The Annals of Statistics
1980, Vol. 8, No. 6, 1395–1398

ON THE BERRY-ESSEÉN THEOREM FOR RANDOM *U*-STATISTICS[1]

BY IBRAHIM A. AHMAD

Memphis State University

A Berry-Esseén theorem for U-statistics when the sample size is random is presented for the case when the random size is independent of the observations. This result extends the work of Callaert and Janssen. As an application of the special case of sample means, a rate of convergence to normality is obtained for the supercritical Galton-Watson process. Other possible applications are in sequential analysis.

1. Introduction and the main result. Let $X_1, \cdots, X_n$, $n \geq 2$, be independent identically distributed (i.i.d.) random variables with distribution function F. Define a U-statistic by:

$$U_n = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} h(X_i, X_j), \tag{1.1}$$

where $h(.\,,.)$ is a symmetric function of two variables such that $Eh(X_1, X_2) = 0$ and that $g(X_1) = E[h(X_1, X_2) \mid X_1]$ has a finite positive variance σ_g^2. Further let $\{N_n\}$ be a sequence of positive integer-valued random variables independent of the X_i's and such that $E(N_n) = m$ and $\operatorname{Var}(N_n) = m_2$ are finite for all n. A random-sample size U-statistic is defined by:

$$U_{N_n} = \binom{N_n}{2}^{-1} \sum_{1 \leq i < j \leq N_n} h(X_i, X_j). \tag{1.2}$$

Recently, Callaert and Janssen (1978) established the Berry-Esseén theorem for U-statistics. Precisely, they showed that if $E \mid h(X_1, X_2) \mid^3 = \nu_3 < \infty$, then there exists an absolute constant C such that for all $n \geq 2$, $\sup_x \mid P[U_n \leq \sigma_n x] - \Phi(x) \mid \leq Cn^{-1/2}(\nu_3/\sigma_g^3)$, where $\sigma_n^2 = (4/n)\sigma_g^2$. This result can be extended by modifying the proof of Callaert and Janssen (1978) to the case when $E \mid h(X_1, X_2) \mid^{2+\delta} = \nu_{2+\delta} < \infty$, $0 < \delta < 1$ to obtain the bound $Cn^{-\delta/2}(\nu_{2+\delta}/\sigma_g^{2+\delta})$.

On the other hand, Sproule (1974), Theorem 7, proved that if N_n/n converges in probability to a discrete positive random variable, then $P[U_{N_n} \leq 2_g x/(N_n)^{1/2}]$ converges to the distribution function of the standard normal variate as $n \to \infty$. Note that N_n's need not be independent of the X_i's. It would be interesting to investigate the rates of convergence in Sproule's theorem. We obtain, assuming that N_n's and the X_i's are independent, a Berry-Esseén theorem.

Following Callaert and Janssen (1978) we define the projection of U_n by

$$\hat{U}_n = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} [g(X_i) + g(X_j)]. \tag{1.3}$$

Set $V_n = \binom{n}{2} U_n$ and $\hat{V}_n = \binom{n}{2} \hat{U}_n$. The following is the main result of this note.

THEOREM 1.1. *Let* $\{X_n\}$, $n \geq 2$, *be a sequence of* i.i.d. *random variables and let* $\{N_n\}$ *be a sequence of positive integer-valued random variables independent of the* X_i's *and such that* $E(N_n)$

Received December 1978; revised August 1979.

[1] Research conducted while the author was with McMaster University, Hamilton, Ontario, Canada and was partially supported by the National Research Council of Canada, Grant No. A-0819.

AMS 1970 *subject classification.* Primary 60F05; secondary 60J80, 62L10.

Key words and phrases. Berry-Esseén theorem, U-statistics, random indicies, Galton-Watson process, supercritical, sequential analysis.

$= m$ *and* $\mathrm{Var}(N_n) = m_2$ *for all* n. *If* $E|h(X_1, X_2)|^{2+\delta} = \nu_{2+\delta} < \infty$, $0 < \delta \le 1$, *and if* $m_2 < \infty$, *then*

$$\text{(1.4)} \quad \text{(i)} \quad \sup_x |P[U_{N_n} \le 2x(\mathrm{Var}\hat{V}_{N_n})^{1/2} N_n^{-1}] - \Phi(x)| \le C\left\{\left(\frac{\nu_{2+\delta}}{\sigma_g^{2+\delta}}\right) m^{-\delta/2} + \left(\frac{m_2^{1/2}}{m}\right)\right\}$$

$$\text{(1.5)} \quad \text{(ii)} \quad \sup_x |P[U_{N_n} \le 2\sigma_g x/(N_n)^{1/2}] - \Phi(x)| \le C\left[\left(\frac{\nu_{2+\delta}}{\sigma_g^{2+\delta}}\right) m^{-\delta/2} + \frac{m_2^{1/2}}{m} + \left(\frac{m_2^{1/2}}{m}\right)^{1/2}\right]$$

where C *is an absolute constant.*

An important special case of a U-statistics is the sample mean, thus as a direct consequence of Theorem 1 above we obtain a Berry-Esseén theorem for random sums. This problem was also discussed by Landers and Rogge (1976) where they obtain analogous results without assuming the independence between X_i's and N_n's but under more stringent conditions on the indices $\{N_n\}$.

As pointed out by Sproule (1974) random U-statistics are useful in various sequential analysis applications such as sequential fixed width confidence intervals for population functionals. Thus Theorem 1.1 should prove useful in obtaining rates of convergence of sequential procedures that may be special cases of a random U-statistic.

An application of the sample mean special case to obtain rates of convergence in the central limit theorem of an estimate of generation size in the supercritical Galton-Watson process is presented to illustrate further uses of random Berry-Esseén theorems.

2. Proof of the theorem.

(i) Let $p_{k,n} = P[N_n = k]$, $k = 2, 3, \ldots$ and all n. Also let $V_{N_n} = \binom{N_n}{2} U_{N_n}$ and $\hat{V}_{N_n} = \binom{N_n}{2} \hat{U}_{N_n}$ and consider $\sup_x |P[V_{N_n} \le x\,(\mathrm{Var}\hat{V}_{N_n})^{1/2}] - \Phi(x)| = \Delta_n$, say. But

$$\begin{aligned} \Delta_n &= \textstyle\sum_{k=2}^{\infty} p_{k,n} \sup_x |P[V_k \le x(\mathrm{Var}\hat{V}_{N_n})^{1/2}] - \Phi(x)| = \sum_k p_{k,n}\tilde{\Delta}_k \\ &= \textstyle\sum_{k:|k-m|\le m/2} p_{k,n}\tilde{\Delta}_k + \sum_{k:|k-m|>m/2} p_{k,n}\tilde{\Delta}_k \qquad (2.1) \\ &= I_1 + I_2, \quad \text{say.} \end{aligned}$$

Let us evaluate bounds for I_1 and I_2 separately. In what follows C will always denote generic constants not necessarily the same.

$$\begin{aligned} I_1 &= \textstyle\sum_{k:|k-m|\le m/2} p_{k,n} \sup_x |P[V_k/(\mathrm{Var}\,\hat{V}_k)^{1/2} \\ &\le x(\mathrm{Var}\hat{V}_{N_n}/\mathrm{Var}\,\hat{V}_k)^{1/2}] - \Phi(x)| \qquad (2.2) \\ &\le \textstyle\sum_{k:|k-m|\le m/2} p_{k,n} \sup_y |P[U_k \le x(\mathrm{Var}\hat{U}_k)^{1/2}] - \Phi(y)| \\ &\quad + \textstyle\sum_{k:|k-m|\le m/2} p_{k,n} \sup_x \left|\Phi(x) - \Phi\left(x\left(\frac{\mathrm{Var}\hat{V}_{N_n}}{\mathrm{Var}\hat{V}_k}\right)^{1/2}\right)\right| = J_1 + J_2, \quad \text{say.} \end{aligned}$$

But it follows from Callaert and Janssen (1978) that

$$\begin{aligned} J_1 &\le \textstyle\sum_{k:|k-m|\le m/2} p_{k,n}\, C\left(\frac{\nu_{2+\delta}}{\sigma_g^{2+\delta}}\right) k^{-\delta/2} \le C\left(\frac{\nu_{2+\delta}}{\sigma_g^{2+\delta}}\right) \sum_{k:|k-m|\le m/2} p_{k,n}(m/2)^{-\delta/2} \qquad (2.3) \\ &\le C\left(\frac{\nu_{2+\delta}}{\sigma_g^{2+\delta}}\right) m^{-\delta/2}. \end{aligned}$$

Note also that since $|\Phi(x) - \Phi(\epsilon x)| \le C|\epsilon - 1|$ for all ϵ and x we have

$$(2.4)\qquad \begin{aligned} J_2 &\le \textstyle\sum_{k:|k-m|\le m/2} p_{k,n} \left| \left(\dfrac{\operatorname{Var} \hat{V}_{N_n}}{\operatorname{Var} \hat{V}_k} \right)^{1/2} - 1 \right| \\ &\le \textstyle\sum_{k:|k-m|\le m/2} p_{k,n} \dfrac{|\operatorname{Var} \hat{V}_{N_n} - \operatorname{Var} \hat{V}_k|}{\operatorname{Var} \hat{V}_k}. \end{aligned}$$

Note that $\operatorname{Var} \hat{V}_k = k\sigma_g^2$ and $\operatorname{Var} \hat{V}_{N_n} = E(N_n)\sigma_g^2$. Hence we get

$$(2.5)\qquad \text{RHS of (2.4)} \le C \textstyle\sum_{k:|k-m|\le m/2} p_{k,n} \dfrac{|m-k|}{m} \le \dfrac{C}{m} E|N_n - m| \le C \dfrac{m_2^{1/2}}{m}.$$

From (2.3) and (2.5) we obtain that

$$(2.6)\qquad I_1 \le C\left\{ \left(\frac{\nu_{2+\delta}}{\sigma_g^{2+\delta}} \right) m^{-\delta/2} + \frac{m_2^{1/2}}{m} \right\}.$$

Next, note that,

$$(2.7)\qquad \begin{aligned} I_2 &\le \textstyle\sum_{k:|k-m|>m/2} p_{k,n} \tilde{\Delta}_k \le \sum_{k:|k-m|>m/2} p_{k,n} \dfrac{|k-m|}{(m/2)} \\ &\le \dfrac{C}{m} E|N_n - m| \le C \dfrac{m_2^{1/2}}{m}. \end{aligned}$$

Thus we have proved that

$$(2.8)\qquad \Delta_n \le C\left\{ \left(\frac{\nu_{2+\delta}}{\sigma_g^{2+\delta}} \right) m^{-\delta/2} + \frac{m_2^{1/2}}{m} \right\}, \qquad 0 < \delta \le 1.$$

(ii) Note that

$$(2.9)\qquad \begin{aligned} \sup_x |P[U_{N_n} \le 2\sigma_g x/(N_n^{1/2})] - \Phi(x)| &= \sup_x |P[V_{N_n} \le x\sigma_g(N_n^{1/2})] - \Phi(x)| \\ &= \sup_x \left|P\left[V_{N_n} \le x \frac{\sigma_g(N_n^{1/2})}{(\operatorname{Var} \hat{V}_{N_n})^{1/2}} (\operatorname{Var} \hat{V}_{N_n})^{1/2}\right] - \Phi(x)\right| \\ &= \sup_x |P[V_{N_n}/(\operatorname{Var}(\hat{V}_{N_n}))^{1/2} \le x Y_{N_n}] - \Phi(x)|, \end{aligned}$$

where $Y_{N_n} = (N_n/m)^{1/2}$, since $\operatorname{Var} \hat{V}_{N_n} = m\sigma_g^2$. Using Lemma 1 of Michel and Pfanzagl (1971) we have that for any $r > 0$,

$$(2.10)\qquad \begin{aligned} &\sup_x |P[U_{N_n} \le 2\sigma_g x/(N_n)^{1/2}] - \Phi(x)| \\ &\quad \le \sup_x |P[V_{N_n} \le x(\operatorname{Var}(\hat{V}_{N_n}))^{1/2}] - \Phi(x)| + P\left[\left| \left(\frac{N_n}{m} \right)^{1/2} - 1 \right| > r \right] + r. \end{aligned}$$

Choosing $r = (m_2^{1/2}/m)^{1/2}$ the desired conclusion follows since

$$P\left[\left| \left(\frac{N_n}{m} \right)^{1/2} - 1 \right| \ge \epsilon \right] \le P[|N_n - m| \ge mr] \le \frac{\operatorname{Var}(N_n)}{m^2 r^2} = \left(\frac{m_2}{m^2} \right) r^{-2}. \qquad \square$$

3. An application to branching processes. Let $Z_0 = 1, Z_1, Z_2, \ldots$ denote a supercritical Galton-Watson branching process with $EZ_1 = m > 1$ and $\operatorname{Var} Z_1 = \sigma^2 < \infty$. It is well known that $W_n = m^{-n} Z_n$ converges with probability one to a nondegenerate random variable W as $n \to \infty$ (Harris (1963), page 13). Heyde (1971) showed that $(m^2 - m)^{1/2}\sigma^{-1} Z_n^{-1/2} m^n (W - W_n)$ is asymptotically standard normal. Furthermore Heyde and Brown (1971) derived a rate of convergence in this result, viz.,

$$(3.1)\qquad \begin{aligned} &\sup_x |P[(m^2 - m)^{1/2}\sigma^{-1} Z_n^{-1/2} m^n (W - W_n) \le x \mid Z_n > 0] - \Phi(x)| \\ &\quad \le C\sigma^{-(2+\delta)}(m^2 - m)^{(2+\delta)/2} E(Z_n^{-\delta/2} \mid Z_n > 0) E|W - 1|^{2+\delta}, \qquad 0 < \delta \le 1, \end{aligned}$$

provided $E|Z_1|^{2+\delta} < \infty$. The case $\delta = 1$ is proved in Theorem 2 of Heyde and the case $0 < \delta < 1$ can be proved by the easily adaptable lemma of Heyde and Brown (1971).

On the other hand it is possible to use the above theorem to give another bound which does not involve the evaluation of $EZ_n^{-\delta/2}$. Precisely we prove the following result.

THEOREM 3.1 *If* $E|Z_n|^{2+\delta} < \infty$, $0 < \delta \le 1$, *then*

$$\sup_x |P[(m^2 - m)^{1/2}\sigma^{-1}Z_n^{-1/2}m^n(W - W_n) \le x \mid Z_n > 0] - \Phi(x)| \le C\left\{\left(\frac{E|W-1|^{2+\delta}}{\sigma^{2+\delta}}\right)\frac{(m^2-m)^{2+\delta}}{(E(Z_n \mid Z_n > 0))^{\delta/2}} + \frac{[\mathrm{Var}(Z_n \mid Z_n > 0)]^{1/2}}{E(Z_n \mid Z_n > 0)} + \frac{[\mathrm{Var}(Z_n \mid Z_n > 0)]^{1/2}}{[E(Z_n \mid Z_n \mid 0)]^{1/2}}\right\}. \tag{3.2}$$

PROOF. Note that given $Z_n > 0$, $m^n Z_n^{-1/2}(W - W_n)$ has the same distribution as $(Z_n^*)^{1/2}(X_1 + \cdots + X_{Z_n^*})$, where the X_i are i.i.d. and are independent of Z_n^* which is Z_n under the probability measure conditional on $Z_n > 0$. Thus by the above theorem we get

$$\begin{aligned}
&\sup_x |P[(m^2 - m)^{1/2}\sigma^{-1}Z_n^{-1/2}m^n(W - W_n) \le x \mid Z_n > 0] - \Phi(x)| \\
&= \sup_x |P[S_{Z_n^*} \le x(\mathrm{Var}X_1)^{1/2}(Z_n^*)^{1/2} - \Phi(x)| \qquad (3.3)\\
&\le C\left\{\left(\frac{E|X_1|^{2+\delta}}{[\mathrm{Var}(X_1)]^{2+\delta}}\right)m^{-\delta/2} + \frac{m_2^{1/2}}{m} + \left(\frac{m^{1/2}}{m}\right)^{1/2}\right\},
\end{aligned}$$

where $m = EZ_n^*$. Since the distribution of X_1 is the same as that of $W - 1$, the desired conclusion follows. □

REFERENCES

[1] CALLAERT, H. and JANSSEN, P. (1978). The Berry-Esseén theorem for *U*-statistics. *Ann. Statist.* **6** 417–421.

[2] HARRIS, T. E. (1963). *The Theory of Branching Processes.* Springer, Berlin.

[3] HEYDE, C. C. (1971). Some central limit analogues for supercritical Galton-Watson processes. *J. Appl. Probability* **8** 52–59.

[4] HEYDE, C. C. and BROWN, B. M. (1971). An invariance principle and some convergence rate results for branching processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **20** 271–278.

[5] LANDERS, D. and ROGGE, L. (1976). The exact approximation order in the central limit theorem for random summation. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **36** 264–283.

[6] MICHEL, R. and PFANZAGL, J. (1971). The accuracy of the normal approximation for minimum contrast estimates. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **18** 73–84.

[7] SPROULE, R. N. (1974). Asymptotic properties of *U*-statistics. *Trans. Amer. Math. Soc.* **199** 55–64.

DEPARTMENT OF MATHEMATICAL SCIENCES
MEMPHIS STATE UNIVERSITY
MEMPHIS, TENNESSEE 38152