ON THE RATE OF CONVERGENCE FOR SOME STRONG APPROXIMATION THEOREMS IN EXTREMAL STATISTICS

Dietmar Pfeifer

Abstract. Strong approximation theorems for the logarithms of record and inter-record times of an i.i.d. sequence are investigated with respect to their a.s. rate of convergence which is shown to be exponential. Exact characterizations of this rate are given by means of upper and lower class functions for the Wiener process.

1. Introduction.

For an i.i.d. sequence $\{X_n; n \in \mathbb{N}\}$ of random variables with continuous distribution function let $\{\Delta_n; n \geq 0\}$ and $\{U_n; n \geq 0\}$ denote the inter-record times and record times, resp. defined by

$$(1) \quad \Delta_0 = 1, \ \Delta_{n+1} = \min\{k; X_{U_n+k} > X_{U_n}\} \quad \text{with } U_n = \sum_{k=0}^{n} \Delta_k, \ n \geq 0,$$

which are a.s. well-defined by the assumptions above.

Investigations concerning the a.s. limiting behaviour of these sequences were carried out by different authors (see

---

AMS 1980 subject classification: primary: 60 F 15, 60 J 20,
secondary: 60 J 80, 60 J 85

Key words and phrases: strong approximation, record times, inter-record times, Erdős-type LIL, Wiener process

Deheuvels [2]); however, it has turned out that strong approximation techniques as developed by Deheuvels [1],[2] are most useful in this area. In the present paper, we shall complete his results by a more detailed investigation of the rate of convergence in his basic representation theorems, allowing at the same time for a simple explanation for the different growth behaviour of record and inter-record times which e.g. becomes apparent from the fact that

$$(2)\quad \begin{aligned} &E(\log\Delta_n) = n - C + O(n2^{-n}), \quad Var(\log\Delta_n) = n + \frac{\pi^2}{6} + O(n^2 2^{-n}) \\ &E(\log U_n) = n + 1 - C + O(n^2 2^{-n}), \quad Var(\log U_n) = n+1 - \frac{\pi^2}{6} + O(n^3 2^{-n}) \end{aligned}$$

for $n \to \infty$ (see Pfeifer [5]), where C denotes Euler's constant.

2. Main results. Following Deheuvels [2], the inter-record time sequence can be represented as

$$(3)\quad \Delta_n = \mathrm{int}\{Y_n/-\log(1-\exp(-T_n))\}+1 = Y_n/-\log(1-\exp(-T_n)) + R_n,$$

say where $\{Y_n;\ n \in \mathbb{N}\}$ is an i.i.d. sequence of exponentially distributed random variables with unit mean, and $\{T_n; n \in \mathbb{N}\}$ is the arrival-time sequence of a unit-rate Poisson process, independent of the Y_n's. The following result is implicit in his paper [1], being a simple consequence of (3).

LEMMA 1. For $n \to \infty$, we have

$$(4)\quad \log\Delta_n = \log Y_n + T_n + \exp(-T_n)\left\{\frac{R_n}{Y_n} - \frac{1}{2}\right\} + \exp(-T_n)\,o(1) \text{ a.s.}$$

As will be shown in the sequel, the rate of convergence in (4) is mainly determined by the limiting behaviour of $\exp(-T_n)$ which follows from the result below.

LEMMA 2. For $n \to \infty$, we have

$$(5)\quad \log\left|\frac{R_n}{Y_n} - \frac{1}{2}\right| = O(\log n) \text{ a.s.}$$

Proof. We use a simple Borel-Cantelli argument. In order to prove (5), it suffices to show that in either case

$$(6)\quad \sum_{n=1}^{\infty} P(Y_n \leq n^{-2}) < \infty$$

$$(7)\quad \sum_{n=1}^{\infty} P(a_n \leq R_n/Y_n \leq b_n) < \infty \text{ where } a_n = \tfrac{1}{2} - n^{-2},\ b_n = \tfrac{1}{2} + n^{-2}.$$

While the first convergence is obvious, note that for (7) we have a.s. for $t > 0$

$$(8)\quad P(a_n \leq R_n/Y_n \leq b_n \mid T_n = t) \leq \sum_{k=0}^{\infty} P\left(k+1-b_nY_n \leq \frac{Y_n}{-\log(1-e^{-t})} \leq k+1-a_nY_n\right)$$

$$= \frac{1}{1-\exp\left(-\dfrac{\log(1-e^{-t})}{1-b_n\log(1-e^{-t})}\right)} - \frac{1}{1-\exp\left(-\dfrac{\log(1-e^{-t})}{1-a_n\log(1-e^{-t})}\right)} \leq b_n - a_n.$$

We are now ready to formulate the first main result.

THEOREM 1. For $n \to \infty$, we have

$$(9)\quad \log\Delta_n = \log Y_n + T_n + o(\exp(-n + nH(1/n))) \quad \text{a.s.}$$

where $t\,H(1/t)$ belongs to the upper class of a Wiener process, i.e. $H(t)$ is a positive function defined in some positive neighbourhood of the origin such that $H(t)\uparrow$ and $t^{-1/2}H(t)\downarrow$, and the integral

$$(10)\quad I = \int_{0+}^{\infty} t^{-3/2} H(t) \exp(-H^2(t)/2t)\, dt$$

converges. The above result cannot be extended to lower class functions (i.e. H as above with I being divergent), not even when $o(.)$ is replaced by $O(.)$.

Proof. Just as in Theorem 4 in Deheuvels [2], using Lemma 2 above and the Komlós-Major-Tusnády Theorem [4].

REMARK. A typical choice for upper and lower class functions is

(11) $$n\,H(1/n) = \sqrt{2n\{\log\log n + \tfrac{3}{2}\log^{(3)} n + \log^{(4)} n + \ldots + \log^{(p)} n + (1+\varepsilon)\log^{(p+1)} n\}}$$

for $\varepsilon \in \mathbb{R}$, $p \geq 4$, giving upper class functions for $\varepsilon > 0$ and lower class functions for $\varepsilon \leq 0$. In fact, this choice is closely related with Erdős' form of the LIL [3].

We shall now turn to a corresponding analysis for the record time sequence. Following Williams [7] and Westcott [6], this sequence can be represented by

(12) $$U_0 = 1,\quad U_{n+1} = \mathrm{int}\{U_n \exp Y_{n+1}\} + 1 = U_n \exp(Y_{n+1} + Z_{n+1}) \quad (n \geq 0),$$

say where again $\{Y_n;\ n \in \mathbb{N}\}$ is an i.i.d. sequence of exponentially distributed random variables with unit mean, and

(13) $$0 \leq Z_{n+1} \leq \log\left(1 + \frac{1}{U_n}\exp(-Y_{n+1})\right) \leq \exp(-T_{n+1}),\quad n \geq 0$$

where $T_n = \sum_{k=1}^{n} Y_k$, $n \geq 1$ again defines the arrival-time sequence of a unit-rate Poisson process, and $\sum_{k=1}^{n} Z_k$ converges monotonically to some integrable random variable Z with mean $E(Z) = 1 - C$ as can be concluded from Pfeifer [5]. This gives rise to the following estimations.

THEOREM 2. For $n \to \infty$, we have

(14) $$\log U_n = Z + T_n + o(\exp(-n + nH(1/n)))\quad \text{a.s.}$$

where again $t\,H(1/t)$ is an upper class function of a Wiener process.

<u>Proof.</u> Let $W_n = \sum_{k=n+1}^{\infty} Z_k$. Then $E(W_n) \leq \sum_{k=n+1}^{\infty} E(\exp(-T_n)) = \sum_{k=n+1}^{\infty} 2^{-k} = 2^{-n}$, giving $\sum_{n=1}^{\infty} P(W_n > n^2 2^{-n}) \leq \sum_{n=1}^{\infty} n^{-2} 2^n E(W_n) < \infty$. By the Borel-Cantelli Lemma we therefore have $W_n = O(n^2 2^{-n})$ a.s. for $n \to \infty$. But also, $W_n \leq \sum_{k=n+1}^{2n} \exp(-T_k) + W_{2n} \leq n\exp(-T_n) + W_{2n}$ $= \exp(\log n - T_n) + o(e^{-n})$ a.s. for $n \to \infty$. The result now follows as in the proof of Theorem 1.

We strongly believe that the rate result in Theorem 2 cannot be improved to the case of lower class functions either; for a proof of this, however, a better lower bound for Z_n as in (13) would be necessary.

References.

[1] Deheuvels, P.: Strong approximation in extreme value theory and applications. Coll. Math. Soc. Janos Bolyai 36, Limit Theorems in Probability and Statistics, Veszprém (Hungary). North Holland, Amsterdam (1982).

[2] Deheuvels, P.: The complete characterization of the upper and lower class of the record and inter-record times of an i.i.d. sequence. Z. Wahrscheinlichkeitsth. verw. Geb. 62, 1 - 6 (1983).

[3] Erdös, P.: On the law of the iterated logarithm. Annals of Math. 43, 419 - 436 (1942).

[4] Komlós, J., Major, P. and Tusnády, G.: An approximation of partial sums of independent rv's, and the sample df. II. Z. Wahrscheinlichkeitsth. verw. Geb. 34, 33 - 58 (1976).

[5] Pfeifer, D.: A note on moments of certain record statistics. Z. Wahrscheinlichkeitsth. verw. Geb. 66, 293 - 296 (1984).

[6] Westcott, M.: A note on record times. J. Appl. Prob. 14, 637 - 639 (1977).

[7] Williams, D.: On Rényi's 'record' problem and Engel's series. Bull. London Math. Soc. 5, 235 - 237 (1973).

Dietmar Pfeifer, Institut für Statistik und Wirtschaftsmathematik, RWTH Aachen, Wüllnerstr. 3, D-5100 Aachen, West-Germany