uma distribuição hiperbólica, esta pode ser assimétrica. Uma distribuição como esta tem importantes propriedades tais como: fechada sob condicionado, marginalização e transformações afins. E pelo fato de ser assimétrica, as caudas desta família de distribuição geralmente são pesadas.

Todavia, esta distribuição é raramente utilizada nas aplicações práticas por não ser analiticamente tratável (principalmente para amostras grandes), sendo difícil a estimação do parâmetro de determinação da subclasses como discutido no artigo de Aas et al. (2006) [5] o autor cita Prause (1999) [2]. São muitos os tipos desta classe de distribuição, a seguir será apresentado um caso específico da distribuição hiperbólica.

2.7.2.1 Distribuição Hiperbólica Generalizada Assimétrica T-student (GHsT)

A distribuição hiperbólica generalizada assimétrica T-student é uma distribuição ligada a distribuições hiperbólicas generalizadas como especificada por Prause (1999) [2]. Esta possui propriedades singulares, sendo assim uma importante distribuição para o uso de séries financeiras.

Como citado no artigo de Teixeira (2006) [3], as distribuições hiperbólicas generalizada assimétrica T-student possuem caudas mais pesadas e além disso o decaimento de uma cauda é diferente da outra. Numa é observado decaimento exponencial e na outra decaimento polinomial. Esta distribuição pode ser representada como uma mistura de duas distribuições: Normal e Gaussiana Inversa Gama (GIG), esta última também pode ser representada por uma $Ga(\nu/2, h^2\nu/2)$. Além disso, supondo $\beta = 0$, é obtida uma distribuição T-student não-centrada. Várias formas de densidade são apresentadas na literatura, abaixo é apresentada a densidade criada por Prause citada no artigo de Aas et al. (2006) [5]:

$$f_X(x; \alpha, \beta, \lambda, \delta) = \frac{(\alpha^2 + \beta^2)^{\frac{\lambda}{2}} K_{\lambda-\frac{1}{2}}(\alpha\sqrt{\delta^2 + (x - \mu)^2}) \exp((\beta(x - \mu)))}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta_\lambda K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})(\sqrt{\delta^2 + (x - \mu)^2})^{\frac{1}{2}-\lambda}}. \quad (2.30)$$

Tal que, $x \in (-\infty, \infty)$ e K_j é a função de Bessel modificada do terceiro tipo de ordem j , ver Anexo B 7.

- $\delta \geq 0, |\beta| < \alpha$ se $\lambda > 0$;
- $\delta > 0, |\beta| < \alpha$ se $\lambda = 0$;
- $\delta \geq 0, |\beta| \leq \alpha$ se $\lambda < 0$.

As propriedades abordadas acima é possível notar pelo comportamento das caudas da distribuição GH.

$$f_X(x) \sim \text{const}|x|_{\lambda-1} \exp(-\alpha|x| + \beta x),$$

onde a variável $x \rightarrow \infty$ e para todos os valores de λ . É possível perceber que a distribuição GH possui duas caudas pesadas quando $|\beta| \neq \alpha$.

Usando a definição de que a distribuição hiperbólica assimétrica T-student pode ser representada por uma mistura de normal com normal inversa, é obtida a seguinte densidade identificada por Barndorff-Nielsen (1977) no artigo de Aas et al. (2006) [5]:

$$f(z; \lambda, \delta, \gamma) = \frac{\gamma^\lambda}{\delta} \frac{z^{\lambda-1}}{2K_\lambda(\gamma\delta)} \exp\left\{-\frac{1}{2}(\delta^2 z^{-1} + \gamma^2 z)\right\}. \quad (2.31)$$

Esta densidade (2.31) pode ser representada pela seguinte relação:

$$X = \mu + \beta Z + \sqrt{Z}Y,$$

onde, $Y \sim N(0, 1)$, $Z \sim GIG(\lambda, \delta, \gamma)$ com $Y \perp Z$ e $\gamma = \sqrt{\alpha^2 - \beta^2}$. Considerando $X|Z = z \sim N(\mu + \beta z, z)$ e tomando $\lambda = -\frac{\nu}{2}$ e $\alpha \rightarrow |\beta|$ em (2.30) é obtida a *f.d.p.* dada abaixo:

$$f_X(x) = \begin{cases} \frac{2^{\frac{1-\nu}{2}} \delta^\nu |\beta|^{\frac{\nu+1}{2}} K_{\frac{\nu+1}{2}}(\sqrt{\beta^2(\delta^2 + (x-\mu)^2)}) \exp\{\beta(x-\mu)\}}{\Gamma(\frac{\nu}{2}) \sqrt{\pi} (\sqrt{\delta^2 + (x-\mu)^2})^{\frac{\nu+1}{2}}}, & \beta \neq 0 \\ \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi} \delta \Gamma(\frac{\nu}{2})} \left[1 + \frac{(x-\mu)^2}{\delta^2}\right]^{-\frac{(\nu+1)}{2}}, & \beta = 0. \end{cases}$$

Se $\beta = 0$ a distribuição passa a ser uma T-student não centrada no zero com ν graus de liberdade, espectância μ e variância igual a $\frac{\delta^2}{(\nu-2)}$. Para o caso em que $\beta \neq 0$, média e variância podem ser representadas, respectivamente, por:

$$E(X) = \mu + \frac{\beta\delta^2}{\nu-2}$$

e

$$\text{Var}(X) = \frac{2\beta^2\delta^4}{(\nu-2)^2(\nu-4)} + \frac{\delta^2}{\nu-2},$$

onde, $\nu = -2\lambda$ é o grau de liberdade. Assim, $X \sim GHsT(\mu, \delta, \beta, \nu)$.

Uma observação a ser feita é que a variância é somente finita quando $\nu > 4$. A

assimetria e curtose, que são duas importantes medidas, são dadas abaixo:

$$s = \frac{2(\nu - 4)^{\frac{1}{2}}\beta\delta}{[2\beta^2\delta^2 + (\nu - 2)(\nu - 4)]^{\frac{3}{2}}} \left[3(\nu - 2) + \frac{8\beta^2\delta^2}{\nu - 6} \right].$$

$$K = \frac{6}{[2\beta^2\delta^2 + (\nu - 2)(\nu - 4)]^2} \left[(\nu - 2)^2(\nu - 4) + \frac{16\beta^4\delta^2(\nu - 2)(\nu - 4)}{(\nu - 6)} + \frac{8\beta^4\delta^4(5\nu - 22)}{(\nu - 6)(\nu - 8)} \right].$$

A assimetria e a curtose não existem quando $\nu < 6$ e $\nu < 8$ respectivamente. Depois de especificada a distribuição e suas medidas principais, abaixo são apresentadas a definição de cada parâmetro envolvido na *f.d.p.* como discutido no artigo de Teixeira (2006) [3].

2.7.3 Value at risk

O risco é inevitável num mercado financeiro. Há vários tipos de risco como: risco de mercado, risco de crédito, risco de liquidez, risco operacional, risco legal *etc.* Escolher qual estratégia seguir, é também escolher qual risco ficar exposto. Afim de resolver essa problemática, em 1980 Till Guldman, um dos pesquisadores do J.P. Morgan, criou o *value at risk* no qual foi apresentado para o G-30 e publicado em 1993.

Definição 2.7.1 *O Value at risk (ou VaR) sintetiza a maior (ou pior) perda esperada dentro de determinados períodos de tempo e intervalo de confiança.*

A ideia do cálculo do *value at risk* é encontrar o percentil da distribuição de retornos analisada num dado horizonte, o VaR corresponderá a $(1 - c)$ percentil da distribuição. Por exemplo, consideremos um nível de confiança de $100(1 - c)\%$, o VaR é tal que excede 5% do número total da amostra como tratado no livro de Jorion (2003) [17].

As etapas de cálculo do VaR seguem os seguintes procedimentos:

- Marcar a mercado a atual carteira, isto é, ver a quanto está o preço do ativo;
- Medir a variabilidade de fatores de risco, por exemplo, o quanto varia ao ano;
- Determinar o horizonte temporal;
- Determinar o nível de confiança;
- Reportar a pior perda.

Aqui, será tratado a carteira como um conjunto de ativos financeiros. E de uma carteira qualquer é calculado o VaR. Note que, a carteira pode ter apenas um ativo (que é o caso em questão, pois para cada ativo financeiro que estiver sendo usado aqui será calculado separadamente o VaR). Em geral, a forma genérica do VaR pode ser obtido derivando a distribuição de probabilidade do valor futuro da carteira. Usando um determinado nível de confiança, deseja-se achar a pior realização possível desta carteira, ou seja, encontra a função acumulativa até um valor W^* . O cálculo segue abaixo:

$$c = \int_{W^*}^{\infty} f(w)dw$$

Ou então, uma outra forma de pensar no cálculo da VaR é usar a probabilidade de um valor menor que o W^* , $p = P(w \leq W^*) = p$, ou seja, calcular esta probabilidade significa dizer que a área de $-\infty$ à W^* deve somar $p = 1 - c$. O valor referente a W^* é denominado como quantil da distribuição a ser calculada e corresponde assim a um valor de corte cuja a probabilidade de se exceder ao valor W^* é fixa. Afim de calcular o VaR suposições são feitas nos dados dos preços dos ativos financeiros como:

- Independência, ou seja o valor de hoje não é afetado pelo valor de ontem ou de t dias atrás;
- Identicamente distribuídos, isto é, possuem a mesma distribuição e com o mesmo valor assumido pelo parâmetro. Porém, não necessariamente assumem o mesmo ponto.

O VaR tem muita utilização no mercado financeiro, e, de acordo com a finalidade a ser usada o VaR pode ser calculado para:

- Medida de *Benchmark*;
- Perda Potencial;
- Medida de Capital.

O escopo deste trabalho será lidar com o VaR como Perda Potencial. Ao determinar o horizonte de tempo da carteira, torna-se possível fornecer uma ideia geral da pior perda que esta empresa poderá sofrer. Uma primeira interpretação deste método é de que o horizonte de tempo seja definido pelo período de liquidação. Uma segunda interpretação consiste em considerar que o horizonte deve refletir o tempo essencial para proteger (*hed-gear*) os riscos de mercado. Uma terceira visão sobre o VaR é de que o horizonte deve

corresponder ao período sobre o qual a carteira permanece relativamente constante, esta ideia de interpretação é oposta as demais e este tipo de interpretação pode perder seu significado à medida que o horizonte se estende.

Para esta aplicação, geralmente considera valores diários da carteira. E, a escolha do nível de confiança é relativamente arbitrária. **Deve-se atentar para o fato que o VaR não descreve a pior perda possível, porém o VaR denota uma medida probabilística que deve ser excedida com alguma frequência.** E, note que maiores níveis de confiança acarretarão em maiores medidas de VaR.

2.7.4 Teste de *Kupiec*

Para a verificação de qual método foi o melhor para ser aplicado, é conveniente o uso de testes de comparação para a escolha do melhor método através do VaR. Um dos testes mais usuais é o chamado Teste de Kupiec (1995), seu uso serve para verificar a validação de um modelo para o cálculo do VaR, se realmente o VaR teve uma perda de $(1 - \alpha)\%$ estabelecida. A ideia do método consiste em utilizar exatamente a frequência de vezes em que o VaR foi extrapolado. Gerando intervalos para a frequência de extrapolações e determinando o nível de confiança de um dado ativo ou *portfólio* é definido uma regra de decisão para aceitar ou rejeitar o modelo a ser testado. Torna-se importante observar que a escolha do nível de significância para o teste não tem relação com o proposto para o cálculo do VaR, como mencionado em Goulart (2009) [19].

Para o cálculo do teste de Kupiec é usado as seguintes hipóteses:

$$\begin{cases} H_0 & : \theta = \alpha \\ H_1 & : \theta \neq \alpha \end{cases}$$

Com estatística de teste, sob H_0 , dada por:

$$\Lambda(V) = -2 \ln \frac{L(\alpha|V)}{\sup_{\theta} L(\theta|V) : \theta \in [0, 1]} \sim \chi^2(1)$$

$$\Lambda(V) = \begin{cases} -2 \ln \left(\frac{\alpha^V (1-\alpha)^{(T-V)}}{\hat{\alpha}^V (1-\hat{\alpha})^{(T-V)}} \right), & \text{se } V > 0 \\ -2 \ln((1-\alpha)^T), & \text{se } V = 0 \end{cases},$$

onde $L(\cdot)$ é a função de verossimilhança, $\hat{\alpha} = \frac{V}{T}$ é o estimador de máxima verossimilhança de α .

3 Resultados

3.1 Séries utilizadas na modelagem

O banco de dados, para as três variáveis em análise, está disponível no site <http://kuestertech.com/ktstock/?lang=pt-BR> de análise técnica do mercado de ações no dia 15 de Junho de 2015. Os dados são de cotações diárias, já ajustadas, do mercado brasileiro. A variável escolhida para análises foi o *X.Close* (Fechamento) que trata do último valor negociado da ação num determinado dia, considerando o preço em reais. Para a variável *PETR4* os dados tem início no dia 02 de Janeiro de 2004 e terminam no dia 12 de Junho de 2015 com total de 4812, sendo as últimas 316 observações usadas como dados para verificação da eficiência da previsão. A variável *VALE5* teve seu início em 01 de Junho de 2010 e término na data de 12 de Junho de 2015 com total de 1247, sendo as últimas 277 observações usadas como dados para verificação da eficiência da previsão. E, para o Índice Bovespa a série teve início no dia 16 de Setembro de 2005 e finalizou no dia 12 de Junho de 2015 com total de 2405, sendo as últimas 357 observações usadas como dados para verificação da eficiência da previsão. A ideia de restringir a este período de tempo foi devido ao fato de que, para dados em períodos de tempos mais extremos, teríamos problemas como sazonalidade.

As análises consistiram em dois procedimentos. No primeiro procedimento as séries financeiras foram modeladas utilizando o processo ARMA primeiro e após foram modeladas com o processo GARCH e consideremos os resíduos seguindo distribuição normal. O primeiro método foi uma iniciativa de descobrir como o modelo se comporta para que no segundo método pudesse fazer análises otimizadas.

Já para o segundo método, usando uma função do *software R*, as séries foram modeladas com ARMA e GARCH ao mesmo tempo. E, através do segundo método, foi calculado os coeficientes da distribuição GHsT, afim de obter o VaR para valores futuros. Assim, como primeiro passo, foi feito o teste BDS para verificar se as séries de dados precisariam de modelagem ou não. Os resultados indicam que é preciso a modelagem em todas as

três séries.

3.2 Resultados preliminares

Primeiramente foi feito o gráfico das séries originais para verificar o comportamento das séries. Como podemos observar, há uma leve tendência em todas as séries. Como apresentado na seção 2.1. E, com isso foi aplicado a primeira diferença juntamente com o logaritmo neperiano a fim de tornar as séries estacionárias. Seguindo, foi usado o correlograma para a série de retornos para as três séries de dados.

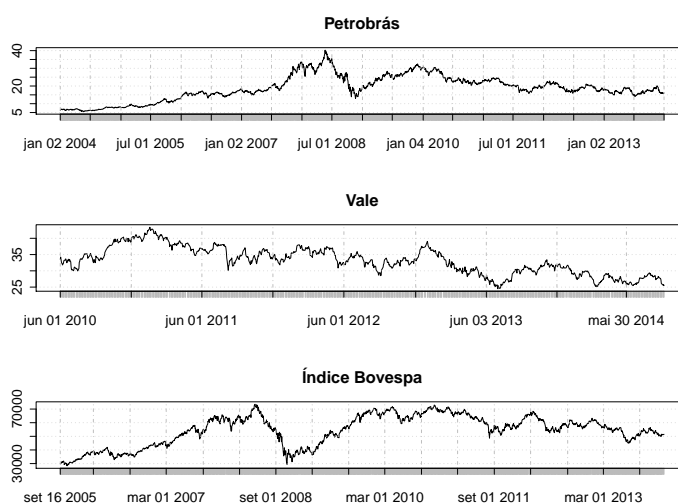


Figura 4: Gráfico das séries PETR4, VALE5 e Índice Bovespa, considerando os valores em reais.

No correlograma dos dados da Petrobrás, é possível verificar no gráfico de autocorrelação um comportamento senoidal o que nos leva a ideia de um auto-regressivo e ao olhar o gráfico da função de autocorrelação parcial é visível ver que os picos que passam o intervalo vai até a *lag* 18. Já para a série de dados da Vale é possível ver no gráfico de autocorrelação senoidal assim como no gráfico de autocorrelação parcial, picos significativos até a defasagem 17. Em seguida, para os dados da série Índice Bovespa a conclusão que se chega é de um auto-regressivo pelo comportamento senoidal e decaimento exponencial no FAC e olhando as defasagens, podemos ver que até a *lag* 20.

3.3 Modelagem e diagnóstico

As análises consistiam em dois procedimentos. No primeiro procedimento as séries financeiras foram modeladas utilizando o processo ARMA e GARCH separadamente,

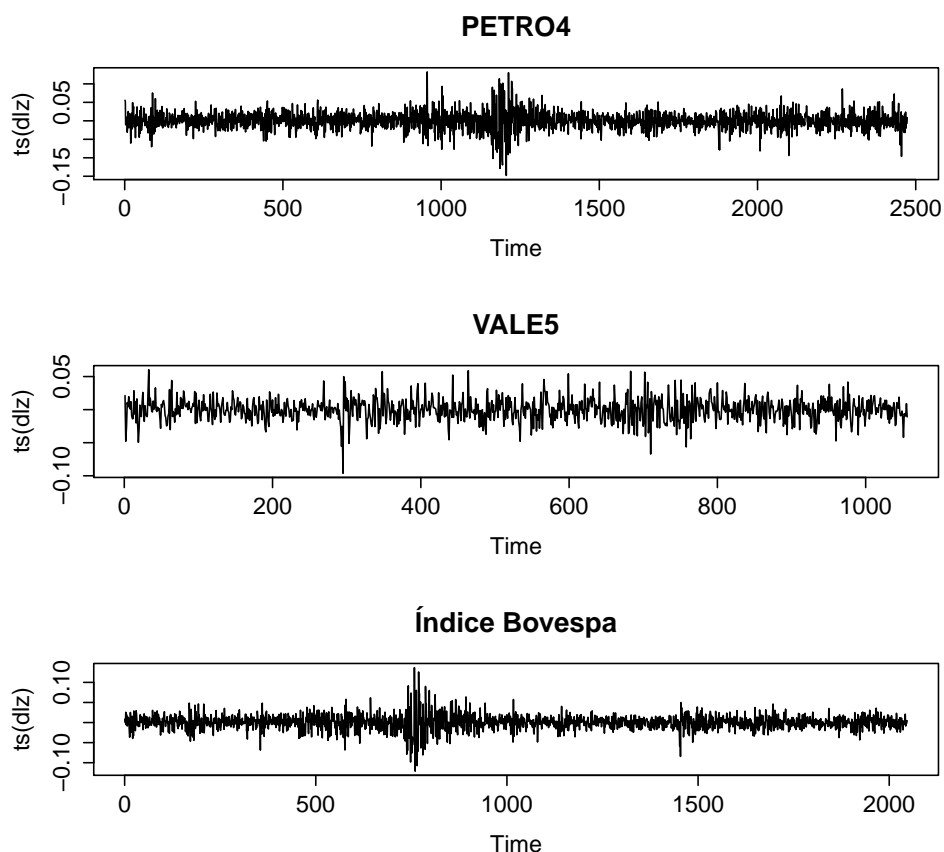


Figura 5: Gráficos das séries Petrobrás (PETR4), Vale (VALE5) e Índice Bovespa (IBOV11) da primeira diferença do logaritmo neperiano.

como modo de otimização. E no segundo procedimento foi feito a modelagem em conjunto. E dando sequência, os resultados dos coeficientes, para estes três métodos seguem nas tabelas 2, 3 e 4. Note que na tabela 4, para o segundo tipo de modelagem foi considerado $ARMA(0,0)$.

PETR4 - BDS Test

data: z

p-value =	[3.5375]	[7.0751]	[10.6126]	[14.1502]
[2]	0	0	0	0
[3]	0	0	0	0

VALE5 - BDS Test

data: z

p-value =	[2.0722]	[4.1445]	[6.2167]	[8.2889]
[2]	0	0	0	0
[3]	0	0	0	0

Índice Bovespa - BDS Test

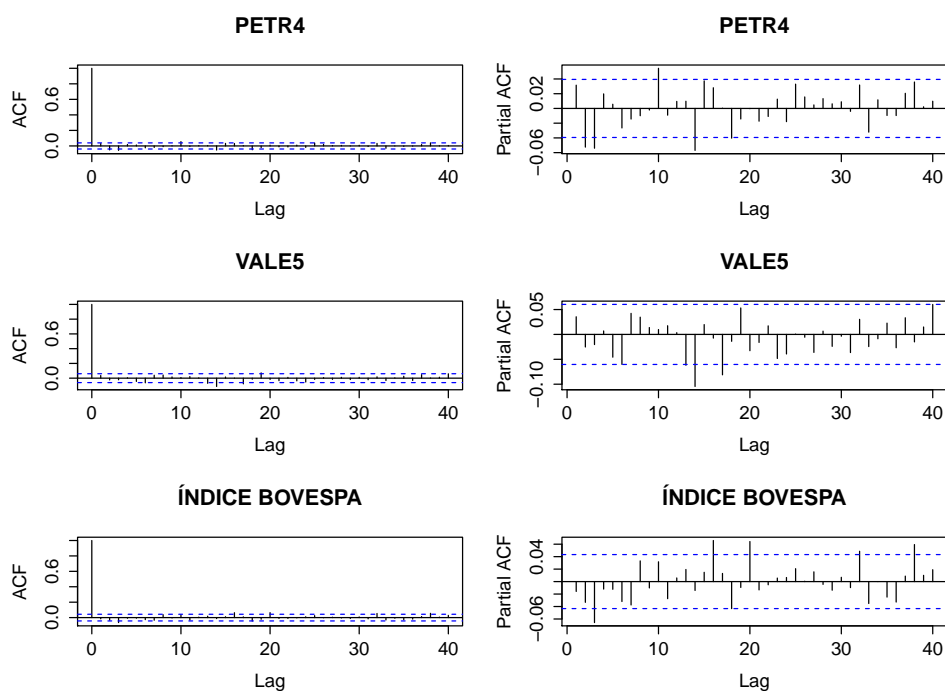


Figura 6: Autocorrelações e autocorrelações parciais amostrais para os dados de retorno das séries PETR4, VALE5 e Índice Bovespa.

```

data: z
p-value = [ 5574.9519 ] [ 11149.9038 ] [ 16724.8557 ] [ 22299.8076 ]
[ 2 ]          0          0          0          0
[ 3 ]          0          0          0          0

```

Tabela 2: PETR4 - Tabela dos valores dos coeficientes estimados para os dois métodos de estimação dos modelos, considerando os resíduos seguindo uma distribuição Normal.

1°					2°	
$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\phi}_{10}$	$\hat{\phi}_{14}$	$\hat{\phi}_{18}$	$\hat{\phi}_3$	$\hat{\phi}_{10}$
-0,0475	-0,0536	0,0558	-0,0531	-0,0447	-0,0409	0,0459

Tabela 3: VALE5 - Tabela dos valores dos coeficientes estimados para os dois métodos de estimação dos modelos, considerando os resíduos seguindo uma distribuição Normal.

1°		2°	
$\hat{\phi}_{14}$	$\hat{\phi}_{17}$	$\hat{\phi}_{14}$	$\hat{\phi}_{17}$
-0,1157	-0,0807	-0,1250	-0,0754

A tabela 8 mostra o ajuste das três séries de dados modeladas pelos dois métodos, onde Desv.Pad. é o erro padrão. É possível perceber a menor quantidade de parâmetros

Tabela 4: Índice Bovespa - Tabela dos valores dos coeficientes estimados pelos dois métodos de estimação dos modelos, considerando os resíduos seguindo uma distribuição Normal.

	1°		2°
	$\hat{\phi}_3$	$\hat{\phi}_{16}$	$\hat{\phi}_{20}$
	-0,0637	0,0656	0,0663

Tabela 5: Ajuste do modelo, com os dois tipos de modelagem aplicadas nas três séries de dados analisadas, considerando os resíduos seguindo uma distribuição Normal.

Dados	Método	Modelo ARMA(p,q)	Valores			
			Coefficiente	Desv. Pad.	Estatística t	p-valor
PETR4	1°	AR(18)	-0,0447	0,0200	-2,2291	0,0258
	2°	AR(10)	0,0458	0,0243	2,2487	0,0245
VALE5	1°	AR(17)	-0,0807	0,0307	-2,6247	0,0088
	2°	AR(17)	-0,0754	0,0304	-2,4674	0,0136
Índice Bovespa	1°	AR(20)	0,0663	0,0220	3,0119	0,0026
	2°	AR(0)	—	—	—	—

no segundo método de modelagem. Para a verificação da ausência de autocorrelação nos resíduos dos modelos escolhidos, foi feito o gráfico do correlograma dado pela figura 7 verificando que todas as defasagens estão dentro do intervalo, ou seja, as séries foram corretamente modeladas pelo primeiro método de modelagem. Com o intuito de analisar se o modelo escolhido é um bom modelo foi usado o teste de Box-Pierce. Nas tabelas 6 e 11, todos os p-valores foram maiores do que o nível de significância definido como $\alpha = 0,05$. Assim, a conclusão que se chega é que os modelos foram bem estimados. Contudo, ao se obter os resíduos ao quadrado a conclusão que se chega, através da figura 8, obtido pelo primeiro método, é da existência ainda da correlação nos resíduos dos modelos. Com o intuito de verificar se a análise do gráfico 8 está correta, foi feito o teste BDS para a verificação da necessidade de modelar com processos GARCH. Como resultado, para as três séries de dados analisadas foi a necessidade de modelagem fazendo uso dos modelos não lineares.

PETR4 - primeiro método - BDS Test

data: res

p-value = [0.0115] [0.023] [0.0345] [0.046]

[2] 0 0 0 0

[3] 0 0 0 0

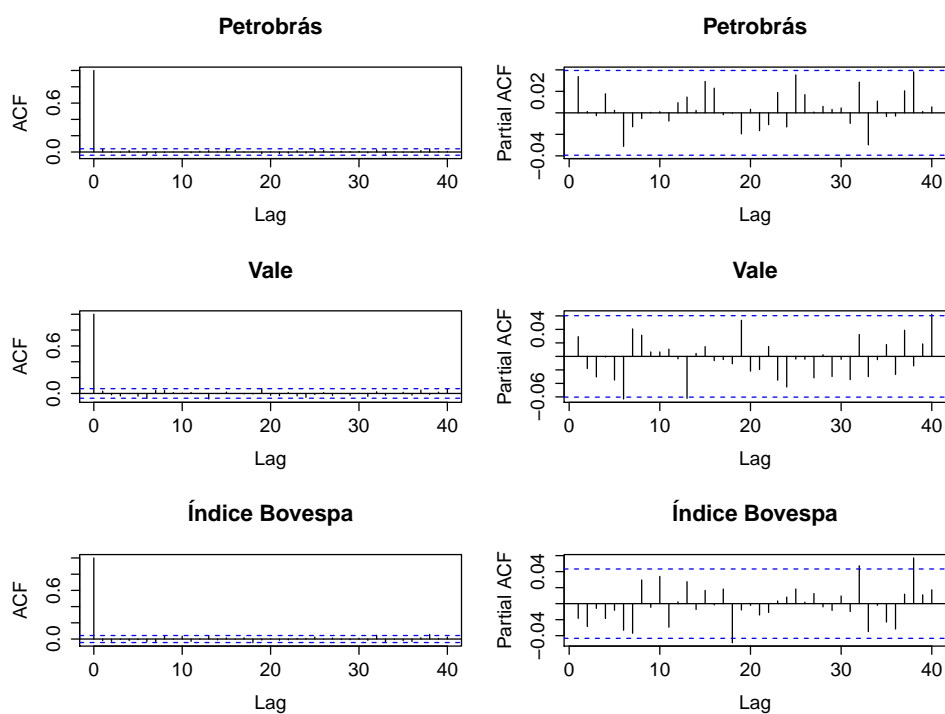


Figura 7: Correlograma dos resíduos dos modelos para as séries de dados dos retornos das PETR4, VALE5 e Índice Bovespa usando o primeiro método de modelagem.

Tabela 6: Primeiro método - Tabela dos p-valores do teste de Box-Pierce para os modelos estimados dos dados de retorno das séries PETR4, VALE5 e Índice Bovespa.

Dados	Valores			
	Lag	χ^2	graus de liberdade	P-valor
PETR4	12	7,0554	7	0,4231
	24	14,7403	19	0,7390
	36	23,3964	31	0,8344
	48	45,2821	43	0,3769
VALE5	12	11,2276	10	0,3401
	24	26,0358	22	0,2501
	36	30,6633	34	0,6320
	48	47,4172	46	0,4146
Índice Bovespa	10	12,3397	8	0,1367
	20	21,5575	18	0,2522
	30	24,4619	28	0,6570
	40	43,5910	38	0,2457

VALE5 - primeiro método -BDS Test

data: res

p-value =

[0.0086] [0.0172] [0.0257] [0.0343]

Tabela 7: Segundo método - Tabela dos p-valores do teste de Box-Pierce para os modelos estimados dos dados dos retornos para as séries PETR4, VALE5 e Índice Bovespa.

Dados	Valores			
	Lag	χ^2	graus de liberdade	P-valor
PETR4	10	9,0064	8	0,3418
	20	17,3365	18	0,5001
	30	24,4965	28	0,6551
	40	32,2852	38	0,7305
VALE5	10	11,2313	8	0,1889
	20	20,4068	18	0,3104
	30	27,6650	28	0,4823
	40	35,6119	38	0,5804
Índice Bovespa	10	5,2029	10	0,8772
	20	16,1343	20	0,7083
	30	19,6388	30	0,9256
	40	30,4413	40	0,8626

[2] 0.2437 0.1014 0.0589 0.0717

[3] 0.0365 0.0060 0.0030 0.0030

Índice Bovespa - primeiro método - BDS Test

data: res

p-value =

[0.0092] [0.0184] [0.0276] [0.0368]

[2] 0 0 0 0

[3] 0 0 0 0

Seguindo, a modelagem com modelos não lineares tiveram dois procedimentos, no primeiro método, considerou $a_t \sim N(0, 1)$ e no segundo segundo método como $a_t \sim N(0, 1)$ calculando depois os coeficientes da $GHS(μ, δ, β, v)$. E, no teste de Box-Pierce, para

Tabela 8: Ajuste de modelos não lineares, com os dois tipos de métodos aplicados nas três séries de dados analisadas, considerando os resíduos seguindo distribuição Normal.

Dados	Método	Modelo
PETR4	1º	GARCH(2,3)
	2º	GARCH(2,3)
VALE5	1º	GARCH(1,1)
	2º	GARCH(1,1)
Índice Bovespa	1º	GARCH(1,1)
	2º	GARCH(2,2)

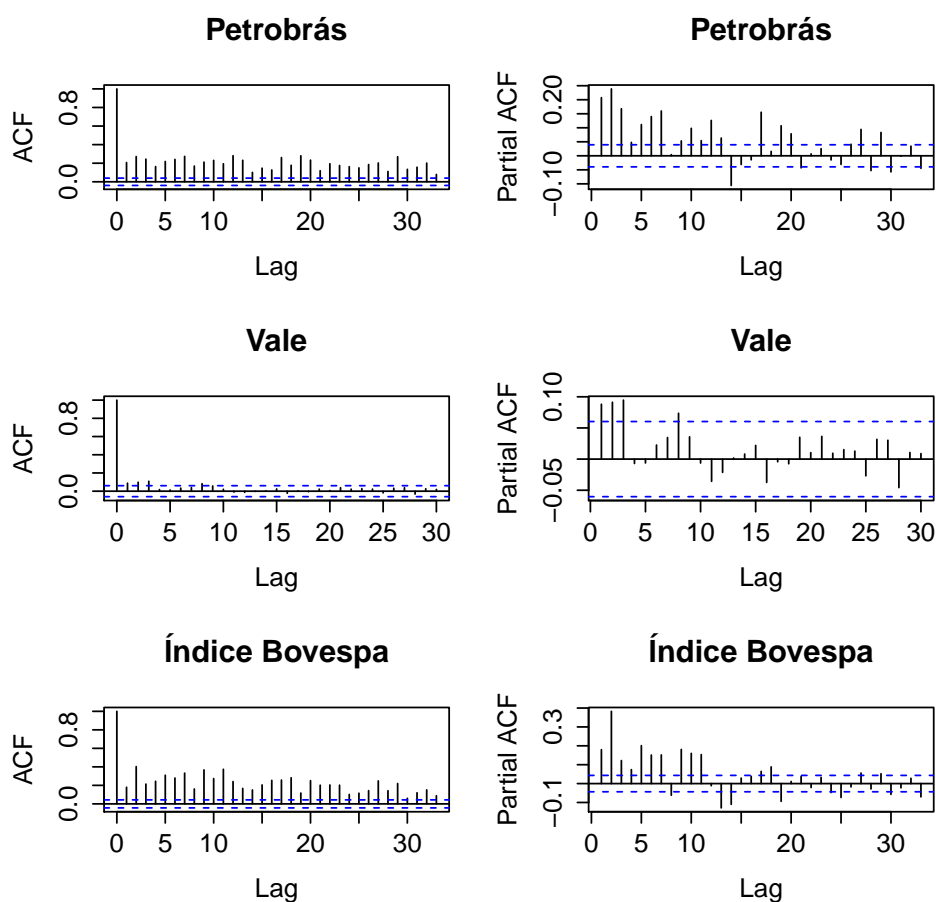


Figura 8: Correlograma dos resíduos ao quadrado dos modelos para as séries de dados dos retornos das séries PETR4, VALE5 e Índice Bovespa considerando o primeiro método de modelagem.

Tabela 9: Segundo método - Tabela dos valores dos coeficientes estimados dos modelos não lineares dos dados de retornos das séries PETR4, VALE5 e Índice Bovespa.

Dados	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
PETR4	0,0001	0,03602	0,0960	0,5816	0	0,2315
VALE5	0,0000	0,06885	0	0,8478	0	0
Índice Bovespa	0,0000	0,0355	0,0979	0,3713	0	0,4550

modelos não lineares também chegou a conclusão de que os coeficientes foram bem escolhidos para o modelo proposto das três séries de dados. Para a verificação de que a modelagem se limita até os modelos não lineares, foi feito o teste BDS com *bootstrap*, os resultados, para os dois métodos de modelagem, foi que as séries foram corretamente modeladas.

Tabela 10: Primeiro método - Tabela dos p-valores do teste de auto-correlação dos modelos não lineares propostos para os dados de retorno das séries PETR4, VALE5 e Índice Bovespa.

Dados	Valores			
	Lag	χ^2	graus de liberdade	P-valor
PETR4	10	6,6614	4	0,1549
	20	21,4111	14	0,0916
	30	33,8128	24	0,0881
	40	40,9742	34	0,1912
VALE5	10	7,0547	7	0,4232
	20	12,2376	17	0,7855
	30	18,5688	27	0,8852
	40	27,1713	37	0,8817
Índice Bovespa	10	13,2071	7	0,0672
	20	20,4456	17	0,2521
	30	26,8198	27	0,4734
	40	39,6270	37	0,3536

Tabela 11: Segundo método - Tabela dos p-valores do teste de auto-correlação dos modelos não lineares propostos dos dados de retorno das séries PETR4, VALE5 e Índice Bovespa.

Dados	Valores			
	Lag	χ^2	graus de liberdade	P-valor
PETR4	10	4,2132	4	0,3779
	20	19,7211	14	0,1392
	30	33,7386	24	0,0894
	40	40,7994	34	0,1963
VALE5	10	11,2313	8	0,1889
	20	20,4068	18	0,3104
	30	27,6650	28	0,4823
	40	35,6119	38	0,5804
Índice Bovespa	10	6,3741	5	0,2715
	20	14,6659	15	0,4757
	30	21,2591	25	0,6781
	40	36,7962	35	0,3857

PETR4 - primeiro método - BDS Test

data: res

p-value = [0.0115] [0.023] [0.0345] [0.046]

[2] 0 0 0 0

[3] 0 0 0 0

PETR4 - segundo método - BDS Test

```

data: data
p-value = [ 0.5007 ] [ 1.0013 ] [ 1.502 ] [ 2.0027 ]
[ 2 ]      0.880      0.510      0.278      0.216
[ 3 ]      0.676      0.454      0.202      0.142
[ 4 ]      0.464      0.240      0.126      0.108
[ 5 ]      0.412      0.250      0.118      0.132
VALE5 - primeiro método - BDS Test
data: res
p-value =
      [ 0.0086 ] [ 0.0172 ] [ 0.0257 ] [ 0.0343 ]
[ 2 ]      0.2437      0.1014      0.0589      0.0717
[ 3 ]      0.0365      0.0060      0.0030      0.0030
VALE5 - segundo método - BDS Test
data: data
p-value = [ 0.5026 ] [ 1.0051 ] [ 1.5077 ] [ 2.0103 ]
[ 2 ]      0.602      0.740      0.904      0.876
[ 3 ]      0.772      0.970      0.846      0.808
[ 4 ]      0.746      0.806      0.676      0.730
[ 5 ]      0.586      0.868      0.760      0.952
Índice Bovespa - primeiro método - BDS Test
data: res
p-value =
      [ 0.0092 ] [ 0.0184 ] [ 0.0276 ] [ 0.0368 ]
[ 2 ]          0          0          0          0
[ 3 ]          0          0          0          0
Índice Bovespa - segundo método - BDS Test
data: data
p-value = [ 0.4999 ] [ 0.9998 ] [ 1.4997 ] [ 1.9996 ]
[ 2 ]      0.278      0.204      0.484      0.890
[ 3 ]      0.424      0.268      0.576      0.968
[ 4 ]      0.746      0.458      0.796      0.818
[ 5 ]      0.994      0.534      0.788      0.842

```

Mas, ao se realizar o teste de *Jarque-Bera* para a verificação de normalidade no segundo método de modelagem mostrado na tabela 13 e considerando um nível de significância de $\alpha = 0,05$, foi rejeitada a hipótese nula assim considerando que os resíduos

não seguem distribuição normal. Com isso este resultado, calculou-se os coeficientes da GHsT e após a volatilidade e o VaR (para um dia de previsão) baseados nesta distribuição abordados nas tabelas 14 e 15.

Tabela 12: Teste de *Jarque-Bera* para o primeiro método dos dados dos retornos das séries analisadas.

Dados	P-valor	Estatística de Teste
PETR4	$2,200 \times 10^{-16}$	152,1486
VALE5	$3,9670 \times 10^{-10}$	38,6388
Índice Bovespa	$2,200 \times 10^{-16}$	97,7276

Tabela 13: Teste de *Jarque-Bera* para o segundo método dos dados dos retornos das séries analisadas.

Dados	P-valor	Estatística de Teste
PETR4	$2,200 \times 10^{-16}$	152,1486
VALE5	$4,071 \times 10^{-9}$	43,2956
Índice Bovespa	$2,200 \times 10^{-16}$	97,7276

3.4 Previsão

Tabela 14: Coeficientes da GHsT para os dados de retornos das séries PETR4, VALE5 e Índice Bovespa, para um dia e considerando o segundo método de modelagem.

Dados	Coeficientes			
	μ	δ	n	β
PETR4	0,0049	0,0545	9,2841	-10,0725
VALE5	-0,0011	0,0666	7,9974	0,8421
Índice Bovespa	0,0030	0,0325	10,1707	-23,1491

Tabela 15: Previsão usando a GHsT para os dados de retornos das séries PETR4, VALE5 e Índice Bovespa, para um dia e considerando o segundo método de modelagem.

Dados	VaR
PETR4	15,53
VALE5	17,11
Índice Bovespa	50530,50

Dando prosseguimento as análises, foram reservadas observações, com tamanhos distintos para cada série analisada, com o intuito de verificar a eficiência do cálculo do VaR baseado nos modelos escolhidos. Os procedimentos foram criar um algoritmo para calcular a previsão para um dia a frente e após recalculiar os parâmetros baseados no novo valor, que de uma previsão foi comparado com a observação deixada para este intuito. Assim, o cálculo do VaR foi feito apenas para os modelos construídos através do segundo método de modelagem, por se tratar de modelos mais precisos. E em seguida foram comparados os modelos através do vetor de VaR de cada série de dados usando o teste de *Kupiec*, consideremos os resíduos com distribuição Normal e com distribuição GHsT. Os resultados seguem abaixo.

Tabela 16: P-valores do Teste de Kupiec considerando os modelos com os resíduos seguindo distribuição Normal e seguindo distribuição GHsT com o *VaR gain*.

Dados	Distribuição	P-valor	Sucesso	Fracasso
PETR4	Normal	0,2993	20	296
PETR4	GHsT	0,2993	20	296
Vale5	Normal	0,0000	44	233
Vale5	GHsT	0,0000	48	229
Índice Bovespa	Normal	0,8075	16	341
Índice Bovespa	GHsT	0,9035	18	339

Tabela 17: P-valores do Teste de Kupiec considerando os modelos com os resíduos seguindo distribuição Normal e seguindo distribuição GHsT com o *VaR loss*.

Dados	Distribuição	P-valor	Sucesso	Fracasso
PETR4	Normal	0,3674	19	297
PETR4	GHsT	0,3674	19	297
Vale5	Normal	0,0000	33	244
Vale5	GHsT	0,0000	34	243
Índice Bovespa	Normal	0,3967	14	343
Índice Bovespa	GHsT	0,3967	14	343

Como é possível observar na tabela 17, para a série de dados da Petrobrás, a modelagem com o uso da Normal e da GHsT deu o mesmo resultado de eficiência. Todavia é válido ressaltar que para dados mais robustos ou oscilações mais bruscas no mercado financeiro, o uso da GHsT se torna um diferencial na modelagem. Para a série de dados da Vale, o teste acusou que a modelagem não foi válida para ambos os casos, porém o maior dos p-valores foi através do uso da Normal. Já para a série de dados do Índice Bovespa, o resultado do teste de hipóteses resultou numa maior eficácia da GHsT. E a seguir foi feito os gráficos para o vetor de VaR, considerando os dados de retornos do

Índice Bovespa. As linhas em azul são as indicações de VaR Loss e VaR Gain e as bolas em verde são os valores que ultrapassaram o VaR Gain e consequentemente as bolas em vermelho são os pontos que estão abaixo do VaR Loss.

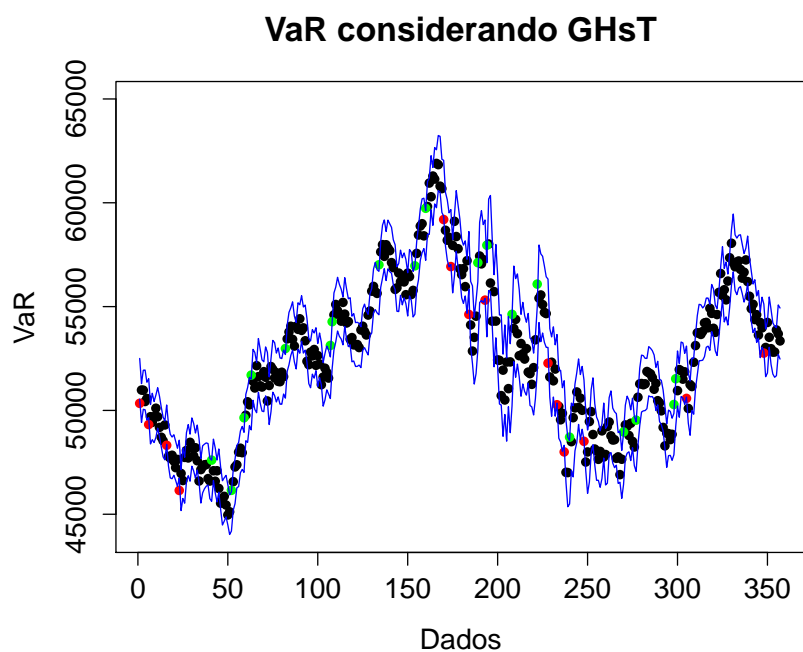


Figura 9: Gráfico dos valores de VaR Loss e o VaR Gain para o Índice Bovespa considerando a GHsT.

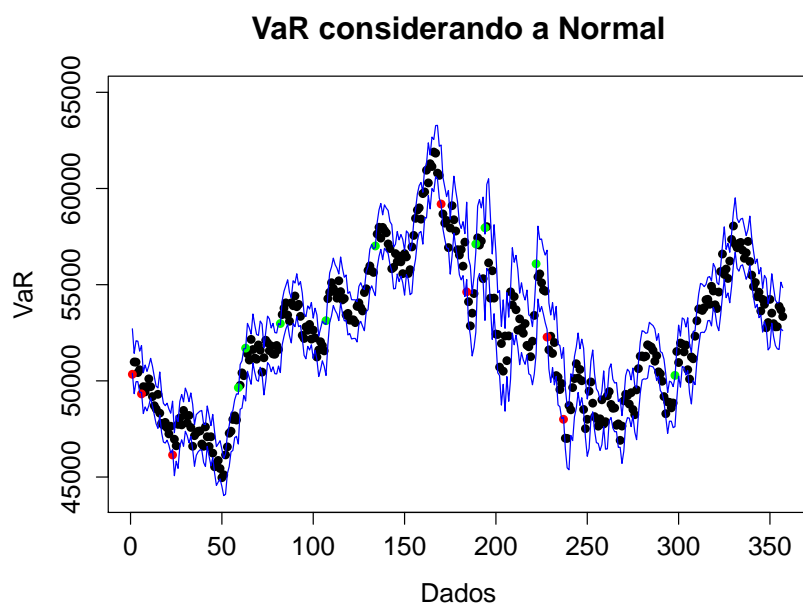


Figura 10: Gráfico dos valores de VaR Loss e o VaR Gain para o Índice Bovespa considerando a Normal.

4 Conclusão

Esse trabalho verifica qual o melhor modelo se ajusta aos dados dos valores de fechamento diário dos ativos da Vale e da Petrobrás e o Índice Bovespa. Para isso, foi feita a modelagem usando dois procedimentos: modelagem com ARMA e GARCH separadamente e em conjunto. E em paralelo a modelagem foi utilizado testes de diagnósticos dos modelos para verificar o uso correto da modelagem. E, com isso a próxima etapa foi de calcular o VaR para um dia a frente. E daí, criar o vetor de VaR para várias observações e verificando quantos sucessos e falhas se ibteve na modelagem de cada série de dados.

Uma das principais conclusões a tirar é que a distribuição Normal, só foi válida para os dados da Petrobrás. A outra importante conclusão a ser feita é a dificuldade em se obter os parâmetros da GHsT devido a problemas como convergência. Mas, com relação a eficácia, a GHsT supera a Normal em todos os aspectos (previsão, por exemplo). Com relação ao ativo da Petrobrás, foi constatado a equivalência da modelagem para as duas distribuições. Contudo, a GHsT ainda se mostra superior quando os dados possam ter maior oscilação no mercado devido a robustes da distribuição. Note que neste ativo foi usado os modelos ARMA(10,0) e GARCH(2,3), para o segundo método de modelagem.

Para a série de dados da Vale foi obtido um resultado de que a modelagem não se aplica em ambos os casos. Porém, ainda assim, temos que o p-valor do modelo considerando a Normal foi maior que o modelo considerando a GHsT, tendo uma exceção neste ativo. Adicionalmente, neste ativo foi usado os modelos ARMA(17,0) e GARCH(1,1), com relação ao segundo método de modelagem. Já para o Índice Bovespa, o resultado foi como o esperado. A distribuição GHsT se mostrou mais adequada para a modelagem dos dados. Possuindo melhor previsibilidade para as observações. Ainda neste ativo foi usado os modelos ARMA(0,0) e GARCH(2,2), pelo segundo método de modelagem.

Comparando os resultados encontrados com o artigo de Aas et al. (2006) [5], pode-se perceber que a distribuição GHsT é a melhor escolha para uma modelagem mais precisa, porém esta conclusão não pode ser generalizada. Os resultados obtidos aqui permitiram

verificar que a Normal também pode ser usada na modelagem e obter bons resultados, porém a GHsT pode se adequar melhor a dados futuros. E, relacionando este teste de Kupiec apresentado aqui com o teste de Kupiec verificado no artigo de Fajardo et al. (2004) [8], foi obtida a verificação da precisão do modelo em termos do cálculo do VaR.

5 Trabalhos futuros

Os valores obtidos e as conclusões retiradas são condicionadas à qualidade do *software* utilizado. No caso, foi utilizado os algoritmos que já estavam implementados no *software R* e também foram criados algoritmos para melhorar a eficiência da modelagem e dos testes (todos implementados no *software R*), contudo existe sempre a possibilidade de melhorar os algoritmos de forma a obter melhores estimações dos parâmetros. Uma sequência natural deste trabalho seria desenvolver a área de programação no que diz respeito automatização das implementações feitas. Assim, a ideia para a automatização seria programar todas as análises feitas de forma automática e de modo que os resultados já fossem dados em uma planilha em excel.

Logo, um trabalho futuro seria o uso de outras séries financeiras com intuito de modelá-las usando a distribuição hiperbólica generalizada assimétrica T-student (GHsT) como modelo para os log retornos dos preços de outros ativos do mercado financeiro brasileiro através da automatização. E, além disso, poder usar estas análises para mercados não só de ativos como também de futuros, por exemplo.

Referências

- [1] COMISSÃO de valores imobiliários. Programa de Treinamento para Professores. Mercado de valores mobiliários brasileiro, 3 th. ed. p. 376, 2014. Disponível em: <<http://www.investidor.gov.br>>.
- [2] PRAUSE, K. *The generalized hyperbolic model: Estimation, financial derivatives, and risk measures*. Tese (Doutorado) — PhD thesis, University of Freiburg, 1999.
- [3] TEIXEIRA, V. M. M. Distribuições hiperbólicas generalizadas: aplicações ao mercado português. Universidade da Madeira, 2006.
- [4] SILVA, A. d. C. Mercados incompletos, distribuições hiperbólicas e aplicações para o caso brasileiro. 1999.
- [5] AAS, K.; HAFF, I. H. The generalized hyperbolic skew student's t-distribution. *Journal of financial econometrics*, Oxford Univ Press, v. 4, n. 2, p. 275–309, 2006.
- [6] EBERLEIN, E.; KELLER, U. Hyperbolic distributions in finance. *Bernoulli*, JSTOR, p. 281–299, 1995.
- [7] ZHOU, J. Modeling S&P 500 stock index using arima asymmetric power arch models. *Hogskolan Dalarna - Monografia de Mestrado, Escola de Economia e Ciências Sociais*, 2009.
- [8] FAJARDO, J.; FARIAS, A. Generalized hyperbolic distributions and brazilian data. *Brazilian review of econometrics*, v. 24, n. 2, p. 249–271, 2004.
- [9] KIM, Y. I.; MCCULLOCH, J. H. The skew-student distribution with application to us stock market returns and the equity premium. *Journal of Financial Econometrics*, 2007.
- [10] STEFFEN, D. Intervalo "bootstrap" para previsões de séries temporais obtidas pelo método theta. 2010.
- [11] ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, The Econometric Society, v. 50, n. 4, p. pp. 987–1007, 1982. ISSN 00129682. Disponível em: <<http://www.jstor.org/stable/1912773>>.
- [12] BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, v. 31, n. 3, p. 307–327, 1986. ISSN 0304-4076. Disponível em: <[http://dx.doi.org/10.1016/0304-4076\(86\)90063-1](http://dx.doi.org/10.1016/0304-4076(86)90063-1)>.
- [13] MORETTIN, P.; TOLOI, C. de C. *Análise de séries temporais*. Edgard Blucher, 2006. (ABE - Projeto Fisher). ISBN 9788521203896. Disponível em: <<http://books.google.com.br/books?id=Q7bJAAAACAAJ>>.

- [14] EHLERS, R. S. *Análises de séries temporais*. 5. ed. [S.l.: s.n.], 2009.
- [15] BROCKWELL, P.; DAVIS, R. *Introduction to Time Series and Forecasting*. Springer, 2002. (Introduction to Time Series and Forecasting, v. 1). ISBN 9780387953519. Disponível em: <<http://books.google.com.br/books?id=VHB40SAmwcUC>>.
- [16] BITENCOURT, W. A. Ensaio empírico sobre a eficiência do mercado futuro de café. Lavras, MG (Brazil), 2007.
- [17] JORION, P. *Value at Risk: A nova fonte de referência para a gestão do risco financeiro*. McGraw-Hill, 2003. (MacGraw-Hill international editions: Finance series). ISBN 9780071228312. Disponível em: <<http://books.google.com.br/books?id=0E1IPAAACAAJ>>.
- [18] STEVENSON, W. J. *Estatística aplicada a administração*. [S.l.: s.n.], 2001.
- [19] GOULART, P. M. Avaliação da metodologia value-at-risk aplicada à análise da volatilidade de títulos públicos federais prefixados. *Revista Gestão & Tecnologia*, v. 9, n. 1, p. 1–11, 2010.

6 Anexo A - Cálculo dos parâmetros da GHsT

Partindo do artigo de AAS et al. (2006) [5], aqui será detalhado o cálculo com o qual o autor chega a média e variância. Primeiro, considere a seguinte relação:

$$X = \mu + \frac{\beta}{V} + \frac{Z}{\sqrt{V}},$$

onde $Z \sim N(0, 1)$ e $V \sim \frac{\chi^2_\nu}{h^2} \sim \text{Gama}(\frac{\nu}{2}, h^2 \frac{\nu}{2})$ e Z e V são independentes.

Notação: $X \sim \text{GHsT}(\mu, \beta, h, \nu)$.

Tomando a esperança de X , encontra-se:

$$\begin{aligned} E(X) &= E(E(X|V)) \\ &= E\left(E\left(\mu + \frac{\beta}{V} \frac{Z}{\sqrt{V}}\right)\right) \\ &= E\left(\mu + \frac{\beta}{V} \frac{E(Z|V)}{\sqrt{V}}\right) \\ &= E\left(\mu + \frac{\beta}{V} \frac{E(Z)}{\sqrt{V}}\right) \\ &= E\left(\mu + \frac{\beta}{V}\right) \\ &= \mu E\left(\frac{\beta}{V}\right) \\ &= \mu \beta E\left(\frac{1}{V}\right) \end{aligned}$$

Sabe-se que $V \sim \frac{\chi_{\nu}^2}{h^2 \frac{\nu}{2}}$, logo:

$$\begin{aligned} E\left(\frac{1}{V}\right) &= \int_0^{\infty} \frac{1}{V} \frac{(h^2 \frac{\nu}{2})^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})} V^{\frac{\nu}{2}-1} \exp\left\{-h^2 \frac{\nu}{2} V\right\} dV \\ &= \frac{1}{V} \frac{(h^2 \frac{\nu}{2})}{\Gamma(\frac{\nu}{2})} \int_0^{\infty} V^{-1} V^{\frac{\nu}{2}-1} \exp\left\{h^2 \frac{\nu}{2} V\right\} dV \\ &= \frac{1}{V} \frac{(h^2 \frac{\nu}{2})}{\Gamma(\frac{\nu}{2})} V^{\frac{\nu}{2}-2} \exp\left\{-h^2 \frac{\nu}{2} V\right\} dV. \end{aligned}$$

É notório que a integral corresponde ao núcleo de uma nova distribuição Gama com $\alpha = \frac{\nu}{2} - 1$ e $\beta = h^2 \frac{\nu}{2}$. Segue-se que,

$$\begin{aligned} E\left(\frac{1}{V}\right) &= \frac{(h^2 \frac{\nu}{2})^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2} - 1)}{\Gamma(\frac{\nu}{2}) (h^2 \frac{\nu}{2})^{\frac{\nu}{2}-1}} \\ &= \left(h^2 \frac{\nu}{2}\right)^{\frac{\nu}{2}} \left(h^2 \frac{\nu}{2}\right)^{\frac{\nu}{2}+1} \frac{1}{\Gamma(\frac{\nu}{2})} \Gamma\left(\frac{\nu}{2} - 1\right) \\ &= \left(h^2 \frac{\nu}{2}\right) \frac{\Gamma(\frac{\nu}{2} - 1)}{(\frac{\nu}{2} - 1) \Gamma(\frac{\nu}{2} - 1)} \\ &= \left(h^2 \frac{\nu}{2}\right) \frac{1}{\frac{\nu}{2} - 1} \\ &= \frac{h^2 \frac{\nu}{2}}{\frac{\nu-2}{2}} \\ &= \frac{h^2}{\nu - 2}. \end{aligned}$$

Logo, $E(\frac{1}{V}) = \frac{h^2 \nu}{\nu-2}$.

Voltando,

$$\begin{aligned} E(X) &= \mu + \beta E\left(\frac{1}{V}\right) \\ &= \mu + \beta \left(\frac{h^2 \nu}{\nu - 2}\right) \\ &= \mu + \beta E\left(\frac{\delta^2}{\nu - 2}\right). \end{aligned}$$

Assim,

$$E(X) = \mu + \beta \frac{\delta^2}{\nu - 2}.$$

Supondo $E(X) = 0$, tem-se:

$$\begin{aligned} 0 &= \mu + \beta \frac{\delta^2}{\nu - 2} \\ \mu &= -\beta \frac{\delta^2}{\nu - 2}. \end{aligned}$$

Agora, será calculado a variância fazendo uso de variância condicional.

$$\begin{aligned} \text{Var}(X) &= E(\text{Var}(X|V)) + \text{Var}(E(X|V)) \\ &= E\left(\frac{1}{V}\text{Var}(Z)\right) + \text{Var}\left(\mu + \frac{\beta}{V}\right) \\ &= E\left(\frac{1}{V}\right) + \beta^2 \text{Var}\left(\frac{1}{V}\right) \\ &= \frac{\frac{h^2\nu}{2}}{\frac{\nu}{2} - 1} + \beta^2 \frac{h^4 \left(\frac{\nu}{2}\right)^2}{\left(\frac{\nu}{2} - 2\right) \left(\frac{\nu}{2} - 1\right)^2} \\ &= \frac{\frac{\delta^2}{2}}{\frac{\nu}{2} - 1} + \beta^2 \frac{\left(\frac{\delta^4}{4}\right)^2}{\left(\frac{\nu}{2} - 2\right) \left(\frac{\nu}{2} - 1\right)^2} \\ &= \frac{\frac{\delta^2}{2}}{\frac{\nu}{2} - \frac{2}{2}} + \beta^2 \frac{\left(\frac{\delta^4}{4}\right)^2}{\left(\frac{\nu}{2} - \frac{4}{2}\right) \left(\frac{\nu}{2} - \frac{2}{2}\right)^2} \\ &= \frac{\delta^2}{\nu - 2} + \beta^2 \frac{\frac{\delta^4}{4}}{\frac{1}{8}(\nu - 4)(\nu - 2)^2} \\ &= \frac{\delta^2}{\nu - 2} + \beta^2 \frac{2\delta^4}{(\nu - 4)(\nu - 2)^2}. \end{aligned}$$

7 Anexo B - Função de Bessel

A função de Bessel, utilizada neste trabalho na seção 2.7.2 é uma solução da seguinte equação, extraída do artigo de Silva (1999) [4].

$$z^2 \frac{d^2 w}{dz^2} + z \frac{dw}{dz} - (z^2 - \nu^2)w = 0.$$

Quando $\nu > -1$ e $z > 0$, $K_\nu(z)$ é de valor real e positiva. $K_\nu(z)$ tende a zero quando o módulo de $z \rightarrow +\infty$ ($z = x + iy; x, y \in \mathbb{R}$ e $i = \sqrt{-1}$). Abaixo segue uma figura ilustrativa da função de Bessel.

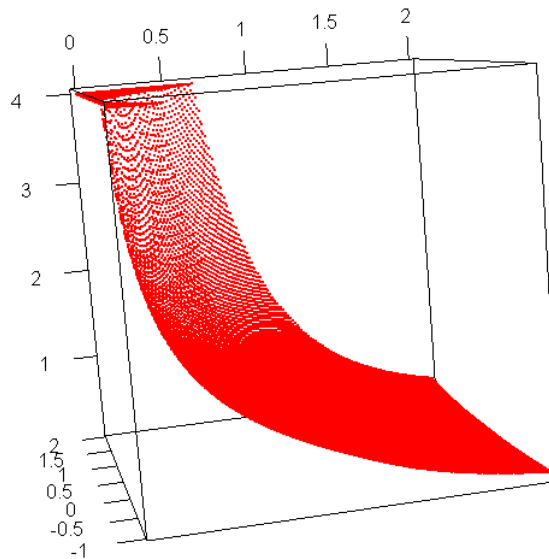


Figura 11: Gráfico da função de Bessel modificada $K_\nu(z)$ de valor real e positiva, com $z > 0$ e $\nu > -1$.

Segue abaixo algumas das funções de Bessel modificadas K_i .

$$K_\nu(z) = \int_0^\infty e^{-z \cos(h)t} \cos(h)(\nu t) dt, \operatorname{Re}(z) > 0$$

$$K_\nu(z) = \frac{1}{2} \int_0^\infty x^{\nu-1} \exp\left(-\frac{1}{2}\nu(x+x^{-1})\right) dx, \operatorname{Re}(z) > 0$$

Tem-se também que: $K_{-\nu}(z) = K_\nu(z)$.