

## ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Pittsburgh, Pennsylvania Meeting of the Institute, March 19-21, 1959.)

### 1. **Mathematical Problems Associated with Measurements Made by Matching with Known Standards.** W. S. CONNOR AND N. C. SEVERO, National Bureau of Standards.

The process of matching can be thought of as measuring the unknown true value of some characteristic of an object, and of comparing the measured value with a series of equally spaced standard values. Let the true value of the unknown be 0, the measured value be  $Y$ , the standard value closest to  $Y$  be  $X$ , and this closest standard value minus the measured value be  $Z$ . Then  $X = Y + Z$ . Further, let  $Y$  be a random variable which is normally distributed with mean zero and variance  $\sigma^2$ , let  $Z$  be uniformly distributed on the interval  $(-a, a)$ , and let  $Y$  and  $Z$  be statistically independent. The probability  $P_A$  that an unknown object will be assigned the standard value which is closest to its true value is shown to be  $P_A = \Pr\{-a < X \leq a\} = [\Phi(2a/\sigma) - \Phi(-2a/\sigma)] + \sigma/a[\varphi(2a/\sigma) - \varphi(0)]$  where

$$\Phi(t) = \int_{-\infty}^t \varphi(\xi) d\xi \equiv \int_{-\infty}^t (1/\sqrt{2\pi})e^{-\xi^2/2} d\xi.$$

The solution is obtained by convenient use of the usual convolution formula for representing the probability density function for the sum of two statistically independent random variables. By similar methods, the probability  $P_B$  that two independent measurements on the same object will result in the same assigned standard value, regardless of whether it is the closest standard value, is shown to be

$$P_B = [\Phi(\sqrt{2}a/\sigma) - \Phi(-\sqrt{2}a/\sigma)] + \sqrt{2}\sigma/a[\varphi(\sqrt{2}a/\sigma) - \varphi(0)].$$

### 2. **The First and Second Moment Structure of the Maximum Likelihood Estimators of the Parameters of a Multivariate Normal Distribution with Double Samples.** JACK NADLER, Bell Telephone Laboratories, Whippany, New Jersey.

Let  $x$  be a  $p$ -dimensional, non-singular, normally distributed, vector random variable whose coordinates are partitioned to form a  $q$ -dimensional vector random variable  $y$  and an  $r$ -dimensional vector random variable  $z$  ( $p = q + r$ ;  $q \geq 1$ ;  $r \geq 1$ ). It is assumed that a random sample of  $n$  observations is taken from the distribution of  $x$  and an independent random sample of  $N - n$  observations is taken from the distribution of  $y$ . The first and second moments of the maximum likelihood estimators of the parameters of the distribution of  $x$  are presented for the combined sample.

### 3. **Some Properties of Stirling's Numbers of the Second Kind.** JOHN L. BAGG, Florida State University.

In the course of investigation of the asymptotic behavior of a probability distribution discovered by Leo Katz and James Powell (*Proc. Am. Math. Soc.*, Vol 5 (1954), pp. 621-626), a relationship was noted between one of Sukhatme's Bipartitional Functions and Stirling's Numbers of the Second Kind. Jordan (*Calculus of Finite Differences*, Chelsea Pub. Co., New York, 1947) deals with many properties of Stirling's Numbers of the Second Kind. Denote the sum on the superscript of these numbers, for fixed subscript  $n$ , by  $f(n)$ . A general formula for the  $m$ th difference of  $f(n)$  is proved.

**4. Minimax Solutions to Trichotomies.** LONNIE L. LASMAN, Florida State University.

This paper presents an explicit method of obtaining minimax solutions to a special class of problems. Let  $x$  be a real-valued random variable with probability density  $f(x, \theta) = b(\theta)c^{2x}g(x)$ , a continuous function of the parameter  $\theta$ , and consider the problem of accepting one out of the three hypotheses  $H_i: \theta = \theta_i, i = 1, 2, 3$ . Given three possible decisions to make,  $a_1, a_2$ , or  $a_3$ , where  $a_i$  means accept  $H_i$ , and a set of losses  $w_{ij}$ , the loss incurred from taking action  $a_i$  when  $H_j$  is the correct hypothesis, then it is shown that if a pair  $(x', x'')$  with  $x' < x''$  exist such that the risks to the statistician from a choice of a test procedure  $\phi$  are the same regardless of the value of  $\theta$ , the minimax solution to the problem is a form of the Bayes solution. Denote such a solution by  $\phi^*$ . In such a case there is a common risk  $V$  regardless of which hypothesis is true. It is then shown that there is a least favorable *a priori* distribution  $\xi$  on  $\theta$  such that the risk when nature uses  $\xi^*$  and the statistician uses  $\phi$  is never less than the risk using  $\xi^*$  and  $\phi^*$ ; this latter risk is also shown to be  $V$ , and thus the minimax solution is established. An example using the normal curve is given.

**5. The Distribution of the Number of Successes in a Sequence of Dependent Trials.** K. R. GABRIEL, University of North Carolina.

A sequence of dependent trials is considered which has the properties of a Markov Chain with two ergodic states and transition probabilities: failure to success  $P(1 - d)$ , success to success  $P(1 - d) + d \cdots P$  being the stationary probability of success. The exact distribution of the number of successes  $S$  in  $n$  trials is derived, and the first four moments obtained exactly. Approximate formulae are suggested as follows:

$$K_2 = nP,$$

$$K_3 = nP(1 - P)(1 + d)/(1 - d), \gamma_3 = (1 - 2P)(1 + 4d + d^2)/(nP(1 - P))^{3/2}(1 + d)(1 - d^2)^{1/2}$$

$$\gamma_4 = (1 - 6P + 6P^2)(1 + 10d + d^2)/nP(1 - P)(1 - d^2).$$

From the theory of recurrent events it is shown that  $S$  is asymptotically normally distributed with mean and variance as above. Numerical computations of exact and approximate cumulants for selected values of  $n, P$  and  $d$  are presented. These give an idea of the characteristics of the distribution and the rapidity of its approach to normality. The distribution is illustrated by an application to the number of rainy days in a month, with  $n = 31, P = .473$  and  $d = .381$ , and the normal approximation is seen to be very good.

**6. Values of Games with Moves in  $[0, 1]$ .** (Preliminary report) MARTIN FOX, University of California, Berkeley.

Consider a zero sum, two person game, in which the players move alternately by choosing points in  $[0, 1]$ . The game will end when a total of  $n$  moves have been made. Let  $X_i$  be the point chosen on the  $i$ th move (by player I if  $i$  is odd, by player II if  $i$  is even). Let the payoff be  $f(X_1, X_2, \dots, X_n)$  where  $f$  is continuous. When the  $i$ th move has been made, the player who makes the  $(i + 1)$ st move will observe  $\varphi_i(X_1, \dots, X_i)$  ( $i = 1, \dots, n - 1$ ). If the  $\varphi_i$  are all constant, we have the case of no information so that the game has a value according to Ville's minimax theorem. If the  $\varphi_i$  are all one-to-one, we have the case of perfect information so the game has a value. In the present paper an example is presented to show that for  $n \geq 3$  these games do not always have values. For the case  $n = 2$  it is proved that these games always have values. Existence of a value is proved for a special case with  $n = 3$ . The author is seeking additional conditions guaranteeing values of these games for arbitrary  $n$ .

**7. Explicit Results for the Dam with Poisson Input.** JOSEPH M. GANI AND N. U. PRABHU, Columbia University and Karnatak University.

Let  $Z(t)$  be the content at time  $0 \leq t < \infty$  of an infinite dam, fed by Poisson inputs of magnitude  $h$  with parameter  $\lambda$ , and subject to a steady continuous release ceasing when  $Z(t) = 0$ . The distribution function  $F(z, t)$  of  $Z(t)$  then satisfies the difference-differential equation  $\partial F(z, t)/\partial t - \partial F(z, t)/\partial z = -\lambda\{F(z, t) - F(z - h, t)\}$  ( $0 \leq z < \infty$ ). This particular case of Takács's integro-differential equation for the d.f. of the waiting time in a single-server queue yields an explicit solution for  $F(z, t)$ . The probability of first emptiness of the dam at time  $T = z_0 + rh$  ( $r = 0, 1, 2, \dots$ ), starting with  $Z(0) = z_0$ , is given by  $g(z_0, T) = e^{-\lambda T} \lambda^r z_0 T^{r-1}/r!$ ; from this, the probability of emptiness at time  $t$  (not necessarily for the first time) may be derived as  $F(0, t) = e^{-\lambda t} \sum_{j=0}^{\lfloor (t-z_0)/h \rfloor} \lambda^j (t - jh)^{j-1}/j!$ . Solving the difference-differential equation directly, the d.f.  $F(z, t)$  is finally found to be

$$F(z, t) = \sum_{r=0}^{\lfloor z/h \rfloor} e^{-\lambda(rh-z)} \{\lambda(rh - z)\}^r F(0, z + t - rh)/r!$$

**8. Some Stochastic Processes with Application to Counter Models.** RONALD PYKE, Columbia University.

Let  $\{Y_n: n > -\infty\}$  be a Renewal process with common distribution function  $H$ , and let  $\{t_j, j > -\infty\}$  be successive time points of discontinuity for a doubly infinite Poisson process. Let  $f$  be any real-valued function defined on  $R_2$  for which  $\int_{R_2} |f(t, y)| dH(y) dt < \infty$ . The processes  $\{\eta(t): t > 0\}$  and  $\{\eta^*(t): -\infty < t < \infty\}$  determined by

$$\eta(t) = \sum_{0 \leq t_j \leq t} f(t - t_j, Y_j), \eta^*(t) = \sum_{-\infty < t_j < \infty} f(t - t_j, Y_j)$$

are studied and their one-dimensional characteristic functions obtained. For given  $0 < a \leq b$ , either of these processes is said to be in state A at time  $t$  if  $\eta(t) \leq b$  and if the process has been less than  $a$  sometime since it last exceeded  $b$ . The expected number of counts (i.e., transitions from state A to state B), as well as the expected time in state A, during  $(0, t)$  are studied. For one case, approximations to the expected number of counts are obtained by the method of steepest descents.

**9. Use of Series Expansion in Estimation Problems for Distributions Involving More Than One Parameter.** (Preliminary report) Y. S. SATHE, University of Alberta. (By title)

Guttman (*Biometrika*, Vol. 45 (1958), pp. 565-567) has given a method of determining an unbiased minimum variance estimator without taking conditional expectation with respect to a sufficient statistics under certain regularity conditions. The same method is extended for distributions depending on more than one parameter. If  $t_1, t_2, \dots, t_l$  are sufficient statistics which assume non-negative integer values with probability

$$k_t \phi_1^{t_1} \cdot \phi_2^{t_2} \cdots \phi_l^{t_l} m(\theta_1, \theta_2, \dots, \theta_l)$$

where  $\phi_i^t$  are functions of  $\theta_i^t$  and  $k_t$  is a function of  $t_i^t$  and if

$$G(\theta_1, \theta_2, \dots, \theta_l) = g(\theta_1, \theta_2, \dots, \theta_l)/m(\theta_1, \theta_2, \dots, \theta_l)$$

can be expanded as a power series of the form  $\sum A_t \phi_1^{t_1} \cdot \phi_2^{t_2} \cdots \phi_l^{t_l}$  where  $A_t$  is a function of  $t_i^t$ , then an unbiased minimum variance estimator of  $g(\theta_1, \theta_2, \dots, \theta_l)$  is  $A_t/k_t$  ( $k_t \neq 0$ ). This method can be used to obtain unbiased minimum variance estimators for certain classes of distributions which involve more than one parameter for e.g. multinomial distribution.

**10. A Comparison of the Effectiveness of Tournaments.** W. A. GLENN, Virginia Polytechnic Institute.

Round robin, replicated knock-out and double elimination tournaments (in which players are eliminated after two losses) are investigated for their effectiveness in selecting the best of four players. Denoting the probability that player  $i$  defeats player  $j$  by  $\pi_{ij}$ , it is assumed that these parameters are constants satisfying a general set of inequality relations. A best player is here defined as one having a probability greater than  $\frac{1}{3}$  of defeating each of the others. It is further assumed that each game results in the selection of a winner, so that  $\pi_{ji} = 1 - \pi_{ij}$ . Since in some tournaments two or three players may receive the same total number of wins, a play-off may be required for the determination of an ultimate winner. The criteria proposed for the comparison are the probability that the best player wins (after play-off if necessary) and the expected number of games required. For general values of the parameters expressions are derived for the evaluation of the criteria, and comparisons are made on the basis of series of assigned parameter values. The special case in which all but one of the players are of equal strength is considered in detail. The possibility of extending this investigation to cases involving a larger number of players is discussed.

**11. Application of the Geometry of Quadrics in Finite Projective Space to the Construction of PBIB Designs.** R. C. BOSE AND D. K. RAY-CHAUDHURI, University of North Carolina.

The geometry of quadrics in finite projective hyperspace has been applied to construct some series of PBIB designs, with two or three associate classes, including some new designs with parameters in the practically useful range. Let (C) and (D) be two classes of linear spaces such that spaces of a given class stand in the same geometrical relation to a quadric  $Q$  in  $PG(n, s)$ ,  $s = p^m$  where  $p$  is prime. Then in many instances, the incidence relationship of (C) and (D), provides a PBIB design. For example, if we take (C) as the class of points on a non-degenerate quadric  $Q$ , and (D) as the class of lines contained in  $Q$ , we get a PBIB design with the following parameters:

$$\begin{aligned} v &= N(0, n), & b &= N(1, n), & r &= N(0, n-2), & k &= s+1, & \lambda_1 &= 1, & \lambda_2 &= 0, \\ n_1 &= sN(0, n-2), & n_2 &= N(0, n) - sN(0, n-2) - 1, & p_{11}^1 &= (s-1) + s^2N(0, n-4), \\ & & & & & & p_{11}^2 &= N(0, n-2), \end{aligned}$$

where  $N(p, n)$  denotes the number of  $p$ -flats in  $Q$  and is given by the formulae,

$$(i) \quad N(p, n) = \prod_{r=0}^p [(s^{n-2p+2r} - 1)/(s^{p+1-r} - 1)],$$

if  $n = 2k$ ,  $p \leq k-1$ ;

$$(ii) \quad N(p, n) = \prod_{r=0}^p [(s^{n-2p+2r} - s^{k-p+r} + s^{k-p+r-1} - 1)/(s^{p+1-r} - 1)],$$

if  $n = 2k-1$ ,  $p \leq k-2$  and  $Q$  is elliptic;

$$(iii) \quad N(p, n) = \prod_{r=0}^p [(s^{n-2p+2r} + s^{k-p+r} - s^{k-p+r-1} - 1)/(s^{p+1-r} - 1)],$$

if  $n = 2k-1$ ,  $p \leq k-1$  and  $Q$  is hyperbolic.

**12. Some New Cases of the Packing Problem in Finite Projective Space with Applications to Fractionally Replicated Designs.** R. C. BOSE, University of North Carolina.

The general packing problem for the finite projective space  $PG(r-1, p^m)$  of  $r-1$  dimensions based on the finite field with  $s = p^m$  elements may be stated as follows: Find the maximum number of points which can be chosen in  $PG(r-1, p^m)$  so that no  $t$  lie on a linear space of dimensions  $t-2$  or less. This number may be denoted by  $n = n_t(r, s)$ , and the associated set of points may be said to give a tight packing of the  $t$ th order for the space. If  $t = 2u$  or  $2u+1$  the coordinates of the points may be used to obtain a  $1/s^{n-t}$  fraction of the  $s^n$  treatments in a factorial experiment with  $n$  factors, so that no  $u$ -factor or lower interaction, is aliased with a  $u$ -factor or lower interaction. In this paper it is shown that  $n_5(6, 3) = 12$ ,  $n_4(5, 3) = 11$ . The associated set of points in the first case is the set of points common to the three quadrics  $Q_1 = 2x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 = 0$ ,  $Q_2 = x_1x_2 + 2x_1x_5 + 2x_1x_6 + x_2x_5 + x_2x_6 + x_5x_6 = 0$ ,  $Q_3 = x_3x_4 + 2x_3x_5 + x_3x_6 + x_4x_5 + 2x_4x_6 + 2x_5x_6 = 0$ . Fractionally replicated designs  $1/36 \times 3^{12}$  and  $1/35 \times 3^{11}$  in which no main effect or two factor interaction is aliased with a main effect or two factor interaction follow.

**13. A Necessary Condition for Existence of a Regular and Symmetrical p.b.i.b. Design of Triangular Type.** J. OGAWA, University of North Carolina.

A necessary condition for existence of a symmetrical b.i.b. design in terms of the Hasse-Minkowski's  $p$ -invariant was obtained by S. S. Shrikhande. Similar necessary conditions for regular symmetrical p.b.i.b. design of group divisible type and for of  $L_2$  type were obtained by R. C. Bose, W. S. Connor and S. S. Shrikhande, respectively. The purpose of this note is to give a similar necessary condition for a regular symmetrical p.b.i.b. design of triangular type.

**14. Use of Partially Balanced Block Designs with Three Associate Classes for Confounded, Asymmetrical Factorial Arrangements.** (Preliminary report) BADRIG M. KURKJIAN, Diamond Ordnance Fuze Laboratories. (By title)

The results and technique, presented by Zelen (*Ann. Math. Stat.* 29 (1958) pp. 22-40) for the case of GD designs, are extended to treat two cases involving partially balanced incomplete block designs with three associate classes when used in conjunction with factorial experiments. The two PBIB designs considered are those that result by (1) replacing each treatment of a BIB by another, complete BIB design and (2) replacing each treatment of a GD design by  $n$  treatments. Vartak (*Ann. Math. Stat.* 26 (1955) pp. 420-438) has shown that each of these designs is at most a PBIB with three associate classes. For each of these designs, the solutions of the reduced normal equations for the treatment estimates are found. With respect to the factorial aspect of the problem, the variances and covariances of the various main effects and interaction terms have been derived for the class of design above. It is shown that these can be written as Kroneker products of matrices which lead directly to the appropriate sums of squares associated with the analysis of variance. In addition, the inter-block analysis is worked out.

**15. Best Linear Estimates by Order Statistics of the Parameters of a Model for Failure Data.** ANDRE G. LAURENT AND ELDON RIEHM, Wayne State University and Bendix Aviation Research Laboratory.

The paper presents tables of 1) the order statistics covariance matrix for samples of size  $n = 1$  to 10 drawn from the population with survival function  $S(t) = \exp[1 + t - \exp(t)]$ ,

where  $t = X/E_0$  or  $t = (X - X_0)/\tau$ ; 2) the minimum variance linear unbiased estimates of the parameters  $E_0$ ,  $X_0$ ,  $\tau$ , based on order statistics for singly censored samples of size 1 to 10 and their variances and covariances. Several aspects (failures, aging, waiting time) of the problem of "estimation cost" versus "precision" are considered and different estimates are compared from the viewpoint of efficiency.

**16. On Testability in Normal ANOVA and MANOVA with All "Fixed Effects."**

S. N. ROY AND J. ROY, University of North Carolina and Indian Statistical Institute.

With the notation of the above paper it can be proved that, for standard form I, Rank  $A \leq \text{Rank} \begin{bmatrix} A \\ C \end{bmatrix} \leq \text{Rank } A + \text{Rank } C$  (both equalities being unattainable at the same time), and, for standard form II,  $0 \leq \text{Rank } A - \text{Rank } AB \leq m - \text{Rank } B$  (both equalities being unattainable at the same time), and, in any case, that what we are generally testing is an  $H_0^*$  such that  $H_0 \subseteq H_0^* \subseteq \text{Model}$ . In this paper it is shown that, for I, (i) if  $\text{Rank } A = \text{Rank} \begin{bmatrix} A \\ C \end{bmatrix} < \text{Rank } A + \text{Rank } C$ , then  $H_0 = H_0^* \subset \text{Model}$ , in which case  $H_0$  is said to be testable in the strong sense, (ii) if  $\text{Rank } A < \text{Rank} \begin{bmatrix} A \\ C \end{bmatrix} < \text{Rank } A + \text{Rank } C$ , then  $H_0 \subset H_0^* \subset \text{Model}$ , in which case  $H_0$  is said to be testable in the weak sense and (iii) if  $\text{Rank } A < \text{Rank} \begin{bmatrix} A \\ C \end{bmatrix} = \text{Rank } A + \text{Rank } C$ ; then  $H_0 \subseteq H_0^* = \text{Model}$ , in which case  $H_0$  is said to be untestable. Likewise, for II, it is shown that (i) if  $0 < \text{Rank } A - \text{Rank } AB = m - \text{Rank } B$ , then  $H_0 = H_0^* \subset \text{Model}$ , (ii) if  $0 < \text{Rank } A - \text{Rank } AB < m - \text{Rank } B$ , then  $H_0 \subset H_0^* \subset \text{Model}$  and (iii) if  $0 = \text{Rank } A - \text{Rank } AB < m - \text{Rank } B$ , then  $H_0 \subseteq H_0^* = \text{Model}$ .

**17. Contributions to Univariate and Multivariate Analysis of Variance with "Fixed Effects," Normal Error and "Random Effects" Not Necessarily Normal.** S. N. ROY AND WHITFIELD COBB, University of North Carolina and The Woman's College of the University of North Carolina (Greensboro).

For simplicity of discussion, a single response two factor experiment under certain broad classes of designs is considered and an additive model is postulated such that one factor goes with "fixed effects," and the other one with "random effects" characterizable in terms of a random sample from an unknown continuous distribution which is assumed to be approximated, in successive stages, by a two valued distribution with probabilities  $\frac{1}{2}$  and  $\frac{1}{2}$ , a three valued distribution with probabilities  $\frac{1}{3}$ ,  $\frac{1}{3}$  and  $\frac{1}{3}$ , and so on, the values in each being assumed to be unknown. The continuous variate is also assumed to be independent of the normal error. Under the model confidence bounds are obtained on the difference between the two values at the first stage, on the two consecutive differences (simultaneously) at the second stage, on the three consecutive differences (simultaneously) at the third stage, and so on. This is, of course, in addition to what is usually done for the "fixed effects." These techniques are then extended, first to the case of an experiment with a single response and more than two factors, and then to multiple response and multifactor experiments. In the latter case, as a step toward this, a generalization has had to be made of the notion of  $m$ -tiles of a univariate distribution to the case of multivariate distributions.

**18. Some Nonparametric Analogues of "Normal" ANOVA and MANOVA and of Studies in "Normal" Association.** S. N. ROY AND V. P. BHAPKAR, University of North Carolina.

In a multifactor multiresponse experiment where some of the responses are assumed to be continuous, some discrete, some categorical with an implied ranking (like good, fair, poor, etc.) and the rest purely categorical, with a similar division into four classes for the factors, a finite number of class intervals are assigned to each continuous response or factor and a probability model is postulated in terms of a product-multinomial distribution with unknown probabilities in the multidimensional cells and preassigned weights or scores along all "marginals," response or factor, that are structured, and, of course, no such scores along the "marginals" that are purely categorical. Under this model, hypotheses are posed analogous to (i) different kinds of hypotheses in "normal" model I ANOVA and MANOVA, including analysis of covariance and regression and (ii) different kinds of independence and regression relations in "normal" multivariate distributions. Large sample tests of such hypotheses are offered, and an indication is given as to how to obtain the asymptotic powers.

**19. On Moments of Order Statistics from Normal Populations.** Z. GOVINDARAJULU, University of Minnesota. (By title)

Let  $\bar{X}_{1/n} < \bar{X}_{2/n} \cdots < \bar{X}_{n/n}$  be the order statistics (os) from a sample of size  $n$  from the standard normal population. Contributions by Tippet (1925), Hastings, et al (1947), Jones (1948), Godwin (1949), Cole (1951), Rosser (1951), Ruben (1956), Bose and Gupta (1956), Teichroew (1956), Sarhan and Greenberg (1956), have been made to the *Problem of Moments of os from Normal Samples*. Exact values of low moments of os for sample size six or less and numerical values for sample size twenty or less are available. In this investigation simple recursion formulae among the first, second, and mixed (linear) moments have been derived. Certain identities among the moments which are true in general (that is without the use of normality) are also obtained. The above will enable one to extend Ruben's table of moments of the largest os in samples of size fifty or less to moments of all os in samples of size fifty or less. It is shown that it is sufficient to know one first moment when  $n$  is even and one second moment when  $n$  is odd, in order to solve for the first and second moments of all os from sample of size  $n$ , in terms of those for the preceding  $n$ . It is also shown that at most  $(n-4)/2$  mixed moments for even  $n$ , and  $(n-3)/2$  mixed ones together with one second moment for odd  $n$  are sufficient to solve for the product-moment matrix of the vector of ordered variables in sample of size  $n$ , in terms of the second and mixed moments of os for the preceding  $n$ .

**20. A Note on J. Roy's "Step-down Procedure in Multivariate Analysis."** V. P. BHAPKAR, University of North Carolina. (By title)

Test criteria in multivariate analysis are usually derived either from the  $\lambda$ -criterion or the largest and/or the smallest characteristic roots. Both of these can be regarded as special cases of the general "union-intersection" principle. An alternative procedure, called the "step-down" procedure, was used by S. N. Roy and R. E. Bargmann (*Ann. Math. Stat.* Vol. 29 (1958), 491-503) to test multiple independence of normal variables. This procedure was recently applied by J. Roy (*Ann. Math. Stat.* Vol. 29 (1958), pp. 1177-1187) to derive test criteria for a large class of hypotheses other than that of multiple independence. In this note, J. Roy's method has been used to test multiple independence of normal variables with means given by a general linear model. Simultaneous confidence bounds on an appropriate set of "deviation-parameters" are also obtained.

**21. On a Class of Problems in Multivariate Analysis of Variance.** S. N. ROY  
AND J. ROY, University of North Carolina and Indian Statistical Institute.  
(By title)

Assuming for a model that  $X = [x_1 \cdots x_n]$  is a set of  $n$  independent stochastic vectors such that  $x_i: N[E(x_i), \Sigma]$  ( $i = 1, 2, \dots, n$ ) and furthermore that  $E(X') = \begin{matrix} p \times n \\ p \times 1 & p \times p \end{matrix} \begin{matrix} A & \xi \\ n \times p & n \times m & m \times p \end{matrix}$ , where  $A$  is a known matrix given by the design and by what the experimental statisticians call the "model," and  $\xi$  is a matrix of unknown parameters, a hypothesis is posed in the form (to be called standard form II)  $H_0: \xi = \begin{matrix} m \times p & m \times k & k \times p \end{matrix} \begin{matrix} B & \eta \end{matrix}$ , where  $B$  is given, but  $\eta$  is unknown. For a hypothesis expressed in this form (which is convenient for a wide class of problems including that of linearity of regression) the matrices  $S^*$  and  $S$  due to the hypothesis and due to the error are computed as a preliminary to the construction of different alternative tests, one of them being that in terms of the largest root of  $S^*S^{-1}$ . The tie-up is then discussed between any test of an hypothesis formulated this way and the corresponding test of an equivalent hypothesis formulated in the form to be called standard form I  $H_0: C \begin{matrix} \xi \\ m \times p \end{matrix} = \begin{matrix} 0 \\ s \times p \end{matrix}$  where  $C$  is given but  $\xi$ , of course, is assumed to be unknown.

**22. A Note on Confidence Bounds Connected with ANOVA and MANOVA for Balanced and Partially Balanced Incomplete Block Designs.** V. P. BHAPKAR, University of North Carolina. (By title)

It is known that confidence bounds can be placed on the "deviation parameter" associated with the test of a linear hypothesis. This "deviation parameter" can be regarded as a measure of departure from the "total" hypothesis under consideration. It is also possible to make simultaneous confidence statements about "partial" deviation parameters which can be regarded as measures of departure from various "partials" of the "total" hypothesis. In this note, the hypothesis considered is that of equality of treatment effects (scalar effects for ANOVA and vector effects for MANOVA) in experimental designs. The ANOVA deviation parameter for BIBD turns out to be  $(\lambda v/k) \sum_{i=1}^v \xi_i^2$  where  $\xi_i = t_i - \bar{t}$ , or  $(\lambda v/k) \xi'_{v \times 1} \xi_{v \times 1}$  where  $\xi = t_{v \times 1} - \bar{t}_{v \times 1}$  [with  $t' = (\bar{t}, \dots, \bar{t})$ ], and the MANOVA deviation parameter  $Ch_{\max} [\xi'_{p \times v} \xi_{v \times p}]$ , where  $\xi_{v \times p} = T_{v \times p} - \bar{T}_{v \times p}$  [with  $\bar{T}'_{v \times p} = (\bar{t}, \bar{t}, \dots, \bar{t})$   $p$ ]. The partial ANOVA deviation parameters for BIBD are found to be  $(\lambda v/k) \sum_{i=1}^v \sum_{j=1}^1 \xi_{a_i}^2$  associated with the partial hypothesis  $ta_1 = ta_2 = \dots = ta_{v_1}$  where  $\xi_{a_i} = t_{a_i} - (\sum_{a_i/v_1} t_{a_i/v_1})$ , with corresponding forms for the partial MANOVA deviation parameters. The ANOVA deviation parameter for PBIBD is  $\sum_i \sum_j a_{ij} \xi_i \xi_j$  where  $a_{ii} = r(k-1)/k$  and  $a_{ij} = -\lambda m/k$  if the  $i$ th and  $j$ th treatments are  $m$ th associates with a corresponding form for MANOVA. In general, the ANOVA deviation parameter is  $\xi' C \xi$  where  $C$  is the matrix of coefficients of the adjusted normal equations with a corresponding form for the MANOVA.

**23. Sufficient Partitions for a Class of Coin-Tossing Problems.** (Preliminary report) T. V. NARAYANA, University of Alberta.

In the following experiment  $G$ , is considered (cf. also Blackwell and Girshick, *Theory of Games and Statistical Decisions*, p. 222): The probability of a coin falling head is  $p_1$  with  $(0 < p_1 < 1)$ , if in the previous trial the outcome was tail; and the probability of its falling head is  $p_2$  with  $(0 < p_2 < 1)$  if in the previous trial the outcome was head. At the first trial the probability of head is  $p_1$ .  $G$ , consists of tossing the coin until that trial when the total number of heads exceeds the number of tails by  $r$  for the first time. The sample space



$\mathcal{Z} = (Z, \Omega, p)$  is considered, where  $Z$  consists of points representing the outcomes of  $G_i$ ,  $\Omega = (0 < p_1 < 1) \times (0 < p_2 < 1)$  with  $p_1 + p_2 > 1$ . Using a combinatorial lemma established by the author (*Comptes Rendus*, t. 240, pp. 1188-89), a sufficient partition for this sample space has been determined, for all  $r \geq 2$ . The problem of estimating  $p_1, p_2$  is being considered.

**24. Estimation of the Mean and Variance of a Quantitative Characteristic in a Polygenic System.** ALLAN G. ANDERSON, Western Kentucky State College.

A system is considered in which a quantitative characteristic (such as yield of corn) is affected by many gene-pairs whose contributions are equal and additive. It is assumed that means and variances are known for a set of inbred parent lines, and formulae are developed for the estimation of mean and variance of any hybrid descended from members of the inbred set by means of a known pedigree. One year of field experimentation is required to obtain needed data on all first generation hybrid crosses possible among the inbreds, but from then on attention can be directed toward those strains for which the prognosis is favorable based on the estimated means and variances.

**25. Selecting a Subset Containing the Best of Several Binomial Populations.** S. S. GUPTA AND M. SOBEL, Bell Telephone Laboratories, Inc.

Let  $x_i$  denote the number of successes in a known number  $n_i$  of observations from a binomial population  $\Pi_i$  with unknown probability  $p_i$  of success in a single trial; let  $y_i = x_i/n_i$  ( $i = 1, 2, \dots, k$ ). The problem is to select a subset of the  $k$  populations  $\Pi_i$  so that the "best" population (i.e., the population with the largest value of  $p$ ) will be included in the selected subset with a preassigned probability  $P^*$ , regardless of the true values of the  $p_i$ . The suggested procedure  $R$  is "Retain a population  $\Pi_i$  if and only if  $y_i \geq \max(y_1, y_2, \dots, y_k) - b$ ". The constant  $b (\geq 0)$  is determined so as to satisfy the given probability requirement. Expressions for the probability of a correct selection (i.e., the selection of a subset containing the best population) for the procedure  $R$  are derived and, in the case of a common number  $n$  of observations, these are used to construct tables of the smallest constant needed to carry out the procedure  $R$  for selected values of  $n, k$  and  $P^*$ . Formulae are obtained for the expected number of populations retained in the selected subset and tables are given for the expected proportion of populations retained. Alternative procedures based on the transformation of variables are briefly discussed.

**26. Hadamard Matrices and a Problem in the Theory of Code Construction.** R. C. BOSE AND S. S. SHRIKHANDE, University of North Carolina.

A sequence  $\alpha = (a_1, a_2, \dots, a_n)$ ,  $a_i = 0$  or  $1$  is called an  $n$ -place message. Hamming distance between two  $n$ -place messages is the number of positions which are different. Let  $A(n, d)$  denote the maximum number of  $n$ -place messages that can be constructed such that the distance between any two messages is greater than or equal to a pre-assigned positive integer  $d < n$ . We prove the result that the following statements are equivalent: (i)  $A(4t, 2t) = 8t$ , (ii)  $A(4t - 1, 4t) = 4t$ , (iii) there exists a symmetric balanced incomplete block design with parameters  $v = b = 4t - 1, r = k = 2t - 1, \lambda = t - 1$ , and (iv) a Hadamard matrix of order  $4t$  exists. This generalizes results of Plotkin (Research Division Report 51-20 (1951), University of Pennsylvania, Moore School of Electrical Engineering), where he showed that  $A(4t, 2t) \leq 8t$  and  $A(4t - 1, 4t) \leq 4t$  and that the maximal codes with  $A(4t, 2t) = 8t$  and  $A(4t - 1, 4t) = 4t$  could be constructed if  $4t - 1$  is a prime. The structure of these maximal codes is also investigated here. These maximal codes can be constructed for all values of  $t \leq 50$ , except possibly for  $t = 23, 29, 39, 46$  and  $47$ .

**27. Modified Neyman-Pearson Methods Which Avoid "Paradoxes" and Tend to Coincide with Other Methods.** (Preliminary report) ALLAN BIRNBAUM, Columbia University.

The methods proposed take the following form in problems of testing between two simple hypotheses  $H_i$  specifying respective densities  $p_i(x)$ ,  $i = 1, 2$ , if  $r(X) = p_2(X)/p_1(X)$  has a continuous c.d.f.  $F_1(u)$  under  $H_1$ : If outcome  $x$  is observed, report the pair of error levels  $\alpha(x)$ ,  $\beta(x)$ , where  $\alpha(x) = 1 - F_1(r(x))$  and  $\beta(x) = F_2(r(x))$ ; if  $q(x) \equiv \beta(x)/\alpha(x) \gg 1$ , the observation  $x$  is considered strong evidence for  $H_2$  as against  $H_1$  (in the usual general Neyman-Pearson sense); if  $q(x)$  is not far from 1, the observation is considered inconclusive evidence for comparing the hypotheses, etc. Such methods avoid usual accept-or-reject formulations which seem over-schematic in scientific research applications. They avoid semblances of reaching strong conclusions from weak data which may appear with more conventional methods (cf. e.g. Cohen, *Ann. Math. Stat.*, Vol. 29 (1958), pp. 947-972). In any experiment providing at least a moderate amount of information, these methods give inferences which tend to agree with those obtained by maximum-likelihood methods; or Bayesian methods (excluding only extremely unequal *a priori* probabilities; cf. Lindley, *Biometrika*, Vol. 44 (1957), pp. 187-192); or Lehmann's approach (excluding only  $k$  very far from 1, on p. 1171 of *Ann. Math. Stat.*, Vol. 29 (1958), pp. 1167-1176).

**28. Hypothesis Tests on the Population Lower Limit (Minimum Life).** PHILLIP G. CARLSON, Arthur Andersen & Co. (introduced by S. B. Littauer). (By title)

Let  $x_1 < x_2 < x_3 < \dots < x_n$  be an ordered sample of  $n$  elements from a population  $F(x, \epsilon, \sigma, K)$  where  $x$  is a random variable,  $\epsilon$  is the lower limit,  $\sigma$  is a scale parameter, and  $K$  is a known family parameter. For certain populations the statistic  $h_n = (x_1 - \epsilon)/(x_n - x_1)$  is independent of the scale parameter  $\sigma$ , and can be used for testing hypotheses on the (unknown) lower limit  $\epsilon$ . For the population  $F(x, \epsilon, \sigma, K) = ((x - \epsilon)/\sigma)^K$ ,  $\epsilon \leq x$ , the  $h_n$  statistic has the probability function  $R(h_n) = 1 - \{1 - [h_n/(h_n + 1)]^K\}^{n-1}$ . The moments are given (for  $K = 1$ ) by  $Eh_n^t = n\beta(t + 1, n)$ . For  $F(x, \epsilon, \sigma, K) = 1 - \exp\{-(x - \epsilon)/\sigma\}^K$ ,  $\epsilon \leq x$ , which is the Third Asymptotic Distribution of Extreme Values,  $h_n$  has the probability function  $R(h_n) = 1 - [n/\{1 - [h_n/(h_n + 1)]^K\}] \beta[n[h_n/(h_n + 1)]^K/\{1 - [h_n/(h_n + 1)]^K\} + 1, n]$ . For  $K = 1$ , the moments are related by

$$Eh_n^t = [t(n - 1)/n(t - 1)] [(n - 1)/n]^{t-1} Eh_{n-1}^{t-1} - Eh_n^{t-1}.$$

Since this family has been used widely to explain fatigue and reliability phenomena, the  $h_n$  statistic can be used to test the minimum life of material, or of a component or set of components.

**29. Truncation and Tests of Hypotheses II.** IRWIN GUTTMAN, Princeton University. (By title)

The distribution of the sum of squares of  $n$  truncated normal variables is derived for the case  $n = 1(1)4$ , where the terminus point is " $a$ " standard deviations on either side of the mean. The difference in power and size of tests of hypotheses concerning the variance (the mean assumed known) is contrasted with the usual procedure, i.e. assuming a random variable has a 'complete' normal distribution, and the correct significance points are obtained. An indication of the corresponding " $F$ " situation is given.

**30. Randomization and Factorial Experiments.** SYLVAIN EHRENFELD, New York University. (By title)

This paper examines several questions relating to the  $2^k$  factorial series of experimental designs. One question that is considered is the effect on the usual estimating and testing procedures of a  $2^k$  factorial experiment when the relevant number of factors are  $k + n$ . If the levels of the remaining  $n$  factors were chosen systematically we have, in effect, a  $1/2^n \times 2^{k+n}$  fractional factorial. This question is further examined in terms of randomized choices of the levels of the  $n$  factors. This is, in effect, a *randomized fractional factorial* where  $2^k$  out of  $2^{k+n}$  experiments are chosen by some randomized procedure. The question of the various methods of carrying out such procedures is examined. A particular procedure is outlined, whereby one can estimate subsets of the effects with randomized fractional factorials without the usual assumptions of negligible interactions. The above procedure lends itself to an approach whereby a larger and larger fraction of the full factorial is used. At each stage, the usual testing and estimation procedures can be carried out. These methods are particularly important when experimentation, for exploratory purposes, with a large number of factors, is carried out sequentially.

**31. Geometrical Methods in the Construction of Group Alphabets.** R. C. BOSE AND ROY R. KUEBLER, JR., University of North Carolina. (Invited paper)

Notions of finite projective geometry are applied to the group alphabet introduced by Slepian (*Bell System Technical Journal*, Vol. 35 (1956), pp. 203-234). This alphabet is a  $2^k$ -element subgroup of the Abelian group of  $2^n$  sequences of  $n$  binary digits. The construction of such an alphabet is equivalent to the distribution of an integral  $W$ -measure over the points of  $PG(k-1, 2)$ , the measure of each point being the *weight* (number of unities) of a certain sequence (letter) of the alphabet. Necessary and sufficient conditions that an integral measure define a group alphabet include the congruences  $W_{k-2,i} = 0 \pmod{2^{k-2}}$  for all  $i$ , where  $W_{k-2,i}$  denotes the sum of the  $W$ -measures of all points on the  $i$ th  $(k-2)$ -flat of  $PG(k-1, 2)$ . The additional congruence conditions  $W_{c,u} = 0 \pmod{2^c}$  for all  $u, c = 1, 2, \dots, k-3$ , are necessary. Geometrical considerations are applied to the problems of (1) satisfying the congruence conditions, (2) finding the relation between  $n$  and  $W$  ( $W$  being defined as the largest integer such that the code associated with the alphabet will correct all transmission errors of multiplicity up to and including  $W$ ), and (3) counting the number of  $(W+1)$ -tuple errors which will be corrected. Specific results are given for  $k = 2, 3, 4$ .

**32. A Simple Minimum-Average-Risk Procedure for the Multiple Comparisons Problem.** (Preliminary report) DAVID B. DUNCAN, University of North Carolina. (Invited paper)

Let  $[y_1, \dots, y_n, s]$  be a sufficient estimator for  $[\mu_1, \dots, \mu_n, \sigma]$  such that, to take a typical simple case,  $[y_1, \dots, y_n]$  is normally distributed with mean  $[\mu_1, \dots, \mu_n]$  and variance  $I\sigma^2$ , and  $s^2$  is the usual form of independent estimate (with  $\nu$  degrees of freedom) for  $\sigma^2$ . Let  $T$  represent the class of  $n(n-1)$  differences  $T = \{\tau: \tau = (\mu_i - \mu_j)/\sqrt{2}\sigma; i, j = 1, \dots, n; i \neq j\}$ . The sub-set system of the multiple comparisons problem considered is that formed as the restricted product of the two-decision component-problem subset pairs  $\tau > 0, \tau \leq 0$  for all  $\tau \in T$ . A Bayes solution of the form indicated in [1] Duncan, *Ann. Math. Stat.*, 1958, p. 622, is developed for each component problem. Their simultaneous application, all  $\tau \in T$ , is shown (see also, [2] Lehmann, *Ann. Math. Stat.*, 1957, pp. 1-25) to be the Bayes solution to the given multiple comparisons problem with respect to a loss function formed as the sum of the component loss functions and to a Bayes function having the component Bayes

functions at its margins. The table of  $t$  in [1] is used for each component solution, the test statistic now also being a function of the residual variance among the  $y$ 's and having  $\nu + n - 2$  instead of  $\nu$  degrees of freedom. (Research jointly supported by the U. S. Air Force through the Office of Scientific Research of the Air Research and Development Command and by the U. S. Public Health Service).

(Abstracts of papers presented at the Cleveland, Ohio Meeting  
of the Institute, April 2-4, 1959.)

### 1. Relation Between Certain Incomplete Block Design. S. S. SHRIKHANDE, University of North Carolina.

The following results are proved. *Theorem (1)*. A partially balanced incomplete block design with two associate classes, and with parameters  $u = (n + 1)_{c_2}$ ,  $b = n_{c_2}$ ,  $r = n - 1$ ,  $k = n + 1$ ,  $n_1 = 2n - 2$ ,  $\lambda_1 = 1$ ,  $\lambda_2 = 2$ ,  $p_{11}^1 = n - 1$ ,  $p_{11}^2 = 4$  has triangular association scheme  $(T_{n+1})$ . (Bose and Shimamoto, *J.A.S.A.*, Vol. 47 (1952), pp. 151-184). *Theorem (2)*. For any value of  $n$ , the existence of the balanced incomplete block design  $D_1$  with parameters  $u = n_{c_2}$ ,  $b = (n + 1)_{c_2}$ ,  $r = n + 1$ ,  $k = n - 1$ ,  $\lambda = 2$  implies the existence of two p.b.i.b. designs,  $D_{11}$  with parameters  $u = b = n_{c_2}$ ,  $r = k = n - 1$ ,  $n_1 = 2n - 4$ ,  $\lambda_1 = 1$ ,  $\lambda_2 = 2$ ,  $p_{11}^1 = n - 2$ ,  $p_{11}^2 = 4$ , and  $D_{12}$  with parameters  $u = n_{c_2}$ ,  $b = n$ ,  $k = n - 1$ ,  $r = 2$ ,  $n_1 = 2n - 4$ ,  $\lambda_1 = 1$ ,  $\lambda_2 = 0$ ,  $p_{11}^1 = n - 2$ ,  $p_{11}^2 = 4$  such that  $D_1 = D_{11} + D_{12}$ , where "+" denotes the fact that blocks of  $D_{11}$  and  $D_{12}$ , taken together give the design  $D_1$ . *Theorem (3)*. The existence of the design  $D_{11}$  implies the existence of the corresponding design  $D_1$ , if the association scheme of  $D_{11}$  is triangular  $(T_n)$ . *Theorem (4)*. If  $n = 5$ , or  $n \geq 9$ , the dual of the design  $D_{11}$  is another p.b.i.b. with the same parameters. A constructive method of embedding the design  $D_1$  into the corresponding symmetric b.i.b.d. with parameters  $u = b = (n^2 + n + 2)/2$ ,  $r = k = n + 1$ ,  $\lambda = 2$  is also given.

### 2. Quasi-Ranges of Samples from an Exponential Distribution. PAUL R. RIDER, Wright Air Development Center.

The distribution of quasi-ranges of samples from an exponential population is given, as are formulas for the cumulants of the distribution.

### 3. Asymptotic Rate of Discrimination for Markov Processes. LAMBERT H. KOOPMANS, Sandia Corporation.

Simple hypotheses  $H_P$  and  $H_Q$  specifying two distinct, positive, transition densities  $p(x | y)$  and  $q(x | y)$  and positive initial densities  $\pi_P(x)$  and  $\pi_Q(x)$  with respect to a finite Lebesgue-Stieltjes measure are assumed for a discrete time parameter Markov process. Let  $R_n$  be the likelihood ratio statistic based on the first  $n + 1$  observations of the process and consider the class of sequences of likelihood ratio tests  $T(a) = \{[R_n > na] : n = 0, 1, 2, \dots\}$  generated by letting  $a$  vary in the interval  $-\infty < a < \infty$ . If the function  $K_t(x, y) = p^{1-t}(x|y)q^t(x|y)$  satisfies a certain regularity condition it is shown that there exist limits  $l_P$  and  $l_Q$ ,  $l_P < 0 < l_Q$ , independent of the initial densities, such that the sequences  $T(a)$  are consistent for  $l_P < a < l_Q$  and inconsistent in the complement of the closure of this interval. Furthermore, the rate at which the error probabilities tend to zero is exponential for  $l_P < a < l_Q$ . An asymptotic rate of discrimination  $\rho(P, Q)$  is defined which is a measure of the limiting behavior of the class of consistent likelihood ratio sequences  $T(a)$ . It is shown that  $\rho(P, Q)$  is the infimum, over the unit interval, of the largest eigenvalue of the integral operator with kernel  $K_t(x, y)$ . Several examples are considered and an extension to Markov processes with respect to arbitrary Lebesgue-Stieltjes measures is indicated.

**4. Distribution of a Quadratic Form in Three Variables.** (Preliminary report)  
 ANDRE G. LAURENT, Wayne State University.

Let  $X = (u, v, w)'$  be normally distributed  $N(0, S), S > 0$ . A series expansion involving confluent hypergeometric functions is proposed for the distribution of  $X'X$ . Further,  $P(X'X \leq R^2) = \sum a_k I_k(R) R^{2(k+1)}$ , where the  $a_k$  are the coefficients of the series expansion corresponding to the two dimensional case and  $I_k$  can be obtained by recursion from  $I_0$  and  $I_1$ , which are easy to compute with the help of a table of normal integrals and ordinates.

**5. On Exponentially-Mapped-Past Statistical Variables.** (Preliminary report)  
 JOSEPH OTTERMAN, Willow Run Laboratories.

The exponentially-mapped-past statistical variables are quantities relating to a set of observations computed in such a way that the recent values of the observations contribute more strongly than the values observed in the more distant past. The relative weighting is a geometrical ratio in the case of discrete (naturally discrete or sampled) data and an exponential function in the case of continuously observed functions. In this paper definitions of some exponentially-mapped-past variables are introduced and certain simple relationships are discussed. The distinct computational advantages of the e.m.p. variables, such as e.m.p. average and e.m.p. variance, are pointed out.

**6. Simultaneous Comparison of the Optimum and Sign Tests of a Normal Mean.** R. R. BAHADUR, Indian Statistical Institute. (By title)

This paper gives a detailed example of a general method of comparing two tests. Consider a sample of  $n$  independent observations from an  $N(\mu, 1)$  population and suppose it is desired to test  $\mu = 0$  against  $\mu > 0$ . Let  $L_0(n)$  and  $L_s(n)$  denote the significance levels actually attained in the given case by the optimum and sign tests respectively. The paper studies the asymptotic joint distribution of  $L_0(n)$  and  $L_s(n)$ . It is shown that if  $\mu = 0$  the limiting distribution of  $L_0$  and  $L_s$  is that of  $G(U)$  and  $G(V)$ , where  $G$  is the  $N(0, 1)$  distribution function, and  $U$  and  $V$  are correlated  $N(0, 1)$  variables, the correlation being  $(2/\pi)^{1/2}$ . In case  $\mu > 0$ ,  $L_0$  and  $L_s$  tend to zero, the asymptotic relationship being (roughly speaking)  $L_0(n) = \tau^n \cdot L_s(n)$ , where  $\tau$  is a constant depending on  $\mu$  such that  $0 < \tau < 1$  and such that  $\tau$  decreases as  $\mu$  increases. The question of estimating or predicting the actual value of  $L_0$  given only  $n$  and  $L_s$  is discussed. It is shown that, with attainment of an assigned level as the criterion,  $\varphi = 2 \log_e [2p^2q^2/\mu^2]$  serves as the asymptotic efficiency of the sign test, where  $p = P(X > 0 | \mu)$ ,  $q = 1 - p$ . If the variance is not known to be 1,  $\mu^2$  is replaced by  $\log(1 + \mu^2)$  in the formula for  $\varphi$ .

**7. Some Estimates of the Binomial Distribution Function.** R. R. BAHADUR, Indian Statistical Institute. (By title)

Let  $p$  be given,  $0 < p < 1$ . Let  $n$  and  $k$  be positive integers such that  $np \leq k \leq n$ , and let  $B_n(k) = \sum_{r=0}^k \binom{n}{r} p^r q^{n-r}$ , where  $q = 1 - p$ . It is shown that  $B_n(k) = \left[ \binom{n}{k} p^k q^{n-k} \right] \cdot q \cdot F(n+1, 1; k+1; p)$ , where  $F$  denotes the hypergeometric function. This representation seems useful for numerical as well as theoretical investigations of small tail probabilities. The representation yields, in particular, the result that with  $A_n(k) = \binom{n}{k} p^k q^{n-k+1} \cdot [(k+1)/(k+1 - (n+1)p)]$ , we have  $1 \leq A_n(k)/B_n(k) \leq 1 + x^{-2}$ , where  $x = (k - np)/(npq)^{1/2}$  (Theorem 1). Next, let  $N_n(k)$  denote the Normal approximation to  $B_n(k)$ , and let  $C_n(k) = (x + (q/np)^{1/2}) \cdot (2\pi)^{1/2} \cdot \exp(x^2/2)$ . It is shown that  $(A_n N_n C_n)/B_n \rightarrow 1$  as  $n \rightarrow \infty$ , provided only

that  $k$  varies with  $n$  so that  $x \geq 0$  for each  $n$  (Theorem 2). It follows that  $A_n/B_n \rightarrow 1$  if and only if  $x \rightarrow \infty$  (i.e.  $B_n \rightarrow 0$ ) (Corollary 1). It also follows that  $N_n/B_n \rightarrow 1$  if and only if  $A_n C_n \rightarrow 1$ . This last condition reduces to  $x = o(n^{1/6})$  for certain values of  $p$ , but is weaker than this for the other values; in particular, there are values of  $p$  for which  $A_n C_n \rightarrow 1$  can hold without requiring even that  $k/n \rightarrow p$  (Corollary 3).

**8. Distribution-Free and Nonparametric Tolerance Regions: The Exponential Case.** (Preliminary report) LEO A. GOODMAN AND ALBERT MADANSKY, University of Chicago and The Rand Corporation. (By title)

Exact one-sided and approximate two-sided  $\alpha$  content tolerance regions at confidence level  $C$  are developed based on the first  $r$  ordered observations from a sample of  $n$  exponentially distributed variables. These regions are compared with nonparametric one-sided and two-sided tolerance regions. Optimal properties of these regions are discussed, as is the asymptotic behavior of the tolerance regions. It is also shown that these regions are distribution-free in the sense defined by Fraser and Guttman, *Ann. Math. Stat.*, Vol. 27 (1956), pp. 162-79. The effect of assuming an exponential distribution, when in fact the distribution is a mixture of two exponentials, is discussed. Also, uniformly most powerful invariant one-sided and two sided  $\beta$ -expectation tolerance regions are derived. Some of the results presented by Goodman, *Ann. Math. Stat.*, Vol. 24 (1953), pp. 139-40, are extended.

**9. Bayesian Lot-by-lot Sampling Inspection.** HERBERT B. EISENBERG, Iowa State College. (By title)

Based on the Work of Arrow, Blackwell and Girshick (*Econometrica*, 1949), this paper derives Bayesian single, double and sequential attribute sampling plans. Lot quality distributions considered are the binomial, two-point, and degenerate one-point. The loss function considered is negative profit, given by  $L_{acc.} = -sd - u_1 N(1 - p) + u_2 (Np - d) + ni$ ,  $L_{rej.} = -sNp - u_1 N(1 - p) + Ni$ ,  $N$  being size of lot. Profit efficiencies of single and double sampling plans relative to sequential plans are computed in specific cases. Partial ignorance is considered by evaluating loss incurred when optimizing with respect to the wrong lot quality distribution. As expected, sampling never pays in the binomial case; in the two-point case, the optimum sequential plan is not necessarily hypergeometric SPRT; indeed, the acceptance portion of the boundary need not be connected.

**10. Combining Inter-block and Intra-block Information in Balanced Incomplete Block Designs.** FRANKLIN A. GRAYBILL AND DAVID L. WEEKS, Oklahoma State University. (By title)

When an Eisenhart Model III (blocks random, error random) is assumed in a balanced incomplete block, two independent estimates of treatment differences have been exhibited by Yates. A combined estimate of treatment differences has also been set forth by Yates but none of the properties of the combined estimate have been given.

It is the purpose of this paper to show that Yates' combined estimate is based on a set of minimal sufficient statistics. A combined estimate is set forth in the paper which is shown to be unbiased and which is also based on a set of minimal sufficient statistics.

**11. Minimal Sufficient Statistics in Incomplete Block Designs, Model II.** DAVID L. WEEKS AND FRANKLIN A. GRAYBILL, Oklahoma State University. (By title)

Under the assumption of an Eisenhart Model II in a balanced incomplete block design, minimal sufficient statistics are exhibited which have dimension six. These six statistics

can be found from the quantities which are used to obtain an analysis of variance by the recovery of inter-block information method. The distribution of the six statistics is discussed. Similar results have been found for certain types of partially balanced incomplete block designs with two associate classes.

**12. Some Theorems Concerning Eisenhart's Model II.** FRANKLIN A. GRAYBILL  
AND ROBERT HULTQUIST, Oklahoma State University. (By title)

Eisenhart's analysis of variance Model II can be written as follows:  $Y = X\beta + e$  or  $Y = \sum_{i=0}^k X_i\beta_i + e$  where  $\beta_0 = \mu$  is a scalar constant,  $\beta_i$ , ( $i \neq 0$ ) is a vector of  $p_i$  random variables such that  $E(\beta_i) = 0$ ;  $E\beta_i\beta_i' = \sigma_i^2 I$ ;  $e$  is a vector of  $n$  random variables such that  $E(e) = 0$ ;  $E(ee') = \sigma_{k+1}^2 I$ ; all random variables are independent. This model is studied with respect to point estimation. Under the assumption that all random variables are *normal* variables, theorems on the following were proved: (1) The maximum and minimum number of distinct characteristic roots of the covariance matrix of  $Y$ ; (2) Conditions on the design matrix  $X$  for complete sufficient statistics to exist; (3) Minimal sufficient statistics when the design matrix satisfies certain conditions; (4) Analysis of variance estimators. When the random variables are not normal, theorems on the following were proved: (1) Uniformly best (minimum variance) unbiased quadratic estimates of the  $\sigma_i^2$ ; (2) Estimable functions of the  $\sigma_i^2$ .