ABSTRACTS OF PAPERS

Presented at the Madison Meeting of the Institute, September 7–10, 1948

1. On Distribution-free Confidence Intervals (Preliminary Report). WASSILY HOEFFDING, University of North Carolina, Chapel Hill.

Let $\theta(F)$ be a functional of a distribution function (d.f.) $F(x)$ (where x is a real number or a vector), defined over a class $\mathfrak{D}$ of d.f.'s; O_n a random sample from a population with d.f. $F(x)$; $\underline{\theta}_n \leq \bar{\theta}_n$ two functions of O_n; and $\alpha_n = Pr\{\underline{\theta}_n \leq \theta(F) \leq \bar{\theta}_n\}$. Conditions are studied under which, given α, $0 < \alpha < 1$, we have either $\alpha_n = \alpha$ or $\alpha_n \geq \alpha$ or $\alpha_n \rightarrow \alpha$, for all $F(x)$ in $\mathfrak{D}$, where $\mathfrak{D}$ is defined independently of the functional form of $F(x)$. Under fairly general conditions we can obtain by "studentization" confidence limits $\underline{\theta}_n$, $\bar{\theta}_n$ such that $\lim_{n\rightarrow\infty} \alpha_n = \alpha$, and $\gamma = \lim_{n\rightarrow\infty} E\sqrt{n}(\bar{\theta}_n - \underline{\theta}_n)$ exists; γ is minimized by using a least variance estimate of $\theta(F)$. If there exists a function $\kappa(\theta)$ such that var $T_n \leq \kappa^2(\theta)n^{-1}$ if $\theta(F) = \theta$, for all F in $\mathfrak{D}$, we can define confidence limits with a positive lower bound for α_n. This applies to a number of population characteristics estimated by rank order statistics, such as the coefficients ρ' and τ (estimated by Spearman's and Lindeberg-Kendall's rank correlation coefficients, respectively). In certain cases (including ρ' and τ), $\theta(F)$ admits a binomially distributed estimate; then exact confidence limits can easily be obtained. This research was done under an Office of Naval Research contract.

2. On Certain Statistics for Samples of 3 from a Normal Population. JULIUS LIEBLEIN, National Bureau of Standards, Washington.

In analytical chemistry three determinations are frequently made. Sometimes the average of only the two *closest* results is reported, the remaining observation being rejected as anomalous. In preparing a critique of this procedure, Dr. W. J. Youden encountered a need for information on certain properties of the distributions of the statistics $(x' - x'')/(x_3 - x_1)$, $(x' + x'')/2$, and $(x' - x'')/2$, where x' and x'' $(x' \geq x'')$ are the two *closest* of the three determinations. This paper shows how these statistics differ from the ones heretofore treated involving "fixed" order statistics; gives the distribution of these statistics in random samples of 3 from a normal universe; and lists values of certain of the moments of their distributions.

3. On Multinomial Distributions with Limited Freedom: A Stochastic Genesis of Pareto's and Pearson's Curves. MARIA CASTELLAIN, University of Kansas City.

The purpose of this paper is to investigate the most probable configuration of N random elements to be distributed in $K(K < N)$ class intervals, where known forces are acting. We shall call these intervals of energy, using the terminology of statistical mechanics.

We will prove that the most probable configuration is a configuration of statistical equilibrium since its probability of occurring converges to 1 as N becomes infinitely large.

The main purpose of this paper is to discover which forces of attraction, operating in the intervals of energy, give Pareto's and Pearson's curves when statistical equilibrium is reached.

We will consider a random variable $Y(t)$, t being an independent variable, obeying a multinomial distribution law with limited freedom, and we will exploit the familiar process of statistical mechanics. The equation of the frequency curves corresponding to the equilibrium stage of the statistical experiment will be shown.

4. Fitting Generalized Truncated Normal Distributions. HAROLD HOTELLING, University of North Carolina, Chapel Hill.

In a sample from a p-dimensional normal distribution only those individuals are supposed to be observed which fall in a specified but arbitrary set A of positive measure. For estimating the parameters the method of moments is proved equivalent to that of maximum likelihood and therefore efficient. The problem is thus reduced to that of expressing the parameters of the normal distribution in terms of the moments of the truncated distribution. This however is not generally possible in simple explicit form. Methods are presented for dealing numerically with several special cases, including those in which A is a linear interval or a parallelogram.

5. On the Distribution of the Two Closest Observations Among a Set of Three Independent Observations. G. R. SETH, Iowa State College.

Let x_1, x_2, x_3 ($x_1 < x_2 < x_3$) be three independent ordered observations from a population having a probability density function $f(x)$. Let x', x'' ($x' < x''$) be the two closest, then the probability density function of x', x'' is given by

$$6 \cdot f(x') \cdot f(x'')[1 + F(2x'' - x') - F(2x' - x'')]$$

where

$$F(x) = \int_x^\infty f(x)\, dx.$$

In the case $f(x)$ is a normal distribution with unit variance, the joint distribution of $y = x'' - x'$ and $z = \dfrac{x'' - x'}{x_3 - x_1}$ is obtained as

$$\frac{2\sqrt{3y^2}}{\pi z^2} \exp\left[-\frac{y^2(1 - z + z^2)}{3z^2}\right].$$

This problem is of interest in cases where the conclusions are to be based on a set of three observations and one of the observations is to be rejected in the analysis of the data.

6. The Derivation of Certain Recurrence Formulae and their Application to the Extension of Existing Published Incomplete Beta Function Tables. T. A. BANCROFT, Alabama Polytechnic Institute, Auburn (presented by title).

The objects of the paper are: (1) to give a number of new recurrence formulae in the incomplete beta function derived by a new method, and (2) to indicate how these new formulae have been used to obtain new tables of the incomplete beta function that are outside the range of the p and q values given in the existing published tables.

The recurrence formulae have been derived by considering the incomplete beta function as a special case of the hypergeometric series, thus

$$B_x(p, q) = \frac{x^p}{p} F(p, 1 - q, p + 1, x),$$

where the usual form of the hypergeometric series is

$$F(a, b, c, x) = 1 + \frac{a \cdot b}{c}\frac{x}{1!} + \frac{a(a+1) \cdot b(b+1)}{c(c+1)}\frac{x^2}{2!} + \frac{a(a+1)(a+2) \cdot b(b+1)(b+2)}{c(c+1)(c+2)}\frac{x^3}{3!} + \cdots.$$

This series converges for $|x| < 1$, and $x = 1$, if and only if $a + b < c$. Certain recurrence formulae for $F(a, b, c, x)$ are then directly converted for use with $B_x(p, q)$, or in the so-called normalized form $I_x(p, q)$, provided $c = a + 1$. All conditions have been satisfied by setting $a = p, b = 1 - q, c = p + 1$, and $q > 0$.

For example, using the above mentioned methods we may obtain, among many others, the recurrence formulae:

(i) $xI_x(p, q) - I_x(p + 1, q) + (1 - x)I_x(p + 1, q - 1) = 0$,

(ii) $(p + q - px)I_x(p, q) - qI_x(p, q + 1) - p(1 - x)I_x(p + 1, q - 1) = 0$,

(iii) $qI_x(p, q + 1) + pI_x(p + 1, q) - (p + q)I_x(p, q) = 0$.

Formula (i) is essentially the basic recurrence formula used to obtain Karl Pearson's tables. An indication of formula (iii) in another form was given by the author in the paper "On Biases in Estimation Due to the Use of Preliminary Tests of Significance," *Annals of Math. Stat.*, Vol. 15 (1944), p. 194, and a direct proof was later given by the author in "Note on an Identity in the Incomplete Beta Function," *Annals of Math. Stat.* Vol. 16 (1945), pp. 98–99. All of the material in the present paper, however, is new, including recurrence formulae and tables and the mathematical method of derivation.

7. Asymptotic Studentization in Testing of Hypotheses. HERMAN CHERNOFF, Cowles Commission for Research in Economics.

If H is a hypothesis for which $t \le c_1(\theta)$ would be a good test if the value of the nuisance parameter θ were known and $\hat{\theta}$ is an estimate of θ, then the following method of asymptotic studentization (obtaining critical regions of almost constant size) was suggested by Wald. Consider $t \le \varphi(\hat{\theta})$ where $\varphi(\hat{\theta}) = c_1(\hat{\theta}) + \cdots + c_s(\hat{\theta})$ and $Pr\{t \le c_1(\theta)\} = \alpha$, $Pr\{t - c_1(\hat{\theta}) \le c_2(\theta)\} = \alpha$, $\cdots$ $Pr\{t - c_1(\hat{\theta}) - \cdots - c_r(\hat{\theta}) \le c_{r+1}(\theta)\} = \alpha$. It is shown that under reasonable conditions this test, and various modifications, designed for those cases where the $c_r(\theta)$ are difficult to obtain exactly have the asymptotic property that $Pr\{t \le \varphi(\hat{\theta})\} = \alpha + O(N^{-s/2})$ where N is the size of the sample involved or an analogous variable. This property can be extended to the case where θ is a k-dimensional variable.

8. Completeness, Similar Regions, and Unbiased Estimation. (Preliminary Report.) ERICH L. LEHMANN AND HENRY SCHEFFÉ, University of California at Los Angeles.

A family $\mathfrak{M}$ of measures M on a space X of points x is defined to be *complete* if $\int_X f(x)\, dM = 0$ for every M in $\mathfrak{M}$ implies $f(x) = 0$ except on a set A for which $M(A) = 0$ for every M in $\mathfrak{M}$. For a given family of measures the question of completeness may be regarded as the question of unicity of a related functional transform. Classical unicity results are applicable to many families of probability distributions that have been studied by statisticians. The notion of completeness throws light on the problem of similar regions and the problem of unbiased estimation. The concept of a *maximal* sufficient statistic—roughly, a sufficient statistic that is a function of all other sufficient statistics—is developed. A constructive method of finding such is given, which seems to apply to all examples ordinarily considered in statistical theory. A relation between completeness and maximality is found.

9. On a Proposed Method for Estimating Populations. CECIL C. CRAIG, University of Michigan, Ann Arbor.

It was proposed to the author by a biologist that a method be devised for estimating the total population in an area which shall utilize the minimum distances between randomly chosen individuals and their neighbors in directions lying in each of the four quadrants. Assuming that the area is a square and that the distribution law over it is rectangular, it turns out that the complete distribution of the lengths of sides of minimum squares which contain a second individual is simpler than that of minimum distances. In both cases a simple estimate is found which uses most but not all of the information in a sample and whose efficiency is comparable to that based on a complete enumeration of a sample area, though such an enumeration is not always possible.

10. Some Results on the Asymptotic Distribution of Maximum- and Quasi-Maximum-likelihood Estimates. HERMAN RUBIN, Institute for Advanced Study.

The author investigates the asymptotic normality of maximum- and quasi-maximum-likelihood estimates of parameters of systems of linear stochastic difference equations. The principal tool is the extension of the Central Limit Theorem to dependent variables previously obtained by the author (presented to the American Mathematical Society in April, 1948). The results obtained are analogous to those in the case in which no differences are present. Some extensions are also made to systems of stochastic difference equations linear in the coefficients but not necessarily in the variables. If the complete system of stochastic difference equations is linear in the jointly dependent variables, asymptotic efficiency is demonstrated for maximum-likelihood estimates.

11. The Probability Points of the Distribution of the Median in Random Samples from Any Continuous Population. CHURCHILL EISENHART, LOLA S. DEMING, and CELIA S. MARTIN, National Bureau of Standards, Washington.

The abscissa of the (one-tail) ϵ-probability point of the distribution of the median in random samples of size $n = 2m + 1$ $(m \geq 0)$ from any continuous population is identical with the abscissa of the corresponding $P_{\epsilon,n}$-probability point of the parent distribution, where $P_{\epsilon,n}$ is determined by

$$\sum_{k=\frac{1}{2}(n+1)}^{n} C_k^n P_{\epsilon,n}^k (1 - P_{\epsilon,n})^{n-k} = \epsilon, \qquad (0 \leq \epsilon \leq 1). \tag{1}$$

From (1) it follows that

$$P_{1-\epsilon,n} = 1 - P_{\epsilon,n} \tag{2}$$

and that

$$P_{\epsilon,n} = x_\epsilon(n+1, n+1) = \frac{1}{1 + F_\epsilon(n+1, n+1)} = \frac{1}{1 + e^{2Z_\epsilon(n+1,n+1)}}, \tag{3}$$

where $x_\epsilon(\nu_1, \nu_2)$, $F(\nu_1, \nu_2)$, and $Z_\epsilon(\nu_1, \nu_2)$ denote the ϵ-probability points of the incomplete-beta-function distribution, Snedecor's F-distribution and Fisher's z-distribution, for $\nu_1(= 2q)$ and $\nu_2(= 2p)$ 'degrees of freedom', respectively. The foregoing results are certainly not "new": Harry S. Pollard implicitly utilized the first equality on the extreme left of (3) in his doctoral dissertation at the University of Wisconsin in 1933 (see *Annals of Math. Stat.*, Vol. 5 (1934), p. 250), and John H. Curtiss has given the generalization of (1) appropriate to the case of the 'rth. position' in random samples from any continuous population (see *Amer. Math. Monthly*, Vol. 50 (1943), p. 103) and utilized (3) explicitly to obtain the 5% point of the distribution of the median in random samples of size $n = 23$. The aim of the present paper is to give these results somewhat greater publicity—they are hardly "well known". To this end a table (Table 1) is given of the values of $P_{\epsilon,n}$ to 5 significant figures for $\epsilon = 0.001, 0.005, 0.01, 0.025, 0.05, 0.10, 0.20, 0.25$ and $n = 3(2)15(10)95$, together with expressions from which $P_{\epsilon,n}$ can be evaluated accurately and conveniently for values of n (and ϵ) not included in the table. Numerical examples illustrate the use of the table and formulas. Concise derivations of the fundamental relations and formulas are given in an appendix.

12. On the Arithmetic Mean and the Median in Small Samples from the Normal and Certain Non-Normal Populations.

CHURCHILL EISENHART, LOLA S. DEMING, and CELIA S. MARTIN, National Bureau of Standards, Washington.

Let $\bar{x}_{\epsilon,n}$ and $\tilde{x}_{\epsilon,n}$ denote the abscissae of the one-tail ϵ-probability points of the arithmetic mean and the median, more specifically, the abscissae *exceeded* with probability ϵ by the mean and the median, respectively, in random samples of size n ($= 2m + 1$) from any specified population, and let $\sigma_{\bar{x}_n}$ and $\sigma_{\tilde{x}_n}$ denote the standard deviations of the mean and the median in such samples, respectively. The following symmetrical populations with zero location parameters and unit scale parameters are considered in this paper:

Type		
normal (Gaussian)	$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$,	$-\infty \leq x \leq \infty$
double-exponential (Laplace)	$\frac{1}{2}e^{-\lvert x \rvert}$,	$-\infty \leq x \leq \infty$
rectangular (uniform)	1,	$-\frac{1}{2} \leq x \leq \frac{1}{2}$
Cauchy	$\frac{1}{\pi}\frac{1}{1+x^2}$,	$-\infty \leq x \leq \infty$
sech	$\frac{1}{\pi}$ sech x,	$-\infty \leq x \leq \infty$
sech² (derivative of "logistic")	$\frac{1}{2}$ sech² x,	$-\infty \leq x \leq \infty$

Using the basic table, relating probability points of the distribution of the median to probability points of the parent distribution, given in Churchill Eisenhart, Lola S. Deming and Celia S. Martin, "The probability points of the distribution of the median in random samples from any continuous population," values of $\tilde{x}_{\epsilon,n}$ for random samples from each of the above distributions have been evaluated, and are tabulated to 5 decimal places in the present paper, for $n = 3(2)15(10)95$ and $\epsilon = 0.001, 0.005, 0.01, 0.025, 0.05, 0.10, 0.20, 0.25$.

In the case of the *normal distribution*, values of $\bar{x}_{\epsilon,n}$ to 5 decimal places are given also for the aforementioned combinations of ϵ and n. Comparison of the values of $\tilde{x}_{\epsilon,n}$ and $\bar{x}_{\epsilon,n}$ gives precise numerical meaning to the well-known lesser accuracy of the median as an estimator of the center of a normal population, for samples of *any* odd size ($n = 2m + 1$). Values of the ratio $R_{\epsilon,n} = \tilde{x}_{\epsilon,n}/\bar{x}_{\epsilon,n}$ are given also for this case (normal population), to 4 decimal places for the above combinations of ϵ and n, together with the best available values of $\sigma_{\tilde{x}_n}/\sigma_{\bar{x}_n}$ for $n = 3(2)15(10)55$. When $0 < \epsilon \leq 0.025$, the ratio $R_{\epsilon,n}$ exceeds the ratio $\sigma_{\tilde{x}_n}/\sigma_{\bar{x}_n}$, showing that the 'tails' of the exact distribution of the median are 'longer' than the tails of the normal distribution with the same mean and standard deviation; and, when $0.05 \leq \epsilon \leq 0.25$, the ratio $R_{\epsilon,n}$ is less than $\sigma_{\tilde{x}_n}/\sigma_{\bar{x}_n}$. (A theoretical argument shows that the point of equality is close to the 0.042-probability point.) A method for computing $\sigma_{\tilde{x}_n}$, based on the foregoing, is given that is believed to be accurate to $.001/\sqrt{n}$, or better for $n \geq 3$.

In the case of the *double-exponential* distribution, values of $\bar{x}_{\epsilon,n}$ are given to 4 decimal places for $n = 3(2)11$, and $\epsilon = 0.005, 0.01, 0.025, 0.05, 0.10, 0.25$, for comparison with the corresponding values of $\tilde{x}_{\epsilon,n}$. It is found that *when* $n = 3$, $\bar{x}_{\epsilon,3} < \tilde{x}_{\epsilon,3}$ for $\epsilon = 0.005, 0.001$, and 0.025, indicating that in random samples of 3 from a double-exponential distribution the arithmetic mean furnishes narrower confidence limits for the center of the distribution at 0.95, 0.98, and 0.99 levels of confidence. When $n = 5$, the mean is 'better' at the .98 and .99 levels of confidence; and, when $n = 7$, at the 0.99 level. For all other combinations of ϵ and n (≥ 3), the median is 'better.'

In the case of the *rectangular* distribution, values of $\bar{x}_{\epsilon,n}$ are tabulated to 4 decimals for $n = 3(2)9$, and values of $\tilde{x}_{\epsilon,n}$, the ϵ-probability point of the mid-range in samples of n, for $n = 3(2)15(10)95$, in each instance for $\epsilon = 0.005, 0.01, 0.025, 0.05, 0.10, 0.25$, and in the case of $\bar{x}_{\epsilon,n}$ for $\epsilon = 0.001$ also. The superiority of the midrange over the mean and the median, well-known but here exhibited numerically for the first time, is truly amazing.

It is planned to provide values of $\bar{x}_{\epsilon,n}$ for samples from the sech and sech^2 distributions in the final paper.

13. The Relative Frequencies with which Certain Estimators of the Standard Deviation of a Normal Population Tend to Underestimate its Value.

CHURCHILL EISENHART and CELIA S. MARTIN, National Bureau of Standards, Washington.

Let $x_1, x_2, \cdots, x_n$ denote a random sample of n independent observations from a normal population with mean μ and standard deviation σ. Common estimators of σ are

$$s_1 = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2/n}, \qquad s_2 = s_1\sqrt{n/(n-1)}, \qquad s_3 = s_1/c_2,$$

$$m_1 = \sqrt{\frac{\pi}{2}}\sum_{i=1}^{n} |x_i - \bar{x}|/n, \qquad m_2 = m_1\sqrt{n/(n-1)},$$

and $R_1 = (x_L - x_S)/d_2$, where $\bar{x} = \sum_{i=1}^{n} x_i/n$, x_L is the largest and x_S the smallest of the x's, $c_2 = \mathsf{E}(s_1)$, and $d_2 = \mathsf{E}(x_L - x_S)$, the symbol $\mathsf{E}(\ )$ denoting "mathematical expectation (or mean value) of." A table is given that shows to 3 decimals the relative frequencies (probabilities) with which these estimators tend to underestimate σ when $n = 2(1)10$, 12, 15, 20, 24, 30, 40, 60. The results show among other things that, for very small samples ($n \leq 10$) such as chemists and physicists commonly use, Bessel's formula for the probable error, which is based on s_2, has a marked downward bias in the probability sense (in addition to its known slight downward bias in the mean value sense), whereas Peter's formula, which is based on m_2, has only a slight downward bias in the probability sense and no bias in the mean value sense. A table of divisors is given by means of which "median estimators" of σ can be computed readily from the basic quantities $\sum_{i=1}^{n} (x_i - \bar{x})$, $\sum_{i=1}^{n} |x_i - \bar{x}|$, and $(x_L - x_S)$, that is, estimators that will over- and underestimate σ equally often in repeated use. An application to control charts is noted. Median estimators, like maximum likelihood estimators ("modal estimators") have the useful property that if $T_{\frac{1}{2}}$ is a median estimator of θ, then $f(T_{\frac{1}{2}})$ is a median estimator of $f(\theta)$, a property unfortunately not possessed by the customary "unbiased" ("mean") estimators.

14. Some Non-Parametric Tests of Whether the Largest Observations of a Set are too Large. (Preliminary Report.)

JOHN E. WALSH, Douglas Aircraft Company, Santa Monica, California.

Let $x(1), \cdots, x(n)$ represent the values of n observations arranged in increasing order of magnitude. By hypothesis these observations have the properties: (1) They are independent and from continuous symmetrical populations (2) For large n the variances of the tail order statistics are either very large or very small compared with the variances of the central order statistics (3) For large n the tail order statistics are approximately independent of the central order statistics (4) Each observation is from a population whose median is either θ or φ, where $x(n - r + 1), \cdots, x(n)$ are from populations with median θ while the central and smaller order statistics are from populations with median φ. The test is: *Accept* $\varphi < \theta$ *if* $\min [x(n - i_k) + x(j_k); 1 \leq k \leq s \leq r] > 2x(t_\alpha)$, where $i_u < i_{u+1}$, $j_v < j_{v+1}$, $i_s = r - 1$, and t_α is defined by $Pr\ [x(t_\alpha) < \varphi \mid \theta = \varphi] = \alpha$. Here

$$\alpha = \Pr \{\min [x(n - i_k) + x(j_k); 1 \leq k \leq s \leq r] > 2\varphi \mid \theta = \varphi\}.$$

For large n the significance level of the test is approximately α while the significance level does not exceed $2\,\alpha$ for any value of n. Suitable values of α can be obtained for $r \geq 4$. As $\theta - \varphi \to -\infty$ the power function tends to zero, while the power function tends to unity as $\theta - \varphi \to \infty$. For $\theta - \varphi < 0$ the power function is monotonically increasing.

15. On the Bounded Significance Level Properties of the Equal-tail Sign Test for the Mean. John E. Walsh, Douglas Aircraft Company, Santa Monica, California, (Presented by Title).

The equal-tail sign test for deciding whether the population mean μ is equal to a given hypothetical value μ_0 is defined by: *Accept* $\mu \neq \mu_0$ *if either* $x_i < \mu_0$ *or* $x_{n+1-i} > \mu_0$, $\left(i > \frac{n+1}{2}\right)$. Here x_j, $(j = 1, \cdots, n)$, is the j^{th} largest of n independent observations drawn from n populations which satisfy the conditions: (i) The mean of each population has the value μ. (ii) Each population is continuous at its mean. (iii) The mean is at a 50% point for each population. This paper investigates how the significance level of the equal-tail sign test varies when (i)–(iii) are not satisfied. It is found that the significance level does not differ noticeably from its hypothetical value under conditions much more general than (i)–(iii). This significance level stability, combined with the properties of being easily applied and reasonably efficient for small samples from a normal population, suggests that the equal-tail sign test be considered for application whenever the population mean is to be tested on the basis of a small number of observations.

16. Infinitely Divisible Distributions. William Feller, Cornell University, Ithaca, New York.

A simple derivation of P. Lévy's formula is given starting from the following definition: a distribution function $F(x)$ is infinitely divisible if for every n it is possible to find finitely many distributions $F_{k,n}(x)$ such that $F(x) = F_{1,n}(x) * \cdots * F_{k_n,n}(x)$ and that $F_{k_1 n}(x)$ tends to the unitary distribution uniformly in n. This definition is more general than the one used by P. Lévy and Khintchine. The equivalence of the two definitions was proved by Khintchine by deep methods. The new approach renders the equivalence obvious. Furthermore, a new characterization of infinitely divisible distributions is given; it is equivalent to Gnedenko's characterization but requires no special analytical tools.

17. Fluctuation Theory of Recurrent Events. William Feller, Cornell University, Ithaca, New York.

Consider a sequence of independent or dependent trials but suppose that each has a discrete sample space. The paper studies recurrent patterns $\mathfrak{E}$ which can be roughly characterized by the property that after every occurrence of $\mathfrak{E}$ the process starts from scratch, the conditional probabilities coinciding with the original absolute probabilities. Typical examples are success runs, returns to equilibrium, zeros of sums of independent variables, passages through a state in a Markov chain. New methods are developed unifying and simplifying previous theories and applying to larger classes of recurrent events. It is shown in an elementary way the probability that $\mathfrak{E}$ occurs at the n-th trial either has a limit or is asymptotically periodic. This theorem has many consequences. For example, the ergodic properties of discrete Markov chains follow in a few lines, and the difference between finite and infinite chains disappears. Several theorems of the renewal type are proved. Weak and strong limit theorems for the number N_n of occurrences of $\mathfrak{E}$ in n trials are derived shedding new light on stable distributions.

18. Formulas for the Percentage Points of the Distributions of the Arithmetic Mean in Random Samples from Certain Symmetrical Universes. UTTAM CHAND, University of North Carolina and National Bureau of Standards.

Using the method of Fisher and Cornish, the $100\epsilon\%$ point of the distribution of the arithmetic mean in random samples of size N from any universe having finite cumulants of the first four orders, κ_1, κ_2, κ_3, κ_4, is expressed to order $1/N^2$ as a function of N, the $100\epsilon\%$ point of a standardized normal deviate and the quantities κ_1, κ_2, $\kappa_3/\kappa_2^{3/2}$, κ_4/κ_2^2. The numerical coefficients are evaluated for the cases of sampling from rectangular, double-exponential, sech and sech² distributions. The application of the resulting formulas is illustrated numerically for ϵ = .001, .005, .010, .025, .050, .100, and .250. In the case of the rectangular and double-exponential distributions, the results obtained for $N = 10$ are compared with accurate values, indicating the accuracy of the formulas.