

# A multi-scale symmetry analysis of daily financial time series using uninterrupted trends returns

Carlos Manuel Rodríguez Martínez

David Hernandez Enríquez

Advisor: Dr. Alejandro Raúl Hernández Montoya

Universidad Veracruzana

16 de agosto de 2019

## Acknowledgments

- Dr. Alejandro Raúl Hernandez Montoya. Universidad Veracruzana.
- Dr. Hector Conoronel Brizio. Universidad Veracruzana.

# Introduction

# Publicly available information

Prices, volume, dividends, beta, etc.

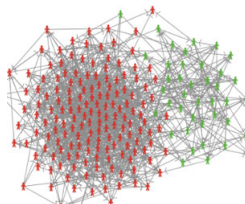
## Intel Corporation (INTC)

Best Bid/Ask	\$ 29.96 / \$ 29.97
1 Year Target	38
Today's High /Low	\$ 30.50 / \$ 29.45
Share Volume	52,122,067
50 Day Avg. Daily Volume	22,872,515
Previous Close	\$ 32.74
52 Week High/Low	\$ 37.03 / \$ 24.87
Market cap	\$ 141,428,430,000
P/E Ratio	12.81
Forward P/E(1y)	13.76
Earnings Per Share (EPS)	\$ 2.34
Annualized dividend	\$ 0.96
Ex Dividend Date	Nov. 4, 2015
Dividend Payment Date	Dec. 1, 2015
Current Yield	2.93 %
Beta	1.11
NASDAQ Official Open Price	\$ 29.63
Date of Open Price	Jan. 15, 2016
NASDAQ Official Close Price	\$ 32.74
Date of Close Price	Jan. 14, 2016
Community Sentiment	 Bullish

# What is relevant to price

- Interaction of many agents that perform buy/sell operations.
- Every agent has access to the same public information to make an informed decision on every operation. They are rational.
- The direct consequence is that current price already contains all the knowledge of future forecasts.

$$E[P_{t+1}|P_0, P_1, \dots, P_t] = P_t.$$



# Efficient market hypothesis (EMH)

- **Bachelier (1900)** proposes that since in a voluntary stock exchange between two agents, each of them believes them to be the winning party. If the game is fair (Martingale) it must behave like a random walk. His work was not well received.
- **Samuelson (1965)** proposes the efficient market hypothesis. If investors were able to incorporate future events in the current price, they would have done so, incorporating their knowledge in the current price.
- **Fama (1995)** finds empirical evidence that markets behave like a random walk.
- The EMH requires agents to be rational.

- Challenger space shuttle. Four contractors involved: Lockheed, Martin Marietta, Morton Thiokol y Rockwell.
- After 5 months of investigation it was found that the misbehaving component came from Morton Thiokol.
- Markets declared their verdict right away.

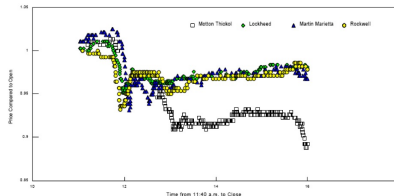
### Efficient Markets – Challenger Example

Efficient Markets - 5

Source: Michael T. Maloney and J. Harold Mulheri, The complexity of price discovery in an efficient market: the stock market reaction to the Challenger crash, *Journal of Corporate Finance*, Volume 9, Issue 4, September 2003, Pages 453-479

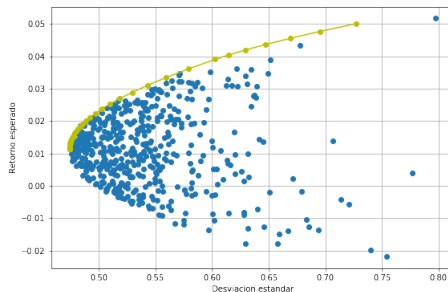
- Here are the stock prices for several major suppliers to NASA (for the day)

Intraday Stock Price Movements Following the Challenger Disaster



# If I can't win there is no point

- **Markowitz (1952)** optimal portfolio theory tell us that we can construct a portfolio such that the relation between expected return and risk is optimal.



- Useful for long term investment.
- EMH implies the reliability of predictive markets.



# There are winners however...

- The existence of agents that can reliably generate profits like Warren Buffet, Peter Lynch and George Soros, implies the possibility that some agents can sometimes predict better than others future outcomes or even affect the market themselves.
- Humans have cognitive biases that systematically makes us averse to lose and have high hopes of winning. We are not always rational.

# Motivation

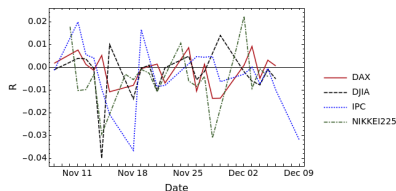
How can we measure if the market really behaves like a random walk?

# Background

# Definitions

- We define the logarithmic returns as

$$R(t, \Delta t) = \log P_{t+\Delta t} - \log P_t.$$



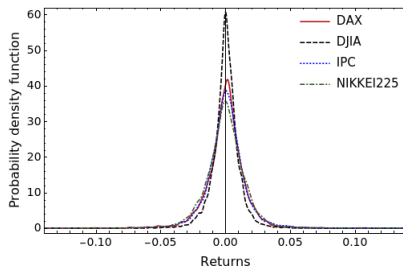
- We can express the efficient market hypothesis in terms of the returns

$$E[R_{t+1} | R_0, R_1, \dots, R_t] = 0.$$

- From this definition of the returns emerge the *stylized facts*, that are statistical properties common to the returns of most markets.

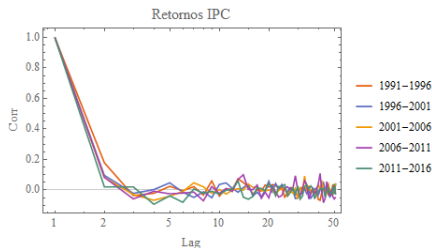
# Stylized facts I

- Win / Loss assymetry. Symmetry is strongly related to the EMH.

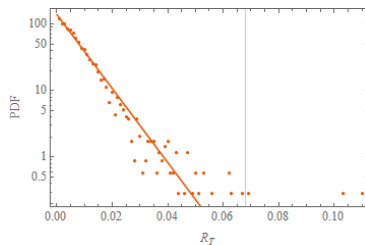


- No auto-correlation on the returns

# Stylized facts II

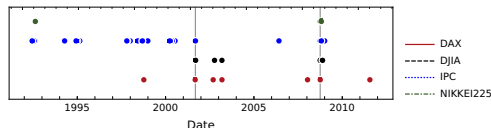


- Returns decay as a power law.



# Stylized facts III

- Some events extremely rare (according to the power law) still happen sometimes. We identify these extreme events as the Dragon Kings described by Sornette. Usually these extreme values signal a big transition in the market like an economic crisis.

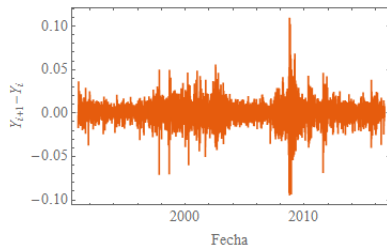
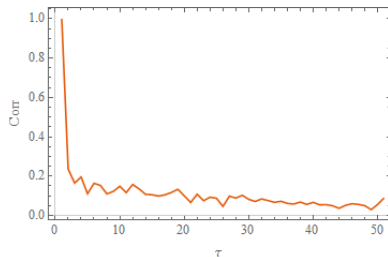


- Volatility clustering

$$C_2(\tau) = \text{Corr}(R(t + \tau, \Delta t)^2, R(t, \Delta t)^2).$$

That is, big changes tend to be followed by big changes of any sign, and small changes tend to be followed also by small changes of any sign.

# Stylized facts IV





## An example of application: Distribution of trends

# Model

- Among all the possible martingale or sub-martingale models that can describe price fluctuations, the geometric random walk is the simplest one.
- At each step there are two possible outcomes: the index either increases or does not increase. In an efficient market, the expected future price depends only on information about the current price, not on its previous history.
- It should be impossible to predict the expected direction of a future price change given the history of the price process. We have

$$\mathbb{E}(S(t + \Delta t) - S(t) | \mathcal{F}_t) = 0; \quad (1)$$

if we consider the sign of the price change

$Y(t, \Delta t) = \text{sign}(S(t + \Delta t) - S(t))$ , which coincides with the sign of returns, we accordingly have

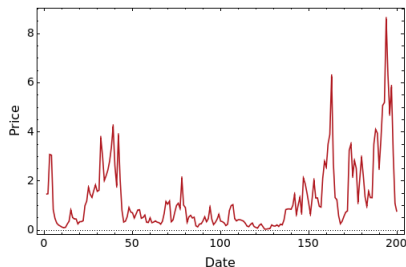
$$\mathbb{E}(Y(t, \Delta t)) = 0. \quad (2)$$

# Model II

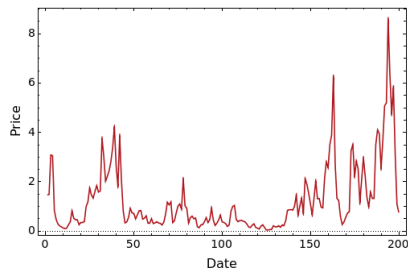
- If the price follows a geometric random walk, then the series of price-change signs can be modeled as a Bernoulli process. Let  $S_0$  be the initial price. The price at time  $t$  will be given by

$$S(t) = S_0 \prod_{i=1}^t Q_i \quad (3)$$

where  $Q_i$  are independent and identically distributed random variables following a log-normal distribution with parameters  $\mu$  and  $\sigma$ .



# Model III



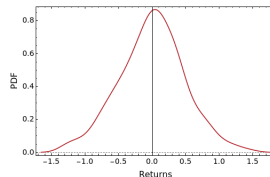
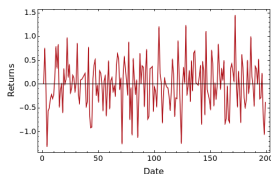
Example of time series:

$$Q = \{0.97, 0.81, 0.71, 1.82, 1.79, 1.02, 0.84, 1.13, 0.76, 2.53\}$$

$$S = \{0.97, 0.79, 0.56, 1.03, 1.84, 1.89, 1.59, 1.80, 1.37, 3.48\}$$

# Model IV

- Note that this price series reproduce a gaussian distribution of returns



- These two parameters come from the corresponding normal distribution for log-returns. As a direct consequence of the EMH we have

$$\mathbb{E}(Q) = 1 + r_F, \quad (4)$$

and for a log-normal distributed random variable, we have also

$$\mathbb{E}(Q) = e^{\mu} e^{\sigma^2/2}. \quad (5)$$

# Model V

- This leads to a dependence between the two parameters

$$\mu = \log(1 + r_F) - \frac{\sigma^2}{2}. \quad (6)$$

- Starting from the cumulative distribution function for a log-normal random variable

$$F_Q(u) = \mathbb{P}(Q \leq u) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{\log(u) - \mu}{\sqrt{2}\sigma} \right), \quad (7)$$

the probability of a negative sign would be given by

$$q = F_Q(1) = \mathbb{P}(Q \leq 1) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{\sigma}{2\sqrt{2}} - \frac{\log(1 + r_F)}{\sigma\sqrt{2}} \right), \quad (8)$$

which yields  $q = 1/2$  for  $r_F = e^{\sigma^2/2} - 1$ .

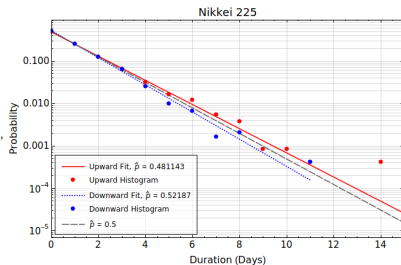
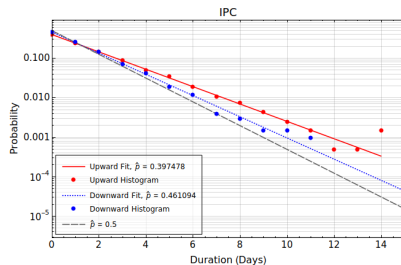
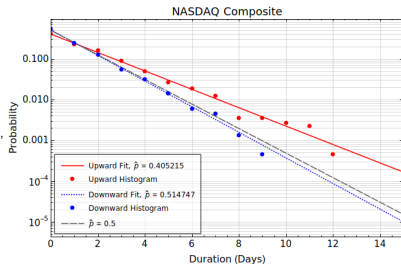
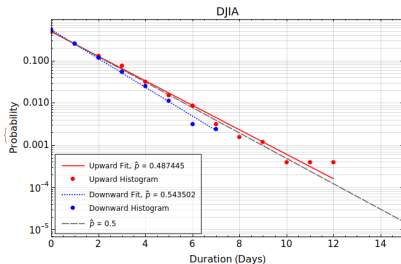
# Model VI

- It becomes natural to use the biased Bernoulli process as the null hypothesis for the time series of signs. It is well known that the distribution of the number  $k$  of failures needed to get one success for a Bernoulli process with success probability  $p = 1 - q$  is the geometric distribution  $\mathcal{G}(p)$ ; the number of failures  $N$  is given by

$$P(k) = \mathbb{P}(N = k) = p(1 - p)^k = pq^k. \quad (9)$$

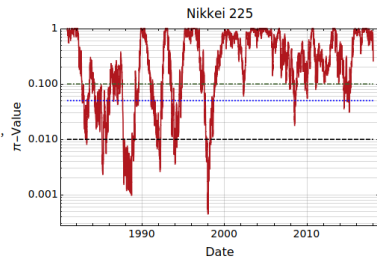
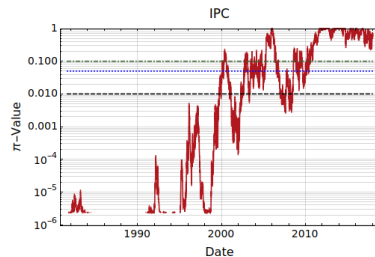
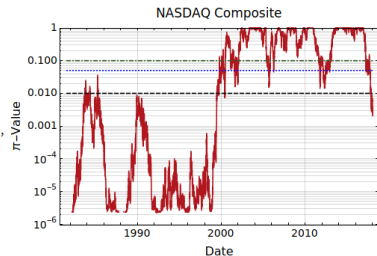
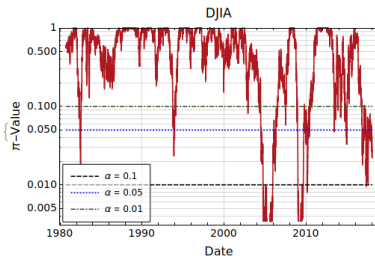
- In order to compare the observed and expected distributions of trend durations, the Anderson-Darling test was used. The Anderson-Darling test was found to be the most suitable for this purpose because it places more weight on the tails of a distribution than other goodness of fit tests.

# Distribution of runs

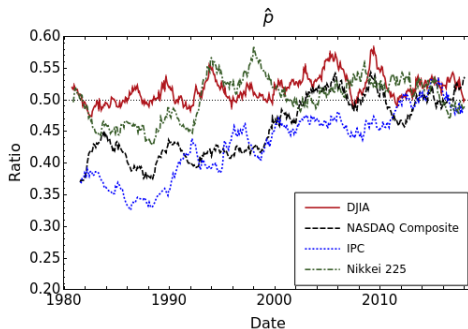




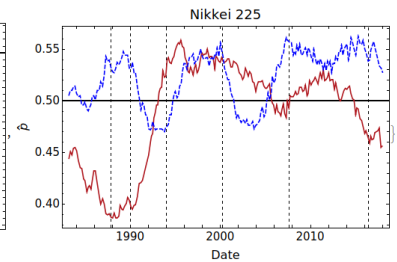
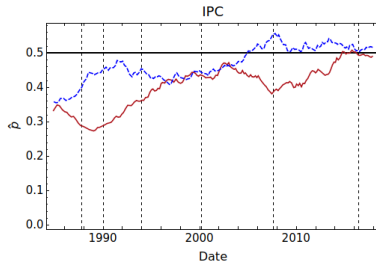
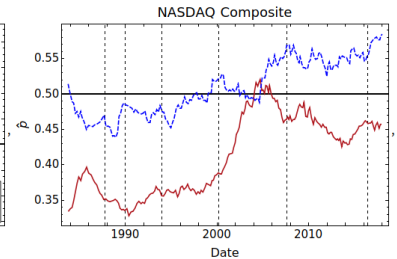
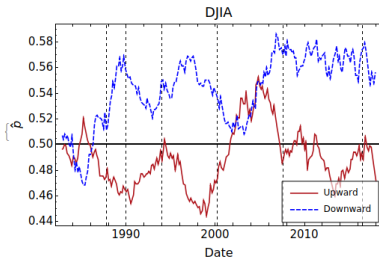
# Test the data with a Geometric Distribution $p = 0.5$



# Estimation of $\hat{p}$ for the runs with a geometric distribution

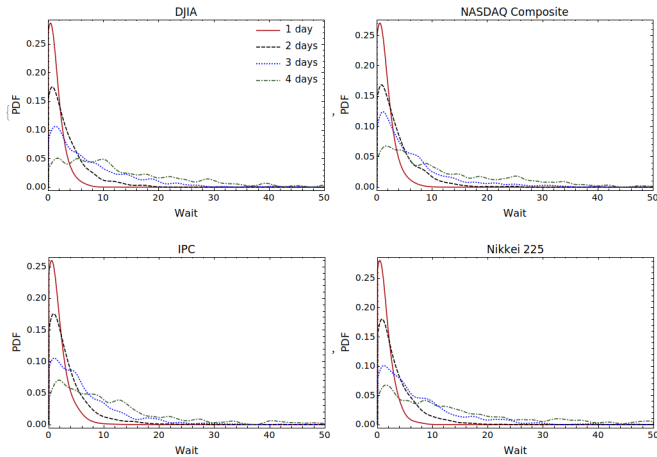


# Estimation of $\hat{p}$ for positive/negative runs with a geometric distribution

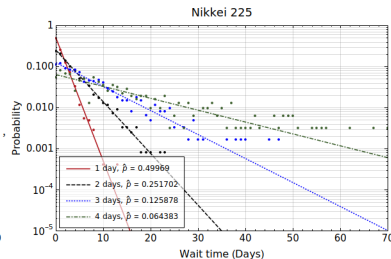
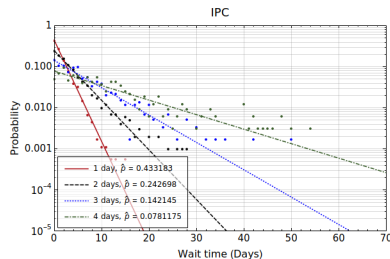
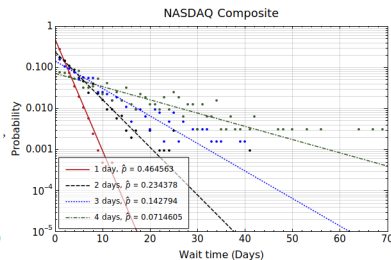
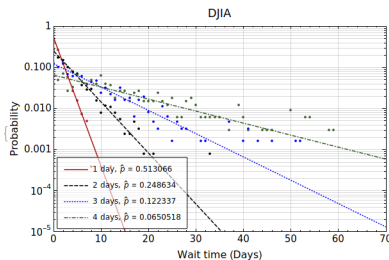


# Waiting time distribution

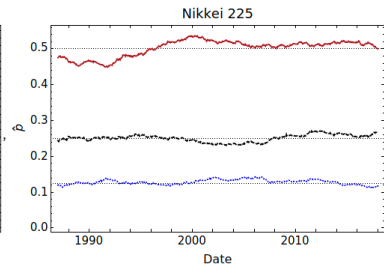
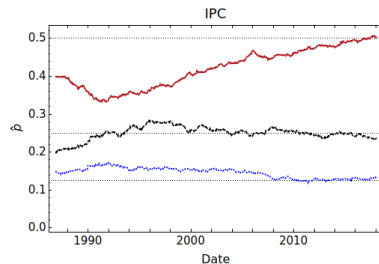
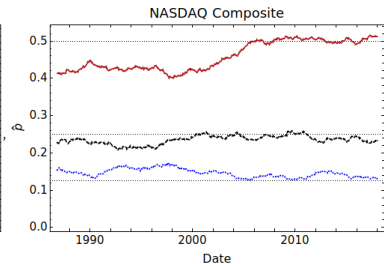
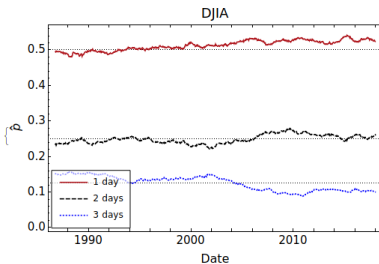
- In order to investigate the cause of the deviation from the random walk we investigated the distribution of waiting times, which are also expected to distribute as a GeometricDistribution.



# Total waiting time parameter estimation

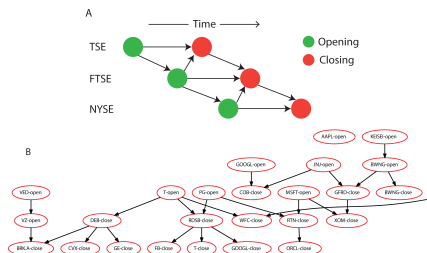


# Waiting time parameter estimation in time



# Next steps

- Right now we have observables based on rigorous statistical tests that may indicate relevant data to build a model from. No technical analysis.
- A model may be constructed using a big database and machine learning techniques.
- For example, having a bayes net model may be useful for prediction:



Thank you all for your attention.