AD-A063 906 IOWA STATE UNIV AMES STATISTICAL LAB F/G 12/1
ORDER STATISTICS AND NONPARAMETRIC STATISTICS.(U)
DEC 78 H A DAVID DA-ARO-D-31-124-73-G55
UNCLASSIFIED ARO-11259.18-M NL

END
DATE FILMED
4-79
DDC

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) ARO 11259.18-M and 13198.19-M

AD A063906

DDC FILE COPY

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
11259-M
13198-M

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
ORDER STATISTICS AND NONPARAMETRIC STATISTICS.

5. TYPE OF REPORT & PERIOD COVERED
Final Report
10/1/1972-9/30/1978

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Herbert A. David
H. A. David

8. CONTRACT OR GRANT NUMBER(s)
DA-ARO-D31-124-73-G55,
DAAG 29-76-G-0057

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Statistical Laboratory
Iowa State University
Ames, Iowa 50011
H. A. David

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Final rept. 1 Oct 72-30 Sep 78

11. CONTROLLING OFFICE NAME AND ADDRESS
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709

12. REPORT DATE
December 1978

13. NUMBER OF PAGES
eight

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)
11259.18-M,
13198.19-M

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

ARO, ARO

DDC RECEIVED JAN 30 1979 C

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

18. SUPPLEMENTARY NOTES
The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Order statistics; Concomitants of order statistics; Outliers; Robustness; Tables, order statistics; Characterizations; Nonparametric statistics; Competing risks; Cyclic designs

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
The research reported falls under the headings: (a) order statistics; (b) competing risks; (c) nonparametric statistics, and (d) cyclic designs.

(a) Concomitants of order statistics were introduced and studied as a class, and applications in selection and double sampling provided. The robustness of various estimators of location and scale in the presence of outliers was examined and the necessary tables prepared. Characterizations of distributions by properties of order statistics were given. A text on order statistics is

→ next page

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

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

79 01 17 011

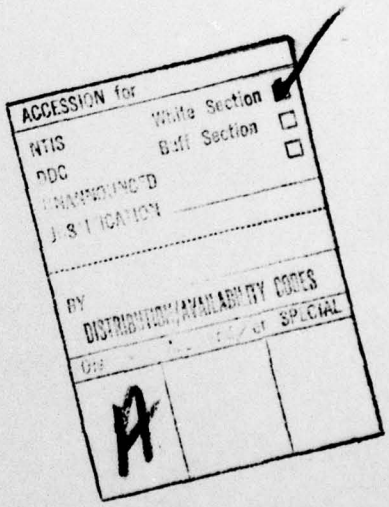
333 300

being revised.

(b) A parametric approach to the theory of competing risks was developed and a monograph on competing risks prepared. Applications to reliability under multiple causes of failure are included.

(c) A number of problems were resolved in nonparametric estimation and in the asymptotic theory of linear rank statistics, U-statistics, and quantiles.

(d) Following earlier work by the author a monograph on cyclic designs was co-authored. This is in a form ready for use, with worked examples provided.



ORDER STATISTICS AND NONPARAMETRIC STATISTICS

Final Report

H. A. David

December 21, 1978

U. S. Army Research Office

Grants
DA-ARO-D31-124-73-G55
DAAG 29 76 G 0057

Iowa State University

1. FOREWORD

This Final Report covers Grants DA-ARO-D31-124-73-G55 and DAAG 29 76 G 0057 for the period October 1, 1972 to September 30, 1978. During this time the following individuals received some salary support:

H. A. David, Principal Investigator

J. Galambos, Visiting Professor (Temple University)

B. C. Arnold, Professor (Iowa State University)

M. Ghosh, Visiting Professor (Indian Statistical Institute)

W. J. Conover, Visiting Professor (Texas Tech University)

M. J. O'Connell, Research Assistant (Ph.D. 1974)

S. S. Yang, Research Assistant (Ph.D. 1976)

V. S. Shu, Research Assistant (Ph.D. 1978)

2. Table of Contents - None

3. List of Appendixes, Illustrations and Tables - None

4. BODY OF REPORT

A. Order Statistics

1. Concomitants of Order Statistics. [B6,8,9,14,16;C1]*. These statistics were named and first studied as a class by the Principal Investigator. In a random sample of n pairs (X_i, Y_i), $i = 1,2,\ldots,n$, drawn from a bivariate distribution let $X_{r:n}$, $r = 1,2,\ldots,n$, be the r-th order statistic among the X_i and let $Y_{[r:n]}$ be the Y- variate paired with $X_{r:n}$. We have called $Y_{[r:n]}$ the concomitant of the rth order statistic. The properties of these concomitants were studied extensively and two main applications were developed. The first, to some double sampling situations, showed that if an inexpensive auxiliary variable X could be measured on n items, then the ordering of the X's could be used to advantage in the selection of a smaller number of items for measurement of a correlated expensive main variable Y of interest. The second application addressed itself to the following selection problem; If an individual A_i has rank r among n candidates scoring $X_1,\cdots, X_n$ in a certain test, what is the probability that A_i will have rank s in some future test with scores $Y_1,\ldots,Y_n$? What is his expected rank? Tables permitting immediate application of the methods advocated were computed when X and Y are bivariate normal. A general theory of concomitants of order statistics has now been developed.

2. Robustness in the Presence of Outliers. [B 7, 10; C2]. Within the broad and active field of robust estimation it is important to know how various estimators of location and scale behave when one or more outliers contaminate the sample of interest. In preparation for such a study an extensive set of tables [B7] was prepared giving the expected values, variances, and covariances of order statistics in normal samples, with one observation differing either in mean or variance. With the help of these tables it was possible to obtain the bias and mean square error, in the presence of an outlier, of estimators expressible as linear functions of order statistics. Further work is in progress, mainly by simulation methods, on the bias and mean square error of more complex estimators, including adaptive ones, in the presence of one or two outliers.

3. Characterizations by Properties of Order Statistics. [B1,11;D8] The subject of characterizations has seen a burst of activity in recent years. These papers are of more theoretical interest in throwing light on the nature of the distributions characterized.

*Entries in square brackets refer to the bibliography in Section 3 of this report.

4. Order Statistics Text. The Principal Investigator's research monograph and text Order Statistics, published by John Wiley & Co. in 1970, is scheduled for a second edition in 1980. Work has been proceeding steadily on the preparation of an updated and extended version. This is a considerable task in view of the large literature on order statistics since 1970. It is expected that the number of references selected for inclusion will be raised from 700 to at least 1000.

B. Competing Risks [A 1; B 4, 5]

Work in this area, begun in 1969 by the Principal Investigator in conjunction with a Ph.D. student M. L. Moeschberger has culminated in the publication of the monograph A 1. The subject has a formal connection with order statistics in that it is concerned with properties of the minimum of a set of random variables (the hypothetical lifetimes associated with the risks under study). The monograph emphasizes parametric approach developed by the authors but provides a much more general account as indicated in the following summary.

The theory of competing risks is concerned with the assessment of a specific risk in the complicating presence of other risks. It is of particular interest to statisticians confronted by reliability or survival analyses, as well as to vital statisticians, demographers, and actuaries. Non-experimental applications to life table data have a long history. The emphasis of this monograph is, however, on experimental situations or observational studies, often of modest size, for which the more traditional actuarial methods are not designed.

An individual's death from one of several competing causes is analogous to the failure of a series system due to the malfunctioning of an identified component. This parallel is exploited to unify applications to biometry and reliability. Parametric assumptions on distributions of lifetime are examined in detail when the risks are assumed to be independent. A number of models for dependent risks are also studied and problems of non-identifiability in this case are discussed. Recent developments incorporating concomitant measurements are reviewed.

Many of the methods described as well as certain graphical techniques are illustrated on two data sets giving survival time and cause of death after a treatment for cancer and after heart surgery. The monograph concludes with some exercises and a historical note.

C. Nonparametric Statistics [B 12, 13, 15;D 1, 2, 3, 4, 5, 6, 9]

Dr. Malay Ghosh, in conjunction with colleagues and students, has been successful in resolving a number of problems in (a) nonparametric estimation, (b) asymptotic distribution theory of linear rank statistics and U-statistics, and (c) asymptotic theory of quantiles. These papers are of considerable interest to mathematical statisticians and several constitute significant advances. Dr. Conover, together with a student, provided tables and approximations to facilitate the use of two well-known nonparametric tests.

D. Cyclic Designs [A 2; B 2]

These items round off research which arose out of the Principal Investigator's work on the method of paired comparisons described in earlier reports. Originally the aim was to fill the need for additional designs for paired-comparison experiments. These designs, of block size two, used in preference testing, were extended later to any block size. An extensive tabulation of such designs, with aids for their analysis is given in A2 which also provides worked examples. The P.I. has been invited to supply the entry on Cyclic Designs for the Wiley Encyclopedia of the Statistical Sciences.

5. BIBLIOGRAPHY

A. Monographs

1. David, H. A. and Moeschberger, M. L. (1978). The Theory of Competing Risks. Griffin, London and McMillan, New York. vii + 103 pp.

2. John, J. A., Wolock, F. W., and David, H. A. (1972). Cyclic Designs, National Bureau of Standards Applied Mathematics Series No. 62. vi + 76 pp.

B. Published Papers

1. Arnold, B. C. and Ghosh, M. (1976). "A Characterization of Geometric Distributions by Distributional Properties of Order Statistics," Scandinavian Actuarial Journal, 232-234.

2. David, H. A. (1972). "Enumeration of Cyclic Graphs and Cyclic Designs,: Journal of Combinatorial Theory B 13, 303-308.

3. David, H. A. (1973). "Waiting Time Paradoxes and Order Statistics," Journal of the American Statistical Association 68, 743-745.

4. David, H. A. (1974). "Parametric Approaches to the Theory of Competing Risks," in F. Proschan and R. J. Serfling (eds.) Reliability and Biometry, Society for Industrial and Applied Mathematics, 275-290.

5. David, H. A. (1976). "The Theory of Competing Risks," (with discussion). Australian Journal of Statistics 18, 101-114.

6. David, H. A. and Galambos, J. (1974). "The Asymptotic Theory of Concomitants of Order Statistics," Journal of Applied Probability, 11, 762-770.

7. David, H. A., Kennedy, W. J., and Knight, R. D. (1977). "Means, Variances, and Covariances of Normal Order Statistics in the Presence of an Outlier," Selected Tables in Mathematical Statistics, sponsored by the Institute of Mathematical Statistics, Vol. 5, 75-204.

8. David, H. A. and Levine, D. N. (1972). "Ranked Set Sampling in the Presence of Judgement Error," Biometrics 28, 553-555.

9. David, H. A., O'Connell, M. J., and Yang, S. S. (1977). "Distribution and Expected Value of the Rank of a Concomitant of an Order Statistics," Annals of Statistics 5, 75-204.

10. David, H. A. and Shu, V. S. (1978). "Robustness of Location Estimators in the Presence of an Outlier," In Contributions to Survey Sampling and Applied Statistics, 235-250.

11. Galambos, J. (1975). "Characterizations of Probability Distributions by Properties of Order Statistics II," Statistical Distributions in Scientific Work, edited by G. P. Patil, S. Kotz, and J. K. Ord, published by D. Reidel Publishing Co., Holland, 89-101.

12. Ghosh, M. and Meeden, G. (1976). "Nonparametric Minimax Estimation of the Mean," Calcutta Statistical Association Bulletin 25, 169-173.

13. Ghosh, M. and Meeden, G. (1977). "On the Non-attainability of Chebyshev Bounds," American Statistician 31, 35-36.

14. O'Connell, M. J. and David, H. A. (1976). "Order Statistics and their Concomitants in some Double Sampling Situations," Essays in Probability and Statistics in Honor of J. Ogawa, 451-466.

15. Sotres, D. A. and Ghosh, M. (1977). "Strong Convergence of Linear Rank Statistics for Mixing Processes," Sankhya B 39, 1-11.

16. Yang, S. S. (1977). "General Distribution Theory of the Concomitants of Order Statistics," Annals of Statistics, 5, 966-1002.

C. Papers read at Army Sponsored Conferences

1. David, H. A. (1974). "Some Uses of Order Statistics," Proceedings of the Twentieth Conference on the Design of Experiments in Army Research Development and Testing, Part 2, 833-844.

2. David, H. A. (1978). "Robust Estimation in the Presence of Outliers," (to appear in the Proceedings of the Army Research Office Workshop on Robustness in Statistics).

D. Submitted Papers

1. Conover, W. J. and Iman, R. L., "Some Exact Tables for the Squared Ranks Test," to appear in Communications in Statistics, Series B.

2. Conover, W. J. and Iman, R. L. "Approximations of the Critical Region for Spearman's RHO With and Without Ties Present," to appear in Communications in Statistics, Series B.

3. Ghosh, M. "Rate of Convergence to Normality for Random Means and U-Statistics: Applications to Sequential Estimation," submitted to the Annals of Statistics.

4. Ghosh, M. and Dasgupta, R. "Limit Laws for U-Statistics in the Non IID Case," submitted to the Australian Journal of Statistics.

5. Ghosh, M. and Mukhopadhyay, N. "Sequential Point Estimation of the Mean When the Distribution is Unspecified," submitted to the Annals of Statistics.

6. Ghosh, M. and Mukhopadhyay, N. "Sequential Point Estimation of the Difference of Two Means," submitted to Communications in Statistics.

7. Ghosh, M. and Sukhatme, S. "On Bahadur's Representation of Quantiles in Nonregular Cases," submitted to Annals of Statistics.

8. Huang, J. S., Arnold, B. C., and Ghosh, M. "On Characterizations of the Uniform Distribution Based on Identically Distributed Spacings," submitted to the Journal of Applied Probability.

9. Sotres, D. A. and Ghosh, M. "Bahadur Representation of Sample Quantiles for Non-stationary Mixing Processes," submitted to the Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete.

10. Sotres, D. A. and Ghosh, M. "Asymptotic Properties of Linear Functions of Order Statistics for Non-stationary Mixing Processes," submitted to the Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete.

6. Appendixes - None