CORRECTION TO "AN APPLICATION OF INFORMATION THEORY TO MULTIVARIATE ANALYSIS, II"

By S. Kullback

The George Washington University

In the paper cited in the title (*Ann. Math. Stat.*, Vol. 27 (1956):

p. 122, line 7, delete 'maximizing information' and replace by 'discriminating between a null hypothesis and the alternative hypothesis by using that distribution corresponding to the alternative hypothesis which for the sample values provides the least information for discrimination';

p. 123, line 3, between 'information' and 'in' insert 'for linear discriminant functions';

p. 124, line 6, same as p. 122, line 7 above;

p. 124, line 11, change 'maximum' to 'minimum';

p. 124, line 13, change 'maximizing' to 'minimizing';

p. 124, (3.8), add '$a = \int \frac{ye^{ty}g_2(y)\,d\gamma(y)}{M_2(t)} = \frac{\frac{dM_2(t)}{dt}}{M_2(t)}$';

p. 125, line 3, delete 'maximum' and replace by 'minimum I^*';

p. 125, immediately following (3.12) add 'It is readily verified that $g^*(y)$ is normal when $g_2(y)$ is normal.';

p. 130, line 23, insert '(' between '=' and '$X_{(1)}$'.

ABSTRACTS OF PAPERS

(*Abstracts of papers submitted for the Seattle meeting of the Institute, August 20–27, 1956*)

1. **On the Studentized Largest and Smallest χ^2**, K. V. Ramachandran, University of Baroda, India.

Let $s_1^2, s_2^2, \cdots, s_k^2$ be k independent χ^2 variables with m d.f. each. Let s^2 be another independent χ^2 variable with n d.f. Then the Studentized largest χ^2 is defined as: $u = s_{\max}^2/s^2$, where $s_{\max}^2 = \max(s_1^2, s_2^2, \cdots, s_k^2)$. Similarly the Studentized smallest χ^2 is defined as: $v = s_{\min}^2/s^2$, where $s_{\min}^2 = \min(s_1^2, s_2^2, \cdots, s_k^2)$. Using methods given in an earlier paper the upper and lower 5 percent points of u and v are given for different values of m, n and k. These statistics have been found to be useful in several situations including control of quality, simultaneous confidence interval estimation, testing for normality against uniform distribution, etc. (Received April 16, 1956.)

2. **The Linear Hypothesis, Information, and the Analysis of Variance, (Preliminary Report)**, Chester H. McCall, Jr., The George Washington University.

The concepts of "information" (designated by "i") and "mean information per observation" (designated by "I") for differentiation between two hypotheses first appeared in

"On Information and Sufficiency" (S. Kullback and R. A. Leibler, *Ann. of Math. Stat.*, Vol. 22 (1951), pp. 79–86). Since that time numerous articles have appeared on the same subject, discussing applications of Information to multivariate analysis as well as analytical developments of distributions of certain of the Information statistics. In this paper, the linear hypothesis and its application to the analysis of variance is considered from the viewpoint of mean information per observation to discriminate between the usual null hypothesis and composite generalized alternative hypothesis. The information measure of divergence between two hypotheses, $J(1:2)$, is introduced. Best unbiased estimates of the parameters are employed in estimating the mean divergence, and we call this estimate $\hat{J}(1:2)$—read "J-caret." The statistic $\hat{J}(1:2)$ is identical to k_1F where k_1 is some appropriate degree of freedom and F has the analysis of variance distribution with k_1 and k_2 degrees of freedom. Applications are made to one-way, two-way, two-way with replication, and Latin Square designs. It is shown that the given method applies to orthogonal and nonorthogonal designs. (Received April 23, 1956.)

3. A Sequential Multiple Decision Procedure for Selecting the Multinomial Event with the Largest Probability, (Preliminary Report), R. E. Bechhofer, Cornell University, and M. Sobel, Bell Telephone Laboratories.

Let $x_j = (x_{1j}, x_{2j}, \cdots, x_{kj})$ be independent vector-observations from a single multinomial population with a common unknown probability vector $p = (p_1, p_2, \cdots, p_k)$; here p_i is the probability of the event $E_i(0 < p_i < 1, \sum_{i=1}^{k} p_i = 1)$ and $x_{ij} = 1$ or 0 according as E_i does or does not occur at the jth stage $(i = 1, 2, \cdots, k; j = 1, 2, \cdots)$. Let $p_{[1]} \leqq p_{[2]} \leqq \cdots \leqq p_{[k]}$ denote the ranked probabilities; let $\Delta = p_{[k]} \div p_{[k-1]} \geqq 1$. A sequential procedure is proposed which guarantees a probability of at least $P^*(1/k \leqq P^* < 1)$ of selecting the event associated with $p_{[k]}$ whenever $\Delta \geqq \Delta^*(1 < \Delta^* < \infty)$; the constants P^* and Δ^* are preassigned. Let $y_{im} = \sum_{j=1}^{m} x_{ij}(i = 1, 2, \cdots, k)$ and let $y_{[1]m} \leqq y_{[2]m} \leqq \cdots \leqq y_{[k]m}$ denote their ranked values. Let $E_{[i]m}$ denote the event associated with $y_{[i]m}$ at the mth stage $(i = 1, 2, \cdots, k)$. *Procedure*: "At the mth stage $(m = 1, 2, \cdots)$ take the vector observation x_m and compute $W_m = \sum_{i=1}^{k-1} [1/\Delta^*]^{y_{[k]m} - y_{[i]m}}$. If $W_m \leqq (1 - P^*)/P^*$, stop and select $E_{[k]m}$ (or if at this stage $y_{[k-t]m} < y_{[k-t+1]m} = \cdots = y_{[k]m}$, select one of $E_{[k-t+1]m}, \cdots, E_{[k]m}$ using a randomized device which assigns probability $1/t$ to each of them); if $W_m > (1 - P^*)/P^*$, take the vector x_{m+1} and compute W_{m+1}." This procedure has probability one of terminating. It can be generalized to handle problems such as obtaining a complete ranking of the k probabilities. (Research supported in part by the U. S. Air Force through the Office of Scientific Research of the ARDC.) (Received May 4, 1956.)

4. On the Existence of Uniformly Efficient Estimates, R. R. Bahadur, University of Chicago.

Suppose that the sample point x is distributed according to some one of a given set P of probability measures p. A real valued function g on P is said to be L_2-estimable if there exists at least one unbiased estimate of g, say $t(x)$, such that the variance of t is finite for each p. Let $y = S(x)$ be a minimal sufficient statistic for P. It is then said that y is L_2-complete (for P) if $t \equiv 0$ is essentially the only unbiased estimate of zero that depends on x only through S and that has finite variance for each p. As may be seen from an argument by Lehmann and Scheffé ("Completeness, similar regions, and unbiased estimation, Part I," *Sankhyā*, 10 (1950), pp. 305–340), if y is L_2-complete, then every L_2-estimable parameter g possesses an unbiased estimate of uniformly minimum variance. The main conclusion of the paper is that, conversely, if every L_2-estimable g possesses an unbiased estimate of uniformly minimum variance, then y must be L_2-complete. It is also shown by an example that in general this converse does not hold with "L_2-estimable" and "L_2-complete" replaced by the parallel L_1 concepts. (Received May 7, 1956.)

5. On the Distribution of Ranks and of Certain Rank Order Statistics, MEYER DWASS, Northwestern University and Stanford University.

Suppose $X_1, \cdots, X_m$ and $X_{m+1}, \cdots, X_N$ are two independent samples from two possibly different populations, $R_1, \cdots, R_m$ are the ranks of the first m observations in the combined sample, and $R_{m+1}, \cdots, R_N$ are the ranks of the remaining observations. Various moment-generating functions involving these ranks are derived, including that of the Wilcoxon statistic. The asymptotic distribution of a finite number of ranks is derived as $N \to \infty$. The remainder of the paper studies certain aspects of the distribution theory of rank order statistics of the form $\sum_{i=1}^{m} f_n(R_i/N)$. The Wilcoxon statistics and Hoeffding c_1-statistic are special cases of such a statistic. Many previous studies have been devoted to showing asymptotic normality. The main purpose here is to show that for certain combinations of sample sizes m, n, the limiting distribution may be non-normal as $m \to \infty$, $n \to \infty$, and $m/N \to 0$. (Work performed under Office of Naval Research Contract Nonr-225(21).) (Received May 8, 1956.)

6. Contributions to Distribution-Free Population Comparisons, WILLIAM E. PERRAULT and WALDO A. VEZEAU, St. Louis University.

Distribution-free techniques based on the median of the combined samples and the number of runs in each sample are employed to test, by means of the chi-square criterion, the null hypothesis that two, or in general k, samples have come from the same population. An extension is carried out to the two-criterion experiment. Procedures are supplied for the following: paired or unpaired replicates or groups of replicates, testing whether the superiority of one material over the other is consistent throughout the range of conditions studied, confidence limits for the mean difference between treatments. The power curve for the two-sample test is given. (Received June 4, 1956.)

7. A Probabilistic Model Describing Drop Count Data for Certain Closed Chamber Experiments, ROBERT R. READ, University of California at Berkeley.

As a fast particle passes through a cloud chamber it collides with the atoms present. Each collision results in the production of some ion pairs. One ion pair is formed directly as a result of the primary collision and some secondary ion pairs may be formed as somewhat of a chain reaction if the transfer of energy in the primary encounter is large enough. Water droplets are made to condense on the ions by expanding the chamber. However, not all of the ions acquire droplets. Furthermore, positive and negative ions are not equally efficient as nuclei of condensation. Let X represent the number of drops counted in a given length of track. Let γ represent the number of primary collisions, and let δ_k be the number of secondary collisions resulting from the kth primary collision. Let $u_{kj}(v_{kj})$ be one or zero according to whether or not the positive (negative) ion in the jth ion pair of the kth primary collision acquires a droplet. Then $X = \sum_{k=1}^{\delta_k} (u_{kj} + v_{kj})$. The distribution of γ is known to be Poisson. The distribution of secondaries is not known, but the use of a Poisson law seems to work. The u_{kj} and v_{kj} are independent and have binomial distributions. The distribution of X is deduced by making use of these assertions. The model was successfully fitted to 21 tracks. Formulae for computing probabilities are displayed. Some refinements and modifications of the model are discussed. A method to estimate the parameters is presented. (Received June 15, 1956.)

8. Efficient Small Sample Nonparametric Median Tests with Bounded Significance Levels, John E. Walsh, Lockheed Aircraft Corporation.

Nonparametric tests and confidence intervals for the median of a continuous statistical population always can be obtained from a sample from that population by a sign test type procedure. However, investigation has indicated that nearly all of these results have moderate or low efficiencies. The only practically important exceptions occur for small sample size cases which are based on the largest and/or smallest of the sample values. Consequently, the sample size, significance level, and confidence coefficient values available for these efficient sign test type results are very limited. This paper presents some additional nonparametric results which appear to be reasonably efficient and which noticeably increase the available numbers of sample sizes, significance levels, and confidence coefficients. These results are for small samples and do not have exactly determined probability properties. That is, the significance level for a test is bounded between two specified numbers but its exact value is unknown; similarly, for confidence coefficient values. One-sided and two-sided tests and confidence intervals are presented. The upper and lower bounds for the one-sided results are only moderately close together. For the two-sided results, the bounds are quite close together unless the continuous population sampled is extremely unsymmetrical. The bound values almost always can be considered sufficiently close for application if a continuous monotonic transformation of variable is available which yields an approximately symmetrical population. (Received June 18, 1956.)

9. Validity of Approximate Normality Values for $\mu \pm k\sigma$ Areas of Practical Type Continuous Populations, John E. Walsh, Lockheed Aircraft Corporation

Let us consider a continuous statistical population with mean μ and standard deviation σ. A useful empirical relation which seems to be approximately valid for many such populations concerns the fraction of the population contained in the interval $\mu - k\sigma$ to $\mu + k\sigma$. This empirical relation states that the fraction contained in this interval is nearly equal to the value obtained by assuming that the population is normal. This paper presents results which show that the stated empirical relation is roughly valid for a rather general class of continuous populations. The populations of this class are referred to as practical type populations. The class considered consists of those populations with probability density functions which can be adequately represented by the first seven terms of their Edgeworth series expansion. The closeness of the approximation obtained by use of the empirical relation changes with the value of k. The preferable values for k are investigated. Some possible applications of the empirical relation are outlined for what appears to be the most desirable value of k. These applications include quality control chart use, confidence intervals for σ, confidence intervals for the probability of success for a binomial population, and joint confidence regions for μ and σ. (Received June 18, 1956.)

10. Bayes Approach to Control of Fraction Defective, John V. Breakwell, North American Aviation, Incorporated.

Following Girshick and Rubin, the optimum quality control rule, when the fraction defective may jump at any time from a known low value to a known higher value, and when the probability of this jump is also known a-priori, belongs to a certain two-parameter family of control rules. The determination of the economic payoff as a function of these two parameters requires the solution of complicated integral-difference equations which may be replaced by differential-difference equations when the fractions defective are sufficiently small. Some asymptotic formulas are obtained from these latter equations by letting one of the parameters grow large. Numerical comparison with some Monte Carlo simulations of the original scheme indicate that these asymptotic formulas may prove useful as approximations to the operating characteristics of the control scheme. (Received June 22, 1956.)

11. Incomplete Sufficient Statistics and Similar Tests, ROBERT A. WIJSMAN, University of California at Berkeley, (introduced by David Blackwell).

If, under the hypothesis, a sufficient statistic is incomplete, the easily constructible similar tests of Neyman structure are not the only similar tests and may, in fact, be useless. For a class of exponential densities it can be proved, under some restrictions, that a sufficient statistic is incomplete if its dimension is greater than the number of parameters which specify the distribution under the hypothesis. The method of proof also provides a method of constructing a large class of similar tests. Applications are possible in the Behrens-Fisher problem and in the problem of tests of hypotheses concerning the ratio of mean to standard deviation in a normal population. In the latter case we have $X_1, \cdots, X_n$: $N(\mu, \sigma^2)$ and independent; the hypothesis specifies the value of μ/σ to be ρ_0, and a minimal sufficient statistic is $T = (T_1, T_2)$ with $T_1 = \sum X_i^2$, $T_2 = \sum X_i$. Similar test functions can be constructed by virtue of the fact that for any function $g(t_1, t_2)$, satisfying some mild conditions, we have

$$\int_0^\infty \int_{-\infty}^\infty \exp\left[-\frac{1}{2\sigma^2}t_1 + \frac{\rho_0}{\sigma}t_2\right]\left(\frac{\partial_g^2}{\partial t_2^2} - 2\rho_0^2\frac{\partial g}{\partial t_1}\right) dt_1\, dt_2 = 0$$

identically in σ. (Received June 25, 1956.)

12. Multi-Decision Problems for the Multivariate Exponential Family, DONALD R. TRUAX, California Institute of Technology.

A study of a class of decision procedures was made when the underlying distribution belongs to the multivariate exponential family. The number of possible actions is assumed finite, and the case of two possible decisions is studied in detail. When the loss functions L_1 and L_2 have the property that the set where $L_1 - L_2$ changes sign is linear, it has been possible to characterize the Bayes solutions and obtain complete classes of decision procedures. Two main problems of this type are considered and then various extensions of these problems given. The first main problem involves deciding whether or not the unknown parameter point lies on a given r-dimensional subspace of an n-dimensional parameter space, and the second problem concerns a decision as to which half space a parameter belongs. Applications of this theory are made to some classical problems of testing composite hypotheses. It is also shown that if the set where $L_1 - L_2$ changes sign is not the union of parallel linear sets, no nice characterization can be given to the Bayes procedures. Some problems where the number of possible actions is greater than two have been considered, and again complete classes were obtained. The question of admissibility has been studied for some of these problems. (Received June 26, 1956.)

13. Some Distributions Related to D_n^+, Z. W. BIRNBAUM and R. PYKE, University of Washington.

Let $U_1, \cdots, U_n$ be an ordered sample of a random variable which, without loss of generality, is assumed to have uniform distribution in (0, 1), and let $D_n^+ = \max_{1 \leq i \leq n}\{i/n - U_i\}$. Consider the random variables i^*, U^* defined by $D_n^+ = i^*/n - U_{i^*}$, $U^* = U_{i^*}$. Explicit formulae are obtained for the probabilities, Prob $\{U^* < u, i^* = j\}, 0 < u \leqq 1, j = 1, \cdots, n$, and $p_j = \text{Prob}\ \{i^* = j\}, j = 1, \cdots, n$. Some of the consequences of these formulae are: (i) $p_1 < p_2 < \cdots < p_n$; (ii) Prob $\{U^* < u\} = u; 0 < u \leqq 1$; (iii) $\lim_{n\to\infty} np_1 = e^{-1}$, $\lim_{n\to\infty} np_n = e$. The asymptotic distribution of i^* is obtained by studying the random variable $\alpha_n = i^*/n$, for which (iv) $\lim_{n\to\infty}$ Prob $\{\alpha_n < u\} = u, 0 < u \leqq 1$. This statement is made more specific by showing that (v) $E(\alpha_n) = 2^{-1}\{1 + n^{-n-1}n! \sum_{i=0}^{n-1} n^i/i!$ and (vi) $\sup_{0\leqq u\leqq 1} (u - \text{Prob}\ \{\alpha_n < u\}) \leqq n^{-n-1}n! \sum_{i=0}^{n-1} n^i/i! = O(n^{-\frac{1}{2}})$. (Received July 2, 1956.)

14. The Distribution of the Extreme Mahalanobis' Distance from the Sample Mean, (Preliminary Report), Yvonne G.M.G. (Mrs. P. M.) Cuttle, University of British Columbia, (introduced by S. W. Nash).

The problem of classification in multivariate analysis is considered. For the special bivariate case of three groups with the common dispersion matrix known, the distribution of the extreme Mahalanobis' distance from the centroid of the groups has been derived, and the cumulative distribution has been partially tabulated. If the three groups are found to be heterogeneous—that is, if the sum of the Mahalanobis' distances from the groups to their centroid is larger than χ^2 with 4 degrees of freedom, it is now possible, by comparing the extreme Mahalanobis' distance with the tabulated value of the cumulative distribution, to test the hypothesis that the group associated with this extreme distance belongs to the same population as the other two groups. For the general case, the characteristic function of the joint distribution of the Mahalanobis' distances from the sample mean has been derived. (Received July 2, 1956.)

15. The Quadratic Birth Process, Peter W. M. John, University of New Mexico.

The simplest divergent birth process is that in which $\lambda_n = \lambda n^2$, $n(0) = 1$. The process is divergent in the sense that for all $t > 0$ there is a positive probability $p_\infty(t)$ that an infinite number of births have occurred. It is shown that this probability is given by $1 - 2 \sum_{r=1}^{\infty} e^{-r^2\lambda t}(-1)^r$, which is the Jacobi Theta function $\Theta_4(0, e^{-\lambda t})$. (Received July 2, 1956.)

16. Sequential Distribution-free Tolerance Regions, Sam C. Saunders, University of Washington.

Let $X_1, X_2, \cdots$ be independent observations of a random variable X, and let $Y_{j,n}$, be the jth order statistic of the first n of these observations $X_1, X_2, \cdots X_n$, with the conventions $Y_{0,n} = -\infty$, $Y_{n+1,n} = +\infty$. Denoting by λ_n some subset of $\{1, 2, \cdots, n+1\}$, for $n = 1, 2, \cdots$, we form the union of closed intervals $A_n = \bigcup_{j\varepsilon\lambda_n} [Y_{j-1,n}, Y_{j,n}]$. Let $a_1, a_2, \cdots$ be a sequence of integers $\geqq 0$. We agree to continue making additional observations until for the first time, after $X_1, \cdots, X_n$ have been obtained and A_n have been determined, the additional observations $X_{n+1}, \cdots, X_{n+a_n}$ are all contained in A_n. If this happens for $n = N$, the set A_N is called a "tolerance region," and the random variable $Q = \text{Prob}\ \{X\ \varepsilon\ A_N\}$ is called the "coverage." The distribution of Q, which depends on the choice of $\{\lambda_n\}$ and $\{a_n\}$, is obtained for various choices of these sequences. Expected sample sizes are found, and criteria are proposed for comparing such sequential procedures. (Received July 2, 1956.)

17. Definite Quadratic Forms and Discontinuous Factor, Andre G. Laurent, Michigan State University.

In many instances, the derivation of the distribution of a positive quadratic form $X'AX$, with X n-dimensional and $f(X)$ distributed, is greatly simplified by using Dirichlet's discontinuous factor for the n-dimensional sphere. $P(X'AX \leqq R^2) = P(Y'Y \leqq R^2) = E(U)$, with $U = (R/2\pi)^{n/2} \int \cdots \int (TT')^{-n/4} e^{iTY} J_{n/2}[R(TT')^{\frac{1}{2}}] \, dT$, where $J_{n/2}$ denotes the Bessel Function of first kind and order $n/2$, and $E(\ \ )$ is taken with respect to the df $g(Y)$ of Y. Under the usual assumptions relative to convergence and order of integration, $P(Y'Y \leqq R^2) = (R/2\pi)^{n/2} \int \cdots \int h(T)(TT')^{-n/4} J_{n/2}[R(TT')^{\frac{1}{2}}] \, dT$, which is a generalised multivariate Hankel transform of the characteristic function $h(T)$ of Y, i.e., $P(Y'Y \leqq R^2) = (R^2/4\pi)^{n/2} \cdot CE\{\sum (-1)^k \ (R^2TT'/4)^k/k! \ \Gamma(n/2 + k + 1)\}$, where $E(\ \ )$ is taken with respect to the pseudo "df" $h(T)/C$. In case $g(Y)$ is $N(0, \Sigma)$, one obtains $P = (R^2/2)^{n/2} \mid \Sigma \mid^{-\frac{1}{2}} E[\sum (-1)^k \cdot (R^2/2)^k (TT'/2)^k / k! \Gamma(n/2 + k + 1)]$, where $E(\ \ )$ is taken with respect to $h(T)/C = N(0, \Sigma^{-1})$. If, further, $n = 2$ and a, b are the eigenvalues of Σ, $E[(TT'/2)^k] = a^{-k} k! \, {}_2F_1(-k, \frac{1}{2}; 1; 1 - a/b)$, i.e., $P = \sum_0^\infty (R^2/2a)^{k+1} (a/b)^{\frac{1}{2}} (-1)^k / (k + 1)! x \sum_0^k (-1)^j \binom{k}{j} \binom{2j}{j} (1 - a/b)^j \cdot 4^{-j}$. In case $g(Y)$ is $N(m, \Sigma)$, $P = R^{n/2} \mid \Sigma \mid^{-\frac{1}{2}} \int e^{iTm} N(0, \Sigma^{-1})(TT')^{-n/4} J_{n/2}[R(TT')^{\frac{1}{2}}] \, dT$, i.e., $R^{n/2} \mid \Sigma \mid^{-\frac{1}{2}} \exp[-\frac{1}{2} m' \Sigma m] \int \cdots \int N(im, \Sigma^{-1})(TT')^{-n/4} J_{n/2}[R(TT')^{\frac{1}{2}}] \, dT$. (Received July 3, 1956.)

18. On the Moments of Order Statistics from a Normal Population, R. C. Bose and Shanti S. Gupta, University of North Carolina.

It is shown that if $x_{(k)}$ is the kth order statistic for a sample of size n from a normal population $N(0, 1)$, then $\mu_r'(n, k)$, the rth moment of $x_{(k)}$ about the origin, can be expressed in terms of lower moments of order $r - 2i$, $(i = 1, 2, \cdots)$ and the integral $\int_{-\infty}^{+\infty} P_{r+1}(x) \cdot e^{-(r+1)x^2/2} \, dx$, where $P_{r+1}(x)$ for $r \geqq 0$ is defined by $P_{r+1}(x) = k \binom{n}{k} d^r/d\Phi^r [\Phi^{k-1}(1 - \Phi)^{n-k}]$, where Φ is replaced after differentiation by $\Phi(x)$, the cdf of $N(0, 1)$. Exact values of all odd order moments can be derived when $n \leqq 5$, and exact values of all even order moments can be derived when $n \leqq 6$. Godwin (*Ann. Math. Stat.*, 1949) has given a table of exact moments for $r = 1$ and 2. The corresponding tables for $r = 3$ and 4 have been provided. In general, numerical evaluation of the integral given above can be expeditiously done by using the Gauss-Jacobi method of mechanical quadrature based on zeros and weight factors corresponding to Hermite polynomials for which tables have been provided by Salzer, Zucker, and Capuano (*J. Research Nat. Bur. Standards*, Vol. 48). (Received July 5, 1956.)

19. Maximum Likelihood Estimation of Restricted Parameters, (Preliminary Report), H. D. Brunk, University of Missouri.

Let the ith of k populations depend exponentially (Blackwell and Girschick, *Theory of Games and Statistical Decisions*, pp. 179–194) on a parameter ω_i. For independent random sampling from these populations, the maximum likelihood (m.l.) estimator of the parameter point $\omega = (\omega_1, \omega_2, \cdots, \omega_k)$, given that it lies in a given closed subregion, S_0, of its "natural" domain, is described. The estimator is shown to possess a property related to sufficiency. If S_0 is bounded by hyperplanes, if A is the interior of S_0, or any of its faces, or edges, etc., if $\hat{\omega}(x)$ is the m.l. estimator of ω, and if E is an event, then for $\hat{\omega}(x) \ \varepsilon \ A$ there is a determination of the conditional probability $p_\omega(E \mid \hat{\omega}(x))$ which is independent of ω in the closure of A. The following consequence is related to the interpretation of sufficient estimator found in Halmos and Savage (*Ann. Math. Stat.*, Vol. 20 (1949), p. 240). If ω is common to all faces, edges, etc. of S_0, then the distribution of an arbitrary statistic can be duplicated by the procedure described by Halmos and Savage; if $\omega \ \varepsilon \ S_0$, the distribution of an arbitrary statistic can be approximated for large sample sizes. (Received July 5, 1956.)

20. A Comparison of the Power Curves of Some Double Sample Tests, Donald B. Owen, Sandia Corporation.

The exact power curves of the double sample tests introduced by the author (*Ann. Math. Stat.*, Vol. 24 (1953), pp. 449–457) for hypotheses on the mean, assuming a normal population and known standard deviation, have been computed. The τ-test procedure given in the referenced article is shown to be more powerful than one chosen from Bowker and Goode's *Sampling Inspection by Variables* (McGraw-Hill Book Co., 1952) after equating expected sample sizes. Tabulations also show that equal size samples at the two stages give a more powerful procedure, except for extreme alternatives, than taking the second sample twice the size of the first sample when expected sample sizes are equated. These differences are small but could become significant if many tests using these procedures were conducted. (Received July 5, 1956.)

21. On a Uniqueness Property not Enjoyed by the Normal Distribution, George P. Steck, Sandia Corporation.

It is known that if X_1 and X_2 are normal random variables with mean zero and variance σ^2, then X_1/X_2 has a Cauchy distribution. However, the fact that the ratio of two independent identically distributed random variables X_1 and X_2 with $EX_1 = EX_2 = 0$ has a Cauchy distribution does not imply that the random variables involved are normal. It can be shown that, for X_1/X_2 to have a Cauchy distribution, the characteristic function of $\log X_i$ must be of the form $e^{i\theta(t)}/(\cosh 2t)^{\frac{1}{2}}$, where $\theta(t)$ is real and odd. Convenient choices of the function $\theta(t)$ give tractable integrals; for example, if $\theta(t) = \arctan \tanh t$, then $p_{X_i}(x) = (2^{\frac{1}{2}}/\pi)x^2/(1 + x^2)$. (Received July 5, 1956.)

22. Confidence Intervals for the Number of Cells in a Multinomial Population with Equal Cell Probabilities, Bernard Harris, Stanford University and Department of Defense.

Consider a sample of n tosses of an r-sided die, with faces labeled $a, a + 1, \cdots, a + r - 1$. The parameter a and r are unknown, and we wish to determine confidence limits for r. The joint distribution of the smallest and largest observation is computed, as well as the distribution of the sample "range" $\hat{r} = x_n - x_1 + 1$. Unless n is very small relative to r, and for confidence coefficient $1 - \alpha$, (α small), suitable upper and lower confidence limits c_1, c_2 are given by $c_1 = \hat{r}$; c_2 is the largest positive real root of $P(y) = \hat{r}^n + (y - \hat{r})(\hat{r}^n - (\hat{r} - 1)^n) - \alpha y^n = 0$ which can be determined by numerical methods. (Received July 6, 1956.)

23. On Some Non-parametric C-sample Tests, Fred C. Andrews, University of Nebraska.

The Q test statistic (proposed by Terpstra, *Koninklijke Nederlandse Akademie van Wetenschappen*, Series A, v. LVII, 1954, p. 505) is shown to be asymptotically the sum of two quadratic forms Q_1 and Q_2, where Q_1 is a function only of $c-1$ Mann-Whitney statistics and Q_2 is a function of linear combinations of Mann-Whitney statistics with coefficients depending upon the sample sizes. A c-sample test based upon the statistic Q_1 is discussed and is shown to have asymptotic relative efficiency one with respect to the Wallis-Kruskal H test. (Received July 9, 1956.)

24. An Asymptotically Distribution-free Multiple Comparison Method with Application to the Problem of n Rankings of m Objects,

IRENE ROSENTHAL and THOMAS S. FERGUSON, University of California at Berkeley.

The statistic known as Hotelling's T^2 is seen to be asymptotically distribution-free and thus to provide a multiple confidence ellipsoid for the joint population means from which multiple comparisons of the means may be made. If m objects are ranked in order by each of n judges, Friedman's test is used to test the hypothesis that the judges rank the objects at random. However, it is of interest usually to inquire which of the objects are ranked significantly higher than which others, in which case Hotelling's T^2 may be used to provide an asymptotically distribution-free multiple comparison. It is seen also that the computations involved in the multiple comparison are very simple. (Received July 9, 1956.)

25. Idempotent Matrices and Quadratic Forms in the General Linear Hypothesis,

FRANKLIN A. GRAYBILL and GEORGE MARSAGLIA, Oklahoma A. and M. College.

The important role that idempotent matrices play in the general linear hypothesis theory has long been recognized, but their usefulness seems not to have been fully exploited. The purpose of this paper is to state and prove some theorems concerning idempotent matrices and to point out how they might be used in linear hypothesis theory. Let $\chi'^2(p, \lambda)$ represent a non-central Chi-square variate with p degrees of freedom and with non-centrality λ. By an idempotent matrix B we will mean a square matrix such that $B = B'$ and $BB = B$. The following theorem is shown to be true. Let Y be an $n \times l$ random vector which has a multivariate normal distribution with mean equal to the $n \times l$ vector μ and with variance-covariance matrix V (positive definite). Also let B be an $n \times n$ matrix with rank equal to p, and $B_i(i = 1, 2, \cdots, k)$ be an $n \times n$ matrix with rank equal to p_i where $Y'BY = Y'B_1Y + Y'B_2Y + \cdots + Y'B_kY$. Then any one of the six conditions C_1, C_2, C_3, C_4, C_5, C_6, is necessary and sufficient that the $Y'B_iY$ be independently distributed as $\chi'^2(p_i, \lambda_i)$ where $\lambda_i = \frac{1}{2}\mu'B_i\mu$. C_1; BV be idempotent and $p_1 + p_2 + \cdots + p_k = p$. C_2; BV and each B_iV be idempotent. C_3; BV be idempotent and $B_iVB_j = \varphi$ (where φ is the null vector or the null matrix) for all $i \neq j$. C_4; $Y'BY$ be distributed as $\chi'^2(p, \lambda)$ and $p_1 + p_2 + \cdots + p_k = p(\lambda = \frac{1}{2}\mu'B\mu)$. C_5; $Y'BY$ be distributed as $\chi'^2(p, \lambda)$ and B_iV be idempotent (where $\lambda = \frac{1}{2}\mu'B\mu$). C_6; $Y'BY$ be distributed as $\chi'^2(p, \lambda)$ and $B_iVB_j = \varphi$ for all $i \neq j$. From this general theorem many important special cases can be derived, viz., if $B = V = I$ (the identity matrix) and if $\mu = \varphi$, then C_1 is the well known Cochran-Fisher theorem on quadratic forms for normal independent variables. (Received July 9, 1956.)*

26. Some Asymptotic Results on Wald's Approximate Classification Statistic,

M. IQBAL, University of North Carolina.*

In his paper "On a Statistical Problem Arising in the Classification of an Individual into One of Two Groups" published in the *Annals of Mathematical Statistics* (1944), Wald proposed a statistic U for use in classification procedure. In an attempt to find its exact sampling distribution, he ended up with a joint distribution of three variables, m_1, m_2 and m_3, and showed that nm_3 can be taken as an approximate classification statistic. In the first part of this paper, an asymptotic series is obtained for m_3 by starting with the joint distribution of m_1, m_2, and m_3 in the degenerate case $\rho_i = 0 = \zeta_i$, and, by noticing that the region of integration for m_1 and m_2 for fixed m_3 is the one enclosed by two hyperbolas in the (m_1, m_2) plane. In the second part of the paper, asymptotic moments for nm_3 are obtained

* Research under Office of Naval Research Contract No. NR-042-031.

by using the fact that m_1, m_2, and m_3 are of order $1/n$ in the probability sense, and the corresponding asymptotic distributions are obtained both for even and odd values of p. (Received July 9, 1956.)

26a. On Infinitely Divisible Random Vectors, Meyer Dwass, Northwestern and Stanford Universities, and Henry Teicher, Purdue and Stanford Universities.

A normally distributed random vector X is well known to be representable by $A - Y$ (in the sense of having identical distributions) where A is a matrix of constants and Y is a random vector whose component random variables are independent. A necessary and sufficient condition for any infinitely divisible to be so representable is given. The limiting case is discussed as are connections with the multivariate Poisson distribution and stochastic processes. (Received July 10, 1956.)

26b. A Further Contribution to the Theory of Univariate Sampling on Successive Occasions, (Preliminary Report), B. D. Tikkiwal, University of North Carolina and Karnatak University.

The general theory of Sampling on Successive Occasions for a single variate has been studied independently by Patterson (J.R.S.S., 1950, B 12) and the author (Theory of successive sampling. Unpublished thesis submitted towards partial fulfilment of requirements for Diploma, I.C.A.R., New Delhi, 1951). Both the above authors obtained the variance of the best linear unbiassed estimator under the assumptions that the weight θ_h and the regression coefficient occurring in the estimator are not calculated from the sample, but are known in advance. Now, the variance of the estimator has been obtained in this paper without these assumptions. It is shown, that, when the common units on tth occasion is the sub-sample of n''_{t-1}, the new units on $(t - 1)$th occasion for $t = 3, \cdots h$, the variance of the estimator on the h-th occasion is given by $E(\hat{\phi}_h \sigma_h^2 / n''_h)$, where $\hat{\phi} h$ is the estimate of the weight ϕ_h. A consistent estimator of the variance-expression has also been obtained. It is further shown, that, the modification, suggested by Narain (*J. Ind. Soc. Agric. Stat.* 1953) in the weighting procedure given by Patterson and the author, results in the increase of the variance of the estimator and thus making it less efficient. (Received July 10, 1956.)

26c. Invariance, Sequential Decision Functions, and Continuous Time Processes, Professor J. Kiefer, Cornell University.

It is shown that in sequential decision problems where certain conditions on a sequence of sufficient statistics hold, there exist fixed sample-size invariant procedures which are minimax in the class of all decision functions. Examples are problems of sequential estimation of an unknown real scale and/or location parameter for the normal, gamma, and rectangular distributions (excepting the location parameter alone in the latter), certain multivariate estimation problems for these distributions, etc. The method applies also to estimation problems for the Wiener and gamma processes in one or several dimensions, and a modification of the result is obtained for processes such as the Poisson process where there is an invariance in time. (Some of these results were obtained for special weight functions by various authors using the Bayes or Cramér-Rao techniques.) The method of proof uses an invariance theorem which is a slight generalization of one due to Peisakoff and which may also be applied to many nonsequential problems. Various sets of conditions on the weight function and group, under which this theorem holds, are given. (Received July 12, 1956.)

26d. A Stochastic Model for the Tunnelling and Retunnelling of the Flour Beetle,

MOHAMED S. AHMED, University of California at Berkeley.

The stochastic model developed to describe the tunnelling and retunnelling of the flour beetle is a Markov chain with only three states: s_1, tunnelling, s_2, stationary, and s_3, retunnelling, and with continuous time parameter and stationary transition intensities $q_{ij}(i, j = 1, 2, 3)$. Transition from one state to any other is visualized with the restriction that the beetle cannot move from the tunnelling state to the retunnelling state, or vice-versa, except by passing through the stationary state.

For this model: 1) the probability that the beetle is in state s_j at the end of the time interval t given that it was in state s_i at the beginning of this time interval, as well as 2) the expected time spent in a state s_i out of the total exposure time T, is found explicitly in terms of the four unknown parameters q_{12}, q_{21}, q_{23} and q_{32}. In addition, a scheme for the estimation of these parameters is given.

This model can be used to study the differences in the behavior of, say, the male and the female beetles with regard to the proportion of time they spend in the various states. (Received July 13, 1956.)

26e. The *r*th Brightest Star in a Galaxy as a Distance Indicator,

MANDAKINI SANE, University of California at Berkeley.

The astronomical problem mentioned in the title reduces to that of estimating the location parameter θ of an absolutely continuous distribution of a r.v. X from the rth smallest observation X_r in a sample of size N. Three cases are considered: (i) N fixed, moderate, (ii) N fixed but large, (iii) N random tending to $+\infty$ in law.

Principal results are (a) the three Smirnov's limit laws for X_r hold if the random N tends to $+\infty$ suitably. (b) In two of Smirnov's limiting cases with observables $X_1, X_2, \cdots, X_r$ the last, X_r is asymptotically sufficient for θ. In the third case X_1 is asymptotically sufficient. Normal and Cauchy distributions belong to the first two types. Normal distribution truncated from left belongs to the third type. (c) Given an integer p, $0 < p < r$, the method of best linear unbiased estimates can be used to determine those p of the r available observables $X_1, X_2, \cdots, X_r$, which yield the least variance. (Received July 13, 1956.)

26f. Effect of Expansion of the Universe on the Serial Correlations of Counts of Images of Galaxies in Regularly Spaced Squares—A Simplified Model,

MARTIN FOX, University of California at Berkeley.

The Neyman-Scott theory of clustering of galaxies in a static universe (*Ap. J.*, 1952), and Neyman's extension to an expanding universe (*Ann. Inst. Henri Poincaré*, 1955) are so complex that non-numerical study of the effect of expansion on serial correlations of counts of galaxies seems prohibitive. In order to obtain an insight into the situation it is justifiable to consider a simpler although less realistic model. We assume: (i) Galaxies occur in clusters. (ii) Cluster centers are Poisson distributed. (iii) Clusters appear as equal and similarly oriented squares perpendicular to the line of sight. (iv) Clusters are visible up to a limiting distance. (v) The "number of images of galaxies" in a cell on a photograph equals the sum of areas of projections of all "visible" clusters overlapping this cell. (vi) If the universe is expanding, the velocity of recession of clusters is directly proportional to their distance from the observer.—The present model differs from earlier models by the contents of hypotheses (iii), (iv) and (v).—Under the present model, explicit formulae are given for the serial correlation between "numbers of images of galaxies" in regularly spaced equal squares. These correlations depend upon the dimensions of the squares, the limit of "visibility" of clusters, and whether the universe is expanding. (Received July 16, 1956.)

26g. Contributions to Univariate and Multivariate Components of Variance Analysis, S. N. Roy and R. Gnanadesikan, University of North Carolina.

Assume a 'normal' univariate linear hypothesis model which involves m unknown parameters in K groups: $\xi_{11}, \cdots, \zeta_{1m_1}; \xi_{21}, \cdots, \xi_{2m_2}; \cdots; \xi_{K1}, \cdots, \xi_{Km_K}$ (with $\sum_i m_i = m$), denoted by σ^2 the common unknown variance of the observations and by $(S.S.)_i$ the sum of squares due to $H_{0i}: \xi_{11} = \cdots = \xi_{im_i} (i = 1, 2, \cdots k)$ and assume that the design matrix is such that σ^2 $(S.S.)_i$'s are independent χ^2's with d.f. $(m_i - 1)(i = 1, 2, \cdots k)$. Then switching over to the case where the first m_1 parameters are independent, $N(\mu_1, \sigma_1^2)$, the second m_2 are independent $N(\mu_2, \sigma_2^2)$, and also independent of the first m, and so on, and all independent of the "error" normal variates it is possible to obtain, with a joint confidence coefficient $\geqq 1 - \alpha$, simultaneous confidence bounds on $\sigma_1^2 \mid \sigma_1^2\sigma_2^2 \mid \sigma^2, \cdots, \sigma_K^2 \mid \sigma^2$. Multivariate extensions of this involving the roots of $\Sigma_1\Sigma_2^{-1}, \cdots, \Sigma_K\Sigma^{-1}$ is discussed next. What happens for the case of design matrices for which independent χ^2's are not available (which would mean that a similar simplification for the multivariate problem would also not be available) is also discussed. (Received July 16, 1956.)

26h. Further Contributions to Multivariate Confidence Bounds, S. N. Roy and R. Gnanadesikan, University of North Carolina.

Assuming S to be the dispersion matrix of a random sample of size n from a p variate normal population (with $p < n$) having a dispersion matrix Σ and denoting by $S^{(i)}$, $\Sigma^{(i)}$, $S^{(i,j)}\Sigma^{(i,j)}$, etc. the corresponding matrices obtained by cutting out the ith variate, the ith and the jth variate, and so on, it is possible to find constants $\lambda_{1\alpha}$ and $\lambda_{2\alpha}$ such that there exist with a joint confidence coefficient $\geqq 1 - \alpha$, confidence bounds

$$\lambda_{1\alpha} C(S)_{\max} \leqq C(\Sigma)_{\max} \leqq \lambda_{2\alpha} C(S)_{\max}, \qquad \lambda_1 \alpha C(S)_{\min} \leqq C(\Sigma)_{\min} \leqq \lambda_{2\alpha} C(S)_{\min};$$

$$\lambda_{1\alpha} C(S^{(i)})_{\max} \leqq C(\Sigma^{(i)})_{\max} \leqq \lambda_{2\alpha} C(S^{(i)})_{\max}, \qquad \lambda_{1\alpha} C(S^{(i)})_{\min} \leqq C(\Sigma^{(i)})_{\min} \leqq \lambda_{2\alpha} C(S^{(i)})_{\min};$$

$$\lambda_{1\alpha} C(S^{(i,j)})_{\max} \leqq C(\Sigma^{(i,j)})_{\max} \leqq \lambda_{2\alpha} C(S^{(i,j)})_{\max};$$

$$\lambda_{1\alpha} C(S^{(i,j)})_{\min} \leqq C(\Sigma^{(i,j)})_{\min} \leqq \lambda_{2\alpha} C(S^{(i,j)})_{\min}; \qquad \text{and so on.}$$

$C_{\max}(M)$ and $C_{\min}(M)$ stand for the largest and smallest characteristic roots of a square matrix with non negative characteristic roots $C(M)$ (say). With slight modifications similar results are true for $C(\Sigma_1\Sigma_2^{-1})$ in relation to $C(S_1S_2^{-1})$ in the case of two samples and two populations, and in the case of the regression matrix of p variates on q variates, and also in the case of the regression matrix which measures the deviation from the customary multivariate linear hypothesis on means. (Received July 16, 1956.)

(Additional abstract for the Princeton meeting of the Institute, April 20–21, 1956)

27. Incomplete Block Rank Analysis: 2^{p-q} Fractional Factorials Using a Method of Paired Comparisons, Otto Dykstra, Jr., General Foods Corporation, (introduced by C. Daniel).

The estimation and tests of significance for main effects and interactions in 2^{p-q} fractional factorials are developed and illustrated using the 2×2 factorial used by Abelson and Bradley. The results are comparable. The procedure for handling 4-level factors is also discussed. An example shows the partitioning of the 3 degrees of freedom for a 4-level quantitative variable into linear, quadratic, and cubic effects. (Received April 30, 1956).

(Abstracts of papers submitted for the Detroit meeting of the Institute, September 7–10, 1956)

28. Some Results on the Analysis of Random Signals by Means of a Cut-Counting Process, Irwin Miller, Graduate Fellow, and John E. Freund, Virginia Polytechnic Institute.

The variance of the number of zeros of a Gaussian signal on a short time interval was derived in a recent paper by Steinberg, *et al.* This result is generalized to include the covariance of the number of zeros of a Gaussian signal at the values θ_1 and θ_2, using a somewhat different mathematical approach. A test of normality in a random process is proposed, using the general result. (Received May 21, 1956.)

29. Some Results on the Distribution of the Peaks of a Gaussian Process, Irwin Miller, Graduate Fellow, and John E. Freund, Virginia Polytechnic Institute.

The behavior of the relative extrema (peaks) of a Gaussian random process, $x(t)$, inside the band $\theta_1 \leqq x(t) \leqq \theta_2$ is discussed in this paper. The expected number and the variance of such peaks is derived. The method of derivation uses the concept of functionals, which reduces the mathematics involved and adds to the intuitive appeal of the argument. The results are specialized to include half-infinite and infinite band widths. These results will be useful in studying problems of gust loads and turbulence which arise in aircraft design. (Received May 21, 1956.)

30. A Continuous Time Treatment of the Waiting-time in a Queueing System Having Poisson Arrivals, a General Distribution of Service-time, and a Single Service Unit, (Preliminary Report), Václav Edvard Beneš, Bell Telephone Laboratories.

Let $W(t)$ be the time that a customer would have to wait if he arrived at time t; $W(t)$ is a mixed-type Markov process, *i.e.*, both discontinuous and continuous changes occur in $W(t)$. The integro-differential equation for the distribution of $W(t)$ is solved by the use of Laplace and Laplace-Stieltjes transforms. Let $P(t) = \text{pr}\,\{W(t) = 0\}$. The Laplace-Stieltjes transform of $W(t)$ is expressed as a functional of $P(t)$; the Laplace transform of $P(t)$ is determined to be $\eta^{-1}\varphi(\eta)$, where φ is the Laplace-Stieltjes transform of $W(0)$, and where $\eta(\tau)$ is the unique root of the equation $\tau - \eta + \lambda = \lambda\beta^*(\eta)$, in which λ is the Poisson arrival rate, and B^* is the Laplace-Stieltjes transform of the service-time distribution. The transform of the time taken to return to 0 from $W(0)$ is $\varphi(\eta)$. It is shown that any analytic function of the root η can be expanded in a Burmann-Lagrange series. These results are used to verify the asymptotic properties of $W(t)$. A functional relation is obtained between the expectation $E\{W(t)\}$ and $P(t)$; from this relation the covariance function R of $W(t)$ is determined, when R is defined, by means of the solution for $P(t)$. It is shown that if the service-time distribution has a finite fourth moment, then R is absolutely integrable, and the spectral distribution of $W(t)$ is absolutely continuous, with a continuous density. (Received June 5, 1956.)

31. On a Multivariate Tchebycheff Inequality, (Preliminary Report), Ingram Olkin and John W. Pratt, University of Chicago.

Let x be a random p-vector with mean 0, variances $1/k^2$, and correlation matrix R. A matrix A is admissible if $xAx' \geqq 1$ for $x \varepsilon S = \{x\colon \text{some } |\, x_i \,| \geqq 1\}$, $xAx' \geqq 0$ for all x. Then for A admissible, $P(x \varepsilon S) \leqq E(xAx') = \text{tr } AR/k^2$. Let $A^{-1} = (1 - t) I + te'e$ where

$$e = (1, \cdots, 1), \qquad -(p-1)^{-1} < t < 1.$$

Then A is admissible and minimizing tr AR over t yields the sharpest inequality for these A. This generalizes the bivariate inequality of Berge (*Biometrika*, 29(1937)); for the admissible A minimizing tr AR, xAx' has a minimum of 1 on each plane $x_i = 1$, and therefore $A^{-1} = B$ has unit diagonal. Since tr $B^{-1}R$ is a strictly convex function of B, it has a unique minimum determined by the condition that $BR^{-1}B$ be diagonal. B reflects certain properties of R. If $R^{\frac{1}{2}}$ has diagonal elements d, $B = d^{-1}R^{\frac{1}{2}}$, the bound is d^2p and is achieved if each row of $\pm B$ has probability $d^2/2$ and 0 has probability $1 - d^2p$. If $R = (1 - \rho) I + \rho e'e$, the bound is $[(p-1)\sqrt{1-\rho} + \sqrt{1+(p-1)\rho}]^2/pk^2$, which reduces to the univariate Tchebycheff inequality for $\rho = 1$. For p uncorrelated variables the bound is pk^{-2}, whereas the Tchebycheff bound for p independent variables is $1 - (1 - k^{-2})^p$. A simple transformation generalizes all these results to arbitrary variances and sets $\{x\text{: some } | x_i | \geqq k_i\}$. (Received June 15, 1956.)

32. Unbiased Estimation of Correlation Coefficients, Ingram Olkin and John W. Pratt, University of Chicago.

Unbiased estimators of certain multivariate normal correlation coefficients are given, namely: (1) bivariate correlation; (2) intraclass correlation, i.e., the common correlation coefficient of a distribution with equal variances and equal covariances; (3) partial correlation; (4) squared multiple correlation. In each case, the estimator is a function of a complete sufficient statistic and is therefore the unique minimum variance unbiased estimator. It is a strictly increasing function of the usual estimator differing from it only by terms of order $1/n$ and consequently having the same asymptotic distribution. The range of the unbiased estimator is the region of possible values of the estimated quantity, except in (4), where any unbiased estimator must assume negative values. The underlying method is that of inverting Laplace transforms. For (1) and (2), the unbiased estimators are $G(r) = rF(\frac{1}{2}, \frac{1}{2}; (n-1)/2; 1-r^2)$ and $H(r') = -H(-r') = 2F(1, (2-n)/2; n/2; (1-r')/(1+r')) - 1$, $r' \geqq 0$, where n is the number of degrees of freedom, r and r' are the usual estimates, and F is the hypergeometric function. Tables of these functions are given. (Received June 15, 1956.)

33. Unbiased Estimation of the Normal Distribution Function, (Preliminary Report), William C. Healy, Jr., Ethyl Corporation Research Laboratories.

Let X be normally distributed with distribution function $\Phi(x; \theta, \sigma^2) = \Pr(X \leqq x \mid \theta, \sigma^2)$, θ and σ^2 being unknown. The problem considered is the estimation of Φ for a given value of x, from a sample $X_1, X_2, \cdots, X_n$. Possible estimates include (a) m/n, where m is the number of observations X_i not exceeding x; (b) $\Phi(x; \bar{X}, s^2)$, where $\bar{X} = \sum X_i/n$ and $s^2 = \sum(X_i - \bar{X})^2/(n-1)$; and (c) plots on probability paper. Baker (*Ann. Math. Stat.*, Vol. 20 (1949), p. 123) has given asymptotic properties of (b) and a comparison with (a), but an estimate optimal in small samples seems not to have been discussed. The present paper derives the uniformly-minimum-variance-unbiased estimate of Φ, by computing the conditional probability $\Pr(X_1 \leqq x \mid \bar{X}, s^2)$. The result is $\tilde{\Phi}(x) = [1 + u \operatorname{sgn}(x - \bar{X})]/2$, where $u = I_v(\frac{1}{2}, (n-2)/2$, $v = \min[1, n(x - \bar{X})^2/(n-1)^2s^2]$, and $I_v(p, q)$ is the incomplete Beta-function ratio. This estimate $\tilde{\Phi}(x)$ and (b) are asymptotically equivalent. Empirical investigation of the sampling variance of $\tilde{\Phi}(x)$ is underway, for small-sample comparisons with other methods. (Received June 25, 1956.)

34. On the Construction of Fractional Factorial Designs, (Preliminary Report), ROBERT C. BURTON, National Bureau of Standards.

For a $1/s^p$ fraction of an s^n factorial design, the identity relationship contains $(s^p - 1)/(s - 1)$ "words." For example, if $s = 3$, $n = 6$, and $p = 2$, the factors may be denoted by $A, B, \cdots, G$ and a possible identity relationship is $I = ABC^2D^2 = CDEF = ABEF = ABCDE^2F^2$, where I is the identity element and ABC^2D^2, $CDEF$, $ABEF$, and $ABCDE^2F^2$ are the three words. Necessary and sufficient conditions are given for constructing an identity relationship containing words having prescribed numbers of letters of each power. The argument consists of regarding the words as sets of elements, and showing that they may be expressed as unions of certain disjoint sets. (Received July 9, 1956.)

NEWS AND NOTICES

Readers are invited to submit to the Secretary of the Institute news items of interest

Personal Items

The Mathematics Department of Catholic University is pleased to announce a lecture series in the field of Mathematical Statistics. The lecture series will start with the beginning of the academic year 1956–57. It is planned to invite prominent mathematical statisticians from the East Coast or visiting the East Coast to give addresses. The cooperation of other universities and interested organizations in the area will be sought. Enquiries concerning the lecture series may be addressed to the Department of Mathematics of Catholic University.

Dr. Raymond H. Burros has accepted an appointment as operations research analyst with Technical Operations, Inc. and is attached to the Combat Operations Research Group, Fort Monroe, Virginia. During the academic year 1954–1955, Dr. Burros was visiting associate professor of psychology in the University of Houston. In the fall of 1955 he held a temporary appointment as director, Validation Department, Institute for Motivational Research, Inc.

Herbert T. David, lecturer in the Committee on Statistics of the University of Chicago, has accepted a position at Iowa State College, beginning in September 1956, as assistant professor in the Statistical Laboratory and the Department of Statistics for teaching, research and consulting work in industrial statistics.

Richard De Lancie, now a partner in Broadview Research and Development, has returned to the home office in Burlingame, California, after establishing a Washington, D. C. office for the firm.

James R. Duffett, formerly at White Sands Proving Ground, is now at Radio Plane Company, Van Nuys, California.

The Copley Medal has been awarded by the Royal Society to Sir Ronald Fisher, F.R.S., Arthur Balfour Professor of Genetics, Cambridge University, for contributions to developing the theory and application of statistics for making quantitative a vast field of biology.