

FACULDADE DE CIÊNCIAS DA
UNIVERSIDADE DE LISBOA
Departamento de Matemática

INSTITUTO SUPERIOR DE CIÊNCIAS DO
TRABALHO E DA EMPRESA
Departamento de Finanças



**Ciências
ULisboa**

ISCTE  **IUL**
Instituto Universitário de Lisboa
Lisbon University Institute

AVALIAÇÃO EMPÍRICA DO RISCO DE MERCADO: EXPECTED SHORTFALL vs. VALUE-AT-RISK

Ana Reis Gameiro

Mestrado em Matemática Financeira

Dissertação orientada por:
Professora Doutora Diana Aldea Mendes

AGRADECIMENTOS

Em primeiro lugar, agradeço de forma particular à Prof.^a Dr.^a Diana Mendes, por toda a dedicação, disponibilidade e orientação que teve durante todo este trabalho.

Agradeço aos meus pais por me proporcionarem esta oportunidade que, sem dúvida, sem eles não teria sido possível. Agradeço também à minha irmã por toda a ajuda e paciência que teve durante toda esta fase, e ao Nélio pela persistência e companheirismo.

Por fim, agradeço aos meus amigos mais próximos e colegas de mestrado que me deram todo o apoio e força para prosseguir neste projeto.

RESUMO

O risco financeiro tem influenciado cada vez mais as decisões tomadas pelas instituições financeiras. Como tal, o preço de um ativo influencia os mercados financeiros, o que atinge diretamente as ações das instituições financeiras, uma vez que estas correm um risco, dado o investimento efetuado previamente. Neste seguimento, as instituições financeiras têm implementado algumas estratégias de prevenção e de gestão do risco, tais como o *Value-at-Risk* (VaR) e o *Expected Shortfall* (ES).

O *Value-at-Risk* é um método de risco bastante utilizado pelas instituições financeiras. Contudo, a desvantagem deste método, ao não ser uma medida de risco coerente, propiciou a procura de um método alternativo, o *Expected Shortfall*, por parte das instituições financeiras. Embora, o *Expected Shortfall* tenha surgido para colmatar as lacunas do *Value-at-Risk*, este método também apresenta as suas desvantagens, pois é considerado uma medida não eliciável, segundo Roccioletti (2016) e Osmundsen (2016). Embora não seja relevante para o *backtesting* de acordo com Acerbi e Szekely (2014). Atualmente, ambos os métodos são utilizados para a avaliação do risco de mercado.

Nesta dissertação apresenta-se uma análise de como são aplicadas, em séries financeiras, as várias formas de implementação do *Value-at-Risk* e do *Expected Shortfall* associado ao principal índice bolsista de Portugal, o PSI20.

Recorreu-se a modelos paramétricos e não-paramétricos: o modelo com distribuição Normal e o modelo com distribuição *T-student*, e o modelo *Kernel*, respetivamente. E, foi, também, utilizado o modelo *GARCH* com distribuição Normal.

Por fim, foram aplicadas e analisadas metodologias de *backtesting*: o teste de Kupiec (1995) e o teste de Christoffersen (1998), para a avaliação dos diferentes modelos de previsão do *Value-at-Risk* e, ainda, os testes Z_1 e Z_2 desenvolvidos por Acerbi e Szekely (2014) para a previsão do *Expected Shortfall*. Para a série financeira em estudo conclui-se que o modelo paramétrico com distribuição Normal determina o melhor desempenho na previsão do risco, relativamente ao método *Expected Shortfall*.

PALAVRAS CHAVE:

Expected Shortfall, *Value-at-Risk*, Simulação Histórica, *GARCH*, *T-student*,

ABSTRACT

The financial risk has increasingly affect the decisions taken by financial institutions. As such, the price of an asset affects the financial markets, which directly affects the actions of financial institutions, since these are at risk, given an investment made previously. Following this, financial institutions have implemented some risk prevention and risk management strategies, such as Value-at-Risk (VaR) and Expected Shortfall (ES).

Value-at-Risk is a risk method widely used by financial institutions. However, the disadvantage of this method, not being a coherent measure, led to search for an alternative method, the Expected Shortfall, by financial institutions. Although, Expected Shortfall has emerged to rectify the gaps of Value-at-Risk, this method also has its drawbacks, is it considered a non-elicitable measure, according to Roccioletti (2016) and Osmundsen (2016). Although it is not relevant for backtesting according to Acerbi e Szekely (2014). Currently, both methods are used for market risk evaluation.

This dissertation present an analysis of how several forms of implementation of Value-at-Risk and Expected Shortfall associated with the main stock index of Portugal, the PSI20, are applied to financial series.

We used parametric and non-parametric models: the model with Normal distribution and the model with T-student distribution, and the Kernel model, respectively. And , the GARCH model with Normal distribution was also used.

Finally, we applied and analyzed backtesting methodologies: the Kupiec (1995) test and the Christoffersen (1998) test, for the evaluation of different Value-at-Risk forecasting models, as well as the tests Z_1 e Z_2 developed by Acerbi e Szekely (2014) for the Expected Shortfall forecasting. For the financial serie under study we conclude that the parametric model with Normal distribution determines the best performance in the risk forecast, in relation to the Expected Shorfall method.

KEYWORDS:

Expected Shortfall, Value-at-Risk, Historical Simulation, GARCH, T-student

ÍNDICE

INTRODUÇÃO	1
1 RISCO FINANCEIRO: ENQUADRAMENTO DO TEMA	3
1.1 PSI20	3
1.2 REVISÃO BIBLIOGRÁFICA	5
2 MODELOS ECONOMETRÍCOS E METODOLOGIAS DE ESTIMAÇÃO DO VAR E ES 9	
2.1 SIMULAÇÃO HISTÓRICA	9
2.2 VALUE-AT-RISK	10
2.2.1 CÁLCULO DO VAR	11
2.3 EXPECTED SHORTFALL	13
2.3.1 CÁLCULO DO ES	14
2.3.2 COMPARAÇÃO: VAR VS ES	15
2.4 MODELOS DE SÉRIES TEMPORAIS	17
2.4.1 SÉRIE TEMPORAL	17
2.4.2 ESTACIONARIDADE E TESTE DA RAIZ UNITÁRIA	18
2.4.3 ARMA	19
2.4.4 GARCH	19
2.5 BACKTESTING	20
2.5.1 BACKTESTING VAR	21
2.5.1.1 TESTES DE COBERTURA INCONDICIONAL	22
2.5.1.2 TESTES DE COBERTURA CONDICIONAL	23
2.5.2 BACKTESTING ES	25
2.5.2.1 TESTE 1	26
2.5.2.2 TESTE 2	27
2.5.2.3 TESTE 3	27
3 ESTIMAÇÃO E BACKTESTING DO VAR E ES	31
3.1 ANÁLISE DA SÉRIE E DOS SEUS RETORNOS	31
3.2 ESTIMAÇÃO DO VAR E DO ES	35
3.2.1 RESULTADOS DO BACKTESTING AO VAR	39
3.2.2 RESULTADOS DO BACKTESTING AO ES	41
CONCLUSÃO	43
REFERÊNCIAS BIBLIOGRÁFICAS	44

ÍNDICE DE FIGURAS

Figura 1.1: PSI20	5
Figura 2.1: Distribuições: Normal e Leptocúrtica	17
Figura 3.1: Evolução da cotação do PSI20 entre 02/01/2006 e 31/01/2017	31
Figura 3.2: Estatística descritiva da série PSI20.....	32
Figura 3.3: Série logaritmizada dos retornos do índice PSI20.....	33
Figura 3.4: Estatística descritiva da série logaritmizada dos retornos do índice PSI20.....	34
Figura 3.5: PSI20 vs Distribuição Normal	35
Figura 3.6: PSI20 vs Distribuição T-student	36
Figura 3.7: PSI20 vs Distribuição Kernel.....	36
Figura 3.8: Observação das caudas	37
Figura 3.9: Estimação VaR e ES com a distribuição Normal.....	37
Figura 3.10: Estimação VaR e ES com a distribuição <i>T-student</i>	38
Figura 3.11: Estimação VaR e ES com a distribuição Kernel.	38
Figura 3.12: Estimação VaR e ES com distribuição GARCH Normal.....	39

ÍNDICE DE TABELAS

Tabela 3.1: Teste ADF aplicado ao índice PSI20.....	32
Tabela 3.2: Teste PP aplicado ao índice PSI20	33
Tabela 3.3: Teste ADF aplicado à série logaritmizada dos retornos do índice PSI20	34
Tabela 3.4: Teste PP aplicado à série logaritmizada dos retornos do índice PSI20.....	35
Tabela 3.5: Número de exceções esperadas e observadas	39
Tabela 3.6: Resultados do Teste de <i>Kupiec</i>	40
Tabela 3.7: Resultados do Teste de <i>Christoffersen</i>	40
Tabela 3.8: Resultados do teste <i>Z1</i>	41
Tabela 3.9: Resultados do teste <i>Z2</i>	41

INTRODUÇÃO

A necessidade das instituições financeiras conseguirem prever a dimensão do seu risco tem aumentado, o que leva a evitar elevados prejuízos para estas. Portanto, a perceção prévia deste valor é cada vez mais precisa para que se tome decisões mais conformes.

O risco financeiro está associado à incerteza do retorno de um investimento perante um evento futuro. O comportamento dos ativos financeiros é bastante variável ao longo do tempo, pode existir uma enorme variação como a variação pode ser praticamente nula. Logo, de forma a minimizar perdas inesperadas nas instituições, a compreensão e previsão do risco presente nas séries financeiras é de grande interesse por parte das instituições financeiras.

Com isto, surgiu a necessidade de existirem metodologias que avaliassem o risco associado às séries financeiras tal como o *Value-at-Risk* (VaR), sendo definido, atualmente pelas instituições financeiras, como um dos melhores métodos para quantificar o risco. No entanto, este método apresenta algumas desvantagens, Osmundsen (2016) refere a falta de coerência e o facto deste método ignorar, na sua totalidade, a amplitude das perdas presentes nas caudas da distribuição. Devido a estas desvantagens presentes no modelo VaR, criou-se a necessidade de ter outro método que pudesse colmatar as suas desvantagens. Neste contexto, surge o método *Expected Shortfall* (ES). O que não implica que este método não falhe noutros aspetos, tanto Roccioletti (2016) como Osmundsen (2016) afirmam que o ES não é uma medida eliciável.

Desta forma, este trabalho tem como principal objetivo avaliar o risco utilizando em ambos os métodos apresentados agregado ao principal índice bolsista português: o PSI20. A base de dados, recolhida no DataStream, é constituída por 2892 observações, indicadas entre o período de 02 de janeiro de 2006 e 31 de janeiro de 2017.

Existem várias formas de poder medir o risco, contudo, neste trabalho irá ser avaliado através de distribuições de probabilidade. Para a realização do *backtesting* são utilizados alguns testes para cada um dos métodos. Para o VaR recorreremos ao teste de Kupiec (1995) e ao teste de Christoffersen (1998) de forma a concluir sobre a sua *performance*. Para o ES recorreremos aos testes Z_1 e Z_2 desenvolvidos por Acerbi e Szekely (2014). Os modelos utilizados nesta análise são distinguidos entre os modelos paramétricos e não-paramétricos. Os modelos com distribuição Normal e *T-student* representam os modelos paramétricos e o modelo *Kernel* representa os modelos não paramétricos. É também utilizado o modelo *GARCH* com distribuição Normal. Todos estes modelos irão ser apresentados com níveis de significâncias de 97.5% e de 99%.

No primeiro capítulo é apresentado um breve enquadramento sobre o PSI20: conceito, a sua fórmula de cálculo e alguns marcos históricos. É apresentada também uma revisão bibliográfica onde é possível observar alguns estudos similares ao que irá ser realizado.

No segundo capítulo são abordados conceitos base sobre as séries temporais, assim como os modelos econométricos para a estimação do VaR e do ES.

No terceiro capítulo é desenvolvido a análise empírica do risco para a série em questão: o PSI20. Recorreremos ao software *EViews* para análise inicial da série PSI20 e o software *MATLAB* para realizar a previsão do risco e o *backtesting* aplicando os testes de Kupiec e Christoffersen e os testes Z_1 e Z_2 para o VaR e ES, respetivamente.

1 RISCO FINANCEIRO: ENQUADRAMENTO DO TEMA

1.1 PSI20

As ações são prováveis de gerar ganhos superiores a qualquer outro tipo de investimentos. As ações são caracterizadas como vulneráveis a oscilações na volatilidade, podendo os preços subir ou descer rapidamente. Em bolsa, as ações negociadas não têm um valor fixo, variando as suas cotações, a longo prazo, em função de vários fatores, tais como os resultados alcançados pela empresa, o seu desempenho, perspectivas futuras, situação económica do momento e a evolução dos mercados financeiros.

Os índices são apresentados na forma de número ou percentagem que ilustram a evolução de um determinado mercado, acabando por nos fornecer informação relativamente à tendência dos mercados, considerando, que os mercados são importantes para avaliar o desempenho de portfólios.

O valor do índice é descrito em “pontos” e a forma de interpretar o valor é em comparação com o valor anterior. O seu valor numérico não tem grande relevância, apenas a sua variação, que indica o sentido da evolução do mercado, sendo que o preço de cada ação varia consoante as alterações do preço e o número de ações em circulação no mercado (Francisco, 2015).

A Comissão do Mercado de Valores Mobiliários (CMVM), em Portugal, é a entidade supervisora e reguladora dos mercados de valores mobiliários, dos instrumentos financeiros e da atividade em que atuam, sendo um organismo público independente, com autonomia administrativa e financeira.

O índice PSI20, *Portuguese Stock Index*, é o principal índice de referência do mercado de capitais português e o principal índice da *Euronext* Lisboa. Este, é composto pelas ações das 20 maiores empresas cotadas na bolsa de valores de Lisboa, refletindo a evolução do preço das ações, que são as de maior liquidez entre as negociadas no mercado português.

O PSI20 foi criado a 31 de dezembro de 1992, com valor base de 3000 pontos.

Segundo a Euronext (2003), o PSI20 foi lançado com uma dupla finalidade, em primeiro serve de indicador da evolução do mercado acionista português, em segundo serve de suporte à negociação de contratos de futuros e opções.

Devido às suas características, o índice PSI20 tem vindo a ser escolhido pelo mercado de forma a servir de subjacente a produtos estruturados, cuja rentabilidade depende do comportamento do mercado bolsista português.

De acordo com a CMVM (2017) as empresas constituintes atualmente são: Altri, BCP, Corticeira Amorim, CTT, EDP Renováveis, GALP, Energias de Portugal, Ibersol, Jerónimo Martins, Montepio, Mota-Engil, NOS, Novabase, Pharol, REN, Semapa, Sonae, Sonae Capital e The Navigator. A seleção das empresas que compõem o PSI20 baseia-se nos seguintes critérios:

- i. dimensão (quantidade de ações negociadas);
- ii. dispersão acionista (*free float* ou quantidade de ações dispersas por pequenos acionistas);
- iii. liquidez previsível (quantidade de negócios realizados em bolsa);
- iv. profundidade;
- v. regularidade e continuidade de formação de preços dessas emissões.

Segundo a *New York Stock Exchange* (NYSE Euronext, 2015, citado em Francisco, 2015), consoante os critérios acima mencionados, nem todas as empresas têm o mesmo peso no cálculo do Índice PSI20. O índice é calculado e divulgado pela *Euronext Lisbon*, a cada 15 segundos, durante a sessão de bolsa, isto é, o período de funcionamento estabelecido pela *Euronext Lisbon* (das 8h às 16h35, de Segunda a Sexta-feira).

Contudo, até ao final de 2014, as empresas presentes no índice eram escolhidas com base no valor negociado em bolsa, no que diz respeito ao já referido *free float market capitalization*. Desde 2013, caso uma empresa tencione pertencer no PSI20 é necessário que esta cumpra os seguintes requisitos:

- a) Valor da sua capitalização bolsista efetivamente dispersa terá de ser superior a 100 milhões de euros;
- b) A dispersão do seu capital não poderá ser inferior a 15%.

Foram, ainda, propostas alterações no que diz respeito à liquidez das empresas que entram para o PSI20 e ao cálculo do respetivo *free float velocity*, ou seja, o cálculo da razão entre o número de ações negociadas e o número de ações dispersas. Esta razão passou de um limite mínimo de 10% para o limite de 25%. Por outro lado, o peso máximo de cada empresa do PSI20 passou de 15% para 12%.

De acordo com o que é apresentado pela *Euronext Lisbon* (2014), a fórmula geral de cálculo para o índice de preços é dada por:

$$I_t = \frac{\sum_{i=1}^N Q_{i,t} F_{i,t} f_{i,t} C_{i,t} X_{i,t}}{d_t} \quad (1.1)$$

onde t representa o momento de tempo do cálculo, N o número de empresas constituintes; $Q_{i,t}$ o número de ações da constituinte i utilizadas no cálculo do índice no dia t ; $F_{i,t}$ o fator de *free float* da constituinte i ; $f_{i,t}$ o fator limite de peso da constituinte i ; $C_{i,t}$ o preço da constituinte i no momento t ; $X_{i,t}$ a taxa de câmbio atual no momento t e, por fim, d_t representa o divisor do índice no dia t .

No entanto existe alguma preocupação relativamente ao risco de ações.

Segundo Associação BM&FBOVESPA (2012), o risco é definido como a possibilidade de que algum acontecimento desfavorável ocorra, ou seja, é a possibilidade de perda financeira. Sendo que o risco é uma das principais variáveis que afeta os resultados dos investimentos.

Bacic, Silveira e Sousa (2013, citado em Augusto, Santos & Vieira, 2016) designam que “o risco financeiro está associado à possibilidade de perdas no mercado financeiro conduzindo à não ocorrência ou dispersão dos resultados esperados” (p.3).

Dentro do risco financeiro, existe o risco de mercado que está relacionado com a variação dos preços. Os principais elementos deste risco relacionado ao mercado financeiro são: taxas de juro, taxas de câmbio, preço das ações e preço das *commodities*.

Com isto, no gráfico seguinte (Figura 1.1) podemos observar a evolução do PSI20 ao longo dos anos:



Figura 1.1: PSI20

Fonte: <https://tradingeconomics.com/portugal/stock-market>

Historicamente, o PSI20 atingiu o seu valor máximo de 14822.59 em março de 2000 e o seu valor mínimo de 2917.56 em janeiro de 1993.

Podemos observar alguns decréscimos relativamente ao índice em análise. A ANACOM (2008) afirma que entre o início do 2º semestre de 2007 e o final do 1º semestre de 2008, o índice PSI20 registou um decréscimo de 33.6%, acompanhando uma tendência de diminuição geral dos mercados internacionais. O que se deveu à crise financeira internacional que se iniciou em 2008 nos Estados Unidos da América e se alastrou aos seus principais parceiros comerciais, nomeadamente os países europeus.

A crise financeira referida teve efeitos devastadores sobre a economia mundial e conduziu a restrições severas no mercado. Assistiu-se a uma descida dos níveis de preços nos mercados, ao encerramento de muitas empresas que, por sua vez, levou ao aumento do desemprego e a uma queda de impostos (Costa, 2014).

Em Portugal teve um impacto bastante significativo para as empresas portuguesas, nomeadamente a nível das suas condições de financiamento, tendo como principal problema o crescimento da dívida (Carvalho, 2013).

Bernanke (2008, citado em Matias, 2012) refere que “os fatores importantes que contribuíram para a crise financeira foram: a inadequada gestão de risco, as fraquezas em geral das instituições financeiras mundiais que criaram e mantiveram produtos de crédito complexos, que resultaram na má gestão do risco de algumas instituições financeiras e na ausência de divulgação do risco” (p.13).

1.2 REVISÃO BIBLIOGRÁFICA

Existem diversas obras (livros, teses, artigos) a nível internacional e nacional que abordam a avaliação do risco de mercado, utilizando vários métodos e técnicas. Vejamos uma breve revisão de alguns destes trabalhos.

Harmantzis, Miao e Chien (2005) realizaram uma comparação do desempenho dos diferentes modelos na estimativa do VaR e do ES utilizando dados históricos onde são utilizados os índices S&P500, DAX, CAC, Nikkei, TSE e FTSE, por dez anos com três tamanhos de janela distintos. São utilizadas as distribuições Gaussiana e Pareto, entre outras. É dada alguma ênfase ao nível de confiança

e ao tamanho da janela. Os autores afirmam que ao nível de significância de 95%, o modelo Gaussiano possui estimativas do VaR mais precisas, sendo que este modelo aparenta ser adequado para calcular o VaR com um baixo nível de confiança. A um nível de confiança superior (99%) os modelos de cauda pesada produzem estimativas melhores do que os que tem a cauda menos pesada, sendo que a distribuição Gaussiana e o modelo histórico apresentam um desempenho mais fraco. A este nível quanto maior o tamanho da janela maior será a precisão dos resultados. No entanto, para o ES o método histórico é o mais preciso para ambos os níveis de confiança. Por outro lado, para o método Gaussiano, as discrepâncias são sempre positivas, ou seja, este modelo subestima o ES. Concluem, que na estimativa do VaR, os modelos de caudas pesadas podem prever o risco de forma mais precisa enquanto na estimativa do ES o método histórico é, de facto, o que fornece estimativas mais corretas.

Kouroma *et al.* (2010) utilizam a distribuição Pareto na teoria de Valor Extremo para medir o VaR e o ES para os índices CAC40 e S&P500 durante a crise financeira de 2008. Consideram os horizontes temporais de 1, 5 e 10 dias. Realizam uma comparação a estes resultados com os obtidos pela Simulação Histórica através de um período de 250 dias, apresentando uma subestimação do risco de perda nos modelos VaR, sendo que o ES indica uma certa confiabilidade. Concluem que o risco de perda de um índice de ações é superior à ocorrência de ganhos esperados. E os resultados apresentam que para todos os níveis de confiança e horizontes diários, a estimativa do VaR baseado na abordagem da Simulação Histórica subestima o risco da perda. Sendo que a estimativa diminui quando o nível de risco diminui ou quando o horizonte temporal aumenta. O estudo evidência também que em períodos de crise financeira o ES, baseado na teoria do Valor Extremo, fornece uma medida confiável para conter ou reduzir perdas excessivas, mostrando uma pequena diferença entre as perdas estimadas e as perdas realizadas. Por fim, concluem que o ES garante a sua superioridade, em relação ao VaR, através da cobertura das perdas reais.

Brandolini e Colucci (2011) efetuam uma comparação entre a Simulação Histórica e as Simulações de Monte Carlo de modo a estimar o VaR, realizando três testes: cobertura incondicional, cobertura condicional e independência recorrendo a Christoffersen. São apresentados resultados ao nível de 1% e 5% com um dia de horizonte temporal para alguns índices tais como S&P500, DAX, MSCI entre outros. Concluem, que a Simulação Histórica é bastante fraca relativamente ao teste de independência pois este demora a ajustar-se à estimativa do VaR e têm bastantes casos de rejeição quanto ao teste de cobertura condicional.

Amaral (2015) faz uma análise de modelos de previsão do VaR aplicado ao índice PSI20 implementando os modelos *Riskmetrics*, *GARCH*, *EGARCH*, *ARCH* segundo as distribuições Normal e *T-student*. Foram, ainda, aplicadas técnicas de *backtesting* como os testes de Kupiec (1995) e Christoffersen (1998). O autor conclui que o modelo que apresentou um melhor desempenho, no geral, foi o modelo *Riskmetrics*. Contudo, para algumas das estatísticas utilizadas o melhor modelo é o *GARCH* (1,1) com distribuição *T-student* e para outras é preferível o modelo *GARCH* (1,1) com distribuição Normal.

Martins (2015) relata algumas abordagens a séries financeiras utilizando o ES, para prever os retornos diários de EUR/CHF e EUR/GBP e outra para prever diretamente as taxas de câmbio. Os métodos usados são a Simulação Histórica, Simulações de Monte Carlo, *ARCH/GARCH*, *EWMA* e *GeneticPred*, usando níveis de confiança 1%, 2.5% e 5%. Provou-se que o valor do ES é superior ao VaR e que cresce com o aumento do horizonte de estimação. A simulação via *RiskMetrics* e a Simulação Histórica geram valores de ES bastante semelhantes podendo ser considerados como os mais baixos.

Karlsson e Zakrisson (2016) analisam o método da estimação paramétrica para a estimação do VaR e ES junto com o modelo de Simulação Histórica para descobrir se a Simulação Histórica pode produzir estimativas precisas em períodos com maior ou menor agitação. Utilizam seis versões diferentes de *GARCH*, duas versões diferentes de *EWMA* e outras duas diferentes de Simulação Histórica. Considerando dois períodos distintos de estabilidade financeira, são utilizados dez índices diferentes na análise implementada. Concluem, que o método *EWMA* é o melhor para estimar ambos, o VaR e o ES. Embora, a Simulação Histórica produza resultados aceitáveis no período pós-crise, ou seja, para períodos de agitação a Simulação Histórica não será de todo uma boa opção.

2 MODELOS ECONOMÉTRICOS E METODOLOGIAS DE ESTIMAÇÃO DO VAR E ES

Existem duas principais abordagens para a previsão do risco financeiro: modelos não paramétricos e paramétricos.

Os modelos não paramétricos, geralmente, referem-se à Simulação Histórica, que usa a distribuição empírica de dados para calcular as previsões de risco. Não é assumido nenhum modelo estatístico e não é necessária nenhuma estimativa de parâmetros. Em contrapartida, os métodos paramétricos baseiam-se na estimativa da distribuição subjacente aos retornos (dados) e na obtenção de previsões de risco a partir da distribuição estimada (Danielsson, 2011).

Nos modelos não-paramétricos é necessário definir o tamanho da janela de observações. A vantagem deste método é que utiliza diretamente os dados observados, não estando sujeito a erros de estimação e consegue demonstrar diretamente a dependência não linear. Contudo tem a desvantagem de, no caso de existir algum período de agitação no mercado, após um longo período de volatilidade reduzida, o VaR calculado através da Simulação Histórica adapta-se lentamente às novas condições e tende a subestimar o seu valor.

Contrariamente, os modelos paramétricos baseiam-se na estimativa de uma distribuição dos dados o que cria, inevitavelmente, a problemática do erro de estimativa e do risco subjacente ao modelo.

2.1 SIMULAÇÃO HISTÓRICA

A Simulação Histórica (HS – *Historical Simulation*) é um método bastante simples para a previsão do risco. Baseado no pressuposto de que o histórico se repete, um dos retornos anteriormente observados é expectável que seja o retorno do período seguinte (Danielsson, 2011).

Este método é caracterizado como um método não-paramétrico, isto é, não tem de obedecer a qualquer regra, logo são livres de qualquer distribuição estatística.

De acordo com Gschöpf (2014) a estrutura i.i.d. (independente e identicamente distribuída) é assumida, apesar de ser uma técnica não paramétrica, o que fornece a vantagem de minimizar erros de falta de especificação do modelo.

Holá (2012) refere a mesma vantagem de ser um método não-paramétrico, pelo facto, de não ser requerido a suposição de uma distribuição de probabilidade. E, como desvantagem o autor afirma que o passado pode ter características de risco distintas do futuro. O método tem como princípio, a base das mudanças históricas nos fatores de mercado, para determinar os lucros e perdas futuras, ou seja, é necessário recolher os dados e identificar os fatores que influenciam toda a carteira.

Cada observação histórica acarreta o mesmo peso para este tipo de previsão. No entanto, pode ser uma desvantagem quando ocorre uma quebra estrutural na volatilidade. O efeito das correlações que é, inicialmente, considerado perde o seu valor. Por outro lado, quando esta quebra não ocorre, o método tende a apresentar melhores resultados do que outros métodos alternativos (Danielsson, 2011).

Uma das vantagens presente neste método é de quando se verifica um *outlier* atípico, ou seja, quaisquer observações com valores díspares não são significativas, não apresentam erros de estimativa como os métodos paramétricos. No que diz respeito aos portfólios, a vantagem presente é a de reter diretamente a dependência não linear, de uma forma que outros os métodos não conseguem (Danielsson, 2011).

De acordo com Gustafsson e Lundberg (2009), a desvantagem nesta abordagem é que são necessários bastantes dados para que exista uma melhoria em níveis de confiança elevados. E, assumindo que o passado é idêntico ao futuro, irá fazer com que os riscos passados sejam os mesmos que os riscos futuros.

Para que o método funcione é necessário que exista um histórico, caso contrário, este método não é de todo aplicável.

Para Karlsson e Zakrisson (2016), a maior vantagem deste método é que, para além da sua simplicidade, evidencia as características das caudas gordas dos retornos dos ativos. Embora percam com o facto de a distribuição ser assumida como igual ao longo do tempo. Os autores referem, ainda, como vantagem que as janelas de observação no caso de serem pequenas reagem rapidamente às mudanças da volatilidade, mas tem a desvantagem de ser menos representativas do verdadeiro retorno.

Segundo Holá (2012), o risco é medido com a variação dos preços.

Seja P_t o preço no momento t . Para o período de um dia são apresentadas as seguintes fórmulas:

- Mudança relativa no preço:

A taxa de retorno relativo (R_t) entre o momento t e $t - 1$ é dado pela fórmula:

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

- Mudança absoluta no preço:

A taxa de retorno absoluto (Ra_t) entre o momento t e $t - 1$ é dado pela fórmula:

$$Ra_t = P_t - P_{t-1} \quad (2.2)$$

- Mudança logarítmica no preço:

A taxa de retorno logarítmica (Rg_t) entre o momento t e $t - 1$ é dado pela fórmula:

$$Rg_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(1 + R_t) \quad (2.3)$$

Para o período de k dias, a taxa de retorno relativo (R_t) é dado pela fórmula:

$$R_t = \frac{P_t - P_{t-k}}{P_{t-k}} \quad (2.4)$$

O retorno relativo r_t é dado por:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (2.5)$$

onde P_t representa o valor da ação no momento t . Referindo que um dos principais problemas é o pressuposto de estacionaridade devido à mesma questão que Danielsson (2011) menciona relativamente à quebra estrutural.

2.2 VALUE-AT-RISK

Value at Risk (VaR) teve origem nos anos 90 quando os bancos mundiais e as instituições financeiras começaram a necessitar de medidas para medir o risco cambial (Holá 2012).

De momento o VaR é das medidas de risco mais utilizadas em Finanças. Têm se tornado uma medida clássica que os executivos financeiros utilizam para quantificar o risco de mercado. É uma extensão direta e inteligente da perda máxima: a ideia é simplesmente substituir a perda máxima pela perda máxima que não excede uma dada probabilidade elevada, o designado nível de confiança (Roccioletti, 2016).

Roccioletti (2016) define o *Value at Risk* com um dado nível de confiança $\alpha \in (0,1)$, o VaR do portfólio ao nível de confiança α é apresentado pelo menor valor l tal que a probabilidade da perda L exceder l não seja superior a $(1 - \alpha)$.

Apresentando:

$$VaR_{\alpha}(L) = \inf \{ l \in \mathbb{R} : P(L > l) \leq 1 - \alpha \} = \inf \{ l \in \mathbb{R} : F_L(l) > \alpha \} \quad (2.6)$$

onde $F_L(l)$ representa a função distribuição.

Já Danielsson (2011) define VaR como sendo a perda no portfólio transacionado, tal que existe uma probabilidade p de perdas igual ou superior ao VaR no dado período de transação e uma probabilidade $1 - p$ de perdas que são inferiores ao VaR.

Sendo Q uma variável aleatória que indica o lucro ou perda do investimento no portfólio. E caso o valor do portfólio seja ϑ então a lucro/perda do portfólio é dado por: $Q = \vartheta Y$, onde Y representa os retornos.

Então o VaR é dado por:

$$P[Q \leq -VaR(p)] = p \quad (2.7)$$

ou

$$p = \int_{-\infty}^{-VaR(p)} f_q(x) dx \quad (2.8)$$

onde $f_q(\cdot)$ representa a densidade de lucro/perdas. O operador aritmético usado, no cálculo, é o sinal menos pois o VaR é um número positivo, mas estamos a lidar com perdas.

Gustafsson e Lundberg (2009) afirmam que o VaR consiste em três parâmetros, nível de confiança, horizonte temporal e um valor. O nível de confiança determina com que certeza a medição é feita, níveis de confiança elevados implica maiores perdas. O horizonte temporal informa-nos sobre o intervalo de tempo em que se está a avaliar, assim, quanto maior for o horizonte temporal maior será a perda. Por fim, a componente de valor é o valor monetário que é arriscado para a perda durante o horizonte temporal e o nível de confiança propostos.

Roccioletti (2016) afirma que é uma medida bastante intuitiva e clara. Contudo, Osmundsen (2016) afirma que esta medida, não preenche algumas propriedades necessárias e é insensível à cauda, pois não refere o tamanho da perda em $(1 - \alpha) * 100\%$ dos casos.

Outra das grandes desvantagens deste método é que ignora na totalidade a gravidade dos prejuízos na cauda da distribuição das perdas (Emmer, Kratz & Tasche, 2015).

Segundo Danielsson (2011) existem três passos no cálculo do VaR. Inicialmente, é necessário que a probabilidade de perdas (p) que excedem o VaR seja especificada, com o respetivo nível. De seguida, refere que o período de detenção, isto é, o período durante o qual ocorrem as perdas, quanto mais longo for, mais realístico é o valor e maior será o VaR. Por fim, é identificada a distribuição de probabilidade do lucro e da perda do portfólio, sendo que a prática standard é estimar a distribuição usando observações passadas e um modelo estatístico.

Recorrendo a Basel II, Gschöpf (2014), para estimar o VaR os participantes do mercado devem usar $\alpha = 1\%$, um horizonte de previsão de 10 dias e com, pelo menos, um ano de dados históricos.

2.2.1 CÁLCULO DO VAR

Nos modelos não paramétricos é usado o método da Simulação Histórica.

Segundo Danielsson (2011), o VaR na probabilidade p é, simplesmente, o valor negativo $T * p$ do vetor de retornos ordenado multiplicado pelo valor monetário do portfólio.

Goorbergh e Vlaar (1999) afirmam que os dados históricos são recolhidos de um dado período de tempo e, de seguida, os preços históricos são considerados uma boa avaliação para os preços futuros. Apresentando a seguinte função (citado em Gustafsson e Lundberg, 2009):

$$\widehat{VaR}_{t+1|t} = -W_0 R_t^p \quad (2.9)$$

o $\widehat{VaR}_{t+1|t}$ é o valor do VaR no momento $t + 1$ onde W_0 é o valor inicial do ativo e R_t^p é o percentil de ordem p de cada observação.

No entanto Izmaylov (2014) apresenta outra fórmula de cálculo, indicando que o pressuposto é apenas que os retornos são i.i.d.. Os retornos são organizados por tamanho, dependendo de um nível de confiança α e de um número de observações n , a observação após os primeiros $\alpha\%$ de retornos é usada no cálculo:

$$VaR = -P_0(r^* - \mu) \quad (2.10)$$

onde P_0 é o valor inicial da posição, μ é a média esperada do retorno e r^* é o *cutoff return* para o menor $\alpha * n$ dos retornos.

Quanto à janela de observações, no caso de a janela ser grande, o VaR baseia-se em valores históricos bastante antigos que podem não representar a situação atual do mercado. Além disso, o ajuste a qualquer período de agitação é lento, como já foi referido. Por outro lado, quando a janela é pequena o cálculo pode não ser suficientemente consistente, especialmente para quantis com baixo nível de probabilidade.

Nos modelos paramétricos são geralmente utilizadas duas distribuições para o cálculo do VaR: a distribuição Normal ou a distribuição *T-student*.

Relativamente à distribuição Normal, Roccioletti (2016) afirma que a distribuição Normal oferece uma solução direta para a estimação da distribuição da perda de portfólios financeiros. Define que o VaR para a distribuição Normal é dada por:

$$VaR_\alpha = \mu + \sigma * \Phi^{-1}(\alpha) \quad (2.11)$$

onde Φ representa a função distribuição Normal standard.

Segundo Duda e Schmidt (2009), a distribuição Normal é uma abordagem que assume que os retornos são normalmente distribuídos e o VaR é calculado aplicando quantis da distribuição Normal standard.

Os autores supracitados apresentam a seguinte fórmula de cálculo do VaR sob a distribuição Normal:

$$VaR_\alpha = -\mu + \sigma * q_z(\alpha) \quad (2.12)$$

onde μ representa a média, σ é uma amostra do desvio padrão e $q_z(\alpha)$ é o valor da variável aleatória Normal.

Danielsson (2011) apresenta a seguinte fórmula do cálculo do VaR:

$$VaR(p) = -\sigma * \gamma(p) * P_{t-1} \quad (2.13)$$

onde $\gamma(p)$ representa a função inversa da distribuição e P_{t-1} representa o valor do portfólio no momento $t - 1$.

Contudo, existem algumas críticas a esta abordagem uma vez que não permite que ocorram grandes movimentos no mercado devido às caudas pesadas. E o uso desta distribuição neste caso leva a uma

subestimação do valor do VaR, de acordo com Duda e Schmidt (2009, citado por Glasserman, Heidelberger, & Shahabuddin, 2000)

Quanto à distribuição *T-student*, Roccioletti (2016) afirma que a distribuição é selecionada devido a ter média zero e a variância é determinada através do parâmetro dos graus de liberdade. Define que o VaR para a distribuição *T-student* é dada por:

$$VaR_{\alpha} = \mu + \sigma * t_v^{-1}(\alpha) \quad (2.14)$$

onde t_v representa a função distribuição *t* standard e v os graus de liberdade.

De acordo com Duda e Schmidt (2009) e Danielsson (2011) a distribuição *T-student* tem um ponto crucial que é a determinação dos graus de liberdade (v) porque é apresentado, através desse parâmetro, o quão gordas são as caudas.

No caso em que $v = \infty$, a distribuição *T-student* transforma-se na distribuição Normal.

Duda e Schmidt (2009) apresentam a seguinte fórmula de cálculo do VaR sob a distribuição *T-student*:

$$tVaR(\alpha) = -\mu + \sqrt{\frac{v-2}{v}} * \sigma * t_{v,\alpha} \quad (2.15)$$

onde μ representa a média, σ é uma amostra do desvio padrão e $t_{v,\alpha}$ é a distribuição de quantil α e v graus de liberdade.

Encontram-se outras críticas a esta abordagem, tais como a incapacidade de capturar a assimetria de distribuição dos retornos (Chu-Hsiung & Shan-Shan, 2006, citando Duda & Schmidt, 2009), a produção de estimativas demasiado elevadas e, um caso em que a distribuição *T-student* não é estável porque caso duas variáveis sigam uma distribuição *T-student*, a soma dessas variáveis não segue, necessariamente, uma distribuição *T-student* (Duda & Schmidt, 2009).

2.3 EXPECTED SHORTFALL

Expected Shortfall (ES) é um modelo apelativo quando o VaR se torna incapaz de preencher alguns dos requisitos necessários, tais como, a distinção entre carteiras com diferentes níveis de risco (Roccioletti, 2016). No entanto, Danielsson (2011), refere que ES surgiu devido à falta de subaditividade e de informação nas caudas que o modelo VaR fornece e que, responde à principal questão, de qual é a perda esperada quando as perdas excedem ao VaR.

Roccioletti (2016) define o *Expected Shortfall* como para uma dada perda L com $E(|L|) < \infty$ e função distribuição F_L , ao nível de confiança $\alpha \in (0,1)$:

$$ES_{\alpha}(L) = \frac{1}{1-\alpha} [E(L; L \geq q_{\alpha}) + q_{\alpha}(1-\alpha - P(L \geq q_{\alpha}))] \quad (2.16)$$

onde q_{α} representa o α -quantil de F_L .

O termo $q_{\alpha}(1-\alpha - P(L \geq q_{\alpha}))$ é interpretado como a parte excedente que é necessária adicionar ao valor esperado $E(L; L \geq q_{\alpha})$ quando o evento $\{L \geq q_{\alpha}\}$ tem probabilidade maior que $1-\alpha$.

Caso contrário, quando $P(L \geq q_{\alpha}) = 1-\alpha$, se a distribuição de probabilidade for contínua, o termo desaparece e a equação é reduzida para:

$$ES_{\alpha}(L) = \frac{1}{1-\alpha} E(L; L \geq q_{\alpha}) = E(L|L \geq VaR_{\alpha}) = TCE_{\alpha}(L) \quad (2.17)$$

onde TCE representa a *Tail Conditional Expectation* que se define como sendo a perda esperada sustentada pela porção de possibilidades mais infelizes, sendo esta a resposta à questão de Danielsson (2011).

Danielsson (2011), apresenta a definição sendo a perda/lucro esperado Q quando é inferior ao VaR negativo, isto é:

$$\begin{aligned} ES &= -E [Q | Q \leq -VaR(p)] \\ &= - \int_{-\infty}^{-VaR(p)} x f_{VaR}(x) dx \end{aligned} \quad (2.18)$$

onde $f_{VaR}(x)$ representa a densidade da cauda.

Roccioletti (2016), apresentada ainda uma definição equivalente em que o ES_{α} é a combinação de valores esperados no intervalo do nível de significância. Tendo que apresentar, primeiramente, a inversa da função distribuição:

$$F_L^{\leftarrow}(u) = \inf \{u \in \mathbb{R} : F_L(u) \geq \alpha\} \quad (2.19)$$

Então ES_{α} pode ser apresentado como a média de F_L^{\leftarrow} :

$$ES_{\alpha}(L) = \frac{1}{1-\alpha} * \int_{\alpha}^1 q_{\alpha}(F_L) du \quad (2.20)$$

Portanto, ES está assim ligado ao VaR através da seguinte fórmula:

$$ES_{\alpha}(L) = \frac{1}{1-\alpha} * \int_{\alpha}^1 VaR_u(F) du \quad (2.21)$$

Sendo que esta é a fórmula principal e a mais utilizada para o cálculo do VaR.

O ES é medido com maior incerteza do que o VaR. O primeiro passo na estimativa de ES é verificar o VaR, sendo o segundo passo obter a expectativa de observações de cauda. O que significa que existem pelo menos dois erros no ES. Com maior relevância, é o fato do ES ser mais difícil de testar do que o VaR, porque o procedimento do ES requer estimativas da expectativa da cauda para comparar com a previsão ES. Portanto, quanto ao *backtesting*, o ES só pode ser comparado através de um modelo enquanto o VaR pode ser comparado com as observações (Danielsson, 2011).

Medina e Munari (2016, citado em Osmundsen, 2016) apresenta como desvantagem o comportamento da cauda através do cálculo da média. E, evidencia que as médias são indicadores de risco fracos, logo, ES é considerada um método de risco potencialmente enganoso.

2.3.1 CÁLCULO DO ES

Nos modelos não paramétricos, como já foi referido, é utilizado o método da Simulação Histórica. Danielsson (2011) especifica o quão fácil é calcular o valor do ES através do método da Simulação Histórica: primeiro, o VaR é obtido através da Simulação Histórica, de seguida ES é calculado através das médias de todas as observações que são iguais ou mais negativas do que o $-VaR$.

É referido que o ES é igual à média das perdas não excedendo o quantil da distribuição (usando sinal positivo).

O autor supracitado menciona a importância do tamanho da janela de observações. Caso seja utilizado uma janela pequena os movimentos da Simulação Histórica serão maiores, isto é, as observações mais próximas dos extremos estão mais suscetíveis a mudanças de um dia para o outro do que observações centrais.

Contrariamente, a janela maior apresenta tanto vantagens como desvantagens ao serem menos sensíveis nos extremos e ao demorarem tempo a ajustar-se a mudanças estruturais no risco, respetivamente.

Izmaylov (2014) reforça a necessidade de existirem muitas observações para obter estimativas confiáveis, uma vez que são necessárias mais observações nos extremos da cauda da distribuição.

Nos modelos paramétricos são utilizadas duas distribuições para o cálculo do VaR: a distribuição Normal ou a distribuição *T-student*.

Relativamente à distribuição Normal, Roccioletti (2016) define que o ES para a distribuição Normal é dada por:

$$ES_{\alpha} = \mu + \sigma * \frac{\phi(\Phi^{-1}(\alpha))}{1 - \alpha} \quad (2.22)$$

onde Φ representa a função distribuição Normal standard e ϕ representa a sua densidade.

Segundo a distribuição Normal e Danielsson (2011), o ES é calculado através da seguinte fórmula:

Seja ϑ o valor do portfólio.

$$ES = -\vartheta * \frac{\sigma^2 * \phi(-VaR(p))}{p} \quad (2.23)$$

onde $\phi(x) = \frac{1}{\sqrt{2\pi}} * e^{-\frac{x^2}{2}}$ é a densidade da distribuição Normal.

Quanto à distribuição *T-student*, Roccioletti (2016) define que o ES para a distribuição *T-student* é dada por:

$$ES_{\alpha} = \mu + \sigma * \frac{g_v(t_v^{-1}(\alpha))}{1 - \alpha} * \left(\frac{v + (t_v^{-1}(\alpha))^2}{v - 1} \right) \quad (2.24)$$

onde t_v representa a função distribuição t standard, g_v representa a função densidade standard e v os graus de liberdade.

2.3.2 COMPARAÇÃO: VAR VS ES

Osmundsen (2016) afirma que as duas principais diferenças entre estes dois métodos são a subaditividade e o risco presente na cauda.

Defina-se, primeiramente, medida de risco coerente. A medida de risco ρ associa um número $\rho(X)$ a uma variável aleatória X representando os retornos dos ativos (números positivos para perdas). Este número indica a posição de risco, para cada uma das variáveis aleatórias X e Y (dependentes ou não) assim como para cada número n e para cada número positivo t , uma medida coerente de risco deve satisfazer as seguintes propriedades:

- I. Subaditividade: $\rho(X + Y) \leq \rho(X) + \rho(Y)$
- II. Homogeneidade: $\rho(t * X) = t * \rho(X)$
- III. Monotonicidade: $\rho(X) \leq \rho(Y)$, se $X \leq Y$

IV. Condição livre de risco: $\rho(X - n) = \rho(X) - n$

A propriedade que importa para o caso em avaliação é a propriedade I. que garante que a medida de risco se comporta razoavelmente quando se adiciona duas posições. E o VaR apenas satisfaz I. quando a distribuição da perda pertence à família da distribuição elíptica e tem variância finita. Ou seja, nestes casos o VaR fornece a mesma informação sobre as caudas que o ES. Por outro lado, ES prevalece sobre o VaR pois é uma medida de risco coerente.

ES oferece um valor de perda esperada em casos extremos, enquanto que o VaR apenas dá informação sobre o valor limite.

De acordo com Emmer, Kratz e Tasche (2015) para aplicações práticas o VaR ao não ser uma medida coerente pode não ser uma problemática, desde que os riscos subjacentes contenham variância finita ou média finita. Mas o fato de não cobrir os riscos da cauda, que passam além do seu valor, é realmente uma desvantagem. No entanto, o ES é uma medida que combate essa falta de subaditividade que o VaR apresenta.

Autores como Osmundsen (2016) e Karlsson e Zakrisson (2016) referem o ES como um modelo que expande a teoria implícita no modelo VaR e explicam o que pode acontecer nos casos em que a perda excede o VaR num determinado dia. Consequentemente, afirmam que o ES contém mais informação do que o VaR.

Agora, defina-se medida de risco eliciável:

Roccioletti (2016) define risco eliciável como, num quadro teórico-decisão, as medidas de risco para os quais é possível realizar uma validação e comparação do desempenho.

Por outro lado, Osmundsen (2016), define eliciável como, seja a estatística $\phi(Y)$ de uma variável aleatória Y , é designada como eliciável se minimiza o valor esperado de uma função pontuação (medida de erro) S , isto é:

$$\phi(Y) = \operatorname{argmin}_x E[S(x, Y)] \quad (2.25)$$

Para comparar os diferentes métodos de previsão é utilizada a função acima apresentada. Assim, o critério de desempenho é o seguinte:

$$\bar{S} = \frac{1}{n} \sum_{i=1}^n S(x_i, y_i) \quad (2.26)$$

onde os valores x_i representam os pontos de previsão, y_i representam as realizações correspondentes e S a função pontuação. Quanto menor for o valor da função, melhor é o processo que prevê a minimização.

Tanto Roccioletti (2016) como Osmundsen (2016) referem que o ES não é uma medida eliciável, pois esta característica está relacionada com um modelo de seleção e não com um modelo de teste. Segundo Acerbi e Szekely (2014), para realizar o *backtesting*, esta não é necessária. Por outro lado, Roccioletti (2016) afirma que a ausência desta característica pode originar uma dificuldade de estimação no *backtesting*. Embora, o VaR seja eliciável.

2.4 MODELOS DE SÉRIES TEMPORAIS

2.4.1 SÉRIE TEMPORAL

Uma série temporal pode ser definida como um conjunto de observações feitas de forma sequencial ao longo do tempo. A ordem dos dados recolhidos é importante pois, as observações próximas acabam por se influenciar de certa forma, apresentando assim uma tendência ao longo do tempo.

O estudo das séries temporais pode ser realizado através de uma série de valores anteriormente observados, de modo a proporcionar uma previsão de valores futuros.

Relativamente a séries temporais financeiras e os seus retornos existem duas características que as diferencia das restantes: a volatilidade e a leptocurtose elevada.

A volatilidade pode ser definida como sendo uma medida de dispersão em relação à média, num determinado intervalo de tempo, dos preços de um ativo, sendo usada para quantificar o risco desse mesmo ativo (Amaral, 2015).

Normalmente é calculada através do desvio-padrão anualizado da variação percentual das variações diárias, semanais ou mensais, sendo expressa em percentagem. Quanto menor for o valor da volatilidade, mais os valores se aproximam da média.

Nas séries temporais financeiras existe a possibilidade de ocorrência de *clusters* de volatilidade, estes caracterizam-se, segundo Campbell et al. (1997, citado em Araújo, 2010), pelo fato de períodos de grande volatilidade serem seguidos por outros períodos de grandes oscilações, da mesma forma, espera-se que pequenas alterações nos preços sejam consecutivamente seguidas por modificações de reduzida dimensão nos mesmos.

A leptocurtose (curtose elevada) é uma propriedade de uma distribuição estatística que se apresenta com caudas mais pesadas. Veja-se a Figura 2, onde é apresentada a distribuição Normal e a distribuição leptocúrtica (proveniente da Normal):

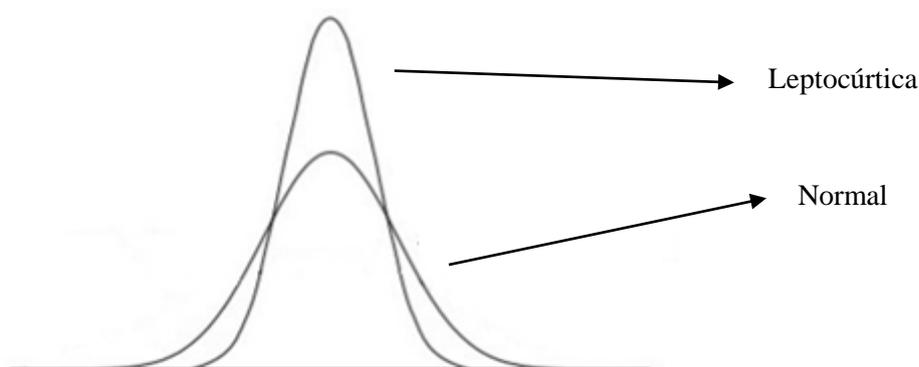


Figura 2.1: Distribuições: Normal e Leptocúrtica

Fonte: <https://www.slideshare.net/anselmorj/coeficiente-de-curtose-da-Normal>

2.4.2 ESTACIONARIDADE E TESTE DA RAIZ UNITÁRIA

A estacionaridade de uma série temporal representa que a distribuição dos dados demonstra estabilidade ao longo do tempo.

Diniz (1998, citando em Albuquerque, 2015) afirma que uma série temporal é estacionária caso o processo aleatório oscile em torno de um nível médio constante, mais precisamente tem média, variância e co-variância constante ao longo do tempo. Uma série cuja média e variância não sejam constantes, definem um processo não-estacionário. As séries temporais com tendência linear ou exponencial são exemplos de séries temporais com comportamentos não estacionários.

Para avaliar se uma série temporal financeira é estacionária recorre-se aos testes de raiz unitária onde a hipótese nula é a série não ser estacionária. Isto é:

$$\begin{aligned} H_0: \rho &= 1 \\ H_1: \rho &< 1 \end{aligned} \quad (2.27)$$

Portanto, se a hipótese nula for rejeitada então a série temporal é estacionária (não tem uma raiz unitária). Caso contrário, a série em estudo é não-estacionária (tem uma raiz unitária).

Iremos abordar testes para verificar a hipótese mencionada: ADF (Dickey-Fuller Aumentado, 1979) e o teste PP (Phillips-Perron, 1988), sendo que este é uma generalização do primeiro.

O teste ADF baseia-se no seguinte modelo:

$$\Delta y_t = y_t - y_{t-1} = \mu + \lambda t + (\rho - 1)y_{t-1} + \sum_{i=1}^p \gamma_i \Delta y_{t-i} + \varepsilon_t \quad (2.28)$$

O teste ADF deve ser aplicado considerando as seguintes hipóteses nulas:

- y_t é um passeio aleatório:

$$y_t = \rho y_{t-1} + \varepsilon_t \quad (2.29)$$

- y_t é um passeio aleatório com *drift* :

$$y_t = \rho y_{t-1} + \mu + \varepsilon_t \quad (2.30)$$

- y_t é um passeio aleatório com *drift* em torno de uma tendência determinística:

$$y_t = \rho y_{t-1} + \mu + \lambda t + \varepsilon_t \quad (2.31)$$

Segundo Amaral (2015), a utilização de dois testes diferentes pode ser útil uma vez que a hipótese nula do teste ADF pode não conduzir à rejeição da hipótese nula do teste de raiz unitária na presença de uma quebra estrutural no processo de geração de dados.

Na prática, têm-se verificado que as séries temporais financeiras são, na sua grande maioria, não-estacionárias, sendo então necessário trabalhar com as diferenças logaritmizadas das observações. Abaixo apresenta-se a fórmula de cálculo da primeira diferença dos logaritmos (retornos, rendibilidades) de uma série temporal de preços (P_t), isto é:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (2.32)$$

A importância em estabilizar (estacionarizar) uma série temporal não-estacionária consta no facto que a maioria dos modelos e métodos econométricos usuais só são aplicáveis a séries estacionárias.

2.4.3 ARMA

Os modelos *Autoregressive Moving Average (ARMA)* surgem da junção de outros dois modelos: *Autoregressive Model (AR)* e *Moving Averages Model (MA)*. Esta junção foi criada com a necessidade de modelar séries temporais estacionárias desde que as ordem p , que representam o número de termos de *AR*, e q , que representam o número de termos de *MA*, estivessem especificadas.

Segundo Amaral (2015), qualquer série Y_t pode ser modelada como uma combinação linear de valores passados de Y_t e de erros passados ε_t :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (2.33)$$

onde $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_p$ são constantes reais.

Este modelo assume que a volatilidade é constante ao longo do tempo. No entanto, raramente se verifica porque os períodos de instabilidade sucedem-se entre si, isto é, um período de variação elevada de rendimentos tende a ser seguido por um período de amplitude idêntica, mas em sentido contrário.

De acordo com Araújo (2010), a estrutura do processo *ARMA* (p, q) tem p termos auto-regressivos, que representam a ordem de defasagens do termo auto-regressivo (histórico da variável), e q termos de médias móveis (histórico do termo erro), que representa a ordem de defasagens do termo de médias móveis.

Algumas das vantagens deste modelo dizem respeito à possibilidade de efetuar previsões a partir dos dados observados, à simplicidade de procedimentos estatísticos e computacionais, o que resulta adequadamente como primeira aproximação.

No entanto, tem as suas limitações, uma delas reside em assumir que as séries temporais em análise são estacionárias, ou seja, média zero e variância constante, o que é errado, pois a grande maioria das séries revelam ser não-estacionárias. Servem, no entanto, para a equação de média dos modelos *GARCH*, que constituem um dos modelos base para a estimação do VaR e ES.

2.4.4 GARCH

Bollerslev (1986) apresentou o modelo *GARCH (Generalized Autoregressive Conditional Heteroskedasticity Model)* como uma extensão do modelo *ARCH*.

Segundo Amaral (2015), o modelo *GARCH* considera que a variância condicionada do processo de erro está relacionada, não só com os quadrados dos valores passados da série, mas também com as variâncias condicionadas passadas. Ao incluir a volatilidade dos defasamentos no modelo *ARCH*, este novo modelo tem potencial para capturar o impacto histórico dos retornos na volatilidade atual.

O modelo *GARCH* (p, q) considera que a variância condicional é parametrizada como uma função linear dos quadrados dos erros passados e das variâncias condicionais passadas, sendo então definido por:

$$\begin{cases} Y_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \alpha_0 + \alpha_1 Y_{t-1}^2 + \alpha_2 Y_{t-2}^2 + \dots + \alpha_q Y_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_p \sigma_{t-p}^2 \end{cases} \quad (2.34)$$

onde ε_t são variáveis aleatórias i.i.d. com média zero e variância igual a 1, e independentes das realizações passadas de Y_{t-i} e $\alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p$ são constantes reais.

Em regra geral, utiliza-se o método de máxima verosimilhança para estimar os parâmetros α_0, α_i ($i = 1, \dots, q$) e β_i ($i = 1, \dots, p$) do modelo.

Oliveira e Andrade (s.d.), consideram que se ε_t segue uma distribuição Normal, temos a seguinte função verosimilhança associada:

$$L = (2\pi)^{-\frac{(T-q-p)}{2}} \prod_{t=q+p+1}^T \left(\frac{1}{\sigma_t^2}\right)^{1/2} \exp\left(-\frac{Y_t^2}{2\sigma_t^2}\right) \quad (2.35)$$

Caso ε_t siga uma distribuição *T-student*, ou seja:

$$\varepsilon_t = \frac{t_v}{\sqrt{v/(v-2)}} \quad (2.36)$$

tal que t_v segue uma distribuição *T-student* com v graus de liberdade, a função verosimilhança associada é:

$$L = \prod_{t=q+p+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}} \left(\frac{1}{\sigma_t^2}\right)^{1/2} \left(1 + \frac{Y_t^2}{\sigma_t^2(v-2)}\right)^{-\frac{(v+1)}{2}} \quad (2.37)$$

Assim, as funções verosimilhança apresentadas podem ser maximizadas relativamente aos respetivos parâmetros desconhecidos.

Este modelo apresenta algumas desvantagens tais como: as restrições de não-negatividade podem não ser cumpridas e não consegue identificar os efeitos *leverage*.

2.5 BACKTESTING

Osmundsen (2016) define *backtesting* como um método utilizado para validar um modelo, onde os procedimentos estatísticos são usados para comparar as perdas reais com as previsões anteriores.

Danielsson (2011) afirma que o *backtesting* é útil para identificar os pontos fracos dos modelos de previsão de risco e fornece ideias para melhorar esses pontos fracos.

Segundo Roccioletti (2016), o termo *backtesting* é utilizado de diferentes formas em finanças:

1. Uma avaliação do desempenho teórico, pressuposto de uma estratégia de negociação planeada;
2. A avaliação dos modelos de risco financeiro, através de dados históricos sobre previsões de risco e realizações de lucros e perdas.

Como a escolha da metodologia do *backtesting* depende do tipo de previsões disponíveis, Emmer, Kratz e Tasche (2013) apresentam os seguintes métodos de *backtesting* usados para os diferentes tipos de previsão:

- i. Previsões de ponto para o valor de uma variável:

Geralmente, definidos em termos da expectativa condicional:

$$\mathbb{E}[Y_{t+k} \mid \mathcal{F}(Y_s, s \leq t)] \quad (2.38)$$

onde $\mathcal{F}(Y_s, s \leq t)$ representa a informação disponível até ao momento t da série temporal Y .

- ii. Previsões de alcance de probabilidade ou previsões de intervalos:

Como exemplo as previsões do VaR e do ES. É projetado um intervalo no qual o valor da previsão tem uma determinada probabilidade p . Usando como exemplo:

$$(-\infty, VaR_p(Y_{t+k})) \quad (2.39)$$

onde $VaR_p(Y_{t+k})$ é o p -quantil projetado de Y_{t+k} .

iii. Previsões da distribuição de probabilidade completa:

$$\mathbb{P}[Y_{t+k} \leq \mathcal{F}(Y_s, s \leq t)] \quad (2.40)$$

Ou a sua função densidade de probabilidade, caso exista.

De acordo com Roccioletti (2016) o VaR e o ES estão contidos no segundo grupo.

2.5.1 BACKTESTING VAR

Após o aparecimento do VaR surgiram as primeiras pesquisas e os primeiros resultados sobre o *backtesting*, de Christoffersen (1998), Kupiec (1995) e Hendricks (1996).

Segundo a abordagem de Christoffersen (1998), existe um *backtesting* popular baseado no *violation process* ou *hit sequence*. O *violation process* remete para a falha associada ao intervalo de confiança, isto é, quando excede o limite colocado. Por outro lado, o *hit sequence* é uma sequência de sucessos. Para este tipo, consideremos uma distribuição de perda contínua. Pela definição do VaR ao nível de confiança α , a probabilidade de *violation* de um valor do VaR é dado através da seguinte forma:

$$P(L > VaR_\alpha(L)) = 1 - \alpha \quad (2.41)$$

Consequentemente, podemos delinear o processo de *violation* como:

$$I_{t+1}(\alpha) = \mathbb{I}_{\{L(t+1) > VaR_\alpha(L(t+1))\}} \quad (2.42)$$

Esta função indicatriz apresenta que o *hit sequence* retorna valor 1 caso a perda no dia $t + 1$ seja maior do que o valor do VaR previsto, e retorna 0 caso contrário. Para testar o modelo é necessário construir uma sequência $\{I_{t+1}\}_{t=1}^T$ (onde T representa o número de dias do período de teste) que revele quando é que as excedências passadas são alcançadas. Contudo, deve ser esperado 1 com probabilidade $1 - \alpha$ e 0 com probabilidade α .

Seguindo esta linha de pensamento, é referido que um modelo de risco para estimação do VaR tem cobertura incondicional correta (*correct unconditional coverage*) se:

$$P(I_{t+1} = 1) = \mathbb{E}(I_{t+1}) = 1 - \alpha \quad (2.43)$$

e tem cobertura condicional correta (*correct conditional coverage*) se:

$$P_t(I_{t+1} = 1) = \mathbb{E}_t(I_{t+1}) = 1 - \alpha \quad (2.44)$$

É de notar que a cobertura condicional correta implica a cobertura incondicional correta, mas o inverso não é verificado. Além disso, este modelo satisfaz a condição de independência, isto é:

$$I_{t+1}(\alpha) \text{ e } I_{s+1}(\alpha) \text{ são independentes para momentos } t \text{ e } s \text{ diferentes.}$$

Os *hits* do VaR são independentes e identicamente distribuídos pela distribuição de Bernoulli, com probabilidade de sucesso $1 - \alpha$. Podendo então questionar a hipótese:

$$H_0: I_{t+1} \sim i. i. d. \text{ Bernoulli } (1 - \alpha) \quad (2.45)$$

Ou seja é necessário avaliar duas hipóteses distintas:

1. Hipótese de cobertura incondicional:

$$H_0: \mathbb{E}(I_{t+1}) = \pi = 1 - \alpha \quad (2.46)$$

2. Hipótese de cobertura condicional:

$$H_0: \mathbb{E}_t(I_{t+1}) = \pi_{t+1|t} = 1 - \alpha \quad (2.47)$$

onde π representa a medida da amostra. Uma das grandes questões é a escolha do nível de significância do teste, devendo ser escolhido tendo em conta os dois seguintes tipos de erros:

- i. Erro tipo I: probabilidade de rejeitar o modelo correto.
- ii. Erro tipo II: probabilidade de não rejeitar o modelo errado.

Aumentar o nível de significância implica um maior erro do tipo I mas, em oposição, um menor erro do tipo II e vice-versa. Ou seja, o ideal seria uma minimização destes dois erros.

2.5.1.1 Testes de Cobertura Incondicional

Primeiramente, é necessário testar se a probabilidade incondicional da *violation* no modelo de risco, π , difere significativamente da probabilidade, p . Neste caso, basta que se verifique o número de *violations*, isto é, se a primeira sequência, segue uma distribuição Binomial:

$$f(x) = \binom{T}{x} p^x (1-p)^{T-x} \quad (2.48)$$

Assim que o número de observações aumenta, a distribuição Binomial pode ser aproximada à distribuição Normal. Portanto, para avaliar a hipótese nula, podemos simplesmente usar o teste médio simples:

$$z = \frac{x - pT}{\sqrt{p(1-p)T}} \approx N(0,1) \quad (2.49)$$

onde pT é o número esperado de exceções e $p(1-p)T$ representa a sua variância.

Podemos também utilizar testes de razão de verossimilhança (*likelihood ratio tests*) recorrendo a Kupiec (1995). Um deles denominado por *POF-test*, *Proportion of Failure*.

Segundo Karlsson e Zakrisson (2016), este teste mede se a quantidade de exceções está de acordo com o nível de probabilidade, ou seja, se um nível de probabilidade $1-p$ for utilizado para o cálculo do VaR, a perda irá exceder o VaR estimado em $p\%$ dos períodos. Caso a perda raramente exceda o VaR estimado, a estimativa está sistematicamente a sobrevalorizar o risco de mercado, logo não é uma medida precisa. O mesmo acontece se a perda exceder o VaR mais do que $p\%$ dos dias, isso indicaria uma subvalorização do risco.

Portanto, sob a hipótese nula, a proporção de vezes que a perda real excede a estimativa será igual a p .

Roccioletti (2016) apresenta a hipótese nula da seguinte forma:

$$H_0: p = \hat{p} = \frac{x}{T} \quad (2.50)$$

onde x representa o número de dias em que a perda excedeu o VaR estimado e T o número de dias totais que o VaR foi estimado. Remetendo para a ideia de que é necessário verificar se existe uma grande discrepância entre a taxa de falha observada, \hat{p} e a taxa de falha teórica p .

O teste é construído de maneira simples tendo em consideração que este verifica se há probabilidade de as estimativas serem confiáveis, dado um intervalo de confiança:

$$LR_{uc} = -2 \ln \left(\frac{(1-p)^{T-x} p^x}{[1 - (\frac{x}{T})]^{T-x} (\frac{x}{T})^x} \right) \quad (2.51)$$

Sob a hipótese nula, LR_{uc} é assintoticamente χ_1^2 (distribuição Qui-Quadrado) com um grau de liberdade. Caso o valor da estatística de teste seja superior ao valor da Qui-Quadrado, a hipótese nula será rejeitada e afirma-se que o modelo é incorreto. É de sublinhar que o poder deste teste aumenta quando o tamanho da amostra aumenta. Assim, quando existe uma quantidade considerável de observações disponíveis, é possível rejeitar um modelo que não é preciso sem grandes dificuldades.

Uma desvantagem do *POF-test* é de que este desconsidera o momento em que ocorrem as perdas. E, como consequência, pode não rejeitar o modelo que produz *violations* agrupadas do VaR, sendo o que se pretende evitar, uma vez que as grandes perdas que acontecem de forma rápida são mais propícias a levar a eventos indesejáveis. Sendo esta a razão que nos leva aos testes de cobertura condicional.

2.5.1.2 Testes de Cobertura Condicional

Andersen *et al.* (2005, citado em Roccioletti, 2016) afirma que existe uma forte evidência da volatilidade variável no tempo dos retornos diários dos ativos. Se o modelo de risco não reconhecer esse comportamento, o VaR responderá demasiado tarde às mudanças nas condições do mercado e as ruturas do VaR aparecerão agrupadas ao longo do tempo.

Portanto, antes de formular um teste de cobertura condicional, é necessário abordar o problema de independência: num modelo preciso, a exceção hoje não deve depender da ocorrência do dia anterior.

A maneira mais simples de testar a dinâmica na análise de séries temporais, é avaliando a função de autocorrelação e, em seguida, recorrer aos testes Portmanteau ou Ljung-Box¹¹.

Então, seja γ_k a autocorrelação no intervalo k para o processo de *violation*.

Traçando a função de correlação (para $k = 1, \dots, m$) obtêm-se o grau de conexão entre o excedente num dos últimos m dias de transação e o *hit* hoje. E assim, a hipótese nula será:

$$H_0: \gamma_k = 0 \text{ para } k = 1, \dots, m \quad (2.52)$$

Sendo que a estatística de teste é:

$$LB(m) = T(T+2) \sum_{k=1}^m \frac{\gamma_k^2}{T-k} \sim \chi_m^2 \quad (2.53)$$

onde χ_m^2 representa a distribuição Qui-Quadrado com m graus de liberdade. De acordo com Berkowitz, Christoffersen e Pelletier (2007, citado em Roccioletti, 2016) ao definir $m = 5$ obtêm-se um teste viável para um cenário realista de *backtesting* ao VaR diário.

Também é possível adotar uma abordagem de verossimilhança para a realização de testes de independência. Recorrendo a Christoffersen (1998), apresenta-se um teste denominado por *Interval Forecast Test*.

¹ Para definição dos testes Portmanteau ou Ljung-Box ver: Ana Amaro - Uma introdução à Metodologia Box-Jenkins (2012).

Suponhamos que a estrutura de dependência de uma *hit sequence* pode ser descrita como uma cadeia de Markov de primeira ordem com uma matriz de probabilidade de transação:

$$\Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix} \quad (2.54)$$

Podemos interpretar os números apresentados na matriz como:

- i. $1 - \pi_{01}$ é a probabilidade de uma *non-violation* após uma *non-violation*.
- ii. π_{01} é a probabilidade de ocorrer uma *violation* amanhã, condicionado a hoje não ocorrerem *violations*.
- iii. $1 - \pi_{11}$ é a probabilidade de uma *non-violation* sucedendo uma *violation*.
- iv. π_{11} é a probabilidade de amanhã ocorrer uma *violation*, sabendo que hoje ocorreu uma *violation*.

Se está disponível uma amostra de T observações então podemos escrever a função de verossimilhança do processo de Markov de primeira ordem como:

$$L(\Pi_1) = (1 - \pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1 - \pi_{11})^{T_{10}} \pi_{11}^{T_{11}} \quad (2.55)$$

onde T_{ij} com $i, j = 0, 1$ é o número de observações com um j sucedendo um i .

Encontrar as estimativas de Máxima Verossimilhança significa encontrar as primeiras derivadas em relação a π_{01} e a π_{11} , e igualá-las a zero. Obtendo como resultado:

$$\hat{\pi}_{01} = \frac{T_{01}}{T_{00} + T_{01}}, \hat{\pi}_{11} = \frac{T_{11}}{T_{10} + T_{11}} \quad (2.56)$$

E dada a definição de probabilidade temos:

$$\hat{\pi}_{00} = 1 - \hat{\pi}_{01}, \hat{\pi}_{10} = 1 - \hat{\pi}_{11} \quad (2.57)$$

Com isto, interessa descobrir se $\hat{\pi}_{01}$ é estatisticamente diferente de $\hat{\pi}_{11}$ e, mais especificamente, se $\hat{\pi}_{11}$ é maior do que $\hat{\pi}_{01}$. O que implicaria que seria mais provável existir duas excedências consecutivas do que existir uma *violation* após de uma *non-violation*.

Para testar a hipótese de independência temos:

$$H_0: \hat{\pi}_{01} = \hat{\pi}_{11} \quad (2.58)$$

Com um teste de razão de verossimilhança subsequente:

$$LR_{ind} = -2 \ln \left(\frac{(1 - \pi)^{T_{00} + T_{10}} \pi^{T_{01} + T_{11}}}{(1 - \pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1 - \pi_{11})^{T_{10}} \pi_{11}^{T_{11}}} \right) \quad (2.59)$$

onde,

$$\pi = \frac{T_{01} + T_{11}}{T_{00} + T_{01} + T_{10} + T_{11}} \quad (2.60)$$

Assintoticamente, este teste segue uma distribuição Qui-Quadrado com um grau de liberdade, $LR_{ind} \sim \chi_1^2$.

Contudo, ao realizar o teste LR_{ind} é possível que se encontre amostras em que $T_{11} = 0$. E, neste caso, substitui-se a função verossimilhança acima descrita por:

$$L(\Pi_1) = (1 - \hat{\pi}_{01})^{T_{00}} \hat{\pi}_{01}^{T_{01}} \quad (2.61)$$

Em última análise, é importante testar simultaneamente as propriedades de um modelo apropriado do VaR: a taxa de falha correta e a independência das exceções. É possível testar a independência e a cobertura correta usando o teste de cobertura condicional:

$$LR_{cc} = LR_{uc} + LR_{ind}, LR_{cc} \sim \chi_2^2 \quad (2.62)$$

A estrutura de Christoffersen (1998) permite verificar qual o motivo de não passar no teste, se é causado pela cobertura imprecisa, se pelas falhas agrupadas ou até ambos os motivos.

Recorremos então ao teste denominado por *Mixed Kupiec Test*.

O teste de Christoffersen (1998) não consegue alcançar todas as formas de dependência, uma vez que apenas considera a agregação de resultados de dois dias sucessivos.

Com isto, Hass (2001) apresenta um teste melhorado, tanto para a independência como para a cobertura, fazendo uma mistura de ideias entre Kupiec (1995) e Christoffersen (1998). Este teste mede o tempo entre exceções, podendo então alcançar as várias formas de dependência.

Para construir este teste é necessário construir uma estatística de razão de verossimilhança para cada exceção i .

$$LR_i = -2 \ln \left(\frac{p(1-p)^{v_i-1}}{\left(\frac{1}{v_i}\right) \left(1 - \frac{1}{v_i}\right)^{v_i-1}} \right) \quad (2.63)$$

onde v_i é o tempo entre as exceções i e $i - 1$.

Após o cálculo das estatísticas LR , realizamos um teste de independência assumindo que existem n *violations*.

$$LR_{ind^n} = \sum_{i=2}^n \left[-2 \ln \left(\frac{p(1-p)^{v_i-1}}{\left(\frac{1}{v_i}\right) \left(1 - \frac{1}{v_i}\right)^{v_i-1}} \right) \right] - 2 \ln \left(\frac{p(1-p)^{v-1}}{\left(\frac{1}{v}\right) \left(1 - \frac{1}{v}\right)^{v-1}} \right) \quad (2.64)$$

$LR_{ind^n} \sim \chi_n^2$, ou seja, segue a distribuição Qui-Quadrado com n graus de liberdade.

Contudo, podemos combinar dois testes para obter este teste:

$$LR_{mix} = LR_{un} + LR_{ind^n} \quad (2.65)$$

e a estatística $LR_{mix} \sim \chi_{n+1}^2$, segue a distribuição Qui-Quadrado com $(n + 1)$ graus de liberdade.

2.5.2 BACKTESTING ES

De acordo com Roccioletti (2016) e Acerbi e Szekely (2014), para realizar um *backtesting* ao ES, apresenta-se três metodologias não paramétricas e independentes.

Para começar é essencial caracterizar a estrutura dos testes: tendo em conta apenas os testes de cobertura incondicional, assumimos que a independência dos acontecimentos presentes na cauda é testada separadamente através das falhas agrupadas do VaR.

Portanto, seja L_t a perda do banco no dia t para cada t pertencente ao período de teste ($t = 1, \dots, T$). Essas perdas seguem uma distribuição real, no entanto desconhecida, F_L e previstas segundo uma distribuição de previsão, P_L .

As variáveis aleatórias $\vec{L} = \{L_t\}$ são assumidas como independentes mas não igualmente distribuídas. Além disso, não há restrição sobre a variação de F_L e P_L ao longo do tempo.

Assume-se que as distribuições são contínuas e estritamente crescentes. A razão para essa suposição deve-se, sob estes pressupostos, ao ES pode ser visto como o *Tail Conditional Expectation* (TCE).

Por fim, a hipótese nula será a previsão exata enquanto as hipóteses alternativas serão a suposição de que as previsões da medida de risco são muito baixas.

2.5.2.1 Teste 1

Este teste segue, exatamente, a ideia do TCE. De facto, dado que:

$$ES_{\alpha,t}(L) = E[L_t | L_t > VaR_{\alpha,t}(L)] \quad (2.66)$$

Obtêm-se assim:

$$E \left[\frac{L_t}{ES_{\alpha,t}(L)} - 1 \mid L_t - VaR_{\alpha,t}(L) > 0 \right] = 0 \quad (2.67)$$

Considere-se agora que o modelo $VaR_{\alpha,t}(L)$ já foi testado. Quer-se investigar a magnitude das *violations* verificadas contra as previsões do modelo.

Seja $I_t = (L_t - VaR_{\alpha,t}(L) > 0)$ o indicador da função de uma falha (*violation*) do VaR. Se $N_t = \sum_{t=1}^T I_t > 0$, pode-se facilmente construir a seguinte estatística de teste:

$$Z_1(\vec{L}) = \frac{\sum_{t=1}^T \frac{L_t \cdot I_t}{ES_{\alpha,t}}}{N_t} - 1 \quad (2.68)$$

Construímos apenas uma média de *violations* do VaR normalizadas $\left(\frac{L_t \cdot I_t}{ES_{\alpha,t}}\right)$.

A hipótese nula subjacente é dada por:

$$H_0: P_t^{[1-\alpha]} = F_t^{[1-\alpha]}, \forall t \quad (2.69)$$

onde $P_t^{[1-\alpha]}(l) = \min\left(1, \frac{1-P_t(l)}{1-\alpha}\right)$, é a distribuição da cauda para $l > VaR_{\alpha,t}$.

As alternativas são:

$$\begin{aligned} H_1: ES_{\alpha,t}^F &= ES_{\alpha,t}, \forall t \\ VaR_{\alpha,t}^F &= VaR_{\alpha,t}, \forall t \end{aligned} \quad (2.70)$$

Tendo em consideração que $ES_{\alpha,t}^F$ e $VaR_{\alpha,t}^F$ representam o valor das medidas de risco quando $L \sim F_L$.

Note-se que o modelo VaR_{α} previsto está ainda correto sob H_1 . Isto significa que este teste recorre a um teste secundário preliminar (não rejeitado).

Sendo uma média dos excedentes do VaR, este teste é indiferente para um número exageradamente grande de exceções então é importante, em primeiro lugar, validar o modelo VaR. Sob essas condições podemos afirmar que:

- I. $ES_{H_0}[Z_1 | N_T > 0] = 0$
- II. $ES_{H_1}[Z_1 | N_T > 0] > 0$

Assim, é de esperar que o valor alcançado $Z_1(\vec{l})$ seja zero, sendo indicativo de que existe um problema quando é positivo.

Este teste é construído com base nos resíduos do modelo que, sob a hipótese nula das previsões corretas do ES, devem comportar-se como uma amostra i.i.d. com média zero. Para investigar a hipótese nula, é usado um teste de *bootstrap*² que não faz qualquer suposição sobre a distribuição subjacente dos resíduos.

2.5.2.2 Teste 2

Este segundo teste deriva da representação do ES como uma expectativa incondicional:

$$ES_{\alpha,t}(L) = E \left[\frac{L_t \cdot I_t}{ES_{\alpha,t}} \right] \quad (2.71)$$

Podemos então delinear a seguinte estatística de teste:

$$Z_2(\vec{L}) = \frac{\sum_{t=1}^T \frac{L_t \cdot I_t}{ES_{\alpha,t}}}{T(1-\alpha)} - 1 \quad (2.72)$$

A hipótese nula e alternativa serão, respetivamente:

$$\begin{aligned} H_0: P_t^{[1-\alpha]} &= F_t^{[1-\alpha]}, \forall t \\ H_1: ES_{\alpha,t}^F &\geq ES_{\alpha,t}, \forall t \\ VaR_{\alpha,t}^F &\geq VaR_{\alpha,t}, \forall t \end{aligned} \quad (2.73)$$

Da mesma forma anterior, temos:

- I. $ES_{H_0}[Z_2] = 0$
- II. $ES_{H_1}[Z_2] > 0$

A última desigualdade vem do facto de que a razão média será estritamente superior a 1. Podemos observar que neste teste se avalia tanto a frequência como o peso dos $(1-\alpha)$ eventos presentes na cauda.

Tendo assim que:

$$Z_2(\vec{L}) + 1 = \left(1 + Z_1(\vec{L})\right) \frac{N_t}{T(1-\alpha)} \quad (2.74)$$

Mantendo também a seguinte relação:

$$Z_2(\vec{L}) = \left(1 + Z_1(\vec{L})\right) \frac{N_t}{T(1-\alpha)} - 1 \quad (2.75)$$

O teste Z_2 apenas exige que se registre dois números por dia: o peso $(L_t \cdot I_t)$ da *violation* do $VaR_{\alpha,t}$ e a previsão do $ES_{\alpha,t}$.

2.5.2.3 Teste 3

Este teste remete para a ideia de que é possível fazer um teste às caudas de um modelo ao verificar se as classificações observadas $U_t = P_t(L_t)$ são $U(0,1)$ i.i.d., como deveriam ser caso a distribuição

² O *bootstrap* é um método de estimação de parâmetros e distribuições.

prevista esteja certa. A classificação U_t é simplesmente a probabilidade acumulada associada à perda observada L_t .

Considera-se um modelo que, ao fim de cada dia, gera uma estimativa da distribuição acumulada para o retorno do dia seguinte (P_t). Isto significa que, dada a performance atual, é possível calcular a probabilidade (implícita pelo modelo de risco) de experienciar uma perda abaixo da perda corrente. Denotamos essa probabilidade por U_t .

Caso se esteja a utilizar o modelo de risco certo para aproximar a distribuição de perda real desconhecida, não deve ser possível prever. Por esta razão, as séries temporais das classificações observadas devem ser distribuídas independentemente ao longo do tempo como uma variável $U(0,1)$.

Por outras palavras, para realizar o *backtest* às caudas da distribuição de perda, é necessário verificar se as variáveis $\vec{U} = \{U_t\}$ são distribuídas uniformemente no intervalo (0,1). É necessário ainda, para o *backtesting* ao ES, adaptar esta ideia para criar um teste específico para a métrica de risco usada.

Deve-se atribuir a cada quantil o seu peso em termos de dinheiro e depois é preciso construir um estimador do ES baseado no aspeto i.i.d. $\vec{Y} = \{Y_t\}$ de uma distribuição geral (contínua e estritamente crescente):

$$\widehat{ES}_\alpha^{(N)} = \frac{1}{[N(1-\alpha)]} \sum_i^{[N(1-\alpha)]} Y_{i:N} \quad (2.76)$$

onde $Y_{i:N}$ é as estatísticas ordenadas.

Podemos então delinear a seguinte estatística de teste:

$$Z_3(\vec{X}) = \frac{1}{T} \sum_{t=1}^T \frac{\widehat{ES}_\alpha^{(T)}(P_t^{-1}(\vec{U}))}{E_V(\widehat{ES}_\alpha^{(T)}(P_t^{-1}(\vec{V})))} - 1 \quad (2.77)$$

onde \vec{V} são $U(0,1)$ i.i.d..

A ideia é recalculer o ES como a média acima do quantil $P_t^{-1}(\vec{U})$, para cada dia passado $t = 1, \dots, T$ e então retirar uma média do resultado. Nesta estatística de teste não se normaliza por $ES_{\alpha,t}$ como nos testes anteriores mas para compensar divide-se por uma estimativa de amostra finita.

As hipóteses subjacentes envolvem agora todas as distribuições:

$$\begin{aligned} H_0: P_t &= F_t, \forall t \\ H_1: P_t &\geq F_t, \forall t \end{aligned} \quad (2.78)$$

Da mesma forma anterior, temos:

- I. $ES_{H_0}[Z_3] = 0$
- II. $ES_{H_1}[Z_3] > 0$

Comparando com os outros dois testes, este é, de facto, o menos intuitivo, mas é bastante generalizado.

Com tudo isto, Danielsson (2011) apresenta dois problemas com o *backtesting*:

- i. Pressupõe que não existam interrupções estruturais nos dados ao longo do período em avaliação. Os mercados financeiros, por outro lado, estão em constante evolução e as novas tecnologias, os ativos, os mercados e as instituições afetam as propriedades estatísticas dos preços de mercado. É improvável que uma propriedade do mercado há dez anos seja igual

há de hoje, implicando que o modelo que funcionou bem há dez anos atrás poderá não funcionar bem hoje.

- ii. *Data mining* e a integridade intelectual. Isto é, em teoria, o *backtesting* é apenas estatisticamente válido se não existir conhecimento prévio dos dados na janela de teste (ou seja, é criado um modelo de risco, executa-se uma vez e toma-se uma decisão relativamente ao modelo). Se iterarmos o processo, o modelo será ajustado aos resultados, falhando nos pressupostos subjacentes. Os modelos que são criados desta forma têm uma probabilidade maior de ter uma *performance* pobre na previsão futura do risco.

3 ESTIMAÇÃO E BACKTESTING DO VAR E ES

Neste capítulo é descrito estatisticamente a evolução da série das rendibilidades do PSI20, de modo a estimar o risco que lhe está associado, recorrendo aos modelos enunciados para previsão do VaR e ES.

A série em estudo é uma série diária (5 dias por semana) de cotação de fecho do PSI20. Compreendida entre 02 de janeiro de 2006 a 31 de janeiro de 2017, tendo um total de 2892 observações, retirados da base de dados: DATASTREAM.

3.1 ANÁLISE DA SÉRIE E DOS SEUS RETORNOS

Recorrendo ao programa *Eviews* e utilizando as observações disponíveis, foi obtida a representação gráfica da série temporal em níveis do PSI20 para o período acima referido, como se ilustra na Figura 3.1.



Figura 3.1: Evolução da cotação do PSI20 entre 02/01/2006 e 31/01/2017

Podemos observar (Figura 3.1) que a série apresenta uma evolução não linear. Podemos também realçar a grande queda do índice no ano de 2008, sendo em concordância com a crise económica *Subprime* iniciada a meados de 2007. E em 2009, verifica-se uma forte tendência de crescimento, embora entre 2009 e meados de 2011 se verifique uma oscilação do valor do índice. Observamos ainda, alguns grandes decréscimos em meados de 2011, meados de 2014 e em meados de 2015. As cotações do PSI20 atingem os seus valores extremos em 17 de julho de 2007 com 13702.00 pontos (valor máximo) e em 27 de junho de 2016 com 4260.13 pontos (valor mínimo).

Na Figura 3.2 apresenta-se a análise da estatística descritiva para a série considerada:

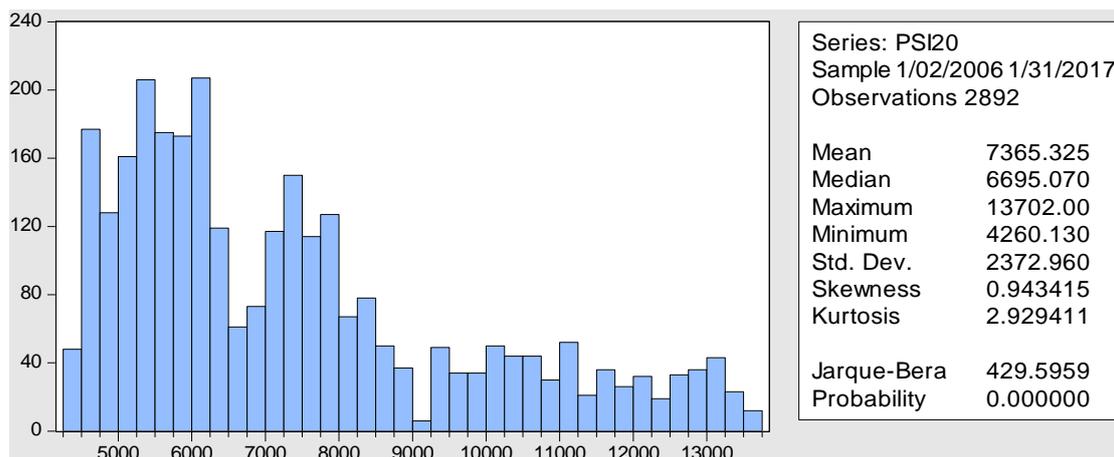


Figura 3.2: Estatística descritiva da série PSI20

Podemos concluir que a série é mesocúrtica pois o coeficiente de curtose é bastante próximo de três. Apresenta-se ainda uma média não nula e uma assimetria positiva moderada. Relativamente ao teste de Normalidade de Jarque-Berra (que apresenta como hipótese nula a Normalidade), verifica-se um *p-value* inferior ao nível de significância de 5%, o que implica que a hipótese nula seja rejeitada, logo, a série não segue uma distribuição Normal.

Como a série apresenta uma grande variabilidade, aponta para que não seja estacionária, portanto recorremos ao teste de raiz unitária para poder confirmar a sua estacionaridade.

Aplica-se então os testes ADF e PP à série do PSI20 de forma a avaliar a estacionaridade da mesma. Para a realização do teste ADF foi utilizado o critério de informação de Schwarz (SIC) com um número (automático) máximo de 27 desfasamentos. Para o teste PP, o método de estimação espectral é o de Bartlett Kernel com o critério de seleção dos desfasamentos automáticos de Newey-West Bandwidth.

Tabela 3.1: Teste ADF aplicado ao índice PSI20

Null Hypothesis: PSI20 has a unit root		
Exogenous: Constant		
Lag Length: 1 (Automatic - based on SIC, maxlag=27)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.862386	0.8004
Test critical values: 1% level	-3.432422	
5% level	-2.862341	
10% level	-2.567241	
*MacKinnon (1996) one-sided p-values.		

Analisando os resultados obtidos do teste ADF (Tabela 3.1), conclui-se que o *p-value* obtido é superior aos níveis de significância de referência (sendo estes 1%, 5% e 10%), não se rejeita a hipótese nula, ou seja, a série tem raiz unitária, consequentemente a série é considerada não estacionária.

Agora recorrendo ao teste PP (Tabela 3.2) e observando o output abaixo conclui-se, igualmente, que a série não é estacionária devido ao seu *p-value* ser superior aos valores de significância de referência.

Tabela 3.2: Teste PP aplicado ao índice PSI20

Null Hypothesis: PSI20 has a unit root		
Exogenous: Constant		
Bandwidth: 8 (Newey-West automatic) using Bartlett kernel		
	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-0.843832	0.8059
Test critical values: 1% level	-3.432421	
5% level	-2.862341	
10% level	-2.567241	

*MacKinnon (1996) one-sided p-values.

Concluimos então que a série do PSI20 não é estacionária em nível. Para estacionarizar a série vamos recorrer aos seus retornos (a primeira diferença dos logaritmos).

Na Figura 3.3 é apresentado o gráfico da série dos retornos logaritmizados (rendibilidades) do índice PSI20.

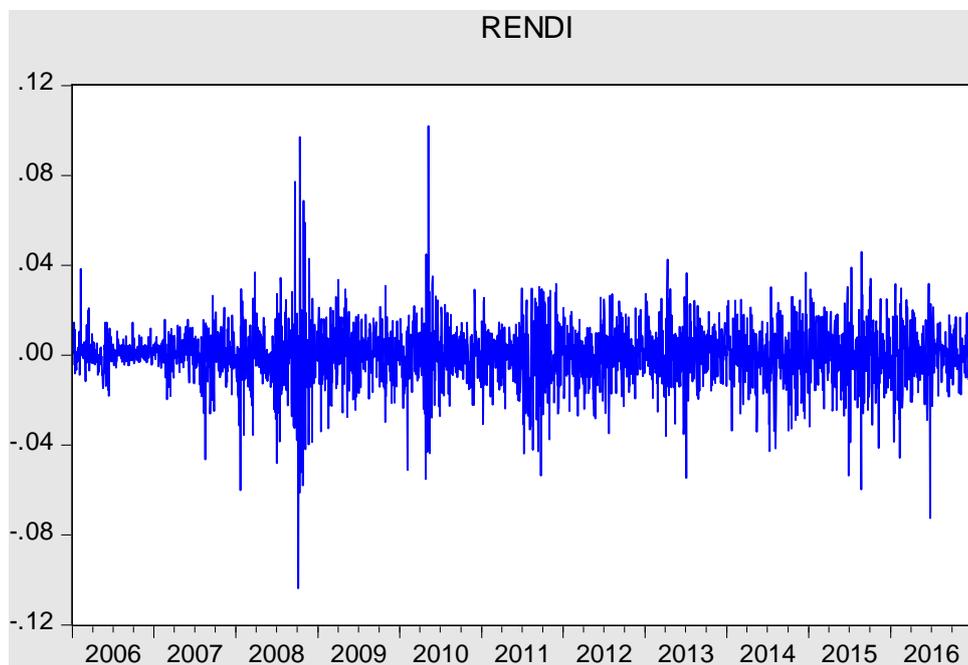


Figura 3.3: Série logaritmizada dos retornos do índice PSI20

Esta transformação nas observações permitiu que se criasse uma estabilização no valor médio da série e que se diminuísse a variabilidade dos dados. Observam-se *clusters* na volatilidade e alguns valores extremos, fenómenos bastante comuns nas séries financeiras.

A estatística descritiva da série dos retornos do PSI20 é representada na Figura 3.4.

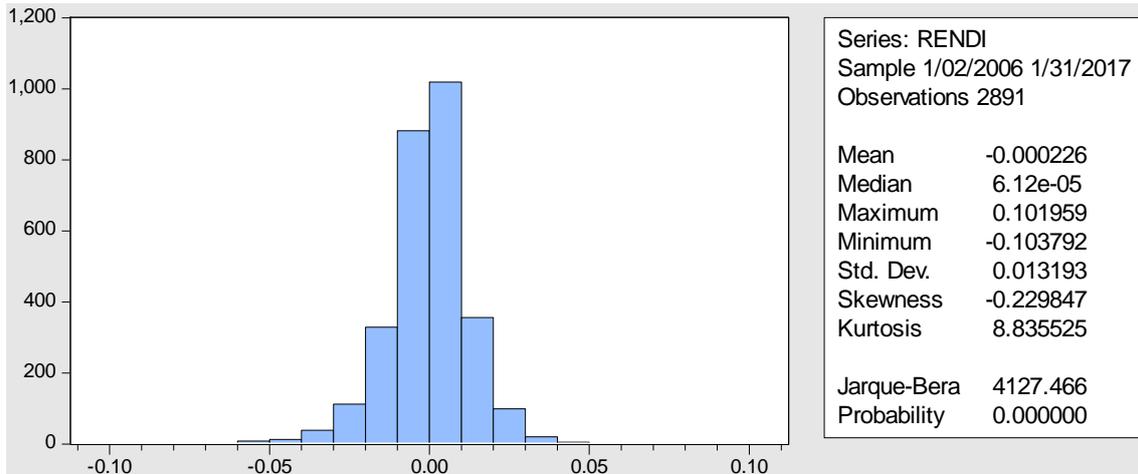


Figura 3.4: Estatística descritiva da série logaritmizada dos retornos do índice PSI20

Conclui-se, através desta análise descritiva, que se rejeita a hipótese de Normalidade, pois o *p-value* permanece próximo de 0, segundo o teste de Jarque-Berra. Contudo, o valor médio está bastante próximo de zero e o desvio padrão é significativamente menor do que na série em níveis. Ao contrário da série original apresenta uma assimetria negativa e o coeficiente da curtose é superior a três, logo concluímos que a série é leptocúrtica.

Da mesma forma, são realizados os testes de raiz unitária para as rendibilidades (ADF e PP), obtendo os seguintes resultados (Tabela 3.3 e 3.4):

Tabela 3.3: Teste ADF aplicado à série logaritmizada dos retornos do índice PSI20

Null Hypothesis: RENDI has a unit root		
Exogenous: None		
Lag Length: 0 (Automatic - based on SIC, maxlag=27)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-49.48728	0.0001
Test critical values:		
1% level	-2.565762	
5% level	-1.940933	
10% level	-1.616626	
*MacKinnon (1996) one-sided p-values.		

Através do teste de ADF, conclui-se que a série dos retornos logaritmizados é estacionária, uma vez que o *p-value* obtido é inferior aos níveis de significância de referência (sendo estes 1%, 5% e 10%), portanto rejeita-se a hipótese nula, ou seja, a série não tem raiz unitária.

Tabela 3.4: Teste PP aplicado à série logaritmizada dos retornos do índice PSI20

Null Hypothesis: RENDI has a unit root Exogenous: None Bandwidth: 7 (Newey-West automatic) using Bartlett kernel		
	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-49.35489	0.0001
Test critical values: 1% level	-2.565762	
5% level	-1.940933	
10% level	-1.616626	

*MacKinnon (1996) one-sided p-values.

Da Tabela 3.4 (Teste de PP), verificamos que o *p-value* é inferior aos valores de significância de referência logo, rejeita-se a hipótese nula, conclui-se que a série dos retornos do índice PSI20 é estacionária.

3.2 ESTIMAÇÃO DO VAR E DO ES

Nesta secção vão ser calculadas as duas medidas de risco já apresentadas: VaR e ES, utilizando o software MATLAB e parte das rotinas de Roccioletti (2016) aplicadas à série financeira em estudo (PSI20). Após a estimação do risco associado a este índice, procede-se ao estudo da *performance* dos modelos utilizando alguns métodos de *backtesting*.

São consideradas duas distribuições distintas nos modelos paramétricos, nomeadamente, a distribuição Normal e a distribuição *T-student* e, a distribuição *Kernel* associada ao modelo não-paramétrico do tipo Simulação Histórica.

Começa-se por apresentar, nas Figuras 3.5-3.7, as distribuições Normal, *T-student* e *Kernel* ajustadas ao histograma dos retornos do índice PSI20.

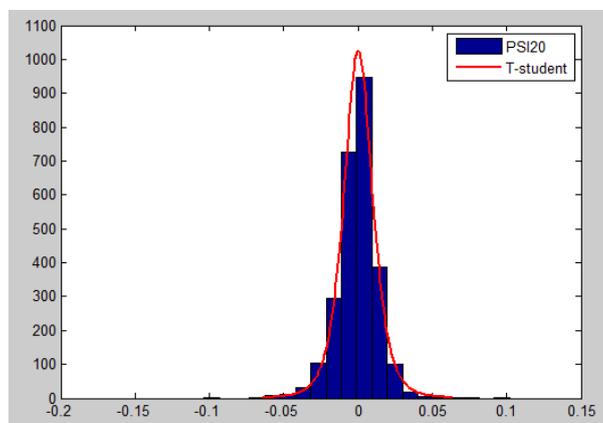


Figura 3.5: PSI20 vs Distribuição Normal

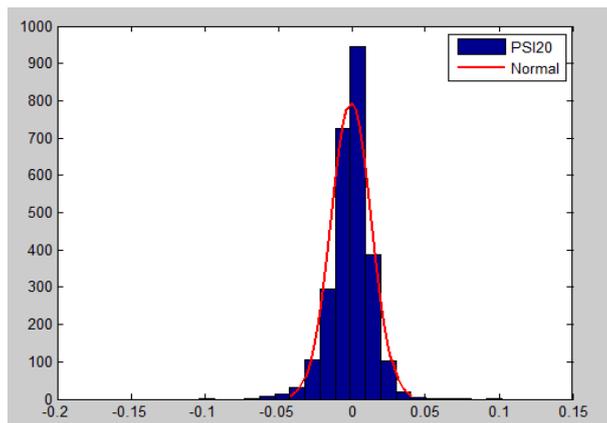


Figura 3.6: PSI20 vs Distribuição T-student

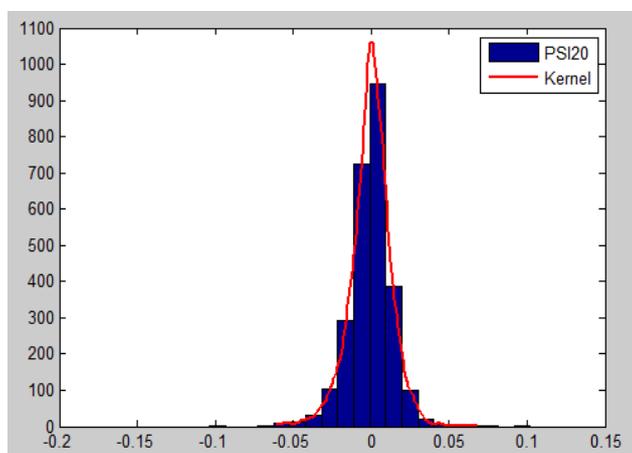


Figura 3.7: PSI20 vs Distribuição Kernel

É possível entender o quanto a distribuição Normal subestima o risco da cauda, como já verificamos a distribuição de perda de séries financeiras dos retornos logaritmizados tem uma cauda mais pesada que a da distribuição Normal.

Dado que a série dos retornos logaritmizados é leptocúrtica é necessária uma distribuição com maior probabilidade de massa nas caudas, surgindo a distribuição *T-student*. No entanto, a distribuição *T-student* continua, de longe, a não ser uma fonte de viabilidade para perdas reais.

A distribuição de *Kernel* pertence à classe de métodos não paramétricos, sendo o seu aspeto bastante similar ao da distribuição *T-student*, mas mais leptocúrtica e com caudas mais gordas, ajustando-se melhor à distribuição das rendibilidades do índice PSI20.

Podemos também observar as caudas de cada uma das distribuições em estudo, na Figura 3.8. De forma análoga, é observável que o pior desempenho se encontra na distribuição Normal. Isto significa que a distribuição Normal apresenta alguma fraqueza nas caudas enquanto as outras duas distribuições apresentam uma melhor aderência nas caudas quando se trata de séries financeiras.

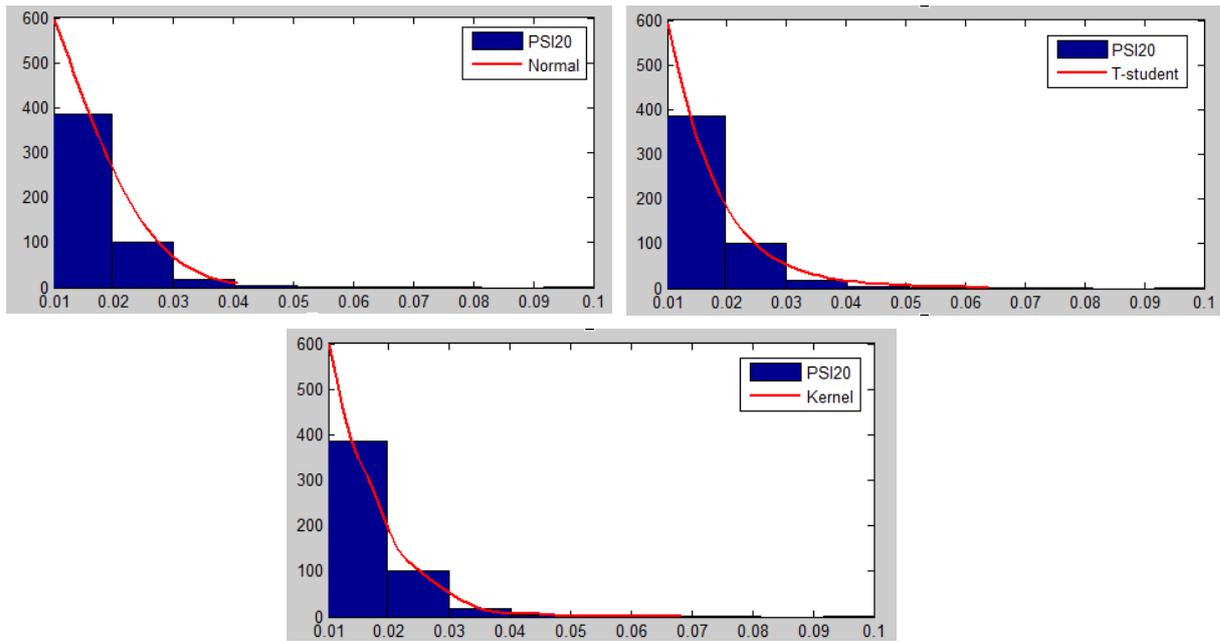


Figura 3.8: Observação das caudas

Segue-se a estimação das funções perda usando quatro modelos distintos e uma janela de 250 dias, resultando assim uma série com 2641 estimações para o VaR e o ES.

Nas figuras 3.9-3.12 apresenta-se, de forma gráfica, os valores estimados para o VaR e para o ES utilizando os modelos paramétricos e não paramétricos com as distribuições Normal, *T-student*, *Kernel* e o modelo *GARCH(1,1)* com distribuição Normal.

Inicialmente, apresenta-se o VaR e o ES estimados segundo a distribuição Normal, sendo que o VaR é avaliado com dois níveis de confiança: 97.5% e 99% e o ES com o nível de confiança de 97.5% (Figura 10).

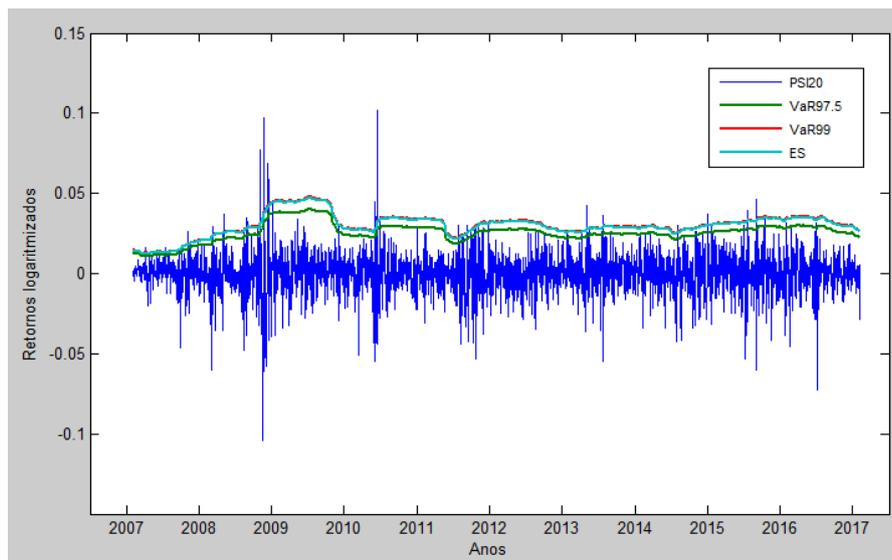


Figura 3.9: Estimação VaR e ES com a distribuição Normal

Na Figura 3.10 e 3.11, é ilustrado o VaR e o ES quando é considerada a distribuição *T-student* e *Kernel* para os mesmos níveis de confiança indicados. Observa-se que as estimativas do risco têm uma variância mais suave e não acompanham os eventos extremos.

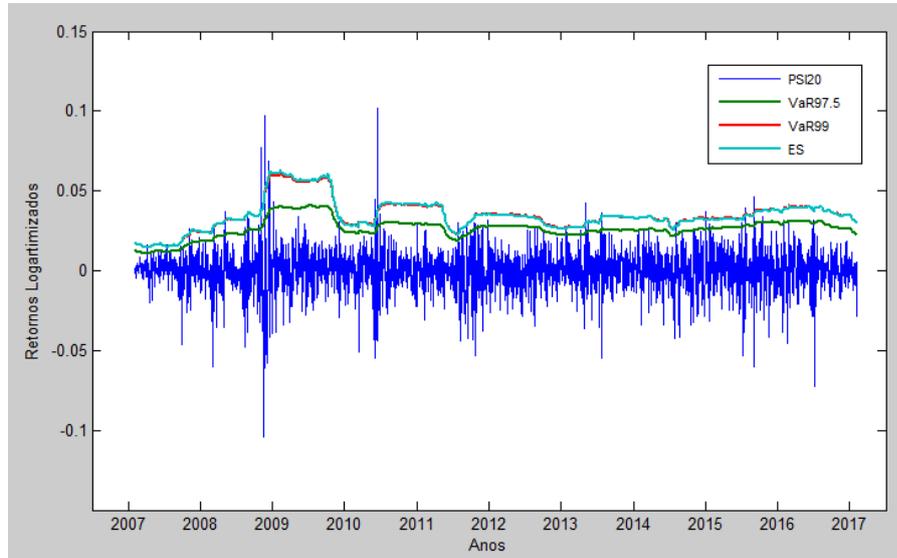


Figura 3.10: Estimação VaR e ES com a distribuição *T-student*

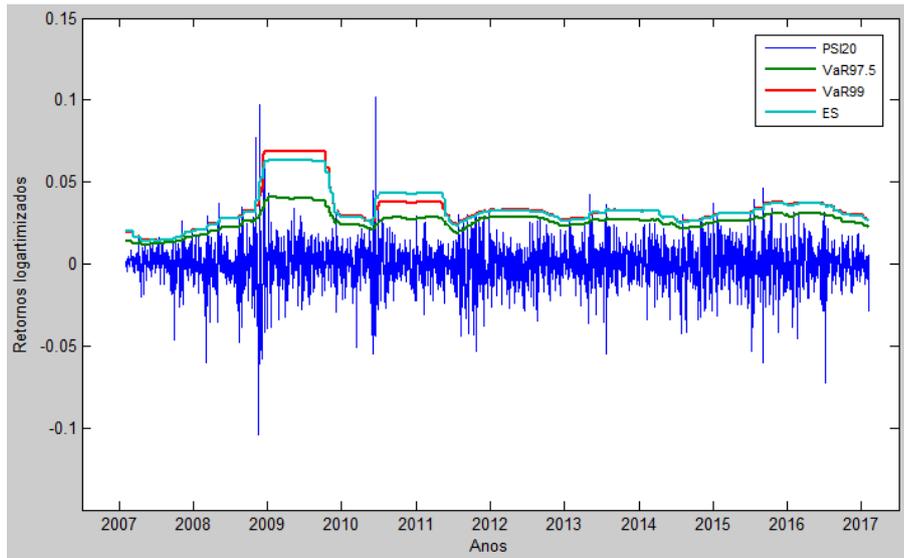


Figura 3.11: Estimação VaR e ES com a distribuição Kernel.

Por fim, na Figura 3.12 temos o VaR e o ES estimados com base no modelo *GARCH(1,1)* com distribuição Normal.

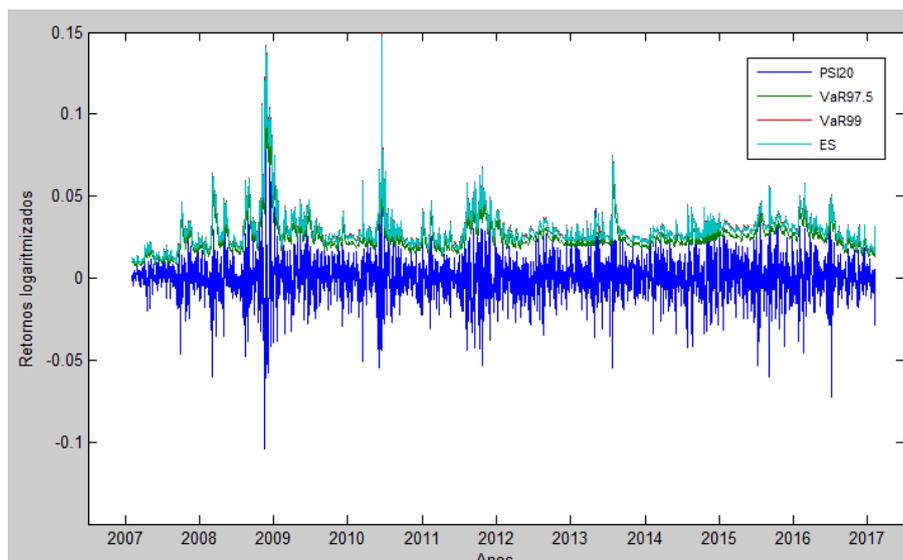


Figura 3.12: Estimação VaR e ES com distribuição GARCH Normal

Observamos que o ES e o VaR com nível de significância de 99%, estão bastante próximos ao longo do tempo. Segundo Roccioletti (2016) o ES foi ajustado para igualar ao VaR devido às caudas pesadas, a mesma situação é evidenciada nos modelos com distribuição *T-student*, *Kernel* e *GARCH* com distribuição Normal.

Podemos observar que o modelo *GARCH* se adapta bem às alterações na volatilidade, uma vez que apresenta um comportamento distinto dos restantes modelos para as distribuições utilizadas. Quanto ao modelo com distribuição Normal é possível verificar que são apresentadas estimativas inferiores às apresentadas pelos modelos com distribuição *T-student* e *Kernel*, aparentando apresentar um baixo nível de risco. É ainda visível a existência de *clusters* de volatilidade (Figura 3.12).

3.2.1 RESULTADOS DO BACKTESTING AO VAR

Nesta secção são analisados os testes de *backtesting* associados aos modelos de estimação do VaR com o objetivo de distinguir o melhor desempenho dos mesmos.

Na Tabela 3.5 apresenta-se o número de exceções observadas para a série do PSI20, a cada um dos níveis de confiança 97.5% e 99%, para o valor do VAR relativamente às distribuições em análise.

Tabela 3.5: Número de exceções esperadas e observadas

Distribuição	Nível de Confiança	Número de observações	Número de exceções esperadas	Número de exceções observadas
Normal	97.5%	2641	66	75
	99%	2641	26	37
<i>T-student</i>	97.5%	2641	66	68
	99%	2641	26	20
<i>GARCH</i> Normal	97.5%	2641	66	62
	99%	2641	26	18

O número de exceções esperadas para um nível de significância de 97.5% é de 66 observações e para um nível de significância de 99% é de, aproximadamente, 26 observações.

É possível observar que o número de exceções esperadas, ao nível de 99%, é inferior apenas da distribuição Normal. Enquanto, ao nível 97.5%, é superior apenas na distribuição *GARCH* Normal.

De forma a verificar se a quantidade de exceções é significativamente grande recorreremos ao teste de Kupiec (Tabela 3.6). Estatisticamente, caso sejam encontradas observações extremistas o modelo é rejeitado.

Tabela 3.6: Resultados do Teste de Kupiec

PSI20	Teste de Kupiec		
	Normal	<i>T-student</i>	<i>GARCH</i> Normal
Distribuição	Normal	<i>T-student</i>	<i>GARCH</i> Normal
	Normal	<i>T-student</i>	<i>GARCH</i> Normal
Estatística de Teste	1.1995	0.0600	0.2568
	3.8139	1.7153	3.0457
<i>P-value</i> (Logical Value)	1	1	1
	1	1	1
Resultado	Rejeita H_0	Rejeita H_0	Rejeita H_0
	Rejeita H_0	Rejeita H_0	Rejeita H_0

O resultado apresentado no teste de Kupiec (1995) é avaliado de acordo com o nível de significância de 5%. Dado os resultados apresentados na Tabela 3.6 e avaliando o valor do *p-value* de cada uma das distribuições é possível verificar quais os modelos que são ou não rejeitados.

É observável que em todos os casos se rejeita a hipótese nula. A rejeição da hipótese nula presente no modelo *GARCH* com distribuição Normal deriva do facto de o número de exceções ser abaixo do valor esperado, ou seja, o nível de risco é superestimado. Com isto, concluímos que nenhum dos modelos apresentados é adequado.

De seguida, recorreremos ao teste de Christoffersen (1998) em que são igualmente avaliados os resultados apresentados na Tabela 3.7.

Tabela 3.7: Resultados do Teste de Christoffersen

PSI20	Teste de Christoffersen		
	Normal	<i>T-student</i>	<i>GARCH</i> Normal
Distribuição	Normal	<i>T-student</i>	<i>GARCH</i> Normal
	Normal	<i>T-student</i>	<i>GARCH</i> Normal
Estatística de Teste	10.4151	9.7645	1.3298
	10.0725	0.3053	0.2471
<i>P-value</i> (Logical Value)	0	0	1
	0	1	1
Resultado	Não se rejeita H_0	Não se rejeita H_0	Rejeita H_0
	Não se rejeita H_0	Rejeita H_0	Rejeita H_0

Realizando a respectiva análise, o modelo *GARCH* com distribuição Normal rejeita-se a hipótese nula tal como no teste de Kupiec (1995). Já o oposto ocorre para os resultados do modelo com distribuição Normal, aparentando ter um melhor desempenho que as restantes. Porém, o modelo com distribuição *T-student* mostra uma melhoria significativa, no sentido em que, para o nível de confiança de 97.5%, não se rejeita a hipótese nula.

Para esta análise é necessário ter em consideração que o *backtesting* é realizado incluindo períodos de crise, como a crise de 2008, dado que todo o mercado é afetado por estes.

Com esta avaliação de resultados, conclui-se que o modelo baseado na distribuição Normal é a que apresenta uma melhor *performance* na previsão do VaR relativo ao índice PSI20, apesar de ser observado graficamente que a distribuição dos dados é leptocúrtica e, tanto o modelo *Kernel*, como o modelo com distribuição *T-student* são os que melhor se ajustam.

3.2.2 RESULTADOS DO BACKTESTING AO ES

Na Tabela 3.8 e 3.9 são apresentados os resultados obtidos na realização dos testes Z_1 e Z_2 , associados às estimativas do ES, sendo que são avaliados através do valor da estatística de teste.

Tabela 3.8: Resultados do teste Z_1

PSI20	Teste Z_1		
	Normal	<i>T-student</i>	<i>GARCH</i> Normal
Distribuição			
Estatística de Teste	0.1398	0.0013	0.0221
Resultado	Rejeita H_0	Não se rejeita H_0	Não se rejeita H_0

Tabela 3.9: Resultados do teste Z_2

PSI20	Teste Z_2		
	Normal	<i>T-student</i>	<i>GARCH</i> Normal
Distribuição			
Estatística de Teste	0.2947	0.0313	-0.0402
Resultado	Rejeita H_0	Não se rejeita H_0	Não se rejeita H_0

Nesta análise não se verificam diferenças nos resultados obtidos entre ambos os testes.

O modelo com distribuição Normal é o único modelo que não é válido em qualquer um dos testes porque rejeita a hipótese nula. Contrariamente, ao modelo Normal, o modelo *GARCH* com distribuição Normal revela um bom desempenho dado que não rejeita a hipótese nula.

De acordo com Roccioletti (2016), a sobrevalorização do risco de défice ocorre quando:

$$P(Z < z^*) < \phi \quad (3.1)$$

onde ϕ representa o nível de significância, sendo nesta análise 5%.

Portanto, o valor negativo presente na Tabela 9 referente à distribuição *GARCH* Normal, pode indicar a existência de uma sobrevalorização do risco de défice.

Contudo, deve-se ponderar ao avaliar o resultado da distribuição *T-student*, visto que esta distribuição passa em ambos os testes, embora com alguma margem de dúvida subjacente, dado ser um valor próximo de zero.

Segundo Roccioletti (2016), como nenhum dos modelos passa o teste de Kupiec (1995) no cálculo do VaR ao nível de confiança de 97.5%, na maioria dos casos é possível considerar que os resultados do teste Z_1 são inválidos e, portanto, deve-se confiar nos resultados apresentados no teste Z_2 . Embora isto seja insignificante no caso apresentado, dado que os resultados são iguais.

Concluimos, graficamente, que foi observado que o modelo com distribuição Normal era o modelo que menos se ajustava aos dados em análise dado a sua fraqueza nas caudas. A série dos retornos logaritmizados sendo leptocúrtica exige uma distribuição com maior probabilidade de massa nas caudas, surgindo daqui o melhor ajuste do modelo com distribuição *T-student* e do modelo *Kernel* pois ambos têm um aspeto similar e caudas mais gordas.

No entanto, no *backtesting* realizado ao VaR pelo teste de Kupiec (1995) nenhum dos modelos é apresentado como válido. No entanto, o teste de Christoffersen (1998) vai contra ao que foi analisado graficamente, pelo facto se ser o modelo com distribuição Normal o que apresenta melhor desempenho, embora ao nível de significância de 97.5%, o modelo com distribuição *T-student* tenha relevado alguma melhoria, sendo que a hipótese nula deixou de ser rejeitada.

Quanto ao *backtesting* realizado ao ES, em ambos os testes Z_1 e Z_2 , foram obtidos os mesmos resultados. Sendo estes resultados, os que vão ao encontro da análise gráfica realizada, uma vez que apenas o modelo com distribuição Normal é que não é válido. Concluindo que o melhor desempenho na previsão de risco remete para o método: *Expected Shortfall* (ES).

CONCLUSÃO

Com os dados relativos ao principal índice bolsista português foi realizada a análise descritiva da série do PSI20 e os testes ADF e PP. Concluindo-se que a série não é estacionária, sendo que existiu a necessidade de estacionarizar a série recorrendo aos seus retornos. O que permitiu que se criasse uma estabilidade no valor médio da série e que a sua variabilidade diminuísse.

De seguida, foi realizado um estudo gráfico relativamente as distribuições Normal, *T-student* e *Kernel* ajustadas ao histograma dos retornos do índice PSI20, de forma a ser possível observar qual a distribuição que melhor se ajusta às rendibilidades do índice PSI20. Obteve-se as distribuições *T-student* e *Kernel*, como um bom ajuste aos retornos. No entanto, a distribuição *Kernel* é mais leptocúrtica e apresenta caudas mais gordas, ajustando-se melhor aos retornos do índice em questão.

Seguindo as rotinas de Roccioletti (2016), realizou-se também o estudo gráfico dos valores estimados para o VaR e o ES utilizando os modelos paramétricos e não paramétricos com as distribuições Normal, *T-student*, *Kernel* e o modelo *GARCH* (1,1) com distribuição Normal. A conclusão é a mesma para todos os gráficos apresentados: o ES e o VaR com nível de significância de 99% encontram-se bastante próximos ao longo do tempo e, Roccioletti (2016) evidencia que o ES foi ajustado para igualar ao VaR devido às caudas pesadas.

Procedeu-se, por fim, ao *backtesting*, aplicando-se o teste de Kupiec (1995) e o teste de Christoffersen (1998) para a avaliação das *performance* dos diferentes modelos de previsão do *Value-at-Risk* e os testes Z_1 e Z_2 para a previsão do *Expected Shortfall*.

De acordo com o teste de Kupiec (1995) não existe nenhum modelo adequado. No entanto, no teste de Christoffersen (1998) o modelo com distribuição Normal é o que apresenta melhor desempenho, embora que, ao nível de significância de 97.5%, o modelo com distribuição *T-student* tenha relevado alguma melhoria, sendo que a hipótese nula deixou de ser rejeitada.

Relativamente aos testes Z_1 e Z_2 , os resultados foram iguais para ambos, onde apenas o modelo com distribuição Normal se apresenta como não sendo válido. Concluindo-se que a melhor *performance* na previsão de risco é apresentada no método: *Expected Shortfall* (ES).

Seria possível melhorar estes resultados e, possivelmente, avaliá-los de forma mais detalhada se se dividisse a amostra em 3 conjuntos, por exemplo: pré-crise, durante a crise e pós-crise. Onde, possivelmente, conseguiríamos ter uma melhor visibilidade sobre como a crise afeta os mercados financeiros, o quanto variavam os resultados devido à crise e o quanto é necessário a existência de métodos de previsão de risco.

Poder-se-ia, também, fazer horizontes temporais distintos, ou seja, formar subconjuntos de 250 dias e avaliar, ano a ano, os resultados obtidos de forma a prevenir perdas nos anos subsequentes. Estas seriam alternativas para uma observação mais minuciosa dos resultados, podendo verificar algumas das causas implícitas na dimensão do risco e até uma explicação mais detalhada dos resultados obtidos.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Acerbi, C. & Szekely, B. (2014). *Backtesting expected shortfall*. Risk Magazine.
- [2] Albuquerque, R. C. (2015). *Modelagem em séries temporais aplicadas a dados climatológicos no sertão paraibano*. Trabalho de conclusão de curso. Disponível em: <http://dspace.bc.uepb.edu.br/jspui/handle/123456789/8188>.
- [3] Amaral, C. (2015). *Análise de modelos de previsão de Value-at-Risk aplicados ao principal índice de ações de mercado português*. Tese de mestrado. Disponível em: <http://repositorio.ul.pt/handle/10451/20656>.
- [4] ANACOM (2008). *Anuário do Sector das Comunicações em Portugal*. Lisboa: ANACOM.
- [5] Araújo, M. (2010). *Análise de clusters e volatilidade dos índices de ações*. Tese de mestrado. Disponível em: <http://repositorio.ipl.pt/bitstream/10400.21/2613/1/disserta%C3%A7%C3%A3o%20volatilidade.pdf>.
- [6] Associação BM&FBOVESPA (2012). *Introdução e gestão de risco*. Disponível em: <http://lojavirtual.bmf.com.br/LojaIE/portal/pages/pdf/Apostila-PQO-Capitulo8.pdf>.
- [7] Augusto, L., Santos J. & Vieira, E. (2016). Cobertura do risco financeiro nas empresas do PSI20. *Estudos do ISCA*, Série IV, 13. Disponível em: <http://revistas.ua.pt/index.php/estudosdoisca/article/view/3767/3465>.
- [8] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 309-328.
- [9] Brandolini, D. & Colucci, S. (2011). Backtesting Value-at-Risk: A comparison between Filtered Bootstrap and Historical Simulation. *Journal of Risk Model Validation*. 6, 4, 3-16.
- [10] Carvalho, H. (2013). *Evolução dos spreads de risco das empresas não financeiras do índice PSI20*. Tese de mestrado. Disponível em: <http://www.repository.utl.pt/handle/10400.5/6345>.
- [11] Christoffersen, P. (1998). Evaluating Interval Forecasts. *International Economic Review*, 39, 841-862.
- [12] CMVM (agosto, 2017). *Relatório de Indicadores mensais do mercado de capitais português*. Lisboa: CMVM. Disponível em: <http://www.cmvm.pt/pt/Estatisticas/EstatisticasPeriodicas/IndicadoresMensaisDoMercadoDeCapitaisPortugues/Pages/Agosto2017.aspx?shpage=IndicadoresMensaisDoMercadoDeCapitaisPortugues>.
- [13] Costa, S. (2014). *Impacto da crise na performance económica-financeira das empresas*. Tese de mestrado. Disponível em: <https://comum.rcaap.pt/bitstream/10400.26/7946/1/Disserta%C3%A7%C3%A3o%20Sandra%20PDF%20corrigido.pdf>.

- [14] Danielsson, J. (2011). *Financial Risk Forecasting*. Wiley Finance.
- [15] Dickey, D. A. & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427-431.
- [16] Duda, M. & Schmidt, H. (2009). *Evaluation of various approaches to Value-at-Risk*. Tese de mestrado. Disponível em:
<http://lup.lub.lu.se/luur/download?func=downloadFile&recordOid=1436923&fileOid=1646971>.
- [17] Emmer, S., Kratz, M. & Tasche D. (2015). What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, 18(2), 31–60.
- [18] Euronext. (2003). *Regras de calculo dos índices PSI*. Disponível em:
<http://www1.eeg.uminho.pt/economia/caac/pagina%20pessoal/Disciplinas/Disciplinas%2005/bancaria/psi20.pdf>.
- [19] Euronext. (2014). *Livro de regras do índice PSI20*. Disponível em:
https://indices.euronext.com/sites/indices.nyx.com/files/regras_psi_20_-_traducao_-_14-03.pdf.
- [20] Francisco, P. (2015). *Impacto das Cimeiras Europeias no PSI20*. Tese de mestrado. Disponível em: <https://comum.rcaap.pt/handle/10400.26/14255>.
- [21] Gschöpf, P. (2014). *Measuring risk with expectile based expected shortfall estimates*. Tese de mestrado. Disponível em: <https://edoc.hu-berlin.de/bitstream/handle/18452/14867/gschoeff.pdf?sequence=1&isAllowed=y>.
- [22] Gustafsson, M. & Lundberg, C. (2009). *Empirical evaluation of Value-at-Risk*. Tese de mestrado. Disponível em:
https://gupea.uu.se/bitstream/2077/19299/1/gupea_2077_19299_1.pdf.
- [23] Haas, M. (2001). *New methods in backtesting*. *Financial Engineering Research center caesar*. Disponível em: <https://www.ime.usp.br/~rvicente/risco/haas.pdf>.
- [24] Harmantzis, C., Linyan, M. & Chien Y. (2005). *Empirical Study of Value-at-Risk and Expected Shortfall*. Disponível em:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=788624
- [25] Hendricks, D. (1996). Evaluation of Value-at-Risk models using historical data. *FRBNY Economic Policy Review*.
- [26] Holá, A. (2012). *Mathematical Models of Value-at-Risk*. Tese de licenciatura.
<https://otik.uk.zcu.cz/handle/11025/7137>
- [27] Izmaylov, B. (2014). *Value-at-Risk: Strengths, Caveats and Considerations for risk managers and regulators*. Tese de mestrado. Disponível em:
http://pure.au.dk/portal/files/71175799/Master_Thesis_Bogdan_Izmaylov.pdf

- [28] Karlsson, J. & Zakrisson, H. (2016). *Practical estimation of Value-at-Risk and Expected Shortfall: are complex methods really necessary?*. Tese de licenciatura. Disponível em: <https://lup.lub.lu.se/student-papers/search/publication/8875765>.
- [29] Kourouma, L., Dupre, D., Sanfilippo, G. & Taramasc, O. (2010). *Extreme Value-at-Risk and Expected Shortfall during financial crisis*. Disponível em: <https://halshs.archives-ouvertes.fr/halshs-00658495/document>.
- [30] Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*. New York, 3. Winter.
- [31] Martins, V. (2015). *Expected Shortfall: Algumas abordagens de implementação em séries financeiras*. Tese de mestrado. Disponível em: http://repositorio.ul.pt/bitstream/10451/23025/1/ulfc117345_tm_V%C3%ADtor_Manuel_Martins.pdf.
- [32] Matias, F. (2012). *Impacto da gestão do risco nas instituições financeiras*. Tese de mestrado. Disponível em: <https://comum.rcaap.pt/bitstream/10400.26/3995/1/Disserta%C3%A7%C3%A3o%20-%20Impacto%20da%20gest%C3%A3o%20do%20risco%20nas%20instui%C3%A7%C3%B5es%20financeiras..pdf>.
- [33] Models with Heavy Tails. *The Journal of Risk Finance*, 7, 2, 117-135.
- [34] Oliveira, S. & Andrade, M. (s.d.). *Abordagem Bayesiana para modelos estocásticos com heteroscedasticidade para os retornos Ibovespa*. Disponível em: <http://www2.ime.unicamp.br/sinape/sites/default/files/Artigo%20-%20Oliveira%20&%20Andrade.pdf>.
- [35] Osmundsen, K. (2016). *Credit risk modeling with expected shortfall – a simulation based portfolio analysis*. Tese de mestrado. Disponível em: https://brage.bibsys.no/xmlui/bitstream/handle/11250/2407353/14708_FULLTEXT.pdf?sequence=1.
- [36] Pavol Jurca, Lecture Notes – Chapter 2, <http://www.iam.fmph.uniba.sk/institute/jurca/qrm/Chapter2.pdf>.
- [37] Phillips, P. & Perron, P. (1988). Testing for unit root in time series *regression*. *Biométrica*, 75(2), 335-346.
- [38] Roccioletti, S. (2016). *Backtesting Value-at-Risk and Expected Shortfall*. Wiesbaden: Springer Gable.