The Canadian Journal of Statistics
Vol. 13, No. 1, 1985, Pages 79–82
La Revue Canadienne de Statistique

The Berry-Esseen bound for Studentized *U*-statistics

R. HELMERS

Centre for Mathematics and Computer Science, Amsterdam

Key words and phrases: Berry-Esseen bound, Studentized *U*-statistics, Student t-statistic, jackknifing, rate of convergence.
AMS 1980 *subject classifications*: Primary 60F05; secondary 62E20.

ABSTRACT

Callaert and Veraverbeke (1981) recently obtained a Berry-Esseen-type bound of order $n^{-\frac{1}{2}}$ for Studentized nondegenerate *U*-statistics of degree two. The condition these authors need to obtain this order bound is the finiteness of the 4.5th absolute moment of the kernel h. In this note it is shown that this assumption can be weakened to that of a finite $(4 + \varepsilon)$th absolute moment of the kernel h, for some $\varepsilon > 0$. Our proof resembles part of Helmers and van Zwet (1982), where an analogous result is obtained for the Student t-statistic. The present note extends this to Studentized *U*-statistics.

RÉSUMÉ

Callaert et Veraverbeke (1981) ont obtenu récemment un limite du type Berry-Esseen de l'ordre $n^{-\frac{1}{2}}$ pour *U*-statistique non-degéneré de degré deux avec un variance estimé. Le condition exigée par ces auteurs pour obtenir un résultat de cet ordre est l'existence du 4.5 moment absolu du noyau h. Ce note montre que cette condition peut être affaibli à l'existence du $4 + \varepsilon$ moment absolu du noyau h, pour $\varepsilon > 0$. Notre épreuve est comparable à une partie de Helmers et van Zwet (1982) où un résultat analogue est obtenu pour la statistique t de Student. Ce note extends ce résultat à *U*-statistique avec un variance estimé.

RESULTS

Let $X_1, X_2, \ldots, X_n$, $n \geq 2$ be independently and identically distributed random variables with common distribution function F. Let $h(x,y)$ be a real-valued function, symmetric in its arguments, and with $Eh(X_1,X_2) = \nu$. Define a *U*-statistic

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

and suppose that $g(X_1) = \mathscr{E}[h(X_1,X_2) - \nu \,|\, X_1]$ has a positive variance σ_g^2. Let

$$S_n^2 = 4(n-1)(n-2)^{-2} \sum_{i=1}^{n} \left[(n-1)^{-1} \sum_{\substack{j=1 \\ j \neq i}}^{n} h(X_i, X_j) - U_n \right]^2,$$

and note that $n^{-1}S_n^2$ is the jackknife estimator of the variance of U_n; i.e., S_n^2 is the sample variance of the "pseudovalues" $nU_n - (n-1)U_{n-1}^i$, where

$$U_{n-1}^i = \binom{n-1}{2}^{-1} \sum\sum_{\substack{1 \leq j < k \leq n \\ j \neq i, k \neq i}} h(X_j, X_k),$$

for $i = 1, 2, \ldots, n$.

THEOREM. *If* $\mathscr{E}|h(X_1, X_2)|^{4+\varepsilon} < \infty$ *for some* $\varepsilon > 0$, *and* $\sigma_g^2 > 0$, *then for* $n \to \infty$

$$\sup_x |P(\{n^{\frac{1}{2}} S_n^{-1}(U_n - \nu) \le x\}) - \Phi(x)| = O(n^{-\frac{1}{2}}). \tag{2}$$

Callaert and Veraverbeke (CV) (1981) proved the theorem for the special case $\varepsilon = \frac{1}{2}$. The purpose of this note is to show that the theorem is also valid in its present form. Our proof will rely heavily on the proof given by Callaert and Veraverbeke. However, to deal with the part of their proof which required the full force of their 4.5th absolute moment assumption, we will modify their proof and employ the following lemma to obtain a sharper result.

LEMMA. *Let*

$$V_n = \binom{n}{2}^{-1} \sum\sum_{1 \le i < j \le n} h_n(X_i, X_j) \tag{3}$$

be a U*-statistic with a varying kernel* h_n *of the form*

$$h_n = \alpha + n^{-1}\beta, \tag{4}$$

where α *and* β *are symmetric functions of their two arguments with* $\mathscr{E}\alpha(X_1, X_2) = \nu$ *and* $\mathscr{E}\beta(X_1, X_2) = 0$. *Suppose that* $\gamma(X_1) = \mathscr{E}[\alpha(X_1, X_2) - \nu | X_1]$ *has a positive variance* σ_γ^2. *If* $\mathscr{E}|\gamma(X_1)|^3 < \infty$ *and, for some* $\eta > 0$,

$$\mathscr{E}|\alpha(X_1, X_2)|^{\frac{5}{3}+\eta} < \infty, \qquad \mathscr{E}|\beta(X_1, X_2)|^{1+\eta} < \infty \tag{5}$$

then for $n \to \infty$

$$\sup_x |P(\{\tau_n^{-1}(V_n - \nu) \le x\}) - \Phi(x)| = O(n^{-\frac{1}{2}}), \tag{6}$$

where $\tau_n^2 = 4\, n^{-1}\sigma_\gamma^2$.

Proof. The lemma is a simple consequence of Theorem 4.1 of Helmers and van Zwet (1982). Q.E.D.

Proof of the theorem. As in CV (1981), we write

$$\frac{n^{\frac{1}{2}}(U_n - \nu)}{S_n} = \frac{n^{\frac{1}{2}}(U_n - \nu)}{2\sigma_g} 2\sigma_g S_n^{-1} \tag{7}$$

and establish a stochastic expansion for $2\sigma_g S_n^{-1}$. Using nothing more than the finiteness of $\mathscr{E}|h(X_1, X_2)|^{4+\varepsilon}$ for some $\varepsilon > 0$, it is proved in CV (1981) that

$$2\sigma_g S_n^{-1} = 1 - \tfrac{1}{8}\sigma_g^{-2} n^{-1} \sum_{i=1}^{n} f(X_i) + R_n, \tag{8}$$

where the function f is given by

$$f(x) = 4\{g(x) - \sigma_g^2\} + 8\int_{-\infty}^{\infty} g(y)\{h(x,y) - \nu - g(x) - g(y)\}\, dF(y) \tag{9}$$

for real x, and R_n is a remainder term which is of order $n^{-\frac{1}{2}}(\ln n)^{-1}$, except on a set with probability $O(n^{-\frac{1}{2}})$, as $n \to \infty$. It follows directly from (7) and (8) (cf. CV, 1981, p. 197) that

$$P(\{|n^{\frac{1}{2}}(U_n - \nu)R_n| \geq 2\sigma_g n^{-\frac{1}{2}}\}) \leq P(\{|R_n| \geq n^{-\frac{1}{2}}(\ln n)^{-1}\}) + P(|\{n^{\frac{1}{2}}(U_n - \nu)| \geq 2\sigma_g \ln n\}) = O(n^{-\frac{1}{2}}), \tag{10}$$

where we have applied the lemma (with $\alpha = h$ and $\beta = 0$) to obtain the order bound in the last line. As in CV (1981), (7), (8), and (10) together imply that it suffices now to establish a Berry-Esseen bound for

$$W_n = 2^{-1}\sigma_g^{-1}n^{\frac{1}{2}}(U_n - \nu)\left(1 - \tfrac{1}{8}\sigma_g^{-2}n^{-1}\sum_{i=1}^{n} f(X_i)\right) \tag{11}$$

instead of obtaining such a bound for $n^{\frac{1}{2}}S_n^{-1}(U_n - \nu)$. By slightly modifying the decomposition of W_n employed in CV (1981), we write

$$W_n = W_{n1} + W_{n2}, \tag{12}$$

where $2\sigma_g n^{-\frac{1}{2}}W_{n1} + \nu$ is a U-statistic with varying kernel h_n of the form V_n [cf. (3)] with $h_n = \alpha + n^{-1}\beta$, where α and β are given by

$$\alpha(x, y) = h(x, y) - \tfrac{1}{8}\sigma_g^{-2}\{g(x)f(y) + g(y)f(x)\} \tag{13}$$

and

$$\beta(x, y) = -\tfrac{1}{8}\sigma_g^{-2}\{(h(x, y) - \nu)(f(x) + f(y)) -2\{g(x)f(y) + g(y)f(x)\} - 2\mu\} \tag{14}$$

with $\mu = \int_{-\infty}^{\infty} g(x)f(x)\,dF(x)$ and where W_{n2} is a remainder term satisfying $\mathscr{E}W_{n2} = O(n^{-\frac{1}{2}})$ and

$$P(\{|W_{n2} - \mathscr{E}W_{n2}| \geq n^{-\frac{1}{2}}\}) = O(n^{-\frac{1}{2}}). \tag{15}$$

We note in passing that W_{n1} and W_{n2} are precisely equal to the terms $(n^{\frac{1}{2}}/2\sigma_g)U_n^* + Z_{n1} - \mathscr{E}Z_{n1} + Z_{n2}$ and $\mathscr{E}Z_{n1} + Z_{n3}$ in CV (1981), which together form the decomposition of W_n employed in that paper. The order bound (15) was proved in CV (1981), requiring $\sigma_g^2 > 0$ and the finiteness of $\mathscr{E}h^4(X_1, X_2)$. Thus W_{n2} is also of negligible order of magnitude under our present assumptions. It remains to consider W_{n1}. The statistic $2\sigma_g n^{-\frac{1}{2}}W_{n1} + \nu$ is a U-statistic of the form V_n [cf. (3)] with varying kernel $h_n = \alpha + n^{-1}\beta$, where α and β are given by (13) and (14) and satisfy the requirements $\mathscr{E}\alpha(X_1, X_2) = \nu$ and $\mathscr{E}\beta(X_1, X_2) = 0$. It follows that, if the assumptions of the lemma are satisfied, we have the Berry-Esseen bound

$$\sup_x |P(\{W_{n1} \leq x\}) - \Phi(x)| = O(n^{-\frac{1}{2}}). \tag{16}$$

To check the assumptions needed for (16) we note first that in this case $\gamma(X_1) = \mathscr{E}[\alpha(X_1, X_2) - \nu|X_1] = \mathscr{E}[h(X_1, X_2) - \nu|X_1] = g(X_1)$ and an application of Jensen's inequality for conditional expectations yields $\mathscr{E}|g(X_1)|^3 \leq \mathscr{E}|h(X_1, X_2) - \nu|^3 < \infty$, so that the assumptions $\sigma_\gamma^2 > 0$ and $\mathscr{E}|\gamma(X_1)|^3 < \infty$ of the lemma are clearly satisfied. Secondly we verify the assumption (5) of the lemma. By the independence of X_1 and X_2, the c_r-inequality, and the relations (13) and (14) we see that it suffices to show that the $(\frac{5}{3} + \eta)$th moment of $h(X_1, X_2)$, $g(X_1)$, and $f(X_1)$ and the $(1 + \eta)$th absolute moment of $h(X_1, X_2)\cdot f(X_1)$ are all finite, for some $\eta > 0$. In view of the remark following (16) we need to only consider the last two of these moments. Application of the Schwarz inequality, the c_r-inequality, and the relation (9) easily leads to the requirements

$E(g(X_1))^{4+4\eta} < \infty$, $E(h(X_1, X_2)^{2+2\eta} < \infty$. Jensen's inequality for conditional expectations can be applied once more to find that we only need $\mathscr{E} h(X_1, X_2)^{4+4\eta} < \infty$ to guarantee this. As $\eta > 0$ is arbitrary, the proof of (16) is now complete. Combining (16) with (15), the remark preceeding (15), and the argument leading to (11) completes the proof of the theorem. Q.E.D.

REMARKS

(1) The idea behind the present modification of the proof given in CV (1981) is that by applying the Berry-Esseen bound (6) to W_{n1} we implicitly use rather delicate characteristic-function methods, whereas in CV (1981) crude moment bounds are employed to deal with part of W_{n1}. As a consequence it is possible to relax their 4.5th absolute moment assumption—which CV (1981) really need only in their treatment of the W_{n1}-term—to that of a finite $(4 + \varepsilon)$th absolute moment for the kernel h, for some $\varepsilon > 0$.

(2) If we take $h(x, y) = \frac{1}{2}(x, y)$, the statistic $n^{\frac{1}{2}}S_n^{-1}(U_n - \nu)$ reduces to the one-sample Student t-statistic. For this very special case the theorem was proved in Helmers and van Zwet (1982) in a similar fashion. Note, however, that in this case W_{n1} simplifies, whereas W_{n2} even becomes nonrandom, so that the relation (15) is superfluous. The theorem yields the rate $n^{-\frac{1}{2}}$ for the accuracy of the normal approximation for Student's t, provided $0 < \mathscr{E}|X_1|^{4+\varepsilon} < \infty$ for some $\varepsilon > 0$, whereas CV (1981) need a finite and positive 4.5th absolute moment for F to prove this.

(3) In a recent paper of Ahmad (1983) it was stated (see his Theorem 2.1) that the Berry-Esseen bound for nondegenerate Studentized U-statistics of degree 2 is also valid under the weaker assumption of a finite 4th absolute moment for the kernel h. However, the proof given obviously fails at a crucial point [in the relation (2.10) of Ahmad (1983), T_N^2 must be replaced by T_N; the resulting probability bound is only $O(1)$, instead of the required $O(n^{-\frac{1}{2}})$]. In fact, it is quite clear that Ahmad's approach cannot produce a Berry-Esseen-type bound of order $n^{-\frac{1}{2}}$: the middle term on the r.h.s. of the basic inequality (2.6) in Ahmad (1983) is typically of the order of 1 if we take $\varepsilon_n = n^{-\frac{1}{2}}$.

Note added in proof. Since this work was completed, a paper of Zhao Lincheng ((1983), *Science Exploration*, Changsha, China, **3**, no. 2, 45–52) has appeared. In this paper the Berry-Esseen bound for non-degenerate Studentized U-statistics of degree 2 is proved under the slightly weaker assumption of a finite fourth moment for the kernel h. Zhao Lincheng's proof resembles ours, but his treatment of the remainder R_n in the stochastic expansion (8) is somewhat different.

REFERENCES

Ahmad, I.A. (1983). On the normal approximation of an estimate of the mean residual life of a multicomponent system. *J. Statist. Plann. Inference*, 7, 195–207.

Callaert, H., and Veraverbeke, N. (1981). The order of the normal approximation for a Studentized U-statistic, *Ann. Statist.*, 9, 194–200.

Helmers, R., and van Zwet, W.R. (1982). The Berry-Esseen bound for U-statistics. *Statistical Decision Theory and Related Topics III. Volume 1* (S.S. Gupta and J.O. Berger, *eds.*) Academic Press, New York, 497–512.

Received 20 December 1982
Revised 25 July 1984
Accepted 1 November 1984

Centre for Mathematics and Computer Science
Kruislaan 413
1098 SJ Amsterdam
The Netherlands