Hurst exponent consistency for fat-tailed return distributions

Dario Bovina
Dipartimento di Fisica 'G. Galilei', Universitá di Padova - via Marzolo 8 - 35131 Padova - Italy
Tel: +39 049 8277243. E-mail: bovinad@pd.infn.it

Abstract: The generalized Hurst exponent H is often used to establish multiscaling in financial time series. In computing H(q) we are concerned with the q-order empirical moment of returns' distribution. Since power-law tails are among the most robust features observed financial data, our aim is to check if such analysis can be carried out when the returns' PDF has divergent moments. At an order q for which the theoretical moment doesn't exist, the behaviour of H(q) seriously suffers from the finiteness of the sample.

We show how Pareto tails and a finite data set can result in a spurious multiscaling even for a strictly self-affine process. A process with independent increments distributed according to a Cauchy law should have a constant exponent H; on the contrary we empirically observe bifractality. We find such behaviour is due to power-law tails only, suggesting that this spurious feature may be common to every scaling process whose increments' PDF decays according to a power law.

Dealing with independent increments we neglect another robust 'stylized fact': the volatility clustering of returns; so we wonder if our above argument still holds when returns are correlated. We take into account some of the oldest and most important Stock Market indexes and, by reshuffling the whole series, we find that Pareto tails give rise to a strong empirical multiscaling, but volatility clustering can attenuate this spurious effect.

In computing H(q) from a finite sample one should care about uncertainty. We propose a 'bootstrap' method to evaluate the reliability of H(q), i.e. its sample dependence. Using this resampling method on real financial series we find such uncertainty suddenly start to increases as the order q reaches the value q_0 for which theoretical moments are expected to diverge. The Hurst analysis is no longer reliable in the range of divergent moments. Since the q-dependence of H has been observed for orders close to q_0 or even larger, in our opinion the robustness of this feature is quite doubtful.

1 Introduction

The generalized Hurst exponent H is the most common tool used to detect multiscaling properties of the real data series. Examples come from fluid dynamics, geophysics and finance: these multiscaling features were found almost everywhere. However in spite of this enthusiasm in so strong and widespread features, the physical mechanisms able to explain multiscaling are often not well understood. Furthermore recently some doubts about the actual stability of H when it is estimated on a finite sample have been raised [3].

Also in finance the Hurst exponent has often been used to study the scaling properties of the empirical time series [12, 13, 14]. In attempting to catch the multifractal features of prices' time evolution many theoretical models were built just starting from a multifractal measure [10, 11] or a multiplicative cascade [9]. The exponent H can be helpful in classifing stock-market maturity too [4]. Hence it's a very widespred and powerful tool of investigation which can influence theoretical model's construction. Like every statistical tool, the determination of the exponent H should be followed by an estimation of its uncertainty, or at least by a search of the factors that could somehow falsify the result. This investigation becomes more and more urgent as the importance of H grows. It's worth noting that some remarks about the reliability and the sample dependence of the Hurst exponent H were already made [20, 21, 6]. However the literature still lacks a systematic study of the effects of a finite sample size, Pareto tails and volatility clustering. This paper aims at filling such a gap.

A well-established 'stylized fact' in all financial time series is the power law decay of return distributions at large absolute returns, at least within the range explored until now [8]. We wonder whether a finite sample combined with fat-tails could give rise to some spurious multiscaling in the empirical estimation of H. To this end we first consider a simple and well-known process: the Levy random walk. The return distributions of such process are strictly self-similar and the process itself is perfectly simple scaling, but the empirical estimate of H, performed on any given data sample, leads to biscaling, an extreme form of multiscaling [5]. We are going to prove that such multiscaling is merely spurious in character. Collecting returns along the path of our time series[1] gives rise to a strong finite sample effect if these returns are distributed according to a power law. This is very important as most of financial time series display power law tails in the density function of their returns. Moreover, when a single empirical time series is only available, the sliding windows method is often the main statistical tool to collect the returns, assuming a stationary process.

Then we take into account some important stock market indexes and, in order to estimate the uncertainty associated to a finite sample we use a non-parametric statistical estimator known as 'bootstrap' which belongs to the wider class of resampling methods [15, 16]. It is based on the random extraction of many samples from the whole time series [17, 23]. To our knowledge there wasn't any effort to estimate the uncertainty associated to Hurst exponent before. We find that fat tails are also responsible for a large amount of uncertainty in the empirical determination of H and this uncertainty makes the multiscaling behaviour rather doubtful.

2 The generalized Hurst exponent analysis

Consider the graph x(t) of the detrended logarithmic price of some asset or financial index. In many applications we are concerned with returns $r_\tau(t) = x(t+\tau) - x(t)$ for different time windows τ. Then, for scaling invariance to hold, we ask such returns to be equally distributed up to a scale factor depending on τ. In particular the process r_τ is self-similar if, upon varying τ, all the stochastic variables $\tau^{-\alpha} r_\tau$ have the same distribution:

$$r_\tau \sim \tau^\alpha r_1 \tag{1}$$

where $\sim$ means 'equally distributed'. It's worth stressing here that we collect returns $r_\tau(t)$ along the path $x(t)$ for every instant t, in a so-called sliding windows method, and regard them as many realizations of a single process r_τ thus assuming homogeneity in time or stationarity.

[1] the so-called sliding windows method

Going back to eq. (1) we see that PDFs p_τ of returns in a time window τ must all be equal up to a scale factor:

$$p_\tau(u) = \frac{1}{\tau^\alpha} p_1(\frac{u}{\tau^\alpha}) \tag{2}$$

and this in turn implies that the q-order moment of p_τ, provided it exists, behaves according to

$$\begin{aligned} M_\tau(q) &= \int_{\mathcal{R}} |u|^q p_\tau(u)\, du \qquad (3)\\ &= \tau^{q\alpha} M_1(q) \\ \ln M_\tau(q) &= c(q) + q\alpha \ln \tau \qquad (4) \end{aligned}$$

For a general processes, not necessarily a self-similar one, the scaling function $D(q) = qH(q)$ characterizes the dependence of the q-order moment of the returns r_τ from the time windows τ:

$$M_\tau(q) \propto \tau^{qH}$$

where H is the Hurst exponent. It can be shown that a constant $H = \frac{1}{2}$ is an asymptotical value for independent, identically distributed flows with finite variance [1]. This value of H is distinctive of normal scaling, whereas for $H \neq \frac{1}{2}$ we have anomalous scaling, as in the Fractional Brownian Motion [2]. Taking into account a strictly self-similar process for which all theoretical moments exist, from eq. (4) we find a constant Hurst exponent

$$H = \alpha$$

i.e. a linear scaling function $D(q)$.

Actually self-similarity in returns distributions is a rather strong condition. When the previous analysis is performed on most of empirical time series one finds a scaling function $D(q)$ which is not linear on the moment's order q. For the sake of precision, in many financial time series, the $D(q)$ curve is no longer a straight line, but bends downward as the order q increases. Such behaviour is generally regarded as multiscaling [18].

The exponent H is quite easy to obtain: empirically the q-order moment of a rich enough data set ($N \gg 1$) is[2]

$$M^e_\tau(q) = \frac{1}{N} \sum_{t=1}^{N} |r_\tau(t)|^q \tag{5}$$

Then we plot $\ln M^e_\tau(q)$ towards $\ln \tau$ and take the slope $D(q) = qH(q)$ of the line which corresponds to each order q. The convergence of the sum is granted as long as the theoretical moment exists. However here a problem arises because strong evidences from data analysis suggest that the tails of the return PDF decay as a power law. In this case $D(q)$ is ill-defined as it is based on theoretical moments' existence. In next section we are going to see what happens to $D(q)$ when moments don't exist.

The main problem concerning $D(q)$ is its sample dependence, even if theoretical moments exist for every order q. We need a statistical tool to evaluate the confidence interval of the outcoming function D(q) when it is calculated starting from a single finite series which contains N returns. In the empirical calculus the uncertainty which affects the sum (eq. 5) is due to the finiteness of N: for an increasing sample size N such sum approaches the theoretical value (eq. 3) better and better provided that the latter exists. If the theoretical moment does not converge, for a given order q, we empirically get a finite value anyway, but the robustness of the result should be carefully verified because the non-existence of a theoretical limit could make it very sample dependent: a few large returns can dominate the sum in eq. (5). The 'bootstrap' is a powerful statistical estimator belonging to a wider class called 'resampling methods' that is often used when a single time series is available.

Usually we estimate the uncertainty associated to a certain quantity by repeating its determination on many simulations and looking at the spreading of the set of values obtained. Unfortunately,

[2] again we stress that returns are collected along the path in a so-called 'sliding window' way

however, in finance we cannot restart the process. Instead we can extract many samples from the whole history. By repeating the determination on every sample we get a set of values for D(q) and we can evaluate the uncertainty simply by looking at the dispersion of these values around the curve D(q) obtained from the whole series. Essentially a small dispersion suggest a great robustness, on the contrary a wider range of values means that H is more sample dependent and unreliable. Thus the area covered by this bundle of curves D(q) provides a confidence interval for D(q).

To test the correctness of this estimator we proceed to apply it to a well-known process: the gaussian random walk; steps are independent increments identically distributed according to a Gauss law with zero mean and variance σ^2. i.e. with a PDF: $\mathcal{N}(0,\sigma)$. Since the sum of τ independent gaussian distributed increments[3] is again a gaussian distributed variable (stability):

$$\sum_{i=1}^{\tau} \mathcal{N}_i(0,\sigma) = \mathcal{N}(0, \sqrt{\tau}\sigma)$$

we know that $H = \frac{1}{2}$; so we expect D(q) to be a straight line with slope very close to 0.5.

We can estimate the accuracy of our method by observing the spreading of curves D(q) calculated for many subsamples. Moments always exist for gaussian distributed increments, so we are confident to find a narrow bundle of curves D(q) around the theoretical line $D(q) = \frac{1}{2}q$, even for large q. Fig. (1(a)) shows a simulation with $N = 25000$ returns; and fig. (2(a)) the theoretical line together with the curves D(q) obtained from many randomly chosen samples of size $N = 5000$. Clearly, regardless of the sample extracted, our expectation is fulfilled within a very small error.

3 Lorentzian random walk)

In the previous example the Hurst exponent analysis led correctly to simple scaling. This is primarily due to the circumstance that Brownian motion increments are distributed according to a Gauss law whose tails decay faster than exponentially. We wonder whether the Hurst exponent still leads to a correct view when large returns follow a Pareto tail. We consider again a process with independent increments, but now we draw them out from a Cauchy distribution:

$$L_\lambda(u) = \frac{\lambda}{\pi} \frac{1}{\lambda^2 + u^2} \tag{6}$$

where λ defines the distribution's width. Cauchy law belongs to the class of Levy stable distributions and has power-law tails which decay as a power law:

$$L_\lambda(u) \propto u^{-2} \quad \text{for} \quad u \gg 1$$

Let $\mathcal{L}_\lambda(i) \quad i = 1 \ldots \tau$ be random variables distributed according to eq. (6), then their sum[4] is a return r_τ and, due to stability:

$$r_\tau = \sum_{i=1}^{\tau} \mathcal{L}_\lambda(i) \sim \mathcal{L}_{\tau\lambda}$$

Hence the returns r_τ are distributed according to a Cauchy law with width parameter equal to $\lambda\tau$ (scaling):

$$\begin{aligned} p_\tau(u) &= \frac{\lambda\tau}{\pi} \frac{1}{(\lambda\tau)^2 + u^2} \\ &= \frac{1}{\tau} L_\lambda(\frac{u}{\tau}) \end{aligned}$$

This means that our process, which we call a Lorentzian Random Walk, is perfectly self-similar with scaling exponent $\alpha = 1$, but has diverging q-order moments for $q \geq 1$.

[3]that is a return r_τ

[4]again $\sim$ stands for 'equally distributed'

Since the return PDFs p_τ are all equal up to a scale factor τ, we could forecast $D(q)$ to be a straight line with slope $H = 1$, as in the case of the Gaussian Random Walk, no matter how large the order q is, provided that the sample size N is large enough:

$$D(q) = q \qquad \forall\, q \geq 0$$

Surely the above equation is true for $q < 1$, when theoretical moments exist, see eq. (3).

For $q \geq 1$ theoretical moments diverge, but for a given sample size N we empirically get a finite value $M^e_\tau(q)$ all the same, see eq. (5). Then we can regard $\ln M^e_\tau(q)$ as a stochastic variable and calculate its density function $\mathcal{P}^{(q)}_\tau$ starting from that of r_τ. It is very important to stress here that in order to get $\mathcal{P}^{(q)}_\tau$ we are not going to use a sliding windows method, i.e. we do not restrict the analysis to a single time series; instead we generate N returns for every value of the order q and of the time window τ.

The stochastic variables $|r_\tau(i)|^q$ are distributed according to[5]

$$\frac{2}{q}\,\frac{p_\tau(u^{1/q})}{u^{1-1/q}}$$

hence their characteristic function is

$$f^{(q)}_\tau(k) = \frac{2}{q}\int_0^\infty du \,\exp ik u^q p_\tau(u)$$

Using the self-similarity of p_τ one easily get[6]:

$$f^{(q)}_\tau(k) = f^{(q)}_1(\tau^{\alpha q} k)$$

Since $|r_\tau(i)|^q$ are independent, the characteristic function of their sum is the product of all characteristic functions, so the CF of $M^e_\tau(q)$ is:

$$\begin{aligned}\left[f^{(q)}_\tau(\frac{k}{N})\right]^N &= \left[f^{(q)}_1(\tau^{\alpha q}\,\frac{k}{N})\right]^N \\ &= \left[f^{(q)}_1(\tau^{\alpha q}\,\frac{k}{N})\right]^N\end{aligned}$$

This means that the PDFs of $M^e_\tau(q)$ also are self-similar:

$$M^e_\tau(q) \sim \tau^{q\alpha} M_1(q)$$

This in turn implies[7] that the PDFs $\mathcal{P}^{(q)}_\tau$ of the logarithm of q-order moments are shifted by a quantity $\alpha q \ln\tau$:

$$\mathcal{P}^{(q)}_\tau(u) = \mathcal{P}^{(q)}_1(u - q\alpha\ln\tau)$$

so

$$\langle \ln M^e_\tau(q)\rangle \;=\; \langle \ln M^e_1(q)\rangle \;+\; D(q)\,\ln\tau \tag{7}$$

where

$$D(q) = \alpha q$$

and $\alpha = 1$ in our case. The ensemble averages above $\langle\cdot\rangle$ are taken over many empirical realizations of our Lorentzian Random Walk each containing N returns.

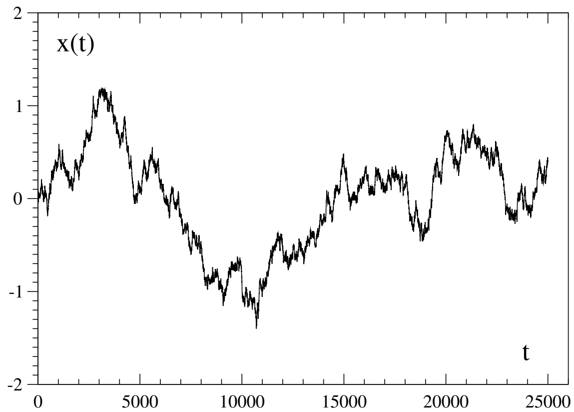
It's worth noting that we made not any hypothesis on the moments' existence. Consider a large amount N of returns. First take an order $q < 1$, thus the moments $M_\tau(q)$ theoretically exist. In

---

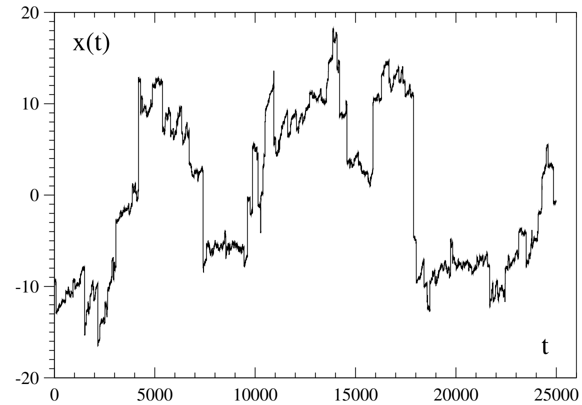
[5] if x is a stochastic variable with PDF p_x then the variable $y = x^q$ has PDF $p_y(u) = \frac{1}{q}\frac{p_x(u^{1/q})}{u^{1-1/q}}$. The factor 2 in front comes from the symmetry of $L_\lambda(u)$ for the transformation $u \rightarrow -u$

[6] in what follows we left the scaling exponent α unspecified for the sake of generality, but it should be borne in mind that $\alpha = 1$ in the present case

[7] if the stochastic variable x has distribution p_x then $y = \ln x$ is distributed according to $p_y(u) = e^u\, p(e^u)$



(a) Gaussian Random Walk.



(b) Lorentzian Random Walk.

Figure 1: random walk simulations with N=25000 independent increments. In (1(a)) the parameter $\sigma = 0.0113$ was set equal to the unconditioned standard deviation of Dow Jones returns; similarly in (1(b)) the parameter $\gamma = 0.00614$ was set to fit the width of Dow Jones returns PDF.

such a case, due to the Law of Large Numbers, $\mathcal{P}_\tau^{(q)}$ are very close to Dirac Delta functions centered in $\ln M_\tau(q)$; then eq. (7) simply become:

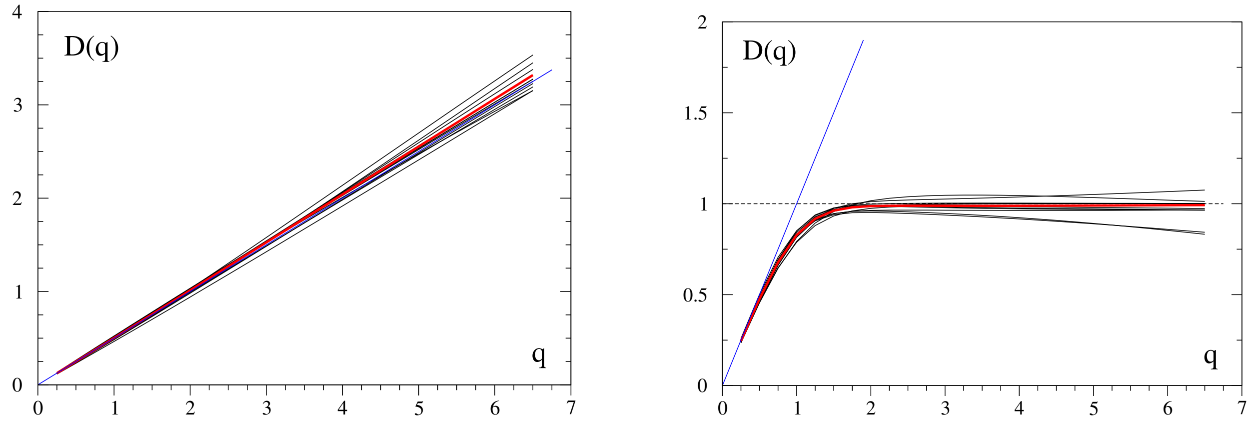
$$M_q(\tau) = M_q(1) + D(q)\ln\tau \quad \text{with} \quad D(q) = q$$

as expected. Moreover, since we are dealing with Delta functions, we expect deviations from linearity of the curve D(q) to be negligible for a sample with lots of data. Now look at an order $q \geq 1$. Whilst eq. (7) still hold we don't expect deviations to be negligible, no matter how large N is. This is why $\lim_{N\to\infty}\langle \ln M_\tau^e(q)\rangle = \infty$ and $\mathcal{P}_\tau^{(q)}$ cannot converge to a delta. Actually it cannot converge to any density function. Anyway the dependence on τ of $\mathcal{P}_\tau^{(q)}$, once N is given, still implies that the empirical curve $D(q)$ follows a straight line[8] with slope $\alpha = 1$, even if deviations may be important when $q \geq 1$.

In conclusion a self-affine process with independent increments has a linear scaling function D(q) no matter whether theoretical moments exist or not. A simulation of a lorentzian random walk is shown in fig. (1(b)); as before $N = 25000$. When the generalized Hurst exponent calculation is performed on the simulated series and on many randomly extracted samples of size $N = 5000$ we get a somehow unexpected result: within a small error all curves strongly bend downward at $q = 1$. It is very important to stress that now we use the sliding windows method to get returns from a single history, as we have to do for any financial index.

For smaller orders the slope is $H = 1$ and then, beyond $q = 1$, it vanishes and the curves D(q) become horizontal, see fig. (2(b)). Furthermore such behaviour does not depend on the choice of the sample. This is generally accepted as multiscaling, or bifractality, indeed it is so according to the previous definition based on moments. However the process has independent returns with perfectly self-similar Probability Density Functions. We just saw that in this case D(q) should follow a straight line if collecting returns along a path (the 'sliding window') would be equivalent to ensemble averaging. We also stress that the curve D(q) becomes suddenly horizontal when the order q of moments reaches unity, i.e. at the threshold for diverging moments. In conclusions we suggest that the generalized Hurst exponent analysis gives rise to spurious multiscaling even for self-similar processes, when returns are distributed according to a power law.

[8] in particular for what follows it's enough to state that D(q) strictly grows with q



(a) Gaussian Random Walk.

(b) Lorentzian Random Walk.

Figure 2: Scaling functions D(q) for several samples (N=5000, *black curves*) extracted from the whole series (*red curves*) plotted in fig. (1). The theoretical lines are drawn in *blue*.

4 Financial indexes

Since the finite size effect viewed in the last section begins at the threshold order q for which the relative theoretical moment diverge we may worry about the reliability of D(q) when it's estimated on a single financial time series. Indeed the data analysis performed on such series, like the stock indexes, clearly show fat tails.

Let's now proceed to study some important Stock Market indexes. Data are collected daily for each index starting from its birth to the present, so the availability of a large sized series led our choice towards the oldest ones, namely Dow Jones and Standard & Poor. Since we need a stationary stochastic process x(t) we start by linearly detrending the logarithm of prices S(t) of each index [19]:

$$x(t) = \ln S(t) - At - B$$

The parameter A and B came from a linear best fit interpolation of $\ln S(t)$. In the following table we show the parameters calculated for each index taken into account, see fig. (3(a)).

Index	from	until	data set size N	$A \times 10^{-4}$	B
Dow Jones Ind.	02/01/1900	14/11/2005	26589	1.90	3.52
Standard & Poor 500	03/01/1928	29/12/2006	19837	2.59	1.82
Dow Jones Trans.	02/02/1897	31/08/2005	27260	1.26	3.47
Standard & Poor (1950)	03/01/1950	18/12/2006	14333	2.82	3.06

Table 1: Stock Market indexes.

Fig. (3(b)) shows the return PDFs for a time window τ ranging from 1 to 32, rescaled according to eq. (2) with $\alpha = 0.52$. The very good collapse of all distributions warrants for their self-affinity [22], at least inside this range $1 \leq \tau \leq 32$. It's worth to stress that scaling properties hold for small time windows only because for τ so large that correlations become negligible the process is fairly independent with finite variance and, due to Central Limit Theorem, the distribution of r_τ converges to a Gauss law.

Furthermore, even if tails in fig. (3(b)) are worse sampled than the central bulk, we clearly observe the presence of a power law decay:

$$p_\tau(u) \propto u^{-\gamma} \qquad \text{for } u \gg 1$$

empirical estimations suggest a value

$$\gamma \approx 4$$

so the existence of q-order theoretical moments is only warranted for an order q strictly less than $q_0 = \gamma - 1$. We remark that the above estimation is affected by great uncertainty, nevertheless the finiteness of the variance is a robust stylized fact.

Fig. (4) shows the behaviour of D(q) for our indexes. The generalized Hurst analysis was carried out on the whole series (*red thick line*) and on many uniformly extracted samples, giving a bundle of curves for each index (*black lines*). In order to test the robustness of the curve D(q) we used a bootstrap method. When dealing with correlated series the subsample extracted should not overlap too much otherwise they cannot regarded as distincts. Their separation should be greater than the decorrelation time, i.e. the time window needed to consider two returns as independents. However the reliability of the bootstrap estimator is as greater as the extracted samples are longer and more numerous; so the size of samples we drawed from the whole series was set to $N = 5000$ and in the extraction procedure very overlapping samples were avoided.

The main features which stand out are the following.

- A downward bending of all curves, or rather a downward deviation from the straight line $D(q) = \frac{1}{2}q$: this is generally accepted as multiscaling.
- The spreading of D(q) curves is very different from those of previous examples, i.e. Gaussian and Lorentzian Random Walk. Here we distinguish quite clearly two regions: before a certain order q_* the curves are close to one another, but beyond q_* much of them start to diverge. The behaviour of D(q) become very erratic beyond q_*: it changes very much when the sample is changed. One can approximately value

$$q_* \sim 3$$

- A kind of clustering among D(q) curves. Their spreading doesn't span a continous area, it tends to occupy discrete sites. D(q) stay the same for some extracted samples and then changes abruptly.

The approximative value of q_* agrees with $q_0 = \gamma - 1$ the order at which one expect the theoretical moment to diverge. There is no any reason for such a coincidence: we believe it's a clue suggesting that the non-existence of a limit for the sum in eq. (5) can give rise to a strong sample dependence in the determination of the Hurst exponent. The uncertainty associated to $D(q)$ beyond the threshold order q_0 is so high that any effort to state its form precisely is useless. Furthermore we can observe many sample to behave according biscaling, like the Lorentzian Random Walk in section (2). The strong bending which make $D(q)$ horizontal just begins when $q = q_0$, i.e. when theoretical moments diverge. Hence, remembering the spurious effect which affects the Lorentzian random walk, see section (3), we may ask whether this multiscaling is real or merely an illusion.

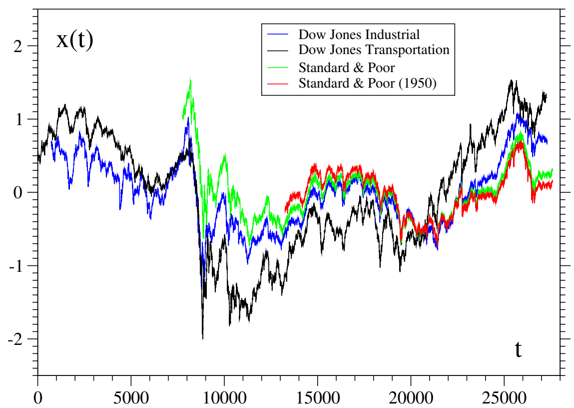
5 Volatility clustering effect

The normalized autocorrelation A(s) among absolute values of returns r_1, or volatility clustering, is a decreasing function of the time gap s between two returns:

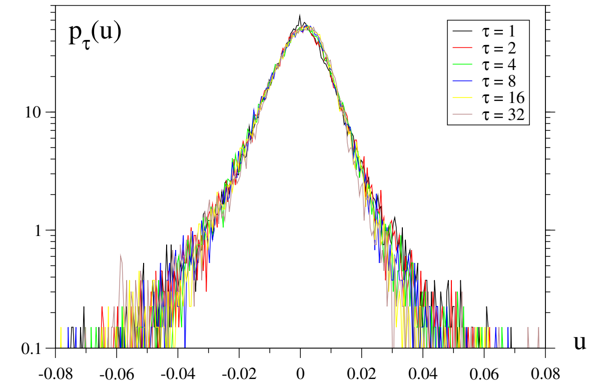
$$A(s) = \frac{\langle r_1(t+s)\, r_1(t)\rangle \; - \; \langle r_1(t+s)\rangle\langle r_1(t)\rangle}{\sigma^2}$$

where the mean $\langle\cdot\rangle$ is a sum over t, for $1 \leq t \leq N - s$. Clearly $A(0) = 1$, then A(s) usually decays as a power law until it vanishes around $s \approx 2000$, the decorrelation time.

In section (3) we found a very strong spurious multiscaling due to Pareto tails in a process with independent increments, beyond the threshold order q_0 of diverging moments; in the last section (4) we found a great erraticity in the scaling function D(q), again beyond the threshold order $q = q_0$. Now we proceed in taking into account correlations among the returns. To understand the influence of the volatility clustering we would need a correlated process which display fat tails in return PDFs



(a)



(b)

Figure 3: (3(a)) histories of some Stock Market indexes. Data have been collected daily, at closure. (3(b)) Dow Jones Ind. return PDFs for several time window $\tau = 2^i \quad i = 0 \ldots 5$. Each density function was rescaled by a factor τ^H with H=0.52

and which is analitically soluble in order to state self-similarity exactly. Otherwise we may consider our real financial indexes. Of course we cannot state self-similarity here nor solve analitically the problem, but we can destroy any correlation by reshuffling data randomly. A permutation on the set of returns r_1 available doesn't change the empirical sums (eq. 5), neither their distribution; but surely it will make volatility clustering negligible. Then a comparison between the curve D(q) obtained by real data and by the same reshuffled data should provided us with an understanding of volatility clustering effects.

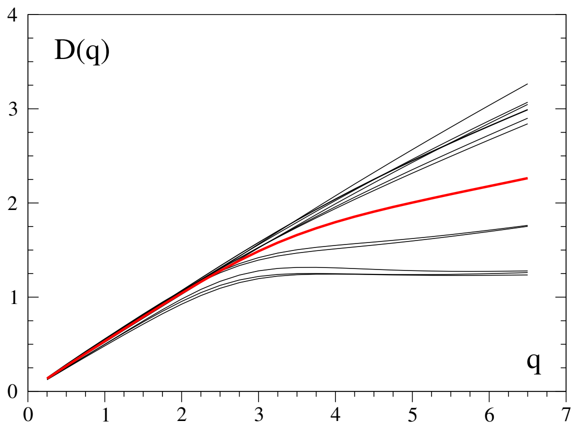
We randomly reshuffled the whole series of the Dow Jones Industrial daily returns and then performed the same analysis as in the case of Gaussian and Lorentzian random walk. It's important to remember that for the PDF of returns variance σ^2 is very strongly suggested to exist, see section (4), so returns $r_\tau \to \mathcal{N}(0, \sigma\sqrt{\tau})$ as $\tau \to \infty$. Correlations avoid such convergence until the time window τ become so large to make them negligible. If the correlations are destroyed one may worry about the existence of scaling properties themselves, i.e. the validity of eq. (2) which implies the same functional form for return PDFs regardless of the time window τ. Fortunately it seems that sums of variables distributed according to a law whith long tails converge to the Gaussian very slowly [7], so we can approximatively regard scaling as still holding in the range $1 \leq \tau \leq 32$.

In fig. (5) we show the result obtained. The bundle of curves D(q) should be compared with the bundle from the original series (fig. 4(a)). Here we get another somehow singular result: volatility clustering attenuates the multiscaling behaviour of our indexes; whilst correlation among returns was often found to be a cause of multiscaling. When there is not any dependence between returns D(q) behaves as if it's affected by the same spurious effect seen in section (3). In our opinion this is another fact supporting that the multiscaling observed in the financial indexes considered is fictious.

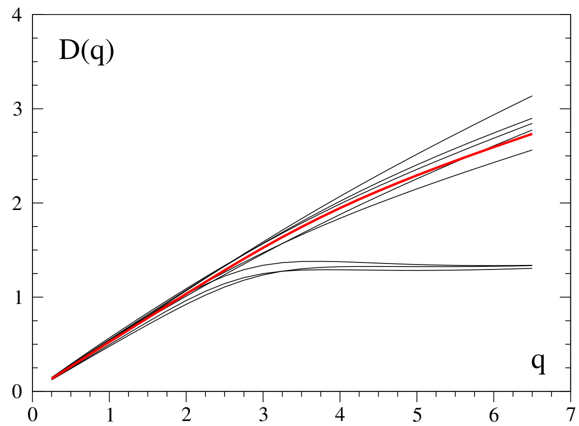
6 Conclusions

First we studied a well-known analytically soluble process: the Lorentzian Random Walk, a process belonging to the Levy Flights. In spite of its simplicity we found a rather singular result. Whilst this process is perfectly self-similar and the scaling function should be linear with slope 1, collecting increments using a sliding windows method from a single history leads to a multiscaling behaviour. We prove such spurious multiscaling to be due to Pareto tails as it begins when moments ceases to exist.

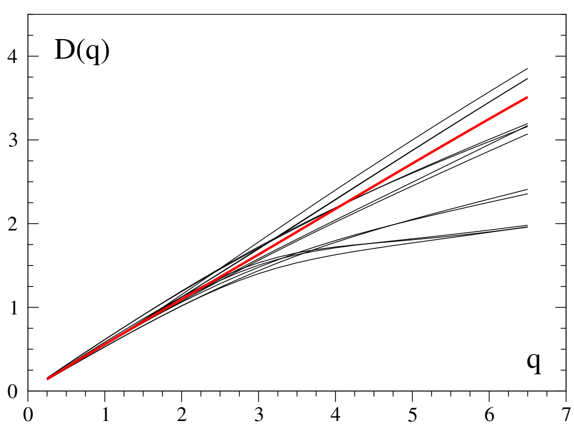
As a further step we proceed in estimating the uncertainty associated to the determination of Hurst exponent for some of the most famous financial indexes. To this end we used a non-parametric



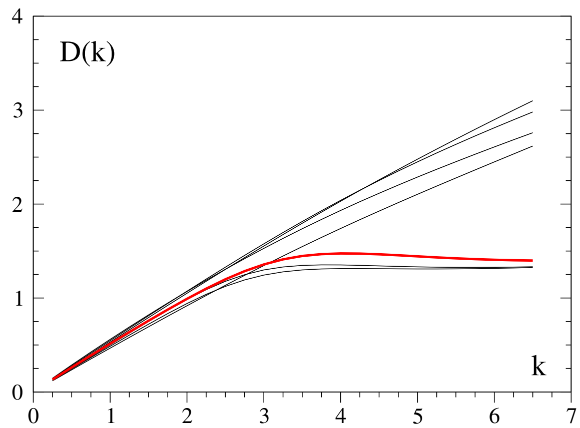
(a) Dow Jones Industrial.



(b) Standard & Poor.



(c) Dow Jones Transportation.



(d) Standard & Poor (1950).

Figure 4: Scaling functions D(q) for several samples (N=5000, *black curves*) extracted from the whole series (*red curves*) plotted in fig. (3(a)).

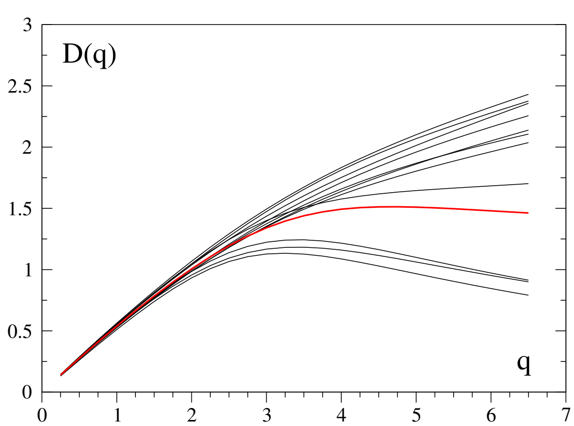


Figure 5: Scaling functions D(q) for several samples (N=5000, *black curves*) extracted from the reshuffled Dow Jones Industrial index. The curve for the whole reshuffled series is also shown (*red curves*)

bootstrap method. The main empirical results are shown in fig. (4). Clearly the scaling function D(q) becomes very erratic and sample dependent beyond the threshold q_0 of divergent moments. Actually the uncertainty is so high for $q > q_0$ to make $D(q)$ completely unreliable in this range.

The last step took into account correlations. We saw in section (4) that a positive correlation among increments can raise the curve D(q) towards linearity, compare fig. (4(a)) and (5). When volatility clustering vanishes, after a data reshuffling, we get a multiscaling behaviour of the curve $D(q)$ related to Dow Jones Industrial index which is very similar to the one related to the Lorentzian random walk. However we saw the latter multiscaling features to be fictious. Since ultimately this spurious features are due to the sliding windows method and to Pareto tails; and since both are also present in the analysis performed on all our financial indexes; we cannot exclude that the multiscaling properties of these financial indexes are spurious too.

In conclusion there are many evidences that the multiscaling behaviour observed in the financial indexes considered is merely a finite sample size effect due to the sliding windows method of collecting returns and to the power law tails.

7 Bibliografia

References

[1] W. Feller (1951). 'The asymptotic distribution of the range of sums of independent random variables'. *Ann. Mat. Statist*, **22**, 427–432

[2] B.B. Mandelbrot and J.W. Van Ness (1968). 'Fractional brownian motions, fractional noises and applications'. *Soc. Ind. Appl. Math. Rev.*, **10**, 422–437

[3] M. Eneva (1994). 'Monofractal of multifractal: a case study of spatial distribution of mining induced seismic activity'. *Nonlinear Processes in Geophysics*, **1**, 182–190

[4] T. Aste, M.M. Dacorogna and T. Di Matteo (2005). 'Long-term memories of developed and emerging markets: using the scaling analysis to characterize their stage of development'. *Journal of Banking & Finance*, **29**, 827–851

[5] A.V. Chechkin and V.Y. Gonchar (2000). 'Self and spurious multi-affinity of ordinary Levy motion, and pseudo-Gaussian relations'. *Chaos Solitons & Fractals*, **11**, 2379–2390

[6] B. LeBaron (2001). 'Stochastic volatility as a simple generator of apparent financial power laws and long memory'. *Quantitative Finance*, **1**, 621–631

[7] R.N. Mantegna and H.E. Stanley (1994). 'Stochastic Process with Ultraslow Convergence to a Gaussian: The Truncated Levy Flight'. *Physical Review Letters*, **23**, 2946–2949

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

[9] J.F. Muzy, D. Sornette, J. Delour and A. Arneodo (2001). 'Multifractal returns and Hierarchical Portfolio Theory'. *Quantitative Finance*, **1**, 131–148

[10] B.B. Mandelbrot, A. Fisher and L. Calvet (1997). 'A Multifractal Model of Asset Returns'. Working Paper

[11] B.B. Mandelbrot (2001). 'Scaling in financial prices: IV. Multifractal concentration'. *Quantitative Finance*, **1**, 641–649

[12] J. Fillol (2003). 'Multifractality: Theory and Evidence an Application to the French Stock Market'. *Economics Bulletin*, **31**, 1–12

[13] F. Schmitt, D. Schertzer and S. Lovejoy (2000). 'Multifractal fluctuations in finance'. *International Journal of Theoretical and Applied Finance*, **3**, 361–364

[14] K. Matia, Y. Ashkenazy and H.E. Stanley (2003). 'Multifractal properties of price fluctuations of stocks and commodities'. *Europhysics Letters*, **61**, 422–428

[15] B. Efron (1979). 'Bootstrap Methods: Another Look at the Jackknife'. *The Annals of Statistics*, **7**, 1–26

[16] B. Efron and R. Tibshirani (1986). 'Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy'. *Statistical Science*, **1**, 54–75

[17] H.R. Kunsch (1989). 'The Jackknife and the Bootstrap for General Stationary Observations'. *The Annals of Statistics*, **17**, 1217–1241

[18] L. Calvet, A. Fisher and B. Mandelbrot (1997). 'Large Deviations and the Distribution of Price Changes'. *Cowles Foundation*, 1165

[19] J.W. Kantelhardt (2008). 'Fractal and Multifractal Time Series'. *ArXiv: physics.data-an* 0804.0747v1

[20] J.P. Bouchaud, M. Potters and M. Meyer (2000). 'Apparent multifractality in financial time series'. *The European Physical Journal B*, **13**, 595–599

[21] Z.Q. Jiang and W.X. Zhou (2008). 'Multifractality in stock indexes: Fact or Fiction?'. *Physica A*, **387**, 3605–3614

[22] C.J.G. Evertsz (1995). 'Fractal Geometry of Financial Time Series'. *Fractals*, **3**, 609–616

[23] E. Carlstein, K.A. Do, P. Hall, T. Hesterberg and H.R. Kunsch (1998). 'Matched-block bootstrap for dependent data'. *Bernoulli*, **4**, 305–328