ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Annual Meeting of the Institute, Seattle, Washington, June 14–17, 1961. Additional abstracts appeared in the June, 1961 issue.)

18. Tables for the Reliability of Repairable Systems with Time Constraints

(Preliminary report). ROQUEZ BEJARANO AND RONALD S. DICK, International Electric Corp., Paramus, N. J.

Tables have been prepared to solve for the reliability of systems composed of A similar subsystems of which at most N can be inoperable for periods exceeding t_0 time units. A second time constraint is introduced into the model so that for at least time t_1 following a return of the system from state $N + 1$ to N machines inoperative, the system is only in states 0 to N or the system fails.

The mixed difference-differential equations solved are of the forms:

$$P_j^1(t) = -[\mu_j + \lambda_j]P_j(t) + \lambda_{j-1}P_{j-1}(t) + \mu_{j+1}P_{j+1}(t) + \mu_{N+1}P_{(N)(j)}(t_1)[P_{N+1}(t - t_1)]$$

$$P_j^1(t) = -[\mu_j + \lambda_j]P_j(t) + \lambda_{j-1}P_{j=1}(t) + \mu_{j+1}P_{j+1}(t) - \lambda_N[P_N(t - t_0)]P_{(N+1)(j)}(t_0).$$

or

$$P_j^{*1}(t) = -\mu_{N+1}P_{(N)(j)}(t_1)[P_{N+1(t-t_1)}] - (\mu_j + \lambda_j)P_j^*(t) + \mu_{j+1}P_{j+1}^*(t) + \lambda_{j-1}P_{j-1}^*(t)$$

where appropriate boundary conditions are applied. Reliability is defined as $R(t) = \sum_{j=0}^{A} P_j(t) + \sum_{j=0}^{N} P_j^*(t)$. For $A = 1$ (1) 5, and $N = 0(1)A - 1$, the tables give for 81 combinations of λ and μ the approximate time at which $R(t) = .001, .005, .01, .05, .10$ as well as the MTBF. The Cornish-Fisher equation and Weibull approximations are used in finding the reliability points. The MTBF is found by evaluating the Laplace Transforms of the mixed-differential difference equations and is exact. Reference should be made to "The Reliability of Repairable Complex Systems, Part A: The Similar Machine Case" by R. S. Dick, *5th Mil-E-Con National Convention on Military Electronics*, 1961 for a complete set of equations solved in this paper and the details of the model.

19. Mutual Information and Maximal Correlation as Measures of Dependence.

C. B. BELL, San Diego State College.

Kramer (1961) asks if Shannon's mutual information, C_P, is equivalent to Kramer's generalization (to arbitrary σ-algebras) of Gebelein's (1939) Maximal Korrelation, S_P, which satisfies Rényi's (1959) postulates for a dependence measure of pairs of random variables. It is found that for two normalizations C_P' and C_P'' of C_P: (1) $0 \leqq S_P, C_P', C_P'' \leqq 1$; (2) $S_P = 0$ iff $C_P = 0$ iff $C_P = 0$ iff the algebras are independent. For strictly positive probability spaces, (3) the algebras are set independent iff there exists a probability function P_1 such that $S_{P_1} = C_{P_1} = C_{P_1} = 0$; (4) $C_P' = 1$ iff one algebra contains the other; (5) $C_P'' = 1$ iff the algebras are equal; (6) $S_P = 1$ if the algebras have a non-trivial intersection; (in the finite case, the converse of (6) holds;) (7) there exists a probability space such that no two of the dependence measures are equivalent. *Open Problems*: Which of (3)–(6) are valid for (a) the Gelfand-Yaglom (1957) mutual information for non-atomic algebras generated by random variables; and (b) the Lloyd mutual information for arbitrary algebras?

20. On a Necessary and Sufficient Condition for a Set of Jointly Normal Variables to have a Common Variance and a Common Covariance (Preliminary report). B. R. Bhat, University of California, Berkeley.

The following *theorem* is proved. Let $x_i (i = 1, 2, \cdots, n)$ have a joint n-variate normal distribution with mean 0. Then the necessary and sufficient condition that $n\bar{x}^2 = (\sum x_i)^2/n$ and $\sum (x_i - \bar{x})^2$ are distributed independently and the latter as $c\chi^2$, where c is a constant, is that Var $x_i = \sigma^2$ and Cov $(x_i, x_{i'}) = v (i, i' = 1, 2, \cdots, n)$. The sufficiency of this theorem is well known. The necessity follows from the facts that if $\mathbf{X}$ is $\mathbf{N}(\mathbf{0}, \mathbf{\Sigma})$ (i) $\mathbf{X'AX}$ and $\mathbf{X'BX}$ are distributed independently if and only if $\mathbf{A \Sigma B} = \mathbf{0}$ and (ii) $\mathbf{X'AX}$ has a $c\chi^2$ distribution if and only if $c\mathbf{A} = \mathbf{A \Sigma A}$ (cf., C. R. Rao, *Advanced Statistical Methods in Biometric Research*, p. 56). It is also proved that, if $x_i, y_j (i, j = 1, 2, \cdots, n)$ have a joint $2n$-variate normal distribution, then $Q = \sum (x_i - \bar{x})^2 + \sum (y_j - \bar{y})^2$ is distributed as $c\chi^2$ independently of $\bar{x}^2$ and $\bar{y}^2$ if and only if Var $x_i = \sigma_1^2$, Var $y_j = \sigma_2^2$, Cov $(x_i, x_{i'}) = v_1$, Cov $(y_j, y_{j'}) = v_2$, Cov $(x_i, y_j) = v_3$ $(i, i', j, j' = 1, 2, \cdots, n)$. In particular, $(\bar{x} - \bar{y})[n(n-1)/Q]^{\frac{1}{2}}$ has a t-distribution with $2n - 2$ d.f., if further $v_3 = \frac{1}{2}(v_1 + v_2)$.

21. A Property of Least Squares Estimator in Regression Analysis when the Independent Variables are Stochastic. P. K. Bhattacharya, University of North Carolina. (Introduced by S. N. Roy.)

$(X_1, \cdots, X_p, Y)$ follows a $(p + 1)$ variate distribution which is assumed to satisfy the following conditions: (i) for every non-null $(a_0, a_1, \cdots, a_p)$, the set

$$\{(x_1, \cdots, x_p, y): a_0 + a_1x_1 + \cdots + a_px_p = 0\}$$

has probability zero, (ii) $E(X_jX_{j'})$ is finite, $j, j' = 0, 1, \cdots, p$, $X_0 \equiv 1$, (iii) $E[Y \mid X_1, \cdots, X_p]$ is a linear function of $X_1, \cdots, X_p$, (iv) $V[Y \mid X_1, \cdots, X_p]$ is a finite constant. $n \geqq p + 1$ independent observations are made on $(X_1, \cdots, X_p, Y)$ and the loss in estimating the true regression function $\phi(x_1, \cdots, x_p) = E[Y \mid x_1, \cdots, x_p]$ by another function $\psi(x_1, \cdots, x_p)$ is $W(\phi, \psi) = \int [\phi - \psi]^2 \, dF$ where $F(x_1, \cdots, x_p)$ is the marginal distribution function of $X_1, \cdots, X_p$. Let $\mathcal{C}$ be the class of all estimators which are linear in Y's and have bounded risk. Then the estimator obtained by the method of least squares belongs to $\mathcal{C}$ and has uniformly minimum risk in $\mathcal{C}$ if and only if all the elements of the inverse of the matrix of normal equations, have finite expectations. This last condition is not satisfied in general, and in particular, for $p = 1$ and for a normal distribution of X_1, it is satisfied if and only if $n \geqq 4$.

22. Selecting the "Best" t out of k Populations. P. K. Bhattacharya, University of North Carolina. (By title) (Introduced by S. N. Roy.)

$F(x, \theta)$ is a family of continuous distribution functions admitting density functions $f(x, \theta)$ and $g(\theta)$ is a real valued function satisfying the following conditions: (i) for $\theta_2 > \theta_1$, $f(x, \theta_2)/f(x, \theta_1)$ is a monotonically increasing function of x, (ii) $g(\theta)$ is a monotonically increasing function of θ. Suppose $X_1, \cdots, X_k$ have distribution functions

$$F_1(x) = F(x, \theta), \cdots, F_k(x) = F(x, \theta_k)$$

respectively, each of which belongs to the above family and one observation is made on each of $X_1, \cdots, X_k$, the observations being independent. Let $C(\theta_1, \cdots, \theta_k)$ be the sum of the largest t of the quantities $g(\theta), \cdots, g(\theta_k)$. A vector $(d_1, \cdots, d_k)$, $d_j = 0$ or 1, $\sum_1^k d_j = t$, represents the decision for selecting the random variable X_j if and only if $d_j = 1$, and the loss in taking the decision $(d_1, \cdots, d_k)$ when $(\theta_1, \cdots, \theta_k)$ obtains, is

$C(\theta_1, \cdots, \theta_k) - \sum_1^k d_j\theta_j$. It has been shown that the decision function δ^* defined below is admissible and minimax, $-\delta^*(x) = (\delta_1^*(x), \cdots, \delta_k^*(x))$, where $\delta_j^*(x) = 1$ if x_j is one of the largest t of $x_1, \cdots, x_k = 0$ otherwise.

23. Approximations for the Entropy of Functions of Markov Chains. John J. Birch, University of Nebraska. (By title)

If $\{Y_n\}$ is a stationary ergodic Markov process taking on values in a finite set $\{1, 2, \cdots, A\}$, then its entropy can be calculated directly. If ϕ is a function defined on $1, 2, \cdots, A$, with values $1, 2, \cdots, D$, no comparable formula is available for the entropy of the process $\{X_n = \phi(Y_n)\}$. However, the entropy of this functional process can be approximated by the monotonic functions

$$\bar{G}_n = h(X_n \mid X_{n-1}, \cdots, X_1) \quad \text{and} \quad \underline{G}_n = h(X_n \mid X_{n-1}, \cdots, X_1, Y_0),$$

the conditional entropies. Furthermore, if the underlying Markov process $\{Y_n\}$ has strictly positive transition probabilities, these two approximations converge exponentially to the entropy H, where the convergence is given by $0 \leqq \bar{G}_n - H \leqq B\rho^{n-1}$ and $0 \leqq H - \underline{G}_n \leqq B\rho^{n-1}$ with $0 < \rho < 1$, ρ being independent of the function ϕ.

24. Some Properties of a Large Set of Random Signals (Preliminary report). Nelson M. Blachman, Sylvania Electronic Defense Labs. (Introduced by Emanuel Parzen.)

For a communication channel that accepts n-tuples of real numbers of mean-square value P as input signals and delivers them with each component perturbed by the addition of independent, zero-mean normal noise of variance N, M different signals can be distinguished with an error probability approaching zero as $n \to \infty$ provided $\tan^2 \theta > N/P$, where $\sin \theta = M^{-1/n}$. To achieve this result, it suffices to choose for the signals the rectangular coordinates of independent random points $s_1, \cdots, s_M$ on the surface of an n-sphere of radius $(nP)^{\frac{1}{2}}$ centered at the origin O. When a perturbed n-tuple r is received, the most likely signal is that corresponding to the nearest s_i. Thus, the space of all r is divided into M convex, pyramidal regions $R_1, \cdots, R_M$, with R_i consisting of all points closer to s_i than to any other s_j. We find, e.g., that nearly every s_j within an angular distance 2θ of s_i contributes a face to R_i. In a random direction from s_i, with probability approaching 1, R_i extends out very nearly just to the circular cone of generating angle θ with axis Os_i. This cone very closely circumscribes R_i and nearly every one of its faces, edges, etc. Nearly all of R_i's surface is accounted for by faces approximately arc sin $(2^{-\frac{1}{2}} \sin \theta)$ from s_i. The nearest edge of R_i of dimensionality $n - k$ is approximately ϕ_k from s_i, with $\sin^k \phi_k \cos \phi_k = k^{\frac{1}{2}k}(k+1)^{-\frac{1}{2}(k+1)} \sin^k \theta$. From such results, we obtain lower bounds on the noise variance that could result in a large error probability if a portion of the noise should be dependent on the signal being transmitted and on $s_1, \cdots, s_M$.

25. On a Problem in Hilbert Space with Applications. J. R. Blum and D. L. Hanson, Sandia Corporation.

Let $\{X_n, n = 0, \pm 1 \cdots\}$ be a stationary stochastic process. Then it is known that a necessary and sufficient condition that the process be pure nondeterministic is that the spectral distribution of the process be absolutely continuous and that the logarithm of the spectral density be integrable. In this paper we obtain necessary and sufficient conditions directly on the covariance sequence. Several related problems are discussed.

26. Length of the Longest Run of Consecutive Successes. E. J. Burr, University of New England, Armidale, N.S.W., Australia. (Introduced by D. B. DeLury.)

In an ordered sequence of n observations, let those possessing a specified attribute be called "successes", and those not possessing it "failures". A conspicuous feature of such a sequence is the length k of the longest run of consecutive successes observed. On the hypothesis that the successes and failures occur in random order, all permutations being equally probable, we derive formulae for the probability that the statistic k should exceed any specified value (i) when the probability of success in each trial is given, (ii) when the numbers of successes and failures are given, (iii) when the sequence is circular with no preferred initial point. The joint distribution of the lengths of the longest success run and longest failure run is also derived. The treatment is greatly simplified by introducing the concept of a success run of length zero.

27. Comparing Distances between Multivariate Normal Populations, I (Preliminary report). Theophilos Cacoullos, Columbia University. (By title)

Let π_i be p-variate normal populations with means $\mu^{(i)}i = 0, 1, \cdots, k$, respectively, and with the same known covariance matrix Σ. The $\mu^{(i)}$, $i = 1, \cdots, k$, are known and $\mu^{(0)}$ is unknown. Let $\Delta_{ij}^2 = (\mu^{(i)} - \mu^{(j)})'\Sigma^{-1}(\mu^{(i)} - \mu^{(j)})$ denote the generalised (Mahalanobis) distance between π_i and π_j. On the basis of a sample $x_1, \cdots, x_n$ from π_0 a population π_i, $i = 1, \cdots, k$, is to be selected so that $\Delta_{0i}^2 = \min_{1 \leq j \leq k} \Delta_{0j}^2$. Let d_i be the decision of selecting π_i. (1) Assume that the $\mu^{(i)}$, $i = 1, \cdots, k$, are collinear. Then by invariance under linear transformations on the p-space the problem reduces to locating the mean of a normal variable with unit variance into one of k consecutive intervals covering the real line. Hence the theory of monotone procedures for the exponential class of distributions (Girshick and Blackwell, *Theory of Games and Statistical Decisions*, pp. 179–193, John Wiley and Sons, New York, 1954) applies. Let $\bar{x}$ be the sample mean and $\delta_{ij}(\bar{x}) = (2\bar{x} - \mu^{(i)} - \mu^{(j)})'\Sigma^{-1}(\mu^{(j)} - \mu^{(i)})$, $i, j = 1, \cdots, k$. Then, e.g., for $k = 2$ the family of decision rules: take d_1 if $\delta_{12}(\bar{x}) < c$, take d_2 otherwise, $-\infty < c < +\infty$, is minimal complete for a wide class of loss functions. (2) Suppose that the k points $\mu^{(1)}, \cdots, \mu^{(k)}$ are vertices of a $(k - 1)$-simplex in p-space ($p \geqq k - 1$). Define $\delta(\bar{x}) = (\delta_{12}(\bar{x}), \cdots, \delta_{1k}(\bar{x}))'$ and similarly $\delta(\mu^{(0)})$. Then $\delta(\bar{x})$ has a $(k - 1)$-variate normal distribution with mean $\delta(\mu^{(0)})$ and known covariance matrix Δ, say. If $\chi_{k-1}^2(\alpha)$ denotes the 100α percentage point of a χ^2-distribution with $k - 1$ degrees of freedom, then of all level α tests of the hypothesis $\Delta_{01}^2 = \Delta_{02}^2 = \cdots = \Delta_{0k}^2$ with power depending only on $n\delta'(\mu^{(0)})\Delta^{-1}\delta(\mu^{(0)})$ the test with critical region $n\delta'(\bar{x})\Delta^{-1}\delta(\bar{x}) > \chi_{k-1}^2(\alpha)$ is uniformly most powerful. If $\delta_{ij}(\mu^{(0)}) \leqq -\lambda\Delta_{ij}^2$, for all $j \neq i, 0 < \lambda \leqq 1$, is the region where d_i is the correct decision, $i = 1, \cdots, k$, then a unique minimax solution is found for constant loss functions.

28. Subsamples and Order Statistics (Preliminary report). J. T. Chu and Kamal Ya'Coub, University of Pennsylvania.

Suppose that a random sample of size mn is drawn from a given distribution and the sample is divided into m subsamples each of size n. The observations in each subsample may be arranged in order of magnitude. In this way, we obtain m order statistics each of size n. For subsequent analysis, a subset of observations may be selected, as representatives, from each of the m order statistics. Furthermore, by combining the jth order statistic of each subsample, one obtains a random sample of size m from the population of the jth order statistics in samples of size n drawn from the parent distribution. Various types of statistical inference based on such divisions, orderings, and selections of a random sample are being investigated. A number of devices have been found where savings in time and computation compare favorably against loss of accuracy. Methods which improve the efficiencies of existing ones are also found.

29. Percentile Estimators for the Parameters of the Exponential Failure Law.

SATYA D. DUBEY, Procter and Gamble Co. (By title)

For the 2-parameter exponential failure law, the percentile estimators of the location and the scale parameters, based on at most two percentiles, have been derived under three different possible cases. The sampling and the asymptotic distributions and the expressions of the kth moments of these percentile estimators have been obtained. The choices for the cumulative probabilities have been made in such a manner that the corresponding percentiles insure asymptotic minimum variance unbiased percentile estimators of the location and the scale parameters. In case both the location and the scale parameters are unknown, the concept of the generalized variance, which is defined as the determinant of the variance-covariance matrix, has been used to determine two cumulative probabilities ensuring minimum generalized variance. The smallest sample observation and the 80th percentile seem to provide asymptotically most efficient percentile estimators for both the parameters of the exponential distribution.

30. On Separating a Deterministic Component from a Stochastic Sequence.

FRIEDHELM EICKER, University of North Carolina.

In the separation of a deterministic component of the form of a linear regression $Y\underline{\delta}$ from a stochastic sequence $y_1, y_2, \cdots$ the attention has been focussed almost exclusively on the estimation of $\underline{\delta}$. It can easily be seen, however, that often some methods applied for this estimation cannot be used at the same time for an estimation of the stationary sequence $\{x_t\}$ in the assumed model $\underline{y} = Y\underline{\delta} + \underline{x}$. So, for instance, the least squares estimators $\hat{\underline{\delta}}$ of $\underline{\delta}$, though consistent, may asymptotically not even allow a stationary sequence at all.

In order to make an estimation of $\underline{\delta}$ possible $\{x_t\}$ must be submitted to some assumptions such as (to stay quite general) (a) weak stationarity, with a possibly non-zero mean function $E(x_t) = \mu_t$, (b) finite second moments $E(x_t^2) <$ const $< \infty$ only. Through (b) in a sense the most general class, say S, applicable in the model is described. A sufficient condition for consistency (in the sense of mean square convergence) of $\hat{\underline{\delta}}$, given a certain stationary sequence with covariances $R(m)$, is $\sum_{|m|\leqq N} |R(m)|/\lambda_{\min}(Y_N'Y_N) \to 0$. For consistency over the whole class (a) (or even over S), $N^{-1}\lambda_{min}(Y_N'Y_N) \to \infty$ is sufficient. However, one easily finds regression matrices Y_N of this kind, yet the "covariance function" $\hat{R}_t(m) = E(\hat{x}_{t+m}\hat{x}_t)$ of the residuals does not tend to the true one (not to mention an estimation of the entire sequence $\{x_t\}$ at all).

31. On Pairwise Independence.

SEYMOUR GEISSER AND NATHAN MANTEL, National Institutes of Health.

It is well known in statistical theory that pairwise independence is necessary but not sufficient for a set of p variables to be mutually independent. The example that is usually cited in the statistical literature is due to S. Bernstein and involves discrete variables. An example of continuous variables that exhibit this peculiarity is produced and is simply the joint distribution of correlation coefficients from a multivariate normal distribution with a diagonal variance-covariance matrix.

32. On Tests with Likelihood Ratio Criteria in Some Problems of Multivariate Analysis (Preliminary report). N. C. Giri, Stanford University. (Introduced by Charles Stein.)

Let X be a p-dimensional column vector having multivariate normal distribution with unknown mean ξ and unknown non-singular covariance matrix Σ. In this paper we have considered two different testing problems concerning mean ξ and Σ viz.,

(i) to test the hypothesis that ξ lies in $\mathcal{Z}$ against the alternative that ξ lies in $\mathcal{Y}$ where $\mathcal{Z}$ and $\mathcal{Y}$ are subspaces of the parametric space of dimensions $p - p'$ and $p - q$ respectively $(p > p' > q)$;

(ii) to test the hypothesis that $\Sigma^{-1}\cdot\xi$ lies in $\mathcal{Z}'$ against the alternative that $\Sigma^{-1}\cdot\xi$ lies in $\mathcal{Y}'$ where $\mathcal{Z}'$ and $\mathcal{Y}'$ are subspaces of the adjoint space $\mathcal{X}'$ of the space of x's, of dimensions q and p' $(p > p' > q)$ respectively.

It has been shown that the likelihood ratio test for problem (ii) is uniformly most powerful invariant similar; whereas if the sample size N and p' and q are large then the likelihood ratio test for problem (i) is nearly uniformly most powerful invariant.

33. Circular Probability Problems. William C. Guenther, The Martin Co. and University of Wyoming.

When a circle C_1 of radius R is dropped upon a fixed circle C_2 of radius D, several interesting and useful problems arise. Two of these involve (a) the probability that C_1 covers a randomly selected point within C_2, and (b) the probability that C_1 covers a randomly selected point on the circumference of C_2. The first problem has been considered by others and results may be found in Rand RM 330. The second problem is considered and the relationship between the two problems is observed. The three dimensional counterparts are also considered. Tables are included.

34. An Application of the Sequential Probability Ratio Test to Finite Populations. Paul Gunther, Armour Research Foundation.

Let $x_1, \cdots, x_n$ be the values assumed by a finite population consisting of n members. It is desired to derive a sequential procedure to predict whether $y_n = \sum_{i=1}^n x_i$ is $\geqq C$ or $< C$, where C is specified. It is assumed that the finite population is in turn a random sample from a normally distributed superpopulation $f(x_i; \theta)$ with unknown mean θ and known standard deviation σ. (This can be considered also in the sense of an *a priori* distribution.) Further, θ is assumed to be equal either to θ_1 or θ_0 $(\theta_1 > \theta_0)$, each with (*a priori*) probability $\frac{1}{2}$. Define α = Prob (predicting $y_n \geqq C/y_n < C$), as determined from the *a priori* distributions, and similarly for β. The SPRT leads to withholding a prediction and taking further observations if $B < \{f(y_i/y_n \geqq C)/[f(y_i/y_n < C)]\} < A$ where $y_i = \sum_{j=1}^i x_j$; A and B are determined in the usual manner; and $f(\cdot)$ is weighted by θ_0 and θ_1. If $C = n\frac{1}{2}(\theta_1 + \theta_0)$, the acceptance numbers A_i are determined from the equation $\exp(-2T_iD_i) = ((1 + A)^{-1} - N(-T_i - D_i))/(N(-T_i + D_i) - (1 + A)^{-1})$; where

$$T_i = (A_i - i\bar{\theta})/((n - i)^{\frac{1}{2}}\sigma), \qquad D_i = (n - i)^{\frac{1}{2}}\Delta/\sigma, \qquad \bar{\theta} = (\theta_1 + \theta_0)/2,$$

$$\Delta = (\theta_1 - \theta_0)/2, \qquad N(\cdot) = \text{cumulative normal distribution.}$$

This equation is easily solved graphically. As $n \to \infty$, the test approaches the usual Wald SPRT. A numerical application is made to a problem in "budget control." A derivation is also possible without resorting to the a *priori* probabilities of θ_0 and θ_1.

35. On Step-Down Procedure in Simultaneous Multivariate Analysis of Variance (Preliminary report). P. R. KRISHNAIAH, Remington Rand UNIVAC. (By title)

Let X denote an $n \times p$ matrix whose rows form n independent vectors having p variate normal distributions with a common covariance matrix Σ and means given by $E(X) = M\theta$, $M:n \times m$, $\theta = m \times p$, rank $M = r \leqq m \leqq n$, where M has known elements and is called the "design matrix," and θ is unknown. Now, consider the K orthogonal hypotheses $H_i:C_i\theta = 0, i = 1, 2, \cdots, K$, against the alternatives $A_i:C_i\theta = \eta_i$. If the variates can be arranged in some order according to their importance, the hypotheses $H_1, \cdots, H_K$ can be tested simultaneously by using the "Step-Down Procedure". This procedure was first used by Roy and Bargman (these *Annals*, 1958) for testing the hypothesis of multiple independence of sets of variates when the parent population is multivariate normal. In the present paper, the tests of significance are derived for testing the hypotheses $H_1, \cdots, H_K$ simultaneously by using the "Step-Down Procedure". The confidence bounds on meaningful parametric functions are also derived. The extension of these results to random models is under investigation.

36. Linear Hypothesis with Linear Restrictions. ANDRÉ G. LAURENT, Wayne State University.

Let Y be weakly spherical (or make it so by changing the definition of the inner product) with $E(Y) = \mu = A\theta$, where A is $n \times k$ of rank $r \leqq k$, with column space Ω, restricted by $L\theta = O$, $L' = (R', P')$, with $R\theta$ nonestimable, where R is $s \times k$ of rank $s \leqq k - r$ and $P\theta$ estimable where P is $t \times k$ of rank $t \leqq r$ and restricts μ to ω. Completing $R\theta = O$ to nonestimable $R^*\theta = O$, where R^* is $(k - r) \times k$ of rank $k - r$, makes θ (hence any $M\theta$) "pseudo estimable" i.e., with structure (C, μ), $C \in \Omega$ (or ω) under Ω (or ω), with best estimate $\hat{\theta} = (C, Y)$. Let $L^{*\prime} = (R^{*\prime}, P')$.

$$\begin{pmatrix} \hat{\theta}_\omega \\ O \end{pmatrix} = \begin{pmatrix} A'A & L^{*\prime} \\ L^* & O \end{pmatrix}^{-1} \begin{pmatrix} A'Y \\ O \end{pmatrix} \qquad \hat{\theta}_\omega = B_1^1 A'Y \qquad \widehat{M\theta_\omega} = M\hat{\theta}_\omega$$

with $A'AB_1^1 + L^{*\prime}B_2^1 = I$ and $B_1^1L^{*\prime} = O$. Several proofs of the above equations are given as well as geometrical interpretations. If A is of full rank, $R = O$, obtaining $\hat{\theta}_\omega$ is straightforward. These results extend and generalize those of P. Dwyer and others.

37. The Effect of Convergence to Normality on Tests of Hypotheses. LLOYD J. MONTZINGO AND NORMAN C. SEVERO, University of Buffalo. (By title)

Let X be a random variable with mean μ_x and standard deviation σ_x, and let the distribution of X tend to normality as some function of the parameters $\eta(\mu_x, \sigma_x) \to \eta_0$. The result of applying normal theory tests of hypotheses on the mean or variance to a sample from the distribution of X is considered. Denote by P_x the power of a test based on a sample of size n (n fixed) from the distribution of X, and by P the power of the test if X were normally distributed. Then sufficient conditions are given for which $P_x \to P$ as $\eta(\mu_x, \sigma_x) \to \eta_0$.

38. On the Asymptotic Normality and Independence of the Sample Partial Autocorrelations for an Autoregressive Process. V. K. MURTHY, Stanford University. (By title)

For a stationary autoregressive model of order s, the partial autocorrelation coefficients of order j, $j = 0, 1, 2, \cdots, s - 1$ are defined; the partial autocorrelation coefficient of order zero being the same as the autocorrelation coefficient of order one. Denoting these s parameters by $\rho_1, \pi_1, \cdots, \pi_{s-1}$, it is shown in this paper that their sample images namely $r_1, p_1, \cdots, p_{s-1}$ are asymptotically independently normally distributed with means equal to the corresponding population values and asymptotic variances given by

$$\text{Var}(r_1) = n^{-1}(1 - \rho_1^2)(1 - \pi_1^2) \cdots (1 - \pi_{s-1}^2),$$

$$\text{Var}(p_j) = n^{-1}(1 - \pi_j^2)(1 - \pi_{j-1}^2) \cdots (1 - \pi_{s-1}^2), \quad j = 1, 2, \cdots, s-1,$$

where n is the size of the sample from the autoregressive process of order s. The partial correlogram of the model and application of the result are discussed.

39. On Fitting a Linear Trend and Testing Independence when the Residuals Form a Markov Process. V. K. Murthy, Stanford University. (By title)

In this note we are studying the problem of fitting a linear trend when the residuals are serially correlated according to a first order Markov scheme. An iteration method for solving the maximum likelihood equations is proposed and an explicit criterion for the convergence of the iteration process is obtained. It is incidentally shown that for large samples the serial dependence may be neglected and ordinary least squares analysis used for estimating the trend. A general result in this direction was proved by Herman Wold and Juréen Lars, [*Demand Analysis*, John Wiley and Sons, New York (1953)]. The asymptotic variance-covariance matrix of the maximum likelihood estimates and the likelihood ratio criterion for testing $\rho = 0$ are obtained. Extending a result of Ogawara [*Ann. Math. Stat*, Vol 22 (1951), pp. 115–118], the problem of regression when the residuals are serially correlated according to the first order Markov scheme, is reduced to the classical case; in the case of fitting a linear trend exact tests for the regression parameters and the hypothesis $\rho = 0$ are derived. Illustrating the iteration method an application to the data on the average yield per acre of potatoes from 1903 to 1932 of the United States is worked out.

40. On the Cumulants of a General Renewal Process. V. K. Murthy, Stanford University.

In this paper the results of Smith on the cumulants of a Renewal Process are extended to the case of a General Renewal Process. After establishing the asymptotic representation theorems for the Φ-moments and Ψ-cumulants of a General Renewal Process, the table of the first eight cumulants of a Renewal Process has been extended to the case of a General Renewal Process. A theorem is proved leading to a check on the calculations. As a particular case of the General Renewal Process, the cumulants of the "Equilibrium Process" are obtained.

41. Some Distribution-Free Multiple Comparison Procedures in the Asymptotic Case. Peter Nemenyi, S.U.N.Y. College of Medicine at Brooklyn.

By means of a generalization of Stuart's transformation for correlated variables [*J. Roy. Stat. Soc.*, Ser. B, Vol. 20 (1958), 373–378] and by an alternative method, it is shown that existing tables for multiple comparisons of normal means also apply to a large family of asymptotic permutation procedures, including Steel's rank and sign tests, some median tests, and multiple comparisons based on Kruskal-Wallis rank totals. The tests, designed for translation alternatives, can also be adapted to the problem of differences in spread (but in this case it is more difficult to obtain confidence intervals).

The tabulation of $(1 - \alpha)^{1/k}$ and $\frac{1}{2}[(1 - \alpha)^{1/k} + 1]$ points of various one- and two-sample statistics (e.g., for setting simultaneous sign-test confidence intervals on k median treatment effects) is also advocated, and some tables are provided.

42. Formulation of a Model Containing a Chance Mechanism according to which Observations are Missed: The Randomized Block Design. Junjiro Ogawa, Nihon University, Tokyo, Japan and Bernard S. Pasternack, New York University Medical Center. (By title).

In this paper an attempt has been made to introduce a chance mechanism according to which observations are missed into the model and subsequent analysis of the randomized block design. By partitioning the "design matrix" for the randomized block design into $(\Phi:\Psi)$, it is possible to incorporate the process of randomization for this design directly into the theoretical model. The exact distribution of the sum of squares due to treatments contingent upon the incidence matrix for treatments, Φ, being random, and the incidence matrix for blocks, Ψ, being fixed, can then be rigorously obtained.

When there is an *a priori* probability that observations may be missing in a randomized block design, this may be accounted for in the model by introducing a set of mutually independent chance variables

$$m_u = \begin{cases} 0 \text{ with probability } p_u \text{ if } x_u \text{ is missing} \\ 1 \text{ with probability } 1 - p_u = q_u \text{ if } x_u \text{ is not missing.} \end{cases}$$

The vector $m' = (m_1, m_2, \cdots, m_n)$ is called the *missing observation vector*. The probability distribution of m' is given by $\prod_{u=1}^{n} p_u^{1-m_u} q_u^{m_u}$. On the basis of this model, a modified (F) test statistic is obtained. The extension of this approach to other more complex designs is formal. Whether or not it is possible to obtain the exact or approximate distribution of this statistic is, at the moment, an open question.

43. On Some Methods of Estimation for the Logarithmic Series Distribution. G. P. Patil, University of Michigan.

Applications of logarithmic series distribution have been discussed among others by Fisher (1943), Williams (1943, 1944), Harrison (1945) and Kendall (1948). Problem of estimation, however, does not seem to have been thoroughly investigated. This paper provides different estimates for the parameter of the logarithmic series distribution and investigates their efficiency and the amount of bias in certain special cases.

44. Some Asymptotic Properties of the Negative Binomial Distribution. Vivian Pessin, Children's Hospital, Buffalo, N. Y. (By title) (Introduced by Norman C. Severo.)

The following two theorems have been proved:

Theorem 1. The negative binomial frequency function is asymptotically normal as $\lambda/\alpha \to \infty$. Let

$$P(\nu = x) = \left(\frac{\alpha}{1+\alpha}\right)^{\lambda} \left(\frac{1}{1+\alpha}\right)^{x} \binom{\lambda + x - 1}{x}, \qquad x = 0, 1, 2 \cdots; \alpha > 0; \lambda > 0.$$

Let $m_0 = [(\lambda - 1)/\alpha] - \mu((\lambda - 1)(1 + \alpha))^{\frac{1}{2}}/\sigma\alpha$, $\sigma_0 = ((\lambda - 1)(1 + \alpha))^{\frac{1}{2}}/\sigma\alpha$, where σ is an arbitrary constant > 0, and μ is an arbitrary constant. Let $y = (x - m_0)/\sigma_0$. Then, for fixed x, and for α bounded away from 0 and from ∞,

$$\lim_{\lambda/\alpha \to \infty} Q(\eta = y) = \sigma(2\pi)^{-\frac{1}{2}} \exp\left(-(y - \mu)^2/2\sigma^2\right), \qquad -\infty < y < \infty.$$

Theorem 2. The negative binomial frequency function is asymptotically the Gamma frequency function as $\alpha \to 0$, for fixed λ such that $0 < \lambda < 1$. Using the same notation as in the first theorem, let $m_1 = \lambda - 1$, $\sigma_1 = k((1 + \alpha)/\alpha)$, where k is an arbitrary constant > 0. Then $\lim_{\alpha\to 0} Q(\eta = y) = (k^\lambda e^{-ky}y^{\lambda-1})/(\Gamma(\lambda - 1)) = g(y)$, which becomes the gamma frequency function when $g(y)$ is defined to be 0, for $y \leqq 0$.

45. **On Horvitz and Thompson's T-Class Estimators** (Preliminary report). S. G. Prabhu-Ajgaonkar and B. D. Tikkiwal, Karnatak University, Dharwar, India. (By title)

Horvitz and Thompson (*J.A.S.A.*, 1952), while discussing sampling with varying probability and without replacement, have given three classes of estimators. If an empty class is that where the unbiased estimators independent of population values do not exist, it is shown that their T_1-class is in general an empty class when sampling with varying probability is adopted. However, when Midzuno's system of sampling is adopted with replacement, such a class of estimators exists and has a minimum variance unbiased estimator in the class independent of the population values. The T_3-class, which is non-empty, has no minimum variance unbiased estimator independent of the population values even when simple random sampling is adopted. The non-empty T_2-class is known to have only one unbiased estimator and so is a minimum variance unbased estimator. It is noted that for Midzuno's system of sampling with replacement T_2-class estimator has a smaller variance than that of the minimum-variance unbiased estimator of T_1-class. It is further noted for sampling with varying probability and with replacement that the minimum variance ununbiased estimator in the over all class consisting of the classes T_1 and T_2 is the unbiased estimator in the T_2 class. However, the relative efficiency of T_2-class estimator and an estimator in T_3-class depends upon the probability system adopted.

46. **The Role of the Multivariate Edgeworth Series in the Random Walk Problem.** J. F. Price and W. M. Stone, Boeing Airplane Co., Seattle, Washington, and J. D. Wheelock, Oregon State University.

For the random walk in N-space denote the vth step by the random vector $\rho_v = (\cos \varphi_{1v}, \cos \varphi_{2v}, \cdots, \cos \varphi_{Nv})$ where $\cos \varphi_{kv}$ $(k = 1, 2, \cdots, N)$ is a direction cosine. The point x_n attained after n steps from the origin is then the resultant $x_n = \rho_1 + \rho_2 + \cdots + \rho_n$. The vectors ρ_v are independent (and identically distributed) so that $x = \lim_{n\to\infty} x_n$ follows the N-dimensional normal distribution. The paper formally expresses the probability density function of x_n as a multivariate Edgeworth series and deduces therefrom, for $N = 2$, the asymptotic "modified Pearson" series discussed by Greenwood and Durand (*Ann. Math. Stat.*, 26: 233–246 and 28: 978–986) for the distance r_n from the origin. This approach has the advantage of using known moments (or cumulants) with maximum efficiency, requiring only those moments necessary in the determination of the polynomial coefficients of a given power of $1/n$, in contrast to the Laguerre series approach previously employed. A table of probabilities $P(R < r_n)$ for $n = 4, 5, \cdots, 24$ was constructed by quadrature from the known exact distribution for $N = 3$, and is presented for comparison with results from the modified Pearson series.

47. **On a Mathematical Model for Poliomyelitis Vaccine Effectiveness** (Preliminary report). Dana Quade, Communicable Disease Center, Atlanta, Georgia.

A linear relationship between the logarithm of the attack rate of paralytic poliomyelitis and the number of doses of killed-virus vaccine received, which "indicates that each successive dose reduced the remainder of the susceptibles by the same proportion as did the first dose", was discovered by Dr. Jonas Salk (*The Lancet*, October 1, 1960, pp. 715–723). This model is made explicit and various mathematical and statistical problems which it entails are considered.

48. The Distribution of the Ratio of the Variances of Variate Differences in the Circular Case. J. N. K. RAO AND G. TINTNER, Iowa State University.

In time series analysis, the variate difference method is used to test the order of the finite difference at which the trend or the systematic part in the time series is approximately eliminated. There is no exact test available in the literature except for the one proposed by Tintner ("The variate difference method," Bloomington, Indiana, 1940) based on a method of selection which uses only a portion of the observations. In this paper, the statistic V_{k+1}/V_k is proposed to test that the trend is approximately eliminated at the kth finite differencing of the series where V_k is the variance of the series of the kth differences. Its exact distribution assuming that the observations are $NI(0, \sigma^2)$ is derived under a circular definition of the universe. The lower 5% and 1% points of the statistics V_2/V_1 and V_3/V_2 are tabulated for various values of N, the size of the sample. In practice, one uses the non-circular statistic with these percentage points for the circular statistic as an approximation, especially with long time series.

49a. Estimation of Failure Rates of Systems in Development. DAVID RUBINSTEIN, General Electric Company.

Given an $m \times \infty$ matrix (C_{ij}) of populations of components with the corresponding matrices: (λ_{ij}) of failure rates, (T_{ij}) of test times, (X_{ij}) of the number of failures, (a_{ij}) of acceptance numbers. X_{ij} are independent Poisson random variables with parameter $\lambda_{ij} T_{ij} \cdot a_{ij}$ are nonnegative integers or ∞. Components from population C_i^* with failure rate λ_i^* will be used in the system if $C_i^* = C_{ij}$ where $X_{ij} \leqq a_{ij}$ and $j < j'$ for any j' for which $X'_{ij} \leqq a'_{ij}$.

Let $\delta_{ij} = 0$ if $X_{ij} > a_{ij}$, 1 if $X_{ij} \leqq a_{ij}$. Let $\hat{\lambda}_{ij} = 0$ if $X_{ij} > a_{ij} + 1$, X_{ij}/T_{ij} if $X_{ij} \leqq n_{ij} + 1$. Under rather general conditions, $\sum_{i=1}^m \sum_{j=1}^p \hat{\lambda}_{ij} \prod_{k-1}^{j-1} (1 - \delta_{ik})$ is an unbiased estimate of the system failure rate of $\sum_{i=1}^m \lambda_i^*$ in the sense that difference of the two random variables has the expected value zero. Let $\hat{\sigma}_{ij}^2 = 0$ if $X_{ij} > a_{ij} + 2$, $X_{ij}(X_{ij} - 1)/T_{ij}^2$ if $X_{ij} = a_{ij} + 2$, X_{ij}/T_{ij}^2 if $X_{ij} \leqq n_{ij} + 1$. $\sum_{i=1}^m \sum_{j=1}^p \hat{\sigma}_{ij}^2 \prod_{k=1}^{j-1} (1 - \delta_{ik})$ is an unbiased estimate of

$$E\left[\sum_{i=1}^{m} \sum_{j=1}^{\infty} \hat{\lambda}_{ij} \prod_{k=1}^{j=1} (1 - \delta_{ik}) - \sum_{i=1}^{m} \lambda_i^*\right]^2.$$

49b. Determining Bound on Expected Values of Certain Functions. BERNARD HARRIS, University of Nebraska.

This extends some results given by the author in "Determining Bounds on Integrals with Applications to Cataloging Problems" (*Ann. Math. Stat.* Vol. 30, 1959). Let $g(x)$ be a continuous function, not linearly dependent on the first k monomials, whose first k derivatives exist and are monotonic; $\mu_1, \mu_2, \cdots, \mu_k$ are known constants and the first k moments of an unknown distribution function $F(x)$. The sup(inf) $E\{g(x)\}$ is computed, the sup(inf) being taken over all distribution functions, whose first k moments are given by $\mu_1, \mu_2, \cdots, \mu_k$. The extremal distributions are characterized, and computed explicitly for $k \leqq 3$. In addition, some applications are given.

50. Convergence to Normality of Functions of a Normal Random Variable. NORMAN C. SEVERO AND LLOYD J. MONTZINGO, University of Buffalo. (By title)

The asymptotic distributions of functions of a normal random variable are investigated as some function of the parameters tends to a limit. It is assumed that the functions of the normal variate will be defined in such a way as to be real for all values of the variate. In particular, if Y is a normal random variable with mean μ_y and standard deviation σ_y, and if $X = Y^p$, where $p > 0$, has mean μ_x and standard deviation σ_x, then X is asymptotically normally distributed with mean μ_x and standard deviation σ_x as $\eta \equiv \mu_x/\sigma_x \to \infty$. If $p < 0$, X is asymptotically normally distributed with mean $\mu_y^p[1 + O(\eta^{-2})]$ and standard deviation $(p\mu_y^p/\eta)[1 + O(\eta^{-2})]^{\frac{1}{2}}$. Sufficient conditions on a function h are given for the transformed variate $X = h(Y)$ to be asymptotically normal with mean $h(\mu_y)$ and standard deviation $h'(\mu_y)\sigma_y$ as $\eta \to \infty$.

It is shown that, for a large class of transformations h, the variate $X = h(Y)$ is asymptotically normal as $\sigma_y \to 0$ with mean μ_x and standard deviation σ_x providing they exist. If they do not exist, the asymptotic mean and standard deviation are $h(\mu_y) + O(\sigma_y^2)$ and $h'(\mu_y)\sigma_y[1 + O(\sigma_y^2)]^{\frac{1}{2}}$. The condition $h'(\mu_y) \neq 0$ is shown to be necessary for convergence to normality. Furthermore, the asymptotic distribution of $h(Y)$ is characterized when the first m derivatives of h, at μ_y, are 0.

51. **A Probability Model for Couple Fertility.** S. N. Singh, University of California, Berkeley.

A probability distribution for the number of conceptions to a couple (a male and a female leading a married life), during a given time interval T, is derived on the assumptions: (a) the probability of a virtual conception in a unit of time is p, independently of virtual conceptions in any other units of time, where T is assumed to contain T units of time. The probability of conception in the first unit of time is p. (b) if there is a conception in a certain unit of time, then there is no conception during next h units of time. h is constant. (c) a couple belongs to one of two mutually exclusive groups A and B during time T. Group A consists of sterile couples and couples who choose to be so, group B consists of couples not belonging to group A. The group B is homogeneous in the sense that any couple of group B has the same p, the probability of conception in a unit of time. Estimates of parameters are based on sample mean and zero cell frequency. The asymptotic variances of the estimates are derived. The distribution has been applied to two examples given by Dandekar (*Sankhya,* Vol 15 (1955), pp. 237–250).

52. **A Method for Computing the Cumulative Distribution Function of the Product of Two Dependent *t*-Variables.** Rosedith Sitgreaves, Teachers College, Columbia University.

We suppose we have three variables, y_1, y_2, and a, distributed independently of each other; the first two are normal with zero means and unit variances, and the third has a chi-square distribution with n degrees of freedom. The marginal distribution of each of the variables $t_1 = y_1(n/a)^{\frac{1}{2}}$ and $t_2 = y_2(n/a)^{\frac{1}{2}}$, is thus Student's t-distribution with n degrees of freedom, but the two t-variables are not independent. In some problems we are interested in computing the cumulative distribution function of the product t_1t_2. An integral representation is found for the probability that this product is less than a specified value. This integral can be evaluated relatively easily by numerical integration to any desired accuracy.

53. **A Monte Carlo Analysis of the Serial Correlation Coefficient.** John S. White, General Motors Research Labs.

Let (x_t) be a discrete stochastic process satisfying the auto regressive equation $x_t = \alpha x_{t-1} + u_t$ where the u's are NID (0, 1). The limiting distribution of $\hat{\alpha}$, the MLE for α, is known (J. S. White, *Ann. Math. Stat.*, Vol. 30 (1959), 831–834) except when $\alpha = \pm 1$. In this paper the results of a Monte Carlo analysis of the distribution of $\hat{\alpha}$ is given. Samples of size $n = 10, 20, 50, 100$, and 500 were drawn from populations having $\alpha = -2., -1.1, -1.0, -.9, -.5, 0, .5, .9, 1.0, 1.1$ and 2.

54. Distribution Function for Randomized Factorial Experiments. S. ZACKS, The Technion, Israel Institute of Technology AND S. EHRENFELD, New York University.

In a previous paper on Randomization and Factorial Experiments [S. Ehrenfeld and S. Zacks, These *Annals*, Vol. 32 (1961), pp. 270–297] two randomization procedures, for choosing fractional replications, were studied. These procedures have been designed to yield information on a subgroup of preassigned parameters. Schemes of the analyses of variance, associated with each of the proposed randomization procedures, were also given. The objective of the present paper is to study the distribution functions of the associated test statistics, and to establish procedures for the determination of test criterions for given levels of significance, as well as the power of the tests.

The distribution functions of the test statistics, for testing the significance of the chosen parameters, depend on the nuisance parameters (those which do not belong to the preassigned subgroup) in a manner that is determined by the randomization procedure. Since the experimenter generally lacks detailed information on the nuisance parameters, the problem is to appraise the sensitivity of the test functions (criterions) to variations in the nuisance parameters.

It is shown that the effect of the nuisance parameters on the distribution function of the test statistics is through statistics of non-centrality, analogous to the parameters of non-centrality of the F-statistics in the non-randomized case. The low order moments of the statistics of noncentrality are studied, and the distribution functions of the test statistics are approximated by linear contrasts of double non-central F-distributions multiplied by the central moments of the statistics of non-centrality.

(Abstract not connected with any meeting of the Institute.)

1. Some Property of a Sequence of Random Events. MAREK FISZ, University of Warsaw, Poland and University of Washington. (By title)

As the author is aware, the following simple theorem has never been published. Denote by $A_n (n = 1, 2, \cdots)$ a sequence of random events, $B = \cap \bar{A}_n$, $p_n = P(A_n)$, $v_n = P(A_{n+1} | \bar{A}_1 \cdots \bar{A}_n)$. Assume that $0 < p_n < 1, 0 < v_n < 1$ $(n = 1, 2, \cdots)$. Then $P(B) > 0$ if and only if (*) $\sum_1^\infty v_n < \infty$. It is known that if both of the relations $\sum_1^\infty p_n = \infty$ and (**) $\sum_1^\infty v_n = \infty$ hold, then $P(\lim_n \sup A_n) = 1$. If, however, (***) $\sum_1^\infty p_n < \infty$ and (**) hold, then by virtue of the Borel-Cantelli Lemma, $P(\lim_n \sup A_n) = 0$ while the probability of occurrence of at least one of the A_n is positive. If the A_n are independent, relations (*) and (***) are equivalent and the author's theorem asserts $P(B) > 0$ while the Borel-Cantelli Lemma asserts the weaker relation $P(\lim_n \sup A_n) = 0$.

CORRECTION TO ABSTRACT "MOMENTS OF THE RADIAL ERROR"

BY ERNEST M. SCHEUER

The following corrections should be made in the above-titled abstract (*Ann. Math. Stat.*, Vol. 32 (1961), p. 638). Replace sentences two and three by the following:

Let $\sigma_1^2 = \frac{1}{2}\{(\sigma_{11} + \sigma_{22}) + [(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}^2]^{\frac{1}{2}}\}$, $\sigma_2^2 = \frac{1}{2}\{(\sigma_{11} + \sigma_{22}) - [(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}^2]^{\frac{1}{2}}\}$, $k^2 = (\sigma_1^2 - \sigma_2^2)/\sigma_1^2$. Then the moments about the origin μ_n' of the radial error $R = [x_1^2 + x_2^2]^{\frac{1}{2}}$ are $\mu_n' = 2^{\frac{1}{2}n}\sigma_1^n \Gamma[\frac{1}{2}(n + 2)]F(-\frac{1}{2}n, \frac{1}{2}, 1; k^2)$ where $F(a, b, c; z)$ is the hypergeometric function.