REFERENCES

[1] J. H. CADWELL, "The distribution of quasi-ranges in samples from a normal population," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 603–613.
[2] H. O. HARTLEY, "The range in normal samples," *Biometrika*, Vol. 32 (1942), pp. 334–348.
[3] K. PEARSON, *Tables for Statisticians and Biometricians*, Part II, Biometrika Office, University College, London, 1931.

ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Montreal meeting of the Institute, September 10–13, 1954)

1. On Quadratic Estimates of Variance Components in Balanced Models, A. W. Wortham, Chance Vought Aircraft and Oklahoma A and M College.

A balanced model is defined as a model whose analysis of variance mean squares are symmetric in the squares of the observations. Included in this class of models are: (1) Completely Randomized, (2) Randomized Blocks, (3) Latin Squares, (4) Graeco-Latin Squares, (5) Split Plots, (6) Factorial Arrangements, etc.

The "analysis of variance estimates" of the variance components are the estimates obtained by solving the system of equations which result when the observed and expected mean squares in the analysis of variance table are equated. For any infinite population let the general balanced model be $y_{i_1 i_2 \cdots i_n} = \mu + \sum_{k=1}^{n} A_{kik} + e_{i_1 i_2 \cdots i_n}$, where μ is a constant, A_{kik} and $e_{i_1 i_2 \cdots i_n}$ are independent random variables with zero means, finite fourth moments, and variances σ_k^2 and σ_0^2 respectively. Let $\hat{\sigma}_k^2$ and $\hat{\sigma}_0^2$ be "the analysis of variance estimates" of the variance components σ_k^2 and σ_0^2. It is shown that the quadratic estimate of $\sum_{k=0}^{p} g_k \sigma_k^2$ (g_k known) which is unbiased, independent of μ, and has minimum variance is given by $\sum_{k=0}^{p} g_k \hat{\sigma}_k^2$. That is, the best quadratic unbiased estimate of the linear combination of the variance components is given by the same linear combination of "the analysis of variance estimates" of the variance components.

2. The Coefficients in the Best Linear Estimate of the Mean in Symmetric Populations, A. E. Sarhan, University of North Carolina.

In a previous paper ("Estimation of the Mean and Standard Deviation by Order Statistics" by A. E. Sarhan, *Ann. Math. Stat.* Vol. 25 (1954), pp. 317–328) the best linear estimate of the mean of a rectangular, triangular and double exponential population were worked out. By considering some other symmetric distributions with different shapes, it is found that the coefficients in the estimates form a sequence. From the sequence, it is observed that the coefficients in the estimates are influenced by the shape of the distribution. The variances of the estimates are also so affected.

3. Distribution of Linear Contrasts of Order Statistics, Jacques St. Pierre, University of North Carolina.

Consider $n + 1$ independent normal populations with unknown means, $m_0, m_1, \cdots, m_n$, respectively, and with a common known variance $\sigma^2 = 1$ (say). Suppose a sample of size N is available from each population; and let $x_{(0)} > x_{(1)} > \cdots > x_{(n)}$ be the ordered sample means. Consider the linear contrasts $z = x_{(0)} - c_1 x_{(1)} - \cdots - c_n x_{(n)}$, where $\sum_{i=1}^{n} c_i = 1$, $c_i \geqq 0$, $(i = 1, 2, \cdots, n)$. The probability density function of the contrasts z is derived under the null hypothesis H_0: $m_0 = m_1 = \cdots = m_n$. The density of the contrasts z is also obtained in the case of three populations, under the hypothesis $-\infty < m_2 \leqq m_1 \leqq m_0 < +\infty$. Particular hypotheses are considered and tables are given. Finally, the particular contrast $y = x_{(0)} - x_{(1)}$ is considered in the general case.

4. Note on Fourier Periodogram Analyses of Time Series, B. F. Kimball, New York State Public Service Commission.

R. A. Fisher's treatment of the probability distribution of the squares of the amplitudes of the Fourier harmonics R_n^2 is followed. One deals with a time series y_i of N observations. The null hypothesis is taken as the hypothesis that $E(y_i) = 0$ and that the y_i are independently and normally distributed with constant variance. Let the index n of R_n^2 denote the index of the fitted harmonic such that N/n denotes the period of this harmonic. If N/n is an integer one can replace the series of N terms by one of N/n *means* $\sum y_i/n$ where the y_i are of the same phase in period N/n. The harmonics of this series are selected harmonics of the original series. This paper examines the implications of such a breakdown for the testing of the significance of the short period harmonics relative to the null hypothesis.

5. Univariate Two-Population Distribution-Free Discrimination, David S. Stoller, RAND Corporation.

A univariate random variable, z, is defined by the composite cumulative distribution function, $F(z) = \theta F_1(z) + (1 - \theta)F_2(z); 0 \leqq \theta \leqq 1$. Restrict F_1 and F_2 to be such that the optimum a priori discriminating regions are $S_1 = \{z \mid z \leqq \zeta\}$ and $S_2 = \{z \mid z > \zeta\}$, where ζ is unique. Optimum discrimination is defined as that which maximizes the probability of correctly classifying z. Denote the above maximum probability by $Q(\zeta)$. Given an independent random sample of size N from $F(z)$, each member of which is classifiable, a distribution-free estimate of ζ, denoted by ζ^*, is constructed as follows. Let $t(z) = k(z) - h(z)$, where $k(z)$ is the number of observations from the first population (i.e., that defined by F_1) which are less than or equal to z, and $h(z)$ is similarly defined for the second population. Then ζ^* is any value of z that maximizes $t(z)$. The estimate, ζ^*, possesses two asymptotically optimum properties: (1) the probability of correct classification induced by using ζ^* instead of ζ converges in probability to $Q(\zeta)$, and (2) the quantity, $t(\zeta^*)$, may be used to construct an estimate of $Q(\zeta)$ which converges in probability to $Q(\zeta)$.

6. New Types of Easily Constructed Partially Balanced Incomplete Block Designs, John Mandel and Marvin Zelen, National Bureau of Standards.

In the planning of experiments in the physical sciences one is often confronted with natural limitations on the size of experimental blocks. Therefore, the use of incomplete blocks is becoming ever more widespread in this type of application. In this paper a type of partially balanced incomplete block design is introduced, the construction of which consists in replacing each treatment of a balanced design with a group of treatments which themselves form a balanced design. A large class of designs thus becomes at once available by combining Latin Squares with Youden Squares, or Youden Squares with Youden Squares. An important property of these designs is the possibility of two-way elimination of error (according to rows and columns). A general formula is given for the Least-squares estimation of corrected treatment effects. Because of the flexibility of the proposed designs, their ease of construction, and simplicity of analysis, they are well adapted to experiments in physical and chemical laboratories. Investigations are in progress to extend the results to designs formed from combinations of chain block and other partially balanced incomplete block designs.

7. The Stochastic Convergence of a Function of Sample Successive Differences, Lionel Weiss, University of Virginia.

Let $f(x)$ be a bounded density function with at most a finite number of discontinuities, and such that there are two finite numbers, A and B $(A < B)$, with $f(x)$ nondecreasing in the interval $(-\infty, A)$ and nonincreasing in the interval (B, ∞). Let $X_1, X_2, \cdots, X_n$ be independent chance variables each with the density $f(x)$. Define $Y_1 \leqq Y_2 \leqq \cdots \leqq Y_n$ as the ordered values of $X_1, X_2, \cdots, X_n$; T_i as $Y_{i+1} - Y_i$; and $R_n(t)$ as the proportion of the values $T_1, \cdots, T_{n-1}$ not greater than $t/(n-1)$. $S(t)$ denotes $\left[1 - \int_{-\infty}^{\infty} f(x)e^{-tf(x)}\,dx\right]$; and $V(n)$ denotes $\sup_{t \geqq 0} | R_n(t) - S(t) |$. Then it is shown that $V(n)$ converges stochastically to zero as n increases. This result can be used to demonstrate the stochastic convergence of various functions of $T_1, \cdots, T_{n-1}$, including some which have been discussed in the literature.

8. On a Modified T^2 Problem, Ingram Olkin, Michigan State College and S. S. Shrikhande, College of Science, Nagpur.

Consider two independent random vectors $X = (X_1, \cdots, X_p)$, $Y = (Y_1, \cdots, Y_p)$, each obeying a p-variate normal probability law with $EX = (\theta_1, \cdots, \theta_k, \mu_{k+1}, \cdots, \mu_p)$, $EY = (\theta_1, \cdots, \theta_k, \nu_{k+1}, \cdots, \nu_p)$, and same covariance matrix Σ, with all the parameters unknown. On the basis of a sample of n and m observations from X and Y, respectively, the hypothesis $H_0 : \mu_i = \nu_i$ against $H_1: \mu_i \neq \nu_i$ $(i = k+1, \cdots, p)$ is to be tested. The problem is equivalent to the case where X and Y are random vectors with means $EX = (0, \cdots, 0, \phi_{k+1}, \cdots, \phi_p)$, $EY = (0, \cdots, 0)$, and same covariance matrix Σ. On the basis of one and n observations from X and Y, respectively, H_0: $\phi_i = 0$ against H_1: $\phi_i \neq 0$ $(i = k+1, \cdots, p)$ is to be tested. The likelihood ratio statistic is obtained and its distribution under H_0 and H_1 derived. If $k = 0$, the statistic reduces to Hotelling's T^2 statistic.

9. The Validity of Sheppard's Corrections for Grouping, F. J. Anscombe, University of Cambridge and Princeton University.

The moments of an absolutely continuous one-dimensional distribution are to be compared with the moments of the same distribution when it has been "grouped" with constant grouping interval. The characteristic function $\theta^*(t)$ of the grouped distribution may be expressed as a Fourier series in terms of the characteristic function $\theta(t)$ of the original distribution. The expansion is similar to those for moments given by R. A. Fisher (*Philos. Trans. Roy. Soc. London, Ser. A.*, Vol. 222 (1922), pp. 309–368, section 5), but requires no condition on the original distribution other than absolute continuity of the distribution function $F(x)$. Sheppard's formulas are obtained when the periodic terms in the series are neglected. The periodic terms are small if $| \theta(t) |$ is small for large t, and this condition is related to the differentiability of $F(x)$ for all values of x. The emphasis that has often been placed on the differentiability of $F(x)$ at infinity or at the ends of a finite range is misleading, because these points are not specially important.

10. Unbiased Tests Based on Unbiased Estimators, Reed B. Dawson, Department of Defense.

A test of a point-hypothesis $\theta = \theta_0$ of a distribution parameter will be said to be *strongly unbiased* when the power depends on θ alone and exceeds the size against all alternatives. For any α, $0 < \alpha < 1$, there exists a strongly unbiased test of size α if and only if there exists a real-valued function $f(\theta)$ which is zero at θ_0, strictly positive elsewhere, and possesses a bounded unbiased estimator. For, if $\omega(x)$ is the rejection probability corresponding to an outcome x, let $f = E\omega - \alpha$; if f is given, take $\omega = \alpha + K\hat{f}$, where $\hat{f}(x)$ is the estimator and K is a suitable positive constant. One application concerns a sample of n items from the family of all distributions over the unit interval. The possible strongly unbiased tests of a point-hypothesis on the rth moment form a bounded convex body in E_n over which the power is a linear functional. A second application (Mosteller's suggestion) concerns the hypothesis of independence of two attributes in a 2 x 2 table where sampling proceeds until a chosen cell attains a fixed quota. Powers of the determinant of the underlying probabilities admit bounded unbiased estimation, giving unbiased tests without the Neyman structure.

11. **The Mean Square Error of the Sample Median,** Harold Hotelling, University of North Carolina.

For random samples of any odd number from an arbitrary population, the ratio of the mean square error in the sample median, regarded as an estimate of the population median, to the corresponding population parameter, is shown never to be less than unity. This lower bound is actually attained for the familiar two-point distribution with equal probabilities. The fact that in this case the accuracy, however measured, of the median of a large number of observations is no better than that of one random observation destroys the argument sometimes given that the median should be used in the absence of knowledge of the form of the underlying distribution. (Research sponsored by the Office of Naval Research at Chapel Hill, North Carolina).

12. **The Moments of the Sample Median,** J. T. Chu and Harold Hotelling, University of North Carolina.

Moments of medians of random samples are studied by a method involving expansion about $\frac{1}{2}$ of the inverse of the cumulative distribution function, and in other ways. Readily calculable approximations are found, both for large and for small samples, with close upper and lower bounds on the errors of approximation. The asymptotic behvaior for large samples is examined. Calculations are carried out for the Laplace, Cauchy and normal distributions. (Research sponsored by the Office of Naval Research at Chapel Hill, North Carolina).

13. **Distribution of the Largest Vote in Unstructured Random Balloting,** Leo Katz, Michigan State College

The exact distributions of the maximum vote are obtained for two balloting arrangements. In both, each person votes once at random without prior reduction of the field of choice by a nomination process. In the first arrangement, a person may, if he (randomly) wishes, vote for himself; in the second, voting is gentlemanly. The second case has direct application to determination of "stars" in sociometric testing. An approximation is given; it is shown to be reasonably accurate for moderate-sized groups.

14. **Statistical Programming,** D. F. Votaw, Jr., Yale University.

Statistical programming problems arise when some of the constants in a programming problem are unknown but statistical information about them is available. In this paper several methods of statistical programming are compared in connection with a special linear programming problem. The application of simultaneous confidence interval estimation is discussed. (Work sponsored by the Office of Naval Research.)

15. Exact Tests of Significance for Combining Inter- and Intra-Block Information in Incomplete Block Designs (Preliminary Report), Marvin Zelen, National Bureau of Standards.

Consider an incomplete block design where the number of blocks is greater than the number of treatments ($b > v$). It is then shown under the usual assumptions for the recovery of inter-block information that two independent F tests of the null hypothesis (all treatments are the same) exist; one using only inter-block information and the other using the intra-block information. Let $F_i (i = 1, 2)$ represent the F ratio obtained for each test; $1 + r_i\lambda/\sigma_i^2$ represent the expected value of the numerator to the denominator of respective F ratios, where $\lambda = \Sigma(t_i - \bar{t})^2/v - 1$ is a measure of the departure from the null hypothesis (i.e., $\lambda = 0$ if H_0 is true); also let $p_i = P\{F \geqq F_i \mid H_0\}$. Then a combined test which seems to adjust for the differences in power of the two independent tests is given by the region $\{p_1 p_2^\theta \leqq Q\}$, where Q is chosen such that $P\{p_1 p_2^\theta \leqq Q\} = Q - \theta^{1/\theta}/1 - \theta = \alpha$ (level of significance), and $\theta = r_2\sigma_1^2/r_1\sigma_2^2$. For example, $\theta = 1 - \mathrm{E}/\mathrm{E}[\sigma^2/\sigma^2 + k\sigma_b^2]$ for balanced incomplete block designs where σ^2 is the "within block" variance, σ_b^2 the "between blocks" variance component, E is the efficiency factor and k is the plot size. Approximations to the power function of the test have been derived and preliminary calculations indicate that the above critical region seems to have greater power as compared to weighting the individual p_i's equally as in Fisher's method.

16. Moments and Related Quantities of the Null Distribution of Linear Contrasts of Order Statistics in the Case of Three Populations, Jacques St. Pierre, University of North Carolina.

Consider three independent normal populations with unknown means, m_0, m_1, m_2 respectively, and with a common known variance $\sigma^2 = 1$ say. Suppose a sample of size N is available from each population. Let $x_{(0)} > x_{(1)} > x_{(2)}$ be the ordered sample means. Consider the linear contrasts $z = x_{(0)} - cx_{(1)} - (1 - c)x_{(2)}$, where $0 \leqq c \leqq 1$. An expression for the kth moment about the origin is obtained. Properties of the moments and related quantities (skewness and kurtosis coefficients) are established, considering these quantities as functions of the nonstochastic parameter c. Tables of moments of low order are given in cases of special interest.

17. Application of Faà di Bruno's Formula in Mathematical Statistics, Eugene Lukacs, Office of Naval Research.

Let $z = G(y)$ and $y = f(x)$ be two functions such that all the derivatives of $G(y)$ and $f(x)$ up to order p exist. We denote by $D_t^k\{\ \}$ the operation of determining the kth derivative of the function in the braces with respect to t and we write $f_v = D_t^v\{f(t)\}/v!$, $f = f_0 = f(t)$ then $D_t^p z = D_t^p\{G[f(t)]\} = \Sigma p! D_y^k\{G(y)\} f_{k_1}^{i_1} \cdots f_{k_s}^{i_s}/(i_1! \cdots i_s!)$ where the summation is to be extended over all partitions of p such that $i_1 + i_2 + \cdots + i_s = k$ and $i_1k_1 + i_2k_2 + \cdots + i_sk_s = p$. This formula is due to F. Faà di Bruno [Sullo Sviluppo delli Funzioni, *Annali di science mathematiche e fisiche 6* (1855), pp. 479–480,]. Faà di Bruno's formula can be applied in mathematical statistics. The relations between the cumulants and the moments of a distribution are derived easily by means of this formula. It is also useful in the study of R. A. Fisher's k-statistics. For instance, the explicit formula, expressing the k-statistic of order p in terms of the observations, can be obtained. In addition to these familiar results, the following theorem is proven. Let $x_1, x_2, \cdots, x_n$ be n independent observations taken from a population with distribution function $F(x)$ and denote by p an integer greater than one. Assume that the pth moment of $F(x)$ exists. The population is normal if, and only if, the k-statistic of order p is independent of the sample mean.

18. On Simultaneous Minimax Point Estimation, Waldo A. Vezeau and Koichi Ito, St. Louis University.

This paper is concerned with simultaneous minimax point estimation of all the parameters in the multivariate distribution function of a parent population on the basis of a sample of fixed size. Extending results due to K. Miyasawa (*Bull. Math. Stat.*, Vol. 5(1953), pp. 1–17), it is shown that if the risk is a bounded function of s parameters, $\theta_1, \cdots, \theta_s$ and their point estimates, $d_1, \cdots, d_s$, and a convex, measurable function of $d_1, \cdots, d_s$ for any fixed $\theta_1, \cdots, \theta_s$, and if the space D of $d_1, \cdots, d_s$ is compact and convex, then there exists a set of simultaneous minimax point estimates of $\theta_1, \cdots, \theta_s$ in D. Applications of this theorem are made to simultaneous minimax point estimation of the parameters in a multinomial distribution, the mean and variance (or standard deviation) of a univariate normal distribution, and the means, variances and covariances of a multivariate normal distribution.

19. Estimation of Structural Parameters when the Number of Incidental Parameters is Unbounded, J. Wolfowitz, Cornell University.

Let $\prod_{i=1}^{n} \prod_{j=1}^{m_i} f(z_{ij} \mid \theta, \alpha_i)$ be the frequency function of the observed chance variables $\{X_{ij}\}$, $i = 1, \cdots, n$; $j = 1, \cdots, m_i$, which depends upon the unknown (structural) parameter θ and the unknown (incidental) parameters $\{\alpha_i\}$. The author proves that in general there exists no estimator of θ which is efficient for all sequences $\{\alpha_i\}$. This verifies a conjecture of the author's, described in the *Proc. Roy. Dutch Acad. Sci.*, Ser. A, Vol. 56, No. 2, and *Indag. Math.*, Vol. 15, 1953, where a heuristic supporting argument was given.

20. On Power Properties of Certain Simultaneous Tests, K. V. Ramachandran, University of North Carolina.

(1) Let $y_1, y_2, \cdots, y_K$ be k independent normal variates with $E(y_i) = \mu_i$ and $v(y_i) = \sigma^2 (i = 1, 2, \cdots, K)$. μ_i and σ^2 are unknown but an independent estimate s^2 of σ^2 with v d.f. is available. To test the hypothesis: $\mu_1 = \mu_2 = \cdots = \mu_K$ we have a short cut test of Tukey based on the studentized range. (2) Let $y_1, y_2, \cdots, y_K$ be k independent normal variates with variances $\sigma_1^2, \sigma_2^2, \cdots, \sigma_K^2$ respectively. To test the hypothesis: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2$ we have the $F_{\max}$ ratio test of Hartley. In this paper the following properties of the tests are proved. The power function of the tests depend only on $k - 1$ parameters, namely, $\delta_{i-1} = \mu_i - \mu_1 (i = 2, 3, \cdots, K)$ in case (1) and $\eta_{i-1} = \sigma_i^2/\sigma_1^2$ $(i = 2, 3, \cdots, K)$ in case (2). The tests are completely unbiased but the power functions do not have the monotonicity property. A set of useful lower bounds are obtained for the power in the two situations. Power properties of multivariate and other generalizations of these tests are being investigated.

21. On Tests of Normality and Other Tests of Goodness of Fit Based on Distance Methods, M. Kac, J. Kiefer, and J. Wolfowitz, Cornell University.

The authors study the problem of testing whether the common distribution function (d.f.) of the observed independent chance variables $x_1, \cdots, x_n$ is a member of a given class. A classical problem is concerned with the case where this class is the class of all normal d.f.'s, and for the sake of brevity the description in this abstract will be limited to some of the results for this problem. For any two d.f.'s $F(y)$ and $G(y)$ let $\delta(F, G) = \sup_y \mid F(y) - G(y) \mid$. Let $N(y \mid \mu, \sigma^2)$ be the normal d.f. with mean μ and variance σ^2. Define $\bar{x} = n^{-1}\Sigma_1^n x_i$, $s^2 = n^{-1}\Sigma_1^n x_1^2 - \bar{x}^2$. Let $G_n^*(y)$ be the empiric d.f. of $x_1, \cdots, x_n$. The authors consider, inter alia, tests of normality based on $v_n = \delta(G_n^*(y), N(y \mid \bar{x}, s^2))$ and on $w_n = \int (G_n^*(y) - N(y \mid \bar{x}, s^2))^2 \, d_y N(y \mid \bar{x}, s^2)$. It is shown that the asymptotic power of these tests is considerably greater than that of the optimum χ^2 test. The covariance function of a certain Gaussian process $Z(t)$, $0 \leqq t \leqq 1$, is found. It is shown that the sample functions of $Z(t)$ are continuous with probability one, and that, as $n \rightarrow \infty$, $\lim P\{nw_n < a\} = P\{W < a\}$, where $W = \int_0^1 [Z(t)]^2 \, dt$. Tables of the distribution of W and of the limiting distribution of $\sqrt{n}\, v_n$ are given. The role of various metrics is discussed.

22. Tolerance Regions (Preliminary Report), D. A. S. Fraser and I. Guttman, University of Toronto.

Three definitions are considered for tolerance regions. A "distribution-free tolerance region" has the distribution of its probability content independent of the parameter. A "β-content tolerance region" has probability content at least β with an assigned level of confidence. A "β-expectation tolerance region" has probability content on the average equal to β. For the first definition a necessary and sufficient condition has been obtained for the characteristic function of the region. For sampling from univariate distributions for which the order statistics are complete, the nonexistence of distribution-free tolerance regions was obtained in the discontinuous case and some results on distribution-free tolerance bounds were obtained in the continuous case. For the third definition an analogy with hypothesis testing has been established by introducing a density function to indicate the desirability that different points of a distribution be included in the region. For normal distributions the center of the distribution was weighted more heavily than the tails and most stringent tolerance regions obtained. For univariate distributions they were $[\bar{X} \pm \lambda\sigma]$ and $[\bar{X} \pm \lambda s_x]$ depending on whether or not σ was known. In the multivariate case they are based on Hotelling's T^2 statistic.

23. Comparison of the Power of Nonparametric Two Sample Tests against Normal Alternatives, Benjamin Epstein, Wayne University.

This is a sampling study in which we compare the power of run, rank sum, exceedance, and truncated maximum deviation two sample tests. The particular case studied involves normal alternatives whose distance apart is measured by the difference in population means. Two hundred random samples of size 10 are drawn from each population. These results are related to recent work of Dixon and Teichroew [Abstract, *Ann. Math. Stat.* Vol. 25, (1954), p. 175]. There are, however, these differences: (i) in the present study we assume that the two samples have been placed (simultaneously) on life test, thus making the times to failure available in an ordered way and (ii) include exceedance and truncated maximum deviation rules among the nonparametric tests. Such rules are particularly useful in life test situations. Experimental sampling assigns the following order to the power (best to worse): rank sum, untruncated maximum deviation, truncated maximum deviation, exceedance, and run. The first four power curves are fairly close together and are all substantially better than the power curve for the run test. Also included in the paper is experimental information on the expected number of items failed in reaching a decision when an exceedance or truncated maximum deviation rule is used. Substantial savings in this direction are possible. (Research sponsored by the Office of Ordnance Research, U. S. Army).

24. On the Distribution of Radial Errors Having Normally Distributed Components, A. C. Cohen, Jr., University of Georgia.

For a set of p independent random variables x_j $(j = 1, 2, \cdots, p)$, each of which is normal $(0, \sigma)$, the radial error defined as $r = (x_1^2 + x_2^2 + \cdots + x_p^2)^{1/2}$ is considered. It is well known that the distribution of r is given by $[2r/\sigma^2]f_p(r^2/\sigma^2)$ where $f_p(\chi^2)$ is the χ^2 frequency function with p degrees of freedom. This paper is concerned with the problem of estimating the scale parameter σ from unrestricted (complete), truncated, and censored samples of r. Maximum likelihood estimators are developed for each of these cases, and asymptotic estimate variances are given. In the case of unrestricted samples, $(pn\hat{\sigma}^2/\sigma^2)$ has a χ^2 distribution with pn degrees of freedom, where n is the number of sample observations and $\hat{\sigma}^2$ is the maximum likelihood estimate. Tables and graphs of functions necessary for solving the maximum likelihood estimating equations for truncated samples are given for $p = 2$ and $p = 3$. Illustrative examples relating to target analysis studies are included.

25. Confidence Bounds on Departures from a Particular Kind of Multi-Collinearity of Means, S. N. Roy, University of North Carolina.

For k $(p + q)$-variate $N(\underset{\sim}{\xi}_i, \Sigma)$, where $\Sigma((p + q) \times (p + q))$ is symmetric p.d. with submatrices $\Sigma_{11}(p \times p)$, $\Sigma_{22}(q \times q)$ and $\Sigma_{12}(p \times q)$, and $\underset{\sim}{\xi}_i((p + q) \times 1)$ has column subvectors $\underset{\sim}{\xi}_{1i}(p \times 1)$ and $\underset{\sim}{\xi}_{2i}(q \times 1)$, we can set, in the following way, confidence bounds on $\underset{\sim}{\xi}_{1i} - \Sigma_{12}\Sigma_{22}^{-1}\underset{\sim}{\xi}_{2i}$ which are departures from the hypothesis $\underset{\sim}{\xi}_{1i} - \Sigma_{12}\Sigma_{22}^{-1}\underset{\sim}{\xi}_{2i} = 0$ $(i = 1, 2, \cdots, k)$. Let S_{11}, S_{22} and S_{12} stand for the submatrices of the "within" covariance matrix pooled from k samples of size n each and $\underset{\sim}{x}_{1i}(p \times 1)$ and $\underset{\sim}{x}_{2i}(q \times 1)$ $(i = 1, \cdots, k)$ for the subvectors of the k sample mean vectors. Then setting $S_{1.2} = S_{11} - S_{12}S_{22}^{-1}S_{12}'$, and $B(p \times k)$ and $\beta(p \times k)$ for the matrices with respective column vectors $\underset{\sim}{x}_{1i} - S_{12}S_{22}^{-1}\underset{\sim}{x}_{2i}$ and $\underset{\sim}{\xi}_{1i} - \Sigma_{12}\Sigma_{22}^{-1}\underset{\sim}{\xi}_{2i}$ $(i = 1, \cdots, k)$, we have, with a confidence coefficient, say $1 - \alpha$, the following set of simultaneous confidence bounds (for all arbitrary nonnull $\underset{\sim}{a}'(1 \times p)$ and unit-length $\underset{\sim}{b}(k \times 1)$): $\underset{\sim}{a}'B\underset{\sim}{b} - [k(\underset{\sim}{a}'S_{1.2}\underset{\sim}{a})c_\alpha(p, k, nk - k)]^{1/2} \leqq \underset{\sim}{a}'\beta\underset{\sim}{b} \leqq \underset{\sim}{a}'B\underset{\sim}{b} + [k(\underset{\sim}{a}'S_{1.2}\underset{\sim}{a})c_\alpha(p, k, nk - k)]^{1/2}$, where $c_\alpha(p, k, nk - k)$ is the upper α-point of the distribution of the (central) largest determinantal root based on p, k and $nk - k$ D. F. Test for the associated hypothesis is also easily obtained.

26. The Efficiency of Tests, Wassily Hoeffding and Joan R. Rosenblatt, University of North Carolina.

The efficiency of a family of tests is defined. Let $\{X_n\}$ be a sequence of random variables such that for every n the vector $(X_1, \cdots, X_n)$ has cdf G_n in some class $\mathcal{C}_n$. Let $\mathcal{C}_{1n}$, $\mathcal{C}_{2n}$ be disjoint subsets of $\mathcal{C}_n$ such that we prefer one or the other of two alternatives A_1, A_2 according as $G_n \in \mathcal{C}_{in}$ $(i = 1, 2)$. Given α_1, α_2, we say that the problem $(\{\mathcal{C}_{1n}\}, \{\mathcal{C}_{2n}\}, \alpha_1, \alpha_2)$ is solved by a test (general nonsequential two-decision rule) ϕ_n such that $P(\phi_n \text{ selects } A_i \mid G_n) \geqq 1 - \alpha_i$ for all $G_n \in \mathcal{C}_{in}$ $(i = 1, 2)$. The index of efficiency of a family of tests $\mathfrak{I}$ for the problem $(\{\mathcal{C}_{1n}\}, \{\mathcal{C}_{2n}\}, \alpha_1, \alpha_2)$ is $N(\mathfrak{I}) = N(\mathfrak{I}, \{\mathcal{C}_{1n}\}, \{\mathcal{C}_{2n}\}, \alpha_1, \alpha_2)$, the least sample size with which the problem can be solved by a member of the family $\mathfrak{I}$. If $\mathfrak{I}_1$, $\mathfrak{I}_2$ are two families of tests, the efficiency of $\mathfrak{I}_2$ relative to $\mathfrak{I}_1$ is given by eff $(\mathfrak{I}_2/\mathfrak{I}_1) = N(\mathfrak{I}_1)/N(\mathfrak{I}_2)$. The determination of $N(\mathfrak{I})$ is closely related to finding a test which maximizes the minimum power. Let $\theta(G_n)$ be a real-valued function of G_n and suppose $\mathcal{C}_{1n} = \{G_n : \theta(G_n) \leqq \theta_1\}$, $\mathcal{C}_{2n} = \{G_n : \theta(G_n) \geqq \theta_2\}$, $\theta_1 < \theta_2$. Under suitable assumptions, we derive asymptotic expressions for $N(\mathfrak{I})$ as $\delta = \theta_2 - \theta_1$ tends to zero while α_1, α_2 remain fixed.

27. On a Decision Procedure to Select the Population with the Largest Mean (Preliminary Report), R. C. Bose and Jacques St. Pierre, University of North Carolina.

Consider $n + 1$ independent normal populations with unknown means $m_0 \geqq m_1 \geqq m_2 \cdots \geqq m_n$, respectively, and with known or unknown common variance σ^2. Suppose a sample of size N is available from each population, and a decision procedure is required to select the population with the largest mean, with the following properties. (a) Either a decision is made that the population from which the ith sample was drawn has the largest mean, or no decision is made. (b) The probability of making a wrong decision (if a decision is made) is less than a pre-assigned number α_0 (independent of the unknown means

$m_0, m_1, \cdots, m_n$. Subject to the requirements (a) and (b), the decision rule must control the chance of indecision. The case of three populations with known σ^2 is considered in detail, and the properties of a decision rule based on the auxiliary statistic $y = x_{(0)} - x_{(1)}$ are studied, where $x_{(0)} \geqq x_{(1)} \geqq x_{(2)}$ are the ordered sample means, the rule being to decide that $x_{(0)}$ comes from the population with the largest mean if $y > k$, and not to take a decision if $y \leqq k$. The general case when $n > 2$ and σ^2 is unknown is under consideration.

28. **Most Economical Multiple-Decision Rules,** William Jackson Hall, University of North Carolina.

Suppose x has an unknown distribution function F, belonging to one of m disjoint classes $\omega_1, \cdots, \omega_m$, and suppose $A_1, \cdots, A_m$ are corresponding alternative decisions. A decision rule D_N, based on a sample of size N, is said to be a "most economical multiple-decision rule (M.E. d.r.) relative to $(\alpha_1, \cdots, \alpha_m)$, $0 \leqq \alpha_i < 1$, for choosing among $A_1, \cdots, A_m$" if it satisfies (1) $\Pr(D_n \text{ chooses } A_i \mid F) \geqq \alpha_i$ for all $F \in \omega_i (i = 1, \cdots, m)$ and if N is the least integer n for which (1) can be satisfied. It is proved that to obtain M.E. d.r.'s one need only consider d.r.'s in the sequence $\{D_n^0\}$, $n = 0, 1, 2, \cdots$, where D_n^0 denotes a minimax solution w.r.t. a certain weight function for samples of fixed size n. If ω_i contains but one distribution F_i $(i = 1, \cdots, m)$, D_n^0 is of the form: (2) choose A_i if $a_iL_i \geqq a_jL_j$ $(j = 1, \cdots, m)$ where $L_1, \cdots, L_m$ are the likelihood functions of the sample corresponding to $F_1, \cdots, F_m$ and $a_1, \cdots, a_m$ are positive constants. In the general case, D_n^0 is of a similar form where now $F_1, \cdots, F_m$ are "average" distribution functions, averaged w.r.t. least favorable conditional distributions over $\omega_1, \cdots, \omega_m$ (if existent). Similar results are obtained for "M.E. d.r.'s relative to (β_{ij}), $0 < \beta_{ij} \leqq 1$," defined as above with (1) replaced by (1′) $\Pr(D_n \text{ chooses } A_i \mid F) \leqq \beta_{ij}$ for all $F \in \omega_j (i \neq j; j = 1, \cdots, m)$; and (2) is replaced by: (2′) choose A_i if $\Sigma_{k \neq i} b_k L_k \leqq \Sigma_{k \neq j} b_k L_k$ $(j = 1, \cdots, m)$, for some positive constants $b_1, \cdots, b_m$. Other properties of the d.r.'s are derived and various extensions and examples given.

NEWS AND NOTICES

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

Personal Items

Archie Blake is now employed as an Advisory Engineer in the Systems Analysis of Westinghouse Electric Corporation, Baltimore, Maryland.

E. L. Cox has left Operations Research Group, Case Institute of Technology, to take a position with Chemical Corps Biological Laboratories, Frederick, Maryland.

Harold Davis has transferred from Headquarters, United States Air Force to The Operations Analysis Office, Hq. Far East Air Forces.

Professor Hilda Geiringer is on leave of absence from Wheaton College in order to complete and prepare for publication on behalf of Harvard University some of the post-humous work of Richard von Mises.

Dr. S. G. Ghurye has accepted the position of Reader in Statistics, Department of Mathematics and Statistics, University of Lucknow, Lucknow, U.P., India.