stichting
mathematisch
centrum

MC

DEPARTMENT OF MATHEMATICAL STATISTICS SW 38/76 APRIL

P.J. BICKEL & W.R. VAN ZWET

# ASYMPTOTIC EXPANSIONS FOR THE POWER OF DISTRIBUTIONFREE TESTS IN THE TWO-SAMPLE PROBLEM

Prepublication

2e boerhaavestraat 49 amsterdam

*Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.*

*The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.*

---

AMS(MOS) subject classification scheme (1970): 62G10, 62G20; Secondary 60F05.

---

Asymptotic expansions for the power of distributionfree tests in the two-sample problem *)

by

P.J. BICKEL & W.R. VAN ZWET

ABSTRACT

Asymptotic expansions are established for the power of rank tests in the two-sample problem.

KEY WORDS & PHRASES: *Distributionfree tests, linear rank test, power, contiguous alternatives, Edgeworth expansions.*

---

*) This paper is not for review; it is meant for publication elsewhere

## 1. INTRODUCTION

Let $X_1, X_2, \ldots, X_N$, $N = m+n$, be independent random variables such that $X_1, \ldots, X_m$ are identically distributed with common distribution function F and density f and $X_{m+1}, \ldots, X_N$ are identically distributed with distribution function G and density g. For $N = 1,2,\ldots$ , $0 < \varepsilon \leq m/N \leq 1-\varepsilon < 1$, consider the problem of testing the hypothesis $F = G$ against a sequence of alternatives that is contiguous to the hypothesis. The level $\alpha$ of the sequence of tests is fixed in (0,1). Standard tests for this two-sample problem are linear rank tests and expressions for the limiting powers of such tests are well-known. In this paper we shall establish asymptotic expansions to order $N^{-1}$ for the powers $\pi_N$ of such 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 finding 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. Extension of the results to general contiguous alternatives is straightforward but messy.

The paper is thus the natural counterpart of the first four sections of ALBERS, BICKEL and VAN ZWET (ABZ) [1976] where the same program is carried out for the one-sample problem. In view of the strong similarity between the one- and two-sample cases, it is not surprising that many of the techniques developped in ABZ [1976] also play an important part in the present paper. However, there also exist significant differences that make the two-sample problem essentially more complicated. Two of these differences are worth mentioning at this stage. In the one-sample case a conditioning argument reduces the rank statistic to a sum of independent Bernoulli random variables. A similar argument in the two-sample case leads to a much more complicated random variable connected with rejective sampling. A second difference is that in the one-sample location problem the terms of order $N^{-\frac{1}{2}}$ in the expansions vanish because of certain symmetries that are present in this case. In general, there are no such symmetries in the two-sample location problem and the presence of terms of order $N^{-\frac{1}{2}}$ creates several additional technical difficulties.

A number of authors have computed formal expansions for the distributions of two-sample rank statistics without proof of their validity. For an account of this work we refer to a review paper of BICKEL [1974]. ROGERS [1971] has given a proof for the two-sample Wilcoxon statistic under the hypothesis.

In section 2 we point out that for arbitrary F and G, the conditional distribution of the two-sample linear rank statistic given the order statistics of the combined sample is the same as the distribution of the sample sum in a rejective sampling scheme. We establish an expansion for the distribution of such a sample sum, which may be of interest in its own right. As a corollary we immediately obtain an expansion for the distribution function of the rank statistic under the hypothesis. In section 3 we return to general F and G and obtain an unconditional expansion for the distribution function of the rank statistic. We specialize to contiguous location alternatives in section 4 and derive an expansion for the power of the rank test. In section 5 we deal with the important case where the scores are exact or approximate scores generated by a smooth function.

As was explained in the introduction of ABZ [1976], an important application of asymptotic expansions of the power of rank tests is the computation of asymptotic deficiencies in the sense of HODGES and LEHMANN [1970] of these rank tests with respect to their parametric competitors. For the one-sample case such deficiency computations for rank tests and for the associated estimators were given in sections 6 and 7 of ABZ [1976]. For the two-sample case these computations will be given in a seperate paper.

## 2. AN EXPANSION FOR THE CONDITIONAL DISTRIBUTION OF TWO-SAMPLE RANK STATISTICS AND ITS APPLICATION TO REJECTIVE SAMPLING

Let $X_1, X_2, \ldots, X_N$, $N = m + n$, be independent random variables (r.v.'s) such that $X_1, \ldots, X_m$ are identically distributed (i.d.) with common distribution function (d.f.) $F$ and density $f$ and $X_{m+1}, \ldots, X_N$ are i.d. with common d.f. $G$ and density $g$. Let $Z_1 < Z_2 < \ldots < Z_N$ denote the order statistics of $X_1, \ldots, X_N$, define the antiranks $D_1, D_2, \ldots, D_N$ by $X_{D_j} = Z_j$ and let

$$(2.1) \qquad V_j = \begin{cases} 1 & \text{if } m+1 \le D_j \le N \\ 0 & \text{otherwise.} \end{cases}$$

For a specified vector of scores $a = (a_1, a_2, \ldots, a_N)$ define a two-sample rank statistic by

$$(2.2) \qquad T = \sum_{j=1}^{N} a_j V_j .$$

Our aim is to obtain an asymptotic expansion as $N \to \infty$ for the distribution of $T$ for suitable sequences of pairs of d.f.'s $(F_N, G_N)$, arrays of scores $\{a_{j,N}\}$, $1 \le j \le N$, and sample sizes $(m_N, n_N)$. As in ALBERS, BICKEL and VAN ZWET (ABZ) (1976) we shall suppress dependence on $N$ whenever possible and formally present our results in terms of error bounds for fixed, but arbitrary, values of $N$.

Under the null-hypothesis that $F = G$,

$$P(V_1=v_1, \ldots, V_N=v_N) = \frac{1}{\binom{N}{n}}$$

for any vector $(v_1, \ldots, v_N)$ with $m$ co-ordinates equal to 0 and $n$ co-ordinates equal to 1. In general, conditional on $Z = (Z_1, \ldots, Z_N)$,

$$(2.3) \qquad P(V_1=v_1, \ldots, V_N=v_N \mid Z) = c^{-1}(P) \prod_{j=1}^{N} P_j^{v_j} (1-P_j)^{1-v_j} ,$$

where

$$(2.4) \qquad P_j = \frac{\lambda g(Z_j)}{(1-\lambda) f(Z_j) + \lambda g(Z_j)} ,$$

(2.5) $$\lambda = \frac{n}{N} ,$$

(2.6) $$c(P) = \sum \prod_{j=1}^{N} P_j^{w_j}(1-P_j)^{1-w_j} ,$$

and the summation is over all vectors $(w_1,\ldots,w_N)$ consisting of m zeros and n ones.

Let $W_1,W_2,\ldots,W_N$ be independent r.v.'s with $P(W_j=1) = 1 - P(W_j=0) = p_j$, $1 \le j \le N$. Suppose that

(2.7) $$\begin{aligned} p_j &= 0 \qquad \text{for at most m indices j} \\ p_j &= 1 \qquad \text{for at most n indices j} \end{aligned}$$

and consider the conditional distribution of $\sum a_j W_j$ given that $\sum W_j = n$. Note that if we replace $p = (p_1,\ldots,p_N)$ by $P = (P_1,\ldots,P_N)$, then this is the distribution of T given Z. For general p this distribution is of interest in its own right since $\sum a_j W_j$ given $\sum W_j = n$ is the sample sum we obtain when we use a rejective sampling scheme with parameters $p_1,\ldots,p_N$ in selecting a sample of size n from the sampling frame $\{a_1,a_2,\ldots,a_N\}$ (see HÁJEK (1964) for details).

Define

(2.8) $$\rho(t,p) = E(\exp\{itN^{-\frac{1}{2}} \sum_{j=1}^{N} a_j(W_j-p_j) \mid \sum_{j=1}^{N} W_j=n),$$

(2.9) $$R(x,p) = P(N^{-\frac{1}{2}} \sum_{j=1}^{N} a_j(W_j-p_j) \le x \mid \sum_{j=1}^{N} W_j=n).$$

Our program for obtaining an Edgeworth expansion for the d.f. of T parallels in part that of ABZ (1976). We obtain a formula for $\rho$. From this formula we obtain an expansion for $\rho$ which we can rigorously translate into an Edgeworth expansion for R. Because of the connection with rejective sampling we isolate this result as the only theorem in this section. In the next section we proceed with our main program and obtain an expansion for the d.f. of T by replacing p by P and taking the expectation of the resulting expression. We begin with

LEMMA 2.1. *Define*

$$(2.10)\qquad \psi(s,t,p) = \exp\{isN^{-\frac{1}{2}}\sum_{j=1}^{N}(p_j-\lambda)\}\prod_{j=1}^{N}[p_j\exp\{iN^{-\frac{1}{2}}(1-p_j)(s+a_jt)\} + (1-p_j)\exp\{-iN^{-\frac{1}{2}}p_j(s+a_jt)\}],$$

$$(2.11)\qquad \nu(t,p) = \int_{-\pi N^{\frac{1}{2}}}^{\pi N^{\frac{1}{2}}}\psi(s,t,p)ds,$$

$$(2.12)\qquad c(p) = \sum\prod_{j=1}^{N}p_j^{w_j}(1-p_j)^{1-w_j},$$

*where the last summation is over all vectors* $(w_1,\ldots,w_N)$ *consisting of* m *zeros and* n *ones. Then, if* (2.7) *is satisfied,*

$$(2.13)\qquad \rho(t,p) = \frac{1}{2\pi c(p)N^{\frac{1}{2}}}\int_{-\pi N^{\frac{1}{2}}}^{\pi N^{\frac{1}{2}}}\psi(s,t,p)ds = \frac{\nu(t,p)}{\nu(0,p)}.$$

PROOF. Begin with the identity

$$E(\exp\{iN^{-\frac{1}{2}}[s\sum(W_j-p_j) + t\sum a_j(W_j-p_j)]\})$$

$$= \sum_{k=0}^{N} E(\exp\{itN^{-\frac{1}{2}}\sum a_j(W_j-p_j)\}\,|\,\sum W_j=k)P(\sum W_j=k)\exp\{isN^{-\frac{1}{2}}(k-\sum p_j)\}.$$

Because the system $\{(2\pi N^{\frac{1}{2}})^{-1}\exp(iksN^{-\frac{1}{2}}): k = 0, \pm 1,\ldots\}$ is orthonormal on $[-\pi N^{\frac{1}{2}}, \pi N^{\frac{1}{2}}]$ this implies

$$\rho(t,p) = (2\pi N^{\frac{1}{2}}P(\sum W_j=n))^{-1}\int_{-\pi N^{\frac{1}{2}}}^{\pi N^{\frac{1}{2}}}\exp\{isN^{-\frac{1}{2}}\sum(p_j-\lambda)\} \times E(\exp\{iN^{-\frac{1}{2}}\sum(s+a_jt)(W_j-p_j)\})ds.$$

Elementary considerations now yield (2.13). □

Note that if $p_j = \lambda$ for all j - which corresponds to the null-hypothesis in the two-sample problem - our formula agrees with that of ERDÖS and RÉNYI for random sampling without replacement (cf. RÉNYI (1970) p. 462).

In fact their result motivated our approach.

In our asymptotic study of $\psi$, $\nu$ and $\rho$ we shall repeatedly come across the following functions of p.

$$(2.14) \qquad \omega(p) = N^{-\frac{1}{2}} \sum_{j=1}^{N} (p_j - \lambda),$$

$$(2.15) \qquad \sigma^2(p) = N^{-1} \sum_{j=1}^{N} p_j(1-p_j),$$

$$(2.16) \qquad \bar{a}(p) = \sum_{j=1}^{N} p_j(1-p_j)a_j \Big/ \sum_{j=1}^{N} p_j(1-p_j),$$

$$(2.17) \qquad \tau^2(p) = N^{-1} \sum_{j=1}^{N} p_j(1-p_j)(a_j-\bar{a}(p))^2 = N^{-1} \sum_{j=1}^{N} p_j(1-p_j)a_j^2 - \sigma^2(p)\bar{a}^2(p),$$

$$(2.18) \qquad \kappa_{3,i}(p) = N^{-1} \sum_{j=1}^{N} p_j(1-p_j)(1-2p_j)(a_j-\bar{a}(p))^i, \qquad i = 0,1,2,3,$$

$$(2.19) \qquad \kappa_{4,i}(p) = N^{-1} \sum_{j=1}^{N} p_j(1-p_j)(1-6p_j+6p_j^2)(a_j-\bar{a}(p))^i, \qquad i = 0,1,\ldots,4.$$

In this notation we shall suppress the dependence on p when this is convenient. Let $\ell$ denote Lebesgue measure on $R^1$ and define

$$(2.20) \qquad \gamma(\varepsilon,\zeta,p) = \ell\{x: \exists_j \ |x-a_j| < \zeta, \ \varepsilon \le p_j \le 1-\varepsilon\}.$$

LEMMA 2.2. *Suppose that positive numbers* c, C, $\delta$ *and* $\varepsilon$ *exist such that*

$$(2.21) \qquad \tau^2(p) \ge c, \qquad \frac{1}{N}\sum_{j=1}^{N} a_j^4 \le C,$$

$$(2.22) \qquad \gamma(\varepsilon,\zeta,p) \ge \delta N\zeta \qquad \textit{for some } \zeta \ge N^{-3/2}\log N.$$

*Thus there exist positive numbers* b, B *and* $\beta$ *depending only on* c, C, $\delta$ *and* $\varepsilon$ *such that*

$$(2.23) \qquad |\psi(s,t,p)| \le BN^{-\beta\log N}$$

*for all pairs* (s,t) *such that* $|s| \le \pi N^{\frac{1}{2}}$, $|t| \le bN^{3/2}$ *and either* $|s| \ge \log(N+1)$ *or* $|t| \ge \log(N+1)$.

PROOF.

$$(2.24)\qquad |\psi(s,t,p)| = \prod_{j=1}^{N} [1 - 2p_j(1-p_j)\{1 - \cos(N^{-\frac{1}{2}}(s+a_jt))\}]^{\frac{1}{2}}$$

$$\le \exp\{-\sum_{j=1}^{N} p_j(1-p_j)[\tfrac{1}{2}N^{-1}(s+a_jt)^2 - \tfrac{1}{24}N^{-2}(s+a_jt)^4]\}$$

$$\le \exp\{-\tfrac{1}{2}[\tau^2t^2+\sigma^2(s+\bar{a}t)^2] + \tfrac{1}{12}N^{-1}[N^{-1}\sum_{j=1}^{N}(a_j-\bar{a})^4t^4 + (s+\bar{a}t)^4]\}.$$

Now (2.21) ensures that

$$(2.25)\qquad \sigma^2(p) \ge N\tau^4(p) \Big/ \sum_{j=1}^{N} a_j^4 \ge C^{-1}c^2,$$

$$(2.26)\qquad |\bar{a}(p)| \le [N^{-1}\sum_{j=1}^{N} a_j^4]^{\frac{1}{4}} \Big/ \sigma^2(p) \le c^{-2}C^{5/4},$$

and by (2.21), (2.24), (2.25) and (2.26) we conclude that there exist positive $b_1$, $B$ and $\beta$ depending only on $c$ and $C$ such that for $|s| \le b_1N^{\frac{1}{2}}$ and $|t| \le b_1N^{\frac{1}{2}}$

$$(2.27)\qquad |\psi(s,t,p)| \le B\exp\{-\beta(s^2+t^2)\}.$$

Next note that (2.25) and (2.21) imply that the number of indices $j$ for which $p_j(1-p_j) \ge \frac{1}{2}c^2/C$ is at least $2Nc^2/C$ and the number of $j$ for which $|a_j| \le (C/c)^{\frac{1}{2}}$ is at least $N - Nc^2/C$. Hence the number of indices $j$ for which $|a_j| \le (C/c)^{\frac{1}{2}}$ and $p_j(1-p_j) \ge \frac{1}{2}c^2/C$ is at least $Nc^2/C$. Put $b_2 = \frac{1}{2}b_1(c/C)^{\frac{1}{2}}$ and we see that if $b_1N^{\frac{1}{2}} \le |s| \le \pi N^{\frac{1}{2}}$ and $|t| \le b_2N^{\frac{1}{2}}$, then for at least $Nc^2/C$ indices $j$

$$[1 - 2p_j(1-p_j)\{1 - \cos(N^{-\frac{1}{2}}(s+a_jt))\}] \le 1 - c^2C^{-1}\{1 - \cos(\tfrac{b_1}{2})\}.$$

is an asymptotic expansion for $\nu(t,p)$.

LEMMA 2.3. *Suppose that positive numbers* c, C, $\delta$ *and* $\varepsilon$ *exist such that* (2.21) *and* (2.22) *are satisfied. Then there exist positive numbers* b, B *and* $\beta$ *depending only on* c, C, $\delta$ *and* $\varepsilon$ *such that for* $|t| \le bN^{3/2}$,

$$(2.31) \qquad |\nu(t,p) - \tilde{\nu}(t,p)| \le B[(N^{-3/2}+N^{-5/4}|t|^5)\exp\{-\frac{ct^2}{8}\} + N^{-\beta\log N}].$$

PROOF. In this proof b, $b_i$, $B_i$, $\beta_i$ and $N_0$ denote appropriately chosen positive numbers depending only on c, C, $\delta$ and $\varepsilon$.

Arguing as in the proof of theorem 2.1 in ABZ (1976) we find by Taylor expansion of log $\psi$ that if $|s + a_jt| \le \frac{1}{2}\pi N^{\frac{1}{2}}$ for all j, then

$$(2.32) \qquad \psi(s,t,p) = \exp\{i\omega s - \frac{\tau^2t^2}{2} - \frac{\sigma^2(s+\bar{a}t)^2}{2} - \frac{iN^{-3/2}}{6}\sum p_j(1-p_j)(1-2p_j)(s+a_jt)^3 + \frac{N^{-2}}{24}\sum p_j(1-p_j)(1-6p_j+6p_j^2)(s+a_jt)^4 + M_1(s,t,p)\},$$

where

$$|M_1(s,t,p)| \le C_1N^{-5/2}\sum |s+a_jt|^5 \le 16C_1(N^{-5/2}|t|^5\sum |a_j - \bar{a}|^5 + N^{-3/2}|s + \bar{a}t|^5)$$

for some absolute constant $C_1$. Now (2.21) and (2.26) imply that $N^{-1}\sum |a_j-\bar{a}|^3$, $N^{-1}\sum |a_j - \bar{a}|^4$, $N^{-1/4}\max|a_j|$ and $N^{-5/4}\sum |a_j - \bar{a}|^5$ are bounded. Using (2.21) and (2.25) we find that for all $|s| \le b_1N^{\frac{1}{2}}$ and $|t| \le b_1N^{\frac{1}{4}}$

$$\frac{N^{-3/2}}{6}\sum |s + a_jt|^3 + \frac{N^{-2}}{24}\sum (s+a_jt)^4 + |M_1(s,t,p)| \le \frac{\tau^2t^2 + \sigma^2(s+\bar{a}t)^2}{4}.$$

Hence further expansion of part of the exponential in (2.32) shows that

$$(2.33) \qquad \psi(s,t,p) = \tilde{\psi}(s,t,p) + M_2(s,t,p)$$

for $|s| \le b_1 N^{\frac{1}{2}}$ and $|t| \le b_1 N^{\frac{1}{4}}$, where

$$(2.34)\qquad \tilde{\psi}(s,t,p) = \exp\{i\omega s - \frac{\tau^2 t^2}{2} - \frac{\sigma^2(s+\bar{a}t)^2}{2}\}$$
$$\times [1 - \frac{iN^{-3/2}}{6} \sum p_j(1-p_j)(1-2p_j)(s+a_j t)^3$$
$$+ \frac{N^{-2}}{24} \sum p_j(1-p_j)(1-6p_j+6p_j^2)(s+a_j t)^4$$
$$- \frac{N^{-3}}{72} (\sum p_j(1-p_j)(1-2p_j)(s+a_j t)^3)^2] ,$$

$$(2.35)\qquad |M_2(s,t,p)| \le (N^{-3/2}+N^{-5/4}|t|^5)M_3(t,s+\bar{a}t)\exp\{- \frac{\tau^2 t^2 + \sigma^2(s+\bar{a}t)^2}{4}\}$$

and $M_3$ is a polynomial in t and $(s+\bar{a}t)$ of fixed degree with coefficients depending only on c and C. Therefore, for $|t| \le b_1 N^{\frac{1}{4}}$,

$$(2.36)\qquad \int_{-b_1 N^{\frac{1}{2}}}^{b_1 N^{\frac{1}{2}}} |\psi(s,t,p) - \tilde{\psi}(s,t,p)|ds \le B_1(N^{-3/2}+N^{-5/4}|t|^5)\exp\{- \frac{ct^2}{8}\}.$$

Next we show that for $|t| \le b_1 N^{1/4}$,

$$(2.37)\qquad \int_{b_1 N^{\frac{1}{2}} \le |s| \le \pi N^{\frac{1}{2}}} |\psi(s,t,p)|ds \le B_2\, N^{-\beta_2 \log N} ,$$

$$(2.38)\qquad \int_{|s| \ge b_1 N^{\frac{1}{2}}} |\tilde{\psi}(s,t,p)|ds \le B_3\, N^{-\beta_3 \log N} .$$

For $N \ge N_0$, (2.37) is a consequence of lemma 2.2 and since $|\psi| \le 1$ we can choose $B_2$ so that (2.37) holds for all N. Because for all s and t

$$(2.39)\qquad |\tilde{\psi}(s,t,p)| \le \exp\{- \frac{\tau^2 t^2 + \sigma^2(s+\bar{a}t)^2}{2}\}M_4(t,s+\bar{a}t)$$

where $M_4$ is a polynomial depending only on c and C, (2.38) follows. Combining (2.11), (2.36), (2.37) and (2.38) we see that for $|t| \le b_1 N^{1/4}$

$$(2.40)\qquad |\nu(t,p) - \int_{-\infty}^{\infty} \tilde{\psi}(s,t,p)ds| \le B_4[(N^{-3/2}+N^{-5/4}|t|^5) \exp\{- \frac{ct^2}{8}\} + N^{-\beta_4 \log N}].$$

A direct application of lemma 2.2, the fact that $|\psi| \leq 1$ and (2.39) show that we can choose $B_4$ and $\beta_4$ so that (2.40) continues to hold for $b_1 N^{1/4} \leq |t| \leq bN^{3/2}$ with b as in lemma 2.2.

It remains to be shown that for all s and t

$$(2.41)\qquad \tilde{\nu}(t,p) = \int_{-\infty}^{\infty} \tilde{\psi}(s,t,p)ds.$$

This follows by straightforward but tedious computation using the fact that

$$(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \left(\frac{z}{\sigma(p)} + \frac{i\omega(p)}{\sigma^2(p)}\right)^k e^{-\frac{1}{2}z^2}\, dz = \mu_k(p) \quad \text{for even } k$$
$$= i\mu_k(p) \quad \text{for odd } k. \quad \square$$

We now turn to our asymptotic expansion for rejective sampling. For $1 \leq k \leq 6$, define functions $Q_k(p)$ by

$$(2.42)\qquad Q_1 = -\frac{N^{-\frac{1}{2}}}{2}\kappa_{3,1}\,\mu_2 + \frac{N^{-1}}{6}\left[\kappa_{4,1}\mu_3 - 3\kappa_{3,0}\kappa_{3,1}\left(2\,\frac{\omega}{\sigma^6} - \frac{\omega^3}{\sigma^8}\right)\right],$$

$$Q_2 = -\frac{N^{-\frac{1}{2}}}{2}\kappa_{3,2}\,\mu_1 + \frac{N^{-1}}{8}\left[-2\kappa_{4,2}\mu_2 + 2\kappa_{3,0}\kappa_{3,2}\left(\frac{1}{\sigma^4} - \frac{\omega^2}{\sigma^6}\right) + \kappa_{3,1}^2\mu_4\right],$$

$$Q_3 = \frac{N^{-\frac{1}{2}}}{6}\kappa_{3,3} + \frac{N^{-1}}{12}\left[-2\kappa_{4,3}\,\mu_1 + 3\kappa_{3,1}\,\kappa_{3,2}\,\mu_3\right],$$

$$Q_k = A_k, \qquad k = 4,5,6.$$

Let $\Phi$ and $\phi$ denote the standard normal d.f. and its density and let $H_k$ denote the Hermite polynomial of degree k, thus

$$(2.43)\qquad H_0(x) = 1, \quad H_1(x) = x, \quad H_2(x) = x^2 - 1, \quad H_3(x) = x^3 - 3x,$$

$$H_4(x) = x^4 - 6x^2 + 3, \qquad H_5(x) = x^5 - 10x^3 + 15x.$$

We shall show that expansions for (2.8) and (2.9) are given by

$$(2.44)\qquad \tilde{\rho}(t,p) = \exp\left\{-\frac{\tau^2(p)t^2}{2} - i\omega(p)\bar{a}(p)t\right\}\left[1 + \sum_{k=1}^{6} Q_k(p)(it)^k\right],$$

$$(2.45)\qquad \tilde{R}(x,p) = \Phi\left(\frac{x+\omega(p)\bar{a}(p)}{\tau(p)}\right) - \phi\left(\frac{x+\omega(p)\bar{a}(p)}{\tau(p)}\right)\sum_{k=1}^{6} \frac{Q_k(p)}{(\tau(p))^k} H_{k-1}\left(\frac{x+\omega(p)\bar{a}(p)}{\tau(p)}\right).$$

Note that $\tilde{\rho}$ is the Fourier-Stieltjes transform of $\tilde{R}$, i.e. $\tilde{\rho}(t,p) = \int e^{itx} d\tilde{R}(x,p)$.

THEOREM 2.1. *Suppose that positive numbers* c, C, D, $\delta$ *and* $\varepsilon$ *exist such that* (2.21) *and* (2.22) *are satisfied and*

$$(2.46)\qquad |\omega(p)| \le D.$$

*Then there exist positive numbers* $N_0$ *and* B *depending only on* c, C, D, $\delta$ *and* $\varepsilon$ *such that for* $N \ge N_0$, R(x,p) *is well-defined and*

$$(2.47)\qquad \sup_x |R(x,p) - \tilde{R}(x,p)| \le BN^{-5/4}.$$

PROOF. In this proof b, $B_i$, $\beta$, $\eta$ and $N_0$ denote appropriately chosen positive numbers depending only on c, C, D, $\delta$ and $\varepsilon$.

By (2.21), (2.25), (2.26), (2.46) and lemma 2.3 we have for $N \ge N_0$,

$$(2.48)\qquad |\tilde{\nu}(0,p)| \ge \eta, \qquad |\nu(0,p) - \tilde{\nu}(0,p)| \le \frac{\eta}{2},$$

so that $|\nu(0,p)| \ge \eta/2 > 0$. In the first place it follows that for $N \ge N_0$, $c(p) > 0$ and hence (2.7) is satisfied and R(x,p) is properly defined. We assume that $N \ge N_0$ and we shall show that, with b as in lemma 2.3,

$$(2.49)\qquad \int_{-bN^{3/2}}^{bN^{3/2}} \left|\frac{\rho(t,p) - \tilde{\rho}(t,p)}{t}\right| dt \le B_1 N^{-5/4}.$$

By Esseen's smoothing lemma (Esseen (1945)) this suffices to prove the theorem because $\tilde{R}(-\infty,p) = 0$, $\tilde{R}(\infty,p) = 1$ and the derivative of $\tilde{R}$ with respect to x is bounded.

By (2.21), (2.25), (2.26) and (2.46), $\tilde{\rho}$ has a bounded derivative with respect to t. Also

$$\left|\frac{d\rho(t,p)}{dt}\right| \le N^{-\frac{1}{2}} E(|\sum a_j(W_j-p_j)|\,|\sum W_j=n) \le N^{-\frac{1}{2}} \sum |a_j| \le C^{\frac{1}{4}} N^{\frac{1}{2}}.$$

Since $\rho(0,p) = \tilde{\rho}(0,t) = 1$, it follows that

$$\int_{-N^{-2}}^{N^{-2}} \left|\frac{\rho(t,p) - \tilde{\rho}(t,p)}{t}\right| dt \le B_2 N^{-3/2}. \tag{2.50}$$

Next we note that (2.21), (2.25) and (2.26) ensure that for all t

$$|\tilde{\nu}(t,p)| \le B_3 \exp\left\{-\frac{ct^2}{4}\right\}. \tag{2.51}$$

Together with (2.13), (2.48) and lemma 2.3 this implies that for $|t| \le bN^{3/2}$

$$\begin{aligned}\left|\rho(t,p) - \frac{\tilde{\nu}(t,p)}{\tilde{\nu}(0,p)}\right| &\le \frac{2}{\eta}\,|\nu(t,p) - \tilde{\nu}(t,p)| + \frac{2}{\eta^2}\,|\tilde{\nu}(t,p)| \\ &\qquad \cdot |\nu(0,p) - \tilde{\nu}(0,p)| \\ &\le B_4\left[(N^{-3/2}+N^{-5/4}|t|^5)\exp\left\{-\frac{ct^2}{8}\right\} + N^{-\beta\log N}\right].\end{aligned} \tag{2.52}$$

Again with the aid of (2.21), (2.25), (2.26) and (2.46) one can easily check that, for $1 \le k \le 6$, $Q_k$ is obtained from $A_k/A_0$ by expanding the denominator and discarding all terms of order $N^{-3/2}$, i.e. that $|Q_k - A_k/A_0| \le B_5 N^{-3/2}$. It follows that

$$\left|\tilde{\rho}(t,p) - \frac{\tilde{\nu}(t,p)}{\tilde{\nu}(0,p)}\right| \le B_6 N^{-3/2} \exp\left\{-\frac{ct^2}{4}\right\} \tag{2.53}$$

and combined with (2.52) this yields

$$\int_{N^{-2}\le|t|\le bN^{3/2}} \left|\frac{\rho(t,p)-\tilde{\rho}(t,p)}{t}\right| dt \le B_7(N^{-3/2}\log N+N^{-5/4}) \le B_8 N^{-5/4}. \tag{2.54}$$

Together with (2.50) this proves (2.49) and the theorem. □

Two remarks should be made with regard to theorem 2.1. The first one concerns condition (2.46) that does not occur in the preceding lemmas. The meaning of this condition is perhaps obscured by the fact that we make it do some odd jobs in the proof for which it is not really needed. We use it to show that (2.7) is satisfied for $N \geq N_0$, but (2.25) ensures that the number of indices j with $p_j = 0$ (or $p_j=1$) cannot exceed $m - C^{-1}c^2N + |\omega(p)|N^{\frac{1}{2}}$ (or $n - C^{-1}c^2N + |\omega(p)|N^{\frac{1}{2}}$) so that $|\omega(p)| \leq C^{-1}c^2N^{\frac{1}{2}}$ already implies (2.7) for all N. Condition (2.46) is also used to obtain (2.50), but in (2.50) we may replace $N^{-2}$ by an arbitrarily high power of $N^{-1}$ without doing any damage to the proof, and then the trivial bound $|\omega(p)| \leq N^{\frac{1}{2}}$ suffices. Finally we note that since

$$\text{(2.55)} \qquad \min(\lambda, 1-\lambda) \geq \sigma^2(p) - N^{-\frac{1}{2}}|\omega(p)|,$$

(2.46) forces $\lambda$ to be bounded away from 0 and 1 for large N, which is obviously important although it does not show up explicitly in the proof. However, here $|\omega(p)| \leq \frac{1}{2}C^{-1}c^2N^{\frac{1}{2}}$ would be sufficient.

The basic function of assumption (2.46), however, is to avoid a large (or intermediate) deviation situation that the condition $\sum W_j = n$ would get us into if $\omega(p) = N^{-\frac{1}{2}}(E\sum W_j - n)$ would not be bounded. Technically speaking this is reflected in the proof at the point where (2.46) is used to show that $\nu(0,p)$ is bounded away from zero. Also (2.46) ensures that (2.45) provides an expansion in powers of $N^{-\frac{1}{2}}$ to the required order.

To see what happens when condition (2.46) is relaxed, we prefer not to try to adapt the proof of theorem 2.1 but to answer this question more directly by remarking that the conditional distribution of $\sum a_jW_j$ given $\sum W_j = n$ remains unchanged if we replace p by $\tilde{p}$ where $\tilde{p}_j/(1-\tilde{p}_j) = \xi p_j/(1-p_j)$ $1 \leq j \leq N$, for some $0 \leq \xi \leq \infty$. If (2.7) is satisfied there exists a unique $\xi$ for which $\sum \tilde{p}_j = N\lambda$. Since $\omega(\tilde{p}) = 0$ it follows that if (2.21) and (2.22) are satisfied with p replaced by $\tilde{p}$, then (2.47) holds with $\tilde{R}(x,\tilde{p})$ instead of $\tilde{R}(x,p)$. Of course the snag is that in general $\tilde{p}$ can only be expressed analytically in terms of p as an infinite series. However, if $\omega(p) = O(N^{\alpha})$ for some $\alpha < \frac{1}{2}$, then a finite number of terms of this series will yield the required degree of accuracy and an explicit expansion for R(x,p) can be obtained. If $\alpha = 0$ this is expansion (2.45) but for $0 < \alpha < \frac{1}{2}$ more terms have to be included.

The second remark concerns the remainder $O(N^{-5/4})$ of our expansion. It is clear that by requiring that $\sum |a_j|^5 \leq CN$ in theorem 2.1 one obtains

$|R - \tilde{R}| \le BN^{-3/2} \log(N+1)$. Of course the "natural" order of the remainder is $O(N^{-3/2})$ and the factor $\log(N+1)$ is due only to technical difficulties in finding the conditional expectation of $\sum a_j W_j$ given $\sum W_j = n$.

The special case $p_j = \lambda$, $1 \le j \le N$, which is random sampling without replacement, is worth singling out because it corresponds to the null-hypothesis in the two-sample problem. Let $\bar{\lambda}$ denote the vector $(\lambda,\ldots,\lambda)$. For $p = \bar{\lambda}$, (2.45) simplifies to

$$(2.56)\qquad \tilde{R}(x,\bar{\lambda}) = \Phi\left(\frac{x}{\tau(\bar{\lambda})}\right) - \frac{\phi\left(\frac{x}{\tau(\bar{\lambda})}\right)}{\lambda(1-\lambda)}\left[\frac{\lambda(1-\lambda)}{2N} H_1\left(\frac{x}{\tau(\bar{\lambda})}\right)\right.$$
$$+ \frac{\{\lambda(1-\lambda)\}^{\frac{1}{2}}(1-2\lambda)}{6} \frac{\Sigma(a_j-a_.)^3}{\{\Sigma(a_j-a_.)^2\}^{3/2}} H_2\left(\frac{x}{\tau(\bar{\lambda})}\right)$$
$$+ \left\{\frac{1-6\lambda+6\lambda^2}{24} \frac{\Sigma(a_j-a_.)^4}{\{\Sigma(a_j-a_.)^2\}^2} - \frac{(1-2\lambda)^2}{8N}\right\} H_3\left(\frac{x}{\tau(\bar{\lambda})}\right)$$
$$\left. + \frac{(1-2\lambda)^2}{72} \frac{\{\Sigma(a_j-a_.)^3\}^2}{\{\Sigma(a_j-a_.)^2\}^3} H_5\left(\frac{x}{\tau(\bar{\lambda})}\right)\right],$$

where

$$(2.57)\qquad \tau^2(\bar{\lambda}) = \frac{\lambda(1-\lambda)}{N} \sum_{j=1}^{N} (a_j-a_.)^2,$$

$$(2.58)\qquad a_. = \bar{a}(\bar{\lambda}) = \frac{1}{N} \sum_{j=1}^{N} a_j .$$

Define, with $\ell$ denoting Lebesgue measure on $R^1$,

$$(2.59)\qquad \gamma(\zeta) = \ell\{x: \exists_j |x - a_j| < \zeta\}.$$

For $p = \bar{\lambda}$, theorem 2.1 yields

COROLLARY 2.1. *Suppose that positive numbers* c, C, $\delta$ *and* $\varepsilon$ *exist such that*

$$(2.60)\qquad \varepsilon \le \lambda \le 1 - \varepsilon,$$

$$(2.61)\qquad \frac{1}{N} \sum_{j=1}^{N} (a_j-a_.)^2 \ge c, \qquad \frac{1}{N} \sum_{j=1}^{N} a_j^4 \le C,$$

(2.62) $\gamma(\zeta) \geq \delta N\zeta$ *for some* $\zeta \geq N^{-3/2} \log N$.

*Then there exists* $B > 0$ *depending only on* c, C, $\delta$ *and* $\varepsilon$ *such that*

$$\sup_x |R(x,\bar{\lambda}) - \tilde{R}(x,\bar{\lambda})| \leq BN^{-5/4}.$$

Note that there is considerable further simplification in (2.56) if we either have almost equal sample sizes, i.e. $\lambda = \frac{1}{2} + O(N^{-3/4})$, or antisymmetric scores, i.e. $a_j + a_{N-j+1}$ is constant for all j. The latter happens for the locally most powerful rank test against shift alternatives when the underlying distribution is symmetric. In either case the $H_2$ and $H_5$ terms disappear so that the correction to the leading normal term is of order $N^{-1}$ only and is due solely to a correction to the variance, the $H_1$ term, and a kurtosis correction corresponding to $H_3$.

## 3. AN UNCONDITIONAL EXPANSION

We encounter several difficulties on the way to a usable unconditional expansion.

(i) The distribution of Z is awkward to handle analytically;

(ii) As in ABZ (1976), the random variables obtained by substituting P for p in $\tilde{\rho}$ or $\tilde{R}$ are generally not summable;

(iii) Again as in ABZ (1976), final simplification is not possible with our present techniques unless we assume that the sequence of alternatives is contiguous to the hypothesis as $N \to \infty$.

In this section we shall deal with the first two difficulties. Although we do not assume contiguity we shall be governed in the form of our expansion, which will involve polynomials in $(P_j-\lambda)$, in the number of terms that we calculate and in what we relegate to the remainder by the consideration that we expect $P_j = \lambda + \mathcal{O}_P(N^{-\frac{1}{2}})$ and $\sum (P_j-\lambda) = \mathcal{O}_P(1)$.

Recall that we assumed that $X_1,\ldots,X_N$ are independent, $X_1,\ldots,X_m$ having common density f and $X_{m+1},\ldots,X_N$ having density g. We shall write P for probabilities and E for expectations calculated under this model. In addition we need to consider an auxiliary model where $X_1,\ldots,X_N$ are i.i.d. with common density $h = (1-\lambda)f + \lambda g$ and d.f. $H = (1-\lambda)F + \lambda G$. We shall write $P_H$ for probabilities, $E_H$ for expectations and $\sigma_H^2$ for variances calculated under this second model.

To simplify our notation we assume from this point on that

$$(3.1) \qquad \sum_{j=1}^{N} a_j = 0.$$

Since $T = \sum (a_j-a_.)V_j + na_.$ it is obvious how all expansions need to be modified if (3.1) does not hold.

We meet difficulty (i) through

<u>LEMMA 3.1.</u>

$$(3.2) \qquad E \exp\{itN^{-\frac{1}{2}}T\} = \frac{E_H \nu(t,P)\exp\{itN^{-\frac{1}{2}}\sum a_j P_j\}}{2\pi N^{\frac{1}{2}} B_{N,n}(\lambda)},$$

*where*

$$B_{N,n}(\lambda) = \binom{N}{n}\lambda^n(1-\lambda)^{N-n}.$$

PROOF. Under our original model the density of Z at the point $z = (z_1,\ldots,z_N)$ with $z_1 < z_2 < \ldots < z_N$ is given by

$$\sum \prod_{j=1}^{m} f(z_{i_j}) \prod_{j=m+1}^{N} g(z_{i_j}) ,$$

where the sum ranges over all permutations $i_1,\ldots,i_N$ of $1,\ldots,N$. Under our second model this density is

$$N! \prod_{j=1}^{N} [(1-\lambda)f(z_j) + \lambda g(z_j)] .$$

By the Radon-Nikodym theorem and lemma 2.1,

$$E \exp\{itN^{-\frac{1}{2}}T\} = E \frac{\nu(t,P)}{\nu(0,P)} \exp\{itN^{-\frac{1}{2}} \sum a_j P_j\}$$

$$= E_H \frac{\nu(t,P)}{\nu(0,P)} \exp\{itN^{-\frac{1}{2}} \sum a_j P_j\} \sum \prod_{j=1}^{m} \frac{f(Z_{i_j})}{h(Z_{i_j})} \prod_{j=m+1}^{N} \frac{g(Z_{i_j})}{h(Z_{i_j})} \frac{1}{N!}$$

$$= [B_{N,n}(\lambda)]^{-1} E_H \frac{\nu(t,P)}{\nu(0,P)} \exp\{itN^{-\frac{1}{2}} \sum a_j P_j\} c(P),$$

where c is defined by (2.6) or (2.12). The lemma follows from (2.11) and (2.13). □

Lemma 3.1 shows that we are concerned with $\tilde{\nu}$ rather than $\tilde{\rho}$, but since $\tilde{\nu}$ as a function of P is no more summable than $\tilde{\rho}$, we still have to face difficulty (ii). We do this by showing that $\tilde{\nu}$ may be replaced by a summable function $\nu^*$ outside a set that will later be seen to have sufficiently small probability. Define

(3.3) $$\nu^*(t,p) = \left[\frac{2\pi}{\lambda(1-\lambda)}\right]^{\frac{1}{2}} \exp\left\{-\frac{\lambda(1-\lambda)}{2N} \sum a_j^2 t^2\right\} \sum_{k=0}^{6} A_k^*(p)(it)^k,$$

where

$$(3.4)\qquad A_0^*(p) = 1 + \frac{1}{2\lambda(1-\lambda)N}\left[\sum (p_j-\lambda)^2 - \left\{\sum (p_j-\lambda)\right\}^2 - \frac{1-\lambda+\lambda^2}{6}\right],$$

$$A_1^*(p) = N^{-3/2}\sum a_j p_j\left[1 - \frac{1-2\lambda}{\lambda(1-\lambda)}\sum (p_j-\lambda)\right],$$

$$A_2^*(p) = \frac{(1-2\lambda)}{2N}\sum a_j^2(p_j-\lambda) - \frac{\Sigma a_j^2}{2N^2}\left[(1-2\lambda)\sum(p_j-\lambda) - \lambda(1-\lambda)\right]$$

$$- \frac{1}{2N}\sum a_j^2(p_j-\lambda)^2 - \frac{(1-2\lambda)^2}{2\lambda(1-\lambda)N^2}\left\{\sum a_j p_j\right\}^2,$$

$$A_3^*(p) = \frac{N^{-3/2}}{6}\left[\lambda(1-\lambda)(1-2\lambda)\sum a_j^3 + (1-6\lambda+6\lambda^2)\sum a_j^3(p_j-\lambda)\right.$$

$$\left. - \frac{3}{N}(1-2\lambda)^2\sum a_j^2\sum a_j p_j\right],$$

$$A_4^*(p) = \frac{\lambda(1-\lambda)(1-6\lambda+6\lambda^2)}{24N^2}\sum a_j^4 - \frac{\lambda(1-\lambda)(1-2\lambda)^2}{8N^3}\left\{\sum a_j^2\right\}^2$$

$$+ \frac{(1-2\lambda)^2}{8N^2}\left\{\sum a_j^2(p_j-\lambda)\right\}^2,$$

$$A_5^*(p) = \frac{N^{-5/2}}{12}\lambda(1-\lambda)(1-2\lambda)^2\sum a_j^3\sum a_j^2(p_j-\lambda),$$

$$A_6^*(p) = \frac{\lambda^2(1-\lambda)^2(1-2\lambda)^2}{72N^3}\left\{\sum a_j^3\right\}^2.$$

<u>LEMMA 3.2</u>. *Suppose that* (3.1) *holds and that positive numbers* c, C *and* $\varepsilon$ *exist such that* (2.21) *is satisfied and*

$$(3.5)\qquad \varepsilon \le \lambda \le 1-\varepsilon.$$

*Then there exist positive numbers* B *and* $\beta$ *depending only on* c, C *and* $\varepsilon$ *such that*

$$(3.6)\qquad |\tilde{\nu}(t,p) - \nu^*(t,p)| \le B\exp\{-\beta t^2\}\left[\{N^{-3/2} + N^{-5/4}|t|\}\{1 + N\sum(p_j-\lambda)^4\}\right.$$

$$\left. + N^{-3/2}\left\{\sum (p_j-\lambda)^4\right\}\right].$$

PROOF. For simplicity we make use of order symbols in this proof and $O(x)$ will denote a quantity that is bounded by $B_1|x|$ where $B_1$ depends only on c, C and $\varepsilon$.

Suppose first that $|\omega(p)| > 1$. Then (2.21) and (3.5) are easily seen to imply that $|\nu^*(t,p)| = O(\omega^2(p)\exp\{-\varepsilon(1-\varepsilon)ct^2/4\})$, whereas for $\tilde{\nu}(t,p)$ we have the bound (2.51). The right-hand side of (3.6), however, contains a term $BN^{\frac{1}{2}}\omega^4(p)\exp\{-\beta t^2\}$ so that the lemma is trivial for $|\omega(p)| > 1$.

We therefore assume that $|\omega(p)| \le 1$. Noting that $\sigma^2(p)$ is bounded away from zero (c.f. (2.25)), we expand $\sigma^{-2}$, $\bar{a}$, $\tau^2$ and $\kappa_{r,i}$ about the point $p_j = \lambda$, $1 \le j \le N$, using elementary inequalities to bound the remainders in terms of N and

$$M_1 = N^{-1} \sum (p_j-\lambda)^4, \qquad M_2 = N^{-1} \left|\sum(p_j-\lambda)\right|.$$

We find

$$\frac{1}{\sigma^2(p)} = \frac{1}{\lambda(1-\lambda)}\left[1 - \frac{(1-2\lambda)}{\lambda(1-\lambda)N}\sum (p_j-\lambda) + \frac{1}{\lambda(1-\lambda)N}\sum (p_j-\lambda)^2\right]$$

$$+ O(M_1+M_2^2) = \frac{1}{\lambda(1-\lambda)} + O(M_1^{\frac{1}{2}}+M_2) ,$$

$$\bar{a}(p) = \frac{(1-2\lambda)}{\lambda(1-\lambda)N}\sum a_j p_j + O(M_1^{\frac{1}{2}}) = O(M_1^{\frac{1}{4}}) ,$$

$$\tau^2(p) = \frac{\lambda(1-\lambda)}{N}\sum a_j^2 + \frac{(1-2\lambda)}{N}\sum a_j^2(p_j-\lambda) - \frac{1}{N}\sum a_j^2(p_j-\lambda)^2$$

$$- \frac{(1-2\lambda)^2}{\lambda(1-\lambda)N^2}\left\{\sum a_j p_j\right\}^2 + O(M_1^{3/4}) ,$$

$$\kappa_{3,0}(p) = \lambda(1-\lambda)(1-2\lambda) + O(M_1^{\frac{1}{2}}+M_2) ,$$

$$\kappa_{3,1}(p) = -\frac{2\lambda(1-\lambda)}{N}\sum a_j p_j + O(M_1^{\frac{1}{2}}) ,$$

$$\kappa_{3,2}(p) = \frac{\lambda(1-\lambda)(1-2\lambda)}{N}\sum a_j^2 + O(M_1^{\frac{1}{4}}) ,$$

$$\kappa_{3,3}(p) = \frac{\lambda(1-\lambda)(1-2\lambda)}{N} \sum a_j^3 + \frac{(1-6\lambda+6\lambda^2)}{N} \sum a_j^3(p_j-\lambda)$$

$$- \frac{3(1-2\lambda)^2}{N^2} \sum a_j^2 \sum a_j p_j + O(N^{\frac{1}{4}}M_1^{\frac{1}{2}})(=O(1)) ,$$

$$\kappa_{4,0}(p) = \lambda(1-\lambda)(1-6\lambda+6\lambda^2) + O(M_1^{\frac{1}{4}}) ,$$

$$\kappa_{4,1}(p) = O(M_1^{\frac{1}{4}}) ,$$

$$\kappa_{4,2}(p) = \frac{\lambda(1-\lambda)(1-6\lambda+6\lambda^2)}{N} \sum a_j^2 + O(M_1^{\frac{1}{4}}) ,$$

$$\kappa_{4,3}(p) = O(1) ,$$

$$\kappa_{4,4}(p) = \frac{\lambda(1-\lambda)(1-6\lambda+6\lambda^2)}{N} \sum a_j^4 + O(N^{\frac{1}{4}}M_1^{\frac{1}{4}})(=O(1)).$$

To illustrate the computations involved we present the argument for $\kappa_{3,3}$. By (2.21), the result for $\bar{a}(p)$ and the fact that $0 \le M_1 \le 1$, we have

$$\kappa_{3,3}(p) = N^{-1} \sum p_j(1-p_j)(1-2p_j)a_j^3 - 3N^{-1}\bar{a}(p) \sum p_j(1-p_j)(1-2p_j)a_j^2$$

$$+ O(M_1^{\frac{1}{2}}) = N^{-1}\lambda(1-\lambda)(1-2\lambda) \sum a_j^3 + N^{-1}(1-6\lambda+6\lambda^2) \sum a_j^3(p_j-\lambda)$$

$$- 3N^{-2}(1-2\lambda)^2 \sum a_j^2 \sum a_j p_j + O(M_1^{\frac{1}{2}}+N^{-1}\sum|a_j|^3(p_j-\lambda)^2$$

$$+ N^{-1}M_1^{\frac{1}{4}}\sum a_j^2|p_j-\lambda|).$$

Hölder's inequality and (2.21) imply that

$$N^{-1}M_1^{1/4}\sum a_j^2|p_j - \lambda| \le N^{-1}M_1^{1/4}\left(\sum|a_j|^{8/3}\right)^{3/4}(NM_1)^{1/4} = O(M_1^{1/2}) ,$$

$$N^{-1}\sum|a_j|^3(p_j-\lambda)^2 \le N^{-1}(CN)^{1/4}\sum a_j^2(p_j-\lambda)^2 \le C^{1/4}N^{-3/4}(NM_1\sum a_j^4)^{1/2}$$

$$= O(N^{1/4}M_1^{1/2}).$$

As $\bar{a}(p)$ is bounded, $\kappa_{3,3}(p)$ is obviously also $O(1)$. Note that the a-typical order of the remainder $O(N^{\frac{1}{4}}M_1^{\frac{1}{2}})$ originates from the term $O(N^{-1}\sum|a_j|^3(p_j-\lambda)^2)$ where we have to sacrifice a factor $O(N^{-\frac{1}{4}})$ in order to apply Hölder's inequality and (2.21). The same thing occurs for $\kappa_{4,4}(p)$.

For $\mu_k(p)$ defined by (2.28) we find

$$\mu_1(p) = \frac{1}{\lambda(1-\lambda)N^{\frac{1}{2}}} \sum (p_j-\lambda) + O(M_1^{\frac{1}{2}}+N^{\frac{1}{2}}M_2^2) ,$$

$$\mu_2(p) = \frac{1}{\lambda(1-\lambda)} + O(M_1^{\frac{1}{2}}+M_2+NM_2^2) ,$$

$$\mu_3(p) = \frac{3}{\lambda^2(1-\lambda)^2N^{\frac{1}{2}}} \sum (p_j-\lambda) + O(M_1^{\frac{1}{2}}+NM_2^2) ,$$

$$\mu_4(p) = \frac{3}{\lambda^2(1-\lambda)^2} + O(M_1^{\frac{1}{2}}+M_2+NM_2^2) ,$$

$$\mu_5(p) = O(N^{\frac{1}{2}}M_2) ,$$

$$\mu_6(p) = \frac{15}{\lambda^3(1-\lambda)^3} + O(M_1^{\frac{1}{2}}+M_2+NM_2^2) .$$

Straightforward but tedious calculation now yields

(3.7)
$$\sum_{k=0}^{6} A_k(p)(it)^k = \left[1 + \frac{(1-2\lambda)}{2\lambda(1-\lambda)N} \sum (p_j-\lambda) - \frac{(1-\lambda+\lambda^2)}{12\lambda(1-\lambda)N}\right]$$
$$+ \frac{\Sigma a_j p_j}{N^{3/2}} it - \frac{\Sigma a_j^2}{2N^2}\left[(1-2\lambda)\sum(p_j-\lambda) - \lambda(1-\lambda)\right](it)^2$$
$$+ \frac{1}{6N^{3/2}}\left[\lambda(1-\lambda)(1-2\lambda)\sum a_j^3+(1-6\lambda+6\lambda^2)\sum a_j^3(p_j-\lambda) - \frac{3(1-2\lambda)^2}{N}\sum a_j^2 \sum a_j p_j\right](it)^3$$
$$+ \frac{\lambda(1-\lambda)}{24N^2}\left[(1-6\lambda+6\lambda^2) \sum a_j^4 - \frac{3(1-2\lambda)^2}{N}\left\{\sum a_j^2\right\}^2\right](it)^4$$
$$+ \frac{\lambda^2(1-\lambda)^2(1-2\lambda)^2}{72N^3}\left\{\sum a_j^3\right\}^2 (it)^6 + O((|t|^3+t^4)[N^{-5/4} + N^{-1/4}M_1^{1/2}]$$
$$+ (1+t^6)[N^{-3/2} + N^{-1/2} M_1^{1/2} + N^{1/2} M_2^2]) .$$

Next we expand the remaining factor in (2.30). Because both $\tau^2(p)$ and its leading term $\lambda(1-\lambda)N^{-1}\sum a_j^2$ are bounded away from zero, there exists $\beta > 0$ depending only on c, C and D, such that

$$\frac{(2\pi)^{\frac{1}{2}}}{\sigma(p)} \exp\left\{-\frac{\omega^2(p)}{2\sigma^2(p)} - \frac{\tau^2(p)t^2}{2} - i\omega(p)\bar{a}(p)t\right\}$$

$$= \left[\frac{2\pi}{\lambda(1-\lambda)}\right]^{\frac{1}{2}} \exp\left\{-\frac{\lambda(1-\lambda)}{2N}\sum a_j^2 t^2\right\}$$

$$\times \left[1 - \frac{1}{2\lambda(1-\lambda)N}\left\{(1-2\lambda)\sum(p_j-\lambda) - \sum(p_j-\lambda)^2 + \{\sum(p_j-\lambda)\}^2\right\}\right.$$

$$- \frac{(1-2\lambda)}{\lambda(1-\lambda)N^{3/2}}\sum(p_j-\lambda)\sum a_j p_j (it) + \frac{1}{2N}\left\{(1-2\lambda)\sum a_j^2(p_j-\lambda)\right.$$

$$\left.- \sum a_j^2(p_j-\lambda)^2 - \frac{(1-2\lambda)^2}{\lambda(1-\lambda)N}\{\sum a_j p_j\}^2\right\}(it)^2$$

$$\left.+ \frac{(1-2\lambda)^2}{8N^2}\{\sum a_j^2(p_j-\lambda)\}^2(it)^4\right]$$

$$+ O\left(\exp\{-\beta t^2\}\left[N^{-3/2} + N^{1/2}M_1 + N^{5/2}M_2^4\right]\right) .$$

Multiplication by (3.7) yields (3.6). □

Here is our first unconditional expansion. Define

$$(3.8)\qquad \rho(t) = E\exp\{itN^{-\frac{1}{2}}T\} ,$$

$$(3.9)\qquad \rho^*(t) = \exp\left\{-\frac{\lambda(1-\lambda)}{2N}\sum a_j^2 t^2\right\} E_H\left[\exp\{itN^{-\frac{1}{2}}\sum a_j P_j\}\right.$$

$$\left.\times\left\{1 + \frac{1}{2\lambda(1-\lambda)N}\left(\sum(P_j-\lambda)^2 - \{\sum(P_j-\lambda)\}^2\right) + \sum_{k=1}^{6} A_k^*(P)(it)^k\right\}\right] .$$

<u>LEMMA 3.3</u>. *Suppose that* (3.1) *holds and that positive numbers* c, C, δ, δ' *and* ε *exist with* $\delta' < \min(\frac{1}{2}, \delta/2, c^2C^{-1}/4)$ *and such that* (2.62) *is satisfied and*

$$(3.10)\qquad \frac{1}{N}\sum a_j^2 \ge c, \qquad \frac{1}{N}\sum a_j^4 \le C,$$

$$(3.11)\qquad P_H\left(\varepsilon \le \frac{\lambda g(X_1)}{h(X_1)} \le 1-\varepsilon\right) \ge 1 - \delta' .$$

*Then there exist positive numbers* b, B, $\beta_1$ *and* $\beta_2$ *depending only on* c, C, $\delta$, $\delta'$ *and* $\varepsilon$ *such that for* $|t| \le bN^{3/2}$,

$$(3.12)\qquad |\rho(t) - \rho^*(t)| \le B\left[\exp\{-\beta_1 t^2\}(N^{-3/2}+N^{-5/4}|t|)\left\{1 + N^2 E_H\left(\frac{g(X_1)}{h(X_1)} - 1\right)^4\right\} + N^{-\beta_2 \log N}\right] .$$

PROOF. In this proof we again use $O$ symbols that are uniform for fixed c, C, $\delta$, $\delta'$ and $\varepsilon$. Note that $E_H\{g(X_1)/h(X_1)\} = 1$, so that (3.11) and Markov's inequality ensure that $\min(\lambda, 1-\lambda) \ge \varepsilon(1-\delta')$.

Take a number $\delta'' \in (\delta', \min(1/2, \delta/2, c^2C^{-1}/4)$ and define the event E by

$$E = \{\varepsilon \le P_j \le 1 - \varepsilon \text{ for at least } (1-\delta'')N \text{ indices } j\} =$$
$$= \{\varepsilon \le \frac{\lambda g(X_j)}{h(X_j)} \le 1 - \varepsilon \text{ for at least } (1-\delta'')N \text{ indices } j\} .$$

Applying an exponential bound for binomial probabilities (c.f. Okamoto (1958)) we find that (3.11) implies

$$(3.13)\qquad P_H(E) \ge 1 - \exp\{-2N(\delta''-\delta')^2\} .$$

Because $\lambda$ and $(1-\lambda)$ are bounded away from 0, the same is true for $N^{\frac{1}{2}} B_{N,n}(\lambda)$. Also, (2.10) and (2.11) imply that $|\nu(t,p)| \le 2\pi N^{\frac{1}{2}}$ for all t and p. Hence application of lemma 3.1 shows that

$$(3.14)\qquad \rho(t) = \frac{E_H \nu(t,P)\exp\{itN^{-\frac{1}{2}}\sum a_j P_j\}\chi_E}{2\pi N^{\frac{1}{2}} B_{N,n}(\lambda)} + O(\exp\{-N(\delta''-\delta')^2\}) ,$$

where $\chi_E$ denotes the indicator of E.

Since $\delta'' < \delta/2$, (2.62) ensures the validity of (2.22) on the set E with $\delta$ replaced by $\delta - 2\delta''$. If $\sum'$ denotes summation over those indices j for which $P_j \notin [\varepsilon, 1-\varepsilon]$ and k denotes the number of these indices, then $k \le \delta'' N$ on E and as a result

$$\tau^2(P) \ge \frac{\varepsilon(1-\varepsilon)}{N}\left[\sum_{j=1}^{N} (a_j - \bar{a}(P))^2 - \sum{}' (a_j - \bar{a}(P))^2\right]$$

$$\ge \frac{\varepsilon(1-\varepsilon)}{N}\left[\sum_{j=1}^{N} a_j^2 + N\{\bar{a}(P)\}^2 - 2\sum{}' a_j^2 - 2k\{\bar{a}(P)\}^2\right]$$

$$\ge \frac{\varepsilon(1-\varepsilon)}{N}[cN - 2\{k\sum{}' a_j^4\}^{\frac{1}{2}}] \ge \varepsilon(1-\varepsilon)[c - 2\{\delta'' C\}^{\frac{1}{2}}] > 0$$

on E, because $\delta'' < \min(\frac{1}{2}, c^2 C^{-1}/4)$.

We have shown that on the set E, a and P satisfy the conditions on a and p in lemmas 2.3 and 3.2. Combining (3.14), (2.31) and (3.6) we obtain

$$(3.15) \qquad \rho(t) = \frac{E_H \nu^*(t,P)\exp\{itN^{-\frac{1}{2}}\sum a_j P_j\}\chi_E}{2\pi N^{\frac{1}{2}} B_{N,n}(\lambda)} + O(N^{-\beta_2 \log N}$$

$$+ \exp\{-\beta_1 t^2\}[\{N^{-3/2} + N^{-5/4}|t|\}\{1 + NE_H \sum (P_j - \lambda)^4\}$$

$$+ N^{-3/2} E_H\{\sum (P_j - \lambda)\}^4])$$

for $|t| \le bN^{3/2}$, where b, $\beta_1$ and $\beta_2$ depend on c, C, $\delta$, $\delta'$ and $\varepsilon$ only.

Because of (3.13) and the fact that $\nu^*(t,p) = O(N)$, (3.15) remains valid if we delete $\chi_E$. Using

$$2\pi N^{\frac{1}{2}} B_{N,n}(\lambda) = \left[\frac{2\pi}{\lambda(1-\lambda)}\right]^{\frac{1}{2}}\left(1 - \frac{1-\lambda+\lambda^2}{12\lambda(1-\lambda)N}\right) + O(N^{-2})$$

one easily verifies that in (3.15) the first term on the right may be replaced by $\rho^*(t)$ without changing the order of the remainder. Since

$$E_H \sum (P_j - \lambda)^4 = E_H \sum \left(\frac{\lambda g(X_i)}{h(X_i)} - \lambda\right)^4 = \lambda^4 N E_H \left(\frac{g(X_1)}{h(X_1)} - 1\right)^4 ,$$

$$E_H\{\sum (P_j - \lambda)\}^4 = \lambda^4 E_H\left\{\sum \left(\frac{g(X_i)}{h(X_i)} - 1\right)\right\}^4 \le 3\lambda^4 N^2 E_H \left(\frac{g(X_1)}{h(X_1)} - 1\right)^4 ,$$

the proof of the lemma is complete. □

Define

$$(3.16)\qquad \pi_j = E_H P_j, \qquad \pi = (\pi_1,\ldots,\pi_N) .$$

In the remaining part of this section we obtain a further expansion for $\rho(t)$ and convert this expansion into one for the d.f. of T. Although we still do not assume contiguity, we shall be guided in what terms we include in the remainder by the fact that under contiguous alternatives we expect $(P_j-\pi_j)$ to behave roughly like $O_{P_H}(N^{-1})$. Let

$$(3.17)\qquad K(x) = \Phi(x) - \phi(x) \sum_{k=0}^{5} \alpha_k H_k(x) ,$$

where $\Phi$ and $\phi$ denote the standard normal d.f. and its density, the Hermite polynomials $H_k$ are given by (2.43) and

$$(3.18)\qquad \alpha_0 = \frac{\sum a_j\pi_j}{\{\lambda(1-\lambda)\sum a_j^2\}^{\frac12}N} ,$$

$$\alpha_1 = \frac{\sigma_H^2(\Sigma a_jP_j)-\Sigma a_j^2E_H(P_j-\lambda)^2+(1-2\lambda)\Sigma a_j^2(\pi_j-\lambda)}{2\lambda(1-\lambda)\Sigma a_j^2} - \frac{(1-2\lambda)^2\{\Sigma a_j\pi_j\}^2}{2\lambda^2(1-\lambda)^2N\Sigma a_j^2} + \frac{1}{2N},$$

$$\alpha_2 = \frac{[\lambda(1-\lambda)(1-2\lambda)\Sigma a_j^3+(1-6\lambda+6\lambda^2)\Sigma a_j^3(\pi_j-\lambda)-3(1-2\lambda)^2N^{-1}\Sigma a_j^2\Sigma a_j\pi_j]}{6\{\lambda(1-\lambda)\Sigma a_j^2\}^{3/2}} ,$$

$$\alpha_3 = \frac{\lambda(1-\lambda)(1-6\lambda+6\lambda^2)\Sigma a_j^4-3\lambda(1-\lambda)(1-2\lambda)^2N^{-1}\{\Sigma a_j^2\}^2+3(1-2\lambda)^2\{\Sigma a_j^2(\pi_j-\lambda)\}^2}{24\{\lambda(1-\lambda)\Sigma a_j^2\}^2},$$

$$\alpha_4 = \frac{(1-2\lambda)^2\Sigma a_j^3\Sigma a_j^2(\pi_j-\lambda)}{12\{\lambda(1-\lambda)\}^{3/2}\{\Sigma a_j^2\}^{5/2}} ,$$

$$\alpha_5 = \frac{(1-2\lambda)^2\{\Sigma a_j^3\}^2}{72\lambda(1-\lambda)\{\Sigma a_j^2\}^3} .$$

THEOREM 3.1. *Suppose that* (3.1) *holds and that positive numbers* c, C, $\delta$ *and* $\varepsilon$ *exist such that* (3.10) *and* (2.62) *are satisfied and*

(3.19) $\qquad \varepsilon \le \lambda \le 1 - \varepsilon$ .

*Then there exists* B > 0 *depending only on* c, C, δ *and* ε *such that*

$$(3.20) \qquad \sup_x \left| P\left(\frac{T}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac12}} \le x\right) - K\left(x - \frac{\Sigma a_j\pi_j}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac12}}\right)\right|$$
$$\le B\left\{N^{-5/4} + N^{3/4}E_H\left(\frac{g(X_1)}{h(X_1)} - 1\right)^4 + N^{-\frac12}\left[\sum\{E_H|P_j-\pi_j|^3\}^{4/9}\right]^{9/4}\right\} .$$

PROOF. In this proof $B_i$ and $\beta_i$ denote appropriately chosen positive numbers depending only on c, C, δ and ε. We shall have to consider the r.v.

$$(3.21) \qquad U = N^{-\frac12} \sum a_j(P_j-\pi_j)$$

and we note that

$$(3.22) \qquad E_H|U|^3 \le N^{-3/2}\left[\sum|a_j|\{E_H|P_j-\pi_j|^3\}^{1/3}\right]^3$$
$$\le C^{3/4}N^{-3/4}\left[\sum\{E_H|P_j-\pi_j|^3\}^{4/9}\right]^{9/4} .$$

Since $\sup_x(1+|K(x)|) \le B_1(1+E_HU^2) \le B_1(2+E_H|U|^3)$ we may assume without loss of generality that $E_H|U|^3 \le 1$, because otherwise (3.20) is satisfied trivially for $B = 3B_1C^{3/4}$. Hence $\sup_x(1+|K(x)|) \le 3B_1$ and similar bounds $|\alpha_k| \le B_2(1+E_HU^2) \le 3B_2$ and $\sup_x|K'(x)| \le 3B_3$ hold for $\alpha_0,\ldots,\alpha_5$ and for the derivative K' of K.

Take $\delta' = \min(1/4, \delta/4, c^2C^{-1}/8)$. In view of $1 + |K| \le 3B_1$ it is again no loss of generality to assume that $E_H(g(X_1)/h(X_1)-1)^4 \le \delta'\varepsilon^4/16$, because otherwise (3.22) with $B = 48B_1/(\delta'\varepsilon^4)$ is trivially true. Hence by (3.19) and Markov's inequality

$$P_H\left(\tfrac12\varepsilon \le \frac{\lambda g(X_1)}{h(X_1)} \le 1-\tfrac12\varepsilon\right) \ge P_H\left(\left|\frac{g(X_1)}{h(X_1)} - 1\right| \le \tfrac12\varepsilon\right) \ge 1 - \delta' ,$$

so that the conditions of lemma 3.3 are satisfied and (3.12) holds.

The proof hinges on the expansion

$$\exp\{itN^{-\frac{1}{2}}\sum a_jP_j\} = \exp\{itN^{-\frac{1}{2}}\sum a_j\pi_j\}[1+itU+\tfrac{1}{2}(itU)^2] + O(|tU|^3)$$

and its truncation to fewer terms. We apply this expansion to (3.9) and in the resulting expression we replace P by $\pi$ wherever this is possible without giving rise to remainder terms that would be awkward to handle at this point. Using elementary inequalities to seperate out and bound those parts of the remainder that depend on the $(P_j-\lambda)$ rather than on the $(P_j-\pi_j)$, we arrive at

$$(3.23)\qquad |\rho^*(t)-\bar\rho(t)| \le B_4|t|\exp\{-\beta_3t^2\}\Big[N^{-3/2} + N^{1/2}E_H\Big(\frac{g(X_1)}{h(X_1)} -1\Big)^4 + E_H|U|^3 + N^{-1}E_H|U\sum a_j^2(P_j-\pi_j)| + N^{-2}E_H\{\sum a_j^2(P_j-\pi_j)\}^2\Big] ,$$

$$(3.24)\qquad \bar\rho(t) = \exp\Big\{itN^{-\frac{1}{2}}\sum a_j\pi_j - t^2\frac{\lambda(1-\lambda)}{2N}\sum a_j^2\Big\}\Big[1 + \sum_{k=1}^{6}\alpha_{k-1}\Big(\frac{\lambda(1-\lambda)\Sigma a_j^2}{N}\Big)^{\frac{1}{2}k}(it)^k\Big] .$$

Because $\max_j|a_j| \le (CN)^{\frac{1}{4}}$ we find by the same reasoning as in (3.22),

$$N^{-1}E_H|U\sum a_j^2(P_j-\pi_j)| + N^{-2}E_H\{\sum a_j^2(P_j-\pi_j)\}^2 \le B_5N^{-5/4}E_H\{\sum|a_j(P_j-\pi_j)|\}^2$$

$$\le B_5N^{-5/4}[1 + E_H\{\sum|a_j(P_j-\pi_j)|\}^3]$$

$$\le B_5N^{-5/4} + B_6N^{-1/2}\Big[\sum\{E_H|P_j - \pi_j|^3\}^{4/9}\Big]^{9/4} .$$

Together with (3.22) this shows that (3.23) may be reduced to

$$(3.25)\qquad |\rho^*(t) - \bar\rho(t)| \le B_7|t|\exp\{-\beta_3t^2\}\Big\{N^{-5/4} + N^{1/2}E_H\Big(\frac{g(X_1)}{h(X_1)} - 1\Big)^4 + N^{-1/2}\Big[\sum\{E_H|P_j-\pi_j|^3\}^{4/9}\Big]^{9/4}\Big\} .$$

As $\alpha_0,\ldots,\alpha_5$ are bounded and $N^{-\frac{1}{2}}|\sum a_j\pi_j| \le C^{\frac{1}{4}}N^{\frac{1}{2}}$, we have $|\bar\rho'(t)| \le B_8N^{\frac{1}{2}}$ for all t. Since $|\rho'(t)| \le N^{-\frac{1}{2}}E|T| \le C^{\frac{1}{4}}N^{\frac{1}{2}}$ for all t and $\rho(0) = \bar\rho(0) = 1$, (3.26) $$|\rho(t) - \bar{\rho}(t)| \le B_9 N^{\frac{1}{2}} |t| \qquad \text{for all } t.$$

Combining lemma 3.3, (3.25) and (3.26) we find

(3.27) $$\int_{-bN^{3/2}}^{bN^{3/2}} \left|\frac{\rho(t)-\bar{\rho}(t)}{t}\right| dt \le B_9 N^{-3/2} + \int_{N^{-2} \le |t| \le bN^{3/2}} \left|\frac{\rho(t)-\bar{\rho}(t)}{t}\right| dt$$

$$\le B_{10}\left\{N^{-5/4} + N^{3/4} E_H\left(\frac{g(X_1)}{h(X_1)} - 1\right)^4 + N^{-\frac{1}{2}}\left[\sum \{E_H |P_j - \pi_j|^3\}^{4/9}\right]^{9/4}\right\}.$$

Now $\bar{\rho}(t)$ is the Fourier-Stieltjes transform of $K(\{N^{\frac{1}{2}}x - \sum a_j \pi_j\}\{\lambda(1-\lambda)\sum a_j^2\}^{-\frac{1}{2}})$ as a function of x. This is a function of bounded variation assuming the values 0 and 1 at $-\infty$ and $+\infty$ and having a derivative that is bounded by $3B_3 c^{-\frac{1}{2}}\{\varepsilon(1-\varepsilon)\}^{-1}$ in absolute value. It follows from the smoothing lemma (ESSEEN (1945)) that

$$\sup_x \left| P(N^{-\frac{1}{2}} T \le x) - K\left(\frac{N^{\frac{1}{2}}x - \Sigma a_j \pi_j}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac{1}{2}}}\right)\right|$$

is bounded above by the right-hand side of (3.20). A change of scale completes the proof. □

Theorem 3.1 provides the basic expansion for the distribution of T under contiguous alternatives. Only first and second moments of functions of order statistics remain to be determined. In section 4 we shall be concerned with a further simplification of the expansion and a precise evaluation of the order of the remainder. With regard to this remainder we are in a seemingly less favorable position than we were at the same stage in the one-sample problem (c.f. ABZ (1976), theorem 2.3), because the third remainder term in (3.20) is larger than the corresponding term in the one-sample case by a factor $N^{\frac{1}{4}}$. This is due to the appearance of the remainder term $N^{-1} E_H |U \sum a_j^2 (P_j - \pi_j)|$ that does not occur for the one-sample statistic. It will turn out, however, that we shall need only a slightly stronger condition than before to show that the remainder is still $\mathcal{O}(N^{-5/4})$.

The conditions of theorem 3.1 concern only the sample ratio $\lambda$ and the scores a. There are no assumptions about the underlying densities f and g but this is merely a trick; obviously something like contiguity is needed to make the expansion meaningful in the sense that the remainder is at all small. With regard to the conditions on the scores, (3.10) acts as a safeguard against too rapid growth and (2.62) ensures that the $a_j$ do not cluster too much around too few points, thus preventing a too pronounced lattice character of the distribution of T, as was pointed out in ABZ (1976). It was also noted there that in the important case of exact scores $a_j = EJ(U_{j:N})$, with $U_{1:N} < U_{2:N} < \dots < U_{N:N}$ order statistics from the uniform distribution on (0,1), both (3.10) and (2.62) will be satisfied for all N with fixed c, C and $\delta$ if J is a continously differentiable, non-constant function on (0,1) with $\int J^4 < \infty$. The same is true for approximate scores $a_j = J(j/(N+1))$ provided that J is monotone near 0 and 1.

## 4. 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.

We recall some assumptions and notation from section 3 of ABZ (1976). Let F be a d.f. with a density f that is positive on $R^1$ and four times differentiable with derivatives $f^{(i)}$, $i = 1,\dots,4$. Define

$$\psi_i = \frac{f^{(i)}}{f}, \qquad i = 1,\dots,4, \tag{4.1}$$

and suppose that positive numbers $\varepsilon'$ and $C'$ exist such that for

$$\begin{gathered} m_1 = 6, \quad m_2 = 3, \quad m_3 = \tfrac{4}{3}, \quad m_4 = 1, \\ \sup\left\{ \int_{-\infty}^{\infty} |\psi_i(x+y)|^{m_i} f(x)dx : |y| \le \varepsilon' \right\} \le C', \qquad i = 1,\dots,4 . \end{gathered} \tag{4.2}$$

So far, we have studied the distribution of T under the assumption that $X_1,\dots,X_N$ are independent, $X_1,\dots,X_m$ having common d.f. F and $X_{m+1},\dots,X_N$ having d.f. G. We now add the assumptions that

$$G(x) = F(x-\theta) \tag{4.3}$$

for all x and that

$$0 \le \theta \le DN^{-\frac{1}{2}} \tag{4.4}$$

for some $D > 0$. Probabilities under this particular model will still be denoted by P. Note that (4.2), (4.3) and (4.4) together imply contiguity.

In section 3 we also introduced an auxiliary model where $X_1,\dots,X_N$ are supposed to be i.i.d. with common d.f. $H = (1-\lambda)F + \lambda G$. In view of (4.3) this common d.f. now becomes $H(x) = (1-\lambda)F(x) + \lambda F(x-\theta)$. Probabilities, expectations and variances under this model will be denoted by $P_H$, $E_H$ and $\sigma_H^2$ as before. Similarly, $P_F$, $E_F$ and $\sigma_F^2$ will indicate probabilities, expectations and variances under a third model where $X_1,\ldots,X_N$ are i.i.d. with common d.f. F. Note that for $\theta = 0$ these three models coincide.

Define

$$\tilde{K}(x) = \Phi(x) - \phi(x) \sum_{k=0}^{5} \tilde{\alpha}_k H_k(x) , \tag{4.5}$$

where

$$\begin{aligned}
\tilde{\alpha}_0 = \frac{1}{6}\left(\frac{\lambda(1-\lambda)}{\Sigma a_j^2}\right)^{\frac{1}{2}} \Big[ & 3(1-2\lambda)\theta^2 \sum a_j E_F \psi_2(Z_j) - 6N^{-1}\theta \sum a_j E_F \psi_1(Z_j) \\
& - \theta^3 \sum a_j E_F\{(1-3\lambda+3\lambda^2)\psi_3(Z_j) - 6\lambda(1-\lambda)\psi_1(Z_j)\psi_2(Z_j) \\
& + 3\lambda(1-\lambda)\psi_1^3(Z_j)\}\Big] ,
\end{aligned} \tag{4.6}$$

$$\begin{aligned}
\tilde{\alpha}_1 = \frac{1}{8\Sigma a_j^2}\Big[ & -4(1-2\lambda)\theta \sum a_j^2 E_F \psi_1(Z_j) + 2(1-2\lambda)^2\theta^2 \sum a_j^2 E_F \psi_2(Z_j) \\
& - 4\lambda(1-\lambda)\theta^2 \sum a_j^2 E_F \psi_1^2(Z_j) + 4\lambda(1-\lambda)\theta^2 \sigma_F^2(\sum a_j \psi_1(Z_j)) \\
& - 4(1-2\lambda)^2 N^{-1}\theta^2 \{\sum a_j E_F \psi_1(Z_j)\}^2 \\
& + \lambda(1-\lambda)(1-2\lambda)^2\theta^4 \{\sum a_j E_F \psi_2(Z_j)\}^2\Big] + \frac{1}{2N} ,
\end{aligned}$$

$$\begin{aligned}
\tilde{\alpha}_2 = \frac{1}{12\{\lambda(1-\lambda)\}^{\frac{1}{2}}(\Sigma a_j^2)^{3/2}}\Big[ & 2(1-2\lambda)\sum a_j^3 - 2(1-6\lambda+6\lambda^2)\theta \sum a_j^3 E_F \psi_1(Z_j) \\
& + 6(1-2\lambda)^2 N^{-1}\theta \sum a_j^2 \sum a_j E_F \psi_1(Z_j) \\
& - 3\lambda(1-\lambda)(1-2\lambda)^2\theta^3 \sum a_j^2 E_F \psi_1(Z_j) \sum a_j E_F \psi_2(Z_j)\Big] ,
\end{aligned}$$

$$\begin{aligned}
\tilde{\alpha}_3 = \frac{1}{24\lambda(1-\lambda)(\Sigma a_j^2)^2}\Big[ & (1-6\lambda+6\lambda^2)\sum a_j^4 + 3\lambda(1-\lambda)(1-2\lambda)^2\theta^2 \{\sum a_j^2 E_F \psi_1(Z_j)\}^2 \\
& + 2\lambda(1-\lambda)(1-2\lambda)^2\theta^2 \sum a_j^3 \sum a_j E_F \psi_2(Z_j)\Big] - \frac{(1-2\lambda)^2}{8\lambda(1-\lambda)N} ,
\end{aligned}$$

$$\tilde{\alpha}_4 = -\frac{(1-2\lambda)^2\theta\Sigma a_j^3\Sigma a_j^2 E_F\psi_1(Z_j)}{12\{\lambda(1-\lambda)\}^{1/2}(\Sigma a_j^2)^{5/2}} ,$$

$$\tilde{\alpha}_5 = \frac{(1-2\lambda)^2(\Sigma a_j^3)^2}{72\lambda(1-\lambda)(\Sigma a_j^2)^3} ,$$

and let

$$(4.7) \qquad \eta = -\left(\frac{\lambda(1-\lambda)}{\Sigma a_j^2}\right)^{\frac{1}{2}} \theta \sum a_j E_F\psi_1(Z_j).$$

We shall show that $\tilde{K}(x-\eta)$ is an expansion for the d.f. of $\{\lambda(1-\lambda)\sum a_j^2\}^{-\frac{1}{2}}T$. The expansion will be established in theorem 4.1 and an evaluation of the order of the remainder will be given in theorem 4.2.

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

$$(4.8) \qquad \varepsilon'' \le \alpha \le 1 - \varepsilon'' ,$$

for some $\varepsilon'' > 0$. We shall prove that an expansion for $\pi(F,\theta)$ is given by

$$(4.9) \qquad \tilde{\pi}(F,\theta) = 1 - \Phi(u_\alpha-\eta) + \phi(u_\alpha-\eta)\sum_{k=0}^{5} \tilde{\beta}_k H_k(u_\alpha-\eta) ,$$

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

$$(4.10) \qquad \tilde{\beta}_0 = \tilde{\alpha}_0 - \frac{(1-2\lambda)\Sigma a_j^3}{6\{\lambda(1-\lambda)\}^{1/2}(\Sigma a_j^2)^{3/2}}(u_\alpha^2-1) + 2\tilde{\alpha}_5(2u_\alpha^3-5u_\alpha) - \frac{u_\alpha}{2N}$$
$$- \left\{\frac{(1-6\lambda+6\lambda^2)\Sigma a_j^4}{24\lambda(1-\lambda)(\Sigma a_j^2)^2} - \frac{(1-2\lambda)^2}{8\lambda(1-\lambda)N}\right\}(u_\alpha^3-3u_\alpha) ,$$

$$\tilde{\beta}_1 = \tilde{\alpha}_1 + \tilde{\alpha}_5(u_\alpha^2-1)^2 - \frac{(1-2\lambda)^2}{12(\Sigma a_j^2)^2}\theta^2\sum a_j^3\sum a_j E_F\psi_2(Z_j)(u_\alpha^2-1) ,$$

$$\tilde{\beta}_2 = \tilde{\alpha}_2 - \tilde{\alpha}_4(u_\alpha^2-1) ,$$

$$\tilde{\beta}_3 = \tilde{\alpha}_3 - 2\tilde{\alpha}_5(u_\alpha^2-1) ,$$

$$\tilde{\beta}_k = \tilde{\alpha}_k \qquad \text{for } k = 4,5.$$

THEOREM 4.1. *Suppose that* (3.1) *and* (4.3) *hold and that positive numbers* c, C, C', D, δ, ε *and* ε' *exist such that* (3.10), (2.62), (3.19), (4.2) *and* (4.4) *are satisfied. Define*

$$(4.11) \qquad M = N^{-5/4} + N^{-1/2}\theta^3\Big[\sum\{E_F|\psi_1(Z_j) - E_F\psi_1(Z_j)|^3\}^{4/9}\Big]^{9/4} + N^{-3/4}\theta^3\Big[\sum\{E_F(\psi_2(Z_j) - E_F\psi_2(Z_j))^2\}^{2/3}\Big]^{3/2} .$$

*Then there exists* B > 0 *depending only on* c, C, C', D, δ, ε *and* ε' *such that*

$$(4.12) \qquad \sup_x \left| P\left(\frac{T}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac{1}{2}}} \le x\right) - \tilde{K}(x-\eta) \right| \le BM .$$

*If, in addition,* (4.8) *is satisfied there exists* B' > 0 *depending only on* c, C, C', D, δ, ε, ε' *and* ε'' *such that*

$$(4.13) \qquad |\pi(F,\theta) - \tilde{\pi}(F,\theta)| \le B'M.$$

PROOF. The proof of (4.12) hinges on Taylor expansion with respect to $\theta$ of the moments under $P_H$ of functions of $P = (P_1,\ldots,P_N)$ occurring in expansion (3.20). Since both H and P depend on $\theta$ the argument is highly technical and laborious and it is therefore given in the appendix. Theorem 3.1, corollary A.1, (A.12) and (A.13) immediately yield (4.12).

The one-sided level $\alpha$ test based on T rejects the hypothesis if $T\{\lambda(1-\lambda)\sum a_j^2\}^{-\frac{1}{2}} \ge \xi_\alpha$ with possible randomization if equality occurs. Using (4.12) for $\theta = 0$ (or corollary 2.1), (3.10), (3.19) and (4.8) we easily show that

$$(4.14) \qquad \xi_\alpha = u_\alpha + \frac{(1-2\lambda)\Sigma a_j^3}{6\{\lambda(1-\lambda)\}^{1/2}(\Sigma a_j^2)^{3/2}}(u_\alpha^2-1) - 2\tilde{\alpha}_5(2u_\alpha^3-5u_\alpha) + \frac{u_\alpha}{2N} + \left\{\frac{(1-6\lambda+6\lambda^2)\Sigma a_j^4}{24\lambda(1-\lambda)(\Sigma a_j^2)^2} - \frac{(1-2\lambda)^2}{8\lambda(1-\lambda)N}\right\}(u_\alpha^3-3u_\alpha) + O(N^{-5/4}) ,$$

where, in this proof, $O(x)$ denotes a quantity bounded by $B_1|x|$ with $B_1$ depending only on c, C, C', D, $\delta$, $\varepsilon$, $\varepsilon'$ and $\varepsilon''$. Because of (4.12),

$$\pi(F,\theta) = 1 - \tilde{K}(\xi_\alpha - \eta) + O(M) .$$

Using (4.14), (4.8) and the bounds provided by corollary A.1, we now expand $\tilde{K}(\xi_\alpha-\eta)$ about the point $(u_\alpha-\eta)$ and arrive at (4.13). □

Define

$$(4.15) \qquad \Psi_i(t) = \psi_i(F^{-1}(t)) = \frac{f^{(i)}(F^{-1}(t))}{f(F^{-1}(t))} , \qquad i = 1,\dots,4 .$$

THEOREM 4.2. *Let* M *be defined by* (4.11) *and suppose that positive numbers* C *and* $\delta$ *exist such that* $|\Psi_1'(t)| \le C\{t(1-t)\}^{-5/4+\delta}$ *and* $|\Psi_2'(t)| \le C\{t(1-t)\}^{-3/2+\delta}$. *Then there exists* B > 0 *depending only on* C *and* $\delta$ *such that*

$$M \le BN^{-5/4}$$

PROOF. The proof is similar to that of corollary A2.1 in ABZ (1976). To deal with the second term of M we take $h = \Psi_1$ and replace 4/3 by 5/4 in the proof of that corollary. For the third term of M we take $h = \Psi_2$, replace 4/3 by 3/2, appeal to condition $R_2$ instead of $R_3$ and otherwise proceed as in the proof of corollary A2.1 of ABZ (1976). □

$\mathcal{F}$ is the class of d.f.'s F on $R^1$ with positive and four times differentiable densities f and such that, for $\psi_i = f^{(i)}/f$, $\Psi_i = \psi_i(F^{-1})$, $m_1 = 6$, $m_2 = 3$, $m_3 = \frac{4}{3}$, $m_4 = 1$,

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

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

Note that one can argue as in the proof of corollary A2.1 of ABZ (1976) to show that, in conjunction with (5.5), condition (5.4) is weaker than the assumption $\int J^6(t)dt < \infty$. Define

(5.8) $$\bar{\alpha}_0 = \frac{1}{6}\left(\frac{\lambda(1-\lambda)}{N\int J^2(t)dt}\right)^{\frac{1}{2}}\Big[3(1-2\lambda)N\theta^2\int J(t)\Psi_2(t)dt - 6\theta\int J(t)\Psi_1(t)dt$$

$$- N\theta^3\int J(t)\{(1-3\lambda+3\lambda^2)\Psi_3(t)-6\lambda(1-\lambda)\Psi_1(t)\Psi_2(t)+3\lambda(1-\lambda)\Psi_1^3(t)\}dt\Big],$$

$$\bar{\alpha}_1 = \frac{1}{8\int J^2(t)dt}\Big[-4(1-2\lambda)\theta\int J^2(t)\Psi_1(t)dt+2(1-2\lambda)^2\theta^2\int J^2(t)\Psi_2(t)dt$$

$$-4\lambda(1-\lambda)\theta^2\int J^2(t)\Psi_1^2(t)dt$$

$$+ 4\lambda(1-\lambda)\theta^2\iint J(s)J(t)\Psi_1'(s)\Psi_1'(t)[s\wedge t - st]dsdt$$

$$- 4(1-2\lambda)^2\theta^2\{\int J(t)\Psi_1(t)dt\}^2$$

$$+ \lambda(1-\lambda)(1-2\lambda)^2N\theta^4\{\int J(t)\Psi_2(t)dt\}^2\Big] + \frac{1}{2N} ,$$

$$\bar{\alpha}_2 = \frac{1}{12\{\lambda(1-\lambda)N\}^{1/2}\{\int J^2(t)dt\}^{3/2}}\Big[2(1-2\lambda)\int J^3(t)dt$$

$$- 2(1-6\lambda+6\lambda^2)\theta\int J^3(t)\Psi_1(t)dt +$$

$$+ 6(1-2\lambda)^2\theta\int J^2(t)dt\int J(t)\Psi_1(t)dt$$

$$- 3\lambda(1-\lambda)(1-2\lambda)^2N\theta^3\int J^2(t)\Psi_1(t)dt\int J(t)\Psi_2(t)dt\Bigg] ,$$

$$\bar{\alpha}_3 = \frac{1}{24\lambda(1-\lambda)N\{\int J^2(t)dt\}^2}\Bigg[(1-6\lambda+6\lambda^2)\int J^4(t)dt$$

$$+ 3\lambda(1-\lambda)(1-2\lambda)^2N\theta^2\{\int J^2(t)\Psi_1(t)dt\}^2$$

$$+ 2\lambda(1-\lambda)(1-2\lambda)^2N\theta^2\int J^3(t)dt\int J(t)\Psi_2(t)dt\Bigg] - \frac{(1-2\lambda)^2}{8\lambda(1-\lambda)N} ,$$

$$\bar{\alpha}_4 = - \frac{(1-2\lambda)^2\theta}{12\{\lambda(1-\lambda)N\}^{1/2}}\,\frac{\int J^3(t)dt\int J^2(t)\Psi_1(t)dt}{\{\int J^2(t)dt\}^{5/2}} ,$$

$$\bar{\alpha}_5 = \frac{(1-2\lambda)^2}{72\lambda(1-\lambda)N}\,\frac{\{\int J^3(t)dt\}^2}{\{\int J^2(t)dt\}^3} ,$$

(5.9) $$\bar{K}_1(x) = \Phi(x) - \phi(x)\Bigg[\sum_{k=0}^{5}\bar{\alpha}_kH_k(x)$$

$$+ \tfrac{1}{2}\left(\frac{\lambda(1-\lambda)}{N\int J^2(t)dt}\right)^{\frac{1}{2}}\theta\Bigg\{2\sum_{j=1}^{N}\mathrm{cov}(J(U_{j:N}),\Psi_1(U_{j:N}))$$

$$- \frac{\int J(t)\Psi_1(t)dt}{\int J^2(t)dt}\sum_{j=1}^{N}\sigma^2(J(U_{j:N}))\Bigg\}\Bigg] ,$$

(5.10) $$\bar{K}_2(x) = \Phi(x) - \phi(x)\Bigg[\sum_{k=0}^{5}\bar{\alpha}_kH_k(x)$$

$$+ \tfrac{1}{2}\left(\frac{\lambda(1-\lambda)}{N\int J^2(t)dt}\right)^{\frac{1}{2}}\theta\Bigg\{2\int_{N^{-1}}^{1-N^{-1}} J'(t)\Psi_1'(t)t(1-t)dt$$

$$- \frac{\int J(t)\Psi_1(t)dt}{\int J^2(t)dt}\int_{N^{-1}}^{1-N^{-1}} (J'(t))^2t(1-t)dt\Bigg\}\Bigg] ,$$

(5.11) $$\bar{\eta} = - \left(\frac{\lambda(1-\lambda)N}{\int J^2(t)dt}\right)^{\frac{1}{2}}\theta\int J(t)\Psi_1(t)dt ,$$

$$(5.16)\qquad |\pi(F,\theta) - \bar{\pi}_1(F,\theta)| \le \delta_N N^{-1} ,$$

$$(5.17)\qquad |\pi(F,\theta) - \bar{\pi}_2(F,\theta)| \le \delta_N N^{-1} + BN^{-3/2} \int_{N^{-1}}^{1-N^{-1}} |J'(t)|(|J'(t)| + |\Psi_1'(t)|)\{t(1-t)\}^{\frac{1}{2}}dt.$$

PROOF. In the first part of the proof we shall not need requirement (5.4) but only the weaker assumption $\int J^4(t)dt < \infty$. We proceed as in the proof of theorem 4.1 in ABZ (1976), drawing heavily on the results in appendix 2 of ABZ (1976). Note that these results remain valid in the present context even though the definition of the functions $\Psi_i$ is slightly different here. Throughout the proof we shall make use of $O$ and $o$ symbols that are uniform for fixed F, J, D, $\varepsilon$ and $\varepsilon'$.

Because $\sum a_j = N\int J(t)dt = 0$ and in view of the remark made at the end of section 3, the assumptions of theorem 4.1 are satisfied. The proof of corollary A2.1 of ABZ (1976) shows that (5.6) and (5.7) imply that

$$(5.18)\qquad \Psi_1'(t) = o(\{t(1-t)\}^{-7/6}) \qquad \text{for } t \to 0,1 .$$

Hence, because of (5.7), $\Psi_1''(t) = o(\{t(1-t)\}^{-13/6})$ and $\Psi_1(t) = o(\{t(1-t)\}^{-1/6})$ for $t \to 0,1$. Since $f(F^{-1})$ has a summable derivative $\Psi_1$ on (0,1), $f(F^{-1})$ must have limits at 0 and 1; as f is positive on $R^1$, these limits must be equal to 0. It follows that $f(F^{-1}(t)) = o(\{t(1-t)\}^{5/6})$ for $t \to 0,1$. Combining these facts with the identity $\Psi_2'(t) = \Psi_1''(t)f(F^{-1}(t)) + 3\Psi_1(t)\Psi_1'(t)$, we find that

$$(5.19)\qquad \Psi_2'(t) = o(\{t(1-t)\}^{-4/3}) \qquad \text{for } t \to 0,1 .$$

Thus the assumptions of theorem 4.2 are also satisfied and we can take the expansions of section 4 as a starting point for proving theorem 5.1.

In $\tilde{\alpha}_0,\ldots,\tilde{\alpha}_5$, $\tilde{\beta}_0,\ldots,\tilde{\beta}_5$ defined by (4.6) and (4.10) we may replace $E_F$ $\sigma_F^2$ and $\psi_i(Z_j)$ by E, $\sigma^2$ and $\psi_i(F^{-1}(U_{j:N})) = \Psi_i(U_{j:N})$ without changing anything. Next, arguing as in corollary A2.2 of ABZ (1976), we see that for all sums of the form $\sum a_j^k$ and $\sum a_j^k Eh(U_{j:N})$ occurring in $\tilde{\alpha}_0,\ldots,\tilde{\alpha}_5$, $\tilde{\beta}_0,\ldots,\tilde{\beta}_5$ we may write

$$(5.20)\qquad \frac{1}{N}\sum a_j^k = \int J^k(t)dt + o(1) ,$$

$$(5.21)\qquad \frac{1}{N}\sum a_j^k\, Eh(U_{j:N}) = \int J^k(t)h(t)dt + o(1) ,$$

and also

$$(5.22)\qquad \frac{1}{N}\,\sigma^2(\sum a_j\Psi_1(U_{j:N})) = \iint J(s)J(t)\Psi_1'(s)\Psi_1'(t)[s\wedge t - st]dsdt + o(1).$$

We note that $\bar{\alpha}_0,\ldots,\bar{\alpha}_5$, $\bar{\beta}_0,\ldots,\bar{\beta}_5$ are obtained from $\tilde{\alpha}_0,\ldots,\tilde{\alpha}_5$, $\tilde{\beta}_0,\ldots,\tilde{\beta}_5$ by replacing every expression of the form (5.20) - (5.22) by the corresponding integral on the right in (5.20) - (5.22). Since $\int J^2(t)dt > 0$, we know that for those terms in $\tilde{\alpha}_0,\ldots,\tilde{\alpha}_5$, $\tilde{\beta}_0,\ldots,\tilde{\beta}_5$ that are $O(N^{-1})$, this substitution can only introduce errors that are $o(N^{-1})$.

The first terms in $\tilde{\alpha}_0$, $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ as well as the second term in $\tilde{\beta}_0$ are generally not $O(N^{-1})$ but only $O(N^{-\frac{1}{2}})$, and here the substitution of integrals for sums gives rise to more complicated remainder terms. This creates problems we did not encounter in the one-sample case where certain symmetries prohibit the occurrence of $O(N^{-\frac{1}{2}})$ terms. We have

$$\frac{1}{N}\sum a_j^2 = \int J^2(t)dt - \frac{1}{N}\sum\sigma^2(J(U_{j:N})) ,$$

$$\frac{1}{N}\sum a_j^3 = \int J^3(t)dt - \frac{1}{N}\sum \mathrm{cov}(J(U_{j:N}),J^2(U_{j:N})) - \frac{1}{N}\sum EJ(U_{j:N})\sigma^2(J(U_{j:N})),$$

$$\frac{1}{N}\sum a_j E\Psi_2(U_{j:N}) = \int J(t)\Psi_2(t)dt - \frac{1}{N}\sum \mathrm{cov}(J(U_{j:N}),\Psi_2(U_{j:N}))$$

$$\frac{1}{N}\sum a_j^2 E\Psi_1(U_{j:N}) = \int J^2(t)\Psi_1(t)dt - \frac{1}{N}\sum \mathrm{cov}(J^2(U_{j:N}),\Psi_1(U_{j:N}))$$

$$- \frac{1}{N}\sum E\Psi_1(U_{j:N})\sigma^2(J(U_{j:N})) .$$

By (A2.22) in ABZ (1976), $N^{-3/2}\sum\sigma^2(J(U_{j:N})) = o(N^{-1})$. It follows that for $k = 0,\ldots,5$,

$$(5.23)\qquad \tilde{\alpha}_k - \bar{\alpha}_k = o(N^{-1}) + O(M_1), \qquad \tilde{\beta}_k - \bar{\beta}_k = o(N^{-1}) + O(M_1) ,$$

$$(5.24)\qquad M_1 = (1-2\lambda)N^{-3/2}\Big[\,|\textstyle\sum \operatorname{cov}(J(U_{j:N}),\, J^2(U_{j:N}))|$$

$$+\ |\textstyle\sum EJ(U_{j:N})\sigma^2(J(U_{j:N}))| + |\sum \operatorname{cov}(J(U_{j:N}),\Psi_2(U_{j:N}))|$$

$$+\ |\textstyle\sum \operatorname{cov}(J^2(U_{j:N}),\, \Psi_1(U_{j:N}))| + |\sum E\Psi_1(U_{j:N})\sigma^2(J(U_{j:N}))|\Big]\ .$$

By (A2.17), (A2.22) and (A2.23) in ABZ (1976) we have

$$(5.25)\qquad \eta = \bar{\eta} + \tfrac{1}{2}\left(\frac{\lambda(1-\lambda)}{N\int J^2(t)dt}\right)^{\frac{1}{2}}\theta\Big[2\textstyle\sum \operatorname{cov}(J(U_{j:N}),\, \Psi_1(U_{j:N}))$$

$$-\ \frac{\int J(t)\Psi_1(t)dt}{\int J^2(t)dt}\textstyle\sum \sigma^2(J(U_{j:N}))\Big] + o(N^{-1}) = \bar{\eta} + o(N^{-\frac{1}{2}}).$$

Hence, uniformly in x,

$$\tilde{K}(x-\eta) = \Phi(x-\bar{\eta}) - \phi(x-\bar{\eta})\Big[(\eta-\bar{\eta}) + \sum_{k=0}^{5}\bar{\alpha}_k H_k(x-\bar{\eta})\Big] + o(N^{-1}) + O(M_1)$$

$$= \bar{K}_1(x-\bar{\eta}) + o(N^{-1}) + O(M_1),$$

and similarly

$$\tilde{\pi}(F,\theta) = \bar{\pi}_1(F,\theta) + o(N^{-1}) + O(M_1).$$

It follows that, in order to prove (5.14) and (5.16), it suffices to show that $M_1 = o(N^{-1})$. Since (5.15) and (5.17) are immediate consequences of (5.14) and (5.16) on the one hand and (A2.22) and (A2.23) in ABZ (1976) on the other, the proof of the theorem will then be complete.

At this point we finally need condition (5.4) rather than the weaker assumption $\int J^4(t)dt < \infty$. Using (5.4), (5.18) and (5.19) and proceeding as in the proof of corollary A2.1 in ABZ (1976), we find that each term of $M_1$ is

$$(5.26)\qquad o(N^{-3/2}\int_{N^{-1}}^{1-N^{-1}} \{t(1-t)\}^{-3/2}dt) = o(N^{-1})\ . \qquad \square$$

REMARK. In the above we have stressed the fact that the only reason for requiring (5.4) rather than assuming $\int J^4(t)dt < \infty$, is that we have to show that $M_1 = o(N^{-1})$. However, there are special cases of interest where $\int J^4(t)dt < \infty$ suffices. If either $\lambda = \frac{1}{2}$, or f is a symmetric density and J(t) is antisymmetric about $t = \frac{1}{2}$, then $M_1 = 0$. Less trivially, since $\int J^4(t)dt < \infty$ and (5.5) imply that $J'(t) = o(\{t(1-t)\}^{-5/4})$, we can follow the reasoning leading to (5.26) while retaining the factor $(1-2\lambda)$, to arrive at

$$(5.27) \qquad M_1 = o(|1-2\lambda|N^{-3/2}\int_{N^{-1}}^{1-N^{-1}} \{t(1-t)\}^{-7/4}dt) = o(|1-2\lambda|N^{-3/4}) .$$

Hence in the special cases where either $\lambda = \frac{1}{2} + O(N^{-\frac{1}{4}})$, or f is a symmetric density and J is antisymmetric about the point $\frac{1}{2}$, the conclusions of theorem 5.1 will hold if condition (5.4) is replaced by the assumption $\int J^4(t)dt < \infty$. Comparison with ABZ (1976) shows that in these special cases the conditions under which theorem 5.1 holds are essentially the same as the conditions of the comparable theorem 4.1 in ABZ (1976) for the one-sample problem. This is not surprising as one may think of the one-sample case under contiguous alternatives as a two-sample situation with $\lambda = \frac{1}{2} + O_P(N^{-\frac{1}{2}})$.

We now turn to the special case $J = -\Psi_1$. For $F \in \mathcal{F}$ we obtain by partial integration

$$(5.28) \qquad \int \Psi_1(t)\Psi_2(t)dt = \tfrac{1}{2}\int \Psi_1^3(t)dt ,$$

$$\int \Psi_1^2(t)\Psi_2(t)dt = \frac{2}{3}\int \Psi_1^4(t)dt ,$$

$$\int \Psi_1(t)\Psi_3(t)dt = \frac{2}{3}\int \Psi_1^4(t)dt - \int \Psi_2^2(t)dt ,$$

$$\iint \Psi_1(s)\Psi_1(t)\Psi_1'(s)\Psi_1'(t)[s\wedge t-st]dsdt = \frac{1}{4}\int \Psi_1^4(t)dt - \frac{1}{4}\left(\int \Psi_1^2(t)dt\right)^2 .$$

Substitution of $J = -\Psi_1$ and application of (5.28) considerably simplifies the expressions (5.8) and (5.12) for $\bar{\alpha}_k$ and $\bar{\beta}_k$. Note that $\bar{\eta}$ defined by (5.11) reduces to

(5.29) $$\eta^* = \theta\{\lambda(1-\lambda)N \int \Psi_1^2(t)dt\}^{\frac{1}{2}} .$$

The expressions for $\bar{\alpha}_k$ and $\bar{\beta}_k$ simplify somewhat further if we express $\theta$ in terms of $\eta^*$ throughout. Finally we rearrange the terms in $\sum \bar{\alpha}_k H_k(x-\eta^*)$ and $\sum \bar{\beta}_k H_k(u_\alpha-\eta^*)$ according to the integrals involved and substitute the explicit expressions (2.43) for the Hermite polynomials $H_k$. In this way we find after laborious but straightforward calculations that for $J = -\Psi_1$,

(5.30) $$\Phi(x-\bar{\eta}) - \phi(x-\bar{\eta}) \sum_{k=0}^{5} \bar{\alpha}_k H_k(x-\bar{\eta}) = L_0(x) ,$$

$$1 - \Phi(u_\alpha-\bar{\eta}) + \phi(u_\alpha-\bar{\eta}) \sum_{k=0}^{5} \bar{\beta}_k H_k(u_\alpha-\bar{\eta}) = \pi_0^*(F,\theta) ,$$

where

(5.31) $$L_0(x) = \Phi(x-\eta^*)$$

$$- \frac{\phi(x-\eta^*)}{288}\Bigg[\frac{24(1-2\lambda)}{\{\lambda(1-\lambda)N\}^{1/2}} \frac{\int\Psi_1^3(t)dt}{\{\int\Psi_1^2(t)dt\}^{3/2}} \{-2(x^2-1) - 2\eta^* x + \eta^{*2}\}$$

$$+ \frac{4}{\lambda(1-\lambda)N} \frac{\int\Psi_1^4(t)dt}{\{\int\Psi_1^2(t)dt\}^2} \{3(1-6\lambda+6\lambda^2)(x^3-3x+\eta^*(x^2-1))-3(1-5\lambda+5\lambda^2)\eta^{*2}x$$

$$+ 5(1-3\lambda+3\lambda^2)\eta^{*3}\} - \frac{48}{\lambda(1-\lambda)N} \frac{\int\Psi_2^2(t)dt}{\{\int\Psi_1^2(t)dt\}^2} (1-3\lambda+3\lambda^2)\eta^{*3}$$

$$+ \frac{(1-2\lambda)^2}{\lambda(1-\lambda)N} \frac{\{\int\Psi_1^3(t)dt\}^2}{\{\int\Psi_1^2(t)dt\}^3} \{4(x^5-10x^3+15x) + 4\eta^*(x^4-6x^2+3) - 8\eta^{*2}(x^3-3x)$$

$$- 4\eta^{*3}(x^2-1) + 5\eta^{*4}x - \eta^{*5}\} + \frac{144x}{N} + \frac{36}{\lambda(1-\lambda)N} \{-(1-2\lambda)^2(x^3-3x+\eta^* x^2)$$

$$+ \eta^* + (1-5\lambda+5\lambda^2)\eta^{*2}x + (1-3\lambda+3\lambda^2)\eta^{*3}\}\Bigg] ,$$

$$\pi_0^*(F,\theta) = 1 - \Phi(u_\alpha-\eta^*) +$$

$$+ \frac{\eta^*\phi(u_\alpha-\eta^*)}{288}\left[\frac{24(1-2\lambda)}{\{\lambda(1-\lambda)N\}^{1/2}}\frac{\int\Psi_1^3(t)dt}{\{\int\Psi_1^2(t)dt\}^{3/2}}(-2u_\alpha+\eta^*)\right.$$

$$+ \frac{4}{\lambda(1-\lambda)N}\frac{\int\Psi_1^4(t)dt}{\{\int\Psi_1^2(t)dt\}^2}\{3(1-6\lambda+6\lambda^2)(u_\alpha^2-1)-3(1-5\lambda+5\lambda^2)\eta^*u_\alpha$$

$$+ 5(1-3\lambda+3\lambda^2)\eta^{*2}\} - \frac{48}{\lambda(1-\lambda)N}\frac{\int\Psi_2^2(t)dt}{\{\int\Psi_1^2(t)dt\}^2}(1-3\lambda+3\lambda^2)\eta^{*2}$$

$$+ \frac{(1-2\lambda)^2}{\lambda(1-\lambda)N}\frac{\{\int\Psi_1^3(t)dt\}^2}{\{\int\Psi_1^2(t)dt\}^3}\{-8(2u_\alpha^2-1)+4\eta^*(u_\alpha^3+3u_\alpha)-8\eta^{*2}(u_\alpha^2-1)$$

$$+ 5\eta^{*3}u_\alpha - \eta^{*4}\} + \frac{36}{\lambda(1-\lambda)N}\{-(1-2\lambda)^2u_\alpha^2+1+(1-5\lambda+5\lambda^2)\eta^*u_\alpha$$

$$\left. + (1-3\lambda+3\lambda^2)\eta^{*2}\}\right].$$

Define

(5.32)
$$L_1(x) = L_0(x) + \frac{\eta^*\phi(x-\eta^*)}{2N\int\Psi_1^2(t)dt}\sum_{j=1}^{N}\sigma^2(\Psi_1(U_{j:N})),$$

$$L_2(x) = L_0(x) + \frac{\eta^*\phi(x-\eta^*)}{2N\int\Psi_1^2(t)dt}\int_{N^{-1}}^{1-N^{-1}}(\Psi_1'(t))^2t(1-t)dt,$$

$$L_3(x) = L_0(x) + \frac{\eta^*\phi(x-\eta^*)}{2N\int\Psi_1^2(t)dt}\sum_{j=1}^{N}E(\Psi_1(U_{j:N})-\Psi_1(\tfrac{j}{N+1}))^2,$$

$$\pi_1^*(F,\theta) = \pi_0^*(F,\theta) - \frac{\eta^*\phi(u_\alpha-\eta^*)}{2N\int\Psi_1^2(t)dt}\sum_{j=1}^{N}\sigma^2(\Psi_1(U_{j:N})),$$

$$\pi_2^*(F,\theta) = \pi_0^*(F,\theta) - \frac{\eta^*\phi(u_\alpha-\eta^*)}{2N\int\Psi_1^2(t)dt}\int_{N^{-1}}^{1-N^{-1}}(\Psi_1'(t))^2t(1-t)dt,$$

$$\pi_3^*(F,\theta) = \pi_0^*(F,\theta) - \frac{\eta^*\phi(u_\alpha-\eta^*)}{2N\int\Psi_1^2(t)dt}\sum_{j=1}^{N}E(\Psi_1(U_{j:N})-\Psi_1(\tfrac{j}{N+1}))^2.$$

Note that (5.9), (5.10), (5.11), (5.13), (5.30) and (5.31) imply that for $J = -\Psi_1$, $\bar{K}_i(x-\bar{\eta}) = L_i(x)$ and $\bar{\pi}_i(F,\theta) = \pi_i^*(F,\theta)$ for $i = 1,2$. The expansions

$L_3$ and $\pi_3^*$ are connected only with approximate scores that were not considered so far.

THEOREM 5.2. *Let* $F \in \mathcal{F}$, $J = -\Psi_1$, $G(x) = F(x-\theta)$, $0 \le \theta \le DN^{-\frac{1}{2}}$, $\varepsilon \le \lambda \le 1-\varepsilon$ *and* $\varepsilon' \le \alpha \le 1-\varepsilon'$ *for positive* D, $\varepsilon$ *and* $\varepsilon'$. *Then, for every fixed* F, J, D, $\varepsilon$ *and* $\varepsilon'$, *there exist positive numbers* B, $\delta_1, \delta_2, \ldots$ *with* $\lim_{N\to\infty} \delta_N = 0$ *such that the following statements hold for every* N.

(i) *For exact scores* $a_j = -E\Psi_1(U_{j:N})$,

$$(5.33)\qquad \sup_x \left| P\left(\frac{T}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac{1}{2}}} \le x\right) - L_1(x)\right| \le \delta_N N^{-1} ,$$

$$(5.34)\qquad \sup_x \left| P\left(\frac{T}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac{1}{2}}} \le x\right) - L_2(x)\right| \le \delta_N N^{-1} + BN^{-3/2}\int_{N^{-1}}^{1-N^{-1}} (\Psi_1'(t))^2\{t(1-t)\}^{\frac{1}{2}}dt ,$$

$$(5.35)\qquad |\pi(F,\theta) - \pi_1^*(F,\theta)| \le \delta_N N^{-1} ,$$

$$(5.36)\qquad |\pi(F,\theta) - \pi_2^*(F,\theta)| \le \delta_N N^{-1} + BN^{-3/2}\int_{N^{-1}}^{1-N^{-1}} (\Psi_1'(t))^2\{t(1-t)\}^{\frac{1}{2}}dt;$$

(ii) *For approximate scores* $a_j = -\Psi_1\left(\frac{j}{N+1}\right)$,

$$(5.37)\qquad \sup_x \left| P\left(\frac{T-\lambda\Sigma a_j}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac{1}{2}}} \le x\right) - L_3(x)\right| \le \delta_N N^{-1} ,$$

$$(5.38)\qquad \sup_x \left| P\left(\frac{T-\lambda\Sigma a_j}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac{1}{2}}} \le x\right) - L_2(x)\right| \le \delta_N N^{-1} + BN^{-3/2}\int_{N^{-1}}^{1-N^{-1}} (\Psi_1'(t))^2\{t(1-t)\}^{\frac{1}{2}}dt,$$

$$(5.39)\qquad |\pi(F,\theta) - \pi_3^*(F,\theta)| \le \delta_N N^{-1}$$

*and* (5.36) *continues to hold.*

PROOF. For $F \in \mathcal{F}$, $\Psi_1$ is not constant on (0,1), $\int \Psi_1(t)dt = 0$ and $\Psi_1^6$ is summable. In view of the remark following definition 5.1, this implies that $J \in \mathcal{J}$. We have already noted that $K_i(x-\bar{\eta}) = L_i(x)$ and $\bar{\pi}_i(F,\theta) = \pi_i^*(F,\theta)$ for $i = 1,2$, if $J = -\Psi_1$. Part (i) of the theorem is therefore an immediate consequence of theorem 5.1.

To prove part (ii) we retrace the proof of theorem 5.1 for $J = -\Psi_1$ and approximate scores $a_j = -\Psi_1(j/(N+1))$. The first difficulty we encounter is that in general $\sum a_j \neq 0$. However, lemma A2.3 of ABZ (1976), (5.7) and (5.18) yield

$$(5.40) \qquad a_{\cdot} = \frac{1}{N}\sum_{j=1}^{N} a_j = -\int_0^1 \Psi_1(t)dt + O(N^{-1}\int_{N^{-1}}^{1-N^{-1}} |\Psi_1'(t)|dt) = o(N^{-5/6}) ,$$

and one easily verifies that the conditions of theorem 4.1 hold for the reduced scores $a_j - a_{\cdot}$. Since the assumptions of theorem 4.2 are also satisfied, we have

$$(5.41) \qquad \sup_x \left| P\left(\frac{T-\lambda\Sigma a_j}{\{\lambda(1-\lambda)\Sigma(a_j-a_{\cdot})^2\}^{\frac{1}{2}}} \le x\right) - \hat{K}(x-\hat{\eta})\right| = O(N^{-5/4}) ,$$

where $\hat{K}$ and $\hat{\eta}$ are obtained from $\tilde{K}$ and $\eta$ by replacing $a_j$ by $a_j - a_{\cdot}$ throughout. Because, by (3.10) and (5.40) ,

$$(5.42) \qquad \sum (a_j-a_{\cdot})^2 = \sum a_j^2(1+o(N^{-5/3})) ,$$

we can change the norming constant $\sum(a_j-a_{\cdot})^2$ of T in (5.41) back to $\sum a_j^2$ with impunity. As $\int \Psi_1(t)dt = 0$, (5.42) also ensures that $|\hat{\eta} - \eta| = o(N^{-5/3})$. Finally (A2.16) of ABZ (1976) and (5.18) imply that $\sigma_F^2(\sum a_j\psi_1(Z_j)) = O(N)$ for $J = -\Psi_1$ and, together with (5.42), (3.10), (5.6) and (5.40), this yields $\sup_x|\hat{K}(x) - \tilde{K}(x)| = o(N^{-4/3})$. Combining these results we find

$$(5.43) \qquad \sup_x \left| P\left(\frac{T-\lambda\Sigma a_j}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac{1}{2}}} \le x\right) - \tilde{K}(x-\eta)\right| = O(N^{-5/4})$$

and similarly

$$(5.44) \qquad |\pi(F,\theta) - \tilde{\pi}(F,\theta)| = O(N^{-5/4}) .$$

The remainder of the proof parallels that of theorem 5.1 for the special case $J = -\Psi_1$. We replace all sums as well as $\sigma^2(\sum a_j \Psi_1(U_{j:N}))$ by the appropriate integrals. The reasoning of corollary A2.2 of ABZ (1976) shows that for those terms in the expansions that are $O(N^{-1})$, this substitution will only lead to errors that are $o(N^{-1})$. For the $O(N^{-\frac{1}{2}})$ terms the error committed is $O(M_1) + O(M_2)$, where $M_1$ is given by (5.24) with $J = -\Psi_1$ and $M_2$ originates from the difference between exact and approximate scores. It was shown in the proof of theorem 5.1 that $M_1 = o(N^{-1})$. With regard to $M_2$, (5.7), lemma A2.3 of ABZ (1976), (5.18) and (5.19) imply that, uniformly in j,

$$(5.45) \qquad |\{E\Psi_1(U_{j:N})\}^k - \Psi_1^k(\tfrac{j}{N+1})| = O(N^{-1}) + o\left(N^{-1}\left\{\frac{j(N-j+1)}{(N+1)^2}\right\}^{-1-k/6}\right)$$

$$|E\Psi_1(U_{j:N})| = O(1) + o\left(\left\{\frac{j(N-j+1)}{(N+1)^2}\right\}^{-1/6}\right),$$

$$|E\Psi_2(U_{j:N})| = O(1) + o\left(\left\{\frac{j(N-j+1)}{(N+1)^2}\right\}^{-1/3}\right),$$

where $k = 1,2,3$. It follows that $M_2$ is of the form (5.26) and is therefore $o(N^{-1})$.

It remains to replace $\eta$ by $\eta^*$. Because of (5.7), (5.18) and lemma A2.3 of ABZ (1976), $N^{-1} \sum \sigma^2(\Psi_1(U_{j:N})) = o(N^{-2/3})$, and in view of (5.45),

$$\frac{1}{N} \sum \Psi_1(\tfrac{j}{N+1}) E\Psi_1(U_{j:N}) - \int \Psi_1^2(t)dt$$

$$= -\frac{1}{N} \sum E\Psi_1(U_{j:N})\left[\Psi_1(U_{j:N}) - \Psi_1(\tfrac{j}{N+1})\right] = o(N^{-2/3}),$$

$$\frac{1}{N} \sum \Psi_1^2(\tfrac{j}{N+1}) - \int \Psi_1^2(t)dt = -\frac{1}{N} \sum \left[E\Psi_1^2(U_{j:N}) - \Psi_1^2(\tfrac{j}{N+1})\right] = o(N^{-2/3}).$$

Hence, for $J = -\Psi_1$,

$$(5.46) \qquad \eta = \eta^* - \frac{\eta^*}{2N\int \Psi_1^2(t)dt} \sum E\{\Psi_1(U_{j:N}) - \Psi_1(\tfrac{j}{N+1})\}^2 + o(N^{-4/3})$$

$$= \eta^* + o(N^{-2/3}),$$

and a comparison with (5.25) for $J = -\Psi_1$ shows that (5.37) and (5.39) will hold if $L_3$ and $\pi_3^*$ can be obtained from $L_1$ and $\pi_1^*$ by replacing $\sum \sigma^2(\Psi_1(U_{j:N}))$ by $\sum E\{\Psi_1(U_{j:N}) - \Psi_1(j/(N+1))\}^2$. Since this is true, (5.37) and (5.39) are proved. The validity of (5.38) and (5.36) for approximate scores is a consequence of (5.37), (5.39) and corollary A2.2 of ABZ (1976). The proof of the theorem is complete. □

At this point it is appropriate to repeat some remarks made in ABZ (1976). The correspondence between expansions (5.34) and (5.38) and the fact that (5.36) holds for both exact and approximate scores seem to be typical for the case $J = -\Psi_1$. In the general case where $J \neq -\Psi_1$, expansions (5.15) and (5.17) will not hold for approximate scores even if T is replaced by $T - \lambda \sum a_j$ in (5.15). A second remark is that the growth conditions on $J'$ and $\Psi_1'$ implicit in our assumptions (viz. (5.4) and (5.18)) do not guarantee that the right-hand side in (5.15), (5.17), (5.34), (5.36) and (5.38) is indeed $o(N^{-1})$ as is our aim. For this we would need $J'(t) = o(\{t(1-t)\}^{-1})$ and $\Psi_1'(t) = o(\{t(1-t)\}^{-1})$. This may explain the presence of the remaining expansions in theorems 5.1 and 5.2, which are less explicit but do have remainder $o(N^{-1})$ under the conditions stated. Note that their presence in theorem 5.2 also indicates that even for $J = -\Psi_1$, expansions for exact and approximate scores are not necessarily identical to $o(N^{-1})$. Finally we should point out that similar expansions with remainder $o(N^{-1})$ might have been given in theorem 4.2 of ABZ (1976) where they were unfortunately omitted.

We conclude this section with a few examples of the power expansions in theorems 5.1 and 5.2. First we consider the powers $\pi_W(\Phi,\theta)$ and $\pi_W(\Lambda,\theta)$ of Wilcoxon's two-sample test (W) against normal and logistic location alternatives $(\Phi(x),\Phi(x-\theta))$ and $(\Lambda(x),\Lambda(x-\theta))$ respectively, where $\Lambda(x) = (1+\exp\{-x\})^{-1}$ and $\theta = O(N^{-\frac{1}{2}})$. We find

$$(5.47)\qquad \pi_W(\Phi,\theta) = 1 - \Phi(u_\alpha - \bar{\eta}) + \frac{\bar{\eta}\phi(u_\alpha-\bar{\eta})}{N}\Big[-\tfrac{1}{2} - \frac{37-217\lambda+217\lambda^2}{20\lambda(1-\lambda)}(u_\alpha^2-1)$$

$$+ \frac{1}{\lambda(1-\lambda)}\left\{\frac{\sqrt{3}}{6} + \frac{67-437\lambda+437\lambda^2}{20}\right\}u_\alpha\bar{\eta} +$$

$$- \frac{1}{\lambda(1-\lambda)} \left\{ \frac{\sqrt{3}}{6} + \frac{\pi}{36} + \frac{29-219\lambda+219\lambda^2}{20} \right\} \bar{\eta}^2$$

$$+ \frac{(1-6\lambda+6\lambda^2)}{\lambda(1-\lambda)} \frac{6 \operatorname{arctn}\sqrt{2}}{\pi} \{u_\alpha^2 - 1 - 2u_\alpha\bar{\eta} + \bar{\eta}^2\} \Big] + o(N^{-1}),$$

where $\bar{\eta} = \left(\frac{3\lambda(1-\lambda)N}{\pi}\right)^{\frac{1}{2}} \theta$, and

$$(5.48) \qquad \pi_W(\Lambda,\theta) = 1 - \Phi(u_\alpha - \eta^*) + \frac{\eta^*\phi(u_\alpha-\eta^*)}{N} \Big[ -\tfrac{1}{2} - \frac{1-\lambda+\lambda^2}{20\lambda(1-\lambda)} (u_\alpha^2 - 1)$$

$$+ \frac{1-5\lambda+5\lambda^2}{20\lambda(1-\lambda)} u_\alpha\eta^* - \frac{1-3\lambda+3\lambda^2}{20\lambda(1-\lambda)} \eta^{*2} \Big] + o(N^{-1}),$$

where $\eta^* = \left(\frac{\lambda(1-\lambda)N}{3}\right)^{\frac{1}{2}} \theta$.

As a second example we compute expansions for the powers $\pi_{NS}(\Phi,\theta)$ and $\pi_{NS}(\Lambda,\theta)$ of the two-sample normal scores test against the normal and logistic location alternatives described above. The result is

$$(5.49) \qquad \pi_{NS}(\Phi,\theta) = 1 - \Phi(u_\alpha - \eta^*) + \frac{\eta^*\phi(u_\alpha-\eta^*)}{N} \Big[ \tfrac{1}{2} - \tfrac{1}{4}(u_\alpha^2 - 1)$$

$$- \int_0^{\Phi^{-1}(1-N^{-1})} \frac{\Phi(x)(1-\Phi(x))}{\phi(x)} dx \Big] + o(N^{-1}) ,$$

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

$$(5.50) \qquad \pi_{NS}(\Lambda,\theta) = 1 - \Phi(u_\alpha - \bar{\eta}) + \frac{\bar{\eta}\phi(u_\alpha-\bar{\eta})}{N} \Big[ -\frac{3}{2} - \frac{1-3\lambda+3\lambda^2}{12\lambda(1-\lambda)} (u_\alpha^2 - 1)$$

$$+ \int_0^{\Phi^{-1}(1-N^{-1})} \frac{\Phi(x)(1-\Phi(x))}{\phi(x)} dx$$

$$+ \left\{ \frac{\sqrt{3}(1-2\lambda)^2}{4\lambda(1-\lambda)} - \frac{\pi}{6} - \frac{4-21\lambda+21\lambda^2}{12\lambda(1-\lambda)} \right\} u_\alpha\bar{\eta}$$

$$+ \Big\{ \frac{6(1-5\lambda+5\lambda^2)}{\lambda(1-\lambda)} \operatorname{arctn}\sqrt{2} - \frac{\sqrt{3}(1-2\lambda)^2}{4\lambda(1-\lambda)} - \frac{11\pi(1-5\lambda+5\lambda^2)}{6\lambda(1-\lambda)}$$

$$+ \frac{5-21\lambda+21\lambda^2}{12\lambda(1-\lambda)} \Big\} \bar{\eta}^2 + o(N^{-1}) ,$$

where now $\bar{\eta} = \left(\frac{\lambda(1-\lambda)N}{\pi}\right)^{\frac{1}{2}}\theta$. Note that theorem 5.2 ensures that expansion (5.49) is also valid for van der Waerden's two-sample test which is based on the approximate scores $a_j = \Phi^{-1}(j/(N+1))$. To evaluate the integral in (5.49) and (5.50) we write (cf. (4.25) in ABZ (1976))

$$(5.51)\qquad \int_0^{\Phi^{-1}(1-N^{-1})} \frac{\Phi(x)(1-\Phi(x))}{\phi(x)}\,dx = \tfrac{1}{2}\log\log N + \tfrac{1}{2}\log 2 - 2\int_0^{\infty}\phi(x)\log x\,dx$$

$$+ \int_0^{\infty} \frac{(2\Phi(x)-1)\{x(1-\Phi(x))-\phi(x)\}}{x\phi(x)}\,dx + \int_0^{\infty}\frac{(1-\Phi(x))^2}{\phi(x)}\,dx + o(1) =$$

$$= \tfrac{1}{2}\log\log N + \tfrac{1}{2}\log 2 + 0.40489\ldots + o(1)\ ,$$

where the final result is obtained by numerical integration.

## APPENDIX. EXPANSIONS FOR THE CONTIGUOUS LOCATION CASE

In this appendix we provide the tools for deriving theorem 4.1 from theorem 3.1. The quantities appearing in the expansion of theorem 3.1 are expected values under $P_H$ of functions of $P_1,\ldots,P_N$ and in the set-up of section 4 both H and $P_1,\ldots,P_N$ depend on $\theta$. Our task is to provide Taylor expansions in $\theta$ with error bounds for these quantities, thus reducing expectations $E_H$ to expectations $E_F$ while at the same time expanding the r.v.'s involved. Since we are only concerned with the models $P_H$ and $P_F$ under the assumptions of section 4, we suppose throughout that $X_1,\ldots,X_N$ are i.i.d. with common density h under $P_H$ and f under $P_F$, where $h(x) = (1-\lambda)f(x) + \lambda f(x-\theta)$ and f is positive and four times differentiable on $R^1$. Define $\xi(x,t)$, $p(x,t)$ and $\tilde{p}(x,t)$ by

$$(A.1)\qquad (1-\lambda)F(\xi(x,t)) + \lambda F(\xi(x,t)-t) = F(x) ,$$

$$(A.2)\qquad p(x,t) = \frac{\lambda f(x-t)}{(1-\lambda)f(x)+\lambda f(x-t)} ,$$

$$(A.3)\qquad \tilde{p}(x,t) = p(\xi(x,t),t) .$$

As in appendix 1 of ABZ (1976), these functions are introduced because $\tilde{p}(Z_1,\theta),\ldots, \tilde{p}(Z_N,\theta)$ under $P_F$ have the same joint distribution as $P_1,\ldots,P_N$ under $P_H$. Our main problem is therefore to expand $\tilde{p}(x,t)$ as a function of t around $t = 0$.

With $\psi_i = f^{(i)}/f$ as in (4.1), we define for $i = 1,\ldots,4$,

$$(A.4)\qquad \chi_i(x,t) = |\psi_i(\xi(x,t))| + |\psi_i(\xi(x,t)-t)|$$

and for any function q of two variables we write

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

Then elementary but tedious computations yield

(A.5) $\tilde{p}(x,0) = \lambda,\ \tilde{p}_{0,1}(x,0) = -\lambda(1-\lambda)\psi_1(x),\ \tilde{p}_{0,2}(x,0) = \lambda(1-\lambda)(1-2\lambda)\psi_2(x),$

$$\tilde{p}_{0,3}(x,0) = -\lambda(1-\lambda)\{(1-3\lambda+3\lambda^2)\psi_3(x)-6\lambda(1-\lambda)\psi_1(x)\psi_2(x) + 3\lambda(1-\lambda)\psi_1^3(x)\} ,$$

(A.6) $|\tilde{p}_{0,1}| \le b_1\chi_1,\ |\tilde{p}_{0,2}| \le b_2(\chi_2+\chi_1^2),\ |\tilde{p}_{0,3}| \le b_3(\chi_3+\chi_2^{3/2}+\chi_1^3)$ ,

$$|\tilde{p}_{0,4}| \le b_4(\chi_4+\chi_3^{4/3}+\chi_2^2+\chi_1^4) ,$$

where $b_1,\ldots,b_4$ are positive constants.

Define $\pi_j = E_H P_j$ as in (3.16).

THEOREM A.1. *Suppose that positive numbers* C, C' *and* $\varepsilon'$ *exist such that* $\sum a_j^4 \le CN$, $0 \le \theta \le \varepsilon'$ *and* (4.2) *is satisfied. Then there exists* $B > 0$ *depending only on* C, C' *and* $\varepsilon'$ *such that*

(A.7) $$\sum a_j(\pi_j-\lambda) = \lambda(1-\lambda)\{-\theta \sum a_jE_F\psi_1(Z_j)+(1-2\lambda)\frac{\theta^2}{2} \sum a_jE_F\psi_2(Z_j) - \frac{\theta^3}{6} \sum a_jE_F[(1-3\lambda+3\lambda^2)\psi_3(Z_j)-6\lambda(1-\lambda)\psi_1(Z_j)\psi_2(Z_j)+3\lambda(1-\lambda)\psi_1^3(Z_j)]\}+M_1,$$

$$|M_1| \le BN^{5/4}\theta^4 ;$$

(A.8) $$\sum a_j^2(\pi_j-\lambda) = \lambda(1-\lambda)\{-\theta \sum a_j^2E_F\psi_1(Z_j)+(1-2\lambda)\frac{\theta^2}{2} \sum a_j^2E_F\psi_2(Z_j)\} + M_2,$$

$$|M_2| \le BN^{5/4}\theta^3 ;$$

(A.9) $$\sum a_j^3(\pi_j-\lambda) = -\lambda(1-\lambda)\theta \sum a_j^3E_F\psi_1(Z_j) + M_3 ,$$

$$|M_3| \le BN^{13/12}\theta^2 ;$$

(A.10) $$\sum a_j^2E_H(P_j-\lambda)^2 = \lambda^2(1-\lambda)^2\theta^2 \sum a_j^2E_F\psi_1^2(Z_j) + M_4 ,$$

$$|M_4| \le BN^{5/4}\theta^3 ;$$

$$\text{(A.11)}\qquad \sigma_H^2(\sum a_jP_j) = \lambda^2(1-\lambda)^2\theta^2\sigma_F^2(\sum a_j\psi_1(Z_j)) + M_5 ,$$

$$|M_5| \le B\{N^2\theta^{24/5} + N\theta^{19/5}[E_F|\sum a_j(\psi_1(Z_j) - E_F\psi_1(Z_j))|^3]^{1/3}$$

$$+ \theta^3\sigma_F(\sum a_j\psi_1(Z_j))\sigma_F(\sum a_j\psi_2(Z_j)) + \theta^4\sigma_F^2(\sum a_j\psi_2(Z_j))\} ;$$

$$\text{(A.12)}\qquad E_H\left(\frac{\lambda g(X_1)}{h(X_1)} - \lambda\right)^4 \le B\theta^4 ;$$

$$\text{(A.13)}\qquad \left[\sum\{E_H|P_j - \pi_j|^3\}^{4/9}\right]^{9/4} \le \theta^3\left[\sum\{E_F|\psi_1(Z_j)-E_F\psi_1(Z_j)|^3\}^{4/9}\right]^{9/4} + BN^{9/4}\theta^6.$$

PROOF. Although the proof is very similar to that of theorem A1.1 and the relevant part of corollary A1.1 in ABZ (1976), there are additional complications due to the fact that now $\tilde{p}_{0,2}(x,0) \not\equiv 0$. We begin by noting that the distribution of $\xi(X_1,t)$ under F is that of $X_1$ under $\lambda F(x) + (1-\lambda)F(x-t)$, so that (4.2) and (A.6) imply the existence of $B_1 > 0$ depending only on C' and such that

$$\text{(A.14)}\qquad \sup\{E_F|\tilde{p}_{0,i}(X_1,\nu\theta)|^{m_i}: 0 \le \nu \le 1\} \le B_1, \qquad i = 1,\dots,4,$$

where $m_1 = 6$, $m_2 = 3$, $m_3 = 4/3$, $m_4 = 1$.

Using lemma A1.1 of ABZ (1976) together with $\sum a_j^4 \le CN$ and (A.14), we find that

$$|M_1| \le \frac{\theta^4}{24}\sup\{\sum |a_j|E_F|\tilde{p}_{0,4}(Z_j,\nu\theta)|: 0 \le \nu \le 1\}$$

$$\le \frac{(CN)^{\frac14}\theta^4}{24}\sup\{NE_F|\tilde{p}_{0,4}(X_1,\nu\theta)|: 0 \le \nu \le 1\} \le \frac{B_1C^{\frac14}}{24}N^{5/4}\theta^4,$$

$$|M_2| \le \frac{\theta^3}{6}\sup\{\sum a_j^2E_F|\tilde{p}_{0,3}(Z_j,\nu\theta)|: 0 \le \nu \le 1\}$$

$$\le \frac{\theta^3}{6}\left(\sum a_j^8\right)^{\frac14}\sup\{[NE_F|\tilde{p}_{0,3}(X_1,\nu\theta)|^{4/3}]^{3/4}: 0 \le \nu \le 1\}$$

$$\le \frac{B_1^{\frac34}C^{\frac12}}{6}N^{5/4}\theta^3 ,$$

$$|M_3| \le \frac{\theta^2}{2} \sup\{\sum |a_j|^3 E_F|\tilde{p}_{0,2}(Z_j,\nu\theta)| : 0 \le \nu \le 1\}$$

$$\le \frac{\theta^2}{2}\left(\sum |a_j|^{9/2}\right)^{2/3}(NB_1)^{1/3} \le \frac{B_1^{1/3}C^{3/4}}{2} N^{13/12}\theta^2 ,$$

$$|M_4| \le \frac{\theta^3}{6} \sup\{\sum a_j^2 E_F[2|\tilde{p}_{0,3}(Z_j,\nu\theta)|$$

$$+ 6|\tilde{p}_{0,1}(Z_j,\nu\theta)\tilde{p}_{0,2}(Z_j,\nu\theta)|] : 0 \le \nu \le 1\}$$

$$\le \frac{\theta^3}{6}\left[2\left(\sum a_j^8\right)^{\frac{1}{4}}(NB_1)^{\frac{3}{4}} + 6\left(\sum a_j^4\right)^{\frac{1}{2}}(NB_1)^{\frac{1}{2}}\right] \le (B_1^{\frac{1}{2}}+B_1^{\frac{3}{4}})C^{\frac{1}{2}}N^{5/4}\theta^3 ,$$

$$E_H\left(\frac{\lambda g(X_1)}{h(X_1)} - \lambda\right)^4 \le \theta^4 \sup\{E_F\tilde{p}^4_{0,1}(X_1,\nu\theta) : 0 \le \nu \le 1\} \le B_1^{2/3}\theta^4 ,$$

which proves (A.7) - (A.10) and (A.12). To establish (A.13) we note that

$$|\tilde{p}(Z_j,\theta) - E_F\tilde{p}(Z_j,\theta)| \le \theta|\tilde{p}_{0,1}(Z_j,0) - E_F\tilde{p}_{0,1}(Z_j,0)|$$
$$+ \frac{\theta^2}{2}\int_0^1 2(1-\nu)\{|\tilde{p}_{0,2}(Z_j,\nu\theta)| + E_F|\tilde{p}_{0,2}(Z_j,\nu\theta)|\}d\nu.$$

Hence

$$E_H|P_j-\pi_j|^3 \le \frac{\theta^3}{16} E_F|\psi_1(Z_j)-E_F\psi_1(Z_j)|^3+4\theta^6\int_0^1 2(1-\nu)E_F|\tilde{p}_{0,2}(Z_j,\nu\theta)|^3 d\nu,$$

$$\sum\{E_H|P_j-\pi_j|^3\}^{4/9} \le \theta^{4/3}\sum\{E_F|\psi_1(Z_j)-E_F\psi_1(Z_j)|^3\}^{4/9}+2(B_1+1)N\theta^{8/3} ,$$

and (A.13) follows.

It remains to prove (A.11). We have

$$\tilde{p}(x,t) - \lambda + \lambda(1-\lambda)t\psi_1(x) - \tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)t^2\psi_2(x)$$
$$= \frac{t^2}{2}\int_0^1 2(1-\nu)(\tilde{p}_{0,2}(x,\nu t) - \tilde{p}_{02}(x,0))d\nu = \frac{t^3}{6}\int_0^1 3(1-\nu)^2\tilde{p}_{0,3}(x,\nu t)d\nu,$$

and as a result

$$(\tilde{p}(x,t) - \lambda + \lambda(1-\lambda)t\psi_1(x) - \tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)t^2\psi_2(x))^2$$

$$\le \left|\frac{t^2}{2}\int_0^1 2(1-\nu)(\tilde{p}_{0,2}(x,\nu t)-\tilde{p}_{0,2}(x,0))d\nu\right|^{6/5}$$

$$\times \left|\frac{t^3}{6}\int_0^1 3(1-\nu)^2\tilde{p}_{0,3}(x,\nu t)d\nu\right|^{4/5}$$

$$\le |t|^{24/5}\{\left|\tfrac{1}{2}\int_0^1 2(1-\nu)(\tilde{p}_{0,2}(x,\nu t)-\tilde{p}_{0,2}(x,0))d\nu\right|^3$$

$$+ \left|\frac{1}{6}\int_0^1 3(1-\nu)^2\tilde{p}_{0,3}(x,\nu t)d\nu\right|^{4/3}\}$$

$$\le |t|^{24/5}\int_0^1 \{|\tilde{p}_{0,2}(x,\nu t)|^3 + |\tilde{p}_{0,2}(x,0)|^3 + |\tilde{p}_{0,3}(x,\nu t)|^{4/3}\}d\nu \ .$$

Similarly,

$$|\tilde{p}(x,t) - \lambda + \lambda(1-\lambda)t\psi_1(x) - \tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)t^2\psi_2(x)|^{3/2}$$

$$\le |t|^{21/5}\int_0^1 \{|\tilde{p}_{0,2}(x,\nu t)|^3 + |\tilde{p}_{0,2}(x,0)|^3 + |\tilde{p}_{0,3}(x,\nu t)|^{4/3}\}d\nu.$$

It follows that

$$\sigma_F^2\Big(\sum a_j\{\tilde{p}(Z_j,\theta)+\lambda(1-\lambda)\theta\psi_1(Z_j)-\tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)\theta^2\psi_2(Z_j)\}\Big)$$

$$\le N\sum a_j^2E_F(\tilde{p}(X_1,\theta)-\lambda+\lambda(1-\lambda)\theta\psi_1(X_1)-\tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)\theta^2\psi_2(X_1))^2$$

$$\le 3B_1C^{\frac{1}{2}}N^2\theta^{24/5} \ ,$$

$$\Big|\mathrm{cov}_F\Big(\sum a_j\{\tilde{p}(Z_j,\theta)+\lambda(1-\lambda)\theta\psi_1(Z_j)-\tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)\theta^2\psi_2(Z_j)\},\ \sum a_j\psi_1(Z_j)\Big)\Big|$$

$$\le [E_F|\sum a_j\{\tilde{p}(Z_j,\theta)-\lambda+\lambda(1-\lambda)\theta\psi_1(Z_j)-\tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)\theta^2\psi_2(Z_j)\}|^{3/2}]^{2/3}$$

$$\times [E_F|\sum a_j(\psi_1(Z_j) - E_F\psi_1(Z_j))|^3]^{1/3} \le$$

$$\leq \left[\left(\sum|a_j^3|\right)^{\frac{1}{2}} NE_F|\tilde{p}(X_1,\theta)-\lambda+\lambda(1-\lambda)\theta\psi_1(X_1)-\tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)\theta^2\psi_2(X_1)|^{3/2}\right]^{2/3}$$

$$\times\ [E_F|\sum a_j(\psi_1(Z_j) - E_F\psi_1(Z_j))|^3]^{1/3}$$

$$\leq (3B_1)^{2/3}C^{1/4}N\theta^{14/5}[E_F|\sum a_j(\psi_1(Z_j) - E_F\psi_1(Z_j))|^3]^{1/3} ,$$

$$|\mathrm{cov}_F\left(\sum a_j\{\tilde{p}(Z_j,\theta)+\lambda(1-\lambda)\theta\psi_1(Z_j)-\tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)\theta^2\psi_2(Z_j)\},\ \sum a_j\psi_2(Z_j)\right)|$$

$$\leq (3B_1)^{1/2}C^{1/4}N\theta^{12/5}\sigma_F(\sum a_j\psi_2(Z_j)) .$$

These inequalities ensures that there exists $B_2 > 0$ depending only on $B_1$ and C such that

$$|\sigma_H^2\left(\sum a_jP_j\right)-\sigma_F^2\left(\sum a_j\{\lambda(1-\lambda)\theta\psi_1(Z_j)-\tfrac{1}{2}\lambda(1-\lambda)(1-2\lambda)\theta^2\psi_2(Z_j)\}\right)|$$

$$\leq B_2\{N^2\theta^{24/5}+N\theta^{19/5}[E_F|\sum a_j(\psi_1(Z_j)-E_F\psi_1(Z_j))|^3]^{1/3}$$

$$+ N\theta^{22/5}\sigma_F(\sum a_j\psi_2(Z_j))\} .$$

Since $N\theta^{22/5}\sigma_F(\sum a_j\psi_2(Z_j)) \leq N^2\theta^{24/5} + \theta^4\sigma_F^2(\sum a_j\psi_2(Z_j))$, (A.11) follows immediately and the proof of the theorem is complete. □

<u>COROLLARY A.1</u>. *Suppose that* (3.1) *and* (4.3) *hold and that positive numbers* c, C, C', D, ε *and* ε' *exist such that* (3.10), (3.19), (4.2) *and* (4.4) *are satisfied. Let* K, $\alpha_i$, $\tilde{K}$, $\tilde{\alpha}_i$ *and* η *be defined by* (3.17), (3.18), (4.5), (4.6) *and* (4.7). *Then there exists* B > 0 *depending only on* c, C, C', D, ε *and* ε' *such that*

$$\text{(A.15)}\qquad \sup_x\left|K\left(x - \frac{\Sigma a_j\pi_j}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac{1}{2}}}\right) - \tilde{K}(x-\eta)\right| \leq B\Big\{N^{-5/4}$$

$$+ N^{-\frac{1}{2}}\theta^3\left[\sum\{E_F|\psi_1(Z_j)-E_F\psi_1(Z_j)|^3\}^{4/9}\right]^{9/4}$$

$$+ N^{-3/4}\theta^3\left[\sum\{E_F(\psi_2(Z_j)-E_F\psi_2(Z_j))^2\}^{2/3}\right]^{3/2}\Big\} ,$$

(A.16) $$\theta^2 \frac{|\Sigma a_j E_F \psi_2(Z_j)|}{(\Sigma a_j^2)^{\frac{1}{2}}} \le BN^{-\frac{1}{2}}, \quad \theta \frac{|\Sigma a_j^2 E_F \psi_1(Z_j)|}{\Sigma a_j^2} \le BN^{-\frac{1}{2}}, \quad \frac{|\Sigma a_j^3|}{(\Sigma a_j^2)^{3/2}} \le BN^{-\frac{1}{2}},$$

(A.17) $$\theta^2 \frac{\sigma_F^2(\Sigma a_j \psi_1(Z_j))}{\Sigma a_j^2} \le B\left\{N^{-1} + N^{-\frac{1}{4}}\theta^3\left[\sum \{E_F|\psi_1(Z_j) - E_F\psi_1(Z_j)|^3\}^{4/9}\right]^{9/4}\right\},$$

*and all other terms occurring in* $\tilde{\alpha}_0,\ldots,\tilde{\alpha}_5$ *are bounded in absolute value by* $BN^{-1}$.

PROOF. In this proof $O(x)$ will denote a quantity that is bounded by $B_1|x|$ with $B_1$ depending only on c, C, C', D, $\varepsilon$ and $\varepsilon'$.

We begin by noting that (A.16) and the last statement in corollary A.1 are immediate consequences of Hölder's inequality, (3.10), (4.2) and (4.4). Also

(A.18) $$\theta^2\sigma_F^2\left(\sum a_j\psi_1(Z_j)\right) \le 1 + \theta^3\sigma_F^3\left(\sum a_j\psi_1(Z_j)\right)$$

$$\le 1 + \theta^3 E_F \left| \sum a_j(\psi_1(Z_j) - E_F\psi_1(Z_j))\right|^3$$

$$\le 1 + \theta^3\left[\sum |a_j|\{E_F|\psi_1(Z_j) - E_F\psi_1(Z_j)|^3\}^{1/3}\right]^3$$

$$\le 1 + \theta^3\left(\sum a_j^4\right)^{3/4}\left[\sum \{E_F|\psi_1(Z_j) - E_F\psi_1(Z_j)|^3\}^{4/9}\right]^{9/4},$$

and in view of (3.10) and (4.4), this implies (A.17). For later use we note that similarly

(A.19) $$\sigma_F^2\left(\sum a_j\psi_2(Z_j)\right) \le C^{\frac{1}{2}}N^{\frac{1}{2}}\left[\sum\{E_F(\psi_2(Z_j) - E_F\psi_2(Z_j))^2\}^{2/3}\right]^{3/2}.$$

It remains to prove (A.15). Since (A.15) is trivially satisfied for $N < (D/\varepsilon')^2$, we may assume that $0 \le \theta \le \varepsilon'$ so that theorem A.1 applies. Because of (3.1), $\sum a_j\pi_j = \sum a_j(\pi_j-\lambda)$. In view of the bounds obtained above, we can truncate expansions (A.7) and (A.8) to

(A.20) $$\sum a_j\pi_j = \lambda(1-\lambda)\{-\theta\sum a_jE_F\psi_1(Z_j)+(1-2\lambda)\frac{\theta^2}{2}\sum a_jE_F\psi_2(Z_j)\} + O(N\theta^3)$$

$$= -\lambda(1-\lambda)\theta\sum a_jE_F\psi_1(Z_j) + O(N\theta^2) = O(N\theta),$$

(A.21) $$\sum a_j^2(\pi_j-\lambda) = -\lambda(1-\lambda)\theta\sum a_j^2E_F\psi_1(Z_j) + O(1) = O(N^{\frac{1}{2}}).$$

Using (A.8) - (A.11), (A.20), (A.21), (3.10), (3.19) and (4.4) we expand $\alpha_0,\ldots,\alpha_5$ and find

(A.22) $$\sup_x|K(x) - \hat{K}(x)| = O\Big(N^{-5/4}+\theta^{19/5}[E_F|\sum a_j(\psi_1(Z_j) - E_F\psi_1(Z_j))|^3]^{1/3}$$

$$+ N^{-1}\theta^3\sigma_F\Big(\sum a_j\psi_1(Z_j)\Big)\sigma_F\Big(\sum a_j\psi_2(Z_j)\Big) + N^{-1}\theta^4\sigma_F^2\Big(\sum a_j\psi_2(Z_j)\Big)\Big),$$

where

(A.23) $$\hat{K}(x) = \Phi(x) - \phi(x)\sum_{k=0}^{5}\hat{\alpha}_kH_k(x),$$

(A.24) $$\hat{\alpha}_0 = -\left(\frac{\lambda(1-\lambda)}{\Sigma a_j^2}\right)^{\frac{1}{2}} N^{-1}\theta\sum a_jE_F\psi_1(Z_j),$$

$$\hat{\alpha}_1 = \tilde{\alpha}_1 - \frac{1}{8\Sigma a_j^2}\lambda(1-\lambda)(1-2\lambda)^2\theta^4\{\sum a_jE_F\psi_2(Z_j)\}^2,$$

$$\hat{\alpha}_2 = \tilde{\alpha}_2 + \frac{\{\lambda(1-\lambda)\}^{\frac{1}{2}}}{4(\Sigma a_j^2)^{3/2}}(1-2\lambda)^2\theta^3\sum a_j^2E_F\psi_1(Z_j)\sum a_jE_F\psi_2(Z_j),$$

$$\hat{\alpha}_3 = \tilde{\alpha}_3 - \frac{1}{12(\Sigma a_j^2)^2}(1-2\lambda)^2\theta^2\sum a_j^3\sum a_jE_F\psi_2(Z_j),$$

$$\hat{\alpha}_k = \tilde{\alpha}_k \quad \text{for } k = 4,5,$$

with $\tilde{\alpha}_k$ as given by (4.6). By applying elementary inequalities (A.22) may be simplified to

(A.25) $$\sup_x|K(x) - \hat{K}(x)| = O\Big(N^{-5/4}+N^{-5/4}\theta^3E_F|\sum a_j(\psi_1(Z_j) - E_F\psi_1(Z_j))|^3$$

$$+ N^{-5/4}\theta^3\sigma_F^2\Big(\sum a_j\psi_2(Z_j)\Big)\Big).$$

With the aid of (A.7), (A.20) and the bounds obtained in the first part of the proof we now expand $\hat{K}(x-\sum a_j\pi_j\{\lambda(1-\lambda)\sum a_j^2\}^{-\frac{1}{2}})$ about the point $(x-\eta)$ and obtain

(A.26) $$\sup_x\left|\hat{K}\left(x - \frac{\Sigma a_j\pi_j}{\{\lambda(1-\lambda)\Sigma a_j^2\}^{\frac{1}{2}}}\right) - \tilde{K}(x-\eta)\right| = \mathcal{O}\left(N^{-5/4}+N^{-1}\theta^3\sigma_F^2\left(\sum a_j\psi_1(Z_j)\right)\right)$$

with $\tilde{K}$ as given by (4.5). Combining (A.25), (A.26), (A.18) and (A.19) we see that (A.15) and corollary A.1 are proved. □

REFERENCES

ALBERS, W., BICKEL, P.J. & VAN ZWET, W.R. [1976], *Asymptotic expansions for the power of distributionfree tests in the one-sample problem*, Ann. Statist., 4, 108-156.

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

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

HÁJEK, J. [1964], *Asymptotic theory of rejective sampling with varying probabilities from finite populations*, Ann. Math. Statist., 35, 1491-1523.

HODGES, J.L. & LEHMANN, E.L. [1970], *Deficiency*, Ann. Math. Statist., 41, 783-801.

OKAMOTO, M. [1958], *Some inequalities relating to the partial sum of binomial probabilities*, Ann. Inst. Statist. Math., 10, 29-35.

RÉNYI, A. [1970], *Probability Theory*, North Holland, Amsterdam.

ROGERS, W.F. [1971], *Exact null distributions and asymptotic expansions for rank test statistics*, Technical Report 145, Departments of O.R. and Statistics, Stanford University.