

EDGEWORTH EXPANSION FOR ONE-SAMPLE U -STATISTICS

Maesono, Yoshihiko
Department of Mathematics, Faculty of Science, Kagoshima University

<https://doi.org/10.5109/13386>

出版情報 : Bulletin of informatics and cybernetics. 22 (3/4), pp.189-197, 1987-03. Research Association of Statistical Sciences

バージョン :

権利関係 :



EDGEWORTH EXPANSION FOR ONE-SAMPLE U -STATISTICS

By

Yoshihiko MAESONO*

Abstract

Under some regularity conditions on kernel, an asymptotic expansion with remainder term $o(N^{-1})$ is established for one-sample U -statistics with kernel of arbitrary degree. This is an extension of the result by Callaert, Janssen and Veraverbeke [1].

1. Introduction

Let X_1, X_2, \dots, X_N , be independently and identically distributed random variables with common distribution function F . Let h be symmetric function of its arguments, satisfying $Eh(X_1, X_2, \dots, X_r)=0$ with $r \leq N$, h is called a kernel and r is called its degree. We shall define a one-sample U -statistic with a kernel of degree r , h , by

$$U_N = \binom{N}{r}^{-1} \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq N} h(X_{i_1}, X_{i_2}, \dots, X_{i_r}).$$

In the case of degree two, Callaert, Janssen and Veraverbeke [1] have obtained the asymptotic expansion of the distribution of U_N with the remainder term $o(N^{-1})$. In this paper, using the forward martingale characterization of U_N , we obtain an asymptotic expansion of U_N with a kernel of arbitrary degree r . In Section 2 we obtain a representation for U_N in terms of forward martingales, and get the bounds of absolute moments of martingales. We state the main theorem in Section 3 and prove it in Section 4.

2. Preliminaries

We shall represent U_N in terms of forward martingales. This representation is due to Hoeffding [5] (cf. Serfling [7] p.178). Under the assumption $E|h(X_1, X_2, \dots, X_r)| < \infty$, let us define the following notations: for $1 \leq k \leq r$

[illegible]

* Department of Mathematics, Faculty of Science, Kagoshima University, Kagoshima 890, Japan.

and

$$A_{k,N} = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq N} g_k(X_{i_1}, X_{i_2}, \dots, X_{i_k}), \quad \text{for } 1 \leq k \leq r.$$

Then U_N can be rewritten as

$$U_N = \binom{N}{r}^{-1} \sum_{k=1}^r \binom{N-k}{r-k} A_{k,N}.$$

It is shown in the proof of Lemma 2 that $\{A_{k,N}\}_{N \geq k}$ is a forward martingale for each $k=1, 2, \dots, r$.

By the definition of g_k , it is easily shown that if one of $\{i_1, i_2, \dots, i_k\}$ is not contained in $\{j_1, j_2, \dots, j_s\}$, then

$$E\{g_k(X_{i_1}, X_{i_2}, \dots, X_{i_k}) | X_{j_1}, X_{j_2}, \dots, X_{j_s}\} = 0. \quad (2.1)$$

Using this property, we can prove the useful two lemmas.

LEMMA 1. *If $E|g_k(X_1, X_2, \dots, X_k)| < \infty$ and one of $\{i_1, i_2, \dots, i_k\}$ is not contained in $\{j_1, j_2, \dots, j_s\}$, then for f satisfying $E|fg_k| < \infty$,*

$$E\{f(X_{j_1}, X_{j_2}, \dots, X_{j_s}) g_k(X_{i_1}, X_{i_2}, \dots, X_{i_k})\} = 0.$$

PROOF. Taking the conditional expectation, we have the desired result from (2.1).

Before describing the next lemma, we prepare the notations. For $1 \leq N_1 < N_2 < \dots < N_k \leq N$ and $1 \leq k \leq r$, let us define

$$B_k(N_1, N_2, \dots, N_k) = \sum_{i_1=1}^{N_1} \sum_{i_2=i_1+1}^{N_2} \dots \sum_{i_k=i_{k-1}+1}^{N_k} g_k(X_{i_1}, X_{i_2}, \dots, X_{i_k}).$$

Then we have the upper bound of the p th absolute moment of B_k .

LEMMA 2. *Given the existence of the p th ($p \geq 2$) absolute moment of kernel h , there exist a positive constant C such that*

$$E|B_k(N_1, N_2, \dots, N_k)|^p \leq C(\prod_{i=1}^k N_i)^{p/2}. \quad (2.2)$$

If the second moment of kernel h is finite, the inequality (2.2) holds also with $p=1$.

PROOF. The latter part of the lemma immediately follows from the former. Therefore we consider the case $p \geq 2$. By induction on s we prove the following inequality for $1 \leq s \leq k$ and $1 \leq m_1 < m_2 < \dots < m_s < i_{s+1}, \dots, i_k$,

$$\begin{aligned} E|\sum_{i_1=1}^{m_1} \sum_{i_2=i_1+1}^{m_2} \dots \sum_{i_s=i_{s-1}+1}^{m_s} g_k(X_{i_1}, X_{i_2}, \dots, X_{i_s}, \dots, X_{i_k})|^p \\ \leq (C_p)^s E|g_k(X_1, X_2, \dots, X_k)|^p (\prod_{i=1}^s m_i)^{p/2} \end{aligned} \quad (2.3)$$

where $C_p = \{8(p-1) \max(1, 2^{p-3})\}^p$.

When $s=1$, let $Y_j = \sum_{i_1=1}^{j_1} g_k(X_{i_1}, X_{i_2}, \dots, X_{i_k})$ for $j=1, 2, \dots, m_1$. Then for $j=1, 2, \dots, m_1$, we have $Y_j - Y_{j-1} = g_k(X_j, X_{i_2}, \dots, X_{i_k})$ and $j < i_2, i_3, \dots, i_k$, where $Y_0=0$. Since Y_1, Y_2, \dots, Y_{j-1} are functions of $X_1, X_2, \dots, X_{j-1}, X_{i_2}, \dots, X_{i_k}$, we find from (2.1) that

$$\begin{aligned} E\{Y_j - Y_{j-1} | Y_1, Y_2, \dots, Y_{j-1}\} \\ = E(E\{g_k(X_j, X_{i_2}, \dots, X_{i_k}) | X_1, X_2, \dots, X_{j-1}, X_{i_2}, \dots, X_{i_k}\} | Y_1, Y_2, \dots, Y_{j-1}) = 0. \end{aligned}$$

Therefore $\{Y_j\}_{0 \leq j \leq m_1}$ is a forward martingale. Applying an upper bound for moments of martingales obtained by Dharmadhikari, Fabian and Jogdeo [2], we have the inequality (2.3) when $s=1$.

Assume that (2.3) holds for $1, 2, \dots, s-1$ instead of s . For $s \leq j \leq m_s$, put $\tilde{m}_i(j) = \min(m_i, j-s+i)$ (for $i=1, 2, \dots, s-1$) and

$$Z_j = \sum_{i_1=1}^{\tilde{m}_1(j)} \dots \sum_{i_{s-1}=i_{s-2}+1}^{\tilde{m}_{s-1}(j)} \sum_{i_s=i_{s-1}+1}^j g_k(X_{i_1}, \dots, X_{i_{s-1}}, X_{i_s}, \dots, X_{i_k}).$$

Then in the same way of $s=1$, $\{Z_j\}_{0 \leq j \leq m_s}$ is a forward martingale, where $Z_j=0$ for $0 \leq j < s$. Hence using the result of Dharmadhikari et al. [2], we get the inequality (2.3) for s . Thus we have the desired result.

3. Main Theorem

Before we state the main theorem, we define the following notations:

$$\xi_k^2 = E g_k^2(X_1, X_2, \dots, X_k) \quad \text{for } 1 \leq k \leq r,$$

$$\sigma_N^2 = E\{U_N\}^2 = \frac{r^2}{N} \xi_1^2 + \frac{\{r(r-1)\}^2}{2N(N-1)} \xi_2^2 + \dots + \frac{r!}{N(N-1) \dots (N-r+1)} \xi_r^2,$$

$$\eta(t) = E(\exp\{itg_1(X_1)\}),$$

$$\zeta(x, y) = w_2(x, y) - \frac{r-2}{r-1} \{g_1(x) + g_1(y)\} \quad \text{for } r \geq 2,$$

$$\kappa_3 = \xi_1^{-3} \{E g_1^3(X_1) + 3(r-1)E\{g_1(X_1)g_1(X_2)g_2(X_1, X_2)\}\},$$

$$\begin{aligned} \kappa_4 = & \xi_1^{-4} \{E g_1^4(X_1) - 3\xi_1^4 + 12(r-1)E\{g_1^2(X_1)g_1(X_2)g_2(X_1, X_2)\} \\ & + 12(r-1)^2 E\{g_1(X_2)g_1(X_3)g_2(X_1, X_2)g_2(X_1, X_3)\} \\ & + 4(r-1)(r-2)E\{g_1(X_1)g_1(X_2)g_1(X_3)g_3(X_1, X_2, X_3)\}\}, \end{aligned}$$

and

$$Q_N(x) = \Phi(x) - \phi(x) \left\{ \frac{\kappa_3}{6N^{1/2}}(x^2 - 1) + \frac{\kappa_4}{24N}(x^3 - 3x) + \frac{\kappa_5}{72N}(x^5 - 10x^3 + 15x) \right\}$$

where $\Phi(x)$ and $\phi(x)$ denote the distribution function and the density of the standard normal distribution.

THEOREM. *If the following conditions are satisfied (when $r=1$, the condition (C) must be omitted)*

(A) $E|h(X_1, X_2, \dots, X_r)|^5 < \infty$

(B) $\limsup_{|t| \rightarrow \infty} |\eta(t)| < 1$

(C) *there exist positive constants $c < 1$ and $\alpha < 1/8$ such that for $m = [N^\alpha]$,*

$$\begin{aligned} & P\left(\left|E\left(\exp\left\{it\sigma_N^{-1} \frac{r(r-1)}{N(N-1)} \sum_{j=m+1}^N \zeta(X_1, X_j)\right\} \middle| X_{m+1}, \dots, X_N\right)\right| \leq c\right) \\ & \geq 1 - o\left(\frac{1}{N \log N}\right) \end{aligned}$$

uniformly for all $t \in [N^{3/4}/\log N, N \log N]$ then

$$\sup_x |P(\sigma_N^{-1} U_N \leq x) - Q_N(x)| = o(N^{-1}).$$

REMARK 1. Instead of condition (A), the asymptotic expansion is valid under the existence of a fourth moment of kernel h . Lin [6] has proved it in the case of one-sample U -statistics with kernel of degree two. For arbitrary degree, we can similarly prove it by the way of Lin [6].

REMARK 2. The asymptotic expansion with remainder term $o(N^{-1/2})$ is valid without condition (C).

4. Proof of the Theorem

For $r=1$, U_N is a sum of independently and identically distributed random variables and the expansion of U_N has been obtained already (cf. Gnedenko and Kolmogorov [4]). Then we consider the case $r \geq 2$.

Let

$$\Psi_N(t) = E\{\exp(it\sigma_N^{-1}U_N)\}$$

and for $s=1, 2, 3$,

$${}_s\Psi_N(t) = E\{\exp(it\sum_{k=1}^s d_{k,N} A_{k,N})\}$$

where

$$d_{k,N} = \sigma_N^{-1} \binom{N}{r}^{-1} \binom{N-k}{r-k}.$$

Then for ${}_2\Psi_N(t)$, we have the following lemma.

LEMMA 3. If (A) is satisfied, then there exist positive constants $K_s (s=1, 2, \dots, 6)$ such that for all $t (-\infty < t < \infty)$, all integers N and m with $6 < m < N-2$

$$\begin{aligned} |{}_2\Psi_N(t)| \leq & \left| \eta \left(-\frac{rt}{N\sigma_N} \right) \right|^{m-6} (K_1 \sum_{s=0}^3 |t|^s d_{2,N}^s m^s N^s + K_2 |t| d_{3,N} m N^2) \\ & + K_3 t^4 d_{2,N}^4 m^2 N^2 + K_4 t^2 d_{2,N} d_{3,N} m N^{3/2} \\ & + K_5 t^2 d_{3,N}^2 m N^2 + K_6 |t| d_{4,N} m^{1/2} N^{3/2}. \end{aligned} \quad (4.1)$$

Note that if $r=2$, the terms which include $d_{3,N}$ or $d_{4,N}$ are omitted. Similarly we omit $K_6 |t| d_{4,N} m^{1/2} N^{3/2}$, if $r=3$.

PROOF. See Appendix 1.

Furthermore, for ${}_1\Psi_N(t)$ we have Lemma 4.

LEMMA 4. If (A) is satisfied, then there exist positive constants $M_s (s=1, 2, \dots, 12)$ such that for all $t (-\infty < t < \infty)$, all integers N and m with $8 < m < N-3$

$$\begin{aligned} |{}_1\Psi_N(t)| \leq & E(|E(\exp(itd_{2,N} \sum_{j=m+1}^N \zeta(X_1, X_j)) | X_{m+1}, \dots, X_N)|^{m-8}) \\ & \times (M_1 \sum_{s=0}^1 |t|^{s+1} d_{2,N}^s d_{4,N} m^{2s+1} N^3 + M_2 \sum_{s=0}^1 |t|^{s+2} d_{2,N}^s d_{3,N}^2 m^{2(s+1)} N^4 \\ & + M_3 \sum_{s=0}^2 |t|^{s+1} d_{2,N}^s d_{3,N} m^{2s+1} N^2 + M_4 \sum_{s=0}^3 |t|^s d_{2,N}^s m^{2s}) \\ & + M_5 |t|^3 d_{2,N}^2 d_{4,N} m^{9/2} N^{3/2} + M_6 t^2 d_{3,N} d_{4,N} m N^{5/2} \\ & + M_7 t^4 d_{2,N}^2 d_{3,N}^2 m^5 N^2 + M_8 t^4 d_{2,N}^3 d_{3,N} m^{13/2} N \\ & + M_9 t^4 d_{2,N}^4 m^8 + M_{10} |t|^3 d_{3,N}^3 m^{3/2} N^3 \\ & + M_{11} t^2 d_{3,N}^2 m N^3 + M_{12} |t| d_{5,N} m^{1/2} N^2. \end{aligned} \quad (4.2)$$

Note that if $r=2$, the terms which include $d_{3,N}$, $d_{4,N}$ or $d_{5,N}$ are omitted. Similarly if

$r=3$, we omit the terms which include $d_{4,N}$ or $d_{5,N}$, and if $r=4$, omit $M_{12}|t|d_{5,N}m^{1/2}N^2$.

PROOF. See Appendix 2.

Now applying Lemma 3, 4 and Esseen's smoothing lemma [3] we shall prove the Theorem. In the sequel, we consider the proof for the case $r \geq 5$. When $r=2, 3, 4$, we can prove the Theorem more easily.

Let

$$\begin{aligned}\tilde{\Psi}_N(t) &= \int_{-\infty}^{\infty} \exp(itx) dQ_N(x) \\ &= \exp\left(-\frac{t^2}{2}\right) \left\{ 1 + \frac{\kappa_3}{6N^{1/2}}(it)^3 + \frac{\kappa_4}{24N}(it)^4 + \frac{\kappa_5^2}{72N}(it)^6 \right\}.\end{aligned}$$

From smoothing lemma, we have

$$\sup_x |P(\sigma_N^{-1}U_N \leq x) - Q_N(x)| \leq \frac{1}{\pi} \int_{-N \log N}^{N \log N} |t|^{-1} |\Psi_N(t) - \tilde{\Psi}_N(t)| dt + o(N^{-1}).$$

Since the proof for the negative part of t is similar to that for the positive one, we shall show that

$$\int_0^{N \log N} t^{-1} |\Psi_N(t) - \tilde{\Psi}_N(t)| dt = o(N^{-1}).$$

Since $d_{k,N} = O(N^{1/2-k})$ and $E|A_{k,N}| \leq O(N^{k/2})$ from (2.2) in Lemma 2,

$$\begin{aligned}\int_0^{N \log N} t^{-1} |\Psi_N(t) - \tilde{\Psi}_N(t)| dt &\leq \sum_{k=6}^r d_{k,N} E|A_{k,N}| \int_0^{N \log N} dt \\ &= o(N^{-1}).\end{aligned}$$

Similary we have

$$\int_0^{N^{3/4}/\log N} t^{-1} |{}_1\Psi_N(t) - {}_2\Psi_N(t)| dt = o(N^{-1})$$

and

$$\int_0^{N^{1/4}/\log N} t^{-1} |{}_2\Psi_N(t) - {}_3\Psi_N(t)| dt = o(N^{-1}).$$

Then, putting

$$(I) = \int_0^{N^{1/4}/\log N} t^{-1} |{}_8\Psi_N(t) - \tilde{\Psi}_N(t)| dt,$$

$$(II) = \int_{N^{1/4}/\log N}^{N^{3/4}/\log N} t^{-1} |{}_2\Psi_N(t)| dt,$$

$$(III) = \int_{N^{3/4}/\log N}^{N \log N} t^{-1} |{}_1\Psi_N(t)| dt,$$

and

$$(IV) = \int_{\log N}^{\infty} t^{-1} |\tilde{\Psi}_N(t)| dt,$$

we have

$$\int_0^{N \log N} t^{-1} |\Psi_N(t) - \tilde{\Psi}_N(t)| dt \leq (I) + (II) + (III) + (IV) + o(N^{-1}). \quad (4.3)$$

It immediately follows from condition (A) that $(IV) = o(N^{-1})$. Next, we shall prove that (I) , (II) and (III) are $o(N^{-1})$.

Order of (I): Let us define

$$\begin{aligned}\Psi_N^*(t) = & E\left(\exp(itd_{1,N}A_{1,N})\left\{1+itd_{2,N}A_{2,N}+\frac{(it)^2}{2}d_{2,N}^2A_{2,N}^2\right\}\right) \\ & + E(itd_{3,N}A_{3,N}\exp(itd_{1,N}A_{1,N})).\end{aligned}$$

Then,

$$\begin{aligned}& \int_0^{N^{1/4}/\log N} t^{-1} |\Psi_N^*(t) - \Psi_N(t)| dt \\ & \leq \int_0^{N^{1/4}/\log N} t^{-1} \left| E\left(\exp(itd_{1,N}A_{1,N})\left\{\exp(itd_{2,N}A_{2,N}) - \left(1+itd_{2,N}A_{2,N} + \frac{(it)^2}{2}d_{2,N}^2A_{2,N}^2\right)\right\}\right) \right| dt \\ & \quad + \int_0^{N^{1/4}/\log N} t^{-1} |E(itd_{3,N}A_{3,N}\exp(itd_{1,N}A_{1,N})) \times \{\exp(itd_{2,N}A_{2,N}) - 1\}| dt + o(N^{-1}).\end{aligned}$$

Using the similar way which has been described in Callaert et al. [1] pp 304-306, we can establish that the first term is $o(N^{-1})$. Applying Schwartz's inequality and (2.2) in Lemma 2, we can easily obtain the order of the second term.

Now we evaluate the difference of $\Psi_N^*(t)$ and $\tilde{\Psi}_N(t)$. From Lemma 1 $\Psi_N^*(t)$ can be rewritten as

$$\begin{aligned}\Psi_N^*(t) = & \eta^N(d_{1,N}t) + it\eta^{N-2}(d_{1,N}t)d_{2,N}\frac{N(N-1)}{2}E(\exp\{itd_{1,N}(g_1(X_1)+g_1(X_2))\}g_2(X_1, X_2)) \\ & + \frac{(it)^2}{2}\eta^{N-2}(d_{1,N}t)d_{2,N}^2\frac{N(N-1)}{2}E(\exp\{itd_{1,N}(g_1(X_1)+g_1(X_2))\}g_2^2(X_1, X_2)) \\ & + \frac{(it)^2}{2}\eta^{N-3}(d_{1,N}t)d_{2,N}^2N(N-1)(N-2)E(\exp\{itd_{1,N}(g_1(X_1)+g_1(X_2) \\ & + g_1(X_3))\}g_2(X_1, X_2)g_2(X_1, X_3)) + \frac{(it)^2}{2}\eta^{N-4}(d_{1,N}t)d_{2,N}^2\frac{N(N-1)(N-2)(N-3)}{4} \\ & \times \{E(\exp\{itd_{1,N}(g_1(X_1)+g_1(X_2))\}g_2(X_1, X_2))^2 + it\eta^{N-3}(d_{1,N}t)d_{3,N}\frac{N(N-1)(N-2)}{6} \\ & \times E(\exp\{itd_{1,N}(g_1(X_1)+g_1(X_2)+g_1(X_3))\}g_3(X_1, X_2, X_3))\}.\end{aligned}$$

Let us denote $\Psi_N^*(t)$ by

$$\begin{aligned}\Psi_N^*(t) = & I_0^* + itd_{2,N}\frac{N(N-1)}{2}I_2^*E_1^* + \frac{(it)^2}{2}d_{2,N}^2\frac{N(N-1)}{2}I_2^*E_2^* \\ & + \frac{(it)^2}{2}d_{2,N}^2N(N-1)(N-2)I_3^*E_3^* + \frac{(it)^2}{2}d_{2,N}^2\frac{N(N-1)(N-2)(N-3)}{4}I_4^*E_4^* \\ & + itd_{3,N}\frac{N(N-1)(N-2)}{6}I_3^*E_5^*.\end{aligned}$$

Then approximations of I_k^* and E_k^* are given as follows:

$$\begin{aligned}I_k = & \exp\left(-\frac{t^2}{2}\right)\left(1 - \frac{(it)^2}{2N}\left\{\frac{(r-1)^2\xi_2^2}{2\xi_1^2} + k\right\} + \frac{(it)^3}{6N^{1/2}\xi_1^3}Eg_1^3(X_1)\right. \\ & \left. + \frac{(it)^4}{24N\xi_1^4}\{Eg_1^4(X_1) - 3\xi_1^4\} + \frac{(it)^6}{72N\xi_1^6}\{Eg_1^6(X_1)\}^2\right),\end{aligned}$$

$$\begin{aligned}
E_1 &= \frac{(it)^2}{N\xi_1^2} E\{g_1(X_1)g_1(X_2)g_2(X_1, X_2)\} \\
&\quad + \frac{(it)^3}{N^{3/2}\xi_1^3} E\{g_1^2(X_1)g_1(X_2)g_2(X_1, X_2)\}, \\
E_2 &= E g_2^2(X_1, X_2) \\
E_3 &= \frac{(it)^2}{N\xi_1^2} E\{g_1(X_2)g_1(X_3)g_2(X_1, X_2)g_2(X_1, X_3)\}, \\
E_4 &= \left(\frac{(it)^2}{N\xi_1^2} E\{g_1(X_1)g_1(X_2)g_2(X_1, X_2)\} \right)^2, \\
E_5 &= \frac{(it)^3}{N^{3/2}\xi_1^3} E\{g_1(X_1)g_1(X_2)g_1(X_3)g_2(X_1, X_2, X_3)\}.
\end{aligned}$$

By the same way of Lemma 2 as Callaert et al. [1], there exist positive constants ε and a such that for $0 \leq t \leq \varepsilon N^{1/2}$,

$$|I_k^* - I_k| \leq o(N^{-1})P(t) \exp(-at^2)$$

where $P(t)$ is a polynomial in t .

Furthermore, from condition (A), and $\sigma_N^{-1} = N^{1/2}(\xi_1 r)^{-1}(1 + O(N^{-1}))$, we have the following inequalities:

$$\begin{aligned}
|E_1^* - E_1| &\leq t^4 O(N^{-2}) + t^2 O(N^{-2}) + |t|^3 O(N^{-5/2}), \\
|E_2^* - E_2| &\leq |t| O(N^{-1/2}), \\
|E_3^* - E_3| &\leq |t|^3 O(N^{-3/2}) + t^2 O(N^{-2}), \\
|E_4^* - E_4| &\leq |t|^5 O(N^{-5/2}) + t^6 O(N^{-3}) + t^4 O(N^{-3}), \\
|E_5^* - E_5| &\leq t^4 O(N^{-2}) + |t|^3 O(N^{-5/2}).
\end{aligned}$$

Define

$$\begin{aligned}
\tilde{\Psi}_N^*(t) &= I_0 + it d_{2,N} \frac{N(N-1)}{2} I_2 E_1 + \frac{(it)^2}{2} d_{2,N}^2 \frac{N(N-1)}{2} I_2 E_2 \\
&\quad + \frac{(it)^2}{2} d_{2,N}^2 N(N-1)(N-2) I_3 E_3 + \frac{(it)^2}{2} d_{2,N}^2 \frac{N(N-1)(N-2)(N-3)}{4} I_4 E_4 \\
&\quad + it d_{3,N} \frac{N(N-1)(N-2)}{6} I_3 E_5.
\end{aligned}$$

Then,

$$\begin{aligned}
&\int_0^{N^{1/4}/\log N} t^{-1} |\Psi_N^*(t) - \tilde{\Psi}_N(t)| dt \\
&\leq \int_0^{N^{1/4}/\log N} t^{-1} |\Psi_N^*(t) - \tilde{\Psi}_N^*(t)| dt + \int_0^{N^{1/4}/\log N} t^{-1} |\tilde{\Psi}_N^*(t) - \tilde{\Psi}_N(t)| dt.
\end{aligned}$$

Since $\sigma_N^{-1} = N^{1/2}(\xi_1 r)^{-1}(1 + O(N^{-1}))$ and for any k ,

$$\int_0^\infty t^k \exp(-t^2/2) dt < \infty,$$

we get that the last term is $o(N^{-1})$.

For the first term, we have

$$\begin{aligned}
|\Psi_N^*(t) - \tilde{\Psi}_N^*(t)| &\leq |I_0^* - I_0| + |t| \frac{r(r-1)}{2\sigma_N} \{|E_1^*(I_2^* - I_2)| + |I_2(E_1^* - E_1)|\} \\
&\quad + \frac{t^2}{2} \frac{\{r(r-1)\}^2}{2\sigma_N N(N-1)} \{|E_2^*(I_2^* - I_2)| + |I_2(E_2^* - E_2)|\} \\
&\quad + \frac{t^2}{2} \frac{\{r(r-1)\}^2(N-2)}{2\sigma_N N(N-1)} \{|E_3^*(I_3^* - I_3)| + |I_3(E_3^* - E_3)|\} \\
&\quad + \frac{t^2}{2} \frac{\{r(r-1)\}^2(N-2)(N-3)}{4\sigma_N N(N-1)} \{|E_4^*(I_4^* - I_4)| + |I_4(E_4^* - E_4)|\} \\
&\quad + |t| \frac{r(r-1)(r-2)}{6\sigma_N} \{|E_5^*(I_3^* - I_3)| + |I_3(E_5^* - E_5)|\}.
\end{aligned}$$

Therefore, using the previous discussions, we can establish that the first term is $o(N^{-1})$.

Order of (II): Applying (4.1) in Lemma 3, condition (B) and the same arguments which have been described in Callaert et al. [1] pp 308-309, we get the order of (II).

Order of (III): Combining (4.2) in Lemma 4 and condition (C), we can easily establish that the order of (III) is $o(N^{-1})$.

Thus we showed that (I), (II) and (III) are $o(N^{-1})$ and therefore by (4.3) we have the desired result.

Appendix

1. Proof of Lemma 3

We have

$$\begin{aligned}
{}_2\Psi_N(t) &= E(\exp\{itd_{1,N}B_1(m)\}\exp\{itd_{1,N}(A_{1,N}-B_1(m))\}\exp\{itd_{2,N}B_2(m,N)\} \\
&\quad \times \exp\{itd_{2,N}(A_{2,N}-B_2(m,N))\}\exp\{itd_{3,N}B_3(m,N-1,N)\} \\
&\quad \times \exp\{itd_{3,N}(A_{3,N}-B_3(m,N-1,N))\}\exp\{itd_{4,N}B_4(m,N-2,N-1,N)\} \\
&\quad \times \exp\{itd_{4,N}(A_{4,N}-B_4(m,N-2,N-1,N))\}).
\end{aligned}$$

Let us define $B_k^* = B_k(m, \dots)$ and $R_k = \exp\{itd_{k,N}(A_{k,N} - B_k^*)\}$ ($k=1, 2, 3, 4$). Then expanding $\exp\{itd_{k,N}B_k^*\}$ ($k=3, 4$), we have

$$\begin{aligned}
|{}_2\Psi_N(t)| &\leq |E(\exp\{itd_{1,N}B_1^*\}\exp\{itd_{2,N}B_2^*\}\prod_{j=1}^4 R_j)| \\
&\quad + |t|d_{3,N}|E(\exp\{itd_{1,N}B_1^*\}\prod_{j=1}^4 R_j B_3^*)| \\
&\quad + t^2 d_{2,N} d_{3,N} E|B_2^* B_3^*| + t^2 d_{3,N}^2 E|B_3^*|^2 + |t|d_{4,N} E|B_4^*|.
\end{aligned}$$

Therefore using (2.2) in Lemma 2, we can obtain (4.1) in the same way of Lemma 4 of Callaert et al. [1].

2. Proof of Lemma 4

We get

$$\begin{aligned}
{}_1\Psi_N(t) &= E(\exp\{itd_{2,N}\sum_{1\leq i < j \leq N}\zeta_i(X_i, X_j)\}\exp\{itd_{3,N}B_3(m,N-1,N)\} \\
&\quad \times \exp\{itd_{3,N}(A_{3,N}-B_3(m,N-1,N))\}\exp\{itd_{4,N}B_4(m,N-2,N-1,N)\}
\end{aligned}$$

$$\begin{aligned} & \times \exp\{itd_{4,N}(A_{4,N} - B_4(m, N-2, N-1, N))\} \\ & \times \exp\{itd_{5,N}B_5(m, N-3, N-2, N-1, N)\} \\ & \times \exp\{itd_{5,N}(A_{5,N} - B_5(m, N-3, N-2, N-1, N))\}. \end{aligned}$$

Let us define $D = \exp\{itd_{2,N} \sum_{1 \leq i < j \leq N} \zeta(X_i, X_j)\}$, $B_3^* = B_3(m, N-1, N)$, $B_4^* = B_4(m, N-2, N-1, N)$, $B_5^* = B_5(m, N-3, N-2, N-1, N)$ and $R_k = \exp\{itd_{k,N}(A_{k,N} - B_k^*)\}$ ($k=3, 4, 5$). Then expanding $\exp\{itd_{k,N}B_k^*\}$ ($k=3, 4, 5$), we have

$$\begin{aligned} |\Psi_N(t)| & \leq \sum_{s=0}^2 |t|^s d_{3,N}^s |E\{DB_3^* \prod_{j=3}^5 R_j\}| + |t|^3 d_{3,N}^3 E|B_3^*|^3 \\ & \quad + |t| d_{4,N} |E\{DB_4^* \prod_{j=3}^5 R_j\}| + t^2 d_{3,N} d_{4,N} E|B_3^* B_4^*| \\ & \quad + t^2 d_{4,N}^2 E(B_4^*)^2 + |t| d_{5,N} E|B_5^*| \end{aligned}$$

Using (2.2) in Lemma 2 and applying the factorization of D which has been discussed in Lemma 5 of Callaert et al. [1], we have the desired result.

References

- [1] CALLAERT, H., JANSSEN, P. and VERAVERBEKE, N.: *An Edgeworth Expansion for U-statistics*, Ann. Statist., **8** (1980), 299-312.
- [2] DHARMADHIKARI, S.W., FABIAN, V. and JOGDEO, K.: *Bounds of Moments of Martingales*, Ann. Math. Statist., **39** (1968), 1719-1723.
- [3] ESSEEN, C.F.: *Fourier Analysis of Distribution Functions. A Mathematical Study of the Laplace-Gaussian Law*, Acta Math., **77** (1945), 1-125.
- [4] GNEDENKO, B.V. and KOLMOGOROV, A.N.: *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading., (1968).
- [5] Hoeffding, W.: *The Strong Law of Large Numbers for U-statistics*, Univ. of North Carolina Institute of Statistics Mimeo Series. No. 302, (1961).
- [6] LIN ZHENGYAN: *A Note on the Asymptotic Expansion for U-statistics*, (in Chinese) Acta Mathematica Sinica, **27** (1984), 595-598.
- [7] SERFLING, R.J.: *Approximation Theorems of Mathematical Statistics*, Wiley, New York, (1980).

Communicated by Ch. Asano

Received May 18, 1985