Statistica Sinica **20** (2010), 613-636

A UNIFIED APPROACH TO EDGEWORTH EXPANSIONS FOR A GENERAL CLASS OF STATISTICS

Bing-Yi Jing and Qiying Wang

Hong Kong University of Science and Technology and University of Sydney

Abstract: This paper is concerned with Edgeworth expansions for a general class of statistics under very weak conditions. Our approach unifies the treatment for both standardized and studentized statistics that have been traditionally studied separately under usually different conditions. These results are then applied to several special classes of well-known statistics: U-statistics, L-statistics, and functions of sample means. Special attention is paid to the studentized statistics. We establish Edgeworth expansions under very weak or minimal moment conditions.

Key words and phrases: Edgeworth expansion, function of sample means, L-statistic, standardization, studentization, U-statistic.

1. Introduction

Suppose we are interested in the distribution of a statistic, $T = T(X_1, \ldots, X_n)$, where $X_1, \ldots, X_n$ is a sequence of independent and identically distributed (i.i.d.) random variables (r.v.'s). Typically, one can use the delta method to show that T converges in distribution to a normal distribution. The rates of convergence to normality are usually of the order $n^{-1/2}$ and can be described by Berry-Esséen bounds. To get a better approximation than asymptotic normality, one can develop higher-order Edgeworth expansions under appropriate conditions.

The theory of Edgeworth expansions dates back a long way, the simplest case being the Edgeworth expansion for the sample mean. Much effort has been devoted to Edgeworth expansions for other classes of statistics, such as functions of means, U-, L-statistics, and others. On the other hand, Edgeworth expansions for their studentized counterparts have also gained much momentum, partly due to their usefulness in statistical inference. It is worth pointing out that each of the methods for deriving Edgeworth expansions for the above-mentioned statistics was tailored to the individual structures of these statistics. A general unifying approach is to consider *symmetric statistics*, which include all the aforementioned, as in Lai and Wang (1993), Bentkus, Götze and van Zwet (1997), and Putter and van Zwet (1998) for instance.

A quick glance at the literature reveals that the moment conditions in Edgeworth expansions for the studentized statistics are typically stronger than for the corresponding standardized statistics; see Section 3 for more discussions. One notable exception is the case of the sample mean, where the third moment is enough for both the standardized mean and Student's t-statistic, see Hall (1987) and Bentkus and Götze (1996), for instance. This begs the question whether the same phenomenon is also true for U-, L-statistics, and other classes of statistics.

In this paper we address this issue. We consider Edgeworth expansions for a very general class of statistics, and then apply them to some special cases of interest. In particular, we consider statistics of the form T_n/S_n, where

$$\begin{aligned} T_n &= n^{-1/2}\sum_{j=1}^{n}\alpha(X_j) + n^{-3/2}\sum_{i<j}\beta(X_i,X_j) + V_{1n}, \\ S_n^2 &= 1 + \frac{1}{n(n-1)}\sum_{i<j}\gamma(X_i,X_j) + V_{2n}, \end{aligned}$$

with V_{1n} and V_{2n} as remainder terms. One can think of T_n as the statistic of interest and S_n^2 as the normalizing variable. We refer to T_n/S_n as the *studentized* statistic when S_n is random, and as the *standardized* statistic when $S_n = 1$.

There are several reasons to consider such a class of statistics. First, a great many of commonly-seen statistics belong to this class. These include (standardized or studentized) U-, L-statistics, and function of sample means. Second, the approach taken in this paper unifies the treatment for both standardized and studentized statistics. For example, if $\gamma(x,y) = 0$ and V_{2n} is sufficiently small in the normalizing variable S_n^2, then the studentized statistic T_n/S_n will reduce to the standardized statistic T_n. Therefore, it is possible to derive asymptotic results for both cases under the same set of conditions. Finally, the conditions required for our main results are very weak, and often minimal; see Section 3.

Section 2 gives the main results of the paper. They will be applied in Section 3 to several well-known examples. Proofs and technical details are given in Sections 4 and 5.

Throughout, we use C to denote some positive constants, independent of n, which may be different at each occurrence. For a set B, let $I_{(B)}$ denote an indicator function of B. The standard normal density and distribution function are denoted by $\phi(x)$ and $\Phi(x)$, respectively. Finally, write

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

2. Main Results

Let $X, X_1, \ldots, X_n$ be a sequence of i.i.d. r.v.s. Let $\alpha(x)$, $\beta(x,y)$, $\gamma(x,y)$ be some Borel measurable functions of x, y, and z. Let $V_{in} \equiv V_{in}(X_1, \ldots, X_n)$ $(i = 1, 2)$ be functions of $\{X_1, \ldots, X_n\}$. Let

$$T_n = n^{-1/2} \sum_{j=1}^{n} \alpha(X_j) + n^{-3/2} \sum_{i<j} \beta(X_i, X_j) + V_{1n}, \tag{2.1}$$

$$S_n^2 = 1 + \frac{1}{n(n-1)} \sum_{i<j} \gamma(X_i, X_j) + V_{2n}. \tag{2.2}$$

The dominant term in T_n/S_n is $n^{-1/2} \sum_{j=1}^{n} \alpha(X_j)$, which converges in distribution to a normal distribution as $n \to \infty$ under weak conditions. We first give an Edgeworth expansion for T_n/S_n with remainder $o(n^{-1/2})$ under very weak conditions.

Theorem 2.1. *Assume the following.*
(a) $\alpha(X_1)$ *is nonlattice;* $\beta(x,y)$ *and* $\gamma(x,y)$ *are symmetric in* x, y.
(b) $E\alpha(X_1) = 0, \quad E\alpha^2(X_1) = 1, \quad E|\alpha(X_1)|^3 < \infty,$
$E[\beta(X_1, X_2)|X_1] = 0, \quad E|\beta(X_1, X_2)|^{5/3} < \infty,$
$E\gamma(X_1, X_2) = 0, \quad E|\gamma(X_1, X_2)|^{3/2} < \infty.$
(c) $P\left(|V_{jn}| \geq \delta_n n^{-1/2}\right) = o\left(n^{-1/2}\right)$, $j = 1, 2$, *where* $\delta_n > 0$ *and* $\delta_n \to 0$.

Then we have $\sup_x |P(T_n/S_n \leq x) - E_n(x)| = o\left(n^{-1/2}\right)$, *where*

$$E_n(x) = \Phi(x) - \frac{\Phi^{(3)}(x)}{6\sqrt{n}} \left(E\alpha^3(X_1) + 3E\alpha(X_1)\alpha(X_2)\beta(X_1, X_2)\right) - \frac{x\Phi^{(2)}(x)}{2\sqrt{n}} E\alpha(X_1)\gamma(X_1, X_2).$$

The next corollary may be easier to use in some applications.

Corollary 2.1. *Assume the conditions of Theorem* 2.1, *except with* S_n^2 *replaced by* $S_n^2 = 1 + n^{-3} \sum_{i \neq j \neq k} \eta(X_i, X_j, X_k) + V_{2n}$, *where* $E\eta(X_1, X_2, X_3) = 0$, $E|\eta(X_1, X_2, X_3)|^{3/2} < \infty$. *Then* $\sup_x |P(T_n/S_n) - E_n(x)| = o\left(n^{-1/2}\right)$, *where*

$$E_n(x) = \Phi(x) - \frac{\Phi^{(3)}(x)}{6\sqrt{n}} \left(E\alpha^3(X_1) + 3E\alpha(X_1)\alpha(X_2)\beta(X_1, X_2)\right) - \frac{x\Phi^{(2)}(x)}{2\sqrt{n}} \sum_{i=1}^{3} E\alpha(X_i)\eta(X_1, X_2, X_3).$$

Remark 2.1. If $\gamma(x,y)=0$ in Theorem 2.1, we get $\sup_x |P(T_n \le x) - E_{n0}(x)| = o\left(n^{-1/2}\right)$, where $E_{n0}(x)=\Phi(x)-6^{-1}n^{-1/2}\Phi^{(3)}(x)(E\alpha^3(X_1)+3E\alpha(X_1)\alpha(X_2)\beta(X_1,X_2))$. The Edgeworth expansions for both types of statistics thus hold under the same conditions if one can show that conditions for $\gamma(X_1,X_2)$ in T_n/S_n are implied by those imposed on $\alpha(X_1)$ and $\beta(X_1,X_2)$. Such examples can be seen in Sections 3.2 and 3.3.

Remark 2.2. The underlying distribution F of X_i's is typically unknown in practice, so one cannot directly apply Theorem 2.1 and Corollary 2.1 since they involve unknown quantities such as $E\alpha^3(X_1)$ and $E\alpha(X_1)\alpha(X_2)\beta(X_1,X_2)$, etc. In such situations, one could use empirical Edgeworth expansions (EEE) $\hat{E}_n(x)$ obtained by estimating the unknown quantities in $E_n(x)$ by their empirical versions (e.g., the jackknife).

Remark 2.3. Singh (1981) used a second-order Edgeworth expansion for the mean of a sample from a nonlattice distribution to show, for the first time, that the bootstrap distribution approximates the true distribution of the standardized statistic better than the normal approximation under a finite third moment. This classical result has been extended by Bloznelis and Putter (2003) to Student's t-statistic under the same optimal conditions. It would be of great interest to find out whether similar results hold under the conditions for the more general statistics considered in this paper. We hope to be able to report on this.

3. Some Important Applications

In this section, we apply the main results in Section 2 to three well-known examples: U- and L-statistics and functions of the sample mean, which results in second-order Edgeworth expansions under very weak, often optimal conditions.

3.1. *U*-statistics

Let $h(x,y)$ be a real-valued Borel measurable function, symmetric in its arguments, with $Eh(X_1,X_2)=\theta$. Define a U-statistic of degree 2 with kernel $h(x,y)$ by

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

Let $g(X_j)=E\left(h(X_i,X_j)-\theta \mid X_j\right)$, $\sigma_g^2 = Var\left(g(X_1)\right)$, and

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

Note that $n^2R_n^2$ is the jackknife estimator of $4\sigma_g^2$. Let the distributions of the standardized and studentized U-statistics be, respectively,

$$G_{1n}(x) = P\left(\frac{\sqrt{n}(U_n-\theta)}{2\sigma_g} \le x\right), \text{ and } G_{2n}(x) = P\left(\frac{\sqrt{n}(U_n-\theta)}{R_n} \le x\right).$$

Asymptotic normality of $G_{1n}(x)$ and $G_{2n}(x)$ was established by Hoeffding (1948) and Arvesen (1969), respectively, provided that $Eh^2(X_1,X_2) < \infty$ and $\sigma_g^2 > 0$. Berry-Esséen bounds for $G_{1n}(x)$ were studied by many authors, see Bentkus, Götze and Zitikis (1994) and references therein. Also see Wang and Weber (2006) for exact convergence rates and leading terms in the Central Limit Theorem. Berry-Esséen bounds for $G_{2n}(x)$ have been given by many authors; see Wang, Jing and Zhao (2000) for references. On the other hand, Edgeworth expansions for U-statistics have also been intensively studied in recent years. For $G_{1n}(x)$, under the conditions that $\sigma_g^2 > 0$, the d.f. of $g(X_1)$ is nonlattice, $E|g(X_1)|^3 < \infty$, and $E|h(X_1,X_2)|^{2+\epsilon} < \infty$ for some $\epsilon > 0$, Bickel Götze and van Zwet (1986) showed that

$$\sup_x |G_{1n}(x) - E_{n0}(x)| = o(n^{-1/2}), \tag{3.1}$$

where

$$E_{n0}(x) = \Phi(x) - \frac{\phi(x)}{6\sqrt{n}\sigma_g^3}\left(x^2-1\right)\left\{Eg^3(X_1) + 3Eg(X_1)g(X_2)h(X_1,X_2)\right\}.$$

Jing and Wang (2003) weakened the moment condition $E|h(X_1,X_2)|^{2+\epsilon} < \infty$ to $E|h(X_1,X_2)|^{5/3} < \infty$. On the other hand, under the conditions that $\sigma_g^2 > 0$, the d.f. of $g(X_1)$ is nonlattice, and $E|h(X_1,X_2)|^{4+\epsilon} < \infty$ for some $\epsilon > 0$, Helmers (1991) showed that

$$\sup_x |G_{2n}(x) - E_n(x)| = o(n^{-1/2}), \tag{3.2}$$

where

$$E_n(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{n}\sigma_g^3}\left\{(2x^2+1)Eg^3(X_1) + 3(x^2+1)Eg(X_1)g(X_2)h(X_1,X_2)\right\}.$$

Putter and van Zwet (1998) weakened the $(4+\epsilon)$-*th* moment condition of Helmers (1991) to $E|h(X_1,X_2)|^{3+\epsilon} < \infty$.

The following result can be obtained from Corollary 2.1.

Theorem 3.1. *Suppose that the d.f. of $g(X_1)$ is nonlattice and $\sigma_g^2 > 0$.*
(a) *If $E|g(X)|^3 < \infty$ and $E|h(X_1,X_2)|^{5/3} < \infty$, then* (3.1) *holds.*

(b) *If* $E|h(X_1,X_2)|^3<\infty$*, then* (3.2) *holds.*

Remark 3.1. The moment conditions $E|g(X_1)|^3<\infty$ and $E|h(X_1,X_2)|^{5/3}<\infty$ are optimal for Edgeworth expansion of standardized U-statistics; see Jing and Wang (2003). The conditions for Edgeworth expansion of studentized U-statistics are suboptimal in the following sense: the condition $E|h(X_1,X_2)|^3<\infty$ cannot be weakened to $E|h(X_1,X_2)|^{3-\delta}<\infty$ for any $\delta>0$, but it might be weakened to $E|g(X_1)|^3<\infty$, $E|h(X_1,X_2)|^{5/3}<\infty$, and $E|h(X_1,X_2)h(X_1,X_3)|^{3/2}<\infty$.

Remark 3.2. If $h(x,y)=(x+y)/2$, then $\sqrt{n}(U_n-\theta)/R_n$ reduces to Student's t-statistic. From Theorem 3.1, the moment condition for Edgeworth expansions of error size $o(n^{-1/2})$ for Student's t-statistic is $E|X_1|^3<\infty$, which is optimal; see Hall (1987).

Proof of Theorem 3.1. Part (a) can be proved by applying Theorem 2.1 directly. We prove part (b). Similar to (A_3) in Callaert and Veraverbeke (1981) (also see Serfling (1980)), we may write

$$\frac{\sqrt{n}(U_n-\theta)}{R_n}=\frac{\sqrt{n}(U_n-\theta)/(2\sigma_g)}{R_n/(2\sigma_g)}\equiv\frac{T_n}{S_n},$$

where T_n and S_n^2 are defined as in Corollary 2.1, with

$$\begin{aligned}
\alpha(X_j)&=\sigma_g^{-1}g(X_j),\\
\beta(X_i,X_j)&=\sigma_g^{-1}\left[h(X_i,X_j)-\theta-g(X_i)-g(X_j)\right],\\
\eta(X_i,X_j,X_k)&=\sigma_g^{-2}\left[h(X_i,X_j)-\theta\right]\left[h(X_i,X_k)-\theta\right]-1,\\
V_{1n}&=n^{-3/2}(n-1)^{-1}\sum_{i<j}\beta(X_i,X_j),\\
V_{2n}&=\frac{2\sigma_g^{-2}}{(n-1)(n-2)^2}\sum_{i<j}(h(X_i,X_j)-\theta)^2-\frac{n(n-1)\sigma_g^{-2}}{(n-2)^2}(U_n-\theta)^2\\
&\quad+\frac{(n-2)^2+4(n-1)}{n^3(n-1)(n-2)^2}\sum_{i\neq j\neq k}\eta(X_i,X_j,X_k)+\frac{2}{n-2}.
\end{aligned}$$

By the properties of conditional expectation, it can easily shown that

$$\begin{aligned}
E\alpha(X_i)\alpha(X_2)\beta(X_i,X_j)&=\sigma_g^{-3}Eg(X_1)g(X_2)h(X_1,X_2), &&\text{if } 1\le i\neq j\le 2,\\
E\alpha(X_i)\eta(X_1,X_2,X_3)&=\sigma_g^{-3}Eg^3(X_1), &&\text{if } i=1,\\
&=\sigma_g^{-3}Eg(X_1)g(X_2)h(X_1,X_2), &&\text{if } i=2,3.
\end{aligned}$$

In view of these estimates and the relations $\Phi^{(2)}(x)=-x\phi(x)$ and $\Phi^{(3)}(x)=(x^2-1)\phi(x)$, one can apply Corollary 2.1 to obtain

$$E_n(x)=\Phi(x)+\frac{\phi(x)}{6\sqrt{n}\sigma_g^3}\left\{(2x^2+1)Eg^3(X_1)+3(x^2+1)Eg(X_1)g(X_2)h(X_1,X_2)\right\}.$$

On the other hand, condition (a) of Corollary 2.1 can be easily checked. Then Theorem 3.1 follows from Corollary 2.1 if we can show $P\left(|V_{jn}| \geq n^{-1/2}(\log n)^{-1}\right) = o(n^{-1/2})$ for $j = 1, 2$, but this follows from Lemma 4.1 below.

3.2. *L*-statistics

Let $X_1, \ldots, X_n$ be i.i.d. r.v.s with distribution function F. Denote the empirical distribution by $F_n(x) = n^{-1}\sum_{j=1}^n I_{(X_i \leq x)}$. Let $J(t)$ be a real-valued function on $[0, 1]$ and $T(G) = \int xJ(G(x))\, dG(x)$. Then $T(F_n)$ is called an L-statistic; see Serfling (1980). Write $s \wedge t = \min\{s, t\}$, $s \vee t = \max\{s, t\}$, and

$$\sigma^2 \equiv \sigma^2(J, F) = \int\int J\left(F(s)\right) J\left(F(t)\right) F\left(s \wedge t\right)\left[1 - F\left(s \vee t\right)\right] dsdt,$$

A natural estimate of σ^2 is $\widehat{\sigma}^2 \equiv \sigma^2(J, F_n)$. Let the distributions of the standardized and studentized L-statistic $T(F_n)$ be, respectively,

$$L_{1n}(x) = P\left(\frac{\sqrt{n}(T(F_n) - T(F))}{\sigma} \leq x\right), \text{ and}$$
$$L_{2n}(x) = P\left(\frac{\sqrt{n}(T(F_n) - T(F))}{\widehat{\sigma}} \leq x\right).$$

Asymptotic normality of $L_{1n}(x)$ and $L_{2n}(x)$ holds if $E|X_1|^2 < \infty$, $\sigma^2 > 0$, and some smoothness conditions on $J(t)$; see Serfling (1980). Berry-Esséen bounds were given for them by Helmers (1977), van Zwet (1984), Helmers, Janssen and Serfling (1990), Helmers (1982), and Wang, Jing and Zhao (2000). Also see Wang and Weber for exact convergence rates and leading terms in the Central Limit Theorem. On the other hand, Edgeworth expansions for $L_{1n}(x)$ were given by Helmers (1982), Lai and Wang (1993), among others. Edgeworth expansions for $L_{2n}(x)$ were studied by Putter and van Zwet (1998) under some smoothness conditions on $J(t)$ and the moment condition $E|X_1|^{3+\epsilon} < \infty$ for some $\epsilon > 0$. The next theorem shows that this moment condition can be weakened.

Theorem 3.2. *Assume the following.*
(a) *$J''(t)$ is bounded on $t \in [0, 1]$.*
(b) *$\sigma^2 > 0$ and $\int[F(t)(1 - F(t))]^{1/3}dt < \infty$.*
(c) *the d.f. of $Y = \int J(F(s))\left(I_{(X_1 \leq s)} - F(s)\right)\, ds$ is nonlattice.*

Then we have

$$\sup_x \left|L_{1n}(x) - \widetilde{E}_{n0}(x)\right| = o(n^{-1/2}) \text{ and } \sup_x \left|L_{2n}(x) - \widetilde{E}_n(x)\right| = o(n^{-1/2}), \quad (3.3)$$

where $\widetilde{E}_{n0}(x) = \Phi(x) + 6^{-1}\sigma^{-3}n^{-1/2}\left(-3a_1\Phi^{(1)}(x) + a_3\Phi^{(3)}(x)\right)$ *and*

$$\widetilde{E}_n(x) = \Phi(x) + \frac{1}{6\sigma^3\sqrt{n}}\left(-3a_1\Phi^{(1)}(x) + 3a_2x\Phi^{(2)}(x) + a_3\Phi^{(3)}(x)\right).$$

Here, $J_0(t) = J(F(t))$,

$$\begin{aligned}
a_1 &= \sigma^2 \int J_0'(x)F(x)[1-F(x)]dx,\\
a_2 &= \iiint J_0(y)J_0(z)\left[J_0(x)K_1(x,y,z) + J_0'(x)K_2(x,y,z)\right] dxdydz,\\
a_3 &= \iiint J_0(x)J_0(y)\left[J_0(z)K_3(x,y,z) + 3J_0'(z)K_4(x,z)K_4(y,z)\right] dxdydz,\\
K_1(x,y,z) &= \left[F(x\wedge y\wedge z) - F(x\wedge y)F(z)\right]\left[(1-F(x\vee y)\right]\\
&\quad + F(x\wedge y)F(x\vee y)\left[F(z)-1\right],\\
K_2(x,y,z) &= F(x\wedge y)\left[1-F(x\vee y)\right]\left[F(y\wedge z) - F(y)F(z)\right],\\
K_3(x,y,z) &= F(x\wedge y\wedge z) - F(x)F(y\wedge z) - F(y)F(x\wedge z)\\
&\quad - F(z)F(x\wedge y) + 2F(x)F(y)F(z),\\
K_4(x,y) &= F(x\wedge y) - F(x)F(y).
\end{aligned}$$

Remark 3.3. Note that $\int[F(t)(1-F(t)]^{1/3}dt < \infty$ is weaker than $E|X_1|^{3+\epsilon} < \infty$ for every $\epsilon > 0$. To see this, first apply Markov's inequality to get $F(t)\{1 - F(t)\} \le E|X_1|^{3+\epsilon}/|t|^{3+\epsilon}$. Then,

$$\begin{aligned}
\int \{F(t)[1-F(t)]\}^{1/3}dt &\le \int_{|t|\le 1} 1dt + \int_{|t|>1} \{F(t)[1-F(t)]\}^{1/3}dt\\
&\le 2 + \left(E|X_1|^{3+\epsilon}\right)^{1/3}\int_{|t|>1} |t|^{-1-\epsilon/3}dt < \infty.
\end{aligned}$$

Proof of Theorem 3.2. We only prove the second relation in (3.3), the first can be done similarly. Write

$$\begin{aligned}
J_n(t) = J(F_n(t)), \qquad & Z(s,t,F) = F(s\wedge t)(1-F(s\vee t)),\\
\xi(X_i,X_j) &= \sigma^{-2}\iint J_0(s)J_0(t)\left(I_{(X_i\le s\wedge t)}I_{(X_j>s\vee t)} - Z(s,t,F)\right) dsdt,\\
\varphi(X_i,X_j,X_k) &= \sigma^{-2}\iint J_0'(s)J_0(t)\left[I_{(X_i\le t)} - F(t)\right] I_{(X_j\le s\wedge t)}I_{(X_k>s\vee t)}\, dsdt.
\end{aligned}$$

From Lemma B of Serfling (1980, p.265), we have $T(F_n)-T(F) = -\int[K_1(F_n(x)) - K_1(F(x))]dx$, where $K_1(t) = \int_0^t J(u)du$. We now can write

$$\frac{\sqrt{n}\left(T(F_n)-T(F)\right)}{\widehat{\sigma}} = \frac{\sqrt{n}\left(T(F_n)-T(F)\right)/\sigma}{\widehat{\sigma}/\sigma} \equiv \frac{T_n + n^{-1/2}E\beta(X_1,X_1)}{S_n},$$

where T_n and S_n^2 are defined as in Corollary 2.1, with

$$\begin{aligned}
\alpha(X_j) &= -\sigma^{-1}\int J_0(t)\left[I_{(X_j\le t)} - F(t)\right]dt,\\
\beta(X_i,X_j) &= -\sigma^{-1}\int J_0'(t)\left[I_{(X_i\le t)} - F(t)\right]\left[I_{(X_j\le t)} - F(t)\right]dt,\\
\eta(X_i,X_j,X_k) &= \xi(X_i,X_j) + \varphi(X_i,X_j,X_k),\\
|V_{1n}| &\le \frac{1}{n^{3/2}}\Big|\sum_{j=1}^n\left[\beta(X_j,X_j) - E\beta(X_1,X_1)\right]\Big| + Cn^{1/2}\int|F_n(t)-F(t)|^3dt,\\
V_{2n} &= Q_{1n}+Q_{2n}+Q_{3n},
\end{aligned}$$

where the Q_{in}'s are

$$\begin{aligned}
Q_{1n} &= 2\sigma^{-2}\iint\left[J_n(s)-J_0(s)-J_0'(s)(F_n(s)-F(s))\right]J_0(t)Z(s,t,F_n)dsdt,\\
Q_{2n} &= \sigma^{-2}\iint\left[J_n(s)-J_0(s)\right]\left[J_n(t)-J_0(t)\right]Z(s,t,F_n)dsdt,\\
Q_{3n} &= n^{-3}\sum_{j\ne k}\left[\xi(X_j,X_k)+\varphi(X_j,X_j,X_k)+\varphi(X_k,X_j,X_k)\right]\\
&\quad -n^{-1}\sigma^{-2}\iint F(s\wedge t)\left[1-F(s\vee t)\right]dsdt.
\end{aligned}$$

Conditions (a) and (b) in Corollary 2.1 can be easily checked. Let us check condition (c). It suffices to show that

$$\begin{aligned}
&P\left(|V_{1n}|\ge n^{-1/2}(\log n)^{-1}\right) = o(n^{-1/2}),\\
&P\left(|Q_{jn}|\ge n^{-1/2}(\log n)^{-1}\right) = o(n^{-1/2}),\quad \text{for } j=1,2,3,
\end{aligned}\tag{3.4}$$

For illustration, we only prove (3.4) for $j=1$. Others can be shown similarly. Noting that $Z(s,t,F_n)\le[F_n(s)(1-F_n(s))]^{1/2}[F_n(t)(1-F_n(t))]^{1/2}$, we have

$$\begin{aligned}
|Q_{1n}| &\le \sigma^{-2}\sup_{x,y}|J_0''(x)J_0(y)|\iint(F_n(s)-F(s))^2Z(s,t,F_n)dsdt\\
&\le C\sigma^{-2}\int(F_n(s)-F(s))^2ds\int F_n^{1/2}(t)(1-F_n(t))^{1/2}dt\\
&=: C\sigma^{-2}Q_{6n}\,Q_{7n},\qquad \text{say.}
\end{aligned}\tag{3.5}$$

Using the inequality $E|F_n(t)-F(t)|^k\le Cn^{-k/2}F(t)(1-F(t))$, we get

$$EQ_{6n}^3 = \iiint E\ \left\{(F_n(s)-F(s))^2(F_n(t)-F(t))^2(F_n(v)-F(v))^2\right\}dsdtdv$$

$$\leq \left(\int \left(E|F_n(s) - F(s)|^6 \right)^{1/3} ds \right)^3 \leq Cn^{-3} \left(\int F^{1/3}(s)(1 - F(s))^{1/3} ds \right)^3.$$

Similarly, we can show $EQ_{7n}^3 \leq \left(\int F^{1/3}(s)(1 - F(s))^{1/3} ds \right)^3$. Therefore,

$$\begin{aligned} P\left(|Q_{1n}| \geq n^{-1/2} (\log n)^{-1} \right) &\leq \left(n^{1/2} \log n \right)^{3/2} E|Q_{1n}|^{3/2} \\ &\leq C^{3/2} \sigma^{-3} \left(n^{1/2} \log n \right)^{3/2} \left(EQ_{6n}^3 \right)^{1/2} \left(EQ_{7n}^3 \right)^{1/2} \\ &\leq Cn^{-3/4} (\log n)^{3/2} \left(\int [F(t)(1 - F(t)]^{1/3} dt \right)^3 \\ &= o(n^{-1/2}). \end{aligned}$$

Finally, we can apply Corollary 2.1 to get $\sup_x |P(T_n/S_n \leq x) - E_n(x)| = o(n^{-1/2})$, From this, and using the similar method as in proof of Theorem 2.1, we can get

$$\begin{aligned} &\sup_x \left| P\left(\frac{T_n + n^{-1/2} E\beta(X_1, X_1)}{S_n} \leq x \right) - n^{-1/2} \phi(x) E\beta(X_1, X_1) - E_n(x) \right| \\ &= o(n^{-1/2}). \end{aligned}$$

This reduces to (3.3) after some tedious but routine calculations.

3.3. Functions of the sample mean

Let $X_1, \ldots, X_n$ be i.i.d. r.v.s with $EX_1 = \mu$ and $Var(X_1) = \sigma^2 < \infty$; let f be differentiable in a neighborhood of μ with $f'(\mu) \neq 0$. The asymptotic variance of $\sqrt{n} f(\overline{X})$ is $\sigma_f^2 = (f'(\mu))^2 \sigma^2$. Write $\overline{X} = n^{-1} \sum_{i=1}^n X_i$ and $\widehat{\sigma}^2 = n^{-1} \sum_{i=1}^n (X_i - \overline{X})^2$. A simple estimator of σ_f^2 is $|f'(\overline{X})|^2 \, \widehat{\sigma}$, and the jackknife variance estimator of σ_f^2 is

$$\widehat{\sigma}_f^2 = \frac{n-1}{n} \sum_{j=1}^n \left(f(\overline{X}^{(j)}) - f(\overline{X}) \right)^2, \qquad \text{where} \quad \overline{X}^{(j)} = \frac{1}{n-1} \left(\sum_{i=1}^n X_i - X_j \right).$$

Write the distributions of the standardized and studentized $f(\overline{X})$, respectively, as

$$H_{1n}(x) = P\left(\frac{\sqrt{n}(f(\overline{X}) - f(\mu))}{\sigma_f} \leq x \right), \quad H_{2n}(x) = P\left(\frac{\sqrt{n}(f(\overline{X}) - f(\mu))}{\widehat{\sigma}_f} \leq x \right).$$

Asymptotic properties of $H_{1n}(x)$ have been well studied (see Bhattacharya and Ghosh (1978) for instance). On the other hand, Miller (1964) showed that $\widehat{\sigma}_f^2$ is a consistent estimator of σ_f^2 and hence proved that $H_2(x)$ is asymptotically normal. Applying Bai and Rao (1991), Edgeworth expansions for $H_{1n}(x)$ with error size $o(n^{-1/2})$ hold under the minimal moment condition $E|X_1|^3 < \infty$. Putter and van Zwet (1998) gave Edgeworth expansions for $H_{2n}(x)$ under $E|X_1|^{3+\epsilon} < \infty$ for some $\epsilon > 0$. The next theorem gives the optimal moment condition.

Theorem 3.3. *Assume that $f^{(3)}(x)$ is bounded in a neighborhood of μ and $f'(\mu) \neq 0$, $E|X_1|^3 < \infty$, and the d.f. of X_1 is nonlattice. Then we have*

$$\sup_x \left| H_{1n}(x) - \widetilde{E}_{n0}(x) \right| = o\left(n^{-1/2}\right), \ \sup_x \left| H_{2n}(x) - \widetilde{E}_n(x) \right| = o\left(n^{-1/2}\right), \tag{3.6}$$

where $\rho = E(X_1 - \mu)^3/\sigma^3$, $b = 2^{-1}\sigma f''(u)/f'(u)$, and

$$\widetilde{E}_{n0}(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{n}} \left((1 - x^2)\rho + 6(2 - x^2)b\right),$$
$$\widetilde{E}_n(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{n}} \left((2x^2 + 1)\rho + 6b\right).$$

Proof. We only prove the second relation in (3.6), the first can be done similarly. Applying Taylor's expansion to $f(\overline{X}) - f(\mu)$ and $n\widehat{\sigma}_f^2$, we get

$$\frac{\sqrt{n}(f(\overline{X}) - f(\mu))}{\widehat{\sigma}_f} = \frac{T_n + n^{-1/2}b}{S_n},$$

where T_n and S_n^2 are defined in (2.1) and (2.2), with

$$\alpha(X_j) = \frac{X_j - \mu}{\sigma}, \qquad \beta(X_i, X_j) = 2b\alpha(X_i)\alpha(X_j),$$
$$\gamma(X_i, X_j) = \alpha^2(X_i) + \alpha^2(X_j) - 2 + 2b\sigma^{-1}\alpha(X_i)\alpha(X_j)\left[\alpha(X_i) + \alpha(X_j)\right],$$
$$|V_{1n}| \leq \frac{C}{n^{3/2}} \Big| \sum_{j=1}^n \left((X_j - \mu)^2 - \sigma^2\right) \Big| + C\sqrt{n}\left|\overline{X} - \mu\right|^3,$$
$$|V_{2n}| \leq C \sum_{k=2}^4 |\overline{X} - \mu|^k + \frac{C}{n^2} \sum_{j=1}^n \left((X_j - \mu)^2 + |X_j - \mu|^3\right) + \frac{C}{n^3} \sum_{j=1}^n (X_j - \mu)^4$$
$$+ \frac{C}{n}(\overline{X} - \mu)^2 \sum_{j=1}^n (X_j - \mu)^2 + \frac{C}{n^3} \Big| \sum_{i \neq j} \gamma(X_i, X_j) \Big|.$$

Conditions (a) and (b) in Theorem 2.1 can be easily checked. Condition (c) can be verified by applying Lemmas 4.1–4.2 in the Appendix. Applying Theorem 2.1, we get

$$\sup_x \left| P\left(\frac{T_n}{S_n} \leq x\right) - E_n(x) \right| = o(n^{-1/2}), \tag{3.7}$$

where $E_n(x)$ is given in Theorem 2.1. It follows from (3.7), and a similar method to the one used in the proof of Theorem 2.1, that

$$\sup_x \left| P\left(\frac{T_n + b/\sqrt{n}}{S_n} \le x \right) - \frac{b}{\sqrt{n}} \Phi^{(1)}(x) - E_n(x) \right| = o(n^{-1/2}).$$

Theorem 3.3 then follows from $\Phi^{(2)}(x) = -x\phi(x)$, $\Phi^{(3)}(x) = (x^2-1)\phi(x)$, and

$$E\alpha^3(X_1) = \rho, \quad E\alpha(X_1)\alpha(X_2)\beta(X_1, X_2) = 2b, \quad E\alpha(X_1)\gamma(X_2, X_1) = \rho + 2b.$$

4. Proof of Theorem 2.1

4.1. Some useful lemmas

Lemma 4.1. *Let $g(x_1, \ldots, x_k)$ be symmetric in its arguments. Assume that*

$$U_n(g) = \binom{n}{k}^{-1} \sum_{1 \le i_1 < \cdots < i_k \le n} g(X_{i_1}, \ldots, X_{i_k})$$

is a degenerate U-statistic of order m, i.e., $Eg(x_1, \ldots, x_m, X_{m+1}, \ldots, X_k)=0$.

(a) *If $E|g(X_1, \ldots, X_k)|^p < \infty$,*

$$E|U_n(g)|^p \le C\, n^{(m+1)(1-p)}, \quad \text{for } 1 \le p \le 2,$$
$$E|U_n(g)|^p \le C\, n^{-(m+1)p/2}, \quad \text{for } p \ge 2.$$

(b) *If $E|g(X_1, \ldots, X_k)|^2 < \infty$, then $P\left(|U_n(g)| \ge Cn^{-1/2}\right) = o(n^{-1/2})$ for $m \ge 1$.*

(c) *If $E|g(X_1, \ldots, X_k)|^{3/2} < \infty$, then*

$$P\left(|U_n(g)| \ge Cn^{1/2}(\log n)^{-1}\right) = o(n^{-1/2}), \qquad \text{for} \quad m = 0,$$
$$P\left(|U_n(g)| \ge Cn^{-3/10}\right) = o(n^{-1/2}), \qquad \text{for} \quad m = 1,$$
$$P\left(|U_n(g)| \ge Cn^{-1/2}(\log n)^{-1}\right) = o(n^{-1/2}), \quad \text{for} \quad m = 2.$$

(d) *If $E|g(X_1, \ldots, X_k)|^3 < \infty$, then for all $m \ge 0$ and $k > 0$,*

$$P\left(|U_n(g)| \ge Cn^{-1/4}(\log n)^{-k}\right) = o(n^{-1/2}).$$

The proof of (a) can be found in Theorems 2.1.3 and 2.1.4 of Koroljuk and Borovskich (1994). Others can be shown by (a) of the present lemma and Markov's inequality.

Lemma 4.2.

(a) *If $E|X_1| < \infty$, there exists a $\delta_n \to 0$ such that $P\left(\left|\bar{X}\right| \ge \delta_n \sqrt{n}\right) = o(n^{-1/2})$.*

(b) *If* $EX_1 = 0$ *and* $E|X_1|^{3/2} < \infty$*, there exists a* $\delta_n \to 0$ *such that* $P\left(\left|\bar{X}\right| \geq \delta_n\right) = o(n^{-1/2})$.

(c) *If* $E|X_1|^{3/4} < \infty$*, then* $P\left(\left|\bar{X}\right| \geq Cn\right) = o(n^{-1/2})$.

(d) *If* $EX_1 = 0$*,* $EX_1^2 = 1$*, and* $E|X_1|^3 < \infty$*, then* $\sup_{x^2 \geq 10 \log n} P(\left|\bar{X}\right| \geq |x|n^{-1/2}/3) = o(n^{-1/2})$.

Proof. To prove (a), let $\delta_n^3 = \max\left\{E|X_1|I_{(|X_1| \geq n^{1/4})}, \left(2n^{-1/2}E|X_1|\right)^3\right\}$. Since $E|X_1| < \infty$, we have $\delta_n \to 0$ and

$$\begin{aligned}
&P\left(\left|\bar{X}\right| \geq \delta_n\sqrt{n}\right)\\
&\quad \leq nP(|X_1| \geq n^{3/2})\\
&\qquad + P\Bigg(\frac{1}{n^{3/2}}\Big|\sum_{j=1}^{n}\left(X_jI_{(|X_j| \leq n^{3/2})} - EX_jI_{(|X_j| \leq n^{3/2})}\right)\Big| \geq \delta_n - \frac{E|X_1|}{n^{1/2}}\Bigg)\\
&\quad \leq n^{-1/2}E|X_1|I_{(|X_1| \geq n^{3/2})} + \frac{8}{n^2\delta_n^2}\left(EX_1^2I_{(|X_1| \leq n^{1/4})} + EX_1^2I_{(n^{1/4} < |X_1| \leq n^{3/2})}\right)\\
&\quad \leq o(n^{-1/2}) + \frac{8}{n(E|X_1|)^2}EX_1^2I_{(|X_1| \leq n^{1/4})} + 8\delta_n n^{-1/2} = o\left(n^{-1/2}\right).
\end{aligned}$$

This proves (a). Similarly, we can prove (b) and (c). We prove (d) next. From Chapter V of Petrov (1975), $\sup_x(1+|x|^3)\left|P\left(\sqrt{n}\bar{X} \leq x\right) - \Phi(x)\right| \leq Cn^{-1/2}E|X_1|^3$. From this and the inequality

$$1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \leq \frac{2}{1+|x|^3}, \quad \text{for} \quad x \geq 1, \tag{4.1}$$

we have $P\left(\sqrt{n}\bar{X} \geq |x|/3\right) \leq 1 - \Phi(|x|/3) + Cn^{-1/2}(\log n)^{-1} = o(n^{-1/2})$ for $n \geq 3$ and $x^2 \geq 10 \log n$. Similarly, $P\left(\sqrt{n}\bar{X} \leq -|x|/3\right) = o(n^{-1/2})$ for $n \geq 3$ and $x^2 \geq 10 \log n$. We have proved (d).

The proof of the next lemma can be found in Jing and Wang (2003).

Lemma 4.3. *Let* $V_n(x)$ *and* $W_n(x, y)$ *be real Borel-measurable functions and* $W_n(x, y)$ *be symmetric in its arguments. Assume the following.*

(a) *the d.f. of* $V_n(X_1)$ *is nonlattice for sufficiently large* n.

(b) $EV_n(X_1) = 0$, $EV_n^2(X_1) = 1$, $\sup_{n \geq 1} E|V_n(X_1)|^3 < \infty$.

(c) $E[W_n(X_1, X_2)|X_1] = 0$, $\sup_{n \geq 2} E|W_n(X_1, X_2)|^{5/3} < \infty$.

Then,

$$\sup_x \left|P\Big(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}V_n(X_j) + \frac{1}{n^{3/2}}\sum_{i<j}W_n(X_i, X_j) \leq x\Big) - E_n(x)\right| = o(n^{-1/2}),$$

where $\mathcal{L}_n(x) = n\left\{E\Phi\left(x - n^{-1/2}V_n(X_1)\right) - \Phi(x)\right\} - \Phi^{(2)}(x)/2$ *and*

$$E_n(x) = \Phi(x) + \mathcal{L}_n(x) - \frac{\Phi^{(3)}(x)}{2\sqrt{n}} EV_n(X_1)V_n(X_2)W_n(X_1, X_2).$$

The next lemma may be of independent interest.

Lemma 4.4. *Let $\xi_j(x)$, $\varphi_j(x,y)$, and $V_n \equiv V_n(X_1,\ldots,X_n)$ be real Borel measurable functions in their arguments, and $\varphi_j(x,y) = \varphi_j(y,x)$. Assume the following.*

(a) *The d.f. of $\xi_1(X_1)$ is nonlattice and $E\xi_1(X_1) = 0$, $E\xi_1(X_1)^2 = 1$, $E|\xi_1(X_1)|^3 < \infty$.*

(b) $E\xi_2(X_1) = 0$, $E|\xi_2(X_1)|^{3/2} < \infty$, $E|\xi_3(X_1)|^{3/4} < \infty$.

(c) $E(\varphi_j(X_1,X_2)|X_1) = 0$, $j = 1,2$; $E|\varphi_1(X_1,X_2)|^{5/3} < \infty$, $E|\varphi_2(X_1,X_2)|^{3/2} < \infty$.

(d) $P(|V_n| \geq o(n^{-1/2})) = o(n^{-1/2})$.

Let $\varsigma_{nj} = \xi_2(X_j) + n^{-1}\xi_3(X_j)$ *and* $\psi_{nij}(x) = \varphi_1(X_i,X_j) + xn^{-1/2}\varphi_2(X_i,X_j)$,

$$\begin{aligned} K_n(x) &= \frac{1}{\sqrt{n}}\sum_{j=1}^n \xi_1(X_j) + \frac{x}{n}\sum_{j=1}^n \varsigma_{nj} + \frac{1}{n^{3/2}}\sum_{i<j}\psi_{nij}(x), \\ E_K(x) &= \Phi(x) - 6^{-1}n^{-1/2}\Phi^{(3)}(x)\left(E\xi_1^3(X_1) + 3E\xi_1(X_1)\xi_1(X_2)\varphi_1(X_1,X_2)\right) \\ &\qquad + n^{-1/2}x\Phi^{(2)}(x)E\xi_1(X_1)\xi_2(X_1). \end{aligned}$$

Then, $\sup_x |P(K_n(x) \leq x(1+V_n)) - E_K(x)| = o(n^{-1/2})$ *as* $n \to \infty$, .

Proof. Without loss of generality, we assume that

$$|\varphi_2(X_i,X_j)| \leq 4n^2 \quad \text{for all} \ \ i, \ j. \tag{4.2}$$

For, if not, we can define

$$\begin{aligned} \varphi_3(X_i,X_j) &= \varphi_2(X_i,X_j)I_{(|\varphi_2|\leq n^2)} - E\varphi_2(X_i,X_j)I_{(|\varphi_2|\leq n^2)}, \\ \varphi_4(X_i,X_j) &= \varphi_3(X_i,X_j) - E(\varphi_3(X_i,X_j)|X_i) - E(\varphi_3(X_i,X_j)|X_j), \\ \widetilde{\psi}_{nij}(x) &= \varphi_1(X_i,X_j) + \frac{x}{\sqrt{n}}\varphi_4(X_i,X_j). \end{aligned}$$

Then we have

$$\frac{1}{n^{3/2}}\sum_{i<j}\psi_{nij}(x) = \frac{1}{n^{3/2}}\sum_{i<j}\widetilde{\psi}_{nij}(x) + xR_n^*, \quad \textit{say}.$$

Write $\delta_n^2 = E|\varphi_2(X_1, X_2)|^{3/2} I_{(|\varphi_2| \geq n^2)}$. Since $E\left[\varphi_2(X_1, X_2)|X_1\right] = 0$, we have

$$\begin{aligned} P\left(|R_n^*| \geq \frac{\delta_n}{\sqrt{n}}\right) &= P\Big(\frac{1}{n^2}\Big|\sum_{i<j}\left[\varphi_2(X_i, X_j) - \varphi_4(X_i, X_j)\right]\Big| \geq \frac{\delta_n}{\sqrt{n}}\Big) \\ &\leq 4\sqrt{n}\delta_n^{-1} E|\varphi_2(X_1, X_2)| I_{(|\varphi_2| \geq n^2)} \\ &\leq (4\sqrt{n}\delta_n^{-1}) n^{-1} E|\varphi_2(X_1, X_2)|^{3/2} I_{(|\varphi_2| \geq n^2)} \\ &\leq 4\delta_n n^{-1/2}. \end{aligned}$$

It is easy to show that $\delta_n \to 0$ and that $\varphi_4(x, y)$ is a symmetric function satisfying $E(\varphi_4(X_1, X_2)|X_1) = 0$, $E|\varphi_4(X_1, X_2)|^{3/2} < \infty$, and $|\varphi_4(X_i, X_j)| \leq 4n^2$. Therefore, if (4.2) does not hold, we can replace $\varphi_2(X_i, X_j)$ by $\varphi_4(X_i, X_j)$, and V_n by $V_n - R_n^*$.

For simplicity, take $V_n = 0$. Clearly, this will not affect the result since V_n only makes contribution of size $o(n^{-1/2})$ to the Edgeworth expansion. Write $\xi_2^*(X_j) = \xi_2(X_j) I_{(|\xi_2(X_j)| \leq n/(1+x^2))}$, $\xi_3^*(X_j) = \xi_3(X_j) I_{(|\xi_3(X_j)| \leq n^2/(1+x^2))}$, $\varsigma_{nj}^* = \xi_2^*(X_j) + \xi_3^*(X_j)/n$, and

$$K_n^*(x) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \xi_1(X_j) + \frac{x}{n} \sum_{j=1}^n \varsigma_{nj}^* + \frac{1}{n^{3/2}} \sum_{i<j} \psi_{nij}(x).$$

Then we have

$$\begin{aligned} &\sup_x \left|P\left(K_n(x) \leq x\right) - E_K(x)\right| \\ &\quad \leq \sup_{x^2 \geq 10 \log n} \left|P\left(K_n(x) \leq x\right) - E_K(x)\right| + \sup_{x^2 \leq 10 \log n} \left|P\left(K_n^*(x) \leq x\right) - E_K(x)\right| \\ &\qquad + \sup_{x^2 \leq 10 \log n} \left|P\left(K_n^*(x) \leq x\right) - P\left(K_n(x) \leq x\right)\right| \\ &\quad =: P_1 + P_2 + P_3, \qquad \text{say.} \end{aligned}$$

It suffices to show that $P_i = o(n^{-1/2})$ for $i = 1, 2, 3$.

(i) First, we prove $P_1 = o(n^{-1/2})$. Write $P_1 = P_1^+ + P_1^-$, where

$$\begin{aligned} P_1^+ &= \sup_{x \geq (10 \log n)^{1/2}} \left|P\left(K_n(x) \leq x\right) - E_K(x)\right| \\ &\leq \sup_{x \geq (10 \log n)^{1/2}} P\left(K_n(x) \geq x\right) + \sup_{x \geq (10 \log n)^{1/2}} |1 - E_K(x)| \\ &\leq \sup_{x \geq (10 \log n)^{1/2}} P\Big(\frac{1}{\sqrt{n}} \sum_{j=1}^n \xi_1(X_j) \geq \frac{x}{3}\Big) + P\Big(\frac{1}{n} \sum_{j=1}^n \varsigma_{nj} \geq \frac{1}{3}\Big) \\ &\qquad + \sup_{x \geq 1} P\Big(\frac{1}{n^{3/2}} \sum_{i<j} \psi_{nij}(x) \geq \frac{x}{3}\Big) + \sup_{x \geq (10 \log n)^{1/2}} |1 - E_K(x)| \end{aligned}$$

$$= o(n^{-1/2}),$$

where the last equality follows from Lemmas 4.1−4.2 and the inequality (4.1). Similarly, we can show that $P_1^- = o(n^{-1/2})$. This proves $P_1 = o(n^{-1/2})$.

(ii) We show that $P_3 = o(n^{-1/2})$ next. Clearly

$$\begin{aligned}
P_3 \le\ & \sup_{x^2\le 1} \left| P\left(K_n^*(x) \le x\right) - P\left(K_n(x) \le x\right) \right| \\
& + \sup_{1\le x^2\le 10\log n} P\left(K_n(x) \ge x,\ \varsigma_{nj} \ne \varsigma_{nj}^*,\ \text{for some } j\right) \\
& + \sup_{1\le x^2\le 10\log n} P\left(K_n^*(x) \ge x,\ \varsigma_{nj} \ne \varsigma_{nj}^*,\ \text{for some } j\right) \\
=:\ & \Omega_{n0} + \Omega_{n1} + \Omega_{n2}, \qquad \text{say.}
\end{aligned}$$

First consider Ω_{n0}. It is easy to see that

$$\begin{aligned}
P\left(\varsigma_{nj} \ne \varsigma_{nj}^*\right) &\le P\left(|\xi_2(X_1)| \ge \frac{n}{1+x^2}\right) + P\left(|\xi_3(X_1)| \ge \frac{n^2}{1+x^2}\right) \\
&\le \frac{1+|x|^{3/2}+|x|^3}{n^{3/2}} \Big(E|\xi_2(X_1)|^{3/2} I_{(|\xi_2(X_1)|\ge n/(1+x^2))} \\
&\qquad\qquad + E|\xi_3(X_1)|^{3/4} I_{(|\xi_3(X_1)|\ge n^2/(1+x^2))} \Big). \qquad (4.3)
\end{aligned}$$

It follows from (4.3) that $\Omega_{n0} \le \sup_{x^2\le 1} \sum_{j=1}^n P(\varsigma_{nj} \ne \varsigma_{nj}^*) = o(n^{-1/2})$.

Next we investigate Ω_{n1}. Without loss of generality, we assume that $x \ge 1$. Then in view of (4.3) and independence of X_k, we obtain

$$\begin{aligned}
& \sup_{1\le x^2\le 10\log n} P\Big(\frac{1}{\sqrt{n}} \sum_{k=1}^n \xi_1(X_k) \ge \frac{x}{3},\ \varsigma_{nj} \ne \varsigma_{nj}^*,\quad \text{for some } j\Big) \\
&\le \sum_{j=1}^n \sup_{1\le x^2\le 10\log n} P\Big(\frac{1}{\sqrt{n}} \sum_{k=1}^n \xi_1(X_k) \ge \frac{x}{3},\ \varsigma_{nj} \ne \varsigma_{nj}^*\Big) \\
&\le \sum_{j=1}^n P\Big(\frac{1}{\sqrt{n}} \xi_1(X_j) \ge \frac{1}{6}\Big) + \sum_{j=1}^n \sup_{1\le x^2\le 10\log n} P\Big(\frac{1}{\sqrt{n}} \sum_{\substack{k=1\\k\ne j}}^n \xi_1(X_k) \ge \frac{x}{6},\ \varsigma_{nj} \ne \varsigma_{nj}^*\Big) \\
&= o(n^{-1/2}) + \sum_{j=1}^n \sup_{1\le x^2\le 10\log n} P\Big(\frac{1}{\sqrt{n}} \sum_{\substack{k=1\\k\ne j}}^n \xi_1(X_k) \ge \frac{x}{6}\Big) P\Big(\varsigma_{nj} \ne \varsigma_{nj}^*\Big) \\
&= o(n^{-1/2}),
\end{aligned}$$

where we used the inequality: $P(n^{-1/2} \sum_{\substack{k=1\\k\ne j}}^n \xi_1(X_k) \ge x/6) \le Cx^{-3}$ for all

$1 \leq j \leq n$. From this and Lemmas 4.1-4.2, we get

$$\begin{aligned}\Omega_{n1} &\leq \sup_{x^2\geq 1} P\Big(\frac{1}{n^{3/2}}\sum_{i<j}\psi_{nij}(x) \geq \frac{x}{3}\Big) + P\Big(\frac{1}{n}\sum_{j=1}^{n}\varsigma_{nj} \geq \frac{1}{3}\Big) \\ &\quad + \sup_{1\leq x^2\leq 10\log n} P\Big(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\xi_1(X_k) \geq x/3,\ \varsigma_{nj} \neq \varsigma_{nj}^*,\ \text{for some } j\Big) \\ &= o(n^{-1/2}). \end{aligned} \tag{4.4}$$

Similarly, we can show $\Omega_{n2} = o(n^{-1/2})$. Thus, we have shown $P_3 = o(n^{-1/2})$.

(iii) Finally, we prove $P_2 = o(n^{-1/2})$. Write

$$\begin{aligned} Y_{nj}(x) &= \xi_1(X_j) + \frac{x}{\sqrt{n}}\left(\varsigma_{nj}^* - E\varsigma_{nj}^*\right), \\ \sigma_n^2(x) &= EY_{n1}^2(x), \qquad \theta_n(x) = \frac{x}{\sigma_n(x)}\left(1 - E\varsigma_{n1}^*\right). \\ \mathcal{L}_n(y) &= n\left\{E\Phi\left(y - \frac{Y_{n1}(x)}{\sqrt{n}\sigma_n(x)}\right) - \Phi(y)\right\} - \frac{1}{2}\Phi^{(2)}(y), \end{aligned}$$

and define

$$\begin{aligned} \widetilde{K_n^*}(x) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{1}{\sigma_n(x)}Y_{nj}(x) + \frac{1}{n^{3/2}}\sum_{i<j}\frac{1}{\sigma_n(x)}\psi_{nij}(x), \\ E_n(y) &= \Phi(y) + \mathcal{L}_n(y) - \frac{\Phi^{(3)}(y)}{2\sqrt{n}\sigma_n^3(x)}EY_{n1}(x)Y_{n2}(x)\psi_{n12}(x), \\ E_n^*(y) &= \Phi(y) - \frac{\Phi^{(3)}(y)}{6\sqrt{n}\sigma_n^3(x)}\left(EY_{n1}^3(x) + 3EY_{n1}(x)Y_{n2}(x)\psi_{n12}(x)\right). \end{aligned}$$

Then we have

$$\begin{aligned} P_2 &\leq \sup_{x^2\leq 10\log n}\sup_y\Big|P\left(\widetilde{K_n^*}(x) \leq y\right) - E_n(y)\Big| + \sup_{x^2\leq 10\log n}\sup_y|E_n(y) - E_n^*(y)| \\ &\quad + \sup_{x^2\leq 10\log n}|E_n^*(\theta_n(x)) - E_K(x)| \\ &=: I_{1n} + I_{2n} + I_{3n}, \qquad \text{say.} \end{aligned}$$

Thus, $P_2 = o(n^{-1/2})$ follows if we can show

$$I_{jn} = o(n^{-1/2}), \qquad \text{for } j = 1, 2, 3. \tag{4.5}$$

Under condition (b), we obtain that for all $x^2 \leq 10\log n$,

$$|E\xi_2^*(X_1)| \leq E|\xi_2(X_1)|I_{(|\xi_2(X_1)|\geq n/(1+x^2))} = o\left(\frac{1+|x|}{\sqrt{n}}\right), \tag{4.6}$$

$$\begin{aligned} E|\xi_2^*(X_1)|^\alpha &\le E|\xi_2(X_1)|^\alpha I_{(|\xi_2(X_1)|\le\sqrt{n})} + \left(\frac{n}{1+x^2}\right)^{\alpha-3/2} E|\xi_2(X_1)|^{3/2} I_{(|\xi_2(X_1)|\ge\sqrt{n})} \\ &= o\left(\frac{n}{1+x^2}\right)^{\alpha-3/2}, \qquad \text{for } \alpha > \tfrac{3}{2}. \end{aligned} \tag{4.7}$$

Similarly, we have

$$E|\xi_3^*(X_1)|^\alpha = o\left(\frac{n^2}{1+x^2}\right)^{\alpha-3/4}, \qquad \text{for } \alpha > \tfrac{3}{4}. \tag{4.8}$$

Recalling $\varsigma_{nj}^* = \xi_2^*(X_j) + n^{-1}\xi_3^*(X_j)$, it follows from (4.6)−(4.8) that for $x^2 \le 10\log n$,

$$|E\varsigma_{n1}^*| = \left|E\xi_2^*(X_1) + \frac{1}{n}E\xi_3^*(X_1)\right| = o\left(\frac{1+|x|}{\sqrt{n}}\right), \tag{4.9}$$

$$E|\varsigma_{n1}^*| \le E|\xi_2^*(X_1)| + \frac{1}{n}E|\xi_3^*(X_1)| = O(1), \tag{4.10}$$

$$\left(\frac{|x|}{\sqrt{n}}\right)^2 E(\varsigma_{n1}^*)^2 \le \frac{2x^2}{n}\left(E(\xi_2^*(X_1))^2 + \frac{E(\xi_3^*(X_1))^2}{n^2}\right) = o\left(\frac{1+|x|}{\sqrt{n}}\right), \tag{4.11}$$

$$\left(\frac{|x|}{\sqrt{n}}\right)^3 E|\varsigma_{n1}^*|^3 \le 6\left(\frac{|x|}{\sqrt{n}}\right)^3 \left(E|\xi_2^*(X_1)|^3 + \frac{1}{n^3}E|\xi_3^*(X_1)|^3\right) = o(1). \tag{4.12}$$

By using (4.9)−(4.12), together with Hölder's inequality, we get that if $x^2 \le 10\log n$, then

$$\sigma_n^2(x) = 1 + \frac{2x}{\sqrt{n}}E\xi_1(X_1)\xi_2(X_1) + o\left(\frac{1+|x|}{\sqrt{n}}\right), \tag{4.13}$$

$$\sigma_n^{-1}(x) = 1 - \frac{x}{\sqrt{n}}E\xi_1(X_1)\xi_2(X_1) + o\left(\frac{1+|x|}{\sqrt{n}}\right), \tag{4.14}$$

$$EY_{n1}^3(x) = E\xi_1(X_1)^3 + o(1), \qquad E|Y_{n1}(x)|^3 = O(1), \tag{4.15}$$

$$EY_{n1}(x)Y_{n2}(x)\psi_{n12}(x) = E\xi_1(X_1)\xi_1(X_2)\varphi_1(X_1,X_2) + o(1). \tag{4.16}$$

We only check (4.16) below, others can be checked similarly. Let $\mu_{nj}(x) = xn^{-1/2}\left(\varsigma_{nj}^* - E\varsigma_{nj}^*\right)$. It follows from (4.12) that $E|\mu_{n1}(x)|^3 = o(1)$. Then,

$$EY_{n1}(x)Y_{n2}(x)\psi_{n12}(x) = E\xi_1(X_1)\xi_1(X_2)\varphi_1(X_1,X_2) + B_1 + B_2$$

where, by noting independence of X_k and (4.15),

$$\begin{aligned} |B_1| &\le |E\{\xi_1(X_1)\mu_{n2}(x) + \mu_{n1}(x)Y_{n2}(x)\}\varphi_1(X_1,X_2)| \\ &\le 3\left(E|\mu_{n1}(x)|^3\right)^{1/3}\left(E|\xi_1(X_1)|^3 + E|Y_{n1}(x)|^3\right)^{1/3}\left(E|\varphi_1(X_1,X_2)|^{3/2}\right)^{2/3} \end{aligned}$$

$$
\begin{aligned}
&= o(1),\\
|B_2| &\le E|Y_{n1}(x)Y_{n2}(x)\varphi_2(X_1,X_2)| \le \left(E|Y_{n1}(x)|^3\right)^{2/3}\left(E|\varphi_2(X_1,X_2)|^{3/2}\right)^{2/3}\\
&= O(1).
\end{aligned}
$$

This proves (4.16).

We turn back to the proof of (4.5). To show $I_{3n} = o(n^{-1/2})$, it suffices to show

$$
\sup_{x^2\le 10\log n}\left|\Phi(\theta_n(x)) - \Phi(x) - \frac{x\Phi^{(2)}(x)}{\sqrt{n}}E\xi_1(X_1)\xi_1(X_2)\right| = o(n^{-1/2}), \tag{4.17}
$$

$$
\sup_{x^2\le 10\log n}\left|\Phi^{(3)}(\theta_n(x)) - \Phi^{(3)}(x)\right| = O(n^{-1/2}), \tag{4.18}
$$

$$
\sup_{x^2\le 10\log n}\left|\frac{EY_{n1}^3(x)}{\sigma_n^3(x)} - E\xi_1(X_1)^3\right| = o(1), \tag{4.19}
$$

$$
\sup_{x^2\le 10\log n}\left|\frac{EY_{n1}(x)Y_{n2}(x)\psi_{n12}(x)}{\sigma_n^3(x)} - E\xi_1(X_1)\xi_1(X_2)\varphi_1(X_1,X_2)\right| = o(1). \tag{4.20}
$$

Clearly, (4.18)$-$(4.20) follow easily from (4.13)$-$(4.16). Now let us check (4.17). Using (4.9) and (4.14), we have that, for all $x^2 \le 10\log n$,

$$
\theta_n(x) = \frac{x}{\sigma_n(x)}\left(1 - E\varsigma_{n1}^*\right) = x\left(1 - \frac{x}{\sqrt{n}}E\xi_1(X_1)\xi_1(X_2)\right) + o\left(\frac{1+x^2}{\sqrt{n}}\right).
$$

Hence, for sufficiently large n, we have $x/2 \le \theta_n(x) \le 3x/2$. From these and Taylor's expansion, there exists $1/2 \le \delta \le 3/2$ such that for all $x^2 \le 10\log n$,

$$
\begin{aligned}
\Phi(\theta_n(x)) &= \Phi(x) + (\theta_n(x) - x)\phi(x) + \frac{(\theta_n(x)-x)^2}{2}\Phi^{(2)}(\delta x)\\
&= \Phi(x) - \frac{x^2\phi(x)}{\sqrt{n}}E\xi_1(X_1)\xi_2(X_1) + o(n^{-1/2})f(x)\phi(\frac{x}{2}),
\end{aligned}
$$

where $f(x)$ is a polynomial in x. Since $\Phi^{(2)}(x) = -x\phi(x)$, (4.17) is shown. Thus, $I_{3n} = o(n^{-1/2})$.

Next we show that $I_{2n} = o(n^{-1/2})$. Note that

$$
\begin{aligned}
&\sup_y |E_n(y) - E_n^*(y)|\\
&= \sup_y\left|n\left\{E\Phi\left(y - \frac{Y_{n1}(x)}{\sqrt{n}\sigma_n(x)}\right) - \Phi(y)\right\} - \frac{1}{2}\Phi^{(2)}(y) + \frac{EY_{n1}^3(x)}{6\sqrt{n}\sigma_n^3(x)}\Phi^{(3)}(y)\right|\\
&\le \frac{C}{\sqrt{n}\sigma_n^3(x)}E|Y_{n1}(x)|^3 I_{(|Y_{n1}(x)|\ge\sqrt{n}\sigma_n(x))} + \frac{C}{n\sigma_n^4(x)}E|Y_{n1}(x)|^4 I_{(|Y_{n1}(x)|\le\sqrt{n}\sigma_n(x))},
\end{aligned}
$$

where the last inequality follows from Theorem 3.2 of Hall (1982). It follows from (4.13) that for sufficiently large n and all $x^2 \leq 10 \ \log n$,

$$\frac{1}{2} < |\sigma_n(x)| < \frac{3}{2}. \tag{4.21}$$

It follows from (4.9) that for sufficiently large n and all $x^2 \leq 10 \ \log n$,

$$\begin{aligned} |Y_{n1}(x)| &= \left|\xi_1(X_1) + \frac{x}{\sqrt{n}}\left(\varsigma_{n1}^* - E\varsigma_{n1}^*\right)\right| \\ &\leq 1 + |\xi_1(X_1)| + |\xi_2(X_1)|^{1/2} + |\xi_3(X_1)|^{1/4} =: \kappa(X_1), \qquad \text{say.}(4.22) \end{aligned}$$

Noting (4.21) and $E\kappa^3(X_1) < \infty$, we get for sufficiently large n,

$$\begin{aligned} I_{2n} &= \sup_{x^2 \leq 10 \log n} \sup_y |E_n(y) - E_n^*(y)| \\ &\leq Cn^{-1/2} \sup_{x^2 \leq 10 \log n} \left(E|Y_{n1}(x)|^3 I_{(|Y_{n1}(x)|\geq n^{1/4})} + \frac{1}{\sqrt{n}} E|Y_{n1}(x)|^4 I_{(|Y_{n1}(x)|\leq n^{1/4})} \right) \\ &\leq Cn^{-1/2} \left(E\kappa^3(X_1) I_{(\kappa(X_1)\geq n^{1/4})} + \frac{1}{n^{1/4}} E\kappa^3(X_1) \right) = o(n^{-1/2}). \end{aligned}$$

Finally we use Lemma 4.3 to show that $I_{1n} = o(n^{-1/2})$ by taking $V_n(X_j) = Y_{nj}(x)/\sigma_n(x), W_n(X_i, X_j) = \psi_{nij}(x)/\sigma_n(x)$. First, we check condition (a) of Lemma 4.3. By Theorem 1.3 of Petrov (1995), a d.f. with the characteristic function $f(t)$ is a nonlattice d.f. if and only if for every fixed number $s_0 \neq 0$, we have $|f(s_0)| < 1$ or, equivalently, if and only if $\sup_{\delta \leq t \leq t_0} |f(t)| < 1$ for any $t_0 > 0$ and $\delta > 0$. Hence, to show that $V_n(X_j)$ is nonlattice uniformly for $x^2 \leq 10 \log n$ for sufficiently large n, we only need to prove $b_n =: \sup_{x^2 \leq 10 \log n} \sup_{\delta \leq |t| \leq t_0} \left|Ee^{itV_n(X_j)}\right| < 1$ for sufficiently large n and each $\delta > 0$. Noting that $\sup_{\delta \leq |t| \leq t_0} \left|Ee^{it\xi_1(X_1)}\right| < d < 1$, and from (4.10) and (4.14), we have

$$\begin{aligned} b_n &\leq \sup_{x^2 \leq 10 \log n} \sup_{\delta \leq |t| \leq t_0} \left|Ee^{itV_n(X_j)} - Ee^{it\xi_1(X_1)}\right| + \sup_{\delta \leq |t| \leq t_0} \left|Ee^{it\xi_1(X_1)}\right| \\ &\leq t_0 \sup_{x^2 \leq 10 \log n} E\left|V_n(X_j) - \xi_1(X_1)\right| + d \\ &\leq t_0 \sup_{x^2 \leq 10 \log n} \left(\left|\sigma_n(x)^{-1} - 1\right| + \frac{x}{\sqrt{n}} E\left|\varsigma_{n1}^* - E\varsigma_{n1}^*\right| \right) + d \\ &= d + o(1) < 1, \qquad \text{for sufficiently large } n. \end{aligned}$$

Condition (a) of Lemma 4.3 also holds here.

To check condition (b) of Lemma 4.3, it is easy to see that $EV_n(X_1) = 0$ and $EV_n^2(X_1) = 1$. For $x^2 \leq 10 \log n$, from (4.21)–(4.22), we have $\sup_n E|V_n(X_1)|^3 < \infty$.

To check condition (c) of Lemma 4.3, it is easy to see that $E[W_n(X_1, X_2)|X_1] = 0$. For $x^2 \le 10\log n$, by Minkowski's inequality and the assumptions, we get

$$\begin{aligned} E|\psi_{nij}(x)^{5/3}| &\le CE|\varphi_1(X_i, X_j)|^{5/3} + \frac{Cx^{5/3}}{n^{5/6}} E|\varphi_2(X_i, X_j)|^{5/3} I\{|\varphi_2(X_i, X_j)| \le 4n^2\} \\ &\le CE|\varphi_1(X_i, X_j)|^{5/3} + \frac{Cx^{5/3}}{n^{1/2}} E|\varphi_2(X_i, X_j)|^{3/2} \le C, \end{aligned}$$

which, together with (4.21), yields $\sup_{n\ge 2} E|W_n(X_i, X_j)^{5/3}| < \infty$.

Hence, applying Lemma 4.3, we have that, for all sufficiently large n,

$$I_{1n} = \sup_{x^2\le 10\log n} \sup_y \left| P\left(\widetilde{K_n^*}(x) \le y\right) - E_n(y)\right| = o(n^{-1/2}).$$

Then, $P_2 = o(n^{-1/2})$. The proof of Lemma 4.4 is complete.

4.2. Proof of Theorem 2.1

Without loss of generality, we assume $V_{jn} = 0$. It will be clear that this assumption does not affect the proof of the main results since their contributions to the Edgeworth expansion is only of size $o(n^{-1/2})$. Let $\gamma_1(x) = E\gamma(x, X_1)$ and $\gamma_2(x, y) = \gamma(x, y) - \gamma_1(x) - \gamma_1(y)$. It is easy to show that

$$S_n^2 = 1 + \frac{1}{n}\sum_{j=1}^n \gamma_1(X_j) + \frac{1}{n(n-1)}\sum_{i<j}\gamma_2(X_i, X_j) =: 1 + Z_n + R_n, \quad \text{say.} \quad (4.23)$$

Noting that $1 + u/2 - u^2/6 \le (1+u)^{1/2} \le 1 + u/2 + u^2/6$ for $|u| \le 1/9$, if $|Z_n + R_n| \le 1/9$, we have

$$1 + \frac{1}{2}(Z_n + R_n) - \frac{1}{3}\left(Z_n^2 + R_n^2\right) \le S_n \le 1 + \frac{1}{2}(Z_n + R_n) + \frac{1}{3}\left(Z_n^2 + R_n^2\right). (4.24)$$

Put $\Delta_n(s) = Z_n/2 + (n - 1/2n)R_n + sZ_n^2$. Then from (4.23) and (4.24), we have

$$\begin{aligned} P\left(\frac{T_n}{S_n} \le x\right) &\le P\left(\frac{T_n}{S_n} \le x, |Z_n + R_n| \le \frac{1}{9}\right) + P(|Z_n + R_n| \ge \frac{1}{9}) \\ &\le P\left(T_n \le x\left\{1 + \Delta_n(\frac{1}{3}) + \frac{R_n}{2n} + \frac{R_n^2}{3}\right\}\right) + P(|Z_n + R_n| \ge \frac{1}{9}) \\ &\le P\left(T_n \le x\left\{1 + \Delta_n(\frac{1}{3}) + n^{-3/5}\right\}\right) \\ &\quad + P(|Z_n + R_n| \ge \frac{1}{9}) + P\left(\left|\frac{R_n}{2n} + \frac{R_n^2}{3}\right| \ge n^{-3/5}\right). \end{aligned} \quad (4.25)$$

Similarly, we get

$$P\left(T_n/S_n \le x\right) \ge P\left(T_n \le x\left\{1 + \Delta_n(-\frac{1}{3}) - n^{-3/5}\right\}\right)$$

$$-P(|Z_n + R_n| \geq \frac{1}{9}) - P\left(\left|\frac{R_n}{2n} - \frac{R_n^2}{3}\right| \geq n^{-3/5}\right). \quad (4.26)$$

Using Jensen's inequality, we can easily see that $E|\gamma_1(X_1)|^{3/2} < \infty$ and $E|\gamma_2(X_1, X_2)|^{3/2} < \infty$. Since $2R_n$ is a degenerate U-statistic of order 1, it follows from Lemmas 4.1−4.2 that

$$P\left(\left|\frac{1}{2n}R_n \pm \frac{1}{3}R_n^2\right| \geq n^{-3/5}\right) \leq P(|R_n| \geq 1) + P(|R_n| \geq n^{-3/10}) = o(n^{-1/2}),$$

$$P(|Z_n + R_n| \geq \frac{1}{9}) \leq P\Big(|Z_n| \geq \frac{1}{18}\Big) + P\Big(|R_n| \geq \frac{1}{18}\Big) = o(n^{-1/2}).$$

In view of these inequalities, (4.25) and (4.26), Theorem 2.1 follows if

$$\sup_x |P\left(T_n \leq x\left(1 + \Delta_n(s) + A_n\right)\right) - E_n(x)| = o(n^{-1/2}), \quad (4.27)$$

where $|A_n| \leq n^{-3/5}$ and $|s| \leq 1/3$. To prove (4.27), let

$$\varsigma_{nj}(x) = -\left(\frac{1}{2}\gamma_1(X_j) + \frac{s}{n}\gamma_1^2(X_j)\right),$$

$$\psi_{nij}(x) = \beta(X_i, X_j) - \frac{x}{\sqrt{n}}\left(2s\gamma_1(X_i)\gamma_1(X_j) + \frac{1}{2}\gamma_2(X_i, X_j)\right).$$

An elementary calculation shows that

$$P\left(T_n \leq x\left(1 + \Delta_n(s) + A_n\right)\right)$$
$$= P\bigg(\frac{1}{\sqrt{n}}\sum_{j=1}^n \alpha(X_j) + \frac{x}{n}\sum_{j=1}^n \varsigma_{nj} + \frac{1}{n^{3/2}}\sum_{i<j}\psi_{nij}(x) \leq x\left(1 + A_n\right)\bigg).$$

It is easy to check that conditions of Lemma 4.4 are satisfied with $|s| \leq 1/3$, $V_n = n^{-3/5}$, and with $\xi_1(X_j) = \alpha(X_j)$, $\xi_2(X_j) = -\gamma_1(X_j)/2$, $\xi_3(X_j) = -s\gamma_1^2(X_j)$, and

$$\varphi_1(X_i, X_j) = \beta(X_i, X_j), \quad \varphi_2(X_i, X_j) = -\left(2s\gamma_1(X_i)\gamma_1(X_j) + \frac{1}{2}\gamma_2(X_i, X_j)\right).$$

So (4.27) follows from Lemma 4.4 and the relation $E\alpha(X_1)\gamma_1(X_1) = E\alpha(X_1)\gamma(X_1, X_2)$.

Acknowledgements

Jing's research is partially supported by Hong Kong RGC grants HKUST6011/07P and HKUST6015/08P. Wang's research is partially supported by an Australian Research Council discovery project. We thank the Editors and three referees for very useful comments and constructive criticisms which have greatly improved the paper.

References

Arvesen, J. N. (1969). Jackknifing U-statistics. *Ann. Math. Statist.* **40**, 2076-2100.

Bai, Z. and Rao, C. R. (1991). Edgeworth expansions of a function of sample means. *Ann. Statist.* **19**, 1295-1315.

Bentkus, V. and Götze, F. (1996). Optimal rates of convergence in the CLT for quadratic forms. *Ann. Probab.* **24**, 466-490.

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

Bentkus, V., Götze, F. and van Zwet, W. R. (1997). An Edgeworth expansion for symmetric statistics. *Ann. Statist.* **25**, 851-896.

Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6**, 434-451.

Bickel, P. J. Götze, F. and van Zwet, W. R. (1986). The Edgeworth expansion for U-Statistics of degree two. *Ann. Statist.* **14**, 1463-1484.

Bloznelis, M. and Putter, H. (2003). Second order and bootstrap approximation to Student's t statistic. *Theory Probab. Appl.* **47**, 300–307

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

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

Hall, P. (1987). Edgeworth expansion for student's t-statistic under minimal conditions. *Ann. Probab.* **15**, 920-931.

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

Helmers, R. (1982). *Edgeworth Expansions for Linear Combinations of Order Statistics.* Mathematical Centre Tracts 105, Amsterdam.

Helmers, R. (1991). On the Edgeworth expansion and the Bootstrap approximation for a studentized U-statistic. *Ann. Statist.* **19**, 470-484.

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

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

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

Koroljuk, V. S. and Borovskich, Yu. V. (1994). *Theory of U-statistics, Mathematics and Its Applications*, Vol. 273. Kluwer Academic Publishers.

Lai, T. L. and Wang, J. Q. (1993). Edgeworth Expansions for symmetric statistics with applications to bootstrap methods. *Statistica Sinica* **3**, 517-542.

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

Petrov, V. V. (1975). *Sums of Independent Random Variables.* Springer-Verlag, Berlin.

Petrov, V. V. (1995). *Limit Theorems of Probability Theory.* Clarendon Press, Oxford.

Putter, H. and van Zwet, W. R. (1998). Empirical Edgeworth expansions for symmetric statistics. *Ann. Stat.* **26**, 1540-1569.

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

Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9**, 1187-1195.

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. (2006). Exact convergence rate and leading term in the central limit theorem for U-statistics, *Statist. Sinica* **16**, 1409-1422.

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

Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong.

E-mail: majing@ust.hk

School of Mathematics and Statistics, University of Sydney, NSW, Australia.

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

(Received December 2007; accepted December 2008)