TEST (2015) 24:302–321
DOI 10.1007/s11749-014-0408-0

ORIGINAL PAPER

On the point process of near-record values

Raúl Gouet · F. Javier López · Gerardo Sanz

Received: 10 March 2014 / Accepted: 2 October 2014 / Published online: 18 October 2014

Abstract Let (X_n) be a sequence of independent and identically distributed random variables, with common absolutely continuous distribution F. An observation X_n is a near-record if $X_n \in (M_{n-1} - a, M_{n-1}]$, where $M_n = \max\{X_1, \ldots, X_n\}$ and $a > 0$ is a parameter. We analyze the point process η on $[0, \infty)$ of near-record values from (X_n), showing that it is a Poisson cluster process. We derive the probability generating functional of η and formulas for the expectation, variance and covariance of the counting variables $\eta(A)$, $A \subset [0, \infty)$. We also obtain strong convergence and asymptotic normality of $\eta(t) := \eta([0, t])$, as $t \to \infty$, under mild tail-regularity conditions on F. For heavy-tailed distributions, with square-integrable hazard function, we show that $\eta(t)$ grows to a finite random limit $\eta(\infty)$ and compute its probability generating function. We apply our results to Pareto and Weibull distributions and include an example of application to real data.

Keywords Record · Near-record · Poisson cluster process · Law of large numbers · Central limit theorem

Mathematics Subject Classification 60G70 · 60G55 · 60F05 · 60F15

R. Gouet
Dpto. Ingeniería Matemática y Centro de Modelamiento Matemático (UMI 2807, CNRS), Universidad de Chile, Casilla 170/3, Santiago, Chile
e-mail: rgouet@dim.uchile.cl

F. J. López (✉) · G. Sanz
Dpto. Métodos Estadísticos and BIFI, Facultad de Ciencias, Universidad de Zaragoza, C/ Pedro Cerbuna, 12, 50009 Zaragoza, Spain
e-mail: javier.lopez@unizar.es

G. Sanz
e-mail: gerardo.sanz@unizar.es

1 Introduction

Given a sequence of random variables (X_n), indexed by the positive integers, X_n is a (upper) record if it exceeds all preceding observations $X_1, \ldots, X_{n-1}$. Mathematical properties of records have been under study for decades, mostly for independent and identically distributed (iid) random variables, with continuous parent distribution. The interested reader may consult the monograph by Arnold et al. (1998) for a complete account of the theory up to the late 1990s.

A near-record is an observation which is not a record but is close to being one, and as such can be seen as natural extension of the idea of record. Informally, a near-record fails to be a record but is located within distance a of the current record value.

Near-records were defined in Balakrishnan et al. (2005) and their potential interest in actuarial mathematics was pointed out. More recently, the idea of combining records and near-records in statistical procedures was first proposed in Gouet et al. (2012c), as a way of compensating for the relative exiguity of records, without radically changing the experimental design. Maximum likelihood estimation was developed and a natural strategy for simultaneously collecting records and near-records was described, in the context of destructive stress testing. See Gulati and Padgett (2003) for information on record-data-based inference.

Previous work on near-records has mainly focused on "time-axis" counting variables such as the number of near-records associated to the n-th record, first studied in Balakrishnan et al. (2005), with additional results in Pakes (2007). The asymptotic behavior of the total number of near-records, among the first n observations, was investigated in Gouet et al. (2012a, b), where central limit theorems and laws of large numbers were established. Extensions to bivariate observations and to the Pfeifer model have been analyzed by Bose and Gangopadhyay (2011) and Bairamov and Stepanov (2011), respectively.

We consider the aforementioned results as being of time-axis kind because near-records are counted as time n goes by. The aim of the present paper is to study near-records in the spatial axis of their values. To that end, we define and completely characterize the point process η of near-record values as a Poisson cluster process, in a way similar to Shorrock's characterization of record values; see Arnold et al. (1998, p. 42). Together with the characterization of η, we give the explicit expression of its probability generating functional and formulas for the probability generating function, mean, variance and covariance of counting variables related to η.

When the common distribution F of the observations is heavy-tailed, the total number of near-records along the whole sequence (X_n), denoted by $\eta(\infty)$, is shown to be finite. We are able to characterize the distribution of $\eta(\infty)$ and, in the particular case of the Pareto distribution ($F(x) = 1 - 1/x^r$, for $x > 1$), we answer a question posed in Gouet et al. (2012b). For moderate- and light-tailed F, $\eta(t)$ grows to infinity as $t \to \infty$ and, under further regularity assumptions, we establish a law of large numbers and a central limit theorem.

The paper is organized as follows. Notations, definitions and a technical Lemma are presented in Sect. 2. Results on the structure of η and about the limiting behavior of $\eta(t)$ are stated and proved in Sect. 3. Examples of application of the results to particular distributions are given in Sect. 4.

2 Notation and preliminaries

Let (X_n) denote an iid sequence of nonnegative random variables, indexed by the positive integers. Let F be their common distribution function, which is assumed absolutely continuous, with positive density f and hazard function $\lambda(x) := f(x)/\bar{F}(x)$, where $\bar{F} := 1 - F$. The nonnegativity of the random variables and the strict positivity of f can be dropped but we assume them to ease the exposition.

Records. Denote by $M_n := \max\{X_1, \ldots, X_n\}, n \geq 1$, the sequence of partial maxima and call X_n a (upper) record if $X_n > M_{n-1}$, with $M_0 = 0$ (so that X_1 is conventionally counted as record). Record times and record values are, respectively, defined by $L_1 = 1$, $L_n = \min\{k > L_{n-1} | X_k > M_{k-1}\}, n > 1$, and $R_n = X_{L_n}, n \geq 1$. For any Borel subset A of the line, let $F(A) := \int_A f(t)\mathrm{d}t$.

Let ξ denote the point process of record values on the line (Shorrock's process), with counting function given by

$$\xi(t) = \text{card}\{n \geq 1 \mid R_n \leq t\}, t \in [0, \infty).$$

It is well known that, under the assumptions stated above, ξ is a non-homogeneous Poisson process on $[0, \infty)$, with intensity measure $\lambda(t)\mathrm{d}t$.

Near-records. Let a be a fixed positive parameter and call X_n a (upper) near-record if $X_n \in (M_{n-1} - a, M_{n-1}]$. Additionally, we say that the near-record is associated to the m-th record if $L_m < n < L_{m+1}$.

The number of near-records associated to R_m is denoted by S_m and near-record values associated to R_m by $Y_{m,j},\ j = 1, \ldots, S_m$. Observe that $0 \leq S_m < L_{m+1} - L_m$ and also that $R_m - a < Y_{m,j} \leq R_m$, for $j = 1, \ldots, S_m$.

Lemma 1 *Under the hypotheses at the beginning of this section:*

(i) *conditional on $R_m = x$, the distribution of S_m is geometric (starting at 0) with success parameter $p(x) := \bar{F}(x)/\bar{F}(x - a)$.*

(ii) *Conditional on $R_m = x$ and $S_m = k$, the random variables $Y_{m,j},\ j = 1, \ldots, k$, are independent, with common density*

$$\frac{f(y)\mathbf{1}_{(x-a,x]}(y)}{F((x-a,x])}. \tag{1}$$

(iii) *Conditional on the sequence (R_n), the random variables S_n are independent. Moreover, near-records associated to different record values are mutually independent.*

Proof (i) S_m counts the number of independent observations X_j, with $j > L_m$, such that $X_j > R_m - a$, until one observation is greater than R_m (success). Clearly, given $R_m = x$, the probability of success is $P[X_j > x | X_j > x - a] = p(x)$.

(ii) The random variables $Y_{m,j}$ are just X observations constrained to be in $(R_m - a, R_m]$, hence, given $R_m = x$, they are independent with density (1).

(iii) The conclusion follows from the mutual independence of the sets of random variables $\{X_j | \ j \in [L_m + 1, L_{m+1}]\}, m \geq 1$, given (R_n). □

Let R_n^a be the n-th near-record of the sequence (X_n). We define η, the point process of near-record values, through its counting function as

$$\eta(t) = \text{card}\{n \geq 1 \mid R_n^a \leq t\}, t \in [0, \infty),$$

In Sect. 3, $\eta := \{\eta(t), t \geq 0\}$ is shown to be a Poisson cluster process (PCP), with center process given by Shorrock's process ξ.

Poisson cluster process. They are among the most important and versatile types of point process models, with applications in many fields. A PCP is the superposition of a family of independent point processes $\{N(\cdot \mid x), x \in \mathbb{R}\}$ over the points of a Poisson process N_c (the center process). The points x_i of N_c act as the centers of the clusters and, for each x_i, a realization of $N(\cdot \mid x_i)$ (the component process) is observed. The PCP is the result of superposing the points of all clusters $N(\cdot \mid x_i)$.

The probability generating functional G of a point process M is defined as $G_M(h) := E\left[\prod_i h(x_i)\right]$, where the product is taken over the points of each realization of M and h is a measurable function such that $0 \leq h \leq 1$ and $1 - h$ vanishes outside some bounded set.

We recall (Daley and Vere-Jones 2003, Proposition 6.3.III) that the superposition of processes $N(\cdot \mid x)$ over the points of N_c is a PCP if the following conditions hold:

(i) the family $\{N(\cdot \mid x), x \in [0, \infty)\}$ is measurable (i.e., for each h, its probability generating functional $G_N(h \mid x)$ is a measurable function of x);
(ii) $\{N(\cdot \mid x), x \in [0, \infty)\}$ is an independent family and $N([0, \infty) \mid x) < \infty$ almost surely (a.s.), for each $x > 0$;
(iii) the center process N_c is Poisson and
(iv) $\int_0^\infty P[N(A|x) > 0]\mu(dx) < \infty$, for every bounded Borel set $A \subset [0, \infty)$, where μ is the intensity measure of N_c.

3 Results

3.1 Structure of the point process of near-record values

Theorem 1 *Let η be the point process of near-record values. Then, for any measurable function $0 \leq h \leq 1$, with $1-h$ vanishing outside a bounded set, and A, B Borel subsets of $[0, \infty)$,*

(a) *η is a PCP with centers given by Shorrock's process ξ and component processes $\{K(\cdot \mid x), x \in (0, \infty)\}$, such that $K(A|x)$ is geometrically distributed (starting at 0) with parameter*

$$r(x) := \frac{\bar{F}(x)}{\bar{F}(x) + F(A \cap (x - a, x])}. \tag{2}$$

(b) *The probability generating functional G of* η *is given by*

$$G(h) = \exp\left\{-\int_0^\infty \left(1 - \frac{\bar{F}(x)}{\bar{F}(x-a) - \int_{x-a}^x h(t)f(t)\mathrm{d}t}\right)\lambda(x)\mathrm{d}x\right\}. \tag{3}$$

(c) *Moments of* $\eta(A)$*:*

$$\mu(A) := E[\eta(A)] = \int_0^\infty \frac{F(A \cap (x-a,x])}{\bar{F}(x)}\lambda(x)\mathrm{d}x, \tag{4}$$

$$\sigma^2(A) := \mathrm{Var}[\eta(A)] = 2\int_0^\infty \left(\frac{F(A \cap (x-a,x])}{\bar{F}(x)}\right)^2 \lambda(x)\mathrm{d}x + E[\eta(A)] \tag{5}$$

and, if $A \cap B = \emptyset$*,*

$$\mathrm{Cov}[\eta(A), \eta(B)] = 2\int_0^\infty \frac{F(A \cap (x-a,x])}{\bar{F}(x)} \frac{F(B \cap (x-a,x])}{\bar{F}(x)}\lambda(x)\mathrm{d}x. \tag{6}$$

(d) *Probability generating function of* $\eta(A)$*: for* $\alpha \in [0,1]$ *and A bounded*

$$\varphi_A(\alpha) := E[\alpha^{\eta(A)}] = \exp\left(-\int_0^\infty \frac{(1-\alpha)F(A \cap (x-a,x])}{(1-\alpha)F(A \cap (x-a,x]) + \bar{F}(x)}\lambda(x)\mathrm{d}x\right). \tag{7}$$

(e)

$$\int_0^\infty E[K(A|x)^3]\lambda(x)\mathrm{d}x = 6\int_0^\infty \left(\frac{F(A \cap (x-a,x])}{\bar{F}(x)}\right)^3 \lambda(x)\mathrm{d}x + 3\sigma^2(A) - 2\mu(A). \tag{8}$$

Proof (a) Clearly, η is the superposition of near-records associated to the sequence (R_n). On the other hand, near-records associated to R_n form a point process on $[0,\infty)$, which can be characterized as follows: from Lemma 1 (i), the number S_n of near-records associated to R_n, conditional on $R_n = x$, is geometric with parameter $p(x) = \bar{F}(x)/\bar{F}(x-a)$. Each near-record lands in A with probability $\frac{F(A\cap(x-a,x])}{F((x-a,x])}$, independently of other near-records. Therefore, the number of near-records associated to R_n, landing in A, is geometric with parameter

$$\frac{p(x)}{p(x) + (1-p(x))\frac{F(A\cap(x-a,x])}{F((x-a,x])}} = r(x).$$

Also, from Lemma 1 (ii), near-records in A, associated to R_n, are (conditionally on $R_n = x$) independent with common density $f(y)/F(A \cap (x-a,x])$, for $y \in$

$A \cap (x-a, x]$. Let us calculate the probability generating functional $G_K(\cdot \mid x)$ of the process $K(\cdot \mid x)$ by conditioning on S_n. Then,

$$\begin{aligned} G_K(h \mid x) &= E\Big[\prod_i E\Big[h(x_i)\Big|S_n\Big]\Big] \\ &= \sum_{k=0}^{\infty}\prod_{i=1}^{k} E[h(x_i)]P[S_n = k] \\ &= \sum_{k=0}^{\infty}\left(\int_{x-a}^{x}\frac{h(y)f(y)}{F((x-a,x])}\mathrm{d}y\right)^k (1-p(x))^k p(x) \\ &= \frac{\bar F(x)}{\bar F(x-a) - \int_{x-a}^{x} h(t)f(t)\mathrm{d}t}. \end{aligned} \tag{9}$$

The series above is convergent because $h \le 1$ and so, $\int_{x-a}^{x} h(y)f(y)\mathrm{d}y < \bar F(x-a)$.

We now verify that the superposition of the processes $K(\cdot \mid x)$ defines a PCP, by checking conditions (i) to (iv) of Sect. 2.

Condition (i) clearly follows from (9). For (ii), the independence of processes $K(\cdot \mid x)$ follows from Lemma 1 (iii) while finiteness holds because of its geometric distribution. Condition (iii) is immediate since the center process ξ is Poisson. Finally, for (iv) let $A \subseteq [c,d]$, then $x \notin [c, d+a]$ implies $P[K(A|x) > 0] = 0$. So, recalling that λ is the intensity of ξ,

$$\int_0^{\infty} P[K(A|x)>0]\lambda(x)\mathrm{d}x = \int_c^{d+a} P[K(A|x)>0]\lambda(x)\mathrm{d}x \le \int_c^{d+a} \lambda(x)\mathrm{d}x < \infty.$$

Therefore, (a) is proved.

(b) By (6.3.9) in Daley and Vere-Jones (2003), $G(h) = \mathrm{e}^{-\int_0^{\infty}(1-G_K(h|x))\lambda(\mathrm{d}x)}$. Replacing (9) into the previous formula, we obtain the result.

(c) From (6.3.10) in Daley and Vere-Jones (2003), $E[\eta(A)] = \int_0^{\infty} E[K(A|x)]\lambda(x)\mathrm{d}x$. Then, (4) follows by noting that $K(A|x)$ is geometric with expectation $\frac{1-r(x)}{r(x)} = \frac{F(A\cap(x-a,x])}{\bar F(x)}$. For the variance, we apply (6.3.11) in Daley and Vere-Jones (2003) to obtain

$$\mathrm{Var}[\eta(A)] = \int_0^{\infty} E[K(A|x)^2]\lambda(x)\mathrm{d}x.$$

We observe that $E[K(A|x)^2] = 2\left(\frac{1-r(x)}{r(x)}\right)^2 + \frac{1-r(x)}{r(x)}$ and so (5) follows. For the covariance, we note that $2\mathrm{Cov}[\eta(A), \eta(B)] = \mathrm{Var}[\eta(A)+\eta(B)] - \mathrm{Var}[\eta(A)] - \mathrm{Var}[\eta(B)]$. So, if $A \cap B = \emptyset$,

$$\begin{aligned}
&\mathrm{Var}[\eta(A)+\eta(B)] = \mathrm{Var}[\eta(A\cup B)] \\
&\quad = 2\int_0^\infty \left(\frac{F((A\cup B)\cap(x-a,x])}{\bar F(x)}\right)^2 \lambda(x)\mathrm{d}x + E[\eta(A\cup B)] \\
&\quad = 2\int_0^\infty \left(\frac{F(A\cap(x-a,x])}{\bar F(x)}\right)^2 \lambda(x)\mathrm{d}x + 2\int_0^\infty \left(\frac{F(B\cap(x-a,x])}{\bar F(x)}\right)^2 \lambda(x)\mathrm{d}x \\
&\qquad + 4\int_0^\infty \frac{F(A\cap(x-a,x])F(B\cap(x-a,x])}{\bar F(x)^2}\lambda(x)\mathrm{d}x + E[\eta(A\cup B)] \\
&\quad = 4\int_0^\infty \frac{F(A\cap(x-a,x])F(B\cap(x-a,x])}{\bar F(x)^2}\lambda(x)\mathrm{d}x + \mathrm{Var}[\eta(A)] + \mathrm{Var}[\eta(B)],
\end{aligned}$$

which yields (6).

(d) For (7), we define $h(x) = \alpha^{\mathbf{1}_A(x)}$ and calculate

$$\int_{x-a}^{x} h(t)f(t)\mathrm{d}t = \bar F(x-a) - \bar F(x) - (1-\alpha)F(A\cap(x-a,x]),$$

which is replaced in (3).

(e) For (8), we use $E[K(A|x)^3] = 6\left(\frac{1-r(x)}{r(x)}\right)^3 + 6\left(\frac{1-r(x)}{r(x)}\right)^2 + \frac{1-r(x)}{r(x)}$. □

We consider below the random variables $N(s,t) := \eta(t) - \eta(s)$, for $0 \le s < t$, corresponding to the number of near-record values in the interval $(s,t]$. For simplicity, we write $\eta(t)$ instead of $N(0,t)$.

Closed-form expressions for $E[\eta(t)]$, $\mathrm{Var}[\eta(t)]$, $\mathrm{Cov}[N(s,t),N(t,u)]$ and the probability generating function $E[\alpha^{\eta(t)}]$ are shown below. For the covariance between $N(s_1,s_2)$ and $N(t_1,t_2)$, note first that if $s_2 < t_1 - a$, $N(s_1,s_2)$ and $N(t_1,t_2)$ are independent and so the covariance is 0. Also, due to the properties of the covariance and having computed $\mathrm{Var}[\eta(t)]$, it is easy to see that the different relative positions of s_1, s_2, t_1, t_2 can be reduced to $s_1 < s_2 = t_1 < t_2$. Therefore, we compute the covariance explicitly only in this situation. We use the notations $r\wedge s := \min\{r,s\}$, $r \vee s := \max\{r,s\}$.

Corollary 1 *Moments and probability generating function.*

$$\mu(t) := E[\eta(t)] = \int_0^{t+a} \frac{\bar F(x-a) - \bar F(x\wedge t)}{\bar F(x)}\lambda(x)\mathrm{d}x, \quad (10)$$

$$\sigma^2(t) := \mathrm{Var}[\eta(t)] = 2\int_0^{t+a}\left(\frac{\bar F(x-a)-\bar F(x\wedge t)}{\bar F(x)}\right)^2\lambda(x)\mathrm{d}x + \mu(t). \quad (11)$$

Let $s < t < u$, then

$$\begin{aligned}
&\mathrm{Cov}[N(s,t),N(t,u)] \\
&\quad = 2\int_t^{t+a} \frac{\left(\bar F(t) - \bar F(x\wedge u)\right)\left(\bar F(s\vee(x-a)) - \bar F(t)\right)}{\bar F(x)^2}\lambda(x)\mathrm{d}x \qquad (12)
\end{aligned}$$

and, for $\alpha \in [0, 1]$,

$$\varphi_t(\alpha) := E[\alpha^{\eta(t)}] = \exp\left(-\int_0^{t+a} \frac{(1-\alpha)(\bar{F}(x-a) - \bar{F}(x \wedge t))}{(1-\alpha)(\bar{F}(x-a) - \bar{F}(x \wedge t)) + \bar{F}(x)} \lambda(x)\mathrm{d}x\right). \tag{13}$$

Proof Formulas for $\mu(t)$ and $\sigma^2(t)$ are obtained from (4) and (5), respectively, with $A = [0, t]$, noting that $F([0, t] \cap (x-a, x]) = (\bar{F}(x-a) - \bar{F}(x \wedge t))\mathbf{1}_{\{x<t+a\}}$. For the covariance, we use (6) with $A = (s, t]$, $B = (t, u]$, noting that $F((s, t] \cap (x-a, x]) = (\bar{F}(s \vee (x-a)) - \bar{F}(t \wedge x))\mathbf{1}_{\{s<x<t+a\}}$.

For (13), it suffices to take $A = [0, t]$ in (7). □

Remark 1 Observe that while the center process ξ has independent increments, η only has a-dependent increments, in the sense that for $t_0 < t_1 < \cdots < t_{2n+1}$, with $t_{2i} - t_{2i-1} > a$, $i = 1, \ldots, n$, the random variables $\eta(t_{2i+1}) - \eta(t_{2i})$, $i = 0, \ldots, n$, are independent. This fact is exploited in the proof of the strong law of large numbers. See Propositions 1 and 2.

Remark 2 A notion closely related to near-record is that of δ-record. Given a fixed real parameter δ, X_n is called a δ-record if $X_n > M_{n-1} + \delta$. So, if we take $\delta = -a < 0$, then a δ-record is either a record or a near-record (of parameter a). Therefore, for negative δ, the process of δ-record values differs from the process of near-record values only in that the centers are counted as points of the respective component processes. So, the point process of δ-record values is also a PCP and the formulas in Theorem 1 and Corollary 1 can be easily adapted. For distributional results on δ-records, see López-Blázquez and Salamanca-Miño (2013).

3.2 Asymptotic analysis

In this section, we study the limit behavior of $\eta(t)$, as $t \to \infty$. For simplicity, we assume that the density f is continuous and strictly decreasing on $(0, \infty)$. But since we deal with asymptotics of upper extremes, where only the upper-tail behavior matters, we could replace these conditions by weaker ones such as f being ultimately continuous and strictly decreasing. It is straightforward to check that all theorems in this section hold under the weaker versions of the hypotheses; see Remark 5. Note that conditions above imply that the right-end point of F, defined by $\sup\{x \mid F(x) < 1\}$, is infinite.

Before stating the main results, we give two technical lemmas. The first one deals with heavy- and exponential-tailed distributions, while the second applies to light-tailed ones. Observe that tail properties are stated in terms of the hazard function λ. We define $v(x) := (\bar{F}(x-a) - \bar{F}(x))/\bar{F}(x)$, which is the expected number of near-records associated to a record with value x.

Lemma 2 *Suppose* $\int_{x-a}^{x} \lambda(t)\mathrm{d}t < D$, *for some* $D > 0$ *and every* $x > 0$. *Then, there exist positive constants* A, C *such that*

(a) $0 < v(x) < A$, *for all* $x > 0$, *and*
(b) $a\lambda(x) < v(x) < C\lambda(x-a)$, *for all* $x > a$.

Proof (a) Since $\bar{F}(x) = \mathrm{e}^{-\int_0^x \lambda(t)\mathrm{d}t}$, we have $0 < v(x) = \frac{\bar{F}(x-a)-\bar{F}(x)}{\bar{F}(x)} = \mathrm{e}^{\int_{x-a}^x \lambda(t)\mathrm{d}t} - 1 < \mathrm{e}^D - 1$.

(b) Let $x > a$ then, by the mean value theorem, $\bar{F}(x-a) - \bar{F}(x) = af(\theta(x))$, where $\theta(x) \in (x-a, x)$. Hence, since f is decreasing,

$$a\lambda(x) = a\frac{f(x)}{\bar{F}(x)} < a\frac{f(\theta(x))}{\bar{F}(x)} = v(x).$$

Also, from (a), we have

$$a\frac{f(\theta(x))}{\bar{F}(x)} < a\frac{f(x-a)}{\bar{F}(x-a)}\frac{\bar{F}(x-a)}{\bar{F}(x)} = a\lambda(x-a)(v(x)+1) < a\lambda(x-a)\mathrm{e}^D,$$

thus (b) is proved. □

Remark 3 Note from Lemma 2 that $\int_{x-a}^x \lambda(t)\mathrm{d}t$ is bounded if and only if $\lambda(x)$ is bounded. In this equivalence, the hypothesis of f being decreasing is crucial.

Lemma 3 *Suppose that* $\lambda(x) \to \infty$ *as* $x \to \infty$ *and that* $|\lambda'(x)| < M, \forall x > x_0$, *for some* $x_0 > a, M > 0$. *Then, there exist constants* $A, B, C > 0$ *and* $D > 1$ *such that, for all* $x \geq a$,

(a) $\frac{\bar{F}(x-a)}{\bar{F}(x)} > D$ *and*
(b) $A\mathrm{e}^{a\lambda(x)} < B\frac{\bar{F}(x-a)}{\bar{F}(x)} < v(x) < C\mathrm{e}^{a\lambda(x)}$.

Proof (a) Let $m = \inf\{\lambda(x); x > 0\}$, which is positive since f is continuous and decreasing, and $\lambda(x) \to \infty$. Then, $\frac{\bar{F}(x-a)}{\bar{F}(x)} = \mathrm{e}^{\int_{x-a}^x \lambda(t)\mathrm{d}t} > \mathrm{e}^{am} > 1$.

(b) Let $x > x_0$ and $z \in [x-a, x]$ then, by the mean value theorem, $|\lambda(z) - \lambda(x)| < M(x-z) < Ma$. Therefore,

$$\left|\int_{x-a}^x \lambda(t)\mathrm{d}t - a\lambda(x)\right| = \left|\int_{x-a}^x (\lambda(t) - \lambda(x))\mathrm{d}t\right| < Ma\int_{x-a}^x \mathrm{d}t = Ma^2,$$

so

$$a\lambda(x) - Ma^2 < \int_{x-a}^x \lambda(t)\mathrm{d}t < a\lambda(x) + Ma^2$$

and

$$\mathrm{e}^{a\lambda(x)} < \mathrm{e}^{Ma^2}\mathrm{e}^{\int_{x-a}^x \lambda(t)\mathrm{d}t} = \mathrm{e}^{Ma^2}\frac{\bar{F}(x-a)}{\bar{F}(x)}.$$

On the other hand,

$$\frac{D-1}{D}\frac{\bar{F}(x-a)}{\bar{F}(x)} < v(x) < \frac{\bar{F}(x-a)}{\bar{F}(x)} < \mathrm{e}^{Ma^2}\mathrm{e}^{a\lambda(x)},$$

so (b) is proved in the case of $x > x_0$. Observe that all inequalities in (b) also hold (with possibly different constants) for $x \in [a, x_0]$ since the functions involved are all positive and continuous. □

In the next result, we show that, for distributions F with very heavy tails, $\eta(\infty) := N(0,\infty)$ is finite a.s. and give its probability generating function.

Theorem 2 *Suppose that* $\int_a^\infty \lambda^2(x)\mathrm{d}x < \infty$*, then* $\eta(\infty)$ *is a.s. finite, with probability generating function given by*

$$\varphi_\infty(\alpha) := E[\alpha^{\eta(\infty)}] = \exp\left\{-\int_0^\infty \frac{(1-\alpha)v(x)}{(1-\alpha)v(x)+1}\lambda(x)\mathrm{d}x\right\}. \tag{14}$$

Proof Let $\alpha \in [0,1]$ then $\alpha^{\eta(t)} \to \alpha^{\eta(\infty)}$ a.s. and, by the monotone convergence theorem, $\varphi_t(\alpha) \to \varphi_\infty(\alpha)$, as $t \to \infty$. Thus, to prove (14), we take limits in (13). In fact, the integral in (13) can be written as $\int_0^\infty h(x,t)\mathrm{d}x$, with

$$h(x,t) := \frac{(1-\alpha)(\bar{F}(x-a) - \bar{F}(x\wedge t))}{(1-\alpha)(\bar{F}(x-a) - \bar{F}(x\wedge t)) + \bar{F}(x)}\lambda(x)\mathbf{1}_{\{x<t+a\}}.$$

Then, $h(x,t)$ converges increasingly to $h(x) := \frac{(1-\alpha)v(x)}{(1-\alpha)v(x)+1}\lambda(x)$, as $t \to \infty$. So, by the monotone convergence theorem, we have $\int_0^\infty h(x,t)\mathrm{d}x \to \int_0^\infty h(x)\mathrm{d}x$. Finally, we check that $\eta(\infty)$ is finite by showing that the integral in (14) is finite. To apply Lemma 2 we observe, from the Cauchy–Schwarz inequality, that

$$\left(\int_{x-a}^{x} \lambda(t)\mathrm{d}t\right)^2 \le a\int_{x-a}^{x} \lambda^2(t)\mathrm{d}t \le a\int_a^\infty \lambda^2(t)\mathrm{d}t < \infty,$$

for all $x > 2a$. Hence, by Lemma 2 (b) and the Cauchy–Schwarz inequality

$$\int_{2a}^\infty \frac{v(x)}{(1-\alpha)v(x)+1}\lambda(x)\mathrm{d}x \le \int_{2a}^\infty v(x)\lambda(x)\mathrm{d}x \le C\int_{2a}^\infty \lambda(x-a)\lambda(x)\mathrm{d}x < \infty.$$

Thus, the result is proved. □

Remark 4 In the absence of some regularity condition on λ, we do not know whether $\int_a^\infty \lambda^2(x)\mathrm{d}x = \infty$ implies $\eta(\infty) = \infty$ or not but we are able to show that $E[\eta(\infty)] = \infty$. Indeed, $E[\eta(\infty)] = \int_0^\infty v(x)\lambda(x)\mathrm{d}x > a\int_a^\infty \lambda^2(x)\mathrm{d}x = \infty$, because the first inequality in Lemma 2 (b) is valid for all functions λ.

Lemma 4 *Suppose that either* $\liminf_{x\to\infty}\lambda(x) > 0$ *or* $\limsup_{x\to\infty}\lambda(x) < \infty$*. Then,* $\int_a^\infty \lambda^2(x)\mathrm{d}x = \infty$ *implies* $\eta(\infty) = \infty$ *a.s.*

Proof We show that $\varphi_\infty(\alpha) = 0$. From Lemma 2 (b), we obtain

$$\int_a^\infty \frac{v(x)}{(1-\alpha)v(x)+1}\lambda(x)\mathrm{d}x > \int_a^\infty \frac{\lambda^2(x)}{(1-\alpha)\lambda(x)+a^{-1}}\mathrm{d}x. \tag{15}$$

If $A > 0$ is a lower bound of λ, then $\frac{\lambda^2(x)}{(1-\alpha)\lambda(x)+a^{-1}} \geq \frac{A^2}{(1-\alpha)A+a^{-1}}$ and so the integral in the rhs of (15) diverges. If λ is bounded above by $B > 0$, then $\frac{\lambda^2(x)}{(1-\alpha)\lambda(x)+a^{-1}} \geq \frac{\lambda^2(x)}{(1-\alpha)B+a^{-1}}$ and again, the integral diverges. □

In Theorem 3, we establish a strong law of large numbers for $\eta(t)$ under regularity conditions on λ, which are satisfied by many well-known distributions. Note that hypothesis (a) is related to heavy- and moderate-tailed distributions, while (b) concerns light-tailed ones. We first prove two propositions dealing with the discretized process $\eta(na), n = 1, 2, \ldots$, under (a) and (b), respectively.

Proposition 1 *If* $\lambda(x) \leq M, \forall x > 0$, *for some* $M > 0$, *and* $\int_a^\infty \lambda^2(x)\mathrm{d}x = \infty$, *then* $\eta(na)/\mu(na) \to 1$ *a.s.*

Proof Let $Z_n = N((n-1)a, na), n \geq 1$. The a-dependence of η, commented in Remark 1, implies that the Z_n are 1-dependent. We apply Lemma 5 in the Appendix, with $Y_n = Z_n - E[Z_n]$ and $b_n = \mu(na)$, which grows to ∞ as shown in Remark 4. So, we have to establish

$$\sum_{n=1}^\infty \mathrm{Var}[Z_n]/\mu(na)^2 < \infty. \tag{16}$$

We calculate and bound the variance of $N(t-a, t)$ as follows. From (4), (5) and Lemma 2 we have, for $t \geq 3a$,

$$\begin{aligned} \mathrm{Var}[N(t-a,t)] =& 2\int_{t-a}^{t+a} \left(\frac{F((t-a,t] \cap (x-a,x])}{\bar{F}(x)}\right)^2 \lambda(x)\mathrm{d}x \\ &+ \int_{t-a}^{t+a} \frac{F((t-a,t] \cap (x-a,x])}{\bar{F}(x)}\lambda(x)\mathrm{d}x \\ &\leq \int_{t-a}^{t+a} (2v^2(x) + v(x))\lambda(x)\mathrm{d}x \\ &\leq \int_{t-a}^{t+a} (2C^2\lambda^2(x-a) + C\lambda(x-a))\lambda(x)\mathrm{d}x \\ &\leq \frac{C^2(2aM+1)}{a}\int_{t-a}^{t+a} \lambda^2(x-a)\mathrm{d}x = C'\int_{t-2a}^{t} \lambda^2(x)\mathrm{d}x. \end{aligned} \tag{17}$$

Now, letting $t = na$, with C', C'', C''' positive constants, we obtain from Lemma 2

$$\int_{(n-2)a}^{na} \lambda^2(x)\mathrm{d}x \le \int_{(n-2)a}^{(n-1)a} \lambda^2(x)\mathrm{d}x + C'' \int_{(n-1)a}^{na} \lambda^2(x-a)\mathrm{d}x$$
$$= (1+C'') \int_{(n-2)a}^{(n-1)a} \lambda^2(x)\mathrm{d}x,$$

for $n \ge 3$. Hence, $\mathrm{Var}[Z_n] \le C''' \int_{(n-2)a}^{(n-1)a} \lambda^2(x)\mathrm{d}x$. Also, from (10) and Lemma 2, $\mu(na) \ge \int_a^{na} v(x)\lambda(x)\mathrm{d}x \ge a \int_a^{(n-1)a} \lambda^2(x)\mathrm{d}x$. Therefore, (16) holds if we prove

$$\sum_{n\ge 4} \frac{c_n}{S_n^2} < \infty, \tag{18}$$

with $c_n = \int_{(n-2)a}^{(n-1)a} \lambda^2(y)\mathrm{d}y$ and $S_n = \sum_{k=3}^n c_k, n \ge 3$. Note that the series (18) is bounded above by

$$\sum_{n\ge 4} \frac{c_n}{S_n S_{n-1}} = \sum_{n\ge 4} \left(\frac{1}{S_{n-1}} - \frac{1}{S_n} \right) = \frac{1}{\int_a^{2a} \lambda^2(y)\mathrm{d}y} < \infty$$

and so (16) follows. □

Proposition 2 *If* $\lambda(x) \to \infty$ *as* $x \to \infty$ *and* $|\lambda'(x)| < x^{-r}, \forall x > x_0$, *for some* $x_0 > a, r > 1/2$, *then* $\eta(na)/\mu(na) \to 1$ *a.s.*

Proof Without loss of generality, let $r \in (1/2, 1)$. We proceed as in Proposition 1 and prove convergence of the series (16). From (10) and Lemma 3, we have

$$\mu(t) \ge \int_a^t v(x)\lambda(x)\mathrm{d}x > A \int_a^t \mathrm{e}^{a\lambda(x)}\lambda(x)\mathrm{d}x \tag{19}$$

and, from (17),

$$\mathrm{Var}[N(t-a,t)] \le \int_{t-a}^{t+a} (2v^2(x) + v(x))\lambda(x)\mathrm{d}x$$
$$\le (2+B^{-1}) \int_{t-a}^{t+a} v^2(x)\lambda(x)\mathrm{d}x$$
$$\le (2+B^{-1})C \int_{t-a}^{t+a} \mathrm{e}^{2a\lambda(x)}\lambda(x)\mathrm{d}x. \tag{20}$$

Also, for t large enough, knowing that λ' is bounded, there is a constant $H > 0$ such that

$$\int_{t-a}^{t+a} \mathrm{e}^{2a\lambda(x)}\lambda(x)\mathrm{d}x \le H\mathrm{e}^{2a\lambda(t)}\lambda(t). \tag{21}$$

So, to establish (16), it suffices to take $t = na$ in (19) and (20) and prove convergence of the series $\sum_{n\geq 3} h(na)$, where $h(t) = \mathrm{e}^{2a\lambda(t)}\lambda(t)/\left(\int_0^t \mathrm{e}^{a\lambda(x)}\lambda(x)\mathrm{d}x\right)^2$. It can be shown (Gouet et al. 2012c, p. 202), that $t^{2r}h(t) \to 0$ as $t \to \infty$. Hence, for n large enough, $h(na) < (na)^{-2r}$ and, because $r \in (1/2, 1)$, we obtain $\sum_{n\geq 3} h(na) < \infty$. □

Theorem 3 *Suppose that either*

(a) $\lambda(x) \leq M, \forall x > 0$, *for some* $M > 0$, *and* $\int_a^\infty \lambda^2(x)\mathrm{d}x = \infty$ *or*
(b) $\lambda(x) \to \infty$ *as* $x \to \infty$ *and* $|\lambda'(x)| < x^{-r}, \forall x > x_0$, *for some* $x_0 > a, r > 1/2$.

Then, as $t \to \infty$, $\eta(t)/\mu(t) \to 1$ *a.s.*

Proof Note that, for all $t > a$,

$$\frac{\mu(\lfloor t/a\rfloor a)}{\mu(\lceil t/a\rceil a)}\frac{\eta(\lfloor t/a\rfloor a)}{\mu(\lfloor t/a\rfloor a)} \leq \frac{\eta(t)}{\mu(t)} \leq \frac{\eta(\lceil t/a\rceil a)}{\mu(\lceil t/a\rceil a)}\frac{\mu(\lceil t/a\rceil a)}{\mu(\lfloor t/a\rfloor a)}, \quad \text{a.s.}$$

So, by Propositions 1 and 2, we only need to prove that $\mu(\lfloor t/a\rfloor a)/\mu(\lceil t/a\rceil a) \to 1$, as $t \to \infty$, which is equivalent to $E[Z_n]/\mu(na) \to 0$.

Suppose (a) holds then, by Lemma 2,

$$E[Z_n] = \int_0^\infty \frac{F(((n-1)a, na] \cap (x-a, x])}{\bar{F}(x)}\lambda(x)\mathrm{d}x \leq \int_{(n-1)a}^{(n+1)a} v(x)\lambda(x)\mathrm{d}x \leq 2A^2$$

and the conclusion follows. If (b) holds, by Lemma 3, we have

$$\int_{(n-1)a}^{(n+1)a} v(x)\lambda(x)\mathrm{d}x \leq C\int_{(n-1)a}^{(n+1)a} \mathrm{e}^{a\lambda(x)}\lambda(x)\mathrm{d}x \leq H\mathrm{e}^{a\lambda(na)}\lambda(na),$$

for some $H > 0$. Then, from (19) and using L'Hôpital's rule, we obtain

$$\frac{E[Z_n]}{\mu(na)} \leq H\frac{\mathrm{e}^{a\lambda(na)}\lambda(na)}{\int_a^{na} \mathrm{e}^{a\lambda(x)}\lambda(x)\mathrm{d}x} \to 0.$$

□

The last result of the paper is a central limit theorem for $\eta(t)$. As in Theorem 3, we first consider heavy- and exponential-tailed distributions and then the light-tailed ones. For the former, we impose the same condition as in Theorem 2 (essentially $\eta(\infty) = \infty$), while for the latter the condition is weaker than that for the strong law of large numbers. In particular, our central limit theorem yields a weak law of large numbers for some light-tailed distributions which are not covered by Theorem 3. We use the notation $\xrightarrow{d} N(0, \sigma^2)$ to indicate convergence in distribution to the gaussian distribution with variance σ^2.

Theorem 4 *Suppose that either*

(a) $\lambda(x) \le M, \forall x > 0$, *for some* $M > 0$, *and* $\int_a^\infty \lambda^2(x)\mathrm{d}x = \infty$ *or*
(b) $\lambda(x) \to \infty$ *as* $x \to \infty$ *and* $|\lambda'(x)| < M, \forall x > 0$, *for some* $M > 0$.

Then, as $t \to \infty$,

$$\frac{\eta(t) - \mu(t)}{\sigma(t)} \xrightarrow{d} N(0, 1). \tag{22}$$

Proof Note first that under either (a) or (b), $\mu(\infty) = \infty$ (see Remark 4). Hence, from (11), $\sigma^2(t) > \mu(t) \to \infty$, as $\to \infty$. We apply Corollary 4 of Lane (1984), since the PCP is a particular case of the Poisson shot-noise process. Our result will follow if we prove, under (a) and (b), the following Lyapunov-type condition: as $t \to \infty$,

$$L(t) := \int_0^\infty E\big[K((0, t)|x)^3\big]\lambda(x)\mathrm{d}x/\sigma^3(t) \to 0. \tag{23}$$

Observe from (8) that $L(t) \le \int_0^{t+a} v(x)^3\lambda(x)\mathrm{d}x/\sigma^3(t) + 3/\sigma(t)$ and so (23) follows if we prove

$$\int_0^{t+a} v(x)^3\lambda(x)\mathrm{d}x/\sigma^3(t) \to 0. \tag{24}$$

Note that $\sigma^2(t) \ge 2\int_0^t v^2(x)\lambda(x)\mathrm{d}x$. Then, assuming that (a) holds, from Lemma 2 we have $\int_0^{t+a} v(x)^3\lambda(x)\mathrm{d}x \le A\int_0^t v(x)^2\lambda(x)\mathrm{d}x + A^4$. In view of these bounds, we get (24).

Under (b), we get $\sigma^2(t) \ge 2A^2\int_a^t \mathrm{e}^{2a\lambda(x)}\lambda(x)\mathrm{d}x$ and also, reasoning as for inequality (21),

$$\begin{aligned}\int_a^{t+a} v(x)^3\lambda(x)\mathrm{d}x &\le C^3\left(\int_a^t \mathrm{e}^{3a\lambda(x)}\lambda(x)\mathrm{d}x + \int_t^{t+a} \mathrm{e}^{3a\lambda(x)}\lambda(x)\mathrm{d}x\right)\\ &\le C^3\left(\int_a^t \mathrm{e}^{3a\lambda(x)}\lambda(x)\mathrm{d}x + H\mathrm{e}^{3a\lambda(t)}\lambda(t)\right),\end{aligned}$$

with $H > 0$, for t large enough. Hence, for (24), we only need to show that

$$\frac{\int_a^t \mathrm{e}^{3a\lambda(x)}\lambda(x)\mathrm{d}x + H\mathrm{e}^{3a\lambda(t)}\lambda(t)}{\left(\int_a^t \mathrm{e}^{2a\lambda(x)}\lambda(x)\mathrm{d}x\right)^{3/2}} \to 0,$$

as $t \to \infty$, which is achieved by applying L'Hôpital's rule. □

Remark 5 All theorems above remain valid under the (weaker) hypothesis of f being continuous and decreasing on an interval (x^*, ∞), for some $x^* > 0$. In this situation, the integral $\int_a^\infty \lambda^2(x)\mathrm{d}x$ has to be replaced by $\int_{a+x^*}^\infty \lambda^2(x)\mathrm{d}x$ in Theorems 2–4. Also, the boundedness conditions for λ and λ' in Theorems 3 and 4 should be stated for $x > x^*$.

4 Examples

Example 1 (*Pareto distribution*) Let $r > 0$, $\bar{F}(x) = x^{-r}$ and $f(x) = rx^{-r-1}$, for $x > 1$. Since $\lambda(x) = r/x$, for $x > 1$, we have $\int_1^\infty \lambda(x)^2 \mathrm{d}x < \infty$ and, by Theorem 2, $\eta(\infty) < \infty$. We first analyze the case $r = 1$ and then $r \neq 1$.

(a) ($r = 1$) The behavior of $\eta(\infty)$ was studied in Gouet et al. (2012b) via Monte Carlo simulation. Since the geometric distribution fitted very well the simulated data, it was conjectured that the distribution of $\eta(\infty)$ was in fact geometric. Now, we answer the conjecture in the positive.
Note that $\lambda(x) = x^{-1}\mathbf{1}_{\{x>1\}}$ and $v(x) = (x-1)\mathbf{1}_{\{1<x\leq 1+a\}} + a(x-a)^{-1}\mathbf{1}_{\{x>1+a\}}$. So, the integral in (14) can be written as

$$\int_1^{a+1} \frac{(1-\alpha)(x-1)}{((1-\alpha)(x-1)+1)x}\mathrm{d}x + \int_{a+1}^\infty \frac{(1-\alpha)a}{((1-\alpha)a+x-a)x}\mathrm{d}x$$
$$= \log(1+a-a\alpha).$$

Therefore, $\varphi_\infty(\alpha) = \mathrm{e}^{-\log(1+a-a\alpha)} = \frac{(1+a)^{-1}}{1-(1+a)^{-1}\alpha}$, that is, $\eta(\infty)$ has a geometric distribution with parameter $(1+a)^{-1}$.

(b) ($r \neq 1$) In this case, the integral in (14) cannot be calculated in closed form. For $r = 2, 3, \ldots$, it is still possible to find an explicit expression but the calculations become cumbersome. For illustration's sake, we pick $r = 2$ and compute $E[\eta(\infty)] = a^2 + 4a$ and $\mathrm{Var}[\eta(\infty)] = a^4 + \frac{16}{3}a^3 + 9a^2 + 4a$. These values reveal that $\eta(\infty)$ is not geometrically distributed.

Example 2 (*Weibull distribution*) For $\alpha, \beta > 0$, let $\bar{F}(x) = \mathrm{e}^{-(x/\alpha)^\beta}$ and $\lambda(x) = \beta\alpha^{-\beta}x^{\beta-1}$, for $x > 0$. We analyze the limit behavior of $\eta(t)$ as $t \to \infty$, depending on the value of β; detailed calculations are provided only for $\beta = 1/2$. The notation $g(t) \sim h(t)$ stands for $\lim g(t)/h(t) = 1$, as $t \to \infty$.

(a) ($\beta < 1/2$) In this case, $\int_a^\infty \lambda(y)^2 \mathrm{d}y < \infty$, so $\eta(\infty) < \infty$ and the moment generating function can be obtained from Theorem 2.

(b) ($\beta = 1/2$) Here, $\lambda(y) = (\alpha y)^{-\frac{1}{2}}/2 \to 0$ and condition (a) of both Theorems 3 and 4 hold. We have $v(y) \sim \int_{y-a}^y \lambda(x)\mathrm{d}x = \alpha^{-\frac{1}{2}}\left(y^{\frac{1}{2}} - (y-a)^{\frac{1}{2}}\right) \sim a(\alpha y)^{-\frac{1}{2}}/2$ and so, $\int_0^t v(y)\lambda(y)\mathrm{d}y \sim \frac{a}{4\alpha}\log t$. Moreover,

$$\int_t^{t+a} \frac{\bar{F}(y-a) - \bar{F}(t)}{\bar{F}(y)}\lambda(y)\mathrm{d}y \leq \int_t^{t+a} v(y)\lambda(y)\mathrm{d}y \to 0, \tag{25}$$

so $\mu(t) \sim \frac{a}{4\alpha}\log t$. Further, since $v(y) \to 0$, we have $\int_0^t v(y)^2\lambda(y)\mathrm{d}y / \int_0^t v(y)\lambda(y)\mathrm{d}y \to 0$ and

$$\int_t^{t+a} \left(\frac{\bar{F}(y-a) - \bar{F}(t)}{\bar{F}(y)}\right)^2 \lambda(y)\mathrm{d}y \leq \int_t^{t+a} v(y)^2\lambda(y)\mathrm{d}y \to 0.$$

Then, by (11), $\sigma^2(t) \sim \mu(t)$. Finally, observe that, for $t > a$,

$$\left|E[N(a,t)] - \frac{a}{4\alpha}\log t\right| \le \left|\int_a^t \left(v(y) - \int_{y-a}^{y} \lambda(x)\mathrm{d}x\right)\lambda(y)\mathrm{d}y\right| + \left|\int_a^t \left(\alpha^{-\frac{1}{2}}\left(y^{\frac{1}{2}} - (y-a)^{\frac{1}{2}}\right) - a(\alpha y)^{-\frac{1}{2}}/2\right)\lambda(y)\mathrm{d}y\right| + \frac{a}{4\alpha}|\log a| + \int_t^{t+a} \frac{\bar{F}(y-a) - \bar{F}(t)}{\bar{F}(y)}\lambda(y)\mathrm{d}y. \tag{26}$$

We check that all summands in (26) are bounded (as functions of t) so they tend to 0 when divided by $\sigma(t)$; $C, C', \ldots$ are generic positive constants. For the first one note that, for $y > a$,

$$\left|v(y) - \int_{y-a}^{y} \lambda(x)\mathrm{d}x\right| = \left|\mathrm{e}^{\int_{y-a}^{y} \lambda(x)\mathrm{d}x} - 1 - \int_{y-a}^{y} \lambda(x)\mathrm{d}x\right| \le C\left(\int_{y-a}^{y} \lambda(x)\mathrm{d}x\right)^2.$$

Therefore, the first term is bounded above by $C' \int_a^t y^{-3/2}\mathrm{d}y < C''$. For the second, we have

$$\int_a^t \left|\alpha^{-\frac{1}{2}}\left(y^{\frac{1}{2}} - (y-a)^{\frac{1}{2}}\right) - a(\alpha y)^{-\frac{1}{2}}/2\right|\lambda(y)\mathrm{d}y \le C''' \int_a^t y^{-2}\mathrm{d}y < C'''/a.$$

The last term is bounded because of (25). Thus, for $\beta = 1/2$, we have $\mu(t) \sim \sigma^2(t) \sim \gamma \log t$ and $(\mu(t) - \gamma \log t)/\sigma(t) \to 0$, where $\gamma = a/(4\alpha)$, so

$$\frac{\eta(t)}{\log t} \to \gamma \quad \text{a.s.} \qquad \text{and} \qquad \frac{\eta(t) - \gamma \log t}{\sqrt{\log t}} \xrightarrow{d} N(0, \gamma).$$

(c) ($\beta \in (1/2, 1)$) As in (b), $\lambda(y) \to 0$ so condition (a) of Theorems 3 and 4 holds. We find that $\mu(t) \sim \sigma^2(t) \sim \gamma t^{2\beta-1}$, with $\gamma = a(\beta/\alpha^\beta)^2/(2\beta - 1)$, but $(\mu(t) - \gamma t^{2\beta-1})/\sigma(t) \to 0$ only if $\beta < 3/4$. In this case, we have

$$\frac{\eta(t)}{t^{2\beta-1}} \to \gamma \quad \text{a.s.} \qquad \text{and} \qquad \frac{\eta(t) - \gamma t^{2\beta-1}}{t^{\beta-1/2}} \xrightarrow{d} N(0, \gamma).$$

When $\beta \in [3/4, 1)$, a finer estimate of $\mu(t)$ is needed for the asymptotic normality. It can be shown that if $\beta \in [3/4, 5/6)$ then

$$\frac{\eta(t) - \gamma t^{2\beta-1} - \delta t^{3\beta-2}}{t^{\beta-1/2}} \xrightarrow{d} N(0, \gamma), \tag{27}$$

where $\delta = a^2(\beta/\alpha^\beta)^3/(6\beta - 4)$. The estimate of $\mu(t)$ in (27) can be further refined so that the range of validity of β gets as close to $(1/2, 1)$ as desired.

(d) ($\beta = 1$: exponential distribution) Since $\lambda(x)$ is constant, condition (a) of Theorems 3 and 4 is fulfilled. In this case, $\mu(t) = \gamma t$ and $\sigma^2(t) \sim \delta t$, with $\gamma = (\mathrm{e}^{a/\alpha} - 1)/\alpha$ and $\delta = \gamma(2\mathrm{e}^{a/\alpha} - 1)$, are readily obtained and we have

$$\frac{\eta(t)}{t} \to \gamma \quad \text{a.s.} \qquad \text{and} \qquad \frac{\eta(t) - \gamma t}{\sqrt{t}} \xrightarrow{d} N(0, \delta).$$

In this case, the process η is a Neyman–Scott process, as defined in page 662 of Cressie (1993). In fact, the sizes of the clusters are iid geometric with parameter $\mathrm{e}^{-a/\alpha}$ and the distance of a point of a cluster to its center has density

$$f_d(t) = \frac{\mathrm{e}^{t/\alpha}}{\alpha(\mathrm{e}^{a/\alpha} - 1)}, \qquad t \in [0, a], \tag{28}$$

regardless of the position of the center. Moreover, the process is stationary, since the center process is homogeneous Poisson with rate $1/\alpha$. Note also that the exponential is the only distribution where η is a Neyman–Scott process since under any other distribution the sizes of the clusters are not identically distributed.

(e) ($\beta \in (1, 3/2)$) In this case, conditions (b) of Theorems 3 and 4 hold and we find the following estimations $\mu(t) \sim \gamma t \mathrm{e}^{\delta t^{\beta-1}}$ and $\sigma^2(t) \sim \gamma t \mathrm{e}^{2\delta t^{\beta-1}}$, with $\gamma = 1/(a(\beta - 1))$ and $\delta = a\beta\alpha^{-\beta}$. However, it turns out that $(\mu(t) - \gamma t \mathrm{e}^{\delta t^{\beta-1}})/\sigma(t) \not\to 0$, which means that we need a better approximation of $\mu(t)$ to be used as centering in the central limit theorem. We do not carry out such calculation here and so we obtain

$$\frac{\eta(t)}{t\mathrm{e}^{\delta t^{\beta-1}}} \to \gamma \quad \text{a.s.} \qquad \text{and} \qquad \frac{\eta(t) - \mu(t)}{\sqrt{t}\mathrm{e}^{\delta t^{\beta-1}}} \xrightarrow{d} N(0, \gamma).$$

(f) ($\beta \in [3/2, 2)$) In this case (b) of Theorem 4 is satisfied. However, Theorem 3 does not apply and we do not have a strong convergence for $\eta(t)$. Nevertheless, from the central limit theorem, we can recover a weak law of large numbers. The estimations of $\mu(t)$ and $\sigma^2(t)$ are as in (e) and so

$$\frac{\eta(t)}{t\mathrm{e}^{\delta t^{\beta-1}}} \to \gamma \quad \text{in probability} \qquad \text{and} \qquad \frac{\eta(t) - \mu(t)}{\sqrt{t}\mathrm{e}^{\delta t^{\beta-1}}} \xrightarrow{d} N(0, \gamma).$$

(g) ($\beta = 2$) Condition (b) of Theorem 4 is satisfied and, as in (f), we do not have strong convergence, while $\mu(t)$ is calculated exactly as

$$\mu(t) = \mathrm{e}^{(a/\alpha)^2}\left(\frac{\mathrm{e}^{2at/\alpha^2}}{a}\left(t - \frac{\alpha^2}{2a}\right) + \frac{\alpha^2}{2a^2}\right) - \frac{t^2}{\alpha^2} \sim \frac{t}{a}\mathrm{e}^{(a/\alpha)^2}\mathrm{e}^{2at/\alpha^2}$$

and $\sigma^2(t) \sim \frac{t}{a}\mathrm{e}^{2(a/\alpha)^2}\mathrm{e}^{4at/\alpha^2}$. So $(\mu(t) - \frac{t}{a}\mathrm{e}^{(a/\alpha)^2}\mathrm{e}^{2at/\alpha^2})/\sigma(t) \to 0$ and we obtain

$$\frac{\eta(t)}{t\mathrm{e}^{\delta t}} \to \gamma \quad \text{in probability} \qquad \text{and} \qquad \frac{\eta(t) - \gamma t \mathrm{e}^{\delta t}}{\sqrt{t}\mathrm{e}^{\delta t}} \xrightarrow{d} N(0, a\gamma^2),$$

where $\gamma = \mathrm{e}^{(a/\alpha)^2}/a$ and $\delta = 2a/\alpha^2$.

(h) ($\beta > 2$) Theorems 3 or 4 do not apply and we have no asymptotic result for $\eta(t)$ in this situation.

Example 3 (*Exponentiality test*) We exploit structural properties of η for testing exponentiality of the parent distribution F. Given the dataset $\{x_1, \ldots, x_N\}$ of near-record values, observed in an interval $[0, t]$, we develop a procedure for testing the hypothesis that F is the exponential distribution.

This is an illustrative example intended to show the statistical applicability of structural properties of η and, consequently, the discussion of the merits and technical details of the proposed strategy are kept to a minimum. The reader interested in methods for assessing the fit of a point process model can consult, for instance, Diggle (1983), Karr (1991) and Cressie (1993).

As mentioned in Example 2(d), η is a Neyman–Scott process if and only if $\bar{F}(x) = \mathrm{e}^{-x/\alpha}$, $x > 0$, for some $\alpha > 0$. Hence, the quality of the fit of the Neyman–Scott model to the data is informative about the exponentiality of F. To assess the fit of the theoretical model, we use a Cramér-Von Mises-type statistic, based on Ripley's $K(h)$ function, given by

$$D = \int_0^{h_0} \left(K(h)^{0.5} - \hat{K}(h)^{0.5}\right)^2 \mathrm{d}h, \tag{29}$$

where

$$\hat{K}(h) = \frac{t}{N^2} \sum_{i \neq j} w^{-1}(x_i, x_j) \mathbf{1}_{\{|x_i - x_j\} \leq h\}},$$

is Ripley's estimator, with w a weight function defined to minimize edge effects. See section 5.3 in Diggle (1983) and sections 8.4, 8.5 in Cressie (1993) for further information and discussion. An explicit formula for $K(h)$ is obtained by adapting (8.5.39) in Cressie (1993) to our case, since the sizes of the clusters are geometric and the distance of a point to its parent has density given in (28), yielding

$$K(h) = \begin{cases} 2h + 2\alpha \frac{1-\mathrm{e}^{-h/\alpha} + \mathrm{e}^{-2a/\alpha}(1-\mathrm{e}^{h/\alpha})}{(1-\mathrm{e}^{-a/\alpha})^2} & \text{for } 0 \leq h \leq a, \\ 2h + 2\alpha & \text{for } h > a. \end{cases}$$

For the practical implementation of the proposed test, we require the null distribution of D, which is not known, but can be approximated by means of Monte Carlo simulation. Also, the parameter α of the exponential model has to be estimated, and this is done by

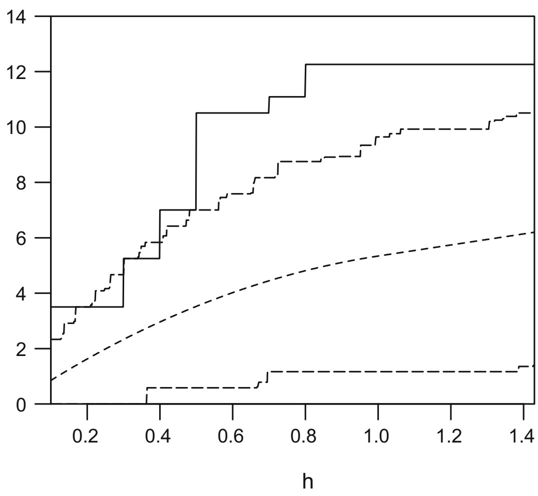

Fig. 1 Values of Ripley's K function for Wilmington snowfall data. *Solid line* $\hat{K}(h)$; *dotted line* K(h); *dashed lines* 5, 95 confidence bounds

equating the expected number of points of η in $[0, t]$ with the number N of data points actually observed; that is, α is estimated as the unique solution of $(\mathrm{e}^{a/\alpha} - 1)t/\alpha = N$, see Example 2(d).

For illustration, we consider the Wilmington NC snowfall database, recording snowfall measurements in inches from 1870, available at www.weather.gov. The series has 114 observations and, setting $a = 1$, there are 7 near-record values given by 1.5, 2, 1.5, 1.6, 2, 2 and 2.3. Also, based on the observed data, we can pick $t = 14.3$ and $h_0 = 0.143$, and formula (29) yields the value $D = 1.56$. On the other hand, from simulations under the null hypothesis, we obtain the 5 % critical value $D^* = 1.55$ and thus, the claim of exponentiality for the snowfall dataset is rejected, although not by much. A better understanding of the poor fit of the Neyman–Scott model to the snowfall data can be obtained from Fig. 1, where K and $\hat{K}$ are plotted along with confidence bounds for the fitted model estimated from simulations.

Acknowledgments We thank the editor and the referees for their useful comments and suggestions. Financial support from project PFB-03-CMM, Fondecyt grant 1120408 and project MTM2010-15972 of MINECO is gratefully acknowledged. The authors are members of the research group Modelos Estocásticos (DGA).

Appendix

Lemma 5 *Let (Y_n) be a sequence of centered, square-integrable, m-dependent random variables and (b_n) a sequence of real numbers growing to ∞. Then, $\sum \frac{E[Y_n^2]}{b_n^2} < \infty$ implies $\frac{1}{b_n}\sum_{k=1}^n Y_k \to 0$ a.s.*

Proof For $1 \le k \le m+1$, $\left(\frac{Y_{k+(m+1)n}}{b_{k+(m+1)n}}\right)_{n\ge 0}$ is a sequence of independent random variables. By hypothesis $\sum_{n\ge 0} \frac{E[Y_{k+(m+1)n}^2]}{b_{k+(m+1)n}^2} < \infty$ and so, by the Khintchin–Kolmogorov convergence theorem (Chow and Teicher 1988, p. 113), $\sum_{n\geq 0} \frac{Y_{k+(m+1)n}}{b_{k+(m+1)n}} < \infty$ a.s. Adding up the series for $k = 1, \ldots, m+1$, we obtain $\sum_{n\geq 1} \frac{Y_n}{b_n} < \infty$ a.s. and the conclusion follows from Kronecker's lemma. □

References

Arnold BC, Balakrishnan N, Nagaraja HN (1998) Records. Wiley, New York

Bairamov I, Stepanov A (2011) Numbers of near bivariate record-concomitant observations. J Multivar Anal 102:908–917

Balakrishnan N, Pakes AG, Stepanov A (2005) On the number and sum of near-record observations. Adv Appl Probab 37:765–780

Bose A, Gangopadhyay S (2011) Asymptotic properties of near Pfeifer records. Extremes 14:253–265

Chow YS, Teicher H (1988) Probability theory: independence, interchangeability and martingales, 2nd edn. Springer, New York

Cressie N (1993) Statistics for spatial data. Wiley, New York

Diggle PJ (1983) Statistical analysis of spatial point patterns. Academic Press, London

Gouet R, López FJ, Sanz G (2012a) Central limit theorem for the number of near-records. Commun Stat Theory Methods 41:309–324

Gouet R, López FJ, Sanz G (2012b) Limit theorems for the counting process of near-records. Commun Stat Simul Comput 41:820–832

Gouet R, López FJ, Sanz G (2012c) On δ-record observations: asymptotic rates for the counting process and elements of maximum likelihood estimation. Test 21:188–214

Gulati S, Padgett WJ (2003) Parametric and nonparametric inference from record-breaking data. Lect Notes in Statist 172, Springer

Karr A (1991) Point processes and their statistical inference. Marcel Dekker, New York

Lane JA (1984) The central limit theorem for the Poisson shot-noise process. J Appl Probab 21:287–301

López-Blázquez F, Salamanca-Miño B (2013) Distribution theory of δ-record values. Case $\delta < 0$. Test 22:715–738

Pakes AG (2007) Limit theorems for numbers of near-records. Extremes 10:207–224

Daley DJ, Vere-Jones D (2003) An introduction to the theory of point processes, vol. I elementary theory and methods, 2nd edn. Springer, New York