The Annals of Statistics
1977, Vol. 5, No. 2, 357-369

ERROR BOUNDS FOR LINEAR COMBINATIONS OF ORDER STATISTICS[1]

BY STEINAR BJERVE

University of Oslo

A Berry-Esseen bound is obtained for trimmed linear combinations of order statistics. These linear combinations are written as the sum of a linear and a quadratic combination of independent exponentially distributed random variables plus a remainder term. The remainder term is shown to be of negligible order and Fourier methods are then employed to handle the linear and quadratic terms.

The main theorem is also given in a version that more easily lends itself to applications.

1. Introduction. Suppose that we observe a sample, $\{X_i\}_{i=1}^n$, of independent and identically distributed random variables with common distribution F (denoted i.i.d. (F)). The order-statistics from this sample are given as $\{X_{in}\}_{i=1}^n$, where $X_{1n} \leqq X_{2n} \leqq \cdots \leqq X_{nn}$.

Linear combinations of functions of order statistics (L-statistics) are statistics of the form

$$T_n = n^{-1} \sum_{i=1}^n c_{in} h(X_{in}), \tag{1.1}$$

where h is some measurable function. We will prove that L-statistics appropriately normalized under certain conditions admit a uniform error bound of the order $n^{-\frac{1}{2}}$, i.e., that

$$\sup_x |\Phi(x) - P(s_n^{-1} n(T_n - \mu_n) < x)| < K \cdot n^{-\frac{1}{2}}, \tag{1.2}$$

for some constant K. Here, s_n and μ_n are normalizing terms and Φ denotes the standard normal distribution. To this end we will apply elementary techniques as in Chernoff et al. (1967) and Fourier techniques as in Bickel (1974).

The results obtained in this work are obtained for trimmed L-statistics, i.e., a certain proportion of the observations among the smallest and the largest are discarded. The remainder of the observations then stay within finite bounds with probability one minus an exponentially small quantity. This will be seen using the following lemma derived from Bernstein's inequality which one can find in Hoeffding (1963), formula (2.13).

LEMMA 1.1. *Suppose* $X_1, \cdots, X_n$ *are* i.i.d. (F). *Let* $m = \alpha \cdot n + O(1)$. *Then*,

Received February 1974; revised October 1975.

[1] This research is based on a part of the author's Ph. D. dissertation submitted at the University of California, Berkeley, June 1974. The research was supported by National Science foundation Grant GP-33697x1 and by the Norwegian Research Council for Science and the Humanities.

AMS 1970 *subject classifications.* Primary 62E20, 62G30.

Key words and phrases. Error bounds, linear combinations of order statistics, elementary methods, Fourier methods.

if $F(\Delta) < \alpha < 1$, *there is a constant* $c > 0$ *such that*

$$P(X_{mn} < \Delta) \leqq e^{-c \cdot n} .$$

PROOF. We have $X_{mn} < \Delta$ if and only if N, the number of observations that fall to the left of Δ, is m or larger. N is obviously binomial with parameters n and $p = F(\Delta)$ and

$$P(X_{mn} < \Delta) = P(N > m) .$$

Choose $t > 0$ such that $p + t < \alpha$. Then, for n large enough,

$$P(N > m) \leqq P(N > n \cdot (p + t)) = P\left(\frac{N - np}{n} > t\right) .$$

Bernstein's inequality states that the right-hand side is less than or equal e^{-cn} for some $c > 0$.

In the case where zero weight is given to the tail order statistics, Rosenkrantz and O'Reilly (1972) obtained error bounds of the order $n^{-\frac{1}{4}}$.

In [1] Bickel obtained Berry–Esseen bounds of the type (1.2) for 2nd order U-statistics with bounded kernel. His results cover nonparametric statistics such as the Wilcoxon 2-sample statistic and Kendall's τ but do not cover the L-statistics considered in the present paper.

2. Error bounds. T_n has a distribution identical to that of

$$\frac{1}{n} \sum_{i=1}^{n} c_{in} H(Z_{in}) , \qquad H = h \circ F^{-1} \circ G , \tag{2.1}$$

where $Z_{1n} \leqq \cdots \leqq Z_{nn}$ is the ordered version of a random sample $Z_1, \cdots, Z_n$, from the exponential distribution, $G(x) = 1 - e^{-x}$. Furthermore, it is well known (David (1970), page 17), that

$$Z_{in} \sim \frac{Z_1}{n} + \cdots + \frac{Z_i}{n - i + 1} \tag{2.2}$$ [2]

and

$$EZ_{in} = \nu_{in} = \frac{1}{n} + \cdots + \frac{1}{n - i + 1} .$$

Let $c_{in} = 0$ for $i \leqq \alpha \cdot n$ or $i > \beta \cdot n$, $0 < \alpha < \beta < 1$.

Assuming H is twice differentiable and Taylor-expanding $H(Z_{in})$ about ν_{in}, we get

$$n \frac{T_n - \mu_n}{s_n} = L_n + Q_n + R_n , \tag{2.3}$$

where

$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} c_{in} H(\nu_{in}) ,$$

$$L_n = \frac{1}{s_n} \sum_{i=1}^{n} c_{in} H'(\nu_{in})(Z_{in} - \nu_{in}) ,$$

[2] $\sim$ denotes "distributed as."

$$Q_n = \frac{1}{s_n} \sum_{i=1}^n \tfrac{1}{2} c_{in} H''(\nu_{in})(Z_{in} - \nu_{in})^2 ,$$

$$R_n = \frac{1}{s_n} \sum_{i=1}^n c_{in} G_{in}(Z_{in})(Z_{in} - \nu_{in})^3 ,$$

and

$$G_{in}(z) = \left\{\left[\frac{H(z) - H(\nu_{in})}{z - \nu_{in}} - H'(\nu_{in})\right] \frac{1}{z - \nu_{in}} - \tfrac{1}{2} H''(\nu_{in})\right\} \frac{1}{z - \nu_{in}} .$$

Denote

$$g_n = \sup_{a \leq z \leq b; i: c_{in} = 0} |G_{in}(z)| \tag{2.4}$$

where a and b are constants such that

$$0 < a < -\log(1 - \alpha) \quad \text{and} \quad -\log(1 - \beta) < b < \infty .$$

The normalizing factor, s_n, is determined so that the variance of L_n is 1.

Using (2.2) we get

$$L_n \sim \frac{1}{s_n} \sum_{i=1}^n \alpha_{in}(Z_i - 1) , \tag{2.5}$$

where

$$\alpha_{in} = \frac{1}{n - i + 1} \sum_{j=i}^n c_{jn} H'(\nu_{jn}) . \tag{2.6}$$

It is seen that we must put $s_n^2 = \sum_{i=1}^n \alpha_{in}^2$.

Let F_n denote the distribution of $n(T_n - \mu_n)/s_n$. The theorem we will prove is the following.

THEOREM 1. *Let H'' satisfy a first order Lipschitz condition on $[a, b]$ and let the following hold for all n.*

$$s_n^2 > \theta \cdot n \quad \textit{for some} \quad \theta > 0 , \tag{2.7}$$

and

$$\sum_{i=1}^n |c_{in}| < c \cdot n , \tag{2.8}$$

for some $c < \infty$. Then there is a constant K such that

$$\sup_x |F_n(x) - \Phi(x)| \leq K \cdot n^{\frac{1}{2}} . \tag{2.9}$$

PROOF. We will show that a Berry–Esseen bound holds for the distribution of $M_n = L_n + Q_n$, and that R_n is of negligible order, i.e., that

$$P(|R_n| > n^{-\frac{1}{2}}) = O(n^{-\frac{1}{2}}) . \tag{2.10}$$

The result, (2.9), then follows as is seen by the following calculations:

$$\begin{aligned} F_n(x) &= P(M_n + R_n \leq x \cap |R_n| \leq n^{-\frac{1}{2}}) + P(M_n + R_n \leq x \cap |R_n| > n^{-\frac{1}{2}}) \\ &\leq P(M_n \leq x + n^{-\frac{1}{2}}) + O(n^{-\frac{1}{2}}) \\ &\leq \Phi(x + n^{-\frac{1}{2}}) + O(n^{-\frac{1}{2}}) \leq \Phi(x) + O(n^{-\frac{1}{2}}) , \end{aligned}$$

where we have used the fact that $|\Phi'(x)|$ is bounded.

A similar argument may be used to establish the inequality $F_n(x) \geqq \Phi(x) - O(n^{-\frac{1}{2}})$ from which we conclude that

$$(2.11) \qquad |F_n(x) - \Phi(x)| = O(n^{-\frac{1}{2}}) \qquad \text{uniformly in} \quad x\,.$$

We now proceed to show that $P(|R_n| > n^{-\frac{1}{2}}) = O(n^{-\frac{1}{2}})$.

The probability that Z_{in} falls outside the interval $[a, b]$ for any $\alpha \cdot n < i \leqq \beta \cdot n$ is exponentially small. This is seen by transforming the Z_{in} to uniform order statistics and using Lemma 1.1. Let $U_{in} = 1 - e^{-Z_{in}}$. Then,

$$\begin{aligned} P(Z_{in} \notin [a, b] \text{ for some } \alpha \cdot n < i \leqq \beta \cdot n) &= P(Z_{[\alpha \cdot n+1],n} < a) + P(Z_{[\beta \cdot n],n} > b) \\ &= P(U_{[\alpha \cdot n+1],n} < 1 - e^{-a}) + P(U_{[\beta \cdot n],n} > 1 - e^{-b}) \\ &= o(e^{-\varepsilon \cdot n}) \qquad \text{for some} \quad \varepsilon > 0\,, \end{aligned}$$

since $1 - e^{-a} < \alpha$ and $1 - e^{-b} > \beta$.

This, with (2.4) and the fact that $c_{in} = 0$ for $i \leqq \alpha \cdot n$ or $i > \beta \cdot n$, implies that

$$P(|R_n| > n^{-\frac{1}{2}}) \leqq P\left(\frac{g_n}{s_n} \sum_{i=1}^n |c_{in}||Z_{in} - \nu_{in}|^3 > n^{-\frac{1}{2}}\right) + o(e^{-\varepsilon \cdot n})\,.$$

It is easily seen that

$$E(Z_{in} - \nu_{in})^4 = O(n^{-2}) \qquad \text{uniformly for} \quad \alpha \cdot n < i \leqq \beta \cdot n$$

(cf. Lemma 2.1 below). Thus,

$$E|Z_{in} - \nu_{in}|^3 \leqq [E(Z_{in} - \nu_{in})^4]^{\frac{3}{4}} = O(n^{-\frac{3}{2}}) \qquad \text{uniformly for} \quad \alpha n < i \leqq \beta \cdot n\,.$$

From the assumptions on H it follows that

$$G_{in}(z) = \tfrac{1}{2}(H''(\nu'_{in}) - H''(\nu_{in}))/(z - \nu_{in})$$

where ν'_{in} is some number between z and ν_{in}. It also follows (cf. (2.4)) that the g_n are uniformly bounded. Note that $\nu_{in} \in [a, b]$ for all $i : c_{in} \neq 0$ and all n large enough. By Markov's inequality, (2.7) and (2.8) we now have

$$P(|R_n| > n^{-\frac{1}{2}}) \leqq n^{\frac{1}{2}} \cdot O(n^{-\frac{3}{2}}) \cdot \frac{g_n}{s_n} \sum |c_{in}| = O(n^{-\frac{1}{2}})\,.$$

It remains to show that a Berry–Esseen bound holds for the distribution of $L_n + Q_n$.

Let

$$\Phi_n(t) = Ee^{it(L_n + Q_n)}\,,$$

$$\eta(t) = Ee^{it(Z-1)}\,,$$

where Z is exponentially distributed with parameter 1 and let

$$\tilde{\Phi}_n(t) = Ee^{itL_n} = \prod_{j=1}^n \eta\left(\frac{\alpha_{jn} t}{s_n}\right).$$

From Esseen (1944), page 45, it is seen that one can find constants ε_1 and c_1

such that

$$\int_{-n^{\frac12}\varepsilon_1}^{n^{\frac12}\varepsilon_1} \frac{1}{|t|} |\tilde{\Phi}_n(t) - e^{-t^2/2}| \, dt < \frac{c_1}{n^{\frac12}} .$$

if

$$\frac{(1/n)(|\alpha_{1n}|^3 + \cdots + |\alpha_{nn}|^3)}{(s_n{}^2/n)^{\frac32}} = n^{\frac12} \frac{(|\alpha_{1n}|^3 + \cdots + |\alpha_{nn}|^3)}{s_n{}^3}$$

is uniformly bounded. This expression is ρ_{3n} in Esseen's notation. From (2.5) and (2.6) we in fact get

$$n^{\frac12} \frac{(|\alpha_{1n}|^3 + \cdots + |\alpha_{nn}|^3)}{s_n{}^3} \leqq \frac{|\alpha_{1n}|^3 + \cdots + |\alpha_{nn}|^3}{\theta^{\frac12} s_n{}^2} \leqq \theta^{-\frac12} \sup_i |\alpha_{in}| . \tag{2.12}$$

Since $c_{in} = 0$ when $i \leqq \alpha \cdot n$ or $i > \beta \cdot n$ and since $H'(\nu_{in})$ remain bounded for these values of i and all n, we see from the expression (2.6) that the α_{in} themselves are uniformly bounded. If we now are able to show that

$$\int_{-n^{\frac12}\varepsilon_2}^{n^{\frac12}\varepsilon_2} \frac{1}{|t|} |\Phi_n(t) - \tilde{\Phi}_n(t)| \, dt < c_2 n^{-\frac12} , \tag{2.13}$$

for some constants c_2 and $\varepsilon_2 > 0$, we may conclude that for $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ and $C = c_1 + c_2$ we have

$$\int_{-n^{\frac12}\varepsilon}^{n^{\frac12}\varepsilon} \frac{1}{|t|} |\Phi_n(t) - e^{-t^2/2}| \, dt < C \cdot n^{-\frac12} . \tag{2.14}$$

The validity of a Berry–Esseen bound for $L_n + Q_n$ then follows from Lemma 2 of Feller (1966), page 512. From the fact that

$$\left| e^{it} - 1 - it - \cdots - \frac{(it)^{k-1}}{(k-1)!} \right| \leqq \frac{|t|^k}{k!} ,$$

we have

$$\begin{aligned} |\Phi_n(t) - \tilde{\Phi}_n(t)| &= |Ee^{itL_n}(e^{itQ_n} - 1)| \\ &\leqq \textstyle\sum_{j=1}^{m-1} \left| \frac{t^j}{j!} Ee^{itL_n} Q_n{}^j \right| + \frac{|t|^m}{m!} E|Q_n|^m . \end{aligned} \tag{2.15}$$

Introducing

$$\beta_{jn} = \frac{1}{n-j+1} \textstyle\sum_{i=j}^n c_{in} H''(\nu_{in}) ,$$

it is seen that for Q_n we have

$$Q_n \sim \frac{1}{s_n} \textstyle\sum_{p,q} \beta_{p\vee q,n} \frac{(Z_p - 1)(Z_q - 1)}{n - (p \wedge q) + 1} .$$

To determine the order of magnitude of $|\Phi(t) - \tilde{\Phi}(t)|$, we need the following two lemmas.

LEMMA 2.1. *There is a constant c_0 such that*

$$E(Z_{in} - \nu_{in})^{2m} \leqq c_0{}^m n^{-m} m^m \quad \textit{for} \quad i < \beta \cdot n \quad \textit{and} \quad m < n .$$

PROOF. First, let Z, Z' be independent with equal distributions. Using Jensen's inequality we then find

$$\begin{aligned} E(Z - Z')^{2m} &= E[E(Z - Z')^{2m} \mid Z] \\ &\geqq E[E(Z - Z' \mid Z)]^{2m} \\ &= E(Z - EZ)^{2m} . \end{aligned}$$

If $Z, Z' \sim e^{-z}$, then $W = Z - Z' \sim \frac{1}{2}e^{-|z|}$ and we find that

$$E(Z - 1)^{2m} \leqq \int_{-\infty}^{\infty} z^{2m}\tfrac{1}{2}e^{-|z|}\, dz = (2m)!$$

Now, let Z, Z' be distributed as Z_{in}. Then $Z - Z' \sim W_1/n + \cdots + W_i/(n - i + 1)$ where $W_1, \cdots, W_n$ are independent with common double exponential distribution. We get:

$$\begin{aligned} E(Z_{in} - \nu_{in})^{2m} &\leqq E\left(\frac{W_1}{n} + \cdots + \frac{W_i}{n - i + 1}\right)^{2m} \\ &= \sum_{n_1+\cdots+n_i=2m} \binom{2m}{n_1, \cdots, n_i} \frac{EW_1^{n_1} \cdots W_i^{n_i}}{n^{n_1} \cdots (n - i + 1)^{n_i}} . \end{aligned}$$

Now,

$$\begin{aligned} EW_1^{n_1} \cdots W_i^{n_i} &= n_1! \cdots n_i! \qquad &&\text{if all } n\text{'s are even} \\ &= 0 &&\text{otherwise.} \end{aligned}$$

Thus,

$$\begin{aligned} E(Z_{in} - \nu_{in})^{2m} &\leqq \sum_{n_1+\cdots+n_i=2m;\, n_j \text{ even } \forall j} \binom{2m}{n_1, \cdots, n_i} \frac{n_1! \cdots n_i!}{n^{n_1} \cdots (n - i + 1)^{n_i}} \\ &\leqq \frac{(2m)!}{(n - i + 1)^{2m}} \sum_{m_1+\cdots+m_i=m} 1 \\ &= \frac{(2m)!}{(n - i + 1)^{2m}} \binom{m+i-1}{i-1} . \end{aligned}$$

Using Stirling's formula, $(2\pi)^{\frac{1}{2}}n^{n+\frac{1}{2}}\exp[-n + 1/(12n+1)] < n! < (2\pi)^{\frac{1}{2}}n^{n+\frac{1}{2}}\exp[-n + 1/(12n)]$ (Feller (1957), page 52), we can bound the right-hand side by

$$\begin{aligned} c_1 \cdot \frac{(2m)^{2m+\frac{1}{2}}(m + i - 1)^{m+k-\frac{1}{2}}e^{-2m}}{(n - i + 1)^{2m}(i - 1)^{i-\frac{1}{2}}m^{m+\frac{1}{2}}} & \qquad i > 1 \\ \leqq c_2 \cdot \frac{(m + i - 1)^m}{(n - i + 1)^{2m}} \left(1 + \frac{m}{i - 1}\right)^{i-\frac{1}{2}} m^m \left(\frac{2}{e}\right)^{2m} & \\ \leqq c_2 n^{-m} \left[\frac{1 + \beta}{1 - \beta}\right]^m m^m \cdot 2^{2m} . & \end{aligned}$$

The last inequality holds since $i < \beta \cdot n$, $m < n$ and $(1 + m/(i - 1))^{i-\frac{1}{2}} < e^{2m}$.

LEMMA 2.2. *Under the assumptions of the theorem, we can find constants c_1, c_2, c_3 such that the following holds.*

(i) $E|Q_n|^m \leqq n^{-m/2}m^m c_1^m$, $m < n$.
(ii) $|Ee^{itL_n}Q_n| \leqq c_2 n^{-\frac{1}{2}}|\tilde{\Phi}_n(t)|(t^2 + 1)$.

(iii) $|Ee^{itL_n}Q_n{}^m| \leqq c_3{}^m m^{2m} n^{m/2} \sup_{S\subset\{1,\cdots,n\};\#S\geqq n-2m} \left|\prod_{j\in S} \eta\left(\frac{\alpha_{jn}\cdot t}{s_n}\right)\right|$, $m < n$.[3]

PROOF. We may write

$$\begin{aligned} s_n{}^m \cdot E|Q_n|^m &= E|\textstyle\sum_{i=1}^n c''_{in}(Z_{in} - \nu_{in})^2|^m \\ &\leqq \textstyle\sum_{i_1,\cdots,i_m}^n E\prod_{j=1}^m |c''_{i_j}|(Z_{i_jn} - \nu_{i_jn})^2 \\ &\leqq \textstyle\sum \prod_{j=1}^m |c''_{i_j}|[E(Z_{i_jn} - \nu_{i_jn})^{2m}]^{1/m}, \end{aligned}$$

where $c''_{in} = \frac{1}{2}c_{in}H''(\nu''_{in})$.

Using the result of Lemma 2.1, we get for $m < n$,

$$\begin{aligned} E|Q_n|^m &\leqq c_0{}^m n^{-m} m^m s_n{}^{-m} \textstyle\sum \prod |c''_{i_j}| \\ &\leqq \theta^{-m/2} c_0{}^m n^{-m/2} m^m \left(\frac{1}{n}\textstyle\sum |c''_{i_j}|\right)^m \\ &\leqq n^{-m/2} m^m c_1{}^m, \quad \text{which proves (i).} \end{aligned}$$

Now, write

$$|E^{itL_n}Q_n| = \left|Ee^{itL_n}s_n{}^{-1}\textstyle\sum_{p,q}\beta_{p\vee q,n}\frac{(Z_p-1)(Z_q-1)}{n-(p\wedge q)+1}\right|.$$

We have

$$\begin{aligned} Ee^{itL_n}(Z_k - 1)(Z_1 - 1) &= \textstyle\prod_{j=1}^n \eta\left(\frac{\alpha_{jn}\cdot t}{s_n}\right)\cdot h_1\left(\frac{\alpha_{kn}\cdot t}{s_n}\right) \\ &\quad \times h_1\left(\frac{\alpha_{1n}\cdot t}{s_n}\right) \qquad \text{for} \quad k \neq 1 \\ &= \textstyle\prod_{j=1}^n \eta\left(\frac{\alpha_{jn}\cdot t}{s_n}\right)\cdot h_2\left(\frac{\alpha_{kn}\cdot t}{s_n}\right) \qquad \text{for} \quad k = 1, \end{aligned}$$

where

$$h_1(t) = \frac{it}{1-it} \qquad \text{and} \qquad h_2(t) = \frac{1-t^2}{(1-it)^2}.$$

Note that $\eta(t) = [e^{it}(1-it)]^{-1}$. Differentiating $\eta(t)$ we find that $E(Z-1)^k e^{it(Z-1)} = h_k(t)\eta(t)$, $k = 1, 2$.

It is easily seen that $|h_1(t)| \leqq |t|$ and that $|h_1|$ and $|h_2|$ are both bounded by 1 for all real t. Thus we get,

$$|Ee^{itL_n}Q_n| \leqq |\tilde{\Phi}_n(t)|\cdot s_n{}^{-1}\left|\textstyle\sum_{p\neq q}\frac{t^2\alpha_{pn}\cdot\alpha_{qn}\beta_{p\vee q,n}}{s_n^2(n-(p\wedge q)+1)} + \textstyle\sum_{p=1}^n\frac{\beta_{pn}}{n-p+1}\right|.$$

The α's and the β's remain bounded by, say, A and B respectively. This is seen for the β's the same way as for the α's in (2.12). Note also that $\beta_{in} = 0$ for $i > \beta\cdot n$. The quantity inside the last set of absolute value signs is therefore bounded by $t^2A^2B/\theta(1-\beta) + B\sum_{p=[(1-\beta)\cdot n]}^n p^{-1}$. This proves (ii) since the last sum is bounded. We proceed to show (iii). Now,

$$\begin{aligned} |Ee^{itL_n}Q_n{}^m| &\leqq [n(1-\beta)s_n]^{-m}\textstyle\sum_{i_1,\cdots,i_m=1;j_1,\cdots,j_m=1}^n E|\prod_{j=1}^n \exp[it\alpha_{jn}(Z_j-1)s_n{}^{-1}] \\ &\quad \times \textstyle\prod_{k=1}^m \beta_{i_k\vee j_k,n}(Z_{i_k}-1)(Z_{j_k}-1)| \end{aligned}$$

[3] $\#S$ = number of elements in S.

$$\leq [n(1-\beta)s_n]^{-m} \sum_{i_1,\cdots,i_m=1;j_1,\cdots,j_m=1}^{n} \prod_{j\neq i_l,j_l;l=1,\cdots,m}^{n} \left|\eta\left(\frac{\alpha_{jn}t}{s_n}\right)\right|$$

$$\times E|\prod_{k=1}^{m}(Z_{i_k}-1)(Z_{j_k}-1)| \prod_{k=1}^{m}|\beta_{i_k \vee j_k,n}|$$

$$\leq [n(1-\beta)s_n]^{-m} \sup_{S\subset\{1,\cdots,n\};\#S\geq n-2m} \prod_{j\in S} \eta\left(\frac{\alpha_{jn}t}{s_n}\right)\Big|$$

$$\times E|Z_1-1|^{2m}[\sum_{p,q=1}^{n}|\beta_{p\vee q,n}|]^m$$

$$\leq [(1-\beta)^{-1}\theta^{-\frac{1}{2}}B]^m(2m)^{2m}n^{m/2} \sup_{S\subset\{1,\cdots,n\};\#S\geq n-2m} \prod_{j\in S}\left|\eta\left(\frac{\alpha_{jn}t}{s_n}\right)\right| .$$

We have here exploited the independence of the Z's. Note also that $[\sum_{p\cdot q}\beta_{p\vee q,n}|]^m \leq B^m n^{2m}$ and that $E|Z_1-1|^{2m} \leq 2m!$. Since

$$\log\eta(t) = -\frac{t^2}{2} - \frac{(it)^3}{3} + \cdots + \frac{(-it)^k}{k} + \cdots ,$$

we see that for $|t|$ smaller than some $\varepsilon > 0$, $\mathrm{Re}\,(\log\eta(t)) < -\tau^2t^2/2$, $\tau > 0$. This immediately gives

$$\mathrm{Re}\,(\log\tilde{\Phi}(t)) < \sum_{i=1}^{n} -\frac{\tau^2\alpha_{in}^2t^2}{2s_n^{\,2}} = -\frac{\tau^2t^2}{2} .$$

Also, since the α's are bounded by A, we have

$$\log \sup_{S\subset\{1,\cdots,n\};\#S\geq n-2m} \prod_{j\in S}\left|\eta\left(\frac{\alpha_{jn}\cdot t}{s_n}\right)\right|$$

$$\text{(2.16)}\qquad < \sup_S \sum_{j\in S} -\frac{\tau^2\alpha_{jn}^2t^2}{2s_n^{\,2}}$$

$$= \sup_S -\frac{\tau^2t^2}{2}\left(1-\sum_{j\notin S}\frac{\alpha_{jn}^2}{s_n^{\,2}}\right) \leq \left(1-\frac{2mA^2}{\theta\cdot n}\right)\frac{\tau^2t^2}{2} .$$

From (2.15) with $m = 2$ and Lemma 2.2, we now have

$$|\Phi_n(t)-\tilde{\Phi}_n(t)| \leq |t|c_2n^{-\frac{1}{2}}e^{-\tau^2t^2/2}(t^2+1) + 2\cdot t^2\cdot n^{-1}\cdot c_1^{\,2} .$$

We conclude that

$$\text{(2.17)}\qquad \int_{-n^{\frac{1}{2}}}^{n^{\frac{1}{2}}} \frac{|\Phi_n(t)-\tilde{\Phi}_n(t)|}{|t|}\,dt = O(n^{-\frac{1}{2}}) .$$

On $\langle -\varepsilon\cdot n^{\frac{1}{2}}, \varepsilon\cdot n^{\frac{1}{2}}\rangle$, we have from Lemma 2.2,

$$\frac{|t|^m}{m!}E|Q_n|^m \leq \frac{\varepsilon^m n^{m/2}}{m!}\cdot n^{-m/2}m^mc_1^{\,m}$$

$$\leq (e\cdot\varepsilon\cdot c_1)^m , \qquad \text{since}\quad m! \geq m^m\cdot e^{-m} .$$

Choose $\varepsilon = p/(e\cdot c_1)$, $p < 1$ and let $m = [\log n/|\log p| + 1] \wedge n$. Then

$$\text{(2.18)}\qquad \frac{|t|^m}{m!}E|Q_n|^m = O(n^{-1}) \qquad \text{on}\quad \langle -\varepsilon\cdot n^{\frac{1}{2}}, \varepsilon\cdot n^{\frac{1}{2}}\rangle .$$

We also have on $\langle -\varepsilon \cdot n^{\frac{1}{2}}, \varepsilon \cdot n^{\frac{1}{2}}\rangle$, using Lemma 2.2 combined with (2.16), that

$$\begin{aligned}\sum_{j=1}^{m-1} \left|\frac{t^j}{j!} Ee^{itL_n}Q_n^{\ j}\right| &\leqq \sum_{j=1}^{m-1} \frac{|t|^j}{j!} \cdot c_3^{\ j} \cdot j^{2j} n^{j/2} \exp\left[-\left(1 - \frac{A^2 2m}{\theta \cdot n}\right)\frac{\tau^2 t^2}{2}\right] \\ &\leqq \sum_{j=1}^{m-1} (c_3 \cdot e \cdot j \cdot \varepsilon \cdot n)^j \exp\left[-\left(1 - \frac{A^2 2m}{\theta \cdot n}\right)\frac{\tau^2}{2} \cdot n\right] \qquad (2.19) \\ &\leqq (c_3 \cdot e \cdot m \cdot \varepsilon \cdot n)^m \exp\left[-\left(1 - \frac{A^2 2m}{\theta \cdot n}\right)\frac{\tau^2}{2} \cdot n\right] \\ &= o(n^{-1}) .\end{aligned}$$

Combining (2.17), (2.18) and (2.19), we see that (2.13) holds, and the theorem is proved.

As a corollary to Theorem 1, we will prove another version of this theorem that better lends itself to applications.

We now assume that T_n can be written in the following form, which is a special case of (2.1):

$$T_n = n^{-1} \sum_{j=1}^{n} J\left(\frac{j}{n+1}\right) h(X_{jn}) + \sum_{i=1}^{r} a_i h(X_{[p_i \cdot h], n}) . \qquad (2.20)$$

A finite number, r, of quantiles are here given special weights $a_i \neq 0$. (2.20) includes statistics such as the trimmed mean, the Winsorized mean and systematic statistics.

Chernoff et al. (1967), among others, showed under certain conditions, that T_n is asymptotically normal with mean

$$\mu = \int_0^1 J(u)G(u)\, du + \sum_{i=1}^{r} a_i G(p_i)$$

and variance

$$\sigma^2 = \int_0^1 \alpha^2(u)\, du ,$$

where $G = h \circ F^{-1}$ and

$$\alpha(u) = (1-u)^{-1}\{\int_u^1 J(w)G'(w)(1-w)\, dw + \sum_{i:p_i \geqq u} a_i(1-p)G'(p_i)\} .$$

We will prove the following corollary, notation being the same as in Theorem 1.

COROLLARY. *If J and G'' satisfy a first order Lipschitz condition on an open interval containing $I = [\alpha, \beta]$ $0 < \alpha < \beta \lessdot 1$, J vanishes outside I and if $p_i \in I$; $i = 1, \cdots, r$, then*

$$\sup_x \left| P\left(n^{\frac{1}{2}} \frac{T_n - \mu}{\sigma} < x\right) - \Phi(x)\right| \leqq \frac{K}{n^{\frac{1}{2}}} \qquad (2.21)$$

for some constant K. If $r = 0$ it is assumed that $J(x) > 0$ on some open interval.

PROOF. T_n may certainly be written in the form (1.1), and μ_n, α_{in} and $s_n^{\ 2}$ are defined as previously.

First, we assure that the conditions of this theorem imply those of Theorem

1. With $H(x) = G(1 - e^{-x})$, H'' will satisfy a first order Lipschitz condition, on an interval $[a, b]$ where $0 < a < -\log(1-\alpha) < -\log(1-\beta) < b$. As will be shown below, $|s_n^2/n - \sigma^2| \leqq c/n$, so that (2.7) is satisfied at least if n is large enough. J is bounded on I since it satisfies a first order Lipschitz condition. We therefore have constants c_1 and c_2 such that

$$\begin{aligned} c_{in} &= J\left(\frac{i}{n+1}\right) \leqq c_1\,, \qquad i \neq [p_j h]\,,\quad j = 1, \cdots, r\,; \quad n = 1, 2, \cdots \\ &= J\left(\frac{i}{n+1}\right) + n \cdot a_i \leqq n \cdot c_2\,, \\ &\qquad\qquad i = [p_j h]\,,\quad j = 1, \cdots, r\,; \quad n = 1, 2, \cdots\,. \end{aligned}$$

We can conclude that (2.8) holds for the c_{in} above. Thus, (2.9) holds for the distribution F_n, of $n(T_n - \mu_n)/s_n$.

LEMMA 2.3. *Under the conditions of the theorem,*

$$\mu_n = \mu + O\left(\frac{1}{n}\right) \tag{2.22}$$

and

$$\frac{s_n^2}{n} = \sigma^2 + O\left(\frac{1}{n}\right). \tag{2.23}$$

PROOF. Define $\tilde{\nu}_{jn} = -\log(1 - j/(n+1))$. As in Chernoff et al. (1967), it is easy to see that

$$0 \leqq \nu_{jn} - \tilde{\nu}_{jn} < \frac{j}{(2n+1)(n-j+\frac{1}{2})}\,.$$

Thus, there is a constant c, for which

$$|\nu_{jn} - \tilde{\nu}_{jn}| \leqq \frac{c}{n} \qquad \text{for all } j \text{ such that } \frac{j}{n+1} \in I\,. \tag{2.24}$$

Therefore, since H has a bounded derivative on I, $H(\nu_{jn}) - H(\tilde{\nu}_{jn}) = H(\nu_{jn}) - G(j/(n+1)) = O(n^{-1})$. Define

$$\tilde{\mu}_n = \frac{1}{n} \sum_{j=1}^{n} J\left(\frac{j}{n+1}\right) G\left(\frac{j}{n+1}\right) + \sum_{i=1}^{r} a_i G\left(\frac{[p_i \cdot n]}{n+1}\right).$$

Rewriting

$$\mu_n = n^{-1} \sum c_{in} H(\nu_{in}) \quad \text{as}$$
$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) H(\nu_{in}) + \sum_{i=1}^{r} a_i H(\nu_{[p_i \cdot n], n})\,,$$

we see that

$$\mu_n - \tilde{\mu}_n = O(n^{-1})\,. \tag{2.25}$$

We have further,

$$\tilde{\mu}_n - \mu = \frac{1}{n} \sum_{j=1}^{n} \int_{(j-1)/n}^{j/n} \left(J\left(\frac{j}{n+1}\right) G\left(\frac{j}{n+1}\right) - J(u)G(u) \right) du$$
$$+ \sum_{i=1}^{r} a_i \left(G\left(\frac{[p_i \cdot n]}{n+1}\right) - G(p_i) \right).$$

Since J and G both satisfy a first order Lipschitz condition, we have

$$J\left(\frac{j}{n+1}\right) G\left(\frac{j}{n+1}\right) - J(u)G(u)$$
$$= \frac{1}{2}\left(J\left(\frac{j}{n+1}\right) - J(u)\right)\left(G\left(\frac{j}{n+1}\right) + G(u)\right)$$
$$+ \frac{1}{2}\left(G\left(\frac{j}{n+1}\right) - G(u)\right)\left(J\left(\frac{j}{n+1}\right) + J(u)\right)$$
$$= O(n^{-1}) \qquad \text{for} \quad u \in \left[\frac{j-1}{n}, \frac{j}{n}\right] \quad \text{uniformly in} \quad j$$

and

$$G\left(\frac{[p_i \cdot n]}{n+1}\right) - G(p_i) = O(n^{-1}).$$

Thus,

$$\tilde{\mu}_n - \mu = O(n^{-1}), \tag{2.26}$$

and we can conclude (2.22).

We now prove the second part of the lemma.

$$s_n^2 = \sum_{i=1}^{n} \alpha_{in}^2,$$

where

$$\alpha_{in} = \frac{1}{n-i+1} \sum_{j=1}^{n} c_{jn} H'(\nu_{jn})$$
$$= \frac{1}{n-i+1} \left\{ \sum_{j=i}^{n} J\left(\frac{j}{n+1}\right) H'(\nu_{jn}) + n \cdot \sum_{j: p_j \cdot n \geq i}^{r} a_j H'(\nu_{jn}) \right\}.$$

Replacing ν_{jn} by $\tilde{\nu}_{jn}$ and noting that $H'(u) = G'(u)(1-u)$, we get

$$\alpha'_{in} = \frac{1}{n-i+1} \left\{ \sum_{j=i}^{n} J\left(\frac{j}{n+1}\right) G'\left(\frac{j}{n+1}\right)\left(1 - \frac{j}{n+1}\right) \right.$$
$$\left. + n \cdot \sum_{j: p_j n \geq i}^{r} a_j G'\left(\frac{j}{n+1}\right)\left(1 - \frac{j}{n+1}\right) \right\}.$$

Since G' also satisfies a first order Lipschitz condition, (2.24) gives

$$\alpha'_{in} - \alpha_{in} = O(n^{-1}) \qquad \text{uniformly in} \quad i. \tag{2.27}$$

Also,

$$\alpha'_{in} - \alpha\left(\frac{i}{n}\right) = \frac{n}{n-i}\left[\sum_{j=1}^{n-1} \int_{j/n}^{(j+1)/n} \left\{ \frac{n-i}{n-i+1} J\left(\frac{j}{n+1}\right) G'\left(\frac{j}{n+1}\right) \right.\right.$$
$$\left. \times \left(1 - \frac{j}{n+1}\right) - J(u)G'(u)(1-u) \right\} du$$

$$+ \frac{n-i}{n-i+1} \cdot \frac{1}{n} J\left(\frac{n}{n+1}\right) G'\left(\frac{n}{n+1}\right)\left(1 - \frac{n}{n+1}\right)$$

$$+ \frac{n-i}{n-i+1} \sum_{j:p_j \cdot n \geqq i}^{r} a_j G'\left(\frac{j}{n+1}\right)\left(1 - \frac{j}{n+1}\right)$$

$$- \sum_{j:p_j \geqq i/n}^{r} a_j G'(p_j)(1-p_j)\Big] .$$

The functions J, G' and $(1-x)$ each satisfy a first order Lipschitz condition, and therefore $J(x)G'(x)(1-x)$ also does. We can conclude that

$$\alpha'_{in} - \alpha\left(\frac{i}{n}\right) = O(n^{-1}) \qquad \text{uniformly in } i . \tag{2.28}$$

We can now write

$$\frac{s_n^2}{n} - \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (\alpha_{in}^2 - \alpha_{in}'^2) + \frac{1}{n} \sum_{i=1}^{n} \left(\alpha_{in}'^2 - \alpha\left(\frac{i}{n}\right)^2\right)$$

$$+ \sum_{i=1}^{n} \int_{(i-1)/n}^{i/n} \left(\alpha(u)^2 - \alpha\left(\frac{i}{n}\right)^2\right) du .$$

Since the α_{in}, α'_{in} and the function $\alpha(u)$ are all bounded, we have, for some constant K,

$$\frac{1}{K}\left|\frac{s_n^2}{n} - \sigma^2\right| \leqq \frac{1}{n} \sum_{i=1}^{n} |\alpha_{in} - \alpha'_{in}| + \frac{1}{n} \sum_{i=1}^{n} \left|\alpha'_{in} - \alpha\left(\frac{i}{n}\right)\right|$$

$$+ \sum_{i=1}^{n} \int_{(i-1)/n}^{i/n} \left|\alpha(u) - \alpha\left(\frac{i}{n}\right)\right| du .$$

It follows from (2.27) and (2.28), that the first two summands are of the order $1/n$. $\alpha(u)$ is continuous with uniformly bounded derivative, except at the points p_i; $i = 1, \cdots, r$, where it has jumps of size $a_i(1-p_i)G'(p_i)$. These jumps are all less than some $D < \infty$. The last sum, taken over all indices i, for which $p_j \notin [(i-1)/n, i/n]$; $j = 1, \cdots, r$, is of the order $1/n$ and the remainder is less than $r/n \cdot D$. We can thus conclude (2.23) and the lemma is proved.

The equalities that follow are valid uniformly in x. Using (2.22), the result of Theorem 1 and defining $\sigma_n = n^{-\frac{1}{2}} s_n$, we can write

$$P\left(n^{\frac{1}{2}} \frac{T_n - \mu}{\sigma} < x\right) = \Phi\left(x + n^{\frac{1}{2}} \frac{\mu_n - \mu}{\sigma_n}\right) + O\left(\frac{1}{n^{\frac{1}{2}}}\right)$$

$$= \Phi(x) + O\left(\frac{1}{n^{\frac{1}{2}}}\right).$$

Secondly, we have

$$P\left(n^{\frac{1}{2}} \frac{T_n - \mu}{\sigma} < x\right) = \Phi\left(\frac{\sigma}{\sigma_n} x\right) + O\left(\frac{1}{n^{\frac{1}{2}}}\right)$$

$$= \Phi(x) + x \frac{\sigma - \sigma_n}{\sigma_n} \varphi\left(x + \theta \frac{\sigma - \sigma_n}{\sigma_n}\right) + O\left(\frac{1}{n^{\frac{1}{2}}}\right)$$

where $\varphi = \Phi'$.

Since $x\varphi(x + y)$ is bounded for all x and y, (2.23) gives

$$P\left(n^{\frac{1}{2}} \frac{T_n - \mu}{\sigma} < x\right) = \Phi(x) + O\left(\frac{1}{n^{\frac{1}{2}}}\right).$$

Acknowledgments. I am greatly indebted to Professor Peter J. Bickel for having suggested the problems and for his patient and continual guidance throughout this investigation. Thanks are also due to the referees for many valuable comments.

REFERENCES

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

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

[3] DAVID, H. A. (1970). *Order Statistics*. Wiley, New York.

[4] ESSEEN, C. G. (1944). Fourier analysis of distribution functions. *Acta Math.* **77** 30-80.

[5] FELLER, W. (1957). *An Introduction to Probability Theory and Its Applications* **1**, 2nd ed. Wiley, New York.

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

[7] HOEFFDING, W. (1963). Probability inequalities for sums of independent random variables. *J. Amer. Statist. Assoc.* **58** 13-30.

[8] ROSENKRANTZ, W. and O'REILLY, N. (1972). Application of the Skorohod representation theorem to rates of convergence for linear combinations of order statistics. *Ann. Math. Statist.* **43** 1204-1212.

INSTITUTE OF MATHEMATICS
UNIVERSITY OF OSLO
BLINDERN, OSLO 3
NORWAY