Test (2012) 21:188–214
DOI 10.1007/s11749-011-0242-6

ORIGINAL PAPER

# On $\delta$-record observations: asymptotic rates for the counting process and elements of maximum likelihood estimation

**Raúl Gouet · F. Javier López · Gerardo Sanz**

Received: 25 September 2009 / Accepted: 24 March 2011 / Published online: 9 April 2011


**Abstract** Given a sequence of random variables $\{X_n, n \geq 1\}$ and $\delta \in \mathbb{R}$, an observation $X_n$ is a $\delta$-record if $X_n > \max\{X_1, \dots, X_{n-1}\} + \delta$. We obtain, for $\delta \leq 0$, weak and strong laws of large numbers for the counting process of $\delta$-records among the first $n$ observations from a sequence of independent identically distributed random variables, with common distribution $F$, possibly discontinuous. We provide examples of our results in the context of common probability distributions. Finally, we show how $\delta$-records can be used for maximum likelihood estimation.

**Keywords** Record · Near-record · Extreme · Law of large numbers

**Mathematics Subject Classification (2000)** Primary 60G70 · Secondary 60F15 · 62F10

## 1 Introduction

Given a sequence of random variables $\{X_n, n \geq 1\}$, an observation $X_n$ is a record if it is greater than all previous observations; that is, if $X_n > M_{n-1}$, where $M_n =$

In Memoriam. Miguel San Miguel.

R. Gouet
Dpto. Ingeniería Matemática y Centro de Modelamiento Matemático (UMI 2807, CNRS), Universidad de Chile, Av. Blanco Encalada 2120, 837-0459 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: gerardo.sanz@unizar.es

F.J. López
e-mail: javier.lopez@unizar.es

$\max\{X_1, \ldots, X_n\}$. On the other hand, $X_n$ is said to be a $\delta$-record if $X_n > M_{n-1} + \delta$, where $\delta$ is a fixed real parameter.

The mathematical theory of records is a well-established branch of extreme value theory, with interesting applications; see Arnold et al. (1998), Nevzorov (2001) for general theory and applications of records and Gulati and Padgett (2003) for inferential procedures based on record-breaking data. In recent years generalizations of records, usually known as near-records, have attracted interest because of their relevance in, among other fields, stochastic models of large insurance claims; see Balakrishnan et al. (2005), Hashorva (2003) or Hashorva and Hüsler (2005). There is however no generally agreed definition of near-record but, among those proposed in the literature, $\delta$-records seem to be a natural and tractable alternative which includes, as particular cases, usual records ($\delta = 0$) and weak records ($\delta = -1$), for integer-valued observations. Besides being a natural generalization of records, $\delta$-records are related to other existing concepts such as $\delta$-exceedance records (Balakrishnan et al. 1996), $\epsilon$-repeated records (Khmaladze et al. 1997) and near-records (Balakrishnan et al. 2005; Pakes 2007).

In the present paper we obtain laws of large numbers for $N_n$, the number of $\delta$-records among the first $n$ observations from a sequence of nonnegative, independent and identically distributed (iid) random variables, with common distribution $F$, in the case of $\delta \leq 0$. Our results are a natural complement of the central limit theorem for $N_n$ given in Gouet et al. (2007) but here, in order to widen the field of potential applications, we allow $F$ to be general, unlike in Gouet et al. (2007) where observations are integer valued or in Balakrishnan et al. (1996), Balakrishnan et al. (2005), Khmaladze et al. (1997) and Pakes (2007), cited above, where $F$ is assumed to be continuous.

It is well known that the growth rate of usual records for continuous distributions is $\log n$, regardless of $F$, while in the discrete case the rate depends on the tail of $F$. Our results show that the growth of the number of $\delta$-records $N_n$ depends on the tail behavior of $F$, both for continuous and discrete distributions. In fact, when $F$ is heavy-tailed (heavier than exponential), the growth rate of $N_n$ is $\log n$, the same than usual records; when $F$ has moderately heavy tails (such as exponential) the growth rate is proportional to $\log n$, the factor depending both on $F$ and $\delta$; when $F$ is light-tailed (lighter than exponential) the growth rate of $N_n$ is faster than $\log n$, the specific rate depending both on $F$ and $\delta$.

Some particular cases of our results are worth highlighting. First, taking $\delta = 0$ we have a new strong law of large numbers for the number of records when $F$ has moderately heavy tails; notice that, besides the classical continuous case (Rényi 1962), only results for discrete (Gouet et al. 2001; Key 2005) and heavy-tailed distributions (Proposition 3.1 in Gouet et al. 2001) have been reported. On the other hand, in the case of light-tailed continuous distributions and $\delta < 0$, a direct application of our results complements those of Balakrishnan et al. (2005) and Pakes (2007) on the growth of near-records (see Remark 4).

For negative $\delta$, an observation is a $\delta$-record if it is a record or a near-record in the sense of Balakrishnan et al. (2005). Near-records are potentially useful in several fields such as insurance theory (see Balakrishnan et al. 2005 and references therein) and industrial stress-testing. It is therefore interesting to analyze how these data can be used in statistical inference (statistical inference based on records values has been extensively studied; see for instance Arnold et al. 1998 and Gulati and Padgett 2003). We provide a maximum likelihood estimation procedure for the parameters of the distribution of the observations, when $\delta$-records are used, and explore its potential in examples based on real and simulated data.

The paper is organized as follows: in Sect. 2 we give some notation and preliminaries, relating $N_n$ to the sum of partial minima of random variables; in Sect. 3 we first state and prove our main result (Theorem 1) giving conditions for the strong and the weak law of large numbers for $N_n$. Then we apply the result to general distributions with heavy or moderately heavy tails (Sect. 3.1) or light tails (Sect. 3.2) including, as examples, the most common discrete and continuous distributions. The last section is devoted to maximum likelihood estimation based on $\delta$-record values.

## 2 Notation and preliminaries

Let $F$ be a distribution function concentrated on $\mathbb{R}_+ = [0, \infty)$. We define $\overline{F}(x) = 1 - F(x)$, $F^-(x) = \lim_{t \uparrow x} F(t)$, for $x \geq 0$, and the generalized inverse $F^{\leftarrow}(y) = \inf\{x \geq 0 : F(x) \geq y\}$, for $y \in [0, 1)$. Note that, for all $a \in \mathbb{R}$, $b \in (0, 1)$, $F(a) \geq b$ if and only if $a \geq F^{\leftarrow}(b)$.

Sequences are denoted by $\{a_n, n \geq 1\}$ or $\{a_n\}$ but braces are sometimes omitted for simplicity. Convergence (divergence) of $a_n$ to a finite (infinite) limit $a$, as $n \to \infty$, is denoted $\lim_n a_n = a$, $a_n \to a$ or $a_n \uparrow a$ when $a_n$ is increasing. We write $a_n \sim b_n$ if either both sequences diverge to infinity or converge to zero, with $\lim_n a_n/b_n = 1$, or both converge to nonzero, possibly different, finite limits. The $O(\cdot)$ and $o(\cdot)$ notations have their usual meanings. Analogous conventions are used for real functions defined on $[0, \infty)$, regarding limits and the $\sim$, $O(\cdot)$, $o(\cdot)$ notations. In general, limits for real functions are taken as the argument goes to infinity. For the floor of $x$ (greatest integer less than or equal to $x$) we use the notation $\lfloor x \rfloor$; for the indicator function we use the symbol $\mathbf{1}_{\{\cdot\}}$.

Random variables are defined on a common space $(\Omega, \mathcal{F}, P)$. Probabilities and expectations are denoted respectively by $P[\cdot]$ and $E[\cdot]$; conditional probabilities and conditional expectations with respect to a sub $\sigma$-algebra $\mathcal{G} \subseteq \mathcal{F}$ are denoted by $P[\cdot|\mathcal{G}]$ and $E[\cdot|\mathcal{G}]$. For sequences of random variables, convergence in distribution, in probability and almost sure (a.s.) are denoted respectively by superscripted arrows $\xrightarrow{D}$, $\xrightarrow{P}$ and $\xrightarrow{\text{a.s.}}$. The $\sim$, $O(\cdot)$, $o(\cdot)$ notations are understood in the a.s. sense.

In the sequel, $\{X_n\}$ stands for a sequence of nonnegative, independent and identically distributed (iid) random variables, with common distribution function $F$, which is assumed to have an infinite right endpoint, that is, $\overline{F}(x) > 0$, for all $x \geq 0$. Notice that there is no loss of generality in assuming the $X_n$ nonnegative since we deal with upper extremes.

For $\delta \leq 0$, let $I_1 = 1$ and $I_n = \mathbf{1}_{\{X_n > M_{n-1} + \delta\}}$, $n \geq 2$, be the indicators of $\delta$-records and $N_n = \sum_{i=1}^n I_i$, $n \geq 1$, the counting process of $\delta$-records among $X_1, \ldots, X_n$, where $M_n = \max\{X_1, \ldots, X_n\}$. Let $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$, $n \geq 1$, and $\mathcal{F}_0 = \{\emptyset, \Omega\}$. It is easy to see that $E[I_n|\mathcal{F}_{n-1}] = P[X_n > M_{n-1} + \delta|\mathcal{F}_{n-1}] = \overline{F}(M_{n-1} + \delta)$, a.s., for $n \geq 2$. But, since $\overline{F}$ is decreasing, we have $\overline{F}(M_n + \delta) = \min\{Y_1, \ldots, Y_n\}$, where $Y_n = \overline{F}(X_n + \delta)$, $n \geq 1$.

On the other hand, it is well known (see Corollary VII-2-6 in Neveu 1972) that sums of indicators and sums of their conditional expectations, with respect to an increasing family of $\sigma$-fields, are a.s. asymptotically equivalent in the sense that they either both converge or diverge simultaneously; besides, under divergence their ratio tends a.s. to 1. Since $\delta \leq 0$, there are more $\delta$-records than records whose number is unbounded because $F(x) < 1$ for all $x$. Therefore $N_n \uparrow \infty$ a.s. and $N_n = \sum_{i=1}^n I_i \sim \sum_{i=1}^n E[I_i | \mathcal{F}_{i-1}]$ a.s. We have

**Lemma 1** *Let $N_n$, $Y_n$ be as defined above and $S_n = \sum_{k=1}^n \min\{Y_1, \ldots, Y_k\}$, $n \geq 1$. Then, for every $\delta \leq 0$,*

$$N_n / S_n \xrightarrow{\text{a.s.}} 1. \tag{1}$$

Convergence in (1) means that the law of large numbers for $N_n$ can be obtained from the corresponding result for sums of partial minima of nonnegative iid random variables. The asymptotic behavior of this process is well known. Deheuvels (1974) established weak and strong convergence results which are useful here.

In what follows, $X$ denotes a generic nonnegative random variable with distribution $F$. For $t \geq 1$ define $m(t) = \sup\{x \geq 0 : P[X \geq x] \geq 1/t\}$.

It is clear that, since $P[X \geq x] = 1 - F^-(x)$,

$$1 - F^-\big(m(t)\big) \geq 1/t \quad \text{and} \quad 1 - F^-\big(m(t) + \epsilon\big) < 1/t, \tag{2}$$

for all $t \geq 1$, $\epsilon > 0$. Formulas for the distribution function of $\overline{F}(X + \delta)$ and its generalized inverse are given in the following lemma.

**Lemma 2** *Let $\delta \leq 0$ and $G(y) = P[\overline{F}(X + \delta) \leq y]$. Then*

(a) $G(y) = 1 - F^-(F^{\leftarrow}(1 - y) - \delta)$, $y \in (0, 1)$,
(b) $G^{\leftarrow}(z) = \overline{F}(m(1/z) + \delta)$, $z \in (0, 1)$.

*Proof* Immediate. □

The function

$$H(x) = \int_1^{e^x} G^{\leftarrow}(1/t)\, dt, \quad x \geq 0, \tag{3}$$

plays a key role in our main result. Observe that since $G^{\leftarrow}$ is increasing, as $n \to \infty$,

$$H(\log n) = \int_1^n G^{\leftarrow}(1/t)\, dt \sim \sum_{k=2}^n G^{\leftarrow}(1/k). \tag{4}$$

Two useful formulas for $H$ are presented in Lemmas 3 and 4, for discrete and absolutely continuous distributions respectively.

### 2.1 Discrete distributions

Consider $X$ with distribution $F$ concentrated on $\mathbb{Z}_+ := \{0, 1, \ldots\}$ and let $p_k = P[X = k] > 0$, $y_k = P[X > k]$ and $r_k = P[X = k | X \geq k] = p_k / y_{k-1}$ (the discrete hazard rate), for $k \in \mathbb{Z}_+$. Let also $p_k = r_k = 0$ and $y_k = 1$, for $k < 0$. Then $F(x) = 1 - y_k$, for $k \leq x < k + 1$ and $F^{\leftarrow}(y) = k$, for $1 - y_{k-1} < y \leq 1 - y_k$, with $k \in \mathbb{Z}_+$. Observe that $m(t)$ takes integer values, for $t \geq 1$, and is characterized by $y_{m(t)} < 1/t \leq y_{m(t)-1}$.

For each $\delta \in \mathbb{Z}_- := \{0, -1, -2, \ldots\}$, the random variable $\overline{F}(X + \delta)$ takes values $y_{k+\delta}$ with probabilities $p_k$, for $k \geq -\delta$ and the value 1 with probability $\sum_{i<-\delta} p_i$. The corresponding distribution function $G$ and its generalized inverse $G^{\leftarrow}$ are given by $G(y) = y_{k-1}$, for $y_{k+\delta} \leq y < y_{k+\delta-1}$, $k \geq -\delta$, and $G^{\leftarrow}(z) = y_{k+\delta}$, for $y_k < z \leq y_{k-1}$, $k \geq 0$, $z \in (0, 1)$. Notice that, by Lemma 2(b),

$$G^{\leftarrow}(1/t) = y_{m(t)+\delta}, \quad t > 1. \tag{5}$$

**Lemma 3** *Let $F$ be concentrated on $\mathbb{Z}_+$ and $\delta \in \mathbb{Z}_-$. Then, for $t \geq 1$,*

$$H(\log t) = \sum_{k=0}^{m(t)} y_{k+\delta}\, r_k / y_k - \rho(t), \tag{6}$$

*where* $0 \leq \rho(t) := y_{m(t)+\delta}(y_{m(t)}^{-1} - t) \leq y_{m(t)+\delta}\, r_{m(t)} / y_{m(t)} \leq y_{m(t)+\delta} / y_{m(t)}$.

*Proof* From (3) and (5),

$$\begin{aligned} H(\log t) &= \int_1^t y_{m(x)+\delta}\, dx = \sum_{k=0}^{m(t)} y_{k+\delta}\left(y_k^{-1} - y_{k-1}^{-1}\right) - y_{m(t)+\delta}\left(y_{m(t)}^{-1} - t\right) \\ &= \sum_{k=0}^{m(t)} y_{k+\delta}\, r_k / y_k - \rho(t). \end{aligned}$$

The inequalities for $\rho(t)$ follow from $y_{m(t)} < 1/t \leq y_{m(t)-1}$ and $r_k \leq 1$. □

### 2.2 Absolutely continuous distributions

Let $X$ be nonnegative and absolutely continuous, with distribution $F$ and density $f$, such that $\overline{F}(x) > 0$, for $x \geq 0$. The hazard and the cumulative hazard functions are defined respectively by

$$\lambda(x) = f(x) / \overline{F}(x) \quad \text{and} \quad \Lambda(x) = -\log \overline{F}(x) = \int_0^x \lambda(t)\, dt,$$

for $x \geq 0$.

Assuming, moreover, that $f$ is strictly positive on $(0, \infty)$, we have $m(t) = F^{-1}(1 - 1/t)$, for $t \geq 1$, where $F^{-1}$ denotes the usual inverse of $F$ (restricted to $[0, \infty)$, which exists since $F$ is strictly increasing). Clearly, $\overline{F}(m(t)) = 1/t$. Thus, for

$\delta < 0$, $\overline{F}(X+\delta)$ takes values in $(0, 1]$, with distribution $G(y) = \overline{F}(F^{-1}(1-y)-\delta)$, $y \in (0, 1)$, and generalized inverse given by $G^{\leftarrow}(z) = \overline{F}(m(1/z)+\delta) = \overline{F}(F^{-1}(1-z)+\delta)$, $z \in (0, 1)$.

**Lemma 4** *Let $F$ be concentrated on $[0,\infty)$, with strictly positive density $f$ and hazard function $\lambda$, and let $\delta \le 0$. Then, for $t \ge 1$,*

$$H(\log t) = \int_0^{m(t)} \lambda(u) e^{-\int_u^{u+\delta} \lambda(v)\,dv}\,du. \tag{7}$$

*Proof* Since $f$ is strictly positive, the function $m$ is differentiable and its derivative is calculated by differentiation of the identity $\overline{F}(m(x)) = 1/x$, for $x > 1$:

$$m'(x) = \frac{1}{x\lambda(m(x))}. \tag{8}$$

A simple change of variable in (3) yields

$$H(\log t) = \int_1^t \overline{F}\big(m(x)+\delta\big)\,dx = \int_0^{m(t)} \lambda(u)\frac{\overline{F}(u+\delta)}{\overline{F}(u)}\,du. \qquad \square$$

## 3 Main results

In the following theorem we establish the weak and strong laws of large numbers for $N_n$, the counting process of $\delta$-records, using Deheuvels' results and Lemma 1.

**Theorem 1** *Let $\{X_n\}$ be an iid sequence with common (general) distribution function $F$ such that $\overline{F}(x) > 0$, for $x \ge 0$, and let $\delta \le 0$.*

(a) *If*

$$\lim_{x\to\infty} H(x+\log x)/H(x) = 1 \quad \textit{and} \tag{9}$$

$$\sum_{n=2}^{\infty}\left[ nG^{\leftarrow}(1/n)^2 \Big/ \left(\sum_{k=2}^{n} G^{\leftarrow}(1/k)\right)^2\right] < \infty \tag{10}$$

*hold, then*

$$N_n/H(\log n) \xrightarrow{\text{a.s.}} 1.$$

(b) *If there exists a sequence of real numbers $x_n \uparrow \infty$ such that*

$$\lim_n H(x_n+\log n)/H(\log n) = 1 \quad \textit{and} \tag{11}$$

$$\lim_n \sum_{k=2}^{n} kG^{\leftarrow}(1/k)^2 \Big/ \left(\sum_{k=2}^{n} G^{\leftarrow}(1/k)\right)^2 = 0 \tag{12}$$

*hold, then*

$$N_n/H(\log n) \xrightarrow{P} 1.$$

*Proof* First note that $\lim_{t\to\infty} H(t)=\infty$ because otherwise, by Theorem 7 in Deheuvels (1974), we have $\lim_{n\to\infty} S_n<\infty$ and then, the conditional Borel–Cantelli lemma (see comments before Lemma 1) yields $\lim_{n\to\infty} N_n<\infty$, which is a contradiction. Now Corollary 4 in Deheuvels (1974) gives $S_n/H(\log n)\xrightarrow{\text{a.s.}}1$ and $S_n/H(\log n)\xrightarrow{P}1$ under (a) and (b) respectively. Finally, Lemma 1 allows us to replace $S_n$ by $N_n$. □

In the next subsection we consider families of distributions $F$ in terms of their tail properties and apply Theorem 1 to obtain the asymptotic behavior of $N_n$ by studying conditions (9)–(12).

### 3.1 Heavy and exponential-like tails

**Definition 1** (a) $F$ is in the class $\mathcal{OL}$ ($F\in\mathcal{OL}$) if

$$0<\liminf_{x\to\infty}\frac{\overline{F}(x+y)}{\overline{F}(x)}\le\limsup_{x\to\infty}\frac{\overline{F}(x+y)}{\overline{F}(x)}<\infty,\qquad y\in\mathbb{R}. \tag{13}$$

(b) $F$ is long-tailed ($F\in\mathcal{L}$) if

$$\lim_{x\to\infty}\frac{\overline{F}(x+y)}{\overline{F}(x)}=1,\qquad y\in\mathbb{R}. \tag{14}$$

(c) $F$ has exponential-like tails ($F\in\mathcal{L}_\alpha$) if there exists $\alpha>0$ such that

$$\lim_{x\to\infty}\frac{\overline{F}(x+y)}{\overline{F}(x)}=e^{-\alpha y},\qquad y\in\mathbb{R}, \tag{15}$$

if $F$ is non-lattice and

$$\lim_{n\to\infty}\frac{\overline{F}(n+1)}{\overline{F}(n)}=e^{-\alpha},$$

if $F$ is concentrated on $\mathbb{Z}_+$ (lattice of span 1).

*Remark 1* If $F\in\mathcal{OL}$ then $e^{-\alpha x}/\overline{F}(x)\to 0$, for some $\alpha>0$ while if $F\in\mathcal{L}$ then $e^{-\alpha x}/\overline{F}(x)\to 0$, for all $\alpha>0$. Clearly $\mathcal{L}\subset\mathcal{OL}$ and $\mathcal{L}_\alpha\subset\mathcal{OL}$. The notion of long tail is also known as moderate growth; see Embrechts et al. (1997). Common distributions belonging to the $\mathcal{L}$ class are zeta, Pareto, log-normal, Cauchy and Weibull (with shape parameter less than 1) distributions. On the other hand, the exponential, gamma, geometric and negative binomial distributions belong to $\mathcal{L}_\alpha$. See Examples 1 and 2.

**Theorem 2** *Let $F\in\mathcal{OL}$ and $\delta\le 0$. Then $H(\log n)=O(\log n)$ and $\frac{N_n}{H(\log n)}\xrightarrow{\text{a.s.}}1$.*

*Proof* First, note that $(1-F^-(x))/\overline{F}(x)\le\overline{F}(x-1)/\overline{F}(x)$, for all $x>0$. Thus, from (13), there exists $a>0$ such that, for all $x>0$,

$$1\le\frac{1-F^-(x)}{\overline{F}(x)}\le a.$$

Then, from (2), (13) and Lemma 2(b), there exist $A, B > 0$ such that, for all $t > 1$, $\epsilon > 0$,

$$A < \frac{\overline{F}(m(t)+\delta)}{1-F^{-}(m(t))} \leq tG^{\leftarrow}(1/t) < \frac{\overline{F}(m(t)+\delta)}{1-F^{-}(m(t)+\epsilon)} < B. \tag{16}$$

This implies $H(\log n) = O(\log n)$ by (4). On the other hand, for all $x > 1$, (16) implies

$$0 < \frac{H(x+\log x) - H(x)}{H(x)} = \frac{\int_{e^x}^{xe^x} G^{\leftarrow}(1/t)\,dt}{\int_1^{e^x} G^{\leftarrow}(1/t)\,dt} < \frac{B\log x}{Ax},$$

and (9) follows. For (10) note that (16) implies $G^{\leftarrow}(1/n) = O(1/n)$ and $\sum_{k=2}^{n} G^{\leftarrow}(1/k) = O(\log n)$. Hence (10) is obtained from the convergence of the series $\sum_{n=2}^{\infty}(n(\log n)^2)^{-1}$. The conclusion follows from Theorem 1(b). □

**Corollary 1** *Let $F \in \mathcal{L}$ and $\delta \leq 0$. Then $N_n/\log n \xrightarrow{\text{a.s.}} 1$.*

*Proof* Since $\mathcal{L} \subset \mathcal{OL}$, Theorem 2 applies. By (2) and Lemma 2(b),

$$\frac{\overline{F}(m(t)+\delta)}{1-F^{-}(m(t))} \leq tG^{\leftarrow}(1/t) < \frac{\overline{F}(m(t)+\delta)}{1-F^{-}(m(t)+\epsilon)} \tag{17}$$

for all $t > 1$, $\epsilon > 0$, and the left and right sides above converge to 1 by (14). Therefore $G^{\leftarrow}(1/t) \sim 1/t$, as $t \to \infty$, and $H(\log n) \sim \log n$ by (4). □

*Remark 2* Observe that for $F \in \mathcal{L}$ the growth rate of $\delta$-record counts is $\log n$, so, in this case, $\delta$-records for $\delta < 0$ appear as frequently as records ($\delta = 0$). We stress that the above result applies to general distributions (not only discrete or continuous). In the discrete or continuous cases we have the next corollary.

**Corollary 2** (a) *Let $F$ be concentrated on $\mathbb{Z}_+$, with hazard rates $r_k \to 0$, and let $\delta \in \mathbb{Z}_-$. Then $N_n/\log n \xrightarrow{\text{a.s.}} 1$.*

(b) *Let $F$ be absolutely continuous, with hazard function $\lambda(x) \to 0$, and let $\delta \leq 0$. Then $N_n/\log n \xrightarrow{\text{a.s.}} 1$.*

*Proof* (a) For $x \geq 0$, $\overline{F}(x) = \overline{F}(\lfloor x \rfloor) = \prod_{k=0}^{\lfloor x \rfloor}(1-r_k)$. Then, for every $y \in \mathbb{R}$, $\overline{F}(x+y)/\overline{F}(x) \to 1$ as $x \to \infty$ since $r_k \to 0$. Hence, $F \in \mathcal{L}$ and the conclusion follows from Corollary 1.

(b) The result follows immediately from

$$\overline{F}(x+y)/\overline{F}(x) = \exp\left(-\int_x^{x+y} \lambda(t)\,dt\right) \to 1$$

as $x \to \infty$, and Corollary 1. □

*Example 1* Consider the following distributions in $\mathcal{L}$: the zeta distribution on $\mathbb{Z}_+$, with $\overline{F}(x) = C(\lfloor x \rfloor + 1)^{-a}, x \geq 0, \ a > 0$; the Pareto distribution, with density $f(x) = K/x^r$, for $x \geq a$, $a > 0$, $r > 1$; and the heavy-tailed Weibull distribution, with $f(x) = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^\alpha}$, for $x \geq 0$, $\alpha \in (0, 1)$ and $\beta > 0$. In all the cases we obtain $N_n/\log n \xrightarrow{\text{a.s.}} 1$.

We now turn our attention to exponential-like distributions.

**Corollary 3** *Let $F \in \mathcal{L}_\alpha$ non-lattice and $\delta \leq 0$. Then $N_n/\log n \xrightarrow{\text{a.s.}} e^{-\alpha\delta}$. In particular, if $F$ is absolutely continuous, with hazard function $\lambda(x) \to a \in (0, +\infty)$, then $N_n/\log n \xrightarrow{\text{a.s.}} e^{-a\delta}$.*

*Proof* Since $\mathcal{L}_\alpha \subset \mathcal{OL}$, Theorem 2 applies. The left- and right-hand sides of (17) tend respectively to $e^{-\alpha\delta}$ and $e^{-\alpha(\delta-\epsilon)}$, by (15). Therefore $G^{\leftarrow}(1/t) \sim e^{-\alpha\delta}/t$ and $H(\log n) \sim e^{-\alpha\delta}\log n$, by (4).

If $F$ is absolutely continuous, from $\overline{F}(x + y)/\overline{F}(x) = \exp(-\int_x^{x+y} \lambda(t)\,dt) \to e^{-ay}$, as $x \to \infty$, we obtain that $F \in \mathcal{L}_\alpha$, with $\alpha = a$. □

**Corollary 4** *Let $F$ be concentrated on $\mathbb{Z}_+$, with hazard rates $r_k \to r \in (0, 1)$, and let $\delta \in \mathbb{Z}_-$. Then $N_n/\log n \xrightarrow{\text{a.s.}} -r(1-r)^\delta/\log(1-r)$.*

*Proof* $F$ is lattice with span 1. For $x \geq 0$, $\overline{F}(x) = \overline{F}(\lfloor x \rfloor) = \prod_{k=0}^{\lfloor x \rfloor}(1 - r_k)$. Then, as $x \to \infty$, $\overline{F}(x+1)/\overline{F}(x) = 1 - r_{\lfloor x \rfloor+1} \longrightarrow 1 - r$, so $F \in \mathcal{L}_\alpha$, with $\alpha = -\log(1-r)$, and $N_n/H(\log n) \xrightarrow{\text{a.s.}} 1$ by Theorem 2. Also, from (6), we have

$$H(\log n) \sim \sum_{k=0}^{m(n)} \frac{r_k y_{k+\delta}}{y_k} \sim r(1-r)^\delta m(n).$$

The result follows from $m(n) \sim -\log n/\log(1-r)$; see Proposition 3.3 in Gouet et al. (2001). □

*Example 2* (a) The exponential distribution with parameter $\mu > 0$ is in $\mathcal{L}_\alpha$ since $\lambda(x) = \mu$, for all $x \geq 0$. The same conclusion applies to the gamma distribution, with density $f(x) = \mu^p e^{-\mu x}x^{p-1}/\Gamma(p)$, for $x \geq 0, \mu, p > 0$, since $\lambda(x) \to \mu$. In both cases $N_n/\log n \xrightarrow{\text{a.s.}} e^{-\mu\delta}$, for $\delta \leq 0$.

(b) The geometric distribution on $\mathbb{Z}_+$, with parameter $p \in (0, 1)$, is in $\mathcal{L}_\alpha$ since $r_k = p$, for all $k \in \mathbb{Z}_+$. The negative binomial distribution, with $p_k = (-1)^k\binom{-a}{k}p^a(1-p)^k$, for $k \geq 0, \ p \in (0, 1)$ and $a > 1$, is also in $\mathcal{L}_\alpha$ since $r_k \to p$. In both cases we obtain $N_n/\log n \xrightarrow{\text{a.s.}} -p(1-p)^\delta/\log(1-p)$, for $\delta \in \mathbb{Z}_-$.

## 3.2 Light tails

**Definition 2** $F$ is light-tailed if there exists $a > 0$ such that

$$\lim_{x\to\infty} \frac{\overline{F}(x+t)}{\overline{F}(x)} = 0, \quad t \geq a.$$

For these distributions we do not have a result as general as Theorem 2, valid for all $F \in \mathcal{OL}$. We consider discrete and absolutely continuous distributions separately.

**Lemma 5** *Let $F$ be concentrated on $\mathbb{Z}_+$ with hazard rates $r_k \to 1$. Then $F$ is light-tailed.*

*Proof* It is clear that, for $x \geq 0$ and $t \geq 1$,

$$\frac{\overline{F}(x+t)}{\overline{F}(x)} \leq \frac{y_{\lfloor x \rfloor +1}}{y_{\lfloor x \rfloor}} = 1 - r_{\lfloor x \rfloor +1} \to 0,$$

and the result is proved. □

*Remark 3* Common distributions with light tails are the normal, Weibull (with shape parameter greater than 1) and Poisson distributions. See Examples 3 and 4.

**Theorem 3** *Let $F$ be concentrated on $\mathbb{Z}_+$, with hazard rates $r_k \to 1$, and $\delta \in \mathbb{Z}_- \setminus \{0\}$ (the negative integers). Let $c_n = \sum_{k=0}^{m(n)} a_k$, with $a_k = (1 - r_k)^{\delta}$.*

(a) *If $(1 - r_k)/(1 - r_{k-1}) \to 1$ then $N_n/c_n \xrightarrow{P} 1$.*
(b) *If $k^{\alpha}(r_k - r_{k-1})/(1 - r_{k-1}) \to 0$, for some $\alpha > 1/2$, then $N_n/c_n \xrightarrow{\text{a.s.}} 1$.*

*Proof* (a) Convergence in probability is directly deduced from Theorem 3.1(b) in Gouet et al. (2007), which asserts that, if $r_k \to 1$ and $(1 - r_k)/(1 - r_{k-1}) \to 1$, then

$$\frac{N_n - \sum_{k=0}^{m(n)} p_{k+\delta}/y_{k-1}}{\sqrt{\sum_{k=0}^{m(n)} a_k^2}} \xrightarrow{D} N(0,1).$$

The hypothesis on $r_k$ implies $\sum_{k=0}^{m(n)} p_{k+\delta}/y_{k-1} \sim c_n$. Also, by Lemma A.1 in Gouet et al. (2007), we have $\sum_{k=0}^{m(n)} a_k^2/c_n^2 \to 0$ and the conclusion follows.

(b) The case $\delta = -1$ is Theorem 2.1(b) in Gouet et al. (2008). It is not difficult to check that the proof of that result is valid for $\delta < -1$ with obvious modifications, provided $k^{\alpha}(1 - \frac{a_{k-1}}{a_k}) \to 0$, for some $\alpha > 1/2$. Thus, it suffices to check that $k^{\alpha}\varphi_k \to 0$ implies $k^{\alpha}(1 - \frac{a_{k-1}}{a_k}) \to 0$, with $\varphi_k = (r_k - r_{k-1})/(1 - r_{k-1})$. In fact,

$$k^{\alpha}\left(1 - \frac{a_{k-1}}{a_k}\right) = k^{\alpha}\big(1 - (1 - \varphi_k)^{-\delta}\big) = (-\delta)k^{\alpha}\varphi_k\big(1 + o(1)\big) \to 0.$$

□

*Example 3* The Poisson distribution, with $p_k = e^{-\lambda}\lambda^k/k!$, $k \in \mathbb{Z}_+$, $\lambda > 0$, is light-tailed and the hazard rates satisfy $r_k = 1 - \lambda/k + o(1/k)$ (see p. 328 in Vervaat 1973). Clearly, $(r_k - r_{k-1})/(1 - r_{k-1}) < C/k$ for some $C > 0$ and every $k \geq 1$. Thus, condition (b) of Theorem 3 holds with $\alpha = 3/4$. Moreover, since $a_k \sim \lambda^{\delta}/k^{\delta}$ and $m(n) \sim \log n/\log\log n$, we have

$$\frac{N_n}{(\log n/\log\log n)^{1-\delta}} \xrightarrow{\text{a.s.}} \lambda^{\delta}/(1-\delta).$$

The above convergence also holds for records ($\delta = 0$); see Gouet et al. (2001).

In the sequel, $F$ is assumed to be absolutely continuous, with differentiable density $f$, such that $\lambda(x) \to \infty$. Since the latter condition implies $f(x) > 0$, for $x$ large enough, and the asymptotic behavior of $N_n$ depends only on the tail of $F$, we assume, without loss of generality, that $f(x) > 0$, for all $x > 0$. Derivatives of a function $g$ are denoted $g'$, $g''$, etc.

**Lemma 6** *Let $F$ be absolutely continuous, with hazard function $\lambda(x) \to \infty$. Then $F$ is light-tailed.*

*Proof* The conclusion follows from $\overline{F}(x+y)/\overline{F}(x) = e^{-\int_x^{x+y} \lambda(t)\, dt}$, $y \geq 0$. □

**Theorem 4** *Let $F$ be absolutely continuous, with differentiable hazard function $\lambda(x) \to \infty$, and let $\delta < 0$. Let also $c_t = \int_0^{m(t)} \lambda(z) a(z)\, dz$, for $t \geq 1$, with $a(z) = e^{\int_{z+\delta}^{z} \lambda(u)\, du}$.*

(a) *If $\lambda'$ is bounded then $N_n/c_n \xrightarrow{P} 1$.*
(b) *If $|\lambda'(x)| < 1/x^r$, for some $r > 1/2$ and all $x$ large enough, then $N_n/c_n \xrightarrow{\text{a.s.}} 1$.*

*Proof* From (7) we know that $H(\log t) = c_t$, for $t \geq 1$, and we have to prove (11) and (12) for the weak law or (9) and (10) for the strong law.

(a) If $z > 0$, $t \in (z+\delta, z)$ and $M$ is an upper bound for $|\lambda'|$, then

$$\left|\lambda(t) - \lambda(z)\right| \leq \int_t^z \left|\lambda'(u)\right| du < M|\delta|.$$

So,

$$\left| \int_{z+\delta}^{z} \lambda(t)\, dt + \delta\lambda(z) \right| < M\delta^2$$

and there exist $A, B > 0$ such that, for all $z > 0$,

$$A e^{-\delta\lambda(z)} < a(z) < B e^{-\delta\lambda(z)}. \tag{18}$$

From (7) and (18) we obtain that (11) is equivalent to the existence of $u_n \uparrow \infty$ such that

$$\frac{\int_{m(n)}^{m(nu_n)} \lambda(z) e^{-\delta\lambda(z)}\, dz}{\int_0^{m(n)} \lambda(z) e^{-\delta\lambda(z)}\, dz} \to 0.$$

In order to prove the convergence stated above, we let $u(x) = \log \lambda(m(x))$ and show that

$$\frac{\int_0^{m(xu(x))} \lambda(z) e^{-\delta\lambda(z)}\, dz}{\int_0^{m(x)} \lambda(z) e^{-\delta\lambda(z)}\, dz} \to 1, \tag{19}$$

by using L'Hôpital's rule and (8). First, notice that

$$\frac{(u(x)+xu'(x))m'(xu(x))\lambda(m(xu(x)))e^{-\delta\lambda(m(xu(x)))}}{m'(x)\lambda(m(x))e^{-\delta\lambda(m(x))}}$$
$$=\left(1+\frac{xu'(x)}{u(x)}\right)e^{-\delta(\lambda(m(xu(x)))-\lambda(m(x)))}.$$

It is clear that

$$\frac{xu'(x)}{u(x)}=\frac{\lambda'(m(x))}{\lambda^2(m(x))\log\lambda(m(x))}\to 0$$

since $\lambda'$ is bounded. Therefore, (19) follows if $\lambda(m(xu(x)))-\lambda(m(x))\to 0$, which is implied by $m(xu(x))-m(x)\to 0$, since $\lambda'$ is bounded. Let us then check that $m(xu(x))-m(x)\to 0$. We have

$$\big|m\big(xu(x)\big)-m(x)\big|\le m'\big(\theta(x)\big)xu(x)=\frac{xu(x)}{\theta(x)\lambda(m(\theta(x)))}, \tag{20}$$

where $\theta(x)\in(x,xu(x))$ and the last equality follows from (8). Now, since $|\lambda'(t)|<M$ for every $t>0$, we have, for $x$ large enough,

$$\big|\lambda\big(m\big(\theta(x)\big)\big)-\lambda\big(m(x)\big)\big|<M\big(m\big(\theta(x)\big)-m(x)\big)=Mm'\big(\psi(x)\big)\big(\theta(x)-x\big), \tag{21}$$

with $\psi(x)\in(x,\theta(x))\subset(x,xu(x))$. So, using again (8), we find that (21) is bounded above by

$$\frac{Mxu(x)}{\psi(x)\lambda(m(\psi(x)))}<\frac{Mu(x)}{\lambda(m(\psi(x)))}.$$

Thus

$$\frac{|\lambda(m(\theta(x)))-\lambda(m(x))|}{\lambda(m(x))}<\frac{Mu(x)}{\lambda(m(\psi(x)))\lambda(m(x))}=\frac{M\log\lambda(m(x))}{\lambda(m(\psi(x)))\lambda(m(x))}\to 0.$$

Therefore, $\lambda(m(\theta(x)))\sim\lambda(m(x))$ and the right-hand side of (20) is equivalent to

$$\frac{xu(x)}{\theta(x)\lambda(m(x))}<\frac{u(x)}{\lambda(m(x))}=\frac{\log\lambda(m(x))}{\lambda(m(x))}\to 0$$

and (19) follows. Now, defining $u_n=\min_{k\ge n}\{u(k)\}$, which is increasing, we have

$$\frac{\int_{m(n)}^{m(nu_n)}\lambda(z)e^{-\delta\lambda(z)}\,dz}{\int_0^{m(n)}\lambda(z)e^{-\delta\lambda(z)}\,dz}\le\frac{\int_{m(n)}^{m(nu(n))}\lambda(z)e^{-\delta\lambda(z)}\,dz}{\int_0^{m(n)}\lambda(z)e^{-\delta\lambda(z)}\,dz}\to 0$$

and (11) is proved.

For (12) recall that $G^{\leftarrow}(1/k)=\overline{F}(m(k)+\delta)$ and $kG^{\leftarrow}(1/k)=a(m(k))$. Thus, the sequence in (12) is equal to

$$\frac{\sum_{k=2}^{n}a(m(k))^2/k}{(\sum_{k=2}^{n}a(m(k))/k)^2}.$$

For $k \geq 2$, let $\mu(k) = \lfloor m(k) \rfloor$. As $|\mu(k) - m(k)| < 1$ and $\lambda'$ is bounded, from (18) there exist $\alpha, \beta > 0$ such that $\alpha < a(\mu(k))/a(m(k)) < \beta$, for all $k \geq 2$. Thus (12) is equivalent to

$$\frac{\sum_{k=2}^{n} a(\mu(k))^2/k}{(\sum_{k=2}^{n} a(\mu(k))/k)^2} \to 0. \tag{22}$$

Note that, as $i \to \infty$,

$$\sum_{k=2}^{\infty} \frac{1}{k} \mathbf{1}_{\{\mu(k)=i\}} \sim \log \frac{\overline{F}(i)}{\overline{F}(i+1)} = \int_i^{i+1} \lambda(t)\, dt \sim \lambda(i), \tag{23}$$

since $\lambda(t) \to \infty$ and $\lambda'$ is bounded. Thus, the numerator in (22) is bounded above by

$$\sum_{i=0}^{\mu(n)} \sum_{k=2}^{\infty} \frac{a(\mu(k))^2}{k} \mathbf{1}_{\{\mu(k)=i\}} \sim \sum_{i=1}^{\mu(n)} a(i)^2 \lambda(i) < B^2 \sum_{i=1}^{\mu(n)} \lambda(i) e^{-2\delta\lambda(i)},$$

as $n \to \infty$, where the last inequality comes from (18). Analogously, the square root of the denominator in (22) is bounded below by

$$\sum_{i=0}^{\mu(n)-1} \sum_{k=2}^{\infty} \frac{a(\mu(k))}{k} \mathbf{1}_{\{\mu(k)=i\}} \sim \sum_{i=1}^{\mu(n)-1} a(i)\lambda(i) > A \sum_{i=1}^{\mu(n)-1} \lambda(i) e^{-\delta\lambda(i)}.$$

Thus, we have to prove that

$$\frac{\sum_{k=1}^{n+1} \lambda(k) e^{-2\delta\lambda(k)}}{(\sum_{k=1}^{n} \lambda(k) e^{-\delta\lambda(k)})^2} \to 0. \tag{24}$$

We decompose the sequence above as

$$\frac{\sum_{k=1}^{n} \lambda(k) e^{-2\delta\lambda(k)}}{(\sum_{k=1}^{n} \lambda(k) e^{-\delta\lambda(k)})^2} + \frac{\lambda(n+1) e^{-2\delta\lambda(n+1)}}{(\sum_{k=1}^{n} \lambda(k) e^{-\delta\lambda(k)})^2}.$$

For the first term, let $\epsilon > 0$ and $k_0$ be such that $\lambda(k) > 2/\epsilon$ for all $k \geq k_0$; we have

$$\begin{aligned}\frac{\sum_{k=1}^{n} \lambda(k) e^{-2\delta\lambda(k)}}{(\sum_{k=1}^{n} \lambda(k) e^{-\delta\lambda(k)})^2} &< \frac{\sum_{k=1}^{k_0} \lambda(k) e^{-2\delta\lambda(k)}}{(\sum_{k=1}^{n} \lambda(k) e^{-\delta\lambda(k)})^2} + \frac{(\epsilon/2) \sum_{k=1}^{n} (\lambda(k) e^{-\delta\lambda(k)})^2}{(\sum_{k=1}^{n} \lambda(k) e^{-\delta\lambda(k)})^2} \\ &< \frac{\sum_{k=1}^{k_0} \lambda(k) e^{-2\delta\lambda(k)}}{(\sum_{k=1}^{n} \lambda(k) e^{-\delta\lambda(k)})^2} + \frac{\epsilon}{2} < \epsilon\end{aligned}$$

for $n$ large enough. Last,

$$\frac{\lambda(n+1) e^{-2\delta\lambda(n+1)}}{(\sum_{k=1}^{n} \lambda(k) e^{-\delta\lambda(k)})^2} < \frac{\lambda(n+1) e^{-2\delta\lambda(n+1)}}{\lambda(n)^2 e^{-2\delta\lambda(n)}} \to 0,$$

since, as $|\lambda'(t)| < M$, we have $|\lambda(n+1) - \lambda(n)| < M$. Thus, (24) follows.

(b) For $z > 0$ and $t \in (z+\delta, z)$, we have $|\lambda(t) - \lambda(z)| \leq \int_{z+\delta}^{z} |\lambda'(u)|\, du$. It follows that $\lambda(t) - \lambda(z) \to 0$, as $z \to \infty$, for $t \in (z+\delta, z)$. Thus, $\int_{z+\delta}^{z} \lambda(t)\, dt + \delta\lambda(z) = \int_{z+\delta}^{z} (\lambda(t) - \lambda(z))\, dt \to 0$, so

$$a(z) = e^{\int_{z+\delta}^{z} \lambda(t)\, dt} \sim e^{-\delta\lambda(z)} \tag{25}$$

and $c_t \sim \int_0^{m(t)} \lambda(z) e^{-\delta\lambda(z)} dz$. Therefore, (9) is equivalent to

$$\frac{\int_0^{m(u \log u)} \lambda(z) e^{-\delta\lambda(z)}\, dz}{\int_0^{m(u)} \lambda(z) e^{-\delta\lambda(z)}\, dz} \to 1.$$

In order to apply L'Hôpital's rule to the above, note that, by (8),

$$\frac{m'(u\log u)(\log u + 1)\lambda(m(u\log u))e^{-\delta\lambda(m(u\log u))}}{m'(u)\lambda(m(u))e^{-\delta\lambda(m(u))}} \sim e^{-\delta(\lambda(m(u\log u)) - \lambda(m(u)))},$$

as $u \to \infty$. To see that $\lambda(m(u\log u)) - \lambda(m(u)) \to 0$, note that

$$\left|\lambda\big(m(u\log u)\big) - \lambda\big(m(u)\big)\right| = \left|\lambda'\big(\theta(u)\big)\right|\big(m(u\log u) - m(u)\big), \tag{26}$$

with $\theta(u) \in (m(u), m(u\log u))$, so the right-hand side of (26) is bounded above, for $u$ large enough, by $(m(u\log u) - m(u))/(m(u))^r$. We now prove that

$$\frac{m(u\log u) - m(u)}{m(u)^r} \to 0, \tag{27}$$

for any $r > 0$. Since $\lambda'(t) \to 0$, there exists $x_0 > 0$ such that $|\lambda'(x)| < 1$ for all $x > x_0$. Then, for all $x > x_0$, we have $\lambda(x) < \lambda(x_0) + (x - x_0) < \lambda(x_0) + x$, so

$$\log \frac{\overline{F}(x_0)}{\overline{F}(x)} = \int_{x_0}^{x} \lambda(z)\, dz < \lambda(x_0)x + x^2/2 < Kx^2,$$

for some $K > 0$ and all $x$ large enough. Therefore,

$$\log n = -\log\big(\overline{F}\big(m(n)\big)\big) < -\log\big(\overline{F}(x_0)\big) + Km(n)^2$$

and there exists $C > 0$ such that $m(n) > C\sqrt{\log n}$, for $n$ large enough. On the other hand, we have $\lambda(x) > 1$, for $x$ large enough, since $\lambda(x) \to \infty$. Thus, $\log\log n = \int_{m(n)}^{m(n\log n)} \lambda(x)\, dx > m(n\log n) - m(n)$. Then, for $n$ large enough,

$$\frac{m(n\log n) - m(n)}{m(n)^r} < \frac{\log\log n}{C^r(\log n)^{r/2}} \to 0$$

and (27) follows.

For (10) note first that, without loss of generality, we may assume $r \in (1/2, 1)$. Since, by (4), $\sum_{k=2}^{n} G^{\leftarrow}(1/k) \sim H(\log n)$, (10) holds if

$$\sum_{n=2}^{\infty} \frac{a(m(n))^2/n}{(\int_0^{m(n)} \lambda(z)a(z)\, dz)^2} < \infty$$

or, letting (as in part (a)) $\mu(k)=\lfloor m(k)\rfloor$, if

$$\sum_{k=0}^{\infty}\frac{a(k)^2}{(\int_0^k a(z)\lambda(z)\,dz)^2}\sum_{n=2}^{\infty}\frac{1}{n}\mathbf{1}_{\{\mu(n)=k\}}<\infty, \tag{28}$$

since $a(m(n))=O(a(\mu(n)))$ and $\int_0^{\mu(n)}a(z)\lambda(z)\,dz\le\int_0^{m(n)}a(z)\lambda(z)\,dz$. From (23) and (25), we have, for $k\to\infty$,

$$\frac{a(k)^2}{(\int_0^k a(z)\lambda(z)\,dz)^2}\sum_{n=2}^{\infty}\frac{1}{n}\mathbf{1}_{\{\mu(n)=k\}}\sim\frac{\lambda(k)e^{-2\delta\lambda(k)}}{(\int_0^k\lambda(z)e^{-\delta\lambda(z)}\,dz)^2}.$$

In order to compare the right-hand side above with $k^{-2r}$, we compute

$$\lim_{x\to\infty}\frac{x^r\sqrt{\lambda(x)}e^{-\delta\lambda(x)}}{\int_0^x\lambda(z)e^{-\delta\lambda(z)}\,dz}$$

by using L'Hôpital's rule. Note that

$$\left(rx^{r-1}\sqrt{\lambda(x)}+\frac{x^r\lambda'(x)}{2\sqrt{\lambda(x)}}-\delta\lambda'(x)x^r\sqrt{\lambda(x)}\right)/\lambda(x)\sim\frac{x^{r-1}(r-\delta x\lambda'(x))}{\sqrt{\lambda(x)}}\to 0,$$

since $|\lambda'(x)|<x^{-r}$ for $x$ large enough. Therefore, for $k$ large enough,

$$\frac{\lambda(k)e^{-2\delta\lambda(k)}}{(\int_0^k\lambda(z)e^{-\delta\lambda(z)}\,dz)^2}<k^{-2r}$$

and, since $r\in(1/2,1)$, the series in (28) converges. □

*Remark 4* Our results complement those of Balakrishnan et al. (2005) and Pakes (2007) for the number $\xi_n(a)$ of near-records, which is defined as the number of observations $X_i$ in $(X(n)-a,X(n)]$, for $i\in(L(n),L(n+1))$, where $L(n)$ is the $n$th record time, $X(n)$ is the $n$th record value and $a>0$ is fixed. For continuous distributions with $\lambda(x)\to\infty$, it is shown in Theorem 4.1 of Balakrishnan et al. (2005) that $\xi_n(a)\xrightarrow{P}\infty$ and a weak law of large numbers is given for $\log\xi_n(a)$ in Sect. 6 of Pakes (2007). If we add the near-record counts for the first $n$ inter-record intervals, we have $\sum_{k=1}^n\xi_k(a)=N_{L(n+1)}^{-a}-N_{L(n+1)}^0=N_{L(n+1)}^{-a}-(n+1)$, where $N_n^{-a}$ denotes the number of $(-a)$-records and $N_n^0$ the number of records among the first $n$ observations. From Theorem 4 we have that, if $\lambda(x)\to\infty$, with $|\lambda'(x)|<x^{-r}$, for some $r>1/2$,

$$\frac{N_n^{-a}}{\int_0^{m(n)}\lambda(x)\overline{F}(x-a)/\overline{F}(x)\,dx}\xrightarrow{\text{a.s.}}1;\qquad\frac{N_n^0}{N_n^{-a}}\xrightarrow{\text{a.s.}}0,$$

so

$$\frac{\sum_{k=1}^n\xi_k(a)}{\int_0^{m(L(n+1))}\lambda(x)\overline{F}(x-a)/\overline{F}(x)\,dx}\xrightarrow{\text{a.s.}}1.$$

The above result gives the precise growth rate of the cumulative number of near-records in terms of record times.

*Example 4* Consider the family of distributions with density $f(x) = ax^{\gamma}e^{-bx^{\nu}}$, for $x > 1$, with $a, b > 0$, $\nu > 1$, $\gamma \in \mathbb{R}$ (note that the values of $f(x)$ for $x < 1$ have no effect on the asymptotic behavior of $N_n$). Clearly, this family contains the normal distribution $N(0, 1)$ ($\nu = 2$, $b = 1/2$ and $\gamma = 0$) and the light-tailed Weibull distribution, with density $f(x) = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^{\alpha}}$, for $x \geq 0$, $\alpha > 1$ and $\beta > 0$ ($\nu = \alpha$, $\gamma = \alpha - 1$ and $b = \beta^{-\alpha}$). It is easy to see that

$$\overline{F}(x) = \frac{ax^{\gamma+1}e^{-bx^{\nu}}}{b\nu}\left(x^{-\nu} + \frac{\gamma - \nu + 1}{b\nu}x^{-2\nu} + o\big(x^{-2\nu}\big)\right).$$

Therefore $\lambda(x) \sim b\nu x^{\nu-1} \to \infty$. Also, simple calculations show that $\lambda'(x) \sim b\nu(\nu - 1)x^{\nu-2}$ as $x \to \infty$.

Fix $\delta < 0$. Then, since $\lambda'(x) \sim b\nu(\nu - 1)x^{\nu-2}$, condition (a) of Theorem 4 holds, for $\nu \in (1, 2]$, while condition (b) holds for $\nu \in (1, 3/2)$.

Let $c_\nu = 1$ for $\nu \in (1, 2)$ and $c_\nu = e^{-b\delta^2}$ for $\nu = 2$. We have $\overline{F}(x + \delta)/\overline{F}(x) \sim c_\nu e^{-b\nu\delta x^{\nu-1}}$. Moreover, it is easy to see, using L'Hôpital's rule, that

$$\int_0^t b\nu x^{\nu-1}e^{-b\nu\delta x^{\nu-1}}\,dx \sim \frac{te^{-b\nu\delta t^{\nu-1}}}{-\delta(\nu - 1)}.$$

Therefore,

$$\frac{N_n}{m(n)e^{-b\nu\delta(m(n))^{\nu-1}}} \to \frac{c_\nu}{-\delta(\nu - 1)},$$

a.s. for $\nu \in (1, 3/2)$ and in probability for $\nu \in (1, 2]$.

In the particular case of the normal distribution we have $\nu = 2$ and $m(n) - \sqrt{2\log n} \to 0$, so

$$\frac{N_n}{\sqrt{2\log n}\; e^{-\delta\sqrt{2\log n}}} \xrightarrow{P} \frac{e^{-\delta^2/2}}{-\delta}.$$

For the Weibull distribution we have $m(n) = \beta(\log n)^{1/\alpha}$, thus obtaining

$$\frac{N_n}{(\log n)^{1/\alpha}e^{-\delta\alpha(\log n)^{(\alpha-1)/\alpha}/\beta}} \to \frac{\beta c_\alpha}{-\delta(\alpha - 1)},$$

a.s. for $\alpha \in (1, 3/2)$ and in probability for $\alpha \in (1, 2]$.

## 4 Maximum likelihood estimation

In this section we show how $\delta$-record statistics can be used to estimate the parameters of the parent distribution $F_\theta$, which is assumed to be absolutely continuous, with density $f_\theta$ and hazard function $\lambda_\theta$, depending on a $k$-dimensional parameter $\theta \in \Theta \subseteq \mathbb{R}^k$. For additional information on inference using record data, see Gulati and Padgett (2003).

It is well known (see Arnold et al. 1998, p. 10) that the likelihood function of the $n$ first record values $R_1, \ldots, R_n$, in a sequence of iid observations, is given by

$$L(r_1, \ldots, r_n; \theta) = \overline{F}_\theta(r_n) \prod_{i=1}^{n} \lambda_\theta(r_i). \tag{29}$$

Fix $\delta < 0$ and suppose that instead of collecting only record values, we collect $\delta$-record values. That is, after a record value $R_i$ is observed, we also keep all observations at a distance less than $|\delta|$ of $R_i$, until the next record value $R_{i+1}$ is observed. Let $K_i$ be the number of such observations and $Y_i^1, \ldots, Y_i^{K_i}$ their values. Recall that $K_i$ is the number of near-records as defined in Balakrishnan et al. (2005), with $a = -\delta$. Thus, our sample consists of the $n$ first record values $(R_1, \ldots, R_n)$ plus, for each record value $R_i$, the number $K_i$ and the values $Y_i^1, \ldots, Y_i^{K_i}$ of near-records associated with $R_i$; in other words, $R_1, Y_1^1, \ldots, Y_1^{K_1}, \ldots, R_n, Y_n^1, \ldots, Y_n^{K_n}$ are the $\delta$-record values before the arrival of the $(n+1)$-th record. Note that, when collecting the data, near-records associated with a record $R_i$ need not be ordered, that is, we do not necessarily have $Y_i^j < Y_i^{j+1}$. Moreover, if two consecutive records are distance less than $|\delta|$ apart, their associated near-records may be not ordered either, that is, we do not necessarily have $Y_i^j < Y_{i+1}^k$.

In the following proposition we give the likelihood function for the sample of $\delta$-records $(\mathbf{R}_n, \mathbf{K}_n, \mathbf{Y}_n)$, where $\mathbf{R}_n = (R_1, \ldots, R_n)$, $\mathbf{K}_n = (K_1, \ldots, K_n)$ and $\mathbf{Y}_n = (Y_1^1, \ldots, Y_1^{K_1}, \ldots, Y_n^1, \ldots, Y_n^{K_n})$.

**Proposition 1** *The likelihood function of* $(\mathbf{R}_n, \mathbf{K}_n, \mathbf{Y}_n)$ *is given by*

$$L(\mathbf{r}_n, \mathbf{k}_n, \mathbf{y}_n; \theta) = \overline{F}_\theta(r_n) \prod_{i=1}^{n} \frac{f_\theta(r_i)}{\overline{F}_\theta(r_i + \delta)^{k_i+1}} \prod_{j=1}^{k_i} f_\theta\left(y_i^j\right), \tag{30}$$

*with* $0 < r_1 < \cdots < r_n < \infty$, $k_i \in \mathbb{Z}_+$ *and* $y_i^j \in (r_i + \delta, r_i)$, *for* $j = 1, \ldots, k_i$, $i = 1, \ldots, n$.

*Proof* Given a record value $t$, the number of near-records has a geometric distribution (starting at 0) with success probability equal to $p_\theta(t) := \overline{F}_\theta(t)/\overline{F}_\theta(t+\delta)$. Also, given that $k_i$ near-records have been observed, their values are independent with density $f_\theta(x)/(\overline{F}_\theta(t+\delta) - \overline{F}_\theta(t))$, for $x \in (t+\delta, t)$, and 0 otherwise. Moreover, conditional on a record value $R_i$, the number and values of its associated near-records are independent of the values of records $R_j$, $j > i$, and the number and values of their associated near-records. Therefore, by successively conditioning on the record values $R_i$ and recalling that, given a record value $R_i = r_i$, the density of $R_{i+1}$ is $f_\theta(x)/\overline{F}_\theta(r_i)$ on $x > r_i$, we obtain

$$L(\mathbf{r}_n, \mathbf{k}_n, \mathbf{y}_n; \theta) = f_\theta(r_1)\big(1 - p_\theta(r_1)\big)^{k_1} p_\theta(r_1) \prod_{j=1}^{k_1} \frac{f_\theta(y_1^j)}{\overline{F}_\theta(r_1+\delta) - \overline{F}_\theta(r_1)}$$
$$\times \cdots$$
$$\times \frac{f_\theta(r_n)}{\overline{F}_\theta(r_{n-1})}\big(1 - p_\theta(r_n)\big)^{k_n} p_\theta(r_n) \prod_{j=1}^{k_n} \frac{f_\theta(y_n^j)}{\overline{F}_\theta(r_n+\delta) - \overline{F}_\theta(r_n)},$$

which yields the result. □

Expression (29) can be used to find the maximum likelihood estimator (MLE) of functions $h(\theta)$ of the parameter, based on the sample of record values. While in most situations maximization of (29) must be done numerically, there are cases where an explicit expression of the MLE can be obtained. For instance, in the exponential distribution with parameter $\lambda > 0$, the MLE of its mean $1/\lambda$ based on the first $n$ record values is given by $R_n/n$. In a similar way, expression (30) can be used for maximum likelihood estimation.

In the following subsections we explore the potential of $\delta$-records in estimation, by carrying out series of simulations for the Weibull and the exponential distributions. We also consider two illustrative examples using real data.

## 4.1 Simulated data

### *4.1.1 Weibull distribution*

We consider here the MLE for the parameters $\alpha, \beta$ of the Weibull distribution, with density $f(x) = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^\alpha}$, for $x \geq 0$, $\alpha > 0$ and $\beta > 0$. The logarithm of the likelihood function of $(\mathbf{R}_n, \mathbf{K}_n, \mathbf{Y}_n)$, obtained from (30), is given by

$$l(\alpha, \beta) = -\beta^{-\alpha} g(\alpha) + N(\log\alpha - \alpha\log\beta) + (\alpha - 1)S, \tag{31}$$

where $g(\alpha) = r_n^\alpha + \sum_{i=1}^n (r_i^\alpha - ((r_i+\delta)^+)^\alpha) + \sum_{i=i}^n \sum_{j=1}^{k_i} ((y_i^j)^\alpha - ((r_i+\delta)^+)^\alpha)$, $S = \sum_{i=1}^n (\log r_i + \sum_{j=1}^{k_i} \log y_i^j)$ and $N = n + \sum_{i=1}^n k_i$.

We now analyze (31) and show how the MLE of the parameters can be computed. Observe that, for fixed $\alpha$, function $l$ is maximized at $\hat\beta(\alpha) = (g(\alpha)/N)^{1/\alpha}$. Hence, for all $\alpha$,

$$\max_\beta l(\alpha, \beta) = l(\alpha, \hat\beta) = N(\log N - 1) + N\big(\log\alpha - \log g(\alpha)\big) + (\alpha - 1)S.$$

Now, in order to obtain the MLE of $\alpha$, we maximize $l(\alpha, \hat\beta)$ or, equivalently, $h(\alpha) = \log\alpha - \log g(\alpha) + (\alpha - 1)\overline{S}$, with $\overline{S} = S/N$. Thus, the problem of finding the MLE $\hat\alpha, \hat\beta$ is reduced to the maximization of a function on the line. Furthermore, it is easy to find a compact interval which contains the optimum, noting that $r_n^\alpha < g(\alpha) < (N+1)r_n^\alpha$ holds, so $h_2(\alpha) \leq h(\alpha) \leq h_1(\alpha)$, where $h_1(\alpha) = \log\alpha - \alpha\log r_n + (\alpha - 1)\overline{S}$ and $h_2(\alpha) = h_1(\alpha) - \log(N+1)$. Next, we observe that $h_1$ and $h_2$ are

**Table 1** Estimation of parameters $\alpha, \beta$ in the Weibull distribution from 10,000 simulation runs, with $\alpha = 2$, $\beta = 1$ and $n = 10$ records

| $\delta$ | | $\alpha, \beta$ unknown | | $\beta$ known | $\alpha$ known |
|---|---|---|---|---|---|
| | | $\alpha$ | $\beta$ | $\alpha$ | $\beta$ |
| 0 | min | 0.661 | 0.091 | 1.206 | 0.502 |
| | max | 12.404 | 3.098 | 4.349 | 1.626 |
| | mean | 2.489 | 1.169 | 2.096 | 0.989 |
| | sdev | 0.924 | 0.422 | 0.304 | 0.158 |
| | rmse | 1.046 | 0.455 | 0.319 | 0.158 |
| −0.5 | min | 1.023 | 0.374 | 1.328 | 0.762 |
| | max | 5.188 | 2.473 | 2.766 | 1.363 |
| | mean | 2.155 | 1.084 | 1.998 | 1.005 |
| | sdev | 0.431 | 0.247 | 0.101 | 0.060 |
| | rmse | 0.458 | 0.261 | 0.101 | 0.060 |

concave and attain their maxima at $\alpha^* = 1/K$, where $K = \log r_n - \overline{S} > 0$. Therefore, $h_2(\alpha^*) \leq \max_\alpha h(\alpha) \leq h_1(\alpha^*)$.

Since $h_1$ is concave and $\lim_{\alpha \to 0^+} h_1(\alpha) = \lim_{\alpha \to +\infty} = -\infty$, we conclude that there are two unique real numbers $L < R$ such that $h_1(L) = h_1(R) = h_2(\alpha^*) = -\log K - 1 - \overline{S} - \log(N+1)$, so $\hat{\alpha} = \operatorname{argmax} h(\alpha) \in [L, R]$. Elementary analysis of functions $h_1, h_2$ yields explicit bounds for $L$ and $R$, such as $L' = \alpha^*/(3N+2) < L$ and $R' = 2\alpha^*(1 + \log(N+1)) > R$. Therefore, the computation of the MLE is reduced to the maximization of $h$ in the interval $[L', R']$, which can be done using standard numerical methods.

In order to compare the estimators based on usual records with those using $\delta$-record data, we consider the Weibull distribution with shape parameter $\alpha = 2$ and scale parameter $\beta = 1$, which is light-tailed (see Example 4). We set $\delta = -0.5$ and the number of records $n = 10$. We make 10,000 simulation runs and compute the MLE of $\alpha$ and $\beta$ when both are unknown and of each parameter, when the other one is known. The values of the MLEs based on records ($\delta = 0$), when both parameters are unknown, or when $\alpha$ is known, are obtained from their explicit expressions shown in Soliman et al. (2006). When $\beta$ is known, the MLE of $\alpha$ based on records is computed numerically. The minimum (min), maximum (max), mean, standard deviation (sdev) and root mean square error (rmse) of the estimations are displayed in Table 1. We observe that the estimations based on $\delta$-records have much smaller rmse than estimations based on records.

Results in Sect. 3 give the asymptotic number of $\delta$-records as a function of the number of observations and can be used as a guide to find the number of data available for estimation when $\delta$-records are used. See Example 4.

### 4.1.2 Exponential distribution

The exponential distribution, with density $f(x) = \lambda e^{-\lambda x}, x \geq 0, \lambda > 0$ can be seen as a special case of the Weibull distribution, considered previously, with $\alpha = 1$ and

**Table 2** Estimation of $1/\lambda$ in the exponential distribution, from 10,000 simulation runs with parameter $\lambda = 1$ and $n = 5, 10, 15, 20$ records

| $\delta$ | | $n=5$ | $n=10$ | $n=15$ | $n=20$ |
|---|---|---|---|---|---|
| 0 | $1/\lambda$ | 1.000 | 1.000 | 1.000 | 1.000 |
| | $1/\sqrt{n}\lambda$ | 0.447 | 0.316 | 0.258 | 0.224 |
| $-1$ | min | 0.253 | 0.533 | 0.589 | 0.630 |
| | max | 3.462 | 2.338 | 1.958 | 1.816 |
| | mean | 1.070 | 1.038 | 1.025 | 1.018 |
| | sdev | 0.321 | 0.212 | 0.165 | 0.142 |
| | rmse | 0.329 | 0.215 | 0.167 | 0.143 |
| $-2$ | min | 0.237 | 0.648 | 0.755 | 0.775 |
| | max | 2.899 | 2.273 | 2.074 | 1.554 |
| | mean | 1.055 | 1.028 | 1.019 | 1.014 |
| | sdev | 0.244 | 0.143 | 0.108 | 0.090 |
| | rmse | 0.250 | 0.146 | 0.109 | 0.091 |

$\beta = 1/\lambda$. The log-likelihood function is obtained from (30) as

$$-\lambda\left(\sum_{i=1}^{n}\bigl(r_i-(r_i+\delta)^+\bigr)+\sum_{i=1}^{n}\sum_{j=1}^{k_i}\bigl(y_i^j-(r_i+\delta)^+\bigr)+r_n\right)+\log\lambda\left(\sum_{i=1}^{n}k_i+n\right).$$

Maximization of the expression above yields the following formula for the MLE of $1/\lambda$:

$$\frac{1}{N}\sum_{i=1}^{n}\left(R_i-(R_i+\delta)^+ +\sum_{j=1}^{K_i}\bigl(Y_i^j-(R_i+\delta)^+\bigr)+\frac{R_n}{n}\right). \tag{32}$$

Table 2 shows the results of the estimation of $1/\lambda$ in the exponential distribution, for different values of $\delta$ and $n$ (number of records). For $\delta = 0$ (usual records) the true values are displayed, that is, the expectation $(1/\lambda)$ and the sdev $1/\sqrt{n}\lambda$. Note that in this case the rmse and the sdev are equal, since the MLE is unbiased (see Arnold et al. 1998, p. 122). We observe that the estimations improve when $|\delta|$ grows, and this is consistent with the fact that more data are being used.

Corollary 3 shows that the asymptotic ratio of the number of $\delta$-records and the number of records is $e^{-\delta\lambda}$. This means that we expect to have $e^{-\delta\lambda}$ times more $\delta$-records than records for the exponential distribution.

### 4.2 Real data

In this subsection, maximum likelihood estimation using $\delta$-records is applied to two different sets of real data. The first contains cumulative precipitation data, from September to November, recorded at Castellote (meteorological station located at the Guadalope river in Teruel, Spain), between 1927 and 2000. The second corresponds to fracture-stress data of brittle materials (Guerra Rosa et al. 2006).

Results for the precipitation data are mainly illustrative because, in most cases, weather stations record all observations and not only records. However, this allows us to compare the estimates obtained from records and $\delta$-records, using the estimate from the whole sample as reference.

In the second example, about fracture-stress data, the usefulness of $\delta$-record observations is more easily seen. An important application of inference based on usual records is in destructive stress testing experiments. See Glick (1978) and Gulati and Padgett (2003). In a classical sampling scheme, stress is applied to all specimens of the sample until they break. Alternatively, in a record sampling scheme, the elements arrive sequentially and are stressed only up to the minimum level that some previous element broke at, thus obtaining a sequence of lower-record values, which can be used for inferential purposes.

We propose an improvement over this scheme, which consists in stressing the elements a bit further than the previous lower-record value, by a fixed factor $\gamma > 1$, say. This procedure yields a sequence of multiplicative $\delta$-lower-records, which provide better inferences, as shown below. We say that an observation $X_n$ is a multiplicative $\delta$-lower-record, with parameter $\gamma \geq 1$, if $X_n < \gamma \min\{X_1, \dots, X_{n-1}\}$ (the case $\gamma = 1$ yields lower-records). This concept is analogous to that of $\delta$-records: the additive parameter $\delta$ is replaced by a multiplicative parameter $\gamma$ and the upper-record is replaced by a lower-record. Asymptotic results for the number of multiplicative $\delta$-lower-records can be easily obtained from results in Sect. 3, because $-\log$ of a multiplicative $\delta$-lower-record is a $\delta$-record.

#### 4.2.1 Precipitation data

A Weibull model has been fit to the Castellote precipitation data of Table 3. The corresponding Q–Q plot is displayed in Fig. 1. The MLEs of $\alpha, \beta$, using the complete sample are given by $\hat{\alpha} = 2.04$ and $\hat{\beta} = 130.38$. If only record values are used, the estimations computed from the formulas in Soliman et al. (2006) are $\hat{\alpha} = 3.93$ and $\hat{\beta} = 185.06$.

In order to compare the behavior of estimators using records and $\delta$-records, we carry out the $\delta$-record-based estimation of $\alpha$ and $\beta$, choosing the values $-25, -50, -75$ for $\delta$. Given that the sample may not contain all near-records corresponding to the last record value, formula (30) has to be slightly modified. That is, the probability, $(1 - p_\theta(r_n))^{k_n} p_\theta(r_n)$, of the event $\{K_n = k_n\}$ must be replaced by the probability, $(1 - p_\theta(r_n))^{k_n}$, of $\{K_n \geq k_n\}$. We obtain

$$L(\mathbf{r}_n, \mathbf{k}_n, \mathbf{y}_n; \theta) = \overline{F}_\theta(r_n + \delta) \prod_{i=1}^{n} \frac{f_\theta(r_i)}{\overline{F}_\theta(r_i + \delta)^{k_i+1}} \prod_{j=1}^{k_i} f_\theta\big(y_i^j\big). \tag{33}$$

The samples of $\delta$-records, for each $\delta$, are shown in Table 4 and results are presented in Table 5. As no closed-form expressions for the sdev and rmse are available, their estimations were obtained through simulations of samples of size 74 containing 5 records, under the assumption that the true values of the parameters are the MLEs based on the complete sample, namely $\alpha = 2.04$ and $\beta = 130.38$. For ease of exposition, the estimations of sdev and rmse are also referred to as sdev and rmse. Since

**Table 3** Precipitation data from Castellote meteorological station

| 1927–1939 | 1940–1952 | 1953–1965 | 1966–1978 | 1979–1991 | 1992–2000 |
|---|---|---|---|---|---|
| 164.6 | 74.6 | 88.5 | 123.5 | 182.0 | 119.0 |
| 32.2 | 78.0 | 10.9 | 150.0 | 63.0 | 112.7 |
| 138.5 | 124.0 | 111.6 | 60.0 | 38.0 | 206.0 |
| 108.0 | 154.1 | 84.9 | 173.0 | 167.0 | 54.0 |
| 62.8 | 97.0 | 247.1 | 76.0 | 67.3 | 96.5 |
| 184.9 | 151.1 | 112.0 | 100.0 | 124.0 | 97.6 |
| 164.7 | 52.2 | 278.8 | 230.0 | 60.8 | 37.1 |
| 130.1 | 77.5 | 176.5 | 82.0 | 214.0 | 119.4 |
| 17.3 | 48.0 | 123.0 | 65.0 | 136.0 | 209.1 |
| 224.9 | 42.2 | 204.3 | 102.0 | 148.0 | |
| 179.5 | 65.7 | 112.3 | 73.2 | 112.0 | |
| 145.2 | 157.3 | 51.6 | 75.0 | 96.2 | |
| 131.5 | 30.6 | 143.3 | 62.0 | 98.5 | |

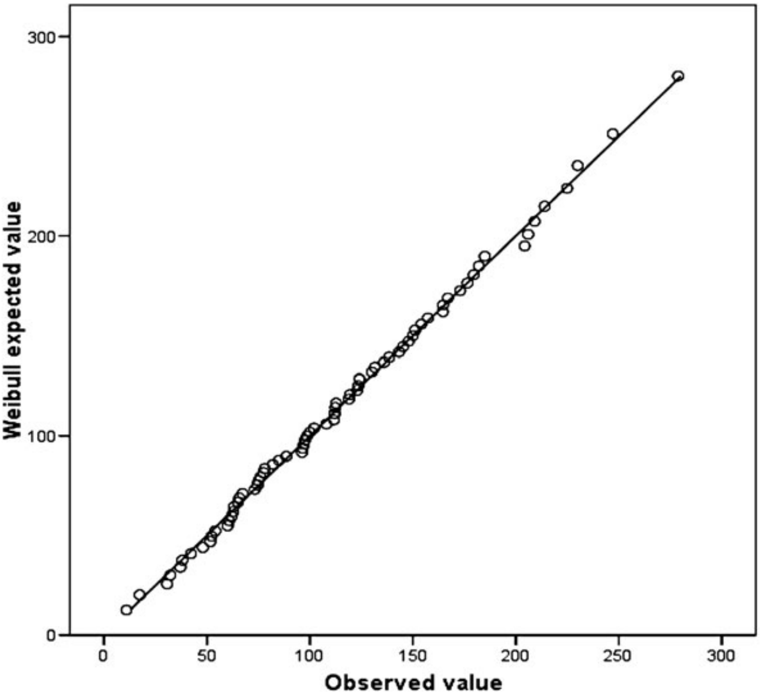


**Fig. 1** Weibull Q–Q plot of Castellote precipitation data

2.04 and 130.38 are not the real values of $\alpha$ and $\beta$, we have also performed simulations using values of $\alpha$ and $\beta$ in a neighborhood of 2.04 and 130.38 and studied how the rmse vary. For illustrative purposes we describe results for records and $\delta$-records

**Table 4** $\delta$-record values for the Castellote precipitation data

| l | $\delta$-record values |
|---|---|
| 0 | 164.6, 184.9, 224.9, 247.1, 278.8 |
| −25 | 164.6, 184.9, 164.7, 224.9, 247.1, 278.8 |
| −50 | 164.6, 138.5, 184.9, 164.7, 224.9, 179.5, 247.1, 278.8, 230.5 |
| −75 | 164.6, 138.5, 108.0, 184.9, 164.7, 130.1, 224.9, 179.5, 154.1, 151.1, 157.3, 247.1, 278.8, 204.3, 230.5, 214.0, 206.0, 209.1 |

**Table 5** Parameter estimation of the Weibull model for Castellote data

| $\delta$ | $\hat{\alpha}$ | rmse $\hat{\alpha}$ | sdev $\hat{\alpha}$ | $\hat{\beta}$ | rmse $\hat{\beta}$ | sdev $\hat{\beta}$ |
|---|---|---|---|---|---|---|
| 0 | 3.93 | 1.86 | 1.69 | 185.06 | 39.70 | 38.26 |
| −25 | 3.46 | 1.80 | 1.65 | 188.32 | 41.67 | 38.66 |
| −50 | 3.44 | 1.50 | 1.35 | 181.69 | 39.74 | 36.35 |
| −75 | 2.90 | 1.10 | 0.98 | 150.17 | 34.52 | 31.75 |
| Complete sample | 2.04 | | | 130.38 | | |

with $\delta = -75$. For $\alpha$ varying in $[1.5, 2.5]$ and $\beta = 130.38$, the rmse of $\hat{\alpha}$ varies from 1.30 to 2.19 using records and from 1.02 to 1.15 using $\delta$-records. The rmse of $\hat{\beta}$ varies from 31.80 to 54.53 using records and from 23.45 to 55.74 using $\delta$-records. When $\beta$ varies in $[100, 160]$, and $\alpha = 2.04$, the rmse of $\hat{\alpha}$ varies from 1.74 to 1.91 using records and from 0.82 to 1.30 using $\delta$-records. The rmse of $\hat{\beta}$ varies from 29.84 to 47.80 using records and from 21.91 to 45.34 using $\delta$-records.

Considering the values of the rmse and how close $\hat{\alpha}$ and $\hat{\beta}$ are to the MLEs based on the whole sample (Table 5), we conclude that, for the scale parameter $\beta$, the estimations based on $\delta$-records yield results similar to those obtained from records, for $\delta = -25, -50$ while, for $\delta = -75$, there is a significant improvement. For the shape parameter $\alpha$, estimations using $\delta$-records, for all values of $\delta$ analyzed, are noticeably better than estimations using only records.

### *4.2.2 Ceramic data*

We analyze the silicon carbide stress data shown in Table 5 in Guerra Rosa et al. (2006). The Weibull distribution fits well the sample of size 69 and the corresponding Q–Q plot is shown in Fig. 2. The MLEs of the parameters, based on the complete sample, are $\hat{\alpha} = 10.62$ and $\hat{\beta} = 362.14$. If the experiment had been conducted using the sequential sampling strategy, as described in Gulati and Padgett (2003), the data set would consist only of lower-record values, shown in the first row of Table 6. Using these record-breaking data, the MLEs of the parameters, computed numerically, are $\hat{\alpha} = 29.77$ and $\hat{\beta} = 305.16$. If the experiment had been conducted using our proposal of stressing each element a bit further than the previous minimum, as explained above, the sample would consist of multiplicative $\delta$-lower-records, as seen in rows 2 to 5 of Table 6.

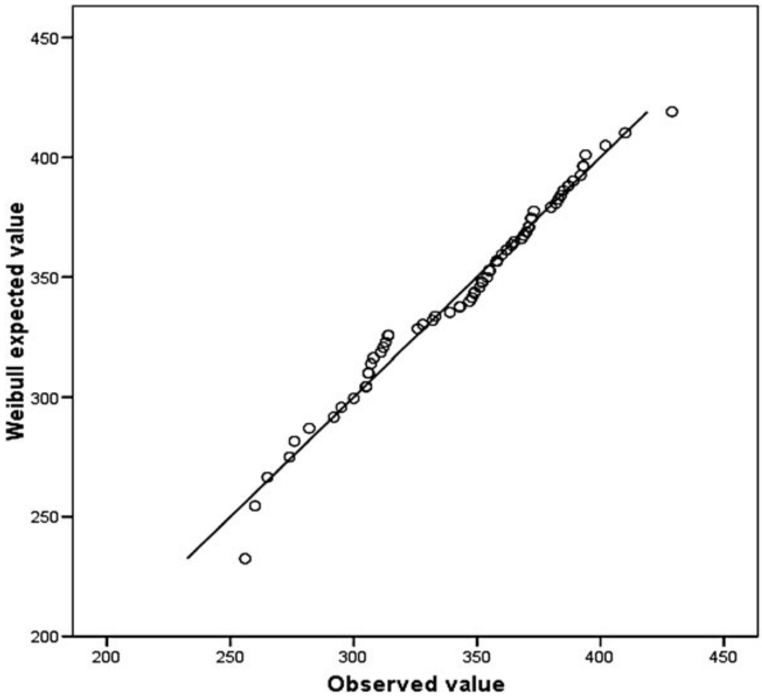


**Fig. 2** Weibull Q–Q plot of silicon carbide stress data

**Table 6** Multiplicative $\delta$-lower-record values for the silicon carbide stress data

| $\gamma$ | Multiplicative $\delta$-lower-record values |
|---|---|
| 1 | 306, 300, 274, 260, 256 |
| 1.05 | 306, 300, 314, 305, 306, 305, 274, 260, 256, 265 |
| 1.10 | 306, 300, 314, 305, 306, 305, 274, 260, 282, 256, 265, 276 |
| 1.15 | 306, 349, 300, 314, 305, 306, 305, 274, 332, 260, 282, 256, 265, 276, 292 |

The likelihood of the multiplicative $\delta$-lower-record sample can be obtained by a similar reasoning to that of Proposition 1. In fact, we only have to replace $\overline{F}_\theta$ by $F_\theta$ and $+\delta$ by $\cdot\gamma$ in (30). Furthermore, as in Sect. 4.2.1, we do not know the exact number of multiplicative near-records associated with the last record, and the likelihood has to be modified as in (33). Finally, we obtain

$$L(\mathbf{s}_n, \mathbf{k}_n, \mathbf{z}_n; \theta) = F_\theta(s_n\gamma) \prod_{i=1}^{n} \frac{f_\theta(s_i)}{F_\theta(s_i\gamma)^{k_i+1}} \prod_{j=1}^{k_i} f_\theta\left(z_i^j\right), \tag{34}$$

with $s_1 > \cdots > s_n$, $k_i \in \mathbb{Z}_+$ and $z_i^j \in (s_i, \gamma s_i)$, for $j = 1, \ldots, k_i$, $i = 1, \ldots, n$.

The maximum likelihood estimation of parameters $\alpha, \beta$ is based on (34), specialized to the Weibull distribution. The maximization is carried out numerically and

**Table 7** Parameter estimation of the Weibull model for the silicon carbide stress data

| $\gamma$ | $\hat{\alpha}$ | rmse $\hat{\alpha}$ | sdev $\hat{\alpha}$ | $\hat{\beta}$ | rmse $\hat{\beta}$ | sdev $\hat{\beta}$ |
|---|---|---|---|---|---|---|
| 1 | 29.77 | 7.92 | 6.88 | 305.16 | 31.08 | 28.69 |
| 1.05 | 26.22 | 5.68 | 5.05 | 309.30 | 31.56 | 28.19 |
| 1.10 | 24.44 | 4.25 | 3.84 | 304.27 | 29.39 | 26.19 |
| 1.15 | 13.29 | 3.29 | 3.06 | 321.32 | 25.59 | 23.13 |
| Complete sample | 10.62 | | | 362.14 | | |

results are shown in Table 7. As for the precipitation data, the rmse and the sdev were estimated through simulation, assuming the parameter values equal to the MLEs based on the complete sample. We have also analyzed the variation in the rmse of $\hat{\alpha}$ and $\hat{\beta}$ when the values of $\alpha$ and $\beta$ vary in a neighborhood of 10.62 and 362.14. For $\alpha$ varying in [10, 11] and $\beta = 362.14$, the rmse of $\hat{\alpha}$ varies from 6.43 to 10.34 using records and from 3.08 to 3.42 for $\gamma = 1.15$. The rmse of $\hat{\beta}$ varies from 29.64 to 33.04 using records and from 23.71 to 28.98 for $\gamma = 1.15$. When $\beta$ varies in [320, 400] and $\alpha = 10.62$, the rmse of $\hat{\alpha}$ varies from 6.83 to 8.85 using records and from 3.12 to 3.42 for $\gamma = 1.15$. The rmse of $\hat{\beta}$ varies from 27.87 to 35.82 using records and from 22.86 to 29.70 for $\gamma = 1.15$.

For the scale parameter $\beta$ we observe similar performances of estimators based on records and estimators based on multiplicative $\delta$-records when $\gamma = 1.05$ or $\gamma = 1.10$, while estimations clearly improve when $\gamma = 1.15$. For the shape parameter $\alpha$, estimations based on multiplicative $\delta$-records are better than estimations based only on records, especially when $\gamma = 1.15$.

### 4.3 Concluding remarks

One important issue of statistical inference based on record values is their scarceness. In fact, the number of trials needed to observe a reasonable number of records may be very large, which makes small values of the rmse unattainable by estimators using only records. In this section we show that, if $\delta$-records are available or the experiment can be conducted in such a way that $\delta$-records are registered, these data can be successfully incorporated in the likelihood, yielding better results than the estimations based on record values only. Table 2 shows that estimations based on $\delta$-records, related to a low number of records, have smaller rmse than estimators based only on a high number of records.

Regarding the practical applications of estimations based on $\delta$-records, we believe that they can play an important role in destructive stress testing, as explained in Sect. 4.2.2. Another field of potential application is actuarial mathematics; for instance, Teugels (1982) describes a procedure based on records to assess the validity of a model for insurance claims, which could be improved by the inclusion of near-records.

Finally, some extensions of $\delta$-record based estimation worth considering are:

- Other sampling schemes: for instance, sampling can be continued until an observation greater than a fixed value is obtained. Another situation to be considered is to register only a random sample of near-records along with each record.

– Take $\delta$ as a function of the record value: throughout this paper we have assumed that $\delta$ is fixed. However, the likelihood function can be rewritten in the case that $\delta$ depends on the corresponding record value.
– Discrete observations: arguments similar to those in Proposition 1 can be used to analyze discrete observations. In this situation, the likelihood is given by

$$L(\mathbf{r}_n, \mathbf{k}_n, \mathbf{y}_n; \theta) = \overline{F}_\theta(r_n) \prod_{i=1}^{n} \frac{\pi_\theta(r_i)}{\overline{F}_\theta(r_i+\delta)^{k_i+1}} \prod_{j=1}^{k_i} \pi_\theta\big(y_i^j\big), \tag{35}$$

with $r_1 < \cdots < r_n$, $k_i \in \mathbb{Z}_+$ and $y_i^j \in (r_i + \delta, r_i]$, for $j = 1, \ldots, k_i$, $i = 1, \ldots, n$ and $\pi_\theta(x) = P_\theta[X_1 = x]$.

**Acknowledgements** We thank the referees and the editor for their useful comments and suggestions. Thanks are also due to J. Abaurrea, J. Asín and A. Cebrián for providing us with the precipitation data used in Sect. 4.2.1. This research was supported by PFB-03-CMM project, Fondecyt Grant 1090216 and projects MTM2007-63769, MTM2010-15972 of MICINN. The authors are members of the research group Modelos Estocásticos (DGA).

## References

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

Balakrishnan N, Balasubramanian K, Panchapakesan S (1996) $\delta$-exceedance records. J Appl Stat Sci 4:123–132

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

Deheuvels P (1974) Valeurs extrémales d'échantillons croissants d'une variable aléatoire réelle. Ann Inst Henri Poincaré X:89–114

Embrechts P, Klüppelberg C, Mikosch T (1997) Modelling extremal events. Springer, Heidelberg

Glick N (1978) Breaking records and breaking boards. Am Math Mon 85:2–26

Gouet R, López FJ, San Miguel M (2001) A martingale approach to strong convergence of the number of records. Adv Appl Probab 33:864–873

Gouet R, López FJ, Sanz G (2007) Asymptotic normality for the counting process of weak records and $\delta$-records in discrete models. Bernoulli 13:754–781

Gouet R, López FJ, Sanz G (2008) Laws of large numbers for the number of weak records. Stat Probab Lett 78:2010–2017

Guerra Rosa L, Lamon J, Figueiredo I, Costa Oliveira FA (2006) A method to distinguish extrinsic and intrinsic fracture-origin populations in monolithic ceramics. J Eur Ceram Soc 26:3887–3895

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

Hashorva E (2003) On the number of near-maximum insurance claim under dependence. Insur Math Econ 32:37–49

Hashorva E, Hüsler J (2005) Estimation of tails and related quantities using the number of near-extremes. Commun Stat, Theory Methods 34:337–349

Key ES (2005) On the number of records in an iid discrete sequence. J Theor Probab 18:99–107

Khmaladze E, Nadareishvili M, Nikabadze A (1997) Asymptotic behaviour of a number of repeated records. Stat Probab Lett 35:49–58

Neveu J (1972) Martingales à temps discret. Masson, Paris

Nevzorov VB (2001) Records: mathematical theory. Translations of mathematical monographs, vol 194. American Mathematical Society, Providence

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

Rényi A (1962) Théorie des éléments saillants d'une suite d'observations. Ann Fac Sci Univ Clermont-Ferrand 8:7–13

Soliman AA, Abd Ellah AH, Sultan KS (2006) Comparison of estimates using record statistics from Weibull model: Bayesian and non-Bayesian approaches. Comput Stat Data Anal 51:2065–2077

Teugels JL (1982) Large claim in insurance mathematics. ASTIN Bull 13:81–88

Vervaat W (1973) Limit theorems for records from discrete distributions. Stoch Process Appl 1:317–334