Z. Wahrscheinlichkeitstheorie verw. Gebiete
58, 1-6 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

The Strong Approximation of Extremal Processes

Paul Deheuvels

7, avenue du Château, F-92340 Bourg-la-Reine

Summary. If $X_1, X_2, \ldots,$ are i.i.d. random variables and $Y_n = \operatorname{Max}(X_1, \ldots, X_n)$; if for some sequences A_n, B_n, $n=1,2,\ldots$, $E_n(t)=A_n Y_{[nt]}+B_n$ is such that $E_n(1)$ weakly converges to a non degenerate limit distribution, then we prove that it is possible to construct a sequence of replicates of extremal processes $E^{(n)}(t)$ on the same probability space, such that $d(E_n(.), E^{(n)}(.))\to 0$ a.s., with the Levy metric. We give the rates of consistency of the approximations.

1. Introduction

If $X_1, X_2, \ldots,$ is an i.i.d. sequence with common d.f. $F(x)=P(X\leqq x)$, noting $\Phi(x)=1-F(x)$, and if for some sequences $a_n>0$, b_n, $n=1,2,\ldots$, $\operatorname{Lim}_{n\infty} F^n(a_n x+b_n) =\exp(-H(x))$, $H(x)$ being one among the possible limit functions e^{-x}, x^{-a}, $(-x)^a$, then, if $Y_n=\operatorname{Max}(X_1, \ldots, X_n)$, it is well known that the process $E_n(t) =a_n^{-1}(Y_{[nt]}-b_n)$ has all its finite distributions converging to those of an extremal process $E(t)$ (see [4], p. 304).

The aim of the following is to give a strong approximation result in the following sense: we shall show that there exists a sequence of replications of the extremal process $E^{(n)}(t)$, $n=1,2,\ldots$, defined on the same probability space (eventually enlarged) as the initial sequence, and such that for arbitrary $t\geqq 0$,

$$\begin{aligned} E_n(t+u_n^1(t)) &\leqq E^{(n)}(t)+v_n(t) \\ E_n(t+u_n^2(t)) &\geqq E^{(n)}(t)+w_n(t) \end{aligned} \tag{1}$$

where $u_n^i(t)=n^{-1}\,O(\operatorname{Log} nt)$, $nt\uparrow\infty$, $i=1,2$, and v_n, w_n going to zero with rates depending on H.

To obtain (1), we will first transform the sequence by the use of the classical representation $X_n=G(U_n)$, $G(u)=\operatorname{Max}\{v; \Phi(v)>u\}$, U_n, $n=1,2,\ldots$, being an i.i.d. sequence of $U(0,1)$ random variables, so that $Y_n=G(V_n)$, with $V_n =\operatorname{Min}(U_1, \ldots, U_n)$.

0044-3719/81/0058/0001/$01.20

Taking now $e_n(t)=nV_{[nt]}$ and $E_n(t)=a_n^{-1}(Y_{[nt]}-b_n)$, it can be seen that

$$E_n(t)=a_n^{-1}(G(n^{-1}e_n(t))-b_n). \tag{2}$$

Furthermore $F^n(a_n x+b_n)\to\exp(-H(x))$, when $n\to\infty$ iff exists sequences $\alpha_n(.)$, c_n, such that

$$\begin{aligned}&\text{(i) } c_n\downarrow c=H^{-1}(+\infty)\\ &\text{(ii) If } d=H^{-1}(0), \text{ and } n\Phi(a_n x+b_n)=H(x)+\alpha_n(x), \ \operatorname*{Lim}_{n\infty}\{\sup_{c_n\leqq x\leqq d}|\alpha_n(x)|\}=0.\end{aligned} \tag{3}$$

This can be derived easily from the fact that $n\,\mathrm{Log}(1-\Phi(a_n x+b_n))\to H(x)$, $n\to\infty$, and that the consistency of a d.f. to a continuous d.f. is always uniform.

From (2) and (3), we can deduce, for a continuous F, that

$$E_n(t)=H^{-1}(e_n(t)-\alpha_n(E_n(t))), \tag{4}$$

where $\alpha_n(.)$ is defined in (3)(ii).

It is thus enough to derive a strong approximation result for $e_n(.)$ in order to obtain a corresponding version for $E_n(.)$ by (4).

We shall now do so, restricting the study to the sequences $U_1, U_2, \ldots$, and $V_n=\mathrm{Min}(U_1,\ldots,U_n)$.

2. Strong approximation of the extreme sequences. Let $(\Omega,\mathfrak{A},P)$ be a probability space, on which is defined an i.i.d. sequence $U_1, U_2, \ldots$, of $U(0,1)$ random variables. Put for $t>0$, $\tau_t=\mathrm{Inf}\{k\geqq 1,\ U_k\leqq t\}$, $V_n=\mathrm{Min}(U_1,\ldots,U_n)$; define the sequence of downcrossing levels by

$$v_0=1, \qquad v_k=\mathrm{Inf}\{t>v_{k-1};\tau_{1/t+0}-\tau_{1/t-0}\neq 0\}, \quad k=1,2,\ldots,$$

and the sequence of downcrossing points by

$$n_1=1, \qquad n_k=\mathrm{Inf}\{m>n_{k-1}; V_m<V_{m-1}\}, \quad k=2,3,\ldots.$$

It is to be noted that $v_1<v_2<\ldots$ is the ordered set of values taken by $\{1/V_n, n\geqq 1\}$, and that precisely

$$v_k=1/V_{n_k}, \qquad \tau_{1/v_k}=n_k, \quad k=1,2,\ldots; \qquad \tau_{t+0}=\operatorname*{Lim}_{\varepsilon\searrow 0}\tau_{t+\varepsilon}=\tau_t. \tag{5}$$

Lemma 1. *If* $z_k=\mathrm{Log}(1/v_k)$, $k=0,1,\ldots$, *then* $z_0=0<z_1<\ldots<z_k<\ldots$ *are the arrival times of a normalized Poisson process* $\{N(t), t\geqq 0\}$.

Proof. $\{\mathrm{Log}(1/U_i), i\geqq 1\}$ is an i.i.d. sequence of exponential $E(1)$ r.v., so that $\{z_k, k\geqq 1\}$ is the ordered set of the successive maxima of this sequence. It can then be readily seen that $\{z_k-z_{k-1}, k\geqq 0\}$ is an i.i.d. sequence of exponential $E(1)$ r.v., so that the result follows.

Knowing $z_{k-1}=\mathrm{Log}(1/v_{k-1})$ for some $k\geqq 1$, the conditional distribution of n_k-n_{k-1} is independent of n_{k-1}, and such that

$$P(n_k-n_{k-1}\geqq r|z_{k-1})=\left(1-\frac{1}{v_{k-1}}\right)^{r-1}, \qquad r=1,2,\ldots. \tag{6}$$

Lemma 2. *If I is a discrete random variable such that $P(I>0)=1$, and G is a random variable defined on the same probability space, and such that if*

$P(I=\alpha)>0$, $P(G\geqq r|I=\alpha)=\alpha^{r-1}$, $r=1,2,\ldots$, then there exists on the same probability space eventually enlarged an exponential $E(1)$ random variable Y, independent of I, and such that

$$G=\left[\frac{Y}{\mathrm{Log}(1/I)}\right]+1, \quad ([u] \textit{ standing for the integer part of } u). \tag{7}$$

Proof. First note that the converse assertion of the lemma is trivial, since if Y is $E(1)$ and independent of I, $P([Y/\mathrm{Log}(1/I)]+1\geqq r|I=\alpha)$ $=\exp(-(r-1)\mathrm{Log}(1/\alpha))=\alpha^{r-1}$.

To get the result, assume that $I=\alpha$, and consider the conditional distribution of $(G-1)\mathrm{Log}(1/I)$; we have $P((G-1)\mathrm{Log}(1/I)\geqq t|I=\alpha)=e^{-t}$ for all t belonging to $\{n\mathrm{Log}(1/\alpha), n=0,1,\ldots\}$. It is thus possible to construct for $I=\alpha$ an exponentially $E(1)$ distributed r.v. Y such that $G=[Y/\mathrm{Log}(1/I)]+1$, and that $P(Y\geqq t|I=\alpha)=e^{-t}$, $t>0$. Repeating the construction for all possible values of α, which amounts each time to show that a discrete r.v. can be considered in an enlarged probability space as a function of a continuous r.v., we obtain finally a r.v. Y whose distribution does not depend of I, giving the result.

As a consequence of Lemmas 1 and 2, we get

Theorem 1. *If $(\Omega,\mathfrak{A},P)$ is rich enough, it is possible to define on it a normalized Poisson process $\{N(t),t\geqq 0\}$ with arrival times denoted by $z_0=0<z_1<\ldots<z_k<\ldots$, and an i.i.d. sequence $\{\omega_n, n\geqq 1\}$ of exponential $E(1)$ random variables independent of $\{N(t),t\geqq 0\}$, and such that if*

$$Z^*(t)=1+\sum_{k=1}^{N(t)}\left(\left[\frac{\omega_k}{-\mathrm{Log}(1-e^{-z_k})}\right]+1\right), \quad t\geqq 0, \tag{8}$$

then

$$\begin{aligned} &\text{(i)}\ \ \tau_{1/t}=Z^*(\mathrm{Log}\,t), \quad t\geqq 1,\\ &\text{(ii)}\ \ v_k=\exp(z_k), \quad k=0,1,\ldots,\\ &\text{(iii)}\ \ n_k=Z^*(z_k)=1+\sum_{i=1}^{k-1}\left(\left[\frac{\omega_i}{-\mathrm{Log}(1-e^{-z_i})}\right]+1\right), \quad k=1,2,\ldots. \end{aligned} \tag{9}$$

$\left(\text{With the usual assumption that } \sum_{i=1}^{0}(.)=0\right)$. Furthermore, the process $\{Z^*(t),t\geqq 0\}$ and the process $\{\tau_{1/t},t\geqq 1\}$ have independent increments.

Proof. It is possible to deduce (9)(iii) from (6) and (7). Now, if $t\geqq 1$ is arbitrary, $\tau_{1/t}$ is equal to $\tau_{1/v_k}=n_k$, where v_k is the smallest $v_i\geqq t$, $i=1,2,\ldots$. The number of the $v_i<t$ is $N((\mathrm{Log}\,t)-0)=N(\mathrm{Log}\,t)$, so that $k=N(\mathrm{Log}\,t)+1$, and the result follows. Note that here $N(t)=N(t-0)$, and in particular, that $N(z_k)=k-1$, $k=1,2,\ldots$; this leads logically to $N(0)=-1$, which does not interfere with (8) and (9), since the corresponding summations give 0. The fact that $\{\tau_{1/t},t\geqq 1\}$, and hence $\{Z(t),t\geqq 0\}$, is a process with independent increments was proved in [2] (see also [5]).

Theorem 2. *Under the hypothesis of Theorem 1, $\{N(t),t\geqq 0\}$ being a normalized Poisson process, and $\{\omega_n,n\geqq 1\}$ being an i.i.d. independent $E(1)$ sequence, if the process $\{N(t),t\geqq 0\}$ is completed as a normalized Poisson process $\{N(t), -\infty<t<+\infty\}$ on the whole line, and likewise $\{\omega_n,n\geqq 1\}$ is completed by $\{\omega_n,$*

$-\infty<n<+\infty\}$, the extra variables being independent of the original sequence $\{U_n, n\geqq 1\}$, then if

$$Z(t)=\sum_{k=-\infty}^{N(t)} \omega_k e^{z_k}, \tag{10}$$

where $\ldots<z_{-1}<z_0<0<z_1<\ldots$ are the arrival points of $N(.)$,

$$\begin{aligned}&\operatorname*{LimSup}_{t\uparrow\infty}(\operatorname{Log} t)^{-1}(\tau_{1/t}-Z(\operatorname{Log} t))\leqq\frac{3}{2} \quad a.s.,\\ &\operatorname*{LimInf}_{t\uparrow\infty}(\operatorname{Log} t)^{-1}(\tau_{1/t}-Z(\operatorname{Log} t))\geqq\frac{-1}{2} \quad a.s.\end{aligned} \tag{11}$$

Proof. It is clear that Z(.) is defined, and that $\left|Z(t)-1-\sum_{k=1}^{N(t)}\omega_k e^{z_k}\right|=O(1)$ a.s.; now, if $s_k=\left[\frac{\omega_k}{-\operatorname{Log}(1-e^{-z_k})}\right]+1$,

$$\omega_k e^{z_k}-\tfrac{1}{2}\omega_k(1+O(e^{-z_k}))\leqq s_k\leqq\omega_k e^{z_k}-\tfrac{1}{2}\omega_k(1+O(e^{-z_k}))+2,$$

and, remarking that $z_k/k\to 1$ a.s., $(\omega_1+\ldots+\omega_k)/k\to 1$ a.s., and that $N(t)/t\to 1$ a.s., for $t, k\to+\infty$, (11) follows.

Before going further, we shall prove some facts concerning the $Z(\operatorname{Log} t)$ process.

Theorem 3. *If $Z(.)$ is defined by* (10), *and if $Y(t)=Z(\operatorname{Log} t)$, $t\geqq 0$, then*

1°) *The process $\{Y(t), t\geqq 0\}$ has independent, mean stationary, increments, $E(Y(t))=t$, $t\geqq 0$, and is a left continuous $(Y(t-0)=Y(t))$ increasing step process;*

2°) *For $0\leqq s\leqq t$, the characteristic function of $Y(t)-Y(s)$ is $\frac{1-ius}{1-iut}$; the distribution of $Y(t)-Y(s)$ being also given by*

$$\begin{aligned}&P(Y(t)-Y(s)=0)=s/t,\\ &P(Y(t)-Y(s)>u)=\left(1-\frac{s}{t}\right)e^{-u/t}, \quad u>0;\end{aligned} \tag{12}$$

3°) *For $t>0$, $Y(t)$ is an exponentially $E(1/t)$ distributed random variable;*

4°) *If $\lambda>0$, the process $\{Y(\lambda t)/\lambda, t\geqq 0\}$ is identical in distribution to the process $\{Y(t), t\geqq 0\}$.*

Proof. Following [6] p. 146, Theorem 5.A,

$$E(e^{iu(Z(t)-Z(s))})=\exp\left(\int_s^t E(e^{iu\omega e^v}-1)\,dv\right)=\frac{1-iue^s}{1-iue^t},$$

ω being in the preceding an $E(1)$ r.v. A similar calculus enables to obtain, for $0<s\leqq t$,

$$E(e^{iuZ(s)+iv(Z(t)-Z(s))})=\frac{1}{1-iue^s}\cdot\frac{1-ive^s}{1-ive^t};$$

the rest follows likewise. We may now derive a strong approximation result for the $\{\tau_{1/t}, t \geqq 1\}$ process.

Theorem 4. *If* $\delta_n(t) = n^{-1}\tau_{1/nt}$, $t > 0$, *then* $\varepsilon_n(t) = n^{-1}\tau_{1/nt} - n^{-1}Y(nt)$ *is such that*

$$(1+o(1))\frac{-\operatorname{Log}(nt)}{2n} \leqq \varepsilon_n(t) \leqq \frac{3\operatorname{Log}(nt)}{2n}(1+o(1)), \quad a.s., \quad when\ nt\uparrow\infty. \tag{13}$$

Proof. (13) merely restates (11), noting that by Theorem 3, 4°, $n^{-1}Y(nt)$ defines a process identical in distribution for each n to $Y(t)$.

To get a strong approximation of the extremal process itself, we may note that

$$\tau_{1/t} \leqq n \Leftrightarrow V_n \leqq 1/t; \qquad \tau_{1/t} > n \Leftrightarrow V_n > 1/t. \tag{14}$$

Taking (11) as $Y(t) - \phi(t) < \tau_{1/t} \leqq Y(t) + \psi(t)$, with $Y - \phi$ and $Y + \psi$ integers, and such that $\phi(t) \leqq (\operatorname{Log} t)(\frac{1}{2} + o(1))$, $\psi(t) \leqq (\operatorname{Log} t)(\frac{3}{2} + o(1))$, a.s., $t \to \infty$, we get by (14)

$$V_{Y(t)+\psi(t)} \leqq \frac{1}{t} < V_{Y(t)-\phi(t)}. \tag{15}$$

Let us now inverse $Y(.)$ in the following way

Definition. If $Y(t) = Z(\operatorname{Log} t)$, $Z(.)$ *being defined in* (10), *we define the inverse process* $\{W(u), u \geqq 0\}$ *by*

$$W(u) = \operatorname{Min}\{t; Y(t) \geqq u\} = \operatorname{Max}\{t; Y(t) < u\}. \tag{16}$$

Using (16), we can see that $W(.)$ is left continuous, $W(u-0) = W(u)$, that $W(0) = 0$, and that for $u \geqq 0$, $u \leqq Y(W(u))$, so that putting $t = W(u)$ in (15), we get

$$\frac{1}{W(u)} < V_{Y(W(u)) - \phi(W(u))} \leqq V_{[u - \phi(W(u))]}. \tag{17}$$

To get a bound on the other side, note that ϕ and ψ can be assumed to be constant in each interval such as (v_{k-1}, v_k), $Y(.)$ and $\tau_{1/.}$ having same points of increase, so that, by an eventual good choice of these functions,

$$V_{[u + \psi(W(u))]} \leqq \frac{1}{W(u)}. \tag{18}$$

Lemma 3. *For any* $\varepsilon > 0$, *almost surely for t large enough,*

$$\frac{t}{(\operatorname{Log} t)^{1+\varepsilon}} \leqq \tau_{1/t} \leqq t(1+\varepsilon)\operatorname{LogLog} t. \tag{19}$$

Proof. See [3], or equivalently [1].

It follows from (11) and (19), that $\displaystyle\operatorname*{Lim}_{u\downarrow\infty}\left(\frac{\operatorname{Log} W(u)}{\operatorname{Log} u}\right) = 1$ a.s.

Note that (15) is valid for $t \geqq 1$, and thus, that (17) and (18) assume implicitely that $W(u) \geqq 1$. By the preceding relation, $\operatorname*{Lim}_{u\downarrow\infty} W(u) = +\infty$, $W(.)$

being an increasing step function, and this will always be satisfied for u large enough a.s.

We have now obtained our final result.

Theorem 5. *If* $e_n(t)=nV_{[nt]}$, *where* $V_n=\mathrm{Min}(U_1,\ldots,U_n)$, *then there exists two functions of* $t\geqq 0$, $\rho(t)$ *and* $\gamma(t)$, *such that almost surely for* nt *large enough, say* $t\geqq C/n$,

$$e_n(t+\rho(nt))\leqq\frac{n}{W(nt)}\leqq e_n(t-\gamma(nt)),$$

with

$$0\leqq\rho(nt)\leqq\frac{3(\mathrm{Log}\,nt)(1+o(1))}{2n},\qquad 0\leqq\gamma(nt)\leqq\frac{(\mathrm{Log}\,nt)(1+o(1))}{2n},\quad a.s.$$

when $nt\uparrow+\infty$.

Proof. It is a direct consequence of (11), (17), (18), (19).

We can see here that $\{n/W(nt), t\geqq 0\}$ is a process with distribution independent of n. It defines in that case the approximating sequence of extremal processes.

Since by (3) and (4) we can deduce the general approximating sequences for an arbitrary extremal process, we can define such processes by

$$E_n^\infty(t)=H^{-1}(n/W(nt)). \tag{20}$$

This formula gives the general structure of extremal processes.

References

1. Barndorff-Nielsen, O.: On the rate of growth of the partial maxima of a sequence of independent identically distributed random variables. Math. Scand. **9**, 383–394 (1961)
2. Deheuvels, P.: Valeurs extremales d'échantillons croissants d'une variable aléatoire réelle. Ann. Inst. H. Poincaré Ser. B**10**, 89–114 (1974)
3. Deheuvels, P.: Majoration et minoration presque sûre optimale des éléments de la statistique ordonnée d'un échantillon croissant de variables aléatoires indépendantes. Rend. Acad. Nazionale dei Lincei Ser. VIII Vol. **LVI** 707–719 (1974)
4. Galambos, J.: The asymptotic theory of extreme order statistics. New York: Wiley 1978
5. Hebna-Grabowska, H., Szynal, D.: An almost sure invariance principale for the partial sums of infima of independent random variables. Ann. Probability **7**, 1036–1045 (1979)
6. Parzen, E.: Stochastic processes. San Francisco: Holden Day 1962

Received February 13, 1981