Characteristic function approach to the sum of stochastic variables

Annibal Figueiredo and Iram Gleria and Raul Matsushita and Sergio Da Silva

Federal University of Santa Catarina

2006

Characteristic function approach to the sum of stochastic variables

Annibal Figueiredo[a], Iram Gleria[b*], Raul Matsushita[c], Sergio Da Silva[d]

[a]*Department of Physics, University of Brasilia, 70910-900 Brasilia DF, Brazil*
[b]*Department of Physics, Federal University of Alagoas, 57072-970 Maceio AL, Brazil*
[c]*Department of Statistics, University of Brasilia, 70910-900 Brasilia DF, Brazil*
[d]*Department of Economics, Federal University of Santa Catarina, 88049-970 Florianopolis SC, Brazil*

Abstract

This paper puts forward a technique based on the characteristic function to tackle the problem of the sum of stochastic variables. We consider independent processes whose reduced variables are identically distributed, including those that violate the conditions for the central limit theorem to hold. We also consider processes that are correlated and analyze the role of nonlinear autocorrelations in their convergence to a Gaussian. We demonstrate that nonidentity in independent processes is related to autocorrelations in nonindependent processes. We exemplify our approach with data from foreign exchange rates.

Keywords: Central limit theorem; characteristic function; reduced variables; autocorrelation

1 Introduction

The problem of the limit of sums of random variables attracted great interest in the second half of the 19th century and the first one of the 20th century. At the time mathematicians extending Bernoulli's and Moivre-Laplace's theorems were also pioneering the modern theory of the sum of random variables.

Hot topics in the research agenda included searching for general conditions under which the sum of random variables converges to a Gaussian. Liapunoff, Lindberg, and Levy are among those who contributed to clarify the problem.

Liapunoff suggested his statistical moments approach (Liapunoff, 1900); Lindberg came up with his convolution technique (Lindberg, 1922), and Levy focused on the classic approach to the characteristic function (CF) (Levy, 1924, 1929).

Levy not only analyzed the convergence of sums of random variables but also extended the classic approach to consider infinite first and second moments. He also examined the role of stable distributions in characterizing the limits of sums of random variables. In particular, he put forward 'extraordinary laws' to show how his stable distributions (today's Levy-stable distributions) play a role similar to that of the Gaussian when the second moment is infinite.

* Corresponding author.
E-mail address: iram@df.ufal.br or iram@pesquisador.cnpq.br .

Levy's central limit theorem settles the conditions under which sums of random variables converge to a Levy-stable distribution. Here the Gaussian collapses to a special case of the entire family of Levy-stable distributions.

Roughly Levy's approach can be seen as an application of Kolmogorov's triangular scheme (Gnedenko and Kolmogorov, 1954), which encompasses previous results in terms of limit theorems. In what follows we will discuss the triangular scheme in greater detail since it helps to put our results in this paper into perspective.

A triangular arrangement of random variables takes the form

$$\mathbf{X}^{(n)} = (x_1^{(n)}, x_2^{(n)}, ..., x_n^{(n)}); \quad n = 1, 2, ..., \infty \tag{1}$$

where the random variables $x_i^{(n)}$ ($i = 1, \ldots, n$) are defined in the same probability space so as to satisfy two properties. (1) Random variables $x_1^{(n)}, \ldots, x_n^{(n)}$ are independent for every n, and (2) for all $\varepsilon \in R$, it holds true that $\lim_{n \to \infty} \sup_{1 \le i \le n} \left\{ \left| 1 - f_i^{(n)}(\varepsilon) \right| \right\} = 0$, where $f_i^{(n)}$ is the CF of $x_i^{(n)}$. These are the conditions of independence and infinitesimality respectively.

We consider sum sequence

$$S_n = x_1^{(n)} + ... + x_n^{(n)} + a_n \tag{2}$$

that is related to the n^{th} row of the triangular arrangement in Eq. (1) (the a_n is a constant of fine tuning).

Our problem is to examine the limit of the distribution of S_n assuming the probability distributions of variables $x_i^{(n)}$ ($i = 1, \ldots, n$) to be known. If the limit of the distribution of S_n does exist, we are able to tackle the problem of its slow convergence. Here our two main tasks are to characterize the class of each possible limit of sums in Eq. (2), and to devise convergence criteria for sum S_n as a function of the properties of the distributions of $x_i^{(n)}$ ($i = 1, \ldots, n$).

Kolmogorov hypothesized that the class of possible limits of S_n matches the class of Finetti's (Finetti, 1929) infinitely divisible distributions. Kolmogorov's student Bavly (Bavly, 1936) confirmed the hypothesis to the case where second moments of $x_i^{(n)}$ ($i = 1, \ldots, n$) are finite. And a broader proof was presented soon afterwards (Khintchine, 1938). (A comprehensive discussion can be found elsewhere (Zolotarev, 1990)). And Gnedenko and Kolmogorov (Gnedenko and Kolmogorov, 1954), and Levy (Levy, 1937) are primers on the classic theory of the sum of random variables.)

Another key question is to address the conditions for sum S_n to converge to a Levy-stable distribution under the independence and infinitesimality conditions (Feller, 1935), (Feller, 1937). Lindberg-Feller condition is currently known as the central limit theorem. Yet this condition is valid only for sums satisfying independence and infinitesimality.

By the end of the 1940s, Kolmogorov's problem vanished. Subsequent work was devoted to refine the characterization of convergence of sums under independence and infinitesimality. Meanwhile Levy alone tackled the convergence problem in the absence of independence and infinitesimality. He showed convergence of S_n to a Gaussian without relying on infinitesimality (Levy, 1937). Unfortunately this finding made negligible impact on literature.

From the mid-1960s to the 1970s and 1980s, Russian mathematician Zolotarev made a significant breakthrough (Zolotarev, 1965, 1997). He examined convergence of the sum of random variables without relying on the classic assumptions. Zolotarev dubbed 'non-classic' the limit theorems lacking the independence and infinitesimality conditions. His work gave rise to a broader theory of the existence of limits in sums of random variables. The infinitesimality condition was left out and asymptotic independence was assumed (Zolotarev, 1990). One key result is his Theorem 5 (at page 131) showing necessary and sufficient conditions for convergence of sums of random variables.

Our own work in this paper elaborates further on such a non-classic framework. Yet our interest goes beyond theory in that we also devise applications to the statistical analysis of actual time series.

Our previous work is motivated by Mantegna and Stanley's (Mantegna and Stanley, 1994, 1995). We show the general approach of the CF to the sum of random variables to be useful to tracking convergence in time series of financial returns (Figueiredo *et al*., 2003). And also how correlations are key in curbing convergence to the Gaussian (Figueiredo *et al*., 2004). We put forward that Mantegna and Stanley's truncated Levy flights can be explained by linear and nonlinear autocorrelation in data. Moreover we show how departures from infinitesimality are related to financial volatility in that they can explain lack of convergence to the Gaussian (Figueiredo *et al*., 2005).

In particular, our aim in this paper is to develop a general technique to approaching convergence without relying on independence and infinitesimality. Our technique belongs to the class of non-classic methods that are based on the analysis of empirical CFs (Feuerverger, 1977, 1981). An important role is played by a function $\omega(z)$ univocally associated with a reduced distribution, which is the canonical form of Levy's CF (Levy, 1924). We confine ourselves to the analysis of processes with finite second moment, thereby rendering it useful for applications.

This paper departs from Zolotarev's in two ways. (1) We propose theorems without relying on infinitesimality (like him); the latter is replaced with volatility change in the sum. (2) We put forward general results that alow one to gauge the effect of autocorrelations in convergence without imposing any specific functional form to the CF. Indeed we are able to reckon the part of function $\omega(z)$ that is exclusively related to the autocorrelations.

Employing the canonical form of the CF allows one to get fruitful results based solely on classic theory. And using function $\omega(z)$ enables one to devise statistical gauges of the distance of a given distribution to the Gaussian. What is more, the convergence rate of an actual process can be measured and then compared to that of an independent and identically distributed (IID) one.

2 Existence of limits in the sum of random variables

Let us consider that each variable in the triangular arrangement (1) has finite standard deviation and mean given respectively by $m_i^{(n)}$ and $\mu_i^{(n)}$. Without loss of generality we can consider these variables to be sorted such that $m_i^{(n)} \geq m_j^{(n)}$ for $i > j$ and $i, j = 1, \ldots, n$. We define the reduced variable as $\overline{x}_i^{(n)} = \dfrac{x_i^{(n)} - \mu_i^{(n)}}{m_i^{(n)}}$ and can then rewrite the sum in Eq. (2) as

$$S_n - a_n = \sum_{i=1}^{n} m_i^{(n)} \overline{x}_i^{(n)} \tag{3}$$

where a_n now stands for the mean of S_n.

One of the most important issues concerning the sequences of sums given by Eq. (3) is related to convergence of the probability distribution of S_n as $n \to \infty$. To answer this question we apply the CF method as developed by Levy in his 1924 seminal work (Levy, 1924). The method consists in calculating the CF of the reduced variable

$$\overline{S}_n = \frac{S_n - a_n}{M_n} \tag{4}$$

where M_n is standard deviation of S_n. Assuming variables $x_i^{(n)}$ to be statistically independent one has

$$M_n^2 = \sum_{i=1}^{n} (m_i^{(n)})^2 \tag{5}$$

For a reduced random real variable (i.e. one with zero mean and unit variance), its CF is (Levy, 1922)

$$\psi(z) = \langle e^{Iz} \rangle = e^{-\frac{z^2}{2}(1+\omega(z))}, \omega(z) = \omega_R(z) + I\omega_I(z), \omega(0) = 0$$

Denoting $\psi(z) = \psi_R(z) + I\psi_I(z)$ yields

$$\omega_R = \frac{-2\ln|\psi| - z^2}{z^2}, \omega_I = \frac{2}{z^2}\arctan\left(-\frac{\psi_I}{|\psi|}, \frac{\psi_R}{|\psi|}\right)$$

For Gaussian distributions one has $\omega(z) = 0$ for all $z \in R$. Function $\omega(z)$ uniquely determines the distribution function of a given reduced variable. The CF of reduced variable $\overline{x}_i^{(n)}$ is

$$\psi_i^{(n)}(z) = e^{-\frac{z^2}{2}\left(1+\omega_i^{(n)}(z)\right)} \tag{6}$$

and then the CF of $\overline{S}_n$ is (Levy, 1924)

$$\Psi_n(z) = e^{-\frac{z^2}{2}(1+\Omega_n(z))} \tag{7}$$

where

$$\Omega_n(z) = \sum_{j=1}^{n} \left(\frac{m_j^{(n)}}{M_n} \right) \omega_j^{(n)} \left(\frac{m_j^{(n)}}{M_n} z \right) \tag{8}$$

In the framework of the classic central limit theorem one condition for the CF in (7) to converge to a Gaussian ($\lim_{n\to\infty} \Omega_n(z) = 0$) is precisely the infinitesimality hypothesis (referred to in Section 1). It states that

$$\lim_{n\to\infty} \frac{\mu_n}{M_n} = 0, \ \mu_n = \max_{1\le i\le n} \{ m_i^{(n)} \}$$

Yet infinitesimality does not hold for $m_i^{(n)}$. As a consequence, if we wish to track asymptotical behaviour, we need to develop tools capable of evaluating limit distributions for sums S_n without relying on infinitesimality.

As observed, Zolotarev and co-workers (Zolotarev, 1990) established the conditions for convergence of sum S_n in a non-classic limit theorem. Inspired by the works of Zolotarev, and using Levy's CF formalism, we state a theorem as follows.

Theorem 1. Consider the sum in Eq. (2) and assume

(1) $\lambda_i = \lim_{n\to\infty} \frac{m_i^{(n)}}{M_n}$ to exist for all $i \in N$. And

(2) $\omega_{\mathrm{i}}(z) = \lim_{n\to\infty} \omega_i^{(n)}(z)$ to exist for all $i \in N$.

Thus $\lim_{n\to\infty} \Psi_n(z) \equiv \Psi(z)$ too exists and can be written as

$$\Psi(z) = e^{-\frac{z^2}{2}(1+\Omega(z))} \tag{9}$$

Such a limit defines the limit distribution's CF for $\overline{S}_n$. And function $\Omega(z)$ is given by

$$\Omega(z) = \sum_{i=1}^{\infty} \lambda_i \omega_i(\lambda_i z) \tag{10}$$

Proof. Define the sequence of random variables $y_1, y_2, ..., y_n$ such that
(1) the means of $y_1, y_2, ..., y_n$ are zero,
(2) the standard deviation of y_i is λ_i, and
(3) reduced variables y_i / λ_i have a CF given by $\phi_i(z) = e^{-\frac{z^2}{2}(1+\omega_i(z))}$

Then let us consider the sum $y_1 + y_2 + ... + y_n = Y_n$, and its reduced variable $\bar{Y}_n = \frac{Y_n}{\sigma_n}, \sigma_n = \lambda_1 + \lambda_2 + ... + \lambda_n$. It is clear that $\lim_{n\to\infty} \sigma_n = 1$. According to a Kolmogorov\`s (Kolmogorov, 1921) theorem, variable Y_n converges in probability to a finite value. Its CF is given by (9) and (10). Thus the series in (10) converges to all *z*, and is continuous in the neighborhood of the origin. According hypothesis (1) and (2) the function $\Omega_n(z)$, given by Eq. (8), converges to the function given in Eq. (10), then we conclude that sequence in Eq. (7) converges for all z and its limit is continuous around the origin. Thus, according to Levy's continuity theorem, we conclude that the distribution of S_n/M_n converges to a well-defined limit distribution.

Now we move on to define an identically distributed reduced process (IDRP) as one in which $\omega_i^{(n)}(z) = \omega(z), i = 1,...,n$. In this case the stochastic variables $x_i^{(n)}$ have the same reduced distribution.

Since the $\omega_i^{(n)}(z)$ s are the same, then hypothesis 2 of Theorem 1 holds. We assume hypothesis 1 to hold as well. Considering Eq. (6) yields

$$\psi_i^{(n)}(z) = e^{-\frac{z^2}{2}(1+\omega(z))} \equiv \psi(z) \tag{11}$$

If $\psi(z)$ is analytic, then all statistical moments of the distribution are finite. As a result, one gets Taylor expansion

$$\psi(z) = 1 + I^2 \frac{1}{2!} z^2 + I^3 \frac{\mu_3}{3!} z^3 + I^4 \frac{\mu_4}{4!} z^4 + ..., \mu_p = \langle \left(\bar{X}_i^{(n)} \right)^p \rangle \tag{12}$$

Function $\omega(z) = \omega_R(z) + I\omega_I(z)$ in Eq. (11) can also be expanded in series to produce

$$\begin{aligned} \omega_R(z) &= K_2 z^2 + K_4 z^4 + ... + K_{2p} z^{2p} + ... \\ \omega_I(z) &= K_1 z^1 + K_3 z^3 + ... + K_{2p-1} z^{2p-1} + ... \end{aligned} \tag{13}$$

Employing Eq. (12) one gets the k_i s as a function of μ_i. For instance

$$K_1 = -\frac{1}{3}\mu_3, K_2 = \frac{1}{4} - \frac{1}{12}\mu_4, K_3 = -\frac{1}{6}\mu_3 + \frac{1}{60}\mu_5,$$
$$K_4 = \frac{1}{12} - \frac{1}{36}\mu_3^2 - \frac{1}{24}\mu_4 + \frac{1}{360}\mu_6. \tag{14}$$

Now we expand function $\Omega(z) = \Omega_R(z) + I\Omega_I(z)$ in Eq. (10) in series

$$\Omega_R(z) = L_2 z^2 + L_4 z^4 + L_{2p} z^{2p} + \ldots$$
$$\Omega_I(z) = L_1 z + L_3 z^3 + L_{2p-1} z^{2p-1} + \ldots \tag{15}$$

So

$$\Omega_R(z) = \sum_{n=1}^{\infty} \lambda_i^2 \omega_R(\lambda_i z), \Omega_I(z) = \sum_{n=1}^{\infty} \lambda_i^2 \omega_I(\lambda_i z) \tag{16}$$

Substituting Eq. (13) in Eq. (16) and comparing the output with Eq. (15) produce

$$L_p = K_p \sum_{n=1}^{\infty} \lambda_i^{2+p}, \lambda_i = \lim_{n\to\infty} \frac{m_i}{M_n}, p \geq 1 \tag{17}$$

Thus one gets an expression for the CF of IDRPs with analytical distributions of $\bar{x}_i^{(n)}$. What if this condition is not fulfilled? Here $\bar{x}_i^{(n)}$ will have infinite moments. To examine this case we first show that the CF in Eq. (11) can be expanded in terms of the finite moments of the distribution.

3 Another theorem on the existence of limits to the sum of random variables

Now we consider that a variable $x_i^{(n)}$ has finite moments up to order 4. Its kurtosis will be denoted by $K_i^{(n)}$ and its skewness by $S_i^{(n)}$. The reduced variable's CF satisfies

$$\psi_i^{(n)}(z) = e^{-\frac{z^2}{2}(1+\omega_i^{(n)}(z))}, \omega_i(0) = 0 \tag{18}$$

and

$$\omega_i^{(n)}(z) = -I\frac{S_i^{(n)}}{3}z - \frac{K_i^{(n)}}{12}z^2 + \rho_i^{(n)}(z) \tag{19}$$

where

$$\rho_i^{(n)}(z) = O(z^2), \frac{\rho_i^{(n)}(z)}{z^2} \to 0, z \to 0$$

The CF of reduced variable $\bar{S}_n$ can be obtained using $\Omega_n(z)$. From Eq. (19) we can write

$$\Omega_n = -\frac{1}{12}\left(\sum_{i=1}^{n} K_i^{(n)} \frac{m_i^4}{M_n^4}\right) z^2 - I\frac{1}{3}\left(\sum_{i=1}^{n} S_i^{(n)} \frac{m_i^3}{M_n^3}\right) z + \sum_{i=1}^{n} \frac{\left(m_i^{(n)}\right)^2}{M_n^2} \rho_i \left(\frac{m_i^{(n)}}{M_n^{(n)}} z\right) \tag{20}$$

At this point we make the same hypotheses 1 and 2 of Theorem 1, apart from the fact that now we replace functions $\omega_i^{(n)}(z)$ with functions $\rho_i^{(n)}(z)$ and consider an extra hypothesis as follows.

(3) Both series $\left(\sum_{i=1}^{\infty} K_i \frac{m_i^4}{M_n^4}\right)$ and $\left(\sum_{i=1}^{\infty} S_i \frac{m_i^3}{M_n^3}\right)$ converge.

Thus we can state another theorem.

Theorem 2. If the sum of random variables in Eq. (2) is such that hypotheses 1−3 hold, then there exists a distribution F (associated with the reduced variable) that is the limit of the process as $n \to \infty$.

4 Limits to the sum of random variables when the variance follows a formation law

Now we reckon functions $\Omega(z)$ and $\Omega_n(z)$ to processes following a formation law in the second moment. We consider two cases in IDRPs, namely

(1) Exponential law: $m_i^{(n)} \equiv m_i = Ae^{Bi}, A > 0, B \in \Re$, and

(2) Power law: $m_i^{(n)} \equiv m_i = Ai^B, A > 0, B \in \Re$.

4.1 Analysis of the exponential law

To fully understand $\lim_{n\to\infty}\left(\frac{m_i}{M_n}\right)$, we need to evaluate M_n. For $m_i = Ae^{Bi}$,

$$\frac{m_{i+1}^2}{m_i^2} = e^{2B} \equiv r > 0 \tag{21}$$

holds. Here M_n is given by

$$M_n^2 = m_1^2\left(1 + r + r^2 + ... + r^{n-1}\right) \tag{22}$$

Taking the sum in Eq. (22) yields

$$M_n^2 = m_1^2\left(\frac{1-r^n}{1-r}\right) \tag{23}$$

We assess two possibilities, namely

(1) $\lim_{n\to\infty}\left(\frac{\mu_n}{M_n}\right) = (1-r)^{1/2} > 0$, if $0 < r < 1$

and

(2) $\lim_{n\to\infty}\left(\frac{\mu_n}{M_n}\right) = \left(\frac{r-1}{r}\right)^{1/2}$, if $r > 1$

where $\mu_n = \max_i \{m_i, i = 1,...,n\}$. Thus according to Theorem 1, sum variable $\bar{S}_n$ presents a limit distribution function that fails to be Gaussian.

4.2 Analysis of the power law

Now we turn to the power law. It is appropriate to introduce function

$$Z(n,r) = 1 + 2^r + 3^r + ... + n^r, r \in \Re \tag{24}$$

We can then write

$$m_i^2 = A^2 i^r, M_n^2 = A^2 Z(n,r), r = 2B \tag{25}$$

We consider the cases $r < -1$ and $r \geq -1$. Calculations (not shown) for $r < -1$ take account of the fact that

$$\lim_{n\to\infty}\left(\frac{\mu_n}{M_n}\right) = \lim_{n\to\infty}\frac{1}{\sqrt{Z(n,r)}} \tag{26}$$

where

$$\lim_{n\to\infty} Z(n,r) = \xi(|r|) \to \lim_{n\to\infty}\frac{\mu_n}{M_n} = \frac{1}{\sqrt{\xi(|r|)}} > 0, \xi(|r|) = \sum_{p=0}^{\infty}\frac{1}{p^{|r|}} \tag{27}$$

and $\xi(|r|)$ is Riemman's zeta function. Thus the limit distribution fails to be Gaussian and $\lim_{n\to\infty} M_n$ has a finite value.

For case $r \geq -1$, we have $\lim_{n\to\infty}\frac{\mu_n}{M_n} = 0$, and the limit distribution function is Gaussian.

4.3 Characteristic function of identically distributed reduced processes

Now we tackle the problem of determination of the CF in IDRPs. We take the case of an analytical CF. If the CF is analytic, function $\omega(z)$ can be expanded in series as in Eqs. (13) and (15). Considering these equations we have $L_{np} = K_p \sum_{i=1}^{n}\left(\frac{m_i}{M_n}\right)^{p+2}$. Our task is to calculate an expression for $\sum_{i=1}^{n}\left(\frac{m_i}{M_n}\right)^{p+2}$.

We again illustrate our case with exponential and power laws. As for the exponential law one has

$$m_i = m_1 r^{\frac{i-1}{2}} \, (r = e^{2B}) \rightarrow \left(\frac{m_{i+1}}{m_i}\right)^{p+2} = r^{\frac{2+p}{2}} = R_p \tag{28}$$

Thus sum $m_1^p + m_2^p + ... + m_n^p$ is geometric and accordingly

$$m_1^p + m_2^p + ... + m_n^p = m_1^{p+2}\left(\frac{1-R_p^n}{1-R_p}\right) \tag{29}$$

Using Eq. (23) for M_n yields

$$M_n^p = m_1^{p+2}\left(\frac{1-r^n}{1-r}\right)^{\frac{p+2}{2}} \tag{30}$$

And from Eqs. (29) and (30) one gets

$$L_{np} = K_p \frac{(1-r)^{\frac{p+2}{2}}\left(1-r^{\frac{p+2}{2}n}\right)}{\left(1-r^{\frac{p+2}{2}}\right)(1-r^n)^{\frac{p+2}{2}}}, 0 < r < 1 \tag{31}$$

and

$$L_{np} = K_p \frac{(r-1)^{\frac{p+2}{2}}\left(r^{\frac{p+2}{2}n} - 1\right)}{\left(r^{\frac{p+2}{2}} - 1\right)(r^n - 1)^{\frac{p+2}{2}}}, r > 1 \tag{32}$$

As $n \to \infty$ function $\Omega(z)$ becomes

$$\Omega(z) = \sum_{i=even} L_p z^p + I \sum_{i=odd} L_p z^p$$

where

$$L_p = \lim_{n\to\infty} L_{np} = K_p \frac{(1-r)^{\frac{p+2}{2}}}{\left(1 - r^{\frac{p+2}{2}}\right)}, 0 < r < 1, L_p = \lim_{n\to\infty} L_{np} = K_p \frac{(r-1)^{\frac{p+2}{2}}}{\left(r^{\frac{p+2}{2}} - 1\right)}, r > 1 \tag{33}$$

As for the power law, it can be shown that

$$M_n^{p+2} = A^{p+2} Z(n,r)^{\frac{p+2}{2}}, m_1^{p+2} + ... + m_n^{p+2} = A^{p+2} Z\left(n, \frac{p+2}{2} r\right) \tag{34}$$

and then one gets

$$L_{np} = K_p \frac{Z\left(n, \frac{p+2}{2} r\right)}{Z(n,r)^{\frac{p+2}{2}}} \tag{35}$$

Thus as $n \to \infty$ function $\Omega(z)$ is either

$$L_p = \lim_{n\to\infty} L_{np} = K_p \frac{\xi\left(\left|\frac{p+2}{2} r\right|\right)}{\xi\left(|r|^{\frac{p+2}{2,}}\right)}, r < -1 \tag{36}$$

or

$$L_p = \lim_{n\to\infty} L_{np} = 0 \to \Omega(z) = 0, r \geq -1 \tag{37}$$

In the latter case the process will reach the Gaussian regime. For the distinct processes defined by $r = 2B > -1$ there will be different 'convergence speeds', as measured by L_{np} in Eq. (35).

Now we will examine the asymptotic behavior of both M_n and the terms in the expansion of Ω_n. Our approach relies heavily on analysis of $Z(n,r)$. It can be shown that

$$\begin{aligned} &Z(n,r) \to \xi(r) \quad r < -1 \\ &Z(n,r) \to \log(n) \quad r = -1 \\ &Z(n,r) \to \frac{n^{r+1}}{r+1} \quad r > -1 \end{aligned} \tag{38}$$

As for the analytical IDRPs with second moment $m_i = m_1 i^{r/2}$, we consider results for five distinct values of r.

(1) For $r < -1$, using Eqs. (35) and (38) one can show that a process converges to a non-Gaussian distribution where variance M_n and the asymptotic terms in the series of Ω_n are

$$M_n \to m_1 \xi\left(\left|r\right|\right)^{-1/2}, L_{np} \to K_p \frac{\xi\left(\left|(p+2)/r\right|\right)}{\xi\left(\left|r\right|\right)^{(p+2)/2}}, p \in N \tag{39}$$

(2) For $r = -1$, a process converges to a Gaussian distribution where the variance and the asymptotic terms in the series of Ω_n are

$$M_n \to m_1 \left(\log n\right)^{1/2}, L_{np} \to K_p \xi\left(\left|(p+2)/2\right|\right)(\log n)^{-1}, p \in N \tag{40}$$

(3) For $-1 < r < 0$, a process converges to a Gaussian where the variance and the asymptotic terms in the series of Ω_n are

$$\begin{aligned} &M_n \to \frac{m_1}{\left(r+1\right)^{1/2}} n^{(r+1)/2} \\ &L_{np} \to K_p \frac{2(r+1)^{(p+2)/2}}{2+(p+2)r} n^{-p/2}, (p+2)r/2 > -1 \\ &L_{np} \to K_p (r+1)^{(p+2)/2} \frac{\log n}{n^{(p+2)(r+1)/2}}, (p+2)r/2 = -1 \\ &L_{np} \to K_p (r+1)^{(p+2)/2} \frac{\xi\left(\left|(p+2)r/2\right|\right)}{n^{(p+2)(r+1)/2}}, (p+2)r/2 < -1 \end{aligned} \tag{41}$$

with $p \in N$. The second possibility occurs only if p satisfies $(p+2)r/2 = -1$.

(4) For $r = 0$ (IID process) the second moment is the same for all *i*, i.e. $\forall i \to m_i = m_1$. The process converges to a Gaussian with the variance and asymptotic terms in the series of Ω_n given by

$$M_n \to m_1 n^{1/2}, L_{np} \to K_p n^{-p/2}, p \in N \tag{42}$$

(5) For $r > 0$, one has

$$M_n \to \frac{m_1}{(r+1)} n^{(r+1)/2}, L_{np} \to K_p \frac{2(r+1)^{(p+2)/2}}{2+(p+2)r} n^{-p/2}, p \in N \tag{43}$$

At this point some remarks are worthwhile. As for $r > -1$ the cumulative standard deviation M_n is governed by power law $n^{(r+1)/2}$ as $n \to \infty$. And a logarithm law holds for $r = -1$. The process then bifurcates into two ones, namely (1) convergence to a Gaussian ($r > -1$), and (2) no convergence ($r < -1$). In the latter situation the second moment has a saturation value at $m_1 \xi(|r|^{1/2})$.

As for the power laws in $r \geq 0$, L_{np} (which defines the series for $\Omega_n(z)$) shows an asymptotic decay with same power law $n^{-p/2}$. This process thus converges at the same pace to the Gaussian regime as *n* approaches infinity. But the process can show distinct transient behavior at the onset of aggregation. A large transient phase means a slow convergence, suggesting that $\Omega_n(z)$ varies sluggishly. This kind of behavior can be observed in truncated Lévy flights (TLFs) (Mantegna and Stanley, 1995).

4.4 A class of nonidentically distributed processes

Now we consider a class of stochastic processes with finite fourth moment. We assume every $\omega_i^{(n)}(z)$ to be distinct for alternative *i*. If the fourth moment is finite one has

$$\omega_i^{(n)}(z) = \omega_i(z) = -\frac{K_i}{12} z^2 - I \frac{S_i}{3} z + \rho(z) \tag{44}$$

where $\forall i \Rightarrow \rho_i(z) = o(z^2)$, i.e. $\frac{\rho_i(z)}{z^2} \to 0$ when $z \to 0$. The power laws we take are

$$m_i = m_1 i^{r/2}, S_i = S_1 i^{3y/2}, K_i = K_1 i^{2x}, r, x, y \in \Re, S_1, K_1 \geq 0 \tag{45}$$

In such cases

$$\Omega_n = -\frac{K_n}{12} z^2 - I \frac{S_n}{3} z + o(z^2) \tag{46}$$

where

$$S_n = \left(\sum_{i=1}^{n} S_i \frac{m_i^3}{M_n^3} \right), K_n = \left(\sum_{i=1}^{n} K_i \frac{m_i^4}{M_n^4} \right) \tag{47}$$

Employing Eqs. (21) and (36) yields

$$M_n = m_1 Z(n,r)^{1/2}, S_n = S_1 \frac{Z\left(n, 3(r+y)/2\right)}{Z(n,r)^{3/2}}, K_n = K_1 \frac{Z\left(n, 2(r+x)\right)}{Z(n,r)^2} \tag{48}$$

The IDRP calculated in this section encompasses that of previous section, where $x = 0$ and $y = 0$. And the suggested formulas are suitable for description of a class for which $m_i = m, \forall i$, i.e. an IID process where $x = 0, y = 0$, and $r = 0$.

Now we examine the asymptotic behavior of the above processes. We take into account the standard deviation, skewness, and kurtosis, as defined in Eq. (48). We consider the evolving laws in Eq. (45) with $r > -1, y > 0$ and $x > 0$.

As $n \to \infty$ then

$$\begin{aligned} M_n &\to \frac{m_1}{\left(r+1\right)^{1/2}} n^{(r+1)/2} \\ S_n &\to S_1 \frac{2(r+1)^{3/2}}{2+3(r+y)} n^{(3y-1)/2} \\ K_n &\to K_1 \frac{(r+1)^2}{2(r+x)+1} n^{2x-1} \end{aligned} \tag{49}$$

Quantities M_n, S_n, and K_n follow power laws of distinct exponents. It is intriguing (1) that M_n depends only on *r*, which describes the evolution of m_i, (2) that S_n depends on *y*, which shows how skewness S_i evolves over time, and (3) that K_n depends only on *x*, which gives the evolution of kurtosis K_i.

Note that for $0 < y < 1/3$ one has $S_n \to 0$ as $n \to \infty$. If $y = 1/3$ then the skewness approaches a finite value and the limit distribution will fail to be Gaussian. For $y > 1/3$ one has $S_n \to \infty$ as $n \to \infty$, i.e. a limit process at an infinite distance from the Gaussian.

A similar analysis can be delivered to the kurtosis. As $n \to \infty$,

$$\begin{aligned} &0 < x < 1/2 \to K_n \to 0 \\ &x = 1/2 \to K_n \to C \\ &x > 1/2 \to K_n \to \infty \end{aligned} \tag{50}$$

After a threshold value at bifurcation point ½, the third and fourth moments diverge. The bifurcation parameters are then $x = 1/2$ and $y = 1/3$.

4.5 Probability of return to the origin and scaling properties of statistical moments

Now we sketch an explanation for asymptotic laws in the 'probability of return to the origin' of TLFs (Mantegna and Stanley, 1995). First we introduce the concept of asymptotic Ω-stability.

Definition 1. There exists an n_0 such that, for all $n > n_0$, function Ω_n is nearly constant.

The notion captures the idea of a process whose CF approaches a threshold. This is equivalent to $\Omega_n(z) = \Omega(z), \forall n > n_0$. Thus $P_n(0)$ becomes

$$P_n(0) = \frac{1}{M_n} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2(1+\Omega(z))} dz = \frac{\text{constant}}{M_n} \tag{51}$$

Once we have learned the asymptotic laws governing M_n (from the cases studied here), the asymptotic laws for $P_n(0)$ end up known. For the exponential law one has

$$\begin{aligned} P_n(0) &= \frac{C\left(1-r\right)^{1/2}}{m_1} \quad \text{for } r < 1, \\ P_n(0) &= \frac{C\left(r-1\right)^{1/2}}{m_1} e^{-n} \quad \text{for } r > 1 \end{aligned} \tag{52}$$

And for the power law it holds that

$$\begin{aligned} P_n(0) &= \frac{C}{m_1 \xi\left(|r|\right)} = \text{constant}, \quad \text{for } r < -1 \\ P_n(0) &= \frac{C}{m_1 (\log n)^{1/2}} \quad \text{for } r = -1 \\ P_n(0) &= \frac{C(1+r)^{1/2}}{m_1} \frac{1}{n^{(1+r)/2}} \quad \text{for } r > -1 \end{aligned} \tag{53}$$

So there are distinct evolution laws for the probability of return. These can be either a constant, an exponential law, or a power law. If we take the latter (usually dubbed 'scaling') we get

$$P_n(0) = \frac{\text{constant}}{M_n} = \frac{D}{n^{1/\alpha}} \Rightarrow$$
$$D = \frac{C(1+r)^{1/2}}{m_1}, n^{1/\alpha} = n^{(1+r)/2} \Rightarrow \alpha = \frac{2}{r+1}, r > -1 \quad (54)$$

i.e. $\alpha > 0$. If $-1 < r < 0$ then $2 < \alpha < \infty$. If $r = 0$ then $\alpha = 2$. And if $r > 0$ then $0 < \alpha < 2$. The latter case is that of a TLF. The probability of return in the remaining cases still scales with *n*, but surely they cannot converge to the TLF.

For the asymptotic stable processes the evolution of moments $\langle S_n^p \rangle$ is entirely determined by the evolution law of dispersion M_n. To see this remember that a CF of $\overline{S}_n$ is

$$\overline{\Psi}_n(z) = e^{-\frac{z^2}{2}(1+\Omega(z))} = 1 + \frac{I^2}{2!}z^2 + \frac{I^3}{3!}K_3 z^3 + ..., \forall n > n_0 \quad (55)$$

And that the CF of S_n is

$$\Psi_n(z) = \Psi_n(M_n z) = 1 + \frac{I^2}{2!}M_n^2 z^2 + \frac{I^3}{3!}M_n^3 K_3 z^3 + ... \quad (56)$$

These two equations produce

$$\overline{\Psi}_n(z) = 1 + \frac{I^2}{2!}\langle S_n^2 \rangle z^2 + \frac{I^3}{3!}\langle S_n^3 \rangle z^3 + ..., \langle S_n^p \rangle = K_p M_n^p \quad (57)$$

Thus the evolution law of M_n fully determines other moments' motion laws.

5 Example

This section employs some of the results in previous sections to examine a financial time series, namely the Brazilian *real*–US dollar exchange rate. We take returns *Z* rather than raw data as our stochastic variable, i.e. $Z_{\Delta t}(t) = Y(t+\Delta t) - Y(t)$ where *Y*(*t*) is a rate at day *t*. Returns x_i are ranked, i.e. $m_1 \geq m_2 \geq ... \geq m_n$. We define vector $\mathbf{X}_n = (x_1, x_2, ..., x_n)$ and permutation operator *P* such that $P\mathbf{X}_n = (x_{i_1}, ... x_{i_n})$ with $(i_1, ..., i_n)$ being any permutation of $(1, ..., n)$. Sum variable is $X_n = x_{i_1} + ... + x_{i_n} = S(P\mathbf{X}_n)$.

Then we define an event as an *n*-dimensional, real vector generated by the following rule.

Definition 2. To create event $\mathbf{E}$ we first select vector $\mathbf{X}_n$ and permutation operator P, and then build up another vector in which the k th component is randomly generated by a number from the distribution of x_{i_k} .

The event is a realization of vector $\mathbf{X}_n$ followed by a permutation of any of its components. The sum of N events $\mathbf{E}_i = (E_{i1}, E_{i2}, ...E_{in})$ is defined as $S_i = S(\mathbf{E}_i) = E_{i1} + ... + E_{in}$. For N large enough, sequence $S_1, S_2, ..., S_N$ presents the same distribution as that of sum variable X_n.

For series of events $E_1, E_2, ..., E_N$, a list of $N' = nN$ numbers can be obtained, i.e.

$$\mathbf{u} = \left[u_1, ..., u_n, u_{n+1}, ..., u_{2n}, ..., u_{(N-1)n+1}, ..., u_{nN} \right] \tag{58}$$

where $u_1 = E_{11}, u_2 = E_{12}, ..., u_n = E_{1n}, u_{n+1} = E_{21}$. Eq. (58) can be thought of as an n-periodic stochastic process. If a period is known with certainty, Eq. (58) can be used to produce list $\mathbf{U} = [U_1, ..., U_N]$, where $U_1 = u_1 + ... + u_n, U_2 = u_{n+1} + ... + u_{2n}, ...$, and so on. How to get the standard deviations m_i? One is aware by the very nature of the process that they are not ordered. Besides, it is not generally possible to evaluate the standard variation of a stochastic variable from a single measure. But such problems can still be overcome (see (Figueiredo *et al*. 2005).)

We then employ such a routine to the 15-minute spaced Brazilian *real*-US dollar rate from year 2002 (Fig. 1). Using Eq. (58) we get list $\mathbf{u}$ from this set of data. The list is made up of N' = 6140 numbers. Then we get list $\mathbf{U}$ from a 'daily' set of data, which is built up as follows. A 'day' is considered to have 20 data points from the original 15-minute series. So the process' period is n = 20. Since $N' = nN$, list $\mathbf{U}$ ends up with 307 numbers.

The skewness and kurtosis of the two lists are

$$\begin{aligned} Sk(\mathbf{u}) &= 3.0653, K(\mathbf{u}) = 114.4593 \\ Sk(\mathbf{U}) &= 1.5288, K(\mathbf{U}) = 19.3846 \end{aligned} \tag{59}$$

As can be seen, the hypothesis of an IID is promptly discarded because

$$\begin{aligned} Sk(\mathbf{U}) &= \frac{1}{n^{1/2}} Sk(\mathbf{u}) = \frac{1}{20^{1/2}} 3.0653 = 0.6854 \\ K(\mathbf{U}) &= \frac{1}{n} K(\mathbf{u}) = \frac{1}{20} 114.4593 = 5.2965 \end{aligned} \tag{60}$$

So we evaluate whether the process above can be explained in terms of our suggested IDRP. To apply the technique summed up in (Figueiredo *et al*., 2005) we take N = 309 periods of size $n - 1$ = 19. Standard deviations m_i are shown in Fig. 2. Using the values of m_i produces

$$Sk(\mathbf{U}) = \frac{m_1^3 + ... + m_n^3}{\left(m_1^2 + ... + m_n^2\right)^{3/2}} Sk(\mathbf{u}) = 1.3497$$
$$K(\mathbf{U}) = \frac{m_1^4 + ... + m_n^4}{\left(m_1^2 + ... + m_n^2\right)^2} K(\mathbf{u}) = 25.6804 \tag{61}$$

which is in good agreement with experimental data.

Now we examine whether there is a law governing the standard deviations in Fig. 2. First we try out exponential law $m_i = Ae^{-Bi}$. Fig. 3 displays the data in Fig. 2 together with $m_i = Ae^{-Bi}$. Using this exponential law it can be shown with some algebra that

$$Sk(\mathbf{U}) = Sk(\mathbf{u}) \frac{(1-r)^{\frac{3}{2}}}{\left(1-r^{3/2}\right)} \frac{\left(1-r^{\frac{3}{2}n}\right)}{\left(1-r^n\right)^{3/2}}$$
$$K(\mathbf{U}) = K(\mathbf{u}) \frac{(1-r)^2}{\left(1-r^2\right)} \frac{\left(1-r^{2n}\right)}{\left(1-r^n\right)^2}, r = e^{-2B} < 1 \tag{62}$$

Eq. (62) and *B* (Table 1) together yield

$$Sk(\mathbf{U}) = 1.3086$$
$$K(\mathbf{U}) = 23.9674 \tag{63}$$

which is in agreement with the experimental data.

Power laws $m_i = Ai^{-B}$ and $m_i = (A + iB)^{-1/C}$ are presented in (Figueiredo *et al.* 2005). Table 2 sums up results. So the best fit is found with the assumption of an IDRP together with the exponential law describing second moment's behavior.

6 Characteristic function approach to the sum of autocorrelated variables

So far only independent stochastic variables have been considered. Now we tackle the problem of the sum of nonindependent variables (Figueiredo *et al*., 2004). Here we restrict ourselves to sequence $x_i^{(n)} = x_i$. We denote the moments of order *p* of x_i and S_n as $\mu_{ip} = \langle x_i^{\,p} \rangle$ and $\nu_{np} = \langle S_n^{\,p} \rangle$ respectively. The CFs of x_i and S_n are

$$\psi_i(z) = \langle e^{Ix_i z} \rangle = e^{-\mu_{i2} z^2 (1+\omega_i(\sqrt{\mu_{i2}} z))/2} \tag{64}$$

and

$$\Psi_n(z) = \langle e^{IS_n z} \rangle = e^{-\nu_{n2} z^2 (1+\Omega_n(\sqrt{\nu_{n2}} z))/2} \tag{65}$$

respectively.

The existence of a CF in sum variable as $n \to \infty$ is assured by Levy's continuity theorem. As for independent variables it holds that $\Psi_n(z) = \psi_1(z) \ldots \psi_n(z)$. Yet this does not extend to autocorrelated processes where

$$\Psi_n(z) = C_n(z) \psi_1(z) \ldots \psi_n(z) \tag{66}$$

Written $C_n(z)$ as

$$C_n(z) = e^{-\frac{z^2}{2}(-2C_{n2} + W_n(z))}, \tag{67}$$

The CF in Eq. (66) then becomes

$$\Psi_n(z) = e^{-z^2 \left(\nu_{n2} + \sum_{i=1}^{n} \mu_{i2} \omega_i (\mu_{i2}^{1/2} z) + W_n(z) \right)/2} \tag{68}$$

After writing the CF of the reduced variable as

$$\overline{\Psi}_n(z) = \Psi_n\left(z / \sqrt{\nu_{n2}} \right) = e^{-z^2 \left(1 + \Omega_n^{(1)}(z) + \Omega_n^{(2)}(z) \right)/2} \tag{69}$$

one gets

$$\Omega_n^{(1)}(z) = \frac{1}{\nu_{n2}} \sum_{i=1}^{n} \mu_{i2} \omega_i \left(z \sqrt{\mu_{i2}} / \sqrt{\nu_{n2}} \right) \tag{70}$$

and

$$\Omega_n^{(2)}(z) = \frac{1}{\nu_{n2}} W_n\left(z / \sqrt{\nu_{n2}} \right) \tag{71}$$

Function $\Omega_n^{(1)}(z)$ matches that for uncorrelated series, i.e. as $n \to \infty$ it approaches zero. And term $\Omega_n^{(2)}(z)$ is related to the existence of autocorrelations. This term gives sum variable's CF, which can be used to obtain its PDF as $n \to \infty$.

6.1 Autocorrelation and convergence

Now we put forward an expression for $\Omega_n^{(2)}(z)$ containing only statistical moments. Expanding the CFs in series and assuming

$$C_n(z) = 1 + C_{n2}z^2 + C_{n3}z^3 + C_{n4}z^4 + o(z^4), \quad (72)$$

$W_n(z)$ can be extended to $W_n(z) = IW_{n1}z + W_{n2}z^2 + o(z^2)$. Comparing equal order terms produces

$$C_{n2} = -\frac{1}{2}(\nu_{n2} - \sigma_{n2}), C_{n3} = -\frac{I}{3!}(\nu_{n3} - \sigma_{n3})$$
$$C_{n4} = \frac{1}{4!}(\nu_{n4} - \sigma_{n4}) - \frac{1}{2!2!}(\sigma_{n2}(\nu_{n2} - \sigma_{n2}) + \gamma_n) \quad (73)$$

where $\gamma_n = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mu_{i2}\mu_{j2} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \langle x_i^2 \rangle \langle x_j^2 \rangle$ and $\sigma_{np} = \sum_i \mu_{ip}$. It can also be shown that

$$W_{n1} = \frac{1}{3}(\nu_{n3} - \sigma_{n3}), W_{n2} = \frac{1}{4}(\nu_{n2} - \sigma_{n2})^2 - 2C_{n4} \quad (74)$$

That renders a distribution suitable for practical purposes. First we define nonlinear autocorrelation term

$$\langle p_1 p_2 ... p_k \rangle_n = \sum_{i_1 ... i_k = 1}^{n} \left(\langle x_{i_i}^{p_1} ... x_{i_k}^{p_k} \rangle - \langle x_{i_i}^{p_1} \rangle ... \langle x_{i_k}^{p_k} \rangle \right) \quad (75)$$

where $p_1 p_2 ... p_k$ are positive integers, and $i_1 \neq i_2 \neq ... \neq i_k$.

Using Eq. (69) it can be shown that $\Omega_n^{(2)} = I\Omega_{n1}^{(2)} z + \Omega_{n2}^{(2)} z^2$, where

$$\Omega_{n1}^{(2)} = \frac{1}{3} \frac{\nu_{n3} - \sigma_{n3}}{\nu_{n2}^{3/2}} = \frac{\langle 111 \rangle_n + 3\langle 12 \rangle_n}{\nu_{n2}^{3/2}},$$
$$\Omega_{n2}^{(2)} = -\frac{1}{12} \frac{\nu_{n4} - \sigma_{n4} - 6\gamma_n}{\nu_{n2}^2} + \frac{1}{4}\left(1 - \frac{\sigma_{n2}^2}{\nu_{n2}^2}\right) \quad (76)$$
$$= (-1/12)(\langle 1111 \rangle_n + 6\langle 112 \rangle_n + 4\langle 13 \rangle_n + 3\langle 22 \rangle) / \nu_{n2}^2 + R_n$$

so $\Omega_{n1}^{(2)}$ and $\Omega_{n2}^{(2)}$ are functions of third- and fourth-order correlations respectively.

We conclude that, although linear autocorrelations play a key role in a distribution convergence, it is still necessary to take the nonlinear autocorrelations into account to fully characterize a process. Although that is arguably well known in literature for long, our novel technique presents formulas capturing the nonlinear terms explicitly.

6.2 Example

Now we illustrate our approach with data coming from the daily variations of the British pound against the US dollar. Our set is made up of 8780 data points, covering the time period from 1 April 1971 to 30 December 2005. As before, we take single returns rather than raw data.

Fig. 4a displays kurtosis behavior, and Fig. 4b that of skewness. These are the leading terms in the expansion of S_n. Curves of IID processes are shown for comparison. The skewness is clearly bounded by real values. This limits $\Omega_{n1}^{(2)}$ and prevents convergence to the Gaussian regime. (Figs. 5a and 5b show $\Omega_{n1}^{(2)}$ and $\Omega_{n2}^{(2)}$ respectively.) Further details can be found elsewhere (Figueiredo *et al*., 2004).

7 Relating identically distributed reduced processes to autocorrelated ones

Section 5 dealt with IDRP behavior, and Section 6 examined the sum of autocorrelated variables. The purpose of this section is to relate IDRPs to autocorrelated processes.

Nonidentity of a process is entirely determined by average values $\mu_{i1}, i = 1,...,N$ and standard deviations $\sqrt{\mu_{i2}}, i = 1,...,N$. We define a *reduced random generator* (RRG) with zero mean and unit standard deviation. Every x_i obtained in an IDRP are then given by $x_i = \sqrt{\mu_{i2}} G_r + \mu_{i1}$, where G_r is a random choice from a reduced distribution $g(x)$ of zero mean and unit standard deviation. Time series from an IDRP can be obtained as follows. First we pick a particular RRG, say G_r. Secondly we choose $\mu_{i1}, i = 1,...,N$ to track how the mean evolves over time. And finally we select $\sqrt{\mu_{i2}}, i = 1,...,N$ to capture standard deviation's time evolution.

Time series obtained from an IDRP generator can then be theoretically evaluated. We take a particular RRG derived from a truncated Cauchy (TC) distribution, i.e.

$$f(x) = \begin{cases} \dfrac{A}{1+x^2} & -L_1 \le x \le L_2 \\ 0 & x < -L_1 \text{ or } x > L_1 \end{cases} \tag{77}$$

with $L_1, L_2 > 0$ and $A = \dfrac{1}{\arctan(L_1) + \arctan(\mathrm{L}_2)}$.

Defining a random generator associated with x with distribution function $f(x)$ is a well-known problem. We can relate x to, say, y (which is uniformly distributed within interval $[0,1]$), and use the property of probability conservation to show that

$$x = \tan\left(\left[\arctan\left(L_1\right) + \arctan\left(L_2\right)\right] y - \arctan\left(L_1\right)\right) \tag{78}$$

Since y is uniformly distributed within [0,1], x is distributed in $[-L_1, L_2]$ with a TC. Then reduced variable $\bar{x} = \left(x - \mu_1\right) / \sqrt{\mu_2}$ has a TC distribution. A TC RRG of $\alpha = 1$ can be derived from

$$G_r = \frac{\tan\left(\left[\arctan\left(L_1\right) + \arctan\left(L_2\right)\right]\text{rand}() - \arctan\left(L_1\right)\right) - \mu_1}{\sqrt{\mu_2}} \quad (79)$$

where rand() stands for a uniform random generator within [0,1]. Thus to fully characterize an IDRP generator we should first set μ_{i1} and $\sqrt{\mu_{i2}}$ ($i = 1,2,\ldots,N$).

7.1 Example

We again take the time series from the daily changes of the British pound–US dollar rate to illustrate our case. The heart of our technique is as follows. We divide such a sequence into equal, non-overlapped time periods. Then we compute the means and standard deviations of these periods. For instance, defining a *p*-sized period as a sequence of *p* days means the series of 8780 days will have n_p periods of *p* days that are consecutive and non-overlapped ($n_p \times p = 8780$). We then calculate (for each of these periods) the means and standard deviations using the pound-dollar series.

Once p and n_p are characterized, we are ready to define IIDR random generator

$$\sqrt{(\mu_{i2})}G_r + A\mu_{i1}, i = 1,2,3,\ldots,8615 \quad (80)$$

where A is a real number within $[0,1]$. The μ_{i1} and $\sqrt{\mu_{i2}}$ are given by

$$\mu_{11} = \mu_{21} = \mu_{31} = \ldots = \mu_{p1} = \text{first-period mean,...} \quad (81)$$

and

$$\sqrt{\mu_{12}} = \sqrt{\mu_{22}} = \sqrt{\mu_{32}} = \ldots = \sqrt{\mu_{p2}} = \text{first-period standard deviation,...} \quad (82)$$

If $A = 0$ then we discard weekly mean's time evolution. In this case the generator produces a zero mean for all time periods and gets nonstationary in second moments. If $A = 1$ the generator is able to mimic the actual time series since it presents the same mean profile.

Now we compare the properties of the pound-dollar rate with the IDRP RRG of μ_{i1} and μ_{i2} as above. We show how time evolution of the two series' moments are similar if one selects appropriate A, L_1, and L_2.

If $A = 0$ time evolution of a week's mean is discarded. In this case the generator will feature a zero mean forever and second moments will be nonstationary. If $A = 1$ the generator mimics the actual time series' mean.

When comparing the statistical properties of the pound-dollar time series with those of an RRG (obtained with μ_{i1} and μ_{i2}) we realize that moments over time of the two series are similar (as long as we calibrate using appropriate A, L_1, and L_2). Thus an identically distributed process with autocorrelations can be obtained from an independently, yet nonidentically distributed, random generator.

Both symmetric ($L_1 = L_2$) and asymmetric ($L_1 \neq L_2$) cases are considered. For robustness, the routine is repeated twenty times to take mean values. In every case we pick a different seed for the uniformly distributed generator. Outcomes for processes with $p = 20$ (trading months of 20 days) are also considered.

Fig. 6 displays a symmetric RRG. We get $L_1 = L_2$ from maximum likelihood estimates. Kurtosis behavior in the IIDR process is very similar to actual value. This suggests that kurtosis behavior can be explained in terms of the time evolution characterizing the RRG. Skewness behavior is not that clear-cut, however. This is somewhat expected because the generator is symmetric, and $A = 0$. Standard deviation behaves as if the process had a Hurst exponent of ½.

8 Concluding remarks

This paper approaches the issue of the sum of stochastic variables and take independent processes that are identically distributed in their reduced variables as well as autocorrelated processes that are identically distributed. We extend the classic central limit theorem that features finite variance.

The paper also examines cases where a formation law for series' variance is present. Our suggested reduced variable (that is independent and identically distributed) seems to fit well a financial data set sampled from the intraday Brazilian *real*-US dollar rate of year 2002. We find the reduced variable together with an exponential law to mimic the series' volatility behavior. And we also find the reduced variable to fail reaching zero as sample size approaches infinity.

We too investigate the role of nonlinear autocorrelations in the dynamics of convergence to a Gaussian. We find sluggish convergence to be due to the nonlinear autocorrelations.

Thus some features of the nonlinear autocorrelations can be rationalized in terms of an independently distributed, reduced process. Information about the autocorrelations is already encompassed in mean and standard deviation's time paths. This is in line with the fact that a process is independent but not identically distributed. Nonidentity satisfactorily explains slow convergence to the Gaussian regime as well as greater-than-½ Hurst exponents. And it is possible to observe a non-IID behavior of skewness and kurtosis even when the Hurst equals ½.

Nonconvergence to a Gaussian is thus explained by departures from the infinitesimality hypothesis of independently distributed, reduced processes.

Acknowledgements

Annibal Figueiredo, Iram Gleria, and Sergio Da Silva acknowledge financial support from the Brazilian National Council for Scientific and Technological Development (CNPq).

References

Bavly, G. M. (1936) Über einige Verallgemeinerungen der grenzwertsätze der Wahrsheinlichkeitrenchnung, *Sbornik.*, **1**, 917-930.

Feller, W. (1935) Über den Zentralen Grenzwertsatz der Wahrsheinlichkeitrenchnung, *Math. Zeitshirift*, **40**, 521-529.

Feller, W. (1937) Über den Zentralen Grenzwertsatz der Wahrsheinlichkeitrenchnung, *Math. Zeitshirift*, **42**, 301-312.

Feuerverger, A. and Mureika, R. (1977) The empirical characteristic function and its applications, *The Ann. of Statist.*, **5**, 88-97.

Feuerverger, A. and McDunnough, P. (1981) On the efficiency of empirical characteristic function procedures, *J. R. Statist. Soc. B*, **43 (1)**, 20-27.

Figueiredo, A., Gleria, I., Matsushita and R., Da Silva, S. (2003) On the origins of truncated Levy flights, *Phys. Lett. A*, **326**, 51-60.

Figueiredo, A., Gleria, I., Matsushita, R. and Da Silva, S. (2004) Levy flights, autocorrelation, and slow convergence, *Physica A*, **337**, 369-383.

Figueiredo, A., Gleria, I., Matsushita, R. and Da Silva, S. (2005) Financial volatility and independent and identical distributed variables, *Physica A*, **346** (2005), 484-498.

Finetti, B. (1929) Sulla funzione a incremento aleatorio, *Tai. Acad. Naz. Lincei.,* **6**, 163-168, 325-329, 548-553.

Gnedenko, B. V. and Kolmogorov, A. N. (1954) *Limit distributions for sums of independent random variables*, Reading: Addison-Wesley.

Khintchine, A. J. (1938) *Limit laws for sums of independent random variables*, Moscow: ONTI.

Kolmogorov, A. N. (1921) Über die summen durch Zuffal bestimmter unabhängiger Gröβen, *Mathematishe Annalen,* **99**, 309-319.

Levy, P. (1922) Sur la loi de Gauss, *C. R. Acad. Sc.,* **174**, 1682-1684.

Levy, P. (1924) Théorie des erreurs. La loi de Gauss et les lois exceptionelles, *Bull. Soc. Math. France*, **LII**, 49-85.

Levy, P. (1929) Sur quelques travaux relatifs à la théorie de errers, *Bull. Soc. Math. France*, **LIII**, 1-29.

Levy, P. (1937) *Théorie de l´addition de variables aléatoires*, Paris : Gauthiers-Villars.

Liapunoff, A. M. (1900) Sur une proposition de la théorie dês probabilités, *Bull. De L´Acad. Des Sciences St. Petersbourg*, **13**, 359-386.

Lindberg, J. W. (1922) Eine neue herleitung des Exponentialgesetzes in der Wahrsheinlichkeitrenchnung, *Math. Zeitschrift,* **15**, 211-225.

Mantegna, R. N. and Stanley, H. E. (1994) Stochastic process with ultraslow convergence to a Gaussian, the truncated Levy flight, *Phys. Rev. Lett.*, **73**, 2946-2949.

Mantegna, R. N. and Stanley, H. E. (1995) Scaling behavior in the dynamics of an economic index, *Nature*, **376**, 46-49.

Zolotarev, V. M. (1990) Reflections on the classical theory of limit theorems, *Theory Probab. Appl.,* **36 (1)**, 124-137.

Zolotarev, V. M. (1965) On the closeness of distributions of two sums of independent random variables, *Theory Probab. Appl.,* **10** (1965), 519-526.

Zolotarev, V. M. (1997) *Modern theory of summation of random variables, Modern Probability and Statistics*, Utrecht: VSP.

	Estimated Value ± Standard Error
IIDR	A = 0.0156 ± 0.000599
Exponential law $m_i = Ae^{-Bi}$	B = 0.2070 ± 0.0101
IIDR	A = 0.0149 ± 0.000899
Power law $m_i = Ai^{-B}$	B = 0.7711 ± 0.0521
IIDR	A = 4.2267 ± 1.3063
Power law $m_i = (A + iB)^{-1/C}$	B = 0.4973 ± 0.3191
	C = 0.3607 ± 0.0807

Table 1. Estimated models

	Experimental Data	IID	IIRD	Exponential Law	Power Law	Power Law
Skewness	1.5287	0.6854	1.3497	1.3086	1.4296	1.3493
Kurtosis	19.3846	5.7230	25.6804	23.9674	31.0246	25.5359

Table 2. Skewness and kurtosis of experimental data under alternative assumptions

As can be seen, the exponential law gives a kurtosis that is closer to that of the experimental data

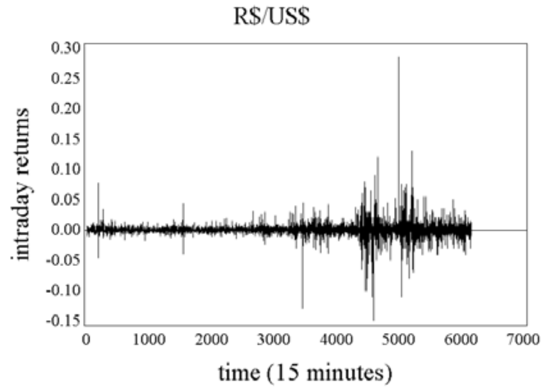


Fig. 1. Brazilian *real*-US dollar 15-minute spaced returns for year 2002

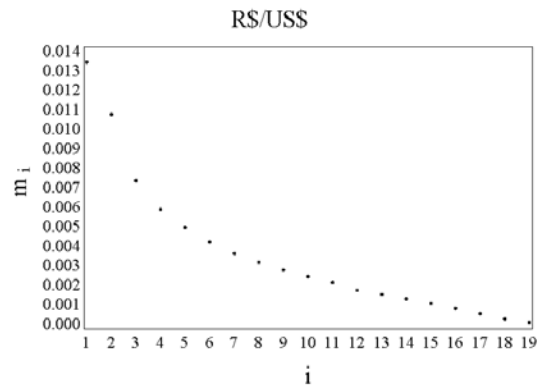


Fig. 2. Standard deviations m_i against i

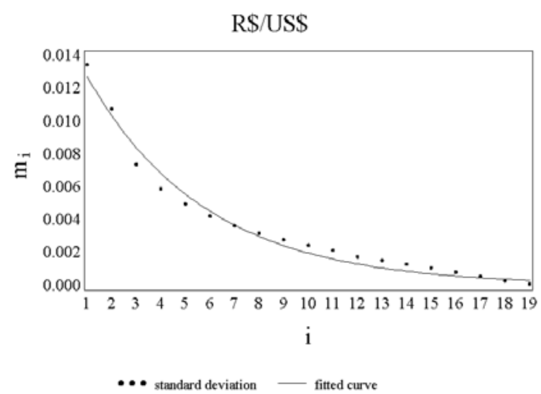
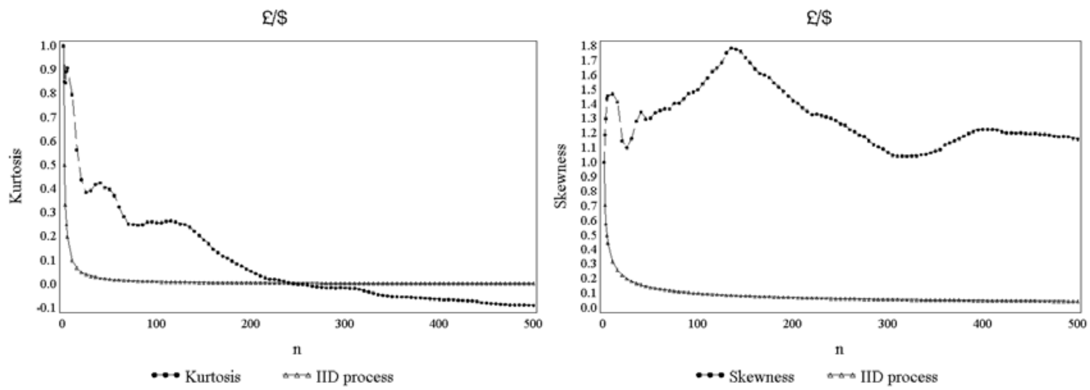
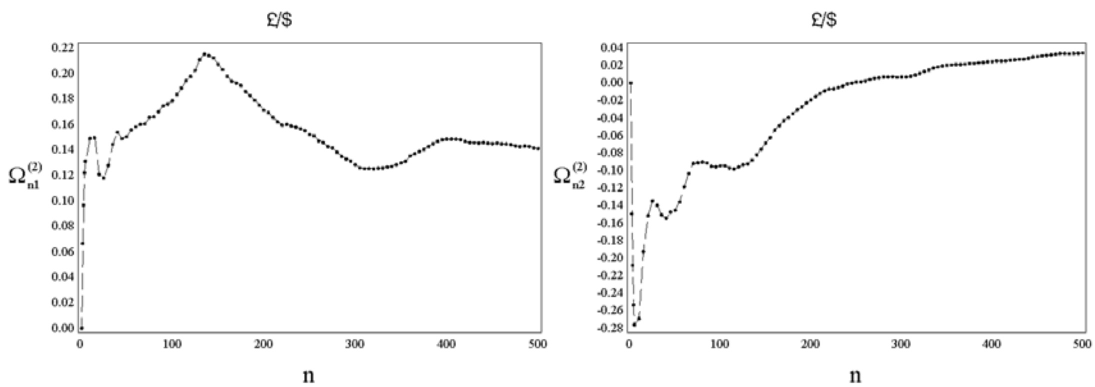


Fig. 3. Fitting exponential law $m_i = Ae^{-Bi}$



(a) (b)

Fig. 4. Behavior of kurtosis (a) and skewness (b) for daily returns of the pound-dollar rate
Curves for an IID process are shown for comparison



(a) (b)

Fig. 5.

(a) Behavior of $\Omega_{n1}^{(2)}$ for the daily pound-dollar rate

(b) Behavior of $\Omega_{n2}^{(2)}$

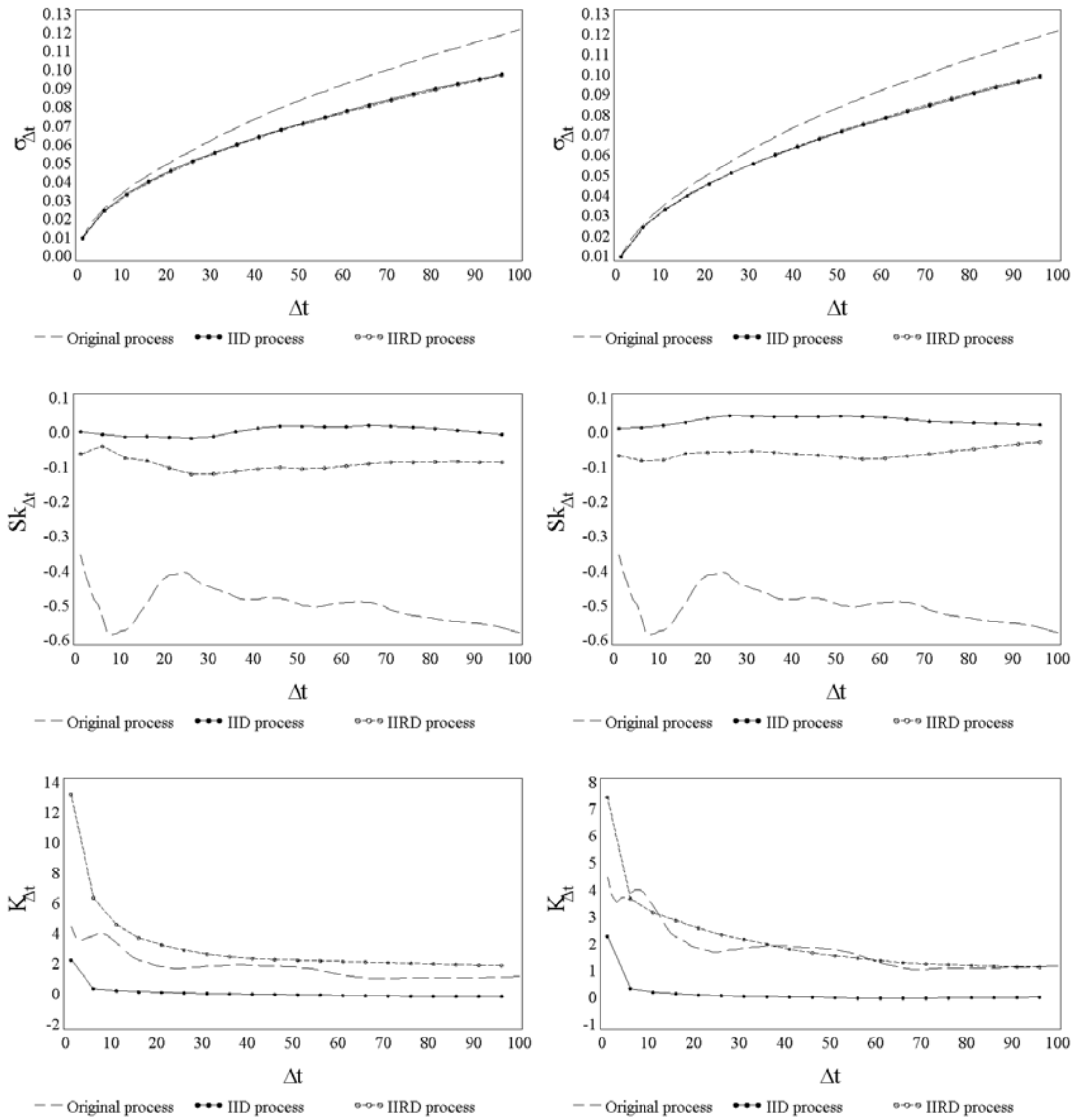


Fig. 6.
Scaling in standard deviations (upper panel), skewness (mid panel) and kurtosis (lower panel) of daily observations of the pound-dollar rate
The IIRD process on left plots is generated with $P = 5$, $A = 0$, and symmetric case $L = L_1 = L_2$
The maximum likelihood estimate of L is 7.52
The IIRD process on the right plots is generated with $P = 20$