# ABSTRACTS OF PAPERS

*(Abstracts of papers presented at the Eastern Regional Meeting of the Institute, April 21–22, 1961. Additional abstracts appeared in the March, 1961 issue.)*

### 6. **Tables of Minimum Functions for Generating Galois Fields** GF ($p^n$). J. D. Alanen, Case Institute of Technology. (Introduced by I. M. Chakravarti.)

A polynomial $f(x)$ of degree $n$ irreducible in the field GF$(p)$ where $p$ is a prime number, is called a minimum function, if a root $\omega$ of the equation $f(x) = 0$, serves as a primitive element of GF$(p^n)$, that is, $\omega^0 = 1, \omega, \omega^2, \cdots, \omega^{p^n-2}$ are the $p^n - 1$ non-zero elements of GF$(p^n)$. It is known that for the GF$(p^n)$, there are $\varphi(p^n - 1)/n$ minimum functions, where $\varphi$ is the Euler function, $p$ a prime, and $n$ an integer. Minimum functions were very successfully used in the past in constructing sets of mutually orthogonal Latin squares, balanced incomplete block designs, confounded and fractional factorial designs. Recently these have found a new application in the construction of error-correcting codes. While searching for a minimum function of GF$(13^3)$, we noticed a lack of comprehensive tables of minimum functions in the published literature. A program has been written and all minimum functions generated for a fairly comprehensive set of values of $p$ and $n$.

### 7. **Testing to Establish a High Degree of Safety or Reliability.** F. J. Anscombe, Princeton University and Bell Telephone Laboratories. (Invited paper)

We are concerned with the possibility of establishing the safety of a weapon or the reliability of a component or device by testing a large number of specimens under some standard operating conditions and demonstrating that the proportion of failures, $p$, is very small. When possible, a fully economic treatment of such a problem is to be desired, in which the expected loss from wrong decisions is assessed and balanced against the cost of testing. But sometimes a noneconomic type of requirement must be considered, such as: (A) *The device will be accepted for service only if the test results permit an assertion with* 99% *confidence that* $p < 1/2000$. Most statisticians will interpret such a requirement, by analogy with the definition of a confidence coefficient, as follows: (B) *The acceptance rule must be such that, for all values of* $p > 1/2000$, *the least upper bound to the chance of acceptance* = 1%. But it is suggested that the following weaker interpretation is sufficiently stringent and more appropriate: (C) *The device will be accepted only if the test results justify fair betting odds of* 99:1 *that* $p < 1/2000$. These odds of 99:1 are to be the final betting odds of an observer who before the trial begins is open-minded and unprejudiced. A suitable prior probability distribution for $p$, relating to such an observer, is proposed, and an acceptance boundary for sequential testing is obtained. In order to complete the specification of a sequential rule of procedure, it is necessary to add a second boundary, for abandoning the trial when the cost of continuing seems excessive. Possible ways of doing this are discussed, and a boundary based on a detailed economic analysis is developed. Because the acceptance requirement (C) is probabilistic (Bayesian), the validity of the acceptance boundary is not affected by the introduction of a boundary for abandoning the trial.

### 8. **Extreme Values in Gaussian Sequences.** Simeon Berman, Columbia University.

The first theorem extends a result of Rényi (1958). Let $(\Omega, \mathcal{A}, P)$ be a probability space and $\{X_n\}$ a sequence of independent and identically distributed random variables defined on the space. For each $n$, let $Z_n = \max(X_1, \cdots, X_n)$; suppose that $Z_n$ has a limiting distribution. If $Q$ is another probability measure on $\mathcal{A}$ which is absolutely continuous with respect to $P$, then $Z_n$ has the same limiting distribution under $Q$ as under $P$. An application to a nonstationary Gaussian sequence is given. The second theorem concerns a stationary Gaussian sequence $\{X_n\}$. Conditions are given on the covariance sequence which are sufficient for the convergence in probability of max $(X_1, \cdots, X_n) - (2 \log n)^{\frac{1}{2}}$ to zero. The conditions are satisfied by the stationary Gaussian Markov process.

9a. **On the Foundations of Statistical Inference II** (Preliminary report). ALLAN BIRNBAUM, New York University. (By title) (Abstract printed in the December, 1960 issue, p. 1216.)

9b. **Some Theory and Techniques for Robust Estimation** (Preliminary report). ALLAN BIRNBAUM, New York University.

Let $f(x, \theta, \gamma)$ be the density function of sample point $x$, depending on real parameter $\theta$ and any (nuisance) parameter $\gamma$ of specified ranges. For any given estimator $\theta^* = \theta^*(x)$ of $\theta$, let $r(\theta, \gamma)$ denote the mean-squared-error (m.s.e.) (or alternatively the variance when it is useful to restrict consideration to unbiased estimators). Admissibility of $\theta^*$ (possibly in a restricted class of estimators) is defined as usual. $\theta^*$ is called robust (over the specified range of $\gamma$) if for each $\gamma'$ there exists a corresponding estimator, admissible when $\gamma = \gamma'$ is known, with m.s.e. $\underline{r}(\theta, \gamma')$, such that $\underline{r}(\theta, \gamma')/r(\theta, \gamma')$ is near unity for all $\theta$. The degree of attainable robustness depends on the form of $f(x, \theta, \gamma)$, and its determination is of interest along with the characterization and construction of robust estimators. It is appropriate here to call an estimator *admissibly robust* if it is simply admissible. (Problems of robust confidence limit estimation and testing, and other formulations of point estimation, can be discussed similarly, taking $r$ as a possibly vector-valued risk function representing relevant error-probabilities.) Admissibly robust invariant estimators of a location parameter $\theta$ are characterized (as having a modified Pitman structure), and problems of computing $r$'s and attainable robustness are discussed. Under restriction to unbiased estimators linear in ordered observations, admissibly robust estimators are characterized, and fairly tractable theoretical and computational methods are illustrated.

10. **Bayes Rules for the Problem of Choosing the Largest Mean** (Preliminary report). RICHARD P. BLAND, University of North Carolina, AND DAVID B. DUNCAN, Johns Hopkins University.

Random samples of equal size $m$ are drawn independently from $n$ normal populations with the same variance $\sigma^2$. A Bayes rule is derived for choosing a superior subset of population means so as to contain the largest. The loss is the sum of losses for each of the means involved, these being: zero for the largest mean if chosen, $k_1\delta$ if not; $k_0\delta$ for an inferior mean if chosen, and zero if not; where $\delta$ is the difference of the mean concerned from the second largest or largest mean respectively and $k_1 > k_0 > 0$. The population means have independent identical normal prior densities with variance $\gamma^2\sigma^2/m$. Rejection from the superior subset depends on the differences of the sample mean concerned from the largest sample mean, the second largest and so on, the dependencies rapidly diminishing. The differences necessary for rejection depend on the loss-ratio $k = k_1/k_0$ and the prior variance ratio $\gamma^2$ and, notably, diminish slightly with $n$. These are tabled for $n = 3$, $\sigma^2$ known. A conservative-near-Bayes rule with rejection depending only on the difference from the largest sample mean is also presented with tables for all $n$ and $\sigma^2$ unknown.

11. **Iterated Steepest Ascent on Ellipsoidal Contours.** R. J. BEUHLER, B. V. SHAH, AND O. KEMPTHORNE, Iowa State University.

An arbitrary quadratic response in $n$ variables $y = c_0 + \sum b_i x_i + \sum \sum a_{ij} x_i x_j$ having a unique minimum or maximum may without loss of generality be replaced by $y = \sum \alpha_i x_i^2$ in which $\alpha_1 \geqq \alpha_2 \geqq \cdots \geqq \alpha_n > 0$; $\alpha_i^{-\frac{1}{2}}$ being equal to the $i$th shortest axis of one of the family of similar ellipsoidal contours centered at the origin. Starting at an arbitrary point $P_1$ proceed rectilinearly in the direction of the gradient at $P_1$ until a minimum of $y$ is found at $P_2$. Iterate to obtain a sequence $P_1, P_2, P_3, \cdots$ of points having responses $y_1, y_2, y_3, \cdots$ converging to zero. The relative success of the $m$th step can be measured by the smallness of $\rho_m = y_{m+1}/y_m$. For $n = 2$ it is shown that: (i) $\rho_1 = \rho_2 = \rho_3 = \cdots$, so that $y_m$ converges in a geometric progression to zero. (ii) The "least favorable" starting points on any contour occur where tangents are at 45 degrees to the axes. (iii) At these points one has $\rho_{\max} = (r^2 - 1)^2/(r^2 + 1)^2$ where $r^2 = \alpha_1/\alpha_2$. For arbitrary $n$ it is shown that: (i) $\rho_1 \leqq \rho_2 \leqq \rho_3 \leqq \cdots$ so that the rate of convergence never improves. (ii) This rate however is never worse than that associated with the "least favorable" starting point, which occurs in the two-dimensional subspace containing the longest and shortest axes. (iii) Starting at this point $\rho_i = \rho_{\max} = (r^2 - 1)^2/(r^2 + 1)^2$ where $r^2 = \alpha_1/\alpha_n$.

### 12. Steepest Ascent Partan on Ellipsoidal Contours. R. J. Buehler, B. V. Shah, and O. Kempthorne, Iowa State University.

In the method of parallel tangents (PARTAN) the ambiguous directions at $P_1$ and $P_2$ (and in $n$ dimensions at $P_4, P_6, \cdots, P_{2n-4}$) can be resolved by proceeding normal to the corresponding tangent planes (steepest ascent). The resulting procedure is invariant under translations ($x' = x + c$) and rotations ($x' = Ux$, $U$ orthogonal) but not under changes of scale ($x' = Ax$, $A$ diagonal). Thus a general quadratic response may be represented in a canonical form $y = \sum_1^n \alpha_j x_j^2 (\alpha_j > 0)$. Let $P_1$ be parametrized by squared direction cosines $l_i^2 = a_i x_i^2/\sum \alpha_j x_j^2$, using the $x_i$ values at $P_1$. Define $S_j = \sum \alpha_i^j l_i^2 (j = 0, 1, \cdots)$ and $\Delta_m = m \times m$ determinant with $(\Delta_m)_{ij} = S_{i+j-2} (m = 2, 3, \cdots)$. It is shown that the response at $P_{2m-2}$ is $y_{2m-2} = y_1 \Delta_m/\Delta_m^{11}$ $(m = 2, 3, \cdots)$, where $\Delta_m^{11}$ is the cofactor of $(\Delta_m)_{11}$. For any given $n$, $m$ and $\alpha$'s the "poorest" (least fortunate) $P_1$ has been determined explicitly by locating all maxima of $\Delta_m/\Delta_m^{11}$. Typical results are (i) In $n = 3$ dimensions if any two $\alpha$'s are equal, $y_4 = \Delta_3 = 0$ for any $P_1$, and (ii) In $n$ dimensions $y_4/y_1 \leqq (r^2 - 1)^4/(r^4 + 6r^2 + 1)^2$ for all $P_1$, $\alpha_i$, where $r^2 = \alpha_{\max}/\alpha_{\min}$, and the bound improves if $\alpha_i \neq \frac{1}{2}(\alpha_{\min} + \alpha_{\max})$ for every $i$.

### 13. Asymptotic Relative Efficiency of Mood's and Massey's Test Against Some Parametric Alternatives. I. M. Chakravarti, F. C. Leone and J. D. Alanen, Case Institute of Technology.

In an earlier paper the expressions for the exact power of Mood's and Massey's tests have been derived. The numerical values of the power for several sample sizes and selected values of the parameters were computed. The asymptotic relative efficiency of Mood's test against normal and rectangular alternatives was derived by Mood and Andrews. In the present paper, the asymptotic relative efficiency of Mood's test against the likelihood ratio for the change in location of exponential distribution, is derived. Further, this is carried out for all three alternatives for Massey's test. The asymptotic powers are compared with the exact powers to find out how large a sample size is needed before one could use the expressions for the asymptotic power.

### 14. Several-Sided Kolmogoroff-Smirnoff Procedures. Herbert T. David, Iowa State University. (By title)

Let $D_n$ be Kolmogoroff's statistic and let $X^1$ and $X^n$ be the extreme order statistics. Consider testing $H_o : F = F_o$ by the non-parametric test based on $\max(D_n, k_1 \cdot F_o(X^n), k_2 \cdot (1 - F_o(X')))$, leading to a "four-sided" acceptance region for the "stairway" portion of the sample CDF. A pertinent distributional fact is as follows. Let the statistic $S_F(x_1, \cdots, x_n)$, $F$ continuous, have *structure* $(d)$. Let $S_{a,b}(x_1, \cdots, x_n)$ be the statistic

$S_F$ corresponding to the uniform distribution over $[a, 1 - b]$, $a$ and $b \geqq 0$. Let $\beta$ be the order in $n$ of the sup, for $x_i$ in $[tn^{-1}, 1 - un^{-1}]$, of

$$|S_{o,o}(x_1, \cdots, x_n) - S_{tn^{-1},un^{-1}}(x_1, \cdots, x_n)|.$$

Let $\Pr\{S_F < s \mid F\} = \phi_n(s)$, with $\lim \phi_n(sn^{-\alpha}) = \phi(s)$ continuous. Then, if $\alpha + \beta < 0$, $F_o(X^n)$, $F_o(X^1)$, and $D_n$ are asymptotically independent when the population CDF is $F_o$. Error bounds are computed for assuming independence for finite $n$. These bounds show that independence sets in quite early. For example, for $n = 38$, $\Pr\{D_{38} < .2347 \mid F_o\} = \Pr\{F_o(X^{38}) < .99932 \mid F_o\} = \sqrt{.95}$, whereas

$$.948 \leqq \Pr\{D_{38} < .2347, F_o(X^{38}) < .99932 \mid F_o\} \leqq .952.$$

### 15. A Generalization of a Simple Test Function for Guarantee Time Associated with the Exponential Failure Law. Satya D. Dubey, Proctor & Gamble Co.

In a previous paper entitled, "A Simple Test Function for Guarantee Time Associated with the Exponential Failure Law" the author has shown that a test function based on the first and the $r(\leqq n$, the sample size)th observations for testing the hypothesis on the guarantee time has several desirable properties (see the abstract of this paper in the same issue of this journal). This has prompted the author to consider a generalized simple test function based on any two sample observations for the same hypothesis. For the use of this simple test function upper 1, 5, and 10 per cent critical values are tabulated up to the sample size 10. Several moment recurrence formulas are established which reveal interesting relationships between its $k$th and $(k - 1)$th order moments. Its power functions are derived. The results applicable to the special cases of the generalized simple test function are mentioned and the useful properties of this test function in some special cases are pointed out.

### 16. A Simple Test Function for Guarantee Time Associated with the Exponential Failure Law (Preliminary report). Satya D. Dubey, Proctor and Gamble Co. (By title)

Let the probability density function (p.d.f.) of a time to failure random variable, $T$ be represented by

$$f_T(t) = \begin{cases} \theta^{-1} e^{-\theta^{-1}(t-G)}, & t > G \text{ and } \theta > 0 \\ 0 \text{ otherwise}. \end{cases}$$

Consider the $H: G = G_0$ against the $A: G \neq G_0$ and assume $\theta$ to be unknown. On the basis of the first $r(\leqq n$, the sample size) ordered observations we have derived the likelihood ratio test of the above hypothesis. Under $H$ this test has an $F$ distribution with 2 and $2r - 2$ degrees of freedom. It is completely *unbiased*, has a monotone power and is a uniformly most powerful unbiased test. The results cover the work of Paulson (*Ann. Math. Stat.*, Vol. 12 (1941) pp. 301–306). For the same hypothesis an alternative test function based on the first and $r$th sample observations is suggested. This test function is completely unbiased and has a monotone power. Furthermore, its power for $G < G_0$, which is of interest in life testing situations, is exactly the same as of the likelihood ratio test and is independent of $r$. This immediately suggests taking $r = 2$ and using only the first two ordered observations of the sample. The results go beyond the work of Carlson (*Skandinavisk Aktuarietidskrift*, Haft 1–2 (1958), pp. 47–54).

### 17. Asymptotically Most Efficient Single Observation Estimator of Expected Life for Exponential Failure Law (Preliminary report). Satya D. Dubey, Proctor and Gamble Co. (By title)

For the exponential failure law with known location parameter, the minimum variance single observation unbiased estimator of the scale parameter is investigated. It is found that if the $r(\leqq n$, the sample size)th observation in order of increasing time is the single observation on which this estimate is based and if we write $r = n\delta_n$, then $\lim_n \delta_n = \delta_0$ which is the positive real root of the equation, $\log (1 - \delta) + 2\delta = 0$. It is about 66 per cent efficient in comparison with the minimum variance unbiased estimator based on all the $n$ observations in the sample. The sample median has only 48 per cent efficiency. The smallest sample observation is $100/n$ per cent efficient and the largest sample observation has asymptotic efficiency of $(600/\pi^2)[(\log^2 n)/n]$. The percentile estimator approach also yields the same results; however, the present investigation has yielded some interesting side results.

## 18. Central Limit Theorem for Sums Over Sets of Random Variables. Friedhelm Eicker, University of North Carolina.

Many estimates are given as linear combinations $\zeta_n = \sum_{k=1}^{n} a_{nk}\epsilon_k$ of independent errors $\epsilon_k$ whose distributions are unknown and non-identical. Some of the asymptotic properties of these estimates are governed by the central limit theorem (CLT). The question is raised as to what restrictions must be placed upon the set $F$ of random variables (r.v.) from which the $\epsilon$'s are taken in any arbitrary order, and also upon the matrix of constants in order that the CLT holds "over $F$." A sequence of functions like $\{\zeta_n\}$ is said to *converge over (or: on)F* (in any sense) if this is true for any possible choice of a sequence $\epsilon_1, \epsilon_2, \cdots$ with elements $\epsilon_i$ in $F$. The following *theorem* is derived: Let $F$ be a set of r.v. with $E\epsilon = 0$, $0 < E\epsilon^2 < \infty$. Then it is necessary and sufficient for the convergence of $\zeta_n$ on $F$ to the normal law $N(0, \operatorname{var} \zeta_n)$ for $n \to \infty$ that simultaneously holds:

(I) $\max_{k=1,\cdots,n} a_{nk}^2 / \sum_r a_{nr}^2 \to 0$ for $n \to \infty$.

(II) There exists a bounded function $g(c)$ for $c \leqq 0$ with $\lim_{c\to\infty} g(c) = 0$ such that for each r.v. in $F$ $\int_{|\epsilon| \geqq c} \epsilon^2 \, dG(\epsilon) < g(c)$ holds where $G(\epsilon)$ is the distribution function of $\epsilon$.

(III) $\operatorname{var} \epsilon > m > 0$ for all r.v. in $F$.

Application can be made to the least squares estimates in linear regression and in autoregressive schemes of time series. Of particular importance is the bearing upon some non-parametric problems and on some nonstationary time series.

## 19. Power of a Non-Parametric Test of Independence. Regina C. Elandt, Case Institute of Technology. (Introduced by N. L. Johnson.)

In a previous paper [R. C. Elandt, *Zastosowania Matematyki*, Vol. 3 (1956), pp. 8–45] the author proposed a test of independence based on the number of pairs, in the two-dimensional sample $(x_1, y_1), \cdots, (x_{2n}, y_{2n})$, for which $(x_i - Me_x)(y_i - Me_y) > 0$, where $Me_x$, $Me_y$ are the sample medians of $x$ and $y$ respectively. In the present paper the power function of this test is investigated. If $Me_x$, $Me_y$ are replaced by the (usually unknown) population medians, it is possible to evaluate the power function exactly; this is used as a basis for a heuristic approximation to the power of the original test. This approximation is shown to give good results when the null hypothesis (of independence) is true. The asymptotic relative efficiency of the test is also evaluated.

## 20. Location and Scale Parameters in Exponential Families of Distributions (Preliminary report). T. S. Ferguson, University of California, Los Angeles.

If a location parameter, $\theta$, is the parameter of a one-parameter exponential family of distributions, then the distribution for fixed $\theta$ is either (1) the distribution of $a^{-1} \log X$, where $X$ has a gamma distribution and $a \neq o$, or (2) (corresponding to the case $a = o$) a normal distribution. This result is extended to the case where the location parameter is a parameter in a $k$-parameter exponential family of distributions, with the aid of the (probably superfluous) assumption that the density has $k + 2$ derivatives. Similar results are derived for scale parameters. If the parameters of a two parameter exponential family of distributions may be taken to be location and scale parameters, then the distribution must be normal. The family of distributions of the first mentioned result is thus seen to be a main class of distributions to which Basu's theorem (on statistics independent of a complete sufficient statistic) applies. Furthermore, this family of distributions provides a natural setting in which to prove certain characterization theorems which have been proved separately for the normal and gamma distributions.

**21. A Double-Ended Queuing Process.** SAMUEL M. GIVEEN, Northeastern University. (Introduced by Lionel Weiss.)

A queuing model is considered in which the possible states are designated by the positive and negative integers and zero. Two types of units, described as "positive" and "negative", enter the system. The arrival of a unit of one type constitutes an addition to the queue if units of the same type are waiting, and a service completion if units of the other type are waiting. Thus, if the system is in any state $n$, the arrival of a "positive" unit will always shift it to state $n + 1$ and the arrival of a "negative" unit will always shift it to state $n - 1$. Arrivals of the two types of units are governed by Poisson probability laws with time-dependent parameters $\lambda(t)$ and $\mu(t)$. An explicit transient solution is obtained for the probability that the system will be in state $n$ at time $t$, but a steady-state solution does not in general exist. In the special case in which $\lambda$ and $\mu$ are constant, this solution is the well-known probability of a difference of $n$ between two Poisson variables. In the general case conditions are determined under which the mean and the variance will remain finite as $t$ becomes infinite.

**22. The Use of Sample Ranges in Setting Exact Confidence Bounds for the Standard Deviation of a Rectangular Population.** H. LEON HARTER, Wright-Patterson Air Force Base.

A discussion is given of point estimates and interval estimates of the population standard deviation $\sigma$, based on the sample range and quasi-ranges. In the case of a rectangular population, the efficient point estimate and the most effective interval estimates are those based on the sample range, so it is not necessary to consider estimates based on sample quasi-ranges. The coefficients of the sample range $w$ in the exact confidence bounds for the population standard deviation $\sigma$ are found by taking the reciprocals of percentage points of the (standardized) range $W = w/\sigma$. The following tables for the rectangular population are included: (1) A six-decimal-place table of the percentage points of the range corresponding to cumulative probabilities $P = 0.0001, 0.0005, 0.001, 0.005, 0.01, 0.025, 0.05, 0.1$ (0.1) 0.9, 0.95, 0.975, 0.99, 0.995, 0.999, 0.9995, 0.9999 for sample sizes $n = 2$ (1) 20 (2) 40 (10) 100; and (2) a table, to seven significant figures or six decimal places, whichever is less accurate, of the coefficients of the sample range $w$ in the exact lower confidence bounds for $\sigma$ for the above values of $P$ and $n$.

**23. The Moments of the Non-Central $t$-Distribution.** D. HOGBEN, R. S. PINKHAM AND M. B. WILK, Rutgers University.

If $X$ is normally distributed with mean $\delta$ and variance 1 independently of $Y^2$ distributed as chi-squared with $f$ degrees of freedom, then the random variable $t = X(f)^{\frac{1}{2}}/Y$ is distributed as non-central $t$. Let $p(x, y^2)$ be the joint density of $X$ and $Y^2$. Then, explicit expressions for the raw moments of $t$ can be obtained readily from an identity which results from continued differentiation, with respect to $\delta$, of $\iint p(x, y^2)\, dy^2\, dx = K(\delta, f)$. A table of numerical values relevant to the first four central moments and cumulants is given. Application of the table is illustrated in finding the approximate values of the expectation of certain functions of the non-central $t$.

### 24. Some Notes on the Investigation of Heterogeneity in Interactions. N. L. Johnson, Case Institute of Technology.

Consider a cross-classification with model $y_{ij} = A + R_i + C_j + z_{ij}$ $(i = 1, \cdots r; j = 1, \cdots c)$, $\sum R_i = \sum C_j = 0$, and $z_{ij}$'s independent $N(0, \sigma_j)$ variables. To investigate possible differences among the $\sigma_j$'s, criteria based on the statistics

$$S_j = \sum_{i=1}^{r} (y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot})^2$$

may be used.

$$(\bar{y}_{i\cdot} = c^{-1} \sum_{j=1}^{c} y_{ij}\, ,\ \bar{y}_{\cdot j} = r^{-1} \sum_{i=1}^{r} y_{ij}\, ,\ \bar{y}_{\cdot\cdot} = (rc)^{-1} \sum_{i=1}^{r} \sum_{j=1}^{c} y_{ij}\, .)$$

The distribution of the ratio $S_j/S_{j'}$ is shown to be related to that of $g = (\sum_{l=1}^{\nu} w_{lj}^2)/(\sum_{l=1}^{\nu} w_{lj'}^2)$ where the $w$'s are multi-normal units variables with correlation $\rho$ between $w_{lj}$ and $w_{lj'}$, all other correlations being zero. The distribution of $g$ is shown to be a mixture of $F$ distributions with $\nu + 2k$, $\nu + 2k$ degrees of freedom, in proportions given by terms in the expansion of the negative binomial $\left(\frac{1}{1-\rho^2} - \frac{\rho^2}{1-\rho^2}\right)^{-\frac{1}{2}\nu}$. Some approximations to the distribution of $S_j/S_{j'}$ are discussed and conjectural extensions to the joint distribution of $S_1, S_2, \cdots, S_c$ are mentioned.

### 25. On the Limiting Distribution of $-2 \log \lambda$ in the Non-regular Case. Donald A. Jones, University of Michigan.

Consider a family, say $F$, of pdfs (Lebesgue) on $E'$ which satisfies a set of regularity conditions w.r.t. a real-valued parameter. A new family of pdfs, say $G$, can be constructed from $F$ by doubly truncating each member of $F$ at unknown points, $a < b$. Let $\mathbf{x} = (x\ , x_2, \cdots, x_m)$ be a random variable satisfying the following conditions:

(i) the real random variables $x_1, x_2, \cdots, x_m$ are independent and,

(ii) the distribution of $x_j$ has a pdf in $G$, $j = 1, 2, \cdots, m$.

Observe that the pdf of $\mathbf{x}$ belongs to a family indexed by a $3m$-dim. parameter, say $\theta$ with range $\Omega$, where $m$ of the dimensions are from $F$ and $2m$ were introduced by truncation. Let $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$, be $n$ random observations of $\mathbf{x}$ and denote their pdf, as a function of $\theta$, by $F_n$. The limiting distribution of $-2 \log \lambda (n \to \infty)$, where

$$\lambda = \text{l.u.b. } [F_n : \theta \varepsilon \omega]/\text{l.u.b. } [F_n : \theta \varepsilon \Omega]$$

and $\omega \subset \Omega$, is shown to be, for some special classes of $\omega$ and assuming that $\theta \varepsilon \omega$ obtains, a chi-square distribution with $2b + d$ d.f. where $b =$ (dim. $\Omega$ − dim. $\omega$) w.r.t. the "truncation parameters" and $d =$ (dim. $\Omega$ − dim. $\omega$) w.r.t. the "regular parameters". The motivation is hypothesis testing by the likelihood ratio test and the methods of proof are standard limit theorems.

### 26. Use of Some a priori Knowledge in the Estimation of Means from Double Samples. S. K. Katti, Florida State University.

When there is no *a priori* knowledge available regarding the values of the mean $\mu$, the overall sample mean $\bar{x}$ has many desirable properties as an estimate of $\mu$. The problem considered here is one in which the statistician has a guessed estimate (guestimate) on $\mu$ either due to his past experience or due to his acquaintance with the behavior of the system. It is found that if $\mu_o$ is the guestimate, the true value is close to $\mu_o$, the variance $\sigma^2$ is known and if it is possible to take samples of size $n_1$, and $n_2$ in a succession, an estimate with an expected mean square smaller than that of the overall mean $\bar{x}$ is obtained by using the mean of the first sample only, ignoring the second if this mean lies in the interval $(\mu_o - \sigma/(2n_1 + n_2)^{\frac{1}{2}}, \mu_o + \sigma/(2n_1 + n_2)^{\frac{1}{2}})$ and by using the overall mean if it does not. If the observations are independent, have a common normal distribution and $\mu$ close to $\mu_o$, this method gives an estimate with an expected mean square 3.2% smaller than that of $\bar{x}$ when $n_1/n_2 = 0.33$ and the sample size used to obtain this estimate is 74.9% of $(n_1 + n_2)$, i.e., there is a saving of 25.1%.

### 27. **On the Expected Value and Variance of a Ratio Estimate.** J. C. KOOP, North Carolina State College. (By title)

From the data of sample surveys various kinds of ratio estimates and their standard errors are computed, often without questioning the validity of the underlying mathematical procedures used in deriving the formulas. The limitations of the classical technique of deriving the expected value and the variance of ratio estimates of finite populations by a series expansion are briefly discussed in this paper. By a new device of expanding the ratio or its powers as a truly convergent series, and then determining the expected values term by term, expressions differing from the classical ones only by a constant multiplying factor are obtained. This constant multiplier is an element of the inequality which shows what minimum value it shall assume to insure that the series expansion is valid.

### 28. **On the Higher Moments of Linear Estimates Based on Multistage Samples from a Finite Population.** J. C. KOOP, North Carolina State College.

The theorem for the derivation of higher moments of linear estimates based on multistage samples from a finite population is derived. An $r$th order moment of a random variable about its mean or expected value is shown to be equal to the expected value (or the first moment) of the conditional moment of the same order, plus the $r$th moment of the conditional expected value, plus a series of product-moments of conditional expected values of various orders (of course not exceeding $r$) each of them with appropriate binomial coefficients. For the case when $r = 2$, we obtain the theorem given in the well-known textbook of Hansen, Hurwitz and Madow. In the context of multistage (or multiphase) sampling the net result of all moments of an estimate about its expected value of order lower than $r$ can be interpreted as increase due to extra steps in sampling beyond the first step (stage or phase). The theorem is applied to obtain the third and fourth moments of linear estimates based on multistage samples from a finite population with the following probability systems which are very general:

(i) first stage units are selected with equal probabilities and without replacement but the probabilities for the selection of units in the subsequent stages are left undefined.

(ii) first stage units are selected with unequal probabilities and with replacement but the probabilities for the subsequent stages are left undefined as above.

When the probabilities beyond the first stage and the structure of the sample (i.e., sample design relevant to those steps in question) are defined, then the conditional moments can all be evaluated in terms of the moments of the underlying variates. With these moments the expressions for the classical coefficients of kurtosis and skewness of an estimate may be obtained. For the above cases (and indeed for all multistage estimates) the approach to approximate normal form is retarded because of the presence of the product moments.

### 29. **On Simultaneous Tests in Nested Designs** (Preliminary report). P. R. KRISHNAIAH, Remington Rand UNIVAC.

Consider the "fixed effects" model $x_{ijkm} = \alpha_{ijk} + e_{ijkm}$ $(i = 1, 2, \cdots, p; j = 1, 2, \cdots, q; k = 1, 2, \cdots, r; m = 1, 2, \cdots, s)$ where the errors $e_{ijkm}$ are distributed independently and identically with zero means and variance $\sigma_i^2$ ; $\alpha_{ijk}$ is the effect of $k$th level of $C$ within $j$th level of $B$ within $i$th level of $A$ and $\sum \sum \sum \alpha_{ijk} = 0$. Now consider the following hypotheses

$$H_{ij} : \alpha_{ijk} = \bar{\alpha}_{ij\cdot} \quad \text{for all } k$$

$$H_i : \bar{\alpha}_{ij\cdot} = \bar{\alpha}_{i\cdot\cdot} \quad \text{for all } j$$

$$H : \bar{\alpha}_{i\cdot\cdot} = \bar{\alpha}_{\cdots} \quad \text{for all } i$$

where $\bar{\alpha}_{ij\cdot} = r^{-1} \sum \alpha_{ijk}$ ; $\bar{\alpha}_{i\cdot\cdot} = q^{-1} \sum \bar{\alpha}_{ij\cdot}$ ; $\bar{\alpha}_{\cdots} = p^{-1} \sum \bar{\alpha}_{i\cdot\cdot}$ . The present paper discusses testing $pq$ hypotheses of the form $H_{ij}$ , $p$ hypotheses of the form $H_i$ and the hypothesis $H$ simultaneously by using (1) the Simultaneous ANOVA Test and (2) the Joint (Overall $F$) Test. The lengths of the simultaneous confidence bounds associated with the above tests are compared. These results are generalized to multivariate situations and to higher order hierarchal classification.

## 30. A Multivariate Analogue of One-Sided Test (Preliminary report). AKIO KUDO, University of Michigan Medical School.

Consider a sample of size $n$ from a $k$-variate normal population with unknown means $m_i$ ($i = 1, \cdots, k$) and a known variance matrix $\Lambda$. The likelihood ratio test of the null hypothesis $H_0$ ; $m_i = 0$ ($i = 1, \cdots, k$) against the alternative hypothesis $H_1$ ; $m_i \geqq 0$ ($i = 1, \cdots, k$) where inequality is strict for at least one of the $k$ will be considered. The test is based on the statistic $K^2 = n[\bar{x}'\Lambda^{-1}\bar{x} - \min (\bar{x} - m)'\Lambda^{-1}(\bar{x} - m)]$, where $\bar{x}$ and $m$ are the sample and the population mean vectors and the minimum is over the region $m_i \geqq 0$ ($i = 1, \cdots, k$). The computation of $K^2$ will be discussed, and we shall prove

$$\Pr(K^2 \geqq K_0^2) \leqq \sum_{\emptyset \subsetneq M \subseteqq K} \Pr(\chi^2_{n(M)} \geqq K_0^2) P_M (1 - P_{M'}),$$

where the summation runs over all non-empty sets $M$ included in or equal to the set $K = \{1, 2, \cdots, k\}$, $n(M)$ is the number of elements in $M$, $P_M = \Pr(x_i \geqq 0, i \in M)$, $P_{M'} = \Pr(x_i \geqq 0, i \notin M \mid x_i = 0, i \in M)$. $P_M$ and $P_{M'}$ are probabilities under the null hypothesis, and the latter is a conditional probability equal to zero when $M = K$.

## 31. Power of Mood's and Massey's Test Against Exponential and Rectangular Alternatives. F. C. LEONE, I. M. CHAKRAVARTI AND J. D. ALANEN, Case Institute of Technology.

When sample observations are originally recorded in order of their magnitude, Mood's test, based on the median of the combined samples, and Massey's extension of Mood's test based on fractiles, have much to commend themselves as quick tests. In Mood's test one need record observations only up to the median of the combined samples, and in Massey's test, up to the highest fractile included in the test. It is known (Mood, (1954) and Andrews, (1954)) that the asymptotic relative efficiency of Mood's test against normal and rectangular alternatives is low. However, it was felt that there was enough gain in Massey's extension to justify this investigation. The objects of the present investigation are:

(i) to derive the exact power functions of Mood's and Massey's tests for two samples against parametric alternatives of exponential and rectangular populations;

(ii) to tabulate them for comparable sample sizes in order to get an idea about their respective performances and also to evaluate if there is any resultant gain in the use of Massey's test (which uses more than one fractile and hence is more elaborate) over Mood's tests.

## 32. On some Properties of Compositions of an Integer and their Application to Probability Theory. T. V. NARAYANA AND S. G. MOHANTY, University of Alberta.

A partial order called "domination" has been defined on the $r$-compositions of $n$ and the $r$-compositions of $n$ form a distributive lattice ($1 \leqq r \leqq n$), (Narayana and Fulton, *Canad. Math. Bull.*, Vol. 1, No. 3). Using an obvious partial order, the distributive lattice formed by the simple symmetric sampling plans of size $2n$ is shown to be isomorphic to the lattice formed by the $n$-compositions of $3n$ dominated by the $n$-compositions $(3, 3, \cdots 3)$ or $3n$. A natural one-to-one correspondence between the simple symmetric sampling plans of size $2n$ and simple sampling plans of size $n$ induces the lattice on the simple sampling plans of size $n$ and as a consequence the number of such plans is found to be $n^{-1}\binom{3n}{n-1}$. Since it is easily seen that this distributive lattice is isomorphic to a distributive lattice formed by the minimal lattice paths in a classical ballot problem ([1] Grossman, *Scripta Mathematica*, Vol. 15, No. 1–2; [2] Dvoretzky and Motzkin, *Duke Math. J.*, 1947), the solution to the problem is rederived and a further generalization of it is obtained. The authors hope to extend the application of this unified approach to various other problems in probability theory.

## 33. Fractional Factorial $2^m$ and $3^n$ Designs with and without Blocks, Preserving the Main Effects and the Two-Factor Interactions. M. S. PATEL, University of North Carolina and Research Triangle Institute.

The object of this paper is to construct fractional factorial designs which have a number of treatment combinations just sufficient to estimate the main effects, the two-factor interactions, and the error, without requiring difficult computation. Starting with orthogonal arrays of strength $d$ as defined by Rao, fractional designs have been built up by combining the treatment combinations belonging to these arrays, the number of arrays to be taken depending upon their strength and the number of effects to be estimated. It has been shown, for instance, that if the treatment combinations regarded as points satisfying the equations $a_{\alpha 1}x_1 + a_{\alpha 2}x_2 + \cdots + a_{\alpha m}x_m = d_\alpha$ ($\alpha = 1, 2, \cdots, r$) in GF(2) form an array $(2^{m-r}, m, 2, 2)$, then the totality of treatment combinations satisfying the $(r + 1)$ such arrays given by the solutions of $\mathbf{AXJ}' = D$ in GF(2) where $\mathbf{A} = ((a_{\alpha i}))$, $\alpha = 1, 2, \cdots, r; i = 1, 2, \cdots, m$; $X' = (x_1, x_2, \cdots, x_m)$, $D = (0, \mathbf{I}_r)$ and $\mathbf{J}' = (1, 1, \cdots, (r + 1)$ times), are sufficient to estimate the main effects and the two-factor interactions, assuming all higher factor interactions to be absent. The same has been worked out with component arrays, each of strength 3. Finally, assigning the treatments belonging to $r + 1$ component arrays to blocks, block designs have been obtained. Following a slightly different procedure, fractional designs for the $3^n$ series have also been obtained.

## 34. On the Performance of the Group-Screening Method. M. S. PATEL, Research Triangle Institute. (By title)

In this paper is discussed the efficiency of two-stage group-screening designs relative to the corresponding single-stage designs. The criterion of efficiency is the expected number of correct decisions that are made using the design. Only orthogonal designs are considered. Let $\gamma$ be the size of the critical region used for decision making in a single-stage design and $\alpha$, $\beta_n$ ($n = 1, 2, \cdots, g$) be sizes of the critical regions in a two-stage design when $n$ out of $g$ group-factors are declared effective. Under certain assumptions, sufficient conditions are developed under which $\max_{\alpha, \beta_n}$ (expected number of correct decisions) exceeds $\max_\gamma$ (expected number of correct decisions) and vice-versa. Some other side results of interest are also given.

### 35. On the Distribution of First Significant Digits. ROGER PINKHAM, Rutgers University. (By title)

Consider an initial distribution on the positive real numbers. If these numbers are expressed base 10, a distribution of first significant digits is thereby induced. It is shown: (1) The only distribution of significant digits which is invariant under scale change of the initial distribution is the logarithmic, (2) A wide range of initial distributions result in the logarithmic law to a high degree of approximation.

### 36. A posteriori Distributions in the Translation Parameter Case (Preliminary report). MARTIN FOX AND HERMAN RUBIN, Michigan State University.

Let $(X, Y)$ have distribution function $F(x - \theta, y)$ with $\theta$ unknown. Assume $X$ and $\theta$ are restricted to the same linear set $S$ of real numbers (such as the integers or the rationals) while $Y$ can take values in any space. Let the *a priori* distribution of $\theta$ be "uniform" on $(-\infty, \infty) \cap S$. Let $\xi_{x,y}$ be the *a posteriori* distribution function of $\theta$ given $X = x, Y = y$. Let $\theta_{x,y;\alpha}$ be the $\alpha$th quantile of $\xi_{x,y}$ and let $\hat{\theta}_{x,y}$ have distribution function $\xi_{x,y}$. Then, (i) given $\theta$, the $(1 - \alpha)$th quantile of $\theta_{X,Y;\alpha}$ is $\theta$; (ii) if it exists, then

$$E_\theta(\hat{\theta}_{X,Y} - \theta)^n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} [\mu_k \mu_{n-k} + \operatorname{cov}(\mu_{k,Y} \mu_{n-k,Y})]$$

where $\mu_k = E_\theta(X - \theta)^k$ and $\mu_{k,y} = E_\theta(X - \theta)^k \mid Y = y]$; and (iii) given $\theta$, the distribution of $\hat{\theta}_{X,Y} - \theta$ is symmetric with absolute maximum at 0. The proofs of (i) and (ii) follow from the observation that $\xi_x(E) = F(x - E)$ for any $E \subset S$. The proof of (iii) consists in showing that $\hat{\theta}_{X,Y} - \theta = U - V$ where $U$ and $V$ are independent and identically distributed. Statement (iii) can be extended to the case of distributions invariant under a locally compact group transitive on the parameter space.

### 37. Asymptotic Relative Efficiency of Mood's Test for Two-Way Classification. Y. S. SATHE, University of North Carolina. (By title)

Let the distribution of $x_{ij} (i = 1, 2, \cdots, r; j = 1, 2, \cdots, c)$ be $F_{ij}(x) = F(x + \nu + \alpha_i + \beta_j)$. Then under the null hypothesis $H_o : \alpha_i = 0$ $(i = 1, 2, \cdots, r)$, all the observations in a given column have same distribution. Let $\tilde{x}_j$ be the median of observations in the $j$th column and in the two way table, let $x_{ij}$ be replaced by 1 if it exceeds $\tilde{x}_j$ or by 0 if it does not. Let $m_i$ be the number of 1's in the $i$th row. Then Mood and Brown show that under the null hypothesis

$$X_M^2 = \frac{\gamma(r - 1)}{ca(r - a)} \sum_{i=1}^{\gamma} \left(m_i - \frac{ca}{\gamma}\right)^2,$$

where $a = r/2$ if $r$ is even or $a = r - \frac{1}{2}$ if $r$ is odd, has a limit $\chi^2$ distribution with $r - 1$ d.f. as $c \to \infty$. In this paper, we show that under $H_c : \alpha_i = \delta_i/c^{\frac{1}{2}}$ where $\sum_{i=1}^{r} \delta_i = 0$ but not all $\delta_i$'s are 0 and $r$ odd, $X_M^2$ has a limit noncentral $\chi^2$ with $r - 1$ d.f. and noncentrality parameter

$$\lambda = \frac{\gamma^3(\gamma - 1)}{a(\gamma - a)} \binom{\gamma - 2}{a - 1}^2 \left[\int_{-\infty}^{\infty} [1 - F(y)]^{a-1} F^a(y) f^2(y)\, dy\right]^2 \sum \delta_i^2$$

where $f(y)$ is the p.d.f. For $r = 3$, $f(y) = 2\pi^{-\frac{1}{2}} e^{-\frac{1}{2}y^2}$, $\lambda = (27/16\pi) \sum \delta_i^2$. For the usual $F$-test, the corresponding noncentrality parameter is $\sum \delta_i^2$. Thus a.r.e. of $X_M^2$ compared with $F$-test for $r = 3$ is $27/16\pi = .54$.

ABSTRACT WITHDRAWN IN PROOF

### 39. The Method of Parallel Tangents (PARTAN) for Finding an Optimum. B. V. Shah, R. J. Buehler and O. Kempthorne, Iowa State University.

Let $y$ be a quadratic function of $n = 3$ variables (or a monotone function thereof) having a minimum; let $P_1P_2P_3P_4P_5P_6$ be a polygonal line such that $P_j$ occurs at the minimum value of $y$ on the (extended) line $P_{j-1}P_j$ ; let $\pi_j$ be the plane tangent to the contour of $y$ at $P_j$ . For any $P_1$ , $P_1P_2$ may be any direction of decreasing $y$. $P_2P_3$ may be any decreasing direction parallel to $\pi_1$ . $P_4$ is on the line $P_1P_3$ beyond $P_3$ . $P_4P_5$ is parallel to both $\pi_1$ and $\pi_2$ . Then $\pi_2$ and $\pi_5$ are parallel and the minimum of $y$ will be at $P_6$ which is on the line $P_2P_5$ beyond $P_5$ . The generalization to higher dimensions is to take $P_{2j-4}$ , $P_{2j-1}$ and $P_{2j}$ colinear. Then $P_{2j}$ is the minimum in the linear space spanned by the preceding points. The proof is given for hyperspherical contours, and is extended to hyperellipsoidal contours by a general affine transformation. The instructions are invariant under change of units and, more generally, under any affine transformation by virtue of the exclusive use of parallel, rather than normal, directions. In two dimensions $P_4$ is also the minimum for a quite general class of contours all of which make the same angle with any given ray from the minimum.

### 40. Selection of the Best Treatment in a Paired-Comparison Experiment. B. J. Trawinski and H. A. David, Virginia Polytechnic Institute.

The results of a balanced paired-comparison experiment in which ties are not permitted can be summarized in the scores (number of preferences) achieved by the treatments being judged. Let $\pi_{ij}$ $(i, j = 1, 2, \cdots, t; i \neq j)$ denote the constant probability that the $i$th treatment $T_i$ is preferred to $T_j$ in any one of $n$ comparisons of the two. Then the best treatment $T_b$ is the one with the highest average probability $\pi_{i\cdot}$ of success. Following the general approach of Gupta and Sobel one can select a subset $S$ of the $T_i$ by including in $S$ the $T_i$ associated with the top score $k_{(t)}$ and with all other scores $\geq k_{(t)} - \nu$. With proper choice of $\nu$ this decision rule has the property that $T_b$ is included in $S$ with at least a specified probability. Tables have been constructed giving $\nu$ as a function of $t$ and $n$. Suppose next that $\pi_{tj} = \pi > \frac{1}{2}$ $(j = 1, 2, \cdots, t - 1)$ and $\pi_{ij} = \frac{1}{2}$ otherwise. For this case the smallest number of replications $n$ has been determined, for given $t$ and $\pi$, which ensures that with at least a specified probability the highest score corresponds to $T_b(= T_t)$. The formulation of the problem is that of Bechhofer but in the present case the procedure does not have the usual conservative properties associated with Bechhofer's approach.

### 41. A Plotting Procedure in MANOVA. M. B. Wilk and R. Gnanadesikan, Bell Telephone Laboratories, Murray Hill. (Invited paper)

A procedure is presented for the generalization and extension, to multiresponse factorial experiments, of the technique of half-normal plotting for uni-response factorials. Consider, for definiteness, two-level factorial experiments wherein, for each treatment combination, $p$ responses are observed. For this multiresponse situation, the analogue of the uni-response single degree of freedom contrasts is a vector of $p$ elements, each element being a single degree of freedom contrast corresponding to one of the responses. A positive semidefinite or definite quadratic form in the elements of each of these vectors is obtained (for example, the squared length of the vector). This is interpretable as a distance function in a meaningful space. The null distribution of the quadratic form is approximated as a gamma with two parameters. Under reasonable experimental assumptions, the quadratic forms are mutually independent. Using only the $m$ smallest of the quadratic forms as a sample of the first $m$ order statistics from a sample of size $k$ (where $m \leqq k \leqq N$, the total number of contrasts) from a gamma distribution, maximum likelihood estimates are obtained for the parameters of the gamma distribution. Using the estimates, a "gamma plot" is made of the ordered quadratic forms. Interpretations and uses of the plot are discussed, with examples.

**42. On the Estimation of Error Variance and Number of Significant Effects in Two-level Factorial Experiments.** M. B. WILK, R. GNANADESIKAN, AND MISS A. R. EBNER, Bell Telephone Laboratories, Murray Hill. (Invited paper)

This paper is concerned with the estimation of the error variance, $\sigma^2$, and the number of "significant" contrasts, $r$, when a collection of $N$ contrasts, each based on one degree of freedom, are analyzed. A typical situation giving rise to this is the set of $N = 2^n - 1$ estimates of effects obtained from the responses in a $2^n$ factorial experiment. Let $y_1^2 < y_2^2 < \cdots < y_N^2$ denote the ordered squared contrasts. Suppose that the $m$ smallest ($m$ specified) are known to "belong to error", and that $k = N - r$, the number of error contrasts, is known. Tables are provided for determining at once the maximum likelihood estimate of $\sigma^2$ based on the first $m$ order statistics $y_1^2, \cdots, y_m^2$. A summary of a Monte Carlo study of the properties of this estimate is given. Though $k$ will in general not be known, it is felt that this procedure, based on a "guessed" $k$ will be less biased than one based on either pre- or post-experiment "assignment" of effects to error or of the "conservative" estimation of error from all the contrasts in the collection. The procedure is buttressed by an informal sequential inference scheme which is specifically directed toward the determination of $k$ from the data. Examples and results of some Monte Carlo studies are given.

**43. Estimation of Parameters of the Gamma Distribution Using Order Statistics.** M. B. WILK, R. GNANADESIKAN, AND MISS M. J. HUYETT, Bell Telephone Laboratories, Murray Hill. (Invited paper)

Using the $m$ smallest observations in a random sample of size $K$ from a gamma distribution whose density function is,

$$f(y; \lambda, \eta) = [\lambda^\eta/\Gamma(\eta)]e^{-\lambda y}y^{\eta-1}, \qquad y > 0, \lambda > 0, \eta > 0,$$

the problem of estimating $\lambda$ and $\eta$ is considered. Let $P$ and $S$ denote, respectively, the ratios of the geometric mean and the arithmetic mean of the $m$ smallest observations to the $m$th observation. Then, $0 \leqq P \leqq S \leqq 1$, and $P$ and $S$ are jointly sufficient for $\lambda$ and $\eta$. Tables are given, for various values of $K/m$, $P$ and $S$, from which the maximum likelihood estimates of $\lambda$ and $\eta$ are easily obtainable.

**44. Probability Plots for the Gamma Distribution.** M. B. WILK, R. GNANADESIKAN, AND MISS M. J. HUYETT, Bell Telephone Laboratories, Murray Hill.

If $y_1 \leqq y_2 \leqq \cdots \leqq y_n$ is an ordered random sample from the general gamma distribution with density

$$f(y; \alpha, \lambda, \eta) = [\lambda^\eta/\Gamma(\eta)](y - \alpha)^{\eta-1}e^{-\lambda(y-\alpha)}; \qquad \alpha \leqq y < \infty,$$

where $-\infty < \alpha < \infty$, $\lambda > 0$ and $\eta > 0$, then a plot of the $y$ values against appropriate quantiles of the standard gamma distribution ($\alpha = 0$, $\lambda = 1$) would be expected to yield a straight line pattern with intercept $\alpha$ and slope $1/\lambda$. This paper describes a systematic numerical method for calculating the quantiles of the standard gamma distribution corresponding to the fractions $b_i = (i - \frac{1}{2})/n$, $i = 1, 2, \cdots, n$, for *given* values of $\eta$. A table of quantiles as a function of $\eta$ is given, together with several samples of gamma plots. Various uses of such plots are discussed, with examples. The entire procedure, viz the calculation of quantiles and the actual plotting, has been programmed for a high-speed computer and that procedure (with instructions for use) is described in an Appendix.

*(Abstracts of papers to be presented at the Annual Meeting of the Institute, Seattle, Washington, June 14–17, 1961. Additional abstracts will appear in the September, 1961 issue.)*

1. **An Optimal Sequential Accelerated Life Test.** Stuart A. Bessler, General Telephone and Electronic Laboratories, Herman Chernoff, Stanford University, and Albert W. Marshall, Stanford University and Institute for Defense Analyses, Princeton.

Suppose that the distribution of lifetime $T$ of a device subjected to stress $x$ is exponential with failure rate (reciprocal of mean) equal to $\theta_1 x + \theta_2 x^2$ for $0 \leqq x \leqq x^*$. It is desired to test $H_1$: $\theta_1 x_0 + \theta_2 x_0^2 \leqq a$ versus $H_2$: $\theta_1 x_0 + \theta_2 x_0^2 > a$. The cost of experimenting at stress level $x$ is assumed to be proportional to the mean lifetime $(\theta_1 x + \theta_2 x^2)^{-1}$. An asymptotically optimal sequential procedure, characterized by Chernoff, Albert, and Bessler, is derived. After each observation it calls for selecting the next observation from one of two levels. These levels correspond to the solution of two person zero sum game whose payoff function is a Kullback-Leiber Information number divided by the cost of experimentation.

2. **A Statistic Related to Kolmogorov's.** H. D. Brunk, University of Missouri.

A distribution-free statistic is proposed for testing the hypothesis that a sample comes from a population with prescribed distribution function $F$. It occupies a position intermediate between Kolmogorov's (*Inst. Ital. Attuari, Giorn.*, Vol. 4 (1933), pp. 1–11) and Sherman's (*Ann. Math. Stat.*, Vol. 21 (1950), pp. 339–361), and the corresponding test appears more powerful than Kolmogorov's against certain alternatives (e.g., different scale parameter, for a symmetric distribution) and less powerful against others (e.g., different location parameter). If $F_n$ is the empiric distribution function of the sample, then the statistic, proposed by Dr. Harold Lischner and the author, is $\max_x [F(x) - F_n(x)] - \min_x [F(x) - F_n(x)]$; more precisely, it bears the same relationship to that mentioned as does Pyke's (*Ann. Math. Stat.*, Vol. 30 (1959), pp. 568–576) modification of Kolmogorov's to Kolmogorov's itself. Asymptotic formulas due to Doob (*Ann. Math. Stat.*, Vol. 20 (1949), pp. 393–403) and Donsker (*Ann. Math. Stat.*, Vol. 23 (1952), pp. 277–281) yield the asymptotic distribution. A theorem of Sparre Andersen (*Skand. Aktuarietidskrift*, Vol. 36 (1953), pp. 123–138) makes possible an essential simplification of the problem of determining the distribution for finite sample size. After this simplification, methods developed for Kolmogorov's statistic by Kolmogorov, Feller, Dempster and others can be used. Tables are in preparation.

3. **Evaluation and Design of Multiple Choice Questionnaires** (Preliminary report). H. Chernoff, Stanford University.

The evaluation of tests based on multiple choice questions is complicated by the fact that subjects who do not know the answer have the opportunity to guess. If only $\frac{1}{3}$ of the subjects answer correctly a question with three choices, one may infer that very few subjects knew the answer. The standard procedure of giving credit to those who answered this question correctly serves no purpose but to introduce "noise" into the test score. In this paper, a method is proposed to give scores which depend not only on the individual's answers but on those of the population. The score $x$ is regarded as an estimate of a value $v$ attached to the subject's knowledge. The method of scoring is to be selected so as to minimize $E(x - v)^2$. Applications to various models are given. The problem of optimal design is discussed. The mathematics of the theory reduces to that of regression or conditional expectation. For example, the best score given the subject's response $z$ is $x(z) = E\{v \mid z\}$. If, in the example described above, $p$ is the proportion of students who answer a one-point question correctly, all correct answers are given a score of $(3p - 1)/2p$ and all incorrect answers a score of zero.

### 4. Optimal Accelerated Life Designs for Estimation. H. Chernoff, Stanford University, (Invited paper).

It is desired to estimate the probability distribution of the lifetime of a device when subjected to a standard environment. If this lifetime tends to be large, the time consumed in testing a sample of devices will be exorbitant. A common approach is that of acceleration. The sample of devices is subjected to environments of greater stress than that of the standard environment. To extrapolate these results a model relating distribution of lifetime to environment is required. Once such a model is proposed the problem of optimal design arises. At what level of stress should experiments be carried out? An answer to this question involves the cost of experimenting at a certain stress level. Optimal designs are derived for five examples. In each of these the distribution of lifetime is assumed to be exponential with failure rate (reciprocal of the mean) equal to a specified function of stress and the cost of experimentation proportional to the mean lifetime at the stress level. Each of the optimal designs consists of experimenting at two stress levels. The method reduces to an application of a technique of Elfving derived for estimating the optimal designs for the coefficients of a linear regression.

### 5. Minimum Risk Estimation: A Non-Parametric Case Involving Percentiles. Herbert B. Eisenberg and John E. Walsh, System Development Corporation.

Minimum risk estimation of 100 $p$-percentile, $\theta_p$, is considered for any continuous univariate population. Sample values are $x_1 \leqq \cdots \leqq x_n$; $x_0 = -\infty$; $x_{n+1} = \infty$. Risk $= \sum_{i=0}^{n} L(\hat{\theta}_p, \theta_p^{(i)} \mid x_i < \theta_p \leqq x_{i+1}) P_r(x_i < \theta_p \leqq x_{i+1})$ and estimate is value $\hat{\theta}_p$ minimizing this function; $\theta_p^{(i)}$ in conditional loss function is representative value for $\theta_p$ given that $\theta_p$ in $(x_i, x_{i+1}]$ and can be randomly selected from a specified distribution over this interval. Method is applicable when the minimizing $\hat{\theta}_p$ is unique. If conditional loss function is $[\hat{\theta}_p - \theta_p^{(i)}]^2$, minimizing $\hat{\theta}_p = \sum_{i=0}^{n} \theta_p^{(i)} \binom{n}{i} p^i (1 - p)^{n-i}$. Conditions for applicability of method are examined. Explicit results are derived for special classes of loss functions. Further nonparametric applications (univariate and multivariate) are being developed through extension of percentile concept to that of general population coverage—specifically, to minimum risk estimation of regions with specified amounts of population coverage. Method is also applicable to estimation of parameters for parametric populations (continuous or discrete and univariate or multivariate). In general, the confidence intervals $x_i < \theta_p \leqq x_{i+1}$ are replaced by mutually exclusive confidence regions that include all the possible values for the quantity estimated. For the continuous parametric case, size of all but tail confidence regions ordinarily can be made arbitrarily small. Then, in the limit risk function is usually expressible as an integral.

**6. Non-Parametric Sequential Tests for the Two-Sample and Several-Sample Problems.** R. R. M. GEOGHAGEN AND JOHN E. WALSH, System Development Corporation.

Null hypothesis asserts that $k$ specified populations ($k \geqq 2$) are equal. Populations can be univariate or multivariate and arbitrary otherwise. Same number of sample values is taken from given population at each sequential step; this number can be different for each population. Each observation is converted to categorical form by subdivision (same for all populations) of space of possible values. Categories are denoted by $1, \cdots, C$ and should have nonzero null probabilities. At each sequential step, observed value is vector with dimensionality equal to total number of observations per step and each coordinate is one of $1, \cdots, C$. A value with all coordinates equal is discarded. Disjoint classes are formed from remaining possible vector values so that each class has a determined null probability (many possible ways unless $C$ very small). In all cases, these classes have a multinomial distribution. The completely determined null distribution can be tested against specified alternative distributions by standard sequential analysis methods. Also some special sequential methods are presented. When $C$ is not very small, alternatives of a broad class can be emphasized by suitable choice of category subdivisions and suitable formation of classes. Procedures for making suitable selections are outlined for several univariate cases and a multivariate case. Also methods for determining category subdivisions with nonzero null probabilities are considered.

**7. Three-Quarter Replicates of $2^n$ Designs.** PETER W. M. JOHN, California Research Corporation and University of California, Berkeley.

Let a three-quarter replicate be formed by omitting the quarter defined by $I = P = Q = PQ$, where $P$, $Q$, $PQ$ are main effects or interactions. Let $R$ be any other effect. Then if one member, say $PQR$, of the alias coset $R$, $PR$, $QR$, $PQR$ is negligible, *a priori*, the least squares estimates of the remaining members are obtained from half replicates confounded only with $PQR$. If, in addition, $QR$ is negligible, then the least squares estimates of $R$ and $PR$ are obtained by averaging their estimates from two half replicates, and are confounded only with $QR$ and $PQR$. In the three-quarter replicate of $2^n$ obtained by omitting $I = AB = AC = BC$, the alias cosets consist either of four effects with odd numbers of letters or else of four effects with even numbers of letters. In each of the alias cosets containing "odd" effects, equate three of these to main effects, and we have a design of Type B (main effects clear of two factor interactions) for $3.2^{n-3}$ factors with $3.2^{n-2}$ points. Other designs obtained include a three-eighths replicate, 48 points, of $2^7$ of Type B′ (both main effects and two factor interactions clear) and a six point first-order design for five factors.

**8. Oddities in Estimating the Scale Parameter of the Weibull Distribution.** EUGENE H. LEHMAN, JR., Purdue University. (Introduced by Irving W. Burr.)

In a recent paper [*Estimation of the Scale Parameter in the Weibull Distribution Using Samples Censored by Time and by Number of Failures*, Ph.D. Thesis, North Carolina State College, Raleigh, (1961)] the author described a test procedure for estimating the scale parameter $\alpha$ in the Weibull distribution, $F(t) = 1 - \exp[-(t^m/a)]$, $t > 0$. The test is: place on test $N$ tubes whose life spans follow the Weibull distribution with known scale parameter $M$; stop test when both $R$ tubes have failed and $T$ hours has elapsed. The maximum likelihood estimator $\hat{\alpha}$ of $\alpha$ from this procedure possesses several mysterious idiosyncrasies. In the present paper some of these strange mannerisms are explained. For instance, the variance, $V$, and mean square error, $D$, (if $N$ and $R$ are small) *increase* with $T$ for a short interval of small values of $T$, thus implying that a short test supplies more information than a long one. The reason for this peculiarity is that if the actual number of failures, $r$, exceeds by only a small integer the minimum $R$, then $\hat{\alpha}$ is very biased and variant. If $T$ is small, the probability of using this particular form of $\hat{\alpha}$ increases with $T$ before it decreases, and thus $V$ and $D$ of $\hat{\alpha}$ as a whole portray this same confusing oddity.

### 9. **The Distribution of Probabilities in a Stochastic Learning Model** (Preliminary report). J. R. McGregor and T. V. Narayana, University of Alberta. (By title)

Bush and Mosteller (1955) have developed a model to describe simple learning behavior in the case of two *subject-controlled events*. They consider a learning experiment in which a subject is presented, on repeated trials, with the same two alternatives $A_1$ and $A_2$. Their probabilistic model specifies the way in which the response probability (i.e., the probability of choice of alternative $A_1$) is modified from trial to trial as a result of earlier responses and outcomes. On the $n$th trial there will be $2^n$ possible values of the response probability (called *p-values*) the $\nu$th value being denoted by $p_{\nu n}$ ($\nu = 1, 2, \cdots, 2^n$). These $p$-values have a probability distribution, say, Prob $(p_{\nu n}) = P_{\nu n}$. In a special case, the transition from trial $n$ to trial $(n + 1)$ results in each old $p$-value $p_{\nu n}$ generating two new $p$-values given by $Q_1p_{\nu n} = \frac{1}{2}p_{\nu n}$, $Q_2p_{\nu n} = \frac{1}{2} + \frac{1}{2}p_{\nu n}$. It is the purpose of this note to show that in the special case considered, the use of binary representations of integers or equivalently of the boolean algebra formed by the subsets of a set yields interesting results about the exact asymptotic distribution.

### 10. **Some Properties of Compositions of an Integer and their Application to Probability and Statistics.** T. V. Narayana and S. G. Mohanty, University of Alberta. (By title)

The authors have developed a unified approach to various problems in probability and statistics, through partial orders defined on compositions of an integer (abstract published in *Ann. Math. Stat.*, June, 1961). This approach yields a new proof of a theorem of Chung and Feller: Let $L_{k,n}$ be the number of minimal lattice paths from the origin to $(n, n)$ such that $2k$ steps lie above the line $x = y$ and $2n$-$2k$ steps below it ($k = 0, 1, \cdots, n$); then $L_{k,n} = (n + 1)^{-1}\binom{2n}{n}$, independently of $k$. A feature of our (non-inductive) proof is the explicit 1:1 correspondences which are provided by it, between any two of the $n + 1$ sets of paths considered. Similar methods may be applied to the problem of setting an upper bound to the number of non-isomorphic scores in a round robin tournament.

### 11. **Some Tests for Outliers.** C. P. Quesenberry, Montana State College, and H. A. David, Virginia Polytechnic Institute.

This paper is concerned with the problem of detecting outlying observations when in addition to the normal sample $x_1, x_2, \cdots, x_n$ at hand an independent mean-square estimate $s_\nu^2$ of the common variance is available. The maximum deviate from the sample mean $(x_{\max} - \bar{x})$ divided by the pooled estimate of standard deviation $S$ and the maximum absolute deviation from the sample mean $\max_i |x_i - \bar{x}|$ divided by $S$ are considered as test statistics. These statistics are known to have optional power properties for testing against alternatives of one outlier on one (specified) side of the sample and for one outlier on either side of the sample, respectively. The distributions of the related statistics $(x_i - \bar{x})/S$ are obtained and used in conjunction with the Bonferroni inequalities to obtain significance points for the one- and two-sided test statistics. Tables of significance points of the test statistics for $\alpha = .01$ and $.05$ are given for selected values of $n$ (sample size) and $\nu$ (the degrees of freedom for the independent estimate of variance). Two examples are given to illustrate use of the tables.

### 12. **Estimation of Location and Scale Parameters by Optimally Selected Observations.** CARL-ERIK SARNDAL, University of North Carolina. (Introduced by Bernard G. Greenberg.)

A large sample of $n$ observations, distributed according to a function of type $F[(z - \alpha_1)/\alpha_2]$, is available. Those $k$ observations $z_{(n_1)} < z_{(n_2)} < \cdots < z_{(n_k)}$ which contain most information with respect to the unknown parameters $\alpha_1$ and $\alpha_2$ are to be selected to form linear estimates $\alpha_r^* = \sum_{i=1}^{k} g_{ri} z_{(n_i)}$, $(r = 1, 2)$. Another question to be answered is to what extent this censoring of the sample can be carried out in order to maintain the quality of joint asymptotical efficiency. If $\lambda_i = n_i/(n+1)$, it is found that, under certain general conditions, the optimum spacings are of the form $\lambda_i = Q[i/(k+1)] + O(k+1)^{-2}$, where $Q$ is a function, the inverse of which is expressible in terms of the density function ts derivatives. Using this result, the joint asymptotical efficiency of $\alpha_1^*$ and $\alpha_2^*$ is shown to be of the form $1 - C(k+1)^{-2} + o(k+1)^{-2}$, where $C$ is a certain constant. In fact, joint asymptotical efficiency is achieved by using only $k = k(n)$ observations, where $k(n) \to \infty$ with $n$ in the weakest possible manner. It is to be expected that, also for rather small values of $k$, satisfactory approximations to the optimum spacings can be obtained by taking $\lambda_i = Q[i/(k+1)]$. For the important case of the normal distribution, calculations show that these "nearly optimum" spacings yield estimates of very high joint asymptotical efficiency by utilizing only 10 observations.

### 13. **On the Jiřina Sequential Tolerance Limits Procedure.** SAM C. SAUNDERS, University of Wisconsin.

This paper contains a short exposition of the necessary assumptions that the Jiřina Sequential Tolerance Limits Procedure (*Selected Trans. in Math. Stat. and Prob.*, I.M.S. & A.M.S., Vol. 1, 1961, pp. 145–156) be generalized to spaces other than the real line. A method of obtaining the confidence level for any parameter values is given. A table is computed for small values and for large values it is shown that tables of the exponential integral can be used. Formulae for the expectation and variance of the random sample size are derived. That the random sample size, appropriately scaled by one of the parameters, has an asymptotic distribution as that parameter increases is proved and the La Place transform of this distribution is found. Also formulae for the asymptotic mean and variance are found as one parameter is increased.

### 14. **Moments of the Radial Error.** ERNEST M. SCHEUER, Space Technology Laboratories.

Let $x_1$ and $x_2$ be normally distributed random variables with zero means, variances $\sigma_{11}$ and $\sigma_{22}$, respectively, and covariance $\sigma_{12}$. Define the radial error $R$ by $R = (x_1^2 + x_2^2)^{\frac{1}{2}}$. Further, let $\sigma_1^2 = \frac{1}{2}\{(\sigma_{11} + \sigma_{22}) + [(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}^2]^{\frac{1}{2}}\}$, $\sigma_2^2 = \frac{1}{2}\{(\sigma_{11} + \sigma_{22}) - [(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}^2]^{\frac{1}{2}}\}$, and $k^2 = (\sigma_1^2 - \sigma_2^2)/\sigma_1^2$. Then for the moments about the origin $\mu_n'$ of $R$, one obtains: $\mu_{2p} = \{[(2\sigma_1^2)^p p!]/\pi^{\frac{1}{2}}\} \sum_{\nu=0}^{p} \binom{p}{\nu} (-k^2)^\nu [\Gamma(\nu + \frac{1}{2})]/\nu!$ and $\mu_{2p+1} = 1 \cdot 3 \cdots (2p+1)[(\pi/2)^{\frac{1}{2}}]$ $F(-\frac{1}{2}(2p+1), \frac{1}{2}, 1; k^2)$ where $F(a, b, c; z)$ is the hypergeometric function. Cf., the results of Rice (in Wax, Ed., *Selected Papers on Noise and Stochastic Processes*, Dover, 1954, p. 238 et. seq.) who treats the case $\sigma_{12} = 0, \sigma_{11} = \sigma_{22}$, $E(x_1) = P$ (generally not zero), $E(x_2) = 0$.

### 15. An Optimal Sequential Accelerated Life Test with Exponential Dependence on Stress. Gideon Schwartz, Stanford University.

Bessler, Chernoff and Marshall present a game-theoretic method for obtaining optimal sequential designs for experiments to test whether a device subjected to standard stress has an expected lifetime exceeding a specified value. They solve explicitly the case in which that function is the inverse of a quadratic. In the present paper, the same method is applied to the case where the dependence of the expected lifetime on the stress is exponential.

### 16. Tests for Regression Coefficients When a Continuous Sample is Available. M. M. Siddiqui, Boulder Laboratories, National Bureau of Standards.

In continuation of previous studies (these *Annals*, Vol. 29 (1958), pp. 1251–56; Vol. 31 (1960), pp. 929–938), tests for regression coefficients are developed when a continuous sample is available and the error process is assumed to be stationary. Bivariate Gram-Charlier series are used to approximate the joint distribution of the least-squares estimate of a regression coefficient and the integral of squared residuals. Finally, approximate $t$ and $F$ tests are suggested. The theory is applied to some cases of interest such as analysis of variance for one way classification, and polynomial trend.

### 17. A Generalization of the Ballot Problem and Its Application in the Theory of Queues (Preliminary report). Lajos Takács, Columbia University. (Invited paper)

In the case of single server queueing processes denote by $G_n(x)$ the probability that the busy period consists of $n$ services and its length is $\leqq x$. The author determines $G_n(x)$ explicitly for types $[M, E_m, 1]$ and $[E_m, M, 1]$ by using the solution of the ballot problem and for types $[M, G, 1]$ and $[G, M, 1]$ by using the following more general theorem: Suppose that an urn contains $n$ cards labeled with nonnegative integers whose sum is $k \leqq n$. All the $n$ cards are drawn without replacement from the urn. The probability that for $r = 1, 2, \cdots, n$ the sum of the first $r$ numbers drawn is less than $r$ is $1 - k/n$.