Exploring new empirical properties of financial index time series: a data analysis of daily short term trends and their corresponding multi-scale returns

H.R Olivares-Sánchez[1]**, C.M. Rodríguez-Martínez**[3]**, E. Scalas**[2]**,H.Tapia-McClung**[3]**, A.R. Hernández-Montoya***[3]

[1]Department of Astrophysics, Radboud University, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands.
[2]Centro de Investigación en Inteligencia Artificial. Universidad Veracruzana. Sebastián Camacho 5, Xalapa Veracruz 91000, México. Tel/Fax: 52-228-8172957/8172855.
[3]Department of Mathematics, School of Mathematical and Physical Sciences, University of Sussex, Brighton, BN1 9QH, United Kingdom.
[4]Facultad de Física. Universidad Veracruzana, Apdo. Postal 475. Xalapa, Veracruz. México.

E-mail: `alhernandez@uv.mx`

Abstract. In financial time series there are time intervals in which market indices values or prices increase or decrease monotonically. We call those events "price runs" or "elemental trends". In the present paper we study the distribution of the duration of elemental trends for the daily indices DJIA, NASDAQ, IPC and Nikkei 225 along the period of time from 10/30/1978 to 26/09/2019. It is found that elemental trend distribution often differs from the one expected under no memory. The expected and observed duration distributions are compared by means of the Anderson-Darling test, making possible to assess and compare different markets efficiency and its time evolution. Additionally, we construct the returns of the elemental trends, obtaining a sample of non arbitrary multi-scale returns whose distribution displays very interesting properties, such as bi-modality for time lags bigger than one day, and an asymptotic power law whose exponent is estimated for all the studied market data samples. Finally, some applications are presented: 1) The construction of a version of volatility observable by means of multi-scale returns, which is independent of any selected and fixed time window as in the case of the usual definition of volatility. And 2) A characterization study of extreme events in the multi-scale returns constructed sample (Dragon Kings).

1. Introduction

One of the goals of financial-market analysis is to predict the future movements of prices and financial indices. In order to achieve this goal, a huge variety of methods to forecast markets behavior have been developed, ranging from complex mathematical models even to astrological pseudo-scientific techniques. An approach that has been recently growing in popularity is the statistical analysis of large sets of data, which has now become possible due to the increasing availability of computer power and high quality data sets. This approach has benefited from the contributions not only from economists, but also from many physicists and mathematicians who have applied methods and ideas of probability theory and statistical physics to finance. As a result of these efforts, a set of nontrivial statistical properties of historical data was observed

and classified as "stylized facts" [3], which are expected to provide a better insight on markets structure and behavior.

When looking at the price values of an asset on a time series chart, it is common to observe "trends" in which most of the values are greater (or smaller) than the previous ones, that can be seen as composed by smaller elemental trends, with periods in which the value increases or decreases monotonically, a popular subject within the so called *technical analysis* of financial markets. Among other things, patterns in the trend direction of financial data are believed to be indicators of changes in the market and an indicative of future prices. The effectiveness of this technical analysis approach to financial markets is disputed and put at a stake by what is known as the Efficient Market Hypothesis (EMH), which indicates that current prices reflect available information of future values. These kind of trends are the main subject of the present work, in which we describe a basic statistical model consistent with the EMH that allows us to quantify trend direction in financial time series. Before going further, it is necessary to give some definitions. In Subsections 1.1, 1.2 and 1.3 of this introduction these definitions and other useful information will be presented. In Section 2, a model for the distribution of trends durations will be developed from the efficient market hypothesis. Section 3 will explain how the data were analyzed and section 4 provides an interpretation of the analysis.

1.1. Definitions

Let $S(t)$ be the price of an asset or an index value at time t and $X(t) = \log S(t)$ its logarithm. A common quantity used to study trends in financial time series is the *log-return* of the price of an asset, $S(t)$ defined at time t as

$$r(t, \Delta t) = X(t + \Delta t) - X(t) \tag{1}$$

for a given time sampling scale Δt. If the price variation is small, the log-return is a good approximation of the return

$$R(t, \Delta t) = \frac{S(t + \Delta t) - S(t)}{S(t)}. \tag{2}$$

In this paper, we consider Δt equal to 1 day and we use the values of the indices corresponding to the close value further in the investigated markets. More details on the data set will be given in section 3.

An *elemental trend* of duration k is defined as a subseries of $k + 1$ values within the series $S(t)$ in which every value is greater (for an uptrend) or smaller or equal (for a downtrend) than the preceding one, an example of which is shown in Figure 1 for the prices of the DJIA, in a time period between October 1978 and January 1979. The duration of an elemental downward/upward trend in daily data is the number of days before the price increases/decreases, as the price varies. Trends like these, and their duration are important because THIS AND THAT AND because MORE AND MORE ...

The aim of this work is to present a statistical study of these type of short elementary trends using market closing price values from four different indices over a time sampling scale of $\Delta t = 1$ day in the period between October 30, 1978 and September 26, 2019.

1.2. The Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH), first stated by Eugene F. Fama in 1970, claims that the market quickly finds the rational price for a traded asset [15], as the current value incorporates all possible information about the price in the future. The most important consequence of this hypothesis was shown by P. Samuelson [17] and it is the fact that the best forecast for the future price of an asset is its present price.

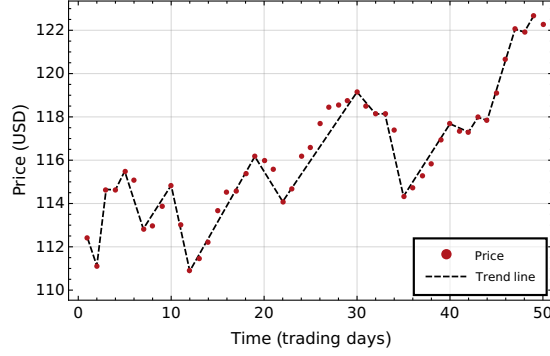$$\mathbb{E}(S(t + \Delta t)|\mathcal{F}_t) = S(t), \tag{3}$$

Figure 1: Elemental trends on the time series for the prices of the DJIA during the period from Oct/30/1978 to Jan/09/1979. The dashed line segments join the starting and ending points of each elemental trend.

where $\mathbb{E}(\cdot|\mathcal{F}_t)$ is the conditional expectation with respect to the filtration $\mathcal{F}_t$, namely with respect to the known history up to time t. Indeed, it is easy to derive the EMH from a simple no-arbitrage argument. Suppose we have two assets, a risky one, with price $S(t)$ and a risk-free one giving a constant interest rate r_F. To avoid arbitrage, one has to require that the expected return of the risky asset is equal to the risk-free interest rate, that is

$$\mathbb{E}(R(t, \Delta t)|\mathcal{F}_t) = r_F; \tag{4}$$

the latter equation immediately yields, for non vanishing $S(t)$,

$$\mathbb{E}(S(t + \Delta t)|\mathcal{F}_t) = (1 + r_F)S(t), \tag{5}$$

which reduces to (3) for $r_F = 0$. Equations (3) and (5), jointly with the integrability of the process $S(t)$, are known as martingale and sub-martingale conditions (remember that under normal conditions $r_F \geq 0$), respectively.

The EMH would invalidate the attempts of technical analysis to predict future prices or trends; in fact, in Samuelson's words, "there is no way of making an expected profit by extrapolating past changes in the futures price, by chart or any esoteric devices of magic or mathematics" [17] as the best forecast of the future price would be the current price.

1.3. Stylized facts

As mentioned before, financial time series share some nontrivial statistical properties called stylized facts. Although those properties are often formulated qualitatively, they are so constraining that it is difficult to reproduce all of them by means of a stochastic process[3]. As a matter of fact, none of the market models, including analytical models, Monte Carlo simulations and multi-agent based models, created before 1990, when awareness of such regularities gradually started to appear, could reproduce all of these stylized facts [16]. As an interesting issue, some studies suggest that stylized facts appear not only in financial time series, but also in other complex systems such as Conway's Game of Life [5]. To fix the ideas, some of the stylized facts, taken from reference [3], are listed below:

Absence of linear autocorrelations: Autocorrelations of returns are often negligible, except for very small time scales, depending on the market and on the time horizon.

Heavy tails: The return distribution is leptokurtic and some authors claim that the tails decay as a power-law.

Gain-loss asymmetry: Large downward jumps in stock prices and stock index values are observed, but not equally large upward movements. (In exchange rates there is a higher symmetry in up/down movements).

Volatility clustering: High volatility events do cluster in time.

2. An 'Efficient Market' model for the duration distribution

Among all the possible martingale or sub-martingale models that can describe price fluctuations, the geometric random walk is the simplest one. A geometric random walk is just a product of independent and identically distributed positive random variables. If the expected value of these variables is 1, then the geometric random walk is a martingale; otherwise, if the expected value is larger than 1, the geometric random walk is a submartingale. However, the geometric random walk hypothesis is neither necessary nor sufficient for an efficient market, as shown by many authors among whom Leroy [6], Lucas [7] and Lo and Mackinlay [8]. To understand this point, it is enough to consider Equation (5) allowing for any martingale model.

At each step of a series of index values, 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. Therefore, it should be impossible to predict the expected direction of a future price change given the history of the price process. In formula, from Equation (3) (after discounting for the risk-free rate), we have

$$\mathbb{E}(S(t + \Delta t) - S(t)|\mathcal{F}_t) = 0; \tag{6}$$

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

$$\mathbb{E}(Y(t, \Delta t)) = 0. \tag{7}$$

If the price follows a geometric random walk, then the series of price-change signs can be modeled as a Bernoulli process. This process could be biased to take the presence of a risk-free interest rate into account. To be more specific, let us consider a log-normal geometric random walk and let us use the assumption $\Delta t = 1$. Let S_0 be the initial price. The price at time t will be given by

$$S(t) = S_0 \prod_{i=1}^{t} Q_i \tag{8}$$

where Q_i are independent and identically distributed random variables following a log-normal distribution with parameters μ and σ. These two parameters come from the corresponding normal distribution for log-returns. As a direct consequence of the EMH in the form (5), we have

$$\mathbb{E}(Q) = 1 + r_F. \tag{9}$$

Whilst for a log-normal distributed random variable the expected value is

$$\mathbb{E}(Q) = e^{\mu} e^{\sigma^2/2}, \tag{10}$$

by combining these two equations the following dependence between the parameters is found:

$$\sigma = \sqrt{2\left(\log(1 + r_f) - \mu\right)}, \tag{11}$$

which allows us to compare the parameters from the distributions measures of the index price returns, as the values of the two parameters μ and σ come from the corresponding normal

distribution for log-returns of the price data. Further, from the cumulative distribution function of a log-normal random variable

$$F_Q(u) = \mathbb{P}(Q \le u) = \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{\log(u) - \mu}{\sqrt{2\sigma^2}}\right), \tag{12}$$

we find that the probability of a negative sign of the return is given by

$$q = F_Q(1) = \mathbb{P}(Q \le 1) = \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{-\mu}{\sqrt{4(\log(1 + r_f) - \mu)}}\right). \tag{13}$$

For typical markets, the value of the risk-free rate of returns oscillates in the range $0 < r_F \le 0.061$ as can be seen in Figure 2, where the values of r_F corresponding to three markets for the period from 1978 to 2018 are shown. Using these values for r_F, the probability of negative returns is fund to be $q = 0.5 \pm 0.05$, which is compatible with the EMH model, thus the Bernoulli process is a good approximation for the probability of a change in sign for the return data.
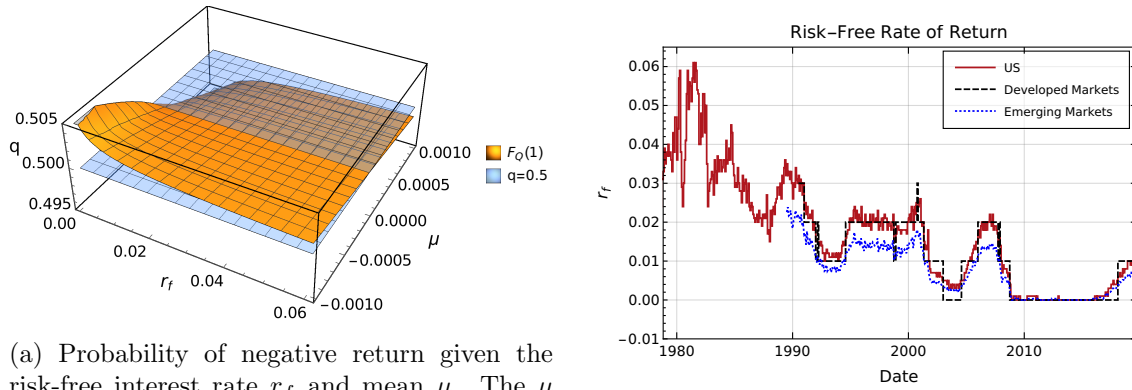
(a) Probability of negative return given the risk-free interest rate r_f and mean μ. The μ parameter can be estimated from the return time series Q_i, and r_f is estimated from a reference asset that depends on the market being studied. Intersection with $q = 0.5$ is also shown.

(b) Risk-Free rate of returns r_f for the US market, developed markets and emerging markets [10]. For the period ranging from 1978 to 2018 the risk-free rate oscillates between the interval $[0, 0.061)$.

Figure 2: For a tipical time series of Q_i the parameter μ is close to zero oscillating on a small interval ranging from -0.001 to 0.001. The risk-free rate of return r_f oscillates between the interval $[0, 0.061)$. For these values the probability of a negative return is $q = 0.5 \pm 0.05$.

Under this framework, it becomes natural to use the biased Bernoulli process as the null hypothesis for the time series of sign changes of the log-returns [9]. 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 is then given by

$$P(k) = \mathbb{P}(N = k) = p(1 - p)^k = pq^k. \tag{14}$$

The duration of a elemental downward trend in daily data is the number of days before the price increases, so the distribution of such trend durations should follow a geometric distribution. An identical argument applies to the duration of a upward trend. Such sequences of identical outcomes are also known as *runs* or *clumps* in the mathematical literature.

In the next section we describe the data used for testing the model presented, as well as a discussion of the goodness of fit test applied to compare the observed and expected distributions, namely the geometric distribution, of the duration of the trends of upwards/downwards price, which coincide with the sign of the log-returns.

3. Methodology

3.1. Data

For this work, data corresponding to closing prices of four indices were analyzed, namely Dow Jones Industrial Average (DJIA), NASDAQ Composite, the Mexican Índice de Precios y Cotizaciones (IPC) and Nikkei 225 during the period between 30 October 1978 - 26 September 2019. The data was obtained from Yahoo Finance.

3.2. Building the sample

For each series of index values, several time windows of 504 trading days (which represent two trading years in average) were created, each one shifted forward by one trading day with respect to the previous one. This procedure resulted in 4830 time windows for the DJIA, 4216 for the NASDAQ, 3901 for the IPC and 4584 for the Nikkei. For every sample in the time windows, the number of days at time t before every sign change $T_d(t)$ (defined as the trend duration) is measured. For instance, given the sequence $- - + + + + + -$, the recorded times will be $T_d = (2, 5, 1)$.

Two histograms were built for each market measuring T_d on the whole sample without time windows. One for upward trends and the other one for downward trends. Examples of both uptrend and downtrend histograms for different time windows are shown in Figure 8.

3.3. The Anderson-Darling goodness of fit test

In order to compare the observed and expected distributions of trend durations, the Anderson-Darling test described in reference [1] was used. The AD test belongs to a family of goodness of fit tests called the Cramr-von Mises tests, which includes the Anderson-Darling test, Watson's test and the Cramr-von Mises test itself. The family was originally developed to test continuous distributions, but a generalization for discrete distributions appeared for the first time in an article by Choulakian *et.al.* [23]. 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.

The principle behind this kind of tests is defining a statistic that serves to measure the distance between a theoretical distribution function $F_0(k)$ and the empirical (cumulative) distribution function for n events, $\mathbf{F}_n(k)$. Every value of the statistic is associated with a p-value, that can be interpreted as the probability of obtaining a value of the statistic at least as large as the one obtained, given that the null hypothesis

$$\mathcal{H}_0 : \mathbf{F}_n(k) = F_0 \tag{15}$$

is true. If the p-value is smaller than a previously defined threshold value α, the null hypothesis is rejected.

Two separate tests were applied. The first is a test of whether the observed data is drawn from a distribution belonging to a parametric family $\mathcal{G}(p)$. This tell us whether the price fluctuations can be modeled as a Bernoulli process. The second is a test of whether the observed data comes from a geometric distribution with $p = 0.5$. Based on the model outlined in section 2 we interpret a rejection of this null hypothesis as evidence that the market is not efficient.

For a more complete discussion on Anderson-Darling Methodology see appendix 5.

4. Discussion

In Figure 3, the ratio of the upward to total price changes in daily data is plotted against time for the years 1978 - 2019. This ratio is calculated over a time window of 504 trading days. It can be seen that variations are greater than those expected for the same time windows in a Bernoulli process with parameter $p = 1/2$ (± 0.05). Plots in figure 8 show the distribution of
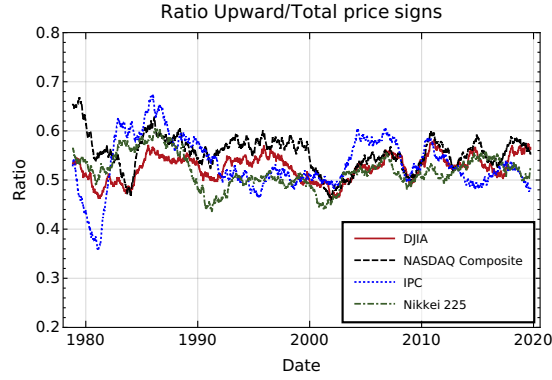
Figure 3: Ratio of upward to total price changes in daily data, plotted against time for the interval from 1978-10-30 to 2019-09-26, calculated over a time window of 504 trading days.

trend durations corresponding to different indices. Figure 6 display the p-values of the Anderson-Darling statistic for different periods. In order to avoid confusion between the parameter of the geometric distribution and the p-values for the distribution of A_n^2, the latter will be referred to as π-values. The meaning of π-values is the probability of obtaining a value of A_n^2 at least as big as the one that was really obtained, given that the probability distribution is actually geometric.
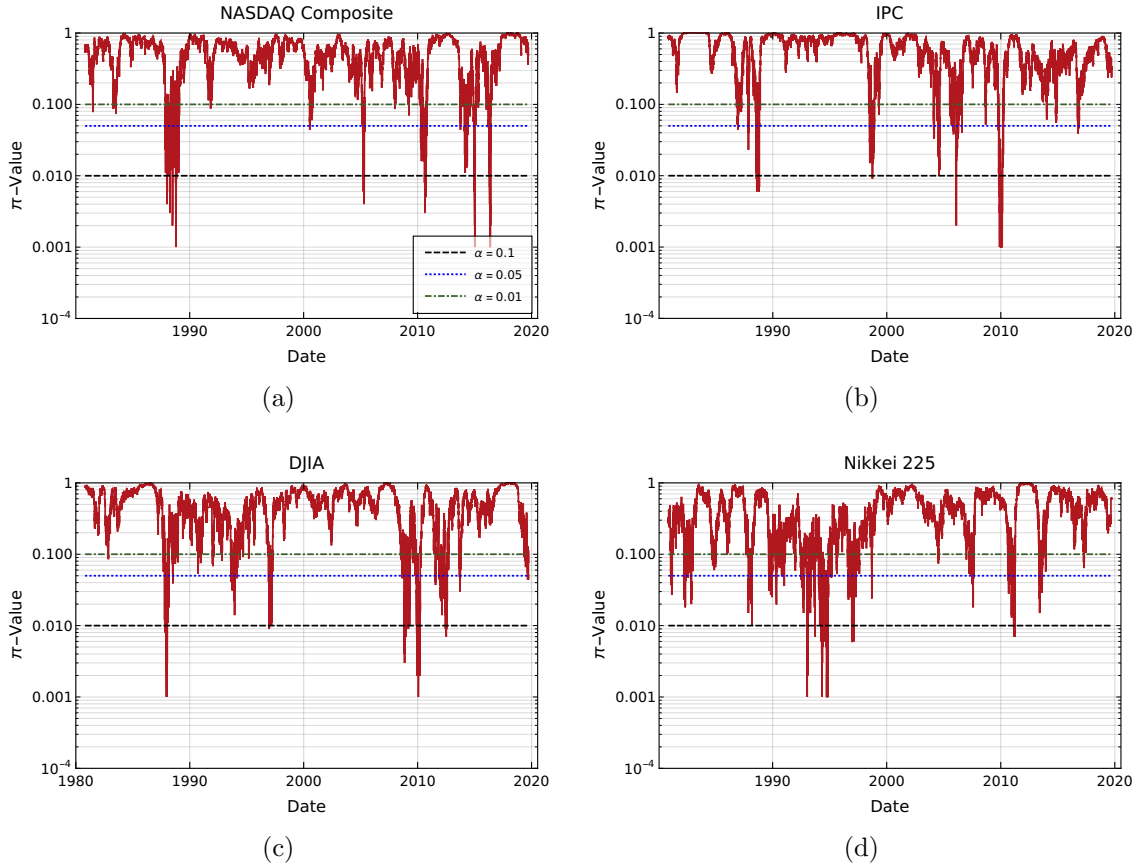
Figure 4: 6a: π-values of the parametric family Anderson-Darling statistic for the studied markets.
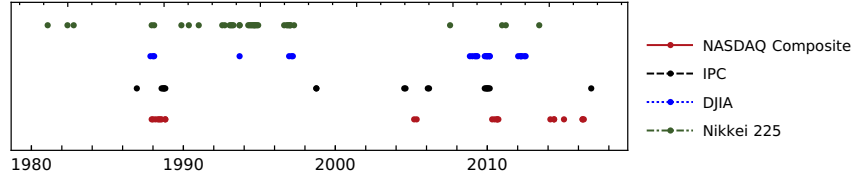
Figure 5: Dates where events from figure 4 have a π-value below the $\alpha = 0.05$ confidence level.

For the parametric family test in figure 4, it was observed that the assumption of a Bernoulli process for price directions holds reasonably well most of the time, except for sporadic deviations that are usually present around extreme market movements like a financial crisis.
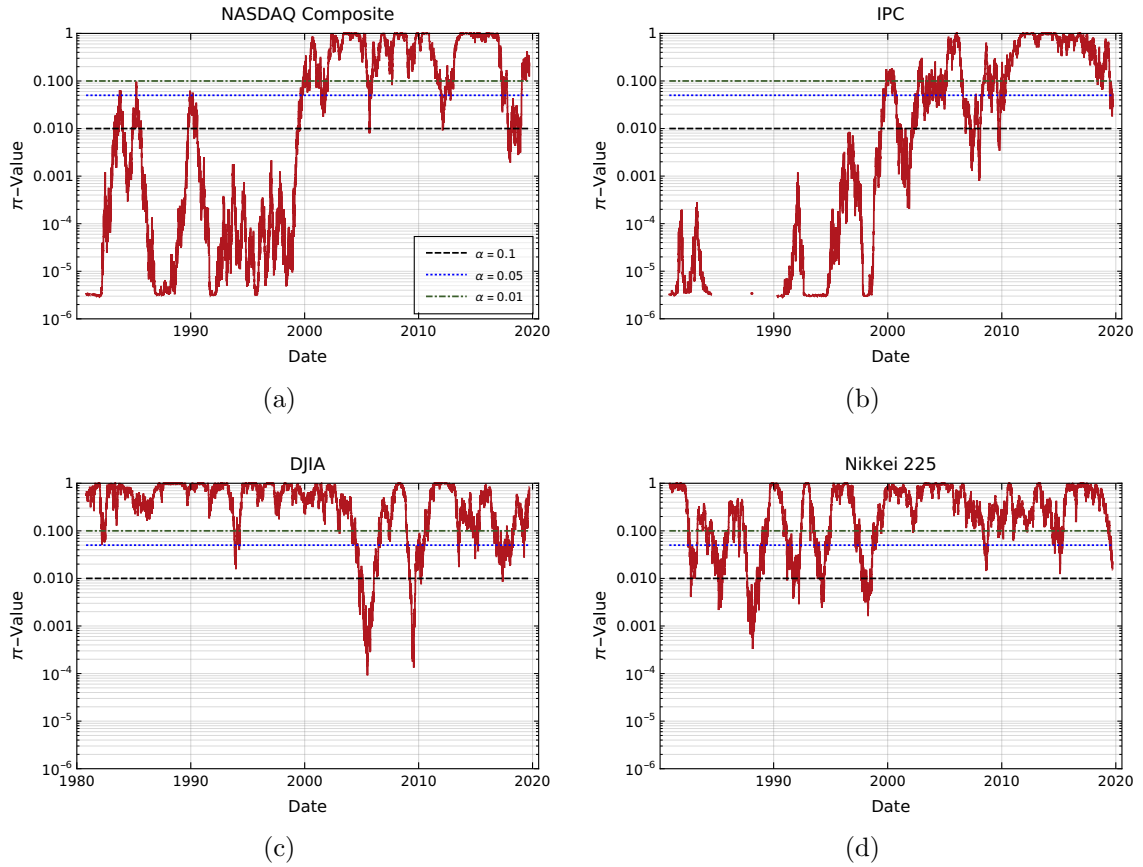

Figure 6: 6a: π-values of the Anderson-Darling statistic for the NASDAQ plotted against time. As time passes, the data agree better with the geometric distribution $\mathcal{G}(0.5)$; 6b: π-values of the Anderson-Darling statistic for the IPC plotted against time. As for NASDAQ, agreement between data and the geometric distribution increases with time; 6c: π-values of the Anderson-Darling statistic for the DJIA plotted against time. The greatest deviations from the geometric distribution occurred between the years 2000-2011. 6d: π-values of the Anderson-Darling statistic for the Nikkei plotted against time. The greatest deviations from the geometric distribution occurred between the years 1980-2000.

It was observed that as time passes, the direction of price changes for the IPC and the NASDAQ is better described by a geometric distribution following the efficient market hypothesis $\mathcal{G}(0.5)$ (Figures 6a and 6b). Meanwhile the distribution of trend durations for the Dow Jones and Nikkei
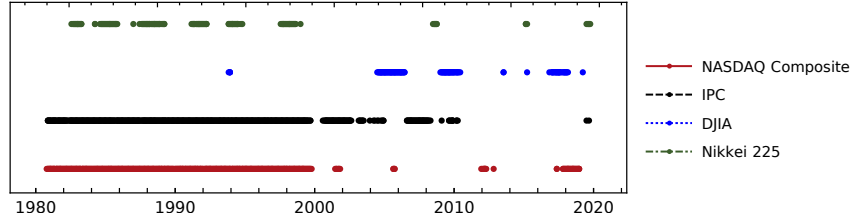
Figure 7: Dates where events from figure 6 have a π-value below the $\alpha = 0.05$ confidence level.

is generally reasonably well fitted by the geometric distribution. This fact can be interpreted as a possible evidence that the Mexican stock market (that has become public and regulated since 1975) has been increasing its efficiency, as reported by previous research [4]. The same claim can be made about the NASDAQ, given that it is also a market of relatively recent creation. In contrast, the Dow Jones Industrial Average index represents a more mature market. However, there is also evidence that the New York Stock Exchange, represented by the Dow Jones, has swiftly increased its efficiency between the beginning of the 1980s and the end of the 1990s [19]. Figure 6c shows that for the Dow Jones, the greatest deviations from the geometric distribution in the studied period occurred between the years 2000-2011. Figure 6d shows that for the Nikkei, the greatest deviations from the geometric distribution in the studied period occurred between the years 1980-2000.

5. Conclusions

The probability distributions for the duration of elemental trends were studied for the market indices Dow Jones Industrial Average (DJIA), NASDAQ Composite, the Mexican Índice de Precios y Cotizaciones (IPC) and for the Japanese Nikkei 225. These distributions are expected to be geometric and memoryless according to the discussion in section 2. The IPC and the NASDAQ present periods in which the memoryless hypothesis must be definitively rejected, while DJIA and Nikkei showed to be fitted reasonably well by geometric distribution with sporadic deviations.
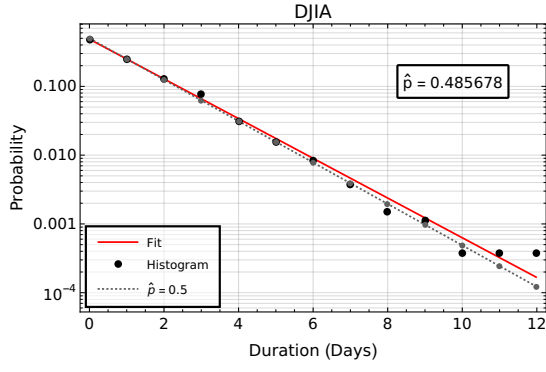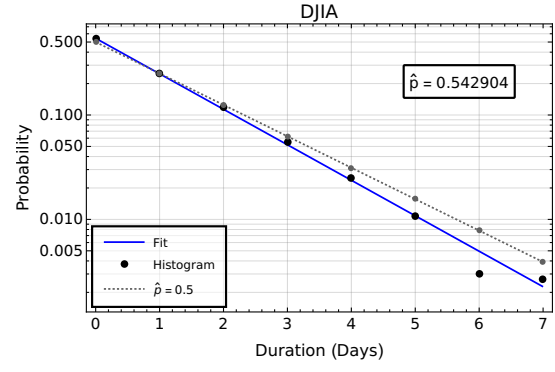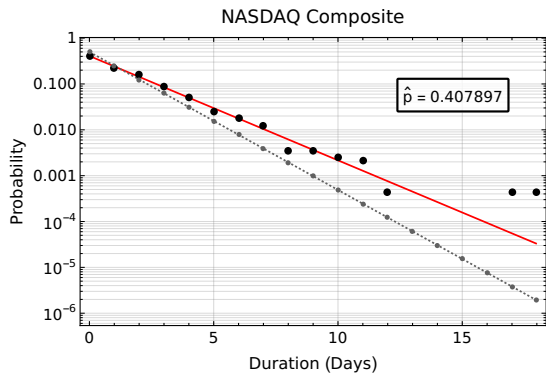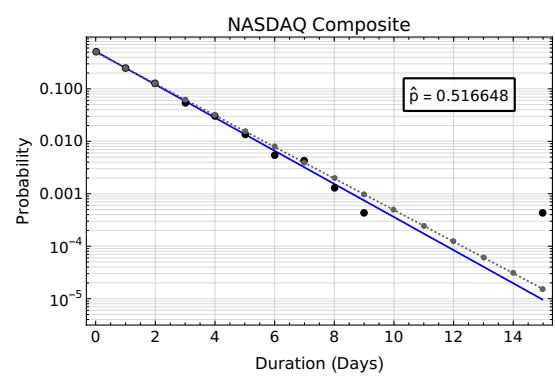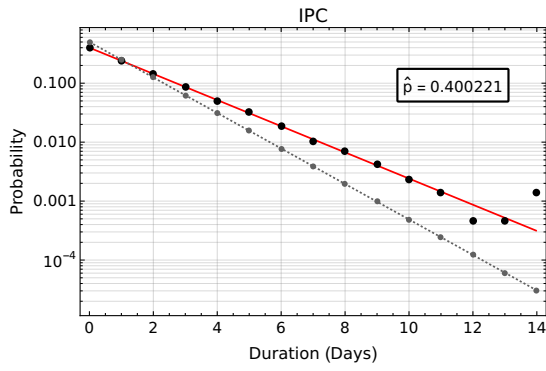
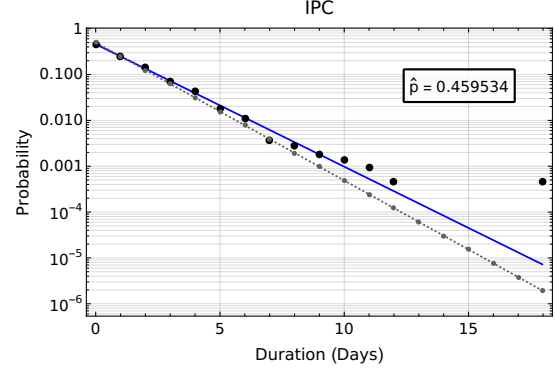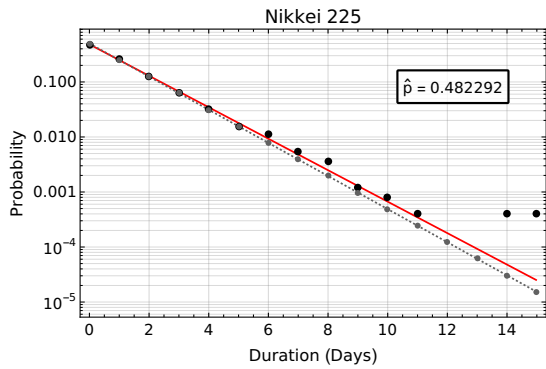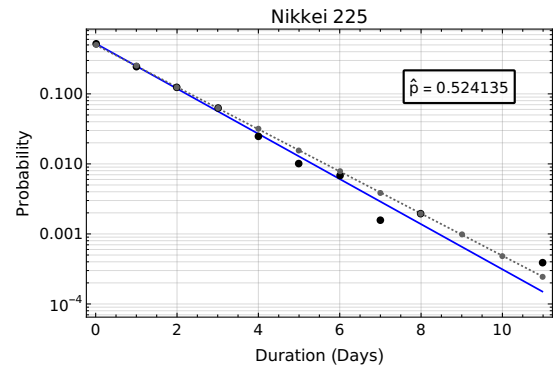Figure 8: Distribution of trend durations T_d for different indices. (a), (c), (e) and (g) present the upward distribution and (b), (d), (f) and (h) the downward distribution. The solid lines are the fitted geometric distributions while the dotted lines are the expected geometric distributions ($p = 0.5$). The parameter $\hat{p}$ is shown on each figure.

Appendices

The discrete version of the Anderson-Darling goodness of fit test
Determining whether or not a given probability model fits the observed data that it attempts to describe is one of the most important problems of applied statistics. Even though extensive research has been done regarding this problem, most of it deals with continuous distributions. Unfortunately, the studies done for fitting discrete distributions are much less numerous, or at least harder to find.

Discrete distributions, however, are important in many fields as medicine, psychology and engineering. Scientific papers and textbooks often prescribe the chi-square test as the option to use when testing goodness of fit for these distributions.

Nevertheless, it is well known that chi-square tests suffer from low power, especially when applied to data in which there are bins with very small content (as a single event), and when the expected distribution predicts a very low probability for one of the categories. Data drawn from a process that can be described by a geometric distribution precisely exhibits both features.

Bracquemond *et. al.* (2002) [22] make a review of eight alternatives to the chi-square test proposed over the past years and perform a simulation-based comparative study specifically for the geometric distribution. It consists first on checking the empirical significance level against the nominal one for each test, and then performing a power study. They analyzed three tests based on the empirical distribution function, three based on the empirical generating function, the Neyman smooth test, and a test by Nikulin (1992) [24] based on the generalized Smirnov transformation.

The tests that had an overall better performance were the Baringhaus-Henze (BH) test, the Anderson-Darling (AD) test and Nikulin's test. Among these, Nikulin's test was considered to have a satisfying power, but they recommend not to use it for small data. The two other tests have the disadvantage of requiring a numerical procedure called parametric bootstrap, that is relatively expensive computationally speaking. However, since the BH test involves by far many more operations than the AD test, we preferred to use the latter for our analysis. The computer power required to carry out the AD tests in this work is reasonable with current technology. The code we made to perform the test gives the results within about one second on an ordinary laptop.

The AD test belongs to a family of goodness of fit tests called the Cramér-von Mises tests, which includes the Anderson-Darling test, Watson's test and the Cramér-von Mises test itself.

The family was originally developed to test continuous distributions, but a generalization for discrete distributions appeared for the first time in an article by Choulakian *et.al.* [23].

The principle behind this kind of tests is defining a statistic that serves to measure the distance between a theoretical distribution function $F_0(k)$ and the empirical (cumulative) distribution function for n events, $\mathbf{F}_n(k)$. Every value of the statistic is associated with a p-value, that can be interpreted as the probability of obtaining a value of the statistic at least as large as the one obtained, given that the null hypothesis

$$\mathcal{H}_0 : \mathbf{F}_n(k) = F_0 \tag{16}$$

is true. If the p-value is smaller than a previously defined threshold value α, the null hypothesis is rejected. For the case of the discrete Anderson-Darling test, this statistic is the *Anderson-Darling statistic*:

$$A_n^2 = n \sum_{k=1}^{\infty} \frac{[\mathbf{F}_n(k) - F_0(k)]^2 p_0(k)}{F_0(k)(1 - F_0(k))}, \tag{17}$$
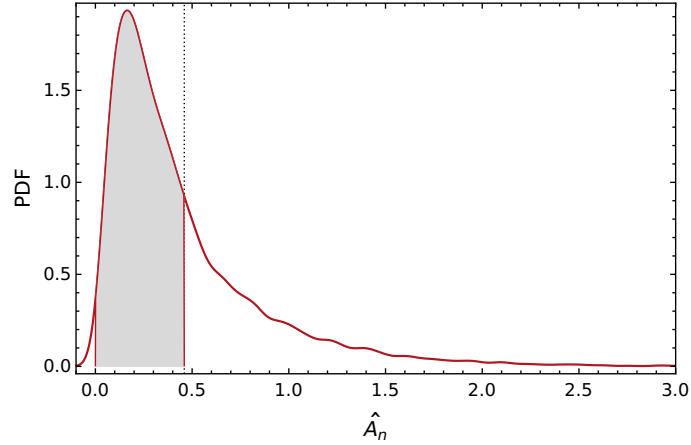
Figure 9: Distribution of $\hat{A}_n$ generated from 10000 replicas of a sample with $\hat{p} = 0.502994$. The area under the curve ranging from $[0, \hat{p}]$ is the π-value that represents the probability of measuring a value as extreme as $\hat{p}$. The measured π-value for this sample is 0.609.

where $p_0 = F_0(k) - F_0(k - 1)$. If instead what is being tested is whether the observed data comes from a distribution belonging to a parametric family $F(\cdot; \theta)$, then the parameter θ must be estimated first.

For the case of the geometric distribution $\mathcal{G}(p)$, this is an additional complication, since the distribution of A_n^2, and therefore the correspondence between it and the p-values, depends both on n and on the parameter p. In order to get around this problem, a numerical technique called parametric bootstrap is used. First, p is estimated using the maximum-likelihood estimator

$$\hat{p}_n = \frac{n}{\sum_{i=1}^n K_i}, \tag{18}$$

where K_i are the values of the random variable in the sample. Then, a large number of copies of the sample is generated and filled with random numbers taken from the actual geometric distribution $\mathcal{G}(\hat{p})$, and the AD statistic is calculated for every one of them (figure 9). The distribution of the statistic found from the samples can be then integrated up to the value of A_n^2 calculated from the empirical data, in order to find the p-value for our case of interest. For the parametric bootstrap, we used 1000 copies of the sample, which was two times the number used by Bracquemond *et. al.* [22] for their tests. As our random number generator, we used the routine *MersenneTwister* from the software *Wolfram Mathematica*, which breaks serial correlations to a great extent and has a period of $\approx 2^{19937} - 1$.

Acknowledgements

References

[1] Bracquemond C, Crétois E and Gaudoin O 2002 'A comparative study of goodness-of-fit tests for the geometric distribution and application to discrete time reliability' *Preprint* http://www-ljk.imag.fr/SMS/preprints.html.

[2] Chen M K, Lakshminarayanan V and Santos L R 2006 'How Basic Are Behavioral Biases? Evidence from Capuchin Monkey Trading Behavior' *J. Political Econ.* **144** 517–537

[3] Cont R 2001 Empirical properties of asset returns: stylized facts and statistical issues' *Quantitative Finance* **1** 223–236

[4] Coronel-Brizio H F, Hernández-Montoya A R, Huerta-Quintanilla R, and Rodríguez-Achach M E 2007 'Evidence of increment of efficiency of the Mexican Stock Market through the analysis of its variations', *Physica A* (380) 391398

[5] Hernández Montoya A R, Coronel-Brizio H F, Rodríguez-Achach M E, Stevens-Ramírez G A, Politi M, and Scalas E 2011 'Emerging properties of financial time series in the Game of Life' *Phys. Rev. E* **84** 066104

[6] Leroy S F 1973, 'Risk aversion and the martingale property of stock prices' *International Economic Review* **14** 436–446.

[7] Lucas R E 1978, 'Asset prices in an exchange economy' *Econometrica* **46** 1429–1445.

[8] Lo A W and Mackinlay A C 1999, *A Non-Random Walk Down Wall Street*, (Princeton: Princeton University Press).

[9] Scalas E 1998, 'Scaling in the market of futures', *Physica* A **253** 394–402.

[10] Mba.tuck.dartmouth.edu. (2015). Kenneth R. French - Data Library. [online] Available at: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html [Accessed 11 Oct. 2019].

[11] Hołyst J A and Sieczka P 2008 'Statistical properties of short term price trends in high frequency stock market data' *Physica* A **387** 1218–1224

[12] Instituto Nacional de Estadística, Geogrfía e Informática. Índice Nacional de Precios al Consumidor. (2011, 5 de noviembre). http://www.inegi.org.mx/est/contenidos/proyectos/inp/inpc.aspx.

[13] Jensen M H, Johansen A and Simonsen I 2002 'Inverse Statistics in Economic: The gain-loss asymmetry', *Physica* A **324** 6.

[14] Levy M, Levy H and Solomon S 2000 *Microsimulations of Financial Markets: From Investor Behavior to Market Phenomena* (London: Academic Press)

[15] Mantegna R N and Stanley H E 2000 *An Introduction to Econophysics.* (Cambridge: University Press)

[16] Samanidou E, Zschischang E, Stauffer D and Lux T 2006 'Microscopic Models of Financial Markets'. Economics Working Paper 2006-15. Department of Economics, University of Kiel.

[17] Samuelson P A 1965 Proof that Properly Anticipated Prices Fluctuate Randomly, *Industrial Management Rev.* **6** 41–45.

[18] Samuelson P A and Nordhaus W D 2005 *Economics* (New York: McGraw-Hill)

[19] Totha B and Kerteszb J 'Increasing market efficiency: Evolution of cross-correlations of stock returns' *Physica* A **320** 505–515

[20] U. S. Bureau of Labor Statistics, Consumer Price Index. [Accessed 5 Nov. 2011]. http://data.bls.gov/timeseries/CUSR0000SA0?output_view=pct_1mth.

[21] Voit J 2003 *The Statistical Mechanics of Financial Market.* (New York: Springer-Verlag)

[22] Bracquemond, C., Crétois, E., Gaudoin O., (2002) "A comparative study of goodness-of-fit tests for the geometric distribution and application to discrete-time reliability" Technical Report, URL: www-ljk.imag.fr/SMS/ftp/BraCreGau02.pdf.

[23] Choulakian, V., Lockhart, R. A., Stephens, M. A., (1994) *The Canadian Journal of Statistics*, 22, 1, pp. 125-137.

[24] Nikulin, M. S. (1992), *C. R. Math. Rep. Acad. Sci. Canada*, 14, 4, pp. 151-156.