UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DE RIBEIRÃO PRETO
DEPARTAMENTO DE CONTABILIDADE
PROGRAMA DE PÓS-GRADUAÇÃO EM CONTROLADORIA E CONTABILIDADE

JOSÉ RAFAEL PEREIRA

Estudo de correlações não lineares entre variações do Índice da Bolsa de Valores de São Paulo (IBOVESPA) e variações de preço de ações

Orientador: Prof. Dr. Evandro Marcos Saidel Ribeiro

RIBEIRÃO PRETO

2010

Prof. Dr. João Grandino Rodas
Reitor da Universidade de São Paulo

Prof. Dr. Sigismundo Bialoskorski Neto
Vice-diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Profa. Dra. Adriana Maria Procópio de Araújo
Chefe do Departamento de Contabilidade

JOSÉ RAFAEL PEREIRA

Estudo de correlações não lineares entre variações do Índice da Bolsa de Valores de São Paulo (IBOVESPA) e variações de preço de ações

Dissertação apresentada ao Programa de Pós-Graduação em Controladoria e Contabilidade da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo como requisito para obtenção do título de Mestre em Ciência.

Orientador: Prof. Dr. Evandro Marcos Saidel Ribeiro

RIBEIRÃO PRETO
2010

FICHA CATALOGRÁFICA

Pereira, José Rafael
Estudo de correlações não lineares entre variações do Índice da Bolsa de Valores de São Paulo (IBOVESPA) e variações de preços de ações. Ribeirão Preto, 2010.
67 p. : il. ; 30cm

Dissertação de Mestrado, apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto/USP. Área de concentração: Controladoria e Contabilidade.
Orientador: Ribeiro, Evandro Marcos Saidel

1. CAPM. 2. Correlação 3. Entropia. 4. Informação mútua.

FOLHA DE APROVAÇÃO

Nome: Pereira, José Rafael
Título: Estudo de correlações não lineares entre variações do Índice da Bolsa de Valores de São Paulo (IBOVESPA) e variações de preço de ações

Dissertação apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo para obtenção do título de Mestre em Ciências

Aprovado em:

Banca Examinadora

Prof. Dr. ________________________________ Instituição: ______________________________

Julgamento: _____________________________Assinatura: _______________________________

Prof. Dr. ________________________________ Instituição: ______________________________

Julgamento: _____________________________Assinatura: _______________________________

Prof. Dr. ________________________________ Instituição: ______________________________

Julgamento: _____________________________Assinatura: _______________________________

Dedico este trabalho a Deus e à minha família.

AGRADECIMENTOS

Agradeço a Deus, por todas as oportunidades e esperanças.

Ao meu pai, Osvaldo Furtado Pereira, e à minha mãe, Marlene Aparecida Tozetti Pereira.

Ao meu orientador, Prof. Dr. Evandro Marcos Saidel Ribeiro.

“Coragem é a resistência ao medo, domínio do medo, e não a ausência do medo”

Mark Twain

RESUMO

Pereira, J. R. **Estudo de correlações não lineares entre variações do Índice da Bolsa de Valores de São Paulo (IBOVESPA) e variações de preço de ações.** 2010. 67 f. Dissertação (Mestrado em Controladoria e Contabilidade) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, São Paulo, 2010.

Estudos de correlação entre variações de preços de ações e variação de índices de mercado são importantes na compreensão da relação entre o retorno e o risco envolvido na alocação de recursos (investimentos). De acordo com o risco envolvido, deve haver um adequado retorno. Esta questão é abordada pelo modelo CAPM – *Capital Asset Pricing Model* –, que parte da premissa de que o risco sistemático de um ativo pode ser mensurado pela sua sensibilidade aos movimentos do mercado, e para isso se supõe que os retornos dos títulos são linearmente relacionados às flutuações de um índice de mercado amplo com um grau conhecido de sensibilidade. No entanto, pode haver relações não lineares entre os retornos dos títulos e as flutuações do índice de mercado. Sendo assim, o presente trabalho analisa uma medida de correlação global vinda da teoria da informação, que mensura qualquer tipo de relação entre duas variáveis, isto é, lineares e não lineares. O objetivo é mostrar a presença de correlações não lineares no mercado de capitais brasileiro. Demonstra-se que a correlação global é expressiva e maior ou igual à correlação linear em toda a amostra constituída de todas as ações que se mantiveram no Índice da Bolsa de Valores de São Paulo (IBOVESPA) de maio de 2001 a abril de 2008, totalizando 84 meses (7 anos).

Palavras-chave: CAPM. Correlação. Entropia. Informação mútua.

ABSTRACT

Pereira, J. R. **Nonlinear correlations among variations of São Paulo Exchange Index (IBOVESPA) and stock price variations.** 2010. 67 f. Dissertação (Mestrado em Controladoria e Contabilidade) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, São Paulo, 2010.

Correlations among stock price variations and stock market indices variations are important in understanding the relationship between return and risk involved in the allocation of resources (investments). According to the risk involved there exists an appropriate return. This issue is addressed by the CAPM – Capital Asset Pricing Model –, based on the premise that the systematic risk of an asset can be measured by its sensitivity to market movements and it is assumed that the returns are linearly related to the fluctuations of a market index with a known degree of sensitivity. However, nonlinear relationships may occur. Thus, the present study analyzes a global measure of correlation of information coming from theory, which measures any type of relationship between two variables, i.e. linear and nonlinear. The goal here is to show the presence of nonlinear correlations in the Brazilian capital market. The overall correlation obtained is expressive and greater than the linear correlation across the sample of 33 stock assets from the theoretical portfolio of São Paulo Exchange Index (IBOVESPA), from May 2001 to April 2008, totaling 84 months (7 years).

Keywords: CAPM. Correlation. Entropy. Mutual information.

LISTA DE FIGURAS

LISTA DE GRÁFICOS

LISTA DE QUADROS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO

1.1 Contextualização

Um dos interesses da área de finanças é compreender a relação entre o retorno e o risco envolvido na alocação de recursos (investimentos). Nesse contexto, a teoria de equilíbrio de mercado em condições de risco, o *Capital Asset Pricing Model* (CAPM) de Sharpe (1964), Lintner (1965) e Mossin (1966) é uma alternativa muito utilizada na contabilidade para mensurar o custo de capital próprio, ou seja, o retorno mínimo exigido que um dado investimento deve proporcionar em função do seu nível de risco.

O modelo CAPM, desenvolvido com base no trabalho de Markowitz (1952), tornou-se consagrado em finanças por sua sólida fundamentação teórica. Sharpe (1964) afirma que os preços dos ativos se ajustam de acordo com o risco, ou seja, o modelo considera que o retorno esperado dos ativos está relacionado ao risco de mercado (risco sistemático). Como demonstrado por Markowitz (1952), parte do risco total inerente em um ativo pode ser eliminado pela diversificação, de forma que seu risco total não é a influência pertinente em seu preço, sendo, dessa forma, relevante o componente do risco específico (não diversificável ou sistemático), pois um ativo que oscila independentemente do mercado (carteira de mercado) terá unicamente risco diversificável (não sistemático), isto é, risco específico da empresa, que pode ser diversificado. Assim, pelo modelo CAPM, o risco sistemático de um ativo pode ser mensurado pela sua sensibilidade aos movimentos do mercado (carteira de mercado estimada com algum índice de mercado, como o Índice da Bolsa de Valores de São Paulo, IBOVESPA); para isso, Sharpe supõe que os retornos dos títulos são linearmente relacionados às flutuações de um índice de mercado amplo com um grau conhecido de sensibilidade. Essa sensibilidade é identificada com o coeficiente angular da reta de regressão linear (beta), que é dado pela covariância dos retornos do ativo com os retornos da carteira de mercado dividido pela variância dos retornos da carteira de mercado (DAMODARAN, 2007).

Para que o cálculo da reta de regressão linear seja coerente, é necessário que exista uma correlação linear significativa entre o ativo e a carteira de mercado; do contrário, a reta de regressão linear não é válida (TRIOLA, 2005). Nesse ponto, Dionísio, Menezes e Mendes (2004) salientam que o coeficiente de correlação linear, comumente utilizado, requer para a sua aplicação uma pura relação linear entre as variáveis em questão. No entanto, essa estatística pode não ser útil quando existir algum tipo de não linearidade entre os dados. Diante disso, este trabalho utilizará uma medida de correlação global para mensurar a relação entre as variáveis retornos do ativo e retornos da carteira de mercado (IBOVESPA). Este coeficiente de correlação global é originado da medida padronizada de informação mútua. A informação mútua é uma medida vinda da teoria da informação ou, mais especificamente, do conceito de entropia. Dionísio, Menezes, e Mendes (2004) afirmam que a principal vantagem da aplicação da informação mútua em séries temporais financeiras é o fato de que essa medida capta, em nível global, a dependência (linear e não linear) sem qualquer requisito sobre a distribuição de probabilidade ou de modelo específico de dependência.

1.2 Formulação do problema

O problema de pesquisa que se levanta neste estudo é: **Correlações não lineares são relevantes no mercado brasileiro?**

1.3 Objetivos

Os objetivos deste trabalho são:

- Mensurar os coeficientes de correlações lineares e não lineares entre as variações de preço de ações e variações do IBOVESPA;
- Comparar os dois tipos de coeficientes, linear (r de Pearson) e não linear (λ), para analisar a discrepância entre eles.

1.4 Hipótese

O presente trabalho está baseado na seguinte hipótese:

A correlação não linear entre variações de preço de ações e variações do IBOVESPA é, em média, maior do que a correlação linear.

1.5 Justificativa

O uso de uma medida que captura qualquer tipo de correlação entre duas variáveis é importante para a compreensão da relação entre estas variáveis. Além disso, o uso de tal medida pode ser aplicado a todo tipo de série temporal financeira, pois não há qualquer requisito sobre a distribuição de probabilidade (o método utilizado é não paramétrico) ou de modelo específico de dependência sobre os dados, como afirmam Dionísio, Menezes, e Mendes (2004).

Pela razão de o modelo CAPM ser muito utilizado na contabilidade para a mensuração do custo do capital próprio, sendo esse modelo embasado na correlação linear entre o retorno de um ativo e do índice de mercado, torna-se relevante o conhecimento da correlação global entre os dados para saber se o uso do CAPM não irá subestimar a verdadeira relação entre as variáveis.

Até o momento, não foi encontrado nenhum estudo que aplique este coeficiente de correlação global ao mercado de capitais brasileiro, o que faz com este trabalho seja uma contribuição original para a pesquisa na área de finanças.

1.6 Estrutura do trabalho

O presente trabalho está organizado em cinco capítulos, sendo o segundo uma exposição sobre a teoria de finanças, na qual são tratados os conceitos de incerteza e risco, e sua forma de mensuração mais comum, dada pelo modelo CAPM, que se baseia na análise de regressão linear.

Serão abordados, também, os conceitos sobre regressão linear e as condições que devem ser satisfeitas para a sua realização. Isso é importante para entender quando realmente é possível ajustar uma reta em um conjunto de pontos dispersos. Uma das condições que precisa ser satisfeita é a existência de uma correlação linear significativa (mensurada pelo coeficiente de correlação linear de Pearson) entre os retornos do ativo e do IBOVESPA; no entanto, pode haver correlações não lineares que não são capturadas pelo coeficiente de Pearson, o que pode comprometer a mensuração do risco sistemático. Esse capítulo ainda faz uma exposição sobre a teoria da informação e o conceito de entropia, necessário para o cálculo da informação mútua e do coeficiente de correlação global. O coeficiente de correlação global mensura qualquer tipo de correlação entre duas variáveis; no caso em questão, variações de um ativo e variações do IBOVESPA. Esse coeficiente permitirá a comparação com o coeficiente de correlação linear para analisar a diferença e a magnitude entre eles. E para finalizar esse capítulo serão mostradas as relações entre as medidas obtidas pela entropia e as medidas obtidas pela regressão linear, suas similaridades e diferenças. O capítulo 3 trata da amostra e de suas propriedades e o método utilizado para os cálculos. Foram consideradas todas as ações que se mantiveram no IBOVESPA de maio de 2001 a abril de 2008, totalizando 84 meses (7 anos). O capítulo 4 apresenta os resultados dos coeficientes de correlação global e linear. É realizado, nesse capítulo, um teste de hipótese para verificar se a correlação não linear entre variações de preço de ações e variações do IBOVESPA é, em média, maior do que a correlação linear. Por fim, o último capítulo apresenta as conclusões.

2 REFERENCIAL TEÓRICO

2.1 Teoria de finanças

O conceito moderno de finanças surgiu nos anos 1950 (até então, só havia uma noção intuitiva da existência do risco, não existia instrumentos que possibilitassem um tratamento analítico não determinista) e vem se tornando uma das áreas mais produtivas em termos de avanços teóricos e empíricos da Economia, pela grande quantidade e pela qualidade da produção científica.

Uma das atuações das finanças é estudar o processo pelo qual certos mercados tratam com o problema da transferência de fluxos de caixa de uma data para outra. O que possibilita essa transferência de recursos são os mercados financeiros, que, de outra forma, permitem aos seus agentes ajustarem seus padrões temporais de consumo e às empresas ajustarem seus padrões de investimento (ROSS, 1995).

A definição de Ross (1995) não fala do risco que está envolvido nessas transferências de fluxos de caixa, ou seja, a ideia da incerteza e risco que permeia toda decisão de investimento. A crescente relevância da incerteza e do risco nas transações econômicas fez com que os estudos em finanças se orientassem no sentido de reduzir tais incertezas e riscos ou compreendê-los com o principal objetivo de auxiliar a tomada de decisão, na busca de maximizar uma determinada função-objetivo, que pode ser a utilidade ou o retorno esperado de um investimento, entre outras. Siqueira (1999, p. 2) enfatiza que "não seria nenhum exagero dizer que a teoria de finanças moderna nasceu em virtude da preocupação dos acadêmicos de buscar uma estrutura conceitual e lógica que auxiliasse a tomada de decisões em condição de incerteza e de risco". A incerteza e o risco envolvidos na tomada de decisão se tornaram um desafio na teoria de finanças contemporânea. Isso se deve à grande capacidade que a incerteza e o risco têm de afetar as decisões dos agentes. Sassatani (1999) afirma que a incerteza pode ser vista como um elemento que influencia a atitude dos agentes econômicos. Como alguns exemplos, têm-se a teoria do portfólio de Markowitz (1952), a teoria do investimento e custo de capital de Modigliani e Miller (1958), a teoria de equilíbrio de mercado em condições de risco (CAPM – *Capital Asset Pricing Model*) de Sharpe (1964), *Value at Risk* (VAR) (JORION, 2003), entre outros. Tais modelos tiveram, e ainda têm, grande relevância pelas suas contribuições na teoria de finanças. Nas palavras de Menezes (2002, p. 48), "esses novos modelos mudaram a abordagem das finanças corporativas, transformando-as, a partir de uma versão eminentemente descritiva, em outra voltada ao rigor analítico e à incorporação de características dos investidores e do mercado".

O trabalho seminal de Markowitz (1952) trouxe um grande avanço em finanças. A sua grande contribuição foi mostrar a distinção entre a variabilidade do retorno de um ativo financeiro e o seu efeito no risco de uma carteira de ativos financeiros. Brealey, Myers e Allen (2008, p. 159) salientam que o texto de Markowitz (1952) "chamou a atenção para a prática comum da diversificação das carteiras, e mostrou como um investidor pode reduzir o desvio-padrão do retorno da carteira mediante a escolha de ações cujas oscilações não sejam exatamente paralelas”. Markowitz (1952) mostrou que se duas ações ou mais forem negativamente correlacionadas, ou seja, se uma valorizar enquanto a outra desvaloriza ao combinarmos ambas numa carteira, o risco da carteira será inferior aos riscos das ações individualmente.

Sharpe (1964) apresentou o modelo CAPM de determinação de preço de ativos causando grande impacto e avanço na área de finanças. Esse modelo estabelece uma relação linear entre risco e retorno para todos os ativos, possibilitando determinar, para cada nível de risco, a correspondente taxa de retorno que o premia. Brealey, Myers e Allen (2008) salientam que o CAPM representou um avanço significativo para a compreensão do efeito do risco no valor de um ativo, no entanto, há muitas incógnitas estatísticas e teóricas em aberto.

O modelo CAPM tem sofrido muitas críticas a respeito de sua eficiência; isso teve início na década de 1990 (BREALEY; MYERS E ALLEN 2008). Fama e French (1992) testaram de várias formas o modelo CAPM e concluíram que a média do retorno das ações não está positivamente relacionada aos betas do mercado. Essa conclusão mostrou que o CAPM não sustenta sua previsão mais básica, que afirma ser a média do retorno das ações positivamente relacionada aos betas do mercado. Apesar disso, não se pode afirmar que o CAPM não é válido, pois todos os testes empíricos têm problemas de aplicação, como, por exemplo, a dificuldade de os retornos dos títulos serem linearmente relacionados às variações de um índice de mercado, entre outras que não satisfazem os pressupostos do modelo CAPM que serão vistas na seção 2.1.4.

Apesar das críticas ao modelo CAPM, ele “permanece como o modelo mais utilizado no mercado de capitais para o cálculo do retorno exigido pelos acionistas de uma empresa, de maneira a compensá-los pelo risco de seu investimento” (FAMÁ; BARROS; SILVERIA, 2001, p. 72).

Nas próximas três seções, serão apresentados, com base na literatura, os conceitos de incerteza e risco, e seus tipos, conceitos estes muito importantes no modelo CAPM, bem como para a compreensão deste trabalho.

2.1.1 Incerteza e risco

Siqueira (2003) afirma que é comum encontrar na Economia o termo risco como sinônimo de incerteza. No entanto, o economista Knight (1921, p. 103, tradução nossa), definiu risco e incerteza de modo distinto:

> A diferença prática entre as duas categorias, risco e incerteza, é que na primeira a distribuição dos resultados em um grupo de casos é conhecida (quer por meio do cálculo a priori, quer a partir de estatísticas de experiências passadas), enquanto que no caso da incerteza isso não acontece; a razão disso, em geral, é a impossibilidade de se formar um grupo de casos, porque a situação tratada é única.

De acordo com a definição de Knight, o que diferencia incerteza de risco é a existência ou não de uma distribuição de probabilidade de um dado evento. Assim, o risco pode ser compreendido nos casos em que se podem determinar os possíveis resultados e mensurar suas respectivas probabilidades de ocorrência. Mais especificamente, a ideia de risco está associada às probabilidades de acontecerem eventos que se desviem de um determinado valor esperado (média). Ao contrário, a incerteza se encontra naqueles casos em que não se conhece a distribuição de probabilidade dos resultados.

Neste trabalho, o risco deve ser sempre entendido do ponto de vista do investidor, ou seja, aquele que está investindo ou interessado em investir em uma empresa, comprando ações dela no mercado. Nesse ponto, Damodaran (2007) argumenta que o risco de um investimento deve ser percebido por aqueles que investirão na empresa. O autor diz ainda que o investidor tem de ser aquele com mais chances de negociar a ação a qualquer tempo, isto é, o investidor marginal.

2.1.2 Risco de mercado

Damodaran (2007) afirma que o retorno esperado sobre um investimento deve ser uma função do “risco de mercado” embutido neste investimento. Assim, empresas mais arriscadas devem oferecer maiores retornos (maiores prêmios pelo risco) como forma de remunerar o seu risco, já que o risco de mercado é um tipo de risco financeiro, e afeta o custo de capital médio ponderado *Weighted Average Cost of Capital* (WACC) de uma empresa (ASSAF NETO, 2005a), o que está de acordo com o CAPM.

O risco de mercado decorre da existência de outros riscos presentes na economia e que afetam os negócios (BREALEY; MYERS e ALLEN, 2008), como as variações em fatores de mercado, como taxas de juros, taxas de câmbio, preços de *commodities* e ações.

Goulart (2003) salienta que o risco de mercado ocorre quando há variação no valor de um ativo decorrente de oscilações nas taxas de juros ou de câmbio, alterações na oferta e na demanda, descasamentos de prazo entre ativos e passivos, indexadores (moedas), entre outros.

2.1.3 Risco sistemático e não sistemático

Em análise de investimento, o risco total inerente aos retornos das ações pode ser decomposto em dois fatores, o risco não sistemático e o risco sistemático (SHARPE, 1964). O risco não sistemático (ou diversificável) é o risco próprio do ativo, e é mensurado pela dispersão dos retornos do título em relação aos movimentos do retorno da carteira de mercado; é peculiar de um ativo, não pode ser atribuído a causas exteriores, e pode ser eliminado pela diversificação. Já o risco sistemático é o risco que não se pode eliminar pela diversificação, pois é determinado por elementos de natureza política, econômica e social, ou seja, esse risco pode estar relacionado a alguns índices ou o retorno de uma carteira de mercado (DAMODARAN, 2007).

2.1.4 Modelo CAPM (*Capital Asset Pricing Model*)

O CAPM surge para mensurar a influência do risco no comportamento dos preços de ativos. Sharpe (1964) afirma que os preços dos ativos se ajustam de acordo com o seu risco. Para isso acontecer, é necessário que exista uma taxa de juros pura (livre de risco) estabelecida pela interação das preferências dos indivíduos do mercado, de forma que um prêmio de risco de mercado seja definido para se incorporar de alguma forma no preço dos ativos conforme o seu risco. A ideia do prêmio pelo risco é que maiores taxas esperadas de retornos podem ser obtidas quando se incorrem em riscos maiores.

Sharpe (1964) estabelece uma teoria de equilíbrio de mercado dos preços dos ativos em condições de risco, obtendo assim a relação entre o preço de um ativo e os vários componentes de seu risco total, que é o próprio modelo CAPM.

Em finanças, é comum mensurar o risco sistemático de um ativo pelo cálculo do beta, pois, conforme já dito, o beta de um ativo mensura a sua exposição a todo o risco de mercado, de acordo com as palavras de Assaf Neto (2005b, p. 366):

> O modelo do CAPM exprime o risco sistemático de um ativo pelo seu coeficiente beta, identificado com o parâmetro angular na reta de regressão linear (reta característica). Admite-se que a carteira de mercado, por conter unicamente risco sistemático (o risco não sistemático foi todo eliminado pela diversificação), apresenta um beta igual a 1,0.

Matematicamente, o beta é calculado pela seguinte expressão algébrica:

$$\beta_i = \frac{Cov_{i,m}}{\sigma_m^2}. \tag{1}$$

Sendo,

$Cov_{i,m}$: covariância do ativo i com a carteira de mercado m.

σ_m^2: variância da carteira de mercado.

Quando o beta de um ativo for exatamente igual a um, implica que a covariância do ativo i com a carteira de mercado m é igual à variância da carteira de mercado m. O que significa que o risco do ativo é igual ao risco sistemático do mercado. Se o beta for maior que 1, quer dizer que o risco sistemático do ativo é maior do que o da carteira de mercado. Se, no entanto, o beta for menor que 1, então o ativo possui um risco sistemático menor do que o da carteira de mercado.

Outra forma de obter o beta de um ativo é por meio da reta de regressão linear simples. Como o coeficiente beta mede o risco sistemático de um título, essa mensuração é dada pela inclinação (parâmetro angular) da reta de regressão linear entre os retornos do título e os retornos da carteira de mercado, que correspondem a algum índice de bolsa, por exemplo, o IBOVESPA (ASSAF NETO, 2005b).

Esta metodologia consiste em determinar uma reta que melhor se ajuste aos pontos que representam os retornos do preço de um determinado ativo (eixo vertical) e os retornos dos pontos de algum índice de mercado (eixo horizontal). A regressão linear tem como objetivo explicar o comportamento de uma variável, definida como variável dependente (retornos dos preços de alguma ação), a partir de outra variável, chamada variável independente (retornos dos pontos de algum índice de mercado). A partir de um gráfico de dispersão em que a variável x é o retorno do mercado e a variável y é o retorno da ação, pode-se observar se existe relação linear entre x e y. A regressão linear ajusta uma reta aos pontos de tal modo a minimizar a soma dos desvios quadrados dos pontos (coordenadas) à linha. A inclinação da reta obtida pela regressão é o beta.

Ao se realizar a regressão, obtém-se também o R^2 (R-quadrado) da regressão. Estatisticamente, o R^2 mensura a precisão de ajuste da regressão, isto é, mede a proporção da variabilidade da variável dependente, que é explicada pela variável independente.

Em finanças, o R^2 pode ser interpretado, de acordo com Damodaran (2002, p. 94), como "uma estimativa da proporção de risco (variância) da empresa que pode ser atribuído ao risco de mercado; o saldo $(1 - R^2)$ pode então ser atribuído ao risco específico da empresa". O risco específico a que Damodaran se refere, é exatamente o risco não sistemático (ou diversificável). Portanto, $(1 - R^2)$ mensura a proporção do risco da empresa, que é diversificável. Assaf Neto (2005b, p. 376) salienta que:

> Em termos financeiros, R^2 permite que se conheça a parte do risco de uma empresa explicada pelas condições de mercado, o denominado risco sistemático (taxas de juros, política econômica etc.), e a parcela decorrente de variáveis específicas de uma empresa $1 - R^2$, conhecida por risco não sistemático ou diversificável.

O coeficiente β mede a relação obtida entre as variações das cotações das ações de uma dada empresa y_i, e de certo índice de mercado x_i, ou seja, o β calcula o efeito da carteira de mercado sobre as ações de uma empresa. Isso significa, como explica Assaf Neto (2005b, p. 367), que "se o $\beta =$1,30, a valorização média de 10% na carteira de mercado determina uma expectativa de rentabilidade de 13% na ação".

O CAPM está fundamentado nos seguintes pressupostos de preferência dos agentes que atuam no mercado de capitais (SHARPE, 1964):

1) Preocupação apenas com o retorno esperado e o desvio-padrão, não importando a liquidez, ou todos os ativos têm a mesma liquidez;
2) Preferência por maior retorno esperado e menor risco;
3) Preferência por carteiras ótimas;

4) Os investidores usam a mesma informação e interpretam da mesma forma; as expectativas são iguais;
5) Os ativos são perfeitamente divisíveis;
6) Existe um ativo sem risco e qualquer um pode comprá-lo ou vendê-lo da forma que lhe convier;
7) Não há tributos ou custos (ou são iguais para todos).

A primeira condição que deve ser verificada para que o modelo CAPM seja válido é se todos os investidores estão preocupados apenas com o retorno esperado e o desvio-padrão. Isso sempre será verdade se a distribuição dos retornos dos ativos for normalmente distribuída, já que a média e o risco são as únicas duas medidas que um investidor precisa conhecer para tomar sua decisão de investimento.

O segundo e o terceiro pressupostos afirmam, respectivamente, que a preferência por maior retorno esperado e menor risco é preferível, e que os investidores têm preferência por carteiras ótimas. Como mostrado por Markowitz (1952), a fronteira eficiente de investimento contém todas as carteiras que apresentam os maiores retornos esperados para cada nível de risco assumido. Essas carteiras serão preferíveis às outras que apresentam o mesmo nível de risco para um retorno esperado menor.

Em um mercado em que os investidores preferem as carteiras ótimas (carteiras diversificadas), isto implica que os agentes se preocupam principalmente com os riscos que não podem ser eliminados por meio da diversificação. Sabe-se que, por causa da diversificação, o risco de uma carteira é inferior à média dos riscos das ações individuais (BREALEY; MYERS; ALLEN, 2008). Esta ideia dá mais respaldo ao modelo CAPM, pois este considera apenas o risco não diversificável. Brealey, Myers e Allen (2008) afirmam que se não fosse assim, sempre que duas empresas se fundissem com o objetivo de reduzir seus riscos, pela diversificação, o preço das suas ações aumentaria. Mas nenhuma fusão com o objetivo de diversificação (diluir os riscos) aumenta o preço das ações.

O pressuposto quatro afirma que os investidores usam a mesma informação e a interpretam da mesma forma, ou seja, as expectativas são homogêneas. Sobre isso, Assaf Neto (2005a) salienta que os investidores possuem a mesma percepção com relação ao desempenho dos ativos, formando carteiras eficientes a partir de expectativas idênticas.

O pressuposto cinco afirma que os ativos são perfeitamente divisíveis, o que está de acordo com a construção da carteira eficiente em que as proporções de cada ativo na carteira são determinadas conforme a obtenção da carteira ótima, e o tamanho do lote em que ele é vendido no mercado.

O pressuposto seis, da existência de um ativo sem risco, é fundamental para a construção do modelo CAPM, pois somente existindo este ativo e sendo possível sua negociação no mercado será possível um referencial de comparação com os ativos de risco. Por último o pressuposto sete considera a inexistência de tributos ou custos, ou que eles são iguais para todos. Esta condição implica que o equilíbrio determinado no modelo CAPM será alcançado sem interferências exteriores ao mercado.

Como visto acima, o risco sistemático (beta) tem várias implicações importantes, no entanto é preciso cuidado ao realizar seu cálculo, pois para um bom ajuste de uma reta entre duas variáveis é necessário que elas se relacionem de forma linear, isto é, a correlação linear entre x e y deve ser significativa.

2.2 Correlação e regressão linear

2.2.1 Coeficiente de correlação linear de Pearson

O coeficiente de correlação linear r de Pearson "mede a intensidade da relação linear entre os valores quantitativos emparelhados x e y em uma amostra" (TRIOLA, 2005, p. 382) e é dado por:

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}. \quad (2)$$

x_i: cada um dos eventos x de 1 a n.

$\bar{x}$: média aritmética dos n eventos x.

y_i: cada um dos eventos y de 1 a n.

$\bar{y}$: média aritmética dos n eventos y.

O coeficiente r varia entre -1 e 1, sendo que quando ele é igual a -1 significa que as variáveis x e y possuem uma correlação negativa perfeita, ou seja, quando x cresce, y decresce de forma perfeitamente previsível. Se r for nulo, significa que não há relação entre x e y, pois conforme x aumenta, y não apresenta tendência para aumentar ou diminuir. E por fim, se r for igual a 1, significa que as variáveis x e y possuem uma correlação positiva perfeita, ou seja, quando x cresce, y cresce de forma perfeitamente previsível. Quando r assumir valores próximos de 1, por exemplo, 0,80, significa que os dados de x e y não se posicionam todos sobre uma mesma reta, formando, portanto, uma relação linear não perfeita.

É importante deixar claro que a relação entre x e y é descrita como associação linear apenas, e não causas e efeitos. Isso, pois somente o cálculo da correlação não prova se existe ou não causa e efeito, isto é, que a variação de uma variável x causou alteração na variável y. Pode acontecer de haver uma grande correlação entre as variáveis consideradas, mas essa ser provocada por uma terceira variável não considerada (LEVINE et al., 2008).

2.2.2 Regressão linear

Esta seção foi baseada em Levine et al. (2008). A análise de regressão possibilita identificar o tipo de relação matemática que há entre uma variável independente e uma variável dependente, e permite também quantificar o efeito que a variável independente exerce sobre a variável dependente.

Pode haver diversos tipos de relação entre duas variáveis, desde uma relação simples e direta até relações extremamente complexas. A regressão linear consiste em uma relação entre uma variável independente numérica x, utilizada para prever a variável dependente numérica y, desde que a influência de x sobre y possa ser captada por uma reta. Ou seja, a relação entre as variáveis x e y, para qualquer observação i, é dada por:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i. \tag{3}$$

Sendo que:

y_i: variável dependente ou de resposta para a observação i.

β_0: intercepto de y quando $x = 0$.

β_1: inclinação da reta.
x_i: variável independente ou explanatória para a observação i.
ε_i: erro aleatório em y para a observação i.

A inclinação da reta β_1 capta a variação esperada em y por variação unitária em x, isto é, β_1 representa a média aritmética do total que y varia para uma variação unitária (de uma unidade) em x. Da mesma forma, β_0 é a média aritmética do valor que y assume quando $x = 0$. Finalmente, ε_i é o erro aleatório em y para cada observação i, ou seja, é a distância vertical do verdadeiro valor de y_i abaixo ou acima do seu valor previsto pela reta. Estas propriedades podem ser vistas na Figura 1 abaixo.

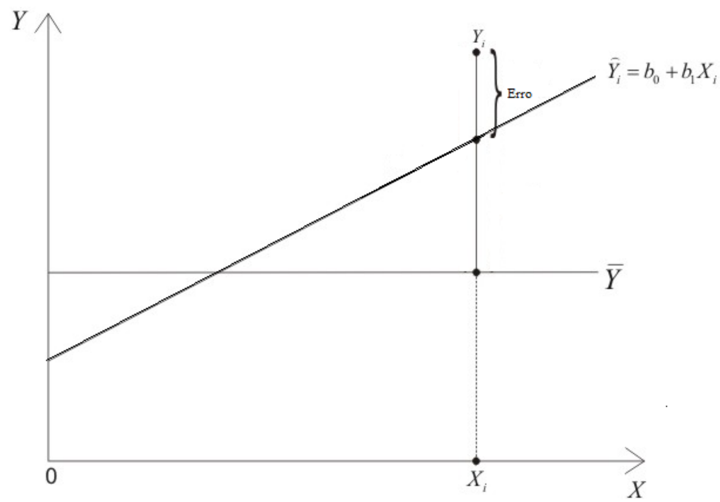

Figura 1 – Erro da estimação do verdadeiro valor de y

2.2.3 Estimativa da equação de regressão linear

A equação (3) representa um modelo estatístico supondo que seja o mais adequado para representar dados que variam conforme uma reta. Para obtê-lo, seria preciso analisar toda a população dos eventos de interesse. Como isso na prática pode inviabilizar ou dificultar a análise de regressão, considera-se que a amostra foi selecionada aleatoriamente dentro dos eventos de interesse para calcular b_0 e b_1 como estimadores dos parâmetros da população β_0 e β_1 da equação (3), respectivamente.

Os dados amostrais devem satisfazer os quatro seguintes pressupostos:

1) Linearidade;
2) Normalidade de erros;
3) Independência de erros;
4) Homocedasticidade ou igualdade de variâncias.

O pressuposto um da linearidade está diretamente suportando o uso da regressão linear simples, pois ele assegura que a relação entre as variáveis é linear.

O pressuposto dois, normalidade de erros, exige que os erros ε_i para cada valor de x tenham distribuição normal. Todavia, desde que a distribuição não seja extremamente diferente da distribuição normal, as inferências de β_0 e β_1 não serão seriamente prejudicadas, pois a regressão se mostra robusta relativamente ao distanciamento da normalidade de seus erros.

O pressuposto três, independência dos erros, é sempre relevante na coleta de dados de um determinado período de tempo, pois ele requer a independência dos erros ε_i entre si e, consequentemente, entre os períodos. Muitas vezes, ocorre de os erros de um determinado período de tempo serem correlacionados com os erros de outro período.

O pressuposto quatro, igualdade de variâncias, requer a igualdade da variância dos erros $\sigma^2(\varepsilon_i)$ para qualquer valor de x. Isso quer dizer que tanto para valores altos como baixos de x, a variabilidade dos valores que y assume é constantes.

Então, satisfazendo os quatro pressupostos acima, pode-se usar b_0 e b_1 como estimadores dos parâmetros da população β_0 e β_1, respectivamente. Logo, a equação da regressão linear simples será:

$$\hat{y}_i = b_0 + b_1 x_i. \tag{4}$$

y_i: valor estimado de y para o evento i.

b_0: intercepto da amostra y quando $x = 0$.

b_1: inclinação da reta da amostra.

x_i: valor de x para a observação i.

2.2.4 Cálculo de b_0 e b_1 – Método dos mínimos quadrados

O método dos mínimos quadrados é uma técnica de otimização matemática que tem como objetivo encontrar o melhor ajuste de uma reta para um conjunto de dados. Isso é obtido minimizando a soma dos resíduos ao quadrado. O resíduo u é a diferença entre os reais valores de y_i e os valores previstos ou estimados de $\hat{y}_i$. Considerando uma amostra de dados com n eventos, então:

$$\sum_{i=1}^{n} u^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2. \tag{5}$$

Como $\hat{y}_i = b_0 + b_1 x_i$, então substituindo em (5) resulta:

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} [y_i - (b_0 + b_1 x_i)]^2. \tag{6}$$

Sendo assim, o método dos mínimos quadrados determina os valores de b_0 e b_1, que minimizam a soma dos resíduos u ao quadrado (u^2). Além disso, o teorema de Gauss-Markov garante que o método dos mínimos quadrados, quando os quatro pressupostos não são violados, é não viesado e de variância mínima, sendo o mais eficiente entre os estimadores lineares. Portanto, $\hat{y}_i = b_0 + b_1 x_i$ será a melhor reta que se ajusta aos dados, supondo que todos os quatro pressupostos sejam satisfeitos.

2.2.5 Medidas de qualidade do ajuste da regressão linear

São três as medidas de qualidade do ajuste da regressão linear, e sempre que se utiliza o método dos mínimos quadrados para obter os coeficientes de regressão para uma amostra, tornam-se necessários seus cálculos na regressão. Essas três medidas são denominadas:

1) Soma dos quadrados da regressão (SQReg) ou variação explicada;
2) Soma dos quadrados dos resíduos ou erros (SQR) ou variação não explicada;
3) Soma total dos quadrados (STQ) ou variação total.

A primeira medida, soma dos quadrados da regressão (SQReg), é a soma das diferenças ao quadrado entre o valor previsto (ou estimado) de y, $\hat{y}_i$, e a média aritmética do valor de y, $\bar{y}$.

Essa medida também é chamada de variação, explicada por medir a parte da variação total, que é explicada pela reta de regressão, matematicamente fica:

$$\text{SQReg} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2. \tag{7}$$

A segunda medida, soma dos quadrados dos resíduos ou erros (SQR), calcula a parte da variação em y que não é explicada pela regressão. É dada pela soma das diferenças ao quadrado entre o valor observado de y e o seu valor previsto, $\hat{y}$. Algebricamente fica:

$$\text{SQR} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2. \tag{8}$$

A terceira medida, a soma total dos quadrados (STQ), calcula a soma ao quadrado das variações dos y_i em relação à sua média aritmética, $\bar{y}$. Essa medida mensura toda a variação, desde a média dos valores de y até o real valor de y_i, e é dada por:

$$\text{STQ} = \sum_{i=1}^{n}(y_i - \bar{y})^2. \tag{9}$$

Na análise de regressão, a STQ ou variação total é composta pela SQReg (variação explicada) e pela SQR (variação não explicada) ou, matematicamente:

$$\text{STQ} = \text{SQReg} + \text{SQR}. \tag{10}$$

Ou ainda

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2. \tag{11}$$

2.2.6 Coeficiente de determinação – R^2

O conhecimento das medidas de qualidade de ajuste da regressão linear, acima discutidas, não permite saber se o ajuste da reta aos pontos da amostra está bom ou não. É necessário ter uma medida relativa, ou seja, que seja capaz de dizer o quanto a variação explicada ou a soma dos quadrados da regressão (SQReg) compõe a variação total ou soma total dos quadrados (STQ). Somente dessa forma haverá parâmetro de comparação e, portanto será possível descobrir o grau de ajuste da reta ou modelo da regressão linear, o qual recebe o nome de coeficiente de determinação, R^2; matematicamente, é:

$$R^2 = \text{SQReg}/\text{STQ}. \tag{12}$$

Ou,

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}. \tag{13}$$

A equação (13) mede o poder explicativo da regressão, sendo $0 \leq R^2 \leq 1$, ou seja, o R^2 varia de 0 a 1. Quanto mais o R^2 se aproxima de 1, melhor será o ajuste da reta, o que indica uma forte relação linear positiva entre as variáveis x e y. Por exemplo, uma reta de regressão que forneça um $R^2 = 90\%$ significa que o modelo de regressão foi capaz de explicar 90% da variabilidade de y, sendo que os outros 10% da variabilidade da amostra devidos a outros fatores que a variável x não abrange.

Na próxima seção, será apresentada uma teoria que permite calcular várias medidas estatísticas semelhantes às que foram vista nesta seção, com base no cálculo da informação mútua.

2.3 Teoria da informação

A teoria da informação permite a mensuração da incerteza. Para isso, o conceito de entropia é utilizado. A entropia está relacionada à quantidade de informação que um sistema apresenta (sem considerar a semântica da informação); quanto menos informação em um sistema, maior será a sua entropia ou a sua incerteza (PHILIPPATOS; WILSON, 1974). Shannon (1948) apresenta a entropia de uma fonte de informação como conceito-chave na teoria da informação, definindo-a como o total de incerteza sobre um evento, a partir de um conjunto de eventos possíveis, que também é conhecida como entropia ou conteúdo esperado de informação.

O conceito de incerteza dentro da teoria da informação, diferentemente da definição de Knight (1921), se refere a uma medida de quanta “escolha” está envolvida na seleção de um dado evento ou da forma como estamos incertos sobre um resultado, definição esta baseada em Shannon (1948). Para Knight (1921), a incerteza se refere àqueles casos em que não se conhece a distribuição de probabilidade dos resultados, e, por conseguinte, não é possível mensurá-la por não se conhecer as probabilidades. No caso da incerteza no contexto da teoria da informação, esta é mensurada pelo conceito de entropia, que está totalmente baseada em probabilidades.

A teoria da informação é um ramo da teoria da probabilidade e da matemática estatística que abrange sistemas de comunicação, transmissão de dados, teoria do ruído, correção de erros, etc. Ela está baseada na mensuração da incerteza ou surpresa, e para isso se fundamenta em dois conceitos próprios, o informacional e a entropia (SIQUEIRA, 1999).

A grande contribuição de Shannon (1948) foi o tratamento dado à teoria da informação, ao considerar a comunicação como um problema matemático fundamentado na estatística. Sua teoria permitiu determinar a capacidade de um canal de comunicação em termos de ocorrência de bits, aspectos relacionados com a perda de informação, transmissão de sinais e ruídos de canal, sendo que a teoria da informação, como adverte Shannon, não se preocupa com a semântica dos dados.

A importância de se utilizar a teoria da informação é ela proporcionar uma medida universal da quantidade de informação ou incerteza em termos de estados de probabilidades, não ser dependente de uma distribuição em particular e poder ser aplicada a qualquer situação que lida com incerteza e ao estudo do risco nos retornos de ativos (PHILIPPATOS; WILSON, 1974).

Vários trabalhos vêm aplicando a teoria da informação nos mercados de capitais como forma de substituir e complementar outras teorias. Por exemplo, no mercado de capitais, Fama (1965) e Philippatos e Nawrocky (1973) usam a teoria da informação para testar a hipótese de eficiência na Bolsa de Nova Iorque (N.Y.S.E.). Philippatos e Wilson (1972) utilizam a teoria da informação para a construção de carteiras eficientes de ativos. Philippatos e Nawrocki (1973) realizaram análises sobre os índices das bolsas de Nova Iorque e da *American Stock Exchanges* utilizando o conceito de entropia. Buckley (1985) discute os princípios da máxima e mínima entropia e suas aplicações na tomada de decisão sob risco. Em seu trabalho, o autor afirma que o princípio da máxima entropia é aplicável quando existe pouca informação sobre as possibilidades dos estados futuros do sistema, ou seja, quando a distribuição de probabilidade exata ao longo dos estados da natureza não é exatamente conhecida, mas alguma informação anterior está disponível sobre as possibilidades desses resultados. Fernholz (1999) analisa a entropia como medida de diversificação nos mercados financeiros. Cassetari (2003) desenvolve uma metodologia de alocação de capital baseada no princípio de máxima entropia, e nesse trabalho é utilizada a entropia de Shannon como medida de risco financeiro, comparando suas vantagens e desvantagens em relação a outras medidas mais comumente utilizadas vindas da moderna teoria das carteiras, como de Markowitz. Nesse caso, o autor conclui que, pelo fato de não se assumir qualquer distribuição de probabilidades, confere um grau muito grande de realismo à metodologia (na solução de Markowitz se assume a distribuição gaussiana). Pelo princípio de máxima entropia, o resultado será o mesmo proporcionado pela solução de Markowitz quando se assume a aproximação gaussiana. Horowitz e Horowitz (1968) utilizaram a entropia como medida de concentração industrial e competitividade de mercado.

2.3.1 O conceito de entropia

A entropia no contexto da mecânica estatística é definida como a quantidade de desordem de um sistema. Isso quer dizer que quanto maior a desordem de um sistema, maior será a sua entropia. A mecânica estatística de Boltzmann afirma que a desordem de determinado sistema é diretamente proporcional à quantidade de diferentes estados. A entropia está associada diretamente à quantidade de possíveis estados para determinado macroestado; isso quer dizer que ela depende diretamente da quantidade de microestados possíveis, o que implica em diferentes combinações do sistema. Portanto, um sistema terá sua entropia aumentada sempre quando houver um aumento da desordem do sistema, e de outro modo a entropia cresce conforme o número de trajetórias possíveis aumenta com a complexidade do sistema (HERSCOVICI, 2005).

Shannon (1948) demonstra que o total de incerteza sobre um evento x, a partir de um conjunto n de eventos possíveis, com probabilidades p_i de ocorrências – também conhecido como entropia da mecânica estatística – é definido da seguinte forma para o caso discreto:

$$H(X) = -\sum_{i=1}^{n} p(x_i) log_2 \, [p(x_i)] \text{ bit / evento} \qquad (14)$$

na qual $0log_2(0) = 0$.

A explicação para colocar $0log_2(0) = 0$ é que o limite de $p(x)log_2 p(x)$ quando $p(x)$ tende a 0 é 0. Entretanto, a função $p(x)\, ln\, p(x)$ não está definida para x igual a zero; assim, como será necessário conhecer a função quando a probabilidade $p(x)$ for zero, define-se a função como sendo igual ao seu limite.

A equação (14) é também conhecida como medida de entropia de Shannon ou conteúdo esperado de informação, e ao obtê-la Shannon tinha como objetivo determinar a capacidade do canal necessária para a transmissão de informação. A base mais comum de se usar o logaritmo é a 2, na qual a unidade de medida é o bit.

O logaritmo na base 2 determina o grau de caoticidade da distribuição de probabilidade de p_i e pode ser usada para determinar a capacidade do canal necessária para a transmissão da informação.

Uma vez que é idêntica na forma de expressão para a entropia em mecânica estatística, $H(X)$ é chamado de entropia da comunicação. Bem como uma medida de informação esperada para ser adquirida a partir do recebimento de uma mensagem x, entropia da comunicação é uma medida da média da incerteza sobre a mensagem antes da sua recepção.

2.3.2 A incerteza mensurada pela entropia

A entropia é utilizada na teoria econômico-financeira como medida de dispersão global, informação e incerteza na avaliação de ativos, carteiras de ativos e mercado. E não necessita da pressuposição de distribuição normal de probabilidade, como comumente alguns modelos necessitam (DIONÍSIO; MENEZES; MENDES, 2003).

Sejam $n_1, n_2, \dots, n_n$ eventos que formam um conjunto com as seguintes probabilidades de ocorrência $p_1, p_2, \dots, p_n$, respectivamente, sendo tais probabilidades de ocorrência tudo o que se conhece sobre o evento que irá se realizar. Sendo assim, Shannon (1948, p. 10, tradução nossa) pergunta: “É possível achar uma medida de quanta ‘escolha’ está envolvida na seleção do evento ou de como incerto se está do resultado?”. A medida que quantifica essa incerteza do conjunto é calculada pela entropia $H = H(p_1, \dots, p_n)$, satisfazendo as seguintes propriedades (SHANNON, 1948).

1) A entropia H deve ser contínua em p_i, para todo i;
2) Se todos são iguais a p_i, $p_i = 1/n$, então a entropia H deve ser uma função monótona crescente em n. Quanto maior for n, ou seja, o número de eventos possíveis, maior será a incerteza;
3) Se uma escolha for separada em duas escolhas sucessivas, a entropia H será a soma ponderada dos valores individuais de H. Como exemplifica a figura abaixo.

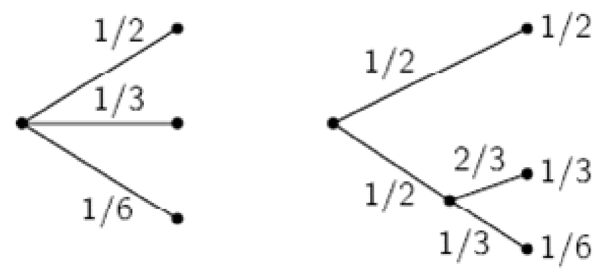

Figura 2 – Decomposição de uma escolha de três possibilidades
Fonte: Shannon (1948)

Na figura 2, o desenho à esquerda, tem três possibilidades $p_1 = 1/2$, $p_2 = 1/3$ e $p_3 = 1/6$. No da direita, primeiro tem de se escolher entre duas possibilidades, cada uma com uma probabilidade 1/2, e depois fazer outra escolha, com probabilidades 2/3 e 1/3. Os dois casos têm o mesmo resultado final, ou seja, a mesma probabilidade, isto é:

$$H(1/2\,,1/3\,,1/6) = H(1/2\,,1/2) + H(2/3\,,1/3). \qquad (15)$$

O coeficiente 1/2 se deve à segunda escolha que ocorre apenas metade do tempo (SHANNON, 1948).

O único H que satisfaz os três pressupostos acima, de acordo com Shannon (1948), isto é, a entropia do conjunto de probabilidades $p_1, p_2, \dots, p_n$ é:

$$H = -K \sum_{i=1}^{n} p_i \, log \, p_i. \qquad (16)$$

onde K é uma constante positiva. A constante K equivale a uma escolha de uma unidade de medida. O valor desta constante será igual a 1 quando a base do logaritmo for 2 (SHANNON, 1948).

A base do logaritmo não interfere no cálculo da entropia, a sua escolha irá apenas mudar a sua unidade. Sendo assim, se a base for 2, a unidade de medida será em bits, se for 10 será em dits e se for e será em nats (DIONÍSIO; MENEZES; MENDES, 2003). Shannon e Weaver (1964) justificam o uso do logaritmo por vários motivos. Em primeiro lugar, como sempre será calculado o logaritmo de uma probabilidade p_i, e como $0 \leq p_i \leq 1$ (exceto quando $p_i = 0$), o seu logaritmo será negativo e assim a entropia será positiva. Em segundo, caso $p_i = 1$, seu logaritmo será igual a 0, e a entropia nula. Isso quer dizer que quando houver um evento certo, não existirá incerteza relativa a ele. Da mesma forma, conforme a quantidade de eventos igualmente possíveis duplicar, a entropia aumentará uma unidade. Em terceiro, o logaritmo permite que eventos difíceis de ocorrer, ou seja, de pequena probabilidade tenham maior ponderação ou mais "peso". Isso permite que eventos extremos ou raros tenham maior atenção.

Em todas as situações em que há o uso de probabilidades existirá incerteza. Sendo assim, sempre que um fenômeno for descrito com probabilidades, o objetivo será saber se é possível quantificar a sua incerteza e, caso seja, de que forma. Já foi visto que a entropia H não é função de uma variável aleatória, mas sim da distribuição de probabilidade da variável aleatória, ou seja, das suas probabilidades. Sendo assim, Fieldman (1998) conclui que a partir da própria definição axiomática da entropia, a entropia de um evento x mede quantitativamente a incerteza total associada à distribuição de probabilidade do evento x.

A entropia, no caso de dois eventos independentes, x e y com probabilidades p e $q = 1 - p$ respectivamente, será:

$$H = -(p\, log_2\, p + q\, log_2\, q). \tag{17}$$

No Gráfico1, pode-se ver como se comporta a entropia dada pela equação (17), para diversos valores de p. Pode ser visto que quando as variáveis aleatórias x e y possuem as mesmas probabilidades de ocorrência, ou seja, quando $p = q = 0{,}5$, a entropia será máxima. Isso quer dizer que a incerteza será máxima quando as probabilidades dos eventos forem iguais, pois "quando todos os resultados possíveis forem equiprováveis, significando que o nível de informação disponível é mínimo" (DIONÍSIO; MENEZES; MENDES, 2003, p. 8).

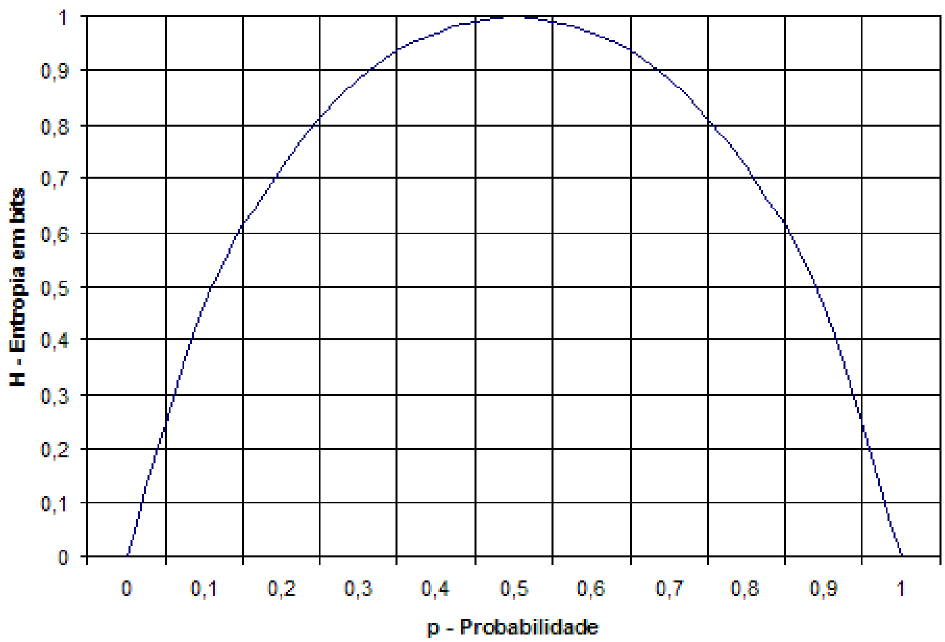

Gráfico 1 – Entropia no caso de dois eventos com probabilidades p e $q = 1 - p$

Shannon (1948) descreve as características da entropia H que reforçam a sua escolha como medida de informação ou de escolha.

1) A entropia será nula ($H = 0$) se, e somente se, todas as probabilidades p_i forem nulas, exceto quando for 1, que será igual a 0. Portanto, um evento certo terá entropia mínima e nível de informação máximo;
2) Quando todas as n variáveis aleatórias tiverem a mesma probabilidade p_i de ocorrência, isto é, $p_i = 1/n$, implicando na situação mais incerta, a entropia será:

$$H = -\sum_{i=1}^{n} \frac{1}{n} log_2 \frac{1}{n} \Leftrightarrow H = -\left[n \frac{1}{n} (log_2 1 - log_2 n)\right] \Leftrightarrow H = log_2 n. \qquad (18)$$

2.3.3 Entropia conjunta e entropia condicional

Considere dois eventos x e y, com m e n possibilidades de ocorrência, respectivamente. Suponha que $p(i, j)$ seja a probabilidade de ocorrência conjunta de i para x e j para y. A entropia do evento conjunto, de acordo com Shannon (1948), é dada por:

$$H(x,y) = -\sum_{i,j} p(i,j)\, log_2\, p(i,j). \tag{19}$$

Sendo que,

$$H(x) = -\sum_{i,j} p(i,j)\, log_2 \sum_j p(i,j) \tag{20}$$

e

$$H(y) = -\sum_{i,j} p(i,j)\, log_2 \sum_i p(i,j). \tag{21}$$

Sabendo que a incerteza de um evento conjunto é menor ou igual à soma das incertezas, verifica-se que:

$$H(x,y) \leq H(x) + H(y). \tag{22}$$

A igualdade somente se verifica na equação (22) quando os eventos x e y são independentes, ou seja, $p(i,j) = p(i)p(j)$.

Ainda considerando dois eventos x e y, não necessariamente independentes, com m e n possibilidades de ocorrência, respectivamente, para cada valor i que x pode assumir, haverá uma probabilidade condicional $p(j|i)$, na qual y tem o valor j, determinado por:

$$p(j|i) = \frac{p(i,j)}{\sum_j p(i,j)}. \tag{23}$$

A medida de incerteza de y quando x é conhecido, ou seja, a entropia condicional de y dado x é definida por Shannon (1948) como:

$$H(y|x) = -\sum_{i,j} p(i,j)\, log_2\, p(j|i). \tag{24}$$

Shannon (1948, p. 12) define a entropia condicional de y dado x, $H(y|x)$ "como a média da entropia de y para cada valor de x, ponderado em função da probabilidade de se obter esse x. Ou quanto incerto se está em média de y quando se conhece x".

Substituindo $p(j|i)$ de (23) na equação (24), obtém-se:

$$H(y|x) = -\sum_{i,j} p(i,j)\, log_2\, p(i,j) + \sum_{i,j} p(i,j)\, log_2 \sum_j p(i,j), \tag{25}$$

ou,

$$H(y|x) = H(x,y) - H(x). \tag{26}$$

A entropia ou incerteza do evento conjunto x e y é a entropia ou incerteza de x mais a entropia ou incerteza de y quando se conhece x.

De (22) e (26) tem se:

$$H(x) + H(y) \geq H(x,y) = H(x) + H(y|x). \tag{27}$$

Portanto,

$$H(y) \geq H(y|x). \tag{28}$$

A entropia ou incerteza de y nunca é aumentada pelo conhecimento de x; será menor, a menos que x e y sejam eventos independentes, caso em que não é alterado e o nível de incerteza será máximo (SHANNON, 1948). Se $H(y|x) = 0$, a incerteza é nula, pois x determina totalmente y.

2.3.4 Entropia de uma distribuição contínua

A entropia de um conjunto discreto de probabilidades foi definida na equação (16) como:

$$H = -\sum_{i=1}^{n} p_i \log p_i.$$

Shannon (1948), de forma análoga, define a entropia de uma distribuição contínua pela função de distribuição de densidade $p(x)$ por:

$$H(x) = -\int_{-\infty}^{\infty} p(x) \log p(x)\, dx. \tag{29}$$

Para uma distribuição n dimensional $p(x_1, \dots, x_n)$ tem-se:

$$H = -\int \dots \int p(x_1, \dots, x_n) \log p(x_1, \dots, x_n)\, dx_1 \dots dx_n. \tag{30}$$

No caso de haver dois argumentos x e y (que podem ser multidimensionais), a entropia conjunta $H(x, y)$ é dada por:

$$H(x,y) = -\int \int p(x,y) \log p(x,y) dxdy. \tag{31}$$

A entropia de distribuições contínuas tem a maioria (mas não todas) das propriedades do caso discreto. Assim como no caso discreto, também se verificam no caso contínuo as seguintes propriedades:

1) Para quaisquer duas variáveis x e y tem se:

$$H(x,y) \leq H(x) + H(y). \tag{32}$$

A igualdade somente valerá se, e somente se, x e y forem independentes, ou seja, $p(x,y) = p(x)p(y)$ (com exceção, eventualmente, de um conjunto de pontos com probabilidade zero).

2) Verifica-se também a seguinte relação para entropia conjunta e condicional:

$$H(x,y) = H(x) + H(y|x) = H(y) + H(x|y). \tag{33}$$

Sendo $H(y|x) \leq H(y)$.

3) Há uma importante diferença entre a entropia para distribuições discretas e distribuições contínuas. No caso discreto, a entropia mensura de forma absoluta a aleatoriedade da variável aleatória. Já no caso contínuo, esta medida é afetada pelo sistema de coordenadas. Em geral, quando há modificações do sistema de coordenadas, a entropia sofre alterações.

O item 3 acima quer dizer, em outras palavras, que a unidade de medida afeta a entropia no caso contínuo, sendo que a entropia no caso discreto nada sofre. Dessa forma, como no caso contínuo, a entropia mede a aleatoriedade relativa a um sistema de coordenadas, logo, se for escolhida uma escala para uma distribuição uniforme na qual a entropia seja nula, a sua entropia será negativa, em qualquer distribuição que tenha um maior nível de concentração (LAZO; RATHIE, 1978).

No entanto, se for utilizada duas distribuições de probabilidade, estando seus valores numa dada dimensão, qualquer alteração em ambas as unidades de medida não irá causar quaisquer mudanças na diferença relativa de entropia entre elas (PHILIPPATOS; WILSON, 1972).

“A entropia das distribuições contínuas é sensível a quaisquer outras alterações que possam ocorrer na variável, por isso se adéqua mais para o estudo das taxas de rentabilidade dos títulos” (DIONÍSIO; MENEZES; MENDES, 2003, p. 13). Neste trabalho será utilizada a entropia contínua.

2.3.5 Informação mútua

A informação mútua é uma medida vinda da teoria da informação ou, mais especificamente, do conceito de entropia.

Um resultado importante obtido anteriormente mostrado na equação (28) que $H(y) \geq H(y|x)$, isto é, o conhecimento de x reduz a entropia ou incerteza de y, exceto quando x e y forem independentes. Essa diminuição da entropia significa que houve uma introdução de informação no sistema (DIONÍSIO; MENEZES; MENDES, 2003). Dessa forma, foi criada uma medida do ganho dessa informação, que recebeu o nome de informação mútua, que é dada por:

$$I(x, y) = H(y) - H(y|x). \tag{34}$$

De acordo com a equação (34), $I(x, y) \geq 0$ pois $H(y|x) \leq H(y)$. Se $I(x, y) = 0$ então $H(y|x) = H(y)$ o que significa que x e y são estatisticamente independentes. Se $I(x, y) = H(y)$, então $H(y|x) = 0$, o que implica y ser totalmente determinado por x.

De acordo com a equação (26), outra forma mais conveniente de calcular a informação mútua, dada pela equação (34), é:

$$I(x, y) = H(x) + H(y) - H(x, y). \tag{35}$$

Georgescu-Roegen (1971) afirma que o nível de entropia de um dado sistema mensura a ignorância da microestrutura desse mesmo sistema. Isso significa dizer que quando todos os eventos de um determinado grupo ou sistema tiverem iguais probabilidades de ocorrência, a entropia desse sistema será máxima, e, portanto o nível de informação disponível de tal sistema será mínimo.

Mayumi (1997) argumenta que a entropia e a informação são equivalentes. Assim, a informação assume um valor bastante elevado, pois conforme a necessidade de informação se torna maior, devido a grandes níveis de entropia, maior será também o seu valor. Logo, a informação pode reduzir a entropia, fazendo-a passar de um estado para outro (GEORGESCU-ROEGEN, 1971).

2.3.6 Coeficiente de correlação global

Granger e Lin (1994), Darbellay e Wuertz (2000) e Dionísio, Menezes e Mendes (2004) utilizaram a medida de correlação global, que é originada da medida padronizada de informação mútua.

A informação mútua pode ser considerada uma medida de dependência entre duas variáveis aleatórias, x e y, ou seja, uma medida da correlação estatística entre x e y (DIONÍSIO; MENEZES; MENDES, 2004). Back e Weigend (1998) utilizaram a medida de informação mútua (ou entropia mútua) na análise de dados de ações do mercado financeiro, com o objetivo de separar seus componentes independentes, ou seja, aqueles componentes não correlacionados.

A informação mútua $I(x, y)$ pode assumir qualquer valor real positivo, ou seja, $0 \leq I(x, y) \leq +\infty$, o que, por sua vez, dificulta a comparação entre amostras diversas. Sendo assim, Granger e Lin (1994), Darbellay e Wuertz (2000) usaram uma medida padronizada da informação mútua, denominada de coeficiente de correlação global definida por:

$$\lambda(x, y) = \sqrt{1 - e^{-2I(x,y)}}. \tag{36}$$

Esta medida varia entre 0 e 1, pois se $I(x, y) \rightarrow +\infty$ então $\lambda(x, y) \rightarrow 1$ e se $I(x, y) = 0$ então $\lambda(x, y) = 0$. Pode ser visto que quanto maior for a informação mútua, maior será o coeficiente de correlação global. A grande vantagem de λ é conseguir captar qualquer tipo de dependência, ou seja, tanto linear quanto não linear, entre x e y. Além disso, λ permite ser comparado ao coeficiente de correlação linear de Pearson r.

Com relação ao coeficiente linear, ele é interpretado tanto para valores positivos quanto para negativos, já o coeficiente de correlação global não assume valores negativos, portanto não é possível comparar o significado, mas sim a magnitude: quanto mais expressivo for o coeficiente global em relação ao linear, mais efeitos não lineares existem. Mas qual o tipo de não linearidade? Não se sabe a resposta, e o coeficiente de correlação global não informa isto. Não existe uma interpretação do tipo: quando uma variável aumenta a outra diminui. Não há linearidade; pode ser um seno, pode ser uma parábola, um círculo, qualquer relação não linear. O importante no tipo de análise que será realizada neste trabalho não é descrever a forma funcional, mas sim deixar claro que existem não linearidades expressas por meio do cálculo de λ.

2.4 Entropia e regressão linear

De acordo com Garner e McGill (1956) e Philippatos e Wilson (1972 e 1974), a entropia é formalmente equivalente à regressão e à análise de variância, correspondendo à variância da variável dependente na análise de regressão. Por exemplo, chamando de y os retornos de algum ativo, e de x os retornos de algum índice que representa o mercado, sendo y a variável dependente e x a variável independente, a entropia $H(y)$ nesse caso medirá a variância de y na regressão.

De forma análoga, a entropia condicional $H(y|x)$ mede a entropia inerente no ativo y dado o índice x, isto é, mede a variação não explicada ou residual de y, quando se considera que os efeitos de x sejam constantes. Similarmente, a medida de informação mútua $H(y) - H(y|x)$ mede a quantidade de informação que as variáveis x e y têm em comum, ou seja, mede a soma explicada dos quadrados da regressão, por mostrar a dependência que existe entre as variáveis x e y.

Por fim, a informação mútua de x e y dividido pela entropia de y é equivalente ao coeficiente de determinação R^2 na análise de regressão, medindo a proporção da variação de y que é explicada pela variação de x e é dada por:

$$\frac{I(x,y)}{H(y)} = \frac{H(y) - H(y|x)}{H(y)}. \tag{37}$$

A expressão $H(y|x)/H(y)$ mensura a proporção de y que não é explicada por x, o que é equivalente a $1 - R^2$. Se as variáveis x e y forem perfeitamente independentes, então a equação (37) será igual a zero, pois $H(y) = H(y|x)$ e a expressão $H(y|x)/H(y)$ será igual a 1, ou equivalentemente $R^2 = 0$, o que significa que x não explica nenhuma variação de y. Se, no entanto, x e y forem perfeitamente dependentes, então $H(y|x) = 0$, pois x determinará completamente y, assim $H(y|x)/H(y) = 0$; logo, (37) será igual 1, ou equivalentemente $R^2 = 1$.

Análise de regressão	Teoria da informação	Similaridades
Soma dos quadrados da regressão SQReg = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	Informação mútua $I(x,y) = H(y) - H(y\|x)$	Ambas as medidas medem a variação da variável dependente, que é explicada pela variável independente.
Soma dos quadrados do erro SQR = $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	Entropia condicional $H(y\|x) = H(x,y) - H(x)$	Ambas as medidas medem a variação da variável dependente, que não é explicada pela variável independente.
Soma total dos quadrados STQ = $\sum_{i=1}^{n}(y_i - \bar{y})^2$	Entropia total de y $H(y)$	Ambas as medidas medem a dispersão total da variável dependente.
Coeficiente de determinação $R^2 = \frac{SQR}{SQT}$	Informação mútua dividida pela entropia total de y $\frac{H(y) - H(y\|x)}{H(y)}$	Ambas as medidas medem a proporção da variação de y, que é explicada pela variação de x.

Quadro 1 – Equivalências entre a análise de regressão e medidas da teoria da informação
Fonte: Adaptado de Dionísio; Menezes; Mendes, 2003

As equivalências entre as medidas somente são significativas quando todas as premissas da regressão linear simples são satisfeitas, as quais são a linearidade, a homocedasticidade, a ausência de autocorrelação e a normalidade das distribuições de probabilidade dos resíduos. Só assim a reta de regressão irá captar a real relação existente entre as variáveis, ou seja, a relação será eficientemente representada pela reta de regressão linear. Caso não forem satisfeitos todos ou algum desses pressupostos, as medidas não terão uma correspondência perfeitamente proporcional. Isso se deve ao fato de a regressão linear se limitar a mensurar as variações na variável dependente que possam ser explicadas pela reta de regressão. Já as medidas da teoria da informação conseguem capturar a relação entre as variáveis como um todo, concentrando-se na maior ou menor independência delas (DIONÍSIO; MENEZES; MENDES, 2003). No quadro 1 são resumidas todas as relações acima discutidas.

3 METODOLOGIA

3.1 A amostra e suas propriedades

Blume (1971) examinou as propriedades estatísticas do coeficiente beta e Philippatos e Wilson (1974) aplicaram a metodologia da teoria da informação ao problema de avaliação e separação do risco do mercado de capital em seus componentes sistemático e não sistemático, sendo que ambos utilizaram uma janela de tempo de 84 meses. Dessa forma, como este trabalho também está examinando uma propriedade estatística (coeficiente de correlação global), foi escolhido o mesmo período (janela) de tempo de 84 meses.

A amostra aqui foi tomada da base de dados Economática® e foram selecionados todos os ativos que se mantiveram no IBOVESPA (carteira teórica) de maio de 2001 a abril de 2008, totalizando 84 meses (7 anos). Todas elas foram ajustadas por proventos (dividendos e mudanças de capital), utilizou-se o preço de fechamento mensal de cada ação e foi utilizado deflator em moeda original.

A razão de escolher as ações do IBOVESPA é a maior liquidez que em geral todas possuem. Se todas as ações da amostra compõem o IBOVESPA, todas terão influência sobre o comportamento do índice e, portanto, todas as acompanharão. Assim, a pergunta que se coloca sobre isso é: existe correlação não linear entre as variações do preço das ações e as variações do IBOVESPA?

3.2 Método

Esta seção descreve os passos até a obtenção do coeficiente de correlação global entre as variações de cada ativo da amostra (33 ativos no total) com as variações do IBOVESPA. Para ilustrar cada etapa, será usado como exemplo o ativo ARCZ6, que faz parte da amostra selecionada.

Inicialmente, com os preços do ativo e dos pontos do IBOVESPA no período de maio de 2001 a abril de 2008, calcula-se a variação mensal do mês t pela expressão:

$$R_t = \frac{R_t - R_{t-1}}{R_{t-1}}. \tag{38}$$

Dessa forma, foram obtidas 83 variações mensais para cada ativo e para o IBOVESPA. De forma a ilustrar, na Tabela 1 abaixo são mostradas 10 variações do total de 83 para os pontos do IBOVESPA e dos preços da ARCZ6, referentes aos meses de maio de 2001 a dezembro de 2002.

Tabela 1 – Retornos para os pontos do IBOVESPA e para os preços da ARCZ6

IBOVESPA			**ARCZ6**		
Data	**Fechamento**	**Retorno**	**Data**	**Fechamento**	**Retorno**
mai/01	14649		mai/01	2,9449	
jun/01	14559	**-0,0061**	jun/01	3,0167	**0,0244**
jul/01	13754	**-0,0553**	jul/01	2,8587	**-0,0524**
ago/01	12840	**-0,0665**	ago/01	3,1747	**0,1106**
set/01	10635	**-0,1717**	set/01	2,8587	**-0,0995**
out/01	11364	**0,0685**	out/01	3,4477	**0,2060**
nov/01	12931	**0,1379**	nov/01	3,1388	**-0,0896**
dez/01	13577	**0,0500**	dez/01	2,8084	**-0,1053**
jan/02	12721	**-0,0630**	jan/02	3,0239	**0,0767**
fev/02	14033	**0,1031**	fev/02	3,5195	**0,1639**
mar/02	13254	**-0,0555**	mar/02	3,4261	**-0,0265**

No CAPM, o valor esperado do retorno (variação média) de uma ação a ($E(R_a)$) é explicado pelo retorno esperado (variação média) do mercado ($E(R_m)$), ou seja, em termos matemáticos:

$$E(R_a) = \alpha + \beta\big(E(R_m)\big). \tag{39}$$

Sendo,

α: constante ou o intercepto de $E(R_a)$.

β: coeficiente angular na reta de regressão linear.

O coeficiente angular da reta de regressão linear, como dito anteriormente, é o chamado coeficiente β no modelo CAPM que representa o risco sistemático do ativo a.

Então como os valores de retorno do ativo (variável dependente) são explicados pelos retornos do IBOVESPA (variável independente), pode-se representar os dados (pontos (R_m, R_a)) num gráfico de dispersão conforme pode ser visto no Gráfico 2.

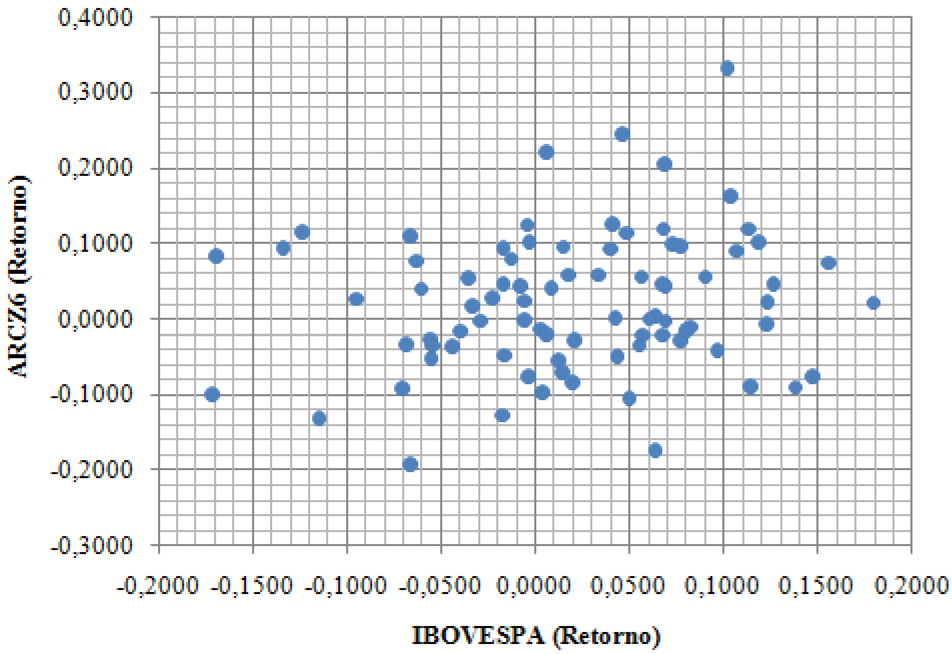

Gráfico 2 – Dispersão dos retornos dos pontos do IBOVESPA e dos preços da ARCZ6

Conforme foi visto, para o cálculo da entropia de uma distribuição contínua, é necessário conhecer a função de densidade de probabilidade. Sendo assim, houve a necessidade de encontrar um método que possibilitasse determinar essa função para os retornos do IBOVESPA $p(x)$, retornos do ativo $p(y)$ e a função de densidade de probabilidade conjunta, isto é, entre os retornos do IBOVESPA e do ativo, $p(x, y)$. O método utilizado aqui foi o da janela de Parzen (Parzen, 1962), método este não paramétrico.

O método da janela de Parzen é um método não paramétrico para a estimação da função de densidade de probabilidade de uma amostra de variáveis aleatórias (PARZEN, 1962). Dado uma amostra de uma população de variáveis aleatórias, a janela de Parzen faz uma estimativa da densidade de probabilidade para qualquer valor da variável utilizando uma soma de funções de base (kernel) centradas nos valores amostrais. Seja x_1, x_2,..., x_n uma amostra independente e identicamente distribuída de uma variável aleatória x, a aproximação de sua função densidade de probabilidade é dada por

$$p(x) = \frac{1}{n} \sum_{i=1}^{n} N(x, x_i, h_x). \tag{40}$$

Para uma variável y também se pode escrever

$$p(y) = \frac{1}{n} \sum_{i=1}^{n} N(y, y_i, h_y). \tag{41}$$

A distribuição conjunta pode ser escrita por

$$p(x, y) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} N(x, x_i, h_x) N(y, y_j, h_y). \tag{42}$$

Nas expressões acima, $N(x, x_i, h_x)$ são funções gaussianas com médias iguais aos valores de x_i e h_x, um valor fixo a ser determinado e assim o mesmo sendo para $N(y, y_i, h_y)$.

Como a área (integral) da função de densidade de probabilidade sempre deve ser igual a 1, integrando $p(x)$ e $p(y)$, o resultado deve ser 1. Assim também para o volume sob a curva $p(x, y)$, o resultado deve ser 1. Analiticamente estas propriedades são satisfeitas diretamente pela própria construção das funções. Numericamente, deve ser escolhida uma região para integração de forma que, para valores determinados de h_x e h_y, a quadratura (integral) resulte em 1. Devido a esses problemas, escalam-se valores para h até tornarem a integral da função igual a 1.

Para ilustrar o método acima, na sequência são apresentados os procedimentos para a estimação da função de densidade de probabilidade da amostra dos retornos dos preços do ativo ARCZ6 $p(y)$, dos retornos do IBOVESPA $p(x)$ e a conjunta deles $p(x, y)$.

Inicialmente, com os retornos obtidos dos preços do ativo ARCZ6 e do IBOVESPA, encontrou-se o máximo e mínimo retorno para cada um e o intervalo entre eles (diferença), obtendo-se os valores apresentados na Tabela 2.

Tabela 2 – Retornos máximo e mínimo para o IBOVESPA e para os preços da ARCZ6

	IBOVESPA	**ARCZ6**
Máximo retorno	0,1792	0,3333
Mínimo retorno	-0,1717	-0,1929
Intervalo	0,3509	0,5262

Com os valores da Tabela 2, pode-se fazer uma grade de valores para os retornos do ativo (ARCZ6) e para o retorno do IBOVESPA. Foram determinados 101 valores no intervalo limitado inferiormente pelo valor:

$$\text{Limitante inferior} = (\text{Mínimo retorno} - \text{Intervalo})/3. \tag{43}$$

E limitado superiormente por:

$$\text{Limitante superior} = (\text{Máximo retorno} + \text{Intervalo})/3. \quad (44)$$

No exemplo em questão, os valores obtidos estão apresentados na Tabela 3.

Tabela 3 – Limitante inferior e superior para o IBOVESPA e ARCZ6

	IBOVESPA	**ARCZ6**
Limitante inferior	-0,2900	-0,3700
Limitante superior	0,3000	0,5100
Variação	0,0059	0,0088

A variação na Tabela 3 é a diferença entre os limitantes superior e inferior dividido por 100. Para obter os 101 valores, pegou-se o limitante inferior como sendo o primeiro valor e outros 100 foram obtidos somando-se a variação até se chegar ao último valor (101), que é o limitante superior, isto é:

$$\text{Retorno } (i) = \text{Retorno } (i\text{-}1) + \text{Variação}. \quad (45)$$

Sendo que i começa no 2, pois o primeiro retorno (1) é o próprio limitante inferior. A título de ilustração, mostram-se, na Tabela 4, os 20 retornos dos 101 para o IBOVESPA x, e ARCZ6 y.

Tabela 4 – Retornos do IBOVESPA e ARCZ6 considerados na determinação das densidades de probabilidades *p*(*x*), *p*(*y*) e *p*(*x*,*y*)

i	x	y
1	-0,2900	-0,3700
2	-0,2841	-0,3612
3	-0,2782	-0,3524
4	-0,2723	-0,3436
5	-0,2664	-0,3348
6	-0,2605	-0,326
7	-0,2546	-0,3172
8	-0,2487	-0,3084
9	-0,2428	-0,2996
10	-0,2369	-0,2908
11	-0,231	-0,282
12	-0,2251	-0,2732
13	-0,2192	-0,2644
14	-0,2133	-0,2556
15	-0,2074	-0,2468
16	-0,2015	-0,238
17	-0,1956	-0,2292
18	-0,1897	-0,2204
19	-0,1838	-0,2116
20	-0,1779	-0,2028

O método da janela de Parzen foi aplicado para os 101 retornos, a fim de encontrar a densidade de probabilidade de cada um. A Tabela 5 mostra 20 densidades de probabilidades do IBOVESPA $p(x)$ e para o ativo ARCZ6 $p(y)$.

Tabela 5 – Densidade de probabilidade do IBOVESPA e dos retornos dos preços da ARCZ6

$\boldsymbol{j}$	$\boldsymbol{p(x)}$	$\boldsymbol{p(y)}$
1	1,1826E-04	7,3208E-06
2	2,5358E-04	1,9128E-05
3	5,2324E-04	4,7673E-05
4	1,0390E-03	1,1334E-04
5	1,9856E-03	2,5710E-04
6	3,6522E-03	5,5653E-04
7	6,4666E-03	1,1498E-03
8	1,1023E-02	2,2677E-03
9	1,8095E-02	4,2709E-03
10	2,8611E-02	7,6841E-03
11	4,3589E-02	1,3213E-02
12	6,4020E-02	2,1725E-02
13	9,0696E-02	3,4185E-02
14	1,2403E-01	5,1531E-02
15	1,6391E-01	7,4517E-02
16	2,0959E-01	1,0356E-01
17	2,5971E-01	1,3868E-01
18	3,1250E-01	1,7953E-01
19	3,6603E-01	2,2566E-01
20	4,1854E-01	2,7688E-01

Conhecendo os pontos $\big(x, p(x)\big)$ e $\big(y, p(y)\big)$ pode-se esboçar o gráfico da função de densidade de probabilidade dos retornos do IBOVESPA e do ativo. A área sob a curva da função de densidade de probabilidade deve ser igual a 1, e para isso o parâmetro h foi ajustado para o IBOVESPA $h_x = 0{,}0300$ e $h_y = 0{,}0400$ para a ARCZ6 (e para todos os outros ativos da amostra). A integral numérica utilizada para encontrar a área foi o método do trapézio. Os Gráficos 3 e 4 mostram o gráfico da função de densidade de probabilidade obtida para o IBOVESPA e ARCZ6, respectivamente.

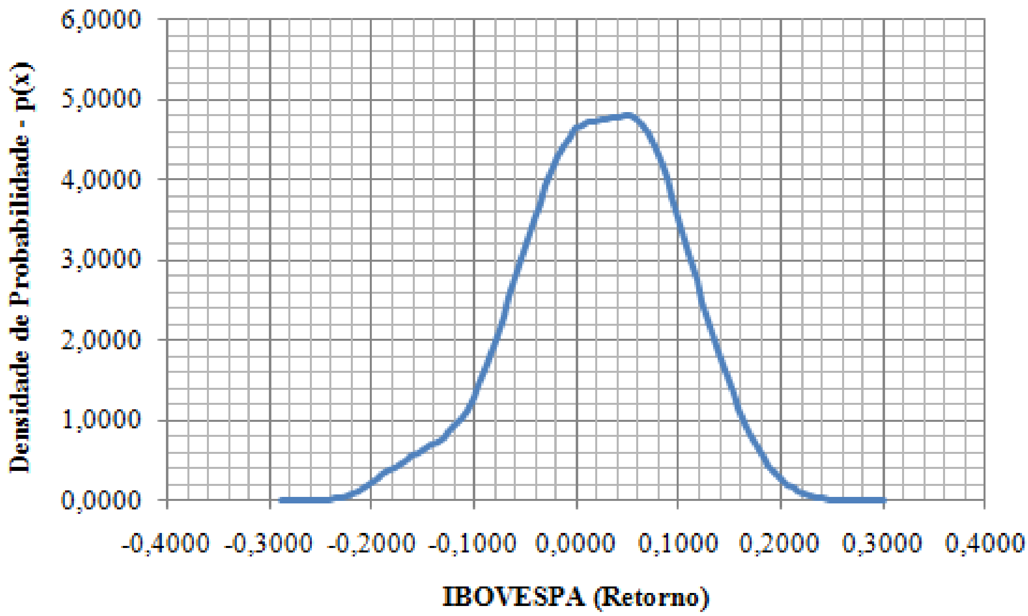

Gráfico 3 – Densidade de probabilidade dos retornos do IBOVESPA

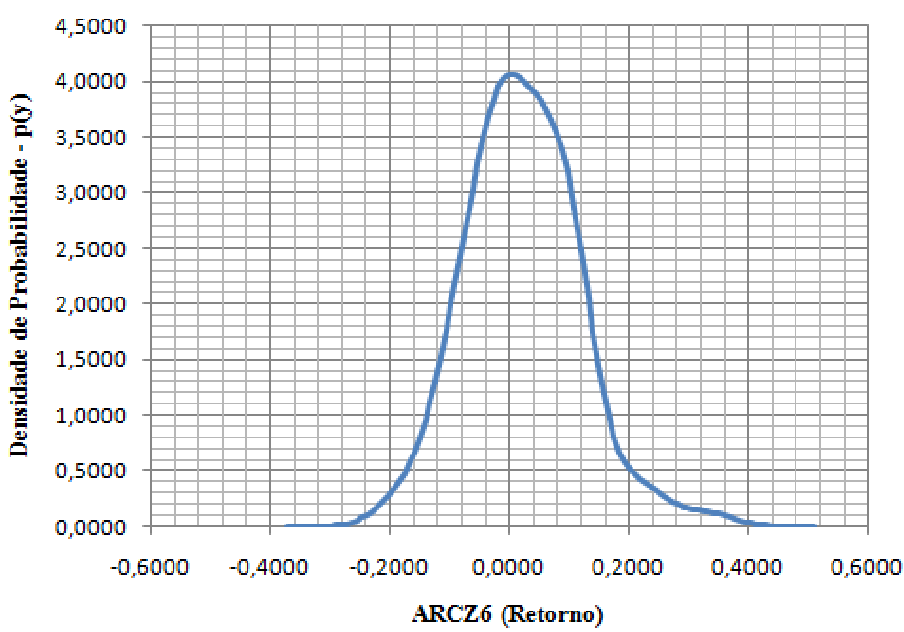

Gráfico 4 – Densidade de probabilidade dos retornos da ARCZ6

Para o caso da função de densidade de probabilidade conjunta $p(x, y)$ (IBOVESPA conjuntamente com outro ativo, aqui mostrado para apenas o ativo ARCZ6), a função utilizada foi a gaussiana conjunta para cada observação da amostra, conforme equação (42).

Da mesma forma, para o cálculo de $p(x)$ e $p(y)$ foi considerada a mesma grade dos valores de x e y para formarem pontos (x, y) dando um total de 10.201 pontos. Ilustrativamente, mostram-se os 20 pontos dos 10.201 pontos conjuntos para o IBOVESPA x e ARCZ6 y.

Tabela 6 – Retornos conjuntos do IBOVESPA e ARCZ6

j	x	y
1	-0,2900	-0,3700
2	-0,2900	-0,3612
3	-0,2900	-0,3524
4	-0,2900	-0,3436
5	-0,2900	-0,3348
6	-0,2900	-0,3260
7	-0,2900	-0,3172
8	-0,2900	-0,3084
9	-0,2900	-0,2996
10	-0,2900	-0,2908
11	-0,2900	-0,2820
12	-0,2900	-0,2732
13	-0,2900	-0,2644
14	-0,2900	-0,2556
15	-0,2900	-0,2468
16	-0,2900	-0,2380
17	-0,2900	-0,2292
18	-0,2900	-0,2204
19	-0,2900	-0,2116
20	-0,2900	-0,2028

Com os pontos (x, y), calculou-se a função de densidade de probabilidade $p(x, y)$ utilizando a equação (42). Ilustrativamente na Tabela 7 são apresentados alguns valores.

Tabela 7 – Densidade de probabilidade conjunta do IBOVESPA e ARCZ6

$\boldsymbol{j}$	$\boldsymbol{x}$	$\boldsymbol{y}$	$\boldsymbol{p(x,y)}$
1	-0,0717	-0,3700	**0,0001**
2	-0,0717	-0,3612	**0,0002**
3	-0,0717	-0,3524	**0,0006**
4	-0,0717	-0,3436	**0,0013**
5	-0,0717	-0,3348	**0,0029**
6	-0,0717	-0,3260	**0,0062**
7	-0,0717	-0,3172	**0,0126**
8	-0,0717	-0,3084	**0,0244**
9	-0,0717	-0,2996	**0,0449**
10	-0,0717	-0,2908	**0,0790**
11	-0,0717	-0,2820	**0,1322**
12	-0,0717	-0,2732	**0,2112**
13	-0,0717	-0,2644	**0,3216**
14	-0,0717	-0,2556	**0,4671**
15	-0,0717	-0,2468	**0,6478**
16	-0,0717	-0,2380	**0,8581**
17	-0,0717	-0,2292	**1,0873**
18	-0,0717	-0,2204	**1,3200**
19	-0,0717	-0,2116	**1,5395**
20	-0,0717	-0,2028	**1,7309**

Conhecendo os pontos $(x, y, p(x, y))$, pode-se esboçar o gráfico da função de densidade de probabilidade conjunta dos retornos conjunto do IBOVESPA e do ativo. Como o volume sob a curva da função deve ser igual a 1, calculou-se a integral numérica pelo método do tronco do prisma para encontrar o volume sob a curva de $p(x, y)$, o que confirmou uma área igual a 1. O Gráfico 5 é o gráfico da função de densidade de probabilidade conjunta obtida para o IBOVESPA e ARCZ6.

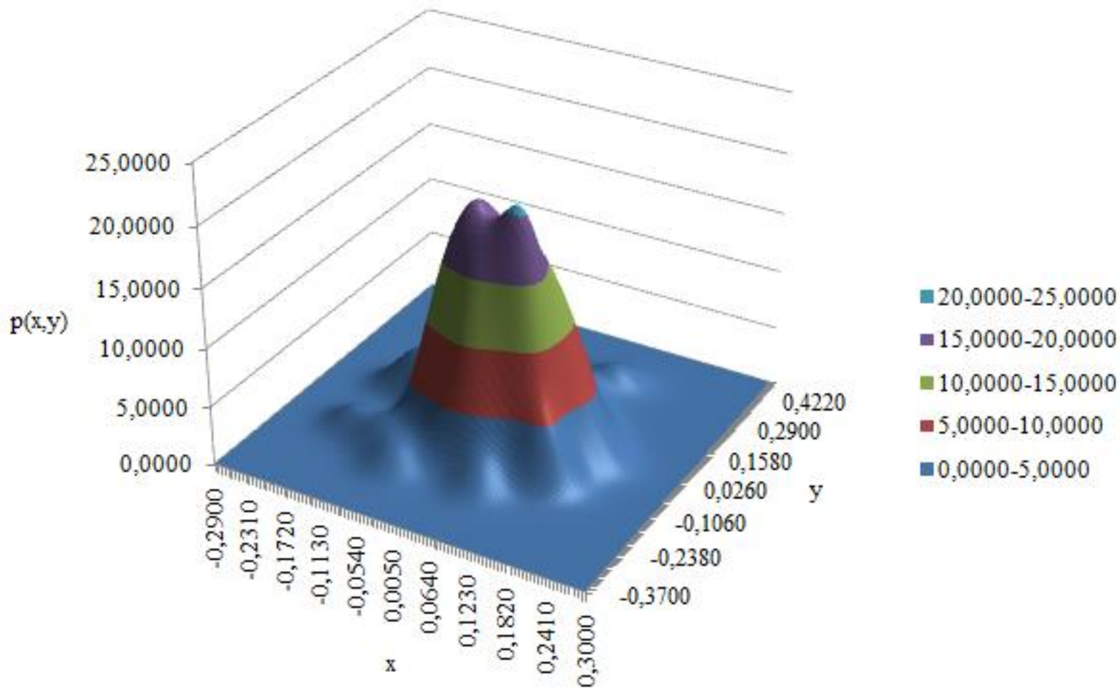

Gráfico 5 – Densidade de probabilidade conjunta dos retornos do IBOVESPA e ARCZ6

A partir dos valores das densidades de probabilidades $p(x)$, $p(y)$ e $p(x, y)$, foi calculada a entropia pontual para cada x, y e (x, y), como mostra a tabela 8, para 20 valores de $p(x)$, $p(y)$ e $p(x, y)$.

Tabela 8 – Densidade de probabilidade e entropia pontual dos retornos do IBOVESPA e ARCZ6

	IBOVESPA		ARCZ6		Conjunta	
j	$p(x)$	$-p(x)\,log_2\,p(x)$	$p(y)$	$-p(y)\,log_2\,p(y)$	$p(x,y)$	$-p(x,y)\,log_2\,p(x,y)$
1	0,0001	0,0015	0,0001	0,0015	0,0001	0,0012
2	0,0003	0,0030	0,0003	0,0031	0,0002	0,0027
3	0,0005	0,0057	0,0006	0,0060	0,0006	0,0060
4	0,0010	0,0103	0,0011	0,0112	0,0013	0,0125
5	0,0020	0,0178	0,0023	0,0199	0,0029	0,0245
6	0,0037	0,0296	0,0043	0,0336	0,0062	0,0455
7	0,0065	0,0470	0,0077	0,0540	0,0126	0,0795
8	0,0110	0,0717	0,0132	0,0825	0,0244	0,1306
9	0,0181	0,1047	0,0217	0,1200	0,0449	0,2011
10	0,0286	0,1467	0,0342	0,1665	0,0790	0,2892
11	0,0436	0,1970	0,0515	0,2205	0,1322	0,3860
12	0,0640	0,2539	0,0745	0,2792	0,2112	0,4738
13	0,0907	0,3141	0,1036	0,3388	0,3216	0,5263
14	0,1240	0,3735	0,1387	0,3953	0,4671	0,5130
15	0,1639	0,4276	0,1795	0,4448	0,6478	0,4058
16	0,2096	0,4725	0,2257	0,4847	0,8581	0,1895
17	0,2597	0,5051	0,2769	0,5130	1,0873	-0,1313
18	0,3125	0,5244	0,3337	0,5284	1,3200	-0,5288
19	0,3660	0,5307	0,3978	0,5290	1,5395	-0,9583
20	0,4185	0,5259	0,4719	0,5113	1,7309	-1,3701

Com as entropias pontuais calculadas, pode-se calcular a entropia total dada pela equação (29), isto é, $H(x) = -\int_{-\infty}^{\infty} p(x)\,log\,p(x)\,dx$. Esta integral foi calculada numericamente pelo método do trapézio. Os valores obtidos para o IBOVESPA foram $H(x) =$ -1,6236, e para o ativo ARCZ6 $H(y) =$ -1,3056. Para o cálculo da entropia conjunta dada pela equação (31), ou seja, $H(x,y) = -\int\int p(x,y)\,log\,p(x,y)dxdy$, utilizou-se o método do tronco do prisma para o cálculo da integral numérica, obtendo-se $H(x,y) =$ -2,9913. A informação mútua dada pela equação (35), isto é, $I(x,y) = H(x) + H(y) - H(x,y)$ pode agora ser calculada, resultando $I(x,y) =$ 0,0620. Agora é possível encontrar o coeficiente de correlação global entre os retornos do IBOVESPA e do ativo, como está sendo ilustrado para o ativo ARCZ6, tem-se que o coeficiente global dado pela equação (36), ou seja, $\lambda(x,y) = \sqrt{1 - e^{-2I(x,y)}}$ é $\lambda(x,y) =$ 0,3416, sendo o coeficiente linear dado pela equação (2), isto é, $r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$, e $r =$ 0,1500.

4 RESULTADOS

Na Tabela 9 abaixo, são apresentados a informação mútua $I(x, y)$, coeficiente de correlação global $\lambda(x, y)$ e o coeficiente de correlação linear $r(x, y)$, sendo x os retornos do IBOVESPA e y os retornos do ativo. Os dados estão dispostos em ordem alfabética dos códigos dos 33 ativos.

Tabela 9 – Resultados: Informação mútua $I(x,y)$, coeficiente de correlação global $\lambda(x,y)$, coeficiente de correlação linear $r(x,y)$ e a diferença entre eles $\lambda(x,y) - r(x,y)$

j	Ativos	$I(x,y)$	Global: $\lambda(x,y)$	Linear: $r(x,y)$	Diferença: $\lambda(x,y) - r(x,y)$
1	ACES4	0,2817	0,6563	0,5316	0,1247
2	AMBV4	0,2384	0,6159	0,5512	0,0647
3	ARCZ6	0,0620	0,3416	0,1500	0,1916
4	BBAS3	0,4000	0,7421	0,7166	0,0255
5	BBDC4	0,4287	0,7588	0,7557	0,0030
6	BRAP4	0,3681	0,7218	0,6679	0,0539
7	BRTP3	0,2524	0,6296	0,5186	0,1110
8	BRTP4	0,2135	0,5895	0,5293	0,0602
9	CLSC6	0,2869	0,6607	0,6090	0,0517
10	CMIG4	0,3674	0,7214	0,7025	0,0189
11	CPLE6	0,3414	0,7034	0,6754	0,0280
12	CRUZ3	0,1451	0,5019	0,4350	0,0669
13	CSNA3	0,3764	0,7273	0,6915	0,0358
14	ELET3	0,3391	0,7018	0,6253	0,0764
15	ELET6	0,3560	0,7137	0,6567	0,0570
16	EMBR3	0,2315	0,6088	0,4971	0,1117
17	GGBR4	0,4019	0,7432	0,7129	0,0303
18	ITSA4	0,4494	0,7700	0,7876	-0,0176
19	KLBN4	0,2683	0,6444	0,4609	0,1835
20	LIGT3	0,3633	0,7186	0,5399	0,1788
21	PETR3	0,3787	0,7288	0,7229	0,0059
22	PETR4	0,3886	0,7351	0,7324	0,0026
23	SBSP3	0,3218	0,6889	0,6620	0,0269
24	TCSL3	0,2621	0,6387	0,5153	0,1235
25	TCSL4	0,2989	0,6708	0,5805	0,0903
26	TLPP4	0,1885	0,5604	0,5073	0,0531
27	TMCP4	0,1919	0,5645	0,4707	0,0939
28	TNLP3	0,2522	0,6294	0,4603	0,1691
29	TNLP4	0,3145	0,6833	0,5313	0,1520
30	TRPL4	0,2766	0,6519	0,6039	0,0479
31	USIM5	0,4738	0,7825	0,7556	0,0269
32	VALE5	0,1701	0,5370	0,3670	0,1700
33	VCPA4	0,0746	0,3723	0,1845	0,1878

De modo geral, o coeficiente de correlação global obtido foi maior do que o linear em todos os casos, exceto para o ativo ITSA4, sugerindo a presença de correlações não lineares não captadas pelo coeficiente linear. O ativo ITSA4, mesmo tendo seu coeficiente de correlação linear maior do que o global, apresenta-os muito próximos um do outro. A razão disso é que os valores dos coeficientes globais e lineares calculados neste trabalho são apenas estimativas pontuais sem um nível de confiança, o que pode significar que, nestes casos, estatisticamente, estes valores sejam iguais. Fica aqui a sugestão de incluir um intervalo de confiança para estes coeficientes para trabalho futuro.

O menor coeficiente global foi 0,3416, referente ao ativo ARCZ6, e o maior foi 0,7825, referente ao ativo USIM5. Para o coeficiente linear, o menor foi 0,1500 e o maior 0,7876. Comparativamente o menor coeficiente linear é praticamente a metade do menor coeficiente global, apresentando uma diferença discrepante, fato que não ocorre para o maior coeficiente linear em relação ao coeficiente global, ambos apresentando valores extremamente próximos. Em relação às diferenças entre os coeficientes lineares e globais, a menor diferença foi de 0,0026 (desconsiderando o caso negativo de -0,0176) para o ativo PETR4, sugerindo que este ativo acompanha o IBOVESPA de forma predominantemente linear. A maior diferença foi de 0,1916 para o ativo ARCZ6, o que sugere um acompanhamento de forma predominantemente não linear em relação ao IBOVESPA. Interessante visualizar o gráfico de dispersão desses dois casos para ver como os retornos estão dispersos. Do ativo ARCZ6, os valores já estão dispostos no Gráfico 2, onde é possível ver pontos bastante dispersos um dos outros. Para o ativo PETR4, é possível ver no Gráfico 6 os pontos mais próximos um dos outros e mais linearmente dispostos, fato que não ocorre com a ARCZ6.

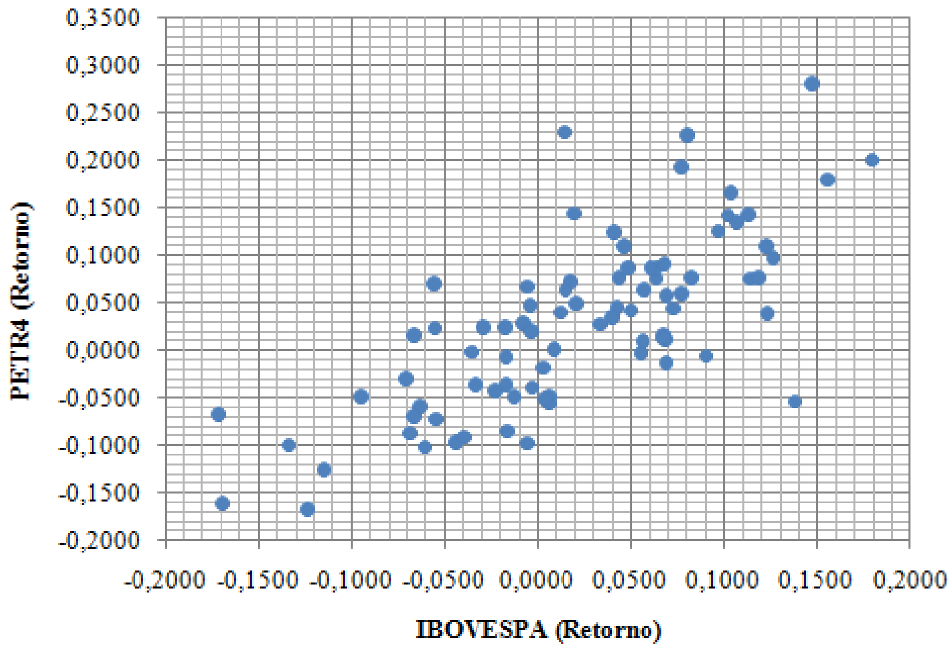

Gráfico 6 – Dispersão dos retornos do IBOVESPA e dos preços da PETR4

O Gráfico 7 apresenta uma comparação.

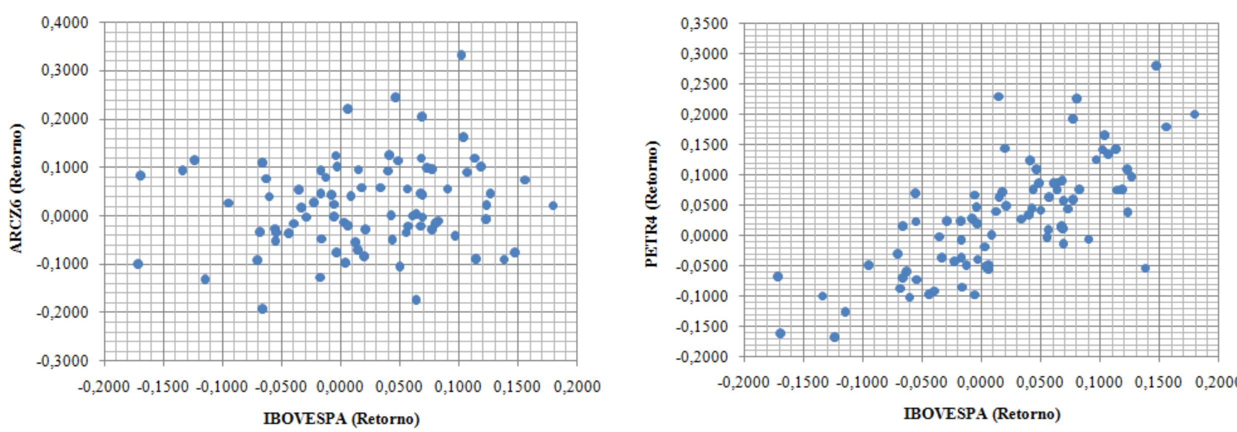

Gráfico 7 – Dispersão dos retornos do IBOVESPA e dos preços da ARCZ6 e PETR4

Com base nos resultados apresentados, pôde se verificar a existência de correlações não lineares significativas entre os retornos de preços de ações e retorno do IBOVESPA. No entanto, de modo a testar quantitativamente a hipótese desta dissertação, realizou-se o teste *t* de diferenças de médias para amostras emparelhadas. Retomando a hipótese do trabalho:

Hipótese: A correlação não linear entre variações de preço de ações e variações do IBOVESPA é, em média, maior do que a correlação linear.

Considerando a média de λ (correlações não lineares) como μ_1 e a média de r (correlações lineares) como μ_2, a hipótese desta pesquisa pode ser escrita, em símbolos matemáticos como:

Afirmativa original: $\mu_1 > \mu_2$

O oposto desta afirmativa original pode então ser escrito como:

Oposto da afirmativa original: $\mu_1 \leq \mu_2$

De forma a realizar o teste estatístico, definimos abaixo as hipóteses estatísticas H_0 e H_1. De acordo com Triola (2005), a hipótese H_0 é a hipótese que contém o sinal de igualdade e a hipótese alternativa é a que não contém o sinal de igualdade; assim, para o teste t de amostras aos pares, temos:

H_0: $\mu_1 - \mu_2 = 0$

H_1: $\mu_1 - \mu_2 > 0$

Considerando a amostra de 33 empresas, o resultado do teste t obtido pelo Excel é apresentado na Tabela 10 abaixo.

Tabela 10 – Teste: Duas amostras em par para médias

	Lambda - λ	**r**
Média	0,6520	0,5730
Variância	0,0106	0,0226
H_0: $\mu_1 - \mu_2 = 0$		
H_1: $\mu_1 - \mu_2 > 0$		
Grau de liberdade	32	
t de teste	7,3229	
Valor-P (uni-caudal)	0,000000013	
t crítico (uni-caudal)	1,6939	

Com base no resultado do teste t apresentado na Tabela 10, rejeita-se a hipótese H_0. Desta forma, considerando a hipótese de pesquisa, como a afirmativa original não contém a igualdade ($\mu_1 > \mu_2$), conclui-se que:

Os dados amostrais apoiam a afirmativa de que “a correlação não linear entre variações de preço de ações e variações do IBOVESPA é, em média, maior do que a correlação linear”.

5 CONCLUSÃO

Os resultados apresentados na seção anterior resumem, de forma quantitativa, os objetivos deste trabalho: mensurar os coeficientes de correlações lineares e não lineares entre as variações de preço de ações e o IBOVESPA, e comparar os dois tipos de coeficientes para analisar a discrepância entre eles.

Com os resultados obtidos (Tabela 9), fica possível responder o problema de pesquisa deste trabalho: As correlações não lineares são relevantes no mercado brasileiro? A resposta é sim, pois a correlações globais são maiores do que as correlações lineares, de acordo com um teste de hipóteses para diferenças de médias para amostras pareadas. Isso demonstra a relevância das correlações não lineares no mercado brasileiro.

As variações do preço dos ativos da amostra (com apenas uma exceção) apresentaram o seu coeficiente de correlação global com as variações do IBOVESPA maiores do que o coeficiente de correlação linear, o que sugere a presença de correlações não lineares. Logo, fica evidente a sustentação da hipótese de pesquisa na qual se baseia este trabalho: A correlação não linear entre variações de preço de ações e variações do IBOVESPA é, em média, maior do que a correlação linear.

A contribuição deste trabalho foi ter trazido mais uma ferramenta na área de finanças para auxiliar o entendimento do risco dado pelo comportamento de ativos no mercado de capitais brasileiro, captado pelo coeficiente de correlação global $\lambda(x,y)$. Sobre isso, Damodaran (2007) argumenta que um ativo que oscila independentemente do mercado (carteira de mercado) terá unicamente risco diversificável (não sistemático), ou seja, risco específico da empresa que pode ser diversificado. Assim, no caso de voltar às atenções apenas para correlações lineares, pode-se achar (no caso de correlações fracas) que o ativo oscila independentemente do mercado e, como argumenta Damodaran, ter apenas risco diversificável. Na verdade, como a correlação não linear podem ser maiores do que a linear observa-se que as oscilações não são independentes, ou seja, pode ocorrer um risco sistemático não linear nos ativos.

REFERÊNCIAS BIBLIOGRÁFICAS

ASSAF NETO, A. **Finanças corporativas e valor**. 2. ed. São Paulo: Altas, 2005a.

______. **Mercado financeiro**. 6. ed. São Paulo: Atlas, 2005b.

BACK, A. D.; WEIGEND, A. S. What drives stock returns? An independent component analysis. In: CONFERENCE ON COMPUTATIONAL INTELLIGENCE FOR FINANCIAL ENGINEERING (CIFE), 1998, New York. **PROCEEDINGS…** New York: IEEE, 1998, p. 12-156.

BLUME, M. E. On the assessment of risk. **Journal of Finance**, Hoboken, v. 26, n. 1, p. 1-10, Mar. 1971.

BREALEY, R. A.; MYERS, S. C.; ALLEN, F. **Princípios de finanças corporativas**. 8. ed. São Paulo: McGraw-Hill, 2008.

BUCKLEY, J. J. Entropy principles in decision making under risk. **Risk Analysis**, Hoboken, v. 5, n. 10, p. 303-313, 1985.

CASSETARI, A. O princípio da máxima entropia e a moderna teoria das carteiras. **Revista Brasileira de Finanças**, Rio de Janeiro, v. 1, n. 2, p. 271-300, dez. 2003.

DAMODARAN, A. **Avaliação de empresas**. 2. ed. São Paulo: Pearson Prentice Hall, 2007.

______. **Finanças corporativas aplicadas:** manual do usuário. Porto Alegre: Bookman, 2002.

DARBELLAY, G. A.; WUERTZ, D. The entropy as a tool for analysing statistical dependences in financial time series. **Physica A: Statistical Mechanics and its Aplications**, Amsterdam, v. 287, n. 3/10, p. 429–439, Dec. 2000.

DIONÍSIO, A.; MENEZES, R.; MENDES, D. A entropia como medida de informação na modelação económica. In: REIS, E.; HILL, M. M. **Temas em métodos quantitativos.** 3. ed. Lisboa: Silabo, 2003.

______. Mutual information: a measure of dependency for nonlinear time series. **Physica A: Statistical Mechanics and its Aplications**, Amsterdam, v. 35, n. 1/2, p. 326-329, Dec. 2004.

FAMA, E. F. Tomorrow on the New York Stock Exchange. **The Journal of Business**, Chicago, v. 38, n. 3, p. 285-299, July 1965.

FAMA, E. F.; FRENCH, K. R. The cross-section of expected stock returns. **Journal of Finance**, Hoboken, v. 8, n. 2, p. 37-75, June 1992.

FAMÁ, R.; BARROS, L. A. B.; SILVEIRA, A. D. M. A estrutura de capital é relevante? Novas evidências a partir de dados norte-americanos e latino-americanos. **Caderno de Pesquisas em Administração**, São Paulo, v. 8, n. 2, p. 71-84, jun. 2001.

FERNHOLZ, R. On the diversity of equity markets, **Journal of Mathematical Economics**, Amsterdam, v. 31, n. 3, p. 393-427, 1999.

FIELDMAN, D. **A brief introduction to: information theory, excess entropy and computational mechanics**. California: University of California, 1998.

GARNER, W. R.; MCGILL, W. J. The relation between information and variance analysis. **Psychometrika**, New York, v. 21, p. 219-228, 1956.

GEORGESCU-ROEGEN, N. **Entropy law and the economic process**. Harvard University Press: Cambridge, 1971.

GOULART, A. M. C. **Evidenciação contábil do risco de Mercado por instituições fnanceiras no Brasil**. 2003. 202 f. Dissertação (Mestrado em Controladoria e Contabilidade) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2003.

GRANGER, C. W. J.; LIN, J. L. Using the mutual information coefficient to identify lags in nonlinear models. **Journal of Time Series Analysis**, Oxford, v. 15, n. 10, p. 371-381, 1994.

HERSCOVICI, A. Historicidade, entropia e não linearidade: algumas aplicações possíveis na Ciência Econômica. **Revista de Economia Política**, São Paulo, v. 25, n. 3, p. 277-294, July 2005.

HOROWITZ, A.; HOROWITZ, I. Entropy, Markov processes and competition in the brewing industry. **The Journal of Industrial Economics**, Oxford, v. 16, n. 3, p. 196-211, July 1968.

JORION, P. **Value-at-risk**: a nova fonte de referência para a gestão do risco financeiro. 2. ed. São Paulo: Bolsa de Mercadorias e Futuros, 2003.

LEVINE, D. M.; STEPHAN, D. F.; KREHBIEL, T. C.; BERENSON, M. L. **Estatística**: teoria e aplicações. 5 ed. Rio de Janeiro: LTC, 2008.

LINTNER, J. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. **The Review of Economics and Statistics**, Cambridge, v. 8, p. 13-37, Feb. 1965.

MARKOWITZ, H. M. Portfolio selection. **Journal of Finance**, Hoboken, v. 7, n. 1, p. 77-91, Mar. 1952.

MAYUMI, K. Information, pseudo measures and entropy: an elaboration on Nicholas Georgescu-Roegen's critique. **Ecological Economics**, Amsterdam, v. 22, n. 3, p. 249-259, 1997.

MENEZES, E. A. Breve história do pensamento teórico em finanças. **Fae business**, Curitiba, n. 10, p. 48-50, dez. 2002.

KNIGHT, F. H. **Risk, uncertainty and profit**. Boston: Houghton Mifflin, 1921. Disponível em: <http://www.econlib.org>. Acesso em: 22 jul. 2010.

MODIGLIANI, F.; MILLER, M. H. The cost of capital, corporation finance and the theory of investment. **The American Economic Review**, Nashville, v. 9, n. 3, p. 261-297, 1958.

MOSSIN, J. Equilibrium in capital asset market. **Econometrica**, Oxford, v. 310, n. 10, p. 768-783, Oct. 1966.

PARZEN, E. On the estimation of probability density function and the mode. **The Annals of Mathematical Statistics**, Ann Arbor, v. 33, n. 3, p. 1065-1076, 1962.

PHILIPPATOS, G. C.; NAWROCKY, D. N. The information inaccuracy of stock market forecasts: some new evidence of dependence on the N.Y.S.E. **The Journal of Financial and Quantitative Analysis**, Stanford, v. 8, n. 3, p. 55-68, June 1973.

PHILIPPATOS, G. C.; WILSON, C. J. Entropy, market risk and the selection of efficient portfolios. **Applied Economics**, Abingdon, v. 4, n. 3, p. 209-220, Sept. 1972.

______. Information theory and risk in capital markets. **Omega- International Journal of Management Science**, Kidlington, v. 2, n. 10, p. 523-532, 1974.

ROSS, S. A.; WESTERFIELD, R. W.; JAFFE, J. F. **Administração financeira**. São Paulo: Atlas, 1995.

SASSATANI, R. **Uma análise empírica do prêmio pelo risco nos contratos futuros de índice Bovespa da BM&F**. 1999. 166 f. Dissertação (Mestrado em Administração) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 1999.

SHANNON, C. E. A mathematical theory of communication. **The Bell System Technical Journal**, New York, v. 27, p. 379-423, July 1948.

SHANNON, C. E.; WEAVER, W. **The mathematical theory of communication**. Urbana: The University of Illinois Press, 1964.

SHARPE, W. F. Capital asset prices: a theory of market equilibrium under conditions of risk. **Journal of Finance**, Hoboken, v. 19, n. 3, p. 35-10, Sept. 1964.

SIQUEIRA, J. O. **Determinação entrópica do preço racional da opção europeia simples ordinária sobre ação e bond:** uma aplicação da teoria da informação em finanças em condição de incerteza. 1999. 309 f. Tese (Doutorado em Administração) – Departamento de Administração, Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 1999.

______. Introdução à gestão do risco. **Revista FACEF Pesquisa,** Franca, v. 6, n. 3, p. 22-30, 2003.

TRIOLA, M. F. **Introdução à estatística**. 9. ed. Rio de Janeiro: LTC, 2005.

LAZO, A. V.; RATHIE, P. On the Entropy of Continuos Probability Distributions. **IEEE Transactions on Information Theory**, Piscataway, v. 34, n. 1, p. 120-122, 1978.