Work supported by (i) the Army Research Office, Durham, Grant ARO-D-DA-31-124-G746 and (ii) Office of Naval Research, Contract No. N00014-68-A-0187.

LIMITING BEHAVIOUR OF THE EXTREMUM OF CERTAIN SAMPLE FUNCTIONS

by

P. K. SEN, B. B. BHATTACHARYYA AND M. W. SUH

University of North Carolina, Chapel Hill, and
North Carolina State University, Raleigh

Institute of Statistics Mimeo Series No. 628

June 1969

1. Introduction and Summary. Consider a bundle of n parallel filaments of equal length and let the non-negative random variables $X_1, X_2, \ldots X_n$ denote the strength of individual filaments and $Y_1 \leq Y_2 \leq \ldots \leq Y_n$ represent the corresponding ordered random variables. Now if we assume that the force of a free load on the bundle is distributed equally on each individual filament and the strength of an individual filament is independent of the number of filaments in a bundle, then the minimum load B_n beyond which all the filaments of the bundle give way is defined to be the strength of the bundle.

Now if a bundle breaks under load L then all the inequalities $L/n \geq Y_1$, $L/n-1 \geq Y_2$, $\ldots$ $L \geq Y_n$ are simultaneously satisfied. Consequently the bundle strength can be represented as

$$B_n = \max\{nY_1, (n-1)Y_2, \ldots Y_n\} \quad . \tag{1.1}$$

Daniels [1] has investigated the probability distribution of B_n and has shown by very elaborate and complicated analysis that the large sample distribution of B_n, properly standardized converges to N(0,1) distribution when $\{X_i\}$ is a sequence of iid (independent and identically distributed) random variables having continuous distribution function and certain other regularity conditions are satisfied.

The major objective of the present paper is to obtain similar result by probabilistic argument for a class of statistics of the form $\sup_x\{\psi(x,S_n(x))\}$ to which B_n belongs (under centain simple conditions) where $S_n(x)$ is the empirical distribution function of a sample of size n from an m-dependent stochastic process and ψ is a member of a suitably chosen family of non-negative valued functions. In sections 2, 3, and 4, the basic regularity conditions and some preliminary results are derived. In section 5, the asymptotic normality of the statistic defined in section 3 is proved. In section 6, under a concavity assumption, moment convergence of the statistic is proved. In section 7, we have discussed the special case when $\psi(x,S_n(x)) = x(1-S_n(x))$ which in large sample is equivalent to B_n/n if $\max_{1 \leq i \leq n} X_i/n$ converges to zero in probability. We have shown that if $\{X_i\}$ is a sequence of non-negative interchangeable random variables then B_n/n forms a reverse semi-martingale sequence if $E(X_i) < \infty$. Moreover if $\{X_i\}$ is a sequence non-negative iid random variables with its moment generating function $M(t) < \infty$ for all $|t| < T$ then for any $0 < \alpha < 1$,

$$n^{\alpha} \left|E\{\sup_x x(1-S_n(x)) - y_o\}\right| \to 0 \text{ as } n \to \infty$$

where y_o is maximum of the function $X(1-F(x))$ attained at a unique point x_o. In section 8, some possible generalizations have been discussed.

2. Statement of the problem. Let $\{X_1, X_2, \ldots\}$ be a sequence of random variables forming an m-dependent stochastic process, not necessarily stationary. The marginal cdf of X_i is assumed to be continuous and is denoted by $F_i(x)$, and the joint cdf of (X_i, X_{i+h}) is denoted by $F_{i,h}(x,y)$, for $h = 1, \ldots, m$, $i = 1,2, \ldots$. The empirical cdf $S_n(x)$

when the sample is $(X_1, \ldots, X_n)$ is defined by

(2.1) $$S_n(x) = n^{-1} \Sigma_{i=1}^n c(x-X_i) \text{ where } c(u) = \{\begin{smallmatrix} 1, & u \geq 0 \\ 0, & u < 0 \end{smallmatrix}\} .$$

Also, the <u>average cdf</u> $\bar{F}_{(n)}(x)$ is defined by

(2.2) $$\bar{F}_{(n)}(x) = n^{-1} \Sigma_{i=1}^n F_i(x); \quad \text{thus } E\{S_n(x)\} = \bar{F}_{(n)}(x), \text{ for all } x.$$

Consider a non-negative and real valued function $h_n(x) = \psi(x, \bar{F}_{(n)}(x))$, where $\psi(x, \bar{F}_{(n)}(x))$ assumes a unique (and finite) maximum h_n^o at $x = x_n^o$. We assume that

(2.3) $$0 < \inf_n h_n^o \leq \sup_n h_n^o < \infty \text{ and } 0 < |\inf_n x_n^o|, |\sup_n x_n^o| < \infty .$$

Further, we let $p_n = \bar{F}_{(n)}(x_n^o)$, and assume that

(2.4) $$0 < \inf_n p_n \leq \sup_n p_n < 1.$$

Our primary concern is to provide a suitable estimator of h_n^o based on the sample $(X_i, \ldots, X_n)$. Since S_n unbiasedly estimates $\bar{F}_{(n)}$ at all points, we are intuitively led to the following estimator of h_n^o. Let

(2.5) $$Z_{ni} = \psi(X_i, S_n(X_i)), \; i = 1, \ldots, n;$$

(2.6) $$Z_n^* = \max_{1 \leq i \leq n} Z_{n,n} \quad \text{where } Z_{n,1} \leq \cdots \leq Z_{n,n} .$$

Our central problem is to derive the asymptotic normality of $n^{\frac{1}{2}}[Z_n^* - h_n^o]$. Since the random variates $Z_{n1}, \ldots, Z_{nn}$ are not all independent (even when $m = 0$), nor necessarily identically distributed, the usual techniques of deriving the distribution theory for sample maximum fails to provide our desired result.

Our task is accomplished here by showing that $n^{\frac{1}{2}}[Z_n^* - h_n^o]$ is proportional to $n^{\frac{1}{2}}[S_n(x_n^o) - \overline{F}_{(n)}(x_n^o)]$, in probability (as $n \to \infty$), and then using the central limit theorem for the later variable.

3. Basic regularity conditions and the main theorem. We assume that for all (x,y): $-\infty < x < \infty$, $0 < y < 1$, $\psi(x,y)$ if non-negative and absolutely continuous in x and y. Also,

(3.1) (i) $\psi(x,y+\delta) - \psi(x,y) = \delta\psi_{01}(x,y+\theta\delta),\ 0 < \theta < 1$,

where for all $|\delta| \leq \delta_0 (> 0)$, $\psi_{01}(x, y+\delta)$ is continuous in δ at $y = \overline{F}_{(n)}$,

(3.2) $|[\psi_{01}(x,y)]_{y = \overline{F}_{(n)}(x)}| \leq g(x),\ \lim_n \sup_{1 \leq i \leq n} \int_{-\infty}^{\infty} g^2(x)\, dF_i(x) < \infty$

and g(x) is continuous and uniformly integrable,

(ii) for all $(x,y) \varepsilon [x_n^o \pm \delta_1,\ p_n \pm \delta_2]$ where (δ_1, δ_2) are arbitrarily small,

(3.3) $\psi(x_n^o + h, p_n + k) = \psi(x_n^o, p_n) + h\,\psi_{10}(x_n^o, p_n) + k\,\psi_{01}(x_n^o, p_n)$
$+ o(|h| + |k|),\quad |h| \leq \delta_1,\ |k| \leq \delta_2$,

where ψ_{10} and ψ_{01} stand for the partial derivatives, and (iii)

(3.4) $0 < \inf_n |\psi_{01}(x_n^o, p_n)| \leq \sup_n |\psi_{01}(x_n^o, p_n)| < \infty$.

Since $\psi(x, \overline{F}_{(n)}(x))$ assumes a maximum at x_n^o, it follows from (2.4) that if $\overline{f}_{(n)}(x_n^o) = (d/dx)\,\overline{F}_{(n)}(x)\,|x_n^o > 0$ and is continuous at x_n^o; then

(3.5) $|\psi_{10}(x_n^o, p_n)| > 0$,

and if $\bar{f}_{(n)}(x)$ is continuous at x_n^o then $h_n(x) = \psi(x, \bar{F}_{(n)}(x))$ has a continuous first order derivative at $x = x_n^o$. Since $h_n(x)$ attains a maximum at x_n^o, for all x in the neighborhood of x_n^o we have

$$h_n(x) = h_n(x_n^o) + (x-x_n^o)\, h_n'\,(x_n^o\theta + (1-\theta)x),\ 0 < \theta < 1$$

(3.6) $$= h_n(x_n^o) + o(x-x_n^o) \quad \text{as } h_n'\,(x_n^o) = 0.$$

We also assume that

(3.7) $$\limsup_{n\to\infty} \{h_n'(x_n^o + \varepsilon_1) - h_n'\,(x_n^o - \varepsilon_2)\} < 0 \text{ for all } 0 < \varepsilon_1, \varepsilon_2 < \varepsilon_0.$$

Concerning $\{F_i\}$, we assume that F_i has a continuous density function $f_i(x)$ at $x = x_n^o$ for all $i = 1, \ldots, n$. Let then

(3.8) $$\bar{f}_{(n)j}(x_n^o) = n_j^{-1} \Sigma_{k=0}^{n_j-1} f_{j+k(m+1)}(x_n^o);\ n_j = [\frac{n+1+m-j}{n+1}],\ j=1,\ldots,m+1,$$

(3.9) $$\bar{f}_{(n)}(x_n^o) = n^{-1} \Sigma_{j=1}^{m+1} n_j \bar{f}_{(n)j}(x_n^o), \quad \text{where } (m+1)n_j \sim n\ .$$

Then we assume that

(3.10) $$0 < \inf_n \bar{f}_{(n)}(x_n^o) \leq \sup_n \bar{f}_{(n)}(x_n^o) < \infty\ .$$

Now, we require the following notations. Let

(3.11) $$p_{n,i} = F_i(x_n^o),\ i=1,\ldots,n;\ \bar{F}_{(n)h}(x,y) = \frac{1}{n-h} \Sigma_{i=1}^{n-h} F_{i,h}(x,y)\ ;$$

(3.12) $$\alpha_{n,h} = \bar{F}_{(n)h}(x_n^o, x_n^o) - p_n^2,\ \beta_{n,h} = \frac{1}{n-h} \Sigma_{i=1}^{n-h} \{p_{n,i}p_{n,i+h} - p_n^2\}$$

for $h = 0, 1, \ldots, m$, where of course $\alpha_{n,0} = p_n(1-p_n)$ and

$\beta_{n,0} = n^{-1} \sum_{i=1}^{n} (p_{n,i} - p_n)^2$. Then, let

(3.13) $$\nu^2_{n,m} = (\alpha_{n,0} - \beta_{n,0}) + 2\, \Sigma^m_{h=1} \{(n-h)/n\} \{\alpha_{n,h} - \beta_{n,h}\} ;$$

(3.14) $$\gamma^2_{n,m} = \nu^2_{n,m} [\psi_{01}(x^o_n, p_n)]^2 .$$

Finally we assume that excepting in the neighborhood of a finite number of points (Say, $\ell_n^{(1)} < \ldots < \ell_n^{(k)}$),

(3.15) $$n^{-1} \Sigma^n_{i=1} F_i(x)[1-F_i(x)] > 0 \text{ for all } 0 < \overline{F}_{(n)}(x) < 1 \text{ and all } n .$$

Then the main theorem of the paper is the following.

Theorem 3.1. Under the conditions stated above, if $\inf_n \gamma_{n,m} > 0$, then

(3.16) $$\lim_{n\to\infty} P\{n^{\frac{1}{2}}[Z^*_n - h^o_n]\, \gamma_{n,m} \le x\} = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp(-\tfrac{1}{2} t^2)\, dt ,$$

uniformly in $-\infty < x < \infty$.

The proof of the theorem is postponed to section 5.

4. Some preliminary results.

Let

(4.1) $$Y_{ni} = \psi(X_i, \overline{F}_{(n)}(x_i)), \quad i = 1, \ldots, n ;$$

$Y_{n,1} \le \ldots \le Y_{n,n}$ the order statistics. Now, $Y_{n1}, \ldots, Y_{nn}$ form an m-independent process with the marginal cdf's $G_1^{(n)}(x), \ldots, G_n^{(n)}(x)$, respectively. We have by assumption

(4.2) $$\sup_x \psi(x, \overline{F}_{(n)}(x)) = h^o_n < \infty \implies G_i^{(n)}(h^o_n) = 1 \text{ for } i = 1, \ldots, n .$$

Lemma 4.1. If $h_n(x) = \psi(x, \overline{F}_{(n)}(x))$ has a continuous first order derivative in the neighborhood of $x = x^o_n$, then for every $\varepsilon > 0$, $\lim_{n\to\infty} P\{Y_{n,n} \ge h^o_n - \varepsilon/n\} = 1$, when (3.6) holds.

Proof. By the hypothesis of the lemma, for any $\{a_n\}$ such that $a_n \to 0$ as $n \to \infty$,

$$(4.3) \qquad h_n(x_n^o + a_n) = h_n(x_n^o) + a_n\, h_n'(x_n^o + \theta a_n) \quad (0 < \theta < 1)$$

$$= h_n(x_n^o) + o(a_n) \text{ as } h_n'(x_n^o) = 0 .$$

Thus, for every $\varepsilon > 0$,

$$(4.4) \qquad G_i^{(n)}(h_n^o - \varepsilon/n) = P\,\{h_n(x) \geq h_n^o - \varepsilon/n\}$$

$$= P\,\{x_i \leq x_n^o - \varepsilon_1(n)\} + P\,\{x_i > x_n^o + \varepsilon_2(n)\}\ ,\ i = 1, \ldots, n,$$

where $\varepsilon_1(n)$ and $\varepsilon_2(n)$ both $\to 0$ as $n \to \infty$, but

$$(4.5) \qquad n\varepsilon_j(n) \to \infty\ ,\ j = 1,2, \text{ as } n \to \infty \Longrightarrow n\ \varepsilon(n) \to \infty \text{ as } n \to \infty ,$$

where $\varepsilon(n) = \varepsilon_1(n) + \varepsilon_2(n)$. From (4.2) and (4.4), we have

$$(4.6) \qquad 1-G_i^{(n)}(h_n^o - \varepsilon/n) = P\,\{x_i \leq x_n^o\} + P\,\{x_i > x_n^o\} - P\,\{x_i \leq x_n^o - \varepsilon_1(n)\}$$

$$- P\,\{x_i \geq x_n^o + \varepsilon_2(n)\} \simeq \varepsilon(n)\ f_i(x_n^o),\ i = 1, \ldots, n .$$

Let then

$$Y_{(n)}^{(j)} = \max\,[Y_{nj}, Y_{nj+m+1}, \ldots, Y_{nj+(n_j-1)(m+1)}],\ j = 1, \ldots, m+1.$$

Then, $Y_{n,n} \geq Y_{(n)}^{(j)}$ for all $j = 1, \ldots, m+1$. Thus,

$$(4.7) \qquad P\,\{Y_{n,n} \leq h_n^o - \varepsilon/n\} \leq \min_j P\,\{Y_{(n)}^{(j)} \leq h_n^o - \varepsilon/n\} .$$

Now, $Y_{(n)}^{(j)}$ is the maximum over n_j independent random variables. Hence,

$$(4.8) \qquad P\,\{Y_{(n)}^{(j)} \leq h_n^o - \varepsilon/n\} = \prod_{k=0}^{n_j-1} G_{j+k(m+1)}^{(n)}(h_n^o - \varepsilon/n)$$

$$\leq \left[\frac{1}{n_j} \sum_{k=0}^{n_j-1} G_{j+k(m+1)}^{(n)} (h_n^o - \varepsilon/n) \right]^{n_j}$$

$$\approx [1 - \varepsilon(n)\, \overline{f}_{(n)j}(x_n^o)]^{n_j}, \quad j = 1, \ldots, m+1 .$$

Since, $\sup_j \overline{f}_{(n)j}(x_n^o) \geq \overline{f}_{(n)}(x_n^o) > 0$, by (3.9) and $n\varepsilon(n) \to \infty$ as $n \to \infty$, we have

(4.9) $$\min_j P\{Y_{(n)}^{(j)} \leq h_n^o - \varepsilon/n\} \leq [1 - \varepsilon(n)\, \overline{f}_{(n)}(x_n^o)]^{n_j} \to 0 \text{ as } n \to \infty .$$

Hence, the lemma follows from (4.7) and (4.9).Q.E.D.

A direct consequence of lemma 4.1 is the following:

(4.10) $$n[h_n^o - Y_{n,n}] = o_p(1), \text{ uniformly in } \{F_i\} .$$

<u>Lemma 4.2</u>. <u>For every</u> $\varepsilon > 0$, there exists <u>a finite</u> $c(\varepsilon) > 0$ <u>such that</u>

$$P\{\sup_x n^{\frac{1}{2}} |S_n(x) - \overline{F}_{(n)}(x)| > c(\varepsilon)\} < \varepsilon ,$$

<u>uniformly in</u> $\{F_i\}$, <u>satisfying</u> (3.15).

<u>Proof</u>. Let us define

(4.11) $$S_{n,j}(x) = \frac{1}{n_j} \sum_{k=0}^{n_j-1} c(x - X_{j+k(m+1)}) \text{ and } \overline{F}_{(n)j}(x) = \frac{1}{n_j} \sum_{k=0}^{n_j-1} F_{j+k(m+1)}(x) ,$$

for $j = 1, \ldots, m+1$. Then, by definition,

(4.12) $$\sup_x |n^{\frac{1}{2}}\{S_n(x) - \overline{F}_{(n)}(x)\}| \leq \Sigma_{j=1}^{m+1} \left(\frac{n}{n_j}\right)^{-\frac{1}{2}} \{\sup_x |n_j^{\frac{1}{2}}\{S_{n,j}(x) - \overline{F}_{(n)j}(x)\}|\}.$$

Since $(m+1)n_j \sim n$, $j = 1, \ldots, m+1$, it suffices to prove the following theorem.

Theorem 4.3. Let $\{X_i\}$ be a sequence of independent random variables with continuous cdf's $\{F_i\}$, and let $S_n(x)$ and $\bar{F}_{(n)}(x)$ be defined as in (2.1) and (2.3). Then, if (3.15) holds, for all $\lambda > 0$,

$$\lim_{n\to\infty} P\{\sup_x n^{\frac{1}{2}}|S_n(x) - \bar{F}_{(n)}(x)| > \lambda\} \leq 2\sum_{k=1}^{\infty}(-1)^{k+1}e^{-2k^2\lambda^2}, \tag{4.13}$$

where the equality sign holds when $F_1 = \ldots = F_n = F$ for all n.

Proof. Let us write $Y_{ni} = \bar{F}_{(n)}(X_i)$[1] and $G_{ni}(t) = P\{Y_{ni} \leq t\}$, $0 \leq t \leq 1$, $i = 1, \ldots, n$. Thus, $n^{-1}\Sigma_{i=1}^{n} G_{ni}(t) = t$: $0 \leq t \leq 1$. Define $G_n^*(t) = n^{-1}\Sigma_{i=1}^{n} c(t - Y_{ni})$ and $G_{ni}(t) = t_i$, $i = 1, \ldots, n$. Also let

$$U_n(t) = n^{\frac{1}{2}}[G_n^*(t) - t],\ 0 \leq t \leq 1. \tag{4.14}$$

Thus, it suffices to show that for every $\lambda > 0$,

$$\lim_{n\to\infty} P\{\sup_{0\leq t\leq 1} |U_n(t)| > \lambda\} \leq 2\sum_{k=1}^{\infty}(-1)^{k+1}e^{-2k^2\lambda^2}, \tag{4.15}$$

uniformly in $\{F_i\}$ satisfying (3.15). Direct computations yield

$$E\{U_n(t)\} = 0,\ E\{U_n(s)U_n(t)|s \leq t\} = n^{-1}\Sigma_{i=1}^{n}s_i(1-t_i),\ 0 \leq s \leq t \leq 1; \tag{4.16}$$

$$E\{U_n(t) - U_n(s)\}^2 = n^{-1}\Sigma_{i=1}^{n}(t_i - s_i)(1 - t_i + s_i) \leq (t - s), \tag{4.17}$$

$$E\{U_n(t) - U_n(s)\}^4 \leq 3(t - s)^2 \text{ for all } 0 \leq s \leq t \leq 1. \tag{4.18}$$

---

[1] Note that Y_{ni} has nothing to do with the Y_{ni} of (4.1) and will not be used later on.

Since the sample paths of $U_n(t)$ are not continuous (a.e.), we define

$$(4.19) \qquad U_n^*(t) = U_n(\tfrac{k-1}{n+1}) + [(n+1)t - (k-1)][U_n(\tfrac{k}{n+1}) - U_n(\tfrac{k-1}{n+1})] \ ,$$

for $(k-1)/(n+1) \le t < k/(n+1)$, $k = 1, \ldots, n+1$. Then $U_n^*(t)$ is a process with continuous sample paths, and using (4.18) we find

$$(4.20) \qquad E\{U_n^*(t) - U_n^*(s)\}^4 \le 27(t-s)^2, \ 0 \le s \le t \le 1 \ .$$

Hence by a theorem of Kolmogorov (cf. Hájek and Šidák [3, pp. 177-179]), it follows that for every $\varepsilon > 0$

$$(4.21) \qquad \lim_{\delta \to 0} P\{\max_{|t-s|<\delta} |U_n^*(t) - U_n^*(s)| < \varepsilon\} = 1.$$

Using (4.14) and (4.19), we have

$$(4.22) \qquad \sup_{0 \le t \le 1} |U_n(t) - U_n^*(t)| \le \max_{1 \le k \le n+1} |U_n^*(\tfrac{k}{n+1}) - U_n^*(\tfrac{k-1}{n+1})| + 2(n+1)^{-\frac{1}{2}}$$

and hence, from (4.21) it follows that for every $\varepsilon > 0$,

$$(4.23) \qquad \liminf_n P\{\sup_{0 \le t \le 1} |U_n(t) - U_n^*(t)| < \varepsilon\} = 1.$$

It therefore follows from (4.15) and (4.23) that we are only to show that

$$(4.24) \qquad \lim_{n \to \infty} P\{\sup_{0 \le t \le 1} |U_n^*(t)| > \lambda\} < 2 \sum_{k=1}^{\infty} (-1)^{k+1} e^{-2k^2\lambda^2} \ .$$

For every $n (\ge 1)$, let now $[Z_n(t);\ 0 \le t \le 1]$ be a Gaussian process with $EZ_n(t) = 0$ and $E[Z_n(s)Z_n(t) | s \le t] = n^{-1} \Sigma_{i=1}^n s_i(1-t_i)$, $0 \le s \le t \le 1$. [This sequence of Gaussian processes can be conceived of as an average of n independent Gaussian processes.] Since, by definition

$E[Z_n(t) - Z_n(s)]^2 = n^{-1} \Sigma_{i=1}^n (t_i - s_i)(1 - t_i + s_i) \le (t - s)$ and

$E[Z_n(t) - Z_n(s)]^4 = 3\{E[Z_n(t) - Z_n(s)]^2\}^2 \leq 3(t - s)^2$, according to the same theorem of Kolmogorov (cf. Hájek and Šidák [3, p. 177]), such a process exists (in the space of all continuous functions). Now, for any $m(\geq 1)$ and real $\underset{\sim}{\lambda} = (\lambda_1, \ldots, \lambda_m)$, let $H_n(x; \underset{\sim}{\lambda}, \underset{\sim}{t})$ be the cdf of $\sum_1^m \lambda_j U_n^*(t_j)$, where $0 \leq t_1 < \ldots < t_m \leq 1$. Also, let $\Phi_n(x; \underset{\sim}{\lambda}, \underset{\sim}{t})$ be the cdf of $\sum_1^m \lambda_j Z_n(t_j)$. Then by (4.23) and the multivariate central limit theorem we have

$$\sup_{-\infty<x<\infty} |H_n(x; \underset{\sim}{\lambda}, \underset{\sim}{t}) - \Phi_n(x; \underset{\sim}{\lambda}, \underset{\sim}{t})| \to 0 \text{ as } n \to \infty, \tag{4.25}$$

for all $\underset{\sim}{\lambda}$ and $\underset{\sim}{t}$. Hence, on making use of (4.23), (4.25) and a theorem in Hájek and Šidék [3, p. 180], we may conclude that $||U_n^*(t) - Z_n(t)||$ converges in the Prohorov-sense to zero i.e.,

$$\sup_{0\leq t\leq 1} |U_n^*(t) - Z_n(t)| \text{ converges in law to zero as } n \to \infty . \tag{4.26}$$

Hence, it suffices to show that for every $\lambda > 0$

$$\lim_n P\{\sup_{0\leq t\leq 1} |Z_n(t)| > \lambda\} \leq 2 \sum_{k=1}^{\infty} (-1)^{k+1} e^{-2k^2\lambda^2} . \tag{4.27}$$

Consider now the Brownian bridge $[Z(t), 0 \leq t \leq 1]$ with $EZ(t) = 0$ and $E[Z(s)Z(t)|s \leq t] = s(1-t)$, $0 \leq s \leq t \leq 1$. For finitely many points $0 < t^{(1)} < \ldots < t^{(m)} < 1$, let $\underset{\sim}{D}_n^{(m)}$ and $\underset{\sim}{D}^{(m)}$ be the covariance matrices for $Z_n(t)$ and $Z(t)$ respectively. It is then easy to verify that

$$\underset{\sim}{D}^{(m)} - \underset{\sim}{D}_n^{(m)} = ((n^{-1} \sum_1^n (t_i^{(j)} - t^{(j)})(t_i^{(\ell)} - t^{(\ell)})))_{j,\ell=1,m} = \underset{\sim}{D}_n^{*(m)}, \tag{4.28}$$

where $\underset{\sim}{D}_n^{*(m)}$ is positive semi-definite (p.s.d.) and $\underset{\sim}{D}^{(m)}$ is p.d.

Also, by the well-known Kolmogorov theorem

$$\lim_{m\to\infty} P\{\max_{1\leq j\leq m} |Z(t^{(j)})| > \lambda\} \geq 2 \sum_{k=1}^{\infty} (-1)^{k+1} e^{-2k^2\lambda^2} . \tag{4.29}$$

So the desired result will follow if we can show that for every

m and $1 \leq t^{(1)} < \ldots < t^{(m)} \leq 1$,

$$P\{\max_{1\leq j\leq m} |Z_n(t^{(j)})| > \lambda\} \geq P\{\max_{1\leq j\leq m} |Z(t^{(j)})| > \lambda\} . \tag{4.30}$$

(4.30) really follows from (4.28) and the following lemma.

Lemma 4.4. Let $\mathcal{C}_p$ be a convex space (in p-dimensions) with the origin origin as an inner point. Let $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ be independent normally distributed random variables (p-vectors) with null means and dispersion matrices $\underset{\sim}{B}_1$ and $\underset{\sim}{B}_2$ respectively, where $\underset{\sim}{B}_1$ is p.d. and $\underset{\sim}{B}_2$ and $\underset{\sim}{B}_3 = \underset{\sim}{B}_1 - \underset{\sim}{B}_2$ are at least p.s.d. Then

$$P\{\underset{\sim}{X} \in \mathcal{C}_p\} \leq P\{\underset{\sim}{Y} \in \mathcal{C}_p\} . \tag{4.31}$$

where the equality holds only when $\underset{\sim}{B}_3$ is a null matrix.

Proof. There exists a non-singular matrix $\underset{\sim}{D}$ such that $\underset{\sim}{D}\underset{\sim}{B}_1\underset{\sim}{D}' = \underset{\sim}{I}_p$ and $\underset{\sim}{D}\underset{\sim}{B}_2\underset{\sim}{D}' = \underset{\sim}{\nu}_p = \text{Diag}(\nu_1, \ldots, \nu_p)$, where $\nu_1, \ldots, \nu_p$ are the characteristic roots of $\underset{\sim}{B}_1^{-1} \underset{\sim}{B}_2$. By hypothesis $0 \leq \nu_1, \ldots, \nu_p \leq 1$. Under the mapping $X \to \underset{\sim}{D}X = X^*$, $Y \to \underset{\sim}{D}Y = Y^*$, let $\mathcal{C}_p^*$ be the image of $\mathcal{C}_p$. Since $\mathcal{C}_p$ is convex so is $\mathcal{C}_p^*$. Let then $\nu_1, \ldots, \nu_q > 0$, $q \leq p$, while the remaining p - q of the ν's be equal to 0. Then, for the q-dimensional distribution of $\underset{\sim}{X}_q^*$ and $\underset{\sim}{Y}_q^*$, let C_{pq}^* be the intersection of C_p^* with the q-dimensional hyperplane $X_{q+1}^* = \ldots = X_p^* = 0$. Since $\nu_1, \ldots, \nu_q$ are all positive and $X_1^*, \ldots, X_q^*$ (or $Y_1^*, \ldots, Y_q^*$) are all independent, well-known results on the multivariate normal distribution yield that

$$P\{\underset{\sim}{X}_q^* \varepsilon \mathcal{C}_{pq}^*\} \leq P\{\underset{\sim}{Y}_q^* \varepsilon \mathcal{C}_{pq}^*\} , \tag{4.32}$$

where the equality sign holds only when $\nu_1 = \ldots = \nu_q = 1$. Since $\mathcal{C}_{pq}^*$ is a sub-space of $\mathcal{C}_p^*$, $P\{\underset{\sim}{X}_q^* \varepsilon \mathcal{C}_{pq}^*\} \geq P\{\underset{\sim}{X}^* \varepsilon \mathcal{C}_p^*\} = P\{\underset{\sim}{X} \varepsilon \mathcal{C}_p\}$, while by virtue of $\nu_{q+1} = \ldots = \nu_p = 0$, $Y_{q+1}^* = \ldots = Y_p^* = 0$, with probability 1, and hence, $P\{\underset{\sim}{Y}_q^* \varepsilon \mathcal{C}_{pq}^*\} = P\{\underset{\sim}{Y}^* \varepsilon \mathcal{C}_p^*\} = P\{\underset{\sim}{Y} \varepsilon \mathcal{C}_p\}$. Hence, (4.31) follows from (4.32). Q.E.D.

Now, by definition, $Y_{n,n}$ in (4.1) is equal to $\psi(X_{(r)}, \overline{F}_{(n)}(X_{(r)}))$, where $1 \leq r \leq n$ and $X_{(1)} \leq \ldots \leq X_{(n)}$ are the order statistics.

<u>Lemma 4.5</u>. $(r/n - p_n) \overset{p}{\to} 0$, as $n \to \infty$.

<u>Proof</u>. By the continuity of $h_n(x)$ at x_n^o, (4.1) , (4.10) and (3.7), we have $|X_{(r)} - x_n^o| \overset{p}{\to} 0$. Hence,

$$\overline{F}_{(n)}(X_{(r)}) - \overline{F}_{(n)}(x_n^o) \overset{p}{\to} 0, \text{ as } n \to \infty . \tag{4.33}$$

Now, upon noting that $S_n(X_{(r)}) = r/n$ and $\overline{F}_{(n)}(x_n^o) = p_n$, we have

$$\overline{F}_{(n)} (X_{(r)}) - \overline{F}_{(n)}(x_n^o) = [\overline{F}_{(n)}(X_{(r)}) - S_n(X_{(r)})] + (r/n - p_n), \tag{4.34}$$

where by lemma 4.2, the first term on the right hand side of(4.34) is $O_p(n^{\frac{1}{2}})$. Hence the lemma follows from (4.33) and (4.34). Q.E.D.

Let us now consider a sequence $\{a_n\}$ of real numbers, where $a_n \to 0$ as $n \to \infty$, and let

$$(4.35) \qquad I_n = \{x: \ x_n^o - a_n \leq x \leq x_n^o + a_n\} \ ;$$

$$(4.36) \qquad G_n(x) = n^{\frac{1}{2}}\{[S_n(x) - \overline{F}_{(n)}(x)] - [S_n(x_n^o) - \overline{F}_{(n)}(x_n^o)]\}, \ x\varepsilon I_n.$$

<u>Theorem 4.6</u>. <u>There exists two sequences</u> $\{\varepsilon_n\}$ <u>and</u> $\{\delta_n\}$ <u>of real and positive numbers, such that</u> (i) $\varepsilon_n \to 0$, $\delta_n \to 0$ as $n \to \infty$, <u>and</u> (ii)

$$P\{\sup[|G_n(x)| \ : x\varepsilon I_n] > \varepsilon_n\} < \delta_n \ .$$

<u>Proof</u>. It follows directly from (4.21) and (4.23).

From lemma 4.5 and theorem 4.6, we have the following.

<u>Lemma 4.7</u>. $n^{\frac{1}{2}}[S_n(X_{(r)}) - \overline{F}_{(n)}(X_{(r)})] \sim_p n^{\frac{1}{2}}[S_n(x_n^o) - \overline{F}_{(n)}(x_n^o)]$.

For later convenience, we define

$$(4.37) \qquad W_n^* = n^{\frac{1}{2}}[S_n(x_n^o) - \overline{F}_{(n)}(x_n^o)].$$

<u>Lemma 4.8</u>. <u>If</u> $\inf_n \nu_{n,m} > 0$, then $\mathcal{L}(W_n^*/\nu_{n,m}) \to \mathcal{N}(0,1)$.

For proof, see lemma 2.3 of Sen [5].

<u>Lemma 4.9</u>. <u>Under</u> (3.2), $\max_{1 \leq i \leq n} g(X_{(i)}) = o_p(n^{\frac{1}{2}})$.

<u>Proof</u>. Let $g_i = g(X_i)$ and the cdf of g_i be H_i, $i = 1, \ldots, n$. Then, (2.2) implies that

(4.38) $$\sup_{1\le i\le n} \int_0^\infty g^2 dH_i(g) < \infty .$$

Now $g^2[1 - H_i(g)] \le \int_g^\infty t^2 dH_i(t) \to 0$ as $g \to \infty$. Hence, for any $\{g_n^*\}$ for which $\lim_{n\to\infty} g_n^* = \infty$, we have

(4.39) $$\sup_i [1 - H_i(g_n^*)] \le \alpha(g_n^*)/[g_n^*]^2, \text{ where } \lim_{n\to\infty} \alpha(g_n^*) = 0 \text{ by (3.2).}$$

Now, let $g_n^{(j)} = \max[g_j, g_{j+m+1}, \ldots, g_{j+(n_j-1)(m+1)}]$, $j = 1, \ldots, m + 1$, so that $\max_i g(X_{(i)}) = \max_{1\le j\le m+1} [g_n^{(j)}, j = 1, \ldots, m + 1]$. Consequently,

(4.40) $$P\{\max_i g(X_{(i)}) \ge g_n^*\} \le \Sigma_{j=1}^{m+1} P\{g_n^{(j)} \ge g_n^*\}$$

$$= \Sigma_{j=1}^{m+1} [1 - P\{g_n^{(j)} < g_n^*\}] = \Sigma_{j=1}^{m+1} [1 - \prod_{k=0}^{n_j-1} H_{j+k(m+1)}(g_n^*)]$$

$$\le \Sigma_{j=1}^{m+1} [1 - \{1 - \alpha(g_n^*)/[g_n^*]^2\}^{n_j}] \quad \text{(by (3.31))} .$$

Now, we let $g_n^* = \varepsilon\sqrt{n}$, where $\varepsilon(>0)$ is arbitrarily small. Then the difference between the right hand side of (4.40) and $(m + 1)[1 - \exp\{-\alpha(g_n^*)/(m + 1)\varepsilon^2\}]$ converges to zero. Also

(4.41) $$(m+1)[1 - \exp\{-\alpha(g_n^*)/(m+1)\varepsilon^2\}] \le \frac{\alpha(g_n^*)}{\varepsilon^2} \to 0, \text{ as } n \to \infty .$$

Thus, $P\{\max_i g(X_{(i)}) \ge \varepsilon\sqrt{n}\} \to 0$ as $n \to \infty$, for every $\varepsilon > 0$. Q.E.D.

Lemma 4.10. Under (3.2)

(4.42) $$\max_{1\leq i\leq n} |\psi(X_i, \bar{F}_{(n)}(X_i)) - \psi(X_i, S_n(X_i))| = o_p(1) .$$

Proof. The L.H.S. of (4.42) is bounded above by

(4.43) $$[\max_{1\leq i\leq n} g(X_i)][\max_{1\leq i\leq n} |\bar{F}_{(n)}(X_i) - S_n(X_i)|] ,$$

and hence, the lemma directly follows from lemma 4.2 and lemma 4.9.

5. The proof of theorem 3.1. Let us define

(5.1) $$V^*_{ni} = \psi(X_{(i)}, S_n(X_{(i)})),\ V_{ni} = \psi(X_{(i)}, \bar{F}_{(n)}(X_{(i)})),\ i=1,\ldots,n.$$

Note that $Z^*_n = \max_i V^*_{ni}$ and $Y_{n,n} = \max_i V_{ni}$. Also let $a^{(1)}_n$, $a^{(2)}_n$ be so defined that

(5.2) $$h_n(x^o_n - a^{(1)}_n) = h_n(x^o_n + a^{(2)}_n) = h^o_n - cn^{-\frac{1}{2}}\log n,\ c > 0.$$

Then, both $a^{(1)}_n$ and $a^{(2)}_n \to 0$ as $n \to \infty$, and by lemma 4.1,

(5.3) $$P\{X_{(r)}\varepsilon[x^o_n - a^{(1)}_n,\ x^o_n + a^{(2)}_n]\} \to 1, \text{ as } n \to \infty .$$

Consider now the two subsets

(5.4) $$S^{(1)}_n = \{X_{(i)}: x^o_n - a^{(1)}_n \leq X_{(i)} \leq x^o_n + a^{(2)}_n\} \text{ and}$$

$$S^{(2)}_n = \{X_{(i)}: X_{(i)}\notin S^{(1)}_n\} .$$

Now, using (2.3), (2.4) and theorem 4.6, we have

(5.5) $$\sup_i [|n^{\frac{1}{2}}(V^*_{ni} - V_{ni}) - \psi_{01}(x^o_n, p_n)W^*_n|: i\varepsilon S^{(1)}_n] = o_p(1).$$

Hence, from (5.3) and (5.5), we have

(5.6) $$Z^*_{n,1} = \max_{i\varepsilon S_n^{(1)}} V^*_{ni} = Y_{n,n} + n^{-\frac{1}{2}}\psi_{01}(x_n^o, p_n)W_n^* + o_p(n^{-\frac{1}{2}}) .$$

Now, by lemma 4.1, $n^{\frac{1}{2}}[h_n^o - Y_{n,n}] = o_p(n^{-\frac{1}{2}})$, and hence, from (5.6), we have

(5.7) $$Z^*_{n,1} = h_n^o + \psi_{01}(x_n^o, p_n)\, n^{-\frac{1}{2}} W_n^* + o_p(n^{-\frac{1}{2}}) .$$

Consequently, by lemma 4.8,

(5.8) $$\mathcal{L}(n^{\frac{1}{2}} [Z^*_{n,1} - h_n^o]/\gamma_{n,m}) \to \mathcal{N}(0,1),$$

and hence, for any $c > 0$,

(5.9) $$P\{Z^*_{n,1} \geq h_n^o - (n^{-\frac{1}{2}} \log n)\, c/2\} \to 1, \text{ as } n \to \infty .$$

So the proof of the theorem will be completed if we can show that

(5.10) $$P\{\sup_{i\varepsilon S_n^{(2)}} V^*_{ni} \leq h_n^o - (n^{-\frac{1}{2}} \log n)\, c/2\} \to 1, \text{ as } n \to \infty ,$$

(as then $n^{\frac{1}{2}}(Z_n^* - Z^*_{n,1}) \overset{p}{\to} 0$). With this end in view, we divide $S_n^{(2)}$ into two subsets

(5.11) $$S^{(2)}_{n,1} = \{X_{(i)}\varepsilon S_n^{(2)}: g(X_{(i)}) < g_o < \infty\},\ S^{(2)}_{n,2} = S_n^{(2)} - S^{(2)}_{n,1} ,$$

where $g(X)$ is defined by (3.2) and $g_o > \sup_n |\psi_{01}(x_n^o, p_n)|$. Note that

(5.12) $$n^{\frac{1}{2}}|V_{ni} - V^*_{ni}| \leq g(X_{(i)})|n^{\frac{1}{2}}\{S_n(X_{(i)}) - \overline{F}_{(n)}(X_{(i)})\}| .$$

Hence, for all $i \varepsilon S_{n,1}^{(2)}$, the right hand side of (5.12) is $0_p(1)$, by lemma 4.2. On the other hand, $V_{ni} \leq h_n^o - cn^{-\frac{1}{2}} \log n$ for all $i \varepsilon S_{n,1}^{(2)}$. Consequently, with probability approaching to unity,

$$(5.13) \qquad V_{ni}^* \leq h_n^o - cn^{-\frac{1}{2}} \log n + 0_p(n^{-\frac{1}{2}}) \leq h_n^o - \frac{c}{2} n^{-\frac{1}{2}} \log n, \text{ for all } i \varepsilon S_{n,1}^{(2)} .$$

Thus,

$$(5.14) \qquad P\{\sup_{i\varepsilon S_{n,1}^{(2)}} V_{ni}^* \leq h_n^o - (c/2)\ n^{-\frac{1}{2}} \log n\} \rightarrow 1, \text{ as } n \rightarrow \infty .$$

Finally, for all $i \varepsilon S_{n,2}^{(2)}$, $g(X) > g_o > \sup_n |\psi_{01}(x_n^o, p_n)|$ where ψ_{01} and $g(X)$ are continuous functions in the neighbourhood of (x_n^o, p_n) and g_o can be an arbitrarily large but finite real number. Therefore, by lemma 4.10 the supremum of V_{ni}^* over $S_{n,2}^{(2)}$ is less than $Z_{n,1}^*$ in probability. Hence, $Z_{n,1}^* = Z_n^*$ in probability. Hence the theorem.

6. Moment convergence of Z_n^*. We impose the following regularity conditions: (a) $\psi(x, y)$ is concave in $y (0 \leq y \leq 1)$ i.e.,

$$(6.1) \qquad \psi(x, y) \leq (1-y)\psi(x, 0) + y\psi(x, 1), \ \forall\, y \varepsilon (0, 1),$$

where (b) $\psi(x, \delta)$, $\delta = 0, 1$ are non-negative,

$$(6.2) \qquad \psi(x, 0) \text{ is } \uparrow \text{ in } x \text{ and } \psi(x, 1) \text{ is } \downarrow \text{ in } x:\ -\infty < x < \infty ,$$

and (c) square integralability of $\psi(x, \delta)$, $\delta = 0, 1$, i.e.,

$$(6.3) \qquad \sup_n \int_0^\infty \psi^2(x, \delta)\, d\bar{F}_{(n)}(x) \leq \mu_2(\delta) < \infty \ \text{ for } \delta = 0, 1.$$

<u>Theorem 6.1</u>. <u>Under</u> (6.1)-(6.3), $|E(Z_n^*) - h_n^o| \to 0$ <u>as</u> $n \to \infty$.

<u>Proof</u>. Let $X_{(1)} \leq \cdots \leq X_{(n)}$ be the order statistics. Then, by (6.1),

(6.4) $$Z_{n,i} = \psi(X_{(i)}, S_n(X_{(i)})) \leq \frac{n-i}{n} \psi(X_{(i)}, 0) + \frac{i}{n} \psi(X_{(i)}, 1), \; 1 \leq i \leq n.$$

Now, by (6.2)

(6.5) $$\frac{n-i}{n} \psi(X_{(i)}, 0) \leq \frac{1}{n} \Sigma_{j=i}^{n} \psi(X_{(j)}, 0) \leq \frac{1}{n} \Sigma_{j=1}^{n} \psi(X_{(j)}, 0)$$
$$= \frac{1}{n} \Sigma_{j=1}^{n} \psi(X_j, 0) \; ;$$

(6.6) $$\frac{i}{n} \psi(X_{(i)}, 1) \leq \frac{1}{n} \Sigma_{j=1}^{i} \psi(X_{(j)}, 1) \leq \frac{1}{n} \Sigma_{j=1}^{n} \psi(X_{(j)}, 1)$$
$$= \frac{1}{n} \Sigma_{j=1}^{n} \psi(X_j, 1) \; .$$

Therefore, on writing $\overline{\psi}_n = n^{-1} \Sigma_{i=1}^{n} \{\psi(X_i, 0) + \psi(X_i, 1)\}$, we have

(6.7) $$0 \leq Z_n^* \leq \overline{\psi}_n \; .$$

Now, Z_n^* converges in probability to h_n^o (by theorem 3.1), while by (6.3) $\overline{\psi}_n$ is uniformly (in n) integrable. Hence, the theorem follows by using the dominated convergence theorem (cf. Loeve [4, p. 125]). Q.E.D.

The next problem is to study the conditions under which in theorem 3.1 we can replace h_n^o by $E(Z_n^*)$. We have failed to provide a general theorem on this. However, in many cases this can be done. In the remaining of the paper, we consider the special case of iidrv and a simple ψ, where all these results apply directly.

7. A special case. In the introduction, we have discussed the statistic B_n^* which describes the average strength of a bundle of n filaments. In the present section we shall show that under certain restriction, $B_n^* = \sup_{0<x<\infty} x(1 - \overline{F}_{(n)}(x))$ in probability as $n \to \infty$ and we shall also study the moment convergence of the statistic B_n^*. We know that

$$B_n^* = \max\{Y_1, (1 - 1/n)Y_2, (1 - 2/n)Y_3, \ldots Y_n/n\} \tag{7.1}$$

where $Y_1, Y_2, \ldots Y_n$ are the ordered random variables obtained from n non-negative random variables $X_1, X_2, \ldots X_n$ which are the individual strength of the filaments in the bundle.

Lemma 7.1. *Let $\{X_i\}$ be a sequence of independent non-negative random variables with distribution functions $\{F_i\}$ and let*

$$\overline{\lim_i} \int_0^\infty x^2 \, dF_i(x) < \infty \tag{7.2}$$

and x^2 be uniformly integrable then $\lim_{n\to\infty} P\{\max_{1\leq i\leq n} \{X_i\} > \varepsilon n^{\frac{1}{2}}\} = 0$ *for every* $\varepsilon > 0$.

Proof. It follows directly from lemma 4.9.

Lemma 7.2. *Under the conditions of lemma 7.1.*

$$\lim_{n\to\infty} n^{\frac{1}{2}} |B_n^* - \sup_x x(1 - S_n(x))| \overset{p}{=} 0$$

Proof. Let $I(u) = 1$ or 0 according as $u > 0$ or not. Then

$$S_n(x) = \frac{1}{n} \sum_{i=1}^n I(x - X_i), \tag{7.3}$$

(7.4) $$B_n^* = \max_{1\le i\le n} \{Y_i(1 - S_n(Y_i))\} = \sup_x \{x(1 - S_n(x)\}$$

since the sup occurs at one of the jump points. Now at all points which are not jump points of the function $S_n(x)$, $x(1 - S_n^-(x)) = x(1 - S_n(x))$ and at jump points

(7.5) $$X_i(1 - S_n^-(X_i)) - X_i(1 - S_n(X_i)) = \frac{X_i}{n}$$

Hence

(7.6) $$n^{\frac{1}{2}}|B_n^* - \sup_x x(1 - S_n(x))| = \max_i X_i/ n^{\frac{1}{2}} \overset{p}{\to} 0|| .$$

This leads us to study the special case when

$$\psi(x, \overline{F}_{(n)}(x)) = x(1 - \overline{F}_{(n)}(x)) .$$

<u>An ineqality</u>: Consider a set of n + m non-negative real numbers $x_1, x_2, \ldots x_{n+m}$ with $y_1 \le y_2 \le \ldots \le y_{n+m}$ as the corresponding ordered set. Let $(x_1^{(i)}, x_2^{(i)}, \ldots x_n^{(i)})$; $i = 1, 2, \ldots \binom{n+m}{n}$ be the all possible combinations of n tuples that can be formed from the n+m x's and let $y_1^{(i)} \le y_2^{(i)} \le \ldots \le y_{(n)}^{(i)}$ denote the corresponding ordered set.

(7.7) $$b_{n+m}(x_1, x_2, \ldots x_{n+m}) = \max\{(n+m)y_1, (n+m-1)y_2, \ldots y_{n+m}\}$$

and

(7.8) $$b_n(x_1^{(i)}, x_2^{(i)}, \ldots x_n^{(i)}) = \max\{ny_1^{(i)}, (n-1)y_2^{(i)}, \ldots y_n\}$$

<u>Lemma 7.3</u>.

$$nb_{n+m}(x_1, x_2, \ldots x_{n+m}) \le \frac{n+m}{\binom{n+m}{n}} \sum_{i=1}^{\binom{n+m}{n}} b_n(x_1^{(i)}, x_2^{(i)}, \ldots x_n^{(i)})$$

Proof. There are $\binom{n+m-j+1}{n-j+k}\binom{j-1}{k-1}$ combinations where

$$b_n(x_1^{(i)}, \ldots x_n^{(i)}) \geq (n-j+k)\, y_j \qquad 1 \leq k \leq j \tag{7.9}$$

Hence

$$\{(n+m) \Big/ \binom{n+m}{n}\} \sum_{i=1}^{\binom{n+m}{n}} b_n(x_1^{(i)}, \ldots x_n^{(i)}) \geq \tag{7.10}$$

$$\frac{(n+m)y_j}{\binom{n+m}{n}} \sum_{k=1}^{j} (n-j+k) \binom{n+m-j+1}{n-j+k} \binom{j-1}{k-1}$$

$$= n(n+m-j+1)\, y_j; \quad (j = 1, 2, \ldots n+m) .$$

Therefore

$$\frac{(n+m)}{\binom{n+m}{m}} \sum_{i=1}^{\binom{n+m}{n}} b_n(x_1^{(i)}, \ldots x_n^{(i)}) \geq n \max\{(n+m-j+1)\, y_j\} \tag{7.11}$$

$$= nb_{n+m}\, (x_1, \ldots x_n) \qquad ||.$$

Theorem 7.4. <u>If the sequence of random variables</u> $\{X_i\}$ <u>is a sequence of iid random variables then the r.v.'s</u> $\{B_n^*\}$ <u>with probability one form a reverse-semi-martingale if</u> $E(X_i) < \infty$.

Proof. Multiplying both sides of (7.11) by the conditional distribution of $X_1, X_2, \ldots X_{n+m}$ given $B_{n+m}, B_{n+m+1}, \ldots B_{n+m+k}$ and integrating and observing the interchangeability of the random variables we get

$$\frac{B_{n+m}}{n+m} = E\left(\frac{B_n}{n} \,\middle|\, \frac{B_{n+m}}{n+m}, \ldots \frac{B_{n+m+k}}{n+m+k} \right) \tag{7.12}$$

Moreover

$$(7.13) \qquad 0 \le \frac{B_n}{n} \le \bar{X}_n \Rightarrow E\left(\frac{B_n}{n}\right) < \infty .$$

Therefore

$$\lim_{k\to\infty} \frac{B_{n+m}}{n+m} \le \lim_{k\to\infty} E\left(\frac{B_n}{n} \,\middle|\, \frac{B_{n+m}}{n+m}, \ldots \frac{B_{n+m+k}}{n+m+k}\right)$$

$$= E\left(\frac{B_n}{n} \,\middle|\, \frac{B_{n+m}}{n+m}, \ldots\ldots\ldots\right)$$

with prob. one. That the limiting process is valid follows from Doob [2, p. 332].

<u>Corollary</u>: <u>Under the conditions of theorem</u> 7.4, $E\left(\frac{B_n}{n}\right)$ <u>is a monotone decreasing function of n</u>.

Proof. Taking m = 1, multiplying both sides of (7.11) by the joint distribution of the interchangeable random variables $(X_1, X_2, \ldots X_{n+1})$ and integrating we obtain the result immediately.

<u>Theorem 7.5</u>. <u>Let</u> $\{X_i\}$ <u>be a sequence of iidrv with</u> $E(X_i^2) < \infty$. <u>Then</u> $B_n^* \xrightarrow{a.s.} y_o$, <u>where</u> y_o <u>by assumption is the unique maximum of the function</u> $x(1 - F(x))$.

<u>Proof</u>. By theorem 3.1 and lemma 7.2 we know that $B_n^* \xrightarrow{p} y_o$. Moreover $\{B_n^*\}$ is a reverse semi-matingale sequence and $B_n^* \le \bar{X}_n$ which is uniformly integrable. Therefore, (cf. [4, p. 393]), there exists a random variable Z_o such that $B_n^* \xrightarrow{a.s.} Z_o$. Hence by the equivalence theorem the result follows ||.

Let $Z_{ni} = X_i[1 - S_n(X_i)]$, $i = 1, \ldots, n$, where X_i's are iidrv with the common cdf $F(x)$, such that $M(t) = E(e^{tx}) < \infty$ for all $|t| < T$.

Theorem 7.6. If $M(t) < \infty$, $\forall\ |t| < T$, then for any α: $0 < \alpha < 1$,

(7.14) $$n^{\alpha}|[E(Z_n^*) - y_o]| \to 0 \text{ as } n \to \infty, \text{ where } y_o = \max_x x[1 - F(x)].$$

Proof. Let $c(u) = 1$ if $u \geq 0$ and 0, otherwise. Then,

(7.15) $$Z_{ni} = n^{-1}[X_i + \Sigma_{j \neq i}^{n} X_i\ c(X_j - X_i)],\ i = 1, \ldots, n.$$

As $E(Z_n^*)$ is $\downarrow$ in n and by theorem 6.1, $E(Z_n^*) \to y_o$ as $n \to \infty$, we must have $E(Z_n^*) \geq y_o$ for all n. Hence, it suffices to show that for any α: $0 < \alpha < 1$,

(7.16) $$n^{\alpha}(E(Z_n^*) - y_o) \to 0 \quad \text{as } n \to \infty .$$

Now, for any positive integer k

(7.17) $$E(Z_{ni}^k) = E(Z_{n1}^k) = n^{-k}\ E\{X_1^k[1 + \Sigma_{j=2}^{n}\ c(X_j - X_1)]^k\}$$

$$= n^{-k}\ \Sigma_{s=0}^{k}\ \binom{k}{s}\ E\{X_1^k[\Sigma_{j=2}^{n}\ c(X_j - X_1)]^{k-s}\}$$

$$= n^{k}\ \Sigma_{s=0}^{k}\ \binom{k}{s}\ \Sigma_{\ell=1}^{k-s}\ \binom{n-1}{\ell}\ \Sigma^*\ E\{X_1^k\ [c(X_2-X_1)]^{r_1} \ldots [c(X_{\ell+1}-X_1)]^{r_\ell}\},$$

where $r_j \geq 1$, $j = 1, \ldots, \ell$, $\Sigma_{j=1}^{\ell} r_j = k - s$, and the summation Σ^* extends over all possible choice of $r_1, \ldots, r_\ell$. Now, conditioned on X_1, the random variables $c(X_j - X_1)$, $j = 2, \ldots, n$ are iid with

$$E\{[c(X_j - X_1)]^r | X_1\} = 1 - F(X_1),\ j = 2, \ldots, n - 1, \text{ and } r \geq 1$$

and hence,

(7.18) $$E(Z_{ni}^{k}) = n^{-k} \sum_{s=0}^{k} \binom{k}{s} \Sigma_{\ell=1}^{k-s} \binom{n-1}{\ell} \Sigma^{*} \int_{0}^{\infty} x^{k}[1 - F(x)]^{\ell}\, dF(x)$$

$$= \Sigma_{s=0}^{k} \binom{k}{s} n^{-s} \Sigma_{\ell=1}^{k-s} n^{-(k-s)} \binom{n-1}{\ell} \Sigma^{*} \int_{0}^{\infty} x^{k-\ell} y^{\ell}\, dF(x)$$

(where $y = x[1 - F(x)] \leq y_o$)

$$\leq \Sigma_{s=0}^{k} \binom{k}{s} n^{-s} \Sigma_{\ell=1}^{k-s} n^{-(k-s)} \binom{n-1}{\ell} \Sigma^{*} y_o^{\ell} \mu'_{k-\ell}$$

$$\leq \Sigma_{s=0}^{k} \binom{k}{s} n^{-s} \Sigma_{\ell=1}^{k-s} y_o^{\ell} \mu'_{k-\ell} n^{-(k-s-\ell)} \left(1 - \frac{\ell+1}{2n}\right)^{\ell} \left(\frac{1}{\ell!} \Sigma^{*} 1\right),$$

as $(n-1)^{[\ell]} = (n-1) \ldots (n-\ell) \leq \left(n - \frac{\ell+1}{2}\right)^{\ell}$. Let us now take

(7.19) $$k = k_n = [\sqrt{2n \log n}\,] \Rightarrow \left(1 - \frac{k_n+1}{2n}\right)^{k_n} \sim e^{-\log n} = n^{-1}.$$

Then, from (7.18) and (7.19), we have

(7.20) $$E(Z_{ni}^{k_n}) \leq n^{-1} y_o^{k_n} \left[1 + \frac{k_n}{2n\, y_o} \{1 + Q_{n1}\}\right] + n^{-1} y_o^{k_n} \left[\frac{k_n}{n y_o} \mu'_1 (1 + Q_{n2})\right],$$

where it can be shown that

(7.21) $$\lim_{n\to\infty} Q_{ni} = 0,\ i = 1, 2, \text{ when } M(t) < \infty.$$

Thus, on using the fact that $(Z_n^*)^k \leq \Sigma_{i=1}^{n} Z_{ni}^{k}$, $k \geq 0$, we have

(7.22) $$E(Z_n^*) \leq [E(Z_n^{*k_n})]^{1/k_n} \leq [E(\Sigma_{i=1}^{n} Z_{ni}^{k_n})]^{1/k_n}$$

$$= [n\, E(Z_{ni}^{k_n})]^{1/k_n}$$

$$\leq y_o \{1 + \frac{3k_n \mu_1'}{2ny_o} [1 + \frac{1}{3} Q_{n1} + \frac{2}{3} Q_{n2}]\}^{1/k_n}$$

$$\leq y_o \{1 + \frac{3\mu_1'}{2ny_o} [1 + \frac{1}{3} Q_{n1} + \frac{2}{3} Q_{n2}] + o(\frac{1}{n})\}.$$

Consequently, for any α: $0 < \alpha < 1$,

$$n^{\alpha}[E(Z_n^*) - y_o] \leq \frac{3 \mu_1'}{2n^{1-\alpha} y_o} [1 + \frac{1}{3} Q_{n1} + \frac{2}{3} Q_{n2}] + o(n^{-1+\alpha}) \to 0 \text{ an } \to \infty . \quad (7.23)$$

Hence the theorem.

8. Possible generalizations of theorem 3.1. (a) Multivariate case. Let $\underset{\sim}{X}_i = (X_{ij}, \ldots, X_{ip})'$ $(p \geq 1)$, $i = 1, \ldots, n$ be independent r.v. with continuous cdf's $F_1, \ldots, F_n$ respectively. For the jth variate, we denote the marginal cdf's by $F_{i[j]}$, $i = 1, \ldots, n$; $j = 1, \ldots, p$; and let

$$\bar{F}_{n(j)}(x) = n^{-1} \Sigma_{i=1}^{n} F_{i[j]}(x), \; S_{n(j)}(x) = n^{-1} \Sigma_{i=1}^{n} c(x - X_{ij}). \quad (8.1)$$

$$j = 1, \ldots, p.$$

Consider now a function (positive valued)

$$h_n(x) = \psi(\underset{\sim}{x}, \bar{F}_{n(1)}(x_1), \ldots, \bar{F}_{n(p)}(x_p)), \; \underset{\sim}{x} \in R^p, \quad (8.2)$$

and assume that $h_n(\underset{\sim}{x})$ assumes a unique maximum at $\underset{\sim}{x}_n^o$, where $h_n(\underset{\sim}{x}_n^o) = h_n^o < \infty$. Let then

$$Z_{ni} = \psi(\underset{\sim}{X}_i, S_{n(1)}(X_{i1}), \ldots S_{n(p)}, (X_{ip})), \; i = 1, \ldots, n \; ; \quad (8.3)$$

$$Z_n^* = \max_i Z_{ni} . \quad (8.4)$$

We denote the partial derivatives of $\psi(\underset{\sim}{x}, y_1, \ldots, y_p)$ with respect to $y_1, \ldots, y_p$ at $\underset{\sim}{x} = \underset{\sim}{x}_n^o$, $y_j = \overline{F}_{n(j)}(x_{nj}^o)$, $j = 1, \ldots, p$ by $\zeta_{ni}, \ldots, \zeta_{np}$, respectively. Finally, let

(8.5) $$W_{n(j)}^* = n^{\frac{1}{2}} [S_{n(j)}(x_{nj}^o) - \overline{F}_{n(j)}(x_{nj}^o)], \; j = 1, \ldots, p .$$

Then, by the same technique as in theorem 3.1, it can be shown that

(8.6) $$Z_n^* \sim p \, h_n^o + n^{-\frac{1}{2}} \Sigma_{j=1}^p \zeta_{nj} W_{n(j)}^* + o_p(n^{-\frac{1}{2}}) .$$

Hence,

(8.7) $$\mathcal{L}(n^{\frac{1}{2}} (Z_n^* - h_n^o)) \to \mathcal{N}(0, \gamma_n^2) ,$$

where $\gamma_n^2 = \Sigma_{j=1}^p \Sigma_{\ell=1}^p \zeta_{nj} \zeta_{n\ell} \operatorname{Cov}[W_{n(j)}^*, W_{n(\ell)}^*]$.

(b) Vector case. Suppose now that $\underset{\sim}{h}_n(x) = [h_n^{(1)}(x), \ldots, h_n^{(b)}(x)]$, where $h_n^{(j)}(x)$ assumes a unique maximum at x_{nj}^o, $j = 1, \ldots, p$. The asymptotic multinormality of $n^{\frac{1}{2}}[\underset{\sim}{h}_n(x) - \underset{\sim}{h}_n^o]$ follows along the same line as in theorem 3.1.

(c) Extension to linear ordered estimators. We let

(8.8) $$Z_{n,r}^* = r\text{th maximum of } Z_{n1}, \ldots, Z_{nn}, \quad r = 1, \ldots, n.$$

Let us consider a set of weights $a_{n1}, \ldots, a_{nn}$ satisfying the following conditions: (a)

(8.9) $$a_{nr} \geq 0, \; \forall\, r \quad \text{and} \quad \Sigma_{r=1}^n a_{nr} = 1 ,$$

and (b) $\Sigma a_{nr} Z^*_{nr}$ is essentially an extreme statistic, so that a_{nr} rapidly converges to zero as r increases. Specifically, we assume that

$$(8.10) \qquad \Sigma_{r>k_n} a_{nr} = o(n^{-\frac{1}{2}}) ,$$

where k_n is so chosen that

$$(8.11) \qquad h_n(x^o_n \pm k_n/n) \geq h^o_n - cn^{-\frac{1}{2}} \log n .$$

Thus, by (3.6), $k_n = o(n^{\frac{1}{2}})$. Then, proceeding as in the proof of theorem 3.1, we have

(a) $Z^*_{n1}, \ldots, Z^*_{nk_n}$ all belong to $S^{(1)}_n$, with probability $\rightarrow 1$, as $n \rightarrow \infty$,

$$(8.12) \qquad \text{(b)} \quad |n^{\frac{1}{2}} \sum_{r=1}^{k_n} a_{nr}[Z^*_{nr} - Y_{nr}] - \sum_{r=1}^{k_n} a_{nr} \psi_{o1}(x^o_n, p_n) W^*_n| = o_p(1) ,$$

and (c) by (8.10), $\Sigma^{k_n}_{r=1} a_{nr} = 1-o(n^{-\frac{1}{2}})$. Further

$$n^{\frac{1}{2}} \Sigma^n_{r=k_n+1} a_{nr} Z^*_{nr} \leq Z^*_{nk_n+1} (n^{\frac{1}{2}} \Sigma_{r>k_n} a_{nr}) = 0_p(1).\ o(1) = o_p(1).$$

Consequently,

$$(8.13) \qquad |n^{\frac{1}{2}} [\Sigma^n_{r=1} a_{nr} Z^*_{nr} - h^o_n] - n^{\frac{1}{2}} [\Sigma^{k_n}_{r=1} a_{nr}(Z^*_{nr} - h^o_n)]| = o_p(1) .$$

Along the same line as in lemma 4.1, it can be shown that

$$(8.14) \qquad n^{\frac{1}{2}}|Y_{n,r} - h^o_n| = o_p(1) \quad \text{for all } r < k_n = o(n^{\frac{1}{2}}) .$$

Hence, from (8.12), (8.13), and (8.14), we have

$$(8.15) \qquad n^{\frac{1}{2}} [\Sigma_{r=1}^{n} a_{nr} Z_{nr}^{*} - h_n^o] \underset{p}{\sim} \psi_{o1}(x_n^o, p_n) W_n^* .$$

Consequently,

$$(8.16) \qquad \mathcal{L}(n^{\frac{1}{2}} [\Sigma_{r=1}^{n} a_{nr} Z_{nr}^{*} - h_n^o]/\gamma_{n,m}) \to \mathcal{N}(0,1) .$$

(e) A Jackknife estimator: As by theorem 7.5, $|E(Z_n^*) - y_o| \leq a/n$, where $a < \infty$, it may be better to consider the alternative estimator

$$(8.17) \qquad Z_n^{**} = \frac{1}{n} \sum_{i=1}^{n} \{n Z_n^* - (n-1) Z_{n-1,i}^*\} ,$$

where $Z_{n-1,i}^*$ is the maximum of $X_i[1 - F_{n-1}(X_i)]$, in a sample of size $n - 1$ where the ith observation of the sample of size n is omitted. Obviously, the bias of Z_n^{**} is less than that of Z_n^*. It is easily seen (on using (5.7)) that $|Z_n^{**} - Z_n^*| = o_p(n^{-\frac{1}{2}})$, and hence, the normality of $n^{\frac{1}{2}} (Z_n^{**} - h_n^o)$ would follow from theorem 3.1.

9. Concluding remarks. In order to obtain the asymptotic normality of the statistic $\sup_x \psi(x, S_n(x))$, we have shown that $\sup_x n^{\frac{1}{2}} |S_n(x) - \overline{F}_{(n)}(x)|$ is bounded in probability. We have relaxed the identity of $F_1, \ldots, F_n$ but not thier continuity. In a subsequent paper various limiting properties of the above statistics for possibly discontinuous cdf's will be presented.

REFERENCES

[1] DANIELS, H. E. (1945) The statistical theory of the strength of bundles of threads. Proc. Roy. Soc. London Ser. A. 183, 405-435.

[2] DOOB, J. L. (1953) Stochastic process, New York: Wiley.

[3] HÁJEK, J. & ŠIDÁK, Z. (1967) Theory of rank tests, New York: Academic Press.

[4] LOÈVE, M. (1960) Probability theory, New York: D. van Nostrand Co. Inc.

[5] SEN, P. K. (1968) Asymptotic normality of sample quantiles for m dependent processes. Ann. Math. Statist. 39, 1724-1730.