Statistica Sinica **16**(2006), 1409-1422

EXACT CONVERGENCE RATE AND LEADING TERM IN THE CENTRAL LIMIT THEOREM FOR U-STATISTICS

Qiying Wang and Neville C Weber

The University of Sydney

Abstract: The leading term in the normal approximation to the distribution of U-statistics of degree 2 is derived. This result is applied to establish the exact rate of convergence in the Central Limit Theorem for U-statistics and to obtain the one-term Edgeworth expansion for the distribution function. Analogous results for more general U-type statistics are also considered.

Key words and phrases: Berry-Esséen theorem, characterisation of rate of convergence, Edgeworth expansion, optimal moments, nonlattice condition, U-statistics, L-statistics.

1. Introduction and Main Results

Let $X, X_1, \ldots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables. Let $h(x, y)$ be a real-valued Borel measurable function, symmetric in its arguments with $Eh(X_1, X_2) = 0$. For $n \geq 2$, a U-statistic of degree 2 with kernel $h(x, y)$ is defined by

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

Write $g(x) = Eh(x, X_1)$ and $\phi(x, y) = h(x, y) - g(x) - g(y)$. The statistic U_n may be represented as

$$U_n = \frac{2}{n} \sum_{j=1}^{n} g(X_j) + \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \phi(X_i, X_j) := U_{1n} + U_{2n}. \tag{2}$$

See, for example, Lee (1990, p.25).

Throughout we assume that $Eg^2(X_1) = 1$. This assumption implies that $\sqrt{n}U_{1n}/2$ is a standard sum of non-degenerate iid random variables and its distribution may be approximated by a standard normal distribution Φ. Indeed, the classical result (see Hall (1982, p.11), for example) shows that

$$\sup_x \Big| P\big(\frac{\sqrt{n}U_{1n}}{2} \leq x\big) - \Phi(x) \Big| + n^{-\frac{1}{2}} \asymp \delta_n + n^{-\frac{1}{2}}. \tag{3}$$

Here and below we define

$$\delta_n = Eg^2(X_1)I_{(|g(X_1)|\geq\sqrt{n})} + n^{-\frac{1}{2}}\left|Eg^3(X_1)I_{(|g(X_1)|\leq\sqrt{n})}\right| + n^{-1}Eg^4(X_1)I_{(|g(X_1)|\leq\sqrt{n})},$$

and we say that two sequences of positive numbers $\{a_n\}$ and $\{b_n\}$ satisfy $a_n \asymp b_n$ if $0 < \liminf_{n\to\infty} a_n/b_n \leq \limsup_{n\to\infty} a_n/b_n < \infty$. Note that (3) yields concise results about the rate of convergence in the Central Limit Theorem.

In past decades, there has been considerable interest in stating the accuracy of the normal approximation to the distribution of $\sqrt{n}U_n/2$ in a manner that is similar to (3). In increasing generality, the upper bound for the accuracy of the normal approximation has been established in a number of papers. We mention only Bickel (1974), Chan and Wierman (1977), Callaert and Janssen (1978), Borovskikh (1996, 2001), Alberink and Bentkus (2001, 2002) and Wang (2002). The result given by Borovskikh (2001) (also see Alberink and Bentkus (2001)), which is closest to the upper bound in (3), states that if $E|h(X_1,X_2)|^{5/3} < \infty$, then

$$\sup_x \left|P\Big(\frac{\sqrt{n}U_n}{2} \leq x\Big) - \Phi(x)\right| \leq A\left[Eg^2(X_1)I_{(|g(X_1)|\geq\sqrt{n})} + n^{-\frac{1}{2}}E|g(X_1)|^3 I_{(|g(X_1)|\leq\sqrt{n})}\right] + O(n^{-\frac{1}{2}}), \quad (4)$$

where A is an absolute positive constant.

In contrast to rich results on the upper bound, there are only a few papers concerned with the lower bound for the accuracy of the normal approximation to the distribution of $\sqrt{n}U_n/2$. Maesono (1988, 1991) obtained a lower bound of order $O(n^{-1/2})$ under the condition $Eh^4(X_1,X_2) < \infty$. Only assuming the existence of $Eh^2(X_1,X_2)$, Wang (1992) derived a result for the distribution of $\sqrt{n}U_n/2$ that is similar to (3). In a slightly different problem Bentkus, Götze and Zitikis (1994) proved that the best bound of order $O(n^{-1/2})$ in (4) cannot be obtained under $E|h(X_1,X_2)|^{5/3-\epsilon} < \infty$ for any $\epsilon > 0$.

In the present paper we give the leading term in a normal approximation to the distribution of $\sqrt{n}U_n/2$. Using the leading term we derive the exact convergence rate in the Central Limit Theorem for U-statistics, up to terms of order $O(n^{-1/2})$, under $E|h(X_1,X_2)|^{5/3} < \infty$. As mentioned above, to get the terms of order $O(n^{-1/2})$, the latter moment condition is the best possible. We also show that, if in addition $E|g(X_1)|^3 < \infty$, the leading term transforms into the conventional first term in an Edgeworth expansion of the distribution of U-statistics.

Our main result is the following.

Theorem 1.1. *If* $E|h(X_1,X_2)|^{5/3}<\infty$*, then*

$$\sup_x \Big|P\Big(\frac{\sqrt{n}U_n}{2}\le x\Big)-\Phi(x)-\mathcal{L}_{1n}(x)+\mathcal{L}_{2n}(x)\Big| = o(\delta_n)+O(n^{-\frac12}), \tag{5}$$

where δ_n *is defined as in* (3),

$$\mathcal{L}_{1n}(x)=n\left[E\Phi\Big\{x-\frac{g(X_1)}{\sqrt{n}}\Big\}-\Phi(x)\right]-\frac12\Phi^{(2)}(x),$$

$$\mathcal{L}_{2n}(x)=\frac{\Phi^{(3)}(x)}{2\sqrt{n}}E\big\{g(X_1)g(X_2)\phi(X_1,X_2)I_{(|\phi(X_1,X_2)|\le n^{\frac32})}\big\}.$$

If in addition $g(X_1)$ *is nonlattice, then the right-hand side of* (5) *may be replaced by* $o(\delta_n+n^{-1/2})$.

As is well-known, $\delta_n\to 0$, as $n\to\infty$, and

$$\sup_x|\mathcal{L}_{1n}(x)|\asymp\delta_n, \tag{6}$$

(see, for example, Chapter 2 of Hall (1982)). We show in Section 3 that

$$\sup_x|\mathcal{L}_{2n}(x)|=o(\delta_n)+O(n^{-\frac12}). \tag{7}$$

Together, (5)−(7) give concise results about the rate of convergence in the Central Limit Theorem for U-statistics. Indeed, if $E|h(X_1,X_2)|^{5/3}<\infty$, then

$$\sup_x\Big|P\Big(\frac{\sqrt{n}U_n}{2}\le x\Big)-\Phi(x)\Big|+n^{-\frac12}\asymp\delta_n+n^{-\frac12}. \tag{8}$$

Note that (8) refines (4) even for the upper bound. One application of (8) is to characterise the rate of convergence. The following theorem gives examples. Generalizations of the examples are readily derived, refer to Theorems 2.9 and 2.10 of Hall (1982) for more details.

Theorem 1.2. *Assume* $E|h(X_1,X_2)|^{5/3}<\infty$*. If* $0\le r<1/2$*, then*

$$\sum_{n=1}^{\infty}n^{r-1}\sup_x\Big|P\Big(\frac{\sqrt{n}U_n}{2}\le x\Big)-\Phi(x)\Big|<\infty \tag{9}$$

if and only if $E|g(X_1)|^{2(r+1)}<\infty$*. If* $0<r<1/2$*, then*

$$\sup_x\Big|P\Big(\frac{\sqrt{n}U_n}{2}\le x\Big)-\Phi(x)\Big|=O(n^{-r}) \tag{10}$$

if and only if $Eg^2(X_1)I_{(|g(X_1)|\ge x)}=o(x^{-2r})$.

It is also interesting to note that the effect of $\mathcal{L}_{2n}(x)$ in (5) on the rate of convergence appears only when $\delta_n = O(n^{-1/2})$. This can be easily seen from (7) and the following corollary, which provides the main result of Jing and Wang (2003), about an Edgeworth expansion of the distribution of U-statistics under optimal conditions. Corollary 1.1 below also is an alternative to Theorem 5 of Borovskikh (1998) with $m = 2$, where the Edgeworth expansion is obtained under $\limsup_{|t|\to\infty} |Ee^{itg(X_1)}| < 1$ instead of the condition that the distribution of $g(X_1)$ is nonlattice. Borovskikh (1998) also uses a weaker moment condition on the kernel $h(x, y)$.

Corollary 1.1. *Assume that $E|h(X_1, X_2)|^{5/3} < \infty$, $E|g(X_1)|^3 < \infty$, and the distribution of $g(X_1)$ is nonlattice. Then, as $n \to \infty$,*

$$\sup_x \left| P\Big(\frac{\sqrt{n}U_n}{2} \le x\Big) - F_n(x) \right| = o(n^{-\frac{1}{2}}), \tag{11}$$

where $F_n(x) = \Phi(x) - (\Phi^{(3)}(x)/(6\sqrt{n}))\left\{Eg^3(X_1) + 3Eg(X_1)g(X_2)h(X_1, X_2)\right\}$.

The proof of all results will be given in Section 3. To conclude this section we mention that the rate of convergence in the Central Limit Theorem for U-statistics depends on the moment conditions for both $h(X_1, X_2)$ and $g(X_1)$. If only $E|h(X_1, X_2)|^p < \infty$, where $4/3 < p < 5/3$, the term of order $O(n^{-1/2})$ in (8) has to be replaced by a term of lower order. This follows from Theorem 2.1 in the next section, which gives an extension of Theorem 1.1 to U-type statistics. Throughout the paper we denote constants by $A, A_1, A_2, \ldots$, which may be different at each occurrence.

2. Extensions to U-type Statistics and L-statistics

Let $\alpha(x)$ and $\beta(x, y)$ be some real-valued Borel measurable functions of x and y. Furthermore, let $V_n \equiv V_n(X_1, \ldots, X_n)$ be real-valued functions of $\{X_1, \ldots, X_n\}$. Define a U-type statistic by

$$T_n = n^{-\frac{1}{2}} \sum_{j=1}^{n} \alpha(X_j) + n^{-\frac{3}{2}} \sum_{i\neq j} \beta(X_i, X_j) + V_n. \tag{12}$$

In this section we derive the leading term in a normal approximation to the distribution of T_n under mild conditions, which gives an extension of Theorem 1.1.

Theorem 2.1. *Assume that*

(a) $E\alpha(X_1) = 0$ *and* $E\alpha^2(X_1) = 1$;

(b) $E[\beta(X_1, X_2)|X_i] = 0$, $i = 1, 2$, *and* $E|\beta(X_1, X_2)|^p < \infty$ *for* $4/3 < p \le 5/3$;

(c) $P(|V_n| \geq C_0 n^{-1/2}) \leq C_1 n^{-1/2}$ *for some constants* $C_0 > 0$ *and* $C_1 > 0$.

Then, as $n \to \infty$,

$$\sup_x \Big| P\Big(T_n \leq x\Big) - \Phi(x) - \tilde{\mathcal{L}}_{1n}(x) + \tilde{\mathcal{L}}_{2n}(x) \Big| = o(\delta_{1n} + n^{\frac{4-3p}{2}}) + O(n^{-\frac{1}{2}}), \tag{13}$$

where

$$\delta_{1n} = E\alpha^2(X_1) I_{(|\alpha(X_1)| \geq \sqrt{n})} + n^{-\frac{1}{2}} \big| E\alpha^3(X_1) I_{(|\alpha(X_1)| \leq \sqrt{n})} \big| + n^{-1} E\alpha^4(X_1) I_{(|\alpha(X_1)| \leq \sqrt{n})},$$

$$\tilde{\mathcal{L}}_{1n}(x) = n \left[E\Phi\{x - \frac{\alpha(X_1)}{\sqrt{n}}\} - \Phi(x) \right] - \frac{1}{2}\Phi^{(2)}(x),$$

$$\tilde{\mathcal{L}}_{2n}(x) = \frac{\Phi^{(3)}(x)}{2\sqrt{n}} E\left\{ \alpha(X_1)\alpha(X_2) \left[\beta(X_1, X_2) I_{(|\beta| \leq n^{\frac{3}{2}})} + \beta(X_2, X_1) I_{(|\beta| \leq n^{\frac{3}{2}})} \right] \right\}.$$

If the condition (c) *is replaced by (c′)* $P\big\{|V_n| \geq o(n^{-1/2})\big\} \leq o(n^{-1/2})$, *and in addition* $\alpha(X_1)$ *is nonlattice, then the right-hand side of* (13) *may be replaced by* $o(\delta_n + n^{(4-3p)/2})$.

Note that the U-type statistic T_n defined by (12) is quite general. We next consider an application to L-statistics. Let $X_1, \ldots, X_n$ be i.i.d. real random variables with distribution function F. Define F_n to be the empirical distribution, i.e., $F_n(x) = n^{-1} \sum_{j=1}^n I\{X_i \leq x\}$, where $I\{\cdot\}$ is the indicator function. Let $J(t)$ be a real-valued function on $[0, 1]$ and $T(G) = \int x J(G(x))\, dG(x)$. The statistic $T(F_n)$ is called an L-statistic (see Chapter 8 of Serfling (1980)). Write

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

and define the distribution function of the standardized L-statistic $T(F_n)$ by

$$H_n(x) = P\big(\sqrt{n}\sigma^{-1}(T(F_n) - T(F)) \leq x\big).$$

As is well-known, $H_n(x)$ converges to $\Phi(x)$ uniformly in x provided $E|X_1|^2 < \infty$, $\sigma^2 > 0$, and some smoothness conditions on $J(t)$ hold, see Serfling (1980) and Helmers, Janssen and Serfling (1990) for example. The upper bounds for the rate of convergence to normality were investigated by Helmers (1977) van Zwet (1984), Helmers, Janssen and Serfling (1990), Wang, Jing and Zhao (2000) and Wang (2002).

As a consequence of Theorem 2.1, the following theorem derives the exact convergence rate (two-sided bound) in the Central Limit Theorem for L-statistics, up to terms of order $O(n^{-1/2})$, under mild conditions.

Theorem 2.2. *Assume that*

(a) $|J(s)-J(t)| \le K|s-t|, 0<s<t<1$, *for some* $K>0$;
(b) $EX_1^2<\infty$ *and* $\sigma^2>0$.
Then, as $n\to\infty$,

$$\sup_x |H_n(x)-\Phi(x)| + n^{-\frac{1}{2}} \asymp \delta_{1n} + n^{-\frac{1}{2}}, \tag{14}$$

where $\alpha(X_1) = -\sigma^{-1}\int J(F(t))(I(X_1\le t)-F(t))\,dt$, *and* δ_{1n} *is defined as in Theorem* 2.1.

3. Proofs

Proof of Theorem 1.1. The result is an immediate consequence of Theorem 2.1.

Proof of Theorem 2.1. Without loss of generality we assume that $\beta(x,y)$ is symmetric. Otherwise it is enough to replace $\beta(X_i,Y_j)$ by $\beta(X_i,Y_j)+\beta(X_j,Y_i)$. The proof is along the lines of Jing and Wang (2003). Write

$$\begin{aligned}
&\tilde\beta(X_i,X_j) = \beta(X_i,X_j) I_{(|\beta(X_i,X_j)|\le n^{\frac{3}{2}})},\\
&\alpha^*(X_j) = E\left(\tilde\beta(X_i,X_j)\mid X_j\right), \quad \alpha^{**}(X_j) = \frac{2(n-1)}{n}\alpha^*(X_j) I_{(|\alpha^*(X_j)|\le\sqrt{n})},\\
&T_n^* = n^{-\frac{1}{2}}\sum_{j=1}^n(\alpha(X_j)+\alpha^{**}(X_j)) + 2n^{-\frac{3}{2}}\sum_{i<j}\left(\tilde\beta(X_i,X_j)-\alpha^*(X_i)-\alpha^*(X_j)\right)\\
&\qquad +V_n.
\end{aligned}$$

Noting $E\beta(X_1,X_2)=0$, it is easily seen that

$$\begin{aligned}
|\alpha^*(X_j)| &= \left|E\left(\beta(X_i,X_j)I_{(|\beta(X_i,X_j)|\le n^{3/2})}|X_j\right)\right|\\
&\le E\left(|\beta(X_i,X_j)|\,I_{(|\beta(X_i,X_j)|\ge n^{3/2})}|X_j\right),
\end{aligned} \tag{15}$$

and, as in (2.24)$-$(2.25) of Jing and Wang (2003),

$$\begin{aligned}
&\sup_x |P(T_n\le x) - P(T_n^*\le x)|\\
&\quad \le nP(|\alpha^*(X_1)|\ge\sqrt{n}) + n^2P\left(|\beta(X_1,X_2)|\ge n^{\frac{3}{2}}\right)\\
&\quad \le 2n^{\frac{4-3p}{2}}E|\beta(X_1,X_2)|^p I_{(|\beta(X_1,X_2)|\ge n^{\frac{3}{2}})}\\
&\quad = o\left(n^{\frac{4-3p}{2}}\right).
\end{aligned} \tag{16}$$

We further let, $m_0 = ([10\log n]+1)/b$, where $b>0$ is a constant to be chosen later,

$$\begin{aligned}
\eta_j &= \alpha(X_j) + \alpha^{**}(X_j) - E\alpha^{**}(X_j), \\
\gamma_{ij} &= 2\left[\tilde{\beta}(X_i, X_j) - \alpha^*(X_i) - \alpha^*(X_j) + E\tilde{\beta}(X_i, X_j)\right], \\
S_n &= n^{-\frac{1}{2}} \sum_{j=1}^{n} \eta_j, \\
\Delta_m &= n^{-\frac{3}{2}} \sum_{j=m+1}^{n} \gamma_{mj} \quad \text{for } 1 \le m \le n-1, \\
\Delta_{n,m} &= \sum_{k=m}^{n-1} \Delta_k \quad \text{if } 0 < m < n, \quad \text{and} \quad \Delta_{n,m} = 0 \quad \text{if } m \ge n.
\end{aligned}$$

It follows immediately that $T_n^* = S_n + \Delta_{n,m_0} + \tilde{V}_n + \sqrt{n}E\alpha^{**}(X_1) - (n-1)n^{-1/2}E\tilde{\beta}(X_1, X_2)$, where $\tilde{V}_n = V_n + \sum_{m=1}^{m_0-1} \Delta_m$. Note that for any fixed $1 \le m < k \le n$ and $1 \le q \le 2$,

$$E|\Delta_{n,m} - \Delta_{n,k}|^q \le 8n^{-\frac{3q}{2}+1}(k-m)E|\gamma_{12}|^q; \tag{17}$$

see Theorem 2.1.3 in Koroljuk and Borovskich (1994). It follows from (17) with $q = p$ and $E|\gamma_{12}|^p < \infty$ (see (21) below) that for $4/3 \le p \le 5/3$,

$$P\Big(\Big|\sum_{m=1}^{m_0-1} \Delta_m\Big| \ge \frac{n^{-\frac{1}{2}}}{\log n}\Big) \le An^{1-p} \log^{1+p} n\, E|\gamma_{12}|^p = o\left(n^{\frac{4-3p}{2}}\right). \tag{18}$$

In terms of the condition (c) (or (c$'$)), (18) and the fact that $|\sqrt{n}E\alpha^{**}(X_1) - (n-1)n^{-1/2}E\tilde{\beta}(X_1, X_2)| \le 3\sqrt{n}E|\beta(X_1, X_2)|I_{(|\beta| \ge n^{3/2})} = o(n^{(4-3p)/2})$, routine calculations show that, to prove (13), it suffices to prove

$$\begin{aligned}
I_n &:= \sup_x \Big|P\Big(S_n + \Delta_{n,m_0} \le x\Big) - \Phi(x) - \tilde{\mathcal{L}}_{1n}(x) + \tilde{\mathcal{L}}_{2n}(x)\Big| \\
&= o(\delta_{1n} + n^{\frac{4-3p}{2}}) + O(n^{-\frac{1}{2}})
\end{aligned} \tag{19}$$

and, if in addition $\alpha(X_1)$ is nonlattice, then the right-hand side of (19) may be replaced by $o(\delta_n + n^{(4-3p)/2})$.

We first establish five lemmas before the proof of (19). The proofs of these lemmas will be omitted. The details can be found in Wang and Weber (2004), on which the present paper is based.

Lemma 3.1. *Write* $\hat\alpha(X_1) = \alpha^{**}(X_1) - E\alpha^{**}(X_1)$. *We have*

$$E|\hat\alpha(X_1)|^\lambda \le 2E|\alpha^{**}(X_1)|^\lambda = o\left(n^{\frac{\lambda+2-3p}{2}}\right), \quad \text{for } 1 \le \lambda \le 2, \tag{20}$$

$$E|\gamma_{12}|^p \le 16|\tilde\beta(X_1, X_2)|^p < \infty, \tag{21}$$

$$E|\gamma_{12}|^q \le 16|\tilde\beta(X_1, X_2)|^q = o\left(n^{\frac{3(q-p)}{2}}\right), \quad \text{for } p < q \le 2, \tag{22}$$

$$|E\eta_1^2 - 1| = o\left(\delta_{1n} + n^{\frac{4-3p}{2}}\right). \tag{23}$$

Lemma 3.2. *We have*

$$E\gamma_{12}e^{\frac{it(\eta_1+\eta_2)}{\sqrt{n}}} = -\frac{2t^2}{n}E\left\{\alpha(X_1)\alpha(X_2)\tilde\beta(X_1, X_2)\right\} + o\left(\delta_{1n} + n^{\frac{4-3p}{2}}\right) n^{-\frac{1}{2}}(t^2 + |t|^3), \tag{24}$$

$$\left|E\gamma_{12}e^{\frac{it(\eta_1+\eta_2)}{\sqrt{n}}}\right| \le A\min\left\{(\frac{t}{\sqrt{n}})^{\frac{4(p-1)}{p}}, n^{-1}(E\gamma_{12}^2)^{\frac{1}{2}}(t^2 + |t|^3)\right\}. \tag{25}$$

Next define, $f(t) = Ee^{it\eta_1/\sqrt{n}}$, $g(t) = Ee^{it\alpha(X_1)/\sqrt{n}}$, and $g_n(t) = e^{-t^2/2}(1 + n(g(t)-1) + t^2/2)$.

Lemma 3.3. *There exists a constant* $c_0 > 0$ *such that for all* $|t| \le c_0 n^{1/2}$ *and all sufficiently large* n,

$$|f(t)| \le e^{-\frac{t^2}{8n}}, \qquad |g(t)| \le e^{-\frac{t^2}{4n}}, \tag{26}$$

$$\left|f^n(t) - e^{-\frac{t^2}{2}}\right| \le A\left(\delta_{1n} + o(n^{\frac{4-3p}{2}})\right)(t^2 + t^4)e^{-\frac{t^2}{16}}, \tag{27}$$

$$\left|f^n(t) - g_n(t)\right| = o\left(\delta_{1n} + n^{\frac{4-3p}{2}}\right)(t^2 + t^8)e^{-\frac{t^2}{16}}. \tag{28}$$

If in addition $\alpha(X_1)$ *is nonlattice, then there exist constants* $b > 0$ *and* $\epsilon_n \to \infty$ *such that for* $c_0 \le |t|/\sqrt{n} \le \epsilon_n$,

$$|f(t)| \le e^{-\frac{b}{2}} \text{ and } |g(t)| \le e^{-b}. \tag{29}$$

Lemma 3.4. *For any* $|t| \le c_0\sqrt{n}$, *where* c_0 *is defined as in Lemma* 3.3,

$$\left|E\Delta_{n,m_0}e^{itS_n} + \frac{t^2e^{-\frac{t^2}{2}}}{2\sqrt{n}}E\left\{\alpha(X_1)\alpha(X_2)\left[\tilde\beta(X_1, X_2) + \tilde\beta(X_2, X_1)\right]\right\}\right| = o\left(\delta_{1n} + n^{\frac{4-3p}{2}}\right)(t^2 + t^6)e^{-\frac{t^2}{16}}. \tag{30}$$

To introduce the next lemma we first define some notation. As in (2.13) and

(2.16) of Jing and Wang (2003), we have

$$\begin{aligned} Z_n(t) &:= Ee^{it(S_n+\Delta_{n,m_0})} - Ee^{itS_n} - itE\Delta_{n,m_0}e^{itS_n} \\ &= Z_{n1}(t) + it\left[Z_{n2}^{(1)}(t) + Z_{n2}^{(2)}(t)\right], \end{aligned}$$

where, $l_{m,k} = n^{-3/2}\sum_{j=k+1}^{n}\gamma_{mj}$, $j(m)$ is the largest integer such that $mj(m) < n$ and

$$\begin{aligned} Z_{n1}(t) &= \sum_{m=m_0}^{n-1} Ee^{it(S_n+\Delta_{n,m+1})}\left(e^{it\Delta_m} - 1 - it\Delta_m\right), \\ Z_{n2}^{(1)}(t) &= \sum_{m=m_0}^{n-1}\sum_{j=1}^{j(m)} El_{m,jm}e^{itS_n}\left(e^{it\Delta_{n,jm+1}} - e^{it\Delta_{n,(j+1)m+1}}\right), \\ Z_{n2}^{(2)}(t) &= \sum_{m=m_0}^{n-1}\sum_{j=1}^{j(m)} E\left(l_{m,jm} - l_{m,(j+1)m}\right)e^{itS_n}\left(e^{it\Delta_{n,(j+1)m+1}} - 1\right). \end{aligned}$$

Lemma 3.5. *For* $4/3 < p \le 5/3$, *we have*

$$\int_{|t|\le c_0\sqrt{n}} \frac{1}{|t|}\left|Z_{n1}(t)\right| dt = o\left(\delta_{1n} + n^{\frac{4-3p}{2}}\right), \tag{31}$$

$$\int_{|t|\le c_0\sqrt{n}} \left(|Z_{n2}^{(1)}(t)| + |Z_{n2}^{(2)}(t)|\right) dt = o\left(\delta_{1n} + n^{\frac{4-3p}{2}}\right), \tag{32}$$

where c_0 *is defined as in Lemma* 3.3.

We are now ready to prove (19). We continue to use the notation defined in Lemmas 3.1–3.5. Furthermore write $\varphi_n(t) = -t^2 B_n e^{-t^2/2}$, where

$$B_n = \frac{1}{2\sqrt{n}}E\left\{\alpha(X_1)\alpha(X_2)\left[\beta(X_1,X_2)I_{(|\beta|\le n^{\frac{3}{2}})} + \beta(X_2,X_1)I_{(|\beta|\le n^{\frac{3}{2}})}\right]\right\}.$$

Using Lemmas 3.3–3.5 we have

$$\begin{aligned} J_1(n) &:= \int_{|t|\le c_0\sqrt{n}} \frac{1}{|t|}\left|Ee^{it(S_n+\Delta_{n,m_0})} - g_n(t) - it\varphi_n(t)\right| dt \\ &\le \int_{|t|\le c_0\sqrt{n}} \frac{1}{|t|}\left|Z_n(t)\right| dt + \int_{|t|\le c_0\sqrt{n}} \frac{1}{|t|}\left|f^n(t) - g_n(t)\right| dt \\ &\qquad + \int_{|t|\le c_0\sqrt{n}} \left|E\Delta_{n,m_0}e^{itS_n} - \varphi_n(t)\right| dt \\ &= o(\delta_{1n} + n^{\frac{4-3p}{2}}). \end{aligned} \tag{33}$$

Note that $\int_{-\infty}^{\infty} e^{itx} d\Big(\Phi(x) + \tilde{\mathcal{L}}_{1n}(x) - \tilde{\mathcal{L}}_{2n}(x)\Big) = g_n(t) + it\varphi_n(t)$. It follows from Esseen's smoothing lemma and (33) that

$$\begin{aligned} I_n &\le \int_{|t|\le c_0\sqrt{n}} \frac{1}{|t|} \left| Ee^{it(S_n+\Delta_{n,m_0})} - g_n(t) - it\varphi_n(t)\right| dt \,+\, \frac{A}{\sqrt{n}} \\ &= o(\delta_{1n} + n^{\frac{4-3p}{2}}) + O(n^{-\frac{1}{2}}). \end{aligned} \tag{34}$$

This proves the first part of (19).

If $\alpha(X_1)$ is nonlatice, it follows from the fact that Δ_{n,m_0} only depends on $X_{m_0+1}, \ldots, X_n$, and (29), that for any $\epsilon_n \to \infty$,

$$\begin{aligned} J_2(n) &:= \int_{c_0\sqrt{n}\le|t|\le\epsilon_n\sqrt{n}} \frac{1}{|t|} \left| Ee^{it(S_n+\Delta_{n,m_0})} - g_n(t) - it\varphi_n(t)\right| dt \\ &\le \int_{c_0\sqrt{n}\le|t|\le\epsilon_n\sqrt{n}} \frac{1}{|t|}|f(t)|^{m_0}\, dt + \int_{c_0\sqrt{n}\le|t|\le\epsilon_n\sqrt{n}} \frac{1}{|t|}|g_n(t) + it\varphi_n(t)|\, dt \\ &= o(\delta_{1n} + n^{\frac{4-3p}{2}}). \end{aligned} \tag{35}$$

Using (33), (35) and Esséen's smoothing lemma again, we obtain for $4/3 \le p \le 5/3$,

$$\begin{aligned} I_n &\le \int_{|t|\le\epsilon_n\sqrt{n}} \frac{1}{|t|} \left| Ee^{it(S_n+\Delta_{n,m_0})} - g_n(t) - it\varphi_n(t)\right| dt \,+\, \frac{A}{\epsilon_n\sqrt{n}} \\ &\le J_1(n) + J_2(n) + o(n^{-\frac{1}{2}}) \\ &= o(\delta_{1n} + n^{\frac{4-3p}{2}}). \end{aligned}$$

This implies the second part of (19) and hence the proof of (19).

The proof of Theorem 2.1 is now complete.

Proof of Theorem 2.2. Write $\eta_j(t) = I\{X_j \le t\} - F(t)$,

$$\alpha(X_j) = -\sigma^{-1}\int J(F(t))\eta_j(t)dt, \quad \beta(X_i, X_j) = K\sigma^{-1}\int \eta_i(t)\eta_j(t)dt.$$

As in (29) of Wang (2002), we have

$$\begin{aligned} & n^{-\frac{1}{2}}\sum_{j=1}^{n}\alpha(X_j) - n^{-\frac{3}{2}}\sum_{i\ne j}\beta(X_i,X_j) - V_n \\ &\quad \le \frac{\sqrt{n}\,(T(F_n) - T(F))}{\sigma} \\ &\quad \le\; n^{-\frac{1}{2}}\sum_{j=1}^{n}\alpha(X_j) + n^{-\frac{3}{2}}\sum_{i\ne j}\beta(X_i,X_j) + V_n, \end{aligned} \tag{36}$$

where $V_n = n^{-3/2}\sum_{j=1}^n Z(X_j)$ with $Z(X_j) = K\sigma^{-1}\int \eta_j^2(t)dt$. It is readily seen that $E\alpha(X_1) = 0, E\alpha^2(X_1) = 1$, $E(\beta(X_i, X_j)\Big|X_i) = 0$, $i \neq j$, and similar to the proof of Lemma A in Serfling (1980, p.288),

$$|\alpha(X_j)| + |\beta(X_i, X_j)| + Z(X_j) \leq A\sigma^{-1}(|X_j| + E|X_1|). \tag{37}$$

In terms of these facts, (14) follows easily from Theorem 2.1. We omit the details.

Proof of Corollary 1.1. Equation (11) follows easily from Theorem 1.1, the classical result

$$\sup_x \Big|\mathcal{L}_{1n}(x) + \frac{\Phi^{(3)}(x)}{6\sqrt{n}} Eg^3(X_1)\Big| = o(n^{-\frac{1}{2}}),$$

and by Hölder's inequality, that

$$\begin{aligned}
&\sup_x \Big|\mathcal{L}_{2n}(x) - \frac{\Phi^{(3)}(x)}{2\sqrt{n}} E\left\{g(X_1)g(X_2)\phi(X_1, X_2)\right\}\Big| \\
&\quad \leq A\, n^{-\frac{1}{2}}\, E\left\{|g(X_1)g(X_2)|\,|\phi(X_1, X_2)| I_{(|\phi(X_1,X_2)|\geq n^{\frac{3}{2}})}\right\} \\
&\quad \leq A\, n^{-\frac{1}{2}} \left(E|g(X_1)|^{\frac{5}{2}}\right)^{\frac{4}{5}} \left(E|\phi(X_1, X_2)|^{\frac{5}{3}} I_{(|\phi(X_1,X_2)|\geq n^{\frac{3}{2}})}\right)^{\frac{3}{5}} = o(n^{-\frac{1}{2}}).
\end{aligned}$$

Proof of (7). It suffices to show that

$$\Big|E\left\{g(X_1)g(X_2)\phi(X_1, X_2) I_{(|\phi(X_1,X_2)|\leq n^{\frac{3}{2}})}\right\}\Big|(\sqrt{n})^{-1} = o(\delta_n) + O(n^{-\frac{1}{2}}). \tag{38}$$

Write $\tilde{\phi}(X_1, X_2) = \phi(X_1, X_2) I_{(|\phi(X_1,X_2)|\leq n^{3/2})}$. It is readily seen that

$$\begin{aligned}
&E\left\{g(X_1)g(X_2)\tilde{\phi}(X_1, X_2)\right\} \\
&\quad = E\left\{g(X_1) I_{(|g(X_1)|\geq\sqrt{n})} g(X_2)\tilde{\phi}(X_1, X_2)\right\} \\
&\qquad + E\left\{g(X_1) I_{(|g(X_1)|<\sqrt{n})} g(X_2) I_{(|g(X_2)|\geq\sqrt{n})}\tilde{\phi}(X_1, X_2)\right\} \\
&\qquad + E\left\{g(X_1) I_{(|g(X_1)|<\sqrt{n})} g(X_2) I_{(|g(X_2)|<\sqrt{n})}\tilde{\phi}(X_1, X_2)\right\} \\
&\quad := I_{6n} + I_{7n} + I_{8n}.
\end{aligned} \tag{39}$$

By Hölder's inequality, and similar to (22), we have

$$\begin{aligned}
\frac{|I_{6n}| + |I_{7n}|}{\sqrt{n}} &\leq 2\left(Eg(X_1)^2 I_{(|g(X_1)|\geq\sqrt{n})}\right)^{\frac{1}{2}} \left(E\frac{\tilde{\phi}(X_1, X_2)^2}{n}\right)^{\frac{1}{2}} \\
&\leq 2\,\delta_n^{\frac{1}{2}} \left[o(n^{-\frac{1}{2}})\right]^{\frac{1}{2}} = o\left(\delta_n + n^{-\frac{1}{2}}\right).
\end{aligned} \tag{40}$$

Similarly, by noting $E|\tilde{\phi}(X_1, X_2)|^{5/3} < \infty$,

$$|I_{8n}|/\sqrt{n} \leq n^{-\frac{1}{2}} \left(E|g(X_1)|^{\frac{5}{2}} I_{(|g(X_1)|<\sqrt{n})}\right)^{\frac{4}{5}} \left(E|\tilde{\phi}(X_1, X_2)|^{\frac{5}{3}}\right)^{\frac{3}{5}}$$
$$\leq A n^{-\frac{1}{2}} \left(E|g(X_1)|^{\frac{5}{2}} I_{(|g(X_1)|<\sqrt{n})}\right)^{\frac{4}{5}}. \tag{41}$$

In terms of (41), if $E|g(X_1)|^{5/2} I_{(|g(X_1)|<\sqrt{n})} < \infty$, then

$$\frac{|I_{8n}|}{\sqrt{n}} = O(n^{-\frac{1}{2}}). \tag{42}$$

We show that if $E|g(X_1)|^{5/2} I_{(|g(X_1)|<\sqrt{n})} = \infty$, then

$$\sqrt{n} E|g(X_1)|^{\frac{5}{2}} I_{(|g(X_1)|<\sqrt{n})} \leq A\, E|g(X_1)|^4 I_{(|g(X_1)|<\sqrt{n})}, \tag{43}$$

and hence it follows from (41), that

$$|I_{8n}|/\sqrt{n} = o(1)\, n^{-\frac{1}{2}}\, E|g(X_1)|^{\frac{5}{2}} I_{(|g(X_1)|<\sqrt{n})}$$
$$= o(1)\, n^{-1}\, E|g(X_1)|^4 I_{(|g(X_1)|<\sqrt{n})} = o(\delta_n). \tag{44}$$

Combining (39)−(40), (42) and (44), we obtain the proof of (38).

We next prove (43). Write $l_\tau(x) = E|g(X_1)|^\tau I_{(|g(X_1)|<x)}$. Note that $l_4(x)$ is a non-decreasing function and $l_4(x) \leq A\, x^2$. It follows from Proposition 2.2.1 of Bingham, Goldie and Teugels (1987) that $\limsup_x l_4(2x)/l_4(x) < \infty$. Now (43) follows easily from question 34 on page 289 of Feller (1971) ((also see Feller (1969) or Theorem 2.6.6 of Bingham, Goldie and Teugels (1987)).

The proof of (7) is now complete.

Acknowledgements

The authors would like to thank two referees and an associate editor for their detailed reading of this paper and valuable comments. The research is partially supported by ARC discovery grant DP0451722 and Sesqui NSSS grant 2004.

References

Alberink, I. B. and Bentkus, V. (2001). Berry-Esseen bounds for von-Mises and U-statistics. *Lithuanian Math. J.* **41**, 1-16.

Alberink, I. B. and Bentkus, V. (2002). Lyapunov type bounds for U-statistics. *Theory Probab. Appl.* **46**, 571-588.

Bentkus, V., Götze, F. and Zitikis, M. (1994). Lower estimates of the convergence rate for U-statistics. *Ann. Probab.* **22**, 1707-1714.

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

Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*. Cambridge University Press.

Borovskikh, Yu. V. (1996). *U-statistics in Banach Space*. VSP, Utrecht.

Borovskikh, Yu. V. (1998). Sharp estimates of the rate of convergence for U-statistics. Report 98-21. School of Mathematics and Statistics, University of Sydney.

Borovskikh, Yu. V. (2001). On a normal approximation of U-statistics. *Theory Probab. Appl.* **45**, 406-423.

Callaert, H. and Janssen, P. (1978). The Berry-Esseen theorem for U-statistics. *Ann. Statist.* **6**, 417-421.

Chan, Y. K. and Wierman, J. (1977). On the Berry-Esseen theorem for U-statistics. *Ann. Probab.* **5**, 136-139.

Feller, W. (1969). One-sided analogues of Karamata's regular variation. *L'Enseignement Mathématique* **15**, 107-121.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications* II, second edition. John Wiley, New York.

Hall, P. (1982). *Rates of Convergence in the Central Limit Theorem.* Research Notes in Mathematics, N62. Pitman Advanced Publishing Program.

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

Jing, B.-Y. and Wang, Q. (2003). Edgeworth expansion for U-statistics under minimal conditions. *Ann. Statist.* **31**, 1376-1391.

Koroljuk, V. S. and Borovskich, Yu. V. (1994). *Theory of U-statistics.* Kluwer Academic Publishers, Dordrecht.

Lee, A. J. (1990). *U-statistics - Theory and Practice*. Marcel Dekker, New York.

Maesono, Y. (1988). A lower bound for the normal approximations of U-statistics, *Metrika* **35**, 255-274.

Maesono, Y. (1991). On the normal approximations of U-statistics of degree two. *J. Statist. Plann. Inference* **27**, 37-50.

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

Wang, Q. (1992). Two-sided bounds on the rate of convergence to normal distribution of U-statistics. *J. Systems Sci. Math. Sci.* **12**, 35-40.

Wang, Q. (2002). Non-uniform Berry-Esséen bound for U-statistics. *Statist. Sinica* **12**, 1157-1169.

Wang, Q., Jing, B.-Y. and Zhao, L. (2000). The Berry-Esséen bound for studentized statistics. *Ann. Probab.* **28**, 511-535.

Wang, Q. and Weber, N. C. (2004). Exact convergence rate and leading term in the Central Limit Theorem for U-statistics. School of Mathematics and Statistics, University of Sydney, Research Report 33, (www.maths.usyd.edu.au/u/pubs/publist/preprints/2004/wang-33.html).

van Zwet, W. R. (1984). A Berry-Esséen bound for symmetric statistics. *Z. Wahrsch. Verw. Gebiete* **66**, 425-440.

School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia.
E-mail: qiying@maths.usyd.edu.au

School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia.
E-mail: neville@maths.usyd.edu.au

(Received April 2004; accepted October 2004)