

**HUGO KENJI DE OLIVEIRA YAGIHARA**

**Análise das Métricas de Value at Risk e Expected Shortfall para Gerenciamento  
de Riscos**

São Paulo

2024

**HUGO KENJI DE OLIVEIRA YAGIHARA**

**Análise das Métricas de Value at Risk e Expected Shortfall para Gerenciamento de Riscos**

Monografia apresentada à Escola  
Politécnica da Universidade de São Paulo  
no MBA de Engenharia Financeira

Orientador: Prof. Dr. Danilo Zucolli  
Figueiredo

São Paulo

2024

**HUGO KENJI DE OLIVEIRA YAGIHARA**

**Análise das Métricas de Value at Risk e Expected Shortfall para Gerenciamento de Riscos**

Monografia apresentada à Escola  
Politécnica da Universidade de São Paulo  
no MBA de Engenharia Financeira

Orientador: Prof. Dr. Danilo Zucolli  
Figueiredo.

São Paulo

2024

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

#### Catálogo-na-publicação

Yagihara, Hugo

Análise das Métricas de Value at Risk e Expected Shortfall para Gerenciamento de Riscos / H. Yagihara – São Paulo, 2024.  
39 p.

Monografia (MBA em Engenharia Financeira) - Escola Politécnica da Universidade de São Paulo. PECE – Programa de Educação Continuada em Engenharia.

1.VaR 2.ES 3.Kupiec 4.GBI 5.Backtesting I.Universidade de São Paulo. Escola Politécnica. PECE – Programa de Educação Continuada em Engenharia II.t

## RESUMO

O presente estudo tem como objetivo avaliar a eficiência dos modelos de *Value at Risk* (*VaR*) e *Expected Shortfall* (*ES*) para gerenciamento de riscos. Para essa análise, foram utilizados o Teste de Basileia e o Teste de Kupiec como ferramentas de *backtesting* para o *VaR*, enquanto para o *ES*, adotou-se a recente abordagem do *Generalized Breach Indicator* (*GBI*), desenvolvida por Costanzino e Curran (2018). Este estudo, até o momento, é o primeiro a testar essa abordagem de *backtesting* para o *ES* considerando ativos emitidos no Brasil. Utilizando dados do índice Ibovespa dos últimos seis anos, e analisando para um intervalo de tempo de 250 dias úteis e um nível de confiança de 97,5%, os resultados indicaram que os modelos testados demonstram resultados satisfatórios em períodos de normalidade, mas apresentam vulnerabilidades em períodos de alta volatilidade.

**Palavras-Chave:** *VaR*. *ES*. Kupiec. *GBI*. *Breach*. *Backtesting*.

## **ABSTRACT**

The present study aims to assess the efficiency of Value at Risk (VaR) and Expected Shortfall (ES) models for risk management. For this analysis, the Basel Test and the Kupiec Test were used as backtesting tools for VaR, while for ES, the recent approach of the Generalized Breach Indicator (GBI), developed by Costanzino and Curran (2018), was adopted. This study, up to this point, is the first to test this backtesting approach for ES considering assets issued in Brazil. Using data from the Ibovespa index over the last six years and analyzing for a time interval of 250 business days and a confidence level of 97.5%, the results indicated that the tested models demonstrate satisfactory outcomes in normal periods but show vulnerabilities in periods of high volatility.

**Keywords:** VaR. ES. Kupiec. GBI. Breach. Backtesting.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>6</b>
<b>2</b>	<b>MÉTRICAS DE RISCO</b> .....	<b>10</b>
<b>2.1</b>	<b>VALUE AT RISK</b> .....	<b>10</b>
<b>2.1.1</b>	<b>Simulação Histórica</b> .....	<b>12</b>
<b>2.1.2</b>	<b>Paramétrico Normal</b> .....	<b>13</b>
<b>2.1.3</b>	<b>Cornish-Fisher</b> .....	<b>14</b>
<b>2.1.4</b>	<b>EWMA</b> .....	<b>15</b>
<b>2.2</b>	<b>EXPECTED SHORTFALL</b> .....	<b>16</b>
<b>3</b>	<b>MÉTODOS DE VALIDAÇÃO DE MODELOS</b> .....	<b>18</b>
<b>3.1</b>	<b>BACKTESTING</b> .....	<b>18</b>
<b>3.1.1</b>	<b>Teste de Kupiec</b> .....	<b>18</b>
<b>3.1.2</b>	<b>Teste de Basileia</b> .....	<b>19</b>
<b>3.1.3</b>	<b>Generalized Breach Indicator</b> .....	<b>21</b>
<b>4</b>	<b>METODOLOGIA</b> .....	<b>24</b>
<b>4.1</b>	<b>DADOS</b> .....	<b>24</b>
<b>4.2</b>	<b>ANÁLISE E MODELAGEM</b> .....	<b>25</b>
<b>5</b>	<b>RESULTADOS</b> .....	<b>28</b>
<b>6</b>	<b>CONCLUSÃO</b> .....	<b>34</b>
	<b>REFERÊNCIAS</b> .....	<b>36</b>

## 1 INTRODUÇÃO

O risco, do ponto de vista financeiro, pode ser definido como a possibilidade de perda ou de variações indesejadas sobre uma decisão de investimento ou uma carteira de ativos. O mesmo é uma característica inerente às operações financeiras e está presente em diferentes contextos. Markowitz (1952) definiu risco como a variabilidade dos retornos de um investimento, argumentando que os investidores podem otimizar seu portfólio maximizando o retorno para um determinado nível de risco ou minimizando o risco para um dado nível de retorno. De acordo com Jorion (2007), o risco pode ser definido como a volatilidade de resultados inesperados que podem representar o valor dos ativos, do patrimônio líquido ou dos lucros. Em uma definição mais recente, Pinho (2019), reitera a definição dada pelo autor anterior, abordando especificamente o risco financeiro, definindo-o como as perdas resultantes das oscilações financeiras que afetam o valor do ativo e passivo das empresas. Portanto, é de extrema importância fazer o gerenciamento desse risco, pois, além de permitir uma melhor alocação de recursos e otimização de portfólios, permite que a instituição financeira esteja preparada para eventuais perdas, podendo assim minimizar seu impacto e garantir a continuidade do negócio, principalmente em períodos de alta volatilidade.

Nas últimas décadas, principalmente em períodos de alta volatilidade, ocorreram alguns casos de crises ou colapsos financeiros que ocasionaram grande desvalorização ou falência de instituições e empresas, que em alguns casos não estavam preparadas ou tratando o risco de forma adequada. Como a Crise Financeira Asiática de 1997, ou Contágio Asiático de 1997, que impactou grande parte dos países asiáticos, em especial os tigres asiáticos, devido a uma desvalorização da moeda fruto de uma política cambial, impactando pouco tempo depois, a Bolsa de Valores de Hong Kong, ocasionando posteriormente quedas acentuadas no valor das moedas, ações e títulos desses países. Outro exemplo famoso é o caso do fundo Long-Term Capital Management (LTCM), gerenciado por alguns dos *Traders* mais famosos de Wall Street (Robert Merton e Myron Scholes), que em 1998 entrou em colapso devido à alavancagem excessiva de suas posições combinadas com a Crise Financeira Russa e com a Crise Financeira Asiática de 1997 citada anteriormente.

Destaca-se, dentre as diversas crises da história, a Crise Financeira Global de 2008 que teve, do ponto de vista regulatório, desdobramentos importantes que

contribuíram com o aperfeiçoamento das ferramentas de gestão de riscos, visando evitar eventos similares no futuro. Essa crise teve origem no mercado imobiliário dos EUA, devido à concessão de empréstimos hipotecários de alto risco a clientes com baixo *score* de crédito. O colapso do banco de investimento Lehman Brothers foi um dos ápices que desencadeou a crise, provocando pânico e levando a população a correr aos bancos para resgatar seus investimentos. No mesmo mês, os bancos Morgan Stanley e Goldman Sachs solicitaram crédito emergencial ao Federal Reserve (FED).

Mais recentemente, houve a Pandemia do Coronavírus (COVID-19), que rapidamente gerou altas volatilidades nos mercados globais no ano de 2020, devido aos *lockdowns*, fechamento de fronteiras, incerteza econômica e o escasso conhecimento sobre o vírus, forçando aos Bancos Centrais em todo o mundo a reduzirem as taxas de juros. Nesse período, muitas empresas aumentaram seu endividamento e sofreram com a redução de sua receita. Conforme dados da Bolsa de Valores de São Paulo (B3), durante o mês de março de 2020, a Bolsa acionou o mecanismo de segurança *Circuit Breaker*, seis vezes, sendo que todas as quedas do índice foram superiores a 10%, e no dia 12, a paralização foi acionada duas vezes. Essa crise, devido ao seu impacto, duração e sequelas geradas, pode ser considerada umas das mais impactantes da história, e, junto com seus impactos, trouxe novamente às pautas de discussões a imprescindível necessidade de haver controles e métricas para um eficiente gerenciamento de riscos, o que reacendeu o debate sobre duas das métricas mais famosas e utilizadas no mercado, o *Value at Risk (VaR)* e o *Expected Shortfall (ES)*.

Nesse contexto regulatório, a principal organização responsável pela supervisão bancária, a fim de garantir solidez e o controle dos riscos, é o Comitê de Basileia para Supervisão Bancária (BCBS, sigla de *Basel Committee on Banking Supervision*). Em 1996, em seu documento publicado *Amendment to the capital accord to incorporate market risks*, introduziu as regras de cálculo de capital para risco de mercado, permitindo aos bancos utilizarem modelos internos para o cálculo de capital regulatório. Isso tinha como intuito assegurar uma quantia mínima que os bancos deveriam reter para cobrir possíveis riscos. Esses modelos deveriam ser baseados na famosa métrica de *VaR*, gerando um grande avanço no seu desenvolvimento e popularização. De acordo com Scaillet (2004), a principal utilização nos bancos é garantir que o risco incorrido pela instituição financeira permaneça

dentro do seu apetite ao risco declarado e que a avaliação desse risco ocorra em tempo hábil para permitir uma ação corretiva imediata, se necessária.

Após a Crise Financeira Global de 2008, o mercado viu a necessidade de uma melhoria das práticas já existentes, principalmente em eventos extremos, também conhecidos como eventos de cauda. Devido a isso, o BCBS agilizou o processo de revisão do acordo de Basileia, após entender algumas falhas e incapacidades das métricas de *VaR*. Em 2013, o BCBS apresentou em um de seus documentos, um conjunto de propostas para um novo requerimento de capital mínimo exigido aos bancos, denominado *Fundamental Review of the Trading Book: A revised market risk framework* (2013). Entre diversas medidas, propõe a substituição do *VaR* por uma métrica de risco coerente, o *ES*. Essa mudança de métrica já tinha sido sugerida anteriormente por alguns autores, como Acerbi e Tasche (2002) que colocaram o *ES* com alternativa ao *VaR*, argumentando que o *ES* mensura o risco quando o *VaR* falha, ou seja, quando o mesmo é superado, além de ser uma medida de fácil entendimento e simples aplicação. Além disso, já havia uma pressão tanto por parte dos gestores de risco como dos acadêmicos especialistas no assunto para que houvesse essa substituição do *VaR* por outra medida que resolvesse suas deficiências (COSTANZINO e CURRAN, 2018).

No entanto, apenas o cálculo dessas métricas não era suficiente para garantir confiança e robustez ao processo de gerenciamento de riscos. Era necessário algum método de testar a acurácia e validade desses modelos através de procedimentos chamados de *backtesting*, que têm como objetivo principal testar a eficiência dos modelos de previsões. No Brasil, o Banco Central do Brasil (BACEN) é a autarquia responsável por supervisionar e regulamentar as instituições financeiras nacionais e, por meio da circular nº 3.646, de 4 de março de 2013, artigo 12, estabelece que devem ser realizados testes de aderência (*Backtests*) de forma a assegurar avaliações consistentes da aderência do modelo utilizado. Entre algumas exigências mínimas que estes testes devem atender, estão que o período de manutenção deve ser de um dia, a periodicidade mínima deve ser mensal e devem ser utilizados diversos intervalos de confiança. Portanto, a importância de utilizar modelos eficientes comprovados através de testes de aderência não é apenas uma necessidade da instituição financeira para garantir maiores retornos ou minimizar suas perdas, mas também é uma exigência regulatória a qual é obrigada a seguir.

Nesse contexto, o objetivo desse trabalho é verificar a eficiência de alguns dos principais modelos de *VaR* e *ES* conhecidos na literatura e utilizados no mercado. Para o *VaR*, serão testados os modelos de Simulação Histórica, Paramétrico Normal (Gaussiano), Expansão de *Cornish Fisher* e *EWMA*. Para o *backtesting* desses modelos serão utilizados o Teste de Kupiec e o Teste de Basileia. Já para o *ES*, serão utilizados os mesmos modelos citados anteriormente para o *VaR*. Porém, o *backtesting* será feito pela abordagem apresentada nos estudos de Costanzino e Curran (2018), denominada *Generalized Breach Indicator (GBI)*. A série de retornos utilizada será a do Índice Ibovespa, e os modelos serão calculados para o horizonte de tempo de 250 dias úteis e nível de confiança de 97,5%.

A escolha do Índice Ibovespa se dá pela alta volatilidade apresentada em seu histórico de pontos. Nos últimos anos, o índice despencou do patamar de 118.000 pontos em janeiro de 2020 para o patamar de 66.000 pontos em março do mesmo ano. Após o auge da pandemia, o índice se recuperou e, em junho de 2021, chegou a superar os 130.000 pontos, voltando a cair em julho de 2022 para o patamar de 98.000 pontos e novamente se recuperar e fechar o ano de 2023 com mais de 134.000 pontos, o que o torna propício para esse estudo.

## 2 MÉTRICAS DE RISCO

### 2.1 VALUE AT RISK

A abordagem do *VaR* tem suas origens na década de 1920, devido a requisitos de capital impostos pela Bolsa de Valores de Nova Iorque (HOLTON, 2002). No entanto, foi somente décadas depois que Leavens (1945) apresentou um exemplo quantitativo de valor em risco, analisando uma carteira de ativos, sendo essa considerada a primeira medida de *VaR* publicada na história. Essa medida só veio a se popularizar por volta de 1994, após o banco J.P. Morgan disponibilizar sua metodologia para mensuração de risco de mercado. Além de disponibilizar todos os métodos, o banco em parceria com a agência de notícias Reuters, disponibilizou aos participantes do mercado financeiro todo o instrumental para o cálculo, denominado *RiskMetrics*. Ao divulgar suas metodologias ao mercado, o J.P. Morgan deu uma contribuição crucial para o gerenciamento de risco de mercado, sendo o *VaR* adotado por diversas instituições, inclusive para o cálculo de capital regulatório, tendo uma metodologia bem similar à adotada posteriormente pelo BCBS em 1996.

O *VaR* pode ser entendido como uma medida estatística de risco que quantifica a perda máxima esperada de um investimento ou carteira, dado um nível de significância em determinado horizonte de tempo. Segundo Linsmeier e Pearson (1996), de forma específica, o valor em risco é uma medida de perdas devido a movimentos “normais” do mercado. Devido a isso, o *VaR* tem a capacidade de expressar de forma simples, em um único número, a perda financeira esperada. Pode ser utilizado também para estabelecer limites de exposição ao risco que a instituição está exposta, permitindo uma análise mais segura, visto que as decisões devem ser tomadas através da relação risco e retorno. Jorion (2007) aborda que parte da popularização e adoção do modelo deve-se à sua facilidade de expressar e calcular seus resultados, citando que mesmo em uma carteira composta por diversos fatores de riscos, o *VaR* consegue fornecer um único valor ao considerar a correlação de seus ativos. Além disso, outra vantagem é que o *VaR* é um conceito amplamente difundido e adotado por diversos órgãos reguladores de todo mundo, dado que o mesmo apresenta resultados razoavelmente satisfatórios como modelo estatístico de previsão de perdas para carteiras lineares.

O  $VaR$  pode ser calculado através de duas abordagens principais: a forma paramétrica e a forma não paramétrica. Na abordagem paramétrica, pressupõe-se que os retornos seguem alguma distribuição específica, geralmente a distribuição normal (Gaussiana) ou, em alguns casos, a distribuição  $t$  de Student. Na forma não paramétrica, não há nenhuma suposição sobre a distribuição dos retornos, baseando apenas nos dados históricos.

O  $VaR$  vem sendo bastante criticado após sua popularização. Uma das críticas refere-se às premissas do cálculo, pois depende de algumas suposições, sejam elas sobre a distribuição dos retornos ou a estabilidade dos mercados, que podem não condizer com a realidade. Conforme Caritá et al.(2015), parte das principais metodologias são calculadas com dados históricos para tentar projetar dados futuros, e em situações em que o futuro tem pouca correlação com o passado, as previsões podem não ser eficientes. Nesse sentido, Pérignon e Smith (2010), utilizando uma amostra de dados de bancos internacionais de 2005, constataram que aproximadamente 73% dos bancos utilizam o método de simulação de histórica para apuração do  $VaR$ , e uma grande desvantagem desse método é que o mesmo é sensível ao intervalo de tempo utilizado e à presença de outliers. Além disso, tradicionalmente, o cálculo do  $VaR$  é realizado com informações de fechamento do dia anterior, desconsiderando posições *day trade*. Em casos de crises financeiras, onde valores extremos tendem aparecer com maior frequência, as críticas se baseiam no fato do  $VaR$  desconsiderar e não fornecer nenhuma informação além do nível de confiança.

Além das críticas citadas anteriormente, umas das mais discutidas na literatura é de que o  $VaR$  não é uma medida de risco coerente. Artzner et al. (1999) estabeleceram os conceitos de medidas coerentes de riscos, definindo um conjunto de axiomas que um modelo de risco deveria seguir para serem considerados coerentes.

Definindo  $\rho$  como a medida de risco,  $X$  e  $Y$  sendo variáveis aleatórias que representam uma carteira sujeita a eventuais perdas e  $G$  o conjunto de todas as possíveis carteiras sujeitas a eventuais perdas, os quatro axiomas apresentados pelos autores foram:

- Monotonicidade: Se em qualquer cenário a carteira  $X$  tem menor perda que a carteira  $Y$ , então a carteira  $X$  deve ter menos risco que a carteira  $Y$ .

$$\forall X e Y \in G / X \leq Y, \text{ temos que } \rho(Y) \leq \rho(X)$$

- Invariância de translação: Sempre que se adicionar um ganho certo à carteira, o risco da mesma deve diminuir na mesma proporção.

$$\forall X \in G \text{ e } \forall \alpha \in R, \text{ temos que } \rho(X + \alpha) = \rho(X) - \alpha$$

- Homogeneidade positiva: O risco deve aumentar proporcionalmente ao tamanho da posição da carteira. Assim, caso o tamanho da posição seja multiplicado por um escalar, o risco também deverá ser multiplicado por esse escalar.

$$\forall \lambda \geq 0 \text{ e } X \in G, \text{ temos que } \rho(\lambda X) = \lambda \rho(X)$$

- Subaditividade: A medida de risco de uma carteira composta por  $n$  ativos deve ser menor que a soma das medidas de risco individuais dos ativos que compõem a carteira, ou seja, a diversificação de uma carteira não pode aumentar o risco.

$$\forall X_1 \text{ e } X_2 \in G, \text{ temos que } \rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$$

De acordo com os autores, desses quatro axiomas, o *VaR* não satisfaz o da subaditividade, sendo que em certos casos, a diversificação pode aumentar o risco, ao invés de diminuir, dessa forma prejudicando uma alocação eficiente dos recursos, visto que será exigida de forma desnecessária uma parcela maior de capital mínimo regulatório. Segundo os autores, uma solução possível para esse problema seria a adoção de um modelo que cumprisse todas essas propriedades, sendo que o modelo conhecido como *Expected Shortfall* foi o sugerido pelos mesmos para substituir o *VaR* como medida de risco coerente. Sugestão essa também apoiada por outros autores, como Yamai e Yoshihara (2002) e Acerbi e Tasche (2002).

### 2.1.1 Simulação Histórica

O *VaR* calculado por Simulação Histórica é um método que utiliza dados passados para estimar dados futuros. Nesse método, não há nenhuma suposição sobre a distribuição dos retornos e pressupõem-se que os retornos futuros serão repetições dos retornos passados. O cálculo é feito a partir de uma janela de retornos

ordenados de forma crescente, a partir de um percentil desejado. Conforme definido por Harmantzis, Miao e Chien (2006), utilizando dados de resultados estatísticos sobre processos empíricos de Van Der Vaart (1998), que dispõem sobre o comportamento da distribuição de um estimador à medida que o tamanho da amostra aumenta indefinidamente (normalidade assintótica), matematicamente, o *VaR* Histórico pode ser definido como:

$$F_n^{-1}(\alpha) = X_{n(i)}, \alpha \in \left(\frac{i-1}{n}, \frac{i}{n}\right) \quad (1)$$

Onde  $X_{n(1)} \leq X_{n(2)} \leq \dots \leq X_{n(n)}$  são os retornos ordenados  $i$  da carteira de investimento cujo risco está sendo medido,  $\alpha$  é o nível de confiança e  $F_n^{-1}$  é a função quantil empírica no tempo  $n$  (tamanho da amostra).

Logo, o *VaR* Histórico é dado por:

$$VaR_n(\alpha) = F_n^{-1}(\alpha) \quad (2)$$

Entre as vantagens desse método estão a sua simplicidade de compreensão e implementação. Além disso, devido ao uso de dados históricos, pode reproduzir resultados mais precisos em eventos de cauda. Em contrapartida, por se tratar de um método não paramétrico, é bastante sensível ao intervalo de tempo utilizado, à presença de outliers e à premissa de correlação com o passado, conforme já citado anteriormente.

### 2.1.2 Paramétrico Normal

O *VaR* Paramétrico Normal, também conhecido como *VaR* Gaussiano, é um modelo que assume a premissa que os retornos dos ativos seguem uma distribuição normal padrão. Segundo Harmantzis, Miao e Chien (2006), dada uma amostra  $X_i$ ,  $i = 1, \dots, n$ , independente e identicamente distribuída com  $N(\mu, \sigma^2)$ , sendo  $\mu$  e  $\sigma$  desconhecidos (geralmente assumindo  $\mu = 0$ ) e  $\alpha$  o nível de confiança, o *VaR* é definido como  $\mu + z_\alpha \sigma$ , onde  $z_\alpha$  é tal que  $P(Z > z_\alpha) = \alpha$  com  $Z \sim N(0, 1)$ . Os autores ainda estimaram o desvio padrão  $\sigma$  com a seguinte fórmula:

$$\sigma_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - M)^2, \quad \text{com } M := \frac{1}{n} \sum_{i=1}^n X_i \quad (3)$$

Portanto, assumindo  $\mu = 0$ , o *VaR* Paramétrico Normal pode ser definido como:

$$VaR_n(\alpha) = z_\alpha \sigma_n \quad (4)$$

Esse método, além de ser fácil implementação, apresenta a vantagem de possibilitar a inclusão de outras variáveis estatísticas no cálculo e também permitir alterar a forma de estimar seus parâmetros, dando origem a outros métodos conhecidos. A desvantagem desse método é suposição de normalidade dos retornos, o que pode ser uma simplificação incoerente, principalmente em situações de alta volatilidade.

### 2.1.3 Cornish-Fisher

O *VaR* calculado a partir da expansão de *Cornish-Fisher* é uma variação do modelo tradicional de *VaR* Paramétrico Normal que estima quantis de distribuições diferentes da distribuição normal em função de quantis da Normal Padrão levando em consideração a assimetria e a curtose da distribuição dos retornos para adequar a estimativa do *VaR*.

A aproximação de quarta ordem da expansão de *Cornish-Fisher*  $\tilde{x}_\alpha$  para um dado quantil  $\alpha$  de uma distribuição empírica com média zero e variância igual a 1 é:

$$\tilde{x}_\alpha \approx z_\alpha + \frac{S}{6}(z_\alpha^2 - 1) + \frac{K}{24}(z_\alpha^3 - 3z_\alpha) - \frac{S^2}{36}(2z_\alpha^3 - 5z_\alpha) \quad (5)$$

Onde  $z_\alpha$  é o valor crítico associado ao nível de confiança  $\alpha$  da distribuição normal padrão,  $S$  é o coeficiente de assimetria e  $K$  o coeficiente de curtose.

Definindo  $z_{CF} = \tilde{x}_\alpha$  como o valor crítico de *Cornish-Fisher* associado ao nível de confiança  $\alpha$  da distribuição normal padrão, pode-se definir o *VaR* por esse modelo como:

$$VaR_n(\alpha) = z_{CF} \sigma_n \quad (6)$$

Esse modelo tem a capacidade de adequar a distribuição e corrigir a subestimação ou superestimação do  $VaR$ , ajustando a assimetria e curtose dos retornos. Isso o torna um modelo coerente para capturar eventos de cauda pesada. Devido à implementação desses coeficientes, esse modelo apresenta maior complexidade em comparação com os modelos anteriores e ainda é sensível às estimativas dos coeficientes.

#### 2.1.4 EWMA

O modelo *Exponentially Weighted Moving Average (EWMA)*, desenvolvido pelo banco J.P. Morgan e publicado no documento *RiskMetrics*, é uma técnica que incorpora de forma mais rápida os choques do mercado, através da aplicação de um peso maior aos retornos mais recentes, atribuindo pesos que decrescem exponencialmente conforme os retornos vão ficando mais antigos. Dessa forma, garante que os retornos mais recentes tenham maior influência no cálculo da volatilidade do que as observações mais antigas.

Considerando os retornos ordenados de forma decrescente em relação ao tempo, a fórmula utilizada para estimação da volatilidade é dada por:

$$\sigma_n = \sqrt{\frac{\sum_{i=0}^{n-1} \lambda^i (r_t - \bar{r})^2}{\sum_{i=0}^{n-1} \lambda^i}}, \quad (0 < \lambda \leq 1) \quad (7)$$

Onde  $r_i$  é o retorno,  $\bar{r}$  é a média dos retornos,  $\lambda$  é o fator de decaimento e  $n$  é a quantidade de retornos.

O parâmetro lambda que foi adicionado ao cálculo da volatilidade, determina a velocidade com que os retornos passados perdem sua relevância na estimação, sendo que quanto menor for esse parâmetro, maior será a velocidade de decaimento. O *RiskMetrics* apresenta um procedimento para estimar um lambda ótimo, porém, recomenda o uso do fator de decaimento de 0,94 para operações *Trading* e 0,97 para operações *Banking*.

Logo, o  $VaR$  calculado com a volatilidade ponderada *EWMA* é dado por:

$$VaR_n(\alpha) = z_\alpha \sigma_n \quad (8)$$

Onde  $z_\alpha$  é o valor crítico associado ao nível de confiança  $\alpha$  da distribuição normal padrão e  $\sigma_n$  é a volatilidade *EWMA*.

## 2.2 EXPECTED SHORTFALL

O *Expected Shortfall (ES)*, também conhecido como *Conditional Value at Risk (CVaR)*, *Expected Tail Loss (ETL)*, *Average Value at Risk (AVaR)* ou *Tail Value at Risk (TVaR)*, é uma medida de risco complementar que surgiu devido às necessidades do mercado perante as limitações apresentadas pelo *VaR*. O ES representa a média das perdas que superem o *VaR* para um mesmo nível de confiança e intervalo de tempo, sendo principalmente útil em situações onde é necessária informação sobre a cauda da distribuição de retornos, dado que se concentra no extremo das perdas.

De forma geral, supondo que  $X$  seja uma variável aleatória que denota a perda de um determinado ativo ou de uma carteira e  $VaR_\alpha(X)$  seja o *VaR* no nível de confiança  $100(1 - \alpha)\%$ , o  $ES_\alpha(X)$  é definido, segundo Yamai e Yoshiba (2002), pela equação:

$$ES_\alpha(X) = E(X|X > VaR_\alpha(X)) \quad (9)$$

No modelo histórico pode ser definido como:

$$ES = E(X|X > VaR) = \left( \frac{\sum_{i=[t\alpha]}^t X_{t(i)}}{t - [t\alpha]} \right) \quad (10)$$

Onde  $i$  é a  $i$ -ésima observação,  $t$  é o horizonte de tempo e  $t\alpha$  é o tamanho do intervalo do quantil do *VaR*.

Caso adote-se que os retornos sigam uma distribuição normal, o ES pode ser definido como:

$$ES = E(X|X > VaR) = E(X|X > Z_\alpha \sigma_t) \quad (11)$$

Onde  $\sigma$  é a volatilidade do ativo ou da carteira,  $t$  é o horizonte de tempo e  $Z_\alpha$  é o valor crítico associado ao nível de confiança  $\alpha$  da distribuição normal padrão.

Artzner et al. (1999) sugerem a utilização do *ES*, pois o mesmo cumpre todos os axiomas de uma medida de risco coerente, requisitos aos quais o *VaR* não cumpre. Além disso, o *VaR* não fornece nenhuma informação caso a perda ultrapasse o intervalo de confiança, sendo que nesses eventos é importante saber a magnitude da perda em relação ao *VaR* calculado. Devido ao *ES* utilizar a cauda esquerda da distribuição de retornos, ou seja, apenas as informações quando as perdas superam o *VaR* para um mesmo nível de confiança, o mesmo fornece uma medida de risco mais conservadora, o que lhe permite capturar com mais facilidade eventos extremos, garantindo uma maior segurança à gestão da carteira e ao processo de alocação de capital mínimo exigido por parte do BCBS.

### 3 MÉTODOS DE VALIDAÇÃO DE MODELOS

#### 3.1 BACKTESTING

Para validar a precisão e confiabilidade dos modelos de risco, é necessário adotar procedimentos que possam testar a validade estatística dos resultados obtidos. Essa validação, geralmente, é feita através do processo denominado *backtesting*. De forma geral, o *backtesting* é um processo estatístico destinado a verificar a eficácia e validade dos modelos testados. Essa prática é feita através da comparação entre as perdas estimadas e as perdas observadas, constituindo-se em uma abordagem essencial para garantir a robustez e utilidade prática desses modelos. Nesse estudo serão utilizados três modelos de *backtesting*: o Teste de Kupiec, Teste de Basileia e o *GBI*. Sendo que os dois primeiros serão utilizados para avaliar os modelos *VaR*, enquanto o último será utilizado para os modelos de *ES*.

##### 3.1.1 Teste de Kupiec

O Teste de Kupiec (1995), também conhecido como teste proporção de falhas (POF), é um método estatístico utilizado para verificar a frequência que as perdas podem superar o *VaR*, de modo que o modelo ainda seja considerado adequado, dado um nível de confiança.

Considera-se uma violação do *VaR*, situações em que o *i*-ésimo retorno,  $r_i$ , for menor que a perda máxima estimada no dia *i*, dado nível de confiança  $\alpha$ . Define-se então, a função indicadora de violação do *VaR*, como:

$$I_i = \begin{cases} 1, & \text{se } r_i < VaR_i(\alpha) \\ 0, & \text{se } r_i \geq VaR_i(\alpha) \end{cases} \quad (12)$$

A hipótese nula e a hipótese alternativa do teste são definidas por:

$$\begin{cases} H_0: p = p_0 \\ H_a: p \neq p_0 \end{cases} \quad (13)$$

Onde  $p$  é a proporção de falhas esperadas e  $p_0 = \frac{T}{N}$  é a proporção de falhas observadas.

A estatística do Teste de Kupiec segue uma distribuição Qui-quadrado com 1 grau de liberdade, probabilidade igual a  $1 - p$ ,  $N$  observações e  $T$  violações, e é definida pela fórmula:

$$LR = -2 \ln((1 - p)^{N-T} (p^T)) + 2 \ln\left(\left(1 - \frac{T}{N}\right)^{N-T} \left(\frac{T}{N}\right)^T\right) \quad (14)$$

Logo, se o resultado do teste for maior que a inversa da distribuição Qui-quadrado com 1 grau de liberdade, dado o nível de confiança utilizado, a hipótese nula será rejeitada e o modelo será considerado não adequado. A tabela abaixo, apresenta as regiões de não rejeição, de acordo com o tamanho da amostra e para diferentes níveis de confiança.

Tabela 1 – Região de não rejeição para teste POF para tamanhos de amostra alternativos

Nível de Probabilidade	Nível de Confiança VaR	Região de Não Rejeição de Violações		
		T = 255 dias	T = 510 dias	T = 1000 dias
0,010	99%	$N < 7$	$1 < N < 11$	$4 < N < 17$
0,025	97,5%	$2 < N < 12$	$6 < N < 21$	$15 < N < 36$
0,050	95%	$6 < N < 21$	$16 < N < 36$	$37 < N < 65$
0,075	92,5%	$11 < N < 28$	$27 < N < 51$	$59 < N < 92$
0,100	90%	$16 < N < 36$	$38 < N < 65$	$81 < N < 120$

Fonte: Kupiec (1995)

### 3.1.2 Teste de Basileia

O *Basel Traffic Light Test*, também conhecido como Teste de Basileia é uma metodologia de *backtesting* do VaR publicado em 1996 pelo BCBS no documento *Supervisory Framework for the Use of "Backtesting" in Conjunction with the Internal Models Approach to Market Risk Capital Requirements*. Esse teste verifica a quantidade de violações do VaR, Equação 12, classificando-as em três zonas distintas. O número de violações, segue uma função de probabilidade binomial definida por:

$$b(x, N, \alpha) = \binom{N}{x} \alpha^x (1 - \alpha)^{N-x} \quad (15)$$

Onde  $x$  é o número de violações do VaR,  $N$  é o total de retornos e  $\alpha$  é o nível de significância. Sendo que o BCBS recomenda utilizar uma janela de 250 retornos para  $N$ .

O BCBS (1996) define três zonas de violações: a zona verde, na qual a distribuição acumulada  $B(x, 250, \alpha) < 95\%$ , em seguida a zona amarela, onde  $95\% \leq B(x, 250, \alpha) < 99,99\%$  e a zona vermelha, onde  $B(x, 250, \alpha) \geq 99,99\%$ .

Abaixo, é apresentada a tabela segundo o Teste de Basileia utilizando o nível de confiança  $\alpha = 0,025$  para o VaR e o intervalo de retornos recomendado.

Tabela 2 - Teste de Basileia para o VaR com  $\alpha = 0,025$

Zona	Número de Violações	Probabilidade Cumulativa
Verde	0	0,18%
Verde	1	1,32%
Verde	2	4,97%
Verde	3	12,70%
Verde	4	24,95%
Verde	5	40,40%
Verde	6	56,57%
Verde	7	71,03%
Verde	8	82,29%
Verde	9	90,05%
Verde	10	94,85%
Amarela	11	97,53%
Amarela	12	98,90%
Amarela	13	99,54%
Amarela	14	99,82%
Amarela	15	99,94%
Amarela	16	99,98%
Vermelha	17	99,99%

Fonte: Elaboração Própria

O modelo será classificado como adequado caso se enquadre na zona verde, ou seja, se o número de violações estiver entre zero e quatro. Caso ocorram de cinco a nove falhas, o modelo é passível de aceitação, porém, a instituição financeira deve comprovar a robustez do modelo para continuar o utilizando, além disso, é adicionado um fator multiplicativo mais punitivo no cálculo de alocação de capital. E, por fim, se o modelo apresentar dez ou mais violações se enquadrará na zona vermelha e será classificado como inadequado, indicando que há um problema e que o mesmo precisa ser melhorado ou substituído.

### 3.1.3 Generalized Breach Indicator

O *Generalized Breach Indicator (GBI)* é um modelo de *backtesting* para o *ES*, desenvolvido por Costanzino e Curran (2018) e baseada na abordagem de zonas coloridas do Teste de Basileia para o *VaR* proposta pelo BCBS. O teste se baseia na função indicadora de violação do *VaR*, Equação 12, aplicada ao *ES*, resultando em um indicador que leva em conta a gravidade da violação.

O *ES* é definido em sua forma contínua por:

$$ES(\alpha) = \frac{1}{\alpha} \int_0^\alpha VaR(p) dp \quad (16)$$

A fórmula representa a média da cauda de diferentes quantis  $p$  da distribuição, dado o nível de significância  $\alpha$  utilizado para o *VaR*.

Similar ao Teste de Basileia do *VaR*, define-se o *GBI* do *ES* por  $X_{ES}^{(t)}$  que toma valores no intervalo  $[0, 1]$ , como:

$$X_{ES}^{(t)}(\alpha) = \frac{1}{\alpha} \int_0^\alpha I_t dp \quad (17)$$

$$X_{ES}^{(t)}(\alpha) = (I_t) \left( 1 - \frac{\phi(r_t)}{\alpha} \right) \quad (18)$$

Onde  $\phi$  é a função de distribuição acumulada da variável aleatória  $r_t$ ,  $\alpha$  o nível de significância e o termo  $1 - \phi(r_t)/\alpha$  determina a gravidade da violação. Supondo, que  $r_t = VaR_t$ , então  $\phi(r_t)/\alpha = 1$  e conseqüentemente  $X_{ES}^{(t)} = 0$ , o que significa que não houve severidade. Por outro lado, conforme  $r_t$  tende a valores mais negativos,  $\phi(r_t)$  tende a zero, o que implica em  $X_{ES}^{(t)} = 1$  e ocasiona uma gravidade severa. Assim, o  $X_{ES}^{(t)}$  monitora se ocorreu uma violação no dia  $t$ , bem como a gravidade dessa violação, logo, a severidade em todos os dias  $N$ , é dada por:

$$X_{ES}^N(\alpha) = \sum_{t=1}^N \frac{1}{\alpha} \int_0^\alpha I_t dp \quad (19)$$

$$X_{ES}^N(\alpha) = \sum_{t=1}^N (I_t) \left(1 - \frac{\phi(r_t)}{\alpha}\right) \quad (20)$$

Definindo o número de violações generalizadas do *ES* por  $x$ , tem-se que a probabilidade cumulativa de obter  $x$  ou menos violações é dada por:

$$P(X_{ES}^N(\alpha) \leq x) = \phi(x) \quad (21)$$

Portanto, para qualquer quantil  $q$ , pode-se calcular o valor de violação generalizada  $x$  correspondente invertendo a equação:

$$\lim_{x \in [0, \infty[} \sup \phi(x) < q \quad (22)$$

Porém, ao contrário do Teste de Basileia do *VaR*, os valores de violação para o *ES* são variáveis contínuas, fazendo mais sentido escolher o quantil e depois inverter para obter o valor da violação correspondente.

Sob a hipótese nula de que o modelo *ES* está corretamente especificado, a distribuição do  $X_{ES}^N(\alpha)$  é fornecida pelos autores com base na distribuição binomial e na distribuição de Irwin-Hall, sendo que a distribuição tende assintoticamente a uma distribuição normal para longos períodos de previsão, onde:

$$\mu_{ES} = \frac{1}{2} \alpha N \quad (23)$$

$$\sigma^2_{ES} = \alpha \left( \frac{4 - 3\alpha}{12} \right) \quad (24)$$

E, portanto:

$$\lim_{N \rightarrow \infty} X_{ES}^N(\alpha) \sim N(\mu_{ES}, \sigma^2_{ES}) \quad (25)$$

Assim, utilizando os quartis das zonas coloridas do Teste de Basileia do *VaR*, pode-se criar um teste similar de zonas coloridas para o *ES*.

A tabela a seguir fornece os quantis para diferentes valores de violação, utilizando os parâmetros  $\alpha = 2,5\%$  e  $N = 250$  retornos. Nota-se que no caso da abordagem de semáforo do *ES*, devido os valores serem contínuos, pequenas alterações podem resultar em uma alteração da zona de aceitação do teste.

Tabela 3 - Abordagem de Semáforo para o Expected Shortfall

Zona	Valor de Violação Generalizada	Probabilidade Cumulativa
Verde	0	0,18%
Verde	1,3929	10%
Verde	2,1131	25%
Verde	3,0276	50%
Verde	4,052	75%
Verde	5,0622	90%
Verde	5,7049	95%
Amarela	6,9844	99%
Amarela	8,5285	99,9%
Amarela	9,8833	99,99%
Vermelha	mais que 9,8833	99,99%

Fonte: Costanzino e Curran (2018)

Portanto, um  $X_{ES}^N(\alpha)$  elevado indica uma taxa correspondentemente alta de violação do *ES*, sugerindo, assim, uma possível subestimação do risco.

## 4 METODOLOGIA

A ferramenta computacional utilizada para implementar todas as simulações deste trabalho foi o Microsoft Excel, utilizando a linguagem de programação incorporada Visual Basic for Applications (VBA).

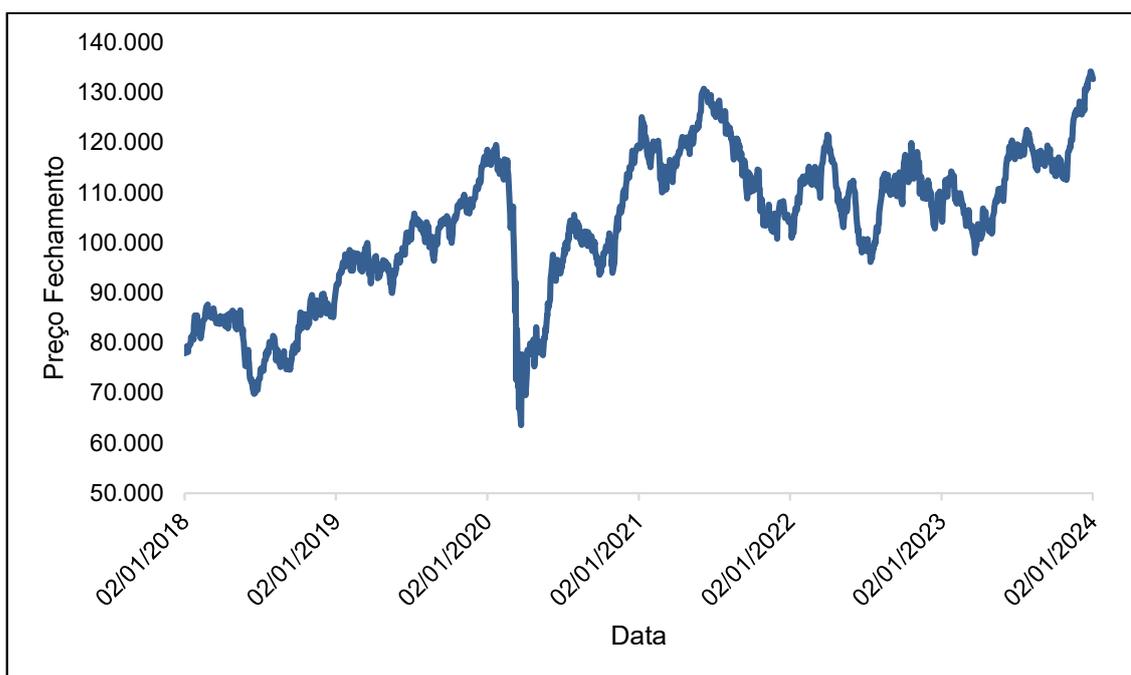
### 4.1 DADOS

Esse estudo utilizará dados diários do preço de fechamento do Índice Ibovespa, o principal indicador do desempenho médio das ações das empresas mais importantes do mercado de capitais brasileiro negociadas na Bolsa de Valores de São Paulo (B3). O índice é expresso em pontos e é reavaliado a cada quatro meses, sendo resultado de uma carteira teórica de ativos, e a escolha de sua composição é definida conforme critérios estabelecidos pela B3. O Ibovespa é amplamente utilizado como indicador do desempenho do mercado de ações nacional, sendo útil para avaliar a tendência do mercado e auxiliar na tomada de decisão de investimentos. Além disso, o índice apresentou alta volatilidade nos últimos anos, o que somado às suas outras características, o torna ideal para avaliar a eficiência dos modelos que são objetos de estudo desse trabalho.

Os dados da série histórica serão extraídos diretamente do site da B3 e será utilizada uma janela de tempo de 6 anos (2018 a 2023), analisada em intervalos de 250 dias úteis, agrupadas em cinco grupos anuais, cada um correspondendo aos anos de 2019 a 2023. Dado que as métricas utilizam uma janela móvel com os 250 retornos anteriores mais recentes, foi necessário a utilização dos dados de 2018 para realização dos cálculos referentes ao ano de 2019.

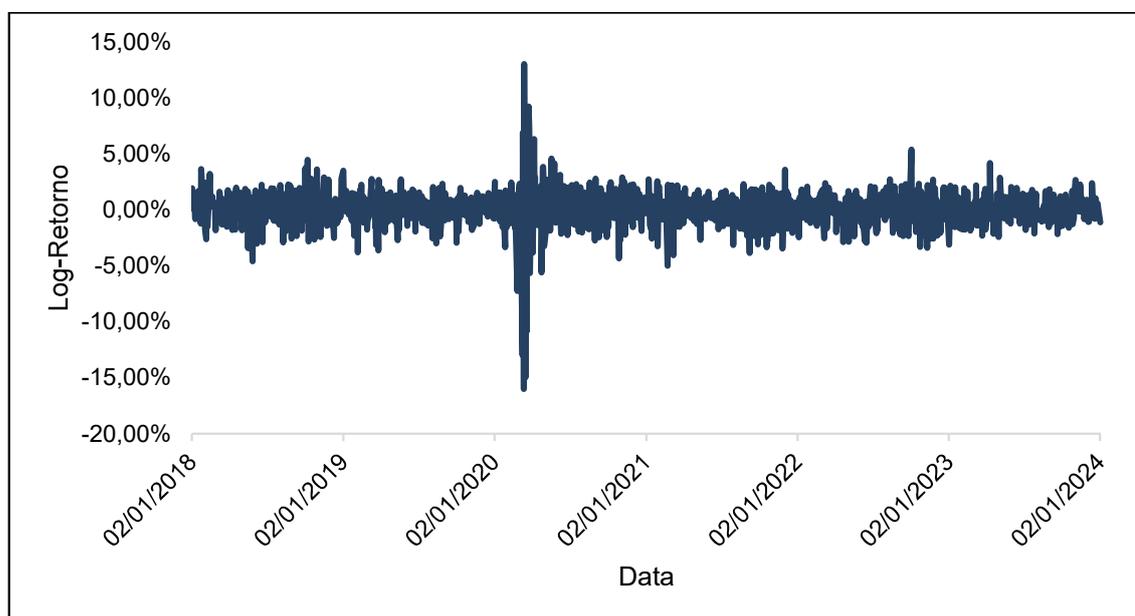
A janela de tempo escolhida engloba o período da pandemia da COVID-19, onde os mercados globais passaram por períodos de alta volatilidade motivadas por significativas quedas de preços, e também o período pós pandemia, onde acompanhou-se posteriormente a recuperação e estabilização dos mercados. Esse cenário possibilita testar os modelos tanto em períodos de normalidade, como em situações de crise.

Figura 1 – Preço Fechamento Ibovespa Jan/2018 a Jan/2024



Fonte: Elaboração Própria

Figura 2 – Retornos Logarítmicos Ibovespa Jan/2018 a Jan/2024



Fonte: Elaboração Própria

## 4.2 ANÁLISE E MODELAGEM

Os retornos serão calculados de forma logarítmica, também conhecido como log-retorno. O retorno logarítmico representa a variação percentual contínua do preço

ao longo do tempo. Esse método possui a vantagem de ser aditivo, facilitando o cálculo de retornos acumulados, bastando somar os retornos diários. A fórmula é dada por:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (26)$$

Onde  $P_t$  é o preço de fechamento do Ibovespa no dia  $t$ .

Utilizando os dados, conforme descrito anteriormente, será realizado cálculo do *VaR* e do *ES* pelas metodologias Histórica, Paramétrico Normal, *EWMA* e *Cornish-Fisher* para a janela de tempo proposta, para posteriormente testar a eficiência dos modelos através dos métodos de *backtesting* adotados nesse estudo.

O horizonte de tempo utilizado para o cálculo do *VaR* e do *ES* será de um dia útil e será adotada uma abordagem de janela móvel para o cálculo diário dessas métricas, ou seja, o *VaR* e o *ES* do dia  $t$  será calculado com base nos 250 retornos mais recentes, começando em  $t - 1$  e terminando em  $t - 250$ . Dessa forma, à medida que se avança um dia no tempo, tanto as métricas quanto os retornos avançam um dia, incorporando sempre os valores mais recentes e removendo os mais antigos. Essa técnica ajuda a suavizar flutuações nos dados, destacando tendências de longo prazo e facilitando a análise de padrões ao longo do tempo.

Em média, um ano compreende 252 dias úteis. Dessa forma, serão constituídos cinco conjuntos anuais, cada um contendo 250 resultados de ambas as métricas, abrangendo o período janeiro de 2019 a dezembro de 2023. Nos casos em que um ano apresentar uma quantidade de observações inferior ao necessário, serão incorporados os resultados das métricas mais recentes do ano anterior para completar a amostra. Em situações em que houver excesso de observações, serão utilizados os 250 primeiros resultados do ano em questão.

Conforme orientações do BCBS apresentadas em seu documento de consultivo *Fundamental Review of the Trading Book: A revised market risk framework* (2013), recomenda-se a adoção do nível de confiança 99% para o *VaR* e seu equivalente para o *ES*, que é de 97,5%. Nos estudos conduzidos por Costanzino e Curran (2018), são fornecidos os quantis para o nível de confiança de 97,5%. De acordo com as orientações do BCBS e considerando a complexidade do procedimento numérico

utilizado para estimar com precisão os quantis do *backtesting* do *ES* para outros níveis de confiança, optou-se por utilizar o intervalo de confiança de 97,5%.

Para avaliar a eficiência dos modelos, serão empregados o Teste de Basileia e o Teste de Kupiec para os modelos de *VaR*, juntamente com o recente indicador proposto por Costanzino e Curran (2018), *GBI*, para o *ES*. Em todos os casos, os testes são fundamentados no número de violações, visando verificar se o número de falhas identificadas pelo modelo está ou não na região de não rejeição.

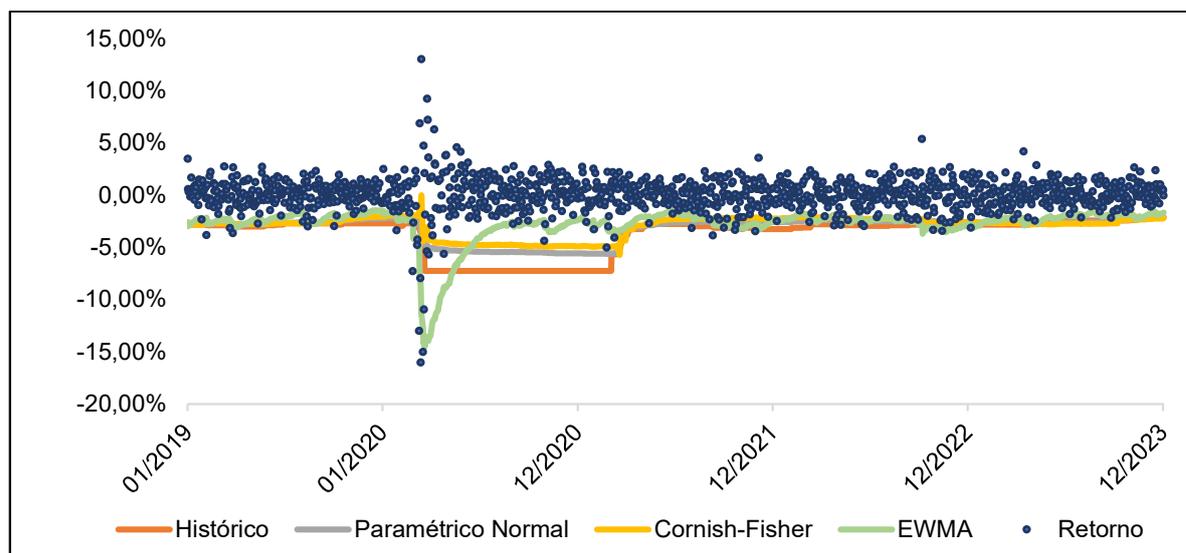
## 5 RESULTADOS

Neste capítulo são apresentados os resultados dos *backtesting* aplicados aos modelos de *VaR* e *ES* para a janela móvel de 250 dias e nível de confiança de 97,5%. Em seguida, os resultados obtidos entre os modelos testados são discutidos. Por fim, é realizada uma análise do impacto dos desvios entre o retorno observado e as métricas de riscos utilizadas para o gerenciamento do risco. A escolha dos parâmetros utilizados para os cálculos e os testes dos modelos foi definida com o intuito de simular de forma semelhante ao que é comumente adotado pelo mercado financeiro atualmente, dadas as limitações do estudo.

Devido os modelos estarem sendo calculados diretamente sobre o índice Ibovespa, e não em termos financeiros ou em uma simulação de uma carteira de investimentos, os valores das métricas serão apresentados em termos percentuais, assim, facilitando a compreensão da magnitude dos valores.

A seguir, apresentam-se os valores dos modelos de *VaR* e os retornos para o período analisado neste estudo. Esta representação possibilita a identificação dos momentos em que ocorreram violações para cada modelo de *VaR*.

Figura 3 – Retornos x Modelos de VaR Ibovespa Jan/2019 a Dez/2023



Fonte: Elaboração Própria

A tabela abaixo apresenta o resultado do número de violações após a aplicação de todos os modelos citados de *VaR* e *ES*. É importante ressaltar que, no caso do *backtesting* dos modelos de *ES*, o número de violações é uma variável contínua,

devido à incorporação do fator de severidade em relação à forma que são contabilizadas as violações do *VaR*.

Tabela 4 – Total de violações por período e modelos de *VaR* e *ES*

Modelo	Ano				
	2019	2020	2021	2022	2023
VaR Histórico	5	9	4	4	1
VaR Paramétrico Normal	5	13	6	8	1
VaR Cornish Fisher	5	13	10	12	1
VaR EWMA	9	8	12	9	5
ES Histórico	3,1890	8,9174	3,2035	2,9155	0,6206
ES Paramétrico Normal	3,1890	10,4819	4,0667	4,2704	0,6206
ES Cornish-Fisher	3,0426	11,8245	5,3152	5,8572	0,6035
ES EWMA	6,4233	6,4223	7,7812	4,8236	2,2549

Fonte: Elaboração Própria

Chama a atenção na tabela o fato de que, embora o modelo *EWMA* tenha acumulado 43 violações do *VaR* ao longo do período, o maior número entre todos os modelos estudados, o mesmo foi o que apresentou o menor número de violações em 2020, ano auge da pandemia da COVID-19. A explicação para esse resultado é que as maiores volatilidades ocorreram no início do ano, e como o *backtesting* é realizado ao final do período testado, o modelo deu mais peso às amostras mais recentes, ou seja, aquelas do final do ano, as quais foram menos severas.

As tabelas seguintes apresentam os resultados dos *backtesting* para todos os modelos de *VaR* e *ES*, com base nos parâmetros detalhados anteriormente.

Tabela 5 – Resultados Backtesting *VaR* para janela de 250 dias e nível de confiança de 97,5%

Modelo	Backtesting	Ano				
		2019	2020	2021	2022	2023
VaR Histórico	Kupiec	Adequado	Adequado	Adequado	Adequado	Não Adequado
	Teste de Basileia	Verde	Verde	Verde	Verde	Verde
VaR Paramétrico Normal	Kupiec	Adequado	Não Adequado	Adequado	Adequado	Não Adequado
	Teste de Basileia	Verde	Amarela	Verde	Verde	Verde
VaR Cornish-Fisher	Kupiec	Adequado	Não Adequado	Adequado	Adequado	Não Adequado
	Teste de Basileia	Verde	Amarela	Verde	Amarela	Verde
VaR EWMA	Kupiec	Adequado	Adequado	Adequado	Adequado	Adequado
	Teste de Basileia	Verde	Verde	Amarela	Verde	Verde

Fonte: Elaboração Própria

Tabela 6 – Resultados Backtesting ES para janela de 250 dias e nível de confiança de 97,5%

Modelo	Ano				
	2019	2020	2021	2022	2023
ES Histórico	Verde	Amarela	Verde	Verde	Verde
ES Paramétrico Normal	Verde	Vermelha	Verde	Verde	Verde
ES Cornish-Fisher	Verde	Vermelha	Verde	Amarela	Verde
ES EWMA	Amarela	Amarela	Amarela	Verde	Verde

Fonte: Elaboração Própria

O modelo Histórico apresentou um desempenho satisfatório de 2019 a 2022, sendo considerado adequado pelo Teste de Kupiec. Entretanto, em 2023, o modelo foi considerado não adequado por apresentar apenas uma falha, ficando de fora do intervalo de violações para os parâmetros utilizados. Quando o número de violações observadas é inferior ao esperado, isso pode indicar que o modelo está subestimando o risco. No entanto, é importante considerar os parâmetros e a magnitude das violações. Neste caso específico, houve apenas uma violação no intervalo de 250 dias úteis, algo próximo do limite inferior de duas violações para o modelo ser considerado adequado por esse teste. Essa situação é comum em modelos históricos com janelas de tempo consideradas pequenas, visto que pode haver uma subestimação do risco devido a pouca quantidade de extremos na distribuição de retornos. Além disso, o Teste de Kupiec (1995) apresenta baixa confiança para amostras pequenas, podendo não rejeitar a hipótese nula quando ela é falsa. No Teste de Basileia o modelo Histórico foi classificado na zona verde em todos os anos analisados, sendo assim, considerado adequado. Em relação ao *ES*, o modelo proposto de backtesting, *GBI*, baseado no Teste de Basileia, enquadrou o modelo na zona verde de aceitação em todos os anos, exceto em 2020, período em que ocorreu o maior número de violações. Essa diferenciação na classificação das zonas coloridas de ambos os testes é devida ao fato de que o Teste de Basileia aborda apenas sobre a frequência das violações, ignorando outras informações relevantes, como a magnitude das perdas. Essa deficiência é corrigida no *backtesting* do *GBI*, que adiciona um fator de severidade às violações, enquadrando de forma mais eficiente o modelo na zona correspondente.

O modelo *EWMA* demonstrou eficiência em todos os anos, segundo o Teste de Kupiec, e no Teste de Basileia manteve-se predominantemente na zona verde do teste, com exceção de 2021, quando foi classificado na zona amarela. Para o *ES*, o *GBI* apresentou novamente resultados ligeiramente diferentes ao Teste de Basileia, demonstrando o impacto do fator de severidade. No caso, do modelo *EWMA*, incorpora-se uma correção adicional em relação ao teste utilizado para o *VaR*. Além da magnitude das perdas, o Teste de Basileia desconsidera o período em que ocorreu essa violação, algo corrigido por esse modelo, pois o mesmo pondera os retornos dando maior peso às observações mais recentes.

Para o modelo *Cornish-Fisher*, o Teste de Kupiec indicou adequação do modelo para os anos de 2019, 2021 e 2022, sendo considerado não adequado nos demais anos. No Teste de Basileia, o modelo foi classificado na zona verde nos anos de 2019, 2021 e 2023 e na zona amarela em 2020 e 2022. No caso do *ES*, as zonas mantiveram-se as mesmas, exceto em 2020, quando o fator de severidade do ano somou 11,8245, levando o modelo à zona vermelha, sendo assim considerado inadequado. Essa alteração na zona entre o Teste de Basileia e o *GBI* ocorreu novamente devido à severidade aplicada no teste, especialmente porque o ano de 2020 foi impactado pelo auge da pandemia da COVID-19. Durante esse período, a Bolsa de valores acionou o mecanismo de segurança, *Circuit Breaker*, seis vezes, com todas as quedas sendo superiores a 10%, caracterizando um período de volatilidade extrema. Além disso, é importante salientar que o modelo de *Cornish-Fisher* tende a ter melhores resultados para janelas de tempo maiores, devido a uma maior suavização da assimetria e curtose da distribuição.

Por fim, o modelo Paramétrico Normal demonstrou um desempenho satisfatório nos anos de 2019, 2021 e 2022, sendo considerado não adequado no Teste de Kupiec em 2023 pelos mesmos motivos explicados no modelo Histórico. Contudo, em 2020, ano mais volátil do período analisado, o modelo foi reprovado pelo Teste de Kupiec e ficou na zona amarela pelo Teste de Basileia, apresentando um alto número de falhas. Neste caso, a não aprovação do modelo pode ser justificada pela fragilidade à sensibilidade aos pressupostos de normalidade, especialmente em anos com eventos atípicos. Como o período analisado apresentou alta volatilidade do índice Ibovespa, a distribuição apresentou um comportamento distante da normal, dificultando os modelos capturarem corretamente o comportamento da mesma e, conseqüentemente comprometendo os resultados. Quanto ao *Expected Shortfall*, a classificação na zona

vermelha nesse mesmo ano reforça a dificuldade do modelo paramétrico normal em lidar com caudas pesadas, confirmando sua limitação na previsão de eventos extremos.

Por último, é calculado os desvios absolutos médios entre o retorno e o *VaR*, assim como entre o retorno e o *ES*, apenas nos casos em que o retorno é menor que o *VaR*, ou seja, quando há uma violação.

No caso do *VaR*, o desvio médio absoluto é dado por:

$$\Delta_{VaR}(\alpha) = \frac{\sum_{t=1}^N (I_t) |r_t - VaR_t(\alpha)|}{\sum_{t=1}^N I_t} \quad (27)$$

Análogo a Equação 26, define-se o desvio médio absoluto para o *ES* como:

$$\Delta_{ES}(\alpha) = \frac{\sum_{t=1}^N (I_t) |r_t - ES_t(\alpha)|}{\sum_{t=1}^N I_t} \quad (28)$$

É importante ressaltar que ambas as fórmulas representam medidas de dispersão em relação às perdas. No entanto, ao compará-las, é crucial lembrar das diferenças de definição entre o *VaR* e o *ES*.

A tabela a seguir apresenta os resultados obtidos dos desvios para ambas as métricas.

Tabela 7 – Desvio Absoluto Médio das violações do retorno para o *VaR* e o *ES*

Modelo	Ano				
	2019	2020	2021	2022	2023
VaR Histórico	0,49%	5,77%	0,44%	0,39%	0,34%
VaR Paramétrico Normal	0,62%	4,40%	0,70%	0,45%	0,51%
VaR Cornish Fisher	0,59%	5,09%	0,66%	0,42%	0,49%
VaR EWMA	0,77%	3,35%	0,51%	0,44%	0,28%
ES Histórico	0,30%	3,26%	0,53%	0,39%	0,03%
ES Paramétrico Normal	0,26%	3,98%	0,31%	0,22%	0,01%
ES Cornish-Fisher	0,29%	4,34%	0,56%	0,26%	0,01%
ES EWMA	0,50%	2,54%	0,53%	0,28%	0,16%

Fonte: Elaboração Própria

Os resultados indicam que durante os períodos de normalidade, o modelo Histórico apresentou os menores desvios entre o *VaR* e o retorno, se aproximando

mais da perda quando esta superou o limite máximo estimado para o dia. Em relação aos desvios do *ES*, o modelo Paramétrico Normal nas mesmas condições registrou os menores valores, se aproximando mais da perda.

## 6 CONCLUSÃO

Devido à crescente necessidade e importância de um gerenciamento de riscos eficiente, esse estudo propôs testar duas das métricas mais amplamente utilizadas no mercado financeiro, o *VaR* e o *ES*, por meio dos modelos Históricos, Paramétrico Normal, *Cornish-Fisher* e *EWMA*, utilizando dados do índice Ibovespa. O momento é particularmente propício, dado que as instituições financeiras passam por revisão das regras de cálculo de capital para risco de mercado, e uma das mudanças é a substituição do *VaR* pelo *ES*. Contudo, uma das críticas a essa mudança é a falta de consenso do mercado e a insuficiente literatura sobre métodos eficazes de *backtesting* para o *ES*, essenciais para garantir a eficiência e a validação dos modelos. Nesse contexto, foi utilizada a abordagem *Generalized Breach Indicator (GBI)* proposta por Costanzino e Curran (2018), baseada no Teste de Basileia, sendo este estudo, salvo melhor juízo, o primeiro a testar essa abordagem de *backtesting* considerando ativos emitidos no Brasil.

Ao analisar os resultados para o *backtesting* do *VaR*, observou-se que o modelo Histórico apresentou desempenho superior, permanecendo na zona verde em todos os anos pelo Teste de Basileia. O Teste de Kupiec reprovou o modelo apenas em 2023 por insuficiência de violações. No entanto, conforme detalhado, o Teste de Kupiec apresenta baixa confiança para amostras pequenas, sendo importante considerar outros parâmetros, como a magnitude da violação, que no caso do modelo Histórico apresentou menor desvio absoluto médio no período. Entre os demais modelos, nenhum ficou situado na zona vermelha segundo o Teste de Basileia, e no Teste de Kupiec, tirando os casos de insuficiência de violações, somente os modelos Paramétrico Normal e *Cornish-Fisher* foram reprovados em 2020, período da pandemia da COVID-19. Em geral, os modelos apresentam resultados satisfatórios em períodos de normalidade, mas demonstraram vulnerabilidades em situações de alta volatilidade.

Para o *ES*, a recente abordagem *GBI* demonstrou coerência, seguindo em linha com os resultados do Teste de Basileia para o *VaR*, destacando a eficiência do fator de severidade incorporado em sua metodologia. Observou-se que em anos nos quais as violações foram mais severas, o *GBI* aplicado ao *ES* diferenciou-se do Teste de Basileia aplicado ao *VaR*, resultando em uma alteração para uma zona de menor

aceitação. Essa alteração foi observada em todos os modelos durante o ano de 2020 e no modelo *EWMA* em 2019.

Em relação os modelos de *backtesting* e sua utilização no mercado financeiro, o Teste de Kupiec e o Teste de Basileia ofereceram métricas úteis para avaliar a adequação dos modelos de *VaR*, enquanto o *GBI* trouxe uma abordagem inovadora, considerando não apenas a ocorrência de violações, mas também a gravidade das mesmas para o *ES*. Essa metodologia mais abrangente, incorporando a severidade das violações, destacou a importância de compreender não apenas a frequência, mas também a magnitude das perdas.

As diferenças nas classificações entre os modelos de *backtesting* ressaltam a complexidade da gestão de riscos, especialmente em períodos de eventos extremos, como a pandemia de COVID-19. A sensibilidade dos modelos às condições de mercado atípicas destaca a necessidade de uma abordagem flexível na escolha e ajuste dos modelos de risco. Em relação aos parâmetros utilizados, há a necessidade de testar para diferentes níveis de significância em janelas de tempo maiores, visto que intervalos menores podem não capturar corretamente retornos significantes da distribuição.

A escolha do modelo mais apropriado dependerá das características específicas do portfólio e das necessidades provenientes da gestão de riscos à qual se aplica. É fundamental reconhecer a dinâmica do mercado financeiro e a possibilidade de cenários excepcionais ao selecionar e ajustar os modelos.

## REFERÊNCIAS

ACERBI, C.; TASCHE, D. **Expected Shortfall**: A Natural Coherent Alternative to Value at Risk. *Economics Notes*, v. 31, p. 379-388, 2002.

ARTZNER, F. et al. **Coherent Measures of Risk**. *Mathematical Finance*, v. 9, n. 3, p. 203-228, 1999.

BANCO CENTRAL DO BRASIL. **Circular nº 3.646**, 2013. Disponível em: <<https://www.bcb.gov.br/htms/Normativ/CIRCULAR3646.pdf>>. Acesso em: 26 de janeiro de 2024.

BASLE COMMITTEE ON BANKING SUPERVISION. **Amendment to the capital accord to incorporate market risks**, 1996. Disponível em: <<https://www.bis.org/publ/bcbs24.pdf>>. Acesso em: 26 de janeiro de 2024.

BASLE COMMITTEE ON BANKING SUPERVISION. **Fundamental review of the trading book**: A revised market risk framework, 2013. Disponível em: <<https://www.bis.org/publ/bcbs265.pdf>>. Acesso em: 26 de janeiro de 2024.

BASLE COMMITTEE ON BANKING SUPERVISION. **Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements**, 2013. Disponível em: <<https://www.bis.org/publ/bcbs22.pdf>>. Acesso em: 26 de janeiro de 2024.

CARITÁ, F. et al. **Teoria dos valores extremos adequa-se ao Ibovespa na crise de 2008?**. Insper Working Paper, 2015.

COSTANZINO, N.; CURRAN, M. **A Simple Traffic Light Approach Backtesting Expected Shortfall**. New York: Risks, 2018.

HARMANTIS, F.; MIAO, L.; CHIEN, Y. **Empirical study of value-at-risk and expected shortfall models with heavy tails**. New Jersey: *The Journal of Risk Finance*, v. 7, n. 2, p. 117-135, 2006.

HOLTON, G. A. **History of Value-at-Risk: 1922-1998**. Boston: Working Paper, 2002.

KUPIEC, P. **Techniques for verifying the accuracy of risk measurement models**. *The Journal of Derivatives*, 1995.

JORION, P. **Value at Risk: The New Benchmark for Managing Financial Risk**. 3<sup>th</sup> ed. New York: The McGraw-Hill Companies, 2007.

JP MORGAN. **Risk Metrics**: Technical Document. 4<sup>th</sup> ed. 1997. Disponível em: <<http://www.jpmorgan.com/RiskManagement/RiskMetrics/RiskMetrics.html>>. Acesso em: 26 de janeiro de 2024.

LEAVENS, D. **Diversification of Investments**. *Trusts and Estates*, 1945. 469-479 p.

LINSMEIER, T.; PEARSON, N. **Risk Measurement**: An Introduction to Value at Risk. University of Illinois at Urbana-Champaign, 1996.

MARKOWITZ, H. **Potfolio Selection**. The Journal of Finance, 1952. v. 7.

PÉRIGNON, C.; SMITH, D. **The Level and quality of Value-at-Risk disclosure by comercial banks**. Journal of Banking and Finance, v. 34, p. 362-377, 2010.

PINHO, C. et al. **Risco Financeiro**: Medida e Gestão. 2<sup>th</sup> ed. Lisboa: Edições Sílabo, 2019.

SCAILLET, O. **Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall**. Mathematical Finance, 2004. v. 14.

VAN DER VAART, A. W. **Asymptotic Statistics**. Cambridge University Press, 1998. v. 3.

YAMAI, Y.; YOSHIBA, T. **Comparative Analyses of Expected Shortfall and Value-at-Risk**: Their Estimation Error, Decomposition, and Optimization. Monetary and Economic Studies, 2002.