ABSTRACTS OF PAPERS

Presented June 22–24, 1948 at the Berkeley Meeting of the Institute

1. Estimation of Parameters for Truncated Multinormal Distributions. Z. W. Birnbaum, E. Paulson and F. C. Andrews, University of Washington.

Let $X_{(N)} = (X_1, \cdots X_p, X_{p+1}, \cdots, X_N)$ be an N-dimensional random variable with a non-singular normal distribution, and let the expectations, variances and covariances of $X_{p+1}, \cdots X_N$ be known. A large sample of $X_{(N)}$ is available, obtained under some side-condition on $(X_{p+1}, \cdots, X_N)$; this side-condition may be a truncation of any kind or, more generally, a selection; i.e. imposing on $X_{p+1}, \cdots, X_N$ a probability-distribution different from the original marginal distribution. A method is developed for estimating, from such a large sample with a side condition, all the missing parameters of the original distribution of $X_{(N)}$, that is the expectations, variances and covariances of $X_1, \cdots, X_p$, and the co-variances $\sigma X_j X_k$ for $j = 1, \cdots, p$ and $k = p+1, \cdots, N$. This method does not require the knowledge of the side-condition. (This paper was prepared under the sponsorship of the Office of Naval Research.)

2. A Test of the Hypothesis that a Sample of Three Came from the Same Normal Distribution. Carl A. Bennett, General Electric Company, Hanford Works, Richland, Washington.

In the control of the precision of chemical analyses performed in duplicate, a test sometimes becomes necessary as to whether three determinations can reasonably be assumed to have arisen from the same normal population. A critical region for testing this hypothesis is given by $R > R_0$, where $R = D/d$, D being the maximum and d the minimum difference between the three values, and R_0 is determined by integration over the upper tail of the Cauchy distribution. It can easily be seen that this test is equivalent to a t-test between a sample of one and a sample of two.

3. A Note on the Application of the Abbreviated Doolittle Solution to Non-Orthogonal Analysis of Variance and Covariance. Carl A. Bennett, General Electric Company, Hanford Works, Richland, Washington.

S. S. Wilks has shown that the sums of squares necessary to the tests commonly made in non-orthogonal analyses of variance or covariance can in general be reduced to the ratio of two determinants. If several determinantal operations are performed to remove the singular principal minors, the abbreviated Doolittle solution yields these sums of squares directly. A combination of this technique and the calculational methods advanced by Wald and Yates greatly reduces the tedium of calculation in this type of analysis.

4. Yield Trials with Backcrossed Derived Lines of Wheat. G. A. Baker and F. N. Briggs, University of California, Davis.

Strains of White Federation 38 and Baart 38 Wheats derived by backcrossing sufficient to insure a high degree of homogeneity for all genetic factors were grown in conventional yield trials. The results were somewhat contradictory and led to a critical examination of such trials. The assumption that the deviations of yields in field trials from the specified pattern are random with uniform variance and expectation zero is not sufficiently realistic. We are led to consider a mathematical model which assumes a set of fertility levels upon which a random element is superimposed. On the basis of this model it is possible to account for the low observed correlations between residuals and plot yields. In such a model the variance ratio F may be approximately unbiased but then its variance is smaller than under conventional assumptions. On the other hand, the expected value of F may be greater than one and sufficiently large so that "significant differences" between strains will always be found due to the differences in fertility levels. In such cases the results of the experiment may be misinterpreted. Transformations, in the ordinary sense of the word, will not bring such data into conformity with the conventional model. In order to bring the correlation between residuals and plot yields down to a sufficiently low level it is necessary to concentrate most of the variation in fertility levels into a few plots. That this is not unreasonable is borne out by agronomic observations. This model also explains the absence of correlation between the yields of strains as determined in two different trials on the same set of strains.

5. The Selection of the Largest of a Number of Means.

CHARLES M. STEIN, University of California, Berkeley.

Suppose X_{ij}, $i = 1, \cdots, p$; $j = 1, 2, \cdots$ are independently normally distributed with means $\xi_i + \eta_j$ and variances σ_j^2 where ξ_i, η_j are unknown but σ_j^2 are known. ϵ, α are fixed numbers with $0 < \epsilon, 0 < \alpha < 1$. It is desired to select, by a sequential procedure, in which we take first the observations with second subscript 1, etc. an integer M among $1, \cdots, p$ such that, for every $k = 1, \cdots, p$ and $\xi_1, \cdots \xi_p, \eta_1, \eta_2, \cdots$ satisfying $\xi_k \geqq \xi_j + \epsilon$ for all $j \neq k, P\{M = k\} \geqq 1 - \alpha$. In accordance with the following rule, one decides at each stage (after the observations with second subscript n) to take no more observations with certain first subscripts. For each $n = 1, 2, \cdots$ and each $l = 1, \cdots, p$ compute

$$\sum_{j=1}^{n} \frac{1}{\sigma_j^2}\left(X_{l,j} - \bar{X}_j - \frac{\epsilon(t_j - 1)}{t_j}\right)$$

where $\bar{X}_j$ is the average of the observations with second subscript j and t_j is the number of such observations. Continue taking observations $X_{l,n+1} \cdots$ for those l for which this expression is greater than $(ln\alpha)/\epsilon$ but not for the others. Eventually there will be at most one subscript $l = 1, \cdots, p$ for which one continues to take observations and if there is one this is chosen to be M. If there is none, the l for which the sum is largest is chosen to be M. This procedure is a straight-forward application of the Lemma on p. 146 of Wald's *Sequential Analysis*, and generalizations can easily be found.

6. The Effect of Inbreeding on Height at Withers in a Herd of Jersey Cattle.

W. C. ROLLINS, S. W. MEAD, and W. M. REGAN, University of California, Davis.

The data consist of measurements of height at withers of about 200 females for various ages from one month to five years. The intensity of inbreeding as measured by Wright's coefficient of inbreeding averaged 15 per cent and reached as high as 44 per cent in a few cases.

An intra-sire covariance analysis of height and per cent of inbreeding was made for various ages from the first month to the fifty-fourth month.

The results of the statistical analysis indicate that the inbred animals are shorter at one month of age and grow more slowly up to about the sixth month than do the outcrossed animals, but that from the sixth month on the inbreds begin to catch up with the outcrossed so that at maturity there is no significant difference in height.

7. An Example of a Singular Continuous Distribution. HENRY SCHEFFÉ, University of California at Los Angeles.

Simple and "natural" examples of singular continuous probability distributions are of pedagogical interest. They are trivially available in the k-variate case for $k > 1$. A univariate example may be obtained from the notion of a sequence of independent trials of an event with constant probability p of success, a notion familiar to the student and indispensable in elementary probability theory. The (real-valued) random variable X is taken to be the dyadic representation of the sequence of results (1 and 0, respectively, for success and failure). It is known that X has a singular continuous distribution for $p \neq 0, \frac{1}{2}, 1$. This result may be proved by using only the Tchebycheff inequality together with the formulas for the mean and variance of the binomial distribution.

8. On the Theory of Some Non-Parametric Hypotheses. ERICH L. LEHMANN and CHARLES STEIN, University of California, Berkeley, California.

For two types of non-parametric hypotheses optimum tests are derived against certain classes of alternatives. The two kinds of hypotheses are related and may be illustrated by the following example: (1) The joint distribution of the variables $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ is invariant under all permutations of the variables; (2) the variables are independently and identically distributed. It is shown that the theory of optimum tests for hypotheses of the first kind is the same as that of optimum similar tests for hypotheses of the second kind. Most powerful tests are obtained against arbitrary simple alternatives, and in a number of important cases most stringent tests are derived against certain composite alternatives. For the example (1), if the distributions are restricted to probability densities, Pitman's test based on $\bar{y} - \bar{x}$ is most powerful against the alternatives that the X's and Y's are independently normally distributed with common variance, and that $E(X_i) = \xi$, $E(Y_i) = \eta$ where $\eta > \xi$. If $\eta - \xi$ may be positive or negative the test based on $| \bar{y} - \bar{x} |$ is most stringent. The definitions are sufficiently general that the theory applies to both continuous and discrete problems, and that tied observations present no difficulties. It is shown that continuous and discrete problems may be combined. Pitman's test for example, when applied to certain discrete problems, coincides with Fisher's exact test, and when $m = n$ the test based on $| \bar{y} - \bar{x} |$ is most stringent for hypothesis (1) against a broad class of alternatives which includes both discrete and absolutely continuous distributions.

9. Concerning Compound Randomization in the Binary System. JOHN E. WALSH, Project Rand, Santa Monica, California.

Consider a set of binary digits. The numerical deviation from $\frac{1}{2}$ of the conditional probability that a specified digit equals 0 is called the bias of that digit for the given conditions on the remaining digits of the set. The maximum bias of the set is defined to be the maximum of the biases of the digits of the set. A set of binary digits is called random if its maximum bias is zero. Now consider an array of $(1 + t_1) \cdots (1 + t_K) \times n$ binary digits such that the rows are statistically independent. A compounding method of obtaining a set of $t_1 \cdots t_K n$ binary digits from the original array is presented. By suitable choices of $K, t_1, \cdots, t_K$ the maximum bias of the compounded set can be made extremely small even if the maximum bias of the original array is not small; this can be done so that $t_1 \cdots t_K / (1 + t_1) \cdots (1 + t_K)$ is moderately large. Also a method is outlined for constructing an approximately random binary digit table. This table has the property that the maximum bias of a set of digits taken from the table is an increasing function of the number of digits in the set.

10. A Multiple Decision Problem Arising in the Analysis of Variance. EDWARD PAULSON, University of Washington, Seattle.

In some applications of the analysis of variance, a procedure is required for classifying varieties into 'superior' and 'inferior' groups. Consider K varieties, with $x_{i\alpha}$ the α^{th} observation on the i^{th} variety ($\alpha = 1, 2, \cdots r; i = 1, 2, \cdots K$), let $\bar{x}_i = \sum_{\alpha=1}^{r} x_{i\alpha}/r$ and let s^2 be an independent estimate of the variance. For the i^{th} variety form the corresponding interval $\left(\bar{x}_i - \frac{\lambda s}{\sqrt{r}}, \bar{x}_i + \frac{\lambda s}{\sqrt{r}}\right)$. The superior group then consists of the variety with greatest sample mean, together with those varieties whose corresponding intervals have at least one point in common with the interval for the variety with the greatest mean. If all varieties fall into one group, this group is labeled 'neutral' and the varieties are considered homogeneous. To select λ, consider the relative importance of different incorrect classifications. For a given λ, an explicit expression is found for $P(A)$, the probability the varieties will not all be classified in one group when $m_1 = m_2 = \cdots = m_k$ where $m_i = E(\bar{x}_i)$; also explicit expressions are found for $P(B_1)$ and $P(B_2)$, where $P(B_1)$ is the probability there will not be a superior group consisting only of the Kth variety and $P(B_2)$ is the probability there will not be a superior group consisting of at least the Kth variety, when $m_1 = m_2 = \cdots = m_{k-1} = m$ and $m_k = m + \Delta(\Delta > 0)$. Similar results are obtained for classifying K processes according to their variances.

11. Recurrence Formulae for the Moments and Semi-variants of the Joint Distribution of the Sample Mean and Variance. OLAV REIERSØL, University of Oslo, Norway.

Let $x_1, x_2, \cdots, x_n$ be independent and having the same distribution. We consider the arithmetic mean m and the variance $v = (1/(n-1)) \Sigma (x_i - m)^2$. Let κ_{rs} denote the seminvariants of the joint distribution of m and v, and let the seminvariant generating operators K be defined by the equations: $\kappa_{r+1,s} = K_1\kappa_{rs}$, $\kappa_{r,s+1} = K_2\kappa_{rs}$ $K_i.1 = 0, K_i(PQ) = P(K_iQ) + Q(K_iP)$. An operator which operates only on the first factor of a product shall be denoted by a prime, and an operator which operates only on the second factor shall be denoted by a double prime. We have the following general formula, valid for any parent distribution: $K_1^r[(n-1)(K_2 + \kappa_{01}' + \kappa_{01}'') - 2n(K_1' + \kappa_{10}')(K_1'' + \kappa_{10}'')]^s[1\cdot(\kappa_{01} - n\kappa_{20}) + n(\kappa_{10}\cdot\kappa_{10} - 1\cdot\kappa_{10}^2)] = 0$. For $s = 0, 1, 2$, we obtain the formulae, $K_1^r(\kappa_{01} - n\kappa_{20}) = 0$, $K_1^r[(n-1)(\kappa_{02} - n\kappa_{21}) - 2n^2\kappa_{20}^2] = 0$, $K_1^r[(n-1)^2(\kappa_{03} - n\kappa_{21}) - 8n^2(n-1)\kappa_{21}\kappa_{20} + 4n^3(n-1)\kappa_{30}^2 - 8n^3(n-1)\kappa_{20}^3] = 0$.

12. The Problem of Identification in Factor Analysis. OLAV REIERSØL, University of Oslo, Norway.

The paper is concerned with the multiple factor analysis of L. L. Thurstone. Thurstone has given criteria which he says are almost certain to constitute sufficient and more than necessary conditions for uniqueness (i.e. identifiability) of a simple structure. It is shown that Thurstone's criteria are not always sufficient, and conditions are derived which are more nearly necessary and sufficient for the identifiability of a simple structure. Let A be the matrix of factor loadings with n rows and r columns. When the communalities are identifiable, the conditions will be: (1) Each column of A should have at least r zeros. (2) Let us consider the submatrix B of A, consisting of all the rows which have zeros in the kth column. Then, for $q = 1, 2, \cdots, r-1$, there should for any combination of q columns different from the kth, exist at least $q + 1$ rows of B containing non-zero elements in the q columns. This should be true for any value of k.

13. Note on Distinct Hypotheses. AGNES BERGER, Columbia University, New York.

As was pointed out by Neyman, one of the difficulties which one may encounter when devising a test to distinguish between two exhaustive and exclusive composite hypotheses referring to the unknown distribution of a random vector X is the following: If H_0 states that the true distribution function of X belongs to a set $\{F\}$ and H_1 that it belongs to a set $\{G\}$, it may happen that to every Borel set W of the sample space there exists an element F_W in $\{F\}$ and an element G_W in $\{G\}$ for which the probability of the sample point x falling on W is the same and therefore independent of whether H_0 or H_1 is true. If this is the case the pair H_0, H_1 is called non-distinct, otherwise they are called distinct. The existence of non-distinct hypotheses is demonstrated by a simple example, H_0 consisting of one, H_1 of three suitably chosen stepfunctions. It is shown however that if the sets $\{F\}$ and $\{G\}$ contain only continuous distribution functions and are at most enumerable then the pair H_0, H_1 is distinct. Necessary and sufficient conditions for H_0 and H_1 to be distinct were obtained jointly with Wald for an important class of hypotheses each containing a continuum of alternatives.

14. Place of Statistical Sampling in the Education of Engineers. E. L. GRANT, Stanford University.

There is convincing evidence that many engineering problems could be solved better with the aid of statistical methods than they are now solved without this aid. However, few practising engineers or teachers of engineering have had any training in statistical methods. As a result, those engineering problems which are in part statistical problems are seldom recognized as such. Even in the field of industrial quality control, in which successful applications of some of the simpler statistical techniques have been made in many different industries, the surface has barely been scratched and a serious obstacle to progress is the lack of a widespread appreciation of the statistics point of view among design engineers, production engineers, inspection personnel, and management.

This condition might gradually be corrected if during the next few years instruction in statistics should be introduced into all undergraduate engineering curricula. However, some recent discussions touching on the subject of statistics instruction for engineering students (e.g., the report on "The Teaching of Statistics" which appeared in the March 1948 issue of the *Annals of Math. Stat.*) have been most unrealistic regarding the amount of statistics instruction which could be added to engineering curricula. These discussions have suggested a full year of basic statistics followed by one or more courses in engineering applications. Desirable as this arrangement might be from the point of view of the most effective instruction in statistics, it is out of the question when considered in the light of the many subjects which are needed in engineering curricula. Although undergraduate engineering curricula have always been tighter than other curricula, the pressures today are greater than ever before—for more time devoted to the humanistic-social stem, for more time in basic mathematics and science, for introductory courses in various economic and management subjects such as engineering economy, accounting, industrial relations, business law, and industrial organization and management, and for more time in the various departmental courses in engineering subjects in order to permit presentation of important recent developments in engineering technology. Under these circumstances the most that can be hoped for in the undergraduate program is a single statistics course for one term, possibly three units for one semester or four units for one quarter. This should be supplemented by additional statistics instruction for some graduate students in engineering. A few engineering graduates should be encouraged to take graduate degrees in statistics and to make careers in the field of applied statistics.

In a successful undergraduate statistics course for engineering students, the problems and illustrations should be selected with two purposes in mind. One purpose, of course, should be to develop the principles of probability and statistical method. The other, equally important if these engineering graduates are to persuade their colleagues and superiors to adopt the statistics point of view in approaching engineering problems, should be to demonstrate how statistical method provides a useful guide to action in many different engineering situations. Applications of statistics to industrial quality control provide particularly good problems and illustrative examples which serve this second purpose.

15. **Statistical Problems of Medical Diagnosis.** Jerzy Neyman, University of California, Berkeley.

"Diagnosis" is used to describe the outcome of a strictly defined test T, such as Wassermann test, which may lead to either of two possible outcomes, "positive" or "negative". Cases contemplated are such that at the time the test T is performed it is impossible to verify its verdict for certain and the best one can do is to repeat the test. It is postulated that to each individual of a population there corresponds a probability p that the test T will give a positive outcome. The value of p may vary from one individual to another. It is presumed that as p increases, the illness in the patient increases. Problem of comparison of two alternative tests and problem of estimating the distribution of p reduces to problems relating to the distribution of X = number of positive outcomes in n independent diagnoses. Statistical machinery suggested is that of *BAN* estimates (*Public Health Report*, Vol. 62, (1947), p. 1449). Principal result reported is that, with the mathematical model used in the paper quoted, the empirical variances of four *BAN* estimates computed for 205 samples of 1000 elements each agreed reasonably with the theoretical asymptotic values. Empirical distributions of three of these estimates did not show deviations from normality. That of the fourth was non-normal. It seems therefore that the asymptotic procedure of *BAN* estimate may be adequate for similar analyses.

16. **Power of Certain Tests Relating to Medical Diagnosis.** C. L. Chiang and J. L. Hodges, Jr., University of California, Berkeley.

Associate with each individual in a population π the probability p that he will be found tubercular when examined by a standard X-ray technique. Yerushalmy and others [*J. Am. Med. Assn.*, Vol. 133, (1947), p. 359] performed 5 independent such diagnoses on each of 1256 persons. Neyman [*Public Health Reports*, Vol. 62, (1947), p. 1449] proposed a simple four-parameter model for the distribution of p in π, estimated the parameters from the data of Yerushalmy and others, and obtained a satisfactory fit. In the present paper, the work of Neyman is paralleled with four new models, all giving satisfactory fit with the same data. The five models differ considerably in shape, and in the number of repeated diagnoses which they indicate to be necessary to detect a high proportion of those individuals having, say, $p \geqq 0.1$. Therefore further preliminary study seems indicated before one can design a mass survey to detect a high proportion of such persons. The approximate power of the χ^2 test of the Neyman model is considered, using one of the other models as alternative. It is found that to obtain power 0.7 with level of significance 0.05, it would be necessary to diagnose 5290 individuals 5 times each.

17. **Iterative Treatment of Continuous Birth Processes.** T. E. Harris, Project Rand, Santa Monica, California.

Random variables z_n are defined by $z_0 = 1$; $P(z_1 = r) = p_r$, $r = 1, 2, \cdots$; if $z_n = k$, z_{n+1} is the sum of k independent variates, each distributed like z_1. Let $x = \sum_1^\infty rp_r < \infty$;

$\sum_{1}^{\infty} r^2 p_r < \infty$; $0 < p_1 < 1$. The generating function $f(s) = \sum_{1}^{\infty} p_r s^r$ is said to be C.I. if there exists a family of generating functions $f(s, t)$ with $f(s, 1) = f(s)$, $f[f(s, t), t'] = f(s, tt')$ for all nonnegative t and t'. A necessary and sufficient condition that $f(s)$ be C.I. is that the numbers a_r, $r = 2, 3, \cdots$, be nonnegative; the a_r are determined recursively by requiring that the power series $\xi(s) = -s + \sum_{2}^{\infty} a_r s^r$ satisfy formally the functional equation $\xi(s)f'(s) = \xi[f(s)]$. The problem is connected with classical works on iteration. If $f(s)$ is C.I., the given Markoff process can be imbedded in a continuous birth process. If $\xi(s)$ is given, the m.g.f. $\phi(s)$ of the asymptotic distribution of the variate z_n/x^n may be determined from the formula $\phi^{-1}(s) = (s - 1) \exp\left\{\int_1^s \left[\frac{\xi'(1)}{\xi(y)} + \frac{1}{1 - y}\right] dy\right\}$. Various properties of the corresponding distribution can be inferred from this expression.

18. Estimation of Means on the Basis of Preliminary Tests of Significance. BLAIR M. BENNETT, University of California, Berkeley.

This paper examines the statistical procedure of pooling two sample means on the basis of the results of one or more preliminary tests of significance. Let x_i, $(i = 1, \cdots, N_1)$, represent a sample of N_1 observations from a normal population $\eta_1(\xi, \sigma_1^2)$, and y_i a sample of N_2 observations from $\eta_2(\eta, \sigma_2^2)$. An estimate of ξ which is commonly used in certain practical situations is given by: $x' = \bar{x}$, or $x' = (N_1\bar{x} + N_2\bar{y})/(N_1 + N_2)$, according as the sample means $\bar{x}$, $\bar{y}$ do or do not differ significantly on the basis of a preliminary test. The distribution of the estimate x' is determined, according as $\sigma_1 = \sigma_2$ are known or unknown. In both situations, the maximum (or minimum) bias is computed as a function of various levels of significance of the preliminary test of equality of means. Also, the mean square error of the estimate x' is calculated in both cases. If now equality of variances cannot be assumed, but an F-test of the sample variances s_1^2, s_2^2 does not indicate any significant difference, then in practice $\bar{x}$, $\bar{y}$ may be pooled, the weights being inversely proportional to the sample variances. Thus, the usual estimate of ξ will be of the form: $x' = \bar{x}$, or $x' = (N_1\bar{x}/s_1^2 + N_2\bar{y}/S_2^2)/(N_1/S_1^2 + N_2/S_2^2)$, according as $\bar{x}$ and $\bar{y}$ do or do not differ significantly on the basis of the Student t-test, subsequent to an F-test. The bias and mean square error of this estimate have been computed with the aid of the conditional power function of the t-test subsequent to an F-test.

19. Note on Power of the F Test. STANLEY W. NASH, University of California, Berkeley.

Assuming "treatment" expectations to be normal random variables, the ratio of the sum of squares due to treatments to the sum of squares due to error has a central F distribution in the cases of randomized blocks, Latin squares, and one-way classifications. The F statistic converges in probability to a constant as the number of treatments is increased. This is one plus a multiple of the variance between treatment expectations. The power of the F test increases monotonely to one as the number of treatments is increased. This power can be calculated using tables of the incomplete beta function.

20. Best Asymptotically Normal Estimates. E. W. BARANKIN and J. GURLAND, University of California, Berkeley.

The methods of minimum χ^2 developed by Neyman for obtaining BAN (best asymptotically normal) estimates of the parameters appearing in the multinomial distribution are generalized to obtain certain optimum types of estimates in the case of an arbitrary distribution under certain restrictions. Let the random vector X have the probability density $v(x; \theta)$ in the absolutely continuous case and let $v(x; \theta) = P\{X = x/\theta\}$ in the discrete case, where θ is a fixed vector in the parameter space. Functions $\phi_i(X)$, $(i = 1, 2, \cdots, r)$ are selected for the purpose of forming estimates; these estimates are taken to be functions of the sample moments $\frac{1}{n} \sum_{j=1}^{n} \phi_i(x_j)$. Certain quadratic forms which depend on the choice of functions $\phi_1(X), \phi_2(X), \cdots, \phi_r(X)$ are minimized with respect to the parameters. In this manner, asymptotically normal estimates are obtained which are consistent, and have minimum asymptotic variances within the class of estimates so determined by the particular functions $\phi_1, \phi_2, \cdots, \phi_r$. It is possible, through a modification of this procedure, to obtain estimates by solving a set of linear equations. If $v(x; \theta)$ has the form

$$v(x; \theta) = \exp \{\beta_0(\theta) + \sum_{i=1}^{s} \beta_i(\theta) y_i(x) + y_0(x)\}$$

it can be shown that the best choice of the ϕ's is $y_1(x), y_2(x), \cdots, y_s(x)$. Maximum likelihood estimates belong to this class of BAN estimates.