ABSTRACTS OF PAPERS

Presented June 17–19, 1947, at the San Diego meeting of the Institute

1. Random Variables with Comparable Peakedness. Z. W. BIRNBAUM, University of Washington.

Let U and V be random variables with symmetrical distributions, i.e. with $P(U \leqq -T) = P(U \geqq T)$ and $P(V \leqq -T) = P(V \geqq T)$ for all $T \geqq 0$. The random variable U shall be called more peaked than V if $P(|U| \geqq T) \leqq P(|V| \geqq T)$ for all $T \geqq 0$. Let X_1, Y_1 and X_2, Y_2 be two pairs of independent random variables such that X_i is more peaked than Y_i for $i = 1, 2$. Then under certain additional conditions $X = X_1 + X_2$ is more peaked than $Y = Y_1 + Y_2$.

2. On Optimum Tests of Composite Hypotheses with One Constraint. ERICH L. LEHMANN, University of California, Berkeley.

The problem studied is that of finding all similar and bisimilar test regions of composite hypotheses, and of obtaining the most powerful of these regions. Various results are obtained for distributions which admit sufficient statistics with respect to their parameters. Applications are made to the hypothesis specifying the value of the circular correlation coefficient in a normal population, and certain hypotheses concerning scale and location parameters in exponential and rectangular populations.

3. Estimation of a Distribution Function by Confidence Limits. FRANK J. MASSEY, JR., University of California, Berkeley.

Let $x_1, x_2, \cdots, x_n$ be the results of n independent observations, having the same cumulative distribution function $F(x)$. Form the function $S_n(x) = k/n$ where k is the number of observations less than or equal to x. A confidence band $S_n(x) \pm \lambda/\sqrt{n}$ will be used to estimate $F(x)$. To determine the confidence coefficient it is necessary to find $Pr\{\max \sqrt{n} \mid S_n(x) - F(x) \mid \leqq \lambda/\sqrt{n}\}$. It is sufficient to consider x uniformly distributed in the interval $(0, 1)$. Let $\lambda\sqrt{n} = s/t$ where s and t are integers. Then $S_n(x)$, to stay in the band $F(x) \pm \lambda/\sqrt{n}$, can only pass through certain lattice points above $x = i/tn$, $i = 1, 2, \cdots, tn$. The probability of $S_n(x)$ passing through a particular sequence of these points is given by the multinomial law, and this can be summed over all permissible sequences. Limiting distributions have been given by A. Kolmogoroff, and by N. Smirnoff. It is desired to test the hypothesis $F(x) = F_0(x)$ against alternatives $F(x) = F_1(x)$. Using the criterion: reject $F_0(x)$ if

$$\max_x \sqrt{n} \mid F_0(x) - S_n(x) \mid > \lambda$$

the probability of first kind of error can be controlled by choice of λ. A lower bound to the probability of second kind of error against alternatives such that $\max \sqrt{n} \mid F_0(x) - F_1(x) \mid \geqq \Delta$ is given. This lower bound approaches one as $n \to \infty$. Thus the test is consistent.

4. A Note on Sequential Confidence Sets. CHARLES STEIN, Columbia University.

This paper generalizes a paper of Stein and Wald, appearing in the *Annals of Math. Stat.*, Sept., 1947.

Let $\{X_i\}$, $(i = 1, 2, \cdots)$, be a sequence of random variables whose distribution depends on an unknown parameter θ. Sequential confidence sets are determined by a rule indicating when to stop sampling and a rule giving the confidence set as a function of the sample. It is desired that, for each sample point, the confidence set should be one of a specified class S, that the probability of covering the true parameter should be $\geqq \alpha$, and that the least upper bound of the expected number of observations should be minimized. If X_i are independent with the rectangular distribution on $(0, \theta)$ and S consists of all intervals of the form $(\theta_0, k\theta_0)$ with k fixed and θ_0 a function of the sample, the optimum sequential procedure is the classical non-sequential procedure. If the X_i are independently and identically distributed in accordance with a multivariate normal distribution with known covariance matrix Σ but unknown mean θ, and the confidence sets are to be of the form $(\theta - \theta_0)' \Sigma^{-1}(\theta - \theta_0) = r$, r fixed, θ_0 a variable p-dimensional vector, a similar result holds, provided the desired confidence coefficient α is not excessively small.

5. Explicit Solution of the Problem of Fitting a Straight Line when Both Variables are Subject to Error for the Case of Unequal Weights. ELIZABETH L. SCOTT, University of California, Berkeley.

Let α, β and ξ_i, $(i = 1, 2, \cdots, s)$, be unknown fixed numbers and let $\eta_i = \alpha + \beta\xi_i$. For each value of i there exist m_i measurements x_{ij} of ξ_i and n_i measurements y_{ik} of η_i, $(j = 1, 2, \cdots, m_i; k = 1, 2, \cdots, n_i)$. The variables x_{ij} and y_{ik} are normally distributed about ξ_i and η_i with variances σ_1^2/u_i and σ_2^2/v_i respectively, where the weights u_i and v_i are known but σ_1^2 and σ_2^2 are unknown. The numbers m_i and n_i are bounded (usually small) while s increases indefinitely. Thus α, β, σ_1^2 and σ_2^2 appear as structural parameters and the ξ_i as incidental parameters. (See paper by J. Neyman and E. L. Scott to appear in *Econometrica.*) Modified maximum likelihood equations (MMLE) yielding consistent estimates of the structural parameters are tedious to solve when the products m_iu_i and n_iv_i depend on i. The main result of this paper consists in proving that the varying m_iu_i and/or n_iv_i can be treated as constants. Let w_1 and w_2 be the harmonic means of m_iu_i and n_iv_i, respectively. Now, MMLE's written with $m_iu_i = w_1$ and $n_iv_i = w_2$ yield consistent estimates of α and β. The asymptotic variances are also found. An application is made to certain problems of astronomy.

6. Unbiased Estimates with Minimum Variance. CHARLES STEIN, Columbia University.

Let X be a random variable distributed in the space R according to one of the p.d.f.'s $\varphi(x \mid \theta)$, where θ is an unknown parameter, and let $g(\theta)$ be a real-valued function of θ. Let $B(\theta)$ be the set of all x such that $\varphi(x \mid \theta) > 0$ but $\varphi(x \mid \theta_0) = 0$, and S the set of all θ such that $B(\theta)$ has probability 0 when θ is the true parameter value. Let

$$\psi(x \mid \theta) = \varphi(x \mid \theta_0)/\varphi(x \mid \theta_0 \text{ and } A(\theta_1, \theta_2) = E\{\psi(X \mid \theta_1)\, \psi(X \mid \theta_2) \mid \theta_0\}$$

for θ_1, θ_2 in S. Suppose $A(\theta_1, \theta_2)$ is everywhere finite and there exists a set function λ of bounded variation over S such that $\int_S A(\theta_1, \theta_2)\, d\lambda(\theta_1) = g(\theta_2)$. Then an estimate of $g(\theta)$, unbiased for all θ in S and having minimum variance at θ_0 is given by $f(x) = \int_S \varphi(x \mid \theta)\, d\lambda(\theta)/\varphi(x \mid \theta_0)$. The minimum variance is $\int_S g(\theta)\, d\lambda(\theta) - [g(\theta_0)]^2$. If the definition of $f(x)$ is modified at a set having probability 0 when $\theta = \theta_0$, the properties on S and at θ_0 remain unchanged. Under mild restrictions this alteration can be carried out so as to make $f(x)$ an unbiased estimate of θ for all S. The results are related to the work of Fisher, Dugué, Rao, and Bhattacharyya on the amount of information.

7. Sufficient Statistics and a System of Partial Differential Equations. (A Contribution to the Neyman-Pearson Theory of Testing Hypotheses.) Preliminary Report. ERICH L. LEHMANN, University of California, Berkeley, AND HENRY SCHEFFÉ, University of California, Los Angeles.

In the Neyman-Pearson theory of testing hypotheses the problem of the existence and determination of similar regions has been treated under two approaches: (1) Assuming the existence of a set of sufficient statistics for the nuisance parameters; (2) Assuming that the probability density satisfies a certain system of partial differential equations. By solving the differential equations it is now shown that they imply the existence of sufficient statistics for the nuisance parameters. Knowledge of the form of the solution of the differential equations permits simplification of the known theory of optimum tests (type B, B_1, etc.) as well as some generalization.

8. Power Function of the Analysis of Variance and Covariance Test of a Normal Bivariate Population. W. M. CHEN, University of California, Berkeley.

The problem of finding the power function of the analysis of variance and covariance test of a normal bivariate population, $\rho = 0$ and $\sigma_1 = \sigma_2$, by means of principle of likelihood was reduced to the determination of the distribution function $P(L)$ of the following moment problem:

$$\int_0^1 L^k \, dP(L) = \frac{(1-a)^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2}\right)} \sum_{r=0}^{\infty} \frac{a^r}{r!} \Gamma\left(\frac{n-1}{2} + r\right) M_{k,r}, \qquad (k = 1, 2, \cdots),$$

where

$$M_{k,r} = \frac{4^k \Gamma(n-1+r)\, \Gamma\left(\frac{n-2+2k}{2}\right) \Gamma\left(\frac{n-1+2k}{2} + r\right)}{\Gamma\left(\frac{n-2}{2}\right) \Gamma\left(\frac{n-1}{2} + r\right) \Gamma(n-1+2k+r)},$$

and a, the argument of the power function, lies in the interval $(0, 1)$ and vanishes only when the hypothesis tested is true. The moment problem was found and solved by rather tricky methods. The result is

$$P(L) = \frac{(1-b)^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2}\right)} \sum_{s=0}^{\infty} \frac{b^s}{s!} \Gamma\left(\frac{n-1}{2} + s\right) I_L\left(\frac{n-2}{2}, s+1\right)$$

where $b = \left(\frac{a}{1-a}\right)^2$.

9. A Mathematical Model of the Relation between White and Yolk Weights of Birds' Eggs. G. A. BAKER, University of California, Davis.

The purpose of such a model is to find a rational method of estimating a "best line" in some sense which will represent the relation between white and yolk weights for some or all species of birds. From data at hand it appears that birds within species may differ in means and variances of weights and that the yolk and white weights are positively correlated. Yolk and white weights within a species are functions of egg number. The standard deviations of yolk and white weights for different species are approximately proportional to mean values. The "true" means for yolk and white weights for different species do not lie on a line because of biological differences between species with the same egg size. The standard deviation of species deviations from a straight line depend on the size of the egg (may be proportional to a weighted sum of the yolk and white weights). If sampling variances are sufficiently small they may be neglected and a straight line fitted assuming both variables subject to error and non-uniform variance. The practicality of maximum likelihood estimates is considered.

10. Statistical Analysis for a New Procedure in Sensitivity Experiments. A. M. MOOD, Iowa State College, AND W. J. DIXON, University of Oregon.

In the language of biological assay the sensitivity experiment investigates the proportion of subjects that respond to a given concentration, x, of a certain chemical. It is assumed that only one test may be made on each subject. The new procedure is characterized by a change in x for each successive test, depending on the result of the preceding test. x is reduced to the next lower of a fixed set of concentrations for the next test if there is no response and is increased to the next higher concentration if there is a response. Observations are thus concentrated near the mean and few tests are made for values of x where a very large or very small proportion of subjects would respond. Assuming x is normally distributed, approximate maximum likelihood estimates are obtained for the mean and standard deviation of x. These assume a form which is simple to compute. Choice of optimum increments of x for various situations is investigated.

11. The Relation of Inbreeding to Calf Mortality. W. M. REGAN, S. W. MEAD, AND P. W. GREGORY, University of California, Davis.

An analysis of calf mortality in the University of California dairy cattle breeding experiment is presented. Calves up to 4 months of age that were born singly are included in the study. Only those stillbirths and abortions from cows free from Brucellosis and health and reproductive abnormalities were considered. A total of 774 Jersey and 258 Holstein calves were included. Calves were classified according to inbreeding coefficients as follows: Class I, the controls 0.0 to 0.1249; Class II, 0.125 to 0.2448; Class III, 0.245 to 0.3749; and Class IV, 0.375 and over. There was no relation between the number of abortions and the degree of inbreeding. The stillbirths, too few to be statistically significant, tend to increase as the coefficient of inbreeding increased. Following birth, however, mortality was correlated with inbreeding of both males and females but for the males it was greater than for the females in Classes III and IV, but the difference is hardly significant. The Jerseys tended to be less viable than the Holsteins. Some of the increased mortality of the more highly inbred animals could be accounted for by the action of two lethal genes; one controlling an anomaly of the liver, the other an anomaly of the heart; there was no plausible explanation for most of it. Within sex, inbreeding class, and breed there was considerable variation in the mortality of the progeny of different sires. Some of these differences were statistically significant.

12. Observations on Designs for Cooperative Field Tests. P. A. MINGES, University of California, Davis.

In California conditions vary so greatly between the principal production areas that it is necessary to establish experimental plots in each of the areas if reliable information is to be obtained regarding cultural practices. Most of these tests must be conducted on ranches in cooperation with growers and local agricultural extension agents. The designs of these tests should be relatively simple, the arrangement should be adjustable to work into the growers' cultural practices and to permit the obtaining of yield records with a minimum of interference to the growers' operations, yet the design must be adequate to yield valid data. The randomized block design has proved the most useful, although paired plots, factorials, split-plots and Latin squares have been used successfully under certain conditions. The Latin square design is useful when a two-way variation is expected, otherwise it is not usually very efficient. Where yield data are of prime importance, for ureplications have been considered most practical. In tests such as variety trials when factors other than yields are important, two replications may be adequate. The size of the plot has been varied to fit the crop, conditions of the field, and known soil variability. Plots two rows wide and 50 to 135 feet long often have been used, frequently without guards between plots. Since it is desirable to include checks (untreated controls) in most tests, small plots will reduce the loss to the growers when the treatments prove beneficial. The information derived from these tests is of most interest to growers and county agents so the data should be presented in tables that are easily read. The variability figure which is confusing to most people probably can best be presented as the least significant difference.

13. Population Genetics. N. H. Horowitz, California Institute of Technology.

Population genetics attempts to describe the effects on the genetical structure of Mendelian populations of factors such as mutation, selection, migration, and random fluctuations due to sampling errors. These diverse elements are brought under a common viewpoint by considering their effects on gene frequencies. Since change in gene frequency is the elementary process of evolution, the above factors are causal agents of evolution. Mathematical models illustrating the interplay of the various elements have been constructed by Wright, Haldane, and Fisher. The nature of Mendelian inheritance is such that gene frequencies remain constant in large populations not subject to net mutation, selection, or migration pressures. Unbalanced pressures initiate evolutionary changes which continue until equilibrium is reached at a new level of gene frequencies. Equilibrium frequencies are determined by opposing pressures—e.g., opposing mutation rates, mutation opposed by selection, etc. Equilibrium, stable or unstable, is also possible under selection alone. In small populations, sampling errors among the gametes produce random fluctuations in gene frequencies which, superimposed on the equilibrium values, result in probable distributions of frequencies. The latter provide a mechanism for the evolution of characters, especially biochemical syntheses, which depend on the simultaneous action of a number of individually non-adaptive genes.

14. The Choice of Inspection Stringency in Acceptance Sampling by Attributes. J. L. Hodges, Jr., University of California, Berkeley.

In acceptance sampling by attributes, the probability p that an item will be defective is taken to be a function $g(x, y)$ of the quality x of the population and the stringency y of inspection. Let n, the number of items inspected, be fixed, and reject if the number of defectives is $\geqq k$. It may then be possible to satisfy a condition on the power function with different values of k, by adjusting y properly. This paper is concerned with the choice of k and y in such situations. A criterion is given, and it is shown that the criterion is approximately satisfied by $k = [ng(x_0, \hat{y})]$ where x_0 separates acceptable and non-acceptable values of x, and $\hat{y}$ maximizes

$$\frac{\partial g(x_0, y)}{\partial x} \bigg/ \sqrt{g(x_0, y)[1 - g(x_0, y)]}.$$

An asymptotic property of this approximation is shown. The method is applied to two examples: (a) testing the mean bacterial density x of a liquid by the dilution method, y being the volume of liquid incubated, and (b) testing the variance x of a normally distributed dimension of known mean m by applying gauges set at $m \pm \frac{1}{y}$. The approximate solution is found to be satisfactory in both cases for $m = 20$.

15. The Application of Learning Curves to Industrial Planning. Preliminary Report. James R. Crawford, Lockheed Aircraft Corporation.

Learning curves are significant factors of analysis in industries producing quantities of less than 20,000 units of a given article. Ship-building and airframe manufacture are the two largest industries in this class. Learning curves occur where job costs are kept either by individual unit or by lot, and also where achievement is measured against a standard. Cost per unit plots against ordinal unit number as a straight line on logarithmic graph-paper. Learning curves are used to supplement time-studies, determine the capacity of tooling, layout of budgets, and for estimating and bidding. The experience of individual workers and management are reflected in these analyses. The slope of the learning curve is related to the amount to be learned. Plateaus occur which are related to the hiring of new workers and to the relaxing of control measures. Other consistent minor patterns occur which are related to specific conditions. Equations have been derived and tables computed for five related forms of the learning curve. Graphic methods are satisfactory except for bidding. This study covers a simple approach to an important problem of industrial management. The findings in the industrial field may benefit research in the field of the psychology of learning.

16. Relative Effects of Inbreeding and Selection in Poultry. W. O. Wilson, University of California, Davis.

Egg production rate, fertility, hatchability, and chick mortality records from the Iowa State College Poultry Department's inbreeding project were studied. Statistics which were calculated from the data included simple and partial regression of traits on inbreeding, estimates of heritability by correlation between paternal half-sibs and by daughter-dam regressions, and selection differentials. The net genetic gain or loss in merit per generation was considered to be the sum of the product of selection differentials and heritability, plus the product of regression of trait on inbreeding and increase in amount of inbreeding. The amount of inbreeding that can be done in each of the traits was estimated when there was no net loss or gain. Of the traits studied, the rank was in the following order: Hatchability, chick mortality, fertility, and egg production.

17. The Rate of Genetic Gain in Egg Production in Progeny-Tested Flocks as a Function of the Interval between Generations. Everett R. Dempster and I. Michael Lerner, University of California, Berkeley.

The rate of genetic gain in a character for which selection is practiced depends in addition to the intensity of selection on (1) the accuracy of selection, and (2) the average interval between generations. These factors are not independent and exercise a pull in opposite directions. Through the application of Wright's technique of path coefficients comparisons can be made between the expected rates of genetic gain in populations containing varying proportions of breeding animals of different ages. The methods used involve the estimation of correlations between genotypes, and various selection indexes based on individual, sib and progeny records in inculled populations as well as in populations whose range has been restricted by previous selection. From these estimates the relative efficiencies of different age distribution schemes of a breeding population can be determined. A specific solution for such a situation in a flock bred for egg production will be presented as an illustration of the problems and methods used in the study of the genetics of populations under artificial selection.

18. Statistical Criteria of the Effectiveness of Selective Procedures. Preliminary Report. R. F. Jarrett, University of California, Berkeley.

The "validity coefficient," the standard error of estimate, the index of predictive efficiency, the "selection ratio" of Taylor and Russell, Johnson's Gamma, and other statistical devices have been suggested as indices of the effectiveness of selective programs. These devices all suffer from the deficiency that they do not permit a satisfactorily precise estimate of the dollar value of the increased output expected from the selection program and thus leave unsettled the question as to whether or not the cost of such a program is justified. The relationship between the correlation coefficient on the one hand and the mean value of Y for an unselected population (Y being an objective output-type criterion), the standard deviation of Y for an unselected population, and the mean value of Y for the upper Np individuals selected on the basis of their high performance on the selective test X on the other hand, provides the basis for estimating the increase in the mean output of a group of workers selected on the basis of a testing program yielding any specified validity coefficient with the criterion Y. Increase in productivity of selected workers is shown to be a function of the validity coefficient, the rigorousness of selection, and the coefficient of variability of the output criterion among "unselected" employees.

19. Approaches to Univocal Factor Scores. Preliminary Report. J. P. Guilford, University of Southern California.

In spite of the fact that univocal factor scores are badly needed for various reasons, it appears to be impossible by present methods to construct pure tests for some common factors. Recourse must therefore be made to statistical control of component variances. It is desirable to derive each factor score from a minimum number of tests. The availability of a few univocal tests makes this requirement fairly easy to satisfy. Such tests serve well as suppression variables for their common-factor variances where not wanted in other tests. Several principles may be invoked as objectives: (1) to maximize the desired variance in the impure test, (2) to reduce the undesired variance to zero, or (3) to minimize the undesired variance without intolerable loss of the desired variance. A secondary objective is to assure a combining weight of +1.00 for the test measuring the desired factor. Equations for achieving the objectives have been derived and the limitations and implications of each procedure have been noted. By means of statistical control, the situation seems hopeful for the achievement of univocal scores for a fairly large number of unique psychological variables. There are implications for experimental psychology as well as for vocational testing.

20. A Note on the Problem of Binary Stars. Elizabeth L. Scott, University of California, Berkeley,

This paper concerns some of the problems of Trumpler (see next abstract). ξ_{ij} is the radial velocity of the i-th star, $i=1, 2, \cdots, s$, at t_j selected at random, $j = 1, 2, \cdots, n$. x_{ij}, measurement of ξ_{ij}, is $N(\xi_{ij}, \sigma_i)$. ξ_{ij} is random with distribution $c\,(k_i^2 - (\xi_{ij} - \xi_{i0})^2)^{-\frac{1}{2}}$ where $k_i \geqq O$ and ξ_{i0} are unknown. (1) Test of hypothesis that $k_i = O$. Case (i) σ_i known. Whatever the exact test T, its power $\beta_T(k)$ has derivative $\beta_T'(0) = 0$. Test maximizing $\beta_T''(0)$ is that of Trumpler with criterion $S^2 = \sum_{j=1}^{n} (x_{ij} - x_{i\cdot})^2 > \chi^2 \sigma_i^2$. Case (ii) σ_i unknown. Whatever the exact test τ, $\beta_\tau^{(m)}(0) = 0$, $m = 1, 2, 3$. Test maximizing $\beta_\tau^{(4)}(0)$ is Trumpler's test $\sum_{j=1}^{n} (x_{ij} - x_{i\cdot})^4 > \left[\sum_{j=1}^{n} (x_{ij} - x_{i\cdot})^2\right]^2 C$. (2) Let $\sum_{j=1}^{n} (\xi_{ij} - \xi_{i0})^2 = 2\lambda_i \sigma^2$. For constant velocity stars $\lambda = 0$. For others it is a random variable. Since, given $\lambda = 0$, S^2 is distributed as non-central χ^2, an integral equation connects the distributions of S^2 and λ. Its solution yields an estimate of the proportion of constant velocity stars. After estimating the distribution of λ, the level of significance can be estimated and also the number n of measurements so that the proportion of constant velocity stars declared variable will be less than p, specified in advance.

21. Statistical Problems of Spectroscopic Binaries. ROBERT J. TRUMPLER, University of California, Berkeley.

Spectroscopic Binaries are stars whose radial velocities, as measured by the Doppler shift of spectral lines, show a periodic variation. The first problem is to obtain a statistical criterion for deciding whether a star with several radial velocity measures, made at different times, has a high probability (larger than a specified limit) of variable velocity and should be announced as an object worthy of further study. The second problem is to find the percentage of variable velocity stars among a large list of stars with several radial velocity measures for each star. From the distribution of standard errors only the percentage of cases where the velocity variation exceeds a certain limit can be ascertained. The third problem is concerned with those stars for which a binary orbit has been determined. The statistical distribution of these binary systems according to mean distance between the two stars and the ratio of their masses can be evaluated within certain limits.