The Annals of Probability
2000, Vol. 28, No. 1, 511–535

THE BERRY–ESSÉEN BOUND FOR STUDENTIZED STATISTICS

BY QIYING WANG,[1] BING-YI JING[1] AND LINCHENG ZHAO[2]

University of Wollongong, Hong Kong University of Science and Technology and University of Science and Technology of China

We derive Berry–Esséen bounds for a class of Studentized statistics. The results are applied to Studentized U-statistics, Studentized L-statistics and Studentized functions of the sample mean to give the Berry–Esséen bounds under conditions weaker than those obtained by alternative methods.

1. Introduction. Suppose we are interested in the distribution F_n of some statistic $T_n = T(X_1, \dots, X_n)$, where $X_1, \dots, X_n$ are n independent identically distributed (i.i.d.) real random variables. Typically, one can "linearize" the statistic T_n and prove that the linearized statistic is equivalent to T_n as $n \to \infty$ in the sense that the difference between the two goes to zero in probability. As a result, F_n can often be approximated by a normal distribution through the use of the central limit theorem. It is then of both theoretical and practical interest to examine the error in the normal approximation.

The rate of convergence to normality has been intensively studied in various situations. The classical Berry–Esséen bound for sample means is contained, for instance, in Feller [(1971), page 543]. Bounds for other statistics are also available. For instance, Bhattacharya and Rao (1976) gave a bound for functions of multivariate sample means. For U-statistics, Berry–Esséen bounds were established under different sets of conditions and in increasing generality by Bickel (1974), Chan and Wierman (1977), Callaert and Janssen (1978), Helmers and van Zwet (1982) and Ghosh (1985) among others. Berry–Esséen bounds for L-statistics were given by Bjerve (1977), Helmers (1977), and Helmers, Janssen and Serfling (1990). Results for R-statistics were provided by Hajek (1968).

It should be pointed out that each of the methods for deriving Berry–Esséen bounds for U-, L- and R-statistics was tailored to the individual structures of these statistics. A general unifying method was proposed by van Zwet (1984), who proved a Berry–Esséen bound theorem for a broad class of statistics, namely symmetric functions of n i.i.d. random variables. Friedrich (1989) removed the symmetry assumption and relaxed the moment conditions. For other extensions to multivariate symmetric statistics and multivariate sampling statistics, we refer to Götze (1991) and Bolthausen and Götze (1993).

The main purpose of the present paper is to establish Berry–Esséen bounds of $O(n^{-1/2})$ for a general class of Studentized statistics under natural and

Received June 1998; revised July 1999.
[1]Supported in part by Hong Kong RGC CERG No. HKUST6162/97P.
[2]Supported in part by National Natural Science Foundation of China (19631040).
AMS 1991 *subject classifications*. Primary 62E20; secondary 60F05.
Key words and phrases. Berry–Esséen bound, Studentized U-statistic, Studentized L-statistic, Studentized functions of the sample mean.

"minimal" or near "minimal" conditions. Berry–Esséen bounds for certain Studentized statistics such as the Studentized U-statistics, Studentized L-statistics and Student's t-statistics have been obtained by various authors. However, moment conditions imposed in these papers are typically stronger than those for their Standardized counterparts. (See Section 3 for literature reviews.) One notable exception is the case of the sample mean, where the third moment assumption is sufficient for both the standardized mean and Student's t-statistic; see Bentkus and Götze (1996). This begs the question whether Berry–Esséen bounds of $O(n^{-1/2})$ are also available for other Studentized statistics under the same moment conditions as those for their standardized counterparts. In the present paper we address the issue and show that it is often the case. To do that, we first derive Berry–Esséen bounds of $O(n^{-1/2})$ for a class of Studentized statistics under weak moment conditions (cf., Theorems 2.1 and 2.2), and then apply these to three special examples: Studentized U-statistics, Studentized L-statistics and Studentized functions of sample means.

We point out that the existing results on Berry–Esséen bounds for general statistics by van Zwet (1984) and Friedrich (1989) are not sufficient for our purpose in the present paper. To appreciate why, let us take the simple case of Student's t-statistics for example. The general result of van Zwet (1984) yields a bound of $O(n^{-1/2})$ provided $E|X_1|^4 < \infty$. Friedrich (1989) improved this result by requiring only $E|X_1|^{10/3} < \infty$. Bentkus, Götze and Zitikis (1994) obtained a lower estimate showing that the result of Friedrich (1989) is final, and therefore offers the best possible result that can be derived from general results. On the other hand, it is well known that the optimal moment condition for Student's t-statistics is $E|X_1|^3 < \infty$; see Bentkus and Götze (1996). The reason why the results of van Zwet (1984) and Friedrich (1989) do not always lead to the best result in certain special cases is that the class of statistics under study is too wide. To get better results, it is necessary to focus on narrower classes of statistics. In this paper, we investigate one such class of studentized statistics [cf. (2.1) and (2.2)], where each term can be examined carefully so that the Berry–Esséen theorem will be valid under natural and "minimal" or near "minimal" conditions.

Section 2 will describe the main results of the paper. Applications of these results to several special cases are presented in Section 3. The proofs of the main results of Section 2 are given in Section 4. Finally, some technical details needed in the proofs of the main theorems are relegated to the Appendix.

Throughout this paper, we denote by $A, A_1, A_2, \dots$ absolute positive contants, which may be different at each occurrence. If a constant depends on a parameter, say u, then we write $A(u)$. Furthermore, we denote the standard normal distribution function by $\Phi(x)$. Finally, we introduce the following notation for simplicity of presentation:

$$\sum_{i<j} \equiv \sum_{1\le i<j\le n}, \qquad \sum_{i<j<k} \equiv \sum_{1\le i<j<k\le n}, \qquad \sum_{i\neq j} \equiv \sum_{\substack{i,\,j=1\\ i\neq j}}^{n}, \qquad \sum_{i\neq j\neq k} \equiv \sum_{\substack{i,\,j,\,k=1\\ i\neq j,\, j\neq k,\, k\neq i}}^{n}.$$

2. Main results. Let $X, X_1, \dots, X_n$ be a sequence of i.i.d. real random variables. Let $\alpha(x)$, $\beta(x, y)$, $\gamma(x, y, z)$ and $\eta(x, y)$ be some real-valued Borel measurable functions of x, y and z. Furthermore, let $V_{in} \equiv V_{in}(X_1, \dots, X_n)$, $i = 1, 2$, be real-valued functions of $\{X_1, \dots, X_n\}$. Define the statistic

$$T = n^{-1/2} \sum_{j=1}^{n} \alpha(X_j) + D_{1n} \sum_{i \neq j} \beta(X_i, X_j) + V_{1n} \tag{2.1}$$

and a normalizing statistic

$$S^2 = 1 + D_{2n} \sum_{i \neq j \neq k} \gamma(X_i, X_j, X_k) + V_{2n}, \tag{2.2}$$

where D_{1n} and D_{2n} are two sequences of real numbers depending only on n. Note that many statistics of interest can be written as special cases of T/S. Our interest here lies in the distribution function of the Studentized statistic T/S, namely, $P(T/S \leq x)$. Under appropriate conditions, the dominant term in T/S is $n^{-1/2} \sum_{j=1}^{n} \alpha(X_j)$, and it has an asymptotic normal distribution by the central limit theorem. In this paper, we study the rate of convergence to normality of T/S. In the first theorem we give a Berry–Esséen bound for the convergence rate of the distribution function $P(T/S \leq x)$ to $\Phi(x)$ under rather natural conditions.

THEOREM 2.1. *Assume that*:

(a) $E\alpha(X_1) = 0$, $E\alpha^2(X_1) = 1$; $E[\beta(X_1, X_2)|X_i] = 0$, $i = 1, 2$; $E\gamma(X_1, X_2, X_3) = 0$.

(b) *The sequences D_{1n} and D_{2n} (depending only on n) satisfy $|D_{1n}| \leq An^{-3/2}$ and $|D_{2n}| \leq An^{-3}$.*

(c) $P(|V_{jn}| \geq C_0 n^{-1/2}) \leq C_j n^{-1/2}$, $j = 1, 2$.

Then, for all $n \geq 3$, we have

$$\sup_x |P(T/S \leq x) - \Phi(x)| \leq [A_1(C_0 + C_1 + C_2) + A_2 \mathscr{L}] n^{-1/2},$$

where $\rho = E|\alpha(X_1)|^3$, $\lambda_s = E|\beta(X_1, X_2)|^s$, $\theta_s = E|\gamma(X_1, X_2, X_3)|^s$ and

$$\mathscr{L} = \rho + \lambda_{5/3} + \theta_{3/2} + [(\rho + \theta_{3/2})\lambda_{3/2}]^{2/3}.$$

Here and below, we denote by $C_0, C_1, C_2, \dots$ positive constants depending on the distribution of X but not on n or other quantities.

In some applications, it is often easier to use the next theorem, which is a simplified version of Theorem 2.1 above. To describe the theorem, define

$$\widetilde{T} = n^{-1/2} \sum_{j=1}^{n} \alpha(X_j) + \widetilde{D}_{1n} \sum_{i<j} \beta(X_i, X_j) + V_{1n}, \tag{2.3}$$

$$\widetilde{S}^2 = 1 + \widetilde{D}_{2n} \sum_{i<j} \eta(X_i, X_j) + V_{2n}. \tag{2.4}$$

We have the following theorem.

THEOREM 2.2. *Assume that*:

(a) $\beta(x, y)$ *and* $\eta(x, y)$ *are symmetric in their arguments.*

(b) $E\alpha(X_1) = 0$; $E\alpha^2(X_1) = 1$; $E[\beta(X_1, X_2)|X_1] = 0; E\eta(X_1, X_2) = 0$.

(c) *The sequences* $\widetilde{D}_{1n}$ *and* $\widetilde{D}_{2n}$ *(depending only on* n*) satisfy* $|\widetilde{D}_{1n}| \le An^{-3/2}$ *and* $|\widetilde{D}_{2n}| \le An^{-2}$.

(d) $P(|V_{jn}| \ge C_0 n^{-1/2}) \le C_j n^{-1/2}$, $j = 1, 2$.

Then, for all $n \ge 2$, *we have*

$$\sup_x \left| P\left(\widetilde{T}/\widetilde{S} \le x\right) = \Phi(x) \right| \le [A_1(C_0 + C_1 + C_2) + A_2 \mathscr{L}] n^{-1/2},$$

where $\mathscr{L}$ *is defined as in Theorem* 2.1 *except that* θ_s *is now replaced by* $\theta'_s = E|\eta(X_1, X_2)|^s$.

3. Some applications. In this section, we shall apply the main results presented in Section 2 to several well-known examples, namely, the Studentized U- and L-statistics and Studentized functions of sample means. Berry–Esséen bounds for these statistics have been studied in recent years by various authors. As can be seen later, applications of Theorems 2.1 and 2.2 to these statistics lead to Berry–Esséen bounds of $O(n^{-1/2})$ under weaker (sometimes minimal) moment conditions.

3.1. *Studentized U-statistics.* Let $h(x, y)$ be a real-valued Borel measurable function, symmetric in its arguments with $Eh(X_1, X_2) = \theta$. Then the U-statistic of degree 2 for estimation of θ with kernel $h(x, y)$ is defined to be

$$U_n = \frac{2}{n(n-1)} \sum_{i<j} h(X_i, X_j).$$

Write

$$g(X_j) = E(h(X_i, X_j) - \theta | X_j), \qquad \sigma_g^2 = \operatorname{Var}(g(X_1)),$$

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

Note that $n^{-1}S_n^2$ is the jackknife estimator of σ_g^2. Define the distributions of the standardized and studentized U-statistic, respectively, by

$$G_1(x) = P\left(\sqrt{n}\sigma_g^{-1}(U_n - \theta) \le x\right),$$

$$G_2(x) = P\left(\sqrt{n}S_n^{-1}(U_n - \theta) \le x\right).$$

It is well known that $G_1(x)$ and $G_2(x)$ converge to the standard normal distribution function $\Phi(x)$ provided $Eh^2(X_1, X_2) < \infty$ and $\sigma_g^2 > 0$ [see Hoeffding (1948) and Arvesen (1969), respectively].

In recent years, there has been considerable interest in obtaining rates of convergence to normality for standardized and Studentized U-statistics. For standardized U-statistics, for instance, this has been investigated by Grams and Serfling (1973), Bickel (1974) and Chan and Wierman (1977). A sharper Berry–Esséen bound was given by Callaert and Janssen (1978), which states that

$$\sup_{t\in R} |G_1(x) - \Phi(x)| \le C\sigma_g^{-3} E|h(X_1, X_2)|^3 n^{-1/2}$$

under the conditions that $E|h(X_1, X_2)|^3 < \infty$ and $\sigma_g^2 > 0$. However, we note that the sharpest Berry–Esséen bound of $O(n^{-1/2})$ for standardized U-statistics comes from Friedrich (1989), who established the ideal bound under the conditions that $E|g(X_1)|^3 < \infty$, $E|h(X_1, X_2)|^{5/3} < \infty$ and $\sigma_g^2 > 0$. Indeed, Bentkus, Götze and Zitikis (1994) showed that the moment conditions of Friedrich (1989) are the weakest for U-statistics.

For Studentized U-statistics, Berry–Esséen bounds were given by Callaert and Veraverbeke (1981) and Helmers (1985) among others. Zhao (1983) sharpened the work of Callaert and Veraverbeke (1981) and obtained the classical rate $O(n^{-1/2})$ under the condition $E|h(X_1, X_2)|^4 < \infty$ and $\sigma_g^2 > 0$. However, it remained an open question whether the moment condition $E|h(X_1, X_2)|^4 < \infty$ can be further weakened to $E|h(X_1, X_2)|^3 < \infty$, as in Callaert and Janssen (1978) for the standardized U-statistics. The answer to this question is affirmative, as the next theorem shows.

THEOREM 3.1. *Assume that $E|h(X_1, X_2)|^3 < \infty$ and $\sigma_g^2 > 0$, then for $n \ge 3$,*

$$\sup_{x} |G_2(x) - \Phi(x)| \le A\sigma_g^{-3} E|h(X_1, X_2)|^3 n^{-1/2}.$$

PROOF. In order to apply Theorem 2.1, we rewrite $\sqrt{n} S_n^{-1}(U_n - \theta) = T/S$, where

$$T = \frac{\sqrt{n}(U_n - \theta)}{2\sigma_g}, \qquad S^2 = \frac{S_n^2}{4\sigma_g^2}.$$

Then from Serfling (1980) and (A_3) in Callaert and Veraverbeke (1981), we have the following expansions:

$$T = n^{-1/2} \sum_{i=1}^{n} \alpha(X_i) + D_{1n} \sum_{i\ne j} \beta(X_i, X_j) + V_{1n},$$

$$S^2 = 1 + D_{2n} \sum_{i\ne j\ne k} \gamma(X_i, X_j, X_k) + V_{2n},$$

where

$$\begin{aligned}\alpha(X_j) &= \sigma_g^{-1} g(X_j),\\ \beta(X_i, X_j) &= \sigma_g^{-1}[h(X_i, X_j) - \theta - g(X_i) - g(X_j)],\\ \gamma(X_i, X_j, X_k) &= \sigma_g^{-2}[h(X_i, X_j) - \theta][h(X_i, X_k) - \theta] - 1,\\ D_{1n} &= n^{-1/2}(n-1)^{-1},\\ D_{2n} &= (n-1)^{-1}(n-2)^{-2},\\ V_{1n} &= 0,\\ V_{2n} &= 2(n-2)^{-1} + Q_{n1} + Q_{n2}\end{aligned}$$

with Q_{n1} and Q_{n2} given by

$$Q_{n1} = \frac{2\sigma_g^{-2}}{(n-1)(n-2)^2} \sum_{i<j} (h(X_i, X_j) - \theta)^2,$$

$$Q_{n2} = -\frac{n(n-1)\sigma_g^{-2}}{(n-2)^2} (U_n - \theta)^2.$$

The conditions (a) and (b) in Theorem 2.1 can be easily checked. By Jensen's inequality, we have

$$\sigma_g^3 \le E|g(X_1)|^3 \le 8E|h(X_1, X_2)|^3,$$

$$E|\beta(X_i, X_j)|^s \le A\sigma_g^{-s} E|h(X_1, X_2)|^s \le A\sigma_g^{-s}\big(E|h(X_1, X_2)|^3\big)^{s/3} \quad \text{for } 0 \le \text{s} \le 3,$$

$$E|\gamma(X_i, X_j, X_k)|^{3/2} \le A\sigma_g^{-3} E|h(X_1, X_2)|^3.$$

From these inequalities, clearly $\mathscr{L} \le A\sigma_g^{-3} E|h(X_1, X_2)|^3$. Now, by Theorem 2.1, we only need to show that $P(|V_{2n}| \ge n^{-1/2}) \le A\sigma_g^{-3} E|h(X_1, X_2)|^3 n^{-1/2}$, which follows from

$$P\big(|V_{2n}| \ge n^{-1/2}\big) \le n^{1/2} E\big|2(n-2)^{-1} + Q_{n1} + Q_{n2}\big| \le A n^{-1/2} \sigma_g^{-3} E|h(X_1, X_2)|^3,$$

where we have used $E|h(X_1, X_2)|^3/\sigma_g^3 \ge 1/8$. The proof of Theorem 3.1 is thus complete. □

Remark 3.1. If $h(x, y) = (x + y)/2$, then the Studentized U-statistic reduces to the Student t-statistic, whose Berry–Esséen bounds have been investigated by Slavova (1985), Hall (1988) and Bentkus and Götze (1996). For other results on Studentized U-statistics, we refer to Maesono (1996, 1997) and Ghosh (1985).

Remark 3.2. It remains an open question whether the moment conditions in Theorem 3.1 can be further weakened to $E|g(X_1)|^3 < \infty$, $E|h(X_1, X_2)|^{5/3} < \infty$ and $\sigma_g^2 > 0$.

3.2. *Studentized L-statistics.* Let $X_1, \dots, X_n$ be i.i.d. real random variables with distribution function F. Define F_n to be the empirical distribution, that is, $F_n(x) = n^{-1} \sum_{j=1}^{n} I\{X_i \le x\}$, where $I\{\cdot\}$ is the indicator function. Let $J(t)$ be a real-valued function on [0,1] and $T(G) = \int xJ(G(x))\, dG(x)$. The statistic $T(F_n)$ is called an L-statistic [see Chapter 8 of Serfling (1980)]. Write

$$\sigma^2 \equiv \sigma^2(J, F) = \int\int J(F(s))J(F(t))F(\min\{s,t\})[1 - F(\max\{s,t\})]\, ds\, dt.$$

Clearly, a natural estimate of σ^2 is given by $\hat{\sigma}^2 \equiv \sigma^2(J, F_n)$. Now let us define the distributions of the standardized and Studentized L-statistic $T(F_n)$ respectively by

$$H_1(x) = P\Big(\sqrt{n}\sigma^{-1}(T(F_n) - T(F)) \le x\Big),$$

$$H_2(x) = P\Big(\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F)) \le x\Big).$$

It is well known that $H_1(x)$ and $H_2(x)$ converge to the standard normal distribution function $\Phi(x)$ provided $E|X_1|^2 < \infty$, $\sigma^2 > 0$ and some smoothness conditions on $J(t)$ [see Serfling (1980) and Helmers, Janssen and Serfling (1990) for references].

For standardized L-statistics, the rates of convergence to normality have been studied by various authors. For instance, assuming that $E|X_1|^3 < \infty$, $\sigma^2 > 0$ and some smoothness conditions on $J(t)$, Helmers (1977) and Helmers, Janssen and Serfling (1990) showed that

$$\sup_{x \in R} |H_1(x) - \Phi(x)| = O(n^{-1/2}).$$

For Studentized L-statistics, on the other hand, a Berry–Esséen bound was also given by Helmers (1982) under the same conditions as those given above except that $E|X_1|^3 < \infty$ is now replaced by a stronger condition $E|X_1|^{4.5} < \infty$. In the following theorem, we show that the condition $E|X_1|^{4.5} < \infty$ can be weakened to $E|X_1|^3 < \infty$.

THEOREM 3.2. *Assume that*:

(a) *$J''(t)$ is bounded on $t \in [0,1]$.*
(b) *$E|X_1|^3 < \infty$ and $\sigma^2 > 0$.*

Then there exists a positive constant $A(J)$ such that for all $n \ge 2$,

$$\sup_x |H_2(x) - \Phi(x)| \le A(J)\sigma^{-3}E|X_1|^3 n^{-1/2}.$$

PROOF. To apply Theorem 2.1, we rewrite $\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F)) = T/S$, where

$$T = \sqrt{n}(T(F_n) - T(F))/\sigma \quad \text{and} \quad S^2 = \hat{\sigma}^2/\sigma^2.$$

For abbreviation, we introduce the following notation:

$$J_0(t) = J(F(t)), \qquad J_n(t) = J(F_n(t)),$$
$$s \wedge t = \min\{s, t\}, \qquad s \vee t = \max\{s, t\},$$
$$\eta_j(t) = I\{X_j \le t\} - F(t), \qquad Z(s, t, F) = F(s \wedge t)(1 - F(s \vee t)),$$
$$\xi(X_i, X_j) = \sigma^{-2} \int\int J_0(s)J_0(t)\big(I\{X_i \le s \wedge t\}I\{X_j > s \vee t\} - Z(s, t, F)\big)\, ds\, dt,$$
$$\varphi(X_i, X_j, X_k) = \sigma^{-2} \int\int J_0'(s)J_0(t)\eta_i(t)I\{X_j \le s \wedge t\}I\{X_k > s \vee t\}\, ds\, dt.$$

From Lemma B of Serfling [(1980), page 265], we have

$$T(G) - T(F) = -\int \big[K_1(G(x)) - K_1(F(x))\big]\, dx,$$

where $K_1(t) = \int_0^t J(u)\, du$. Then, after some algebra, we have

$$T = n^{-1/2} \sum_{j=1}^{n} \alpha(X_j) + n^{-3/2} \sum_{i \ne j} \beta(X_i, X_j) + V_{1n}, \tag{3.1}$$

$$S^2 = 1 + n^{-3} \sum_{i \ne j \ne k} \gamma(X_i, X_j, X_k) + V_{2n}, \tag{3.2}$$

where

$$\alpha(X_j) = -\sigma^{-1} \int J(F(t))\eta_j(t)\, dt,$$
$$\beta(X_i, X_j) = -\tfrac{1}{2}\sigma^{-1} \int J'(F(t))\eta_i(t)\eta_j(t)\, dt,$$
$$\gamma(X_i, X_j, X_k) = \xi(X_i, X_j) + \varphi(X_i, X_j, X_k),$$
$$V_{1n} = n^{1/2}(Q_{1n} + Q_{2n}),$$
$$V_{2n} = Q_{3n} + Q_{4n} + Q_{5n}$$

with Q_{in}, $i = 1, \ldots, 5$, defined by

$$Q_{1n} = n^{-2} \sum_{j=1}^{n} \beta(X_j, X_j),$$
$$|Q_{2n}| \le A(J)\sigma^{-1} \int |F_n(t) - F(t)|^3\, dt,$$
$$Q_{3n} = 2\sigma^{-2} \int\int \big[J_n(s) - J_0(s) - J_0'(s)(F_n(s) - F(s))\big] \times J_0(t)Z(s, t, F_n)\, ds\, dt,$$
$$Q_{4n} = \sigma^{-2} \int\int (J_n(s) - J_0(s))(J_n(t) - J_0(t))Z(s, t, F_n)\, ds\, dt,$$

$$Q_{5n} = n^{-3} \sum_{j \neq k} [\xi(X_j, X_k) + \varphi(X_j, X_j, X_k) + \varphi(X_k, X_j, X_k)]$$
$$-n^{-1}\sigma^{-2} \int\int F(s \wedge t)(1 - F(s \vee t))\, ds\, dt.$$

The condition (a) of Theorem 2.1 can be shown easily by integration by using the assumption (1): $J''(t)$ is bounded on $t \in [0, 1]$. The condition (b) of Theorem 2.1 is trivial as we have $D_{1n} = n^{-3/2}$ and $D_{2n} = n^{-3}$. For the remainder of this section, we shall show that

$$\mathscr{L} \le A(J)\sigma^{-3}E|X_1|^3, \tag{3.3}$$

$$nE|Q_{1n}| \le A(J)\sigma^{-3}E|X_1|^3 \tag{3.4}$$

and, for $i = 1, 2$,

$$P\big(|V_{in}| > 2n^{-1/2}(1 + nE|Q_{1n}|)\big) \le A(J)n^{-1/2}\sigma^{-3}E|X_1|^3. \tag{3.5}$$

Theorem 3.2 will then follow from Theorem 2.1.

Similarly to the proof of Lemma A is Serfling [(1980), page 288], we can show that

$$|\alpha(X_j)| \le A(J)\sigma^{-1}(|X_j| + E|X_1|), \tag{3.6}$$

$$|\beta(X_i, X_j)| \le A(J)\sigma^{-1}(|X_j| + E|X_1|), \tag{3.7}$$

$$|\xi(X_j, X_k)| \le A(J)\sigma^{-2}\big(X_j^2 + X_k^2 + EX_1^2\big), \tag{3.8}$$

$$|\varphi(X_i, X_j, X_k)| \le A(J)\sigma^{-2}\big(X_j^2 + X_k^2\big). \tag{3.9}$$

Noting $E\alpha^2(X_1) = 1$, it follows from (3.6) that

$$A(J) \le \sigma^{-2}EX_1^2 \le \big(\sigma^{-3}E|X_1|^3\big)^{2/3}.$$

Therefore, it is easy to check (3.3) and (3.4).

Let us look at V_{1n} next. Clearly,

$$\begin{aligned} P\big(n^{1/2}|Q_{1n}| \ge n^{-1/2}(1 + nE|Q_{1n}|)\big) &\le P(|nQ_{1n} - E(nQ_{1n})| \ge 1) \\ &\le A(J)n^{-1/2}\sigma^{-2}EX_1^2. \end{aligned} \tag{3.10}$$

Using $E|F_n(t) - F(t)|^3 \le An^{-3/2}F(t)(1 - F(t))$, we have

$$\begin{aligned} P\Big(n^{1/2}|Q_{2n}| \ge n^{-1/2}\Big) &\le A(J)n^{-1/2}\sigma^{-1}\int F(t)(1 - F(t))\, dt \\ &\le A(J)n^{-1/2}\sigma^{-1}E|X_1|. \end{aligned} \tag{3.11}$$

Hence, from (3.10) and (3.11), we obtain (3.5) for $i = 1$.

We shall now investigate V_{2n}. From Taylor's expansion, assumption (a) and the following inequality: $Z(s,t,F_n) \le [F_n(s)(1-F_n(s))]^{1/2}[F_n(t)(1-F_n(t))]^{1/2}$, we have

$$|Q_{4n}| \le \sigma^{-2} \sup_x |J_0'(x)| \Big(\int |F_n(t) - F(t)|\, dt \Big)^2, \tag{3.12}$$

$$\begin{aligned} |Q_{3n}| &\le \sigma^{-2} \sup_{x,y} |J_0''(x) J_0(y)| \int\int (F_n(s) - F(s))^2 Z(s,t,F_n)\, ds\, dt \\ &\le A(J)\sigma^{-2} \int (F_n(s) - F(s))^2\, ds \int F_n^{1/2}(t)(1 - F_n(t))^{1/2}\, dt. \end{aligned} \tag{3.13}$$

From $E|F_n(t) - F(t)|^k \le An^{-k/2}F(t)(1-F(t))$ and $E[F_n(t)(1-F_n(t))] \le F(t)(1-F(t))$, it follows that

$$\begin{aligned} E\Big\{ \Big(\int |F_n(t) - F(t)\, dt \Big)^2 \Big\} &\le \Big(\int \big(E|F_n(t) - F(t)|^2\big)^{1/2} dt \Big)^2 \\ &\le An^{-1} \Big(\int \big[F(t)(1-F(t))\big]^{1/2} dt \Big)^2 \end{aligned} \tag{3.14}$$

and

$$\begin{aligned} &E\Big\{ \int (F_n(s) - F(s))^2\, ds \int F_n^{1/2}(t)\big(1 - F_n(t)\big)^{1/2} dt \Big\} \\ &\quad \le \Big(\int \big[E(F_n(s) - F(s))^4\big]^{1/2} ds \Big) \Big(\int \big[EF_n(t)(1 - F_n(t))\big]^{1/2} dt \Big) \\ &\quad \le An^{-1} \Big(\int \big[F(t)(1-F(t))\big]^{1/2} dt \Big)^2. \end{aligned} \tag{3.15}$$

Combining (3.12)–(3.15) and applying Markov's inequality, we have

$$\begin{aligned} P(\sqrt{n}|Q_{3n} + Q_{4n}| \ge 1) &\le A(J) n^{-1/2} \sigma^{-2} \Big(\int \big[F(t)(1-F(t))\big]^{1/2} dt \Big)^2 \\ &\le A(J) n^{-1/2} \sigma^{-2} \Big(\int (1 - F(t))^{1/2} dt \Big)^2 \\ &\le A(J) n^{-1/2} \sigma^{-2} \Big(\int_{|t| \le \sigma} 1\, dt + E|X_1|^{3/2} \int_{|t| \ge \sigma} t^{-3/2}\, dt \Big)^2 \\ &\le A(J) n^{-1/2} \sigma^{-3} E|X_1|^3. \end{aligned} \tag{3.16}$$

For Q_{5n}, it follows from (3.8) and (3.9) that $nE|Q_{5n}| \le A(J)\sigma^{-2}EX_1^2$. Hence

$$P\big(\sqrt{n}|Q_{5n}| \ge 1\big) \le A(J) n^{-1/2} \sigma^{-2} EX_1^2. \tag{3.17}$$

Consequently it follows from (3.16) and (3.17) that (3.5) holds for $i = 2$. The proof of Theorem 3.2 is thus complete. □

3.3. *Studentized functions of the sample mean.* Let $X_1, \ldots, X_n$ be i.i.d. real random variables with $EX_1 = \mu$ and $\mathrm{Var}(X_1) = \sigma^2 < \infty$. Let f be a real-valued function differentiable in a neighborhood of μ with $f'(\mu) \neq 0$. Thus the asymptotic variance of $\sqrt{n} f(\overline{X})$ is given by $\sigma_f^2 = (f'(\mu))^2\sigma^2$. Denote the sample mean and sample variance by $\overline{X} = n^{-1}\sum_{i=1}^n X_i$ and $\hat{\sigma}^2 = n^{-1}\sum_{i=1}^n (X_i - \overline{X})^2$. Then an obvious estimate of σ_f is simply $|f'(\overline{X})|\hat{\sigma}$. In this paper, however, we shall use an alternative estimate, that is, the jackknife variance estimate given by

$$\hat{\sigma}_f^2 = \frac{n-1}{n}\sum_{j=1}^n \Big(f\Big(\overline{X}^{(j)}\Big) - f(\overline{X})\Big)^2 \quad \text{where } \overline{X}^{(j)} = \frac{1}{n-1}\Big(\sum_{i=1}^n X_i - X_j\Big).$$

Define the distributions of the standardized and Studentized $f(\overline{X})$ respectively by

$$L_1(x) = P\Big(\sqrt{n}\sigma_f^{-1}(f(\overline{X}) - f(\mu)) \le x\Big),$$
$$L_2(x) = P\Big(\sqrt{n}\hat{\sigma}_f^{-1}(f(\overline{X}) - f(\mu)) \le x\Big).$$

Asymptotic properties of $L_1(x)$ (e.g., the asymptotic normality, Berry–Esséen bound and Edgeworth expansion) have been well studied [see Bhattacharya and Ghosh (1976), for instance]. On the other hand, Miller (1964) showed that $\hat{\sigma}_j^2$ is a consistent estimator of σ_f^2 and hence proved that $L_2(x)$ follows the asymptotic standard normal distribution. In this section, we shall give a Berry–Esséen bound for the convergence rate of $L_2(x)$ to normality. The special case where $f(x) = x$ has been studied by Slavova (1985), Hall (1988) and Bentkus and Götze (1996).

THEOREM 3.3. *Assume that $f^{(3)}(x)$ is bounded in a neighborhood of μ and $f'(\mu) \neq 0$, and that $E|X_1|^3 < \infty$. Then for all $n \ge 2$,*

$$\sup_x |L_2(x) - \Phi(x)| \le A(f) n^{-1/2}\phi(\sigma, f)E|X_1|^3,$$

where $A(f)$ depends only on f and $\phi(\sigma, f)$ depends only on σ.

PROOF. Using Taylor's expansion and noting that $\overline{X}^{(j)} - \overline{X} = (\overline{X} - X_j)/(n-1)$, we can get

$$n\hat{\sigma}_f^2 = (n-1)\sum_{j=1}^n \Big\{f'(\mu)\Big(\overline{X}^{(j)} - \overline{X}\Big) + \Big(f'(\overline{X}) - f'(\mu)\Big)\Big(\overline{X}^{(j)} - \overline{X}\Big)$$
$$+\frac{1}{2}f''\Big(\eta_j\overline{X}^{(j)} + (1-\eta_j)\overline{X}\Big)\Big(\overline{X}^{(j)} - \overline{X}\Big)^2\Big\}^2$$
$$(\text{for } 0 \le \eta_j \le 1)$$

$$
\begin{aligned}
&= \frac{f'^2(\mu)}{n-1} \sum_{j=1}^n \Big(X_j - \overline{X}\Big)^2 + \frac{2f'(\mu)f''(\mu)}{n-1} \sum_{j=1}^n \Big(X_j - \overline{X}\Big)^2 \Big(\overline{X} - \mu\Big) + W_{2n} \\
&= \frac{f'^2(\mu)}{n-1} \sum_{j=1}^n \Big(X_j - \mu\Big)^2 + \frac{2f'(\mu)f''(\mu)}{n-1} \sum_{j=1}^n \Big(X_j - \mu\Big)^2 \Big(\overline{X} - \mu\Big) + W_{3n} \\
&= \sigma^2 f'^2(\mu) + \frac{1}{n^2} \sum_{j \neq k} \Big\{ f'^2(\mu) \Big((X_j - \mu)^2 - \sigma^2\Big) \\
&\qquad\qquad + 2f'(\mu)f''(\mu)(X_j - \mu)^2 (X_k - \mu) \Big\} + W_{4n},
\end{aligned}
$$

where $|W_{4n}| \leq A(f)(K_{n1} + K_{n2} + K_{n3} + K_{n4} + K_{n5})$ with K_{ni} being defined by

$$
\begin{aligned}
K_{n1} &= \Big(\overline{X} - \mu\Big)^2 + |\overline{X} - \mu|^3 + \Big(\overline{X} - \mu\Big)^4, \\
K_{n2} &= \frac{1}{n^2} \sum_{j=1}^n |X_j - \overline{X}|^3, \\
K_{n3} &= \frac{1}{n^3} \sum_{j=1}^n \Big(X_j - \overline{X}\Big)^4, \\
K_{n4} &= \frac{1}{n} \Big(\overline{X} - \mu\Big)^2 \sum_{j=1}^n (X_j - \mu)^2, \\
K_{n5} &= \frac{1}{n^2} \sum_{j=1}^n \Big((X_j - \mu)^2 + |X_j - \mu|^3\Big).
\end{aligned}
$$

Similarly, by Taylor's expansion, we have

$$
\begin{aligned}
f\Big(\overline{X}\Big) - f(\mu) &= f'(\mu)\Big(\overline{X} - \mu\Big) + \frac{1}{2} f''(\mu) \Big(\overline{X} - \mu\Big)^2 + K_n^*, \\
&= \frac{1}{n} \sum_{j=1}^n f'(\mu)(X_j - \mu) \\
&\quad + \frac{1}{2n^2} \sum_{j \neq k} f''(\mu)(X_j - \mu)(X_k - \mu) + K_{n1}^*,
\end{aligned} \tag{3.18}
$$

where K_{n1}^* is the remainder term, which will not be given explicitly here.

In order to apply Theorem 2.2, we rewrite $\sqrt{n}\hat{\sigma}_f^{-1}(f(\overline{X}) - f(\mu)) = \tilde{T}/\tilde{S}$, where

$$
\begin{aligned}
\tilde{T} &= n^{-1/2} \sum_{i=1}^n \alpha(X_i) + n^{-3/2} \sum_{i \neq j} \beta(X_i, X_j) + V_{1n}, \\
\tilde{S}^2 &= 1 + n^{-2} \sum_{i<j} (\eta(X_i, X_j) + \eta(X_j, X_i)) + V_{2n},
\end{aligned}
$$

where

$$\alpha(X_j) = (X_j - \mu)/\sigma,$$

$$\beta(X_i, X_j) = \frac{f''(\mu)}{\sigma f'(\mu)}(X_i - \mu)(X_j - \mu),$$

$$\eta(X_i, X_j) = \frac{(X_j - \mu)^2 - \sigma^2}{\sigma^2} + \frac{f''(\mu)}{\sigma^2 f'(\mu)}(X_i - \mu)^2(X_j - \mu),$$

$$V_{1n} = n^{1/2}\sigma^{-1}K^*_{n1},$$

$$V_{2n} = \sigma^{-2}W_{n4}.$$

The conditions (a) and (b) in Theorem 2.2. $\mathscr{L} \le A(f)\phi(\sigma, f)n^{-1/2}E|X_1|^3$ can be easily checked. In the remainder of this section, we shall show that

$$P(|V_{jn}| \ge n^{-1/2}) \le A(f)\phi(\sigma, f)n^{-1/2}E|X_1|^3 \quad \text{for } j = 1, 2.$$

Theorem 3.3 then follows from Theorem 2.2. We shall investigate V_{2n} first. It suffices to show that

$$P\big(\sqrt{n}K_{nj} \ge A(f)\sigma^2\big) \le A(f)\phi(\sigma, f)E|X_1|^3 \quad \text{for } j = 1, 2, 3, 4, 5. \tag{3.19}$$

Here we only show (3.19) in details for the case $j = 4$. Proofs for $j = 1, 2, 3, 5$ are similar but simpler and hence omitted. Note $K_{n4} = W^*_{n1} + W^*_{n2} + W^*_{n3}$, where

$$W^*_{n1} = \frac{1}{n^3}\sum_{j=1}^{n}(X_j - \mu)^4,$$

$$W^*_{n2} = \frac{1}{n^3}\sum_{j\ne k}\big[(X_j - \mu)^2(X_k - \mu)^2 + 2(X_j - \mu)^3(X_k - \mu)\big],$$

$$W^*_{n3} = \frac{1}{n^3}\sum_{j\ne k\ne m}(X_j - \mu)^2(X_k - \mu)(X_m - \mu).$$

By Markov's inequality, it follows that

$$\begin{aligned} P\Big(\sqrt{n}|W^*_{n1}| \ge A(f)\sigma^2\Big) &\le A(f)\sigma^{-3/2}n^{3/8}E|W^*_{n1}|^{3/4} \\ &\le A(f)\phi(\sigma, f)n^{-5/8}E|X_1|^3, \end{aligned} \tag{3.20}$$

$$\begin{aligned} P\Big(\sqrt{n}|W^*_{n2}| \ge A(f)\sigma^2\Big) &\le A(f)\sigma^{-2}\sqrt{n}E|W^*_{n2}| \\ &\le A(f)\phi(\sigma, f)n^{-1/2}E|X_1|^3. \end{aligned} \tag{3.21}$$

For the term W^*_{n3}, note that $W^*_{n3} = 18n^{-3}\sum_{1\le j<k<m\le n} q(X_j, X_k, X_m)$, where

$$q^*(X_j, X_k, X_m) = (X_j - \mu)^2(X_k - \mu)(X_m - \mu)$$

and

$$q(X_j, X_k, X_m) = \tfrac{1}{3}(q^*(X_j, X_k, X_m) + q^*(X_k, X_j, X_m) + q^*(X_m, X_j, X_k)).$$

It is easy to see that q is symmetric in its arguments and satisfies $E(q(X_j, X_k, X_m)|X_j) = 0$. Therefore, by applying a moment inequality for a degenerate U-statistic, we obtain

$$\begin{aligned} P\Big(\sqrt{n}|W^*_{n3}| \geq A(f)\sigma^2\Big) &\leq A(f)\sigma^{-4/3}n^{-3/4}E|q(X_1, X_2, X_3)|^{3/2} \\ &\leq A(f)\phi(\sigma, f)n^{-1/2}E|X_1|^3. \end{aligned} \tag{3.22}$$

Thus, (3.19) for the case $j = 4$ follows from (3.20)–(3.22). Hence, we have shown that $P(|V_{2n}| \geq n^{-1/2}) \leq A(f)\phi(\sigma, f)n^{-1/2}E|X_1|^3$. Similarly, we can show that $P(|V_{1n}| \geq n^{-1/2}) \leq A(f)\phi(\sigma, f)n^{-1/2}E|X_1|^3$. The proof of Theorem 3.3 is complete. □

4. Proofs of the main results.

4.1. *Proof of Theorem* 2.2. The proof of Theorem 2.2 is rather long. We shall therefore first present the main steps of the proof. The technical details are given in the Appendix. Roughly speaking, the main idea of the proof is first to truncate the Studentized statistic on the level $\sqrt{n}/|x|$ then to approximate the statistic by a U-statistic of second order and finally to apply a result due to Friedrich (1989) to the U-statistic.

Without loss of generality, we may assume that with probability 1,

$$|\eta(X_i, X_j)| \leq 2n^2 \quad \text{for } 1 \leq i \leq n,\ 1 \leq j \leq n. \tag{4.1}$$

For, if not, we may replace $\eta(X_i, X_j)$ and V_{2n} by

$$\begin{aligned} \eta^*(X_i, X_j) &= \eta(X_i, X_j)I\Big\{|\eta(X_i, X_j)| \leq n^2\Big\} \\ &\qquad - E\Big(\eta(X_i, X_j)I\Big\{|\eta(X_i, X_j)| \leq n^2\Big\}\Big), \\ V^*_{2n} &= \widetilde{D}_{2n}\sum_{i<j}[\eta(X_i, X_j) - \eta^*(X_i, X_j)] + V_{2n}, \end{aligned}$$

respectively, since we have

$$\widetilde{D}_{2n}\sum_{i<j}\eta(X_i, X_j) + V_{2n} = \widetilde{D}_{2n}\sum_{i<j}\eta^*(X_i, X_j) + V^*_{2n},$$

where $|\eta^*(X_i, X_j)| \leq 2n^2$, and by Markov's inequality,

$$\begin{aligned} &P\Big[|V^*_{2n}| \geq (C_0 + 1)n^{-1/2}\Big] \\ &\quad \leq P\Big(\widetilde{D}_{2n}\Big|\sum_{i<j}[\eta(X_i, X_j) - \eta^*(X_i, X_j)]\Big| \geq n^{-1/2}\Big) \\ &\qquad + P\Big(|V_{2n}| \geq C_0 n^{-1/2}\Big) \\ &\quad \leq A\sqrt{n}E\Big(|\eta(X_1, X_2)|I\Big\{|\eta| > n^2\Big\}\Big) + C_2 n^{-1/2} \\ &\quad \leq (C_2 + A\theta_{3/2})n^{-1/2}. \end{aligned}$$

Next we note that, for any real random variables X, Y, Z_1, Z_2 and any positive constants C_1 and C_2, we have

$$P\left(\frac{X+Z_1}{\sqrt{1+Y+Z_2}} \le x\right) \le P\left(\frac{X-C_1}{\sqrt{1+Y+C_2}} \le x\right) + \sum_{i=1}^{2} P(|Z_i| \ge C_i);$$

$$P\left(\frac{X+Z_1}{\sqrt{1+Y+Z_2}} \le x\right) \ge P\left(\frac{X+C_1}{\sqrt{1+Y-C_2}} \le x\right) - \sum_{i=1}^{2} P(|Z_i| \ge C_i).$$

Therefore, we can replace V_{jn} by $\pm C_0 n^{-1/2}$. For simplicity, we may further assume $V_{jn} = 0$ for $i = 1, 2$ and

$$\widetilde{D}_{1n} = n^{-3/2}, \qquad \widetilde{D}_{2n} = n^{-1}(n-1)^{-1}.$$

It will be clear that this assumption will not affect the proof of the main results.

Coming back to the main proof of the theorem, we note that $\widetilde{S}^2$ can be rewritten as

$$\widetilde{S}^2 = 1 + \frac{1}{n}\sum_{j=1}^{n} g(X_j) + \frac{1}{n(n-1)}\sum_{i<j} \psi(X_i, X_j),$$

where

$$g(X_j) = E[\eta(X_i, X_j)|X_j],$$
$$\psi(X_i, X_j) = \eta(X_i, X_j) - g(X_i) - g(X_j).$$

Now we define the truncated version of $\widetilde{S}^2$ by

$$\widetilde{S}_0^2 = 1 + \frac{1}{n}\sum_{j=1}^{n} g(X_j) I\{|g(X_j)| \le n/(1+x^2)\} + \frac{1}{n(n-1)}\sum_{i<j} \psi(X_i, X_j).$$

Define

$$\Omega_n = \left\{x\colon \frac{1+|x|^3}{\sqrt{n}}(\rho + \theta_{3/2}) \le 1/32\right\}.$$

Then we have

$$\begin{aligned}
&\sup_x \left|P(\tilde{T}/\tilde{S} \le x) - \Phi(x)\right| \\
&\quad \le \sup_{x\in\Omega_n} \left|P(\tilde{T}/\tilde{S}_0 \le x) - \Phi(x)\right| + \sup_{x\notin\Omega_n} \left|P(\tilde{T}/\tilde{S} \le x) - \Phi(x)\right| \\
&\qquad + \sup_{x\in\Omega_n} \left|P(\tilde{T}/\tilde{S} \le x) - P(\tilde{T}/\tilde{S}_0 \le x)\right|.
\end{aligned}$$

The proof of Theorem 2.2 follows immediately from Lemmas 2–4 in the Appendix. □

4.2. *Proof of Theorem* 2.1. Without loss of generality, we assume that $\gamma(x, y, z)$ is a symmetric real-valued Borel measurable function. Let us introduce kernels $\gamma^{(1)}$, $\gamma^{(2)}$ and $\gamma^{(3)}$, which are defined recursively by the equations

$$\gamma^{(1)}(x_1) = E\gamma(x_1, X_2, X_3),$$

$$\gamma^{(2)}(x_1, x_2) = E\gamma(x_1, x_2, X_3) - \sum_{j=1}^{2} \gamma^{(1)}(x_j),$$

$$\gamma^{(3)}(x_1, x_2, x_3) = \gamma(x_1, x_2, x_3) - \sum_{j=1}^{3} \gamma^{(1)}(x_j) - \sum_{1 \le i < j \le 3} \gamma^{(2)}(x_i, x_j).$$

By applying Hoeffding's decomposition for U-statistics [see, e.g., Lee (1990), page 25], we obtain

$$\begin{aligned}
&\frac{1}{n(n-1)(n-2)} \sum_{i \ne j \ne k} \gamma(X_i, X_j, X_k) \\
&\quad = \binom{n}{3}^{-1} \sum_{i<j<k} \gamma(X_i, X_j, X_k) \\
&\quad = \frac{3}{n} \sum_{i=1}^{n} \gamma^{(1)}(X_i) + 3\binom{n}{2}^{-1} \sum_{i<j} \gamma^{(2)}(X_i, X_j) + H_n^{(3)} \\
&\quad = 3\binom{n}{2}^{-1} \sum_{i<j} \eta(X_i, X_j) + H_n^{(3)},
\end{aligned}$$

where $\eta(X_i, X_j) = \gamma^{(2)}(X_i, X_j) + \frac{1}{2}(\gamma^{(1)}(X_i) + \gamma^{(1)}(X_j))$ and

$$H_n^{(3)} = \binom{n}{3}^{-1} \sum_{i<j<k} \gamma^{(3)}(X_i, X_j, X_k).$$

By applying Jensen's inequality and the moment inequality of degenerate U-statistics [see Chen (1980) for example], it is easy to show that

$$E|\gamma^{(t)}(X_1, X_2, X_t)|^{3/2} \le AE|\gamma(X_1, X_2, X_3)|^{3/2} = A\theta_{3/2} \quad \text{for } t = 1, 2, 3;$$

$$P\Big(\sqrt{n}|H_n^{(3)}| \ge 1\Big) \le E\Big[\sqrt{n}|H_n^{(3)}|\Big]^{3/2} \le An^{-3/4}E|\gamma^{(3)}(X_1, X_2, X_3)|^{3/2}.$$

Therefore, from Theorem 2.2, we have the desired result. □

APPENDIX

In this section, we shall give some lemmas which were used in the proof of the main theorems in Section 4. The notation here is as in Section 4. For abbreviation, we write

$$\begin{aligned}
W_j &= g(X_j)I\Big\{|g(X_j)| \le n/\big(1+x^2\big)\Big\},\\
\overline{Y}_n &= \frac{1}{n}\sum_{j=1}^{n} g(X_j),\\
\overline{W}_n &= \frac{1}{n}\sum_{j=1}^{n} W_j,\\
S_n &= n^{-1/2}\sum_{j=1}^{n}\alpha(X_j),\\
R_{1n} &= n^{-3/2}\sum_{i<j}\beta(X_i, X_j),\\
R_{2n} &= \frac{1}{n(n-1)}\sum_{i<j}\psi(X_i, X_j).
\end{aligned}$$

From these, we can rewrite $\widetilde{T}$ and $\widetilde{S}^2$ in (2.3) and (2.4) as

$$\widetilde{T} = S_n + R_{1n}, \qquad \widetilde{S}^2 = 1 + \overline{Y}_n + R_{2n}. \tag{A.1}$$

We now establish some inequalities concerning the random variables defined above.

LEMMA 1. *Under the conditions of Theorem* 2.2, *we have*

$$\begin{aligned}
\rho = E|\alpha\big(X_1\big)|^3 &\ge 1,\\
E|g\big(X_1\big)|^{3/2} &\le \theta_{3/2},\\
E|\psi\big(X_1, X_2\big)|^{3/2} &\le 4\theta_{3/2},\\
P\big(|\overline{Y}_n| \ge A\big) &\le A_1 n^{-1/2}\theta_{3/2},\\
P\big(|R_{1n}| \ge A\big) &\le A_2 n^{-1/2}\lambda_{5/3},\\
P\big(|R_{2n}| \ge An^{-1/4}\big) &\le A_3 n^{-1/2}\theta_{3/2},\\
P\big(|\overline{W}_n| \ge A\big) &\le A_5 n^{-1/2}\theta_{3/2} \quad \textit{if } x \in \Omega_n.
\end{aligned} \tag{A.2}$$

PROOF. We shall only show the last inequality (A.2) since the proofs of others are easier and hence omitted here. It is easy to see that

$$|E\overline{W}_n| \le E\Big(|g(X_1)|I\Big\{|g(X_1)| \ge n/\big(1+x^2\big)\Big\}\Big) \le \frac{1+|x|}{\sqrt{n}}\theta_{3/2}.$$

Then, for $x \in \Omega_n$, we have

$$\big|E\overline{W}_n\big|^{3/2} \le \left(\frac{(1+|x|)^3}{\sqrt{n}\,\theta_{3/2}}\right)^{1/2}\frac{\theta_{3/2}}{\sqrt{n}} \le \frac{\theta_{3/2}}{\sqrt{n}}.$$

Therefore, by Markov's inequality, we obtain

$$\begin{aligned}P(|\overline{W}_n| \geq A) &\leq A^{-3/2} E|\overline{W}_n|^{3/2} \\ &\leq 2A^{-3/2}\Big(E|\overline{W}_n - E\overline{W}_n|^{3/2} + |E\overline{W}_n|^{3/2}\Big) \\ &\leq An^{-1/2}\theta_{3/2}.\end{aligned}$$

The proof is complete. □

LEMMA 2. *Under the conditions of Theorem* 2.2, *we have*

$$\sup_{x\notin\Omega_n} \big|P\big(\tilde{T}/\tilde{S} \leq x\big) - \Phi(x)\big| \leq An^{-1/2}\mathscr{L}. \tag{A.3}$$

PROOF. From the definition of Ω_n, if $x \notin \Omega_n$ and $|x| \leq 1$, then we have $1 \leq 64(\rho + \theta_{3/2})n^{-1/2} \leq 64n^{-1/2}\mathscr{L}$, which implies that

$$\sup_{x\notin\Omega_n,\, |x|\leq 1} \big|P\big(\tilde{T}/\tilde{S} \leq x\big) - \Phi(x)\big| \leq 128n^{-1/2}\mathscr{L}. \tag{A.4}$$

We thus need to prove only (A.3) for $|x| \geq 1$. From the Berry–Esséen theorem for sums of independent random variables [cf. Feller (1971), page 544],

$$\sup_x |P(S_n \leq x) - \Phi(x)| \leq An^{-1/2}\rho \tag{A.5}$$

and

$$1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}}\exp(-x^2/2) \leq \frac{2}{1+|x|^3} \quad \text{for } x \geq 1, \tag{A.6}$$

we have, for $x \geq 1$ and $x \notin \Omega_n$,

$$\begin{aligned}P(S_n \geq x/2) &\leq 1 - \Phi(x/2) + An^{-1/2}\rho \\ &\leq A_1 n^{-1/2}(\rho + \theta_{3/2}).\end{aligned}$$

From this, (A.6) and Lemma 1, if follows that

$$\begin{aligned}\sup_{x\notin\Omega_n,\, x\geq 1} \big|P(\tilde{T}/\tilde{S} \leq x) - \Phi(x)\big| &= \sup_{x\notin\Omega_n,\, x\geq 1} \big|P(\tilde{T}/\tilde{S} \geq x) - (1 - \Phi(x))\big| \\ &\leq P(|R_{1n}| \geq 1/4) + P(|\overline{Y}_n + R_{2n}| \geq 1/2) \\ &\qquad + \sup_{x\notin\Omega_n,\, x\geq 1} [P(S_n \geq x/2) + 1 - \Phi(x)] \\ &\leq An^{-1/2}\mathscr{L}.\end{aligned} \tag{A.7}$$

Similarly, we have

$$\sup_{x\notin\Omega_n,\, x\leq -1} \big|P(\tilde{T}/\tilde{S} \leq x) - \Phi(x)\big| \leq An^{-1/2}\mathscr{L}. \tag{A.8}$$

Now, (A.3) for the case $|x| \geq 1$ follows from (A.4), (A.7) and (A.8). The proof of Lemma 2 is thus complete. □

LEMMA 3. *Under the conditions of Theorem* 2.2, *we have*

$$\sup_{x\in\Omega_n}\left|P(\tilde{T}/\tilde{S}\le x)-P(\tilde{T}/\tilde{S}_0\le x)\right|\le An^{-1/2}\mathscr{L}.$$

PROOF. Let $Z_k=g(X_k)I\{|g(X_k)|\ge n/(1+x^2)\}$, $1\le k\le n$. For $|x|\le 1$, it is clear that

$$\begin{aligned}\left|P(\tilde{T}/\tilde{S}\le x)-P(\tilde{T}/\tilde{S}_0\le x)\right|&\le\sum_{k=1}^n P(Z_k\ne 0)\\&=\sum_{k=1}^n P\Big(|g(X_k)|\ge n/(1+x^2)\Big)\\&\le 3n^{-1/2}\mathscr{L}.\end{aligned}$$

Hence, it remains to show that

$$\sup_{x\in\Omega_n,\,|x|\ge 1}\left|P(\tilde{T}/\tilde{S}\le x)-P(\tilde{T}/\tilde{S}_0\le x)\right|\le An^{-1/2}\mathscr{L}. \tag{A.9}$$

Without loss of generality, we may assume that $x\ge 1$, since, if $x\le -1$, we can replace $\alpha(X_j)$ and $\eta(X_i,X_j)$ by $-\alpha(X_j)$ and $-\eta(X_i,X_j)$, respectively. From $x\ge 1$, it follows that for any $k=1,\dots,n$,

$$P\big(|\alpha(X_k)|\ge x\sqrt{n}/8\big)\le An^{-3/2}\rho\big(1+|x|^3\big)^{-1}.$$

Hence, from (A.5), (A.6) and $x\in\Omega_n$, we have

$$\begin{aligned}P\left(n^{-1/2}\sum_{j\ne k}\alpha(X_j)\ge x/4\right)&\le P(S_n\ge x/8)+P(|\alpha(X_k)|\ge x\sqrt{n}/8)\\&\le 1-\Phi(x/8)+An^{-1/2}\rho+A_1\big(1+|x|^3\big)^{-1}\\&\le A\big(1+|x|^3\big)^{-1}.\end{aligned}$$

Therefore, for $x\ge 1$ and $x\in\Omega_n$,

$$\begin{aligned}&P(S_n\ge x/2,\ Z_k\ne 0\text{ for some }k\ge 1)\\&\quad\le\sum_{k=1}^n P(S_n\ge x/2,\ Z_k\ne 0)\\&\quad\le\sum_{k=1}^n P\left(n^{-1/2}\sum_{j\ne k}\alpha(X_j)\ge x/4,\,Z_k\ne 0\right)+\sum_{k=1}^n P(n^{-1/2}|\alpha(x_k)|\ge 1/4)\\&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(by independence)}\\&\quad\le A(1+|x|^3)^{-1}\sum_{k=1}^n P\Big(|g(X_k)|\ge n/\big(1+x^2\big)\Big)+A_1n^{-1/2}\mathscr{L}\\&\quad\le An^{-1/2}\mathscr{L}.\end{aligned}$$

Next, by applying Lemma 1, it follows that

$$\begin{aligned}
&\sup_{x\in\Omega_n,\, x\geq 1} P\Big(\widetilde{T}/\widetilde{S} \geq x,\, Z_k \neq 0 \text{ for some } k \geq 1\Big)\\
&\quad \leq P(|R_{1n}| \geq 1/4) + P\big(|\overline{Y}_n + R_{2n}| \geq 1/4\big)\\
&\qquad + P(S_n \geq x/2,\, Z_k \neq 0 \text{ for some } k \geq 1)\\
&\quad \leq An^{-1/2}\mathscr{L}.
\end{aligned} \tag{A.10}$$

Similarly, we have

$$\sup_{x\in\Omega_n,\, x\geq 1} P\Big(\widetilde{T}/\widetilde{S}_0 \geq x,\, Z_k \neq 0 \text{ for some } k \geq 1\Big) \leq An^{-1/2}\mathscr{L}. \tag{A.11}$$

Now (A.9) follows from (A.10) and (A.11). This completes the proof of Lemma 3. □

LEMMA 4. *Under the conditions of Theorem* 2.2, *we have*

$$\sup_{x\in\Omega_n} \Big|P\Big(\widetilde{T}/\widetilde{S}_0 \leq x\Big) - \Phi(x)\Big| \leq An^{-1/2}\mathscr{L}. \tag{A.12}$$

PROOF. We shall only show (A.12) for $x > 0$ *and* $\mathscr{L} < \infty$. The other cases are similar and hence omitted. First, we note that for $x \in \Omega_n$,

$$\frac{1+|x|^k}{\sqrt{n}}\big(\rho + \theta_{3/2}\big) \leq 1/16, \qquad k = 0, 1, 2. \tag{A.13}$$

From the inequality $|(1+u)^{1/2} - 1 - u/2| \leq u^2/6$ for $|u| \leq 1/9$, we have, for $|\overline{W}_n + R_{2n}| \leq 1/9$,

$$\Big|\big(1 + \overline{W}_n + R_{2n}\big)^{1/2} - 1 - \tfrac{1}{2}\big(\overline{W}_n + R_{2n}\big)\Big| \leq \tfrac{1}{3}\Big(\overline{W}_n^2 + R_{2n}^2\Big). \tag{A.14}$$

Recall $\widetilde{T}/\widetilde{S}_0 = (S_n + R_{1n})/(1 + \overline{W}_n + R_{2n})^{1/2}$. Also write $\Delta_n(\tau) = \frac{1}{2}\overline{W}_n + \tau\overline{W}_n^2 + \frac{1}{2}R_{2n}$. It follows from (A.14) that

$$\begin{aligned}
P\Big(\widetilde{T}/\widetilde{S}_0 \leq x\Big) &\leq P\Big[S_n + R_{1n} \leq x\big(1 + \Delta_n(1/3) + n^{-1/2}\big)\Big]\\
&\qquad + P\big(|\overline{W}_n + R_{2n}| \geq 1/9\big) + P\Big(R_{2n}^2 \geq 3n^{-1/2}\Big),
\end{aligned} \tag{A.15}$$

$$\begin{aligned}
P\Big(\widetilde{T}/\widetilde{S}_0 \leq x\Big) &\geq P\Big[S_n + R_{1n} \leq x\big(1 + \Delta_n(-1/3) + n^{-1/2}\big)\Big]\\
&\qquad - P\big(|\overline{W}_n + R_{2n}| \geq 1/9\big) - P\Big(R_{2n}^2 \geq 3n^{-1/2}\Big),
\end{aligned} \tag{A.16}$$

In view of (A.15) and (A.16), then (A.12) follows from Lemma 1 and

$$\delta_n \equiv \sup_{x\in\Omega_n} |P[S_n + R_{1n} \leq x(1 + \Delta_n(\tau) + C_n)] - \Phi(x)| \leq An^{-1/2}\mathscr{L}, \tag{A.17}$$

where $|C_n| \le n^{-1/2}$ and $|\tau| \le 1/3$. Let us define

$$K_n(x) = \frac{1}{\sqrt{n}d_n}\sum_{j=1}^{n}(l(X_j) - El(X_j)) + \frac{1}{n^{3/2}d_n}\sum_{i<j} M(X_i, X_j),$$

where

$$b_n = \frac{1}{2} + \frac{2\tau(n-1)}{n^{3/2}}EW_1,$$

$$l(X_j) = \alpha(X_j) - \frac{xb_n}{\sqrt{n}}W_j - \frac{\tau x}{n^{3/2}}W_j^2,$$

$$d_n^2 = \frac{1}{n}\sum_{j=1}^{n} E(l(X_j) - El(X_j))^2,$$

$$M(X_i, X_j) = \beta(X_i, X_j) - \frac{x}{2\sqrt{n}}\psi(X_i, X_j) - \frac{2\tau x}{\sqrt{n}}(W_i - EW_i)(W_j - EW_j).$$

Note that the term $(\overline{W}_n)^2$ in the definition of $\Delta_n(\tau)$ can be written as

$$\begin{aligned}\left(\overline{W}_n\right)^2 &= \frac{1}{n^2}\left(\sum_{j=1}^{n} W_j^2 + \sum_{i\neq j} W_i W_j\right)\\ &= \frac{1}{n^2}\left(\sum_{j=1}^{n} W_j^2 + 2\sum_{i<j}(W_i - EW_i)(W_j - EW_j)\right.\\ &\qquad\left. +2(n-1)EW_1\sum_{j=1}^{n} W_j - n(n-1)(EW_1)^2\right).\end{aligned}$$

Then it is clear that

$$P(S_n + R_{1n} \le x(1 + \Delta_n(\tau) + C_n)) = P(K_n(x) \le \sigma_n(x)), \tag{A.18}$$

where $\sigma_n(x) = d_n^{-1}(1+\theta_n)x - \sqrt{n}d_n^{-1}El(X_1)$ and $\theta_n = C_n - \tau(EW_1)^2(n-1)/n$. Therefore, from (A.17), we have

$$\begin{aligned}\delta_n &= \sup_{x\in\Omega_n} |P(K_n(x) \le \sigma_n(x)) - \Phi(x)|\\ &\le \sup_{x\in\Omega_n}\sup_{y} |P(K_n(x) \le y) - \Phi(y)| + \sup_{x\in\Omega_n} |\Phi(\sigma_n(x) - \Phi(x)|.\end{aligned} \tag{A.19}$$

We shall investigate the first term on the right-hand side of (A.19). It is clear from the definition of W_j that

$$|EW_1| \le E|g(X_1)|I\{|g(X_1)| \ge n/(1+x^2)\} \le \frac{1+|x|}{\sqrt{n}}\theta_{3/2}, \tag{A.20}$$

$$E|\alpha(X_1)W_1| \le \rho^{1/3}\theta_{3/2}^{2/3} \le \rho + \theta_{3/2}, \tag{A.21}$$

$$E|W_1|^{\alpha} \le \left(\frac{n}{1+x^2}\right)^{\alpha-3/2}\theta_{3/2}, \qquad \alpha \ge 3/2. \tag{A.22}$$

Hence, if $x \in \Omega_n$, we have that $|b_n| \le 2/3$,

$$|El(X_1)| \le \frac{|x|}{\sqrt{n}}\left(\frac{2}{3}|EW_1| + \frac{1}{3n}EW_1^2\right) \le \frac{(1+|x|)|x|}{n}\theta_{3/2} \le 1/4, \tag{A.23}$$

$$\begin{aligned} |d_n^2 - 1| &\le |El(X_1)^2 - 1| + (El(X_1))^2 \\ &\le \frac{2|x|}{\sqrt{n}}\left(\frac{2}{3}E|\alpha(X_1)W_1| + \frac{1}{3n}E|\alpha(X_1)|W_1^2 \right. \\ &\qquad\qquad \left. + \frac{|x|}{3n^{3/2}}E|W_1|^3\right) \\ &\quad + \frac{4x^2}{9n}\left(E|W_1|^2 + \frac{1}{4n^2}EW_1^4\right) + \frac{1}{4}|El(X_1)| \\ &\le \frac{5(1+x^2)}{\sqrt{n}}(\rho + \theta_{3/2}) \\ &\le 1/3. \end{aligned} \tag{A.24}$$

Similarly, it follows from $x \in \Omega_n$ that

$$\begin{aligned} E|l(X_1)|^3 &\le A\left(E|\alpha(X_1)|^3 + \left(\frac{|x|}{\sqrt{n}}\right)^3 E|W_1|^3 \right. \\ &\qquad \left. + \left(\frac{|x|}{3n^{3/2}}\right)^3 E|W_1|^6\right) \\ &\le A_1(\rho + \theta_{3/2}), \end{aligned} \tag{A.25}$$

$$\begin{aligned} E|M(X_1, X_2)|^{3/2} &\le A\left(E|\beta(X_1, X_2)|^{3/2} + \left(\frac{|x|}{2\sqrt{n}}\right)^{3/2} E|\psi(X_1, X_2)|^{3/2} \right. \\ &\qquad \left. + 4\left(\frac{|x|}{\sqrt{n}}\right)^{3/2} \left(E|W_1|^{3/2}\right)^2\right) \\ &\le A\left(\lambda_{3/2} + 2\left(\frac{|x|}{\sqrt{n}}\right)^{3/2}\theta_{3/2} + 4\left(\frac{|x|}{\sqrt{n}}\right)^{3/2}\left(\theta_{3/2}\right)^2\right) \\ &\le A_1\left(\lambda_{3/2} + \max\left(1, \theta_{3/2}^{1/2}\right)\right) \text{ [from (A.13)]}, \end{aligned} \tag{A.26}$$

$$\begin{aligned} E|M(X_1, X_2)|^{5/3} &\le A\left(\lambda_{5/3} + \left(\frac{|x|}{2\sqrt{n}}\right)^{5/3} E|\psi(X_1, X_2)|^{5/3} \right. \\ &\qquad \left. + 4\left(\frac{|x|}{\sqrt{n}}\right)^{5/3}\left(E|W_1|^{5/3}\right)^2\right) \\ &\le A_1\left(\lambda_{5/3} + \theta_{3/2}\right), \end{aligned} \tag{A.27}$$

where in the last inequality of (A.27) we have used the following inequalities:

$$E|\psi(X_1, X_2)|^{5/3} \le 6n^{1/3}\theta_{3/2} \qquad \text{[from (4.1)]},$$

$$\left(\frac{|x|}{\sqrt{n}}\right)^{5/3}(E|W_1|^{5/3})^2 \le \frac{|x|}{\sqrt{n}}\big(E|W_1|^{3/2}\big)^2 \le \theta_{3/2} \qquad \text{[from (A.13) and (A.22)]}.$$

From (A.25) and (A.26), we find that

$$\begin{aligned}
&\big(E|l(X_1)|^3\big)^{2/3}\big(E|M(X_1, X_2)|^{3/2}\big)^{2/3} \\
&\quad \le A\big(\rho + \theta_{3/2}\big)^{2/3}\left[\lambda_{3/2} + \max\left(1, \sqrt{\theta_{3/2}}\right)\right]^{2/3} \\
&\quad \le A\big(\rho + \theta_{3/2}\big)^{2/3}\lambda_{3/2}^{2/3} + \big(\rho + \theta_{3/2}\big)^{2/3}\max\left(1, \theta_{3/2}^{1/3}\right) \\
&\quad \le A\mathscr{L}.
\end{aligned} \tag{A.28}$$

Therefore, from Lemma 5 and (A.24)–(A.28), we obtain that, for any fixed $x \in \Omega_n$,

$$\sup_y |P(K_n(x) \le y) - \Phi(y)| \le An^{-1/2}\mathscr{L}. \tag{A.29}$$

Next, we shall study the second term on the right-hand side of (A.19). From (A.13) and (A.20), when $x \in \Omega_n$, we have

$$|\theta_n| \le \frac{1}{\sqrt{n}} + \frac{1 + |x|}{3\sqrt{n}}\theta_{3/2} \le \frac{2(1 + |x|)}{3\sqrt{n}}(\rho + \theta_{3/2}) \le 1/4 \quad \text{as } \rho \ge 1.$$

Thus it follows from (A.23) and (A.24) that

$$\begin{aligned}
\left|\frac{\sigma_n(x)}{x} - 1\right| &\le \frac{\sqrt{n}|El(X_1)|}{|x|d_n} + \frac{|d_n - 1| + \theta_n}{d_n} \\
&\le \frac{2(1 + |x|)}{\sqrt{n}}\theta_{3/2} + \frac{6(1 + x^2)}{\sqrt{n}}(\rho + \theta_{3/2}) \\
&\le \frac{16(1 + |x|^3)}{\sqrt{n}}(\rho + \theta_{3/2}) \\
&\le \frac{1}{2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
|\Phi[\sigma_n(x)] - \Phi(x)| &\le \frac{1}{\sqrt{2\pi}}|\sigma_n(x) - x|\min\left\{\exp\big(-\sigma_n^2(x)/2\big), \exp\big(-x^2/2\big)\right\} \\
&\le A_1 n^{-1/2}\mathscr{L}|x|(1 + |x|^3)\exp(-x^2/18) \\
&\le An^{-1/2}\mathscr{L}.
\end{aligned} \tag{A.30}$$

Thus (A.17) follows from (A.19), (A.29) and (A.30). We therefore complete the proof of Lemma 4. □

Lemma 5 is a corollary of Theorem 3.1 in Friedrich (1989).

LEMMA 5. *Let* $X_1, \dots, X_n$, $n \geq 2$, *be a sequence of independent real r.v.'s. Define*

$$K_n = \left(\sum_{j=1}^{n} E g_{nj}^2(X_j) \right)^{-1} \left(\sum_{j=1}^{n} g_{nj}(X_j) + \sum_{i<j} \psi_{nij}(X_i, X_j) \right),$$

where $g_{nj}(\cdot)$ *and* $\psi_{nij}(\cdot, \cdot)$, $i \neq j$, *are real-valued Borel measurable functions such that for each* $n \geq 2$, $i \neq j$,

$$E g_{nj}(X_j) = 0, \qquad 0 < s_n^2 \equiv \textstyle\sum_{j=1}^{n} E g_{nj}^2(X_j) < \infty,$$

$$\sup_{n,j} E|g_{nj}(X_j)|^3 \leq \rho^*, \qquad E[\psi_{nij}(X_i, X_j)|X_t] = 0, \ t = i, j,$$

$$\sup_{i \neq j} E|\psi_{nij}(X_i, X_j)|^s \leq \lambda_{n,s}^* \qquad \text{for } s = 3/2, \ 5/3.$$

Then there exists an absolute constant $A > 0$ *such that*

$$\sup_y |P(K_n \leq y) - \Phi(y)| \leq A n s_n^{-3} (\rho^* + \mathscr{L}_1 + \mathscr{L}_2),$$

where $\mathscr{L}_1 = n\big(\rho^* \lambda_{n,3/2}^*\big)^{2/3}$ *and* $\mathscr{L}_2 = n s_n^{4/3} \lambda_{n,5/3}^*$.

Acknowledgments. The authors thank a referee and the Editor for their valuable comments, which have led to this much improved version of the paper.

REFERENCES

ARVESEN, J. N. (1969). Jackknifing U-statistics. *Ann. Math. Statist.* **6** 417–421.

BENTKUS, V. and GÖTZE, F. (1996). The Berry–Esséen bound for Student's statistic. *Ann. Probab.* **24** 491–501.

BENTKUS, V., GÖTZE, F. and ZITIKIS, M. (1994). Lower estimates of the convergence rate for U-statistics. *Ann. Probab.* **22** 1707–1714.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–485.

BHATTACHARYA, R. N. and RAO, R. R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.

BICKEL, P. J. (1974). Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **2** 1–20.

BJERVE, X. (1977). Error bounds for linear combinations of order statistics. *Ann. Statist.* **5** 237–255.

BOLTHAUSEN, E. and GÖTZE, F. (1993). The rate of convergence for multivariate sampling statistics. *Ann. Statist.* **21** 1692–1710.

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

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

CHAN, Y. K. and WIERMAN, J. (1977). On the Berry-Esséen theorem for U-statistics. *Ann. Probab.* **5** 136–139.

CHEN, X. R. (1980). The limit properties for the Von-Mises and U-statistics. *Sci. China Ser. A* **10** 522–532 (in Chinese).

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2.** Wiley, New York.

FRIEDRICH, K. O. (1989). A Berry–Esséen bound for functions of independent random variables. *Ann. Statist.* **17** 170–183.

GHOSH, M. (1985). Berry-Esséen bounds for functionals of U-statistic. *Sankhyā* **47** 255–270.

GÖTZE, F. (1991). On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** 724–739.

GRAMS, W. F. and SERFLING, R. J. (1973). Convergence rates for U-statistics and related statistics. *Ann. Statist.* **1** 153–160.

HAJEK, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.* **39** 325–346.

HALL, P. (1988). On the effect of random norming on the rate of convergence in the central limit theorem. *Ann. Probab.* **16** 1265–1280.

HELMERS, R. (1977). The order of the normal approximation for linear combinations of order statistics with smooth weight functions. *Ann. Probab.* **5** 940–953.

HELMERS, R. (1982). *Edgeworth Expansions for Linear Combinations of Order Statistics. Math. Centre Tracts* **105**. *Math. Centrum*, Amsterdam.

HELMERS, R. (1985). The Berry–Esséen bound for Studentized U-statistics. *Canad. J. Statist.* **13** 79–82.

HELMERS, R., JANSSEN, P. and SERFLING, R. (1990). Berry–Esséen and bootstrap results for generalized L-statistics. *Scand. J. Statist.* **17** 65–77.

HELMERS, R. and VAN ZWET, W. R. (1982). The Berry–Esséen bound for U-statistics. In *Statistical Decision Theory and Related Topics III* (S. S. Gupta and J. O. Berger, eds.) **1** 497–512. Academic Press, New York.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

LEE, A. J. (1990). *U-Statistics—Theorem and Practice.* Dekker, New York.

MAESONO, Y. (1996). Edgeworth expansions for functionals of U-statistics. *Kyushu J. Math.* **50** 311–334.

MAESONO, Y. (1997). Edgeworth expansions for a Studentized U-statistic and jackknife estimator of variance. *J. Statist. Plann. Inference* **61** 61–84.

MILLER, R. G. (1964). A trustworthy jackknife. *Ann. Math. Statist.* **35** 1594–1605.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

SLAVOVA, V. V. (1985). On the Berry–Esséen bound for Student's statistics. *Lecture Notes in Math.* **1155** 355-390. Springer, Berlin.

VAN ZWET W. R. (1984). A Berry–Esséen bound for symmetric statistics. *Z. Wahrsch. Verw. Gebiete* **66** 425–440.

ZHAO, L. (1983). The rate of the normal approximation for a Studentized U-statistic. *Science Exploration* **3** 45–52.

Q. WANG
SCHOOL OF MATHEMATICS
AND APPLIED STATISTICS
UNIVERSITY OF WOLLONGONG, NSW 2522
AUSTRALIA

B.-Y. JING
DEPARTMENT OF MATHEMATICS
HONG KONG UNIVERSITY OF
SCIENCE AND TECHNOLOGY
CLEAR WATER BAY, KOWLOON
HONG KONG
E-MAIL: majing@uxmail.ust.hk

L. ZHAO
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
HEFEI, ANHUI 230026
P.R. CHINA