ABSTRACTS OF PAPERS

Presented on August 13, 1944, at the Wellesley meeting of the Institute

1. **Statistical Tests Based on Permutations of the Observations.** A. WALD and J. WOLFOWITZ, Columbia University.

It was pointed out by Fisher that statistical tests of exact size, based on permutations of the observations, can be carried out without assuming anything about the underlying distributions except their continuity. Scheffé has proved that, for an important class of hypotheses, these tests are the only ones with regions of exact size. Texts based on permutations of the observations have been constructed by Fisher, Pitman, Welch, and the present authors. In the present paper, the authors prove a theorem on the limiting normality of the distribution, in the universe of permutations, of a class of linear forms. Application of this theorem gives the limiting normality (in the universe of permutations, of course) of the correlation coefficient, and of a statistic introduced by Pitman to test the difference between two means. The limiting distribution of the analysis of variance statistic in the universe of permutations is also obtained.

2. **Error Control in Matrix Calculation.** FRANKLIN E. SATTERTHWAITE, Aetna Life Insurance Co.

The arithmetic evaluation of matrix expressions is often rather complicated. One of the causes of this is the fact that relatively minor errors (such as rounding errors) introduced in an early step may be magnified to such an extent in succeeding steps that the final result is useless. Iterative methods to meet this difficulty have been reviewed very completely by Hotelling. In this paper a different approach is taken. Conditions on the norm of a matrix are determined so that a Doolittle process will not magnify errors to more than two or three decimal places. It is then pointed out that if an approximation to the inverse of the matrix is available, most problems can be rearranged so that the required norm conditions are met. A Doolittle process may then be used to any number of decimal places with assurance that errors will not accumulate to more than a limited number of decimal places.

3. **On Cumulative Sums of Random Variables.** A. WALD, Columbia University.

Let $\{z_i\}$ $(i = 1, 2, \cdots \text{ad inf.})$ be a sequence of independent random variables each having the same distribution. Denote by Z_j the sum of the first j elements of the sequence. Let $a > 0$ and $b < 0$ be two constants and denote by n the smallest integer for which either $Z_n \geq a$ or $Z_n \leq b$. Neglecting the quantity by which Z_n may differ from a or b (this can be done if the mean value of $|z_i|$ is small), the probability that $Z_n \geq c$ for $c = a$ and $c = b$ is derived, and the characteristic function of n is obtained. The probability distribution of n when z_i is normally distributed is derived. These results have application to various statistical problems and to problems in molecular physics dealing with the random walk of particles in the presence of absorbing barriers.

4. **The Approximate Distribution of the Mean and of the Variance of Independent Variates.** P. L. HSU, National University of Peking.

Let X_k be mutually independent random variables with the same cumulative distribution function; let $E(X_k) = 0$, $E(X_k^2) = 1$ and $E(X_k^4) = \delta$. Finally put $S = n^{-1} \sum_{k=1}^{n} X_k$ and $\eta = n^{-1} \sum_{k=1}^{n} (X_k - S)^2$. The author first gives a new derivation of H. Cramer's well-known asymptotic expansions for $Pr\ (n^{\frac{1}{2}}S \leq x)$. The proof is much more elementary and avoids in particular the use of M. Riesz' singular integrals. Instead a considerably simpler Cesaro-type kernel is used, which has first been introduced by A. C. Berry (*Trans. Amer. Math. Soc.* 49 (1941), pp. 122–136). The same method is then used to derive similar asymptotic expansions for $Pr\ (n^{\frac{1}{2}}(\eta - 1) \leq (\delta - 1)^{\frac{1}{2}}x)$. The method can be extended to the case of unequal components and also for the study of other functions encountered in mathematical statistics.

5. **Ranges and Midranges.** E. J. GUMBEL, New School for Social Research.

The mth range w_m and the mth midrange v_m are defined as the difference and as the sum of the mth extreme value taken in descending magnitude ("from above") and the mth extreme value taken in ascending magnitude ("from below"). The semi-invariant generating functions $L_m(t)$ and ${}_mL(t)$ of the mth extreme values from above and below are simple generalizations of the semi-invariant generating functions of the largest and of the smallest value which have been given by R. A. Fisher and L. H. C. Tippett. If the sample size is large enough the two mth extreme values may be considered as independent variates. Then, the semi invariant generating functions $L_w(t, m)$ and $L_v(t, m)$ of the mth range and of the mth midrange are

$$L_w(t, m) = L_m(t) + {}_mL(-t);\ L_v(t, m) = L_m(t) + {}_mL(t).$$

If the initial distribution is symmetrical the semi invariant generating function of the mth range is twice the semi invariant generating function of the mth extreme value from above. The distribution of the mth range is skew, whereas the distribution of the mth midrange is of the generalized, symmetrical, logistic type. The even semi invariants of the mth midrange are equal to the even semi invariants of the mth range. For increasing indices m the distributions of the mth extremes, of the mth ranges and of the mth midranges converge toward normality.

6. **Statistics of Sensitivity Data, II.** Preliminary report. C. W. CHURCHMAN, Frankford Arsenal, and BENJAMIN EPSTEIN, Westinghouse Electric and Manufacturing Co.

In this paper a study is made of the distribution of the first two moments of sensitivity data as functions of the sample size. The chief results are briefly these:

(a) The distributions of $\bar{x}$ and $\sigma_{\bar{x}}^2$ (for definition of these functions, see "On the Statistics of Sensitivity Data," by the authors in the *Annals of Mathematical Statistics*, Vol. XV, No. 1) approach normality rapidly as functions of the sample size;

(b) $\bar{x}$ and $\sigma_{\bar{x}}^2$ are "almost" independent even for small sample sizes, thus justifying the use of Student's ratio in texts of significance for differences between two sample means.