ON THE CENTERING OF A SIMPLE LINEAR RANK STATISTIC

by

Wassily Hoeffding

University of North Carolina

Institute of Statistics Mimeo Series No. 585

June 1968

This research was supported by the Mathematics Division of the Air Force Office of Scientific Research Grant No. AF-AFOSR-68-1415.

DEPARTMENT OF STATISTICS

University of North Carolina

Chapel Hill, N. C.

SUMMARY

Hájek [3] proved that under weak conditions the distribution of a simple linear rank statistic S (see (1.1)) is asymptotically normal (ES, σ^2), with σ^2 defined in (1.6). He left open the question whether under the same conditions the centering constant ES may be replaced by the simpler constant μ defined by (1.8), as was found to be true in the two-sample case and under different conditions by Chernoff and Savage [1] and Govindarajulu, LeCam and Rhagavachari [2]. In this paper it is shown that the replacement of ES by μ is permissible if one of Hájek's conditions is slightly strengthened. The relation of the problem to one in the theory of polynomial approximation is noted.

1. Introduction and statement of results. Hájek [3] studied the asymptotic distribution of the sum

$$(1.1) \qquad S = \sum_{i=1}^{N} c_i \, a_N \, (R_i),$$

called a simple linear rank statistic, where $c_1, \ldots, c_N$ are real numbers, $R_1, \ldots, R_N$ are the respective ranks of N independent random variables $X_1, \ldots, X_N$ whose distribution functions $F_1, \ldots, F_N$ are continuous, and the so-called scores $a_N(i)$ are generated by a function $\varphi(t)$, $0 < t < 1$, in either of the following two ways:

$$(1.2) \qquad a_N(i) = \varphi(i/(N+1)), \qquad i = 1, \ldots, N,$$

$$(1.3) \qquad a_N(i) = E\varphi(U_N^{(i)}), \qquad i = 1, \ldots, N.$$

Here $U_N^{(i)}$ denotes the i-th order statistic in a sample of size N from the uniform distribution on (0,1). Hájek proved four theorems asserting the asymptotic normality of S under different conditions, of which we quote

Hájek's Theorem 2.3. Let $\varphi(t) = \varphi_I(t) - \varphi_{II}(t)$, $0 < t < 1$, where $\varphi_I(t)$ and $\varphi_{II}(t)$ are both non-decreasing, square integrable, and absolutely continuous inside (0,1). Then for every $\epsilon > 0$ and $\eta > 0$ there exists $N(\epsilon, \eta)$ such that

$$(1.4) \qquad N > N(\epsilon,\eta), \ \text{var } S > \eta N \max_{1 \leq i \leq N} (c_i - \bar{c})^2$$

implies

$$(1.5) \qquad \sup_x |P(S-ES < x(\text{var } S)^{\frac{1}{2}}) - \Phi(x)| < \epsilon.$$

The assertion remains true if we replace var S in (1.4) and (1.5) by

$$(1.6) \quad \sigma^2 = \Sigma_{i=1}^N \text{var } \ell_i(X_i), \quad \ell_i(x) = N^{-1}\Sigma_{j=1}^N (c_j - c_i) \int \{u(y-x) - F_i(y)\} \varphi'(H(y)) dF_j(y).$$

Here $\bar{c} = N^{-1}\Sigma_{i=1}^{N} c_i$, $\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp(-y^2/2)dy$, $u(x) = 1$ or 0 according as $x \geq 0$ or $x < 0$, φ' denotes the derivative of φ, and

$$(1.7) \qquad H(x) = N^{-1}\Sigma_{i=1}^{n} F_i(x).$$

Hájek's Theorem 2.4 states that under the conditions of Theorem 2.3, for every $\epsilon > 0$ and $\eta > 0$ there exist $N(\epsilon, \eta)$ and $\delta(\epsilon, \eta)$ such that the conclusion of Theorem 2.3 holds with var S replaced by

$$d^2 = \Sigma_{i=1}^{N}(c_i - \bar{c})^2 \int_0^1 \{\varphi(t) - \bar{\varphi}\}^2 \, dt,$$

where $\bar{\varphi} = \int_0^1 \varphi(t)dt$, provided that the condition $\max_{i,j,x} |F_i(x) - F_j(x)| < \delta(\epsilon, \eta)$ is added to (1.4).

Hájek's theorems are extensions of the earlier results of Chernoff and Savage [1] and of Govindarajulu, LeCam and Raghavachari [2], which are concerned with the special case $c_1 = \ldots = c_m$, $c_{m+1} = \ldots = c_N$, $F_1 = \ldots = F_m$, $F_{m+1} = \ldots = F_N$ (two-sample case). Apart from different sets of assumptions (which, in essential parts, are more restrictive than Hájek's), the theorems of [1] and [2] differ from Hájek's theorems in that the centering constant ES is replaced by

$$(1.8) \qquad \mu = \Sigma_{i=1}^{N} c_i \int_{-\infty}^{\infty} \varphi(H(x))dF_i(x).$$

The problem of whether ES may be replaced by μ is of interest since, typically, μ is easier to evaluate. Hájek observed ([3], p. 330) that he did not succeed in showing that this replacement is possible under the conditions of Theorems 2.3 and 2.4.

In this paper it is shown that if the condition of square integrability of φ_I and φ_{II} is slightly strengthened, then the conclusions of Theorems 2.3 and 2.4 remain true with ES replaced by μ or by

$$(1.9) \qquad \mu' = \mu + \bar{c}\{\Sigma_{i=1}^{N}\varphi(i/(N+1)) - N\int_0^1 \varphi(t)dt\}.$$

Explicitly, the following result is proved.

<u>Theorem 1</u>. Let $\varphi(t) = \varphi_I(t) - \varphi_{II}(t)$ satisfy the conditions of Hájek's Theorem 2.3 and the additional condition

$$(1.10) \qquad \int_0^1 |\varphi_k(t)|\, t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\, dt < \infty, \qquad k = I, II.$$

Then the conclusions of Hájek's Theorems 2.3 and 2.4 remain true with ES replaced by μ' in the case of the scores (1.2) and by μ in the case of the scores (1.3). Moreover, if $|\bar{c}|/\max_i|c_i - \bar{c}|$ is bounded, ES may be replaced by μ also in the case (1.2).

Condition (1.10) implies square integrability of the monotone functions φ_k (see Lemma 1 below). It has been shown in [4] that if φ_k is monotone and $\int_0^1 \varphi_k^2(t)\{\log\,(1+|\varphi_k(t)|)\}^{2+\delta}\, dt < \infty$ for some $\delta > 0$, then (1.10) is satisfied. In this sense condition (1.10) is not much stronger than square integrability. The author does not know whether square integrability of φ_I and φ_{II} is sufficient for the conclusion of Theorem 1.

Theorem 1 is proved in Section 5. The proof depends on the following two propositions which are demonstrated in Sections 3 and 4.

<u>Proposition 1</u>. There is a numerical constant C_1 such that if φ is non-decreasing, then

$$(1.11) \qquad \Sigma_{i=1}^{N} |E\varphi(U_N^{(i)}) - \varphi(i/(N+1))| \le C_1\, N^{\frac{1}{2}} \int_0^1 |\varphi(t)|\, t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\, dt.$$

Proposition 2. There is a numerical constant C_2 such that if φ is nondecreasing and $F_1, \ldots, F_N$ are any continuous distribution functions, then

$$(1.12) \qquad \Sigma_{i=1}^{N} \left|E\varphi(R_i/(N+1)) - \int_{-\infty}^{\infty} \varphi(H(x))dF_i(x)\right| \leq C_2 N^{\frac{1}{2}} \int_0^1 |\varphi(t)| t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\, dt.$$

Remark. The integral on the right of (1.11) and (1.12) may be replaced by $\inf_{-\infty<a<\infty} \int_0^1 |\varphi(t)-a| t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\, dt$. This is due to the fact that the left sides of the inequalities are not changed if φ is replaced by φ + const.

Proposition 1 gives an upper bound for a distance between the sequences of scores (1.2) and (1.3). Proposition 2 has an obvious bearing on the estimation of $|ES-\mu'|$ with scores (1.2). Taken together with Proposition 1, it will be seen (in Section 5) to imply an inequality analogous to (1.12) for the scores (1.3).

The stated results are closely related to a problem in polynomial approximation. For a function φ which is Lebesgue integrable on (0,1) define the (modified) Bernstein polynomial of order N by

$$P_N^{\varphi}(t) = \Sigma_{i=0}^{N} (N+1) \int_{i/(N+1)}^{(i+1)/(N+1)} \varphi(u)du \binom{N}{i} t^i (1-t)^{N-i}$$

(see Lorentz [6], Chapter II). It is shown in [4] that if φ is nondecreasing then

$$\int_0^1 |P_N^{\varphi}(t) - \varphi(t)|dt \leq C\, N^{-\frac{1}{2}} \int_0^1 |\varphi(t)| t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\, dt,$$ where C is a numerical constant. The proof is similar to that of Proposition 2.

2. Preliminary lemmas. The following lemmas will be used in the proofs. The symbols C_1, C_2, C_3, ... will denote numerical constants.

Lemma 1. If φ is nondecreasing,

$$(2.1)\qquad \left\{\int_0^1 \varphi^2(t)dt\right\}^{\frac{1}{2}} \leq \int_0^1 |\varphi(t)|\, t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\,dt.$$

Lemma 2. If φ is nondecreasing,

$$(2.2)\qquad \Sigma_{j=1}^{N-1}\left\{\varphi(\tfrac{j+1}{N+1}) - \varphi(\tfrac{j}{N+1})\right\}(\tfrac{j}{N})^{\frac{1}{2}}(1-\tfrac{j}{N})^{\frac{1}{2}} \leq C_3 \int_0^1 |\varphi(t)|\, t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\,dt.$$

Lemmas 1 and 2 are identical with Lemmas 1 and 2 of [4] and are proved there.

Let G_{Ni} denote the distribution function of $U_N^{(i)}$ and let the random variable $W_N(u)$ have the binomial (N,u) distribution. Then for $i = 1,\ldots,N$,

$$(2.3)\qquad G_{Ni}(u) = P\{U_N^{(i)} < u\} = P\{W_N(u) \geq i\},\ 0 < u < 1.$$

Lemma 3. For $j = 1,\ldots,N-1$,

$$(2.4)\qquad \Sigma_{i=1}^{j}\{1 - G_{Ni}(\tfrac{j}{N+1})\} < 1 + C_4\, N^{\frac{1}{2}}(\tfrac{j}{N})^{\frac{1}{2}}(1 - \tfrac{j}{N})^{\frac{1}{2}}.$$

Proof. By (2.3), for $0 < u < 1$,

$$\Sigma_{i=1}^{j}\{1 - G_{Ni}(u)\} = \Sigma_{i=1}^{j} P\{W_N(u) \leq i-1\} = \Sigma_{i=0}^{j-1}\ \Sigma_{k=0}^{i} P\{W_N(u) = k\}$$
$$= \Sigma_{k=0}^{j}(j-k)P\{W_N(u) = k\} = jP\{W_N(u) \leq j\} - \Sigma_{k=0}^{j} kP\{W_N(u)=k\}$$

Now

$$\Sigma_{k=0}^{j} kP\{W_N(u) = k\} = \Sigma_{k=1}^{j} k\binom{N}{k}u^k(1-u)^{N-k} = NuP\{W_{N-1}(u) \leq j-1\}$$

and $P\{W_N(u) \leq j\} = P\{W_{N-1}(u) \leq j-1\} + (1-u)\,P\{W_{N-1}(u) = j\}$.

Hence

$$(2.5)\qquad \Sigma_{i=1}^{j}\{1 - G_{Ni}(u)\} = (j-Nu)P\{W_{N-1}(u) \leq j-1\} + j\binom{N-1}{j}u^j(1-u)^{N-j}.$$

For $u = j/(N+1)$ we have $j-Nu = j/(N+1) < 1$. The maximum for $u \in (0,1)$ of the last term in (2.5) is attained at $u=j/N$. Hence

(2.6) $$\Sigma_{i=1}^{j}\{1-G_{Ni}(\tfrac{j}{N+1})\} < 1 + j\binom{N-1}{j}(\tfrac{j}{N})^{j}(1-\tfrac{j}{N})^{N-j} = 1 + N\binom{N}{j}(\tfrac{j}{N})^{j+1}(1-\tfrac{j}{N})^{N-j+1}.$$

By Stirling's formula there is a numerical constant C_4 such that

(2.7) $$\binom{N}{j} \leq C_4\, N^{-\frac{1}{2}}\,(\tfrac{j}{N})^{-j-\frac{1}{2}}(1-\tfrac{j}{N})^{-N+j-\frac{1}{2}}, \qquad 1 \leq j \leq N-1.$$

Inequality (2.4) now follows from (2.6) and (2.7).

3. Proof of Proposition 1. Let $\varphi_j = \varphi(j/(N+1))$, $1 \leq j \leq N$, and let G_{Ni} be defined by (2.3). Since φ is nondecreasing,

$$\begin{aligned}
E\varphi(U_N^{(i)}) &= \int_0^1 \varphi(u)dG_{Ni}(u) = \Sigma_{j=0}^{N} \int_{j/(N+1)}^{(j+1)/(N+1)} \varphi(u)dG_{Ni}(u) \\
&\leq \Sigma_{j=0}^{N-1} \varphi_{j+1}\int_{j/(N+1)}^{(j+1)/(N+1)} dG_{Ni}(u) + \int_{N/(N+1)}^{1}\varphi(u)dG_{Ni}(u) \\
&= \varphi_1 + \Sigma_{j=1}^{N-1} (\varphi_{j+1}-\varphi_j)\{1-G_{Ni}(\tfrac{j}{N+1})\} + \int_{N/(N+1)}^{1}\{\varphi(u)-\varphi_N\}dG_{Ni}(u).
\end{aligned}$$

If in the sum in the last line we majorize the j-th term by $\varphi_{j+1}-\varphi_j$ for $1 \leq j \leq i-1$, we obtain

(3.1) $$E\varphi(U_N^{(i)}) - \varphi_i \leq a_i + b_i, \quad i = 1, \ldots, N,$$

where

(3.2) $$a_i = \Sigma_{j=i}^{N-1} (\varphi_{j+1}-\varphi_j)\{1-G_{Ni}(\tfrac{j}{N+1})\}, \quad b_i = \int_{N/(N+1)}^{1}\{\varphi(u)-\varphi_N\}dG_{Ni}(u),$$

with $a_N = 0$. Since $a_i \geq 0$, $b_i \geq 0$, (3.1) implies

$$|E\varphi(U_N^{(i)})-\varphi_i| \leq 2a_i + 2b_i + \varphi_i - E\varphi(U_N^{(i)}).$$

Hence

$$(3.3)\quad \Sigma_{i=1}^{N} |E\varphi(U_N^{(i)}) - \varphi_i| \leq 2A + 2B + D,$$

where

$$(3.4)\quad A = \Sigma_{i=1}^{N-1} a_i = \Sigma_{j=1}^{N-1} (\varphi_{j+1} - \varphi_j) \Sigma_{i=1}^{j} \{1 - G_{Ni}(\tfrac{j}{N+1})\},$$

$$(3.5)\quad B = \Sigma_{i=1}^{N} b_i, \qquad D = \Sigma_{i=1}^{N} \varphi_i - \Sigma_{i=1}^{N} E\varphi(U_N^{(i)}).$$

By Lemmas 3 and 2,

$$(3.6)\quad \begin{aligned} A &\leq \varphi_N - \varphi_1 + C_4 N^{\frac{1}{2}} \Sigma_{j=1}^{N-1} (\varphi_{j+1} - \varphi_j)(\tfrac{j}{N})^{\frac{1}{2}}(1 - \tfrac{j}{N})^{\frac{1}{2}} \\ &\leq \varphi_N - \varphi_1 + C_3 C_4 N^{\frac{1}{2}} \int_0^1 |\varphi(t)| t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}} dt. \end{aligned}$$

Since $\Sigma_{i=1}^{N} G_{Ni}(u) = Nu$,

$$(3.7)\quad B = N \int_{N/(N+1)}^{1} \varphi(u)du - N(N+1)^{-1}\varphi_N.$$

Since $\Sigma_{i=1}^{N} \varphi_i \leq \Sigma_{i=1}^{N} (N+1)\int_{i/(N+1)}^{(i+1)/(N+1)} \varphi(t)dt = (N+1)\int_{1/(N+1)}^{1} \varphi(t)dt$

and $\Sigma_{i=1}^{N} E\varphi(U_N^{(i)}) = N \int_0^1 \varphi(t)dt$,

$$(3.8)\quad D \leq \int_{1/(N+1)}^{1} \varphi(t)dt - N\int_0^{1/(N+1)} \varphi(t)dt.$$

Now

$$(3.9)\quad (N+1)N^{-1}\int_0^{N/(N+1)} \varphi(t)dt \leq \varphi_N \leq (N+1)\int_{N/(N+1)}^{1} \varphi(t)dt,\ \varphi_1 \geq (N+1)\int_0^{1/(N+1)} \varphi(t)dt.$$

From (3.7) to (3.9) we obtain after repeated application of Schwarz's inequality

$$(3.10)\quad 2(\varphi_N - \varphi_1) + 2B + D \leq C_5 N^{\frac{1}{2}} \{\int_0^1 \varphi^2(t)dt\}^{\frac{1}{2}}.$$

Combining (3.3), (3.6) and (3.10) and applying Lemma 1, we obtain inequality (1.11) of Proposition 1 with $C_1 = 2C_3C_4 + C_5$.

4. Proof of Proposition 2. Let

$$\psi_i(x) = E\{\varphi(R_i/(N+1))\,|X_i = x\}. \tag{4.1}$$

Then

$$\Sigma_{i=1}^{N} \left|E\varphi(R_i/(N+1)) - \int_{-\infty}^{\infty} \varphi(H(x))dF_i(x)\right| \le \Sigma_{i=1}^{N} \int_{-\infty}^{\infty} |\psi_i(x) - \varphi(H(x))|dF_i(x). \tag{4.2}$$

We have

$$\psi_i(x) = \Sigma_{j=1}^{N} \varphi_j\ P\{R_i = j\,|X_i = x\}$$

$$= \varphi_1 + \Sigma_{j=1}^{N-1} (\varphi_{j+1} - \varphi_j)\ P\{R_i \ge j+1\,|X_i = x\}.$$

Now $R_i = \Sigma_{k=1}^{N} u(X_i - X_k)$, where $u(\cdot)$ is defined after (1.6). Hence if we let

$$V(x) = \Sigma_{k=1}^{N} u(x - X_k), \tag{4.3}$$

then $P\{R_i \ge j+1\,|X_i = x\} \le P\{V(x) \ge j\}$ for all i and

$$\psi_i(x) \le \varphi_1 + \Sigma_{j=1}^{N-1} (\varphi_{j+1} - \varphi_j)\ P\{V(x) \ge j\}. \tag{4.4}$$

By Theorem 1 of [5],

$$P\{V(x) \ge j\} \le (\tfrac{j}{N})^{-j}\ (1 - \tfrac{j}{N})^{-N+j}\ H(x)^j\ \{1-H(x)\}^{N-j} \quad \text{for } j \ge NH(x). \tag{4.5}$$

If we majorize $P\{V(x) \ge j\}$ in (4.4) by 1 for $j \le [NH(x)]$ and apply (4.5) to the terms with $j \ge [NH(x)]+1$, we obtain

$$\psi_i(x) - \varphi_{[NH(x)]+1} \le J(H(x)), \quad 0 < H(x) < 1, \tag{4.6}$$

where

$$J(t) = \Sigma_{j=[Nt]+1}^{N-1} (\varphi_{j+1} - \varphi_j)(\tfrac{j}{N})^{-j}(1 - \tfrac{j}{N})^{-N+j}\ t^j(1-t)^{N-j} \tag{4.7}$$

for $0 < t \le 1 - N^{-1}$ and $J(t) = 0$ for $1 - N^{-1} < t < 1$.

Since $J(t) \geq 0$, (4.6) implies

$$|\psi_i(x) - \varphi_{[NH(x)]+1}| \leq 2J(H(x)) + \varphi_{[NH(x)]+1} - \psi_i(x).$$

Therefore, by (1.7),

$$(4.8)\quad \Sigma_{i=1}^{N} \int_{-\infty}^{\infty} |\psi_i(x) - \varphi_{[NH(x)]+1}| dF_i(x) \leq 2N\int_0^1 J(t)dt + N\int_0^1 \varphi_{[Nt]+1}\, dt$$
$$- \Sigma_{i=1}^{N} E\varphi(R_i/(N+1)).$$

Now

$$(4.9)\quad N\int_0^1 \varphi_{[Nt]+1}\, dt = \Sigma_{k=1}^{N} \varphi(k/(N+1)) = \Sigma_{i=1}^{N} E\varphi(R_i/(N+1)$$

and

$$\int_0^1 J(t)dt = \Sigma_{k=1}^{N-1} \int_{(k-1)/N}^{k/N} J(t)dt$$
$$= \Sigma_{k=1}^{N-1} \Sigma_{j=k}^{N-1} (\varphi_{j+1} - \varphi_j)(\tfrac{j}{N})^{-j}(1-\tfrac{j}{N})^{-N+j}\int_{(k-1)/N}^{k/N} t^j(1-t)^{N-j}dt$$
$$= \Sigma_{j=1}^{N-1} (\varphi_{j+1} - \varphi_j)(\tfrac{j}{N})^{-j}(1-\tfrac{j}{N})^{-N+j}\int_0^{j/N} t^j(1-t)^{N-j}dt.$$

By Stirling's formula, for $1 \leq j \leq N-1$,

$$\int_0^{j/N} t^j(1-t)^{N-j}\, dt \leq \int_0^1 t^j(1-t)^{N-j}\, dt = \frac{j!\,(N-j)!}{(N+1)!}$$
$$\leq C_6\, N^{-\frac{1}{2}}(\tfrac{j}{N})^{j+\frac{1}{2}}\,(1-\tfrac{j}{N})^{N-j+\frac{1}{2}}\,.$$

Hence

$$(4.10)\quad \int_0^1 J(t)dt \leq C_6\, N^{-\frac{1}{2}}\, \Sigma_{j=1}^{N-1} (\varphi_{j+1} - \varphi_j)(\tfrac{j}{N})^{\frac{1}{2}}(1-\tfrac{j}{N})^{\frac{1}{2}}.$$

It follows from (4.8) to (4.10) that

$$(4.11)\quad \Sigma_{i=1}^{N}\int_{-\infty}^{\infty} |\psi_i(x) - \varphi_{[NH(x)]+1}| dF_i(x) \leq 2C_6 N^{\frac{1}{2}}\Sigma_{j=1}^{N-1}(\varphi_{j+1} - \varphi_j)(\tfrac{j}{N})^{\frac{1}{2}}(1-\tfrac{j}{N})^{\frac{1}{2}}.$$

Furthermore,

$$\Sigma_{i=1}^{N} \int_{-\infty}^{\infty} |\varphi_{[NH(x)]+1} - \varphi(H(x))| dF_i(x) = N \int_0^1 |\varphi_{[Nt]+1} - \varphi(t)| dt$$

$$\leq N \int_0^1 \{\varphi(\tfrac{Nt+1}{N+1}) - \varphi(\tfrac{Nt}{N+1})\} dt$$

$$= (N+1)\left\{\int_{N/(N+1)}^{1} \varphi(t)dt - \int_0^{1/(N+1)} \varphi(t)dt\right\}$$

$$(4.12) \quad \leq (N+1)^{\frac{1}{2}} \left\{ \left(\int_{N/(N+1)}^{1} \varphi^2(t)dt\right)^{\frac{1}{2}} + \left(\int_0^{1/(N+1)} \varphi^2(t)dt\right)^{\frac{1}{2}}\right\}$$

$$(4.13) \quad \leq 4\, N^{\frac{1}{2}} \left\{\int_0^1 \varphi^2(t)dt\right\}^{\frac{1}{2}} .$$

It now follows from (4.2), (4.11), (4.13) and Lemmas 1 and 2 that inequality (1.12) holds with $C_2 = 2C_6 + 4$.

5. Proof of Theorem 1. The following lemma will be used.

Lemma 4. If φ satisfies the conditions of Theorem 1, then for every $\alpha > 0$ there exists a decomposition

$$(5.1) \qquad \varphi(t) = \psi(t) + \varphi^{(1)}(t) - \varphi^{(2)}(t), \qquad 0 < t < 1,$$

such that ψ is a polynomial, $\varphi^{(1)}$ and $\varphi^{(2)}$ are nondecreasing, and

$$(5.2) \qquad \int_0^1 (|\varphi^{(1)}(t)| + |\varphi^{(2)}(t)|) t^{-\frac{1}{2}} (1-t)^{-\frac{1}{2}} dt < \alpha .$$

Lemma 4 is an analog of Lemma 5.1 of Hájek [3], which differs from Lemma 4 in that φ is assumed to satisfy the conditions of Theorem 2.3 and (5.2) is replaced by $\int_0^1 |\varphi^{(1)}(t)|^2 dt + \int_0^1 |\varphi^{(2)}(t)|^2 dt < \alpha$. Hájek's proof of Lemma 5.1 serves without change to prove Lemma 4.

It will be sufficient to prove the assertion of Theorem 1 concerning Theorem 2.3 since for Theorem 2.4 the proof is analogous. First let S be defined with $a_N(i) = \varphi(i/(N+1))$. To prove the statement of the theorem with centering constant μ', it is enough to show that for every $\beta > 0$ and $\eta > 0$ there exists a number $N' = N'(\beta,\eta)$ such that

(5.3) $$N > N', \quad \text{var } S > \eta N \max_{1\le i\le N}(c_i-\bar{c})^2$$
implies
(5.4) $$|ES - \mu'|/(\text{var } S)^{\frac{1}{2}} < \beta.$$

Indeed, given $\epsilon > 0$ and $\eta > 0$, choose $\beta = \beta(\epsilon)$ so that $\max_x |\Phi(x\pm\beta)-\Phi(x)| < \epsilon/2$. Let $N''(\epsilon,\eta) = \max(N'(\beta(\epsilon),\eta)), N(\epsilon/2,\eta))$, with $N(\cdot,\cdot)$ defined in Hájek's Theorem 2.3. Then (1.4) with $N(\epsilon,\eta)$ replaced by $N''(\epsilon,\eta)$ implies (1.5) with ES replaced by μ'.

We write $S(\varphi)$ $\mu'(\varphi)$, for S, μ' to exhibit the dependence on φ. Since $\Sigma_{i=1}^N \varphi(R_i/(N+1)) = \Sigma_{i=1}^N \varphi(i/(N+1))$ and $\Sigma_{i=1}^N \int\varphi(H(x))dF_i(x) = N\int_0^1 \varphi(t)dt$, we have from (1.9)

$$S(\varphi) - \mu'(\varphi) = \Sigma_{i=1}^n(c_i-\bar{c}) \{\varphi(R_i/(N+1))- \int\varphi(H(x))dF_i(x)\}.$$

Hence

(5.5) $$|ES(\varphi)- \mu'(\varphi)| \le \max_{1\le i\le N}|c_i-\bar{c}|\Sigma_{i=1}^N |E\varphi(R_i/(N+1))-\int\varphi(H(x))dF_i(x)|.$$

We apply Lemma 4 with α to be determined later. Clearly

(5.6) $$|ES(\varphi)-\mu'(\varphi)| \le |ES(\psi)-\mu'(\psi)| + \Sigma_{k=1}^2 |ES(\varphi^{(k)})-\mu'(\varphi^{(k)})|.$$

Since ψ has a bounded second derivative, it follows by a Taylor expansion (see Hájek [3], p. 340) that there is a constant $K(\psi)$ such that

(5.7) $$|E\psi(R_i/(N+1)-\int\psi(H(x))dF_i(x)| < K(\psi)N^{-1}, \quad i = 1,\ldots,N.$$

Hence, from (5.5) with $\varphi = \psi$,

(5.8) $$|ES(\psi)-\mu'(\psi)| \le K(\psi) \max_{1\le i\le N} |c_i - \bar{c}|.$$

From (5.5) with $\varphi = \varphi^{(k)}$, Proposition 2, and (5.2),

(5.9) $$\Sigma_{k=1}^{2} |ES(\varphi^{(k)}) - \mu'(\varphi^{(k)})| \leq C_2 N^{\frac{1}{2}} \alpha \max_{1\leq i\leq N} |c_i - \bar{c}|.$$

If var $S > \eta N \max_i (c_i - \bar{c})^2$, it follows from (5.6), (5.8) and (5.9) that

(5.10) $$|ES(\varphi) - \mu'(\varphi)|/(\text{var } S)^{\frac{1}{2}} \leq \eta^{-\frac{1}{2}} K(\psi) N^{-\frac{1}{2}} + C_2 \eta^{-\frac{1}{2}} \alpha .$$

Now, given $\beta > 0$ and $\eta > 0$, choose α in Lemma 4 so that $C_2 \eta^{-\frac{1}{2}} \alpha = \beta/2$. This choice fixes $K(\psi) = K_1(\beta, \eta)$. Define $N' = N'(\beta, \eta)$ by $\eta^{-\frac{1}{2}} K(\psi)(N')^{-\frac{1}{2}} = \beta/2$. Then (5.3) implies (5.4), as was to be proved.

To prove the last part of Theorem 1 concerning the case (1.2), note that, by (1.9),

$$|\mu' - \mu| \leq |\bar{c}| \, |\Sigma_{i=1}^{N} \varphi(i/(N+1)) - N \int_0^1 \varphi(t)dt| = |\bar{c}| N| \int_0^1 \{\varphi(\tfrac{[Nt]+1}{N+1}) - \varphi(t)\}dt|.$$

Since φ is the difference of two non-decreasing, square integrable functions, it follows from (4.12) that

$$|\mu' - \mu| \leq |\bar{c}| \, N^{\frac{1}{2}} K_N,$$

where $K_N = K_N(\varphi) \to 0$ as $N \to \infty$. Hence if var$S > \eta N \max(c_i - \bar{c})^2$, then

$$|\mu' - \mu| \, /(\text{var}S)^{\frac{1}{2}} < \eta^{-\frac{1}{2}} K_N \, |\bar{c}|/\max|c_i - \bar{c}|,$$

which is arbitrarily small for N large enough if $|\bar{c}|/\max|c_i - \bar{c}|$ is bounded. This implies the last part of the theorem.

Finally consider S with $a_N(i) = E\varphi(U_N^{(i)})$. In this case $\Sigma_{i=1}^{N} a_N(i) = N\int_0^1 \varphi(t)dt$, hence

$$S(\varphi) - \mu(\varphi) = \Sigma_{i=1}^{N} (c_i - \bar{c}) \{a_N(R_i) - \int \varphi(H(x)dF_i(x)\},$$

(5.11) $$|ES(\varphi) - \mu(\varphi)| \leq \max_i |c_i - \bar{c}| \, \Sigma_{i=1}^{N} \, |Ea_N(R_i) - \int \varphi(H(x))dF_i(x)|.$$

Now it is easily seen that

$$(5.12)\quad \Sigma_{i=1}^{N} \left|Ea_N(R_i) - \int \varphi(H(x))dF_i(x)\right| \le \Sigma_{i=1}^{N} \left|E\varphi(R_i/(N+1)) - \int \varphi(H(x))dF_i(x)\right|$$

$$+ \Sigma_{i=1}^{N} \left|E\varphi(U_N^{(i)}) - \varphi(i/(N+1))\right|.$$

For $\varphi = \psi$ we apply Taylor's formula to the last term. Since $EU_N^{(i)} = i/(N+1)$ and var $U_N^{(i)} < N^{-1}$ for all i, we find that there is a constant $K'(\psi)$ such that $|E\psi(U_N^{(i)}) - \psi(i/(N+1))| < K'(\psi)N^{-1}$, $i = 1,\ldots,N$. Together with (5.7) this implies an inequality analogous to (5.8). Applying Propositions 1 and 2 to (5.12) with $\varphi = \varphi^{(k)}$, $k = 1,2$, and using Lemma 4, we obtain an inequality analogous to (5.9). Now the conclusion follows as in the first part of the proof.

REFERENCES

[1] Chernoff, H. and Savage, I.R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. Ann. Math. Statist. 29, 972-994.

[2] Govindarajulu, Z., LeCam, L. and Raghavachari, M. (1966). Generalizations of theorems of Chernoff and Savage on the asymptotic normality of test statistics. Proc. Fifth Berkeley Symp. Math. Statist. Prob. 1, 609-638.

[3] Hájek, Jaroslav (1968). Asymptotic normality of simple linear rank statistics under alternatives. Ann. Math. Statist., 39, 325-346.

[4] Hoeffding, Wassily (1968). Approximation of a monotone function by Bernstein polynomials. Inst. of Statist. Mimeo Series No. 581. University of North Carolina.

[5] Hoeffding, Wassily (1963). Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc., 58, 13-30.

[6] Lorentz, G.G. (1953). Bernstein Polynomials. University of Toronto Press, Toronto.