out to me in personal correspondence, this is actually not the case. However, the theorems and their proofs remain completely valid in their present form if the observations are drawn from a stochastic process satisfying condition (5) of the paper. This chain condition states that the process be such that

$$\text{Prob } (X_n \leq x \mid X_1 \leq x, X_2 \leq x, \cdots, X_{n-1} \leq x) = \text{Prob } (X_n \leq x \mid X_{n-1} \leq x)$$

is satisfied for all x and for all positive integers n.

ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Chicago meeting of the Institute, December 27–29, 1950)

1. Cost Functions for Sample Surveys. (Preliminary Report). Garnet E. McCreary, University of Manitoba and Iowa State College.

Assume: (1) one travels in a rectangular (grid) fashion rather than straight line (airline) path, (2) n random points have a uniform distribution over the region or stratum. Moderate changes in shape of regions have a minor effect on expected distance. Mean airline distance can be prediced from mean grid distance fairly accurately. The following formulas are derived: (1) expected minimum grid distance for $n = 3$ in a square, (2) an upper bound to expected minimum grid distance for all n, (3) expected grid distance for a stratified and unstratified sample, if the path among the points does not reverse a certain direction, (4) expected distance of a random point from (a) the center of the arc of the circle, semicircle or quadrant, (b) any fixed point, inside or outside the rectangular region, (5) mean square distance between any pair of points adjacent in a clockwise direction (6.7 to 9.5 per cent biased upwards over corresponding mean airline distance). Certain conclusions are drawn regarding the most efficient design with respect to total distance. Detailed mileage records of three Iowa farm surveys were compared with theoretical estimates. If the cost is balanced against the losses resulting from errors in estimate, for a particular design, the problem of determining sample size is broached by using Wald's minimax principle.

2. On a Preliminary Test for Pooling Mean Squares in the Analysis of Variance. A. E. Paull, Abitibi Power and Paper Company, Limited, Toronto, Canada.

The consequences of performing a preliminary F-test in the analysis of variance is described. The use of the 5% or 25% significance level for the preliminary test results in disturbances that are frequently large enough to lead to incorrect inferences in the final test. A more stable procedure is recommended for performing the preliminary test, in which the two mean squares are pooled only if their ratio is less than twice the 50% point.

3. Estimation for Sub-Sampling Designs Employing the County as a Primary Sampling Unit. Emil H. Jebe, Iowa State College and North Carolina State College.

This paper summarizes a study of the application of various two-stage designs including the estimation procedures for providing state estimates of agricultural items in North Carolina. Among the principal objectives of the investigation were (1) the comparison of the efficiency of selection of the primary units with equal and with unequal probabilities, and (2) assessment of the relative contributions of the between primary sampling unit and within primary sampling unit error components to the total sampling error. Examination of several linear and ratio estimates indicates a number of advantages for a particular ratio estimate.

4. The Probability Distribution of the Number of Isolates in a Social Group.

LEO KATZ, Michigan State College.

Each of the N members of a well-defined social group is asked to name d others with whom he would prefer to be associated in some specified activity. Under the null hypothesis, his choices are randomly distributed. An isolate is an individual who is not chosen by any of the other members of the group. The probability of exactly i isolates in the group is then given by

$$P_i = \Sigma_{j=1}^{N-1-d} (-1)^{i+j} C(j, i) C(N,i) [C(N - i, d)]^i [C(N - 1 - i, d)]^{N-i} [C(N - 1, d)]^{-N},$$

where $C(N, n) = {}_NC_n$, the binomial coefficient. This expression for P_i is somewhat unwieldy. It is further shown that this probability function is asymptotically a binomial p.f., $P_i' = C(n, i)p^i(1 - p)^{n-i}$, where

$$p = N[(N - 1 - d)/(N - 1)]^{N-1} - (N - 1)[(N - 1 - d)/(N - 1)][(N - 2 - d)/(N - 2)]^{N-2}$$

and $np = N[(N - 1 - d)/(N - 1)]^{N-1}$. The approximation is very good even for moderately small values of N.

5. Estimating Population Size Using Sequential Sampling Tagging Methods.

LEO A. GOODMAN, University of Chicago.

Let $[n_i]$ be a sequence of positive integers and let $S(L, n_i)$ denote the procedure whereby (1) n_1 elements are drawn at random from a population P, then tagged to distinguish them from the remaining elements, and replaced in P, (2) n_2 elements are drawn from P, the number of tagged elements appearing is observed, the n_2 elements are then tagged and replaced in P, (3) $\cdots$, this process is halted when at least $L > 0$ tagged elements have appeared. Given $S(L, n_i)$, there exists a minimum variance unbiased estimator (m.v.u.e.) of the number N of elements in P which may be determined as the quotient of two determinants and simplified, by combinatorial methods, in special cases. If $[n_i]$ is bounded, as N approaches infinity, the limiting distribution of t^2/N, where t is the total number of elements drawn before the procedure ceases, is χ^2 with $2L$ degrees of freedom. Hence the asymptotic m.v.u.e. of N, confidence intervals and tests of hypotheses for N may be obtained as well as the approximate fiducial distribution of N. Similar results may be obtained for the more general cases where (a) information concerning size of some subclasses in P is used and (b) where taggings may or may not be differentiated. The $S(L, n_i)$ compares favorably with other procedures considered.

6. Application of the Distribution of a Linear Form in Chi-square Variates.

ARTHUR GRAD AND HERBERT SOLOMON, Office of Naval Research, Washington, D. C.

The probability of hitting a target depends both on the accuracy with which the position of the target is known and the dispersion of the weapon about the point of aim. Under the assumption that each of these errors has a bivariate Gaussian distribution with known covariance matrix, $\| \sigma(p) \|$ for position prediction error and $\| \sigma(a) \|$ for aiming error, about the point of aim (predicted position), the probability, P, of hitting the target with a weapon having a radius of effectiveness R is shown to be $P = Pr\{k_1^2x_1^2 + k_2^2x_2^2 \leqq R^2/C^2\}$, where $k_1^2 = [\sigma_{11}(p) + \sigma_{11}(a)]/C^2$, $k_2^2 = [\sigma_{22}(p) + \sigma_{22}(a)]/C^2$, $C^2 = \sigma_{11}(p) + \sigma_{22}(p) + \sigma_{11}(a) + \sigma_{22}(a)$, and x_i^2 is a chi-square variate with 1 degree of freedom. When $\sigma_{12}(p) = \sigma_{12}(a) = 0$, then the chi-square variates are independent. If not, a linear transformation exists such that $z = k_1^2x_1^2 + k_2^2x^2 = l_1^2y_1^2 + l_2^2y_2^2$, where $l_1^2 + l_2^2 = k_1^2 + k_2^2$ and y_1^2 and y_2^2 are independently distributed chi-square variates each having one degree of freedom. It is then demonstrated that $P = 2k_1k_2 \int_0^t e^{-z}I_0[z(1 - 4k_1^2k_2^2)^{\frac{1}{2}}]\,dz$, where $t = R^2/4C^2k_1^2k_2^2$, when the chi-square variates are independent; in case of dependence, k_i should be replaced by l_i. A table was constructed which covers the entire range of the parameters.

7. A Large Sample t-statistic which Is Insensitive to Nonrandomness. John E. Walsh. The Rand Corporation.

Most of the well known significance tests and confidence intervals for the population mean are based on the assumption of a random sample. This paper considers how the significance levels and confidence coefficients of the commonly used class of tests and intervals based on the standard Student t-statistic are changed when the random sample requirement is violated and the number of observations is large. It is found that even a slight deviation from the random sample situation can result in a substantial significance level and confidence coefficient change. Thus this class of tests and confidence intervals would seem to be of questionable practical value for large sets of observations. Large sample tests and confidence intervals for the mean which are not sensitive to the random sample requirement are obtained for a situation of practical interest by development of a special type of t-statistic. These results are as efficient (asymptotically) as those based on the standard t-statistic for the case of a random sample.

8. Conditional Expectation and Convex Functions. E. W. Barankin, University of California, Berkeley.

The inequality $E\psi(E(f \mid \cdot)) \leqq E\psi(f)$, (where the conditional expectation is taken with respect to a function t) with f a real- (or complex-) valued function on the fundamental space, was shown by Blackwell to hold in the case $\psi(z) = \mid z \mid^2$, and by the present author to hold in the case $\psi(z) = \mid z \mid^s$, $s \geqq 1$ (*Annals of Math. Stat.*, Vol. 18 (1947), pp. 105–110, and Vol. 21 (1950), pp. 280–284, respectively). More recently Hodges and Lehmann (*Annals of Math. Stat.*, Vol. 21 (1950), pp. 182–197) proved the inequality in the case of f a function to $\mathcal{E}^k$ (Euclidean k-space) and ψ a finite, convex, real-valued function on $\mathcal{E}^k$. Now, both Blackwell and this author exhibited the above inequality, in their cases, as (obvious) consequences of the more fundamental relation: $\psi(E(f \mid \tau)) \leqq E(\psi(f) \mid \tau)$ *for almost all points* τ *in the range of* t. The work of Hodges and Lehmann, however, leaves open the question whether or not the latter inequality holds in the more general case. In the present note this almost-everywhere inequality is established for f to $\mathcal{E}^k$ and ψ convex. The first inequality then obtains by integration.

9. Transformation Parameters. Melvin P. Peisakoff, The Rand Corporation.

Location, scale, and location-scale parameters are examples of *transformation parameters*. Transformation parameters are defined by: (1) the parameter space is a group, (2) the sample space can be factored into the same group and an arbitrary space, (3) the random variable associated with each parameter point, θ, can be generated by drawing from the population associated with the unit of the parameter space and left multiplying the group component of the sample by θ. *Decision function theory* is investigated when the decision space and the cost function are of a special intuitively appealing form. The formulation is broad enough to include sequential analysis. Minimax decision functions are found. Also investigated is *testing and confidence region theory*, using extensively the results on decision functions. Both simple and composite hypotheses are treated. Finally, (Fisher) *information theory* is examined. It is shown that modifications are necessary if information theory is to be useful in estimation problems. One modification is suggested. This modification enlarges the class of standard estimators to include each estimator which is minimax with respect to a certain risk function determined by the estimator itself. The approach is generalized to include inequalities for the mean square error other than the information inequality.

10. A Generalization of the Neyman-Pearson Fundamental Lemma. HENRY SCHEFFÉ, Columbia University.

Given $m + n$ real integrable functions $f_1(x), \cdots, f_m(x), h_1(x), \cdots, h_n(x)$ of a point x in a Euclidean space R, a real function $\varphi(y_1, \cdots, y_n)$ of n real variables, and m constants $c_1, \cdots, c_m$, the problem is to consider the existence of, and to find necessary conditions and sufficient conditions on, a set S maximizing $\varphi\left(\int_S h_1\,dx, \cdots, \int_S h_n\,dx\right)$ subject to the m side conditions $\int_S f_i\,dx = c_i$. In some applications the values of the vector

$$\left(\int_S h_1\,dx, \cdots, \int_S h_n\,dx\right)$$

may also be restricted to a given set. A statistical example in which $\varphi(y_1, \cdots, y_n) = \prod_{i=1}^{n} y_i$ arose in an unpublished paper of Isaacson. The methods of the present paper are suggested by those of an unpublished paper of Dantzig and Wald. Under certain regularity conditions the inequalities appearing in the Neyman-Pearson lemma are replaced by $\sum_{i=1}^{n} a_i^S h_i(x) - \sum_{j=1}^{m} k_j f_j(x) \geq 0$ (a.e. in S), ≤ 0 (a.e. in $R - S$). Here a_i^S and k_j are constants with $a_i^S = \partial\varphi/\partial y_i$ evaluated at $(y_1, \cdots, y_n) = \left(\int_S h_1\,dx, \cdots, \int_S h_n\,dx\right)$.

11. Nonparametric Estimation V, Sequentially Determined Statistically Equivalent Blocks. D. A. S. FRASER, University of Toronto.

In 1943 Wald gave a method for constructing tolerance regions for continuous multivariate distributions. Tukey generalized Wald's procedure and then interpreted the results for discontinuous distributions. In this paper a further generalization of the method is given by which statistically equivalent blocks can be determined sequentially; that is, the particular function used to cut off a block may depend on the shape or structure of previously selected blocks. The results are also extended to the case of discontinuous distributions. Possible advantages for the practitioner are discussed.

12. A Bayes Approach to a Quality Control Model. M. A. GIRSHICK AND HERMAN RUBIN, Stanford University.

A machine producing items of quality characteristic x can be in one of four states. In state $i = 1, 2$ the machine is in production and is characterized by a density $f_i(x)$. In state $j = 3, 4$ the machine is in repair having come from state $j = 2$. When the machine is in state 1 there is a probability g that in the next time unit it enters state 2, remaining in state 2 until brought to repair by some rule R based on observations. The income from items of quality x is $V(x)$; repair cost per unit time in state $j = 3, 4$ is c_j. A rule R^* is Bayes if it maximizes lim I_N as $N \to \infty$ where I_N is the expected income per unit time in N time units. It is proved that for 100% inspection, R^* states that sampling is to continue as long as $Z_n < a$ and sampling is to terminate and the machine placed in repair when

$Z_n \geq a$, where $Z_n = y_n(1 + Z_{n-1})$, $Z_0 = 0$ and $y_n = f_2(x_n)/(1 - g)f_1(x_n)$. R^* is also obtained in case inspection costs are taken into account. It is shown that the above Markoff process approaches a stable distribution and the required integral equations are derived.

13. On the Translation Parameter Problem for Discrete Variables. David Blackwell, Stanford University.

Let $x = (x_1, \cdots, x_N)$ be a vector chance variable, let $y = x + h\epsilon$, where $\epsilon = (1, \cdots, 1)$ and h is an unknown constant, and let $t = t(y)$ be any function of y, considered as an estimate for h when y is observed. Let $f(d)$ be any function of a real variable d, considered as the loss to the statistician when the error of estimate is d, so that the risk from an estimate t is $R_t(h) = Ef[h - t(x + \epsilon h)]$. Extending the work of Pitman, Girshick and Savage have exhibited an estimate t^* for which $R_{t^*}(h) = R$ independent of h, and have shown that t^* is minimax. It is shown here that if x assumes only a finite number of values $v_i = (n_{i1}, \cdots, n_{iN})$ and each n_{ij} is an integer, and if $f(d)$ is strictly convex and assumes its minimum value, then t^* is admissible and is in fact the unique minimax estimate. Two examples in which t^* is not admissible are given. A closely related fact is that if S is a closed bounded strictly convex subset of n-space intersecting the line $x_1 = \cdots = x_n$ at the single point $(w, \cdots, w)$, then the only sequence $\{z_m\}$, $-\infty < m < \infty$, for which $P_m = (z_{m+1}, \cdots, z_{m+n}) \in S$ for all m is $z_m = w$ for all m.

14. On Ratios of Certain Algebraic Forms. Robert V. Hogg, State University of Iowa.

Let x and y be random variables having a continuous cumulative distribution function, and let $M(u, t) = E[\exp(ux + ty)]$ exist in the neighborhood of the origin of the u, t plane. Subject to certain conditions a necessary and sufficient condition for the stochastic independence of y and x/y is $(\partial^k/\partial u^k)M(0, t) \equiv K_k(\partial^k/\partial t^k)M(0, t)(k = 0, 1, 2, \cdots)$, where K_k is evaluated by setting $t = 0$. This result is used in the study of certain ratios of quadratic and linear forms. In dealing with the quadratic forms, the sample arises from a normal population with mean zero. A necessary and sufficient condition is determined for the stochastic independence of Q_2 and Q_1/Q_2, where essentially $Q_1 = a_1x_1^2 + \cdots + a_nx_n^2$ and $Q_2 = b_1x_1^2 + \cdots + b_nx_n^2$. In the linear case however, the distribution is unspecified. Then it is found that the requirement of the stochastic independence of L_2 and L_1/L_2 implies that the sample arose from a gamma type distribution. Here $L_1 = a_1x_1 + \cdots + a_nx_n$ and $L_2 = x_1 + \cdots + x_n$.

15. The Economics of Sampling. Norman Rudy, Sacramento State College.

An optimum single sampling plan for acceptance inspection of attributes is developed by the method of minimizing the maximum risk. The first application is to warehouse or surveillance inspection, in which the value of a good item, g, and the cost of a bad item, b, define a breakeven quality, p_0. It is shown that under these conditions, and with sampling cost a linear function of sample size, s, tn, the optimum sample size is approximately equal to $[(.085 \text{ lot size})/t]^{2/3}(bg)^{1/3}$, the optimum acceptance number is approximately equal to np_0, and the $\min_{n,a} \max_p$ of the risk is approximately equal to $s + .58(tbg)^{1/3}(\text{lot size})^{2/3}$. The more general case, where the breakeven quality p_0 is determined by trade practice or contract, is also worked out, but cannot be presented in completely analytic form. A simple table involving the quotient of the normal integral and the normal density is required. Given this and the cost parameters of the situation, then the sample size and the acceptance number which minimize the maximum risk are determined from relatively simple expressions.

16. Exact Tests of Serial Correlation Using Noncircular Statistics. G. S. WATSON, University of Cambridge, AND J. DURBIN, London School of Economics.

The paper shows how noncircular statistics for testing hypotheses of serial independence may be constructed for which exact distributions can be obtained using results given by R. L. Anderson ("Distribution of the serial correlation coefficient," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 1–13). The statistics are derived by throwing away a small amount of relevant information. As an example the statistic

$$c_1 = (x_1x_2 + \cdots + x_{m-1}x_m + x_{m+1}x_{m+2} + \cdots + x_{2m-1}x_{2m}) \Big/ \sum_1^{2m} x_i^2$$

may be used for testing independence in a series of $2m$ observations whose mean is known to be zero. The quadratic form in the numerator of c_1 is based on a matrix whose roots are pair-wise equal, so that the distribution of c_1 when the x's are normal with the same variance is known from the results of R. L. Anderson. Tests of the errors in certain regression models may be made by fitting separate regressions to the two halves of the series and substituting the residuals in expressions similar to c_1. Exact tests can be obtained in this way for polynomial regressions, one-way, two-way etc. classifications, and periodic regressions. The statistics appear to have power comparable with that of the related circular statistics against alternative hypotheses specified by a stationary Markoff process. In many cases occurring in practice, however, serial correlation of the errors will be due to systematic behaviour arising from the inadequacy of the theoretical model to represent the true relationship. The statistics proposed will often be preferable to circular statistics in such cases.

17. Stochastic Difference Equations with a Continuous Time Parameter. (Preliminary Report). S. G. GHURYE, University of North Carolina.

Given a discrete sequence of observations ordered equidistantly in time, it is often assumed that this discrete process is explained by a stochastic difference equation with a purely random "disturbance". However, this observed discrete process might be the result of observations on a stochastic process $X(t)$ in which t is not discrete, but continuous. Is it possible to have a process $X(t)$, defined for t real, such that given any real t_0 and any real $h > 0$, the sequence $\{X(t_0 \pm jh)\}, j = 0, 1, \cdots$, satisfies the equation

$$X(t_0 + [j + p]h) + \alpha_1(h)X(t_0 + [j + p - 1]h) + \cdots + \alpha_p(h)X(t_0 + jh) = \delta(t_0 + jh),$$

δ being a linear function of mutually independent random variables having a common c.d.f. which is independent of h? The cases $p = 1$ and $p = 2$ are dealt with in detail, and the possible forms of such processes derived; the further problem for any p, as also for a system of equations, is being considered. It is also proposed to tackle the problems of estimation and testing which arise in this connection.

18. Nonsequential Problems in the Case of k Hypotheses. (Preliminary Report). HERMAN CHERNOFF, University of Illinois.

Suppose that there are k possible simple hypotheses $H_1, H_2, \cdots, H_k$ and a possibly infinite set of actions may be taken. To a decision function there corresponds a vector $\rho = (\rho_1, \rho_2, \cdots, \rho_k)$ where ρ_i is the risk if H_i is true. The closure of the range of ρ is convex in the nonatomic case and in the randomized case. In the randomized case the closure of the range of ρ is the convex hull of the closure of the range of ρ in the nonrandomized case. (The randomized case is that one where a number is selected at random from the unit interval before an action is taken.) The range of ρ is closed under suitable closure conditions on the range of the weight function.

19. The Moments of a Multinormal Distribution after One-sided Truncation of Some or All Coordinates. Z. W. Birnbaum and Paul L. Meyer, University of Washington.

Let $X = (X_1, X_2, \cdots, X_p)$ be a multinormal random variable with given first and second moments and the probability density $f(X_1, X_2, \cdots, X_p)$. The random variable $Y = (Y_1, Y_2, \cdots, Y_p)$ is said to be obtained from X by truncation to the set $X_i \geq \tau_i$, $i = 1, 2, \cdots, p$, if its probability density is $g(Y_1, Y_2, \cdots, Y_p) = Cf(Y_1, Y_2, \cdots, Y_p)$ for $Y_1 \geq \tau_1, Y_2 \geq \tau_2, \cdots, Y_p \geq \tau_p$, and $g(Y_1, Y_2, \cdots, Y_p) = 0$ elsewhere. The problem considered is to determine the mathematical expectations $E(Y_i^m Y_j^n)$. Explicit formulae are obtained for the first and second moments $E(Y_i)$ and $E(Y_i Y_j)$, and recursion formulae are given for the general case. (Research done under the sponsorship of the Office of Naval Research.)

20. An Algorithm for the Determination of all Solutions of a Two-Person Zero Sum Game with a Finite Number of Strategies. H. Raiffa, G. L. Thompson, and R. M. Thrall, University of Michigan.

Consider a zero-sum two-person game in which each player has a finite number of strategies. A computational procedure is given for finding the value of the game and all optimal basic strategies for each player. The basic computations required are evaluation of linear forms and solution of linear equations in one unknown. This method, based on geometric reasoning, is a step by step process with no more stages than the total number of strategies for the two players.

21. A Note on the Convolution of Uniform Distributions. Edwin G. Olds, Carnegie Institute of Technology.

Let X_i be independent random variables with probability density functions $[\epsilon(X_i) - \epsilon(X_i - a_i)]/a_i$, where $\epsilon(x - c)$ is unity for $x \geq c$ and zero elsewhere. This paper gives a simple proof that the probability density function for $S = \Sigma_1^n x_i$ is

$$[S^{n-1}\epsilon(S) - \Sigma_1^n (S - a_i)^{n-1}\epsilon(S - a_i) + \Sigma_{i<j} (S - a_i - a_j)^{n-1}\epsilon(S - a_i - a_j) - \cdots + (-1)^n(S - \Sigma a_i)^{n-1}\epsilon(S - \Sigma a_i)]/(n - 1)!\Pi_1^n a_i .$$

A sufficient condition for the asymptotic normality of S is $0 < \alpha \leq a_i \leq \beta$ (finite). For the special case where $a_{i+1} = ra_i$ the necessary and sufficient condition for asymptotic normality is $r = 1$. For $0 \leq r \leq 0.5$ or $r \geq 2$ the probability that S will be outside the interval $\mu_S \pm 3\sigma_S$ is zero. From the Edgeworth Series for the distribution function for the standardized sum it follows that $F(-3) \doteq 0.00135 - 0.004[\Sigma a_i^4/(\Sigma a_i^2)^2]$ where the bracketed expression takes its minimum value n^{-1} when all of the a_i's are equal. These results are useful in connection with the problem of random assembly.

22. On the Consistency of Certain Estimates of the Linear Structural Relation. Elizabeth L. Scott, University of California, Berkeley.

Let $\{x_i, y_i\}$ denote n independent pairs of observations on x, y where $x = \xi + u$ and $y = \alpha + \beta\xi + v$ with ξ, u and v random variables with finite variances, $E(u) = E(v) = 0$ and ξ independent of the pair u, v. *Procedure* (1): Fix $a \leqq b$ such that

$$P\{x \leqq a\} > 0, \qquad P\{x > b\} > 0.$$

Let X_1, Y_1 stand for the arithmetic mean of the x_i's and y_i's, respectively, for $x_i \leqq a$ and X_2, Y_2 for those for which $x_i > b$. As an estimate of β, consider, say, $b_1 =$

$(Y_2 - Y_1)/(X_2 - X_1)$. *Procedure* (2): Let X_1, Y_1 denote arithmetic mean of x_i's and y_i's, respectively, for which x_i is one of the r smallest of the x_i's and X_2, Y_2 for those for which x_i is one of the s largest, with r, s preassigned, $r < n - s + 1$. The corresponding estimate of β is, say, b_2 defined as above. Let (μ, ν) denote the shortest interval such that $P\{\mu \leqq u \leqq \nu\} = 1$. THEOREM 1. *In order that b_1 preserve the property of being a consistent estimate of β irrespective of the value of β, $-\infty < \beta < \infty$, it is n.a.s. that $P\{a - \nu < \xi \leqq a - \mu\} = P\{b - \nu < \xi \leqq b - \mu\} = 0$.* Now let $r = p_1 n$, $s = p_2 n$ and m, M be the corresponding percentile points such that $P\{\xi \leqq m\} = p_1$ and $P\{\xi > M\} = p_2$. THEOREM 2. *If $n \to \infty$ while p_1 and p_2 are held constant, the n.a.s. condition that b_2 preserve the property of being a consistent estimate of β irrespective of the value of β, $-\infty < \beta < \infty$, is that $P\{m - \nu < \xi \leqq m - \mu\} = P\{M - \nu < \xi \leqq M - \mu\} = 0$.* Similar estimates were considered, for $p_1 = p_2 = \frac{1}{2}$, u and v independent, by A. Wald (*Annals of Math. Stat.*, Vol. 11 (1940), pp. 295–297) who showed sufficiency.

23. A 3-decision Problem Concerning the Mean of a Normal Population. R. R. BAHADUR, University of Chicago.

Given n independent observations $x_1, x_2, \cdots, x_n$ from a normal population having an unknown mean $\theta\sigma$ and unknown variance σ^2, suppose that the statistician is asked to say whether the unknown mean is $>c$ or $\leq c$ where c is a given constant (which is supposed henceforth to be zero), or to say that he would rather reserve judgement on the matter. In the present problem (which was suggested by Professor R. C. Bose as a modification of the problems considered in "The Problem of the Greater Mean," [R. R. BAHADUR AND H. ROBBINS, *Annals of Math. Stat.*, Vol. 21 (1950), pp. 469–487]), reserving judgement is considered to be undesirable, and the possibility of doing so is admitted only for the purpose of reducing the probability of the statistician making an incorrect assertion. For any procedure d which associates each sample with one of the three decisions "assert $\theta > 0$", "assert $\theta \leq 0$", and "reserve judgement", let $a(d \mid \theta\sigma, \sigma) = Pr.$("incorrect assertion" using $d \mid \theta\sigma, \sigma$), $b(d \mid \theta\sigma, \sigma) = Pr.$("reserve judgement" using $d \mid \theta\sigma, \sigma$), and set $\alpha(d \mid \theta) = \sup_\sigma \{[a(d \mid \theta\sigma, \sigma) + a(d \mid -\theta\sigma, \sigma)]/2\}$,

$$\beta(d \mid \theta) = \sup_\sigma \{[b(d \mid \theta\sigma, \sigma) + b(d \mid -\theta\sigma, \sigma)]/2\}.$$

The class of procedures $\{d_\tau^*\}$ is defined as follows: for any τ, $0 \leq \tau \leq \infty$, $d_\tau^* \equiv$ "assert $\theta > 0$ if $\bar{x} > \tau s$, assert $\theta \leq 0$ if $\bar{x} \leq -\tau s$, and reserve judgement otherwise", where $\bar{x} = n^{-1}\Sigma_1^n x_i$ and $s^2 = n^{-1}\Sigma_1^n(x_i - \bar{x})^2$. One of the results obtained concerning the class $\{d_\tau^*\}$ is as follows. *Corresponding to any d there exists a d_τ^* such that $\alpha(d_\tau^* \mid \theta) \leq \alpha(d \mid \theta)$ and $\beta(d_\tau^* \mid \theta) \leq \beta(d \mid \theta)$ for all θ.* In particular, given p, $(0 < p < \frac{1}{2})$, there (evidently) exists a $\tau(p)$, $(0 < \tau(p) < \infty)$, such that $\sup_\theta \{\alpha(d_{\tau(p)}^* \mid \theta)\} = p$, and if d is any other procedure such that $\sup_\theta \{\alpha(d \mid \theta)\} \leq p$, then $\beta(d \mid \theta) \geq \beta(d_{\tau(p)}^* \mid \theta)$ for all θ. These results provide a justification of the manner in which the two-sided t test of a normal mean is sometimes used in practice.

24. Consistent Estimate of the Slope of a Linear Structural Relation. J. NEYMAN, University of California, Berkeley, AND CHARLES M. STEIN, University of Chicago.

Let Z_n denote the system of $8n$ independent pairs of measurements (X_{ik}, Y_{ik}), for $i = 1, 2, \cdots, n$ and $k = 1, 2, \cdots, 8$, of two nonobservable random variables ξ_{ik} and $\eta_{ik} = \alpha \operatorname{cosec}\beta - \xi_{ik} \cot\beta$, where α and β are constants. Variable ξ_{ik} is nonnormal. It is assumed that any nonnormal components of the errors of measurement $X_{ik} - \xi_{ik}$ and $Y_{ik} - \eta_{ik}$ are mutually independent, independent of ξ_{ik} and of the normal components of the errors. The normal components of errors may be correlated but as a pair are independent of ξ_{ik}. For every $n \geqq 4$, let $m(n)$ be the greatest integer not exceeding $\sqrt{n}$. Let $\Delta(n) = \pi/(m(n) - 1)$ and $b_{n,j} = -\pi/2 + (j - 1)\Delta(n)$, for $j = 1, 2, \cdots, m(n)$. For every b, $|b| \leqq \pi/2$ and for

$i = 1, 2, \cdots, n$ let $A_i = \exp\{-\frac{1}{2}[(X_{i1} - X_{i2} + X_{i3} - X_{i4}) \cos b + (Y_{i1} - Y_{i2} + Y - Y_{i4}) \sin b]^2 - \frac{1}{2}(X_{i1} - X_{i2} + X_{i5} - X_{i6})^2\}$, $B_i = \exp\{-\frac{1}{2}(Y_{i1} - Y_{i2} + Y_{i7} - Y_{i8})^2\}$ $C_i = \exp\{-\frac{1}{2}(Y_{i1} - Y_{i4} + Y_{i6} - Y_{i7})^2\}$, $D_i = \exp\{-\frac{1}{2}(Y_{i3} - Y_{i4} + Y_{i5} - Y_{i6})^2\}$, and finally, $G(b, Z_n) = [\Sigma_{i=1}^{n} A_i(B_i - 2C_i + D_i)]/n$. Let $g(Z_n)$ be the smallest of the $m(n)$ values of the function $G(b, Z_n)$ computed for $b_{n1}, b_{n2}, \cdots, b_{nm(n)}$ and let $T(Z_n)$ denote the smallest of the b_{nj} for which $G(b_{nj}, Z_n) = g(Z_n)$. THEOREM. *As $n \to \infty$, the function $T(Z_n)$ thus defined is a consistent estimate of β.* The present problem grew out of the problem of identifiability of β studied by Olav Reiersøl (*Econometrica*, Vol. 18 (1950), pp. 375–389). The results obtained here represent a generalization of the previous results of one of the authors presented at the I.M.S. meeting in Boulder, Colorado, as the Second Rietz Memorial Lecture, September, 1949.

25. **A Remark on Almost Sure Convergence.** Michel Loève, University of California, Berkeley.

A criterion for almost sure convergence is given. It contains criteria of Kolmogorov, Marcinkiewicz, and P. Lévy.

26. **A Significance Test for Differences Among Ranked Treatments in an Analysis of Variance.** D. B. Duncan, Virginia Polytechnic Institute.

Given a set of n treatment means (or totals) $x_1, x_2, \cdots, x_n$, it is often desired to decide whether each of the differences $x_j - x_i$ is significant, that is, whether each of the hypotheses $H:\mu_j > \mu_i$, $i, j = 1, 2, \cdots, n$, $i \neq j$ can be accepted. A test is obtained for this purpose under the conditions which usually apply or are taken to apply in many analyses of variance, namely that $x_1, x_2, \cdots, x_n$ is a random sample from n normal populations with means $\mu_1, \mu_2, \cdots, \mu_n$, respectively, and a common unknown variance σ^2 for which the common form of independent estimate s^2 based on n_2 degrees of freedom is available. In approaching the problem the complete Wald multiple decision function form of analysis is found to be too unwieldy for a general case and is waived in favor of a simpler set of requirements. These state that an α level test should provide likelihood ratio tests as closely as possible for each of the ${}_nC_r$ hypotheses that any combination of r of the treatment means are equal. Also satisfactory upper limits should be placed on the significance level of the whole test with respect to each of these particular ${}_nC_r$ hypotheses. The test obtained satisfies the given requirements better than other currently available procedures. It consists of a fairly simple sequence of range-like tests followed by variance tests which are presented in detail together with examples.

27. **On Information and Sufficiency.** S. Kullback, George Washington University, and R. A. Leibler, Washington, D. C.

For probability spaces $(\mathbf{X}, \mathbf{S}, \mu_i)$, $i = 1, 2$, and probability measures λ, μ_1, μ_2 absolutely continuous with respect to each other in pairs, f_i, $i = 1, 2$, is defined by

$$\mu_i(E) = \int_E f_i(x)\, d\lambda(x) \quad \text{for all} \quad E \in \mathbf{S}.$$

Then $I_{1:2}(E) = [1/\mu_1(E)] \int_E f_1(x)[\log f_1(x) - \log f_2(x)]\, d\lambda(x)$ for $\mu_1(E) > 0$, and $I_{1:2}(E) = 0$ for $\mu_1(E) = 0$, is defined as the mean information for discrimination between H_1 and H_2 per observation from $E \in \mathbf{S}$ for μ_1, where H_i is the hypothesis that x is selected from the population with probability measure μ_i. $J_{12}(E)$, the divergence between the populations in E, is defined as $I_{1:2}(E) + I_{2.1}(E)$ or

$$J_{12}(E) = \int_E [f_1(x)/\mu_1(E) - f_2(x)\mu_2(E)][\log f_1(x) - \log f_2(x)]\, d\lambda(x).$$

Properties of I and J are considered and the relations of I to the information notions of Fisher, Shannon and Wiener and J to Mahalanobis' generalized distance are noted. In particular it is proved that a transformation T never increases $I_{1:2}(\mathbf{X})$ and a necessary and sufficient condition that T leave $I_{1:2}(\mathbf{X})$ unchanged is that T be a sufficient statistic.

28. Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. T. W. ANDERSON, Columbia University, AND D. A. DARLING, University of Michigan.

The statistical problem treated is that of testing the hypothesis that a sample of n independent, identically distributed random variables have the common continuous distribution function $F(x)$. If $F_n(x)$ is the empirical cumulative distribution function and $\psi(x)$ is some nonnegative weight function $(0 \leqq x \leqq 1)$, we consider

$$K_n = n^{\frac{1}{2}} \sup_{-\infty<x<\infty} \{| F(x) - F_n(x) | \psi^{\frac{1}{2}}[F(x)]\}$$

and $W_n^2 = n \int_{-\infty}^{\infty} [F(x) - F_n(x)]^2\psi[F(x)]\, dF(x)$. For suitable choices of ψ these tests have been considered by Kolmogorov, Cramér, von Mises, Smirnov, and others. A unified method for calculating the limiting distributions of K_n and W_n^2 is developed by reducing them to corresponding problems in stochastic processes, which in turn lead to more or less classical eigen-value and boundary value problems for special classes of differential equations. For certain weight functions we give explicit limiting distributions. For $\psi \equiv 1$ we obtain, e.g., the Kolmogorov distribution and the ω^2 distribution of Smirnov and von Mises for K_n and W_n^2, respectively. By courtesy of the numerical analysis section of the Rand Corporation a tabulation of the ω^2 distribution has been prepared. (This work was supported by the Rand Corporation.)

29. The Effect of Preliminary Tests of Significance on the Size and Power of Certain Tests of Univariate Linear Hypotheses with Special Reference to the Analysis of Variance. (Preliminary Report). ROBERT E. BECHHOFER, Columbia University.

Let $X_1, \cdots, X_q; Y_1, \cdots, Y_r; Z_1, \cdots, Z_s$ be normally and independently distributed with means $0, \cdots, 0; \mu_1, \cdots, \mu_r; \nu_1, \cdots, \nu_s$, respectively, and variance σ^2. The null hypothesis is $H_0: \nu_1 = \cdots = \nu_s = 0$. The standard test (T_1) of H_0 is an F-test involving $\Sigma_{k=1}^s Z_k^2/\Sigma_{i=1}^q X_i^2$. If $\mu_1 = \cdots = \mu_r = 0$, a more powerful test (T_2) of H_0 is an F-test involving $\Sigma_{k=1}^s Z_k^2/(\Sigma_{i=1}^q X_i^2 + \Sigma_{j=1}^r Y_j^2)$. However, if $\Sigma_{j=1}^r \mu_j^2/\sigma^2$ should be large, T_2 would have low power. When the statistician believes (based on past experience) that the μ's equal zero, but wishes to protect himself against the possibility that they do not, he can use a preliminary F-test (T_0), i.e., he pools (uses T_2) or does not pool (uses T_1) accordingly as $\Sigma_{j=1}^r Y_j^2/\Sigma_{i=1}^q X_i^2$ is less than or greater than some preassigned constant. The power of the composite test $[T = (T_0 \text{ plus } T_1 \text{ or } T_2)]$ depends on q, r, s; the levels of significance $\alpha_0, \alpha_1, \alpha_2$ associated with T_0, T_1, T_2, respectively; and $\lambda_2 = \Sigma_{j=1}^r \mu_j^2/2\sigma^2$ (the nuisance parameter) and $\lambda_3 = \Sigma_{k=1}^s \nu_k^2/2\sigma^2$. Formulae are derived for the size (Type I error) and power of T. The behavior of the size and power as a function of λ_2 and λ_3 is characterized. It is shown that certain choices of $\alpha_0, \alpha_1, \alpha_2$ yield tests T which have desirable properties. (Part of this work was carried out under the sponsorship of the Office of Naval Research.)

30. The Exact Distribution of the Extremal Quotient. E. J. GUMBEL, New York, AND L. H. HERBACH, Columbia University.

The distribution of the extremal quotient q (the ratio of the largest value x_n to the smallest x_1 of n independent observations taken from the same distribution), is obtained in four stages, three special cases: (1) $x_1 \geqq 0, x_n \geqq 0, q \geqq 1$. (2) $x_1 \leqq 0, x_n \leqq 0,$

$0 \leqq q \leqq 1$. (3) $x_1 \leqq 0, x_n \geqq 0, q \leqq 0$, culminating in the general case: (4) $-\omega_1 \leqq x_1 \leqq x_n \leqq \omega_2$, $-\omega_2/\omega_1 \leqq q < \infty$. The common procedure in the first three cases is to integrate out the extreme from the joint distribution of one extreme and the extremal quotient. Geometric considerations give the appropriate regions of integration. The general case is obtained by a composition of cases (3), (2), and (1). For symmetrical initial distributions there exist only two branches which join at $q = 1$, and the probability function may be written in a symmetrical form. When $n = 2$, the distribution of q for a symmetrical distribution is symmetrical about zero and invariant under a reciprocal transformation, and if the initial distribution possesses no moments and does not vanish at $x = 0$, the density of probability becomes infinite at $q = 0$. The distribution of q is not affected by changes in scale but is very sensitive to changes in origin. For a uniform distribution, the extremal quotient of a nonnegative variate has just the opposite qualities of the extremal quotient of a nonpositive variate. For variates changing sign, the extremal quotient is asymptotically negative.

31. **The Distributions of the t and F Statistics for a Class of Nonnormal Populations.** RALPH A. BRADLEY, Virginia Polytechnic Institute.

Series expansions of the cumulative distribution functions of t and of F in powers of t^{-1} and F^{-1} are obtained. The general method of derivation presented is valid for populations with density functions, $f(u)$, such that $f(u) > 0$, $f(u)$ is continuous, and has continuous derivatives for all values, $-\infty < u < \infty$. The coefficients of terms in these expansions are reduced from integrals, of multiplicity equal to the sample size, to products of coefficients, common to all populations of the class defined above, and integrals of no greater multiplicity than the number of groups of observations in the sample. Selected values of the common coefficients are given as well as illustrative examples for the Cauchy and "squared hyperbolic secant" population.

32. **Note on the Behavior of the Characteristic Function of a Random Variable at Zero.** M. ROSENBLATT, University of Chicago.

Let X be a random variable with characteristic function $\phi(z)$. Let $X_n = X$ when $|X| < n^{1/\alpha}$ and let $X_n = 0$ when $|X| \geq n^{1/\alpha}$. The following theorems are proved: (1) $1 - \phi(z) = o(|z|^\alpha)$, $0 < \alpha < 1$, at $z = 0$ if and only if $n \cdot Pr(|X| > n^{1/\alpha}) = o(1)$. (2) $1 - \phi(z) = o(|z|^\alpha)$, $1 \leq \alpha < 2$, at $z = 0$ if and only if $n \cdot Pr(|X| > n^{1/\alpha}) = o(1)$ and $E(X_n) = o(1)$. The results are obtained by making use of W. Feller's necessary and sufficient conditions for the weak law of large numbers (see W. FELLER, *Acta Univ. Szeged*, Vol. 8 (1937), pp. 191-201).

NEWS AND NOTICES

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

Personal Items

Dr. R. R. Bahadur, who received his Ph.D. in mathematical statistics from the University of North Carolina in June, 1950, is now an instructor in the Committee on Statistics of the University of Chicago.

Dr. T. A. Bancroft, Associate Professor of Statistics, Iowa State College, has been appointed Head of the Department of Statistics and Director of the Statistical Laboratory at Iowa State College.