The Annals of Probability
1979, Vol. 7, No. 6, 1092–1096

EXTENSION OF THE DARLING AND ERDÖS THEOREM ON THE MAXIMUM OF NORMALIZED SUMS[1]

BY GALEN R. SHORACK

Australian National University and C.S.I.R.O.

The limiting distribution of $\max_{1 \leqslant k \leqslant n} S_k / k^{\frac{1}{2}}$ is derived via embedding. The theorem also applies to partial sums of certain dependent rv's. Thus the proof of the Darling and Erdös result is brought into line with recent literature; moreover, the scope of its applicability is greatly increased in regard both to dependence and moment assumptions.

1. Normalized Brownian motion. Let $\{\mathbb{S}(t) : t \geqslant 0\}$ denote Brownian motion. Let

$$m(t) \equiv \sup_{1 \leqslant s \leqslant t} \frac{\mathbb{S}(s)}{s^{\frac{1}{2}}} \quad \text{and} \quad M(t) \equiv \sup_{1 \leqslant s \leqslant t} \frac{|\mathbb{S}(s)|}{s^{\frac{1}{2}}}. \tag{1.1}$$

We define normalizing functions b and c by

$$b(t) \equiv (2 \log_2 t)^{\frac{1}{2}} \tag{1.2}$$

and

$$c(t) \equiv 2 \log_2 t + 2^{-1} \log_3 t - 2^{-1} \log(4\pi) \tag{1.3}$$

for $t > e^e$. Let E_v denote the extreme value df defined by

$$E_v(t) = \exp(-\exp(-t)) \quad \text{for} \quad -\infty < t < \infty. \tag{1.4}$$

Note that the kth power E_v^k of E_v is the df of the maximum of k independent rv's having df E_v. Darling and Erdös (1956) prove that

$$b(t)m(t) - c(t) \rightarrow_d E_v \quad \text{as} \quad t \rightarrow \infty \tag{1.5}$$

and

$$b(t)M(t) - c(t) \rightarrow_d E_v^2 \quad \text{as} \quad t \rightarrow \infty. \tag{1.6}$$

If one makes the transformation $X(t) = e^{-t}\mathbb{S}(e^{2t})$ for $t \geqslant 0$, then X is the Uhlenbeck process and

$$m(t) = \sup_{0 \leqslant s \leqslant 2^{-1}\log t} X(s) \quad \text{and} \quad M(t) = \sup_{0 \leqslant s \leqslant 2^{-1}\log t} |X(s)|. \tag{1.7}$$

It was (1.7) that figured in the Darling and Erdös proof of (1.5); see their Lemmas 3.4 and 3.10 for (1.5). They did not prove (1.6), but it is implicit in their Theorem 2.

Received August 21, 1978; revised November 22, 1978.

[1]This work was done while the author was on leave from the University of Washington and supported by NSF grants INT 77-08720 and MPS 75-08557

AMS 1970 *subject classifications.* Primary 60G50, 60B10; secondary 60F05.

Key words and phrases. Maximum of normalized sums, normalized Brownian motion, Uhlenbeck process, extreme value distribution, embedding, dependent rv's.

2. The normalized partial sum process. Let $X, X_1, X_2, \cdots$ be i.i.d. rv's with mean 0 and variance 1. Let $S_n \equiv X_1 + \cdots + X_n$ for $n \geqslant 1$ with $S_0 \equiv 0$. Define the random function S on $[0, \infty)$ by letting $S(t) = S_{\text{int}(t)}$ for $t \geqslant 0$, where int( ) denotes the greatest integer function. If $E|X|^{2+\theta} < \infty$ for some $\theta > 0$ then Major (1976, equation (2′)) shows that $X_1, X_2, \cdots$ and $\mathbb{S}$ can be defined so that

$$(2.1)\quad D(t) \equiv |S(t) - \mathbb{S}(t)|/t^{\frac{1}{2}} = 0(t^{-\delta}) \text{ a.s. as } t \to \infty \qquad \text{for some} \quad \delta > 0.$$

Let $X, X_1, X_2, \cdots$ be i.i.d. with mean 0 and variance 1. Let

$$(2.2)\qquad Y_{n,\varepsilon} \equiv \sup_{1 \leqslant s \leqslant (\log n)^\varepsilon} \frac{|S(s)|}{s^{\frac{1}{2}}} \qquad \text{and} \quad Z_{n,\varepsilon} \equiv \sup_{(\log n)^\varepsilon < s < n} \frac{|S(s)|}{s^{\frac{1}{2}}}.$$

We shall now show that for all big $K > 0$ and small $\varepsilon > 0$

$$(2.3)\quad P(b(n)Y_{n,\varepsilon} - c(n) \geqslant -K) \leqslant g(\varepsilon) \qquad \text{where} \quad g(\varepsilon) \to 0 \qquad \text{as } \varepsilon \to 0$$

provided n exceeds some N_K. The proof: by Lemma 1 of Shorack and Smythe (1976)

$$(2.4)\qquad Y_{n,\varepsilon} = \max_{1 \leqslant k \leqslant (\log n)^\varepsilon} \frac{|S_k|}{k^{\frac{1}{2}}} \leqslant 2 \max_{1 \leqslant k \leqslant (\log n)^\varepsilon} \left| \Sigma_{j=1}^k \frac{X_j}{j^{\frac{1}{2}}} \right| \equiv 2\overline{Y}_{n,\varepsilon},$$

(and note that this lemma is also valid for all dependent rv's). Thus (2.4) and Kolmogorov's inequality show that for all n exceeding some N_K

$$\begin{aligned} P(b(n)Y_{n,\varepsilon} - c(n) \geqslant -K) &= P(Y_{n,\varepsilon} \geqslant (c(n) - K)/b(n)) \\ &\leqslant P(Y_{n,\varepsilon} \geqslant b(n)) \leqslant P\big(\overline{Y}_{n,\varepsilon} \geqslant b(n)/2\big) \\ &\leqslant 4\Sigma_{j=1}^{(\log n)^\varepsilon} j^{-1}/b_n^2 \\ &\leqslant 4\varepsilon. \end{aligned}$$

Thus (2.3) holds. Only finite variance was needed in this paragraph.

Now Philipp and Stout (1975) contains many cases where a partial sum type process S on $[0, \infty)$ satisfies (2.1). Also, (2.3) will hold in many dependent situations, since the only really critical step in its proof is a version of Kolmogorov's inequality.

THEOREM 1. *Let S denote any process on* $[0, \infty)$ *that satisfies* (2.1) *and* (2.3). *Then*

$$(2.5)\qquad m_n \equiv \sup_{1 \leqslant s \leqslant n} \frac{S(s)}{s^{\frac{1}{2}}} \text{ and } M_n \equiv \sup_{1 \leqslant s \leqslant n} \frac{|S(s)|}{s^{\frac{1}{2}}}$$

satisfy

$$(2.6)\qquad b(n)m_n - c(n) \to_d E_v \qquad \text{as} \quad n \to \infty$$

and

$$(2.7)\qquad b(n)M_n - c(n) \to_d E_v^2 \qquad \text{as} \quad n \to \infty.$$

Darling and Erdös prove (2.6) and (2.7) for the case of independent rv's with mean 0, variance 1 and uniformly bounded third moments; condition (2.3) is easy and (2.1) follows from Theorem 7.1 of Phillip and Stout (1975) in that situation.

Darling and Erdös prove their theorem by an invariance principle, but it does not involve embedding. The present paper brings their result into line with recent literature both in terms of methodology and of a greatly increased scope of applicability.

An interesting related result involving the normalized empirical process and normalized Brownian bridge is found in Jaeschke (1976).

3. Proof. Define $Y'_{n,\varepsilon}$ and $Z'_{n,\varepsilon}$ by replacing S by $\mathbb{S}$ in (2.2). A passage to the limit in our proof of (2.3) shows that (2.3) also holds for $Y'_{n,\varepsilon}$. Note that (2.3) says that $b(n)Y_{n,\varepsilon} - c(n)$ and $b(n)Y'_{n,\varepsilon} - c(n)$ are arbitrarily close to $-\infty$ with probability arbitrarily close to 1. Thus

$$b(n)M_n - c(n) = b(n)[Y_{n,\varepsilon} \vee Z_{n,\varepsilon}] - c(n) \tag{3.1}$$

and

$$b(n)M(n) - c(n) = b(n)[Y'_{n,\varepsilon} \vee Z'_{n,\varepsilon}] - c(n) \tag{3.2}$$

will be controlled asymptotically by the $Z_{n,\varepsilon}$ and $Z'_{n,\varepsilon}$ terms respectively.

However, (2.1) implies that for a.e. ω we have $D(t) \leqslant t^{-\delta}$ for all t exceeding some T_ω. Thus $D(t) \leqslant (\text{some } K_\omega)t^{-\delta}$ for all $t \geqslant 1$. Thus

$$b(n)|Z_{n,\varepsilon} - Z'_{n,\varepsilon}| \leqslant b(n)K_\omega[(\log n)^\varepsilon]^{-\delta} \to 0 \text{ a.s.} \tag{3.3}$$

Thus the asymptotic behavior of $b(n)Z_{n,\varepsilon} - c(n)$ will be the same as that of $b(n)Z'_{n,\varepsilon} - c(n)$.

The consequence of (2.3) for $Y'_{n,\varepsilon}$ together with (3.2) and (1.6) show that $b(n)Z'_{n,\varepsilon} - c(n) \to_d E_v^2$. Thus (3.3) gives $b(n)Z_{n,\varepsilon} - c(n) \to_d E_v^2$. The consequences of (2.3) for $Y_{n,\varepsilon}$ together with (3.1) thus show that (2.7) holds. The proof of (2.6) is analogous.

Note that (2.1) does not require t^δ; all we require is enough for the conclusion (3.3). Jain, Jogdeo and Stout (1975, Theorem 4.2, say) does not help us relax (2.1), but it certainly suggests it is possible to get by with less than a $2 + \theta$ moment.

4. Examples. Now (2.1) was shown by Philipp and Stout (1975) to hold for certain Gaussian, lacunary trigonometric and stationary ϕ-mixing sequences. Thus, in these cases, we need only show (see (2.4) and the display following it) that

$$P(T_n \geqslant b(n)) \leqslant g(\varepsilon) \quad \text{where} \quad g(\varepsilon) \to 0 \text{ as } \varepsilon \to 0 \tag{4.1}$$

with

$$T_n \equiv T_{n,\varepsilon} \equiv 2 \max_{1 \leqslant k \leqslant (\log n)^\varepsilon} |\Sigma_{j=1}^k j^{-\frac{1}{2}} X_j|. \tag{4.2}$$

We now state the hypotheses for these three cases precisely; since we've already observed that (2.1) has already been shown in [5], we need only establish (2.3) or

(4.1). (After seeing a preprint of this paper, Philipp and Stout communicated these examples to me. The following quotation is essentially verbatim; I thank them for their very generous permission).

EXAMPLE 1. Let $\{X_n : n \geqslant 1\}$ be a stationary Gaussian sequence with 0 means and $EX_1X_n \ll n^{-2}$. Then (2.6) and (2.7) hold.

PROOF. We may use $g(\varepsilon) = \exp(-K/\varepsilon)$ for some $K > 0$ in (4.1) since Theorem 2.5 of Marcus and Shepp (1972) shows that

$$P(T_n \geqslant b(n)) \leqslant \exp(-(1-a)b^2(n)/4\sigma^2) \qquad \text{for all} \quad a > 0$$

where

$$\sigma^2 = \max_{1 \leqslant k \leqslant (\log n)^\varepsilon} E\Big(\Sigma_{j=1}^k j^{-\frac{1}{2}} X_i\Big)^2$$

$$\ll \Sigma_{j=1}^{(\log n)^\varepsilon}\left[\frac{1}{j} + \frac{1}{j}\Sigma_{j'=j+1}^{(\log n)^\varepsilon}(j-j')^{-2}\right] \ll \varepsilon \log_2 n. \qquad \square$$

EXAMPLE 2. Let $\{X_n : n \geqslant 1\}$ be a lacunary cosine series as in (3.1.1) of [5]. Then (2.6) and (2.7) hold.

PROOF. Apply Lemma 6.2.2 of [5] with $a_\nu = \nu^{-\frac{1}{2}}$. Then $A_k = (\Sigma_{\nu \leqslant k} a_\nu^2)^{\frac{1}{2}}/2 \ll (\log k)^{\frac{1}{2}}$. Hence (6.1.2) holds. Set $\lambda = b(n)/2$. By Lemma 6.2.2 we can choose $g(\varepsilon) \ll \lambda^{-4} A^4_{(\log n)^\varepsilon} \ll (\varepsilon \log_2 n)^2/b^4(n) \ll \varepsilon^2$ in (4.1). $\square$

EXAMPLE 3. Let $\{X_n : n \geqslant 1\}$ be a stationary ϕ-mixing sequence with 0 means and $E|X_1|^{2+\delta} < \infty$ for some $\delta > 0$. Suppose $\Sigma_{n=1}^\infty \phi^{\frac{1}{2}}(n) < \infty$, $\phi(n)\downarrow 0$ and the σ^2 of (4.1.3) of [5] is positive. Then (2.6) and (2.7) hold.

PROOF. See Lemma 1.1.6 of Iosifescu and Theodorescu (1969). $\square$

Though (2.1) holds in the strong mixing case of [5], the necessary maximal inequality is unavailable.

Note the acknowledgement of Philipp and Stout earlier in this section.

REFERENCES

[1] DARLING, D. and ERDÖS, P. (1956). A limit theorem for the maximum of normalized sums of independent random variable. *Duke Math. J.* **23** 143–155.

[1A] IOSIFESCU, M. and THEODORESCU, R. (1969). *Random Processes and Learning*. Springer-Verlag, New York.

[2] JAESCHKE, D. (1976). The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals. Univ. Dortmund technical report.

[3] JAIN, N., JOGDEO, K. and STOUT, W. (1975). Upper and lower functions for martingales and mixing processes. *Ann. Probability* **3** 119–145.

[4] MAJOR, P. (1976). The approximation of partial sums of independent rv's. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **35** 213–220.

[4a] MARCUS, M. and SHEPP, L. (1972). Sample behavior of Gaussian processes. *Proc. Sixth Berkeley Symp. Math. Statist. Probability* Univ. California Press, Berkeley.

[5] PHILIPP, W. and STOUT, W. (1975). Almost sure invariance principles for sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* **161**.

[6] SHORACK, G. and SYMTHE, R. (1976). Inequalities for max $|S_K|/b_k$ where $k \in N^r$. *Proc. Amer. Math. Soc.* **54** 331–336.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98105