The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 770–788

LINEAR FUNCTIONS OF ORDER STATISTICS[1]

BY STEPHEN MACK STIGLER[2]

University of California, Berkeley

0. Introduction and summary. The purpose of this paper is to investigate the asymptotic normality of linear combinations of order statistics; that is, to find conditions under which a statistic of the form $S_n = \sum_{i=1}^n c_{in}X_{in}$ has a limiting normal distribution as n becomes infinite, where the c_{in}'s are constants and $X_{1n}, X_{2n}, \cdots, X_{nn}$ are the observations of a sample of size n, ordered by increasing magnitude. Aside from the sample mean (the case where the weights c_{in} are all equal to $1/n$), the first proof of asymptotic normality within this class was by Smirnov in 1935 [19], who considered the case that nonzero weight is attached to at most two percentiles. In 1946, Mosteller [13] extended this to the case of several percentiles, and coined the phrase "systematic statistic" to describe S_n. Since the publication in 1955 of a paper by Jung [11] concerned with finding optimal weights for S_n in certain estimation problems, interest in proving its asymptotic normality under more general conditions has grown. For example, Weiss in [21] proved that S_n has a limiting normal distribution when no weight is attached to the observations below the pth sample percentile and above the qth sample percentile, $p < q$, and the remaining observations are weighted according to a function J by $c_{in} = J(i/(n + 1))$, where J is assumed to have a bounded derivative between p and q.

Within the past few years, several notable attempts have been made to prove the asymptotic normality of S_n under more general conditions on the weights and underlying distribution. These attempts have employed three essentially different methods. In [1] Bickel used an invariance principle for order statistics to prove asymptotic normality when $\sum_{i<tn} c_{in}$ converges to a function $J(t)$ of bounded variation and the underlying distribution F has a continuous density, positive on the support of F. His method was quite successful in dealing with statistics which put no weight on observations below the pth and above the qth sample percentile, $p < q$, but in other cases he did not allow the more extreme observations to be weighted more than in the sample mean. More recently in [17], Shorack used a more high powered version of the same approach to obtain a stronger result, allowing much more weight on the extremes. Chernoff, Gastwirth, and Johns, in [3], employed a device of Rényi [14] and expressed S_n as a linear combination of independent, exponentially distributed random variables plus a remainder term; they then showed that the sum of independent variables has a limiting normal distribution and the remainder is asymptotically negligible

Received 11 September 1967.

[1] Prepared with the partial support of the National Science Foundation, Grants GP-2593 and GP-5059 and a National Aeronautics and Space Administration Predoctoral Traineeship.

[2] Now at the University of Wisconsin.

in the sense that it tends to zero in probability as n tends to infinity. They proved asymptotic normality if the c_{in}'s give substantial weight to only a fixed number of percentiles and obey certain boundedness conditions elsewhere, and F has a density which is continuous outside a set of measure zero, bounded away from zero on the interior of the support of F, and smooth in the tails. A different approach is used by Govindarajulu in [7]. He adopted a version due to LeCam of the method of Chernoff and Savage [4], which was used by Govindarajulu, LeCam, and Raghavachari in [8] to prove asymptotic normality of linear rank statistics, and for the case $c_{in} = J(i/(n+1))$ he expressed S_n as a linear combination of independent random variables plus a remainder term. Then if J is absolutely continuous and both it and its derivative satisfy certain boundedness conditions at zero and one, the remainder term tends to zero in probability and S_n has a limiting normal distribution. While his conditions on the weights are very restrictive (for example he does not allow substantial weight to be put on sample percentiles), his conditions on the underlying distribution are the weakest yet obtained, requiring only that the inverse of F does not grow very rapidly at zero or one. Recently in [12], Moore gave a short proof of asymptotic normality, also along the lines of Chernoff and Savage [4], which permits quite general F, but again at the expense of stringent conditions on J.

In this investigation we use yet another method of attacking this problem. Using a procedure due to Hájek [9], who applied it to linear rank statistics, we will represent the statistic S_n as a linear combination of independent random variables, to which the usual central limit theory can be applied, plus a remainder term, and then, under quite general conditions on the weights, prove that the remainder converges to zero *in mean square*, rather than in the weaker sense of convergence in probability as in [3], [7], and [12]. This is accomplished by first approximating the statistic $T_n = \sum_{i=b_n}^{n-b_n} c_{in} X_{in}$ by a sum of independent random variables, where b_n is a sequence of integers tending to infinity slower than n but faster than $\log n$ as n increases, and then finding conditions under which T_n approximates S_n in mean square. The result is essentially stronger than that of Bickel, in that much more weight is allowed on the extreme observations; however, a smoothness condition on the distribution similar to that of Chernoff, Gastwirth, and Johns is required. For statistics of the form T_n the result is essentially stronger than that of Chernoff, Gastwirth, and Johns, although slightly more restrictive conditions are required to prove mean square equivalence of T_n and S_n. We treat a much more general class of weights than does Govindarajulu or Moore, but our conditions on the underlying distribution are stronger than their conditions.

In the course of the proof we derive asymptotic expressions for the covariances of the order statistics and the variance of S_n. For the case of the variance of a single order statistic X_{in}, the result is proved under weaker conditions on F and for a wider range of i than by Sen [16], Van Zwet [20], Bickel [1], or Blom [2], although the speed of convergence is slower than in [16], [20], or [2]. The asymptotic expression for covariances is similarly improved over Blom [2].

The first section of the paper describes the method to be used and states two previously known propositions which will be useful in the following sections. Section two contains the calculation of an approximation for an order statistic by a sum of independent random variables and an exact expression for the covariance of two such approximations. Asymptotic expressions for this covariance and the covariance of two order statistics are derived in section three, as well as an expression for the variance of S_n. Section four contains the proof of the asymptotic normality of S_n when the extremes are not included and no weight is allowed for $i/(n+1)$ near a point where the derivative of the inverse of F misbehaves, and conditions are given under which these restrictions can be dropped. Finally, in section five we discuss the limitations of the method used and extend the results to the slightly more general class of statistics of the form $\sum c_{in}h(X_{in})$, also considered in [3], [7], and [17].

1. Preliminaries. Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a population having a continuous cumulative distribution function $F(x)$ and a density $f(x)$ with respect to Lebesgue measure. Let $X_{1n} \leqq X_{2n} \leqq \cdots \leqq X_{nn}$ denote the order statistics of the sample. For $0 < t < 1$ let $F^{-1}(t) = \min\{x:F(x) = t\}$, an inverse of $F(x)$, and denote $\psi(t) = (d/dt)F^{-1}(t) = 1/f(F^{-1}(t))$ when it exists. Let $p_i = i/(n+1)$. We shall consider statistics of the form

$$S_n = \textstyle\sum_{i=1}^n c_{in}X_{in}, \tag{1.1}$$

where $\{c_{in}\}$, $1 \leqq i \leqq n$, $n \geqq 1$, is a double sequence of known constants. Our method of attack shall basically be that used by Hájek [9] in proving the asymptotic normality of simple linear rank statistics under alternatives; that is, we shall find that linear combination of independent random variables (denoted $\hat{S}_n$ and called the projection of S_n) which most closely approximates S_n in the sense of mean square, show that S_n and $\hat{S}_n$ are asymptotically equivalent in mean square, and prove that $\hat{S}_n$ is asymptotically normally distributed $(E\hat{S}_n, \sigma^2(\hat{S}_n))$. To this end we shall need the following simple propositions.

PROPOSITION 1. (Hájek [9].) *Let $X_1, X_2, \cdots, X_n$ be independent random variables and $\mathfrak{F}$ be the Hilbert space of* a.s. *equivalence classes of square integrable statistics depending on $X_1, \cdots, X_n$. Let $\mathfrak{L}$ be the closed linear subspace of $\mathfrak{F}$ consisting of statistics of the form $L = \sum_{i=1}^n l_i(X_i)$, where the l_i are functions such that $El_i^2(X_i) < \infty$. Then if $S \in \mathfrak{F}$, the projection of S on $\mathfrak{L}$ is given by*

$$\hat{S} = \textstyle\sum_{i=1}^n E(S \mid X_i) - (n-1)ES. \tag{1.2}$$

Thus

$$ES = E\hat{S} \tag{1.3}$$

and

$$E(S-\hat{S})^2 = \sigma^2(S) - \sigma^2(\hat{S}). \tag{1.4}$$

The importance of Proposition 1 is that (1.2) permits the straight forward calculation of the approximating statistic $\hat{S}_n$ in terms of the original random variables $X_1, \cdots, X_n$, and (1.4) reduces the problem of proving the mean square equivalence of S_n and $\hat{S}_n$ to that of obtaining asymptotic expressions for the variances of S_n and $\hat{S}_n$ separately.

Since Proposition 1 requires the existence of the second moment of S_n, the following proposition, essentially due to Bickel [1], is of interest.

PROPOSITION 2. *The following four statements are equivalent*:

(i) *There exists some* $\epsilon > 0$ *such that*

$$\lim_{x\to\infty} |x|^{\epsilon}[1 - F(x) + F(-x)] = 0.$$

(ii) *For any finite number* $k > 0$, *there exists a finite* $r = r(k, F)$ *such that if* $r \leqq i \leqq n - r$, *then* $E|X_{in}|^k < \infty$.

(iii) *There exists a finite* $\tau > 0$ *such that*

$$\lim_{s\to 0} s^{\tau}F^{-1}(s) = \lim_{s\to 1} (1 - s)^{\tau}F^{-1}(s) = 0.$$

(iv) *There exists a finite* $m \geqq 0$ *such that*

$$\int_0^1 \psi(u)[u(1 - u)]^m\, du < \infty.$$

The equivalence of (i)–(iii) is contained in the proof of Bickel's Theorem 2.2 (a). The equivalence of (i) and (iv) is trivial. Throughout this paper we shall assume that (i) holds.

2. The projection of an order statistic. We shall now proceed to calculate an explicit expression for the projection $\hat{S}_n$ of $S_n = \sum_{i=1}^n c_{in}X_{in}$. It follows from the linearity of the projection operator that $\hat{S}_n = \sum_{i=1}^n c_{in}\hat{X}_{in}$. Thus it shall be sufficient to find the projections $\hat{X}_{in}$ of the X_{in}.

Let U_{in} denote the ith order statistic of a random sample of size n from a uniform [0, 1] distribution, and let $g_{in}(u)$ denote the density of U_{in}; that is,

$$g_{in}(u) = n\binom{n-1}{i-1}u^{i-1}(1 - u)^{n-i}, \qquad 0 < u < 1. \tag{2.1}$$

Then X_{in} and $F^{-1}(U_{in})$ have the same distribution, and in particular, $EX_{in} = \int_0^1 F^{-1}(u)g_{in}(u)\, du$.

LEMMA 1. *There is some* $n_0 = n_0(F)$ *such that for* $i \geqq n_0$ *and* $n - i + 1 \geqq n_0$,

$$\hat{X}_{in} = n^{-1} \sum_{k=1}^n \int_0^{F(X_k)} \psi(u)g_{in}(u)\, du + nEX_{i-1,n-1} - (n - 1)EX_{in}. \tag{2.2}$$

PROOF. Let $i \geqq n_0$ and $n - i + 1 \geqq n_0$, where n_0 is such that all of the integrals in this proof are convergent, which is possible by Proposition 2. Note that if $\sigma^2(X_k) < \infty$ we can take $n_0 = 1$, in which case we will interpret $EX_{0,n-1} = 0$.

Now $X_{in} = \min(X_{i,n-1}, X_n) - \min(X_{i-1,n-1}, X_n) + X_{i-1,n-1}$ and $\mathcal{L}(E(X_{in} \mid X_k)) = \mathcal{L}(E(X_{in} \mid X_n))$, $k = 1, 2, \cdots, n - 1$, so to prove the lemma we need only show that $E[\min(X_{i,n-1}, X_n) - \min(X_{i-1,n-1}, X_n) \mid X_n] = n^{-1} \int_0^{F(X_n)} \psi(u)g_{in}(u)\, du$. But

$$\begin{aligned} &E[\min(X_{i,n-1}, X_n) - \min(X_{i-1,n-1}, X_n) \mid X_n) \\ &= \int_0^{F(X_n)} F^{-1}(u)[g_{i,n-1}(u) - g_{i-1,n-1}(u)]\, du + X_n \int_{F(X_n)}^1 [g_{i,n-1}(u) - g_{i-1,n-1}(u)]\, du \\ &= n^{-1} \int_0^{F(X_n)} \psi(u)g_{in}(u)\, du \end{aligned}$$

by integration by parts, since

$$(d/du)g_{in}(u) = n[g_{i-1,n-1}(u) - g_{i,n-1}(u)]. \qquad \text{Q.E.D.}$$

The principal usefulness of Lemma 1 is that it permits us to find a tractable expression for the covariance of the projections of two order statistics. Let us denote

$$(2.3) \qquad K(u, v) = \psi(u)\psi(v)[\min(u, v) - uv] \qquad \text{for} \quad 0 < u, v < 1.$$

Let $g_{in}(u)$ be as defined above (2.1).

LEMMA 2. *Let i, j, and n be integers such that $\hat{X}_{in}$ and $\hat{X}_{jn}$ exist and are given by* (2.2). *Then*

$$(2.4) \qquad n \operatorname{cov}(\hat{X}_{in}, \hat{X}_{jn}) = \int_0^1 \int_0^1 K(u, v) g_{in}(u) g_{jn}(v)\, du\, dv.$$

PROOF. By Lemma 1 and the fact that $X_1, \cdots, X_n$ are independent, identically distributed, we have

$$(2.5) \quad n \operatorname{cov}(\hat{X}_{in}, \hat{X}_{jn}) = \operatorname{cov}\left(\int_0^{F(X_1)} \psi(u) g_{in}(u)\, du, \int_0^{F(X_1)} \psi(v) g_{jn}(v)\, dv\right).$$

Now

$$E\{\int_0^{F(X_1)} \psi(u) g_{in}(u)\, du \cdot \int_0^{F(X_1)} \psi(v) g_{jn}(v)\, dv\}$$
$$= E\{\int_0^1 \int_0^1 \psi(u)\psi(v) g_{in}(u) g_{jn}(v) I_{[\max(u,v) < F(X_1)]}\, du\, dv\}$$
$$= \int_0^1 \int_0^1 \psi(u)\psi(v) g_{in}(u) g_{jn}(v)[1 - \max(u, v)]\, du\, dv$$

by Fubini's theorem, and similarly

$$E\{\int_0^{F(X_1)} \psi(u) g_{in}(u)\, du\} \cdot E\{\int_0^{F(X_1)} \psi(v) g_{jn}(v)\, dv\}$$
$$= \int_0^1 \int_0^1 \psi(u)\psi(v) g_{in}(u) g_{jn}(v)(1 - u)(1 - v)\, du\, dv.$$

Since $1 - \max(u, v) - (1 - u)(1 - v) = \min(u, v) - uv$, these expressions together with (2.5) prove the lemma. Q.E.D.

3. Asymptotic expressions for covariances. In this section we shall derive asymptotic expressions for the covariances of order statistics and the covariances of projections of order statistics. The derivation of these expressions shall make much use of the following lemmas. Let $g_{in}(u)$ be given by (2.1) and $K(u, v)$ by (2.3).

LEMMA 3. *Let $f_{kn}(u) = (n - 1)^k u^{k-1} e^{-(n-1)u}/(k - 1)!$ for $u \geqq 0$. Then for any $\epsilon > 0$ there exists an $M > 0$ depending only on ϵ such that $g_{in}(u) \leqq M f_{in}(u)$ for all $u \geqq 0$, $i \leqq (1 - \epsilon)n$.*

PROOF. It suffices to find a constant $M > 0$ such that $g_{in}(u)/f_{in}(u) \leqq M$ all $0 < u < 1$.

Now $g_{in}(u)/f_{in}(u) = [n!/(n - i)!\,(n - 1)^i](1 - u)^{n-i} e^{(n-1)u}$ is maximized where $(1 - u)^{n-i} e^{(n-1)u}$ is maximized, which upon differentiating we see to be at $u_0 = (i-1)/(n-1)$. Thus we need only bound

$$\begin{aligned} g_{in}(u_0)/f_{in}(u_0) &= [n!/(n-i)!\,(n-1)^i][(n-i)/(n-1)]^{n-i}e^{i-1} \\ &\leqq [n/(n-i)]^{\frac{1}{2}}[n/(n-1)]^n \qquad \text{by Stirling's formula} \\ &\leqq 4\epsilon^{-\frac{1}{2}} \qquad \text{for } n > 1. \end{aligned}$$

Q.E.D.

LEMMA 4. *Let $h(u)$ be a positive function such that for some $k \geqq 0$, $\int_0^1 h(u)[u(1-u)]^k\,du < \infty$. Let b_n be any sequence of integers such that $b_n \to \infty$, $b_n/n \to 0$ as $n \to \infty$. Then for any $m \geqq 0$ there exists $\lambda = \lambda(m, k) > 0$ such that*

$$n^m \int_{B_n^c(i)} h(u)g_{in}(u)\,du \to 0 \tag{3.1}$$

uniformly for $b_n \leqq i \leqq n - b_n$, where

$$B_n(i) = [(i-1)/(n-1) - \lambda\, d_n(n-1)^{-1},\ (i-1)/(n-1) + \lambda\, d_n(n-1)^{-1}],$$

and $$d_n = [\min(i-1, n-i-1)\log n]^{\frac{1}{2}}.$$

PROOF. We will show that (3.1) holds uniformly for $b_n \leqq i \leqq n/2$, the lemma will then follow by symmetry. Let us denote $B_n(i) = [\gamma_n, \gamma_n']$, and let $d_n = [(i-1)\log n]^{\frac{1}{2}}$. Let

$$A_n = n^m \int_0^{\gamma_n} h(u)g_{in}(u)\,du \quad \text{and} \quad A_n' = n^m \int_{\gamma'_n}^1 h(u)g_{in}(u)\,du.$$

Let n be large enough so that $b_n > k + 1$. In what follows, C will be used as a generic constant, independent of i for $b_n \leqq i \leqq n/2$. By Lemma 3 we have

$$A_n \leqq Cn^m \int_0^{\gamma_n} h(u)f_{in}(u)\,du \leqq Cn^{m+k} \int_0^{\gamma_n} h(u)u^k f_{i-k,n}(u)\,du.$$

Since $f_{i-k,n}(u)$ is monotonically increasing for $u < (i-k-1)/(n-1)$, we have that for $\lambda\, d_n > k$,

$$A_n \leqq Cn^{m+k}f_{i-k,n}(\gamma_n) \int_0^{\frac{1}{2}} h(u)u^k\,du \leqq Cn^{m+k}f_{i-k,n}(\gamma_n).$$

From Stirling's formula it follows that $f_{i-k,n}(\gamma_n) \leqq Cn(1 - \lambda\, d_n/(i-1))^{i-1}e^{\lambda d_n}$. Using the inequality that for $0 \leqq a \leqq r$, $(1 - a/r)^r \leqq \exp\{-a - a^2/2r\}$, we see

$$\begin{aligned} A_n &\leqq Cn^{m+k+1}\exp\{-(\lambda d_n)^2/2(i-1)\} \\ &= Cn^{m+k+1}\exp\{-\log n^{(\lambda^2/2)}\} \\ &= Cn^{m+k+1-\lambda^2/2}. \end{aligned}$$

It can be shown through similar arguments that $A_n' \leqq Cn^{m+k+1-\lambda^2/6}$. Thus $n^m \int_{B_n^c(i)} h(u)g_{in}(u)\,du \leqq Cn^{m+k+1-\lambda^2/6}$, and the lemma follows upon taking $\lambda = 3(m+k+2)^{\frac{1}{2}}$. Q.E.D.

We shall also need the following definitions and notation.

CONDITION T. *We say the function $\psi(u)$ satisfies condition T at a point $p \in [0, 1]$ if* (a) *for any $\epsilon > 0$ there exist $\tau > 0$ and $0 < q < 1$ such that for any $u_1, u_2 \in (0, 1)$ satisfying $0 < q(p - u_2) < p - u_1 < p - u_2 \leqq \tau$ or $0 < q(u_2 - p) < u_1 -$*

$p < u_2 - p \leqq \tau$, $|\psi(u_2)/\psi(u_1) - 1| < \epsilon$ *holds, and* (b) *there exist* $\gamma > 0$, $M > 0$, $a \geqq 0$ *such that for* $0 < |p - u| < \gamma$, $u \,\varepsilon\, (0, 1)$, $\psi(u) \geqq M|u - p|^a$ *holds.*

REMARK. Part (a) of this condition says that if $\psi(u)$ tends to infinity or to zero as $u \uparrow p$ or $u \downarrow p$ then it does not oscillate too wildly; roughly speaking it behaves as a power of $|u - p|$. Part (b) says $\psi(u)$ cannot approach zero faster than some power of $|u - p|$ as $u \to p$. In particular, (a) is satisfied if for some $\epsilon > 0$, $\psi(u) = |p - u|^b$ for $0 < |p - u| < \epsilon$ and some finite b, or if $\lim_{u \uparrow p} \psi(u)$ and $\lim_{u \downarrow p} \psi(u)$ exist and are finite and nonzero, or more generally if $\psi(u)$ is regularly varying in the sense of Karamata as $u \uparrow p$ and $u \downarrow p$ (see Cibisov [5] or Feller [6]).

Let us denote by D_1 the set of all discontinuities of ψ, and let

$$D_2 = \{x \,\varepsilon\, [0, 1]: \forall\, m > 0, \epsilon > 0 \{y: \psi(y) \geqq m\} \cap \{y: |x - y| < \epsilon \text{ and } x \neq y\} \neq \emptyset\},$$

$$D_3 = \{x \,\varepsilon\, [0, 1]: \forall\, m > 0, \epsilon > 0 \{y: \psi(y) < m\} \cap \{y: |x - y| < \epsilon \text{ and } x \neq y\} \neq \emptyset\}.$$

For any sets $A, B \subset [0, 1]$ we shall write $A^\epsilon = \{y \,\varepsilon\, [0, 1]: \exists\, x \,\varepsilon\, A \text{ with } |x - y| < \epsilon\}$ and $\rho(A, B) = \inf_{x\varepsilon A, y\varepsilon B} |x - y|$. Let $p_i = i/(n + 1)$, $p_j = j/(n + 1)$.

THEOREM 1. *Let* ϵ *and* M *be any positive numbers, and let* B *be any subset of* $[0, 1]$. *Assume there exists some* $\alpha > 0$ *such that* $\lim_{x\to\infty} |x|^\alpha[1 - F(x) + F(-x)] = 0$. *Then*

(i) *If condition T holds uniformly for all points in* $B^\epsilon \cap (D_2 \cup D_3)$, *and for any* $\delta > 0$, ψ *is of bounded variation on* $B^\epsilon \cap (D_2^\delta)^c$, *then for any sequence of integers* $\{b_n\}$ *such that* $b_n/\log n \to \infty$ *and* $b_n/n \to 0$ *as* $n \to \infty$,

$$n \operatorname{cov}(\hat{X}_{in}, \hat{X}_{jn})/K(p_i, p_j) = 1 + o(1) \tag{3.2}$$

and

$$n \operatorname{cov}(X_{in}, X_{jn})/K(p_i, p_j) = 1 + o(1) \tag{3.3}$$

uniformly for p_i, p_j *in* $B \cap [b_n/n, 1 - b_n/n]$ *such that* $\rho(\{p_i, p_j\}, (D_1 \cup D_3) \cap (0, 1)) \geqq Mn^{-\frac{1}{2}} \log n$.

(ii) *If* $p_i \to p$ *and* $p_j \to q$, $p, q \,\varepsilon\, (0, 1)$, *and* f *satisfies a Hölder condition of order* $s \leqq 1$ *in an* ϵ*-neighborhood of* $F^{-1}(p)$ *and an* ϵ*-neighborhood of* $F^{-1}(q)$, *then* (3.2) *and* (3.3) *hold with* $o(1)$ *replaced by* $O(n^{-s/2})$, *if* $f(F^{-1}(p)) \neq 0$ *and* $f(F^{-1}(q)) \neq 0$.

PROOF. Let n be large enough such that $n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}} < \epsilon$, and for $b_n \leqq i$, $j \leqq n - b_n$, $\operatorname{cov}(\hat{X}_{in}, \hat{X}_{jn})$ exists and is given by Lemma 2 and $\operatorname{cov}(X_{in}, X_{jn})$ exists. Let $A_n = \{p \,\varepsilon\, B \cap [b_n/n, 1 - b_n/n]: \rho(p, (D_1 \cup D_3) \cap (0, 1)) \geqq Mn^{-\frac{1}{2}} \log n\}$. Let $B_n(i, j) = B_n(i) \times B_n(j)$, where $B_n(i)$ is given by Lemma 4.

We shall first prove Theorem 1 for the projections of the order statistics. Let $p_i, p_j \,\varepsilon\, A_n$. Now

$$\sup_{(u,v)\varepsilon B_n(i,j)} \{|u/p_i - 1|, |v/p_j - 1|\} \to 0 \quad \text{uniformly for } p_i, p_j \,\varepsilon\, A_n,$$

and since by Lemma 2

$$n \operatorname{cov}(\hat{X}_{in}, \hat{X}_{jn})/K(p_i, p_j) - 1 = \int_0^1 \int_0^1 [K(u, v)/K(p_i, p_j) - 1] g_{in}(u) g_{jn}(v)\, du\, dv,$$

we will by condition T have proved (i) if we show

$$(3.4)\qquad \int\int_{B_n(i,j)^c} [K(u, v)/K(p_i, p_j) - 1]g_{in}(u)g_{jn}(v)\,du\,dv \to 0$$

uniformly for p_i, $p_j \varepsilon A_n$. But $\int\int_{B_n(i,j)^c} g_{in}(u)g_{jn}(v)\,du\,dv \to 0$ uniformly by Lemma 4, and by Condition T and the hypothesis of the theorem,

$$[K(p_i, p_j)]^{-1} \leqq n^2[\psi(p_i)\psi(p_j)]^{-1}, \qquad [\psi(p_i)]^{-1} \leqq Cn^a,$$

and $K(u, v) \leqq \psi(u)\psi(v)$, so

$$\begin{aligned}\int\int_{B_n(i,j)^c} (K(u, v)/K(p_i, p_j))g_{in}(u)g_{jn}(v)\,du\,dv &\leqq n^2 \int\int_{B_n(i,j)^c} (\psi(u)\psi(v)/\psi(p_i)\psi(p_j))g_{in}(u)g_{jn}(v)\,du\,dv \\ &\leqq Cn^{a+2}[\int_{B_n(i)^c} \psi(u)g_{in}(u)\,du + \int_{B_n(j)^c} \psi(u)g_{jn}(u)\,du]\end{aligned}$$

which tends to zero uniformly by Lemma 4.

We next prove (ii) for projections. Under the given conditions, $|K(u, v) - K(u_0, v_0)| \leqq C[|u - u_0|^s + |v - v_0|^s]$ for some constant C and (u, v) and (u_0, v_0) in some neighborhood of (p, q). It is enough to show $n \operatorname{cov}(X_{in}, X_{jn}) - K(p_i, p_j) = O(n^{-s/2})$. As in the proof of (i), it follows from Lemma 4 that we need only show

$$\int\int_{B_n(i,j)} |K(u, v) - K(p_i, p_j)|g_{in}(u)g_{jn}(v)\,du\,dv = O(n^{-s/2}).$$

But

$$\begin{aligned}\int\int_{B_n(i,j)} |K(u, v) - K(p_i, p_j)|g_{in}(u)g_{jn}(v)\,du\,dv &\leqq C\int\int_{B_n(i,j)} (|u - p_i|^s + |v - p_j|^s)g_{in}(u)g_{jn}(v)\,du\,dv \\ &\leqq C[\int_0^1 (u - p_i)^2 g_{in}(u)\,du]^{s/2} + C[\int_0^1 (v - p_j)^2 g_{jn}(v)\,dv]^{s/2} \\ &= C(p_i(1 - p_i)/(n + 2))^{s/2} + C(p_j(1 - p_j)/(n + 2))^{s/2} \\ &= O(n^{-s/2}),\end{aligned}$$

and (ii) is proved.

It remains to prove Theorem 1 for the order statistics themselves. We recall that $F^{-1}(U_{in})$ has the same distribution as X_{in}, where U_{in} is the ith order statistic of a sample of size n from a uniform [0, 1] distribution, and $(F^{-1}(U_{in}), F^{-1}(U_{jn}))$ has the same joint distribution as (X_{in}, X_{jn}). We shall denote the joint density of (U_{in}, U_{jn}) by, for $i < j$,

$$(3.5)\qquad \begin{aligned} g_{ijn}(u, v) &= [n!/(i - 1)!\,(j - i - 1)!\,(n - j)!]u^{i-1}(v - u)^{j-i-1}(1 - v)^{n-j} \quad \text{for } 0 \leqq u \leqq v \leqq 1 \\ &= 0 \quad \text{otherwise,}\end{aligned}$$

and as before, the density of U_{in} by $g_{in}(u)$. Assume without loss of generality that $i \leqq j$ and $i \leqq n - j + 1$. Let $p_i, p_j \varepsilon A_n$. Now

$$\text{cov}(X_{in}, X_{jn}) = \text{cov}(X_{in} - F^{-1}(p_i), X_{jn} - F^{-1}(p_j)),$$

and we claim that both

$$(3.6) \quad nE[(F^{-1}(U_{in}) - F^{-1}(p_i))(F^{-1}(U_{jn}) - F^{-1}(p_j))I_{B_n(i,j)^c}][K(p_i, p_j)]^{-1}$$

and

$$(3.7) \quad nE[(F^{-1}(U_{in}) - F^{-1}(p_i))I_{B_n(i,j)^c}]E[(F^{-1}(U_{jn}) - F^{-1}(p_j))I_{B_n(i,j)^c}] \cdot [K(p_i, p_j)]^{-1}$$

are $O(n^{-1})$ uniformly for $p_i, p_j \varepsilon A_n$. As noted previously, for $p_i, p_j \varepsilon A_n$, $[K(p_i, p_j)]^{-1} \leqq Cn^{2a+2}$, thus it follows from the Schwarz inequality, Proposition 2 (iii), and Lemma 4 that (3.6) is uniformly $O(n^{-1})$. Similarly, by Proposition 2 (iii) and Lemma 4, (3.7) is uniformly $O(n^{-1})$.

Now for large n, ψ exists and is continuous on $B_n(i, j)$ for $p_i, p_j \varepsilon A_n$, so by the mean value theorem, $F^{-1}(u) - F^{-1}(p_i) = (u - p_i)\psi(\theta_i(u))$, where $\theta_i(u)$ is some point between u and p_i, for $u \varepsilon B_n(i)$. Let us denote $\psi_i(u) = \psi(\theta_i(u))$, and define $\psi_i(p_i) = \psi(p_i)$. We note that on $B_n(i)$, $\psi_i(u)$ satisfies condition T with the same uniformity as does $\psi(u)$, if $p_i \varepsilon A_n$. Also, for $i \leqq j$,

$$(n + 2)E(U_{in} - p_i)(U_{jn} - p_j) = p_i(1 - p_j),$$

and from (3.6) with $F^{-1}(x) = x$, it follows that

$$[p_i(1 - p_j)]^{-1}(n + 2)E[(U_{in} - p_i)(U_{jn} - p_j)I_{B_n(i,j)^c}] = O(n^{-1}).$$

We shall now consider two cases. Let R be a very large, but fixed, number.

Case 1. We first show that (3.3) holds uniformly for $p_i, p_j \varepsilon A_n$, $p_i \leqq p_j$, $j/i \leqq R$. By (3.6) and (3.7) it is enough to show that

$$(3.8) \quad [p_i(1 - p_j)]^{-1}nE\{U_{in} - p_i)(U_{jn} - p_j) \cdot [\psi_i(U_{in})\psi_j(U_{jn})[\psi(p_i)\psi(p_j)]^{-1} - 1]I_{B_n(i,j)}\}$$

and

$$(3.9) \quad [p_i(1 - p_j)]^{-1}nE[(U_{in} - p_i)(\psi_i(U_{in})[\psi(p_i)]^{-1} - 1)I_{B_n(i,j)}] E[(U_{jn} - p_j)(\psi_j(U_{jn})[\psi(p_j)]^{-1} - 1)I_{B_n(i,j)}]$$

tend to zero. But it can be easily seen that (3.8) is less in absolute value than

$$(1 + R)\sup_{(u,v)\varepsilon B_n(i,j)} |\psi_i(u)\psi_j(v)[\psi(p_i)\psi(p_j)]^{-1} - 1|,$$

and (3.9) is less in absolute value than

$$(1 + R)\sup_{(u,v)\varepsilon B_n(i,j)} |\psi_i(u)[\psi(p_i)]^{-1} - 1||\psi_j(v)[\psi(p_j)]^{-1} - 1|,$$

both of which tend to zero uniformly for $p_i, p_j \varepsilon A_n$ as $n \to \infty$. Thus (3.3) holds for Case 1.

Case 2. We will now sketch the proof that (3.3) holds uniformly for $p_i, p_j \varepsilon A_n$, $j/i \geqq R$, if R is chosen large enough. See [18] for details. Let us denote $\Gamma_{ij}(u, v) = [\psi_i(u)\psi_j(v)[\psi(p_i)\psi(p_j)]^{-1} - 1]I_{B_n(i,j)}$, and let $q_i = i/n$, $q_j = (j-1)/n$. By (3.6) and (3.7) we can write

$$(n+2)\operatorname{cov}(X_{in}, X_{jn})[K(p_i, p_j)]^{-1} - 1$$
$$= (n+2)[p_i(1-p_j)]^{-1} \int_0^1\int_0^1 (u-p_i)(v-p_j)\Gamma_{ij}(u, v) \cdot [g_{ijn}(u, v) - g_{in}(u)g_{jn}(v)]\, dv\, dv + o(1).$$

It is clear that we may exchange q_i and q_j for p_i and p_j under the integral signs without adding more than $o(1)$ to the expression. Thus it is sufficient to show

$$(3.10)\quad n[p_i(1-p_j)]^{-1}\int\int (u-q_i)(v-q_j)\Gamma_{ij}(u, v) \cdot [g_{ijn}(u, v)[g_{in}(u)g_{jn}(v)]^{-1} - 1]g_{in}(u)g_{in}(v)\, du\, dv$$

tends to zero uniformly. Let us write

$$(u, v) = ((i+a)/n, (j-1+b)/n),$$

and

$$h_n(a, b) = (1 - (a-b)/(j-i-1))^{j-i-1} \cdot (1 - a/(n-i))^{-(n-i)}(1 + b/(j-1))^{-(j-1)}.$$

Now $j - i \geqq i(R-1) \geqq b_n(R-1)$, so by Stirling's formula we have

$$g_{ijn}(u, v)[g_{in}(u)g_{jn}(v)]^{-1} = ((n-i)(j-i-1)[n(j-1)]^{-1})^{\frac{1}{2}}h_n(a, b)e^{\theta/(j-i)},$$

where $|\theta| \leqq 1/12$. We claim that (3.10) tends to zero uniformly with $g_{ijn}(u, v)[g_{in}(u)g_{jn}(v)]^{-1}$ replaced by $h_n(nu - i, nv - j + 1)$, and that this replacement adds no more than $o(1)$ to the expression. This is accomplished by a close study of the behavior of the function $h_n(a, b)$. In particular, it can be shown that $(u - q_i)(v - q_j)(h_n(a, b) - 1) \geqq 0$ throughout most of the range of integration, and the contribution of the remainder of the range is negligible. Thus (3.10) can be bounded in absolute value by

$$o(1)n[p_i(1-p_j)]^{-1} \cdot \int\int (u-p_i)(v-p_j)[g_{ijn}(u, v) - g_{in}(u)g_{jn}(v)]\, du\, dv + o(1),$$

which tends to zero uniformly for $p_i, p_j \varepsilon A_n$ and $j/i \geqq R$, completing the proof of Case 2, and thus of part (i).

Finally, (ii) follows directly from (3.8) and (3.9) by the Schwarz inequality, and noting that for large enough n, some constant C, and $(u, v) \varepsilon B_n(i, j)$,

$$|\psi_i(u)\psi_j(v)[\psi(p_i)\psi(p_j)]^{-1} - 1| \leqq C[|\theta_i(u) - p_i|^s + |\theta_j(v) - p_j|^s]$$
$$\leqq C[|u - p_i|^s + |v - p_j|^s]. \qquad \text{Q.E.D.}$$

A particular case of Theorem 1 which is sufficient for many applications is the following immediate corollary.

COROLLARY 1.1 *Assume that there exists* $\epsilon > 0$ *such that*

$$\lim_{x\to\infty} |x|^\epsilon[1 - F(x) + F(-x)] = 0,$$

and that the density of $F(x)$, $f(x)$, *is continuous and strictly positive on* $F^{-1}\{(0, 1)\}$. *Then*

(i) *For any* $\delta > 0$,

$$n \operatorname{cov}(\hat{X}_{in}, \hat{X}_{jn})[K(p_i, p_j)]^{-1} = 1 + o(1) \tag{3.11}$$

and

$$n \operatorname{cov}(X_{in}, X_{jn})[K(p_i, p_j)]^{-1} = 1 + o(1), \tag{3.12}$$

uniformly for $p_i, p_j \varepsilon [\delta, 1 - \delta]$.

(ii) *If we also assume condition* T *holds at zero and at one, then for any sequence* b_n *such that* $b_n/n \to 0$ *and* $b_n/\log n \to \infty$, (3.11) *and* (3.12) *hold uniformly for* $p_i, p_j \varepsilon [b_n/n, 1 - b_n/n]$.

An important feature of Theorem 1 is that (3.2) and (3.3) hold uniformly in the stated range. It is this uniformity that allows an asymptotic expression for the variance of $S_n = \sum_{i=1}^n c_{in} X_{in}$. Let

$$A_n = \{p \,\varepsilon\, B \cap [b_n/n, 1 - b_n/n] : \rho(p, (D_1 \cup D_3) \cap (0, 1)) \geqq Mn^{-\frac{1}{2}} \log n\}.$$

Corollary 1.2. *Assume that the conditions of Theorem* 1 (i) *hold and that* $c_{in} = 0$ *for* $i/(n + 1) \not\varepsilon A_n$. *If for some constant* C *independent of* n

$$\textstyle\sum_{i,j=1}^n |c_{in} c_{jn}| K(p_i, p_j) \leqq C \sum_{i,j=1}^n c_{in} c_{jn} K(p_i, p_j) \tag{3.13}$$

for all n, *then*

$$n\sigma^2(S_n)[\textstyle\sum_{i,j=1}^n c_{in} c_{jn} K(p_i, p_j)]^{-1} = 1 + o(1)$$

and

$$n\sigma^2(\hat{S}_n)[\textstyle\sum_{i,j=1}^n c_{in} c_{jn} K(p_i, p_j)]^{-1} = 1 + o(1).$$

Proof. The proof is immediate, since

$$\begin{aligned} |n\sigma^2(S_n)[\textstyle\sum c_{in} c_{jn} K(p_i, p_j)]^{-1} - 1| &= |\textstyle\sum c_{in} c_{jn} n \operatorname{cov}(X_{in}, X_{jn})[\sum c_{in} c_{jn} K(p_i, p_j)]^{-1} - 1| \\ &\leqq o(1) \textstyle\sum |c_{in} c_{jn}| K(p_i, p_j)[\sum c_{in} c_{jn} K(p_i, p_j)]^{-1}. \end{aligned}$$ Q.E.D.

Remark. Theorem 1 gives asymptotic expressions for the covariances of order statistics and the covariances of projections of order statistics as long as p_i and p_j are sufficiently far away from the discontinuities and zeros of $\psi(x) = [f(F^{-1}(x))]^{-1}$ (for example at a distance greater than $n^{-\frac{1}{2}} \log n$), and ψ is sufficiently smooth (in the sense of condition T) at these discontinuities and zeros. To go from these asymptotic expressions to ones for $\sigma^2(S_n)$ and $\sigma^2(\hat{S}_n)$, where the c_{in}'s are zero near the discontinuities and zeros of ψ, it has been necessary to add condition (3.13), which says essentially that the variances of $\sum c_{in} X_{in}$ and $\sum |c_{in}| X_{in}$ are of the same order. This condition would thus exclude statistics such as $T_n = X_{in} - X_{jn}$, where $i - j = o(n)$. However, it is evident from (ii) of Theorem 1 that if ψ satisfies a Hölder condition of order s in a neighborhood of the set of p_i's for which $c_{in} \neq 0$, then the C of (3.13) may be replaced by $o(n^{s/2})$.

4. Asymptotic normality. In proving the asymptotic normality of the statistic $S_n = \sum_{i=1}^n c_{in}X_{in}$, we shall first deal with the case that S_n satisfies the conditions of Corollary 1.2; that is, we shall assume the c_{in}'s are zero for $i/(n+1)$ near to discontinuities and zeros of $\psi(x) = [f(F^{-1}(x))]^{-1}$, and the variance of S_n is of the same order as the variance of $\sum_{i=1}^n |c_{in}|X_{in}$. We will then proceed to extend the theorem to more general statistics by finding conditions under which the $c_{in}X_{in}$ are asymptotically negligible in the sense of mean square for $i/(n+1)$ near the discontinuities and zeros of ψ.

Let the set D_1, D_2, and D_3 be as defined previously,

$$A_n = \{p \varepsilon B \cap [b_n/n, 1 - b_n/n] : \rho(p, (D_1 \cup D_3) \cap (0, 1)) \geqq Mn^{-\frac{1}{2}} \log n\}.$$

We shall always assume that for n large enough, $p_i \varepsilon A_n$ for some i; that is, A_n is not virtually empty.

THEOREM 2. *Assume that the conditions of Theorem* 1 (i) *hold and that* $c_{in} = 0$ *for* $i/(n+1) \not\varepsilon A_n$. *If for some constant C independent of n*

$$\sum_{i,j=1}^n |c_{in}c_{jn}|K(p_i, p_j) \leqq C \sum_{i,j=1}^n c_{in}c_{jn}K(p_i, p_j) \tag{4.1}$$

for all n, then $\mathfrak{L}((S_n - ES_n)/\sigma(S_n))$ *converges to the standard normal distribution as* $n \to \infty$.

PROOF. By Proposition 1, $E(S_n - \hat{S}_n)^2 = \sigma^2(S_n) - \sigma^2(\hat{S}_n)$, and by Corollary 1.2, $[\sigma^2(S_n) - \sigma^2(\hat{S}_n)]/\sigma^2(\hat{S}_n) \to 0$, so we need only prove that $(\hat{S}_n - E\hat{S}_n)/\sigma(\hat{S}_n)$ converges to the standard normal distribution. From lemma 1,

$$\hat{S}_n = n^{-1} \sum_{k=1}^n [\sum_{i=1}^n c_{in} \int_0^{F(X_k)} \psi(u)g_{in}(u)\, du] + \Delta_n,$$

where Δ_n is non-random, and if we write $Z_{kn} = \sum_{i=1}^n c_{in} \int_0^{F(X_k)} \psi(u)g_{in}(u)\, du$, then it will be enough to verify the Lindeberg condition for $\sum_{k=1}^n (Z_{kn} - EZ_{kn})$. Let F_n denote the distribution function of $Z_{kn} - EZ_{kn}$, and s_n^2 the variance of $\sum_{k=1}^n Z_{kn}$. Then we must show that for any $\epsilon > 0$,

$$ns_n^{-2} \int_{[|x| \geqq \epsilon s_n]} x^2\, dF_n(x) \to 0.$$

We will in fact show that $P[|Z_{kn} - EZ_{kn}| \geqq \epsilon s_n] = 0$ for n sufficiently large. Now $s_n^2 = n^2\sigma^2(\hat{S}_n)$, and $Z_{kn} - EZ_{kn} = \sum_{i=1}^n c_{in} \int_0^1 \psi(u)g_{in}(u)(u - I_{[F(X_k)<u]})\, du$, therefore, since

$$|u - I_{[F(X_k)<u]}|\, |v - I_{[F(X_k)<v]}|\, [\min(u, v) - uv]^{-1} \leqq \max\{u^{-1}, v^{-1}, (1-u)^{-1}, (1-v)^{-1}\},$$

$$\begin{aligned}(Z_{kn} - EZ_{kn})^2 &\leqq \sum_{i,j=1}^n |c_{in}c_{jn}| \int_0^1 \int_0^1 \psi(u)\psi(v) \\ &\quad \cdot g_{in}(u)g_{jn}(v)|u - I_{[F(X_k)<u]}|\, |v - I_{[F(X_k)<v]}|\, du\, dv \\ &\leqq \sum_{i,j=1}^n |c_{in}c_{jn}| \int_0^1 \int_0^1 K(u, v) \\ &\quad \cdot g_{in}(u)g_{jn}(v) \max\{u^{-1}, v^{-1}, (1-u)^{-1}, (1-v)^{-1}\}\, du\, dv.\end{aligned}$$

But

$$\int_0^1 \int_0^1 K(u, v) g_{in}(u) g_{jn}(v) \max \{u^{-1}, v^{-1}, (1-u)^{-1}, (1-v)^{-1}\}\, du\, dv$$
$$\cdot [K(p_i, p_j) \max \{p_i^{-1}, p_j^{-1}, (1-p_i)^{-1}, (1-p_j)^{-1}\}]^{-1} = 1 + o(1)$$

uniformly for $p_i, p_j \varepsilon A_n$ by precisely the same proof as that of Theorem 1 (i) for projections, and on A_n, $\max \{p_i^{-1}, (1-p_i)^{-1}\} \leqq n/b_n$, so we have

$$\begin{aligned}(Z_{kn} - EZ_{kn})^2 &\leqq (1 + o(1))(n/b_n) \textstyle\sum_{i,j=1}^n |c_{in}c_{jn}|\, K(p_i, p_j) \\ &\leqq (1 + o(1))(n/b_n) C \textstyle\sum_{i,j=1}^n c_{in}c_{jn} K(p_i, p_j) \\ &\leqq b_n^{-1}(1 + o(1)) C s_n^2\end{aligned}$$

by Corollary 1.2. Since $b_n \to \infty$, we then have $(Z_{kn} - EZ_{kn})^2 \leqq \epsilon^2 s_n^2$ for n large enough. Q.E.D.

REMARK. As in the remark following Corollary 1.2 it is clear from Theorem 1 (ii) that if ψ satisfies a Hölder condition of order s in some ϵ-neighborhood of $\bigcup_{k=1}^\infty A_k$ and $\max_{A_n} \{p_i^{-1}, (1-p_i)^{-1}\} = na_n^{-1}$, then the C of Theorem 2 can be replaced by $\min (o(a_n), o(n^{s/2}))$.

A particular case of the above, useful in many applications, occurs when the c_{in}'s come from a scores generating function J; that is, when $c_{in} = J(i/(n+1))$. Noting that then

$$(n+1)^{-2} \textstyle\sum_{i,j=1}^n c_{in}c_{jn} K(p_i, p_j) \quad \text{and} \quad (n+1)^{-2} \textstyle\sum_{i,j=1}^n |c_{in}c_{jn}|\, K(p_i, p_j)$$

are Riemann sums over the unit square, we have from Corollary 1.2 and Theorem 2 the following immediate corollary. In this corollary and the remainder of the paper, by "possibly improper Riemann integral" we shall mean that the integral exists as a Lebesgue integral and is the limit of the appropriate one of these sums.

COROLLARY 2.1. *Assume that the conditions of Theorem* 1 (i) *hold and that* $c_{in} = 0$ *for* $i/(n+1) \not\varepsilon A_n$. *If* $c_{in} = J(i/(n+1))$ *for* $i/(n+1) \varepsilon A_n$, *and*

$$\int\int_{B\times B} J(u)J(v)K(u, v)\, du\, dv \tag{4.2}$$

and

$$\int\int_{B\times B} |J(u)J(v)| K(u, v)\, du\, dv \tag{4.3}$$

exist as possibly improper Riemann integrals and are finite and non-zero, then $\mathcal{L}((S_n - ES_n)/\sigma(S_n))$ *converges to the standard normal distribution and* $n^{-1}\sigma^2(S_n) \to \int\int_{B\times B} J(u)J(v)K(u, v)\, du\, dv$ *as* $n \to \infty$.

So far we have assumed that the c_{in}'s put no weight on the X_{in}'s for $i/(n+1)$ near (at a distance of order less than $n^{-\frac{1}{2}} \log n$) to discontinuities or zeros of ψ or near (at a distance less than $n^{-1}b_n$) to zero or one. We now turn to the problem of finding conditions under which the c_{in}'s are allowed to put weight on such X_{in}'s, but such that they remain negligible with respect to the rest of the statistic. Let $C_n = A_n^c \cap B \cap [b_n/n, 1 - b_n/n]$.

THEOREM 3. *Let* B *be a subset of* [0,1], *and assume that for some* $\alpha > 0$, $\lim_{x\to\infty} x^{\alpha}[1 - F(x) + F(-x)] = 0$. *Assume also that for some* $\epsilon > 0$, F^{-1} *satisfies a Lipschitz condition on* B^{ϵ}, *and for some* $\tau > 0$, $n^{-1}\sigma^2(\sum_{p_i \varepsilon A_n} c_{in}X_{in}) \geqq \tau$ *for n sufficiently large.*

(i) *If* $n^{-1}\sum_{p_i\varepsilon C_n} |c_{in}| \to 0$ *as* $n \to \infty$, *then*

$$\sigma^2(\textstyle\sum_{p_i\varepsilon C_n} c_{in}X_{in})/\sigma^2(\sum_{p_i\varepsilon A_n} c_{in}X_{in}) \to 0. \tag{4.4}$$

(ii) *In particular, if the Lebesgue measure of* $A_n^c \cap B$ *converges to zero as* $n \to \infty$ *and* $c_{in} = J(i/(n+1))$, *where the possibly improper Riemann integral* $\int_B |J(u)|\,du$ *exists and is finite, then* (4.4) *holds.*

PROOF. We wish to prove that

$$n^{-1}\sigma^2(\textstyle\sum_{p_i\varepsilon C_n} c_{in}X_{in}) = n^{-2}\sum\sum_{p_i,p_j\varepsilon C_n} c_{in}c_{jn}n \operatorname{cov}(X_{in}, X_{jn}) \to 0.$$

It is clearly enough to show that $n \operatorname{cov}(X_{in}, X_{jn})$ is uniformly bounded for $p_i, p_j \varepsilon C_n$. Let $B_n(i,j)$ be as defined in the proof of Theorem 1. As in that proof we have

$$\begin{aligned} nE(X_{in} - F^{-1}(p_i))(X_{jn} - F^{-1}(p_j)) \\ &= nE[(X_{in} - F^{-1}(p_i))(X_{jn} - F^{-1}(p_j))I_{B_n(i,j)}] + o(1) \\ &\leqq nE|U_{in} - p_i|\cdot|U_{jn} - p_j|\cdot C^2 + o(1) \leqq C^2 + o(1), \end{aligned}$$

for some constant C, by the Lipschitz condition. Similarly, $n^{\frac{1}{2}}E(X_{in} - F^{-1}(p_i)) \leqq C + o(1)$, and $\operatorname{cov}(X_{in}, X_{jn}) = \operatorname{cov}(X_{in} - F^{-1}(p_i), X_{jn} - F^{-1}(p_i)) \leqq 2n^{-1}C^2 + o(n^{-1})$ uniformly for p_i, $p_j \varepsilon C_n$. Thus (i) is proved, and (ii) is immediate. Q.E.D.

THEOREM 4. *Suppose that for some* C *and* $\delta > 0$,

$$|F^{-1}(u)| \leqq C[u(1-u)]^{-\frac{1}{2}+\delta} \qquad \text{for} \quad 0 < u < 1.$$

(i) *If* $n^{-\frac{1}{2}}\sum_{i=1}^{b_n} |c_{in}|\, p_i^{-\frac{1}{2}+\delta} \to 0$ *and* $n^{-\frac{1}{2}}\sum_{i=n-b_n}^{n} |c_{in}|(1-p_i)^{-\frac{1}{2}+\delta} \to 0$, *then*

$$n^{-1}[\sigma^2(\textstyle\sum_{i=1}^{b_n} c_{in}X_{in}) + \sigma^2(\sum_{i=n-b_n}^{n} c_{in}X_{in})] \to 0. \tag{4.5}$$

(ii) *For any* $C, \epsilon > 0$ *a sequence* $\{b_n\}$ *of integers with* $b_n/\log n \to \infty$ *and* $b_n/n \to 0$ *can be found such that if* $|c_{in}| \leqq C[p_i(1-p_i)]^{\epsilon-\delta}$ *for* $i \leqq b_n$ *and* $i \geqq n - b_n$, *then* (4.5) *holds.*

PROOF. We prove that $n^{-1}\sigma^2(\sum_{i=1}^{b_n} c_{in}X_{in}) \to 0$, the other tail will follow by symmetry. We claim that $|F^{-1}(u)| \leqq C[u(1-u)]^{-\frac{1}{2}+\delta}$ implies that $\operatorname{cov}(X_{in}, X_{jn}) \leqq C_1(p_ip_j)^{-\frac{1}{2}+\delta}$ for $i, j \leqq b_n$, C_1 a constant. By the Schwarz inequality it will be sufficient to show $EX_{in}^2 \leqq C_1p_i^{-1+2\delta}$.

$$\begin{aligned} EX_{in}^2 &= \textstyle\int_0^1 [F^{-1}(u)]^2 g_{in}(u)\,du \\ &\leqq C^2 \textstyle\int_0^1 [u(1-u)]^{-1+2\delta} g_{in}(u)\,du \\ &= C^2 n\binom{n-1}{i-1}\Gamma(i-1+2\delta)\Gamma(n-i+2\delta)/\Gamma(n-1+4\delta). \end{aligned}$$

Then by Stirling's formula, for some constant C_1, $EX_{in}^2 \leqq C_1(p_i)^{2\delta-1}$. Thus

$$\begin{aligned} n^{-1}\sigma^2(\sum_{i=1}^{b_n} c_{in}X_{in}) &= n^{-1}\sum_{i,j=1}^{b_n} c_{in}c_{jn}\,\mathrm{cov}\,(X_{in}, X_{jn}) \\ &\leqq n^{-1}C_1\sum_{i,j=1}^{b_n} |c_{in}c_{jn}|\,(p_ip_j)^{-\frac{1}{2}+\delta} \\ &= C_1(n^{-\frac{1}{2}}\sum_{i=1}^{b_n}|c_{in}|\,p_i^{-\frac{1}{2}+\delta})^2, \end{aligned}$$

proving (i). In the case of (ii) we have

$$\begin{aligned} n^{-\frac{1}{2}}\sum_{i=1}^{b_n}|c_{in}|\,p_i^{-\frac{1}{2}+\delta} &= n^{-\frac{1}{2}}C\sum_{i=1}^{b_n} p_i^{-\frac{1}{2}+\epsilon} \\ &\leqq n^{-\frac{1}{2}}C\int_0^{b_n/n} u^{-\frac{1}{2}+\epsilon}\,du \\ &= Cn^{-\epsilon}b_n^{\frac{1}{2}+\epsilon}. \end{aligned}$$

Taking $b_n = [\log n]^2$ we are finished. Q.E.D.

REMARK. In both Theorems 3 and 4 the additional restrictions on F are to a certain degree necessary. The Lipschitz condition in Theorem 3 says essentially that the density f of F is bounded away from zero on the interior of the support of F. The necessity of a condition of this type can be seen by considering $f(x) = 1$ for $x \,\varepsilon\, [-1.5, -1] \cup [1, 1.5]$, and letting μ_n be the sample median, ν_n any other percentile, and $S_n = n^{3/4}\mu_n + \nu_n$. Then the conclusion (4.4) does not hold.

The restriction of F^{-1} in Theorem 4 is very close to the assumption $\sigma^2(X_i) < \infty$ (or alternatively $\int_0^1 |F^{-1}(x)|^2\,dx < \infty$), which would be necessary in any case since we are dealing with mean square convergence. The condition $|c_{in}| \leqq [p_i(1-p_i)]^{\epsilon-\delta}$ is also almost necessary, for if we have $F^{-1}(x) = (1-x)^{-\frac{1}{2}+\delta}$ and $c_{nn} = n^{\delta}$, then it can easily be shown that $n^{-1}\sigma^2(c_{nn}X_{nn})$ tends to $\Gamma(2\delta) - \Gamma^2(\frac{1}{2}+\delta) > 0$.

Theorems 2, 3, and 4 together give quite general conditions under which the statistic S_n has limiting normal distribution. A special case, which is sufficient for most applications, occurs when ψ is continuous and positive on $(0, 1)$.

COROLLARY 4.1. *Suppose that F possesses a continuous density f which is strictly positive on $F^{-1}\{(0, 1)\}$, and that $[f(F^{-1}(u))]^{-1}$ satisfies condition T at zero and one. Assume that for some $\epsilon > 0$, $\delta > 0$, $C > 0$, and sequence of integers b_n with $b_n/\log n \to \infty$ and $b_n/n \to 0$ as $n \to \infty$ the following conditions are satisfied for all n.*

(a) $|F^{-1}(u)| \leqq C[u(1-u)]^{-\frac{1}{2}+\delta}$ *for* $0 < u < 1$.

(b) $|c_{in}| \leqq C[p_i(1-p_i)]^{-\delta+\epsilon}$ *for* $i \leqq b_n$ *or* $n - i \leqq b_n$.

(c) $\sum_{i,j=b_n}^{n-b_n} |c_{in}c_{jn}|\,K(p_i, p_j) \leqq C\sum_{i,j=b_n}^{n-b_n} c_{in}c_{jn}K(p_i, p_j)$.

Then $\mathcal{L}((S_n - ES_n)/\sigma(S_n))$ converges to the standard normal distribution as $n \to \infty$.

If the weights are given by $c_{in} = J(i/(n+1))$, then (b) *and* (c) *can be replaced by*

(b′) $|J(u)| \leqq C[u(1-u)]^{-\delta+\epsilon}$ for $0 < u < 1$.

(c′) $\int_0^1\int_0^1 J(u)J(v)K(u, v)\,du\,dv$ *and* $\int_0^1\int_0^1 |J(u)J(v)|\,K(u, v)\,du\,dv$

exist as possibly improper Riemann integrals and are both finite and non-zero. Then also $n^{-1}\sigma^2(S_n) \to \int_0^1 \int_0^1 J(u)J(v)K(u, v)\, du\, dv$.

5. Extensions and limitations. In [3], [7], and [17], a more general class of statistics than S_n is considered, namely the statistics of the form

$$T_n = \sum_{i=1}^n c_{in} h(X_{in}),$$

where h is some known function defined on the support of F. Despite this apparent greater generality, our Theorems 1–4 and their corollaries apply equally well to T_n, with only minor changes. The condition that there exists some $\epsilon > 0$ such that $\lim_{x\to\infty} x^\epsilon[1 - F(x) + F(-x)] = 0$ should be replaced by the condition that for some $\epsilon > 0$, $\lim_{x\to-\infty} |h(x)|^\epsilon F(x) = 0$ and $\lim_{x\to\infty} |h(x)|^\epsilon [1 - F(x)] = 0$, this then implies the eventual existence of $E\,|h(X_{in})|^2$ for $b_n \leqq i \leqq n - b_n$, but is no longer equivalent to it in general. If we let $\phi(x) = (d/dx)h(F^{-1}(x))$ when it exists, the results remain true with ψ replaced by $|\phi|$ in (b) of condition T, the definitions of D_2 and D_3, the function K in the left hand side of (4.1) and in (4.3), ψ replaced by ϕ otherwise, and the obvious modifications of the statements of (ii) of Theorem 1 and Corollaries 1.1 and 4.1. (Corollaries 1.1 and 4.1 actually do not generalize effectively, since ϕ continuous and strictly positive on $F^{-1}\{(0, 1)\}$ implies that h is strictly monotone, and only serves to change the distribution function.)

In another direction of possible extension, it might be conjectured that the approximation of the projection $\hat{S}_n$ to S_n is effective in the sense of Proposition 2 whenever S_n is in fact asymptotically normal, at least under suitable regularity conditions on F. This is in fact not true even in the simple case $F(x) = x$. It is well known that if F is the uniform [0, 1] distribution, X_{kn} has a limiting normal distribution as long as $k \to \infty$ and $n - k \to \infty$. Also, for $i > j$, $X_{in} - X_{jk}$ has the same distribution as $X_{i-j,n}$.

We will show that while in the first case the projection of X_{kn} is always effective if $k \to \infty$ and $n - k \to \infty$; in the latter case if $i/n \to p, j/n \to p, 0 < p < 1$, the projection of $X_{in} - X_{jn}$ is effective only if $(i - j)n^{-\frac{1}{2}} \to \infty$. Theorem 1 is insufficient to prove these facts, for while part (ii) shows that $\hat{X}_{in} - \hat{X}_{jn}$ is effective if $(i - j)n^{-\frac{1}{2}} \to \infty$, part (i) shows only that $\hat{X}_{kn}$ is effective when k and $n - k$ tend to infinity more rapidly than $\log n$.

We need the following lemma.

LEMMA 5. *If* $i \to \infty$ *as* $n \to \infty$, *such that* $(n - i)/n$ *stays bounded away from* 0, *then*

$$1 - [2(n - i) + 1]^{-1} \sum_{k,l=i}^n \binom{k+l}{k}\binom{2n-k-l}{n-k}\Big/\binom{2n}{n} = o(i/n).$$

This holds uniformly for $i \geqq b_n$, *where* b_n *is an arbitrary sequence of integers tending to infinity.*

Lemma 5 is proved by rewriting the summation as a sum of hypergeometric probabilities and using the inequality $P(|Z| \geqq a) \leqq a^{-4}E\,|Z|^4$. See [18] for details.

We can now prove

THEOREM 5. *Let* b_n *be any sequence of integers tending to infinity as* $n \to \infty$, *such* *that* $n^{-1}b_n \to 0$. *Then in the uniform case* $F(x) = x$, the projection $\hat{X}_{in}$ *is an effective approximation of* X_{in} *in the sense that* $E(X_{in} - \hat{X}_{in})^2/\sigma^2(X_{in}) \to 0$ *as* $n \to \infty$, *uniformly for* $b_n \leqq i \leqq n - b_n$.

PROOF. By Proposition 1, $E(X_{in} - \hat{X}_{in})^2 = \sigma^2(X_{in}) - \sigma^2(\hat{X}_{in})$, and from Lemma 1, after using a well-known identity for the incomplete beta function, we have

$$\hat{X}_{in} = n^{-1} \sum_{j=1}^{n} \sum_{k=i}^{n} \binom{n}{k} X_j^k (1 - X_j)^{n-k} + (2i - n - 1)/(n + 1).$$

Since

$$E\left(\sum_{k=i}^{n} \binom{n}{k} X_j^k (1 - X_j)^{n-k}\right)^2 = (2n + 1)^{-1} \sum_{k,l=i}^{n} \binom{k+l}{k}\binom{2n-k-l}{n-k}\Big/\binom{2n}{n},$$

and $\sigma^2(X_{in}) = i(n - i + 1)/[(n + 1)^2(n + 2)]$, a little algebra gives us

$$nE(X_{in} - \hat{X}_{in})^2 = ((n - i + 1)/(n + 1))(1 - (2/(n + 2))(i/(n + 1))) - (2n + 1)^{-1} \sum_{k,l=i}^{n} \binom{k+l}{k}\binom{2n-k-l}{n-k}\Big/\binom{2n}{n},$$

and the theorem follows from Lemma 5. Q.E.D.

We now show that although $X_{in} - X_{jn}$ has the same distribution as $X_{i-j,n}$, and therefore we could infer its asymptotic normality as long as $i - j \to \infty$, the projection is not effective if $i - j = o(n^{\frac{1}{2}})$.

THEOREM 6. *Let* $F(x) = x$, $i > j$, *and* $i/n \to p$, $j/n \to p$ *as* $n \to \infty$, *where* $0 < p < 1$. *Then if* $i - j = o(n^{\frac{1}{2}})$,

$$\sigma^2(\hat{X}_{in} - \hat{X}_{jn})/\sigma^2(X_{in} - X_{jn}) \to 0.$$

SKETCH OF PROOF. We have

$$\sigma^2(X_{in} - X_{jn}) = \sigma^2(X_{i-j,n}) = (i - j)(n - i + j + 1)(n + 1)^{-2}(n + 2)^{-1}.$$

Thus it is sufficient to show that $n^2(i - j)^{-1}\sigma^2(\hat{X}_{in} - \hat{X}_{jn}) \to 0$.

By Lemma 2 we can write $n^2(i - j)^{-1}\sigma^2(\hat{X}_{in} - \hat{X}_{jn}) = A_{1n} - A_{2n}$, where

$$A_{1n} = n(i - j)^{-1} \int_0^1 \int_0^1 K(u, v)(g_{in}(u) - g_{jn}(u))g_{in}(v)\, du\, dv,$$

$$A_{2n} = n(i - j)^{-1} \int_0^1 \int_0^1 K(u, v)(g_{in}(u) - g_{jn}(u))g_{jn}(v)\, du\, dv,$$

and since $\psi(x) = 1$ in this case, $K(u, v) = \min(u, v) - uv$. Now

$$\begin{aligned} A_{1n} &= n(i - j)^{-1} \int_0^1 \int_0^1 [K(u, v) - K(p_i, v)] \\ &\quad \cdot (g_{in}(u) - g_{jn}(u))g_{in}(v)\, du\, dv \\ &= n(i - j)^{-1} \int \int_{B_n(i,j)} [K(u, v) - K(p_i, v)] \\ &\quad \cdot (g_{in}(u) - g_{jn}(u))g_{in}(v)\, du\, dv + o(1) \end{aligned}$$

by Lemma 4, and making a change of variable $u = n^{-\frac{1}{2}}x + p_i$, $v = n^{-\frac{1}{2}}y + p_i$, and letting $g_{in}^*(x) = n^{-\frac{1}{2}}g_{in}(u)$, $g_{jn}^*(y) = n^{-\frac{1}{2}}g_{jn}(n^{-\frac{1}{2}}y + p_j)$, the first term becomes

$$\int\int_{B_n} n^{\frac{1}{2}}[K(n^{-\frac{1}{2}}x + p_i, n^{-\frac{1}{2}}y + p_i) - K(p_i, n^{-\frac{1}{2}}y + p_i)]$$
$$\cdot (i-j)^{-1} n^{\frac{1}{2}}[g_{in}^*(x) - g_{jn}^*(x + (i-j)n^{\frac{1}{2}}/(n+1))] g_{in}^*(y)\, dx\, dy,$$

where B_n is the transformed range of integration. It can then be shown in a straightforward manner that as $n \to \infty$, on B_n $n^{\frac{1}{2}}[K(n^{-\frac{1}{2}}x + p_i, n^{-\frac{1}{2}}y + p_i) - K(p_i, n^{-\frac{1}{2}}y + p_i)]$ converges to a nice integrable limit, say $M(x, y)$. It can also be shown that as long as $(i - j)n^{-\frac{1}{2}} \to 0$,

$$(i-j)^{-1} n^{\frac{1}{2}}[g_{in}^*(x) - g_{jn}^*(x + (i-j)n^{\frac{1}{2}}/(n+1))] \to (2\pi)^{-\frac{1}{2}}\sigma^{-3} x \exp\{-x^2/(2\sigma^2)\}$$

where $\sigma^2 = p(1-p)$. Lemma 2.2 of Bickel [1] then allows us to use the dominated convergence theorem to say $A_{1n} \to A(p)$, where the function $A(p)$ is the integral of the limits $M(x, y)$ and $(2\pi)^{-\frac{1}{2}}\sigma^{-3}x \exp\{-x/(2\sigma^2)\}$. In precisely the same manner it can be seen that $A_{2n} \to A(p)$, and since $A(p)$ is finite for $0 < p < 1$, the theorem is proved. Q.E.D.

Remark. It can also be shown that the projection of $X_{in} - X_{jn}$ is not effective when $i - j = O(n^{\frac{1}{2}})$.

6. Acknowledgments. I wish to thank Professor Lucien LeCam for his advice and encouragement in the course of this investigation, and Professor Jaroslav Hájek for a number of helpful conversations. I also thank a referee for suggesting a shorter proof of Lemma 1.

REFERENCES

[1] Bickel, P. J. (1965). Some contributions to the theory of order statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 575–591.

[2] Blom, G. (1958). *Statistical Estimates and Transformed Beta-variables.* Wiley, New York.

[3] Chernoff, H., Gastwirth, J. L. and Johns, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Ann. Math. Statist.* **38** 52–72.

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

[5] Cibisov, D. M. (1964). Some theorems on the limiting behavior of the empirical distribution function. Trudy Mat. Inst. Steklov **71** 104–112. In Russian. (Translated in *Selected Translations in Math. Statistics and Probability*; Vol. 6, 1966.)

[6] Feller, W. (1966). *An Introduction to Probability Theory and Its Applications* **2.** Wiley, New York.

[7] Govindarajulu, Z. (1965). Asymptotic normality of linear functions of order statistics in one and multi-samples. Unpublished.

[8] Govindarajulu, Z., LeCam, L. and Raghavachari, M. (1965). Generalizations of theorems of Chernoff and Savage on the asymptotic normality of test statistics. *Proc. Fifth Berkeley Symp. Math. Statist.* **1** 609–638.

[9] Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.* **39** 325–346.

[10] Hardy, G. H., Littlewood, J. E. and Polya, G. (1964). *Inequalities.* University Press, Cambridge.

[11] JUNG, J. (1955). On linear estimates defined by a continuous weight function. *Arkiv Für Matematik* **3** 199–209.

[12] MOORE, D. S. (1968). An elementary proof of asymptotic normality of linear functions of order statistics. *Ann. Math. Statist.* **39** 263–265.

[13] MOSTELLER, F. C. (1946). On some useful "inefficient" statistics. *Ann. Math. Statist.* **17** 377–408.

[14] RÉNYI, A. (1953). On the theory of order statistics. *Acta Math. Acad. Sci. Hungar.* **4** 191–232.

[15] SARHAN, A. E. and GREENBERG, B. G. (1962). *Contributions to Order Statistics.* Wiley, New York.

[16] SEN, P. K. (1959). On the moments of the sample quantiles. *Calcutta Statist. Assoc. Bull.* No. 33–34, 1–21.

[17] SHORACK, G. R. (1968). Asymptotic normality of linear combinations of functions of order statistics. Unpublished.

[18] STIGLER, S. M. (1967). Linear functions of order statistics. Unpublished dissertation.

[19] SMIRNOV, N. (1935). Über die Verteiling des allgemeinen Gliedes in der Variationsreihe. *Metron* **12** 59–81.

[20] VAN ZWET, W. R. (1964). *Convex Transformations of Random Variables.* Mathematical Centre Tracts No. 7., Mathematisch Centrum, Amsterdam.

[21] WEISS, L. (1963). On the asymptotic distribution of an estimate of a scale parameter. *Naval Res. Logistics Quart.* **10** 1–9.