



UNIVERSIDAD CARLOS III DE MADRID

working
papers

Working Paper 03-64
Statistics and Econometrics Series 14
October, 2003

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

ON THE RECORD PROPERTIES OF INTEGRATED TIME SERIES

Felipe M. Aparicio Acosta*

Abstract

This paper compares the statistical properties of the records from *i.i.d.* time series with those of time series containing a single unit root. It is shown that there are important differences in both the limiting distributions and the convergence rates of the associated record counting processes. Since the record properties of *i.i.d.* time series are shared by a large class of stationary time series, the reported differences underline the possibility of using record-based statistics for robust testing procedures of the unit root hypothesis.

Keywords and Phrases: Records, extreme order statistics, counting processes, renewal theory, geometric distribution, integrated time series, test of unit roots.

* Department of Statistics and Econometrics; Universidad Carlos III de Madrid, Getafe (Madrid), Spain. E-mail: aparicio@est-econ.uc3m.es

ON THE RECORD PROPERTIES OF INTEGRATED TIME SERIES

FELIPE M. APARICIO ACOSTA

Abstract. This paper compares the statistical properties of the records from i.i.d. time series with those of time series containing a single unit root. It is shown that there are important differences in both the limiting distributions and the convergence rates of the associated record counting processes. Since the record properties of i.i.d. time series are shared by a large class of stationary time series, the reported differences underline the possibility of using record-based statistics for robust testing procedures of the unit root hypothesis.

1. Introduction

The collection of records is an important aspect of empirical research in science. Records are ubiquitous in daily life. We are already accustomed to hear about records in sport competitions, in stock prices, claim sizes and financial gains; in series of temperature, of earthquake, pressure or star brightness intensities; of wind speeds and river floods; in maximal loads on engineering structures, or as regards the strength of materials, etc. As a consequence, the analysis of the statistical properties of records is extremely important in many fields of endeavour. For example, if we were able to anticipate meteorological disasters, or financial crashes, ruins and insolvency problems, we could avoid the often extremely high human and economic costs associated with such events.

Records are often studied in the context of extreme value theory (see for instance Galambos, 1978; Resnick, 1987), since a record is essentially a temporary maximum or minimum in an ever-increasing sequence of collected samples of data. As time goes on, new maxima and minima do appear in the series of measurements of any variable. If we denote by $X_{i:i}$ and $X_{1:i}$, respectively, the maximum and the minimum up to time i in the time series $\{X_1; X_2; \dots; X_i; \dots; g\}$, that is,

$$X_{i:i} = \max \{X_1; \dots; X_i\};$$

$$X_{1:i} = \min \{X_1; \dots; X_i\};$$

we will have a new record at time $i + 1$ if either $X_{i+1} > X_{i:i}$ or $X_{i+1} < X_{1:i}$: In the former case, we have an upper record, whereas in the latter we have a lower record. Upper records produce jumps in the sequence of relative maxima, $\{X_{i:i}g_i\}$, while lower records do it in the sequence of relative minima, $\{X_{1:i}g_i\}$: Since the lower records in a time series $\{X_1; \dots; X_i\}$ are the upper records of its "specular

Date: October 31, 2003.

Key words and phrases. Records, extreme order statistics, counting processes, renewal theory, geometric distribution, integrated time series, tests of unit roots.

image" around the horizontal axis, that is of $f_i X_1; \dots; X_i g$, we can often restrict our attention to the analysis of upper records.

Whatever the type of records, they result in jumps in the monotonically increasing sequence of ranges, $fR_i g$; where $R_i = X_{i;i} - X_{1;i}$. In other words, a record (either upper or lower) has appeared at time i if $\Phi R_i = R_i - R_{i-1} > 0$:

The time instants at which records occur are commonly called record or ladder times. They will be denoted here by T_i^+ for the i -th upper record time ("ascending ladder time"), and by T_i^- for the i -th lower record time ("descending ladder time"). Formally,

$$T_i^+ = \inf_{k > T_{i-1}^+} X_k > X_{T_{i-1}^+};$$

with $T_0^+ = 0; X_0 = 0$:

The time differences $\zeta_i^+ = T_i^+ - T_{i-1}^+$; and $\zeta_i^- = T_i^- - T_{i-1}^-$, for $i > 0$, where $T_0^+ = T_0^- = 0$; are called the inter-arrival record (upper and lower, respectively) times, or simply, inter-record times.

An important process related to records is the record counting process, denoted as the process $N = fN(t); t \geq 0g$ where $N(t)$ is given by:

$$(1.1) \quad N(t) = \sum_{i=1}^{\infty} 1(\Phi R_i > 0); \quad \text{with } R_0 = 0:$$

Depending on the type of records that are counted, we may distinguish the upper and lower record counting processes, that is $N^+ = fN^+(t); t \geq 0g$ and $N^- = fN^-(t); t \geq 0g$, respectively, where:

$$(1.2) \quad N^+(t) = \sum_{i=1}^{\infty} 1(X_i > X_{i-1});$$

$$(1.3) \quad N^-(t) = \sum_{i=1}^{\infty} 1(X_i < X_{i-1});$$

with $X_{1;1} = X_1; X_{0;0} = X_{1;0} = 0$:

Record counting processes are related renewal counting processes, where records are here referred to as renewals. A renewal counting process $fN(t); t \geq 1g$ is defined as:

$$N(t) = \sup \{i \geq 1 : T_i \leq t\}; \quad t \geq 1;$$

where $T_i = \zeta_1 + \zeta_2 + \dots + \zeta_i$ denotes the instant at which the i -th renewal occurs; and $f\zeta_i g$ is a sequence of i -i.i.d. non-negative random variables representing the inter-renewal times.

A renewal process is often an appropriate model for the successive occurrences of a particular event (renewals). These processes have often been used to model phenomena such as the successive failures of a machinery, the repeated incidence of earthquakes, the emission of radioactive particles, customers and computer arrivals at a server queue, the number of contingent claims received by an insurance company, etc.

Most of the available results in the analysis of records for time series rely on the assumption that the inter-record times are i -i.i.d. random variables with finite mean and variance, or when the underlying time series $fX_1; \dots; X_t; \dots g$ is a sequence of

i.i.d. random variables (see for instance Chandler, 1952; Resnick, 1973; Shorrock, 1973; Glick, 1978; Ballerini and Resnick, 1985; Nagaraja, 1988). Andel (2001) and Arnold, Balakrishnan and Nagaraja (1992) provide a nice survey on the subject for the i.i.d. case.

Fewer results have been reported on the behavior of records in scenarios involving heterogeneity or serial dependence (see Lindgren and Rootzén, 1987; Leadbetter and Rootzén, 1988; Arnold and Balakrishnan, 1989 -pp. 146-149-; Leadbetter and Nandagopalan, 1989). Such results point that, under conditions which restrict the amount of serial dependence and the clustering of extremes, the limit joint distributions of stationary time series are the same as for sequences of i.i.d. variables (Lindgren and Rootzén, 1987). For stationary Gaussian series these conditions can be replaced by a simpler one known as the "Berman's condition", which states that

$$(1.4) \quad \lim_{k \rightarrow \infty} \frac{\text{Cov}(X_i; X_{i+k})}{k} \log k = 0:$$

This condition is satisfied by a wide class of stationary time series, among which are those any whose covariance function decreases exponentially fast with k , as it is the case for all stationary Gaussian ARMA processes.

In time series analysis, one is often concerned with the invariance, or at least the robustness, of test statistics in the face of unknown transformations affecting the variables. For example, the performance of standard unit-root and cointegration testing methods can severely worsen when the variables are nonlinearly transformed. There is a vast literature reporting on such problems (for a favor of it see Granger and Hallman, 1991; Granger and Teräsvirta, 1993; Aparicio and Granger, 1995; Granger, 1995; Aparicio and Escribano, 1998; Breitung and Gouriéroux, 1997; Breitung, 1998; Franses and McAleer, 1998; Aparicio, Escribano and Garcia, 2000a,b).

One of the most interesting properties of record-related statistics is their invariance to monotone transformations. Formally, if $T_i; \zeta_i$ and $N(t)$ are record statistics associated with a time series process $X = fX_t; t \geq 1$, then if we construct a new process $Y = fY_t; t \geq 1$ through a monotone transformation \hat{A} in such a way that $Y_t = \hat{A}(X_t)$ for each $t \geq 1$; the corresponding record statistics $T_i^0; \zeta_i^0$ and $N^0(t)$ will have the same distribution as $T_i; \zeta_i$ and $N(t)$; respectively.

But some record statistics seem also to be robust in the face of other types of phenomena, such as structural breaks and outliers (Aparicio, Escribano and Garcia, 2000a,b). These properties suggest using record statistics for constructing robust procedures to test different aspects of a time series.

An important question is whether the statistical behavior of records can be exploited for testing serial dependence, and in particular, whether the statistical properties of records from sequences of i.i.d. random variables differ or not from those of the important class of nonstationary time series called integrated time series "of order 1", that is, time series with a single unit root.

A time series process $X = fX_t; t \geq 1$ is said to be integrated of order d , or $I(d)$ (with d being a non-negative integer), if the sequence of d -th differences of X is stationary. The integer d is called the order of integration of X , and we write briefly $X \gg I(d)$: When $d = 0$ the series is supposed to be stationary and have finite variance. With $d = 1$ the series has explosive variance and can fluctuate wildly around a constant mean level or around a deterministic trend. An $I(1)$ time series can be represented by an autoregressive model with a unit root:

$$X_t = X_{t-1} + \varepsilon_t;$$

where the model errors, ε_t , will be here assumed to be a sequence of i.i.d. random variables. Depending on whether $\mu = E(\varepsilon_t)$ is nonzero or equal to zero, X will fluctuate around a linear trend or around 0: As we will show later, the records of unit root time series behave quite differently when $\mu = 0$ and when $\mu \neq 0$:

The purpose of this paper is to analyse and compare the behavior of record counting processes and related quantities for i.i.d. and for unit-root time series, in order to provide some theoretical support for using record-based test statistics in the unit-root testing problem. The structure of the paper is as follows. In Section 2 we review the statistical behavior of records of i.i.d. sequences of random variables, while in Section 3, we present some results for the behavior of records from $I(1)$ time series. First, we discuss some classical results for the case of trending $I(1)$ time series, where $\mu \neq 0$: Then we analyse the class of non-trending $I(1)$ time series, that is with $\mu = 0$, and compare our findings with those obtained in the other cases: In particular, we show that both the rates of convergence of the associated counting processes and their limiting distributions differ greatly from those of either i.i.d. or trending $I(1)$ time series. Finally, we discuss the implications of our findings for unit root testing. After the conclusions in Section 4, we devote an Appendix to proving the new results.

2. The record properties of i.i.d. time series

A lot of questions can be answered about records when the time series is completely random. Suppose $\{X_t\}_{t=1}^n$ is an i.i.d. sequence of random variables having a continuous probability distribution function $F_X(\cdot)$, and suppose that we want to compute the probability that at a given time instant, say t , we have an upper record in the series. As we pointed earlier, lower records are upper records of the sign-inverted series, and thus the subsequent discussion will focus on the behavior of upper record properties.

In this case it is straightforward to see that

$$(2.1) \quad P(X_t > X_{t_1}, \dots, X_{t_{i-1}}) = \frac{1}{t};$$

and that the probability of n successive (upper) records in the sequence $\{X_1; X_2; \dots; X_n\}$ is given by:

$$(2.2) \quad P(N^+(n) = n) = \frac{1}{n!};$$

Using Stirling's approximation, we can write $n! \approx n^{n+1/2} e^{-n}$: Thus

$$(2.3) \quad P(N^+(n) = n) \approx \frac{1}{n^{n+1/2} e^{-n}} = \frac{e^n}{n^{n+1/2}};$$

Remark that the probability of getting a new record in the series decreases to zero as time grows to infinity more than exponentially fast, and that whatever the type of process, the distribution of $N^+(t)$ has finite support for finite t since $P(N^+(t) > t) = 0$: Similarly, it can also be shown (Andel, Theorem 4.1, p. 67,

2001) that the probability that there is $n \cdot t$ (upper) records in $(0; t]$ can be expressed recursively:

$$PfN^+(t) = ng = \frac{t-1}{t} PfN^+(t-1) = ng + \frac{1}{t} PfN^+(t-1) = n \cdot \frac{t-1}{t} g;$$

with boundary condition:

$$PfN^+(1) = 1g = 1 \text{ and } PfN^+(0) = ng = 0;$$

For t large and $n \ll t$; we obtain the approximation (see for instance Andel, p. 70, 2001):

$$(2.4) \quad PfN^+(t) = ng \gg \frac{1}{(n-1)!t} (\ln t)^{n-1};$$

On the other hand, by Euler's formula:

$$(2.5) \quad E N^+(t) = \sum_{i=1}^t P(X_i > X_{i-1}; i-1) = \sum_{i=1}^t \frac{1}{i} = O(\ln t);$$

It follows that the long-run probability of a new record, given by $\lim_{t \rightarrow \infty} \frac{1}{t} E fN^+(t)g$, is equal to zero, thus implying that no new record value will appear in the series in the long run: Similarly, we have for the variance:

$$(2.6) \quad \text{Var} N^+(t) = \sum_{i=1}^t \frac{1}{i} \frac{1}{i^2} = O(\ln t);$$

Also following Embrechts, Klüppelberg and Mikosch (1999, pp. 257-8) and Arnold, Balakrishnan and Nagaraja (1992, p. 247-8) the point process of records $N^+(t)$ obeys a Strong Law of Large Numbers (SLLN) and a Central Limit Theorem (CLT):

$$(2.7) \quad \lim_{t \rightarrow \infty} (\ln t)^{i-1} N^+(t) = 1 \text{ a.s. (SLLN),}$$

$$(2.8) \quad (\ln t)^{i-2} N^+(t) \xrightarrow{d} N(0; 1); \text{ (CLT),}$$

where " \xrightarrow{d} " here denotes weak convergence.

As for the record times, T_n^+ , we have the following limit results:

$$(2.9) \quad \lim_{n \rightarrow \infty} n^{i-1} \ln T_n^+ = 1 \text{ a.s. (SLLN)}$$

$$(2.10) \quad n^{i-2} (\ln T_n^+ - i \ln n) \xrightarrow{d} N(0; 1); \text{ (CLT),}$$

which imply that $T_n^+ = O(e^n)$ and that $\zeta_n^+ = T_n^+ - T_{n-1}^+ = O(e^n)$: That is the distance between successive maxima grows exponentially with time, and consequently $E(T_n) = E(\zeta_n^+) = 1 + 8n$:

Among the upper (lower) record times, the time occurrence of the last maximum (minimum) is often of interest. Let L_n^+ (L_n^-) denote the time of the last maximum (respectively, minimum) in $fX_1; X_2; \dots; X_n g$: That is,

$$L_n^+ = \min_{k \leq n} \{X_k = X_n\};$$

$$L_n^- = \min_{k \leq n} \{X_k = X_1\};$$

Using a straightforward counting argument,

$$(2.11) \quad P(L_n^+ = t) = \frac{(t-1)!(n-t)!}{n!} = \frac{t}{n} \frac{1}{t};$$

It follows for the expected value of L_n^+ :

$$\begin{aligned}
 E L_n^+ &= \sum_{t=1}^n P(L_n^+ \geq t) = \sum_{t=1}^n \mu_n^{[t-1]} \\
 &= \frac{1}{n} + \frac{2}{n(n-1)} + \frac{6}{n(n-1)(n-2)} + \dots + 1 \\
 &< \frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \dots + 1; \text{ for large enough } n \\
 (2.12) \quad &= O(\ln n); \text{ as } n \rightarrow \infty
 \end{aligned}$$

Proceeding similarly for the second-order moment we obtain:

$$\begin{aligned}
 E L_n^{+2} &= \sum_{t=1}^n t P(L_n^+ \geq t) \\
 &= \sum_{t=1}^n t \mu_n^{[t-1]} \\
 &= \frac{1}{n} + \frac{4}{n(n-1)} + \frac{18}{n(n-1)(n-2)} + \dots + n \\
 (2.13) \quad &= O(n), \text{ as } n \rightarrow \infty
 \end{aligned}$$

from what follows:

$$(2.14) \quad \text{Var}(L_n) = E(L_n^2) - [E(L_n)]^2 = O(n); \text{ as } n \rightarrow \infty.$$

Thus the average position of the last maximum grows more slowly than n , while the size of its fluctuations around this mean, grow at the faster rate of $n^{1/2}$.

3. The records of I(1) time series

Time series that are completely random exclude most real life data sets. Kotz and Nadarajah (2000, p. 60) point that for m -dependent time series (that is, series whose variables are independent if they are separated by more than m units of time) one can expect record values to occur in runs or clusters. Notwithstanding the previous results obtained for i.i.d. sequences apply to the records of stationary time series with some degree of serial dependence. If X is a stationary and Gaussian process then the joint distribution of any fixed set of its extreme order statistics converges to the same limit as if it was a sequence of i.i.d. random variables (see Lindgren and Rootzén, 1987; Leadbetter and Rootzén, 1988). Therefore one need not be too concerned about possible dependencies as long as the time series is Gaussian. Unfortunately, nonlinear transformation of Gaussian variables have non-Gaussian distributions, and thereby the Berman condition may be inappropriate in such cases. We refer the interested reader to Lindgren and Rootzén (1987) and to Leadbetter and Rootzén (1988) for the more general conditions to be imposed on stationary series to guarantee that their records behave as in the i.i.d. case

In this section, we present some of the statistical properties of records from I(1) time series, and show that these properties differ greatly from those of records from sequences of i.i.d. random variables or from Gaussian stationary time series satisfying the Berman condition. As a result, we show that it is possible to use record-based statistics for the robust testing of unit roots.

3.1. Generalities. Let us define the upper records of X as the sequence $S_i^+ = X_{T_i^+}$ where $S_i^+ = X_{T_i^+}$; with T_i^+ representing the i -th (upper) record time, that is the smallest time instant verifying that $X_{T_i^+} > \max(S_1^+; \dots; S_{i-1}^+)$; and $S_1^+ = X_1$. Therefore S_i^+ is the sequence of successive maxima of X_t . Similarly, the lower records of X is the sequence S_i^- whose terms are given by $S_i^- = X_{T_i^-}$; where T_i^- is the i -th lower record, that is the smallest time instant verifying that $X_{T_i^-} < \min(S_1^-; \dots; S_{i-1}^-)$; and $S_1^- = X_1$. Thus S_i^- is the sequence of successive minima of X_t . As we mentioned in a previous section, lower records are upper records of the sequence $-X_1; \dots; -X_{i-1}; \dots$. Thus it is enough to study the statistical properties of upper records.

The time instants T_i^+ are sometimes called ladder indexes, and the difference $T_i^+ - T_{i-1}^+$ represents the waiting time until the arrival of the new (upper) record value S_i^+ at $t = T_i^+$. The pair $(T_i^+; S_i^+)$ is called a ladder point. The sections between ladder points are probabilistic replicas of each other. Therefore, from a statistical point of view, it suffices to study the properties of the first non-trivial ladder point, that is the point $(T_2^+; S_2^+)$ for which $S_2^+ > X_1$. For the sake of notational homogeneity, we will borrow the terminology used in Feller (1971) and write $\zeta_1^+ = T_2^+$; and $H_1^+ = S_2^+$:

First of all, notice that the event $\zeta_1^+ = k$ is the same event as

$$fX_2 \cdot X_1; X_3 \cdot X_1; \dots; X_{k+1} > X_1g:$$

The random variables ζ_1^+ and H_1^+ are called the first ladder epoch and the first ladder height, respectively. Their joint distribution function can be written as $D_k(x) = P(\zeta_1^+ = k; H_1^+ \leq x)$ which yields:

$$(3.1) \quad P(\zeta_1^+ = k) = D_k(1);$$

$$(3.2) \quad G(t) = P(\zeta_1^+ \leq t) = \sum_{k=1}^{\infty} D_k(1);$$

$$(3.3) \quad P(H_1^+ \leq x) = \sum_{k=1}^{\infty} D_k(x) = D(x);$$

The pair $(\zeta_1^+ + \zeta_2^+ + \dots + \zeta_i^+; H_1^+ + H_2^+ + \dots + H_i^+)$ represents the i -th ladder point of the time series X_t . If X_t is a symmetric random walk then for $i \neq j$, $(\zeta_i^+; H_i^+)$ and $(\zeta_j^+; H_j^+)$ are jointly independent random variables. Moreover as both ζ_i^+ and H_i^+ are positive, the random sequences defined as:

$$(3.4) \quad T_{i+1}^+ = \zeta_1^+ + \zeta_2^+ + \dots + \zeta_i^+$$

$$(3.5) \quad S_{i+1}^+ = H_1^+ + H_2^+ + \dots + H_i^+$$

are renewal processes. The vector sequence $(\zeta_1^+ + \zeta_2^+ + \dots + \zeta_i^+; H_1^+ + H_2^+ + \dots + H_i^+)^0$ is therefore a two-dimensional renewal process. From our definition of $N^+(t)$, we have

$$(3.6) \quad N^+(t) = \sum_{i=1}^{\infty} 1(\zeta_1^+ + \zeta_2^+ + \dots + \zeta_i^+ \leq t);$$

and

$$(3.7) \quad E[N^+(t)] = \sum_{n=1}^{\infty} P(\zeta_1 + \zeta_2 + \dots + \zeta_n \leq t) = \sum_{n=1}^{\infty} G^{*n}(t);$$

where $G^{*n}(t)$ denotes the n -fold convolution of $G(\cdot)$:

3.2. I(1) time series with $\mu \neq 0$. When $\mu > 0$ then X exhibits a linear drift (to ∞ if $\mu > 0$ and to $-\infty$ if $\mu < 0$). As a consequence, if $\mu > 0$ we have:

$$(3.8) \quad \lim_{t \rightarrow \infty} t^{-1} E f N^+(t) g > 0; \quad E(\zeta_i^+) < 1;$$

$$(3.9) \quad \lim_{t \rightarrow \infty} t^{-1} E f N^i(t) g = 0; \quad E(\zeta_i^i) < 1;$$

$$(3.10) \quad G^+(1) = P(\zeta_i^+ < 1) = 1 = P(\zeta_i^i < 1) = G^i(1);$$

In this case, $N^+ = f N^+(t)$; $t \rightarrow \infty$ g is an ordinary renewal counting process. Similarly, when $\mu < 0$, we obtain:

$$(3.11) \quad \lim_{t \rightarrow \infty} t^{-1} E f N^i(t) g > 0; \quad E(\zeta_i^i) < 1;$$

$$(3.12) \quad \lim_{t \rightarrow \infty} t^{-1} E f N^+(t) g = 0; \quad E(\zeta_i^+) < 1;$$

$$(3.13) \quad G^+(1) = P(\zeta_i^+ < 1) = 1 = P(\zeta_i^i < 1) = G^i(1);$$

and thus $N^i = f N^i(t)$; $t \rightarrow \infty$ g defines also an ordinary renewal counting process. Although the superposition of renewal processes is not, in general, a renewal process, whatever the sign of the innovations' mean μ , $N^\mu(t) = N^+(t) + N^i(t)$ will be an ordinary renewal process satisfying:

$$(3.14) \quad \lim_{t \rightarrow \infty} t^{-1} E f N^\mu(t) g > 0;$$

$$(3.15) \quad E(\zeta_i^\mu) < 1;$$

$$(3.16) \quad G(1) = 1;$$

where ζ_i^μ represent the time elapsed between the i th and the $(i-1)$ th record (either upper or lower).

As the analysis is identical when $\mu > 0$ and when $\mu < 0$; we may assume in the sequel and without loss of generality that $\mu > 0$: Therefore we have for the upper record counting process $N^+(t)$ the following results from Feller (1971, pp.396-7):

$$(3.17) \quad \lim_{t \rightarrow \infty} t^{-1} N^+(t) = \frac{1}{\mu} > 0;$$

$$(3.18) \quad \frac{\mu}{t^{1-2\mu/\mu}} N^+(t) \Big|_{\frac{t}{\mu}} \Big) N(0; 1);$$

where $\mu = E(\zeta_i^+) < 1$: And since $\lim_{t \rightarrow \infty} t^{-1} N^\mu(t) = N^+(t) = 1$ with probability one; we can also write:

$$(3.19) \quad \lim_{t \rightarrow \infty} t^{-1} E f N^\mu(t) g = \frac{1}{\mu} > 0;$$

$$(3.20) \quad \frac{\mu}{t^{1-2\mu/\mu}} N^\mu(t) \Big|_{\frac{t}{\mu}} \Big) N(0; 1);$$

3.3. $I(1)$ time series with $\mu_2 = 0$. In this case we have from Loève (1987, vol. I, p. 384):

$$X_{i+1} < X_i \text{ with probability one.}$$

This latter result has important implications as regards the distribution of either $N^+(t)$ or $N^i(t)$, since we cannot invoke asymptotic results, available for ordinary renewal processes, that rely on the finiteness of the mean inter-record time, that is when $\mu_2 = E(\zeta_i) < \infty$. Instances of such standard results are the basic renewal theorem, which states that $M(t) - E f N(t) g = O(t^{-1})$ for t large, or Blackwell's theorem, according to which $M(t) - M(t - 1) \sim \mu_2^{-1}$. In fact, for the symmetric random walk case we have $E f \zeta_1 g = 1$; $G(1) = 1$ (see Feller, vol. II, 1971, p. 395); and $\lim_{t \rightarrow \infty} t^{-1} M(t) = 0$ (see Bosq and Nguyen, 1996, p. 152). or the $I(1)$. In Theorem 1 in the Appendix it is shown that, under very general conditions, $M(t) \gg t^{1-2}$ for t large, and

$$(3.21) \quad t^{1-2} N^+(t) \sim \frac{E f \zeta_1^2 g}{\sqrt{2\pi}} Z_1;$$

$$(3.22) \quad t^{1-2} N^i(t) \sim \frac{E f \zeta_1^2 g}{\sqrt{2\pi}} Z_2;$$

where Z_1 and Z_2 are two independent standard Gaussian random variables. It follows that the long-run probability of a new record goes to zero faster for time series of i.i.d. or Gaussian stationary random variables than for $I(1)$ series. The different behavior can be summarised as follows:

$$(3.23) \quad \frac{N(t)}{t} = O(t^{-1} \ln t); \text{ for large } t \text{ and for sequences of i.i.d. random variables,}$$

$$(3.24) \quad \frac{N(t)}{t} = O(t^{-1-2}); \text{ for large } t \text{ and for } I(1) \text{ time series with } \mu_2 = 0,$$

$$(3.25) \quad \frac{N(t)}{t} = O(1); \text{ for large } t \text{ and for } I(1) \text{ time series with } \mu_2 \neq 0;$$

On the other hand, Theorem 2 in the Appendix shows that the associated reliability or survivor function of either record counting process (upper or lower) is given by

$$(3.26) \quad P f N(t) > ng = G(t)^n \frac{G(t)}{M(t)};$$

with $G(t) = P(\zeta_1 < t)$;

where by $N(t)$ and ζ_1 we refer indistinctly to either $N^+(t)$ or $N^i(t)$, and to either ζ_1^+ or ζ_1^i , respectively. When the possibility of infinite inter-record times is allowed; that is when $G(1) < 1$; the inter-record times ζ_i have not a proper probability distribution; and N becomes a terminating renewal process, following Karlin and Taylor (1975, pp. 204-5). In this case, the total number of records, say

$N(1)$; and thereby their expected value, $M(1)$, will be finite. In fact, since

$$\begin{aligned}
 M(1) &= E fN(1)g = \sum_{n=0}^{\infty} P fN(1) > ng \\
 &= \sum_{n=0}^{\infty} P fT_n < 1g = \sum_{n=0}^{\infty} \prod_{i=1}^n P f\zeta_i < 1g \\
 (3.27) \quad &= \sum_{n=0}^{\infty} G^n(1) = \frac{G(1)}{1 - G(1)};
 \end{aligned}$$

The finiteness of $N(1)$ also implies that, for t large enough, the distribution of the record counting process $N(t)$ could be well approximated by a geometric probability law. Note that the random walk with $\zeta_2 = 0$ is not in this category since although $E(\zeta_1) = 1$ we still $G(1) = 1$: The convergence of $G(t)$ towards its asymptotic value must be, however, very slow so that the mean inter-record time diverges to infinity.

The different rates of growth of $N(t)$ allows a classification of the time series that we are considering in terms of their probability distribution tails. Indeed, a first-order Taylor series approximation leads to

$$(3.28) \quad G(t) \sim 1 - \frac{G(t) - G(1)}{M(t)} \approx 1 - \frac{nG(t) - nG(1)}{M(t)};$$

which in turn allows approximating the tail behavior for the component record counting processes when $n \ll t$ as

$$(3.29) \quad P fN(t) > ng \approx 1 - \frac{n}{\ln t}; \quad \text{for sequences of i.i.d. random variables,}$$

$$(3.30) \quad P fN(t) > ng \approx 1 - \frac{n}{t}; \quad \text{for I(1) time series with } \zeta_2 = 0,$$

$$(3.31) \quad P fN(t) > ng \approx 1 - \frac{n}{t}; \quad \text{for I(1) time series with } \zeta_2 \neq 0.$$

It follows from these results that the records of an I(1) time series tend to quickly outgrow in number those from sequences of i.i.d. random variables.

The particular form of the reliability function for $N(t)$ suggests that its probability distribution is related to a geometric distribution with parameter ("success probability") $p(t) = G(t) - M(t)$: Indeed, we have for the mass probability function of $N(t)$:

$$\begin{aligned}
 P fN(t) = ng &= P fN(t) > n - 1g - P fN(t) > ng \\
 &= G(t)p(t)[1 - p(t)]^{n-1} \\
 (3.32) \quad &= P fZ(t) = ngG(t);
 \end{aligned}$$

where $Z(t)$ represents a random variable with geometric distribution having $p(t)$ as a time-varying parameter: The previous equality allows an interpretation of $p(t)$ as the conditional probability:

$$(3.33) \quad p(t) = \frac{P fN(t) = ng}{P fN(t) \geq ng} = P fN(t) = njN(t) \geq ng;$$

which represents the hazard rate of $N(t)$, and gives us roughly the probability of having exactly n records at t when at least n records are known to have occurred.

Note that $p(t)$ is independent of n as far as $0 < n < t$, and that in the especial cases of $t = 0$ and $t = 1$; we have:

$$(3.34) \quad p(0) = \frac{P fN(t) = 0g}{P fN(t) \leq 0g} = 0;$$

$$(3.35) \quad p(1) = \lim_{t \rightarrow 1} p(t) = \lim_{t \rightarrow 1} \frac{P fN(t) = 1g}{P fN(t) \leq 1g} = 1; \quad G(1) = 0;$$

Finally, remark that since

$$P fN(t) = ng = P fZ(t) = ng P f_{\zeta_1} < tg;$$

the event $fN(t) = ng$ can be seen as the intersection of two mutually exclusive events:

$$fN(t) = ng = fZ(t) = ng \setminus f_{\zeta_1} < tg;$$

It follows that conditioned on $f_{\zeta_1} < tg$ (or equivalently, for t large enough), $N(t)$ behaves as a geometric random variable with parameter $p(t)$: Further, since

$$(3.36) \quad E fZ(t)g = \frac{1}{p(t)};$$

$$(3.37) \quad \text{Var} fZ(t)g = \frac{1 - p(t)}{p^2(t)};$$

we obtain:

$$(3.38) \quad \begin{aligned} E fN(t)g &= \sum_{n=0}^t n P fN^+(t) = ng \\ &= G(t) \sum_{n=0}^t np(t)[1 - p(t)]^{n-1} \\ &= G(t) E fZ(t)g = \frac{G(t)}{p(t)} = M(t); \quad \text{as expected.} \end{aligned}$$

and

$$(3.39) \quad \begin{aligned} \text{Var} fN(t)g &= \sum_{n=0}^t [n - M(t)]^2 P fN^+(t) = ng \\ &= G(t) \sum_{n=0}^t n^2 P fZ(t) = ng + G(t) M^2(t) \sum_{n=0}^t P fZ(t) = ng + 2G(t) M(t) E fZ(t)g \\ &= \frac{G(t)[2 - p(t)]}{p^2(t)} + G(t) M^2(t) \frac{1 - [1 - p(t)]^{t+1}}{[1 - p(t)]^{t+1}} - 2 \frac{G(t) M(t)}{p(t)} \\ &= \frac{M^2(t)}{G(t)} + G(t) M^2(t) - 2M^2(t); \quad \text{for large } t, \\ &= \frac{M^2(t)}{G(t)} [1 + [G(t) - 1]^2]; \quad \text{as } t \rightarrow \infty; \\ &= M^2(t); \quad \text{as } t \rightarrow \infty; \end{aligned}$$

Notice that if the probability distribution function of Z_i is symmetric around its zero mean, that is if $F_z(0) = 1/2$; then the probability of n successive upper records

for an $I(1)$ time series with $\mu_2 = 0$ must be 2^{i-n} : Such probability is much larger than in the $i:i:d$ case, for which we obtained:

$$(3.40) \quad P f N^+(n) = n g = \frac{1}{n!} \gg n^{i-n} e^n = \frac{n^i}{e} = o(2^{i-n});$$

Therefore long runs of records tend to occur much more often in $I(1)$ time series than in sequences of $i:i:d$ random variables. As we pointed in a previous session, the clustering of record occurrences is a most remarkable consequence of serial dependence.

At this point, it may be also interesting to compare the asymptotic statistical behavior of the time of arrival of the last maximum (or minimum) in non-trending $I(1)$ series with its behavior in sequences of $i:i:d$ random variables, as studied in a previous section. Under the assumption that the model errors $f^2_i g_i$ are exchangeable random variables (a weaker requirement than the $i:i:d$ assumption), the Andersen equivalence lemma (Loève, vol. I, 1987, pp. 378-9) establishes that L_n has the same distribution as the random variable \circ_n , where \circ_n represents the number of positive terms in $fX_1; X_2; \dots; X_n g$. In addition, if the model errors $f^2_i g_i$ are $i:i:d$ and their distribution probability function is symmetric around zero then we have for any $x \in [0, 1]$:

$$(3.41) \quad \lim_{n \rightarrow \infty} P^{i-n} \circ_n < x = \frac{2}{\pi} \text{Arcsin} \sqrt{\frac{x}{2}};$$

Consequently, for $I(1)$ series with zero-mean $i:i:d$ model errors, \circ_n is a non-degenerate random variable. Thus $\circ_n = O(n)$; and also $L_n = O(n)$: In particular, we have for the average position of the last maximum (or minimum) in $fX_1; X_2; \dots; X_n g$:

$$(3.42) \quad E(L_n) = \frac{n}{\pi} \int_0^1 \frac{x}{x(1-x)} dx = \frac{n}{2};$$

But since the arcsine density probability function, given by

$$(3.43) \quad \frac{1}{\pi} \frac{1}{y(1-y)} = f_{\circ_n}(y)$$

is U-shaped, the random variable \circ_n is more likely to take values near 0 or n than near $n/2$. Moreover, the size of the fluctuations of L_n around this mean will also increase as n . Therefore both the mean and variance of L_n grow faster with increasing n than the corresponding quantities in the $i:i:d$ case. And this occurs in spite that the long-run probability of a new record, that is $n^{-1}M(n)$; goes to zero as n grows to infinity:

3.4. Joint statistical properties of the upper and lower record counting processes. So far we have discussed the univariate statistical behavior of both upper and lower records for $I(1)$ series with $\mu_2 = 0$: In order to further investigate the distribution of the total number of records up to time t , that is of $N^a(t) = N^+(t) + N^-(t)$, define for any non-negative integers n and m , the conditional mass probability function of $N^+(t)$ given $N^-(t)$ as

$$(3.44) \quad \mathcal{N}_t(n; m) = P f N^+(t) = n j N^-(t) = m g;$$

where it is obvious that $\mathbb{1}_t(0; 0) = 0$ for any $t \geq 1$. Now from the total probability theorem we can write:

$$\begin{aligned}
 P fN^+(t) + N^i(t) &= \sum_{m=0}^{\infty} \mathbb{1}_t(k_j - m; m) P fN^i(t) = mg \\
 &= G(t) \sum_{m=0}^{\infty} \mathbb{1}_t(k_j - m; m) P fZ(t) = mg \\
 (3.45) \quad &= \sum_{m=0}^{\infty} \mathbb{1}_{t;k}(m) P fZ(t) = mg;
 \end{aligned}$$

which could be interpreted as a "smoothed geometric probability mass function", with smoothing kernel $\mathbb{1}_{t;k}(m) = G(t)\mathbb{1}_t(k_j - m; m)$. Since

$$\begin{aligned}
 \sum_{m=0}^{\infty} P fN^i(t) = n; N^+(t) = mg = P fN^i(t) = ng; \quad \forall t \geq n \\
) \quad \sum_{m=0}^{\infty} \frac{P(fN^i(t) = n; N^+(t) = mg)}{P(fN^i(t) = ng)} = 1 < 1;
 \end{aligned}$$

we must conclude, from the Borel-Cantelli lemma (see for example Shiryaev, 1996, p. 255), that

$$(3.46) \quad \mathbb{1}_t(n; m) = \frac{P(fN^i(t) = n; N^+(t) = mg)}{P(fN^i(t) = ng)} \neq 0; \text{ as } t \rightarrow \infty;$$

from what follows an important result: that $N^+(t)$ and $N^i(t)$ are asymptotically independent. This allows to derive the limiting distribution of $t^{1-2}N^{\pm}(t)$ as the convolution of the limiting distributions of $t^{1-2}N^+(t)$ and $t^{1-2}N^i(t)$; which from Theorem 1 is given by:

$$(3.47) \quad P f \lim_{t \rightarrow \infty} t^{1-2} N^{\pm}(t) < hg = \frac{1}{2^{1/4}} \int_0^h \exp(-i \frac{v^2+2}{4}) [1 - \Phi(v)] dv;$$

where $\Phi(\cdot)$ is the probability distribution function of a standard Normal random variable. On the other hand, by the Tchebyshev theorem, we have for the upper and lower record counting processes, simultaneously referred to as $N = fN(t); t \geq 1g$, that for any positive real number \pm :

$$(3.48) \quad P f t^{1-j} N(t) - M(t) j \geq \pm g \cdot \frac{Var fN(t)g}{\pm^2 t^2} \gg \frac{t}{t^2} \neq 0; \text{ as } t \rightarrow \infty;$$

This together with $E fN^+(t)g = E fN^+(t) = M(t)$; obtains:

$$(3.49) \quad \lim_{t \rightarrow \infty} \frac{N^+(t)}{N^i(t)} = 1; \text{ with probability one;}$$

and similarly:

$$(3.50) \quad \lim_{t \rightarrow \infty} \frac{N^{\pm}(t)}{2N^+(t)} = \lim_{t \rightarrow \infty} \frac{N^{\pm}(t)}{2N^i(t)} = 1; \text{ with probability one.}$$

Therefore for any positive integer k and t large enough we can write the following approximation:

$$(3.51) \quad \begin{aligned} P\{N^+(t) > k\} &= P\left\{\sum_{i=1}^t \Phi N^+(i) > k\right\} \\ &= G(t) [1 - p(t)]^{hk-2i} \end{aligned}$$

3.5. Statistical properties of the record increments. Our upper and lower record count processes, $N^+(t)$ and $N^-(t)$, can be decomposed in terms of their increments $\Phi N^+(i)$ and $\Phi N^-(i)$ as:

$$(3.52) \quad N^+(t) = \sum_{i=1}^t \Phi N^+(i)$$

$$(3.53) \quad N^-(t) = \sum_{i=1}^t \Phi N^-(i);$$

where $\Phi N^+(i) = N^+(i) - N^+(i-1)$ and $\Phi N^-(i) = N^-(i) - N^-(i-1)$ are binary random variables taking the values 1 or 0, depending on whether a new record occurs a time i or not. Notice that even though $\Phi N^+(i)$ has the same distribution as $\Phi N^-(i)$, this distribution is different at different time instants i . Moreover,

$$(3.54) \quad \begin{aligned} P\{\Phi N^+(t) > 0\} &= P\{\Phi N^+(t) > 0\} + P\{\Phi N^-(t) > 0\} \\ &= P\{\Phi N^+(t) > 0\} = 1 - p(t) \end{aligned}$$

Since $\lim_{t \rightarrow \infty} t^{-1} N^+(t) = \lim_{t \rightarrow \infty} t^{-1} N^-(t) = 0$, both $P\{\Phi N^+(t) > 0\} = 1 - p(t)$ and $P\{\Phi N^-(t) > 0\} = 1 - p(t)$ must be decreasing in t . In fact:

$$(3.55) \quad \begin{aligned} P\{\Phi N^+(t) > 0\} &= 1 - p(t) = P\{\Phi N^+(t) > 0\} = 1 - p(t) = P\{\Phi N^-(t) > 0\} \\ &= E\{\Phi N^+(t)\} = E\{N^+(t) - N^+(t-1)\} \\ &= \lambda(t) \\ &\gg t^{-1-2} - (t-1)^{-1-2}; \quad \text{for large } t \\ &= \frac{1}{t^{1-2} + (t-1)^{1-2}} \gg t^{-1-2} \neq 0; \quad \text{as } t \rightarrow \infty \end{aligned}$$

The function $\lambda(t)$ represents a renewal intensity or density function, and vanishing limit means just that the probability of a new record become very small in the long run.

Theorem 3 in the Appendix shows that the increments $\Phi N^+(i)$ and $\Phi N^-(i)$ of both upper and lower record counting processes are neither independent nor even uncorrelated. In fact, they are positively correlated. On the other hand, the correlation between the sequences of increments $\Phi N^+(i)$ and $\Phi N^-(i)$ is negative, although it vanishes hyperbolically fast to zero as $i \rightarrow \infty$, or more precisely:

$$(3.56) \quad \text{Cov}\{\Phi N^+(i), \Phi N^-(i)\} = -\lambda(i) \gg -i^{-1-2}; \quad \text{for large } i:$$

This negative correlation between contemporaneous increments of the upper and lower counting processes entails a net negative correlation between $N^+(t)$ and $N^-(t)$ for finite t : The slowly decaying correlation between the increments $\Phi N^+(i)$ and $\Phi N^-(i)$ also suggests that a Normal approximation to the distribution of $N^+(t)$ cannot be a good one. Finally, recalling that the partial record counting processes

$N^+(t)$ and $N^i(t)$ are asymptotically (t^{-1}) independent, the fact that the distribution of $N^+(t) + N^i(t)$ cannot be Gaussian entails that neither can be the distribution of either component counting process, $N^+(t)$ or $N^i(t)$, as proved in Theorem 1.

3.6. Some implications for unit root testing in time series. For the purpose of testing and comparing certain properties of time series, it may be of interest to use record-based statistics. We have seen in the previous sections that the statistical properties of record counting processes differ importantly for a large class of stationary time series, and for both trending and non-trending time series with a unit root. In particular, we have seen that either $t^{1-2}N^{\square}(t)$, $t^{1-2}N^+(t)$ and $t^{1-2}N^i(t)$ converge to a non-degenerate random variable in the latter case, while they diverge for trending unit root time series, and vanish to zero for iid sequences and for all the Gaussian stationary series satisfying the Berman condition. This large spectra of behavior makes possible to use any of these statistics for discriminating between Gaussian stationary series (and possible non-Gaussian ones) from unit root time series. The testing of unit roots is an important problem in economics and finance where the optimal forecasting strategy may come to depend critically on whether a time series has or not a unit root.

In principle, any of the three statistics considered previously could be used for testing unit roots. However, $t^{1-2}N^{\square}(t)$ will offer the best small-sample performances, since it takes advantage of the negative correlation between $N^+(t)$ and $N^i(t)$. This produces, first, a gain in efficiency (which would double approximately that of either $t^{1-2}N^+(t)$ or $t^{1-2}N^i(t)$), and, second, a variance for $N^{\square}(t)$ smaller than the sum of the variances of $N^+(t)$ and $N^i(t)$: To see this, remark that as $N^+(t)$ and $N^i(t)$ are positive random variables, the dispersion of the distributions of either of the test statistics can be dimensionlessly measured by the coefficient of variation, CV. We can see that in the non-trending single unit-root case (null hypothesis of a unit root test) we have

$$(3.57) \quad CV_{t^{1-2}N^+(t)}^n \circ = CV_{t^{1-2}N^i(t)}^n \circ = \frac{\rho \overline{\text{varf}N^+(t)g}}{E\text{f}N^+(t)g};$$

while:

$$(3.58) \quad CV_{t^{1-2}N^{\square}(t)}^n \circ = \frac{\rho \overline{2\text{varf}N^+(t)g + 2\text{covf}N^+(t); N^i(t)g}}{2E\text{f}N^+(t)g} < \frac{\rho \overline{\text{varf}N^+(t)g}}{2E\text{f}N^+(t)g} < \frac{\rho \overline{\text{varf}N^+(t)g}}{E\text{f}N^+(t)g};$$

Thus

$$(3.59) \quad CV_{t^{1-2}N^{\square}(t)}^n \circ < CV_{t^{1-2}N^+(t)}^n \circ = CV_{t^{1-2}N^i(t)}^n \circ;$$

which means that the small-sample distribution of $t^{1-2}N^{\square}(t)$ is more concentrated around its mean than the distributions of either $t^{1-2}N^+(t)$ or $t^{1-2}N^i(t)$; thereby offering improved discrimination capabilities over the latter.

4. Conclusions

In this paper we have analysed the statistical properties of records from a class of nonstationary processes: the class of time series with a single unit root. We have

compared their records to those obtained from sequences of i.i.d. random variables, and have shown that these records behave quite differently in each case, exhibiting different limiting distributions as well as different convergence rates. and that for integrated time series their behavior depends crucially on the eventual existence of a drift, that is a nonzero mean value of the model errors or innovations. For time series with a single unit root, this behavior depends crucially on the eventual existence of a drift implied by a nonzero mean of the model errors. The presence of a drift leads to Gaussian asymptotic distributions for the standardised record counting processes, while its absence entails Non-Gaussian limits. This remarkable property of the standardized record counting processes suggest using them as statistics for the robust testing of certain hypotheses such as the presence of unit roots in a single time series (hypothesis of integration), or in the static regression errors of a pair of time series (hypothesis of non-cointegration).

5. Appendix

Theorem 1. Let $X_t = \sum_{i=1}^t \varepsilon_i$ where ε_i 's are continuous i.i.d. random variables with bounded and symmetric pdf, zero mean and finite variance σ^2 . Suppose that X_0 has also a bounded pdf and finite variance. And let $N^a(t) = N^+(t) + N^i(t)$ with $N^+(t) = \sum_{i=1}^t 1(X_i = X_{i-1})$ and $N^i(t) = \sum_{i=1}^t 1(X_i = X_{i+1})$. Then for there exist two standard Normal variables Z_1 and Z_2 such that:

$$\begin{aligned} t^{-1/2} N^+(t) &\xrightarrow{d} \frac{E f_j^2 \mathbb{1}_j}{\sigma} Z_1; \\ t^{-1/2} N^i(t) &\xrightarrow{d} \frac{E f_j^2 \mathbb{1}_j}{\sigma} Z_2; \end{aligned}$$

And since $N^+(1)$ and $N^i(1)$ are independent random variables:

$$P \left(\lim_{t \rightarrow \infty} t^{-1/2} N^a(t) < h \right) = \frac{1}{2} \int_0^h \frac{1}{\sigma} \exp\left(-\frac{v^2 + 2}{4}\right) [1 - \Phi(v)] dv;$$

where $\Phi(\cdot)$ is the probability distribution function of a standard Normal random variable.

Lemma 1. Let $X_t = X_{t-1} + \varepsilon_t$ where ε_t 's are i.i.d. random variables with zero mean and finite variance σ^2 ; and let

$$N^{(b)}(t) = \sum_{i=1}^t [1(X_{i-1} < b; X_i \leq b) + 1(X_{i-1} > b; X_i \leq b)]$$

denote the normalized number of crossings of level b . If X_0 and ε_1 have bounded pdf's with finite variance then we must have:

$$t^{-1/2} N^{(b)}(t) \xrightarrow{d} \frac{E f_j^2 \mathbb{1}_j}{\sigma} Z;$$

where Z is a standard Normal random variable.

Proof. (lemma). See Theorem 1 in Burrigge and Guerre (1996). \square

Proof. (theorem). Given that x_t is an $I(1)$, and noting that for this process a zero "crossing" amounts to a visit to the origin (crossing over the zero level is impossible), it follows from the previous lemma that \square

$$t_i^{1=2} N^+(t) \sim \frac{E f_{jZ_1}^2}{\sqrt{2}} j Z_{1j}$$

$$t_i^{1=2} N^-(t) \sim \frac{E f_{jZ_2}^2}{\sqrt{2}} j Z_{2j}$$

where Z_1 and Z_2 are independent standard Normal random variables.

Since the pdf of the absolute value of any standard Normal random variable Z_i is given by

$$f_{|Z_i|}(u) = \frac{1}{\sqrt{2}} \exp\left(-\frac{u^2}{2}\right); \quad u \geq 0;$$

we can easily obtain for the asymptotic pdf of both $t_i^{1=2} N^+(t)$ and $t_i^{1=2} N^-(t)$ respectively, the following expression:

$$f_{\pm}(u) = f_{|Z_i|}\left(\frac{u}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}} \exp\left(-\frac{u^2}{4}\right); \quad u \geq 0; \quad i = 1; 2;$$

Finally, as the random variables $N^+(1)$ and $N^-(1)$ are independent, the standard Normal variables Z_1 and Z_2 must also be independent. Now letting $f_{jZ_i}(\cdot)$ denote the pdf of the random variable jZ_i we obtain for asymptotic pdf of the standardized record counting process $t_i^{1=2} N^\pm(t)$:

$$\begin{aligned} f_{\pm}(v) &= \int_0^v f_{\pm}(w) f_{|Z_i|}(v-w) dw \\ &= a^{-2} \int_0^v f_{jZ_i}\left(\frac{w}{a}\right) f_{jZ_i}\left(\frac{v-w}{a}\right) dw; \quad \text{where } a = \frac{E f_{jZ_i}^2}{\sqrt{2}} \\ &= a^{-1} \int_0^v f_{jZ_i}(u) f_{jZ_i}\left(\frac{v}{a} - u\right) du; \quad \text{where we let } u = \frac{w}{a} \\ &= \int_0^v f_{jZ_i}(u) f_{jZ_i}(v-u) du \\ &= \frac{1}{\sqrt{2}} \exp\left(-\frac{v^2+2}{4}\right) [1 - \Phi(v)]; \end{aligned}$$

Theorem 2. Let $fN(t); t \geq 0$ a renewal counting process with renewal function $M(t)$, and assume that the renewals interarrival times τ_i are i.i.d: random variables with probability distribution function $G(t) = P\{\tau_i \leq t\}$: Then for each $n \geq 0$ the probability distribution function of $N(t)$ is exponentially bounded and ver...es:

$$P\{fN(t) > ng\} = G(t) \leq \frac{G(t)}{M(t)} \approx \frac{1}{n};$$

Lemma 2. Let $fN(t); t \geq 0$ a renewal counting process Under the assumptions of the theorem, for each t we can find two positive constants $\phi(t) \in (0, 1)$ and $\psi(t)$ such that $P\{fN(t) > ng\} = \psi(t) e^{-n \ln \phi(t)}$:

Proof. (lemma). For any $s > 0$ define by $G_L(s)$ the Laplace transform of $dG(t) = G(t + dt) - G(t)$; where dt is an infinitesimal increment of t : Formally:

$$\begin{aligned} G_L(s) &= \int_0^\infty e^{-st} dG(t) \\ &= \int_0^\infty e^{-st} dG(t) \\ &= G(0) + \int_0^\infty e^{-st} dG(t) \end{aligned}$$

Taking the limit when $s \rightarrow 1$, we can see from the previous equality that:

$$\lim_{s \rightarrow 1} G_L(s) = G(0) = 0:$$

Now let α be any nonzero and finite real number. It is clear that for sufficiently large s we will have $\alpha G_L(s) < 1$: On the other hand, since $P f T_n \cdot t g = P f \lambda_1 + \lambda_2 + \dots + \lambda_n \cdot t g = G^{\alpha n}(t)$; we can write:

$$\begin{aligned} \sum_{n=0}^{\infty} \alpha^n P f T_n \cdot t g &= \sum_{n=0}^{\infty} \alpha^n G^{\alpha n}(t) \\ &= \sum_{n=0}^{\infty} \alpha^n P \left(e^{-s T_n} \cdot e^{s t \alpha} \right); \end{aligned}$$

where s is supposed to be such that $\alpha G_L(s) < 1$: A direct application of Markov's inequality (see for instance Gnedenko, 1988, p. 218), and of the independence of f_i, g_i obtains:

$$\begin{aligned} \sum_{n=0}^{\infty} \alpha^n P \left(e^{-s T_n} \cdot e^{s t \alpha} \right) &= e^{s t} \sum_{n=0}^{\infty} \alpha^n \int_0^\infty e^{-s t} dG(t) \\ &= e^{s t} \sum_{n=0}^{\infty} \alpha^n G_L^n(s) \\ &= e^{s t} \sum_{n=0}^{\infty} [\alpha G_L(s)]^n \\ &< 1: \end{aligned}$$

It follows that $\alpha^n G^{\alpha n}(t) \rightarrow 0$ as $n \rightarrow \infty$: Thus we can find a positive integer n^0 such that for $n > n^0$ and a particular value $\alpha(t) \in (0, 1)$ we get $G^{\alpha n}(t) < \alpha^n(t)$ for each $t > 0$: Thus for any $n > n^0$ we can write:

$$\begin{aligned} P f N(t) > n g &= P f T_n \cdot t g \\ &= G^{\alpha n}(t) \\ &< \alpha^n(t) = \exp[n \ln \alpha(t)]: \end{aligned}$$

Finally notice that for each t we can find a constant $\alpha(t)$ such that for any n we get $G^{\alpha n}(t) < \alpha^n(t)$; and this proves the lemma. \square

Proof. (theorem). To prove the theorem, let the renewal function be $M(t) = E \sum_{n=0}^{\infty} P fN(t) > ng$: We have:

$$\begin{aligned} M(t) &= \sum_{n=0}^{\infty} P fN(t) > ng \\ &= \sum_{n=0}^{\infty} e^{-i n \ln \phi} = \sum_{n=0}^{\infty} e^{-i \ln \phi^n} \\ &= \frac{1}{1 - e^{-i \ln \phi}} \\ &= \frac{1}{1 - \phi^{-i}} \end{aligned}$$

The result of the lemma is valid for all $n \geq 0$. For $n = 0$ we have

$$P fN(t) > 0g = P fN(t) \geq 1g = P(\xi \cdot t) = G(t):$$

Therefore:

$$\begin{aligned} G(t) &= \sum_{n=0}^{\infty} e^{-i n \ln \phi} = [1 - \phi^{-i}]^{-1} M(t) \\ &= \sum_{n=0}^{\infty} e^{-i n \ln \phi} = \sum_{n=0}^{\infty} \frac{G(t)}{M(t)} \phi^{-i n} \\ &= \sum_{n=0}^{\infty} P fN(t) > ng = G(t) \sum_{n=0}^{\infty} \frac{G(t)}{M(t)} \phi^{-i n} \end{aligned}$$

□

Theorem 3. Let $X_t = \sum_{i=1}^t z_i$ where $f^2_{z_i}$ are continuous i.i.d: zero-mean random variables with finite variance σ^2 , and let $N^+(t) = \sum_{i=1}^t \Phi N^+(i)$, and $N^-(t) = \sum_{i=1}^t \Phi N^-(i)$: Then for any i and for finite j we have:

$$\begin{aligned} \text{Cov} f \Phi N^+(i); \Phi N^+(j)g &> 0; \\ \text{Cov} f \Phi N^-(i); \Phi N^-(j)g &> 0; \\ \text{Cov} f \Phi N^+(i); \Phi N^-(j)g &= \phi^{-i} \phi^{-j} < 0; \\ \text{Cov} f N^+(t); N^-(t)g &< 0; \text{ for finite } t; \end{aligned}$$

with

$$\phi^{-i} = P f \Phi N^+(i) = 1g = P f \Phi N^-(i) = 1g \gg i^{-1/2}; \text{ for large } i.$$

Proof. To see this it is enough to remark that the conditional probability that a renewal occurs at time i when a renewal occurred at the previous time instant $i - 1$; verifies:

$$\begin{aligned} P f \Phi N^+(i) = 1j \Phi N^+(i - 1) = 1g &= P f^2_i > 0g \\ \text{and } P f \Phi N^-(i) = 1j \Phi N^-(i - 1) = 1g &= P f^2_i < 0g; \end{aligned}$$

where $z_i = X_i - X_{i-1}$; but

$$\begin{aligned} P f \Phi N^+(i) = 1j \Phi N^+(i - 1) = 0g &\notin P f^2_i > 0g \\ \text{and } P f \Phi N^-(i) = 1j \Phi N^-(i - 1) = 0g &\notin P f^2_i < 0g; \end{aligned}$$

In fact, it is shown here below that the increments $f \Phi N(i)g$ of either component

record counting process are positively correlated. Indeed, for any $i > j$:

$$\begin{aligned} \text{Cov}f\Phi N(i); \Phi N(j)g &= E f\Phi N(i)\Phi N(j)g_i - \Phi M(i)\Phi M(j) \\ &= P f\Phi N(i) = 1; \Phi N(j) = 1g_i - \Phi M(i)\Phi M(j) \\ &= P f\Phi N(i) = 1j\Phi N(j) = 1g P f\Phi N(j) = 1g_i - \Phi M(i)\Phi M(j) \\ &= [P f\Phi N(i) = 1j\Phi N(j) = 1g_i - \Phi M(i)]\Phi M(j); \end{aligned}$$

Now it is straightforward that if the model error distribution is symmetric then:

$$P f\Phi N(i) = 1j\Phi N(j) = 1g = P(z_i + z_{i-1} + \dots + z_{i-j+1} > 0) = 1=2; \text{ for any } j < i:$$

But $\Phi M(i) \neq 0$ as $i \neq 1$, so for fixed j and growing i we have $\text{Cov}f\Phi N(i); \Phi N(j)g >$

0. When the error distribution is not symmetric we still have $P(z_i + z_{i-1} + \dots + z_{i-j+1} > 0) > 0$ as i grows to infinity, and the result also holds. On the other hand,

the increments $\Phi N^+(i)$ and $\Phi N^-(i)$ are negatively correlated. Indeed:

$$E f\Phi N^+(i)\Phi N^-(i)g = P^{\circ} \Phi N^+(i) = 1; \Phi N^-(i) = 1^a = 0; \quad 8i;$$

and therefore:

$$\begin{aligned} \text{Cov}f\Phi N^+(i); \Phi N^-(i)g &= -i E f(\Phi N^+(i))^2g \\ &= -i P f\Phi N^+(i) = 1g = -i P f\Phi N^-(i) = 1g \\ &= -i \rho(i); \end{aligned}$$

As $\rho(i) \gg i^{-1/2}$; this correlation decays too slowly. Notice that for $j < i$ we must

have:

$$E f\Phi N^+(i)\Phi N^-(j)g = P^{\circ} \Phi N^+(i) = 1; \Phi N^-(j) = 1^a > 0;$$

Therefore

$$\begin{aligned} \text{Cov}f\Phi N^+(i); \Phi N^-(j)g &> -i \rho(i) \rho(j) > -i \rho^2(i) > -i \rho(i) = \text{Cov}f\Phi N^+(i); \Phi N^-(i)g \\ \text{while } j \text{Cov}f\Phi N^+(i); \Phi N^-(j)g &< j \text{Cov}f\Phi N^+(i); \Phi N^-(i)g; \end{aligned}$$

So the covariance matrix for the increments $f\Phi N^+(i); \Phi N^-(j)g_{i,j}$ must be negative

definite, entailing for finite t a net negative correlation between the upper and lower record counting processes $N^+(t)$ and $N^-(t)$: \square

6. Bibliography

1. Andel, J. (2001) *Mathematics of Chance*. John Wiley & Sons, New York.
2. Aparicio, F.M. and C.W.J. Granger (1985) *Nonlinear Cointegration and some New Tests for Comovements*. Discussion Paper 95-15, Department of Economics, University of California at San Diego.
3. Aparicio, F.M. and Escribano, A. (1998) Information-theoretic analysis of serial dependence and cointegration. *Studies in Nonlinear Dynamics and Econometrics*, 3, 3, 119-140.
4. Aparicio, F.M., Escribano, A. and Garcia, A. (2000a) A robust test of dominant low-frequency behaviour in time series. 20th International Symposium on Forecasting, Lisbon (Portugal), June 2000.
5. Aparicio, F.M., Escribano, A. and Garcia, A. (2000b) A nonparametric test for strong low-frequency dependence in time series. International Conference on Nonlinear and Long-Memory Time Series, Cardiff (UK), July, 2000.
6. Arnold, B.C. and Balakrishnan, N. (1989) *Relations, Bounds and Approximations for Order Statistics*. Lecture Notes in Statistics 53, Springer-Verlag, New York.
7. Arnold, B.C., Balakrishnan, N. and Nagarja, H.N. (1993) *A First Course in Order Statistics*. John Wiley & Sons, New York.
8. Ballerini, R. and Resnick, S.L. (1985) Records from improving populations. *Journal of Applied Probability*, 22, 487-502.
9. Bosq, D. and Nguyen, H.T. (1996) *A Course in Stochastic Processes*. Kluwer Academic Press, New York.
10. Breitung, J. and Gouriéroux, C. (1997) Rank tests for unit roots. *Journal of Econometrics*, 81, 7-28.
11. Breitung, J. (1998), Nonparametric tests for nonlinear cointegration. In *Decision Technologies for Computational Finance*, 109-123. Kluwer Academic Press, New York.
12. Burridge, P. and Guerre, E. (1996) The Limit Distribution of Level Crossings of a Random Walk. *Econometric Theory*, 12, 705-723.
13. Chandler, K.N. (1952) The distribution and frequency of record values. *Journal of the Royal Statistical Society, B* 14, 220-228.
14. Cox, D.R. (1962) *Renewal Theory*. John Wiley & Sons, New York.
15. Embrechts, P., Klüppelberg, C. and Mikosch, T. (1999) *Modelling Extremal Events (for Insurance and Finance)*. Springer-Verlag, Berlin.
16. Feller, W. (1971) *An Introduction to Probability Theory and its Applications* (vol. 2). John Wiley & Sons, New York.
17. Franses, P.H. and McAleer, M. (1998) Testing for unit roots and nonlinear transformations. *Journal of Time Series Analysis*, 19, 147-164.
18. Galambos, J. (1978) *The Asymptotic Theory of Extreme Order Statistics*. John Wiley & Sons, New York.
19. Glick, N. (1978) Breaking records and breaking boards. *American Mathematical Monthly*, 85, 2-26.
20. Gnedenko, B. (1988) *The Theory of Probability*. Mir Publishers, Moscow.
21. Granger, C.W.J. and Hallman, J. (1991) Nonlinear transformations of integrated time series. *Journal of Time Series Analysis*, 12, 207-224.
22. Granger, C.W.J. (1995) Modelling nonlinear relationships between extended-memory variables. *Econometrica*, 63, 2, 265-269.

23. Granger, C.W.J. and Teräsvirta, T. (1993) *Modelling Economic Nonlinear Relationships*. Oxford University Press.
24. Kallenberg, O. (1997) *Foundations of Modern Probability*. Springer-Verlag, Berlin.
25. Karlin, S. and Taylor (1975) *A Second Course in Stochastic Processes*. Academic Press, New York.
26. Knopp, L. (1990), *Theory of Infinite Series*, Dover, New York.
27. Kotz, S. and Nadarajah, S. (2000) *Extreme Value Distributions*. Imperial College Press, London.
28. Leadbetter, M.R. and Rootzén, H. (1988) Extremal theory for stochastic processes. *The Annals of Probability*, vol. 16, nº2, pp. 431-478.
29. Leadbetter, M.R. and Nandagopalan, S. (1989) On exceedance point processes for stationary sequences under mild oscillation restrictions. In *Extreme Value Theory* (J. Hüßler and R.D. Reiss, eds.), *Lectures Notes in Statistics*, Springer-Verlag, Heidelberg.
30. Lindgren, G. and Rootzén, H. (1987) Extremes values: theory and technical applications. *Scandinavian Journal of Statistics*, 14, 241-279.
31. Loève, M. (1987) *Probability Theory* (vol. I). Springer-Verlag, New York.
32. Nagaraja, H.N. (1988) Record values and related statistics: a review. *Communications in Statistics -Theory and Methods-*, 17, 2223-2238.
33. Resnick, S.I. (1973) Record values and maxima. *Annals of Probability*, 1, 650-652.
34. Resnick, S.I. (1987) *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag, New York.
35. Resnick, S.I. (1992) *Adventures in Stochastic Processes*. Springer-Verlag, New York.
36. Ross, S. (1996) *Stochastic Processes*. John Wiley & Sons, New York
37. Shorrock, R.W. (1973) Record values and inter-record times. *Journal of Applied Probability*, 10, 543-555.
38. Shiryaev, A.N. (1996) *Probability*. Springer-Verlag, Berlin.
39. Zwillinger, D. (1996) *Standard Mathematical Tables and Formulae*. CRC Press, New York.

Universidad Carlos III de Madrid, Avda. de la Universidad, 22, 28270 Colmenarejo, Madrid (Spain)

E-mail address: aparicio@est-econ.uc3m.es