Integrating by parts, we find that $S_n = \frac{1}{2}\int_0^1 R_n(t)\, dt$. By the result proved in Section 2 this last expression converges stochastically to

$$\frac{1}{2}\int_0^1 \left[1 - \int_A^B f(x)\, e^{-tf(x)}\, dx\right] dt = \frac{1}{2}\left[1 + \int_A^B e^{-f(x)}\, dx - (B - A)\right].$$

Therefore Ω_n converges stochastically to $\frac{1}{2}(1 + A - B) + \int_A^B e^{-f(x)}\, dx$. For the special case $A = 0$ and $B = 1$, this is essentially the result contained in theorems 3 and 4 of [1].

REFERENCE

[1] B. Sherman, "A random variable related to the spacing of sample values," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 339–361.

Note added in proof. Professor Julius Blum has pointed out that Lemma 2 holds with the words "converges stochastically" replaced by "converges with probability one." Then it is easily seen that all the results above hold when this replacement is made.

ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Chapel Hill meeting of the Institute, April 22–23, 1955)

1. **Estimation of Location and Scale Parameters by Order Statistics from Singly and Doubly Censored Samples. Part I. The Normal Distribution up to Samples of Size 10.** A. E. Sarhan and B. G. Greenberg, University of North Carolina.

The variances and covariances of the order statistics for samples of sizes ≤ 20 from a normal distribution were calculated to 10 decimal places from Teichroew's tables of the expected value of the product of two order statistics. By the use of these values, and with the table of expected values of Rosser, the best linear estimates of the mean and standard deviation were calculated from singly and doubly censored samples up to samples of size 10. This was accomplished by applying the method of least squares to the linear combination of the ordered known observations to obtain unbiased estimates with minimum variance. The variances of the estimates were also calculated. An alternative linear estimate was derived for larger values of n which can be used to obtain estimates from doubly censored samples.

2. **An Application of Chung's Lemma to the Kiefer-Wolfowitz Stochastic Approximation Procedure.** Cyrus Derman, Syracuse University.

Let $M(x)$ be a strictly increasing regression function for $x < \theta$, and strictly decreasing regression function for $x > \theta$. Kiefer and Wolfowitz (Ann. Math. Stat., Vol. 23 (1952), pp. 462–466) suggested a recursive scheme for estimating θ. They proved, under certain regularity conditions, that their scheme converges stochastically to θ. Their conditions exclude the case $M(x) = K - K' (x - \theta)^2$ where K and K' are constants ($K' > 0$). Conditions, which do not exclude the above case, are given here for their scheme to converge stochastically to θ. Under stronger conditions (the above case still not excluded) convergence to the normal distribution is proved. The main tool used in the analysis is a lemma due to Chung (Ann. Math. Stat., Vol. 25 (1954), pp. 463–483).

3. Simplified Estimators Based on Order Statistics. (Preliminary Report.) Benjamin Epstein, Wayne University.

Best linear unbiased estimates based on order statistics have been given recently by A. E. Sarhan [*Ann. Math. Stat.*, Vol. 25 (1954), pp. 317–328] for the mean and standard deviation of a number of distributions. It is assumed in that paper that all observations are known. In a paper given at the Berkeley meeting in December, 1954, Sarhan considered the same estimation problem in the case where some of the ordered observations may be missing. Here we give unbiased estimators which are much simpler in the sense that they can be expressed in terms of only a few of the order statistics about which we have information. Efficiency of the suggested estimators is high for small sample sizes.

4. Distribution of the Difference Between the Two Largest Sample Values. (Preliminary Report.) A. Zinger and J. St-Pierre, University of Montreal.

A decision procedure to select the population with the largest mean, proposed by R. C. Bose and J. St.-Pierre (Ann. Math. Stat., Vol. 25 (1954), p. 813), involves the auxiliary statistic $y = x_{(0)} - x_{(1)}$, where $x_{(0)}$ and $x_{(1)}$ are respectively the largest and second largest values in a sample of $n + 1$ variates. The distribution of this statistic is obtained with a method simpler than the one already used by the senior author. The final result involves iterated integrals of the normal density over simple limits. The general form can be easily reduced to neat expressions in the case of lower dimensions. The densities of 3, 4, and 5 dimensions have been extensively tabulated and a recursion formula established between the densities. The establishment of a recursion formula in the general case is being worked on.

5. Some Continuous Monte Carlo Methods for the Dirichlet Problem. Mervin E. Muller, Cornell University.

Monte Carlo techniques are introduced using stochastic models which are Markov processes. This material includes the N-dimensional spherical, general spherical, and general Dirichlet domain processes. These processes are proved to converge with probabily 1 and thus yield direct statistical estimates of the solution to the N-dimensional Dirichlet problem. The results are obtained without requiring any further restrictions on the boundary or the function defined on the boundary in addition to those required for the existence and uniqueness of the solution to the Dirichlet problem. A detailed study is made for the N-dimensional spherical process. This includes a study of the order of the average number of steps required for convergence. Asymptotic confidence intervals are obtained. When computing effort is measured in terms of the order of the average number of steps required for convergence, the often made conjecture that the computing effort of a Monte Carlo procedure should be a linear function of the dimensionality of the problem is shown to be true for the cases considered. Comments are included regarding the application of these processes on digital computers. Truncation methods are suggested.

6. On the Distribution of the Number of Successes in Independent Trials.

WASSILY HOEFFDING, University of North Carolina.

Let S be the number of successes in n independent trials. Let p_i be the probability of success in the ith trial. The problem is considered of finding the maximum or the minimum of the expected value of a function of S when $E(S) = np$ is fixed, $0 < p < 1$. It is well known that the variance of S attains its maximum when the p_i are all equal. It is shown: (i) for any two integers b, c such that $0 \leqq b \leqq np \leqq c \leqq n$ the probability $P(b \leqq S \leqq c)$ attains its minimum if and only if all the p_i are equal, unless $b = 0$, $c = n$; (ii) for any strictly convex function g the expected value $Eg(S)$ attains its maximum if and only if all the p_i are equal. The maximum and the minimum of $P(S \leqq c)$, $0 \leqq c \leqq n$, are determined. These results are obtained with the aid of some theorems concerning the extrema of $Eg(S)$, where g is an arbitrary function. For example, the maximum and the minimum of $Eg(S)$ are attained at points $(p_1, \cdots, p_n)$ whose coordinates take on at most three different values, only one of which is distinct from 0 and 1. Statistical applications of (i) and (ii) are pointed out.

7. On the Solution of Truncated and Censored Sample Estimating Equations for Normal Populations.

A. C. COHEN, JR., University of Georgia.

To obtain maximum likelihood estimates of the mean and standard deviation of a normally distributed population from doubly truncated and from doubly censored samples, it is necessary to carry out the simultaneous solution of a pair of rather complicated nonlinear estimating equations. In this paper, iterative techniques for solving these equations are examined, and a procedure is developed which yields solutions of specified accuracy with less computational effort than required by other methods previously employed. A chart has been prepared which, for doubly truncated samples, permits a quick graphic solution to a degree of accuracy that is adequate for many purposes, and which provides a good first approximation for subsequent improvement through iteration when greater accuracy is demanded. A chart has also been devised to permit a quick graphic solution in the case of singly censored samples.

8. The Modified Mean Square Successive Difference and Related Statistics.

SEYMOUR GEISSER, University of North Carolina.

In estimating the variance of a normal population, one uses the sample variance because of its optimum properties. In certain cases where there is an indeterminable trend in the data, it has been thought useful to estimate the variance by another statistic, namely, the mean square successive difference, the mean of the squared first differences, which under certain conditions, eliminates a good deal of the trend and is less biased than the sample variance. An explicit form of the exact distribution of this statistic seems, at least for the present, too difficult to obtain. However, by applying the device of Durbin and Watson, that is, by dropping from the mean square successive difference the middle term for an even number of observations and the two middle terms for the odd case, it is found that the quadratic form has double roots, thus making it possible to obtain the exact distribution in terms of elementary functions. In addition, one defines analogues of the Student t and the Fisher F using similarly modified statistics and proceeds to derive their exact distributions when the observations are independent and in a specific dependency case which has several properties in common with the stationary Markov process.

9. The Distribution of the Ratios of Certain Quadratic Forms in Time Series. (By Title.) SEYMOUR GEISSER, University of North Carolina.

In testing the hypothesis that successive members of a series of observations are serially correlated, a number of statistics have been proposed by various statisticians. R. L. Anderson gave the first exact distribution of a serial correlation coefficient using a circular definition. J. Durbin and G. Watson gave the exact distributions of several other statistics using double root methods. In this paper the work of Durbin and Watson has been extended for a non-null case of one of their statistics. Also, by introducing a new model, the exact distribution of a modified form of the von Neumann ratio has been derived in the non-null case. It has also been shown that this ratio provides a "best" test for the parameter involved.

10. The "Inefficiency" of the Sample Median for many Familiar Symmetric Distributions. J. T. CHU, University of North Carolina.

If the pdf of a certain distribution is symmetric and has an absolute maximum at the point of symmetry, a lower bound for var $\tilde{x}$, the variance of the sample median $\tilde{x}$ of a sample of size $2n + 1$ is $(2n + 1)/(2n + 3)$ multiplied by the variance of the asymptotic distribution of $\tilde{x}$ (which is normal). Therefore if sample size is not too small, the asymptotic variance of $\tilde{x}$ is for all practical purposes a lower bound for var $\tilde{x}$. If $\tilde{x}$ is asymptotically less efficient than $\bar{x}$, it is probable that $\tilde{x}$ is less efficient than $\bar{x}$ for most finite samples as well. For many symmetric distributions familiar to statisticians, such as triangular, Student's t, symmetric β, and Cauchy type distributions $(f(x) = C_\alpha/(1 + |x|^\alpha), -\infty < x < \infty, \alpha \geqq 4.65)$, not counting normal and rectangular distributions, it is shown that $\tilde{x}$ is for most sample sizes less efficient than $\bar{x}$.

11. On Some Stochastic Models of Behavioral Interaction of Organization Theory. DAVID ROSENBLATT, American University.

This paper treats certain stochastic models of behavioral interaction which constitute applications of a general approach to a calculus of behavior. Participants or groups in organizations are viewed as entities provided with a stochastic preference (or threshold) apparatus; entities engage in adaptive or reactive behavior by adjustment of the stochastic processes governing their own activities. Transition probabilities are modified in accord with experience of "relative success" of the entities in accord with certain criteria, e.g., organizational norms or observed actions of other entities. Modes of behavior are sequentially reinforced or inhibited as a result of the moves of entities in the course of interaction. Various types of memory structure are explicitly introduced. The "performance characteristics" of a given structure of interaction may be summarized by the expected individual and joint probability distributions of behavioral activity of each entity at each interaction transaction γ_k $(k = 1, 2, \cdots)$. Algorithms are developed for the determination of these "performance characteristics." For certain parametric characterizations (r entities, n_j decision alternatives and m_j preference valuation alternatives for the jth entity, $j = 1, 2, \cdots, r$), the algorithms lead to closed-form expressions. In the simplest cases, these become systems of linear difference equations. Many of the asymptotic results of stochastic learning theory may be readily obtained by specialization of the present models. (Work supported by the Office of Naval Research.)

12. On Inverting a Class of Patterned Matrices, Part I. S. N. ROY and A. E SARHAN, University of North Carolina.

In this note, inverses are given of a class of patterned matrices that occur in different sectors of statistics, e.g., least squares solutions relating to problems of estimation of population parameters by ordered or unordered observations, analysis of variance and covariance, response surfaces, etc. The actual examples given here are illustrative and will be followed up later by other examples. In the technique given here of obtaining such inverses, use is made of the fact that (i) a non-singular square matrix has a unique inverse and (ii) for the class of patterned matrices considered it is possible to guess a form for the inverse with a few unknown (and thus flexible) parameters which could then be determined by equating to the identity matrix the product of the original matrix and the inverse that is guessed. At the moment the guess is just intuitive, but the authors believe there is a deeper calculus behind the whole thing, which may emerge later and thus make the inversion of such matrices an entirely trivial problem.

13. **Convergence Properties of a General Stochastic Approximation Process. (Preliminary Report.)** Donald Burkholder, University of North Carolina.

Theorem. *Let* $\{R_n\}$ *be a sequence of Borel measurable functions,* θ, σ^2, c, d, x_1 *real numbers,* Q *a function from the positive numbers to the natural numbers, and* $\{a_n\}$ *a positive number sequence such that*: (i) *for each natural number* n *and each real number* x *there is a random variable* $Z_n(x)$ *such that* $EZ_n(x) = R_n(x)$, Var $[Z_n(x)] \leqq \sigma^2$, *and* $|R_n(x)| \leqq c + d\,|x|$; (ii) if $0 < \epsilon < \infty$ *and* $0 < \delta_1 < \delta_2 < \infty$, *then* $(x - \theta)\, R_n(x) > 0$ *for* $|x - \theta| > \epsilon$, $n > Q(\epsilon)$, *and* $\Sigma a_n [\inf_{\delta_1 \leqq |x-\theta| \leqq \delta_2} |R_n(x)|] = \infty$; (iii) $\Sigma a_n^2 < \infty$. *Then the sequence* $\{x_n\}$ *of random variables defined recursively by* $x_{n+1} = x_n - a_n Z_n(x_n)$ *converges to* θ *with probability one.* The proof involves methods similar to those used by Blum, *Ann. Math. Stat.*, Vol. 25 (1954), pp. 382–386. Some immediate corollaries to the theorem are: (1) The Robbins-Munro process converges with probability one (Blum's Theorem 1). (2) The Kiefer-Wolfowitz process converges with probability one (under conditions less restrictive than those heretofore published; for instance, a regression function M where $M(x) = e^{-x^2}$ or $M(x) = -x^2$ is permissible). (3) There exists a strongly consistent sequence of estimates of the mode of a density function under fairly general conditions. (4) There exists a strongly consistent sequence of estimates of a root of a regression equation even when the variances around the regression line may not exist. The theorem, and hence also each of the corollaries (1), (2), and (3), has been generalized to the case where the number θ does not exist uniquely. This permits, for instance, the use of the Robbins-Munro process in the problem of estimating a quantile of a distribution function when the quantile is not unique.

14. **Distribution of Rounding Off Errors in Some Numerical Processes, Part I.** A. E. Sarhan, University of North Carolina.

The distributions of rounding-off errors in a product, quotient, raising to a power process, several combinations of these processes, and other special cases are derived. The moment generating functions and the first four moments are calculated. Expressions for the significance points at a given level are provided.

15. **On a measure of the Information Provided by an Experiment. (Preliminary Report.)** D. V. Lindley, University of Cambridge and University of Chicago.

An experiment consists in the observation of a random variable x with probability density $f(x \mid \theta)$, where θ is an unknown parameter. Let $p(\theta)$ be the probability density of θ, expressing the knowledge of θ prior to performing the experiment. Then the average amount of information provided by the experiment is defined to be $\int\int f(x \mid \theta)p(\theta) \log \{f(x \mid \theta)/p(x)\}$

$\cdot dx\, d\theta$, where $p(x) = \int f(x \mid \theta)p(\theta)\, d\theta$. This definition is suggested by the corresponding definition of Shannon's in connection with the rate of transmission of information in communication engineering. It is shown that it is always nonnegative, is not reduced by consideration of sufficient statistics alone, and if x and y are independent random variables for each θ, then the experiment in which y is observed is more informative if carried out before x is observed than if carried out after x has been observed. The definition enables comparisons to be made of different experiments but these comparisons are unlike those considered by Blackwell in that the losses in pursuing various courses of action are not considered. The ideas are therefore more relevant to the inference problem than the decision problem. Examples of the use of the definition, in particular for multivariate normal densities, are considered.

16. A Comparison Between Alternative Techniques Using Supplementary Information in Sample Survey Design. (Preliminary Report.) El Mahdy Said, North Carolina State College.

Three alternative methods of incorporating the advance information available on a variable X in the design of finite population sample surveys to estimate aggregate or mean values for a variable Y are studied. For a given sample size n and ignoring cost, the systems are (a) stratification with $s = n/2$ strata, (b) sampling without replacement with unequal selection probabilities such that the probability of including units $u_i u_j$ together in the sample is $P(u_i u_j) = n(n - 1)X_i X_j(1/T_x - X_i + 1/T_x - X_j)/2T_x$ where $T_x = \sum_{i=1}^{N} X_i$ and the estimator used is $\bar{y}_p = \sum_{i=1}^{n} y_i/P(u_i)$, (c) ratio estimate. Formulas for the mean square errors of the estimators are derived for both linear and curvilinear relationships between X and Y. Exact comparison for $n = 2$, using discrete counterparts of some Pearson type III distributions for X, showed that (b) is superior to (c) except for ρ_{xy} very close to 1. Approximate comparisons were obtained for $n > 2$ assuming large N and continuous type III distributions. Variance with stratification was approximated by using the uniform distribution for X within the first $(s - 1)$ strata. Method (a) was found to be superior to (b) and (c) except when c_y (the coefficient of variation of Y) is in the neighbourhood of c_x. In certain instances the issue depends on ρ_{xy} alone; in others, on combination of ρ_{xy} and c_y.

17. The Canonical Distribution of the Non-central Rectangular Co-ordinates. Miss Aleyamma George, University of North Carolina and University of Travancore.

This paper is concerned with a matrix method of (a) deriving the canonical distribution of the non-central rectangular coordinates directly from the probability law for random samples from a p-variate normal population for the cases (i) one non-zero root and (ii) two non-zero roots for general p and (b) using this to obtain the canonical non-central Wishart distribution obtained by T. W. Anderson and M. A. Girshick for the same cases.

18. Confidence Interval Estimation for the Parameters of a Rectangular and an Exponential Population in Terms of Complete or Censored Samples. (Preliminary Report.) (By Title.) S. N. Roy and A. E. Sarhan, University of North Carolina.

In previous papers by the second author point estimation in the above situations was discussed. In the present paper, using the techniques given in previous papers by the first author, confidence intervals are given for parameters and certain statistically important parametric functions in the case of rectangular and exponential populations (both general and special forms) in terms of complete samples. A method is also indicated of generalizing to the case of censored samples and to certain other populations as well.

19. Some Generalizations of Analysis of Variance and Covariance to the Case of Discrete Variates or of Grouping in Qualitative Categories. (By Title.) S. N. Roy and Marvin Kastenbaum.

Associated with any design there is a general cell—say an m-dimensional cell—in which we have a number of observations classified into, say, p-dimensional cells where p is the number of "variates" or "number of ways of classification." The whole data can now be regarded as being arranged in an $(m + p)$-way classification such that the m-cells and p-cells are, as it were, two different kinds of marginals. Bearing in mind the nature of this difference, the usual estimation and testing procedures in analysis of variance and covariance are generalized to the situations indicated above. The generalization of multivariate analysis of variance of means to the above situations will be discussed in a later paper.

20. Some Analytic Properties of Markoff Functions: Denumerable Case. (By Title.) Donald G. Austin, Syracuse University.

Let $p_{ij}(t)$, $0 < t < \infty$, $i, j = 1, 2, \cdots$, be the stationary transition matrix of a Markov chain. The author extends his earlier result (to appear in *Proc. Nat. Acad. Sci.*) that $Dp_{ii}(0) = -q_i > -\infty$ implies $p_{ij}(t)$ has a continuous derivative on $[0, \infty]$. It is shown that if $q_j < \infty$, $p_{ij}(t)$ has a continuous derivative on $[0, \infty]$. In either case $\lim_{t\to\infty} Dp_{ij}(t)$ exists and is equal to 0. If $q_i < \infty$, then $Dp_{ij}(t + s) = \Sigma_k Dp_{ik}(t)p_{kj}(s)$; if q_i, $q_j < \infty$, then $Dp_{ij}(t + s) = \Sigma_k p_{ik}(t) Dp_{kj}(t)$ for $t, s > 0$.