

Nome do aluno: Rafael Paulino Veiga

Nome do orientador: Raul Ikeda Gomes da Silva

Título: Estudo da Entropia de Shannon no Mercado de Ações Brasileiro

Palavras-chave: entropia de Shannon, *high frequency trading*.

1 – Descrição do Problema e Revisão de Literatura

Entropia é uma medida utilizada em diversos ramos da ciência para o descobrimento de perdas, seja de massa, temperatura, entre outras (Pineda, 2010, p. 22). Ainda, o conceito de entropia, especificamente a entropia de Shannon, lida com a organização dos dados. Através da medida de entropia é possível estabelecer quão uniforme é a distribuição de probabilidade dos dados observados (Pineda, 2010, p. 27).

O trabalho proposto por Molgedey e Ebeling (2000) utilizou a medida de entropia no índice Dow Jones. Observou-se que no longo prazo há uma correlação entre entropia de Shannon e a entropia condicional (entropia dinâmica) nos cálculos e nas características flutuantes. Maasoumi e Racine (2002) examinaram o comportamento do mercado de ações com uma medida de entropia e identificaram evidências significativas de dependência não-linear no retorno das séries. Zapart (2009) analisou a medida de entropia de Shannon em séries financeiras de moedas no mercado internacional. Observou-se que, quando a entropia normalizada estava abaixo de um certo limiar, poderia implicar em uma formação de tendência.

Contudo, a frequência de amostragem dos dados pode alterar as medidas e levar a diferentes conclusões (Aldridge, 2010, p. 14-35). Os trabalhos citados anteriormente utilizaram janelas de dias ou meses para estabelecer um estudo de longo prazo, porém não foi levada em consideração a utilização de dados em alta frequência, ou *tick-by-tick*.

2- Objetivo

O objetivo desta proposta é reproduzir parte dos trabalhos anteriores, aplicando a entropia de Shannon para o mercado brasileiro e verificar o seu comportamento para uma frequência de tempo maior. Será aplicada a medida em diferentes séries de preços de ações e também será estudado como a frequência de observação dos dados pode influenciar no resultado.

3- Metodologia

Através das séries de preços extraídas de um repositório de dados de alta frequência, será desenvolvido um *framework* em ambiente Excel/VBA para o cálculo das medidas e avaliação de desempenho.

Os dados deverão ser pré-processados e padronizados, descartando eventuais valores errados. Para o cálculo da entropia de Shannon, será necessário construir uma tabela de probabilidades, e sua fórmula é definida por:

$$H_p = - \sum_N p(x) \log_2 p(x)$$

Onde N é a quantidade de classes e p(x) representa a probabilidade de cada classe. Após essa etapa, a medida é normalizada pela máxima entropia, dada por:

$$H_p^{max} = \log_2 N$$

Com a medida normalizada será possível testar o conceito para diversos limiares, estabelecendo um índice de desempenho comum a todas amostras. Ainda será possível analisar o comportamento para diversos agrupamentos no tempo, desde *tick-by-tick* até a resolução de dias. Após a conclusão da implementação, será possível estender a análise para diversas empresas e períodos distintos. O mesmo conceito também poderá ser expandindo para outras medidas de entropia.

4- Resultados esperados

Caso o mercado brasileiro apresente comportamento similar ao descrito nos trabalhos referenciados, poderá ser possível estabelecer uma previsibilidade na formação de tendências das séries. A análise com diferentes janelas de tempo, pode fornecer informações acerca da melhor discretização, dado um certo critério de desempenho. Contudo, como ressaltado pelos autores citados, não há compromisso na previsão da direção ou duração dessa tendência, bem como é possível também obter alarmes falsos.

Referências bibliográficas

Aldridge, I., High-Frequency Trading, 2013. 2a Edição, Wiley trading series.

Maasoumi, E., Racine, J., Entropy and predictability of stock market returns. Journal of Econometrics, 107 (1-2) 2002, p. 291-312.

