ABSTRACTS OF PAPERS

(Presented on September 8, 1942, at the Poughkeepsie meeting of the Institute)

On the Theory of Testing Composite Hypotheses with One Constraint. HENRY SCHEFFÉ, Princeton University.

A composite hypothesis with one constraint specifies the value of one and only one parameter of a set occurring in a distribution function. The theory of testing such hypothesis is not only of direct interest for many important problems, but is intimately related to Neyman's theory of confidence intervals (*Phil. Trans. Roy. Soc. London*, 1937). A method of Neyman (*Bull. Soc. Math. France*, 1935) for finding type B regions for testing these hypotheses is extended to the case of any number of nuisance parameters. Type B_1 regions are defined by generalizing the type A_1 regions of Neyman and Pearson (*Stat. Res. Mem.*, 1936) to the case where nuisance parameters are present, and sufficient conditions are found that a type B region be also of type B_1. An interesting moment problem is encountered, in which the admissible functions are not of constant sign, and is solved for the case where the original distribution is multivariate normal.

On the Consistency of a Class of Non-Parametric Statistics. J. WOLFOWITZ, N. Y. City.

Let X and Y be two stochastic variables about whose distribution nothing is known except that they are continuous and let it be required to test whether their distribution functions are the same. Let V be the observed sequence of zeros and ones constructed as described elsewhere (Wald and Wolfowitz, *Annals of Math. Stat.*, Vol. 11 (1940), p. 148). Suppose that the statistic $S(V)$ used to test the hypothesis is of the form $S(V) = \Sigma\varphi(l_j)$, where l_j is the length of the j-th run and $\varphi(x)$ a suitable function defined for all positive integral x. The notion of consistency, originated by Fisher for parametric problems, has already been extended to the non-parametric case (loc. cit., p. 153). The author now proves that, subject to reasonable conditions on $\varphi(x)$ and statistically unimportant restrictions on the alternatives to the null hypothesis, statistics of the type $S(V)$ are consistent. In particular, a statistic discussed by the author (*Annals of Math. Stat.* September, 1942) and for which $\varphi(x) \equiv \log \left(\frac{x^x}{x\,!}\right)$ belongs to the class covered by the theorem.

Graphical Controls Based on Serial Numbers. E. J. GUMBEL, New School for Social Research.

The index m of the observed value x_m ($m = 1, 2, \cdots n$) is called its serial number. A value x of a continuous statistical variable defined by a probability $W(x) = \lambda$ is called a grade (e.g. the median for $\lambda = \frac{1}{2}$). The coordination of serial numbers with grades furnishes two graphical methods for comparing the observations and the theory, namely the equiprobability test based on $m = n\lambda$, and the return periods based on $m = n\lambda + \frac{1}{2}$.

Starting from the distribution of the mth value, we determine the most probable serial number $\tilde{m} = n\lambda + \Delta$, where Δ depends upon the distribution. For a symmetrical distribution, the corrections Δ for two grades defined by λ and $1 - \lambda$, are equal in absolute value and opposite in sign. Then no correction is needed for the median. For an asymmetrical distribution, we calculate the most probable serial number of the mode considered as an mth value. Thus the mode is obtained from the observations through the theory. In this case the mode is not the most precise mth value.

If m is of the order $\frac{1}{2}n$, the distribution of the mth value converges towards a normal distribution with an expectation given by $m = nW(x)$, and a standard deviation $s(x)$, where $s(x)\sqrt{n} = \sqrt{W(x)(1 - W(x))/W(x)}$. By attributing to eachtheoretical value x its standard deviation, we obtain intervals $x \pm s(x)$ which may be used for the control of the equiprobability test, the comparison of the observed step function with the frequency, and the comparison of the observed with the theoretical return periods. Besides, the standard error of the mth value leads to the precision of the determination of a constant obtained from a grade.

Significance Tests for Multivariate Distributions. D. S. VILLARS, U. S. Rubber Company.

The observed mean of sets of m variates, each normally and independently distributed, is distributed around the population mean according to a χ^2 distribution with m degrees of freedom. The sum of squares of deviations of n observed points from the observed mean is distributed as χ^2 with $m(n - 1)$ degrees of freedom (not with $n - 1$). A much more powerful test for correlation than that by the correlation coefficient is described, which for bivariate distributions, involves comparisons between $n - 1$ and $n - 1$ degrees of freedom. This can be extended to $m - 1$ tests with m variates. Distribution of distance between two means and distribution of fiducial radius is worked out in detail for two variates.

On the Choice of the Number of Class Intervals in the Application of the Chi-Square Test. H. B. MANN and A. WALD, Columbia University.

The distance of two distribution functions is defined as the l.u.b. of the absolute value of the difference between the two cumulative distribution functions. Let $C(\Delta)$ be the class of alternatives with distance $>\Delta$ from the null-hypothesis. Let $f(N, k, \Delta)$ be the g.l.b. of the power with respect to alternatives in $C(\Delta)$ of the chi-square test with sample size N and k equally probable class intervals. A positive integer k is called best with respect to sample, size N if there exists a Δ such that $f(N, k, \Delta) = \frac{1}{2}$ and $f(N, k', \Delta) \leqslant \frac{1}{2}$ for every positive integer k'. The authors show that $k_N = \sqrt[5]{\frac{2(N-1)^2}{C^x}}$ where $\frac{1}{\sqrt{2\pi}} \int_c^\infty e^{-\frac{1}{2}x^2}\, dx$ is equal to the size of the critical region, fulfills approximately the conditions of a best k with $\Delta_N = \frac{5}{k_N} - \frac{4}{k_N^2}$ as the correspondingvalue of Δ. The approximation is shown to be satisfactory for $N \geq 450$ if the 5% level of significance is used and for $N \geq 300$ if the 1% level is used.

Generalized Poisson Distribution. F. E. SATTERTHWAITE, Aetna Life Insurance Company.

In this paper the Poisson distribution is generalized to allow for the assignment of varying weights to a set of events when the number of events follows the Poisson law. The development used brings out the fact that distributions falling in this class do not require that the underlying statistics be homogeneous. The only requirement is that they be independent. Formulas are given for the moments of the generalized distribution as functions of the moments of the underlying distribution of weights. The principles to be observed in the solution of practical problems are outlined.

The Relationship of Fisher's z Distribution to Student's t Distribution. LEO A. AROIAN, Hunter College.

For n_1 and n_2 sufficiently large $W = \frac{1}{\beta}\sqrt{\frac{N}{N+1}}\, z$ is distributed as Student's t with N degrees of freedom, $N = n_1 + n_2 - 1$, $\beta^2 = \frac{1}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$. If the level of significance is α for Student's distribution, the level of significance for z will be $\sqrt{\frac{N}{N+1}}\, \alpha e^{\frac{\beta^2}{3} - \frac{5}{12N}} < \alpha$. As a corollary it follows that the distribution of z approaches normality, $n_1, n_2 \to \infty$, with mean zero and variance $\frac{1}{2}\left(\frac{1}{n_2} + \frac{1}{n_1}\right)$. This simplifies a previous proof of the author. Application of this result is made to finding levels of significance of the z distribution. On the whole R. A. Fisher's formulas for this purpose, n_1 and n_2 large, as modified by W. G. Cochran are superior. The results given by the Fischer-Cochran formulas are compared with those obtained by using the formula recently found by E. Paulson.

On a Statistical Problem Arising in the Classification of an Individual in One of Two Groups. Abraham Wald, Columbia University.

Let π_1 and π_2 be two p-variate normal populations which have a common covariance matrix. A sample of size N_i is drawn from the population $\pi_i (i = 1, 2)$. Denote by $x_{i\alpha}$ the α-th observation on the ith variate in π_1, and by $y_{i\beta}$ the βth observation on the ith variate in π_2. Let $z_i (i = 1, \cdots, p)$ be a single observation on the ith variate drawn from a population π where it is known that π is equal either to π_1 or to π_2. The parameters of the populations π_1 and π_2 are assumed to be unknown. It is shown that for testing the hypothesis $\pi = \pi_1$ a proper critical region is given by $U \geq d$ where $U = \Sigma\Sigma s^{ij} z_i(\bar{y}_j - \bar{x}_j)$, $\| s^{ij} \| = \| s_{ij} \|^{-1}$, $s_{ij} = \left[\sum_\alpha J(x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j) + \sum_\beta J(y_{i\beta} - \bar{y}_i)(y_{j\beta} - \bar{y}_j)\right]/(N_1 + N_2 - 2)$, $\bar{x}_i = \left(\sum_\alpha J x_{i\alpha}\right)/N_1$, $\bar{y}_i = \left(\sum_\beta J y_{i\beta}\right)/N_2$ and d is a constant. The large sample distribution of U is derived and it is shown that U is a simple function of three angles in the sample space whose exact joint sampling distribution is derived.

Modern Statistical Methods in Penology. Saly R. R. Struik, Radcliffe College and Miriam van Waters, Massachusetts Reformatory for Women.

In applying statistical methods to penological problems, so far the best known studies have considered 100, 500, or once in England (to refute Lombroso's theory) 1500 cases. But from the correct statistical standpoint, far more cases are needed to establish a law. Over a period of years, an attempt has been made to use statistical methods in the study of penological problems in the Massachusetts Reformatory for Women, but the results will take on real significance and be conclusive only when similar investigations are made all over the United States.

Regularity of Label-Sequences Under Configuration Transformations. T. N. E. Greville, Bureau of the Census.

There is developed a class of transformations on sequences of arbitrary labels in terms of which a wide variety of problems in the theory of probability can be formulated. It is shown that, with mild restrictions on the transformations used and on the measure function assumed on the label-space, almost every label-sequence produces a transform having the frequency distribution expected. The class of transformations considered is shown to include as special cases the four fundamental operations of von Mises: place selection, partition, mixing, and combination.

On the Ratio of the Variances of Two Normal Populations. HENRY SCHEFFÉ, Princeton University.

Let θ be the above ratio. The two problems considered in this paper are the formulation and comparison of (i) significance tests for the hypothesis $\theta = \theta_0$, and (ii) confidence intervals for θ. The paper is divided into two parts; the first is kept on an elementary level and only solutions based on the F-distribution are considered. Following various approaches, six tests and corresponding sets of confidence intervals are introduced. It turns out that the limits on the F-distribution which yield an unbiased test are the same as those which yield confidence intervals optimum in a certain intuitive sense. The values of these limits are difficult to compute and some numerical data are given to indicate the loss of efficiency in using instead the easily obtained "equal tails" limits. The second part of the paper is concerned with the existence of common best critical regions and type B_1 regions, and the application of Neyman's theory of confidence intervals. No new tests or confidence intervals not already considered in part I are obtained, but those previously judged best of a very narrow class are now shown to be best of all those based on similar regions of the same size.