

Dissertação de Mestrado
Análise Técnica - Sorte ou Realidade?

Pedro Alberto Chauffaille Saffi ¹
EPGE - Fundação Getulio Vargas

Orientador: Marco Bonomo (EPGE / FGV-RJ)
Co-orientador: Marcelo Fernandes (EPGE / FGV-RJ)

¹saffi@fgvmail.br

Resumo

Este trabalho busca testar a validade da hipótese de eficiência dos mercados no mercado futuro do índice Ibovespa através do uso das chamadas estratégias de análise técnica. São utilizados testes de habilidade preditiva para verificar a hipótese de superioridade destas regras de decisão como forma de investimento. Estes testes possuem a vantagem de considerar a possibilidade de *data-snooping* na escolha da melhor estratégia, permitindo identificar se a aparente capacidade preditiva destes modelos é realmente significativa ou mero produto do acaso.

Os resultados indicam que as estratégias de análise técnica não são capazes de gerar retornos estatisticamente significativos quando os efeitos de *data-snooping* são levados em conta. Estes resultados estão de acordo com o previsto pela hipótese fraca de eficiência de mercado.

Sumário

1	Introdução	1
1.1	Eficiência de Mercado	2
1.2	<i>Data-Snooping</i>	3
1.3	Análise Técnica vs. Eficiência de Mercado	4
1.3.1	Brock <i>et al.</i> (1992)	4
1.3.2	Leal e Varanda (2000)	6
1.3.3	Sullivan, Timmermann e White (1997)	6
2	Metodologia	7
2.1	Dados	7
2.2	Estratégias de Análise Técnica (Elder, 1993)	9
2.2.1	Médias Móveis	10
2.2.2	Índice de Força Relativa (RSI)	11
2.2.3	Indicador William's %R	12
2.2.4	Estocástico	13
2.2.5	Indicador MACD	13
2.3	Técnica <i>Bootstrap</i>	14
2.3.1	<i>Bootstrap</i> em Blocos	15
2.4	Testes de Habilidade Preditiva	16
2.4.1	Teste de Realidade (White 2000)	16
2.4.2	Distribuição Assintótica	18
2.4.3	Teste de Habilidade Preditiva Superior	19
3	Resultados Empíricos	21
3.1	Testes Iniciais	21
3.2	P-Valores Nominiais	22
3.3	Teste de Realidade e SPA	24
4	Conclusão	26
5	Referências	26

“October. This is one of the peculiarly dangerous months to speculate in stocks. The others are July, January, September, April, November, May, March, June, December, August and February”.

Mark Twain (1894)

“Markets look a lot less efficient from the banks of the Hudson than from the banks of the Charles”.

Fischer Black

Agradecimentos

A conclusão desta dissertação só foi possível com a ajuda de vários amigos. Agradeço especialmente aos amigos (e nas horas vagas meus professores) Marco Bonomo e Marcelo Fernandes, por todo estímulo intelectual e confiança depositados desde o início. Aos professores Rene Garcia, pela conversa que deu início à minha exploração do tema e a Fernando de Holanda Barbosa e Pedro Cavalcanti Ferreira, por toda ajuda na continuação de meus estudos.

Não posso deixar de agradecer também à minha namorada, Bárbara Boavista, que com o incansável apoio e carinho, tornou estes dois últimos anos ainda mais inesquecíveis. Importantes também são os amigos de toda hora: Carlos Marinho, Guilherme Rocha, Marcelo Pessoa, Bernardo Sá Mota e Aldo Ramos, pela inestimável ajuda no aprimoramento do texto. Por último, agradeço a meus pais, Pedro e Laura, e à minha irmã, Vivian, sempre presentes em todos os momentos especiais.

1 Introdução

A discussão sobre a eficiência dos mercados é um dos assuntos que mais controvérsia gerou (e ainda gera) entre os economistas, sendo até hoje alvo da atenção da comunidade acadêmica.

Dentre as diversas estratégias de investimento à disposição dos agentes, a análise técnica utiliza informações passadas sobre os preços para definir decisões de investimento. Este tipo de estratégia é utilizada por vários autores (Fama, 1966; Sweeney 1988), em busca de evidências sobre sua performance superior na geração de retornos superiores à estratégia de referência “comprar e esperar”.¹ Estas evidências por sua vez implicariam na rejeição da eficiência sob a forma fraca, resultados encontrados mesmo após considerarmos os custos de transação (Levich e Thomas, 1993).

Uma das críticas feitas a estes trabalhos é que suas conclusões baseiam-se na aplicação das estratégias de análise técnica somente a uma realização do processo estocástico do preço do ativo: a série observada dos preços. Estaríamos portanto sujeitos aos problemas relacionados a *data-snooping*, já que é muito provável que, dentre milhares de estratégias diferentes, alguma (ou algumas) apresente performance superior (em termos de excesso de retorno). Isto poderia ter ocorrido simplesmente ao acaso e não devido à uma capacidade preditiva superior da estratégia.

Tentando minimizar este problema, trabalhos posteriores (Brock *et al.*, 1992; Levich e Thomas, 1993; Leal e Varanda, 2000) utilizaram a técnica de *bootstrap* para simular realizações alternativas da série dos preços. Esta técnica permite, sob algumas condições, criar sequências de preços que tenham propriedades similares às da série real em termos de média, dependência serial, etc.... Desta forma, a técnica permite construir a distribuição empírica da performance de uma estratégia de análise técnica particular e assim, verificamos se a sua performance estatisticamente significativa é realmente devido à uma capacidade preditiva superior. Caso contrário, os resultados seriam um mero produto do acaso.

O objetivo deste trabalho é testar a eficiência do mercado brasileiro através da rentabilidade das estratégias de análise técnica. A metodologia utilizada compreende os testes de análise preditiva (White, 2000; Hansen 2001), que utilizam a técnica de *bootstrap* para construir a distribuição empírica dos retornos no mercado futuro do índice Bovespa,

¹Todos os cálculos do retorno das estratégias estão em termos de excesso de rentabilidade em relação ao modelo de referência.

permitindo o teste de 14.630 estratégias para decisão de investimento.

1.1 Eficiência de Mercado

As origens da hipótese de eficiência de mercado (EMH) podem ser traçadas desde o início do século XX com os trabalhos de Bachelier (1900), Cowles (1933) e Cowles e Jones (1937). Podemos resumir esta hipótese através da seguinte frase (Fama, 1970): Um mercado onde os preços sempre refletem as informações disponíveis é dito eficiente. Essa afirmação, contudo, é muito geral para verificação empírica e necessita de maior profundidade. Assim, quando dizemos que um mercado é eficiente, devemos especificar também qual o conjunto de informações disponível aos agentes. Neste caso os preços não se alteram se estas informações são disponibilizadas a todos. A classificação usual dos conjuntos de informação (Roberts, 1967) os distingue da seguinte forma:

1. Eficiência Fraca: O conjunto de informação inclui somente a história passada dos preços.
2. Eficiência Semi-Forte: O conjunto de informação inclui toda informação conhecida por todos os agentes, i.é., toda informação pública disponível.
3. Eficiência Forte: O conjunto de informação inclui toda informação conhecida por *qualquer* agente, i.é., toda informação pública e privada disponível.

Além da definição do conjunto de informação relevante, todo teste da EMH deve estabelecer também um modelo de apreçamento dos ativos. Desta forma, todo teste empírico sobre a EMH é um teste de hipótese conjunto, onde não só a EMH mas também o modelo de apreçamento assumido são testados.

A análise da eficiência de mercado através das estratégias de análise técnica é um teste da forma fraca da EMH. Somente a história passada dos preços (sejam eles os de abertura, máximo ou fechamento) é considerada pelas estratégias. Assim, caso a EMH seja válida, estas estratégias não podem produzir retornos estatisticamente superiores à estratégia “comprar e esperar ”.²

²Nesta estratégia, o ativo é comprado e carregado na carteira do agente até o final do período.

1.2 *Data-Snooping*

Os problemas relacionados a *data-snooping* referem-se ao viés resultante na inferência estatística do uso das informações de um determinado conjunto de dados, quando utilizados para guiar pesquisas futuras que utilizem a mesma amostra inicial (Campbell *et al.*, 1997). A natureza dos processos econômicos torna este tipo de viés quase impossível de ser evitado em trabalhos empíricos, já que normalmente não é possível refazer o experimento e obter um novo conjunto de dados. Desta forma, quanto maior o número de trabalhos utilizando a mesma base de dados, maior a probabilidade de identificação de padrões espúrios na amostra.

Os retornos de ativos financeiros estão entre as séries mais estudadas pelos economistas, sendo especialmente suscetíveis à este tipo de problema. O trabalho de Lo e Mackinlay (1990) é o primeiro a exibir a magnitude deste viés estatístico, examinando o modelo de fatores (APT) através da análise do impacto de determinadas características sobre o retorno de carteiras de ações, dentre as quais o valor de mercado, a relação preço/lucro, etc... Estas características muitas vezes não são escolhidas através de modelos teóricos, mas sim através da análise empírica de suas correlações com o erro de apreçamento do modelo. Os resultados deste trabalho indicam que as evidências de rejeição do modelo são completamente eliminadas quando as estatísticas de teste são corrigidas para considerar *data-snooping*. Este trabalho é importante ao mostrar que os resultados de testes de significância, cujas estatísticas são influenciadas por relações empíricas derivadas da mesma amostra, podem ser drasticamente distorcidos quando a distribuição da estatística de teste não leva em conta *data-snooping*.

Conforme mencionado anteriormente, este tipo de problema encaixa-se perfeitamente no contexto dos testes de performance da análise técnica para validar a EMH. Como a maioria dos trabalhos testa a validade da eficiência de mercado para os principais índice de mercado dos EUA, a possibilidade de diversos padrões identificados na literatura serem espúrios é real.

As sugestões para lidar com este tipo de problema (e seus defeitos) são resumidas por Sullivan *et al.*(1999a):

1. Aguardar o surgimento de novas observações da mesma série utilizada.

Infelizmente, esta solução é inviável em muitos casos, como por exemplo em testes que utilizem dados anuais do consumo agregado.

2. Aplicar os testes à séries similares provenientes de outras fontes, tais como o índice

Dow Jones ou o índice Bovespa. Podemos também efetuar os testes em sub-amostras dos dados.

A observação de um mesmo padrão em séries similares diminui a chance de que os resultados obtidos tenham surgido ao acaso. Porém, as relações encontradas podem ser dependentes das características institucionais de um mercado, impossibilitando o uso de séries provenientes de mercados diferentes para testes de hipótese. Na mesma linha, o teste em sub-amostras dos dados originais permite encontrar resultados mais robustos para os testes.

3. Divulgar todos os modelos investigados no estudo.

Esta solução está ligada à um possível comprometimento da análise devido à inspeção de um grande número de modelos. Isto pode ser diminuído com a divulgação de todos os modelos considerados no trabalho e não somente àqueles que apresentaram resultados “interessantes”, isto é, evidências contrárias à teoria. Esta sugestão porém, não leva em conta a possibilidade de *data-snooping* entre pesquisadores diferentes. É possível que diversos autores apresentem resultados espúrios provenientes de testes de modelos correlacionados, aplicados aos mesmo conjunto de dados (ou pelo menos fortemente correlacionados). Este acúmulo de evidências, aumenta ainda mais o interesse por estas evidências, guiando a pesquisa futura para resultados que podem ser produtos do acaso.

1.3 Análise Técnica vs. Eficiência de Mercado

A literatura inicial sobre a utilidade das estratégias de análise técnica (Fama e Blume, 1966) mostra que este tipo de decisão de investimento não é útil para os investidores como decisão de investimento. Trabalhos posteriores porém, encontram evidências de sua utilidade mesmo após considerar custos de transação (Sweeney, 1988). O primeiro artigo a utilizar a técnica *bootstrap* para testar as estratégias de análise técnica é apresentado por Brock *et al.* (1992).

1.3.1 Brock *et al.* (1992)

Neste artigo, os autores investigam a aplicação de 26 estratégias no índice Dow Jones (DJIA), no período de 1897 a 1996 sem considerar custos de transação. A construção dos

retornos é feita sob diversas hipóteses para o processo gerador dos dados (DGP): passeio aleatório, $AR(1)$, $GARCH$, $GARCH-M$ e $E-GARCH$.

A geração das séries reamostradas do DJIA é feita da seguinte forma. Primeiro, faz-se a estimação dos retornos segundo cada modelo. Os resíduos deste modelo são reamostrados e recombinaados com as estimativas. Isto produz um novo conjunto de retornos que nos permitem, finalmente, encontrar uma série artificial para o DJIA. Repetindo o processo 500 vezes, através do sorteio aleatório com reposição dos resíduos, podemos aplicar as estratégias a cada uma destas séries reamostradas, obtendo a distribuição empírica de seus excessos de retorno. À partir desta distribuição, podemos testar a hipótese de utilidade da análise técnica, comparando a significância do retorno da estratégia aplicada à série real do DJIA com aqueles obtidos nas séries artificiais.

O primeiro modelo testado é o passeio aleatório padrão, onde todas as estratégias exibem retornos estatisticamente significativos. Estes resultados porém, podem ocorrer somente devido a um maior risco inerente a estas estratégias. Isto leva os autores a testar posteriormente se a variância da estratégia é estatisticamente menor na série real do DJIA do que nas séries reamostradas. Surpreendentemente, as estratégias não só exibem retornos médios maiores como também variâncias menores, mostrando que mesmo ponderadas pelo risco, as estratégias de análise técnica têm utilidade.

Apesar de considerarem o risco, os autores argumentam que estes resultados podem ter ocorrido devido a uma especificação incorreta do processo gerador dos retornos. Desta forma, as estratégias são testadas novamente sobre séries artificiais geradas sob a hipótese de processos cada vez mais complexos: $AR(1)$, $GARCH$, $GARCH-M$ e $E-GARCH$. Os resultados ainda assim, mostram a utilidade destas estratégias na tomada de decisão.

Os autores também mencionam a possibilidade de existência de *data-snooping*. O argumento é que inúmeras estratégias diferentes podem ser construídas e por isso, algumas delas certamente irão apresentar resultados estatisticamente significativos. Esta possibilidade é estudada no artigo de Sullivan *et al.* (1997).

Um dos problemas com o tipo de técnica *bootstrap* utilizada neste artigo é a necessidade de parametrização do DGP dos retornos. No trabalho, consideramos um tipo não-paramétrico de *bootstrap*. Esta técnica possui a vantagem de permitir processos mais gerais para construção das séries reamostradas, reduzindo a chance de um erro de especificação na geração das séries.

1.3.2 Leal e Varanda (2000)

A mesma metodologia acima é empregada ao mercado brasileiro por Leal e Varanda (2000) utilizando somente o passeio aleatório padrão. Neste estudo, é usada a série do Ibovespa e de suas ações mais líquidas entre julho de 1994 e novembro de 1998. As estratégias testadas são as de média móvel com filtros (vide pág.10), aplicadas a cada um destes ativos, num total de dez estratégias em cada mercado. Incluiu-se também um custo de 0,5% por transação efetuada, descontados do retorno de cada operação.

Os resultados exibidos para a série do Ibovespa indicam a existência de uma estratégia (MA(1,50) com *filtro* = 0) que produz, mesmo após considerarmos os custos de transação, retornos superiores à estratégia de “comprar e esperar”. Este excesso de retorno porém, não é estatisticamente significativo, indicando a rejeição da análise técnica como forma de investimento no mercado brasileiro.

Este trabalho no entanto, ao utilizar somente um modelo muito simples para o DGP dos retornos, faz com que o emprego da metodologia *bootstrap* às séries de tempo produza resultados inadequados.³ A aplicação desta técnica só pode ser efetuada caso a distribuição dos retornos seja independente, trazendo dúvidas quanto às conclusões obtidas. Deve-se utilizar uma técnica que considere a dependência existente nos dados, como o *bootstrap* em blocos detalhado mais abaixo.

1.3.3 Sullivan, Timmermann e White (1997)

Na tentativa de melhorar os resultados publicados anteriormente, o artigo de Sullivan *et al.* (1997) apresenta inúmeras inovações frente à metodologia utilizada até então.

A primeira mudança é em relação ao tipo de *bootstrap* utilizado. Para levar em conta a dependência dos dados, os autores aplicam o *bootstrap* estacionário (Politis e Romano, 1994), que permite construir séries que levem em conta a dependência dos retornos. Outra inovação é o grande aumento no tamanho do espaço de estratégias, considerando 7.846 tipos diferentes.

A grande contribuição deste trabalho é a aplicação do teste de habilidade preditiva para testar a eficiência das estratégias de análise técnica, conhecido como teste de realidade de White (2000). Este teste considera a possibilidade de problemas causados por *data-snooping*,

³Como mencionado pelo autor, não foram considerados outros processos para os retornos devido ao alto custo computacional.

presente quando aplicamos milhares de estratégias à mesma série de preços. No teste de White compara-se a melhor performance na série real, como por exemplo o maior retorno entre as estratégias consideradas, com a melhor performance em cada uma das séries construídas via *bootstrap*. Neste caso, permite-se que a melhor estratégia em uma dada reamostragem seja diferente da melhor estratégia observada na série real, levando em conta portanto, os problemas causados por *data-snooping*. Na metodologia anterior, a performance de uma estratégia na série real do ativo é comparada com a performance desta mesma estratégia. Para que a análise técnica seja considerada útil como forma de investimento, o artigo testa se o maior excesso de retorno (ou índice de Sharpe) observado na série do Dow Jones é estatisticamente significativo em relação à distribuição gerada pelo maior excesso de retorno (ou índice de Sharpe) observado em cada série reamostrada.⁴

Os autores encontram evidências da utilidade desta forma de investimento (excluindo-se custos de transação) na amostra considerada, obtendo p-valores para a hipótese nula de não-validade da análise técnica que são próximos à 0,02%. Estes resultados são válidos tanto para o excesso de retorno médio quanto para o índice de Sharpe. A melhor estratégia observada porém, não consegue produzir performance superior significativa quando aplicada aos dez anos seguintes à amostra estudada. Os p-valores observados fora da amostra são, respectivamente, iguais a 30,07% e 26,42% para o excesso de retorno médio e o índice de Sharpe.

O teste de realidade porém, contém problemas que podem tornar sua estimação inconsistente. Por isto, usaremos também o teste de habilidade preditiva proposto por Hansen (2001), que corrige estes problemas e encontra uma distribuição mais geral para a estatística do teste.

2 Metodologia

2.1 Dados

O ativo escolhido para análise das estratégias no mercado brasileiro é o contrato futuro do índice Bovespa (Ibovespa) negociado na Bolsa de Mercadorias e Futuros (BM&F). Isto

⁴O índice de Sharpe utilizado em nosso trabalho é calculado por: $\frac{R_k - R_f}{\sigma(R_k)} - \frac{R_{bh} - R_f}{\sigma(R_{bh})}$. Na verdade, ele corresponde à diferença entre o índice de Sharpe da estratégia k e o índice de Sharpe da estratégia “comprar e esperar”.

deve-se primordialmente à grande facilidade de execução de ordens de venda à descoberto, que são muito difíceis (quando não impossíveis) de serem efetuadas no mercado brasileiro de ações. Os contratos futuros possuem prazo de vencimento de três meses, sempre na quarta-feira mais próxima ao dia 15 do mês de vencimento, tendo como objeto de negociação o valor do Ibovespa na data de vencimento. A construção desta série é feita através da união dos preços dos contratos mais negociados em um dado momento. A fim de que os dados sejam consistentes, uma correção é feita através da multiplicação de toda a série antiga pelo percentual esperado de valorização/desvalorização expressa nos preços dos contratos futuros.⁵ Isto é feito através da seguinte fórmula:

$$Fator\ de\ Ajuste(t) = \frac{P_t(Contrato\ Novo)}{P_t(Contrato\ Anterior)} \quad (1)$$

Este fator de ajuste nos dá o grau esperado de variação do Ibovespa pelos agentes. Quando multiplicado pela série antiga, ele faz com que os preços passados passem a fazer referência ao contrato mais líquido, garantindo a integridade dos dados após a união das séries. Estes dados ajustados nos permitem calcular retornos idênticos àqueles que seriam obtidos através da “rolagem” da posição de um contrato para outro.⁶ Os preços diários (máximo, mínimo, abertura e fechamento) utilizados compreendem o período entre 04/06/1992 e 08/01/2002, perfazendo um total de 2.369 observações.⁷

O gráfico abaixo nos mostra a trajetória da série no período. Após quase 10 anos observamos que o índice futuro ainda permanece ao redor do patamar de 14.500 pontos. O mercado futuro teve grande alta até o início do Plano Real, seguido de forte queda até meados de 1995. Depois observa-se a grande valorização com ápice em julho de 1997, meses antes à crise asiática. Após este período, com as sucessivas crises cambiais ao redor do mundo o índice volta aos níveis do final de 1993, flutuando neste patamar até o presente momento.

[Figura 1]

Os preços de fechamento exibem retornos logarítmicos negativos de $-13,60\%$ no período, sendo ligeiramente assimétricos à esquerda e leptocúrticos. A autocorrelação de primeira

⁵A série antiga refere-se à série de preços calculada até o momento.

⁶Não considerados custos de transação. Esta hipótese não altera os resultados obtidos. Além disso, para grandes investidores (fundos de investimento, etc...) os custos são muito próximos a zero.

⁷As séries foram obtidas através do programa Economática.

ordem, apesar de pequena, é estatisticamente significativa conforme mostra a tabela abaixo.⁸ Analisando o quadrado destes retornos, observamos grande autocorrelação, indicando a presença de um componente autoregressivo na variância condicional. Exibimos também os cálculos para os retornos não-justapostos no período de 10 dias. Neste caso, tanto r_{10} quanto $(r_{10})^2$ não exibem autocorrelações estatisticamente significativas, sugerindo que a hipótese de passeio aleatório pode ser aplicada neste caso.

Tabela 1: Estatísticas Descritivas - Retornos

Medidas	r_1	r_{10}	$(r_1)^2$	$(r_{10})^2$
Média	-0,013%	-0,122%	0,118%	0,891%
Desv. Pad.	3,426%	10,425%	0,320%	2,559%
Assimetria	-0,053	-0,508	9,006	6,011
Curtose	7,583	6,589	-	-
Autocorrelações				
$\rho(1)$	-0,0334	0,0027	0,2662**	0,1160
$\rho(2)$	-0,0085	-0,0699	0,2708**	-0,1306
$\rho(3)$	-0,0263	-0,0060	0,2110**	-0,1937
$\rho(4)$	-0,0148	-0,0402	0,1873**	0,0379
$\rho(5)$	-0,0307	-0,0021	0,1603**	-0,0281
$\rho(6)$	-0,0862*	-0,0701	0,1951**	0,0240
$\rho(7)$	-0,0248	0,0210	0,1688**	0,0032

[Figura 2]

2.2 Estratégias de Análise Técnica (Elder, 1993)

A análise técnica de investimentos baseia-se no uso da informação passada dos preços de um ativo para gerar sinais de compra e venda. Ela pode ser dividida em dois tipos básicos de estratégias: análise gráfica e análise computadorizada.

A análise gráfica teve início no começo do século XX e consiste na identificação de padrões recorrentes nos gráficos da série dos preços. Seu maior problema é a subjetividade da tomada

⁸O termo r_i representa o retorno no período de i dias. $\rho(k)$ representa a autocorrelação de ordem k . O símbolo *(**) denota confiança a 95% (99%).

de decisões e por isso tais estratégias não são consideradas neste trabalho, apesar de já terem sido alvo de estudos teóricos (Osler e Chang, 1995).

A análise computadorizada usa funções dos preços passados - os indicadores técnicos - na decisão de investimentos. Estes indicadores tentam expressar um suposto consenso de mercado para o preço do ativo no futuro, usando para isto informações correntes como volume e preços de fechamento, máximo ou mínimo. A facilidade com que estes indicadores podem ser calculados nos permite estabelecer objetivamente critérios de compra e venda de um ativo, tornando extremamente simples sua implementação computacional (Lo *et al.*, 2000).

São considerados cinco tipos básicos de estratégias: médias móveis, o índice de força relativa (RSI), indicador William's %R, Estocástico e o indicador de Média Móvel Convergência-Divergência (MACD). Este conjunto é escolhido por estas estratégias serem muito utilizadas pelos investidores nos mercados e possui a vantagem de ser mais complexo do que àqueles utilizados em trabalhos anteriores. Este espaço de estratégias considera não somente a série dos preços de fechamento, mas também as de máximo e mínimo, usando um conjunto de informações maior à disposição dos agentes. Isto também permite verificar se resultados anteriores, que indicam a rejeição da utilidade das estratégias de análise técnica, ocorrem devido ao tipo de estratégia utilizada.

2.2.1 Médias Móveis

Esta estratégia consiste na compra do ativo quando a média móvel de curto prazo do preço de fechamento (P_t) for superior à uma média móvel de longo prazo, e na venda caso o inverso ocorra. O cálculo da média móvel com p defasagens de uma série é feito através da seguinte fórmula:

$$P_t(p) = \frac{\sum_{i=1}^p P_{t-i}}{p} \quad (2)$$

A explicação dada pelos analistas técnicos para o uso desta estratégia é que a média móvel nos permitem identificar uma opinião dos agentes a respeito dos preços. Suponha dois números c e l , onde $c < l$. Caso uma média móvel de curto prazo ($P_t(c)$) seja maior do que uma média móvel de longo prazo ($P_t(l)$), o mercado está mais otimista (“*bullish*”) e o ativo

deve ser comprado.⁹

As diferentes estratégias são obtidas pela combinação de médias móveis com diferentes defasagens e também com a inclusão de filtros $(1+f)$, que alteram as decisões de investimento. Isto é feito para dificultar o início de uma operação sem impor restrições sobre seu término, diminuindo o número de ordens executadas e com isso, os custos de transação. A inclusão do filtro nos dá a seguinte regra de decisão:¹⁰

$$\begin{cases} \text{Inicia compra, caso } P_t(c) > (1+f)P_t(l) \\ \text{Encerra compra, caso } P_t(c) < P_t(l) \\ \text{Inicia venda à descoberto, caso } P_t(c)(1+f) < P_t(l) \\ \text{Encerra venda à descoberto, caso } P_t(c) > P_t(l) \end{cases}$$

Suponha um filtro de 5%, com $P_1(2) = 100$, $P_1(20) = 95$, $P_{10}(2) = 95$ e $P_{10}(20) = 101$. Como $100 = P_1(2) > (1+f)P_1(20) = 99,75$, compramos o ativo na data $t = 1$. Na data $t = 10$, $95 = P_{10}(2) < P_{10}(20) = 101$ e por isso encerramos a operação e vendemos o ativo à descoberto somente se $(1+f)P_{10}(2) < P_{10}(20)$.

2.2.2 Índice de Força Relativa (RSI)

Este indicador técnico é um dos mais difundidos entre os investidores, buscando utilizar os preços de fechamento para identificar o potencial de incremento de um dado ativo. Seu cálculo é feito através de:

$$RSI_t(p) = 100 - \frac{100}{1 + RS_t(p)}, \text{ onde:} \quad (3)$$

$$RS_t(p) = \frac{\sum_{i=1}^p (P_{t-i} - P_{t-i-1}) I(P_{t-i} > P_{t-i-1})}{\sum_{i=1}^p (P_{t-i-1} - P_{t-i}) I(P_{t-i} \leq P_{t-i-1})} \quad (4)$$

⁹As estratégias técnicas abaixo consideram tanto estratégias de “momento” quanto estratégias que buscam captar uma reversão à média dos preços.

¹⁰Os filtros são aplicados de forma similar em todas as estratégias mencionadas no trabalho.

e $I(\cdot)$ denota a função indicadora.

Logo, $RSI_t(p)$ equivale à razão entre a média de crescimento dos preços nos dias em que o preço do ativo subiu e a média de decréscimo nos dias em que o preço do ativo caiu, implicando em $RSI \in [0, 100)$. A decisão de investimento baseia-se na seguinte regra:

$$\begin{cases} \text{Inicia compra, caso } RSI_t(p)(1 + f) < Inf \\ \text{Encerra compra, caso } RSI_t(p) > Inf \\ \text{Inicia venda à descoberto, caso } RSI_t(p) > (1 + f)Sup \\ \text{Encerra venda à descoberto, caso } RSI_t(p) < Sup \end{cases}$$

A intuição por trás deste indicador é que quanto maior for o seu valor, menor é o potencial de crescimento do ativo. Caso o indicador fique menor do que o limite inferior (assumindo $f = 0$), o ativo tem maior potencial de valorização do que desvalorização e portanto deve ser comprado. Os valores Sup e Inf indicam os pontos à partir dos quais considera-se o ativo subavaliado ou sobreavaliado. Essas duas variáveis são escolhidas pelo agente de forma *ad-hoc* e constituem mais dois parâmetros de escolha.

2.2.3 Indicador William's %R

Este indicador calcula em que ponto está o último preço de fechamento em relação à amplitude dos preços em um período de p dias:

$$W\%R_t(p) = 100 \left(\frac{H_p - P_t}{H_p - L_p} \right) \quad (5)$$

onde $H_p = \text{Max}\{P_t, \dots, P_{t-p}\}$, $L_p = \text{Min}\{P_t, \dots, P_{t-p}\}$.

O indicador vale zero quando o preço de fechamento (P_t) é igual ao preço máximo observado no período de p dias (H_t), e vale 100 quando P_t é igual ao preço mínimo (L_t) neste período. A estratégia para ele é dada por:

$$\begin{cases} \text{Inicia venda à descoberto, caso } W\%R_t(p)(1 + f) < Inf \\ \text{Encerra venda à descoberto, caso } W\%R_t(p) > Inf \\ \text{Inicia compra, caso } W\%R_t(p) > (1 + f)Sup \\ \text{Encerra compra, caso } W\%R_t(p) < Sup \end{cases}$$

2.2.4 Estocástico

Neste caso as decisões de compra e venda baseiam-se em duas séries para identificar os pontos de entrada ou saída do mercado: o Estocástico rápido (%K) e o Estocástico lento (%D). Este indicador utiliza médias móveis que calculam em que ponto está o último preço de fechamento em relação à amplitude dos preços em um período de p dias.

$$\%K_t(p) = 100 \left(\frac{P_t - L_p}{H_p - L_p} \right) \quad (6)$$

$$\%D_t(p, d) = 100 \left(\frac{\sum_{i=0}^{d-1} (P_{t-i} - L_p)}{\sum_{i=0}^{d-1} (H_p - L_p)} \right) \quad (7)$$

O indicador lento busca filtrar o ruído de mercado e com isso diminuir o número de sinais falsos dado pelo William's %R. A estratégias também utilizam valores *ad-hoc* para definir os pontos de sobreavaliação (*Sup*) e subavaliação (*Inf*) do ativo. A decisão baseia-se na seguinte regra:

$$\begin{cases} \text{Inicia compra, caso } \%K_t(p)(1+f) \text{ e } \%D_t(p, d)(1+f) < Inf \\ \text{Encerra venda à descoberto, caso } \%K_t(p, d) \text{ ou } \%D_t(p, d) > Inf \\ \text{Inicia compra, caso } \%K_t(p) \text{ e } \%D_t(p, d) > (1+f)Sup \\ \text{Encerra compra, caso } \%K_t(p) \text{ ou } \%D_t(p, d) < Sup \end{cases}$$

Valores altos de %K indicam que os “comprados” estão dominando o mercado, já que conseguem que o preço de fechamento esteja sempre próximo ao preço máximo observado no período. Esta maior influência dos “comprados” nos sugere comprar o ativo.

2.2.5 Indicador MACD

Nosso último tipo de estratégia usa o indicador MACD. Este indicador é uma extensão das estratégias baseadas em médias móveis, utilizando suavizações das mesmas para reduzir a quantidade de falsos sinais de compra e venda. As decisões de investimento são tomadas através da comparação de médias de curto e longo prazo: quando a média de curto prazo cruza a média de longo prazo, é sinal que o ativo está se valorizando no curto prazo e portanto sugere sua compra. O cálculo do MACD é feito da seguinte forma:

1. Calcular uma média móvel exponencial (EWMA) de k períodos do preço de fechamento
2. Calcular uma média móvel exponencial (EWMA) de s períodos do preço de fechamento
3. $MACD\ Rápido(k, s) = EWMA(k) - EWMA(s)$
4. $MACD\ Lento(k, s, d) = EWMA(d)$ de $MACD\ Rápido$

A variável $MACD\ Rápido$ retrata o consenso de mercado no curto prazo enquanto $MACD\ Lento$ denota o consenso de longo prazo. Se $MACD\ Rápido < MACD\ Lento$ temos indicação que o preço do ativo deve cair e assim vendemos o ativo. A regra geral é dada por:

$$\begin{cases} \text{Inicia compra, caso } MACD\ Rápido(k, s) > (1 + f)MACD\ Lento(k, s, d) \\ \text{Encerra compra, caso } MACD\ Rápido(k, s) < MACD\ Lento(k, s, d) \\ \text{Inicia venda à descoberto, caso } MACD\ Rápido(k, s)(1 + f) < MACD\ Lento(k, s, d) \\ \text{Encerra venda à descoberto, caso } MACD\ Rápido(k, s) > MACD\ Lento(k, s, d) \end{cases}$$

2.3 Técnica *Bootstrap*

O próximo passo é a construção de realizações alternativas do processo estocástico dos preços do contrato futuro do índice Bovespa. Essas realizações são necessárias para calcular os retornos das estratégias de análise técnica e aplicar os testes de habilidade preditiva. Isto é feito através da técnica *bootstrap*, que consiste na reamostragem dos dados para estimação da distribuição de um estimador ou estatística de teste. Para facilitar a compreensão, seguimos ao longo do trabalho a notação de Horowitz (2001).

Seja a amostra de tamanho n , $\{X_i : i = 1, \dots, n\}$, com função de distribuição acumulada (FDA) F_0 . Seja $T_n = T_n(X_1, \dots, X_n)$ uma função da amostra cuja FDA seja dada por $G_n(\tau, F_0) = P(T_n \leq \tau)$.

O *bootstrap* nos permite estimar a distribuição em amostras finitas de G_n substituindo a função de distribuição F_0 pela distribuição empírica F_n , dada por:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x)$$

A dificuldade em encontrar $G_n(\tau, F_n)$ analiticamente torna necessário o uso de técnicas de reamostragem para estimá-la. O procedimento consiste em extrair aleatoriamente diversas amostras de F_n e então calcular a estatística de teste T_n para cada uma delas. Usamos estas estatísticas para encontrar a probabilidade empírica do evento $T_n \leq \tau$ e poder realizar assim os testes de hipótese.

A existência de dependência nos dados faz com que esta reamostragem precise ser feita de forma a capturar adequadamente esta característica do processo gerador dos dados. Para evitar a especificação paramétrica do DGP, utilizamos o *bootstrap* em blocos desenvolvido por Künsch(1989).

2.3.1 *Bootstrap* em Blocos

A técnica não-paramétrica do *bootstrap* em blocos sobrepostos procura reproduzir a dependência através de blocos consecutivos de observações de tamanho q . A idéia é construir blocos que sejam independentes mas preservem a dependência dos dados dentro de um bloco, permitindo a aplicação da reamostragem dos dados.

A escolha do tamanho do bloco é uma questão muito discutida na literatura (Hall *et al.*, 1995). Um valor pequeno para q nos produz amostras que não capturam corretamente a dependência dos dados. Por outro lado, um valor muito grande restringe as possíveis realizações da série, diminuindo a eficiência das estatísticas. O valor de q deve aumentar com o tamanho da amostra para que as estimativas das FDA's sejam consistentes. A escolha ótima do tamanho do bloco é feita através da minimização do erro médio quadrático do estimador gerado pela técnica. Dependendo do estimador desejado, como por exemplo variância, viés ou uma função de distribuição, temos um diferente valor para o tamanho do bloco. Para estimação de funções de distribuição simétricas (ex: $P(T_n \leq \tau)$, trabalhos anteriores (Hall *et al.*, 1995) mostram que $q \sim n^{\frac{1}{4}}$. Consequentemente, com $n = 2.369$ temos blocos de tamanho $q = 6$. A estimação da técnica é feita através do seguinte procedimento:

Sejam $\{X_t : i = 1, \dots, n\}$ um vetor contendo as k séries utilizadas, q o tamanho do bloco e B o número de séries construídas.

1. Formar blocos de tamanho n , onde o $m - \text{ésimo}$ é formado pelas observações $\{X_m, \dots, X_{m+q}\}$.
2. Extrair blocos aleatoriamente com reposição e colá-los na ordem extraída.

3. Repetir o processo B vezes.

4. Denotar as séries construídas por: $\{X_t^{(b)} : t = 1, \dots, n ; b = 1, \dots, B\}$.

O procedimento nos permite calcular B realizações alternativas do processo gerador dos dados. Aplicando a estatística de teste a cada série criada, podemos construir a função de densidade empírica do teste.

2.4 Testes de Habilidade Preditiva

Os testes de habilidade preditiva têm como objetivo comparar a eficiência relativa de modelos estatísticos sob uma determinada medida de performance. Um exemplo seria o erro quadrático médio de uma previsão ou o excesso de retorno de uma estratégia. As hipóteses nulas destes testes envolvem o cálculo da distribuição do máximo de variáveis aleatórias, complicando a derivação da distribuição das estatísticas de teste. A aplicação de técnicas de reamostragem permitem estimar consistentemente a distribuição da estatística de teste via métodos numéricos.

Considere uma situação onde existam $l + 1$ modelos diferentes e desejamos testar se o modelo de referência 0 é inferior a qualquer um dos modelos $k = 1, \dots, l$, onde l é o tamanho do espaço de modelos considerado. Seja μ o vetor l -dimensional do desempenho de cada modelo alternativo em relação ao modelo de referência. Nossa hipótese nula é definida por $H_0 : \mu \leq 0$, equivalente a testar se $\mu_{\max} \equiv \max \mu_k \leq 0$. Já a fronteira de H_0 é dada por: $\left\{ \mu \in \mathfrak{R} : \max_{k=1, \dots, l} \mu_k = 0 \right\}$.

2.4.1 Teste de Realidade (White 2000)

A superioridade de uma estratégia na série original do Ibovespa pode ocorrer simplesmente por acaso e não devido à superioridade da regra de decisão. Ao testar milhares de estratégias diferentes é muito provável que pelo menos uma apresente retornos significativos, sem nos dar algum tipo de informação à respeito de sua capacidade preditiva. O teste de realidade de White tem por objetivo justamente considerar esta possibilidade na avaliação das performances dos modelos.

A utilização indiscriminada da técnica *bootstrap* não é capaz de levar em conta os problemas causados por *data-snooping*. Anteriormente, comparávamos a melhor média de

retorno na série original do Ibovespa com os retornos obtidos por esta mesma regra em todas as reamostragens construídas. Os testes de habilidade preditiva comparam a melhor estratégia na série original com a melhor estratégia observada em *cada uma* das séries reamostradas. Estas estratégias podem ser (e geralmente são) diferentes da melhor estratégia encontrada na série original. Os resultados principais do teste criado por White (2000) são apresentados abaixo:

Seja n o tamanho da série utilizada. A hipótese é testada em um vetor $E(f^*)$ de tamanho $l \times 1$, onde $f^* \equiv f(Z, \beta^*)$ é um vetor $l \times 1$ com elementos $f_k^* \equiv f_k(Z, \beta^*)$ para o vetor aleatório Z e parâmetros $\beta^* \equiv \text{plim } \hat{\beta}_T$. O vetor Z consiste em vetores de variáveis dependentes Y e regressores X .

O teste é baseado na estatística $l \times 1$ denotada por:

$$\bar{f} \equiv n^{-1} \sum_{t=1}^n \hat{f}_t \quad (9)$$

onde $\hat{f}_t \equiv f(Z_t, \hat{\beta}_t)$. A hipótese nula que exprime o teste de ausência de superioridade preditiva em relação ao modelo de referência é dada por:

$$H_0 : E(f^*) \leq 0 \quad (10)$$

Para testar a hipótese de uma determinada estratégia produzir retornos superiores a uma estratégia de referência usamos:

$$\hat{f}_{t+1} = \ln [1 + y_{t+1} S_1(X_{1,t+1}, \beta_1^*)] - \ln [1 + y_{t+1} S_0(X_{0,t+1}, \beta_0^*)], \quad (11)$$

onde y_{t+1} representa o retorno diário e S_0 e S_1 representam funções que convertem os indicadores de análise técnica ($X_{0,t+1}$ e $X_{1,t+1}$) e os parâmetros (β_0^* e β_1^*) em posições de mercado. As funções podem assumir três valores: -1 (vendido a descoberto), 0 (fora do mercado) e 1 (comprado). Em nosso caso, os parâmetros (β_0^* e β_1^*) são estabelecidos *a priori* e não precisam ser estimados (exemplo: o n° de defasagens utilizadas na estratégia de média móvel). Se a estratégia “comprar e esperar” for o modelo de referência, teremos $S_0(X_{0,t+1}, \beta_0^*) = 1$ para todo t .

A hipótese nula que testa o modelo com melhor performance é dada por:

$$H_0 = \max_{k=1, \dots, l} E(f_k^*) \leq 0 \quad (12)$$

A hipótese múltipla, representada por (12), é uma interseção das hipóteses individuais da equação (10) para cada um dos l modelos.

2.4.2 Distribuição Assintótica

A proposição 2.1 de White (2000) mostra que, se algum modelo supera aquele usado como referência, isto acaba sendo revelado por uma estimativa positiva da performance em relação ao modelo de referência na série original.

A proposição seguinte permite a construção de um teste para seleção do melhor modelo, utilizando o extremo de um vetor de variáveis normalmente distribuídas:

Suponha que $n^{\frac{1}{2}} (\bar{f} - E(f^*)) \xrightarrow{d} N(0, \Omega)$ com Ω positiva semi-definida. Então, para $T \rightarrow \infty$, temos:

$$\max_{k=1, \dots, l} \left\{ n^{\frac{1}{2}} (\bar{f}_k - E(f_k^*)) \right\} \xrightarrow{d} V_l \equiv \max_{k=1, \dots, l} \{H_k\} \quad (13)$$

e

$$\min_{k=1, \dots, l} \left\{ n^{\frac{1}{2}} (\bar{f}_k - E(f_k^*)) \right\} \xrightarrow{d} W_l \equiv \min_{k=1, \dots, l} \{H_k\} \quad (14)$$

onde H é um vetor $l \times 1$ com componentes $H_k \sim N(0, \Omega)$.

Desta forma, estabelecida a distribuição assintótica, a conclusão do teste é a mesma independente da validade da hipótese nula. O teste da nula é reforçado utilizando o valor de $E(f_k^*)$ menos favorável à alternativa: $E(f_k^*) = 0, \forall k$. Assim, o comportamento do critério de seleção do modelo que maximiza a estatística $\bar{V}_l \equiv \max_{k=1, \dots, l} n^{\frac{1}{2}} \bar{f}_k$ é conhecido, permitindo a construção de p-valores assintóticos.

Qualquer método que permita encontrar estimativas consistentes para os p-valores do teste $H_0 = \max_{k=1, \dots, l} E(f_k^*) \leq 0$ é chamado de “teste de realidade”. O desafio em sua implementação é que a distribuição da estatística de teste (o máximo de um vetor de variáveis normalmente distribuídas e correlacionadas) é desconhecida, sendo necessário a implementação de métodos numéricos para estimá-la..

Isto é feito através da técnica *bootstrap* descrita anteriormente. Com repetidas extrações de $n^{\frac{1}{2}} (\bar{f}_k - E(f_k^*))$ podemos encontrar uma estimativa da distribuição $N(0, \Omega)$. Os p-valores do teste de realidade podem ser obtidos pela comparação de $\bar{V}_l \equiv \max_{k=1, \dots, l} n^{\frac{1}{2}} \bar{f}_k$, com os percentis de $\bar{V}_l^{(b)} \equiv \max_{k=1, \dots, l} n^{\frac{1}{2}} (\bar{f}_k^{(b)} - \bar{f}_k)$. Estes valores são obtidos através da aplicação desta estatística em cada série artificial gerada pela técnica *bootstrap*.

2.4.3 Teste de Habilidade Preditiva Superior

Vimos acima que o teste de realidade (RC) pode ser usado para testar a hipótese nula de superioridade da estratégia de referência. Entretanto, conforme resultados apresentados por Hansen (2001), o RC possui um problema causado pela existência de um parâmetro de estorvo, a $E(f_k^*)$. Isto torna o RC sensível à inclusão de modelos irrelevantes ao espaço de estratégias alternativas, produzindo p-valores inconsistentes quando a hipótese nula *não* é a de que todos os modelos têm performance idênticas ($H_0 : \mu_k^{RC} = E(f_k^*) = 0$, para todo $k = 1, \dots, l$).

A hipótese nula do RC utiliza o valor zero para $E(f_k^*)$, por ser este valor o menos favorável à hipótese-alternativa. A estatística é dada por $\bar{V}_l = \max_{k=1, \dots, l} n^{\frac{1}{2}} \bar{f}_k$. Porém, $E(f_k^*) = 0$ é somente um dos valores consistentes para H_0 e portanto, assintoticamente, a distribuição pode não ser a correta para \bar{V}_l , justamente devido à existência do parâmetro de estorvo.

O teste de Hansen (2001) propõe-se a utilizar uma estimativa consistente para $E(f_k^*)$, que permite a derivação correta da estatística do teste e o uso da técnica *bootstrap*. A distribuição do teste derivada por Hansen (2001) é a seguinte:

Suponha que a estatística de teste possui a forma $T_n = \varphi(n^{\frac{1}{2}} \bar{f})$, onde $\varphi : \mathfrak{R}^l \rightarrow \mathfrak{R}$ é uma função contínua que satisfaz as seguintes condições:

1. $\varphi(0) = 0$;
2. $\varphi(z) = \varphi(z^+)$ onde $z^+ = \max(0, z_k)$, $k = 1, \dots, l$
3. $\varphi(z) > \varphi(z')$ se $z_k \geq z'_k \geq 0$, para todo $k = 1, \dots, l$

Com estas condições, Hansen (2001) mostra o seguinte resultado:

Suponha que $n^{\frac{1}{2}} (\bar{f} - E(f^*)) \xrightarrow{d} N(0, \Omega)$, onde $E(f^*) \leq 0$. Defina a matriz Ω^0 por:

$$\Omega_{ij}^0 = \begin{cases} \Omega_{ij} & \text{se } E(f_i^*) = E(f_j^*) = 0 \\ 0 & \text{caso contrário} \end{cases}$$

Seja $\varphi(\cdot)$ descrito pela hipóteses acima. Então:

$$\varphi(n^{\frac{1}{2}} \bar{f}) \xrightarrow{d} N(0, \Omega^0)$$

Vale notar que a estatística do RC não satisfaz a hipótese (2), já que \bar{V}_l pode ser negativo. Entretanto, a distribuição de \bar{V}_l é idêntica no eixo positivo à de $\max(0, \bar{V}_l)$ e assim satisfaz as condições requeridas.

O fato da hipótese nula ser composta não nos permite identificar claramente qual o melhor estimador para $E(f_k^*)$. A dificuldade na construção de um estimador consistente para os p-valores ocorre, fundamentalmente, porque o valor de $E(f_k^*)$ está na fronteira de um espaço de parâmetros definido por relações de desigualdade. Isto causa sérios problemas na aplicação do *bootstrap* (Andrews, 2000), que passa a produzir estimativas inconsistentes. Uma primeira aproximação seria usar $\mu_k^l = E(f_k^*) = \min(\bar{f}_k, 0)$. Infelizmente, μ_k^l também causa distorções na estimativa de modelos onde $E(f_k^*) = 0$, fazendo com que os p-valores sejam sempre subestimados. Ainda assim, pode-se usar μ_k^l como uma estimativa do limite inferior dos p-valores dos testes.

A resolução do problema passa pela modificação de μ_k^l para que, assintoticamente, modelos que estejam na fronteira do espaço de parâmetros ($E(f_k^*) = 0$) sejam estimados consistentemente. A estimativa corrigida, denotada por μ_k^c , neste caso é dada por (Hansen, 2001):

$$\mu_k^c = \begin{cases} 0 & \text{se } \bar{f}_k \geq -\frac{1}{4}n^{-\frac{1}{4}}\sqrt{\text{Var}\left(n^{-\frac{1}{2}}\bar{f}_k\right)} \\ \bar{f}_k & \text{se } \bar{f}_k < -\frac{1}{4}n^{-\frac{1}{4}}\sqrt{\text{Var}\left(n^{-\frac{1}{2}}\bar{f}_k\right)} \end{cases}$$

Esta alteração muda a estimativa de $E(f_k^*)$ garantindo que quase certamente, conseguimos estimar corretamente a $E(f_k^*)$ dos modelos que estão na fronteira do espaço de parâmetros, resultando em p-valores consistentes. Outras alterações assintoticamente consistentes podem ser construídas, com diferenças em pequenas amostras. Por isso, usamos os p-valores encontrados pelas três estimativas mencionadas acima (μ_k^c , a estimativa liberal μ_k^l e a conservadora μ_k^{RC}) para nos dar mais informação à respeito da rejeição da hipótese nula.¹¹

¹¹Um estimativa é liberal (conservadora) quando rejeita a hipótese nula mais (menos) vezes do que a estimativa consistente. Desta forma, estes estimadores estabelecem limites inferiores (superiores) para os p-valores.

3 Resultados Empíricos

3.1 Testes Iniciais

A primeira forma utilizada para testar a eficácia da análise técnica é através da estatística-t. Neste caso, testamos se a média do excesso de retorno de uma dada estratégia em relação à estratégia de “comprar e esperar ” é estatisticamente significativa. Dentre as 11.780 estratégias consideradas, a que gera melhor resultado, tanto em termos de excesso de retorno quanto índice de Sharpe, é a *Média Móvel*(25, 30, *Filtro* = 0). Ela tem um retorno diário médio de 0,14% e desvio-padrão 4,71%. A estatística-t correspondente (1,47) não rejeita a hipótese nula de eficácia da análise técnica (e conseqüentemente a EMH). Porém, como mostra a tabela abaixo, existe um grande número de estratégias cujos retornos são estatisticamente diferentes de zero.

Nível de Confiança	# Estratégias	% Significativas
1%	1.134	9,62%
5%	1.862	15,80%
10%	2.527	21,45%

A Figura 3 mostra o excesso de retorno de cada estratégia na série real do Ibovespa. Apesar das estratégias do tipo *RSI* e *Estocástico* apresentarem baixa performance em termos de excesso de retorno, várias do tipo *MACD* obtêm retornos positivos na série real do Ibovespa.

[Figura 3]

O uso de estratégias que utilizem não somente o preço de fechamento, mas também a informação contida nos preços de máximo e mínimo, parece adicionar ganhos de performance. No entanto esta afirmação precisa ser melhor avaliada, já que retornos superiores das estratégias *Média Móvel* e *MACD* podem ter ocorrido por acaso. Por isso, testaremos os retornos obtidos por esta estratégia na série real do Ibovespa *versus* os retornos observados nas séries reamostradas.

3.2 P-Valores Nominais

O uso da estatística-t para avaliar a significância de uma estratégia utiliza apenas uma observação, isto é, sua aplicação na série real, para inferência sobre sua significância. Para resolver este problema, seguimos os passos de Brock *et al.* (1992), comparando agora o excesso de retorno de cada estratégia com sua distribuição empírica, gerada pela aplicação somente desta estratégia nas 500 séries reamostradas. O p-valor nominal encontrado nos mostra se a performance obtida na série real do Ibovespa é um indicador confiável para o valor populacional desta performance.¹² Assim, sob a hipótese nula de que o excesso de retorno médio de uma estratégia de análise técnica é igual a zero, um p-valor de $\alpha\%$ nos diz com confiança de $(1 - \alpha)\%$, se a relação encontrada na série real é apenas devido ao acaso. Caso a performance na série real seja significativa, podemos rejeitar a hipótese que o resultado obtido ocorre somente devido à variabilidade amostral, indicando que a estratégia possui capacidade preditiva superior à do modelo de referência.

Este método resolve o problema mencionado acima, pois a técnica *bootstrap* utilizada não faz nenhuma hipótese a respeito da distribuição da estatística de teste. Além disso, a técnica permite encontrar a distribuição da estatística de performance através de sua aplicação nas séries artificiais.

No espaço de 11.780 estratégias, 560 (5026) obtêm retornos (índice de Sharpe) significativos a 5% de confiança, correspondendo a 4,75% (42,66%) do total. O aumento no número de estratégias significativas pelo índice de Sharpe, sugere que muitas estratégias conseguem diminuir significativamente o risco para os agentes. A tabela abaixo resume os resultados:

Tabela 3: Significância das Estratégias - P-Valores Nominais

Significância	# Retorno	% Sig.	# Sharpe	% Sig.
1%	187	1,58%	996	8,45%
5%	560	4,75%	5.026	42,66%
10%	902	7,65%	8.475	71,94%

A Figura 4 mostra o p-valor nominal para cada estratégia de investimento:

¹²Identifica-se na literatura como p-valor nominal, o p-valor do teste de significância através do uso da distribuição empírica da estratégia.

[Figura 4]

Estes resultados são interessantes ao mostrar a relevância do problema de *data-snooping* na análise dos dados. O grande número de estratégias significativas ao usarmos a estatística-t, mas que após o uso do p-valor nominal deixam de ser significativas, mostra a importância da consideração conjunta das performances para avaliar a utilidade da análise técnica.

Como a decisão de compra de um ativo é feita de forma completamente *ad-hoc*, podemos nos perguntar o que aconteceria caso invertessemos as decisões de investimento em cada estratégia. Ou seja, onde anteriormente comprávamos passamos a vender, e onde vendíamos, passamos a comprar o ativo.

A estratégia onde esta inversão aparenta ser mais útil é a *Williams %R* (denotadas por $W\%R^*$). O melhor resultado obtido com esta reversão é através da regra $W\%R^*(2, Sup = 30, Inf = 70, Filtro = 0)$, com excesso de retorno médio de 1,99% e desvio-padrão 4,27%.

Por questões de factibilidade computacional, os cálculos com regras de decisão invertidas serão refeitos utilizando somente as 2.850 combinações deste tipo de estratégia, aumentando o tamanho do espaço de modelos para 14.630 estratégias.

A Figura 5 exibe os retornos do novo espaço de estratégias, onde percebemos que a reversão dos sinais de compra/venda das estratégias é capaz de produzir retornos significativos na série real.¹³

[Figura 5]

A tabela abaixo exibe o percentual destas 2.850 novas estratégias significativas à 95%, utilizando a estatística-t e o p-valor nominal:

Tabela 4: Significância das Estratégias - P-valores Nominais

	Estatística-t	P-valor Nominal
Excesso de Retorno	16,24%	24,91%
Índice de Sharpe	-	33,64%

¹³Os resultados das estratégias $W\%R^*$ estão localizados nos pontos mais à direita nos gráficos

3.3 Teste de Realidade e SPA

Neste ponto é pertinente a crítica feita por Sullivan *et al.* (1999) de que a distribuição empírica utilizada acima, considera somente a performance de uma estratégia aplicada às séries reamostradas. Este resultado pode ocorrer devido ao acaso (através de *data-snooping*) e não por uma maior capacidade preditiva da estratégia. Os dois métodos que se propõem a corrigir este problema são o teste de realidade de White e a correção de Hansen (2001).

Assim, o excesso de retorno obtido na série real deve ser comparado ao maior retorno observado em cada série reamostrada pela aplicação das 14.630 estratégias. Caso a estratégia que obtenha o maior retorno na série real tenha significância rejeitada, temos indicação que nenhuma estratégia é significativa e assim não rejeitamos a EMH. A tabela abaixo mostra os resultados dos testes de habilidade preditiva.

Tabela 5: Testes de Habilidade Preditiva: P-Valores

Performance	μ^{RC}	μ^c	μ^l
Excesso de Retorno	21,8%	21,8%	21,4%
Índice de Sharpe	73,0%	73,0%	73,0%

Vemos então que a utilidade das estratégias de análise técnica é rejeitada, tanto através do excesso de retorno quanto pelo índice de Sharpe. Isto indica que nenhuma estratégia no espaço considerado possui performance estatisticamente superior à estratégia “comprar e esperar”. Podemos observar que os problemas de inconsistência levantados por Hansen (2001), não afetam os resultados no trabalho, pois $\mu^{RC} = \mu^c$.

A importância do uso de testes que considerem os efeitos de *data-snooping* é explicitado pelo resultado acima. Vimos que, dentre as 14.630 estratégias utilizadas, 5.985 (40,9%) possuem índice de Sharpe estatisticamente significativo. Ao utilizar conjuntamente a informação dada pelas estratégias, vemos que esta performance ocorre devido ao acaso, e não por uma maior capacidade destes modelos gerarem ganhos superiores significativos.

Nas Figuras 6 e 7, mostramos os comportamentos sequenciais do p-valor do teste ao adicionar uma estratégia ao espaço considerado para cada medida de performance. Cada adição impacta o p-valor da seguinte forma. Se a estratégia adicionada não tem performance superior às utilizadas até então, o p-valor da hipótese nula (que a melhor estratégia no espaço não possui performance superior) irá subir, ou no extremo, permanecer estável. Por outro lado, caso o ganho de performance seja grande o suficiente o p-valor irá cair, indicando uma maior chance da performance deste modelo ser significativa.

[Figura 6]

[Figura 7]

Como a ordem das estratégias é arbitrária, o p-valor relevante é o situado mais à direita do gráfico. Porém, é interessante observar os ganhos ocorridos quando um modelo com capacidade preditiva superior àquelas utilizadas até então, é adicionado. Tomando o espaço inicial (11.780 estratégias), podemos observar que a ausência de uma estratégia com performance superior faz com que o p-valor vá subindo lentamente na direção da não-rejeição da nula. Ao incluir o novo conjunto de estratégias, vemos na Figura 6 que quando a estratégia de maior excesso de retorno é adicionada (11.956), o p-valor cai abruptamente de 99% para aproximadamente 22%. Mesmo assim, não podemos rejeitar a hipótese nula de superioridade da estratégia “comprar e esperar”

Quando a estratégia com maior índice de Sharpe é adicionada (idêntica à de maior excesso de retorno), observamos na Figura 7, que o p-valor do teste de realidade (RC) varia de aproximadamente 99% para 73%. Após a inclusão desta estratégia no espaço de amostras, nenhuma outra é capaz de gerar excesso de retorno superior e, por isto, o p-valor segue estável até o final.

O gráfico abaixo mostra o p-valor do teste para o índice de retorno em função do número de séries reamostradas. Podemos observar que um aumento do número de reamostragens não parece ser necessário.¹⁴

[Figura 8]

Os resultados nos mostram que, *ex-post*, sempre podemos encontrar estratégias significativas. Porém, elas só podem ser consideradas úteis, caso esta performance significativa seja observada em diversas realizações do processo estocástico do ativo.

¹⁴Hall (1986) mostra que 500 replicações são suficientes para testes de hipótese com significância de até $\frac{500}{500+1} = 98,20\%$.

4 Conclusão

Após a realização dos testes, concluímos que as estratégias de análise técnica não devem ser utilizadas para auxiliar as decisões de investimento, já que não são capazes de remunerar o investidor adequadamente sob as medidas de performance consideradas. Assim, rejeitamos a hipótese de eficiência de mercado mesmo desconsiderando os custos de transação.

Outro ponto importante é a relevância de considerarmos todo o espaço de estratégias disponível ao investidor no cálculo da significância das estratégias. As conclusões extraídas quando isto não é feito nos leva à rejeição da EMH. Porém, o uso conjunto das informações contidas no espaço de estratégias, através dos testes de habilidade preditiva (o teste de realidade e o de Hansen), nos mostra que elas não são capazes de gerar retornos significativos.

Em relação aos trabalhos feitos anteriormente no país, esperamos ter agregado valor à literatura dos testes de eficiência de mercado. O trabalho é o primeiro a chamar atenção para os problemas causados pelo uso de um mesmo conjunto de dados para comparação da capacidade preditiva de modelos na área de finanças. O teste de realidade é uma ferramenta poderosa que permite ao economista validar de forma mais precisa suas afirmações e teorias.

Nossos resultados são similares aos encontrados nos mercados internacionais, contribuindo para o conjunto de evidências que mostram aos investidores que as estratégias de análise técnica (pelo menos nas formas simples como as utilizadas neste trabalho) não são úteis como investimento.

Em trabalhos posteriores, estratégias cada vez mais complexas serão testadas, especialmente aquelas que envolvem a combinação de sinais provenientes do uso de diferentes ativos, como por exemplo as séries do Ibovespa à vista e à do mercado futuro, ou estratégias, como por exemplo: $W\%R$ e $MACD$).

5 Referências

1. Andrews, D., 2000, Inconsistency of the Bootstrap when a Parameter is on the Boundary of the Parameter Space, *Econometrica* 68, 299-405.
2. Bachelier, L., 1900, Theory of Speculation, em *The Random Character of Stock Market Prices* (MIT Press), *Paul Cootner (ed.)*, 17-78.
3. Brock, W., Lakonishok, J. e LeBaron, B., 1992, Simple Technical Trading Rules and

- the Stochastic Properties of Stock Returns, *Journal of Finance* 49, 153-181.
4. Campbell, J., Lo, A. e Mackinlay, C., 1997, *The Econometrics of Financial Markets* (Princeton University Press).
 5. Cowles, A. 3rd, 1933, Can Stock Market Forecasters Forecast?, *Econometrica* 1, 309-324.
 6. Cowles, A. 3rd e Jones, H., 1937, Some A Posteriori Probabilities in Stock Market Action, *Econometrica* 5, 280-294.
 7. Elder, A., 1993, *Trading for a Living: Psychology, Trading Tactics, Money Management* (John Wiley).
 8. Fama, E. e Blume, M., 1966, Filter Rules and Stock Market Trading Profits, *Journal of Business* 39, 226-241.
 9. Fama, E., 1970, Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance* 25, 383-417.
 10. Hall, P., Horowitz, J. e Jing, B., 1995, On Blocking Rules for the Bootstrap with Dependent Data, *Biometrika* 82, 561-574.
 11. Hall, P., 1986, On the number of bootstrap simulations needed to construct a confidence interval, *Annals of Statistics* 14, 1453-1462.
 12. Hansen, P. R., 2001, An Unbiased and Powerful Test for Superior Predictive Ability, *Brown University Working Paper Series*.
 13. Horowitz, J., 2001, The Bootstrap, *Handbook of Econometrics Vol. 5* (North-Holland).
 14. Künsch, H., 1989, The Jackknife and the Bootstrap for General Stationary Observations, *Annals of Statistics* 17, 1217-1241.
 15. Leal, R. P. C. e Varanda, A. J., 2000, Desempenho de Estratégias Grafistas de Interseção de Médias Móveis através de Simulação *Bootstrap*, *Mercado de Capitais: Análise Empírica no Brasil* (Ed. Atlas), Capítulo 14.

16. Levich, R. e Thomas, L. III, 1993, The Significance of Technical Trading-Rule Profits in Foreign Exchange Market: A Bootstrap Approach, *Journal of International Money and Finance* 12, 451-474.
17. Lo, A., Mamaysky, H. e Wang, J., 2000, Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation, *Journal of Finance* 55, 1705-1765.
18. Osler, C.L. e Chang, P.H. K., 1995, Head and Shoulders: Not Just a Flaky Pattern, *Federal Reserve Bank of New York Staff Report* 4.
19. Politis, D. e Romano, J., 1994, The Stationary Bootstrap, *Journal of the American Statistical Association* 89, 1303-1313.
20. Roberts, H., 1967, Statistical versus Clinical Prediction of the Stock Market, Manuscrito não-publicado, *CRSP, University of Chicago*.
21. Sweeney, R., 1988, Some New Filter Rule Tests: Methods and Results, *Journal of Financial and Quantitative Analysis* 23, 285-300.
22. Sullivan, R., Timmermann, A. e White, H., 1999, Data-Snooping, Technical Trading Rules and the Bootstrap, *Journal of Finance* 54, 1647-1692.
23. —, No prelo, Scientific Progress with Data-Sharing, *International Journal of Forecasting*.
24. White, H., 2000, A Reality Check for Data Snooping, *Econometrica* 68, 1097-1126.

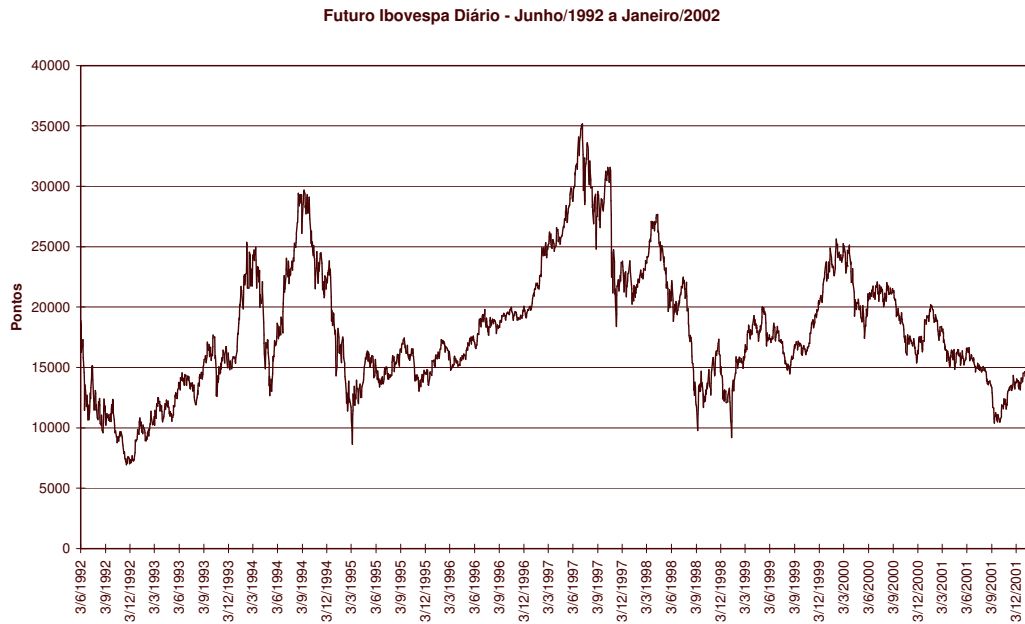


Figura 1: Fechamento do índice Bovespa Futuro - Base: Janeiro/2002

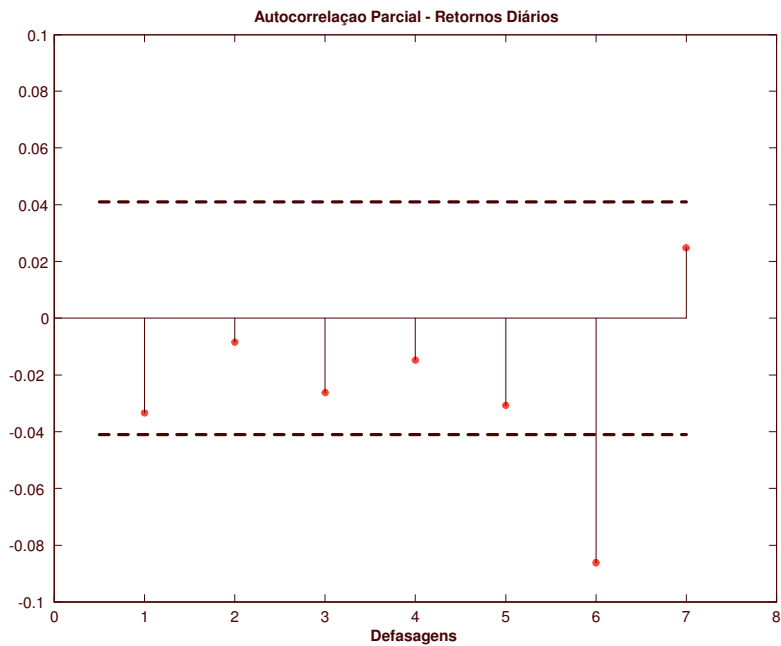


Figura 2: Autocorrelação dos Retornos Diários

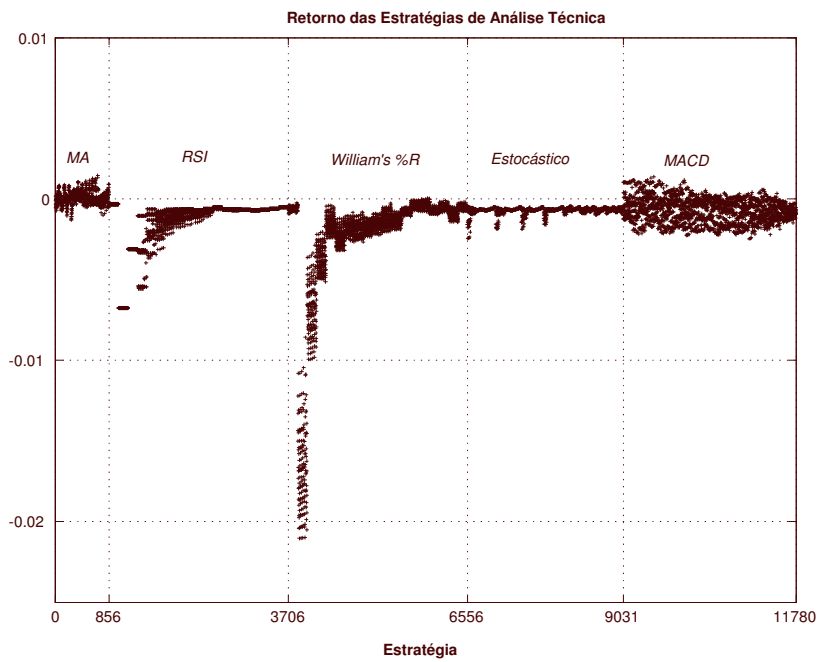


Figura 3: Excesso de Retorno Médio das Estratégias

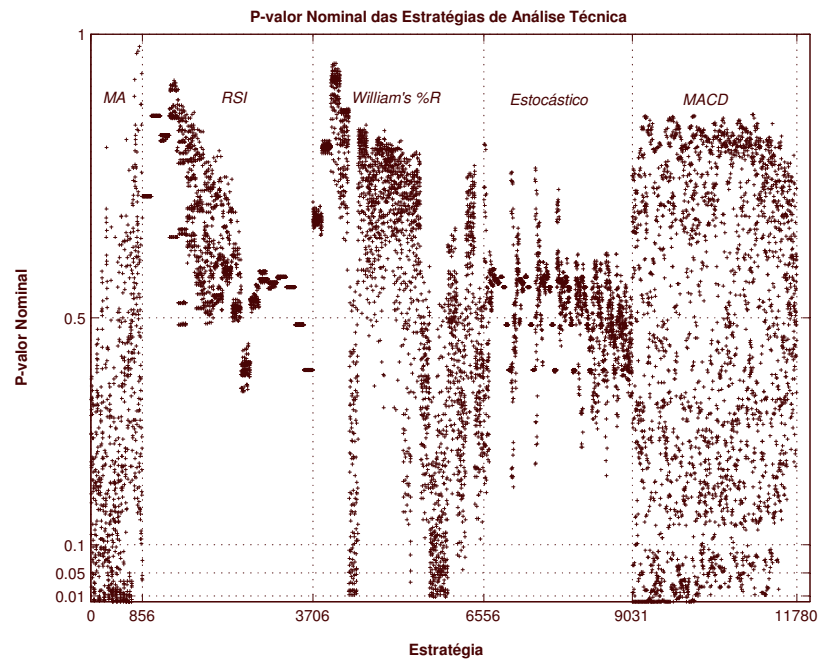


Figura 4: P-Valor Nominal - Retorno

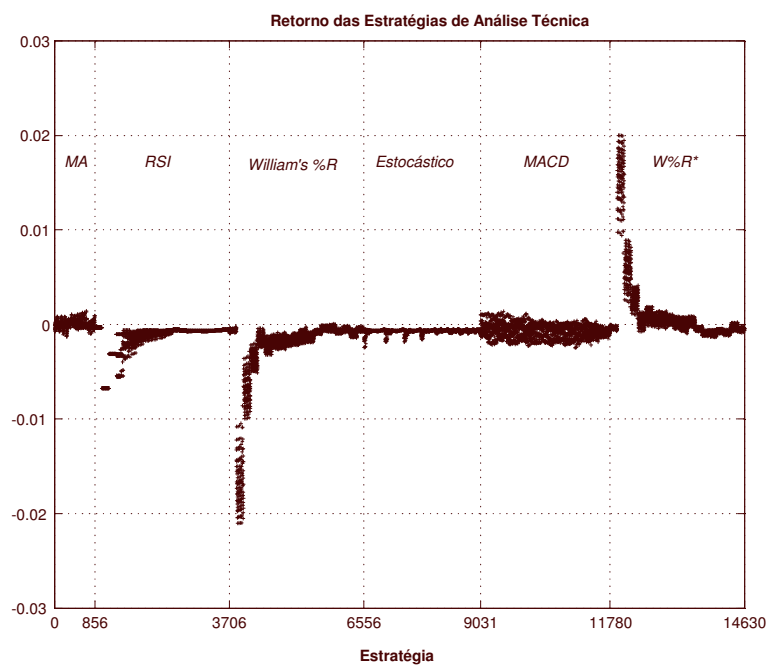


Figura 5: Excesso de Retorno Médio - Espaço de Estratégias Ampliado

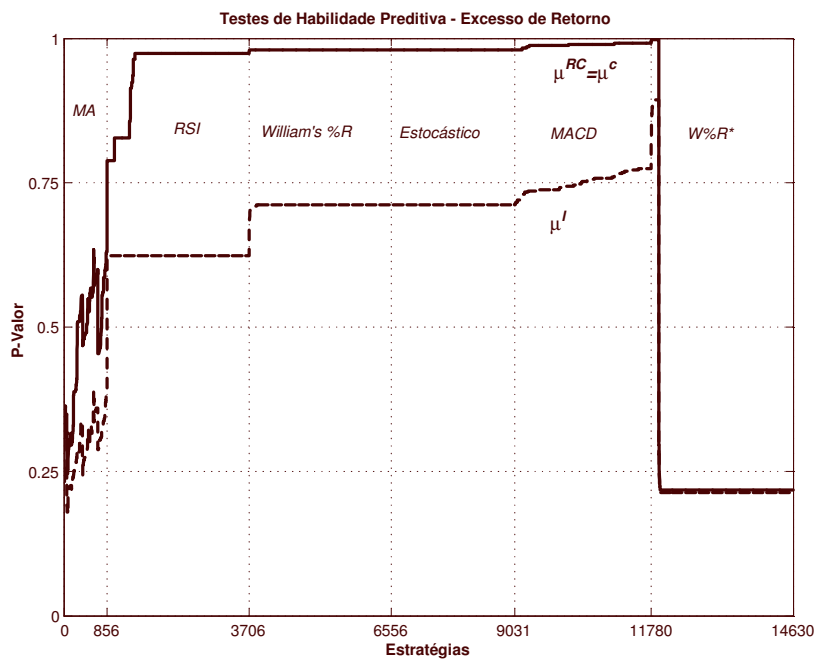


Figura 6: Testes de Habilidade Preditiva - Excesso de Retorno

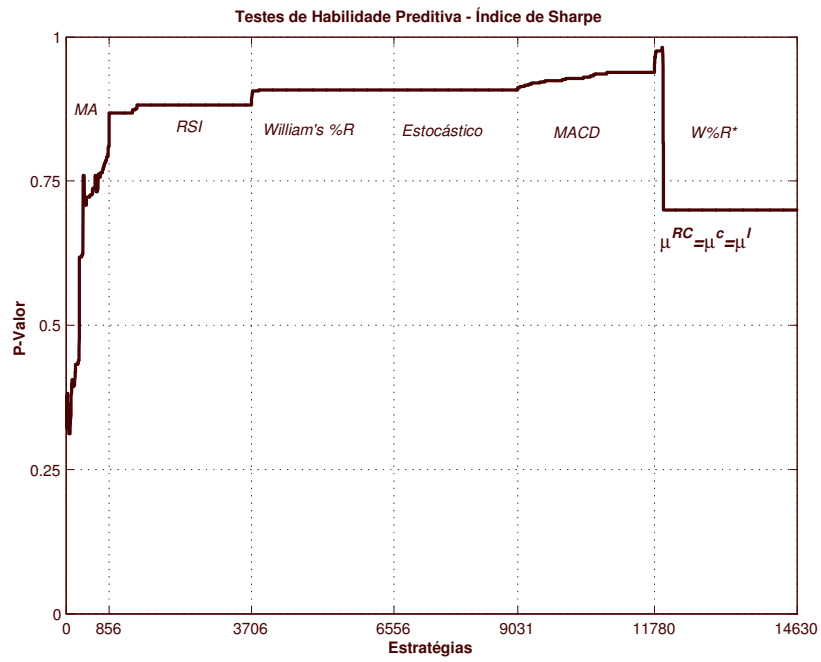


Figura 7: Testes de Habilidade Preditiva - Índice de Sharpe

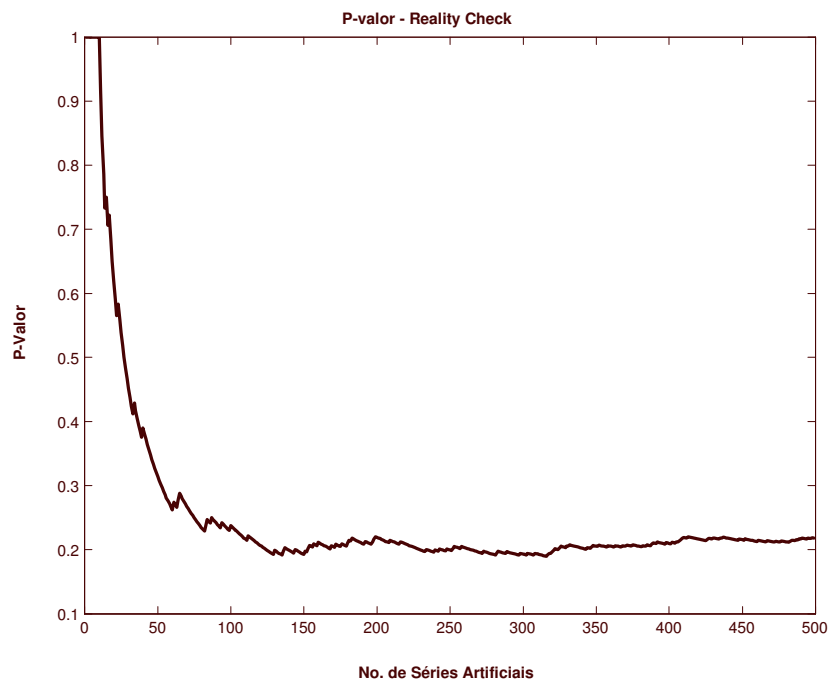


Figura 8: Teste de Realidade x No. de Reamostragens

6 Apêndice

Abaixo seguem os parâmetros utilizados para construção das estratégias de análise técnica:

1. *Média Móvel*

$$c = [1 \ 2 \ 3 \ 5 \ 8 \ 10 \ 13 \ 15 \ 20 \ 21 \ 25 \ 30 \ 40 \ 50 \ 75 \ 100 \ 150 \ 200 \ 250]$$

$$l = [2 \ 3 \ 5 \ 8 \ 10 \ 13 \ 15 \ 20 \ 21 \ 25 \ 30 \ 40 \ 50 \ 75 \ 100 \ 150 \ 200 \ 250]$$

$$f = [0 \ 0.001 \ 0.005 \ 0.01 \ 0.05]$$

2. *RSI, Williams%R e W%R**

$$p = [1 \ 2 \ 3 \ 5 \ 8 \ 10 \ 13 \ 15 \ 20 \ 21 \ 25 \ 30 \ 40 \ 50 \ 75 \ 100 \ 150 \ 200 \ 250]$$

$$f = [0 \ 0.001 \ 0.005 \ 0.01 \ 0.05]$$

$$Inf = [5 \ 10 \ 15 \ 20 \ 25 \ 30] , Sup = [70 \ 75 \ 80 \ 85 \ 90]$$

3. *Estocástico*

$$p = [3 \ 5 \ 8 \ 13 \ 21 \ 50 \ 75 \ 100 \ 150 \ 200 \ 250]$$

$$d = [5 \ 8 \ 13 \ 21 \ 50 \ 75 \ 100 \ 150 \ 200 \ 250]$$

$$f = [0 \ 0.001 \ 0.005 \ 0.01 \ 0.05]$$

$$Inf = [10 \ 20 \ 30] , Sup = [70 \ 80 \ 90]$$

4. *MACD*

$$k = [3 \ 5 \ 8 \ 13 \ 21 \ 50 \ 75 \ 100 \ 150 \ 200 \ 250]$$

$$s = [5 \ 8 \ 13 \ 21 \ 50 \ 75 \ 100 \ 150 \ 200 \ 250]$$

$$d = [5 \ 8 \ 13 \ 21 \ 50 \ 75 \ 100 \ 150 \ 200 \ 250]$$

$$f = [0 \ 0.001 \ 0.005 \ 0.01 \ 0.05]$$