ABSTRACTS OF PAPERS

(Abstracts of papers presented at the European Regional Meeting of the Institute, Dublin, Ireland, September 3–5, 1962. Additional abstracts appeared in the June, 1962 issue, and will appear in the December, 1962 issue.)

5. **Partial Geometries and Partially Balanced Designs.** R. C. BOSE, University of North Carolina.

The concept of partial geometry has been introduced. This is defined as a system of undefined objects, points and lines and a relation incidence satisfying the axioms (i) any two points are incident with utmost one line, (ii) each line is incident with k points, (iii) each point is incident with r lines, (iv) if a point P is not incident with a line 1, there exist exactly t points, $Q_1, Q_2, \cdots, Q_t$ on 1, such that P and Q_i are incident with a line. It is shown that a partial geometry is a PBIB design based on the association scheme with parameters $n_1 = r(k-1)$, $n_2 = (r-1)(k-1)(k-t)/t$, $p_{11}^1 = (t-1)(r-1) + k - 2$, $p_{11}^2 = rt$. Conversely if there is an association scheme with the above parameters a sufficient condition that we can find a PBIB design based on it (which is a partial geometry) is that $k > \frac{1}{2}[r(r-1) + t(r+1)(r^2 - 2r + 2)]$. A general uniqueness theorem is deduced, from which a number of results by Bruck, Shrikhande, Connor and others can be deduced as special cases.

6. **Conditions under which a Given Process is a Function of a Markov Chain** (Preliminary report). MARTIN FOX, Michigan State University.

The notation of this abstract is that of Gilbert (*Ann. Math. Statist.*, **30**, 688–697). Gilbert proved that a necessary condition for $\{Y_n\}$ to be a function of a given finite stationary Markov chain is that $\Sigma_\epsilon\, n(\epsilon) \leqq N$ where N is the number of states of the Markov chain. Gilbert conjectured that if $\Sigma_\epsilon\, n(\epsilon) < N$, then $\{Y_n\}$ is representable as a function of a Markov chain with $\Sigma_\epsilon\, n(\epsilon)$ states. It is now shown that, given $\{Y_n\}$ with $n(\epsilon) \leqq 2$ for each state ϵ, a Markov chain $\{X_n\}$ with $\Sigma_\epsilon\, n(\epsilon)$ states can be constructed such that $\{Y_n\}$ is a function of $\{X_n\}$. Neither stationarity nor finiteness of the state space for $\{Y_n\}$ are required for this construction. An example exists of a stationary chain $\{Y_n\}$ with states $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$ for which $n(\epsilon_1) = 3$ and $n(\epsilon_2) = n(\epsilon_3) = n(\epsilon_4) = 1$ yet $\{Y_n\}$ any representation of $\{Y_n\}$ as a function of a Markov chain requires at least seven states. This example suggests (but does not prove) the falsity of Gilbert's other conjecture that $\Sigma_\epsilon\, n(\epsilon) < \infty$ implies that $\{Y_n\}$ is a function of a finite Markov chain. It appears that a countable state space would be needed for some examples.

7. **Asymptotic Relative Efficiency of the Combined Sign Test and Wilcoxon's Test for Symmetry.** D. J. STOKER, University of Pretoria, Republic of South Africa.

Van Eeden and Benard (*Indag. Math.* **19** (1957) 381–408) suggested a procedure which consists of a combination of the sign test and Wilcoxon's test for symmetry. In the present paper it is proved that the simultaneous probability distribution of the test statistics of the sign test and Wilcoxon's test for symmetry is asymptotically the bivariate normal distribution under a very general class of hypotheses. Using this result the a.r.e. (asymptotic relative efficiency) of the combined test procedure with regard to the standard t-test is derived for observations from a symmetric distribution with continuous frequency function. In general, no explicit solution of the a.r.e. could be obtained. It has to be calculated numerically in special cases. This was done for the uniform and normal distributions, showing that the a.r.e. is a function of the size of the test and of c, increasing with increasing values of c, where c is specified by the sequence of alternatives, viz. $c/N^{\frac{1}{2}}$, used to obtain the a.r.e. For example, if the size of the test lies between 0.004 and 0.08 the a.r.e. is approximately equal to 0.646 (uniform distribution) and 0.818 (normal distribution) for $c \to 0$ and for increasing values of c, it increases to values greater than 0.94 (uniform distribution) and 0.91 (normal distribution).

(Abstracts of papers presented at the Annual Meeting of the Institute, Minneapolis, Minnesota, September 7–10, 1962. Additional abstracts appeared in the June, 1962 issue, and will appear in the December, 1962 issue.

3. The Reliability of Repairable Complex Systems, Part B: The Dissimilar Machine Case. Ronald S. Dick, International Electric Corporation.

A generalization of the reliability model given in Part A of this reliability study is made. The failure and repair rates are assumed to be exponential as in Part A, but the model equations are changed so that the following ramifications are possible.

(1) The label case: it is necessary to know which of the similar machines are broken down in order to decide if a system failure has occurred.

(2) The parameter case: the values of the failure and repair rates are not the same on each machine.

(3) The type of failure case: the system goes from state i to $i + j$ after a failure, $j \geqq 1$, and to $i - k$ after a repair, $k \geqq 1$, whereas in Part A $j = k = 1$ only.

A procedure for finding dissimilar machine model equations with time delays before failure and time restoration constraints after repair are given. Reference is made to Dick, R. S. (1961) The Reliability of Repairable Complex Systems, Part A: The Similar Machine Case. *5th MIL-E-CON Symposium on Military Electronics*, Washington, D. C., and Bejarano, R. and Dick, R. S. (1961) Tables for the Reliability of Repairable Systems with Time Constraints. *Ann. Math. Statist.*, **32**, 914 (Abstract).

4. A Characterization of the Geometric Distribution. Thomas S. Ferguson, University of California at Los Angeles.

A random variable X is said to have a geometric distribution with location parameter α, scale parameter $\beta > 0$ and geometric parameter π, $0 < \pi < 1$, if the distribution of $(X - \alpha)/\beta$ is given by the probability mass function $p_k = (1 - \pi)\pi^k$ for $k = 0, 1, 2, \cdots$. As a characterization of the geometric distribution the following *theorem* may be proved. Let X and Y be independent, non-degenerate, discrete random variables. Then min (X, Y) and $X - Y$ are stochastically independent if, and only if, both X and Y have geometric distributions with the same location and scale parameters (and possibly different geometric parameters). In this paper, the following more general problem is completely solved. What are the possible distributions of independent, discrete random variables X and Y, for which the random variables min (X, Y) and $|X - Y|$ are independent? As a corollary of this solution, it is seen that the only such distributions which are non-degenerate and identical, are the geometric distributions with identical location scale and geometric parameters. Similar characterizations are valid for the negative exponential distribution.

5. **Application of Range in Sequential Analysis.** Bhaskar K. Ghosh, Lehigh University.

Three problems on testing of composite hypotheses have been considered: (i) ratio of variances of two normal populations, (ii) ratio of between-groups and within-group variances in a one-way classification by groups under a random model, (iii) ratio of between main-groups, and within sub-group variances, and of between sub-groups and within sub-group variances in a second-order hierarchical classification by main-groups and sub-groups. Two sequential procedures are constructed for each problem using sample ranges. The properties of the tests are critically examined and possible optimal values of parameters are suggested which will minimize the ASN and cost. The range procedures are compared with the corresponding fixed-sample and sequential variance procedures. It is also shown that when range is used in sequential tests of the type considered here a Chi-approximation to the distribution of range, as proposed by Patnaik (*Biometrika*, 1950), is more adequate than a Chi-square approximation which was suggested by Cox (*J. Roy. Statist. Soc.*, 1949) and subsequently used by Cox (*ibid.*) and Rushton (*Sankhya*, 1952) to construct a sequential test based on range for problem (i). An exact sequential test based on the true distribution of range is also constructed for particular cases of problems (i) and (ii).

6. **Optimum Allocation for Sampling with Replacement in Stratified Sampling.** Sakti P. Ghosh, I.B.M. Research Center, New York.

In a stratified sample when sampling is done with replacement in each strata, a better estimate of the population mean can be achieved by considering the weighted mean of the stratum-means based on distinct units only. An explicit expression for the variance for the mean of a stratified sample based on distinct units only, is obtained. Then the optimum allocations for the different stratum are obtained by minimizing this variance, subject to the condition that expected number of distinct units is fixed. Neyman's solution for optimum allocation follows as a special case.

7. **Inverse Moments.** Zakkula Govindarajulu, Case Institute of Technology.

In recent years inverse moments of positive discrete random variables (for example, the binomial, the Poisson, the negative binomial and the hypergeometric variables truncated at zero) have been of interest. Inverse moments of continuous distributions are also of some interest since the latter can be used as approximations to the inverse moments of the positive discrete random variables. A precise condition for the existence of inverse moments for an arbitrary distribution is given. Liapownoff's inequality for the regular moments has been extended to hold for the inverse moments. The inverse characteristic or moment generating function has been defined and its properties of uniqueness, continuity and convergence have been studied. Levy's inversion formula for the regular characteristic functions has been extended to the inverse characteristic functions. Other results have also been obtained. Some examples are considered.

8. **A Test of Linearity Versus Convexity of a Median Regression Curve.** Bruce M. Hill, University of Michigan.

A test of linearity versus convexity of a median regression curve is presented. The test consists in estimating a line by the Mood-Brown procedure (using medians) from a central subset of the observations, making a weighted count of the number of remaining observations lying above the line, and rejecting the hypothesis of linearity if this number is large.

The asymptotic distributions of the estimated line and of the test statistic are derived, and the test is shown to be consistent against twice differentiable convex alternatives.

9. Small Sample Efficiency for the One Sample Wilcoxon and Normal Scores Tests. JEROME KLOTZ, University of California, Berkeley.

A recursive scheme is given which permits computation of rank order probabilities in the one sample case. The scheme is applied for normal shift alternatives to compute power and efficiency for the Wilcoxon and normal scores tests up to sample size $n = 10$. Local efficiency for the two nonparametric tests is computed for finite sample size using the values of the normal scores statistic. In addition, efficiencies for large shifts are obtained by comparing the rate of exponential convergence to zero of the type II error. The methods used are similar to those used by Hodges and Lehmann in the two sample problem. The efficiencies of the Wilcoxon and normal scores tests are quite high relative to the t-test in the region of interest. Although depending upon the level α, it appears, roughly, that the efficiency for the Wilcoxon decreases with increasing sample size towards its limiting Pitman value of $3/\pi$. The power of the normal scores test is usually only slightly greater than that of the Wilcoxon. However, for some normal alternatives the Wilcoxon is slightly better with generally higher efficiency for large shift. A very strong case can be made to use the nonparametric tests in place of the t-test.

10. Efficiency of the Wilcoxon Two-Sample Test for Randomized Blocks (Preliminary report). GOTTFRIED E. NOETHER, Boston University.

The Wilcoxon two-sample test is often used to compare the effectiveness of two treatments producing observations $x_1, \cdots, x_N$ and $y_1, \cdots, y_N$, respectively. In the Mann-Whitney form, the test is based on the statistic $U = \#(x_i > y_j)$, $i, j = 1, \cdots, N$, where $\#(\ )$ equals the number of times that the relationship indicated within parentheses is satisfied. Proper experimental design requires complete randomization of the $2N$ units involved in the experiment. If the $2N$ units occur in b blocks containing $2n$ units each, a randomized block design may be more sensitive. An appropriate Mann-Whitney type statistic is given by $T = \sum_{m=1}^{b} U_m$, where U_m is the Mann-Whitney statistic for the observations in the mth block. It is suggested to call the quantity $E = (2N + 1)e^2/(2N + b)$ where $e = [(4N - 2)T - nN(2N - 1)]/[4nU - 2T - nN(2N - 1)]$ an estimate of the efficiency of the Wilcoxon two-sample test for randomized blocks. In e, U is computed on the basis of the observations of the randomized block design. If $b = N$, the T-statistic reduces to the sign test statistic. In this case, $E \doteq 2e^2/3$.

11. Dependability Models for a System of N Parallel Elements. RAY E. SCHAFER and J. M. FINKELSTEIN, Hughes Aircraft Company. (Invited paper)

The dependability measures, availability and the probability of survival for time t, are derived for a system consisting of N parallel elements. It is assumed the elements fail independently and are repaired independently according to the one parameter exponential distribution. Failure of the system is defined to occur when, and only when $X \geqq N - K + 1$ elements are in a failed state, $K \geqq 1$. Thus, degraded forms of operation are permitted. The method used is to set up the difference equations describing the stochastic process of system life and death, i.e., state probabilities; convert these to differential equations; and obtain the solutions by means of Laplace transforms. The explicit expression for the Laplace transform of each of the state probabilities is obtained by the Gauss Jordan method of linear algebra. Explicit results were obtained for the time dependent state probabilities. The Laplace transforms, $f_i(s)$, are rational fractions $A(S)/B(S)$ with one pole at the origin, and usual methods of inversion can be used.

12. Some Percentage Points of the Non-Central t-Distribution. ERNEST M. SCHEUER, Rand Corporation.

Percentage points x such that $P\{t/f^{\frac{1}{2}} > x\} = \epsilon$ where t has the non-central t-distribution with f degrees of freedom and non-centrality parameter $(f + 1)^{\frac{1}{2}}K_p$ (K_p is the standard normal deviate exceeded with probability p) are tabulated for $\epsilon = .975, .025$; $p = .2500, .1500, .1000, .0650, .0400, .0250, .0100, .0040, .0025, .0010$; $f = 2\ (1)\ 24\ (5)\ 49$. This supplements percentage points tabulated by Resnikoff and Lieberman in *Tables of the Non-Central t-Distribution,* Stanford University Press (1957).

13. On the Non-Existence of Some Classes of P.B.I.B. Based on Triangular Schemes. ESTHER SEIDEN, Michigan State University.

Consider unsymmetrical P.B.I.B. based on triangular schemes with the following parameters: $v = n(n - 1)/2$; $b = (n - 1)(n - 2)/2$; $r = n - 2$; $\lambda_1 = 1$, $\lambda_2 = 2$. It is shown that a necessary condition for the existence of such designs is the existence of symmetrical P.B.I.B. with the parameters $v = b = (n - 1)(n - 2)/2$; $r = k = n - 2$; $\lambda_1 = 1$, $\lambda_2 = 2$. Furthermore it is shown that there do not exist the above symmetrical P.B.I.B. based on triangular schemes in the following cases:

(1) $n = 4k + 1$, $(4k - 1, 2)_p(4k - 1, -1)_p = -1$, k an odd integer;

(2) $n = 4k + 1$, $(4k - 1, 2)_p = -1$, k an even integer;

(3) $n = 4(k + 1)$, $(4k + 2, 2)_p = -1$, k an odd integer;

(4) $n = 4(k + 1)$, $(4k + 2, 2)_p(-1, 4k + 2)_p = -1$, k an even integer;

where the expressions of the form $(\alpha, \beta)_p$ denote the Hilbert symbols.

14. Optimum Estimators of the Parameters of Exponential Distributions from One or Two Order Statistics. M. M. SIDDIQUI, National Bureau of Standards, Boulder, Colorado.

Let $f_1(x) = \sigma^{-1} \exp(-x/\sigma)$, $x \geqq 0$; $f_2(x) = \sigma^{-1} \exp[-(x - \alpha)/\sigma]$, $x \geqq \alpha$; and zero otherwise. Let x_k denote the kth order statistic of a random sample of size n. Harter (these *Annals* **32** (1961) 1078–1090) discusses the following three problems: (1) Best unbiased estimator of the form c_kx_k for σ of $f_1(x)$. (2) Best unbiased estimator of the form $c_lx_l + c_mx_m$ for σ of $f_1(x)$. (3) Best unbiased estimators of the form $c_lx_l + c_mx_m$ for σ, α and the mean, μ, of $f_2(x)$. For the problem (3) he shows that the optimum l is equal to 1 and that the same m is optimum for all three parameters. He determines k, l and m for n up to 100 by exhaustive numerical computations. In the present paper, using Euler-Maclaurin formula to approximate the efficiency of an estimator, the following solutions are obtained. For problem (1): k is the nearest integer to $0.79681(n + 1) - 0.39841 + 1.16312(n + 1)^{-1} + \cdots$; for problem (3); m is the nearest integer to $0.79681n + 0.60159 + 1.16312n^{-1} + \cdots$. It is found that a three term approximation is accurate enough to determine the exact optimum in these two cases. Only on very rare occasions, when the fractional part is near 0.5, we need a fourth term. For problem (3), only one term approximation for l and m, namely $l \cong 0.639(n + 1)$, $m = 0.927(n + 1)$ could be obtained.

15. A Note on Inflated Poisson Distribution. S. N. SINGH, Pennsylvania State University. (By title)

In many cases, where the number exposed to a certain risk follows Poisson distribution the zero cell is inflated due to the presence of those not exposed to the risk. In such cases truncated Poisson is used to estimate the parameter λ. Cohen, *Biometrics* **16** 203–211 has extended the truncated Poisson which takes account of zero cell. It is easy to find the M.L. estimates $\hat{\alpha}$ and $\hat{\lambda}$ of α, the proportion exposed, and λ from his distribution by a simple substitution, but the distribution of $\hat{\alpha}$ seems involved from his formulation. In this note it is shown that with inflated Poisson distribution $P(X = 0) = 1 - \alpha + \alpha \exp(-\lambda)$, $P(X = k) = \alpha \exp(-\lambda)\lambda^k/k!$ for $k = 1, 2, \cdots, \infty$, $0 < \alpha < 1$, $0 < \lambda < \infty$; the M.L. estimates $\hat{\alpha}$ and $\hat{\lambda}$ are given by $\hat{\alpha} = (n - n_0)/n[1 - \exp(-\hat{\lambda})]$, $\hat{\alpha}\hat{\lambda} = \bar{x}$, with n_0 the number in zero cell. The asymptotic distributions of $\sqrt{n}(\hat{\alpha} - \alpha)$ and $\sqrt{n}(\hat{\lambda} - \lambda)$ are $N(0, \sigma_1)$ and $N(0, \sigma_2)$ with

$$\sigma_1^2 = \alpha[(1 - \alpha)(1 - \lambda e^{-\lambda}) + \alpha e^{-\lambda}]/(1 - e^{-\lambda} - \lambda e^{-\lambda}),$$

$$\sigma_2^2 = \lambda(1 - e^{-\lambda})/\alpha(1 - e^{-\lambda} - \lambda e^{-\lambda}).$$

16. Absolute Continuity of Infinitely Divisible Distributions. HOWARD G. TUCKER, University of California, Riverside.

Theorem. Let F be an infinitely divisible distribution function determined by the constant γ and the bounded non-decreasing function G in the Lévy-Khinchin representation. Then F is absolutely continuous if at least one of the following two conditions is satisfied: (i) G is not continuous at 0, or (ii) $\int_{-\infty}^{\infty} (1/x^2)\, dG(x) = \infty$. An example is given to show that the condition given above is not a necessary one for absolute continuity of F.

17. Aids for the Separate Maximum Likelihood Estimation of Scale or Shape Parameters of the Gamma Distribution Using Order Statistics. M. B. WILK, R. GNANADESIKAN, MISS M. J. HUYETT, Bell Telephone Laboratories.

Tables and graphs are given to facilitate the computation of the separate maximum likelihood estimation of the scale or the shape parameter of the gamma distribution, based on the M smallest observations in a random sample of size $K(M \leqq K)$. The statistical use of the tables and graphs is illustrated. Some possible applications of these estimation procedures are discussed.

(Abstracts not connected with any meeting of the Institute.)

1. Two-Sample Rank Tests: (A) Efficient against Double-Exponential Alternatives; (B) Efficiency-Robust. ALLAN BIRNBAUM, New York University.

In the notation of E. Lehmann's *Testing Statistical Hypotheses* (Wiley, 1959), pp. 236–240: (A) For the double-exponential translation alternatives $f(y - \Delta) = \frac{1}{2} \exp - |y - \Delta|$, when Δ is sufficiently small the most powerful two-sample rank tests are Reject if $\sum_{i=1}^{n} B(s_i, N) >$ const., where $B(s_i, N) = \sum_{j=s_i}^{N} \binom{N}{j}$ and $\binom{N}{j} = N!/j!(N - j)!$ (B) Let $f_0(y - \Delta)$ and $f_1(y - \Delta)$ be specified translation alternatives (e.g., normal and double-exponential). For $j = 0, 1$, let $t_j = t_j(s) = \sum_{i=1}^{n} h_j(s_i)$ be the rank-test statistic giving most powerful tests against alternative f_j when Δ is sufficiently small. For any rank-test t of fixed size α, let $\beta_j'(t)$ denote the slope of the power function for alternative f_j at $\Delta = 0$, $j = 0, 1$. For each g, $0 \leqq g \leqq 1$, let $t_g = gt_1 + (1 - g)t_0$. Then among all rank tests t of size α, t_g uniquely maximizes the quantity $g\beta_1'(t) + (1 - g)\beta_0'(t)$, for each respective g. It follows that for sufficiently small Δ, no rank test, other than t_g, of size α, is equally or more powerful against both alternatives f_0 and f_1. The tests t_g may be called *efficiency-robust* in this sense, with respect to alternatives f_0, f_1. Extension to larger classes of alternatives is immediate. The proofs are adaptations of the method described by Lehmann.

2. **On the Product of Two Independent Beta Random Variables.** B. Ramachandran, The Catholic University of America.

Let X and Y be two independent beta random variables distributed respectively as $B(a, b)$ and $B(c, d)$. In multivariate statistical analysis, we occasionally have to use the fact that if one or the other of the conditions below is satisfied, then the product $Z = XY$ is also a beta random variable. The conditions are (i) $c = a + b$, in which case $Z \sim B(a, b + d)$, (ii) $a = c + d$, in which case $Z \sim B(c, b + d)$. It is possible to establish a (partial) converse result also; namely, that if the product Z is a beta random variable, say $B(e, f)$,—X and Y being beta random variables as before—then one or the other of the above two conditions must be satisfied. This fact is easily proved on noting that for all $t \geqq 0$, $E(X^t) \cdot E(Y^t) = E(Z^t)$ which leads to the relation (valid for all $t \geqq 0$)

$$[(a + b + t)/(a + t)][(c + d + t)/(c + t)] = (e + f + t)/(e + t)$$

which, in turn, implies that (1) $f = b + d$, and (2) either $e = a$ or $e = c$. $e = a$ corresponds to case (i) above, and $e = c$ to case (ii).

3. **Strong Converse of the Coding Theorem for Indecomposable Channels.** J. Wolfowitz, Cornell University.

Indecomposable channels were introduced by Blackwell, Breiman, and Thomasian (*Ann. Math. Statist.* **29** No. 4, 1958, 1209–1220) who proved a coding theorem and weak converse. (For another proof see A: J. Wolfowitz, Coding theorems of information theory, 1961, Prentice-Hall, Englewood Cliffs, N. J., Section 6.6.) The author now proves the strong converse of the coding theorem, which establishes the capacity of the channel. He also gives a theorem on approximation of the capacity which enables one, at least in principle, to compute the capacity to any desired degree of approximation in a bounded number of arithmetical operations.

The same method enables one to weaken considerably the conditions under which the results of A, Section 6.5, were proved. For example, these results hold under the condition that all the matrices are scrambling matrices (Hajnal (1958). *Proc. Cambridge Philos. Soc.* **54** Part 2, 233–246).