AD-A113 027 PITTSBURGH UNIV PA DEPT OF MATHEMATICS AND STATISTICS F/G 12/1

MULTIVARIATE ANALYSIS AND ITS APPLICATIONS.(U)

DEC 81 P R KRISHNAIAH F49620-79-C-0161

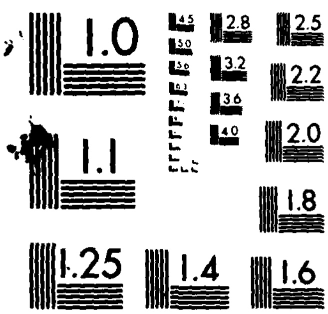
UNCLASSIFIED AFOSR-TR-82-0271 NL

END

DATE FILMED

4 82

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MULTIVARIATE ANALYSIS AND ITS APPLICATIONS

FINAL REPORT

Contract F49620-79-C-0161
June 1, 1979 - Dec. 14, 1981

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 82-0271	2. GOVT ACCESSION NO.: AD A113 027
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): MULTIVARIATE ANALYSIS AND ITS APPLICATIONS	5. TYPE OF REPORT & PERIOD COVERED: Final
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Prof. Krishnaiah	8. CONTRACT OR GRANT NUMBER(s): F49620-79-C-0161
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of Pittsburgh, Dept. of Math & Statistics, Pittsburgh, PA 15260	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS: AFOSR/NM, Bolling AFB, DC 20332	12. REPORT DATE
	13. NUMBER OF PAGES: 24
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS (of this report): unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report covers the work done under the contract F49620-79-C-0161 during the period of June 1, 1979 - Dec. 14, 1981. All this work is reported in various papers and the abstracts of these papers are attached. The work in these papers is supported completely or partially by the above contract.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

This report covers the work done under the contract F49620-79-C-0161 during the period of June 1, 1979 - Dec. 14, 1981. All this work is reported in various papers and the abstracts of these papers are attached. The work in these papers is supported completely or partially by the above contract.

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special

[1] Block, H. W. and Sampson, A. R. (1982). Inequalities on distributions: bivariate and multivariate. To appear in _Encyclopedia of Statistical Sciences_ (N. L. Johnson and S. Kotz, editors). John Wiley and Sons.

Inequalities concerning bivariate and multivariate distributions in statistics are surveyed, as well as historical background. Subjects treated include inequalities arising through positive and negative dependence, Boole, Bonferroni and Frechet inequalities, convex symmetric set inequalities; stochastic ordering; stochastic majorization and inequalities obtained by majorization, Chebyshev and Kolmogorov type inequalities; multivariate moment inequalities; and applications to simultaneous inference, unbiased testing and reliability theory.

[2] Burbea, J. and Rao, C. R. (1980). On the convexity of some divergence measures based on entropy functions. _IEEE Information Theory_, 1982, in press.

Three measures of divergence between vectors in a convex set of n-dimensional real vector space have been defined in terms of certain types of entropy functions, and their convexity property studied. Among other results, a classification of the entropies of degree α is obtained by the convexity of these measures. These results have applications in information theory and biological studies.

[3] Burbea, J. and Rao, C. R. (1980). Entropy differential metric, distance and divergence measures in probability spaces - a unified approach. Technical Report No. 80-18, Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

The paper is devoted to metrization of probability spaces through the introduction of a quadratic differential metric in the parameter space of the probability distributions. For this purpose, a ϕ-entropy functional is defined on the probability space and its Hessian along a direction of the tangent space of the parameter space is taken as the metric. The distance between two probability distributions is computed as the geodesic distance induced by the metric. The paper also deals with three measures of divergence between probability distributions and their inter-relationships.

[6] Burbea, J. and Rao, C. R. (1981). On the convexity of higher order Jensen differences based on entropy functions. Technical Report No. 81-20. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

In an earlier work, the authors introduced a divergence measure, called the first order Jensen difference, or in short, J-divergence, which is based on entropy functions of degree α. This provided a generalization of the measure of mutual information based on Shannon's entropy (corresponding to $\alpha = 1$). It was shown that the first order J-divergence is a convex function only when α is restricted to some range. In this paper we define higher order Jensen differences and show that they are convex functions only when the underlying entropy function is of degree 2. A statistical application requiring the convexity of higher order Jensen differences is indicated.

[7] Burbea, J. and Rao, C. R. (1981). Differential metrics in probability spaces. Technical Report No. 81-29. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

We discuss the construction of differential metrics in probability spaces through entropy functionals and examine their relationship with the information metric introduced by Rao using the Fisher information matrix in the statistical problem of classification and discrimination, and the classical Bergman metric. It is suggested that the scalar and Ricci curvatures associated with the Bergman-information metric may yield results in statistical inference analogous to those of Efron using the Gaussian curvature.

[6] Chikuse, Y. (1980). Invariant polynomials with real and complex matrix arguments and their applications. Technical Report No. 80-3. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

A class of homogeneous polynomials $C_{\phi}^{k_1,k_2,\ldots,k_r}(X_1,X_2,\ldots,X_r)$ with a general number r symmetric matrix arguments ($r \geq 1$), invariant under the orthogonal group, is proposed through the theory of polynomial representations of the linear group. The class of polynomials is a generalization of the invariant polynomials with smaller numbers of matrix arguments developed previously up to the case of three matrix arguments. Fundamental properties and relations of the polynomials $C_{\phi}^{k_1,k_2,\ldots,k_r}$ are shown, and some applications in multivariate distribution theory are indicated. The complex analogue of the above arguments is also discussed.

[7] Chikuse, Y. (1980). Simultaneous pseudo confidence regions for ratios of the discriminant coefficients. Technical Report No. 80-4. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

We consider simultaneous confidence regions for some hypotheses on ratios of the discriminant coefficients of the linear discriminant function when the population means and common covariance matrix are unknown. This problem, involving hypotheses on ratios, yields the so-called "pseudo" confidence regions valid conditionally in subsets of the parameter space. We obtain the explicit formulas of the regions and give further discussion on the validity of these regions. Illustrations of the pseudo confidence regions are given.

[8] Chikuse, Y. (1980). Distributions of some matrix variates and latent roots in multivariate Behrens-Fisher discriminant analysis. Technical Report No. 80-12. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

In this paper, it is shown that some distributions

of the matrix variates and latent roots arising in the multivariate Behrens-Fisher discriminant problem can be explicitly expressed in terms of the invariant polynomials with two matrix arguments, due to A. W. Davis, extending the zonal polynomials of matrix argument.

[9] Chikuse, Y. (1980). Robustness of a test statistic under the Gauss-Markov models. Technical Report No. 80-16, Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

A statistical procedure based on the simple Gauss-Markov model $(Y, X\beta, \sigma^2 I)$ is said to be (β, V)-optimal, if it remains valid under the general model $(Y, X\beta, \sigma^2 V)$. This robustness problem has been considered in an extensive literature concerning the BLUE and simple testing procedures in the least squares of the Gauss-Markov model. In this paper, we shall investigate the robustness of the likelihood ratio test statistic t_V for the test of hypothesis on estimable linear parametric functions and give a representation of V for $t_I = t_V$. This is a generalization of the result due to Ghosh and Sinha (*J. Multivariate Analysis*, 1980) who considered the special case when V has an intraclass correlation structure. We also consider the robustness for a system of seemingly unrelated regression equations as a special case.

[10] Chikuse, Y. (1980). Representations of the covariance matrix for robustness in the Gauss-Markov model. Technical Report No. 80-17. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

This paper considers some extensions of the results of Rao (1967) and Rao and Mitra (1969 and 1971, Chapter 8), who gave a table of general representations of the covariance matrix in terms of the given design matrix, under which various statistical procedures in the least squares theory based on the simple Gauss-Markov model with the spherical covariance matrix are also valid under the general Gauss-Markov. We shall give extended tables adding some more results relating to robustness, especially in connection with the estimation and testing of hypothesis on linear parametric functions.

[11] Chikuse, Y. (1980). Pseudo confidence regions for the solution of a multivariate linear functional relationship. Technical Report No. 80-16. Institute for Statistics and Applications, Mathematics and Statistics Department, University of Pittsburgh.

The problem of finding confidence regions (CR) for a q-variate vector β given as the solution of a linear functional relationship (LFR) $A\beta = \mu$ is investigated. Here an m-variate vector μ and an $m \times q$ matrix $A = (A_1, A_2, \ldots, A_q)$ are unknown population means of an $m(q+1)$-variate normal distribution $N_{m(q+1)}(\xi, \Omega \otimes \Sigma)$, where $\xi' = (\mu', A_1', A_2', \ldots, A_q')$, Σ is unknown, symmetric and positive definite $m \times m$ matrix and Ω be a known, symmetric and positive definite $(q+1) \times (q+1)$ matrix. This problem is a generalization of the univariate problem for the ratio of normal means.

A CR for β with level of confidence $1 - \alpha$, is given by a quadratic inequality, which yields the so-called "pseudo" confidence regions (PCR) valid conditionally in subsets of the parameter space. Our discussion is focused on the "bounded pseudo" confidence region (BPCR) given by the interior of a hyperellipsoid. The two conditions necessary for a BPCR to exist are shown to be the consistency conditions concerning the multivariate LFR. The probability that these conditions hold approaches one under "favourable circumstances" in many practical situations. Hence, we may have a BPCR with confidence approximately $1 - \alpha$. Some simulation results are presented.

[12] Eaton, M. L. and Kariya, T. (1981). On a general condition for null robustness. Accepted for publication in the Journal of Multivariate Analysis subject to revision.

This paper gives sufficient conditions that certain statistics have a common distribution under a wide class of underlying distributions. Invariance methods are the primary technical tools used in establishing the theoretical results. These results are applied to MANOVA problems, problems involving canonical correlations, and certain statistics associated with the complex normal distribution.

[13] Fang, [illegible] (1981). A survey of asymptotic distributions of functions [illegible] the eigenvalues of random matrices for nonnormal populations. Invited paper presented at the Eastern Regional Meeting of the Institute of Mathematical Statistics held in Philadelphia, May 1981.

In this paper, the author gave a review of the literature on various developments on certain aspects of multivariate distribution theory when the underlying distribution is not multivariate normal.

[14] Fang, C. and Krishnaiah, P. R. (1981). Asymptotic distributions of functions of the eigenvalues of the doubly noncentral F matrix and their applications. Submitted for publication.

Let S_1 and S_2 be distributed independently as noncentral Wishart matrices. Then the distribution of $S_1 S_2^{-1}$ is known to be doubly noncentral F matrix. In this paper, asymptotic distributions of various functions of the eigenvalues of $S_1 S_2^{-1}$ are derived. Applications of these results in the area of robustness of tests and other areas are discussed.

[15] Fang, C., Krishnaiah, P. R. and Nagarsenker, B. N. (1981). Asymptotic distributions of the likelihood ratio test statistics for covariance structures of the complex multivariate normal distribution. Accepted for publication, subject to revision in <u>Journal of Multivariate Analysis</u>.

In this paper, the authors obtained asymptotic expressions for the null and nonnull distributions of the likelihood ratio test statistic for multiple independence when the underlying distribution is complex multivariate normal. The authors also derived asymptotic null distribution of the likelihood ratio test statistic for homogeneity of the covariance matrices of several complex multivariate normal populations; asymptotic expressions are also obtained for the nonnull case when we have two populations. The expressions obtained in this paper are in terms of beta series.

[16] Fang, C., Krishnaiah, P. R. and Nagarsenker, B. N. (1981). Asymptotic distribution of the likelihood ratio test statistic for sphericity of complex multivariate normal distribution. Technical Report No. 81-27. Institute for Statistics and Applications, Mathematics and Statistics Department, University of Pittsburgh.

In this report, the authors approximated the distribution of certain power of the likelihood ratio test statistic for sphericity of the complex multivariate normal distribution with a beta type series in the null and non-null cases. Applications of these results to multiple time series are also discussed.

[17] Fang, C. and Krishnaiah, P. R. (1982). Asymptotic distributions of functions of the eigenvalues of some random matrices for nonnormal populations. Journal of Multivariate Analysis, (in press).

In this paper, the authors investigated the asymptotic joint distributions of certain functions of the eigenvalues of the sample covariance matrix, correlation matrix, and canonical correlation matrix in nonnull situations when the population eigenvalues have multiplicities. These results are derived without assuming that the underlying distribution is multivariate normal. In obtaining these expressions, the authors used Edgeworth type expressions.

[18] Fang, C. and Krishnaiah, P. R. (1981). Asymptotic distributions of functions of the eigenvalues of the real and complex noncentral Wishart matrices. In Statistics and Related Topics (Csorgo, M., Dawson, D., Rao, J. N. K. and Saleh, A. K. Md. E., editors). North-Holland Publishing Company.

In this paper, the authors investigated asymptotic distributions of functions of the eigenvalues of the noncentral Wishart matrix. Applications of these results are discussed to a problem in cluster analysis and to the problem of studying the structure of interaction in two-way classification with one observation per cell.

[19] Fang, C. and Krishnaiah, P. R. (1982). Asymptotic joint distributions of functions of the elements of sample covariance matrix. In _Statistics and Probability: Essays in Honor of C. R. Rao_. (G. Kallianpur, P. R. Krishnaiah and J. K. Ghosh, editors). North-Holland Publishing Company.

In this paper, the authors give asymptotic expressions for the joint distribution of the functions of the elements of the sample covariance matrix and sample correlation matrix in the noncentral case when the underlying distribution is multivariate normal. Accuracy of the above expressions is also studied. Also, asymptotic expressions are given for functions of the elements of the sample covariance matrix for nonnormal populations. Finally, some applications of the above results are discussed.

[20] Janardan, K. G. (1981). On characterizing the Markov-Polya distribution. Technical Report No. 81-[illegible]. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

A discrete survival model is considered where an unobservable random variable is subjected to destruction so that what is observed and recorded is only the undamaged part Y of X. Assuming the destruction process is represented by the Markov-Polya distribution, a characterization of the negative binomial distribution is obtained. Utilizing the completeness property of the negative binomial distribution, a characterization of the Markov-Polya distribution is derived. Several other characterization theorems are also proved concerning these probability distributions.

[21] Janardan, K. G. and Rao, B. Raja (1981). Characterizations of generalized Markov-Polya and generalized Polya-Eggenberger distributions. Technical Report No. 81-03. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

A discrete model is considered where the original observation is subjected to partial destruction according to

the generalized Markov-Polya damage model. A characterization of the generalized Polya-Eggenberger distribution is given in the context of the Rao-Rubin condition. Several other characterization theorems are also proved concerning these probability distributions.

[22] Kariya, T. (1980). Robustness of multivariate tests. <u>Annals of Statistics</u>.

This paper gives necessary and sufficient condition for the null distribution of a test statistic to remain the same in the class of left O(n)-invariant distributions. Secondly, it is shown that in certain special cases, the usual MANOVA tests are still uniformly most powerful invariant in a class of left O(n)-invariant distributions.

[23] Kariya, T. (1981). A new concept of second order efficiency and its application to a missing data problem. Technical Report No. 81-16. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

In this paper, following the framework of Kariya, Krishnaiah and Rao (1981), a finite sample concept of second order efficiency is defined and it is applied to the analysis of the problem of estimating bivariate normal parameters with extra data on the first variate. Only the modified maximum likelihood estimator is shown to be second order efficient.

[24] Kariya, T., Krishnaiah, P. R. and Rao, C. R. (1981). Inference on the parameters of multivariate normal population when some data is missing. In <u>Developments in Statistics, Vol. 4</u> (in press). Academic Press, New York.

In this paper, the authors considered the problems of estimation and tests of hypotheses on the parameters of the multivariate normal population when the data is incomplete. Special emphasis is made for the case of the bivariate normal when the variances of the variables are equal. The authors also obtained certain expressions for the asymptotic distributions of a wide class of test statistics useful in testing various hypotheses on the parameters of the multivariate normal when some of the data is missing.

[25] Kariya, T., Sinha, B. K. and Subramanyam, K. (1981). Nearly efficient estimators based on order statistics. Technical Report No. 81-05. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

In this paper, based on 3 or 5 order statistics out of n order statistics, nearly asymptotically efficient estimators are obtained for Cauchy, logistic and normal distributions.

[26] Kariya, T., Sinha, B. K. and Subramanyam, K. (1981). First, second and third order efficiencies of the estimators for a common mean. To be given for typing.

Based on Kariya, Krishnaiah and Rao (1981) and Kariya (1981), this paper considers efficiencies of several estimators proposed in the problem of estimating a common mean of K univariate normal populations. Only the Graybill-Deal (1959) estimator is shown to be third order efficient.

[27] Kariya, T., Sinha, B. K. and Krishnaiah, P. R. (1981). Some properties of left orthogonally invariant distributions. Technical Report No. 81-28. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

In this note, we have considered generalizations of some of the results on spherical distributions to the case of left spherical matrix variate distributions. First, it is shown that the independence of n rows of an $n \times p$ left spherical random matrix implies multivariate normality. Secondly, most results in Eaton (1981) are extended to the matrix variate case.

[28] Kariya, T., Sinha, B. K. and Subramanyam, K. (1981). Berkson's bioassay problem - revisited. Technical Report No. 81-24. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

Consider Berkson's problem of estimating θ on the basis of independent random variables X_i having the binomial distribution $B(n_i, \pi_i(\theta))$ $(i = 1,\ldots,k)$, where

$\pi_i(\theta) = [1 + \exp(-\theta-\beta d_i)]^{-1}$, and β and d_i's are known. Extensive asymptotic and nonasymptotic comparisons of two particular estimates of θ, the MLE $\hat{\theta}_M$ and the Rao-Blackwellized version $\hat{\hat{\theta}}_B$ of Berkson's estimator $\hat{\theta}_B$, have been made in the literature. Our subject in this paper is three fold: (a) to derive a simple necessary condition for $\hat{\hat{\theta}}_B$ to dominate $\hat{\theta}_M$ uniformly in the MSE criterion; (b) to propose three new estimators and compare them with $\hat{\hat{\theta}}_B$ and $\hat{\theta}_M$; (c) to show that $\hat{\theta}_B$ and $\hat{\theta}_M$ are the first order efficient no matter what $\pi_i(\theta) = F(\theta+\beta d_i)$ may be so long as F is strictly increasing.

[29] Khatri, C. G. and Rao, C. R. (1981). Some extensions of the Kantorovich inequality and statistical applications. J. Multivariate Analysis, 11.

Kantorovich gave an upper bound to $(x'Vx)(x'V^{-1}x)$ where x is an n-vecotr of unit length and V is an n×n positive definite matrix. Bloomfield, Watson and Knott found the bound to $|X'V X X'V^{-1}X|$, and Khatri and Rao to the trace and determinant of $X'V Y Y'V^{-1}X$ where X and Y are n×k matrices such that $X'X = Y'Y = 1$. In the present paper we establish bounds for traces and determinants of $X'V Y Y'V^{-1}X$ and $X'BYY'CX$ when X and Y are matrices of different orders. A review of previous results on generalizations of Kantorovich inequality and a number of new results of independent interest are also given.

[30] Kimeldorf, C., May, J. and Sampson, A. R. (1980). Concordant and discordant monotone correlations and their evaluation by nonlinear optimization. TIMS Studies in the Management Sciences: Optimization in Statistics Volume (to appear).

This paper presents four new statistical measures of monotone relationship derived from the concept of monotone correlation. A nonlinear optimization algorithm is employed to evaluate these new measures, as well as the monotone correlation, for ordinal contingency tables. A computer program to implement the algorithm is developed, and is applied to several insightful examples to provide further understanding of the usefulness of these measures.

[31] Kimeldorf, G., May, J. and Sampson, A. (1981). MONCOR - A program to compute concordant and other monotone correlations. Computer Science and Statistics: Proceedings of the 13th Symposium on Interface. Springer-Verlag, New York.

The new interactive FORTRAN program MONCOR is described. MONCOR computes the concordant monotone correlation, discordant monotone correlation, isoconcordant monotone correlation, isodiscordant monotone correlation and their associated monotone variables. Data input can be finite discrete bivariate probability mass functions or ordinal contingency tables, both of which must be given in matrix form. The well-known British Mobility data are used to illustrate the input and output options available in MONCOR.

[32] Kimeldorf, G., Plachky, D. and Sampson, A. R. (1981). A simultaneous characterization of the Poisson and Bernoulli distributions. J. Appl. Probability **18**, 316-320.

Let N, $X_1, X_2, \ldots$ be non-constant independent random variables with $X_1, X_2, \ldots$ being identically distributed and N being nonnegative and integer valued. It is shown that the independence of $\sum_{i=1}^{N} X_i$ and $N - \sum_{i=1}^{N} X_i$ implies that the X_i's have a Bernoulli distribution and N has a Poisson distribution. Other related characterization results are considered.

[33] Krishnaiah, P. R. (1981). Computations of some multivariate distributions. Handbook of Statistics, Vol. 1 (P. R. Krishnaiah, editor), 745-971. North-Holland Publishing Company.

This paper reviews the present state-of-the-art on computations of functions of the roots of several random matrices which arise in multivariate statistical analysis. Many of the available tables useful in statistical analysis of multivariate data are given in the Appendix.

[34] Krishnaiah, P. R. and Lee, J. C. (1981). Likelihood ratio tests for mean vectors and covariance matrices. Handbook of Statistics, Vol. 1. (P. R. Krishnaiah, editor). North-Holland Publishing Company.

This paper discusses computational aspects of the distributions of various likelihood ratio test statistics for mean vectors and covariance matrices.

[35] Krishnaiah, P. R. and Yochmowitz, T. G. (1981). Inference on the structure of interaction in two-way classification model. Handbook of Statistics (P. R. Krishnaiah, editor), 973-994. North-Holland Publishing Company.

In this paper, the authors discussed various tests for studying the structure of interactions in two-way classification model with one observation per cell. These test basically involve various functions of the eigenvalues of the Wishart matrix.

[36] Lau, K. and Rao, C. R. (1981). Integrated Cauchy functional equation and characterizations of the exponential law. Sankhyā, 44.

A general solution of the functional equation

$$\int_0^\infty f(x+y)d\mu(y) = f(x)$$

where f in a nonnegative function and μ is a positive Borel measure on $[0,\infty]$ is shown to be $f(x) = p(x) \exp(\lambda x)$ where p is a periodic function with every $y\epsilon\underline{\mu}$, the support of μ, as a period. The solution is applied in characterizing Pareto, exponential and geometric distributions by properties of integrated lack of memory, record values, order statistics and conditional expectation.

[37] Lau, K. and Rao, C. R. (1981). Solution to the integrated Cauchy functional equation on the whole line. Sankhyā, 44.

A general solution of the integrated Cauchy functional equation

$$\int_{-\infty}^{\infty} f(x+y)d\mu(y) = f(x) \text{ a.e. for } x \quad (-\infty,\infty)$$

where f is a locally integrable positive function and μ is a positive Borel measure on R is shown to be

$$f(x) = p_1(x)e^{\lambda_1 x} + p_2(x)e^{\lambda_2 x} \text{ a.e.}$$

where p_1 and p_2 are positive periodic functions with every u $\underline{\mu}$, the support of μ, as a period. A variant of the Choquet-Deny theorem on μ-harmonic functions is given.

[38] Rao, B. R. and Janardan, K. G. (1981). Characterization theorems involving the generalised Markov-Polya damage model. Technical Report No. 81-04. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

In the present paper, certain random damage models are examined, such as the Generalized Markov-Polya and the Quasi-Binomial, in which an integer-valued random variable N is reduced of N_A. If N_B is the missing part, where $N = N_A + N_B$, the covariance between N_A and N_B is obtained for some general classes of distributions, such as the G.P.S.D. and M.P.S.D. for the random variable N. A characterization theorem is proved that under the generalized Markov-Polya damage model, the random variables N_A and N_B are independent if, and only if, N has the Generalized Polya-Eggenberger distribution. This generalizes the corresponding result fo r the Quasi-Binomial damage model and the generalized Poisson distribution. Finally, some interesting identities are obtained using the independence property and the covariance formulas between the numbers N_A and N_B.

[39] Rao, C. R. (1979). Fisher efficiency and estimation of several parameters (Fisher memorial lecture given at the joint IMS-ASA-BS meeting in Washington, DC August 1979).

In this paper, the role of Minimum Mean Square Error as a general criterion for estimation of parameters is critically examined. It is shown that smaller mean square error does not necessarily imply greater concentration of the estimator around the true value. The empirical Bayes method for simultaneous estimation of parameters introduced by Fisher is shown to provide a good ranking of populations for individual populations.

[40] Rao, C. R. and Kleffe, J. (1979). Estimation of variance components. Handbook of Statistics, Vol. 1 1-40. North-Holland Publishing Company.

This paper provides a comprehensive review of the methods of estimation of variance components. Conditions for identifiability and estimability of variance components have been given. Methods of minimum variance unbiased estimation in the normal case, MINQE in the general case and ML estimation in the normal case have been discussed.

[41] Rao, C. R. (1981). Diversity and dissimilarity coefficients: a unified approach. J. Theoretical Population Biology.

Three general methods for obtaining measures of diversity within a population and dissimilarity between populations are discussed. One is based on an intrinsic notion of dissimilarity between individuals and others make use of the concepts of entropy and discrimination. The use of a diversity measure in apportionment of diversity between and within populations is discussed.

[42] Rao, C. R. (1981). Some comments on the minimum mean square error as a criterion of estimation. In Statistics and Related Topics, 123-144. North-Holland Publishing Company.

It is shown that estimators obtained by MMSE (minimizing the mean square error) may not have optimum properties with respect to other criteria such as PN (probability of nearness ot the true value in the sense of Pitman) or PC (probability of concentration around the true value). In particular, a detailed study is made of estimators obtained by shrinking the minimum variance unbiased estimators to reduce the MSE. It is suggested that because of mathematical convenience and some intuitive considerations, MMSE could be used as a primitive postulate to derive estimators, but their acceptability should be judged on more intrinsic criteria such as PN and PC.

[43] Rao, C. R. (1981). A lemma on g-inverse of a matrix and computation of correlation coefficients in the singular case. Communications in Statistics A 10 1-10.

Formulae for multiple, partial and canonical correlations are generally expressed in terms of the elements of the inverse covariance matrix of the variables. These are not valid when the covariance matrix is singular. Appropriate formulae using g-inverse are developed.

The results depend on some basic lemmas on the idempotent matrix AA^- where $AA^-A = A$ (g-inverse as defined by Rao, 1962) and on the spectral decomposition of a hermitian matrix with respect to a non-negative definite matrix.

[44] Khatri, C. G. and Rao, C. R. (1981). Some generalizations of the Kantorovich inequality. To appear in Sankhyā, 44.

Inequalities are obtained for expressions of the type $|X'AYY'A^{-1}X|$, $|X'AY(Y'BAY)^{-1}Y'BX|/|X'ABX|$, $|X'B^2X|\ |X'C^2X|/|X'BCX|^2$, etc., for variations in matrices X and Y given matrices A, B, C. When X and Y are vectors we have the Kantorovich inequality. The inequalities when X and Y are matrices have applications in determining the efficiencies of estimators of several parameters in linear models.

[45] Rao, C. R. (1979). Perspectives in Statistics. Sankhya 41 129-137. (This is Presidential Address at the meeting of the International Statistical Institute held in Manila.)

The paper refers to the expanding frontiers of knowledge in statistics and the current controversies, and outlines the role that the International Statistical Institute can play in directing future research to make statistics a socially meaningful and viable science. Suggestions are made about the training of statisticians with

a proper blend of theoretical knowledge and skill in applications, development of statistical courses for specialists in other disciplines, the role of government statisticians, and the use of computers in statistical research. Some examples are given to highlight the difficulties involved in defining the efficiency of an estimator and the possible dangers in the uncritical use of methods developed by academic statisticians in practical work.

[46] Ghosh, J. K. and Subramanyam, K. (1981). Estimation in separated families - two parameter case. Technical Report No. 81-32. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

In an earlier paper Ghosh and Subramanyam (1975) which will be referred as GS (1975) studies some properties of the mle (maximum likelihood estimator) of a discrete parameter. An expression for the asymptotic risk of the mle was obtained in the case of a single discrete parameter. But in most practical cases one has one discrete and one or more continuous parameter. For more examples refer to GS (1975) and Cox (1962). In such a case an expression for the asymptotic risk of the mle of the discrete parameter is given in section 2. Two examples are discussed. The first example is that considered by Cox (1962), deciding between two distributions, Poisson or geometric. Our asymptotic theory is different from that of Cox. The second example is B(N,p), both N and p unknown.

[47] Rao, C. R., Sinha, B. K. and Subramanyam, K. (1981). Third order efficiency of the maximum likelihood estimator in the multinomial distribution. Technical Report No. 81-21. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

Consider a multinomial population with k cells. Let $\pi_1(\theta),\ldots,\pi_k(\theta)$ be the population proportions and $p_1,\ldots,p_k$ be the sample proportions out of a sample of size n. We restrict our attention to the class of estimators T_n obtained as the solution of the equation:

$$0 = \sum_1^k f\left(\frac{p_r}{\pi_r(T_n)}\right) \pi_r'(T_n).$$

We show that a n.s. condition for T_n to be second order efficient (SOE) is that $f'(1) \neq 0$, $f''(1) = 0$. In the subclass of SOE estimators let T_n by any estimator such that the corresponding $f'''(1) \neq 0$. It is proved that there exist bias-adjusted versions of $\hat{\theta}_n$, the mle, and T_n, say $\hat{\theta}_n^*$ and T_n^*, such that $\hat{\theta}_n^*$ has a smaller mean squared error up to $o(n^{-2})$ than T_n^*. Here $\hat{\theta}_n^*$ and T_n^* have the same bias up to $O(n^{-2})$. This generalizes the SOE property of the mle proved earlier by Rao (1963)).

[48] Rao, C. R. (1981). Gini-Simpson index of diversity: a characterization, generalization and applications. Utilitas Mathematica, 21.

The Gini-Simpson index of diversity of a multinomial distribution defined by k cell probabilities $p = (p_1, \ldots, p_k)$ is $1 - \Sigma p_i^2 = 1 - p'p$. In this paper, this index [illegible] characterized by a set of postulates based on [illegible] of distance between distributions. Further, [illegible] suggested that in practical applications involving measurement of qualitative variation it is more appropriate to use an index of the type $p'\Delta p$ where Δ is a $k \times k$ matrix of assigned distances between attributes characterizing the cells of the multinomial distribution. The Gini-Simpson index corresponds to a special choice of Δ. The expression $p'\Delta p$ may also be interpreted as a generalized quadratic entropy. An example is worked out illustrating analysis of diversity and cluster analysis based on the measure $p'\Delta p$.

[49] Rao, C. R. (1981). Analysis of diversity: a unified approach. Technical Report No. 81-26. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

A general method of constructing a measure of diversity within a population subject to some postulates is described. It is shown that such a diversity measure, when it satisfies a concavity condition, can be used to analyse diversity in a mixture of populations as due to main effects and interactions of factors, by the levels of which the individual populations are specified. The method is applicable to both quantitative and categorical response data, and provides a generalization of the technique of analysis of variance. Some applications are considered.

[50] Rao, C. R. (1981). Optimization of functions of matrices with applications to statistical problems. Technical Report No. 81-31. Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

The paper reviews some methods of optimizing functions of vectors and matrices subject to some restrictions and developes techniques for solving them without using the calculus of matrix derivatives. Their application to a number of statistical problems in linear estimation and multivariate analysis is illustrated.

[51] Rao, C. R. (1982). Diversity: its measurement, decomposition, apportionment and analysis. Sankhyā, 44 (in press).

Two general methods of obtaining measures of diversity within a population are discussed. One is based on an intrinsic notion of dissimilarity between individuals and the other makes use of the concepts of entropy.

Some examples are given of the decompositon of diversity within a population in terms of given or conceptual factors.

Methods for apportionment of diversity in a hierarchically classified set of populations are discussed.

The concept of analysis of diversity as a generalization of analysis of variance is developed for populations classified by combinations of different levels of chosen factors.

[32] Sampson, A. (1980). Nonnegative Cholesky decomposition and its application to association of random variables. *SIAM J. on Algebraic and Discrete Methods* 1 284-291.

The concept of a multivariate family of distributions indexed by a covariance scale parameter Σ is formally defined and examples are given. The multivariate normal is one such family. Sufficient conditions are given so that a positive definite matrix has a nonnegative Cholesky decomposition. These conditions also yield the association of random variables with a covariance scale parameter distribution. These results are related to other matrix results and to Barlow and Proschan's stronger conditions (Statistical Theory of Reliability and Life Testing: Probability Models, Holt, Rinehart, Winston, New York, 1975), for the association of the multivariate normal, namely, $\lambda_{ij} \le 0$, $i \ne j$ where $\Sigma^{-1} = \Lambda = (\lambda_{ij})$.

[33] Sampson, A. R. (1980). Positive dependence properties of elliptically symmetric distributions. Accepted subject to revision in *Journal of Multivariate Analysis*.

Let $X_1,\ldots,X_p$ have p.d.f. $g(x_1^2+\cdots+x_p^2)$. It is shown that (a) $X_1,\ldots,X_p$ are positively lower orthant dependent or positively upper orthant dependent if, and only if, $X_1,\ldots,X_p$ are i.i.d. $N(0,\sigma^2)$; and (b) the p.d.f. of $|X_1|,\ldots,|X_p|$ is TP_2 in pairs if, and only if, ln $g(u)$ is convex. Let X_1, X_2 have p.d.f. $f(x_1,x_2) = g((x_1,x_2)\Sigma^{-1}(x_1,x_2)')$. Necessary and sufficient conditions are given for $f(x_1,x_2)$ to be TP_2 for fixed correlation ρ. It is shown that if f is TP_2 for all $\rho > 0$, then $(X_1,X_2)' \sim N(0,\Sigma)$. Related positive dependence results and applications are also considered.

[54] Sampson, A. R. (1980). A multivariate correlation ratio. Technical Report No. 80-6, Institute for Statistics and Applications, Department of Mathematics and Statistics, University of Pittsburgh.

A brief review of the historical background and certain known results concerning the univariate correlation ratio are given. A multivariate correlation ratio of a random vector Y upon a random vector X is defined by $\eta_A(Y;X) = [tr(A^{-1} Cov\, E(Y|X))]^{\frac{1}{2}}[tr(A^{-1}\Sigma)]^{-\frac{1}{2}}$, where A is a given positive definite matrix. The properties of η_A are discussed, with particular attention paid to a "correlation-maximizing" property. A number of examples include the multivariate normal, the elliptically symmetric distributions, and the multinomial. The problem of maximizing $\eta_A(BY;X)$ over suitable matrices B is considered and the results that are obtained are related to canonical correlations for the multivariate normal.

[55] Sampson, A. and Smith, R. (1980). Assessing risks through the determination of rare event probabilities. Accepted subject to revision in Operations Research.

We consider the problem in risk assessment of evaluating the probability of occurrence of rare, but potentially catastrophic, events. The lack of historical data due to the sheer novelty of the event makes conventional statistical approaches inappropriate. The problem is compounded by the complex multivariate dependencies that may exist across potential event sites. In order to evaluate the likelihood of one or more such catastrophic events occurring, we provide an information theoretic model for merging a decision maker's opinion with expert judgment. Also provided is a methodology for the reconciling of conflicting expert judgments. This merging approach is invariant to the decision maker's viewpoint in the limiting case of exceptionally rare events. These methods are applied to certain case studies.

[36] Sampson, A. R. (1981). Stochastic approximation. To appear in *Encyclopedia of Statistical Sciences*. John Wiley and Sons.

This article reviews univariate and multivariate stochastic approximation procedures from the Robbins-Monro and Kiefer-Wolfowitz procedures to newly derive [illegible] estimation procedures. Variations of these procedures are also described. Applications of stochastic approximation procedures are also considered.

[37] Sinha, B. K. and Mouqadem, O. (1981). Estimation of the common mean of two univariate normal populations. Technical Report No. 81-07. Institute for Statistics and Applications, Department of Mathematics and Statistics University of Pittsburgh.

The problem of estimating the common mean μ of two univariate normal populations with unknown and unequal variances is considered from a decision-theoretic point of view. We restrict our attention to an appropriate class C and its three subclasses C_0, C_1, C_2 of unbiased estimates of μ. It is proved that the usual estimate of μ, which is the weighted linear combination of the sample means with weights as reciprocals of the sample variances, is admissible in C_0 and extended admissible in C but is neither Bayes nor limiting Bayes in C_2. Admissible estimates in C_1 and C_2 are obtained. The loss is always assumed to be squared error.

[38] Sinha, B. K. and Sarkar, S. K. (1982). Invariant confidence sequences for some parameters in a general linear regression model.

Let $X_1, X_2, \ldots$ be independent p-variate normal vectors with $E X_\alpha = \theta Y_\alpha$, $\alpha = 1, 2, \ldots$ and same p.d. dispersion matrix Σ. Here θ: $p \times q$ and Σ are unknown parameters and Y_α's are known $q \times 1$ vectors. Writing $\theta = (\theta_1' \; \theta_2')' = (\theta_{(1)} \theta_{(2)})$ with θ_i: $p_i \times q (p_1 + p_2 = p)$ and $\theta_{(i)}$: $p \times q_i (q_1 + q_2 = q)$, we have constructed invariant confidence sequences for (i)θ, (ii)$\theta_{(1)}$, (iii)θ_1 when $\theta_2 = 0$ and (iv)$\sigma^2 = |\Sigma|$. This uses the basic ideas of Robbins (1970) and generalizes some of his and Lai's (1976) results. In the process

alternative simpler solutions of some of Khan's results (1978) are obtained.