ABSTRACTS OF PAPERS

(Abstracts of papers to be presented at the Central Regional Meeting, Lafayette, Indiana, March 23–25, 1966. Additional abstracts will appear in the April 1966 issue.)

1. Some asymptotically efficient sequential procedures for ranking and slippage problems. M. S. SRIVASTAVA, Princeton University.

In this paper Chow and Robbins' [*Ann. Math. Statist.* **35** (1964) 457–462] sequential theory has been applied to slippage and ranking problems: (1) A class $\mathcal{C}$ of sequential procedures is given for selecting the population with the largest mean from k populations having fixed distribution functions F_i with unknown but finite variance $\sigma^2 > 0$, so that in each case the probability of making the correct decision exceeds a specified value (say, $1 - \alpha$) when the greatest mean exceeds all the other means by at least a specified amount $2d$ and when $d \rightarrow 0$. The members of the class $\mathcal{C}$ of sequential procedures are asymptotically 'efficient' (in the sense of Chow and Robbins). Other ranking problems can be considered similarly. (2) Slippage problem. In the above problem, the possibility that all parameters (means) of the various populations are equal was ruled out. In this paper, we consider this possibility also and a class $\mathcal{C}$ of efficient sequential procedures is given. (3) A Multivariate Slippage problem. The problem 2 has been extended to multivariate case also.

2. On the asymptotic theory of sequential confidence intervals. M. S. SRIVASTAVA, Princeton University. (By title)

Chow and Robbins [*Ann. Math. Statist.* **35** (1964), 457–462] have considered the problem of finding a confidence interval of prescribed width $2d$ and prescribed coverage probability α for the unknown μ of a univariate population Ω having fixed distribution function F with unknown, but finite, variance $\sigma^2 > 0$. Since no fixed sample procedure can possibly work, they consider a certain class of sequential procedures and show that the members of this class are asymptotically consistent and efficient. It is the object of this paper to investigate the extent to which these results may be extended to find the confidence intervals for the following: (1) Mean of a multivariate population. (2) All normalized linear functions of the p components of the mean vector. (3) Linear regression parameters. (4) Means of K independent univariate populations. (6) All differences between means of k independent populations. (7) For the difference between the means of two univariate populations with unequal unknown variances, Fisher-Behren's problem. (8) All normalized liner functions of means of k independent univariate populations.

(Abstract of a paper to be presented at the Eastern Regional Meeting, Upton, Long Island, New York, April 27–29, 1966. Additional abstracts will appear in the April 1966 issue.)

1. Minimal two-level main-effect-clear plans (preliminary report). CUTHBERT DANIEL, Private Consultant, New York.

Intermediate in resolving power between main-effect plans (which alias main-effect estimates with two-factor interactions) and two-factor-interaction-clear plans (which alias each main-effect and two-factor interaction only with higher order interactions) are main-effect-clear plans (which alias main-effects only with three-factor interactions and two-factor interactions with each other). An estimate of the main-effect of factor A, unbiased by any two-factor interactions with factors $B, C, D, \cdots, Q$, can be obtained from the results of *four* runs, made under conditions: (1), a, $bcd \cdots q$, and $abcd \cdots q$. Similarly unbiased estimates of the main effects of three factors, A, B, C, require *six* runs: (1), a, b, ac, bc, and abc. If five factors are under variation the *ten* runs required to permit main-effect-clear estimation are: (1), a, b, ac, bd, ace, bde, $acde$, $bcde$, and $abcde$. Extensions, blocking, estimation of two-factor interaction sums, and isolation of single interactions are discussed.

(Abstracts not connected with any meeting of the Institute.)

1. Waiting times in tandem queues. P. J. Burke, Bell Telephone Laboratories.

The result of Edgar Reich [*Ann. Math. Statist.* **34** (1963) 338–341] that the waiting times of a customer in a sequence of J stationary $M/M/1$ queuing systems in tandem, with order of arrival service, are mutually independent random variables is extended to $M/M/N_j$ systems, $j = 1, \cdots, J$, with $N_j = 1$ for $j = 2, \cdots, J - 1$. The conditional distribution of the waiting time in an $M/M/N$ system given the state at the departure instant is first calculated and shown to be the same as that given the state at the arrival instant. This result is then used to show that the waiting time is independent of the output process previous to the departure instant of the customer in question. The waiting time in the next system is entirely determined by the previous output process of the first queue and by the independent sequence of service times. Hence the mutual independence of the waiting times follows as a corollary by the same argument used by Reich for the single server case.

2. The nearest neighbor decision rule. Thomas M. Cover and Peter Hart, Stanford Electronics Research Laboratories and Stanford University.

The nearest neighbor decision rule assigns to an unclassified point the classification of its nearest neighbor. In a large sample analysis it is shown that the risk of the nearest neighbor rule is bounded below by the Bayes risk and above by twice the Bayes risk for all suitably smooth underlying distributions and a wide class of loss functions. These loss functions include probability of error in the finite hypothesis problem, and expected distance and mean square error in the estimation problem. In particular, for the case of mean square error, the nearest neighbor risk is shown to be precisely equal to twice the Bayes risk. For many loss functions these bounds are the tightest possible. Roughly speaking it may be said that half of the classification information in an infinite sample set is contained in the nearest neighbor.

3. One-sample normal scores test distribution and power (preliminary report). R. Thompson, Z. Govindarajulu and K. Doksum, Massachusetts Institute of Technology, Case Institute of Technology and University of California, Berkeley.

Some percentage points of the normal scores test for symmetry are extended to sample size 20 and compared with normal and Edgeworth approximations. The powers of the normal scores (N), Wilcoxon (W), sign (S), and t (T) tests are estimated from monte carlo trials for shifts of symmetrical hypothesized populations which are normal, logistic, double exponential, and uniform. Perhaps the most striking feature of the power results is the relative goodness of the t-test, and perhaps also of the normal scores test. The t-test power seemed to come out significantly lower than the best of the four considered only for small double exponential shift, the normal scores test only for uniform shift. In most cases, the powers of T, N, and W were virtually indistinguishable, but there were some tendencies. From best to worst, with a comma marking a definite break, the tests seemed to stack somewhat as follows: for normal shift TNW, S; for logistic shift TWN, S; for small double exponential shift $SWNT$, for large $WTNS$; for uniform shift T, NW, S.

4. Asymptotic normality of linear functions of order statistics in one and multi-samples. ZAKKULA GOVINDARAJULU, Case Institute of Technology.

A class of linear functions of order statistics in a random sample drawn from a continuous population is defined and sufficient conditions for its asymptotic normality are given. This class, in particular, includes the systematic statistics. The case when the sample size is random is also considered. In a natural way, linear functions of order statistics in the pooled sample of several subsamples drawn from continuous, otherwise, arbitrary populations are also considered and their joint asymptotic normality studied.

5. Generalizations of Chernoff-Savage theorems on asymptotic normality of nonparametric statistics (preliminary report). ZAKKULA GOVINDARAJULU, L. LE CAM and M. RAGHAVACHARI, Case Institute of Technology, University of California, Berkeley, and University of California, Berkeley.

Chernoff and Savage (1958) (See *Ann. Math. Statist.* **29** 972–994) provided interesting results on asymptotic normality of a class of two-sample nonparametric test statistics, which covers more situations than earlier contributions to this problem. In this investigation, we have strengthened Chernoff-Savage theorems by relaxing some of the sufficient conditions required by Chernoff and Savage. Further, some new theorems are also obtained. The natural extensions of these results to the c-sample situations are also provided.

6. Selecting a subset containing the population with the largest α th-percentile. MILTON SOBEL and M. HASEEB RIZVI, University of Minnesota and Ohio State University.

A nonparametric solution is developed for the problem of selecting a subset *of* $k \geqq 2$ populations containing the one with the largest α-percentile. It is assumed that the k cdf's F_i $(i = 1, 2, \cdots, k)$ are continuous and that $1 \leqq (n + 1)\alpha \leqq n$ where n is the given common number of observations per population. Let $Y_{j,i}$ denote the jth order statistic from F_i; the proposed procedure R_1 includes F_i iff $Y_{r,i} \geqq \text{Max}_{1 \leqq j \leqq k} Y_{r-c,j}$ where r is an integer defined by $r \leqq (n + 1)\alpha < r + 1$; here c is the smallest integer $(1 \leqq c \leqq r)$ with the property that if the cdf $F_{[k]}(y)$ with the largest α-percentile is not greater than $F_{[j]}(y)$ for all y and each j then the probability of a correct selection $P\{CS \mid R_1\}$ under procedure R_1 will be at least P^*. Asymptotic expressions are developed for the infimum of the $P\{CS \mid R_1\}$ and the expected size of the selected subset $E\{S \mid R_1\}$ and the asymptotic relative efficiency of R_1 relative to other procedures is computed for various alternatives. The procedure R_1 is shown to be unbiased. The dual problem for the smallest α-percentile is briefly mentioned. A table of $r - c$ values for $k = 2$ and $\alpha = \frac{1}{2}$ is included for selected values of n and P^*.

7. An estimate of the mean of a finite population using several auxiliary variables. SURENDRA K. SRIVASTAVA, Lucknow University.

To estimate the population mean of a finite population of a character y, use is made of $(s + t)$ auxiliary characters, s of which, say $x_1, \cdots, x_s$ are positively correlated with y and the remaining t, say $z_1, \cdots, z_t$ are negatively correlated with y. The suggested estimate is $\tilde{y} = \sum_{i=1}^{s} u_i(\bar{X}_i\bar{y}/\bar{x}_i) + \sum_{j=1}^{t} v_j(\bar{y}\bar{z}_j/\bar{Z}_j)$ where $\sum_{i=1}^{s} u_i + \sum_{j=1}^{t} v_j = 1$. The weights u's and v's are to be determined so as to minimise the variance of y. Approximate formulae for bias and variance of the estimate have been obtained. The variance formula takes the form $w'\Sigma w$, where $w' = (u_1, \cdots, u_s, v_1 \cdots, v_s)$ is the weight vector and Σ is a partitioned matrix based on the population covariances and variances of the characters. The minimisation of this function has been dealt by Olkin (*Biometrika* **45** 154–165) and the minimum variance is found to be $1/e'\Sigma^{-1}e$ where e is a unit column vector of order $(s + t)$.