# On the Asymptotic Distribution of the Maximum of Sums of a Random Number of I.I.D. Random Variables

by

PREM S. PURI*

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 265

September, 1971

*This investigation was supported in part by research grant GM-10525 from NIH, Public Health Service, at the University of California, Berkeley.

On the Asymptotic Distribution of the Maximum
of Sums of a Random Number of I.I.D. Random Variables

by

PREM S. PURI*, PURDUE UNIVERSITY

1. INTRODUCTION. Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables taking on real values. Let for $n = 0,1,2,\ldots$, $S_n = \sum_{i=0}^{n} X_i$, where $S_0 = X_0 = 0$. We shall be concerned here with the random variables

(1) $$\eta_n = \max(0, S_1, S_2, \ldots, S_n),\ n = 0,1,2,\ldots \quad .$$

We assume that $E|X_n| < \infty$ and write $a = EX_n$. Let

(2) $$\eta = \lim_{n \to \infty} \eta_n = \sup_{0 \leq n < \infty} S_n \quad .$$

The random variable $\eta$ is nonnegative, but possibly improper. We shall call the process $\{\eta_n\}$ subcritical, critical and supercritical according as $a < 0$, equal to zero and $a > 0$, respectively. We shall assume that $P(X_n = 0) < 1$, for in the trivial case where $P(X_n = 0) = 1$, we have $P(\eta_n = 0) = 1$, for all $n$. We summarize in the following few known asymptotic results concerning $\eta_n$. The exact distribution of $\eta_n$ is of course covered by the celebrated Spitzer's identity [6].

(i) In the subcritical case, $P(\eta < \infty) = 1$, whereas in the remaining cases

*This investigation was supported in part by research grant GM-10525 from NIH, Public Health Service, at the University of California, Berkeley.

$P(\eta = \infty) = 1$. (See Takács [7]).

(ii) If $a = 0$ and $EX_n^2 = 1$, then

$$\lim_{n \to \infty} P(\eta_n/\sqrt{n} \leq x) = \begin{cases} 2\Phi(x) - 1 & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases} , \tag{3}$$

where $\Phi(x)$ is the standard normal distribution function. (See Erdős and Kac [3])

(iii). An asymptotic result for the supercritical case is given below in the form of a theorem. Although, this result is known (See Chung [2]), we provide here a much simpler proof.

THEOREM 1: *If* $a > 0$, *then* $\eta_n/n \overset{a.s.}{\to} a$, *as* $n \to \infty$. *If moreover*, $\text{Var } X_n = 1$, *then*

$$\lim_{n \to \infty} P\left(\frac{\eta_n - na}{\sqrt{n}} \leq x\right) = \Phi(x) \quad . \tag{4}$$

Proof. The almost sure (a.s.) convergence of $\eta_n/n$ follows from the strong law of large number and from the fact that for a sequence $\{b_n\}$ of real numbers

$$\lim_{n \to \infty} \left[\frac{1}{n} \sum_{i=1}^{n} b_i\right] = b \Rightarrow \lim_{n \to \infty} \frac{1}{n} \max(0,b_1,b_1+b_2,\ldots,b_1+\ldots+b_n) = \max(0,b). \tag{5}$$

Again since $X_n$'s are i.i.d., it follows that the distribution of $\eta_n$ is same as that of $\xi_n = \max(0,X_n,X_n+X_{n-1},\ldots, X_n+\ldots+X_1)$. Thus proving (4) for $\eta_n$ is equivalent to proving it for $\xi_n$. On the other hand

$$\xi_n = \max(0,S_n-S_{n-1},\ldots,S_n) = S_n + \max(0,-S_1,-S_2,\ldots,-S_n), \tag{6}$$

so that

$$\frac{\xi_n - na}{\sqrt{n}} = \frac{S_n - na}{\sqrt{n}} + \frac{\max(0,-S_1,-S_2,\ldots,-S_n)}{\sqrt{n}} \quad . \tag{7}$$

Now using the fact that $\max(0,-S_1,-S_2,\ldots,-S_n)$ corresponds to a subcritical process, (taking $-X_n$'s instead of $X_n$'s) it follows from (i) that $\max(0,-S_1,-S_2,\ldots,-S_n)$ tends in law to a proper random variables, so that the last term of (7) tends to zero in probability as $n \to \infty$. Hence the theorem follows from (7) by using the central limit theorem.

The aim of the present paper is to establish the above asymptotic results for $\eta_{\nu(n)}$ as $n \to \infty$ where $\nu(n)$ is a positive integer-valued random variable for $n \geq 1$, which converges in probability to $+\infty$ as $n \to \infty$. For the case, where it is assumed that for any $n \geq 1$, $\nu(n)$ is independent of the random variables $\eta_n$ $(n = 1,2,\ldots)$ the above results are easy to establish. However, in the present work we make no such assumption. A more general result in this direction was originally established by Anscombe [1] under a condition of uniform continuity in probability of the random variables involved. Rényi ([4],[5]) gave a simpler proof of Anscombe's theorem for the special case of simple sums of i.i.d. random variables and established a central limit theorem. While proving the above results here, in a sense, we shall be showing that Anscombe's result holds for the present case of $\eta_{\nu(n)}$ as well.

## 2. SUBCRITICAL CASE.

Here we assume that $a < 0$, and prove the following theorem.

**THEOREM 2:** *Let* $a < 0$. *Let* $\nu(n)$ *denote a positive integer valued random variable for every* $n = 1,2,\ldots,$ *such that as* $n \to \infty$, $\nu(n)/f(n)$ *converges in probability to a constant* $c > 0$, *for some sequence of positive numbers* $f(n)$ *with* $f(n) \to \infty$, *as* $n \to \infty$. *Then* $\eta_{\nu(n)} \xrightarrow{P} \eta$ *as* $n \to \infty$.

**Proof.** Since $\nu(n)/f(n) \xrightarrow{P} c > 0$, there exists a nonincreasing sequence $\varepsilon_n > 0$ with $\varepsilon_n \to 0$ as $n \to \infty$, such that

(8) $$P\left( \left|\frac{\nu(n)}{f(n)} - c\right| > c\,\varepsilon_n\right) \le \varepsilon_n,\ n = 1,2,\ldots \quad .$$

Let $I_{A(n)}$ denote the indicator function of the set

$$A(n) = \{c(1-\varepsilon_n)f(n) \le \nu(n) \le (1+\varepsilon_n)c\ f(n)\} \quad ,$$

and let $N_1(n) = [c(1-\varepsilon_n)f(n)]$ and $N_2(n) = [c(1+\varepsilon_n)f(n)]$, where $[\cdots]$ denotes the integral part of the number in the square bracket. Also we shall occasionally supress the arguments of $N_1$ and $N_2$ for convenience. Since $\eta_{N_1(n)} \overset{a.s.}{\to} \eta$, it suffices to show that $|\eta_{\nu(n)} - \eta_{N_1(n)}| \overset{P}{\to} 0$, as $n \to \infty$. However,

(9) $$|\eta_{\nu(n)} - \eta_{N_1(n)}| = (\eta_{\nu(n)} - \eta_{N_1(n)})I_{A(n)} + |\eta_{\nu(n)} - \eta_{N_1(n)}|\ I_{\tilde{A}(n)} \quad ,$$

where $\tilde{A}(n)$ denotes the complement of $A(n)$. Again since, $I_{\tilde{A}(n)} \overset{P}{\to} 0$, the last term of (9) tends to zero in probability. On the other hand

$$0 \le (\eta_{\nu(n)} - \eta_{N_1(n)})I_{A(n)} \le \sup_{N_1 < k \le N_2} (\eta_k - \eta_{N_1(n)}) = \eta_{N_2(n)} - \eta_{N_1(n)} \le \eta - \eta_{N_1(n)} \quad ,$$

and since the last quantity tends to zero in probability as $n \to \infty$, the theorem follows.

3. CRITICAL CASE. We prove here the following theorem, the analogue of (ii).

**THEOREM 3:** Let $a = 0$ and $EX_n^2$ exist. Also without loss of generality let $EX_n^2 = 1$. Let $\nu(n)$ be as defined in theorem 2. Then

(10) $$\lim_{n \to \infty} P\left(\frac{\eta_{\nu(n)}}{\sqrt{\nu(n)}} \le x\right) = \begin{cases} 2\Phi(x) - 1, & \text{for } x \ge 0 \\ 0 \quad , & \text{for } x < 0. \end{cases}$$

Proof. Consider the following identity

$$(11)\qquad \frac{\eta_{\nu(n)}}{\sqrt{\nu(n)}} \equiv \frac{\eta_{\nu(n)} - \eta_{N_1}}{\sqrt{N_1}}\; I_{A(n)}\,(N_1/\nu(n))^{\frac{1}{2}} + \frac{\eta_{N_1}}{\sqrt{N_1}}\; I_{A(n)}\,(N_1/\nu(n))^{\frac{1}{2}} + \frac{\eta_{\nu(n)}}{\sqrt{\nu(n)}}\; I_{\tilde{A}(n)} \quad .$$

Since $I_{\tilde{A}(n)} \overset{P}{\to} 0$, as $n \to \infty$, the last term of (11) tends to zero in probability. Again since $\nu(n)/N_1(n) \overset{P}{\to} 1$, as $n \to \infty$, it follows from (3) that the second term on the right side of (11) tends in law to that of (10). Thus, since

$$(12)\qquad \left|\frac{\eta_{\nu(n)} - \eta_{N_1}}{\sqrt{N_1}}\right| \cdot I_{A(n)} \le \frac{\eta_{N_2} - \eta_{N_1}}{\sqrt{N_1}} \quad ,$$

in order to complete the proof, it suffices to show that the right side of (12) tends in probability to zero. This we achieve as follows. For any arbitrary constant $b > 0$, we have

$$(13)\qquad \{\eta_{N_2} - \eta_{N_1} \ge b\sqrt{N_1}\} \subset \{\max(S_{N_1+1},\dots,S_{N_2}) \ge \eta_{N_1} + b\sqrt{N_1}\}$$

$$\subset \{\max_{N_1<k\le N_2} \sum_{i=N_1+1}^{k} X_i \ge \eta_{N_1} - S_{N_1} + b\sqrt{N_1}\}$$

$$\subset \{\max_{N_1<k\le N_2} \left|\sum_{i=N_1+1}^{k} X_i\right| \ge b\sqrt{N_1}\}$$

On the other hand, by Kolmogorov inequality

$$(14)\qquad P(\max_{N_1<k\le N_2} \left|\sum_{i=N_1+1}^{k} X_i\right| \ge b\sqrt{N_1}) \le \frac{N_2(n) - N_1(n)}{b^2 N_1(n)} \quad ,$$

and since the right side of (14) tends to zero as $n \to \infty$, we have $\lim_{n\to\infty} P(\eta_{N_2} - \eta_{N_1} \ge b\, N_1^{1/2}) = 0$. This completes the proof of theorem 3.

4. SUPERCRITICAL CASE. We need the following theorem in order to prove the main result of this section.

THEOREM 4. *Let* $a > 0$. *Then*

$$\lim_{n \to \infty} P(\eta_n - S_n \le x) = W(x) \quad , \tag{15}$$

*where* $W(x)$ *is a distribution function of a nonnegative proper random variable. Furthermore, this distribution is same as that of the limit of a subcritical process obtained by replacing* $X_n$ *by* $-X_n$, *for all* $n$.

The proof of this theorem is omitted as it follows along the lines of the proof of Theorem 1 and in particular from (6). Finally, we have the following theorem as the analogue of theorem 1.

THEOREM 5. *Let* $a > 0$ *and* $\mathrm{Var}\, X_n = 1$. *Let* $\nu(n)$ *be as defined in theorem* 2. *Then*

$$\lim_{n \to \infty} P\left(\frac{\eta_{\nu(n)} - a\cdot\nu(n)}{\sqrt{\nu(n)}} \le x\right) = \Phi(x) \quad . \tag{16}$$

PROOF. Consider the following identity.

$$\frac{\eta_{\nu(n)} - \nu(n)a}{\sqrt{\nu(n)}} \equiv \frac{\eta_{\nu(n)} - \eta_{N_1} - (\nu(n)-N_1)a}{\sqrt{N_1}} I_{A(n)} (N_1/\nu(n))^{1/2} + \frac{\eta_{N_1} - N_1 a}{\sqrt{N_1}} I_{A(n)} (N_1/\nu(n))^{1/2} + \frac{\eta_{\nu(n)} - \nu(n)a}{\sqrt{\nu(n)}} \cdot I_{\tilde{A}(n)} \quad . \tag{17}$$

As before since $I_{\tilde{A}(n)} \xrightarrow{P} 0$, the last term tends to zero in probability. Also, since $\nu(n)/N_1 \xrightarrow{P} 1$, it follows from theorem 1 that the second term on the right side of (17) tends in law to a standard normal random variable as $n \to \infty$. Thus to complete the proof, it suffices to prove that the sequence

$$\sup_{N_1 < k \leq N_2} |\eta_k - \eta_{N_1} - (k-N_1)a| \cdot N_1^{-1/2} \quad ,$$

tends in probability to zero, as $n \to \infty$. However, since

$$(18) \qquad \sup_{N_1 < k \leq N_2} |\eta_k - \eta_{N_1} - (k-N_1)a| \cdot N_1^{-1/2} \leq \sup_{N_1 < k \leq N_2} |\eta_k - S_{N_1} - (k-N_1)a| \cdot N_1^{-1/2} + |\eta_{N_1} - S_{N_1}| \cdot N_1^{-1/2} \quad ,$$

and the last term of this by virtue of theorem 4 tends to zero in probability as $n \to \infty$, it is sufficient to show that the first term on the right side of inequality (18) tends to zero in probability. Let $0 < \beta < \alpha$ be two arbitrary constants. Let $\delta > 0$ be another arbitrary constant. Then since $(\eta_{N_1} - S_{N_1})N_1^{-1/2} \xrightarrow{P} 0$, we can find an integer $n_0$ depending upon $\delta$ and $\beta$, such that for $n > n_0$,

$$(19) \qquad P(\eta_{N_1} - S_{N_1} > \beta N_1^{1/2}) \leq \delta \ .$$

Thus we have for $n > n_0$ ,

$$(20) \qquad P(\sup_{N_1<k\le N_2} |\eta_k - S_{N_1} - (k-N_1)a| N_1^{-1/2} \ge \alpha)$$

$$\le \delta + P(\sup_{N_1<k\le N_2} |\max(\eta_{N_1} - S_{N_1}, S_{N_1+1} - S_{N_1}, \ldots, S_k - S_{N_1}) - (k-N_1)a| \ge \alpha N_1^{1/2}, \text{ and } \eta_{N_1} - S_{N_1} \le \beta N_1^{1/2})$$

$$\le \delta + P(\sup_{N_1<k\le N_2} |\max(S_{N_1+1} - S_{N_1}, \ldots, S_k - S_{N_1}) - (k-N_1)a| \ge \alpha N_1^{1/2})$$

$$\le \delta + P(\sup_{N_1<k\le N_2} |(X_{N_1+1} + \ldots + X_k) - (k-N_1)a| \ge \alpha N_1^{1/2})$$

$$\le \delta + P(\sup_{N_1<k\le N_2} |\sum_{i=N_1+1}^{k} (X_i - a)| \ge \alpha N_1^{1/2})$$

$$\le \delta + \frac{N_2(n) - N_1(n)}{\alpha^2 N_1(n)} \quad .$$

Here for going from first inequality to the second, among others, we have used the fact that $\beta < \alpha$. The last step of (20), of course, follows from Kolmogorov inequality. Now since $(N_2(n) - N_1(n))/N_1(n)$ tends to zero as $n \to \infty$, and $\delta$ being arbitrary, the theorem follows.

We close with the remark that all the above results can easily be extended to cover the case where $S_0 = X_0$ is a nonnegative random variable. When $\nu(n)$ is nonrandom and is equal to $n$, this case has recently been considered by Takác [7].

ACKNOWLEDGEMENT. The author is grateful to Professor L. LeCam for some helpful discussions and in particular for drawing his attention to Rényi's work ([4],[5]).

# REFERENCES

[1] Anscombe, F. J. (1952). Large-sample theory of sequential estimation, Proc. Camb. Phil. Soc., 48, 600.

[2] Chung, K. L. (1948). Asymptotic distribution of the maximum cumulative sum of independent random variables, Bul. Am. Math. Soc., 54, 1162-70.

[3] Erdös, P. and Kac, M. (1946). On certain limit theorems of the theory of probability. Bull. Amer. Math. Soc., 52, 292-302.

[4] Rényi, A. (1957). On the asymptotic distribution of the sum of a random number of independent random variables, Acta. Math. Acad. Sc. Hung., 8, 193-199.

[5]R Rényi, A. (1960). On the central limit theorem for the sum of a random number of independent random variables, Acta. Math. Acad. Sc. Hung., 11, 97-102.

[6] Spitzer, F. (1956). A combinatorial lemma and its application to probability theory, Trans. Amer. Math. Soc., 82, 323-339.

[7] Takacs, L. (1970). On the distribution of the maximum of sums of mutually independent and identically distributed random variables, Adv. Appl. Prob., 2, 344-354.