ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Seattle Meeting of the Institute on November 26–27, 1948)

1. Estimation of the Variance of the Bivariate Normal Distribution. HARRY M. HUGHES, University of California, Berkeley.

Let x_1 and x_2 be two random variables normally distributed with known means m_1 and m_2, and with common unknown variance σ^2. Consider an experiment in which the observed variable is $Y = \sqrt{(x_1 - m_1)^2 + (x_2 - m_2)^2}$. This paper considers the problem of estimating the parameter σ when the observations are grouped. By the method of minimum reduced chi-square with linear restrictions, two best asymptotically normal estimates are derived. By minimization of the asymptotic variance of these estimates, the optimum choice of grouping is found as a function of σ. For two and for three groups, when it is known or assumed a priori that σ is on a certain finite interval, the optimum grouping is derived which will minimize the maximum asymptotic variance on that interval. If such interval is moderately small, it is shown that the optimum grouping is the same as if σ were known to have the value at the upper end of the interval. Finally the effect of using non-optimum grouping is analyzed.

2. Derivation of a Broad Class of Consistent Estimates. R. C. DAVIS, Inyokern, California.

Given a chance vextor X with cumulative distribution function $F(X, \theta)$, where θ is an unknown parameter vector, a broad class of estimates of θ is derived having the following properties: a) any estimate in this class is a consistent estimate of θ; b) any estimate is a symmetric function of independent observations of the chance vector X. The novel feature of this class is that no assumptions about the existence of various partial derivatives of a density function with respect to θ are made. As a matter of fact not even the existence of a density function is required, and it is postulated merely that a continuous function of X for each θ (in a certain neighborhood of the true θ_0) and of θ for each X exist which satisfies a Lipschitz condition in θ. For each such function having a finite first moment an estimate of θ is constructed which has the properties a) and b) listed above.

3. Locally Best Unbiased Estimates. EDWARD W. BARANKIN, University of California, Berkeley.

Let $p = \{p_\vartheta(x); \vartheta \in \Theta\}$ be a family of probability densities in the space Ω of points x; and g a function on Θ. Let s be fixed and >1; call an unbiased estimate of g *best at* ϑ_0 if its s-th absolute central moment (s.a.c.m.) under p_{ϑ_0} is (finite and) not greater than the s.a.c.m., under p_{ϑ_0}, of any unbiased estimate of g. With a certain integrability postulate on the p_ϑ's, a necessary and sufficient condition, of finite character, is established for the existence of an unbiased estimate of g having a finite s.a.c.m. under p_{ϑ_0}. When such a one exists, there then exists a *unique* unbiased estimate which is best at ϑ_0. The existence condition defines the s.a.c.m. of the best estimate explicitly as the l.u.b. of a set of numbers; in particular, we obtain immediately a generalization of the Cramér-Rao inequality. Also, when it exists, the best unbiased estimate is explicitly constructed from the function g and the densities p_ϑ. The case $s = 2$ is studied more closely. Also, a detailed example is considered.

4. Some Problems Related to the Distribution of a Random Number of Random Variables. EDWARD PAULSON, University of Washington, Seattle.

Let $\{x_i\}$ $(i = 1, 2, 3 \cdots)$ be a set of independent random variables with identical distributions, with $E(x) = a$ and $\sigma^2(x) = b$ $(0 < b < \infty)$. Let N be a positive integral-valued random variable with distribution $F_\lambda(N)$ depending on a parameter λ, where $E(N) = A_\lambda$, and $\sigma^2(N) = B_\lambda$ $(0 \leq B_\lambda < \infty)$. Now let $T_N = x_1 + x_2 + \cdots + x_N$. The limiting distribution as $\lambda \to \infty$ of $\dfrac{T_N - aA_\lambda}{\sqrt{a^2B_\lambda + bA_\lambda}}$ has been investigated by Robbins (*Proc. Nat. Acad. Science*, Vol. 34 (April 1948), pp. 162–163) for several different sets of conditions on the distribution of N. It can be shown that analogous results will hold if instead of T_N we consider a more general statistic T_N^*, whose conditional distribution with respect to the variable N is such that there exist constants a_1 and b_1 so that

$$\lim_{N \to \infty} E\left\{\exp\left[it\left(\frac{T_N^* - a_1N}{\sqrt{b_1N}}\right)\right]\right\} = h(t)$$

uniformly in any finite t interval, where $h(t)$ is a characteristic function. Returning to the statistic T_N, it can be shown that there exists an asymptotic expansion in powers of $\lambda^{-\frac{1}{2}}$ with remainder $O(\lambda^{-(k/2)+1})$ for $P\left\{\dfrac{T_N - aA_\lambda}{\sqrt{a^2B_\lambda + bA_\lambda}} \leq u\right\}$ when the following conditions are satisfied: (1) the distribution function of x has a non-zero absolutely continuous component, (2) $E(|x|^k) < \infty$, and (3) $\lambda \to \infty$ through integral values, and $F_\lambda(N)$ is the λth convolution of a random variable n such that $E(n^k) < \infty$.

5. Asymptotic Expansions for the Distribution of Certain Likelihood Ratio Statistics. ALBERT H. BOWKER, Stanford University.

Asymptotic expansions of the "Cramerian" type are derived for the distribution of likelihood ratio statistics given by Wilks for testing various hypotheses about means, variances, and covariances of a normal multivariate distribution. The point of departure is Wilks' result that minus twice the logarithm of the likelihood ratio has the χ^2 distribution; terms in $\dfrac{1}{N}, \dfrac{1}{N^2}, \cdots$ may also be expressed in terms of the χ^2 distribution. In addition, asymptotic expansions of the "Fisher-Cornish" type are obtained for the percentage points and for a transformation of the statistic to a χ^2 variate.

6. On a Problem of Confounding in Symmetrical Factorial Design. ESTHER SEIDEN, University of California, Berkeley.

Let $m_3(r, s)$ be the maximum number of factors that is possible to accommodate in symmetrical factorial experiment in which each factor is at $s = p^n$ levels (p being any positive prime number, n a positive integer) and each block is of size s^r, without confounding any degrees of freedom belonging to any interaction involving 3 or lesser number of factors.

R. C. Bose proved in a paper "Mathematical theory of factorial design," *Sankhyā*, Vol. 8 (1947), pp. 107–166, that the following inequality holds:

$s^2 + 1 \leqq m_3(4, s) \leqq s^2 + s + 2$. This gives in case $s = 4, 17 \leqq m_3(4, 4) \leqq 22$. It is now proved that $m_3(4, 4) = 17$.

The proof consists in showing that the maximum number of non three collinear points which can be chosen in a finite projective space $PG(3, 4)$ cannot exceed 17, which according to a proof of R. C. Bose is equivalent to the staement that m_3 (4, 4) cannot exceed 17.

7. Some Bounded Significance Level Tests of Whether the Largest Observations of a Set are Too Small (*Preliminary Report*) John E. Walsh, Santa Monica, California.

A set of n observations are given which satisfy: (1) They are independent and from continuous symmetrical populations; (2) The r largest observations are from populations with median θ while the remaining observations are from populations with median φ. It is required to test whether $\theta < \varphi$. Let $x(1), \cdots, x(n)$ denote the observations arranged in increasing order of magnitude. For $r = 1$ tests of the form: *Accept* $\theta < \varphi$ *if* $x(n) < 2x(w_\alpha) - x(i)$, *where* $\alpha = \Pr[x(n) + x(i) < 2\theta \mid \theta = \varphi]$ *and* w_α *is the smallest integer satisfying* $\Pr[x(w_\alpha) > \theta \mid \theta = \varphi] \leq \alpha$, can be obtained from $n \geq 15$. Exact significance levels can be obtained by assuming a sample from a specified population (e.g. normal). On the basis of (1)–(2) alone, the significance level never exceeds 2α. For large n, tests can be obtained for any r if the observations satisfy the additional weak condition: (3) The tail order statistics are approximately independent of the central order statistics; also the variances of the tail order statistics are either very large or very small compared with the variances of the central order statistics. The test is: *Accept* $\theta < \varphi$ *if* $\max[x(i_i) + x(n - j_k); 1 \leq k \leq s \leq r] < 2x(w_\alpha)$, *where* $i_u < i_{u+1}$, $j_v < j_{v+1}$, $j_s = r - 1$, w_α *is the smallest integer satisfying* $\Pr[x(w_\alpha) > \theta \mid \theta = \varphi] \leqq \alpha$, *and* $\alpha = \Pr\{\max[x(i_k) + x(n - j_k)] < 2\theta \mid \theta = \varphi\}$. For large n the significance level is approximately α but is $\leq 2\alpha$ for all n. The power function $\rightarrow 1$ as $\varphi - \theta \rightarrow \infty$ and $\rightarrow 0$ as $\varphi - \theta \rightarrow -\infty$.

8. Determination of Optimal Test Length to Maximize the Multiple Correlation. Paul Horst, University of Washington, Seattle.

If the lengths of the tests in a battery are altered, their intercorrelations and their validities or correlations with a criterion are also altered. Consequently, the multiple correlation of the battery with the criterion will also be altered. These changes are a function of the reliabilities of the tests. Suppose we have given from a set of experimental data (1) the time allowed for each test in the battery, (2) the reliability of each test, (3) the intercorrelations, and (4) the validities of all the tests. If we specify the overall testing time we are willing to allow for the test in the future, we can determine the amount by which each test must be altered in order to give the maximum multiple correlation with the criterion. The method, together with numerical examples and the mathematical proof, is presented.

9. Some Numerical Comparisons of a Non-Parametric Test with other Tests. F. J. Massey, University of Oregon, Eugene.

Let $F(x)$ be the cumulative distribution function of a R.V.X, and let $x_1 < x_2 \cdots < x_n$ be the results of n independent observations ordered as to size.

Define $S_n(x) = 0$ if $x < x_1$;

$$= \frac{k}{n} \text{ if } x_k \leq x < x_k + 1;$$

$$= 1 \text{ if } x_n < x.$$

To test the hypothesis $H_0 : F(x) = F_0(x)$, where $F_0(x)$ is completely specified, use the criterion: reject H_0 if $\max \mid S_n(x) - F_0(x) \mid > \frac{\lambda}{\sqrt{n}}$. Choice of λ determines the first kind of error. The second kind of error against specified alternatives can be calculated numerically.

10. On the Deviation of Extreme Values. W. J. DIXON, University of Oregon, Eugene.

Let $x(i)$ be the ith observation in order of magnitude in a sample of size n. The distribution of $R = \frac{x(n) - x(2)}{x(n) - x(1)}$ is obtained explicitly for samples from a rectangular distribution and for $n = 3, 4, 5$, for samples from a normal distribution. Percentage values of R for values of n up to 30 are presented. Generalizations of R are indicated.

11. The Optimum Size of Interval for Making Measurements of a Rocket's Angular Velocity. EDWARD A. FAY, University of California, Berkeley.

Over a given range of time $0 \leq \tau \leq T$, the angular velocity of a rocket's spin is adequately represented by a polynomial $\xi(\tau)$ of given degree $s - 1$ but with unknown coefficients. The rocket's angular acceleration and the angle through which it spins in a given time-interval may then be obtained respectively by differentiating and integrating $\xi(\tau)$. Let ν be an integer $\geq s$, let $t = T/\nu$, and let η_i be the angle through which the rocket turns in the interval $(i - 1)t \leq \tau \leq it$. While $\xi(\tau)$ and $\xi'(\tau)$ cannot be directly observed, the angles $\eta_1, \eta_2, \cdots, \eta_\nu$ can. Let Y_i be an observation on η_i, and assume that $Y_1, Y_2, \cdots, Y_\nu$ are independent homoscedastic variables. The Y_i may then be combined by the method of least squares to obtain best linear estimates $X(\tau, t)$ and $X'(\tau, t)$ of $\xi(\tau)$ and $\xi'(\tau)$. The choice of t is at the observer's disposal. For the cases $s = 2, 3, 4$, and for the cases that the common variance of the Y_i is (a) independent of t or (b) proportional to t, an expression is derived for the variance of $X'(\tau, t)$, and the maximum value of that variance over the range of τ is minimized with respect to t. The method is of much more general application.

12. Stationary Time Series Analysis and Common Stock Price Forecasting. ZENON SZATROWSKI, University of Oregon, Eugene.

The objective of this paper is to present a statistical procedure of practical value in the problem of extracting information from the past behaviour of economic time series, information to be used in projecting future patterns. The author feels that his approach yields results closer to reality than the techniques described by Herman Wold, M. C. Kendall, H. T. Davis, and in particular, the technique of "disturbed harmonics" used by G. U. Yule. The idea of the proposed technique can be described by examining the autoregression scheme, which seems to be considered the most desirable by the above men. A simple example of such a scheme is the equation

$$u_{t+2} = -au_{t+1} - bu_t + E_{t+2},$$

where the u's are the time series values and E's are random elements. The above linear relationship, when determined either directly or through an empirical correlogram (for which data is usually inadequate) is a kind of an average relationship. It may be as inappropriate in estimating future values of a time series as would be an average in estimating the level of a series with a pronounced trend.

The author proposes using derived time series to shed light on the nature of the changes in the parameters under consideration. Such derived series could be estimates of the a's and b's for successive time periods. The author has found that projections of common stock price fluctuations were improved considerably when the changing nature of the cyclical pattern was taken into account. This was done by constructing derived time series, "moving" estimates of the amplitude, period and phase of the dominant harmonics.

The author points out that the above approach has shown promise in commodity prices as well as common stocks. The value of this approach in forecasting lies in the facts that (1) it does not require forecasts of other series and (2) it is based on the realistic assumption that history repeats itself but with variations, variations which may be taken into account through appropriate models.

13. Distribution of the Number of Schools of Fish Caught Per Boat. J. NEYMAN, University of California, Berkeley.

Let λ be the average number of schools of fish per unit area of a fishing ground A. Let a be any area partial to A, and let $\Omega(m, a, \lambda)$ denote the probability that exactly n schools of fish will be found within a. At time $t = 0$ a boat begins scouting for fish in A traveling at constant speed v. It is assumed that all schools of fish within distance r of the boat are detected and none is detected at a greater distance. If $s \geqq 1$ schools are detected then they are caught in turn, the catching of one school taking up exactly h hours. $X(t)$ denotes the random variable representing, for each $t \geqq 0$, the number of schools caught up to time t including the one which may be in the process of being caught at the moment t. Probability distribution of $X(t)$ is given by the formula

$$P\{X(t) \leq k\} = \sum_{m=0}^{k} \Omega[m, 2rv(t - kh), \lambda]$$

for $k = 0, 1, 2, \cdots, n - 1$, where $n - 1$ is the greatest integer smaller than t/h. Of course $P\{X(t) \leq n\} = 1$. This result is easily generalized for the joint distribution of catches of several boats fishing in the same area so that their paths do not cross. Assuming specific functions to represent $\Omega(m, a, \lambda)$ formulae may be obtained to estimate the parameters λ and rv.

14. Some Problems in Fishery Research to which Statistical Methods are Applicable (*Preliminary Report*). RALPH P. SILLIMAN, U. S. Fish and Wildlife Service, Seattle.

One of the most difficult problems is the obtaining of a random sample of a fish population. Rarely are such populations randomly distributed over any area, and the samples must often be taken from the catches of fishing vessels which do not uniformly cover even a part of the area of distribution of the population. Many distributions of variables found in fishery research are not normal, and statistical methods based on the normal distribution can be applied only through the use of unsatisfactory transformations. Since fishery research is largely observational in technique, data reflecting the concurrent effect of several variables are usually obtained. Although the present methods of multiple correlation and regression can be used in some instances to measure the relative effect of the separate variables, there are many situations in which these methods do not yield good results. Finally, many data used in fishery research must be adjusted before use, and existing methods do not give good measures of the expected variability of such adjusted data. Examples of specific problems are found in the distribution of deliveries and the variations in catch of Columbia River chinook salmon.

15. The Application of the Hypergeometric Distribution to Problems of Estimating and Comparing Zoological Population Sizes. DOUGLAS CHAPMAN, University of California, Berkeley.

Estimates and tests of the χ^2 type, as developed by Neyman, are adapted to sampling without replacement from a finite population. These results are applied to problems of estimation and comparison of zoological population sizes as determined by sampling procedures. For single samples the bias and variance of different estimators is compared. Finally some numerical calculations are made for various population and sample sizes to determine how different sample sizes and different methods of analysis affect the size of the critical region which is necessarily an approximation to the desired size. For some of these the power of the test is considered.

16. Extension to Multivariate Case of Neyman's Smooth Test with Astronomical Application. Elizabeth L. Scott, University of California, Berkeley.

It is more or less generally accepted that the distribution of extra-galactic nebulae in space is not uniform in the small. In particular, counts in small cubes show distinct signs of contagion. On the other hand, it is not settled whether or not lack of uniformity in the large exists. One way of making this statement precise is to assert that the power series expansion of the logarithm of the probability density of the two angular coordinates of the nebulae within a given large area on the unit sphere does not contain low order terms. In fact, any such low order terms could be interpreted as determining "trends" or what could be described as lack of uniformity in the large. From this point of view, uniformity in the large may be tested by a two dimensional Neyman Smooth Test for goodness of fit.

Let $\{\pi_{ij}(x, y)\}$ be a sequence of polynomials in x and y ortho-normal for $|x| < a$ and $|y| < b$. If x_k and y_k are the coordinates of the kth out of n nebulae counted within the rectangle $(-a, a)$, $(-b, b)$ then the smooth test of mth order consists of rejecting the hypothesis of uniformity in the large when $\sum_{i+j=1}^{m} \left(\sum_{k=1}^{n} \pi_{ij}(x_k, y_k) \right)^2 \geqq n\chi_\epsilon^2$ where χ_ϵ^2 is the tabled value of χ^2 with $m(m+3)/2$ degrees of freedom.

17. A Mathematical Theory of Vitamin A Metabolism in Fish (*Preliminary Report*). Normal E. Cooke, Vancouver, B.C.

Several possible hypotheses for vitamin A metabolism in fish are developed from simple postulates. These hypotheses are tested (by least squares method) against experimental data in an attempt to deduce the correct mechanism.

18. The Interactance Hypothesis between Populations. Stuart C. Dodd, University of Washington, Seattle.

The hypothesis of interacting between human populations, or of demographic gravitation, is that the number of interactions between two communities (or other groups) tends to vary directly with the product of the two populations and their "specific coefficients" and the overall duration and tends to vary inversely with the intervening distance and the average duration of an interact. The hypothesis is tested by isolating factors and measuring their correlation with the amount of interacting in the pairs of a set of N communities.

This hypothesis is supported by studies of telephoning; news circulating; travel by bus, train, or plane; R. R. express; college attendance; intermarrying; etc. Further lists of interhuman actions are suggested for investigation.

A new corroborating bit of data comes from a poll by the Washington Public Opinion Laboratory in a Seattle housing project where negro-white relations threatened violence. The tension units of verbal interaction (defined as one anti-negro opinion asserted by one white person) were observed to decrease inversely with a power of the distance from a rape site. The observed tension correlated with the formulas or curves predicting that tension at $\rho = .94$ and passed the chi-square test at the one per cent level. The tension is dimensionally analyzed as a social force and social energy.

19. The Employment of Marked Members in the Estimation of Animal Populations. Milner B. Schaefer, U. S. Fish and Wildlife Service, Honolulu, T. H.

The estimation of population numbers by marked members is an important technique in fisheries research. The number N of individuals in the population, of which T are known to be marked, may be estimated from a sample of n of which t are found to be marked. Several estimates are available, all of which reduce to $N = \frac{nT}{t}$ when the numbers are all large, but more precise formulae should be used when the numbers are not all large. An estimate of the variance of N has been derived by Karl Pearson (*Biometrika*, Vol. 20 (1928), pp. 149–174) on the basis of inverse probabilities. The sampling error may also be measured by means of confidence intervals. Formulae have been developed for estimating N from repeated samples of the same population, but no very suitable estimates of the sampling error are available in this case. For some migratory fishes marked at a point on their migration path and sampled later at another point, there exists a correlation between time of marking and time of recovery in the subsequent samples. In such case, the total number of fish marked or drawn in the subsequent samples cannot in general be regarded as random samples of the population. Where numbered tags are employed as marks, so the fish may be individually identified both when marked and recovered, a method of estimating N in this case also is suggested.

20. Non-Response and Repeated Call-Backs in Sampling Surveys. Z. W. Birnbaum and Monroe G. Sirken, University of Washington, Seattle.

In opinion-polls and other sampling surveys, a response can only be obtained from those individuals of a sample who are available for interviewing. Let $p_{.1}$ be the probability that an individual chosen at random from the population answers "yes" to a question, $p_{1.}$ that an individual is available for interviewing, and p_{11} that an individual is available and answers "yes." Usually one wishes to estimate the parameter $p_{.1}$, but from a sample it is only possible to estimate $\frac{p_{11}}{p_{1.}} = p' =$ the probability that an individual answers "yes" if he is available. Thus the total error in estimating $p_{.1}$ from a sample contains two components: the bias $p_{.1} - p'$ and the sampling error. In this paper a technique is presented in which individuals not available at a call are called upon repeatedly, up to k times. It is shown how, for a given upper bound of the total error at a prescribed probability level and a given k, it is possible to minimize the cost of the survey by optimizing the relationship between the greatest possible bias and the sampling error.

(Abstracts of papers presented at the Cleveland Meeting of the Institute on December 27–30, 1948.)

21. A Necessary Condition for a Certain Class of Characteristic Functions (*Preliminary report*). Eugene Lukacs, NOTS, Inyokern, California and Our Lady of Cincinnati College, Cincinnati, Ohio.

Let $\varphi(t) = \left\{\left(1 - \frac{t}{v_1}\right)\left(1 - \frac{t}{v_2}\right) \cdots \left(1 - \frac{t}{v_n}\right)\right\}^{-1}$ be the reciprocal of a polynomial without multiple roots. The following necessary condition is derived which $\varphi(t)$ has to satisfy in order to be the Fourier transform (characteristic function) of a distribution.

If $\varphi(t)$ is the Fourier transform of a distribution, then

1) $\varphi(t)$ has no real roots. If $b + ia$ ($a \neq 0, b \neq 0$) is a root then $-b + ia$ is also a root. That is the roots of $\varphi(t)$ are either located on the imaginary axis or are symmetrical to this axis.

2) If $b + ia$ ($a \neq 0$) is a root then there exists also at least one root $i\alpha$ so that sign α = sign a and $|\alpha| \leqq |a|$.

As a particular case one obtains the well known fact that $(1 + t^4)^{-1}$ cannot be a characteristic function.

22. Precision of Estimates from Samples Selected under Marginal Restrictions. *(Preliminary Report)*. Clifford J. Maloney, Camp Detrick, Frederick, Maryland.

Formulas are derived for estimates and for their variances computed from samples drawn at random subject only to marginal restrictions from populations classified by several characters, and estimates are made of the efficiency of such sampling plans compared to sampling with complete stratification or sampling completely at random. By means of two simple but general theorems it is shown that the variances are independent of the individual values of the character being sampled for in the population and in the sample and depend only on the first two moments for each cell of the population. It is shown that in the large sample approximation a practical scheme for actually drawing such samples can be obtained by drawing a sample of size n entirely at random and using the results of Deming and Stephan (*Annals of Math. Stat.*, Vol. 11 (1940), p. 427) to adjust the sample marginal totals to the specified values. Deficient cells will of course be filled up by additional drawings. A measure is given of the relative loss of information in sampling with marginal restrictions on the sample cell numbers compared to sampling with complete stratification. If a_{ij} represents the population mean in the ijth cell, r_i the population mean in the ith row and c_j the population mean in the jth column, and if a_{ij} is of the form $a_{ij} = a + r_i + c_j$, then marginally restricted sampling is as efficient as sampling with complete stratification. For arbitrary a_{ij} a measure of the relative efficiency compared to sampling completely at random is given by the relative degrees of freedom for the sample cell numbers. A comparison with other possible sampling procedures is given.

23. Properties of Maximum- and Quasi-Maximum Likelihood Estimates of Parameters of a System of Linear Stochastic Difference Equations with Serially Correlated Disturbances *(Preliminary Report)*. Herman Rubin, Cowles Commission, The University of Chicago.

Let $A_{ux}x_t' = u_t'$ be a complete system of linear stochastic difference equations, $x_t = (y_t, z_t)$, y_t jointly dependent, z_t predetermined. Let us suppose $u_t' + B_{vu}u_{t-1}' = v_t'$, where the random vectores v_t are serially independent and have mean zero. If the vectors v_t have the same Gaussian distribution, and the system is identified, we can obtain maximum-likelihood estimates; if the distributions are not identical Gaussian, quasi-maximum-likelihood estimates result. The identification problem is a special case of that with independent u_t and bilinear restrictions on some A_{ux}^*, if the restrictions on A_{ux}^* are linear or bilinear. As in that case, we may have multiple identification. However, the special aspects of this type of system yield some help in the discussion of the identification problem. We also observe that if the system is identified, we obtain consistency and asymptotic normality of the estimates under the same conditions as with serially independent u's for A_{ux}.

24. The Computation of Maximum Likelihood Estimates of Parameters of a System of Linear Stochastic Difference Equations with Serially Correlated

Disturbances. HERMAN CHERNOFF, Cowles Commission, The University of Chicago.

Consider the structural equations $A_{ux}x_t' = u_t'$ where the vector $x_t = [y_t, z_t]$, y_t are the jointly dependent, and z_t the predetermined variables and where u_t are serially correlated. In particular assume that the disturbances u_t satisfy the simple Markoff Process $u_t' + B_{vu}u_{t-1}' = v_t'$ where v_t is a stationary serially uncorrelated Gaussian Process with zero mean. Then we have $A_{ux}x_t' + B_{vu}A_{ux}x_{t-1}' = v_t'$. The estimates of B_{vu} and $E\{v_t'v_t\}$ can be simply expressed in terms of those of A_{ux}. It is shown that iterative gradient methods of maximization require about 2 to 3 times as much work per iteration as in the serially uncorrelated case. To apply the Newton Method about 8 times as much work per iteration is required. The Newton Method uses the second order terms of the expansion of the log of the likelihood in terms of the independent parameters of A_{ux} and these can be used to obtain estimates of the asymptotic covariance of the estimates.

25. Test Criteria for Hypotheses of Symmetry and Definiteness of a Regression Matrix for Demand Functions. UTTAM CHAND, University of North Carolina.

The importance of relations between two sets of variates (e.g. the study of relations of the prices to the quantities of several commodities) invariant under linear transformations of one set of variates contragredient to those of the other was first pointed out by Hotelling. In the study of related demand functions no suitable statistical tests have existed for testing the hypotheses of symmetry and negative definiteness of the regression matrix of prices on quantities. The test proposed here for the hypothesis of symmetry is exact and invariant under all contragredient transformations. A separate test studied for both symmetry and negative definiteness satisfies the property of invariance but its distribution depends on a nuisance parameter which is the non-zero root of a certain determinantal equation. The likelihood ratio criterion under the hypothesis of symmetry leads to a multilateral matric equation which represents $\frac{1}{2}p(p+1)$ equations of the third degree in $\frac{1}{2}p(p+1)$ unknown regression coefficients for the p-variate case, and does not admit of a unique solution.

26. The Distribution of Extreme Values in Samples whose Members are Stochastically Dependent. BENJAMIN EPSTEIN, Wayne University, Detroit.

In this paper the following problem is considered. To find the distribution of largest and smallest values in samples of size n drawn from a random process subject to the following conditions:

(i) observations $x_1, x_2, \cdots, x_n$ are taken in order from some random process.

(ii) the random process is such that successive observations x_i and x_{i+1} are jointly dependent. The joint distribution is described analytically, independently of i, by a two-dimensional d.f.

$$F_2(x, y) = \text{Prob}\ (x_i \leq x, x_{i+1} \leq y), \qquad 1 \leq i \leq n - 1.$$

(iii) $F_2(x, y) = F_2(y, x)$

(iv) Any other pairs of observations (x_i, x_{i+j}), $1 \leq i \leq n - 1$, $2 \leq j \leq n - 1$, are assumed to be independent.

The results in this paper generalize the special situation where all observations are independent. More general cases than those covered by (i)–(iv) are briefly considered.

27. On Age-Dependent Stochastic Branching Processes. RICHARD BELLMAN AND THEODORE E. HARRIS, Stanford University and The RAND Corporation, Palo Alto and Santa Monica, California.

An initial particle has a random life length T with c.d.f. $G(t)$. At death it is replaced by a random number N of similar particles; $P(N = n) = q_n$. Particles produced have the same distributions of life-length and replacement as the original one.

Let $z(t)$ = number of particles at time t, $h(s) = \sum_{n=0}^{\infty} q_n s^n$, $F(s, t) = \sum_{n=0}^{\infty} P(z(t) = n)s^n$. The integral equation $F(s, t) = \int_0^t h[F(s, t - y)]\, dG(y) + s[1 - G(t)]$ uniquely determines $F(s, t)$. When suitable restrictions are put on $h(s)$ and $G(t)$, results of Feller can be applied to study the asymptotic behavior of the moments of $z(t)$, which satisfy linear integral equations of the convolution type, and further special results on the moments can be obtained. The condition $\Sigma n q_n > 1$ and certain further restrictions insure that $z(t)/e^{bt}$ converges in probability as $t \to \infty$, where b is a certain constant. The m.g.f. $\phi(s)$ of the limiting distribution satisfies the equation $\phi(s) = \int_0^\infty h[\phi(se^{-by})]\, dG(y)$. Further restrictions imply that $\phi(s)$ is analytic in a neighborhood of $s = 0$, and that the corresponding distribution is absolutely continuous.

28. Cuboidal Lattices. G. S. WATSON, Institute of Statistics, North Carolina.

Yates has given two series of partially balanced incomplete block designs, square and cubic lattices, which enable the experimenter to test respectively k^2 and k^3 varieties in blocks of size k. Harshbarger has recently given a series of designs, rectangular lattices, which supplement Yates' square lattices.

In this paper two series of designs are given called cuboidal lattices, supplementing the cubic lattice series. They may be used to test respectively $k^2(k + 1)$ and $k(k + 1)^2$ varieties in blocks of k, when the number of reflications is a multiple of 3. Interblock information may be recovered. The first series has a relatively simple analysis and should prove useful.

This work was sponsored by the Office of Naval Research.

29. Transformations Induced by Series Approximation of Prior Probability Amplitude. ARCHIE BLAKE, Office of The Surgeon General, U. S. Army.

Consider a class A of mutually exclusive and exhaustive possible outcomes of a test. (We assume A finite; this condition can under suitable conditions be removed at a later stage by a limiting process.) For a hypothesis h, let u be the vector whose value, for each member a of A, is the square root of the prior probability of a and h jointly. This vector is called the probability amplitude; its norm, the scalar product $u'u$, is proportional to the prior probability of h, the constant of proportionality being determined by comparing the norms of the u's for all h. Let the test leave the alternatives of a subclass B of A still possible, while ruling out the members of $A - B$. Represent this test by a vector r having the value 1 on B and 0 on $A - B$. Define d on AA as a matrix equal to r on the main diagonal and zero elsewhere. The posterior probability is proportional to the form value $u'du$, the norm of the projection of u on a subspace determined by suppressing the coordinates of $A - B$. Consider the transformation $u = tv$, t being a matrix on AA and v a vector on A. Then $u'du$ takes the form $v't'dtv$. Denote $t'dt$ by e. If u is approximated as a partial sum of the series tv, i.e. by truncating v with a subclass C of A, the truncation induced on e is that with the minor on CC. (How much of the prior probability norm is

retained with a particular truncation is most easily seen if t is orthogonal, for then the transform of $u'u$ is $v'v$).

For example, in an agricultural experiment, let A be the composite of P, the class of plots, and Y, the class of possible yields on a plot. Then u takes the form of a second order tensor or matrix on PY, while d and t are fourth order tensors. For some member y of Y, it often happens that some of the initially most probable, numerous, and economically consequential hypotheses will be such that for them the values of $u(y)$ are predominantly high on some row of plots, low on another row, etc. The transformation $u = tv$ on P induces the transformation $e = t'dt$; this is R. A. Fisher's transformation, performed, however, on d instead of on the yields themselves. The truncation of v and e corresponds to Fisher's relegating the higher interactions to error. This calculation may be accompanied by a linear transformation on Y, e.g. in series of orthogonal functions. (Such series are not subject to the disadvantage of classical Gram-Charlier series, which are expressed in terms of the probability instead of its square root, that their partial sums can be at places negative.)

30. On the Utilization of Marked Specimens in Estimating Populations of Flying Insects. CECIL C. CRAIG, University of Michigan, Ann Arbor.

The experimenter catches flying insects, say butterflies, marks and immediately releases them. It is assumed that all the insects in a segregated area are equally liable to capture whether unmarked or marked, even several times, and that the population is stable for this period over which a series of captures is made. From the record of insects caught once, twice, three times, and so on, the problem is to estimate the total population. Two mathematical models which seem appropriate are considered and four methods of estimation are compared with respect to the large sample variances of the estimates they give.

31. On a Probability Distribution. MAX A. WOODBURY, University of Michigan, Ann Arbor.

In this paper the probability of x successes in n trials of an exent is computed for the case when the probability of success in a given trial depends only on the number of previous successes. The solution $P(n, x)$ satisfies the equation of partial differences

$$P(n + 1, x + 1) = (q - q_x)P(n, x) + q_{x+1}P(n, x + 1)$$

in the case when $q = 1$. The boundary conditions are obviously $P(0, 0) = 1$ and $P(n, x) = 0$ for $x < 0$ or $>n + x$. The solution of this equation is obtained by use of a generating function and $P(n, x)$ proves to be the xth term in the expansion of q^n by means of Newton's divided difference formula given the values $q_0^n, \cdots, q_x^n, \cdots, q_n^n$. Specifically, by setting $q = 1$, one obtains the result

$$P(n, x) = p_0 p_1 \cdots p_{x-1} \sum_{i=0}^{x} q_i^n / [(q_i - q_0)(q_i - q_1) \cdots (q_i - q_{i-1})(q_i - q_{i+1}) \cdots (q_i - q_x)].$$

In the case $p_x = p_0$ one has the result

$$P(n, x) = \frac{p_0^x}{x!} \frac{d^x}{dq^x} (q^n) \bigg|_{q=q_0}$$

which yields the usual result on simplification.

32. Distribution-Free Tests of Data from Factorial Experiments. G. W. BROWN AND A. M. MOOD, Iowa State College.

A device for avoiding the assumption of normality in analysis of variance problems was developed by M. Friedman (*Am. Stat. Assoc. Jour.*, Vol. 32 (1937), pp. 675–701) in which the values of the observations were replaced by their ranks.

An alternative approach is presented here in which medians are used to construct certain contingency tables, and the various null hypotheses of interest are easily tested by means of the ordinary chi-square criterion applied to such tables. These tests:

(1) Avoid the assumption of normality.

(2) Are particularly sensitive to differences in locations of cell distributions but not to their shapes.

(3) Usually require very little arithmetic computation.

The tests and the relevant distribution theory have been worked out for some of the simpler experimental designs.

33. On Sums of Symmetrically Truncated Normal Random Variables. FRED C. ANDREWS AND Z. W. BIRNBAUM, University of Washington, Seattle.

Let X_a be the random variable with the probability density

$$f_a(X) = Ce^{-X^2/2} \quad \text{for} \quad |X| \leq a, \qquad f_a(X) = 0 \quad \text{for} \quad |X| > a,$$

and let $S_a^{(n)} = \sum_{j=1}^{n} X_a^{(j)}$ where $X_a^{(1)}, \cdots, X_a^{(n)}$ are independent determinations of X_a. The problem considered is: for given n, $T > 0$, $\epsilon > 0$, determine a such that $P(|S_a^{(n)}| \geq T) = \epsilon$. The exact solution of this problem would require laborious computations. In this paper a method is given for obtaining approximate values of a which are "safe" i.e. such that $P(|S_a^{(n)}| \geq T) \leq \epsilon$.

34. On the Foundation of Statistics. MAX A. WOODBURY, University of Michigan, Ann Arbor.

The results on this paper are part of the author's University of Michigan dissertation, "Probability and Expected Values." The work covered by this paper was sponsored by the Office of Naval Research. One may take the notion of an expected value as the basis for the theory of Statistics; i.e. a linear functional on a linear space of random variables (real valued functions defined over a population). The space is called statistical if it contains all constant functions and the expected value of such constant functions is just the constant and if the expected value of a non-negative function is non-negative. A statistical space is called strong if it contains with a random variable also the random variable whose values are the absolute values of the given random variable. Every expected value defines a probability measure over a quorum of subsets of the population and it is shown that the integral of the random variable, if it exists, coincides with the expected value. Further it is shown that if the statistical space is strong the integral necessarily exists and also that a necessary and sufficient condition that the quorum be a field is that the statistical space be strong.

35. Finitely Additive Probability Functions. MAX A. WOODBURY, University of Michigan, Ann Arbor.

The results in this paper are part of the work in the author's University of Michigan dissertation, "Probability and Expected Values." The work covered by this paper was sponsored by the Office of Naval Research. A quorum is a family of sets that contains with each pair of disjoint sets also their union and also the complement of any of its sets. Trivially a quorum is required to contain at least one set and hence at least the universe set or population and the empty set. An extension of the notion of a finitely additive probability measure function to quorums is given and proved to be equivalent to the usual definition in case the quorum is a field of sets. The extension of a quorum of sets relative to the probability measure function is investigated using the properties of the inner and outer measure. The upper and lower integrals are defined and a condition for the existence of the integral is given. When the quorum is a field it is shown that integrability of a function implies the existence of the distribution function. This last result is well known in the case where the probability measure function is completely additive.

36. **On Inverting a Matrix via the Gram-Schmidt Orthogonalization Process.** MAX A. WOODBURY, University of Michigan, Ann Arbor.

The application of the classical Gram-Schmidt orthogonalization process to the factorization of a correlation matrix is accomplished by considering the inner product $[x, y] = E(xy)$ in the linear space determined by the statistical variables $x_1, x_2, \cdots, x_n$. In this way a representation of the original set of statistical variables in terms of an orthonormal set is obtained. (By an orthonormal set we mean a set $\xi_1, \xi_2, \cdots, \xi_n$ such that $E(\xi_i\xi_j) = 0$ for $i \neq j$ and $E(\xi_i^2) = 1$.) The matrix of coefficients $B = (b_{ij})$, where $x_i = \sum_{j=1}^{i} b_{ij}\xi_j$, has the property that $C = BB'$ where $C = (E(x_ix_j))$ and $'$ denotes the transpose. Further the matrix B is triangular hence B^{-1} is readily computed, from which one obtains at once $C^{-1} = (B^{-1})'B^{-1}$. The quantities b_{ij} are readily obtainable by the method of determinants (Dwyer and Waugh, *Annals of Math. Stat.*, Vol. 16 (1945), pp. 259–271, cf. pg. 264) formerly called the method of multiplication and subtraction with division.

37. **Certain Properties of the Multiparameter Unbiased Estimates.** G. R. SETH, Iowa State College.

If $\theta^* = (\theta_1^*, \theta_2^*, \cdots, \theta_q^*)$ is an unbiased vector estimate of $\theta = (\theta_1, \theta_2, \cdots, \theta_q)$ in the density function $p(x_1, x_2, \cdots, x_n; \theta_1, \theta_2, \cdots, \theta_q)$ having the smallest concentration ellipsoid among the class of unbiased estimates of θ, and further if ϵ is any statistic of q components having $E(\epsilon) = 0$ and finite covariance matrix, then ϵ is uncorrelated with θ^*.

If a set of sufficient statistics $(T_1, T_2, \cdots, T_p)$, $p \leq q$, exists for estimating θ, then corresponding to any unbiased vector estimate ϕ^* of θ, there exists an unbiased estimate of θ depending on $T_1, T_2, \cdots, T_p$ alone, where the latter has a concentration ellipsoid equal to or contained in that of the former.

When $q = 1$, and ϕ^* has the smallest variance among the class c formed by unbiased estimates of θ which are functions of θ^* having a finite variance, and the set of polynomials with respect to the distribution function of ϕ^* is complete, then ϕ^* is the only element in the class c. For $q > 1$, the result holds when the "variance" is replaced by the "concentration ellipsoid."

38. **A Class of Lower Bounds for the Variance of Point Estimates.** DOUGLAS CHAPMAN, University of California, Berkeley.

A class of lower bounds for the variance of point estimates is derived by means of the calculus of finite differences under very weak restrictions and it is shown that they give valid lower bounds for certain parameter estimation problems for which the Cramér-Rao formula is invalid. In some cases even when the latter lower bound exists a sharper lower bound may be found in the class here defined. On the other hand when it exists, the Cramér-Rao lower bound is asymptotically superior to any of this class.

39. **Standard Errors and Tests of Significance for Interpolated Medians.** CHURCHILL EISENHART AND MIRIAM L. YEVICK, National Bureau of Standards.

If a sample of N observations is grouped by a sequence of class intervals with boundaries $-\infty, \cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots, +\infty$, where x_0 is the largest boundary point for which the observed 'fraction below', p_B, is less than $\frac{1}{2}$, and x_1 is the smallest boundary point for which the observed 'fraction above', p_A, is less than $\frac{1}{2}$, so that the observed 'central fraction', p_C, between x_0 and x_1 is positive, then, at least for the case of N large, standard textbooks take as the median of the grouped data the interpolated median,

$$m = x_0 + b(x_1 - x_0)$$

where

$$b = (\tfrac{1}{2} - p_B)/p_C .$$

The literature is silent regarding the sampling properties of such medians, and regarding tests of significance appropriate to them. Let P_B and P_C be the population fractions below x_0, and between x_0 and x_1, respectively, and let u and β be the population analogs of m and b obtained by replacing p_B and p_C in the above equations by P_B and P_C, respectively. It is shown that m is asymptotically normally distributed about u so defined with asymptotic variance given by

$$\frac{1}{NY_C^2}\left[P_B(1 - P_B) - 2\beta P_B P_C + \beta^2 P_C(1 - P_C)\right]$$

where

$Y_C = \dfrac{P_C}{x_1 - x_0}$ = ordinate of 'central rectangle' of 'population histogram'. The classical formula for the variance of a median can be obtained as the limit of the above when $(x_1 - x_0) \to 0$ *with* $P_B \to \frac{1}{2}$.

In addition, tests of hypotheses regarding the value of the 'interpolated median of the population', u, and regarding the difference, $u_2 - u_1$, of the interpolated medians of two populations, are developed (1) by utilizing the above asymptotic results, and (2) by utilizing the Neyman-Pearson likelihood-ratio-test approach.

40. Some Efficient Range-Estimates of Variation. NILAN NORRIS, Hunter College, New York.

The commonly used sample range (in the sense of the difference between the largest and smallest of the variates) is one of an unlimited number of range or difference-measures which can be used to scale parent populations. For samples drawn from a Type III universe, the maximum-likelihood estimate of dispersion is given by $A - G$, where A is the sample arithmetic mean and G is the sample geometric mean. For samples drawn from a Type V universe, a 100% efficient estimate of absolute variation is given by $G - H$, where G is the sample geometric mean and H is the sample harmonic mean. Under certain general conditions usually fulfilled, the standard errors of both of these range-measures of absolute dispersion may be estimated from expressions obtained by application of the Laplace-Liapounoff theorem. The two parametric methods of estimating absolute variation as developed in this paper are likely to be most useful when the form of the parent universe is known, and it is either too expensive or impossible to obtain samples large enough to permit the use of inefficient estimates. An example of such a case is the learning curve encountered in the analysis of frequency of occurrence of aircraft accidents by hours of flying experience of pilots in training. E. J. G. Pitman, *Proc. Camb. Phil. Soc.*, Vol. 33 (1937), pp. 217–218, has discussed the scaling of the Type III distribution. The method of scaling given by Pitman differs from the method of estimation developed in this paper for the Type III universe.