ABSTRACTS OF PAPERS

Presented on August 21, 1946, at the Cornell meeting of the Institute

1. A Test of Randomness in Two Dimensions. Howard Levene, Columbia University.

A square of side N is divided into N^2 unit cells, and each cell takes on the characteristics A or B with probabilities p and $q = 1 - p$ respectively, independently of the other cells. A cell is an "upper left corner" if it is A and the cell above and cell to the left are not A. Let V_1 be the total number of upper left corners and let V_2, V_3, V_4 be the number of similarly defined upper right, lower right, and lower left corners respectively. Let $V = (V_1 + V_2 + V_3 + V_4)/4$. It is proved that V is normally distributed in the limit with $E(V) = p(Nq + p)^2$ and $\sigma^2(V) \sim N^2pq^2(4 - 20p + 45p^2 - 27p^3)/4$. The conditional limit distribution of V when p is estimated from the data, and the limit distribution of a related quadratic form are also obtained. These statistics are in a sense a generalization of the run statistics used for testing randomness in one dimension.

2. Asymptotic Distribution of Moments from a System of Linear Stochastic Difference Equations. Herman Rubin, Cowles Commission for Research in Economics.

Let $\sum_{\tau=0}^{\infty} B_\tau y'_{t-\tau} + \Gamma z'_t = u'_t$, $(t = 1, 2, \cdots)$, be a complete system of linear stochastic difference equations determining y_{ti} (the coordinates of y_t), $t > 0$, in terms of y_{ti}, $t \leq 0$, and z_{tk} (the coordinates of z_t), which are assumed to be fixed variates, and the random variables u_{ti} (the coordinates of u_t). Such a system is called a stable if for every bounded set of fixed variates, and $E(u'_t u_t)$ uniformly bounded, $E(y'_t y_t)$ is uniformly bounded. This condition is shown to be equivalent to $\sum | h_{ij\tau} |$ finite, where $y'_t = \sum_{\tau=0}^{\infty} H_\tau(u'_{t-\tau} - \Gamma z'_{t-\tau}) + \sum_{\nu=0}^{\infty} J_{t,\nu} y'_{-\nu}$ is the solution of the above difference equation. Let Q_t be an infinite quadratic form in $y_{t-\tau,i}$ and $z_{t-\nu,k}$ $(\tau, \nu = 0, 1, \cdots)$ with coefficients depending only on i, k, τ, and ν. Such a quadratic form is called convergent if the sum of the absolute values of the coefficients is finite. It is shown under fairly general conditions that the mean of a convergent quadratic form is asymptotically normally distributed with variance $0\left(\frac{1}{T}\right)$.

3. Conditional Expectation and Unbiased Sequential Estimation. David Blackwell, Howard University.

It is shown that $E[f(x_\alpha)E_\alpha y] = E(fy)$ whenever $E(fy)$ is finite, and that $\sigma^2(E_\alpha y) \leq \sigma^2(y)$, with equality holding only if $E_\alpha y = y$, where $E_\alpha y$ denotes the conditional expectation of y with respect to the family of chance variables x_α. These results imply that whenever there is a sufficient statistic u and an unbiased estimate t, not a function of u only, for a parameter p, the function $E_u t$, which is a function of u only, is an unbiased estimate for p with variance smaller than that of t. A sequential unbiased estimate for a parameter is obtained, such that when the sequential test terminates after i observations, the estimate is a function of a sufficient statistic for the parameter with respect to these observations. A special case of this estimate is that obtained by Girshick, Mosteller, and Savage (*Annals of Math. Stat.*, Vol. XVII (1946), pp. 13–23) for the parameter of a binomial distribtion.

4. A Discussion of the Ehrenfest Model. Preliminary report. Mark Kac, Cornell University.

A particle moves along a straight line in steps Δ, the duration of each step being τ. The probabilities that the particle at $k\Delta$ will move to the right or left are $(1/2)(1 - k/R)$

and $(1/2)(1 + k/R)$ respectively. R and k are integers and $| k | \leq R$. M. C. Wang and G. E. Uhlenbeck in their paper *On the theory of Brownian motion II* (*Rev. Mod. Phys.* Vol. 17 (1945), pp. 323–342) discuss this random walk problem and state several unsolved problems. In answer to some of the questions raised the following results are obtained: Let $(1 - z)^{R-j} \cdot (1 + z)^{R+j} = \Sigma C_k^{(j)} z^k$ (j an integer) then, the probability $P(n, m \mid s)$ that a particle starting from $n\Delta$ will come to $m\Delta$ after time $t = s\tau$ is equal to $2^{-2R}(-1)^{R+n}\Sigma(j/R)^s C_{R+j}^{(-n)} C_{R+m}^{(j)}$, where the summation is extended over all j such that $| j | \leq R$. Also, if R is even the probability $P'(n, 0 \mid s)$ that the particle starting from $n\Delta$ will come to 0 at $t = s\tau$ for the first time is calculated. For $n = 0$ this gives a solution of the so-called recurrence time problem first studied on simpler models by Smoluchowski. Through a limiting process in which $\tau \to 0$, $\Delta \to 0$, $\Delta^2/2\tau \to D$, $1/R\tau \to \beta$, $n\Delta \to x_0$, $m\Delta \to x$, $s\tau = t$, one is led to fundamental distributions concerning the velocity of a free Brownian particle. In particular, $P(n, m \mid s)$ approaches the well-known Ornstein-Uhlenbeck distribution.

5. Sampling from Contaminated Distributions. Preliminary report. John W. Tukey, Princeton University.

A contaminated distribution is a nearly normal distribution in which extreme observations are more frequent than in a normal distribution. By studying the bias and variability of several measures of dispersion when applied to samples from particular one-parameter families of contaminated distributions it is shown that: (i) for nearly normal distributions, the mean deviation is often better than the standard deviation; (ii) small changes in the underlying distribution may increase the sampling variance of the standard deviation by a factor of three. This suggests that, in a broad class of cases, the mean deviation is safer than the standard deviation when a single dispersion is estimated from a set of data. This conclusion need not apply in an analysis of variance situation.

6. On the Class of Functions Defined by the Difference Equation $(x + 1)f(x + 1) = (a + bx)f(x)$. Leo Katz, Wayne University.

The difference equation defines only three discrete functions: the binomial, the Poisson and the Pascal functions; the first and third have one parameter (N) slightly generalized. It is shown that the Pascal function with this generalization is identical with the Polya-Eggenburgher distribution, which is a very useful form of the Compound Poisson Law and has been used to explain probability situations involving contagion. Areas for all functions in the class are given in terms of existing tables of the incomplete γ and β-functions. Observed distributions are fitted by two moments. As Carver (*Handbook of Mathematical Statistics*) pointed out, the advantages of fitting by difference equations are many; not the least is the fact that it is unnecessary to discriminate among the various functions in fitting an observed distribution. The problem of discrimination, posed by Frisch (*Metron*, Vol. 10) and others, may be resolved in terms of the sampling distribution of variances for the Poisson function, since the three functions correspond to situations where the variance is less than, equal to, or greater than the mean, respectively.

7. Retention of Decimal Places in Matrix Calculations. Franklin E. Satterthwaite, Aetna Life Insurance Company. (Read by title)

The accumulation of errors in matrix calculations has been studied by the author and others for special types of matrices and for special methods of calculation. In the present paper, error formulae are developed for the standard Doolittle and Waugh-Dwyer Compact routines. These formulae do not place any restrictions on the matrices involved and do not require any extra calculations or initial approximations. Simple rules are developed which give for each step in the calculations the number of decimal places which must be retained. These rules are efficient in the sense that the retention of fewer places will, except for good fortune in balancing of errors, lead to results less accurate than those specified. The rules also assist in choosing that arrangement of the calculations which will lead to the smallest average number of significant figures which must be retained for the calculation as a whole.

8. The Efficiency of the Mean Moving Range. PAUL G. HOEL, University of California at Los Angeles. (Read by title)

The statistic $w = \sum_1^{n-1} | x_{i+1} - x_i | \sqrt{\pi}/2(n - 1)$ is studied as an estimate of σ for a normal variable subject to trend effects. It is shown that the efficiency of w compares favorably with that of the mean square successive difference, δ^2. The proof that w, and also δ^2, is asymptotically normally distributed is made to depend upon a general result that can be derived from a theorem of S. Bernstein on dependent variables.

9. Some Basic Theorems for Developing Tests of Fit for the Case of the Non-Parametic Probability Distribution Function. BRADFORD F. KIMBALL, New York State Department of Public Service. (Read by title)

Given a universe with C.D.F. $P[X \leq x] = F(x)$. Consider a random sample of n values x_i which have been ordered so that $x_i \leq x_{i+1}$. The successive differences of the true c.d.f. values at $X = x_i$ are denoted by u_i. Thus

$$u_1 = f(x_1)$$

$$u_i = F(x_i) - F(x_{i-1}), 2 \leq i \leq n$$

$$u_{n+1} = 1 - F(x_n).$$

THEOREM 1. *The product power moments*

$$E(u_r^p u_s^q u_t^w \cdots)$$

for any or all different indices from 1 *to* $n + 1$, *where the powers are real numbers greater than minus one, are given by*

$$E(u_r^p u_s^q u_t^w \cdots) = \frac{\Gamma(n + 1)\,\Gamma(p + 1)\,\Gamma(q + 1)\,\Gamma(w + 1)\cdots}{\Gamma(n + 1 + p + q + w + \cdots)}$$

COROLLARY. *If a range* $R(k, m)$ *is defined by*

$$R(0, m) = F(x_m),\ R(n + 1, m) = 1 - F(x_{n+1-m})$$

$$R(k, m) = F(X_{k+m}) - F(x_k)$$

where k *and* m *are positive integers such that* $m \leq n$ *and* $k + m \leq n$, *its probability distribution is independent of* k, *and hence equal to that of* $F(x_m)$.

THEOREM 2. *Given a test funtion of* u_i

$$Y = \sum_m u_i^p$$

where p *is a real positive number, and the sum is for* m *indices chosen at random on the range* 1 *to* $n + 1$. *Let* $\bar{Y}$ *and* σ^2 *denote the mean and variance of this test function. Establish a convention for increasing the indices included in the above sum for increasing* m *as* n *increases, such that* $[m/(n + 1)] =$ *constant, to nearest multiple of* $1/(n + 1)$. *Then the asymptotic distribution of* $(Y - \bar{Y})/\sigma$ *for inreasing* n, *subject to the above condition, is the normal distribution with zero mean and unit variance, except in the trivial case* $m = n + 1$, $p = 1$.

10. Confidence Limits for the Fraction of a Normal Population which Lies between Two Given Limits. JACOB WOLFOWITZ, Columbia University. (Read by title)

Let $x_1, \cdots, x_N$ be N independent observations from a normal population with mean μ and variance σ^2, both unknown. Let $N\bar{x} = \Sigma x_i$ and $(N-1)s^2 = \Sigma(x_i - \bar{x})^2$ define $\bar{x}$ and s^2. Let L_1 and L_2 be given constants with $L_1 < L_2$, and let

$$\gamma = (\sqrt{2\pi}\sigma)^{-1} \int_{L_1}^{L_2} \exp -\tfrac{1}{2}\left\{\frac{y-\mu}{\sigma}\right\}^2 dy$$

By a lower confidence limit on γ with confidence coefficient α is meant a function $D(x_1, \cdots, x_N)$ such that the probability is α that $D \leq \gamma$. Since $\bar{x}$ and s^2 are sufficient estimates of μ and σ^2 the restriction that D be a function of $\bar{x}$ and s only is imposed. It is assumed that there exist a) a positive d such that $L_1 + d < \mu < L_2 - d$; b) a positive C such that $\sigma < C$. From these it follows that there exists a lower bound $G = G(d, C)$ on γ. Let $\chi^2_{1-\alpha}$ be that number for which $P\{\chi^2 < \chi^2_{1-\alpha}\} = 1 - \alpha$, where χ^2 has $N - 1$ degrees of freedom, and let $w = \dfrac{\sqrt{N-1}s}{\chi_{1-\alpha}}$. It is shown that if D be defined as follows:

1) if $L_1 \leq \bar{x} \leq L_2$,

$$D = (2\pi)^{-\frac{1}{2}} \int_{L_1-\bar{x}/w}^{L_2-\bar{x}/w} \exp\{-\tfrac{1}{2}y^2\}\, dy$$

2) $D = G$ otherwise, then $| P\{D \leq \gamma\} - \alpha |$ approaches zero as $N \to \infty$. Thus D is a large sample lower confidence limit. The extension to upper and two-sided limits presents no difficulty.

11. The Consolidated Doolittle Technique. PAUL BOSCHAN, Econometric Institute. (Read by title)

The quadratic matrix notation is interpreted as a segment in a sequence of matrices wherein each successor matrix is augmented by a bordering row and column. Extension theorems based on this idea date back into the last century. The step from the original concept to one of higher order is also fruitful in discussing inverse matrices, specifically the inverse of a symmetric matrix. The symmetry of the matrix of normal equations for a set of multiple regression coefficients is restored by adding the transpose of the column on the right side of the equations, i.e. the co-variances with the dependent variable and the variance of the dependent variable itself. The inverse of this matrix can be constructed as partial sum over a series of matrices. Each individual element of this series is in itself meaningful. The solution for the set of multiple regression coefficients relating the k-th variable to the preceding $(k-1)$ variables is a column matrix. The product of this matrix with its transpose expressed in terms of the residual variance forms the k-th term in the matrix series. The summation of the first n products yields the inverse matrix. This characteristic of the inverse can be used to great advantage in the standardization of elementary computational steps.

12. Estimation of Structural Equations through Linear Transformation of Regression Coefficients. THEODORE W. ANDERSON and HERMAN RUBIN, Cowles Commission for Research in Economics.

A method is presented for estimating the coefficients of a single structural equation in a system $By'_t + \Gamma z'_t = u'_t$ $(t = 1, 2, \cdots, T)$, where B and Γ are matrices of coefficients, y_t is a row vector of G observed jointly dependent variables, z_t of K observed predetermined variables and u_t of G random elements. Given the distribution of the random elements, the equations define the distribution of the y_t. Some coordinates of z_t may be coordinates of y_{t-1}, etc. It is assumed that the structural equation to be estimated has at least $G - 1$ coefficients prescribed zero. The part of the population regression matrix corresponding to the predetermined variables with zero coefficients has rank one less than the number of jointly dependent variables with non-zero coefficients. The maximum likelihood estimate of this matrix is a linear transformation of the unrestricted sample regression matrix. The estimated vector of coefficients of y_t is the vector annihilated by this matrix. The vector of coefficients of z_t is estimated by means of this vector and the regression matrix. These estimates are consistent and asymptotically normally distributed. For z_t fixed, small sample confidence regions are given for the coefficients.