ABSTRACTS OF PAPERS

(*Abstracts of papers presented at the Central Regional Meeting of the Institute, Urbana, Illinois, November 24–25, 1961. Additional abstracts appeared in the December, 1961 issue.*)

8. Completeness and the Existence of Unbiased Tests. C. B. BELL, San Diego State College and the University of Washington. (By title)

In investigating extensions of the Lehmann unbiasedness criterion for two-sample tests' and, independently, the concept of completeness, one finds the following relations between the existence of unbiased test functions (UTF) and the classes of cpf's involved. *Theorem* A. If $\mathcal{G}$ is a translation-invariant cone of functions, then (a) implies (b); and (b) implies both (c) and (d), where (a) ω and $\Omega - \omega$ are totally complete with respect to (wrt) $\mathcal{G}_b$, the class of bounded functions of $\mathcal{G}$, and the null classes of ω and $\Omega - \omega$ are equal; (b) every UTF in $\mathcal{G}$ is essentially constant [Ω]; (c) every UTF in $\mathcal{G}$ is of constant power; and (d) Ω is complete wrt $\mathcal{G}_b$. *Corollary* B. For non-randomized tests the last three conditions are (b′) every unbiased critical region (UCR) with indicator function in $\mathcal{G}$ is similar with probability 0 or 1; (c′) every UCR with indicator function in $\mathcal{G}$ is similar; and (d′) every similar region has probability 0 or 1. By noting that no nondegenerate class of power cpf's is complete one establishes *Corrolary* C. For arbitrary ω and $\Omega - \omega$ there exists a nonconstant unbiased test for any random sample of size $n \geqq 2$.

9. Estimating the Parameters of Negative Exponential Populations from One or Two Order Statistics. H. LEON HARTER, Wright-Patterson Air Force Base.

This paper discusses the use of order statistics in estimating the parameters of (negative) exponential populations. For the one-parameter exponential population, the best linear unbiased estimators, $\tilde{\sigma}_k = c_k x_k$ and $\tilde{\sigma}_{lm} = c_l x_l + c_m x_m$, of the parameter σ are given, based on one order statistic x_k and on two order statistics x_l and x_m. For samples of any size up through $n = 100$, a table is given of k, l, and m and of the coefficients c_k, c_l, and c_m, together with the coefficients of σ^2 in the variances V_k and V_{lm} of the estimators, and the corresponding efficiencies E_k and E_{lm} (relative to the best linear unbiased estimator based on all order statistics). For the two-parameter exponential population, the best linear unbiased estimators, $\tilde{\alpha} = c_{\alpha l} x_l + c_{\alpha m} x_m$, $\tilde{\sigma} = c_{\sigma l} x_l + c_{\sigma m} x_m$, and $\tilde{\mu} = c_{\mu l} x_l + c_{\mu m} x_m$, of the parameters α and σ and the mean $\mu = \alpha + \sigma$ are given, based on two order statistics x_l and x_m. For samples of any size up through $n = 100$, a table is given of m (l is always 1 for the best estimator) and of factors c_α and c_σ for computing the coefficients $c_{\alpha l} = 1 + c_\alpha$, $c_{\alpha m} = -c_\alpha$, $c_{\sigma l} = -c_\sigma$, $c_{\sigma m} = c_\sigma$, $c_{\mu l} = 1 + c_\alpha - c_\sigma$, and $c_{\mu m} = c_\sigma - c_\alpha$, together with the coefficients of σ^2 in the variances $V_{\tilde{\alpha}}$, $V_{\tilde{\sigma}}$, and $V_{\tilde{\mu}}$ of the estimators, and the corresponding efficiencies $E_{\tilde{\alpha}}$, $E_{\tilde{\sigma}}$, and $E_{\tilde{\mu}}$ (relative to the best linear unbiased estimators based on all order statistics).

10. Estimation of Variance Components in the Two-Way Classification, Eisenhart's Model II (Preliminary report). LEONE Y. LOW, University of Illinois.

Several methods of Analysis of Variance are used for observations with the model $y_{ij} = \mu + b_i + t_j + e_{ij}$ where μ is constant, and b_i, t_j and e_{ij} are distributed normally, independently, with zero mean, and variances respectively σ_b^2, σ_t^2 and σ^2. Each of the methods yields unbiased estimates which are linear combinations of four quadratic forms. The variances and covariances were obtained. An expression giving all unbiased estimators based on the four quadratic forms is shown, for each variance component, to consist of linear combinations of two estimators. Since the minimum variance estimator of this type is a function of the unknown variance components, the error obtained by using another estimator is discussed. If the design has proportional numbers in the subclasses, there is a unique estimator for each component. Some inequalities involving the variance components are obtained for the balanced incomplete block estimators.

11. A Generalized Partially Balanced Association Scheme. J. N. SRIVASTAVA, University of North Carolina. (By title)

Partially balanced designs were defined by Bose and Nair and the association scheme possessed by them and the corresponding linear associative algebras have been studied by Bose and Mesner. This paper generalizes the concept of a partially balanced association scheme to many dimensions. Its properties have been studied by considering the corresponding linear algebras. Its use in the inversion of a large class of patterned matrices has been discussed. This scheme includes most of the important known association schemes as particular cases. It appears to have great potentialities for the development of new designs, and also in the analysis of known designs. It has been used to develop multidimensional partially balanced designs, and also the theory of fractional replications.

12. Some New Incomplete Factorial Designs. STEPHEN R. WEBB, University of Chicago. (Introduced by William Kruskal.)

An exhaustive search of possible three-quarter replicates of a 2^4 design has revealed that there are nineteen essentially different design configurations, fifteen of which permit estimation of the mean, main effects, and two-factor interactions. Two have been previously discussed in the literature. Two new designs seem to be of practical value. One has variances of $(15/32)\sigma^2$ for main effects and $(1/2)\sigma^2$ for two-factor interactions, and the other has variances of $(7/16)\sigma^2$ for the main effects, but the interactions have somewhat higher variances. An incomplete 2^7 design with 29 runs has been found in which all main effects and two-factor interactions can be estimated with variance $0.201\ \sigma^2$. A modification with 36 runs allows estimates of the main effects with variance $0.138\ \sigma^2$ and of two-factor interaction with variance $0.190\ \sigma^2$.

13. Limiting Behavior of a Sequence of Density Ratios. SUNARDI WIRJOSUDIRDJO, Bandung Institute of Technology, Indonesia.

Let the sequence $\{X_n\}$ of random variables have the following properties: (i) $X_n = f_n(Z_1, \cdots, Z_n)$ where f_n is a symmetric Baire function of n variables, and $Z_1, Z_2, \cdots$ is a sequence of independent and identically distributed random variables; (ii) the family $\mathcal{P}$ of distributions of $X_1, X_2, \cdots$ is indexed by $\theta \in \Theta$, Θ an ordered set; (iii) $\mathcal{P}$ is a homogeneous monotone likelihood ratio family on the σ-field $\mathcal{A}_n$ generated by $X_1, \cdots, X_n$; (iv) X_n is sufficient on $\mathcal{A}_n$. Let $\theta_1 < \theta_2$, let $p_{\theta_i}^{(n)}$ be the corresponding joint densities of $X_1, \cdots, X_n$ with respect to some σ-finite measure, and $R_n = p_{\theta_2}^{(n)}/p_{\theta_1}^{(n)}$ the corresponding density ratio. The following results are obtained: R_n converges to 0 or to ∞ according as $\theta \leqq \theta_1$ or $\geqq \theta_2$. For $\theta_1 < \theta < \theta_2$ $\liminf R_n = 0$ or $\limsup R_n = \infty$ a.e. P_θ, except possibly for one value of θ, say θ_0. If there is such an exceptional θ_0, then $\lim R_n = 0$ or ∞ a.e. P_θ according as $\theta < \theta_0$ or $> \theta_0$. In the case that $\{X_n\}$ is a sequence of non-central t-ratios, David and Kruskal (*Ann. Math. Stat.*, Vol. 27 (1956), pp. 797–804) have shown (implicitly) that there exists θ_0 such that $\lim R_n = 0$ or ∞ a.e. P_θ according as $\theta < \theta_0$ or

$> \theta_0$. We show that θ_0 is between 0 and $(\theta_1 + \theta_2)/2$. Furthermore we show lim inf $R_n = 0$ and lim sup $R_n = \infty$ a.e. P_{θ_0}.

(Abstracts of papers presented at the Eastern Regional Meeting of the Institute, New York City, December 27–30, 1961. Additional abstracts appeared in the December, 1961 issue.)

9. Application of the Tandem Test to a Problem of Two Populations (Preliminary report). LEE R. ABRAMSON, Columbia University.

Given two independent random variables with distributions F_x and F_y, we test H_1: $F_x = F_1$, $F_y = F_2$ vs. $H_2 : F_x = F_2$, $F_y = F_1$, where F_1 and F_2 are known distributions. This paper studies Bayes sequential procedures. A tandem test is a sequence of two Wald tests, the first performed on one of X or Y and the second on the other. Restricting ourselves to decision rules which sample first from one random variable, then from the other, and then make a decision, a heuristic argument leads to the replacement of the restricted Bayes decision rule by a tandem test. Using Wald's approximations (which do not affect our asymptotic results), we find the risk of an arbitrary tandem test and minimize it for small c, the common cost of sampling. From a result of Chernoff's (*Ann. Math. Stat.*, Vol. 30 (1959), pp. 755–770), we find that the minimizing tandem test has asymptotically (as $c \to 0$) minimal risk in the class of *all* decision rules. Moreover, if the Kullback-Leibler information numbers of X and Y are different, the minimizing tandem test has smaller asymptotic risk than a Wald test on either X or Y alone. If the information numbers are equal, the tandem test reduces to a Wald test.

10. The Choice of the Degree of a Polynomial Regression as a Multiple Decision Problem. T. W. ANDERSON, Columbia University.

On the basis of a sample of observations, an investigator wants to determine the appropriate degree of a polynomial in the index, say time, to represent the regression of the observable variable. This multiple decision problem is formulated in terms used in the theory of testing hypotheses. Given the degree of polynomial regression, the probability of deciding a higher degree is specified and does not depend on what the actual polynomial is (except its degree). Within the class of procedures satisfying these conditions and symmetry (or two-sidedness) conditions, the probabilities of correct decisions are maximized. The optimal procedure is to test in sequence whether coefficients are 0, starting with the highest (specified) degree. The procedure holds for other linear regression functions when the independent variates are ordered. The problem and its solution can be generalized to the multivariate case and to other cases with a certain structure of sufficient statistics.

11. Some Confidence Bounds for Determinantal Roots. T. W. ANDERSON, Columbia University. (Invited paper.)

The main problem considered is to find optimal simultaneous confidence bounds for the roots of $|\Sigma_1 - \gamma\Sigma_2| = 0$, where Σ_1 and Σ_2 are the covariance matrices of two multivariate normal distributions, or equivalently bounds for all ratios of quadratic forms, $x'\Sigma_1 x/x'\Sigma_2 x$. It seems reasonable to base the confidence bounds on the smallest and largest corresponding sample roots, that is, roots of the above equation with sample covariance matrices. The shortest such bounds are obtained. The method can be applied to other multivariate problems. It is shown how the bounds on various roots can be used to give bounds on the variances and covariances as well as other relevant functions of the parameters.

12. The Auto-Covariance of Discrete Pulse Trains. RALPH H. BACON, General Precision, Inc.

Expressions for the auto-covariance as a function of the pulse characteristics, and of the delay (or shift), and for the variance of the auto-covariance as a function of sample size, are derived. The Poisson process, and a family of processes related to the Poisson, are treated as illustrative examples.

13. A Normal Approximation to the Beta and Gamma Distributions. R. R. BAHADUR, University of Chicago.

Under suitable quadratic or cubic changes in the variable of integration, the incomplete beta and gamma functions resemble integrals with respect to the normal density function. The method yields certain normal approximations to the beta and gamma distributions, together with bounds for the errors of approximation. These results apply, through known identities, to distributions such as the binomial, negative binomial, Poisson, chi-square, and F. It will suffice to state here the main result concerning the binomial distribution. Let $B = \sum_{r=k}^{n} \binom{n}{r} x^r (1-x)^{n-r}$, where $1 < k < n$, and $0 < x < 1$. Writing $a = k - 1$, $b = n - k$, $c = a + b$, $h = (c/ab)^{\frac{1}{2}}$, and $\alpha = (a-b)h/3c$, let $y = (a - cx)h$, and $z = \alpha^{-1}\{(1 + 3\alpha y)^{1/3} - 1\}$ if $\alpha \neq 0$ and $z = y$ otherwise. Let $N = 1 - \Phi(z)$, where $\Phi(z) = \int_{-\infty}^{z} \phi(t)dt$ with $\phi(t) = (2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}t^2)$. Let $\gamma = (a^{-1} + b^{-1} - c^{-1})/12$, and $\delta = (1 + c^{-1})(1 + \alpha^2)(1 + \gamma)^{-1} - 1$. With $C = (1 + \delta)[N + (1 + \alpha^2)^{-1}(2\alpha + \alpha^2 z)\phi(z)]$, we then have $C - \delta < B < C$. We always have $0 < \delta < 3(a^{-1} + b^{-1})/16$, and $\delta \doteq (36a)^{-1}$ if b is very much larger than a.

14. Some Basic Theorems of Distribution-Free Statistics I. C. B. BELL, San Diego State College and University of Washington. (By title)

A mapping T of $\Omega \times R_n$ into R_1 is a statistic distribution-free with respect to (DF wrt)Ω, a class of cpf's on R_n if there exists cpf Q on R_1 such that for each F in Ω (i) $T_F(\cdot) = T(F, \cdot)$ is measurable, and (ii) $P_F T_F^{-1} = P_Q$. *Example.* $T_1(F^n, x) = (\bar{x} - \mu_F)\sigma_F^{-1}$ and $T_2(F^n, x) = h[F(x_1), \cdots, F(x_n)]$ are DF wrt the power measures of the normals and continuous cpf's, respectively. *Theorem* 1. There exists T DF wrt Ω and having cpf Q with $\sum p_i \epsilon_i$ and qF as its discrete and continuous parts, resp., iff for each F in Ω there exists a partition $\{A_0(F), A_1(F), \cdots\}$ of R_n with P_F nonatomic on $A_0(F)$ and $P_F(A_i(F)) = p_i (i \geqq 1)$. *Corollary* 2. If Ω is a class of continuous cpf's, then for each cpf Q on R_1 there exists T DF wrt Ω and having cpf Q. *Theorem* 3. If $\{T_i\}$ are independent and DF wrt Ω, then each measurable $g(T_1, \cdots, T_k)$ is DF wrt Ω. Generalizing the given examples one finds *Theorem* 4. $T(G_s, x) = f[s^{-1}(x)]$ is DF wrt $\Omega = \{G_s : s \varepsilon \mathcal{G}\}$, where $\mathcal{G}$ is a group of 1–1 transformations of R_n onto R_n and G_s is a cpf such that $P_{G_s} = P_G s^{-1}$.

15. Some Basic Theorems of Distribution-Free Statistics II. C. B. BELL, San Diego State College and University of Washington. (By title)

A measurable mapping T of R_n into R_1 is a statistic nonparametric and distribution-free with respect to (NPDF wrt) Ω, a class of cpf's on R_n, if there exists cpf Q on R_1 such that $P_F T^{-1} = P_Q$ for all F in Ω. *Example.* $T(x_1, \cdots, x_n)$ = number of x_i's > 0 is NPDF wrt $\Omega = \{F^n \cdot F(0) = \frac{1}{2}\}$; and $T(x_1, \cdots, x_n) = (x_1 - x_2)(x_3 - x_4)^{-1}$ is NPDF wrt the normal power measures. *Theorem* 1. There exists T NPDF wrt Ω and having cpf Q with $\sum p_i \epsilon_i$ and qF as its discrete and continuous parts, resp., iff there similar partition $\{A_0, A_1, \cdots\}$ of R_n

and a σ-sub-algebra $\mathfrak{D}$ of the Borels sets $\mathfrak{B}_n$ such that (i) $P_F(A_i) = p_i$ for $i \geqq 1$, and (ii) $A_0 \cap \mathfrak{D}$ is a nonatomic σ-ring of similar sets. *Corollary* 2. If there exists T NPDF wrt Ω and having continuous cpf Q, then for arbitrary cpf H on R_1 there exists V NPDF wrt Ω and having cpf H. *Corollary* 3. If T is independent of S, which is a sufficient statistic for Ω, then T is NPDF wrt Ω. *Theorem* 4. If $\{T_i\}$ are mutually independent and NPDF wrt Ω, then each $(T_1, \cdots, T_k)$ is NPDF wrt Ω.

16. Zero Crossings of the Gaussian Process. SIMEON M. BERMAN, Columbia University. (By title)

The result of Kac (1943) and Rice (1944) on the expected number of zeros of a continuous stationary Gaussian process whose spectral distribution function is a step function with a finite number of jumps is extended to the case of a general spectrum with a finite second moment. The method of proof used here is based on the work of Prokhorov (1956) on the weak convergence of measures in function space. A general weak convergence theorem for second order stationary processes is developed and applied to the distribution of the number of zero-crossings.

17. On the Distribution of Bilinear Forms in Normal Variables. THEOPHILOS CACOULLOS, Stanford University. (By title)

Classification procedures with normal alternative populations as well as procedures for comparing distances between p-variate normal populations (cf., Abstract of the author in the Sept. 1961 issue of the *Annals*) involve the distribution of statistics which can be reduced to the bilinear from $T = X'\Sigma^{-1}Y$ or $T^* = X'S^{-1}Y$, according, as the covariance matrix Σ is known or estimated by the sample covariance matrix S; each of the p-component random vectors X and Y has a p-variate normal distribution with covariance matrix Σ, $E(X) = \mu \neq 0$, $E(Y) = \nu \neq 0$, and the covariance matrix between X and Y is $\rho\Sigma$, $-1 < \rho < 1$; S is a $p \times p$ random positive definite matrix which has the Wishart distribution $W(S \mid \Sigma, p, n)$ with $n \geqq p$ degrees of freedom and is independently distributed of X and Y. The exact density function of T is derived in terms of a double series with terms which involve the Whittaker's confluent hypergeometric function. Extending a result of Bowker (*Contributions to Probability and Statistics*, Stanford University Press, pp. 142–149, 1960) concerning the case $\nu = c\mu$ (c constant), a representation of T^* is obtained in terms of simpler statistics.

18. Comparing Distances Between Multivariate Normal Populations, II (Preliminary report). THEOPHILOS CACOULLOS, Stanford University.

Let π_i be p-variate normal populations with mean μ_i, $i = 0, 1, \cdots, k$, respectively, and with the same known covariance matrix Σ. The μ_i, $i = 1, \cdots, k$ are unknown and μ_0 is known. Let $\Delta_{ij}^2 = (\mu_i - \mu_j)'\Sigma^{-1}(\mu_i - \mu_j)$ denote the generalized (Mahalanobis) distance between π_i and π_j. On the basis of a sample of size n from each π_j, $j = 1, \cdots, k$, a population π_j is to be selected so that $\Delta_{0j}^2 = \min(\Delta_{01}^2, \cdots, \Delta_{0k}^2)$. Let d_j be the decision of selecting π_j. Attention is restricted to decision rules which are invariant under a subgroup of affine transformations on kp-space and the symmetric group of permutations of the sample means $\bar{x}_1, \cdots, \bar{x}_k$. Then, under certain natural symmetry assumptions on the loss functions, the following unique invariant and symmetric Bayes decision rule is obtained: take decision d_j if $D_j^2 \equiv (\bar{x}_j - \mu_0)'\Sigma^{-1}(\bar{x}_j - \mu_0) = \min(D_1^2, \cdots, D_k^2)$. Incidentally, a more general result is shown. *Theorem.* If $x_1, \cdots, x_k$ are independently distributed with densities $f_{\theta_1}(x), \cdots, f_{\theta_k}(x)$, respectively, and $f_\theta(x)$ has the monotone likelihood ratio property in x, θ, then the (unique) symmetric Bayes rule for choosing the smallest parameter θ_j is: choose θ_j if the observation x_j on X_j is $x_j = \min(x_1, \cdots, x_k)$.

19. Exact Power of Some Tests Based on the Mann-Whitney-Wilcoxon *U* Statistic. I. M. CHAKRAVARTI, GEORGE E. HAYNAM AND F. C. LEONE, Case Institute of Technology.

A closed form expression for the distribution of two sample Mann-Whitney-Wilcoxon U statistic under the null hypothesis has been developed. This has been checked computationally against the recursive definition given by Mann and Whitney. There have been developed, by the use of similar techniques, expressions for the exact power of the U test under the alternatives of translation in the exponential population and changes of scale and location in the rectangular population. The power of the test has been tabulated for various values of the population parameters by evaluating the exact expression.

20. Applicability of the Extended Method of Parabolic Curves in the Analysis of Agricultural Data. REGINA C. ELANDT, Case Institute of Technology. (Introduced by Fred C. Leone.)

If in an experiment one suspects some kind of smooth systematic trend in uncontrolled variability, e.g., systematic trend in soil fertility in an agricultural experiment, it may well be useful to apply curvi-linear regression to remove this trend. The theory for this problem in the case of the systematic arrangement of treatments: ABCD, ABCD, $\cdots$ etc., was first presented by Neyman (1929) ("The theoretical basis of different methods of testing cereals"). A more convenient form (using the method of orthogonal polynomials corrected for the treatments) was given by Hald (1948) ("The decomposition of series of observations"). Another systematic balance arrangement: ABCD, DCBA, $\cdots$ etc., is considered by R. Elandt and tables of orthogonal polynomials for this case are calculated (not yet published).

If a systematic trend in two directions, e.g., influence of two uncontrolled factors, is observed, the method of parabolic curves in two-dimensional form can be applied. The general theory and formulae are given in this paper. The case of an agricultural experiment, in which the treatments are located systematically in long strips, is considered in detail. The efficiency of both one-and two-dimensional regression analyses in comparison to the method of randomized blocks has been investigated for agricultural uniformity trial data. Also, some results in oil-plant experiments, carried out in Poland, are presented.

21. On a Characterization of the Gamma Distribution (Preliminary report). THOMAS S. FERGUSON, University of California at Los Angeles.

Under certain conditions, or perhaps none at all, if X and Y are independent random variables, and if the sum $X + Y$ is independent of the ratio XY^{-1}, then both $\pm X$ and $\pm Y$ have gamma distributions, with a common scale parameter. R. G. Laha, *Ann. Math. Stat.* (1954) pp. 784–787, has demonstrated such a characterization of the gamma distribution under the conditions that X and Y be identically distributed with a finite second moment. The author proves this characterization under the conditions that X and Y be positive random variables with continuous densities on the positive real axis. A connection between this characterization of the gamma distribution and a similar characterization of the normal distribution is mentioned.

22. Fiducial Method: Its Consistency and a Suggested Modification. D. A. S. FRASER, Stanford University.

D. V. Lindley (*J. Roy. Stat. Soc.*, Ser. B, Vol. 20 (1958), pp. 102–107) and D. A. Sprott (*J. Roy. Stat. Soc.*, Ser. B, Vol. 22 (1960), pp. 312–318) have proposed some consistency criteria for fiducial probability distributions, and in these terms have investigated consistency for Koopman-Darmois models with a real variable and a real parameter. In this paper, the consistency is investigated without the Koopman-Darmois restriction and is found to hold if and only if the parameter is essentially a location parameter. Also, if fiducial distributions are to have a frequency interpretation and to satisfy consistency requirements, then the range of validity can be extended from the location parameter models by a suggested modification to the fiducial method of derivation.

23. On Solving a Markovian Decision Problem by Linear Programming. MARSHALL FREIMER, Institute for Defense Analyses, Cambridge, Mass.

For a non-discounted Markovian decision process, Howard (*Dynamic Programming and Markov Processes*, Wiley, 1960) derived the recurrence equations

$$v_i + g = \max_k \left(q_i^k + \sum_{j=1}^{N} p_{ij}^k v_j\right) \qquad i = 1, 2, \cdots, N,$$

for the average gain per stage g and the expected value v_i of being in state i, (q_i^k is the expected immediate return, and p_{ij}^k is the transition probability). Howard solved this problem by an iterative scheme. The solution can also be obtained by minimizing g subject to $v_1 + g - \sum_{j=1}^{N} p_{ij}^k v_j \geq q_i^k$ all i, k. The dual of this linear programming problem is to maximize $\sum_{i,k} P_i^k q^k$ subject to $P_i^k \geq 0$, $\sum_k P_j^k = \sum_{i,k} P_i^k P_{ij}^k$, and $\sum_{i,k} P_i^k = 1$. P_i^k may be interpreted as the probability of being in state i and using decision k. Considering just this dual problem, we can show that for any i only one P_i^k is positive. Similar results hold for the discounted process. For this case everything but the direct proof that only one $P_i^k > 0$ can be found in d'Epenoux (*Revue Francoise de Recherche Opérationnelle*, No. 14, 1960, pp. 3–15).

24. A Comparison of Minimum Variance Regression Coefficients with Weighted Least Squares Regression Coefficients. GENE H. GOLUB, Space Technology Labs, Inc. (Introduced by Ernest M. Scheuer.)

Let $y = \Phi\alpha + \epsilon$ where Φ is an $n \times p$ matrix of known constants and of rank p; α is a vector with p components which is to be estimated; and ϵ is a random vector of n components with $E(\epsilon) = 0$ and covariance matrix Σ. The minimum variance estimate of α is $\alpha^* = (\Phi'\Sigma^{-1}\Phi)^{-1}\Phi'\Sigma^{-1}y$. Sometimes Σ is not known or Σ^{-1} is not easily computed and consequently, α is estimated by its weighted least squares estimate $\hat{\alpha} = (\Phi'W\Phi)^{-1}\Phi'Wy$ where W is an $n \times n$ positive definite matrix of weights. Then given the eigenvalues of $F'\Sigma F$ where $F'F = W$ and using the matrix inequalities of Kantorovich, Wielandt and Schopf, it is possible to determine attainable inequalities which compare the covariance matrix of α^* with that of $\hat{\alpha}$, and the generalized variance of α^* with that of $\hat{\alpha}$. This extends results of T. Magness (*Ann. Math. Stat.*, to appear).

25. Statistical Method for the Mover-Stayer Model. LEO A. GOODMAN, University of Chicago. (By title)

The mover-stayer model can be described as follows: Let there by a finite number, I, of states with n_i individuals in state i at time zero ($i = 1, \cdots, I$). Each of the n_i individuals is either a "stayer" or a "mover". The stayers in state i remain with probability one in that state at time $t = 0, 1, \cdots, T$; the state transitions of each mover at $t = 0, 1, \cdots, T$ can be described by a Markov chain with (unknown) transition probability matrix M. Let s_i denote the (unknown) proportion of the n_i individuals who are stayers ($i = 1, \cdots, I$). Estimators of the parameters M and s_i were presented by Blumen, Kogan, and McCarthy

(*The Industrial Mobility of Labor as a Probability Process*, Ithaca, N. Y., Cornell University Press, 1955). In the present paper, we develop several modifications of their estimators, and derive alternative (more direct) estimators, including maximum likelihood estimators based on the data from the $n = \Sigma_i n_i$ series observed at $t = 0, 1, \cdots, T$. The accuracy of the various estimators is compared; e.g., it is shown that the Blumen-Kogan-McCarthy estimators are not consistent, while the estimators recommended herein are. The asymptotic variances of the estimators are computed. In addition, tests of various hypotheses concerning the mover-stayer model are developed; e.g., we present a test of the null hypothesis that $s_i = 0$.

26. Tests Based on the Movements in and the Comovements between m-Dependent Time Series. Leo A. Goodman, University of Chicago.

This paper presents methods for deriving tests for trend in an m-dependent time series $X = \{X_1, \cdots, X_T\}$ and tests of independence between an m_1-dependent series X and an m_2-dependent series $Y = \{Y_1, \cdots, Y_T\}$. A test for trend in X is presented based on the series $W = \{W_1, \cdots, W_S\}$ where $W_i = f(X_{2S+i} - X_i), f(a) = 1$ if $a > 0, f(a) = 0$ if $a < 0$, and $S = T/3$. (Assume T divisible by 3 and that $X_{2S+i} - X_i$ has a continuous distribution.) This test is a modification of the Cox-Stuart test (*Biometrika*, Vol. 42 (1955), pp. 80–95) for 0-dependent series. Tests of independence between X and Y are presented which eliminate the primary effects of trends in X and Y, and which are based on the series $U = \{U_1, \cdots, U_{T-k}\}$ and $V = \{V_1, \cdots, V_{T-k}\}$, where $U_i = f(X_{k+i} - X_i)$, $V_i = f(Y_{k+i} - Y_i)$, and k is a fixed integer. These are generalizations of tests proposed by Goodman (*Biometrika*, Vol. 46 (1959), pp. 525–32) and by Goodman and Grunfeld (*J. Amer. Stat. Assn.*, Vol. 56 (1961), pp. 11–26) for 0-dependent series. Various properties of these tests are studied; e.g., it is shown that the statistic appropriate for testing the null hypothesis of independence between X and Y when min $[m_1, m_2] = 0$ is asymptotically equivalent (under the null hypothesis) to the statistic appropriate when $m_1 = m_2 = 0$. Various measures of dependence between X and Y are also discussed, and confidence intervals for these measures are presented.

27. The First Two Movements of the Reciprocal of a Positive Hypergeometric Variable. Zakkula Govindarajulu, Case Institute of Technology.

The random variable X is said to have a positive hypergeometric distribution if X takes on the value m with probability $\binom{M}{m}\binom{N-M}{n-m}\Big/\left[\binom{N}{n} - \binom{N-M}{n}\right]$, $m = 1, 2, \cdots, n$. In sampling without replacement from finite populations many situations do arise in which one should know the expected value and the variance of $1/X$. Stephan (see *Ann. Math. Stat.*, Vol. 16 (1945) pp. 50–61) considers this problem and obtains expressions for the moments of $1/X$ in the form of infinite series, which do not yield computational ease. Let $a_i(N, M, n) = \sum_{m=1}^{n} m^{-i}\binom{M}{m}\binom{N-M}{n-m}\Big/\binom{N}{n}$, $i = 0, 1, 2$, and $b_i(N, M, n) = E[m^{-i}|N, M, n, m > 0] = a_i(N, M, n)|a_0(N, M, n)$, $i = 1, 2$. The $a_i(N, M, n)$, $i = 0, 1, 2$ have been computed directly. Using these values, the mean and the variance of $1/X$ have been computed to eight decimal places for $N = 2(1)20$; $M = 1(1)N$, $n = 1(1)M$. n need go only to M, since for each N, the mean and the variance of $1/X$ are symmetric in M and n. These computations will be extended up to some large values of N, M and n. Recursion relationships among $a_i(N, M, n)$, $i = 0, 1, 2$ have been noted. Formulae for $E[(m + 1)^{-1} \mid N, M, n, m > 0]$, $E[(n - m + 1)^{-1} \mid N, M, n, m > 0]$, $E[\{(m + 1)(m + 2)\}^{-1} \mid N, M, n, m > 0]$, $E[(n - m)^{-1} \mid N, M, n, m < n]$ have been derived. Various lower and upper bounds for $b_i(N, M, n)$, $i = 1, 2$ have been obtained.

28. Nonparametric Life Test Sampling Plans. SHANTI S. GUPTA, Stanford University and Bell Telephone Laboratories.

Assuming only that the lifetime (failure) distribution $F(t)$ has a monotone increasing failure rate, sampling plans are developed which accept a lot on the basis of a truncated life test only if lower bounds on a specified quantile or the mean life are established at a prescribed confidence level P^*. If ζ_q^0 and μ^0 denote respectively, the specified lower bounds on the quantile of order q and the mean of the distribution, then if the experiment time t is $\geqq \zeta_q^0 (> \mu^0)$ one can use lower bounds on $F(t)$ to derive the smallest sample size n. Under these plans, the decision to accept can take place only at the end of time t and only if the number of failures does not exceed a given acceptance number c. For given c, P^*, q and the ratio $\lambda = t/\zeta_q^0$, the required sample size is the smallest positive integer n which satisfies the inequality $I_{1-(1-q)^\lambda}(c + 1, n - c) > P^*$. A similar inequality gives n for the case when the interest is in establishing $\mu \geqq \mu^0$. Operating characteristic functions of these plans are studied. It is shown that the sampling plan for $\zeta_q \geqq \zeta_q^0$ is the same as the one for establishing the same quantile of an exponential distribution.

29. A Sequential Selection Procedure for the Best Population (Preliminary report). IRWIN GUTTMAN, McGill University.

Consider a set of k populations and suppose there is a member of the set which is defined as best in a certain sense (e.g., *Ann. Math. Stat.*, Vol. 31 (1960), p. 1216). A sequential procedure is described for which there is probability of at least β, say, of isolating the best population. The proofs depends on a multiple comparison type argument (Tukey, 1954, "Multiple Comparisons", Statistical Research Group, Memorandum Report, Princeton University). It is proved that the procedure terminates with probability one. Various illustrative examples are given (binomial, normal, exponential), and for particular cases, the use of different sample sizes is discussed.

30. A Generalized Queueing Model. VINCENT HODGSON, Florida State University. (Introduced by Ralph A. Bradley.)

Consider a single-server queueing system differing from the classical model in that an idle server is not continuously inspecting the system so that service will recommence as soon as a unit arrives. Instead inspections occur at discrete instants and the intervals between successive inspections when the server is idle are statistically distributed. We investigate a system in which arrival times are exponentially distributed and service and inspection times are generally distributed; by an obvious extension of Kendall's familiar notation this is the generalized $E_1/G/G/1$ model. Some connections with the classical theory are discussed.

31. The Moments of a Variate Related to the Non-Central *t*. D. HOGBEN, R. S. PINKHAM AND M. B. WILK, Rutgers—The State University. (By title)

Suppose that a random variable W, normally distributed with mean θ and variance 1 is independently distributed of a random variable X^2 which is distributed as chi-squared with n degrees of freedom. Then, the random variable Q, with non-centrality θ and n degrees of freedom, is defined by $Q = W/(W^2 + X^2)^{\frac{1}{2}}$. The probability density function for Q is obtained by transforming the joint density of W and X to polar coordinates and integrating. Then, closed form analytic expressions and recurrence relations for the raw moments of Q are obtained from the application of two derived lemmas. A table of numerical values of the first four central moments and cumulants for 38 values of Q from 0 to 10 and 36 values of n from 0 to 100 is given with a discussion of the computation and interpolation procedures. Also, there is a discussion of possible applications of the table of moments.

32. An Approximation to the Distribution of Q (a Variate Related to the Non-Central t). D. Hogben, R. S. Pinkham and M. B. Wilk, Rutgers—The State University.

If W is distributed as normal with mean θ and variance 1 and Z^2 is distributed independently as chi-squared with n degrees of freedom, then the random variable Q, with non-centrality θ and n degrees of freedom, is defined by $Q = W/(W^2 + Z^2)^{\frac{1}{2}}$. The probability integral of Q is implicitly defined by that of the non-central t, but the existing tables of the non-central t are not extensive. The purpose of the present paper is to propose as an approximation to the distribution of Q, the distribution of a linearly transformed beta variate. The approximation is obtained by equating the first two central moments of Q to the first two central moments of the linearly transformed beta and then solving for the unknown parameters of the beta-distribution. The accuracy of the approximation is discussed for several values of n and θ. The results provide a useful approximation to the probability integral of the non-central t.

33. Variance Components in Two-Way Classification Models with Interaction. Chandrakant H. Kapadia, University of Georgia and David L. Weeks, Oklahoma State University.

In minimum variance unbiased estimation it is desirable to have a minimal sufficient statistic for the family of distributions under consideration. In this paper, a minimal sufficient statistic is exhibited for the Balanced Incomplete Block design and the Group-Divisible, Partially Balanced Incomplete Block design with two associate classes. In addition, the distribution of the components of the minimal sufficient statistic is found and pairwise independence of the components investigated. A method of computing the minimal sufficient statistic is given in the case of the B.I.B. design and is based on quantities normally computed when an analysis of variance table is to be formed. The quantities calculated in an Analysis of Variance omit one of the components in the minimal sufficient statistic.

34. Theory of Queues with a Single Server. Tatsuo Kawata, Catholic University.

Considering a queue system with a single server and first come-first served discipline, suppose that $\{\tau_n\}$, $(0 < \tau_0)n = 0, 1, 2, \cdots$ are arrival instants of successive customers, $X_n = \tau_n - \tau_{n-1} (n \geq 1)$ and Y_n are interarrival times and service times required by customers respectively and $\{X_n\}$, $\{Y_n\}$ are independent as a whole, each being a sequence of random variables identically distributed. Set $Z_n = Y_n - X_n$, $S_n = \sum_1^n Z_k$. Continuing the previous work (*Bull. Inst. Internat. Stat.*, Vol. 38, 1961) we shall discuss the function $G(y) = \sum_1^\infty F_k(y)$, $F_k(y) = P(S_1 > 0, \cdots, S_{k-1} > 0, y > S_k)(k > 1)$, $F_1(y) = P(S_1 < y)$ which is well defined if $EZ_1 < 0$, and is the limit distribution of the waiting time of customers except a constant factor. We shall discuss the existence of moments of $G(y)$. For instance, if X_n has the characteristic function regular in some upper half-plane which includes the whole real line and $E|Y|^{m+1} < \infty$, then the absolute moment of $G(y)$ of mth order is finite. Also the limit distribution of interdeparture times is found.

35. On the Estimation of the Probability Density. M. R. LEADBETTER, Research Triangle Institute AND G. S. WATSON, University of Toronto.

Estimators of the form $\hat{f}_n(x) = (1/n) \sum_{i=1}^{n} \delta_n(x - x_i)$, of a probability density $f(x)$ are considered, where $x_1, \cdots, x_n$ is a sample of n observations from $f(x)$. The properties of such estimators are discussed on the basis of their mean integrated square errors, $E[\int (\hat{f}_n(x) - f(x))^2 \, dx]$ (M.I.S.E.), and also on the basis of various pointwise consistency criteria. The corresponding development for discrete distributions is sketched and examples are given in both continuous and discrete cases. The definitions and results are analogous to those of Parzen for the spectral density.

36. Averages of Correlated Observations. RAY H. LEE AND ARTHUR YASPAN, Autometric Company.

Let the random variables $x_1, \cdots, x_n$ each be an unbiased observation of the same quantity $\bar{x}$. An "average" of $x_1 \cdots x_n$ is defined to be an unbiased linear combination of $x_1 \cdots x_n$, and an "optimal average" to be an average with minimum variance. When the observations are independent and have the same variance, the optimal average is the arithmetic mean; when they are merely independent, the optimal average is that for which the weight applied to an observation is proportional to the reciprocal of its variance. In this paper, a solution is obtained for the case of a set of observations with arbitrary intercorrelation.

37. Some Sample Function Properties of a Process with Stationary Independent Increments. S. S. MITRA, University of Idaho.

This paper investigates the sample function behavior, for large values of t, of a general additive stochastic process $X(t)$ based on assumptions concerning the Lévy Measure of the process. The author calls attention to two indices θ and θ' defined respectively by $\theta = \sup \{\alpha \geqq 0: |y|^{-\alpha} \operatorname{Re} \psi(y) \to 0 \text{ as } y \to 0\}$ and $\theta' = \inf \{\alpha \geqq 0: |y|^{-\alpha} \operatorname{Re} \psi(y) \to \infty \text{ as } y \to 0\}$ where $\psi(y)$ is the exponent of the characteristic function of $X(1)$. The author proves: (i) $0 \leqq \theta' \leqq 2$; (ii) $t^{-1/a}X(t) \to 0$ a.s. as $t \to \infty$, provided $a < \min(1, \theta)$; (iii) if the process is symmetric about zero and $\theta > 1$, then $\liminf |X(n)| \to 0$ a.s. as $n \to \infty$; (iv) $\limsup t^{-1/a}|X(t)| \to \infty$ a.s. provided $a > \theta$; (v) if $\theta' < 1$, then $t^{-1/a}|X(t)| \to \infty$, provided $a > \theta'/(1 - \theta')$. In particular, if $X(t)$ is a subordinator, i.e., the corresponding Lévy measure is of the form $\int_0^1 u \, d\nu(u) < \infty$, $\nu(-\infty, 0] = 0$, then some other interesting properties are noted. Defining $\delta' = \inf\{\alpha \geqq 0: y^{-\alpha} \int_0^\infty (i - e^{-ry}) \, d\nu(r) \to \infty \text{ as } y \to 0_+\}$ the author proves that $1 > \delta' = \min(1, \theta)$ and that for every $a > \delta'$, $t^{-1/a}X(t) \to \infty$ a.s.

38. Estimation of the Spectrum II. V. K. MURPHY, Stanford University.

In this paper the author extends the results of his previous paper entitled "Estimation of the Spectrum" (*Ann. Math. Stat.*, Vol. 32 (1961), pp. 730–738) to the case of a Stationary Vector process using a different approach.

39. Partially Duplicated Fractional Factorial Designs. M. S. PATEL, Research Triangle Institute and Purdue University. (By title)

In this paper are given two-level fractional factorial designs in which some of the treatment combinations (runs) are duplicated. As is well known, the duplicated runs provide an unbiased estimate of error variance and more pricise estimates of the effects. Designs of this type were suggested by C. Daniel at the 1957 convention of the American Society for Quality Control and were discussed in detail by O. Dykstra, Jr. (*Technometrics*, Vol. 1, No. 1, pp. 63–75). In all the designs that are given, estimates of the main effects and the two-factor interactions are obtained with lesser number of runs than that required by the corresponding designs given by O. Dykstra. Finally the designs are arranged into blocks to give the corresponding block designs.

40. Asymptotic Bias and Variance of Ratio Estimates in Generalized Power Series Distributions and Certain Applications. G. P. PATIL, McGill University.

A discrete probability distribution which forms a generalization of some important discrete distributions like the Binomial, Poisson, Negative Binomial and Logarithmic Series and their truncated forms is introduced. It is called the "generalized power series distribution (gpsd)." In this paper we suggest what we call the "Ratio Method" for estimation of the parameter of the gpsd and investigate properties and study certain applications. The method is applicable not merely for estimating the parameter, but also for its integral powers. The performance of the method is investigated, in particular, in case of truncated Binomial and truncated Poisson distributions and correspondingly certain recommendations are offered.

41. A Sequential Procedure for Selecting the Population with the Largest Mean from k Normal Populations with a Common Known Variance (Preliminary report). EDWARD PAULSON, Queens College, New York.

Let $X_{ij}(i = 1, 2, \cdots, k; j = 1, 2, \cdots)$ be independent observations from normal populations Π_i with means m_i and common known variance σ^2, and let $m_{[1]} \leqq m_{[2]} \leqq \cdots \leqq m_{[k]}$ denote the ranked means. Let $S(i,j) = \sum_{\beta=1}^{j} X_{i\beta}$, let $D = \log_e[(k-1)/\alpha]$, let $A = 2\sigma^2 D/\Delta$, and let N_0 denote the largest integer contained in $2A/\Delta$. A sequential procedure with which inferior populations are eliminated as the experiment procedes is proposed so that the probability is $\geqq 1 - \alpha$ of selecting the population with the greatest mean whenever $m_{[k]} - m_{[k-1]} \geqq \Delta$, where Δ and α are preassigned. *Procedure:* At the first stage of the experiment, take the vector observation $(X_{11}, X_{21}, \cdots, X_{k1})$. Eliminate any population Π_i for which $X_{i1} \leqq \max\{X_{11}, X_{21}, \cdots, X_{k1}\} - A + \Delta/2$. After the $(r-1)$st stage $(r = 2, 3, \cdots, N_0)$ suppose that $t(r) = t$ populations $\Pi_{i_1}, \Pi_{i_2}, \cdots, \Pi_{i_t}$ have not been eliminated. Then at the rth stage take the vector abservation $(X_{i_1 r}, X_{i_2 r}, \cdots, X_{i_t r})$. Eliminate any category $\Pi_{i_\lambda}(\lambda = 1, 2, \cdots, t)$ for which $S(i_\lambda, r) \leqq \max\{S(i_1, r), \cdots, S(i_t, r)\} - A + r\Delta/2$. The experiment is terminated when all but one population has been eliminated, in which case the one remaining population is selected. If more than one population remains after the N_0th stage, then at the next stage take one measurement from each remaining population and then terminate the experiment by selecting the population with the largest cumulative sum.

42. Shorter Confidence Intervals for the Mean of a Normal Distribution with Known Variance (Preliminary report). JOHN W. PRATT, Harvard University.

In "Length of confidence intervals," *J. Amer. Stat. Assoc.*, Vol. 56 (1961), 549–67, the expected length of a confidence interval was shown to equal the integral of the probability of covering false values, and a method was given for minimizing average expected length. If the average is computed with respect to an extreme weighting function, the minimizing procedure may be very non-standard, even in simple problems. The present paper concerns the mean μ of a normal distribution with known variance when the weighting function is also normal. The minimizing procedures (which include the usual procedure as a limiting case) are computed compared with one another as to efficiency, and compared with Bayesian methods using the normal weighting function as a prior distribution.

43. General Limit Theorems for Spacings. Frank Proschan, Boeing Scientific Research Labs and Ronald Pyke, University of Washington.

Let $X_1 < X_2 < \cdots < X_n$ be the ordered observations of a sample of size n from an absolutely continuous distribution function F. Set $D_{ni} = X_i - X_{i-1}$ and $D^*_{ni} = (n - i + 1)D_{ni}$ for $i = 2, \cdots, n$. For a given function g, consider the sum $V_n = \sum_{i=2}^n g(D^*_{ni})$. Under weak assumptions on the functions g and F, it is shown that V_n has the same limiting distribution as does $\sum_{i=1}^n \{g(r_{i/n}Y_i) + (Y_i - 1)(n-1)^{-1}\sum_{j=i}^n G(j/n)(1 - j/n)\}$ where $r_\alpha = (1-\alpha)/F'(F^{-1}(\alpha))$, $G(\alpha) = (d/d\alpha)E[g(Y_1 r_\alpha)]$ and $Y_1, Y_2, \cdots, Y_n$, are independent random variables, each having the exponential distribution function; $1 - e^{-x}$. The classical Lindberg-Feller Central Limit theorem is then applied to this latter sum, to prove asymptotic normality of V_n. Generalizations of these results to the case of higher dimensional functions g are also described. Of particular importance, is the 2-dimensional case, $V_n = \sum_{i=2,j=i+1}^{n,n} g(D^*_{ni}, D^*_{nj})$ which has many applications to hypothesis testing in Reliability Theory, and to the comparative study of various test procedures by means of Asymptotic Relative Efficiency. The methods used in obtaining these limit theorems are used also to provide simple proofs of the corresponding theorems for D_{ni}, thereby generalizing results of Darling (*Ann. Math. Stat.*, Vol. 24 (1953), pp. 239–253) and L. Weiss (*Ann. Math. Stat.*, Vol. 29 (1958), pp. 310–316).

44. On the Order of an Entire Characteristic Function. B. Ramachandran, Catholic University.

It is known that the order of an entire characteristic function (c.f.) cannot be less than unity, unless the function itself is identically equal to one (P. Lévy). Also, a necessary and sufficient condition for a c.f. to be entire, of order one and of exponential type is that the corresponding distribution function (d.f.) be "finite" (G. Pólya). We here investigate necessary and sufficient conditions for a d.f. to have (i) an entire c.f. of given finite order greater than one, and (ii) an entire c.f. of order one and of maximal type. We also note in this connection that there cannot exist an entire c.f. of order one and of minimal type.

45. Multivariate Hierarchical Designs. S. N. Roy and J. N. Srivastava, University of North Carolina.

Consider a multivariate problem involving, say, p variables $x_1, x_2, \cdots, x_p$ and n experimental units. In many situations, because of cost and other considerations, it may not be worthwhile to study each characteristic with the same intensity, and hence each experimental unit may not be studied on all characteristics. A multivariate hierarchical design is such that the n units are divided into p mutually exclusive and exhaustive sets $D_1, D_2, \cdots, D_p$, such that on the set D_j, only the characters $x_1, x_2, \cdots, x_j (j = 1, 2, \cdots, p)$ are observed. Such designs have been introduced in this paper, and the corresponding theory of testing of linear hypothesis on the populations means, and of testing the usual kind of hypotheses on the population dispersion matrix have been developed.

46. Asymptotic Relative Efficiency of Massey's Test for c Samples $c \geqq 2$. Y. S. Sathe, University of North Carolina.

Let $X_{ij}, i = 1, 2, \cdots, c; j = 1, 2, \cdots, n_i$ be $N = \sum_{i=1}^c n_i$ independent random variable, and let $F_i(x)$ be the c.d.f. of X_{ij}. Further, let $Z_1, Z_2, \cdots, Z_h$ $(h \geqq 2)$ denote the ρ_1th,

$(\rho_1 + \rho_2)$th $\cdots$ $(\rho_1 + \rho_2 + \cdots + \rho_h)$th quantiles of the combined sample $(\rho_t > 0)$ and for convenience, let $Z_0 = -\infty$, $Z_{h+1} = \infty$. Denote by $M_{i,t}$ the number of observations from the ith sample which lie in the interval (Z_{t-1}, Z_t), $t = 1, 2, \cdots, h + 1$. Then under $H_0 : F_1(x) = F_2(x) = \cdots = F_c(x)$, Massey has shown that the asymptotic distribution of $X^2_{\text{Ma}} = \sum_{i=1}^{c} \sum_{t=1}^{h+1} (M_{i,t} - \rho_t n_i)^2/\rho_t n_i$, $(\rho_{h+1} = 1 - \sum_{t=1}^{h} \rho_t)$ is a chi-square with $h(c - 1)$ d.f. In this paper, we show that when the hypothesis $H_N : F_i(x) = F(x + N^{-\frac{1}{2}}\theta_i)$ holds, where $F(x)$ is a continuous c.d.f. with a continuous derivative $f(x)$, the limit distribution of X^2_{Ma} as $n_i \to \infty$ such that $n_i/N \to p_i > 0$ is noncentral chi-square with $h(c - 1)$ d.f. and non-centrality parameter, $\lambda^2_{\text{Ma}} = (\sum_{t=1}^{h+1} [f(c_t) - f(c_{t-1})]^2/\rho_t)(\sum_{i=1}^{c} p_i(\theta_i - \bar{\theta})^2)$ where $\bar{\theta} = \sum_{i=1}^{c} p_i\theta_i$, $c_1, c_2, \cdots, c_h$ are the respective quantiles of $F(x)$ and $f(c_0) = f(c_{h+1}) = 0$. A comparison is made of this test with other non-parametric tests and expressions for a.r.e. are obtained for specified choices of $F(x)$.

47. On the Set of Distributions Induced by a Ratio Statistic. Kenzo Seo, Colorado State University.

Let $f_1, \cdots, f_k$ be a set of probability density functions. We define a vector-valued statistic $r(x)$ by $r(x) = (r_1(x), \cdots, r_k(x))$ where $r_i(x) = f_i(x)/[f_1(x) + \cdots + f_k(x)]$. The distributions $P_1, \cdots, P_k$ induced by r may be characterized as follows: If μ is any measure to which all P_i are absolutely continuous, then the derivatives $p_i = dP_i/d\mu$ satisfy $p_i(t)/[p_1(t) + \cdots + p_k(t)]$, $t = (t_1, \cdots, t_k)$ a.e. $(P_1 + \cdots + P_k)$. A k-decision problem may be characterized in terms of the distributions $P_1, \cdots, P_k$. Suppose that the set $f_1, \cdots, f_k$ and the set $g_1, \cdots, g_k$ induce identical distributions $P_1, \cdots, P_k$. Then the problem of selecting the true density of a random variable from among the $\{f_i\}$ and the parallel problem of selecting from among the $\{g_i\}$ may be considered equivalent, since there is a one to one correspondence between the Bayes rules for the former problem and the Bayes rules for the latter.

48. On the Time of First Birth. S. N. Singh, Pennsylvania State University. (By title)

A probability distribution of the time of the first *complete* conception (a conception resulting in completed pregnancy) to a couple after marriage has been derived under the following assumptions. (a) The number of cohabitations during any time-interval $(0, t)$ follows a Poisson distribution $P(\lambda_1 t)$; (b) cohabitations are independent and P_1, the probability that a cohabitation results in a conception, is constant; (c) conceptions are independent and P_2, the probability that a conception is complete, is constant and (d) $\lambda = \lambda_1 P_1 P_2$ follows a type III distribution. Following Neyman, "Contribution to the theory of x^2 test," *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability* (1949), pp 239–273, a procedure to find the B.A.N. (best asymptotically normal) estimates of the parameters is outlined. The distribution has been applied to data, suitably modified, on the time of first birth from Singh, *Demographic Survey of Banaras Tehsil*, India (1956).

49. Analysis of Fractionally Replicated Asymmetrical or Symmetrical Factorial Designs, I. J. N. Srivastava, University of North Carolina. (By title)

In this paper, a general theory has been developed for analyzing a fraction obtained from any general asymmetrical or symmetrical factorial experiment. The development of the theory has been pushed up to the stage of evaluating the matrix which we have to invert for solving the normal equations. The special techniques for inverting such a matrix are discussed in the second paper.

50. Statistical Methods in Reliability Demonstration. ALEXANDER STERNBERG, JOHN S. YOUTCHEFF, AND SIDNEY DEMSKEY, General Electric Company.

This paper presents the statistical methods which have been developed and utilized effectively to provide a thorough reliability demonstration program for satellites and space vehicles. The analytical, temporal, and monetary tradeoff parameters are fully explored and integrated to develop the necessary reliability test and evaluation program for adequate reliability demonstration. Reliability values and associated confidence estimates, project development schedules, and equipment and testing costs are the specific parameters considered in the establishment of an optimum testing program. Such a testing program provides for demonstrating the reliability objectives for any desired confidence level relative to a specified mission time.

51. On Sampling with Replacement: An Axiomatic Approach (Preliminary report). RICHARD C. TAEUBER, C-E-I-R, Inc. AND JOHN C. KOOP, North Carolina State College.

When sampling finite populations, probabilities enter into sampling problems only in the process of selecting the units to be observed, and not in connection with the unit characteristics to be measured. Thus, in examining the criteria which may be applied to determine bestness in an estimator, the criteria developed for infinite populations, for the most part, are not applicable to samples from finite populations. Only consistency (by and large) and minimum mean square error (and their more restrictive counter-parts: unbiasedness and minimum variance) survive. Further, the classical approach of determining criteria for which one can develop classes of estimators is not germane to sample surveys. In sample survey theory the converse is true, i.e., classes of estimators are developed in some systematic, physically-realizable manner, and then criteria are applied to determine bestness within each class. As one possible method of obtaining classes of estimators, this study has applied the axiomatic approach developed earlier by Koop to formulate seven classes of estimators when sampling with replacement and with arbitrary selection probabilities. Estimators are derived whose coefficients satisfy requirements that the estimator shall be unbiased, and the weights be independent of the variate values under observation.

52. Bivariate Limiting Distributions of Maxima. HENRY TEICHER, Purdue University.

Let $X_n = (X_{n1}, X_{n2})$, $n = 1, 2, \cdots$ denote a sequence of independent identically distributed random vectors with common c.d.f. $F(x, y)$ and define $U_{nj} = a_{nj}^{-1}(\max_{1 \leq i \leq n} X_{ij} - b_{nj})$ where b_{nj} and $a_{nj} > 0$ are suitable norming constants, $j = 1, 2$, $n = 1, 2, \cdots$. Employing a result of M. Sibuya (*Ann. Inst. Stat. Math.*, Vol. 11, No. 3, 1960), the class of proper limiting distributions for (U_{n1}, U_{n2}) is obtained.

53. Percentile Points of Order Statistics from a Uniform Distribution. JOHN S. WHITE, General Motors Research Labs. (By title)

Let $x_{1n} \leq x_{2n} \leq \cdots x_{nn}$ be an ordered random sample from a uniform (0, 1) distribution. The distribution of the ith order statistic is then Prob $[x_{in} \leq y] = F_{in}(y) = I_y(i, n - i + 1)$, where $I_x(p, q)$ is Karl Pearson's incomplete Beta function ratio. Tables of the Pth percentile points of $x(i, n)$ are given for $n = 1$, (1), 50; $i \leq n$; and $P = .1\%$, $.5\%$, 1.0%, 2.5%, 5%, 10%, 25%, 50%. These tables may be used in plotting observed data on probability graph paper using for the probability p_i corresponding to y_{in}, the ith largest sample value,

the median of x_{in}. Confidence intervals about the plotted empirical distribution function may be obtained using other percentile points. These tables may also be used to find points on the operating characteristic curves of simple sampling plans. For example the AQL (95% acceptance point) for sample size $= n$, rejection number $= i$, is the 5% point of $x(i, n)$.

(Abstracts of papers to be presented at the Eastern Regional Meeting of the Institute, Chapel Hill, April 12–14, 1962. Additional abstracts will appear in the June, 1962 issue.)

1. Linear Estimation and the Analysis of Gamma Ray Spectra. Bernard S. Pasternack, New York University Medical Center.

Gamma ray spectroscopy is a relatively recent technique which utilizes radioactivity counters for the assaying of a mixture of gamma-emitting radionuclides with minimum sample alteration. This paper is concerned with a method of obtaining best estimates, in a statistical sense, of the amount of each radionuclide present in the mixture. It should be most useful in situations involving low levels of activity—i.e., near the limits of detection. Such problems arise and are of importance, for example, in the determination of radioactivity in the environment and in neutron activation analysis. Since extensive computations are required when more than a few isotopes are involved, the full potential of the estimation procedure suggested requires access to a modern digital computer programmed for handling problems of linear estimation and, in particular, weighted least squares. The solution proposed yields best linear unbiased estimates of these unknown amounts, which also have the property of being best asymptotically normal. In other words, as the counting time increases the estimates approach the true amount of activity of the isotopes with a limiting distribution that is normal, and such that for any given time interval of counting the variances of these estimates are minimal.

2. Testing of Hypotheses Connected With a Certain Class of Patterns in the Population Dispersion Matrix. S. N. Roy and J. N. Srivastava, University of North Carolina.

In many multivariate experimental situations encountered in psychometrics, econometrics, etc. background considerations suggest that the population dispersion matrix has a certain pattern, and this leads to the problem of testing the hypothesis that such a matrix does have the given pattern. On the other hand, we can sometimes assume that the population dispersion matrix has a certain pattern, and, on this model we have the problem of testing for linear hypotheses. In this paper, it has been shown that for a certain wide class of patterns, these problems are related to certain Linear Associative Algebras. Using the properties of such algebras, likelihood ratio tests have been given for each of the two kinds of hypotheses mentioned above. In this connection the use of the stepdown procedure as an alternative approach has also been indicated.

3. The Distribution of Incubation Periods in a Birth-Death Process. G. Trevor Williams, Johns Hopkins University. (Introduced by Allyn W. Kimball.)

We describe a stochastic model which mimics many of the characteristics of infectious diseases. Assume that an inoculum of n organisms follows a homogeneous birth-death process with parameters λ and μ. Place an absorbing barrier at N, the number of organisms required to elicit a defensive response from the host. In those cases where the process actually reaches this upper barrier, the first-passage time will be called the incubation period, t. We find that $\bar{t} \backsim \log N/(\lambda - \mu)$, $(N \rightarrow \infty)$ and further, that the frequency function of t

falls into a well-behaved asymptotic form about this mean. For most diseases one expects even the inoculum to be rather large if there is to be an appreciable probability of infection. Hence we let $n \to \infty$ while letting $\mu \to \lambda$ in such a way as to hold the probability of infection $p = 1 - (\mu/\lambda)^n$ fixed, since p is a physically significant quantity. The distribution of the incubation period then comes out explicitly in terms of a Bessel function, with p as a parameter that determines the shape. For $p = 0$, it is a Fisher-Tippett distribution and it is relatively insensitive to changes in p. The model is shown to have many properties that agree with commonly accepted empirical descriptions of infectious disease phenomena.

(Abstracts of papers to be presented at the Western Regional Meeting of the Institute, Albuquerque, New Mexico, April 19–20, 1962. Additional abstracts will appear in the June, 1962 issue.)

1. Estimating the Parameters in the Model $y_{ijk} = a_i - b_j + e_{ijk}$. AARON S. GOLDMAN, Los Alamos Scientific Laboratory.

Given the model $y_{ijk} = a_i - b_j + e_{ijk}$ where $i = 1, 2, \cdots, n, j = 1, 2, \cdots, m, k = 1, 2, \cdots, t$, y_{ijk} is an observed random variable, e_{ijk} is a random variable with mean 0 and variance σ^2, and a_i and b_j are parameters, then the best estimates of a_i and b_j and their errors are found. Also the computing techniques are presented for $m + n \leqq 1000$ so that accuracy in the results are obtained for as many digits as in the given data. The case when data are missing is also given.

2. Games Associated with a Renewal Process. M. M. SIDDIQUI, Boulder Labs.

Consider a sequence of occurrences of a recurrent event E for which the intervals X_1, X_2, $\cdots$ are identically distributed non-negative random variables (a renewal process) with common distribution function $F(x)$. Robbins (these *Annals*, Vol. 32 (1961), pp. 187–194) considered games when X_1 is an integer-valued random variable. In this paper his results are extended to games when X_1 is not necessarily integer-valued. A continuous analogue of the Petersburg game is also presented. Games associated with a general renewal process are also considered where $X_1, X_2, \cdots$ are preceded by another independent random variable X_0 with distribution function $B(x)$.

(Abstracts of papers not presented at any meeting of the Institute.)

1. A Mathematical Theory of Pattern Recognition. ARTHUR ALBERT, Massachusetts Institute of Technology. (By title)

Let X_0 and X_1 be (unknown) disjoint subsets of a Hilbert space H, such that the convex hulls of X_0 and X_1 are a positive distance apart. Suppose that samples are drawn independently and at random from $X_0 \cup X_1$. After the nth sample, Z_n, it is required to guess whether Z_n came from X_0 or X_1. After each guess, we are told whether we were right or wrong. In this paper, a decision procedure is exhibited, having the property that the probability of making an error on the nth trial converges to zero with increasing n. Furthermore, the guessing rule used on the $n + 1$st trial depends on the past data only through the rule used on the nth trial, the value of Z_n, and whether or not the guess about Z_n was correct. The application to pattern recognition problems of a dichotomous sort is immediate when we identify X_0 and X_1 with two classes of patterns which are observed in temporal succession. The rules for membership in X_0 and X_1 are not known, but we (or a machine) are/is told to which class each pattern belongs, *after* making a guess about that pattern. As the "training period" increases, errors are made with ever decreasing frequency.

2. A Simple Randomization Procedure. MARTIN SANDELIUS, Sandelius Statistiska Byra, Sweden. (By title)

The paper describes a randomization procedure consisting in distributing a deck of cards into 10 decks using random digits and repeating this step with each deck consisting of 3 or more cards. One random digit is used for randomizing a deck of 2 cards. This procedure is called the multistage randomization procedure or MRP. A recursive formula is given for the expected number of random digits required by MRP for the randomization of n different symbols. As measure of the efficiency of a randomization procedure applied to n different symbols the quantity $(\log_{10} n!)$/(Expected number of random digits required) is used. It is shown that it is possible to construct a randomization procedure which on unlimited repetition gives an efficiency which is arbitrarily close to 1. Using a family of upper bounds on the expected number of random digits required by MRP it is shown that MRP is asymptotically efficient. The efficiency of MRP is compared with the efficiencies of two onestage randomization procedures.

3. A New Design for Experiments with Mixtures (Preliminary report). HENRY SCHEFFÉ, University of California at Berkeley. (By title)

A "simplex-centroid" design is proposed for mixture experiments (Scheffé, *JRSS*, Ser. B, 20 (1958), 344–360). If x_i denotes the proportion of the ith component $(i = 1, \cdots, q)$, the factor space is the $(q - 1)$-dimensional simplex $\sum_1^q x_i = 1$, $x_i \geqq 0$. The new design consists of the following $2^q - 1$ points of the simplex: its centroid, and the centroids of all the lower-dimensional simplexes it contains (including the q vertices). A regression equation expected to be useful with this design is $\eta = \sum_i \beta_i x_i + \sum_{i<j} \beta_{ij} x_i x_j + \sum_{i<j<k} \beta_{ijk} x_i x_j x_k + \cdots + \beta_{12\cdots q} x_1 x_2 \cdots x_q$. This is the same as that usually used with factorial experiments in q factors each at two levels (each $x_i = +1$ and -1), except for the lack of the term β_0, which can here be absorbed into the linear terms. A natural correspondence, depending on which components are present and absent, may be set up between the points of the new mixture design and the $2^q - 1$ points of the complete factorial experiment design left after deletion of the point $(-1, -1, \cdots, -1)$. The analogy immediately suggests designs and regression equations for the addition of process (i.e., non-mixture) variables to the mixture variables, each process variable being at two levels, and how one might try to fractionate these designs.