The Annals of Statistics
1980, Vol. 8, No. 2, 299–312

AN EDGEWORTH EXPANSION FOR *U*-STATISTICS

BY H. CALLAERT, P. JANSSEN AND N. VERAVERBEKE

Limburgs Universitair Centrum, Belgium

It is shown that, under some regularity conditions on the kernel, a one-sample U-statistic with kernel of degree two admits an asymptotic expansion with remainder term $o(N^{-1})$.

1. Introduction. Let $X_1, X_2, \cdots, X_N$, $N \geqslant 2$, be i.i.d. random variables with common distribution function F. Define a one-sample U-statistic with kernel of degree two by

$$U_N = \binom{N}{2}^{-1} \Sigma_{1 \leqslant i < j \leqslant N} h(X_i, X_j)$$

where h is a symmetric function of two variables with $Eh(X_1, X_2) = 0$ and such that $g(X_1) = E[h(X_1, X_2)|X_1]$ has a positive variance σ_g^2. The asymptotic normality of the statistic $(\text{Var } U_N)^{-\frac{1}{2}} U_N$ has been obtained by Hoeffding (1948) under the sole condition of the existence of $Eh^2(X_1, X_2)$. The study of the rate of convergence, started by Grams and Serfling (1973), resulted in the Berry-Esseen theorem for U-statistics requiring only the existence of $E|h(X_1, X_2)|^3$ (see Callaert and Janssen (1978)), improving results of Bickel (1974) and Chan and Wierman (1977).

The purpose of this paper is to establish an Edgeworth expansion with remainder term $o(N^{-1})$. In Section 2 we state the main theorem and outline its proof, which essentially reduces to the estimation of the three integrals treated in Sections 3, 4 and 5. Section 6 contains a modification of the main theorem, illustrated by two examples.

2. Main result and outline of the proof. The U-statistic defined in the introduction can be rewritten as

$$U_N = \binom{N}{2}^{-1} \left[(N-1) \Sigma_{i=1}^{N} g(X_i) + \Sigma_{1 \leqslant i < j \leqslant N} \varphi(X_i, X_j) \right]$$

where

$$\varphi(X_i, X_j) = h(X_i, X_j) - g(X_i) - g(X_j).$$

Note that φ is symmetric in its arguments and that $E\varphi(X_1, X_2) = 0$. Moreover for $1 \leqslant i < j \leqslant N$ and $1 \leqslant k \leqslant N$ one has $E[g(X_k)\varphi(X_i, X_j)] = 0$. Also $E[\varphi(X_i, X_j)\varphi(X_k, X_l)]$ equals $E\varphi^2(X_1, X_2)$ if $\{i, j\} = \{k, l\}$ and zero otherwise.

For the variance of U_N, denoted by σ_N^2, we have

$$\sigma_N^2 = \frac{4}{N} \sigma_g^2 + \frac{2}{N(N-1)} E\left[\varphi^2(X_1, X_2) \right].$$

Received February 1978; revised July 1978.
AMS 1970 *subject classifications.* Primary 60F05; secondary 62E20.
Key words and phrases. Asymptotic expansion, Gini's mean difference, sample variance, U-statistic.

Throughout the paper the notations

$$c_N = \sigma_N^{-1}\binom{N}{2}^{-1}$$

and

$$\eta(\theta) = E\big[\exp(i\theta g(X_1))\big]$$

will be used.

Finally let

$$\kappa_3 = \frac{1}{\sigma_g^3}\big[Eg^3(X_1) + 3E\big[g(X_1)g(X_2)\varphi(X_1, X_2)\big]\big]$$

$$\kappa_4 = \frac{1}{\sigma_g^4}\big[Eg^4(X_1) - 3\sigma_g^4 + 12E\big[g^2(X_1)g(X_2)\varphi(X_1, X_2)\big] + 12E\big[g(X_2)g(X_3)\varphi(X_1, X_2)\varphi(X_1, X_3)\big]\big]$$

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

where $\Phi(x)$ (resp. $\phi(x)$) is the distribution function (resp. density) of a standard normal random variable.

THEOREM 1. *If the following conditions are satisfied*

(A) $E|h(X_1, X_2)|^4 < \infty$,

(A′) $E|g^4(X_1)\varphi(X_1, X_2)| < \infty$ *and* $E|g^3(X_k)\varphi(X_1, X_2)\varphi(X_1, X_3)| < \infty$ *for* $k = 1, 2$,

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

(C) *there exists a positive constant* $c < 1$, *such that for* $m = [N^\alpha]$, $0 < \alpha < \frac{1}{8}$,

$$P\big[|E\big[\exp\big(itc_N\Sigma_{j=m+1}^N h(X_1, X_j)\big)|X_{m+1}, \cdots, X_N\big]| \leqslant c\big] \geqslant 1 - o\left(\frac{1}{N\log N}\right)$$

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

$$\sup_x|P\big[\sigma_N^{-1}U_N \leqslant x\big] - K_N(x)| = o(N^{-1}).$$

REMARKS.

1. Although $E|h^5(X_1, X_2)| < \infty$ is sufficient for conditions (A) and (A′) to be fulfilled and leads to a shorter proof of the theorem, we prefer to work with the weaker moment conditions (A) and (A′) for two reasons. First, the existence of the fourth moment seems to be the "natural" condition for the expansion up to $o(N^{-1})$. Therefore as far as moments are concerned we prove the whole theorem

using only condition (A) with one single exception (in Lemma 3) where (A′) is needed. Further for Gini's mean difference, discussed at the end of the paper, condition (A) implies (A′), providing an example where the expansion is valid without imposing the existence of a fifth absolute moment.

2. Possible generalisation of Theorem 1 to the general case of multisample U-statistics with arbitrary degree is not studied here. The extension of the formal expansion will be rather technical and the main difficulty will remain the search for an elegant condition under which the characteristic function is sufficiently small outside a neighbourhood of the origin.

3. It will be seen from the proof that condition (B) may be dispensed with if the range of t in condition (C) is extended to $[\varepsilon N^{\frac{1}{2}}, N \log N]$ where ε is a positive constant chosen in such a way that

$$|\eta(\theta)| \leqslant \exp(-\theta^2\sigma_g^2/3) \text{ for } |\theta| \leqslant \varepsilon\sigma_g^{-1}.$$

4. Condition (C) will enable us to reduce expressions like $\Sigma_{j=m+1}^{N} h(X_1, X_j)$ to $\Sigma_{j=m+1}^{N} h(X_1, x_j)$ via a conditioning argument. Although this seems to simplify somewhat the complicated structure, it will generally be very hard to check the validity of (C) in most of the examples encountered in statistics. We therefore in Section 6 propose a more stringent alternative to (C) which can be checked more easily.

The starting point of the proof is Esseen's smoothing lemma (1945), which may be found in Feller page 512 (1966).

Let

$$\psi_N(t) = E\left[\exp\left(it\sigma_N^{-1}U_N\right)\right]$$

$$\tilde{\psi}_N(t) = \int_{-\infty}^{+\infty} \exp(itx)\, dK_N(x)$$

$$= \exp(-t^2/2)\left[1 + \frac{\kappa_3}{6N^{\frac{1}{2}}}(it)^3 + \frac{\kappa_4}{24N}(it)^4 + \frac{\kappa_3^2}{72N}(it)^6\right].$$

Choosing $T = N \log N$, the smoothing lemma ensures that

$$\sup_x |P\left[\sigma_N^{-1}U_N \leqslant x\right] - K_N(x)| \leqslant \frac{2}{\pi}\int_0^{N \log N} t^{-1}|\psi_N(t) - \tilde{\psi}_N(t)|\, dt + o(N^{-1}).$$

Further

$$\int_0^{N \log N} t^{-1}|\psi_N(t) - \tilde{\psi}_N(t)|\, dt \leqslant \int_0^{N^{\frac{1}{4}}/\log N} t^{-1}|\psi_N(t) - \tilde{\psi}_N(t)|\, dt$$

$$+ \int_{N^{\frac{1}{4}}/\log N}^{N^{\frac{3}{4}}/\log N} t^{-1}|\psi_N(t)|\, dt + \int_{N^{\frac{3}{4}}/\log N}^{N \log N} t^{-1}|\psi_N(t)|\, dt + \int_{\log N}^{\infty} t^{-1}|\tilde{\psi}_N(t)|\, dt$$

$$= \text{(I)} + \text{(II)} + \text{(III)} + \text{(IV)}.$$

The proof that (I), (II) and (III) are $o(N^{-1})$ will be obtained in Sections 3, 4 and 5. The order bound $o(N^{-1})$ for (IV) follows immediately from condition (A).

The following shorthand notation will be useful. For $r = 1, \cdots, N$ let

$$A_r = A_r(X_1, \cdots, X_r) = \Sigma_{i=1}^{r} g(X_i)$$

and for $1 \leqslant r < s \leqslant N$ define

$$B_{r,s} = B_{r,s}(X_1, \cdots, X_s) = \Sigma_{i=1}^{r}\Sigma_{j=i+1}^{s}\varphi(X_i, X_j).$$

We also frequently use the following lemma.

LEMMA 1. *Given the existence of the* pth $(p \geqslant 2)$ *absolute moment of the kernel* h, *there exists a positive constant* C *such that*

$$E|B_{r,s}|^p \leqslant C(rs)^{p/2}.$$

The proof of this lemma is very similar to that on page 420 of Callaert and Janssen (1978). It uses the martingale structure of $B_{r,s}$ and an upper bound for moments of martingales obtained by Dharmadhikari, Fabian and Jogdeo (1968).

3. Estimate for (I). The ideas of this section adhere to a paper on Edgeworth expansions for linear combinations of order statistics by Helmers (1976). The expression $K_N(x)$ for the formal Edgeworth expansion, given in Section 2, results from the approximations for $\psi_N(t)$ which will now be introduced. Noting that $\psi_N(t) = E[\exp(itc_N(N-1)A_N)\exp(itc_N B_{N-1,N})]$ we first construct an approximation $\psi_{1,N}(t)$ for $\psi_N(t)$ by replacing the second factor in the expectation by its Taylor series up to the term in t^2. This yields:

$$\psi_{1,N}(t) = E\left\{\exp(itc_N(N-1)A_N)\left[1 + itc_N B_{N-1,N} + \frac{(it)^2}{2}c_N^2 B_{N-1,N}^2\right]\right\}.$$

In its turn, $\psi_{1,N}(t)$ will be approximated by $\tilde{\psi}_{1,N}(t)$ in the following way. Since A_N is a sum of i.i.d. random variables we remark that $\psi_{1,N}(t)$ can be rewritten as

$$\begin{aligned}
\psi_{1,N}(t) &= \eta^N(c_N(N-1)t) \\
&\quad + it\eta^{N-2}(c_N(N-1)t)c_N\frac{N(N-1)}{2} \\
&\quad \times E\{[\exp(itc_N(N-1)(g(X_1) + g(X_2)))]\varphi(X_1, X_2)\} \\
&\quad + \frac{(it)^2}{2}\eta^{N-2}(c_N(N-1)t)c_N^2\frac{N(N-1)}{2} \\
&\quad \times E\{[\exp(itc_N(N-1)(g(X_1) + g(X_2)))]\varphi^2(X_1, X_2)\} \\
&\quad + \frac{(it)^2}{2}\eta^{N-3}(c_N(N-1)t)c_N^2 N(N-1)(N-2) \\
&\quad \times E\{[\exp(itc_N(N-1)(g(X_1) + g(X_2) + g(X_3)))]\varphi(X_1, X_2)\varphi(X_1, X_3)\} \\
&\quad + \frac{(it)^2}{2}\eta^{N-4}(c_N(N-1)t)c_N^2\frac{N(N-1)(N-2)(N-3)}{4} \\
&\quad \times [E\{[\exp(itc_N(N-1)(g(X_1) + g(X_2)))]\varphi(X_1, X_2)\}]^2
\end{aligned}$$

which will be denoted by

$$\psi_{1,N}(t) = I_0^* + itc_N \frac{N(N-1)}{2} I_2^* E_1^* + \frac{(it)^2}{2} c_N^2 N(N-1)\Big[\tfrac{1}{2} I_2^* E_2^* + (N-2) I_3^* E_3^* + \tfrac{1}{4}(N-2)(N-3) I_4^* E_4^*\Big].$$

We now take the first few terms of the Taylor series for approximating E_i^* by E_i, $1 \leqslant i \leqslant 4$, and put:

$$E_1 = \frac{(it)^2}{N\sigma_g^2} E\big[\, g(X_1)g(X_2)\varphi(X_1, X_2)\big] + \frac{(it)^3}{N^{\frac{3}{2}}\sigma_g^3} E\big[\, g^2(X_1)g(X_2)\varphi(X_1, X_2)\big]$$

$$E_2 = E\big[\varphi^2(X_1, X_2)\big]$$

$$E_3 = \frac{(it)^2}{N\sigma_g^2} E\big[\, g(X_2)g(X_3)\varphi(X_1, X_2)\varphi(X_1, X_3)\big]$$

$$E_4 = \frac{(it)^4}{N^2\sigma_g^4} E^2\big[\, g(X_1)g(X_2)\varphi(X_1, X_2)\big].$$

Further, an approximation of I_k^*, $k = 0, 2, 3, 4$, is obtained by using an expansion for the characteristic function

$$E\Big\{\exp\Big[\, itc_N(N-1)\sigma_g(N-k)^{\frac{1}{2}}\big(\sigma_g^{-1}(N-k)^{-\frac{1}{2}}\textstyle\sum_{i=1}^{N-k} g(X_i)\big)\Big]\Big\}$$

where $\sigma_g^{-1}(N-k)^{-\frac{1}{2}}\sum_{i=1}^{N-k} g(X_i)$ is a normalised sum of i.i.d. random variables. For more details we refer to the proof of Lemma 2.

In this way we are led to the following approximation of $\psi_{1,N}(t)$

$$\tilde{\psi}_{1,N}(t) = I_0 + itc_N \frac{N(N-1)}{2} I_2 E_1 + \frac{(it)^2}{2} c_N^2 N(N-1)\Big[\tfrac{1}{2} I_2 E_2 + (N-2)\, I_3 E_3 + \tfrac{1}{4}(N-2)(N-3) I_4 E_4\Big]$$

with

$$I_k = e^{-t^2/2}\left[1 - \frac{(it)^2}{2N}\left(\frac{2E\varphi^2(X_1, X_2)}{4\sigma_g^2} + k\right) + \frac{(it)^3}{6N^{\frac{1}{2}}\sigma_g^3} Eg^3(X_1) + \frac{(it)^4}{24N\sigma_g^4}\big(Eg^4(X_1) - 3\sigma_g^4\big) + \frac{(it)^6}{72N\sigma_g^6} E^2 g^3(X_1)\right]$$

and E_k as defined above.

We finally remark that

$$\begin{aligned}(\mathrm{I}) &\leqslant \textstyle\int_0^{N^{\frac{1}{4}}/\log N} t^{-1}|\psi_N(t) - \psi_{1,N}(t)|\, dt + \int_0^{N^{\frac{1}{4}}/\log N} t^{-1}|\psi_{1,N}(t) - \tilde{\psi}_{1,N}(t)|\, dt \\ &\quad + \textstyle\int_0^{N^{\frac{1}{4}}/\log N} t^{-1}|\tilde{\psi}_{1,N}(t) - \tilde{\psi}_N(t)|\, dt \\ &= (\mathrm{I}.1) + (\mathrm{I}.2) + (\mathrm{I}.3)\end{aligned}$$

and now prove that (I.1) and (I.2) are $o(N^{-1})$. That (I.3) is $o(N^{-1})$ follows from a straightforward computation by writing down explicitly the difference $\tilde{\psi}_{1,N}(t) - \tilde{\psi}_N(t)$, replacing $c_N(N-1)$ by $N^{-\frac{1}{2}}\sigma_g^{-1}[1 + 0(N^{-1})]$, and noting that one needs only to look at the order of N in the terms involved because

$$\int_0^\infty t^k e^{-t^2/2}\,dt < \infty.$$

Estimate for (I.1). From the definition of $\psi_N(t)$ and $\psi_{1,N}(t)$ it immediately follows that

$$|\psi_N(t) - \psi_{1,N}(t)| \leqslant c_N^3|t|^3 E|B_{N-1,N}|^3.$$

Hence, since $c_N^3 = 0(N^{-\frac{9}{2}})$ and $E|B_{N-1,N}|^3 = 0(N^3)$ (see Lemma 1), we have:

$$\int_0^{N^{\frac{1}{6}}/\log N} t^{-1}|\psi_N(t) - \psi_{1,N}(t)|\,dt \leqslant c_N^3 E|B_{N-1,N}|^3 \int_0^{N^{\frac{1}{6}}/\log N} t^2\,dt = o(N^{-1}).$$

For t in the range $[N^{\frac{1}{6}}/\log N, N^{\frac{1}{4}}/\log N]$ we again use the definition of $\psi_N(t)$ and $\psi_{1,N}(t)$ and now write

$$|\psi_N(t) - \psi_{1,N}(t)| \leqslant |t|^3 c_N^3 \left|E\left[\exp(itc_N(N-1)A_N)B_{N-1,N}^3\right]\right| + t^4 c_N^4 E\left(B_{N-1,N}^4\right).$$

The coefficient of $|t|^3 c_N^3$ is bounded by

$$\begin{aligned}
&\Sigma_{(1)}\left|E\left\{\left[\exp(itc_N(N-1)A_N)\right]\varphi(X_{i_1}, X_{j_1})\varphi(X_{i_2}, X_{j_2})\varphi(X_{i_3}, X_{j_3})\right\}\right| \\
&\quad = \Sigma_{(1)}\left|E\left[\exp(itc_N(N-1)\Sigma_{i\in\Delta'} g(X_i))\right]\right| \\
&\qquad \times \left|E\left\{\left[\exp(itc_N(N-1)\Sigma_{i\in\Delta} g(X_i))\right]\varphi(X_{i_1}, X_{j_1})\varphi(X_{i_2}, X_{j_2})\varphi(X_{i_3}, X_{j_3})\right\}\right| \\
&\quad \leqslant |\eta(tc_N(N-1))|^{N-6} \\
&\qquad \times \Sigma_{(1)}\left|E\left\{\left[\exp(itc_N(N-1)\Sigma_{i\in\Delta} g(X_i))\right]\varphi(X_{i_1}, X_{j_1})\right.\right. \\
&\qquad\qquad \left.\left.\varphi(X_{i_2}, X_{j_2})\varphi(X_{i_3}, X_{j_3})\right\}\right|
\end{aligned}$$

where Δ is the set of different indices among $i_1, j_1, i_2, j_2, i_3, j_3$ and $\Delta' = \{1, 2, \cdots, N\} \setminus \Delta$. The number k of elements in Δ is at least 2 and at most 6 and for each k the number of terms in $\Sigma_{(1)}$ is $0(N^k)$. Now there always exists an $\varepsilon > 0$ such that

$$|\eta(\theta)| \leqslant \exp\left(-\tfrac{1}{3}\theta^2\sigma_g^2\right) \qquad \text{for } |\theta| < \varepsilon/\sigma_g.$$

Since $2/\sigma_N < N^{\frac{1}{2}}/\sigma_g$ we have that $tc_N(N-1) < \varepsilon/\sigma_g$ for $t < \varepsilon N^{\frac{1}{2}}$. Hence for $t < \varepsilon N^{\frac{1}{2}}$

$$|\eta(tc_N(N-1))|^{N-6} \leqslant \exp\left(-\tfrac{1}{3}t^2c_N^2(N-1)^2\sigma_g^2(N-6)\right) \leqslant e^{-\frac{1}{12}t^2} \quad \text{for } N \text{ large.}$$

From this remark we see that the terms for $k = 2, 3$ can be replaced by an absolute constant because $c_N^3 = 0(N^{-\frac{9}{2}})$, the number of terms is at most $0(N^3)$ and $\int_0^\infty t^2 e^{-\frac{1}{12}t^2}\,dt < \infty$.

To treat the terms with $k = 4, 5, 6$ we first of all remark that for each integrable

Borel-measurable function $f(x_{k_1}, \cdots, x_{k_r})$ with $E|f\varphi| < \infty$ we have

$$(*) \qquad E\big[f(X_{k_1}, \cdots, X_{k_r})\varphi(X_i, X_j)\big] = 0$$

if at least one of the indices i or j does not belong to $\{k_1, \cdots, k_r\}$. In fact, if both i and j differ from all k_s, $1 \leqslant s \leqslant r$, then $(*)$ holds by independence together with $E\varphi(X_i, X_j) = 0$. On the other hand, if $i \in \{k_1, \cdots, k_r\}$ and $j \notin \{k_1, \cdots, k_r\}$ we first use a conditioning on the Borel-field generated by $X_{k_1}, \cdots, X_{k_r}$ and then note that $E[\varphi(X_i, X_j)|X_{k_1}, \cdots, X_{k_r}] = E[\varphi(X_i, X_j)|X_i]$ which is zero according to the definition of $\varphi(X_i, X_j)$.

We now give the argument for $k = 4$ and indicate the analogy for $k = 5, 6$. If Δ contains exactly four different indices we encounter expressions of the types $\varphi(X_1, X_2)\varphi(X_1, X_3)\varphi(X_1, X_4)$ and $\varphi^2(X_1, X_2)\varphi(X_3, X_4)$. Note that the first type does not factorize when one takes the expectation. Hence $\varphi(X_1, X_2)\varphi(X_2, X_3)\varphi(X_3, X_4)$ and $\varphi(X_1, X_2)\varphi(X_1, X_3)\varphi(X_3, X_4)$ also belong to this type. To fix the idea we work with $\varphi(X_1,X_2)\varphi(X_1, X_3)\varphi(X_1, X_4)$.

As to the first type one has for each particular term, using $(*)$ and condition (A):

$$\begin{aligned}
&|E\{[\exp(itc_N(N-1)\Sigma_{i=1}^4 g(X_i))]\varphi(X_1, X_2)\varphi(X_1, X_3)\varphi(X_1, X_4)\}| \\
&\quad = |E\{[\exp(itc_N(N-1)\Sigma_{i=1}^4 g(X_i)) - 1]\varphi(X_1, X_2)\varphi(X_1, X_3)\varphi(X_1, X_4)\}| \\
&\quad \leqslant Ktc_N(N-1).
\end{aligned}$$

Remembering that each term has to be multiplied by $t^{-1}|t|^3c_N^3|\eta(tc_N(N-1))|^{N-6}$ and then integrated for $t \in [N^{\frac{1}{6}}/\log N, N^{\frac{1}{4}}/\log N]$ we find that each term is bounded by

$$\begin{aligned}
&Kc_N^4(N-1)\int_{N^{\frac{1}{6}}/\log N}^{N^{\frac{1}{4}}/\log N} t^3 e^{-\frac{1}{12}t^2}\,dt \\
&\qquad \leqslant Kc_N^4(N-1)\tfrac{1}{4}N(\log N)^{-4}\exp\big(-\tfrac{1}{12}N^{\frac{1}{3}}(\log N)^{-2}\big).
\end{aligned}$$

Since there are at most $0(N^4)$ terms we find an order bound of $o(N^{-1})$ for all terms of the first type together.

The method used for the terms of the second type in $k = 4$ will be standard in the analysis of $k = 5, 6$. By independence, $(*)$ and condition (A) we have

$$\begin{aligned}
&|E\{[\exp(itc_N(N-1)\Sigma_{i=1}^4 g(X_i))]\varphi^2(X_1, X_2)\varphi(X_3, X_4)\}| \\
&\quad = |E\{[\exp(itc_N(N-1)\Sigma_{i=1}^2 g(X_i))]\varphi^2(X_1, X_2)\} \\
&\qquad \times E\{[\exp(itc_N(N-1)\Sigma_{i=3}^4 g(X_i)) - 1 - itc_N(N-1)\Sigma_{i=3}^4 g(X_i)]\varphi(X_3, X_4)\}| \\
&\quad \leqslant Lt^2c_N^2(N-1)^2.
\end{aligned}$$

Hence, performing the same operations as above, each term is bounded by

$$Lc_N^5(N-1)^2\int_0^\infty t^4 e^{-\frac{1}{12}t^2}\,dt = 0(N^{-\frac{11}{2}})$$

which provides an order bound of $0(N^{-\frac{3}{2}})$ for all terms of the second type together.

For $k = 5, 6$ we essentially use the same argument. For example, for $k = 5$, we write

$$|E\{[\exp(itc_N(N-1)\Sigma_{i=1}^5 g(X_i))]\varphi(X_1, X_2)\varphi(X_1, X_3)\varphi(X_4, X_5)\}|$$
$$= |E\{[\exp(itc_N(N-1)\Sigma_{i=1}^3 g(X_i)) - 1 - itc_N(N-1)\Sigma_{i=1}^3 g(X_i)]$$
$$\times \varphi(X_1, X_2)\varphi(X_1, X_3)\} E\{[\exp(itc_N(N-1)\Sigma_{i=4}^5 g(X_i)) - 1$$
$$- itc_N(N-1)\Sigma_{i=4}^5 g(X_i)]\varphi(X_4, X_5)\}|$$
$$\leqslant Mt^4 c_N^4 (N-1)^4$$

and analogously for $k = 6$.

Finally, since $c_N^4 E(B_{N-1,N}^4) = 0(N^{-2})$ and $\int_0^{N^{\frac{1}{4}}/\log N} t^3\, dt = 0(N(\log N)^{-4})$, we find that

$$\int_0^{N^{\frac{1}{4}}/\log N} t^{-1}|\psi_N(t) - \psi_{1,N}(t)|\, dt = o(N^{-1}).$$

Estimate for (I.2). Writing η for $\eta(c_N(N-1)t)$ we have

$$|\psi_{1,N}(t) - \tilde{\psi}_{1,N}(t)| \leqslant |\eta^N - I_0|$$
$$+|t|c_N\frac{N(N-1)}{2}\left[|\eta^{N-2}(E_1^* - E_1)| + |E_1(\eta^{N-2} - I_2)|\right]$$
$$+\frac{t^2}{2}c_N^2\frac{N(N-1)}{2}\left[|\eta^{N-2}(E_2^* - E_2)| + |E_2(\eta^{N-2} - I_2)|\right]$$
$$+\frac{t^2}{2}c_N^2 N(N-1)(N-2)\left[|\eta^{N-3}(E_3^* - E_3)| + |E_3(\eta^{N-3} - I_3)|\right]$$
$$+\frac{t^2}{2}c_N^2\frac{N(N-1)(N-2)(N-3)}{4}\left[|\eta^{N-4}(E_4^* - E_4)| + |E_4(\eta^{N-4} - I_4)|\right].$$

That (I.2) $= o(N^{-1})$ now follows immediately from the next two lemmas.

LEMMA 2. *If* (A) *is satisfied then there exist positive constants* K_3, a, ε *and a sequence of positive numbers* $\delta_1, \delta_2, \cdots$, *with* $\delta_N \to 0$ *as* $N \to \infty$ *such that for each fixed* $k = 0, 1, 2, \cdots$ *and for* $N > k$ *and* $0 \leqslant t \leqslant \varepsilon N^{\frac{1}{2}}$

$$|\eta^{N-k}(c_N(N-1)t) - I_k| \leqslant K_3\delta_N N^{-1} tP(t)e^{-at^2}$$

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

PROOF. From Theorem 1, Section 41 in Gnedenko and Kolmogorov (1968) it follows that

$$\left|\eta^{N-k}\left(\frac{t}{(N-k)^{\frac{1}{2}}\sigma_g}\right) - e^{-t^2/2}\left[1 + \frac{(it)^3}{6(N-k)^{\frac{1}{2}}\sigma_g^3}Eg^3(X_1)\right.\right.$$
$$\left.\left. + \frac{(it)^4}{24(N-k)\sigma_g^4}(Eg^4(X_1) - 3\sigma_g^4) + \frac{(it)^6}{72(N-k)\sigma_g^6}E^2g^3(X_1)\right]\right|$$
$$\leqslant c\delta_N N^{-1} tP(t)e^{-t^2/4}.$$

This expression remains valid if we substitute t by $\sigma_g(N-k)^{\frac{1}{2}}c_N(N-1)t$ because $\sigma_g(N-k)^{\frac{1}{2}}c_N(N-1) \leqslant 1$. Since

$$\frac{4\sigma_g^2}{N\sigma_N^2} = 1 - \frac{2}{N-1}\frac{E\varphi^2(X_1, X_2)}{N\sigma_N^2}$$

we have

$$\frac{\left(i\sigma_g(N-k)^{\frac{1}{2}}c_N(N-1)t\right)^3}{6(N-k)^{\frac{1}{2}}\sigma_g^3}Eg^3(X_1) = \frac{(it)^3}{6N^{\frac{1}{2}}\sigma_g^3}Eg^3(X_1)\left[1 + 0(N^{-1})\right]$$

and analogous expressions for the terms in $(it)^4$ and $(it)^6$. The lemma now follows easily if we take into account the fact that

$$\exp\left[\frac{t^2}{2}\left(1 - \sigma_g^2(N-k)c_N^2(N-1)^2\right)\right] = 1 - \frac{(it)^2}{2N}\left(\frac{2E\varphi^2(X_1, X_2)}{4\sigma_g^2} + k\right) + 0(N^{-2}).$$

LEMMA 3. *If* (A) *and* (A′) *are satisfied, then for all* t

$$|E\{[\exp(itc_N(N-1)(g(X_1) + g(X_2)))]\varphi(X_1, X_2)\} - E_1| \leqslant K_4\left(N^{-2}(t^2 + t^4) + N^{-\frac{5}{2}}|t|^3\right)$$

$$|E\{[\exp(itc_N(N-1)(g(X_1) + g(X_2)))]\varphi^2(X_1, X_2)\} - E_2| \leqslant K_4N^{-\frac{1}{2}}|t|$$

$$|E\{[\exp(itc_N(N-1)(g(X_1) + g(X_2) + g(X_3)))]\varphi(X_1, X_2)\varphi(X_1, X_3)\} - E_3| \leqslant K_4\left(N^{-\frac{3}{2}}|t|^3 + N^{-2}t^2\right)$$

$$|E^2\{[\exp(itc_N(N-1)(g(X_1) + g(X_2)))]\varphi(X_1, X_2)\} - E_4| \leqslant K_4\left(N^{-\frac{5}{2}}|t|^5 + N^{-3}t^4\right)$$

where K_4 *is an absolute constant.*

The proof of this lemma follows from remark (*) in the proof of the estimate for (I.1), together with

$$c_N^k(N-1)^k = N^{-k/2}\sigma_g^{-k}\left[1 + 0(N^{-1})\right], \qquad k = 2, 3, 4.$$

4. Estimate for (II). To establish a suitable estimate for (II) one only needs an appropriate upper bound for $|\psi_N(t)|$. We therefore rely on the next lemma.

LEMMA 4. *If* (A) *is satisfied, then there exist positive constants* K_5 *and* K_6 *such that for all* t *and all integers* N *and* m *with* $6 < m < N$

$$|\psi_N(t)| \leqslant |\eta(tc_N(N-1))|^{m-6}\left(1 + K_5\Sigma_{k=1}^3|t|^kc_N^k(mN)^k\right) + K_6|t|^4c_N^4(mN)^2.$$

PROOF.

$$\psi_N(t) = E\left[\exp(itc_N(N-1)A_m)\exp(itc_N(N-1)(A_N - A_m)) \times \exp(itc_N(B_{N-1,N} - B_{m,N}))\exp(itc_NB_{m,N})\right].$$

Hence, using the expansion for $\exp(itc_N B_{m,N})$

$$|\psi_N(t)| \leqslant \Sigma_{k=0}^{3}|t|^k c_N^k |E\big[\exp(itc_N(N-1)A_m)\exp(itc_N(N-1)(A_N - A_m)) \times \exp(itc_N(B_{N-1,N} - B_{m,N}))B_{m,N}^k\big]| + |t|^4 c_N^4 E(B_{m,N}^4).$$

By an independence argument, the term for $k = 0$ is bounded by

$$|E\big[\exp(itc_N(N-1)A_m)\big]| \leqslant |\eta(tc_N(N-1))|^{m-6}.$$

The coefficient of $|t|^k c_N^k$ for $k = 1, 2, 3$, is bounded by

$$\Sigma_{(2)}|E\big[\exp(itc_N(N-1)A_m)\exp(itc_N(N-1)(A_N - A_m)) \times \exp(itc_N(B_{N-1,N} - B_{m,N}))\varphi(X_{i_1}, X_{j_1})\cdots\varphi(X_{i_k}, X_{j_k})\big]|$$

where the number of terms in $\Sigma_{(2)}$ is less then $(mN)^k$.

Now for any arbitrary term in $\Sigma_{(2)}$ let Δ be the set of different indices among $i_1, j_1, \cdots, i_k, j_k$. If $\Delta_1 = \{1, \cdots, m\} \cap \Delta$ and $\Delta_2 = \{1, 2, \cdots, m\} \setminus \Delta_1$ it is easily seen that Δ_1 contains at least 1 and at most $2k$ elements and hence the number of elements in Δ_2 lies between $m - 1$ and $m - 2k$. Therefore

$$\begin{aligned}|E\big[&\exp\big(itc_N(N-1)\Sigma_{i\in\Delta_2} g(X_i)\big)\exp\big(itc_N(N-1)\big(\Sigma_{i\in\Delta_1} g(X_i) + (A_N - A_m)\big)\big) \\ &\times \exp(itc_N(B_{N-1,N} - B_{m,N}))\varphi(X_{i_1}, X_{j_1})\cdots\varphi(X_{i_k}, X_{j_k})\big]| \\ &\leqslant |E\big[\exp\big(itc_N(N-1)\Sigma_{i\in\Delta_2} g(X_i)\big)\big]|E|\varphi(X_{i_1}, X_{j_1})\cdots\varphi(X_{i_k}, X_{j_k})| \\ &\leqslant K_5|\eta(tc_N(N-1))|^{m-6}.\end{aligned}$$

Finally, by Lemma 1, we have

$$E(B_{m,N}^4) \leqslant K_6(mN)^2$$

finishing the proof of the lemma.

The previous lemma and the fact that

$$|\eta(tc_N(N-1))| \leqslant \exp\big(-t^2c_N^2(N-1)^2\sigma_g^2/3\big) \qquad \text{for } 0 \leqslant t \leqslant \varepsilon N^{\frac{1}{2}}$$

enable us to prove that (II) is $o(N^{-1})$.

Choosing $m = [N^{\frac{1}{2}+\delta}]$, $0 < \delta < \frac{1}{4}$, for $t \in [N^{\frac{1}{4}}/\log N, N \log N]$ we have that

$$\begin{aligned}\int_{N^{\frac{1}{4}}/\log N}^{N^{\frac{3}{8}}} t^{-1}|\psi_N(t)|\,dt \leqslant &\int_{N^{\frac{1}{4}}/\log N}^{N^{\frac{3}{8}}} t^{-1}\exp\left(-\frac{4\sigma_g^2}{3N^2\sigma_N^2}(m-6)t^2\right)dt \\ &+ K_5\Sigma_{k=1}^{3}c_N^k(mN)^k\int_{N^{\frac{1}{4}}/\log N}^{N^{\frac{3}{8}}} t^{k-1}\exp\left(-\frac{4\sigma_g^2}{3N^2\sigma_N^2}(m-6)t^2\right)dt \\ &+ K_6c_N^4(mN)^2\int_{N^{\frac{1}{4}}/\log N}^{N^{\frac{3}{8}}} t^3\,dt.\end{aligned}$$

The last term is $0(N^{-1})$ by our choice of m. For the first term we have

$$\int_{N^{\frac{1}{4}}/\log N}^{N^{\frac{3}{8}}} t^{-1}\exp\left(-\frac{4\sigma_g^2}{3N^2\sigma_N^2}(m-6)t^2\right)dt \leqslant \exp\left(-\frac{4\sigma_g^2}{3N\sigma_N^2}(m-6)\frac{N^{\frac{1}{2}}}{\log^2 N}\right)\log N^{\frac{3}{8}} = o(N^{-1}).$$

The other terms are treated similarly. For $t \in [N^{\frac{3}{8}}, \varepsilon N^{\frac{1}{2}}]$, the same argument works with the choice $m = [N^{\delta}]$, $\frac{1}{4} < \delta < \frac{1}{2}$. Finally if (B) is satisfied, then there exists a $c > 0$ such that

$$|\eta(tc_N(N-1))| \leqslant e^{-c}$$

for $t \geqslant \varepsilon N^{\frac{1}{2}}$. Hence applying again the previous lemma with now $m = (2/c)\log N$, we find

$$\int_{\varepsilon N^{\frac{1}{2}}}^{n^{\frac{3}{4}}/\log N} t^{-1}|\psi_N(t)|\,dt = o(N^{-1}).$$

5. Estimate for (III). The proof that (III) is $o(N^{-1})$ relies on (A) and (C) and the following inequality.

LEMMA 5. *If* (A) *and* (C) *are satisfied, then there exist positive constants* K_7 *and* K_8 *such that for all* t *and all integers* N *and* m *with* $6 < m < N$

$$|\psi_N(t)| \leqslant E\Big[\Big|E\Big[\exp\big(itc_N\Sigma_{j=m+1}^{N}h(X_1, X_j)\big)\Big|X_{m+1}, \cdots, X_N\Big]\Big|^{m-6}\Big] \times \big(1 + K_7\Sigma_{k=1}^{3}|t|^k c_N^k m^{2k}\big) + K_8|t|^4 c_N^4 m^8.$$

PROOF. For $1 \leqslant r < s \leqslant N$ let

$$\tilde{B}_{r,s} = \tilde{B}_{r,s}(X_1, \cdots, X_s) = \Sigma_{i=1}^{r}\Sigma_{j=i+1}^{s}h(X_i, X_j).$$

Then

$$\psi_N(t) = E\Big[E\Big[\exp(itc_N(\tilde{B}_{m,N} - \tilde{B}_{m-1,m}))\exp(itc_N\tilde{B}_{m-1,m}) \times \exp(itc_N(\tilde{B}_{N-1,N} - \tilde{B}_{m,N}))\Big|X_{m+1}, \cdots, X_N\Big]\Big].$$

Hence, using a conditioning argument and then an expansion for $\exp(itc_N\tilde{B}_{m-1,m})$

$$\begin{aligned}|\psi_N(t)| &\leqslant E\big|E\big[\exp(itc_N(\tilde{B}_{m,N} - \tilde{B}_{m-1,m}))\exp(itc_N\tilde{B}_{m-1,m})\big|X_{m+1}, \cdots, X_N\big]\big| \\ &\leqslant \Sigma_{k=0}^{3}|t|^k c_N^k E\big|E\big[\exp(itc_N(\tilde{B}_{m,N} - \tilde{B}_{m-1,m}))\tilde{B}_{m-1,m}^k\big|X_{m+1}, \cdots, X_N\big]\big| \\ &\quad + |t|^4 c_N^4 E(\tilde{B}_{m-1,m}^4).\end{aligned}$$

The rest of the proof is essentially the same as in Lemma 4.

Now if (C) is satisfied, then there exists a constant $\gamma > 0$ such that

$$\big|E\big[\exp\big(itc_N\Sigma_{j=m+1}^{N}h(X_1, X_j)\big)\big]\big|X_{m+1}, \cdots, X_N\big]\big| \leqslant e^{-\gamma}$$

uniformly for $t \in [N^{\frac{3}{4}}/\log N, N \log N]$ except on a set A_N^c with $P[A_N^c] = o\Big(\dfrac{1}{N \log N}\Big)$. Hence,

$$\int_{N^{\frac{3}{4}}/\log N}^{N \log N} t^{-1}|\psi_N(t)|\,dt = o(N^{-1}).$$

6. Examples. In this section we replace condition (C) in Theorem 1 by an alternative condition (C′) which is more tractable for some applications. Let $X, X_1, \cdots, X_N$ be i.i.d. random variables with df F and let $U, U_1, \cdots, U_N$ be i.i.d. random variables with a uniform distribution on [0, 1].

Let $D_N \subset [0, 1]^N$ (cartesian product of N unit intervals) be a set such that $P[D_N^c] = o\left(\frac{1}{N \log N}\right)$. For

$$\bar{u}_N = (u_1, \cdots, u_N) \in D_N$$

and

$$h_{N, \bar{u}_N}(F^{-1}(u)) = \frac{1}{N}\Sigma_{j=1}^N h\big(F^{-1}(u), F^{-1}(u_j)\big)$$

we state the alternative condition as follows:

(C′) There exist a positive constant c and an interval $I_F \subset [0, 1]$ of length at least η, where η is a positive constant, such that

(i) $h_{N, \bar{u}_N}(F^{-1}(u))$ is monotone and differentiable w.r.t. u on I_F

(ii) $|(\partial/\partial u)h_{N, \bar{u}_N}(F^{-1}(u))| \geqslant c$ on I_F, where c is a uniform constant w.r.t. N and D_N.

It then follows from the paragraph preceding formula (5.21) in Albers, Bickel, van Zwet (1976) that, under conditions (A) and (C′), for each $\delta > 0$ there exists a $b > 0$ such that

$$|E[e^{ith_{N, \bar{u}}N^{(F^{-1}(U))}}]| \leqslant 1 - b,$$

uniformly in N and D_N for all $t \geqslant \delta$. From this remark we immediately have the following theorem:

THEOREM 2. *If conditions* (A), (A′) *and* (C′) *are satisfied then* $\sup_x|P[\sigma_N^{-1}U_N \leqslant x] - K_N(x)| = o(N^{-1})$.

We now consider two examples where, for simplicity, we assume an underlying distribution F which is uniform on [0, 1].

Example 1. *sample variance.* Let $h(u_1, u_2) = \frac{(u_1 - u_2)^2}{2} - \frac{1}{12}$ and take $0 < \varepsilon < \frac{1}{4}$,

$$I = \left(1 - \varepsilon, 1 - \frac{\varepsilon}{2}\right) \subset [0, 1] \quad \text{and} \quad D_N = \left\{\bar{u}_N: \frac{1}{N}\Sigma_{j=1}^N u_j < 1 - 2\varepsilon\right\}.$$

If $S_N = \Sigma_{j=1}^N U_j$ then the Markov inequality yields

$$P[D_N^c] \leqslant \frac{E\{e^{tS_N}\}}{\exp(tN(1 - 2\varepsilon))} = \exp[-N\{\log t + t(1 - 2\varepsilon) - \log(e^t - 1)\}].$$

It nows follows by elementary calculus that there exists a choice for t, $t > 0$, and ε, $0 < \varepsilon < \frac{1}{4}$, such that $\log t + t(1 - 2\varepsilon) - \log(e^t - 1)$ is positive, which implies that $P[D_N^c] = o\left(\frac{1}{N \log N}\right)$. Hence it only remains to find a positive constant such that (C′) is satisfied.

But for $u \in I$ and $\bar{u}_N \in D_N$ we have

$$\frac{\partial}{\partial u} h_{N, \bar{u}_N}(u) = u - \frac{1}{N}\Sigma_{j=1}^N u_j > \varepsilon$$

indicating the validity of the result of Theorem 2 for the sample variance.

Example 2. *Gini's mean difference.* We now take $h(u_1, u_2) = |u_1 - u_2| - \frac{1}{3}$ and consider for $0 < \varepsilon < \frac{1}{4}$

$$I = \left(1 - \varepsilon, 1 - \frac{\varepsilon}{2}\right) \subset [0, 1]$$

and $D_{N,k} = \{\bar{u}_N: u_{i_1} > 1 - \varepsilon, \cdots, u_{i_k} > 1 - \varepsilon$ and $u_j < 1 - \varepsilon$ for $j \neq i_1, \cdots, i_k\}$. With $D_N = \cup_{k=0}^{[\frac{N(1-\varepsilon)}{2}]} D_{N,k}$ and for $u \in I$ and $\bar{u}_N \in D_N$ we have

$$\left|\frac{\partial}{\partial u} h_{N,\bar{u}_N}(u)\right| = \left|\frac{1}{N}\Sigma_{j=1}^{N}\frac{\partial}{\partial u}|u - u_j|\right| \geqslant \frac{1}{N}[(N-k) - k] = \frac{N - 2k}{N} > \varepsilon.$$

It only remains to prove that $P[D_N^c] = o\left(\frac{1}{N \log N}\right)$, which follows from the Markov inequality as in the previous example.

REMARKS.

1. For Gini's mean difference where X_1 has df F the following general result can be proved in an analogous way. Suppose there exist positive constants $0 < \varepsilon < \frac{1}{4}$, m and M such that F has a density f on $(F^{-1}(1-\varepsilon), F^{-1}(1 - \varepsilon/2))$ with $m \leqslant f \leqslant M$ then (C') is satisfied for $I_F = (1 - \varepsilon, 1 - \varepsilon/2)$ and D_N as before.

2. From the equality (David (1970) page 146)

$$G = \binom{N}{2}^{-1}\Sigma_{i<j}|X_i - X_j| = \frac{4}{N(N-1)}\Sigma_{i=1}^{N}\left(i - \tfrac{1}{2}(N+1)\right)X_{i:N}$$

it follows that Gini's mean difference is also a linear combination of the order statistics $X_{1:N}, \cdots, X_{N:N}$. Hence it is interesting to compare our result with that in Theorem 2.1 of Helmers (1976). His conditions are satisfied with $J_1(s) = J_2(s) = 4(s - \frac{1}{2})$ and $B = \gamma = 2$. Further for the functions $h_1(u)$, $h_2(u, v)$,$h_3(u, v, w)$, defined in Helmers (2.1), (2.2) and (2.3) we have $h_1(F(x)) = 2g(x)$, $h_2(F(x), F(y)) = 2\varphi(x, y)$, $h_3(u, v, w) = 0$, which implies that the two expansions coincide.

Finally we observe that condition (A') may be dispensed with. This follows from remark (ii) on page 11 in Helmers.

Acknowledgments. The authors thank Professor W. R. van Zwet for stimulating and fruitful discussions on this subject. They also acknowledge the useful comments of an associate editor and of the referees.

REFERENCES

[1] ALBERS, W., BICKEL, P. J. and VAN ZWET, W. (1976). Asymptotic expansions for the power of distribution free tests in the one-sample problem. *Ann. Statist.* **4** 108–156.

[2] BICKEL, P. J. (1974). Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **2** 1–20.

[3] CALLAERT, H. and JANSSEN, P. (1978). The Berry-Esseen theorem for *U*-statistics. *Ann. Statist.* **6** 417–421.

[4] CHAN, Y.-K. and WIERMAN, J. (1977). On the Berry-Esseen theorem for *U*-statistics. *Ann. Probability* **5** 136–139.

[5] DAVID, H. A. (1970). *Order Statistics*. Wiley, New York.

[6] DHARMADHIKARI, S. W., FABIAN, V. and JOGDEO, K. (1968). Bounds on the moments of martingales. *Ann. Math. Statist.* **39** 1719–1723.

[7] ESSEEN, C. F. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Math.* **77** 1–125.

[8] FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications*. 2. Wiley, New York.

[9] GNEDENKO, B. V. and KOLMOGOROV, A. N. (1968). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Reading.

[10] GRAMS, W. F. and SERFLING, R. J. (1973). Convergence rates for *U*-statistics. *Ann. Statist.* **1** 153–160.

[11] HELMERS, R. (1976). Edgeworth expansions for linear combinations of order statistics with smooth weight functions. Report SW44/76, Mathematisch Centrum, Amsterdam.

[12] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

DEPARTMENT OF MATHEMATICS
LIMBURGS UNIVERSITAIR CENTRUM
B-3610 DIEPENBEEK, (BELGIUM)