ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Washington meeting of the Institute

April 29–May 1, 1953

1. **Optimum Sample Sizes for Choosing the Largest of $(k + 1)$ Means Using Minimax Methods.** PAUL N. SOMERVILLE, University of North Carolina.

Assume we have $(k + 1)$ normally distributed populations with unknown means $a_0 \geqq a_i \geqq \cdots \geqq a_k$. It is decided to choose N individuals from these populations in such a way that the expected value of their total is as large as possible. A preliminary sample of n is taken from each population with the object of deciding from which population the further sample of size N should be taken. $N(a_i - a_0)$ is then the loss involved in choosing the population with parameter a_i. Assume the cost of the sample is a linear function of the sample size. Using results previously given it is shown that the minimax n is proportional to $N^{(2/3)}$. Explicit results are given for $k = 1, 2, 3, 4, 5$, for a one-stage preliminary sample. For the case $k = 2$, results for a two-stage sample are given. In the first stage, samples of n_1 are taken for each of the three populations. In the second stage, samples of n_2 are taken from each of the two populations with the largest means in the first stage. If $3n_1 + 2n_2 = 3n$, then it is found that the maximum expected loss is less for the two-stage sample than for the one-stage sample provided n_1/n_2 is greater than .37 (approximately). The optimum ratio in this sense is found to be $n_1/n_2 = 1.2$ (approximately). If for $n_1/n_2 = 1.2$, the maximum expected losses are equated by a reduction in the total preliminary sample size, a saving of 6.6 per cent over the one-stage procedure in the preliminary sample size is effected.

2. **The Correspondence Between Two Classes of Balanced Incomplete Block Designs.** W. S. CONNOR, National Bureau of Standards.

Let $\Sigma_1(n)$ denote the problem of constructing the design with parameters $v = \frac{1}{2}n(n + 1)$, $b = \frac{1}{2}(n + 1)\ (n + 2)$, $k = n$, $r = n + 2$, and $\lambda = 2$; and let $\Sigma_2(n)$ denote the problem of constructing the design with parameters $v = b = \frac{1}{2}(n + 1)\ (n + 2) + 1$, $r = k = n + 2$, and $\lambda = 2$, $(n > 1)$. It is shown that $\Sigma_1(n)$ has a solution only if $\Sigma_2(n)$ has a solution.

3. **A Finite Frequency Theory of Probability.** A. H. COPELAND, SR., University of Michigan.

This paper develops a new theory of probability, the finite frequency theory, in which probabilities are regarded as physical hypotheses. Associated with each probability is a system of predictions which can be tested by experiment. An experiment may either confirm or disagree with a given prediction. This theory of probability produces some complications in formal logic. However the theory and its associated deductive and inductive logics are in better agreement with modern scientific reasoning than the conventional probability theories and the conventional logics.

4. **Characterizations of Complete Classes of Tests of Some Multiparametric Hypotheses, with Applications to Likelihood Ratio Tests.** ALLAN BIRNBAUM, Columbia University.

Let H_0 be a simple hypothesis on a density function of the form

$$p_\varphi(e) = \exp \{\varphi_0 + \Sigma_1^k \varphi_i t_i(e) + t_0(e)\}.$$

Let T, the range of the sufficient statistic $t = (t_1, \cdots, t_k)$, be independent of t. Let V' be the class of nonrandomized decision functions $\delta(t)$ such that each $\delta(t) = 0$ just on the intersection of some open convex set with T. Let V be the class of randomized decision functions $\delta(t)$ each of which coincides with a member of V' except on a set of measure zero. Then under certain assumptions it is shown (a) that V' is essentially complete, and (b) that V is complete. Under further assumptions, chiefly requirements that the alternative hypothesis be sufficiently general, it is shown (a) that V' is minimal essentially complete, and (b) that V is minimal complete. Applications are made to likelihood ratio tests of H_0, which are shown to be included in V', to discrete distributions of the form $p_\varphi(e)$, and to tests of composite hypotheses on $p_\varphi(e)$.

5. Confidence Regions for the Location of the Vertex in Quadratic Regression (Preliminary Report.) DAVID L. WALLACE, Princeton University.

Procedures are considered for obtaining confidence regions for the location of the vertex of a regression surface which is a quadratic function of k "determining" variables $x_1, \cdots, x_k$ from a sample with normal homoscedastic error on the dependent variable only. The hypothesis that $(x_1^0, \cdots, x_k^0)$ is the vertex of the regression surface is a general linear hypothesis; a set of k linear homogeneous equations in the regression coefficients in which the coefficients in the equations are linear functions of the $\{x_i^0\}$. For any general linear hypothesis of this form, a confidence region for $(x_1^0, \cdots, x_k^0)$ is obtained by the standard (F) test. This region possesses several "optimum" properties, but is unsatisfactory for practical applications. If each of the k single linear hypotheses making up the general linear hypothesis is tested separately by the standard (t) test, k different confidence regions, whose shapes are usually hyperboloids, are obtained for the $(x_1^0, \cdots, x_k^0)$. The intersection of these is a confidence region for $(x_1^0, \cdots, x_k^0)$ with bounded risk. Approximations to this intersection region by parallelepiped and polyhedra are discussed. Requirements for usable confidence regions are discussed and proposed procedures are rated primarily by these requirements.

6. The Noncentral Wishart Distribution. (Preliminary Report.) A. T. JAMES, Princeton University.

The noncentral Wishart distribution, as T. W. Anderson showed, is the central distribution multiplied by a symmetric function, ψ, of the latent roots α_i of the matrix

$$\Sigma^{-1}T\Sigma^{-\prime}A$$

where Σ is the $k \times k$ variance covariance matrix of the parent normal k-variate distributions, T is the $k \times k$ matrix of sums of squares and products of population means about their averages and A is the sample variance covariance matrix. It is shown that ψ is the average of an exponential function in several variables over the orthogonal group. The exponential function is an eigen value of the Laplace operator Δ, and Δ commutes with the operation of averaging over the group. Hence $\Delta\psi = \psi$. If Δ is expressed in terms of the latent roots α_i a system of second order partial differential equations for ψ is obtained, which can be solved in power series for $k \leqq 3$. For $k > 3$, the partial differential equations yield an effective system of recurrence relations for the coefficients of the multiple power series.

7. On Time-Dependent Waiting Line Processes. A. BRUCE CLARKE, University of Michigan.

A single-server waiting line process with Poisson distributions on the input and service times is considered. The parameters λ and μ of these Poisson distributions are assumed to be arbitrary nonnegative functions of time. An exact formula for the transition probabilities, $P_{\nu,n}(t)$, for the line to have length n at time t, given that it had length ν at time 0, is found. The formula involves a function which is defined as the solution of a certain Volterra type integral equation; this can be determined explicitly for the special case in which the ratio of λ and μ is independent of time, and numerically otherwise. The general method of solution is to use the Kolmogorov equations to obtain a hyperbolic partial differential equation for a modified characteristic function of the distribution, thus reducing the problem to a boundary value problem that can be solved by standard methods. The formula for $P_{\nu,n}(t)$ is used to discuss various properties of the distribution, with special attention to nonstationary processes.

8. Some Estimates Which Minimize the Least Upper Bound of a Probability Together with the Cost of Observation. H. S. Konijn, University of California, Berkeley.

When an investigator aims primarily at insuring a high chance of getting a point estimate T_N of an unknown parameter point θ within a reasonable distance α of θ, the loss function proportional to the distance $d(T_N, \theta)$, which is generally used implicitly or explicitly, is inappropriate, and should be replaced by $W = 1$ if $d > \alpha$ (or $>\alpha_i$ in direction i, $i = 1, 2, \cdots$) and $= 0$ otherwise. Somewhat similar considerations already arose in the theory of confidence intervals. Following in part a procedure of Wolfowitz (*Ann. Math. Stat.*, Vol. 21 (1950), pp. 218–230), the present paper bases the choice of estimating interval $[R_N, S_N]$ on the probability of covering θ, $Pr\{d(R_N, \theta) > \alpha'\}$, and $Pr\{d(S_N, \theta) > \alpha''\}$. When the cost of observation plays a role in the selection of estimates, it usually enters in the form of its mathematical expectation, but other ways may be considered. The paper investigates in detail the (highly manageable) case of a normal variate with known variance. In several instances it obtains explicit results, which allow suggestive comparisons with classical methods as to sample size, unbiasedness, "shortness," etc. A still different point of view is briefly formulated.

9. On a Multivariate Analogue of Student's t-Distribution, with Some Tables for the Bivariate Case. Charles W. Dunnett and Milton Sobel, Cornell University.

We consider the joint distribution of p variates $t_i = x_i/s$, $i = 1, 2, \cdots, p$. The x_i have a joint multivariate normal distribution with means 0, variances σ^2 and correlation matrix (ρ_{ij}); ns^2/σ^2 has a chi square distribution, independent of the x_i, with n degrees of freedom. The joint density function of the t_i is given by $|A|^{\frac{1}{2}}\Gamma(n + p/2)\ \{(n\pi)^{p/2}\Gamma(n/2)\ (1 + \Sigma_{i,j}\, a_{ij}t_it_j/n)^{(n+p)/2}\}^{-1}$, where $|A|$ is the determinant of the positive definite matrix $(a_{ij}) = (\rho_{ij})^{-1}$. This reduces to the Student t-distribution when $p = 1$. For the bivariate case ($p = 2$), the following results were obtained: (a) an exact expression in the form of a finite series for the probability integral from (h, k) to (∞, ∞), (b) an asymptotic series in powers of n^{-1} for this probability integral, (c) an asymptotic series in powers of n^{-1} for the value of $h = k$ for which the probability integral is equal to an arbitrary specified value, and (d) tables of the probability integral and certain percentage points for the special cases $h = k$ and $\rho = \pm\frac{1}{2}$, where ρ is the correlation between x_1 and x_2. These tables are required for certain multiple decision ranking problems involving three population means (*Ann. Math. Stat.*, Vol. 24 (1953), p. 136). (Research sponsored by Air Research and Development Command.)

10. On the Completeness of Classes of Bayes' Solutions. Lucien M. LeCam, University of California, Berkeley.

The terminology used is that in Wald's book, *Statistical Decision Functions*, John Wiley and Sons, 1950. It is shown that assumptions (3.3) and (3.4) of the preceding book can be replaced by the following weaker assumptions. (1) F is a function of an element ω in some arbitrary index set Ω. (2) The space D^t of terminal decisions is a compact metrisable Hausdorff space. (3) The weight function W depends on ω, d^t and possibly on the indices of the random variables actually observed. Moreover, $\inf_{d^t \, D^t, s_1, \cdots, s_k} W(\omega, d^t; s_1, \cdots, s_k) > -\infty$. (4) For each $\omega \varepsilon \Omega$ and each set $\{s_1, \cdots, s_k\}$, the function W is lower semicontinuous on D^t. The assumption of separability (3.2) loses part of its meaning and can be dropped. Assumptions (3.1), (3.5) and (3.6) can also be weakened but not very significantly. Under these weakened assumptions, the class of admissible decision functions is complete and theorems (3.5), (3.7), (3.8), (3.9), (3.17) and (3.18) remain true. In theorems (3.17) and (3.18), the class $\mathfrak{D}_b$ of decision procedures with bounded risk functions can be replaced by the class $\mathfrak{D}$ of all decision procedures.

11. Identification and Estimation of Linear Structures with Symmetric Errors. T. A. Jeeves, University of California, Berkeley.

Consider a vector random variable X with n-components having the following structure: (i) $X = \xi + U$; (ii) ξ is a vector random variable such that $B\xi = A$ where B is a $s \times n$ matrix of constants and A is a constant vector of s components; (iii) ξ and U are independent and (iv) the distribution of U is symmetric about some known point in n-dimensional space. A necessary and sufficient condition for the identifiability of A and the column space of B is that the distribution of ξ should not be symmetric about a point. An estimate based on the sample characteristic function is given which converges almost surely when the parameters are identifiable.

12. The Cramér-Smirnov Test in the Parametric Case. (Preliminary Report.) Donald A. Darling, Columbia University.

Given a set of n data (independent, identically distributed random variables) $X_1, X_2, \cdots, X_n$ we wish to test the hypothesis H that their common continuous cdf is $F(x; \theta)$ for some (unknown) value of the (real) parameter $\theta \varepsilon \Omega$. In modifying the usual chi square test where an auxiliary parameter is to be estimated we consider, following a suggestion of Cramér, the test function $W_n^2 = \sqrt{n} \int_{-\infty}^{\infty} (F_n(x) - F(x; \hat\theta_n))^2 dF(x; \hat\theta_n)$ where $F_n(x)$ is the empirical cdf of the data and θ_n is some estimate of θ. Two essentially distinct cases arise. a) If $\hat\theta_n$ is a superefficient estimator of θ W_n^2 has the same limiting distribution as in the nonparametric case—the Smirnov distribution. b) If $F(x; \theta)$ satisfies Cramér's conditions for regular estimation and an asymptotically efficient unbiased estimator $\hat\theta_n$ (the maximum likelihood estimator essentially) exists we have the following result: let $f = \partial F/\partial x$, $\sigma^2 = \lim_{n\to\infty} n \operatorname{Var}(\theta_n) = E\{(\partial \log f/\partial\theta)^2\}^{-1}$ and put $u = F(x; \theta)$, $h'(u) = \sigma\partial \log f/\partial\theta$, $0 \leqq u \leqq 1$ and let $a_k = \sqrt{2} \int_0^1 h(u) \sin \pi ku \, du$, $k = 1, 2, \cdots$, $G(\lambda) = 1 + \Sigma_1^\infty (\lambda a_k^2)/(1 - \lambda/\pi^2k^2)$. Then the limiting characteristic function of W_n^2 is $\sqrt{2it} \csc \sqrt{2it} (G(2it))^{-\frac{1}{2}}$. The method of proof is by a consideration of a Gaussian process following an idea of Doob. Unlike the corresponding nonparametric case the test is not distribution free, and in general the limiting distribution of W_n^2 will even depend on the true value θ. In important special cases including that where θ is a scale or location parameter the function $h(u)$, and consequently the distribution of W_n^2, does not depend on θ, however. The theory can be extended to the case of several unknown parameters, and it is possible to discuss the corresponding Kolmogoroff test function using these methods. (Research sponsored by Air Research and Development Command of the Air Force.)

13. Asymptotic Solutions of the Compound Problems for Two Completely Specified Populations. James F. Hannan and Herbert Robbins, Catholic University of America and University of North Carolina.

Let v be a vector of arbitrary dimensionality and let $F(v, 0)$ and $F(v, 1)$ be any two distinct distribution functions. Let $X_1, \cdots, X_n$ be independent random vectors such that X_i

has the distribution function $F(v, \theta_i)$. Let $X = (X_1, \cdots, X_n)$ and let $\theta = (\theta_1, \cdots, \theta_n)$. It is required to decide for each i, on the basis of X and the known distribution functions $F(v, 0)$ and $F(v, 1)$, whether θ_i is 0 or 1. The loss of the compound decision $d = (d_1, \cdots, d_n)$ is taken to be $W(d, \theta) = n^{-1} \sum_1^n [a\theta_i(1 - d_i) + b(1 - \theta_i)d_i]$, a and b being positive constants determined by the empirical background of the problem. This problem was previously considered for the special case of $N(-1, 1)$ and $N(1, 1)$ (Herbert Robbins, "Asymptotically subminimax solutions of compound statistical decision problems," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951, pp. 131–148). The present paper constitutes a generalization and amplification of results obtained there.

14. On the Estimation of the Mean Life of a Radioactive Source. (Preliminary Report.) Richard F. Link, Princeton University.

Procedures are given for estimating the mean life of a radioactive source assuming: i) the individual times at which particles disintegrate are recorded for a time interval (T_0, T_1), ii) the number of particles which disintegrate in each of K nonoverlapping time intervals is recorded. Methods for obtaining exact confidence intervals for the estimate of the mean life are presented for two of the procedures. Asymptotic variances are derived for all of the estimates. Comparisons of the asymptotic efficiency of the various methods are given. For two of the methods comparisons of the expected lengths of confidence intervals for the mean life, given that n disintegrations are observed, are presented for $n = 10, 25$.

15. The Use of the Questionnaire to Compare Two Populations for the Purpose of Improving the Course Content in a Mathematics Course for Business Teachers. Mary Goins, Marshall College.

A study was made to determine the relative amount and kind of mathematics required in business teacher training programs in professional schools of business as compared with such programs in teachers colleges. Random samples from the two populations were drawn. An appropriate questionnaire was sent to the administrators of the institutions. Results were tabulated and statistical computations made. On the average, professional collegiate schools of business were found to have stronger required courses in mathematics than teachers colleges having the same type of curriculum. Changes in course content based on the computed statistics are discussed in this paper.

16. Minimax Decisions Regarding Mean of a Normal Variable with Unknown Variance. Manindra N. Ghosh, University of North Carolina.

In a recent paper by the author (*Sankhyā*, Vol. 13) Wald's decision problem has been generalized to the case of unbounded weight function $W(F, d^t)$ and locally compact space D^t of decisions. In this paper some applications of this method to the case of decisions regarding the mean of a normal variable, in the fixed sample or sequential procedure have been made when the variance is unknown.

17. On Two-Stage Estimation Procedures. S. G. Ghurye and Herbert Robbins, University of North Carolina and Institute for Advanced Study, Princeton.

Let P_i, $i = 1, 2$, be two populations and let θ_i be a parameter connected with P_i. Let $t_i(n)$ be statistics (of finite variance) based on samples of size n from P_i and such that $\mathcal{E}t_i(n) = \theta_i$. Samples of sizes n_i from P_i yield the unbiased estimate $t_1(n_1) - t_2(n_2)$ of $\theta_1 - \theta_2$. The total sample size $N = n_1 + n_2$ being prescribed, it is desired to partition N so as to minimize the variance of $t_1(n_1) - t_2(n_2)$. When the variances of the $t_i(n)$ are unknown, a two-stage sampling procedure is utilized. Some particular investigations of such problems have been made by others (e.g., J. Putter), but this paper considers the asymptotic behavior (as $N \to \infty$) under general conditions, and also the situation for finite N in special cases. (This work was supported by the U. S. Air Force under contract AF 18(600)-83.)

18. Estimation of the Location Parameter in the Structural Problem of Neyman. T. A. Jeeves, University of California, Berkeley.

Consider a pair of random variables (X, Y) having the following structure: (i) $X = \xi + U$, $Y = \eta + V$; (ii) (ξ, η) are random variables such that $\xi \cos \theta + \eta \sin \theta = p$ for certain constants θ and p $(-\pi/2 < \theta \leqq \pi/2)$; (iii) (U, V) are independent of (ξ, η). Let $\{X_m, Y_m\}$ be a sequence of random variables such that each pair has the above structure and is independent of every other pair. The basic problem is to use this sequence to construct a pair of statistics $\hat{\theta}_N$ and $\hat{p}_N$ which will converge in some sense to θ and p, respectively. If $U = U_1 + U_2$ and $V = V_1 + V_2$ with (U_1, V_1) jointly normal and independent of (U_2, V_2) and U_2 independent of V_2, then under the assumption that ξ and η are not both normal, Neyman ("Existence of Consistent Estimates of the Directional Parameter in a Linear Structural Relation Between Two Variables," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 497–512) has given a consistent estimate of θ^* $(\theta^* = \theta$ if $\theta \neq \pi/2$, $\theta^* = 0$ if $\theta = \pi/2)$. To date no estimates of p have been given. In fact, even assuming the means of (U_1, V_1) known, without further restrictions on (U_2, V_2), p is not identifiable and hence no consistent estimate exists. Using the sample characteristic function, estimates $(\hat{\theta}_N, \hat{p}_N)$ have been obtained which converge almost surely to (θ, p) under the assumption that U_2 and V_2 have symmetric distributions. In a similar manner, estimates have been obtained for the case in which the first moments of U_2 and V_2 exist and are known.

19. On the Distribution of the Sum of the Roots of a Determinantal Equation. K. C. S. Pillai, University of North Carolina.

In four different situations of testing hypotheses relating to p-variate normal populations we run into the roots (all nonnegative) of the determinantal equation in θ: $| S_1 - \theta S_2 | = 0$ where $S_1(p \times p)$, $S_2(p \times p)$ are sample matrices such that almost everywhere S_1 is at least positive semidefinite of rank q $(\leqq p)$ and $S_2 - S_1$ (and hence necessarily also S_2) is positive definite. Under the null hypothesis, the joint distribution of the q positive roots $\theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_q$ is well known. Starting from this distribution, the successive moments of the sum of the q roots, s_θ, say, have been studied by means of a recurrence relation. The lower order moments indicate that the distribution of s_θ can be approximated by a Beta function of the form: const. $s_\theta^{q[m+\frac{1}{2}(q+1)]-1} (1 - s_\theta/q)^{q[n+\frac{1}{2}(q+1)]-1}$ $(0 \leqq s_\theta \leqq q)$. For small values of q the approximation is satisfactory if $m + n \geqq 30$ and for large values of n this distribution can be further approximated by a Gamma function with $q[m + \frac{1}{2}(q + 1)]$ degrees of freedom. This result has been established in two different ways, namely, using (1) the distribution of s_θ given above and (2) the method of characteristic function on an asymptotic joint distribution of the roots. T. W. Anderson following P. L. Hsu has obtained the asymptotic Gamma function distribution by another method.

20. On a Problem in Multivariate Regression. Thomas S. Ferguson, University of California, Berkeley.

Consider s random variables $\xi_1, \cdots, \xi_s$ and $n + 1$ random variables $\eta_0, \eta_1, \cdots, \eta_n$ such that the η_j are independent of the ξ_k and also independent among themselves, but the ξ_k are not necessarily independent among themselves. We assume that the ξ_k and η_j are nondegenerate and that all have finite first moments which we assume to be zero. Let $X_j = \sum_{k=1}^{s} a_{jk}\xi_k + \eta_j$ for $j = 0, 1, \cdots, n$ where the a_{jk} are arbitrary constants. In the case $n = 1$

the following result is obtained. *Theorem 1: In order that the regression of* X_0 *on* X_1 *be a linear function of* X_1 *irrespective of the values of the constants* a_{jk}, *it is necessary and sufficient that the characteristic functions of* η_1 *and* $(\xi_1, \cdots, \xi_s)$ *be of the form* $\exp\{-K \mid u \mid^{\nu}\}$ *and* $\exp\{-\rho^{\nu} g(v_1/\rho, \cdots, v_s/\rho)\}$ *respectively, where* K *and* ν *are constants,* $K > 0, 1 < \nu \leqq 2, \rho = \sqrt{v_1^2 +, \cdots, + v_s^2}$, *and* g *is an arbitrary real function such that* $\exp\{-\rho^{\nu} g(v_1/\rho, \cdots, v_s/\rho)\}$ *is the joint characteristic function of* s *nondegenerate random variables with zero means.* The general result for $n > 1$ is *Theorem 2: In order that the regression of* X_0 *on* $X_1, X_2, \cdots, X_n$ *be linear in* $X_1, X_2, \cdots, X_n$ *irrespective of the values of the constants* a_{jk}, *it is necessary and sufficient that each* η_j *be normal and that* $\xi_1, \cdots . \xi_s$ *have a multivariate normal distribution.* This paper extends the results of E. Fix ("Distributions which lead to linear regressions," *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1949).

21. On the Problem of Construction of Orthogonal Arrays. Esther Seiden, Universty of Chicago.

The remarks made by O. Kempthorne, (*Biometrika*, Vol. 34 (1947)) and K. A. Brownlee and P. K. Loraine ("The relationship between finite groups and completely orthogonal squares, cubes and hyper-cubes," *Biometrika*, Vol. 35 (1948), pp. 277–282) regarding some of the multifactorial designs constructed by Plackett and Burman can be extended to all of them. In order to avoid confounding of main effects with first order interactions, only arrays of strength at least 3 should be used. It is shown that all the designs of Plackett and Burman, in which each factor takes on only two levels, form a scheme leading to the construction of orthogonal arrays of strength 3 with the maximum possible numbers of constraints. An orthogonal array (36, 13, 3, 2) is constructed. It is known that the upper bound for the number of constraints is in this case 16. The method of construction used could not lead to a number of constraints greater than 13, but it is not known whether one would not do better using another one.

22. The Joint Distribution of *n* Successive Amplitudes. (Preliminary Report.) W. C. Hoffman, U. S. Navy Electronics Laboratory, San Diego.

The joint probability density function for two values of the output $R(t) = \{X^2(t) + Y^2(t)\}^{\frac{1}{2}}$ of a linear detector (Lawson and Uhlenbeck, *Threshold Signals*, McGraw-Hill Book Co., 1950, p. 61, equation (72)) is generalized to the case of n such random variables, assuming a multivariate normal distribution for the input signals. The derivation depends in an essential manner on the following properties of elements of the inverse covariance matrix Λ_n^{-1}: $\lambda^{2j-1,2k-1} = \lambda^{2j,2k}$, $(j, k = 1, 2, \cdots, n)$; $\lambda^{2j-1,2k} = -\lambda^{2j,2k-1}$ $(j \neq k)$; $\lambda^{2j-1,2j} = \lambda^{2j,2j-1} = 0$. The joint probability density function for the n-dimensional case has the form $f(r_1, \cdots, r_n) = |\Lambda_n|^{-\frac{1}{2}} r_1, \cdots, r_n \exp\{-\frac{1}{2}\sum_{j=1}^{n} \lambda^{2j-1,2j-1} r_j^2\} Q(r_1, \cdots, r_n; \Gamma)$, where Λ_n is the $2n \times 2n$ covariance matrix of the input, Γ is the symmetric matrix (γ_{jk}) with $\gamma_{jk} = \{(\lambda^{2j-1,2k-1})^2 + (\lambda^{2j-1,2k})^2\}^{\frac{1}{2}}$, and Q is an infinite series each of whose terms consists of products of modified Bessel functions of the first kind multiplied by the cosine of a weighted sum of the parameters $\varphi_{jk} = \text{Arc tan } (\lambda^{2j-1,2k}/\lambda^{2j-1,2k-1})$. The subscripts of the Bessel functions range over all nonnegative integers but must satisfy certain linear relations.

23. Simultaneous Tests for Regression Coefficients by the Two Stage Procedure. (Preliminary Report.) Manindra nath Ghosh, University of North Carolina.

In setting up a prediction equation of the form $E(\omega) = \alpha + \beta x + \gamma y + \delta z$, the tests of significance of the hypothesis $H_1: \beta = 0$, $H_2: \gamma = 0$, $H_3: \delta = 0$, by the usual method are not independent. Instead of combining these hypotheses and using an F-test, one would prefer to make simultaneous decisions regarding the hypothesis before setting up the final prediction formula. The methods developed by Scheffé-Tukey-Bose-Roy of simultaneous confidence intervals have been employed for this purpose and a two-stage procedure along the lines of Stein ("A two-sample test for a linear hypothesis whose power is independent of the variance," *Ann. Math. Stat.*, Vol. 16 (1945), pp. 243–258) has been developed to keep the probability of a wrong judgment regarding the hypotheses H_1: $|\beta| > \beta_0$, H_2: $|\gamma| > \gamma_0$, H_3: $|\delta| > \delta_0$, less than α per cent, where β_0, γ_0, δ_0 depend upon the relative cost of measuring the variables, and the variables x, y, z can be controlled for the purpose of the experiment.

24. Optimum Sample Size for Choosing the Largest of (k + 1) Parameters from (k + 1) Otherwise Identically Distributed Populations. PAUL N. SOMERVILLE, University of North Carolina.

Assume we have $(k + 1)$ populations, identically distributed except for unknown parameters $a_0 \geqq a_i \geqq, \cdots, \geqq a_k$). Let it be required to take a preliminary sample of size $(k + 1)n$ with the object of deciding which population should be used for a further sample of size N. Let $W(a_i, a_0)$ be the loss involved in choosing the population with parameter a_i where $W(a_i, a_0) \geqq 0$, $W(a_0, a_0) = 0$. Let $C(n)$ be the cost of taking a preliminary sample. Then it is shown that under certain conditions the maximum expected loss over all values of a_i, $i = 0, 1, 2, \cdots, k$, occurs where $a_1 = a_2 = \cdots = a_k$. This enables us to find the maximum expected loss, which can then be minimized with respect to the preliminary sample size.

25. Necessary Conditions for the Existence of Partially Balanced Incomplete Block Designs with Two Associate Classes. W. S. CONNOR AND W. H. CLATWORTHY, National Bureau of Standards.

For a partially balanced incomplete block design with two associate classes and with parameters $v, b, r, k, n_1, n_2, \lambda_1, \lambda_2$, and p^i_{jk} $(i, j, k = 1, 2)$, the following theorem has been proved. If (i) $v > b$, then it is necessary that (a) Δ be a perfect square and (b) either $r - r_1 = 0$, or $r - r_2 = 0$; (ii) $v = b$ and v is even, then it is necessary that (a) Δ be a perfect square and (b) $r - r_u$ be a perfect square when α_u is odd $(u = 1, 2)$; (iii) $v = b$, v is of the form $4t + 3$ $(t = 0, 1, 2, \cdots)$, and Δ is not a perfect square, then it is necessary that $(r - r_1)(r - r_2)$ be a perfect square, and (iv) $v < b$ and v is even, then it is necessary that Δ be a perfect square where $r_u = \frac{1}{2}[(\lambda_1 - \lambda_2)(-\gamma + (-)^u\sqrt{\Delta}) + (\lambda_1 + \lambda_2)]$, $(u = 1, 2)$, $\gamma = p^2_{12} - p^1_{12}$, $\Delta = \gamma^2 + 2\beta + 1$, $\beta = p^1_{12} + p^2_{12}$, and α_1 and α_2 are nonnegative integers such that $\alpha_1 + \alpha_2 = v - 1$. Examples are given of sets of parameters which fail to satisfy these conditions.

26. Estimation in Truncated Bivariate Normal Distributions. (Preliminary Report.) A. C. COHEN, JR., University of Georgia.

Maximum likelihood estimators of the parameters of a bivariate normal population are developed for samples which are subjected to a truncation on one of the variates at known terminals. Both single and double truncations with the number of missing (unmeasured) observations either known or unknown are considered. Asymptotic variances of the estimates are obtained from the likelihood information matrices.

27. On a Class of Optimum Linear Predictors. R. F. DRENICK AND P. NESBEDA, R. C. A. Victor Division, Camden, New Jersey.

Prediction is the problem of projecting into the future a set of observed data in order to obtain an estimate for future observable data. For optimum prediction one assigns, through some considerations which are not part of the method, a loss function representing the penalty for error. An optimum prediction procedure is the one which minimizes, in the long run, this penalty. N. Wiener pointed out that the optimum mean square predictor is linear if the interference affecting the observations has Gaussian probability distribution. By using a method of estimation due to Pitman ("Estimation of the location and scale parameters of a continuous population of any given form," *Biometrika*, Vol. 30 (1939), pp. 391–421) the authors show that the class of linear predictors is characterized by the Gaussian probability distribution and by a loss function more general than r.m.s., namely, one which is symmetric and has continuous derivatives. Most of the loss functions of practical interest are in this category. Furthermore any such loss function leads to the same linear predictor X_p which has also the property: $P(| X_p - x | \leqq k) = \max$ for all $k > 0$, x being the true value. (Work sponsored by the Bureau of Aeronautics.)

28. **Multiple Range Tests and the Multiple Comparisons Test. (Preliminary Report.)** D. B. DUNCAN, Virginia Polytechnic Institute.

Several methods are available for testing differences between treatments in an analysis of variance. The two considered most satisfactory are one by Newman (1952) and Keuls (1952) and the Multiple Comparisons Test by Duncan (1951). Both employ repeated homogeneity tests. The Newman-Keuls test is simpler because it uses repeated range tests instead of F tests as used by the Multiple Comparisons Test. The latter is generally more sensitive owing partly to this reason but mostly to the relaxation of the significance levels of some of the tests considered to be of diminished importance. This paper presents: a new Multiple Range Test which achieves the simplicity of the Newman-Keuls test by using range tests and most of the sensitivity of the Multiple Comparisons Test by using the special significance levels, and an improved set of application rules for the Multiple Comparisons Test. Each of these is recommended for use depending on the relative means for simplicity or sensitivity. The special system of significance levels is discussed in some detail. The author is indebted to W. Beyer in the determination of significance ranges for the new test which is still in progress. (Research under contract No. DA-36-034-ORD-1084 (RD) with the Office of Ordnance Research, Department of the Army.)

29. **A Property of the Normal Distribution Related to a Theorem of S. Bernstein. (Preliminary Report.)** EUGENE LUKACS AND EDGAR P. KING, National Bureau of Standards.

The following theorem is proved. Let $x_1, x_2, \cdots, x_n$ be n independently (but not necessarily identically) distributed random variables and assume that the nth moment of each x_i $(i = 1, 2, \cdots, n)$ exists. The necessary and sufficient conditions for the existence of two statistically independent linear forms $y_1 = \sum_{s=1}^{n} a_s x_s$ and $y_2 = \sum_{s=1}^{n} b_s x_s [a_s \neq 0; b_s \neq 0; a_s/b_s \neq a_t/b_t$ for $s \neq t; s, t = 1, 2, \cdots, n]$ are that each random variable be normally distributed and that $\sum_{s=1}^{n} a_s b_s \sigma_s^2 = 0$. For $n = 2$ this reduces to a theorem of S. Bernstein ("Sur une propriété caractéristique de la loi de Gauss," *Trans. Leningrad Polytechnic Institute*, (1941), pp. 21–22).

30. **An Asymptotically Efficient Formula for Estimating Parameters from Grouped Data. (Preliminary Report.)** M. C. K. TWEEDIE, Virginia Polytechnic Institute.

Suppose that, in a sample from a discrete or grouped distribution, x_i observations fall in group i, whose probability is $\pi_i(\theta_1, \cdots \theta_R)$, with $i = 1$ to N. The total sample size is $n, = \sum_{i=1}^{N} (x_i)$, and may be constant or determined sequentially. Write $G = \sum_{i=1}^{N} x_i g(X_i)$, where $X_i = n\pi_i(T_1, \cdots T_R)/x_i$ and $g(X)$ is an arbitrary function of X approximately quadratic near $X = 1$. An estimate of $(\theta_1, \cdots, \theta_R)$ may be obtained (usually by differentiation) as the set of values of $(T_1, \cdots, T_R)$ for which G is least (or greatest, depending on g). Under normal conditions of regularity, with large samples the consequent estimates are effectively consistent and have minimum variance, and are, in Neyman's terminology, BAN. (Cf. also H. Cramér *Mathematical Methods of Statistics*, Princeton University Press 1946, §30.3). This formulation includes some important methods precisely, such as maximum likelihood $[g(X) = \log X]$ and minimum χ^2 $[g(X) = (1 - X)^2/X]$. Thus, as Fisher (*Statistical Methods for Research Workers*, Chapter IX) has shown, the recombination fraction can be estimated efficiently from F_2 genetical data by both these methods, and also by a product-ratio formula (given by $g(X) = X \log X$). In this problem an efficient *linear* estimation equation results from using $g(X) = (1 - X)^2$, equivalent to minimizing a modification of χ^2 which has the observed frequencies in the denominators.