

AD-A033 581

KENTUCKY UNIV LEXINGTON DEPT OF STATISTICS
LIFE TESTING, RELIABILITY, AND MULTIVARIATE NONPARAMETRIC METHO--ETC(U)
JUL 76 S BLUMENTHAL, V P BHAPKAR

F/G 12/1

AF-AFOSR2841-75

AFOSR-TR-76-1299

NL

UNCLASSIFIED

| OF |

AD
A033581



END

DATE
FILMED

2-77

Life Testing, Reliability, and Multivariate Nonparametric Methods
July 27, 1976

3
B.S.

ADA033581

Interim Scientific Report - Grant AFOSR 75-2841

Principal Investigator: Saul Blumenthal,
Department of Statistics
University of Kentucky

Co- Investigator: V. P. Bhapkar
Department of Statistics
University of Kentucky

DDC
RECEIVED
DEC 9 1976
A

Since the grant involves two distinct projects, this report is divided into two sections, number I dealing with the work of the principal investigator, and number II dealing with the co-investigator's research.

I. Screening and Estimation Procedures for the Unknown Number of Defective Items in a Life Test, and Estimation of the Size of a Finite Population.

A. Research Accomplished Under the Current Grant

1.) Continuation of previous research.

In addition to research which was initiated under the grant, grant-sponsored research time was used to make extensive revisions of several papers which had been submitted for publication. The revised versions of the papers acknowledge the grant for partial support of the research.

These papers are:

Approved for public release;
distribution unlimited.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

**AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC**

This technical report has been reviewed and is
approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.

**A. D. BLOSE
Technical Information Officer**

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM	
1. REPORT NUMBER 18 AFOSR-TR-76-1299 ✓	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) 6 LIFE TESTING, RELIABILITY, AND MULTIVARIATE NONPARAMETRIC METHODS.	5. TYPE OF REPORT & PERIOD COVERED 7 Interim scientific rept.		
7. AUTHOR(s) 19 Saul/Blumenthal • V. P./Bhaskar	8. CONTRACT OR GRANT NUMBER(s) 15 ✓ AF-AFOSR 75-2841-75 ✓		
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Kentucky Department of Statistics ✓ Lexington, Kentucky 40506	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 79769-05 12 45		
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332	12. REPORT DATE 11 27 July 1976		
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 12p.	13. NUMBER OF PAGES 11		
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) Describes progress under the grant from 1 June 1975 to 30 June 1976.			

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 408091-
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

2B

ACCESSION NO. NTIS DOC UNCLASSIFIED BY 68-00000-1

- a) "Sequential Estimation of the Largest Normal Mean When the Variance is Known", to appear in Annals of Statistics, November, 1976.
- b) "A Comparison of the Bad as Old and Superimposed Renewal Models", (joint with J. A. Greenwood and L. H. Herbach), to appear in Management Science.

- c) Robustness of Stein's Two-Stage Procedure for Mixtures of Normal Populations, (joint with Z. Govindarajulu), to appear in Journal of the American Statistical Association.
- d) Estimating the Complete Sample Size from an Incomplete Poisson Sample, (joint with Ram C. Dahiya and Alan J. Gross), submitted to Journal of the American Statistical Association. This work is directly related to the mission of this grant.

a.) Completed. One technical report was started and completed under sponsorship of the grant. This was: "Estimating Population Size with Truncated Sampling" Tech. Report No. 96, Department of Statistics, University of Kentucky. Submitted to Communications in Statistics. An abstract follows:

July 27, 1976

Page 3

Let $X_1, X_2, \dots, X_n$ be i.i.d. r.v.'s with density $f(x, \theta)$ w.r.t. a σ -finite measure μ . Let R be a measurable set in the sample space χ . The value of X is observable if $X \in (\chi - R)$ and not otherwise. The number J of observable X 's is binomial, $N, Q, Q = 1 - P(X \in R)$. On the basis of J observations, it is desired to estimate N and θ . Estimators considered are conditional and unconditional maximum likelihood (c.m.l.e. and u.m.l.e.) and modified maximum likelihood using a prior weight function to modify the likelihood before maximizing. Asymptotic expansions are developed for the $\hat{N}$'s of the form $\hat{N} = N + \alpha\sqrt{N} + \beta + o_p(1)$, where α and β are random variables. All estimators have the same α , which has mean 0, variance σ^2 (a function of θ) and is asymptotically normal. Hence, all are asymptotically equivalent by the usual limit distributional theory. The β 's differ and $E\beta$ can be considered an "asymptotic bias". Formulas are developed to compare the asymptotic biases of the various estimators. For a scale parameter family of absolutely continuous distributions with $\chi = (0, \infty)$, and $R = (T, \infty)$, special formulae are developed and a best estimator is found.

b.) Ongoing. Several projects are currently in progress under grant sponsorship.

1. The major project now under way is the preparation of a monograph to be entitled "Statistical Methods for Truncated and Other Incomplete Data." This is joint with L. Sanathanan. This book will cover the results of the two co-authors on the estimation of sample size from truncated observations (which is

July 27, 1976

Page 4

the subject of this grant) and on estimation of parameters from these same observations. Comprehensive coverage of other research in the literature on inference procedures for truncated data will also be given with the emphasis being put on making this material accessible to the practitioner, and on reliability applications. In addition to estimation from truncated data, techniques for censored data will be discussed. Handling of missing values and of partially classified data in the multivariate normal and multinomial distributions is also to be included.

2. In some truncated sampling situations, such as measuring items from a production line which have been screened by go-nogo gauging, it can be assumed that the number of observations n is fixed. The number of items N needed to give the n values observed is a negative binomial random variable. Good estimators of the parameters are needed for this type of inverse binomial sampling, and general expressions are being developed to allow the selection of best modified maximum likelihood estimators. Work is joint with L. Sanathanan.
3. There are many estimators available for censored data, and these can be adapted to truncated data by using an iterative procedure with estimated sample size. Convergence and asymptotic variances of these procedures are under examination, jointly with L. Sanathanan.

July 27, 1976

Page 5

4. In a previous paper, the author gave a sequential screening procedure for the case in which $F(t)$ was known completely, and results were stated in terms of the exponential distribution. The procedure could be used if the distribution contained an unknown scale parameter, provided that a bound was given on the value of the parameter. However, it would be an inefficient procedure in that case unless the parameter value was actually near the bound. We have been investigating a "staged" sequential procedure which takes observations for a fixed period of time T_0 , then computes an estimate $\hat{\theta}$ of the unknown scale parameter θ and thereafter uses the same sequential stopping rule as in the earlier paper except that $\hat{\theta}$ is used in place of θ . The goal is to guarantee with probability P^* that no more than K (P^*, K given) defectives remain after sampling terminates. Not only must we choose the constants for the stopping rule as in the earlier work, but T_0 must also be chosen. Some analytic results have been obtained, but the problem is not readily amenable to analysis. Currently Monte Carlo studies are underway to determine the probabilities of various numbers K remaining for different choices of T_0 , of the stopping constants, of θ , N , and different estimators of θ . From these, we hope to be able to recommend approximately optimal rules to assure the specified goals. We expect this to be a major

July 27, 1976

Page 6

advance in making these screening procedures accessible for use in realistic situations. It should be noted that J. M. Finkelstein and R. E. Schafer of Hughes Aircraft Co., Fullerton, California, in an unpublished manuscript entitled, "A Sequential Screening Procedure for the Exponential Distribution", have made extensive Monte Carlo studies of the properties of an ad-hoc sequential screening procedure designed to eliminate the effect of an unknown scale parameter. From this study, it can be presumed that good procedures of this type are of potential value in the aircraft industry.

5. For fixed amount of sampling when $F(t)$ is exponential with unknown failure rate λ , the limiting variance of the M.L.E. is an unbounded function of λ . Basically the difficulty lies not with some defect in the maximum likelihood process but with the difficulty of estimating N on the basis of insufficient information. For instance, for fixed time sampling, the estimator behaves quite well if $T > t_0 \lambda$ (t_0 any constant). But if λ is unknown, no fixed T can be assured to be sufficiently large, and what would appear to be needed is either a two-sample or sequential rule which generates an estimate $\hat{\lambda}$ and assures that the ultimate test duration exceeds $K \hat{\lambda}$, for an appropriate K . Such a rule should give good properties to any estimator, including the maximum likelihood one. In particular such a rule should lead to bounds

July 27, 1976

Page 7

on the variance of the estimator. A two stage rule with a fixed duration T_0 in the first stage and a second stage duration proportional to $\hat{\lambda}$ where $\hat{\lambda}$ is based on the first stage observations is now being studied by a graduate student under the supervision of the present investigator.

6. Previous research has concentrated on finding a modified maximum likelihood estimator which minimizes the asymptotic second order bias. Additional terms in our series expansions are being developed to allow selection of an estimator which minimizes asymptotic second order mean squared error. Research is being conducted by a graduate student under the supervision of the principal investigator.
7. A problem of interest in reliability and related to this grant's purposes is the estimation of the parameter of the exponential distribution when the parameter changes with time. In joint research with A. Gross, this problem is under exploration.

B. Consultants. During the past year, the following visitors came to Kentucky under Grant sponsorship to discuss mutual research interests. All of these consultants proved helpful in furthering the research efforts of the investigators. Two of the visitor have become engaged in joint research projects with the principal investigator as a direct result of the invitation (see section A-2-b-1, 2, 3, and 7).

July 27, 1976

Page 8

- 1.) Professor Ram C. Dahiya, Department of Mathematics and Statistics, University of Massachusetts.
- 2.) Professor Alan J. Gross, Department of Biometry, Medical University of South Carolina.
- 3.) Professor Lars Holst, Mathematics Research Center, Madison, Wisconsin, and Uppsala University, Sweden.
- 4.) Professor Arthur Roth, Statistics Department, Carnegie-Mellon University.
- 5.) Professor Lalitha Sanathanan, Department of Quantitative Methods, College of Business Administration, University of Illinois at Chicago Circle.

II. Multivariate Nonparametric Methods for Several Samples

Asymptotically distribution-free tests for comparing several samples in the multivariate case were offered in the statistical literature during the last ten years in view of the need to discard the stringent assumption of multivariate normality. However, in view of their computational complexity, these theoretical advances had not yet filtered down to the application stage.

During the first phase of this research suitable computer programs were developed for the computation of test-criteria for comparing three samples with two or three correlated variables. Simulations were then carried out using random samples generated from three diverse types of populations in order to study the performance of these criteria which are theoretically valid only for large samples.

In the second phase of this research, asymptotically distribution-free test-criteria ^{ve} have been developed for the "profile analysis" of several multivariate samples. Suppose that we have independent random samples of size n_i from the i -th p -variate population with c.d.f. F_i , $i=1,2,\dots,k$. The earlier nonparametric tests have dealt with the overall null-hypothesis

$$H_0: F_1 = F_2 = \dots = F_k,$$

i.e. there are no differences among k populations. However, if there are some differences, one is often interested in finding out whether the pattern of differences among populations is, in some sense, uniform across variables. In other words, one wants to know whether there is any

Page 10

interaction between populations and variables. Under the assumption of multivariate normality with common nonsingular covariance matrix, the hypothesis of no interaction is then formulated as

$$H_1: \mu_i^{(j)} - \mu_i^{(1)} = \mu_1^{(j)} - \mu_1^{(1)},$$

$$i=2, \dots, k, \quad j=2, \dots, p;$$

here $\mu_i^{(j)}$ is the mean of the j -th variable for the i -th population. If we draw the "profile" of the i -th population by plotting its means against the corresponding variable, the hypothesis H_1 can be interpreted as parallelism of profiles (See Figure 1.)

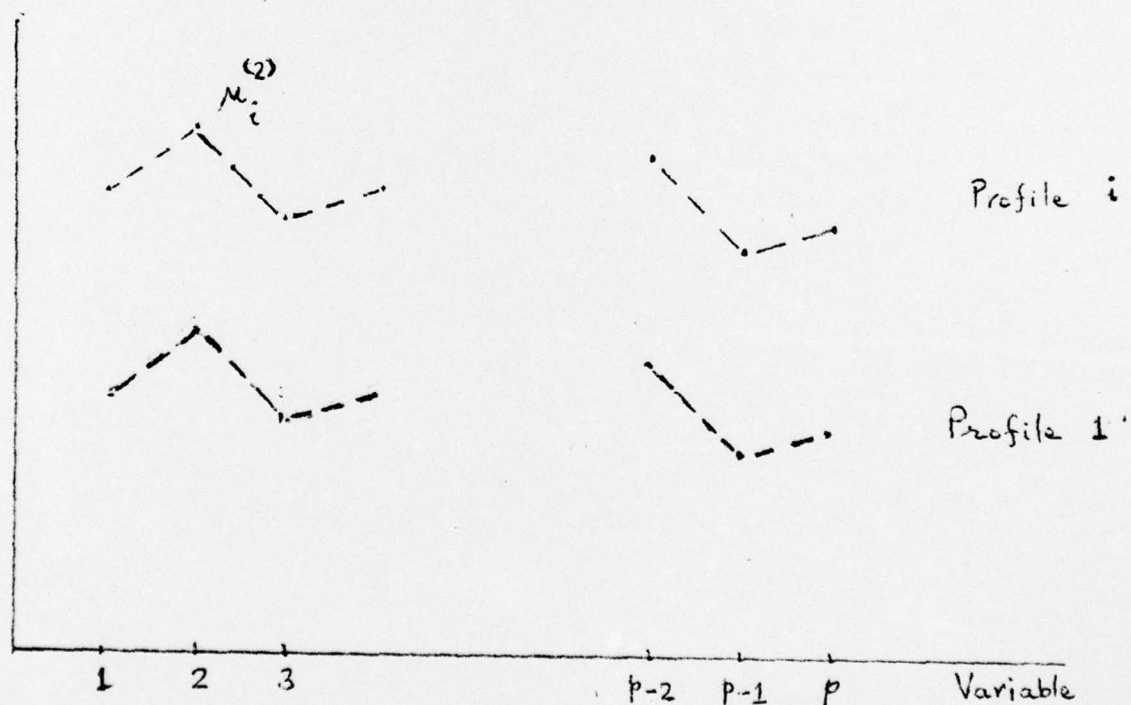


Figure 1

It is then seen that such an hypothesis of parallelism is meaningful (and of interest) only when the variables are commensurable, i.e. are comparable in some strong sense (e.g. repeated measurements at different time points, or test scores in different correlated categories on some common scale).

Discarding the assumption of normality and to a considerable extent that of commensurability, we have developed nonparametric criteria for such a profile analysis. First, a suitable nonparametric analog of H_1 , say H_1^* , has been formulated as the hypothesis of no population-variable interaction. The nonparametric criteria are then developed to test H_1^* by appropriate modifications of the criteria to test the overall H_0 . The mathematical and other details concerning their theoretical properties have been developed and are given in the following technical report:

V.P. Bhapkar and K.W. Patterson, "On Some Nonparametric Tests for Profile Analysis of Several Multivariate Samples," University of Kentucky, Department of Statistics, Technical Report No. 100, March 1976.

Finally, simulation studies similar to these in the first phase were carried out in order to study the performance of these new criteria. The findings of both these simulation studies are reported in the following:

V.P. Bhapkar and K.W. Patterson, "A Monte Carlo Study of Some Multivariate Nonparametric Statistics for Profile Analysis of Several Samples," University of Kentucky, Department of Statistics Technical Report No. 101, March 1976.