The Annals of Probability
1984, Vol. 12, No. 1, 256–261

A NOTE ON THE BEHAVIOR OF SAMPLE STATISTICS WHEN THE POPULATION MEAN IS INFINITE

BY JEESEN CHEN[1] AND HERMAN RUBIN[2]

University of Cincinnati and Purdue University

Let $X_i \geq 0$ be i.i.d. random variables with $E(X_i) = \infty$. Then for suitable functions φ we have $\overline{\varphi(X)}/\varphi(\bar{X}) \to 0$ a.s. We give some applications of this result.

1. The main theorem.

Let us prove the following:

THEOREM 1. *If $X_1, X_2, \cdots$ are i.i.d., $X_i \geq 0$. $EX_1 = \infty$, and if φ is a function such that*

(1.1) *there exist constants A and B such that $a_i \geq B$, $i = 1, 2, \cdots n$, implies*

$$\sum_{i=1}^{n} \frac{\varphi(a_i)}{n} \leq A\varphi\left(\frac{\sum_{i=1}^{n} a_i}{n}\right),$$

(1.2) $$\varphi(x) \to \infty \quad as \quad x \to \infty,$$

(1.3) *there exist constants C, x_0 and α, $\alpha < 1$ such that $\varphi(\lambda x)/\varphi(x) \leq C\lambda^\alpha$ for $\lambda \geq 1$, $x \geq x_0$, and $\varphi(x)$ is bounded for $x \leq x_0$,*

then

$$R_n = \frac{(1/n)\sum_{i=1}^{n} \varphi(X_i)}{\varphi((1/n)\sum_{i=1}^{n} X_i)} \to_{\text{a.s.}} 0.$$

NOTE.

(a) Following the same argument of Mulholland [4], Theorem 1, we have: (1.1) is equivalent to

(1.4) there exist constants A and B and a concave function ψ, such that

$$\psi(x) \leq \varphi(x) \leq A\psi(x) \quad \text{for all} \quad x \geq B.$$

(b) Condition (1.3), according to the terminology of Bingham and Goldie [3], is:

(1.5) The upper Matuszewska index of φ is less than 1.

Received August 1982; revised May 1983.

[1]The research of this author was partially supported by the Taft Fellowship at the University of Cincinnati.

[2]The research of the second mentioned author was partially supported by Army Research Office Contract #DAAG 29-80-K-0043.

AMS 1980 *subject classifications.* Primary 60F99; secondary, 62G99, 62F99.

Key words and phrases. Sums of random variables, nonexistence of moments, moments.

For properties connected with (1.3), see Drasin and Shea [1] and Bingham and Goldie [2], [3].

PROOF. Let d be a positive number. Let p_n be the proportion of i's, $i \le n$, such that $X_i > d$. For n sufficiently large, $p_n > 0$. Let us assume this is the case. Then:

$$R_n = \frac{(1/n)\sum_{X_i \le d} \varphi(X_i) + (1/n)\sum_{X_i > d}\varphi(X_i)}{\varphi((1/n)\sum_{i=1}^n X_i)}$$

$$\le \frac{(1/n)\sum_{X_i \le d} K_d}{\varphi((1/n)\sum_{i=1}^n X_i)} + \frac{(1/j)\sum_{X_i>d}\varphi(X_i)}{\varphi((1/j)\sum_{X_i>d} X_i)} \frac{\varphi((1/j)\sum_{X_i>d} X_i)}{\varphi((1/n)\sum_{i=1}^n X_i)} \cdot \frac{j}{n} \tag{1.6}$$

$$= T_1 + T_2 \cdot T_3 \cdot \frac{j}{n}, \quad \text{say,}$$

where K_d comes from condition (1.3) since (1.3) implies φ is bounded in any finite interval.

Since $E(X_1) = \infty$, T_1 approaches 0 a.s. as $n \to \infty$ by condition (1.2) and the strong law of large numbers.

Let $j = np_n = \#\{i: X_i > d, i = 1, 2, \cdots, n\}$. Then condition (1.1) implies T_2 is bounded by A.

Since

$$(1/n)\sum_{X_i>d} X_i \le (1/n)\sum_{i=1}^n X_i \le (1/j)\sum_{X_i>d} X_i, \tag{1.7}$$

for n sufficiently large with probability 1,

$$1 \le \frac{(1/j)\sum_{X_i>d} X_i}{(1/n)\sum_{i=1}^n X_i} \le \frac{n}{j}. \tag{1.8}$$

Apply (1.3) and (1.8):

$$T_3 = \frac{\varphi((1/j)\sum_{X_i>d} X_i)}{\varphi((1/n)\sum_{i=1}^n X_i)} \le C\left(\frac{n}{j}\right)^{\alpha}. \tag{1.9}$$

Hence

$$T_2 \cdot T_3 \cdot (j/n) \le AC(j/n)^{1-\alpha}. \tag{1.10}$$

Notice that since $(j/n) \to P(X_i > d)$ a.s., it follows that if we choose d large enough then R_n is eventually less than any positive number with probability 1. □

2. Some applications. If $\varphi(x) = x^{\mu}L(x)$ where $0 < \mu < 1$ and $L(x)$ is a slowly varying function, i.e. $\lim_{x\to\infty}(L(\lambda x)/L(x)) = 1$ for all $\lambda > 0$, then $\varphi(x)$ satisfies (1.2), (1.3) and (1.4).

THEOREM 2. *If* $X_1, X_2, \cdots$ *are* i.i.d., $X_i \ge 0$, $EX_1 = \infty$, *and* $\varphi(x) = x^{\mu}L(x)$ *for* *some* $0 < \mu < 1$ *and slowly varying function* L *then*:

$$R_n = \frac{(1/n)\sum_{i=1}^n \varphi(X_i)}{\varphi((1/n)\sum_{i=1}^n X_i)} \to_{\text{a.s.}} 0. \tag{2.1}$$

An easy corollary of Theorem 2 is:

COROLLARY 3. *If* $X_1, X_2, \cdots$ *are* i.i.d., $X_i \geq 0$, $EX_1 = \infty$, *and* $0 < \mu < 1$, *then*

$$\frac{(1/n)\sum_{i=1}^n X_i^\mu}{((1/n)\sum_{i=1}^n X_i)^\mu} \to_{\text{a.s.}} 0. \tag{2.2}$$

COROLLARY 4. *If* $Y_1, Y_2, \cdots$ *are* i.i.d., $p > 1$, *and* $E|Y_1|^p = \infty$, *then*

$$\frac{((1/n)\sum_{i=1}^n Y_i)^p}{(1/n)\sum_{i=1}^n |Y_i|^p} \to_{\text{a.s.}} 0. \tag{2.3}$$

PROOF. Let $\varphi(x) = x^{1/p}$ and apply Theorem 2 to the i.i.d. random variables $|Y_1|^p, |Y_2|^p, \cdots, |Y_n|^p, \cdots$. We have

$$\frac{(1/n)\sum_{i=1}^n |Y_i|}{((1/n)\sum_{i=1}^n |Y_i|^p)^{1/p}} \to_{\text{a.s.}} 0. \quad \square \tag{2.4}$$

For the special case $p = 2$, it is easy to see from Corollary 4 that when the second moment of the population does not exist, the ratio of the sample mean to the sample standard deviation approaches 0 almost surely as the sample size increases.

It is possible to apply the theorem to compare the growth rates of some familiar statistics.

PROPOSITION 5. *Let* $X_1, X_2, \cdots X_n, \cdots$ *be* i.i.d. *random variables. If* $EX_1^2 = \infty$ *then*

$$\frac{\binom{n}{2}^{-1}\sum_{1\le i<j\le n} X_iX_j}{(1/n)\sum X_i^2} \to_{\text{a.s.}} 0. \tag{2.5}$$

PROOF. $(X_1 + X_2 + \cdots + X_n)^2 = \sum X_i^2 + 2\sum_{1\le i<j\le n} X_iX_j$, hence

$$\frac{\left(\dfrac{X_1 + X_2 + \cdots X_n}{n}\right)^2}{(1/n)\sum X_i^2} = \frac{1}{n} + \frac{2\binom{n}{2}\binom{n}{2}^{-1}\sum_{1\le i<j\le n} X_iX_j}{n^2(1/n)\sum X_i^2}.$$

Applying Corollary 4, we get (2.5). $\square$

Notice that $\binom{n}{2}^{-1}\sum_{1\le i<j\le n} X_iX_j$ is the U-statistic of the kernel $\phi(x_1, x_2) = x_1x_2$,

and $\binom{n}{2}^{-1} \sum X_i^2$ is the U-statistic of the kernel $\tilde{\phi}(x) = \phi(x, x) = x^2$. Following the same type of argument, we have the following theorem.

THEOREM 6. *If* $k > 1$,

$$\phi(x_1, x_2, \cdots, x_k) = x_1 x_2 \cdots x_k, \quad \tilde{\phi}(x) = \phi(x, x, \cdots, x),$$

$U_n(\phi)$, $U_n(\tilde{\phi})$ *are the U-statistics of the kernel function* ϕ *and* $\tilde{\phi}$ *respectively, and* $E\tilde{\phi}(X_1) = \infty$, *then*

$$\frac{U_n(\phi)}{U_n(\tilde{\phi})} \to_{\text{a.s.}} 0. \tag{2.6}$$

COROLLARY 7. *Let* $\phi(x_1, \cdots, x_k)$ *be a symmetric polynomial in* $x_1, \cdots, x_k$, *with all coefficients* ≥ 0, *and*

$$\frac{\phi(x, 1, \cdots, 1)}{\tilde{\phi}(x)} \to \mu \quad \text{if} \quad x \to \infty. \tag{2.7}$$

Let $X_1, \cdots, X_n, \cdots$ *be* i.i.d. *non-negative random variables with* $E(\tilde{\phi}(X_1)) = \infty$. *Then*

$$\frac{U_n(\phi)}{U_n(\tilde{\phi})} \to k\mu \leq 1 \quad \text{a.s.} \tag{2.8}$$

Another application of Theorem 1 to compare the growth rates of statistics is:

THEOREM 8. *Let* $X_1, X_2, \cdots, X_n, \cdots$ *be* i.i.d. *with* $E|X_1|^p = \infty$ *for some* $p > 1$. *Then*

$$\frac{\sum_{i=1}^n |X_i - \bar{X}|^p}{\sum_{i=1}^n |X_i|^p} \to_{\text{a.s.}} 1 \tag{2.9}$$

where $\bar{X} \equiv (1/n) \sum_{i=1}^n X_i$.

(The result (2.9) does not always hold for $p = 1$; by a different argument the ratio is asymptotically between $1 - \varepsilon$ and 2 a.s. for all $\varepsilon > 0$).

PROOF. Since

$$\begin{aligned}(\textstyle\sum_{i=1}^n |X_i|^p)^{1/p} - (\sum_{i=1}^n |\bar{X}|^p)^{1/p} &\leq (\textstyle\sum_{i=1}^n |X_i - \bar{X}|^p)^{1/p} \\ &\leq (\textstyle\sum_{i=1}^n |X_i|^p)^{1/p} + (\sum_{i=1}^n |\bar{X}|^p)^{1/p},\end{aligned} \tag{2.10}$$

and

$$\frac{(\sum_{i=1}^n |\bar{X}|^p)^{1/p}}{(\sum_{i=1}^n |X_i|^p)^{1/p}} \leq \left[\frac{((1/n)\sum_{i=1}^n X_i)^p}{(1/n)\sum_{i=1}^n |X_i|^p}\right]^{1/p} \to_{\text{a.s.}} 0, \tag{2.11}$$

the result follows.

COROLLARY 9. *Let $X_1, X_2, \cdots X_n, \cdots$ be* i.i.d., *S_n be the sample standard deviation. $h_n = cS_n n^{-\lambda}$, $\lambda > 0$. Then*

$$h_n \to_{\text{a.s.}} 0 \tag{2.12}$$

if and only if

$$E\,|\,X_1\,|^{2/(1+2\lambda)} < \infty. \tag{2.13}$$

PROOF. For $E\,|\,X_1\,| < \infty$.

$$\frac{\sum (X_i - \bar{X})^2}{n^{1+2\lambda}} = \frac{\sum X_i^2}{n^{1+2\lambda}} - \frac{n(\bar{X})^2}{n^{1+2\lambda}}. \tag{2.14}$$

Hence

$$\frac{\sum (X_i - \bar{X})^2}{n^{1+2\lambda}} \to_{\text{a.s.}} 0 \quad \text{iff} \quad \frac{\sum X_i^2}{n^{1+2\lambda}} \to_{\text{a.s.}} 0. \tag{2.15}$$

For $E\,|\,X_1\,| = \infty$

$$\frac{\sum (X_i - \bar{X})^2}{n^{1+2\lambda}} = \frac{\sum X_i^2}{n^{1+2\lambda}} \frac{\sum (X_i - \bar{X})^2}{\sum X_i^2}. \tag{2.16}$$

Applying Theorem 8 with $p = 2$

$$\frac{\sum (X_i - \bar{X})^2}{n^{1+2\lambda}} \to_{\text{a.s.}} 0 \quad \text{iff} \quad \frac{\sum X_i^2}{n^{1+2\lambda}} \to_{\text{a.s.}} 0. \tag{2.17}$$

Then apply the Marcinkiewicz-Zygmund Strong Law of strong numbers.

$$\frac{\sum X_i^2}{n^{1+2\lambda}} \to_{\text{a.s.}} 0 \quad \text{iff} \quad E\,|\,X_1^2\,|^{1/(1+2\lambda)} < \infty. \quad \square \tag{2.18}$$

Finally, if we regard Corollary 4 as a strengthened result of the Cauchy-Schwartz Inequality under stronger conditions, the following theorem strengthens the familiar arithmetic mean-geometric mean inequality.

THEOREM 10. *Let $X_1, X_2, \cdots, X_n, \cdots$ be* i.i.d., *$X_1 \geq 0$. Then a necessary and sufficient condition for*

$$\frac{(X_1 X_2 \cdots X_n)^{1/n}}{(1/n) \sum_{i=1}^{n} X_i} \to_{\text{a.s.}} 0 \tag{2.19}$$

is

$$E(X_1 - \log X_1) = \infty. \tag{2.20}$$

PROOF. Condition (2.20) is equivalent to

$$EX_1 = \infty \quad \text{or} \quad E \log X_1 = -\infty. \tag{2.21}$$

If $EX_1 = \infty$, then

$$(2.22)\qquad \begin{aligned}(X_1X_2 \cdots X_n)^{1/n} &= [(X_1^{1/2}X_1^{1/2}X_2^{1/2}X_2^{1/2} \cdots X_n^{1/2}X_n^{1/2})^{1/2n}]^2 \\ &\le [(1/2n)\ 2 \textstyle\sum_{i=1}^n X_i^{1/2}]^2 = [(1/n) \sum_{i=1}^n X_i^{1/2}]^2,\end{aligned}$$

and

$$\frac{[(1/n) \sum_{i=1}^n X_i^{1/2}]^2}{(1/n) \sum_{i=1}^n X_i} \rightarrow_{\text{a.s.}} 0.$$

Hence (2.19) holds.

If $EX_1 < \infty$ and $E \log X_1 = -\infty$, apply the strong law of large numbers to the logarithm of the numerator, and (2.19) also holds in this case. Suppose (2.19) is true, then either $(1/n) \sum_{i=1}^n X_i \rightarrow_{\text{a.s.}} \infty$ or $(X_1X_2 \cdots X_n)^{1/n} \rightarrow_{\text{a.s.}} 0$; applying the strong law of large numbers, we get $EX_1 = \infty$ or $E \log X_1 = -\infty$.

Acknowledgements. The authors want to express their thanks to a referee and an associate editor. They pointed out the relevant references [1], [2], [3], and their suggestions led to an improved presentation of this paper.

REFERENCES

[1] BINGHAM, N. H. and GOLDIE, C. M. (1982). Extensions of regular variation, I: uniformity and quantifiers. *Proc. London Math. Soc.* **44** 473–496.

[2] BINGHAM, N. H. and GOLDIE, C. M. (1982). Extensions of regular variation, II: representations and indices. *Proc. London Math. Soc.* **44** 497–534.

[3] DRASIN, D. and SHEA, D. F. (1972). Pólya peaks and the oscillation of positive functions. *Proc. Amer. Math. Soc.* **34** 403–411.

[4] MULHOLLAND, H. P. (1931). The generalization of certain inequality theorems involving powers. *Proc. London Math. Soc. 2nd Ser.* **33** 481–516.

DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF CINCINNATI
CINCINNATI, OHIO 45221

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907