



Universidade de Brasília
Departamento de Estatística

Modelos Heterocedásticos para a Estimação do Valor em Risco - VaR

Hyanka Mayra da Silva

Relatório apresentado para o Departamento de Estatística, Instituto de Ciências Exatas, Universidade de Brasília, como parte dos requisitos necessários para o grau de Bacharel em Estatística.

Brasília
2020

Hyanka Mayra da Silva

Modelos Heterocedásticos para a Estimação do Valor em Risco - VaR

Orientador:

Prof. Dr. **José Augusto Fiorucci**

Relatório Parcial apresentado para o Departamento de Estatística, Instituto de Ciências Exatas, Universidade de Brasília, como parte dos requisitos necessários para o grau de Bacharel em Estatística.

Brasília
2020

Dedicatória

Este trabalho, eu dedico à minha família mais especificamente à minha mãe, ela que é minha maior inspiração de vida e é a pessoa que mais acredita em mim, sem ela eu não estaria onde estou hoje.

Agradecimentos

Agradeço aos meus pais, minha mãe principalmente por ter abdicado de tanto para me proporcionar sempre a melhor educação. Obrigada por toda assistência, paciência e amor que tiveram comigo nessa jornada. Mãe, sem a senhora nada disso seria possível.

Agradeço a minha avó que é a pessoa mais forte que eu conheço e me ensinou a ser forte também. Meu avô que sempre torceu e acreditou em mim e eu sei que aonde quer que esteja está torcendo para meu sucesso.

As minhas tias que me ensinaram a ser a mulher que sou hoje, que me dão os melhores conselhos e que sempre acreditaram no meu potencial.

Ao meu tio Alexandre que me fez escolher a área que estudo hoje e se não fosse por ele não sei o que teria feito na vida.

Minha irmã por ser minha base e a pessoa que me faz querer ser sempre melhor para poder ser sempre um exemplo para ela.

Agradeço muito ao meu namorado André, ele me motiva a ser melhor todos os dias, me passa confiança de que eu posso fazer qualquer coisa, por ser meu melhor amigo, por me escutar nas horas difíceis, pela paciência e pelo companheirismo, sou uma pessoa melhor por sua causa.

De modo geral, agradeço a toda minha família que são tudo para mim e sem eles eu não seria nada.

Agradeço meu gerente Herbert e ao meu amigo Danniel que me auxiliaram muito no entendimento sobre o tema.

Agradeço meu orientador José Augusto, ele que me aceitou praticamente no meio do trabalho sem nem questionar e só me motivou a fazer o que tinha que ser feito, sempre muito paciente e disposto a tirar minhas dúvidas, meu muito obrigada.

E por fim, agradeço todos os professores do departamento, a jornada foi longa e difícil mas só foi possível por cada um de vocês.

Resumo

Modelos Heterocedásticos para a Estimação do Valor em Risco - VaR

Este trabalho visa avaliar a qualidade do *Value at Risk* (VaR) calculado a partir da metodologia de séries temporais, em especial fazendo uso dos modelos condicionalmente heterocedásticos. Para este estudo foram considerados os retornos de duas séries temporais financeiras, o índice Ibovespa e a ação da empresa Vale. Dado o comportamento típico de retornos de séries financeiras, foram considerados os modelos GARCH e EGARCH com duas variações para a componente de erro, a tradicional Normal Padrão e a t-Student Assimétrica Padronizada (SST), totalizando assim quatro abordagens. A qualidade do VaR gerado por cada modelo foi avaliada utilizando o procedimento conhecido como janela rolante, no qual os períodos de avaliação são sempre considerados em tempo posterior ao período utilizado no ajuste dos modelos. Todos os modelos apresentaram resultados satisfatórios, mas com vantagem para o modelo EGARCH com distribuição SST. A análise estatística foi realizada pelo software livre R.

Palavras-chave: *Value at Risk*, *VaR*, GARCH, EGARCH, SST, Ibovesta.

Abstract

Heteroscedastic Models for the Estimation of the Risk Value - VaR

This work aims to evaluate the quality of the Value at Risk (VaR) based on the time series methodology, especially using the conditionally heteroscedastic models. For this study, the returns of two financial time series were considered, the Ibovespa index and the Vale company stock. Given the typical behavior of income from financial series, the GARCH and EGARCH models were considered with two variations for an error component, a traditional Normal Standard and a t-Student Standardized and Asymmetric (SST), thus totaling four approaches. The quality of the VaR generated by each model was evaluated using the procedure known as a rolling window, which the evaluation periods are always considered after the period used in the adjustment of the models. All models have satisfactory results, but with an advantage for the EGARCH model with SST distribution. Statistical analysis was performed using the free software R.

Keywords: Value at Risk, VaR, GARCH, EGARCH, Ibovespa.

Lista de Figuras

2.1	Fechamento Ibovespa indexado no período de 2008 a 2019	15
2.2	log-Retorno Ibovespa	15
2.3	log-Retorno Ibovespa em 2008	16
2.4	Autocorrelação dos log-Retornos - 2008	16
2.5	Estimação de Kernel da densidade dos retornos do Ibovespa	17
3.1	Preço de Fechamento	33
3.2	Autocorrelação do Preço de Fechamento	33
3.3	log-Retorno - Ibovespa	34
3.4	log-Retorno - Vale	34
3.5	Quantis do log-retorno - Ibovespa	35
3.6	Quantis do log-retorno - Vale	35
3.7	Autocorrelação do log-retorno	36
3.8	Autocorrelação do log-retorno ²	36
3.9	Autocorrelação dos resíduos padronizados ² - Ibovespa	38
3.10	Autocorrelação dos resíduos padronizados ² - Vale	38
3.11	Resíduos padronizados do modelo GARCH (1, 1) - Ibovespa	39
3.12	Resíduos padronizados do modelo GARCH (1, 1) - Vale	39
3.13	Histograma dos resíduos padronizados do modelo GARCH (1, 1) - Ibovespa	40
3.14	Histograma dos resíduos padronizados do modelo GARCH (1, 1) - Vale	40
3.15	Quantil dos resíduos padronizados do modelo GARCH (1, 1) - Ibovespa	40
3.16	Quantil dos resíduos padronizados do modelo GARCH (1, 1) - Vale	41
3.17	Autocorrelação dos resíduos padronizados ² - Ibovespa	43
3.18	Autocorrelação dos resíduos padronizados ² - Vale	43
3.19	Resíduos padronizados do modelo EGARCH (1, 1) - Ibovespa	44
3.20	Resíduos padronizados do modelo EGARCH (1, 1) - Vale	44
3.21	Histograma dos resíduos padronizados do modelo EGARCH (1, 1) - Ibovespa	45
3.22	Histograma dos resíduos padronizados do modelo EGARCH (1, 1) - Vale	45
3.23	Quantil dos resíduos padronizados do modelo EGARCH (1, 1) - Ibovespa	45
3.24	Quantil dos resíduos padronizados do modelo EGARCH (1, 1) - Vale	46
3.25	Volatilidade do modelo GARCH (1, 1) - Ibovespa	48
3.26	Volatilidade do modelo GARCH (1, 1) - Vale	48
3.27	log-Retorno e Violações do VaR - Ibovespa	49
3.28	log-Retorno e Violações do VaR - Vale	50
3.29	Volatilidade do modelo EGARCH (1, 1) - Ibovespa	52
3.30	Volatilidade do modelo EGARCH (1, 1) - Vale	52
3.31	log-Retorno e Violações do VaR - Ibovespa	53
3.32	log-Retorno e Violações do VaR - Vale	54

Lista de Tabelas

1	Fechamento Ibovespa	19
2	Medidas de Resumo - log-retorno	34
3	Estimativas do Modelo GARCH (1,1)- Ibovespa	37
4	Estimativas do Modelo GARCH (1,1) - Vale	37
5	Medidas de Resumo - Resíduos Padronizados do modelo GARCH(1,1) . . .	39
6	Estimativas do Modelo EGARCH (1,1) - Ibovespa	42
7	Estimativas do Modelo EGARCH (1,1)- Vale	42
8	Medidas de Resumo - Resíduos Padronizados do modelo EGARCH(1,1) . .	44
9	Comparação entre modelos - Ibovespa	47
10	Comparação entre modelos - Vale	47
11	Coberturas do modelo GARCH(1,1)	51
12	Teste de cobertura do modelo GARCH(1,1)	51
13	Coberturas do modelo GARCH(1,1)	55
14	Teste de cobertura do modelo GARCH(1,1)	55

Sumário

Agradecimentos	iv
Resumo	v
Abstract	vi
1 Introdução	12
1.1 Organização dos Capítulos	13
2 Metodologia	14
2.1 Retorno	14
2.2 Volatilidade	17
2.3 <i>Value at Risk - VaR</i>	18
2.3.1 <i>Backtesting</i>	20
2.4 Modelos Heterocedásticos	21
2.5 Modelo GARCH	24
2.5.1 Estimação	26
2.5.2 Verificação	27
2.6 Modelo EGARCH	28
2.6.1 Estimação	30
2.6.2 Verificação	32
3 Aplicação	33
3.1 Séries Estudadas	33
3.2 Ajuste dos Modelos	37
3.2.1 GARCH	37
3.2.2 EGARCH	41
3.3 Acurácia do VaR	47
3.3.1 GARCH	48
3.3.2 EGARCH	51
4 Considerações Finais	56
5 Apêndice	57
5.1 Teste de Box-Pierce-Ljung	57
5.2 Critério de Akaike - AIC	57
5.3 Critérios de Informação Bayesiano - BIC	58
Referências	60

1 Introdução

Com a recente queda na taxa SELIC, muitas pessoas vêm procurando diversificar seus investimentos afim de ter ganhos de capitais acima de aplicações tradicionais, como poupança, tesouro direto e outras de renda fixa. O mercado de renda variável oferece liquidez e possibilidade de ganhos expressivos, seja através da valorização dos ativos ou por meio de participação nos lucros, no caso, recebimento de dividendos ou juros sobre capital próprio.

O mercado de ações pode ser visto como um vasto mecanismo de avaliação de empresas com ações negociadas em bolsa. Informações sobre economias, setores e empresas viajam a altas velocidades para analistas de investimento e investidores de todos os portes. Na medida em que as variações do preço das ações de uma empresa oferecem evidências confiáveis de alterações em seu nível de crédito, os credores têm a oportunidade de explorar uma ferramenta de gestão de risco de crédito de alcance e poder enorme (Caouette, Altman e Narayanan (2000)).

A possibilidade de ganhos elevados consiste na parte mais atraente do mercado financeiro, no entanto, renda variável está diretamente vinculada ao risco de perdas. Por conta disto, a modelagem estatística da movimentação dos preços de ativos financeiros, em especial dos retornos, está em constante aprimoramento, uma vez que esse tipo de modelagem está envolvida na gestão de risco. É de conhecimento científico que retornos financeiros possuem propriedades específicas como assimetria à esquerda, distribuição de probabilidade leptocúrtica e presença de volatilidade. Como um dos fatos estilizados das séries temporais como por exemplo o agrupamento de volatilidade, a tendência de aumento com as grandes mudanças seriam seguidos por grandes mudanças e as pequenas tendem a ser seguidas por pequenas mudanças.

Fazendo uma linha do tempo sobre a modelagem da volatilidade, temos que, inicialmente foi proposto o modelo Autoregressivo Condicionalmente Heterocedástico (ARCH) por Engle (1982) e em seguida Bollerslev (1986) estendeu-o para o modelo ARCH Generalizado (GARCH). À medida que mais regressores foram sendo adicionados ao modelo GARCH, vários outros modelos derivados do GARCH foram sendo criados, como o modelo do valor absoluto GARCH (AV-GARCH) de Taylor (1986) e Schwert (1990), o modelo exponencial GARCH (EGARCH) de Nelson (1991), modelo Glosten-Jagannathan-Runkle GARCH (GJRGARCH) que leva em conta o fator de assimetria γ de Glosten, Jagannathan e Runkle (1993), o modelo Threshold GARCH (T-GARCH) de Zakoian (1994) e assim por diante.

A suposição de normalidade aplicada ao GARCH não é suficiente para explicar a assimetria à esquerda, distribuição de probabilidade leptocúrtica e caudas pesadas.

Várias distribuições foram implementadas para descrever o conjunto de dados, como a distribuição t-Student e várias outras.

O *VaR* estabeleceu-se como a medida mais proeminente de risco financeiro de mercado. Embora tenha sido criticado por ser teoricamente deficiente (por exemplo, não-aditivo) e numericamente problemático (é não-convexo), também é ainda reconhecido, apesar de suas desvantagens em comparação com medidas coerentes de risco, como a medida de risco mais ampla utilizada na prática, e seu cálculo preciso também é fundamental para o cálculo de outras medidas de risco baseadas em quantis, como deficit esperado (Dowd e Blake (2006)).

1.1 Organização dos Capítulos

A presente monografia está organizada da seguinte forma: no Capítulo 2 é apresentado uma revisão das principais metodologias utilizadas neste estudo, em especial, o conceito de retornos e volatilidade, os modelos condicionalmente heterocedásticos e as equações para o cálculo do VaR. No Capítulo 3 são apresentados os conjuntos de dados, o ajuste dos modelos e a avaliação da medida de risco computada por cada um dos modelos selecionados. A conclusão e os comentários finais são apresentados no Capítulo 4.

2 Metodologia

2.1 Retorno

Segundo Pedro Alberto Morettin (2008), um dos objetivos em finanças é a avaliação de riscos de uma carteira de ativos (instrumentos) financeiros. O risco é frequentemente medido em termos da variação de preços dos ativos.

A modelagem de séries temporais financeiras é um problema complexo e se deve principalmente à existência de regularidades estatísticas (fatos estilizados) que são comuns a um grande número de séries financeiras e são difíceis de reproduzir artificialmente usando modelos estocásticos.

A maioria dos fatos estilizados que serão mostrados a seguir, foram apresentados em um artigo de Mandelbrot (1963). Desde então, estes fatos foram documentados e concluídos por muitos estudos empíricos, sendo potencialmente observados de forma clara dependendo da natureza da série e sua frequência. As propriedades apresentadas a seguir são principalmente de preços diários das ações.

Supondo P_t o preço de um ativo no instante t , então variação relativa de preços ou retorno líquido simples deste ativo entre os mesmo instantes é definida por

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}} \quad (2.1)$$

O retorno bruto simples é definido como:

$$R_t = \frac{P_t}{P_{t-1}} - 1 \Rightarrow R_t + 1 = \frac{P_t}{P_{t-1}} \quad (2.2)$$

usualmente R_t é expresso em porcentagem, relativamente ao período.

O retorno composto continuamente ou simplesmente log-retorno é expresso como:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(1 + R_t) = p_t - p_{t-1} \quad (2.3)$$

Segundo Francq e Zakoian (2019), os retornos ou preços relativos não dependem de unidades monetárias, diferente dos preços ou cotas, o que facilita as comparações entre os ativos. Os principais fatos estilizados podem ser resumidos como:

- i. Não estacionariedade das séries de preços/cotas;
- ii. Ausência de autocorrelação para as variações de preço/cota;

- iii. Autocorrelações dos retornos de preços ao quadrado;
- iv. Agrupamento de volatilidade;
- v. Distribuições de cauda pesadas.

Para ilustrar estes fatos considere a série do índice Ibovespa, na Figura 2.1, é apresentado a série de fechamento e note como a série oscila ao longo do tempo. Essas séries são próximas a um passeio aleatório, não possui média constante e nem variância constante.

Tomando a média e a variância em períodos específicos é possível perceber como variam dependendo do período considerado, isso ilustra o item i. da não estacionariedade dos preços/cotas, no caso do exemplo é considerado o preço de fechamento do Ibovespa.

A Figura 2.2 apresenta a séries dos log-retornos em todo o período e mostra um comportamento compatível com as suposições de estacionariedade, note como os retornos oscilam em torno de zero e não há períodos em que a variação é muito diferente da série no geral. A Figura 2.3 apresenta a série do log-retorno apenas no anos de 2018 que devido a crise financeira há uma maior volatilidade, em magnitude nesse período.

Figura 2.1: Fechamento Ibovespa indexado no período de 2008 a 2019



Figura 2.2: log-Retorno Ibovespa

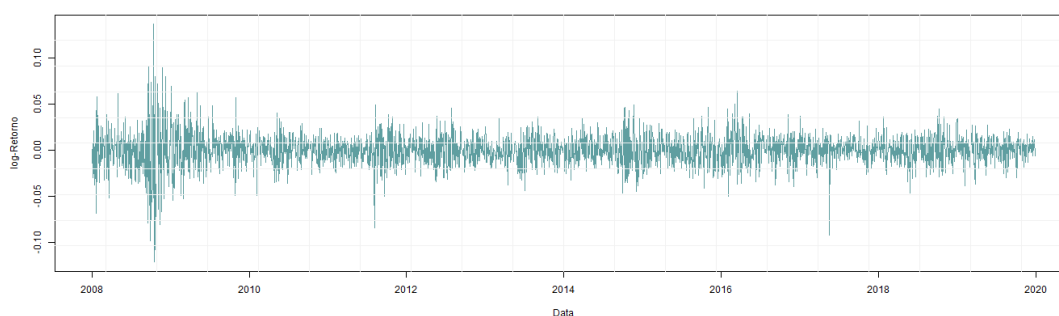
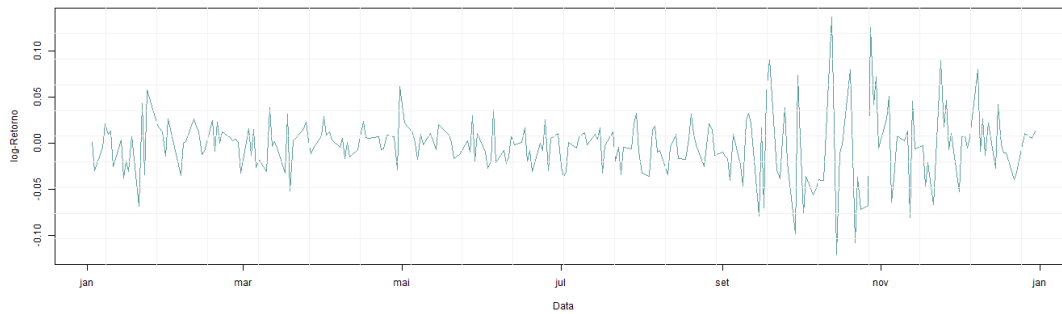
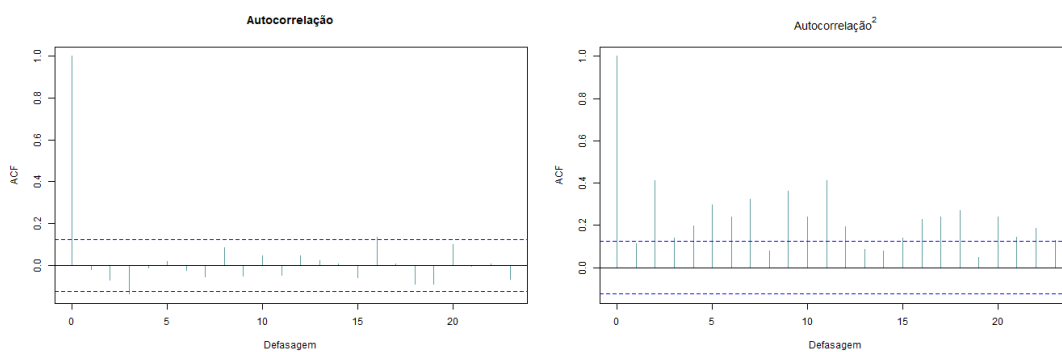


Figura 2.3: log-Retorno Ibovespa em 2008



O primeiro gráfico da Figura 2.4 apresenta a autocorrelação para séries dos log-retornos, no qual observamos o fato da ausência de autocorrelação, note como não há autocorrelação significativa, ou seja, nenhuma defasagem ultrapassou as faixas, tornando-a próxima de um ruído branco. No entanto quando tomamos o gráfico da autocorrelação para o quadrado dos retornos, segundo gráfico da Figura 2.4, podemos notar que existe autocorrelações significativas em diversas defasagens, o que evidencia a presença de autocorrelação na variância condicional (volatilidade).

Figura 2.4: Autocorrelação dos log-Retornos - 2008

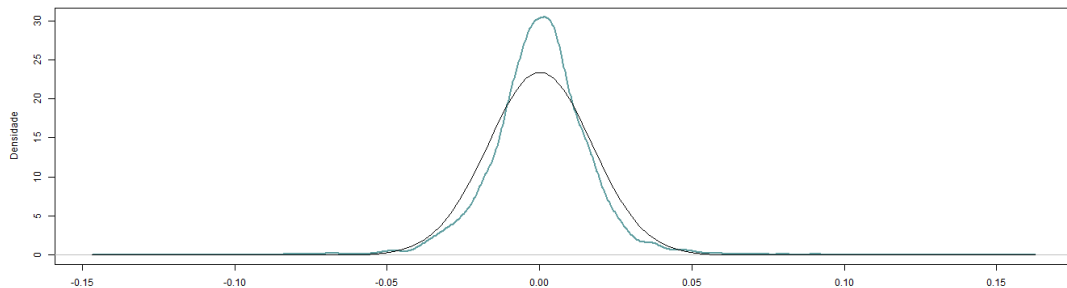


A Figura 2.3 também evidencia o agrupamento de volatilidade, onde grandes retornos tendem a aparecer em clusters ou grupos, os períodos de alta volatilidade são seguidos por períodos de baixa volatilidade.

A Figura 2.5 ilustra o caso das caudas pesadas para os retornos, ela compara o estimador de kernel da densidade dos retornos do Ibovespa (linha verde) com a densidade da distribuição Normal com média e variância dos retornos do Ibovespa (linha preta). Note como a distribuição dos retornos não se assemelha com a distribuição Normal. Como será mostrado a seguir os testes estatísticos levam a rejeição da hipótese nula de que os dados dos retornos seguem uma distribuição Normal. A densidade dos retornos tendem a zero lentamente e atingem o pico bem acentuado em zero, por isso são chamados de

leptocúrticos. A medida de curtose evidencia esse fato, uma vez que para dados gaussianos com observações i.i.d. é igual a 3 assintoticamente e dados financeiros possuem o coeficiente de curtose muito maior do que 3.

Figura 2.5: Estimação de Kernel da densidade dos retornos do Ibovespa



Os modelos tradicionais de séries temporais assumem seus dados correlacionados com média e variância condicional constante no tempo. Já os modelos heterocedásticos assumem que as observações são não serialmente correlacionadas, porém, são dependentes, com média constante, mas a variância condicional podendo variar no tempo.

2.2 Volatilidade

É uma medida de dispersão dos retornos de um ativo financeiro, como exemplificado anteriormente, representa uma variável relativa à intensidade e frequência das oscilações nos preços ou retornos de ativo financeiro e é um dos parâmetros mais utilizados como forma de mensurar o risco de um ativo.

Convencionou chamar de **volatilidade**, que pode ser definida de várias maneiras, mas não é diretamente observável. Para levar em conta a presença de grupos (*clusters*) de volatilidade em uma série financeira é necessário recorrer a modelos heterocedásticos condicionais.

Segundo Shumway e Stoffer (2017), nestes modelos, o desvio padrão (volatilidade) de um retorno num dado instante de tempo, depende de retornos passados e de outras informações disponíveis até àquele instante, de modo a definir uma **variância condicional** que, não sendo constante, não coincide com a variância global ('incondicional') da série observada. Do mesmo modo, é possível que a média varie com o tempo, ou outros momentos da distribuição dos retornos variem com o tempo.

Segundo Pedro Alberto Morettin (2008), há três enfoques para o cálculo de volatilidade, sendo eles:

- i. uma maneira é equacionar um preço de mercado observado com o preço modelado de uma opção. Obtemos o que se chama volatilidade implícita, essa fórmula supõe normalidade dos preços e volatilidade constante;
- ii. outra maneira é modelar diretamente a volatilidade da série de retornos, usando alguma família, como a dos modelos ARCH; obtemos a chamada volatilidade estocástica;
- iii. uma alternativa é modelar a volatilidade por meio de uma média de uma função de k retornos. Obtemos o que se chama de volatilidade histórica. Podemos considerar os quadrados dos retornos ou os valores absolutos dos retornos nesta média móvel.

Há várias outras formas de calcular a volatilidade, que depende bastante do objetivo e de como os dados financeiros foram coletados. No presente estudo a volatilidade utilizada será a (ii), mas de modo geral, o modelo típico para a volatilidade é da forma

$$r_t = \mu_t + \sqrt{h_t}\varepsilon_t,$$

onde h_t é a variância da série, $E_{t-1}[\varepsilon_t] = 0$, $Var_{t-1}(\varepsilon_t) = 1$ e tipicamente ε_t é i.i.d. com distribuição F.

2.3 Value at Risk - VaR

A medida do VaR é estimada em unidades monetárias (\$) da perda máxima esperada em determinado período de tempo com certo nível de confiança, em condições normais de mercado, considerando que os retornos seguem uma distribuição Gaussiana e a partir do valor atual do investimento. De maneira geral, estudos realizados por Jorion (2003) mostram que o VaR é definido como a Equação 2.1,

$$VaR = -I \times Z \times \sigma, \tag{2.1}$$

onde I é o investimento, Z representa o quantil da distribuição Normal e σ é o desvio padrão dos log-retornos. O VaR tem valor negativo, pois quem tem uma posição comprada sofre uma perda se a variação do ativo for menor do que zero.

Para ilustrar uma aplicação do VaR, considere os seguintes dados hipotéticos na tabela abaixo:

Tabela 1: Fechamento Ibovespa

Data	Fechamento	Retorno
04-01-2010	70.045	-
05-01-2010	70.240	0,28%
06-01-2010	70.729	0,70%
07-01-2010	70.451	-0,39%
08-01-2010	70.263	-0,27%
11-01-2010	70.433	0,24%
12-01-2010	70.076	-0,51%
13-01-2010	70.385	0,44%
14-01-2010	69.801	-0,83%
15-01-2010	68.978	-1,18%

O retorno acumulado no decorrer dos oito meses foi de -1,52% e o risco é de 0,59%. Suponha que o desvio padrão projetado para o próximo período seja 0,59, então considerando o nível de confiança de 95%, o quantil da distribuição normal que representa essa confiança é 1,645. Assim, considerando a compra de uma cota a R\$ 68.978, o *VaR* desse investimento é apresentado na Equação 2.1

$$VaR_{95\%} = -68.978 \times 1,645 \times 0,59\% = R\$ - 668,12 \quad (2.2)$$

O *VaR* calculado mostra que, com 95% de confiança, a perda máxima estimada no próximo mês, considerando a compra de um título no último dia no valor de R\$ 68.978 é de R\$668,12.

O *VaR* (*Value at Risk*) é um método de mensuração de risco baseado em modelagem estocástica, o qual mede potenciais perdas ao longo de um intervalo de tempo, sob condições normais de mercado e dentro de um determinado nível de significância α . De maneira geral, o VaR é descrito conforme a Equação 2.3,

$$VaR_{\alpha}(L) = \inf\{\ell \in \mathbb{R} : P(L > \ell) \leq 1 - \alpha\}, \quad (2.3)$$

onde L representa a perda observada e ℓ é a perda esperada.

Em Jorion (2003), o VaR é a estimação da maior perda no próximo instante de tempo, considerando retornos passados, um grau de confiança e a variação dos retornos. A abordagem aqui utilizada é simples, considerando na sua estimação o desvio padrão dos retornos passados. Esse método não apresenta bons resultados quando há crises ou movimentos inesperados de mercado, quando isso ocorre essa estimação do VaR acaba sofrendo de viés (Assaf (2015)).

Um outro problema com essa abordagem ocorre quando os dados são assimétricos, pois o VaR utiliza a série histórica dos retornos, logo, se esta for assimétrica é necessária uma abordagem diferente para tratamento da variação dos dados.

2.3.1 *Backtesting*

Para especificar uma métrica do valor em risco é importante identificar o período, um quantil de perda, por exemplo 0,90, 0,95 e 0,99, que corresponde a um α de 10%, 5% e 1%, respectivamente.

Segundo Holton (2003), a tradição empírica nos diz que um modelo não deve ser avaliado com base na razoabilidade das suas suposições ou na sofisticação de sua análise. Deve ser avaliado com base na utilidade de suas previsões. O *backtesting* é o processo de comparar as perdas previstas por um modelo de valor de risco (VaR) com experiências reais durante o período de teste. Isso é feito para garantir que os modelos VaR sejam razoavelmente precisos.

O objetivo geral do *backtesting* é garantir que as perdas reais não excedam as perdas esperadas em um determinado nível de confiança. Excedências são os números de observações reais acima do nível esperado. No contexto do VaR, o número de violações não deve exceder nível de significância, por exemplo, as violações devem ocorrer apenas para 1% do tempo se o nível de confiança for de 99%. Para um modelo perfeitamente ajustado, o número de excedências observadas deve ser aproximadamente o mesmo que o nível de significância VaR.

As técnicas para testar estatisticamente as medidas de valor em risco seguem a metodologia de *backtesting*, onde os testes de cobertura avaliam se a frequência de violações é consistente com o quantil de perda que uma medida do VaR pretende refletir, os testes de aderência são utilizados para verificar se a distribuição de probabilidade hipotética dos retornos condiz com os observados e os testes de independência avaliam se o modelo empregado conseguiu captar toda a volatilidade dos retornos.

Quando muitas violações são observadas, o modelo empregado provavelmente não é adequado e pode estar subestimando o risco. A importância do *backtesting* se dá por oferecer uma verificação de acurácia nas previsões do VaR.

2.3.1.1 Testes de Cobertura

Holton (2003), considera uma medida de valor em risco de quantil q de perda e define um processo de excedência univariada I_t com termos:

$$I_t(\alpha) = \begin{cases} 0, & r_t > -VaR_{t|t-1}(\alpha) \text{ se a perda é menor ou igual ao VaR} \\ 1, & r_t < -VaR_{t|t-1}(\alpha) \text{ se a perda excede o VaR} \end{cases} \quad (2.4)$$

Definimos a cobertura q^* da medida do VaR como a frequência de violações por período. Isso pode ser expresso como uma esperança incondicional

$$q^* = E[I_t] \quad (2.5)$$

Os testes de cobertura verificam a hipótese nula $q = q^*$. Seja X o número observado de violações observadas nos dados:

$$x = \sum_{t=0}^n I \quad (2.6)$$

Sob a hipótese nula de um modelo corretamente ajustado (H_0 : modelo bem ajustado), o número de violações (X) segue uma distribuição de probabilidade binomial:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

onde n é o tamanho da amostra e p é a probabilidade de violações.

A esperança da distribuição binomial é dada por:

$$E[X] = np$$

e a variância por:

$$Var(x) = np(1 - p)$$

A suposição inerente aqui presente é que as violações (falhas) são variáveis aleatórias i.i.d. Considerando X para representar o número de excedências, a taxa de cobertura será dada por X/n ou a média das excedências que são os retorno que são menores que o VaR.

2.4 Modelos Heterocedásticos

Modelos heterocedásticos permitem que a variância da série seja tratada como sendo um processo dinâmico. Mais especificamente para o cálculo do VaR, os modelos de inte-

resse são destinados a modelar a variância condicional dos retornos, sendo assim conhecidos como condicionalmente heterocedásticos.

Considerando a definição de retorno mencionada na Seção 2.1, seja $X_t = (r_1, r_2, \dots, r_t)$ uma série de retornos de uma variável P_t , considere que $\mu_t = E[X_t] = 0$ e $\sigma_t^2 = Var(X_t) = \sigma^2$, ou seja, a média e a variância não condicional são constantes.

Para melhor compreensão de algumas propriedades que serão apresentadas no decorrer desta monografia, segue abaixo algumas definições.

Definição 2.1 A *função de autocorrelação (FAC)* é definida como

$$\rho(s, t) = cor(x_s, x_t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}} = \frac{\gamma(s, t)}{\sigma_s \sigma_t}. \quad (2.1)$$

A FAC mede a previsibilidade linear da série no tempo t , digamos x_t , usando apenas o valor x_s . Podemos mostrar facilmente que $-1 \leq \rho(s, t) \leq 1$ usando o método da desigualdade de Cauchy - Schwarz. Se pudermos prever x_t perfeitamente de x_s através de um relacionamento linear, $x_t = \beta_0 + \beta_1 x_s$, a correlação será $+1$ quando $\beta_1 > 0$ e -1 quando $\beta_1 < 0$. Portanto, temos uma medida aproximada da capacidade de prever a série no tempo t do valor no tempo s .

Definição 2.2 *Séries temporais estritamente estacionárias* são aquelas para as quais a probabilidade comportamental de toda a coleção de valores

$$\{x_{t_1}, x_{t_2}, \dots, x_{t_k}\}$$

é idêntico ao conjunto do tempo alterado

$$\{x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}\}$$

para quaisquer índices t_1, t_2, \dots, t_n e qualquer defasagem $h = 0, 1, 2, \dots$. Isso é,

$$P(x_{t_1} \leq c_1, \dots, x_{t_k} \leq c_k) = P(x_{t_1+h} \leq c_1, \dots, x_{t_k+h} \leq c_k) \quad (2.2)$$

para todos $k = 1, 2, \dots$, todos os pontos do tempo t_1, t_2, \dots, t_k , todos os números c_1, c_2, \dots, c_k , e toda mudança de tempo $h = 0, \pm 1, \pm 2, \dots$.

Se uma série temporal é estritamente estacionária, todas as funções de distribuição multivariada para os subconjuntos de variáveis devem concordar com suas contrapartes no conjunto deslocado para todos os valores do parâmetro alterado h . Por exemplo, quando $k = 1$, em 2.2, implica que

$$P(x_s \leq c) = P(x_t \leq c) \quad (2.3)$$

para qualquer ponto do tempo s e t .

Além disso, se a função média, μ_t , da série existe, Equação 2.3 implica que $\mu_s = \mu_t$ para todos os s e t , portanto, μ_t deve ser constante.

Quando $k = 2$, 2.3 pode ser escrita como

$$P(x_s \leq c_1, x_t \leq c_2) = P(x_{s+h} \leq c_1, x_{t+h} \leq c_2) \quad (2.4)$$

para quaisquer pontos do tempo s , e t e mudança de h . Assim, se a função de variância do processo existe, Equação 2.3 - Equação 2.4 implica que a função de autocovariância da série x_t satisfaz

$$\gamma(s, t) = \gamma(s + h, t + h)$$

para todo s e t e h . Pode interpretar esse resultado dizendo que a função de autocovariância do processo depende apenas da diferença de tempo entre s e t , e não dos tempos reais.

A estacionaridade estrita é muito forte para a maioria das aplicações. Além disso, é difícil avaliar a estacionariedade estrita a partir de um único conjunto de dados. Ao invés de impor condições a todas as distribuições possíveis de uma série temporal, usa-se uma versão que impõe condições apenas nos dois primeiros momentos da série.

Definição 2.3 *Uma série temporal fracamente estacionária ou somente estacionária, x_t , é um processo de variação finita onde*

- i a função de valor médio, $\mu_t = E[x_t] = \mu$, $\forall t = 1, 2, \dots$ é constante e não depende do tempo t ;*
- ii variância finita, $Var(x_t) < \infty$ para todo $t = 1, 2, \dots$ e*
- iii a função de autocovariância, $\gamma(s, t)$ depende apenas de s e t através da diferença $|s - t|$.*

Estacionaridade requer regularidade nas funções de média e autocorrelação para que essas quantidades (pelo menos) possam ser estimadas pela média. Deve ficar claro a partir da estacionariedade estrita, que uma série temporal estritamente estacionária, com variância finita também é estacionária. O inverso não é verdadeiro, a menos que haja outras condições. Um caso importante em que estacionariedade implica estacionariedade estrita é se a série temporal for gaussiana (o que significa todas as distribuições finitas da série são gaussianos).

Definição 2.4 *A função de autocorrelação parcial FACP de um processo esta-*

cionário, x_t denotada por ϕ_{hh} , para $h = 1, 2, \dots$, é

$$\phi_{hh} = \text{cor}(x_{t+h}, x_t) = \rho(1) \quad (2.5)$$

e

$$\phi_{hh} = \text{cor}(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t), \quad h \geq 2. \quad (2.6)$$

O FACP, ϕ_{hh} , é uma correlação entre x_{t+h} e x_t com dependência linear de $\{x_{t+1}, \dots, x_{t+h-1}\}$ em cada um, removidos. Se o processo x_t for gaussiano, então $\phi_{hh} = \text{cor}(x_{t+h}, x_t | x_{t+1}, \dots, x_{t+h-1})$, isto é, ϕ_{hh} é o coeficiente de correlação entre x_{t+h} e x_t na distribuição bivariada de (x_{t+h}, x_t) condicional a $\{x_{t+1}, \dots, x_{t+h-1}\}$. Para mais detalhes, consultar Shumway e Stoffer (2017).

2.5 Modelo GARCH

Motivado pelo estudo da volatilidade ou variabilidade de uma série temporal, onde a suposição de uma variância condicional constante é violada, modelos como heterocedásticos condicionalmente autorregressivos ou ARCH, introduzido pela primeira vez por Engle, foi desenvolvido para modelar mudanças na volatilidade. Esses modelos foram posteriormente estendidos ao modelo ARCH generalizado ou GARCH por Bollerslev (1986).

Um modelo ARMA pode ser mais parcimonioso, no sentido de apresentar menos parâmetros do que o AR ou MA puro. Do mesmo modo, um modelo GARCH pode ser usado para descrever a volatilidade, sendo mais flexível no sentido de descrever uma memória mais longa e com menos parâmetros do que um modelo ARCH. A idéia foi introduzir na fórmula da variância condicional regressores da variância condicional passada.

Definição 2.5 Um modelo GARCH(p, q) é definido por

$$r_t = \sqrt{h_t} \varepsilon_t \quad (2.1)$$

onde

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \quad (2.2)$$

em que ε_t é uma sequência de variáveis aleatórias independentes e identicamente distribuídas, com média zero e variância constante, $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, p-1, \beta_j \geq 0, j = 1, \dots, q-1, \alpha_p > 0, \beta_q > 0, \sum_{i=1}^m (\alpha_i + \beta_i) < 1, m = \max(p, q)$.

Em Bollerslev (1986) é mostrado que a condição $\sum_{i=1}^q \beta_i + E[\varepsilon_0^2] \sum_{j=1}^p \alpha_j < 1$ é imposta para que o modelo seja estacionário.

O modelo GARCH possui algumas vantagens e desvantagens, como falado anteriormente nos fatos estilizados dos retornos. Grandes retornos tendem a aparecer em *clusters* ou grupos, os períodos de alta volatilidade são seguidos por retornos de alta volatilidade, as caudas da distribuição dos retornos são mais pesadas do que a da distribuição normal, de modo que a curtose é maior que 3. Os retornos são incorporados na fórmula elevados ao quadrado, r_{t-i}^2 , ou seja, sendo positivo ou negativo serão tratados da mesma maneira, porém, empiricamente a volatilidade reage de maneira diferente dependendo do sinal do retorno, sendo mais influenciada por retornos negativos, os modelos tendem a super prever a volatilidade porque respondem lentamente a grandes retornos isolados.

Da Definição 2.5 é possível perceber que se $q = 0$, o modelo reduz para um ARCH (p) e, se p e q são nulos, $r_t = \sqrt{\alpha_0}\varepsilon_t$ é um ruído branco. De maneira análoga à relação entre os modelos ARMA e MA, um modelo GARCH pode ser escrito como um modelo ARCH de ordem infinita, ARCH (∞), sob certas condições de regularidade, isso é mostrado em Bollerslev (1986).

Supomos que os ε_t são normais ou seguem uma distribuição t-Student padronizada, considere

$$\nu_t = r_t^2 - h_t, \quad (2.3)$$

de modo que, substituindo em 2.2 obtemos

$$r_t^2 = \alpha_0 + \sum_{i=1}^m (\alpha_i + \beta_i) r_{t-i}^2 - \sum_{j=1}^q \beta_j \nu_{t-j} + \nu_t, \quad (2.4)$$

onde $m = \max(p, q)$ e temos um modelo ARMA(m, q) para r_t^2 , mas ν_t não é, em geral, um processo i.i.d. Na realidade, ν_t é uma diferença de martingale, pois, para todo t ,

$$E[\nu_t] = E[r_t^2 - h_t] = E[h_t \varepsilon_t^2 - h_t] = E[h_t] [E[\varepsilon_t^2] - 1] = 0,$$

$$E[\nu_t | F_{t-1}] = E[r_t^2 | F_{t-1}] - E[h_t | F_{t-1}] = h_t - h_t = 0.$$

Segue que, em particular,

$$E[r_t^2] = \frac{\alpha_0}{1 - \sum_{i=1}^m (\alpha_i + \beta_i)}.$$

A longo prazo, a volatilidade convergirá para essa média.

Pedro Alberto Morettin (2008) mostra que um modelo bastante usado na prática é o GARCH (1,1), para o qual a volatilidade é expressa como:

$$h_t = \alpha_0 + \alpha_1 r_t^2 + \beta_1 h_{t-1}, \quad (2.5)$$

com $0 < \alpha_1, \beta_1 < 1$, $\alpha_1 + \beta_1 < 1$. Para o modelo 2.5 obtemos

$$K = \frac{E[r_t^4]}{[E[r_t^2]]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3, \quad (2.6)$$

dado que o denominador seja positivo, o que novamente mostra que se r_t segue um modelo GARCH, as caudas de r_t são mais longas do que as da normal.

2.5.1 Estimação

Para a estimação dos parâmetros do modelos GARCH(p, q), 2.1 e 2.2, são obtidos pelo método de máxima verossimilhança condicional, onde a função de verossimilhança é dada por:

$$\begin{aligned} \mathcal{L}(r_1, \dots, r_T | \alpha, \beta) &= f(r_T | F_{T-1}) f(r_{T-1} | F_{T-2}) \dots f(r_{p+1} | F_p) f(r_1, \dots, r_p | \alpha, \beta) \\ &= \prod_{t=p+1}^T f(r_t | F_{t-1}) f(r_{t-1} | F_{t-2}) f(r_1, \dots, r_p | \alpha, \beta), \end{aligned} \quad (2.7)$$

Os erros ε_t de modelo GARCH podem seguir qualquer distribuição com média zero e variância um, segundo a definição do modelo.

Distribuição Normal

Supondo normalidade dos $\varepsilon_t \sim N(0, 1)$ temos que a função de densidade condicional de r_t segue uma distribuição Normal com média zero e variância h_t dada por:

$$f(r_t) = \frac{1}{\sqrt{2\pi h_t}} \exp \left\{ \frac{-r_t^2}{2h_t} \right\}.$$

Assim a função de verossimilhança condicional às p primeiras observações é dada por

$$\mathcal{L}(r_{p+1}, \dots, r_T | \alpha, \beta, r_1, \dots, r_p) = \prod_{p+1}^T \left(\sqrt{2\pi h(t)} \right)^{-1} \exp \left\{ \frac{-r_t^2}{2h_t} \right\}, \quad (2.8)$$

onde a volatilidade h_t é obtida recursivamente.

Tomando o logaritmo da Equação 2.8

$$\ell(r_{p+1}, \dots, r_T | \alpha, \beta, r_1, \dots, r_p) \propto -\frac{1}{2} \sum_{t=p+1}^T \ln(h_t) - \frac{1}{2} \sum_{t=p+1}^T \frac{r_t^2}{h_t}. \quad (2.9)$$

Bollerslev (1986) utiliza em 2.9, $h_t = \hat{\sigma}^2$, $t = 1, \dots, n$, com $\hat{\sigma}^2 = \sum_{t=1}^T \frac{r_t^2}{T}$. As

estimativas dos parâmetros são obtidas por meio de métodos numéricos de maximização, utilizando a função de verossimilhança condicional (Pedro Alberto Morettin (2008)).

Como feito nos artigos de Engle (1982) e Bollerslev (1986), utiliza a log-verossimilhança média com erros seguindo uma distribuição normal padrão para estimar os parâmetros do modelo dada por:

$$\ell(\alpha, \beta) = \frac{1}{T} \sum_{t=1}^T -\frac{1}{2} \ln(h_t) - \frac{r_t^2}{h_t}.$$

Distribuição t-Student

Suponha que $\varepsilon_t \sim t_\nu$, distribuição t-Student com ν graus de liberdade, a função de densidade condicional de r_t é dada por:

$$f(r_t) = \frac{\Gamma((\nu + 1)/2)}{\Gamma(\nu/2)\sqrt{(\nu - 2)}\pi h_t} \left(1 + \frac{r_t^2}{(\nu - 2)h_t}\right)^{-(\nu+1)/2}.$$

A verossimilhança condicionada aos p primeiros retornos é dada por:

$$\mathcal{L}(r_{p+1}, \dots, r_T | \alpha, \beta, r_1, \dots, r_p) = \prod_{p+1}^T \frac{\Gamma((\nu + 1)/2)}{\Gamma(\nu/2)\sqrt{(\nu - 2)}\pi h_t} \left(1 + \frac{r_t^2}{(\nu - 2)h_t}\right)^{-(\nu+1)/2}. \quad (2.10)$$

onde a volatilidade h_t é obtida recursivamente.

Para encontrar os estimadores de máxima verossimilhança para os parâmetros é preciso maximizar a função de verossimilhança condicional. Neste caso, é possível estimar os parâmetros fixando um grau de liberdade ou estimar os graus de liberdade juntamente com os parâmetros. Em Bollerslev (1986), é utilizado o processo iterativo BHHH para obter as estimativas de máxima verossimilhança.

2.5.2 Verificação

Para verificar a adequabilidade do modelo GARCH(p, q), com ε_t normal ou t-Student, onde os resíduos padronizados são dados por

$$\tilde{r}_t = \frac{r_t}{\sqrt{h_t}}$$

e são variáveis aleatórias i.i.d. com distribuição normal padrão ou t-Student, será calculado a estatística Q de Ljung-Box para a sequência dos resíduos padronizados, os coeficientes de assimetria e curtose estimados e fazer um gráfico qq (quantil \times quantil) para avaliar a suposição de normalidade. Para avaliar se ainda existe heteroscedasticidade condicional nos resíduos, pode-se aplicar o teste de Ljung-Box (verificar Apêndice 5.1 para

mais informações sobre o teste) para a sequência dos resíduos ao quadrado (\tilde{r}_t^2).

2.6 Modelo EGARCH

Para superar alguns pontos fracos do modelo GARCH em lidar com série financeiras, como a suposição de que a variância condicional não leva em consideração se o retorno é positivo ou negativo, isso pode ser visto na formulação do modelo pela Equação 2.2, onde os retornos são elevados ao quadrado na equação da variância. Devido a essa garantia de que os coeficientes serão não negativos, o modelo adquire uma limitação, a formulação da variância permanece não negativa e com probabilidade um, entretanto, muitas vezes esta imposição não é satisfeita ou dificulta muito a estimação dos parâmetros.

Visando solucionar esses problemas, Nelson (1991) estendeu o modelo GARCH para o modelo heteroscedástico condicional autoregressivo generalizado exponencial, exponencial GARCH (EGARCH), que trata os efeitos assimétricos entre retornos positivos e negativos, para isso ele considerou a inovação ponderada.

Definição 2.6 Um modelo EGARCH(p, q) é dado por:

$$r_t = \sqrt{h_t} \varepsilon_t, \quad (2.1)$$

$$\ln(h_t) = \alpha_0 + \frac{1 + b_1 B + \dots + b_p B^p}{1 - a_1 B - \dots - a_q B^q} g(\varepsilon_{t-1}), \quad (2.2)$$

onde α_0 é uma constante, B é o operador de retrocesso com raízes fora do círculo unitário, ou seja, os valores absolutos de $a(B)$ e $b(B)$ são, em módulo, maiores que um.

A função da curva de impacto de informação, $g(\varepsilon_t)$, é dada por:

$$g(\varepsilon_t) = \theta \varepsilon_t + \gamma [|\varepsilon_t| - E[|\varepsilon_t|]]. \quad (2.3)$$

Aqui, θ e γ são parâmetros reais, e $|\varepsilon_t| - E[|\varepsilon_t|]$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média zero. Então $E[g(\varepsilon_t)] = 0$ e a assimetria de $g(\varepsilon_t)$ pode ser escrita como

$$g(\varepsilon_t) = \begin{cases} (\theta + \gamma) \varepsilon_t - \gamma E[|\varepsilon_t|], & \text{se } \varepsilon_t \geq 0 \\ (\theta - \gamma) \varepsilon_t - \gamma E[|\varepsilon_t|], & \text{se } \varepsilon_t < 0 \end{cases}$$

A Equação 2.2 usa a parametrização usual ARMA para descrever a evolução da variância condicional de r_t . Algumas diferenças entre os modelos GARCH e EGARCH são, primeiro, note que em 2.2 usa-se o logaritmo de h_t , o que implica que os coeficientes do modelo não necessitam ser não negativos. O uso de $g(\varepsilon_t)$ permite que o modelo responda

de forma simétrica a valores positivos e negativos de r_t .

O modelo EGARCH original utiliza α_t ao invés de α_0 permitindo a variação no tempo. A forma utilizada no artigo de Nelson (1991) foi $\alpha_t = \alpha + \ln(1 + N_t \delta)$, onde N_t é o número de dias sem negociação entre $t - 1$ e t dia de negociação, e α e δ são parâmetros constantes. Dado isso, o modelo agora consegue captar a maior variação após os dias sem negociação (Tsay (2005)).

A Equação 2.2 pode ser escrita como:

$$\ln(h_t) = \alpha_* + \sum_{i=0}^p \beta_i g(\varepsilon_{t-i-1}) + \sum_{j=1}^q \alpha_j \ln(h_{t-j}), \quad (2.4)$$

onde $\alpha_* = \alpha_0(1 - \alpha_1 - \alpha_2 - \dots - \alpha_T)$ e $\beta_0 = 1$.

Substituindo a Equação 2.3 na Equação 2.4, supondo normalidade de ε_t , temos uma outra forma do modelo EGARCH(p,q),

$$\begin{aligned} \ln(h_t) &= \alpha_* + \sum_{i=0}^p \beta_i (\theta \varepsilon_{t-i-1} + \gamma_i [|\varepsilon_{t-i-1}| - E[|\varepsilon_{t-i-1}|]]) + \sum_{j=1}^q \alpha_j \ln(h_{t-j}) \\ &= \varphi + \sum_{i=0}^p \beta_i^* \left(\frac{\gamma_i^* r_{t-i-1} + |r_{t-i-1}|}{\sqrt{h_{t-i-1}}} \right) + \sum_{j=1}^q \alpha_j \ln(h_{t-j}), \end{aligned}$$

onde $\varphi = \alpha_* - \sum_{i=0}^p \beta_i \gamma_i \sqrt{2/\pi}$, $\gamma_i^* = \theta/\gamma_i$ e $\beta_i^* = \beta_i \gamma_i$.

Para que retorno negativos tenham maior impacto na volatilidade, esperamos $\gamma < 0$ na Equação 2.3. Para ε_t seguindo uma distribuição normal padrão, temos que:

$$\begin{aligned} E[|\varepsilon_t|] &= \int_{-\infty}^{+\infty} |\varepsilon_t| f(\varepsilon_t) d\varepsilon_t \\ &= \int_{-\infty}^{+\infty} |\varepsilon_t| \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{-\varepsilon_t^2}{2}\right\} d\varepsilon_t \\ &= 2 \int_0^{+\infty} \varepsilon_t \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{-\varepsilon_t^2}{2}\right\} d\varepsilon_t \\ \text{Substituindo } u &= \frac{-\varepsilon_t^2}{2} \rightarrow \frac{du}{d\varepsilon_t} = -\varepsilon_t \rightarrow d\varepsilon_t = -\frac{1}{\varepsilon_t} du \\ &= -2 \int_0^{-\infty} \frac{1}{\sqrt{2\pi}} \exp\{u\} du \\ &= \sqrt{2/\pi}. \end{aligned}$$

Essa assimetria permite que a volatilidade responda mais rapidamente a retornos negativos do que a positivos, fato este conhecido como "efeito alavancagem".

Para ε_t seguindo uma distribuição t-Student padrão com ν graus de liberdade, temos que:

$$\begin{aligned}
E[|\varepsilon_t|] &= \int_{-\infty}^{+\infty} |\varepsilon_t| f(\varepsilon_t) d\varepsilon_t \\
&= \int_{-\infty}^{+\infty} |\varepsilon_t| \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{(\nu-2)}\pi} \left(1 + \frac{\varepsilon_t^2}{(\nu-2)}\right)^{-(\nu+1)/2} d\varepsilon_t \\
&= 2 \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{(\nu-2)}\pi} \int_0^{+\infty} \varepsilon_t \left(1 + \frac{\varepsilon_t^2}{(\nu-2)}\right)^{-(\nu+1)/2} d\varepsilon_t \\
&= \frac{2\sqrt{\nu-2}\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi}(\nu-1)}.
\end{aligned}$$

Defina $\mathcal{R}_s = \{r_s, r_{s-1}, r_{s-2}, \dots\}$, a média e a variância do modelo EGARCH são dadas por:

$$E[r_t|\mathcal{R}_{t-1}] = E[\sqrt{h_t}\varepsilon_t|\mathcal{R}_{t-1}] = \sqrt{h_t} E[\varepsilon_t|\mathcal{R}_{t-1}] = 0$$

e

$$Var(r_t|\mathcal{R}_{t-1}) = Var(\sqrt{h_t}\varepsilon_t|\mathcal{R}_{t-1}) = E[r_t^2|\mathcal{R}_{t-1}] = E[h_t\varepsilon_t^2|\mathcal{R}_{t-1}] = h_t E[\varepsilon_t^2|\mathcal{R}_{t-1}] = h_t.$$

A covariância de r_t e r_{t+h}

$$\begin{aligned}
cov(r_t, r_{t+h}) &= E[r_t r_{t+h}] = E[E[r_t r_{t+h}|\mathcal{R}_{t+h-1}]] \\
&= E[r_t E[r_{t+h}|\mathcal{R}_{t+h-1}]] = E[r_t E[\sqrt{h_{t+h}}\varepsilon_{t+h}|\mathcal{R}_{t+h-1}]] \\
&= E[r_t \sqrt{h_{t+h}} E[\varepsilon_{t+h}|\mathcal{R}_{t+h-1}]] = 0
\end{aligned}$$

para $h > 0$, pois r_t é uma sequência de variáveis não correlacionadas (ruído branco), com média zero e variância condicional dada por h_t e r_t e $\sqrt{h_{t+h}}$ estão em \mathcal{F}_{t+h-1} e $E[\varepsilon_{t+h}|\mathcal{F}_{t+h-1}] = 0$

Um modelo EGARCH (1,1) é dado por:

$$\ln(h_t) = \alpha_0 + \alpha_1 g(\varepsilon_{t-1}) + \beta_1 \ln(h_{t-1}). \quad (2.5)$$

em que ε_t são variáveis aleatórias i.i.d. com média zero e $g(\cdot)$ como em 2.3.

2.6.1 Estimação

A estimação dos parâmetros do modelo EGARCH (p,q) dadas pelas equações 2.1 e 2.2 seguem o mesmo princípio da estimação para os parâmetros do modelo GARCH(p,q).

A função de máxima verossimilhança é dada por:

$$\begin{aligned}\mathcal{L}(r_1, \dots, r_T | \alpha, \beta, \gamma) &= f(r_T | F_{T-1}) f(r_{T-1} | F_{T-2}) \dots f(r_{p+1} | F_p) f(r_1, \dots, r_p | \alpha, \beta, \gamma) \\ &= \prod_{t=p+1}^T f(r_t | F_{t-1}) f(r_1, \dots, r_p | \alpha, \beta, \gamma),\end{aligned}\quad (2.6)$$

Os erros ε_t podem seguir qualquer distribuição com média zero e variância um, segundo a definição do modelo.

Distribuição Normal

Supondo que $\varepsilon_t \sim N(0, 1)$ temos que r_t segue uma distribuição com média zero e variância h_t dada por:

$$f(r_t) = \frac{1}{\sqrt{2\pi h_t}} \exp\left\{\frac{-r_t^2}{2h_t}\right\}.$$

Utilizando a verossimilhança dada por 2.6 e assumindo T grande, $f(r_1, \dots, r_p | \alpha, \beta, \gamma)$ pode ser desprezado. Então a verossimilhança condicionada aos p primeiros retornos é dada por:

$$\mathcal{L}(r_{p+1}, \dots, r_T | \alpha, \beta, \gamma, r_1, \dots, r_p) = \prod_{p+1}^T \left(\sqrt{2\pi h(t)}\right)^{-1} \exp\left\{\frac{-r_t^2}{2h_t}\right\}, \quad (2.7)$$

onde a volatilidade h_t é obtida recursivamente.

Tomando o logaritmo da Equação 2.7

$$\ell(r_{p+1}, \dots, r_T | \alpha, \beta, \gamma, r_1, \dots, r_p) \propto -\frac{1}{2} \sum_{t=p+1}^T \ln(h_t) - \frac{1}{2} \sum_{t=p+1}^T \frac{r_t^2}{h_t}. \quad (2.8)$$

Como feito nos artigos de Engle (1982) e Bollerslev (1986), utiliza o log-verossimilhança média com erros seguindo uma distribuição normal padrão para estimar os parâmetros do modelo (θ) dado por:

$$\ell(\theta) = \frac{1}{T} \sum_{t=1}^T -\frac{1}{2} \ln(h_t) - \frac{r_t^2}{2h_t}.$$

Distribuição t-Student

Suponha que $\varepsilon_t \sim t_\nu$, distribuição t-Student padronizada com ν graus de liberdade,

a função de densidade condicional de r_t é dada por:

$$f(r_t) = \frac{\Gamma((\nu + 1)/2)}{\Gamma(\nu/2)\sqrt{(\nu - 2)}\pi h_t} \left(1 + \frac{r_t^2}{(\nu - 2)h_t}\right)^{-(\nu+1)/2}.$$

Dada a verossimilhança dada pela Equação 2.6 e assumindo um T grande, $f(r_1, \dots, r_T, \alpha, \beta, \gamma)$ pode ser desprezado. Então a verossimilhança condicionada aos p primeiros retornos é dada por:

$$\mathcal{L}(r_{p+1}, \dots, r_T | \alpha, \beta, r_1, \dots, r_p) = \prod_{p+1}^T \frac{\Gamma((\nu + 1)/2)}{\Gamma(\nu/2)\sqrt{(\nu - 2)}\pi h_t} \left(1 + \frac{r_t^2}{(\nu - 2)h_t}\right)^{-(\nu+1)/2}. \quad (2.9)$$

onde a volatilidade h_t é obtida recursivamente.

Para encontrar os estimadores de máxima verossimilhança para os parâmetros é preciso maximizar a função de verossimilhança condicional. Neste caso, é possível estimar os parâmetros fixando um grau de liberdade ou estimar os graus de liberdade juntamente com os parâmetros. Em Bollerslev (1986), é utilizado o processo iterativo BHHH para obter as estimativas de máxima verossimilhança.

2.6.2 Verificação

Para verificar a adequabilidade do modelo EGARCH (p, q), com ε_t normal ou t-Student, onde os resíduos padronizados são dados por:

$$\tilde{r}_t = \frac{r_t}{\sqrt{h_t}}$$

e são variáveis aleatórias i.i.d. com distribuição normal padrão ou t-Student, será calculado a estatística Q de Ljung-Box para a sequência dos resíduos padronizados, os coeficientes de assimetria e curtose estimados e fazer um gráfico qq (quantil \times quantil) para avaliar a suposição de normalidade. Para avaliar se ainda existe heteroscedasticidade condicional nos resíduos, pode-se aplicar o teste de Ljung-Box (verificar Apêndice 5.1 para mais informações sobre o teste) para a sequência dos resíduos ao quadrado (\tilde{r}_t^2).

3 Aplicação

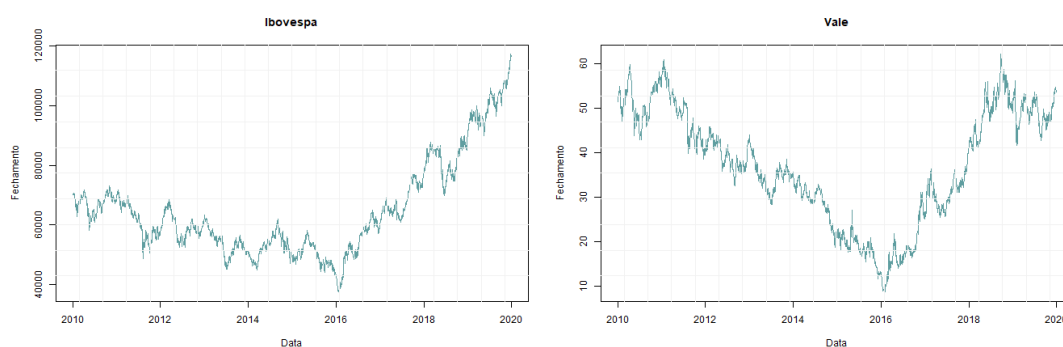
3.1 Séries Estudadas

Neste capítulo apresentamos um estudo sobre a acurácia do cálculo do VaR para os retornos do Ibovespa e para os retornos do preço das ações ordinárias da empresa Vale (ativo VALE3). O estudo contempla um período de 11 anos, de 04/01/2010 até 27/12/2019, totalizando 2479 observações.

Para o cálculo da volatilidade consideramos os modelos GARCH e EGARCH, ambos com duas variações para distribuição de probabilidade do termo aleatório, no caso, a distribuição Normal e a distribuição t-Student Assimétrica e Padronizada, essa última representada pela sigla SST.

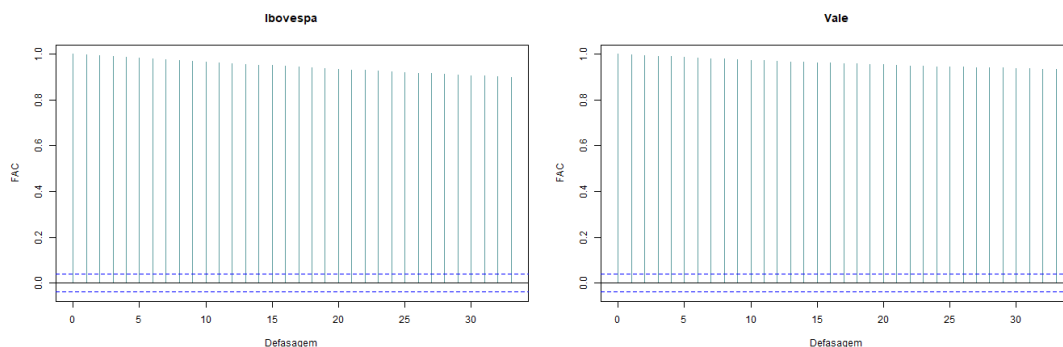
A Figura 3.1 mostra a não estacionaridade da séries, note como as séries oscilam bastante e a média e variância não são constantes ao longo de toda a série.

Figura 3.1: Preço de Fechamento



As Figuras a seguir apresentam os valores da função de autocorrelação e autocorrelação parcial para o conjunto de dados de fechamento.

Figura 3.2: Autocorrelação do Preço de Fechamento



Na Figura 3.2, tem-se que o decaimento da função de autocorrelação ocorre de forma bastante lenta, o que indica que o processo não é estacionário, como já mencionado anteriormente. Dessa forma, os modelos serão ajustados ao log-retorno dos dados. O resultado dessa transformação pode ser observado a seguir:

Tabela 2: Medidas de Resumo - log-retorno

Medida	Ibovespa	Vale
Mínimo	-0,092	-0,282
Máximo	0,064	0,138
Média	0,000	0,000
Variância	0,000	0,001
Desvio Padrão	0,014	0,026

Figura 3.3: log-Retorno - Ibovespa

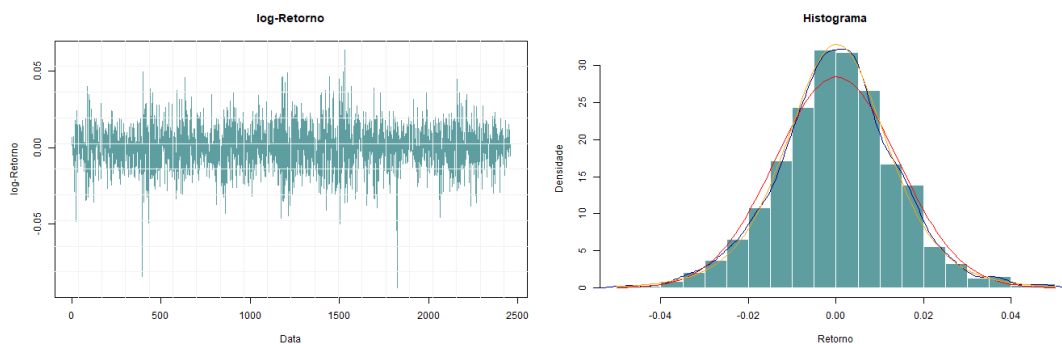
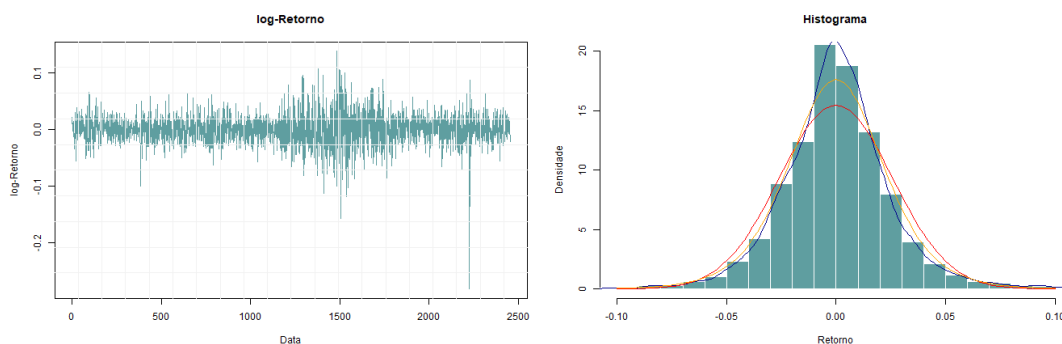


Figura 3.4: log-Retorno - Vale



No histograma das Figuras 3.3 e 3.4, a linha azul representa a densidade dos dados log-retorno e a linha vermelha e laranja representam as densidades das distribuições Normal e SST, respectivamente, com média e variância igual às apresentadas na Tabela 2. Nos histogramas nota-se que a distribuição SST parece se ajustar melhor aos dados

pois estes possuem uma distribuição mais leptocúrtica do que a distribuição Normal consegue capturar. Como os retornos possuem volatilidade, ou seja, não são independentes e identicamente distribuídas (i.i.d.) os ajustes não ficaram muito bons.

Abaixo seguem os gráficos dos quantis para ajudar na visualização do ajuste da distribuição Normal e SST nos dados.

Figura 3.5: Quantis do log-retorno - Ibovespa

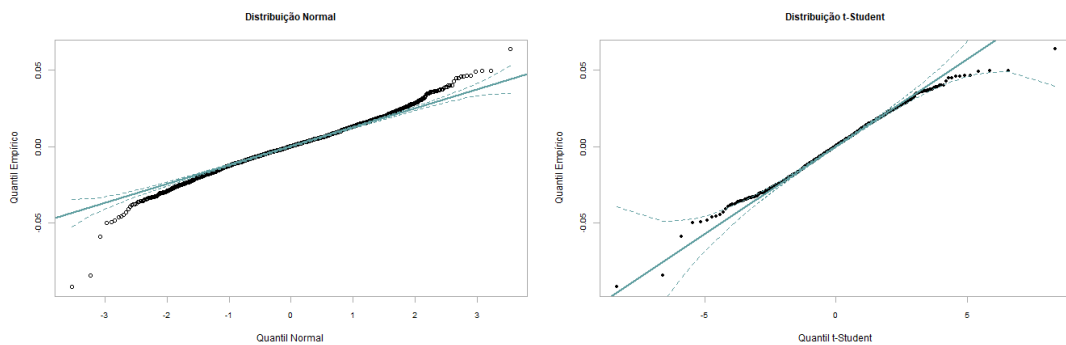
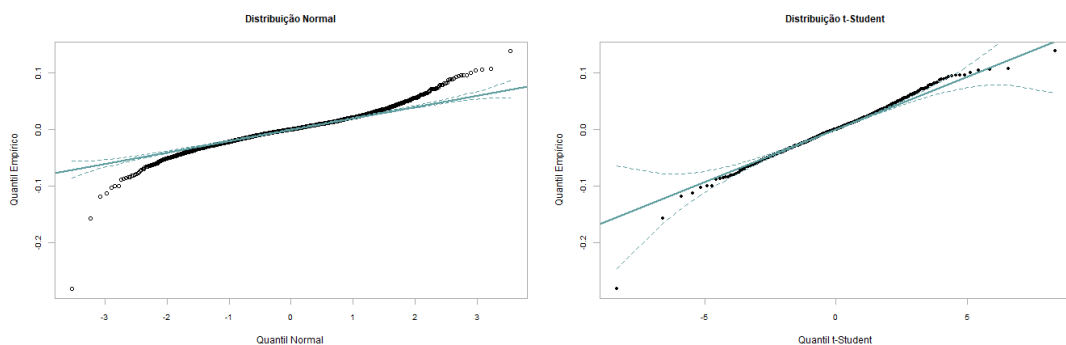
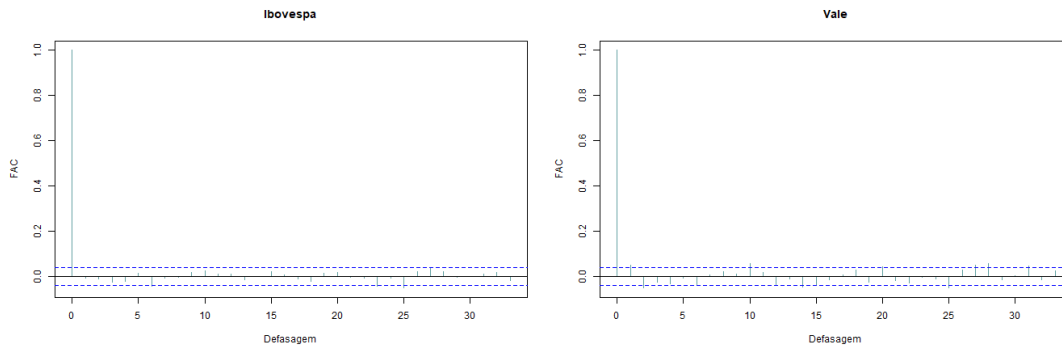


Figura 3.6: Quantis do log-retorno - Vale

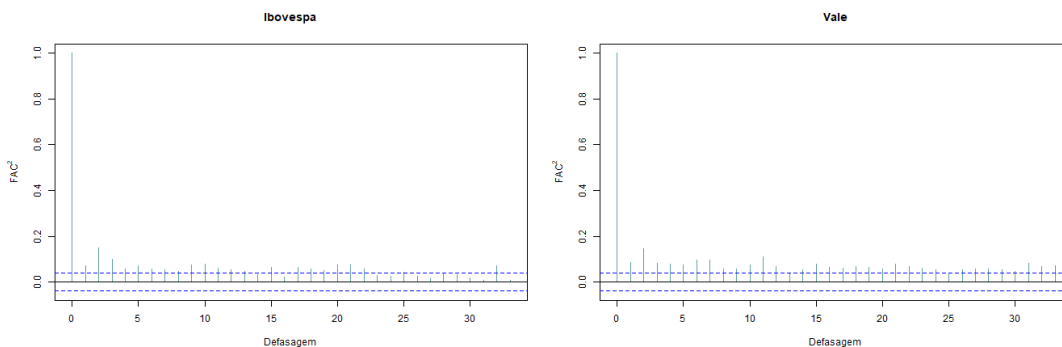


Nas Figuras 3.5 e 3.6 podemos ver que nenhuma das distribuições possuem um bom ajuste para as caudas, vale ressaltar que a distribuição Normal não consegue capturar as caudas por subestimar seu peso. Para avaliar a estacionariedade, construiu-se os gráficos das funções de autocorrelação e autocorrelação parcial do log-retorno, como segue a Figura 3.7 abaixo.

Figura 3.7: Autocorrelação do log-retorno



É possível perceber que não parece ser necessário ajustar uma estrutura autorregressiva ao log-retorno dos dados, pois não mostra autocorrelação, já a Figura 3.8 mostra uma estrutura de autocorrelação nas defasagens iniciais mas é amortizado ao longo do tempo.

Figura 3.8: Autocorrelação do log-retorno²

Para verificar se há efeito heterocedástico nos dados de log-retorno é realizado o teste de Ljung-Box, para maiores informações sobre o teste, consultar Apêndice 5.1. Considerando as seguintes hipóteses:

$$\begin{cases} H_0 : \text{Não há autocorrelação, dados independentes} \\ H_1 : \text{Há autocorrelação, dados não são independentes.} \end{cases}$$

Para o log-retorno o p-valor foi igual a 0,6591 para a série do log-retorno Ibovespa e de 0,05127 para a série do log-retorno Vale, indicando que não há autocorrelação nos dados ao nível de significância de 5%, já para o log-retornos ao quadrado, resultou em p-valor igual a $2,2 \times 10^{-16}$, isso para os dois banco de dados. Ao nível de confiança de 5%, rejeita-se a hipótese nula, ou seja, o efeito de correlação nos dados do log-retorno ao quadrado é estatisticamente significativo, como mostra a Figura 3.8. Assim, modelos da família heterocedástica são apropriados para modelar os dados do log-retorno.

3.2 Ajuste dos Modelos

3.2.1 GARCH

O primeiro modelo ajustado será o modelo GARCH (1, 1). As Tabelas 3 e 4 a seguir apresenta as estimativas dos parâmetros e seus respectivos desvio-padrão para o modelo.

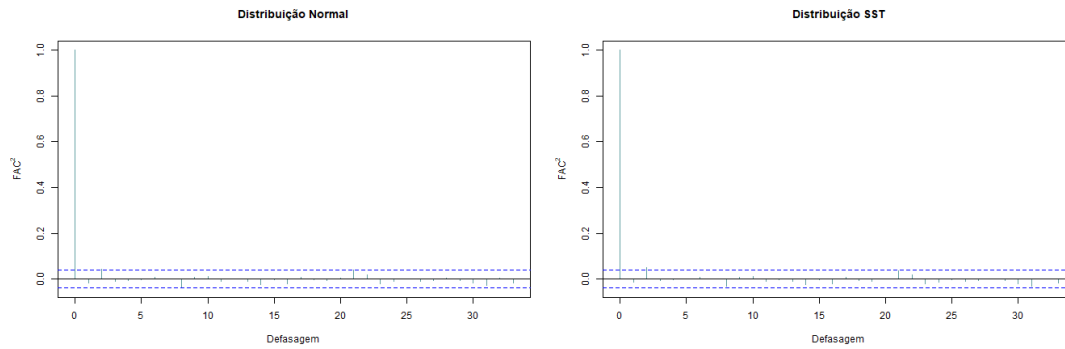
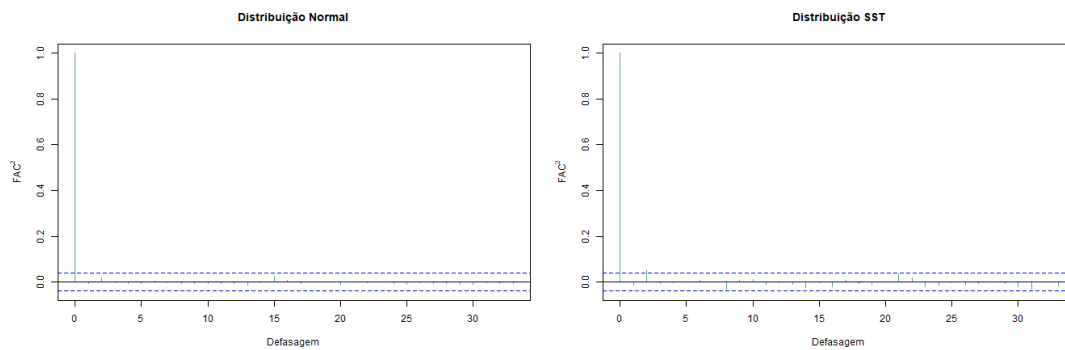
Tabela 3: Estimativas do Modelo GARCH (1,1)- Ibovespa

Parâmetros	Distribuição Normal		Distribuição SST	
	Estimativa	Desvio Padrão	Estimativa	Desvio Padrão
μ	0,00033	0,00026	0,00032	0,00026
α_0	0,00001	0,00000	0,00001	0,00000
α_1	0,06203	0,00432	0,05509	0,00456
β_1	0,90150	0,00725	0,91712	0,00871
Inclinação			0,96846	0,02778
Forma			10,33786	1,82541

Tabela 4: Estimativas do Modelo GARCH (1,1) - Vale

Parâmetros	Distribuição Normal		Distribuição SST	
	Estimativa	Desvio Padrão	Estimativa	Desvio Padrão
μ	-0,00011	0,00043	0,00003	0,00042
α_0	0,00000	0,00000	0,00000	0,00000
α_1	0,03798	0,00375	0,04584	0,01063
β_1	0,95580	0,00357	0,94711	0,01272
Inclinação			1,00043	0,02773
Forma			7,06932	0,96088

A Figura 3.9 apresenta os gráficos de autocorrelação para os resíduos padronizados do modelo GARCH com distribuição de probabilidade Normal e distribuição de probabilidade SST, respectivamente para os dados Ibovespa. E a Figura 3.10 para o dados Vale.

Figura 3.9: Autocorrelação dos resíduos padronizados² - IbovespaFigura 3.10: Autocorrelação dos resíduos padronizados² - Vale

Podemos notar pelos gráficos que para as duas distribuições não há autocorrelação presente na variância, visto que nenhuma defasagem foi significativa, ou seja, o modelo conseguiu captar toda a volatilidade. Este fato é comprovado ao aplicar o teste de Ljung-Box para o quadrado dos resíduos padronizados, que resulta em p-valor igual a 0,5879 para a distribuição Normal e p-valor igual a 0,5222 para a distribuição SST, para os dados do Ibovespa e para os dados da Vale, resultam em p-valores iguais a 0,9927 e 0,5222 para a distribuição Normal e SST, respectivamente. Ou seja, não há evidências estatísticas para a rejeição da hipótese nula, ao nível de significância de 5%. Assim, não há autocorrelação presente nos resíduos padronizados.

A Tabela a seguir contém medidas descritivas sobre os resíduos que podem auxiliar na validação da adequabilidade das supostas distribuições.

Tabela 5: Medidas de Resumo - Resíduos Padronizados do modelo GARCH(1,1)

Parâmetros	Ibovespa		Vale	
	Normal	SST	Normal	SST
Média	-0,018	-0,002	-0,002	-0,009
Variância	0,999	1,005	0,996	1,04
Desvio Padrão	1	1,002	0,998	1,02

Figura 3.11: Resíduos padronizados do modelo GARCH (1, 1) - Ibovespa

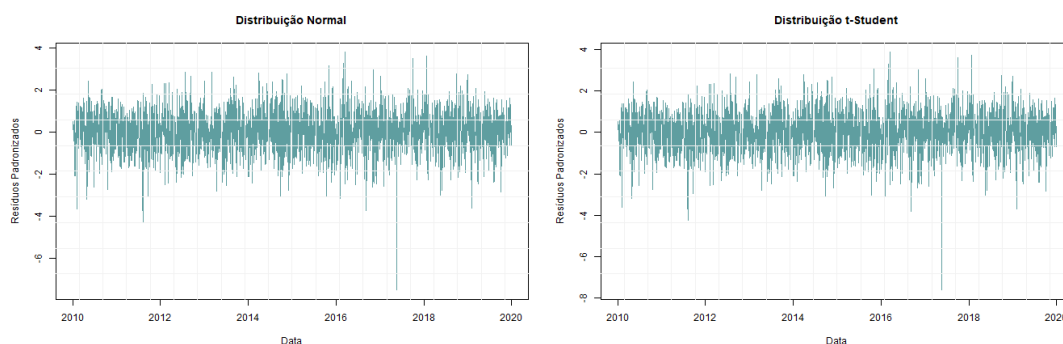
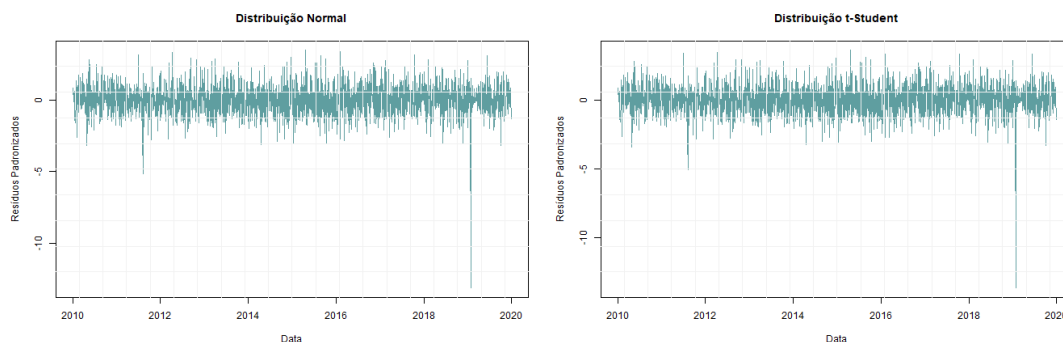


Figura 3.12: Resíduos padronizados do modelo GARCH (1, 1) - Vale



A suposição do modelo é que a média é zero e a variância é igual a um na distribuição dos resíduos. Observando a Tabela 5 podemos ver que tanto para a distribuição Normal quanto para a SST a média e a variância ficaram bem próximos dos valores supostos. Nos histogramas abaixo é possível visualizar tal informação.

Figura 3.13: Histograma dos resíduos padronizados do modelo GARCH (1, 1) - Ibovespa

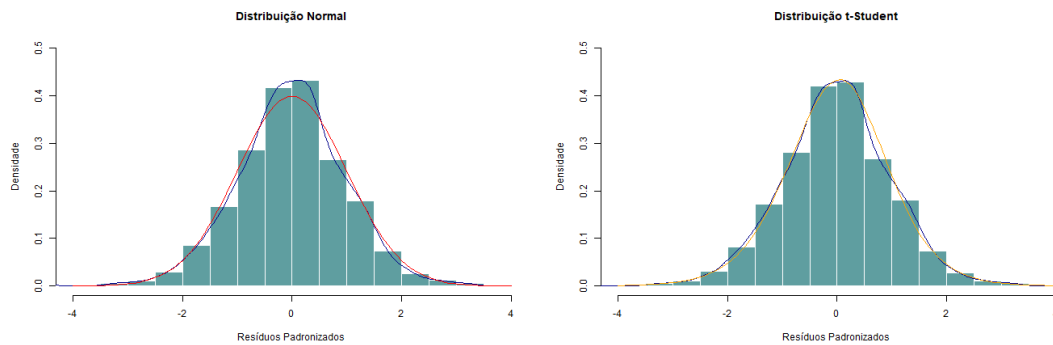
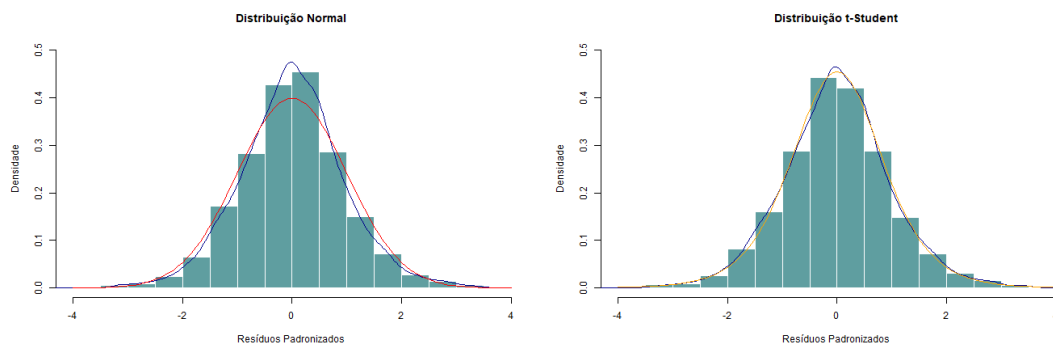


Figura 3.14: Histograma dos resíduos padronizados do modelo GARCH (1, 1) - Vale



A linha azul representa a densidade dos dados e as linhas vermelha e laranja representam as distribuições Normal e SST, respectivamente.

Figura 3.15: Quantil dos resíduos padronizados do modelo GARCH (1, 1) - Ibovespa

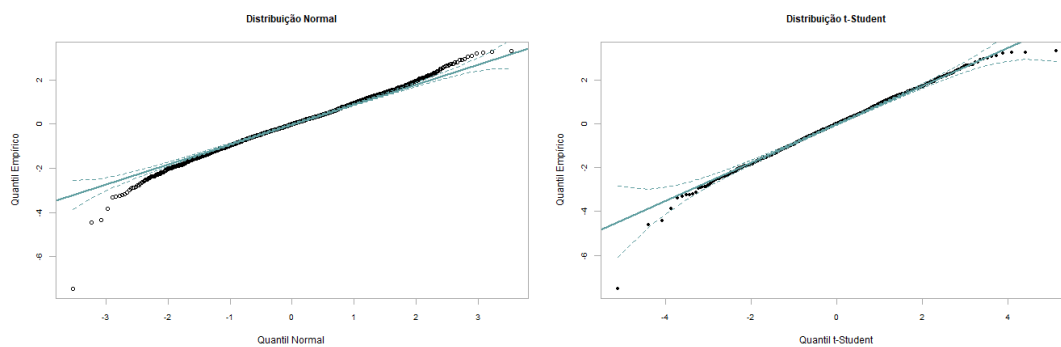
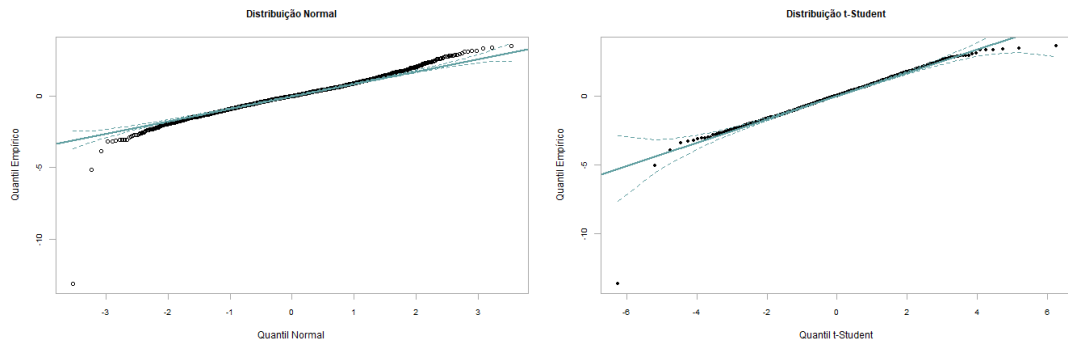


Figura 3.16: Quantil dos resíduos padronizados do modelo GARCH (1, 1) - Vale



As Figuras 3.15 e 3.16 mostram que o ajuste ficou melhor para a distribuição SST, a distribuição Normal parece não conseguir capturar as caudas dos dados. Abaixo segue o teste de Kolmogorov-Smirnov para auxiliar na interpretação dos gráficos.

Fazendo o teste de Kolmogorov-Smirnov que pode ser utilizado para avaliar as hipóteses:

$$\begin{cases} H_0 : \text{os dados seguem a distribuição proposta} \\ H_1 : \text{os dados não seguem a distribuição proposta} \end{cases}$$

Este teste observa a máxima diferença absoluta entre a função de distribuição acumulada assumida para os dados, no caso a Normal e a SST, e a função de distribuição empírica dos dados. Como critério, comparamos esta diferença com um valor crítico, para um dado nível de significância.

O teste supondo que os resíduos padronizados seguem uma distribuição Normal, considerando os dados do Ibovespa, resultou no p-valor igual a 0,03188 ou seja, ao nível de significância $\alpha = 0,05$ há evidência estatística para rejeitar a hipótese nula, concluindo que os dados não seguem a distribuição normal. Realizando o mesmo processo, mas agora supondo que os dados seguem uma distribuição SST, resultou em um p-valor 0,372, ou seja, ao nível de significância $\alpha = 0,05$ não há evidência estatística para rejeitar a hipótese nula, concluindo que os dados seguem a distribuição SST.

Para os dados Vale, os testes resultaram nos seguintes p-valores, 0,002187 e 0,4417, ou seja, há evidência estatística para rejeitar a hipótese nula e considerar que os dados não seguem uma distribuição Normal e há evidências estatísticas para concluir que os dados seguem a distribuição SST, respectivamente.

3.2.2 EGARCH

O segundo modelo ajustado será o modelo EGARCH (1, 1). As Tabelas a seguir apresentam as estimativas dos parâmetros e seus respectivos desvios-padrão para o mo-

delo.

Tabela 6: Estimativas do Modelo EGARCH (1,1) - Ibovespa

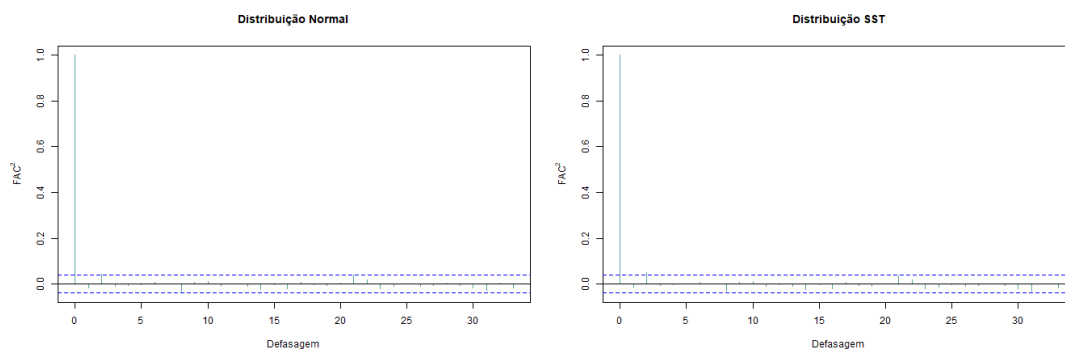
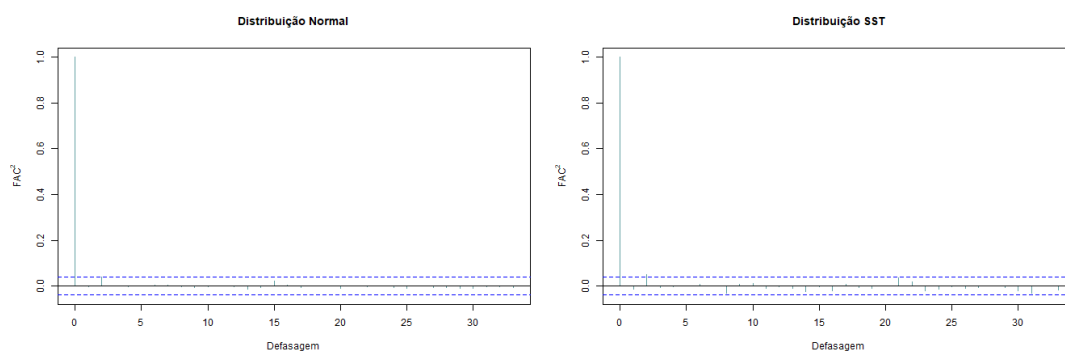
Parâmetros	Distribuição Normal		Distribuição SST	
	Estimativa	Desvio Padrão	Estimativa	Desvio Padrão
μ	0,00004	0,00026	0,00010	0,00026
α_0	-0,27799	0,00728	-0,22910	0,00186
α_1	-0,06530	0,01057	-0,07005	0,01162
β	0,96739	0,00097	0,97344	0,00023
γ	0,10565	0,00863	0,09476	0,00415
Inclinação			0,96497	0,02792
Forma			10,14668	1,79281

Tabela 7: Estimativas do Modelo EGARCH (1,1)- Vale

Parâmetros	Distribuição Normal		Distribuição SST	
	Estimativa	Desvio Padrão	Estimativa	Desvio Padrão
μ	-0,00032	0,00043	-0,00010	0,00042
α_0	-0,03947	0,00041	-0,03501	0,00105
α_1	-0,01654	0,00697	-0,02133	0,00889
β	0,99425	0,00010	0,99540	0,00026
γ	0,068635	0,00329	0,08486	0,00775
Inclinação			0,99593	0,02777
Forma			7,12842	0,68997

Pelas estimativas é possível notar que não há efeito de alavancagem, pois o parâmetro γ é positivo e significativo para as duas distribuições. As estimativas indicam que retornos passados tem forte impacto na volatilidade futura. O efeito de assimetria é negativo, ou seja, retornos negativos tem um maior efeito sobre a volatilidade do que retornos positivos.

A Figura 3.17 apresenta os gráficos de autocorrelação para os resíduos padronizados do modelo EGARCH com distribuição de probabilidade Normal e distribuição de probabilidade SST, respectivamente para os dados Ibovespa. E a Figura 3.18 para o dados Vale.

Figura 3.17: Autocorrelação dos resíduos padronizados² - IbovespaFigura 3.18: Autocorrelação dos resíduos padronizados² - Vale

Podemos notar pelos gráficos que para as duas distribuições não há autocorrelação presente na variância condicional, visto que nenhuma defasagem foi significativa, ou seja, o modelo conseguiu captar toda a volatilidade. Este fato é comprovado ao aplicar o teste de Ljung-Box para o quadrado dos resíduos padronizados, que resultaram em p-valor iguais a 0,5879 e 0,5222 para as distribuições Normal e SST, respectivamente para os dados do Ibovespa.

Para os dados da Vale resultaram em p-valores iguais a 0,9462 e 0,5222 para as distribuições Normal e SST, respectivamente, ou seja, não há evidências estatísticas para a rejeição da hipótese nula, não há autocorrelação presente nos resíduos padronizados.

A Tabela a seguir contém medidas descritivas sobre os resíduos que podem auxiliar na validação da adequabilidade das supostas distribuições.

Tabela 8: Medidas de Resumo - Resíduos Padronizados do modelo EGARCH(1,1)

Parâmetros	Ibovespa		Vale	
	Normal	SST	Normal	SST
Média	0,003	-0,002	0,005	-0,005
Variância	0,999	1,005	0,996	1,046
Desvio Padrão	1	1,002	0,998	1,023

Figura 3.19: Resíduos padronizados do modelo EGARCH (1, 1) - Ibovespa

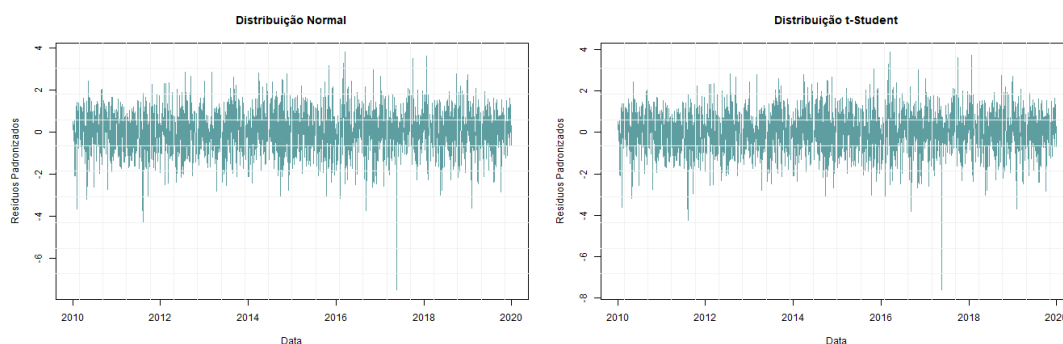
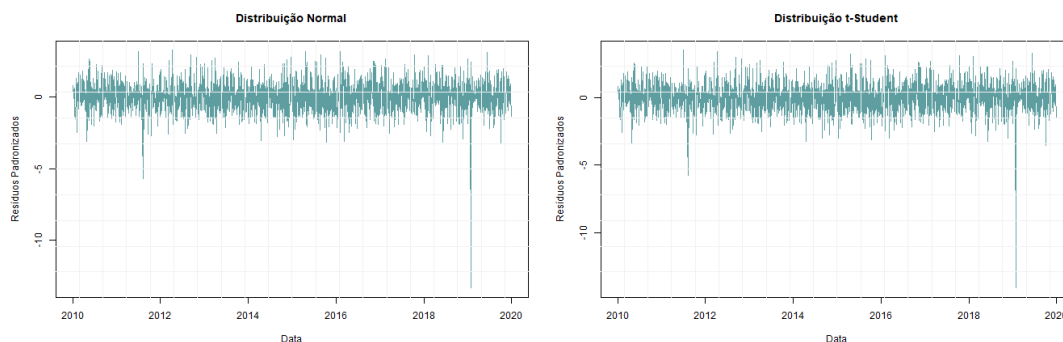


Figura 3.20: Resíduos padronizados do modelo EGARCH (1, 1) - Vale



A suposição do modelo é que a média é zero e a variância é igual a um na distribuição dos resíduos. Observando a Tabela 8 podemos ver que tanto para a distribuição Normal quanto para a SST a média e a variância ficaram bem próximos dos valores supostos. Nos histogramas abaixo é possível visualizar tal informação.

Figura 3.21: Histograma dos resíduos padronizados do modelo EGARCH (1, 1) - Ibovespa

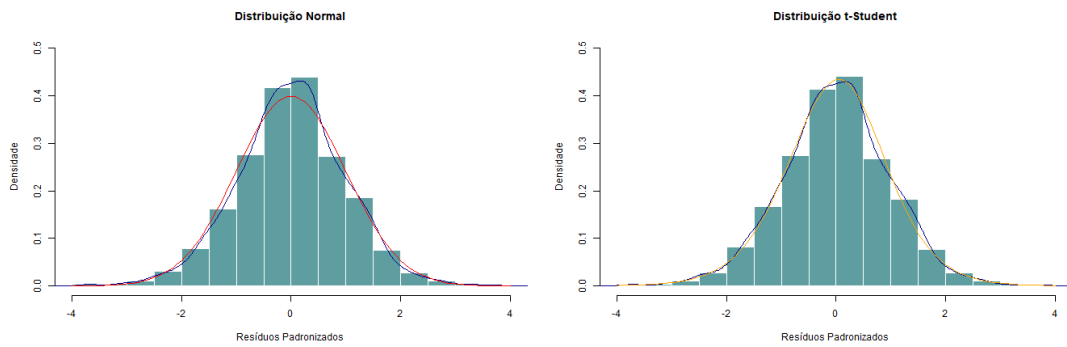
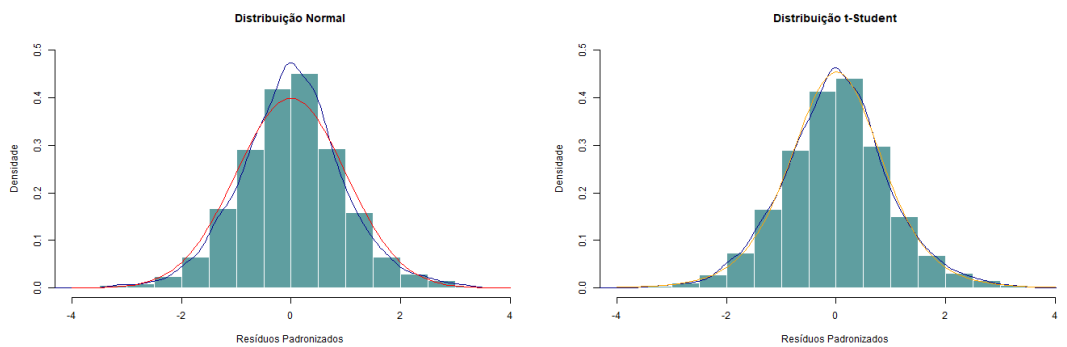


Figura 3.22: Histograma dos resíduos padronizados do modelo EGARCH (1, 1) - Vale



A linha azul representa a densidade dos dados e as linhas vermelha e laranja representam as distribuições Normal e SST, respectivamente.

Figura 3.23: Quantil dos resíduos padronizados do modelo EGARCH (1, 1) - Ibovespa

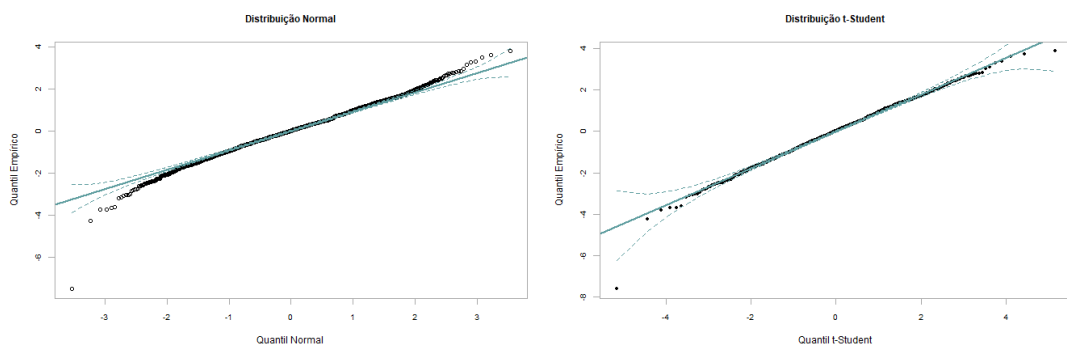
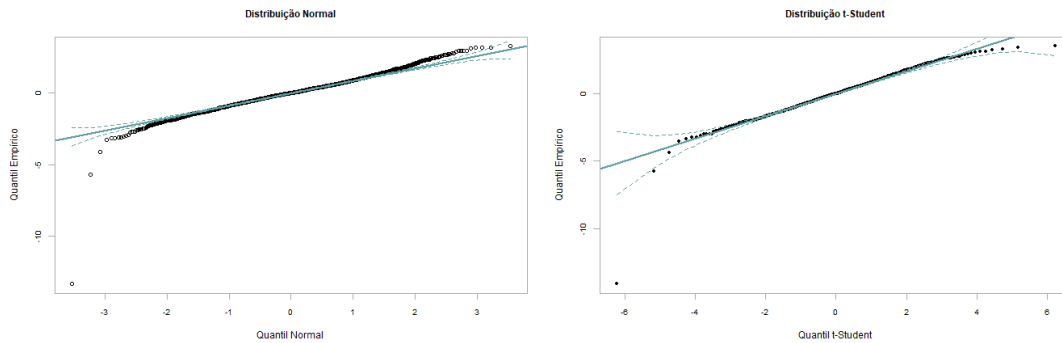


Figura 3.24: Quantil dos resíduos padronizados do modelo EGARCH (1, 1) - Vale



As Figuras 3.23 e 3.24 mostram que o ajuste ficou melhor para a distribuição SST, a distribuição Normal parece não conseguir capturar as caudas dos dados. Abaixo segue o teste de Kolmogorov-Smirnov para auxiliar na interpretação dos gráficos.

Fazendo o teste de Kolmogorov-Smirnov que pode ser utilizado para avaliar as hipóteses:

$$\begin{cases} H_0 : \text{os dados seguem a distribuição proposta} \\ H_1 : \text{os dados não seguem a distribuição proposta} \end{cases}$$

Este teste observa a máxima diferença absoluta entre a função de distribuição acumulada assumida para os dados, no caso a Normal e a SST, e a função de distribuição empírica dos dados. Como critério, comparamos esta diferença com um valor crítico, para um dado nível de significância.

O teste supondo que os resíduos padronizados seguem uma distribuição Normal, considerando os dados do Ibovespa, resultou no p-valor igual a 0,04523 ou seja, ao nível de significância $\alpha = 0,05$ há evidência estatística para rejeitar a hipótese nula, concluindo que os dados não seguem a distribuição normal. Realizando o mesmo processo, mas agora supondo que os dados seguem uma distribuição SST, resultou em um p-valor 0,6738, ou seja, ao nível de significância $\alpha = 0,05$ não há evidência estatística para rejeitar a hipótese nula, concluindo que os dados seguem a distribuição SST.

Para os dados Vale, os testes resultaram nos seguintes p-valores, 0,002582 e 0,4609, ou seja, há evidência estatística para rejeitar a hipótese nula e considerar que os dados não seguem uma distribuição Normal e há evidências estatísticas para concluir que os dados seguem a distribuição SST, respectivamente.

Com o intuito de comparar os modelos e as distribuições apresentadas, construiu-se as seguintes tabelas.

Tabela 9: Comparação entre modelos - Ibovespa

Medidas	GARCH (1, 1)		EGARCH (1, 1)	
	Normal	SST	Normal	SST
Log-verossimilhança	7070,31	7100,22	7083,129	7113,884
AIC	-5,7543	-5,7771	-5,7639	-5,7874
BIC	-5,7449	-5,7629	-5,7521	-5,7708

Tabela 10: Comparação entre modelos - Vale

Medidas	GARCH (1, 1)		EGARCH (1, 1)	
	Normal	SST	Normal	SST
Log-verossimilhança	5722.223	5823.078	5724.4	5830.508
AIC	-4.6603	-4.7409	-4.6613	-4.7461
BIC	-4.6509	-4.7267	-4.6495	-4.7296

Observando as Tabelas 9 e 10 é possível perceber que os resultados foram bastante semelhantes, mas como mencionado anteriormente, os dados não seguem uma distribuição Normal para nenhum dos modelos, assim, o "melhor" modelo é o EGARCH (1, 1) pois apresenta os menores AIC e BIC e o maior valor para o log-verossimilhança.

3.3 Acurácia do VaR

A terceira parte é a estimativa do VaR. Foi utilizado a função *ugarchroll* do pacote 'rugrarch' que fornece o ajuste do processo. Na função foram utilizados os seguintes parâmetros:

- *n.start*: considerando um total de 2456 observações, começamos no ponto 1456, ficando com um número de previsões igual a 1000.
- *refit.every*: esse parâmetro determina a frequência em que o modelo será reestimado. Se a escolha do número for muito grande, por exemplo 30, o modelo se ajustará aos dados de forma muito lenta e pode não conseguir capturar a volatilidade em períodos curtos. Se o valor escolhido for muito pequeno o processo será bastante lento dado a quantidade de vezes que o modelo precisará ser reestimado. Um número em torno de 10 a 15 é o mais adequado.
- *refit.window*: foi usado a forma 'moving', ou seja, todos os dados serão usados na primeira estimativa do modelo e ir se movendo de acordo com o tamanho definido. Tem também a forma recursiva 'recursive', em que a estimação é feita em uma janela

incluindo todos os dados.

- *solve*: usando a opção 'hybrid' oferece um processo ergódico, tentando todos os solucionadores, incluindo 'solnp', 'nlminb', 'gosolnp' e 'Nloptr', o que resultar na melhor chance de ajustar o modelo aos dados.

3.3.1 GARCH

Primeiramente, considere a análise para o modelo GARCH. As Figuras 3.25 e 3.26 abaixo mostram a volatilidade estimada utilizando as duas distribuições Normal e SST para os dois banco de dados. O comportamento da volatilidade das duas distribuições é bastante parecido.

Figura 3.25: Volatilidade do modelo GARCH (1, 1) - Ibovespa

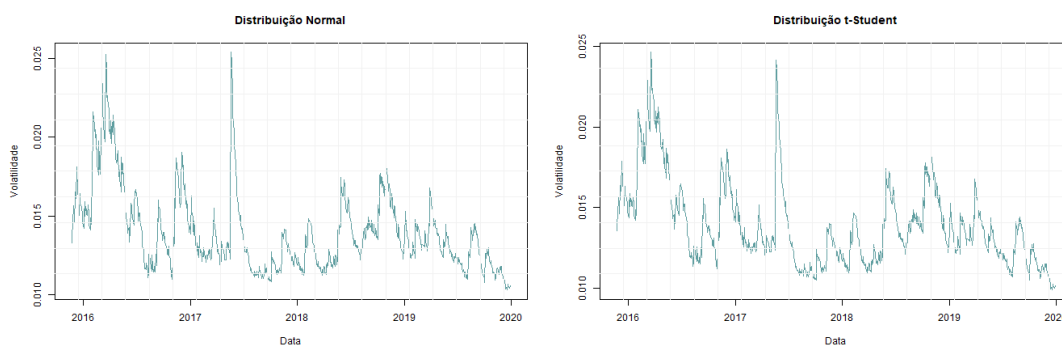
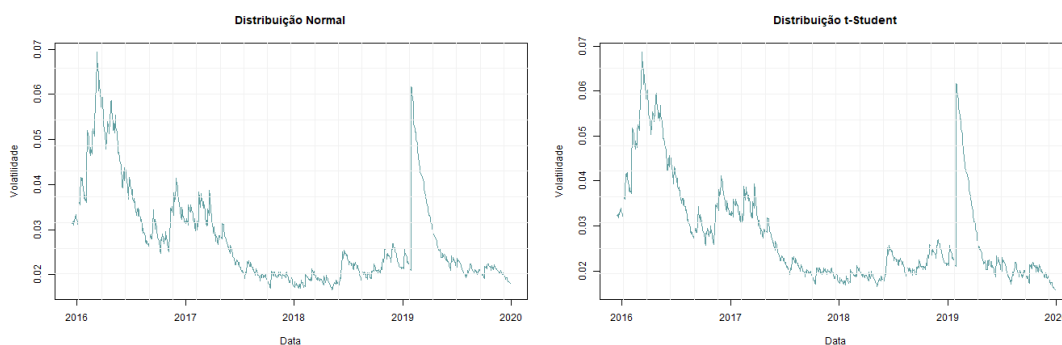


Figura 3.26: Volatilidade do modelo GARCH (1, 1) - Vale



Abaixo seguem os gráficos da cobertura do VaR considerando três níveis de significância.

Figura 3.27: log-Retorno e Violações do VaR - Ibovespa

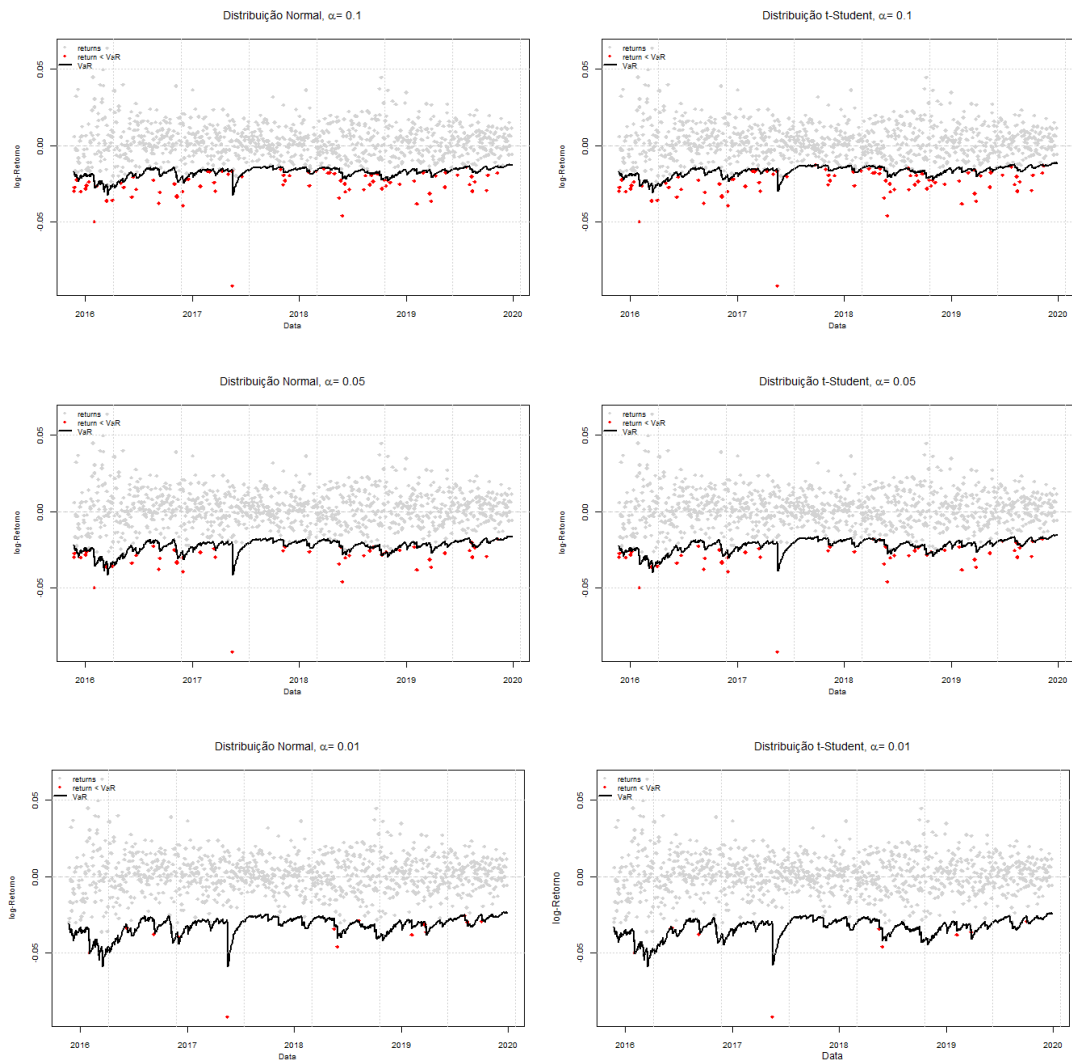
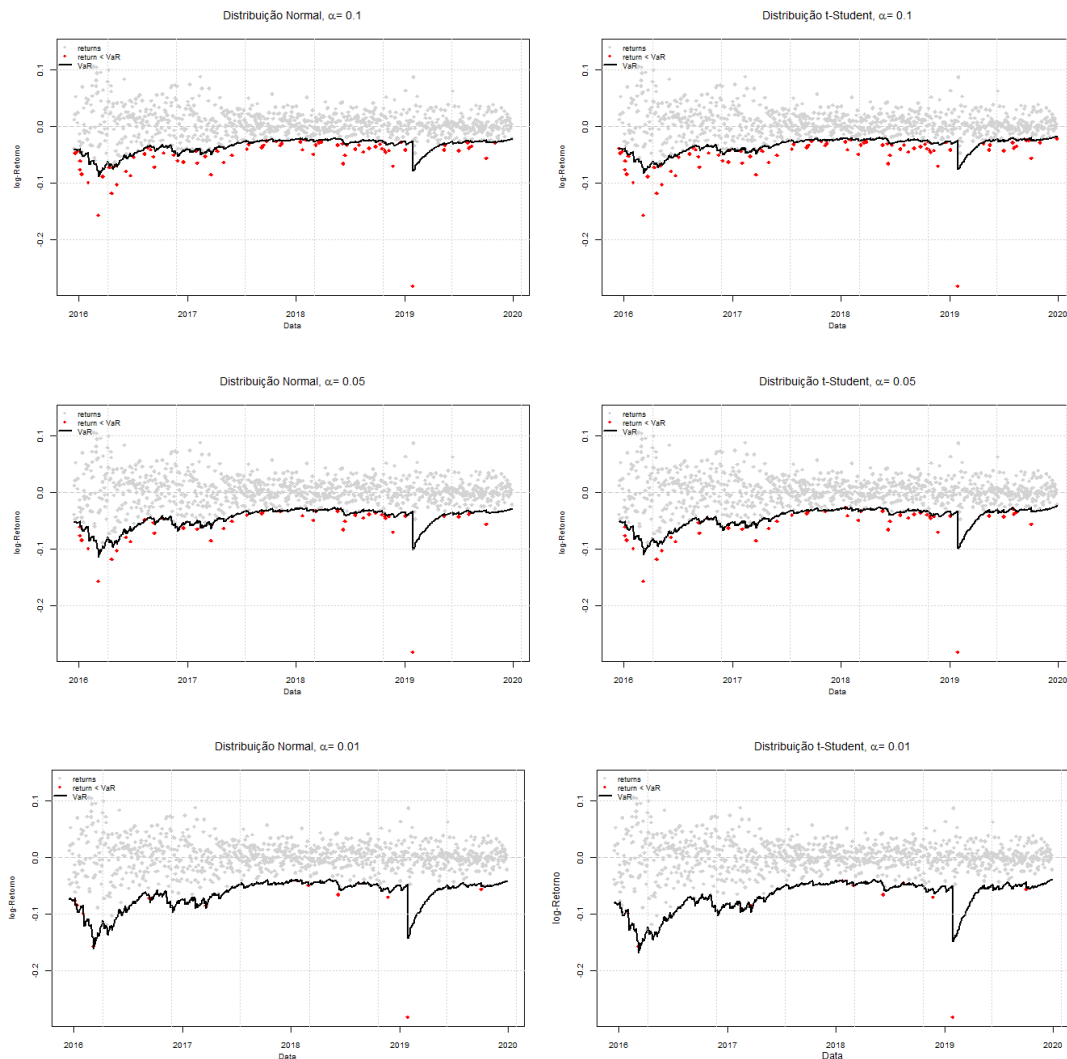


Figura 3.28: log-Retorno e Violações do VaR - Vale



A Figura 3.27 representa os lucros e perdas diárias do Ibovespa contra o VaR (negativo) e a Figura 3.28 representa os lucros e perdas diárias da Vale contra o VaR (negativo). Os gráfico não apenas resumem o desempenho do ativo e a evolução de seu risco de mercado como fornece uma análise gráfica simples do desempenho da medida de valor em risco.

Com uma métrica do VaR de 90%, 95% e 99% em um dia, espera-se que as perdas excedam o VaR em aproximadamente 10%, 5% e 1% do tempo, respectivamente.

Portanto, tomando o grupo de séries com o mesmo tamanho amostral e avaliando a cobertura do VaR, que é definida como a frequência das excedências do VaR, quando o real retorno é menor do que o VaR previsto. Para um modelo de previsão ser considerado válido, a cobertura VaR tem que ficar em torno da probabilidade de perda α . Abaixo segue Tabela 11 com os valores de cobertura para as duas distribuições e para os α 's.

Tabela 11: Coberturas do modelo GARCH(1,1)

α	Ibovespa		Vale	
	Normal	SST	Normal	SST
10%	0,084	0,09	0,076	0,083
5%	0,048	0,05	0,044	0,046
1%	0,013	0,009	0,013	0,01

De acordo com os valores de cobertura apresentados na Tabela 11, tanto para os dados Ibovespa quanto para os dados da Vale, a distribuição SST apresentou melhores resultados. Para os dados do Ibovespa, o nível de significância $\alpha = 0,05$ foi o único em que o número de violações observadas foi igual ao nível de significância do VaR. Já para os dados da Vale, o nível de significância $\alpha = 0,01$ foi o único em que o número de violações observadas foi igual ao nível de significância do VaR. Os outros resultados ficaram bem próximo do nível de significância do valor de risco para os dois banco de dados.

Realizando o teste de cobertura, considerando as seguintes hipóteses:

$$\begin{cases} H_0 : q = q^* \\ H_1 : q \neq q^* \end{cases}$$

onde q é o quantil do VaR e q^* é a frequência de violações. A hipótese nula testa se o modelo foi bem ajustado. A Tabela 13 a seguir contém os resultados.

Tabela 12: Teste de cobertura do modelo GARCH(1,1)

α	Ibovespa		Vale	
	Normal	SST	Normal	SST
10%	Rejeita H_0	Não rejeita H_0	Rejeita H_0	Rejeita H_0
5%	Não rejeita H_0	Não rejeita H_0	Não Rejeita H_0	Não Rejeita H_0
1%	Não rejeita H_0	Não rejeita H_0	Não Rejeita H_0	Não Rejeita H_0

A distribuição SST possui os valores de cobertura mais próximos do que a distribuição Normal. Para os dados do Ibovespa não rejeita H_0 para nenhum valor de α , para os dados da Vale rejeita apenas ao nível de significância $\alpha = 0,1$.

3.3.2 EGARCH

Agora, considere a análise para o modelo EGARCH. As Figuras 3.29 e 3.30 abaixo mostram a volatilidade estimada utilizando as duas distribuições Normal e SST para os dois banco de dados. O comportamento da volatilidade das duas distribuições são bastante

parecidos.

Figura 3.29: Volatilidade do modelo EGARCH (1, 1) - Ibovespa

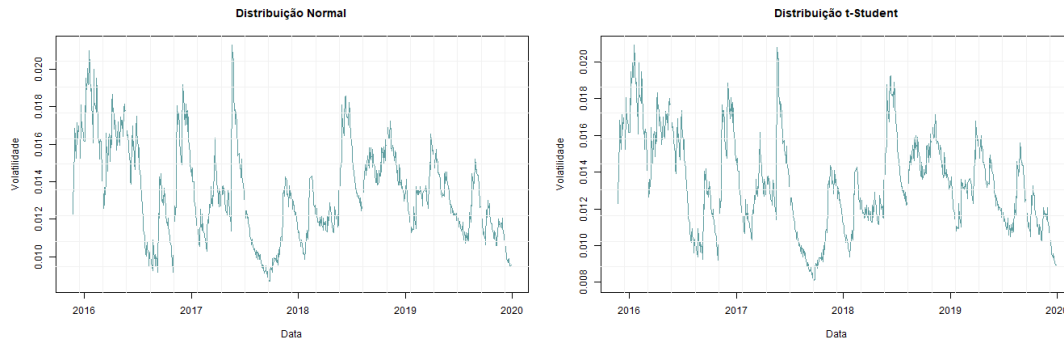
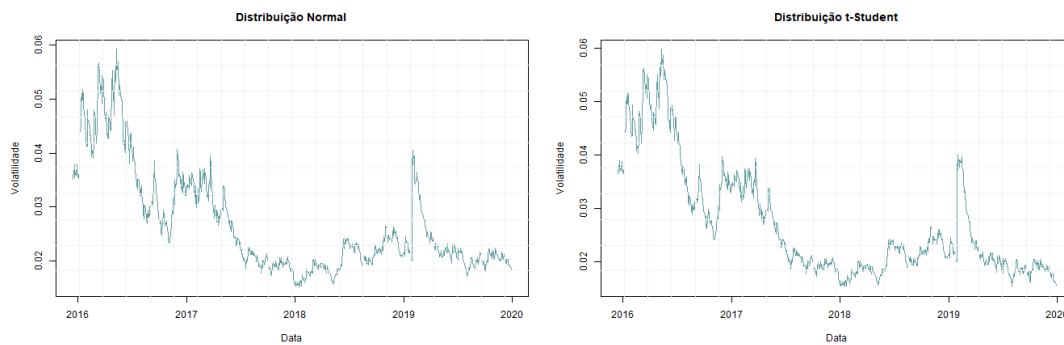


Figura 3.30: Volatilidade do modelo EGARCH (1, 1) - Vale



Abaixo seguem os gráficos da cobertura do VaR considerando três níveis de significância.

Figura 3.31: log-Retorno e Violações do VaR - Ibovespa

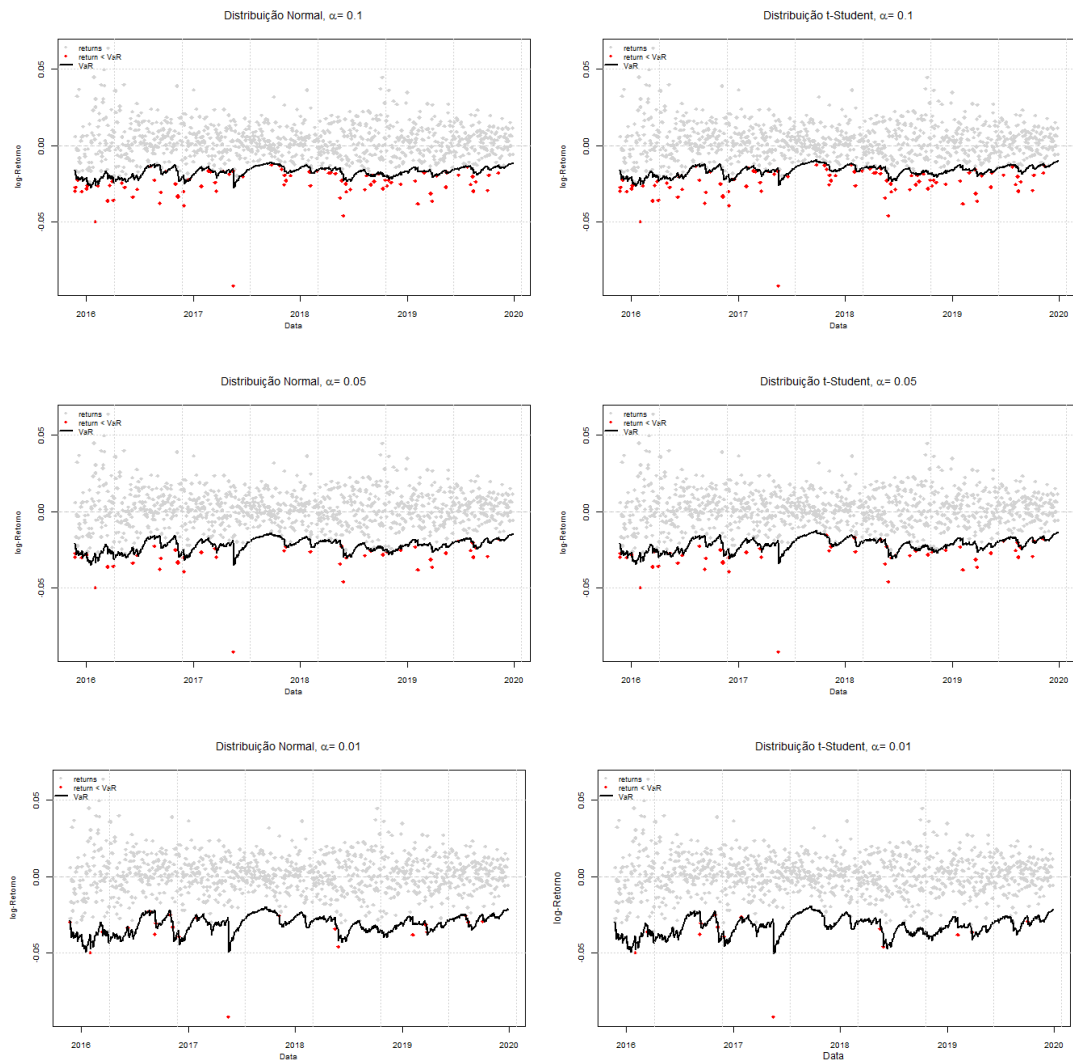
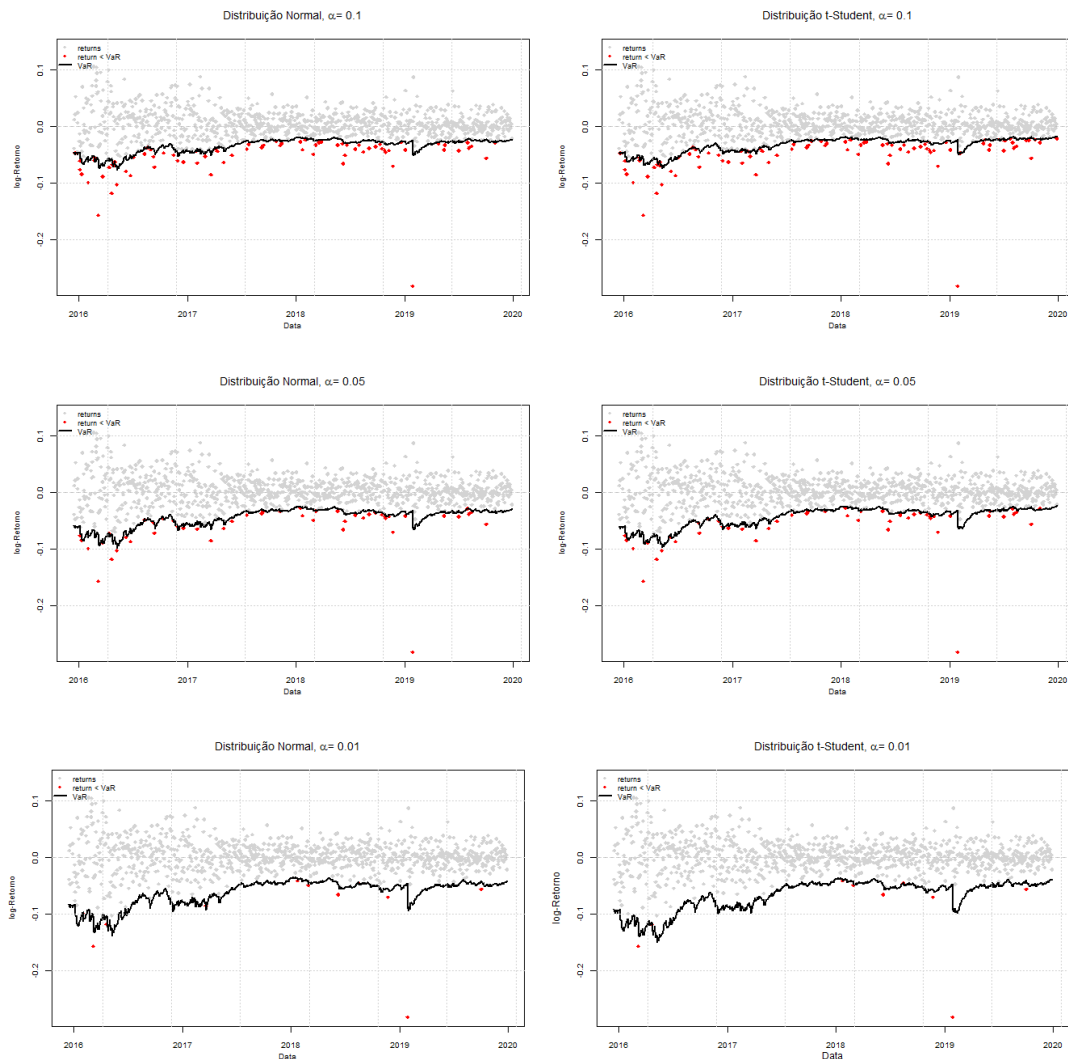


Figura 3.32: log-Retorno e Violações do VaR - Vale



A Figura 3.31 representa os lucros e perdas diárias do Ibovespa contra o VaR (negativo) e a Figura 3.32 representa os lucros e perdas diárias da Vale contra o VaR (negativo). Os gráficos não apenas resumem o desempenho dos ativos e a evolução de seus riscos de mercado como fornecem uma análise gráfica simples do desempenho da medida de valor em risco.

Com uma métrica do VaR de 90%, 95% e 99% em um dia, espera-se que as perdas excedam o VaR em aproximadamente 10%, 5% e 1% do tempo, respectivamente.

Portanto, tomando o grupo de séries com o mesmo tamanho amostral e avaliando a cobertura do VaR, que é definida como a frequência das excedências do VaR, quando o real retorno é menor do que o VaR previsto. Para um modelo de previsão ser considerado válido, a cobertura VaR tem que ficar em torno da probabilidade de perda α . Abaixo segue Tabela 13 com os valores de cobertura para as duas distribuições e para os α 's.

Tabela 13: Coberturas do modelo GARCH(1,1)

α	Ibovespa		Vale	
	Normal	SST	Normal	SST
10%	0,087	0,094	0,079	0,092
5%	0,048	0,053	0,05	0,05
1%	0,021	0,014	0,011	0,009

De acordo com os valores de cobertura apresentados na Tabela 13, tanto para os dados Ibovespa quanto para os dados da Vale, a distribuição SST apresentou melhores resultados. Para os dados da Vale, o nível de significância $\alpha = 0,05$ foi o único em que o número de violações observadas foi igual ao nível de significância do VaR. Os outros resultados ficaram bem próximo do nível de significância do valor de risco para os dois banco de dados.

Realizando o teste de cobertura, considerando as seguintes hipóteses:

$$\begin{cases} H_0 : q = q^* \\ H_1 : q \neq q^* \end{cases}$$

onde q é o quantil do VaR e q^* é a frequência de violações. A hipótese nula testa se o modelo foi bem ajustado. A Tabela 14 a seguir contém os resultados.

Tabela 14: Teste de cobertura do modelo GARCH(1,1)

α	Ibovespa		Vale	
	Normal	SST	Normal	SST
10%	Não Rejeita H_0	Não rejeita H_0	Rejeita H_0	Não Rejeita H_0
5%	Não rejeita H_0	Não rejeita H_0	Não Rejeita H_0	Não Rejeita H_0
1%	Rejeita H_0	Não rejeita H_0	Não Rejeita H_0	Não Rejeita H_0

A distribuição SST possui os valores de cobertura mais próximos do que a distribuição Normal. Para os dados do Ibovespa não rejeita H_0 para nenhum valor de α , para os dados da Vale rejeita apenas ao nível de significância $\alpha = 0,1$.

4 Considerações Finais

Uma das principais dúvidas de quem utiliza o VaR como ferramenta de suporte ao gerenciamento de risco de carteiras de investimentos, é a confiabilidade. Ainda que existam outras abordagens para o cálculo do VaR, a utilização dos modelos conhecidos como condicionalmente heterocedásticos consiste em uma das estratégias mais utilizadas, em especial os modelos derivados da família GARCH, proposto por Bollerslev (1986) e depois expandido para o EGARCH em Nelson (1991).

O principal objetivo que motivou esse estudo foi avaliar o nível acurácia do VaR quando calculado pelos modelos GARCH e EGARCH. Para tanto foi escolhido um período de aproximadamente 11 anos de observação de duas séries temporais financeiras, o índice Ibovespa e a série formada pelos fechamentos dos preços da ação VALE3 da empresa Vale. Uma vez que o VaR está relacionado ao quantil da distribuição de probabilidade condicional dos retornos, este estudo também contemplou duas alternativas, no caso a distribuição de probabilidade Normal, a qual é tradicionalmente utilizada, e a distribuição SST, a qual possibilita um comportamento leptocúrtico e assimétrico. Totalizando assim, quatro abordagens distintas e duas séries temporais para avaliação.

O estudo foi realizado através de um processo conhecido como janela deslizante, no qual, os modelos são sempre ajustados em períodos anteriores aos períodos que são utilizados para avaliar a precisão do VaR. Os resultados mostraram que as quatro abordagens tiveram resultados satisfatórios em ambas as séries, no entanto a utilização de uma distribuição de probabilidade leptocúrtica e assimétrica melhora os resultados, uma vez que as abordagens com distribuição SST foram melhores que as abordagens com distribuição Normal, tendo níveis de cobertura fidedignos aos níveis teóricos. Os resultados também revelaram uma ligeira vantagem do modelo EGARCH.

Vale destacar que as abordagens tradicionais pressupõem que o modelo ajustado é verdadeiro, de modo que a variabilidade dos estimadores dos parâmetros não é levada em consideração no cálculo do VaR, o que teoricamente poderia gerar previsões mais afuniladas que a ideal. Estudos futuros podem focar no emprego de abordagens por inferência Bayesiana ou por Bootstrap para obter métodos mais precisos, ainda que com consideravelmente maior custo computacional.

5 Apêndice

5.1 Teste de Box-Pierce-Ljung

Box e Pierce em 1970 sugeriram um teste para as autocorrelações dos resíduos estimados, que apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se os valores são muito altos. Uma modificação deste teste foi proposto por Ljung e Box em 1978, que é baseado em uma série de ordens de defasagem para determinar se a relevância ou aleatoriedade existe entre a sequência da população. A hipótese nula do teste de Ljung-Box é que os dados originais são independentes, ou seja, o coeficiente de autocorrelação geral é zero, e as autocorrelações observadas são causadas pelos erros de amostragem aleatória, enquanto a hipótese alternativa é que os dados não são distribuídos de forma independente, ou seja, existência de correlação serial.

Se o modelo for apropriado, a estatística

$$Q = n(n+2) \sum_{t=1}^K \frac{\hat{r}_t^2}{n-t}$$

terá distribuição qui-quadrado χ_r^2 com $r = K - p - q$ graus de liberdade. A hipótese de ruído branco para os resíduos é rejeitada para valores grandes de Q .

5.2 Critério de Akaike - AIC

Esse critério foi proposto por Akaike (1974) que utilizou a informação de Kullback-Laiber (K-L) para definir o critério AIC_{p+1} . A informação de Kullback-Leiber é uma medida da distância entre o modelo verdadeiro e um modelo candidato. Akaike desenvolveu uma estimativa da medida de K-L, baseada no logaritmo da função de verossimilhança no ponto de máximo, acrescida de uma penalidade associada ao número de parâmetros. A função de penalidade tem a finalidade de corrigir um viés proveniente da comparação de modelos com diferentes números de parâmetros. A medida denominada informação de Akaike é definida por:

$$AIC = -2 \log [\mathcal{L}(\hat{\theta})] + 2m,$$

em que m é o número de parâmetros estimado no modelo. Para dados normalmente distribuídos e utilizando as estimativas de máxima verossimilhança dos coeficientes, temos que

$$AIC = n \log(\hat{\sigma}_\varepsilon^2) + 2m,$$

onde $\hat{\sigma}_\varepsilon^2 = \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t^2$ e $\hat{\varepsilon}_t$ é o resíduo do modelo.

O modelo escolhido deve apresentar **menor** valor de *AIC* dentre todos os modelos considerados para determinado problema. Burnham e Anderson (2002) recomendam o uso do *AIC* apenas quando $\frac{n}{m} \geq 40$.

5.3 Critérios de Informação Bayesiano - BIC

O critério de informação bayesiano (*BIC*), proposto por Schwarz (1978), é definido por:

$$BIC = -2 \log [\mathcal{L}(\hat{\theta})] + 2m + m \log(n),$$

onde n é o tamanho da amostra. O modelo escolhido deve apresentar **menor** valor de *BIC* dentre todos os modelos considerados para determinado problema. O critério *BIC* penaliza mais modelos com maior número de parâmetros do que o critério *AIC*. Tendendo dessa forma, a selecionar modelos com um número menor de parâmetros.

Referências

- Mandelbrot, Benoit (1963). “New methods in statistical economics”. Em: *Journal of political economy* 71(5), pp. 421–440.
- Engle, Robert F (1982). “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation”. Em: *Econometrica: Journal of the Econometric Society*, pp. 987–1007.
- Bollerslev, Tim (1986). “Generalized autoregressive conditional heteroskedasticity”. Em: *Journal of econometrics* 31(3), pp. 307–327.
- Nelson, Daniel B (1991). “Conditional heteroskedasticity in asset returns: A new approach”. Em: *Econometrica: Journal of the Econometric Society*, pp. 347–370.
- Jorion, Philippe (1997). *Value at risk: the new benchmark for controlling market risk*. Irwin Professional Pub.
- Levine, David M et al. (1999). *Statistics for managers using Microsoft Excel*. Vol. 660. Prentice Hall Upper Saddle River, NJ.
- Caouette, John B, Edward I Altman e Paul Narayanan (2000). *Gestão do risco de crédito: o próximo grande desafio financeiro*. Qualitymark Editora Ltda.
- Holton, Glyn (jan. de 2003). “Value-at-Risk: Theory and Practice”. Em:
- Jorion, Philippe (2003). *Value at risk: a nova fonte de referência para a gestão do risco financeiro*. Bolsa de Mercadorias & Futuros.
- Tsay, Ruey S (2005). *Analysis of financial time series*. Vol. 543. John wiley & sons.
- Dowd, Kevin e David Blake (2006). “After VaR: the theory, estimation, and insurance applications of quantile-based risk measures”. Em: *Journal of Risk and Insurance* 73(2), pp. 193–229.
- Morettin, Pedro A e Clélia Tolo (2006). “Análise de séries temporais”. Em: *Análise de séries temporais*.
- Ishizawa, Danilo Kenji (2008). “Modelos de volatilidade estatística”. Em:
- Morettin, Pedro Alberto (2008). “Econometria financeira: um curso em séries temporais financeiras”. Em:
- Assaf, Ata (2015). “Value-at-Risk analysis in the MENA equity markets: Fat tails and conditional asymmetries in return distributions”. Em: *Journal of Multinational Financial Management* 29, pp. 30–45.

- Morettin, Pedro Alberto e WILTON OLIVEIRA BUSSAB (2017). *Estatística básica*. Editora Saraiva.
- Shumway, Robert H e David S Stoffer (2017). *Time series analysis and its applications: with R examples*. Springer.
- Franco, Christian e Jean-Michel Zakoian (2019). *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons.