Chapter 1
Rank-Based Nonparametrics

Willem R. van Zwet

1.1 Introduction to Two Papers on Higher Order Asymptotics

1.1.1 Introduction

Peter Bickel has contributed substantially to the study of rank-based nonparametric statistics. Of his many contributions to research in this area I shall discuss his work on second order asymptotics that yielded surprising results and set off more than a decade of research that deepened our understanding of asymptotic statistics. I shall restrict my discussion to two papers, which are Albers et al. (1976) "Asymptotic expansions for the power of distribution free tests in the one-sample problem" and Bickel (1974) "Edgeworth expansions in nonparametric statistics" where the entire area is reviewed.

1.1.2 Asymptotic Expansions for the Power of Distribution Free Tests in the One-Sample Problem

Let $X_1, X_2, \cdots$ be i.i.d. random variables with a common distribution function F_θ for some real-valued parameter θ. For $N = 1, 2, \cdots$, let A_N and B_N be two tests of level $\alpha \in (0,1)$ based on $X_1, X_2, \cdots, X_N$ for the null-hypothesis $H : \theta = 0$ against a contiguous sequence of alternatives $K_{N,c} : \theta = cN^{-1/2}$ for a fixed $c > 0$. Let $\pi_{A,N}(c)$ and $\pi_{B,N}(c)$ denote the powers of A_N and B_N for this testing problem and suppose

W.R. van Zwet (✉)
Leiden University, P.O.Box 9512, 2300 RA, Leiden, The Netherlands
e-mail: vanzwet@math.leidenuniv.nl

J. Fan et al. (eds.), *Selected Works of Peter J. Bickel*, Selected Works in Probability and Statistics 13, DOI 10.1007/978-1-4614-5544-8_1,

that A_N performs at least as well as B_N, i.e. $\pi_{A,N}(c) \geq \pi_{B,N}(c)$. Then we may look for a sample size $k = k_N \geq N$ such that B_k performs as well against alternative $K_{N,c}$ as A_N does for sample size N, i.e. $\pi_{B,k}(c(k/N)^{1/2}) = \pi_{A,N}(c)$. For finite sample size N it is generally impossible to find a usable expression for $k = k_N$, so one resorts to large sample theory and defines the asymptotic relative efficiency (ARE) of sequence $\{B_N\}$ with respect to $\{A_N\}$ as

$$e = e(B,A) = \lim_{N\to\infty} N/k_N.$$

If $\pi_{A,N}(c) \to \pi_A(c)$ and $\pi_{B,N}(c) \to \pi_B(c)$ uniformly for bounded c, and π_A and π_B are continuous, then e is the solution of

$$\pi_B(ce^{-1/2}) = \pi_A(c).$$

Since we assumed that A_N performs at least as well as B_N, we have $e \leq 1$.

If $e < 1$, the ARE provides a useful indication of the quality of the sequence $\{B_N\}$ as compared to $\{A_N\}$. To mimic the performance of A_N by B_k we need $k_N - N = N(1-e)/e + o(N)$ additional observations where the remainder term $o(N)$ is relatively unimportant. If $e = 1$, however, all we know is that the number of additional observations needed is $o(N)$, which may be of any order of magnitude, such as 1 or $N/\log\log N$. Hence in Hodges and Lehmann (1970) the authors considered the case $e = 1$ and proposed to investigate the asymptotic behavior of what they named the deficiency of B with respect to A

$$d_N = k_N - N,$$

rather than k_N/N. Of course this is a much harder problem than determining the ARE. To compute e, all we have to show is that $k_N = N/e + o(N)$, and only the limiting powers π_A and π_B enter into the solution. If $e = 1$, then $k_N = N + o(N)$, but for determining the deficiency, we need to evaluate k_N to the next lower order, which may well be $O(1)$ in which case we have to evaluate k_N with an error of the order $o(1)$. To do this, one will typically need asymptotic expansions for the power functions $\pi_{A,N}$ and $\pi_{B,N}$ with remainder term $o(N^{-1})$. For this we need similar expansions for the distribution functions of the test statistics of the two tests under the hypothesis as well as under the alternative.

In their paper Hodges and Lehmann computed deficiencies for some parametric tests and estimators, but they clearly had a more challenging problem in mind. When Frank Wilcoxon introduced his one- and two-sample rank tests (Wilcoxon 1945) most people thought that replacing the observations by ranks would lead to a considerable loss of power compared to the best parametric procedures. Since then, research had consistently shown that this is not the case. Many rank tests have ARE 1 when compared to the optimal test for a particular parametric problem, so it was not surprising that the first question that Hodges and Lehmann raised for further research was: "What is the deficiency (for contiguous normal shift alternatives) of the normal scores test or of van der Waerden's X-test with respect to the t-test?".

In the paper under discussion this question is generalized to other distributions than the normal and answered for the appropriate one-sample rank test as compared with the optimal parametric test. Let $X_1, X_2, \cdots, X_N$ be i.i.d. with a common distribution function G and density g, and let $Z_1 < Z_2 < \cdots < Z_N$ be the order statistics of the absolute values $|X_1|, |X_2|, \cdots, |X_N|$. If $Z_j = |X_{R(j)}|$, define $V_j = 1$ if $X_{R(j)} > 0$ and $V_j = 0$ otherwise. Let $a = (a_1, a_2, \cdots, a_N)$ be a vector of scores and define

$$T = \sum_{1 \le j \le N} a_j V_j. \tag{1.1}$$

T is the linear rank statistic for testing the hypothesis that g is symmetric about zero. Note that the dependence of G, g and a on N is suppressed in the notation. Conditionally on Z, the random variables $V_1, V_2, \cdots, V_N$ are independent with

$$P_j = P(V_j = 1|Z) = g(Z_j)/\{g(Z_j) + g(-Z_j)\}. \tag{1.2}$$

Under the null hypothesis, $V_1, V_2, \cdots, V_N$ are i.i.d. with $P(V_j = 1) = 1/2$. Hence the obvious strategy for obtaining an expansion for the distribution function of T is to introduce independent random variables $W_1, W_2, \cdots, W_N$ with $p_j = P(W_j = 1) = 1 - P(W_j = 0)$ and obtain an expansion for the distribution function of $\sum_{1 \le j \le N} a_j W_j$. In this expansion we substitute the random vector $P = (P_1, P_2, \cdots, P_N)$ for $p = (p_1, p_2, \cdots, p_N)$. The expected value of the resulting expression will then yield an expansion for the distribution function of T.

This approach is not without problems. Consider i.i.d. random variables $Y_1, Y_2, \cdots, Y_N$ with a common continuous distribution with mean $EY_j = 0$, variance $EY_j^2 = 1$, third and fourth moments $\mu_3 = EY_j^3$ and $\mu_4 = EY_j^4$, and third and fourth cumulants $\kappa_3 = \mu_3$ and $\kappa_4 = \mu_4 - 3\mu_2^2$. Let $S_N = N^{-1/2} \sum_{1 \le j \le N} Y_j$ denote the normalized sum of these variables. In Edgeworth (1905) the author provided a formal series expansion of the distribution function $F_N(x) = P(S_N \le x)$ in powers of $N^{-1/2}$. Up to and including the terms of order 1, $N^{-1/2}$ and N^{-1}, Edgeworth's expansion of $F_N(x)$ reads

$$F_N^*(x) = \Phi(x) - \phi(x) \cdot [(\kappa_3/6)(x^2 - 1)N^{-1/2} + \{(\kappa_4/24)(x^3 - 3x) + (\kappa_3^2/72)(x^5 - 10x^3 + 15x)\}N^{-1}]. \tag{1.3}$$

We shall call this the three-term Edgeworth expansion. Though it was a purely formal series expansion, the Edgeworth expansion caught on and became a popular tool to approximate the distribution function of any sequence of continuous random variables U_N with expected value 0 and variance 1 that was asymptotically standard normal. As $\lambda_{3,N} = \kappa_3 N^{-1/2}$ and $\lambda_{4,N} = \kappa_4 N^{-1}$ are the third and fourth cumulants of the random variable S_N under discussion, one merely replaced these quantities by the cumulants of U_N in (1.3). Incidentally, I recently learned from Professor Ibragimov that the Edgeworth expansion was first proposed in Chebyshev (1890), which predates Edgeworth's paper by 15 years. Apparently this is one more example of Stigler's law of eponymy, which states that no scientific discovery – including Stigler's law – is named after its original discoverer (Stigler 1980).

A proof of the validity of the Edgeworth expansion for normalized sums S_N was given by Cramér (cf. 1937; Feller 1966). He showed that for the three-term Edgeworth expansion (1.3), the error $F_N^*(x) - F_N(x) = o(N^{-1})$ uniformly in x, provided that $\mu_4 < \infty$ and the characteristic function $\psi(t) = E\exp\{itY_j\}$ satisfies Cramér's condition

$$\limsup_{|t|\to\infty} |\psi(t)| < 1. \tag{1.4}$$

Assumption (1.4) can not be satisfied if Y_1 is a discrete random variable as then its characteristic function is almost periodic and the lim sup equals 1. In the case we are discussing, the summands a_jW_j of the statistic $\sum_{1\le j\le N} a_jW_j$ are independent discrete variables taking only two values 0 and a_j. However, the summands are not identically distributed unless the a_j as well as the p_j are equal. Hence the only case where the summands are i.i.d. is that of the sign test under the null-hypothesis, where $a_j = 1$ for all j, and the values 0 and 1 are assumed with probability 1/2. In that case the statistic $\sum_{1\le j\le N} a_jW_j$ has a binomial distribution with point probabilities of the order $N^{-1/2}$ and it is obviously not possible to approximate a function F_N with jumps of order $N^{-1/2}$ by a continuous function F_N^* with error $o(N^{-1})$.

In all other cases the summands a_jW_j of $\sum_{1\le j\le N} a_jW_j$ are independent but not identically distributed. Cramér has also studied the validity of the Edgeworth expansion for the case that the Y_j are independent by not identically distributed. Assume again that $EY_j = 0$ and define S_N as the normalized sum $S_N = \sigma^{-1}\sum_{1\le j\le N} Y_j$ with $\sigma^2 = \sum_{1\le j\le N} EY_j^2$. As before $F_N(x) = P(S_N \le x)$ and in the three-term Edgeworth expansion $F_N^*(x)$ we replace $\kappa_3 N^{-1/2}$ and $\kappa_4 N^{-1}$ by the third and fourth cumulants of S_N. Cramér's conditions to ensure that $F_N^*(x) - F_N(x) = o(N^{-1})$ uniformly in x, are uniform versions of the earlier ones for the i.i.d. case: $EY_j^2 \ge c > 0, EY_j^4 \le C < \infty$ for $j = 1, 2, \cdots, N$, and for every $\delta > 0$ there exists $q_\delta < 1$ such that the characteristic functions $\psi_j(t) = E\exp\{itY_j\}$ satisfy

$$\sup_{|t|\ge\delta} |\psi_j(t)| < q_\delta \quad \text{for all } j. \tag{1.5}$$

As the a_jW_j are lattice variables (1.5) does not hold for even a single j and the plan of attack of this problem is beginning to look somewhat dubious. However, Feller points out, condition (1.5) is "extravagantly luxurious" for validating the three-term Edgeworth expansion and can obviously be replaced by $\sup_{|t|\ge\delta} |\Pi_{1\le j\le N}\psi_j(t)| = o(N^{-1})$ (cf. Feller 1966, Theorem XVI.7.2 and Problem XVI.8.12). This, in turn, is slightly too optimistic but it is true that the condition

$$\sup_{\delta\le|t|\le N} |\Pi_{1\le j\le N}\psi_j(t)| = o((N\log N)^{-1}) \tag{1.6}$$

is sufficient and the presence of $\log N$ is not going to make any difference. Hence (1.6) has to be proved for the case where $Y_j = a_j(W_j - p_j)$ and $S_N = \sum_{1\le j\le N} a_j(W_j - p_j)/\tau(p)$ with $\tau(p)^2 = \sum_{1\le j\le N} p_j(1-p_j)a_j^2$ and $\rho(t) = \prod_{1\le j\le N}\psi_j(t)$ is the characteristic function of S_N.

This problem is solved in Lemma 2.2 of the paper. The moment assumptions (2.15) of this lemma simply state that $N^{-1}\tau(p)^2 \ge c > 0$ and $N^{-1}\sum_{1\le j\le N} a_j^4 \le C < \infty$, and assumption (2.16) ensures the desired behavior of $|\prod_{1\le j\le N}\psi_j(t)|$ by requiring that there exist $\delta > 0$ and $0 < \varepsilon < 1/2$ such that

$$\lambda\{x : \exists j : |x - a_j| < \zeta,\ \varepsilon \le p_j \le 1-\varepsilon\} \ge \delta N\zeta \quad \text{for some } \zeta \ge N^{-3/2}\log N, \tag{1.7}$$

where λ is Lebesgue measure. This assumption ensures that the set of the scores a_j for which p_j is bounded away from 0 and 1, does not cluster too much about too few points. As is shown in the proof of Lemma 2.2 and Theorem 2.1 of the paper, assumptions (2.15) and (2.16) imply

$$\sup_{\delta\le|t|\le N} |\prod_{1\le j\le N}\psi_j(t)| \le \exp\{-d(\log N)^2\} = N^{-d\log N}, \tag{1.8}$$

which obviously implies (1.6). Hence the three-term Edgeworth expansion for $S_N = \sum_{1\le j\le N} a_j(W_j - p_j)/\tau(p)$ is valid with remainder $o(N^{-1})$, and in fact $O(N^{-5/4})$. This was a very real extension of the existing theory at the time.

To obtain an expansion for the distribution of the rank statistic $T = \sum_{1\le j\le N} a_j V_j$, the next step is to replace the probabilities p_j by the random quantities P_j in (1.2) and take the expectation. Under the null-hypothesis that the density g of the X_j is symmetric this is straightforward because $P_j = 1/2$ for all j. The alternatives discussed in the paper are contiguous location alternatives where $G(x) = F(x-\theta)$ for a specific known F with symmetric density f and $0 \le \theta \le CN^{-1/2}$ for a fixed $C > 0$. Finding an expansion for the distribution of T under these alternatives is highly technical and laborious, but fairly straightforward under the assumptions $N^{-1}\sum_{1\le j\le N} a_j^2 \ge c$, $N^{-1}\sum_{1\le j\le N} a_j^4 \le C$,

$$\lambda\{x : \exists j : |x - a_j| < \zeta\} \ge \delta N\zeta \quad \text{for some } \zeta \ge N^{-3/2}\log N \tag{1.9}$$

and some technical assumptions concerning f and its first four derivatives. Among many other things, the latter ensure that $\varepsilon \le P_j \le 1-\varepsilon$ for a substantial proportion of the P_j. Having obtained expansions for the distribution function of $(2T - \sum a_j)/(\sum a_j^2)^{1/2}$ both under the hypothesis and the alternative, an expansion for the power is now immediate.

It remains to discuss the choice of the scores $a_j = a_{j,N}$. For a comparison between best rank tests and best parametric tests we choose a distribution function F with a symmetric smooth density f and consider the locally most powerful (LMP) rank test based on the scores

$$a_{j,N} = E\Psi(U_{j:N}) \quad \text{where } \Psi(t) = -f'F^{-1}((1+t)/2)/fF^{-1}((1+t)/2) \tag{1.10}$$

and $U_{j:N}$ denotes the j-th order statistic of a sample of size N from the uniform distribution on $(0,1)$. Since $F^{-1}((1+t)/2)$ is the inverse function of the distribution function $(2F-1)$ on $(0,\infty)$, $F^{-1}((1+U_{j:N})/2)$ is distributed as the j-th order statistic V_j of the absolute values $|X_1|,|X_2|,\cdots,|X_N|$ of a sample $X_1,X_2,\cdots,X_N$ from F. Hence $a_j=-Ef'(V_j)/f(V_j)$. As f is symmetric, the function f'/f can only be constant on the positive half-line if f is the density $f(x)=1/2\gamma e^{-\gamma|x|}$ of a Laplace distribution on R^1 for which the sign test is the LMP rank test. We already concluded that this test can not be handled with the tools of this paper, but for every other symmetric four times differentiable f, the important condition (1.9) will hold.

If, instead of the so-called exact scores $a_{j,N}=E\Psi(U_{j:N})$, one uses the approximate scores $a_{j,N}=\Psi(j/(N+1))$, then the power expansions remain unchanged. This is generally not the case for other score generating functions than Ψ.

The most powerful parametric test for the null-hypothesis F against the contiguous shift alternative $F(x-\theta)$ with $\theta=cN^{1/2}$ for fixed $c>0$ will serve as a basis for comparison of the LMP rank test. Its test statistic is simply $\sum_{1\leq j\leq N}\{\log f(X_j-\theta)-\log f(X_j)\}$ which is a sum of i.i.d. random variables and therefore its distribution function under the hypothesis and the alternative admit Edgeworth expansions under the usual assumptions, and so does the power. Explicit expressions are found for the deficiency of the LMP rank test and some examples are:

Normal distribution (Hodges-Lehmann problem). For normal location alternatives the one-sample normal scores test as well as van der Waerden's one-sample rank test with respect to the most powerful parametric test based on the sample mean equals

$$d_N=1/2\log\log N+1/2(u_\alpha^2-1)+1/2\gamma+o(1),$$

where $\Phi(u_\alpha)=1-\alpha$ and $\gamma=0.577216$ is Euler's constant. Note that in the paper there is an error in the constant (cf. Albers et al. 1978). In this case the deficiency does tend to infinity, but no one is likely to notice as $1/2\log\log N=1.568\cdots$ for $N=10^{10}$ (logarithms to base e).

It is also shown that the deficiency of the permutation test based on the sample mean with respect to Student's one-sample test tends to zero as $O(N^{-1/2})$.

Logistic distribution. For logistic location alternatives the deficiency of Wilcoxon's one-sample test with respect to the most powerful test for testing $F(x)=(1+e^{-x})^{-1}$ against $F(x-bN^{-1/2})$ tends to a finite limit and equals

$$d_N=\{18+12u_\alpha^2+(48)^{1/2}bu_\alpha+b^2\}/60+o(1).$$

It came as somewhat of a surprise that Wilcoxon's test statistic admits a three-term Edgeworth expansion, as it is a purely lattice random variable. As we pointed out above, the reason that this is possible is that its conditional distribution is that of a sum of independent but not identically distributed random variables. Intuitively the reason is that the point probabilities of the Wilcoxon statistic are of the order $N^{-3/2}$ which is allowed as the error of the expansion is $o(N^{-1})$.

The final section of the paper discusses deficiencies of estimators of location. It is shown that the deficiency of the Hodges-Lehmann type of location estimator associated with the LMP rank test for location alternatives with respect to the maximum likelihood estimator for location, differs by $O(N^{-1/4})$ from the deficiency of the parent tests.

The paper deals with a technically highly complicated subject and is therefore not easy to read. At the time of appearance it had the dubious distinction of being the second longest paper published in the Annals. With 49 pages it was second only to Larry Brown's 50 pages on the admissibility of invariant estimators (Brown 1966). However, for those interested in expansions and higher order asymptotics it contains a veritable treasure of technical achievements that improve our understanding of asymptotic statistics. I hope this review will facilitate the reading. While I'm about it, let me also recommend reading the companion paper (Bickel and van Zwet 1978) where the same program is carried out for two-sample rank tests. With its 68 pages it was regrettably the longest paper in the Annals at the time it was published, but don't let that deter you! Understanding the technical tricks in this area will come in handy in all sorts of applications.

1.1.3 Edgeworth Expansions in Nonparametric Statistics

This paper is a very readable review of the state of the art at the time in the area of Edgeworth expansions. It discusses the extension of Cramér's work to sums of i.i.d. random vectors, as well as expansions for M-estimators. It also gives a preview of the results of the paper we have just discussed on one-sample rank tests and the paper we just mentioned on two-sample rank tests. There is also a new result of Bickel on U-statistics that may be viewed as the precursor of a move towards a general theory of expansions for functions of independent random variables. As we have already discussed Cramér's work as well as rank statistics, let me restrict the discussion of the present paper to the result on U-statistics.

First of all, recall the classical Berry-Esseen inequality for normalized sums $S_N = N^{-1/2} \cdot \sum_{1\leq j\leq N} X_j$ of i.i.d. random variables $X_1, \cdots, X_N$, with $EX_1 = 0$ and $EX_1^2 = 1$. If $E|X_1|^3 < \infty$, and Φ denotes the standard normal distribution function, then there exists a constant C such that for all N,

$$\sup_x |P(S_N \leq x) - \Phi(x)| \leq CE|X_1|^3 N^{-1/2}. \tag{1.11}$$

In the present paper a bound of Berry-Esseen-type is proved for U-statistics. Let $X_1, X_2, \cdots$ be i.i.d. random variables with a common distribution function F and let ψ be a measurable, real-valued function on R^2 where it is bounded, say $|\psi| \leq M < \infty$, and symmetric, i.e. $\psi(x,y) = \psi(y,x)$. Define

$$\gamma(x) = E(\psi(X_1, X_2)|X_1 = x) =!_{(0,1)} \psi(x,y)\, dF(y)$$

and suppose that $E\psi(X_1,X_2) = E\gamma(X_1) = 0$. Define a normalized U-statistic T_N by

$$T_N = \sigma_N^{-1} \sum_{1\le i<j\le N} \psi(X_i,X_j) \quad \text{with } \sigma_N^2 = E\{\sum_{1\le i<j\le N} \psi(X_i,X_j)\}^2, \tag{1.12}$$

and hence $ET_N = 0$ and $ET_N^2 = 1$. In the paper it is proved that if $E\gamma^2(X_1) > 0$, then there exists a constant C depending on ψ but not on N such that

$$\sup_x |P(T_N \le x) - \Phi(x)| \le CN^{-1/2}. \tag{1.13}$$

When comparing this result with the Berry-Esseen bound for the normalized sum S_N, one gets the feeling that the assumption that ψ is bounded is perhaps a bit too restrictive and that it should be possible to replace it by one or more moment conditions. But it was a good start and improvements were made in quick succession. The boundedness assumption for ψ was dropped and Chan and Wierman (1977) proved the result under the conditions that $E\gamma^2(X_1) > 0$ and $E\{\psi(X_1,X_2)\}^4 < \infty$. Next Callaert and Janssen (1978) showed that $E\gamma^2(X_1) > 0$ and $E|\psi(X_1,X_2)|^3 < \infty$ suffice. Finally Helmers and van Zwet (1982) proved the bound under the assumptions $E\gamma^2(X_1) > 0$, $E|\gamma(X_1)|^3 < \infty$ and $E\psi(X_1,X_2)^2 < \infty$.

Why is this development of interest? The U-statistics discussed so far are a special case of U-statistics of order k which are of the form

$$T = \sum_{\substack{1\le j(1)<j(2)< \\ \cdots<j(k)\le N}} \psi_k(X_{j(1)},X_{j(2)},\cdots,X_{j(k)}), \tag{1.14}$$

where ψ_k is a symmetric function of k variables with $E\psi_k(X_1,X_2,\cdots,X_k) = 0$ and the summation is over all distinct k-tuples chosen from $X_1,X_2,\cdots,X_N$. Clearly the U-statistics discussed above have degree $k = 2$, but extension of the Berry-Esseen inequality to U-statistics of fixed finite degree k is straightforward. In an unpublished technical report (Hoeffding 1961) Wassily Hoeffding showed that any symmetric function $T = t(X_1,\cdots,X_N)$ of N i.i.d. random variables $X_1,\cdots,X_N$ that has $ET = 0$ and finite variance $\sigma^2 = ET^2 - \{ET\}^2 < \infty$ can be written as a sum of U-statistics of orders $k = 1,2,\cdots,N$ in such a way that all terms involved in this decomposition are uncorrelated and have several additional desirable properties. Hence it seems that it might be possible to obtain results for symmetric functions of N i.i.d. random variables through a study of U-statistics. For the Berry-Esseen theorem this was done in van Zwet (1984) where the result was obtained under fairly mild moment conditions that reduce to the best conditions for U-statistics when specialized to this case. A first step for obtaining Edgeworth expansions for symmetric functions of i.i.d. random variables was taken in Bickel et al. (1986) where the case of U-statistics of degree $k = 2$ was treated. More work is needed here.

References

Albers W, Bickel PJ, van Zwet WR (1976) Asymptotic expansions for the power of distribution free tests in the one-sample problem. Ann Stat 4:108–156

Albers W, Bickel PJ, van Zwet WR (1978) Correction to asymptotic expansions for the power of distributionfree tests in the one-sample problem. Ann Stat 6:1170

Bickel PJ (1974) Edgeworth expansions in nonparametric statistics. Ann Stat 2:1–20

Bickel PJ, van Zwet WR (1978) Asymptotic expansions for the power of distribution free tests in the two-sample problem. Ann Stat 6:937–1004

Bickel PJ, Goetze F, van Zwet WR (1986) The Edgeworth expansion for U-statistics of degree two. Ann Stat 14:1463–1484

Brown LD (1966) On the admissibility of invariant estimators of one or more location parameters. Ann Math Stat 37:1087–1136

Callaert H, Janssen P (1978) The Berry-Esseen theorem for U-statistics. Ann Stat 6:417–421

Chan YK, Wierman J (1977) The Berry-Esseen theorem for U-statistics. Ann Probab 5:136–139

Chebyshev PL (1890) Sur deux théorèmes relatifs aux probabilités. Acta Math 14:305–315

Cramér H (1937) Random variables and probability distributions. Cambridge tracts in mathematics, vol 36. The University press, Cambridge

Edgeworth FY (1905) The law of error. Proc Camb Philos Soc 20:36–65

Feller W (1966) An introduction to probability theory and its applications, vol II. Wiley, New York

Helmers R, van Zwet WR (1982) The Berry-Esseen bound for U-statistics. In: Gupta SS, Berger JO (eds) Statistical decision theory and related topics III, vol 1. Academic, New York, pp 497–512

Hodges JL Jr, Lehmann EL (1970) Deficiency. Ann Math Stat 41:783–801

Hoeffding W (1961) The strong law of large numbers for U-statistics. Institute of statistics, University of North Carolina mimeograph series, vol 302

Stigler SM (1980) Stiglers law of eponymy. Trans New York Acad Sci 39:147–158

van Zwet WR (1984) A Berry-Esseen bound for symmetric statistics. Z. Wahrscheinlichkeitstheorie verw Gebiete 66:425–440

Wilcoxon F (1945) Individual comparisons by ranking methods. Biometrics Bulletin 1(6):80–83

The Annals of Statistics
1976, Vol. 4, No. 1, 108–156

ASYMPTOTIC EXPANSIONS FOR THE POWER OF DISTRIBUTION FREE TESTS IN THE ONE-SAMPLE PROBLEM[1]

By W. Albers, P. J. Bickel[2] and W. R. van Zwet[3]

University of Leiden and University of California, Berkeley

Asymptotic expansions are established for the power of distribution free tests in the one-sample problem. These expansions are then used to obtain deficiencies in the sense of Hodges and Lehmann (1970) for distribution free tests with respect to their parametric competitors and for the estimators of location associated with these tests.

1. Introduction. Let $X_1, \cdots, X_N$ be independent and identically distributed random variables with a common absolutely continuous distribution. For $N = 1, 2, \cdots$, consider the problem of testing the hypothesis that this distribution is symmetric about zero against a sequence of alternatives that is contiguous to the hypothesis as $N \to \infty$. The level α of the sequence of tests is fixed in (0, 1). Standard tests for this problem are linear rank tests and linear permutation tests and expressions for the limiting powers of such tests are of course well-known. In this paper we shall be concerned with obtaining asymptotic expansions to order N^{-1} for the powers π_N of these tests, i.e. expressions of the form $\pi_N = c_0 + c_1 N^{-\frac{1}{2}} + c_{2,N} N^{-1} + o(N^{-1})$. Of course this involves establishing similar expansions for the distribution function of the test statistic under the hypothesis as well as under contiguous alternatives. For simplicity we shall eventually limit our discussion to contiguous location alternatives and in this case terms of order $N^{-\frac{1}{2}}$ do not occur in the expansions.

One reason to consider these problems would be to obtain better numerical approximations for the critical value of the test statistic and the power of the test than can be provided by the usual normal approximation. A number of authors have investigated this possibility, usually dealing only with the hypothesis in order to obtain critical values and more often for the two-sample case than for the one-sample tests we are concerned with here. For an account of this work we refer to a review paper of Bickel (1974), which incidentally also contains a preview of the present study. Here we merely note that, with the exception of a recent paper of Rogers (1971), all previous work is based on formal Edgeworth

Received July 1974; revised June 1975.

[1] Report SW 30/74 Mathematisch Centrum, Amsterdam.

[2] Research supported by the National Science Foundation, Grant GP-38485 A1, and by the Office of Naval Research, Contract N0014-69-A-0200-1038.

[3] Research supported by the National Science Foundation, Grants GP-29123 and GP-31091X, and by the Office of Naval Research, Contract N00014-69-A-0200-1036.

AMS 1970 *subject classifications.* Primary 62G10, 62G20; Secondary 60F05.

Key words and phrases. Distribution free tests, linear rank tests, permutation tests, power, contiguous alternatives, Edgeworth expansions, deficiency.

expansions. One of the purposes of the present paper is to give a rigorous proof of the validity of such expansions. Rogers (1971) has given such a proof for the two-sample Wilcoxon test under the hypothesis. In a companion paper (Bickel and van Zwet (1975)) expansions will be derived for the general two-sample linear rank test under the hypothesis as well as under contiguous location alternatives.

Here we shall not dwell on the numerical aspects of the expansions we obtain. Numerical results are contained in the Ph. D. thesis of Albers (1974). We only mention that the expansions for the power seem to behave as might be expected. In those cases where the normal approximation already produces reasonably good results, the expansions perform even better and often much better. On the other hand, in cases where the normal approximation is known to be disastrous—the Wilcoxon test for Cauchy alternatives for instance—the expansion is as bad or even worse.

We shall concentrate on a different aspect of the expansions for the power. Consider two sequences of tests $\{T_N\}$ and $\{T_N'\}$ for the same hypothesis at the same fixed level α. Let $\pi_N(\theta_N)$ and $\pi_N'(\theta_N)$ denote the powers of these tests against the same sequence of contiguous alternatives parametrized by a parameter θ. If T_N is more powerful than T_N' we search for a number $k_N = N + d_N$ such that $\pi_N(\theta_N) = \pi'_{k_N}(\theta_N)$. Here k_N and d_N are treated as continuous variables, the power π_N' being defined for real N by linear interpolation between consecutive integers. The quantity d_N was named the deficiency of $\{T_N'\}$ with respect to T_N by Hodges and Lehmann (1970), who introduced this concept and initiated its study. Of course, in many cases of interest, d_N is analytically intractable and one can only study its asymptotic behavior as N tends to infinity.

Suppose that for $N \to \infty$, the ratio N/k_N tends to a limit e, the asymptotic relative efficiency of $\{T_N'\}$ with respect to $\{T_N\}$. If $0 < e < 1$, we have $d_N \sim (e^{-1} - 1)N$ and further asymptotic information about d_N is not particularly revealing. On the other hand, if $e = 1$, the asymptotic behavior of d_N, which may now be anything from $o(1)$ to $o(N)$, does provide important additional information. Of special interest is the case where d_N tends to a finite limit, the asymptotic deficiency of $\{T_N'\}$ with respect to $\{T_N\}$ (cf. Hodges and Lehmann (1970)).

Of course, an asymptotic evaluation of d_N is a more delicate matter than showing that $e = 1$. What is needed is an expansion for the power of the type we discussed above. With the aid of such expansions we arrive at the following results. Let F be a distribution function with a density f that is symmetric about zero and let b be a positive real number. Consider the problem of testing the hypothesis F against the sequence of alternatives $F(x - bN^{-\frac{1}{2}})$ at level α. Let d_N denote the deficiency of the locally most powerful rank test with respect to the most powerful test for this problem. Under certain regularity conditions on F we establish an expression for d_N with remainder $o(1)$ and show that this expression remains unchanged if the exact scores in the locally most powerful rank test are replaced by the corresponding approximate scores. The asymptotic behavior of d_N is found to be governed by that of

$$I_N = \int_{1/N}^{1-1/N} \left(\frac{d^2}{dt^2} f\left(F^{-1}\left(\frac{1+t}{2}\right)\right)\right)^2 t(1-t)\,dt \tag{1.1}$$

in the sense that $d_N = O(I_N)$ as $N \to \infty$. By taking F to be the normal distribution we find that the deficiency of both Fraser's normal scores test and van der Waerden's test with respect to the $\bar{X}$-test for contiguous normal alternatives tends to ∞ at the rate of $\frac{1}{2} \log \log N$. For logistic alternatives the deficiency of Wilcoxon's signed rank test with respect to the most powerful parametric test tends to a finite limit. Another typical result is that for contiguous normal alternatives the deficiency of the permutation test based on $\sum X_i$ with respect to Student's test tends to zero for $N \to \infty$.

Combining numerical and Monte Carlo methods, Albers (1974) has evaluated the deficiency of the normal scores test with respect to the $\bar{X}$-test for $N = 5 - (1) - 10$, 20 and 50. The results agree reasonably well with the asymptotic expression for d_N.

To every linear rank test with nonnegative and nondecreasing scores, there corresponds an estimator of location due to Hodges and Lehmann (1963). A similar correspondence exists between the locally most powerful parametric test and the maximum likelihood estimator. We shall exploit this correspondence to obtain asymptotic expansions for the distribution functions of these estimators. We shall show that, when suitably defined, the deficiency of the Hodges–Lehmann estimator associated with the locally most powerful rank test with respect to the maximum likelihood estimator is asymptotically equivalent to the deficiency of the parent tests.

In Section 2 we establish an asymptotic expansion for the distribution function of the general linear rank statistic for the one-sample problem under the hypothesis as well as under alternatives. We specialize to contiguous location alternatives in Section 3 and derive an expansion for the power of the linear rank test. In Section 4 we deal with the important case where the scores are exact or approximate scores generated by a smooth function J. Linear permutation tests are discussed in Section 5. The results on deficiencies of distribution free tests are contained in Section 6. Finally, Section 7 is devoted to estimators.

Although the basic ideas underlying this paper are simple, the proofs are a highly technical matter. The most laborious parts are dealt with in two appendices. We have omitted the proofs of Theorem 5.1 and Lemma 6.1 because we felt that their inclusion would entail much repetition without essentially new ideas. Some relevant results have been left out altogether for much the same reasons. We are referring to a treatment of contiguous alternatives other than location alternatives for linear rank tests, to expansions for the power of locally most powerful parametric tests, most powerful permutation tests and randomized rank score tests. These missing parts may all be found in the Ph. D. thesis of Albers (1974).

2. The basic expansion. Let $X_1, \cdots, X_N$ be independent and identically distributed (i.i.d.) random variables (rv's) with common distribution (df) G and density g, and let $0 < Z_1 < Z_2 < \cdots < Z_N$ denote the order statistics of the absolute values of $X_1, \cdots, X_N$. If $|X_{R_j}| = Z_j$, define

$$V_j = 1 \quad \text{if} \quad X_{R_j} > 0 \qquad = 0 \quad \text{otherwise.} \tag{2.1}$$

We introduce a vector of scores $a = (a_1, \cdots, a_N)$ and define the statistic

$$T = \sum_{j=1}^N a_j V_j \,. \tag{2.2}$$

We shall be concerned with obtaining an asymptotic expansion for the distribution of T as $N \to \infty$.

Our notation strongly suggests that we are considering a fixed underlying df G and perhaps also a fixed infinite sequence of scores as $N \to \infty$. However, this is merely a matter of notational convenience and our main concern will in fact be the case where the df depends on N and the scores form a triangular array $a_{j,N}, j = 1, \cdots, N, N = 1, 2, \cdots$. Since we are suppressing the index N throughout our notation we shall formally present our results in terms of error bounds for a fixed, but arbitrary, value of N. However, as we shall point out following the proof of Theorem 2.2, these results are really asymptotic expansions in disguise.

The rv T is of course the general linear rank statistic for testing the hypothesis that g is symmetric about zero. Under this hypothesis, $V_1, \cdots, V_N$ are i.i.d. with $P(V_j = 1) = \frac{1}{2}$. For general G, $V_1, \cdots, V_N$ are not independent. However, one easily verifies that, conditional on $Z = (Z_1, \cdots, Z_N)$, the rv's $V_1, \cdots, V_N$ are independent with

$$P_j = P(V_j = 1 \mid Z) = \frac{g(Z_j)}{g(Z_j) + g(-Z_j)} \,. \tag{2.3}$$

As independence allows us to obtain expansions of Edgeworth type, we shall carry out the following program to arrive at an expansion for the distribution of T. First we obtain an Edgeworth expansion for the distribution of $\sum a_j W_j$, where $W_1, \cdots, W_N$ are independent with $p_j = P(W_j = 1) = 1 - P(W_j = 0)$. Having done this we substitute the random vector $P = (P_1, \cdots, P_N)$ defined in (2.3) for $p = (p_1, \cdots, p_N)$ in this expansion. The expected value of the resulting expression will then give us an expansion for the distribution of T.

In carrying out the first part of this program we shall indicate any dependence on $p = (p_1, \cdots, p_N)$ in our notation. Consider the rv

$$\frac{\sum_{j=1}^N a_j(W_j - p_j)}{\tau(p)} \,, \tag{2.4}$$

where

$$\tau^2(p) = \sum_{j=1}^N p_j(1 - p_j)a_j^2 \tag{2.5}$$

denotes the variance of $\sum a_j W_j$. Obviously (2.4) has expectation 0 and variance 1; its third and fourth cumulants, multiplied by $N^{\frac{1}{2}}$ and N respectively, are

(2.6) $$\kappa_3(p) = -N^{\frac{1}{2}} \frac{\sum p_j(1-p_j)(2p_j-1)a_j^3}{\tau^3(p)},$$

(2.7) $$\kappa_4(p) = N \frac{\sum p_j(1-p_j)(1-6p_j+6p_j^2)a_j^4}{\tau^4(p)}.$$

Let R and ρ denote the df and the characteristic function (ch.f.) of (2.4), thus

(2.8) $$R(x, p) = P\left(\frac{\sum a_j(W_j - p_j)}{\tau(p)} \leqq x\right),$$

(2.9) $$\rho(t, p) = \prod_{j=1}^N \left[p_j \exp\left\{i(1-p_j)\frac{a_j t}{\tau(p)}\right\} + (1-p_j)\exp\left\{-ip_j \frac{a_j t}{\tau(p)}\right\}\right].$$

A formal Edgeworth expansion to order N^{-1} for the df R is given by (Cramér (1946), page 229)

(2.10) $$\tilde{R}(x, p) = \Phi(x) + \phi(x)\{N^{-\frac{1}{2}}Q_1(x, p) + N^{-1}Q_2(x, p)\},$$

where Φ and ϕ denote the df and the density of the standard normal distribution, and

(2.11) $$Q_1(x, p) = -\frac{\kappa_3(p)}{6}(x^2 - 1),$$

$$Q_2(x, p) = -\frac{\kappa_4(p)}{24}(x^3 - 3x) - \frac{\kappa_3^2(p)}{72}(x^5 - 10x^3 + 15x).$$

Let $\tilde{r}(x, p)$ be the derivative of $\tilde{R}(x, p)$ with respect to x. In what follows we shall need an expression for the Fourier transform $\tilde{\rho}(t, p) = \int \exp(itx)\tilde{r}(x, p)\,dx$ of $\tilde{r}$ and one easily verifies that

(2.12) $$\tilde{\rho}(t, p) = e^{-\frac{1}{2}t^2}\left\{1 - \frac{\kappa_3(p)it^3}{6N^{\frac{1}{2}}} + \frac{3\kappa_4(p)t^4 - \kappa_3^2(p)t^6}{72N}\right\}.$$

To justify a formal Edgeworth expansion like (2.10), i.e. to show that $|\tilde{R} - R|$ is indeed $o(N^{-1})$, one usually invokes the following result (Feller (1966), page 512).

LEMMA 2.1. *Let R be a* df *with vanishing expectation and* ch.f. ρ. *Suppose that $R - \tilde{R}$ vanishes at $\pm\infty$ and that $\tilde{R}$ has a derivative $\tilde{r}$ such that $|\tilde{r}| \leqq m$. Finally, suppose that $\tilde{r}$ has a continuously differentiable Fourier transform $\tilde{\rho}$ such that $\tilde{\rho}(0) = 1$ and $\tilde{\rho}'(0) = 0$. Then for all x and $T > 0$,*

(2.13) $$|R(x) - \tilde{R}(x)| \leqq \frac{1}{\pi}\int_{-T}^{T}\left|\frac{\rho(t) - \tilde{\rho}(t)}{t}\right| dt + \frac{24m}{\pi T}.$$

To prove that $|R - \tilde{R}| = o(N^{-1})$, it therefore suffices to show that e.g. for $T = bN^{\frac{3}{2}}$, the integral in (2.13) is $o(N^{-1})$. For the case we are considering this may be done in the standard manner (Feller (1966), Chapter 16) with one important modification at the point where it is shown that $|\rho(t, p)/t|$ is sufficiently small when $|t|$ is of the order $\tau(p)$ or larger. Here one usually makes what Feller calls the extravagantly luxurious assumption that the ch.f.'s of all summands are uniformly bounded away from 1 in absolute value outside every neighborhood of 0. Obviously this condition is not satisfied in our case where the summands $a_j W_j$ are lattice rv's. Weaker sufficient conditions of this type are known, but all seem to imply at the very least that the sum itself is nonlattice. In our case this would exclude for instance both the sign test and the Wilcoxon test.

Although the assumptions mentioned above may be unnecessarily strong, it is clear that one has to exclude cases where the sum (2.4) can only assume relatively few different values. As $\tilde{R}$ is continuous, one can not allow R to have jumps of order N^{-1} or larger. Thus the sign test where jumps of order $N^{-\frac{1}{2}}$ occur, will certainly have to be excluded. However, it is exactly the simple lattice character of this statistic that makes it easily amenable to other methods of expansion (see for instance Albers (1974)). For the Wilcoxon statistic on the other hand, all jumps are $O(N^{-\frac{3}{2}})$ and the assumptions we shall make will not rule out this case.

For $0 < \varepsilon < \frac{1}{2}$ and $\zeta > 0$ consider the set of those a_j for which the corresponding p_j satisfies $\varepsilon \leqq p_j \leqq 1 - \varepsilon$, and let $\gamma(\varepsilon, \zeta, p)$ denote the Lebesgue measure λ of the ζ-neighborhood of this set, thus

$$\gamma(\varepsilon, \zeta, p) = \lambda\{x \mid \exists_j |x - a_j| < \zeta, \varepsilon \leqq p_j \leqq 1 - \varepsilon\} . \tag{2.14}$$

LEMMA 2.2. *Suppose that positive numbers c, C, δ and ε exist such that*

$$\frac{1}{N} \sum_{j=1}^N p_j(1 - p_j)a_j^2 \geqq c , \qquad \frac{1}{N} \sum_{j=1}^N a_j^4 \leqq C , \tag{2.15}$$

$$\gamma(\varepsilon, \zeta, p) \geqq \delta N\zeta \qquad \textit{for some} \quad \zeta \geqq N^{-\frac{3}{2}} \log N . \tag{2.16}$$

Then there exist positive numbers b, B and β depending on N, a and p only through c, C, δ and ε, such that

$$\int_{\log(N+1) \leqq |t| \leqq bN^{\frac{3}{2}}} \left| \frac{\rho(t, p) - \tilde{\rho}(t, p)}{t} \right| dt \leqq BN^{-\beta \log N} .$$

PROOF. Since (2.15) implies that $|\kappa_3(p)| \leqq (Cc^{-2})^{\frac{3}{4}}$ and $|\kappa_4(p)| \leqq Cc^{-2}$,

$$\int_{|t| \geqq \log(N+1)} \left| \frac{\tilde{\rho}(t, p)}{t} \right| dt \leqq B_1 N^{-\beta_1 \log N} ,$$

where B_1, $\beta_1 > 0$ depend only on c and C. Also, for all t,

$$\begin{aligned} |\rho(t, p)| &= \prod_{j=1}^N \left\{1 - 2p_j(1 - p_j)\left(1 - \cos \frac{a_j t}{\tau(p)}\right)\right\}^{\frac{1}{2}} \\ &\leqq \exp\left\{-\sum p_j(1 - p_j)\left[\frac{1}{2}\left(\frac{a_j t}{\tau(p)}\right)^2 - \frac{1}{24}\left(\frac{a_j t}{\tau(p)}\right)^4\right]\right\} \\ &\leqq \exp\left\{-\tfrac{1}{2}t^2 + \frac{Ct^4}{96c^2N}\right\} . \end{aligned} \tag{2.17}$$

For $|t| \leqq 4cC^{-\frac{1}{2}}N^{\frac{1}{2}}$ this is $\leqq \exp(-t^2/3)$. Hence, if $b' = 4cC^{-\frac{1}{2}}$, there exist positive constants B_2 and β_2 such that

$$\int_{\log(N+1)\leqq|t|\leqq b'N^{\frac{1}{2}}} \left|\frac{\rho(t,p)}{t}\right| dt \leqq B_2 N^{-\beta_2 \log N} .$$

As $\gamma(\varepsilon, \zeta, p)/\zeta$ is nonincreasing in ζ, we may assume that $\zeta \leqq 1$ in (2.16). Because of (2.15), for any $M > \zeta$ the number of $|a_j| \geqq M - \zeta$ can be at most $CN(M-\zeta)^{-4}$; choosing $M = (8C/\delta)^{\frac{1}{4}} + 1$ we have $CN(M-\zeta)^{-4} \leqq \delta N/8 \leqq \gamma(\varepsilon, \zeta, p)/8\zeta$. It follows that

$$\lambda\{x \mid \exists_j |a_j| \geqq M - \zeta, |x - a_j| < \zeta\} \leqq 2\zeta \frac{\gamma(\varepsilon, \zeta, p)}{8\zeta} = \frac{\gamma(\varepsilon, \zeta, p)}{4} .$$

Together with (2.16) this implies that for every real t

$$\lambda\left\{z \mid \exists_j |a_j| \leqq M - \zeta, \left|z - \frac{a_j t}{\tau(p)}\right| < \frac{\zeta|t|}{\tau(p)}, \varepsilon \leqq p_j \leqq 1 - \varepsilon\right\} \geqq \frac{3|t|\gamma(\varepsilon, \zeta, p)}{4\tau(p)} .$$

Take $b = \delta[(32M/\pi c^{\frac{1}{2}}) + (16/b')]^{-1}$. Then, for every $|t| \in [b'N^{\frac{1}{2}}, bN^{\frac{3}{2}}]$

$$\lambda\left\{z \mid |z| \leqq \frac{M|t|}{\tau(p)}, |z - k\pi| \leqq \frac{2\zeta bN^{\frac{3}{2}}}{\tau(p)} \text{ for some integer } k\right\}$$
$$\leqq \left(\frac{2M|t|}{\pi\tau(p)} + 1\right)\frac{4\zeta bN^{\frac{3}{2}}}{\tau(p)} \leqq \left(\frac{2M|t|}{\pi(cN)^{\frac{1}{2}}} + \frac{|t|}{b'N^{\frac{1}{2}}}\right)\frac{4bN^{\frac{3}{2}}}{\tau(p)} \frac{\gamma(\varepsilon, \zeta, p)}{\delta N} = \frac{|t|\gamma(\varepsilon, \zeta, p)}{4\tau(p)} ,$$

and hence

$$\lambda\left\{z \mid |z| \leqq \frac{M|t|}{\tau(p)}, \exists_j |a_j| \leqq M - \zeta, \left|z - \frac{a_j t}{\tau(p)}\right| < \frac{\zeta|t|}{\tau(p)}, \varepsilon \leqq p_j \leqq 1 - \varepsilon;\right.$$
$$\left. |z - k\pi| > \frac{2\zeta bN^{\frac{3}{2}}}{\tau(p)} \text{ for every integer } k\right\} \geqq \frac{|t|\gamma(\varepsilon, \zeta, p)}{2\tau(p)} .$$

As $\zeta|t| \leqq \zeta bN^{\frac{3}{2}}$, this implies that the number of indices j for which $|(a_j t/\tau(p)) - k\pi| > \zeta bN^{\frac{3}{2}}/\tau(p)$ for every integer k and $\varepsilon \leqq p_j \leqq 1 - \varepsilon$, is at least equal to

$$\frac{\tau(p)}{2\zeta|t|} \cdot \frac{|t|\gamma(\varepsilon, \zeta, p)}{2\tau(p)} \geqq \frac{\delta N}{4} .$$

For such an index j we have for all $|t| \in [b'N^{\frac{1}{2}}, bN^{\frac{3}{2}}]$,

$$\left\{1 - 2p_j(1 - p_j)\left(1 - \cos\frac{a_j t}{\tau(p)}\right)\right\}^{\frac{1}{2}} \leqq \left\{1 - 2\varepsilon(1 - \varepsilon)\frac{\zeta^2 b^2 N^3}{(\pi\tau(p))^2}\right\}^{\frac{1}{2}}$$
$$\leqq \exp\left\{-\frac{\varepsilon(1 - \varepsilon)\zeta^2 b^2 N^3}{(\pi\tau(p))^2}\right\}$$

and hence, as $4\tau^2(p) \leqq C^{\frac{1}{2}}N$ and $\zeta \geqq N^{-\frac{3}{2}} \log N$,

$$|\rho(t, p)| \leqq \exp\left\{-\frac{\delta\varepsilon(1 - \varepsilon)b^2 N^4 \zeta^2}{4\pi^2\tau^2(p)}\right\} \leqq \exp\left\{-\frac{\delta\varepsilon(1 - \varepsilon)b^2}{\pi^2 C^{\frac{1}{2}}} (\log N)^2\right\} .$$

This implies that for some B_3, $\beta_3 > 0$ depending on c, C, δ and ε,

$$\int_{b'N^{\frac{1}{2}} \leqq |t| \leqq bN^{\frac{3}{2}}} \left| \frac{\rho(t, p)}{t} \right| dt \leqq B_3 N^{-\beta_3 \log N} ,$$

which completes the proof. □

We now justify expansion (2.10).

THEOREM 2.1. *Suppose that positive numbers c, C, δ and ε exist such that* (2.15) *and* (2.16) *are satisfied. Then there exists $A > 0$ depending on N, a and p only through c, C, δ and ε such that*

$$\sup_x |R(x, p) - \tilde{R}(x, p)| \leqq AN^{-\frac{5}{4}} . \tag{2.18}$$

PROOF. For $0 \leqq y \leqq 1$ and $-\pi/2 \leqq z \leqq \pi/2$, $\mathrm{Re}\,[y \exp\{i(1-y)z\} + (1-y)\exp\{-iyz\}] \geqq \frac{1}{2}$, and hence we have the following Taylor expansion (mod. $2\pi i$)

$$\begin{aligned} &\log (ye^{i(1-y)z} + (1-y)e^{-iyz}) \\ &\quad = -\tfrac{1}{2}y(1-y)z^2 + \tfrac{1}{6}y(1-y)(2y-1)iz^3 \\ &\quad\quad + \tfrac{1}{24}y(1-y)(1-6y+6y^2)z^4 + M_1(y, z) , \end{aligned} \tag{2.19}$$

where $|M_1(y, z)| \leqq C_1|z|^5$ for some fixed $C_1 > 0$. If $|a_j t/\tau(p)| \leqq \pi/2$ for all j, we can apply this expansion to the logarithm of every factor in (2.9) which yields

$$\rho(t, p) = \exp\left\{-\tfrac{1}{2}t^2 - \frac{\kappa_3(p)it^3}{6N^{\frac{1}{2}}} + \frac{\kappa_4(p)t^4}{24N} + M_2(t, p)\right\} , \tag{2.20}$$

where $|M_2(t, p)| \leqq C_1|t/\tau(p)|^5 \sum |a_j|^5$.

Condition (2.15) implies that $\max |a_j| \leqq (CN)^{\frac{1}{4}}$ and hence that $|a_j t/\tau(p)| \leqq (Cc^{-2})^{\frac{1}{4}}N^{-\frac{1}{4}}|t|$ for all j. We have already seen that $|\kappa_3(p)| \leqq (Cc^{-2})^{\frac{3}{4}}$ and $|\kappa_4(p)| \leqq Cc^{-2}$; because $\max |a_j| \leqq (CN)^{\frac{1}{4}}$ we also have $\tau^{-5}(p) \sum |a_j|^5 \leqq (Cc^{-2})^{\frac{5}{4}}N^{-\frac{5}{4}}$. It follows from these remarks that there exists $c_1 > 0$, depending only on c and C, such that for $|t| \leqq c_1 N^{\frac{1}{4}}$ expansion (2.20) is valid and also

$$\left| -\frac{\kappa_3(p)it^3}{6N^{\frac{1}{2}}} \right| + \left| \frac{\kappa_4(p)t^4}{24N} \right| + |M_2(t, p)| \leqq \tfrac{1}{4}t^2 .$$

Hence, for $|t| \leqq c_1 N^{\frac{1}{4}}$, Taylor expansion of (2.20) yields

$$\rho(t, p) = \tilde{\rho}(t, p) + M_3(t, p) , \tag{2.21}$$

where $\tilde{\rho}$ is given by (2.12), $|M_3(t, p)| \leqq (N^{-\frac{3}{2}} + N^{-\frac{5}{4}} \sum |a_j|^5)|t|^5 Q(|t|) \exp(-t^2/4)$, and Q is a polynomial with coefficients depending on c and C. This implies the existence of $A_1 > 0$ depending on c and C and such that

$$\int_{|t| \leqq c_1 N^{\frac{1}{4}}} \left| \frac{\rho(t, p) - \tilde{\rho}(t, p)}{t} \right| dt \leqq A_1 N^{-\frac{5}{4}} . \tag{2.22}$$

As c_1 depends only on c and C we may assume without loss of generality that N is so large that $\log (N + 1) \leqq c_1 N^{\frac{1}{4}}$. The theorem is now proved by combining

(2.22) and Lemma 2.2, noting that $\tilde{r}(x, t) = (\partial/\partial x)\tilde{R}(x, t)$ is bounded by a number depending only on c and C and applying Lemma 2.1. □

It will be clear that by requiring that $\sum |a_j|^5 \leqq CN$ in Theorem 2.1 one obtains $|R - \tilde{R}| \leqq AN^{-\frac{3}{2}}$ which is the "natural" order of the remainder.

Before we replace p by the random vector $P = (P_1, \cdots, P_N)$ defined in (2.3) and compute the unconditional distribution of T by taking the expected value, we first have to change the standardization of $\sum a_j W_j$ into one that does not involve p. As before, let $W_1, \cdots, W_N$ be independent with $P(W_j = 1) = 1 - P(W_j = 0) = p_j$, let $\bar{p} = (\bar{p}_1, \cdots, \bar{p}_N)$ be a vector with $0 \leqq \bar{p}_j \leqq 1$ for all j, and consider the df $R^*(x, p, \bar{p})$ of the rv $\tau^{-1}(\bar{p}) \sum a_j(W_j - \bar{p}_j)$, thus

$$R^*(x, p, \bar{p}) = P\left(\frac{\sum a_j(W_j - \bar{p}_j)}{\tau(\bar{p})} \leqq x\right). \tag{2.23}$$

Here $\tau^2(\bar{p}) = \sum \bar{p}_j(1 - \bar{p}_j)a_j^2$ in accordance with (2.5); similarly $\kappa_3(\bar{p})$, $\kappa_4(\bar{p})$, $Q_1(x, \bar{p})$, $Q_2(x, \bar{p})$ and $\tilde{R}(x, \bar{p})$ are defined by replacing p by $\bar{p}$ in (2.6), (2.7), (2.11) and (2.10).

For reasons that will become clear in the sequel we shall also at this stage expand $\tau(\bar{p})/\tau(p)$ in powers of $(\tau^2(p) - \tau^2(\bar{p}))/\tau^2(\bar{p})$; at the same time the numerators of $\kappa_3(p)$ and $\kappa_4(p)$ will be expanded about the point $p = \bar{p}$. Later on, when p_j is replaced by P_j, we shall e.g. take $\bar{p}_j = EP_j$ thus ensuring that $P_j - \bar{p}_j$ is roughly speaking a rv of order $N^{-\frac{1}{2}}$. At the moment, however, we do not make any assumptions about $p - \bar{p}$ and as a result Lemma 2.3 provides only a formal expansion in the sense that we do not claim that the remainder term is at all small.

The expansion for $R^*(x, p, \bar{p})$ that we shall establish is

$$\begin{aligned} \tilde{R}^*(x, p, \bar{p}) = \tilde{R}(x - u, \bar{p}) - \phi(x - u) \Big\{ & \frac{1}{2} \frac{\tau^2(p) - \tau^2(\bar{p})}{\tau^2(\bar{p})}(x - u) \\ & + \frac{1}{6} \frac{\sum (p_j - \bar{p}_j)(1 - 6\bar{p}_j + 6\bar{p}_j^2)a_j^3}{\tau^3(\bar{p})}[(x - u)^2 - 1] \\ & + \frac{1}{8}\left(\frac{\tau^2(p) - \tau^2(\bar{p})}{\tau^2(\bar{p})}\right)^2 [(x - u)^3 - 3(x - u)] \\ & + \frac{\kappa_3(\bar{p})}{12N^{\frac{1}{2}}} \frac{\tau^2(p) - \tau^2(\bar{p})}{\tau^2(\bar{p})}[(x - u)^4 - 6(x - u)^2 + 3]\Big\}, \end{aligned} \tag{2.24}$$

where $\tilde{R}$ is given by (2.10) and

$$u = \frac{\sum (p_j - \bar{p}_j)a_j}{\tau(\bar{p})}. \tag{2.25}$$

LEMMA 2.3. *Let $\bar{p} = (\bar{p}_1, \cdots, \bar{p}_N)$ be a vector of real numbers in* $[0, 1]$ *and suppose that positive numbers c, C, δ and ε exist such that* (2.15) *and* (2.16) *are satisfied and that*

$$\frac{1}{N} \sum_{j=1}^N \bar{p}_j(1 - \bar{p}_j)a_j^2 \geqq c. \tag{2.26}$$

Then there exists $A > 0$ *depending on* N, a, p *and* $\tilde{p}$ *only through* c, C, δ *and* ε *and such that*

$$\text{(2.27)} \qquad \sup_x |R^*(x, p, \tilde{p}) - \tilde{R}^*(x, p, \tilde{p})| \leqq A\{N^{-\frac{5}{4}} + N^{-\frac{3}{2}} \sum (p_j - \tilde{p}_j)^2 |a_j|^3 + N^{-3}|\tau^2(p) - \tau^2(\tilde{p})|^3\} .$$

PROOF. Changing the standardization in Theorem 2.1 we find

$$\text{(2.28)} \qquad \sup_x \left| R^*(x, p, \tilde{p}) - \tilde{R}\left((x - u)\frac{\tau(\tilde{p})}{\tau(p)}, p\right)\right| \leqq AN^{-\frac{5}{4}} .$$

The assumptions of the lemma ensure that $\tau^2(\tilde{p})/\tau^2(p) \geqq cC^{-\frac{1}{2}}$, $\tau^2(p)/\tau^2(\tilde{p}) \geqq cC^{-\frac{1}{2}}$, $|\kappa_3(p)| \leqq (c^{-2}C)^{\frac{3}{4}}$, $|\kappa_3(\tilde{p})| \leqq (c^{-2}C)^{\frac{3}{4}}$, $|\kappa_4(p)| \leqq c^{-2}C$ and $|\kappa_4(\tilde{p})| \leqq c^{-2}C$. It follows that the derivatives of $\tilde{R}((x - u)y, p)$ with respect to y are bounded for $y^2 \geqq cC^{-\frac{1}{2}}$ and all $x - u$, and hence

$$\tilde{R}\left((x - u)\frac{\tau(\tilde{p})}{\tau(p)}, p\right)$$

$$\text{(2.29)} \qquad = \tilde{R}(x - u, p) + \tilde{R}'(x - u, p)\left(\frac{\tau(\tilde{p})}{\tau(p)} - 1\right)(x - u) + \tfrac{1}{2}\tilde{R}''(x - u, p)\left(\frac{\tau(\tilde{p})}{\tau(p)} - 1\right)^2 (x - u)^2 + O\left(\left(\frac{\tau(\tilde{p})}{\tau(p)} - 1\right)^3\right),$$

where $\tilde{R}'(x, p)$ and $\tilde{R}''(x, p)$ denote first and second derivatives of $\tilde{R}(x, p)$ with respect to x. Since $(\tau^2(p) - \tau^2(\tilde{p}))/\tau^2(\tilde{p}) \geqq -1 + cC^{-\frac{1}{2}}$,

$$\text{(2.30)} \qquad \frac{\tau(\tilde{p})}{\tau(p)} = 1 - \frac{1}{2}\,\frac{\tau^2(p) - \tau^2(\tilde{p})}{\tau^2(\tilde{p})} + \frac{3}{8}\left(\frac{\tau^2(p) - \tau^2(\tilde{p})}{\tau^2(\tilde{p})}\right)^2 - \cdots ,$$

where the remainder is of the order of the first term omitted. As $\kappa_3(\tilde{p})$ and $\kappa_4(\tilde{p})$ are bounded, we obtain the following one and two term expansions with remainder for $\kappa_3(p)$ and $\kappa_4(p)$.

$$\kappa_3(p) = \left[\kappa_3(\tilde{p}) - N^{\frac{1}{2}}\,\frac{\sum \{p_j(1-p_j)(2p_j-1) - \tilde{p}_j(1-\tilde{p}_j)(2\tilde{p}_j-1)\}a_j^3}{\tau^3(\tilde{p})}\right]\left(\frac{\tau(\tilde{p})}{\tau(p)}\right)^3$$

$$= \kappa_3(\tilde{p}) + O(N^{-1}|\tau^2(p) - \tau^2(\tilde{p})| + N^{-1} \sum |p_j - \tilde{p}_j|\,|a_j|^3)$$

$$\text{(2.31)} \qquad = \kappa_3(\tilde{p})\left[1 - \frac{3}{2}\,\frac{\tau^2(p) - \tau^2(\tilde{p})}{\tau^2(\tilde{p})}\right] + N^{\frac{1}{2}}\,\frac{\sum (p_j - \tilde{p}_j)(1 - 6\tilde{p}_j + 6\tilde{p}_j^2)a_j^3}{\tau^3(\tilde{p})} + O(N^{-2}(\tau^2(p) - \tau^2(\tilde{p}))^2 + N^{-1} \sum (p_j - \tilde{p}_j)^2|a_j|^3 + N^{-2}|\tau^2(p) - \tau^2(\tilde{p})| \sum |p_j - \tilde{p}_j|\,|a_j|^3) ,$$

$$\text{(2.32)} \qquad \kappa_4(p) = \kappa_4(\tilde{p}) + O(N^{-1}|\tau^2(p) - \tau^2(\tilde{p})| + N^{-1} \sum |p_j - \tilde{p}_j| a_j^4) .$$

In (2.29) we may now replace $\tilde{R}$, $\tilde{R}'$ and $\tilde{R}''$ by explicit expressions and substitute (2.32) and appropriate versions of (2.31) and (2.30). The algebra is straightforward and will be omitted. Combining the result with (2.28) we find that (2.27) holds if a term

$$O(N^{-2} \sum |p_j - \tilde{p}_j|(|a_j|^3 + a_j^4) + N^{-\frac{5}{2}}|\tau^2(p) - \tau^2(\tilde{p})| \sum |p_j - \tilde{p}_j|\,|a_j|^3 + N^{-2}|\tau^2(p) - \tau^2(\tilde{p})| + N^{-\frac{5}{2}}(\tau^2(p) - \tau^2(\tilde{p}))^2)$$

is added to the right-hand side. Here, as well as above, the order symbol is uniform for fixed c and C. The lemma is now proved by noting that

$$N^{-2} \sum |p_j - \tilde{p}_j| |a_j|^3 \leqq N^{-\frac{5}{2}} \sum |a_j|^3 + N^{-\frac{3}{2}} \sum (p_j - \tilde{p}_j)^2 |a_j|^3 ,$$
$$N^{-2} \sum |p_j - \tilde{p}_j| a_j^4 \leqq N^{-\frac{5}{2}} \sum |a_j|^5 + N^{-\frac{3}{2}} \sum (p_j - \tilde{p}_j)^2 |a_j|^3 ,$$
$$N^{-\frac{5}{2}} |\tau^2(p) - \tau^2(\tilde{p})| \sum |p_j - \tilde{p}_j| |a_j|^3 \leqq N^{-\frac{3}{2}} \sum (p_j - \tilde{p}_j)^2 |a_j|^3 + N^{-\frac{7}{2}} (\tau^2(p) - \tau^2(\tilde{p}))^2 \sum |a_j|^3 ,$$
$$N^{-2} |\tau^2(p) - \tau^2(\tilde{p})| + N^{-\frac{5}{2}} (\tau^2(p) - \tau^2(\tilde{p}))^2 \leqq N^{-\frac{3}{2}} + N^{-3} |\tau^2(p) - \tau^2(\tilde{p})|^3 ,$$

and that $\sum |a_j|^3 \leqq C^{\frac{3}{4}} N$ and $\sum |a_j|^5 \leqq (CN)^{\frac{5}{4}}$. □

We shall now replace p by $P = (P_1, \cdots, P_N)$ in $\tilde{R}^*(x, p, \tilde{p})$ and take expectations. Define the vector $\pi = (\pi_1, \cdots, \pi_N)$ by

$$\pi_j = EP_j , \qquad j = 1, \cdots, N ; \tag{2.33}$$

it will play the role of $\tilde{p}$. Furthermore, for $\zeta > 0$ we let $\gamma(\zeta)$ denote the Lebesgue measure λ of the ζ-neighborhood of the set $\{a_1, \cdots, a_N\}$, thus

$$\gamma(\zeta) = \lambda\{x \mid \exists_j |x - a_j| < \zeta\} . \tag{2.34}$$

THEOREM 2.2. *Let* $X_1, \cdots, X_N$ *be* i.i.d. *with common* df G *and density* g, *and let* T, P *and* π *be defined by* (2.2), (2.3) *and* (2.33). *Suppose that positive numbers* c, C, δ, δ' *and* ε *exist with* $\delta' < \min(\delta/2, c^2 C^{-1})$ *and such that*

$$\frac{1}{N} \sum_{j=1}^N a_j^2 \geqq c , \qquad \frac{1}{N} \sum_{j=1}^N a_j^4 \leqq C , \tag{2.35}$$

$$\gamma(\zeta) \geqq \delta N \zeta \qquad \text{for some} \quad \zeta \geqq N^{-\frac{3}{2}} \log N , \tag{2.36}$$

$$P\left(\varepsilon \leqq \frac{g(X_1)}{g(X_1) + g(-X_1)} \leqq 1 - \varepsilon\right) \geqq 1 - \delta' . \tag{2.37}$$

Then there exists $A > 0$ *depending on* N, a *and* G *only through* c, C, δ, δ' *and* ε, *and such that*

$$\sup_x \left| P\left(\frac{T - \sum a_j \pi_j}{\tau(\pi)} \leqq x\right) - E\tilde{R}^*(x, P, \pi)\right| \leqq A\{N^{-\frac{5}{4}} + N^{-\frac{3}{4}}[\sum \{E(P_j - \pi_j)^2\}^{\frac{5}{2}}]^{\frac{2}{5}} + N^{-\frac{3}{2}}[\sum \{E|P_j - \pi_j|^3\}^{\frac{2}{3}}]^{\frac{3}{2}}\} . \tag{2.38}$$

PROOF. We start by showing that a, P and π satisfy the conditions for a, p and $\tilde{p}$ in Lemma 2.3 with large probability.

The number of P_j that lie in $[\varepsilon, 1 - \varepsilon]$ is equal to the number of $g(X_j)/(g(X_j) + g(-X_j))$ in that interval. Applying an exponential bound for binomial probabilities (Okamoto (1958)) we find that for $\delta'' \in (\delta', \min(\delta/2, c^2 C^{-1}))$, (2.37) implies

$$P(\varepsilon \leqq P_j \leqq 1 - \varepsilon \;\text{ for at least }\; (1 - \delta'')N \;\text{ indices }\; j) \geqq 1 - e^{-2N(\delta'' - \delta')^2} .$$

Suppose that $\varepsilon \leqq P_j \leqq 1 - \varepsilon$ for at least $(1 - \delta'')N$ values of j. It then follows from (2.36) that a and P satisfy condition (2.16) if δ is replaced by $\delta - 2\delta'' > 0$.

For $\eta \in (0, 1)$, suppose that $a_j^2 \leqq \eta c$ for exactly k indices j and let $\sum'$ indicate summation over the remaining $N - k$ indices. Because of (2.35)

$$c \leqq \frac{1}{N} \sum a_j^2 \leqq \frac{k}{N} \eta c + \frac{1}{N} \sum' a_j^2 \leqq \eta c + \frac{N-k}{N} \left(\frac{1}{N-k} \sum' a_j^4 \right)^{\frac{1}{2}}$$
$$\leqq \eta c + \left(\frac{N-k}{N} C \right)^{\frac{1}{2}} ,$$

and hence the number of $a_j^2 > \eta c$ is at least $(1 - \eta)^2 c^2 C^{-1} N$. By choosing η sufficiently small we can ensure that $(1 - \eta)^2 c^2 C^{-1} > \delta''$. This implies that $N^{-1}\tau^2(P) \geqq \tilde{c}$, where $\tilde{c} = ((1 - \eta)^2 c^2 C^{-1} - \delta'')\varepsilon(1 - \varepsilon)\eta c > 0$. This in turn ensures that $N^{-1}\tau^2(\pi) \geqq N^{-1}E\tau^2(P) \geqq c^*$, where $c^* = \tilde{c}(1 - \exp\{-2(\delta'' - \delta')^2\}) > 0$.

Thus we have shown that if c, C, δ and ε are replaced by positive numbers c^*, C, $\delta - 2\delta''$ and ε depending only on c, C, δ, δ' and ε, then a and π satisfy (2.26) and the second part of (2.15), whereas a and P satisfy (2.16) and the first part of (2.15) except on a set E with $P(E) \leqq \exp\{-2N(\delta'' - \delta')^2\} = O(N^{-\frac{5}{4}})$. Hence a, P and π satisfy the assumptions of Lemma 2.3 on the complement of E. In dealing with the set E it will suffice to note that $\tilde{R}^*(x, P, \pi)$ is bounded since (2.26) and the second part of (2.15) ensure the boundedness of $\kappa_3(\pi)$, $\kappa_4(\pi)$, $(\tau^2(P) - \tau^2(\pi))/\tau^2(\pi)$ and $\sum |a_j|^3/\tau^3(\pi)$. Of course $R^*(x, P, \pi)$, being a probability, is also bounded.

As

$$P\left(\frac{T - \sum a_j \pi_j}{\tau(\pi)} \leqq x \right) = ER^*(x, P, \pi) ,$$

the left-hand side of (2.38) is bounded above by

$$E \sup_x |R^*(x, P, \pi) - \tilde{R}^*(x, P, \pi)| . \tag{2.39}$$

Applying Lemma 2.3 on the complement of E and using the boundedness of $|R^*(x, P, \pi) - \tilde{R}^*(x, P, \pi)|$ together with $P(E) = O(N^{-\frac{5}{4}})$ we find that (2.39) is

$$O(N^{-\frac{5}{4}} + N^{-\frac{3}{2}} \sum E(P_j - \pi_j)^2 |a_j|^3 + N^{-3} E|\tau^2(P) - \tau^2(\pi)|^3) ,$$

where the order symbol is uniform for fixed c, C, δ, δ' and ε. Now

$$N^{-\frac{3}{2}} \sum E(P_j - \pi_j)^2 |a_j|^3 \leqq N^{-\frac{3}{2}} [\sum \{E(P_j - \pi_j)^2\}^{\frac{5}{2}}]^{\frac{2}{5}} (\sum |a_j|^5)^{\frac{3}{5}} ,$$
$$N^{-3} E|\tau^2(P) - \tau^2(\pi)|^3 \leqq N^{-3} E[\sum |P_j - \pi_j| a_j^2]^3 \leqq N^{-3} [\sum \{E|P_j - \pi_j|^3\}^{\frac{1}{3}} a_j^2]^3$$
$$\leqq N^{-3} [\sum \{E|P_j - \pi_j|^3\}^{\frac{2}{3}}]^{\frac{3}{2}} (\sum a_j^4)^{\frac{3}{2}} ,$$

and since $\sum |a_j|^5 \leqq (CN)^{\frac{5}{4}}$ and $\sum a_j^4 \leqq CN$, this completes the proof. □

We note that the boundedness of $\tilde{R}^*(x, P, \pi)$ on E plays an important role in the above proof. Because $\tau(P)$ may be arbitrarily small on E, this explains why we had to remove $\tau(p)$ from the denominator of the expansion in Lemma 2.3 by means of (2.30).

Although Theorem 2.2 is formally stated as a result for a fixed, but arbitrary value of N, it is of course meaningless for fixed N because we do not investigate the way in which A depends on c, C, δ, δ' and ε. In fact the theorem is a purely asymptotic result. Let us for a moment indicate dependence on N by a superscript. Thus, for $N = 1, 2, \cdots$, consider the distribution of the statistic $T^{(N)}$ based on a vector of scores $a^{(N)} = (a_1^{(N)}, \cdots, a_N^{(N)})$ when the underlying df is $G^{(N)}$. Fix positive values of c, C, δ, δ' and ε with $\delta' < \min(\delta/2, c^2C^{-1})$. The theorem asserts that if for every N, $a^{(N)}$ and $G^{(N)}$ satisfy (2.35)—(2.37) for these fixed c, C, δ, δ' and ε, then the error of the approximation $E\tilde{R}^*(x, P^{(N)}, \pi^{(N)})$ is

$$O(N^{-\frac{5}{4}} + N^{-\frac{3}{4}}[\sum \{E(P_j^{(N)} - \pi_j^{(N)})^2\}^{\frac{5}{2}}]^{\frac{2}{5}} + N^{-\frac{3}{2}}[\sum \{E|P_j^{(N)} - \pi_j^{(N)}|^3\}^{\frac{2}{3}}]^{\frac{3}{2}})$$

as $N \to \infty$. Moreover, the order of the remainder is uniform for all such sequences $a^{(N)}$, $G^{(N)}$, $N = 1, 2, \cdots$.

Assumption (2.36) may need some clarification. It is clear from the proof of Lemma 2.2 that the role of conditions (2.16) and (2.36) in Theorems 2.1 and 2.2 is to ensure that the a_j do not cluster too much around too few points. Assumption (2.36) is certainly satisfied if for some $k \geqq \delta N/2$, indices $j_1, j_2, \cdots, j_k$ exist such that $a_{j_{i+1}} - a_{j_i} \geqq 2N^{-\frac{3}{2}} \log N$ for $i = 1, \cdots, k-1$. Under condition (2.35) this will typically be the case. Consider for instance the important case $a_j = EJ(U_{j:N})$, where $U_{1:N} < U_{2:N} < \cdots < U_{N:N}$ are order statistics from the uniform distribution on (0, 1) and J is a continuously differentiable, nonconstant function on (0, 1) with $\int J^4 < \infty$. Here both (2.35) and (2.36) are satisfied for all N with fixed c, C and δ. The same is true if $a_j = J(j/(N+1))$ provided that J is monotone near 0 and 1.

For a large class of underlying df's G, the right-hand side of (2.38) is uniformly $o(N^{-1})$. Still Theorem 2.2 does not yet provide an explicit expansion to order N^{-1} for the distribution of T since we are still left with the task of computing the expected value of $\tilde{R}^*(x, P, \pi)$. This is of course a trivial matter under the hypothesis that g is symmetric about zero and, more generally, in the case where, for some $\eta > 0$, $g(x)/g(-x) = \eta$ for all $x > 0$. In this case $P_j = \eta(1+\eta)^{-1}$ with probability 1 for all j and an expansion for the distribution of T is already contained in Theorem 2.1. For fixed alternatives in general, however, the computation of $E\tilde{R}^*(x, P, \pi)$ presents a formidable problem that we shall not attempt to solve here. It would seem that what is needed, is an expansion for the distribution of a linear combination of functions of order statistics.

In the remaining part of this paper we shall restrict attention to sequences of alternatives that are contiguous to the hypothesis. Heuristically the situation is now as follows. Since $g(x)/(g(x) + g(-x)) = \frac{1}{2} + O(N^{-\frac{1}{2}})$, $P_j - \frac{1}{2}$ and $\pi_j - \frac{1}{2}$ will be $O(N^{-\frac{1}{2}})$, whereas $P_j - \pi_j$ will be $O(N^{-1})$ instead of $O(N^{-\frac{1}{2}})$ as before. In the first place this allows us to simplify $E\tilde{R}^*(x, P, \pi)$ considerably as a number of terms may now be relegated to the remainder and functions of π_j may be expanded about the point $\pi_j = \frac{1}{2}$. Much more important, however, is the fact that $U^* = \tau^{-1}(\pi) \sum (P_j - \pi_j)a_j$ will now be $O(N^{-\frac{1}{2}})$ and that we may therefore expand $\tilde{R}^*(x, P, \pi)$ in powers of U^*. This means that we shall be dealing with low moments of linear combinations of functions of order statistics rather than with their distributions. We need hardly point out that a heuristic argument like this can be entirely misleading and that the actual order of the remainder in our expansion will of course have to be investigated. The unduly complicated form of the remainder terms in the preceeding theorem is, of course, preparatory to such further expansion.

Define

$$\tilde{K}(x) = \Phi(x) + \phi(x) \left\{ \frac{\sum a_j^2 E(2P_j - 1)^2 - 4\sigma^2(\sum a_j P_j)}{2 \sum a_j^2} x + \frac{\sum a_j^3(2\pi_j - 1)}{3(\sum a_j^2)^{\frac{3}{2}}} (x^2 - 1) + \frac{\sum a_j^4}{12(\sum a_j^2)^2} (x^3 - 3x) \right\}, \tag{2.40}$$

where $\sigma^2(Z)$ denotes the variance of a rv Z. Carrying out the type of computation outlined above we arrive at the following simplified version of Theorem 2.2.

THEOREM 2.3. *Theorem* 2.2 *continues to hold if* (2.38) *is replaced by*

$$\sup_x \left| P\left(\frac{2T - \sum a_j}{(\sum a_j^2)^{\frac{1}{2}}} \leq x\right) - \tilde{K}\left(x - \frac{\sum a_j(2\pi_j - 1)}{(\sum a_j^2)^{\frac{1}{2}}}\right)\right| \leq A\{N^{-\frac{5}{4}} + \sum \{E(2P_j - 1)^4\}^{\frac{5}{4}} + N^{-\frac{3}{4}}[\sum \{E|P_j - \pi_j|^3\}^{\frac{4}{9}}]^{\frac{9}{4}}\}. \tag{2.41}$$

PROOF. The proof of this theorem becomes somewhat shorter if we use a modification of Theorem 2.2 as a starting point rather than Theorem 2.2 itself. We recall that Theorem 2.2 was proved by an application of Lemma 2.3 for $p = \pi$. However, the proof clearly goes through for any other choice of p that satisfies (2.26). Because of (2.35), we may therefore replace π in (2.38) by a vector $\bar{p}$ with $\bar{p}_j = \frac{1}{2}$ for all j. Noting that for this choice of $\bar{p}$, $\kappa_3(\bar{p}) = 0$, $\kappa_4(\bar{p}) = -2N \sum a_j^4/(\sum a_j^2)^2$, $\tau^2(P) - \tau^2(\bar{p}) = -\frac{1}{4} \sum (2P_j - 1)^2 a_j^2$, and adding the last two terms in $\tilde{R}^*(x, P, \bar{p})$ to the remainder, we obtain

$$\begin{aligned} P\left(\frac{2T - \sum a_j}{(\sum a_j^2)^{\frac{1}{2}}} \leq x\right) &= E\Phi(x - \tilde{U}) + E\phi(x - \tilde{U}) \left\{ \frac{\sum a_j^4}{12(\sum a_j^2)^2} [(x - \tilde{U})^3 - 3(x - \tilde{U})] \right. \\ &\quad + \frac{\sum a_j^2(2P_j - 1)^2}{2 \sum a_j^2} (x - \tilde{U}) \\ &\quad \left. + \frac{\sum a_j^3(2P_j - 1)}{3(\sum a_j^2)^{\frac{3}{2}}} [(x - \tilde{U})^2 - 1] \right\} \\ &\quad + O(N^{-\frac{5}{4}} + N^{-\frac{3}{4}}[\sum \{E(2P_j - 1)^2\}^{\frac{5}{2}}]^{\frac{2}{5}} + N^{-\frac{3}{2}}[\sum \{E|2P_j - 1|^3\}^{\frac{2}{3}}]^{\frac{3}{2}} \\ &\quad + N^{-2}E[\sum a_j^2(2P_j - 1)^2]^2 + N^{-\frac{3}{2}} \sum a_j^2 E(2P_j - 1)^2), \end{aligned} \tag{2.42}$$

where $\tilde{U} = \sum a_j(2P_j - 1)/(\sum a_j^2)^{\frac{1}{2}}$. All order symbols in this proof are uniform for fixed c, C, δ, δ' and ε. The remainder in (2.42) may be simplified by noting that

$$\begin{aligned} N^{-\frac{3}{4}}[\sum \{E(2P_j - 1)^2\}^{\frac{5}{2}}]^{\frac{2}{5}} &+ N^{-\frac{3}{2}}[\sum \{E|2P_j - 1|^3\}^{\frac{2}{3}}]^{\frac{3}{2}} \\ &\leq N^{-\frac{5}{4}} + \sum \{E(2P_j - 1)^2\}^{\frac{5}{2}} + N^{-1} \sum E|2P_j - 1|^3 \\ &\leq N^{-\frac{5}{4}} + N^{-\frac{3}{2}} + 2 \sum \{E(2P_j - 1)^4\}^{\frac{5}{4}}, \end{aligned}$$

$$
\begin{aligned}
N^{-\frac{3}{2}}E[\textstyle\sum a_j^2(2P_j - 1)^2]^2 + N^{-\frac{3}{2}} \textstyle\sum a_j^2 E(2P_j - 1)^2
&\leqq 2N^{-\frac{3}{2}}E[\textstyle\sum a_j^2(2P_j - 1)^2]^2 + N^{-\frac{3}{2}} \\
&\leqq 2N^{-\frac{3}{2}} \textstyle\sum a_j^4 \sum E(2P_j - 1)^4 + N^{-\frac{3}{2}} \\
&\leqq 2C \textstyle\sum \{E(2P_j - 1)^4\}^{\frac{5}{4}} + (2C + 1)N^{-\frac{3}{2}} .
\end{aligned}
$$

Define $U = \sum a_j(P_j - \pi_j)/(\sum a_j^2)^{\frac{1}{2}}$, so $x - \tilde{U} = x - \sum a_j(2\pi_j - 1)/(\sum a_j^2)^{\frac{1}{2}} - 2U$. By expanding in powers of U under the expectation sign in (2.42) we find

$$
\begin{aligned}
P&\left(\frac{2T - \sum a_j}{(\sum a_j^2)^{\frac{1}{2}}} \leqq x\right) \\
&= \tilde{K}\left(x - \frac{\sum a_j(2\pi_j - 1)}{(\sum a_j^2)^{\frac{1}{2}}}\right) + O(N^{-\frac{5}{4}} + \textstyle\sum \{E(2P_j - 1)^4\}^{\frac{5}{4}} + E|U|^3 \\
&\qquad + E|U|\{N^{-1} + N^{-1} \textstyle\sum a_j^2(2P_j - 1)^2 + N^{-\frac{3}{2}} \sum |a_j|^3|2P_j - 1|\}) .
\end{aligned}
\tag{2.43}
$$

Now

$$
\begin{aligned}
N^{-\frac{3}{2}} \textstyle\sum |a_j|^3|2P_j - 1| &\leqq N^{-2} \textstyle\sum a_j^4 + N^{-1} \sum a_j^2(2P_j - 1)^2 , \\
N^{-1}E|U| &\leqq N^{-\frac{3}{2}} + E|U|^3 , \\
N^{-1}E|U| \textstyle\sum a_j^2(2P_j - 1)^2 &\leqq N^{-\frac{1}{2}}EU^2 + N^{-\frac{3}{2}}E[\textstyle\sum a_j^2(2P_j - 1)^2]^2 \\
&\leqq N^{-\frac{3}{2}} + E|U|^3 + C \textstyle\sum \{E(2P_j - 1)^4\}^{\frac{5}{4}} + CN^{-\frac{3}{2}} ,
\end{aligned}
$$

where the last inequality is based on a bound obtained earlier in this proof. It follows that the remainder in (2.43) is of the order of the sum of its first three terms. The proof is completed by noting that

$$
\begin{aligned}
E|U|^3 &\leqq (cN)^{-\frac{3}{2}}E[\textstyle\sum |a_j|\,|P_j - \pi_j|]^3 \leqq (cN)^{-\frac{3}{2}}[\textstyle\sum |a_j|\{E|P_j - \pi_j|^3\}^{\frac{1}{3}}]^3 \\
&\leqq (cN)^{-\frac{3}{2}}(\textstyle\sum a_j^4)^{\frac{3}{4}}[\textstyle\sum \{E|P_j - \pi_j|^3\}^{\frac{4}{9}}]^{\frac{9}{4}} . \ \square
\end{aligned}
$$

Theorem 2.3 provides the basic expansion for the distribution of T under contiguous alternatives. In Section 3 we shall be concerned with a further simplification of this expansion and a precise evaluation of the order of the remainder term.

3. Contiguous location alternatives. The analysis in this section will be carried out for contiguous location alternatives rather than for contiguous alternatives in general. The general case can be treated in much the same way as the location case but the conditions as well as the results become more involved. The interested reader is referred to Albers (1974).

Let F be a df with a density f that is positive on R^1, symmetric about zero and four times differentiable with derivatives $f^{(i)}$, $i = 1, \cdots, 4$. Define functions

$$
\psi_i = \frac{f^{(i)}}{f} , \qquad i = 1, \cdots, 4 , \tag{3.1}
$$

and suppose that positive numbers ε and C exist such that for

$$
\begin{aligned}
&m_1 = 6 , \qquad m_2 = 3 , \qquad m_3 = \tfrac{4}{3} , \qquad m_4 = 1 , \\
&\sup \{\textstyle\int_{-\infty}^{\infty} |\psi_i(x + y)|^{m_i} f(x)\, dx : |y| \leqq \varepsilon\} \leqq C , \qquad i = 1, \cdots, 4 .
\end{aligned}
\tag{3.2}
$$

Let $X_1, \cdots, X_N$ be i.i.d. with common df $G(x) = F(x - \theta)$ where

$$0 \leqq \theta \leqq CN^{-\frac{1}{2}} \tag{3.3}$$

for some positive C. Note that (3.2) and (3.3) together imply contiguity. Let $0 < Z_1 < Z_2 < \cdots < Z_N$ denote the order statistics of $|X_1|, \cdots, |X_N|$ and let T be defined by (2.2). Probabilities, expected values and variances under G will be denoted by P_θ, E_θ and σ_θ^2; under F they will be indicated by P_0, E_0 and σ_0^2. Define

$$\begin{aligned} K_\theta(x) = \Phi(x) + \phi(x) \Bigg\{ & \frac{\sum a_j^4}{12(\sum a_j^2)^2}(x^3 - 3x) - \theta \frac{\sum a_j^3 E_0 \psi_1(Z_j)}{3(\sum a_j^2)^{\frac{3}{2}}}(x^2 - 1) \\ & + \frac{\theta^2}{2 \sum a_j^2}[\sum a_j^2 E_0 \psi_1^2(Z_j) - \sigma_0^2(\sum a_j \psi_1(Z_j))]x \\ & + \frac{\theta^3}{6(\sum a_j^2)^{\frac{1}{2}}} \sum a_j E_0[3\psi_1^3(Z_j) - 6\psi_1(Z_j)\psi_2(Z_j) + \psi_3(Z_j)] \Bigg\} , \end{aligned} \tag{3.4}$$

and

$$\eta = -\theta \frac{\sum a_j E_0 \psi_1(Z_j)}{(\sum a_j^2)^{\frac{1}{2}}} . \tag{3.5}$$

We shall show that $K_\theta(x - \eta)$ is an expansion to order N^{-1} for the df of $(2T - \sum a_j)/(\sum a_j^2)^{\frac{1}{2}}$. The expansion will be established in Theorem 3.1 and an evaluation of the order of the remainder will be given in Theorem 3.2.

Let $\pi(\theta)$ denote the power of the one-sided level α test based on T for the hypothesis of symmetry against the alternative $G(x) = F(x - \theta)$. Suppose that for some $\varepsilon > 0$,

$$\varepsilon \leqq \alpha \leqq 1 - \varepsilon . \tag{3.6}$$

We prove that an expansion for $\pi(\theta)$ is given by

$$\tilde{\pi}(\theta) = 1 - K_\theta(u_\alpha - \eta) + \phi(u_\alpha - \eta) \frac{\sum a_j^4}{12(\sum a_j^2)^2}(u_\alpha^3 - 3u_\alpha) , \tag{3.7}$$

where $u_\alpha = \Phi^{-1}(1 - \alpha)$ denotes the upper α-point of the standard normal distribution.

THEOREM 3.1. *Suppose that positive numbers, c, C, δ and ε exist such that* (2.35), (2.36), (3.2) *and* (3.3) *are satisfied. Then there exists $A > 0$ depending on N, a, F and θ only through c, C, δ and ε and such that*

$$\begin{aligned} \sup_x \left| P_\theta\left(\frac{2T - \sum a_j}{(\sum a_j^2)^{\frac{1}{2}}} \leqq x\right) - K_\theta(x - \eta)\right| \\ \leqq A\{N^{-\frac{5}{4}} + N^{-\frac{3}{4}}\theta^3[\sum \{E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^3\}^{\frac{4}{9}}]^{\frac{9}{4}}\} , \end{aligned} \tag{3.8}$$

$$|\eta| \leqq A , \tag{3.9}$$

$$\theta \frac{|\sum a_j^3 E_0 \psi_1(Z_j)|}{(\sum a_j^2)^{\frac{3}{2}}} \leqq AN^{-1} , \qquad \theta^2 \frac{\sum a_j^2 E_0 \psi_1^2(Z_j)}{\sum a_j^2} \leqq AN^{-1} , \tag{3.10}$$

$$\frac{\theta^3}{(\sum a_j^2)^{\frac{1}{2}}} |\sum a_j E_0[3\psi_1^3(Z_j) - 6\psi_1(Z_j)\psi_2(Z_j) + \psi_3(Z_j)]| \leqq AN^{-1} .$$

If, in addition, (3.6) *is satisfied there exists* $A' > 0$ *depending on* N, a, F, θ *and* α *only through* c, C, δ *and* ε *and such that*

$$(3.11)\qquad |\pi(\theta) - \tilde{\pi}(\theta)| \leqq A'\{N^{-\frac{5}{4}} + N^{-\frac{3}{4}}\theta^3[\sum \{E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^3\}^{\frac{4}{9}}]^{\frac{9}{4}}\} .$$

PROOF. We begin by checking assumption (2.37). One easily verifies that

$$\left|\frac{\partial}{\partial\theta}\frac{f(x-\theta) - f(x+\theta)}{f(x-\theta) + f(x+\theta)}\right| \leqq \tfrac{1}{2}|\psi_1(x-\theta)| + \tfrac{1}{2}|\psi_1(x+\theta)| .$$

Hence the symmetry of f and an application of Markov's inequality and Fubini's theorem yield

$$\begin{aligned} P_\theta\left(\varepsilon \leqq \frac{g(X_1)}{g(X_1) + g(-X_1)} \leqq 1 - \varepsilon\right) &= P_\theta\left(\left|\frac{f(X_1-\theta) - f(X_1+\theta)}{f(X_1-\theta) + f(X_1+\theta)}\right| \leqq 1 - 2\varepsilon\right) \\ &\geqq P_\theta(\textstyle\int_0^\theta \{|\psi_1(X_1 - t)| + |\psi_1(X_1 + t)|\}\, dt \leqq 2(1 - 2\varepsilon)) \\ &\geqq 1 - \frac{1}{2(1-2\varepsilon)} E_\theta \textstyle\int_0^\theta \{|\psi_1(X_1 - t)| + |\psi_1(X_1 + t)|\}\, dt \\ &\geqq 1 - \frac{\theta}{1 - 2\varepsilon} \sup_{|t|\leqq\theta} E_\theta|\psi_1(X_1 + t)| . \end{aligned}$$

Take $\varepsilon < \frac{1}{2}$ and choose $\delta' = \frac{1}{2}\min(\delta/2, c^2C^{-1})$. Because of (3.3) there exists $N_0 > 0$ depending only on c, C, δ and ε such that for $N \geqq N_0$, $2\theta \leqq \varepsilon$ and $\theta \leqq (1 - 2\varepsilon)C^{-\frac{1}{3}}\delta'$. Then (3.2) implies that (2.37) is satisfied for $N \geqq N_0$. This is of course sufficient to ensure that the conclusion of Theorem 2.3 holds.

The passage from (2.41) to (3.8) is achieved by Taylor expansion with respect to θ. Since this part of the proof is highly technical and laborious it will not be given in the body of the text. Instead we refer the interested reader to Appendix 1 where the results we shall need are stated in Corollary A1.1. Using parts (A1.27), (A1.31) and (A1.32) of Corollary A1.1 together with the inequality $\sum \{E_\theta(2P_j - 1)^4\}^{\frac{5}{4}} \leqq \sum E_\theta|2P_j - 1|^5$ we see that the left-hand side of (3.8) is bounded by the right-hand side of (3.8) plus a term

$$(3.12)\qquad O(\theta^{\frac{19}{8}}\{E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3\}^{\frac{1}{3}} + N^{-\frac{1}{2}}\theta^5\sigma_0^{\,2}(\sum a_j\psi_1(Z_j))) .$$

Here, and later in this proof all order symbols are uniform for fixed c, C, δ and ε. Now

$$\begin{aligned} \theta^{\frac{19}{8}}\{E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3\}^{\frac{1}{3}} &+ N^{-\frac{1}{2}}\theta^5\sigma_0^{\,2}(\sum a_j\psi_1(Z_j)) \\ &\leqq \theta^{\frac{27}{16}} + \theta^6 E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3 \\ &\qquad + N^{-\frac{1}{2}}\theta^3 + N^{-\frac{1}{2}}\theta^6\sigma_0^{\,3}(\sum a_j\psi_1(Z_j)) \\ &= O(N^{-\frac{5}{4}} + N^{-\frac{3}{2}}\theta^3 E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3) , \\ E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3 &\leqq [\sum |a_j|\{E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^3\}^{\frac{1}{3}}]^3 \\ &\leqq (CN)^{\frac{3}{4}}[\sum \{E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^3\}^{\frac{4}{9}}]^{\frac{9}{4}} , \end{aligned}$$

which proves (3.8). In view of (2.35) and (3.3) it is clear that (3.9) and (3.10) are merely restating parts (A1.28)—(A1.30) of Corollary A1.1.

The one-sided level α test based on T rejects the hypothesis if $(2T - \sum a_j)(\sum a_j^2)^{-\frac{1}{2}} \geqq \xi_\alpha$ with possible randomization if equality occurs. Taking $\theta = 0$ in (3.8) we find that

$$1 - \Phi(\xi_\alpha) - \phi(\xi_\alpha) \frac{\sum a_j^4}{12(\sum a_j^2)^2} (\xi_\alpha^3 - 3\xi_\alpha) = \alpha + O(N^{-\frac{5}{4}}) ,$$

and hence because of (2.35) and (3.6),

$$\xi_\alpha = u_\alpha - \frac{\sum a_j^4}{12(\sum a_j^2)^2} (u_\alpha^3 - 3u_\alpha) + O(N^{-\frac{5}{4}}) . \tag{3.13}$$

The power of this test against the alternative $F(x - \theta)$ is

$$\pi(\theta) = 1 - K_\theta(\xi_\alpha - \eta) + O(N^{-\frac{5}{4}} + N^{-\frac{3}{4}}\theta^3[\sum \{E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^3\}^{\frac{4}{9}}]^{\frac{9}{4}}) . \tag{3.14}$$

In (3.14) we expand $K_\theta(\xi_\alpha - \eta)$ around $u_\alpha - \eta$. Noting that $|\xi_\alpha - u_\alpha| = O(N^{-1})$ and using (2.35) and (3.10) we arrive at the conclusion that the left-hand side of (3.11) is bounded by the right-hand side of (3.11) plus a term

$$O(N^{-2}\theta^2\sigma_0^2(\sum a_j\psi_1(Z_j))) = O(N^{-3} + N^{-\frac{3}{2}}\theta^3 E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3) .$$

As we have already shown earlier in this proof that such a term does not change the order of the remainder in (3.11), the proof of Theorem 3.1 is completed. □

For $i = 1, 2, 3$, define functions Ψ_i on (0, 1) by

$$\Psi_i(t) = \psi_i\left(F^{-1}\left(\frac{1+t}{2}\right)\right) = \frac{f^{(i)}\left(F^{-1}\left(\frac{1+t}{2}\right)\right)}{f\left(F^{-1}\left(\frac{1+t}{2}\right)\right)} . \tag{3.15}$$

THEOREM 3.2. *Suppose that positive numbers C and δ exist such that* (3.3) *is satisfied and that $|\Psi_1'(t)| \leqq C(t(1-t))^{-\frac{4}{3}+\delta}$ for all $0 < t < 1$. Then there exists $A'' > 0$ depending on N, F and θ only through C and δ and such that*

$$N^{-\frac{3}{4}}\theta^3[\sum \{E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^3\}^{\frac{4}{9}}]^{\frac{9}{4}} \leqq A''N^{-\frac{5}{4}} .$$

For the highly technical proof of this result the reader is referred to Appendix 2. Theorem 3.2 follows at once from Corollary A2.1 in this appendix by taking $h = \Psi_1$.

4. Exact and approximate scores. The expansions given in Section 3 can be simplified further if we make certain smoothness assumptions about the scores a_j. Consider a continuous function J on (0, 1) and let $U_{1:N} < U_{2:N} < \cdots < U_{N:N}$ denote order statistics of a sample of size N from the uniform distribution on (0, 1). For $N = 1, 2, \cdots$ we define the exact scores generated by J by

$$a_j = a_{j,N} = EJ(U_{j:N}) , \qquad j = 1, \cdots, N , \tag{4.1}$$

and the approximate scores generated by J by

$$a_j = a_{j,N} = J\left(\frac{j}{N+1}\right), \qquad j = 1, \cdots, N. \tag{4.2}$$

For almost all well-known linear rank tests the scores are of one of these two types. The locally most powerful rank test against location alternatives of type F is based on exact scores generated by the function $-\Psi_1$, where Ψ_1 is defined in (3.15).

So far, we have systematically kept the order of the remainder in our expansions down to $O(N^{-\frac{5}{4}})$. From this point on, however, we shall be content with a remainder that is $o(N^{-1})$, because otherwise we would have to impose rather restrictive conditions. In the previous sections we have also consistently stressed the fact that the remainder depends on a and F only through certain constants occurring in our conditions, thus in effect indicating classes of scores and distributions for which the expansion holds uniformly. As the number of these constants is becoming rather large, we prefer to formulate our results from here on for a fixed score function J and a fixed df F. The reader can easily construct uniformity classes for himself by using the results of Section 3 and tracing the development of Appendix 2.

DEFINITION 4.1. $\mathscr{J}$ is the class of functions J on (0, 1) that are twice continuously differentiable and nonconstant on (0, 1), and satisfy

$$\int_0^1 J^4(t)\, dt < \infty. \tag{4.3}$$

$$\limsup_{t\to 0,1} t(1-t)\left|\frac{J''(t)}{J'(t)}\right| < \tfrac{3}{2}. \tag{4.4}$$

$\mathscr{F}$ is the class of df's F on R^1 with positive densities f that are symmetric about zero, four times differentiable and such that, for $\psi_i = f^{(i)}/f$, $\Psi_i(t) = \psi_i(F^{-1}((1+t)/2))$, $m_1 = 6$, $m_2 = 3$, $m_3 = \frac{4}{3}$, $m_4 = 1$,

$$\limsup_{y\to 0} \int_{-\infty}^{\infty} |\psi_i(x+y)|^{m_i} f(x)\, dx < \infty, \qquad i = 1, \cdots, 4, \tag{4.5}$$

$$\limsup_{t\to 0,1} t(1-t)\left|\frac{\Psi_1''(t)}{\Psi_1'(t)}\right| < \tfrac{3}{2}. \tag{4.6}$$

For $J \in \mathscr{J}$ and $F \in \mathscr{F}$, let

$$\begin{aligned} \tilde{K}_\theta(x) = \Phi(x) + \phi(x)\Bigg\{ & N^{-1}\frac{\int_0^1 J^4(t)\,dt}{12(\int_0^1 J^2(t)\,dt)^2}(x^3-3x) \\ & - N^{-\frac{1}{2}}\theta\frac{\int_0^1 J^3(t)\Psi_1(t)\,dt}{3(\int_0^1 J^2(t)\,dt)^{\frac{3}{2}}}(x^2-1) + \frac{\theta^2}{2\int_0^1 J^2(t)\,dt} \\ & \times \Big[\int_0^1 J^2(t)\Psi_1^2(t)\,dt - \int_0^1\int_0^1 J(s)\Psi_1'(s)J(t)\Psi_1'(t)(s\wedge t - st)\,ds\,dt\Big]x \\ & + \frac{N^{\frac{1}{2}}\theta^3}{6(\int_0^1 J^2(t)\,dt)^{\frac{1}{2}}}\int_0^1 J(t)[3\Psi_1^3(t) - 6\Psi_1(t)\Psi_2(t) + \Psi_3(t)]\,dt\Bigg\}, \end{aligned} \tag{4.7}$$

(4.8) $$K_{\theta,1}(x) = \tilde{K}_\theta(x) + \phi(x)\frac{N^{-\frac{1}{2}}\theta}{2(\int_0^1 J^2(t)\,dt)^{\frac{1}{2}}}\Big\{\frac{\int_0^1 J(t)\Psi_1(t)\,dt}{\int_0^1 J^2(t)\,dt}\sum_{j=1}^N \sigma^2(J(U_{j:N})) - 2\sum_{j=1}^N \operatorname{Cov}(J(U_{j:N}),\Psi_1(U_{j:N}))\Big\},$$

(4.9) $$K_{\theta,2}(x) = \tilde{K}_\theta(x) + \phi(x)\frac{N^{-\frac{1}{2}}\theta}{2(\int_0^1 J^2(t)\,dt)^{\frac{1}{2}}}\Big\{\frac{\int_0^1 J(t)\Psi_1(t)\,dt}{\int_0^1 J^2(t)\,dt}\int_{1/N}^{1-1/N}(J'(t))^2 t(1-t)\,dt - 2\int_{1/N}^{1-1/N} J'(t)\Psi_1'(t)t(1-t)\,dt\Big\},$$

(4.10) $$\tilde{\eta} = -N^{\frac{1}{2}}\theta\frac{\int_0^1 J(t)\Psi_1(t)\,dt}{(\int_0^1 J^2(t)\,dt)^{\frac{1}{2}}},$$

(4.11) $$\pi_i(\theta) = 1 - K_{\theta,i}(u_\alpha - \tilde{\eta}) + \phi(u_\alpha - \tilde{\eta})N^{-1}\frac{\int_0^1 J^4(t)\,dt}{12(\int_0^1 J^2(t)\,dt)^2}(u_\alpha^3 - 3u_\alpha),$$

for $i = 1, 2$. Then, in the notation of Section 3, we have for contiguous location alternatives and exact scores

THEOREM 4.1. *Let $F \in \mathscr{F}$, $J \in \mathscr{J}$, $a_j = EJ(U_{j:N})$ for $j = 1, \cdots, N$, and let $0 \leqq \theta \leqq CN^{-\frac{1}{2}}$, $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$ for positive C and ε. Then, for every fixed J, F, C and ε, there exist positive numbers A, δ_1, δ_2, $\cdots$ such that $\lim_{N\to\infty}\delta_N = 0$ and for every N*

(4.12) $$\sup_x\left|P_\theta\Big(\frac{2T - \sum a_j}{(\sum a_j^2)^{\frac{1}{2}}} \leqq x\Big) - K_{\theta,1}(x - \tilde{\eta})\right| \leqq \delta_N N^{-1},$$

(4.13) $$\sup_x\left|P_\theta\Big(\frac{2T - \sum a_j}{(\sum a_j^2)^{\frac{1}{2}}} \leqq x\Big) - K_{\theta,2}(x - \tilde{\eta})\right| \leqq \delta_N N^{-1} + AN^{-\frac{3}{2}}\int_{1/N}^{1-1/N}|J'(t)|(|J'(t)| + |\Psi_1'(t)|)(t(1-t))^{\frac{1}{2}}\,dt,$$

(4.14) $$|\pi(\theta) - \pi_1(\theta)| \leqq \delta_N N^{-1}$$

(4.15) $$|\pi(\theta) - \pi_2(\theta)| \leqq \delta_N N^{-1} + AN^{-\frac{3}{2}}\int_{1/N}^{1-1/N}|J'(t)|(|J'(t)| + |\Psi_1'(t)|)(t(1-t))^{\frac{1}{2}}\,dt.$$

PROOF. For fixed $J \in \mathscr{J}$, positive constants c, C and δ exist for which (2.35) and (2.36) hold for all N (cf. one of the remarks following the proof of Theorem 2.2). Similarly, for fixed $F \in \mathscr{F}$, (3.2) is satisfied and it follows that the conclusions of Theorem 3.1 hold with A and A' depending only on F, J, C and ε. Also (4.5) ensures that Ψ_1^6 is summable and together with (4.6) and the second part of Corollary A2.1, this implies that the conclusion of Theorem 3.2 holds with A'' depending only on F and C.

To complete the proof we now apply the results collected in Corollary A2.2 to the expansions $K_\theta(x - \eta)$ and $\tilde{\pi}(\theta)$ in Theorem 3.1 and then expand these functions of η around the point $\eta = \tilde{\eta}$, while noting that $\eta - \tilde{\eta} = o(N^{-\frac{1}{2}})$ by (A2.22) and (A2.23). □

In general, the expansions given in Theorem 4.1 will not hold if the exact scores are replaced by approximate scores $a_j = J(j/(N+1))$, because $\eta - \tilde{\eta}$ will then give rise to a different term of order N^{-1}. If $J = -\Psi_1$, however, it is clear from Corollary A2.2 and the proof of Theorem 4.1 that expansions (4.13) and (4.15) are valid for approximate as well as exact scores. Also for $J = -\Psi_1$, these expansions may be simplified because $F \in \mathscr{F}$ implies that by partial integration

$$\int_0^1 \int_0^1 \Psi_1(s)\Psi_1'(s)\Psi_1(t)\Psi_1'(t)(s \wedge t - st)\, ds\, dt = \tfrac{1}{4} \int_0^1 \Psi_1^4(t)\, dt - \tfrac{1}{4}\left(\int_0^1 \Psi_1^2(t)\, dt\right)^2 ,$$
$$\int_0^1 \Psi_1(t)[6\Psi_1(t)\Psi_2(t) - \Psi_3(t)]\, dt = \tfrac{10}{3} \int_0^1 \Psi_1^4(t)\, dt + \int_0^1 \Psi_2^2(t)\, dt .$$

It follows that in this case $\tilde{\eta}$, $K_{\theta,2}(x - \tilde{\eta})$ and $\pi_2(\theta)$ reduce to

$$\eta^* = N^{\frac{1}{2}}\theta\left(\int_0^1 \Psi_1^2(t)\, dt\right)^{\frac{1}{2}} , \tag{4.16}$$

$$\begin{aligned} L_\theta(x) = \Phi(x - \eta^*) &+ \frac{\phi(x - \eta^*)}{72N} \\ &\times \left\{ \frac{\int_0^1 \Psi_1^4(t)\, dt}{(\int_0^1 \Psi_1^2(t)\, dt)^2} [6(x^3 - 3x) + 6\eta^*(x^2 - 1) - 3\eta^{*2}x - 5\eta^{*3}] \right. \\ &+ \frac{12 \int_0^1 \Psi_2^2(t)\, dt}{(\int_0^1 \Psi_1^2(t)\, dt)^2} \eta^{*3} + 9\eta^{*2}(x - \eta^*) \\ &\left. + \frac{36 \int_{1/N}^{1-1/N} (\Psi_1'(t))^2 t(1-t)\, dt}{\int_0^1 \Psi_1^2(t)\, dt} \eta^* \right\} , \end{aligned} \tag{4.17}$$

$$\begin{aligned} \pi^*(\theta) = 1 - \Phi(u_\alpha - \eta^*) &+ \frac{\eta^*\phi(u_\alpha - \eta^*)}{72N} \\ &\times \left\{ \frac{\int_0^1 \Psi_1^4(t)\, dt}{(\int_0^1 \Psi_1^2(t)\, dt)^2} [-6(u_\alpha^2 - 1) + 3\eta^* u_\alpha + 5\eta^{*2}] \right. \\ &- \frac{12 \int_0^1 \Psi_2^2(t)\, dt}{(\int_0^1 \Psi_1^2(t)\, dt)^2} \eta^{*2} - 9\eta^*(u_\alpha - \eta^*) \\ &\left. - \frac{36 \int_{1/N}^{1-1/N} (\Psi_1'(t))^2 t(1-t)\, dt}{\int_0^1 \Psi_1^2(t)\, dt} \right\} . \end{aligned} \tag{4.18}$$

Finally we note that for $F \in \mathscr{F}$, $-\Psi_1$ can not be constant on $(0, 1)$ because the density $f(x) = \frac{1}{2}\lambda e^{-\lambda|x|}$ of the double exponential distribution is not differentiable at zero. It follows that $-\Psi_1 \in \mathscr{J}$ for every $F \in \mathscr{F}$. We have proved

THEOREM 4.2. *Let $F \in \mathscr{F}$ and let either $a_j = -E\Psi_1(U_{j:N})$ for $j = 1, \cdots, N$ or $a_j = -\Psi_1(j/(N+1))$ for $j = 1, \cdots, N$. Suppose that $0 \leqq \theta \leqq CN^{-\frac{1}{2}}$ and $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$ for positive C and ε. Then, for every fixed F, C and ε, there exist positive numbers $A, \delta_1, \delta_2, \cdots$ such that $\lim_{N\to\infty} \delta_N = 0$ and for every N*

$$\sup_x \left| P_\theta\left(\frac{2T - \sum a_j}{(\sum a_j^2)^{\frac{1}{2}}} \leqq x\right) - L_\theta(x) \right| \leqq \delta_N N^{-1} + AN^{-\frac{5}{4}} \int_{1/N}^{1-1/N} (\Psi_1'(t))^2 (t(1-t))^{\frac{1}{2}}\, dt , \tag{4.19}$$

$$|\pi(\theta) - \pi^*(\theta)| \leqq \delta_N N^{-1} + AN^{-\frac{5}{4}} \int_{1/N}^{1-1/N} (\Psi_1'(t))^2 (t(1-t))^{\frac{1}{2}}\, dt . \tag{4.20}$$

At this point it may be useful to make some remarks concerning the assumptions in Theorems 4.1 and 4.2. Conditions (4.4) and (4.6) ensure that J' and Ψ_1' do not oscillate too wildly near 0 and 1. They also limit the growth of these functions near 0 and 1, but in this respect conditions (4.3) and (4.5) for $i = 1$ are typically much stronger. Together with (4.4) and (4.6) they imply that $J'(t) = o((t(1-t))^{-\frac{5}{4}})$ and $\Psi_1'(t) = o((t(1-t))^{-\frac{7}{8}})$ near 0 and 1 (cf. the proof of Corollary A2.1).

For expansions (4.13), (4.15), (4.19) and (4.20) to be meaningful rather than just formally correct, even stronger growth conditions have to be imposed. Consider, for example, expansion (4.20) and suppose, as is typically the case, that Ψ_1' remains bounded near 0. If $\Psi_1'(t) = o((1-t)^{-1})$ near 1, then the right-hand side in (4.20) is $o(N^{-1})$ and the expansion makes sense. However, if $\Psi_1'(t)$ is of exact order $(1-t)^{-1}$, the expansion reduces to

$$\pi(\theta) = 1 - \Phi(u_\alpha - \eta^*) - \frac{\eta^*\phi(u_\alpha - \eta^*)}{2N} \frac{\int_0^{1-1/N} (\Psi_1'(t))^2 t(1-t)\,dt}{\int_0^1 \Psi_1^2(t)\,dt} + O(N^{-1}) .$$

Finally, if $\Psi_1'(t) \sim (1-t)^{-1-\delta}$ for $t \to 1$ and some $0 < \delta < \frac{1}{6}$, then all we have left in (4.20) is $\pi(\theta) = 1 - \Phi(u_\alpha - \eta^*) + O(N^{-1+2\delta})$. Of course, in these cases too, more exact results can be obtained by paying careful attention to the behavior of the extreme order statistics.

We conclude this section with a few applications of Theorems 4.1 and 4.2. The tedious computations will be omitted. First we consider the power $\pi_{W,N}(\theta)$ and $\pi_{W,L}(\theta)$ of Wilcoxon's signed rank test (W) against normal (N) and logistic (L) location alternatives $G(x) = \Phi(x-\theta)$ and $G(x) = (1 + \exp\{-(x-\theta)\})^{-1}$ respectively, where $\theta = O(N^{-\frac{1}{2}})$. We find

$$\begin{aligned}\pi_{W,N}(\theta) = 1 - \Phi(u_\alpha - \tilde\eta) - \frac{\tilde\eta\phi(u_\alpha - \tilde\eta)}{N} \Big\{ &\tfrac{26}{5} - 2^{\frac{1}{2}} - \tfrac{69}{20}u_\alpha^2 \\ &+ \left(\tfrac{169}{20} - \frac{2(3)^{\frac{1}{2}}}{3}\right) u_\alpha\tilde\eta - \left(\tfrac{103}{20} - \frac{2(3)^{\frac{1}{2}}}{3} - \frac{\pi}{9}\right)\tilde\eta^2 \\ &+ \frac{12 \operatorname{arctn} 2^{\frac{1}{2}}}{\pi}(-1 + u_\alpha^2 - 2u_\alpha\tilde\eta + \tilde\eta^2)\Big\} + o(N^{-1}) ,\end{aligned} \tag{4.21}$$

where $\tilde\eta = (3N/\pi)^{\frac{1}{2}}\theta$, and

$$\begin{aligned}\pi_{W,L}(\theta) = 1 - \Phi(u_\alpha - \eta^*) - \frac{\eta^*\phi(u_\alpha - \eta^*)}{20N}\{2 + 3u_\alpha^2 + u_\alpha\eta^* + \eta^{*2}\} \\ + o(N^{-1}) ,\end{aligned} \tag{4.22}$$

where $\eta^* = (N/3)^{\frac{1}{2}}\theta$.

As a second example we consider the one-sample normal scores test which is based on the scores $a_j = E\Phi^{-1}((1 + U_{j:N})/2)$. Its power $\pi_{NS,N}(\theta)$ and $\pi_{NS,L}(\theta)$ against the normal and logistic location alternatives described above satisfies

$$\begin{aligned}\pi_{NS,N}(\theta) = 1 - \Phi(u_\alpha - \eta^*) - \frac{\eta^*\phi(u_\alpha - \eta^*)}{4N} \Big\{ &-1 + u_\alpha^2 \\ &+ 2\int_0^{\Phi^{-1}(1-1/2N)} \frac{(2\Phi(x)-1)(1-\Phi(x))}{\phi(x)}\,dx \Big\} + o(N^{-1}) ,\end{aligned} \tag{4.23}$$

where now $\eta^* = N^{\frac{1}{2}}\theta$, and

$$\begin{aligned}\pi_{NS,L}(\theta) = 1 &- \Phi(u_\alpha - \tilde{\eta}) \\ &- \frac{\tilde{\eta}\phi(u_\alpha - \tilde{\eta})}{12N}\Big\{23 - 12(2)^{\frac{1}{2}} + u_\alpha^2 + (2\pi - 5)u_\alpha\tilde{\eta} \\ &+ (72 \operatorname{arctn} 2^{\frac{1}{2}} - 22\pi + 1)\tilde{\eta}^2 \\ &- 6 \int_0^{\Phi^{-1}(1-1/2N)} \frac{(2\Phi(x) - 1)(1 - \Phi(x))}{\phi(x)}\, dx\Big\} + o(N^{-1}), \end{aligned} \tag{4.24}$$

where now $\tilde{\eta} = (N/\pi)^{\frac{1}{2}}\theta$. We note that Theorem 4.2 ensures that (4.23) will also hold for van der Waerden's one-sample test which is based on the approximate scores $a_j = \Phi^{-1}((N + j + 1)/2(N + 1))$. To evaluate the integral in (4.23) and (4.24) we write

$$\begin{aligned}&\int_0^{\Phi^{-1}(1-1/2N)} \frac{(2\Phi(x) - 1)(1 - \Phi(x))}{\phi(x)}\, dx \\ &\quad = \tfrac{1}{2} \log\log N + \tfrac{1}{2}\log 2 - 2\int_0^\infty \log x\, \phi(x)\, dx \\ &\qquad + \int_0^\infty \frac{(2\Phi(x) - 1)\{x(1 - \Phi(x)) - \phi(x)\}}{x\phi(x)}\, dx + o(1) \\ &\quad = \tfrac{1}{2}\log\log N + \tfrac{1}{2}\log 2 + 0.05832\cdots + o(1), \end{aligned} \tag{4.25}$$

where the final result is obtained by numerical integration.

5. Permutation tests. In this section we consider distribution free tests other than rank tests, viz. permutation tests. We limit our discussion to linear permutation tests that reject the hypothesis of symmetry if

$$\sum_{i=1}^N h(X_i) \geqq \xi_\alpha(Z) \tag{5.1}$$

with possible randomization if equality occurs. Here h is a function on R^1, $Z = (Z_1, \cdots, Z_N)$ denotes the vector of order statistics of $|X_1|, \cdots, |X_N|$ as before and ξ_α is chosen in such a way that under the hypothesis of symmetry

$$P(\sum_{i=1}^N h(X_i) \geqq \xi_\alpha(Z) \mid Z) = \alpha \qquad \text{a.s.} \tag{5.2}$$

with an obvious modification if there is randomization.

Since (5.1) is equivalent to $\sum \{h(X_i) - h(-X_i)\} \geqq 2\xi_\alpha(Z) - \sum \{h(Z_j) + h(-Z_j)\}$, we assume without loss of generality that h is antisymmetric about the origin, i.e.

$$h(x) = -h(-x) \qquad \text{for all } x. \tag{5.3}$$

But then, under G and conditional on Z, $\sum h(X_i)$ is distributed as $2\sum a_j(V_j - \frac{1}{2})$ with V_j as in (2.3) and $a_j = h(Z_j)$. This means that we can obtain an expansion for this conditional distribution of $\sum h(X_i)$ if we can apply Theorem 2.1.

Under the hypothesis of symmetry, $P_j = \frac{1}{2}$ in (2.3) for all j. Hence in this case Theorem 2.1 yields an expansion for the conditional df of $\sum h(X_i)/(\sum h^2(Z_j))^{\frac{1}{2}}$ that holds uniformly on the set of all values of Z for which the $a_j = h(Z_j)$ satisfy

(2.35) and (2.36) for fixed c, C and δ. If α satisfies (3.6), this immediately leads to an expansion for $\xi_\alpha(Z)$. We find (cf. (3.13))

$$\frac{\xi_\alpha(Z)}{(\sum h^2(Z_j))^{\frac{1}{2}}} = u_\alpha - \frac{\sum h^4(Z_j)}{12(\sum h^2(Z_j))^2}(u_\alpha^3 - 3u_\alpha) + O(N^{-\frac{5}{4}}) \tag{5.4}$$

uniformly on the set E_0^c where, for fixed positive c, C and δ, $\sum h^2(Z_j) \geqq cN$, $\sum h^4(Z_j) \leqq CN$ and $\lambda\{x \mid \exists_j |x - h(Z_j)| < \zeta\} \geqq \delta N \zeta$ for some $\zeta \geqq N^{-\frac{3}{2}} \log N$.

Next we consider the contiguous location alternatives $G(x) = F(x - \theta)$ of Section 3. Under these alternatives, Theorem 2.1 yields an expansion for the conditional df of $\frac{1}{2}\{\sum h(X_i) - \sum (2P_j - 1)h(Z_j)\}/\{\sum P_j(1 - P_j)h^2(Z_j)\}^{\frac{1}{2}}$ uniformly on the set E_θ^c where, for fixed positive c, C and δ, $\sum P_j(1 - P_j)h^2(Z_j) \geqq cN$, $\sum h^4(Z_j) \leqq CN$ and $\lambda\{x \mid \exists_j |x - h(Z_j)| < \zeta, \varepsilon \leqq P_j \leqq 1 - \varepsilon\} \geqq \delta N \zeta$ for some $\zeta \geqq N^{-\frac{3}{2}} \log N$.

Since $E_0 \subset E_\theta$ it suffices to show that $P_\theta(E_\theta) = O(N^{-\frac{5}{4}})$ in order to obtain an expansion to $O(N^{-\frac{5}{4}})$ for the conditional power given Z of the permutation test. The unconditional power is then obtained by taking the expectation. This is done in very much the same way as in Sections 2 and 3 for linear rank tests, the only difference being that now not only the P_j but also the a_j depend on Z.

This program is carried out in Albers (1974) for the special case of the locally most powerful permutation test where $h = -\psi_1 = -f'/f$. In Theorem 5.1 we reproduce a version of this result without further proof. Of course a similar result may be obtained for the general linear permutation test (5.1) with $h \neq -\psi_1$.

Suppose that F is a df with a density f that is positive, symmetric about zero and five times differentiable. Define ψ_i and Ψ_i by (3.1) and (3.15) and take $h = -\psi_1$. Let $\pi_P(\theta)$ be the power of the permutation test (5.1) against the alternative $F(x - \theta)$ and define

$$\begin{aligned} \pi_P^*(\theta) = 1 &- \Phi(u_\alpha - \eta^*) \\ &+ \frac{\eta^*\phi(u_\alpha - \eta^*)}{72N}\left\{\frac{\int_0^1 \Psi_1^4(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2}[-6u_\alpha^2 - 3 + 3u_\alpha\eta^* + 5\eta^{*2}] \right. \\ &\left. - \frac{12\int_0^1 \Psi_2^2(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2}\eta^{*2} + 9(1 - u_\alpha\eta^* + \eta^{*2})\right\}, \end{aligned} \tag{5.5}$$

where η^* is given by (4.16).

THEOREM 5.1. *Let F satisfy (4.5) for $i = 1, \cdots, 5$ and $m_1 = 10$, $m_2 = \frac{5}{2}$, $m_3 = \frac{5}{3}$, $m_4 = \frac{5}{4}$, $m_5 = 1$ and suppose that positive numbers C and ε exist such that $0 \leqq \theta \leqq CN^{-\frac{1}{2}}$ and $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$. Take $h = -\psi_1$. Then there exists $A > 0$ depending on N, F, θ and α only through F, C and ε and such that*

$$|\pi_P(\theta) - \pi_P^*(\theta)| \leqq AN^{-\frac{5}{4}}.$$

For $F = \Phi$, we have $-\psi_1(x) = x$ and Theorem 5.1 provides an expansion for the power of the permutation test based on $\sum X_i$ against normal shift alternatives

$\Phi(x - \theta)$ with $0 \leqq \theta \leqq CN^{-\frac{1}{2}}$ and $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$. We find that this power equals

$$(5.6) \qquad 1 - \Phi(u_\alpha - N^{\frac{1}{2}}\theta) - \frac{\theta u_\alpha^2 \phi(u_\alpha - N^{\frac{1}{2}}\theta)}{4N^{\frac{1}{2}}} + O(N^{-\frac{5}{4}}) .$$

But (5.6) is also the power of Student's one-sided one-sample test for Φ against $\Phi(x - \theta)$ (cf. Hodges and Lehmann (1970)). It follows that for testing the hypothesis Φ against contiguous normal shift alternatives for fixed $0 < \alpha < 1$, the powers of the permutation test based on $\sum X_i$ and of Student's test differ by only $O(N^{-\frac{5}{4}})$ as $N \to \infty$. In fact, this difference is $O(N^{-\frac{3}{2}})$, since Φ satisfies the stronger regularity conditions needed to replace $N^{-\frac{5}{4}}$ by $N^{-\frac{3}{2}}$ in Theorem 5.1.

The remainder of this section will be devoted to a further investigation of this rather striking phenomenon. Roughly speaking, we shall show that for testing any given symmetric distribution against near alternatives, the permutation test (5.1) is almost equivalent to Student's test applied to $h(X_1), \cdots, h(X_N)$ with the correct level of significance for the given null-distribution. Our proof differs from the one outlined above in that we do not use power expansions to establish the near equivalence of the two tests. Instead, we show that the critical regions of the tests are almost identical. This more direct approach has the additional advantage of providing a simple explanation of our result.

Let F be the df of a distribution that is symmetric about zero and consider the problem of testing the hypothesis that $X_1, \cdots, X_N$ have df F against the alternative that they have another df G. For this testing problem and an arbitrary h satisfying (5.3) we compare the permutation test (5.1) with Student's test applied to $h(X_1), \cdots, h(X_N)$ that rejects the hypothesis if

$$(5.7) \qquad \tilde{T} = \frac{\sum h(X_i)}{[\sum h^2(X_i) - N^{-1}(\sum h(X_i))^2]^{\frac{1}{2}}} (1 - N^{-1})^{\frac{1}{2}} \geqq t_\alpha$$

with possible randomization if equality occurs. Here t_α depends on α, h, F and N and is chosen in such a way that the test (5.7) has level α.

THEOREM 5.2. *Suppose there exist positive numbers* $c, C, \varepsilon, \eta, \delta_1, \delta_2, \cdots$ *with* $\lim_{N\to\infty} \delta_N = 0$ *and* $m > 8$, *such that* hF^{-1} *and* hG^{-1} *are monotone and differentiable on intervals* I_F *and* I_G *of length at least* η *where*

$$(5.8) \qquad \left|\frac{d}{dt} h(F^{-1}(t))\right| \geqq c , \qquad \left|\frac{d}{dt} h(G^{-1}(t))\right| \geqq c ,$$

and such that $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$, *and*

$$(5.9) \qquad \int_{-\infty}^{\infty} |h(x)|^m \, dF(x) \leqq C , \qquad \int_{-\infty}^{\infty} |h(x)|^m \, dG(x) \leqq C ,$$

$$(5.10) \qquad \left|\int_{-\infty}^{\infty} h^{2k}(x)\, dF(x) - \int_{-\infty}^{\infty} h^{2k}(x)\, dG(x)\right| \leqq \delta_N \qquad \textit{for } k = 1, 2 .$$

Then there exist $A > 0$ *depending on* N, F, G, h *and* α *only through* c, C, η *and* ε, *and* $\beta > 0$ *depending only on* m, *such that the powers of the tests* (5.1) *and* (5.7) *for* F *against* G *differ by at most* $A(N^{-\beta} + \delta_N)N^{-1}$.

PROOF. We denote probabilities and expected values under $G(F)$ by $P_G(P_F)$ and $E_G(E_F)$. By (5.9) and (5.8) we have

$$\sigma_G^2(h(X_1)) \leqq E_G h^2(X_1) \leqq [E_G h^4(X_1)]^{\frac{1}{2}} \leqq C^{2/m}, \tag{5.11}$$

$$\sigma_G^2(h(X_1)) \geqq 2 \int_0^{\eta/2} (ct)^2\, dt = \frac{c^2\eta^3}{12}, \tag{5.12}$$

so that these moments are bounded away from 0 and ∞. For positive integer $k \leqq 4$, Markov's inequality, the Marcinkievitz–Zygmund–Chung inequality (Chung (1951)) and (5.9) yield

$$\begin{aligned} P_G(|\textstyle\sum (h^k(X_i) - E_G h^k(X_i))| \geqq \tau N) & \\ \leqq \frac{E_G|\sum (h^k(X_i) - E_G h^k(X_i))|^{m/k}}{(\tau N)^{m/k}} & \\ \leqq B_m(\tau^2 N)^{-m/(2k)} E_G|h^k(X_1) - E_G h^k(X_1)|^{m/k} & \\ \leqq B_m C\left(\frac{2}{\tau}\right)^{m/k} N^{-m/(2k)}, & \end{aligned} \tag{5.13}$$

where B_m depends only on m. Choose

$$\beta = \min\left(\frac{m-8}{2m+8}, \tfrac{1}{4}\right). \tag{5.14}$$

Taking $\tau = N^{-\beta}$ in (5.13) and using (5.3) we find that

$$\frac{1}{N} \sum h^{2k}(Z_j) = \frac{1}{N} \sum h^{2k}(X_i) = E_G h^{2k}(X_1) + O(N^{-\beta}), \qquad k = 1, 2, \tag{5.15}$$

$$\frac{1}{N} \sum h^2(X_i) - \left[\frac{1}{N} \sum h(X_i)\right]^2 = \sigma_G^2(h(X_1)) + O(N^{-\beta}), \tag{5.16}$$

uniformly on a set with probability $1 - O(N^{-1-\beta})$ under G.

Assumption (5.3) implies that

$$\lambda\{x \,|\, \exists_j |x - h(Z_j)| < \zeta\} \geqq \tfrac{1}{2}\lambda\{x \,|\, \exists_i |x - h(X_i)| < \zeta\},$$

and under G the right-hand side is distributed like

$$\tfrac{1}{2}\lambda\{x \,|\, \exists_j |x - h(G^{-1}(U_{j:N}))| < \zeta\},$$

where $U_{1:N} < \cdots < U_{N:N}$ are order statistics from a uniform distribution on (0, 1). Now for $n \geqq 1$

$$\begin{aligned} P(U_{j+n:N} - U_{j:N} \leqq z) & \\ = \iint_{0<s<t<1,\, t-s\leqq z} \frac{N!}{(j-1)!\,(n-1)!\,(N-j-n)!} s^{j-1}(t-s)^{n-1}(1-t)^{N-j-n}\, ds\, dt & \\ \leqq \frac{(Nz)^{n-1}}{(n-1)!} \iint_{0<s<t<1} \frac{(N-n+1)!}{(j-1)!\,(N-j-n)!} s^{j-1}(1-t)^{N-j-n}\, ds\, dt & \\ = \frac{(Nz)^{n-1}}{(n-1)!}. & \end{aligned}$$

Taking $n = 6$ and $z = 2c^{-1}N^{-\frac{5}{6}} \log N$ we see that

$$P\left(U_{6(k+1):N} - U_{6k:N} \geqq 2c^{-1}N^{-\frac{5}{6}} \log N \text{ for all } 1 \leqq k \leqq \left[\frac{N}{6}\right] - 1\right)$$
$$\geqq 1 - \frac{N}{6}(2c^{-1}N^{-\frac{1}{6}} \log N)^5 = 1 - O(N^{-1-\beta}) .$$

Together with (5.8) this implies that for $\zeta = N^{-\frac{5}{6}} \log N$

$$\lambda\{x \mid \exists_j |x - h(Z_j)| < \zeta\} \geqq \tfrac{1}{2}\eta N\zeta \tag{5.17}$$

with probability $1 - O(N^{-1-\beta})$ under G.

Now (5.11), (5.12), (5.15) and (5.17) ensure that expansion (5.4) holds uniformly except on a set E_0 with $P_G(E_0) = O(N^{-1-\beta})$. Simplifying this expansion by using (5.11), (5.12) and (5.15) once more, we arrive at the conclusion that the power against G of the test (5.1) is given by

$$\pi_P(G) = P_G\left(\frac{\sum h(X_i)}{(\sum h^2(X_i))^{\frac{1}{2}}} \geqq u_\alpha - \frac{E_G h^4(X_1)}{12N(E_G h^2(X_1))^2}(u_\alpha^{\ 3} - 3u_\alpha) + O(N^{-1-\beta})\right) + O(N^{-1-\beta}) . \tag{5.18}$$

Here the first remainder term depends on Z but may now be taken to be uniformly $O(N^{-1-\beta})$.

The inequality $\sum h(X_i)/(\sum h^2(X_i))^{\frac{1}{2}} \geqq a$ is algebraically equivalent with

$$\frac{\sum h(X_i)}{[\sum h^2(X_i) - N^{-1}(\sum h(X_i))^2]^{\frac{1}{2}}} \geqq \frac{a}{(1 - a^2/N)^{\frac{1}{2}}}$$

on the set where $\sum h^2(X_i) - N^{-1}(\sum h(X_i))^2 \neq 0$ and provided that $a^2 < N$. We may apply this to (5.18) in view of the condition $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$, (5.11), (5.12) and (5.16). At the same time we may replace E_G by E_F in (5.18), and by (5.10) this only involves adding $O(\delta_N N^{-1})$ to the first remainder term in (5.18). In this way we obtain

$$\pi_P(G) = P_G\left(\tilde{T} \geqq u_\alpha + \frac{u_\alpha^{\ 3} - u_\alpha}{2N} - \frac{E_F h^4(X_1)}{12N(E_F h^2(X_1))^2}(u_\alpha^{\ 3} - 3u_\alpha) + O\left(\frac{N^{-\beta} + \delta_N}{N}\right)\right) + O(N^{-1-\beta}) , \tag{5.19}$$

where $\tilde{T}$ is the statistic in (5.7).

By (5.11), (5.12) and (5.16) we have for $B \geqq 0$,

$$\sup_t P_G(t \leqq \tilde{T} \leqq t + BN^{-1}(N^{-\beta} + \delta_N))$$
$$\leqq \sup_t P_G\left(t \leqq \frac{N^{-\frac{1}{2}} \sum h(X_i)}{\sigma_G(h(X_1))} \leqq t + 2BN^{-1}(N^{-\beta} + \delta_N)\right) + O(N^{-1-\beta}) . \tag{5.20}$$

Now (5.8) ensures that under G the distribution of $h(X_1)$ has an absolutely continuous part; in fact, this distribution may be written as a mixture $Q = \eta\tilde{Q}_1 + (1 - \eta)\tilde{Q}_2$ where $\tilde{Q}_1$ is an absolutely continuous distribution with density $\tilde{q}_1 \leqq (c\eta)^{-1}$. Moreover, (5.9) and Markov's inequality imply that $\tilde{Q}_1([-C_1, C_1]) \geqq \frac{1}{2}$

where $C_1 = \max(1, (2C/\eta)^{\frac{1}{3}})$. It follows that $Q = (\eta/2)Q_1 + (1 - \eta/2)Q_2$ where $Q_1([-C_1, C_1]) = 1$ and Q_1 is absolutely continuous with density $q_1 \leqq c_1 = 2(c\eta)^{-1}$.

Let ρ_1 be the ch.f. of Q_1. Obviously, for any fixed $t \neq 0$, $|\rho_1(t)| \leqq |\bar{\rho}_1(t)|$ where $\bar{\rho}_1$ is the ch.f. of the distribution with density

$$\bar{q}_1(y) = c_1 \qquad \text{for} \quad y \in \bigcup_{k=0}^{n}\left[-C_1 + \frac{2k\pi}{|t|}, -C_1 + \frac{2k\pi + 2\xi}{|t|}\right]$$
$$= 0 \qquad \text{elsewhere,}$$

with $n = [C_1|t|/\pi]$ and $(n+1)c_1 2\xi/|t| = 1$. An easy calculation yields $|\bar{\rho}_1(t)| = (\sin\xi)/\xi$; for $|t| \geqq \pi/C_1$ we have $\xi \geqq \pi/(4c_1C_1)$. It follows that there exists $b > 0$ depending only on η, c and C, such that the ch.f. of $h(X_1)$ under G satisfies

$$(5.21) \qquad |E_G e^{ith(X_1)}| \leqq 1 - b \qquad \text{for } |t| \geqq \pi .$$

Because of (5.9), (5.12), (5.21) and Lemma 1 in Cramér (1962), page 27, the df of $\sigma_G^{-1}(h(X_1))N^{-\frac{1}{2}} \sum (h(X_i) - E_G h(X_i))$ under G has an Edgeworth expansion; uniformly for all G satisfying (5.8) and (5.9) for fixed c, C and η, the derivative of this expansion is bounded and its remainder term is $O(N^{-3})$. Applying this result and (5.20) to (5.19) we find

$$(5.22) \qquad \pi_P(G) = P_G(\tilde{T} \geqq \tilde{t}_\alpha) + O(N^{-1}(N^{-\beta} + \delta_N))$$

uniformly for fixed c, C, η and ε, where

$$(5.23) \qquad \tilde{t}_\alpha = u_\alpha + \frac{u_\alpha^3 - u_\alpha}{2N} - \frac{E_F h^4(X_1)}{12N(E_F h^2(X_1))^2}(u_\alpha^3 - 3u_\alpha) .$$

Let t_α be as defined in (5.7). Since F satisfies all assumptions imposed on G, (5.22) will hold under F as well as under G. We have $\pi_P(F) = \alpha$ and hence $\tilde{t}_\alpha = t_{\tilde{\alpha}}$ where $|\tilde{\alpha} - \alpha| = O(N^{-1}(N^{-\beta} + \delta_N))$ uniformly for $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$, but of course also uniformly for $\varepsilon/2 \leqq \alpha \leqq 1 - \varepsilon/2$. Because t_α is decreasing in α and $\tilde{t}_\alpha$ has a bounded derivative with respect to α for $\varepsilon/2 \leqq \alpha \leqq 1 - \varepsilon/2$, it follows that

$$(5.24) \qquad t_\alpha = \tilde{t}_\alpha + O(N^{-1}(N^{-\beta} + \delta_N))$$

uniformly for $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$. In view of (5.22) and the preceding part of the proof this implies that

$$(5.25) \qquad \pi_P(G) = P_G(\tilde{T} \geqq t_\alpha) + O(N^{-1}(N^{-\beta} + \delta_N))$$

uniformly for fixed c, C, η and ε. This completes the proof. □

It may be useful to comment briefly on assumption (5.10) in Theorem 5.2. Of course this assumption is satisfied for a sequence of alternatives G_N that tends to F in an appropriate manner. It is easy to see, for instance, that if the sequence G_N^N is contiguous to F^N, (5.9) implies (5.10) with $\delta_N = O(N^{-\frac{1}{2}})$. Similarly, (5.9) will imply (5.10) for some sequence $\delta_N = o(1)$ if h is continuous and G_N converges weakly to F.

6. Deficiencies of distribution free tests. Let F be a fixed df with density f that is positive, symmetric about zero and five times differentiable. Consider the problem of testing, on the basis of $X_1, \cdots, X_N$, the hypothesis $G = F$ against the alternative $G(x) = F(x - \theta)$ at level α. For any particular θ, the maximum power $\pi^+(\theta)$ is attained by the test based on the statistic $\sum \{\log f(X_i - \theta) - \log f(X_i)\}$. This statistic is a sum of i.i.d. random variables and therefore its df admits an Edgeworth expansion under the usual conditions. By expanding the cumulants of the statistic Albers (1974) obtains an expansion for $\pi^+(\theta)$. Define Ψ_i by (3.15) and take

$$\tilde{\pi}^+(\theta) = 1 - \Phi(u_\alpha - \eta^*) + \frac{\eta^*\phi(u_\alpha - \eta^*)}{72N}\left\{\frac{\int_0^1 \Psi_1^4(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2}[3(u_\alpha^2 - 1) - 3\eta^* u_\alpha + 2\eta^{*2}] - \frac{3\int_0^1 \Psi_2^2(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2}\eta^{*2} - 9[(u_\alpha^2 - 1) - \eta^* u_\alpha]\right\}, \tag{6.1}$$

where η^* is given by (4.16). Lemma 6.1 is a version of Albers' result.

LEMMA 6.1. *Let F satisfy (4.5) for $m_i = 5/i$, $i = 1, \cdots, 5$, and suppose that positive numbers C and ε exist such that $0 \leqq \theta \leqq CN^{-\frac{1}{2}}$ and $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$. Then there exists $A > 0$ depending on N, F, θ and α only through F, C and ε and such that*

$$|\pi^+(\theta) - \tilde{\pi}^+(\theta)| \leqq AN^{-\frac{3}{2}}. \tag{6.2}$$

For the same testing problem Theorem 4.2 provides an expansion for the power $\pi(\theta)$ of the locally most powerful rank test. Together, Theorem 4.2 and Lemma 6.1 will enable us to find the deficiency d_N of the locally most powerful rank test with respect to the most powerful parametric test. To ensure that F satisfies the assumptions of both Theorem 4.2 and Lemma 6.1, we require that $F \in \mathscr{F}_1$, where

DEFINITION 6.1. $\mathscr{F}_1$ is the class of df's F on R^1 with positive densities f that are symmetric about zero, five times differentiable and such that (4.5) is satisfied for $i = 1, \cdots, 5$ with $m_1 = 6$, $m_2 = 3$, $m_3 = \frac{5}{3}$, $m_4 = \frac{5}{4}$, $m_5 = 1$, and such that (4.6) holds.

Furthermore, define

$$\tilde{d}_N = \frac{1}{12}\left\{\frac{\int_0^1 \Psi_1^4(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2}[3(u_\alpha^2 - 1) - 2\eta^* u_\alpha - \eta^{*2}] + \frac{3\int_0^1 \Psi_2^2(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2}\eta^{*2} - 3[(u_\alpha^2 - 1) - 2\eta^* u_\alpha + \eta^{*2}] + 12\frac{\int_{1/N}^{1-1/N} (\Psi_1'(t))^2 t(1-t)\,dt}{\int_0^1 \Psi_1^2(t)\,dt}\right\}, \tag{6.3}$$

with η^* as in (4.16).

THEOREM 6.1. *Let d_N be the deficiency of the locally most powerful rank test* *with respect to the most powerful parametric test for testing* $G = F$ *against* $G(x) = F(x - \theta)$ *on the basis of* $X_1, \cdots, X_N$ *and at level* α. *Suppose that* $F \in \mathscr{F}_1$ *and that* $cN^{-\frac{1}{2}} \leqq \theta \leqq CN^{-\frac{1}{2}}$, $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$ *for positive* c, C *and* ε. *Then, for every fixed* F, c, C *and* ε, *there exist positive numbers* $A, \delta_1, \delta_2, \cdots$ *such that* $\lim_{N\to\infty} \delta_N = 0$ *and for every* N

$$|d_N - \tilde{d}_N| \leqq \delta_N + AN^{-\frac{1}{2}} \int_{1/N}^{1-1/N} (\Psi_1'(t))^2(t(1-t))^{\frac{1}{2}}\, dt\,. \tag{6.4}$$

This result continues to hold if the locally most powerful rank test is replaced by the rank test with the corresponding approximate scores $a_j = -\Psi_1(j/(N+1))$.

PROOF. As $\mathscr{F}_1 \subset \mathscr{F}$, the remark following Theorem 4.2 shows that

$$\int_{1/N}^{1-1/N} (\Psi_1'(t))^2(t(1-t))^{\nu}\, dt = o(N^{\frac{3}{2}-\nu}) \qquad \text{for } \nu = 1, \tfrac{1}{2}\,. \tag{6.5}$$

Theorem 4.2 and Lemma 6.1 provide expansions for $\pi(\theta)$ and $\pi^+(\theta)$. In view of (6.5), the boundedness of u_α and the fact that $c \leqq N^{\frac{1}{2}}\theta \leqq C$, it is clear from these expansions that $d_N = o(N^{\frac{1}{2}})$. To find d_N we replace N by $N + d_N$ and η^* by $\eta^*(1 + d_N N^{-1})^{\frac{1}{2}}$ in the expansion for $\pi(\theta)$ and equate the result to the expansion for $\pi^+(\theta)$. Taylor expansion with respect to $d_N N^{-1}$ in (4.18) yields

$$\begin{aligned} \frac{\eta^*\phi(u_\alpha - \eta^*)}{24N} \Bigg\{ 12 d_N &+ \frac{\int_0^1 \Psi_1^4(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2} [-3(u_\alpha^2 - 1) + 2\eta^* u_\alpha + \eta^{*2}] \\ &- \frac{3\int_0^1 \Psi_2^2(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2}\eta^{*2} + 3(u_\alpha^2 - 1) - 6\eta^* u_\alpha + 3\eta^{*2} \\ &- \frac{12\int_{1/N}^{1-1/N} (\Psi_1'(t))^2 t(1-t)\,dt}{\int_0^1 \Psi_1^2(t)\,dt}\Bigg\} \\ &= o(N^{-1}) + O(N^{-\frac{3}{2}} \int_{1/N}^{1-1/N} (\Psi_1'(t))^2(t(1-t))^{\frac{1}{2}}\,dt)\,, \end{aligned} \tag{6.6}$$

uniformly for fixed $F \in \mathscr{F}_1$, c, C and ε. As $\eta^*\phi(u_\alpha - \eta^*)$ is bounded away from zero, (6.4) follows. The last assertion of the theorem is an immediate consequence of Theorem 4.2. □

Obviously (6.3) and (6.4) imply that under the conditions of Theorem 6.1

$$d_N = O(\int_{1/N}^{1-1/N} (\Psi_1'(t))^2 t(1-t)\,dt) \tag{6.7}$$

for $N \to \infty$. Hence d_N remains bounded as $N \to \infty$ if $\int_0^1 (\Psi_1'(t))^2 t(1-t)\,dt$ converges. Fortunately, in most cases of interest Theorem 6.1 provides more detailed information than (6.7) and remarks similar to those following Theorem 4.2 apply. Typically Ψ_1' will be bounded near 0 and the asymptotic behavior of d_N will be determined by the rate of growth of Ψ_1' near 1. If $\Psi_1'(t) = o((1-t)^{-1})$ near 1, then $d_N = \tilde{d}_N + o(1)$. If $\Psi_1'(t)$ is of exact order $(1-t)^{-1}$, then

$$d_N = \frac{\int_0^{1-1/N} (\Psi_1'(t))^2 t(1-t)\,dt}{\int_0^1 \Psi_1^2(t)\,dt} + O(1)$$

and d_N will be of the order $\log N$. Finally, if $\Psi_1'(t) \sim (1-t)^{-1-\delta}$ for $t \to 1$ and some $0 < \delta < \frac{1}{6}$, then the expansion (6.4) reduces to $d_N = O(N^{2\delta})$, which is nothing but (6.7).

We shall give two applications of Theorem 6.1. First we consider the problem of testing the hypothesis $G = \Phi$ against the alternative $G(x) = \Phi(x - \theta)$, where $cN^{-\frac{1}{2}} \leqq \theta \leqq CN^{-\frac{1}{2}}$. Let d_N be the deficiency of the normal scores test (or van der Waerden's test) with respect to the most powerful parametric test based on $\bar{X}$. Computations similar to those in Section 4 yield

$$(6.8) \qquad d_N = \tfrac{1}{2}(u_\alpha^2 - 1) + \int_0^{\Phi^{-1}(1-1/2N)} \frac{(2\Phi(x) - 1)(1 - \Phi(x))}{\phi(x)}\, dx + o(1)$$
$$= \tfrac{1}{2}\log\log N + \tfrac{1}{2}(u_\alpha^2 - 1) + \tfrac{1}{2}\log 2 + 0.05832\cdots + o(1)\,.$$

In this case $d_N \sim \frac{1}{2}\log\log N \to \infty$ for $N \to \infty$. Note that there is no dependence on θ in this expansion for d_N and that the leading term is also independent of α.

As a second example we take the logistic df $F(x) = (1 + e^{-x})^{-1}$ and consider the testing problem $G = F$ against $G(x) = F(x - bN^{-\frac{1}{2}})$, where $b > 0$ is fixed. Now d_N is the deficiency of Wilcoxon's signed rank test with respect to the most powerful parametric test for this problem. We find

$$(6.9) \qquad d_N = \tfrac{1}{60}\{18 + 12u_\alpha^2 + 4(3)^{\frac{1}{2}}bu_\alpha + b^2\} + o(1)$$

and here d_N tends to a finite limit for $N \to \infty$.

Having shown that the deficiency of a distribution free test with respect to the best parametric test may tend to a finite limit, we now address ourselves to the intriguing question whether this limit can be zero. To answer this question we first have to decide what is meant by the best parametric test. So far, we have compared the performance of a distribution free test with that of the most powerful parametric test for known scale against a simple location alternative, thus in effect comparing with envelope power. Of course this comparison is not quite fair. Computed in this way, the deficiency of a distribution free test reflects the losses incurred by using (i) the same test against every location alternative $\theta > 0$; (ii) a scale invariant test; (iii) a distribution free test. Since our main interest is the deficiency due to (iii), it is more appropriate to compare with the uniformly most powerful scale invariant test, if such a test exists. Unfortunately, invariant tests are in general rather intractable, the main exception being Student's test for the normal location case. We note that Hodges and Lehmann (1970) have shown that the deficiency of Student's test with respect to the most powerful parametric test based on $\bar{X}$ tends to a finite but positive limit, so that it does indeed matter whether one compares with Student's test or with envelope power.

We are thus led to consider the normal location case with Student's test as the best parametric test. To establish the existence of a distribution free test with deficiency tending to zero, the obvious candidate is the permutation test based on $\sum X_i$. Theorem 6.2 is an immediate consequence of Theorem 5.1 and the remark following it.

THEOREM 6.2. *Let d_N be the deficiency of the permutation test based on $\sum X_i$ with respect to Student's test for testing $G = \Phi$ against $G(x) = \Phi(x - \theta)$ on the* *basis of* $X_1, \cdots, X_N$ *and at level* α. *Suppose that positive numbers* c, C *and* ε *exist such that* $cN^{-\frac{1}{2}} \leqq \theta \leqq CN^{-\frac{1}{2}}$ *and* $\varepsilon \leqq \alpha \leqq 1 - \varepsilon$. *Then there exists* $A > 0$ *depending on* N, θ *and* α *only through* c, C *and* ε *and such that*

$$d_N \leqq AN^{-\frac{1}{4}} . \tag{6.10}$$

Hence in this case we do find that d_N tends to zero for $N \to \infty$. Perhaps the most surprising thing about this example is that asymptotically one has to pay a certain price for scale invariance, but that once this price has been paid, there is no additional penalty for using a distribution free test. We note that the remark following Theorem 5.1 implies that (6.10) may be replaced by $d_N \leqq AN^{-\frac{1}{2}}$.

Theorem 6.2 may of course be generalized considerably by taking Theorem 5.2 for $h(x) \equiv x$ as a starting point instead of Theorem 5.1. For d_N as in Theorem 6.2, it is clear that $d_N = o(1)$ for a much larger class of testing problems than the normal location problem of Theorem 6.2. Although Student's test is generally not optimal for these problems, this shows how closely the two tests resemble one another.

7. Expansions and deficiencies for related estimators. Let $T = T(X_1, \cdots, X_N)$ be given by (2.2) and suppose that the scores a_j are nonnegative and nondecreasing in $j = 1, \cdots, N$. Define the statistic M by

$$\begin{aligned} M(X_1, \cdots, X_N) = {} & \tfrac{1}{2} \sup \{t : 2T(X_1 - t, \cdots, X_N - t) > \textstyle\sum a_j\} \\ & + \tfrac{1}{2} \inf \{t : 2T(X_1 - t, \cdots, X_N - t) < \textstyle\sum a_j\} . \end{aligned} \tag{7.1}$$

Suppose that $X_1, \cdots, X_N$ are i.i.d. with common df $G(x) = F(x - \mu)$, where F has a density f that is symmetric about zero. Then M is the midpoint of the interval between the upper and lower 0.5 confidence bounds for μ induced by the statistic T. Hodges and Lehmann (1963) proposed M as an estimator for μ and studied its connection with T. They showed that the normal approximation to the power of the level $\frac{1}{2}$ test based on T for contiguous location alternatives could be used to establish asymptotic normality of M. We shall show that, similarly, power expansions for level $\frac{1}{2}$ yield expansions for the df of $N^{\frac{1}{2}}(M - \mu)$. We restrict attention to the case where the scores are generated by a smooth function J.

Let $\mathscr{J}$ and $\mathscr{F}$ be given by Definition 4.1, let $\pi(\theta, \frac{1}{2})$ denote the power of the level $\frac{1}{2}$ right-sided test based on T against the alternative $F(x - \theta)$ and define $K_{\theta,i}$ and $\tilde{\eta}$ as in (4.8)—(4.10).

THEOREM 7.1. *Let* $F \in \mathscr{F}$, $J \in \mathscr{J}$, *suppose that* J *is nonnegative and nondecreasing and let* $a_j = EJ(U_{j:N})$. *Take* $\theta = \xi N^{-\frac{1}{2}}$. *Then, for every fixed* J, F *and* $C > 0$,

$$\sup_{|\xi| \leqq C} |P_\mu(N^{\frac{1}{2}}(M - \mu) \leqq \xi) - \pi(\theta, \tfrac{1}{2})| = O(N^{-\frac{5}{4}}) , \tag{7.2}$$

$$\sup_{|\xi| \leqq C} |P_\mu(N^{\frac{1}{2}}(M - \mu) \leqq \xi) - \{1 - K_{\theta,1}(-\tilde{\eta})\}| = o(N^{-1}) , \tag{7.3}$$

$$\begin{aligned} & \sup_{|\xi| \leqq C} |P_\mu(N^{\frac{1}{2}}(M - \mu) \leqq \xi) - \{1 - K_{\theta,2}(-\tilde{\eta})\}| \\ & \quad = o(N^{-1}) + O(N^{-\frac{3}{2}} \textstyle\int_{1/N}^{1-1/N} |J'(t)|(|J'(t)| + |\Psi_1'(t)|)(t(1 - t))^{\frac{1}{2}} \, dt) . \end{aligned} \tag{7.4}$$

PROOF. It follows from Hodges and Lehmann (1963) that M is translation invariant and that its distribution is absolutely continuous and symmetric about μ. Thus, for $\theta = \xi N^{-\frac{1}{2}}$,

$$P_\mu(N^{\frac{1}{2}}(M-\mu) \leqq \xi) = P_\theta(M \geqq 0)\,, \tag{7.5}$$

and, in view of (7.1),

$$P_\theta(2T > \sum a_j) \leqq P_\theta(M \geqq 0) \leqq P_\theta(2T \geqq \sum a_j)\,. \tag{7.6}$$

According to the proof of Theorem 4.1, the conclusions of Theorems 3.1 and 3.2 hold, which implies that $P_\theta(2T = \sum a_j) = O(N^{-\frac{5}{4}})$ uniformly for $|\theta| \leqq CN^{-\frac{1}{2}}$. This proves (7.2). The remaining part of Theorem 7.1 is now an immediate consequence of Theorem 4.1. □

The case where $J = -\Psi_1$, with Ψ_1 as in (3.15), is of course of special interest. Theorem 7.2 deals with this case for exact as well as approximate scores. Note that for $F \in \mathscr{F}$, the condition that $-\Psi_1$ is nonnegative and nondecreasing is equivalent to concavity of $\log f$, i.e. to strong unimodality of f.

THEOREM 7.2. *Let $F \in \mathscr{F}$, suppose that f is strongly unimodal and let either $a_j = -E\Psi_1(U_{j:N})$ for $j = 1, \cdots, N$ or $a_j = -\Psi_1(j/(N+1))$ for $j = 1, \cdots, N$. Then, for every fixed F and $C > 0$,*

$$\sup_{|\xi|\leq C} |P_\mu(N^{\frac{1}{2}}(M-\mu) \leqq \xi) - \pi(\xi N^{-\frac{1}{2}}, \tfrac{1}{2})| = O(N^{-\frac{5}{4}})\,, \tag{7.7}$$

$$\begin{aligned} &P_\mu((N \textstyle\int_0^1 \Psi_1^2(t)\,dt)^{\frac{1}{2}}(M-\mu) \leqq x) \\ &= \Phi(x) + \frac{x\phi(x)}{72N}\Big\{x^2\Big[\frac{5\int_0^1 \Psi_1^4(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2} - \frac{12\int_0^1 \Psi_2^2(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2} + 9\Big] \\ &\quad + \frac{6\int_0^1 \Psi_1^4(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2} - \frac{36\int_{1/N}^{1-1/N} (\Psi_1'(t))^2 t(1-t)\,dt}{\int_0^1 \Psi_1^2(t)\,dt}\Big\} \\ &\quad + o(N^{-1}) + O(N^{-\frac{3}{2}} \textstyle\int_{1/N}^{1-1/N} (\Psi_1'(t))^2 (t(1-t))^{\frac{1}{2}}\,dt) \end{aligned} \tag{7.8}$$

uniformly for $|x| \leqq C$.

PROOF. The proof of (7.7) is identical to the proof of (7.2) in Theorem 7.1. Expansion (7.8) follows from (7.7) and Theorem 4.2. □

The estimators in Theorem 7.2 are efficient and their natural competitor is the maximum likelihood estimator M' which solves

$$\sum_{j=1}^N \psi_1(X_j - M') = 0 \tag{7.9}$$

with ψ_1 as in (3.1). The performance of M' is connected with that of the locally most powerful test for F against $F(x-\theta)$, which is based on the statistic $-\sum \psi_1(X_j)$. Let $\pi'(\theta, \frac{1}{2})$ be the power of the level $\frac{1}{2}$ right-sided test based on $-\sum \psi_1(X_j)$ for F against $F(x-\theta)$.

LEMMA 7.1. *Suppose that f is positive, symmetric about zero and strongly unimodal and that* (4.5) *is satisfied for $m_i = 5/i$, $i = 1, \cdots, 5$. Then, for every fixed F and*

$C > 0$,

$$\sup_{|\xi| \leqq C} |P_\mu(N^{\frac{1}{2}}(M' - \mu) \leqq \xi) - \pi'(\xi N^{-\frac{1}{2}}, \tfrac{1}{2})| = O(N^{-\frac{5}{4}}), \tag{7.10}$$

$$\begin{aligned} P_\mu((N \textstyle\int_0^1 \Psi_1^2(t)\,dt)^{\frac{1}{2}}(M' - \mu) \leqq x) \\ = \Phi(x) + \frac{x\phi(x)}{72N} \Big\{ x^2 \Big[\frac{5 \int_0^1 \Psi_1^4(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2} - \frac{12 \int_0^1 \Psi_2^2(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2} + 9 \Big] \\ - \frac{3 \int_0^1 \Psi_1^4(t)\,dt}{(\int_0^1 \Psi_1^2(t)\,dt)^2} + 9 \Big\} + O(N^{-\frac{5}{4}}) \end{aligned} \tag{7.11}$$

uniformly for $|x| \leqq C$.

PROOF. The estimator M' is translation invariant and its distribution is symmetric about μ. Thus, for $\theta = \xi N^{-\frac{1}{2}}$, (7.5) holds with M replaced by M', and in view of (7.9),

$$P_\theta(-\textstyle\sum \psi_1(X_j) > 0) \leqq P_\mu(N^{\frac{1}{2}}(M' - \mu) \leqq \xi) \leqq P_\theta(-\sum \psi_1(X_j) \geqq 0). \tag{7.12}$$

Since f is everywhere positive and ψ_1 is everywhere differentiable, the distribution of $\psi_1(X_1)$ under θ contains a fixed absolutely continuous component for all θ in a neighborhood of zero. Together with (4.5) for $m_1 = 5$, this ensures that the df of $\sum \psi_1(X_j)$ under θ possesses an Edgeworth expansion with remainder $O(N^{-\frac{5}{4}})$ uniformly for $|\theta| \leqq CN^{-\frac{1}{2}}$. This implies that $P_\theta(-\sum \psi_1(X_j) = 0) = O(N^{-\frac{5}{4}})$ uniformly for $|\theta| \leqq CN^{-\frac{1}{2}}$, which proves (7.10).

The expansion for the df of $\sum \psi_1(X_j)$ is used in Albers (1974) to establish an expansion for the power of the locally most powerful test under the conditions of Lemma 6.1. Specializing to the case where $\alpha = \frac{1}{2}$ and using (7.10) we obtain (7.11). □

There is no unique natural measure of scale to assess the performance of an estimator $\hat{\mu}$ admitting an expansion of the form (7.8) or (7.11). One possibility is to consider a family of measures determined by the quantiles of $\hat{\mu}$. We can define $\sigma(\hat{\mu}, s)$ to be the s-quantile of $(\hat{\mu} - \mu)$ divided by $u_{1-s} = \Phi^{-1}(s)$. As we are only considering estimators that are distributed symmetrically about μ, $\sigma(\hat{\mu}, s)$ may serve as a measure of scale for any $\frac{1}{2} < s < 1$. If we fix a value of s, we can define the deficiency $D_N(s)$ of a sequence of estimators $\{\hat{\mu}_{2,N}\}$ with respect to an estimator $\hat{\mu}_{1,N}$ by equating $\sigma(\hat{\mu}_{2,N+D_N}, s)$ and $\sigma(\hat{\mu}_{1,N}, s)$, with the usual convention that σ is determined by linear interpolation for nonintegral values of $N + D_N$. Similarly, for two sequences of level α tests, $d_N(\alpha, s)$ will denote the deficiency as defined in Section 1 for the case where the alternative θ is chosen in such a way that the common power equals s.

Let $\mathscr{F}_1$ be given by Definition 6.1.

THEOREM 7.3. *Let* $d_N(\frac{1}{2}, s)$ *be the deficiency for level* $\frac{1}{2}$ *and power* s *of the locally most powerful rank test with respect to the locally most powerful test for testing* F *against* $F(x - \theta)$. *Let* $D_N(s)$ *be the deficiency of the Hodges–Lehmann estimator associated with the locally most powerful rank test with respect to the maximum likelihood estimator for estimating* μ *in* $F(x - \mu)$. *Suppose that* $F \in \mathscr{F}_1$ *and that* f *is* *strongly unimodal. Then, for fixed F and* $\frac{1}{2} < s < 1$,

$$(7.13) \qquad |D_N(s) - d_N(\tfrac{1}{2}, s)| = O(N^{-\frac{1}{4}}) ,$$

$$(7.14) \qquad D_N(s) = \frac{\int_{1/N}^{1-1/N} (\Psi_1'(t))^2 t(1-t)\, dt}{\int_0^1 \Psi_1^2(t)\, dt} - \frac{1}{4} \frac{\int_0^1 \Psi_1^4(t)\, dt}{(\int_0^1 \Psi_1^2(t)\, dt)^2} + \tfrac{1}{4} + o(1) + O(N^{-\frac{1}{2}} \textstyle\int_{1/N}^{1-1/N} (\Psi_1'(t))^2 (t(1-t))^{\frac{1}{2}}\, dt) .$$

This result continues to hold if in the locally most powerful rank test and the associated estimator, the exact scores are replaced by the approximate scores $a_j = -\Psi_1(j/(N+1))$.

PROOF. The conditions of Theorem 7.2 and Lemma 7.1 are satisfied. Writing M_N and M_N' for M and M', we see that for some ξ

$$(7.15) \qquad P_\mu(N^{\frac{1}{2}}(M_N' - \mu) \leqq \xi) = s + O(N^{-\frac{3}{4}}) ,$$

$$(7.16) \qquad P_\mu(N^{\frac{1}{2}}(M_{N+d_N} - \mu) \leqq \xi) = s + O(N^{-\frac{3}{4}}) .$$

By the remark following Theorem 4.2 we have $\Psi_1'(t) = o((t(1-t))^{-\frac{3}{2}})$ near 0 and 1, and combining this with (7.8) and (7.11) we find that (7.15) and (7.16) imply (7.13). The proof of (7.14) is now the same as that of Theorem 6.1. □

An interesting property of the expansion (7.14) is that it is independent of s. Thus, to the order considered, the deficiency $D_N(s)$ is asymptotically independent of the particular choice of the quantile used to measure the performance of the estimators. Of course, this reflects the fact that the deficiency $d_N(\frac{1}{2}, s)$ is independent of the power in the same asymptotic sense. Algebraically, the reason for this phenomenon is that the term involving $x^3\phi(x)$ is the same in (7.8) and (7.11).

We also note that upon formal substitution of $\alpha = \frac{1}{2}$ and $\theta = 0$ in (6.3), the expansion for d_N in Theorem 6.1 reduces to the expansion for $D_N(s)$ in Theorem 7.3. This shows that if the remainder in (7.14) is $o(1)$, then $D_N(s)$ will tend to a nonnegative but possibly infinite limit.

In Section 6 we have already pointed out that an expansion like (7.14) may or may not be of interest, depending on the behavior of the remainder term. We should stress that, even if the expansion (7.14) is useless, (7.13) still establishes the asymptotic equivalence of $D_N(s)$ and $d_N(\frac{1}{2}, s)$.

We conclude our discussion with one example of Theorem 7.3. For estimating normal location, the deficiency of either one of the Hodges–Lehmann estimators associated with the normal scores test and with van der Waerden's test with respect to $\bar{X}$ is asymptotic to $\frac{1}{2} \log\log N$. The deficiency of one of these Hodges–Lehmann estimators with respect to the other tends to zero for $N \to \infty$.

APPENDIX

1. Expansions for the contiguous case. Our purpose in this appendix will be the justification of the passage from (2.41) to (3.8) under the assumptions stated in Section 3. Thus we shall suppose throughout that f is positive and symmetric about 0 and that $g(x) = f(x - \theta)$.

Begin by defining a function $\xi(x, t)$ for $x \geqq 0$, $t \geqq 0$, by

$$F(\xi(x, t) - t) + F(\xi(x, t) + t) = 2F(x) . \tag{A1.1}$$

Introduce also two other functions of two variables, p and $\bar{p}$, by

$$p(x, t) = \frac{f(x - t)}{f(x - t) + f(x + t)} , \tag{A1.2}$$

$$\bar{p}(x, t) = p(\xi(x, t), t) . \tag{A1.3}$$

The basic property of the function ξ is, of course, that the joint distribution of $(\xi(Z_1, \theta), \cdots, \xi(Z_N, \theta))$ under F is the same as the joint distribution of $(Z_1, \cdots, Z_N)$ under G. It follows that the joint distribution of $(\bar{p}(Z_1, \theta), \cdots, \bar{p}(Z_N, \theta))$ under F is the same as the joint distribution of $(P_1, \cdots, P_N)$ under G. It is evident therefore that our task is essentially that of expanding $\bar{p}(x, t)$ around 0 as a function of t and giving suitable estimates of the remainder terms. We begin by differentiating formally. For convenience we shall, for any function of two variables $q(x, t)$, write

$$q_{i,j}(x, t) = \frac{\partial^{i+j} q(x, t)}{\partial x^i \, \partial t^j} .$$

Differentiating (A1.1) with respect to t we get

$$\xi_{0,1} = 2\bar{p} - 1 . \tag{A1.4}$$

It is now easy though tedious to obtain $\bar{p}_{0,j}(x, t)$ in terms of the $p_{i,k}(\xi(x, t), t)$ by replacing $\xi_{0,1}$ by $2\bar{p} - 1$ after each differentiation. Thus, for example,

$$\bar{p}_{0,1}(x, t) = [p_{0,1} + p_{1,0}(2p - 1)](\xi(x, t), t) , \tag{A1.5}$$

$$\begin{aligned} \bar{p}_{0,2}(x, t) = {} & [p_{0,2} + 2p_{1,1}(2p - 1) + p_{2,0}(2p - 1)^2 + 2p_{1,0}p_{0,1} \\ & + 2p_{1,0}^2(2p - 1)](\xi(x, t), t) . \end{aligned} \tag{A1.6}$$

Calculation of the $p_{i,j}$ is also tedious. Again we list the first few. Define

$$_1\psi_k(x, t) = \psi_k(x - t) , \qquad _2\psi_k(x, t) = \psi_k(x + t) , \tag{A1.7}$$

where $\psi_k = f^{(k)}/f$ as defined in (3.1), and let

$$_1\tilde{\psi}_k(x, t) = \psi_k(\xi(x, t) - t) , \qquad _2\tilde{\psi}_k(x, t) = \psi_k(\xi(x, t) + t) . \tag{A1.8}$$

Then

$$\begin{aligned} & p_{0,1} = -p(1 - p)[_1\psi_1 + {}_2\psi_1] , \qquad p_{1,0} = p(1 - p)[_1\psi_1 - {}_2\psi_1] , \\ & p_{0,2} = p(1 - p)[_1\psi_2 - {}_2\psi_2 - 2p \cdot {}_1\psi_1^2 + 2(1 - p)_2\psi_1^2 \\ & \qquad + 2(1 - 2p)_1\psi_1 \cdot {}_2\psi_1] , \end{aligned} \tag{A1.9}$$

$$\begin{aligned} & p_{1,1} = p(1 - p)[-{}_1\psi_2 - {}_2\psi_2 + 2p \cdot {}_1\psi_1^2 + 2(1 - p)_2\psi_1^2] , \\ & p_{2,0} = p(1 - p)[_1\psi_2 - {}_2\psi_2 - 2p \cdot {}_1\psi_1^2 + 2(1 - p)_2\psi_1^2 \\ & \qquad - 2(1 - 2p)_1\psi_1 \cdot {}_2\psi_1] . \end{aligned} \tag{A1.10}$$

Substituting (A1.9) and (A1.10) into (A1.5) and (A1.6) at $t = 0$ and employing similar manipulations with the third order derivatives we obtain

$$\text{(A1.11)} \qquad \tilde{p}(x, 0) = \tfrac{1}{2}, \qquad \tilde{p}_{0,1}(x, 0) = -\tfrac{1}{2}\psi_1(x), \qquad \tilde{p}_{0,2}(x, 0) = 0,$$
$$\tilde{p}_{0,3}(x, 0) = -\tfrac{1}{2}\psi_3(x) + 3\psi_1(x)\psi_2(x) - \tfrac{3}{2}\psi_1^3(x).$$

Moreover, from (A1.9), (A1.10) and the boundedness of p it is easy to see that constants b_1 and b_2 exist such that

$$\text{(A1.12)} \qquad |\tilde{p}_{0,1}| \leqq b_1 \textstyle\sum_{i=1}^2 |{}_i\tilde{\psi}_1|, \qquad |\tilde{p}_{0,2}| \leqq b_2 \textstyle\sum_{i=1}^2 \{|{}_i\tilde{\psi}_2| + {}_i\tilde{\psi}_1^2\}.$$

Similarly bounding first the $p_{i,k}$ and expressing $\tilde{p}_{0,j}$ appropriately, and invoking the inequality $|ab| \leqq r^{-1}|a|^r + s^{-1}|b|^s$, $r^{-1} + s^{-1} = 1$, we obtain for suitable b_3 and b_4

$$\text{(A1.13)} \qquad |\tilde{p}_{0,3}| \leqq b_3 \textstyle\sum_{i=1}^2 \{|{}_i\tilde{\psi}_3| + |{}_i\tilde{\psi}_2|^{\frac{3}{2}} + |{}_i\tilde{\psi}_1|^3\},$$
$$|\tilde{p}_{0,4}| \leqq b_4 \textstyle\sum_{i=1}^2 \{|{}_i\tilde{\psi}_4| + |{}_i\tilde{\psi}_3|^{\frac{4}{3}} + {}_i\tilde{\psi}_2^2 + {}_i\tilde{\psi}_1^4\}.$$

We need the following application of Taylor's formula with Cauchy's form of the remainder.

LEMMA A1.1. *Let $q(x, t)$ be a function of two variables possessing derivatives of order $\leqq k + 1$ in t in a neighborhood of* 0. *Then if S is any* rv *and $m \geqq 1$,*

$$\text{(A1.14)} \qquad E\left|q(S, t) - \textstyle\sum_{j=0}^k q_{0,j}(S, 0)\dfrac{t^j}{j!}\right|^m$$
$$\leqq \left[\frac{|t|^{k+1}}{(k+1)!}\right]^m \sup\{E|q_{0,k+1}(S, \nu t)|^m : 0 \leqq \nu \leqq 1\}.$$

Suppose moreover that for $j = 0, \cdots, k$, $Eq_{0,j}(S, 0)$ exists and is finite. Then

$$\text{(A1.15)} \qquad E\left|\{q(S, t) - Eq(S, t)\} - \textstyle\sum_{j=0}^k \{q_{0,j}(S, 0) - Eq_{0,j}(S, 0)\}\dfrac{t^j}{j!}\right|^m$$
$$\leqq 2^m \left[\frac{|t|^{k+1}}{(k+1)!}\right]^m \sup\{E|q_{0,k+1}(S, \nu t)|^m : 0 \leqq \nu \leqq 1\}.$$

PROOF. We have (cf. Dieudonné (1960), page 186, Titchmarsh (1939), page 368)

$$\text{(A1.16)} \qquad q(S, t) = \textstyle\sum_{j=0}^k q_{0,j}(S, 0)\dfrac{t^j}{j!}$$
$$+ \frac{t^{k+1}}{(k+1)!} \int_0^1 (k+1)(1-\nu)^k q_{0,k+1}(S, \nu t)\, d\nu$$

provided that the integral converges. Hence the left-hand side of (A1.14) is bounded by

$$\left[\frac{|t|^{k+1}}{(k+1)!}\right]^m E\left|\int_0^1 (k+1)(1-\nu)^k q_{0,k+1}(S, \nu t)\, d\nu\right|^m.$$

This obviously remains true even if the integral diverges for some values of S. An application of Ljapunov's inequality and Fubini's theorem complete the proof of (A1.14) and a similar argument disposes of (A1.15). □

Note that by using the same device one can show that the left-hand side of (A1.14) and (A1.15) is $o(|t|^{mk})$ for $t \to 0$ if q is k times continuously differentiable and

$$\text{(A1.17)} \qquad \lim_{t \to 0} E|q_{0,k}(S, t)|^m = E|q_{0,k}(S, 0)|^m .$$

Of course (A1.17) holds if $q_{0,k}(S, \cdot)$ is continuous at 0 and

$$\text{(A1.18)} \qquad \sup \{E|q_{0,k}(S, t)|^{m+\delta} : |t| \leqq \delta\} < \infty$$

for some $\delta > 0$.

We introduce two final pieces of notation. If $d_1, \cdots, d_N$ is a sequence of numbers we write

$$\text{(A1.19)} \qquad ||d|| = \frac{1}{N} \sum_{j=1}^N |d_j| .$$

If χ is a function of one variable and $\varepsilon > 0$ is fixed we define

$$\text{(A1.20)} \qquad ||\chi|| = \sup \{\textstyle\int_{-\infty}^{\infty} |\chi(x + y)| f(x)\, dx : |y| \leqq \varepsilon\} .$$

THEOREM A1.1. *Suppose that f is four times differentiable, that $E_0\psi_3(|X_1|)$, $E_0\psi_1(|X_1|)\psi_2(|X_1|)$ and $E_0\psi_1^3(|X_1|)$ exist and are finite and that $0 \leqq 2\theta \leqq \varepsilon$. Then if $r \geqq 1$, $r^{-1} + s^{-1} = 1$, there exists a constant B such that*

$$\text{(A1.21)} \qquad \begin{aligned} \sum_{j=1}^N a_j(2\pi_j - 1) = -\theta \sum_{j=1}^N a_j E_0\psi_1(Z_j) - \frac{\theta^3}{6} \sum_{j=1}^N a_j E_0[\psi_3(Z_j) \\ - 6\psi_1(Z_j)\psi_2(Z_j) + 3\psi_1^3(Z_j)] + M_1 , \\ |M_1| \leqq BN\theta^4 ||a^r||^{1/r}[||\psi_4^s|| + ||\psi_3^{4s/3}|| + ||\psi_2^{2s}|| + ||\psi_1^{4s}||]^{1/s} ; \end{aligned}$$

$$\text{(A1.22)} \qquad \begin{aligned} \sum_{j=1}^N a_j^3(2\pi_j - 1) = -\theta \sum_{j=1}^N a_j^3 E_0\psi_1(Z_j) + M_2 , \\ |M_2| \leqq BN\theta^3 ||a^{3r}||^{1/r}[||\psi_3^s|| + ||\psi_2^{3s/2}|| + ||\psi_1^{3s}||]^{1/s} ; \end{aligned}$$

$$\text{(A1.23)} \qquad \begin{aligned} \sum_{j=1}^N a_j^2 E_\theta(2P_j - 1)^2 = \theta^2 \sum_{j=1}^N a_j^2 E_0\psi_1^2(Z_j) + M_3 , \\ |M_3| \leqq BN\theta^3 ||a^{2r}||^{1/r}[||\psi_3^s|| + ||\psi_2^{3s/2}|| + ||\psi_1^{3s}||]^{1/s} ; \end{aligned}$$

$$\text{(A1.24)} \qquad \begin{aligned} \sigma_\theta^2\Big(\sum_{j=1}^N a_j P_j\Big) = \frac{\theta^2}{4} \sigma_0^2\Big(\sum_{j=1}^N a_j \psi_1(Z_j)\Big) + M_4 , \\ |M_4| \leqq BN^2\theta^{\frac{24}{8}} ||a^2||[||\psi_3^{\frac{4}{3}}|| + ||\psi_2^3|| + ||\psi_1^6||] + BN\theta^{\frac{19}{8}} ||a^3||^{\frac{1}{3}} \\ \times [||\psi^{\frac{4}{3}}|| + ||\psi_2^3|| + ||\psi_1^6||]^{\frac{2}{3}}[E_0|\textstyle\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3]^{\frac{1}{3}} . \end{aligned}$$

Moreover, for $m \geqq 1$ and $\rho > 0$ there exist B' and B'' depending only on m and on m and ρ respectively, and such that

$$\text{(A1.25)} \qquad \sum_{j=1}^N E_\theta|2P_j - 1|^m \leqq B'N\theta^m ||\psi_1^m|| ;$$

$$\text{(A1.26)} \qquad \begin{aligned} [\textstyle\sum_{j=1}^N \{E_\theta|P_j - \pi_j|^m\}^\rho]^{1/\rho} \\ \leqq \theta^m[\textstyle\sum \{E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^m\}^\rho]^{1/\rho} \\ + B''N^{1/\rho}\theta^{2m}[||\psi_2^{m(\rho \vee 1)}|| + ||\psi_1^{2m(\rho \vee 1)}|| + 1]^{1/\rho} , \end{aligned}$$

where $\rho \vee 1$ denotes the larger of ρ and 1.

PROOF. In (A1.14) we take $E = E_0$, $q(Z, \theta) = \sum a_j(2\bar{p}(Z_j, \theta) - 1)$, $k = 3$,

$m = 1$, and find

$$|M_1| \leqq \frac{\theta^4}{4!} \sup \{E_0 | 2 \sum a_j \tilde{p}_{0,4}(Z_j, \nu\theta)| : 0 \leqq \nu \leqq 1\}$$

$$\leqq \frac{N\theta^4}{12} ||a^r||^{1/r} \sup \left\{ \left[\frac{1}{N} \sum E_0 |\tilde{p}_{0,4}(Z_j, \nu\theta)|^s \right]^{1/s} : 0 \leqq \nu \leqq 1 \right\} ,$$

by Hölder's and Ljapunov's inequalities. Since $\sum |\tilde{p}_{0,4}(Z_j, \nu\theta)|^s$ is symmetric in $Z_1, \cdots, Z_N$, we have

$$\frac{1}{N} \sum E_0 |\tilde{p}_{0,4}(Z_j, \nu\theta)|^s = E_0 |\tilde{p}_{0,4}(|X_1|, \nu\theta)|^s .$$

Now we apply (A1.13) and use the fact that the distribution of ${}_i\tilde{\psi}_j(|X_1|, \nu\theta)$ under $F(x)$ is the same as that of ${}_i\psi_j(|X_1|, \nu\theta)$ under $F(x - \nu\theta)$ to obtain

$$E_0 |\tilde{p}_{0,4}(|X_1|, \nu\theta)|^s \leqq b_4{}^s E_{\nu\theta} [\textstyle\sum_{i=1}^2 \{|{}_i\psi_4(|X_1|, \nu\theta)| + |{}_i\psi_3(|X_1|, \nu\theta)|^{\frac{4}{3}}$$
$$+ {}_i\psi_2{}^2(|X_1|, \nu\theta) + {}_i\psi_1{}^4(|X_1|, \nu\theta)\}]^s .$$

Because $s \geqq 1$ and $0 \leqq 2\nu\theta \leqq \varepsilon$ for $0 \leqq \nu \leqq 1$, this implies that

$$E_0 |\tilde{p}_{0,4}(|X_1|, \nu\theta)|^s \leqq 8^{s-1} b_4{}^s [||\psi_4{}^s|| + ||\psi_3{}^{4s/3}|| + ||\psi_2{}^{2s}|| + ||\psi_1{}^{4s}||] ,$$

which proves (A1.21).

The proof of (A1.22), (A1.23) and (A1.25) is similar. In each case we can apply (A1.14), taking $q(Z, \theta) = \sum a_j{}^3(2\tilde{p}(Z_j, \theta) - 1)$, $k = 2$, $m = 1$ to prove (A1.22), and $q(Z, \theta) = \sum a_j{}^2(2\tilde{p}(Z_j, \theta) - 1)^2$, $k = 2$, $m = 1$ to prove (A1.23). In (A1.25) the symmetry in $Z_1, \cdots, Z_N$ is already present from the start, so here we use (A1.14) with $q(|X_1|, \theta) = 2\tilde{p}(|X_1|, \theta) - 1$, $k = 0$ and the value of m as in (A1.25).

A rather delicate argument is needed to deal with (A1.24). Because $\tilde{p}_{0,2}(x, 0) = 0$,

$$\left(\tilde{p}(x, t) - \tfrac{1}{2} + \frac{t}{2} \psi_1(x)\right)^2$$
$$= \left| \frac{t^2}{2} \textstyle\int_0^1 2(1 - \nu)\tilde{p}_{0,2}(x, \nu t)\, d\nu \right|^{\frac{6}{5}} \left| \frac{t^3}{6} \textstyle\int_0^1 3(1 - \nu)^2 \tilde{p}_{0,3}(x, \nu t)\, d\nu \right|^{\frac{4}{5}}$$
$$\leqq |t|^{\frac{24}{5}} \{|\tfrac{1}{2} \textstyle\int_0^1 2(1 - \nu)\tilde{p}_{0,2}(x, \nu t)\, d\nu|^3 + |\tfrac{1}{6} \textstyle\int_0^1 3(1 - \nu)^2 \tilde{p}_{0,3}(x, \nu t)\, d\nu|^{\frac{4}{3}}\}$$
$$\leqq |t|^{\frac{24}{5}} \textstyle\int_0^1 \{|\tilde{p}_{0,2}(x, \nu t)|^3 + |\tilde{p}_{0,3}(x, \nu t)|^{\frac{4}{3}}\}\, d\nu ,$$

and similarly,

$$\left| \tilde{p}(x, t) - \tfrac{1}{2} + \frac{t}{2} \psi_1(x) \right|^{\frac{3}{2}} \leqq |t|^{\frac{21}{5}} \textstyle\int_0^1 \{|\tilde{p}_{0,2}(x, \nu t)|^3 + |\tilde{p}_{0,3}(x, \nu t)|^{\frac{4}{3}}\}\, d\nu .$$

By now familiar manipulations yield

$$\left| \sigma_\theta{}^2(\sum a_j P_j) - \frac{\theta^2}{4} \sigma_0{}^2(\sum a_j \psi_1(Z_j)) \right|$$
$$\leqq \sigma_0{}^2 \left(\sum a_j \left\{ \tilde{p}(Z_j, \theta) + \frac{\theta}{2} \psi_1(Z_j) \right\} \right)$$
$$+ \theta \left| \mathrm{Cov}_0 \left(\sum a_j \left\{ \tilde{p}(Z_j, \theta) + \frac{\theta}{2} \psi_1(Z_j) \right\} , \sum a_j \psi_1(Z_j) \right) \right|$$

$$\leqq N^2||a^2||E_0\left\{\tilde{p}(|X_1|, \theta) - \tfrac{1}{2} + \frac{\theta}{2}\psi_1(|X_1|)\right\}^2 + N\theta||a^3||^{\frac{1}{3}}\left[E_0\left|\tilde{p}(|X_1|, \theta) - \tfrac{1}{2}\right.\right.$$

$$\left.\left. + \frac{\theta}{2}\psi_1(|X_1|)\right|^{\frac{3}{2}}\right]^{\frac{2}{3}}[E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3]^{\frac{1}{3}}$$

$$\leqq BN^2\theta^{\frac{24}{8}}||a^2||[||\psi_3^{\frac{4}{3}}|| + ||\psi_2^3|| + ||\psi_1^6||] + BN\theta^{\frac{19}{8}}||a^3||^{\frac{1}{3}}$$

$$\times [||\psi_3^{\frac{4}{3}}|| + ||\psi_2^3|| + ||\psi_1^6||]^{\frac{2}{3}}[E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3]^{\frac{1}{3}} .$$

It remains to consider (A1.26). Since

$$\tilde{p}(Z_j, \theta) - E_0\tilde{p}(Z_j, \theta) = \theta[\tilde{p}_{0,1}(Z_j, 0) - E_0\tilde{p}_{0,1}(Z_j, 0)]$$
$$+ \frac{\theta^2}{2}\int_0^1 [|\tilde{p}_{0,2}(Z_j, \nu\theta)| + E_0|\tilde{p}_{0,2}(Z_j, \nu\theta)|]2(1 - \nu)\, d\nu ,$$

and $m \geqq 1$, we have

$$E_\theta|P_j - \pi_j|^m \leqq 2^{m-1}\theta^m E_0|\tilde{p}_{0,1}(Z_j, 0) - E_0\tilde{p}_{0,1}(Z_j, 0)|^m$$
$$+ \frac{\theta^{2m}}{2}E_0\int_0^1 \{|\tilde{p}_{0,2}(Z_j, \nu\theta)| + E_0|\tilde{p}_{0,2}(Z_j, \nu\theta)|^m\}2(1 - \nu)\, d\nu$$
$$\leqq \frac{\theta^m}{2}E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^m$$
$$+ 2^{m-1}\theta^{2m}\int_0^1 E_0|\tilde{p}_{0,2}(Z_j, \nu\theta)|^m 2(1 - \nu)\, d\nu .$$

Hence

$$\sum \{E_\theta|P_j - \pi_j|^m\}^\rho \leqq \theta^{m\rho}\sum \{E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^m\}^\rho$$
$$+ 2^{m\rho}N\theta^{2m\rho}[1 + \sup\{E_0|\tilde{p}_{0,2}(|X_1|, \nu\theta)|^{m(\rho \vee 1)} : 0 \leqq \nu \leqq 1\}] .$$

Proceeding as before we prove (A1.26) and the theorem. □

COROLLARY A1.1. *Suppose that positive numbers c, C and ε exist such that* (2.35), (3.2) *and* (3.3) *are satisfied. Let $\tilde{K}$, K_θ and η be defined by* (2.40), (3.4) *and* (3.5). *Then there exists $A > 0$ depending on N, a, F and θ only through c, C and ε, and such that*

(A1.27) $$\sup_x\left|\tilde{K}\left(x - \frac{\sum a_j(2\pi_j - 1)}{(\sum a_j^2)^{\frac{1}{2}}}\right) - K_\theta(x - \eta)\right|$$
$$\leqq A\{N^{-\frac{5}{4}} + \theta^{\frac{19}{8}}[E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3]^{\frac{1}{3}}$$
$$+ N^{-\frac{1}{2}}\theta^5\sigma_0^2(\sum a_j\,\psi_1(Z_j))\} ,$$

(A1.28) $$|\sum a_j^m E_0\psi_1(Z_j)| \leqq AN \qquad \textit{for} \quad m = 1, 3 ,$$

(A1.29) $$|\sum a_j^2 E_0\psi_1^2(Z_j)| \leqq AN ,$$

(A1.30) $$|\sum a_j E_0[\psi_3(Z_j) - 6\psi_1(Z_j)\psi_2(Z_j) + 3\psi_1^3(Z_j)]| \leqq AN ,$$

(A1.31) $$\sum E_\theta|2P_j - 1|^m \leqq AN^{1-m/2} \qquad \textit{for} \quad 1 \leqq m \leqq 6 ,$$

(A1.32) $$[\sum \{E_\theta|P_j - \pi_j|^3\}^{\frac{4}{3}}]^{\frac{3}{4}} \leqq \theta^3[\sum \{E_0|\psi_1(Z_j) - E_0\psi_1(Z_j)|^3\}^{\frac{4}{3}}]^{\frac{3}{4}} + AN^{-\frac{3}{4}} .$$

PROOF. Since the corollary is trivially true for $N \leqq (2C/\varepsilon)^2$, we may assume that $2\theta \leqq 2CN^{-\frac{1}{2}} \leqq \varepsilon$ and use the results in Theorem A1.1. We note that (2.35) implies that $||a^r|| \leqq [C^r \max(1, N^{r-4})]^{\frac{1}{4}}$. In the notation of this appendix (3.2) asserts that $||\psi_i^{m_i}|| \leqq C$ for $m_1 = 6$, $m_2 = 3$, $m_3 = \frac{4}{3}$ and $m_4 = 1$. All order symbols in this proof are uniform for fixed c, C and ε.

(A1.28)—(A1.30) follow from (2.35) and (3.2) by Hölder's and Ljapunov's inequalities, e.g.

$$|\sum a_j^3 E_0 \psi_1(Z_j)| = O(N\,{}'a'\,.\,\psi_1^4\|^{\frac{1}{4}}) = O(N)\,.$$

(A1.31) and (A1.32) are immediate consequences of (A1.25) and (A1.26).

Taking $r = 4$, $s = \frac{4}{3}$ in (A1.22)—(A1.24) we find

(A1.33)
$$M_2 = O(1)\,, \qquad M_3 = O(N^{-\frac{1}{4}})\,,$$
$$M_4 = O(N^{-\frac{2}{5}} + N\theta^{\frac{19}{8}}[E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3]^{\frac{1}{3}})\,.$$

Hence, uniformly in x,

$$\tilde{K}(x) = \Phi(x) + \phi(x)\left\{\frac{\sum a_j^4}{12(\sum a_j^2)^2}(x^3 - 3x) - \theta\frac{\sum a_j^3 E_0\psi_1(Z_j)}{3(\sum a_j^2)^{\frac{3}{2}}}(x^2 - 1)\right.$$

(A1.34)
$$\left. + \frac{\theta^2}{2\sum a_j^2}[\sum a_j^2 E_0\psi_1^2(Z_j) - \sigma_0^2(\sum a_j\psi_1(Z_j))]x\right\}$$
$$+ O(N^{-\frac{5}{4}} + \theta^{\frac{19}{8}}[E_0|\sum a_j(\psi_1(Z_j) - E_0\psi_1(Z_j))|^3]^{\frac{1}{3}})\,.$$

Taking $r = \infty$, $s = 1$ in (A1.21) we have

(A1.35)
$$\frac{\sum a_j(2\pi_j - 1)}{(\sum a_j^2)^{\frac{1}{2}}} = \eta - \frac{\theta^3}{6(\sum a_j^2)^{\frac{1}{2}}}\sum a_j E_0[\psi_3(Z_j)$$
$$- 6\psi_1(Z_j)\psi_2(Z_j) + 3\psi_1^3(Z_j)] + O(N^{\frac{3}{4}}\theta^4)\,,$$

where the second term on the right is $O(N^{\frac{1}{2}}\theta^3)$ by (A1.30). Now we substitute $x - (\sum a_j^2)^{-\frac{1}{2}}\sum a_j(2\pi_j - 1)$ for x in (A1.34) and expand the right-hand side around $x - \eta$. It follows from (A1.35), (A1.28) for $m = 3$ and (A1.29) that in this way we obtain (A1.27).

2. Asymptotic behavior of moments of functions of order statistics. Our aim in this appendix is twofold. In the first place we provide a proof of Theorem 3.2 where the order of the remainder in expansion (3.8) is evaluated. Secondly, we obtain asymptotic expressions for the leading terms in the expansion for the case where exact or approximate scores are used, thus in effect proving Theorems 4.1 and 4.2.

Let $U_{1:N} < U_{2:N} < \cdots < U_{N:N}$ be order statistics of a sample of size N from the uniform distribution on (0, 1).

LEMMA A2.1. *If $\lambda = j/(N+1)$ then for all $N = 1, 2, \cdots, j = 1, \cdots, N$ and $t \geqq 0$,*

$$P\left(|U_{j:N} - \lambda|\left(\frac{N}{\lambda(1-\lambda)}\right)^{\frac{1}{2}} \geqq t\right) \leqq 2\exp\left\{-\frac{3t^2}{6t + 8}\right\}.$$

PROOF. The probability on the left is equal to

(A2.1)
$$B\left(j, N, \lambda - t\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}}\right)$$
$$+ B\left(N - j + 1, N, 1 - \lambda - t\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}}\right)$$

where

$$B(j, N, p) = \sum_{k=j}^{N}\binom{N}{k}p^k(1-p)^{N-k}\,.$$

For $j > Np$ Bernstein's inequality (cf. Hoeffding (1963) page 17) yields

$$B(j, N, p) \leqq \exp\left\{-\frac{j - Np}{1 - p}\, h\left(\frac{j - Np}{Np}\right)\right\}$$

with $h(s) = 3s(2s + 6)^{-1}$. Application of this result gives after some algebra

$$B\left(j, N, \lambda - t\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}}\right)$$

$$\leqq \exp\left\{-\frac{3}{2}\frac{[t + (\lambda/N(1-\lambda))^{\frac{1}{2}}]^2}{(3 + N^{-1}) + t(N\lambda(1-\lambda))^{-\frac{1}{2}}[\lambda(5 + N^{-1}) - 2] - 2N^{-1}t^2}\right\}.$$

Noting that $\lambda \leqq N(N+1)^{-1}$ and $(N\lambda(1-\lambda))^{-\frac{1}{2}} \leqq 1 + N^{-1}$, we see that $\exp\{-3t^2(6t + 8)^{-1}\}$ is an upper bound for the first term in (A2.1). By interchanging j and $(N - j + 1)$ we find that the same is true for the second term in (A2.1) which proves the lemma. □

LEMMA A2.2. *If $\lambda = j/(N+1)$, k is a positive real number, ν_k is the kth absolute moment of the standard normal distribution and $I_{(a,b)}$ is the indicator of (a, b), then uniformly for $j = 1, \cdots, N$ and $\eta \geqq \frac{1}{2}\lambda(1-\lambda)$ we have for $N \to \infty$,*

$$\left(\frac{N}{\lambda(1-\lambda)}\right)^{\frac{1}{2}k} E(\lambda - U_{j:N})^k I_{(\lambda-\eta,\lambda)}(U_{j:N}) = \tfrac{1}{2}\nu_k + O((N\lambda(1-\lambda))^{-\frac{1}{2}}),$$

$$\left(\frac{N}{\lambda(1-\lambda)}\right)^{\frac{1}{2}k} E(U_{j:N} - \lambda)^k I_{(\lambda,\lambda+\eta)}(U_{j:N}) = \tfrac{1}{2}\nu_k + O((N\lambda(1-\lambda))^{-\frac{1}{2}}).$$

PROOF. Let f be the density of $Z = (N/\lambda(1-\lambda))^{\frac{1}{2}}(U_{j:N} - \lambda)$. Application of Stirling's formula in the form $\log n! = (n + \frac{1}{2})\log(n+1) - (n+1) + \frac{1}{2}\log 2\pi + O(n^{-1})$ followed by expansion of logarithms yields

$$\log f(z) = -\tfrac{1}{2}\log 2\pi + \frac{2\lambda - 1}{(N\lambda(1-\lambda))^{\frac{1}{2}}}z - \frac{1}{2}\left[1 - \frac{\lambda^3 + (1-\lambda)^3}{N\lambda(1-\lambda)}\right]z^2$$

$$+ O\left(\frac{|z|^3}{(N\lambda(1-\lambda))^{\frac{1}{2}}} + \frac{1}{N\lambda(1-\lambda)}\right)$$

for $z^2 < N\min(\lambda/(1-\lambda), (1-\lambda)/\lambda)$. Hence, for $|z| \leqq (N\lambda(1-\lambda))^{\frac{1}{6}} < [N\min(\lambda/(1-\lambda), (1-\lambda)/\lambda)]^{\frac{1}{2}}$,

$$f(z) = \frac{1}{(2\pi)^{\frac{1}{2}}}e^{-\frac{1}{2}z^2}\left[1 + O\left(\frac{|z| + |z|^3}{(N\lambda(1-\lambda))^{\frac{1}{2}}} + \frac{1}{N\lambda(1-\lambda)}\right)\right] \qquad \text{(A2.2)}$$

uniformly in j. Since $\eta(N/\lambda(1-\lambda))^{\frac{1}{2}} \geqq \frac{1}{2}(N\lambda(1-\lambda))^{\frac{1}{2}}$, (A2.2) and Lemma A2.1 imply that

$$EZ^k I_{(\lambda,\lambda+\eta)}(U_{j:N}) = \frac{1}{(2\pi)^{\frac{1}{2}}}\int_0^{\frac{1}{2}(N\lambda(1-\lambda))^{\frac{1}{6}}} z^k e^{-\frac{1}{2}z^2}\left[1 + O\left(\frac{1 + |z| + |z|^3}{(N\lambda(1-\lambda))^{\frac{1}{2}}}\right)\right]dz$$

$$+ O\left(\int_{\frac{1}{2}(N\lambda(1-\lambda))^{\frac{1}{6}}}^{\infty} z^k e^{-\frac{1}{2}z}\,dz\right) = \tfrac{1}{2}\nu_k + O((N\lambda(1-\lambda))^{-\frac{1}{2}}),$$

which proves the second part of the lemma. The first part now follows by noting that $U_{j:N}$ and $1 - U_{N-j+1:N}$ have the same distribution. □

REMARK. One easily verifies that Lemma A2.2 continues to hold when η is taken as small as $[c(\lambda(1-\lambda)/N)|\log N\lambda(1-\lambda)|]^{\frac{1}{2}}$ for any $c > 1$. It should also be noted that when j or $(N-j+1)$ remains bounded as $N \to \infty$, Lemma A2.2 merely states that $E|U_{j:N} - \lambda|^k = O(N^{-k})$.

Condition R_r. For real $r > 0$, a function h on (0, 1) is said to satisfy condition R_r if h is twice continuously differentiable on (0, 1) and

$$\limsup_{t\to 0,1} t(1-t)\left|\frac{h''(t)}{h'(t)}\right| < 1 + \frac{1}{r}\,.$$

LEMMA A2.3. *Let $r_1, \cdots, r_m, k_1, \cdots, k_m$ be positive real numbers, $j = 1, \cdots, N$, $\lambda = j/(N+1)$ and ν_k the* kth *absolute moment of the standard normal distribution. Suppose that $h_1, \cdots, h_m$ satisfy conditions $R_{r_1}, \cdots, R_{r_m}$ respectively and that $\sum k_i/r_i \leqq 1$. Define*

$$M = \left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}\sum k_i}\left\{\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}} + (N\lambda(1-\lambda))^{-\frac{1}{2}}\prod_{i=1}^m |h_i'(\lambda)|^{k_i}\right\}.$$

Then, uniformly in j, we have for $N \to \infty$

$$E\prod_{i=1}^m |h_i(U_{j:N}) - h_i(\lambda)|^{k_i} = \left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}\sum k_i}\nu_{\sum k_i}\prod_{i=1}^m |h_i'(\lambda)|^{k_i} + O(M)$$

and for integer $k_1, \cdots, k_m$

$$\begin{aligned} &E\prod_{i=1}^m (h_i(U_{j:N}) - h_i(\lambda))^{k_i} \\ &\quad = O(M) \qquad \text{if } \textstyle\sum k_i \text{ is odd,} \\ &\quad = \left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}\sum k_i}\nu_{\sum k_i}\prod_{i=1}^m (h_i'(\lambda))^{k_i} + O(M) \qquad \text{if } \textstyle\sum k_i \text{ is even.}\end{aligned}$$

PROOF. For reasons of symmetry it is sufficient to consider only $j \leqq (N+1)/2$, i.e. $\lambda \leqq \frac{1}{2}$. Since h_i satisfies condition R_{r_i}, there exists $0 < \varepsilon < \frac{1}{6}$, $\tau > 1$ and $C > 0$ such that for $i = 1, \cdots, m$

$$\left|\frac{h_i''(t)}{h_i'(t)}\right| \leqq \left(1 + \frac{1}{r_i\tau}\right)t^{-1} \qquad \text{for} \quad 0 < t \leqq 3\varepsilon\,, \tag{A2.3}$$

$$|h_i''(t)| \leqq C \qquad \text{for} \quad \varepsilon \leqq t \leqq 1-\varepsilon\,, \tag{A2.4}$$

$$\left|\frac{h_i''(t)}{h_i'(t)}\right| \leqq \left(1 + \frac{1}{r_i\tau}\right)(1-t)^{-1} \qquad \text{for} \quad 1-3\varepsilon \leqq t < 1\,. \tag{A2.5}$$

Suppose first that $\lambda \leqq 2\varepsilon$. Integration of (A2.3) shows that for $0 < t \leqq \lambda$ and $i = 1, \cdots, m$,

$$\left(\frac{t}{\lambda}\right)^{1+1/r_i\tau} \leqq \frac{h_i'(t)}{h_i'(\lambda)} \leqq \left(\frac{\lambda}{t}\right)^{1+1/r_i\tau},$$

$$\frac{r_i\tau}{2r_i\tau+1}\lambda\left[1 - \left(\frac{t}{\lambda}\right)^{2+1/r_i\tau}\right] \leqq \frac{h_i(\lambda) - h_i(t)}{h_i'(\lambda)} \leqq r_i\tau\lambda\left[\left(\frac{\lambda}{t}\right)^{1/r_i\tau} - 1\right].$$

It follows that

$$\frac{h_i(\lambda) - h_i(t)}{h_i'(\lambda)} = (\lambda - t) + O\left(\frac{(\lambda-t)^2}{\lambda}\right) \qquad \text{for} \quad \tfrac{1}{2}\lambda \leqq t \leqq \lambda\,, \tag{A2.6}$$

$$\left|\frac{h_i(\lambda) - h_i(t)}{h_i'(\lambda)}\right| \leqq r_i\tau\lambda\left(\frac{\lambda}{t}\right)^{1/r_i\tau} \qquad \text{for} \quad 0 < t \leqq \tfrac{1}{2}\lambda\,.$$

Application of Lemma A2.2 with $\eta = \frac{1}{2}\lambda$ yields

$$\begin{aligned} &E \prod_{i=1}^m \left(\frac{h_i(\lambda) - h_i(U_{j:N})}{h_i'(\lambda)}\right)^{k_i} I_{(0,\lambda)}(U_{j:N}) \\ \text{(A2.7)} \quad &= \frac{1}{2}\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}\Sigma k_i} \nu_{\Sigma k_i}[1 + O((N\lambda(1-\lambda))^{-\frac{1}{2}})] \\ &\quad + O\left(\lambda^{\Sigma k_i} E\left(\frac{\lambda}{U_{j:N}}\right)^{1/\tau} I_{(0,\frac{1}{2}\lambda)}(U_{j:N})\right), \end{aligned}$$

where we have made use of $\sum k_i/r_i \leqq 1$. For $2 \leqq j \leqq \frac{1}{2}(N+1)$,

$$\begin{aligned} \lambda^{\Sigma k_i} E\left(\frac{\lambda}{U_{j:N}}\right)^{1/\tau} I_{(0,\frac{1}{2}\lambda)}(U_{j:N}) &= \lambda^{\Sigma k_i + 1/\tau} \frac{N}{j-1} EU_{j-1:N-1}^{1-1/\tau} I_{(0,\frac{1}{2}\lambda)}(U_{j-1:N-1}) \\ \text{(A2.8)} \quad &\leqq 2\lambda^{\Sigma k_i} P(U_{j-1:N-1} < \tfrac{1}{2}\lambda) \\ &= O\left(\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}\Sigma k_i} (N\lambda(1-\lambda))^{-\frac{1}{2}}\right) \end{aligned}$$

by Lemma A2.1. For $j = 1$ we have

$$\begin{aligned} &\lambda^{\Sigma k_i} E\left(\frac{\lambda}{U_{j:N}}\right)^{1/\tau} I_{(0,\frac{1}{2}\lambda)}(U_{j:N}) \\ \text{(A2.9)} \quad &= (N+1)^{-\Sigma k_i - 1/\tau} N \int_0^{1/2(N+1)} u^{-1/\tau}(1-u)^{N-1}\, du \\ &= O(N^{-\Sigma k_i}) = O\left(\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}\Sigma k_i} (N\lambda(1-\lambda))^{-\frac{1}{2}}\right). \end{aligned}$$

Together, (A2.8) and (A2.9) ensure that the second remainder term in (A2.7) may be omitted.

A similar analysis based on (A2.3)—(A2.5) shows that for $\lambda \leqq 2\varepsilon$ but $t \geqq \lambda$, (A2.6) holds for $\lambda \leqq t \leqq 3\lambda/2$ and

$$\begin{aligned} \left|\frac{h_i(t) - h_i(\lambda)}{h_i'(\lambda)}\right| &\leqq r_i \tau \lambda \left(\frac{t}{\lambda}\right)^{2 + 1/r_i\tau} \quad \text{for} \quad \frac{3\lambda}{2} \leqq t \leqq 3\varepsilon, \\ &= O(\lambda^{-1-1/r_i\tau}(1-t)^{-1/r_i\tau}) \quad \text{for} \quad 3\varepsilon \leqq t < 1. \end{aligned}$$

Hence by Lemmas A2.2 and A2.1 and a change from $U_{j:N}$ to $U_{j:N-1}$ as in (A2.8),

$$\begin{aligned} &E \prod_{i=1}^m \left(\frac{h_i(U_{j:N}) - h_i(\lambda)}{h_i'(\lambda)}\right)^{k_i} I_{(\lambda,1)}(U_{j:N}) \\ &= \frac{1}{2}\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}\Sigma k_i} \nu_{\Sigma k_i}[1 + O((N\lambda(1-\lambda))^{-\frac{1}{2}})] \\ \text{(A2.10)} \quad &\quad + O(\lambda^{\Sigma k_i} \exp\{-\tfrac{1}{4}(N\lambda)^{\frac{1}{2}}\} \\ &\quad + \lambda^{-\Sigma k_i - 1/\tau} E(1 - U_{j:N-1})^{1-1/\tau} I_{(3\varepsilon,1)}(U_{j:N-1})) \\ &= \frac{1}{2}\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}\Sigma k_i} \nu_{\Sigma k_i}[1 + O((N\lambda(1-\lambda))^{-\frac{1}{2}})]. \end{aligned}$$

Combining (A2.7)—(A2.10) and noting that (A2.7) and (A2.10) remain valid when absolute values are taken inside the expectation signs, we see that the lemma is proved for $\lambda \leqq 2\varepsilon$.

If $2\varepsilon < \lambda \leqq \frac{1}{2}$, (A2.3)—(A2.5) imply that

$$h_i(t) - h_i(\lambda) = h_i'(\lambda)(t - \lambda) + O((t - \lambda)^2) \qquad \text{for} \quad \varepsilon \leqq t \leqq 1 - \varepsilon\,,$$
$$|h_i(t) - h_i(\lambda)| = O((t(1 - t))^{-1/r_i\tau}) \qquad \text{for} \quad t < \varepsilon \quad \text{or} \quad t > 1 - \varepsilon\,,$$

and the proof of the lemma for $2\varepsilon < \lambda \leqq \frac{1}{2}$ follows by noting that $h_i'(\lambda)$ is bounded and arguing as e.g. in (A2.10). □

REMARK. Although the remainder M in Lemma A2.3 consists of two terms, only one of these plays a role for any particular value of λ. For $2\varepsilon < \lambda < 1 - 2\varepsilon$, $h_i'(\lambda)$ and $(\lambda(1 - \lambda))^{-1}$ are bounded and we need only retain the first term of M. It follows from (A2.7)—(A2.10) that for $\lambda \leqq 2\varepsilon$ or $\lambda \geqq 1 - 2\varepsilon$ only the second term of M is needed.

LEMMA A2.4. *Lemma* A2.3 *continues to hold for central moments*, i.e. *if* $h_i(\lambda)$ *is replaced by* $Eh_i(U_{j:N})$ *for* $i = 1, \cdots, m$, *provided only that* $r_i \geqq 1$ *for* $i = 1, \cdots, m$.

PROOF. As $r_i \geqq 1$, Lemma A2.3 contains as a special case

$$\text{(A2.11)} \qquad |Eh_i(U_{j:N}) - h_i(\lambda)| = O\left(\frac{\lambda(1 - \lambda) + |h_i'(\lambda)|}{N}\right).$$

The lemma is proved by expanding the central moments in terms of moments centered at the $h_i(\lambda)$ and applying (A2.11), Lemma A2.3 and the remark following it. □

We also note the following extension of a result of Hoeffding (1953).

LEMMA A2.5. *Let* $h_1, \cdots, h_m$ *be continuous functions on* $(0, 1)$, q *a continuous function on* R^m *and* Q *a convex function on* R^m *such that* $|q| \leqq Q$. *Suppose that* $\int_0^1 |h_i(t)|\, dt < \infty$ *for* $i = 1, \cdots, m$ *and that* $\int_0^1 Q(h_1(t), \cdots, h_m(t))\, dt < \infty$. *Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^N q(Eh_1(U_{j:N}), \cdots, Eh_m(U_{j:N})) = \int_0^1 q(h_1(t), \cdots, h_m(t))\, dt\,.$$

PROOF. Because h_i is continuous and summable, Lemma 2.2 of Bickel (1967) implies that for any $\varepsilon > 0$, $Eh_i(U_{j_N:N}) - h_i(j_N(N + 1)^{-1}) \to 0$ uniformly for $\varepsilon \leqq j_N(N + 1)^{-1} \leqq 1 - \varepsilon$ as $N \to \infty$. Since q is continuous and $q(h_1, \cdots, h_m)$ is summable, the lemma is proved if we show that

$$\lim_{\varepsilon\downarrow 0} \limsup_N \frac{1}{N} \left(\sum_{j=1}^{[\varepsilon(N+1)]} + \sum_{j=[(1-\varepsilon)(N+1)]}^{N}\right)|q(Eh_1(U_{j:N}), \cdots, Eh_m(U_{j:N}))| = 0\,.$$

It is obviously sufficient to prove this for Q instead of q, but as Q has the same properties as q and is moreover nonnegative, this is equivalent to showing that

$$\limsup_N \frac{1}{N} \sum_{j=1}^N Q(Eh_1(U_{j:N}), \cdots, Eh_m(U_{j:N})) \leqq \int_0^1 Q(h_1(t), \cdots, h_m(t))\, dt\,.$$

As Q is convex this follows from Jensen's inequality. □

LEMMA A2.6. *Let* $k_1, \cdots, k_m$ *be positive integers and* $r_1, \cdots, r_m$ *positive real numbers such that* $\sum k_i/r_i \leqq 1$. *Suppose that* $h_1, \cdots, h_m$ *are continuous functions* *on* $(0, 1)$ *for which* $\int_0^1 |h_i(t)|^{r_i}\, dt < \infty$ *for* $i = 1, \cdots, m$. *Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^N \prod_{i=1}^m (Eh_i(U_{j:N}))^{k_i} = \int_0^1 \prod_{i=1}^m (h_i(t))^{k_i}\, dt\, .$$

If, in addition, h_1 *is monotone in neighborhoods of* 0 *and* 1, *then also*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^N \left(h_1\left(\frac{j}{N+1}\right)\right)^{k_1} \prod_{i=2}^m (Eh_i(U_{j:N}))^{k_i} = \int_0^1 \prod_{i=1}^m (h_i(t))^{k_i}\, dt\, .$$

PROOF. The first part of the lemma is a special case of Lemma A2.5, obtained by taking $q(x_1, \cdots, x_m) = \prod x_i^{k_i}$ and $Q(x_1, \cdots, x_m) = 1 + \sum |x_i|^{r_i}$. To establish the second part we follow the proof of Lemma A2.5 for these choices of q and Q but with $Eh_1(U_{j:N})$ replaced by $h_1(j(N+1)^{-1})$, until we arrive at the point where it suffices to show that

$$\limsup_N \frac{1}{N} \sum_{j=1}^N \left[\left|h_1\left(\frac{j}{N+1}\right)\right|^{r_1} + \sum_{i=2}^m |Eh_i(U_{j:N})|^{r_i}\right] \leqq \int_0^1 \sum_{i=1}^m |h_i(t)|^{r_i}\, dt\, .$$

As $|h_1|^{r_1}$ is continuous and summable, its monotonicity near 0 and 1 amply guarantees that $N^{-1} \sum |h_1(j(N+1)^{-1})|^{r_1} \to \int_0^1 |h_1(t)|^{r_1}\, dt$. Application of Jensen's inequality to the remaining terms completes the proof. □

We now state the results needed to prove Theorems 3.2, 4.1 and 4.2 in the form of two corollaries.

COROLLARY A2.1. *Suppose that positive numbers* C *and* δ *exist such that* $|h'(t)| \leqq C(t(1-t))^{-\frac{4}{3}+\delta}$ *for all* $0 < t < 1$. *Then there exists* $A > 0$ *depending on* N *and* h *only through* C *and* δ *and such that*

$$\sum_{j=1}^N \{E|h(U_{j:N}) - Eh(U_{j:N})|^3\}^{\frac{4}{9}} \leqq AN^{\frac{4}{9}}\, .$$

The above condition is fulfilled if h *satisfies condition* R_1 *and* $\int_0^1 h^6(t)\, dt < \infty$.

PROOF. Define $\lambda = j/(N+1)$. For all $0 < t < 1$, $|h(t) - h(\lambda)|$ is maximized by taking $h'(t) \equiv C(t(1-t))^{-\frac{4}{3}+\delta}$ and for this particular choice of h' the function h satisfies condition R_3. Hence, by Lemma A2.3, we have in general

$$E|h(U_{j:N}) - h(\lambda)|^k = O\left(\left(\frac{\lambda(1-\lambda)}{N}\right)^{\frac{1}{2}k} (\lambda(1-\lambda))^{-k(\frac{4}{3}-\delta)}\right)$$

for $0 < k \leqq 3$. It follows that

$$\begin{aligned}\sum_{j=1}^N \{E|h(U_{j:N}) - Eh(U_{j:N})|^3\}^{\frac{4}{9}} &= O(\textstyle\sum_{j=1}^N \{N^{-\frac{3}{2}}(\lambda(1-\lambda))^{-\frac{5}{2}}\}^{\frac{4}{9}}) \\ &= O(N^{\frac{1}{3}} \textstyle\int_{1/N}^{1-1/N} (t(1-t))^{-\frac{10}{9}}\, dt) = O(N^{\frac{4}{9}})\, .\end{aligned}$$

Condition R_1 ensures that for ε as in (A2.3) and $0 < t < \frac{1}{2}u < \varepsilon$, $|h(t) - h(2\varepsilon)| \geqq \frac{1}{4}u|h'(u)|$ and hence for $u \to 0$,

$$u^7(h'(u))^6 \leqq 2^{13} \int_0^{\frac{1}{2}u} (h(t) - h(2\varepsilon))^6\, dt \to 0\, .$$

In the same way one shows that $|h'(u)| = o((1-u)^{-\frac{7}{6}})$ for $u \to 1$, which completes the proof. □

For $i = 1, 2, 3$, let $\psi_i = f^{(i)}/f$ and $\Psi_i(t) = \psi_i(F^{-1}((1+t)/2))$ as in (3.1) and (3.15). Let J be a function on (0, 1).

COROLLARY A2.2. *Suppose that* (3.2) *holds, that* $0 < \int_0^1 J^4(t)\,dt < \infty$ *and that both* J *and* Ψ_1 *satisfy condition* R_2. *Let either* $a_j = a_{j,N} = EJ(U_{j:N})$ *for* $j = 1, \cdots, N$ *or* $a_j = a_{j,N} = J(j/(N+1))$ *for* $j = 1, \cdots, N$. *Then, as* $N \to \infty$,

$$\frac{1}{N}\sum_{j=1}^N a_j^2 = \int_0^1 J^2(t)\,dt + o(1)\,, \tag{A2.12}$$

$$\frac{1}{N}\sum_{j=1}^N a_j^k E\Psi_1^{4-k}(U_{j:N}) = \int_0^1 J^k(t)\Psi_1^{4-k}(t)\,dt + o(1)\,, \qquad k = 1, \cdots, 4\,, \tag{A2.13}$$

$$\frac{1}{N}\sum_{j=1}^N a_j E\Psi_1(U_{j:N})\Psi_2(U_{j:N}) = \int_0^1 J(t)\Psi_1(t)\Psi_2(t)\,dt + o(1)\,, \tag{A2.14}$$

$$\frac{1}{N}\sum_{j=1}^N a_j E\Psi_3(U_{j:N}) = \int_0^1 J(t)\Psi_3(t)\,dt + o(1)\,, \tag{A2.15}$$

$$\begin{aligned}\frac{1}{N}&\sigma^2\Big(\sum_{j=1}^N a_j \Psi_1(U_{j:N})\Big)\\ &= \int_0^1\int_0^1 J(s)J(t)\Psi_1'(s)\Psi_1'(t)[s\wedge t - st]\,ds\,dt + o(1)\,.\end{aligned} \tag{A2.16}$$

If $a_j = EJ(U_{j:N})$ *for* $j = 1, \cdots, N$, *then also*

$$\begin{aligned} N^{-\frac12}&\frac{\sum_{j=1}^N a_j E\Psi_1(U_{j:N})}{(\sum_{j=1}^N a_j^2)^{\frac12}}\\ &= \frac{\int_0^1 J(t)\Psi_1(t)\,dt}{(\int_0^1 J^2(t)\,dt)^{\frac12}} - \frac{1}{N}\frac{\sum_{j=1}^N \operatorname{Cov}(J(U_{j:N}), \Psi_1(U_{j:N}))}{(\int_0^1 J^2(t)\,dt)^{\frac12}}\\ &\quad + \frac{1}{2N}\frac{\int_0^1 J(t)\Psi_1(t)\,dt}{(\int_0^1 J^2(t)\,dt)^{\frac32}}\sum_{j=1}^N \sigma^2(J(U_{j:N})) + o(N^{-1})\\ &= \frac{\int_0^1 J(t)\Psi_1(t)\,dt}{(\int_0^1 J^2(t)\,dt)^{\frac12}} - \frac{1}{N}\frac{\int_{1/N}^{1-1/N} J'(t)\Psi_1'(t)t(1-t)\,dt}{(\int_0^1 J^2(t)\,dt)^{\frac12}}\\ &\quad + \frac{1}{2N}\frac{\int_0^1 J(t)\Psi_1(t)\,dt}{(\int_0^1 J^2(t)\,dt)^{\frac32}}\int_{1/N}^{1-1/N}(J'(t))^2 t(1-t)\,dt + o(N^{-1})\\ &\quad + O\Big(N^{-\frac32}\int_{1/N}^{1-1/N}|J'(t)|(|J'(t)| + |\Psi_1'(t)|)(t(1-t))^{\frac12}\,dt\Big)\,.\end{aligned} \tag{A2.17}$$

If $J = -\Psi_1$ *and either* $a_j = -E\Psi_1(U_{j:N})$ *for* $j = 1, \cdots, N$ *or* $a_j = -\Psi_1(j/(N+1))$ *for* $j = 1, \cdots, N$, *then*

$$\begin{aligned} N^{-\frac12}&\frac{\sum_{j=1}^N a_j E\Psi_1(U_{j:N})}{(\sum_{j=1}^N a_j^2)^{\frac12}}\\ &= -\Big(\int_0^1 \Psi_1^2(t)\,dt\Big)^{\frac12} + \frac{\int_{1/N}^{1-1/N}(\Psi_1'(t))^2 t(1-t)\,dt}{2N(\int_0^1 \Psi_1^2(t)\,dt)^{\frac12}}\\ &\quad + o(N^{-1}) + O\Big(N^{-\frac32}\int_{1/N}^{1-1/N}(\Psi_1'(t))^2(t(1-t))^{\frac12}\,dt\Big)\,.\end{aligned} \tag{A2.18}$$

PROOF. The assumptions imply that Ψ_1, Ψ_2, Ψ_3 and J are continuous, that $\Psi_1{}^6$, $\Psi_2{}^3$, $|\Psi_3|^{\frac{4}{3}}$ and J^4 are summable and that J is monotone near 0 and 1. Hence (A2.12)—(A2.15) follow from Lemma A2.6.

For $a_j = J(j/(N+1))$ a proof of (A2.16) is essentially contained in Stigler (1969). Our condition R_2 for Ψ_1 ensures that $\Psi_1{}'$ will satisfy Stigler's condition T at 0 and 1. As in the proof of Corollary A2.1, one can argue that near 0 and 1

$$\Psi_1{}'(t) = o((t(1-t))^{-\frac{7}{6}}) , \qquad J'(t) = o((t(1-t))^{-\frac{5}{4}}) . \tag{A2.19}$$

Inspection of Stigler's conditions for (A2.16) shows that in our case the only missing ingredient is that Ψ_1 is not necessarily increasing on (0, 1). However, Ψ_1 is monotone where it matters, that is in a neighborhood of 0 and 1.

To prove that (A2.16) remains valid for $a_j = EJ(U_{j:N})$ we note that by Lemma A2.4 and (A2.19)

$$\begin{aligned}\sigma^2\left(\textstyle\sum_{j=1}^N \left(EJ(U_{j:N}) - J\left(\frac{j}{N+1}\right)\right) \Psi_1(U_{j:N})\right) & \\ \leqq \left[\textstyle\sum_{j=1}^N \left|EJ(U_{j:N}) - J\left(\frac{j}{N+1}\right)\right| \sigma(\Psi_1(U_{j:N}))\right]^2 & \\ = o(N^{-1}[\textstyle\int_{1/N}^{1-1/N} (t(1-t))^{-\frac{23}{12}}\, dt]^2) = o(N^{\frac{5}{6}}) . &\end{aligned}$$

For $a_j = EJ(U_{j:N})$ we have

$$\frac{1}{N} \sum_{j=1}^N a_j{}^2 = \int_0^1 J^2(t)\, dt - \frac{1}{N} \sum_{j=1}^N \sigma^2(J(U_{j:N})) , \tag{A2.20}$$

$$\begin{aligned}\frac{1}{N} \sum_{j=1}^N a_j E\Psi_1(U_{j:N}) & \\ = \int_0^1 J(t)\Psi_1(t)\, dt - \frac{1}{N} \sum_{j=1}^N \operatorname{Cov}(J(U_{j:N}), \Psi_1(U_{j:N})) . &\end{aligned} \tag{A2.21}$$

By Lemma A2.4, condition R_2 for J, and (A2.19)

$$\begin{aligned}\frac{1}{N} \sum_{j=1}^N \sigma^2(J(U_{j:N})) & \\ = \frac{1}{N} \int_{1/N}^{1-1/N} (J'(t))^2 t(1-t)\, dt + O(N^{-2} \int_{1/N}^{1-1/N} (J'(t))^2\, dt + N^{-\frac{3}{2}} & \\ + N^{-\frac{3}{2}} \int_{1/N}^{1-1/N} (J'(t))^2 (t(1-t))^{\frac{1}{2}}\, dt) & \\ = \frac{1}{N} \int_{1/N}^{1-1/N} (J'(t))^2 t(1-t)\, dt & \\ + O(N^{-\frac{3}{2}} + N^{-\frac{3}{2}} \int_{1/N}^{1-1/N} (J'(t))^2 (t(1-t))^{\frac{1}{2}}\, dt) = o(N^{-\frac{1}{2}}) . &\end{aligned} \tag{A2.22}$$

Similarly

$$\begin{aligned}\frac{1}{N} \sum_{j=1}^N \operatorname{Cov}(J(U_{j:N}), \Psi_1(U_{j:N})) & \\ = \frac{1}{N} \int_{1/N}^{1-1/N} J'(t)\Psi_1{}'(t) t(1-t)\, dt & \\ + O(N^{-\frac{3}{2}} + N^{-\frac{3}{2}} \int_{1/N}^{1-1/N} |J'(t)\Psi_1{}'(t)|(t(1-t))^{\frac{1}{2}}\, dt) = o(N^{-\frac{7}{12}}) . &\end{aligned} \tag{A2.23}$$

Together (A2.20)—(A2.23) are sufficient to prove (A2.17).

If $J = -\Psi_1$ and $a_j = -E\Psi_1(U_{j:N})$, then (A2.17) reduces to (A2.18). To prove that (A2.18) also holds if $a_j = -\Psi_1(j/(N+1))$, it suffices to show that

$$\begin{aligned}&\textstyle\sum_{j=1}^N \Psi_1\left(\frac{j}{N+1}\right) E\Psi_1(U_{j:N}) \\ \text{(A2.24)}\qquad &- \left[\textstyle\sum_{j=1}^N \Psi_1^2\left(\frac{j}{N+1}\right) \sum_{j=1}^N (E\Psi_1(U_{j:N}))^2\right]^{\frac{1}{2}} \\ &= o(1) + O(N^{-\frac{1}{2}} \textstyle\int_{1/N}^{1-1/N} (\Psi_1'(t))^2(t(1-t))^{\frac{1}{2}}\, dt)\,.\end{aligned}$$

It follows from Lemma A2.3 and condition R_2 for Ψ_1 that

$$\begin{aligned}\textstyle\sum_{j=1}^N \left\{E\Psi_1(U_{j:N}) - \Psi_1\left(\frac{j}{N+1}\right)\right\}^2 &= O(N^{-1} + N^{-1} \textstyle\int_{1/N}^{1-1/N} (\Psi_1'(t)^2\, dt) \\ &= O(N^{-1} + N^{-\frac{1}{2}} \textstyle\int_{1/N}^{1-1/N} (\Psi_1'(t))^2(t(1-t))^{\frac{1}{2}}\, dt)\,,\end{aligned}$$

which suffices to establish (A2.24) and complete the proof. □

REFERENCES

ALBERS, W. (1974). Asymptotic expansions and the deficiency concept in statistics. *Mathematical Centre Tracts* **58**, Amsterdam.

BICKEL, P. J. (1967). Some contributions to the theory of order statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 575–591. Univ. of California Press.

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

BICKEL, P. J. and VAN ZWET, W. R. (1975). Asymptotic expansions for the power of distribution free tests in the two-sample problem. In preparation.

CHUNG, K. L. (1951). The strong law of large numbers. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 341–352. Univ. of California Press.

CRAMÉR, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

CRAMÉR, H. (1962). *Random Variables and Probability Distributions*, 2nd ed. Cambridge Univ. Press.

DIEUDONNÉ, J. (1960). *Foundations of Modern Analysis.* Academic Press, New York.

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

HODGES, J. L. and LEHMANN, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.

HODGES, J. L. and LEHMANN, E. L. (1970). Deficiency. *Ann. Math. Statist.* **41** 783–801.

HOEFFDING, W. (1953). On the distribution of the expected values of the order statistics. *Ann. Math. Statist.* **24** 93–100.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

OKAMOTO, M. (1958). Some inequalities relating to the partial sum of binomial probabilities. *Ann. Inst. Statist. Math.* **10** 29–35.

ROGERS, W. F. (1971). Exact null distributions and asymptotic expansions for rank test statistics. Technical Report No. 145, Departments of O.R. and Statistics, Stanford Univ.

STIGLER, S. M. (1969). Linear functions of order statistics. *Ann. Math. Statist.* **40** 770–788.

TITCHMARSH, E. C. (1939). *The Theory of Functions*, 2nd ed. Oxford Univ. Press.

W. ALBERS AND P. J. BICKEL
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

W. R. VAN ZWET
CENTRAAL REKEN-INSTITUUT
DER RIJKSUNIVERSITEIT TE LEIDEN
WASSENAARSEWEG 80
POSTBUS 2060
LEIDEN, THE NETHERLANDS

The Annals of Statistics
1974, Vol. 2, No. 1, 1-20

SPECIAL INVITED PAPER

EDGEWORTH EXPANSIONS IN NONPARAMETRIC STATISTICS[1]

BY P. J. BICKEL

University of California, Berkeley

This is a survey of recent work on Edgeworth expansions for (M) estimates, rank tests and some other statistics arising in nonparametric models. A Berry-Esséen theorem for U-statistics which seems to be new is also proved.

1. Introduction. During the past 25 years various procedures which are not sensitive to certain departures from normality have been evolved and investigated. The study of such methods is loosely referred to as nonparametric statistics. One broad category of such procedures is that of the distribution free tests such as the permutation t test, the rank tests of Wilcoxon, Kruskal-Wallis, Spearman and Kendall, and the omnibus tests such as the two sample Smirnov test. All of these are discussed in the monograph of Hájek and Šidák [26]. Another major category is that of the various robust estimates such as those discussed in the recent Princeton study [2].

Most of the theoretical work done on these procedures has been devoted to obtaining large sample properties by establishing first order limit theorems for the statistics on which these procedures are based. In this paper I intend to discuss what is known about higher order approximations to the distribution of these statistics. In the main I shall limit myself to discussion of results obtained since the general review paper by D. Wallace which appeared in this journal in 1958, [57].

Suppose that we are given a sequence of statistics $\{T_N\}$, $N \geqq 1$, where N usually denotes sample size. In accordance with [57] we shall say that the distribution function F_N of T_N possesses an asymptotic expansion valid to $(r+1)$ terms if there exist functions $A_0, \cdots, A_r$ such that

$$(1.1) \qquad \left|F_N(x) - A_0(x) - \sum_{j=1}^{r} \frac{A_j(x)}{N^{j/2}}\right| = o(N^{-r/2})\,.$$

If,

$$(1.2) \qquad \sup_x \left|F_N(x) - A_0(x) - \sum_{j=1}^{r} \frac{A_j(x)}{N^{j/2}}\right| = o(N^{-r/2})$$

Received January 1973; revised May 1973.

[1] This research was supported by the Office of Naval Research, Contract N00014-69-A-0200-1038.

AMS 1970 *subject classifications.* Primary 62G05, 10, 20, 30, 35; Secondary 60F05.

Key words and phrases. Edgeworth, Cornish-Fisher, expansions, Berry-Esseen bounds, rank tests, (M) estimates, goodness of fit tests, U-statistics, deficiency.

we shall say the expansion is uniformly valid to $(r + 1)$ terms. (This is not quite in accord with Wallace who requires the remainder to be $O(N^{-(r+1)/2})$ but is more convenient and in accord with [19].) An expansion valid to one term is just an ordinary limit theorem. It is sometimes convenient to consider expansions in which the A_j also depend on N. They are then, of course, no longer uniquely defined.

These higher order terms are of interest on various grounds.

(1) Taking one or two terms of the expansion frequently improves the basic approximation A_0 strikingly. Examples of this phenomenon may be found in Hodges and Fix [28] and Thompson, Govindarajulu and Doksum [55].

(2) The higher order terms give some qualitative insight into regions of unreliability of first order results. For instance, when the limit A_0 is normal the higher order terms A_1 and A_2 typically correct for skewness and kurtosis.

(3) The expansions can be used to discriminate between procedures equivalent to first order, as for example in Hodges and Lehmann's work on deficiency [30].

(4) Last but not least the probabilistic problems involved are very challenging.

Expansions of the type (1.1) and (1.2) are not the only ones of interest. Density functions and frequency functions of lattice random variables can sometimes be expanded. Extreme and intermediate tail probabilities can also sometimes be expanded (see for example [21], pages 517–520, [13] and [37]), and as P. Huber pointed out to me, the approximation to the power function of tests so obtained can be much more satisfactory than that based on the Edgeworth expansion. However, at least to date, the principal method used has been that of saddle point approximation which seems to require more intimate knowledge of the characteristic function of F_N than is usually available. In any case few if any such expansions appear to be available in nonparametric problems. Thus, we limit ourselves to discussion of expansions of types (1.1) ("Edgeworth") and the related expansions of F_N^{-1} ("Cornish–Fisher"). We shall deal primarily with expansions in which A_0 is the normal distribution. General results are available here for linear rank statistics (Section 2) and M estimates (Section 3) and partial results for linear combinations of order statistics and U-statistics (Section 4). What is known in nonnormal limiting situations is discussed briefly in Section 5.

2. The Berry–Esséen method and linear rank statistics. Suppose that a sequence $\{T_N\}$, $N \geqq 1$, of random variables tends to a standard normal distribution. If we let

$$\rho_N(t) = E(e^{itT_N}) \tag{2.1}$$

then we are asserting that there is a version of $\log \rho_N$ such that as $N \to \infty$,

$$\log \rho_N(t) \to -\frac{t^2}{2}\,. \tag{2.2}$$

Suppose that we have an asymptotic expansion of $\log \rho_N$ of the form,

$$\log \rho_N(t) = -\frac{t^2}{2} + \frac{P_1(it)}{N^{\frac{1}{2}}} + \cdots + \frac{P_r(it)}{N^{r/2}} + o(N^{-r/2}), \tag{2.3}$$

where the P_j are polynomials of order $\leqq j + 2$ which vanish at 0. Such a development is plausible if the T_N have cumulants $K_{j,N}$, such that $K_{1,N} = 0$, $K_{2,N} = 1$, $K_{j,N} = O(N^{-(j-2)/2})$, $j \geqq 3$, and which themselves admit asymptotic expansions in powers of $N^{-\frac{1}{2}}$. Thus if

$$K_{j,N} = \sum_{l=0}^{r-j+2} \frac{K_j^{(l)}}{N^{(j+l-2)/2}} + o(N^{-r/2}) \tag{2.4}$$

we should have,

$$P_k(it) = \sum_{j=3}^{k+2} \frac{K_j^{(k+2-j)}}{j!} (it)^j . \tag{2.5}$$

This is typically true although it sometimes requires a separate proof. The prototypical such T_N are, of course, standardized sums of independent identically distributed random variables. For more on expansions of the log characteristic function in terms of cumulants we refer the reader to the discussion in [57] and on pages 221–230 of [12]. Now, (2.3) corresponds to

$$\rho_N(t) = e^{-t^2/2}\left(1 + \sum_{j=1}^{r} \frac{Q_j(it)}{N^{j/2}}\right) + o(N^{-r/2}) \tag{2.6}$$

where

$$Q_1(it) = P_1(it)$$

$$Q_2(it) = P_2(it) + \frac{[P_1(it)]^2}{2}$$

and so on.

Normal Fourier inversion suggests that if

$$Q_j(it) = \sum_{k\geq 1} a_{jk}(it)^k$$

then

$$F_N(x) = \Phi(x) - \phi(x)\left[\sum_{j=1}^{r} \frac{1}{N^{j/2}} \sum_{k\geq 1} a_{jk} N_{k-1}(x)\right] + o(N^{-r/2}) \tag{2.7}$$

where Φ is the standard normal cdf, ϕ is the standard normal density and the N_k are Hermite polynomials defined by

$$\frac{d^k\phi(x)}{dx^k} = (-1)^k N_k(x)\phi(x) . \tag{2.8}$$

This formal step cannot, of course, be justified in general. It fails for instance if T_N is the standardized sum of independent identically distributed lattice random variables. The passage *is* valid if the weak (2.6) can be replaced by

$$\int_{-MN^{r/2}}^{MN^{r/2}} \left\{\left|\rho_N(t) - e^{-t^2/2}\left(1 + \sum_{j=1}^{r} \frac{Q_j(it)}{N^{j/2}}\right)\right| \Big/ |t| \right\} dt = o(N^{-r/2}) \tag{2.9}$$

for every $M < \infty$. An equivalent useful form of (2.9) is

$$\int_{-\varepsilon N^{1/2}}^{\varepsilon N^{1/2}} \left\{ \left| \rho_N(t) - e^{-t^2/2}\left(1 + \sum_{j=1}^{r} \frac{Q_j(it)}{N^{j/2}}\right)\right| \Big/ |t| \right\} dt = o(N^{-r/2}) \tag{2.10}$$

and

$$\int_{\{\varepsilon N^{1/2} \leq |t| \leq M N^{r/2}\}} \frac{|\rho_N(t)|}{|t|}\, dt = o(N^{r/2})$$

for some $\varepsilon > 0$ and every $M < \infty$. That (2.9) suffices follows from a famous lemma of Berry and Esséen whose statement and proof may be found in Feller [21], Chapter 16, page 510.

The validity of (2.9) and hence of (2.7) to order $1/N$ ($r = 2$) has been established for linear rank statistics both under the hypothesis of symmetry and under contiguous location alternatives by Albers, Bickel, and van Zwet [1]. A similar expansion for the two sample Wilcoxon statistic under the null hypothesis was established earlier by Rogers [48]. Expansions for general two sample rank statistics to order $1/N$ both under the hypothesis and contiguous location alternatives are in preparation [6]. Here is a selection of the results of these papers.

Let $X_1, \cdots, X_N$ be independent identically distributed with common cdf G and density g. Let $Z_{1:N} < \cdots < Z_{N:N}$ denote the ordered $|X_j|$. Define ranks $R_1, \cdots, R_N$ by

$$|X_{R_j}| = Z_{j:N} .$$

Let

$$\begin{aligned} \varepsilon_j &= 1 \qquad \text{if } X_{R_j} > 0 \\ &= -1 \qquad \text{otherwise}, \end{aligned}$$

and suppose that $a_{1N}, \cdots, a_{NN}$ are given constants.

Define

$$T_N = \sum_{j=1}^{N} \frac{a_{jN}\varepsilon_j}{\sigma_N} \tag{2.11}$$

where

$$\sigma_N^2 = \sum_{j=1}^{N} a_{jN}^2 . \tag{2.12}$$

For simplicity suppose there exists a function J on (0, 1) such that

$$a_{jN} = E(J(U_{j:N})) \tag{2.13}$$

where $U_{1:N} < \cdots < U_{N:N}$ are the order statistics of a sample of size N from the uniform distribution on (0, 1). All of the usual statistics for testing the hypothesis that g is symmetric about 0, including the sign, Wilcoxon and normal scores tests can be put in this form. Hájek and Šidák [26] provide an extensive discussion of these procedures as well as the two sample tests we shall mention.

If g is symmetric about 0 the ε_j are independent with $P[\varepsilon_j = 1] = \frac{1}{2}$. The statistic T_N is then a sum of independent nonidentically distributed random variables, and

$$\rho_N(t) = \prod_{j=1}^{N} \cos \frac{t a_{jN}}{\sigma_N} . \tag{2.14}$$

If $\int_0^1 J^4(t)\,dt < \infty$, Taylor expansion of (2.14) yields

$$\log \rho_N(t) = -\frac{t^2}{2} - 2\frac{(it)^4}{4!}\sum_{j=1}^N \frac{a_{jN}^4}{\sigma_N{}^4} + o\left(\frac{1}{N}\right) \tag{2.15}$$

$$= -\frac{t^2}{2} - \frac{(it)^4}{12N}\frac{\int_0^1 J^4(t)\,dt}{(\int_0^1 J^2(t)\,dt)^2} + o\left(\frac{1}{N}\right).$$

If J is in addition continuously differentiable and nonconstant it is shown in [1] that (2.10) holds and hence that

$$\Phi(x) + \frac{\int_0^1 J^4(t)\,dt}{12N(\int_0^1 J^2(t)\,dt)^2}\,\phi(x)H_3(x)$$

is a uniformly valid expansion for F_N to three terms. In particular this proves the validity of the expansions used by Fellingham and Stoker [22] for the Wilcoxon test and by Thompson *et al.* [55] for the normal scores test up to terms of order smaller than $1/N$. Thompson *et al.* noted that the approximation using exact cumulants suggested by the first identity in (2.15) is better than the expansion suggested by the second identity while Fellingham and Stoker only considered the approximation using exact cumulants, with continuity correction. The exact cumulant Edgeworth expansion in both cases did provide substantial improvement over the normal approximation for $N = 10 - 20$ although the latter seems satisfactory for all practical purposes. It is not yet known whether the Edgeworth expansion for statistics such as the normal scores is valid to more than three terms. It seems clear that the expansion to order $1/N^2$ for the Wilcoxon with continuity correction used by Fellingham and Stoker can be justified by a local limit expansion and application of the Euler-Maclaurin formula. Local limit theorems for the two sample Wilcoxon statistic were developed by Rogers [48].

If g is not symmetric about 0 the ε_j are no longer independent. However by conditioning on $|X_1|, \cdots, |X_N|$ Albers, Bickel and van Zwet arrive at the following representation for ρ_N,

$$\rho_N(t) = E\{\textstyle\prod_{j=1}^N [P_{jN}\exp[ita_{jN}/\sigma_N] + (1 - P_{jN})\exp[-ita_{jN}/\sigma_N]]\} \tag{2.16}$$

where

$$P_{jN} = \frac{g(Z_{j:N})}{g(Z_{j:N}) + g(-Z_{j:N})}.$$

From this representation it may be shown that if $\int_0^1 J^4(t)\,dt < \infty$ and J is continuously differentiable and nonconstant then

$$\int_{-bN^{\frac{1}{2}}}^{bN^{\frac{1}{2}}} \{|\rho_N(t) - \tilde{\rho}_N(t)|/|t|\}\,dt \leqq cN^{-\frac{3}{2}}$$

for b, c depending on g where

$$\tilde{\rho}_N(t) = E\left\{\exp\left[itK_{1N} - \frac{t^2}{2}K_{2N}\right]\left(1 + \frac{(it)^3}{6}K_{3N} + \frac{(it)^4}{24}K_{4N} + \frac{(it)^6}{72}K_{3N}^2\right)\right\} \tag{2.17}$$

and

$$K_{1N} = \sum_{j=1}^{N} \frac{a_{jN}}{\sigma_N}(2P_{jN} - 1)$$

$$K_{2N} = 4\sum_{j=1}^{N} \frac{a_{jN}^3}{\sigma_N^2} P_{jN}(1 - P_{jN})$$

$$K_{3N} = 8\sum_{j=1}^{N} \frac{a_{jN}^3}{\sigma_N^3} P_{jN}(1 - P_{jN})(1 - 2P_{jN})$$

$$K_{4N} = 16\sum_{j=1}^{N} \frac{a_{jN}^4}{\sigma_N^4} P_{jN}(1 - P_{jN})(1 - 6P_{jN} + 6P_{jN}^2)$$

are the cumulants of T_N.

Further expansion for fixed alternatives appears to depend on the development of the theory of Edgeworth expansion for linear combinations of order statistics. However, if we permit g to depend on N in such a way that g is contiguous to a symmetric density, then K_{1N} is to first order a constant, and further expansion is possible. Specifically suppose that

$$g_N(x) = f(x - \theta_N) \tag{2.18}$$

where f is a fixed density symmetric about 0 and $\theta_N = \theta/N^{\frac{1}{2}}$. It is then shown in [1] under some regularity conditions on f, as well as the previously specified conditions on J, that for some b, c depending on f and J

$$\int_{-bN^{\frac{3}{2}}}^{bN^{\frac{3}{2}}} \{|\tilde{\rho}_N(t) - \gamma_N(t)|/|t|\}\, dt \leq cN^{-\frac{3}{2}} \tag{2.19}$$

where

$$\gamma_N(t) = \exp\left[it\tilde{K}_{1N} - \frac{t^2}{2}\tilde{K}_{2N}\right]\left(1 + \frac{(it)^3}{6}\tilde{K}_{3N} + \frac{(it)^4}{24}\tilde{K}_{4N}\right) \tag{2.20}$$

and

$$\tilde{K}_{1N} = -\theta_N \sum_{j=1}^{N} \frac{a_{jN}}{\sigma_N} E_0(\psi_1(Z_{j:N}))$$
$$\qquad - \frac{\theta_N^3}{3\sigma_N}\sum_{j=1}^{N} a_{jN} E_0[\tfrac{1}{2}\psi_3(Z_{j:N}) - 3\psi_1\psi_2(Z_{j:N}) + \tfrac{3}{2}\psi_1^3(Z_{j:N})]$$

$$\tilde{K}_{2N} = 1 - \theta_N^2 \sum_{j=1}^{N} \frac{a_{jN}^2}{\sigma_N^2} E_0(\psi_1(Z_{j:N}))^2 + \frac{\theta_N^2}{\sigma_N^2}\operatorname{Var}_0\left(\sum_{j=1}^{N} a_{jN}\psi_1(Z_{j:N})\right)$$

$$\tilde{K}_{3N} = 2\theta_N \sum_{j=1}^{N} \frac{a_{jN}^3}{\sigma_N^3} E_0(\psi_1(Z_{j:N}))$$

$$\tilde{K}_{4N} = -2\sum_{j=1}^{N} \frac{a_{jN}^4}{\sigma_N^4}$$

where

$$\psi_j(x) = \frac{f^{(j)}}{f}(x)$$

and the subscript 0 indicates that calculation is carried out under f. The $\tilde{K}_{jN}$ may be shown to be the leading terms in the expansion of the cumulants of T_N

under g_N. Berry's lemma can be applied to yield as a uniformly valid expansion for $F_N(t)$ to three terms

$$\Phi(y_N) - \phi(y_N)\left\{\frac{\tilde{K}_{3N}}{6} N_2(y_N) + \frac{\tilde{K}_{4N}}{24} N_3(y_N)\right\} \tag{2.21}$$

where

$$y_N = \frac{t - \tilde{K}_{1N}}{(\tilde{K}_{2N})^{\frac{1}{2}}} .$$

This is not strictly speaking an expansion of the type we have been considering since N enters into the approximation in a complicated fashion. However, the expansion can be used in this form, for instance, to study power under normal alternatives since in this case

$$\psi_j(x) = (-1)^j H_j(x)$$

and moments of order statistics from the half normal distribution are available (cf. [34]).

If J' is defined and continuous on [0, 1] and f satisfies some mild regularity conditions, integral approximations to the $\tilde{K}_{jN}$ can be shown to hold, and a uniformly valid expansion to three terms as defined in Section 1 can be provided. This is adequate for the Wilcoxon but not the normal scores test. If we consider the distribution of the latter under normal alternatives it turns out that the $\tilde{K}_{1N}$ term does not admit an expansion of the form $A + B/N$ with A, B fixed, but rather requires a term of the form $(B \log\log N)/N$. As noted by Wallace, expansions of the type (2.7) can validly be inverted to yield expansions for percentiles (Cornish–Fisher) and hence expansions for the power functions of the rank statistics T_N. Agreement between the power function expansions for the normal scores and Wilcoxon tests obtained from (2.21) and (2.15) for normal and logistic alternatives appears to agree well with the Monte Carlo figures of Thompson *et al.* [55]. However, agreement with the Monte Carlo figures of Arnold [3] for the power function of the Wilcoxon test under Cauchy alternatives seems unsatisfactory.

In [30] Hodges and Lehmann introduced the notion of *deficiency* of a procedure with respect to an equally efficient competitor. For tests of equal level α, the deficiency is crudely defined as the limit of the difference in sample sizes required to reach equal power for the same alternative. The power functions expansions obtained in [1] are used to calculate the deficiency of the normal scores test with respect to the t test for normal alternatives. This turns out to be infinite but of the order of $\log\log N$. The results of [1] can also be used to establish that the permutation t test has deficiency 0 with respect to the t test under normal alternatives.

Suppose now that we have two samples $X_1, \cdots, X_m, Y_1, \cdots, Y_n, N = m + n$, the first sample being distributed with common density f, the second with common density g. Let $Z_{1:N} < \cdots < Z_{N:N}$ be the order statistics of the pooled sample and define

$$\begin{aligned} \varepsilon_j &= 1 \quad \text{if} \quad Z_{j:N} = Y_k \quad \text{for some} \quad k \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

A two sample linear rank statistic standardized under the null hypothesis is then given by

$$T_N = \sum_{j=1}^N a_{jN}\left(\varepsilon_j - \frac{n}{N}\right)\Big/\tau_N^2 \tag{2.22}$$

where the a_{jN} are specified scores

$$\tau_N^2 = [\sum_{j=1}^N (a_{jN} - \bar{a}_N)^2] \frac{mn}{N(N-1)} \tag{2.23}$$

and

$$\bar{a}_N = \frac{1}{N} \sum_{j=1}^N a_{jN} .$$

Suppose again that the a_{jN} are given by (2.13). Using a representation of the characteristic function ρ_N of T_N related to one due to Erdös and Rényi [20] and the Berry lemma, Bickel and van Zwet [6] obtain a uniformly valid expansion for the distribution function F_N of T_N to three terms if $f = g$, n/N stays bounded away from 0 and 1, $\int_0^1 J^4(t)\,dt < \infty$, and J is nonconstant and has continuous derivative. In this case,

$$F_N(x) = \Phi(x) - \phi(x)\left\{\frac{K_{3N}^*}{6} H_2(x) + \frac{K_{4N}^*}{24} H_3(x) + \frac{[K_{3N}^*]^2}{72} H_5(x)\right\} + o\left(\frac{1}{N}\right) \tag{2.24}$$

where the K_{jN}^* are the cumulants of T_N. Essentially this result was obtained by Rogers in [48] for the Wilcoxon statistic. Formal expansions were previously considered by Hodges and Fix [28]. A Berry-Esséen bound was obtained by Stoker [53]. Expansions of the power function and deficiency calculations are in progress [6]. Formal expansions of the power function were considered by Witting [58] using moment expansions due to Sundrum [54]. More Monte Carlo studies of the power functions of the two sample tests are desirable. Figures are available for the Savage test [17] when f and g are exponential densities and for the Wilcoxon and normal scores test under normal alternatives [34], [35], [41].

There are several open problems in this area. Two which I find interesting are:

(1) The extension of these results to tests of independence such as Spearman's ρ and Kendall's τ.

(2) The establishment of valid expansions for fixed alternatives.

3. Multivariate Edgeworth expansions and (M) estimates. A significant development in the theory of asymptotic expansions occurred in 1961 with the appearance of Ranga Rao's thesis on Edgeworth expansions and Berry-Esséen bounds for sums of independent random vectors. Since then there has been considerable development in the field. Some results typical of the most recent state of the art and many references to older work may be found in Bhattacharya's paper [5] in which the following theorem is announced.

Let $\{X^{(r)} = (X_1^{(r)}, \cdots, X_k^{(r)})\}$ be a sequence of independent identically distributed k dimensional random vectors. Suppose that

$$(3.1) \qquad E(X_i^{(1)}) = 0\,, \qquad i = 1, \cdots, k$$
$$E(X_i^{(1)}X_j^{(1)}) = \delta_{ij}\,, \qquad 1 \leqq i \leqq j \leqq k\,.$$

Let

$$(3.2) \qquad \rho(u) = E(e^{iuX^{(1)}})$$

where $u = (u_1, \cdots, u_k)$ and $uX^{(1)}$ is the inner product of u and $X^{(1)}$. As usual consider the formal expansion of $\rho^N(u/N^{\frac{1}{2}})e^{|u|^2/2}$ where $|u|^2 = \sum_{i=1}^k u_i^2$, as a power series in $N^{-\frac{1}{2}}$

$$(3.3) \qquad e^{|u|^2/2}\rho^N\left(\frac{u}{N^{\frac{1}{2}}}\right) = 1 + \sum_{j=1}^\infty \frac{P_j(iu)}{N^{j/2}}$$

where the P_j are polynomials whose coefficients depend on the cumulants of $X^{(1)}$. Define polynomials $\tilde{P}_j$ on R^k by the property that $(2\pi)^{-k/2}e^{-|t|^2/2}\tilde{P}_j(t)$ has $e^{-|u|^2/2}P_j(iu)$ as its Fourier transform. For any $A \subset R^k$, let $(\partial A)^\varepsilon$ be the set of all points within a distance ε of the boundary of A, i.e.,

$$(3.4) \qquad (\partial A)^\varepsilon = \{x \in R^k : \exists y \in A,\ z \notin A \ni |x - y| < \varepsilon,\ |x - z| < \varepsilon\}\,.$$

Let $\mathscr{A}(\Phi : d, \varepsilon_0)$ be the class of all Borel sets A such that

$$\Phi((\partial A)^\varepsilon) \leqq d\varepsilon\,, \qquad 0 < \varepsilon \leqq \varepsilon_0$$

where Φ is the standard multivariate normal product probability measure on R^k.

We need Cramér's condition

$$(\mathrm{C}) \qquad \limsup_{|u|\to\infty} |\rho(u)| < 1\,.$$

THEOREM (Remark 1, page 255 of [5]). *Suppose that* $E|X_j^{(1)}|^s < \infty$, $1 \leqq j \leqq k$, *for some* $s \geqq 3$, *the* $X^{(j)}$ *are as above and that condition* (C) *holds. Let* $S_N = \sum_{j=1}^N X^{(j)}$. *Then, for every* $d > 0$,

$$(3.5) \qquad \sup\left\{\left|P\left[\frac{S_N}{N^{\frac{1}{2}}} \in A\right] - (2\pi)^{-k/2}\int \cdots \int_A e^{-|t|^2/2} \times\left[1 + \sum_{j=1}^{s-2}\frac{\tilde{P}_j(t)}{N^{j/2}}\right]dt\right| : A \in \mathscr{A}(\Phi : d, \varepsilon_0)\right\} = o(N^{(s-2)/2})\,.$$

By making a linear transformation of the variables this result can obviously be extended to the case that $X^{(1)}$ has a specified nonsingular covariance matrix. These results have been applied in a variety of problems involving expansions of multivariate distributions connected with normal variables. An interesting paper along these lines which also faces the problem of computation of the $\tilde{P}_j(t)$ is that of Chambers [10].

In this section we review the work of Linnik and Mitrofanova [38], [56] and Čibišov [11] who employed results of this type to obtain asymptotic expansions for maximum likelihood estimates, and the related work of Pfanzagl [45], [46] and Michel and Pfanzagl [40]. The work is of interest from the point of view of robust estimation since the same technique yields expansions for Huber's (M) estimates [32], [33].

Let

$$X_j = \theta + E_j\,, \qquad 1 \leqq j \leqq N$$

where the E_j are independent identically distributed with density f. An (M) estimate (scale known) of θ, for given ψ, is by definition, any solution $\hat{\theta}$ of the equation

$$\sum_{j=1}^N \psi(X_j - \hat{\theta}) = 0\,. \tag{3.6}$$

For the estimation to make sense we suppose

$$E_0(\psi(X_1 - \theta)) = 0\,. \tag{3.7}$$

Condition for consistency and asymptotic normality of such estimates are given in [32] and [33].

Linnik and Mitrofanova [38], in the tradition of Cramér [12], obtained expansions for a solution of (3.6) when $\psi = -f'/f$. It is easy to see in the light of [33] how their conditions should be modified to yield expansions for (M) estimates. It should be noted that [38] has many obscure points and, in particular, it seems to me that the appeal to Ranga Rao's theorem [47] at a crucial point in [38] is inadequate. However, I believe application of the more sophisticated theorem of Bhattacharya that was stated above will carry the proof through.

The main idea which was already used by Haldane and Smith [27] and Shenton and Bowman [9] for formal cumulant expansions of maximum likelihood estimates is to expand the likelihood equation beyond the customary two terms.

$$\begin{aligned} 0 = N^{-\frac{1}{2}} \textstyle\sum_{j=1}^N \psi(X_j - \theta) - \left\{\frac{1}{N} \sum_{j=1}^N \psi'(X_j - \theta)\right\} N^{\frac{1}{2}}(\hat{\theta} - \theta) + \cdots \\ + N^{-(k-1)/2} \frac{(-1)^k}{k!} \left\{\frac{1}{N} \textstyle\sum_{j=1}^N \psi^{(k)}(X_j - \theta)\right\} N^{k/2}(\hat{\theta} - \theta)^k + R_{Nk}\,. \end{aligned} \tag{3.8}$$

Using the expansion to two terms and suitable conditions on the derivatives of ψ the first step is to show that large deviations of a suitable root of (3.6) are very unlikely and hence that R_{Nk} which is governed by $N^{\frac{1}{2}}(\hat{\theta} - \theta)^{k+1}$ can be bounded by something only slightly larger than $N^{-k/2}$. The next step is to consider the equation

$$\begin{aligned} 0 = N^{-\frac{1}{2}} \textstyle\sum_{j=1}^N \psi(X_j - \theta) - \left\{\frac{1}{N} \sum_{j=1}^N \psi'(X_j - \theta)\right\} N^{\frac{1}{2}}(t - \theta) + \cdots \\ + N^{-(k-1)/2} \frac{(-1)^k}{k!} \left\{\frac{1}{N} \textstyle\sum_{j=1}^N \psi^{(k)}(X_j - \theta)\right\} N^{k/2}(t - \theta)^k\,. \end{aligned} \tag{3.9}$$

The solution $t = \hat{\theta}_1$ of this equation can be expanded in an asymptotic expansion in $N^{-\frac{1}{2}}$ whose leading term is $N^{-\frac{1}{2}} \sum_{j=1}^N \psi(X_j - \theta)/E_\theta(\psi'(X_1 - \theta))$ and whose coefficients are polynomials in $\xi_0, \cdots, \xi_k$ where

$$(3.10) \qquad \xi_r = \frac{1}{N^{\frac{1}{2}}} \sum_{j=1}^N [\psi^{(r)}(X_j - \theta) - E_\theta(\psi^{(r)}(X_j - \theta))] .$$

Then one shows that $\hat{\theta}$ and $\hat{\theta}_1^{(k)}$, the sum of the first k terms in the expansion of $\hat{\theta}_1$, differ to an order that matters only on a set of relatively negligible probability. Then one applies a theorem such as Bhattacharya's to the event $[N^{\frac{1}{2}}(\hat{\theta}_1^{(k)} - \theta) < x]$ which indeed depends only on $(\xi_0, \cdots, \xi_k)$. Finally there is the problem of expanding the multivariate integrals appearing in the multivariate Edgeworth theorem since these depend on N (since $\hat{\theta}_1^{(k)}$ is a polynomial in powers of $N^{-\frac{1}{2}}$ as well as in the ξ_j). The result is an expansion of the type (2.7). It is formally clear that the coefficients should agree with those obtained by using the formal expansions of the cumulants in powers of $N^{-\frac{1}{2}}$ from [27] and then proceeding to get a formal Edgeworth expansion from the formal Charlier expansion as in (2.4) and (2.5). However, this has not been checked to my knowledge.

Mitrofanova [42] extended the work of [38] to maximum likelihood estimates of a vector parameter. Unfortunately, as was noted by Pfanzagl [46], her proof contains very serious gaps. A salvage operation however seems both possible and worthwhile. In particular this should yield valid expansions for (M) estimates when scale is estimated (as it normally would be). Čibišov's announcement [11] is essentially an extension of the work of [38] to maximum likelihood estimation of a single parameter under rather simple conditions.

Pfanzagl [46] and Michel and Pfanzagl [40] have used a different approach which though much simpler for the case of a single parameter does not appear to generalize. The idea similar to that used by Huber in [32] and earlier by H. E. Daniels [14] is to compare the events $[\hat{\theta} < x]$ and $[\sum_{j=1}^N \psi(X_j - x) < 0]$. For increasing ψ the two events are essentially the same. In general even for functions of the form $\psi(x, \theta)$, under suitable conditions, one can argue that the difference of the two events has negligible probability for $x = \theta + a/N^{\frac{1}{2}}$ with $|a|$ bounded. But to $P[\sum_{j=1}^N \psi(X_j - x) < 0]$ one can apply the classical univariate expansions for sums of independent identically distributed random variables and then use suitable expansions in $(x - \theta)/N^{\frac{1}{2}}$ of the cumulants of $\psi(X_1 - x)$. This method has the advantage of enabling one to deal with ψ functions which are not very smooth such as those introduced by Huber [32]. There seems at present, however, to be no way of dealing with (M) estimates in which scale is estimated simultaneously when the functions defining the estimates cannot be expanded along the lines of [33].

Pfanzagl [46] gives a variety of applications to parametric models of the univariate expansions mentioned above. There have been hardly any numerical studies of the applicability of these expansions. An interesting example, however, is Barnett's work [4] in which he shows that the (formal) expansion is relatively poor when applied to the maximum likelihood estimate of location for a Cauchy sample.

4. Other classes of asymptotically normal statistics. There has been little success so far in validating expansions or even establishing Berry–Esséen bounds of order $1/N^{\frac{1}{2}}$ for general classes of statistics known to be asymptotically normally distributed, other than the ones we have discussed.

Mr. S. Bjerve in work towards a Berkeley thesis has shown that trimmed means admit valid Edgeworth expansions and is in the process of explicitly calculating the coefficients for comparison with the published distributions of the Princeton project [2]. His method employs special properties of the trimmed means and does not carry over to more general estimates. Further work on systematic statistics which can also be handled by elementary means is intended. Even formal work seems surprisingly scarce here. In this connection I would like to mention [16] in which expansions are obtained for the cumulants of single order statistics.

The only theoretical result on rates of convergence for general linear combinations of order statistics known to me is due to Rosenkrantz and O'Reilly [43] who establish various bounds of Berry–Esséen type for the error committed by using the normal approximation to the distribution of a linear combination of order statistics. None of these bounds is of smaller order than $N^{-\frac{1}{4}}$ where N is the sample size. This limitation appears due to the Skorokhod embedding method which they employ. This order is, of course, incorrect for all cases in which sharp bounds are available, i.e., trimmed means (including the mean) and systematic statistics. I conjecture that under mild conditions the "right" order is $N^{-\frac{1}{2}}$.

In 1948 Hoeffding [31] introduced the interesting class of U-statistics, which includes among its members the Wilcoxon two sample statistic. As another illustration of the power of the Fourier technique in a nonstandard situation we shall prove under rather strong conditions that the normal approximation to the distribution of a U-statistic of order 2 is valid to order $N^{-\frac{1}{2}}$. Our method can be adapted to yield the $N^{-\frac{1}{2}}$ bound for the one and two sample Wilcoxon statistic as well as Kendall's τ. (In fact fixed alternative asymptotic expansions for these statistics can be obtained using a combination of the methods of the appendix and those of [1].) The method should also extend to von Mises statistics [56] of order 1 and hence to linear combinations of order statistics. However we are unable to get $N^{-\frac{1}{2}}$ bounds for U-statistics with unbounded kernels. Bounds of order $N^{-r/2}$, $r < 1$, have been obtained by Grams and Serfling in [25] by a different technique. Asymptotic expansions in general seem out of reach. Here is the statement of our theorem. The proof is given in an appendix.

Let $R_1, \cdots, R_N$ be a sample from the uniform distribution on $(0, 1)$. Let ψ be a measurable real-valued function on the closed unit square such that $|\psi| \leqq M < \infty$ (say). Suppose moreover that ψ is symmetric, $\psi(u, v) = \psi(v, u)$ and that

$$\int_0^1 \int_0^1 \psi(u, v)\, du\, dv = 0\,. \tag{4.1}$$

Let

$$T_N = \frac{1}{\sigma_N} \sum_{i<j} \psi(R_i, R_j) \tag{4.2}$$

where

$$\sigma_N^2 = \frac{N(N-1)}{2} \int_0^1 \int_0^1 \psi^2(u, v)\, du\, dv + N(N-1)(N-2) \int_0^1 \gamma^2(u)\, du \tag{4.3}$$

and

$$\gamma(u) = \int_0^1 \psi(u, v)\, dv\,. \tag{4.4}$$

THEOREM 4.1. *If the preceding assumptions hold and γ does not vanish identically, then there exists a constant C depending on ψ but not N such that*

$$\sup_x |P[T_N \leqq x] - \Phi(x)| \leqq \frac{C}{N^{\frac{1}{2}}}$$

where Φ is the standard normal cumulative distribution function.

A new approach has recently been advanced by Stein [52] which does not rely on Fourier analytic methods. Using his method he is able to show that the error committed in applying the normal approximation to the sum of the first N of a stationary sequence of bounded m dependent random variables is of order $N^{-\frac{1}{2}}$. The possibility of applying his method to some of the classes we have considered should be investigated.

5. Expansions for statistics with nonnormal limiting distributions. The omnibus goodness of fit and two sample tests such as those of Kolmogorov–Smirnov and Cramér–von Mises and the Pearson χ^2 test do not have limiting normal distributions. The Russian school of probability theorists has had considerable success in obtaining expansions for the distribution of the Kolmogorov–Smirnov test statistics under the null hypothesis. The methods employed at first used explicit representations of the null distribution. An account of results of this type due to Chan Li-Tsien may be found in Gnedenko, Korolyuk, Skorokhod [23]. The most definitive expansion for the one-sided goodness of fit statistic was given by Lauwerier [36]. Subsequently, the problems were treated as special cases of more general problems of first passage times of random walks (cf. for example Borovkov [7] in which the two sample Smirnov statistic is treated). An account of the latest results and extensive references may be found in Borovkov [8]. Since none of the first order limiting distributions under contiguous alternatives for these statistics have been tabled or extensively studied it is not surprising that there has been no work on asymptotic expansions for the power.

There has recently been some interest in obtaining Berry–Esséen type bounds for the difference between the distribution of the Cramér–von Mises goodness of fit statistic under the null hypothesis and its well known limit distribution. However, the methods used by Rosenkrantz in [49] and Sawyer in [50] (cf. also Orlov [44]) use the Skorokhod embedding and not surprisingly obtain bounds which are of order strictly worse than $N^{-\frac{1}{2}}$ where N is the sample size. In an announcement of results without proofs [15] D. Darling obtained a representation for the characteristic function of the von Mises statistic which he employed to get an asymptotic expansion of the characteristic function to two terms for fixed argument. I do not know whether this approach can be refined to yield the kind of estimates which permit us to apply Berry's lemma.

Finally, I want to mention the recent Chicago thesis of Yarnold [59] in which he obtained asymptotic expansions for the distribution of Pearson's χ^2 statistic. Since χ^2 is a smooth function of the multinomial frequencies we might expect that the theorems on multivariate Edgeworth series should apply. Unfortunately the vector of multinomial frequencies is a normalized sum of independent identically distributed random vectors taking their values in a lattice, Cramér's condition (C) does not hold and in fact the formal Edgeworth expansion is invalid. However, it is possible to use the well-known local limit expansion for the multinomial probability and then sum up over all points in the appropriate region. This is an improvement over the χ^2 approximation but almost as complicated as calculation of the exact probabilities. Moreover, it does not yield a form which is sufficiently tractable analytically to settle long outstanding questions about the relative performance of the χ^2 and likelihood ratio tests. Results which are manageable in this area would be interesting but seem hard.

6. Appendix (Proof of Theorem 4.1). Let

$$S_N = \frac{(N-1)}{\sigma_N} \sum_{i=1}^N \gamma(R_i) \tag{6.1}$$

$$\Delta_N = T_N - S_N \tag{6.2}$$

$$\phi_N(t) = E(e^{itT_N}) \tag{6.3}$$

$$\eta(t) = E(e^{it\gamma(R_1)}) \tag{6.4}$$

$$\tilde{\phi}_N(t) = E(e^{itS_N}) = \eta^N\left(\frac{t(N-1)}{\sigma_N}\right). \tag{6.5}$$

The crux of the argument is to show that there exists $\varepsilon_1 > 0$ and a constant D_1 both independent of N such that

$$\int_{-\varepsilon_1 N^{\frac{1}{2}}}^{\varepsilon_1 N^{\frac{1}{2}}} \frac{|\phi_N(t) - \tilde{\phi}_N(t)|}{|t|}\, dt \leqq D_1 N^{-\frac{1}{2}}. \tag{6.6}$$

Since it is well known that there exists $\varepsilon_2 > 0$ and a constant D_2 both independent of N such that

$$\int_{-\varepsilon_2 N^{\frac{1}{2}}}^{\varepsilon_2 N^{\frac{1}{2}}} \frac{|\tilde{\phi}_N(t) - e^{-t^2/2}|}{|t|}\, dt \leqq D_2 N^{-\frac{1}{2}},$$

it follows that if $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$, $D = D_1 + D_2$,

$$\int_{-\varepsilon N^{\frac{1}{2}}}^{\varepsilon N^{\frac{1}{2}}} \frac{|\phi_N(t) - e^{-t^2/2}|}{|t|}\, dt \leqq D N^{-\frac{1}{2}}, \tag{6.7}$$

and the theorem follows from (6.6) and the usual Berry–Esséen argument.

To prove (6.6) we need the following lemmas.

LEMMA 6.1. *Let* $\{\xi_j\}$, $1 \leqq j \leqq n$ *be a sequence of martingale summands, i.e.,*

$$E(\xi_j \mid \xi_1, \cdots, \xi_{j-1}) = 0\,, \qquad 1 \leqq j \leqq n\,.$$

Let $W_n = \sum_{j=1}^n \xi_j$. *Define* $m_{n,k} = \max_{1\leq j\leq n} E(\xi_j^{2k})$, $k \geqq 1$. *Then, for* $k \leqq n$,

$$E(W_n^{2k}) \leqq n^k m_{n,k}(4ek)^k\,. \tag{6.8}$$

REMARKS. (1) An estimate similar to (6.8) has been obtained by Dharmadhikari, Fabian and Jogdeo [18] with $m_{n,k}$ replaced by $(1/n)\sum_{j=1}^n E(\xi_j^{2k})$. However, their bound grows with k as 2^{k^2} which is quite inadequate for our purposes. We note that our technique readily establishes,

$$E(W_n^{2k}) \leqq n^k m_{n,k}(k)^{2k}$$

for all k, n but even this is inadequate.

(2) The example of ξ_j i.i.d. normal random variables with mean 0 shows that our bound is comparatively sharp. Also see the remark on Lemma 6.2.

Our main interest in Lemma 6.1 is in its application to

LEMMA 6.2. *Under the conditions of Theorem* 4.1, *if* $k \leqq N$,

$$E(\Delta_N^{2k}) \leqq \sigma_N^{-2k} N^{2k}(3M)^{2k}(4ek)^{2k}\,. \tag{6.9}$$

REMARK. The order of magnitude of the coefficient of $\sigma^{N-2k}N^{2k}$ in (6.9) is quite sharp. Thus if $\psi(x, y) = \frac{3}{4}$ if x and y are both $\geqq \frac{1}{2}$, $= -\frac{1}{4}$ otherwise

$$\sigma_N \Delta_N = \sum_{i<j} \eta_i \eta_j = \frac{1}{2}\left[(\textstyle\sum_{i=1}^N \eta_i)^2 - \frac{N}{4}\right] \tag{6.10}$$

where the η_i are independent and equal $\pm\frac{1}{2}$ with equal probability $\frac{1}{2}$. It is easy to see that

$$E(\sigma_N \Delta_N)^{2k} \geqq 8^{-2k}\{2^{-2k+1}E(U_N^{4k}) - N^{2k}\} \tag{6.11}$$

where $U_N = \sum_{i=1}^N \varepsilon_i$ and $\varepsilon_i = \pm 1$ with probability $\frac{1}{2}$. Since,

$$E(U_N^{4k}) = \sum_{t_1+\cdots+t_N=2k} \frac{4k!}{2t_1! \cdots 2t_N!}\,,$$

$$E(U_N^{4k}) \geqq \binom{N}{2k}\frac{4k!}{2^{2k}} \geqq A(kN)^{2k}\left(1 - \frac{(2k-1)}{N}\right)^{2k}\left(\frac{4}{e}\right)^{2k}$$

for some universal constant A and hence,

$$(kN)^{-2k}E(\sigma_N \Delta_N)^{2k} \geqq c\rho^k$$

for all N and $k \leqq aN$, $a < \frac{1}{2}$ where c and ρ depend on a but not on k and N. Then the ratio between $E(\sigma_N \Delta_N)^{2k}$ and the estimate given by (6.9) is (relatively) negligible.

PROOF OF LEMMA 6.1. The proof is by induction on n for fixed k. Note first that

$$E(\xi_1 + \cdots + \xi_k)^{2k} \leqq k^{2k} m_{k,k} \tag{6.12}$$

and hence the induction hypothesis holds for $n = k$. Suppose it is true for $n = l \geqq k$. Then

$$E(W_{l+1}^{2k}) = E(W_l^{2k}) + \sum_{j=2}^{2k} \binom{2k}{j} E(W_l^{2k-j}\xi_{l+1}^j) \tag{6.13}$$

by the martingale hypothesis. By induction and the Hölder inequality we obtain

$$\begin{aligned} E(W_l^{2k-j}\xi_{l+1}^j) &\leqq [c_k l^k m_{l,k}]^{1-j/2k}[m_{l+1,k}]^{j/2k} \\ &\leqq (c_k l^k m_{l+1,k})(c_k^{1/2k} l^{\frac{1}{2}})^{-j} \end{aligned} \tag{6.14}$$

where $c_k = (4ek)^k$. By elementary estimates (6.13) and (6.14) yield

$$\begin{aligned} E(W_{l+1}^{2k}) &\leqq c_k l^k m_{l+1,k}\left(1 + \frac{4k^2}{lc_k^{1/k}} \sum_{j=0}^{2k-2} \binom{2k-2}{j}(c_k^{1/2k} l^{\frac{1}{2}})^{-j}\right) \\ &\leqq c_k l^k m_{l+1,k}\left(1 + \frac{k}{le}\left(1 + \frac{1}{2(ekl)^{\frac{1}{2}}}\right)^{2k-2}\right) \\ &\leqq c_k l^k m_{l+1,k}\left(1 + \frac{k}{l}\right) \end{aligned} \tag{6.15}$$

for $k \leqq l$. Since $(1 + k/l) \leqq ((l + 1)/l)^k$ the hypothesis is verified for $n = l + 1$ and the result follows.

PROOF OF LEMMA 6.2. Begin by noting that

$$\sigma_N \Delta_N = \sum_{j=1}^N \xi_j \tag{6.16}$$

where

$$\xi_j = \sum_{i=1}^{j-1} [\psi(R_i, R_j) - \gamma(R_i) - \gamma(R_j)] \tag{6.17}$$

and that the ξ_j are martingale summands. Moreover, note that

$$E(\xi_j^{2k}) = E(E[\textstyle\sum_{i=1}^{j-1} (\psi(R_i, R_j) - \gamma(R_i) - \gamma(R_j))^{2k} \mid R_j]) \tag{6.18}$$

and that given R_j the summands $\eta_i = (\psi(R_i, R_j) - \gamma(R_i) - \gamma(R_j))$, $i = 1, \cdots, j - 1$ are also martingale summands (in fact i.i.d.). Since

$$E(\psi(R_1, R_2) - \gamma(R_1) - \gamma(R_2))^{2k} \leqq (3M)^{2k} \tag{6.19}$$

we can apply Lemma 6.1 twice in succession to obtain Lemma 6.2.

LEMMA 6.3. *Under the conditions of the theorem,*

$$|E(e^{itS_N}\Delta_N)| \leqq 3M^3t^2 \frac{N^4}{\sigma_N^3} |\eta|^{N-2}\left(\frac{t}{\sigma_N}(N - 1)\right) \tag{6.20}$$

$$|E(e^{itS_N}\Delta_N^j)| \leqq \left(\frac{N^2}{\sigma_N}\right)^j \left(\frac{3M}{2}\right)^j |\eta|^{N-2j}\left((N - 1)\frac{t}{\sigma_N}\right) \quad \textit{for } j \geqq 1\,. \tag{6.21}$$

PROOF. To prove (6.20) we calculate

$$\begin{aligned} E(\Delta_N e^{itS_N}) = {} & \frac{N(N-1)}{2\sigma_N} \eta^{N-2}\left(\frac{t}{\sigma_N}(N - 1)\right) \\ & \times E\left(\exp\left[\frac{it(N-1)}{\sigma_N}(\gamma(R_1) + \gamma(R_2))\right] \right. \\ & \left. \times (\psi(R_1, R_2) - \gamma(R_1) - \gamma(R_2))\right). \end{aligned} \tag{6.22}$$

Since $\psi(R_1, R_2) - \gamma(R_1) - \gamma(R_2)$ and $\gamma(R_1), \gamma(R_2)$ are uncorrelated we can write

$$\left|E\left(\exp\left[\frac{it(N-1)}{\sigma_N}(\gamma(R_1)+\gamma(R_2))\right](\psi(R_1, R_2) - \gamma(R_1) - \gamma(R_2))\right)\right|$$

$$(6.23)\qquad \begin{aligned} &= \left|E\left[\left(\exp\left[\frac{it(N-1)}{\sigma_N}(\gamma(R_1)+\gamma(R_2))\right]-1\right)\right.\right. \\ &\qquad \left.\left.\times\,(\psi(R_1, R_2) - \gamma(R_1) - \gamma(R_2))\right]\right| \\ &\leqq \frac{t^2}{2}\frac{(N-1)^2}{{\sigma_N}^2} E[(\gamma(R_1)+\gamma(R_2))^2|\psi(R_1, R_2) - \gamma(R_1) - \gamma(R_2)|] \\ &\leqq 6M^3t^2\frac{(N-1)^2}{{\sigma_N}^2}, \end{aligned}$$

and (6.20) follows.

Similarly,

$$(6.24)\qquad \begin{aligned} &{\sigma_N}^j E({\Delta_N}^j e^{itS_N}) \\ &= \textstyle\sum_{\{(a_1,b_1),\cdots,(a_j,b_j)\}} E(e^{itS_N}[\prod_{i=1}^j (\psi(R_{a_i}, R_{b_i}) - \gamma(R_{a_i}) - \gamma(R_{b_i}))]) . \end{aligned}$$

Applying elementary inequalities we obtain

$$(6.25)\qquad \begin{aligned} &|{\sigma_N}^j E({\Delta_N}^j e^{itS_N})| \\ &\leqq \frac{N^{2j}}{2^j}|\eta|^{N-2j}\left((N-1)\frac{t}{\sigma_N}\right)E|\psi((R_1, R_2) - \gamma(R_1) - \gamma(R_2))|^j \\ &\leqq \left(\frac{3M}{2}\right)^j N^{2j}|\eta|^{N-2j}\left(\frac{(N-1)t}{\sigma^N}\right). \end{aligned}$$

The lemma follows.

We proceed with the proof of (6.6). Since

$$|\phi_N(t) - \tilde{\phi}_N(t)| = |E(e^{itS_N}(e^{it\Delta_N} - 1))|$$

we have for any k,

$$(6.26)\qquad |\phi_N(t) - \tilde{\phi}_N(t)| \leqq \left|\textstyle\sum_{j=1}^{2k-1}\frac{(it)^j}{j!}E(e^{itS_N}{\Delta_N}^j)\right| + \frac{t^{2k}}{(2k)!}E({\Delta_N}^{2k}) .$$

From (6.26), (6.9) and (6.20),

$$(6.27)\qquad |\phi_N(t) - \tilde{\phi}_N(t)| \leqq \left(3M^3\frac{N^4}{{\sigma_N}^3}|\eta|^{N-2}\left(\frac{t}{\sigma_N}(N-1)\right)t + 8e^2\frac{N^2}{{\sigma_N}^2}M^2\right)t^2 .$$

Since there exists $\theta > 0$ such that ${\sigma_N}^2 \geqq \theta^2 N^3$ for all N we conclude that

$$(6.28)\qquad \begin{aligned} &\int_{-N^{\frac{1}{2}}}^{N^{\frac{1}{2}}}\frac{|\phi_N(t) - \tilde{\phi}_N(t)|}{|t|}dt \\ &\leqq \frac{3N^{-\frac{1}{2}}M^3}{\theta^3}\int_{-N^{\frac{1}{2}}}^{N^{\frac{1}{2}}}|t|^2|\eta|^{N-2}\left(\frac{tN^{-\frac{1}{2}}}{\theta}\right)dt + \frac{8e^2M^2}{\theta^2}N^{-\frac{1}{2}} \\ &\leqq FN^{-\frac{1}{2}} \end{aligned}$$

where F is a constant depending on ψ but not N.

Let

$$\varepsilon = \frac{\theta p}{24Me}, \qquad p < 1 \tag{6.29}$$

$$k = \left\{\left(\left[\frac{1}{2}\frac{\log N}{|\log p|}\right] + 1\right) \wedge N\right\}.$$

If $|t| \leqq \varepsilon N^{\frac{1}{2}}$, by Lemma 6.2 for this k and N sufficiently large,

$$\begin{aligned}\frac{t^{2k}}{(2k)!} E(\Delta_N^{2k}) &\leqq \frac{\varepsilon^{2k} N^k}{(2k)!} \cdot \frac{N^{2k}}{\sigma_N^{2k}} k^{2k}(12eM)^{2k} \\ &\leqq \binom{4k}{2k} 2^{-4k} p^{2k} < N^{-1}.\end{aligned} \tag{6.30}$$

To complete the argument note that for p sufficiently small, there exists $\tau > 0$ such that for $|t| \leqq \varepsilon N^{\frac{1}{2}}$,

$$\log |\gamma| \left(\frac{(N-1)t}{\sigma_N}\right) \leqq -\frac{\tau t^2}{N}. \tag{6.31}$$

Applying (6.31) and (6.21) we conclude that for $N^{\frac{1}{4}} \leqq |t| \leqq \varepsilon N^{\frac{1}{2}}$, $j < 2k$,

$$|E(e^{itS_N} \Delta_N^{\ j})| \leqq \theta^{-j} N^{j/2} \left(\frac{3M}{2}\right)^j \exp\left[-\tau N^{\frac{1}{2}}\left(1 - \frac{4k}{N}\right)\right]. \tag{6.32}$$

Hence for $N^{\frac{1}{4}} \leqq |t| \leqq \varepsilon N^{\frac{1}{2}}$ with k, ε given by (6.29),

$$\begin{aligned}\left|\sum_{j=1}^{2k-1} \frac{(it)^j}{j!} E(e^{itS_N} \Delta_N^{\ j})\right| &\leqq e N^{2k} \left(\frac{3M}{2\theta}\right)^{2k} \exp\left[-\tau N^{\frac{1}{2}}\left(1 - \frac{4k}{N}\right)\right] \\ &= o\left(\frac{1}{N}\right)\end{aligned} \tag{6.33}$$

uniformly for $|t|$ as above. Combining (6.28), (6.30) and (6.33), (6.6) and the theorem follows.

REFERENCES

[1] ALBERS, W., BICKEL, P. J. and VAN ZWET, W. R. (1972). Asymptotic expansions for the power of distribution free tests in the one sample problem. To be submitted to *Ann. Statist.*

[2] ANDREWS, D. F., BICKEL, P. J., HAMPEL, F. R., ROGERS, W. H. and TUKEY, J. W. (1972). *Robust Estimation: Survey and Advances.* Princeton Univ. Press.

[3] ARNOLD, H. J. (1965). Small sample power for the one sample Wilcoxon test. *Ann. Math. Statist.* **36** 1767–1778.

[4] BARNETT, V. (1966). Evaluation of maximum likelihood estimate when likelihood has multiple roots. *Biometrika* **53** 151–165.

[5] BHATTACHARYA, R. N. (1971). Rates of convergence and asymptotic expansions. *Ann. Math. Statist.* **42** 241–259.

[6] BICKEL, P. J. and VAN ZWET, W. R. (1973). Asymptotic expansions for the power of distribution free tests in the two sample problem. In preparation.

[7] BOROVKOV, A. A. (1962). On the problem of two samples. *Selected Transl. Math. Statist. Prob.* **5** 285–307.

[8] BOROVKOV, A. A. (1970). Limit theorems for random walks with boundaries. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **3** 19–30.

[9] BOWMAN, K. and SHENTON, L. R. (1963). Higher moments of maximum likelihood estimates. *J. Roy. Statist. Soc. Ser. B* **25** 305-317.

[10] CHAMBERS, J. M. (1967). On methods of asymptotic approximation for multivariate distributions. *Biometrika* **54** 367-384.

[11] ČIBIŠOV, D. M. (1972). Asymptotic expansions for maximum likelihood estimates. *Teor. Verojatnost. i Primenen.* **17** 387-388.

[12] CRAMÉR, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

[13] DANIELS, H. E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Statist.* **25** 631-650.

[14] DANIELS, H. E. (1960). The asymptotic efficiency of a maximum likelihood estimator. *Proc. Fourth Berkeley Symp. Math. Statist.* **1** 151-164.

[15] DARLING, D. (1960). Sur Les Théorèmes de Kolmogorov-Smirnov. *Theor. Probability Appl.* **5** 356-361.

[16] DAVID, F. N. and JOHNSON, N. L. (1954). Statistical treatment of censored data, Part I. *Biometrika* **41** 228-240.

[17] DAVIES, R. B. (1971). Rank tests for Lehmann's alternative. *J. Amer. Statist. Assoc.* **66** 879-883.

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

[19] ERDELYI, A. (1956). *Asymptotic Expansions.* Dover, New York.

[20] ERDÖS, P. and RÉNYI, A. (1950). On the central limit theorem for samples from a finite population. *Publ. Math. Inst. Hungar. Acad. Sci.* **4** 49-61.

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

[22] FELLINGHAM, S. A. and STOKER, D. J. (1964). An approximation for the exact distribution of Wilcoxon test for symmetry. *J. Amer. Statist. Assoc.* **59** 899-905.

[23] GNEDENKO, B. V., KOROLYUK, V. and SKOROKHOD, A. V. (1960). Asymptotic expansions in probabilty theory. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **2** 153-169.

[24] GOVINDARAJULU, Z. and HAYNAM, G. (1956) Exact power of Mann-Whitney test for exponential and rectangular alternatives. *Ann. Math. Statist.* **37** 945-953.

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

[26] HÁJEK, J. and ŠIDÁK, A. (1967). *Theory of Rank Tests.* Academic Press, New York.

[27] HALDANE, J. B. S. and SMITH, S. M. (1956). Sampling distribution of a maximum likelihood estimate. *Biometrika* **43** 96-103.

[28] HODGES, J. L. and FIX, E. (1955). Significance probabilities of the Wilcoxon test. *Ann. Math. Statist.* **26** 301-312.

[29] HODGES, J. L. and LEHMANN, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598-611.

[30] HODGES, J. L. and LEHMANN, E. L. (1970). Deficiency. *Ann. Math. Statist.* **41** 783-801.

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

[32] HUBER, P. J. (1964). Robust estimation of location. *Ann. Math. Statist.* **35** 73-101.

[33] HUBER, P. J. (1965). Behavior of maximum likelihood estimates under non-standard conditions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 221-233.

[34] KLOTZ, J. (1963). Small sample power and efficiency of the one sample Wilcoxon and normal scores tests. *Ann. Math. Statist.* **34** 624-632.

[35] KLOTZ, J. (1964). On the normal scores two sample rank test. *J. Amer. Statist. Assoc.* **59** 652-654.

[36] LAUWERIER, H. A. (1963). The asymptotic expansions of the statistical distribution of N. V. Smirnov. *Z. Warscheinlichkeitstheorie und Verw. Gebiete* **2** 61-68.

[37] LINNIK, J. (1960). On the probability of large deviations for the sums of independent random variables. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **2** 289-306.

[38] LINNIK, J. and MITROFANOVA, N. (1965). Some asymptotic expansions for maximum likelihood estimates. *Sankhyā A* **27** 73-82.

[39] MAAG, U. and DICAIRE, G. (1971). On Kolmogorov-Smirnov type one sample statistics. *Biometrika* **58** 653-656.

[40] MICHEL, R. and PFANZAGL, J. (1971). Accuracy of normal approximations for minimum contrast estimates. *Z. Warscheinlichkeitstheorie und Verw. Gebiete* **18** 73-84.

[41] MILTON, R. C. (1970). *Rank Order Probabilities.* Wiley, New York.

[42] MITROFANOVA, N. (1967). Asymptotic expansions for maximum likelihood estimates of a vector parameter. *Theor. Probability Appl.* **12** 364-372.

[43] O'REILLY, N. and ROSENKRANTZ, W. (1972). Applications of the Skorokhod representation to rates of convergence for linear combinations of order statistics. *Ann. Math. Statist.* **43** 1204-1212.

[44] ORLOV, A. I. (1971). Estimates of speed of convergence to their limit distributions of certain statistics. *Theor. Probability Appl.* **16** 526.

[45] PFANZAGL, J. (1971) Berry-Esséen bound for minimum contrast estimates. *Metrika* **17** 82-91.

[46] PFANZAGL, J. (1973). Asymptotic expansions related to minimum contrast estimators. *Ann. Statist.* **1** 993-1026.

[47] RANGA RAO, R. (1961). On the central limit theorem in R^k. *Bull. Amer. Math. Soc.* **67** 359-361.

[48] ROGERS, W. F. (1971). Exact null distributions and asymptotic expansions for rank test statistics. Technical Report, Stanford Univ.

[49] ROSENKRANTZ, W. (1969). A rate of convergence for the von Mises functional. *Trans. Amer. Math. Soc.* **139** 329-337.

[50] SAWYER, S. (1972). Rates of convergence for some functionals in probability. *Ann. Math. Statist.* **43** 273-284.

[51] SAZONOV, V. (1969). Improvement of a convergence rate estimate. *Theor. Probability Appl.* **14** 649-651.

[52] STEIN, C. (1970). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2**.

[53] STOKER, D. J. (1954). An upper bound for the deviation between the distribution of Wilcoxon's test statistic for the two sample problem and its limiting normal distribution for finite samples, I, II. *Indag. Math.* **16** 599-606, 607-614.

[54] SUNDRUM, R. (1954). A further approximation to the distribution of Wilcoxon's statistic. *J. Roy. Statist. Soc. Ser. B* **16** 266-268.

[55] THOMPSON, R., GOVINDARAJULU, Z. and DOKSUM, K. (1967). Distribution and power of the absolute normal scores test. *J. Amer. Statist. Assoc.* **62** 966-975.

[56] VON MISES, R. (1947). Differentiable statistical functions. *Ann. Math. Statist.* **18** 309-348.

[57] WALLACE, D. (1958). Asymptotic approximations to distributions. *Ann. Math. Statist.* **29** 635-654.

[58] WITTING, H. (1960). A generalized Pitman efficiency. *Ann. Math. Statist.* **31** 405-414.

[59] YARNOLD, J. (1968). The accuracy of seven approximations for the null distribution of the chi-square goodness of fit statistic. Thesis, Univ. of Chicago.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720