ma
the
ma
tisch

cen
trum

MC

AFDELING MATHEMATISCHE STATISTIEK SW 44/78 JUNI
(DEPARTMENT OF MATHEMATICAL STATISTICS)

R. HELMERS

EDGEWORTH EXPANSIONS FOR LINEAR COMBINATIONS OF ORDER STATISTICS WITH SMOOTH WEIGHT FUNCTIONS

Preprint

Revised edition

amsterdam 1978

stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK SW 44/78 JUNI
(DEPARTMENT OF MATHEMATICAL STATISTICS)

R. HELMERS

EDGEWORTH EXPANSIONS FOR LINEAR COMBINATIONS
OF ORDER STATISTICS WITH SMOOTH WEIGHT FUNCTIONS

Preprint

Revised edition

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).

AMS(MOS) subject classification scheme (1970): Primary 62G30; Secondary 62E20

First Printing 1976
Second Edition 1978

Edgeworth expansions for linear combinations of order statistics with smooth weight functions*)

by

R. Helmers

ABSTRACT

Edgeworth expansions for linear combinations of order statistics with smooth weight functions are established.

KEY WORDS & PHRASES: *Edgeworth expansions, linear combinations of order statistics.*

*) This report will be submitted for publication elsewhere.

1. INTRODUCTION

Statistics of the form $T_n = n^{-1} \Sigma_{i=1}^{n} c_{in} X_{in}$, $n \geq 1$, where X_{in} $i = 1, 2,\ldots,n$ denotes the ith order statistic of a random sample $X_1,\ldots,X_n$ of size n from a distribution with distribution function (d.f.) F and the c_{in}, $i = 1,2,\ldots,n$ are known real numbers (weights), are said to be linear combinations of order statistics. In the last decade there has been considerable interest in these statistics with regard to the problem of their asymptotic normality, which has been investigated under different sets of conditions by many authors in this area. We refer to the important papers of SHORACK (1972) and STIGLER (1974) and the references given in these papers. More recently attention has been paid to the rate of convergence problem. Berry-Esseen type bounds for linear combinations of order statistics were established by BJERVE (1977) and HELMERS (1977a), (1977b). An account of the first two of these results is given by VAN ZWET (1977).

The purpose of this paper is to establish Edgeworth expansions for linear combinations of order statistics with remainder $o(n^{-1})$ for the case of smooth weights. Our method of proof was outlined by VAN ZWET (1977). In his paper he obtained a bound on the characteristic function of a linear combination of order statistics which solves a crucial part of our problem. An important drawback of the approach followed in the present paper is that our results do not include trimmed means. However BJERVE (1974) has shown that trimmed means admit asymptotic expansions. His method employs special properties of the trimmed means and does not carry over to the more general linear combinations of order statistics we consider. In HELMERS (1978) Edgeworth expansions for trimmed linear combinations of order statistics are established. Though there is a strong similarity between the proofs given in HELMERS (1978) and the present paper a major difference is that totally different representations of a linear combination of order statistics are used.

The paper is organized as follows. In section 2 we state our results in the form of two theorems. Section 3 contains a number of preliminaries. Theorem 2.1 is proved in section 4 and theorem 2.2 in section 5.

2. THE RESULTS

Let J be a bounded function on (0,1), which is three times differentiable with first, second and third derivative J', J'' and J''' on (0,1). Let J''' be bounded on (0,1) and let F be a d.f. with finite fourth moment. The inverse of a d.f. will always be the left-continuous one. χ_E denotes the indicator of a set E. Let $\|h\| = \sup_{0<s<1} |h(s)|$ for any function h on (0,1). Introduce functions h_1, h_2 and h_3 by

(2.1) $$h_1(u) = -\int_0^1 J(s)(\chi_{(0,s]}(u)-s)dF^{-1}(s)$$

(2.2) $$h_2(u,v) = -\int_0^1 J'(s)(\chi_{(0,s]}(u)-s)(\chi_{(0,s]}(v)-s)dF^{-1}(s)$$

(2.3) $$h_3(u,v,w) = -\int_0^1 J''(s)(\chi_{(0,s]}(u)-s)(\chi_{(0,s]}(v)-s)(\chi_{(0,s]}(w)-s)dF^{-1}(s)$$

for $0 < u,v,w < 1$. Furthermore define, for each $n \geq 1$ and real x, the function K_n by

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

where Φ and ϕ denote the d.f. and the density of the standard normal distribution. The quantities $\kappa_3 = \kappa_3(J,F)$ and $\kappa_4 = \kappa_4(J,F)$ are given by

(2.5) $$\kappa_3 = \kappa_3(J,F) = \frac{1}{\sigma^3(J,F)}\left[\int_0^1 h_1^3(u)du+3\int_0^1\int_0^1 h_1(u)h_1(v)h_2(u,v)dudv\right]$$

and

(2.6) $$\kappa_4 = \kappa_4(J,F) = \frac{1}{\sigma^4(J,F)}\left[\int_0^1 h_1^4(u)du-3\sigma^4(J_1,F) +\right.$$

$$+ 12 \int_0^1\int_0^1 h_1^2(u)h_1(v)h_2(u,v)dudv + \int_0^1\int_0^1\int_0^1 (4h_1(u)h_1(v)h_1(w)h_3(u,v,w) +$$

$$+ 12h_1(u)h_1(v)h_2(u,w)h_2(v,w))dudvdw\Big]$$

where

(2.7) $$\sigma^2 = \sigma^2(J,F) = \int_0^1 h_1^2(u)du.$$

In our first theorem we shall establish an asymptotic expansion with remainder $o(n^{-1})$ for the d.f. $F_n^*(x) = P(T_n^* \le x)$ for $-\infty < x < \infty$ where

(2.8) $$T_n^* = (T_n - E(T_n))/\sigma(T_n)$$

for the case of smooth weights.

<u>THEOREM 1</u>. *Suppose that positive numbers* c, m, M_1, M_2 *and numbers* $0 \le t_1 < t_2 \le 1$ *exist such that*

(2.9) $$c_{in} = J(\tfrac{i}{n+1}) \qquad \text{for } i = 1,2,\ldots,n,\ n = 1,2,\ldots;$$

J *is three times differentiable on* $(0,1)$ *with first, second and third bounded derivative* J', J'' *and* J''' *on* $(0,1)$. *Suppose further that*

(2.10) $$J(s) \ge c \qquad \text{for } t_1 < s < t_2.$$

and that on $(F^{-1}(t_1), F^{-1}(t_2))$ F *is twice differentiable with density* f *and second derivative* f' *such that on* $(F^{-1}(t_1), F^{-1}(t_2))$

(2.11) $$m \le f \le M_1 \quad \text{and} \quad |f'| \le M_2.$$

Suppose also that F *possesses a finite fourth moment* $\beta_4 = EX_1^4$. *Then we have that*

$$\lim_{n\to\infty} n \sup_x |F_n^*(x) - K_n(x)| = 0.$$

Our second theorem is a modification of theorem 2.1 which lends itself better to applications. We shall establish an asymptotic expansion with remainder $o(n^{-1})$ for the d.f. $G_n(x) = P(n^{\frac{1}{2}}(T_n-\mu)/\sigma\le x)$ for $-\infty < x < \infty$ where

$$\mu = \mu(J,F) = \int_0^1 F^{-1}(s)J(s)ds \tag{2.12}$$

and $\sigma^2 = \sigma^2(J,F)$ as in (2.7). Introduce a function h_4 by

$$h_4(u) = -\int_0^1 (\tfrac{1}{2}-s)J'(s)(\chi_{(0,s]}(u)-s)dF^{-1}(s) \tag{2.13}$$

for $0 < u < 1$. Furthermore quantities $a = a(J,F)$ and $b = b(J,F)$ are given by

$$a = a(J,F) = \frac{1}{\sigma(J,F)}\Big[2^{-1}\int_0^1 s(1-s)J'(s)dF^{-1}(s) - \int_0^1 F^{-1}(s)(\tfrac{1}{2}-s)J'(s)ds\Big] \tag{2.14}$$

and

$$b = b(J,F) = \frac{1}{2\sigma^2(J,F)}\Big[\int_0^1 (h_1(u)h_2(u,u) + 2h_1(u)h_4(u))du + \int_0^1\int_0^1 (2^{-1}h_2^2(u,v) + h_1(u)h_3(u,v,v))dudv\Big]. \tag{2.15}$$

Finally define, for each $n \ge 1$ and real x, the function L_n by

$$L_n(x) = K_n(x) - \phi(x)\Big[-\frac{a}{n^{\frac{1}{2}}} + \frac{(a\kappa_3+a^2+2b)}{2n}x - \frac{a\kappa_3}{6n}x^3\Big]. \tag{2.16}$$

THEOREM 2.2. *Suppose that the assumptions of theorem 2.1 are satisfied. Then we have that*

$$\lim_{n\to\infty} n \sup_x |G_n(x) - L_n(x)| = 0$$

It may be useful to comment briefly on these results. In the first place we remark that it is not difficult to check from our proofs that assumption (2.9) (see CHERNOFF et al. (1967) and STIGLER (1974) were the same type of weights are considered) can be replaced by the weaker condition that

$$\max_{1\le i\le n} |c_{in} - n \int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)ds - \int_{\frac{i-1}{n}}^{\frac{i}{n}} M(s)ds| = \mathcal{O}(n^{-\gamma}), \quad \text{as } n \to \infty,$$

for some $\gamma > 1$ and smooth functions J and M(J(M) must have bounded third (first) derivative on (0,1)),provided we replace the factor $(\frac{1}{2} - s)J'(s)$ appearing in the integrands of (2.13) and (2.14) by M(s). In particular this condition is satisfied in either one of the following cases: $c_{in} = J(\frac{i}{n})$ (see MOORE (1968)) or

$$c_{in} = n \int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)ds$$

(see BICKEL (1967)) with $M(s) = \frac{1}{2} J'(s)$ and $M(s) = 0$ respectively.

Secondly we note that the assumptions (2.10) and (2.11) are needed to ensure sufficient smoothness of F_n^* and G_n, which is what CRAMER's condition (C) (see CRAMER (1962)) does in the classical case of sums of independent and identically distributed random variables (cf. the proof of relation (4.2); see also VAN ZWET (1977)).

Next we give a numerical example which indicates that the expansions given in this paper perform well as approximations of the finite sample exact d.f.'s. It also shows that they are better than the usual normal approximation.

We consider the asymptotically first order efficient estimator, based on linear combinations of order statistics, of the centre θ of the logistic distribution

$$F(x) = \left[1 + e^{-(x-\theta)}\right]^{-1} \quad \text{for } -\infty < x < \infty$$

which is (see e.g. DAVID (1970) p.224) given by the weight function

$$J(s) = 6s(1-s), \qquad 0 < s < 1.$$

As in this case the conditions of theorem 2.2 are easily verified we find after long but straightforward computations

$$L_n(x) = \Phi(x)-\phi(x)\left[\frac{1}{20n}(x^3-3x) + \frac{(\pi^2-9)}{n} x\right].$$

In the following table we give the numerical results. The exact d.f. $G_n(x)$ is computed by numerical integration of the multiple integrals involved in this computation for n = 3 and n = 4 and by Monte-Carlo simulation based on 25.000 samples for n = 10 and n = 25. The agreement between G_n and L_n is already reasonable for n = 3.

x	G_3	L_3	G_4	L_4	G_{10}	L_{10}	G_{25}	L_{25}	Φ
0.0	.5000	.5001	.5000	.5000	.5000	.5000	.4991	.5000	.5000
0.2	.5640	.5604	.5663	.5652	.5734	.5736	.5758	.5770	.5793
0.4	.6262	.6197	.6307	.6286	.6445	.6447	.6492	.6511	.6554
0.6	.6850	.6766	.6919	.6889	.7089	.7110	.7152	.7198	.7257
0.8	.7391	.7301	.7469	.7446	.7680	.7707	.7728	.7812	.7881
1.0	.7875	.7793	.7963	.7948	.8196	.8227	.8295	.8339	.8413
1.2	.8248	.8234	.8391	.8388	.8629	.8665	.8756	.8776	.8849
1.4	.8658	.8621	.8752	.8764	.8985	.9021	.9100	.9124	.9192
1.6	.8958	.8951	.9049	.9076	.9275	.9302	.9376	.9392	.9452
1.8	.9202	.9223	.9287	.9327	.9486	.9515	.9580	.9591	.9641
2.0	.9397	.9441	.9474	.9524	.9646	.9673	.9732	.9733	.9772
2.2	.9550	.9611	.9618	.9673	.9764	.9786	.9830	.9831	.9861
2.4	.9669	.9738	.9726	.9783	.9845	.9864	.9895	.9896	.9918
2.6	.9758	.9829	.9807	.9860	.9905	.9916	.9942	.9938	.9953
2.8	.9825	.9892	.9865	.9913	.9937	.9950	.9963	.9965	.9974
3.0	.9875	.9935	.9907	.9948	.9959	.9971	.9982	.9980	.9987

To conclude this section it may be mentioned that an important application of the asymptotic expansions established in this paper lies in the computation of asymptotic deficiencies in the sense of HODGES and LEHMANN (1970) for estimators and tests based on linear combinations of order statistics. These computations will be given in a separate paper. Here we note only that in the case that F is symmetric about its expectation (say, zero) and J is symmetric about $\frac{1}{2}$ there is no term of order $n^{-\frac{1}{2}}$ in the expansions. We also note (we omit the details) that in the asymmetric case the phenomenon, first noted by PFANZAGL (1977), that "first order efficiency implies second order efficiency" (see also BICKEL & VAN ZWET (1977), p.4 and p.74), also holds true for linear combinations of order statistics.

3. PRELIMINARIES

In this section we present a number of preliminary results which will be needed in our proofs.

Let, for each $n \geq 1$, $U_1,\ldots,U_n$ be independent uniform (0,1) r.v.'s and let U_{in} $(1 \leq i \leq n)$ denote the ith order statistic of $U_1,\ldots,U_n$. It is well-known that the joint distribution of $X_1,\ldots,X_n$ is the same as that of $(F^{-1}(U_1),\ldots,F^{-1}(U_n))$ for any d.f.F. Therefore we shall identify X_i with $F^{-1}(U_i)$ and also X_{in} with $F^{-1}(U_{in})$. The empirical d.f. based on $U_1,\ldots,U_n$ will be denoted by Γ_n. Throughout this paper we shall assume that all r.v.'s are defined on the same probability space (Ω,A,P). For any r.v.X with $0 < \sigma(X) < \infty$ we write $\hat{X} = X - E(X)$ and $X^* = \hat{X}/\sigma(X)$. For any positive number ℓ the ℓth absolute moment of F will be denoted by β_ℓ.
We start by stating a very simple but useful lemma.

<u>LEMMA 3.1</u>. *Let $\{X_n\}$ and $\{Y_n\}$ be two sequences of r.v.'s and let there exist a number $\eta > 1$ such that*
(i) $\sigma^2(X_n-Y_n) = O(n^{-\eta})$ *and*
(ii) $\sigma^{-2}(X_n) = O(n)$, *as* $n \to \infty$.
Then we have that $\sigma^2(X_n^*-Y_n^*) = O(n^{-\eta+1})$, *as* $n \to \infty$.

<u>PROOF</u>. Note first that

$$\sigma^2(X_n-Y_n) = (\sigma(X_n)-\sigma(Y_n))^2 + 2(1-\rho_n)\sigma(X_n)\sigma(Y_n) \tag{3.1}$$

where ρ_n denotes the correlation coefficient of X_n and Y_n. Because of assumption (i) and the fact that each of the terms on the right of (3.1) is nonnegative we find that

$$(3.2) \qquad \sigma(X_n)-\sigma(Y_n) = \mathcal{O}(n^{-\frac{\eta}{2}}), \qquad \text{as } n \to \infty.$$

and also that

$$(3.3) \qquad 2(1-\rho_n)\sigma(X_n)\sigma(Y_n) = \mathcal{O}(n^{-\eta}), \qquad \text{as } n \to \infty.$$

Using now the assumption (ii) and (3.2) and noting that $\eta > 1$ we see that $\sigma^{-2}(Y_n) = \mathcal{O}(n)$ as $n \to \infty$. Combining this and assumption (ii) with (3.3) we find that $2(1-\rho_n) = \mathcal{O}(n^{-\eta+1})$ as $n \to \infty$. Because $\sigma^2(X_n^*-Y_n^*) = 2(1-\rho_n)$ we have proved the lemma. □

Secondly we present an obvious result concerning the finiteness of certain integrals.

LEMMA 3.2. a. *Let* ℓ *be a number* >1 *and let, for some* $\delta > 0$, $\beta_{\ell+\delta} < \infty$. *Then there exists* $A > 0$, *depending only on* ℓ *and* δ, *such that*

$$(3.4) \qquad \int_0^1 (s(1-s))^{\frac{1}{\ell}}\, dF^{-1}(s) \le A\, \beta_{\ell+\delta}^{\frac{1}{\ell+\delta}} < \infty.$$

b. *If* $\ell = 1$ *and* $\delta = 0$ *then* (3.4) *holds with* $A = 1$.

PROOF. Applying integration by parts we obtain

$$(3.5) \qquad \int_0^1 (s(1-s))^{\frac{1}{\ell}}\, dF^{-1}(s) = (s(1-s))^{\frac{1}{\ell}} F^{-1}(s)\Big|_0^1 -$$

$$- \ell^{-1} \int_0^1 F^{-1}(s)(s(1-s))^{\frac{1}{\ell}-1}(1-2s)ds.$$

Both under the assumptions a. and b. the first term on the right of (3.5) is easily seen to be zero. To conclude the proof of part a. we apply Hölder's inequality to the second term on the right of (3.5):

$$|\ell^{-1}\int_0^1 F^{-1}(s)(s(1-s))^{\frac{1}{\ell}-1}ds| \le \int_0^1 |F^{-1}(s)|(s(1-s))^{\frac{1}{\ell}-1}ds \le$$

$$\le \beta_{\ell+\delta}^{\frac{1}{\ell+\delta}} \cdot (\int_0^1 (s(1-s))^{-1+\frac{\delta}{\ell(\ell+\delta-1)}}ds)^{\frac{\ell+\delta-1}{\ell+\delta}} < \infty .$$

The proof of part b. is immediate from (3.5) and the remark made after it. This completes the proof of the lemma. □

The third lemma of this section will enable us to estimate certain moments.

LEMMA 3.3. *Let* ℓ *be a positive integer and let, for some* $\delta > 0$, $\beta_{\ell+\delta} < \infty$. *Then for any number* p *for which* $p\,\ell \ge 2$, *there exists* $A > 0$ *depending only on* p, ℓ *and* δ, *such that*

$$E(\int_0^1 |\Gamma_n(s)-s|^p dF^{-1}(s))^{\ell} \le A\, \beta_{\ell+\delta}^{\frac{\ell}{\ell+\delta}}\, n^{-\frac{p\ell}{2}} . \tag{3.6}$$

PROOF. By Fubini's theorem we have

$$E(\int_0^1 |\Gamma_n(s)-s|^p dF^{-1}(s))^{\ell} =$$

$$= \int_0^1 \cdots \int_0^1 E \prod_{i=1}^{\ell} |\Gamma_n(s_i) - s_i|^p \, dF^{-1}(s_1)\ldots dF^{-1}(s_\ell).$$

An application of Hölder's inequality shows that

$$E \prod_{i=1}^{\ell} |\Gamma_n(s_i) - s_i|^p \le \prod_{i=1}^{\ell} (E|\Gamma_n(s_i) - s_i|^{p\ell})^{\frac{1}{\ell}}$$

for all $0 < s_1,\ldots,s_\ell < 1$. Hence we know that

$$E(\int_0^1 |\Gamma_n(s)-s|^p dF^{-1}(s))^{\ell} \le (\int_0^1 (E|\Gamma_n(s)-s|^{p\ell})^{\frac{1}{\ell}} dF^{-1}(s))^{\ell}.$$

Since $\Gamma_n(s) = n^{-1} \Sigma_{i=1}^{n} \chi_{(0,s]}(U_i)$ for all $0 < s < 1$ and $n \geq 1$ the MARCINKIEVITZ, ZYGMUND, CHUNG inequality (see CHUNG (1951)) yields for $p\ell \geq 2$, $n \geq 1$ and $0 < s < 1$

$$E|\Gamma_n(s)-s|^{p\ell} \leq B\, n^{-\frac{p\ell}{2}}\, s(1-s)$$

where $B > 0$ depends only on p and ℓ. It follows that

$$E(\int_0^1 |\Gamma_n(s)-s|^p dF^{-1}(s))^{\ell} \leq B\, n^{-\frac{p\ell}{2}} (\int_0^1 (s(1-s))^{\frac{1}{\ell}}\, dF^{-1}(s))^{\ell}.$$

An application of lemma 3.2 completes our proof. □

To formulate the following lemma we introduce functions H_1, H_2 and H_3 by

$$(3.7) \qquad H_1(u) = \int_0^1 |J(s)|\cdot|\chi_{(0,s]}(u)-s|\,dF^{-1}(s)$$

$$(3.8) \qquad H_2(u) = \int_0^1 |J'(s)|\cdot|\chi_{(0,s]}(u)-s|\,dF^{-1}(s)$$

$$(3.9) \qquad H_3(u) = \int_0^1 |J''(s)|\cdot|\chi_{(0,s]}(u)-s|\,dF^{-1}(s)$$

for $0 < u < 1$. Note that the integrand of H_i majorize the integrand of h_i and hence that H_i majorize h_i $(1 \leq i \leq 3)$. Remark also that h_2 and h_3 are symmetric in their respectively arguments.

LEMMA 3.4.

(a) *Take* $\ell \geq 1$, *suppose that* J *has a bounded second derivative on* (0,1) *and* $\beta_\ell = E|X_1|^{\ell} < \infty$. *Then*

$$(3.10) \qquad EH_1^{\ell}(U_1) < \infty$$

$$(3.11) \qquad EH_2^{\ell}(U_1) < \infty$$

(3.12) $$EH_3^{\ell}(U_1) < \infty.$$

(b) *Let* J *be twice differentiable with bounded second derivative on* (0,1) *and let* $\beta_1 = E|X_1| < \infty$. *Then* $Eh_i(U_i) = 0$ *for any* i, *and with probability one* $E(h_2(U_i,U_j)|U_j) = 0$ *for* $i \neq j$ *and* $E(h_3(U_i,U_j,U_k)|U_j,U_k) = 0$ *if* $i \neq j$ *and* $i \neq k$.

PROOF. We first prove (3.10). It is immediate from (3.7) that

$$H_1(U_1) \leq \|J\| \cdot \left(\int_{(0,U_1)} s dF^{-1}(s) + \int_{[U_1,1)} (1-s) dF^{-1}(s) \right).$$

Applying the c_r-inequality we find

$$E\, H_1^{\ell}(U_1) \leq 2^{\ell-1} \cdot \|J\|^{\ell} \cdot \Big[E\Big(\int_{(0,U_1)} s dF^{-1}(s) \Big)^{\ell} + $$

$$+ E\Big(\int_{[U_1,1)} (1-s) dF^{-1}(s) \Big)^{\ell} \Big].$$

Using integration by parts, the finiteness of β_ℓ, and applying the c_r-inequality once more we see that

$$E\Big(\int_{(0,U_1)} s dF^{-1}(s) \Big)^{\ell} = E\Big|U_1 F^{-1}(U_1) - \int_0^{U_1} F^{-1}(s) ds\Big|^{\ell} \leq$$

$$\leq 2^{\ell-1} \cdot \Big(E(|F^{-1}(U_1)|^{\ell} + \Big(\int_0^1 |F^{-1}(s)| ds \Big)^{\ell} \Big) \leq 2^{\ell-1} (E|X_1|^{\ell} + (E|X_1|)^{\ell})$$

$$\leq 2^{\ell} E|X_1|^{\ell}.$$

Similarly we can show that

$$E\Big(\int_{[U_1,1)} (1-s) dF^{-1}(s) \Big)^{\ell} \leq 2^{\ell} E|X_1|^{\ell}$$

so that $E\, H_1^{\ell}(U_1) < \infty$ which proves (3.10). The other statements of part a. can be proved in essentially the same way.

b. We shall prove that with probability one $E(h_3(U_i,U_j,U_k)|U_j,U_k) = 0$ for $i \neq j$ and $i \neq k$. Note first that using Fubini's theorem for non-negative integrands and applying (3.12) with $\ell = 1$ we see that with probability one

$$E\Big(\int_0^1 |J''(s)|\,|\chi_{(0,s]}(U_i)-s|\,|\chi_{(0,s]}(U_j)-s|\,|\chi_{(0,s]}(U_k)-s|$$

$$dF^{-1}(s)\,|U_j,U_k\Big) \leq E\, H_3(U_1) < \infty,$$

Therefore the conditional expectation $E(h_3(U_i,U_j,U_k)|U_j,U_k)$ is well-defined and Fubini's theorem can be applied to see that $E(h_3(U_i,U_j,U_k)|U_j,U_k) = 0$ with probability one. The other two statements of part b are easier and can be proved in essentially the same way. □

<u>REMARK</u>. Lemma 3.4 will be applied frequently in the following sections. In particular the proof of lemma 4.6 depends heavily on it. In this remark we give two typical examples of the way we shall use lemma 3.4.

(i) Suppose that $\beta_2 < \infty$ and J' is bounded on (0,1). Then $Eh_1(U_1)h_2(U_1,U_2) = Eh_1^2(U_1)h_2(U_1,U_2) = 0$.

(ii) Suppose that J has a bounded second derivative on (0,1) and $\beta_4 < \infty$. Then $E(h_1(U_1)+h_1(U_2))^4|h_2(U_1,U_2)| < \infty$.

<u>PROOF</u>.

(i) We first prove that $Eh_1^2(U_1)h_2(U_1,U_2) = 0$. It follows directly from (2.1), (2.2), (3.10) and (3.11) and the independence of U_1 and U_2 that $|Eh_1^2(U_1)h_2(U_1,U_2)| \leq E\, H_1^2(U_1)\, E\, H_2(U_2) < \infty$. Hence we can write

$$E\, h_1^2(U_1)h_2(U_1,U_2) = E[E(h_1^2(U_1)h_2(U_1,U_2)|U_1)] =$$

$$= E[h_1^2(U_1)E\{h_2(U_1,U_2)|U_1\}] = 0,$$

because of lemma 3.4.b. This proves the assertion. The other statement can

can be proved in essentially the same way.

(ii) We remark that

$$E(h_1(U_1) + h_1(U_2))^4 |h_2(U_1,U_2)| \le$$

$$\le 8\,E\,h_1^4(U_1)|h_2(U_1,U_2)| + 8\,E\,h_1^4(U_2)|h_2(U_1,U_2)| =$$

$$= 16\,E\,h_1^4(U_1)|h_2(U_1,U_2)| \le 16\,E\,H_1^4(U_1)H_2(U_2) =$$

$$= 16\,E\,H_1^4(U_1)\,E\,H_2(U_2) < \infty,$$

using lemma 3.4 and the independence of U_1 and U_2. This completes the proof. □

The fifth and last lemma of this section gives conditions which guarantee that $\sigma^2 = \sigma^2(J,F)$ (c.f. (2.7)) is bounded away from zero.

LEMMA 3.5. *Let* J *be bounded on* (0,1) *and* $\beta_1 < \infty$. *Suppose that a positive number* c *and numbers* $0 \le t_1 < t_2 \le 1$ *exist such that on* $(F^{-1}(t_1), F^{-1}(t_2))$ F *possesses a bounded density and on* (t_1,t_2) *assumption* (2.10) *is satisfied. Then* $\sigma^2(J,F) > 0$.

PROOF. Note first that because J is bounded on (0,1) and $\beta_1 < \infty$ the function h_1 is well-defined and finite for every $0 < u < 1$. Secondly we remark that $\sigma^2(J,F) = \int_0^1 h_1^2(u)du \ge \int_{t_1}^{t_2} h_1^2(u)du$. It follows directly from (2.1) and the assumptions of the lemma that $h_1(u_2) - h_1(u_1) \ge c\,M_1^{-1}(u_2-u_1)$ for $t_1 < u_1 < u_2 < t_2$ and some constant $M_1 > 0$. The geometry of the situation ensures now that $\int_{t_1}^{t_2} h_1^2(u)du$ is minimized for $h_1(u) = (u - \frac{t_1}{2} - \frac{t_2}{2})\frac{c}{M_1}$. Hence

$$\sigma^2(J,F) \ge \frac{c^2(t_2-t_1)^3}{12\,M_1^2}\,.$$

This completes the proof of the lemma. □

4. PROOF OF THEOREM 2.1.

The purpose of this section is to provide a proof of theorem 2.1. Since our proofs will depend on characteristic function (c.f.) arguments let us denote by $\rho_n^*(t)$ the c.f. of T_n^* and by $\tilde{\rho}_n(t)$ the Fourier-Stieltjes transform

$$\tilde{\rho}_n(t) = \int_{-\infty}^{\infty} \exp(itx)dK_n(x)$$

of K_n (see (2.4)).
We shall show that for some sufficiently small $\varepsilon > 0$

$$(4.1) \qquad \int_{|t|\le n^{\varepsilon}} |\rho_n^*(t) - \tilde{\rho}_n(t)|\,|t|^{-1}dt = o(n^{-1})$$

and that

$$(4.2) \qquad \int_{n^{\varepsilon}<|t|\le n^{3/2}} |\rho_n^*(t)|\,|t|^{-1}dt = o(n^{-1})$$

and

$$(4.3) \qquad \int_{|t|>\log(n+1)} |\tilde{\rho}_n(t)|\,|t|^{-1}dt = o(n^{-1}),$$

hold as $n \to \infty$. An application of Esseen's smoothing lemma (ESSEEN (1945)) will then complete our proof.

We first prove (4.1). We shall essentially have to expand $\rho_n^*(t)$ for these "small" values of $|t|$. To start with we define for $0 < u < 1$

$$(4.4) \qquad \psi_1(u) = \int_u^1 J(s)ds - (1-u)\bar{J}_1$$

and

$$\psi_2(u) = \int_u^1 (\tfrac{1}{2} - s)J'(s)ds - (1-u)\bar{J}_2$$

where $\bar{J}_1 = \int_0^1 J(s)ds$ and $\bar{J}_2 = \int_0^1 (\frac{1}{2} - s)J'(s)ds$. Then it is easy to check (see SHORACK (1972) for a similar approach) that with probability one

$$(4.5) \qquad T_n = \int_0^1 (\psi_1(\Gamma_n(s)) + n^{-1}\psi_2(\Gamma_n(s)))dF^{-1}(s) +$$

$$+ (\bar{J}_1 + n^{-1}\bar{J}_2)n^{-1} \sum_{i=1}^n F^{-1}(U_i) +$$

$$+ n^{-1} \sum_{i=1}^n (c_{in} - n \int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)ds - \int_{\frac{i-1}{n}}^{\frac{i}{n}} (\frac{1}{2} - s)J'(s)ds)F^{-1}(U_{in}).$$

Let J be twice differentiable with first and second derivative J' and J" on (0,1). Let J" be bounded on (0,1) and let $\beta_1 = E|X_1| < \infty$. Introduce for each $n \geq 1$ the r.v. S_n by (a prime denoting differentiation)

$$(4.6) \qquad S_n = \int_0^1 \Big\{\psi_1(s) + n^{-1}\psi_2(s) + (\Gamma_n(s)-s)(\psi_1'(s) + n^{-1}\psi_2'(s)) +$$

$$+ \frac{(\Gamma_n(s)-s)^2}{2} \psi_1''(s) + \frac{(\Gamma_n(s)-s)^3}{6} \psi_1'''(s)\Big\}dF^{-1}(s) +$$

$$+ (\bar{J}_1 + n^{-1}\bar{J}_2)n^{-1} \sum_{i=1}^n F^{-1}(U_i).$$

Note that $|\psi_1(u)| \leq 4\|J_1\|u(1-u)$ and $|\psi_2(u)| \leq 4\|J'\|u(1-u)$ for $0 < u < 1$, and that $\psi_1'(s) = -J(s) + \bar{J}_1$, $\psi_2'(s) = (s - \frac{1}{2})J'(s) + \bar{J}_2$, $\psi_1''(s) = -J'(s)$ and $\psi_1'''(s) = -J''(s)$ on (0,1) so that it easily verified that S_n is a well-defined r.v. Later on in this section it will become clear that $T_n^* - S_n^*$ is, under appropriate conditions, of negligible order for our purposes.

It is convenient to introduce some more notation. Define r.v.'s I_{mn} for m = 1,2,3,4 and $n \geq 1$ by

$$(4.7) \qquad I_{1n} = - \int_0^1 J(s)(\Gamma_n(s)-s)dF^{-1}(s) = n^{-1} \sum_{i=1}^n h_1(U_i)$$

$$(4.8)\qquad I_{2n} = -\int_0^1 J'(s)\,\frac{(\Gamma_n(s)-s)^2}{2}\,dF^{-1}(s) = 2^{-1}n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} h_2(U_i,U_j)$$

$$(4.9)\qquad I_{3n} = -\int_0^1 J''(s)\,\frac{(\Gamma_n(s)-s)^3}{6}\,dF^{-1}(s) = 6^{-1}n^{-3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} h_3(U_i,U_j,U_k)$$

$$(4.10)\qquad I_{4n} = -n^{-1}\int_0^1 (\tfrac{1}{2}-s)J'(s)(\Gamma_n(s)-s)\,dF^{-1}(s) = n^{-2}\sum_{i=1}^{n} h_4(U_i)$$

where the functions h_1, h_2, h_3 and h_4 are given by (2.1) - (2.3) and (2.16). It is easily checked that

$$(4.11)\qquad \hat{S}_n = S_n - ES_n = \sum_{m=1}^{4}\hat{I}_{mn} = \sum_{m=1}^{4}(I_{mn} - EI_{mn}).$$

Furthermore define r.v.'s J_{mn} for $m = 1,2,3,4$, and $n \geq 1$ by

$$(4.12)\qquad J_{mn} = \hat{I}_{mn}/\sigma(S_n) = (I_{mn} - EI_{mn})/\sigma(S_n),$$

so that

$$(4.13)\qquad S_n^* = \sum_{m=1}^{4} J_{mn}.$$

The proof of (4.1) will now be split up in a number of lemma's. In the first lemma of this section we give an asymptotic expansion for the variance of S_n.

<u>LEMMA 4.1</u>. *Suppose* J *has a bounded second derivative on* (0,1) *and that* $\beta_2 < \infty$. *Then we have that*

$$(4.14)\qquad |\sigma^2(S_n) - n^{-1}\sigma^2 - 2n^{-2}\sigma^2 b| = O(n^{-\frac{5}{2}}), \qquad \text{as } n \to \infty.$$

where $\sigma^2 = \sigma^2(J,F)$ *is as in* (2.7) *and* $b = b(J,F)$ *as in* (2.15).

PROOF. In view of (4.11) $\sigma^2(S_n) = \sigma^2(\Sigma_{m=1}^4 I_{mn})$. It follows directly from (2.1) and (4.7) that $\sigma^2(I_{1n}) = n^{-1}\sigma^2$. Also note that it is immediate from (4.7), (4.8) and an application of lemma 3.4 that

$$2\,\mathrm{cov}(I_{1n},I_{2n}) = 2EI_{1n}I_{2n} = n^{-3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} Eh_1(U_i)h_2(U_j,U_k) =$$

$$= n^{-2}\int_0^1 h_1(u)h_2(u,u)du.$$

Next we consider $\sigma^2(I_{2n})$. Using lemma 3.2 and lemma 3.4 once more we directly find that

$$E\,I_{2n}^2 = 4^{-1}\,n^{-2}\,(E\,h_2(U_1,U_1))^2 + 2^{-1}\,n^{-2}\,E\,h_2^2(U_1,U_2) +$$

$$+ O(n^{-3}),\text{ as } n \to \infty.$$

Because we know also that $(E\,I_{2n})^2 = 4^{-1}\,n^{-2}(E\,h_2(U_1,U_1))^2$ we have shown that

$$\sigma^2(I_{2n}) = 2^{-1}\,n^{-2}\int_0^1\int_0^1 h_2^2(u,v)\,dudv + O(n^{-3}),\text{ as } n \to \infty.$$

Similarly we can prove that

$$2\,\mathrm{cov}(I_{1n},I_{3n}) = n^{-2}\int_0^1\int_0^1 h_1(u)\,h_3(u,v,v)\,dudv + O(n^{-3}),\text{ as } n \to \infty$$

and also that

$$2\,\mathrm{cov}(I_{1n},I_{4n}) = 2n^{-2}\int_0^1 h_1(u)\,h_4(u)\,du.$$

Finally we remark that it is easy to prove by using similar arguments as above that

$$\sigma^2(I_{3n}) + \sigma^2(I_{4n}) = \mathcal{O}(n^{-3}), \qquad \text{as } n \to \infty$$

and also that

$$|\mathrm{cov}(I_{2n},I_{3n}) + \mathrm{cov}(I_{2n},I_{4n}) + \mathrm{cov}(I_{3n},I_{4n})| = \mathcal{O}(n^{-\frac{5}{2}}), \text{ as } n \to \infty.$$

Combining all the results we have proved (4.14). □

LEMMA 4.2. *Suppose that* J *has a bounded second derivative on* (0,1) *and* $\beta_2 < \infty$. *Then* $\sigma^2(J,F) > 0$ *implies that for any fixed real number* m

(4.15) $$|\sigma^{-m}(S_n) - n^{\frac{m}{2}} \sigma^{-m}| = \mathcal{O}(n^{\frac{m}{2}-1}), \qquad \textit{as } n \to \infty.$$

where $\sigma^2 = \sigma^2(J,F)$ *is as in* (2.7).

PROOF. The statement is immediate from lemma 4.1. □

The next lemma will enable us to show that $T_n^* - S_n^*$ is of negligible order for our purposes. Let τ_n^* denote the c.f. of S_n^*.

LEMMA 4.3. *Suppose that* J *has a bounded third derivative on* (0,1), *assumption* (2.9) *is satisfied and* $\beta_{2+\delta} < \infty$ *for some* $\delta > 0$. *Then* $\sigma^2(J,F) > 0$ *implies that for every* $\varepsilon > 0$.

(4.16) $$\int_{|t|\le n^{\varepsilon}} |\rho_n^*(t) - \tau_n^*(t)|\,|t|^{-1}dt = \mathcal{O}(n^{-\frac{3}{2}+\varepsilon}), \qquad \text{as } n \to \infty.$$

PROOF. It follows from lemma X.V.4.1 of FELLER (1966) that

(4.17) $$|\rho_n^*(t) - \tau_n^*(t)| \le |t| E|T_n^* - S_n^*|$$

for all t and $n \ge 1$. Using (4.5), (4.6), the boundedness of J''' on (0,1) and applying Taylor's theorem we see directly that

$$(4.18)\qquad \sigma^2(T_n - S_n) \le 3\|J'''\|^2 E\Big(\int_0^1 (\Gamma_n(s)-s)^4 dF^{-1}(s)\Big)^2 +$$

$$+\ 3\|J''\|^2 n^{-2} E\Big(\int_0^1 (\Gamma_n(s)-s)^2 dF^{-1}(s)\Big)^2 +$$

$$+\ 3\sigma^2\Big(n^{-1}\sum_{i=1}^n \Big(J\big(\tfrac{i}{n+1}\big) - n\int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)ds - \int_{\frac{i-1}{m}}^{\frac{i}{n}} (\tfrac{1}{2}-s)J'(s)ds\Big)F^{-1}(U_{in})\Big),$$

Application of lemma 3.3 with $\ell = 2$ and $p = 4$ and $p = 2$ respectively implies that the sum of the first two terms on the right of (4.18) is $O(n^{-4})$, as $n \to \infty$. To treat the third term on the right of (4.18) we need the following simple inequality: $\sigma^2(\Sigma_{i=1}^n a_i X_{in}) \le \sigma^2(\Sigma_{i=1}^n b_i X_{in})$, provided $a_i a_j \le b_i b_j$ for all $1 \le i, j \le n$. (The proof of this inequality is immediate from the well-known fact that the covariance of any two order statistics is always non-negative.) Using this and the fact that it is easily verified that

$$\max_{1\le i\le n} \Big|J\big(\tfrac{i}{n+1}\big) - n\int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)ds - \int_{\frac{i-1}{n}}^{\frac{i}{n}} (\tfrac{1}{2}-s)J'(s)ds\Big| = O(n^{-2}),\quad \text{as } n \to \infty$$

under the assumptions of the lemma, we find that

$$(4.19)\qquad \sigma^2\Big(n^{-1}\sum_{i=1}^n \Big(J\big(\tfrac{i}{n+1}\big) - n\int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)ds - \int_{\frac{i-1}{n}}^{\frac{i}{n}} (\tfrac{1}{2}-s)J'(s)ds\Big)F^{-1}(U_{in})\Big) =$$

$$= O(n^{-5}),\qquad \text{as } n \to \infty.$$

Combining these results it is easy to conclude that

$$(4.20)\qquad \sigma^2(T_n - S_n) = O(n^{-4})\qquad \text{as } n \to \infty.$$

To complete our proof we remark that it follows from an application of the lemma's 3.1 and 4.2 (with m = -2) that $\sigma^2(T_n^* - S_n^*) = O(n^{-3})$ as $n \to \infty$. This combined with (4.17) proves (4.16). □

Next define for real t and $n \geq 1$

$$\tau_{1n}(t) = Ee^{itJ_{1n}}(1 + it(J_{2n} + J_{3n} + J_{4n}) + \frac{(it)^2}{2} J_{2n}^2). \tag{4.21}$$

In the following lemma we shall approximate τ_n^* by τ_{1n} for all $|t| \leq n^{\varepsilon}$.

LEMMA 4.4. *Suppose that* J *has a bounded second derivative on* (0,1), *and* $\beta_{3+\delta} < \infty$ *for some* $\delta > 0$. *Then* $\sigma^2(J,F) > 0$ *implies that for every* $\varepsilon > 0$

$$\int_{|t| \leq n^{\varepsilon}} |\tau_n^*(t) - \tau_{1n}(t)| \, |t|^{-1} dt = O(n^{-\frac{3}{2} + 3\varepsilon}), \qquad \text{as } n \to \infty. \tag{4.22}$$

PROOF. Application of lemma X.V.4.1. of FELLER (1966) yields

$$|\tau_n^*(t) - \tau_{1n}(t)| = |Ee^{itJ_{1n}}(e^{it(J_{2n}+J_{3n}+J_{4n})} - 1 - it(J_{2n}+J_{3n}+J_{4n}) -$$

$$- \frac{(it)^2}{2} J_{2n}^2)| \leq t^2 \, (E|J_{2n} \, J_{3n}| + E|J_{2n} \, J_{4n}| + E|J_{3n} \, J_{4n}| + E \, J_{3n}^2 +$$

$$+ E \, J_{4n}^2) + |t|^3 \, E|J_{2n} + J_{3n} + J_{4n}|^3,$$

for all t and $n \geq 1$. It is not difficult to verify from the proof of lemma 4.1 and from lemma 4.2 that the coefficient of t^2 in the above inequality is $O(n^{-3/2})$, as $n \to \infty$. An application of the c_r-inequality, lemma 3.3 with $\ell = 3$ and p = 2,3 and 4 respectively, and of lemma 4.2 shows that also $E|J_{2n} + J_{3n} + J_{4n}|^3 = O(n^{-3/2})$, as $n \to \infty$. Combining these results we easily check that (4.22) is proved. □

We continue with the analysis of $\tau_{1n}(t)$. For convenience we write σ_n^2 to indicate $n\sigma^2(S_n)$ and we denote the c.f. of $h_1(U_1)$ by ρ. To start with we remark that it follows from (4.21) that

(4.23)
$$\tau_{1n}(t) = \rho^n\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) +$$

$$+ \frac{it}{2n^{\frac{3}{2}}\sigma_n}\,\rho^{n-2}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) n(n-1)\; E\, e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2))} \cdot h_2(U_1,U_2) +$$

$$+ \frac{it}{2n^{\frac{3}{2}}\sigma_n}\,\rho^{n-1}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) n\; E\, e^{\frac{it}{n^{\frac{1}{2}}\sigma_n} h_1(U_1)}\, \hat{h}_2(U_1,U_1) +$$

$$+ \frac{it}{6n^{\frac{5}{2}}\sigma_n}\,\rho^{n-3}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) n(n-1)(n-2)\; E\, e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2)+h_1(U_3))} \cdot$$

$$\cdot h_3(U_1,U_2,U_3) +$$

$$+ \frac{it}{6n^{\frac{5}{2}}\sigma_n}\,\rho^{n-2}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) 3n(n-1)\; E\, e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2))}\, h_3(U_1,U_1,U_2) +$$

$$+ \frac{it}{6n^{\frac{5}{2}}\sigma_n}\,\rho^{n-1}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) n\; E\, e^{\frac{it}{n^{\frac{1}{2}}\sigma_n} h_1(U_1)}\, \hat{h}_3(U_1,U_1,U_1) +$$

$$+ \frac{it}{n^{\frac{3}{2}}\sigma_n}\,\rho^{n-1}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) n\; E\, e^{\frac{it}{n^{\frac{1}{2}}\sigma_n} h_1(U_1)}\, h_4(U_1) +$$

$$+ \frac{(it)^2}{8n^3\sigma_n^2}\,\rho^{n-4}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) n(n-1)(n-2)(n-3)\cdot$$

$$\cdot \left(E e^{\frac{it}{n^{\frac12}\sigma_n}(h_1(U_1)+h_1(U_2))} h_2(U_1,U_2)\right)^2 +$$

$$+ \frac{(it)^2}{8n^3\sigma_n^2}\,\rho^{n-3}\left(\frac{t}{n^{\frac12}\sigma_n}\right)4n(n-1)(n-2)\ \cdot$$

$$\cdot\ E e^{\frac{it}{n^{\frac12}\sigma_n}(h_1(U_1)+h_1(U_2)+h_1(U_3))} h_2(U_1,U_2)h_2(U_1,U_3) +$$

$$+ \frac{(it)^2}{8n^3\sigma_n^2}\,\rho^{n-3}\left(\frac{t}{n^{\frac12}\sigma_n}\right)2n(n-1)(n-2)\ \cdot$$

$$\cdot\ E e^{\frac{it}{n^{\frac12}\sigma_n}(h_1(U_1)+h_1(U_2)+h_1(U_3))} \hat{h}_2(U_1,U_1)h_2(U_2,U_3) +$$

$$+ \frac{(it)^2}{8n^3\sigma_n^2}\,\rho^{n-2}\left(\frac{t}{n^{\frac12}\sigma_n}\right)4n(n-1)\ \cdot$$

$$\cdot\ E e^{\frac{it}{n^{\frac12}\sigma_n}(h_1(U_1)+h_1(U_2))} \hat{h}_2(U_1,U_1)\cdot h_2(U_1,U_2) +$$

$$+ \frac{(it)^2}{8n^3\sigma_n^2}\,\rho^{n-2}\left(\frac{t}{n^{\frac12}\sigma_n}\right)2n(n-1)\ E e^{\frac{it}{n^{\frac12}\sigma_n}(h_1(U_1)+h_1(U_2))} (h_2(U_1,U_2))^2 +$$

$$+ \frac{(it)^2}{8n^3\sigma_n^2}\,\rho^{n-2}\left(\frac{t}{n^{\frac12}\sigma_n}\right)n(n-1)\left(E e^{\frac{it}{n^{\frac12}\sigma_n}h_1(U_1)} \hat{h}_2(U_1,U_1)\right)^2 +$$

$$+ \frac{(it)^2}{8n^3\sigma_n^2} \rho^{n-1}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) n \, \mathcal{E} e^{\frac{it}{n^{\frac{1}{2}}\sigma_n} h_1(U_1)} (\hat{h}_2(U_1,U_1))^2 .$$

In the next lemma we derive an asymptotic expansion for the factors $\rho^{n-m}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right)$ appearing in the terms on the right of (4.23).

LEMMA 4.5. *Suppose that* J *has a bounded second derivative on* (0,1) *and* $\beta_4 < \infty$. *Then* $\sigma^2(J,F) > 0$ *implies that there exists* $a > 0$, *a sequence of positive numbers* $\delta_1, \delta_2, \ldots$ *with* δ_n *depending only on* n *and with* $\lim_{n\to\infty}\delta_n = 0$, *and a fixed polynomial* P *in* t, *such that for any fixed* $m \geq 0$ *and all* $|t| \leq an^{\frac{1}{2}}$ *and* $n \geq 1$.

$$(4.24) \qquad \left|\rho^{n-m}\left(\frac{t}{n^{\frac{1}{2}}\sigma_n}\right) - e^{-\frac{t^2}{2}}\left(1 - \frac{(it)^2}{n}\left(\frac{m}{2}+b\right) + \frac{(it)^3 \int_0^1 h_1^3(u)du}{6n^{\frac{1}{2}}\sigma^3} + \right.\right.$$

$$\left.\left. + \frac{(it)^4(\int_0^1 h_1^4(u)du - 3\sigma^4)}{24\, n\sigma^4} + \frac{(it)^6(\int_0^1 h_1^3(u)du)^3}{72\, n\sigma^6}\right)\right| \leq$$

$$= \mathcal{O}(\delta_n n^{-1}|t|P(t)e^{-\frac{t^2}{4}}),$$

where $\sigma^2 = \sigma^2(J,F)$ *is as in* (2.7) *and* $b = b(J,F)$ *as in* (2.18).

PROOF. Since $(\sigma(n-m))^{-\frac{1}{2}} \Sigma_{i=1}^{n-m} h_1(U_i)$ is a properly standardized sum of independent identically distributed r.v.'s with expectation zero, variance one, and finite fourth moment, it follows directly from the classical theory of Edgeworth expansions for such sums (see e.g. GNEDENKO-KOLMOGOROV (1954), §41, theorem 2.1, inequality (b)) that there exist $a > 0$ and a sequence of positive numbers $\delta_1, \delta_2, \ldots$ satisfying the assumptions of the lemma such that for all $|t| \leq an^{\frac{1}{2}}$ and $n \geq 1$

(4.25) $$\left|\rho^{n-m}\left(\frac{t}{(n-m)^{\frac{1}{2}}\sigma}\right) - e^{-\frac{t^2}{2}}\left(1 + \frac{(it)^3\int_0^1 h_1^3(u)du}{6n^{\frac{1}{2}}\sigma^3} + \frac{(it)^4\left(\int_0^1 h_1^4(u)du - 3\sigma^4\right)}{24n\sigma^4} + \frac{(it)^6\left(\int_0^1 h_1^3(u)du\right)^2}{72n\sigma^6}\right)\right| =$$

$$= O(\delta_n\, n^{-1}|t|\, P(t)\, e^{-\frac{t^2}{4}}), \text{ as } n \to \infty,$$

where P is a fixed polynomial in t. We perform now a change of variables $t_n = tn^{\frac{1}{2}}\sigma_n/(n-m)^{\frac{1}{2}}\sigma)$. It follows after expanding $e^{-t_n^2/2}$ around t and using the result of lemma 4.1 that we obtain (4.24). □

The expectations appearing on the right of (4.23) are expanded in the following lemma.

LEMMA 4.6. *Suppose that* J *has a bounded second derivative on* (0,1) *and* $\beta_4 < \infty$. *Then* $\sigma^2(J,F) > 0$ *implies that for all* t

(4.26) $$\left|Ee^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2))} h_2(U_1,U_2) - \frac{(it)^2}{n\sigma^2}\int_0^1\int_0^1 h_1(u)h_1(v)h_2(u,v)dudv - \frac{(it)^3}{n^{\frac{3}{2}}\sigma^3}\int_0^1\int_0^1 h_1^2(u)h_1(v)h_2(u,v)dudv\right| = O(n^{-2}(t^2+t^4) + n^{-\frac{5}{2}}|t|^3).$$

(4.27) $$\left|Ee^{\frac{it}{n^{\frac{1}{2}}\sigma_n}h_1(U_1)}\hat{h}_2(U_1,U_1) - \frac{it}{n^{\frac{1}{2}}\sigma}\int_0^1 h_1(u)h_2(u,u)du\right| = O(n^{-1}t^2+n^{-\frac{3}{2}}|t|).$$

(4.28) $$\left| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2)+h_1(U_3))} h_3(U_1,U_2,U_3) - \right.$$

$$\left. - \frac{(it)^3}{n^{\frac{3}{2}}\sigma^3} \int_0^1\int_0^1\int_0^1 h_1(u)h_1(v)h_1(w)h_3(u,v,w)\,dudvdw \right| = O(n^{-2}t^4+n^{-\frac{5}{2}}|t|^3).$$

(4.29) $$\left| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2))} h_3(U_1,U_1,U_2) - \right.$$

$$\left. - \frac{it}{n^{\frac{1}{2}}\sigma} \int_0^1\int_0^1 h_1(u)h_3(u,v,v)\,dudv \right| = O(n^{-1}t^2+n^{-\frac{3}{2}}|t|).$$

(4.30) $$\left| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n} h_1(U_1)} \hat{h}_3(U_1,U_1,U_1) \right| = O(n^{-\frac{1}{2}}|t|).$$

(4.31) $$\left| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n} h_1(U_1)} h_4(U_1) - \frac{it}{n^{\frac{1}{2}}\sigma} \int_0^1 h_1(u)h_4(u)\,du \right| = O(n^{-1}t^2+n^{-\frac{3}{2}}|t|).$$

(4.32) $$\left| \left(E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2))} h_2(U_1,U_2) \right)^2 - \right.$$

$$\left. - \frac{(it)^4}{n^2\sigma^4} \left(\int_0^1\int_0^1 h_1(u)h_1(v)h_2(u,v)\,dudv \right)^2 \right| = O(n^{-\frac{5}{2}}|t|^5+n^{-3}t^4).$$

(4.33) $$\left| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2)+h_1(U_3))} h_2(U_1,U_2)h_2(U_1,U_3) - \right.$$

$$- \frac{(it)^2}{n\sigma^2} \int_0^1 \int_0^1 \int_0^1 h_1(u)h_1(v)h_2(u,w)h_2(v,w)\,dudvdw \Big| = O(n^{-\frac{3}{2}}|t|^3 + n^{-2}t^2).$$

(4.34) $$\Big| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2)+h_1(U_3))} \hat{h}_2(U_1,U_1)h_2(U_2,U_3) \Big| = O(n^{-\frac{3}{2}}|t|^3).$$

(4.35) $$\Big| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2))} \hat{h}_2(U_1,U_1)h_2(U_1,U_2) \Big| = O(n^{-\frac{1}{2}}|t|).$$

(4.36) $$\Big| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2))} (h_2(U_1,U_2))^2 - \int_0^1 \int_0^1 h_2^{\,2}(u,v)\,dudv \Big| = O(n^{-\frac{1}{2}}|t|).$$

(4.37) $$\Big| (E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n} h_1(U_1)} \hat{h}_2(U_1,U_1))^2 \Big| = O(n^{-1}t^2).$$

(4.38) $$\Big| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n} h_1(U_1)} (\hat{h}_2(U_1,U_1))^2 \Big| = O(1), \qquad \text{as } n \to \infty$$

PROOF. Because the statements (4.26) - (4.38) are all proved in essentially the same manner we shall only prove, by way of an example, (4.26). Expanding $\exp\left(\frac{it}{h^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2))\right)$ around $t = 0$ and applying part (i) of the remark made after lemma 3.4 we find that for all t and $n \geq 1$

(4.39) $$\Big| E e^{\frac{it}{n^{\frac{1}{2}}\sigma_n}(h_1(U_1)+h_1(U_2))} h_2(U_1,U_2) - \frac{(it)^2}{n\sigma_n^2} \int_0^1 \int_0^1 h_1(u)h_1(v)h_2(u,v)\,dudv \; -$$

$$-\frac{(it)^3}{n^{\frac{3}{2}}\sigma_n^3}\int_0^1\int_0^1 h_1^2(u)h_1(v)h_2(u,v)dudv\,| \le$$

$$\le \frac{t^4}{n^2\sigma_n^4}\,E|h_1(U_1)+h_1(U_2)|^4|h_2(U_1,U_2)|.$$

Application of part (ii) of the remark made after lemma 3.4 shows that the term on the right of (4.39) is $\mathcal{O}(n^{-2}\sigma_n^{-4}t^4)$, as $n \to \infty$. Next we remark that lemma 4.2 implies that $\sigma_n^{-1} = \sigma^{-1} + \mathcal{O}(n^{-1})$, as $n \to \infty$. Inserting this result in (4.39) we have proved (4.26). □

We are now in a position to prove (4.1). We first apply lemma 4.3 with $0 < \varepsilon < \frac{1}{2}$ to see that the integral on the left of (4.16) is $o(n^{-1})$, as $n \to \infty$. Secondly we use lemma 4.4 with $0 < \varepsilon < \frac{1}{6}$ to find that the integral on the left of (4.22) is also $o(n^{-1})$, as $n \to \infty$. To proceed let us note that we can write down $\tilde{\rho}_n(t)$ explicitly as

$$(4.40)\qquad \tilde{\rho}_n(t) = e^{-\frac{t^2}{2}}\left(1-\frac{it^3\kappa_3}{6n^{\frac{1}{2}}}+\frac{3\kappa_4t^4-\kappa_3^2t^6}{72\;n}\right).$$

Next we apply (4.40) and the results of the lemma's 4.5 and 4.6 to check that for all $n \ge 1$

$$\int_{|t|\le an^{\frac{1}{2}}} |\tau_{1n}(t) - \tilde{\rho}_n(t)|\,|t|^{-1}dt = \mathcal{O}(\delta_n n^{-1})$$

with a and δ_n as in lemma 4.5. Hence we can conclude that (4.1) holds for $0 < \varepsilon < \frac{1}{6}$ under the assumptions of theorem 2.1.

Next we consider (4.2) and (4.3). Finding sufficient conditions for (4.2) is a problem of an entirely different nature, which was solved by VAN ZWET (1977). In his theorem 4.1 he obtains a bound for the characteristic function for a linear combination of order statistics. This result of VAN ZWET (1977) provides the argument at exactly the same place in our proof where Cramer's condition (C) (see CRAMER (1962)) is used in the classical proof for sums of independent identically distributed r.v.'s.

To prove (4.2) we remark first that application of theorem 4.1 of VAN ZWET (1977) shows that his bound applies to our situation, under the conditions of theorem 2.1. It is also clear from VAN ZWET (1977) that the only missing ingredient to complete the proof of (4.2) is the requirement that there exist positive numbers e and E such that $e \le n^{\frac{1}{2}}\sigma(T_n) \le E$ for all $n \ge 1$. To see this we first use the lemma's 3.5 and 4.1 to find that $n^{\frac{1}{2}}\sigma(S_n)$ is bounded away from zero and infinity and then apply (4.20) (c.f. also (5.10)). Hence (4.2) is shown to hold under the assumption of theorem 2.1.

To prove (4.3) we simply use (4.40) and note that, under the assumptions theorem 2.1 and lemma 3.5 there exist positive constants A_3 and A_4 such that $|\kappa_3| \le A_3$ and $|\kappa_4| \le A_4$. Since the assumptions of theorem 2.1 imply those of lemma 3.5 this completes the proof of theorem 2.1.

To conclude this section it may be useful to mention that if we suppose that, for some $\delta > 0$, $\beta_{4+\delta} = E|X_1|^{4+\delta} < \infty$ and that the assumptions of theorems 2.1 are all satisfied the expansion K_n established in theorem 2.1 is in fact an Edgeworth expansion; i.e. $\kappa_3 n^{-\frac{1}{2}}$ and $\kappa_4 n^{-1}$ are the first order terms in the asymptotic expansions for the third and fourth cumulant of T_n^*, whereas the higher order terms in these expansions can be proved to be of order $o(n^{-1})$.

5. PROOF OF THEOREM 2.2.

In this section we present a proof of theorem 2.2. To start with we remark that for each $n \ge 1$ and real x

$$(5.1) \qquad G_n(x) = F_n^*(x\sigma n^{-\frac{1}{2}}\sigma^{-1}(T_n) + (\mu - E(T_n))\sigma^{-1}(T_n)).$$

Using this identity and applying theorem 2.1 we find that

(5.2) $$\sup_x |G_n(x) - K_n(x\sigma n^{-\frac{1}{2}}\sigma^{-1}(T_n) + (\mu - E(T_n))\sigma^{-1}(T_n))| = O(\delta_n n^{-1}),$$

as $n \to \infty$ holds under the assumptions of theorem 2.1. To proceed we shall need expansions for $\sigma n^{-\frac{1}{2}}\sigma^{-1}(T_n)$ and $(\mu - E(T_n))\sigma^{-1}(T_n)$. In the following lemma we give these expansions.

<u>LEMMA 5.1</u>. *Suppose that* J *has a bounded third derivative on* (0,1) *and* $\beta_{2+\delta} < \infty$ *for some* $\delta > 0$. *Then* $\sigma^2(J,F) > 0$ *implies that*

(5.3) $$|(\mu - E(T_n))\sigma^{-1}(T_n) - an^{-\frac{1}{2}}| = O(n^{-\frac{3}{2}})$$

and

(5.4) $$|\sigma n^{-\frac{1}{2}}\sigma^{-1}(T_n) - 1 + bn^{-1}| = O(n^{-\frac{3}{2}}), \qquad \text{as } n \to \infty.$$

with $a = a(J,F)$ *and* $b = b(J,F)$ *as in* (2.17) *and* (2.18).

<u>PROOF</u>. We first prove (5.4). Application of lemma 4.1, (4.20) and the Cauchy-Schwarz inequality yields

(5.5) $$\frac{\sigma^2}{n\sigma^2(T_n)} = \frac{\sigma^2}{n\sigma^2(S_n)}(1 + O(n^{-\frac{3}{2}})), \qquad \text{as } n \to \infty.$$

Lemma 4.1 implies that

(5.6) $$\frac{\sigma^2}{n\sigma^2(S_n)} = 1 - 2\frac{b}{n} + O(n^{-\frac{3}{2}}), \qquad \text{as } n \to \infty.$$

Combining (5.5) and (5.6) we find

(5.7) $$\frac{\sigma^2}{n\sigma^2(T_n)} = 1 - 2\frac{b}{n} + O(n^{-\frac{3}{2}}), \qquad \text{as } n \to \infty.$$

Inequality (5.4) is an immediate consequence of (5.7). To prove (5.3) we first use (4.20) again to see that

(5.8) $$ET_n = ES_n + O(E|T_n - S_n|) = ES_n + O(n^{-2}), \qquad \text{as } n \to \infty.$$

Using the definition of S_n (cf. (4.6)) and noting that $E(\Gamma_n(s)-s)^3 = n^{-2}s(1-s)(1-2s)$ we can easily check that

$$ES_n = \mu - a\sigma n^{-1} + O(n^{-2}), \qquad \text{as } n \to \infty$$

so that (5.8) implies that

(5.9) $$\mu - ET_n = a\sigma n^{-1} + O(n^{-2}), \qquad \text{as } n \to \infty.$$

Because (5.7) directly implies that

(5.10) $$\sigma^{-1}(T_n) = n^{+\frac{1}{2}}\sigma^{-1} + O(n^{-\frac{1}{2}}), \qquad \text{as } n \to \infty,$$

we have proved (5.3). □

To complete the proof of theorem 2.2 we use (2.4), (2.19), (5.3), (5.4) and lemma 3.5 and apply a simple Taylor expansion argument to find that

(5.11) $$K_n(xn^{-\frac{1}{2}}\sigma^{-1}(T_n)\sigma + (\mu-E(T_n))\sigma^{-1}(T_n)) = L_n(x) + O(n^{-\frac{3}{2}}), \text{ as } n \to \infty,$$

uniformly in x. Combining this with (5.2) completes the proof of theorem 2.2.

ACKNOWLEDGEMENT

The author is very grateful to W.R. Van Zwet for suggesting the problem and for his kind and essential help during the preparation of this paper. He also provided a proof for relation (4.2), a crucial step in solving our problem. The author also wishes to thank a referee for his helpful comments and M. Bakker of the Mathematical Centre, Amsterdam, for writing the computer programs which were necessary to obtain the numerical results given in section 2.

REFERENCES

BICKEL, P.J. (1967), *Some contributions to the theory of order statistics,* Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1.

BICKEL, P.J. (1974), *Edgeworth expansions in non-parametric statistics,* Am. Statist. 2, 1-20.

BICKEL, P.J. & W.R. VAN ZWET (1977), *Asymptotic expansions for the power of distribution free tests in the two-sample problem,* Mathematical Centre Report SW 38/77, Amsterdam; to appear in the Annals of Statistics.

BJERVE, S. (1974), *Error bounds and Asymptotic expansions for linear combinations of order statistics,* unpublished Ph.D. thesis, Berkeley.

BJERVE, S. (1977), *Error bounds for linear combinations of order statistics,* Ann. Statist. 5, 237-255.

CHERNOFF, H., J. GASTWIRTH & M.V. JOHNS (1967), *Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation,* Ann. Math. Statist. 38, 52-72.

CHUNG, K.L. (1951), *The strong law of large numbers,* Proceedings of the second Berkeley Symposium on Mathematical Statistics and Probability, Vol. 2, 341-352.

CRAMER, H. (1962), *Random variables and Probability distributions,* Second Edition, Cambridge University Press.

DAVID, H.A. (1970), *Order statistics,* Wiley, New York.

ESSEEN, C.G. (1945), *Fourier analysis of distribution functions*, A mathematical study of the Laplace-Gaussian law, Acta Mathematica 77, 1-125.

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

GNEDENKO, B.V. & A.N. KOLMOGOROV (1954), *Limit distributions for sums of independent random variables,* Addison-Wesley Publishing Company, Inc., Cambridge, Mass.

HELMERS, R. (1977a), *The order of the normal approximation for linear combinations of order statistics with smooth weight functions*, Ann. Prob. 5, 940-953.

HELMERS, R. (1977b), *A Berry-Esseen theorem for linear combinations of order statistics*, Mathematical Centre Report SW 54/77, Amsterdam, to appear in the Annals of Probability.

HELMERS, R. (1978), *Edgeworth expansions for trimmed linear combinations of order statistics*, Mathematical Centre Report SW 55/77, Amsterdam, (in preparation).

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

MOORE, D.S. (1968), *An elementary proof of asymptotic normality of linear combinations of order statistics*, Ann. Math. Statist. 39, 263-265.

PFANZAGL, J. (1977), *First order efficiency implies second order efficiency*, to appear in J. Hajek Memorial Volume, Academia Press; North-Holland, Amsterdam.

SHORACK, G.R. (1972), *Functions of order statistics*, Am. Math. Statist. 40, 2041-2050.

STIGLER, S.M. (1974), *Linear functions of order statistics with smooth weight functions*, Ann. Stat. 2, 676-693.

ZWET, W.R. VAN (1977), *Asymptotic expansions for the distribution functions of linear combinations of order statistics*, Statistical Decision Theory and Related Topics II, p. 421-437.

Department of Statistics
Mathematical Centre
2e Boerhaavestraat 49
1091 AL AMSTERDAM
THE NETHERLANDS