On Order Statistics and Some Applications of Combinatorial Methods in Statistics*

by

Shanti S. Gupta and S. Panchapakesan

Purdue University

and

Southern Illinois University

Department of Statistics
Division of Mathematical Science
Mimeograph Series #261

August 1971

*Research supported in part by the Office of Naval Research Contract N00014-67-A-0226-00014 at Purdue University.

ERRATA

Page	Line	For	Read
2	3↑	Anderson	Andersen
19	Th. 4.9 (ii)	$P\{X_{(N_n)} \leq a_{N_n} x + b_{N_n}\} \to \Lambda(x)$ for all $x \in C_\Lambda$.	
23	9↑	$k_n \leq n$.	$k_n = k \leq n$

1. Introduction

Research in the area of order statistics has been steadily and rapidly growing especially during the last two decades. The extensive role of order statistics in several areas of statistical inference has made it imperative and useful to gather these results and present them in varied manner to suit diverse interests. The present paper is an instance of such an attempt.

Historically, formal investigation in the sampling theory of order statistics dates back to 1902 when Karl Pearson solved the problem of finding the mean of the difference between the rth and the (r+1)th order statistics in a sample of n observations from a continuous population. Tippett (1925) found the mean of the sample range and tabulated for certain sample sizes ranging from 3 to 1000, the cumulative distribution function (cdf) of the largest order statistic in a sample from a standard normal population. Asymptotic results were first obtained by Fisher and Tippett (1928), who determined under certain regularity conditions the limiting distributions of the largest and the smallest order statistics as the sample size increases indefinitely by a method of functional equations. These early developments and subsequent research over a period of nearly a quarter of a century have been nicely summarized by Wilks (1948) in a survey paper. Since then, a huge volume of research has been accomplished in this field dealing with several aspects of the problems involving order statistics. Besides the basic distribution theory and limit laws, attention has been focussed by several authors on problems involving order statistics in the theory of estimation and testing of hypotheses and in multiple decision and multiple comparison procedures. Many of these results are embodied in books and monographs; mention should be made of Gumbel (1958), Sarhan and Greenberg (1962), Miller (1966) and David (1970).

The modest objective of the present authors is to state some of the basic results in the thoery of order statistics, describe the trend of the work done in certain areas by referring to what might be called 'land mark' papers and indicate some of the recent results. In doing so, certain areas where order statistics play an important role have not been considered with no reflection on the nature of their importance in applications; for example, multiple comparison problems and slippage tests. A few topics have been treated to a very limited extent. The basic theory (Section 2), results concerning moments and inequalities (Section 3) and problems concerning estimation and hypotheses testing (Section 7) come under this category. Section 4 discusses some important asymptotic results relating to the papers of Gnedenko (1943), Smirnov (1949), Rényi (1953), Berman (1962), Pyke (1965) and Kiefer (1970a). Applications of combinatorial methods in the general distribution theory and fluctuation theory have been described in Sections 5 and 6. These results are mainly concerned with the applications of the ballot lemma and its generalizations and the use of the equivalence principle proved by Anderson (1953). The last section discusses the role of order statistics in the subset selection problems and the algebraic structure involved in identification problems.

2. Basic Distribution Theory

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables each having an absolutely continuous distribution function $F(x)$ and the corresponding density function $f(x)$. Let the ordered variables be denoted by

(2.1) $$X_{(1)} < X_{(2)} < \ldots < X_{(n)} \quad .$$

If the situation demands more clarity, $X_{(r)}$ will be denoted by $X_{r,n}$. It is well-known that $f_r(x)$ and $F_r(x)$, namely, the density and the cdf of $X_{(r)}$ are given by

(2.2) $$f_r(x) = r\binom{n}{r} F^{r-1}(x)\,[1-F(x)]^{n-r} f(x)$$

and

(2.3) $$F_r(x) = \sum_{i=r}^{n} \binom{n}{i} F^i(x)\,[1-F(x)]^{n-i}$$

$$= I_{F(x)}(r,\ n-r+1)$$

where $I_p(a,b)$ is the incomplete beta function defined by

(2.4) $$I_p(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} \int_0^p t^{a-1}(1-t)^{b-1}dt,$$

$$a,b > 0,\ 0 \leq p \leq 1 \ .$$

The joint density function $f_{r,s}(x,y)$ of $X_{(r)}$ and $X_{(s)}$ $(1 \leq r < s \leq n)$ is given by independent and exponentially distributed with cdf = $1-e^{-x} (x > 0)$. An important and very useful fact when dealing with order statistics is that they form a Markov process. To be precise, $\{X_{(r)}: 1 \leq r \leq n\}$ is a non-homogeneous discrete-parameter, real-valued Markov process whose initial measure is $F_1(x) = 1 - [1-F(x)]^n$ and whose transition distribution function $P\{X_{(r+1)} \leq x | X_{(r)} = y\}$ is the distribution of the minimum of (n-r) independent observations on the distribution F truncated at y, that is,

$$(2.11)\quad P\{X_{(r+1)} \leq x | X_{(r)} = y\} = 1 - [1-F(x)]^{n-r} [1-F(y)]^{-n+r}, \quad x > y \quad .$$

This Markov property was first pointed out by Kolmogorov (1933). Further it is clear from (2.10) that $Y_{(r)}/Y_{(r+1)} = e^{-W_{n+1-r}/r}$, $r = 1,\ldots, n$ are all independent $[Y_{(n+1)} \equiv 1]$. Hence $[Y_{(r)}/Y_{(r+1)}]^r$, $r = 1,\ldots, n$, are independent and uniformly distributed on (0,1).

Now, the joint density of $X_{(1)},\ldots, X_{(n)}$ is given by

$$(2.12)\quad f(x_1,\ldots, x_n) = \begin{cases} n!\ f(x_1)\ldots f(x_n)\ , & x_1 < x_2 < \ldots < x_n, \\ 0 & \text{otherwise.} \end{cases}$$

Define

$$(2.13)\qquad D_r = X_{(r)} - X_{(r-1)},\ r = 2,\ldots, n \quad .$$

Then $D_2,\ldots,D_n$ are called the spacing. For some distributions we may define $X_{(0)}$ and $X_{(n+1)}$ suitably depending on F and let $D_1=X_{(1)}-X_{(0)}$ and $D_{n+1}=X_{(n+1)}-X_{(n)}$. For example, if F is the uniform distribution on (0,1), then $X_{(0)}=0$ and $X_{(n+1)}=1$.

If $F(x) = \lambda e^{-\lambda x}$, $x > 0$, we will just define $X_{(0)} = 0$. Thus, depending on the particular F, the number of spacings considered could be different. For the general spacing we see that the joint density of $D_2, \ldots, D_n$ is given by

$$(2.14) \qquad f_{\underline{D}}(d_2, \ldots, d_n) = \begin{cases} n! \int_{-\infty}^{\infty} \prod_{r=2}^{n} f(x+d_2+\ldots+d_r)dx, & d_2, \ldots, d_n > 0, \\ 0 & \text{otherwise.} \end{cases}$$

and the density of D_r is

$$(2.15) \quad f_{D_r}(y) = \frac{n!}{(r-2)!(n-r)!} \int_{-\infty}^{\infty} F^{r-2}(x)\, [1-F(x+y)]^{n-r} f(x)f(x+y)dx.$$

In the case of the exponential spacings $D_1, \ldots, D_n$ are independent exponential random variables with parameters $\lambda n, \lambda(n-1), \ldots, \lambda$. Equivalently, the normalized spacings $\lambda(n-r+1)D_r$, $1 \leq r \leq n$, are independent and identically distributed exponentially with mean unity. The exponential spacings can be looked upon as holding times of a continuous parameter Markov process. The first unified approach to the distribution theory of uniform spacings is given by Darling (1953). A good discussion of spacings can be found in Pyke (1965). The use of spacings in tests of hypotheses is discussed in a subsequent section.

3. Moments of Order Statistics and Bounds

Some important results are concerned with moments of order statistics from specific distributions, particularly, the normal distribution, and bounds for the moments under certain assumptions on the parent distribution. When f(x) is symmetric about the origin, we have

$$E(X_{r,n}) = -E(X_{n-r+1,n}) \tag{3.1}$$

and

$$\mathrm{Cov}(X_{r,n}\ X_{s,n}) = \mathrm{Cov}(X_{n-s+1,n},\ X_{n-r+1,n}) \ . \tag{3.2}$$

In the case of the standard normal distribution, the means, variances and covariances have been calculated for different ranges of values of n by Sarhan and Greenberg (1956), Teichroew (1956) and Harter (1961a). Bose and Gupta (1959) have discussed the evaluation of the exact moments of order statistics in the normal case. By defining

$$I_n(a) = \int_{-\infty}^{\infty} [\Phi(ax)]^n\, e^{-x^2}\, dx \quad , \tag{3.3}$$

they have obtained the recurrence relation

$$I_{2m+1}(a) = \sum_{r=1}^{2m+1} \frac{(-1)^{r+1}\binom{2m+1}{r} I_{2m-r+1}(a)}{2^r} \tag{3.4}$$

which is used to obtain the moments up to $n = 5$.

Moments of order statistics from other continuous distributions have been considered by several authors and tables are available to varying extents. Some of the distributions considered are uniform [Hastings et al (1947)], gamma

[Gupta (1960,1962), Breiter and Krishnaiah (1968)], double exponential [Govindarajulu (1966)], logistic [Gupta and Shah (1965), Shah (1966), Gupta et al (1967)] and Cauchy [Barnett (1966)]. As for the discrete distributions, the mean and variance of the smaller of two binomial variates are considered by Craig (1962) and Shah (1966a). Gupta and Panchapakesan (1967) have discussed order statistics arising out of binomial population and have tabulated the first two moments of the largest and the smallest of M independent and identical binomial random variables, each denoting the number of successes in N independent trials with p as the probability of a success, for $N = 1(1)20$, $M = 1(1)10$ and $p = .05(.05).50$.

In some cases we are interested in inequalities concerning the moments of order statistics from distributions F and G which are partially ordered in a certain sense in the space of probability distributions. Van Zwet (1964) considers convex ordered and s-ordered distributions. Some special cases of partial ordering and some properties of order statistics from partially ordered distributions are of interest in selection problems. Suppose that $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$ are two independent random samples from continuous distributions F and G respectively. Let $X_{r,n}$ and $Y_{r,n}$ $(r = 1,\ldots,n)$ denote the order statistics based on each of the two sets of observations. If F is star-shaped with respect to G, that is, $F(0) = G(0) = 0$ and $G^{-1}F(x)/x$ is increasing in $x > 0$, then it has been shown by Barlow and Gupta (1969) that the distribution of $X_{r,n}$ is star-shaped with respect to that of $Y_{r,n}$. Further, for $0 < c < 1$,

$$P\{\max(X_n/X_1,\ldots,X_n/X_{n-1}) \geq c\} \geq P\{\max(Y_n/Y_1,\ldots,Y_n/Y_{n-1}) \geq c\} , \tag{3.5}$$

a result which is used to obtain a lower bound on the probability of a correct selection. Comparisons between linear combinations of order statistics from F and G have been studied by Barlow and Proschan (1966), where (a) F is star-shaped w.r.t. G and (b) F is convex-ordered w.r.t. G. It is known that (b) implies (a). These results have applications in life testing where the underlying distribution has monotone failure rate or monotone failure rate on the average. For illustrating the nature of the results, we state the following theorem proved by Barlow and Proschan.

Theorem 3.1. Let F be star-shaped w.r.t. G. Then $E X_{r,n}/E Y_{r,n}$ is (i) decreasing in r, (ii) increasing in n, and (iii) $E X_{n-r,n}/E Y_{n-r,n}$ is decreasing in n.

If $G(x)=1-e^{-x}$, $x \geq 0$, then F is an IFRA (increasing failure rate on the average) distribution and from the above theorem it follows that

$E X_{r,n}/ \sum_{j=1}^{r} \frac{1}{n-j+1}$ is decreasing in r and increasing in n.

Theorem 3.1 can also be used to obtain bounds on $E X_{r,n}$. If we assume that F and G have the same mean θ, we obtain

$$\text{(3.6)} \qquad \theta E Y_{r,n}/E Y_{r,r} \leq E X_{r,n} \leq \theta E Y_{r,n}/E Y_{1,n-r+1} .$$

Barlow and Proschan have also obtained a number of interesting special results when G is exponential and $G^{-1} F(x)$ is convex.

<u>Theorem 4.4.</u> If $\lim \frac{r}{n} = q$, $0 < q < 1$, as n tends to infinity, then, for sufficiently large n,

(a) $F(E(X_{r,n}))$ exists and

$$F(E(X_{r,n})) = \frac{r}{n+1} + \frac{r(n+1-r)}{2(n+1)^3} \frac{G''(\frac{r}{n+1})}{G'(\frac{r}{n+1})} + o(n^{-1}), \tag{4.4}$$

(b) $\mu_{2k+1}(X_{r,n})/\sigma^{2k+1}(X_{r,n})$ exists $(k = 1,2,\dots.)$ and

$$\mu_{2k+1}(X_{r,n})/\sigma^{2k+1}(X_{r,n}) \tag{4.5}$$

$$= \mu_{2k+1}(r,n)(\mu_2(r,n))^{-k-\frac{1}{2}} +$$

$$2^{-k} \frac{(2k+1)!}{(k-1)!} \left[\frac{r(n+1-r)}{(n+1)^3}\right]^{\frac{1}{2}} \frac{G''(\frac{r}{n+1})}{G'(\frac{r}{n+1})} + o(n^{-\frac{1}{2}}),$$

(c) $\dfrac{E(X_{r,n}) - m(X_{r,n})}{\sigma(X_{r,n})}$ exists and

$$\frac{E(X_{r,n}) - m(X_{r,n})}{\sigma(X_{r,n})} = -\frac{2r-n-1}{3\sqrt{r(n+1-r)(n+1)}} + \tag{4.6}$$

$$\frac{1}{2}\left[\frac{r(n+1-r)}{(n+1)^3}\right]^{\frac{1}{2}} \frac{G''(r/n+1)}{G'(r/n+1)} + o(n^{-\frac{1}{2}}).$$

The result for $E(X_{r,n})$ is well-known [see David and Johnson (1954)]; the result for $F(E(X_{r,n}))$ derived from it closely resembles the corresponding expression given by Blom (1958), who obtains his result under slightly different conditions.

One of the important areas where fruitful research has been accomplished is the theory of extreme order statistics. Contributions have been made in this area nearly over a period of five decades by several authors among whom notably are Fisher and Tippett (1928), Gumbel (1958, 1962) and Gnedenko (1943).. The important problem is to find $\mathcal{L}_k$, the family of all possible (nondegenerate) limit distributions for sequences of the form $b_n^{-1}(X_{k,n} - a_n)$, where a_n and $b_n (b_n > o)$ are constants. For $k = n$, a complete solution with specification of domains of attraction was given by Gnedenko (1943). His results were generalized by Smirnov (1949) who obtained the following theorem.

<u>Theorem 4.5.</u> The family $\mathcal{L}_k$ is given by

$$\Lambda_1^{(k)}(x) = \begin{cases} 0 & x \leq 0,\ \alpha > 0 \\ \frac{1}{(n-k)!} \int_{x^{-\alpha}}^{\infty} e^{-t}\, t^{n-k}\, dt & x > 0,\ \alpha > 0. \end{cases}$$

$$\Lambda_2^{(k)}(x) = \begin{cases} \frac{1}{(n-k)!} \int_{(-x)^{\alpha}}^{\infty} e^{-t}\, t^{n-k}\, dt & x \leq 0,\ \alpha > 0 \\ 0 & x > 0,\ \alpha > 0 \end{cases} \tag{4.7}$$

$$\Lambda_3^{(k)}(x) = \frac{1}{(n-k)!} \int_{e^{-x}}^{\infty} e^{-t}\, t^{n-k}\, dt \qquad -\infty < x < \infty .$$

Berman (1962) shows that the limiting distribution for the maximal order statistic of a random number of independent indentically distributed random variables under certain general conditions is a mixture of distributions of $\mathcal{L}_n$.

Consider a sequence consisting of the sets of random variables $X_{n,1}, \ldots, X_{n,N_n}$; $n = 1,2,\ldots\ldots$ Assume that $E(X_{n,k}) = 0$, $E(X_{n,k}^2) < \infty$ and the random variables in any set are independent.

Let

$$F_{n,k}(x) = P\{X_{n,k} < x\} ,$$

(4.8)
$$S_{n,k}(x) = \sum_{\nu=1}^{k} X_{n,\nu} ,$$

$$B_n^2 = V(S_{n,N_n}) = \sum_{k=1}^{N_n} V(X_{n,k}) .$$

Suppose that

(4.9)
$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^{N_n} \int_{|x|>\varepsilon B_n} x^2 \, dF_{n,k}(x) = 0 , \quad \varepsilon > 0.$$

Under the above conditions the following results have been obtained by Rényi (1953).

<u>Theorem 4.6.</u>

(a)
$$\lim_{n\to\infty} P\{\max_{1\le k\le N_n} S_{n,k} < xB_n\} = \begin{cases} \sqrt{\frac{2}{\pi}} \int_0^x e^{-t^2/2}\, dt , & x > 0 \\ 0 , & x \le 0 . \end{cases}$$

(b)
$$\lim_{n\to\infty} P\{\max_{1\le k\le N_n} |S_{n,k}| < xB_n\} = \begin{cases} \frac{4}{\pi} \sum_{k=0}^{\infty} (-1)^k \frac{e^{-(2k+1)^2\pi^2/8x^2}}{(2k+1)} , & x > 0, \\ 0 , & x \le 0. \end{cases}$$

(c)
$$\lim_{n\to\infty} P\{-yB_n \le \min_{1\le k\le N_n} S_{n,k} \le \max_{1\le k\le N_n} S_{n,k} < xB_n\}$$

$$= \begin{cases} \frac{4}{\pi} \sum_{k=0}^{\infty} e^{-(2k+1)^2\pi^2/2(x+y)^2} \frac{\mathrm{Sin}[(2k+1)\pi x/x+y]}{2k+1} & , x > 0 \text{ and } y \geq 0, \\ 0 & , x \leq 0 \text{ or } y < 0. \end{cases}$$

(d) Let $A_n^2 = V(S_{n, M_n})$ with $1 \leq M_n < N_n$ and $\lim_{n\to\infty} A_n/B_n = \lambda (0 \leq \lambda < 1)$. Then

$$\lim_{n\to\infty} P\{ \max_{M_n < k \leq N_n} |S_{n,k}| < y B_n \}$$

$$= \begin{cases} \frac{4}{\pi} \sum_{k=0}^{\infty} (-1)^k \frac{e^{-(2k+1)^2\pi^2/8y^2}}{2k+1} \left(1 - \sqrt{\frac{2}{\pi}} \int_{y/\lambda}^{\infty} e^{-u^2/2} du + \rho_k\right), & y > 0, \\ 0 & , y \leq 0, \end{cases}$$

where $\rho_k = \frac{2\lambda e^{-y^2/2\lambda^2}}{\sqrt{2\pi}\, y} \int_0^{(2k+1)\pi/2} e^{\lambda^2 u^2/2y^2} \sin u \, du.$

If $y = x$, (c) reduces to (b). In the special case $M_n = 1$ (i.e. for $\lambda = 0$), (d) is identical with (b). For the case where all the variables $X_{n,k}$ have the same distribution, the parts (a) and (b) were proved by Erdös and Kac (1946).

The classical theory of the limiting distribution of the maximum in sequences of independent random variables has been generalized in two directions, namely, (1) when the random variables are exchangeable and (2) when the number of random variables considered in the determination of the maximum is itself a random variable N_n, depending on a non-negative integer-valued parameter n. Let $\{X_n : n \geq 1\}$ be a sequence of exchangeable random varialbes defined on (Ω, G, P), i.e., the joint df denoted by $G_m(x_1, \ldots, x_m)$ for each m is given according to the fundamental theorem of de Finetti (See Loéve (1960) p. 365) by

$$(4.10) \qquad G_m(x_1, \ldots, x_m) = \int_\Omega G_\omega(x_1) \ldots G_\omega(x_m) \, dP(\omega)$$

where for fixed x, $G_\omega(x)$ is a random variable and for each $\omega \in \Omega$, $G_\omega(x)$ is a df in x. For any sequence $X_1, \ldots, X_n$, $P\{X_{(n)} \leq x\} = EG_\omega^n(x)$. The problem is to find sequences $\{a_n\}$ and $\{b_n\}$ and a df $L(x)$ such that $a_n > 0$ and

$$(4.11) \qquad \lim_{n \to \infty} P\{a_n^{-1}(X_{(n)} - b_n) \leq x\} = \lim_{n \to \infty} EG_\omega^n(a_n x + b_n)$$

$$= L(x)$$

for all $x \in C_L$, the set of continuity points of L. Let $\Lambda_i(x) = \Lambda_i^{(n)}(x), i=1,2,3$. The following results are due to Berman (1962).

<u>Theorem 4.7.</u> Suppose that there exists a sequence of positive numbers $\{a_n\}$ and a df $F(x)$ in the domain of attraction of $\Lambda_1(x)$ such that $\lim_{n \to \infty} F^n(a_n x) = \Lambda_1(x)$. Then (a) there exists a nondegenerate df $L(x)$ such that for all $x \in C_L$

$$\lim_{n \to \infty} P\{a_n^{-1} X_{(n)} \leq x\} = \lim_{n \to \infty} EG_\omega^n(a_n x) = L(x)$$

iff there exists a df $A(y)$ such that

$$\lim_{u \to \infty} P\left\{\frac{\log G_\omega(u)}{\log F(u)} \leq y\right\} = A(y), \quad \text{for all } y \in C_A$$

where $A(y)$ satisfies the conditions

$A(\infty) - A(0-) = 1$; $A(0+) - A(0-) < 1$.

(b) $L(x)$ is necessarily of the form

$$L(x) = \begin{cases} 0 & x < 0, \\ \int_0^\infty [\Lambda_1(x)]^y \, dA(y) & x \geq 0. \end{cases}$$

Berman has obtained similar results for the case where we have a sequence of positive numbers $\{a_n\}$ and a real number x_o and a df $F(x)$ in the domain of attraction of $\Lambda_2(x)$ such that $\lim_{n\to\infty} F^n(a_n x + x_o) = \Lambda_2(x)$ and for the case where there exist sequences $\{a_n\}$ and $\{b_n\}$ $(a_n > 0)$ and df $F(x)$ in the domain of attraction of $\Lambda_3(x)$ such that $\lim_{n\to\infty} F^n(a_n x + b_n) = \Lambda_3(x)$. The limiting distribution $L(x)$ is a mixture of $\Lambda_2(x)$ and $\Lambda_3(x)$, in each case, respectively.

Berman has also investigated the case of random number of random variables. Let $\{X_n : n \geq 1\}$ be a sequence of independent random variables with common df $F(x)$ which is in the domain of attraction of $\Lambda(x)$, oneof the three extreme value df's $\Lambda_i(x), i=1,2,3$. Let $\{N_n, n \geq 1\}$ be a sequence of nonnegative, interger-valued random variables distributed independently of the sequence $\{X_n\}$. Let N_n have thedistribution given by $P\{N_n = k\} = p_n(k), k \geq 0$, where for fixed n, $p_n(k) \geq 0$, $\sum_{k=0}^{\infty} p_n(k) = 1$. Define a sequence of random variables W_n as follows.

$$W_n = \begin{cases} -\infty & , N_n = 0 \, , \\ X_{(N_n)} & , N_n > 0 \, . \end{cases} \tag{4.12}$$

Then the df of W_n is

(4.13) $$P\{W_n \leq x\} = \sum_{k=0}^{\infty} p_n(k)\, F^k(x) \;.$$

This df is not necessarily proper:

(4.14) $$\lim_{x\to-\infty} P\{W_n \leq x\} = p_n(0) \geq 0.$$

Suppose $N_n \to \infty$ in probability as $n \to \infty$. Then we have the following theorem.

Theorem 4.8. There exists a df $L(x)$ such that

$$\lim_{n\to\infty} P\{a_n^{-1}(W_n - b_n) \leq x\} = \lim_{n\to\infty} \sum_{k=0}^{\infty} p_n(k)\, F^k(a_n x + b_n)$$

$$= L(x), \quad \text{for all } x \in C_L$$

iff there exists a df $A(y)$ such that

$$\lim_{n\to\infty} P\{n^{-1}N_n \leq y\} = A(y) \quad \text{for all } y \in C_L$$

where $A(y)$ satisfies the conditions

(i) $A(\infty) - A(0-) = 1;\ A(0+) - A(0-) < 1$

or (ii) $A(0+) - A(0-) = 0;\ 0 < A(\infty) - A(0-) \leq 1$

or (iii) $A(0+) - A(0-) = 0;\ A(\infty) - A(0-) = 1,$

according as $\Lambda(x)$ is $\Lambda_1(x)$ or $\Lambda_2(x)$ or $\Lambda_3(x)$. Further $L(x)$ is a mixture of the appropriate $\Lambda(x)$ in each case..

If $\{N_n\}$ and $\{X_n\}$ are not necessarily independent of each other and if there exists a positive number c such that $n^{-1}N_n \to c$ in probability, then it has been shown that, for every x,

(4.15) $$\lim_{n\to\infty} P\{a_n^{-1}(W_n - b_n) \leq x\} = \Lambda^c(x) \quad ,$$

where $\Lambda^c(x)$ is of the same type as $\Lambda(x)$.

In the above set-up, let $\frac{N_n}{n} \to N$ in probability where N is a random variable satisfying $P\{N \leq 0\} = 0$. It is known that, if $E(X_1) = 0$, $V(X_1) = 1$, then as $n \to \infty$

$$(4.16) \qquad P\{\frac{S_{N_n}}{\sqrt{N_n}} \leq x\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \quad \text{for all } x .$$

This result was proved by Anscombe (1952) in the case where N is constant with probability 1. Renyi (1960) extended Anscombe's result to discrete random variables N. Finally, Blum, Hanson and Rosenblatt (1963) and Mogyoródi (1962), independently obtained a proof for arbitrary positive N. Barndorff-Neilson (1964) proves the following result for $X_{(n)}$.

Theorem 4.9. Let $\{a_n\}$ and $\{b_n\}$ be sequences of constants with $a_n > 0$ for all n and Λ be a nondegenerate df The following three statements are equivalent.

(i) $P\{X_{(n)} \leq a_n x + b_n\} \to \Lambda(x)$ for all $x \in C_\Lambda$

(ii) $P\{X_{(N_n)} \leq a_{N_n}$

(iii) $P\{X_{(N_n)} \leq a_n x + b_n\} \to \int_0^\infty [\Lambda(x)]^s dP[N \leq s]$, for all $x \in C_\Lambda$

as $n \to \infty$.

The above theorem was independently discovered about the same time by Lamperti. As is well-known, if (i) holds, then Λ is one of the three extreme value distributions Λ_i, i=1,2,3. Equivalence of (i) and (iii) when N is constant with probability 1 is the result of Berman (1964).

For a study of $X_{(k)}$ as a stochastic process with emphasis on limit theorems, the reader is referred to Dwass (1964) and Lamperti (1964). Dwass discusses the three possible extremal processes and Lamperti studies the joint limiting behavior of $X_{(n)}$ and $X_{(n-1)}$ considered as a two-dimensional process. Limit laws for maxima of a sequence of random variables defined on a Markov Chain have been studied by Fabens and Neuts (1970), and Resnick and Neuts (1970).

Another area of research under asymptotic results is the theory of spacings. Let us first consider n independent observations from a continuous distribution F(x). Define

(4.17) $$U_{r,n} = F(X_{r,n}) \quad .$$

Then $U_{r,n}$ (r=1,..., n) are order statistics from the uniform distribution on (0,1). For the purpose of notational convenience, let us define slightly modified spacings

(4.18) $$D'_{nr} = (n+1)(U_{r,n} - U_{r-1,n}).$$

Let $\{g_n: n \geq 1\}$ be a sequence of real Borel-measurable functions and consider the random variable

(4.19) $$G_n = \sum_{r=1}^{n+1} g_n (D'_{nr}) \quad .$$

Many of the tests based on spacings considered in the literature are of this form. As pointed out by Pyke (1965), prior to 1953 there was no unified approach to the problem of finding the limiting df of a statistic of the form (4.19). Earlier the asymptotic normality of G_n was obtained for special forms of $g_n(x)$ by Moran (1947), Sherman (1950) and Kimball (1950) using different methods. It was Darling (1953) who provided the first general method of deriving limit theorems for G_n by applying the method of steepest descent to a simple formula for the characteristic function of G_n. Le Cam (1958) gave a more easily applied general approach to this problem. Suppose g_n is defined on $[0,\infty)$ and $\{Y_r: r \geq 1\}$ is a sequence of independent exponential random variables with mean 1. Set $S_n = n^{-\frac{1}{2}} \sum_{r=1}^{n+1} (Y_r - 1)$.

Then the df of G_n is the same as the conditional df of $J_n = \sum_{r=1}^{n+1} g_n(Y_r)$ given that $S_n = 0$. The approach of Le Cam is to use information about the joint limiting behavior of (J_n, S_n) to derive the desired conditional limiting distribution of J_n, given $S_n = 0$. For the details of this approach and other results concerning the weak convergence of the empirical distribution function for uniform spacings and limit theorems for functions of general spacings, one may refer to Pyke (1965).

In conclusion of this section we briefly state some recent large sample results concerning sample quantiles and the deviation between the sample quantile process and the sample df. It is fitting here to mention some of the remarks made by Weiss (1970) and Kiefer (1970b). In deriving the asymptotic distribution of a set of sample quantiles, the usual approach is to study the joint probability density function as the sample size increases. This technique gets complicated enough when each element of the sample is itself a k-dimensional random variable and we seek the joint asymptotic distribution of a quantile of the first co-ordinates in the sample, a quantile of the second co-ordinates in the sample,...., a quantile of the k-th co-ordinates in the sample. The simple approach used by Weiss studies limit probabilities of events concerning sample quantiles by rewriting them as events concerning multinomial rando variables. Weiss (1970) illustrates the use of this method in some nonstandard case

Let $X_1, X_2, \ldots$ be independent and identically distributed with common twice differentiable univariate df F on the unit interval I. Assume that $\inf_{x \in I} F'(x) > 0$ and $\sup_{x \in I} F''(x) < \infty$ and let $\xi_p = F^{-1}(p)$.

Also let S_n and $Y_{p,n}$ denote the sample df and the sample quantile of order p, respectively, both based on $(X_1, \ldots, X_n)$; i.e.

(4.20) $$nS_n(x) = [\text{number of } X_i \leq x, \quad 1 \leq i \leq n]$$

and

(4.21) $$Y_{p,n} = \inf\{x: S_n(x) = p\}.$$

Define

$$(4.22) \qquad R_n(p) = Y_{p,n} - \xi_p + [S_n(\xi_p) - p]/F'(\xi_p) .$$

The study of $R_n(p)$ was initiated by Bahadur (1966). Later Kiefer (1967) showed that, for $u > 0$,

$$(4.23) \qquad \lim_{n\to\infty} P\{n^{3/4}F'(\xi_p)R_n(p) \le u\} = 2\int_0^\infty \Phi(k^{-\frac{1}{2}}u)\, d_k\Phi(k/\sigma_p)$$

and that,

$$(4.24) \qquad \limsup_{n\to\infty} \pm F'(\xi_p)R_n(p)/[2^5 3^{-3}\sigma_p^2 n^{-3}(\log\log n)^3]^{\frac{1}{4}} = 1 \text{ with probability } 1$$

where Φ is the standard normal df, $\sigma_p = [p(1-p)]^{\frac{1}{2}}$. Let

$$(4.25) \qquad \begin{cases} R_n^{\pm} = \sup_{p\in I} \pm F'(\xi p)\, R_n(p) \\ R_n^{*} = \max(R_n^{+}, R_n^{-}) \end{cases} .$$

Kiefer (1970a) proves the following results.

<u>Theorem 4.10.</u> For $Q_n = R_n^+$ or R_n^- or R_n^* ,

(a) $n^{3/4} Q_n/(D_n \log n)^{\frac{1}{2}} \to 1$ in probability as $n \to \infty$ where

$$D_n = n^{\frac{1}{2}} \sup_x |S_n(x) - F(x)| .$$

(b) $\limsup_{n\to\infty} n^{3/4}(\log n)^{-\frac{1}{2}}(\log\log n)^{-\frac{1}{4}} Q_n = 2^{-\frac{1}{4}}$ with probability 1.

The consequence of part (a) is that, for $t > 0$,

$$(4.26) \qquad \lim_{n\to\infty} P\{n^{3/4}(\log n)^{-\frac{1}{2}} Q_n > t\} = 2\sum_{m=1}^{\infty} (-1)^{m+1} e^{-2m^2 t^4} .$$

For some of the consequences of part (b) and a list of open problems, the reader is referred to Kiefer (1970a).

5. Combinatorial Methods in Order Statistics

One may see that some elementary combinatorial arguments are always involved in the study of order statistics, for example, in writing the cdf of the r-th order statistic based on n observations. But in order to throw light on applications of combinatorial methods of deeper significance we interpret order statistics in a broad sense to include Kolmogorov - Smirnov statistic which requires knowledge of the actual ordered observations in the sample only up to a monotonic increasing transformation. Many combinatorial problems arise when we want to compare theoretical and empirical distribution functions. A fundamental theorem of much application in this area is a generalized version of the classical ballot theorem. A brief but interesting summary of the historical development of the classical ballot theorem and some of its generalizations is given in Takács (1970) to whom the following theorem is due.

Theorem 5.1. Let $k_1, k_2, \ldots, k_n$ be nonnegative integers with sum $k_1 + k_2 + \ldots + k_n \leq n$. Among the n! permutations of $(k_1, k_2, \ldots, k_n)$ there are exactly $(n-1)!(n-k)$ for which the r-th partial sum $\sum_1^r k_i$ is less than r for all $r = 1,2,\ldots,n$.

The above theorem was first obtained by Takács in 1960 and the proofs first given by him (1961, 1962) were based on mathematical induction. Later in 1967 he gave a direct combinatorial proof of the theorem. Recently this theorem has been formulated by Takács (1970) in the following slightly more general form.

Theorem 5.2. Let $X_1, X_2, \ldots, X_n$ be exchangeable random variables taking on nonnegative integer values. Set $S_r = X_1 + X_2 + \ldots X_r$ for $r = 1,2,\ldots,n$.

Then

$$(5.1) \qquad P\{S_r < r \text{ for } r = 1,\dots,n \mid S_n = k\} = \begin{cases} 1-\frac{k}{n} & \text{for } k = 0,1,\dots,n, \\ 0 & \text{otherwise,} \end{cases}$$

where the conditional probability is defined up to an equivalence.

As an example of a useful application of Theorem 5.1, consider n random points which are distributed independently and uniformly on the interval $(0,t)$. Let $\chi(u)$ $(0 \le u \le t)$ be c times the number of points in the interval $(0,u]$ where c is a positive constant. Then, by using Theorem 5.1 [see Takács (1970)], we can show that

$$(5.2) \qquad P\{\chi(u) \le u \text{ for } 0 \le u \le t\} = \begin{cases} 1-\frac{nc}{t} & \text{for } 0 \le nc \le t, \\ 0 & \text{otherwise.} \end{cases}$$

Another important combinatorial theorem which together with Theorem 5.2 leads to many applications is the following theorem which is due to Andersen (1953) and Feller (1959).

<u>Theorem 5.3</u>. Let $X_1, X_2,\dots,X_n$ be interchangeable random variables taking on real values. Define $S_r = X_1 + \dots + X_r$ for $r = 1,2,\dots,n$ and $S_0 = 0$. Denote by N_n and N_n^* respectively the number of positive and nonnegative members in the sequence $S_1, S_2,\dots,S_n$. Denote by L_n and L_n^*, the subscripts of the first and the last maximal members in the sequence $S_0, S_1,\dots,S_n$. We have

$$(5.3) \qquad P\{N_n = j\} = P\{L_n = j\}$$

and

$$(5.4) \qquad P\{N_n^* = j\} = P\{L_n^* = j\}$$

for $j = 0,1,\dots,n$.

Theorems 5.2 and 5.3 can be combined to yield the following interesting result.

Theorem 5.4 (Takács 1970). Let $X_1, X_2, \ldots, X_n$ be interchangeable random variables taking on nonnegative integers. Set $S_r = X_1 + \ldots + X_r$ for $r = 1,2,\ldots,n$ and $S_0 = 0$. Denote by Δ_n the number of subscripts $r = 1,2,\ldots,n$ for which $S_r < r$ holds. If $P\{S_n = n-1\} > 0$ then we have

$$(5.5) \qquad P\{\Delta_n = j \mid S_n = n-1\} = \frac{1}{n}$$

for $j = 1,2,\ldots,n$.

Takács (1970) has proved a number of auxiliary theorems which can be used in the theory of order statistics Allthese theorems are consequences of Theorems 5.2 and 5.3 and are concerned with the distributions of $\Delta_n^{(c)}$, the number of subscripts $r = 1,2,\ldots,n$ for which $S_r < r+c$ where $c = 0, \pm 1, \pm 2, \ldots$. In particular, $\Delta_n = \Delta_n^{(0)}$.

We shall now indicate the applications of these results to the problem of comparing a theoretical and an empirical distribution function . Let $X_1, X_2, \ldots, X_n$ be mutually independent random variables having common df $F(x)$. Let $F_n(x)$ be the empirical df , i.e., $n F_n(x)$ = the number of variables $\leq x$. Consider

$$(5.6) \qquad D_n^+ = \sup_{-\infty<x<\infty} [F_n(x) - F(x)]$$

$$= \max_{1 \leq r \leq n} [F_n(X_{(r)}) - F(X_{(r)})].$$

If we assume that $F(x)$ is continuous df , then the joint distribution of $\delta_n(r) = F_n(X_{(r)}) - F(X_{(r)})$ $(r = 1,2,\ldots,n)$ does not depend on $F(x)$ and consequently the distributions of D_n^+, ρ_n and ρ_n^* are also independent of $F(x)$, where ρ_n denotes the number of non-negative elements among $\delta_n(r)$ $(r = 1,2,\dots,n)$ and ρ_n^* denotes the largest r for which $\delta_n(r)$ attains its maximum. By using the auxiliary theorems we can obtain the distributions of D_n^+, ρ_n and ρ_n^*. These distributions have been obtained earlier by several authors [see Takács (1970)] and are given below.

Theorem 5.5.

(a) If $0 < x \leq 1$, then

$$P\{D_n^+ \leq x\} = 1 - \sum_{nx \leq j \leq n} \frac{nx}{n+nx-j} \binom{n}{j} \left(\frac{j}{n} - x\right)^j \left(1+x-\frac{j}{n}\right)^{n-j} . \tag{5.7}$$

(b) For $j = 1,2,\dots,n$,

$$P\{\rho_n=j\} = P\{\rho_n^*=j\} = \frac{1}{n} \sum_{i=1}^{j} \frac{1}{i} \binom{n}{i-1} \left(\frac{i}{n}\right)^{i-1} \left(1-\frac{i}{n}\right)^{n-i}. \tag{5.8}$$

Also of interest is the problem of comparing two empirical distribution functions. Let $X_1, X_2,\dots, X_m$ and $Y_1, Y_2,\dots, Y_n$ be independent random samples from the distributions $F(x)$ and $G(x)$, respectively. Denote by $F_m(x)$ and $G_n(x)$ the empirical distribution functions of the two samples. Define

$$\begin{aligned} D^+(m,n) &= \sup_{-\infty<x<\infty} [F_m(x) - G_n(x)] \\ &= \max_{1\leq r\leq n} [F_m(Y_{(r)}) - G_n(Y_{(r)}-0)]. \end{aligned} \tag{5.9}$$

Let $\gamma_c(m,n)$ denote the number of subscripts $r = 1,2,\dots,n$ for which $F_m(Y_{(r)}) < G_n(Y_{(n)}) - c/n$, where $c = 0, \pm 1,\dots, \pm(n-1)$ and let $\tau(m,n)$ denote the smallest $r = 1,2,\dots,n$ for which $F_m(Y_{(r)}) - G_n(Y_{(r)}-0)$ attains its maximum. If F and G are identical continuous distributions, then the distributions of $D^+(m,n)$, $\gamma_c(m,n)$ and $\tau(m,n)$ do not depend on F. When $n = mp$, these distributions can be derived easily by appealing to the auxiliary theorems discussed earlier and a simple probability result relating to the drawing of the ith white ball at the (i+s)th draw when the balls are drawn without replacement from a box containing m black and n white balls. We state below the results relating to $D^+(m,n)$.

Theorem 5.6.

(a) If $n = mp$ where p is a positive integer, and $c = 0,1,\ldots,n$, then

$$(5.10)\quad P\{D^+(m,n) \le c/n\} = 1 - \frac{1}{\binom{m+n}{n}} \sum_{(c+1)/p \le s \le m} \frac{c+1}{n+c+1-sp} \binom{sp+s-c-1}{s} \binom{m+n+c-sp-s}{m-s} .$$

(b) For $0 < x \le 1$ and $n=mp$,

$$(5.11)\quad \lim_{p\to\infty} P\{D^+(m,n) \le c/n\} = P\{D_m^+ \le x\}$$

where $c = [nx]$.

For more details, the reader is referred to Takács (1970).

Vincze (1970) observes that proofs of Kolmogorov - Smirnov type distribution theorems can be simplified by using a certain generalization of the ballot lemma by G. Tusnády stated below.

Theorem 5.7. Let $A_0, A_1, \ldots, A_n$ be a complete system of events, for which $P(A_0)=q$ and $P(A_j)=p$, $j = 1,2,\ldots,n$ holds. Making n independent observations, let ν_i be the number of cases in which A_i occurred. Then

$$(5.12) \qquad P\{\sum_{i=1}^{j} \nu_i < j \ : \ j = 1,2,\ldots,n\} = q.$$

The above theorem is equivalent to the following theorem of Daniels (1945).

Theorem 5.8. If $F_n(x)$ denotes the empirical distribution function corresponding to a sample of size n taken on a random variable with distribution F(x) which is uniform in (0,1), then

$$(5.13) \qquad P\{\frac{F_n(x) - F(x)}{F(x)} < y, \ 0 \leq x \leq 1\} = \frac{y}{y+1} \ (0 \leq y \leq \infty).$$

The equivalence of Theorems 5.7 and 5.8 was utilized by K. Sarkadi to give an independent proof of Theorem 5.7. Vincze (1970) also refers to a generalized ballot lemma of E. Csáki which is closely related to results of Nef (1964) and gives the following formulation of the generalized ballot lemma in terms of the empirical distribution function.

Let $F_n(x)$ be the empirical distribution function belonging to a sample of size n taken on a random variable distributed uniformly in (0,1). Let $\bar{\lambda}$ denote the number of (horizontal) intersections of the graph of $F_n(x)$ with the straight line $y = \frac{1}{np} x$ $(np \leq 1)$. Then

$$(5.14) \qquad P\{\bar{\lambda} \geq \ell\} = \ell! \binom{n}{\ell} p^{\ell} \quad (\ell = 0,1,2,\ldots,n).$$

6. Some Combinatorial Methods in Fluctuation Theory and the Distribution of the Maxima.

In this section we describe some results concerning the partial sums of a sequence of random variables. Although fluctuations of partial sums of random variables have been investigated in special cases for a long time, and even in more general cases for the purpose of finding limit theorems, the idea of using combinatorial methods for analyzing the partial sums of a fixed finite set of more general random variables goes to the credit of E. Sparre Andersen who made a fundamental contribution in this area.

Let $\{X_k\}$, $k = 1,2,\ldots, n$, be a sequence of independent and identically distributed random variables, with partial sums $S_0 = 0$, $S_1 = X_1, \ldots,$ $S_n = X_1 + \ldots + X_n$. Let

$$(6.1)\qquad \begin{cases} N_n = \text{the number of positive } S_n \text{ among } S_1, \ldots, S_n. \\ \\ L_n = \text{the smallest index } k(= 0,1,\ldots,n) \text{ with } S_k = \max\limits_{0\leq m\leq n} S_m. \end{cases}$$

The variable N_n serves in a way as a "measure" of the ups and downs of the sequence $S_0, S_1, \ldots, S_n$. For any permutation σ: $i_1, \ldots, i_n$ of the integers $1,2,\ldots,n$ define $N_n(\sigma)$ and $L_n(\sigma)$ as in (6.1) in terms of the partial sums $S_k(\sigma) = X_{i_1} + \ldots + X_{i_k}$ of the permuted variables $\sigma(X_1,\ldots,X_n) = (X_{i_1},\ldots,X_{i_n})$.

By the basic assumption it is implied that $N_n(\sigma)$ and $L_n(\sigma)$ have the same distributions as N_n and L_n. As a matter of fact, if we consider the whole class $\{N_n(\sigma)\}_{(\sigma)}$ of $n!$ variables, each element of the class has the same distribution as N_n. By successfully seeking properties of the whole class which do not depend upon the particular values of the variables $X_1,\ldots,X_n$, we can as well carry out the analysis for a set of numbers $x_1,\ldots,x_n$ instead of the variables $X_1,\ldots,X_n$. The following theorem due to Andersen (1953) gives the equivalence principle.

Theorem 6.1 $$\{N_n(\sigma)\}_{(\sigma)} \equiv \{L_n(\sigma)\}_{(\sigma)} \quad .$$

There are two essential facts connected with the above theorem. First, $N_n(\sigma)$ and $L_n(\sigma)$ are integers between 0 and n, so that there will be multiplicities among the integers which are assumed by the $n!$ terms in each set. The theorem asserts that these multiplicities are exactly the same in the two sets $\{N_n(\sigma)\}_{(\sigma)}$ and $\{L_n(\sigma)\}_{(\sigma)}$. Secondly, the identity holds for all sets of numbers $x_1,\ldots,x_n$ and, therefore, it is not directly concerned with probability theory.

Theorem 6.1, restated in terms of the sequence of random variables, gives

$$P\{N_n = j\} = P\{L_n = j\} \quad . \tag{6.2}$$

This is exactly (5.3) of Theorem 5.3. Thus, the equivalence principle permits us to translate statements concerning the position of maximal terms into statements concerning the number of positive terms: usually the statements of the first kind are more readily proved whereas those of the second kind are more important. Further, Andersen (1953) has shown that, if $X_1,\ldots,X_n$ are independent and identically distributed with continuous and symmetric distributions,

$$P\{L_n = m\} = \binom{2m}{m}\binom{2n-2m}{n-m} 2^{-2n}, \quad 0 \le m \le n \quad . \tag{6.3}$$

It should be pointed out that the joint distribution of N_n and L_n is <u>not</u> distribution-free.

Let $R_{n0} \ge R_{n1} \ge \cdots \ge R_{nn}$ be an ordering of the partial sums $S_0,S_1,\ldots,S_n$. Since the distribution of X_1 is continuous, there is a unique index m such that $R_{nk} = S_m$, with probability one.We define $L_{nk} = m$ if $R_{nk} = S_m$. Darling (1951) found the distribution of L_{nk} in terms of products of binomial coefficients, but he gave no results for joint distributions. Baxter (1962) has proved the following theorem.

Theorem 6.2. For all $0 \le m, k \le n$, $(n \ge 1)$,

$$(6.4)\quad P\{L_{nm} = 0, L_{nk} = n\} = \begin{cases} (1/2n)\binom{2m}{m}\binom{2n-2k}{n-k}2^{-2n-2m+2k}, & m < k, \\ 0, & m = k, \\ (1/2n)\binom{2k}{k}\binom{2n-2m}{n-m}2^{-2n-2k+2m}, & m > k \end{cases}.$$

We note that $L_{nm} = 0$ is equivalent to $N_n = m$. Also, $L_{nk} = n$ means that there are exactly k partial sums greater than S_n. Thus Theorem 6.2 provides the joint distribution of the number of partial sums less than $S_0 (= 0)$ and the number of partial sums greater than S_n. In particular, for $k = 0$,

$$(6.5)\quad P\{N_n = m, L_n = n\} = (1/2n)\binom{2n-2m}{n-m} 2^{-2n+2m}, \quad 1 \le m \le n \quad .$$

Now, we consider again a sequence of mutually independent random variables with a common distribution function and define

$$(6.5)\quad \begin{cases} a_n = P\{S_n > 0\}, \quad a_n^* = P\{S_n \ge 0\}, \\ \mu_n = P\{S_n > S_j, j = 0,\dots, n-1\}, \quad n \ge 1, \\ \mu_n^* = P\{S_n \ge S_j, j = 0,\dots, n-1\}, \quad n \ge 1 \end{cases}.$$

Then, the generating functions $\mu(t) = \Sigma\mu_n t^n$ and $\mu^*(t) = \Sigma\mu_n^* t^n$ have been obtained by Andersen (1953, 1954) and also by Spitzer (1956) in the form:

(6.6) $$\mu(t) = \exp \sum_{1}^{\infty} \frac{a_n}{n} t^n \quad \text{and}$$

(6.7) $$\mu^*(t) = \exp \sum_{1}^{\infty} \frac{a_n^*}{n} t^n \quad .$$

The above result shows that the knowledge of the sequences $\{a_n\}$ and $\{a_n^*\}$ suffices for the calculation of the distribution of the position of the maximal term and of the number of positive terms in $\{S_0,\ldots,S_n\}$.

The probability

(6.8) $$\nu_n^* = P\{S_1 \leq 0,\ldots, S_n \leq 0\}$$

has the associated generating function

(6.9) $$\nu^*(t) = \exp \sum_{1}^{\infty} \frac{1-a_n}{n} t^n = [(1-t)\mu(t)]^{-1} \quad .$$

Let

(6.10) $$p_{k,n} = P\{S_k > S_j \text{ for } j < k,\ S_k \geq S_j \text{ for } k < j \leq n\} \quad .$$

Then Feller (1959) gives the following theorem.

<u>Theorem 6.3.</u> $$p_{k,n} = \mu_k \nu_{n-k}^* \quad .$$

As we can see, $p_{k,n}$ is the probability that the first maximum in $(S_0,S_1,\ldots,S_n)$ occurs at the place numbered k. If $p_{k,n}^*$ denotes the probability that the last maximum occurs at the place numbered k, then

(6.11) $$p_{k,n}^* = \mu_k^* \nu_{n-k} \quad .$$

Instead of $a_n = P\{S_n > 0\}$, let us consider more generally the truncated distribution function

(6.12) $$F_n(x) = P\{0 < S_n \leq x\} \ , \quad x > 0$$

with the Laplace transform

(6.13) $$\phi_n(\lambda) = \int_{0+}^{\infty} e^{-\lambda x} \, dF_n(x) \quad .$$

Define

(6.14) $$H_n(x) = P\{S_j < S_n \leq x, \ j = 0,\ldots, n-1\}, \quad x > 0$$

and

(6.15) $$h_n(\lambda) = \int_0^{\infty} e^{-\lambda x} dH_n(x), h(\lambda,t) = 1 + \sum_1^{\infty} h_n(\lambda) \ t^n \quad .$$

The following theorem is due to Spitzer (1956) .

Theorem 6.4. $$h(\lambda,t) = \exp \sum_1^{\infty} \frac{t^n}{n} \phi_n(\lambda) \quad .$$

For $\lambda = 0$, this theorem gives (6.6)

Again considering a sequence $X_1, X_2, \ldots, X_n, \ldots$ of independent and identically distributed random variables, having continuous distribution function, let us define X_n to be outstanding if it is larger than all previous observations, that is $X_n > \max_{1 \leq k \leq n-1} X_k$. Let A_n be the event that X_n is an outstanding observation $(n = 1,2,\ldots)$. Renyi (1962) has obtained some results concerning the outstanding observations based on the simple but surprising fact that the events $A_1, \ldots, A_n, \ldots$ are independent and $P(A_n) = 1/n$. The results of Renyi are contained in the following theorem.

Theorem 6.5. Let $X_{\nu_1}, X_{\nu_2}, \ldots, X_{\nu_k}, \ldots$ be all the outstanding observations of $\{X_k\}$, $k = 1,2,\ldots$. Then

(a) $\lim_{k \to \infty} \nu_k^{1/k} = e$ with probability 1 and

(b) $(\log \nu_k - k)/k^{1/2}$ is asymptotically $(k \to \infty)$ normal with mean 0 and variance 1.

If we define α_N as the number of outstanding observations among $X_1,\ldots,X_N$, then Theorem 6.5 says that $\lim_{N \to \infty} \frac{\alpha_N}{\log N} = 1$ with probability 1 and the distribution of $(\alpha_N - \log N)/(\log N)^{1/2}$ is asymptotically standard normal.

7. Some Estimation and Hypothesis Testing Problems Based on Order Statistics.

Order statistics have been employed in many problems of estimation and testing of hypotheses. The usual methods, in some cases, lead to estimators involving order statistics. An example of practical interest where the observations arise in an ordered sequence is a life test experiment where a certain number of units are put on test and their failure times are observed. The literature has grown so enormously in this area that any attempt to survey all the results will be beyond the aim of the present paper. We will be content with a brief outline of some of the problems investigated.

An important paper is that of Lloyd (1952) in which he considers the least-squares estimates of location and scale parameters using order statistics. Suppose that $X_1, X_2, \ldots, X_n$ are independent observations on X having a continuous distribution $F(\frac{x-\mu}{\sigma})$, $\sigma > 0$, where μ and σ are not necessarily the mean and standard deviation respectively. Define $U = (X-\mu)/\sigma$. Then $U_{(r)}=(X_{(r)}-\mu)/\sigma$ $(r=1,2,\ldots,n)$ can be regarded as ordered observations on U. Let $E(U_{(r)})=\alpha_r$, $V(U_{(r)})=v_{rr}$, $Cov(U_{(r)}, U_{(s)})=v_{rs}$ and V be the matrix (v_{rr}). Under the generalized Gauss-Markov linear model, one can obtain μ^* and σ^*, the least-squares (ℓ.s.) estimates of the parameters μ and σ. The formulas for the estimates and their dispersion matrix simplify considerably when X has a symmetric distribution, in which case we can take μ to be the center of the distribution and σ a symmetric measure of dispersion. Since the ℓ.s. estimates are linear compounds of the ordered observations with minimal variance, $V(\mu^*) \leq \frac{\sigma^2}{n}$, where $\sigma^2=V(X)$. Lloyd has obtained conditions to determine when $V(\mu^*) < \sigma^2/n$, i.e., the ℓ.s. estimate is more efficient than the sample mean. It turns out that $V(\mu^*)< \sigma^2/n$ unless μ^* is the sample mean, a result due to Downton (1953).

Blom (1956, 1958, 1962) addressed himself to the problem of unbiased nearly best linear estimates where one settles for an estimate with nearly minimum variance. He investigated how such an approximation to the best linear estimate can be found. He has also dealt briefly with relaxing the unbiasedness, seeking nearly unbiased, nearly best estimates.

Bennett (1952), in his unpublished thesis, studied the asymptotic properties of estimates which are linear functions of the order statistics with continuous weight functions. Following Bennett (1952) and Jung (1955, 1962), let $T_n = n^{-1} \sum_{j=1}^{n} J(\frac{j}{n+1}) X_{j,n}$ where $J(\cdot)$ is a well-behaved function. Bennett obtained asymptotically optimal J's for both the uncensored and multi-censored cases, but did not derive the asymptotic normality of the estimates. Some of his results were independently obtained by Jung (1955) under rather restrictive conditions. Plackett (1958) and Weiss (1963) independently considered the case where all observations below the p-th and above the q-th sample quantiles $(0 < p < q < 1)$ are censored and obtained asymptotic normality for suitable linear combinations of the available order statistics. Chernoff, Gastwirth and Johns (1967) obtain a quite general theorem concerning the conditions under which the statistics of the form $T_n = n^{-1} \Sigma c_{j,n} h(X_{j,n})$ are asymptotically normally distributed. They specialize their results to the case where $c_{j,n} = J(\frac{j}{n+1})$. These theorems involve the decomposition $T_n = \mu_n + Q_n + R_n$, where μ_n is non-random, $Q_n = n^{-1} \Sigma \alpha_{j,n}(Z_j - 1)$ where Z_j's are independent and identically distributed exponential random variables, $n^{1/2} Q_n$ is asymptotically normal and R_n is asymptotically negligible. Results overlapping with those of Chernoff et al have been obtained by Govindarajulu (1965) whose technique is based on some unpublished results of Le Cam and whose main result requires bounds on $J(u)$ and $J'(u)$ as $u \to 0$ or 1, which is not necessary for the results of Chernoff et al.

Problems of estimation of parameters using censored data from normal as well as non-normal distributions have been studied by several authors. Among the non-normal distributions considered are Gamma [Harter and Moore (1967)], Log normal [Harter and Moore (1966)], Double Exponential [Govindarajulu (1966)], Weibull [Cohen (1965), Gumbel (1958)] and Logistic [Gupta, Shah and Qureishi (1967)]just to mention a few. The published literature on life testing and reliability problems is quite vast and the reader is referred to the bibliographies of Mendenhall (1958), Govindarajulu (1964) and a short classified list of David (1970, p. 124).

In the problems of estimation using only some of the order statistics, an interesti question is how to choose or 'space' the order statistics to obtain good estimates. Let us choose $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_k < 1$. The sample quantiles are $X_{(n_j)}$, $j=1,\ldots,k$, where $n_j = [n\lambda_j] + 1$. Ogawa (1951) considered estimation of the location (μ) and scale (σ) parameters based on sample quantiles in large samples. In these cases, the relative efficiency of an estimate which is a function of the chosen order statistics as compared to those which are based on the whole sample is defined by the ratio of the amounts of information in Fisher's sense in the two cases. The best linear unbiased estimator in each case is found to be efficient for a given spacing $\lambda_1,\ldots, \lambda_k$. However, the efficiency can be raised by suitably choosing the values of $\lambda_1,\ldots, \lambda_k$ for which the relative efficiency of an estimator attains its maximum. Such a set of $\lambda_1,\ldots, \lambda_k$ is called an optimum spacing. Ogawa (1962a) has shown that, in the case of normal distribution, the optimum spacing for the location parameter μ is necessarily a symmetric one. Ogawa (1962c) has also considered optimum spacing for the scale parameter of the exponential distribution. The problem of optimum spacing for the asymptotically best linear estimate (ABLUE) of μ when σ is known has so far been considered for three symmetric distributions with support

$(-\infty,\infty)$. For the normal and logistic distributions it has been proved respectively by Higuchi (1954) and Gupta and Gnanadesikan (1966) that the optimum spacing is symmetric. The question of whether the optimum spacing for the ABLUE of μ when σ is known is symmetric for any distribution which is symmetric and has the support $(-\infty,\infty)$ has been raised and answered in the negative by Kulldorff (1971b) who gives a counter-example. Optimum spacings for the ABLUE of the location parameter μ of an extreme value distribution of Type I ($\Lambda_3^{(n)}(x-\mu)$ given by (4.7)) and the scale parameter σ of an extreme value distribution of Type II or III ($\Lambda_1^{(n)}(\frac{x}{\sigma})$ or $\Lambda_2^{(n)}(\frac{x}{\sigma})$ given by (4.7)) have been considered by Kulldorff (1971a) by making use of the previous results for the scale parameter of an exponential distribution. The problem of determining the optimum choice of the ranks $n_1 < n_2 < \ldots < n_k$ of order statistics in a small sample of size n for estimating the parameters of the exponential distribution $F(x) = 1 - e^{-(x-\alpha)/\sigma}$ $(x \geq \alpha,\ \sigma > 0)$ has been dealt with by Harter (1961b) and Siddiqui (1963) for the case $k > 1$, and by Ukita (1955), Harter (1961b) Sarhan, Greenberg and Ogawa (1963) and Siddiqui (1963) for $k = 2$. The case of general k has been investigated by Kulldorff (1963).

The problems of testing of hypotheses in life test models illustrate the use of order statistics. Some quick tests based on order statistics have been used in several situations; see David and Johnson (1956). Tests for outliers and slippage are further specific problems where test statistics are based on ordered observations.

As regards the use of spacings in testing of hypotheses, Ogawa (1962b) considered the test for the hypothesis H: $\mu = \mu_0$ for the normal mean and the test of the homogeneity of several means. He also discusses selection of the optimum spacing for testing purposes. In another paper (1962d) he discusses the test for

H: $\sigma = \sigma_0$ where σ is the scale parameter of the exponential. Pyke (1965) gives limit theorems for general spacings which are useful in obtaining asymptotic results on the power of goodness-of-fit tests based on spacings. Proschan and Pyke (1967) have discussed tests for monotone failure rate using test statistics based on spacings. Recently, Sethuraman and Rao (1970) have discussed the Pitman efficiencies of tests based on spacings in the goodness-of-fit problem.

8. Multiple Decision (Selection and Ranking) Problems.

The goal in selection and ranking problems can be roughly described as follows. Suppose there are k populations $\pi_1,\ldots,\pi_k$ which are ranked in a certain sense. There are two basic approaches. We may either want to select one of them as the 'best' or select a subset of the given populations so that the selected subset contains the 'best'. Obviously, basing our inference on a sample, we will be content if we can say that the probability of our selection being a correct selection (CS) is at least $P^*(1/k < P^* < 1)$. In the case of selecting a subset we can achieve this regardless of the true states of the distributions. In the case of selecting one of them as the best, we shall require that the condition on the minimum probability of a correct selection be met whenever the true best population is sufficiently apart from the second best. This is the indifference zone formulation of Bechhofer (1954) in its simplest form. The former, known as the subset selection formulation, is due to Gupta (1956).

For a detailed account of subset selection formulation one can refer to Gupta (1965), and Gupta and Panchapakesan (1969, 1971). The monograph of

Bechhofer, Kiefer and Sobel (1968) describes the basic formulation of selection and ranking problems using indifference zone approach.

Many of the multiple decision problems encountered in practice have a common algebraic structure and these problems are called identification problems. Let $\{X_{ij}\}$, $i = 1,\ldots,k$, be k independent sequences of random variables. For $i = 1,\ldots,k$, the X_{ij} have a common distribution $F^0_{j_i}$, where $(F^0_{j_1},\ldots,F^0_{j_k})$ is a permutation of k known probability laws $F^0_1,\ldots,F^0_k$. The space $\Omega=\{(F^0_{j_1},\ldots,F^0_{j_k})\}$ can be viewed as the permutation group S_k on k elements. We denote a typical element of S_k by $(\alpha(1),\ldots,\alpha(k)) = (\alpha_1,\ldots,\alpha_k)$, which is the result of the permutation α on $(1,\ldots,k)$. We may briefly use α to denote an element of S_k. Now, S_k can be regarded either as the space of all possible states of nature or as a group of transformations (permutations) operating on Ω. As a result of this dual interpretation of S_k, if $\alpha, \beta \in S_k$, then $\beta\alpha$ can be considered as the element of Ω arising from the permutation β operating on element α of Ω.

We say α is the true element of Ω if the sequences $\{X_{\alpha_i j}\}$ have the distributions F^0_i, $i = 1,\ldots,k$. If α^{-1} denotes the inverse permutation of α, then we can also say (when α is true) that X_{ij} has the distribution $F^0_{\alpha^{-1}(i)}$ $(i=1,2,\ldots,k)$.

It is also convenient to think in terms of k numbered populations $\pi_1, \pi_2,\ldots,\pi_k$ with the sequence $\{X_{ij}\}$ coming from π_i. By saying that α is the true element of Ω we mean that the population π_i has the distribution $F^0_{\alpha^{-1}(i)}$, or more briefly, $\alpha^{-1}(i)$, $i = 1,2,\ldots,k$. Thus the correct pairing of the populations with the distribution functions can be written in two equivalent ways,

$$(8.1) \qquad \begin{pmatrix} F^0_1 & F^0_2 & \ldots & F^0_k \\ \\ \pi_{\alpha_1} & \pi_{\alpha_2} & \ldots & \pi_{\alpha_k} \end{pmatrix} = \begin{pmatrix} F^0_{\alpha^{-1}(1)} & F^0_{\alpha^{-1}(2)} & \ldots & F^0_{\alpha^{-1}(k)} \\ \\ \pi_1 & \pi_2 & \ldots & \pi_k \end{pmatrix} .$$

An identification problem must satisfy certain requirements. Let D^t denote the space of possible decisions and d a typical element of D^t. Then we require that there exists a group Γ of transformations homomorphic to S_k and operating on D^t. In other words, if g_α is the element of Γ corresponding under the homomorphism $\aleph$ to α, then $g_\alpha g_\beta d = g_{\alpha\beta} d$ for $\alpha,\beta \in S_k$ and $d \in D^t$. Further it is also required that the loss function $W(\alpha,d)$ has the invariance property, that is, $W(\alpha,d) = W(\beta\alpha, g_\beta d)$ for α, $\beta \in S_k$ and $d \in D^t$.

Finally, many decision procedures which are used in conjunction with an identification problem have a corresponding invariant structure under a group of transformations g_α isomorphic to S_k. Let $\alpha \underline{x}_j = \alpha(x_{1j},\ldots,x_{kj}) = (x_{\alpha_1 j},\ldots,x_{\alpha_k j})$, where x_{ij} is a realization of X_{ij}. Let Δ denote a subset of D^t and $P_m\{\Delta|\underline{x}\}$, the probability of arriving at one of the decisions belonging to Δ on the basis of the observations $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_m)$. Then for an invariant procedure, $P_m\{\Delta|\underline{x}\} = P_m\{g_\alpha(\Delta)|\alpha\underline{x}\}$, where $\alpha\underline{x} = (\alpha\underline{x}_1,\ldots,\alpha\underline{x}_m)$.

The description of the basic structure of identification problems given above is on the lines of Bechhofer et al. They have provided several examples to illustrate the basic structure and the properties of minimal invariant sets.

Before we pass on to discuss the role of order statistics in the context of subset selection procedures, we will briefly explain an identification problem and its connection with a ranking problem. Let $\pi_1, \pi_2,\ldots,\pi_k$ be k populations and F_{θ_i} be the distribution associated with $\pi_i (i = 1,2,\ldots,k)$. For the identification problem, we assume that ranked values of the θ_i, denoted by $\theta^o_{[1]} \leq \theta^o_{[2]} \leq \cdots \leq \theta^o_{[k]}$, are known a priori. However, the true pairing of the π_i with the $\theta^o_{[j]}$ is unknown to the experimenter and he has no apriori knowledge relevant to the true pairing of the π_i with the $\theta^o_{[j]} (i, j = 1,2,\ldots,k)$. Suppose that $\theta^o_{[k-1]} < \theta^o_{[k]}$. Then an identification goal would be "to identify the population π_i associated with $\theta^o_{[k]}$". For a ranking problem, we assume that the ordered values of the θ_i, denoted now by $\theta_{[1]} \leq \theta_{[2]} \leq \cdots \leq \theta_{[k]}$, are unknown a priori and that the true pairing or any knowledge relevant to the true pairing of the π_i with the $\theta_{[j]}$ (i, j = 1,2,...,k) is not available to the experimenter. The ranking goal corresponding to the identification goal stated above would be "to select a population π_i associated with $\theta_{[k]}$." But the formulation of the ranking problem will be complete only if the experimenter specifies certain constants, and then states an associated probability requirement involving these constants which must be guaranteed. Several ranking procedures have been investigated for specific cases under this formulation by several authors, notably, Bechhofer and Sobel among them. Generally these procedures have been proposed on heuristic grounds. As one can intuitively see that the decisions in all these cases depend on the sample observations through the ordered values of statistics $T_i (i = 1,2,...,k)$.

We now discuss the role of order statistics in subset selection problems. Let $\pi_1, \pi_2,...,\pi_k$ be k independent populations with continuous distributions $F_{\theta_i} (i = 1,...,k)$, $\theta_i \in \Theta$, an interval on the real line. We assume that $\{F_\theta\}$ is a stochastically increasing family. The ordered values of the unknown parameters θ_i are denoted by $\theta_{[1]} \leq \theta_{[2]} \leq \cdots \leq \theta_{[k]}$. The population associated with θ_k (or $\theta_{[1]}$) is called the best. In the case of a tie, we assume that one of the contenders is tagged as the best. The objective is to select a subset of the given populations and claim that the probability that the best population is included in the selected subset is at least P* (1/k < P* < 1) regardless of the true configuration of the parameters. Let T_i be a suitable statistic based on n independent observations from $\pi_i (i = 1,2,...,k)$. The selection rule proposed in most of the specific situations for selecting a subset containing the population associated with $\theta_{[k]}$ $(\theta_{[1]})$ is of the form:

R: Select π_i iff $T_i \geq T_{max} - d_1$ $(T_i \leq T_{min} + d_2)$ or

R: Select π_i iff $c_1 T_i \geq T_{max}$ $(T_i \leq c_2 T_{min})$

where $T_{max} = \max(T_1,\ldots,T_k)$, $T_{min} = \min(T_1,\ldots,T_k)$ and the constants $c_1, c_2 \geq 1$ and $d_1, d_2 \geq 0$ are to be determined so that the basic probability requirement is satisfied. The above rules are particular cases of a general class of rules discussed by Gupta and Panchapakesan (1970). Usually, T_i is a sufficient statistic for θ_i and preserves the stochastic ordering. In all these cases, the standard technique is to obtain the expression for $P\{CS|R\}$, the probability of a correct selection using the rule R, evaluate its infimum over all parametric configurations and determine the constant of the procedure by equating this infimum to P*. Because of the stochastic ordering of the T_i, the infimum of $P\{CS|R\}$ is to be found over only the equal parameter configuration, namely, $\theta_1 = \theta_2 = \ldots = \theta_k = \theta$ (say). Exceptional situations arise in certain procedures for multinomial cells and in some rules using rank sums. In most of the problems we can establish the monotonic behavior of $P\{CS|R\}$ in θ by verifying certain sufficient condition [see Gupta and Panchapakesan (1970)]. Thus we obtain the infimum of $P\{CS|R\}$ and depending upon the type of the procedure used, we get one of the following relations:

$$(8.2) \quad \begin{cases} P\{T_k \geq T_{max} - d_1\} = P^* , \\ P\{T_1 \leq T_{min} + d_2\} = P^* , \\ P\{c_1 T_k \geq T_{max}\} = P^* , \\ P\{T_1 \leq c_2\, T_{min}\} = P^* , \end{cases}$$

where $T_1,\ldots,T_k$ are independent and identically distributed random variables with a common distribution G(say). Define

$$(8.3) \quad \begin{cases} U_1 = \max\{T_1-T_k,\ T_2-T_k,\ldots,T_{k-1}-T_k\} , \\ U_2 = \min\{T_2-T_1,\ T_3-T_1,\ldots,T_k-T_1\} , \\ V_1 = \max\{T_1/T_k,\ T_2/T_k,\ldots,T_{k-1}/T_k\} , \\ V_2 = \min\{T_2/T_1,\ T_3/T_1,\ldots,T_k/T_1\} . \end{cases}$$

It should be pointed out that V_1 and V_2 arise when the T_i are non-negative random variables. We see that the constants in (8.2) are either appropriate percentage points of the distribution of the random variables in (8.3) or related to these percentage points. The constants are given by the appropriate equation of the following set:

$$(8.4)\qquad \begin{cases} A(k, d_1) = \int_{-\infty}^{\infty} G^{k-1}(x+d_1)\, dG(x) = P^* , \\ B(k, d_2) = \int_{-\infty}^{\infty} [1 - G(x-d_2)]^{k-1}\, dG(x) = P^* , \\ I(k, c_1) = \int_{0}^{\infty} G^{k-1}(c_1 x)\, dG(x) = P^* , \\ J(k, c_2) = \int_{0}^{\infty} [1 - G(x/c_2)]^{k-1}\, dG(x) = P^* . \end{cases}$$

In each problem we know the specific form of G. Tables of constants are available in the literature in several special cases of G for selected values of k and P^*.

Gupta (1963) discusses among other things, the integral

$$(8.5)\qquad F_N(H;\rho) = \int_{-\infty}^{\infty} F^N(x\rho^{1/2} + H)\ (1-\rho)^{-1/2})\ dF(x) ,$$

where F denotes the cdf of a standard normal variable. It can be seen that $F_N(H;\rho)$ is the probability that the maximum of a set of N equally correlated standardized normal random variables does not exceed H or the probability that the minimum of this set exceeds $-H$.

Consider, for example, the rule

R: Select π_i iff $T_i \geq T_{max} - d_1$,

in the case of k normal populations with unknown means $\theta_1,\ldots,\theta_k$ and a common known variance σ^2, where the T_i are the sample means. One can see clearly the possibilities of using procedures which involve more order statistics than just T_{max}. If $T_{[1]} \leq T_{[2]} \leq \cdots \leq T_{[k]}$ are the ordered means, then Seal (1955) considered a class of procedures $D_{\underline{c}}$ defined below.

$D_{\underline{c}}$: Select the population corresponding to $T_{[i]}$ if and only if

$$T_{[i]} \geq c_1 T_{[1]} + \cdots + c_{i-1} T_{[i-1]} + c_i T_{[i+1]} + \cdots + c_{k-1} T_{[k]} - \sigma t(P^*,\underline{c})/\sqrt{n} \tag{8.6}$$

where $\underline{c} = (c_1,\ldots,c_{k-1})$ is a vector whose components are arbitrary real numbers such that $c_i \geq 0$ and $\sum_{i=1}^{k-1} c_i = 1$ and, $t(P^*,\underline{c})$ is chosen so as to satisfy the basic probability requirement. It is possible to propose procedures involving other functions of all the order statistics $T_{[1]},\ldots,T_{[k]}$, but these essentially present difficult distribution problems in terms of evaluating the infimum of $P\{CS|R\}$ and explains to an extent the absence of complete investigations of such procedures in the literature.

The authors wish to thank Mr. T. Santner for a critical reading of the paper and some helpful comments.

REFERENCES

Andersen , E.S. (1953): On sums of symmetrically dependent random variables. Skand. Aktuarietidskr. 36, 123-138.

Andersen , E.S. (1954): On the fluctuations of sums of random variables, II. Math. Scand. 2, 195-223.

Anscombe, F.J. (1952): Large sample theory of sequential estimation. Proc. Cambridge Philos. Soc. 48, 600-607.

Bahadur, R.R. (1966): A note on quantiles in large samples. Ann. Math. Statist. 37, 577-580.

Barlow, R.E. and Gupta, S.S. (1969): Selection procedures for restricted families of probability distributions. Ann Math. Statist. 40, 905-917.

Barlow, R.E. and Proschan, F. (1966): Inequalities for linear combinations of order statistics from restricted families. Ann. Math. Statist. 37, 1574-1592.

Barndorff-Neilson, O. (1964): On the limit distribution of the maximum of a random number of independent random variables. Acta Math. Acad. Sci. Hungar. 15, 399-403.

Barnett, V.D. (1966): Order statistics estimators of the location of the Cauchy distribution. J. Amer. Statist. Assoc. 61, 1205-1218. Correction. 63, 383-38!

Baxter, G. (1962): On a generalization of the finite arcsine law. Ann. Math. Statist. 33, 909-915.

Bechhofer, R.E. (1954): A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist. 25, 16-39.

Bechhofer, R.E., Kiefer, J. and Sobel, M. (1968): Sequential Identification and Ranking Procedures, with Special Reference to Koopman-Darmois Populations. Statistical Research Monographs, Vol. 3, The University of Chicago Press, Chicago.

Bennett, C.A. (1952): Asymptotic properties of ideal linear estimators. Ph.D. Thesis, University of Michigan.

Berman, S.M. (1962): Limiting distribution of the maximum term in sequences of dependent random variables. Ann. Math. Statist. 33, 894-908.

Berman, S.M. (1964): Limiting distribution of the maximum of a diffusion process. Ann. Math. Statist. 35, 319-329.

Bickel, P. J. (1967): Some contributions to the theory of order statistics. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (L. LeCam and J. Neyman, Ed.), Vol. I, pp. 575-591. University of California Press, Berkeley, California.

Blom, G. (1956): On linear estimates with nearly minimum variance. Ark. Mat. 3,365-36

Blom, G. (1958): Statistical Estimates and Transformed Beta-Variables. Almqvist and Wiksell, Uppsala, Sweden; Wiley, New York.

Gupta, S.S. and Panchapakesan, S. (1967): Order statistics arising from independent binomial populations. Mimeo. Ser. No. 120, Department of Statistics, Purdue University, Lafayette, Ind.

Gupta, S.S. and Panchapakesan, S. (1969): Selection and ranking procedures. In The Design of Computer Simulation Experiments (T. H. Naylor, Ed.), pp. 132-160. Duke University Press, Durham, N.C.

Gupta, S.S. and Panchapakesan, S. (1970): On a class of subset selection procedures. Mimeo. Ser. No. 225, Department of Statistics, Purdue University, Lafayette, Ind.

Gupta, S.S. and Panchapakesan, S. (1971): Contributions to multiple decision (subset selection) rules, multivariate distribution theory and order statistics. Mimeo. Ser. No. 260, Department of Statistics, Purdue University, Lafayette, Ind.

Gupta, S.S., Qureishi, A.S. and Shah, B.K. (1967): Best linear unbiased estimators of the parameters of the logistic distribution using order statistics. Technometrics 9, 43-56.

Gupta, S.S. and Shah, B.K. (1965): Exact moments and percentage points of the order statistics and the distribution of the range from the logistic distribution. Ann. Math. Statist. 36, 907-920.

Harter, H.L. (1961a): Expected values of normal order statistics. Biometrika 48, 151-165. Correction. 48, 476.

Harter, H.L. (1961b): Estimating the parameters of negative exponential populations from one or two order statistics. Ann. Math. Statist. 32, 1078-1090.

Harter, H.L. and Moore, A.H. (1966): Local-maximum likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples. J. Amer. Statist. Assoc. 61, 842-855.

Harter, H.L. and Moore, A.H. (1967): Asymptotic variances and covariances of maximum likelihood estimators, from censored samples, of the parameters of Weibull and gamma populations. Ann. Math. Statist. 38, 557-570.

Hastings, C., Jr., Mosteller, F., Tukey, J.W. and Winsor, C.P. (1947): Low moments for small samples: a comparative study of order statistics. Ann. Math. Statist. 18, 413-426.

Higuchi, I. (1954): On the solutions of certain simultaneous equations in the theory of systematic statistics. Ann. Inst. Statist. Math. 5, 77-90.

Jung, J. (1955): On linear estimates defined by a continuous weight function. Ark. Mat. 3, 199-209.

Jung, J. (1962): Approximation of least-squares estimates of location and scale parameters. In Contributions to Order Statistics (A.E. Sarhan and B.G. Greenberg, Ed.), pp. 28-33. Wiley, New York.

Kiefer, J. (1967): On Bahadur's representation of sample quantiles. Ann. Math. Statist. 38, 1323-1342.

Kiefer, J. (1970): (a) Deviations between the sample quantile process and the sample df. (b) Old and new methods for studying order statistics and sample quantiles. In Nonparametric Techniques in Statistical Inference (M.L. Puri, Ed.), pp. 299-319 and 349-357. Cambridge University Press, London.

Kimball, B.F. (1950): On the asymptotic distribution of the sum of powers of unit frequency differences. Ann. Math. Statist. 21, 263-271.

Kolmogorov, A.N. (1933): Sulla determinazionne empirica di une legge di distribuzionne. Giorn. Ist. Ital. Attuari. 4, 83-91.

Kulldorff, G. (1963): Estimation of one or two parameters of the exponential distribution on the basis of suitably chosen order statistics. Ann. Math. Statist. 34, 1419-1431.

Kulldorff, G. (1971a): A note on the optimum spacing of sample quantiles from the six extreme value distributions. Tech. Report No. 26, Texas A & M University, College Station, Texas.

Kulldorff, G. (1971b): Private communication.

Lamperti, J. (1964): On extreme order statistics. Ann. Math. Statist. 35, 1726-1737.

LeCam, L. (1958): Une théoreme sur la division d'une intervalle par des points près au hasard. Publ. Inst. Statist. Univ. Paris 7, 7-16.

Lloyd, E.H. (1952): Least-squares estimation of location and scale parameters using order statistics. Biometrika 39, 88-95.

Loéve, M. (1960): Probability Theory (2nd edition), Van Nostrand, New York.

Mendenhall, W. (1958): A bibliography on life testing and related topics. Biometrika 45, 521-543.

Miller, R.G., Jr. (1966): Simultaneous Statistical Inference. McGraw-Hill, New York.

Mogyoródi, J. (1962): A central limit theorem for the sum of a random number of independent random variables. Magyar Tud. Akad. Mat. Kutató Int. Közl. 7, 409-424.

Moran, P.A.P. (1947): The random division of an interval. J. Roy. Statist. Soc. Ser. B 9, 92-98.

Mosteller, F. (1946): On some useful "inefficient" statistics. Ann. Math. Statist. 17, 377-408.

Nef, W. (1964): Über die Differenz zwischen theoretischer and empirischer Verteilungsfunktion. Z. Wahrsheinlichkeitstheorie und Verw. Gebeite 3, 154-162

Ogawa, J. (1951): Contributions to the theory of systematic statistics, I. Osaka J. Math. 3, 175-213.

Ogawa, J. (1962): (a) Determinations of optimum spacings in the case of normal distribution, (b) Tests of significance using sample quantiles, (c) Optimum spacing and grouping for the exponential distribution, (d) Tests of significance and confidence intervals. In Contributions to Order Statistics (A.E. Sarhan and B.G. Greenberg, Ed.), pp. 272-283, 291-299, 371-380 and 380-382. Wiley, New York.

Pearson, K. (1902): Note on Francis Galton's problem. Biometrika 1, 390-399.

Plackett, R.L. (1958): Linear estimation from censored data. Ann. Math. Statist. 29, 131-142.

Proschan, F. and Pyke, R. (1967): Tests for monotone failure rate. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (L. LeCam and J. Neyman, Ed.) pp. 293-312. University of California Press, Berkeley.

Pyke, R. (1965): Spacings (with discussion). J. Roy. Statist. Soc. Ser. B 27, 395-449.

Rényi, A. (1953): On the theory of order statistics. Acta Math. Acad. Sci. Hungar. 4, 191-231.

Rényi, A. (1960): On the central limit theorem for the sum of a random number of random variables. Acta Math. Acad. Sci. Hungar. 11, 97-102.

Rényi, A. (1962): Théorie des éléments sailants d'une suite d'observations. In Combinatorial Methods in Probability Theory, pp. 104-115. Mathematical Institute, Aarhus University, Denmark.

Resnick, S.I. and Neuts, M.F. (1970): Limit laws for maxima of a sequence of random variables defined on a Markov chain. Adv. Appl. Probability 2, 323-343.

Sarhan, A.E. and Greenberg, B.G. (1956): Estimation of location and scale parameters by order statistics from singly and doubly censored samples. Part I: The normal distribution up to samples of size 10. Ann. Math. Statist. 27, 427-451. Correction. 40, 325.

Sarhan, A.E. and Greenberg, B.G. (Ed.) (1962): Contributions to Order Statistics. Wiley, New York.

Sarhan, A.E., Greenberg, B.G. and Ogawa, J. (1963): Simplified estimates for the exponential distribution. Ann. Math. Statist. 34, 102-116.

Seal, K.C. (1955): On a class of decision procedures for ranking means of normal populations. Ann. Math. Statist. 26, 387-398.

Sen, P.K. (1959): On the moments of the sample quantiles. Calcutta Statist. Assoc. Bull. 9, 1-20.

Sethuraman, J. and Rao, J.S. (1970): Pitman efficiencies of tests based on spacings. In Nonparametric Techniques in Statistical Inference (M.L. Puri, Ed.), pp. 405-415. Cambridge University Press, London.

Shah, B.K. (1966a): A note on Craig's paper on the minimum of binomial variates. Biometrika 53, 614-615.

Shah, B.K. (1966b): On the bivariate moments of order statistics from a logistic distribution. Ann. Math. Statist. 37, 1002-1010.

Sherman, B. (1950): A random variable related to the spacing of sample values. Ann. Math. Statist. 21, 339-361.

Siddiqui, M.M. (1963): Optimum estimators of the parameters of negative exponential distributions from one or two order statistics. Ann. Math. Statist. 34, 117-121.

Smirnov, N.V. (1949): Limit distributions for the terms of a variational series. Translation (1952): Amer. Math. Soc. Transl. Ser. 1, No. 67.

Spitzer, F. (1956): A combinatorial lemma and its application to probability theory. Trans. Amer. Math. Soc. 82, 323-339.

Sukhatme, P.V. (1937): Tests of significance for samples of the χ^2 population with two degrees of freedom. Ann. Eugen. 8, 52-56.

Takács L. (1961): The probability law of the busy period for two types of queueing process. Operations Res. 9, 402-407.

Takács L. (1962): A generalization of the ballot problem and its application in the theory of queues. J. Amer. Statist. Assoc. 57, 327-337.

Takács L. (1967): Combinatorial Methods in the Theory of Stochastic Processes. Wiley, New York.

Takács L. (1970): Combinatorial methods in order statistics. In Nonparametric Techniques in Statistical Inference (M.L. Puri, Ed.), pp. 359-384. Cambridge University Press, London.

Teichroew, D. (1956): Tables of expected values of order statistics and products of order statistics for samples of size twenty and less from the normal distribution. Ann. Math. Statist. 27, 410-426.

Tippett, L.H.C. (1925): On the extreme individuals and the range of samples from a normal population. Biometrika 17, 264-387.