ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Eastern Regional meeting, Upton, Long Island, New York, April 27–29, 1966. Additional papers appeared in the April 1966 issue.)

7. Ratio and regression estimators as minimax procedures for estimating the mean of a population. Om Aggarwal, Iowa State University.

The author has considered earlier the cases of simple random sampling and stratified sampling [*Ann. Math. Statist.* **30** (1959) 206–218] as well as the case of two-stage sampling [*Ann. Math. Statist.* **36** (1965) 1596, abstract #10] from Bayes and minimax point of view by assuming that the loss in estimating the mean is a linear function of the squared error of the estimator and the cost of observations. By using similar approach in this paper, conditions have been derived under which the usual ratio method of estimation is a minimax procedure. It is further shown that if X and Y are distributed jointly, and only their variances, the covariance between X and Y, and the mean of Y are known, then a minimax estimator for the mean of X is the usual regression estimator. The results are extended to the case of sampling from finite populations in each case. (Received 7 March 1966.)

8. k-sample nonparametric life tests (preliminary report). A. P. Basu, University of Minnesota.

Let X_{ij} $(j = 1, 2, \cdots, n_i)$ be a sample of size n_i from the ith population with continuous cdf $F_i(x)$ $(i = 1, 2, \cdots, k; \sum_{i=1}^{k} n_i = N)$. Let the combined N observations be ordered and define $Z_\alpha^{(i)} = 1$, if the αth ordered observation is from the ith population and 0, otherwise. To test the hypothesis H: $F_1 = F_2 = \cdots = F_k$ against the alternative that they are all different we propose the statistic $B = \{12N^3(N-1)/r(r+1)[2N(2r+1) - 3r(r+1)]\} \cdot \sum_{i=1}^{k} n_i^{-1}(S_i + r(r+1)n_i/2N^2)^2$ based only on the first r ordered observations from the combined sample, where $S_i = \sum_{\alpha=1}^{r}[(\alpha - r - 1)/N]Z_\alpha^{(i)}$. For $r = N$, the above statistic reduces to the Kruskal statistic H (*Ann. Math. Statist.* **23** (1952) 525–540), and for $k = 2$ it is equivalent to the Sobel statistic $V_r^{m,n}$ (Technical Report No. 69, University of Minnesota, (1965)). The mean and variance of B under the null hypothesis have been derived. The asymptotic distribution of B in the non-null case is shown to be the noncentral χ^2-distribution with $(k-1)$ df. A second k-sample life test statistic, generalizing the Jonckheere statistic S (*Biometrika* **41** (1954) 133–145), has also been proposed. (Received 28 February 1966.)

9. Method of maximum empirical likelihood for the two-sample location parameter problem. P. K. Bhattacharya, University of Arizona.

$X_1, \cdots, X_m, Y_1, \cdots, Y_n$ are independent random variables, X's having common cdf F and Y's having common cdf G. $G(x) = F(x - \theta)$, but F and θ are unknown. Choose $a_1 < \cdots < a_k$. Let $\chi_1, \cdots, \chi_{k+1}$ be the indicator functions of $(-\infty, a_1]$, $(a_{j-1}, a_j]$, $j = 2, \cdots, k$, and (a_k, ∞) respectively, $\nu_{jm} = \sum_{i=1}^{m} \chi_j(X_i)$, $q_{jn}(t) = \sum_{i=1}^{n} \chi_j(Y_i - t)/n$, $j = 1, \cdots, k+1$, $Z_{m,n}(t) = \sum_{j=1}^{k+1} \nu_{jm} \log q_{jn}(t)$. Choose $d > 0$ and a real number t_0. Consider the inequality $Z_{m,n}[t_0 + rd/(m+n)^{\frac{1}{2}}] > \max\{Z_{m,n}[t_0 + (r-1)d/(m+n)^{\frac{1}{2}}], Z_{m,n}[t_0 + (r+1)d/(m+n)^{\frac{1}{2}}]\}$. If $N = m + n \to \infty$ such that $m/N \to u$ in $(0, 1)$, then under some conditions on F, (i) the probability that the above inequality is satisfied at more than one point of the form $t_0 + rd/N^{\frac{1}{2}}$ (r integer) in the interval $(\theta - A/N^{\frac{1}{2}}, \theta + A/N^{\frac{1}{2}})$ tends to 0 for arbitrary A, (ii) the above inequality has a sequence of solutions T_N such that $N^{\frac{1}{2}}(T_N - \theta)$

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

is asymptotically roughly normal with mean 0 and variance $1/u(1-u)I_k$ for small d, where I_k is the Fisher-information in each classified X. As the class-intervals are made finer and finer, I_k tends to the Fisher-information for the problem. A consistent estimate of I_k is given and large sample tests of hypotheses about θ are constructed. (Received 7 March 1966.)

10. On multivariate exponential-type distributions. SHEELA BILDIKAR and G. P. PATIL, McGill University and Pennsylvania State University.

Certain structural properties of the class of multivariate exponential-type distributions are studied in this paper. A random vector $x = (x_1, x_2, \cdots, x_s)$ is said to have s-variate exponential-type distribution with the parameter vector $\theta = (\theta_1, \theta_2, \cdots, \theta_s)$, if its probability function is given by $f(x, \theta) = a(x) \exp \{\sum_1^s \theta_i x_i - q(\theta)\}$ where $a(x)$ is a function of x and not of θ, and $q(\theta)$ is a bounded analytic function of θ, for θ belonging to an s-dimensional open interval. Distributions, like multivariate normal, multivariate power series, etc., belong to this class. A few properties, known to be "locally" true for the class, are shown to be "globally" true over the parameter space. The problem of characterization of certain members of the class is studied under suitable restrictions on the cumulants. The results obtained include: (i) The random variables $x_1, x_2, \cdots, x_s$, following jointly s-variate exponential-type distribution, are statistically independent if and only if they are pairwise uncorrelated. (ii) In the class of s-variate exponential-type distributions, s-variate normal is the only symmetric distribution. (iii) An s-variate exponential-type distribution is s-variate normal if and only if the regression of one of the variables on the rest is linear and the rest of the $(s-1)$ variables are distributed normally in pairs. (Received 15 March 1966.)

11. A dictionary of distributions. E. A. BLAKE and E. B. FOWLKES, Bell Telephone Laboratories, Inc., Murray Hill, New Jersey.

A study has been made of plots of distribution quantiles for a wide variety of probability distributions against quantiles of each of the normal, uniform, and exponential distributions. This provides a dictionary of quantile plots which shows the "closeness" of distribution families to each of the above three. Within a family, changes in the shape parameter denote different members of the family. Quantile versus quantile plots (Q-Q plots) have the property that they are invariant under linear transformation. The plots are also bordered by the probabilities corresponding to the quantile grid lines. Plots for individual members of a family and composite plots of the selected members taken together are given. A concise and instructive portrayal of the "closeness" of two distributions is presented by plots of this kind, and the relevant information can be assessed at a glance. Examination of the Q-Q plots might influence the statistician in his choice of assumptions for some theoretical or empirical investigation where such choice would lead to mathematical simplifications. The plots may also be of particular use in studies of robustness of statistical procedures. (Received 7 March 1966.)

12. A Bayesian study of the multinomial distribution. D. A. BLOCH and G. S. WATSON, Johns Hopkins University.

Lindley [*Ann. Math. Statist.* (1964)] studied the topic in our title. By using Fisher's conditional-Poisson approach to the multinomial and the logarithmic transformation of gamma variables to normality, he showed heuristically that linear contrasts in the logarithms of the cell probabilities θ_i are asymptotically jointly normal and suggested that the approximation can be improved by applying a "correction" to the sample. By studying the asymptotic series for the joint distribution, we have verified this assertion and found an improved correction procedure. A more detailed expansion is given for the distribution of a single contrast. In many problems linear functions of the θ_i are of interest. The exact distribution for these is obtained. This has a density of a form familiar in the theory of serial correlation coefficients. A beta approximation is given. For three cells, a numerical study showed the merit of this approximation. (Received 2 March 1966.)

13. Organization of machine independent statistical computer systems. ROALD BUHLER, Princeton University.

Large systems of statistical programs are practical if a computer has a large random access bulk storage device. One uses such a system through its control language. A user may also wish to add new programs to the system using the internal programming conventions of the system. These conventions exist as Fortran subroutines that do just one task per subroutine. Therefore any machine (or monitor) dependent task is isolated and can be changed without affecting the control language or the statistical programs. The P-STAT system (written for the I.B.M. 7094) is now being implemented on the C.D.C. 6600, and some of the problems encountered in changing it over will be discussed. (Received 2 March 1966.)

14. Asymptotic distributions for a generalized Banach matchbox problem. T. CACOULLOS, New York University.

Observations are taken one at a time from a multinomial distribution with k cells C_1, $C_2, \cdots, C_k$ with probabilities $p_1, p_2, \cdots, p_k$. Let E_i denote the event that N_i observations are taken from cell C_i before taking N_j observations from cell C_j, $j \neq i$. Suppose $N_i = N_i(N)$ and as N tends to infinity N_i/N tends to a positive finite limit λ_i, $i = 1, \cdots, k$. Set $\theta_i = p_i/\lambda i$ and let $\theta_{[1]} \leqq \theta_{[2]} \leqq \cdots \leqq \theta_{[k]}$ denote the ordered θ_i. Let $C_{(i)}$ denote the cell associated with $\theta_{[i]}$, and $p_{(i)}$ and $N_{(i)}$ the corresponding probability and N_i. THEOREM. *Let $E_{(i)}$ denote the event that $N_{(i)}$ observations are taken from cell $C_{(i)}$ before $N_{(j)}$ observations are taken from cell $C_{(j)}$, $j \neq i$. Let $X_{(j)}$ denote the number of observations drawn from cell $C_{(j)}$ when $E_{(i)}$ occurs ($j \neq i$). Suppose $\theta_{[s]} < \theta_{[s+1]} = \cdots = \theta_{[k]}$. Theo the conditional asymptotic distribution of $X_{(1)}, \cdots, X_{(s)}$ and the $R_{(t)} = N_{(t)} - X_{(t)}$, $t = s+1, \cdots, k$, $t \neq i$ (properly standardized) given $E_{(i)}$ is a restricted (truncated) normal distribution. For $s = k - 1$ the asymptotic distribution is unconditionally normal; for $s = 0$, the conditional asymptotic distribution of the $R_{(j)}$, $j \neq i$, given $E_{(i)}$ is a normal restricted to the positive orthant.* The asymptotic theory is used in setting up approximate tests and confidence limits for the largest p_i. (Received 2 March 1966.)

15. The lattice of ordered partitions and its relation to generalized symmetric means and generalized polykays. EDWARD J. CARNEY, Iowa State University.

An ordered partition of weight m is defined as follows: Let $a = (a_1 \cdots a_m)$ be an m-tuple; an ordered partition of a is a set of $m(m-1)/2$ consistent statements which specify for each pair a_i, a_j whether $a_i = a_j$ or $a_i \neq a_j$. An ordered partition is completely specified by any m-tuple, a, of symbols when one can determine for each pair a_i, a_j, whether $a_i = a_j$. Any particular realization of the statements specifying an ordered partition is called an ordered partition. If a and b are representations of some ordered partitions of weight m, a is said to equal b if and only if $a_i = a_j$ implies, and is implied by, $b_i = b_j$, for every pair $i, j = 1, 2, \cdots, m$. If a and b are ordered partitions of equal weight, a is an ordered subpartition of b, if and only if $b_i = b_j$ implies $a_i = a_j$ for all i, j. It can be shown that the system of ordered partitions of weight m forms a modular lattice with respect to the subpartition partial ordering. A given generalized symmetric mean of degree m for an n-factor crossed, balanced complete structure can be defined as the sum over all mth power products satisfying n given ordered partitions of weight m. The generalized polykays are in 1–1 correspondence with the generalized symmetric means and may be defined implicitly in terms of sums over certain sublattices of the lattice of ordered partitions. The polykays are obtained explicitly by inverting this relationship. The inverse has a simple form in terms of matrices which are nilpotent of index m. A dual of the relationship between polykays and symmetric means relates symmetric functions and certain unrestricted sums. This relationship can be exploited for computation of polykays. (Dayhoff, E. E. (1966). Generalized polykays, an extension of simple polykays and bipolykays, *Ann. Math. Statist.* **37**.) (Received 7 March 1966.)

16. On multivariate Edgeworth expansions. JOHN M. CHAMBERS, Harvard University and Bell Telephone Laboratories.

The Edgeworth expansion for the probability density of a multivariate random variable du is discussed. The expansion is justified when du behaves as a standardized mean of a sample of size n, in the sense that the cumulants which exist are of the correct order of magnitude in n. The conditions for validity are that sufficient cumulants exist and that the inversion of the characteristic function of du outside regions of width Kn^{α}, $0 < \alpha < \frac{1}{2}$, be asymptotically negligible. Conditions for constructing valid expansions for the probability density of functions of du are given. The Wishart and non-central Wishart distributions are considered as examples. (Received 7 March 1966.)

17. Tolerance setting in multivariate surveys. SAMPRIT CHATTERJEE, Harvard University.

A solution of the allocation problem in multivariate surveys is obtained by determining the allocation such that the estimates have stated levels of precision at minimum cost. This involves the solution of a non-linear programming problem which can be solved conveniently by an algorithm developed by the author. In this paper the question of tolerance setting is taken up. A discussion is given of this problem from the decision-theoretic and Bayesian point of view. The limitations of these approaches are examined in the context of sample surveys. Two empirical procedures which are more practical are then developed. These procedures enable the sampler to study the cost implications of the various tolerances and provide him with a systematic method for altering them. The first method involves the calculation of several plans with altered tolerances, the altering being done in a specified manner. The second involves determining the changes in cost for differential changes in the tolerances. A method is then presented for calculating the sampling cost with altered tolerances using these differential changes. (Received 7 March 1966.)

18. On the queue length distribution with balking in a single server queueing process. RONALD S. DICK, G. W. Post College and Columbia University.

The writer develops a model where after the waiting line exceeds a given length m, customers refuse to enter with probability $q = 1 - p$. Theorems on the queue after a departure are proven, and Theorem 18 of Takács (1) is corrected. The unrestricted steady state queue length distribution for the model is found by methods of Keilson and Kooharian (2). Future work and applications of the model are indicated. (1) L. Takács, The Transient

Behavior of a Single Server Queueing Process with a Poisson Input, *Fourth Berkeley Symp. Math. Stat. Prob.* **2** 535–567. University of California Press, (1961). (2) J. Keilson and A. Kooharian, On Time Dependent Queuing Processes, *Ann. Math. Statist.* **31** (1960) 104–112. (Received 4 March 1966.)

19. On a multivariate generalization of the Behrens-Fisher distributions. J. M. Dickey, Yale University.

We consider the random p-dimensional vector, $\boldsymbol{\delta} = \sum \mathbf{B}_k \mathbf{t}_k$, a linear combination of K independent standard p-dimensional multivariate-t vectors $\mathbf{t}_k$ with degrees-of-freedom ν_k (each $\mathbf{t}_k \sim \mathbf{z}_k/[(\nu_k^{-1}u_k)]^{\frac{1}{2}}$ where $\mathbf{z}_k \sim N(\mathbf{0}, \mathbf{I})$ independently of $u_k \sim \chi^2(\nu_k)$). Any linear combination of the mean vectors $\boldsymbol{\mu}_k$ (or regression vectors $\boldsymbol{\beta}_k$, $\boldsymbol{\mu}_k = \mathbf{X}_k\boldsymbol{\beta}_k$) of K unknown multivariate-normal populations is seen to be distributed *a posteriori* like $\boldsymbol{\delta}$, given K independent "ignorance" or conjugate prior distributions for the parameters of the populations. We express the $p(K - 1)$-dimensional integral which is the density function of $\boldsymbol{\delta}$ as a $(K - 1)$-dimensional integral, thus generalizing Ruben's integral forms for the single-variable Behrens-Fisher densities [*J. Roy. Statist. Soc B* **22** (1960)]. Equivalently, $\boldsymbol{\delta} \sim \mathbf{Bt}$, where $\mathbf{t}$ is a standard multivariate-t vector with degrees-of-freedom $\sum \nu_k$, and $\mathbf{B}$ is any square root of $\mathbf{BB}' = (\sum \nu_k)^{-1} \sum \nu_k w_k^{-1} \mathbf{B}_k \mathbf{B}_k'$, with $(w_1, \cdots, w_{K-1})$ having the Dirichlet distribution $D(\frac{1}{2}\nu_1, \cdots, \frac{1}{2}\nu_{K-1}; \frac{1}{2}\nu_K)$ and $\sum w_k = 1((w_1, \cdots, w_K) \sim (u_1/\sum u_k, \cdots, u_K/\sum u_k))$. (Received 15 February 1966.)

20. A stratification approach for estimating nonresponses without recall. S. P. Ghosh, IBM-Research Center, New York.

A method for estimating the variable under analysis for the nonresponded sampling unit has been developed. An auxiliary variable is used to stratify the population into heterogenous groups before the sampling units are drawn. After the sampling units have been surveyed the responded units of a particular group are again classified into subgroups on the basis of the variable under analysis. The subgroup structure is then used to estimate the value of the nonresponded units. It has been shown that under certain conditions the estimates are consistent. It has also been shown that by estimating the nonresponses by the proposed scheme a reduction in variance is also achieved. (Received 1 March 1966.)

21. A Monte Carlo study of the estimators of the parameters of a quantal response curve. M. Gnanadesikan, Bell Telephone Laboratories.

A Monte Carlo study was carried out to study the estimators of the parameters of a response curve with the probability of response p_i to a stimulus level x_i defined by $p_i = \int_{-\infty}^{\alpha+\beta x_i} (2\pi)^{-\frac{1}{2}} e^{-t^2/2}\, dt$. Four sets of parameter values of (α, β) viz. (0, .2), (0, .3), (1, .3), (0, .4) were studied in conjunction with four sets of ten levels $\{x_i\}$ of the stimulus, viz., $-2(1)7$, $-4(1)5$, $-5(1)4$, $-8(1)1$. For each (α, β, x_i) fifty samples of forty observations were generated. For each of the fifty samples, the forty observations at each of the 10 stimulus levels were subdivided into 8, 4 and 2 subsamples of sizes 5, 10 and 20 respectively. The maximum likelihood and minimum normit χ^2 estimators were computed from the three kinds of subsamples and the complete sample. The four subsampling schemes and the two methods of estimation were compared with respect to the bias, error variance and mean square error of the estimators. None of the sampling schemes or the methods of estimation were found to be uniformly better, the relative merits depending on the true values of (α, β), the number of levels and the nature of the levels. (Received 7 March 1966.)

22. On characterizing dependence in joint distributions. W. J. HALL, University of North Carolina.

Ways of characterizing the dependence of one random variable on another (or several others) are investigated. In particular, an *index of dependence* of X on Y is introduced which (i) always exists; (ii) lies in [0, 1]; (iii) is zero if and only if X and Y are independent; (iv) is unity if X is a function of Y (and only if whenever X has finite variance); (v) may assume every value in [0, 1] by varying the joint distribution but holding the marginal distributions fixed (assuming Y continuous); (vi) is invariant under linear transformation of X and one-to-one transformation of Y; and (vii) equals k/m whenever X and Y are sums of IID random variables $Z_1, Z_2, \cdots, X$ being the sum of the first m Z's and Y the sum of the first k Z's $(m \geqq k)$. When the correlation ratio exists, its square cannot exceed the dependence index, and when (X, Y) is bivariate normal the index equals ρ^2. The index is derived by first introducing and investigating a *dependence characteristic*, defined as the squared correlation ratio of exp (itX) on Y, as a function of t. A brief survey of correlation and regression theory for complex-valued random variables is included. (Received 7 March 1966.)

23. Multi-stage sampling on successive occasions where first-stage units are drawn with unequal probabilities and with replacement. JOHN C. KOOP, University of North Carolina. (By title)

In the current practice of sampling on successive occasions, partial replacement of units is made only at the last stage. In this paper, in sampling for p occasions with a constant first-stage sample size of n, and with a replacement rate of μn first-stage units per occasion $(0 < \mu < 1)$, $n + (p - 1)\mu n$ first-stage units are drawn from a universe of N units with replacement and with probabilities $p_i > 0$, $(\sum_1^N p_i = 1)$, observing the order of appearance of the units. On every occasion after the first, the first μn units are rejected and are replaced by the set of μn units following the current sample. Within each first-stage unit at the second and succeeding stages, predetermined numbers of units are selected. Linear and ratio estimators are used to estimate the universe total, utilizing information from the matched and unmatched parts for the present and the past occasion. The estimator with the best linear combination of the sample data from each first-stage unit, is found to be of the form given by Patterson in his 1950 paper with a similar variance function. With this approach a more substantial reduction in variance is possible than in the case when partial replacement is made at the last stage. (Received 10 March 1966.)

24. On sampling a universe by overlapping clusters of units with unequal probabilities. JOHN C. KOOP, University of North Carolina. (By title)

From a universe U of N distinct elements or units $u_1, u_2, \cdots, u_N$ overlapping clusters or sets s, each containing varying numbers of elements $n(s)$ are formed, such that there is at least one set which contains a given pair of elements and no set contains elements appearing more than once. The cardinality of $\mathcal{S} = \{s\}$ is known, and the probability of selecting s, $p(s) > 0$, is given, and is such that $\sum_{s \varepsilon \mathcal{S}} p(s) = 1$. A value x_i is defined on u_i for all $u_i \varepsilon U$, and it is desired to estimate the universe total $T = \sum_1^N x_i$ on the basis of a single sample (set) s, with x_i-values for all $i \varepsilon s$, selected from $\mathcal{S}$ with probability $p(s)$. In this approach only *two features of sample formation* viz. (a) the presence or absence of an element in a given s, (b) the set of elements composing s considered as one of the possible samples, are recognizable in contrast to the approach where an additional feature, i.e. order of appearance, is recognizable when the sample is formed by drawing elements one at a time [Koop (1963) *Metrika* **7** 87–88]. On the basis of (a) and (b), it is demonstrated that only three $(=2^2-1)$ classes of linear estimators are possible for estimating the universe total. The coefficients for these estimators are determined from the condition $E(\text{estimator}) = T$ for all $\{x\}$. The practical implications of this approach, in the multi-stage sampling of preferred clusters of units, are explored. (Received 10 March 1966.)

25. Estimation of the parameters of the Pareto distribution. HENRICK JOHN MALIK, Western Reserve University. (By title)

In this paper, sufficient statistics for the parameters a and v are obtained. It is shown that $Y_1 = \min(X_1, \cdots, X_N)$ is sufficient for a when v is known; the sample geometric mean g is sufficient for v when a is known, and $(Y_1, \sum_{i=1}^N \log(Y_i/Y_1))$ is a joint set of sufficient statistic for (a, v), when both are unknown. Using sufficiency, it is shown that the statistic $Z = \log(Y_1 \cdots Y_N/Y_1)$ is stochastically independent of the sufficient statistic Y_1. Using sufficiency and stochastic independence of Z and Y_1, the exact distribution of the maximum likelihood estimator $\hat{v}$ is derived. Bayesian estimation of the parameter v is also considered. (Received 28 February 1966.)

26. Distribution of product statistics from a Pareto population. HENRICK JOHN MALIK, Western Reserve University. (By title)

Distributions are derived of the product of sample values, the sample geometric mean, the product of two minimum values from sample of unequal size and product of k minimum values from samples of equal size from a Pareto population. The distributions can be conveniently transformed to χ^2. (Received 28 February 1966.)

27. Exact moments of the order statistics of the geometric distribution and their application to a stage-dependent binomial sampling scheme. BARRY H. MARGOLIN and HERBERT S. WINOKUR, JR., Harvard University.

A method for calculating the exact moments of the order statistics from a geometric distribution is presented. A stagewise sampling scheme is investigated in which at each stage a random sample is taken from a binomial $B(N, p)$ distribution whose parameter N is dependent on the sample observed at the previous stage. The distribution of the number of stages until the scheme terminates corresponds to the distribution of a specific order statistic from a geometric distribution. Hence, the exact moments of the distribution of the number of stages until termination can be found by the procedure described earlier. Brief tables and specific applications of the binomial sampling scheme are included. (Received 7 March 1966.)

28. Some Tchebycheff typed inequalities for matrix valued random variables. GOVIND S. MUDHOLKAR, University of Rochester.

Let $x_1, x_2, \cdots, x_n$ be n jointly distributed random p-vectors with $Ex_i = 0$ and $Ex_ix_i' = \Sigma_i$, $i = 1, 2, \cdots, n$. Write $X(p \times n) = (x_1x_2, \cdots, x_n)$ and $\Sigma = \Sigma_1 + \Sigma_2 + \Sigma_n$. Let $c_1, c_2, \cdots, c_p$ denote the p characteristic roots of $XX'(p \times p)$ and $\sigma_1, \sigma_2, \cdots, \sigma_p$ those of Σ. Let $f(c_1, c_2, \cdots, c_p) \geqq 0$ be a concave, symmetric function of (c_1, c_2, c_p). Then it has been proved in this paper that $\Pr[f(c_1, c_2, \cdots, c_p) \geqq K] \leqq K^{-1}f(\sigma_1, \sigma_2, \cdots, \sigma_p)$. As applications, inequalities involving $Ch_{\min}(XX')$, def (XX'), $\sum c_i^r$, $r \leqq 1$, elementary symmetric functions of c's, etc., have been obtained. (Received 7 March 1966.)

29. Detecting outliers in a two-way table. JANE F. MUNK and M. B. WILK, Bell Telephone Laboratories.

Some techniques are presented for the detection of outliers in a two-way table, where the basic assumption is that of additivity. While the usual residuals are an adequate basis for detection of a single outlier, two or more outliers can interact so as to make the usual residuals quite bland. Procedures for probability plotting of certain revised residuals and for empirical cumulative distribution plotting of sums of squares are suggested, as a basis for display and judgment, for the two-outlier case. (Received 7 March 1966.)

30. A multivariate sign test. KANTILAL M. PATEL, Clemson University.

Consider N blocks of size two. Let treatment A and treatment B be assigned at random within each block. Moreover, suppose the yields are obtained in the form of paired, p elements, column vector observations $(x_{A\alpha}, x_{B\alpha})$, $\alpha = 1, \cdots, N$. We consider a test of the null hypothesis of no treatment difference against the alternative that the treatments are different. Let $z_\alpha = x_{A\alpha} - x_{B\alpha}$, $\alpha = 1, \cdots, N$. Suppose for each j, we assign $+1$ if $z_\alpha{}^j > 0$ or -1 if $z_\alpha{}^j \leqq 0$. The following test is proposed and discussed: compute $P^2 = (4/N)(N_+ - N/2)'R^{-1}(N_+ - N/2)$ where N_+ is the vector of sums of $+1$ and R is a sign correlation matrix with the entries $R^{jk} = 2(N_{++}^{jk} + N_{--}^{jk} - N/2)/N$ with N_{++}^{jk} = number of pairs $(z_\alpha{}^j, z_\alpha{}^k) > (0, 0)$ and N_{--}^{jk} = number of pairs $(z_\alpha{}^j, z_\alpha{}^k) \leqq (0, 0)$, $j, k = 1, \cdots, p$. Reject the null hypothesis if P^2 is too large. To develop the null distribution of P^2 we multiply each vector z_α by a random variable e_α which takes the value $+1$ with probability $\frac{1}{2}$ and -1 with probability $\frac{1}{2}$. e_α are independent. It has been shown that use of P^2 is equivalent to use of the standard Hotelling T^2, for the paired-sample problem, with signs. The null distribution of P^2 is approximated by a beta distribution by equating the corresponding mean and variance. It has been shown that P^2 sign test is consistent and it's asymptotic relative efficiency with respect to the Hotelling T^2 test under the assumptions of normality and uncorrelated variates is $2/\pi$. Using the same technique, we have developed a multivariate normal scores test. The test procedures are illustrated by numerical examples. (Received 1 March 1966.)

31. On a class of rank order estimators of contrasts in MANOVA. M. L. PURI and P. K. SEN, New York University and University of North Carolina.

This paper is concerned with the multivariate generalizations of a class of estimators of contrasts in linear models. These estimators are all based on a class of rank order statistics and may be regarded as the multivariate generalizations of similar estimates considered by Hodges and Lehmann [(1963), *Ann. Math. Statist.*], Lehmann [(1963 a), (1963 b), (1963 c), ibid], Sen [(1963), *Biometrics*], Bhuchongkul and Puri [(1965), *Ann. Math. Statist.*], among others. Further, some results on a class of multivariate rank order tests and estimates, derived by the present authors [Abstracts: *Ann. Math. Statist.* (1966)] are utilized here for the study of the asymptotic properties of these estimators. (Received 4 March 1966.)

32. Monte Carlo analysis of the rate of convergence of some empirical Bayes point estimators (preliminary report). J. R. RUTHERFORD, Royal Military College of Canada.

Robbins (1964) *Ann. Math. Statist.* **35** 1–20, describes the empirical Bayes decision problem. In a Ph.D. thesis "Some Parametric Empirical Bayes Techniques", Virginia Polytechnic Institute (1965), the present author developed some asymptotically optimal empirical Bayes point estimators. This paper investigates the rate of convergence to optimality for several conditional density functions: normal pdf with known variance and unknown mean; normal pdf with known mean and unknown variance; and normal pdf with both mean and variance unknown. The rates of convergence are studied using Monte Carlo methods on an IBM 1620. (Received 8 March 1966.)

33. Maximization with respect to partition of an interval and its application to the best systematic estimators of the exponential distribution. M. SIBUYA, University of Western Ontario.

Let $\{X_i\}$ be a sequence of uncorrelated random variables with $E(X_i) = a_i\mu$ and $V(X_i) = b_i\sigma^2$, where $\{a_i\}$ and $\{b_i\}$ are known sequences of real numbers. Partition the natural numbers into $k + 1$ 'intervals,' $I_j = \{n_j, n_j + 1, \cdots, n_{j+1} - 1\}$, $j = 0, 1, \cdots k$, $l = n_0 < \cdots < n_{k+1} = \infty$, and estimate μ by a linear combination of k partial sums $\sum_{i\varepsilon I_j} X_i$, $j = 0, \cdots, k - 1$. The minimum variance unbiased estimator is obtained by the partition which maximizes $\sum_j ((\sum_{i\varepsilon I_j} a_i)^2/\sum_{i\varepsilon I_j} b_i)$. This is one of the dynamic programming problems, and so is its asymptotic case. Estimation of the exponential distribution by linear combination of a subset of order statistics is a special case of the problem. (Received 7 March 1966.)

34. Statistics connected with the uniform distribution; percentage points and application to testing for randomness of directions. M. A. STEPHENS, McGill University.

The paper gives percentage points of the distributions of the statistics $P = \bar{x}$; $Q = (\sum (1 - x_i^2)^{\frac{1}{2}}/N$, $T = (\sum x_i^2)/N$, where $x_1, x_2, \cdots, x_N$ is a random sample from the continuous uniform distribution with limits 0, 1. The statistics arose in testing a sample of directions, in three dimensions, for randomness, against certain proposed alternatives; this application is treated in the paper. (Received 10 March 1966.)

35. Easy distribution-free symmetry tests, and power against non-symmetrical alternatives. RORY THOMPSON, Massachusetts Institute of Technology.

A new one-sample rank test for symmetry is proposed. The test is easier to use than the rival Wilcoxon and normal scores tests, as it involves only inspection of the rank-order, with no summations. The test performs well against short-tailed shift alternatives, but poorly against long-tailed shift alternatives. Maximum power occurs for rectangular shift for monotone rank tests. Minimum power occurs for Birnbaum alternatives. A general formula for power against rectangular shift is given. The test proposed is locally most powerful for rectangular shift. (Received 7 March 1966.)

36. Bayesian analysis of hierarchical design model. GEORGE C. TIAO, Harvard University.

We consider the model $y_{ijk} = \mu + a_i + b_{ij} + e_{ijk}$, $(i = 1, \cdots, I; j = 1, \cdots, J; k = 1, \cdots, K)$ where y_{ijk} are the observations, μ is a common location parameter, a_i and b_{ij} are two different kinds of random effects and e_{ijk} are random disturbances. It is assumed that a_i b_{ij}, and e_{ijk} are normally and independently distributed with zero means and variances σ_a^2, σ_b^2 and σ_e^2, respectively. Employing a non-informative prior distribution, the joint posterior distribution of the variance-components $(\sigma_a^2, \sigma_b^2, \sigma_e^2)$ is obtained. Several features of the joint distribution are discussed. In particular, it is shown that the marginal distributions of the individual components can be closely approximated by that of scaled χ^2 variables. The connection between the Bayesian procedure and the problem of "pooling" variances in sampling theory is discussed. In addition, the distribution of the ratios $(\sigma_a^2/\sigma_y^2, \sigma_b^2/\sigma_y^2)$ where $\sigma_y^2 = \sigma_a^2 + \sigma_b^2 + \sigma_e^2$, is derived and illustrated in detail by an example. It is shown that there may be situations in which the two ratios are highly negatively correlated, and thus conclusions drawn by considering merely the individual marginal distributions of the variance components in such cases could be very misleading. (Received 9 March 1966.)

37. The distributions and moments of some variance components estimators (preliminary report). YING YAO WANG, Cornell University.

In the usual balanced Model II analysis of variance problem, an estimator for a variance component may be given in the form of a linear difference of two mean squares, while its truncated version provides a non-negative estimator. The distributions and moments of these two classes of estimators are given, and a study is made of their behavior under certain conditions of the parameter space. (Received 2 March 1966.)

38. Moment relations for some discrete distributions. J. K. WANI, McGill University and Lethbridge Junior College.

In this paper moments of some discrete distributions are studied for recursion relations. The set of moments obtained through these relations uniquely determines the respective distributions. This solves the famous Hamburger moment problem for these distributions. It is shown that the functional form of the moments of the binomial, logarithmic series and Poisson distributions can be obtained from that of the negative binomial distribution. The distributions studied belong to the class of series distributions but not all the discrete distributions in this class can have their moments in elegant functional form. This depends upon the particular series function. A table is given which can be used to write down the functional form of the above mentioned distributions. The functional form of the cumulants of the negative binomial distribution is identical with that of the moments of the logarithmic series distribution both in the univariate and the multivariate case. On account of this, recursion relations for the cumulants of the negative binomial distribution are automatically obtained. (Received 2 March 1966.)

39. The effect of rounding on a derived distribution. J. L. WARNER and R. G. KRUTCHKOFF, Bell Telephone Laboratories and Virginia Polytechnic Institute.

The sum of squares of observations from a standard normal distribution is known to have a chi-square distribution. A discrete distribution derived by taking the sum of squares of observations from a discretized (rounded) standard normal distribution is, however, found to be multimodal and ragged. For small degrees of freedom there is a large difference between the histograms of the derived distribution and the "associated" discretized chi-square distribution both in shape and in scale. Two rounding intervals and various small degrees of freedom are chosen to study the difference between the histograms. Specifically, the interval 0.5 for degrees of freedom one to five and the interval 0.1 for degrees of freedom one to three are studied. By comparing the cumulative probabilities for like abscissa values it has also been found that the discrete derived distribution has larger tail areas than the corresponding continuous chi-square distribution. For example, with two degrees of freedom and a rounding interval of 0.1 the discrete derived distribution has 17% more area to the left of $x = 0.10$ and 0.07% more area to the right of $x = 7.20$. (Received 7 March 1966.)

40. Linear minimax estimation for linear regression models. SAMUEL ZAHL, University of Connecticut.

Linear minimax estimates of points on a regression are found for a three parameter linear regression of the usual type (independent observations, common variance) when the parameters of the regression are bounded, these bounds together with the variance are known, and a quadratic loss is used. The method of solution is first to convert the minimax problem into a quadratic programming problem and then to solve the latter problem algebraically using the deformation method algorithm [Zahl, *J. Roy. Statist. Soc. Ser. B* (1964) (1965)] together with a partial dual of the primal algorithm. A simple special case, in which the regression is of the form $\theta_1 + \theta_2 t$, with θ_1, θ_2 the parameters, was earlier given in (Zahl, *IEEE Trans. of the PTGIT*, IT-10, July 1964). In principle the method can be used for estimating arbitrary linear parametric functions and also for regressions involving more than three parameters, but the tediousness of the algebra may limit these extensions for algebraic results; numerical estimates are readily found however. Possible applications are to process control and to trajectory prediction of moving objects since often, in these applications, only the last few observations are useful for estimation, while to compensate for the few observations, prior information of the type needed is available. (Received 7 March 1966.)

(Abstracts of papers to be presented at the Western Regional meeting, Los Angeles, California, August 15–17, 1966. Additional abstracts will appear in future issues.)

1. On pseudo games. ALFREDO BAÑOS, University of California, Los Angeles.

In the definition of a two-person zero-sum game both players know the rules of the game. In a pseudo-game at least one of the players does not. Consider a pseudo-game in which player I, say, is aware only of his pure strategy choices (assumed to be finite) and not of player II's strategy choices or of the distribution of the payoff with given strategy choices (assumed to have uniformly bounded second moments). Player II has complete information. Suppose that players I and II play a sequence of identical pseudo-games. Let v denote the value of a single game given that both players have complete information and let X_j denote the loss incurred by player I at the jth place of the sequence. When playing a "sequence of games" a strategy for player I (II) would be a rule P (Q) that would tell him as a function of his past plays and losses what mixed strategy to play at the jth place in the sequence. For the case being considered, a rule P_0 has been constructed such that for all Q, Pr $(\limsup_n \sum_{j=1}^{n} X_j/n \leqq v \mid P_0, Q) = 1$, i.e. the player with incomplete information can do as well asymptotically as he could if he had complete information. (Received 21 February 1966.)

2. Gaussian distributions for random lines, planes, and hyperplanes. BRADLEY EFRON, Stanford University.

A family of distributions called "Gaussian" is introduced for random linear varieties in a Euclidean space. This family is shown to result when the random varieties are generated from ordinary Gaussian vectors by any one of several different "natural" procedures. Estimation of parameters and the testing of linear hypotheses are shown to be straightforward generalizations of the usual least squares techniques for random normal vectors. (Received 10 February 1966.)

3. Detection of an unknown waveform which is randomly recurring in Gaussian noise. MELVIN J. HINICH, Hudson Laboratories, Columbia University, Dobbs Ferry, New York.

Matched-filtering is a widely used signal processing technique for detecting a known signal imbedded in noise. In the pulsed sonar case, the medium often distorts the received signal to such an extent that matched-filtering is not optimal. This paper deals with the case where a record is obtained of Gaussian noise containing many repetitions of the distorted signal waveform. After obtaining an estimate of the waveform autocorrelation based upon non-overlapping sections from the first part of the record, a test procedure is developed for determining whether or not the waveform is present in a subsequent section. The test procedure is derived from Taylor series approximation of the likelihood-ratio, where the signal-to-noise ratio is quite low and the probability that two or more waveforms are in the section, is negligible. This suboptimal test involves cross-correlating the sample autocorrelation with the waveform autocorrelation. An approximation is given to the difference in power between the likelihood-ratio test and its approximation. A comparison is made with the matched-filter approach. (Received 17 February 1966.)

4. Asymptotic efficiency of certain rank tests for comparative experiments. K. L. MEHRA and J. SARANGI, University of Alberta.

Consider a complete experimental design with k objects ($j = 1, \cdots k$), n blocks ($i = 1, 2, \cdots n$) and $m_{ij} = m_j \geqq 1$ independent observations X_{ijl}, $l = 1, 2, \cdots m_j$ in the (i, j)th cell. Assume that X_{ijl} is distributed according to a continuous distribution function $F_{ij}(x) = F_j(x + \xi_i)$ with unknown ξ_i's (block effects). For testing the hypothesis $H_0: F_1 = F_2 = \cdots = F_k$, conditional rank tests, based on a combined ranking of the totality of cell observations after "alignment," were proposed by Hodges and Lehmann in (1962) (*Ann. Math. Statist.* **33** 482–487) on the premise that the interblock comparisons would bring in improvements in the relative efficiency. We show in this paper, using Hoeffding's results on U-statistics, that this is indeed the case for translation alternatives when the "alignment" is done on mean and the parent distribution F is normal. Specifically, we derive explicit expressions for the asymptotic efficiency of the above conditional test and show that relative to the classical F-test, it is never less than $3/\pi$ under normality and relative to the Friedman's separate-rankings test it is always greater than unity, whatever be k. Certain results of the above paper, on the asymptotic distribution of the conditional test statistics for the case $k = 2$, are also extended to general complete designs. (Received 21 March 1966.)

(*Abstracts of papers to be presented at the Annual meeting, New Brunswick, New Jersey, August 30–September 2, 1966. Additional abstracts will appear in future issues.*)

1. Further studies on the robustness of exponential life testing procedures (preliminary report). A. P. BASU, University of Minnesota. (By title)

In this paper the robustness of the exponential life testing procedures has been studied when the true distribution is the Weibull or the gamma distribution. For this purpose we have used a simple approximation for the distribution of sums of observations following the Weibull distribution. The findings are in agreement with that of Zelen and Dannemiller [*Technometrics* **3** (1961)] viz. that the exponential life testing procedures are non-robust. Moreover, use of the above simple transformation has enabled us to consider quite a large number of interesting cases not covered by them. (Received 4 April 1966.)

2. Queues with state-dependent stochastic service rates. CARL M. HARRIS, Western Electric Company, Inc., Princeton.

The standard $M/G/1$ queueing system is generalized so that the service time parameter becomes a stochastic process, $\{M_n, n = 1, 2, \cdots\}$, indexed on the length of the queue at the moment service is begun. The queue length, i, at that instant generates a member, M_i, of the family of independent random variables, $\{M_n\}$, with distribution function, $F_{M_i}(\mu_i)$. The service time is determined by selecting an element, μ_i, from the sample space of M_i, and then, given μ_i, sampling the service time population according to its known conditional distribution function. Three specific examples are explored. Results are obtained which characterize queue size and waiting time using the imbedded Markov chain approach. In the first case, all conditional service time distribution functions, $\{B_n(t \mid \mu_n)\}$, are equal and the service time parameters are identical random variables. For case two, the conditional distribution functions are assumed exponential and identical for $n = 2, 3, \cdots$, that is, when there is a lone customer in the system, the service procedure differs from that at all other times. For the final example, the conditional distribution functions are again exponential, but all distinct. (Received 8 March 1966.)

3. On the probability of large deviations of functions of several empirical cdf's (preliminary report). A. BRUCE HOADLEY, University of California, Berkeley.

Sanov [On the probability of large deviations of random variables. *Select. Transl. Math. Statist. Prob.* **1** (1961) 213–244] has shown that if F_N is the empirical cdf of a sample whose true cdf is F_0, and V is a set of cdf's which satisfies certain conditions and does not contain F_0, then $\lim_{N\to\infty} N^{-1} \ln P\{F_N \varepsilon V\} = -\inf_{F \varepsilon V} \int \ln (dF/dF_0)\, dF$. In the present paper, Sanov's result is generalized and extended to the c-sample case as follows: Let F_{i,n_i} be the empirical cdf of the ith sample whose true cdf is $F_{0,i}$; let $N = \sum_{i=1}^{c} n_i$, $S_N = (F_{1,n_1}, \cdots, F_{c,n_c})$, and $Q_0 = (F_{0,1}, \cdots, F_{0,c})$; and let D be the set of c-vectors of cdf's. If $|n_i/N - \rho_i| = O(\ln N/N)$ as $N \to \infty$, $F_{0,i}$ is continuous, and $W_N \subset D$ satisfies certain conditions, then $\lim_{N\to\infty} N^{-1} \ln P\{S_N \varepsilon W_N\} = -\lim_{N\to\infty} I(W_N, Q_0)$, where $I(W_N, Q_0) = \inf_{(F_1,\cdots,F_c)\varepsilon W_N} \sum_{i=1}^{c} \rho_i \int \ln (dF_i/dF_{0,i})\, dF_i$. It is shown that the above result is true for $W_N(r) = \{Q \varepsilon D: T(Q) \geqq r_N\}$ when T is a uniformly continuous function wrt the metric defined by $d(Q, R) = \max_{1\leqq i \leqq c} \{\sup_x |F_i(x) - G_i(x)|\}$, and $\lim_{N\to\infty} r_N = r > T(Q_0)$, with r being a continuity point of $I(W_N(r), Q_0)$. This form of the theorem has many applications; and usable formulas are derived for the two sample Wilcoxon statistic when the two true cdf's are equal, and for the one sample Kolmogorov-Smirnoff D_N^- statistic both under a particular hypothesis and a wide class of stochastically larger alternatives. It is also shown that the above theorem can be applied to any statistic T_N which can be suitably approximated by a statistic of the form $U_N(F_{1,n_1}, \cdots, F_{c,n_c})$, where U_N is a uniformly continuous function. (Received 7 February 1966.)

4. Estimation of the mean of the multivariate normal distribution with an empirical Bayes application. MICHAEL KANTOR, Columbia University.

Let X be a $p \times 1$ vector which is normally distributed with mean vector θ, and covariance matrix $\sigma^2 I_p$. The problem considered is to estimate θ with a loss function $L(\hat{\theta}, \theta) = \sum (\hat{\theta}_i - \theta_i)^2$ where $\hat{\theta}$ is the estimator of θ. This paper considers minimax estimators $\varphi(X) = (1 - (p-3)\sigma^2/\sum(X_i - \bar{X})^2)(X - \bar{X}e) + \bar{X}e$ for σ known and σ unknown,

$$\varphi^*(X) = \left(1 - \frac{(p-3)S}{(n+2)\sum(X_i - \bar{X})^2}\right) \cdot (X - \bar{X}e) + \bar{X}e)$$

where S/σ^2 is distributed as a chi-square variable with n degrees of freedon, S is independent of X, and $e = (1, 1, \cdots 1)'$. The risks of these estimates are compared with those given by James and Stein in the Fourth Berkeley Symposium. Let $\varphi_p{}^*$ be the pth component of φ^*. Then $\{\varphi_p{}^*\}$ is shown to be a subminimax sequence of estimators for the empirical Bayes sequence of univariate normal mean problems. (Received 21 March 1966.)

(Abstract of a paper to be presented at the European Regional meeting, London, England, September 5–10, 1966. Additional abstracts will appear in future issues.)

1. Some Bayesian stratified two-stage sampling results (preliminary report). NORMAN R. DRAPER and IRWIN GUTTMAN, University of Wisconsin.

The problem of selecting sample sizes in the second stage of a two-stage procedure (the first stage of which can take various forms) when the variançes of the strata are *unknown* is considered. Using a straightforward Bayesian analysis provides an allocation procedure which minimizes the posterior variance of the mean of the stratified population. An algorithm for practical implementation of the procedure is provided. A feature of this method is that it takes into account "oversampling" in the first stage. Another approach to the problem can be made through a pre-posterior analysis. This leads to an alternative procedure which is compared to, and contrasted with, the first. The work is being extended. (Received 28 March 1966.)

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

1. An example in the Darmois-Koopman problem (preliminary report). J. L. DENNY, University of California, Riverside.

For a family of probability distributions on an interval, Lawrence Brown [*Ann. Math. Statist.* **35** (1964) 1456–1474] gave a corrected version of a Darmois-Koopman type theorem of Dynkin. We consider the consequences of relaxing the continuously differentiable density condition in the Brown-Dynkin theorem. More precisely, we first construct on an interval I a family of densities, each of which is Lipschitz (1) and bounded away from zero. The family is not representable as an exponential family on any set of positive Lebesgue measure. For each sample size n there is a minimal sufficient statistic T so that for each $x \varepsilon I^n$ and each neighborhood U of x the sigma-field induced by T on U is strictly smaller, modulo equivalences, then the sigma-field induced by the order statistics on U. (Received 18 March 1966.)

2. Maximal 2^{k-p} designs of resolution V (preliminary report). NORMAN R. DRAPER and TOBY J. MITCHELL, University of Wisconsin.

This paper presents a method for finding 2^{k-p} designs of Resolution V which incorporate the maximum possible number of variables for a given number of runs. For the 256 run case, Addelman (*Technometrics* **7** 439–443) constructed a specific design which included $k = 17$ variables and speculated that this was the maximum number possible. We have shown that this is true and, moreover, that only one such design exists. The corresponding solution in the 512 run case is $k = 23$. The work is being continued. (Received 11 April 1966.)

3. The empirical Bayes shortest confidence intervals for estimating the mean of a finite population. V. P. GODAMBE, Johns Hopkins University.

Earlier the author (1961, *ISI, Paris* meetings) defined Bayes shortest confidence intervals as an interval of parametric values which for a given length maximises the *posteriori* probability of containing the true parameter. These Bayes shortest confidence intervals were subsequently used by Box and Tiao [*Ann. Math. Statist.* **36** (1965) 1468–1482] and Watson [(1965) *Technical Report* No. 34, The Johns Hopkins Univ.] to obtain goodness of fit tests. Now the above Bayes shortest confidence intervals are called empirical Bayes shortest, if the prior itself is estimated from the sample. With this definition it is now proved that for sampling from a finite population $\bar{x} \pm k[(1/n - 1/N)s^2]^{\frac{1}{2}}$ (notation as usual, k being any arbitrary constant) are empirical Bayes shortest confidence intervals for the population mean. Though these confidence intervals are the same as the one in common use, for sufficiently large n and N, the *posteriori* probability associated with them would be numerically equal to the frequency with which the intervals $\bar{x} \pm k[(N/n)(1/n - 1/N)]^{\frac{1}{2}}$, ($k$ being the same as before) cover the population mean, in repeated sampling, with equal probabilities and without replacement. This in a way also generalizes the author's previous [(1965), *Technical Report* No. 41, The Johns Hopkins Univ. and *Ann. Math. Statist.* **37** 522 (abstract)] result establishing the empirical Bayes character of the sample mean, while estimating the mean of a finite population. (Received 1 April 1966.)

4. Some aspects of stochastic simultaneous equations estimation theory. D. G. Kabe, Northern Michigan University.

Let $By_t + \Gamma x_t = u_t$ be a simultaneous equations model, where B is a $G \times G$ positive definite symmetric matrix, Γ, a $G \times K$ matrix of rank $G(<K)$, y_t, x_t, and u_t are column vectors of G, K, and G elements respectively. The error vector u_t has mean zero and a positive definite symmetric matrix as its covariance matrix or u_t has a G variate normal distribution. Then several results in the simultaneous equations estimation theory under the assumption that x_t is a vector of predetermined variables have been reexamined under the assumption that x_t has a K variate normal distribution independent of the distribution of u_t and that for $t = 1, 2, \cdots, N$, the distributions of x_t are independent each x_t having the same distribution. Under this assumption we consider the estimation of the parameters of the first equation of the above model by using two stage least squares and generalize our results to k class estimators. Forecasting theory with the reduced form estimated structural model has also been treated.

5. One sided univarite general linear hypothesis with linear restrictions. D. G. Kabe, Northern Michigan University.

Let $Y = X\beta + e$ be the usual univariate normal linear regression model, where the $N \times q$ matrix X is of rank $q(<N)$ and β is subjected to g estimable linear restrictions $F\beta = W$. The $g \times q$ known matrix F is of rank $g(<q)$, and W is known. The problem considered in this paper is the testing of $M\beta = 0$ against the one sided alternative $M\beta \geqq 0$. We assume that the known $g_1 \times q$ matrix M is of rank $g_1(<g)$. The test criteria and their distributions are given when σ^2 is known and σ^2 is unknown, where e as usual is $N(0, \sigma^2 I)$. This generalizes a result given by P. S. Dwyer [*Ann. Math. Statist.* **29** (1958) 106–117]. (Received 7 March 1966.)

6. Asymptotic efficiencies for the empty cell and run tests. Jerome H. Klotz, University of Wisconsin.

Asymptotic efficiencies for the two sample empty cell test and the two sample run test are derived by comparing the exponential rate of convergence to zero of the type I error (α) while keeping the type II error (β) fixed ($0 < \beta < 1$). Using the closed form expressions for the null distributions and equal samples of size n, it is shown for both tests that $\lim_{n\to\infty} (-1/n) \log \alpha_n = 2Q \log 2Q + 2(1 - Q) \log 2(1 - Q)$ where $Q = \int_{-\infty}^{\infty} (f^2(x)/(f(x) + g(x)))\, dx$

and f and g are the densities under the alternative for the two samples. As a consequence, the limiting relative efficiency for the two tests is one. Further, for either test, the corresponding limiting efficiency relative to the t-test under normal shift alternatives becomes zero as the shift goes to zero. (Reccived 7 March 1966.)

7. The non-central F distribution-extensions of Tang's tables (preliminary report). PETER A. LACHENBRUCH, University of North Carolina.

This paper provides two extensions of Tang's tables of the error of the second kind for the F test. The first part considers higher numerator degrees of freedom which are easily obtained using Tang's formulas. The second extension is tables of percentage points of the non-central F distribution for values of .01, .025, .05, .10, .50, .90, .95, .975 and .99. These are obtained for a variety of numerator and denominator degrees of freedom and for values of the non-centrality parameter $\lambda = 1(1)10$. The values are computed by the Newton-Raphson method to a tolerance of .0001 for the percentage point. (Received 11 April 1966.)

8. Nonparametric ranking procedures for comparison with a control, I (preliminary report). M. HASEEB RIZVI, MILTON SOBEL and GEORGE WOODWORTH, Ohio State University and University of Minnesota.

A population with cdf F_i is called better than one with cdf F_j if the α-fractile $x_\alpha(F_i)$ of F_i is larger than that of F_j; the ordered cdf's are then denoted by $F_{[i]}$ $(i = 1, 2, \cdots, k)$ and these are compared with the standard whose cdf is F_0. A nonparametric procedure R_1 is developed for selecting all populations as good or better than the control. It is assumed that the $k + 1$ cdf's are continuous and that $1 \leqq (n + 1)\alpha \leqq n$ where n is the common number of observations per population; the standard F_0 may be unknown and sampled like the others (or it may be known in which case it is not sampled). Let $Y_{j,i}$ denote the jth order statistic from F_i; the proposed procedure R_1 selects all F_i for which $Y_{r,i} \geqq Y_{r,0}$ (or $x_\alpha(F_0)$ if F_0 is known) where r is an integer defined by $r \leqq (n + 1)\alpha < r + 1$. Fix $\epsilon^* > 0, d^* > 0$ and $P^* < 1$, let I denote the interval $[x_{\alpha-\epsilon^*}(F_0), x_{\alpha+\epsilon^*}(F_0)]$. Let d_j denote the min $(F_0(x) - F_j(x))$ for $x \varepsilon I$. Then n is determined so that the Prob {Correct Selection $|$ R_1} $\geqq P^*$ when all the cdf's F_j better than the control are such that $d_j \geqq d^*$. The asymptotic efficiency of R_1 relative to other procedures is studied. (Received 5 April 1966.)

9. Nonparametric ranking procedures for comparison with a control, II (preliminary report). M. HASEEB RIZVI, MILTON SOBEL and GEORGE WOODWORTH, Ohio State University and University of Minnesota.

Using the notation of the preceding abstract, a nonparametric procedure R_2 is developed for selecting a subset containing all those populations that are as good or better than the control. Under R_2, we put F_i in the selected subset iff $Y_{r,i} \geqq Y_{r-c,0}$ (or $x_{\alpha-\beta}(F_0)$ if F_0 is known); here $c = (n + 1)\beta$ is the smallest integer $\leqq r$ for which P{Correct Selection $|$ R_2} $\geqq$ P^* whenever for each j we have either $F_j(y) \leqq F_0(y)$ for all y or $F_j(y) \geqq F_0(y)$ for all y and $x_\alpha(F_0) < x_\alpha(F_j)$. The asymptotic efficiency of R_2 relative to other procedures is studied. In a secondary problem a definition for a population to be ϵ-inferior to F_0 is given and n is then determined so that E{Number of ϵ-inferiors in the selected subset} $< \eta$ for small specified $\eta > 0$. (Received 5 April 1966.)

10. Testing the homogeneity of truncated Poisson distributions. S. N. SINGH, Banaras Hindu University.

Let the random variables X and Y have truncated Poisson distributions $P(\lambda)$ and $P(\mu)$ (truncated at zero). Following Lehmann, 'Testing statistical hypothesis' the UMP unbiased test for testing the hypothesis H: $(\mu \leqq a\lambda)$ against the alternative $(\mu > a\lambda)$ has been worked out The hypothesis is rejected for large values by Y, where Y has a binomial distribution truncated at zero and z for given $X + Y = z$. The UMP unbiased test has also been found for the above hypothesis when samples $(X_1, X_2, \cdots, X_m)$, $(Y_1, Y_2, \cdots Y_n)$ are taken from $P(\lambda)$ and $P(\mu)$ respectively. (Received 28 March 1966.)

11. A test of inflation in Poisson distribution. S. N. SINGH, Banaras Hindu University.

Singh [*J. Indian Statist. Assoc.* (1963)] has defined the inflated Poisson distribution as $P[X = 0] = 1 - \alpha + \alpha e^{-\lambda}$, $P[X = x] = \alpha e^{-\lambda}(\lambda^x/x!)$ for $x = 1, 2, \cdots, 0 \leqq \alpha \leqq 1; 0 < \lambda < \infty$. This distribution is applicable in situations where there are more observations with zero than can be expected on the basis of simple Poisson. The hypothesis that there is no inflation is equivalent to the hypothesis H: $(\alpha = 1)$. The alternative $(\alpha < 1)$, that is, there is inflation, is appropriate. We have obtained the UMP unbiased test for this situation. The test rejects H when N_0, the number of observations with zero, is large for given $\sum X_i = z$, where $X_1, X_2, \cdots, X_n$, is a random sample. The exact distribution of N_0 given $\sum X_i = z$ has been obtained. (Received 6 April 1966.)

12. On ratio and linear regression methods of estimation using several auxiliary variables. SURENDRA K. SRIVASTAVA, Lucknow University.

For estimating the mean $\bar{Y}$ of a finite population when information on p (>1) auxiliary variables is available, Olkin (*Biometrika* **45** (1958) 154–165) used an estimator $\bar{y} = \sum_{i=1}^{p} w_i t_i$ where t_i is the ratio estimator of $\bar{Y}$ based on ith auxiliary variable only and w_i's are the weights to be determined by minimizing the variance of the estimator $\bar{y}$ subject to the condition $\sum_{i=1}^{p} w_i = 1$. It is shown that if instead of ratio, linear regression estimator based on ith auxiliary variable is used for t_i, the resulting estimator does not necessarily have smaller variance. A condition has been obtained, in a special case, when the estimator $\bar{y}$ based on ratio estimators has smaller variance than that based on linear regression estimators. (Received 16 February 1966.)