

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE FÍSICA  
GRADUAÇÃO EM BACHARELADO EM FÍSICA

Análise do comportamento do Mercado de Capitais  
através da Teoria da Informação

*Trabalho de Conclusão de Curso apresentado ao curso de  
Bacharelado em Física - Pesquisa Básica, do Instituto de Física  
da Universidade Federal do Rio Grande do Sul, como um  
pré-requisito para a obtenção do grau de Bacharel em Física.*

Cássio Murilo Ávila

Orientador: Prof. Dr. David Dominguez

2013

# Sumário

<b>1</b>	<b>Introdução</b>	<b>4</b>
<b>2</b>	<b>Mercado de Capitais</b>	<b>6</b>
2.1	Mercado Primário e Mercado Secundário . . . . .	6
2.2	Bolsa de Valores . . . . .	6
2.3	S&P 500 e o comportamento do mercado . . . . .	8
<b>3</b>	<b>Probabilidade</b>	<b>10</b>
3.1	Probabilidade Marginal e Definições . . . . .	10
3.2	Probabilidade Multivariacional . . . . .	11
3.3	Teorema de Bayes . . . . .	12
3.4	Covariância e Correlação . . . . .	13
<b>4</b>	<b>Teoria da Informação</b>	<b>14</b>
4.1	Conteúdo Informacional de Shannon . . . . .	14
4.2	Entropia Informacional de Shannon . . . . .	15
4.3	Entropia Multivariacional . . . . .	16
4.4	Informação Mútua . . . . .	18
<b>5</b>	<b>Cálculos e Resultados</b>	<b>19</b>
5.1	Método Numérico . . . . .	19
5.2	Correlação x Informação Mútua . . . . .	20
5.3	Informação Mútua Média . . . . .	26
<b>6</b>	<b>Conclusão</b>	<b>29</b>

## Resumo

Sabendo que características marcantes no comportamento do mercado de capitais como *crashes* e *bubbles* podem ser identificadas pela entropia informacional, neste trabalho pretendemos estudar o mercado utilizando algumas variáveis da teoria da informação. Primeiro fazemos uma comparação entre a correlação e a informação mútua de uma série temporal da bolsa de valores de Nova Iorque a fim de confirmar que a informação mútua representa melhor o comportamento do mercado. Por fim nos focamos exclusivamente na informação mútua a fim de entender melhor seu comportamento em relação ao comportamento do mercado.

## **Abstract**

Knowing that remarkable features in the capital market's behavior such as *crashes* and *bubbles* can be identified by information entropy, in this work we intend to study the market using some information theory variables. First, we make a comparison between the correlation and the mutual information of a time series of the New York stock exchange to confirm that mutual information best represents the market's behavior. Last, we focus exclusively on the mutual information to better understand it's behavior in relation to the market's one.

# 1 Introdução

O mercado financeiro pode ser visto como um sistema complexo onde a interação entre agentes leva a uma auto-organização que define a direção do mercado. Aqui interpretamos *interação* num sentido mais amplo, de mútua ou recíproca influência, e os agentes podem ser tanto as ações como os investidores. Sistemas complexos são de grande interesse por exibirem um comportamento coletivo emergente causado, a priori, por regras de interação simples entre os muitos elementos que constituem o sistema. Sabe-se que o mercado exibe características como sincronização [2], reorganização estrutural [3, 4], leis de potência [5, 6], hierarquia e não-aleatoriedade [7]. Foi demonstrado também que a estrutura do mercado pode ser explicada através da interação entre pares [8].

Dadas essas considerações, faz sentido pensar em uma modelagem estatística como abordagem natural para estudar o mercado. De fato, este é o objetivo do presente trabalho. Aqui fizemos uso de variáveis da teoria da informação para investigar o comportamento do mercado. Em [8] mostrou-se que a entropia informacional pode servir para identificar irregularidades no mercado financeiro, como *crashes* ou *bubbles* da bolsa. Neste trabalho fizemos uma análise de uma série temporal dos valores de alta de ações das empresas constituintes do índice S&P500 da bolsa de valores de Nova Iorque através da informação mútua, uma variável da teoria da informação definida a partir da entropia informacional de duas variáveis aleatórias.

Primeiramente fizemos uma comparação da informação mútua com a correlação entre os dias da série temporal analisada para verificar que a informação mútua nos dá uma melhor representação do comportamento do mercado. Posteriormente fizemos um estudo mais aprofundado da informação mútua em diferentes intervalos da série temporal a fim de identificar parâmetros que caracterizem a transmissão da informação e como ela está relacionada com o comportamento do mercado.

O trabalho está dividido da seguinte maneira: primeiro é feita uma breve introdução ao mercado de capitais, explicando seu funcionamento até che-

gamos à bolsa de valores e ao índice estudado. Em segundo lugar, fizemos uma breve introdução a algumas variáveis da teoria de probabilidades que são úteis para o maior entendimento do trabalho. Depois fizemos uma apresentação mais detalhada da teoria da informação e suas variáveis de maior importância. Por fim, explicamos o método numérico utilizado, apresentamos alguns resultados e nossa conclusão.

## 2 Mercado de Capitais

O mercado de capitais é um sistema de distribuição de valores mobiliários que tem o objetivo de proporcionar liquidez aos títulos de emissão de empresas e viabilizar seu processo de capitalização. É constituído pelas bolsas, corretoras e outras instituições financeiras autorizadas. Nele os principais títulos negociados são as ações (representativos do capital de empresas), debêntures (empréstimos tomados via mercado por empresas) e derivativos (contratos no qual se estabelecem pagamentos futuros), o que permite a circulação de capital para custear o desenvolvimento econômico [1].

O mercado de capitais é subdividido em Mercado Primário e Mercado Secundário.

### 2.1 Mercado Primário e Mercado Secundário

No mercado primário é onde ocorre o lançamento de novas ações ao público, isto é, onde os valores mobiliários circulam pela primeira vez e onde a empresa obtém o capital para seus empreendimentos, pois o dinheiro da venda vai para a empresa. Já no mercado secundário são feitas as demais negociações dos títulos colocados em circulação pelo mercado primário, como compra, venda ou simples troca dos títulos entre os possuidores, e a empresa já não tem mais contato com o dinheiro proveniente dessas negociações.

As negociações no mercado secundário se dão nas bolsas de valores e no mercado de balcão (onde ocorrem as negociações de ações de empresas não registradas na bolsa de valores).

Nosso foco no presente trabalho é na bolsa de valores.

### 2.2 Bolsa de Valores

As bolsas de valores são locais que oferecem as condições e os sistemas necessários para a realização de negociações de compra e venda de títulos,

valores mobiliários e outros ativos de forma transparente. Além de oferecer um ambiente para a negociação dos títulos nelas registrados, orientar e fiscalizar os serviços prestados por seus intermediários, facilitar a divulgação constante de informações sobre as empresas e sobre os negócios que se realizam sob seu controle, as bolsas propiciam liquides às aplicações de médio e longo prazos por intermédio de um mercado contínuo representado por seus pregões diários [1].

Os movimentos de uma bolsa de valores são captados através de índices (ou indicadores) chamados *índice de bolsa de valores*. Cada bolsa de valores tem seu(s) índice(s).

Alguns exemplos das principais bolsas de valores e seus respectivos índices:

- Bolsa de Valores de Nova Iorque (New York Stock Exchange - NYSE)
  - Dow Jones
  - S&P 500
- Bolsa de Valores de Tóquio
  - Nikkei 225
- Bolsa de Valores de Londres
  - FTSE 100
- Bolsa de Valores de Frankfurt
  - DAX-30
- Bolsa de Valores de São Paulo (BM&FBovespa)
  - Ibovespa

Neste trabalho nos focamos no índice S&P 500 da bolsa de valores de Nova Iorque.



## 2.3 S&P 500 e o comportamento do mercado

O índice S&P 500, ou Standard & Poor's 500, é baseado na capitalização do mercado das 500 empresas líderes do mercado americano como definido pela Standard & Poor. As 500 empresas são escolhidas por um comitê baseadas em critérios bem definidos, como capitalização de mercado, liquidez, duração do tempo de capital aberto e classificação setorial, entre outros, somando oito critérios no total. Sendo assim, as 500 empresas não são constantes e ocorrem variações de ano para ano sem data específica para a alteração, podendo ocorrer em qualquer mês, embora, de fato, poucas empresas sejam alteradas de um ano para o outro.

É um dos índices comumente mais seguidos e muitos o consideram a melhor representação do mercado americano.

Neste trabalho usamos dados de 349 empresas referentes a 495 dias de mercado, abrangendo o período de 25 de agosto de 2008 a 20 de agosto de 2010. Para fazer os cálculos precisamos que todas as empresas tenham participado dos mesmos dias de mercado, portanto, como dito acima, havendo substituições de algumas empresas de um ano para o outro, isso nos impede de usar todas as 500 presentes do índice. Ainda assim, mesmo algumas empresas que foram constantes no período que analisamos não tinham dados disponíveis em todos os 495 dias, portanto tivemos que fazer uma filtragem afim de manter apenas aquelas com dados disponíveis em toda a série temporal utilizada, restando assim as 349.

A base de dados utilizada [9] continha os valores das variáveis Open, High, Low, Close e Volume, referentes respectivamente aos valores de abertura, alta, baixa, fechamento e volume de ações negociadas da empresa. Escolhemos por nos focar no valor de alta das ações.

Podemos observar na Figura 1 que já no começo da série temporal, por volta de  $t = 6$ , representando o começo de setembro de 2008, os valores das ações começam a diminuir, com uma queda abrupta começando aproximadamente em  $t = 25$ , representando o fim do mesmo mês.

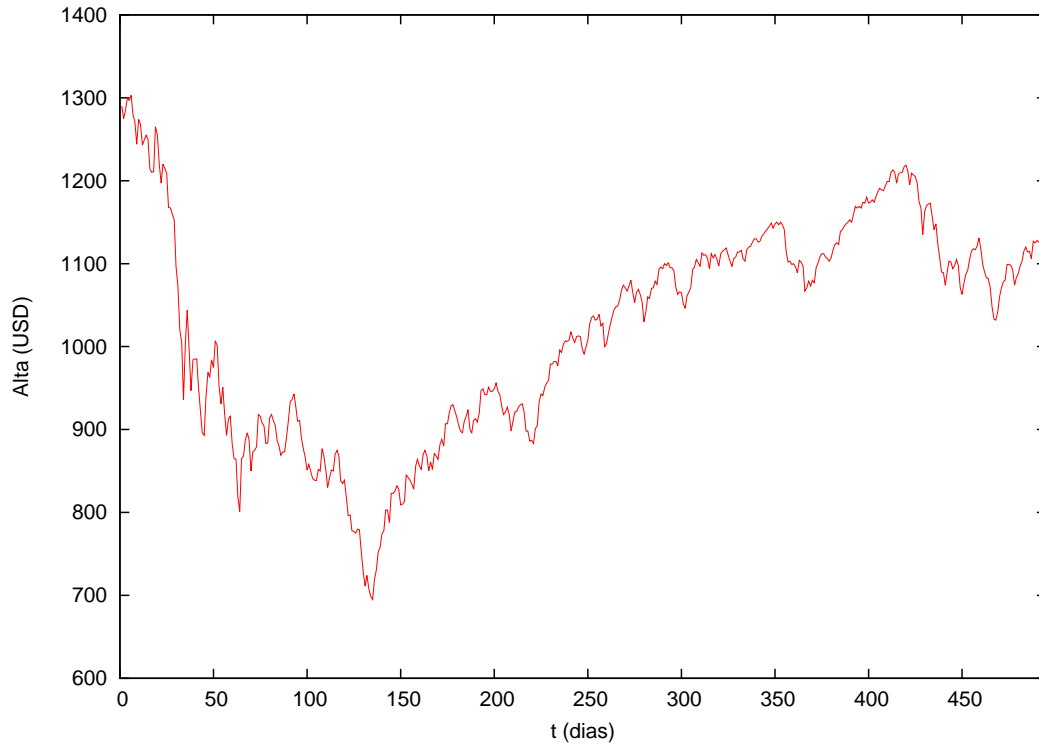


Figura 1: Valores de alta em dólares americanos do índice S&P500, abrangendo o período de 495 dias de mercado entre 25/08/2008 e 20/08/2010.

O que estamos vendo aqui foi o começo da crise do *subprime* de 2008-2009, ou *subprime mortgage crisis*, que foi uma crise financeira desencadeada pela quebra das instituições de crédito dos Estados Unidos, que concediam empréstimos hipotecários de alto risco. Há um mínimo global aproximadamente em  $t = 135$ , que é o dia 9 de março de 2009, onde as ações das empresas alcançaram os menores valores durante a crise. Desde então podemos notar uma tendência de recuperação do mercado.

### 3 Probabilidade

Aqui começaremos a definir a base teórica e algumas das variáveis [10, 11] usadas neste trabalho. Será apresentado apenas o caso discreto, pois foi o usado no trabalho devido aos cálculos computacionais, mas todas podem também ser definidas no caso contínuo, apenas trocando o somatório por uma integral.

#### 3.1 Probabilidade Marginal e Definições

Dada uma variável aleatória  $X$ ,  $x$  é o resultado que a variável aleatória assume de um dos possíveis valores do conjunto  $A_X = \{a_1, a_1, \dots, a_I\}$ , com probabilidade  $P_X = \{p_1, p_2, \dots, p_I\}$ , de maneira que  $P(x = a_i) = p_i$  com  $p_i \geq 0$  e  $\sum_{a_i \in A_X} P(x = a_i) = 1$ .

Por simplicidade de notação muitas vezes usaremos  $P(x = a_i) = P(x)$  e  $\sum_{a_i \in A_X} P(x = a_i) = \sum_x P(x)$ .

A partir disto dizemos que o valor esperado (ou média ponderada) da variável aleatória  $X$  é

$$\langle X \rangle = \sum_x x P(x), \quad (1)$$

e a variância de  $X$  é

$$Var(X) \equiv \sigma_X^2 = \langle (X - \langle X \rangle)^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2, \quad (2)$$

de maneira que o desvio padrão de  $X$  fica

$$\sigma_X \equiv \sqrt{Var(X)} = \sqrt{\langle X^2 \rangle - \langle X \rangle^2}. \quad (3)$$

## 3.2 Probabilidade Multivariacional

**Probabilidade Conjunta** - Dadas duas variáveis aleatórias  $X$  e  $Y$ , onde  $x$  e  $y$  são, respectivamente, os resultados que as v.a.  $X$  e  $Y$  assumem, com  $x \in A_X = \{a_1, a_1, \dots, a_I\}$  e  $y \in A_Y = \{b_1, b_1, \dots, b_J\}$ , dizemos que  $P(x = a_i, y = b_j) = P(x, y)$  é a probabilidade conjunta.

$X$  e  $Y$  não são necessariamente independentes, de maneira que  $P(x, y) = P(x)P(y)$  se, e somente se,  $X$  e  $Y$  são variáveis aleatórias independentes.

**Probabilidade Marginal** - Da probabilidade conjunta obtemos a probabilidade marginal  $P(x)$  pela soma:

$$P(x = a_i) = \sum_{y \in A_Y} P(x = a_i, y). \quad (4)$$

De maneira similar, usando a notação simplificada, a probabilidade marginal  $P(y)$  é:

$$P(y) = \sum_x P(x, y). \quad (5)$$

Onde concluímos que

$$\sum_{xy} P(x, y) = 1.$$

**Probabilidade Condicional** - A probabilidade condicional é dada por:

$$P(x = a_i | y = b_j) = \frac{P(x = a_i, y = b_j)}{P(y = b_j)}, \quad (6)$$

se  $P(y = b_j) \neq 0$ .

Lê-se  $P(x = a_i | y = b_j)$  como "a probabilidade que  $x$  seja igual a  $a_i$  dado que  $y$  é igual a  $b_j$ ".

Na notação simplificada temos que

$$P(y|x) = \frac{P(x,y)}{P(x)}, \quad (7)$$

se  $P(x) \neq 0$ .

As equações (6) e (7) não estão definidas para  $P(y) = P(x) = 0$ . A partir delas podemos reescrever a probabilidade conjunta (5) como:

$$P(x,y) = P(x|y)P(y) = P(y|x)P(x). \quad (8)$$

A partir de (4), (5) e (8) reescrevemos as probabilidades marginais como:

$$P(x) = \sum_y P(x,y) = \sum_y P(x|y)P(y), \quad (9)$$

e também

$$P(y) = \sum_x P(x,y) = \sum_x P(y|x)P(x). \quad (10)$$

### 3.3 Teorema de Bayes

Dos resultados acima derivados em (7), (8) e (9), segue diretamente que:

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)} = \frac{P(x|y)P(y)}{\sum_{y'} P(x|y')P(y')},$$

onde  $y$  assume um valor fixo, no qual estamos interessados, e  $y'$  é a variável que soma sobre todos os valores possíveis do conjunto  $A_Y$ .

Este resultado é conhecido como o Teorema de Bayes. Não usamos explicitamente o teorema de Bayes neste trabalho, mas é um resultado importante que segue diretamente de tudo que precisamos definir até aqui e que pode ser

usado recursivamente para melhorarmos nossa estimativa de uma distribuição de probabilidades. No próximo capítulo será apresentada uma variável que pode ser muito útil em combinação com este resultado.

### 3.4 Covariância e Correlação

Dadas duas variáveis aleatórias  $X$  e  $Y$ , de (2) temos que a covariância entre elas é

$$Cov(X, Y) \equiv \sigma_{XY} = \langle X, Y \rangle = \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle = \langle XY \rangle - \langle X \rangle \langle Y \rangle, \quad (11)$$

de maneira que  $\sigma_{XX} = \sigma_X^2$ .

A correlação [10] entre  $X$  e  $Y$  é definida, a partir de (3) e (11), como:

$$Corr(X, Y) \equiv \rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\langle XY \rangle - \langle X \rangle \langle Y \rangle}{\sqrt{\langle X^2 \rangle - \langle X \rangle^2} \sqrt{\langle Y^2 \rangle - \langle Y \rangle^2}}, \quad (12)$$

onde  $-1 \leq \rho_{XY} \leq 1$ . Para  $\rho_{XY} = 1$ ,  $X$  e  $Y$  estão completamente correlacionadas e possuem o mesmo comportamento. Para  $\rho_{XY} = -1$ ,  $X$  e  $Y$  estão correlacionadas, mas possuem comportamentos em sentidos opostos. Se  $\rho_{XY} = 0$ ,  $X$  e  $Y$  são independentes.

## 4 Teoria da Informação

Neste capítulo começamos a definição das variáveis que são de maior importância para o presente trabalho, através das quais faremos as análises para a investigação do comportamento do mercado de capitais. De novo nos focaremos apenas no caso discreto, que é o de interesse devido aos cálculos computacionais, mas todas as variáveis também estão definidas no caso contínuo, com a simples substituição do somatório pela integral.

### 4.1 Conteúdo Informacional de Shannon

O significado teórico da palavra *informação* não é muito diferente do que se tem conhecimento pelo senso-comum: uma mensagem é considerada informativa se ela diz algo que era desconhecido ou pouco esperado. Isto significa que as mensagens menos prováveis transportam mais informação. A partir desta ideia podemos chegar à importante conclusão de que se for possível medir a probabilidade de ocorrência de uma mensagem, então, de forma similar, também podemos medir seu conteúdo de informação [12].

O conteúdo informacional de Shannon de um resultado, medido em *bits*, é definido como sendo

$$s(x) \equiv \log_2 \frac{1}{P(x)}, \quad (13)$$

onde a referência nominal é feita a Claude Shannon [13], que desenvolveu sozinho praticamente toda a Teoria da Informação.

Esta fórmula corrobora a noção de que a quantidade de informação aumenta se a probabilidade de uma mensagem tende a zero, enquanto no outro extremo, quando a probabilidade de uma mensagem é  $P(x) = 1$ , ou seja, quando se sabe exatamente o que será transmitido, a quantidade de informação recebida é nula:  $s(x) = 0$ .

Espera-se que o conteúdo de informação associado a duas mensagens independentes, com probabilidades  $P(x_1) = p_1$  e  $P(x_2) = p_2$ , seja dado pela soma dos conteúdos de informação dessas mensagens. Ou seja, espera-se que  $s(x_1x_2) = s(x_1) + s(x_2)$ . Esta propriedade também é verificada pela definição do conteúdo informacional:

$$s(x_1x_2) = \log_2 \frac{1}{p_1p_2} = \log_2 \frac{1}{p_1} + \log_2 \frac{1}{p_2} = s(x_1) + s(x_2).$$

## 4.2 Entropia Informacional de Shannon

A entropia informacional (ou entropia marginal) [11, 13] de uma variável aleatória  $X$  é definida como sendo a média (1) do conteúdo informacional (13) de um resultado:

$$S^{bits}(X) \equiv \sum_{x \in A_X} s(x)P(x) = \sum_x P(x) \log_2 \frac{1}{P(x)},$$

com a convenção que  $P(x) = 0 \Rightarrow 0 \times \log 1/0 = 0$ , visto que  $\lim_{t \rightarrow 0^+} t \log 1/t = 0$ .

Aqui a entropia, assim como o conteúdo informacional, é medida em bits.

Duas propriedades importantes da entropia:

- $S(X) \geq 0$ , com igualdade se, e somente se,  $P(x) = 1$  para um  $x$ ;
- A entropia é maximizada se  $P_X$  é uniforme:

$$S^{bits}(X) \leq \log_2(|A_X|), \text{ com igualdade se, e somente se, } P(x) = 1/|A_X| \text{ para todo } x.$$

Aqui  $|A_X|$  denota o número de elementos do conjunto  $A_X$ .



No caso de nosso interesse  $|A_X| = n$ , onde  $n$  é o número de bins que usamos para dividir os dados para a obtenção das probabilidades. Disto, e da segunda propriedade acima listada, segue que:

$$S(X) \equiv \frac{S^{bits}(X)}{\log_2(n)} \leq 1,$$

onde

$$S(X) = \sum_x P(x) \log_n \frac{1}{P(x)} \quad (14)$$

é a entropia normalizada, pois  $\frac{\log_2(1/p_i)}{\log_2(n)} = \log_n(1/p_i)$ .

**Redundância** - A partir da segunda propriedade da entropia, ainda podemos definir a redundância da variável aleatória  $X$  como:

$$R \equiv 1 - \frac{S^{bits}(X)}{\log_2(|A_X|)},$$

onde, no caso de nosso interesse, para  $|A_X| = n$ ,  $R = 1 - S(X)$ .

A entropia de  $X$  também é chamada de incerteza de  $X$ . De agora em diante adotaremos apenas a entropia normalizada neste trabalho.

### 4.3 Entropia Multivariacional

Aqui seguimos com a definição de entropia, expandindo o conceito para duas variáveis aleatórias [11].

**Entropia Conjunta** - Dadas duas variáveis aleatórias  $X$  e  $Y$ , substituindo a probabilidade marginal  $P(x)$  pela probabilidade conjunta  $P(x, y)$  em (14), e somando sobre  $x$  e  $y$ , segue que a entropia conjunta é:

$$S(X, Y) = \sum_{x \in A_X} \sum_{y \in A_Y} P(x, y) \log_n \frac{1}{P(x, y)}. \quad (15)$$

A entropia conjunta é aditiva para variáveis aleatórias independentes:

$$S(X, Y) = S(X) + S(Y) \text{ se, e somente se, } P(x, y) = P(x)P(y).$$

**Entropia Condicional e Equivocação** - A entropia condicional de  $X$  dado que  $y = b_j$  é a entropia da distribuição de probabilidades  $P(x|y = b_j)$ . Usando  $P(x|y)$  no lugar de  $P(x)$  em (14) e somando sobre  $x$  temos:

$$S(X|y = b_j) = \sum_{x \in A_X} P(x|y = b_j) \log_n \frac{1}{P(x|y = b_j)}. \quad (16)$$

A entropia condicional de  $X$  dado  $Y$  é a média (1) sobre  $y$  da entropia condicional (16) de  $X$  dado  $y$ :

$$\begin{aligned} S(X|Y) &= \sum_{y \in A_Y} P(y)S(X|y) = \sum_{y \in A_Y} P(y) \left[ \sum_{x \in A_X} P(x|y) \log_n \frac{1}{P(x|y)} \right] \\ &= \sum_{xy} P(y)P(x|y) \log_n \frac{1}{P(x|y)} = \sum_{xy} P(x, y) \log_n \frac{1}{P(x|y)}, \end{aligned}$$

onde também foi usada a equação (8).

Este resultado também é conhecido como a equivocação de  $X$  sobre  $Y$ . Ela mede a incerteza média que resta sobre  $X$  quando  $Y$  é conhecida.

As entropias conjunta, condicional e marginal estão relacionadas por:

$$S(X, Y) = S(X) + S(Y|X) = S(Y) + S(X|Y). \quad (17)$$

**Entropia Relativa ou Divergência de Kullback-Leibler** - A divergência de Kullback-Leibler [14, 15], também chamada de entropia relativa ou ganho de informação, entre duas distribuições de probabilidade  $P(x)$  e  $Q(x)$ , definidas sobre o mesmo conjunto  $A_X$ , é

$$D_{KL}(P||Q) = \sum_x P(x) \log_n \frac{P(x)}{Q(x)}.$$

Ela satisfaz a *desigualdade de Gibbs*:

$$D_{KL}(P||Q) \geq 0$$

com igualdade apenas se  $P = Q$ . Note que no geral a entropia relativa não é simétrica sob a troca das distribuições  $P$  e  $Q$ : no geral  $D_{KL}(P||Q) \neq D_{KL}(Q||P)$ , então, embora algumas vezes  $D_{KL}$  também seja chamada de "distância KL", ela não é estritamente uma distância.

A  $D_{KL}$  não foi usada neste trabalho, mas pode ser muito útil em conjunto com o teorema de Bayes (Seção 3.3). Com a  $D_{KL}$  podemos fazer uma comparação entre uma nova distribuição de probabilidades, gerada a partir do teorema de Bayes, e a velha, e assim verificar se houve ganho de informação. Também pode ser aplicada recursivamente.

#### 4.4 Informação Mútua

Uma quantidade muito importante na teoria da informação, que também é a mais importante no presente trabalho, a qual usaremos para estudar o comportamento do mercado de capitais, é a informação mútua [16, 17]. Ela mede a redução média na incerteza sobre  $x$  quando sabemos o valor de  $y$ ; ou, em outras palavras, a quantidade média de informação que pode ser obtida de  $X$  quando conhecemos  $Y$ . Ela é definida como:

$$I(X; Y) \equiv S(X) - S(X|Y) = S(X) + S(Y) - S(X, Y), \quad (18)$$

onde foi usada a equação (17). Ela satisfaz  $I(X; Y) = I(Y; X)$ , ou seja, é simétrica: a quantidade média de informação que obtemos de  $X$  ao conhecer  $Y$  é a mesma que obtemos de  $Y$  ao conhecer  $X$ .

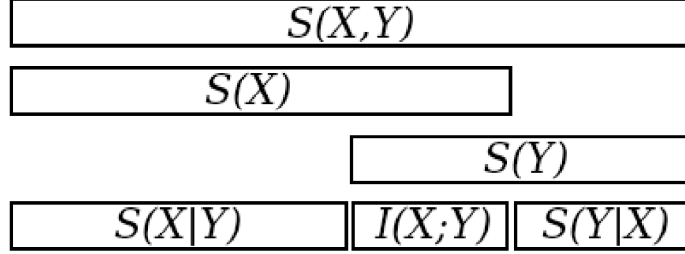


Figura 2: Relação entre Entropia Conjunta, Entropia Marginal, Entropia Condicional e Informação Mútua.

## 5 Cálculos e Resultados

### 5.1 Método Numérico

O estudo foi realizado sobre os valores de alta das ações de  $N = 349$  empresas durante  $T = 495$  dias. Depois de escolhido o número de bins  $n$ , que precisa ser o mesmo para todos os dias, para cada dia  $t \in \{1, 2, \dots, T\}$  calculamos o tamanho do bin como  $b_t = [h_{max}^\mu(t) - h_{min}^\mu(t)]/n$ , onde denominamos como  $h^\mu(t)$  o valor de alta no tempo  $t$  das ações de uma empresa  $\mu \in \{1, 2, \dots, N\}$ . De agora em diante nos referiremos a  $h^\mu(t)$  apenas como o valor das ações da empresa  $\mu$  no tempo  $t$ , ficando subentendido que são os valores de alta.

Assim, a média e a média do quadrado dos valores das ações no dia  $t$  são

$$\langle h(t) \rangle = \frac{1}{N} \sum_{\mu=1}^N h^\mu(t), \quad \langle h(t)^2 \rangle = \frac{1}{N} \sum_{\mu=1}^N [h^\mu(t)]^2, \quad (19)$$

e a variância (2) é

$$\sigma_t^2 = \frac{1}{N} \sum_{\mu=1}^N [h^\mu(t) - \langle h(t) \rangle]^2. \quad (20)$$

De (12), (19) e (20) calculamos a correlação entre dois dias  $t$  e  $t'$  como

$$Corr(t, t') = \frac{\langle h(t)h(t') \rangle - \langle h(t) \rangle \langle h(t') \rangle}{\sqrt{\langle h(t)^2 \rangle - \langle h(t) \rangle^2} \sqrt{\langle h(t')^2 \rangle - \langle h(t') \rangle^2}}, \quad (21)$$

onde  $\langle h(t)h(t') \rangle = \frac{1}{N} \sum_{\mu=1}^N h^\mu(t)h^\mu(t')$ . Repetindo o cálculo para todos os pares de dias obtemos uma matriz de correlação entre todos os dias.

A IM pode ser calculada diretamente através das entropias (18), e as entropias são variáveis que dependem apenas das distribuições de probabilidades, então para calcularmos a matriz da IM entre todos os dias precisamos apenas construir uma matriz de probabilidades conjuntas.

Para construir a matriz de probabilidades conjuntas, a cada par de dias calculamos a frequência relativa com que os valores das ações estavam contidos em um determinado bin. Para cada valor de  $h^\mu(t)$  calculamos  $[h^\mu(t) - h_{min}^\mu(t)]/b_t$  para verificar a qual bin pertence o valor  $h^\mu(t)$ . Dessa maneira criamos uma distribuição de frequências de onde tiramos a frequência relativa de cada bin. As entropias são calculadas somando sobre todos os bins.

## 5.2 Correlação x Informação Mútua

A informação mútua pode ser vista como uma correlação generalizada, mas sensível a qualquer relação entre as variáveis, e não apenas à dependência linear [18]. Aqui faremos uma comparação entre a informação mútua e a correlação entre os dias da série temporal estudada. Usamos  $n = 36$  bins.

Em todos os gráficos abaixo  $t_0$  foi um parâmetro mantido fixo e foram plotadas as curvas das correlações e IM de todos os outros dias  $t$  em relação a  $t_0$ , por isso em todos os gráficos ocorre um pico em  $t = t_0$ , pois a correlação dele com ele mesmo é  $Corr(t_0, t_0) = 1$ , e a partir de (18) vemos que a IM dele com ele mesmo é  $I(t_0; t_0) = S(t_0)$ , que é a própria informação contida naquele dia, pois nenhum outro dia compartilha mais informação com ele do que ele próprio, ou seja,  $S(t_0) \geq I(t_0; t)$  para todo  $t$  com igualdade apenas

em  $t = t_0$ .

Cada gráfico possui cinco curvas superpostas de cinco dias adjacentes, pois, para dias suficientemente próximos, tanto a correlação quanto a IM variam de maneira semelhante a medida que o tempo passa, e aqui não estamos interessados nos valores absolutos da correlação e da IM, mas apenas numa comparação qualitativa de seus comportamentos.

Abaixo segue os gráficos com  $t_0$  situado em quatro diferentes momentos da série temporal que julgamos importantes para a comparação, quais sejam eles: o começo da série temporal, momentos antes da crise estourar (Figura 3); em torno do dia de menor valor das ações (Figura 4); no meio da série temporal (Figura 5); e por último no final da série temporal (Figura 6).

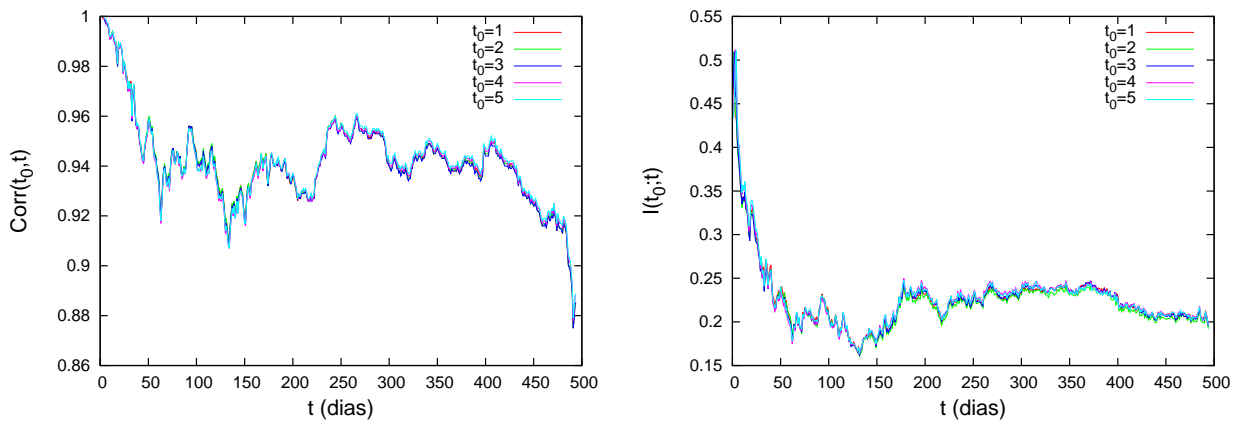


Figura 3: Esquerda: correlação. Direita: informação mútua.  $t_0 \in \{1, 2, 3, 4, 5\}$ , situado no começo da série temporal, pouco antes da crise.

Na Figura 3 podemos perceber que tanto a correlação quanto a IM apresentam um mínimo em  $t \approx 135$ , de acordo com o mínimo global dos valores das ações na Figura 1. Podemos observar também que em ambos os gráficos ocorre um primeiro mínimo por volta de  $t \approx 60$ , que em comparação com a Figura 1 vemos que é o segundo mínimo mais evidente da série temporal, sendo a primeira grande queda após o *crash* da bolsa. Há ainda um terceiro mínimo em  $t \approx 220$ , ainda de acordo com uma ligeira queda nos valores das ações após o começo da recuperação do mercado.

Neste primeiro intervalo vemos que tanto a correlação quanto a IM apre-

sentam mínimos de acordo com os mínimos dos valores das ações, no entanto fica evidente que, enquanto a correlação oscila bastante com o tempo, tornando difícil identificar quando ocorre um valor mínimo significativo, a IM apresenta apenas três mínimos bem definidos, oscilando pouco com a passagem do tempo. Da IM fica também claro que há um vale aproximadamente entre  $t = 90$  e  $t = 170$ , de novo em perfeito acordo com o que nos mostra a Figura 1. Do gráfico da correlação não conseguimos tirar esta mesma informação.

Outro ponto que fica evidente nesta comparação é a escala. Enquanto a correlação se mantém durante praticamente todo o tempo oscilando entre 1 e 0.9, ou seja, mesmo os menores valores ainda apresentam uma alta correlação com os primeiros dias, a IM em  $t = 60$  já baixou de 0.5 para menos de 0.2, mostrando que junto com a queda da bolsa houve também uma grande queda na informação compartilhada com os primeiros dias.

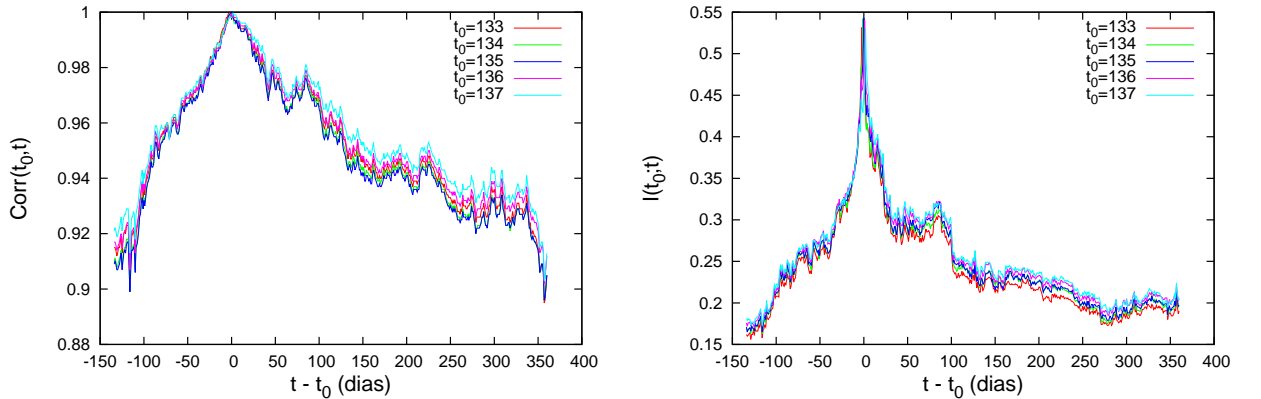


Figura 4: Esquerda: correlação. Direita: informação mútua.  $t_0 \in \{133, \dots, 137\}$ , em torno do dia de menor valor das ações. As ações atingiram seus menores valores em  $t = 135$ .

Aqui e nas próximas duas figuras plotamos as curvas em função de  $\Delta t \equiv t - t_0$  para estabelecer uma distância a partir do ponto de máximo e assim compararmos como a IM e a correlação se comportam para  $t > t_0$  e  $t < t_0$ , sendo o sentido de  $\Delta t$  crescente a recuperação do mercado, onde o mercado tende a se comportar mais uniformemente, e o sentido de  $\Delta t$  decrescente o

começo da série temporal, onde aconteceu a crise e houve grandes variações nos preços das ações.

Na Figura 4 podemos perceber claramente que no intervalo  $-30 \leq \Delta t \leq 30$ , referente ao período entre  $t = 105$  e  $t = 175$  da Figura 1, onde ocorreu os menores valores das ações, a IM cai bastante rápido, descendo de  $I \approx 0.55$  para  $I \approx 0.3$ , enquanto que a correlação varia muito pouco, permanecendo entre 1 e 0.98. Portanto, mesmo em meio à crise, a correlação nos diz que os dias próximos a  $t_0$  parecem bastante correlacionados com ele, o que pode ser verdade para qualquer outro período da bolsa, enquanto que a IM apresenta uma queda grande, o que é atípico para dias adjacentes, como veremos a seguir. Tanto a IM quanto a correlação diminuem mais rapidamente para  $\Delta t < 0$ , em direção à crise, do que para  $\Delta t > 0$ , em direção à recuperação do mercado.

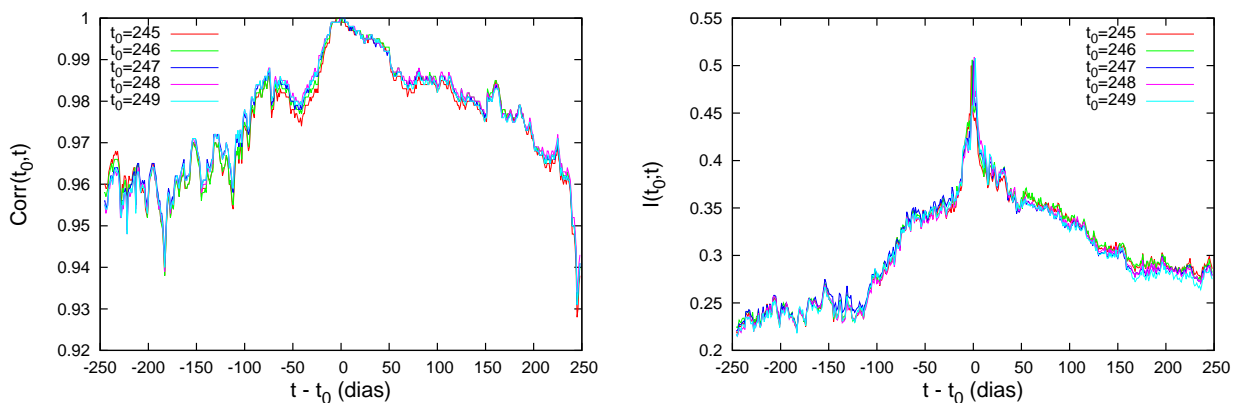


Figura 5: Esquerda: correlação. Direita: informação mútua.  $t_0 \in \{245, \dots, 249\}$ , situado no meio da série temporal.

Na Figura 5 plotamos as cinco curvas centradas em torno do meio da série temporal, para tentarmos comparar melhor o período da crise com o período de recuperação do mercado. Aqui fica evidente a assimetria entre ambos os lados,  $\Delta t < 0$  e  $\Delta t > 0$ , em ambos os gráficos.

No gráfico da correlação vemos que para  $\Delta t < 0$  há uma oscilação maior, enquanto que para  $\Delta t > 0$  o comportamento é mais uniforme. Para  $\Delta t < 0$  há um mínimo que parece se destacar em  $\Delta t \approx -180$ , referente à  $t \approx 67$  na



Figura 1, perto do primeiro mínimo da crise. Apesar disso, se não tivéssemos a priori a informação da crise, devido à grande oscilação da correlação, poderia ser difícil assinalar algum comportamento importante neste ponto.

Para a IM podemos ver que em torno de  $\Delta t = 0$  ela decresce mais lentamente do que no cenário anterior, variando entre 0.5 e 0.35 (contra 0.55 e 0.3). Entretanto, o que fica mais evidente, é que para  $\Delta t > 0$  a IM decresce mais uniformemente, lembrando uma lei de potência, enquanto que para  $\Delta t < 0$  há uma queda abrupta de  $\Delta t \approx -75$  até  $\Delta t \approx -110$ , este último sendo referente à  $t = 135$  na Figura 1, o dia de menor valor das ações, e permanecendo próximo a este valor  $I \approx 0.23$  em todo o intervalo  $-250 \leq \Delta t \leq -110$ , que é o período que caracteriza a crise,  $1 \leq t \leq 135$  na Figura 1, onde os valores das ações estiveram em queda.

Mesmo que não tivéssemos conhecimento da crise a IM poderia indicar um acontecimento importante no intervalo  $-250 \leq \Delta t \leq -110$ .

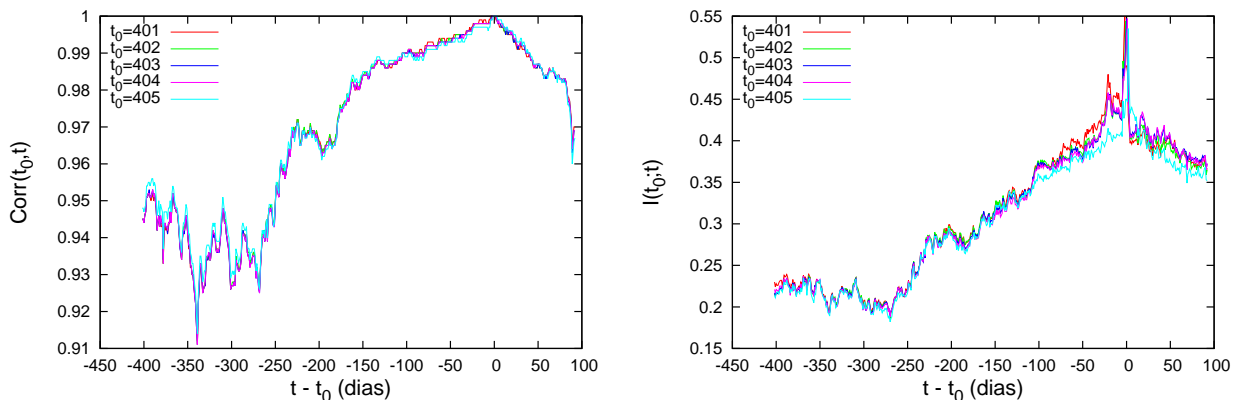


Figura 6: Esquerda: correlação. Direita: informação mútua.  $t_0 \in \{401, \dots, 405\}$ , situado perto do fim da série temporal.

Na Figura 6 plotamos as cinco curvas centradas em torno de um dia no final da série temporal, para vermos como a correlação e a IM decrescem estando afastadas das grandes variações dos preços causadas pela crise.

Aqui podemos ver que, ao contrário dos outros gráficos, a correlação decresce mais uniformemente para ambos os lados até  $\Delta t \approx -175$  e finalmente caindo bruscamente em  $\Delta t \approx -230$ , oscilando bastante no período

$-400 \leq \Delta t \leq -230$ , que é referente aos dias  $1 \leq t \leq 175$  na Figura 1, sendo este o período de grande oscilação nos preços das ações, englobando já o início da recuperação do mercado, que aconteceu a partir de  $t = 135$  (na Figura 1). Há um mínimo que se destaca em  $\Delta t \approx -345$ , referente ao dia  $t \approx 60$ , que é o segundo dia de menor valor das ações, o primeiro grande mínimo após o começo da queda. Como foi destacado também nos outros gráficos, apesar deste mínimo mais evidente, há uma grande oscilação durante esse período, o que dificulta identificar se alguma destas variações está relacionada a algum evento importante.

No gráfico da IM podemos ver que no intervalo  $-100 \leq \Delta t \leq 100$  ela decresce de maneira bastante semelhante para ambos os lados, o que não era observado nos casos anteriores, onde era sempre perceptível uma assimetria mesmo em intervalos pequenos em torno de  $\Delta t = 0$ . Podemos ver também que para  $\Delta t < 0$  a IM decresce de maneira uniforme até  $\Delta t \approx -230$ , onde fica evidente uma queda brusca, assim como na correlação. Diferente do que acontece na correlação, o mínimo da IM é em  $\Delta t \approx -270$ , referente ao dia  $t \approx 135$  na Figura 1, que é o dia de menor valor das ações. Assim como nos casos anteriores, a IM se mantém sem grandes oscilações, mesmo no período da crise, tornando assim mais fácil a identificação de um mínimo e sua possível relação com algum evento importante.

### 5.3 Informação Mútua Média

Nesta seção nos focamos na IM para tentar estabelecer parâmetros que relacionem seu comportamento com o comportamento do mercado de capitais. Primeiro estabelecemos uma janela temporal de  $\Delta T = 50$  dias, e então dividimos a série temporal em 8 intervalos de  $\Delta T$  dias, sendo eles:

- Intervalo 1: dia  $t_i = 1$  ao  $t_f = 50$ ;
- Intervalo 2: dia  $t_i = 51$  ao  $t_f = 100$ ;
- Intervalo 3: dia  $t_i = 101$  ao  $t_f = 150$ ;
- Intervalo 4: dia  $t_i = 151$  ao  $t_f = 200$ ;
- Intervalo 5: dia  $t_i = 201$  ao  $t_f = 250$ ;
- Intervalo 6: dia  $t_i = 251$  ao  $t_f = 300$ ;
- Intervalo 7: dia  $t_i = 301$  ao  $t_f = 350$ ;
- Intervalo 8: dia  $t_i = 351$  ao  $t_f = 400$ .

Após isso calculamos  $I(t_0; t')$  onde  $t' \in \{t_0, t_0 + 1, \dots, t_0 + \Delta T\}$ . Em cada intervalo calculamos uma curva com  $t_0 = t_i$ , então a janela  $\Delta T$  é deslocada de  $\delta t = 1$  dia e calculamos uma nova curva com  $t_0 = t_i + 1$ , repetindo isto até que  $t_0 = t_f$ . Assim, geramos 50 curvas diferentes em cada intervalo e fazemos a média entre elas a fim de obtermos uma curva suavizada representando cada intervalo.

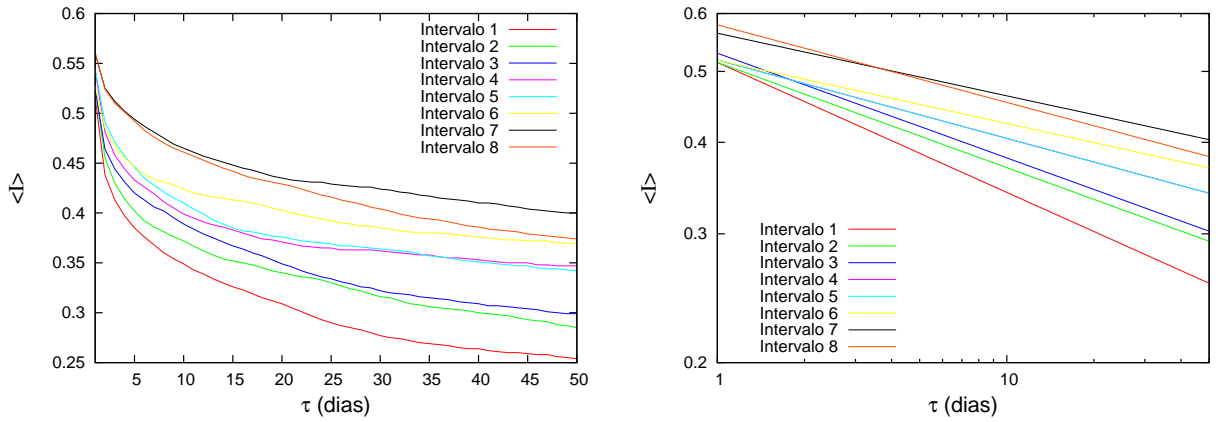


Figura 7: Curvas representando a IM média dos diferentes intervalos da série temporal. Na esquerda as curvas suavizadas reais, e na direita o fit em escala log-log.

Para as curvas da Figura 7 fitamos uma lei de potência do tipo  $f(\tau) = a\tau^{-b}$  a fim de determinarmos os parâmetros  $a$  e  $b$  com que varia a IM nos diferentes intervalos. Os parâmetros estão na Figura 8.

Como o parâmetro  $a$  é a constante de proporcionalidade, e a média da IM nos pontos de máximo, onde  $I(t_0; t_0) = S(t_0)$ , é a própria entropia média do intervalo, esperaríamos que houvesse alguma relação entre  $a$  e  $\langle S \rangle$ . Plotamos a entropia média dos mesmos intervalos na Figura 9 para comparação. De fato se percebe uma semelhança entre ambos, mas há algumas diferenças, sendo as mais explícitas no intervalo 3, onde a entropia média é menor que no intervalo 2, mas o parâmetro  $a$  cresce do intervalo 2 para o 3, e no intervalo 4, onde  $\langle S \rangle$  aumenta em relação ao 3, mas  $a$  diminui. Mesmo onde o comportamento é semelhante também há uma diferença, embora não muito grande, nos valores de  $a$  e  $\langle S \rangle$ . Como  $a$  e  $b$  estão relacionados, era de se esperar que o parâmetro  $a$  sozinho não nos desse uma representação exata de  $\langle S \rangle$ .

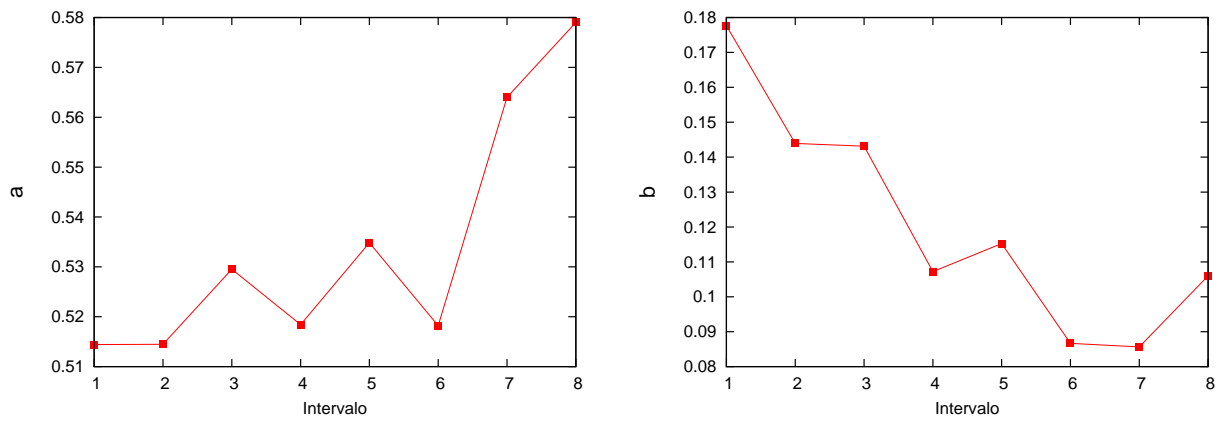


Figura 8: Parâmetros  $a$  (esquerda) e  $b$  (direita) nos diferentes intervalos.

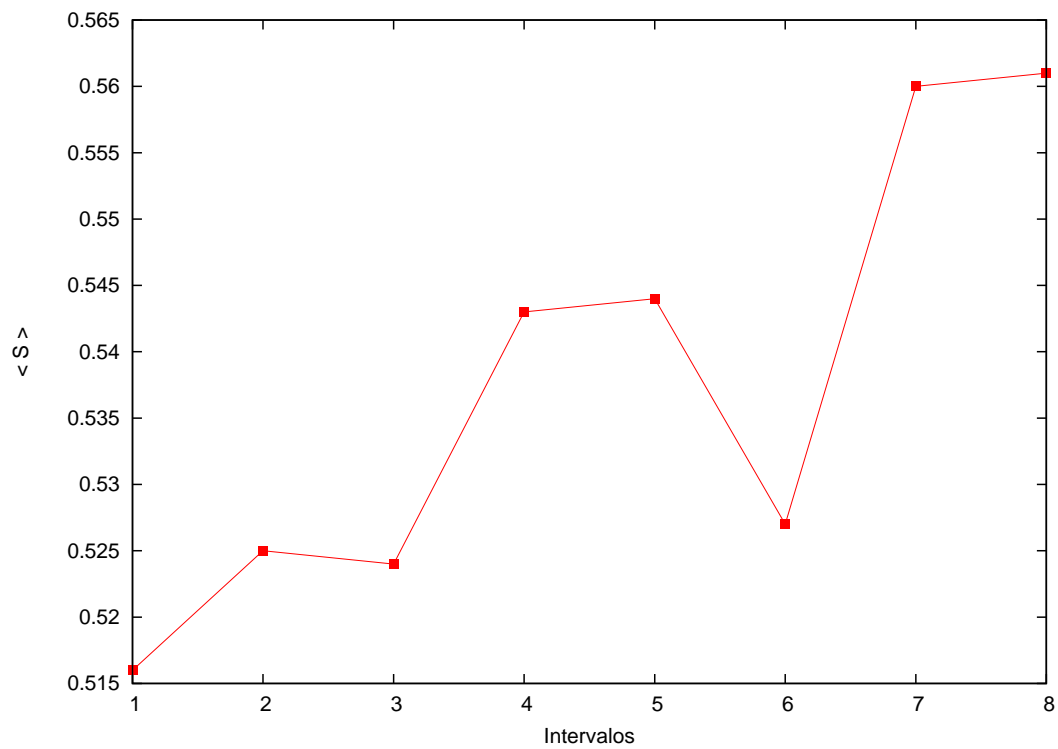


Figura 9: Entropia média dos diferentes intervalos da série temporal.

O parâmetro  $b$  está relacionado com a taxa de decaimento da IM com o

tempo. Como vimos na Seção 5.2, a IM parece decrescer mais rapidamente no período da crise, também apresentando mínimos mais acentuados nos menores valores das ações. Aqui podemos verificar isto através de  $b$ . Podemos ver que o maior valor de  $b$  é no intervalo 1, onde houve a maior queda no valor das ações, mostrando que de fato a IM decresceu mais rapidamente no começo da crise do que em qualquer outro período. Acompanhando na Figura 1, vemos que nos intervalos 2 e 3, apesar dos preços continuarem caindo, a amplitude na oscilação do valor das ações foi menor, e acompanhando o comportamento de  $b$ , vemos que ele diminui e permanece praticamente constante nesses dois intervalos, indicando que a IM decresceu mais lentamente em relação ao intervalo 1, mas ainda assim mais rápido do que em outro intervalo. Após estes intervalos o mercado começa a se recuperar, há uma menor oscilação e o valor das ações começa a subir. Podemos ver que há uma grande queda em  $b$ , portanto a IM decresce mais lentamente. O valor volta a crescer um pouco nos intervalos 5 e 8, onde, na Figura 1, podemos ver que houve ligeiras quedas no valor das ações, mas  $b$  continua abaixo do valor dos primeiros intervalos. Os menores valores, que são pelo menos a metade do valor no intervalo 1, ocorrem nos intervalos 6 e 7, que são os intervalos onde o valor das ações continua crescendo e oscilando pouco.

Disto vemos  $b$  pode estar relacionado com a volatilidade do mercado, mostrando que a IM é sensível à variação dos valores das ações, decrescendo mais rapidamente quando há grandes oscilações.

## 6 Conclusão

Este trabalho de conclusão de curso teve como objetivo o estudo de uma série temporal dos valores de alta de ações das empresas constituintes do índice S&P500 da bolsa de valores de Nova Iorque. Segundo o estudo feito em [8], é possível identificar um comportamento extremo do mercado como *bubbles* ou *crashes* a partir da entropia informacional. No estudo o autor utilizou um modelo semelhante a spins interagentes em um modelo de campo médio, onde era verificado a cada dia, para cada empresa, se o valor de suas

ações subiam (spin up) ou desciam (spin down). Outro ponto demonstrado em [8] é que o mercado pode ser explicado através da interação entre pares de empresas. A partir disso, nossa abordagem para o cálculo da entropia informacional foi diferente, se baseando diretamente no valor das ações em cada dia e construindo uma distribuição de probabilidades para cada par de empresas. Como as variáveis da teoria da informação dependem apenas das distribuições de probabilidade, com isto podemos calcular uma matriz da entropia conjunta (15) para todos os pares de empresas e utilizando a equação (18) segue diretamente o cálculo de uma matriz para a informação mútua para todos os mesmos pares.

Primeiro fizemos uma comparação entre a correlação temporal e a IM temporal, que pode ser vista como uma correlação generalizada sensível à dependência não-linear entre duas variáveis aleatórias [18]. Desta comparação pudemos concluir que a IM de fato é uma medida superior em comparação com a correlação para o comportamento do mercado. Enquanto a correlação tem uma tendência a variar bastante com o tempo, tornando difícil identificar quando ocorre algum acontecimento mais importante, a IM mostra um decrescimento mais uniforme, apresentando grandes variações apenas quando o mercado também apresenta um comportamento anômalo.

Depois, estudando mais diretamente apenas o comportamento da IM, tentamos determinar como seu comportamento se relaciona com o comportamento do mercado. Para isto calculamos os parâmetros que caracterizam uma lei de potência para diferentes curvas da IM média em diferentes intervalos da série temporal. Disto vimos que o parâmetro referente a constante de proporcionalidade está relacionado com a entropia média do intervalo, embora sozinho não consiga explicar o comportamento da entropia. O parâmetro referente ao expoente de decaimento mostrou uma relação maior com o comportamento do mercado, mostrando que de fato a IM decresce mais rapidamente em períodos de grande volatilidade do mercado, o que faz sentido se pensarmos principalmente na analogia usada em [8], pois em momentos de grande volatilidade as empresas tendem a se comportar mais individualmente, mudando sua orientação com maior frequência, e assim dificultando uma visualização da tendência geral do mercado, enquanto que em perío-

dos normais ou de pós-crise, as empresas se comportem de maneira mais uniforme.

Alguns pontos que poderiam ser usados para melhorar nossas estimativas seriam um estudo do número ótimo de bins para a maximização da entropia, o uso do teorema de Bayes junto com a divergência de Kullback-Leibler [14, 15] para a obtenção de uma distribuição de probabilidades que nos desse um ganho de informação, e também o método apresentado em [18] para estimarmos os erros em nossos cálculos da entropia e IM.

O uso da teoria da informação em estudos do mercado de capitais é recente, mas já é bastante difundido no estudo de redes neurais [16, 17]. Futuramente poderíamos trocar nossa análise temporal por uma análise estrutural (espacial), utilizando a IM para estabelecer uma matriz de conexões entre as empresas e então aplicar métodos já conhecidos do estudo de redes neurais, redes complexas e física estatística.



## Referências

- [1] BM&FBovespa, *Introdução ao Mercado de Capitais*,  
<http://www.bmfbovespa.com.br/pt-br/a-bmfbovespa/download/merccap.pdf>
- [2] T. Dal'Maso Peron, F. Rodrigues, *Collective behavior in financial markets*, EPL (Europhysics Letters) 96 (2011) 48004.
- [3] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertész, A. Kanto, *Dynamics of market correlations: taxonomy and portfolio analysis*, Physical Review E 68 (2003) 056110.
- [4] T.K.D. Peron, L. da Fontoura Costa, F.A. Rodrigues, *The structure and resilience of financial market networks*, Chaos: An Interdisciplinary Journal of Nonlinear Science 22 (2012) 013117.
- [5] H. Stanley, X. Gabaix, P. Gopikrishnan, V. Plerou, *Economic fluctuations and statistical physics: quantifying extremely rare and less rare events in finance*, Physica A 382 (2007) 286-301.
- [6] H. Stanley, V. Plerou, X. Gabaix, *A statistical physics view of financial fluctuations: evidence for scaling and universality*, Physica A 387 (2008) 3967-3981.
- [7] P. Vértes, R. Nicol, S. Chapman, N. Watkins, E. Bullmore, *Frontiers: topological isomorphisms of human brain and financial market networks*, Frontiers in Systems Neuroscience (2011).
- [8] Thomas Bury, *Market structure explained by pairwise interactions*, Physica A 392 (2013) 1375-1385.
- [9] <http://finance.yahoo.com/>
- [10] C. Scherer, *Métodos Computacionais da Física*, Cap. 4 - Probabilidade, Variável Aleatória e Processo Estocástico, Editora Livraria da Física (2010).

- [11] David MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press (2003).
- [12] Luiz Henrique Alvez Monteiro, *Sistemas Dinâmicos*, Seção 10.5 - Entropia Informacional de Shannon, Editora Livraria da Física (2011).
- [13] C. E. Shannon, *A Mathematical Method of Communication*, Bell System Technical Journal, v. 27, (1948).
- [14] S. Kullback; R. A. Leibler; *On Information and Sufficiency*, Annals of Mathematical Statistics 22, (1951).
- [15] Adom Giffin, *From Physics to Economics: An Econometric Example Using Maximum Relative Entropy*, Physica A, v.388, 1610-1620 (2009).
- [16] David R. Carreta Dominguez, Elka Korutcheva, *Three-state neural network: From mutual information to the Hamiltonian*, Phys. Rev. E 62, 2620-2628 (2000).
- [17] D. R. C. Dominguez and D. Bollé, *Self-Control in Sparsely Coded Networks*, Phys. Rev. Lett. 80, 2961-2964 (1998).
- [18] Mark S. Roulston, *Estimating the errors on measured entropy and mutual information*, Physica D, 125, 285-294 (1999).