



Trabalho de Conclusão de Curso

Construção de portfólios de longo prazo

Rafael Bernardoni Chaves

19 de maio de 2022

Rafael Bernardoni Chaves

Construção de portfólios de longo prazo

Trabalho de Conclusão apresentado à comissão de Graduação do Departamento de Estatística da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientador: Prof. Dr. Cleiton Guollo
Taufemback

Porto Alegre
Maio de 2022

Rafael Bernardoni Chaves

Construção de portfólios de longo prazo

Este Trabalho foi julgado adequado para obtenção dos créditos da disciplina Trabalho de Conclusão de Curso em Estatística e aprovado em sua forma final pela Orientador e pela Banca Examinadora.

Orientador: _____
Prof. Dr. Cleiton Guollo Taufemback, UFRGS
Universidade Federal do Rio Grande do Sul –
Porto Alegre, RS

Banca Examinadora:

Prof. Dr. Hudson da Silva Torrent, UFRGS
Universidade Federal do Rio Grande do Sul – Porto Alegre, RS

Profa. Dra. Taiane Schaedler Prass, UFRGS
Universidade Federal do Rio Grande do Sul – Porto Alegre, RS

Porto Alegre
Maio de 2022

“Faça da pedra de tropeço, um degrau de subida.” (Bruce Lee)

Agradecimentos

Agradeço aos meus pais Alvaro e Nilce pelo seu amor e apoio, por sempre me ajudarem a manter a cabeça erguida perante os desafios da vida. Vocês são o meu maior porto seguro, agradeço muito a vocês pelo que sou e até onde eu cheguei.

Ao meu irmão Eduardo, que sempre foi um exemplo de foco e resiliência para mim. Mesmo longe, sempre estive próximo de mim em diversos momentos nos últimos anos. Também à minha cunhada Marcela, que sempre fez o papel de irmã mais velha pra mim.

Aos meus colegas e amigos de faculdade, em especial Rafaela, Gustavo e Gabriel que compartilharam comigo ao longo desses anos de graduação das alegrias e das dificuldades dentro e fora do curso.

Ao meu orientador Cleiton Guollo Taufemback. Sou muito grato pelo seu apoio, pelos ensinamentos e orientações. Você foi um professor muito além dos conteúdos do currículo, por isso sou grato.

A professora Taiane e ao professor Hudson, por serem banca deste trabalho e se disponibilizarem nas orientações e correções a serem feitas.

A todos os professores do departamento de Matemática e Estatística da UFRGS, que compartilharam de seus conhecimentos e ensinamentos para a minha formação profissional e acadêmica.

Por fim, gostaria de agradecer aos meus amigos e familiares que de alguma forma sempre contribuíram nos meus dias para que a jornada da vida fosse mais prazerosa.

Resumo

Portfólios de longo prazo são geralmente construídos considerando ativos de baixo risco e com um bom fluxo de dividendos. Entretanto, a metodologia proposta neste trabalho é uma alteração do modelo de Markowitz, considerando a remoção da variabilidade de curto prazo dos retornos, com o objetivo de construir um portfólio que seja menos sensibilizado por variações rápidas e temporárias. Utilizando dados de 20 anos dos ativos presentes na composição histórica do índice Dow Jones Industrial Average, comparamos os resultados do método proposto com o modelo tradicional de Markowitz e o modelo Naive, para horizontes de três, seis e doze meses. Também avaliamos quatro frequências máximas para o “low-pass filter”. Os resultados indicam que portfólios construídos com a metodologia proposta, geralmente, resultam em retornos superiores em relação aos modelos de benchmark. Análises de drawdown e de dominância estocástica também indicam a superioridade do nosso método.

Palavras-Chave: Portfólio, Séries Temporais, Investimentos, Filtragem, Markowitz, Dow Jones, Domínio da frequência.

Abstract

Long-term portfolios are usually built considering low-risk assets with a high payment dividend flow. However, the methodology proposed in this work is an alteration of the Markowitz model, considering the removal of short-term variability in the returns, aiming to build a portfolio that is less sensitive to rapid and temporary variations. Using data from 20 years of assets present in the historic composition of the Dow Jones Industrial Average, we compared the results of the proposed method with the traditional Markowitz model and the Naive model, for horizons of three, six and twelve months. We also implemented four maximum frequencies for the low-pass filter. The results indicate that portfolios built with the proposed methodology, generally, result in superior returns in relation to benchmark models. Drawdown and stochastic dominance analyzes also indicate the superiority of our method.

Keywords: Time series, Low-pass filter, Markowitz, Dow Jones.

Sumário

| | | |
|-----|---|----|
| 1 | Introdução | 11 |
| 2 | Referencial Teórico | 12 |
| 2.1 | Markowitz e Teoria Moderna de Portfólio | 12 |
| 2.2 | Índice de Sharpe | 13 |
| 2.3 | Low-pass Filter | 15 |
| 2.4 | Dominância Estocástica | 16 |
| 2.5 | Drawdown | 16 |
| 2.6 | Teste de Ledoit-Wolf | 17 |
| 2.7 | Taxa de <i>Turnover</i> | 18 |
| 3 | Modelo Proposto | 19 |
| 4 | Fontes e Pré Processamento de Dados | 20 |
| 4.1 | Aquisição de Dados | 20 |
| 4.2 | Períodos Atípicos | 22 |
| 4.3 | Merge e Fechamento de Empresas nos Períodos | 22 |
| 5 | Estudo Empírico | 24 |
| 6 | Conclusão | 34 |
| | Referências Bibliográficas | 34 |

Lista de Figuras

| | |
|---|----|
| Figura 2.1: Fronteira eficiente. | 14 |
| Figura 2.2: Fronteira eficiente e Linha Tangencial. | 14 |
| Figura 4.1: Dow Jones e períodos de recesso. Fonte: Yahoo Finance. | 23 |
| Figura 5.1: Histograma dos Retornos dos Portfólios para o $h = 3$ Meses. | 27 |
| Figura 5.2: Histograma dos Retornos dos Portfólios para o $h = 6$ Meses. | 30 |
| Figura 5.3: Histograma dos Retornos dos Portfólios para o $h = 12$ Meses. | 33 |

Lista de Tabelas

| | |
|---|----|
| Tabela 4.1: Lista das empresas pertencentes índice Dow Jones entre 2000 e 2022. | 21 |
| Tabela 5.1: Frequências utilizadas para o Low-pass Filter. | 24 |
| Tabela 5.2: Comparativo descritivos dos resultados para $h = 3$ | 26 |
| Tabela 5.3: Dominância Estocástica de primeira ordem para $h = 3$ | 26 |
| Tabela 5.4: Dominância Estocástica de segunda ordem para $h = 3$ | 27 |
| Tabela 5.5: P-valores para o teste de Ledoit & Wolf, $h = 3$ | 28 |
| Tabela 5.6: Comparativo descritivos dos resultados para $h = 6m$ | 28 |
| Tabela 5.7: Dominância Estocástica de primeira ordem para $h = 6m$ | 29 |
| Tabela 5.8: Dominância Estocástica de segunda ordem para $h = 6m$ | 29 |
| Tabela 5.9: P-valores para o teste de Ledoit & Wolf, $h = 6$ | 29 |
| Tabela 5.10: Comparativo descritivos dos resultados para $h = 12m$ | 31 |
| Tabela 5.11: Dominância Estocástica de primeira ordem para $h = 12m$ | 31 |
| Tabela 5.12: Dominância Estocástica de segunda ordem para $h = 12m$ | 31 |
| Tabela 5.13: P-valores para o teste de Ledoit & Wolf, $h = 12$ | 32 |
| Tabela 5.14: <i>Turnover Ratio</i> para os três horizontes. | 32 |

1 Introdução

Investidores precisam lidar com o fato de que todo o retorno possui um risco associado. Porém, como apontado em [Markowitz \(1952\)](#), apesar dessa relação entre risco e retorno parecer inseparável, diversificação de ativos pode ajudar a minimizar o risco de um portfólio. Seguindo a teoria proposta por Markowitz, ainda é possível para o investidor construir um portfólio eficiente com uma relação ótima entre risco e retorno.

De acordo com [Cochrane \(2021\)](#), para a construção de um portfólio de longo prazo o investidor deve ir atrás de ativos que possuem uma baixa volatilidade de longo prazo, de forma que possa garantir o retorno a longo prazo do investimento. Ativos de renda variável como ações possuem alta volatilidade quando observados em frequência diária, mas como apontado por [Adrian e Rosenberg \(2008\)](#), essas volatilidades diárias não são relevantes ao investidor de longo prazo já que este só pretende resgatar o valor investido em um horizonte maior de investimento.

Apesar disso, conforme [Pastor e Stambaugh \(2012\)](#), a volatilidade de ações é maior para horizontes maiores, tornando expectativas de retorno mais imprecisas. Contudo, como demonstrado e concluído em [Engle e Lee \(1999\)](#), modelos de retorno de ativos variáveis já são construídos com componentes que levam em consideração essa variabilidade de longo prazo.

O objetivo principal deste trabalho é apresentar uma metodologia de construção de portfólios que é baseada no risco associado dos ativos, porém reduzindo a volatilidade de longo prazo no comportamento da série de retornos. Para a construção do trabalho, as principais referências literárias foram a teoria moderna de portfólio de Markowitz e a teoria de filtro passa-baixa, apresentadas nas obras de [Markowitz \(1952\)](#), [Christiano e Fitzgerald \(2003\)](#) e [Müller e Watson \(2018\)](#).

Esse trabalho é dividido da seguinte forma: No Capítulo 2 são apresentados os principais referenciais teóricos utilizados. No Capítulo 3 são apresentados o modelo proposto e como foram feitos os cálculos para a construção do mesmo. No Capítulo 4, são mostradas as fontes de dados utilizadas. O Capítulo 5 é dedicado às análises dos resultados e, por fim, o Capítulo 6 apresenta as conclusões feitas sobre o estudo.

2 Referencial Teórico

Nesse capítulo serão abordados tópicos da literatura pertinentes ao desenvolvimento, implementação e avaliação do modelo proposto para portfólios de longo prazo, conforme Capítulo 3. Iniciando com a teoria moderna de portfólio, passando pelo desenvolvimento posterior da teoria por Sharpe (1998), seguindo com o modelo de filtro passa-baixa, ou ‘low-pass filter’ e finalizando com o conceito de dominância estocástica e drawdowns.

2.1 Markowitz e Teoria Moderna de Portfólio

Conforme dito em Markowitz (1952), o processo de construção de um portfólio é composto de duas etapas: a primeira é a de seleção de ativos baseando-se em observações e experiências que reforçam crenças em relação à performance de retornos futuros. A segunda etapa é combinar essas crenças sobre diferentes ativos pra construção de um portfólio que possa atingir o retorno esperado com o mínimo de risco.

De acordo com Markowitz (1999), é implícito nessa metodologia que uma maneira de minimizar a variabilidade dos retornos é através da diversificação de ativos em uma carteira, contando com um retorno esperado que possua um baixo valor de risco associado. Essa forma de lidar com o risco associado ao investimento foi criticada em Stoyanov et al. (2011), porém ainda se mostra eficaz para gerenciar o risco total do portfólio.

A metodologia da teoria moderna de portfólio explicada utilizando álgebra matricial é discutida na literatura sobre precificação de ativos e teoria de portfólio, como demonstrado em Cochrane (2008) e Constantinides e Malliaris (1995).

O desenvolvimento do trabalho foi feito utilizando o R no ambiente *Rstudio*. Foi usada a função `fPortfolio` para calcular os pesos para cada ativo dentro do portfólio, a função `PARMA` é usada para calcular a matriz de covariância associada aos pesos do portfólio estimado.

De acordo com Zivot (2021), assumindo que o investidor procura a melhor relação entre retorno esperado e risco, podemos caracterizar esse problema de duas maneiras. Da primeira maneira, os investidores procuram portfólios que maximizam o retorno esperado considerando um nível de risco, mensurado pela variância do portfólio.

Demonstrado com Ingersoll (1987) e Cochrane (2001), seja Y um conjunto de séries em log-retornos, i.e., $Y = \{y_{1,t}, y_{2,t}, \dots, y_{n,t}\}_{t=1}^T$, então o conjunto de médias é dado por $\mu = \{\mu_1, \dots, \mu_n\}$, sendo n o total de ativos, e a sua matriz de covariância é dada por Σ , tal que:

$$\Sigma = \begin{pmatrix} \text{Var}(y_1) & \dots & \text{Cov}(y_n, y_1) \\ \vdots & \ddots & \vdots \\ \text{Cov}(y_1, y_n) & \dots & \text{Var}(y_n) \end{pmatrix}$$

Assuma que σ_p represente o nível de risco esperado para um portfólio qualquer p e $\mathbf{1}$ seja um vetor $n \times 1$, onde n é o total de ativos no portfólio. O problema de maximização para encontrar o portfólio com a melhor relação entre risco, σ_p^2 , e retorno médio do portfólio, $\mu_p = \mathbf{w}'\boldsymbol{\mu}$ é dado por

$$\mathbf{w} = \arg \max_{\boldsymbol{\omega}} \boldsymbol{\omega}'\boldsymbol{\mu} \tag{2.1}$$

$$\text{sujeito a } \sigma_p^2 = \boldsymbol{\omega}'\Sigma\boldsymbol{\omega} \text{ e } \boldsymbol{\omega}'\mathbf{1} = 1.$$

A segunda maneira é minimizar o risco associado a um portfólio, considerando um retorno esperado. Dado μ_p o valor de retorno esperado pelo investidor, o problema de maximização para encontrar o portfólio com a melhor relação entre risco, $\sigma_{p,w} = \mathbf{w}'\Sigma\mathbf{w}$, e retorno, μ_p , será dado por

$$\mathbf{w} = \arg \min_{\boldsymbol{\omega}} \boldsymbol{\omega}'\Sigma\boldsymbol{\omega}$$

$$\text{sujeito a } \mu_p = \boldsymbol{\omega}'\boldsymbol{\mu} \text{ e } \boldsymbol{\omega}'\mathbf{1} = 1.$$

Na prática, o segundo problema de maximização é o mais utilizado. Isso se dá pelo fato de ser mais conveniente para um investidor especificar o valor de retornos de interesse do seu investimento ao invés de especificar um grau de risco, [Zivot \(2021\)](#).

Segundo a teoria de portfólio de Markowitz, todo portfólio que possui o valor máximo de retorno dado um grau de risco é chamado de eficiente, e o conjunto de portfólios que otimizam essa relação risco-retorno são representados pelo que [Markowitz \(1952\)](#) chamava de fronteira eficiente, que compõe todas as combinações ótimas entre risco e retorno, conforme a [Figura 2.1](#).

O gráfico da fronteira eficiente é criado pela relação entre o retorno esperado μ_p e o risco σ_p associados para todas as combinações possíveis de portfólios dado um conjunto de ativos. A linha contempla todos os portfólios considerados eficientes de acordo com Markowitz, já que possuem o retorno máximo dado os portfólios com aquele risco associado. Portfólios abaixo ou a direita da curva são considerados ineficientes, já que a sua taxa de retorno não é alta o suficiente para justificar o risco.

2.2 Índice de Sharpe

Conforme [Sharpe \(1998\)](#), o índice de sharpe é uma medida do excesso de rendimento por unidade de risco de um investimento, definido por

$$S = \frac{E[r - r^*]}{\sigma} \tag{2.2}$$

onde r é o retorno do investimento realizado, r^* é o retorno de um investimento de referência e σ representa o risco associado ao investimento realizado. $E[r - r^*]$

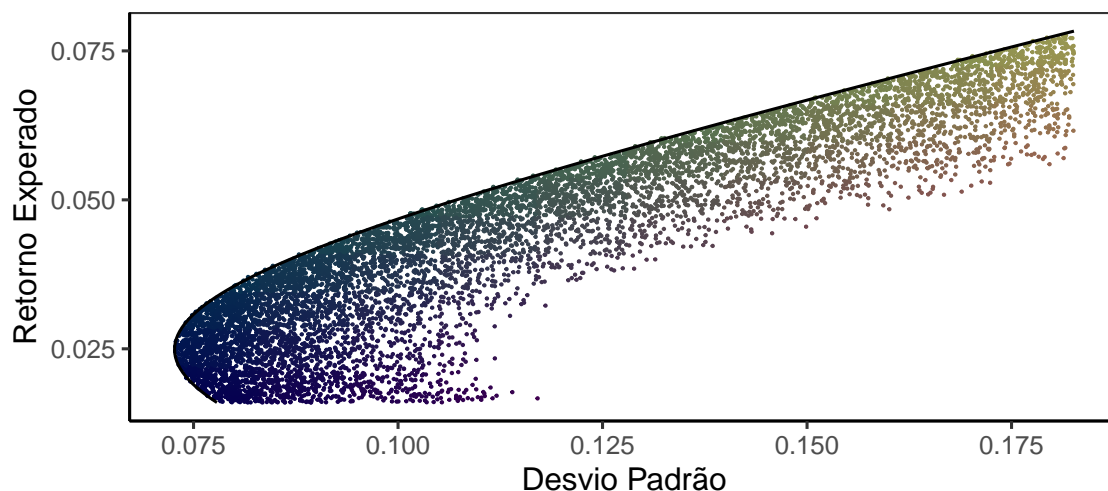


Figura 2.1: Fronteira eficiente.

representa o valor esperado do excesso de rendimento do investimento realizado em relação ao de referência. Subtraindo a taxa de retorno de um ativo livre de risco do retorno total do portfólio pode ilustrar ao investidor o retorno obtido em comparação a uma atividade de investimento livre de risco.

Considerando um ativo de baixo risco para o retorno de referência, podemos avaliar como positivos os retornos esperados dos nossos ativos investidos que tiverem um valor alto para o índice de sharpe.

De acordo com [Sharpe \(1966\)](#), o melhor portfólio será aquele que tangencia a fronteira eficiente, que é aquele para o qual a expressão (2.2) obtém o maior valor. Como visto na Figura 2.2, o ponto laranja representa o portfólio tangente, que possui o maior valor para o índice de sharpe. Para que mais de uma carteira seja eficiente, todas devem estar ao longo de uma linha comum e fornecer valores idênticos desse índice. Consoante com [Bailey e Lopez de Prado \(2012\)](#), ao definir uma Fronteira Eficiente com Índice de Sharpe estamos considerando um conjunto de portfólios que fornecem o maior excesso de retorno esperado em relação ao risco.

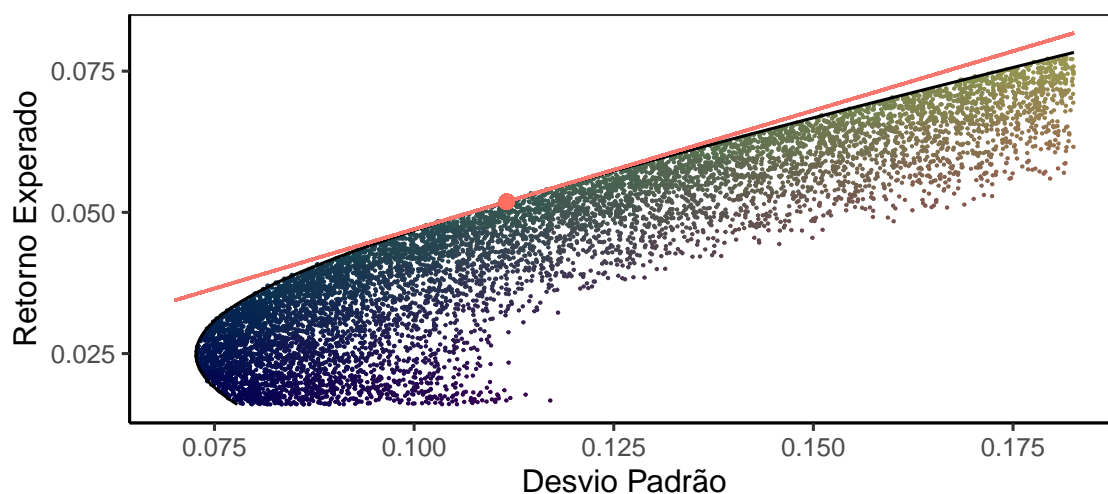


Figura 2.2: Fronteira eficiente e Linha Tangencial.

Dado um conjunto de ativos e um ativo livre de risco, o portfólio construído que maximiza o índice de sharpe é chamado de portfólio tangencial, [Ingersoll \(1987\)](#). Os pesos associados do portfólio tangencial são denotados por $\mathbf{t} = (t_1, \dots, t_n)'$, e resolve o seguinte problema de maximização restrita

$$\mathbf{t} = \arg \max_{\tau} \frac{\tau' \boldsymbol{\mu} - r^*}{(\tau' \boldsymbol{\Sigma} \tau)^{\frac{1}{2}}} \text{ tal que } \tau' \mathbf{1} = 1 \quad (2.3)$$

onde $\mu_{p,t} = \mathbf{t}' \boldsymbol{\mu}$ e $\sigma_{p,t} = (\mathbf{t}' \boldsymbol{\Sigma} \mathbf{t})^{\frac{1}{2}}$. Aqui vemos que a expressão (2.3) é uma combinação da equação (2.1) e (2.2). O Lagrangiano, veja [Zivot \(2021\)](#), associado ao problema é dado por:

$$L(\tau, \gamma) = (\tau' \boldsymbol{\mu} - r_f)(\tau' \boldsymbol{\Sigma} \tau)^{-\frac{1}{2}} + \gamma(\tau' \mathbf{1} - 1).$$

Através da regra da cadeia, as condições de primeira ordem são dadas por:

$$\begin{aligned} \frac{\partial L(\tau, \gamma)}{\partial \tau} &= \boldsymbol{\mu}(\tau' \boldsymbol{\Sigma} \tau)^{-\frac{1}{2}} - (\tau' \boldsymbol{\mu} - r_f)(\tau' \boldsymbol{\Sigma} \tau)^{-3/2} \boldsymbol{\Sigma} \tau + \gamma \mathbf{1} = \mathbf{0}, \\ \frac{\partial L(\tau, \gamma)}{\partial \gamma} &= \tau' \mathbf{1} - 1 = 0. \end{aligned}$$

Logo, o valor de τ que resolve o problema de otimização (2.3) é dado por:

$$\mathbf{t} = \frac{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r^* \cdot \mathbf{1})}{\mathbf{1}' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r^* \cdot \mathbf{1})}.$$

2.3 Low-pass Filter

Demonstrado em [Christiano e Fitzgerald \(2003\)](#), um “low-pass filter” é uma transformação linear dos dados que resulta em apenas os componentes principais dentro de uma faixa de frequência determinada, eliminando o restante de informação.

Conforme [Christiano e Fitzgerald \(1998\)](#), existe um método simples para extrair componentes de uma banda de frequência de uma série temporal. Dado a série temporal $\{y_t\}_{t=1}^n$, a mesma pode ser expressa por

$$y_t = \int_0^{\pi} [a(\omega) \cos(\omega t) + b(\omega) \sin(\omega t)] d\omega, \quad \forall t$$

ou como

$$y_t = a_0 + a_1 \cos(t\omega_1) + b_1 \sin(t\omega_1) + \dots + a_{T/2} \cos(t\omega_{T/2}) + b_{T/2} \sin(t\omega_{T/2}),$$

para $t = 1, \dots, T$, para T sendo o número total de observações da série temporal e onde

$$\omega_j = 2\pi j/T, \quad j = 1, \dots, T/2. \quad (2.4)$$

Aqui os ω_j representam a frequência de oscilação da série temporal. Logo, se desejamos filtrar a série $\{y_t\}_{t=1}^T$ permitindo apenas as frequências abaixo de ω_f , temos que para dado

$$X = \begin{bmatrix} 1 & \cos(\omega_1) & \sin(\omega_1) & \cdots & \cos(\omega_f) & \sin(\omega_f) \\ 1 & \cos(2\omega_1) & \sin(2\omega_1) & \cdots & \cos(2\omega_f) & \sin(2\omega_f) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \cos(T\omega_1) & \sin(T\omega_1) & \cdots & \cos(T\omega_f) & \sin(T\omega_f) \end{bmatrix}.$$

os coeficientes para o filtro de baixas frequências são obtidos via

$$\hat{\beta} = (X'X)^{-1}X'Y$$

onde $\beta = \{a_0, a_1, b_1, \dots, a_f, b_f\}'$. Assim a série filtrada, \hat{Y}_f , é estimada da seguinte forma

$$\hat{Y}_f = X\hat{\beta}.$$

Ademais, Müller e Watson (2020) propôs o estudo da variabilidade de longo prazo utilizando um método utilizando a filtragem das séries apenas por cossenos. Porém como foi utilizado apenas algumas poucas frequências e vários ativos nesse trabalho, o método de Müller e Watson (2020) não se demonstrou adequado a este estudo, pois resultava em problemas de $rank(\Sigma_f) < n$.

2.4 Dominância Estocástica

De acordo com Barrett e Donald (2003), considerando duas opções de investimento com suas respectivas funções de distribuição acumulada, dadas por G e F . Conforme a definição demonstrada em Deaton (1997), dominância estocástica de primeira ordem de G sobre F corresponde a

$$G(z) \leq F(z) \text{ para todo } z.$$

ou seja, o gráfico da distribuição de probabilidade acumulada $F(z)$ sempre está a esquerda da distribuição de probabilidade acumulada $G(z)$. Economicamente, isto significa que, para qualquer nível de probabilidade, o retorno esperado relacionado a $G(z)$ é maior que o retorno esperado relacionado a $F(z)$.

Por outro lado, dominância estocástica de segunda ordem de G sobre F corresponde a

$$\int_0^z G(t)dt \leq \int_0^z F(t)dt \text{ para todo } z.$$

Assim, segundo Mattos (1998), para um investidor avesso ao risco, o retorno esperado relacionado a $G(z)$ domina o relacionado a $F(z)$ se a área acumulada sob o gráfico da distribuição de probabilidade acumulada $F(z)$ for maior que a área acumulada de $G(z)$, dado qualquer valor de z .

Conforme Levy (1992), dentre as aplicações mais comuns de Dominância Estocástica no contexto econômico, as mais relevantes para o método proposto apresentado nesse estudo são as que estão associadas na mensuração e definição de risco relacionado a uma fonte de renda ou de um conjunto de ativos. Os cálculos de Dominância Estocástica foram feitos no software R utilizando as funções disponíveis no pacote `generalCorr`.

2.5 Drawdown

Conforme Chekhlov et al. (2005), Drawdown é um parâmetro de volatilidade que resume a queda do valor de um ativo em relação a sua cotação máxima mais recente. Assim, o investidor pode avaliar de forma percentual o quanto de seu investimento

inicial foi perdido, bem como quanto tempo levou até que a sua carteira se recuperasse dessa queda.

Dado um portfólio construído no intervalo de tempo $[0, T]$ e $r(t)$ o valor de retorno do portfólio no momento t onde $t \in [0, T]$. A avaliação feita na comparação de Drawdown do portfólio é expressado pela função

$$D = \max_{t_1, t_2} [r(t_1) - r(t_2)], \quad D_t = t_2 - t_1, \quad 0 \leq t_1 < t_2 \leq T \quad (2.5)$$

que na prática, D expressa a sequência de quedas ou de aumentos percentuais no valor de retorno de um portfólio. Temos também que D_t expressa o tempo t que o portfólio precisou apresentar ganhos para que compensasse a perda ocorrida em D .

Como descrito em [Cherny e Obloj \(2013\)](#), o uso de análise de Drawdowns em carteiras de longo prazo serve como métrica do risco associado a um conjunto de investimentos, bem como do quanto de retorno é preciso para que a carteira se recupere da perda anterior.

Ao comparar Drawdowns entre portfólios diferentes, estamos analisando qual deles possui um histórico de queda menor, o valor percentual dessa queda e o tempo de recuperação, conforme explicado em [Lopez de Prado \(2013\)](#). Assim, auxiliando a mensurar o risco de determinados investimentos.

2.6 Teste de Ledoit-Wolf

De acordo com [Ledoit e Wolf \(2008\)](#), sejam duas estratégias de investimento i e j cujo retorno adicional em comparação a um retorno de benchmark em um momento t são r_{ti} e r_{tj} respectivamente. O retorno de benchmark em questão é tipicamente dado pelo retorno de um ativo livre de risco, mas nesse caso pode também ser o ticker de algum outro ativo de referência.

Dado um total de T pares de retornos nos momentos t , dados por $(r_{1i}, r_{1j})', \dots, (r_{Ti}, r_{Tj})'$ sendo observados. É assumido que essas observações constituem uma série temporal estritamente estacionária, de modo que o a distribuição multivariada de retornos não se altera ao longo do tempo. Essa distribuição possui média μ e matriz de covariância Σ dada por

$$\mu = \begin{pmatrix} \mu_i \\ \mu_j \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} \sigma_i^2 & \sigma_{ij} \\ \sigma_{ij} & \sigma_j^2 \end{pmatrix}.$$

As médias e variâncias amostrais dos retornos observados são denotadas por $\hat{\mu}_i$, $\hat{\mu}_j$ e $\hat{\sigma}_i^2$, $\hat{\sigma}_j^2$, respectivamente. A diferença entre os índices de sharpe de i e j é dado por

$$\Delta = \frac{\mu_i}{\sigma_i} - \frac{\mu_j}{\sigma_j}$$

e a estimativa da diferença é dada por

$$\hat{\Delta} = \frac{\hat{\mu}_i}{\hat{\sigma}_i} - \frac{\hat{\mu}_j}{\hat{\sigma}_j}$$

Para mais, considerando $u = (\mu_i, \mu_j, \sigma_i^2, \sigma_j^2)'$ e $\hat{u} = (\hat{\mu}_i, \hat{\mu}_j, \hat{\sigma}_i^2, \hat{\sigma}_j^2)'$. [Mermel \(2003\)](#) calcula um erro para $\hat{\Delta}$ baseado na seguinte relação

$$\sqrt{T}(\hat{u} - u) \xrightarrow{d} N(0; \Omega),$$

onde \xrightarrow{d} se refere à convergência em distribuição e na aplicação do método delta. Porém, Ω se baseia no retorno i.i.d. de uma distribuição normal bivariada, assumindo

$$\Omega = \begin{pmatrix} \sigma_i^2 & \sigma_{in} & 0 & 0 \\ \sigma_{ij} & \sigma_n^2 & 0 & 0 \\ 0 & 0 & 2\sigma_i^4 & 2\sigma_{ij}^2 \\ 0 & 0 & 2\sigma_{ij}^2 & 2\sigma_j^4 \end{pmatrix}.$$

A fórmula acima não é válida caso a distribuição dos retornos não siga uma distribuição normal ou se as observações são correlacionadas ao longo do tempo. Para atender essa condição, [Ledoit e Wolf \(2008\)](#) sugere o uso de estimadores *Heteroskedasticity- and autocorrelation-consistent*, ou HAC, para a matriz de covariância, assim como o uso de técnicas de bootstrap.

2.7 Taxa de *Turnover*

Turnover de portfólio é uma taxa do quão rápido um título financeiro é vendido ou comprado, dado um determinado período de tempo. Essa taxa é importante para investidores interessados em um fundo de investimento, já que altas taxas de *turnover* também possuem pagamentos altos.

Conforme [Tucker \(2017\)](#), essa taxa serve de métrica para investidores terem ideia da quantidade de movimentações que estão ocorrendo em um portfólio de investimento, dentro de um período de tempo. A fórmula do *turnover ratio* para um período T , é dada por

$$\text{Turnover Ratio} = \frac{1}{T} \sum_{t=2}^T \sum_{j=1}^n |w_{j,t} - w_{j,t-1}| \quad (2.6)$$

onde $w_{j,t}$ é o peso do ativo j no momento t e n é o total de ativos.

3 Modelo Proposto

Neste capítulo descrevemos o modelo proposto neste trabalho. Tal modelo pode ser entendido como uma extensão a metodologia de Markowitz-Sharpe citada na Seção 2, em especial da equação (2.3). Para tal, a matriz de covariância é construída fazendo uso das séries filtradas conforme visto na Seção 2.3, pois uma vez removendo a covariabilidade de curto prazo dos retornos, é esperado que os pesos que constituem o portfólio sejam menos sensibilizado por variações rápidas e temporárias.

Então seja Y_f um conjunto de séries em log-retornos após a projeção em senos e cossenos, da frequência 0 até $\omega_f = 2\pi f/T$, i.e., $Y_f = \{y_{1,t}^f, \dots, y_{n,t}^f\}_{t=1}^T$, então Σ_f é a matriz de covariância entre os diferentes ativos de Y_f e n é o número de ativos no conjunto. Logo, a obtenção do portfólio com um ativo livre de risco e utilizando a matriz de covariância filtrada é dada por

$$\mathbf{t}_f = \arg \max_{\tau} \frac{\tau' \mu - r^*}{(\tau' \Sigma_f \tau)^{\frac{1}{2}}} \text{ tal que } \tau' \mathbf{1} = 1 \quad (3.1)$$

onde a equação (3.1) define o que chamamos de “Long-Run Markowitz” ou LRM.

Assim, a escolha dos pesos no portfólio de longo prazo, dado um ativo livre de risco, é obtido maximizando os pesos \mathbf{t}_f tal que a diferença entre o retorno esperado e do ativo sem risco seja a maior possível, ponderada pela matriz de covariância filtrada. É importante frisar que utilizar o valor esperado das séries filtradas não seria condizente pois estamos interessados em apenas filtrar as relações entre os ativos, eliminando os componentes de variação mútua de alta frequência.

Conforme demonstrado em Pástor e Stambaugh (2012), ativos como ações costumam apresentar uma volatilidade mais alta em horizontes maiores, então consoante com o estudo presente em Müller e Watson (2018), aqui, o “low-pass filter” é utilizado para reduzir essa variabilidade e garantir que o investidor possa ter mais segurança na sua carteira a longo prazo.

4 Fontes e Pré Processamento de Dados

Nesse capítulo abordaremos as fontes de dados utilizadas no desenvolvimento do trabalho, bem como as formas de tratamento de dados utilizadas. A versão do **R** utilizada foi a 4.1.2, com base em [R Core Team \(2020\)](#).

4.1 Aquisição de Dados

Os dados foram adquiridos pelo site [Yahoo Finance](#), que é a propriedade de mídia do Yahoo voltada para o mercado financeiro. O site conta com informações oficiais dos órgãos financeiros responsáveis pelas bolsas de valores internacionais, garantindo a credibilidade dos dados que fornece. O site é gratuito e a pesquisa de preços de ações dentro dele pode ser feita via integração de api utilizando a função `tq_get()`, do pacote `tidyquant` do R, de acordo com [Dancho e Vaughan \(2021\)](#). Para uso da função, basta fornecer uma lista de códigos de ações de interesse e um período que queremos visualizar o histórico de preços.

Os códigos utilizados foram as ações que já compuseram o índice Dow Jones Industrial average dentro do intervalo de 01/01/2000 até 01/01/2022. O índice foi criado em 1896 pelo editor do The Wall Street Journal e fundador do Dow Jones Company, Charles Dow. O cálculo do índice é baseado na cotação das ações de 30 das maiores e mais importantes empresas dos Estados Unidos, que são escolhidas pelos editores dos jornal financeiro norte-americano The Wall Street Journal. As empresas que compõe o índice são ocasionalmente substituídas para acompanhar as mudanças do mercado.

Dentro do intervalo de interesse de 01/01/2000 até 01/01/2022 o índice teve 15 alterações de seus componentes, com um total de 63 empresas fazendo parte do índice em algum momento nesse período, tal como apresentado na Tabela [4.1](#).

Tabela 4.1: Lista das empresas pertencentes índice Dow Jones entre 2000 e 2022.

| Industria | Código da ação | Adicionado em | Derrubado em | Readicionado em | Derrubado em | Fechada/Renomeada/Fusão |
|---|----------------|-----------------|--------------|-----------------|--------------|-------------------------|
| International Paper Company | IIPC | Apr 01 1901 | Apr 08 2004 | - | - | - |
| General Electric Company | GE | Sat Dec 07 1907 | Jun 26 2018 | - | - | - |
| General Motors Corporation | GMCO | Mar 16 1915 | Jun 08 2009 | - | - | - |
| AlliedSignal Incorporated/Honeywell International | ALD-HONB | Dec 07 1925 | Feb 19 2008 | - | - | Jan 27 2021 |
| Exxon Corporation/Exxon Mobil Corporation | XOM-EXXO | Sat Dec 01 1928 | Aug 31 2020 | - | - | Jan 27 2003 |
| International Business Machines Corporation | IBM | May 26 1932 | - | - | - | - |
| AT&T Corporation | T | Sat Mar 04 1939 | Apr 08 2004 | Sep 21 2005 | Mar 19 2015 | May 06 1991 |
| American Express Company | AXPB | Aug 30 1982 | - | - | - | - |
| Caterpillar Inc. | CAT | May 06 1991 | - | - | - | - |
| E.I. du Pont de Nemours & Company/DowDuPont Inc. | DDNB | Mar 17 1997 | Apr 02 2019 | - | - | Sep 01 2017 |
| Hewlett-Packard Company | HPQB | Mar 17 1997 | Sep 23 2013 | - | - | - |
| J.P. Morgan & Company/JPMorgan Chase & Co. | JPM | Mar 17 1997 | - | - | - | - |
| Johnson & Johnson | JNJ | Mar 17 1997 | - | - | - | - |
| Alcoa Inc. | AA | Nov 01 1999 | Sep 23 2013 | - | - | - |
| Citigroup Inc. | CTGP | Nov 01 1999 | Jun 08 2009 | - | - | - |
| Intel Corporation | ITLC | Nov 01 1999 | - | - | - | - |
| Microsoft Corporation | MSFT | Nov 01 1999 | - | - | - | - |
| SBC Communications Inc./AT | T Inc & T | Nov 01 1999 | Sep 21 2005 | - | - | - |
| The Home Depot, Inc. | HOME | Nov 01 1999 | - | - | - | - |
| Eastman Kodak Company | KODK | Jan 27 2003 | Apr 08 2004 | - | - | - |
| American International Group, Inc. | AIGB | Apr 08 2004 | Sep 22 2008 | - | - | - |
| Pfizer Inc. | PFIZ | Apr 08 2004 | Aug 31 2020 | - | - | - |
| Verizon Communications Inc. | VERZ | Apr 08 2004 | - | - | - | - |
| Bank of America Corporation | BOAC | Feb 19 2008 | Sep 23 2013 | - | - | - |
| Chevron Corporation | CVX | Feb 19 2008 | - | - | - | - |
| Kraft Foods Inc. | MDLZ | Sep 22 2008 | Sep 24 2012 | - | - | - |
| Cisco Systems, Inc. | CSCO | Jun 08 2009 | - | - | - | - |
| The Travelers Companies, Inc. | TRV | Jun 08 2009 | - | - | - | - |
| UnitedHealth Group Incorporated | UNH | Sep 24 2012 | - | - | - | - |
| Nike, Inc. | NKE | Sep 23 2013 | - | - | - | - |
| The Goldman Sachs Group, Inc. | GS | Sep 23 2013 | - | - | - | - |
| Visa Inc. | V | Sep 23 2013 | - | - | - | - |
| Apple Inc. | AAPL | Mar 19 2015 | - | - | - | - |
| Walgreens Boots Alliance, Inc. | WBA | Jun 26 2018 | - | - | - | - |
| Dow Inc. | DOW | Apr 02 2019 | - | - | - | - |
| Raytheon Technologies Corporation | RYTT | Apr 06 2020 | Aug 31 2020 | - | - | - |
| Amgen Inc. | AMGN | Aug 31 2020 | - | - | - | - |
| Honeywell International Inc. | HON | Aug 31 2020 | - | - | - | - |
| salesforce.com, inc. | CRM | Aug 31 2020 | - | - | - | - |
| McDonald's Corporation | MCD | Oct 30 1985 | - | - | - | - |
| Merck & Co., Inc. | MRK | Jun 29 1979 | - | - | - | - |
| Minnesota Mining & Manufacturing Company/3M Company | MMM | Aug 09 1976 | - | - | - | - |
| Philip Morris Companies Inc./Altria Group, Incorporated | PHMO - MOOO | Oct 30 1985 | Feb 19 2008 | - | - | Jan 27 2021 |
| The Boeing Company | BA | Mar 12 1987 | - | - | - | - |
| The Coca-Cola Company | KO | May 26 1932 | - | - | - | - |
| The Procter & Gamble Company | PG | Aug 15 1933 | - | - | - | - |
| The Walt Disney Company | DIS | Mar 17 1997 | - | - | - | - |
| United Technologies Corporation | UTX | Sat Mar 04 1939 | Apr 06 2020 | - | - | Aug 09 1976 |
| Wal-Mart Stores, Inc./Walmart Inc. | WMT | Mar 17 1997 | - | - | - | Feb 01 2018 |

Nota: O símbolo “-” representa que a empresa não saiu do índice.

Nesse trabalho foi utilizado o valor de fechamento ajustado ao invés do valor de fechamento. Em concordância com O'Neil e Ryan (2002), a principal vantagem do uso do preço de fechamento ajustado é que o mesmo não tem o seu valor afetado por movimentações como por exemplo o split de ações, diferente do preço de fechamento simples. Outra vantagem é que o preço de fechamento ajustado contempla os dividendos pagos, fornecendo uma informação mais precisa sobre valor de retorno de investimento. A crítica apresentada sobre o uso do preço de fechamento ajustado é que o mesmo não é sensibilizado por especulação e interesse público, fazendo com que o acompanhamento de seu valor histórico não reflita na relevância da empresa na época.

4.2 Períodos Atípicos

No intervalo analisado, temos a ocorrência de períodos marcantes na economia americana que resultaram em comportamentos atípicos no histórico de preços das ações das empresas selecionadas. A figura 4.1 mostra o acompanhamento de retornos do Dow Jones Industrial Average durante o intervalo analisado e durante os períodos atípicos, estes representados pelas faixas cinzas na figura.

O primeiro período foi a crise internacional financeira que ocorreu entre mar/2001 e nov/2001. De acordo com Hall et al. (2001), apesar do início da recessão ter mostrado início em março do ano, os ataques de 11 de setembro ao World Trade Center agravaram a situação econômica do ano. Antes do ataque, a queda econômica desse intervalo não estava sendo grave o suficiente para configurar um período de recessão.

O segundo período foi Dez/2007 até Jun/2009 e corresponde à crise imobiliária decorrente do aumento dos valores imobiliários, que não foi acompanhado por um aumento de renda da população. Conforme Stock e Watson (2012), nesse período, diversos bancos passaram a oferecer mais créditos, expandindo o crédito imobiliário e atraindo os consumidores, causando uma valorização dos imóveis. Com a alta procura, a taxa de juros subiu, derrubando os preços dos imóveis. Porém, como muitos desses empréstimos foram de alto risco, muitas pessoas não tinham como pagá-los e diversos bancos foram descapitalizados devido a alta na inadimplência.

O período atípico mais recente foi de Fev/2020 até Abr/2020, marcado pela pandemia de COVID-19. Descrito em Jomo e Chowdhury (2020), com os casos iniciais da doença serem surgidos da China, esta sendo uma potência econômica e um grande centro de produção industrial, o surto constituiu uma ameaça de desestabilização a economia global. O impacto na cadeia de fornecimento global e a queda no preço de petróleo devido a diminuição da procura por parte da China são exemplos dos vários efeitos surgidos na economia oriundos da doença.

4.3 Merge e Fechamento de Empresas nos Períodos

Dentro do período de 01/01/2000 até 01/01/2022, temos também a ocorrência de alguns códigos de ações que deixaram de apresentar histórico. Pesquisando os casos, vemos que algumas empresas fecharam ou foram adquiridas por outras empresas (processo também chamado de merge). Nesses casos, foi considerado o último preço



Figura 4.1: Dow Jones e períodos de recessão. Fonte: Yahoo Finance.

registrado no histórico da empresa, o que na prática corresponderia a venda da ação do portfólio.

Um exemplo disso é a empresa AlliedSignal Incorporated, que respondia pelo código de ação ALD e fazia parte do Dow Jones Industrial Average durante o período de 01/01/2000 até 27/01/2003. Tendo sido comprada em 1999 pela empresa Honeywell International, em 27/01/2003 foi substituída no Índice e passou a ser representada pelo ticket da sua empresa pai HONB.

5 Estudo Empírico

Para a avaliação do método proposto, foram considerados três horizontes de pesquisa para o investimento de longo prazo: 3, 6 e 12 meses. Cada um desses horizontes representa o retorno do valor investido no momento t , onde $t \in (1260, T)$.

Como dito na Seção 3, para a construção do conjunto de séries filtradas foi utilizado o método apresentado na Seção 2.3. Então, para cada momento t , 5 anos de dados são utilizados para a construção da matriz de covariância e vetor de retornos esperados. Assumindo que cada ano seja composto por 252 dias úteis, que nos dá 12 meses de 21 dias cada, são utilizados 1260 observações para a obtenção dos vetores de pesos dos portfólios. Foi utilizado a série *Interest Rates, Discount Rate for United States*, como ativo livre de risco para a composição do portfólio tangencial. A série original é anualizada dividindo pelos 252 dias úteis e convertida para ganhos diários, sendo a mesma obtida no FED St. Louis.

Para determinar a frequência de oscilação máxima permitida para a série que será filtrada, utilizamos a fórmula apresentada em (2.4), onde ω_f é a frequência de oscilação máxima permitida, T sendo o número total de observações e $j = 1, \dots, f$. Seja os valores de $f = \{15, 20, 30, 60\}$ escolhidos a relação com o filtro passa baixa é dada pela Tabela 5.1.

Tabela 5.1: Frequências utilizadas para o Low-pass Filter.

| f | ω_f | Low-pass Filter band |
|-----|------------|---|
| 15 | 0.0748 | De longo prazo até movimentos não inferiores a 84 dias ou 4 meses |
| 20 | 0.0997 | De longo prazo até movimentos não inferiores a 63 dias ou 3 meses |
| 30 | 0.1496 | De longo prazo até movimentos não inferiores a 42 dias ou 2 meses |
| 60 | 0.2992 | De longo prazo até movimentos não inferiores a 21 dias ou 1 mês |

Nota: Para um tamanho de amostra igual a $T = 1260$ e $\omega_f = 2\pi f/T$.

Podemos ver que diminuindo os valores para f e por consequência os de ω_f , vamos diminuindo cada vez mais a banda de frequência permitida. O intuito disso é ter cada vez menos volatilidades de curto prazo na nossa série de retornos.

O modelo proposto foi comparado com outros dois métodos de construção de portfólio. O primeiro deles, o método Naive, é caracterizado pela distribuição dos pesos de maneira uniforme para os ativos dentro da carteira, para qualquer tempo t . Sendo n o total de ativos, os pesos para o método Naive são dados por $w_i = 1/n$, $i = 1, \dots, n$. O outro método comparado foi de Markowitz-Sharpe ou Tangencial, conforme apresentado na Seção 2.2. Todos os modelos atuam apenas na opção

“comprado”, isto é, além da restrição que a soma dos pesos dos portfólios seja igual a um, cada peso precisa ser igual ou maior que zero.

Dado o retorno para o momento t , de um determinado método m e para um horizonte h descrito por $r_{t,m,h}$, as métricas de comparação consideradas foram:

$$\Lambda_{m,h} = \sum_{t=1}^T r_{t,m,h}, \quad \hat{\mu}_{m,h} = \frac{\Lambda_{m,h}}{T}, \quad \hat{\sigma}_{m,h}^2 = \frac{1}{T-1} \sum_{t=1}^T (r_{t,m,h} - \hat{\mu}_{m,h})^2, \quad (5.1)$$

e

$$I_m = \frac{1}{T} \sum_{t=1}^T I_{m,t}, \quad (5.2)$$

onde $I_{m,t}$ é igual a 1 se $r_{t,m,h} = \max\{r_{t,1,h}, \dots, r_{t,M,h}\}$, sendo M o total de métodos avaliados, e 0 caso contrário.

Temos que $\Lambda_{m,h}$ é o valor de retorno acumulado para o método m considerando o horizonte h , $\hat{\mu}_{m,h}$ é o valor de retorno médio para o método m , considerando o horizonte h e $\hat{\sigma}_{m,h}^2$ é a variância para o método m , considerando o horizonte h , D_t é o tempo que demora para um ativo se recuperar da sua maior queda, D , como descrito na equação (2.5).

Retornos trimestrais

Podemos ver pela Tabela 5.10 que o modelo Naive apesar de apresentar o maior percentual de I_m , também apresentou os menores valores para a média dos retornos, $\hat{\mu}_m$ e o retorno acumulado, Λ_m . O método Tangencial apresentou o segundo maior percentual de I_m , também possui o segundo maior desvio padrão. Mas possui o segundo menor valor de $\hat{\mu}_m$ e Λ_m .

Dentre as frequências para o “Long-Run Markowitz”, vemos que todas apresentam os menores valores de desvio padrão em relação aos modelos de benchmark, tal qual era o esperado. Nota-se também que, mesmo não tendo percentuais altos de I_m , todas as frequências apresentam retorno médio, $\hat{\mu}_m$, maior do que os modelos Naive e Tangencial. Isso mostra a robustez do método em relação a volatilidade, que apesar de não apresentar os maiores valores de retorno, em geral, frente ao Naive e o Markowitz, apresenta retornos altos de maneira mais consistente do que os demais.

Pela análise de Drawdown, vemos que o método Naive obteve os maiores valores de D e D_t da Tabela 5.10. O modelo LRM_{20} obteve o menor tempo de recuperação da sua maior queda, levando apenas 29 dias. Apesar de ambos os métodos de benchmark obterem os maiores valores para D_t , o valor de D para o método Tangencial ainda foi um dos menores.

As Tabelas 5.11 e 5.14 apresentam a dominância estocástica de primeira e segunda ordem respectivamente do métodos avaliados. Podemos ver que o modelo LRM_{15} domina estocasticamente os modelos Naive e Tangencial tanto em primeira ordem quanto em segunda. E tende a dominar os outros modelos LRM também em primeira e segunda ordem. Isso nos diz que a distribuição de LRM_{15} tende a gerar portfólios mais eficientes, pois sua distribuição está deslocada para mais a direita e mais restrita que os demais métodos.

A Figura 5.1 nos mostra o histograma dos retornos para cada um dos métodos propostos. Podemos ver que a distribuição dos valores de retorno dos modelos de

benchmark foi mais dispersa do que os métodos *LRM*. Dentre os métodos *LRM*, o que apresentou a maior concentração de retornos foi o *LRM*₆₀, ao mesmo tempo que apresentou o maior valor médio dos mesmos.

Pela tabela 5.5 do teste de Ledoit & Wolf, vemos que há diferença significativa entre a covariância de todos os métodos em relação ao Naive. Dos métodos *LRM*, apenas o *LRM*₆₀ apresentou diferença significativa em comparação ao método Tangencial.

Tabela 5.2: Comparativo descritivos dos resultados para $h = 3$.

| | $\hat{\mu}_m$ | $\hat{\sigma}_m$ | Λ_m | I_m | D | D_t |
|--------------------------|---------------|------------------|-------------|--------|-------|-------|
| Naive | 0.0186 | 0.0887 | 75.36 | 0.3335 | -2.47 | 1042 |
| Tangencial | 0.0279 | 0.0672 | 113.28 | 0.1746 | -0.66 | 363 |
| <i>LRM</i> ₁₅ | 0.0298 | 0.0625 | 120.98 | 0.1463 | -1.01 | 53 |
| <i>LRM</i> ₂₀ | 0.0298 | 0.0625 | 120.92 | 0.1224 | -1.50 | 29 |
| <i>LRM</i> ₃₀ | 0.0299 | 0.0614 | 121.29 | 0.1015 | -0.58 | 344 |
| <i>LRM</i> ₆₀ | 0.0300 | 0.0630 | 121.80 | 0.1217 | -0.59 | 350 |

Nota: Onde $\hat{\mu}_m$, $\hat{\sigma}_m$ e Λ_m representam a média dos retornos, o desvio padrão e o retorno acumulado usando cada método, ver equação (5.1). A definição de I_m é dado pela Equação (5.2). O valor D é dado pela equação (2.5) e D_t é o tempo de recuperação de queda associada a D . Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Tabela 5.3: Dominância Estocástica de primeira ordem para $h = 3$.

| | Naive | Tangencial | <i>LRM</i> ₁₅ | <i>LRM</i> ₂₀ | <i>LRM</i> ₃₀ | <i>LRM</i> ₆₀ |
|--------------------------|--------|------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Naive | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Tangencial | 1.0000 | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| <i>LRM</i> ₁₅ | 1.0000 | 1.0000 | - | 0.9299 | 0.7660 | 0.8165 |
| <i>LRM</i> ₂₀ | 1.0000 | 1.0000 | 0.0701 | - | 0.7043 | 0.8112 |
| <i>LRM</i> ₃₀ | 1.0000 | 1.0000 | 0.2340 | 0.2957 | - | 0.6723 |
| <i>LRM</i> ₆₀ | 1.0000 | 1.0000 | 0.1835 | 0.1888 | 0.3277 | - |

Nota: Leia-se que, o método descrito nas linhas domina estocasticamente o método apresentado nas colunas em x por cento dos casos. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Retornos Semestrais

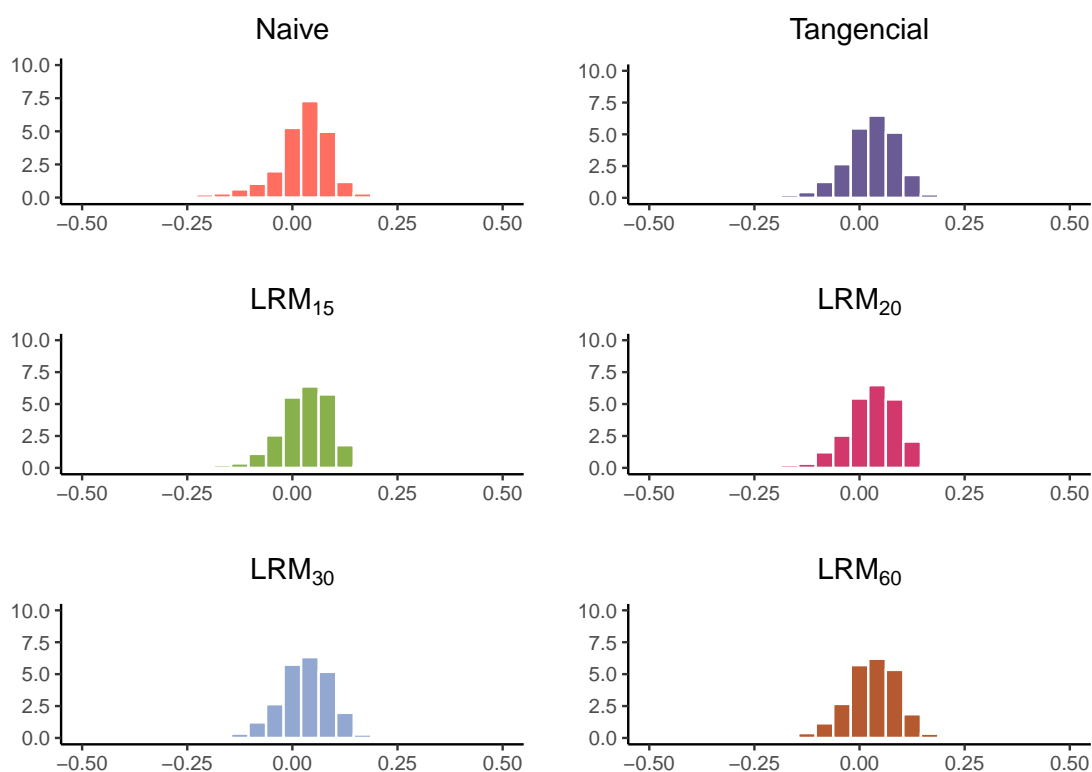
Semelhante a análise trimestral, a Tabela 5.6 nos mostra que o modelo Naive apresenta os menores valores para a média dos retornos, $\hat{\mu}_m$ e o retorno acumulado, Λ_m . O método Tangencial possui o segundo maior desvio padrão, mas possui o segundo menor valor de $\hat{\mu}_m$ e Λ_m .

Dentre os modelos *LRM*, vemos que todos apresentam os menores valores de desvio padrão em relação aos modelos de benchmark. Repara-se também que todas

Tabela 5.4: Dominância Estocástica de segunda ordem para $h = 3$.

| | Naive | Tangencial | LRM_{15} | LRM_{20} | LRM_{30} | LRM_{60} |
|------------|--------|------------|------------|------------|------------|------------|
| Naive | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Tangencial | 1.0000 | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| LRM_{15} | 1.0000 | 1.0000 | - | 0.9388 | 0.6773 | 0.7583 |
| LRM_{20} | 1.0000 | 1.0000 | 0.0612 | - | 0.6027 | 0.7002 |
| LRM_{30} | 1.0000 | 1.0000 | 0.3227 | 0.3973 | - | 0.9283 |
| LRM_{60} | 1.0000 | 1.0000 | 0.2417 | 0.2998 | 0.0717 | - |

Nota: Leia-se que, o método descrito nas linhas domina estocasticamente o método apresentado nas colunas em x por cento dos casos. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Figura 5.1: Histograma dos Retornos dos Portfólios para o $h = 3$ Meses.

as frequências apresentam retorno médio, $\hat{\mu}_m$, maior do que os modelos Naive e Markowitz. O retorno acumulado Λ_m para todos os modelos propostos foram maiores do que os de benchmark.

Pela análise de Drawdown, vemos que o método Naive obteve o menor valor para D e o maior valor para D_t , o que indica que esse método apresentou a maior queda percentual e a maior demora para se recuperar dessa queda. O modelo LRM_{20} obteve o menor tempo de recuperação da sua maior queda, essa sendo a maior queda dentre os modelos LRM .

Verifica-se pelas Tabelas 5.11 e 5.14 que o modelo LRM_{20} domina estocasticamente os modelos Naive e Tangencial tanto em primeira ordem quanto em segunda,

Tabela 5.5: P-valores para o teste de Ledoit & Wolf, $h = 3$.

| | Naive | Tangencial | LRM_{15} | LRM_{20} | LRM_{30} | LRM_{60} |
|------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Naive | - | 0.0365 | 0.0021 | 0.0039 | 0.0078 | 0.0095 |
| Tangencial | 0.0365 | - | 0.2087 | 0.1953 | 0.0637 | 0.0267 |
| LRM_{15} | 0.0021 | 0.2087 | - | 0.9978 | 0.7986 | 0.9915 |
| LRM_{20} | 0.0039 | 0.1953 | 0.9978 | - | 0.7794 | 0.9915 |
| LRM_{30} | 0.0078 | 0.0637 | 0.7986 | 0.7794 | - | 0.5610 |
| LRM_{60} | 0.0095 | 0.0267 | 0.9915 | 0.9915 | 0.5610 | - |

Nota: Veja Seção 2.6 sobre o teste de Ledoit & Wolf. Valores em negrito representam p-valor menor que $\alpha = 0.05$. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

e ainda tende a dominar os outros modelos LRM também em primeira e segunda ordem. Podemos ver que o modelo LRM_{15} tende a dominar estocasticamente todos os outros modelos, com exceção do LRM_{20} . Isso nos mostra que esse método tem condições de gerar portfólios eficientes tanto para o horizonte de $h = 6m$ quanto para $h = 3m$.

Pela Figura 5.2 podemos ver que a distribuição dos valores de retorno dos modelo Naive foi mais dispersa do que os outros métodos. Dentre os métodos LRM , o que apresentou valores de retorno maior, dado que está mais a direita do que os demais.

Vemos pela tabela 5.9 do teste de Ledoit & Wolf resultados semelhantes aos do horizonte $h = 3$, onde há diferença significativa apenas entre a covariância de todos os métodos em relação ao Naive, porém sem diferença significativa dos métodos LRM em comparação ao método Tangencial.

Tabela 5.6: Comparativo descritivos dos resultados para $h = 6m$.

| | $\hat{\mu}_m$ | $\hat{\sigma}_m$ | Λ_m | I_m | D | D_t |
|------------|---------------|------------------|-------------|--------|-------|-------|
| Naive | 0.0379 | 0.1359 | 154.04 | 0.3663 | -2.28 | 1022 |
| Tangencial | 0.0533 | 0.0914 | 216.56 | 0.1224 | -0.72 | 443 |
| LRM_{15} | 0.0573 | 0.0871 | 232.55 | 0.1197 | -0.93 | 491 |
| LRM_{20} | 0.0575 | 0.0879 | 233.26 | 0.1404 | -1.35 | 42 |
| LRM_{30} | 0.0581 | 0.0826 | 236.05 | 0.1064 | -0.65 | 345 |
| LRM_{60} | 0.0582 | 0.0854 | 236.42 | 0.1448 | -0.69 | 407 |

Nota: Onde $\hat{\mu}_m$, $\hat{\sigma}_m$ e Λ_m representam a média dos retornos, o desvio padrão e o retorno acumulado usando cada método, ver Equação (5.1). A definição de I_m é dado pela Equação (5.2). Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Retornos Anuais

Novamente, nota-se pela Tabela 5.10 que o modelo Naive apresenta o maior percentual de I_m , porém os menores valores para a média dos retornos, $\hat{\mu}_m$ e o retorno acumulado, Λ_m . Também apresenta o maior valor para $\hat{\sigma}_m$. O método Tangencial apresentou valores de desvio padrão próximos aos modelos LRM_{15} e LRM_{20} , mas ainda um $\hat{\mu}_m$ menor do que todos os modelos LRM .

Tabela 5.7: Dominância Estocástica de primeira ordem para $h = 6m$.

| | Naive | Tangencial | LRM_{15} | LRM_{20} | LRM_{30} | LRM_{60} |
|------------|--------|------------|------------|------------|------------|------------|
| Naive | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Tangencial | 1.0000 | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| LRM_{15} | 1.0000 | 1.0000 | - | 0.3877 | 0.9175 | 0.9330 |
| LRM_{20} | 1.0000 | 1.0000 | 0.6123 | - | 0.8977 | 0.8990 |
| LRM_{30} | 1.0000 | 1.0000 | 0.0825 | 0.1023 | - | 0.9144 |
| LRM_{60} | 1.0000 | 1.0000 | 0.0670 | 0.1010 | 0.0856 | - |

Nota: Leia-se que, o método descrito nas linhas domina estocasticamente o método apresentado nas colunas em x por cento dos casos. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Tabela 5.8: Dominância Estocástica de segunda ordem para $h = 6m$.

| | Naive | Tangencial | LRM_{15} | LRM_{20} | LRM_{30} | LRM_{60} |
|------------|--------|------------|------------|------------|------------|------------|
| Naive | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Tangencial | 1.0000 | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| LRM_{15} | 1.0000 | 1.0000 | - | 0.3404 | 0.8756 | 0.8942 |
| LRM_{20} | 1.0000 | 1.0000 | 0.6596 | - | 0.8325 | 0.8600 |
| LRM_{30} | 1.0000 | 1.0000 | 0.1244 | 0.1675 | - | 0.9999 |
| LRM_{60} | 1.0000 | 1.0000 | 0.1058 | 0.1400 | 0.0001 | - |

Nota: Leia-se que, o método descrito nas linhas domina estocasticamente o método apresentado nas colunas em x por cento dos casos. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Tabela 5.9: P-valores para o teste de Ledoit & Wolf, $h = 6$.

| | Naive | Tangencial | LRM_{15} | LRM_{20} | LRM_{30} | LRM_{60} |
|------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Naive | - | 0.0008 | 0.0010 | 0.0025 | 0.0001 | 0.0001 |
| Tangencial | 0.0008 | - | 0.3903 | 0.4542 | 0.0824 | 0.1011 |
| LRM_{15} | 0.0010 | 0.3903 | - | 0.8092 | 0.4109 | 0.6809 |
| LRM_{20} | 0.0025 | 0.4542 | 0.8092 | - | 0.3832 | 0.6405 |
| LRM_{30} | 0.0001 | 0.0824 | 0.4109 | 0.3832 | - | 0.1701 |
| LRM_{60} | 0.0001 | 0.1011 | 0.6809 | 0.6405 | 0.1701 | - |

Nota: Veja Seção 2.6 sobre o teste de Ledoit & Wolf. Valores em negrito representam p-valor menor que $\alpha = 0.05$. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Vemos que todos os modelos de LRM apresentam os maiores valores de retorno médio e retorno acumulado, Λ_m , em relação aos modelos de benchmark. Porém, apenas os modelos LRM_{30} e LRM_{60} apresentaram um valor de desvio padrão abaixo do modelo Tangencial, algo distinto dos resultados para $h \in \{3, 6\}$.

Pela análise de Drawdown, vemos que o modelo LRM_{30} apresentou um valor alto para D e baixo para D_t , o que indica que obteve uma queda percentual baixa e teve uma recuperação rápida dessa queda. Já o modelo Tangencial obteve uma

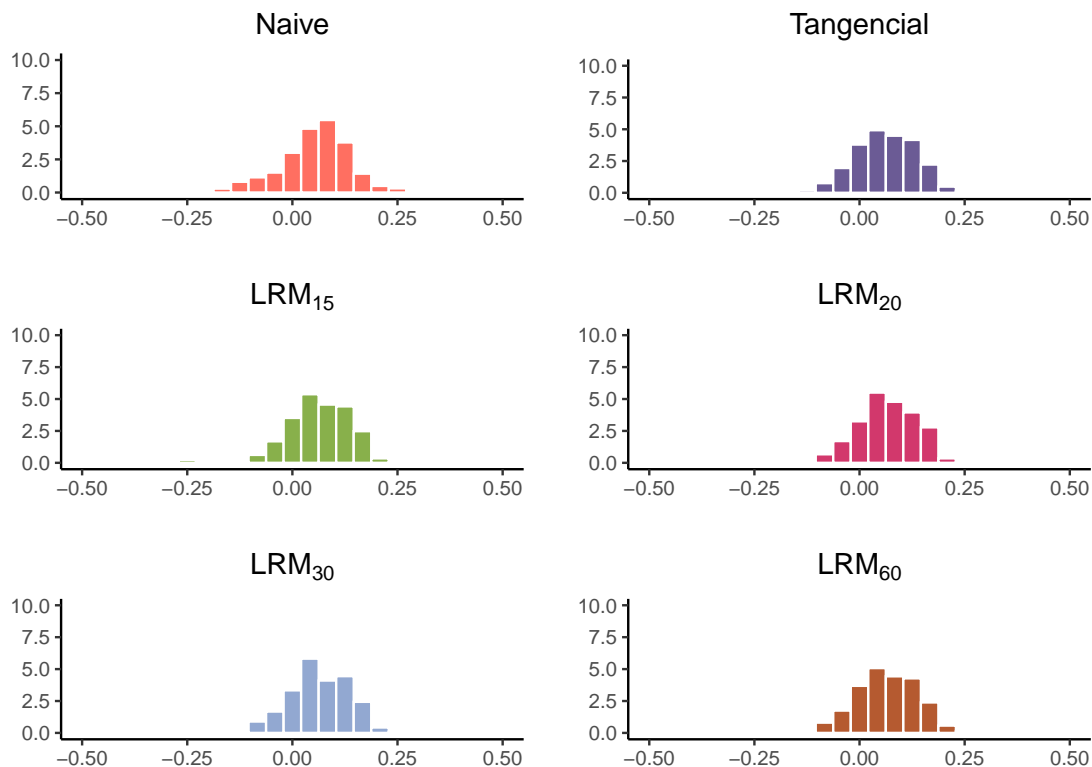


Figura 5.2: Histograma dos Retornos dos Portfólios para o $h = 6$ Meses.

performance melhor do que os métodos LRM_{15} e LRM_{20} .

Em contrapartida, segundo as Tabelas 5.11 e 5.14, os modelos LRM_{15} e LRM_{20} tendem a dominar estocasticamente os outros modelos tanto em primeira ordem quanto em segunda. Se compararmos com os resultados da Tabela 5.10, podemos dizer que alguns resultados dos modelos LRM_{15} e LRM_{20} foram extremamente negativos, mas que a maioria se concentrou a direita, conforme pode ser visto na Figura 5.3.

A tabela 5.13 do teste de Ledoit & Wolf nos mostra diferença significativa entre a covariância de todos os métodos em relação ao Naive, mas nenhuma diferença significativa entre os métodos LRM e o Tangencial. Entre os métodos LRM , vemos diferença significativa entre os métodos LRM_{15} , LRM_{20} e LRM_{30} .

Turnover Ratio

Como descrito na Seção 2.7, quanto maior a taxa de *Turnover* mais eficiente esse portfólio é em gerar pagamentos a partir dos seus ativos. Como podemos ver na tabela 5.14, para o horizonte $h = 3$ todos os métodos mostraram a mesma eficiência em pagamentos. Considerando $h = 6$, vemos a superioridade dos métodos LRM_{60} e LRM_{30} em relação aos demais, já que apresentaram as taxas maiores. Para o horizonte $h = 12$, todos os métodos apresentaram valores baixos. Porém, vemos que os métodos LRM , em particular os LRM_{15} e LRM_{30} apresentaram um desempenho maior que os demais.

Tabela 5.10: Comparativo descritivos dos resultados para $h = 12m$.

| | $\hat{\mu}_m$ | $\hat{\sigma}_m$ | Λ_m | I_m | D | D_t |
|------------|---------------|------------------|-------------|--------|-------|-------|
| Naive | 0.0774 | 0.1901 | 314.25 | 0.3005 | -2.18 | 1021 |
| Tangencial | 0.1104 | 0.1286 | 448.26 | 0.1576 | -0.82 | 465 |
| LRM_{15} | 0.1148 | 0.1331 | 465.89 | 0.1229 | -1.03 | 531 |
| LRM_{20} | 0.1159 | 0.1354 | 470.53 | 0.1502 | -1.13 | 501 |
| LRM_{30} | 0.1177 | 0.1211 | 478.04 | 0.1234 | -0.78 | 434 |
| LRM_{60} | 0.1183 | 0.1211 | 480.30 | 0.1453 | -0.80 | 464 |

Nota: Onde $\hat{\mu}_m$, $\hat{\sigma}_m$ e Λ_m representam a média dos retornos, o desvio padrão e o retorno acumulado usando cada método, ver Equação (5.1). A definição de I_m é dado pela Equação (5.2). Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Tabela 5.11: Dominância Estocástica de primeira ordem para $h = 12m$.

| | Naive | Tangencial | LRM_{15} | LRM_{20} | LRM_{30} | LRM_{60} |
|------------|--------|------------|------------|------------|------------|------------|
| Naive | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Tangencial | 1.0000 | - | 0.0817 | 0.0837 | 0.0000 | 0.0000 |
| LRM_{15} | 1.0000 | 0.9183 | - | 0.5015 | 0.6621 | 0.8372 |
| LRM_{20} | 1.0000 | 0.9163 | 0.4985 | - | 0.6512 | 0.7746 |
| LRM_{30} | 1.0000 | 1.0000 | 0.3379 | 0.3488 | - | 0.9043 |
| LRM_{60} | 1.0000 | 1.0000 | 0.1628 | 0.2254 | 0.0957 | - |

Nota: Leia-se que, o método descrito nas linhas domina estocasticamente o método apresentado nas colunas em x por cento dos casos. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Tabela 5.12: Dominância Estocástica de segunda ordem para $h = 12m$.

| | Naive | Tangencial | LRM_{15} | LRM_{20} | LRM_{30} | LRM_{60} |
|------------|--------|------------|------------|------------|------------|------------|
| Naive | - | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Tangencial | 1.0000 | - | 0.1106 | 0.1240 | 0.0000 | 0.0000 |
| LRM_{15} | 1.0000 | 0.8894 | - | 0.6252 | 0.5424 | 0.7399 |
| LRM_{20} | 1.0000 | 0.8760 | 0.3748 | - | 0.5054 | 0.6933 |
| LRM_{30} | 1.0000 | 1.0000 | 0.4576 | 0.4946 | - | 0.9948 |
| LRM_{60} | 1.0000 | 1.0000 | 0.2601 | 0.3067 | 0.0052 | - |

Nota: Leia-se que, o método descrito nas linhas domina estocasticamente o método apresentado nas colunas em x por cento dos casos. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Discussões Adicionais

De maneira geral, para os três horizontes propostos vemos que o efeito da redução da variabilidade via “low-pass filter” foi atingido, já que os métodos LRM tendem a apresentar os menores valores de desvio padrão em comparação aos modelos de benchmark. Sendo os modelos LRM_{15} e LRM_{20} os que apresentaram resultados mais satisfatórios, considerando todos os horizontes.

Tabela 5.13: P-valores para o teste de Ledoit & Wolf, $h = 12$.

| | Naive | Tangencial | LRM_{15} | LRM_{20} | LRM_{30} | LRM_{60} |
|------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Naive | - | 0.0000 | 0.0009 | 0.0019 | 0.0000 | 0.0000 |
| Tangencial | 0.0000 | - | 0.9629 | 0.9778 | 0.1483 | 0.2746 |
| LRM_{15} | 0.0009 | 0.9629 | - | 0.7400 | 0.0265 | 0.1876 |
| LRM_{20} | 0.0019 | 0.9778 | 0.7400 | - | 0.0081 | 0.1206 |
| LRM_{30} | 0.0000 | 0.1483 | 0.0265 | 0.0081 | - | 0.0628 |
| LRM_{60} | 0.0000 | 0.2746 | 0.1876 | 0.1206 | 0.0628 | - |

Nota: Veja Seção 2.6 sobre o teste de Ledoit & Wolf. Valores em negrito representam p-valor menor que $\alpha = 0.05$. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Tabela 5.14: *Turnover Ratio* para os três horizontes.

| | 3m | 6m | 12m |
|------------|--------|--------|--------|
| Naive | - | 0.0000 | 0.0001 |
| Tangencial | 1.0000 | - | 0.0584 |
| LRM_{15} | 1.0000 | 0.8894 | 0.0674 |
| LRM_{20} | 1.0000 | 0.8760 | 0.0592 |
| LRM_{30} | 1.0000 | 1.0000 | 0.0632 |
| LRM_{60} | 1.0000 | 1.0000 | 0.0585 |

Nota: Veja na Seção 5 sobre o cálculo do *Turnover Ratio*. Resultados para o intervalo entre 03/01/2005 e 19/02/2021.

Para $h = 12m$, o modelo Tangencial obteve melhores valores de $\hat{\mu}_m$, $\hat{\sigma}_m$, D e D_t que alguns dos modelos LRM . O que não invalida o modelo aqui proposto, mas sugere que, para horizontes mais longos, estudos adicionais são necessários.

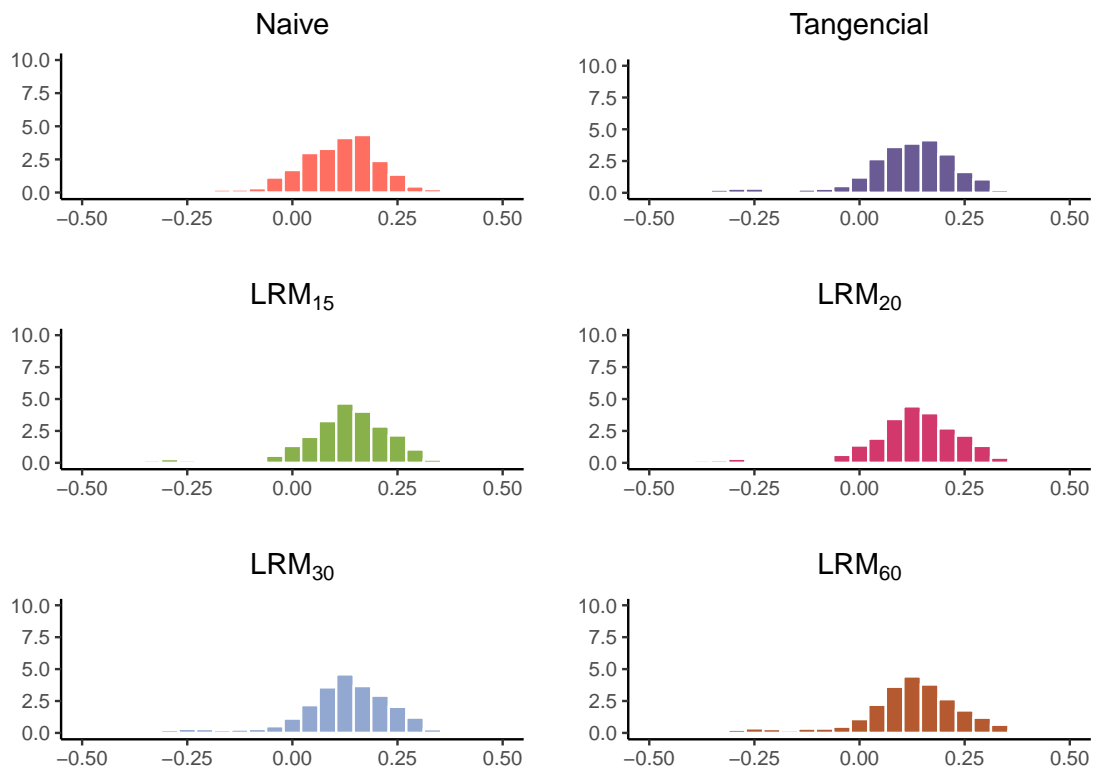


Figura 5.3: Histograma dos Retornos dos Portfólios para o $h = 12$ Meses.

6 Conclusão

Neste estudo foi proposto uma alternativa ao modelo de Markowitz-Sharpe, via a redução da variabilidade dos retornos através da filtragem de flutuações de curto prazo. Como não há uma definição do que caracteriza um investimento de longo-prazo ou de qual o risco ideal a se considerar ao construir uma carteira de investimento, nesse estudo foi utilizado horizontes de 3, 6 e 12 meses, e a taxa básica de juros dos EUA como ativo livre de risco.

Modelos de construção de investimento que consideram risco e variabilidade como fatores principais, podem se beneficiar com a aplicação do modelo proposto dado que algumas das variabilidades do histórico de ações podem se dar por conta de flutuações econômicas que não afetam o retorno a longo prazo.

Referências Bibliográficas

- Adrian, T. e Rosenberg, J. (2008). Stock returns and volatility: Pricing the short-run and long-run components of market risk. *The journal of Finance*, 63(6):2997–3030.
- Bailey, D. H. e Lopez de Prado, M. (2012). The sharpe ratio efficient frontier. *Journal of Risk*, 15(2):13.
- Barrett, G. F. e Donald, S. G. (2003). Consistent tests for stochastic dominance. *Econometrica*, 71(1):71–104.
- Chekhlov, A., Uryasev, S., e Zabarankin, M. (2005). Drawdown measure in portfolio optimization. *International Journal of Theoretical and Applied Finance*, 8(01):13–58.
- Cherny, V. e Obloj, J. (2013). Optimal portfolios of a long-term investor with floor or drawdown constraints. *arXiv preprint arXiv:1305.6831*.
- Christiano, L. J. e Fitzgerald, T. J. (1998). The business cycle: it’s still a puzzle. *Economic Perspectives-Federal Reserve Bank Of Chicago*, 22:56–83.
- Christiano, L. J. e Fitzgerald, T. J. (2003). The band pass filter. *International Economic Review*, 44(2):435–465.
- Cochrane, J. H. (2001). *Asset pricing*. Princeton University Press, Princeton, NJ.
- Cochrane, J. H. (2008). A mean-variance benchmark for intertemporal portfolio theory. *Manuscript, University of Chicago*.
- Cochrane, J. H. (2021). Portfolios for long-term investors. *SSRN Electronic Journal*.
- Constantinides, G. M. e Malliaris, A. G. (1995). Portfolio theory. *Handbooks in operations research and management science*, 9:1–30.
- Dancho, M. e Vaughan, D. (2021). *tidyquant: Tidy Quantitative Financial Analysis*. R package version 1.0.3.
- Deaton, A. (1997). *The analysis of household surveys: a microeconomic approach to development policy*. World Bank Publications.
- Engle, R. F. e Lee, G. (1999). A long-run and short-run component model of stock return volatility. *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, pages 475–497.

- Hall, R., Feldstein, M., Bernanke, B., Frankel, J., Gordon, R., e Zarnowitz, V. (2001). The business-cycle peak of march 2001. *Business Cycle Dating Committee, National Bureau of Economic Research*.
- Ingersoll, J. E. (1987). *Theory of financial decision making*. Rowman and Littlefield Studies in Financial Economics. Rowman & Littlefield, Lanham, MD.
- Jomo, K. e Chowdhury, A. (2020). Covid-19 pandemic recession and recovery. *Development*, 63(2):226–237.
- Ledoit, O. e Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.
- Levy, H. (1992). Stochastic dominance and expected utility: Survey and analysis. *Management Science*, 38(4):555–593.
- Lopez de Prado, M. (2013). How long does it take to recover from a drawdown? *Available at SSRN 2254668*.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.
- Markowitz, H. M. (1999). The early history of portfolio theory: 1600–1960. *Financial analysts journal*, 55(4):5–16.
- Mattos, A. E. K. d. (1998). O uso do downside risk (media - momentos parciais) como medida de risco na seleção de portfólios. Master's thesis, Escola de Administração/UFRGS.
- Müller, U. K. e Watson, M. W. (2018). Long-run covariability. *Econometrica*, 86(3):775–804.
- Müller, U. K. e Watson, M. W. (2020). Low-frequency analysis of economic time series. *preparation. In Handbook of Econometrics. Elsevier*.
- O'Neil, W. J. e Ryan, C. (2002). *How to make money in stocks: A winning system in good times or bad*. McGraw-Hill New York.
- Pastor, L. e Stambaugh, R. F. (2012). *Are Stocks Really Less Volatile in the Long Run?*
- Pástor, L. e Stambaugh, R. F. (2012). Are stocks really less volatile in the long run? *The Journal of Finance*, 67(2):431–478.
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1):119–138.
- Sharpe, W. F. (1998). The sharpe ratio. *Streetwise—the Best of the Journal of Portfolio Management*, pages 169–185.
- Stock, J. H. e Watson, M. W. (2012). Disentangling the channels of the 2007-2009 recession. Technical report, National Bureau of Economic Research.

- Stoyanov, S. V., Rachev, S. T., Racheva-Yotova, B., e Fabozzi, F. J. (2011). Fat-tailed models for risk estimation. *The Journal of Portfolio Management*, 37(2):107–117.
- Tucker, A. M. (2017). The long and the short: Portfolio turnover ratios & mutual fund investment time horizons. *J. Corp. L.*, 43:581.
- Zivot, E. (2021). Introduction to computational finance and financial econometrics with r. <https://bookdown.org/compfinezbook/introcompfinr/>. Acessado em 12/04/2022.