

Normal approximation for nonlinear statistics using a concentration inequality approach

LOUIS H. Y. CHEN¹ and QI-MAN SHAO²

¹*Institute for Mathematical Sciences, National University of Singapore, Singapore 118402, Republic of Singapore. E-mail: matchyl@nus.edu.sg*

²*Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China; Department of Mathematics, University of Oregon, Eugene, OR 97403, USA; Department of Mathematics, Zhejiang University, Hangzhou, Zhejiang 310027, China. E-mail: maqmshao@ust.hk*

Let T be a general sampling statistic that can be written as a linear statistic plus an error term. Uniform and non-uniform Berry–Esseen type bounds for T are obtained. The bounds are the best possible for many known statistics. Applications to U-statistics, multisample U-statistics, L-statistics, random sums and functions of nonlinear statistics are discussed.

Keywords: concentration inequality approach; L-statistics; multisample U-statistics; non-uniform Berry–Esseen bound; nonlinear statistics; normal approximation; U-statistics; uniform Berry–Esseen bound

1. Introduction

Let X_1, X_2, \dots, X_n be independent random variables and let $T := T(X_1, \dots, X_n)$ be a general sampling statistic. In many cases T can be written as a linear statistic plus an error term, say $T = W + \Delta$, where

$$W = \sum_{i=1}^n g_i(X_i), \quad \Delta := \Delta(X_1, \dots, X_n) = T - W$$

and $g_i := g_{n,i}$ are Borel measurable functions. Typical cases include U-statistics, multisample U-statistics, L-statistics and random sums. Assume that

$$E(g_i(X_i)) = 0 \quad \text{for } i = 1, 2, \dots, n \quad \text{and} \quad \sum_{i=1}^n E(g_i^2(X_i)) = 1. \quad (1.1)$$

It is clear that if $\Delta \rightarrow 0$ in probability as $n \rightarrow \infty$, then we have the central limit theorem

$$\sup_z |P(T \leq z) - \Phi(z)| \rightarrow 0, \tag{1.2}$$

where Φ denotes the standard normal distribution function, provided that the Lindeberg condition holds:

$$\forall \varepsilon > 0, \quad \sum_{i=1}^n E g_i^2(X_i) I(|g_i(X_i)| > \varepsilon) \rightarrow 0.$$

If in addition, $E|\Delta|^p < \infty$ for some $p > 0$, then by the Chebyshev inequality, one can obtain the rate of convergence

$$\sup_z |P(T \leq z) - \Phi(z)| \leq \sup_z |P(W \leq z) - \Phi(z)| + 2(E|\Delta|^p)^{1/(1+p)}. \tag{1.3}$$

The first term on the right-hand side of (1.3) is well understood via the Berry–Esseen inequality. For example, using Stein’s method, Chen and Shao [9] obtained

$$\begin{aligned} & \sup_z |P(W \leq z) - \Phi(z)| \\ & \leq 4.1 \left(\sum_{i=1}^n E g_i^2(X_i) I(|g_i(X_i)| > 1) + \sum_{i=1}^n E |g_i(X_i)|^3 I(|g_i(X_i)| \leq 1) \right). \end{aligned} \tag{1.4}$$

However, the bound $(E|\Delta|^p)^{1/(1+p)}$ is, in general, not sharp for many commonly used statistics. Many authors have worked toward obtaining better Berry–Esseen bounds. For example, sharp Berry–Esseen bounds have been obtained for general symmetric statistics by van Zwet [24] and Friedrich [12]. An Edgeworth expansion with remainder $O(n^{-1})$ for symmetric statistics was proved by Bentkus, Götze and van Zwet [3].

The main purpose of this paper is to establish uniform and non-uniform Berry–Esseen bounds for general nonlinear statistics. The bounds are the best possible for many known statistics. Our proof is based on a randomized concentration inequality approach to bounding $P(W + \Delta \leq z) - P(W \leq z)$. Because proofs of uniform and non-uniform bounds for sums of independent random variables can be proved via Stein’s method (Chen and Shao [9]), which is much neater and simpler than the traditional Fourier analysis approach, this paper provides a direct and unifying treatment toward the Berry–Esseen bounds for general nonlinear statistics.

This paper is organized as follows. The main results are stated in next section, three applications are presented in Section 3 and an example is given in Section 4 to show the sharpness of the main results. Proofs of the main results are given in Section 5. For the proofs of other results, including Example 4.1, the reader is referred to our technical report (Chen and Shao [10]).

Throughout this paper, C will denote an absolute constant whose value may change at each appearance. The L_p norm of a random variable X is denoted by $\|X\|_p$, that is, $\|X\|_p = (E|X|^p)^{1/p}$ for $p \geq 1$.

2. Main results

Let $\{X_i, 1 \leq i \leq n\}$, T , W and Δ be defined as in Section 1. In the following theorems, we assume that (1.1) is satisfied. Put

$$\beta = \sum_{i=1}^n E|g_i(X_i)|^2 I(|g_i(X_i)| > 1) + \sum_{i=1}^n E|g_i(X_i)|^3 I(|g_i(X_i)| \leq 1) \tag{2.1}$$

and let $\delta > 0$ satisfy

$$\sum_{i=1}^n E|g_i(X_i)| \min(\delta, |g_i(X_i)|) \geq 1/2. \tag{2.2}$$

Theorem 2.1. *For each $1 \leq i \leq n$, let Δ_i be a random variable such that X_i and $(\Delta_i, W - g_i(X_i))$ are independent. Then*

$$\sup_z |P(T \leq z) - P(W \leq z)| \leq 4\delta + E|W\Delta| + \sum_{i=1}^n E|g_i(X_i)(\Delta - \Delta_i)| \tag{2.3}$$

for δ satisfying (2.2). In particular, we have

$$\sup_z |P(T \leq z) - P(W \leq z)| \leq 2\beta + E|W\Delta| + \sum_{i=1}^n E|g_i(X_i)(\Delta - \Delta_i)| \tag{2.4}$$

and

$$\sup_z |P(T \leq z) - \Phi(z)| \leq 6.1\beta + E|W\Delta| + \sum_{i=1}^n E|g_i(X_i)(\Delta - \Delta_i)|. \tag{2.5}$$

The next theorem provides a non-uniform bound.

Theorem 2.2. *For each $1 \leq i \leq n$, let Δ_i be a random variable such that X_i and $(\Delta_i, \{X_j, j \neq i\})$ are independent. Then, for δ satisfying (2.2) and for $z \in R^1$,*

$$|P(T \leq z) - P(W \leq z)| \leq \gamma_z + e^{-|z|/3} \tau, \tag{2.6}$$

where

$$\begin{aligned} \gamma_z &= P(|\Delta| > (|z| + 1)/3) + \sum_{i=1}^n P(|g_i(X_i)| > (|z| + 1)/3) \\ &+ \sum_{i=1}^n P(|W - g_i(X_i)| > (|z| - 2)/3) P(|g_i(X_i)| > 1), \end{aligned} \tag{2.7}$$

$$\tau = 22\delta + 8.5\|\Delta\|_2 + 3.6 \sum_{i=1}^n \|g_i(X_i)\|_2 \|\Delta - \Delta_i\|_2. \tag{2.8}$$

In particular, if $E|g_i(X_i)|^p < \infty$ for $2 < p \leq 3$, then

$$\begin{aligned} & |P(T \leq z) - \Phi(z)| \\ & \leq P(|\Delta| > (|z| + 1)/3) \\ & \quad + C(|z| + 1)^{-p} \left(\|\Delta\|_2 + \sum_{i=1}^n \|g_i(X_i)\|_2 \|\Delta - \Delta_i\|_2 + \sum_{i=1}^n E|g_i(X_i)|^p \right). \end{aligned} \tag{2.9}$$

A result similar to (2.5) was obtained by Friedrich [12] for $g_i = E(T|X_i)$ using the method of characteristic function. Our proof is direct and simpler, and the bounds are easier to calculate. The non-uniform bounds in (2.6) and (2.9) for general nonlinear statistics are new.

Remark 2.1. Assume $E|g_i(X_i)|^p < \infty$ for $p > 2$. Let

$$\delta = \left(\frac{2(p-2)^{p-2}}{(p-1)^{p-1}} \sum_{i=1}^n E|g_i(X_i)|^p \right)^{1/(p-2)}. \tag{2.10}$$

Then (2.2) is satisfied. This follows from the inequality

$$\min(a, b) \geq a - \frac{(p-2)^{p-2} a^{p-1}}{(p-1)^{p-1} b^{p-2}} \tag{2.11}$$

for $a \geq 0$ and $b > 0$.

Remark 2.2. If $\beta \leq 1/2$, then (2.2) is satisfied with $\delta = \beta/2$.

Remark 2.3. Let $\delta > 0$ be such that

$$\sum_{i=1}^n E g_i^2(X_i) I(|g_i(X_i)| > \delta) \leq 1/2.$$

Then (2.2) holds. In particular, if X_1, X_2, \dots, X_n are independent and identically distributed (i.i.d.) random variables and $g_i = g_1$, then (2.2) is satisfied with $\delta = c_0/\sqrt{n}$, where c_0 is a constant such that $E(\sqrt{n}g_1(X_1))^2 I(|\sqrt{n}g_1(X_1)| > c_0) \leq 1/2$.

Remark 2.4. In Theorems 2.1 and 2.2, the choice of Δ_i is flexible. For example, one can choose $\Delta_i = \Delta(X_1, \dots, X_{i-1}, 0, X_{i+1}, \dots, X_n)$ or $\Delta_i = \Delta(X_1, \dots, X_{i-1}, \hat{X}_i, X_{i+1}, \dots, X_n)$, where $\{\hat{X}_i, 1 \leq i \leq n\}$ is an independent copy of $\{X_i, 1 \leq i \leq n\}$. The choice of g_i is also flexible. It can be more general than $g_i(x) = E(T|X_i = x)$, which is commonly used in the literature.

Remark 2.5. Let X_1, \dots, X_n be independent normally distributed random variables with mean zero and variance $1/n$, and let W, T and Δ be as in Example 4.1. Then

$$\begin{aligned}
 & E|W\Delta| + \sum_{i=1}^n E|X_i|^3 + \sum_{i=1}^n E|X_i(\Delta(X_1, \dots, X_i, \dots, X_n) - \Delta(X_1, \dots, 0, \dots, X_n))| \\
 & \leq C\varepsilon^{2/3}
 \end{aligned}
 \tag{2.12}$$

for $(1/\varepsilon)^{4/3} \leq n \leq 16(1/\varepsilon)^{4/3}$. This together with (4.5) shows that the bound in (2.4) is achievable. Moreover, the term $\sum E|g_i(X_i)(\Delta - \Delta_i)|$ in (2.4) cannot be dropped.

3. Applications

Theorems 2.1 and 2.2 can be applied to a wide range of different statistics and provide the bounds of the best possible order in many instances. To illustrate the usefulness and the generality of these results, we give three applications in this section. The uniform bounds refine many existing results by specifying absolute constants, while the non-uniform bounds are new for many cases. Because the results are direct applications of Theorems 2.1 and 2.2, and the proofs are more or less routine verifications of assumptions, the proofs are omitted. Refer to Chen and Shao [10] for detailed proofs. One may also refer to Chen and Shao [10] for applications to random sums of independent random variables with non-random centering and to functions of nonlinear statistics.

3.1. U-statistics

Let X_1, X_2, \dots, X_n be a sequence of i.i.d. random variables, and let $h(x_1, \dots, x_m)$ be a real-valued Borel measurable symmetric function of m variables, where m ($2 \leq m < n$) may depend on n . Consider the Hoeffding [17] U-statistic

$$U_n = \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \dots < i_m \leq n} h(X_{i_1}, \dots, X_{i_m}).$$

The U-statistic elegantly and usefully generalizes the notion of a sample mean. Numerous investigations on the limiting properties of the U-statistics have been done during the last few decades. A systematic presentation of the theory of U-statistics was given by Koroljuk and Borovskich [19]. We refer to the studies on uniform Berry–Esseen bounds for U-statistics by Filippova [11], Grams and Serfling [13], Bickel [4], Chan and Wierman [8], Callaert and Janssen [7], Serfling [21], van Zwet [24] and Friedrich [12]. One can also refer to the work of Wang, Jing and Zhao [26] on uniform Berry–Esseen bounds for studentized U-statistics.

Applying Theorems 2.1 and 2.2 to the U-statistic yields the following result:

Theorem 3.1. Assume that $Eh(X_1, \dots, X_m) = 0$ and $\sigma^2 = Eh^2(X_1, \dots, X_m) < \infty$. Let $g(x) = E(h(X_1, X_2, \dots, X_m) | X_1 = x)$ and $\sigma_1^2 = Eg^2(X_1)$. Assume that $\sigma_1 > 0$. Then

$$\sup_z \left| P\left(\frac{\sqrt{n}}{m\sigma_1} U_n \leq z\right) - P\left(\frac{1}{\sqrt{n}\sigma_1} \sum_{i=1}^n g(X_i) \leq z\right) \right| \leq \frac{(1 + \sqrt{2})(m - 1)\sigma}{(m(n - m + 1))^{1/2}\sigma_1} + \frac{c_0}{\sqrt{n}}, \tag{3.1}$$

where c_0 is a constant such that $Eg^2(X_1)I(|g(X_1)| > c_0\sigma_1) \leq \sigma_1^2/2$. If in addition $E|g(X_1)|^p < \infty$ for $2 < p \leq 3$, then

$$\sup_z \left| P\left(\frac{\sqrt{n}}{m\sigma_1} U_n \leq z\right) - \Phi(z) \right| \leq \frac{(1 + \sqrt{2})(m - 1)\sigma}{(m(n - m + 1))^{1/2}\sigma_1} + \frac{6.1E|g(X_1)|^p}{n^{(p-2)/2}\sigma_1^p} \tag{3.2}$$

and, for $z \in R^1$,

$$\begin{aligned} \left| P\left(\frac{\sqrt{n}}{m\sigma_1} U_n \leq z\right) - \Phi(z) \right| &\leq \frac{9m\sigma^2}{(1 + |z|)^2(n - m + 1)\sigma_1^2} + \frac{13.5e^{-|z|/3}m^{1/2}\sigma}{(n - m + 1)^{1/2}\sigma_1} \\ &\quad + \frac{CE|g(X_1)|^p}{(1 + |z|)^pn^{(p-2)/2}\sigma_1^p}. \end{aligned} \tag{3.3}$$

Moreover, if $E|h(X_1, \dots, X_m)|^p < \infty$ for $2 < p \leq 3$, then for $z \in R^1$,

$$\begin{aligned} &\left| P\left(\frac{\sqrt{n}}{m\sigma_1} U_n \leq z\right) - \Phi(z) \right| \\ &\leq \frac{Cm^{1/2}E|h(X_1, \dots, X_m)|^p}{(1 + |z|)^p(n - m + 1)^{1/2}\sigma_1^p} + \frac{CE|g(X_1)|^p}{(1 + |z|)^pn^{(p-2)/2}\sigma_1^p}. \end{aligned} \tag{3.4}$$

Note that the error in (3.1) is $O(n^{-1/2})$ only under the assumption of the finite second moment of h . The result appears to be previously unknown. The uniform bound given in (3.2) is not new, but the specifying constant for general m is new. The finite second moment of h is not the weakest assumption for the uniform bound. Friedrich [12] obtained an $O(n^{-1/2})$ when $E|h|^{5/3} < \infty$. We refer the reader to Bentkus, Götze and Zitikis [2] and Jing and Zhou [18] for discussions on the necessity of the moment condition.

For the non-uniform bound, Zhao and Chen [27] proved that if $m = 2$, $E|h(X_1, X_2)|^3 < \infty$, then

$$\left| P\left(\frac{\sqrt{n}}{m\sigma_1} U_n \leq z\right) - \Phi(z) \right| \leq An^{-1/2}(1 + |z|)^{-3} \tag{3.5}$$

for $z \in R^1$, where the constant A depends not on n and z , but on the moment of h . Clearly, (3.4) refines Zhao and Chen’s result specifying the relationship of the constant A with the moment condition. After we finished proving Theorem 3.1, Wang [25] informed the second author that he also obtained (3.4) for $m = 2$ and $p = 3$.

Theorems 2.1 and 2.2 can also be applied to derive uniform and non-uniform bounds for the non-i.i.d. case. Especially the uniform bound of Alberink [1] can be easily recovered.

Remark 3.1. Equation (3.3) implies that

$$\left| P\left(\frac{\sqrt{n}}{m\sigma_1}U_n \leq z\right) - \Phi(z) \right| \leq \frac{Cm^{1/2}\sigma^2}{(1+|z|)^3(n-m+1)^{1/2}\sigma_1^2} + \frac{CE|g(X_1)|^p}{(1+|z|)^p n^{(p-2)/2}\sigma_1^p} \quad (3.6)$$

for $|z| \leq ((n-m+1)/m)^{1/2}$. For $|z| > ((n-m+1)/m)^{1/2}$, a bound like (3.6) can be easily obtained by using the Chebyshev inequality. On the other hand, if (3.6) holds for any $z \in R^1$, then it appears necessary to assume $E|h(X_1, \dots, X_m)|^p < \infty$.

Remark 3.2. Theorem 3.1 shows that the central limit theorem can still hold even when $m = m(n) \rightarrow \infty$. However, the explicit Berry–Esseen bounds in (3.1) and (3.2) may not be optimal in terms of m . For the special case of elementary symmetric polynomials (i.e., $h(x_1, x_2, \dots, x_m) = x_1 x_2 \dots x_m - \mu^m$, where $\mu = E(X_1) \neq 0$), discussed in van Es and Helmers [23], σ/σ_1 in (3.1) and (3.2) is of order r^m , where $r > 1$, which is equivalent to requiring $m \leq c \log n$ for some $c > 0$ so that the bound tends to 0, whereas in the result of van Es and Helmers [23], the bound tends to 0 as long as $m = o(n^{1/2})$. The Berry–Esseen bound obtained by van Es and Helmers [23] gives an optimal dependence on m .

3.2. Multisample U-statistics

Consider k independent sequences $\{X_{j1}, \dots, X_{jn_j}\}$ of i.i.d. random variables, $j = 1, \dots, k$. Let $h(x_{jl}, l = 1, \dots, m_j, j = 1, \dots, k)$ be a measurable function symmetric with respect to m_j arguments of the j th set, $m_j \geq 1, j = 1, \dots, k$. Let

$$\theta = Eh(X_{jl}, l = 1, \dots, m_j, j = 1, \dots, k).$$

The multisample U-statistic is defined as

$$U_{\bar{n}} = \left\{ \prod_{j=1}^k \binom{n_j}{m_j}^{-1} \right\} \sum h(X_{jl}, l = i_{j1}, \dots, i_{jm_j}, j = 1, \dots, k),$$

where $\bar{n} = (n_1, \dots, n_k)$ and the summation is carried out over all $1 \leq i_{j1} < \dots < i_{jm_j} \leq n_j, n_j \geq 2m_j, j = 1, \dots, k$. Clearly, $U_{\bar{n}}$ is an unbiased estimate of θ . The two-sample Wilcoxon statistic and the two-sample ω^2 -statistic are two typical examples of multisample U-statistics. Without loss of generality, assume $\theta = 0$. For $j = 1, \dots, k$, define

$$h_j(x) = E(h(X_{11}, \dots, X_{1m_1}; \dots; X_{k1}, \dots, X_{km_k}) | X_{j1} = x)$$

and let $\sigma_j^2 = Eh_j^2(X_{j1})$ and

$$\sigma_{\bar{n}}^2 = \sum_{j=1}^k \frac{m_j^2}{n_j} \sigma_j^2.$$

A uniform Berry–Esseen bound with $O((\min_{1 \leq j \leq k} n_j)^{-1/2})$ for the multisample U-statistics was obtained by Helmers and Janssen [15] and Borovskich [6] (see Koroljuk and Borovskich [19], pages 304–311). The next theorem refines their results.

Theorem 3.2. *Assume that $\theta = 0$, $\sigma^2 := \mathbb{E}h^2(X_{11}, \dots, X_{1m_1}; \dots; X_{k1}, \dots, X_{km_k}) < \infty$ and $\max_{1 \leq j \leq k} \sigma_j > 0$. Then for $2 < p \leq 3$,*

$$\sup_z |P(\sigma_{\bar{n}}^{-1} U_{\bar{n}} \leq z) - \Phi(z)| \leq \frac{(1 + \sqrt{2})\sigma}{\sigma_{\bar{n}}} \sum_{j=1}^k \frac{m_j^2}{n_j} + \frac{6.6}{\sigma_{\bar{n}}^p} \sum_{j=1}^k \frac{m_j^p}{n_j^{p-1}} \mathbb{E}|h_j(X_{j1})|^p \tag{3.7}$$

and for $z \in \mathbb{R}^1$,

$$\begin{aligned} |P(\sigma_{\bar{n}}^{-1} U_{\bar{n}} \leq z) - \Phi(z)| &\leq \frac{9\sigma^2}{(1 + |z|)^2 \sigma_{\bar{n}}^2} \left(\sum_{j=1}^k \frac{m_j^2}{n_j} \right)^2 + 13.5e^{-|z|/3} \frac{\sigma}{\sigma_{\bar{n}}} \sum_{j=1}^k \frac{m_j^2}{n_j} \\ &+ \frac{C}{(1 + |z|)^p \sigma_{\bar{n}}^p} \sum_{j=1}^k \frac{m_j^p \mathbb{E}|h_j(X_{j1})|^p}{n_j^{p-1}}. \end{aligned} \tag{3.8}$$

3.3. L-statistics

Let X_1, \dots, X_n be i.i.d. random variables with a common distribution function F and let F_n be the empirical distribution function defined by

$$F_n(x) = n^{-1} \sum_{i=1}^n I(X_i \leq x), \quad \text{for } x \in \mathbb{R}^1.$$

Let $J(t)$ be a real-valued function on $[0, 1]$ and define

$$T(G) = \int_{-\infty}^{\infty} x J(G(x)) dG(x)$$

for non-decreasing measurable function G . Put

$$\sigma^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(F(s)) J(F(t)) F(\min(s, t)) (1 - F(\max(s, t))) ds dt$$

and

$$g(x) = \int_{-\infty}^{\infty} (I(x \leq s) - F(s)) J(F(s)) ds.$$

The statistic $T(F_n)$ is called an L-statistic (see Serfling [21], Chapter 8). Uniform Berry–Esseen bounds for the L-statistic that smooths J were given by Helmers [14] and Helmers, Janssen and Serfling [16]. Applying Theorems 2.1 and 2.2 yields the following uniform and non-uniform bounds for L-statistics.

Theorem 3.3. Let $n \geq 4$. Assume that $EX_1^2 < \infty$ and $E|g(X_1)|^p < \infty$ for $2 < p \leq 3$. If the weight function $J(t)$ is Lipschitz of order 1 on $[0, 1]$, that is, there exists a constant c_0 such that

$$|J(t) - J(s)| \leq c_0|t - s|, \quad \text{for } 0 \leq s, t \leq 1, \tag{3.9}$$

then

$$\sup_z |P(\sqrt{n}\sigma^{-1}(T(F_n) - T(F)) \leq z) - \Phi(z)| \leq \frac{(1 + \sqrt{2})c_0\|X_1\|_2}{\sqrt{n}\sigma} + \frac{6.1E|g(X_1)|^p}{n^{(p-2)/2}\sigma^p} \tag{3.10}$$

and

$$\begin{aligned} &|P(\sqrt{n}\sigma^{-1}(T(F_n) - T(F)) \leq z) - \Phi(z)| \\ &\leq \frac{9c_0^2EX_1^2}{(1 + |z|)^2n\sigma^2} + \frac{C}{(1 + |z|)^p} \left(\frac{c_0\|X_1\|_2}{\sqrt{n}\sigma} + \frac{E|g(X_1)|^p}{n^{(p-2)/2}\sigma^p} \right). \end{aligned} \tag{3.11}$$

4. An example

In this section we give an example to show that the bound of (2.4) in Theorem 2.1 is achievable. Moreover, the term $\sum E|g_i(X_i)(\Delta - \Delta_i)|$ in (2.4) cannot be dropped. The example also provides a counterexample to a result of Shorack [22] and of Bolthausen and Götze [5]. We refer to Chen and Shao [10] for a detailed proof.

Example 4.1. Let X_1, \dots, X_n be independent normally distributed random variables with mean zero and variance $1/n$. Define

$$W = \sum_{i=1}^n X_i, \quad T := T_\varepsilon = W - \varepsilon|W|^{-1/2} + \varepsilon c_0 \quad \text{and} \quad \Delta = T - W = -\varepsilon|W|^{-1/2} + \varepsilon c_0,$$

where $c_0 = E(|W|^{-1/2}) = \sqrt{2/\pi} \int_0^\infty x^{-1/2} e^{-x^2/2} dx$. Let $\{\hat{X}_i, 1 \leq i \leq n\}$ be an independent copy of $\{X_i, 1 \leq i \leq n\}$ and define

$$\alpha = \frac{1}{n} \sum_{i=1}^n E|\Delta(X_1, \dots, X_i, \dots, X_n) - \Delta(X_1, \dots, \hat{X}_i, \dots, X_n)|. \tag{4.1}$$

Then $ET = 0$ and for $0 < \varepsilon < 1/64$ and $n \geq (1/\varepsilon)^4$,

$$P(T \leq \varepsilon c_0) - \Phi(\varepsilon c_0) \geq \varepsilon^{2/3}/6, \tag{4.2}$$

$$E|W\Delta| + E|\Delta| \leq 7\varepsilon, \tag{4.3}$$

$$E|\Delta| + \sum_{i=1}^n E|X_i|^3 + \sqrt{\alpha} \leq C\varepsilon, \tag{4.4}$$

where C is an absolute constant.

Clearly, (4.2) implies that

$$\sup_z |P(T_\varepsilon \leq z) - \Phi(z)| \geq \frac{\varepsilon^{2/3}}{6}. \tag{4.5}$$

A result from Shorack [22] (see Lemma 11.1.3, page 261) states that for any random variables W and Δ ,

$$\sup_z |P(W + \Delta \leq z) - \Phi(z)| \leq \sup_z |P(W \leq z) - \Phi(z)| + 4E|W\Delta| + 4E|\Delta|. \tag{4.6}$$

Another result, which is in Theorem 2 of Bolthausen and Götze [5], states that if $ET = 0$, then

$$\sup_z |P(T \leq z) - \Phi(z)| \leq C \left(E|\Delta| + \sum_{i=1}^n E|g_i(X_i)|^3 + \sqrt{\alpha} \right), \tag{4.7}$$

where C is an absolute constant and α is defined in (4.1).

In view of (4.3), (4.4) and (4.5), the results of Shorack and of Bolthausen and Götze can be shown to lead to a contradiction.

5. Proofs of main theorems

Proof of Theorem 2.1. Equation (2.5) follows from (2.4) and (1.4). When $\beta > 1/2$, (2.4) is trivial. For $\beta \leq 1/2$, (2.4) is a consequence of (2.3) and Remark 2.2. Thus, we need only to prove (2.3). Note that

$$-P(z - |\Delta| \leq W \leq z) \leq P(T \leq z) - P(W \leq z) \leq P(z \leq W \leq z + |\Delta|). \tag{5.1}$$

It suffices to show that

$$P(z \leq W \leq z + |\Delta|) \leq 4\delta + E|W\Delta| + \sum_{i=1}^n E|g_i(X_i)(\Delta - \Delta_i)| \tag{5.2}$$

and

$$P(z - |\Delta| \leq W \leq z) \leq 4\delta + E|W\Delta| + \sum_{i=1}^n E|g_i(X_i)(\Delta - \Delta_i)|, \tag{5.3}$$

where δ satisfies (2.2). Let

$$f_\Delta(w) = \begin{cases} -|\Delta|/2 - \delta, & \text{for } w \leq z - \delta, \\ w - \frac{1}{2}(2z + |\Delta|), & \text{for } z - \delta \leq w \leq z + |\Delta| + \delta, \\ |\Delta|/2 + \delta, & \text{for } w > z + |\Delta| + \delta. \end{cases} \tag{5.4}$$

Let

$$\begin{aligned} \xi_i &= g_i(X_i), & \hat{M}_i(t) &= \xi_i \{I(-\xi_i \leq t \leq 0) - I(0 < t \leq -\xi_i)\}, \\ M_i(t) &= E\hat{M}_i(t), & \hat{M}(t) &= \sum_{i=1}^n \hat{M}_i(t), & M(t) &= E\hat{M}(t). \end{aligned}$$

Since ξ_i and $f_{\Delta_i}(W - \xi_i)$ are independent for $1 \leq i \leq n$ and $E\xi_i = 0$, we have

$$\begin{aligned} E\{Wf_{\Delta}(W)\} &= \sum_{1 \leq i \leq n} E\{\xi_i(f_{\Delta}(W) - f_{\Delta}(W - \xi_i))\} \\ &\quad + \sum_{1 \leq i \leq n} E\{\xi_i(f_{\Delta}(W - \xi_i) - f_{\Delta_i}(W - \xi_i))\} \\ &:= H_1 + H_2. \end{aligned} \tag{5.5}$$

Using the fact that $\hat{M}(t) \geq 0$ and $f'_{\Delta}(w) \geq 0$, we have

$$\begin{aligned} H_1 &= \sum_{1 \leq i \leq n} E\left\{\xi_i \int_{-\xi_i}^0 f'_{\Delta}(W+t) dt\right\} \\ &= \sum_{1 \leq i \leq n} E\left\{\int_{-\infty}^{\infty} f'_{\Delta}(W+t) \hat{M}_i(t) dt\right\} \\ &= E\left\{\int_{-\infty}^{\infty} f'_{\Delta}(W+t) \hat{M}(t) dt\right\} \\ &\geq E\left\{\int_{|t| \leq \delta} f'_{\Delta}(W+t) \hat{M}(t) dt\right\} \\ &\geq E\left\{I(z \leq W \leq z + |\Delta|) \int_{|t| \leq \delta} \hat{M}(t) dt\right\} \\ &= \sum_{1 \leq i \leq n} E\{I(z \leq W \leq z + |\Delta|) |\xi_i| \min(\delta, |\xi_i|)\} \\ &\geq H_{1,1} - H_{1,2}, \end{aligned} \tag{5.6}$$

where

$$\begin{aligned} H_{1,1} &= P(z \leq W \leq z + |\Delta|) \sum_{1 \leq i \leq n} E\eta_i, \\ H_{1,2} &= E\left|\sum_{1 \leq i \leq n} \eta_i - E\eta_i\right|, \quad \eta_i = |\xi_i| \min(\delta, |\xi_i|). \end{aligned}$$

By (2.2),

$$\sum_{1 \leq i \leq n} E\eta_i \geq 1/2.$$

Hence

$$H_{1,1} \geq (1/2)P(z \leq W \leq z + |\Delta|). \tag{5.7}$$

By the Cauchy–Schwarz inequality,

$$\begin{aligned} H_{1,2} &\leq \left(E \left(\sum_{1 \leq i \leq n} \eta_i - E\eta_i \right)^2 \right)^{1/2} \\ &\leq \left(\sum_{1 \leq i \leq n} E\eta_i^2 \right)^{1/2} \leq \delta. \end{aligned} \tag{5.8}$$

As to H_2 , it is easy to see that

$$|f_\Delta(w) - f_{\Delta_i}(w)| \leq ||\Delta| - |\Delta_i||/2 \leq |\Delta - \Delta_i|/2.$$

Hence

$$|H_2| \leq (1/2) \sum_{i=1}^n E|\xi_i(\Delta - \Delta_i)|. \tag{5.9}$$

Combining (5.5), (5.7), (5.8) and (5.9) yields

$$\begin{aligned} P(z \leq W \leq z + |\Delta|) &\leq 2 \left\{ E|Wf_\Delta(W)| + \delta + (1/2) \sum_{i=1}^n E|\xi_i(\Delta - \Delta_i)| \right\} \\ &\leq E|W\Delta| + 2\delta E|W| + 2\delta + \sum_{i=1}^n E|\xi_i(\Delta - \Delta_i)| \\ &\leq 4\delta + E|W\Delta| + \sum_{i=1}^n E|\xi_i(\Delta - \Delta_i)|. \end{aligned}$$

This proves (5.2). Similarly, one can prove (5.3) and hence Theorem 2.1. □

Proof of Theorem 2.2. First, we prove (2.9). For $|z| \leq 4$, (2.9) holds by (2.5). For $|z| > 4$, consider two cases.

Case 1: $\sum_{i=1}^n E|g_i(X_i)|^p > 1/2$. By the Rosenthal [20] inequality, we have

$$\begin{aligned} P(|W| > (|z| - 2)/3) &\leq P(|W| > |z|/6) \leq (|z|/6)^{-p} E|W|^p \\ &\leq C(|z| + 1)^{-p} \left\{ \left(\sum_{i=1}^n E g_i^2(X_i) \right)^{p/2} + \sum_{i=1}^n E|g_i(X_i)|^p \right\} \end{aligned}$$

$$\leq C(|z| + 1)^{-p} \sum_{i=1}^n \mathbb{E}|g_i(X_i)|^p. \tag{5.10}$$

Hence

$$\begin{aligned} |P(T \leq z) - \Phi(z)| &\leq P(|\Delta| > (|z| + 1)/3) + P(|W| > (|z| - 2)/3) + P(|N(0, 1)| > |z|) \\ &\leq P(|\Delta| > (|z| + 1)/3) + C(|z| + 1)^{-p} \sum_{i=1}^n \mathbb{E}|g_i(X_i)|^p, \end{aligned}$$

which shows that (2.9) holds.

Case 2: $\sum_{i=1}^n \mathbb{E}|g_i(X_i)|^p \leq 1/2$. Similar to (5.10), we have

$$\begin{aligned} P(|W - g_i(X_i)| > (|z| - 2)/3) &\leq C(|z| + 1)^{-p} \left\{ \left(\sum_{j=1}^n \mathbb{E}g_j^2(X_j) \right)^{p/2} + \sum_{j=1}^n \mathbb{E}|g_j(X_j)|^p \right\} \\ &\leq C(|z| + 1)^{-p} \end{aligned}$$

and hence

$$\begin{aligned} \gamma_z &\leq P(|\Delta| > (|z| + 1)/3) + \sum_{i=1}^n (|z| + 1)^{-p} \mathbb{E}|g_i(X_i)|^p + \sum_{i=1}^n C(|z| + 1)^{-p} \mathbb{E}|g_i(X_i)|^p \\ &\leq P(|\Delta| > (|z| + 1)/3) + C(|z| + 1)^{-p} \sum_{i=1}^n \mathbb{E}|g_i(X_i)|^p. \end{aligned}$$

By Remark 2.1, we can choose

$$\begin{aligned} \delta &= \left(\frac{2(p-2)^{p-2}}{(p-1)^{p-1}} \sum_{i=1}^n \mathbb{E}|g_i(X_i)|^p \right)^{1/(p-2)} \\ &\leq \frac{2(p-2)^{p-2}}{(p-1)^{p-1}} \sum_{i=1}^n \mathbb{E}|g_i(X_i)|^p. \end{aligned}$$

Combining the above inequalities with (2.6) and the non-uniform Berry–Esseen bound for independent random variables yields (2.9).

Next we prove (2.6). The main idea of the proof is first to truncate $g_i(X_i)$ and then adopt the proof of Theorem 2.1 to the truncated sum. Without loss of generality, assume $z \geq 0$, because we can simply apply the result to $-T$. By (5.1), it suffices to show that

$$P(z - |\Delta| \leq W \leq z) \leq \gamma_z + e^{-z/3} \tau \tag{5.11}$$

and

$$P(z \leq W \leq z + |\Delta|) \leq \gamma_z + e^{-z/3} \tau. \tag{5.12}$$

Since the proof of (5.12) is similar to that of (5.11), we prove only (5.11). It is easy to see that

$$P(z - |\Delta| \leq W \leq z) \leq P(|\Delta| > (z + 1)/3) + P(z - |\Delta| \leq W \leq z, |\Delta| \leq (z + 1)/3).$$

Now (5.11) follows directly by Lemmas 5.1 and 5.2 below. This completes the proof of Theorem 2.2. \square

Lemma 5.1. *Let*

$$\xi_i = g_i(X_i), \quad \bar{\xi}_i = \xi_i I(\xi_i \leq 1), \quad \bar{W} = \sum_{i=1}^n \bar{\xi}_i.$$

Then

$$\begin{aligned} &P(z - |\Delta| \leq W \leq z, |\Delta| \leq (z + 1)/3) \\ &\leq P(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3) \\ &\quad + \sum_{i=1}^n P(\xi_i > (z + 1)/3) + \sum_{i=1}^n P(W - \xi_i > (z - 2)/3) P(|\xi_i| > 1). \end{aligned} \tag{5.13}$$

Proof. We have

$$\begin{aligned} &P(z - |\Delta| \leq W \leq z, |\Delta| \leq (z + 1)/3) \\ &\leq P\left(z - |\Delta| \leq W \leq z, |\Delta| \leq (z + 1)/3, \max_{1 \leq i \leq n} |\xi_i| \leq 1\right) \\ &\quad + P\left(z - |\Delta| \leq W \leq z, |\Delta| \leq (z + 1)/3, \max_{1 \leq i \leq n} |\xi_i| > 1\right) \\ &\leq P(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3) + \sum_{i=1}^n P(W > (2z - 1)/3, |\xi_i| > 1) \end{aligned}$$

and

$$\begin{aligned} &\sum_{i=1}^n P(W > (2z - 1)/3, |\xi_i| > 1) \\ &\leq \sum_{i=1}^n P(\xi_i > (z + 1)/3) + \sum_{i=1}^n P(W > (2z - 1)/3, \xi_i \leq (z + 1)/3, |\xi_i| > 1) \\ &\leq \sum_{i=1}^n P(\xi_i > (z + 1)/3) + \sum_{i=1}^n P(W - \xi_i > (z - 2)/3, |\xi_i| > 1) \\ &= \sum_{i=1}^n P(\xi_i > (z + 1)/3) + \sum_{i=1}^n P(W - \xi_i > (z - 2)/3) P(|\xi_i| > 1), \end{aligned}$$

as desired. □

Lemma 5.2. *We have*

$$P(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3) \leq e^{-z/3} \tau. \tag{5.14}$$

Proof. Noting that $E\bar{\xi}_i \leq 0$, that $e^s \leq 1 + s + s^2(e^a - 1 - a)a^{-2}$ for $s \leq a$ and $a > 0$ and that $a\bar{\xi}_i \leq a$, we have for $a > 0$,

$$\begin{aligned} Ee^{a\bar{W}} &= \prod_{i=1}^n Ee^{a\bar{\xi}_i} \\ &\leq \prod_{i=1}^n (1 + aE\bar{\xi}_i + (e^a - 1 - a)E\bar{\xi}_i^2) \\ &\leq \exp\left((e^a - 1 - a) \sum_{i=1}^n E\bar{\xi}_i^2\right) \\ &\leq \exp\left((e^a - 1 - a) \sum_{i=1}^n E\xi_i^2\right) \\ &= \exp(e^a - 1 - a). \end{aligned} \tag{5.15}$$

In particular, we have $Ee^{\bar{W}/2} \leq \exp(e^{1/2} - 1.5)$. If $\delta \geq 0.07$, then

$$\begin{aligned} P(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3) \\ &\leq P(\bar{W} > (2z - 1)/3) \leq e^{-z/3+1/6} Ee^{\bar{W}/2} \\ &\leq e^{-z/3} \exp(e^{0.5} - 4/3) \leq 1.38e^{-z/3} \leq 20\delta e^{-z/3}. \end{aligned}$$

This proves (5.14) when $\delta \geq 0.07$.

For $\delta < 0.07$, let

$$f_\Delta(w) = \begin{cases} 0, & \text{for } w \leq z - |\Delta| - \delta, \\ e^{w/2}(w - z + |\Delta| + \delta), & \text{for } z - |\Delta| - \delta \leq w \leq z + \delta, \\ e^{w/2}(|\Delta| + 2\delta), & \text{for } w > z + \delta. \end{cases} \tag{5.16}$$

Put

$$\bar{M}_i(t) = \xi_i \{I(-\bar{\xi}_i \leq t \leq 0) - I(0 < t \leq -\bar{\xi}_i)\}, \quad \bar{M}(t) = \sum_{i=1}^n \bar{M}_i(t).$$

By (5.5) and similar to (5.6), we have

$$\begin{aligned} E\{Wf_{\Delta}(\bar{W})\} &= E\left\{\int_{-\infty}^{\infty} f'_{\Delta}(\bar{W} + t)\bar{M}(t) dt\right\} \\ &\quad + \sum_{i=1}^n E\{\xi_i(f_{\Delta}(\bar{W} - \bar{\xi}_i) - f_{\Delta_i}(\bar{W} - \bar{\xi}_i))\} \\ &:= G_1 + G_2. \end{aligned} \tag{5.17}$$

It follows from the fact that $\bar{M}(t) \geq 0$, $f'_{\Delta}(w) \geq e^{w/2}$ for $z - |\Delta| - \delta \leq w \leq z + \delta$ and $f'_{\Delta}(w) \geq 0$ for all w that

$$\begin{aligned} G_1 &\geq E\left\{\int_{|t|\leq\delta} f'_{\Delta}(\bar{W} + t)\bar{M}(t) dt\right\} \\ &\geq E\left\{e^{(\bar{W}-\delta)/2}I(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3) \int_{|t|\leq\delta} \bar{M}(t) dt\right\} \\ &= E\left\{e^{(\bar{W}-\delta)/2}I(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3)\right\} \int_{|t|\leq\delta} E\bar{M}(t) dt \\ &\quad + E\left\{e^{(\bar{W}-\delta)/2}I(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3) \int_{|t|\leq\delta} (\bar{M}(t) - E\bar{M}(t)) dt\right\} \\ &\geq G_{1,1} - G_{1,2}, \end{aligned} \tag{5.18}$$

where

$$\begin{aligned} G_{1,1} &= e^{z/3-1/6-0.035} P(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3) \int_{|t|\leq\delta} E\bar{M}(t) dt, \\ G_{1,2} &= E\left\{\int_{|t|\leq\delta} e^{\bar{W}/2}|\bar{M}(t) - E\bar{M}(t)| dt\right\}. \end{aligned}$$

By (2.2) and the assumption that $\delta \leq 0.07$,

$$\begin{aligned} \int_{|t|\leq\delta} E\bar{M}(t) dt &= \sum_{i=1}^n E|\xi_i| \min(\delta, |\bar{\xi}_i|) \\ &= \sum_{i=1}^n E|\xi_i| \min(\delta, |\xi_i|) \geq 1/2. \end{aligned}$$

Hence

$$G_{1,1} \geq (1/2)e^{z/3-1/6-0.035} P(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3). \tag{5.19}$$

By (5.15), we have $Ee^{\bar{W}} \leq \exp(e - 2) < 2.06$. It follows from the Cauchy–Schwarz inequality that

$$\begin{aligned}
 G_{1,2} &\leq 0.5 \int_{|t| \leq \delta} (0.5Ee^{\bar{W}} + 2E|\bar{M}(t) - E\bar{M}(t)|^2) dt \\
 &\leq 0.5 \left\{ 2.06\delta + 2 \sum_{i=1}^n \int_{|t| \leq \delta} E\xi_i^2 (I(-\bar{\xi}_i \leq t \leq 0) + I(0 < t \leq -\bar{\xi}_i)) dt \right\} \\
 &= 0.5 \left\{ 2.06\delta + 2 \sum_{i=1}^n E\xi_i^2 \min(\delta, |\bar{\xi}_i|) \right\} \\
 &\leq 0.5 \left\{ 2.06\delta + 2\delta \sum_{i=1}^n E\xi_i^2 \right\} \leq 2.03\delta.
 \end{aligned} \tag{5.20}$$

With regard to G_2 , it is easy to see that

$$|f_\Delta(w) - f_{\Delta_i}(w)| \leq e^{w/2} ||\Delta| - |\Delta_i|| \leq e^{w/2} |\Delta - \Delta_i|.$$

Hence, by the Hölder inequality, (5.15) and the assumption that ξ_i and $\bar{W} - \bar{\xi}_i$ are independent,

$$\begin{aligned}
 |G_2| &\leq \sum_{i=1}^n E|\xi_i e^{(\bar{W} - \bar{\xi}_i)/2} (\Delta - \Delta_i)| \\
 &\leq \sum_{i=1}^n (E\xi_i^2 e^{\bar{W} - \bar{\xi}_i})^{1/2} (E(\Delta - \Delta_i)^2)^{1/2} \\
 &= \sum_{i=1}^n (E\xi_i^2 Ee^{\bar{W} - \bar{\xi}_i})^{1/2} \|\Delta - \Delta_i\|_2 \\
 &\leq 1.44 \sum_{i=1}^n \|\xi_i\|_2 \|\Delta - \Delta_i\|_2.
 \end{aligned} \tag{5.21}$$

Following the proof of (5.15) and by using $|e^s - 1| \leq |s|(e^a - 1)/a$ for $s \leq a$ and $a > 0$, we have

$$\begin{aligned}
 EW^2 e^{\bar{W}} &= \sum_{i=1}^n E\xi_i^2 e^{\bar{\xi}_i} Ee^{\bar{W} - \bar{\xi}_i} + \sum_{1 \leq i \neq j \leq n} E\xi_i (e^{\bar{\xi}_i} - 1) E\xi_j (e^{\bar{\xi}_j} - 1) Ee^{\bar{W} - \bar{\xi}_i - \bar{\xi}_j} \\
 &\leq 2.06e \sum_{i=1}^n E\xi_i^2 + 2.06(e - 1)^2 \sum_{1 \leq i \neq j \leq n} E\xi_i^2 E\xi_j^2 \\
 &\leq 2.06e + 2.06(e - 1)^2 < 3.42^2.
 \end{aligned}$$

Thus, we obtain

$$\begin{aligned} \mathbb{E}\{Wf_{\Delta}(\bar{W})\} &\leq \mathbb{E}|W|e^{\bar{W}/2}(|\Delta| + 2\delta) \\ &\leq \{\|\Delta\|_2 + 2\delta\}(\mathbb{E}(W^2e^{\bar{W}}))^{1/2} \\ &\leq 3.42(\|\Delta\|_2 + 2\delta). \end{aligned} \tag{5.22}$$

Combining (5.17), (5.19), (5.20), (5.21) and (5.22) yields

$$\begin{aligned} &P(z - |\Delta| \leq \bar{W} \leq z, |\Delta| \leq (z + 1)/3) \\ &\leq 2e^{-z/3+1/6+0.035} \left\{ 3.42(\|\Delta\|_2 + 2\delta) + 2.03\delta + 1.44 \sum_{i=1}^n \|\xi_i\|_2 \|\Delta - \Delta_i\|_2 \right\} \\ &\leq e^{-z/3} \left\{ 22\delta + 8.5\|\Delta\|_2 + 3.6 \sum_{i=1}^n \|\xi_i\|_2 \|\Delta - \Delta_i\|_2 \right\} \\ &= e^{-z/3} \tau. \end{aligned}$$

This proves (5.14). □

Acknowledgements

The authors are grateful to Xuming He for his contribution to the construction of the example in Section 4. The authors thank two referees, an associate editor and the editor for their valuable comments. The research of L.H.Y. Chen was partially supported by National University of Singapore grant R-146-000-013-112, and the research of Q.M. Shao was partially supported by NSF grant DMS-0103487, National University of Singapore grant R-146-000-038-101 and HKUST grant DAG 05/06.SC27 and RGC 602206.

References

- [1] Alberink, I.B. (2000). A Berry–Esseen bound for U-statistics in the non-i.i.d. case. *J. Theoret. Probab.* **13** 519–533. [MR1778586](#)
- [2] Bentkus, V., Götze, F. and Zitikis, M. (1994). Lower estimates of the convergence rate for U-statistics. *Ann. Probab.* **22** 1701–1714. [MR1331199](#)
- [3] Bentkus, V., Götze, F. and van Zwet, W.R. (1997). An Edgeworth expansion for symmetric statistics. *Ann. Statist.* **25** 851–896. [MR1439326](#)
- [4] Bickel, P. (1974). Edgeworth expansion in nonparametric statistics. *Ann. Statist.* **2** 1–20. [MR0350952](#)
- [5] Bolthausen, E. and Götze, F. (1993). The rate of convergence for multivariate sampling statistics. *Ann. Statist.* **21** 1692–1710. [MR1245764](#)
- [6] Borovskich, Yu.V. (1983). Asymptotics of U-statistics and von Mises’ functionals. *Sov. Math. Dokl.* **27** 303–308.

- [7] Callaert, H. and Janssen, P. (1978). The Berry–Esseen theorem for U-statistics. *Ann. Statist.* **6** 417–421. [MR0464359](#)
- [8] Chan, Y.K. and Wierman, J. (1977). On the Berry–Esseen theorem for U-statistics. *Ann. Probab.* **5** 136–139. [MR0433551](#)
- [9] Chen, L.H.Y. and Shao, Q.M. (2001). A non-uniform Berry–Esseen bound via Stein’s method. *Probab. Theory Related Fields* **120** 236–254. [MR1841329](#)
- [10] Chen, L.H.Y. and Shao, Q.M. (2006). Normal approximation for non-linear statistics using a concentration inequality approach. Available at <http://www.math.nus.edu.sg/~lhychen>.
- [11] Filippova, A.A. (1962). Mises’s theorem of the asymptotic behavior of functionals of empirical distribution functions and its statistical applications. *Theory Probab. Appl.* **7** 24–57. [MR0150870](#)
- [12] Friedrich, K.O. (1989). A Berry–Esseen bound for functions of independent random variables. *Ann. Statist.* **17** 170–183. [MR0981443](#)
- [13] Grams, W.F. and Serfling, R.J. (1973). Convergence rates for U-statistics and related statistics. *Ann. Statist.* **1** 153–160. [MR0336788](#)
- [14] Helmers, R. (1977). The order of the normal approximation for linear combinations of order statistics with smooth weight functions. *Ann. Probab.* **5** 940–953. [MR0458716](#)
- [15] Helmers, R. and Janssen, P. (1982). On the Berry–Esseen theorem for multivariate U-statistics. *Math. Cent. Rep. SW 90/82* 1–22. Amsterdam: Mathematisch Centrum.
- [16] Helmers, R., Janssen, P. and Serfling, R.J. (1990). Berry–Esseen bound and bootstrap results for generalized L-statistics. *Scand. J. Statist.* **17** 65–77. [MR1062846](#)
- [17] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325. [MR0026294](#)
- [18] Jing, B.-Y. and Zhou, W. (2005). A note on Edgeworth expansions for U-statistics under minimal conditions. *Lithuanian Math. J.* **45** 353–358. [MR2192994](#)
- [19] Koroljuk, V.S. and Borovskich, Yu.V. (1994). *Theory of U-statistics*. Boston: Kluwer Academic Publishers. [MR1472486](#)
- [20] Rosenthal, H.P. (1970). On the subspaces of L_p ($p > 2$) spanned by sequences of independent random variables. *Israel J. Math.* **8** 273–303. [MR0271721](#)
- [21] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley. [MR0595165](#)
- [22] Shorack, G.R. (2000). *Probability for Statisticians*. New York: Springer. [MR1762415](#)
- [23] van Es, A.J. and Helmers, R. (1988). Elementary symmetric polynomials of increasing order. *Probab. Theory Related Fields* **80** 21–35. [MR0970469](#)
- [24] van Zwet, W.R. (1984). A Berry–Esseen bound for symmetric statistics. *Z. Wahrsch. Verw. Gebiete* **66** 425–440. [MR0751580](#)
- [25] Wang, Q. (2002). Non-uniform Berry–Esseen Bound for U-Statistics. *Statist. Sinica* **12** 1157–1169. [MR1947069](#)
- [26] Wang, Q., Jing, B.Y. and Zhao, L. (2000). The Berry–Esseen bound for studentized statistics. *Ann. Probab.* **28** 511–535. [MR1756015](#)
- [27] Zhao, L.C. and Chen, X.R. (1983). Non-uniform convergence rates for distributions of U-statistics. *Sci. Sinica Ser. A* **26** 795–810. [MR0725649](#)

Received November 2005 and revised November 2006