stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK SW 95/83 APRIL
(DEPARTMENT OF MATHEMATICAL STATISTICS)

W.R. VAN ZWET

A BERRY-ESSEEN BOUND FOR SYMMETRIC STATISTICS

Preprint

kruislaan 413 1098 SJ amsterdam

Printed at the Mathematical Centre, Kruislaan 413, Amsterdam, The Netherlands.

The Mathematical Centre, founded 11 February 1946, is a non-profit institution for the promotion of pure and applied mathematics and computer science. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

1980 Mathematics subject classification: Primary: 60F05
Secondary: 62E20

A Berry-Esseen bound for symmetric statistics *)

by

W.R. van Zwet **)

ABSTRACT

The rate of convergence of the distribution function of a symmetric function of N independent and identically distributed random variables to its normal limit is investigated. Under appropriate moment conditions the rate is shown to be $\mathcal{O}(N^{-\frac{1}{2}})$. This theorem generalizes many known results for special cases and two examples are given. Possible further extensions are indicated.

KEY WORDS & PHRASES: *Berry-Esseen bounds, speed of convergence, symmetric statistics.*

---

*) This report will be submitted for publication elsewhere.

Research supported by the U.S. Office of Naval Research, Contract N 00014-80-C-0163.

**) *University of Leiden.*

1. INTRODUCTION.

During the past decade a good deal of effort has been devoted to extending the theory of Berry-Esseen bounds and Edgeworth expansions to more complicated sequences of random variables than normalized sums of independent and identically distributed (i.i.d.) random variables or vectors. From a statistical point of view, this study of higher order asymptotics for large classes of test statistics and estimators has proved extremely fruitful: it has yielded much that is significant for statistical theory as well as useful in practical applications. To the probabilist, however, most test statistics and estimators occurring in statistical theory appear to be strange artefacts, which are neither particularly interesting objects for study in themselves nor very promising starting points for developing a general probabilistic theory.

There is, perhaps, one exception which is the class of U - statistics introduced by Hoeffding (1948). Though it is usually studied for its statistical applications, it surely constitutes a large class of random variables which would seem to be a natural extension of sums of i.i.d. random variables. Let $X_1, X_2, \ldots$ be i.i.d. random variables and let $h : \mathbb{R}^k \to \mathbb{R}$ be a symmetric function of its k arguments. For $N \geq k$, a U - statistic of degree k is defined as

$$(1.1) \qquad U = \sum \cdots \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq N} h(X_{i_1}, X_{i_2}, \ldots, X_{i_k})$$

and the idea is to study its asymptotic behavior for a fixed h as $N \to \infty$. For $k = 1$, we are back in the case of sums of i.i.d. random variables. As soon as $k \geq 2$, the degree doesn't play an important role any more except, of course, for the fact that it stays fixed as $N \to \infty$. Many authors therefore discuss only the case of degree two, on the understanding that the case $k > 2$ is similar. Let us follow this tradition for a moment and take

$$(1.2) \qquad U = \sum\sum_{1 \leq i < j \leq N} h(X_i, X_j) ,$$

where $h(x,y) = h(y,x)$. Assume that

$$(1.3) \qquad E\, h(X_1, X_2) = 0 , \quad E\, h^2(X_1, X_2) < \infty ,$$

and define

$$(1.4) \qquad g(x) = E(h(X_1, X_2) | X_1 = x) , \qquad \psi(x,y) = h(x,y) - g(x) - g(y) ,$$

$$(1.5)\qquad \hat{U} = (N-1)\sum_{i=1}^{N} g(X_i)\ , \quad \Delta = \sum\sum_{1\le i<j\le N} \psi(X_i,X_j)\ .$$

Clearly, $E(\psi(X_1,X_2)|X_1) = 0$ a.s. so that the random variables $g(X_i)$ and $\psi(X_i,X_j)$ are pairwise uncorrelated and since $U = \hat{U} + \Delta$,

$$(1.6)\qquad \sigma^2(U) = \sigma^2(\hat{U}) + \sigma^2(\Delta) = N(N-1)^2 Eg^2(X_1) + \tfrac{1}{2}N(N-1)E\psi^2(X_1,X_2)\ .$$

If it is assumed that

$$(1.7)\qquad E\, g^2(X_1) > 0\ ,$$

then $\sigma^2(\hat{U})$ dominates the right-hand side of (1.6) and $U\,\sigma^{-1}(U)$ is asymptotically normal (cf. Hoeffding (1948)).

The speed of convergence to normality was investigated by a number of authors who proved in increasing generality that

$$(1.8)\qquad \sup_x \left| P\left(\frac{U}{\sigma(U)} \le x\right) - \Phi(x)\right| = \mathcal{O}(N^{-\frac{1}{2}})\ ,$$

where Φ denotes the standard normal distribution function (d.f.). Suppose that (1.3) and (1.7) are satisfied so that asymptotic normality is ensured. Bickel (1974) established the Berry-Esseen bound (1.8) under the additional assumption that h is bounded. Chan and Wierman (1977) and Callaert and Janssen (1978) successively reduced this assumption first to $E\, h^4(X_1,X_2) < \infty$ and then to $E|h(X_1,X_2)|^3 < \infty$. Helmers and Van Zwet (1982) showed that $E|g(X_1)|^3 < \infty$ suffices. They also proved that the assumption $E\, h^2(X_1,X_2) < \infty$ in (1.3) may be relaxed, provided $\sigma(U)$ is replaced by $\sigma(\hat{U})$ in (1.8). This need not concern us here, however, since we shall concentrate on the case of finite variance in the present paper.

Let us consider the more general case of a symmetric statistic. As before, let $X_1,\ldots,X_N$ be i.i.d. and let $\tau : \mathbb{R}^N \to \mathbb{R}$ be a symmetric function of its N arguments.
Define

$$(1.9)\qquad T = \tau(X_1,\ldots,X_N)$$

and assume that

$$(1.10)\qquad E\,T = 0\ , \qquad E\,T^2 = 1\ .$$

We wish to study the asymptotic behavior of T as $N \to \infty$. The difference with the previous problem is that then we were dealing with a kernel function h that remains fixed as $N \to \infty$, or perhaps with uniformity classes of such functions of a fixed degree k. Now the degree of the kernel τ equals the sample size N and both tend to infinity together.

Define

$$(1.11) \quad T_j = E(T|X_j) , \qquad \hat{T}_1 = \sum_{j=1}^{N} T_j ,$$

then $\hat{T}_1$ and $(T-\hat{T}_1)$ are again uncorrelated. It follows that if $\sigma^2(T) \sim \sigma^2(\hat{T}_1)$ as $N \to \infty$ and the summands T_j satisfy the Lindeberg condition, then $T\,\sigma^{-1}(T)$ is asymptotically normal.

The aim of this paper is to prove the following theorem of Berry-Esseen type.

<u>THEOREM 1.1.</u>

Suppose that (1.10) *is satisfied and that positive numbers* A *and* B *exist such that*

$$(1.12) \quad E|E(T|X_1)|^3 \le A\, N^{-3/2} ,$$

$$(1.13) \quad 1 + E\{E(T|X_1,\ldots,X_{N-2})\}^2 - 2E\{E(T|X_1,\ldots,X_{N-1})\}^2 \le B\, N^{-3} .$$

Then

$$(1.14) \quad \sup_x |P(T \le x) - \Phi(x)| \le C(A+B)N^{-\frac{1}{2}} ,$$

where C *denotes a universal constant.*

Note that although we have formulated the theorem as a uniform error bound for a fixed but arbitrary N and T, it is a purely asymptotic result because the constant C is not specified. It applies to sequences of symmetric statistics $T_N = \tau_N(X_{N,1},\ldots,X_{N,N})$ where, for every fixed N, $X_{N,1},\ldots,X_{N,N}$ are i.i.d. with a common d.f. F_N, provided (1.10), (1.12) and (1.13) are satisfied for every N and fixed values of A and B.

The theorem will be proved in sections 2 and 3. In section 2 we collect some facts concerning L_2 - projections and in section 3 we provide a proof of the theorem based on these facts. Some examples and possible extensions are discussed in sections 4 and 5.

2. L_2 - PROJECTIONS.

L_2 - projections were introduced in statistics by Hoeffding (1948, 1961) and have been used effectively by many authors since then. Most recently Efron and Stein (1981) and Karlin and Rinott (1982) have used these orthogonal projections to establish certain variance inequalities. To indicate decomposition by repeated orthogonal projection, these authors have introduced the descriptive term *ANOVA - type decomposition*, but we prefer to speak of *Hoeffding's decomposition* instead. What follows are some simple and well-known facts concerning L_2 - projections written down in an easy notation.

Let $X_1,\dots,X_N$ be independent random variables and let $T = \tau(X_1,\dots,X_N)$ have $E\,T^2 < \infty$. Note that at this point we do *not* assume that $X_1,\dots,X_N$ are identically distributed, that τ is symmetric in its N arguments, or that $E\,T = 0$ and $E\,T^2 = 1$. Define $\Omega = \{1,2,\dots,N\}$. For any $D \subset \Omega$, let

$$(2.1)\qquad E(T|D) = E(T|X_i,\ i \in D)$$

denote the conditional expectation given all X_i with indices in D. Define

$$(2.2)\qquad T_D = \sum_{A\subset D} (-1)^{|D|-|A|} E(T|A),$$

where the summation is over all subsets A of D, including the empty set, and $|\cdot|$ denotes the cardinality of a set. Of course $T_\phi = E(T|\phi) = E\,T$ a.s. and for convenience we shall write

$$(2.3)\qquad T_j = T_{\{j\}} = E(T|X_j) - E\,T,\quad j = 1,\dots,N.$$

The basic property of T_D is that

$$(2.4)\qquad E(T_D|D') = 0 \text{ a.s. unless } D \subset D'.$$

To see this, write $C = D \cap D'$ and note that, if $|D|-|C| = k > 0$,

$$E(T_D|D') = \sum_{A\subset D} (-1)^{|D|-|A|} E(T|A\cap C) = \sum_{B\subset C} E(T|B) \sum_{j=0}^{k} (-1)^{|D|-|B|-j}\binom{k}{j} = 0 \text{ a.s.}.$$

It follows in particular that $E\,T_D = 0$ if $D \neq \phi$ and that the random variables T_D, $D \subset \{1,\dots,N\}$ are pairwise uncorrelated, i.e.

(2.5) $\quad E\, T_D\, T_{D'} = 0 \qquad \text{if} \quad D \neq D'$.

Since the order of the two operations in $E(T_D|D')$ may be interchanged with impunity, we have $E(T_D|D') = [E(T|D')]_D$. Hence (2.4) also yields that if T depends only on X_i for $i \in D'$, then

(2.6) $\quad T_D = 0 \ \text{a.s.} \qquad \text{unless} \quad D \subset D'$.

For $m = 0,1,\ldots,N$, let L_m denote the linear space of random variables with finite variance that is spanned by functions of at most m of the variables $X_1,\ldots,X_N$, thus

$$L_m = \{Z : Z = \sum_{1\le i_1<i_2<\ldots<i_m\le N} \psi_{i_1,\ldots,i_m}(X_{i_1},\ldots,X_{i_m}) \ , \ E\, Z^2 < \infty\} .$$

We define $\hat{T}_m$ to be the L_2 - projection of T on L_m if $\hat{T}_m \in L_m$ and $E(T-\hat{T}_m)^2$ is minimal, or equivalently, if $\hat{T}_m \in L_m$ and $E(T-\hat{T}_m)Z = 0$ for all $Z \in L_m$. We have

(2.7) $\quad \hat{T}_0 = E\, T \ , \quad \hat{T}_1 - \hat{T}_0 = \sum_{j=1}^{N} T_j \ , \quad \hat{T}_m - \hat{T}_{m-1} = \sum_{|D|=m} T_D \ , \quad \hat{T}_N = T$.

To check this, note that $\hat{T}_m \in L_m$ and that $E\, T_D\, Z = 0$ if $|D| \ge m+1$ and $Z \in L_m$ by (2.4). Hence we have Hoeffding's decomposition

(2.8) $\quad T = \hat{T}_0 + (\hat{T}_1-\hat{T}_0)+\ldots+(\hat{T}_N-\hat{T}_{N-1}) = \sum_{D\subset\Omega} T_D$

and since all terms are pairwise uncorrelated,

(2.9) $\quad E\, T^2 = \sum_{D\subset\Omega} E\, T_D^2$.

If we apply (2.8) to $E(T|A)$ instead of T , (2.6) yields

(2.10) $\quad E(T|A) = \sum_{D\subset A} T_D$

which is the inverse of relation (2.2).

For $m = 0,1,\ldots,N$, let us write

(2.11) $\quad W_m = E(T|X_{m+1},\ldots,X_N)$,

$$(2.12)\quad T = \sum_{j=1}^{m} T_j + W_m + \Delta_m .$$

Clearly $\sum_{j=1}^{m} T_j + W_m$ is the best approximation of T in L_2 by a random variable which depends on $X_1,\ldots,X_m$ only through a sum of functions of each one of these variables separately. We shall need some information concerning the error Δ_m of this approximation. For $r = 0,1,\ldots,N$, define

$$(2.13)\quad \Omega_r = \{1,2,\ldots,r\} , \quad \Omega_r^c = \Omega-\Omega_r = \{r+1,\ldots,N\} .$$

By (2.10) and (2.8) ,

$$(2.14)\quad W_m = \sum_{D\subset\Omega_m^c} T_D ,$$

$$(2.15)\quad \Delta_0 = 0 , \quad \Delta_m = \sum_{\substack{D\cap\Omega_m\neq\phi \\ |D|\geq 2}} T_D = \sum_{\substack{k=1 \\ k+\ell\geq 2}}^{m} \sum_{\ell=0}^{N-m} \sum_{\substack{A\subset\Omega_m \\ |A|=k}} \sum_{\substack{B\subset\Omega_m^c \\ |B|=\ell}} T_{A\cup B} .$$

Now let us assume that $X_1,\ldots,X_N$ are identically distributed, that $T = \tau(X_1,\ldots,X_N)$ is a symmetric function of these variables and that $E\,T = 0$, $E\,T^2 = 1$, so that we are back in the situation of section 1. Then (2.15) and (2.5) imply that

$$(2.16)\quad E\,\Delta_m^2 = \sum_{r=2}^{N}\left\{\binom{N}{r} - \binom{N-m}{r}\right\} E\,T_{\Omega_r}^2 , \quad m = 0,1,\ldots,N .$$

If $D(E\,\Delta_m^2) = E\,\Delta_{m+1}^2 - E\,\Delta_m^2$ and $D^{s+1}(E\,\Delta_m^2) = D\,D^s(E\,\Delta_m^2)$, then (2.16) yields

$$(-1)^{s+1}\, D^s(E\,\Delta_m^2) = \sum_{r=2}^{N-m}\binom{N-m-s}{r-s} E\,T_{\Omega_r}^2 \geq 0 , \quad s \geq 1 ,$$

(cf. Karlin and Rinott (1982) who show that $E\,W_{N-m}^2 = 1 - (N-m)\,E\,T_1^2 - E\,\Delta_{N-m}^2$ is absolutely monotone). In particular, $E\,\Delta_m^2$ is nondecreasing and concave for $m = 0,1,\ldots,N$. Also

$$(2.17)\quad \begin{aligned} 0 \leq -D^2(E\,\Delta_0^2) &= 2\,E\,\Delta_1^2 - E\,\Delta_2^2 = 2(1 - E\,T_1^2 - E\,W_1^2) - (1-2E\,T_1^2 - E\,W_2^2) = \\ &= 1 + E\{E(T|X_1,\ldots,X_{N-2})\}^2 - 2\,E\{E(T|X_1,\ldots,X_{N-1})\}^2 \end{aligned}$$

and under the conditions of theorem 1.1 we therefore have

$$(2.18)\quad 0 \le 2\,E\,\Delta_1^2 - E\,\Delta_2^2 = \sum_{r=2}^{N} \binom{N-2}{r-2} E\,T_{\Omega_r}^2 \le B\,N^{-3} .$$

It follows that

$$(2.19)\quad 0 \le E\,\Delta_1^2 = \sum_{r=2}^{N} \binom{N-1}{r-1} E\,T_{\Omega_r}^2 \le B\,N^{-2} ,$$

$$(2.20)\quad 0 \le E\,\Delta_N^2 = \sum_{r=2}^{N} \binom{N}{r} E\,T_{\Omega_r}^2 \le \tfrac{1}{2}\,B\,N^{-1} ,$$

$$(2.21)\quad 0 \le E\,\Delta_m^2 \le m\,E\,\Delta_1^2 \le B\,m\,N^{-2} ,\quad m = 0,\dots,N ,$$

because of the concavity of $E\,\Delta_m^2$.

So far we have implicitly assumed that the random variable T is real valued, but of course everything in this section goes through for complex valued T with appropriate modifications. In (2.5), $E\,T_D\,T_{D'}$, should be replaced by $E\,T_D\,\overline{T}_{D'}$, where $\overline{T}_{D'}$ denotes the complex conjugate of $T_{D'}$; furthermore, in all expectations of squares such as $E\,T^2$, $E\,T_D^2$, $E\,W_m^2$, $E\,\Delta_m^2$ etc., the squares should be replaced by their moduli $E|T^2|$, $E|T_D^2|$, $E|W_m^2|$, $E|\Delta_m^2|$ etc.. Thus in particular (2.9) becomes

$$(2.22)\quad E|T^2| = \sum_{D\subset\Omega} E|T_D^2| .$$

3. PROOF OF THEOREM 1.1.

Let us agree to take $C \ge 3$. For $1 \le N \le 3\,B$, we have $C(A+B)N^{-\frac{1}{2}} \ge C\,B\,N^{-\frac{1}{2}} \ge C\,N^{\frac{1}{2}}/3 \ge 1$, so that (1.14) is trivially satisfied. We therefore assume that $N > 3B$.

In view of (2.12) and (2.20),

$$(3.1)\quad |E\,T_1^2 - N^{-1}| \le \tfrac{1}{2}\,B\,N^{-2} \le \frac{1}{6N}$$

and hence, under the conditions of the theorem,

$$(3.2)\quad A \ge N^{3/2}\,E|T_1|^3 \ge (N\,E\,T_1^2)^{3/2} \ge (1 - \tfrac{1}{2}\,B\,N^{-1})^{3/2} \ge \left(\frac{5}{6}\right)^{3/2} .$$

Let

$$(3.3)\qquad \gamma(t) = E\, e^{it\, T_1}$$

be the characteristic function of T_1. By (3.1) and (1.12),

$$|\gamma(t) - 1 - \frac{t^2}{2N}| \le \frac{1}{4} B N^{-2} t^2 + \frac{1}{6} A N^{-3/2} |t|^3 \le \frac{t^2}{6N}$$

for all $|t| \le H = \frac{1}{2} A^{-1} N^{\frac{1}{2}}$. For $|t| \le H$, we have $t^2 \le (6/5)^3 N/4 \le \frac{1}{2} N$ and

$$(3.4)\qquad 0 < 1 - \frac{2t^2}{3N} \le |\gamma(t)| \le 1 - \frac{t^2}{3N} \le \exp\{-\frac{t^2}{3N}\},$$

$$(3.5)\qquad \frac{t^2}{2N} \le 1 - |\gamma^2(t)| \le \frac{4t^2}{3N}.$$

Let

$$(3.6)\qquad \psi(t) = E\, e^{it\, T}$$

denote the characteristic function of T. According to Esseen's smoothing lemma (cf. Feller (1971), p. 538)

$$\sup_x |P(T \le x) - \Phi(x)| \le \frac{1}{\pi} \int_{-H}^{H} \left|\frac{\psi(t) - e^{-\frac{1}{2}t^2}}{t}\right| dt + \frac{4}{H}.$$

Define $h = \min(2 N^{1/4}, H)$ and let $C_1, C_2, \ldots$ denote universal constants throughout the proof. From (1.12), (3.1) and the proof of the classical Berry-Esseen theorem we conclude that

$$\int_{-h}^{h} \left|\frac{\gamma^N(t) - e^{-\frac{1}{2}t^2}}{t}\right| dt \le C_1 A N^{-\frac{1}{2}}.$$

Because of (3.2)

$$\int_{|t| \ge h} \left|\frac{e^{-\frac{1}{2}t^2}}{t}\right| dt \le \frac{1}{2e^2} N^{-\frac{1}{2}} \le A N^{-\frac{1}{2}}$$

and combining these results we find

$$(3.7)\qquad \begin{aligned} \sup_x |P(T \le x) - \Phi(x)| &\le \frac{1}{\pi} \int_{-h}^{h} \left|\frac{\psi(t) - \gamma^N(t)}{t}\right| dt + \\ &+ \frac{1}{\pi} \int_{h \le |t| \le H} \left|\frac{\psi(t)}{t}\right| dt + C_2 A N^{-\frac{1}{2}}. \end{aligned}$$

To analyze $\psi(t)$ for $|t| \le h$, we employ decomposition (2.10) for $m = N$, i.e. $T = \hat{T}_1 + \Delta_N$, to obtain

$$(3.8) \quad \psi(t) = E\, e^{it\hat{T}_1}(1 + it\,\Delta_N) + R_N = \gamma^N(t) + it\,E\, e^{it\hat{T}_1}\Delta_N + R_N ,$$

$$(3.9) \quad |R_N| \le \tfrac{1}{2} t^2 E\,\Delta_N^2 \le \frac{Bt^2}{4N}$$

in view of (2.20). Similarly,

$$(3.10) \quad |t\,E\, e^{it\hat{T}_1}\Delta_N| \le |t|\{E\,\Delta_N^2\}^{\frac{1}{2}} \le (\tfrac{1}{2}B)^{\frac{1}{2}}\,|t|\,N^{-\frac{1}{2}} .$$

A more delicate analysis starts with noting that

$$E\, e^{it\hat{T}_1}\Delta_N = \sum_{k=2}^{N}\sum_{|D|=k} E\, e^{it\hat{T}_1} T_D =$$

$$= \sum_{r=2}^{N}\binom{N}{r}\gamma^{N-r}(t)\,E\,T_{\Omega_r}\prod_{j=1}^{r} e^{it\,T_j} =$$

$$= \sum_{r=2}^{N}\binom{N}{r}\gamma^{N-r}(t)\,E\,T_{\Omega_r}\prod_{j=1}^{r}(e^{it\,T_j} - \gamma(t))$$

where the final step follows from (2.4). For $2 \le r \le N$,

$$\binom{N}{r}^2 \le 6\binom{N-2}{r-2}\binom{N+2}{r+2}$$

and since

$$(3.11) \quad E|e^{it\,T_j} - \gamma(t)|^2 = 1 - |\gamma^2(t)| ,$$

repeated application of Schwarz's inequality yields

$$|E\, e^{it\hat{T}_1}\Delta_N| \le 6^{\frac{1}{2}}\sum_{r=2}^{N}\binom{N-2}{r-2}^{\frac{1}{2}}(E\,T_{\Omega_r}^2)^{\frac{1}{2}} \cdot$$

$$\cdot \binom{N+2}{r+2}^{\frac{1}{2}}|\gamma^2(t)|^{\frac{1}{2}(N-r)}(1 - |\gamma^2(t)|)^{\frac{1}{2}r} \le \frac{6^{\frac{1}{2}}}{1-|\gamma^2(t)|}\cdot\left[\sum_{r=2}^{N}\binom{N-2}{r-2}E\,T_{\Omega_r}^2\right]^{\frac{1}{2}} \cdot$$

$$\cdot\left[\sum_{r=2}^{N}\binom{N+2}{r+2}|\gamma^2(t)|^{N-r}(1 - |\gamma^2(t)|)^r\right]^{\frac{1}{2}} \le \frac{6^{\frac{1}{2}}}{1-|\gamma^2(t)|}\left[\sum_{r=2}^{N}\binom{N-2}{r-2}E\,T_{\Omega_r}^2\right]^{\frac{1}{2}} .$$

Invoking (2.18) and (3.5), we see that for $|t| \le H$

$$(3.12)\quad |t\, E\, e^{it\hat{T}_1} \Delta_N| \le (24B)^{\frac{1}{2}} |t|^{-1} N^{-\frac{1}{2}} .$$

Combining (3.8), (3.9), (3.10) and (3.12) and then using (3.2), we arrive at

$$(3.13)\quad \int_{-h}^{h} \left|\frac{\psi(t)-\gamma^N(t)}{t}\right| dt \le (B + 8B^{\frac{1}{2}})\, N^{-\frac{1}{2}} \le 6(A+B)\, N^{-\frac{1}{2}} .$$

It remains to consider $\psi(t)$ for $h \le |t| \le H$ in order to bound the second integral in (3.7). For any fixed $|t|$ in this interval we take

$$(3.14)\quad m = \left[\frac{3\, N \log N}{t^2}\right] ,$$

where $[x]$ denotes the integer part of x. For $|t| \ge h$, we have $0 \le m \le N$, and using decomposition (2.12) for this value of m, we obtain

$$(3.15)\quad \psi(t) = E \exp\{it(\sum_{j=1}^{m} T_j + W_m)\}\cdot(1 + it\,\Delta_m) + R_m ,$$

$$(3.16)\quad |R_m| \le \tfrac{1}{2}\, t^2\, E\, \Delta_m^2 \le \frac{B\, m\, t^2}{2\, N^2} \le \frac{3\, B \log N}{2N}$$

because of (2.21). Since $|t| \le H$, (3.4) and (3.2) imply

$$(3.17)\quad \begin{aligned} &|E \exp\{it(\sum_{j=1}^{m} T_j + W_m)\}| \le |\gamma(t)|^m \le \exp\{-\frac{mt^2}{3N}\} \le \\ &\le \exp\{-\log N + \frac{t^2}{3N}\} \le N^{-1} \exp\{\frac{1}{12\, A^2}\} \le \frac{2A}{N} . \end{aligned}$$

Let us define the complex valued random variable $Z = \exp\{it\, W_m\}$ which depends on $X_{m+1},\ldots,X_N$ only. By (2.15) and two applications of (2.4),

$$E \exp\{it(\sum_{j=1}^{m} T_j + W_m)\}\, \Delta_m =$$

$$(3.18)\quad \begin{aligned} &= \sum_{\substack{k=1\\ k+\ell\ge 2}}^{m} \sum_{\ell=0}^{N-m} \sum_{\substack{A\subset\Omega_m\\ |A|=k}} \sum_{\substack{B\subset\Omega_m^c\\ |B|=\ell}} \gamma^{m-k}(t) \cdot E\Big[T_{A\cup B} \prod_{j\in A} e^{it\,T_j}\, E(Z|B)\Big] = \\ &= \sum_{\substack{k=1\\ k+\ell\ge 2}}^{m} \sum_{\ell=0}^{N-m} \sum_{\substack{A\subset\Omega_m\\ |A|=k}} \sum_{\substack{B\subset\Omega_m^c\\ |B|=\ell}} \gamma^{m-k}(t) \cdot E\Big[T_{A\cup B} \prod_{j\in A} (e^{it\,T_j} - \gamma(t))\, Z_B\Big] . \end{aligned}$$

It follows from (2.22) and (2.6) that

$$(3.19)\qquad \sum_{B\subset\Omega_m^c} E|Z_B^2| = E|Z^2| = 1 .$$

By Schwarz's inequality and (3.11),

$$E|T_{A\cup B} \prod_{j\in A} (e^{itT_j} - \gamma(t)) Z_B| \le (E\, T^2_{A\cup B})^{\frac12}(1-|\gamma^2(t)|)^{\frac12|A|}(E|Z_B^2|)^{\frac12}$$

for every $A\subset\Omega_m$ and $B\subset\Omega_m^c$. Another application of Schwarz's inequality to the terms in (3.18) with $k = 1$ and $k\ge 2$ separately, followed by the use of (2.18) and (2.19) yields

$$(3.20)\qquad \begin{aligned} &|E \exp\{it(\sum_{j=1}^m T_j + W_m)\}\Delta_m| \le m|\gamma(t)|^{m-1}(1-|\gamma^2(t)|)^{\frac12}\cdot \\ &\cdot\left[\sum_{\ell=1}^{N-m}\sum_{\substack{B\subset\Omega_m^c\\|B|=\ell}} E\,T^2_{\Omega_{\ell+1}}\right]^{\frac12}\left[\sum_{\ell=1}^{N-m}\sum_{\substack{B\subset\Omega_m^c\\|B|=\ell}} E|Z_B^2|\right]^{\frac12}+\left[\sum_{k=2}^{m}\sum_{\ell=0}^{N-m}\sum_{\substack{A\subset\Omega_m\\|A|=k}}\sum_{\substack{B\subset\Omega_m^c\\|B|=\ell}}\frac{k(k-1)}{m(m-1)} E\,T^2_{A\cup B}\right]^{\frac12}\cdot \\ &\cdot\left[\sum_{k=2}^{m}\sum_{\substack{A\subset\Omega_m\\|A|=k}}\frac{m(m-1)}{k(k-1)}|\gamma^2(t)|^{m-k}(1-|\gamma^2(t)|)^k\sum_{B\subset\Omega_m^c}E|Z_B^2|\right]^{\frac12}\le \\ &\le m|\gamma(t)|^{m-1}(1-|\gamma^2(t)|)^{\frac12}\left[\sum_{r=2}^{N-m+1}\binom{N-m}{r-1} E\,T^2_{\Omega_r}\right]^{\frac12}+ \\ &+6^{\frac12}\left[\sum_{r=2}^{N}\binom{N-2}{r-2} E\,T^2_{\Omega_r}\right]^{\frac12}\left[\sum_{k=2}^{m}\binom{m+2}{k+2}|\gamma^2(t)|^{m-k}(1-|\gamma^2(t)|)^k\right]^{\frac12}\le \\ &\le B^{\frac12}\left[\frac{m}{N}|\gamma(t)|^{m-1}(1-|\gamma^2(t)|)^{\frac12}+6^{\frac12}N^{-3/2}(1-|\gamma^2(t)|)^{-1}\right] . \end{aligned}$$

Hence, by (3.4), (3.5), (3.14) and (3.2),

$$(3.21)\qquad \begin{aligned} &|t\, E \exp\{it(\sum_{j=1}^m T_j + W_m)\}\Delta_m| \le \\ &\le (3B)^{\frac12}\left[2N^{-3/2}\log N \exp\{\frac{2t^2}{3N}\} + 2^{3/2}N^{-\frac12}|t|^{-1}\right]\le \\ &\le 5B^{\frac12}[N^{-3/2}\log N + N^{-\frac12}|t|^{-1}] \end{aligned}$$

for $h\le |t|\le H$. Combining (3.15) - (3.17) and (3.21) and again using (3.2), we arrive at

$$\int_{h \le |t| \le H} \left|\frac{\psi(t)}{t}\right| dt \le \frac{3\, B(\log N)^2}{4N} + \frac{A \log N}{N} + \frac{5\, B^{\frac{1}{2}}(\log N)^2}{2\, N^{3/2}} + \frac{5\, B^{\frac{1}{2}}}{N^{3/4}} \le \tag{3.22}$$

$$\le 7(A + B)\, N^{-\frac{1}{2}} .$$

Together (3.7), (3.13) and (3.22) establish theorem 1.1. □

4. EXAMPLES.

In this section we apply theorem 1.1 to two special cases - U-statistics and linear functions of order statistics - to see whether we can obtain results comparable to the best available ones for these well-studied special cases.

Let $X_1,\ldots,X_N$ be i.i.d. random variables and let h be a function of $k(\le N)$ variables satisfying

$$E\, h(X_1,\ldots,X_k) = 0 \;, \quad E\, h^2(X_1,\ldots,X_k) < \infty \;. \tag{4.1}$$

Define the U-statistic U by (1.1), the function g by

$$g(x) = E(h(X_1,\ldots,X_k)\,|\,X_1 = x) \tag{4.2}$$

and suppose that

$$E\, g^2(X_1) > 0 \;, \qquad E|g_1(X_1)|^3 < \infty \;. \tag{4.3}$$

We shall show that theorem 1.1 implies

COROLLARY 4.1.

There exists a universal constant C *such that*

$$\sup_x \left| P\left(\frac{U}{\sigma(U)} \le x\right) - \Phi(x)\right| \le C \left[\frac{E|g(X_1)|^3}{\{Eg^2(X_1)\}^{3/2}} + \frac{(k-1)^2\, Eh^2(X_1,\ldots,X_k)}{Eg^2(X_1)}\right] N^{-\frac{1}{2}}$$

whenever $1 \le k \le N$ *and provided* (4.1) *and* (4.3) *are satisfied.*

For $k = 2$ this is the best result known for the case where $E\, h^2(X_1,\ldots,X_k) < \infty$, as was pointed out in section 1. Since the assumption of finite variance is a natural limitation of the results in this paper, we conclude that theorem 1.1

performs as well as might be expected for this special case. This is not really surprising, as theorem 1.1 and its proof are modeled after the earlier work on U-statistics.

To prove the corollary, we begin by noting that (2.6) implies that

$$(4.4)\qquad U_D = 0 \qquad \text{if} \quad |D| \geq k+1 .$$

For $r = 0,1,\ldots,k$, define

$$(4.5)\qquad g_r(X_1,\ldots,X_r) = (h(X_1,\ldots,X_k))_{\Omega_r} = \sum_{A\subset\Omega_r} (-1)^{r-|A|} E(h(X_1,\ldots,X_k)|A) .$$

In particular, $g_0 = 0$ and $g_1 = g$ as defined in (4.2). It follows from (2.9) that

$$(4.6)\qquad E\,h^2(X_1,\ldots,X_k) = \sum_{r=0}^{k} \binom{k}{r} E\,g_r^2(X_1,\ldots,X_r) .$$

Obviously, for $r = 0,1,\ldots,k$,

$$(4.7)\qquad U_{\Omega_r} = \binom{N-r}{k-r} g_r(X_1,\ldots,X_r)$$

and because of (2.7), (4.4) and (4.6) we have

$$(4.8)\qquad E\,\hat{U}_1^2 = N\,E\,U_1^2 = N\binom{N-1}{k-1}^2 E\,g^2(X_1) ,$$

$$(4.9)\qquad E|U_1|^3 = \binom{N-1}{k-1}^3 E|g(X_1)|^3 ,$$

$$(4.10)\qquad \begin{aligned} \sum_{r=2}^{N} \binom{N-2}{r-2} E\,U_{\Omega_r}^2 &= \sum_{r=2}^{k} \binom{N-2}{r-2}\binom{N-r}{k-r}^2 E\,g_r^2(X_1,\ldots,X_r) = \\ &= \binom{N}{k} \sum_{r=2}^{k} \frac{r(r-1)}{N(N-1)} \binom{N-r}{k-r} \cdot \binom{k}{r} E\,g_r^2(X_1,\ldots,X_r) \leq \binom{N-2}{k-2}^2 E\,h^2(X_1,\ldots,X_k) . \end{aligned}$$

Define $T = U/\sigma(U)$, so that $E\,T^2 = 1$. Take

$$(4.11)\qquad A = \frac{E|g(X_1)|^3}{\{Eg^2(X_1)\}^{3/2}} , \quad B = 4(k-1)^2 \frac{Eh^2(X_1,\ldots,X_k)}{Eg^2(X_1)} .$$

By (4.8) - (4.10),

$$E|T_1|^3 = \frac{E|U_1|^3}{\{EU^2\}^{3/2}} \leq \frac{E|U_1|^3}{\{E\hat{U}_1^2\}^{3/2}} = A\,N^{-3/2} ,$$

$$(4.12)\quad \sum_{r=2}^{N} \binom{N-2}{r-2} E\, T_{\Omega_r}^2 \le \frac{\binom{N-2}{k-2}^2 Eh^2(X_1,\ldots,X_k)}{E\, \hat{U}_1^2} \le B\, N^{-3} .$$

Note that the results of these computations are correct also for $k = 1$. In view of (2.17) and (2.18), it follows that assumptions (1.12) and (1.13) of theorem 1.1 are satisfied with A and B as in (4.11). The corollary follows.

We now turn to our second example. Let $X_1,X_2,\ldots,X_N$ be i.i.d. random variables with a common distribution function F , which is not assumed to be continuous. Let $X_{(1)} \le X_{(2)} \le \ldots \le X_{(N)}$ denote the corresponding order statistics. For real numbers $c_1,c_2,\ldots,c_N$, we consider a normed linear function of order statistics

$$(4.13)\quad L = N^{-\frac{1}{2}} \sum_{j=1}^{N} c_j (X_{(j)} - E\, X_{(j)}) .$$

Suppose that

$$(4.14)\quad E|X_1|^3 < \infty , \qquad \sigma^2(L) > 0 ,$$

and let

$$(4.15)\quad \max_{1\le j\le N} |c_j| = a , \qquad N \max_{2\le j\le N} |c_j - c_{j-1}| = b .$$

Theorem 1.1 implies

<u>COROLLARY 4.2.</u>

There exists a universal constant C *such that*

$$\sup_x \left|P\left(\frac{L}{\sigma(L)} \le x\right) - \Phi(x)\right| \le C \left[\frac{a^3 E|X_1|^3}{\sigma^3(L)} + \frac{b^2\{E|X_1|\}^2}{\sigma^2(L)}\right] N^{-\frac{1}{2}}$$

whenever (4.14) *and* (4.15) *are satisfied.*

If $\sigma^2(L)$ is bounded below and $E|X_1|^3$, a and b are bounded above as $N \to \infty$, then corollary 4.2 provides a Berry-Esseen bound of order $N^{-\frac{1}{2}}$. In view of (4.15) we are then dealing with the case of smooth weights $c_1,\ldots,c_N$, but not necessarily smooth underlying distribution function F . For this case, the best result to date has been obtained by Helmers (1981; 1982) and this result is essentially equivalent to corollary 4.2. Thus once again, theorem 1.1 appears to perform in a satisfactory manner.

To prove corollary 4.2 we adopt some additional notation. For $n \le N$, $X_{1:n} \le X_{2:n} \le \dots \le X_{n:n}$ will denote the order statistics corresponding to $X_1, X_2, \dots, X_n$; we take $X_{0:n} = -\infty$, $X_{n+1:n} = +\infty$. We shall find it convenient to introduce i.i.d. random variables $U_1, U_2, \dots, U_N$ with a common uniform distribution on $(0,1)$ and pretend that $X_i = F^{-1}(U_i)$ for $i = 1, \dots, N$. Clearly this does not affect the distribution of L. The rank of U_i among $U_1, \dots, U_N$ will be denoted by R_i,

$$R_i = \sum_{k=1}^{N} 1_{(0,U_i]}(U_k) ,$$

and we define

$$(4.16) \quad K_1 = R_{N-1} \wedge R_N , \qquad K_2 = R_{N-1} \vee R_N ,$$

where $x \wedge y = \min(x,y)$ and $x \vee y = \max(x,y)$. Furthermore we let $b_{j,N}$ be the beta density

$$b_{j,N}(y) = \frac{N!}{(j-1)!(N-j)!} y^{j-1}(1-y)^{N-j} , \quad 0 < y < 1 ,$$

and we define the function G, H and M by

$$(4.17) \quad G(x) = \int_{-\infty}^{x} F(y)dy , \quad H(x) = \int_{x}^{\infty} (1-F(y))dy , \quad M(x) = \int_{-\infty}^{x} F(y)(1-F(y))dy .$$

Obviously G, H and M are monotone and by (4.14), M is bounded. Finally we introduce the random variable

$$(4.18) \quad Z = L - E(L|U_1,\dots,U_{N-1}) - E(L|U_1,\dots,U_{N-2},U_N) + E(L|U_1,\dots,U_{N-2})$$

and note that

$$(4.19) \quad E Z^2 = E L^2 + E\{E(L|U_1,\dots,U_{N-2})\}^2 - 2 E\{E(L|U_1,\dots,U_{N-1})\}^2 .$$

Straightforward but somewhat tedious computations show that with probability 1

$$(4.20) \quad N^{\frac{1}{2}} L_1 = N^{\frac{1}{2}} E(L|U_1) = \frac{1}{N} \sum_{j=1}^{N} c_j \int_0^1 \{1_{(0,U_1)}(y) - (1-y)\} b_{j,N}(y) dF^{-1}(y) ,$$

$$N^{\frac{1}{2}} Z = \sum_{j=1}^{N-1} (c_{j+1}-c_j) \; (M(X_{j:N-2}) - M(X_{j-1:N-2})) +$$

$$(4.21) \qquad - \sum_{j=1}^{K_1} (c_{j+1}-c_j) \; (G(X_{j:N}) - G(X_{j-1:N})) +$$

$$+ \sum_{j=K_2}^{N} (c_j-c_{j-1}) \; (H(X_{j+1:N}) - H(X_{j:N})) \; .$$

By (4.15), $\sum |c_j| b_{j,N}(y) \leq a\, N$ and hence

$$(4.22) \qquad \begin{aligned} N^{\frac{1}{2}}|L_1| &\leq a \left\{ \int_0^{U_1} y dF^{-1}(y) + \int_{U_1}^{1} (1-y) dF^{-1}(y) \right\} \leq \\ &\leq a \left\{ |F^{-1}(U_1)| + \int_0^1 |F^{-1}(y)| dy \right\} = a\{|X_1| + E|X_1|\} \; . \end{aligned}$$

Because of (4.15) and the monotonicity of M, G and H,

$$(4.23) \qquad |Z| \leq b\, N^{-3/2} \left[M(\infty) + G(X_{N-1} \wedge X_N) + H(X_{N-1} \vee X_N) \right] .$$

Define $T = L/\sigma(L)$. Combining (4.14), (4.22) and (4.23) we find after elementary calculations

$$(4.24) \qquad E|T_1|^3 \leq \frac{4a^3\, E|X_1|^3}{\sigma^3(L)} N^{-3/2} \; ,$$

$$(4.25) \qquad \frac{E\, Z^2}{\sigma^2(L)} \leq \frac{25\, b^2 \{E|X_1|\}^2}{\sigma^2(L)} N^{-3} \; .$$

Corollary 4.2 follows from (4.19), (4.24), (4.25) and theorem 1.1.

We should perhaps point out that (4.20) and (4.21) are valid under the sole assumption that $E|X_1| < \infty$ and can therefore be used to treat other cases than the one of smooth weights. Any set of assumptions ensuring that $E|T_1|^3 = \mathcal{O}(N^{-3/2})$ and $E\, Z^2/\sigma^2(L) = \mathcal{O}(N^{-3})$ as $N \to \infty$, will produce a Berry-Esseen bound of order $N^{-\frac{1}{2}}$. Smoothness of the underlying distribution function F can clearly replace smoothness of the weights c_j and intermediate versions are also possible.

5. POSSIBLE EXTENSIONS.

Theorem 1.1 provides a Berry-Esseen bound for a symmetric function τ of i.i.d. random variables $X_1,\ldots,X_N$ under the relatively simple moment assumptions (1.12) and (1.13). For a particular case it may be laborious to check these assumptions, but the work involved is basically straightforward. The technical intricacies of the proof of a Berry-Esseen-type result have been dispensed with and what remains can be done by brute force. Of course this only makes sense up to a point: if too much brute force is needed, one may prefer to tackle the intricacies directly instead.

It would seem that this might be the deciding factor in judging how far the present result can usefully be generalized. There doesn't seem to be a reason, a priori, why one should need the symmetry of τ or the fact that $X_1,\ldots,X_N$ are identically distributed. Hoeffding's decomposition (2.9) works without these assumptions and it should be possible to adapt the remainder of the proof. In short, one should be able to generalize theorem 1.1 to arbitrary functions of independent random variables. Of course the assumptions needed to replace (1.12) and (1.13) will not look nearly as pleasant; worse still, they will probably be almost impossible to check in most nontrivial cases.

One would guess, however, that there is one slight but significant generalization that would still be feasible. This is the k - sample situation, where the independent random variables $X_1,\ldots,X_N$ are split into a fixed number (k) of groups. Within each group the variables are i.i.d. and τ is a symmetric function of the variables in such a group.

Another possible type of extension is to relax the moment assumptions $E\,T^2 < \infty$ and $E|N^{\frac{1}{2}}\,T_1|^3 < \infty$ by the following standard argument. Let $T = \tilde{T} + R$. If we have a Berry-Esseen bound for $\tilde{T}$,

$$(5.1) \qquad \sup_x |P(\tilde{T} \le x) - \Phi(x)| \le c\,N^{-\frac{1}{2}}$$

and R satisfies

$$(5.2) \qquad P(|R| \ge a\,N^{-\frac{1}{2}}) \le b\,N^{-\frac{1}{2}} ,$$

then we have a Berry-Esseen bound for T,

$$(5.3) \qquad \sup_x |P(T \le x) - \Phi(x)| \le (a + b + c)N^{-\frac{1}{2}} .$$

In principle, no moments of R - and therefore of T - are needed, but we note that (5.2) is often established with the aid of a moment of low order and the Markov inequality. We have not incorporated this idea in theorem 1.1 because it is well-known and may be applied ad hoc whenever needed.

The above argument may be used for other purposes than merely to relax the moment assumptions. As we have noted before (cf. (2.17) and (2.18)), assumption (1.13) of theorem 1.1 is equivalent to

$$(5.4) \qquad 2\, E\, \Delta_1^2 - E\, \Delta_2^2 = \sum_{r=2}^{N} \binom{N-2}{r-2} E\, T^2_{\Omega_r} \le B\, N^{-3} .$$

However, if we require that for some positive integer $N' \le N$,

$$(5.5) \qquad E(T - \hat{T}_{N'})^2 = \sum_{r=N'+1}^{N} \binom{N}{r} E\, T^2_{\Omega_r} \le B\, N^{-3/2} ,$$

then

$$P(|T - \hat{T}_{N'}| \ge N^{-\frac{1}{2}}) \le B\, N^{-\frac{1}{2}}$$

and by (5.3) and (3.2) the conclusion of theorem 1.1 will hold for T if it holds for $\hat{T}_{N'}$. But for $\hat{T}_{N'}$ instead of T , assumption (5.4) reduces to

$$(5.6) \qquad \sum_{r=2}^{N'} \binom{N-2}{r-2} E\, T^2_{\Omega_r} \le B\, N^{-3}$$

because of (2.7), (2.6) and (2.4). It follows that (5.5) and (5.6) together may replace assumption (1.13) in theorem 1.1.

We may even go one step further and replace assumption (5.6) in its turn by the requirement that for some N'' with $1 \le N'' \le N'$,

$$(5.7) \qquad \sum_{r=N''+1}^{N'} \binom{N-1}{r-1} E\, T^2_{\Omega_r} \le B(N \log N)^{-2} ,$$

$$(5.8) \qquad \sum_{r=2}^{N''} \binom{N-2}{r-2} E\, T^2_{\Omega_r} \le B\, N^{-3} .$$

To see this, we go over the proof of theorem 1.1 and find that the full force of assumption (5.4) (or (2.18)), as opposed to (2.19), is used only in (3.12) and (3.20). In both places, a strengthened version of (2.19), viz.

$$(5.9) \qquad \sum_{r=2}^{N} \binom{N-1}{r-1} E\, T^2_{\Omega_r} \le B(N \log N)^{-2}$$

would also have been sufficient. Alternatively, we could have required a mixture of (5.4) and (5.9), such as (5.8) combined with

$$(5.10) \qquad \sum_{r=N''+1}^{N} \binom{N-1}{r-1} E\, T^2_{\Omega_r} \le B(N \log N)^{-2} ,$$

and the proof would still have gone through with minor modifications. Applying (5.10) to $\hat{T}_{N'}$ instead of T, we obtain (5.7).

Thus we have shown that (5.5), (5.7) and (5.8) together may replace assumption (1.13) in theorem 1.1. These conditions may be substantially weaker than (1.13), especially if N' and N'' are taken to be of the order of $N^{\frac{1}{2}}(\log N)^{-2}$ and $(\log N)^2$ respectively. In general, however, these assumptions will be hard to check.

REFERENCES.

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

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

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

[4] EFRON, B. and STEIN C. (1981). The jackknife estimate of variance, *Ann. Statist.* 9, 586-596.

[5] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications.* Vol. II, 2nd Ed., Wiley, New York.

[6] HELMERS, R. (1981). A Berry-Esseen theorem for linear combinations of order statistics. *Ann. Probability* 9, 342-347.

[7] HELMERS, R. (1982). *Edgeworth Expansions for Linear Combinations of Order Statistics.* Mathematical Centre Tracts 105. Mathematisch Centrum, Amsterdam.

[8] HELMERS, R. and VAN ZWET, W.R. (1982). The Berry-Esseen bound for U - statistics. *Statistical Decision Theory and Related Topics*, III Vol. 1, S.S. Gupta and J.O. Berger (eds), 497-512. Academic Press, New York.

[9] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distributions. *Ann. Math. Statist.* 19, 293-325.

[10] HOEFFDING, W. (1961). The strong law of large numbers for U - statistics. Inst. of Statist., Univ. of North Carolina, Mimeograph Series No. 302.

[11] KARLIN, S. and RINOTT, Y. (1982). Applications of ANOVA type decompositions for comparisons of conditional variance statistics including jackknife estimates. *Ann. Statist.* 10, 485-501.