ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Stanford Annual Meeting of the Institute, August 23–26, 1960.)

39. On the Number of Distinct Values in a Large Sample from an Infinite, Discrete Distribution. R. R. Bahadur, Indian Statistical Institute (By title).

Let $A_1, A_2, \cdots$ be an infinite sequence of events defined on the sample space of some experiment such that $P(A_j) > 0$ for each j, $P(A_jA_k) = 0$ for $j \neq k$, and $\sum_j P(A_j) = 1$. Consider a sequence of independent repetitions of the experiment, and let T_n be the number of distinct events A_j observed in the first n trials. The paper studies the rate at which $T_n \to \infty$ as $n \to \infty$. Let $\mu_n = E(T_n)$, where E denotes expected value. *Theorem* 1: $T_n/\mu_n \to 1$ in probability as $n \to \infty$. Suppose (with no loss in generality) that $P(A_j) \geqq P(A_{j+1})$ for all j. Let $f(x) = \max \{j : P(A_j) \geqq x\}$ for $x \leqq P(A_1)$ and $f(x) = 0$ (say) otherwise. *Theorem* 2: $\mu_n = n\int_0^\infty e^{-nx} f(x)\, dx + o(1)$ as $n \to \infty$. It follows, e.g., that if $P(A_j) = cj^{-\alpha}$, where $1 < \alpha < \infty$, then $\mu_n = \Gamma(1 - \beta)(cn)^\beta - \theta_n + o(1)$, where $0 \leqq \theta_n \leqq 1$ and $\beta = 1/\alpha$; and that if $P(A_j) = e^{-\lambda}\lambda^j/j!$, where $0 < \lambda < \infty$, then $\mu_n \sim \log n/\log\log n$. There is, however, no attainable maximum or minimum rate of increase of μ_n. *Theorem* 3: Given P, there exist probability distributions P^* and P^{**} such that, with $\mu_n^* = E(T_n \mid P^*)$ and $\mu_n^{**} = E(T_n \mid P_n^{**})$, $\mu_n^* = o(\mu_n)$ and $\mu_n = o(\mu_n^{**})$ as $n \to \infty$.

40. Expansions for Convolutions. Reed Dawson, American Systems Inc.

Asymptotic expansions for the ordinate and tail area in the distribution of the standardized sum of a large number of independent and identically distributed random variables are developed from the Edgeworth series of Cramér. The formula for the ordinate extends results of Daniels (*Ann. Math. Stat.*, Vol. 25 (1954), pp. 631–650) and Good (*Ann. Math. Stat.*, Vol. 28 (1957), pp. 861–881). The expansion of the tail area is believed to be new.

41. On Sufficient Conditions for Consistent Parameter-Estimates in a Stochastic Difference Equation with Regression on Several Lagged and Non-Stochastic Variables. Friedhelm Eicker, University of North Carolina. (By title).

The least squares estimates a_i of α_i and b_i of β_i in the stochastic difference equation $y_t = \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \beta_1 x_{1t} + \cdots + \beta_q x_{qt} + \epsilon_t$, $t = 1, 2, \cdots$, where $y_0, y_{-1}, \cdots, y_{-p+1}$ and all x_{it} are given constants, are shown to be consistent if conditions (A)–(C2) hold: (A) The disturbances ϵ_t are independent with 0 means, and 2nd and 4th moments bounded between two positive constants uniformly in t. (B) The roots of $\rho^p - \alpha_1\rho^{p-1} - \cdots - \alpha_p = 0$ are within the unit-circle (stationarity). With $X(N, q) = (x_{jt})$ and $\lambda_1(P) =$ largest eigenvalue of $P = X'X$ it is assumed further: (C1) If the set $\{N\}$ of those N for which $N^{-2}\lambda_1(P) \geqq g(N)$, with some $g(N) \to 0$ as $N \to \infty$, is infinite, let

$$(N + N[\lambda_1(P)]^{\frac{1}{2}} + \lambda_1(P))^{-\frac{1}{2}} \cdot \lambda_{\min}(K'K) \to \infty, \qquad N \text{ in } \{N\},$$

where $K(N, (p + 1)q) = (X, LX, \cdots, L^pX)$, $L(N, N)$ containing 1's in its left subdiagonal and 0 elsewhere. (C2) If the complement $\{\bar{N}\}$ to $\{N\}$ is infinite, $\lambda_{\min}(P) \to \infty$ suffices for $a_i \to \alpha_i$ i.p. on $\{\bar{N}\}$. For $b_i \to \beta_i$ on $\{\bar{N}\}$ suffices to have in addition that the elements of $P^{-1}X'L^jX$ are uniformly bounded for all j. This theorem is proved using only elementary matrix theory. Simplified conditions are easily derived. They allow even exponentially increasing regression vectors. This case is not included in what seem to be the broadest conditions known so far (Grenander, *Ann. Math. Stat.*, 1954). The methods applied lead also to statements about asymptotic distributions, the explosive case and other questions.

42. **Multivariate Extremal Distributions.** E. J. GUMBEL, Columbia University. (By title).

A bivariate probability function $F(x, y)$ with margins $F_1(x)$ and $F_2(y)$ is obtained by writing $[-\log F\ (x, y)]^m$ equal to the sum of the corresponding expressions for $F_1(x)$ and $F_2(y)$ where $m \geqq 1$. It is shown that the asymptotic bivariate probability function of largest values $\Phi(x, y)$ taken from $F(x, y)$ has the same form, provided that the marginal distributions possess asymptotic extremal distributions. By analogy, a bivariate probability function $\Pi(x, y)$ for values exceeding x and y is linked to $\Phi(x, y)$ by $\Pi(x, y) = \Phi(-x, -y)$. These expressions can easily be generalized to n dimensions.

43. **Tolerance Regions.** IRWIN GUTTMAN, McGill University. (Invited Paper).

A discussion of Distribution-Free and β-expectation tolerance regions for fixed sample cases and sequential sampling schemes is given. Let $X_1, \cdots, X_n$ be a sample of n independent observations on X. *Definition.* $S(X_1, \cdots, X_n)$ is a distribution-free tolerance region if the induced probability distribution of the coverage of S is independent of the probability distribution of X. *Definition.* $S(X_1, \cdots, X_n)$ is a β-expectation tolerance region if the expected value of its coverage is β.

A connection between tolerance regions and the concept of Best Population (c.f., the work of Gupta, et al.) is indicated. *Definition.* Suppose there are k populations that are distributed by $P_x^{\theta_i}$, $i = 1, \cdots, k$ respectively. Let $\mathcal{C}_j = \int_A dP_x^{\theta_j}$, where A is fixed, and known in advance. Then the k populations are said to contain a best population if and only if there exists an ordering of the $\mathcal{C}_j$ such that $\mathcal{C}_{[k]} > \mathcal{C}_{[k-1]} \geqq \cdots \geqq \mathcal{C}_{[1]}$.

44. **Estimation of the Scale Parameter in the Weibull Distribution by Means of a Life Test with Censoring both by Time and by Number of Failures.** EUGENE H. LEHMAN, JR., North Carolina State College. (Introduced by R. L. Anderson).

The distribution of life spans of certain classes of individuals is assumed to be Weibull with two parameters—a shape parameter assumed known and a scale parameter to be estimated. The maximum likelihood estimator of the scale parameter is derived in a test in which N items are subjected to test. The test continues until R (less than N) items have failed *and* a minimum time T has elapsed. The bias, small sample variance, mean square error, asymptotic mean square error, cost and price of the estimator are determined, where cost is defined as a linear combination of N and the expected duration of the test, and price is the product of cost and mean square error. By means of calculations on an IBM 650, actual values of these characteristics are determined for certain values of R, N, T and the shape parameter.

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

1. **On the Foundations of Statistical Inference, II.** (Preliminary Report). ALLAN BIRNBAUM, New York University.

Let $E = (p_{ij})$ be any stochastic matrix ($\sum_j p_{ij} = 1$, for each i), $i = 1, \cdots k$ (k finite), $j = 1, \cdots m$ (possibly infinite). Then E is the mathematical model of a statistical experiment for k simple hypotheses: Prob $(X = j \mid H_i) = p_{ij}$. For any fixed k, E is *simple* if it is (or is equivalent to) such a matrix with $m = k$. A simple experiment is *cyclic-symmetric* (c.s.) if it can be represented by a c.s. square matrix ($p_{i,j} = p_{i-1,j-1}$, with any subscript 0 replaced by k). Any experiment is called c.s. if it has a representation $E = (p_{ij}) = (Q_1, Q_2, \cdots)$ where each Q_h is square c.s. *Lemma* 1: Each c.s. experiment is equivalent to a mixture of c.s. simple experiments. *Lemma* 2: Every experiment is a component of some c.s. experiment. Hence for typical purposes of informative inference, any outcome $X = j$ of any experiment $E = (p_{ij})$ can and should be interpreted as an outcome of the essentially unique simple c.s. experiment having a column proportional to the jth column of E. The structure of E is irrelevant to such interpretations except through its jth column, which is the likelihood function determined by outcome j. Such interpretations can be expressed exclusively, if desired, in terms of error-probabilities defined in the simple c.s. experiment and admitting frequency interpretations; such interpretations include point and confidence-set estimates. A formal correspondence exists between some such interpretations and inferences based on formally postulating equal prior probabilities; this gives a constructive explication of the traditional Bayesian "principle of indifference." A formal correspondence exists also between such interpretations in confidence-set form and the statements obtainable by formal application of Fisher's "fiducial argument" (which is possible in any simple c.s. experiment).

2. **Trees and Negative Estimates of Variance Components.** W. A. THOMPSON, JR., University of Delaware. (By title)

This paper provides an algorithm for solving the problem of negative estimates of variance components (see Abstract 69, *Ann. Math. Stat.*, 1960) for all random effects models whose expected mean square column may be thought of as forming a mathematical tree in a certain sense. The algorithm is as follows. Consider the minimum mean square in the entire array; if this mean square is the root of the tree than equate it to its expectation. If the minimum mean square is not the root then pool it with its predecessor. In either case the problem is reduced to an identical one having one fewer variable and hence in a finite number of steps the process will yield estimates of the variance components. These estimates are non-negative and have a maximum likelihood property.