UNIFORM LIMIT THEOREMS FOR THE MAXIMUM CUMULATIVE SUM IN PROBABILITY(1)

BY
S. SAWYER

1. **Introduction.** Let $\{x_k : k=1, 2, 3, \ldots\}$ be a sequence of independent random variables with mean zero, variance one, and for example bounded absolute third moments. Set $s_n^+ = \max_{1 \le k \le n} s_k$ for $s_k = x_1 + x_2 + \cdots + x_k$ and choose $\lambda \geqq 0$. Then, it was shown by Erdös and Kac [5] in 1946 that

$$\lim_{N\to\infty} P(s_N^+ < \lambda(N)^{1/2}) = M(\lambda) = (2/\pi)^{1/2} \int_0^\lambda e^{-u^2/2}\, du,$$

where $M(\lambda)=0$ for $\lambda<0$. The method was an ingenious use of the multi-dimensional central limit theorem; Kai Lai Chung [3] estimated the errors involved to calculate

$$\sup_\lambda |P(s_N^+ < \lambda(N)^{1/2}) - M(\lambda)| = O\left(\frac{\log N}{N^{1/26}}\right). \tag{1.1}$$

In general, if $\{x_k : k=1, 2, 3, \ldots\}$ is an arbitrary sequence of independent random variables with means zero, let $q_N^2 = E(x_1^2)+E(x_2^2)+\cdots+E(x_N^2)$ and $\{x_t : 0 \leqq t \leqq 1\}$ be Brownian motion. In §2 we show

THEOREM 1. *For all $p>2$ and all N,*

$$\sup_\lambda |P(s_N^+ < \lambda q_N) - M(\lambda)| \leqq 300\left(\frac{1}{q_N^p}\sum_1^N E(|x_k|^p)\right)^{1/p+1}. \tag{1.2}$$

THEOREM 2. *For all $p>2$ and all N,*

$$\begin{aligned} \sup_\alpha \sup_\beta \Big| P(s_N^+ \leqq \alpha q_N, s_N \leqq \beta q_N) - P\Big(\max_{0\le t\le 1} x_t \leqq \alpha, x_1 \leqq \beta\Big)\Big| \\ \leqq 800\left(\frac{1}{q_N^p}\sum_1^N E(|x_k|^p)\right)^{(2/3)/[p+(2/3)]}. \end{aligned} \tag{1.3}$$

For example, if $E(x_k^2)\equiv 1$ and $E(|x_k|^p)\leqq C$,

$$\sup_\lambda |P(s_N^+ < \lambda(N)^{1/2}) - M(\lambda)| = O\left(\frac{1}{N^{(p-2)/(2p+2)}}\right) \tag{1.4}$$

or $O(1/n^{1/8})$ under the hypotheses of (1.1). Similarly, if $N=O(q_N^2)$ and

$$\sup_k P(|x_k| > \lambda) \leqq Ce^{-\beta\lambda^\epsilon}$$

for $\beta>0$, $\varepsilon>0$, then by arguing as in [11]

$$\sup_\lambda |P(s_N^+ < \lambda q_N) - M(\lambda)| = O\left(\frac{(\log N)^{1/\varepsilon}}{(N)^{1/2}}\right). \tag{1.5}$$

Received by the editors March 30, 1967.

(1) This paper represents results obtained at the Courant Institute of Mathematical Sciences, New York University, under a Sloan Foundation Post-Doctoral Grant.

The limit distribution in (1.3) is of course a result of the Invariance Principle ([4], [9]), and is essentially equivalent (see §2) to its special case

$$\sup_{\alpha\geqq 0}\sup_{\beta\geqq 0}|P(s_N^+ > \alpha q_N, s_N \leqq (\alpha-\beta)q_N)-\Phi^c(\alpha+\beta)| \leqq 1600\left(\frac{1}{q_N^p}\sum_1^N E(|x_k|^p)\right)^{(2/3)/[p+(2/3)]}$$

where $\Phi(\lambda)=(1/(2\pi)^{1/2})\int_{-\infty}^{\lambda} e^{-u^2/2}\,du$ and $\Phi^c(\lambda)=1-\Phi(\lambda)$. A uniform rate of convergence here was first obtained by Hirsch [7](2), with, relative to (1.3), restrictions on $\{x_k\}$, α, β, and a weaker rate of convergence. This was used to prove a theorem of iterated logarithm type. In particular, the limit theorem of Hirsch can be extended to classes of $\{x_k\}$ without third moments or essentially constant variances.

As another example, consider the random walk on the real line given by $s_0=0$, $s_N=x_1+x_2+\cdots+x_N$. The probability of moving initially to the left, and of remaining in the left half line for at least N time units, can be estimated by the right hand side of (1.2). In particular, if $|x_k|\leqq C$ a.s., this probability is bounded by $300\, C/q_N$. This last result seems to be due initially to Burkholder(3), by the use of martingale methods.

Added in proof. Burkholder's approach apparently works only if $x_k=\pm a_k$ a.s. for constants $\{a_k\}$, $|a_k|\leqq C$, and uses Theorem 8 in [2]. Theorem 1 would then give a generalization.

For $p=3$, Theorem 1 can be obtained within a logarithmic factor, with the same bound for (1.2) and (1.3), by the methods of [9, §5]. For q_N^2 essentially a constant times N, the estimate (1.2) (multiplied by $(\log N)^{1/2}$) follows by arguing as in [10] for $2<p<4$ or as in [11] for $5<p<\infty$. However none of these arguments are elementary; it is one of the purposes here to show that these results can be obtained from the Reflection Principle of Désirée André and Lévy.

Finally, it should be mentioned that no attempt was made to find the optimal constants in (1.2) and (1.3). If $|x_k|\leqq C$ and $\sum_1^N E(x_k^2)=1$, it follows from (1.2) or Lemma 3, §2 that

(1.6) $$\sup_{\lambda}|P(s_N^+ < \lambda)-M(\lambda)| \leqq 300\, C/(N)^{1/2}.$$

By [1] or e.g. $x_k=\pm N^{-1/2}$, this is the best possible rate of convergence. Whether or not (1.2), or in particular (1.5), is best possible seems to be an open question.

2. Proofs.

Lemma 1 (Centering). *Let $x_1, x_2, \ldots, x_N$ be independent random variables such that*

(2.1) $$E(x_k) = 0,\ E(|x_k|^3) \leqq \Omega E(x_k^2), \qquad 1 \leqq k \leqq N.$$

Then, if $s_N=x_1+x_2+\cdots+x_N$ and $\mu=50\Omega$,

(2.2) $$P(s_N \leqq -\mu) < \tfrac{1}{2}.$$

(2) See also [1] for theorems of Cramér type.

(3) Colloquium talk 1-9-67, discussing results in [2].

Proof. By the Berry-Esseen theorem [6, p. 521]

$$\left| P(s_N \leqq \lambda) - \Phi\left(\frac{\lambda}{q_N}\right) \right| \leqq \frac{33}{4} \frac{\Omega}{q_N}, \qquad \text{all } \lambda,$$

$$\begin{aligned} P(s_N \leqq -\mu) &\leqq \Phi\left(-\frac{\mu}{q_N}\right) + \frac{33}{4}\frac{\Omega}{q_N} \\ &= \frac{1}{2} + \frac{33}{4}\frac{\Omega}{q_N} - \frac{1}{(2\pi)^{1/2}} \int_0^{\mu/q_N} e^{-u^2/2}\, du. \end{aligned}$$

Now, if $0 \leqq q_N < \mu/(2)^{1/2}$, then $P(s_N \leqq -\mu) \leqq (1/\mu^2)q_N^2 < \frac{1}{2}$ by Tchebysheff's inequality. But if $q_N \geqq \mu/(2)^{1/2}$, by the above

$$P(s_N \leqq -\mu) \leqq \frac{1}{2} + \frac{\Omega}{q_N}\left(\frac{33}{4} - \frac{50}{(2\pi)^{1/2}} \int_0^1 e^{-u^2}\, du\right) < \frac{1}{2}.$$

LEMMA 2 (REFLECTION). *Let $x_1, x_2, \ldots, x_N$ be independent random variables, such that for a fixed $\mu \geqq 0$ and $1 \leqq k \leqq N$*

$$P(s_N - s_k \geqq \mu) \leqq \tfrac{1}{2}, \qquad p(s_N - s_k \leqq -\mu) \leqq \tfrac{1}{2}. \tag{2.3}$$

Then, for all $\varepsilon > 0$, $\lambda > 0$,

$$\begin{aligned} 2P(s_N > \lambda + \varepsilon + \mu) - \sum_1^N P(x_k > \varepsilon) &\leqq P(s_N^+ > \lambda) \\ &\leqq 2P(s_N > \lambda - \mu). \end{aligned} \tag{2.4}$$

Proof. This is standard; see e.g. [8, p. 247, p. 558] or argue as in Lemma 5.

LEMMA 3. *Let $x_1, x_2, \ldots, x_N$ be independent random variables such that*

$$E(x_k) = 0, \quad |x_k| \leqq C \text{ a.s.}, \quad \sum_1^N E(x_k^2) = 1 \tag{2.5}$$

for some constant C and $1 \leqq k \leqq N$. Then, for all λ,

$$|P(s_N^+ < \lambda) - M(\lambda)| < 60C. \tag{2.6}$$

Proof. Clearly (2.5) implies (2.1) for $\Omega = C$, and hence (2.3) for $\mu = 50C$. Also, $|P(s_N > \lambda) - \Phi^c(\lambda)| \leqq 33C/4$ by Berry-Esseen. Hence by (2.4) with $\varepsilon = C$, $\lambda \geqq 0$

$$2\Phi^c(\lambda + \varepsilon + \mu) - \tfrac{33}{2}C \leqq P(s_N^+ > \lambda) \leqq 2\Phi^c(\lambda - \mu) + \tfrac{33}{2}C,$$

$$|P(s_N^+ > \lambda) - 2\Phi^c(\lambda)| \leqq \tfrac{33}{2}C + (2/\pi)^{1/2}(\mu + C) < 60C.$$

Proof of Theorem 1. First, note that it is sufficient to assume $q_N = 1$ and $C = (\sum_1^N E(|x_k|^p))^{1/(p+1)} < \frac{1}{2}$. Define $y_k = x_k$ if $|x_k| \leqq C$; otherwise set $y_k = 0$. Then $q_c^2 = \sum_1^N \sigma^2(y_k) \geqq \frac{1}{4}$, and $|y_k - E(y_k)|/q_c \leqq 4C$. Applying Lemma 3 to the $\{y_k\}$, and in general arguing as in [11, §5], we arrive at (1.2).

REMARK. If one does not object to an extra condition of the form $\max_{1\le k\le n} E(|x_k|^3)/E(x_k^2) \leqq Cq_n^{1-\theta}$, then (1.2) can be obtained directly from Lemma 1 by setting

$$\varepsilon = \left(q_N \sum_1^N E(|x_k|^p)\right)^{1/(p+1)} \text{ in (2.4).}$$

LEMMA 4. *Let $x_1, x_2, \dots x_N$ be as in (2.5), and for some fixed $\lambda > 0$ set $\tau = \min\{k : s_k > \lambda\}$ if $s_N^+ > \lambda$ and set $\tau = \infty$ if $s_N^+ \leqq \lambda$. Assume $1 \leqq k \leqq N$ and $\frac{1}{2} \leqq q_k \leqq q_N = 1$. Then*

$$P(k < \tau \leqq N) \leqq 1 - q_k + 200C. \tag{2.7}$$

Proof. We conclude (2.3) for $\mu = 50C$ as before, and by (2.4)

$$\begin{aligned} P(k < \tau \leqq N) &= P(s_N^+ > \lambda) - P(s_k^+ > \lambda) \\ &\leqq 2\Phi^c(\lambda - \mu) - 2\Phi^c\left(\frac{\lambda+\mu+C}{q_k}\right) + \frac{33}{2} C\left(1 + \frac{1}{q_k}\right) \\ &\leqq (2/\pi e)^{1/2} \frac{1-q_k}{q_k} + 200C, \text{ etc.} \end{aligned}$$

LEMMA 5. *For $x_1, x_2, \dots, x_N$ as in Lemma 4,*

$$\sup_{\alpha\ge 0}\sup_{\beta\ge 0} |P(s_N^+ > \alpha, s_N \leqq \alpha - \beta) - \Phi^c(\alpha+\beta)| \geqq 200C^{2/3}. \tag{2.8}$$

Proof. First, by the Berry-Esseen theorem, for all $\beta \geqq 0$, $\alpha \geqq 0$,

$$P(s_N - s_k \leqq -\beta) \leqq P(s_N - s_k \geqq \beta) + \frac{33}{2} \frac{C}{(1-q_k^2)^{1/2}},$$

$$\begin{aligned} P(s_N^+ > \alpha, s_N \leqq \alpha - \beta) &\leqq \sum_1^{N-1} P(s_1 \leqq \alpha, \dots, s_{k-1} \leqq \alpha, s_k > \alpha) P(s_N - s_k \leqq -\beta) \\ &\leqq \sum_1^{N-1} P(s_1 \leqq \alpha, \dots, s_{k-1} \leqq \alpha, s_k > \alpha) P(s_N - s_k \geqq \beta) + Q \\ &\leqq P(s_N > \alpha + \beta) + Q. \end{aligned}$$

Here

$$Q = \frac{33}{2} \frac{C}{(1-q_l^2)^{1/2}} + P(l < \tau \leqq N) \leqq \frac{33}{2} \frac{C}{(1-q_l)^{1/2}} + 1 - q_l + 200C$$

for some l to be determined, $q_l \geqq \frac{1}{2}$. Similarly, since $|x_k| \leqq C$,

$$\begin{aligned} P(s_N^+ > \alpha, s_N \leqq \alpha - \beta) &\geqq \sum_1^{N-1} P(s_1 \leqq \alpha, \dots, s_{k-1} \leqq \alpha, s_k > \alpha) P(s_N - s_k \leqq -\beta - C) \\ &\geqq \sum_1^{N-1} P(s_1 \leqq \alpha, \dots, s_{k-1} \leqq \alpha, s_k > \alpha) P(s_N - s_k \geqq \beta + C) - Q \\ &\geqq P(s_N > \alpha + \beta + 2C) - Q, \end{aligned}$$

and

$$|P(s_N^+ > \alpha, s_N \leqq \alpha - \beta) - \Phi^c(\alpha+\beta)| \leqq \frac{33}{4} C + \frac{2}{(2\pi)^{1/2}} C + Q.$$

Note that unless $C<1/8$, (2.8) is trivial. If now $k=\max\{j : q_j<1-C^{2/3}\}$, then $C^{2/3}<1-q_k<C^{2/3}+C^2\leqq 2C^{2/3}<\frac{1}{2}$. Thus $Q<180C^{2/3}$ and (2.8).

Proof of Theorem 2. I claim that under the hypotheses of Lemma 5, for all α and all β

$$\left|P(s_N^+ \leqq \alpha, s_N \leqq \beta)-P\left(\max_{0\leq t\leq 1} x_t \leqq \alpha, x_1 \leqq \beta\right)\right| \leqq 210C^{2/3}. \tag{2.9}$$

Clearly we can assume $C<1$. Now, if $\alpha<0$, then $P(s_N^+\leqq\alpha, s_N\leqq\beta)\leqq P(s_N^+\leqq 0) <60C<210C^{2/3}$, while $P(\max_{0\leq t\leq 1} x_t\leqq 0)=0$. Also if $0\leqq\alpha\leqq\beta$,

$$P(s_N^+ \leqq \alpha, s_N \leqq \beta) = P(s_N^+ \leqq \alpha) \leqq M(\alpha)+60C,$$

and since $P(\max_{0\leq t\leq 1} x_t\leqq\alpha)=M(\alpha)$ [6, p. 171], (2.9) follows from Lemma 3. Finally, if $\alpha\geqq 0$, $\beta\leqq\alpha$, then

$$P(s_N^+ \leqq \alpha, s_N \leqq \beta) = P(s_N \leqq \beta)-P(s_N^+ > \alpha, s_N \leqq \beta),$$

and (2.9) follows from (2.8), the Berry-Esseen theorem, and the reflection principle for Brownian motion. Finally, given (2.9) for the hypotheses (2.5), Theorem 2 follows by a truncation argument as in Theorem 1 or [11, §5].

References

1. A. Borovkoff, *New limit theorems for boundary problems for sums of independent variables*, Siberian Math. J. **3** (1962), 645–695. (Russian)

2. D. L. Burkholder, *Martingale transforms*, Ann. Math. Statist. **37** (1966), 1494–1504.

3. K. L. Chung, *Asymptotic distribution of the maximum cumulative sum of independent random variables*, Bull. Amer. Math. Soc. **54** (1948), 1162–1170.

4. M. Donsker, *An invariance theorem for certain probability limit theorems*, Mem. Amer. Math. Soc. **6** (1951).

5. P. Erdös and M. Kac, *On certain limit theorems of the theory of probability*, Bull. Amer. Math. Soc. **52** (1946), 292–302.

6. W. Feller, *An introduction to probability theory and its applications*, Volume II, Wiley, New York, 1966.

7. W. Hirsch, *A strong law for the maximum cumulative sum of independent random variables*, Comm. Pure Appl. Math. **18** (1965), 109–127.

8. M. Loève, *Probability theory*, 2nd ed., Van Nostrand, Princeton, N. J., 1960.

9. Y. Prokhoroff, *Convergence of random processes and limit theorems in probability*, Theor. Probability Appl. **1** (1956), 157–214.

10. W. Rosenkrantz, *On rates of convergence in the invariance principle*, Trans. Amer. Math. Soc. **129** (1967), 542–552.

11. S. Sawyer, *A uniform rate of convergence for the maximum absolute value of partial sums in probability*, Comm. Pure Appl. Math. **20** (1967), 647–658.

Courant Institute, New York University
New York, New York