The Annals of Statistics
1974, Vol. 2, No. 4, 676–693

LINEAR FUNCTIONS OF ORDER STATISTICS WITH SMOOTH WEIGHT FUNCTIONS[1]

BY STEPHEN M. STIGLER

University of Wisconsin, Madison

This paper considers linear functions of order statistics of the form $S_n = n^{-1} \sum J(i/(n+1))X_{(i)}$. The main results are that S_n is asymptotically normal if the second moment of the population is finite and J is bounded and continuous a.e. F^{-1}, and that this first result continues to hold even if the unordered observations are not identically distributed. The moment condition can be discarded if J trims the extremes. In addition, asymptotic formulas for the mean and variance of S_n are given for both the identically and non-identically distributed cases. All of the theorems of this paper apply to discrete populations, continuous populations, and grouped data, and the conditions on J are easily checked (and are satisfied by most robust statistics of the form S_n). Finally, a number of applications are given, including the trimmed mean and Gini's mean difference, and an example is presented which shows that S_n may not be asymptotically normal if J is discontinuous.

1. Introduction. The class of statistics which are linear functions of order statistics, which appears to have been first studied extensively by Percy Daniell in 1920 (see Stigler, 1973b), has received considerable attention in recent years, including much work looking into conditions under which such statistics are asymptotically normal. A principal motivation for this research has been that linear functions of order statistics such as $S_n = n^{-1} \sum c_{in} X_{(i)}$ often exhibit desirable robustness, particularly to heavy-tailed distributions or outlying observations, and these statistics are fairly easy to calculate. However, despite the fact that the very consideration of the robustness of a statistic implies an imprecise knowledge of the underlying population distribution, most of the published work on this problem has put quite severe restrictions on the population distribution function (virtually requiring the existence of a continuous density with smooth tails.) The strongest results in this vein are those given by Chernoff, Gastwirth, and Johns (1967), by Stigler (1969), and by Shorack (1969). The results of these three papers are of approximately equal strength, despite the use of three completely different methods of proof.

The result of Stigler (1969), as given by Theorems 2 and 4 there, can be summarized as follows: S_n is asymptotically normal $(ES_n, \sigma^2(S_n))$ if (1) the extremal order statistics do not contribute too much to S_n, (2) the tails of the population

Received February 1972; revised August 1973.

AMS 1970 *subject classifications.* Primary 62G30; Secondary 60F05, 62E20, 62G35.

Key words and phrases. Order statistics, robust estimation, moments of order statistics, trimmed means.

[1] This research was supported in part by the Wisconsin Alumni Research Foundation; the work was completed while the author was a visitor at the University of Wisconsin Mathematics Research Center, sponsored by the United States Army under Contract No.: DA-31-124-Aro-D-462.

distribution are smooth and the population density is continuous and positive over its support, and (3) the variance of S_n is of the same order as the variance of $n^{-1} \sum |c_{in}| X_{(i)}$. It is the purpose of this present paper to show that if the very weak condition (3) (which, for example, permits *any* system of positive weights) is replaced by the stronger condition that the weights are given by $c_{in} = J(i/(n+1))$, where J is a smooth bounded function, then (2) can be dispensed with entirely, and (1) is automatically satisfied. What is more, we shall see that this result continues to hold when we drop the assumption that the original (unordered) observations are identically distributed, retaining only their independence.

The advantage of such a result lies in its wide applicability. The conditions are remarkably easy to check, and the theorems apply to almost every robust statistic with a smooth weight function. In particular, the conditions permit samples from discrete populations and the use of grouped data, and a wide variety of nonidentically distributed observations. This lack of restriction on the population distribution is particularly appealing from the point of view of robustness. The force of the conditions is taken away from the population distribution, where verification is difficult, and placed on the weight function, where verification is easy.

In addition to results on asymptotic normality, we give here asymptotic formulas for the mean and variance of S_n which, for the first time, are shown valid for fairly arbitrary distributions. The key to the asymptotic treatment of $\sigma^2(S_n)$ is Proposition 5, which permits a fairly simple proof for distributions without densities or non-identically distributed variables. These results should permit the evaluation of the loss of efficiency due to grouping for the statistics S_n (see David and Mishriky (1968), for example).

In Section 2, a number of results (some known) which are needed in subsequent proofs are collected, and the paper's notation introduced. Section 3 treats the independent identically distributed case. The asymptotic normality results of this section overlap those of Shorack (1972), where a stochastic process approach is used. Shorack permits some unbounded weight functions, while we do not, but his moment condition is slightly stronger than ours for some of the statistics we consider. Also, it is doubtful that unbounded weight functions will prove useful in robust inference (because of the resulting sensitivity of S_n to outliers). This section also contains proofs of asymptotic formulas for the mean and variance of S_n; Theorem 4 gives conditions under which S_n is asymptotically normal about its asymptotic mean. Section 4 extends all of the results of Section 3 to the case where the observations are independent but not identically distributed. Finally, Section 5 presents a number of applications and examples; the final example (5.6) shows why the theorem of Moore (1968) is false as stated, why a valid theorem must require that J and F^{-1} possess no common discontinuities.

2. Preliminaries.

2.1. *Notation.* The most general situation we shall consider will be described by the following notation. For each $n \geqq 1$, let $X_{1n}, X_{2n}, \cdots, X_{nn}$ be n independent random variables with (possibly different) cumulative distribution functions F_{1n}, $F_{2n}, \cdots, F_{nn}$, where $F_{in}(x) = P\{X_{in} \leqq x\}$. For the present, nothing at all is assumed about the $\{F_{in}\}$: they may be discrete, continuous, or any combination of the two. Let $X_{(1)} \leqq X_{(2)} \leqq \cdots \leqq X_{(n)}$ be the order statistics of the sample $X_{1n}, \cdots, X_{nn}$. We shall be concerned with statistics of the form

$$S_n = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) X_{(i)}, \tag{1}$$

where J is a suitably chosen weight function. We note in passing that the dependence of each $X_{(i)}$ on n has been suppressed for convenience, and that this notation differs slightly from that of Stigler (1969) where the X_{in}'s represented order statistics. In the special case that the X_{in}'s are identically distributed for all i, n, we shall denote their common distribution function by F. Whenever an inverse of a cumulative distribution function appears, it may be taken as the left continuous version.

2.2. *Some useful results.* In this section we shall present a number of results which will provide the basic tools for the proofs that follow. This section may be skimmed and referred back to when necessary.

The point of departure for our attack will be the same as that used in Stigler (1969); that is, we shall use Hájek's projection lemma to find a sum of independent random variables which well approximates S_n in mean square and show that this sum and S_n are mean square equivalent. For completeness, we restate this lemma as Proposition 1.

PROPOSITION 1. (Hájek (1968)). *Let $Z_1, Z_2, \cdots, Z_n$ be independent random variables and $\mathscr{T}$ be the Hilbert space of* a.s. *equivalence classes of square integrable statistics depending on $Z_1, \cdots, Z_n$. Let $\mathscr{L}$ be the closed linear subspace of $\mathscr{T}$ consisting of statistics of the form $L = \sum_{i=1}^{n} l_i(Z_i)$, where the l_i are functions such that $El_i^2(Z_i) < \infty$. Then if $S \in \mathscr{T}$, the projection of S on $\mathscr{L}$ is given by*

$$\hat{S} = \sum_{i=1}^{n} E(S \mid Z_i) - (n-1)ES. \tag{2}$$

Thus

$$ES = E\hat{S} \tag{3}$$

and

$$E(S - \hat{S})^2 = \sigma^2(S) - \sigma^2(\hat{S}). \tag{4}$$

Our approach in this paper differs from that used in Stigler (1969) in that here the proof of mean square equivalence of the statistic and its projection will be accomplished directly through asymptotic expressions for their variances, rather than through the covariances of the individual order statistics as in the previous paper.

As Proposition 1 will require that the second moment of S_n be finite, there will be cases where the following proposition (given by Bickel (1967) for the continuous identically distributed case) will be useful.

PROPOSITION 2. *Suppose that for some* $\varepsilon > 0$,

$$x^{\varepsilon}[1 - F_{in}(x) + F_{in}(-x)] \to 0 \quad \text{as} \quad x \to \infty\,, \quad \text{uniformly in } i \text{ and } n\,.$$

Then for any finite $k > 0$ *there exists a finite* r *(depending only on* k *and* ε*) such that* $E|X_{(i)}|^k < \infty$ *for* $r \leqq i \leqq n - r$.

PROOF. Integration by parts gives

$$EX_{(i)}^k = k \int_0^\infty x^{k-1} P(X_{(i)} > x)\, dx - k \int_{-\infty}^0 x^{k-1} P(X_{(i)} \leqq x)\, dx\,.$$

Now for any fixed x,

$$P(X_{(i)} > x) \leqq \sum_{j=n-i+1}^n \binom{n}{j} (1 - G_n(x))^j 1^{n-j}$$

where $G_n(x) = \min\{F_{sn}(x) : 1 \leqq s \leqq n\}$. Therefore

$$\int_0^\infty x^{k-1} P(X_{(i)} > x)\, dx \leqq \sum_{j=n-i+1}^n \binom{n}{j} \int_0^\infty x^{k-1}(1 - G_n(x))^j\, dx$$

which is finite, at least for $i \leqq n - k/\varepsilon$, by the condition of the proposition. Similar reasoning applied to the second integral in the above expression for $EX_{(i)}^k$ completes the proof. □

In order to describe the projections $\hat{X}_{(i)}$ of the order statistics and their covariances, we shall need some further notation. For any fixed n, let $P_{ik}^n(y) = P(\text{exactly } i - 1 \text{ of } X_{1n}, \cdots, X_{k-1,n}, X_{k+1,n}, \cdots, X_{nn} \text{ are} \leqq y)$ be defined for $i = 2, 3, \cdots, n - 1$.

PROPOSITION 3. *Suppose* $E(X_{(i)}^2 + X_{(j)}^2) < \infty$, *where* $1 < i, j < n$. *Then the projection of* $X_{(i)}$ *(as defined by* (2)*) is given by*

$$\hat{X}_{(i)} = \sum_{k=1}^n \int_{-\infty}^\infty I_{[y < X_{kn}]} P_{ik}^n(y)\, dy + h \tag{5}$$

where h *is a constant (depending on* i, n, *and* $\{F_{jn}\}$*), and*

$$\begin{aligned} &\operatorname{Cov}(\hat{X}_{(i)}, \hat{X}_{(j)}) \\ &\quad = \sum_{k=1}^n \int_{-\infty}^\infty \int_{-\infty}^\infty [F_{kn}(\min(y, z)) - F_{kn}(y)F_{kn}(z)] P_{ik}^n(y) P_{jk}^n(z)\, dy\, dz\,. \end{aligned} \tag{6}$$

PROOF. For any fixed n, let $Y_{(1)} \leqq \cdots \leqq Y_{(n-1)}$ be the order statistics of X_{1n}, $X_{2n}, \cdots, X_{n-1,n}$, and let $H_i(x)$ be the distribution function of $Y_{(i)}$. To begin with we find $E(X_{(i)} \mid X_{nn})$. Now for $1 < i < n$,

$$X_{(i)} = \min(Y_{(i)}, X_{nn}) - \min(Y_{(i-1)}, X_{nn}) + Y_{(i-1)}\,.$$

So $E(X_{(i)} \mid X_{nn} = x) = E[\min(Y_{(i)}, x) - \min(Y_{(i-1)}, x)] + \text{constant}$. Integration by parts gives

$$\begin{aligned} E[\min(Y_{(i)}, x)] &= \int_0^x (1 - H_i(y))\, dy - \int_{-\infty}^0 H_i(y)\, dy \qquad &\text{for } x \geqq 0 \\ &= x - \int_{-\infty}^x H_i(y)\, dy &\text{for } x < 0\,. \end{aligned}$$

Since $H_{i-1}(y) - H_i(y) = P_{in}^n(y)$, we find

$$E(X_{(i)} \mid X_{nn} = x) = \int_{-\infty}^\infty I_{[y<x]} P_{in}^n(y)\, dy + \text{constant}\,.$$

A symmetric argument gives

$$E(X_{(i)} \mid X_{kn} = x) = \int_{-\infty}^{\infty} I_{[y<x]} P_{ik}^n(y)\, dy + \text{constant}\,, \qquad k = 1, \cdots, n\,,$$

and (2) implies (5). Equation (6) then follows from the independence of the X_{kn}, Fubini's theorem, and the fact that

$$\operatorname{Cov}(I_{[y<X_{kn}]}, I_{[z<X_{kn}]}) = F_{kn}(\min(y, z)) - F_{kn}(y)F_{kn}(z)\,. \ \square$$

We remark that in the identically distributed case ($F_{kn} \equiv F$ all k, n),

$$P_{ik}^n(y) = \binom{n-1}{i-1} F(y)^{i-1}(1 - F(y))^{n-i}\,,$$

and (6) can be written explicitly in a very simple form.

It will later prove convenient to deal separately with the two extremal order statistics. The following result will imply that their contribution to S_n is negligible.

PROPOSITION 4. *Suppose that for some distribution function $G(y)$ of a random variable Y with $EY^2 < \infty$ it is true that whenever $y \leqq -M$, $F_{kn}(y) \leqq G(y)$, and whenever $y \geqq M$, $F_{kn}(y) \geqq G(y)$, where M is some finite constant. Then $n^{-1}E(X_{(1)}^2 + X_{(n)}^2) \to 0$ as $n \to \infty$.*

PROOF. It is clearly enough to prove the result for the case $F_{kn} \equiv G$ all k, n. Then we have (by integration by parts)

$$n^{-1}EX_{(n)}^2 = 2 \int_0^{\infty} xn^{-1}(1 - G(x)^n)\, dx - 2 \int_{-\infty}^{0} xn^{-1}G(x)^n\, dx\,.$$

But $n^{-1}(1 - G(x)^n) \to 0$ and $n^{-1}G(x)^n \to 0$ as $n \to \infty$ for each x, and since $n^{-1}(1 - G(x)^n) \leqq 1 - G(x)$ and $n^{-1}G(x)^n \leqq G(x)$ for all x, it follows from the Dominated Convergence Theorem that $n^{-1}EX_{(n)}^2 \to 0$, because $\int_0^{\infty} x(1 - G(x))\, dx$ and $\int_{-\infty}^{0} xG(x)\, dx$ are finite by hypothesis. $n^{-1}EX_{(1)}^2 \to 0$ by a symmetric argument. $\square$

The simplicity of a number of our later derivations depends on the representation of the covariance of two random variables given by the following proposition. This result, together with a short proof, is given in Lehmann (1966), where both are attributed to Hoeffding.

PROPOSITION 5. *Let X and Y be any two real-valued random variables with joint distribution given by $F(x, y) = P(X \leqq x, Y \leqq y)$ and marginal distributions $G(x) = F(x, \infty)$ and $H(y) = F(\infty, y)$. Then if $E(XY)$, $E(X)$, and $E(Y)$ exist,*

$$\operatorname{Cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(x, y) - G(x)H(y)]\, dx\, dy\,. \tag{7}$$

We will also need the following fact in later proofs.

PROPOSITION 6. (Esary, Proschan, and Walkup (1967)). *For any x, y, i, j, n, $\{F_{kn}\}$, we have $P(X_{(i)} \leqq x, X_{(j)} \leqq y) \geqq P(X_{(i)} \leqq x)P(X_{(j)} \leqq y)$.*

3. The independent identically distributed case. We shall now consider the important special case where all of the independent random variables $\{X_{kn}\}$ have the same distribution; that is, $F_{kn} \equiv F$, all k, n. To conform with standard notation we shall, for this section only, consider these random variables as a singly indexed sequence $X_1, X_2, \cdots, X_n, \cdots$. Then each X_i has distribution function $F(x)$, and $X_{(i)}$ represents the ith order statistic among $X_1, \cdots, X_n$. As before, we make no assumptions yet about F.

Our main aim in this section is to present an extremely simply stated theorem (Theorem 2) which is nonetheless of sufficient generality for most applications. Our method of proof will be fairly straightforward. We wish to prove that $S_n = n^{-1} \sum_{i=1}^n J(i/(n+1))X_{(i)}$ is asymptotically normally distributed. Because the expressions for $\hat{X}_{(i)}$ (the projection of $X_{(i)}$) for $i = 1$ or n are slightly different from (5), it will be convenient to consider

$$T_n = \frac{1}{n} \sum_{i=2}^{n-1} J\left(\frac{i}{n+1}\right) X_{(i)}$$

rather than S_n. That S_n and T_n are asymptotically equivalent in mean square will follow from Proposition 4. Theorem 1 below and Proposition 1 will then allow us to conclude that S_n and $\hat{T}_n$ are asymptotically equivalent in mean square. Theorem 2 will then establish the asymptotic normality of $\hat{T}_n$, thus of S_n.

THEOREM 1. *Assume that* $E|X_i|^2 < \infty$, *and that* $J(u)$ *is bounded and continuous* a.e. F^{-1}. *Then*

$$\lim_{n\to\infty} n\sigma^2(\hat{T}_n) = \sigma^2(J, F) \tag{8}$$

and

$$\lim_{n\to\infty} n\sigma^2(S_n) = \sigma^2(J, F) , \tag{9}$$

where

$$\sigma^2(J, F) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(F(x))J(F(y))[F(\min(x, y)) - F(x)F(y)]\, dx\, dy . \tag{10}$$

PROOF. We first prove (8). Recall from (6) of Proposition 3 that for the i.i.d. case we have

$$\mathrm{Cov}(\hat{X}_{(i)}, \hat{X}_{(j)}) = n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(\min(x, y)) - F(x)F(y)]P_i(x)P_j(y)\, dx\, dy$$

where

$$P_i(x) = \binom{n-1}{i-1} F(x)^{i-1}(1 - F(x))^{n-i} \qquad \text{for } 1 < i < n .$$

Then

$$n\sigma^2(\hat{T}_n) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(\min(x, y)) - F(x)F(y)]Q_n(x)Q_n(y)\, dx\, dy ,$$

where

$$Q_n(x) = \sum_{i=2}^{n-1} J\left(\frac{i}{n+1}\right) P_i(x) .$$

Now by the Bernoulli weak law of large numbers, $Q_n(x) \to J(F(x))$ for every x such that $J(u)$ is continuous at $u = F(x)$ and $1 > F(x) > 0$. Since $|Q_n(x)| \leq \sup |J(u)|$ and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(\min(x, y)) - F(x)F(y)]\, dx\, dy = \sigma^2(X_i)$$

(by Proposition 5) is assumed finite, it follows from the Dominated Convergence Theorem that $n\sigma^2(\hat{T}_n) \to \sigma^2(J, F)$.

In order to prove (9), let

$$G_i(x) = P(X_{(i)} \leqq x)$$
$$G_{ij}(x, y) = P(X_{(i)} \leqq x, X_{(j)} \leqq y) .$$

Then

$$n\sigma^2(S_n) = n^{-1} \sum_{i=1}^n \sum_{j=1}^n J\left(\frac{i}{n+1}\right) J\left(\frac{j}{n+1}\right) \operatorname{Cov}(X_{(i)}, X_{(j)}) ,$$

and by Proposition 5,

$$\operatorname{Cov}(X_{(i)}, X_{(j)}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [G_{ij}(x, y) - G_i(x)G_j(y)]\, dx\, dy ,$$

so

$$n\sigma^2(S_n) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_n(x, y)\, dx\, dy$$

where

$$H_n(x, y) = n^{-1} \sum_{i=1}^n \sum_{j=1}^n J\left(\frac{i}{n+1}\right) J\left(\frac{j}{n+1}\right) [G_{ij}(x, y) - G_i(x)G_j(y)] .$$

We first claim that

$$H_n(x, y) \to J(F(x))J(F(y))[F(\min(x, y)) - F(x)F(y)]$$

for almost every x and y, as $n \to \infty$. Without loss of generality, fix x and y such that $J(u)$ is continuous at $u = F(x)$ and $u = F(y)$. Define

$$B_n = \{(i, j) : |i/(n+1) - F(x)| \leqq n^{-\frac{1}{3}}, |j/(n+1) - F(y)| \leqq n^{-\frac{1}{3}}\} .$$

Now $G_i(x)$ is the probability that the number of X_k's less than or equal to x is at least i, the upper tail of a binomial $(n, F(x))$ distribution, and similarly $G_j(y)$ is the upper tail of a binomial $(n, F(y))$ distribution, while $G_{ij}(x, y)$ is the joint upper quadrant probability for the two binomial variables. By applying Chebychev's inequality (for fourth powers) to these binomial probabilities, we find that unless $(i, j) \in B_n$, $|G_{ij}(x, y) - G_i(x)G_j(y)| \leqq Cn^{-\frac{2}{3}}$, where C is independent of i and j (but may depend on F, x and y). It then follows that (since J is bounded) the terms of the double summation corresponding to (i, j) not in B_n contribute nothing asymptotically, and so (since J is continuous at $F(x)$ and $F(y)$) $J(i/(n+1))$ and $J(j/(n+1))$ can be replaced by $J(F(x))$ and $J(F(y))$; that is,

$$|H_n(x, y) - n^{-1} \sum_{i=1}^n \sum_{j=1}^n J(F(x))J(F(y))[G_{ij}(x, y) - G_i(x)G_j(y)]| \to 0$$

as $n \to \infty$. But

$$\begin{aligned} n^{-1} \textstyle\sum_{i=1}^n \sum_{j=1}^n [G_{ij}(x, y) - G_i(x)G_j(y)] & \\ &= n^{-1} \textstyle\sum_{i=1}^n \sum_{j=1}^n \operatorname{Cov}(I_{[X_{(i)} \leqq x]}, I_{[X_{(j)} \leqq y]}) \\ &= n^{-1} \operatorname{Cov}(\textstyle\sum_{i=1}^n I_{[X_{(i)} \leqq x]}, \sum_{j=1}^n I_{[X_{(j)} \leqq y]}) \\ &= n^{-1} \operatorname{Cov}(\textstyle\sum_{i=1}^n I_{[X_i \leqq x]}, \sum_{j=1}^n I_{[X_j \leqq y]}) \\ &= \operatorname{Cov}(I_{[X_i \leqq x]}, I_{[X_j \leqq y]}) = F(\min(x, y)) - F(x)F(y) , \end{aligned}$$

and we see that

$$H_n(x, y) \to J(F(x))J(F(y))[F(\min(x, y)) - F(x)F(y)] .$$

To prove (9) it only remains to show that the Dominated Convergence Theorem can be applied. But if $\sup|J| = M < \infty$, then since $G_{ij}(x, y) = G_i(x)G_j(y)$ all i, j, x, y by Proposition 6, we have that

$$\begin{aligned}|H_n(x, y)| &\leqq M^2 n^{-1} \textstyle\sum_{i=1}^n \sum_{j=1}^n [G_{ij}(x, y) - G_i(x)G_j(y)] \\ &= M^2[F(\min(x, y)) - F(x)F(y)]\end{aligned}$$

which is integrable by hypothesis. □

We are now able to give conditions under which S_n is asymptotically normally distributed.

THEOREM 2. *Assume that* $E(X_i^2) < \infty$, *and that* $J(u)$ *is bounded and continuous* a.e. F^{-1}. *Then* $\sigma^2(J, F) > 0$ *implies*

$$\mathscr{L}\left(\frac{S_n - E(S_n)}{\sigma(S_n)}\right) \to N(0, 1) \qquad \text{as } n \to \infty .$$

$\sigma^2(J, F)$ *is given by* (10).

PROOF. Proposition 4 implies that $n^{\frac{1}{2}}(S_n - ES_n)$ and $n^{\frac{1}{2}}(T_n - ET_n)$ are asymptotically equivalent in mean square, and Theorem 1 and Proposition 1 imply that $n^{\frac{1}{2}}(T_n - ET_n)$ and $n^{\frac{1}{2}}(\hat{T}_n - E\hat{T}_n)$ are asymptotically equivalent in mean square. Therefore, it only remains to show that $n^{\frac{1}{2}}(\hat{T}_n - E\hat{T}_n)$ is asymptotically normally distributed. From (5) of Proposition 3, we can write $\hat{T}_n - E\hat{T}_n = n^{-1} \sum_{k=1}^n Z_{kn}$, where

$$\begin{aligned}Z_{kn} &= \textstyle\sum_{i=2}^{n-1} J\left(\frac{i}{n+1}\right) \int_{-\infty}^{\infty} (F(y) - I_{[X_k \leqq y]})P_i(y)\, dy \\ &= \textstyle\int_{-\infty}^{\infty} (F(y) - I_{[X_k \leqq y]})Q_n(y)\, dy ,\end{aligned}$$

with $P_i(y)$ and $Q_n(y)$ as defined in the proof of Theorem 1. Let

$$Z_k = \int_{-\infty}^{\infty} (F(y) - I_{[X_k \leqq y]})J(F(y))\, dy .$$

Then $n^{-\frac{1}{2}} \sum_{k=1}^n Z_{kn}$ and $n^{-\frac{1}{2}} \sum_{k=1}^n Z_k$ are asymptotically equivalent in mean square since

$$\begin{aligned}\sigma^2(n^{-\frac{1}{2}} \textstyle\sum_{k=1}^n (Z_{kn} - Z_k)) &= \sigma^2(Z_{kn} - Z_k) \\ &= \textstyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(\min(x, y)) - F(x)F(y)] \\ &\qquad \times [Q_n(x) - J(F(x))][Q_n(y) - J(F(y))]\, dx\, dy \to 0\end{aligned}$$

by the proof of (8) of Theorem 1. But the Z_k's are i.i.d. with finite variance $\sigma^2(J, F)$, so the Central Limit Theorem tells us that $n^{-\frac{1}{2}} \sum_{k=1}^n Z_k$ is asymptotically normal, so $n^{\frac{1}{2}}(\hat{T}_n - E\hat{T}_n)$ is also, and the theorem follows. □

In proving the asymptotic normality of S_n, it was not necessary to say anything about the behavior of $E(S_n)$ for large n. However, since statistics of the form S_n are usually employed as estimates of their expectations, the following two results are of interest. Theorem 4, which gives conditions under which $E(S_n)$ approaches its limit faster than $n^{-\frac{1}{2}}$, is of greater statistical relevance than Theorem 3, which only asserts convergence. However, the weaker conditions of Theorem 3 help illustrate the trade-off between regularity conditions and the conclusion which is available with the present method of proof, and its simple proof helps to illustrate the idea behind the more complex argument needed for Theorem 4.

THEOREM 3. *Assume that* $E|X_i| < \infty$, *and that* $J(u)$ *is bounded and continuous* a.e. *with respect to Lebesgue measure. Then as* $n \to \infty$, $E(S_n) \to \mu(J, F)$, *where*

$$\mu(J, F) = \int_0^\infty \int_{F(x)}^1 J(u)\,du\,dx - \int_{-\infty}^0 \int_0^{F(x)} J(u)\,du\,dx = \int_0^1 J(u)F^{-1}(u)\,du\,. \tag{11}$$

PROOF. Integration by parts leads directly to the expression

$$E(S_n) = \int_0^\infty \left[n^{-1} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) P(X_{(i)} > x)\right] dx - \int_{-\infty}^0 \left[n^{-1} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) P(X_{(i)} \leqq x)\right] dx\,.$$

We claim that for each fixed $x \geqq 0$ such that $F(x) < 1$,

$$n^{-1} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) P(X_{(i)} > x) \to \int_{F(x)}^1 J(u)\,du\,. \tag{12}$$

To see this, define a random variable Y_n taking values in $[0, 1]$ by

$$P\left(Y_n = \frac{i}{(n+1)}\right) = (1 - F(x))^{-1} n^{-1} P(X_{(i)} > x)\,, \qquad i = 1, \cdots, n\,.$$

This is a probability distribution since

$$\begin{aligned} \sum_{i=1}^n P(X_{(i)} > x) &= E(\sum_{i=1}^n I_{[X_{(i)} > x]}) \\ &= E(\sum_{i=1}^n I_{[X_i > x]}) \\ &= n(1 - F(x))\,, \end{aligned}$$

and the left-hand side of (12) is just $(1 - F(x))EJ(Y_n)$. But $P(X_{(i)} > x)$ is a tail probability of a binomial $(n, F(x))$ distribution, and the Chebychev inequality implies that the distribution function of Y_n converges weakly to that of a continuous uniform distribution on $[F(x), 1]$. This gives us (12) (see Billingsley (1968), page 31), and applying the Dominated Convergence Theorem (since $1 - F(x)$ is integrable for $x > 0$) and using a similar argument for $x < 0$ proves (11). Fubini's theorem gives the final equality. □

THEOREM 4. *Assume that* $\int [F(x)(1 - F(x))]^{\frac{1}{2}}\,dx < \infty$ *and that* $J(u)$ *is bounded and satisfies Hölder condition with* $\alpha > \frac{1}{2}$ *(except possibly at a finite number of points of* F^{-1} *measure zero). Then* $n^{\frac{1}{2}}(E(S_n) - \mu(J, F)) \to 0$, *where* $\mu(J, F)$ *is given by* (11).

PROOF. Without loss of generality assume zero is a median of F; we can do this since $n^{\frac{1}{2}}[n^{-1} \sum_{i=1}^n J(i/(n+1)) - \int_0^1 J(u)\,du] \to 0$. Integration by parts leads directly to the expression

$$E(S_n) = \int_0^\infty \left[n^{-1} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) P(X_{(i)} > x) \right] dx$$
$$- \int_{-\infty}^0 \left[n^{-1} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) P(X_{(i)} \leqq x) \right] dx .$$

We shall prove that the first integral equals

$$\int_0^\infty \int_{F(x)}^1 J(u)\, du\, dx + o(n^{-\frac{1}{2}}) ;$$

the second integral can be dealt with in a similar manner. Now let

$$I_n = n^{\frac{1}{2}} \int_0^\infty \left\{ n^{-1} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) P(X_{(i)} > x) - \int_{F(x)}^1 J(u)\, du \right\} dx$$
$$= \int_0^\infty \int_0^1 J(u)\, dH_n(u; x)\, dx ,$$

where

$$\begin{aligned} H_n(u; x) &= n^{\frac{1}{2}}(1-u) - n^{-\frac{1}{2}} \textstyle\sum_{i>(n+1)u} P(X_{(i)} > x) && u \geqq F(x) \\ &= n^{-\frac{1}{2}} \textstyle\sum_{i \leqq (n+1)u} P(X_{(i)} > x) && u < F(x) , \\ &= n^{-\frac{1}{2}} \textstyle\sum_{i>(n+1)u} P(X_{(i)} \leqq x) + a_n(u) && u \geqq F(x) \\ &= n^{-\frac{1}{2}} \textstyle\sum_{i \leqq (n+1)u} P(X_{(i)} > x) && u < F(x) , \end{aligned}$$

with $a_n(u) = n^{-\frac{1}{2}}[(n+1)u] - n^{\frac{1}{2}}u$ (take $a_n(1) = 0$). Define $H_n^*(u; x) = H_n(u; x) - a_n(u)I_{[u \geqq F(x)]}$. Since for every x, n, $\int_0^1 dH_n^*(u; x) = 0$, we can write $I_n = I_{1n} + I_{2n}$, where

$$I_{1n} = \int_0^\infty \int_0^1 [J(u) - J(F(x))]\, dH_n^*(u; x)\, dx ,$$
$$I_{2n} = \int_0^\infty \{\int_{F(x)}^1 J(u)\, da_n(u) + J(F(x))a_n(F(x))\}\, dx .$$

We shall show that both I_{1n} and $I_{2n} \to 0$.

First I_{1n}. We show that for fixed x such that $J(u)$ is continuous at $F(x)$, $\int_0^1 [J(u) - J(F(x))]\, dH_n^*(u; x) \to 0$, and that this integral is bounded a.e. by $K[F(x)(1 - F(x))]^{\frac{1}{2}}$. It will then follow from the Dominated Convergence Theorem that $I_{1n} \to 0$.

Now for fixed x, H_n^* is monotone increasing for $u < F(x)$ and decreasing for $u \geqq F(x)$, with a supremum at $u = F(x)$. Define $A_n = \{u: |u - F(x)| \leqq n^{-\frac{1}{4}}$ and $0 \leqq u \leqq 1\}$. Then as in the proof of Theorem 1, Chebychev's Inequality and J bounded imply that

$$\begin{aligned} &|\textstyle\int_0^1 [J(u) - J(F(x))]\, dH_n^*(u; x)| \\ &\qquad = |\textstyle\int_{A_n} [J(u) - J(F(x))]\, dH_n^*(u; x)| + o(1) \\ &\qquad \leqq 2 \sup_{u \in A_n} |J(u) - J(F(x))| \cdot \sup_u H_n^*(u; x) + o(1) . \end{aligned}$$

Now $\sup_{A_n} |J(u) - J(F(x))| \to 0$ if J is continuous at $F(x)$, and it remains to show $\sup H_n^*(u; x) = H_n^*(F(x), x)$ is bounded. We will use the fact that for any non-negative random variable Y, $E(Y) = \int_0^\infty P(Y > y)\, dy = \sum_{k=1}^\infty P(Y \geqq k)$. Let Y_n be the number of X_j's $\leqq x$ (so Y_n has a binomial $(n, F(x))$ distribution). Then

$X_{(i)} \leqq x$ if and only if $Y_n \geqq i$, and

$$H_n^*(F(x); x) = n^{-\frac{1}{2}} \sum_{i > F(x)(n+1)} P(Y_n \geqq i)$$
$$= n^{-\frac{1}{2}} E\{\max(nF(x) - [(n+1)F(x)], nF(x) - Y_n)\}.$$

Now for $F(x) \geqq .5$, $\max(nF(x) - [(n+1)F(x)], nF(x) - Y_n) \leqq |nF(x) - Y_n|$, and so

$$H_n^*(F(x), x) \leqq n^{-\frac{1}{2}} E|nF(x) - Y_n|$$
$$\leqq (F(x)(1 - F(x)))^{\frac{1}{2}} \qquad \text{(by Schwarz)}.$$

Thus $|\int [J(u) - J(F(x))]\, dH_n^*(u; x)| \to 0$ for a.e. x. Since

$$|\int [J(u) - J(F(x))]\, dH_n^*(u; x)| \leqq 4M(F(x)(1 - F(x)))^{\frac{1}{2}}$$

for all x, where $\sup |J(u)| = M$, the Dominated Convergence Theorem applies and $I_{1n} \to 0$.

Finally, we show that $I_{2n} \to 0$. We shall give the proof for the case where J satisfies the Hölder condition at all points u; the more general case is easily handled by looking at I_{2n} as a sum of integrals over intervals where the condition holds. Consider the function

$$g_n(x) = \int_{F(x)}^1 J(u)\, da_n(u) + J(F(x)) a_n(F(x)),$$

for fixed n and $x > 0$. First suppose $a_n(F(x)) \geqq 0$. Let U be a random variable with a uniform distribution on $[F(x), 1]$. Define V_n in terms of U by $V_n = i/(n+1)$ if $i - 1 < nU \leqq i$ and $[(n+1)F(x)] < i \leqq n$, and let $V_n = F(x)$ for $nU \leqq [(n+1)F(x)]$. Then $|U - V_n| \leqq n^{-1}$ always, and we have

$$|g_n(x)| = |n^{\frac{1}{2}}(1 - F(x))(EJ(V_n) - EJ(U))|$$
$$\leqq n^{\frac{1}{2}}(1 - F(x)) E|J(V_n) - J(U)|$$
$$\leqq n^{\frac{1}{2}}(1 - F(x)) \cdot K \cdot E|V_n - U|^{\alpha}$$
$$\leqq K(1 - F(x)) n^{\frac{1}{2} - \alpha}.$$

The same idea will work for the case $a_n(F(x)) < 0$, although the "distribution" of V_n will be a signed measure with a small negative mass at $F(x)$, and some care is required at this point. We omit the details. It is clear that K can be taken independent of n and x. Thus $|g_n(x)| \to 0$ all x, and $|g_n(x)| \leqq K(1 - F(x))$ integrable, so $I_{2n} \to 0$. □

The condition that $\int [F(x)(1 - F(x))]^{\frac{1}{2}}\, dx < \infty$ is almost the same as the existence of a finite second moment. In fact, if the distribution function F has regularly varying tails (see Feller (1966) page 268) with a finite exponent, the two conditions are equivalent.

Theorems 1 and 2 both require that the second moment of F be finite. It is easy to see that the proofs will not work without this condition if J puts positive weight on the extremes, as we are dealing with mean square equivalence. However, if J puts no weight on the extremes, i.e. "trims" them, this condition can be relaxed considerably, as we shall now see.

THEOREM 5. *Assume that for some* $\varepsilon > 0$, $\lim_{x\to\infty} x^\varepsilon[1 - F(x) + F(-x)] = 0$, *and that* $J(u)$ *is bounded and continuous* a.e. F^{-1}. *If in addition,* $J(u) = 0$ *for* $0 < u < \alpha$ *and* $1 - \alpha < u < 1$, *then* $n\sigma^2(S_n) \to \sigma^2(J, F)$ *(given by* (10)), *and if* $\sigma^2(J, F) > 0$,

$$\mathscr{L}\left(\frac{S_n - E(S_n)}{\sigma(S_n)}\right) \to N(0, 1) \qquad \textit{as } n \to \infty .$$

Furthermore, the assumptions that $E|X_i| < \infty$ *and that* $\int [F(x)(1 - F(x))]^{\frac{1}{2}} dx$ *can be dropped from Theorems* 3 *and* 4 *and the conclusions will still hold.*

PROOF. The proofs of Theorems 1, 2, 3, 4 apply with little change. Proposition 2 ensures the existence of the relevant moments of S_n, at least for n sufficiently large, and allows us to use Proposition 1. The only change in the proofs comes where we invoke the Dominated Convergence Theorem. But in all cases, the following procedure allows us to pass to the limit under the integral. First, consider the integral for $F^{-1}(\alpha) \leqq x, y \leqq F^{-1}(1 - \alpha)$. Here the Dominated Convergence Theorem can be easily applied. The asymptotic negligibility of the remainder follows from a suitable bound on a binomial tail probability. As the procedure is essentially the same in all cases, we will only illustrate its use in the proof of (8). First, for $x < F^{-1}(\alpha)$, $|Q_n(x)| \leqq M \cdot P(W \geqq n\alpha)$, where W has a binomial $(n - 1, F(x))$ distribution and $M = \sup |J(u)|$. Now from Chebychev's Inequality for the kth factorial moment

$$P(W > n\alpha) \leqq (n\alpha)^{-[k]} EW^{[k]} = (n\alpha)^{-[k]}(n - 1)^{[k]} F(x)^k \leqq CF(x)^k$$

for n large enough, where C depends only on α. But for $k > 2/\varepsilon$, this bound is uniformly integrable in n for $x < F^{-1}(\alpha)$ by hypothesis, and so for n sufficiently large $|Q_n(x)|$ is uniformly bounded by an integrable function, and the Dominated Convergence Theorem applies for this part of the integral also. Proceed analogously for $x > F^{-1}(1 - \alpha)$. □

Clearly if only one tail of F is "heavy," a similar result holds assuming only that J trims the extremes in the "heavy" tail.

Remarks on the assumptions.

REMARK 1. The hypotheses of Theorems 1 through 5 are not the weakest possible for the given conclusions. However, some of them—in particular the moment conditions and the continuity assumption on J—are necessary to the present method of proof. That the continuity assumption is to some degree necessary to the conclusions will be evident from example 5.6 of Section 5. The moment conditions of Theorems 1, 2, 3, and 4 are the weakest yet obtained in the literature. The assumption that J is bounded is another matter. Using a different method of proof, Shorack (1972) has shown S_n to be asymptotically normal for some unbounded J, but at the expense of stronger conditions on the tails of the distribution. From the point of view of applications the restriction to bounded J is relatively innocuous since statistics S_n with unbounded J will be extremely sensitive to outliers, defeating the very purpose of robust inference which led to the consideration of these statistics.

REMARK 2. It could be argued that the restriction to statistics of the form S_n is rather severe; that we are often interested in statistics

$$S_n' = \frac{1}{n} \sum_{i=1}^{n} J_n\left(\frac{i}{n+1}\right) X_{(i)}, \tag{13}$$

where $J_n \to J$ in some sense. However, inspection of the above proofs reveals that Theorems 1 through 5 apply equally well to S_n' as long as the J_n are uniformly bounded, and for every continuity point p of J there is an open neighborhood of p such that $J_n(u) \to J(u)$ uniformly in this neighborhood.

REMARK 3. Finally, it can be easily shown that the statistic S_n of Theorem 5 is still asymptotically normal (normalized by $\mu(J, F)$ and $\sigma^2(J, F)$) without any condition on the tails of F, provided J satisfies the conditions of Theorem 4 (the statistic S_n is equivalent in probability to one derived from observations truncated at $F^{-1}(\alpha/2)$ and $F^{-1}(1-\alpha/2)$). However, the tail condition $x^\varepsilon[1 - F(x) + F(-x)] \to 0$ is necessary for any moment of S_n to be finite for any n.

4. The independent case. With few exceptions (Weiss (1969), Shorack (1973)), the published literature on linear functions of order statistics has been concerned with the case where the unordered observations are not only independent, but identically distributed. In this section we will see that all of the results for the i.i.d. case given in Section 3 carry over in some form to the general independent case, with little change in the proofs. In this section we shall return to the notation of Section 2.

THEOREM 6. *Suppose that for some distribution function $G(y)$ of a random variable Y with $EY^2 < \infty$ it is true that whenever $y \leqq -M$, $F_{kn}(y) \leqq G(y)$, and whenever $y \geqq M$, $F_{kn}(y) \geqq G(y)$, where M is some finite constant. Assume also that both*

$$\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} F_{kn}(x) = \bar{F}(x) \tag{14}$$

and

$$\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} [F_{kn}(\min(x, y)) - F_{kn}(x)F_{kn}(y)] = K(x, y) \tag{15}$$

exist for a.e. x, y *with respect to Lebesgue measure. Then if $J(u)$ is bounded and continuous* a.e. $\bar{F}^{-1}$, $n\sigma^2(S_n) \to \sigma^2(J, \bar{F}, K)$ *(given below), and if* $\sigma^2(J, \bar{F}, K) > 0$, *then*

$$\mathscr{L}\left(\frac{S_n - E(S_n)}{\sigma(S_n)}\right) \to N(0, 1) \qquad \text{as } n \to \infty .$$

Here

$$\sigma^2(J, \bar{F}, K) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(\bar{F}(x))J(\bar{F}(y))K(x, y)\, dx\, dy .$$

If in addition J satisfies the conditions of Theorem 4 *and* $\int [G(y)(1 - G(y))]^{\frac{1}{2}}\, dy < \infty$, *then* $n^{\frac{1}{2}}(E(S_n) - \mu(J, \bar{F})) \to 0$, *where μ is given by* (11).

PROOF. We shall follow the proofs of Theorems 1 and 2. First, from Proposition 3,

$$n\sigma^2(\hat{T}_n) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_n(x, y)\, dx\, dy ,$$

where

$$K_n(x, y) = \frac{1}{n} \sum_{k=1}^{n} [F_{kn}(\min(x, y)) - F_{kn}(x)F_{kn}(y)]Q_k^n(x)Q_k^n(y) ,$$

$$Q_k^n(x) = \sum_{i=2}^{n-1} J\left(\frac{i}{n+1}\right) P_{ik}^n(x) .$$

Fix x, y so that (14) and (15) hold and $J(u)$ is continuous at $u = \bar{F}(x)$ and $u = \bar{F}(y)$, with $0 < \bar{F}(x), \bar{F}(y) < 1$. For each fixed k, $Q_k^n(x) \to J(\bar{F}(x))$ and $Q_k^n(y) \to J(\bar{F}(y))$ by the weak law of large numbers, since the $P_{ik}^n(x)$ are still binomial probabilities, but with unequal probabilities. In fact, it follows easily from Chebychev's Inequality that the convergence is uniform in k, for fixed x and y. But then (15) implies that $K_n(x, y) \to K(x, y)J(\bar{F}(x))J(\bar{F}(y))$. The hypotheses of the theorem imply that

$$F_{kn}(\min(x, y)) - F_{kn}(x)F_{kn}(y) \leqq C[G(\min(x, y)) - G(x)G(y)] \tag{16}$$

for some constant C, all k, n; thus if

$$\sup|J| = M , \qquad |K_n(x, y)| \leqq M^2 \cdot C \cdot [G(\min(x, y)) - G(x)G(y)]$$

and the Dominated Convergence Theorem gives $n\sigma^2(\hat{T}_n) \to \sigma^2(J, \bar{F}, K)$.

The proof that $n\sigma^2(S_n) \to \sigma^2(J, \bar{F}, K)$ follows as in Theorem 1, except that now

$$n^{-1} \sum \sum [G_{ij}(x, y) - G_i(x)G_j(y)] \\ = n^{-1} \sum [F_{kn}(\min(x, y)) - F_{kn}(x)F_{kn}(y)] \to K(x, y) .$$

Then Chebychev's Inequality for the generalized binomial distribution and (16) give the desired conclusion as before.

The proof that $\hat{T}_n$ is asymptotically normal is also straightforward. We now have $\hat{T}_n - E\hat{T}_n = n^{-1} \sum_{k=1}^{n} Z_{kn}$, with

$$Z_{kn} = \int_{-\infty}^{\infty} (F_{kn}(y) - I_{[X_{kn} \leq y]})Q_k^n(y)\, dy ,$$

$$Q_k^n(y) = \sum_{i=2}^{n-1} J\left(\frac{i}{n+1}\right) P_{ik}^n(y) .$$

Then since $Q_k^n(y) \to J(\bar{F}(y))$ a.e., it follows as above that $\hat{T}_n$ is asymptotically equivalent in mean square to $n^{-1} \sum Z'_{kn}$,

$$Z'_{kn} = \int_{-\infty}^{\infty} (F_{kn}(y) - I_{[X_{kn} \leq y]})J(\bar{F}(y))\, dy .$$

The asymptotic normality of $n^{-1} \sum Z'_{kn}$ follows easily from the Lindeberg-Feller Theorem (in particular, the hypotheses of our theorem imply the Z'_{kn}'s are uniformly square integrable.)

The final statement of Theorem 6 follows in an equally straightforward manner. □

The following result follows easily along the lines of Theorem 5, using Proposition 2.

THEOREM 7. *If the moment condition on G of Theorem 6 is replaced by the condition that for some* $\varepsilon > 0$, $\lim_{x\to\infty} x^\varepsilon[1 - G(x) + G(-x)] = 0$, *and it is also required that* $J(u) = 0$ *for* $0 < u < \alpha$ *and* $1 - \alpha < u < 1$, *the conclusions of Theorem* 6 *continue to hold.*

Clearly Theorem 3 can also be easily modified for the independent case, and the remarks at the end of Section 3 remain true in the independent case. In particular, if S_n' is given by (13), and for every continuity point p of J there is an open neighborhood of p such that $J_n(u) \to J(u)$ uniformly in this neighborhood, and the J_n are uniformly bounded, then Theorems 6 and 7 apply to S_n' also.

Finally, we remark that slightly modified versions of the asymptotic normality parts of Theorems 6 and 7 continue to hold without hypotheses (14) and (15), since the theorems can be seen to hold uniformly in $\{F_{kn}\}$, satisfying the first hypothesis. In particular, the assumption $\sigma^2(J, \bar{F}, K) > 0$ is replaced by the condition $\liminf n\sigma^2(S_n) > 0$, and J is assumed bounded and continuous except possibly at a set of points of measure zero under each F_{kn}^{-1}, and such that the $\{F_{kn}^{-1}\}$ are equicontinuous at these points.

5. Applications and examples.

5.1. *The sample mean.* If we take $J(u) \equiv 1$, we see that $S_n = n^{-1} \sum_{k=1}^n X_{kn}$, the sample mean. Then Theorem 2 for the i.i.d. case says that S_n is asymptotically normal as long as $E(X_{kn}^2) < \infty$, the usual form of the Central Limit Theorem. Theorem 6 can be viewed as a special case of the Lindeberg-Feller theorem.

5.2. *The trimmed mean.* Let $S_n(\alpha)$ denote the α-trimmed mean,

$$S_n(\alpha) = (n - 2[\alpha n])^{-1} \sum_{i=[\alpha n]+1}^{n-[\alpha n]} X_{(i)} .$$

Then in the i.i.d. case, Theorem 5 applies as long as $x^\varepsilon[1 - F(x) + F(-x)] \to 0$ for some $\varepsilon > 0$ as $x \to \infty$, and the αth and $(1 - \alpha)$th percentiles of F are unique. Here $J(u) = (1 - 2\alpha)^{-1}$ for $\alpha \leqq u \leqq 1 - \alpha$, $J(u) = 0$ otherwise. In the more general independent case, Theorem 7 applies if $x^\varepsilon[1 - G(x) + G(-x)] \to 0$ for some $\varepsilon > 0$ as $x \to \infty$, and the αth and $(1 - \alpha)$th percentiles of $\bar{F}$ are unique. As remarked at the end of Section 3, the tail condition on F (resp. G) is not necessary for asymptotic normality, only for the existence of moments. See Stigler (1973) for a necessary and sufficient condition that $S_n(\alpha)$ be asymptotically normal.

5.3. *Gini's mean difference.* Gini's mean difference, given by

$$G_n = [n(n - 1)]^{-1} \sum_{i,j=1}^n |X_{in} - X_{jn}| ,$$

can be written as a multiple of a statistic S_n:

$$G_n = 4\left(\frac{n + 1}{n - 1}\right) S_n$$

where $J(u) = u - \frac{1}{2}$ (see David (1970), page 146). Theorem 2 applies in the i.i.d. case if the second moment is finite, and Theorem 6 applies in the independent case, giving a result which holds under a weaker moment condition than that obtained by applying Theorem 8.1 of Hoeffding (1948).

5.4. *Other robust statistics.* Many robust statistics which have been considered are already of the form of S_n (see Chernoff, Gastwirth, and Johns (1967) and Jaeckel (1971) for example). Some authors have also considered statistics of the form $n^{-1} \sum J_n(i/(n+1))X_{(i)}$, where

$$(n+1)^{-1}J_n(i/(n+1)) = \int_{(i-1)/(n+1)}^{i/(n+1)} J(u)\, du + o(n^{-1}),$$

uniformly for $i/(n+1) \in (0, 1)$, (see Bickel (1967)). Since $J_n(u) \to J(u)$ uniformly in some neighborhood of each continuity point of J, our results apply to these statistics too.

5.5. *Nonidentically distributed observations.* As an example of a situation in which Theorem 6 could be applied, suppose that the F_{kn} all belong to the same scale-location family, $F_{kn}(x) = F((x - \theta_k)/a_k)$, where the second moment of F is finite, F is continuous, and all of the parameters are in some bounded set, say $|\theta_k| + a_k \leqq C$ all k. Let

$$H_n(\theta, a) = \frac{1}{n}\,(\#(\theta_k, a_k) \leqq (\theta, a),\ k = 1, \cdots, n),$$

the "distribution function" of $(\theta_1, a_1), \cdots, (\theta_n, a_n)$. Then if $H_n \to_w$ some H, Theorem 6 applies to S_n, with

$$\bar{F}(x) = \iint F\left(\frac{x-\theta}{a}\right) dH(\theta, a),$$

$$K(x, y) = \bar{F}(\min(x, y)) - \iint F\left(\frac{x-\theta}{a}\right) F\left(\frac{y-\theta}{a}\right) dH(\theta, a),$$

as long as J is bounded and continuous a.e. $\bar{F}^{-1}$. A similar result holds if each X_{kn} has a binomial (n_k, p_k) distribution or a Poisson (λ_k) distribution, or if the $X_{(i)}$ are based on grouped data.

5.6. *Discontinuous weight function.* The following example shows the necessity of requiring that J be continuous a.e. F^{-1} in the theorems of Section 3. Shorack (1972) has given a similar example based on the Bernoulli distribution; this present example demonstrates that the non-normal behavior of S_n is in no way connected with the "discreteness" of F, rather it has to do with the nonuniqueness of a percentile of F. Let $X_1, X_2, \cdots, X_n$ be i.i.d. with distribution $F(x)$ having density

$$f(x) = \begin{cases} .5 & x \in [0, 1] \cup [2, 3] \\ 0 & \text{otherwise.} \end{cases}$$

Let $J(u) = I_{[u>.5]}$. Let $Z_n = (\# X_i \leqq 1)$ and $Y_n = \max(Z_n, [n/2])$. Then we have

$$\begin{aligned} nS_n &= \sum_{i=[n/2]+1}^{Y_n} X_{(i)} + \sum_{i=Y_n+1}^{n} X_{(i)} \\ &= [\sum_{i=[n/2]+1}^{Y_n} (X_{(i)} + 1) + \sum_{i=Y_n+1}^{n} X_{(i)}] - (Y_n - [n/2]) \\ &= V_n - (Y_n - [n/2]) , \qquad \text{say.} \end{aligned}$$

Now $n^{-1}V_n$ has the same distribution as a statistic S_n based on a sample from a $R[1, 3]$ distribution, since if 1 is added to each of the original X_i's which is $\leqq 1$ the "new" X_i's will have the same joint distribution as a random sample from $R[1, 3]$, so Theorem 2 says that $n^{-\frac{1}{2}}(V_n - EV_n)$ is asymptotically normal. Clearly $n^{-\frac{1}{2}}(Z_n - [n/2])$ is asymptotically normal. It can be further shown that they are asymptotically jointly normal. But then S_n is not asymptotically normal, since $n^{-\frac{1}{2}}(Y_n - [n/2])$ is asymptotically "half-normal," with an atom of probability $\frac{1}{2}$ at zero. A more general and detailed proof is given in Stigler (1973).

Acknowledgment. I wish to thank the referee for his careful reading of the manuscript and for valuable comments.

REFERENCES

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

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

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.

DAVID, HERBERT A. (1970). *Order Statistics.* Wiley, New York.

DAVID, H. A. and MISHRIKY, R. S. (1968). Order statistics for discrete populations and for grouped samples. *J. Amer. Statist. Assoc.* **63** 1390-1398.

ESARY, J. D., PROSCHAN, F. and WALKUP, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38** 1466-1474.

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

HÁJEK, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.* **39** 325-346.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293-325.

JAECKEL, L. A. (1971). Robust estimates of location: symmetry and asymmetric contamination. *Ann. Math. Statist.* **42** 1020-1034.

LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137-1153.

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

SHORACK, G. R. (1969). Asymptotic normality of linear combinations of functions of order statistics. *Ann. Math. Statist.* **40** 2041-2050.

SHORACK, G. R. (1972). Functions of order statistics. *Ann. Math. Statist.* **43** 412-427.

SHORACK, G. R. (1973). Convergence of reduced empirical and quantile processes with applications to functions of order statistics in the non i.i.d. case. *Ann. Statist.* **1** 146-152.

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

STIGLER, S. M. (1973a). The asymptotic distribution of the trimmed mean. *Ann. Statist.* **1** 472-477.

STIGLER, S. M. (1973b). Simon Newcomb, Percy Daniell, and the history of robust estimation 1885–1920. *J. Amer. Statist. Assoc.* **68** 872–879.

WEISS, L. (1969). The asymptotic distribution of quantiles from mixed samples. *Sankhyā Ser. A* **31** 313–318.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706