

# Limit theorems for counting variables based on records and extremes

Allan Gut<sup>1</sup> · Ulrich Stadtmüller<sup>2</sup>

Received: 19 November 2015 / Revised: 31 July 2016 / Accepted: 3 August 2016 /  
Published online: 10 September 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** Hsu and Robbins (Proc. Nat. Acad. Sci. USA **33**, 25–31, 1947) introduced the concept of complete convergence as a complement to the Kolmogorov strong law, in that they proved that  $\sum_{n=1}^{\infty} P(|S_n| > n\varepsilon) < \infty$  provided the mean of the summands is zero and that the variance is finite. Later, Erdős proved the necessity. Heyde (J. Appl. Probab. **12**, 173–175, 1975) proved that, under the same conditions,  $\lim_{\varepsilon \searrow 0} \varepsilon^2 \sum_{n=1}^{\infty} P(|S_n| \geq n\varepsilon) = EX^2$ , thereby opening an area of research which has been called *precise asymptotics*. Both results above have been extended and generalized in various directions. Some time ago, Kao proved a pointwise version of Heyde’s result, viz., for the counting process  $N(\varepsilon) = \sum_{n=1}^{\infty} \mathbb{I}\{|S_n| > n\varepsilon\}$ , he showed that  $\lim_{\varepsilon \searrow 0} \varepsilon^2 N(\varepsilon) \xrightarrow{d} EX^2 \int_0^{\infty} \mathbb{I}\{|W(u)| > u\} du$ , where  $W(\cdot)$  is the standard Wiener process. In this paper we prove analogs for extremes and records for i.i.d. random variables with a continuous distribution function.

**Keywords** Record times · Records · Extremes · Counting process · Weak convergence

---

✉ Allan Gut  
allan.gut@math.uu.se;  
<http://www.math.uu.se/~allan>

Ulrich Stadtmüller  
ulrich.stadtmueller@uni-ulm.de;  
<http://www.mathematik.uni-ulm.de/en/mawi/institute-of-number-theory-and-probability-theory/people/stadtmueller.html>

<sup>1</sup> Department of Mathematics, Uppsala University, Box 480, 751 06 Uppsala, Sweden

<sup>2</sup> Department of Number Theory and Probability Theory, Ulm University, 89069 Ulm, Germany

**AMS 2000 Subject Classifications** Primary—60F15 · 60G50 · 60G60;  
Secondary—60F05

## 1 Introduction

As a complement to the classical Kolmogorov strong law of numbers Hsu and Robbins (1947) introduced, in their seminal paper, the concept of *complete convergence*, and proved that the sequence of arithmetic means of independent, identically distributed (i.i.d.) random variables converges completely (which means that the Borel–Cantelli sum of certain tail probabilities converges) to the expected value of the summands, provided their variance is finite. The necessity was proved by Erdős (1949, 1950).

**Theorem 1.1** *Let  $X, X_1, X_2, \dots$  be i.i.d. random variables with partial sums  $S_n = \sum_{k=1}^n X_k, n \geq 1$ . Then*

$$\sum_{n=1}^{\infty} P(|S_n| > n\varepsilon) < \infty \quad \text{for all } \varepsilon > 0 \iff EX^2 < \infty \quad \text{and} \quad EX = 0.$$

This theorem provides, i.a. information about the rate of convergence in the LLN as  $n \rightarrow \infty$ . Another rate problem is what happens as  $\varepsilon \searrow 0$ . Toward that end Heyde (1975) proved that

$$\lim_{\varepsilon \searrow 0} \varepsilon^2 \sum_{n=1}^{\infty} P(|S_n| \geq n\varepsilon) = EX^2,$$

whenever  $EX = 0$  and  $EX^2 < \infty$ , thereby initiating an area which later has been coined “precise asymptotics”. Introducing the counting variable

$$N(\varepsilon) = \text{Card} \{n : |S_n| > n\varepsilon\} = \sum_{n=1}^{\infty} \mathbf{1}\{|S_n| > n\varepsilon\},$$

we note that Heyde’s result is equivalent to

$$\varepsilon^2 EN(\varepsilon) \rightarrow EX^2 \quad \text{as } n \rightarrow \infty.$$

In his paper, Kao (1978) discusses i.a. possible limits of the counting variable itself, that is, possible limit distributions of  $N(\varepsilon)$ , that is, of  $\sum_{n=1}^{\infty} \mathbf{1}\{|S_n| > n\varepsilon\}$ , as  $\varepsilon \searrow 0$ . In summary, his result runs as follows.

**Theorem 1.2** *Let  $X, X_1, X_2, \dots$  be i.i.d. random variables with mean 0, variance  $\sigma^2$ , and partial sums  $S_n, n \geq 1$ , and let  $\{W(t), t \geq 0\}$  denote a standard Wiener process. Then, for the counting process as defined above,*

$$\varepsilon^2 N(\varepsilon) \xrightarrow{d} \sigma^2 \int_0^{\infty} \mathbf{1}\{|W(u)| > u\} du \quad \text{as } \varepsilon \searrow 0.$$

*Remark 1.1* A one-sided analog also holds.

The purpose of the present paper is to prove analogs for the partial maxima of i.i.d. random variables, their record times, the associated counting process, and the corresponding record values. More precisely, in Section 2 we prove an analog where the counting variable in Kao’s theorem is replaced by  $\sum_{k=1}^{\infty} \frac{1}{k} \mathbb{I}\{M_k > (1 + \varepsilon)Q(k)\}$ , where  $M_k = \max_{1 \leq l \leq k} X_l$ ,  $n \geq 1$ ,  $Q(\cdot)$  is a function that will be properly defined in Theorem 2.1, and where the limiting Wiener process is replaced by the Gumbel process. In Section 3 we first investigate the analog for the record times  $\{L(k) \mid k \geq 1\}$ , in which case the counting process will be  $\sum_{k=1}^{\infty} \mathbb{I}\{|\log L(k) - k| > k\varepsilon\}$ . For the counting process  $\{\mu(k), k \geq 1\}$  (the number of records so far) the indicator sum will be  $\sum_{k=1}^{\infty} \frac{1}{k} \mathbb{I}\{|\mu(k) - \log k| > \varepsilon \log k\}$ , and, finally, for the records,  $\{X_{L(k)}, k \geq 1\}$ , the analog is  $\sum_{k=1}^{\infty} \mathbb{I}\{X_{L(k)} > Q(e^{(1+\varepsilon)k})\}$ .

In the three record problems the limiting integral is throughout the same as in Kao’s result above. This is basically due to the fact that sums of independent random variables enter the discussion in those cases.

The proof of Kao’s theorem, and hence to a large extent also of our results, is based on three basic steps: Let  $M$  be large, let  $\varepsilon_n \searrow 0$ , and define

$$N_M(\varepsilon_n) = \sum_{k=1}^{\lfloor M/\varepsilon_n^2 \rfloor} \mathbb{I}\{|S_k| > \varepsilon_n k\}. \tag{1.1}$$

Then

$$\varepsilon_n^2 N_M(\varepsilon_n) \xrightarrow{d} \int_0^M \mathbb{I}\{|W(y)| > y\} dy \quad \text{as } n \rightarrow \infty, \tag{1.2}$$

$$\varepsilon_n^2 (N(\varepsilon_n) - N_M(\varepsilon_n)) \xrightarrow{p} 0 \quad \text{as } \varepsilon_n \searrow 0, M \rightarrow \infty, \tag{1.3}$$

$$\int_M^\infty \mathbb{I}\{|W(y)| > y\} dy \xrightarrow{p} 0 \quad \text{as } M \rightarrow \infty. \tag{1.4}$$

Moreover, since the first two conclusions build on a weak invariance principle and a continuous mapping theorem, the real work amounts, “more or less”, to verifying that the prerequisites for an application of those results are fulfilled. Some tools in this respect can be found in the [Appendix](#).

Finally, since the record sequence is the subsequence of the partial maxima that selects a given maximum when it appears the first time, that is, the records can be seen as a compression of the partial maxima, we devote Section 4 to a discussion of this observation.

## 2 Extremes

Weak and strong limit theorems for partial extrema  $M_n = \max_{1 \leq k \leq n} X_k$ , typically have a very slow rate of convergence and a strong law for  $M_n$  holds only in special cases, namely, when we are in the domain of attraction of the Gumbel law, denoted  $F \in \mathcal{A}(\Lambda)$ . We therefore confine ourselves to such distribution functions  $F$  with  $t_F := \sup\{x : 1 - F(x) > 0\}$ . Our main result can be considered as a refinement of a strong law under these assumptions.

Before going into further details, we need some notation and relevant quantities. For simplicity, we assume that  $t_F = \infty$ . If  $F \in \mathcal{A}(\Lambda)$  we have (see, e.g., Embrechts et al. 1997; Resnick 1987)

$$R(x) = \left( \frac{1}{1 - F} \right) (x) \in \Gamma(g) \quad \text{with an auxiliary function } g,$$

which satisfies (see Geluk and de Haan (1987), p. 41),  $g(x)/x \rightarrow 0$  as  $x \rightarrow \infty$ ,

i.e.,  $\frac{R(x + tg(x))}{R(x)} \rightarrow e^t$  as  $x \rightarrow \infty$  for  $t \in \mathbb{R}$ .

For the inverse function that means

$$Q(x) = R^{-1}(x) \in \Pi(a) \text{ with an auxiliary function, (see Geluk and de Haan (1987), p. 36),}$$

$$a(x) = g(Q(x)),$$

i.e.,  $\frac{Q(xt) - Q(x)}{a(x)} \rightarrow \log t$  as  $x \rightarrow \infty$  for  $t \in \mathbb{R}$ .

With these notations we have

$$\frac{M_n - b(n)}{a(n)} \xrightarrow{d} \Lambda \quad \text{as } n \rightarrow \infty, \tag{2.1}$$

where

$$b(n) = Q(n) \quad \text{and} \quad a(n) = g(b(n)) = g(Q(n)).$$

Finally, denote by  $Y_{0,\Lambda}(t) = Y_\Lambda(t) - \log t$  the centered extremal process associated with the Gumbel process, in the sense that, for all  $t > 0$  and  $x \in \mathbb{R}$ , we have  $P(Y_\Lambda(t) - \log t \leq x) = e^{-e^{-x}}$ . For more on extremal processes, see, e.g., Resnick (1987). Note that the reparametrized process  $Y_{0,\Lambda}(e^t) = Y_\Lambda(e^t) - t$  is a stationary Markov process (see e.g Fahrner and Stadtmüller 2003).

From Eq. 2.1 it follows that  $M(n)/Q(n) \xrightarrow{p} 1$  as  $n \rightarrow \infty$ . As for strong laws one may consult e.g. Embrechts et al. (1997), Section 3.5, which in turn is based on earlier work from e.g. Barndorff-Nielsen (1963) and de Haan and Hordijk (1972). There one finds conditions entailing  $M(n)/Q(n) \xrightarrow{a.s.} 1$  as  $n \rightarrow \infty$ . In the standard normal case this amounts to  $M_n/\sqrt{2 \log n} \xrightarrow{a.s.} 1$  as  $n \rightarrow \infty$ . See also Remark 2.1(a) and (b) below.

Given this setting we are now ready to investigate how often the threshold  $(1 + \varepsilon)Q(n)$  is surpassed by  $M_n$  in terms of an  $\varepsilon$ -rate as  $\varepsilon$  decreases.

**Theorem 2.1** *Let*

$$N(\varepsilon) = \sum_{k=1}^{\infty} \frac{1}{k} \mathbf{1}\{M_k > (1 + \varepsilon)Q(k)\},$$

*and assume that the following conditions are met:*

- (i):  $F \in \mathcal{A}(\Lambda)$  and  $t_F = \infty$ ;
- (ii): *There exist positive constants  $\beta, k_0$ , such that  $R((1 + \varepsilon)Q(k)) \geq k^{1+\beta\varepsilon}$  for all  $\varepsilon \in (0, 1/10]$  and  $k \geq k_0$ ;*

(iii):  $\lim_{x \rightarrow \infty} \frac{Q(e^x)}{x a(e^x)} = c$  with some  $c > 0$ .

Then,

$$\varepsilon N(\varepsilon) \xrightarrow{d} \int_1^\infty \mathbb{I}\{Y_{0,\Lambda}(t) > c \log t\} \frac{dt}{t} = \int_0^\infty \mathbb{I}\{Y_{0,\Lambda}(e^v) > c v\} dv \quad \text{as } \varepsilon \searrow 0.$$

*Remark 2.1* (a): Under Assumption (ii) we have  $\limsup_{n \rightarrow \infty} M_n/Q(n) \leq 1$  a.s., since

$$\sum_k P(X_k > (1 + \varepsilon)Q(k)) = \sum_k \frac{1}{R((1 + \varepsilon)Q(k))} < \infty \quad \text{for all } \varepsilon > 0.$$

Hence  $N(\varepsilon)$  exists a.s.

(b): If, in particular,  $F$  has a positive and monotone density,  $f$ , then  $g(x) \sim (1 - F(x))/f(x)$  (see e.g. de Haan and Geluk 1980, Proposition 1.31,4 and Corollary 1.29), in which case Condition (iii) can be rewritten as

$$\frac{f(x)}{(1 - F(x))(-\log(1 - F(x)))} = \frac{c + o(1)}{x} \quad \text{and thus } \log \log(1 - F(x)) = \int^x \frac{c + o(1)}{u} du,$$

which, in turn, implies that  $\log(1 - F(x))$  belongs to a subclass of the regularly varying functions.

In this specific case  $M_n/Q(n) \xrightarrow{a.s.} 1$  as  $n \rightarrow \infty$  due to Resnick and Tomkins (1973), Theorem 1.

(c): Under Conditions (i)–(iii) above, together with

(iv): There exist  $c, k_0 > 0$ , such that  $R((1 - \varepsilon)Q(k)) \leq c \frac{k}{\varepsilon \log k}$  for all  $\varepsilon \in (0, 1/10]$  and  $k \geq k_0$ , there is also a two-sided version of the result, viz., setting  $\tilde{N}(\varepsilon) = \sum_{k=1}^\infty \frac{1}{k} \mathbb{I}\{|M_k - Q(k)| > Q(k)\varepsilon\}$ , we have

$$\varepsilon \tilde{N}(\varepsilon) \xrightarrow{d} \int_1^\infty \mathbb{I}\{|Y_{0,\Lambda}(t)| > c \log t\} \frac{dt}{t} = \int_0^\infty \mathbb{I}\{|Y_{0,\Lambda}(e^v)| > c v\} dv \quad \text{as } \varepsilon \searrow 0.$$

(d): Since the auxiliary function  $g$  can be chosen as  $g(t) = 1/(\log \int^t R(v) dv)'$  (see e.g. Bingham et al. 1987, bottom p. 177 or Geluk and de Haan 1987, Corollary 1.29), setting  $t = Q(e^{x/\varepsilon})$ , Assumption (iii) in our Theorem is equivalent to

(iii)':  $\frac{t}{g(t) \log(R(t))} \rightarrow c$  as  $t \rightarrow \infty \iff \frac{t \left( \log \left( \int^t R(v) dv \right) \right)'}{\log R(t)} \rightarrow c$  as  $t \rightarrow \infty$ .

*Proof* The key point is that, for a nullsequence  $\{\varepsilon_n\}$ , and some large  $M > 0$  with  $\tau_{k,n} = k^{\varepsilon_n}$  and  $\Delta_n = \tau_{k,n} - \tau_{k-1,n} = \tau_{k,n} \frac{\varepsilon_n}{k} (1 + o(1))$  as  $n \rightarrow \infty$ , we have,

$$\begin{aligned} N_M(\varepsilon_n) &:= \varepsilon_n \sum_{e^{\delta/\varepsilon_n} \leq k \leq e^{M/\varepsilon_n}} \frac{1}{k} \mathbf{I}\{M_k > (1 + \varepsilon_n)Q(k)\} \\ &= \varepsilon_n \sum_{e^{\delta/\varepsilon_n} \leq k \leq e^{M/\varepsilon_n}} \frac{1}{k} \mathbf{I}\left\{ \frac{M_{k^{\varepsilon_n/\varepsilon_n}} - Q(k^{\varepsilon_n/\varepsilon_n})}{a(k^{\varepsilon_n/\varepsilon_n})} > \varepsilon_n \frac{Q(k^{\varepsilon_n/\varepsilon_n})}{a(k^{\varepsilon_n/\varepsilon_n})} \right\} \\ &= \sum_{e^{\delta} \leq \tau_{k,n} \leq e^M} \frac{1 + o(1)}{\tau_{k,n}} \mathbf{I}\left\{ \frac{M_{e^{(\log \tau_{k,n})/\varepsilon_n}} - Q(e^{(\log \tau_{k,n})/\varepsilon_n})}{a(e^{(\log \tau_{k,n})/\varepsilon_n})} > \frac{Q(e^{(\log \tau_{k,n})/\varepsilon_n})}{a(e^{(\log \tau_{k,n})/\varepsilon_n})} \right\} \Delta_n \\ &= \sum_{e^{\delta} \leq \tau_{k,n} \leq e^M} \frac{1 + o(1)}{\tau_{k,n}} \mathbf{I}\left\{ \frac{M_{e^{(\log \tau_{k,n})/\varepsilon_n}} - Q(e^{(\log \tau_{k,n})/\varepsilon_n})}{a(e^{(\log \tau_{k,n})/\varepsilon_n})} > c \log(\tau_{k,n})(1 + o(1)) \right\} \Delta_n, \end{aligned}$$

where we used Assumption (iii) in the last step; note that the  $o(\cdot)$ -terms are uniform in  $k$ .

Next we note that, by a modified version of Resnick (1987), Proposition 4.20, and the continuous mapping theorem,

$$N_M(\varepsilon_n) \xrightarrow{d} \int_{e^{\delta}}^{e^M} \mathbf{I}\{Y_{0,\Lambda}(e^{\log t}) > \log t\} \frac{dt}{t}.$$

To deal with the now two-sided remainders note that, on the one hand,

$$\varepsilon_n \sum_{k \leq e^{\delta/\varepsilon_n}} \frac{1}{k} \mathbf{I}\{M_k > (1 + \varepsilon_n)Q(k)\} \leq \varepsilon_n \sum_{k \leq e^{\delta/\varepsilon_n}} \frac{1}{k} \leq 2\delta,$$

and, similarly, that

$$\int_1^{e^{\delta}} \mathbf{I}\{Y_{0,\Lambda}(e^{\log t}) > \log t\} \frac{dt}{t} \leq \delta,$$

being small for small  $\delta$ . On the other hand, under Assumption (ii),

$$\begin{aligned} E \left( \varepsilon_n \sum_{k \geq e^{M/\varepsilon_n}} \frac{1}{k} \mathbf{I}\{M_k > (1 + \varepsilon_n)Q(k)\} \right) &= \varepsilon_n \sum_{k \geq e^{M/\varepsilon_n}} \frac{1}{k} P(M_k > (1 + \varepsilon_n)Q(k)) \\ &\leq \varepsilon_n \sum_{k \geq e^{M/\varepsilon_n}} \frac{1}{k} \left( 1 - \left( 1 - \frac{1}{R(Q(k)(1 + \varepsilon_n))} \right)^k \right) \\ &\leq \varepsilon_n \sum_{k \geq e^{M/\varepsilon_n}} \frac{m}{k^{1+\beta\varepsilon_n}} \leq m' e^{-M\beta}, \end{aligned}$$

which decreases as  $M$  increases.

Since  $P(Y_{0,\Lambda}(t) > c \log t) \leq t^{-c}$ , the analogous conclusions hold for the limiting random variable. Hence, the remainders converge in probability to zero as  $\delta \rightarrow 0$  and  $M \rightarrow \infty$ . □

In order to prove the claims made in the remark we note that

$$\begin{aligned} E \left( \varepsilon_n \sum_{k \geq e^{M/\varepsilon_n}} \frac{1}{k} \mathbb{I}\{M_k < (1 - \varepsilon_n)Q(k)\} \right) &= \varepsilon_n \sum_{k \geq e^{M/\varepsilon_n}} \frac{1}{k} P(M_k < (1 - \varepsilon_n)Q(k)) \\ &\leq \varepsilon_n \sum_{k \geq e^{M/\varepsilon_n}} \frac{1}{k} \left( 1 - \frac{1}{R((1 - \varepsilon_n)Q(k))} \right)^k \\ &\leq \varepsilon_n \sum_{k \geq e^{M/\varepsilon_n}} \frac{m}{k^{1+\varepsilon_n/c}} \leq m' e^{-M/c}. \end{aligned}$$

The rest follows as before.

*Example 2.1* Let  $\bar{F}(x) = \kappa x^\gamma e^{-(\delta x)^\alpha}$  with constants  $\alpha, \kappa, \delta > 0$  and  $\gamma \in \mathbb{R}$ . Hence,  $F \in \Gamma(g)$  with  $g(t) = \frac{1}{\alpha\delta^\alpha} x^{1-\alpha}$  and thus  $F \in \mathcal{A}(\Lambda)$ . Furthermore,  $R(x) = \kappa^{-1} x^{-\gamma} e^{(\delta x)^\alpha}$  and  $Q(y) = \delta^{-1} (\log y)^{1/\alpha} (1 + \mathcal{O}(\log \log y / \log y))$  as  $y \rightarrow \infty$ .

It follows that

$$\begin{aligned} R((1 + \varepsilon)Q(x)) &= R(Q(x))^{(1+\varepsilon)^\alpha} \kappa^{-1+(1+\varepsilon)^\alpha} Q(x)^{-\gamma(1-(1+\varepsilon)^\alpha)} (1 + \varepsilon)^{-\gamma} \\ &\geq x^{(1+\varepsilon)^\alpha} \kappa^{-1+(1+\varepsilon)^\alpha} Q(x)^{-\gamma(1-(1+\varepsilon)^\alpha)} (1 + \varepsilon)^{-\gamma} \geq x^{1+\alpha\varepsilon/2} \end{aligned}$$

for  $x \geq x_0$  and all  $\varepsilon \in (0, 1]$ , as, in essence,  $1 - (1 + \varepsilon)^\alpha = -\alpha\varepsilon$  for small  $\varepsilon > 0$ , and, moreover,  $\sqrt{x}Q(x)^\gamma \geq \kappa^{-1}e^{\gamma/\alpha}$  for large enough  $x$ .

Next we choose the auxiliary function  $a(y) = \frac{1}{\alpha\delta} (\log y)^{1/\alpha-1}$ , and conclude that

$$\frac{\varepsilon Q(e^{x/\varepsilon})}{x a(e^{x/\varepsilon})} \rightarrow \alpha \quad \text{as } \varepsilon \searrow 0.$$

So this example satisfies the assumptions of the theorem.

Condition (iv) is satisfied as well, since there exist constant  $c, c' > 0$ , such that

$$R((1 - \varepsilon)Q(x)) \leq cx^{1-\alpha\varepsilon/2} \leq c' \frac{x}{\varepsilon \log x}$$

uniformly in  $\varepsilon \in (0, 1/10]$  for large enough  $x$ .

The normal distribution ( $\alpha = 2$ ) can be dealt with in the same manner, since  $\bar{F}(x) \sim \frac{1}{\sqrt{2\pi x}} e^{-x^2/2}$  as  $x \rightarrow \infty$ .

Things become simpler if we readjust the right hand side of the counting variable.

**Theorem 2.2** *Let  $F \in \mathcal{A}(\Lambda)$ , and set*

$$\tilde{N}(\varepsilon) = \sum_{k=1}^{\infty} \frac{1}{k} \mathbb{I}\{M_k > Q(k^{1+\varepsilon})\}.$$

Then

$$\varepsilon \tilde{N}(\varepsilon) \xrightarrow{d} \int_1^{\infty} \mathbb{I}\{Y_{0,\Lambda}(t) > c \log t\} \frac{dt}{t} = \int_0^{\infty} \mathbb{I}\{Y_{0,\Lambda}(e^v) > cv\} dv \quad \text{as } \varepsilon \searrow 0.$$

*Proof* The proof follows the above pattern. Note that Assumption (ii) in Theorem 2.1 is now automatically granted, so the remainders can be dealt with as before. As for the limit behavior of the main part we have

$$\begin{aligned} \frac{Q(k^{\varepsilon_n/\varepsilon_n}(1+\varepsilon_n)) - Q(k^{\varepsilon_n/\varepsilon_n})}{a(k^{\varepsilon_n/\varepsilon_n})} &= \frac{Q(e^{(\log \tau_{k,n})(1+\varepsilon_n)/\varepsilon_n}) - Q(e^{(\log \tau_{k,n})/\varepsilon_n})}{a(e^{(\log \tau_{k,n})/\varepsilon_n})} \\ &= e^{\log \tau_{k,n}}(1 + o(1)) \quad \text{as } n \rightarrow \infty, \end{aligned}$$

locally uniformly, since  $Q \in \Pi(a)$ . This completes the proof of that part and we are done. □

*Remark 2.2* The latter result also includes the case  $t_F < \infty$  as in Gnedenko’s example  $F(x) = 1 - \exp\{-x/(1-x)\}$  for  $0 \leq x < 1$ .

We have not been able to prove a corresponding distributional limit theorem for the counting variable  $N_0(\varepsilon) = \sum_{k=1}^\infty \mathbb{1}\{M_k > (1 + \varepsilon)Q(k)\}$ . This is related to the fact that the expectation of  $N_0(\varepsilon)$  will in general not exist (check e.g. the exponential distribution). However, the following moments do exist.

**Proposition 2.1** *Let  $T_0(\varepsilon) = \operatorname{argmax}\{k : M_k > (1 + \varepsilon)Q(k)\}$  be the last exit time. Under Assumptions (i) and (ii) of Theorem 2.1 we have, for any  $\varepsilon > 0$ ,*

- (i):  $N_0(\varepsilon) \leq T_0(\varepsilon) < \infty$  a.s.;
- (ii):  $E(\log N_0(\varepsilon)) \leq E(\log T_0(\varepsilon)) < \infty$ ;
- (iii):  $E(N(\varepsilon)) < \infty$ .

*Proof* The proofs are all based on calculations of series as in the last part of the proof of Theorem 2.1 using Assumption (ii) there. As for (i), the first inequality is obvious, whereas the a.s. finiteness of  $T_0(\varepsilon)$  follows from the consideration in Remark 2.1(a) using Assumption Theorem 2.1(ii).

Finiteness of the logarithmic moment of  $T_0$  is a consequence of the fact that

$$\sum_{k=1}^\infty \frac{\log k}{k} P(M_k > (1 + \varepsilon)Q(k)) \leq c \sum_{k=1}^\infty \frac{\log k}{k^{1+\beta\varepsilon}} < \infty,$$

together with Proposition A.3 (put  $U_k = M_k$  and  $a_k = Q(k)$  being slowly varying), after which the same holds for  $N_0$  via (i). For (iii) one uses the same calculations as above, which yields

$$E(N(\varepsilon)) = \sum_k \frac{1}{k} P(M_k > (1 + \varepsilon)Q(k)) < \infty.$$

□



### 3 Records

Let  $X, X_1, X_2, \dots$  be i.i.d. continuous random variables. The *record times* are  $L(1) = 1$  and, recursively,

$$L(n) = \min\{k : X_k > X_{L(n-1)}\}, \quad n \geq 2.$$

The associated *counting process*  $\{\mu(n), n \geq 1\}$  is defined by

$$\mu(n) = \# \text{ records among } X_1, X_2, \dots, X_n = \max\{k : L(k) \leq n\}.$$

For a fairly recent and comprehensive introduction to the area we refer to Nevzorov (2001) and further references given there (some basics and references are provided in Gut 2013).

A first observation is that  $\mu(n) = \sum_{k=1}^n I_k$ , where  $I_k = 1$  if  $X_k$  is a record and 0 otherwise, where, in turn,  $P(I_k = 1) = 1 - P(I_k = 0) = 1/k$ , and  $\{I_k, k \geq 1\}$  are independent. One then easily checks that

$$\begin{aligned} m_n = E \mu(n) &= \sum_{k=1}^n \frac{1}{k} = \log n + \gamma + o(1) \text{ and} \\ \text{Var } \mu(n) &= \sum_{k=1}^n \frac{1}{k} \left(1 - \frac{1}{k}\right) = \log n + \gamma - \frac{\pi^2}{6} + o(1), \end{aligned} \tag{3.1}$$

as  $n \rightarrow \infty$ , where  $\gamma = 0.5772156649015328606\dots$  is Euler’s constant.

Following are well-known strong laws and central limit theorems:

$$\begin{aligned} \frac{\mu(n)}{\log n} &\xrightarrow{\text{a.s.}} 1 \quad \text{and} \quad \frac{\mu(n) - \log n}{\sqrt{\log n}} \xrightarrow{d} N(0, 1) \quad \text{as } n \rightarrow \infty, \\ \frac{\log L(n)}{n} &\xrightarrow{\text{a.s.}} 1 \quad \text{and} \quad \frac{\log L(n) - n}{\sqrt{n}} \xrightarrow{d} N(0, 1) \quad \text{as } n \rightarrow \infty. \end{aligned}$$

#### 3.1 Record times

**Theorem 3.1** *Let*

$$N(\varepsilon) = \sum_{k=1}^{\infty} \mathbb{I}\{|\log L(k) - k| > k\varepsilon\}.$$

*Then*

$$\varepsilon^2 N(\varepsilon) \xrightarrow{d} \int_0^{\infty} \mathbb{I}\{|W(y)| > y\} dy \quad \text{as } \varepsilon \searrow 0.$$

*Proof* The key tool here is Williams’ representation (Williams 1973): Let  $\{E_k, k \geq 1\}$  be i.i.d. standard exponential random variables, and set  $\Gamma_n = \sum_{k=2}^n E_k, n \geq 1$ . Then

$$E_n < \log L(n) - \log L(n - 1) \leq E_n + \frac{1}{L(n - 1)} \leq E_n + \frac{1}{n - 1}, \tag{3.2}$$

from which it i.a. follows that

$$\Gamma_n < \log L(n) \leq \Gamma_n + \log(n - 1) + \gamma_n, \tag{3.3}$$

where  $\gamma_n = \sum_{k=1}^{n-1} \frac{1}{k} - \log(n - 1) = \gamma + \mathcal{O}(1/n)$ . Thus, (almost) any property of the gamma distribution carries over to the logarithm of the record times, since  $\Gamma_n \in \Gamma(n - 1, 1)$ .

The conclusion is now immediate in view of Theorem 1.2, together with Proposition A.2, since  $\gamma_n/n \rightarrow 0$  as  $n \rightarrow \infty$ . □

*Remark 3.1* In Gut (2002), Theorem 4.1, it was (i.a.) shown that

$$\lim_{\varepsilon \searrow 0} \varepsilon^2 E N(\varepsilon) = \sum_{k=1}^{\infty} P(|\log L(k) - k| > k\varepsilon) = 1 = E \left( \int_0^{\infty} \mathbb{I}\{|W(y)| > y\} dy \right).$$

### 3.2 The counting process

**Theorem 3.2** *Let*

$$N(\varepsilon) = \sum_{k=1}^{\infty} \frac{1}{k} \mathbb{I}\{|\mu(k) - \log k| > \varepsilon \log k\}.$$

*Then*

$$\varepsilon^2 N(\varepsilon) \xrightarrow{d} \int_0^{\infty} \mathbb{I}\{|W(y)| > y\} dy \quad \text{as } \varepsilon \searrow 0.$$

*Proof* The basic pattern of the proof is the same as that of Theorem 2.1.

The random variables  $I_k - 1/k, k \geq 1$ , are uniformly bounded by 1 in absolute value. An evaluation of the characteristic function shows that a central limit theorem holds for the partial sums,  $S_n = \sum_{k=1}^n (I_k - 1/k), n \geq 1$ , viz.,  $S_n/\sqrt{\log n} \xrightarrow{d} \mathcal{N}(0, 1)$  as  $n \rightarrow \infty$ . Hence, the finite dimensional limit distributions of  $\{S_{n^t}/\sqrt{\log n}, n \geq 1\}$ , converge to those of a Wiener process, where now  $t \in [1, 2]$ .

Tightness follows by Theorem 13.5 and inequality (13.14) in Billingsley (1999), via the moment inequality

$$E \left( \left( \sum_{k=n^r+1}^{n^s} (I_k - 1/k)/\sqrt{\log n} \right)^2 \left( \sum_{k=n^s+1}^{n^t} (I_k - 1/k)/\sqrt{\log n} \right)^2 \right) \leq c(t - r)^2,$$

for  $n \geq 1, 1 \leq r \leq s \leq t \leq 2$ . Therefore,

$$0 \leq \sum_{k=n^s+1}^{n^t} \frac{1}{k} \left( 1 - \frac{1}{k} \right) \leq \sum_{k=n^s+1}^{n^t} \int_{k-1}^k \frac{1}{u} du \leq (t - s) \log n \quad \text{for } t \geq s.$$

Hence  $S_{n^t}/\sqrt{\log n} \xrightarrow{\mathcal{D}} W(t)$  as  $n \rightarrow \infty$  on  $\mathcal{D}[1, 2]$ .

Since Eq. 1.4 is automatic, it remains to check Eqs. 1.2 and 1.3. For the first relation, consider a sequence  $\{\varepsilon_n\} \searrow 0$  and the subsequence  $e^{1/\varepsilon_n^2}, n \geq 1$ . In order

to approximate a sum by an integral we define intermediate points  $\tau_{k,n} = k^{\varepsilon_n^2}$  with increments  $\Delta\tau_{k,n} = \tau_{k+1,n} - \tau_{k,n} = \tau_{k,n}\varepsilon_n^2/k(1+o(1))$  uniformly for  $1 \leq k \leq e^M$  for some large  $M$ . Then, by the weak invariance principle and the continuous mapping theorem, we obtain, as  $n \rightarrow \infty$ ,

$$\begin{aligned} \varepsilon_n^2 \sum_{k=1}^{e^{M/\varepsilon_n^2}} \frac{1}{k} \mathbb{I}\{|S_k| > \varepsilon_n \log k\} &= \varepsilon_n^2 \sum_{k=1}^{e^{M/\varepsilon_n^2}} \frac{1}{k} \mathbb{I}\{|S_{k^{\varepsilon_n^2}/\varepsilon_n^2}|/\sqrt{1/\varepsilon_n^2} > \varepsilon_n^2 \log k\} \\ &= \varepsilon_n^2 \sum_{\tau_{k,n} \leq e^M} \frac{1}{k \Delta \tau_{k,n}} \mathbb{I}\{|S_{\tau_{k,n}^{\varepsilon_n^2}}|/\sqrt{1/\varepsilon_n^2} > \log \tau_{k,n}\} \Delta \tau_{k,n} \\ &= \sum_{\tau_{k,n} \leq e^M} \frac{1}{\tau_{k,n}} (1 + o(1)) \mathbb{I}\{|S_{e^{(\log \tau_{k,n})/\varepsilon_n^2}}|/\sqrt{1/\varepsilon_n^2} > \log \tau_{k,n}\} \Delta \tau_{k,n} \\ &\xrightarrow{d} \int_1^{e^M} \mathbb{I}\{|W(\log v)| > \log v\} \frac{dv}{v} = \int_0^M \mathbb{I}\{|W(v)| > v\} dv. \end{aligned}$$

For Eq. 1.3 we exploit the exponential bound (5.4) from Gut (1990):

$$\begin{aligned} P(|\mu(n) - m_n| > \varepsilon \log n) &\leq \exp\left\{-\frac{1}{2}\varepsilon^2(1 - \frac{1}{2}\varepsilon) \log n\right\} \\ &\leq 2n^{-\varepsilon^2/4} \text{ (for } \varepsilon \in (0, 1)), \end{aligned} \tag{3.4}$$

and obtain

$$\begin{aligned} E\left(\varepsilon_n^2 \sum_{k \geq e^{M/\varepsilon_n^2}} \frac{1}{k} \mathbb{I}\{|S_k| > \varepsilon_n \log k\}\right) &= \varepsilon_n^2 \sum_{k \geq e^{M/\varepsilon_n^2}} \frac{1}{k} P(|S_k| > \varepsilon_n \log k) \\ &\leq \varepsilon_n^2 \sum_{k \geq e^{M/\varepsilon_n^2}} \frac{1}{k} \exp\{-\frac{\varepsilon_n^2}{4} \log k\} \leq \varepsilon_n^2 \sum_{k \geq e^{M/\varepsilon_n^2}} \frac{1}{k^{1+\varepsilon_n^2/4}} \leq 2e^{-M/4}, \end{aligned}$$

which decreases as  $M$  increases. Hence Eq. 1.3 is satisfied.

Finally, replacing  $m_n$  by  $\log n$  in the centering is achieved via Proposition A.2, since  $(m_n - \log n)/\log n \rightarrow 0$  as  $n \rightarrow \infty$ . □

*Remark 3.2* In Gut (2002), Theorem 3.1, it was (i.a.) shown that

$$\lim_{\varepsilon \searrow 0} \varepsilon^2 E N(\varepsilon) = \sum_{k=1}^{\infty} \frac{1}{k} P(|\mu(k) - \log k| > \varepsilon k \log k) = 1.$$

There also exist results on the integrability of the counting variable analogous to those of Section 3. Technically, with

$$N(\varepsilon) = \sum_{k=1}^{\infty} \frac{1}{k} \mathbf{I}\{|\mu(k) - \log k| > \varepsilon \log k\} \quad \text{and}$$

$$N_0(\varepsilon) = \sum_{k=1}^{\infty} \mathbf{I}\{|\mu(k) - \log k| > \varepsilon \log k\},$$

the following result holds.

**Proposition 3.1** *For any  $\varepsilon > 0$ ,*

(i):  $E(\log N_0(\varepsilon)) < \infty$  ;

(ii):  $E(N(\varepsilon)) < \infty$ .

*Proof* For (i) we refer to Gut (1990), Theorem 8, and since  $E N(\varepsilon) = \sum_{n=1}^{\infty} \frac{1}{n} P(|\mu(n) - \log n| > \varepsilon \log n)$ , (ii) follows from the fact that  $m_n - \log n \rightarrow \gamma$  as  $n \rightarrow \infty$  and the exponential bound (3.4). □

As in the previous section, we have, however, no distributional results for  $N_0$ .

### 3.3 Record values

So far in this section we have discussed record times and the corresponding counting process. A third sequence of interest is the sequence of *record values*:

$$X_{L(n)}, \quad n \geq 1.$$

The strong law runs as follows, cf. e.g. Resnick (1987), p. 172:

$$\frac{\log R(X_{L(n)})}{n} \xrightarrow{a.s.} 1 \quad \text{as } n \rightarrow \infty.$$

Moreover, under the additional assumption that

$$\lim_{n \rightarrow \infty} \frac{Q(\exp(n + t\sqrt{n \log \log n}))}{Q(\exp(n))} = 1 \quad \text{for all } t \in \mathbb{R}, \tag{3.5}$$

we have

$$\frac{X_{L(n)}}{Q(e^n)} \xrightarrow{a.s.} 1 \text{ as } n \rightarrow \infty.$$

As for distributional asymptotics, the class of limit laws is of the form

$$\mathcal{N}(-\log(-\log B(x))),$$

where, again,  $\mathcal{N}$  is the standard normal distribution and  $B$  is one of the extreme value distributions, cf. e.g. Resnick (1987), p. 176. If  $F \in \mathcal{A}(\Lambda)$  then the limit distribution is normal.

Whereas the sequence of partial maxima  $\{M_n, n \geq 1\}$  describes the largest value so far, the record values arise as the subsequence that selects the successive maximal values the first time they appear.

In the special case of the exponential distribution, the record values are gamma distributed, so that their asymptotics are immediate from Theorem 1.2. The general case follows from the fact that  $\log R(X) \in \text{Exp}(1)$  and continuity. This is the content of our first result below.

**Theorem 3.3** *Suppose that  $F$  is a continuous distribution, and set*

$$N(\varepsilon) = \sum_{k=1}^{\infty} \mathbb{I}\{X_{L(k)} > Q(e^{(1+\varepsilon)k})\}.$$

Then

$$\varepsilon^2 N(\varepsilon) \xrightarrow{d} \int_0^{\infty} \mathbb{I}\{W(y) > y\} dy \quad \text{as } \varepsilon \searrow 0.$$

*Proof* As hinted at a few lines ago, if  $X \in \text{Exp}(1)$ , then  $X_{L(k)} \in \Gamma(k, 1)$  for all  $k$ , and the conclusion follows from Theorem 1.2.

In the general case  $\log R(X) \in \text{Exp}(1)$  and  $\log R(\cdot)$  is continuous, which tells us that the conclusion of the theorem holds for  $\log R(X_{L(k)})$ , viz., that

$$\varepsilon^2 \sum_{k=1}^{\infty} \mathbb{I}\{\log R(X_{L(k)}) > (1 + \varepsilon)k\} \xrightarrow{d} \int_0^{\infty} \mathbb{I}\{W(y) > y\} dy \quad \text{as } \varepsilon \searrow 0.$$

The proof is complete upon noticing that  $Q(x) = R^{-1}(x)$ , and that

$$\mathbb{I}\{\log R(X_{L(k)}) > (1 + \varepsilon)k\} = \mathbb{I}\{R(X_{L(k)}) > e^{(1+\varepsilon)k}\} = \mathbb{I}\{X_{L(k)} > Q(e^{(1+\varepsilon)k})\}.$$

□

*Remark 3.3* Comparing with Theorem 2.1 we observe that the factor  $\frac{1}{k}$  there is missing here. This can (intuitively) be explained by the fact that the sequence of record values is a contraction of the sequence of partial maxima. This might roughly be explained by the fact that

$$M_n = X_{L(k)} \quad \text{for } L(k) \leq n < L(k + 1), \tag{3.6}$$

which, due to Williams’ representation (Williams 1973), recall (3.2) tells us that

$$\sum_{n=L(k)}^{L(k+1)-1} \frac{1}{n} \sim \log \left( \frac{L(k + 1)}{L(k)} \right) \sim E_k, \tag{3.7}$$

where we also note that  $E(E_k) = 1$ .

We shall delve on this a bit further (and more clearly) in Section 4.

*Remark 3.4* There also exists a two-sided version of the theorem. We leave the formulation and proof to the readers.

For distributions with support on the whole positive axis we obtain the following result, in which the norming is the same as for the partial maxima in Theorem 2.1.

**Theorem 3.4** *If, in particular,  $t_F = \infty$  and, for some  $c > 0$  and any fixed  $M$ ,*

$$(1 + \varepsilon)Q(e^k) = Q(e^{(1+c\varepsilon+o(\varepsilon))k}) \quad \text{uniformly in } k \leq M/\varepsilon^2 \text{ as } \varepsilon \searrow 0, \tag{3.8}$$

*and there exists some  $\beta > 0$ , such that*

$$(1 + \varepsilon)Q(e^k) \geq Q(e^{(1+\beta\varepsilon)k}) \quad \text{for all small } \varepsilon \text{ and large } k, \tag{3.9}$$

*then*

$$\varepsilon^2 \tilde{N}(\varepsilon) = \varepsilon^2 \sum_{k=1}^{\infty} \mathbb{I}\{X_{L(k)} > (1 + \varepsilon)Q(k)\} \xrightarrow{d} \int_0^{\infty} \mathbb{I}\{W(y) > cy\} dy \quad \text{as } \varepsilon \searrow 0.$$

*Proof* The first condition is needed to prove that

$$\begin{aligned} \varepsilon^2 \sum_{k \leq M/\varepsilon^2} \mathbb{I}\{X_{L(k)} > (1 + \varepsilon)Q(e^k)\} &= \varepsilon^2 \sum_{k \leq M/\varepsilon^2} \mathbb{I}\{X_{L(k)} > Q(e^{(1+c\varepsilon+o(\varepsilon))k})\} \\ &= \varepsilon^2 \sum_{k \leq M/\varepsilon^2} \mathbb{I}\{\log R(X_{L(k)}) > (1 + c\varepsilon + o(\varepsilon))k\} \\ &\xrightarrow{d} \int_0^M \mathbb{I}\{W(v) > cv\} dv \\ &\quad \text{as } \varepsilon \searrow 0 \quad \text{as before.} \end{aligned}$$

The second condition is needed to show that

$$P(X_{L(k)} > (1+\varepsilon)Q(e^k)) \leq P(X_{L(k)} > Q(e^{(1+\beta\varepsilon)k})) = P(\Gamma_k > (1+\beta\varepsilon)k) \leq \frac{\kappa\beta}{k^2\varepsilon^4},$$

which suffices to deal with the remaining sum in the usual manner. □

*Remark 3.5* Condition (3.8) is related to Condition (3.5), since, with  $s = \varepsilon\sqrt{k}$ ,

$$\frac{Q(\exp(k + c\varepsilon k))}{(1 + \varepsilon)Q(\exp(k))} = \frac{Q(e^{k+c s\sqrt{k}})}{(1 + s/\sqrt{k})Q(e^k)} = (1 + o(1)).$$

*Example 3.1* If  $1 - F(x) = e^{-\beta x^\alpha}$ ,  $x \geq 0$ , for some  $\alpha, \beta > 0$  then  $Q(e^y) = (y/\beta)^{1/\alpha}$  and both assumptions of Theorem 3.4 are satisfied; for the constant  $c$  we find  $c = \alpha$ . In particular,  $c = 2$  for the normal distribution.

### 4 A connection between extremes and record values

We now return to Remark 3.3 for further elucidation. Recalling (3.6), we thus know that, for all  $n$ ,

$$M_n = X_{L(k)}, \quad \text{for } L(k) \leq n < L(k + 1) \quad \text{for some } k, \quad (4.1)$$

which implies that

$$\sum_{n=1}^{\infty} \mathbb{I}\{M_n > x\} = \sum_{k=1}^{\infty} \sum_{j=L(k)}^{L(k+1)-1} \mathbb{I}\{X_{L(k)} > x\}.$$

We begin with the standard exponential case, when, in addition,  $L(k + 1) - L(k)$  and  $X_{L(k)}$  are independent. Moreover, and  $Q(x) = \log x$ , so that  $Q(L(k)) = \log L(k) \sim k$ , where  $\sim$  is a consequence of the strong law cited earlier.

Recalling Williams’ representation (3.2) we then obtain

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n} \mathbb{I}\{M_n > (1 + \varepsilon)Q(n)\} &= \sum_{n=1}^{\infty} \frac{1}{n} \mathbb{I}\{M_n > (1 + \varepsilon) \log n\} \\ &= \sum_{k=1}^{\infty} \sum_{j=L(k)}^{L(k+1)-1} \frac{1}{j} \mathbb{I}\{X_{L(k)} > (1 + \varepsilon) \log j\} \\ &\left\{ \begin{aligned} &\leq \sum_{k=1}^{\infty} \left( \sum_{j=L(k)}^{L(k+1)-1} \frac{1}{j} \right) \mathbb{I}\{X_{L(k)} > (1 + \varepsilon) \log L(k)\}, \\ &\geq \sum_{k=1}^{\infty} \left( \sum_{j=L(k)}^{L(k+1)-1} \frac{1}{j} \right) \mathbb{I}\{X_{L(k)} > (1 + \varepsilon) \log L(k + 1)\} \end{aligned} \right. \\ &\left\{ \begin{aligned} &\leq \sum_{k=1}^{\infty} (\log(L(k + 1)/L(k)) + 1/k) \mathbb{I}\{X_{L(k)} > (1 + \varepsilon) \log L(k)\}, \\ &\geq \sum_{k=1}^{\infty} (\log(L(k + 1)/L(k)) - 1/k) \mathbb{I}\{X_{L(k)} > (1 + \varepsilon) \log L(k + 1)\} \end{aligned} \right. \\ &\left\{ \begin{aligned} &\leq \sum_{k=1}^{\infty} (E_k + 1/k) \mathbb{I}\{X_{L(k)} > (1 + \varepsilon) \log L(k)\}, \\ &\geq \sum_{k=1}^{\infty} (E_k - 1/k) \mathbb{I}\{X_{L(k)} > (1 + \varepsilon) \log L(k + 1)\} \end{aligned} \right. \end{aligned}$$

The left-hand side is of order  $\mathcal{O}_p(1/\varepsilon)$  by Theorem 2.1. Observing that a series with weight  $1/k$  needs roughly one order  $\varepsilon^2$  less for normalization than one with weight 1, the right-hand side must be of the same order. However, Theorem 3.3 tells us that

$\sum_{k=1}^{\infty} \mathbf{I}\{X_{L(k)} > Q(e^{(1+\varepsilon)k})\}$  is of order  $\mathcal{O}_p(1/\varepsilon^2)$ —but note, in the latter sum we have a deterministic threshold whereas in the sums above it is random.

This can be understood a little better by the following considerations. Applying the strong law for record times,  $\log L(k)/k \xrightarrow{a.s.} 1$  as  $k \rightarrow \infty$ , we continue as follows; recall that  $Q(x) = \log x$ :

$$\begin{aligned} & \sum_{n=1}^{\infty} \frac{1}{n} \mathbf{I}\{M_n > (1 + \varepsilon)Q(n)\} \\ &= \begin{cases} \leq \sum_{k=1}^{\infty} (E_k + 1/k) \left( \mathbf{I}\{X_{L(k)} > (1 + \frac{1}{2}\varepsilon)k\} + \mathbf{I}\{\log(L(k)) < (1 - \frac{1}{2}\varepsilon)k\} \right), \\ \geq \sum_{k=1}^{\infty} (E_k - 1/k) \mathbf{I}\{X_{L(k)} > (1 + 3\varepsilon)k, \log(L(k+1)) \leq (1 + \varepsilon)(k+1)\} \end{cases} \\ & \begin{cases} \leq \sum_{k=1}^{\infty} (E_k + 1/k) \left( \mathbf{I}\{X_{L(k)} > (1 + \frac{1}{2}\varepsilon)k\} + \mathbf{I}\{\log(L(k)) < (1 - \frac{1}{2}\varepsilon)k\} \right), \\ \geq \sum_{k=1}^{\infty} (E_k - 1/k) \left( \mathbf{I}\{X_{L(k)} > (1 + 3\varepsilon)k\} - \mathbf{I}\{\log(L(k+1)) > (1 + \varepsilon)(k+1)\} \right) \end{cases} \\ & \begin{cases} \leq \sum_{k=1}^{\infty} (E_k + 1/k) \left( \mathbf{I}\{X_{L(k)} > Q(e^{(1+\frac{1}{2}\varepsilon)k})\} + \mathbf{I}\{\log(L(k)) < (1 - \frac{1}{2}\varepsilon)k\} \right), \\ \geq \sum_{k=1}^{\infty} (E_k - 1/k) \left( \mathbf{I}\{X_{L(k)} > Q(e^{(1+3\varepsilon)k})\} - \mathbf{I}\{\log(L(k+1)) > (1 + \varepsilon)(k+1)\} \right). \end{cases} \end{aligned}$$

Interpreting the  $E_k$ :s by their average 1 we observe that the sums dealing with  $\log(L(\cdot))$  are of order  $\mathcal{O}_p(1/\varepsilon^2)$  in view of Theorem 3.1; recall also Remark 3.1, which states that this is also true “on average”. Hence, using the deterministic threshold  $Q(e^{(1+\varepsilon)k})$  instead of the random  $\log(L(k))$  for  $X_{L(k)}$  generates a larger variability. The reason for the different  $\varepsilon$ -rates as  $\varepsilon \searrow 0$  in Theorems 2.1 and 3.3 can be understood from the fact, mentioned in Remark 3.3, that the record values are contractions of the partial maxima, in the sense that the latter realizations consist of long constant stretches between jumps, whereas the former consist of the jumps only (recall (4.1)). This leads to substantially more variability in the distribution of the record values, or equivalently, a much smoother behavior in the sequence of partial maxima. As for the general case we remember that  $\tilde{X} = \log R(X) \in \text{Exp}(1)$ . Similar considerations apply using this transformation. We omit the details.

**Acknowledgments** The results of this paper were initiated during A.G:s visit at the Department of Number Theory and Probability Theory at the University of Ulm in May 2014. He wishes to thank his coauthor, U.S., for providing a great stay in Ulm, as well for his generous hospitality. A.G. also wishes to thank the University of Ulm and Kungliga Vetenskapssamhället i Uppsala for financial support. We finally wish to express our thanks to an anonymous referee for several very thoughtful comments and remarks which improved the presentation of paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



## Appendix A

In order not to disturb the flow of the text we collect, in this appendix, three auxiliary results that have been used above.

### A.1 A continuous mapping theorem

The following result gives a continuous mapping theorem which is suitable for our purposes.

**Proposition A.1** *Let  $v : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a measurable bounded mapping for which the set of discontinuities has zero Lebesgue measure in  $\mathbb{R}^2$ . Let  $M > 0$  and  $(D, \mathcal{D})$  be the space  $(D[0, M], \mathcal{D}[0, M])$  endowed with the Skorohod topology  $d$  and the associated Borel  $\sigma$ -algebra. Then*

$$\int_0^M v(x(t), t) dt,$$

*as a mapping from  $(D, \mathcal{D})$  to  $\mathbb{R}$ , is measurable and such that its set of discontinuities has measure zero with respect to the Gumbel process  $Y_{0,\Lambda}$  in  $(D, \mathcal{D})$ .*

*Proof* If  $x$  is a continuous function on  $[0, M]$ , then  $(x, t) \rightarrow x(t)$ , as a mapping from  $(D, \mathcal{D}) \times [0, M]$  to  $(D, \mathcal{D})$ , endowed with the product  $\sigma$ -algebra, to  $\mathbb{R}_+$ , endowed with the Borel  $\sigma$ -algebra, is measurable. The integral  $\int_0^M v(x(t), t) dt$  is thus well defined. Furthermore, we find that  $\psi(x, t) = v(x(t), t)$ , as a mapping from  $(D, \mathcal{D}) \times [0, M]$  to  $\mathbb{R}_+$ , is measurable and bounded, which implies that the integral is measurable as a mapping from  $(D, \mathcal{D})$  to  $\mathbb{R}$ . Next, let  $D_v$  be the set of discontinuities of  $v$ , let  $E \subset (D, \mathcal{D}) \times [0, M]$  be the set defined as  $E = \{(x, t), t \in D_v\}$ , and let  $\lambda_2$  denote Lebesgue measure in the plane. Then  $\lambda_2(D_v) = 0$ , from which it follows that

$$\Lambda\{x(\cdot) : (x(t), t) \in E\} = \int_{\{u \geq 0 : u=t\}} e^{-u} e^{-e^{-u}} du = 0 \quad \text{for any fixed } t > 0.$$

For the product measure  $W \times \lambda$  on  $(D, \mathcal{D}) \times [0, M]$  we therefore conclude that  $(\Lambda \times \lambda)(E) = 0$ , and, hence, that

$$\lambda\{t : \psi(x, t) \in E\} = 0 \quad \text{for all } x \in (D, \mathcal{D}),$$

except, possibly, for a  $\Lambda$ -nullset  $A$ . Now, if  $(x_n)$ ,  $x_n \in (D, \mathcal{D})$ , converges uniformly to  $x \in (D, \mathcal{D})$ , then  $v(x_n(t), t) \rightarrow v(x(t), t)$  for each  $t$  such that  $(x(t), t) \notin D_v$ . If  $x \notin A$ , the latter conclusion is true for almost all  $t$ . An application of the bounded convergence theorem therefore tells us that

$$\int_0^M \psi(x_n, t) dt \rightarrow \int_0^M \psi(x, t) dt \quad \text{as } n \rightarrow \infty,$$

which establishes the desired continuity statement. □

### A.2 A Cramér-theorem

In this subsection we present an analog of the Cramér–Slutsky theorem related to the present setting. Let  $(D, \mathcal{D})$  denote the Skorohod space on  $[0, M]$  for some  $M > 0$ .

**Proposition A.2** *Suppose that  $U_n \in (D, \mathcal{D})$  and  $V_n \in (D, \mathcal{D})$  for all  $n \geq 1$ , that  $U_n \implies U \in (D, \mathcal{D})$  and  $V_n \implies 0$  as  $n \rightarrow \infty$ . Finally, assume that the marginal distribution functions of  $U$  are continuous. If*

$$\int_0^M \mathbb{I}\{|U_n(y)| > y\} dy \xrightarrow{d} \int_0^M \mathbb{I}\{|U(y)| > y\} dy \quad \text{as } n \rightarrow \infty,$$

then

$$\int_0^M \mathbb{I}\{|U_n(y) + V_n(y)| > y\} dy \xrightarrow{d} \int_0^M \mathbb{I}\{|U(y)| > y\} dy \quad \text{as } n \rightarrow \infty.$$

*Proof* Since  $U_n + V_n \implies U$  as  $n \rightarrow \infty$ , the conclusion follows by arguing as in the proof of Proposition A.1. □

### A.3 Slicing

The next result is a useful tool for considering sums of  $P(\sup_{k \geq n} \frac{S_k}{k} > x)$ . The proof consists of a modification of the Baum–Katz “slicing device” (Baum and Katz 1965, cf. also e.g. Gut 2013, Section 6.12).

**Proposition A.3** *Suppose that  $U_n, n \geq 1$ , is a positive, non-decreasing sequence of random variables and that  $a_k, k \geq 1$ , is a positive, non-decreasing sequence in  $RV_\alpha$  with  $\alpha \geq 0$ . Then,*

$$\sum_{n=1}^{\infty} \frac{\log n}{n} P(U_n > a_n \varepsilon) < \infty \quad \text{for all } \varepsilon > 0, \tag{A.1}$$

$\implies$

$$\sum_{n=1}^{\infty} \frac{1}{n} P(\sup_{k \geq n} \frac{U_k}{a_k} > \varepsilon) < \infty \quad \text{for all } \varepsilon > 0. \tag{A.2}$$

Moreover, convergence of the second series is equivalent to existence of the logarithmic moment of the last exit time, viz.,  $E \log T(\varepsilon)$ , where  $T(\varepsilon) = \sup\{n : U_n > a_n \varepsilon\}$ .

*Proof* We thus suppose that (A.1) holds. Taking advantage of the fact that  $a$  is regularly varying, we argue as follows:

$$\begin{aligned}
 \sum_{n=1}^{\infty} \frac{1}{n} P\left(\sup_{k \geq n} \frac{U_k}{a_k} > \varepsilon\right) &= \sum_{i=0}^{\infty} \sum_{j=2^i}^{2^{i+1}-1} \frac{1}{j} \cdot P\left(\sup_{k \geq j} \frac{U_k}{a_k} > \varepsilon\right) \\
 &\leq \sum_{i=0}^{\infty} 1 \cdot P\left(\sup_{k \geq 2^i} \frac{U_k}{a_k} > \varepsilon\right) \\
 &\leq \sum_{i=0}^{\infty} \sum_{j=i+1}^{\infty} P\left(\max_{2^{j-1} \leq k < 2^j} \frac{U_k}{a_k} > \varepsilon\right) \\
 &\leq \sum_{i=0}^{\infty} \sum_{j=i+1}^{\infty} P(U_{2^j} > a_{2^{j-1}} \varepsilon) \\
 &= \sum_{j=1}^{\infty} \left(\sum_{i=0}^{j-1} 1\right) \cdot P(U_{2^j} > a_{2^{j-1}} \varepsilon) \\
 &\leq \sum_{j=0}^{\infty} j \cdot P(U_{2^j} > a_{2^{j-1}} \varepsilon) \\
 &\leq \frac{1}{\log 2} \sum_{j=0}^{\infty} \frac{\log 2^j}{2^j} \sum_{i=2^j}^{2^{j+1}-1} P(U_{2^j} > a_{2^{j-1}} \varepsilon) \\
 &\leq C \sum_{i=1}^{\infty} \frac{\log i}{i} P(U_i > c a_i \varepsilon) < \infty.
 \end{aligned}$$

The equivalence of (A.2) with the logarithmic moment of the last exit times is immediate from the fact that  $\{T(\varepsilon) \geq n\} = \sup_{k \geq n} \{U_k/a_k > \varepsilon\}$ . □

*Remark A.1* (i): In contrast to the case of sums, (Baum and Katz 1965), the inverse conclusion (A.1)  $\implies$  (A.2) is not true in general. Consider e.g.  $U_k \equiv U$  where  $\sup_{k \geq n} U_k/a_k = U/a_n = U_n/a_n$ .

(ii): The result also holds true for a sequence  $\{U_k\}$  of random variables being not necessarily non-decreasing, such that with some  $\kappa, c_1 > 0$

$$P\left(\sup_{1 \leq k \leq n} U_k > x\right) \leq \kappa P(U_n > c_1 x).$$

## References

- Barndorff-Nielsen, O.: On the limit behaviour of extreme order statistics. *Ann. Math. Statist.* **34**, 992–1002 (1963)
- Baum, L.E., Katz, M.: Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120**, 108–123 (1965)
- Bingham, N.H., Goldie, C.M., Teugels, J.L.: *Regular Variation*. Cambridge University Press (1987)
- Billingsley, P.: *Convergence of Probability Measures*. Wiley (1999)
- de Haan, L., Geluk, J.: *Regular Variation, Extensions and Tauberian Theorems*. CWI tracts 40, Amsterdam (1980)
- de Haan, L., Hordijk, A.: The rate of growth of sample maxima. *Ann. Math. Statist.* **43**, 1185–1196 (1972)
- Embrechts, P., Klüppelberg, C., Mikosch, T.: *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, Heidelberg (1997)
- Erdős, P.: On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20**, 286–291 (1949)
- Erdős, P.: Remark on my paper “On a theorem of Hsu and Robbins”. *Ann. Math. Statist.* **21**, 138 (1950)
- Fahrner, I., Stadtmüller, U.: Strong approximation of maxima by extremal processes. *Ann. Probab.* **31**, 926–947 (2003)
- Geluk, J.L., de Haan, L.: *Regular variation, extensions and tauberian theorems*. CWI tracts 40, Amsterdam (1987)
- Gut, A.: Convergence rates for record times and the associated counting process. *Stochastic Process. Appl.* **36**, 135–151 (1990)
- Gut, A.: Precise asymptotics for record times and the associated counting process. *Stochastic Process. Appl.* **101**, 233–239 (2002)
- Gut, A.: *Probability: a Graduate Course*, 2nd edn. Springer, New York (2013)
- Heyde, C.C.: A supplement to the strong law of large numbers. *J. Appl. Probab.* **12**, 173–175 (1975)
- Hsu, P.L., Robbins, H.: Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. USA* **33**, 25–31 (1947)
- Kao, C.-S.: On the time and the excess of linear boundary crossings of sample sums. *Ann. Statist.* **6**, 191–199 (1978)
- Nevzorov, V.B.: *Records: Mathematical Theory*. Translation of Mathematical Monographs. Amer. Math. Soc 194, Providence, RI (2001)
- Resnick, S.I.: *Extreme Values, Regular Variation, and Point Processes*. Springer (1987)
- Resnick, S.I., Tomkins, R.J.: Almost sure stability of maxima. *J. Appl. Probab.* **10**, 387–401 (1973)
- Williams, D.: On Rényi’s “record” problem and Engel’s series. *Bull. London Math. Soc.* **5**, 235–237 (1973)