NON-UNIFORM BERRY-ESSÉEN BOUND FOR U-STATISTICS

Qiying Wang

Australian National University

Abstract: Non-uniform Berry-Esséen bounds for *U*-statistics are derived by using a new and simple method. Extensions to *U*-type statistics and *L*-statistics are also discussed.

Key words and phrases: L-statistic, Non-uniform Berry-Esséen bound, *U*-statistic.

1. Introduction and Main Results

Let $X_1, \ldots, X_n$ 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

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

Introducing functions $g(x) = Eh(x, X_1)$ and $\phi(x, y) = h(x, y) - g(x) - g(y)$, we may rewrite the statistic U_n as

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

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

Throughout this paper, we assume that $0 < \sigma_g^2 = Eg^2(X_1) < \infty$. This assumption implies that $\sum_{j=1}^{n} g(X_j)$ is a sum of non-degenerate i.i.d. random variables and its distribution, when properly normalized, may be approximated by the standard normal distribution Φ. Indeed, the classical Berry-Esseen inequality shows that if $E|g(X_1)|^3 < \infty$, then

$$\left| P\left(\frac{1}{\sqrt{n}\sigma_g} \sum_{j=1}^{n} g(X_j) \leq x \right) - \Phi(x) \right| \leq A n^{-1/2} \sigma_g^{-3} E|g(X_1)|^3, \tag{3}$$

uniformly in $x \in R$, where $A > 0$ is an absolute constant. Under some further conditions, i.e., $E|g(X_1)|^p < \infty$, $p \geq 3$, the bound given by (3) can be refined by

the so-called non-uniform Berry-Esseen inequality

$$\left| P\left(\frac{1}{\sqrt{n}\sigma_g} \sum_{j=1}^{n} g(X_j) \le x \right) - \Phi(x) \right| \le A(p)(1+|x|)^{-p} n^{-1/2} \sigma_g^{-p} E|g(X_1)|^p, \quad (4)$$

uniformly in $x \in R$, where $A(p) > 0$ is a constant depending only on p. It should be noted that the bound given by (4) reflects dependence on x and n as well as on $E|g(X_1)|^p$, and the power of $|x|$ and n in (4) is optimal under the assumed moment conditions; see Michel (1976).

In recent years, there has been considerable interest in obtaining results that are similar to (3) and (4) for U-statistics. The Berry-Esseen inequality for U-statistics has been investigated, for instance, by Grams and Serfling (1973), Bickel (1974) and Chan and Wierman (1977). A sharper Berry-Esseen bound was given by Callaert and Janssen (1978), which states that

$$\left| P\left(\sqrt{n}U_n/(2\sigma_g) \le x \right) - \Phi(x) \right| \le A n^{-1/2} \sigma_g^{-3} E|h(X_1, X_2)|^3$$

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

A non-uniform Berry-Esseen bound for U-statistic was given by Zhao and Chen (1982), who shows that if $E|h(X_1, X_2)|^3 < \infty$, then

$$(1 + |x|)^3 \left| P\left(\sqrt{n}U_n/(2\sigma_g) \le x \right) - \Phi(x) \right| = O(n^{-1/2}). \quad (5)$$

To my knowledge, this is the only known result in this direction. It remains an open and more challenging question whether a result similar to (4) holds for U-statistics. As mentioned before, this question is more interesting because the bound reflects dependence on x and n as well as on a moment condition. The answer to this question is affirmative, as the following theorem shows.

Theorem 1.1. *Assume that $E|h(X_1, X_2)|^p < \infty$, $p \ge 3$. Then for any $n \ge 2 + (E|g(X_1)|^3/\sigma_g^3)^3$,*

$$\left| P\left(\sqrt{n}U_n/(2\sigma_g) \le x \right) - \Phi(x) \right|$$
$$\le A(p)(1+|x|)^{-p} n^{-1/2} \left\{ \frac{E|g(X_1)|^p}{\sigma_g^p} + \frac{E|h(X_1, X_2)|^2}{\sigma_g^2} + \frac{n^{-(p-1)/2} E|h(X_1, X_2)|^p}{\sigma_g^p} \right\}, \quad (6)$$

uniformly in $x \in R$, where $A(p) > 0$ is a constant depending only on p.

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

$$T_n = n^{-1/2} \sum_{j=1}^{n} \alpha(X_j) + D_n \sum_{i \neq j} \beta(X_i, X_j) + V_n, \tag{7}$$

where D_n is a real number depending only on n. In terms of (2), Theorem 1.1 is an easy corollary of the following theorem.

Theorem 1.2. *Assume that*
(a) $E\alpha(X_1) = 0$, $E\alpha^2(X_1) = 1$. $E[\beta(X_1, X_2)|X_i] = 0$, $i = 1, 2$.
(b) $|D_n| \leq An^{-3/2}$ *for some constant* $A > 0$.
(c) $P(|V_n| \geq C_0(1 + |x|)n^{-1/2}) \leq C_1(1 + |x|)^{-p}n^{-1/2}$ *for some constants* $C_0 > 0$ *and* $C_1 > 0$.
Then for all $p \geq 3$ and $n \geq 2 + \rho_3^3$,

$$\left| P\left(T_n \leq x\right) - \Phi(x) \right| \leq A(p)(1 + |x|)^{-p}n^{-1/2}\left\{C_0 + C_1 + \mathcal{L}\right\}, \tag{8}$$

uniformly in $x \in R$, where $A(p)$ is a constant depending only on p; $\rho_s = E|\alpha(X_1)|^s$, $\lambda_s = E|\beta(X_1, X_2)|^s$ and $\mathcal{L} = \rho_p + \lambda_2 + n^{-(p-1)/2}\lambda_p$.

We remark that Theorem 1.2 is quite general. Consider its 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 $T(G) = \int xJ(G(x))\, dG(x)$. The statistic $T(F_n)$ is called an L-statistic (see Chapter 8 of Serfling (1980)). Write

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

and define the distributions of the standardized L-statistic $T(F_n)$ by $H_n(x) = P(\sqrt{n}\sigma^{-1}(T(F_n) - T(F)) \leq x)$. It is well-known that $H_n(x)$ converges to the standard normal distribution function $\Phi(x)$ provided $E|X_1|^2 < \infty$ and $\sigma^2 > 0$, along with some smoothness conditions on $J(t)$ (see Serfling (1980) and Helmers, Janssen and Serfling (1990) for references). The uniform rates of convergence to normality have also been studied by various authors. For instance, assuming that $E|X_1|^3 < \infty$ and $\sigma^2 > 0$, and some smoothness conditions on $J(t)$, Helmers (1977) and Helmers, Janssen and Serfling (1990) showed that $\sup_{x \in R} |H_n(x) - \Phi(x)| = O(n^{-1/2})$.

The following theorem gives a new non-uniform Berry-Esseen bound for standard L-statistics.

Theorem 1.3. *Assume that*
(a) $\left|J(s) - J(t)\right| \leq K|s - t|, 0 < s < t < 1$, *for some* $K > 0$, *so that* $J(t)$ *is bounded,* $|J(t)| \leq M < \infty$, *say;*
(b) $E|X_1|^p < \infty$, *for some* $p \geq 3$ *and* $\sigma^2 > 0$.
Then, for all $n \geq 2 + (E|\alpha(X_1)|^3)^3$ *with* $\alpha(X_1) = -\sigma^{-1} \int J(F(t))(I(X_1 \leq t) - F(t))dt$,

$$|H_n(x) - \Phi(x)| \leq A(J, p, K, M)(1 + |x|)^{-p}n^{-1/2}\sigma^{-p}E|X_1|^p,$$

uniformly in $x \in R$, *where* $A(J, p, K, M)$ *is a positive constant only depending on* J, p, K *and* M.

The proofs of Theorems 1.1−3 will be given in the next section. For convenience, throughout this paper, we denote by A, A_1, ... absolute positive constants, which may be different at each occurrence. If a constant A depends on a parameter, say u, then we write $A(u)$. Furthermore, we introduce the following notation for ease of presentation: $\sum_{i<j} \equiv \sum_{1 \leq i < j \leq n}, \sum_{i \neq j} \equiv \sum_{\substack{i,j=1 \\ i \neq j}}^{n}$.

2. Proofs of Main Results

Proof of Theorem 1.1. In view of (2), Theorem 1.1 is an immediate corollary of Theorem 1.2. The details are omitted.

Proof of Theorem 1.2. It should be pointed out that the proof of (5) given by Zhao and Chen (1982) is very technical and complex, and therefore it is hard to extend the proof to the case for $p \geq 3$. Here we give a totally new and simple proof for the non-uniform Berry-Esseen bound for U-statistics.

Without loss of generality, assume that $x \geq 0$. For simplicity, we further assume $V_n = 0$ and $D_n = n^{-3/2}$. As shown in Wang, Jing and Zhao (2000), these assumptions will not affect the proof of the main results.

The proof of (8) breaks up into three parts: $0 \leq x \leq 1$, $1 \leq x^2 \leq 8 \log n$, and $x^2 \geq 8 \log n$.

In fact, for $0 \leq x \leq 1$, (8) is a direct corollary of Theorem 2.1 given by Wang, Jing and Zhao (2000), by noting that $1 \leq \rho_3 \leq \rho_p$ and $\lambda_{5/3} \leq (\lambda_2)^{5/6} \leq 1 + \lambda_2$.

If $x^2 \geq 8 \log n$, it can be easily shown that

$$1 - \Phi(x/2) \leq \frac{1}{\sqrt{2\pi}}e^{-x^2/8} \leq A(1 + x)^{-p}n^{-1/2}.$$

Hence, from the classical non-uniform Berry-Esseen bound for sums of independent random variables [cf. Michel (1976)],

$$P\Big(n^{-1/2}\sum_{j=1}^{n}\alpha(X_j) \geq x/2\Big) \leq 1 - \Phi(x/2) + A(1+x)^{-p}n^{-1/2}\rho_p$$

$$\leq A_1(p)(1+x)^{-p}n^{-1/2}\rho_p.$$

On the other hand, by noting that $\sum_{i\neq j}\beta(X_i, X_j) = \sum_{i<j}[\beta(X_i,X_j)+\beta(X_j,X_i)]$ is a degenerate U-statistic, it follows from the moment inequality for degenerate U-statistics [cf. Wang (1998)] that $P\Big(n^{-3/2}\sum_{i\neq j}\beta(X_i,X_j) \geq x/2\Big) \leq A(p)(1+x)^{-p}n^{-p/2}\lambda_p$. By using these estimates, for $x^2 \geq 8\log n$, (8) follows from

$$\Big|P\Big(T_n \leq x\Big) - \Phi(x)\Big| = \Big|P\Big(T_n > x\Big) - (1 - \Phi(x))\Big|$$

$$\leq (1 - \Phi(x)) + P\Big(n^{-1/2}\sum_{j=1}^{n}\alpha(X_j) \geq x/2\Big) + P\Big(n^{-3/2}\sum_{i\neq j}\beta(X_i,X_j) \geq x/2\Big)$$

$$\leq A(p)(1+x)^{-p}n^{-1/2}\Big\{\rho_p + n^{-(p-1)/2}\lambda_p\Big\}.$$

So it remains to show that if $1 \leq x^2 \leq 8\log n$ and $n \geq 2 + \rho_3^3$, then

$$\Delta_n(x) \equiv \Big|P\Big(n^{-1/2}\sum_{j=1}^{n}\alpha(X_j) + n^{-3/2}\sum_{i\neq j}\beta(X_i,X_j) \leq x\Big) - \Phi(x)\Big|$$

$$\leq A(p)(1+x)^{-p}n^{-1/2}(\rho_p + \lambda_2). \tag{9}$$

Let $\alpha_j = \alpha(X_j)$ and $\eta_{ij} = \beta(X_i,X_j) + \beta(X_j,X_i)$. As mentioned before, we can rewrite $\sum_{i\neq j}\beta(X_i,X_j) = \sum_{i<j}\eta_{ij}$ with $E\Big(\eta_{12} \mid X_1\Big) = 0$ and $E\eta_{12}^2 \leq 4\lambda_2$. For the rest of this section, we use the following notations: $i = \sqrt{-1}$,

$$g(t) = Ee^{it\alpha_1/\sqrt{n}}, \quad S_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\alpha_j, \quad \Lambda_{n,m} = \frac{1}{n^{3/2}}\sum_{k=1}^{m-1}\sum_{j=k+1}^{n}\eta_{kj},$$

$$f_n(t) = \Big\{1 + n(g(t)-1) + \frac{t^2}{2}\Big\}e^{-t^2/2}, \quad \varphi_n(t) = \frac{(it)^3 e^{-t^2/2}}{\sqrt{n}}E\alpha_1\alpha_2\eta_{12}.$$

The proof of (9) is based on the following lemmas.

Lemma 2.1. If $|t| \leq \sqrt{n}/(4\rho_3)$, we have

$$|g(t)| \leq e^{-t^2/3n}, \qquad \Big|g^n(t) - e^{-\frac{t^2}{2}}\Big| \leq 16n^{-1/2}\rho_3|t|^3 e^{-t^2/3}. \tag{10}$$

If in addition $|t| \leq (\sqrt{n}/\rho_3)^{1/3}$, then

$$|g^n(t) - f_n(t)| \leq An^{-1}(\rho_3)^2 t^4 e^{-t^2/6}. \tag{11}$$

This lemma is well-known. See, for example, Hall (1982).

Lemma 2.2. *If $|t| \leq \sqrt{n}/(4\rho_3)$ and $n \geq \max\{2, \rho_3^2\}$, then*

$$\left| E e^{it(S_n + \Lambda_{n,n})} - e^{-t^2/2} \right| \leq A n^{-1/2} (\lambda_2^{1/2} + \rho_3)(t^2 + t^6)e^{-t^2/3} + 16 n^{-1} \lambda_2 t^2, \quad (12)$$

and for any $2 \leq m \leq n$,

$$\left| E e^{it(S_n + \Lambda_{n,n})} \right| \leq \left(1 + \frac{2m\lambda_2^{1/2}|t|}{\sqrt{n}} \right) e^{-\frac{(m-2)t^2}{3n}} + 16 m n^{-2} \lambda_2 t^2. \quad (13)$$

If in addition $|t| \leq (\sqrt{n}/\rho_3)^{1/3}$, then

$$\left| E e^{it(S_n + \Lambda_{n,n})} - f_n(t) - \varphi_n(t) \right| \leq A n^{-3/4} (\lambda_2 \rho_3)^{1/2}(t^2 + t^8)e^{-t^2/3} + 16 n^{-1} \lambda_2 t^2. \quad (14)$$

Proof. We first prove that for any $|t| \leq \sqrt{n}/(4\rho_3)$ and $n \geq \max\{2, \rho_3^2\}$,

$$\left| E \Lambda_{n,n} e^{itS_n} + \frac{E\alpha_1 \alpha_2 \eta_{12}}{\sqrt{n}} t^2 e^{-\frac{t^2}{2}} \right| \leq A n^{-3/4} (\lambda_2 \rho_3)^{1/2}(t^2 + t^6)e^{-t^2/3}. \quad (15)$$

Recalling $E(\eta_{12}|X_1) = 0$, it can be easily shown that

$$E \eta_{12} e^{it(\alpha_1 + \alpha_2)/\sqrt{n}} = -\frac{t^2}{n} E\alpha_1 \alpha_2 \eta_{12} + K_{1n} + K_{2n}, \quad (16)$$

where $K_{1n} = E\eta_{12} \left(e^{it\alpha_1/\sqrt{n}} - 1 - \frac{it\alpha_1}{\sqrt{n}} \right) \left(e^{it\alpha_2/\sqrt{n}} - 1 \right)$ and $K_{2n} = \frac{it}{\sqrt{n}} E\alpha_1 \eta_{12}$ $\left(e^{it\alpha_2/\sqrt{n}} - 1 - \frac{it\alpha_2}{\sqrt{n}} \right)$.

By $|e^{ix} - 1 - ix| \leq 2|x|^{3/2}$ and Holder's inequality, we obtain that

$$|K_{1n}| + |K_{2n}| \leq 2n^{-5/4}|t|^{5/2} \left(E|\eta_{12}||\alpha_1|^{3/2}|\alpha_2| + E|\eta_{12}||\alpha_1||\alpha_2|^{3/2} \right)$$
$$\leq 8n^{-5/4}(\lambda_2 \rho_3)^{1/2}(t^2 + t^4).$$

Therefore, it follows from independence of the α_j, (10) and (16) that

$$\begin{aligned} E\Lambda_{n,n} e^{itS_n} &= \frac{1}{n^{3/2}} \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \eta_{kj} e^{itS_n} \\ &= \frac{n-1}{\sqrt{n}} g^{n-2}(t) E\eta_{12} e^{it(\alpha_1 + \alpha_2)/\sqrt{n}} \\ &= -\frac{(n-1)E\alpha_1 \alpha_2 \eta_{12}}{n^{3/2}} t^2 g^{n-2}(t) + \frac{(n-1)g^{n-2}(t)}{\sqrt{n}} t^2 (K_{1n} + K_{2n}) \\ &= -\frac{E\alpha_1 \alpha_2 \eta_{12}}{\sqrt{n}} t^2 e^{-\frac{t^2}{2}} + K_{3n}, \end{aligned}$$

where simple calculation shows that (recall $\rho_3 \geq 1$ and $1 \geq (\rho_3^2/n)^{1/4}$)

$$
\begin{aligned}
|K_{3n}| &\leq \sqrt{n} t^2 |g(t)|^{n-2} \Big(|K_{1n}| + |K_{2n}| + n^{-3/2}(\lambda_2)^{1/2} \Big) + \frac{t^2 (\lambda_2)^{1/2}}{\sqrt{n}} \Big| g^{n-2}(t) - e^{-\frac{t^2}{2}} \Big| \\
&\leq A(t^2 + t^6) e^{-t^2/3} \Big(n^{-3/4}(\lambda_2 \rho_3)^{1/2} + n^{-1} \lambda_2^{1/2} \rho_3 \Big) \\
&\leq A n^{-3/4}(\lambda_2 \rho_3)^{1/2}(t^2 + t^6) e^{-t^2/3}.
\end{aligned}
$$

This proves (15).

Let us turn back to the proofs of (12)−(14). Put $\Lambda_{n,m}^* = \Lambda_{n,n} - \Lambda_{n,m} = \frac{1}{n^{3/2}} \sum_{k=m+1}^{n-1} \times \sum_{j=k+1}^{n} \eta_{kj}$ and $\Lambda_{n,m}^* = 0$, if $m \geq n$. By $|e^{iz} - 1 - iz| \leq 2|z|^2$, we have that

$$
\Big| E e^{it(S_n + \Lambda_{n,n})} - E e^{it(S_n + \Delta_{n,m}^*)} - it E \Lambda_{n,m} e^{it(S_n + \Lambda_{n,m}^*)} \Big| \leq 2t^2 E |\Lambda_{n,m}|^2 \leq 16 m n^{-2} \lambda_2 t^2
\tag{17}
$$

By letting $m = n$ in (17), the proofs of (12) and (14) follow easily from (10), (11) and (15) respectively. In view of independence of the α_j, on the other hand, we have that

$$
\begin{aligned}
E |\Lambda_{n,m} e^{it(S_n + \Lambda_{n,m}^*)}| &= E \Big| \frac{1}{n^{3/2}} \sum_{k=1}^{m-1} \sum_{j=k+1}^{n} \eta_{kj} e^{it(S_n + \Lambda_{n,m}^*)} \Big| \\
&\leq m E |\eta_{12}| |g(t)|^{m-2}/\sqrt{n} \leq 2m \lambda_2^{1/2} |g(t)|^{m-2}/\sqrt{n}.
\end{aligned}
$$

This, together with (10) and (17), implies that for any $2 \leq m \leq n$,

$$
\begin{aligned}
\Big| E e^{it(S_n + \Lambda_{n,n})} \Big| &\leq |g(t)|^m + \frac{2m \lambda_2^{1/2} |t|}{\sqrt{n}} |g(t)|^{m-2} + 16 m n^{-2} \lambda_2 t^2 \\
&\leq \Big(1 + \frac{2m \lambda_2^{1/2} |t|}{\sqrt{n}} \Big) e^{-\frac{(m-2)t^2}{3n}} + 16 m n^{-2} \lambda_2 t^2.
\end{aligned}
$$

This provides (13). The proof of Lemma 2.2. is complete.

Lemma 2.3. *Let F be a distribution function with characteristic function f. Then for all $y \in R$ and $T > 0$ it holds that*

$$
\lim_{z \downarrow y} F(z) \leq \frac{1}{2} + V.P. \int_{-T}^{T} exp(-iyt) \frac{1}{T} K\Big(\frac{t}{T}\Big) f(t) dt,
\tag{18}
$$

$$
\lim_{z \uparrow y} F(z) \geq \frac{1}{2} - V.P. \int_{-T}^{T} exp(-iyt) \frac{1}{T} K\Big(-\frac{t}{T}\Big) f(t) dt,
\tag{19}
$$

where $V.P. \int_{-T}^{T} = \lim_{h \downarrow 0} \Big(\int_{-T}^{-h} + \int_{h}^{T} \Big)$, and $2K(s) = K_1(s) + iK_2(s)/(\pi s)$, $K_1(s) = 1 - |s|$, $K_2(s) = \pi s(1 - |s|) \cot \pi s + |s|$, for $|s| < 1$, and $K(s) \equiv 0$ for $|s| \geq 1$.

The proof of Lemma 2.3 can be found in Prawitz (1972).

Lemma 2.4. *It holds that for any $y \in R$ and $n \geq 2 + n_0$,*

$$|I^+|, \quad |I^-| \leq A n^{-1/2} \rho_3 e^{-y^2/2} + A_1 n^{-2/3}(\lambda_2 + \rho_3^{4/3}), \tag{20}$$

where $n_0 = \max\{k : 6 \log k \geq (\sqrt{k}/\rho_3)^{2/3}\}$, $K_1(s)$ is defined as in Lemma 2.3,

$$I^+ = \frac{1}{T} \int_{-T}^{T} e^{-iyt} K_1\Big(\frac{t}{T}\Big) E e^{it(S_n + \Lambda_{n,n})} dt,$$

$$I^- = \frac{1}{T} \int_{-T}^{T} e^{-iyt} K_1\Big(-\frac{t}{T}\Big) E e^{it(S_n + \Lambda_{n,n})} dt, \qquad T = \sqrt{n}/(4\rho_3).$$

Proof. We only prove (20) for $|I^+|$. We first note that $n \geq \max\{2, \rho_3^2\}$ when $n \geq 2 + n_0$. Let $T_1 = (\sqrt{n}/\rho_3)^{1/3}$. Rewrite $I^+ = I_1 + I_2$, where

$$I_1 = \frac{1}{T} \int_{-T_1}^{T_1} e^{-iyt} K_1\Big(\frac{t}{T}\Big) E e^{it(S_n + \Lambda_{n,n})} dt,$$

$$I_2 = \frac{1}{T} \int_{T_1 \leq |t| \leq T} e^{-iyt} K_1\Big(\frac{t}{T}\Big) E e^{it(S_n + \Lambda_{n,n})} dt.$$

It is easy to see that $\left[\frac{6n \log n}{t^2}\right] \leq n - 2$ if $|t| \geq T_1$. Hence, by (13) with $m = \left[\frac{6n \log n}{t^2}\right] + 2$,

$$|I_2| \leq \frac{1}{T} \int_{T_1 \leq |t| \leq T} |E e^{it(S_n + \Lambda_{n,n})}| dt \leq A(\rho_3 + \lambda_2) n^{-2/3}. \tag{21}$$

Noting $K_1(s) = 1 - |s|$, for $|s| < 1$, we obtain $|I_1| \leq |I_{11}| + |I_{12}|$, where

$$I_{11} = \frac{1}{T} \int_{-T_1}^{T_1} e^{-iyt} E e^{it(S_n + \Lambda_{n,n})} dt, \qquad I_{12} = \frac{2}{T^2} \int_0^{T_1} t |E e^{it(S_n + \Lambda_{n,n})}| dt.$$

It is obvious that $|I_{12}| \leq \frac{2}{T^2} \int_0^{T_1} t\, dt \leq 8 n^{-2/3} \rho_3^{4/3}$. Noting that $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-iyt - t^2/2} dt = e^{-y^2/2}$, it follows from (12) that

$$|I_{11}| \leq \frac{1}{T}\Big| \int_{-\infty}^{\infty} e^{-iyt - t^2/2} dt \Big| + \frac{1}{T} \int_{|t| \geq T_1} e^{-t^2/2} dt + \frac{1}{T} \int_{-T_1}^{T_1} \Big| E e^{it(S_n + \Lambda_{n,n})} - e^{-t^2/2} \Big| dt$$

$$\leq A n^{-1/2} \rho_3 e^{-y^2/2} + A_1 n^{-1}(\rho_3^2 + \lambda_2).$$

Collecting all these estimates, we conclude the proof of Lemma 2.4.

Lemma 2.5. *The integrals*

$$J^+ \equiv \frac{i}{\pi} V.P. \int_{-T}^{T} e^{-iyt} K_2\Big(\frac{t}{T}\Big) E e^{it(S_n + \Lambda_{n,n})} \frac{dt}{t},$$

$$J^- \equiv \frac{i}{\pi} V.P. \int_{-T}^{T} e^{-iyt} K_2\Big(-\frac{t}{T}\Big) E e^{it(S_n + \Lambda_{n,n})} \frac{dt}{t}, \qquad T = \sqrt{n}/(4\rho_3),$$

satisfy: for any $y \in R$ and $n \geq 2 + n_0$,

$$\left| J^+ + 1 - 2\Phi(y) - 2\mathcal{L}_n(y) - 2\mathcal{L}_{1n}(y) \right| \leq An^{-2/3}(\lambda_2 + \rho_3^{4/3}), \qquad (22)$$

$$\left| J^- + 1 - 2\Phi(y) - 2\mathcal{L}_n(y) - 2\mathcal{L}_{1n}(y) \right| \leq An^{-2/3}(\lambda_2 + \rho_3^{4/3}), \qquad (23)$$

where $n_0 = \max\{k : 6\log k \geq (\sqrt{k}/\rho_3)^{2/3}\}$, $K_2(s)$ is defined as in Lemma 2.3, $\mathcal{L}_n(y) = n\left\{E\Phi\left(y - \frac{\alpha_1}{\sqrt{n}}\right) - \Phi(y)\right\} - \frac{1}{2}\Phi^{(2)}(y)$ and $\mathcal{L}_{1n}(y) = \frac{E\alpha_1\alpha_2\eta_{12}}{\sqrt{n}}\Phi^{(3)}(y)$.

Proof. We only prove (22). We can write $J^+ = J_{11} + J_{12} + J_{13} + J_2$, where

$$J_{11} = \frac{i}{\pi}V.P.\int_{-T_1}^{T_1} e^{-iyt}\left(f_n(t) + \varphi_n(t)\right)\frac{dt}{t},$$

$$J_{12} = \frac{i}{\pi}V.P.\int_{-T_1}^{T_1} e^{-iyt}\left(Ee^{it(S_n+\Lambda_{n,n})} - f_n(t) - \varphi_n(t)\right)\frac{dt}{t},$$

$$J_{13} = \frac{i}{\pi}V.P.\int_{-T_1}^{T_1} e^{-iyt}\left(K_2(\frac{t}{T}) - 1\right)Ee^{it(S_n+\Lambda_{n,n})}\frac{dt}{t},$$

$$J_2 = \frac{i}{\pi}V.P.\int_{T_1 \leq |t| \leq T} e^{-iyt}K_2(\frac{t}{T})Ee^{it(S_n+\Lambda_{n,n})}\frac{dt}{t}$$

and $T_1 = (\sqrt{n}/\rho_3)^{1/3}$. Similar to (21), it follows that $|J_2| \leq A(1 + \lambda_2)n^{-2/3}$. By using (14), we have

$$|J_{12}| \leq \int_{-T_1}^{T_1}\left|Ee^{it(S_n+\Lambda_{n,n})} - f_n(t) - \varphi_n(t)\right|\frac{dt}{|t|}$$

$$\leq An^{-2/3}\lambda_2 + A_1 n^{-3/4}(\lambda_2\rho_3)^{1/2} \leq An^{-2/3}(\lambda_2 + \rho_3).$$

Noting that $|K_2(s) - 1| \leq As^2$, for $|s| \leq 1/2$ (cf., e.g., Lemma 2.1 in Bentkus (1994)), it can be easily shown that

$$|J_{13}| \leq AT^{-2}\int_{-T_1}^{T_1}|t|\left|Ee^{it(S_n+\Lambda_{n,n})}\right|dt \leq AT^{-2}\int_{-T_1}^{T_1}|t|dt \leq An^{-2/3}\rho_3^{4/3}.$$

On the other hand, simple calculation shows that

$$\frac{i}{2\pi}V.P.\int_{-\infty}^{\infty} e^{-iyt}\left(f_n(t) + \varphi_n(t)\right)\frac{dt}{t} = -\frac{1}{2} + \Phi(y) + \mathcal{L}_n(y) + \mathcal{L}_{1n}(y).$$

Therefore, it follows from all these estimates that (recall $\rho_3 \geq 1$)

$$\left| J^+ + 1 - 2\Phi(y) - 2\mathcal{L}_n(y) - 2\mathcal{L}_{1n}(y) \right|$$

$$\leq \left| J_{11} + 1 - 2\Phi(y) - 2\mathcal{L}_n(y) - 2\mathcal{L}_{1n}(y) \right| + |J_{12}| + |J_{13}| + |J_2|$$

$$\leq \int_{|t| \geq T_1}\frac{1}{|t|}\left|f_n(t) + \varphi_n(t)\right|dt + An^{-2/3}(\lambda_2 + \rho_3^{4/3})$$

$$\leq A_1 n^{-2/3}(\lambda_2 + \rho_3^{4/3}).$$

This also completes the proof of Lemma 2.5.

We are now ready to prove (9). It suffices to show that for $1 \leq x^2 \leq 8 \log n$ and $n \geq 2 + \rho_3^3$,

$$P(S_n + \Lambda_{n,n} \leq x) \leq \Phi(x) + A(p)(1+x)^{-p}n^{-1/2}(\rho_p + \lambda_2), \qquad (24)$$

$$P(S_n + \Lambda_{n,n} \geq x) \leq 1 - \Phi(x) + A(p)(1+x)^{-p}n^{-1/2}(\rho_p + \lambda_2). \qquad (25)$$

We first prove (24). We may assume that $\rho_3^3 \geq \max\{n_1, n_2\}$, where $n_1 = \max\{k : (6 \log k)^9 \geq k\}$ and $n_2 = \max\{k : (\log k)^{9p} \geq k\}$. Otherwise, we have $x^2 \leq 8 \log \left(2 + \max\{n_1, n_2\}\right)$, where $\max\{n_1, n_2\}$ is a constant depending only on p. The result follows immediately from Wang, Jing and Zhao (2000). Taking account of $n \geq 2 + \rho_3^3$ and $\rho_3^3 \geq \max\{n_1, n_2\}$, it is easy to see that $n \geq 2 + n_0$, where $n_0 = \max\{k : 6 \log k \geq (\sqrt{k}/\rho_3)^{2/3}\}$ as defined in Lemmas 2.4 and 2.5, and for $x^2 \leq 8 \log n$,

$$\begin{aligned}
n^{-2/3}(\lambda_2 + \rho_3^{4/3}) &\leq (1+x)^{-p}n^{-2/3}\rho_3^{1/3}(\lambda_2 + \rho_3)(1 + 3 \log^{1/2} n)^p \\
&\leq (1+x)^{-p}n^{-1/2}(\lambda_2 + \rho_3)(1 + 3 \log^{1/2} n)^p n^{-1/18} \\
&\leq A(p)(1+x)^{-p}n^{-1/2}(\lambda_2 + \rho_3).
\end{aligned}$$

Hence, by using (18) in Lemma 2.3 with $y = x$ and $T = \sqrt{n}/(4\rho_3)$, and then Lemmas 2.4−2.5, we have for $1 \leq x^2 \leq 8 \log n$ and $n \geq 2 + \rho_3^3$

$$\begin{aligned}
P(S_n + \Lambda_{n,n} \leq x) &\leq \frac{1}{2}(|I^+| + |J^+ + 1|) \\
&\leq \Phi(x) + |\mathcal{L}_n(x)| + |\mathcal{L}_{1n}(x)| + + A(p)(1+x)^{-p}n^{-1/2}(\rho_3 + \lambda_2).
\end{aligned}$$

Recalling $\rho_3 \geq 1$ and $\rho_p \geq \rho_3$, it is obvious that $|\mathcal{L}_{1n}(x)| \leq n^{-1/2}\lambda_2^{1/2}|\Phi^{(3)}(x)| \leq A(p)(1+x)^{-p}n^{-1/2}(\rho_3 + \lambda_2)$. So, to prove (24), it remains to show

$$|\mathcal{L}_n(x)| \leq A(p)(1+x)^{-p}n^{-1/2}\rho_p. \qquad (26)$$

Let $\alpha_1^* = \alpha_1 I\{|\alpha_1| \leq \sqrt{n}(1+x)/8\}$ and $\mathcal{L}_{jn}^*(x) = E\Phi\left(x - \frac{\alpha_1^*}{\sqrt{n}}\right) - \Phi(x) - \frac{E\alpha_1^*}{\sqrt{n}}\Phi^{(1)}(x) - \frac{E\alpha_1^{*2}}{2n}\Phi^{(2)}(x)$. By using a Taylor expansion of $\Phi(x)$, we have

$$\begin{aligned}
|\mathcal{L}_{jn}^*(x)| &\leq \frac{1}{6n^{3/2}}E|\alpha_1^*|^3\Phi^{(3)}\left(x + \theta\frac{|\alpha_1^*|}{\sqrt{n}}\right) \qquad \text{(where } |\theta| \leq 1) \\
&\leq \frac{A(x^2+1)e^{-x^2/2}}{n^{3/2}}E|\alpha_1^*|^3 \exp\left(x|\alpha_1^*|/\sqrt{n}\right) \\
&\leq An^{-3/2}\rho_3(x^2+1)e^{-x^2/4}
\end{aligned}$$

Hence, it follows that

$$|\mathcal{L}_n(x)| \le \left|\mathcal{L}_n(x) - \sum_{j=1}^{n}\mathcal{L}_{jn}^*(x)\right| + \sum_{j=1}^{n}|\mathcal{L}_{jn}^*(x)|$$

$$\le n\left|E\Phi\left(x - \frac{\alpha_1^*}{\sqrt{n}}\right) - E\Phi\left(x - \frac{\alpha_1}{\sqrt{n}}\right)\right| + An^{-1/2}\rho_3(x^2+1)e^{-x^2/4}$$

$$\le 2nP(|\alpha_1| \ge \sqrt{n}(1+x)/8) + An^{-1/2}\rho_3(x^2+1)e^{-x^2/4}$$

$$\le A(p)(1+x)^{-p}n^{-1/2}\rho_p.$$

This proves (26) and hence (24).

Now we prove (25). Similar to the proof of (24), by using (19) in Lemma 2.3 with $y = x$ and $T = \sqrt{n}/(4\rho_3)$, and then Lemmas 4-5, we have for $1 \le x^2 \le 8\log n$ and $n \ge 2 + \rho_3^3$

$$P(S_n + \Lambda_{n,n} \ge x) \le \frac{1}{2}(|I^-| + |J^- - 1|)$$

$$\le 1 - \Phi(x) + |\mathcal{L}_n(x)| + |\mathcal{L}_{1n}(x)| + + A(p)(1+x)^{-p}n^{-1/2}(\rho_3 + \lambda_2)$$

$$\le 1 - \Phi(x) + + A(p)(1+x)^{-p}n^{-1/2}(\rho_p + \lambda_2).$$

This proves (25). The proof of Theorem 1.2 is now complete.

Proof of Theorem 1.3. As in Serfling (1980, p.265), we have

$$T(F_n) - T(F) = -\int [\psi(F_n(x)) - \psi(F(x))]dx, \tag{27}$$

where $\psi(t) = \int_0^t J(u)du$. The Lipschitz condition on J implies that

$$\left|\psi(F_n(x)) - \psi(F(x)) - (F_n(x) - F(x))J(F(x))\right| \le K(F_n(x) - F(x))^2. \tag{28}$$

Write $\eta_j(t) = I\{X_j \le t\} - F(t)$. It follows from (27) and (28) that

$$n^{-1/2}\sum_{j=1}^{n}\alpha(X_j) - n^{-3/2}\sum_{i \ne j}\beta(X_i, X_j) - V_n$$

$$\le \sqrt{n}(T(F_n) - T(F))/\sigma$$

$$\le n^{-1/2}\sum_{j=1}^{n}\alpha(X_j) + n^{-3/2}\sum_{i \ne j}\beta(X_i, X_j) + V_n, \tag{29}$$

where $\alpha(X_j) = -\sigma^{-1}\int J(F(t))\eta_j(t)dt$, $\beta(X_i, X_j) = K\sigma^{-1}\int \eta_i(t)\eta_j(t)dt$, $V_n = n^{-3/2}\sum_{j=1}^{n}Z(X_j)$ with $Z(X_j) = K\sigma^{-1}\int \eta_j^2(t)dt$. It can be easily shown that

$E\alpha(X_1) = 0, E\alpha^2(X_1) = 1$ and $E(\beta(X_i, X_j)\big|X_i) = 0$, $i \neq j$. For the rest of this section, we prove that

$$E|\alpha(X_1)|^p + E|\beta(X_1, X_2)|^p + E|Z(X_1)|^p \leq A(J,K)\sigma^{-p}E|X_1|^p, \qquad (30)$$

$$P(|V_n| > (1+|x|)n^{-1/2}(1+E|Z(X_1)|)) \leq A(J,K,p)(1+|x|)^{-p}n^{-1/2}\sigma^{-p}E|X_1|^p. \qquad (31)$$

Theorem 1.3 will then follow from Theorem 1.2.

In fact, similar to the proof of Lemma A in Serfling (1980, p.288), we can show that

$$|\alpha(X_j)| + |\beta(X_i, X_j)| + Z(X_j) \leq A(J,K)\sigma^{-1}(|X_j| + E|X_1|), \qquad (32)$$

Noting $E\alpha^2(X_1) = 1$, (30) follows easily from (32). For (31), we have that,

$$P(|V_n| \geq (1+|x|)n^{-1/2}(1+E|Z(X_1)|)) \leq P\Big(\Big|\sum_{j=1}^{n}(Z(X_j) - EZ(X_j))\Big| \geq n(1+|x|)\Big)$$

$$\leq A(J,K,p)(1+|x|)^{-p}n^{-1/2}\sigma^{-p}E|X_1|^p.$$

The proof of Theorem 1.3 is now complete.

Acknowledgements

References

Bentkus, V. (1994). On the asymptotical behavior of the constant in the Berry-Esseen inequality. *J. Theoret. Probab.* **7**, 211-224.

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.

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

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

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

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

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.

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

Lee, A. J. (1990). *U-statistics. – Theorem and Practice.* Statistics Textbooks Monographs **110**.

Michel, R. (1976). Nonuniform central limit bounds with applications to probabilities of deviations. *Ann. Probab.* **4**. 102-106.

Prawitz, H. (1972). Limits for a distribution, if the characteristic function is given in a finite domain. *Skand. Aktuar. Tidskr.* 138-154.

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

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

Wang, Q. (1998). Bernstein type inequalities for degenerate U-statistics with applications, *Chinese Ann. Math. Ser. B* **2**, 157-166.

Zhao, L. and Chen, X. (1982). Non-uniform Berry-Esseen bound for U-statistics, *Sci. China Ser. A* **12**, 1066-1078.

Centre for Mathematics and its Applications, School of Mathematical Science, The Australian National University, Canberra, ACT 0200, Australia.

E-mail: qiying@maths.anu.edu.au