AD NUMBER

AD820242

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; May 1967. Other requests shall be referred to U.S. Army Research Office, Durham, Durham, NC.

AUTHORITY

USARO ltr, 15 Mar 1971

ARO-D Report 67-2

PROCEEDINGS OF THE TWELFTH CONFERENCE ON THE DESIGN OF EXPERIMENTS IN ARMY RESEARCH DEVELOPMENT AND TESTING

This document contains blank pages that were not filmed.

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of the U. S. Army Research Office—Durham, Durham, North Carolina.

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorised documents.

Sponsored by

The Army Mathematics Steering Committee

on Behalf of

[illegible] RCH AND DEVELOPMENT

U. S. Army Research Office - Durham
Report No. 67-2
May 1967

PROCEEDINGS OF THE TWELFTH CONFERENCE ON THE DESIGN OF EXPERIMENTS IN ARMY RESEARCH, DEVELOPMENT AND TESTING

Sponsored by the Army Mathematics Steering Committee

Host

Harry Diamond Laboratories
and
National Bureau of Standards

19-21 October 1966

U. S. Army Research Office - Durham
Box CM, Duke Station
Durham, North Carolina

FOREWORD

At the Eleventh Conference on the Design of Experiments Drs. Joseph Cameron and Walter Foster discussed the possibility of holding the next meeting at the National Bureau of Standards. Talks with Dr. Badrig Kurkjian brought out the fact that he would be willing to investigate the possibility of the Harry Diamond Laboratories serving in the role of the second host. The efforts of these three individuals brought about the desired results. The Army Mathematics Steering Committee, the sponsor of these meetings on behalf of the Office of Chief of Research and Development, was pleased to hear from Dr. Allen V. Astin, Director of the National Bureau of Standards, and Lt. Colonel M. S. Hochmuth, Commanding Officer of Harry Diamond Laboratories, that their organizations would serve as joint hosts for the Twelfth Conference. Both Messrs. Astin and Hochmuth graciously agreed to give welcoming addresses at the start of the conference. Their talks set the stage for this interesting scientific meeting. Incidentally, the Harry Diamond Laboratories and the National Bureau of Standards served as joint hosts for the first three conferences of this series. At those meetings, as well as this one, Mr. John Wheeler, Chairman on Local Arrangements, is well remembered by those in attendance for his excellent execution of the many details which must be handled for smooth running symposia.

The conference was conducted at the new quarters of the National Bureau of Standards at Gaithersburg, Maryland. This afforded the attendees an opportunity to become acquainted with these new facilities, and some of the many scientific experiments being conducted by the staff of the Bureau. They also learned of some of the types of data which NBS could furnish that would be helpful in the conduction of their own research. For the benefit of those who did not get to this meeting, we mention here some of the special equipment now on the Gaithersburg campus. There are three 35-foot grating spectrographs. One operates in the vacuum ultraviolet region, another in the visible region, while the third is used for the short wave ultraviolet region. The NBS LINAC is a 100 Mec linear electron accelerator capable of producing one of the world's most intense high-energy electron beams. Neutron irradiation experiments can be conducted with the new 10-megawatt nuclear research facility. The world's largest testing machine, a 12-million pound compression and tension tester, is about ready for use. This monster rises almost 100 feet above its base. These and many other new scientific machines are to be found at these well-equipped laboratories.

The program of the Twelfth Conference featured the following four invited addresses:

Operations Research
Professor Brian W. Conolly, Saclant ASW Research Centre

Statistics as a Diagnostic Tool in Data Analysis
Dr. John Mandel, National Bureau of Standards

Planning and Analysis of Observational Studies
Professor W. G. Cochran, Harvard University

Sample Censoring
Professor Norman L. Johnson, University of North Carolina at Chapel Hill

Besides these talks, the members of the audience were able to select from 24 contributed scientific papers topics that best suited their own needs. These papers were presented in eight technical and two clinical sessions. We are pleased to say that Dr. Frederick F. Stephan, Presient of the American Statistical Association, was able to attend the banquet. He was called on to present the second Wilks Memorial Medal to General Leslie E. Simon.

This volume of the Proceedings contains 24 of the papers which were presented at this meeting. The Army Mathematics Steering Committee has asked that these articles on modern principles on the design of experiments, together with the application of these ideas, be made available in the form of this technical manual. Members of this committee take this opportunity to express their thanks to the many speakers and other research workers who participated in the conference.

The conference had an attendance of 125 scientists; and 50 organizations were represented. Speakers and panelists came from George Washington University, Harvard University, the National Bureau of Standards, the National Institutes of Health, North Atlantic Treaty Organization, North Carolina State University at Raleigh, Phillips Petroleum Company, Stanford University, University of California at Los Angeles, University of Chicago, University of Georgia, University of Michigan, University of North Carolina at Chapel Hill, University of Wisconsin, Virginia Polytechnic Institute and thirteen Army facilities.

The Chairman wishes to express his appreciation to his Advisory Committee (Joseph Cameron, F. G. Dressel, Walter D. Foster, Bernard Greenberg, Boyd Harshbarger, J. S. Hunter, H. L. Lucas, Jr., Clifford Maloney and Henry B. Mann) for their assistance in formulating the program and selecting the invited speakers.

Frank E. Grubbs
Conference Chairman

TABLE OF CONTENTS

*This paper was presented at the conference. It does not appear in the Proceedings.

*This paper was presented at the conference. It does not appear in these Proceedings.

*This paper was presented at the conference. It does not appear in these Proceedings.

TWELFTH CONFERENCE ON THE DESIGN OF EXPERIMENTS
IN ARMY RESEARCH, DEVELOPMENT AND TESTING

19-21 October 1966

Wednesday, 19 October

0900-0930 REGISTRATION - Administration Building, Main Foyer

0930-1000 OPENING OF THE CONFERENCE - Admin. Bldg., Green Auditorium

John Wheeler, Chairman on Local Arrangements,
Harry Diamond Laboratories, Washington, D. C.

WELCOME

Dr. Allen V. Astin, Director
National Bureau of Standards

Colonel M. S. Hochmuth, Commanding Officer
Harry Diamond Laboratories

1000-1230 GENERAL SESSION 1, Green Auditorium

Chairman: Professor Boyd Harshbarger, Department of Statistics, Virginia Polytechnic Institute, Blacksburg, Va.

OPERATIONS RESEARCH
Professor Brian W. Conolly, North Atlantic Treaty Organization, Saclant ASW Research Centre

STATISTICS AS A DIAGNOSTIC TOOL IN DATA ANALYSIS
Dr. John Mandel, Materials Evaluation Laboratory, National Bureau of Standards, Gaithersburg, Maryland

1230-1330 LUNCH

Technical Sessions I and II will start at 1330 and run to 1500. After a break Technical Sessions III and IV will convene at 1530 and end at 1700. The social hour will begin at 1730. The banquet is scheduled for 1830.

1330-1500 TECHNICAL SESSION I - Lecture Room A

Chairman: Bert Levy, Harry Diamond Laboratories, Washington, D. C.

STEPWISE MULTIPLE REGRESSION STATISTICAL THEORY AND COMPUTER PROGRAM DESCRIPTION
Harold J. Breaux, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

ESTIMATION OF ERROR RATES IN DISCRIMINANT ANALYSIS
Peter A. Lachenbruch and M. Ray Mickey, University of North Carolina at Chapel Hill, North Carolina and University of California, Los Angeles, California Representing the Army Research Office-Durham

1330-1500 TECHNICAL SESSION II - Lecture Room B

Chairman: Henry A. Dihm, U. S. Army Missile Command, Redstone Arsenal, Alabama

SOME STATISTICAL APPLICATIONS IN THE TESTING OF MILITARY RUBBER PRODUCTS
Emil H. Jebe, Willow Run Laboratories, Institute of Science and Technology, The University of Michigan. Representing the U. S. Army Tank-Automotive Center, Warren, Michigan

A STATISTICAL ANALYSIS OF PROVISIONING PROCESSES ON FOUR ARMY MISSILE SYSTEMS
Robert G. Provost, U. S. Army Missile Command, Redstone Arsenal, Alabama

1500-1530 BREAK

1530-1700 TECHNICAL SESSION III - Lecture Room A

Chairman: Henry Ellner, Directorate for Quality Assurance, U. S. Army Edgewood Arsenal, Edgewood Arsenal, Md.

OPTIMAL ECONOMY IN PLANNING EXPERIMENTS
Regina C. Elandt-Johnson, University of North Carolina at Chapel Hill, North Carolina. Representing the Army Research Office-Durham.

TECHNICAL SESSION III (continued)

ON A CLASS OF NONPARAMETRIC TESTS FOR MANOVA IN TWOWAY LAYOUTS
Pranab Kumar Sen, University of North Carolina at Chapel Hill, North Carolina and the University of Calcutta. Representing the Army Research Office-Durham

1530-1700 TECHNICAL SESSION IV - Lecture Room B

Chairman: David Hogben, Statistical Engineering Laboratory, National Bureau of Standards, Gaithersburg, Maryland

TESTS FOR OUTLIERS
H. A. David, University of North Carolina at Chapel Hill, North Carolina. Representing the Army Research Office-Durham

THE PROBABILITY OF SURVIVAL OF A SUBTERRANEAN TARGET UNDER INTENSIVE ATTACK
Bernard Harris, Herman F. Karreman, and J. Barkley Rosser, Mathematics Research Center, U. S. Army, University of Wisconsin, Madison, Wisconsin

1730-1830 SOCIAL HOUR - Country Squire Room, Washingtonian Country Club, Gaithersburg, Maryland

1830- BANQUET - (As above)

Presentation of the Samuel S. Wilks Memorial Award

Thursday, 20 October

Technical Sessions V and VI will run from 0900-1020. After the break Technical Session VII and Clinical Session A will start at 1050 and run to 1230. After lunch Technical Session VIII and Clinical Session B will convene at 1330 and end at 1520. After a half hour break General Session 2 is scheduled for 1550 to 1700.

0900-1020 TECHNICAL SESSION V - Lecture Room A

Chairman: Selig Starr, Mathematics Branch, Office of the Chief of Research and Development, Washington, D. C.

TECHNICAL SESSION V (continued)

SINGLE DEGREE OF FREEDOM ORTHOGONAL COMPONENTS OF A FACTOR AT 2^k LEVELS IN TERMS OF LINEAR COMBINATIONS OF THE 2K CONTRASTS OF K FACTORS AT 2 LEVELS

Joseph Weinstein, Electronic Components Laboratory, U. S. Army Electronics Command, Fort Monmouth, N. J.

CONDITIONAL EFFECTS AND INTERACTION IN SYMMETRICAL FACTORIAL CONFOUNDING WITH APPLICATIONS TO BIOLOGY

N. R. Bohidar, Biomathematics Division, Fort Detrick, Frederick, Maryland

0900-1020 TECHNICAL SESSION VI - Lecture Boom B

Chairman: James B. Duff, U. S. Army Engineering Research and Development Laboratory, Fort Belvoir, Virginia

THE NEGATIVE BINOMIAL DISTRIBUTION APPLIED TO ATMOSPHERIC PARAMETERS

Oskar M. Essenwanger, U. S. Army Missile Command, Redstone Arsenal, Alabama

TRIAL VARIABILITY INTERPRETED AS DIFFERENCES IN TRANSLATION OR ROTATION IN FUNCTION ANALYSIS OF VARIANCE

Walter D. Foster, U. S. Army Biological Laboratories, Fort Detrick, Frederick, Maryland

1020-1050 BREAK

1050-1230 TECHNICAL SESSION VII - Lecture Room A

Chairman: A. Bulfinch, U. S. Army Munitions Command, Picatinny Arsenal, Dover, New Jersey

SIMULATION OF BIO-CELLULAR ANIMAL SYSTEMS

George I. Lavin, Terminal Ballistic Laboratory, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

TECHNICAL SESSION VII (continued)

A METHOD FOR ADJUSTING FOR PARTICLE SIZE IN THE X-RAY FLOURESCENCE ANALYSIS OF A MULTICOMPONENT MIXTURE
R. H. Myers, Virginia Polytechnic Institute, Blacksburg, Virginia, and Donald E. Womeldorph, Phillips Petroleum Company. Representing the Army Research Office-Durham

1050-1230 CLINICAL SESSION A - Lecture Room B

Chairman: Fred Frishman, Mathematics Branch, Office, Chief of Research and Development, Washington, D. C.

Panelists:
Mr. O. P. Bruno, Surveillance & Reliability Laboratory, U. S. Army Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

Professor A. C. Cohen, Jr., Institute of Statistics, University of Georgia, Athens, Georgia

Professor Boyd Harshbarger, Statistical Laboratory, Virginia Polytechnic Institute, Blacksburg, Virginia

Dr. Joan R. Rosenblatt, Statistical Engineering Laboratory, National Bureau of Standards, Gaithersburg, Maryland

Professor Herbert Solomon, George Washington University Washington, D. C. and Stanford University, Stanford, California

THE PROBLEM OF DETERMINING THE CONFIDENCE LEVEL FOR SOME TIME INDEPENDENT SYSTEM RELIABILITY ESTIMATES WHEN ATTRIBUTE DATA FOR THE SYSTEM SUB-COMPONENTS ARE GIVEN (A PROPOSED SOLUTION AND APPROXIMATION TECHNIQUE)
Eugene F. Dutoit, Picatinny Arsenal, Dover, New Jersey

STATISTICS, PROBABILITY, AND DETERMINISM IN A RELIABILITY IMPROVEMENT PROGRAM
Woodie R. Jenkins, Jr., National Range Operations, White Sands Missile Range, New Mexico

1230-1330 LUNCH

1330-1520 TECHNICAL SESSION VIII - Lecture Room A

Chairman: Cyrus Martin, U. S. Army Engineering Research and Development Laboratory, Fort Belvoir, Virginia

A COMPUTERIZED PROCEDURE FOR WRITING MATHEMATICAL MODELS FOR SYSTEMS RELIABILITY
Anthony J. Ricciardi and John G. Mardo, Nuclear Reliability Division, Picatinny Arsenal, Dover, N. J.

1330-1520 CLINICAL SESSION B - Lecture Room B

Chairman: Albert Parks, Harry Diamond Laboratories, Washington, D. C.

Panelists:

Professor Bernard Greenberg, Department of Biostatistics, University of North Carolina, Chapel Hill, N. C.

Dr. Frank E. Grubbs, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

Professor William Kruskal, Department of Statistics, The University of Chicago, Chicago, Illinois

Professor H. L. Lucas, Jr., Institute of Statistics, North Carolina State College, Raleigh, North Carolina

Dr. Henry B. Mann, Mathematics Research Center, U. S. Army, University of Wisconsin, Madison, Wis.

BIOLOGICAL APPLICATION OF GRUBB'S TECHNIQUE
Clifford J. Maloney, National Institutes of Health, Bethesda, Maryland

A BEST FIT PROBLEM
R. T. Thrall, Project Michigan, University of Michigan, Ann Arbor, Michigan

1520-1550 BREAK

1550-1700 GENERAL SESSION 2 - Green Auditorium

Chairman: Dr. Walter D. Foster, Biomathematics Division, U. S. Army Biological Laboratories, Fort Detrick, Frederick, Maryland

PLANNING AND ANALYSIS OF OBSERVATIONAL STUDIES
Professor W. G. Cochran, Department of Statistics, Harvard University, Cambridge, Massachusetts

Friday, 21 October

Technical Sessions IX and X will run from 0900-1030. General Session 3 will start at 1100 and end at 1215.

0900-1030 TECHNICAL SESSION IX - Lecture Room A

Chairman: Robert Eissner, U. S. Army Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

A MODERATELY DISTRIBUTION FREE APPROACH TO RELIABILITY ESTIMATION BASED ON THE FIRST ORDER STATISTIC
Michael G. Billings, Dugway Field Office, C-E-I-R, Inc. Representing the U. S. Army Chemical Corps, Dugway Proving Ground, Dugway, Utah

ON THE RELIABILITY OF COMPLEX SYSTEMS
A. C. Cohen, Jr., Department of Statistics, University of Georgia, Athens, Georgia

0900-1030 TECHNICAL SESSION X - Lecture Room B

Chairman: Erwin Biser, Research Analyst, U. S. Army Electronics Command, Fort Monmouth, New Jersey

THE APPLICATION OF SOME NONLINEAR LEAST SQUARES METHODS IN THE ESTIMATION OF A PREDICTION EQUATION
Weldon F. Willoughby, U. S. Army Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

TECHNICAL SESSION X (continued)

OBSERVATIONS ON THE SELECTION OF PREDICTORS
H. L. Lucas, Jr., Institute of Statistics, North Carolina State University at Raleigh, North Carolina

1030-1100 BREAK

1100-1215 GENERAL SESSION 3 - Green Auditorium

Chairman: Dr. Frank E. Grubbs, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

SAMPLE CENSORING
Professor Norman L. Johnson, Department of Statistics, University of North Carolina at Chapel Hill, North Carolina

OPERATIONS RESEARCH

Professor Brian W. Conolly*
North Atlantic Treaty Organization
Saclant ASW Research Center

INTRODUCTION.

1. It is a privilege and a pleasure to be invited to make a presentation on Operational Research at a specialist statistical conference. Those individuals who choose to make Operational Research their profession come from the ranks of engineers, physicists, chemists, mathematicians as well as statisticians. All have a contribution to make to Operational Research. I myself, for example, am a mathematician by basic training, with a pronounced interest in obtaining practical and verifiable solutions to real life problems.

2. The name Operational Research is itself perhaps not a very good description of the type of activity that O. R. workers usually undertake. I do not propose to be so controversial as to suggest an alternative. My theme is rather to suggest that, as it has developed, modern O. R. has come to depend more and more heavily on the science and techniques of statistics and probability theory. And it is not difficult to see why this is so.

3. In O. R. we are usually concerned with studying the workings of a complex system or process such as the manufacture of an automobile; the organization of an airport; the routing of city traffic; a telephone exchange; the detection, classification and destruction of an enemy target. If we like to call these systems or processes "operations", and the study we make of them "research", then we arrive at the name Operations Research by which O.R. is designated in the U. S. The fact that O. R. is called "Operational Research" in Europe is presumably by analogy with our practice to call research in physics physical research, and research in mathematics mathematical research.

4. The objective of O.R. studies is normally to discover how to optimize in some sense the output of the process: e.g. produce an adequate automobile at a minimal cost; achieve an airport organization which maximizes passenger flow with a minimum of incovenience and the best employment of facilities; maximize the probability of destruction of the enemy target. In order to do this we have to try to understand the structure of the process.

*Now at Virginia Polytechnic Institute, Blacksburg, Virginia

5. The complex processes which O.R. investigates are normally decomposable into a number of subsidiary processes on each of which the ultimate output depends. If one regards these as parameters of the system as a whole, then the study consists first in determining their inter-relation, and the way they affect the output. This leads to the creation of a more or less mathematical model - a set of equations which characterises the process. If the model is verified in the sense that it can be used to predict measurable outputs, then the analysis of the process and its optimization reduces to the application of appropriate mathematical techniques to the model.

6. One reason why modern O.R. has come to be heavily dependent on probability and statistics is the greater recognition of the need to assess the effect of chance on the outcome of a process; rather than to work throughout with average values. Nowadays we are interested in the probability distributions of the outcomes of the subsidiary processes in order to discover the probability distribution of the overall outcome. Under these circumstances we have to deal with stochastic processes and our analysis depends on the specialized techniques developed by the experts.

7. I think that in fact O.R. and statistics have much to offer each other. Erlang was a Danish engineer and an O.R. worker whose interest was the Danish telephone service. His work in the early 20th century founded queueing theory which in all its increasing complexity is the subject of many research papers published in both statistical and O.R. research periodicals. Those who are concerned with military exercies know that one has to deal with experiments whose design cannot be altogether controlled, that the samples are small, and the variables many: a situation shocking to a classical statistician, but a challenge.

8. During the remainder of this presentation I intend to be more specific. In order to illustrate my thesis of the statistical interest which is to be found in O.R. studies and the dependence of the analysis on statistical expertise I shall describe two problems from a military O.R. context, which I hope you will find entertaining.

PROBLEM 1.

9. During anti-submarine operations there inevitably occur events which have a nuisance value, and which one would like to eliminate. The elimination is partly a matter of equipment design, and partly of training in its use.

10. For the purposes of this presentation I am concerned only with finding a simple stochastic process which describes the occurrence of the

events in time in the hope that such a description may throw light on the basic phenomenon. I have no wartime records of the events, though I know they have always occurred. I am therefore dependent on naval exercises for data.

11. Suppose, then, that I have obtained from the records of one ship during a recent exercise the times t_n of occurrence of the events which I shall denote by $E_n (n \geq 1)$. I measure time from the beginning of the exercise. An immediate difficulty arises out of the fortunate fact that the E_n do not occur at a tremendously high rate. Three per day might be a typical average taken over all ships. Exercises of the right sort do not take place frequently and, when they do, they are of a limited duration. Thus, typically, at the end of a week I might have a few tens of events for each ship. At the beginnig I want to consider each ship's records separately, so my sample is not very great.

12. Adopting the good practice of making a simple initial hypothesis I look at the time series (t_n) for each ship and ask if there are indications that the events (which a priori might be thought of as having random origin) occur in a Poisson stream. The answer is that they do not appear to do so, but rather that in all cases there is evidence of cluster (a preponderance of short inter-event time intervals as compared with a Poisson stream with the same mean). Moreover the mean intervals of the event distributions seem to be quite different from each other, and I do not find evidence which supports the hypothesis that the E_n for ship A could be generated by adjustment of the mean from the stream of E_n for ship B.

13. Since there appears to be clustering I next ask myself if a particular stream E_n could have been generated by a contagious process, and for this purpose I choose a Pólya process defined in the following form: Suppose that the process begins at time t=0 and that no event E occurs at that time. The instantaneous probability that the $(n+1)^{th}$ Pólya event takes place in the small interval $(t, t+dt)$ is given by $\beta_{n+1}(t)dt$ where

$$\beta_{n+1}(t) = \frac{\lambda(1+an)}{(1+a\lambda t)} . \tag{13.1}$$

The parameters λ and $\underline{a}$ are supposed to be real, λ is positive and $\underline{a}$ non-negative. There is apparently no other restriction on $\underline{a}$, though I shall make some more remarks on that subject later. When $\underline{a}$ is zero the Pólya process clearly becomes a Poisson process with mean interval length $1/\lambda$ time units.

14. I now give without proof a few key theoretical results of a Pólya process as defined. Let $P_n(T)$ be the probability that exactly n Pólya events

occur during the interval $(0, T)$ and define $p_o(0)=1$. The generating function $P(x, T)$ of $p_n(T)$ is then

(14.1) $$P(x, T) = [\ 1+a\lambda T(1-x)\ ]^{-1/a} \ .$$

The mean and variance of n are

(14.2) $$E(n) = \lambda T,$$

(14.3) $$Var(n) = a\lambda^2 T^2 + \lambda T \ .$$

The exercise data give me a series of time intervals between events. The likelihood of a series of n events occurring at the instants $t_1, t_2, \ldots, t_n$ is

(14.4) $$P_n(t_1, t_2, \ldots t_n) = \frac{\prod_1^{n-1} (1+ra)}{(1+a\lambda t_n)^{n+1/a}} \ .$$

The maximum likelihood estimator of λ is simply n/t_n, but that of <u>a</u> is more complicated.

15. The distribution of the intervals between events is of particular interest. It turns out that, for any n, the p.d.f. $g_n(\tau)$ of the interval between the n^{th} and the $(n+1)^{th}$ events, <u>whenever the first n events took place</u>, is

(15.1) $$g_n(\tau) = \frac{\lambda}{(1+a\lambda\tau)^{1+1/a}}$$

and is independent of n. The interpretation of this is that if we generate a lot of Pólya processes, each having the same parameters <u>a</u> and λ, and then examine the time intervals between, say, the second and third events in all the processes, we should find that they all are distributed according to (15.1).

16. The r^{th} moment μ_r about zero of an interval between two given events is

(16.1) $$\mu_r = \frac{r!}{\lambda^n(1-a)(1-2a)\ldots\ldots(1-ra)}$$

and this clearly exists only if none of the terms in the denominator is zero. Thus if a=1 the interval length distribution has all its moments infinite: if $a = \frac{1}{2}$, the first moment is finite, but none of the higher moments is.

17. I return to my fundamental problem. I have an observed series of events and I want to make statements about the hypothesis that they are generated according to a Pólya process. How do I estimate $\underline{a}$ and λ from the observations?

18. In order to throw light on this problem sequences of Pólya intervals were generated, each having the same $\underline{a}$ and λ. The idea was to compare estimates of $\underline{a}$ and λ obtained by various means with their known values. In fact the problem of estimation remains open, but some features of Pólya processes have been revealed which were a surprise to me.

19. The digital computer generation of the Pólya intervals was carried out as follows. Suppose that n events have been generated and that they occurred at times $t_1, t_2, \ldots t_n$. We require the probability $h_{n+1}(t_{n+1}/t_1, t_2, \ldots t_n)dt_{n+1}$ that the $(n+1)^{th}$ event occurs in the interval $(t_{n+1}, t_{n+1}+dt_{n+1})$, given that the first n occurred at times $t_1, t_2, \ldots t_n$. Clearly,

(19.1) $$h_{n+1}(t_{n+1}/t_1, \ldots t_n) = \frac{p_{n+1}(t_1, t_2, \ldots t_{n+1})}{p_n(t_1, t_2, \ldots t_n)}$$

where the p are given by (14.4). This says that the conditional probability density of the $(n+1)^{th}$ interval τ is

(19.2) $$h_{n+1}(\tau/t_1, t_2, \ldots t_n) = \frac{\lambda(1+na)(1+a\lambda t_n)^{n+1/a}}{(1+a\lambda t_n+a\lambda\tau)^{n+1+1/a}}\ .$$

The conditional probability that the $(n+1)^{th}$ interval is less than T is:

(19.3) $$H_{n+1}(T/t_1, t_2, \ldots t_n) = 1 - \left[\frac{1 + a\lambda t_n}{1+a\lambda t_n + a\lambda T}\right]^{n+1/a} .$$

To obtain the intervals one generates a sequence of random numbers r_n independently and uniformly distributed on $0 \leq r_n < 1$, and then solves

(19.4) $$r_{n+1} = \left[\frac{1+a\lambda t_n}{1+a\lambda t_n + a\lambda \tau_{n+1}}\right]^{n+1/a}$$

for τ_{n+1}.

EXPERIMENT 1.

20. I am now going to describe briefly some of the experiments which were carried out. For the first we generated 5 independent sequences of 1000 Pólya intervals, for each of which we assigned $\lambda = 4/30$, $a = \frac{1}{4}$. This was to give $1/\lambda\,(1-a)$, the mean interval length, the value 10 which corresponded with observation. The means and standard deviations of the interval lengths were as follows.

Means and Standard Deviations of 5 Independent Sequences of 1000 Pólya Intervals with the same Parameters

Sequence No.	Mean	Standard Deviation
1	2.708	3.164
2	21.804	22.964
3	6.428	6.874
4	7.349	8.304
5	9.519	10.290

This Table was the first surprise. We expected each sequence to have a mean and standard deviation reasonably close to the theoretical values of 10 and $10\sqrt{2}$ respectively. The first sequence was also "looked at" just after the 100th event and the means and standard deviation were found to be 2.450 and 2.615 respectively. Thus it appeared that the processes were settling down to a steady state quite rapidly, but a steady state which could be vastly different from one process to another, even though the parameters were the same. On the face of it, then, it appears that estimation of a and λ based on a perfectly valid sequence might well give completely different values from those obtained from another equally valid sequence.

21. The measure used of the mean interval after the n^{th} event at time t_n was

$$x = t_n/n \quad . \tag{21.1}$$

The sampling density function of x is:

$$\frac{(n\lambda)\, x^{n-1} \cdot \prod_{r=1}^{n-1} (1+ra)}{(n-1)!\, (1+na\lambda x)^{n+1/a}} \tag{21.2}$$

and the expectation and variance of x are respectively

$$\begin{aligned} E(x) &= 1/\lambda\,(1-a)\,; \\ Var(x) &= (1-a+an)\,/\,[n\lambda^2(1-a)^2(1-2a)] \quad . \end{aligned} \tag{21.3}$$

Now the mean of the five sequence means is 9.561 and the standard deviation is about 7. The variation of the sequence means is thus less surprising, but no less discomforting.

22. The theoretical reason for the stability of Pólya sequences about widely differing means seems to be that <u>the whole pattern of a sequence is on the average governed by the first interval</u>. This can be seen by considering the conditional expected value $E(\tau/t_1, t_2 .. t_n)$ of the $(n+1)^{th}$ interval τ, given the times $t_1, t_2 ... t_n$ of occurrence of the first n events. Then

$$E(\tau/t_1, t_2 ... t_n) = \frac{1+a\lambda t_n}{\lambda\{1+a(n-1)\}} = \frac{1}{\beta_n(t_n)} \quad . \tag{22.1}$$

Thus the conditional expected value of the second interval is $(1+a\lambda t_1)/\lambda$. If t_1 is greater than its expected value $1/\lambda\,(1-a)$ then

$$E(\tau/t_1) > \frac{1}{\lambda} + \frac{a}{\lambda(1-a)} = \frac{1}{\lambda(1-a)}$$

i.e. the second interval also tends to be greater than its expected value. And so on for all successive intervals.

EXPERIMENT 2.

23. The second experiment was an extension of the first. 500 independent sequences of 500 Pólya intervals were generated, all having the same parameters $a=\frac{1}{4}$, $\lambda = 4/30$. We were looking for something constant in all the sequences. Since the value of the instantaneous probability density of an event, just after the generation of the n^{th} event is

$$\beta_{n+1}(t_n) = \frac{\lambda(1+an)}{(1+a\lambda t_n)}$$

and since we were measuring the mean interval length by the estimator t_n/n we felt that the product $\beta_{n+1}(t_n) \cdot (t_n/n)$ should be constant (1) for long enough sequences. This turns out to be the case. The table [See Table 1 near the end of this article.] shows some typical values corresponding to the 500^{th} event in each sequence. The products are all very close to the theoretical value $1/\lambda$, in this case $30/4 = 7.5$. Unfortunately this constancy is not of much practical use. It does provide some feeling that the computer program is working as it should.

24. It was also decided to group all the 250 000 intervals into a histogram which is shown in Table 2. If we make the hypothesis that this represents a random sample from the event-independent distribution of Pólya intervals

$$P_r(d\tau) = \lambda d\tau/(1+a\lambda\tau)^{1+1/a}$$

the mean and standard deviations are theoretically 10 and $10\sqrt{2}$, and the observed values look close. But are they close enough on the basis of 250 000 observations? I cannot answer that question at the moment.

25. The observed frequencies in cells of one time unit long are tabulated in the column "observed", while the "expected" frequencies were calculated on the basis of the event-independent distribution. The last column gives χ^2. Overall this is enormous. There is a deficiency of

observed short intervals and an excess of long ones. There are also other oddities. A Poisson process with the same mean (9.866...) would give a frequency of about 24 000 in the (0, 1) range, so at least there is evidence of the clustering one expects in a Pólya process. I think perhaps that the sample cannot be considered random and independent, and this may be the explanation for the poor agreement. We also produced a histogram of the 500 process means and this is available if anyone is interested.

EXPERIMENT 3.

26. Our faith in the theory, of the event independent interval distribution was a little shaken by the previous experiment. The next experiment was conducted in order to restore confidence. 1000 independent Pólya process (with $\lambda = 4/30$ and $a = \frac{1}{4}$, as usual) were generated as far as the 12^{th} interval. For each process the lengths of the 4^{th} and of the 11^{th} intervals were grouped into histograms. These are shown in Tables 3 and 4. We did not instruct the computer to group cells with low frequencies, but even so there is satisfactory behaviour according to the hypothesis of the event independent distribution.

EXPERIMENT 4.

27. We have carried out various other experiments. The last which I will mention concerns the correlation between intervals in a Pólya process. Theoretically we appeared to find that the correlation between any pair of intervals is a, provided that $a < \frac{1}{2}$. For $a \geq \frac{1}{2}$ there is trouble over the convergence of the integrals for the second moments.

28. Table 5 concerns sequences of Pólya intervals for fixed $\lambda = 4/30$ and a varying from 0.1 to 0.9. For each a, 1000 sequences were generated and the Table gives the mean and standard deviation of the first and tenth intervals, the mean value of the product of these intervals ("prod"), and finally the correlation coefficient calculated from observed values.

29. Without information on the sampling distribution of the correlation coefficient it is difficult to make meaningful statements about these results. There are signs of agreement between theory and observation for a=0.1, 0.2 and 0.3 . For $a \geq 0.5$ the second moments do not exist, in theory, and a certain wildness will be observed in the results.

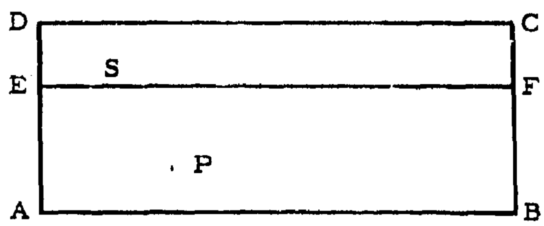
30. This concludes my description of some experiments with Pólya processes. We have subsequently formed the opinion that the Pólya process is not a good model for the natural phenomenon, but we do feel that it has been interesting to study the behaviour of the processes. I feel there is room for a good deal more statistical investigation of these processes.

For example the problems of parameter estimation and sampling distributions are still open, not to mention the interpretation of the apparent restriction on $\underline{a}$ which results from the non-existence of some of the moments for certain values. Perhaps some of you know the answers to these questions, and if I, as a representative of Operational Research, have called your attention to a typical O.R. investigation where expert statistical advice is needed I have succeeded in my objective.

PROBLEM 2.

31. My second example concerns a tactical problem. We were interested in a situation in which a tactical unit has the task of penetrating a barrier patrolled by opposing forces. For the purposes of the example, the barrier forces will be regarded as a point which moves according to the general rules along a line perpendicular to the general expected direction of penetration of the opposing forces.

32. The situation is illustrated in the next figure.



The area of interest is the rectangle ABCD. The line EF is patrolled by the barrier forces S. Its opponent P has the task of moving from some point on the boundary AB to CD. That is to say, P wants to traverse EF without being intercepted by S.

33. S, the intercepter, is provided with exact information about the whereabouts of P either

(a) continuously;
(b) at regularly spaced intervals;
(c) at random intervals having a negative exponential distribution.

P, the penetrator, is supposed to have a number of penetration strategies, for instance;

(a) a straight unvarying track from A to C;

(b) a track composed of a straight portion and one change of course at an arbitrarily selected moment before reaching EF;

(c) a random zig-zag.

Strategies are also postulated for S. It can, for example,

(a) Predict the track of P on the basis of the most recent information, and strive to reach the point of intersection of that track with EF in order to intercept P.

(b) Attempt to equate its x-coordinate with the last reported coordinate of P.

(c) Attempt to reach a point such that, whatever P does, the interception time is a minimum.

34. With three information categories, three strategies for P and three strategies for S, we have a total of 27 combinations to study. What is a suitable criterion of effectiveness? One obvious choice which will be considered here is the shortest distance between P and S during an attempted penetration. If necessary this can later be translated into probabilities of detection and kill.

35. We found in fact that the major part of this study could be carried out analytically. The combination of random information with any of the other possibilities defied analysis, however, and for these cases we resorted to a digital computer simulation. Now it is particularly important when employing digital computer simulation to invoke a check on what one is doing. What, then, would be a suitable check?

36. I would like you to consider the situation in which at time t=0 S is at E and P at A. P's strategy is pursue the diagonal track AC, while S, when it receives information as to P's position, attempts to equate its x-coordinate with the last reported x-coordinate of P. Assuming that the distances between A and E are large we then are naturally led to consider a situation which, evolving over a sufficiently long time consists of a chase of P by S.

37. In projection along an x-axis parallel to AB, P moves continuously from left to right at a speed v, say. S, when it receives information about the x-coordinate of P, tries to equate its own x-coordinate with the last reported x-coordinate of P. It moves with constant speed u. If S arrives

at the last reported x-coordinate of P before further information arrives it stops and waits. Otherwise it continues.

38. It turns out that one can obtain theoretically the statistical distribution of the distance between P and S parallel to the x-axis at an "information instant", assuming a steady state has come about. This quantity can be output very simply from the computer program and, if it conforms with theory, it gives a measure of confidence in the random mechanisms which the computer has been programmed to simulate. I would not like to consider the theoretical problem of the distribution of the difference between the x-coordinates of P and S at an information instant.

39. Let δ_m be the distance measured parallel to the x-axis between P and S at the instant Σ_m when information is transmitted to S for the m^{th} time since t=0. With obvious notation, since

$$x_p(\Sigma_m) = x_p(\Sigma_{m-1}) + v\,T_m$$

and

$$\begin{cases} x_s(\Sigma_m) = x_p(\Sigma_{m-1}) & \text{if } x_s(\Sigma_{m-1}) + u\,T_m \geq x_p(\Sigma_{m-1}) \\ x_s(\Sigma_m) = x_s(\Sigma_{m-1}) + u\,T_m & \text{if } x_s(\Sigma_{m-1}) + u\,T_m < x_p(\Sigma_{m-1}); \end{cases}$$

we have

$$\delta_m = v\,T_m \text{ if } x_s(\Sigma_{m-1}) + u\,T_m \geq x_p(\Sigma_{m-1});$$

$$\delta_m = \delta_{m-1} + (v-u)T_m \text{ if } x_s(\Sigma_{m-1}) + u\,T_m < x_p(\Sigma_{m-1})$$

or

$$\delta_m = v\,T_m \text{ if } \delta_{m-1} \leq u\,T_m;$$

$$\delta_m = \delta_{m-1} - u\,T_m + v\,T_m \text{ if } \delta_{m-1} > u\,T_m.$$

Writing $r_{m+1} = \delta_m - uT_{m+1}$, we have

(a) a straight unvarying track from A to C;

(b) a track composed of a straight portion and one change of course at an arbitrarily selected moment before reaching EF;

(c) a random zig-zag.

Strategies are also postulated for S. It can, for example,

(a) Predict the track of P on the basis of the most recent information, and strive to reach the point of intersection of that track with EF in order to intercept P.

(b) Attempt to equate its x-coordinate with the last reported coordinate of P.

(c) Attempt to reach a point such that, whatever P does, the interception time is a minimum.

34. With three information categories, three strategies for P and three strategies for S, we have a total of 27 combinations to study. What is a suitable criterion of effectiveness? One obvious choice which will be considered here is the shortest distance between P and S during an attempted penetration. If necessary this can later be translated into probabilities of detection and kill.

35. We found in fact that the major part of this study could be carried out analytically. The combination of random information with any of the other possibilities defied analysis, however, and for these cases we resorted to a digital computer simulation. Now it is particularly important when employing digital computer simulation to invoke a check on what one is doing. What, then, would be a suitable check?

36. I would like you to consider the situation in which at time t=0 S is at E and P at A. P's strategy is pursue the diagonal track AC, while S, when it receives information as to P's position, attempts to equate its x-coordinate with the last reported x-coordinate of P. Assuming that the distances between A and E are large we then are naturally led to consider a situation which, evolving over a sufficiently long time consists of a chase of P by S.

37. In projection along an x-axis parallel to AB, P moves continuously from left to right at a speed v, say. S, when it receives information about the x-coordinate of P, tries to equate its own x-coordinate with the last reported x-coordinate of P. It moves with constant speed u. If S arrives

at the last reported x-coordinate of P before further information arrives it stops and waits. Otherwise it continues.

38. It turns out that one can obtain theoretically the statistical distribution of the distance between P and S parallel to the x-axis at an "information instant", assuming a steady state has come about. This quantity can be output very simply from the computer program and, if it conforms with theory, it gives a measure of confidence in the random mechanisms which the computer has been programmed to simulate. I would not like to consider the theoretical problem of the distribution of the difference between the x-coordinates of P and S at an information instant.

39. Let δ_m be the distance measured parallel to the x-axis between P and S at the instant Σ_m when information is transmitted to S for the m^{th} time since t=0. With obvious notation, since

$$x_p(\Sigma_m) = x_p(\Sigma_{m-1}) + v\,T_m$$

and

$$\begin{cases} x_s(\Sigma_m) = x_p(\Sigma_{m-1}) \text{ if } x_s(\Sigma_{m-1}) + u\,T_m \geq x_p(\Sigma_{m-1}) \\ x_s(\Sigma_m) = x_s(\Sigma_{m-1}) + u\,T_m \text{ if } x_s(\Sigma_{m-1}) + u\,T_m < x_p(\Sigma_{m-1}); \end{cases}$$

we have

$$\delta_m = v\,T_m \text{ if } x_s(\Sigma_{m-1}) + u\,T_m \geq x_p(\Sigma_{m-1});$$

$$\delta_m = \delta_{m-1} + (v - u)T_m \text{ if } x_s(\Sigma_{m-1}) + u\,T_m < x_p(\Sigma_{m-1})$$

or

$$\delta_m = v\,T_m \text{ if } \delta_{m-1} \leq u\,T_m;$$

$$\delta_m = \delta_{m-1} - u\,T_m + v\,T_m \text{ if } \delta_{m-1} > u\,T_m.$$

Writing $r_{m+1} = \delta_m - uT_{m+1}$, we have

$$\delta_m = v\,T_m \quad \text{if} \quad r_m \leq 0$$

$$\delta_m = v\,T_m + r_m \quad \text{if} \quad r_m > 0$$

i.e. $$\delta_m = \max\,[v\,T_m,\ v\,T_m + r_m]\ .$$

40. The last equation is extremely reminiscent of the equation for waiting time in a conventional queueing process. In fact the distribution function of δ_m can be easily derived theoretically. The agreement of independent calculations of this theoretical result (in the steady state) with the empirical distribution derived directly from the computer program inspires confidence in the latter as a representation of the real-life situation which it was desired to simulate.

41. I would now like to point out that the situation I have described here is formally a rather unusual single server queueing set-up in which arrival and service intervals are correlated. The connection was observed by Mr. Cruon when a paper on this subject was presented to the NATO Conference on Queueing Theory in 1965.

42. Denote the m^{th} piece of information by I_m. It arrives at time Σ_m. Let us now interpret I_m as a customer who demands as service that S be moved from wherever it is to a position with x coordinate equal to that of P at time Σ_m. Since the distance between P and S at time Σ_m is δ_m then obviously if waiting time includes time to complete service, and since S moves with speed u, the waiting time of I_m is δ_m/u.

43. The arrival intervals T δ the customers I_m are by definition distributed according to a negative exponential distribution with mean τ. If we say that service on I_m cannot begin until S reaches the position specified by I_{m-1} then the actual service time of I_m is $\frac{1}{u}[x_p(\Sigma_m) - x_p(\Sigma_{m-1})] = \frac{vT}{u}$, say. Thus, service time in this model is also negative exponentially distributed with mean $\lambda\tau$, where $\tau = v/u$.

44. Writing

$$a_n = \frac{(-)^{n-1}\,\lambda^{\frac{1}{2}n(n-1)}}{(1-\lambda)\,(1-\lambda^2)\ldots(1-\lambda^n)}$$

we have for the steady state distribution of δ

$$P_r\ [0 < \delta \leq x] = 1 - \sum_{n \geq 1} a_n \exp\left[\frac{-x}{v\tau}\left(1 + \frac{1}{\lambda} + \ldots\ldots + \frac{1}{\lambda^{n-1}}\right)\right] .$$

We have constructed a table which shows the comparison between theory and simulation of the distribution of δ for 1000 trials. It can be seen from this table that the agreement is satisfactory. Consequently one can have confidence that the random mechanisms employed in the simulation of the major problem are in fact behaving as they should. Equally we have an instance of how an Operational Research problem in an apparently completely unrelated field led, as a by-product, to an unusual queueing situation.

EXPERIMENT II

(a)

Int = 500

beta/lambdz=	.34230956	p mean=	22.025328
beta/lambdz=	1.2002621	p mean=	5.0002622
beta/lambdz=	1.0346396	p mean=	7.2430105
beta/lambdz=	1.0340420	p mean=	7.2511035
beta/lambdz=	.31698440	p mean=	34.731089
beta/lambdz=	.64124910	p mean=	11.729519
beta/lambdz=	.80192582	p mean=	8.7120959
beta/lambdz=	1.1020123	p mean=	8.0071046
beta/lambdz=	1.2941994	p mean=	5.7314638
beta/lambdz=	1.390113	p mean=	5.3794073
beta/lambdz=	1.3926331	p mean=	5.3665787
beta/lambdz=	1.1265391	p mean=	6.6089427
beta/lambdz=	1.3909330	p mean=	5.3752141
beta/lambdz=	1.3033840	p mean=	5.7101045
beta/lambdz=	.99689322	p mean=	7.0004222
beta/lambdz=	1.4500189	p mean=	5.1537031
beta/lambdz=	1.4700853	p mean=	5.0025714
beta/lambdz=	1.5515275	p mean=	4.0120205
beta/lambdz=	1.3926333	p mean=	4.4064245
beta/lambdz=	.41007421	p mean=	10.075733
beta/lambdz=	1.0538099	p mean=	7.1104427
beta/lambdz=	1.0170102	p mean=	7.0705200
beta/lambdz=	1.1710097	p mean=	8.3044397
beta/lambdz=	.59640976	p mean=	10.015080
beta/lambdz=	1.4200012	p mean=	5.0550431
beta/lambdz=	.45400925	p mean=	10.577759
beta/lambdz=	.33205466	p mean=	22.727003
beta/lambdz=	.84572053	p mean=	11.047044
beta/lambdz=	.50663477	p mean=	10.027006
beta/lambdz=	.73055369	p mean=	9.5515737
beta/lambdz=	1.4632072	p mean=	5.1067452
beta/lambdz=	1.0003527	p mean=	8.0094544
beta/lambdz=	1.1073027	p mean=	8.0001044
beta/lambdz=	.58515302	p mean=	14.040050
beta/lambdz=	.55006902	p mean=	10.989759
beta/lambdz=	.46340010	p mean=	11.100049
beta/lambdz=	1.1196354	p mean=	8.0022140
beta/lambdz=	.34054060	p mean=	30.357041
beta/lambdz=	.76415753	p mean=	9.0032700
beta/lambdz=	.43100780	p mean=	17.477002
beta/lambdz=	.41440506	p mean=	10.179545
beta/lambdz=	1.3121128	p mean=	5.7017109
beta/lambdz=	.91100062	p mean=	9.0592905
beta/lambdz=	.37417799	p mean=	20.144309
beta/lambdz=	1.3002913	p mean=	5.6246000
beta/lambdz=	1.0215031	p mean=	7.3400770
beta/lambdz=	.89002013	p mean=	10.095503
beta/lambdz=	.40319098	p mean=	8.3103435
beta/lambdz=	1.3164202	p mean=	5.6000266
beta/lambdz=	.56200046	p mean=	13.375200
beta/lambdz=	.59400724	p mean=	10.343171
beta/lambdz=	1.2400001	p mean=	6.0350025
beta/lambdz=	1.1920205	p mean=	6.0770340
beta/lambdz=	.50009563	p mean=	10.777621
beta/lambdz=	.70003100	p mean=	10.605465
beta/lambdz=	.62450075	p mean=	12.044150
beta/lambdz=	1.0059785	p mean=	7.4550893
beta/lambdz=	.54202473	p mean=	13.887733
beta/lambdz=	1.0199246	p mean=	4.0068953

TABLE 1

EXPERIMENT II (b)

beta/lambdz=	.38392241	p mean=	19.631528
beta/lambdz=	.80699907	p mean=	9.3080638
beta/lambdz=	.93074320	p mean=	8.0625614
beta/lambdz=	1.3062971	p mean=	5.7273652
beta/lambdz=	.43771174	p mean=	17.615509
beta/lambdz=	.34335117	p mean=	31.006393
beta/lambdz=	.82886100	p mean=	9.0609729
beta/lambdz=	.89427377	mean=	8.3938080
beta/lambdz=	.61824096	p mean=	12.138787
beta/lambdz=	.78861413	p mean=	9.5264612
beta/lambdz=	.61882436	p mean=	12.156744
beta/lambdz=	1.1318846	p mean=	6.6191429
beta/lambdz=	.42163116	p mean=	17.870408
beta/lambdz=	1.3984032	p mean=	5.3461793
beta/lambdz=	.75173390	p mean=	9.3961063
beta/lambdz=	1.1763884	p mean=	6.3664650
beta/lambdz=	1.1508356	p mean=	6.5091562
beta/lambdz=	1.1409923	p mean=	6.5658277
beta/lambdz=	1.2597965	p mean=	5.9409840
beta/lambdz=	1.4880486	p mean=	5.9179797
beta/lambdz=	1.5370756	p mean=	4.8584430
beta/lambdz=	.80438158	p mean=	9.8359780
beta/lambdz=	.65387223	p mean=	11.501921
beta/lambdz=	.73513244	p mean=	10.223607
beta/lambdz=	1.4768048	p mean=	5.9612533
beta/lambdz=	.72625886	p mean=	10.349537
beta/lambdz=	.73825341	p mean=	10.208228
beta/lambdz=	2.0427892	p mean=	3.6408467
beta/lambdz=	.67530189	p mean=	11.133995
beta/lambdz=	1.2428834	p mean=	6.0226453
beta/lambdz=	1.3535083	p mean=	5.5254986
beta/lambdz=	.93217366	p mean=	8.0590969
beta/lambdz=	.44003813	p mean=	16.779854
beta/lambdz=	1.0782371	p mean=	8.0650609
beta/lambdz=	1.4943938	mean=	4.9989800
beta/lambdz=	.70404610	p mean=	9.5823141
beta/lambdz=	.37835738	p mean=	19.921136
beta/lambdz=	.95579134	p mean=	7.3409917
beta/lambdz=	.81001450	p mean=	9.2731804
beta/lambdz=	1.0539094	p mean=	7.9929484
beta/lambdz=	1.3184849	p mean=	5.6738683
beta/lambdz=	.43218244	p mean=	17.432656
beta/lambdz=	.34069742	p mean=	9.9325561
beta/lambdz=	1.1428698	p mean=	6.5549459
beta/lambdz=	.51959638	p mean=	14.499791
beta/lambdz=	1.7900517	p mean=	4.1970791
beta/lambdz=	.69573642	p mean=	10.806210
beta/lambdz=	.68192464	p mean=	11.026297
beta/lambdz=	.40286062	p mean=	18.705842
beta/lambdz=	.99820711	p mean=	7.0836724
beta/lambdz=	1.1045914	p mean=	6.7341767
beta/lambdz=	.96224521	p mean=	7.7260439
beta/lambdz=	1.4554341	p mean=	5.1343392
beta/lambdz=	2.0317961	p mean=	3.6967150
beta/lambdz=	1.5701821	p mean=	4.7547402
beta/lambdz=	.31697625	p mean=	23.790427
beta/lambdz=	.71664559	p mean=	10.489174
beta/lambdz=	1.9189414	p mean=	3.8815804
beta/lambdz=	.78095111	p mean=	9.6205276
beta/lambdz=	.82550630	p mean=	9.1995221
beta/lambdz=	.54056368	p mean=	13.925436
beta/lambdz=	1.1219812	p mean=	6.6799977

TABLE 1 (continued)

EXPERIMENT II (c)

beta/lambdz=	**2.5662482**	**p**	**mean=**	**2.8859421**
beta/lambdz=	1.1377[illegible]		mean=	[illegible]
beta/lambdz=	.[illegible]		mean=	11.153395
beta/lambdz=	1.1[illegible]94614		mean=	[illegible].4045302
beta/lambdz=	1.1[illegible]17		mean=	[illegible]
beta/lambdz=	[illegible]35176		mean=	[illegible].324[illegible]037
beta/lambdz=	.54338[illegible]		mean=	1[illegible]
beta/lambdz=	.[illegible]414[illegible]	p	mean=	[illegible]
beta/lambdz=	.[illegible]1661088	p	mean=	[illegible].1377[illegible]
beta/lambdz=	.4[illegible]036523		mean=	17.5[illegible]519
beta/lambdz=	1.[illegible]51[illegible]		mean=	7.46[illegible]715
beta/lambdz=	.74747[illegible]33		mean=	1[illegible].[illegible]54[illegible]31
beta/lambdz=	.[illegible]177414	p	mean=	11.30[illegible]
beta/lambdz=	.5[illegible]		mean=	13.3[illegible]
beta/lambdz=	1.[illegible]10477		mean=	[illegible]
beta/lambdz=	.60[illegible]		mean=	1[illegible].343[illegible]
beta/lambdz=	[illegible]	p	mean=	[illegible].4[illegible]30471
beta/lambdz=	.[illegible]451[illegible]		mean=	1[illegible]
beta/lambdz=	1.5[illegible]		mean=	[illegible]
beta/lambdz=	1.[illegible]		mean=	7.11[illegible]75
beta/lambdz=	.[illegible]		mean=	[illegible].772[illegible]
beta/lambdz=	.[illegible]		mean=	1[illegible]
beta/lambdz=	1.[illegible]		mean=	7.[illegible]
beta/lambdz=	.[illegible]1556[illegible]		mean=	[illegible].197[illegible]
beta/lambdz=	1.[illegible]546[illegible]		mean=	7.1[illegible]
beta/lambdz=	1.[illegible]		mean=	[illegible]
beta/lambdz=	.[illegible]		mean=	[illegible]
beta/lambdz=	.[illegible]		mean=	[illegible]
beta/lambdz=	1.[illegible]		mean=	[illegible]
beta/lambdz=	1.51[illegible]0011		mean=	4.[illegible]
beta/lambdz=	1.1[illegible]414	p	mean=	[illegible]
beta/lambdz=	.[illegible]074		mean=	1[illegible]
beta/lambdz=	.7[illegible]		mean=	.7[illegible]15[illegible]
beta/lambdz=	1.[illegible]		mean=	7.[illegible]
beta/lambdz=	.[illegible]		mean=	1[illegible].[illegible]531
beta/lambdz=	1.[illegible]37		mean=	7.[illegible]
beta/lambdz=	.4[illegible]11		mean=	1[illegible].4[illegible]47
beta/lambdz=	1.1[illegible]		mean=	[illegible]
beta/lambdz=	1.[illegible]		mean=	[illegible].1[illegible]
beta/lambdz=	1.1[illegible]		mean=	[illegible].[illegible]11[illegible]7
beta/lambdz=	.[illegible]71[illegible]		mean=	[illegible].14[illegible]
beta/lambdz=	[illegible]		mean=	[illegible].4114577
beta/lambdz=	.[illegible]		mean=	11.[illegible]
beta/lambdz=	1.1[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	1.[illegible]		mean=	3.15[illegible]
beta/lambdz=	1.47[illegible]		mean=	[illegible]
beta/lambdz=	.[illegible]4		mean=	11.5[illegible]31
beta/lambdz=	.[illegible]		mean=	[illegible]
beta/lambdz=	.5[illegible]7		mean=	1[illegible]
beta/lambdz=	.4[illegible]		mean=	1[illegible]
beta/lambdz=	.751[illegible]	p	mean=	1[illegible]1741
beta/lambdz=	.7[illegible]45		mean=	[illegible]
beta/lambdz=	1.[illegible]5		mean=	[illegible]
beta/lambdz=	1.[illegible]54[illegible]		mean=	7.1[illegible]
beta/lambdz=	.[illegible]71		mean=	1[illegible].4[illegible]
beta/lambdz=	1.[illegible]37[illegible]		mean=	7.[illegible]
beta/lambdz=	.45[illegible]		mean=	1[illegible].6[illegible]
beta/lambdz=	.[illegible]		mean=	[illegible].5[illegible]

TABLE I (continued)

EXPERIMENT II (d)

beta/lambdz=	.14126939	p	mean= 53.454910
beta/lambdz=	.17476715	p	mean= 4[illegible].1[illegible]7065
beta/lambdz=	1.41[illegible]741	p	mean= 5.[illegible]04[illegible]
beta/lambdz=	.4372435[illegible]	p	mean= 1[illegible].[illegible]30177
beta/lambdz=	.[illegible]4[illegible]	p	mean= [illegible].[illegible]573248
beta/lambdz=	.[illegible]123374	p	mean= 7.5[illegible]77[illegible]
beta/lambdz=	1.1577243	p	mean= 6.47[illegible]652
beta/lambdz=	.61837460	p	mean= 1[illegible].135630
beta/lambdz=	.[illegible]	p	mean= 1[illegible].[illegible]
beta/lambdz=	.[illegible]83311[illegible]	p	mean= [illegible].[illegible]4[illegible]
beta/lambdz=	.[illegible]759420	p	mean= 1[illegible].444[illegible]54
beta/lambdz=	1.[illegible]210330	p	mean= 7.[illegible]1355
beta/lambdz=	.63[illegible]64[illegible]3	p	mean= 11.7[illegible]4
beta/lambdz=	1.14[illegible]576	p	mean= 6.531[illegible]41[illegible]
beta/lambdz=	[illegible].[illegible]7[illegible]485	p	mean= [illegible].[illegible]249
beta/lambdz=	1.[illegible]1[illegible]28	p	mean= [illegible].1[illegible]6122
beta/lambdz=	1.[illegible]976[illegible]2	p	mean= 6.[illegible]72[illegible]3
beta/lambdz=	.5[illegible]123130		mean= 11.444153
beta/lambdz=	.7[illegible]4[illegible]	p	mean= 1[illegible].[illegible]40592
beta/lambdz=	.611[illegible]8337	p	mean= 1[illegible].[illegible]
beta/lambdz=	1.[illegible]111976	p	mean= 4.11[illegible]442
beta/lambdz=	.7[illegible]1[illegible]251	p	mean= 1[illegible].[illegible]43320
beta/lambdz=	1.[illegible]113164	p	mean= 6.1[illegible]11[illegible]5
beta/lambdz=	.[illegible]4[illegible]	p	mean= 21.[illegible]
beta/lambdz=	.5[illegible]64[illegible]		mean= 1[illegible].[illegible]14[illegible]
beta/lambdz=	1.1[illegible]62771	p	mean= 6.77[illegible]4[illegible]
beta/lambdz=	.75[illegible]41[illegible]	p	mean= [illegible].[illegible]1[illegible]
beta/lambdz=	.[illegible]	p	mean= [illegible].1[illegible]342
beta/lambdz=	1.417[illegible]	p	mean= [illegible].[illegible]47[illegible]
beta/lambdz=	.[illegible]332464	p	mean= 7[illegible].[illegible]34[illegible]1
beta/lambdz=	.5[illegible]135[illegible]7	p	mean= 1[illegible].7[illegible]4[illegible]1[illegible]
beta/lambdz=	.[illegible]8371336	p	mean= 7.[illegible]31[illegible]
beta/lambdz=	.[illegible]1[illegible]375		mean= 1[illegible].[illegible]7271
beta/lambdz=	1.14[illegible]3[illegible]	p	mean= [illegible].[illegible]1443
beta/lambdz=	.[illegible]71[illegible]1[illegible]	p	mean= [illegible].[illegible]4[illegible]
beta/lambdz=	1.[illegible]1[illegible]4[illegible]1	p	mean= 7.[illegible]45[illegible]73
beta/lambdz=	.5[illegible]671[illegible]	p	mean= 1[illegible].[illegible]4[illegible]
beta/lambdz=	1.[illegible]77[illegible]	p	mean= 5.[illegible]1[illegible]5113
beta/lambdz=	1.[illegible]7[illegible]1[illegible]	p	mean= 4.[illegible]7[illegible]74[illegible]
beta/lambdz=	1.[illegible]	p	mean= 7.[illegible]1[illegible]
beta/lambdz=	.[illegible]5525	p	mean= 11.[illegible]5351
beta/lambdz=	1.4531137	p	mean= 5.131[illegible]7[illegible]
beta/lambdz=	1.[illegible]1[illegible]	p	mean= [illegible].1[illegible]4[illegible]
beta/lambdz=	.[illegible]77[illegible]	p	mean= 11.4[illegible]
beta/lambdz=	.[illegible]77[illegible]		mean= 1[illegible].[illegible]21[illegible]
beta/lambdz=	1.1[illegible]4[illegible]	p	mean= [illegible].4[illegible]541
beta/lambdz=	.[illegible]2454[illegible]1[illegible]		mean= [illegible].1143[illegible]
beta/lambdz=	[illegible].[illegible]134733	p	mean= [illegible].[illegible]554553
beta/lambdz=	1.[illegible]27[illegible]	p	mean= 7.[illegible]537127
beta/lambdz=	.5[illegible]164465	p	mean= 15.[illegible]1[illegible]487
beta/lambdz=	.4[illegible]1[illegible]43[illegible]	p	mean= 1[illegible].7[illegible]1763
beta/lambdz=	.5[illegible]211[illegible]4	p	mean= 1[illegible].705734
beta/lambdz=	1.[illegible]45[illegible]	p	mean= 5.735[illegible]57
beta/lambdz=	1.[illegible]245719	p	mean= 7.[illegible]1[illegible]7117
beta/lambdz=	.[illegible]1[illegible]56559	p	mean= 34.[illegible]71719
beta/lambdz=	1.12[illegible]56[illegible]	p	mean= 6.6[illegible]4[illegible]4
beta/lambdz=	1.[illegible]700691	p	mean= 7.[illegible]49817
beta/lambdz=	1.35[illegible]4487	p	mean= 5.5[illegible]51[illegible]51
beta/lambdz=	.55205413	p	mean= 1[illegible].[illegible]34343
beta/lambdz=	.3[illegible]1[illegible]8[illegible]1	p	mean= [illegible].441479
beta/lambdz=	.50[illegible]4[illegible]1[illegible]		mean= 14.7[illegible]

Best Available Copy

TABLE 1 (continued)

EXPERIMENT II

(e)

beta/lambdz=	1.1174040	p	mean= 6.7056988
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]	p	mean= [illegible]
beta/lambdz=	[illegible]	p	mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]	p	mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	1.3[illegible]71777		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]	p	mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	[illegible]		mean= [illegible]
beta/lambdz=	1.4457[illegible]		mean= 5.1[illegible]
beta/lambdz=	1.377[illegible]		mean= 5.4513[illegible]
beta/lambdz=	1.4[illegible]274		mean= 5.[illegible]
beta/lambdz=	.5[illegible]47500		mean= 1[illegible]

TABLE 1 (continued)

Best Available Copy

EXPERIMENT II (f)

beta/lambdz=	.30063233	p mean=	25.087059
beta/lambdz=	.87749743	p mean=	8.5554312
beta/lambdz=	.68546319	p mean=	10.[illegible]
beta/lambdz=	.71[illegible]	mean=	10.552763
beta/lambdz=	.57285319	p mean=	[illegible].601[illegible]71[illegible]
beta/lambdz=	.[illegible]5449432	p mean=	[illegible]1.[illegible]
beta/lambdz=	.711[illegible]	p mean=	1[illegible].[illegible]722
beta/lambdz=	.[illegible]87[illegible]045	p mean=	7.771[illegible]
beta/lambdz=	.63792254	mean=	11.[illegible]73[illegible]
beta/lambdz=	[illegible]	mean=	1[illegible].[illegible]5
beta/lambdz=	1.0407133	p mean=	7.[illegible]43347
beta/lambdz=	.[illegible]7[illegible]	p mean=	7.7[illegible]1773
beta/lambdz=	1.[illegible]3	p mean=	7.4[illegible]
beta/lambdz=	1.7[illegible]	mean=	4.1[illegible]1399
beta/lambdz=	.61134615	p mean=	1[illegible].[illegible]139
beta/lambdz=	1.[illegible]7[illegible]553	mean=	[illegible].[illegible]7750
beta/lambdz=	1.[illegible]	mean=	5.[illegible]34543
beta/lambdz=	2.7545513	mean=	[illegible].[illegible]45553
beta/lambdz=	1.51[illegible]	mean=	4.[illegible]7[illegible]53[illegible]
beta/lambdz=	.[illegible]071[illegible]	mean=	1[illegible].[illegible]5435
beta/lambdz=	.79153[illegible]51	mean=	1[illegible].713127
beta/lambdz=	.[illegible]4[illegible]73	mean=	1[illegible].[illegible]75
beta/lambdz=	1.[illegible]1195	p mean=	5.[illegible]15[illegible]7[illegible]
beta/lambdz=	.34[illegible]011	mean=	1[illegible].7[illegible]
beta/lambdz=	1.11[illegible]	p mean=	[illegible].7[illegible]1[illegible]3
beta/lambdz=	.84964[illegible]1	p mean=	[illegible].[illegible]
beta/lambdz=	.59412741	p mean=	1[illegible].[illegible]243
beta/lambdz=	.67[illegible]	p mean=	11.[illegible]75713
beta/lambdz=	.[illegible]55[illegible]5	p mean=	11.[illegible]
beta/lambdz=	1.1711[illegible]	mean=	[illegible].[illegible]1729
beta/lambdz=	.[illegible]5	p mean=	[illegible].[illegible]73471
beta/lambdz=	1.3324315	p mean=	5.[illegible]
beta/lambdz=	.52498[illegible]	p mean=	1[illegible].[illegible]5449
beta/lambdz=	1.[illegible]	p mean=	5.[illegible]37[illegible]
beta/lambdz=	1.[illegible]77[illegible]	p mean=	5.[illegible]
beta/lambdz=	.6[illegible]04	mean=	1[illegible].[illegible]3543
beta/lambdz=	.[illegible]1[illegible]	mean=	[illegible].[illegible]7[illegible]
beta/lambdz=	1.[illegible]	p mean=	5.[illegible]1[illegible]4[illegible]
beta/lambdz=	.7711147[illegible]	mean=	[illegible].740130
beta/lambdz=	1.[illegible]1[illegible]	mean=	[illegible].[illegible]13
beta/lambdz=	.7[illegible]5[illegible]1	p mean=	[illegible].[illegible]
beta/lambdz=	.5[illegible]	mean=	1[illegible].[illegible]41[illegible]7[illegible]
beta/lambdz=	.[illegible]	mean=	11.[illegible]4[illegible]77
beta/lambdz=	.3[illegible]7113	mean=	1[illegible].[illegible]
beta/lambdz=	1.1[illegible]	mean=	1.[illegible]37[illegible]1[illegible]
beta/lambdz=	.[illegible]1[illegible]	mean=	[illegible].1[illegible]
beta/lambdz=	.71[illegible]	mean=	1[illegible].74550[illegible]
beta/lambdz=	1.[illegible]1[illegible]1[illegible]	p mean=	5.[illegible]71[illegible]
beta/lambdz=	.[illegible]1[illegible]7[illegible]	mean=	[illegible].[illegible]4[illegible]3[illegible]
beta/lambdz=	.[illegible]3	mean=	[illegible]1.154755
beta/lambdz=	1.[illegible]7[illegible]1	mean=	5.457[illegible]173
beta/lambdz=	.[illegible]4[illegible]	mean=	7.[illegible]64562
beta/lambdz=	.[illegible]1[illegible]	mean=	1[illegible].[illegible]367
beta/lambdz=	[illegible]	mean=	[illegible].3411134
beta/lambdz=	1.[illegible]713[illegible]	mean=	[illegible].774[illegible]76[illegible]
beta/lambdz=	.[illegible]2777773	p mean=	[illegible].[illegible]5233
beta/lambdz=	1.[illegible]1[illegible]1	p mean=	7.345[illegible]57[illegible]
beta/lambdz=	1.7728915	p mean=	4.[illegible]4[illegible]
beta/lambdz=	1.[illegible]54440	mean=	5.[illegible]57[illegible]
beta/lambdz=	.[illegible]713704	p mean=	[illegible].[illegible]79[illegible]34
beta/lambdz=	.49302602	mean=	1[illegible].[illegible]13[illegible]
beta/lambdz=	.49163147	p mean=	15.633698
beta/lambdz=	.[illegible]1496223	p mean=	8.2935599

TABLE 1 (continued)

EXPERIMENT II (f)

beta/lambdz=	.68586243	p mean=	10.962646
beta/lambdz=	[illegible]	p mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	p mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	p mean=	[illegible]
beta/lambdz=	[illegible]	p mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	.45504345	p mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	p mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	p mean=	[illegible]
beta/lambdz=	[illegible]	p mean=	[illegible]
beta/lambdz=	[illegible]	p mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	p mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]
beta/lambdz=	[illegible]	mean=	[illegible]

TABLE 1 (continued)

EXPERIMENT II (b)

beta/lambdz=	.84705169	p	mean=	8.8650974
beta/lambdz=	[illegible]	p	mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	[illegible]
beta/lambdz=	2.5247977	p	mean=	2.9343066
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	9.7128520
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	4.9177515
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]	p	mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]
beta/lambdz=	[illegible]		mean=	[illegible]

TABLE 1 (continued)

EXPERIMENT II (i)

a= .250000001= .13333300

number of trials= 250000 mean= 9.8668640 variance= 206.09936 std dev=14.35616[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
-	1	[illegible]	[illegible]	[illegible]
1-	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	7	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	11	[illegible]	[illegible]	[illegible]
11-	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1-	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	1.7
-	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
1-	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
-	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]

TABLE 2

EXPERIMENT II (j)

under		0.0000
0-	0	0.1175
1-	1	[illegible]
2-	[illegible]	[illegible]
3-	3	[illegible]
4-	4	[illegible]
5-	[illegible]	[illegible]
6-	[illegible]	[illegible]
7-	7	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
11-	11	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible].7747
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
17-	17	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	31	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
34-	34	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	45	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	47	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
over		[illegible]

TABLE 2 (continued)

EXPERIMENT II (k)

under		0
0-	0	0
1-	1	[illegible]
2-	[illegible]	5
3-	3	1[illegible]
4-	4	24
5-	[illegible]	7[illegible]
6-	[illegible]	31
7-	7	4[illegible]
[illegible]	[illegible]	4[illegible]
[illegible]	[illegible]	4[illegible]
1[illegible]-	1[illegible]	[illegible]
11-	11	[illegible]
1[illegible]-	1[illegible]	[illegible]
1[illegible]-	1[illegible]	1[illegible]
14-	14	11
1[illegible]-	15	6
16-	1[illegible]	1[illegible]
17-	17	[illegible]
1[illegible]-	1[illegible]	10
19-	1[illegible]	5
over		[illegible]

Best Available Copy

TABLE 2 (continued)

EXPERIMENT III

No. of repetitions = 1000

a= .250000001= .13333330

number of trials= 1000

mean= 9.8830000 variance= 276.75607 stddev= 16.635987

mean= 10.4750[illegible] variance= 29[illegible].55093 stddev= 17.045[illegible]

RANGE		OBSERVED	EXPECTED	CHISQ
0-	1	136	122.9	1.39
1-	2	103	104.6	0.02
2-	3	[illegible]	89.5	0.02
3-	4	[illegible]	76.9	2.24
4-	5	51	66.4	3.56
5-	6	70	57.5	2.71
6-	7	[illegible]	[illegible]	[illegible]
7-	8	42	43.7	0.07
8-	9	[illegible]	[illegible]	[illegible]
9-	10	33	33.7	[illegible].41
10-	11	24	[illegible]	1.11
11-	12	31	26.3	[illegible]
12-	13	25	23.4	[illegible].11
13-	14	14	[illegible]	2.23
14-	15	[illegible]	[illegible]	[illegible]
15-	16	14	16.6	0.41
16-	17	23	14.9	1.74
17-	18	11	13.4	0.43
18-	19	14	12.1	0.31
19-	20	11	[illegible]	[illegible]
20-	21	3	9.9	4.7[illegible]
21-	22	11	[illegible]	0.47
22-	23	[illegible]	[illegible]	[illegible]
23-	24	7	7.4	0.02
24-	25	4	6.7	1.11
25-	26	5	6.2	.22
26-	27	2	5.6	2.34
27-	28	6	5.2	0.14
28-	29	[illegible]	4.[illegible]	2.26
29-	30	2	4.3	1.27
30-	31	2	4.	1.[illegible]
31-	32	4	3.7	[illegible]3
32-	33	4	3.4	.11
33-	34	2	3.1	.41
34-	35	3	2.9	[illegible]
35-	36	3	[illegible]	.1[illegible]
36-	37	2	2.5	.1
37-	38	1	2.3	.75
38-	39	3	2.1	.34
39-	40	2	2.[illegible]	[illegible]
40-	41	1	1.9	.4
41-	42	3	1.7	[illegible]
42-	43	2	1.6	[illegible]
43-	44		1.5	1.51
44-	45		1.4	1.41
45-	46	2	1.3	[illegible].35
46-	47	1	1.2	[illegible]5
47-	48	2	1.2	.61
48-	49		1.1	1.[illegible]
49-	50	1	1.	[illegible]
over		22	1[illegible].	4.3

CHISQUARE = 43.9424

TABLE 3

Best Available Copy

Experiment III (Contn)

0-	1	113	122.9	0.80
1-	2	96	1[illegible]4.6	[illegible].71
2-	3	1[illegible]5	[illegible].5	2.70
3-	4	7[illegible]	76.9	[illegible].62
4-	5	62	66.4	[illegible].29
5-	6	51	57.5	[illegible].74
6-	7	44	5[illegible].1	[illegible].73
7-	[illegible]	44	43.7	[illegible]
[illegible]-	9	41	3[illegible].3	[illegible].1[illegible]
9-	1[illegible]	39	33.7	[illegible]3
1[illegible]-	11	27	29.[illegible]	[illegible].26
11-	12	19	26.3	2.[illegible]5
12-	13	25	23.4	[illegible].11
13-	14	23	2[illegible]	[illegible].23
14-	15	1[illegible]	1[illegible].6	[illegible]1
15-	16	14	16.6	.41
16-	17	15	14.9	[illegible]
17-	1[illegible]	16	13.4	[illegible].59
1[illegible]-	1[illegible]	1[illegible]	12.1	.36
1[illegible]-	2[illegible]	11	1[illegible].9	[illegible]
2[illegible]-	21	14	[illegible].9	1.73
21-	22	4	[illegible]	2.74
22-	23	12	[illegible].1	1.[illegible]5
2[illegible]-	24	7	[illegible].4	[illegible]
24-	25	1	6.7	1.5[illegible]
25-	26	5	6.2	[illegible].22
26-	27	[illegible]	5.6	1.[illegible]
27-	2[illegible]	12	5.2	[illegible]
2[illegible]-	2[illegible]	[illegible]	4.7	[illegible].11
2[illegible]-	[illegible]	6	4.3	.[illegible]3
3[illegible]-	31	5	4.	.25
31-	32	5	3.7	.4[illegible]
32-	33	4	3.4	[illegible].11
33-	34	4	3.1	.24
[illegible]-	35	[illegible]	2.[illegible]	3.[illegible]
35-	[illegible]	4	2.7	.64
36-	37	4	2.5	.[illegible]1
[illegible]-	[illegible]	2	2.3	[illegible]4
[illegible]-	3[illegible]	3	2.1	[illegible].34
3[illegible]-	4[illegible]	4	2.[illegible]	[illegible]
4[illegible]-	41	1	1.[illegible]	.4
41-	42	1	1.7	.[illegible]1
42-	43	1	1.6	.24
4[illegible]-	44	3	1.5	1.47
4[illegible]-	45	[illegible]	1.4	1.7[illegible]
45-	46	1	1.3	[illegible]
46-	47	2	1.2	.47
47-	4[illegible]	1	1.2	[illegible]2
4[illegible]-	4[illegible]	1	1.1	[illegible]1
4[illegible]-	5[illegible]		1.[illegible]	1.[illegible]2
vo[illegible]	1[illegible]		1[illegible].	[illegible].3

CHI[illegible] = [illegible]

TABLE 4

Best Available Copy

EXPERIMENT IV

No. of repetitions- 1000

a= .10000000
mean of first int= ?.1461653mean of tenth int= ?.6574292
sd of first int= ?.7?41161sd of tenth int= 9.6225436
prod= ?5.14141C
corr.coeff.= .17292435

a= .2?
mean of first int= 1 .15?C45mean of tenth int= 9.?224272
sd of first int= 14.17745?sd of tenth int= 13.115371
prod= 142.5?937
corr.coeff.= .22973229

a= .3?
mean of first int= 1 .?15?C1mean of tenth int= 11.461651
sd of first int= 1?.31 319sd of tenth int= 17.?37146
prod= 226.?9299
corr.coeff.= .35???664

a= .4???????
mean of first int= 14.1?8271mean of tenth int= 11.29?154
sd of first int= 44.??3251sd of tenth int= 2?.7?93?7
prod= 714.?237?
corr.coeff.= .397631?6

a= .5?
mean of first int= 1?.4? ?59mean of tenth int= 14.??3?71
sd of first int= 22.? 405sd of tenth int= 34.4?373
prod= 39?.1?3?
corr.coeff.= .25??15?2

a= .6?
mean of first int= 21.3?54?1mean of tenth int= 2 .?.?1?5
sd of first int= 75.435??7?sd of tenth int= 123.?313?
prod= 5560.?745
corr.coeff.= .35213174

a= .7?
mean of first int= 2 .956??6mean of tenth int= 2?.44?7?2
sd of first int= 1,431959sd of tenth int= 72.756??7
prod= 331 . C72
corr.coeff.= .47?43??

a= .?
mean of first int= 3?.?6?3?7mean of tenth int= 2?.35?731
sd of first int= 137.4357?sd of tenth int= 112.58274
prod= 7147.1?6?
corr.coeff.= .4?73??.

a= .9?
mean of first int= 6?.72142mean of tenth int= 37.??4662
sd of first int= 6?.?6672sd of tenth int= 104.?53?1
prod= ?0544.131
corr.coeff.= .71???373

TABLE 5

EXPERIMENT V (a)

No. of repetitions= 1

a= .10000000 l= 10.000000

0	0.055	56	7.138
1	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	7.350
3	[illegible]	[illegible]	[illegible]
4	[illegible]	[illegible]	[illegible]
5	[illegible]	61	7.664
[illegible]	[illegible]	[illegible]	[illegible]
7	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	64	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
11	[illegible]	[illegible]	[illegible]
[illegible]	1.305	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	71	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
17	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	77	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
27	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		

Best Available Copy

TABLE 5 (continued)

EXPERIMENT V (b)

a= .30000000 l= 10.000000

0	0.322
1	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
3	1.[illegible]
4	[illegible]
5	[illegible]
[illegible]	3.[illegible]27
7	[illegible]
[illegible]	4.127
[illegible]	4.722
10	4.72[illegible]
11	5.3[illegible]1
12	5.41[illegible]
13	5.627
14	[illegible]51
15	[illegible]
1[illegible]	[illegible]
17	[illegible]
1[illegible]	7.[illegible]
1[illegible]	7.[illegible]
[illegible]	[illegible]
21	[illegible].113
[illegible]	[illegible].13[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible].540
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
31	[illegible].54[illegible]
[illegible]	1[illegible].34[illegible]
[illegible]	1[illegible].35[illegible]
[illegible]	1[illegible].4[illegible]
[illegible]	1[illegible].5[illegible]
[illegible]	[illegible]
[illegible]	11.[illegible]
[illegible]	11.[illegible]
[illegible]	11.1[illegible]
[illegible]	11.1[illegible]
[illegible]	11.4[illegible]1
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]41
[illegible]	[illegible]
[illegible]	[illegible]

56	14.229
57	14.[illegible]40
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	14.[illegible]93
60	15.535
61	15.7[illegible]
[illegible]	15.75[illegible]
[illegible]3	1[illegible].11[illegible]
[illegible]	1[illegible].45[illegible]
[illegible]	1[illegible].476
[illegible]	17.068
[illegible]	17.[illegible]
[illegible]	17.250
[illegible]	17.707
[illegible]	17.[illegible]
71	1[illegible].[illegible]97
[illegible]	1[illegible].43[illegible]
[illegible]	1[illegible].[illegible]44
[illegible]4	[illegible]
75	1[illegible].[illegible]25
[illegible]	[illegible]
77	[illegible]
[illegible]	1[illegible].351
[illegible]	1[illegible].377
[illegible]	1[illegible].374
[illegible]1	1[illegible].[illegible]24
[illegible]	[illegible].335
[illegible]	[illegible]
[illegible]	[illegible]1.[illegible]40
[illegible]	[illegible]1.[illegible]
[illegible]	[illegible]1.157
[illegible]	[illegible]
[illegible]	[illegible].543
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]1	[illegible]
[illegible]	[illegible].700
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]5.11[illegible]
[illegible]7	[illegible]5.511
[illegible]	[illegible]
[illegible]	[illegible].4[illegible]5

Best Available Copy

TABLE 5 (continued)

EXPERIMENT V (c)

a= .50000000 l= 10.000000

0	0.397	56	29.213
1	[illegible]	57	[illegible]
2	1.657	58	[illegible]
3	2.702	59	30.154
4	[illegible]	60	[illegible]
5	3.372	61	[illegible]
6	[illegible]	62	[illegible]
7	3.417	63	[illegible]
8	[illegible]	64	[illegible]
9	[illegible]	65	[illegible]
10	[illegible]	66	[illegible]
11	[illegible]	67	[illegible]
12	[illegible]	68	[illegible]
13	7.752	69	[illegible]
14	[illegible]	70	[illegible]
15	[illegible]	71	[illegible]
16	[illegible]	72	[illegible]
17	[illegible]	73	[illegible]
18	[illegible]	74	[illegible]
19	[illegible]	75	[illegible]
20	[illegible]	76	[illegible]
21	[illegible]	77	[illegible]
22	[illegible]	78	[illegible]
23	[illegible]	79	[illegible]
24	[illegible]	[illegible]	[illegible]
25	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
51	[illegible]		
[illegible]	[illegible]		
[illegible]	[illegible]		
54	[illegible]		
[illegible]	[illegible]		

TABLE 5 (continued)

EXPERIMENT V (d)

a= .70000000 l= 10.000000

0	0.014
1	[illegible]
2	0.047
3	[illegible]
4	[illegible]
5	[illegible]
6	0.220
7	[illegible]
8	0.297
9	0.455
10	0.475
11	0.553
12	[illegible]
13	[illegible]
14	[illegible]
15	[illegible]
16	[illegible]
17	0.828
18	[illegible]
19	[illegible]
20	[illegible]
21	[illegible]
22	[illegible]
23	[illegible]
24	[illegible]
25	[illegible]
26	[illegible]
27	[illegible]
28	[illegible]
29	1.377
30	[illegible]
31	[illegible]
32	[illegible]
33	[illegible]
34	[illegible]
35	[illegible]
36	[illegible]
37	[illegible]
38	[illegible]
39	[illegible]
40	[illegible]
41	[illegible]
42	2.327
43	[illegible]
44	[illegible]
45	[illegible]
46	[illegible]
47	[illegible]
48	[illegible]
49	2.751
50	[illegible]
51	[illegible]
52	[illegible]
53	[illegible]
54	[illegible]
55	[illegible]
56	2.921
57	[illegible]
58	[illegible]
59	[illegible]
60	[illegible]
61	[illegible]
62	[illegible]
63	3.244
64	[illegible]
65	3.503
66	[illegible]
67	3.547
68	3.636
69	[illegible]
70	[illegible]
71	[illegible]
72	[illegible]
73	[illegible]
74	[illegible]
75	[illegible]
76	[illegible]
77	[illegible]
78	[illegible]
79	[illegible]
80	[illegible]
81	[illegible]
82	[illegible]
83	[illegible]
84	[illegible]
85	[illegible]
86	[illegible]
87	[illegible]
88	[illegible]
89	[illegible]
90	4.711
91	[illegible]
92	[illegible]
93	[illegible]
94	[illegible]
95	[illegible]
96	5.075
97	[illegible]
98	[illegible]
99	5.171

TABLE 5 (continued)

EXPERIMENT V

(e)

a= .90000000 l= 10.000000

0	0.010	51	2.928
1	0.025	52	2.950
2	0.047	53	2.982
3	0.105	54	3.004
4	0.131	55	3.038
5	0.21[illegible]	56	3.[illegible]33
[illegible]	[illegible].286	57	3.49[illegible]
7	0.292	5[illegible]	3.513
[illegible]	0.431	5[illegible]	3.752
[illegible]	0.530	60	3.[illegible]21
1[illegible]	0.573	61	3.[illegible]37
11	0.[illegible]97	62	3.[illegible]53
1[illegible]	[illegible].77[illegible]	63	4.[illegible]7
1[illegible]	[illegible].[illegible]95	64	4.15[illegible]
1[illegible]	[illegible].[illegible]1	65	4.17[illegible]
15	[illegible]	66	[illegible]
1[illegible]	[illegible].[illegible]13	67	4.[illegible]
17	1.[illegible]47	[illegible]	[illegible]
1[illegible]	1.[illegible]1	[illegible]	4.4[illegible]
1[illegible]	1.[illegible]	7[illegible]	[illegible]
[illegible]	1.1[illegible]7	71	[illegible]
[illegible]1	1.1[illegible]	7[illegible]	[illegible]
[illegible]	1.[illegible]	7[illegible]	[illegible]
[illegible]	1.[illegible]	7[illegible]	[illegible]
[illegible]	1.31[illegible]	7[illegible]	[illegible]
[illegible]	1.4[illegible]1	7[illegible]	[illegible]
[illegible]	1.4[illegible]1	77	5.1[illegible]
[illegible]	1.[illegible]	7[illegible]	5.[illegible]
[illegible]	1.56[illegible]	7[illegible]	5.3[illegible]
[illegible]	1.5[illegible]	[illegible]	5.217
[illegible]	1.6[illegible]	[illegible]1	5.3[illegible]
[illegible]1	1.71[illegible]	[illegible]	[illegible]
[illegible]	1.[illegible]1	[illegible]	5.431
[illegible]	[illegible]	[illegible]	5.[illegible]
[illegible]	[illegible]	[illegible]	5.77[illegible]
[illegible]	[illegible].31[illegible]	[illegible]	5.[illegible]
[illegible]	[illegible].1[illegible]1	[illegible]7	5.[illegible]
[illegible]	[illegible].1[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	6.002
[illegible]	[illegible]	[illegible]1	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible].1[illegible]	[illegible]	[illegible].040
[illegible]	[illegible]	[illegible]	[illegible].11[illegible]
[illegible]	[illegible]	[illegible]	[illegible].4[illegible]
[illegible]	[illegible]	[illegible]	[illegible].447
[illegible]	[illegible]	[illegible]7	[illegible]
[illegible]	[illegible]	[illegible]	[illegible].547
4	[illegible].7[illegible]		[illegible]
[illegible]	[illegible]		
[illegible]	[illegible].11		

TABLE 5 (continued)

Best Available Copy

EXPERIMENT V (f)

a= 1.1000000 l= 10.000000

0	0.114	51	16.205
1	[illegible].172	52	[illegible]
2	0.217	53	17.070
[illegible]	[illegible]	54	17.[illegible]
4	0.512	55	1[illegible].447
5	[illegible]	56	1[illegible].7[illegible]
6	0.745	57	[illegible]
7	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	1[illegible].650
[illegible]	[illegible]	60	[illegible].775
10	1.[illegible]	[illegible]	[illegible]
11	[illegible].144	[illegible]	[illegible]1.4[illegible]7
1[illegible]	[illegible]	[illegible]	[illegible]1.744
1[illegible]	[illegible]	64	[illegible]1.[illegible]75
14	[illegible]	[illegible]	[illegible]
15	[illegible]	[illegible]	[illegible]
1[illegible]	[illegible]	[illegible]	[illegible]
1[illegible]	[illegible]	[illegible]	[illegible]
1[illegible]	[illegible]	[illegible]	[illegible]
1[illegible]	4.[illegible]1	7[illegible]	[illegible]
[illegible]	[illegible]	71	[illegible]
21	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]7	[illegible]	[illegible]	[illegible]
[illegible]	7.7[illegible]1	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible].217
[illegible]	[illegible]	[illegible]1	[illegible]
[illegible]1	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible].73[illegible]	[illegible]	[illegible]7.351
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]7	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	11.[illegible]53	[illegible]1	[illegible]
[illegible]1	[illegible]	[illegible]	[illegible]
4[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible].[illegible]71
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]5	[illegible]	[illegible]	[illegible]
[illegible]	1[illegible].74[illegible]	[illegible]7	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]		
[illegible]	[illegible]		

TABLE 5 (continued)

EXPERIMENT V (c)

a= 1.3000000 l= 10.000000

0	0.046
1	[illegible]
2	[illegible]
3	[illegible]
4	0.130
5	[illegible]
6	[illegible]
7	[illegible]
8	[illegible]
9	[illegible]
10	[illegible]
11	[illegible]
12	[illegible]
13	[illegible]
14	[illegible]
15	0.414
16	[illegible]
17	[illegible]
18	[illegible]
19	[illegible]
20	[illegible]
21	[illegible]
22	[illegible]
23	[illegible]
24	[illegible]
25	[illegible]
26	[illegible]
27	[illegible]
28	[illegible]
29	[illegible]
30	[illegible]
31	[illegible]
32	[illegible]
33	[illegible]
34	[illegible]
35	[illegible]
36	[illegible]
37	[illegible]
38	[illegible]
39	[illegible]
40	[illegible]
41	[illegible]
42	[illegible]
43	[illegible]
44	[illegible]
45	[illegible]
46	[illegible]
47	[illegible]
48	[illegible]
49	[illegible]
50	[illegible]
51	1.360
52	[illegible]
53	1.374
54	1.411
55	[illegible]
56	1.457
57	[illegible]
58	[illegible]
59	[illegible]
60	[illegible]
61	[illegible]
62	[illegible]
63	1.627
64	[illegible]
65	[illegible]
66	[illegible]
67	[illegible]
68	[illegible]
69	[illegible]
70	[illegible]
71	[illegible]
72	[illegible]
73	1.906
74	[illegible]
75	[illegible]
76	[illegible]
77	[illegible]
78	[illegible]
79	[illegible]
80	[illegible]
81	[illegible]
82	[illegible]
83	[illegible]
84	[illegible]
85	[illegible]
86	[illegible]
87	[illegible]
88	[illegible]
89	[illegible]
90	[illegible]
91	[illegible]
92	[illegible]
93	[illegible]
94	[illegible]
95	[illegible]
96	[illegible]
97	[illegible]
98	[illegible]
99	[illegible]

TABLE 5 (continued)

COMPUTATIONAL CONSIDERATIONS IN MULTIPLE LINEAR REGRESSION*

Harold J. Breaux
U. S. Army Ballistic Research Laboratories
Aberdeen Proving Ground, Maryland

INTRODUCTION

The statistical theory concerned with multiple linear regression and simple, partial and multiple correlation is highly developed and has been one of the most useful tools of analysis provided by statistics. The widespread availability of modern high speed computing machinery makes practical the solution of many regression problems which beforehand might not have been attempted due to the inherent computational difficulties. High speed computing machinery enhances the value of multiple linear regression by removing the computational drudgery and making possible more sophisticated procedures of analysis. Despite the tremendous speed and computing capabilities of modern computers, much can be gained by the skillful design of computer programs designed to solve the normal equations and provide the associated statistical data for estimating significance of variables and prediction intervals. The computational labor associated with multiple linear regression arises in the formation and solution of the normal equations. Efficient algorithms for solving the normal equations are described in the commonly used texts of statistics and numerical analysis, however, only recently has any widespread effort been made to fully take advantage of the capabilities of computers for doing "exploratory" type regression computations. In problems where many variables are involved the analyst may have only intuitive suspicion regarding those variables which are significant. When this is true it is desirable to define a "candidate" linear model which includes all the variables which are conceivably significant. The exploratory experiment then would consist of entering this candidate model and the appropriate available data to a computer program specifically designed to analyse this model, and output a reduced model containing only significant variables.

One way to design such a program is to have it obtain the solution to all the "sub-set" models that can be formed from the collection of variables in the candidate and choose the one which best meets the significance criteria.

*This paper has been reproduced photographically.

If this model contains N variables there are $2^N - 1$ sub-set models. This method is made practical for as many as 20 variables by a "binary algorithm" decribed by Lotto [1], 1961, and Garside [2], 1965. This binary algorithm defines the optimum path of elimination so that the Gauss-Jordan algorithm goes through the fewest recursions when generating the $2^N - 1$ solutions. The method has the advantage of being always able to identify the "optimum model". For the purpose of this paper the optimum model is defined as that model containing only variables which are statistically significant at a chosen level of significance and which has the minimum variance of residuals among the sub-models that have all terms significant at that level.

The scope of some regression problems is such, however, that more than twenty variables are required in the candidate model. Such a problem is one described by the author in BRL Report No. 1348*, "The Computation of Firing Tables for Guided Missiles", [3]. In this problem it is desirable to define a candidate model containing 100 or more terms. A very practical solution was obtained using "Stepwise Multiple Linear Regression". The program was patterned after the computational scheme described by M. A. Efroymson [4] and is documented in BRL Report No. 1330 [5]. For documentation of similar type programs see References [6], [7] and [8].

Stepwise Multiple Regression takes advantage of the fact that the Gauss-Jordan algorithm, when used to solve the normal equations with N variables, yields intermediate solutions to N regression problems containing respectively 1, 2, ... and N variables. The procedure advances in stages. In the "forward" version the variable which enters into the regression is the one which at that stage results in the greatest reduction in the sum of squares of residuals. The power of the procedure is further enhanced by removing variables at later stages that may have become insignificant. The decision to add or remove variables is made by use of "t" or "f" tests of significance. The procedure advances until an equilbrium point is reached where no significant reduction in the sum of squares of residuals is to be gained by adding variables into the regression and where a significant increase arises if a variable is removed. The "backward" version of the procedure begins with all variables in regression and proceeds in the opposite direction to achieve the equilibrium stage. The relative advantage and disadvantages of the two procedures is dependent upon the application however, it seems desirable for a well designed computer program to contain a capability for both.

---

*Copies of this report are available to qualified requestors.

MATHEMATICAL BASIS OF THE STEPWISE REGRESSION

The mathematical basis of the stepwise regression is that the transformation rules of the Gauss-Jordan algorithm correspond to recurrence relations that exist between covariances of residuals, regression coefficients, and inverse elements of partitions of the covariance matrix. These relations are conveniently expressed by taking advantage of Yule's notation [9]. In this notation the regression equation is written in the form

$$X_n = b_{n1.23\ldots n-1} X_1 + b_{n2.13\ldots n-1} X_2 + \ldots + b_{n,n-1.12\ldots n-2} X_{n-1} \quad (1)$$

The first subscript of b is that corresponding to the dependent variable X_n, the second subscript corresponds to the independent variable attached to the regression coefficient. These two subscripts are called the primary subscripts. The remaining subscripts on the right of the period are those of the remaining independent variables and are called secondary subscripts. For a particular observation equation (1) takes the form

$$X_{jn} = b_1 X_{j1} + b_2 X_{j2} + \ldots + b_{n-1} X_{j,n-1} + e_j \quad (2)$$

e_j is a residual and is the difference between the predicted value and the observed value of X_n*. In Yule's notation the residuals are denoted as $X_{n.12\ldots n-1}$. Since regressions containing fewer than the (n-1) independent variables are of interest it is convenient to introduce the notation

$$q = 1, 2, \ldots (i-1), (k+1), \ldots p \quad (3)$$

Note that q is a set of subscripts containing the digits 1 through p, excluding i, j, and k. Furthermore q is a sub-set of the (n-1) subscripts of the independent variables.

---

*It should be noted that the variables X_i are assumed to be measured without error.

The covariance of the variables X_i and X_j is defined as

$$s_{ij} = \sum X_i X_j / f$$

where f is the degrees of freedom and the summation extends over the m data points. Any variable can be considered as the dependent variable e.g., the residuals $X_{i.q}$ and $X_{j.q}$ will be of interest. The covariance of residuals is defined as

$$s_{ij.q} = \sum X_{i.q} \; X_{j.q} / f$$

Using the above notation, the normal equations can be written in the form

$$\sum X_{n.12\ldots n-1} \; X_k = 0, \quad k = 1, 2, \ldots, n-1 \tag{4}$$

or equivalently

$$s_{1k} \; b_1 + s_{2k} \; b_2 + \ldots + s_{n-1,\, k} \; b_{n-1} = s_{nk}, \qquad k = 1, 2, \ldots, n-1 \tag{5}$$

The complete covariance matrix is

$$S = \begin{matrix} s_{11} & s_{12} & \ldots s_{1n} \\ s_{21} & s_{22} & \ldots s_{2n} \\ \ldots & \ldots & \ldots \\ s_{n1} & s_{n2} & \ldots s_{nn} \end{matrix} \tag{6}$$

This matrix corresponds to the augmented matrix of coefficients usually considered in solving a system of linear equations with the addition of the nth row. The nth row is added so that the variance of residuals, $s_{nn.q}$ will be made available through matrix manipulations, thus avoiding the need for computing residuals at each stage.

The matrix element $X_{ij.q\; ij\; k}$ is defined as the ij' th element of the inverse of the partition of the covariance matrix formed by taking all the rows and columns of indices q, i, j, k.

The recurrence relations between the b's, c's and s's that are of interest in stepwise multiple regression are tabulated in Table 1.

The solution of the normal equations by the Gauss-Jordan algorithm is equivalent to the successive application of linear transformations to transformed matrices, the initial matrix being the covariance matrix. The successive matrices that are generated by the recursive equations can be denoted as A_o, A_1, ... A_{n-1}. A_k (k = 1, 2, ...n-1) is the matrix formed by applying the transformation.

$$
\begin{aligned}
a_{ij}^{k} &= a_{ij}^{k-1} - a_{ik}^{k-1}\, a_{kj}^{k-1} / a_{kk}^{k-1}, && i = 1, 2, \ldots, (k-1)\ (k+1) \ldots, n \\
& && j = 1, 2, \ldots, (k-1)\ (k+1) \ldots, n \\
a_{ik}^{k} &= -\, a_{ik}^{k-1} / a_{kk}^{k-1} && i = 1, 2, \ldots, (k-1)\ (k+1) \ldots, n \\
a_{kj}^{k} &= a_{kj}^{k-1} / a_{kk}^{k-1} && j = 1, 2, \ldots, (k-1)\ (k+1) \ldots, n \\
a_{kk}^{k} &= 1 / a_{kk}^{k-1} && i = j = k
\end{aligned}
\tag{7}
$$

to the matrix A_{k-1}. This transformation is denoted as T_k. The superscripts denote the fact that the matrix A_{k-1} is being operated on to yield A_k. The sequence corresponds to the introduction of the variables into the regression in the order 1, 2, ... n-1. In general the sequence would be different, however, no loss of generality arises, since one can renumber the variables in any arbitrary fashion. By use of the recurrence formulas one can prove the following theorem:

TABLE 1
RECURRENCE FORMULAS

1. $c_{ij.qi\,jk} = c_{ij.qi\,j} - b_{ki.qj}\, d_{kj.qi} / s_{kk.qi\,j}$

2. $c_{ik.qi\,jk} = -b_{ki.qj} / s_{kk.qij}$

3. $b_{ji.qk} = b_{ji.q} - b_{ki.q}\, s_{kj.qi} / s_{kk.qi}$

4. $c_{kj.qi\,jk} = d_{kj.qi} / s_{kk.qi\,j}$

5. $c_{kk.qi\,jk} = 1 / s_{kk.qi\,j}$

6. $b_{jk.q} = s_{kj.q} / s_{kk.q}$

7. $d_{ij.qk} = d_{ij.q} - d_{kj.q}\, s_{ik.qj} / s_{kk.qj}$

8. $d_{ik.q} = -s_{ik.q} / s_{kk.q}$

9. $s_{ij.qk} = s_{ij.q} - s_{ik.q}\, s_{kj.q} / s_{kk.q}$

10. $c_{ij.qi\,j} = c_{ij.qi\,jk} - c_{ik.qi\,jk}\, c_{jk.qi\,jk} / c_{kk.qi\,jk}$

11. $b_{ki.qj} = -c_{ki.qj} / c_{kk.qi\,jk}$

12. $b_{ji.q} = b_{ji.qk} - c_{ik.qi}\, b_{jk.qi} / c_{kk.qi\,k}$

13. $d_{kj.qi} = c_{kj.qi\,jk} / c_{kk.qi\,jk}$

14. $s_{kk.qi\,j} = 1 / c_{kk.qi\,jk}$

15. $s_{kj.q} = b_{jk.q} / c_{kk.qk}$

16. $d_{ij.q} = d_{ij.q} - d_{ik.qj}\, c_{jk\ \ jk} / c_{kk.qjk}$

17. $s_{ik.q} = -d_{ki.q} / c_{kk.qk}$

18. $s_{ij.q} = s_{ij.qk} - d_{ik.q}\, b_{jk.q} / c_{kk.qk}$

THEOREM: *

The matrix A_K, defined above, contains four partitions, the respective partitions having elements as follows:

$$
\begin{array}{llll}
a_{ij} = c_{ij.12\ldots k}, & i = 1,2,\ldots k, & j = 1,2,\ldots k & \\
a_{ij} = b_{ji.12\ldots i-1,\ i+1\ldots k}, & i = 1,2,\ldots k, & j = k+1,k+2,\ldots n & \\
 & & & (8) \\
a_{ij} = d_{ij.12\ldots i-1,\ i+1\ldots k}, & i = k+1,\ k+2,\ldots n, & j = 1,2,\ldots n & \\
a_{ij} = a_{ij.12\ldots k}, & i = k+1,\ k+2,\ldots n, & j = k+1,k+2,\ldots n &
\end{array}
$$

The consequence of the above theorem can be generalized as follows: The collection of variables whose subscripts are represented by the values taken by k in the successive application of T_K are said to be in regression if k appears an odd number of times in the collection. Alternatively, a variable is said not to be in regression if its subscript does not appear in the collection, or if it appears an even number of times. If the subscript appears twice, e.g., the corresponding variable was entered into the regression and then removed. The nine recurrence formulas, 10. through 18. can be used to prove that the application of the transformation T_k to A_k generates the matrix A_{k-1}, i.e., the variable is removed from regression by the same algorithm with which it is entered.

---

*The derivation of the eighteen recurrence formulas and the proof of this theorem are contained in the author's Masters' Thesis, soon to be presented to the Graduate School, Department of Statistics and Computer Science, University of Delaware, Newark, Delaware. The thesis also contains a discussion of storage saving considerations in the programming of the procedure.

The content of the matrix at any stage is as follows:

$a_{ij} \quad = \quad s_{ij}$ when neither X_i nor X_j are in regression

$a_{ij} \quad = \quad b_{ji}$ when X_i is in regression but not X_j

$a_{ij} \quad = \quad d_{ij}$ when X_j is in regression but not X_i

$a_{ij} \quad = \quad c_{ij}$ when both X_i and X_j are in regression.

CHOOSING THE KEY ELEMENT

In forward stepwise regression the variable which is entered into regression is the one which yields the greatest reduction in the variance of residuals at that stage. For an arbitrary variable X_i that is not in regression it is seen from the recurrence formula 9. that the variance reduction is given by the quantity.

$$V_i \quad = \quad a_{in}\, a_{ni} / a_{ii} = s_{in.q}\; s_{ni.q} / s_{ii.q} \tag{9}$$

For an arbitrary variable X_i that is in regression the variance increase resulting from the removal of X_i from regression is given by 18.

$$V_i \quad = \quad a_{in}\, a_{ni} / a_{ii} = d_{ni.q}\; b_{ni.q} / c_{ii.qi} \tag{10}$$

For X_i not in regression V_i is positive and for X_i in regression V_i is negative.

After determining the key element it is necessary to test whether the variance reduction due to entering the key variable is statistically significant. By inspection of 9. it is seen that for $i = j = n$

$$s_{nn.qk} = s_{nn.q}\,(1 - s_{nk.q}\; s_{kn.q} / s_{nn.q}\; s_{kk.q}) \tag{11}$$

The quantity $(s_{nk.q}\; s_{kn.q} / s_{nn.q}\; s_{kk.q})^{\frac{1}{2}}$ is defined as the product moment coefficient of correlation between $X_{n.q}$ and $X_{k.q}$.

This quantity is denoted as $r_{nk.q}$ and is often referred to as a partial correlation coefficient. Equation (11) can be written in the form

$$r^2_{nk.q} = s_{nk.q}\, s_{kn.q} / s_{nn.q}\, s_{kk.q} = (s_{nn.q} - s_{nn.qk}) / s_{nn.q} \qquad (12)$$

By inspection $r^2_{nk.q}$ gives the fractional variance reduction obtained by adding X_k into the regression. If $r_{nk.q}$ is statistically different from zero, then we observe that the fractional variance reduction due to X_k is significant and that X_k should be brought into regression. For forward recursion $r^2_{nk.q}$ can be computed directly from the first expression of (12). For backwards recursion, i.e., to test whether a variable X_k can be removed from regression, $r^2_{nk.q}$ can be computed from the formula

$$r^2_{nk.q} = V_k / (s_{nn.qk} + V_k) \qquad (13)$$

A test of significance for $r_{nk.q}$ is listed by Graybill [10]. If the true coefficient $\bar{r}_{nk.q}$, for which $r_{nk.q}$ is an estimate, is zero the quantity

$$t = r_{nk.q} (f-2)^{\frac{1}{2}} / (1 - r^2_{nk.q})^{\frac{1}{2}} \qquad (14)$$

is distributed as the Student t distribution. A test of the hypothesis $r_{nk.q} \neq 0$ against the alternative $r_{nk.q} = 0$ is performed as follows: The quantity t is compared against the one-tailed t statistic, $t(f-2, c)$ appropriate to the degrees of freedom, f, and the confidence level, c. The hypothesis is accepted if $t > t(f-2, c)$.

The test is used in two ways:

(A) At the beginning of a stage V_i is computed for all subscripts, $i = 1, 2, \ldots n-1$. The largest positive V_i identifies the key variable which should be tested for entering into the regression. The quantity $r_{nk.q}$ is computed using equation (12) and the t test described above is performed. If $t > t(f-2, c)$ the variable X_k is entered into regression by performing the transformation T_k.

(B) The second part of the stage begins by again computing V_i for all i. The negative V_i identify the variables that are not in regression. The negative V_i of smallest magnitude identifies the key variable to test for removal. $r_{nk.q}$ is computed using equation (13). If $t > t\ (f-2, c)$ the correlation is significant and the variable X_k should remain in regression. If $t < t\ (f-2, c)$ the variable can be removed from regression without significantly increasing the variance of residuals. X_k is removed from the regression by applying T_k. The procedure is repeated until all insignificant variables have been removed.

The modification of (A) and (B) above for backward regression is quite simple. Initially the recursion is controlled to proceed all the way forward, yielding the inverse of the covariance matrix. On the way back, after any variable is removed, the determination is made as to whether a variable removed previously has become significant. If not, then the least significant variable in regression is removed, provided again that the resulting variance increase is not significant.

REFERENCES

1. Lotto, G., On the Generation of All Possible Stepwise Combinations, Mathematics of Computation, Vol. 16, 1962.

2. Garside, M. J., The Best Sub-Set in Multiple Regression Analysis, Applied Statistics, Journal of the Royal Statistical Society, Vol. XIV, 1965.

3. Breaux, H. J., The Computation of Firing Tables for Guided Missiles, BRL Report No. 1348, November 1966, Aberdeen Proving Ground, Maryland.

4. Efroymson, M. A., Multiple Regression Analysis, Mathematical Methods for Digital Computers, Edited by Ralston and Wilf, John Wiley and Sons, Inc. 1960.

5. Breaux, H. J., Campbell, L. W., Torrey, J. C., Stepwise Multiple Regression - Statistical Theory and Computer Program Description, BRL Report No. 1330, July 1966, Aberdeen Proving Ground, Maryland

6. Wilcoxin, W. L., Wohlever, J. R., An Improved Stepwise Regression Analysis Procedure, Report No. Y-F015-15-06-513, U. S. Naval Civil Engineering Laboratory, Port Hueneme, California.

7. Dixon, W. J., Biomedical Computer Programs Health Sciences Computing Facility, University of California, Los Angeles, September 1, 1965.

8. Abt, K., Gemmill, G., Herring, T., Shade, R., DA-MRCA: A Fortran IV Program for Multiple Linear Regression, Technical Report No. 2035, U. S. Naval Weapons Laboratory, Dahlgren, Virginia, March 1966.

9. Yule, G. U., Kendall, M. G., An Introduction to the Theory of Statistics, Charles Griffin and Company, London, 1940.

10. Graybill, F. A., An Introduction to Linear Statistical Models, Vol. I, McGraw Hill Book Company, Inc., 1961.

ESTIMATION OF ERROR RATES IN DISCRIMINANT ANALYSIS*

Peter A. Lachenbruch and M. Ray Mickey
University of North Carolina
Chapel Hill, North Carolina and
University of California, Los Angeles, California

ABSTRACT. Several methods of estimating error rates in Discriminant Analysis are evaluated by sampling methods. Multivariate normal samples are generated on a computer which have various true probabilities of misclassification for different combinations of sample sizes and different numbers of parameters. The two methods in most common use are found to be significantly poorer than some new methods that are proposed.

*This article is to appear in Technometrics.

SOME STATISTICAL APPLICATIONS IN THE TESTING OF MILITARY VEHICLE RUBBER COMPONENTS*

Emil H. Jebe
Willow Run Laboratories
The University of Michigan, Ann Arbor, Michigan

SUMMARY. This paper utilizes the results of four test programs for rubber components of military vehicles to illustrate a variety of statistical applications. Two of the programs were concerned with the testing of rubber bushings, an element of the track for track-laying vehicles. A third program was conducted to evaluate experimental types of track pads while the fourth example discussed reliability evaluation for track pads and track shoes.

Two of the test programs were based on experimental designs suggested by the author while the other two may be described as:

(1) A factorial arrangement for two factors with missing treatment combinations.

(2) A "road test" without controls or any basis for comparative evaluation.

The statistical applications described for these test programs include the following:

a. Unweighted least squares analysis,

b. Orthogonal polynomials for unequal spacing of a factor,

c. Use of the Kronecker or Direct Product of matrices to form the Contrast or Design Matrix,

d. Weighted Least Squares analysis,

e. Use of a single replicate with confounding in a 3x3x2x2 experiment for four factors,

f. Estimation of experimental error by a number of techniques, e.g., regression residuals, Half Normal Plot, etc.,

g. Use of "uniformity trial" analyses of data from previous tests to design a new experiment,

*Willow Run Laboratories, Project 07312, Institute of Science and Technology, The University of Michigan. Prepared under Contract No. DA-20-113-AMC-05927(T) with USATAC, Warren, Michigan. Revised 10 February 1967.

h. Reliability estimation based on
   (1) The binomial distribution, and
   (2) Johns and Lieberman (1966) (Technometrics 8, 135, February issue).

INTRODUCTION. Since 1962 the author has had a unique opportunity to participate in a number of investigations of rubber products for military applications. These studies have been conducted by engineers of the Components Research and Development Laboratories (CRDL), Research and Engineering Directorate, USATAC. My participation has been through several contracts between The University of Michigan and USATAC. Among the types of products investigated have been bushings, pads, shoes and tires. The latter needs no definition, but the other three are components or elements of the track for our tracked vehicles, e.g., tanks and personnel carriers.

A few words of non-military explanation may be helpful for these components. The rubber bushing is bonded to a track link pin. A close fitting metal tube is squeezed over the rubber bushings which are bonded in clusters of 2, 3 or more on the pin. This assembly is then inserted into a cylindrical opening in the track shoe. Addition of center guides and end connectors to a group of shoes makes possible the assembly of a complete track. The rubber bushing is a key element in this complete assembly in that it provides a non-lubricated bearing and a load taking element such that the vehicle can travel at high speed without prohibitive noise. Another key part of the track is the friction and load bearing surface between the vehicle and the road. The outer face of the track shoe provides this surface. Again this face of the shoe is made of rubber but it may be provided in two ways. One way is to bond and mold rubber to the desired shape directly on the steel surface of the track shoe. Another way is to make shoe pads of desired shape and bolt them to the track shoe. The pad is made by bonding rubber on a metal plate with welded bolt attached.

As the author understands the situation, polymer science and rubber technology are not yet able to predict reliably the outcome of many military applications. The outcome of interest is durability or life of the component. Hence, various laboratory and field tests need to be undertaken to investigate the suitability and durability of specific applications. Our participation in these tests has comprised:

(1) Analysis of laboratory tests (without an experimental design imposed),

(2) Design of experiments for laboratory and field tests,

(3) Analysis of previous field tests to obtain information for designing new field tests,

(4) Analysis of results from designed experiments,

(5) Estimation of reliability from road test results,

In presenting this paper the author wishes to acknowledge the contributions of his colleagues and co-workers, R. A. King and J. W. Curtis. Further, the strong support, encouragement and active interest of USATAC personnel has made it possible to present this report*.

Least Squares Analysis of a 6x3 Factorial Arrangement for Rubber Bushings with Missing Treatment Combinations. The first problem presented to me concerned the analysis of results of fatigue testing a large number of rubber bushings on a laboratory test machine. This machine is designed to simulate the actual field applications of the bushings. Adjustments of the machine permit variations of (1) the radial load (in psi) on the bushing, (2) the angle of torsional twist (plus or minus in degrees), and (3) the cycling rate for the selected load and angle. During fatigue testing the rubber deteriorates so that the load squeezed the bushing and permits a carefully positioned microswitch to close and stop the machine. A counter mounted on the machine permits recording the torsional cycles to failure at the time the switch closes.

Engineers charged with analysis of these data on cycles to failure were disturbed or baffled by the tremendous spread or variability of the results. Further, plotting of average results showed a non-linear response (Figures 1 and 2) which made prediction appear extremely hazardous [1] **. Table 1 indicates the variability for two groups of tests. Table 2 provides a general summary of these results.

In approaching the analysis of these data, one found that no experimental design had been imposed on the test sequence. Although it appeared

*In this regard the author wishes to mention Messrs. P. L. Goud, C. Banton, C. D. Rose, F. Spencer, E. Kvet, R. Westerman, and Miss C. Cicillini. Statements and opinions expressed in this paper, however, are those of the author and do not express USATAC position or policy. The author also wishes to express his appreciation for the comments of Professor H. B. Mann, Army Mathematics Research Center, Univ. of Wisconsin, made after the presentation of the paper on 19 October 1966.

**Numbers in brackets refer to references. These Figures 1 and 2 are reproduced from Figures 6 and 8 of Reference 1.

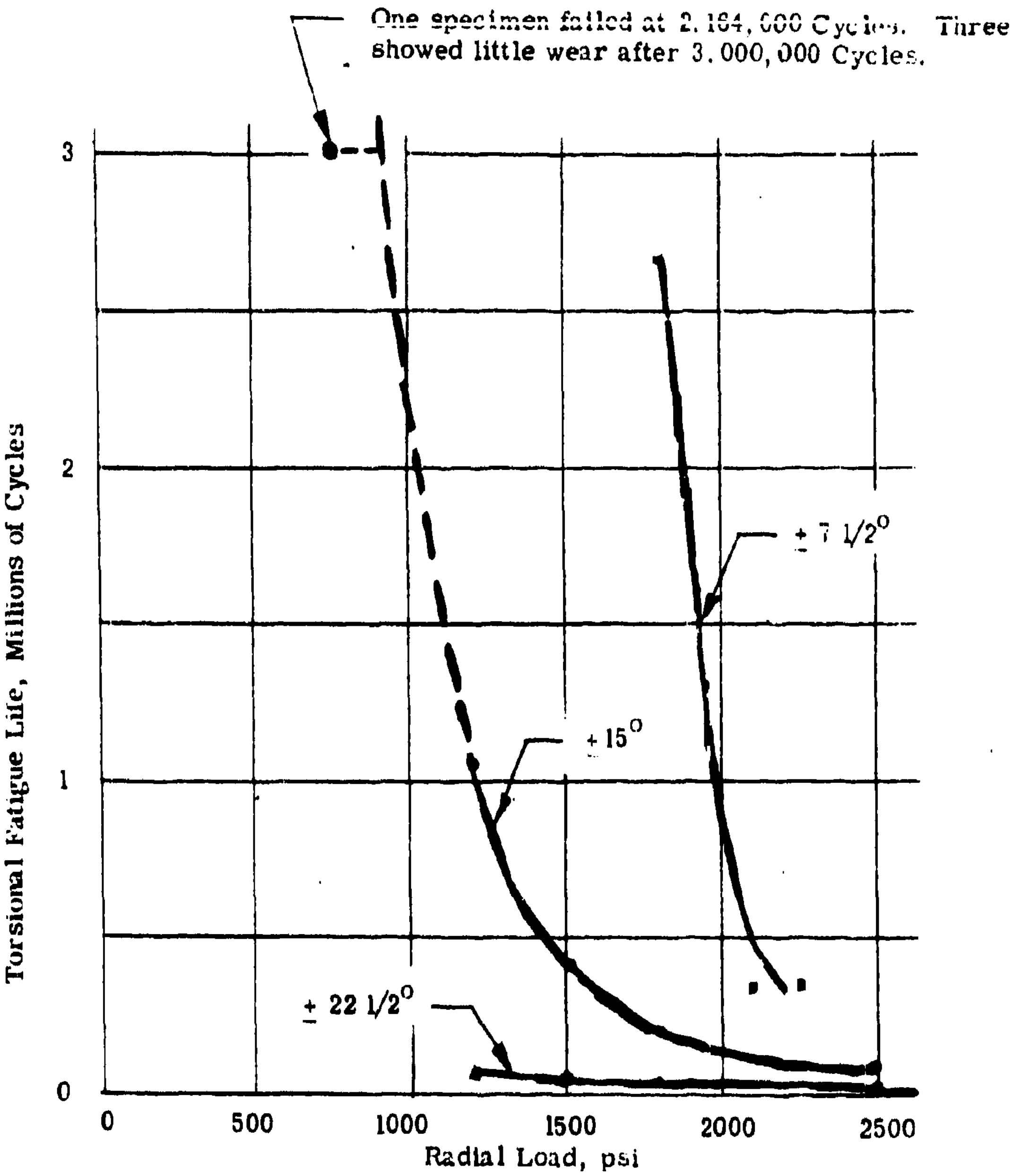


Figure 1. Fatigue Life of Rubber Bushing Track Pins as a Function of Radial Loading, at Different Degrees of Torsional Twist (Figure 6 of [1]).

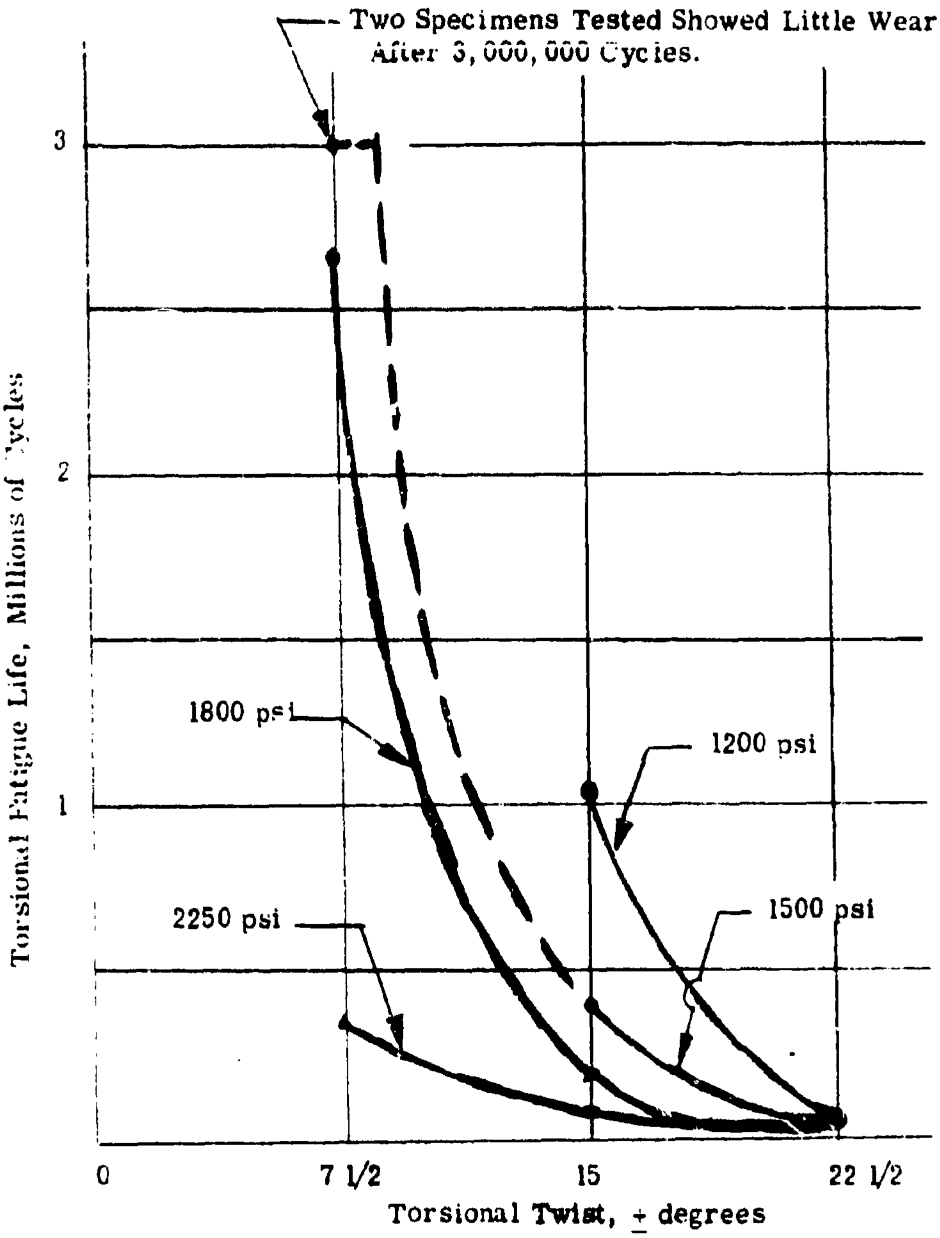


Figure 2. Fatigue Life of Rubber Bushing Track Pins as a Function of Torsional Twist, at Different Levels of Radial Loading. (Figure 8 of [1]).

TABLE 1.
Cycles to Failure for Rubber Bushings Tested at Two Conditions

Test No. 20		Test No. 24	
1	44,900	1	800,000
2	42,700	2	1,326,600
3	34,900	3	1,334,900
4	32,600	4	1,372,100
5	41,500	5	200,000*
6	40,200	6	1,638,800
7	83,500		Load 1950 psi
8	35,000		Angle ± 7.5 degrees

Load 1500 psi

Angle ± 22.5 degrees

Cycling Rates for Both Groups - 255 cpm

Source: Table II of [1].

*Rejected later as an outlier.

TABLE 2. Layout of Rubber Bushing Test Conducted at USATAC, Warren, Michigan, 1962

Radial Load (psi)	Angle of Torsional Twist (Degrees) ± 7.5	± 15.0	± 22.5
1200	$n_{11} = 0$ (no data)	$n_{12} = 6$ $\bar{C} = 1052$* $R = 423$** $\bar{y} = 30175$ $s^2 = 0.00505$	$n_{13} = 8$ $\bar{C} = 56.6$ $R = 40$ $\bar{y} = 1.7373$ $s^2 = 0.01503$
1500	$n_{21} = 0$ (no data)	$n_{22} = 10$ $\bar{C} = 403$ $R = 127$ $\bar{y} = 2.4539$ $s^2 = 0.00922$	$n_{23} = 8$ $\bar{C} = 44.4$ $R = 51$ $\bar{y} = 1.6284$ $s^2 = 0.01643$
1800	$n_{31} = 6$ $\bar{C} = 2661$ $R = 2288$ $\bar{y} = 3.4066$ $s^2 = 0.01906$	$n_{32} = 8$ $\bar{C} = 188$ $R = 164$ $\bar{y} = 2.2586$ $s^2 = 0.01771$	$n_{33} = 8$ $\bar{C} = 25.5$ $R = 8.8$ $\bar{y} = 1.4035$ $s^2 = 0.00222$
1950	$n_{41} = 5^{+}$ $\bar{C} = 1294$ $R = 1439$ $\bar{y} = 3.1006$ $s^2 = 0.01361$	$n_{42} = 0$ (no data)	$n_{43} = 0$ (no data)
2100	$n_{51} = 8$ $\bar{C} = 339$ $R = 317$ $\bar{y} = 2.5021$ $s^2 = 0.02778$	$n_{52} = 0$ (no data)	$n_{53} = 0$ (no data)

Table 2 continued

2250	$n_{61} = 8$	$n_{62} = 6$	$n_{63} = 8$
	$\overline{C} = 342$	$\overline{C} = 76$	$\overline{C} = 14.5$
	$R = 490$	$R = 55$	$R = 4.4$
	$\bar{y} = 2.5014$	$\bar{y} = 1.8697$	$\bar{y} = 1.1582$
	$s^2 = 0.02930$	$s^2 = 0.01296$	$s^2 = 0.00201$

---

*$\overline{C}$ = Average cycles to failure in cell x10^{-3}. Cycles recorded are Torsional Cycles for the Bushing.

**R = Observed range for Cycles to failure in cell x 10^{-3}.

***$\bar{y}_{ij} = \sum_k (\log C_{ijk})/n_{ij}, \quad k = 1, 2, \ldots n_{ij}$

****$s^2 = \sum_k (y_{ijk} - \bar{y})^2/(n_{ij}-1)$

+In Cell 4,1, one test result was rejected as an "outlier".

that a factorial arrangement had been desired for the factors Load (L) and Angle (A), such a program was not completed. Table 2 shows a 6x3 layout but six cells are empty; either no failures were obtained or no tests were conducted. Thus, a least squares analysis became necessary. Next, the question of homogeneity of variance had to be considered. Clearly, differences in dispersion for treatment combinations as shown by Table 1 should be removed. Without previous experience in this field, the writer selected the log normal distribution as a plausible model for the within cell results. Cells 3, 3 and 6, 3 were selected to take a first look at the results of the log transformation. In Table 2 the respective ranges in original scale were 8,800 and 4,400; the s^2 shown for the transform are 0.00202 and 0.00201. Corresponding results for cells 3, 1 and 6, 1 were 2,288,000 and 490,000 for ranges and s^2 of 0.01906 and 0.02930, respectively. Somewhat encouraged by these results the log transformation was accepted* [2].

Plotting the transformed data further showed the usefulness of the transformation. Figures 3 and 4 show the transformed results**. It is seen that the response is approximately linear for either factor for a selected level of the other factor. Some interaction between the factors Load and Angle was indicated by the non-parallelism of the straight lines sketched in the figures.

The next step was selection of a specific regression model and writing out of the X matrix. As a preliminary model, it was assumed that a cell mean, $\bar{y}_{ij}$, could be represented as:

$$\bar{y}_{ij} = \beta_o X_o + \beta_1 L_i + \beta_{11} L_i^2 + \beta_2 A_j + \beta_{22} A_j^2 + \beta_{12}(L_i A_j) + \bar{\epsilon}_{ij}$$

with $i = 1, 2, \ldots, 6$ and $j = 1, 2, 3$, but not over all i, j. The missing cells reduced the data vector, $\bar{y}$, to dimensions 12x1 (the 12 values are shown in Table 2). Small variations in the n_{ij} and variations in s^2 values were ignored at this stage so that the $\bar{\epsilon}_{ij}$ were assumed to have uniform variance.

---

*Later the author became acquainted with some of the relevant literature and concluded that the procedure adopted is reasonably robust against certain alternative models [2].

**Figures 7 and 9 from Reference [1].

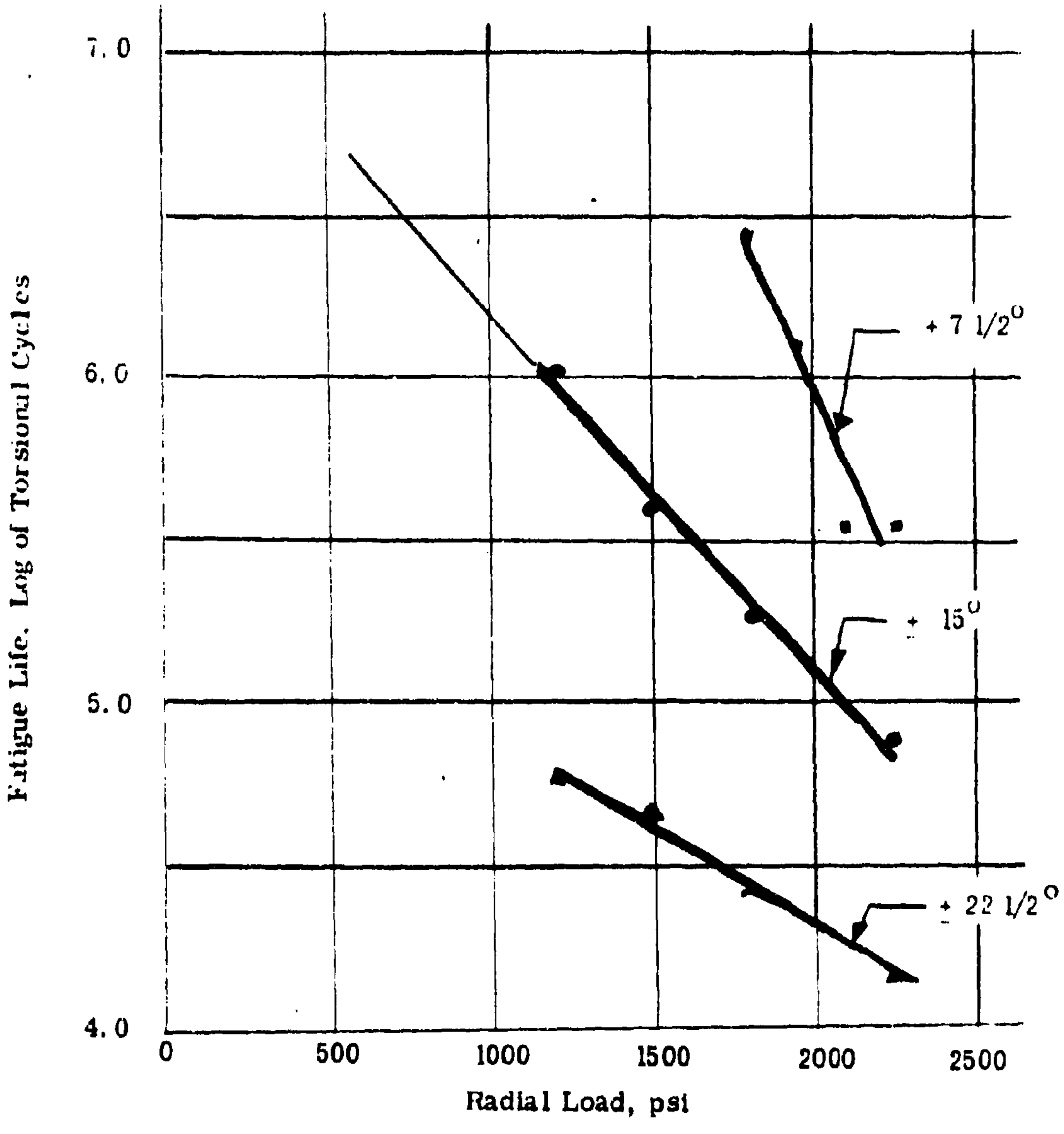


Figure 3. Logarithm of Torsional Fatigue Life as a Function of Radial Load, at Different Degrees of Torsional Twist (Figure 7 of [1]).

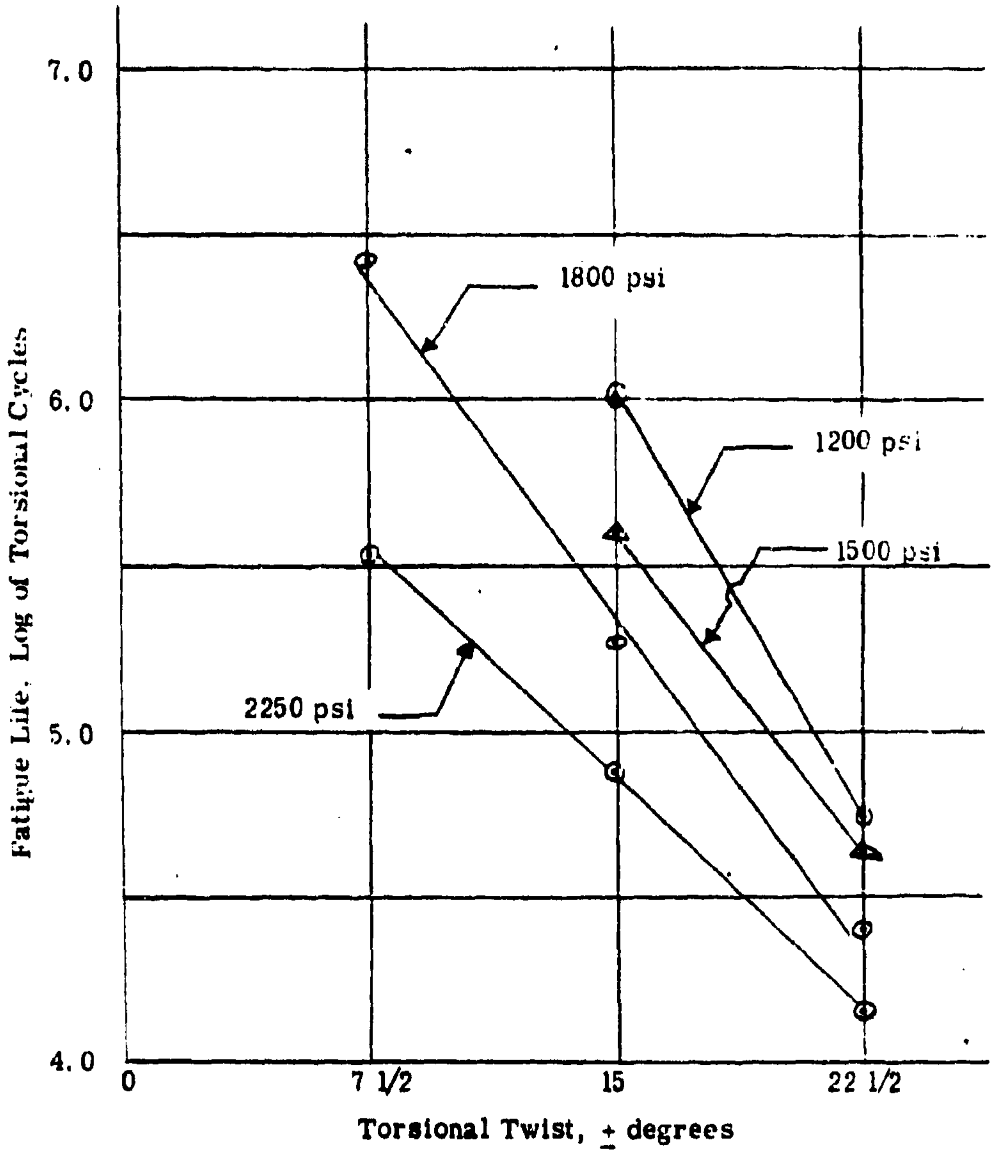


Figure 4. Logarithm of Torsional Fatigue Life as a Function of Torsional Twist, at Different Levels of Radial Loading. (Figure 9 of [1]).

A convenient coding for Angle is seen to be -1, 0, +1 since the spacing was uniform at intervals of ± 7.5 degrees. This coding for A also uses the orthogonal polynomial coefficients for the linear effect of A. Similarly, a convenient coding for Load was found by taking 150 lbs as the unit and centering on 1800 as zero. The coded values became -4, -2, 0, +1, +2, +3. With these values of coded A and L, our first X matrix appears as in Table 3.

TABLE 3. X Matrix for Preliminary Model Fitted to Rubber Bushing Data (Response = Average of Log Cycles to Failure)

X_o	L	L_2	A	A^2	LA
1	0	0	-1	1	0
1	+1	1	-1	1	-1
1	+2	4	-1	1	-2
1	+3	9	-1	1	-3
1	-4	16	0	0	0
1	-2	4	0	0	0
1	0	0	0	0	0
1	+3	9	0	0	0
1	-4	16	+1	1	-4
1	-2	4	+1	1	-2
1	0	0	+1	1	0
1	+3	9	+1	1	+3

From Table 3, the A matrix = X^TX is obtained as shown in Table 4.

TABLE 4. Cross-Product Matrix A for Solution of Normal Equations: AB = G*

X_o	L	L^2	A	A^2	AxL
12	0	72	0	8	-9
0	72	-54	-9	3	15
72	-54	804	15	43	-81
0	-9	15	8	0	3
8	3	43	0	8	-9
-9	15	-81	3	-9	43

From the solution of the normal equations AB = G, which is given by B = CG, where C = A^{-1}, the regression coefficients obtained are shown in Table 7 under First Equation. The summary analysis of variance appears in Table 5.

TABLE 5. Analysis of Variance for Fitting Preliminary Model to Mean of Log of Cycles to Failure

Source of Variation	Degrees of Freedom	Sum of Squares
Total	12	66.4389
Mean (correction term)	1	60.9171
Reduction in Sum of Squares for Regression	5	5.3943
Remainder	6	0.1275
Within Cells (from Table 2)	77	0.18268**
Extension of Table 5 (based upon fitting the Z model with design matrix given in Table 8):		
Add Reduction in S. S.	3	0.07387
Remainder	3	0.05363

*In this compact notation, G = $X^T Y$ where Y is the vector of means given in Table 2 in six rows and three columns.

**The actual within cells sum of squares was 1.096091; a divisor of 6 has been used to place the Remainder SS and Within Cells SS on a comparable basis.

The results presented above are incomplete or inadequate in three respects. First, the regression coefficients are correlated; one would like orthogonal estimates of the effects of Load and Angle and their interaction. Second, the Remainder Mean Square, 0.0212, obtained from Table 5, when compared with the Within Cells Mean Square, 0.00237, indicates a lack of fit for the regression equation used (F value = 8.9 with 6 and 77 degrees of freedom). The Within Cells Mean Square used here may be an underestimate of the proper experimental error due to the lack of randomization in this test program. Third, there is the homogeneity of variance problem already noted in relation to Table 2.

In considering the first point, non-orthogonality of the estimates, one possible approach might be to use a "Missing Value" formula and fill in the six empty cells. Without blocking applied in the experiment, the standard formula for any experimental design could not be used to fill in the missing treatment combinations. Rather naively at the time, I assumed that plausible estimates might be obtained by applying the Randomized Complete Blocks formula for a missing datum to the rows and columns of the two-way layout for the factors Load and Angle. By iterative application of this formula, the six empty cells were filled. Then a second regression equation was obtained. It was found, however, that predictions from this second equation were much worse than for the first equation. For the same 12 observed points, the sum of squares of deviations was 0.33025, about three times the remainder sum of squares of 0.12750 shown in Table 5 for the First Equation.

Why was this decrease in "goodness of fit" observed even though we now had orthogonally estimated regression coefficients (given in Table 7 under the column headed Second Equation)? If there had been only one or two missing treatment combinations, perhaps, the results would have been satisfactory. The consequence of the application of the Randomized Complete Blocks missing value formula to be Load-Angle two-way table was to minimize the Load x Angle interaction. This interaction has 10 degrees of freedom in this Load-Angle table but due to the six empty cells only four degrees of freedom can be estimated. Filling in the empty cells by minimizing these four degrees of freedom apparently had distorted the response surface so that the goodness of fit achieved by the First Equation was destroyed. This view of the problem is supported by a re-examination of Figures 3 and 4, which indicate some interaction that may be largely the Load linear by Angle linear component, and Table 7. In the latter, the values for the b_{AL} (linear by linear) regression coefficient are +0.1038 and +0.0278, respectively, for the equations being compared. This reduction, by a factor of four almost, in this component of interaction regression coefficient appears to be due to the minimization of the overall interaction.

These unsatisfactory results for the second equation posed a dilemma for me. How could more information be extracted from these data? Discussions with Professor Paul Dwyer* brought out two suggestions from him. He did remark that trying to supply one-third of the observations by the missing value approach is "too much like trying to pull yourself up by your own bootstraps". Essentially, his suggestions were to make sub-analyses using subsets of the twelve observed points to form orthogonal structures. The data points used for these analyses are shown in Table 6.

TABLE 6. Data Points Used for Orthogonal Sub-analyses of Rubber Bushing Fatique Life Data

Angle of Torsional Twist

		First Sub-analysis (8 points)	
Load	$\pm$ 7.5	$\pm$ 15.0	$\pm$ 22.5
1200	0	x	x
1500	0	x	x
1800	-	x	x
1950	-	0	0
2100	-	0	0
2250	-	x	x
		Second Sub-analysis (6 points)	
1200	0	-	-
1500	0	-	-
1800	x	x	x
1950	-	0	0
2100	-	0	0
2250	x	x	x

Code: 0 indicates missing values
- datum observed but not used
x datum used for analysis

Observed values appear in Table 2.

*Department of Mathematics and Statistical Research Laboratory, The University of Michigan, Ann Arbor, Michigan.

The results of these sub-analyses are presented in terms of regression coefficients for the "Third Equation" and "Fourth Equation" in Table 7.* A study of Table 7 shows that these sub-analyses support the results for the First Equation. Perhaps, one should be criticized at this point for not presenting standard errors of the regression coefficients. The regression model of the First Equation gave such a good fit and signs of the coefficients were proper so this model was accepted and a report written [1]. Further, extrapolations attempted by the test engineer from these accelerated test results and the regression model gave plausible results.

Personally, I was not yet satisfied and I continued to think about how to improve the analysis. If the data had been complete, one could have worked out the orthogonal polynomial values for the unequal spacing on Radial Load [3]. Forming the Kronecker Product of the Contrast Matrices for Radial Load and Angle of Torsional Twist would then have given an 18 x 18 contract matrix for a complete analysis in terms of single degrees of freedom. From this view, it occurred to me, "Why not proceed in this way to obtain the design matrix for the 12 observed points?" Details are omitted but the resulting matrix is given in Table 8. Here is it seen that additional interaction terms have been added to the model over the First Equation whose design matrix was given in Table 3. If we designate this matrix in Table 8 as Z, then a comparison of Z^TZ with X^TX, given above in Table 4, provides some basis for evaluating the fifth approach to the analysis. The matrix Z^TZ in terms of its first 6 rows and 6 columns is given in Table 9, for making this comparison.

It appears that most of the off-diagonal elements shown in Table 9 have been reduced in relative magnitude. Transformation of Tables 4 and 9 to the correlation matrices shows explicitly that the dependence among predictors has been reduced**. What this means is that use of the Z matrix will give regression coefficients that are less correlated than the coefficients obtained in the first equation. The regression coefficients obtained by use of Z appear in Table 7 under "Fifth Equation". Extension of Table 5 to include the Z model shows an added reduction in sum of squares of 0.07387 with 3 degrees of freedom leaving a new Remainder S. S. of 0.05363 with 3 degrees of freedom.

Setting aside temporarily the inadequacy of goodness of fit noted on page 64, we consider the homogeneity of variance situation. Even though much improved by the logarithmic transformation, it is still apparent in

*Table X of Enclosure 23 [1].

**These correlation matrices have been omitted from this paper.

TABLE 7. List of Regression Coefficients Obtained by the Various Analyses

Term	First Equation	Second Equation	Third Equation	Fourth Equation	Fifth Equation
b_o	2.2494	2.2364	1.9539	2.0997	2.5243
b_L	-0.1759	-0.1526	-0.1231	-0.1710	-0.2950
b_A	-0.9395	-0.8768	-0.9442	-0.8366	-1.2321
b_{LL}	+0.0025	-0.0083	+0.0042	---	+0.0294
b_{AA}	+0.0994	-0.0225	---	+0.0178	+0.1272
b_{AL}	+0.1038	+0.0278	+0.0753	+0.1100	+0.2780
	linear by linear				
b_{LAA}	---	+0.0109	---	---	-0.0682
	linear by quadratic				
b_{LLA}	---	+0.0105	---	-0.0207	-0.0391
	quadratic x linear		-0.0120		
b_{LLAA}	---	+0.0071	---	---	+0.0078
	quadratic x quadratic				

TABLE 8. Design Matrix Based on Forming Orthogonal Polynomials For Load and Angle - Rubber Bushing Experiment

	Lin.	Quad.	Lin.	Quad.	LxL*	LxQ	QxL	QxQ
X_o	X_1	X_{11}	X_2	X_{22}	X_1X_2	X_1X_{22}	$X_{11}X_2$	$X_{11}X_{22}$
	Load		Angle		Interactions			
1	0	-5.667**	-1	+1	0	0	5.667	-5.667
1	1	-3.608	-1	+1	-1	+1	3.608	-3.608
1	2	+0.451	-1	+1	-2	+2	-0.451	+0.451
1	3	+6.510	-1	+1	-3	+3	-6.510	+6.510
1	-4	+6.098	0	-2	0	+8	0	-12.196
1	-2	-3.784	0	-2	0	+4	0	+7.568
1	0	-5.667	0	-2	0	0	0	+11.334
1	+3	+6.510	0	-2	0	-6	0	-13.020
1	-4	+6.098	+1	+1	-4	-4	6.098	+6.098
1	-2	-3.784	+1	+1	-2	-2	-3.784	-3.784
1	0	-5.667	+1	+1	0	0	-5.667	-5.667
1	+3	+6.510	+1	+1	+3	+3	+6.510	+6.510

* LxL = linear by linear

LxQ = linear by quadratic

QxL = quadratic by linear

QxQ = quadratic by quadratic

**See reference 3 for computation of values in this column.

TABLE 9. The Matrix $Z^T Z$ with Last Three Rows and Columns Deleted

12.0	0	4.0	0	0	-9.0
0	72.0	22.24	-9.0	+9.0	+15.0
4.0	22.24	339.71	+5.47	-5.47	-14.12
0	-9.0	+5.47	+8.0	0	+3.0
0	+9.0	-5.47	0	+24.0	-9.0
-9.0	+15.0	-14.12	+3.0	-9.0	43.0

Table 2 that the $\bar{y}_{ij}$ differ considerably in precision. Therefore, a weighted least squares analysis is indicated. Two methods of weighting were used. One method used the individual s^2_{ij} shown in Table 2, i.e., 0.01906, 0.01361, etc., with $w_{ij} = n_{ij}/s^2_{ij}$. The other method pooled sums of squares for the cells where the s^2_{ij} were similar in magnitude, and obtained a set of three values of s^2_p to be used with the n_{ij} of Table 2 to obtain the set of $w_{ij} = n_{ij}/s^2_{p(ij)}$.*

Calculations using these two sets of weights just described were repeated with the Z model already given above. The regression coefficients obtained are presented in Table 7A (Sixth Equation results are for use of the individual s^2_{ij}; Seventh Equation results refers to the pooling method to obtain only three different values of s^2_p used in forming the weights). It is seen that the regression coefficients obtained by the two different weighted least squares analyses are quite similar; differences observed are less than or of the order of the standard errors of the differences. Further comparison of the regression coefficients with those obtained for the Fifth Equation, given in Table 7, reveals some differences that may be judged statistically significant. In terms of practical application for making predictions of bushing fatigue life there may be little to choose between these three equations. In view of the somewhat more reliable weights used to obtain the Seventh Equation results, a statistical choice would lead to this equation, other things being equal.

The results for the weighted regression analyses also permit comment on the goodness of fit issue, which was deferred above. The lower section of Table 7A displays the Residual Mean Squares for the Fifth, Sixth and Seventh Equations. With only three degrees of freedom available for estimating these quantities, no sharp judgments can be made. Qualitatively, the weighted analysis has reduced the residual variation by more than a factor of two, and the goodness of fit has clearly been improved. Yet the ratio of Residual to Within Cells is still large ($P < 0.05$). If one does regard the Within Cells as an underestimate of the experimental error as noted above, then one may conclude that a satisfactory fit has been obtained with

---

*The $s^2_{p(ij)}$ indicates the value of s^2 used for each cell after the pooling operation. The author is indebted to Ralph A. King for this suggestion for obtaining more reliable weights. The sample weights obtained by the two methods appear to be the best surrogates available for the σ^2_{ij} which are unknown.

TABLE 7A. Regression Coefficients Obtained by Weighted Least Squares Analysis Using Model Z and Residual Mean Squares for Three Equations

Regression Coefficient	Sixth Equation	Seventh Equation
b_o	+2.4982	+2.4980
b_L	-0.2830	-0.2805
b_A	-1.2093	-1.2002
b_{LL}	+0.0282	+0.0272
b_{AA}	+0.1217	+0.1140
b_{AL}	+0.2596	+0.2552
b_{LAA}	-0.0618	-0.0601
b_{LLA}	-0.0331	-0.0328
b_{LLAA}	+0.0052	+0.0056

Residual Mean Squares

Source	Degrees of Freedom	Mean Square
Fifth Equation	3	0.01787
Sixth Equation	3	0.00693
Seventh Equation	3	0.00740
Within Cells	77	0.00237

the weighted analysis.* This concludes the story on the first rubber bushing analysis.

Design and Analysis of a 3x3x2x2 Experiment on Rubber Bushings. After completion of the earlier work described above, an opportunity arose to design an experimental program for learning more about rubber bushings. At first, a rather ambitious program was considered which would have involved "experiments with mixtures" (Refer Scheffe [4] and [5] and more recent papers in Technometrics). Suitable bushings prepared from mixtures of natural and synthetic rubbers could not be obtained at the time. Other parameters to be varied in the experiment may be described as Process variables and Test variables. It was desired to retain two levels each of Radial Load and of Angle of Torsional Twist to provide a check on the results for these factors as reported above. These were the Test variables. As Process variables, three levels each of Cure Temperature and Cure Time for production of the bushings were to be tested. Thus, the factorial arrangement became a 3x3x2x2 which requires 36 tests for a single replicate. Two replicates would have required 72 bushings to be tested which I regarded as too large an experiment. After some thought I recommended a single replicate to be carried out in a completely randomized design. At this point the problems began. Complete randomization for the production and testing of the 36 rubber bushings was regarded

*The author believes that some comment on rows 2 and 3 versus row 1 of the lower part of Table 7A may be helpful. Many texts describe weighted regression analysis but none with which I am familiar include a discussion on comparison with the unweighted analysis. With the w_{ij} values defined as explained above, the author was confronted with residual sums of squares for the Sixth and Seventh Equations that apparently provided no basis for comparison with the figure given in Table 5 as 0.05363 with 3 degrees of freedom. Understanding came finally in appreciating the difference in metric. While all three sums of squares represent Euclidean distances in n-space, the scale was different for each. The so-called "unweighted" least squares analysis in reality has a sum of weights equal to n, 12 in this problem. Hence, it was necessary to re-scale the residual sums of squares for the weighted analyses by the factor $n/\Sigma\Sigma w_{ij}$ or $12/\Sigma\Sigma w_{ij}$. These open problems, the choice of scale, the estimation of weights, and more generally, the broader problem of transformation of response to obtain an optimal analysis appear to merit continuing attention.

as impractical, too costly, and too time consuming. My arguments for complete randomization did not convince the engineers that it should be adopted. Then we started to examine possible compromises. Complete randomization for Cure Time, Radial Load and Angle of Torsional Twist could be carried out. Cure Temperature involved bringing the cure press (heated Platten Press) to the desired temperature and holding it there for the ncessary Cure Time. The engineers wanted to reduce the number of times for a press cycle to a minimum. Now, the Cure Temperature could have been made a Main Plot treatment in a Split-Plot design with the 3x2x2 arrangement utilizing 12 split-plots. Replication on Cure Temperature would then have forced the total size of the experiment back to at least 72 bushings.

At this stage it appeared to me that some type of replication for Cure Temperature must be included in the test program. A study of Kempthorne's book provided a possible solution (Reference 6). The 12 split-plot treatments were divided into two main plots of six split-plots each by confounding the Load by Angle (linear x linear) interaction with the main plots*. This contruction of the design required only six press cycles with six bushings cured in each run, two at each of the three Cure Times.

Analysis of the resulting data when this test program had been completed was, of course, considerably more complicated than that outlined by Kempthorne since we imposed an added factor at three levels (refer pp. 351-355 of (6)). Details are given in Reference 7.

Here, I shall only try to describe some of the major features of the analysis and interpretation. The actual layout of the program for the 36 experimental units is given in Table 10. It is to be noted that I insisted on equal spacings for the three levels factors: Cure Temperature at 306, 315 and 324 degrees F. and Cure Time at 15, 30 and 45 minutes. Such equal spacings make the analysis much easier but should not be required for all test programs.

Our first approach to the analysis was to write out a Design Matrix that included the General Mean, Blocks, all main effects and all two-factor interactions. Full column rank was maintained for this matrix by the usual devices, orthogonal polynomials for the contrasts and subtraction of the

*See Table 18.5, p. 350, [6]. The three Cure Temperatures were randomly assigned to the Replicates shown in the table and the main plots in each block for Cure Temperature became the sets shown as "Block 1" and "Block 2" by Kempthorne.

TABLE 10. STRUCTURE AND RANDOMIZATION LAYOUT FOR THE SPLIT-PLOT DESIGN WITH CONFOUNDING. Numbers in parentheses refer to replicates and blocks in Kempthorne. [6]

Block 1

Temperature 306 (2-2)

t	L	A	
1	0	1	13
1	1	0	14
2	0	0	18
2	1	1	17
0	1	1	15
0	0	0	16

Temperature 315 (3-1)

t	L	A	
0	0	1	3
2	1	1	5
1	1	0	1
0	1	0	2
2	0	0	4
1	0	1	6

Temperature 324 (1-2)

t	L	A	
2	1	1	10
1	0	0	12
2	0	0	7
0	1	0	9
0	0	1	8
1	1	1	11

Block 2

Temperature 315 (3-2)

t	L	A	
1	1	1	27
0	1	1	28
0	0	0	25
2	0	1	29
2	1	0	26
1	0	0	30

Temperature 324 (1-1)

t	L	A	
2	0	1	33
0	0	0	35
0	1	1	34
2	1	0	31
1	1	0	36
1	0	1	32

Temperature 306 (2-1)

t	L	A	
1	0	0	19
1	1	1	20
0	1	0	23
0	0	1	24
2	1	0	22
2	0	1	21

Each group of six rubber bushing receives the cure temperature indicated. Two groups at the same temperature form the complete set of 12 = 3 x 2 x 2 for the split-plot treatments. For t, L, A the symbols designate, respectively, 0 = 15 min., 1 = 30 min., 2 = 45 min.; 0 = 180#, 1 = 220#; 0 = $\pm$ 7.5 degrees, 1 = $\pm$ 9 degrees. Note t = Cure Time, L = Radial Load, and $\bar{A}$ = Angle of Torsional Twist. The 4th column in each grouping shows the randomization order for taking the observations over the entire experiment.

column for block 2 from the column for block 1. The resulting matrix was 36 x 21. Least Squares was then applied to estimate these 21 effects or their regression coefficients. Again, the transformation $Y_j = \log C_j$ (where C_j is cycles to Failure) was employed. Original data and logarithms to base 10 appear in Table 11. The regression coefficients obtained are listed in Table 12.

An interpretation of these results is given by quoting three paragraphs from reference 7:

"From the analysis of variance we may deduce that the regression equation comprising blocks, main effects and two-factor interactions provides a good fit to the data. A little over 97% of the total variation about the mean is associated with these effects leaving only about 3% of this total as residual variation.

"From the inverse matrix (obtained in the course of the regression computations) it is found that all of the effects listed in Table 3 are orthogonal (i.e., independent) except Blocks and the Load x Angle interaction. These two effects have a small correlation and are independent of the other 19 effects listed. Further, the diagonal elements of the inverse matrix, c_{ii}, are the elements needed for obtaining the standard errors of the regression coefficients. Specifically, the standard errors are given by $(c_{ii})^{1/2} s_e$, where s_e is the standard deviation of the residuals, given as 0.0908. These standard errors range from about 0.012 to 0.023. Thus, it is found that Radial Load, Angle of Torsional Twist and Cure Time (Linear) which show the largest effects in relation to their sampling errors, should be regarded as real or significant effects. On the other hand, the Cure Temperature (Linear) coefficient is slightly smaller than twice its standard error; thus, it may be regarded as a significant effect. Interestingly enough, two of the interaction coefficients are fairly large in relation to their sampling errors. These are Temperature x Time (Linear) and Time (Linear) x Angle.

"From the signs of the regression coefficients, one may obtain the direction of the effect. Cure Temperature has a positive coefficient so we conclude that a higher temperature, i.e., 324 degrees F., is to be preferred. The quadratic coefficient for temperature is negative which is to be expected. Turning to Cure Time, the coefficient is negative so that a shorter cure time is best, i.e., 15 minutes. Here the quadratic coefficient is positive, but not reliably estimated. The Load and the Angle coefficients are both negative as expected; thus, increasing the level of either shortens the fatigue life."

TABLE 11. FATIGUE LIFE OF RUBBER BUSHINGS. Original Data-Cycles to Failure and Logarithms of these Values.

Bushing No.	Cycles	Logarithm	Bushing No.	Cycles	Logarithm
1*	114,300	5.0580**	19	279,900	5.4470
2	173,300	5.2388	20	36,000	4.5563
3	134,100	5.1274	21	72,600	4.8609
4	246,900	5.3925	22	109,700	5.0402
5	36,600	4.5635	23	154,500	5.1889
6	119,200	5.0763	24	123,300	5.0910
7	194,400	5.2887	25	459,100	5.6619
8	127,000	5.1038	26	59,500	4.7745
9	231,100	5.3638	27	42,000	4.6232
10	32,000	4.5052	28	67,800	4.8312
11	27,700	4.4425	29	79,000	4.8976
12	279,200	5.4459	30	257,100	5.4101
13	60,200	4.7796	31	77,500	4.8893
14	98,800	4.9948	32	132,000	5.1206
15	41,300	4.6160	33	88,700	4.9479
16	416,000	5.6191	34	63,600	4.8035
17	34,700	4.5403	35	722,200	5.8587
18	254,600	5.4059	36	165,700	5.2193

*Treatments applied to each of these bushings were shown in Table 10 - refer corresponding numbers, column 4 of each main plot set.

**Tabulated here with only four decimals in the mantissae. The computer obtained natural logarithms which were converted to common logarithms for ease of interpretation.

TABLE 12. REGRESSION COEFFICIENTS OBTAINED FROM LEAST SQUARES ANALYSIS OF LOGARITHMS OF CYCLES TO FAILURE FOR RUBBER BUSHINGS

	Coefficient*	Name of Effect	
1	5.049560	General Mean	
2	0.019315	Blocks	
3	0.035382	Cure Temperature	- Linear
4	-0.002519		- Quadratic
5	-0.141561	Cure Time	- Linear
6	0.017546		- Quadratic
7	-0.202377	Radial Load - Linear	
8	-0.244741	Angle of Torsional Twist	
9	-0.051940	Temp. (Linear) x time (Linear)	
10	-0.010474	Temp. (Linear) x time (Quad)	
11	0.006171	Temp. (Quad) x time (Linear)	
12	0.005603	Temp. (Quad) x time (Quad)	
13	-0.011462	Temp. (Linear) x Load	
14	0.002001	Temp. (Quad) x Load	
15	-0.002537	time (Linear) x Load	
16	-0.001801	time (Quad) x Load	
17	0.004561	Temp (Linear) x Angle	
18	-0.021679	Temp. (Quad) x Angle	
19	0.036779	time (Linear) x Angle	
20	0.001657	time (Quad) x Angle	
21	0.008773	Load x Angle	

*It is to be noted that the magnitude of these coefficients depends on the scale of the effect used in fitting the regression. Thus, Blocks were coded as -1 and +1; Temperature was coded as one unit = 9 degrees F.; time was coded as one unit = 15 minutes; one unit of Radial Load = 20(20) and one unit of Angle of Torsional Twist = 0.75 degrees.

Further effort in the analysis of these data was devoted to: (1) Estimation of the main plot experimental error for Cure Temperature, and (2) estimation of the split-plot experimental error by various methods. It is true that the regression residual sum of squares 0.123566 with 15 degrees of freedom giving a mean square of 0.008238 is an estimator of experimental variation under appropriate assumptions but it is still a mixture of the main plot and split-plot components just mentioned. Hence, the sentences just quoted may not be valid statements for judging the Cure Temperature effects.

If blocks are ignored, it is possible to estimate each of the 35 individual degree of freedom effects because of the balanced structure for the factorial arrangement. Actually somewhat more is ignored because of the structure of the main plots for Cure Temperature; some of the higher order interactions for Cure Temperature are confounded with blocks. To obtain the sum of squares for each of these 35 effects the full contrast matrix was prepared on the computer by forming the Kronecker or Direct Product of the individual contrast matrices for Temperature, Time, Load and Angle [8] [9]. Our next step was to obtain some estimates of experimental error by applying several techniques that have been suggested in the literature in recent years[10], [11], [12]. Among those used were Daniel's "Half Normal Plot" and the "Gamma Plots" and "smallest ordered contrasts" by Wilk, et al.

While the details about the application of these techniques would be informative and interesting, only the results are shown in Table 13. This table shows the source for the estimate, degrees of freedom (actual or approximate), s^2 and s values, and how or where obtained by a reference. Comparison values from the earlier analysis also are given. Among the problems encountered in making these analyses were the rather large values of the sums of squares associated with certain 3 and 4 factor interactions. No satisfactory explanation has been found for such results*.

Returning to the problem of improving the assessment of the Cure Temperature effects, the analysis of variance shown in Table 14 was prepared.

From this Table 14 it could be judged that the levels of Cure Temperature used in this experiment did not affect the Fatigue Life of the Rubber Bushings. The presence of the confounding with Blocks already

*It is now clear to the author that the complete design matrix should have been constructed in order that the matrix product, $X^T X$ (39 x 39) could have been examined for the nature and degree of confounding present.

TABLE 13. Summary of Estimates of Experimental Error

Source	d. f.	s^2	s	Reference
Regression Residuals (appears too large)	15	0.008238	.0908	(7)
Three and Four Factor Interactions	13	0.006628	.0812	(7)
Half Normal Plot	30*	0.0067	.082	(10)
Gamma Plot (conservative value determined)	31*	0.0071	.084	(11)
Smallest Ordered Contrasts	24	0.004736	.0688	(12)
Average of 24 Smallest Ordered Contrasts (optimistic, appears too small)	24	0.002879	.0536	(7)
Comparison Values				
Regression Residuals**	6	0.1275	.3571	(1)
Within Cells**	77	0.0142	0.1192	(1)

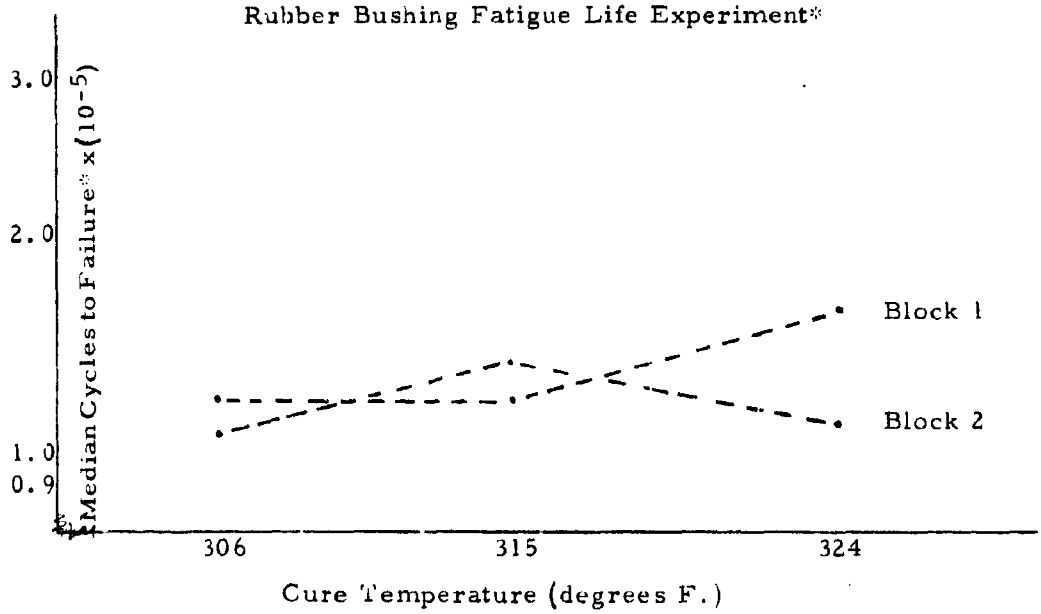
*Approximate.

**Refer Table 5 of this paper. Multiplied up by 6 for comparison with the data above.

TABLE 14. Analysis of Variance for Studying the Cure Temperature Effect

Source of Variation	d. f.	S. S.	Mean Square
Blocks	1	0.012272	.012272
Temperature (Linear)	1	0.030044	.030044
(Quad.)	1	0.000457	.000457
Error (from Temperature by Blocks Interaction)	2	0.037401	.018700
Other Effects (by subtraction from regression analysis)	17	4.261991	xxxx
Remainder (3 and 4 factor interactions)	13	0.086165	.006628

FIGURE 5. Cure Temperature Effects in Blocks 1 and 2. Rubber Bushing Fatigue Life Experiment*



*Points plotted are anti-logs of average logarithms of cycles to failure.

mentioned which enters into this error and the small degrees of freedom raise doubts about such a conclusion. The Cure Temperature results by Blocks are shown in Figure 5. This writer's present opinion is that further experimentation is needed with adequate replication of the Cure Temperature levels and with wider spread, perhaps, 300 to 335 degrees F.

Evaluation of Experimental Types of Track Pads. Attention is directed to another component of the track, the track pad of the Personnel Carrier. We were asked to design a test program for evaluating 14 types of experimental composition pads. Two types of production pads were available as controls. Hence, 16 treatments were to be evaluated. Only one M113 vehicle would be available for carrying out an accelerated road test program. A further restriction was that only seven pads of each of the experimental types could be made available for this program. After several conferences with the interested engineers, the following resume was recorded (quoted from reference [3]):

"Two objective responses could be measured for each individual track pad:

(1) Decrease in thickness (due to wear) of the pad in respect to its height above the grouser shoe to which it is bolted (later referred to as "height loss"), and

(2) Weight loss of the individual pad from its initial weight.

Both of these responses had been measured in previous Army tests with principal dependence placed on the weight loss. Other responses could be considered such as volume loss of the pad from its initial volume and subjective 'scores' or 'ratings' based on chunking and cracking or pieces of material broken off of the pad during use.

"A recent test conducted by the Food Machinery Corporation (FMC), San Jose, California, had utilized the height loss measurement for evaluating the results. Obtaining these measurements had been facilitated by the construction of a special caliper. The level surface of the grouser lug formed the reference for this caliper which was really a type of 'depth gauge'. CRDL constructed a similar gauge for this test program.

"In considering these responses it was pointed out that it should be useful to examine the response data in relation to physical and chemical properties of the pad material compositions for the various type of pads. Examples might be tensile strength, hardness and laboratory abrasion resistance.

"It was expected that total test driving of about 500 miles would be required to reveal differences, if any existed, among the experimental type pads. Test driving would be terminated or test pads would be replaced if wear had progressed to the point that the metal grousers would come into contact with the road surface. Replacement of pads, however, could affect wear of the pads on adjacent shoes. Hence, it was recommended that replacement pads be production type pads whose height above grouser had been gound down or worn to that of the pads on adjacent shoes. Height measurements of pads were to be made: 1) After initial run-in, 2) Each 100 miles thereafter, and 3) At termination. It was suggested also that initial and final weights for individual pads be obtained for all pad types.*"

The real problem encountered in setting up the test program was agreement on the selection of an "experimental unit". Based on their "experience", ATAC engineers tended to favor an experimental unit or plot comprising a cluster of 10 consecutive pads of the same type. The basis for this opinion was that averaging of results from 10 pads would provide a fairly stable average. The left and right sides of the vehicle seemed to form natural blocks for the experimental design. Obviously, with only seven pads on hand for the experimental types, this cluster of 10 could not be obtained. Putting all seven in a cluster would not permit replication.

At this point, it was found that data from previous tests conducted by the Army at several sites was on hand. These data were obtained and analyzed from the "uniformity trial" point of view**. Cluster sizes of 2, 3, 4, 6, 8, 12, 13, and 18 were studied in these analyses with the smaller clusters formed from the larger clusters. It was found that size of cluster did not affect conclusions for any of the previous tests. A peculiar feature of the M113 vehicle added interest to the problem of determining the cluster size; one side of the vehicle has 64 track shoes and the other side has 63.

With the uniformity analyses information available, a cluster size of 4 was established on one side so that 16 x 4 = 64 and on the other side a cluster size of 3 was used with 16 x 3 = 48. The remaining 15 pads on this side were filled in with standard pads and limited supplies of a few other experimental pads. Hence, the experimental design may be described as a Randomized

*Unfortunately, the 14 experimental types of pads were weighed in groups before installation so that data from this experiment do not provide sufficient information for correlation of height loss and weight loss.

**A more sophisticated approach would have calculated auto-correlations of weight losses for adjacent pads and pads separated by 1, 2, 3 or more up to K-2 pads, where K was the number used in a cluster.

Complete Block for 16 treatments in two replicates with each track of the vehicle forming a block. While normal driving provides a natural randomization on the wear of the pads, a different randomization was used for the treatments on each track.

No difficulties were experienced in carrying out the 500 mile accelerated test program. There were some doubts in my mind about the scheduled 500 miles being sufficient to show up differences among the treatments since other Army tests had comprised total mileages of 1000, 1500, or 2000 miles. The program could not be extended for this test, however, because the vehicle had to be returned to another agency.

With respect to analysis, we followed the suggestions of George Box (1950) on analysis of growth and wear curves [14]. Differences between successive measurements of pad height above grouser were formed, e.g., $H_i - H_{i-1}$ for i = 1 through 5. These differences appeared to be reasonably distributed so the analysis of variance was applied directly to these differences without transformation. In order to help understand the analysis of variance (Table 15), Figure 6 explains the structural arrangement of these differences.

It will be noted that Table 15 shows only 14 degrees of freedom for treatments (Types of Pads); this happened because only one Control Type was available when the driving program was started. This one control Type was duplicated on each track. To simplify the computer programming, only one cluster in each block was used for the Control Type of Pad. A list of means for the Pad Types in each block and overall is given in Table 16.

Now what about interpretation? Statistically, I was quite pleased with these results. We used Multiple Comparisons Procedures to group the experimental pad types into significantly different groups [19]. Our next step was to try to relate the values of these means to other available physical and chemical data on the experimental Types of Pads. Unfortunately, no significant regressions could be obtained. Hence, it is my personal opinion that there is still room for a lot of research on military track pads in order that we can find the determinants of longer life for this element of the vehicle track.

Reliability Analysis for Track Components. As a final example in this paper I shall present briefly some attempts at reliability evaluation. Recently, the Army has conducted some "road Testing" of a new track design for the tank. This new design comprised a track made up with track shoes whose grouser shape was formed by replaceable pads. Road testing was conducted at three sites using three vehicles at one site and two each at the other sites, or a total of seven vehicles. Total distance driven

FIGURE 6. Structural Arrangement of Differences for Analysis of Height Loss of Track Pads

1. Block = Side of Vehicle

2. Treatments (16) randomized over plots in each block.

3. Plot = Cluster of 3 or 4 consecutive pads (3 on left side; 4 on right side).

4. Split-Plot = Unit of travel (100 miles) (labeled as Period in Analysis of Variance).

5. Individual Pad = Subsampling unit within the split-plot.

6. Height of Pad recorded at 0, 100, 200, 300, 400, 500 miles.

7. Differences taken for each pad for each increment of wear (100 miles) giving a total of 7 x 16 x 5 = 560 differences. Difference = Height Loss.

TABLE 15. Analysis of Variance of Track Pad Test Results: 15 Pad Types Mounted on Both Tracks of M113 Personnel Carrier with Height Loss Measured for 7 Pads of Each Type over 5 Periods of 100 Miles Each*

Source of Variation	Degrees of Freedom	Sum of Squares**	Mean Square	F**** Ratios
Uncorrected Total	525	299566	—	—
Correction Term for Overall Mean	1	209121	209121	—
Sides of Vehicle	1	392	392	
Types of Pads	14	50696	3621	~ 91
Error (a)	14	555	39.6	
Periods	4	7655	1914	~23
Types x Periods	56	7681	137	~ 1.7
Periods x Sides	4	250	62	
Types x Periods x Sides	56	4675	83	
Error (b)***	60	4925	82.1	
Pads Within Types				
Left Side	30	2406	80	
Right Side	45	2203	49	
Pads Within Types x Periods				
Left Side	120	9588	80	
Right Side	180	4344	24	

*Variable analyzed is Height Loss for a single Period of each individual pad within a Type, i.e., 7 pads for each Type x 15 Types x 5 Periods gives 525 measurements of Height Loss. Units are the same as in Table 16, but squared here.

**Addition may not check in this column because of rounding to whole numbers in sums of squares for each source of variation.

***Error (b) is sum of two preceding sources which appear to be homogeneous.

****F ratios are computed using Error (a) for Sides and Types, and Error (b) for Periods and Types x Periods. This procedure conforms to the split-plot structure of the experimental plan with Periods considered as the split-plot treatments.

TABLE 16. Average Height Losses of Track Pad Types for 500 Mile Test Program* Based on 3 Pads on Left Side and 4 Pads on Right Side. Averages are in Thousandths of an Inch: 14.63 = .01463"

Type No.	CRDL Code	Left Side	Right Side	Overall	Rank**
1	S131C2F2	16.33	13.35	14.63	6
2	Z138	17.67	17.75	17.71	11
3	Z138C	13.87	13.60	13.71	3
4	Z138C1F	14.60	16.95	15.94	8
5	Z138CF1	14.53	11.00	12.51	2
6	Z138F2	16.07	12.55	14.06	4
7	Z121F	30.33	28.10	29.06	17
8	Z140	20.13	18.35	19.11	12
9	S131C2F22	52.53	46.40	49.03	19
10	Z138C2	15.87	14.05	14.83	7
11	Z138C3DF3	16.93	16.80	16.86	10
12	Z116CF2	35.87	33.25	34.37	18
13	S131C2F2BD	16.73	15.65	16.11	9
14	Z128CF	12.93	9.35	10.89	1
15	Comm'l SBR	19.93	21.00	20.54	14
16	A-0	13.60 (2)***	15.20 (2)***	14.40	5
17	C-0	21.80 (4)	17.90 (2)	20.50	13
18	C-10	25.12 (5)	—	25.12	16
15L	Comm'l SBR****	20.71 (7)	—	20.71	15

*Averages are calculated on a per Period basis. Multiplication by 5 gives Average Total Height Loss. Standard deviation of a Type Average = $[\text{Error (a)}/(35)]^{1/2} = (39.6/35)^{1/2} = 1.063$ units (refer Table 15).

**Rank is in order from lowest to highest Height Loss.

***Numbers of Pads averaged in last four rows.

****These added seven pads for Comm'l SBR were omitted from Type 15 in the variance analysis to simplify the programming and weighting of data problems. Comm'l SBR is designated as the principal control. Type C-10 is considered a secondary control.

TABLE 17. Order Statistics for Pad Set Miles with Some Summary Statistics

Line No.	Order Statistic No.	Miles Completed	Summary Statistics
1		389	First three values not counted
2		904	
3		939	
4	1	1667	M_1, smallest value
5	2	1994	
6	3	2096	
7	4	2250	Q_1, first quartile
8	5	2250	
9	6	2300	
10	7	2338	
11	8	2430	Median
12	9	2496	
13	10	2570	
14	11	2628	
15	12	2677	Q_3, third quartile
16	13	2706	
17	14	2750	
18	n = 15	2813	M_{15}, largest value
	Total	35,965	Mean about 2400 miles

TABLE 18. Reliability Estimation for Pad Sets Under the Test Conditions of the Program

	Estimate for m_o = 2000 miles	Lower Confidence Limit (γ = 0.90)
A. Binomial	0.933	0.764
B. Johns and Lieberman $M_o = 0$	0.914	0.811
C. Three parameter approximation $\hat{M}_o = 1667$	0.955	0.881

exceeded 35,000 miles. In covering this distance, complete pad sets were replaced when worn out, and some individual track shoes were replaced although no complete track set was replaced or judged completely worn out.

After much thought about the problem, it appeared to us that a reliability statement might be made about the pad sets and for the first track shoe replacement on each vehicle. Table 17 shows the order statistics data for the 18 pad sets used [15]. Three short mileages were omitted from our analysis for obvious reasons and one value of 1994 miles was counted as a "success" in attaining 2,000 miles. It is to be noted that our reliability estimates apply to the conditions of the road test and not to Army use in general. In Table 18 the results are shown for three approaches to the problem [15] *. Johns and Lieberman refer to their recent Technometrics paper [16]. The binomial result is for 14/15 = 0.933 and use of a binomial table [17]. The third result is a crude approximation that I obtained from the Johns and Lieberman approach.

Information about the first track shoe replacements is given in Table 19 [15] **. Again, an observation has been omitted in the analysis.

TABLE 19. Mileage at Replacement of First Track Shoe During Road-Testing of New Track Design of Seven Vehicles

Vehicle Number	Replacement Mileage	Number of Shoes Replaced
1	904***	30***
2	2745	1
3	2992	2
1	3000	4
4	3315	1
5	3686	2
6	3925	1
7	4894	1

Median = 3315; Average = 3508

*These Tables 17 and 18 are based on Tables 4.1 and 4.3 of [15].

**This Table 19 is derived from Table 6.2 of [15].

***Datum not used because entire track was thrown; damaged shoes were replaced.

Table 19 shows the mileages arranged as order statistics; vehicle numbers are arbitrary designations. From the lowest value, 2745 miles, and the sample size, n = 7, we may estimate with 50% confidence that 90% of vehicles road-tested under similar accelerated conditions will have their first track shoe replacements after 2745 miles. This result is a non-parametric tolerance limit [18]. If a higher confidence statement is desired, then the tolerance proportion or reliability stated must be lowered. For 90% confidence, the figure becomes 72% first track shoe replacements after 2745 miles, which is a one-sided binomial limit [17]. One would like to apply the Johns and Lieberman technique to these track shoe date but the smallest sample size for which they have worked out their tables is n = 10 [16].

REFERENCES

1. C. D. Rose, "Laboratory Investigation on Fatigue Life of Rubber Bushing Track Pin Assemblies," Report No. 7908, USOTAC, 16 July 1963. (Supplement on "Analysis Procedure" by Emil H. Jebe.)

2. M. Zelen, "Factorial Experiments in Life Testing," Technometrics 1, 269 (1959). Also, see Zelen and Dannemiller, "The Robustness of Life Testing Procedures...", same journal, 3, 29 (1961).

3. D. S. Robson, "A Simple Method for Constructing Orthogonal Polynomials when the Independent Variable is Unequally Spaced," Biometrics 15, 187 (1959).

4. H. Scheffe, "Experiments with Mixtures," Jour. Royal Stat. Soc., Series B, 20, 344 (1958).

5. H. Scheffe, "The Simplex-Centroid Design for Experiments with Mixtures," Jour. Royal Stat. Soc., Series B, 25, 235 (1963).

6. O. Kempthorne, "Design and Analysis of Experiments." J. Wiley and Sons, New York, (1952).

7. E.H. Jebe and R. A. King, "Investigation of Cure Time and Cure Temperature Effects on Fatigue Life of Rubber Bushings for Track Link Pins," Report prepared for Projects 05895 and 07312, Willow Run Laboratories, The University of Michigan under Contract DA-20-018-AMC-0954T with USATAC.

8. P. R. Halmos, "Finite Dimensional Vector Spaces," Princeton University Press, Princeton, 6th Printing (1953) (see Appendix II for discussion of the Kronecker or Direct Product, pp. 170-182).

9. E. H. Jebe and R. O. Bennett, Jr., "The MATRIX MACROS Program," (in process). This set of macros or subroutines includes two special Kronecker Product routines.

10. C. Daniel, "Use of Half Normal Plots in Interpreting Factorial Two-Level Experiments," Technometrics 1, 311 (1959).

11. M. B. Wilk, et al, "Probability Plots for the Gamma Distribution," Technometrics 4, 1 (1962).

12. M. B. Wilk, et al, "Estimation of Error Variance from Smallest Ordered Contrasts," Journ. Amer. Stat. Assn., 58, 152 (1963).

13. E. H. Jebe and R. A. King and J. W. Curtis, "T-130 Track Pad Test Program: Experimental Design and Analysis of Test Results," Informal report prepared under Projects 05895 and 07312, Willow Run Laboratories, The University of Michigan under Contract No. DA-20-113-AMC-05927(T).

14. G. E. P. Box, "Problems in the Analysis of Growth and Wear Curves," Biometrics 6, 362 (1950).

15. Report to: OCRDL, USATAC, Subject: "Preliminary Report on Examination of Data from the Track Improvement Program," by E. H. Jebe, J. W. Curtis and R. A. King, 28 June 1966, Prepared under Project 07312, Willow Run Laboratories, The University of Michigan under Contract No. DA-20-113-AMC-05927(T).

16. M. V. Johns and G. J. Lieberman, "An Exact Asymptotically Efficient Confidence Bound for Reliability in the Case of the Weibull Distribution," Technometrics, 8, 135 (1966).

17. J. R. Cooke, M. T. Lee and J. P. Vanderbeck, "Binomial Reliability Table (LCL) for the Binomial Distribution," NAVWEPS Report No. 8090, NOTS TP 3140, China Lake, California, Jan. 1964.

18. D. B. Owens, "Handbook of Statistical Tables," Addison-Wesley, Reading, Mass. (1962).

19. G. W. Snedecor, "Statistical Methods," Ames, Iowa, ISU Press (1956) Refer section 10.6, p. 251.

A STATISTICAL ANALYSIS OF PROVISIONING PROCESSES ON FOUR ARMY MISSILE SYSTEMS

Robert G. Provost
U. S. Army Missile Command, AMSMI-WR
Redstone Arsenal, Alabama

[The author presented a series of slides at the conference. These slides, with the information about each, are reproduced in this article.]

SLIDE 1 - Title slide

2 - Schematic of PDS routes

3 - Station ident

4 - Matrix

5 - Route ident

6 - Bottom of PDS

7 - Blow-up of matrix - 1 cell and title blocks

8 - Figure of head!

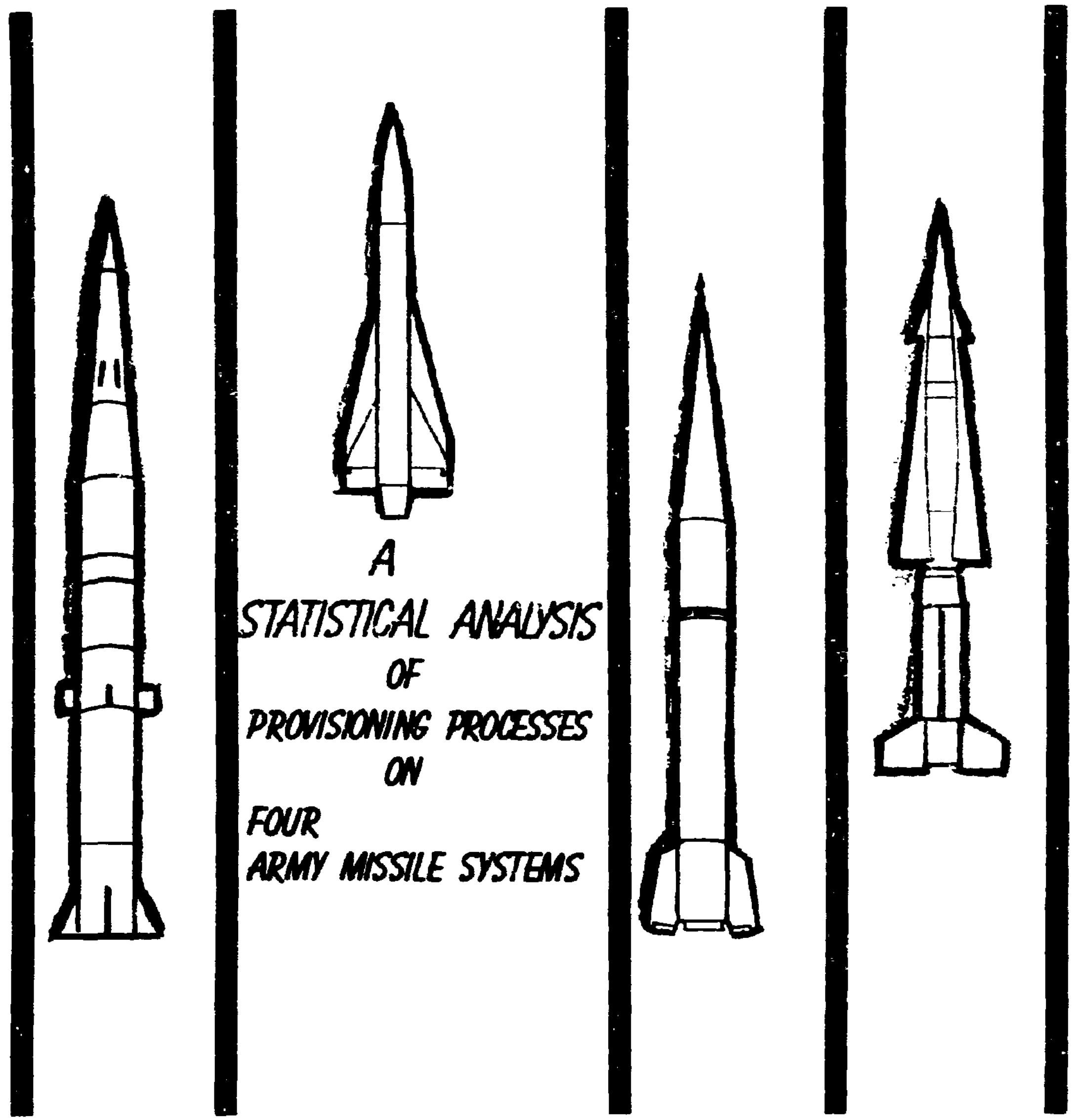
9 - Tukey - Dixon - Snedecor

10 - Matrix w/avg station lengths

11 - 400 PDS sample

12 - Axe head cutting time in half

13 - The end


A
STATISTICAL ANALYSIS
OF
PROVISIONING PROCESSES
ON
FOUR
ARMY MISSILE SYSTEMS

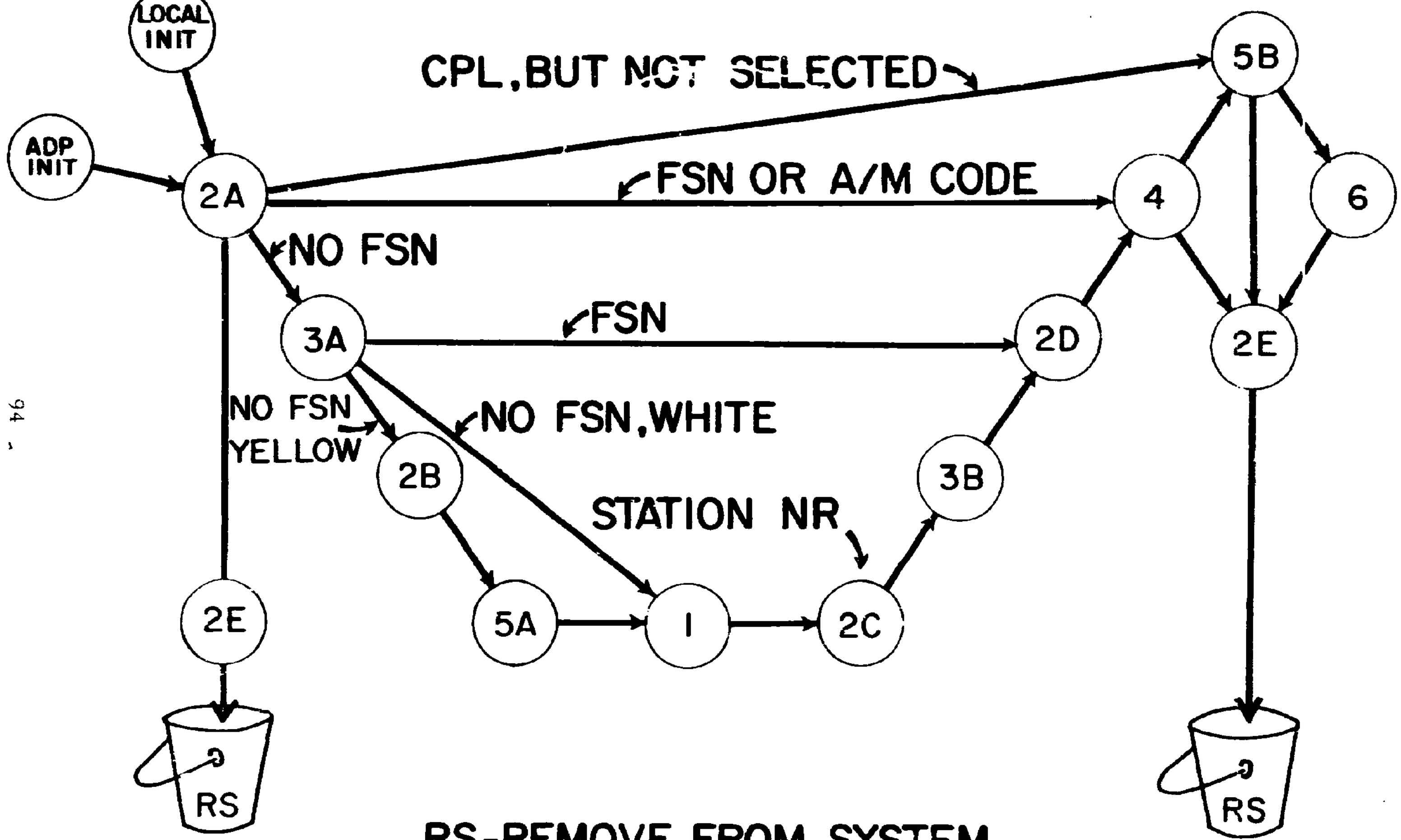
SLIDE 1. Title Slide. As a missile system proceeds from the R&D stage to the production stage, changes will occur as a result of value engineering applications, changes in technology, improved materials and hardware items, preproduction engineering, and the discovery of inadvertent errors. These changes are incorporated through the use of and Engineering Order (an EO). In one stage of the EO, a Provisioning List (PL) is generated which subsequently ends up in our Supply and Maintenance Directorate as a PDS (Provisioning Data Sheet). A Provisioning List contains all the parts needed to support the change, whereas a PDS is a computer-produced sheet for each part listed on the PL. It is used as a worksheet to identify that part within the Federal Cataloging System.

Due to the different types and classes of parts, and to the priority required, a PDS will flow through this portion of the S&M Directorate along different routes. Also due to the lack of different kinds of information, a PDS will take still other routes.

Current regulations provided a 90-day time limit to process a PDS through the S&M Directorate. PDS's which exceeded this time limit were considered delinquent.

Since a PDS represents a single line item in a PL and a PL could contain from one to 1000 or more line items, any PDS which exceeded the time limit caused the entire PL to become delinquent. Management was concerned about these delinquencies and wanted to know, since each PDS flowed through various routes and stations, what the average length of each station and route actually was, and could the 90-day time limit be reduced.

PDS ROUTES



RS-REMOVE FROM SYSTEM

2

SLIDE 2. Schematic of PDS routes. The first step involved the determination of the various routes. This slide is a schematic of these routes. Six basic stations were identified: station number 1 through station number 6. Each station performs one or more functions during the flow cycle of the PDS. These are identified by the alpha characters after each station number. For example, Station 4 has but a single function whereas Station 2E indicates that this is the fifth function performed by that station. By counting all possible combinations of routes in this schematic, one can easily determine that there are fifteen different routes for an ADP initiation, and the same number of routes for a local initiation. This is true for a single type of a PDS, but due to priorities, there are three types of PDS's to consider. These are colored for easy identification: a white PDS for routine or low priority items, a yellow PDS for high priority items, and a green PDS for emergency items. Adding these various types of PDS's into the schematic, a maximum of forty-five different routes is now possible for each initiation. To complicate matters still further, a green PDS, used for emergencies, is also used as a delinquency flag. Should any PDS remain in this portion of the processing cycle beyond a specified period of time, a delinquent green PDS is initiated locally. This is rushed through the system until it reaches that station in which the original PDS is bogged down. Since this delinquent green PDS can travel along any route, our total maximum number of possible routes now stands at sixty. Add to these routes the fact that occasionally a yellow PDS, during the processing cycle, can be downgraded to a lower priority, that is, downgraded to the status of a white PDS. The processors when confronted with this action would hand stamp the yellow PDS not with a "downgrade" stamp, but with one called "PEPSODENT" -- you wonder where the yellow went!

STATION IDENTIFICATIONS & FUNCTIONS

STATIONS	IDENTIFICATION	FUNCTIONS
1	INDUSTRIAL	VALIDATE PART NUMBERS & CHECK AGAINST DRAWINGS FOR ACCURACY
2	MAINT ENG'G	MAINTAINS CONTROL. DISTRIBUTES PDS INTO SYSTEM, VERIFIES MAINTENANCE DATA, ASSIGNS PACK DATA, REVIEWS FOR COMPLETENESS AND ACCURACY, AND REMOVES FROM SYSTEM.
3	FED CATALOGING	OBTAINS FSN FROM PROPER SOURCE & ASSIGNS TO PART
4	PUBLICATIONS	EXTRACTS DATA FROM PDS FOR INCLUSION IN SUPPLY & TECH MANUALS. UPDATES MASTER FILES.
5	SUPPLY CONTROL	MAKES SUPPLY STUDIES, PREPARES AND SUBMITS REQUISITIONS AND SETS UP PURCHASING OF REQUIRED PARTS
6	CATALOGING	BROADCASTS TO DEPOTS ON MICOM MANAGED ITEMS ONLY

3

SLIDE 3. Station ident. That Pepsodent action was generally performed in Station 2. Perhaps at this point we should examine the various functions of each station. This slide depicts each of the stations and identifies their function(s).

Station 1, our Industrial station, is actually located in the Procurement and Production Directorate. This station performs the validation of each part number to facilitate the finding of the proper FSN for that part. This station also checks the part against the drawings for accuracy, and, as need be, obtains new drawings as required.

Station 2, Maintenance Engineering, is the control station for this procedure. Upon receipt from the computer they review each PDS for completeness and accuracy, distribute them into system, verify maintenance data and assign pack data as required, and after all the work has been accomplished remove each completed PDS from system. They also prepare the delinquent green PDS, whenever any PDS is not removed from the system on time.

Station 3 is Federal Cataloging. It is this station which obtains the FSN for each PDS from proper sources, either locally or from outside agencies, and assigns this Federal Stock Number to each part on receipt.

Station 4, the Publications station, extracts the pertinent data from each PDS for inclusion in Supply and Technical Manuals. They also update the master files.

Station 5, our Supply Control station, which makes the necessary supply studies, prepares and submits requisitions and sets up the purchasing of required parts.

Station 6, Cataloging. This function of cataloging involves the advance notification to our supply depots of these various parts that are coming through our system.

It is at this point I should mention that all PDS's do not lack an FSN. Some PDS's do not require an FSN since, for example, the part is fabricated or modified in place. Other PDS's, the bulk in fact, had the proper FSN located by the computer when the PL was converted into the various types of PDS's.

MATRIX FOR PDS SAMPLES

STATIONS / ROUTES	A_1 MAINT ENGINEERING (PROVISIONING) STA2A	A_2 CATALOG STA3A	A_3 MAINT ENGINEERING STA2B	A_4 SUPPLY CONTROL STA5A	A_5 INDUSTRIAL STA1	A_6 MAINT ENGINEERING (PROVISIONING) STA2C	A_7 CATALOG STA3B	A_8 MAINT ENGINEERING (PROVISIONING) STA2D	A_9 PUBLICATION STA4	A_{10} SUPPLY CONTROL STA5B	A_{11} BROADCAST STA6
	K P H S	K P H S	K P H S	K P H S	K P H S	K P H S	K P H S	K P H S	K P H S	K P H S	K P H S
B_1 W											
B_2 W [illegible]											
B_3 [illegible]											
B_4 [illegible]											
B_5 [illegible]											
B_6 [illegible]											
B_7 W 3a											
B_8 W 3aA											
B_9 W 3b											
B_{10} W 3bA											
B_{11} G 3											
B_{12} G 3A											
B_{13} Y 4											
B_{14} Y 4A											

W = WHITE PDS, ROUTINE
Y = YELLOW PDS, HI PRIORITY
G = GREEN PDS, EMERGENCY

K = HAWK H = HERCULES
P = PERSHING S = SERGEANT

4

SLIDE 4. Matrix. Returning to our sixty possible routes for each initiation, each type PDS was carefully examined in relation to all of its possible routes through this portion of the S&M Directorate and instead of some sixty possible routes, a total of twelve basic routes emerged. Four of these basic routes were eliminated for reasons such as: the item was fabricated, not purchased; infrequent use, like once in two years; and, sundry other reasons leaving us with eight basic routes. These routes divided into pairs of routes, with each pair having one broadcasting function: Either the PDS was broadcasted (because it represented a MICOM-managed item), or it was not broadcasted. These routes were further subdivided, by segregating the three colors of PDS's into sub-pairs of routes within the basic routes. An additional pair of routes was developed after discussion with the personnel of one station because either the PDS could be handled in a relatively short period of time (30 days or less) or an extremely long period of time (90 days or more). Since we knew the reason for this long period of time the data collected for this pair of routes were subsequently omitted. This slide shows the final configuration as well as the matrix developed to handle this problem. The columns which identify the stations are coded A_1 through A_{11} for subsequent use in a computer program. The rows which identify the routes are coded B_1 through B_{14} for the same purpose. Routes B_9 and B_{10}, which represent the long cycle time of one station, were subsequently dropped for the reason stated before. The intersection of a row and a column is designated as a cell. Each cell is divided into four columns, one column for each missile system under consideration. The shaded areas represent those stations which are not in that specific route. Thus Route B_1 only contains four stations, A_1, A_9, A_{10}, and A_{11}, whereas Route B_{13} contains all eleven stations, A_1 through A_{11}.

ROUTE	DEFINITION			
B_1	WHITE PDS	WITH FSN;	BROADCASTED	
B_2	WHITE PDS	WITH FSN;	NOT BROADCASTED	
B_3	WHITE PDS	NO FSN;	BROADCASTED	
B_4	WHITE PDS	NO FSN;	NOT BROADCASTED	
B_4	GREEN PDS	NO FSN;	BROADCASTED	
B_5	GREEN PDS	NO FSN;	NOT BROADCASTED	
B_6	WHITE PDS	NO FSN;	BROADCASTED	SHORT CYCLE (30 DAYS OR LESS)
B_7	WHITE PDS	NO FSN;	NOT BROADCASTED	SHORT CYCLE (30 DAYS OR LESS)
B_8	WHITE PDS	NO FSN;	BROADCASTED	LONG CYCLE (90 DAYS OR MORE)
B_9	WHITE PDS	NO FSN;	NOT BROADCASTED	LONG CYCLE (90 DAYS OR MORE)
B_{10}	GREEN PDS	NO FSN;	BROADCASTED	
B_{11}	GREEN PDS	NO FSN;	NOT BROADCASTED	
B_{13}	YELLOW PDS	NO FSN;	BROADCASTED	
B_{14}	YELLOW PDS	NO FSN;	NOT BROADCASTED	

WHITE PDS----ROUTINE *YELLOW PDS----PRIORITY* *GREEN PDS----EMERGENCY OR DELINQUENT*

5

SLIDE 5. Route ident. This slide identifies each pair of routes.

Routes B_1 and B_2 are for white PDS's which have an FSN.

Routes B_3 and B_4 are for white PDS's which do not have an FSN.

Routes B_5 and B_6 are for green PDS's which do not have an FSN.

Routes B_7 and B_8 are for white PDS's without an FSN and require validation. This validation is accomplished in 30 days or less.

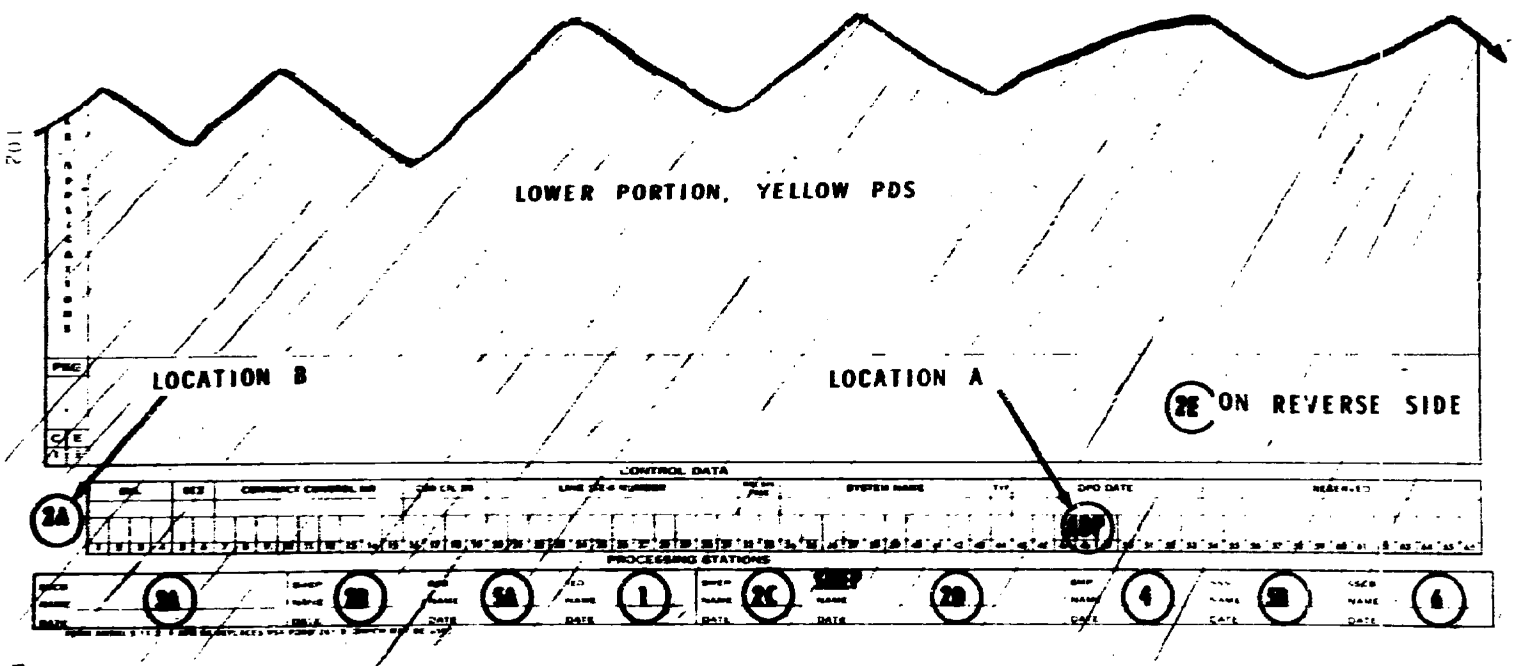
Routes B_9 and B_{10} are also for white PDS's without an FSN which require validation but this validation required 90 days or more to accomplish. For the reason mentioned before, these routes were removed from the analysis.

Route B_{11} and B_{12} are for green PDS's without an FSN. These differ from routes B_5 and B_6 in that the FSN was not immediately located and requires outside agencies assistance.

Routes B_{13} and B_{14} are for yellow PDS's without an FSN.

The difference between each pair of routes is that the first route of a pair contains a broadcasting function.

Are there any questions up to this point?



LOWER PORTION, YELLOW PDS
LOCATION B
LOCATION A
2E ON REVERSE SIDE
CONTROL DATA
PROCESSING STATIONS
2A
2B
5A
1
2C
2D
4
5B
6

6

SLIDE 6. Bottom of PDS. This slide depicts the lower portion of a yellow PDS. As you can see, there are sections for each station to record its completion date. When the PDS emerges from the computer, it is signed off by the comput er at location A. All PDS's that lack an FSN or has questionable FSN's are then sent to Station 2A where, after screening for initial dissemination and other actions which are dependent upon which missile system is involved. Station 2A records its completion date at location B. Each PDS is then sent to the next station as determined by Station 2. In a similar fashion each station records its completion date at its appropriate place on the PDS. The elapsed time between successive dates is indicative of the amount of time that a PDS remained in that station including the transportation time to that station. Since this transportation time is essentially the same for all stations, no effort was made to remove this small amount of time involved. The last station to handle a PDS is Station 2E which removes the PDS from the system. The PDS's thus removed are filed, by missile system, in order of their removal. Since the elapsed time varies greatly from the initiation of a PDS to its completion date, the stack of completed PDS's in each missile system could be considered to be in a random sequence. However, to preclude any possibility of bias, when these stacks of PDS's were sampled, the PDS's were randomly selected.

CELL CONFIGURATION

STATIONS / ROUTES	A_1 MAINTENANCE ENGINEERING (PROVISIONING) STA 2A				STA
	K	P	H	S	K
B_1 W	1	4	1	5	
	1/2	3	1	6	
	1/2	5	1	6	
1	1	5	1	12	
B	1	2	1	3	

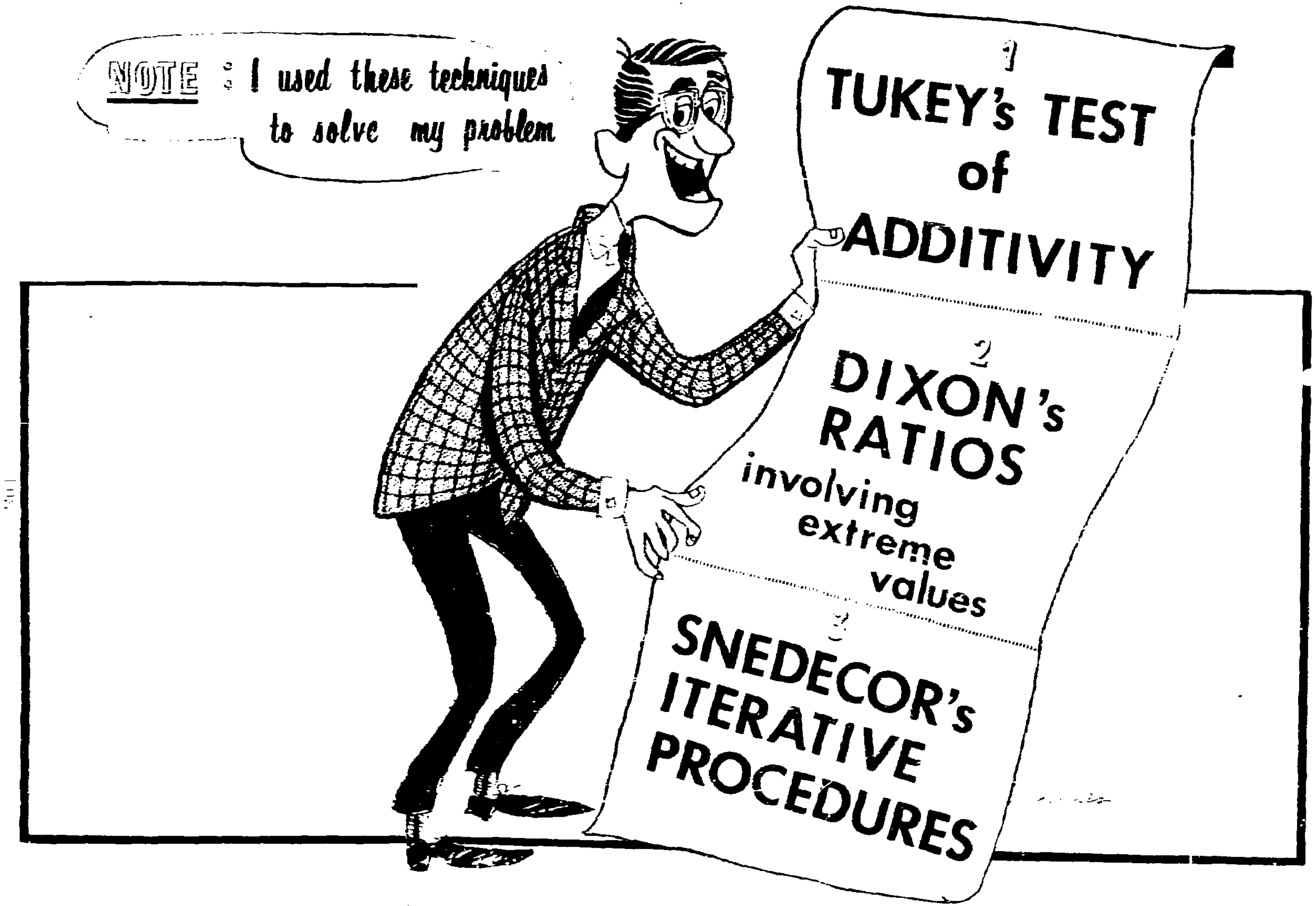
7

SLIDE 7. Blow-up of matrix - 1 cell and title blocks. A worksheet was developed to record the completion date(s) of each station for each PDS selected for each route previously identified. These samples were replicated four times for each missile system. After the range of these completion dates were established through inspection, a pseudo-Julian calendar (one which omitted all Saturdays, Sundays, and holidays) was developed to permit the transfer of each recorded date into the pseudo-Julian date. Subtraction of these converted sign-off dates indicated the elapsed number of days that each PDS remained in each station. This procedure provided a maximum of 56 measurements per missile per station, with a total of more than 1600 measurements taken to fill the matrix. From the slide one can observe how each cell was filled with these real time measurements.

The first attempt to analyze this recorded data was made through the use of an analysis of variance program, which was borrowed from the UCLA Medical Center, for two main determinations: (1) to determine if there were significant differences between each pair of routes (by omitting the broadcast function - Station 6), and (2) to determine if there were significant differences between missiles as well as colors. The results from the analysis of variance program run were tested after proper conversions against the "F" test for significance at the 95% and 99% confidence limits. All were essentially negative, which subsequently, permitted the combination of measurements for larger samples. Unfortunately, some difficulty was experienced during the computer run of this program (conflicting statements in the program and a faulty printer) which delayed the computation of the analysis of variance for these data. Furthermore, this particular analysis of variance program was incomplete in that it was not programmed to compute nor print out the mean, the variance, and the standard deviation for each row, column, sum of rows, and sum of columns nor the required two-way tables for analyzing significant differences. During this delay, utilizing the original matrix, the mean, variance, and standard deviation, at 95% confidence limits, were hand-computed for each missile, for each station, and for each route (each route having been identified for a single color) as well as for the combined group of missiles. All of these resulted in extremely large standard deviations.

8

SLIDE 8. Figure of head! This is how I felt! I needed help! So, at this point, I contacted Dr. Harshbarger, our MICOM Consultant from VPI. After a review of the data in which he agreed to its abnormality, he suggested a transformation to reduce the variability and that, in his opinion, the data followed a Poissonian, and possibly, a logarithmic distribution instead of being normally distributed. A few samples were tested by computing the variance and the mean to determine if the variance was proportional to the mean. The variance was found to be approximately proportional to the mean which indicated the transformation to be utilized could be the square root of the sample value. However, the results of this transformation after the necessary computations were completed approximated the original results. Since some of the data were less than unity, one was added to each sample value and the square root transformation was again attempted. Once again, although variability did decrease significantly, the results still approximated the original results; variations (standard deviations) were still too large. And, I still felt like this!


NOTE : I used these techniques to solve my problem
1 TUKEY's TEST of ADDITIVITY
2 DIXON's RATIOS involving extreme values
3 SNEDECOR's ITERATIVE PROCEDURES

9

SLIDE 9. Tukey - Dixon - Snedecor. Research and consultation with local statisticians produced "Tukey's Test of Additivity," a procedure which is fourfold in nature: It (1) helps decide if a transformation is necessary, (2) indicates a suitable transformation, (3) indicates if the transformation was successful, (4) gets evidence about aberrant observations. Application of this test on a few selected cells by an experienced statistician indicated the transformation required was logarithmic.

By this time, the analysis of variance results had been received from the Computation Center. As previously stated, "F" tests at the 95% and 99% levels, revealed there were no significant differences between missiles, between colors, and between routes. There were highly significant differences between stations but these were to be expected since the work content in each station is different and does require different intervals of time to perform. These results from the analysis of variance permitted the combination of like routes for each missile system to obtain larger samples. However, for comparison purposes, each route was calculated singly as well as combined for each missile system.

To return to the second result of the application of Tukey's Test, several observations were found to be aberrant. Unfortunately Tukey's Test merely indicates aberrance but does not correct them. Through the use of Dixon's "Ratios Involving Extreme Values," a technique which permits one to determine if a value is aberrant, the original matrix was reentered and all values in each cell were tested for aberrance. As each aberrant value was discovered, the remaining values for that particular cell were tested for aberrance until all data were purified. The removal of these aberrant values left the matrix with several missing values. A review of techniques to replace these missing data led to Snedecor's Iterative Procedure which was subsequently utilized.

With the matrix again complete, all recorded data were transferred into logarithmic values. The necessary computations were then performed by hand for all routes and stations for each missile system. These computations resulted in significantly lower variations and, more significantly, after converting the derived values back to normal values, truly approximated the real situation: a highly skewed curve to the left without negative times.

AVERAGE TIME (DAYS) PER STATION FOR ALL MISSILES
($\overline{X}$ = 2.2 DAYS – ALL STATIONS)

STATIONS / ROUTES	A_1 2A SMEP	A_2 3A SSC	A_3 2B SMEP	A_4 5A SSS	A_5 1 IDI	A_6 2C SMEP	A_7 3B SSC	A_8 2D SMEP	A_9 4 SMP	A_{10} 5B SSS	A_{11} 6 SSCB	TOTAL
B_1 W 1	1.5								2.2	2.9	5.0	11.6
B_2 W 1a	1.5								2.2	2.9		6.6
B_3 W 2	1.5						2.9	1.5	2.2	2.9	5.0	19.5
B_4 W 2a	1.5						2.9	1.5	2.2	2.9		14.5
B_5 G 2	1.5						2.9	1.5	2.2	2.9	5.0	19.5
B_6 G 2a	1.5						2.9	1.5	2.2	2.9		14.5
B_7 W 3	1.5	1.7			4.3	1.8	2.9	1.5	2.2	2.9	5.0	27.3
B_8 W 3a	1.5	1.7			4.3	1.8	2.9	1.5	2.2	2.9		22.3
B_{11} G 3	1.5	1.7			4.3	1.8	2.9	1.5	2.2	2.9	5.0	27.3
B_{12} G 3a	1.5	1.7			4.3	1.8	2.9	1.5	2.2	2.9		22.3
B_{13} Y 4	1.5	1.7	1.3	1.6	4.3	1.8	2.9	1.5	2.2	2.9	5.0	30.2
B_{14} Y 4a	1.5	1.7	1.3	1.6	4.3	1.8	2.9	1.5	2.2	2.9		25.2

10

SLIDE 10. Matrix w/avg station lengths. Since the mean times were now available for each station for each missile, these stations were synthesized into routes and subsequently correlated against the average times of the original routes. This was accomplished for each missile system as well as the combined group of missile systems. This slide depicts the average station length for all missiles and the synthesized total for each route. The coefficient of correlation as calculated proved to be +0.87 which is highly significant (well above the 99% level of +0.708 in a significance table for my number of degrees of freedom).

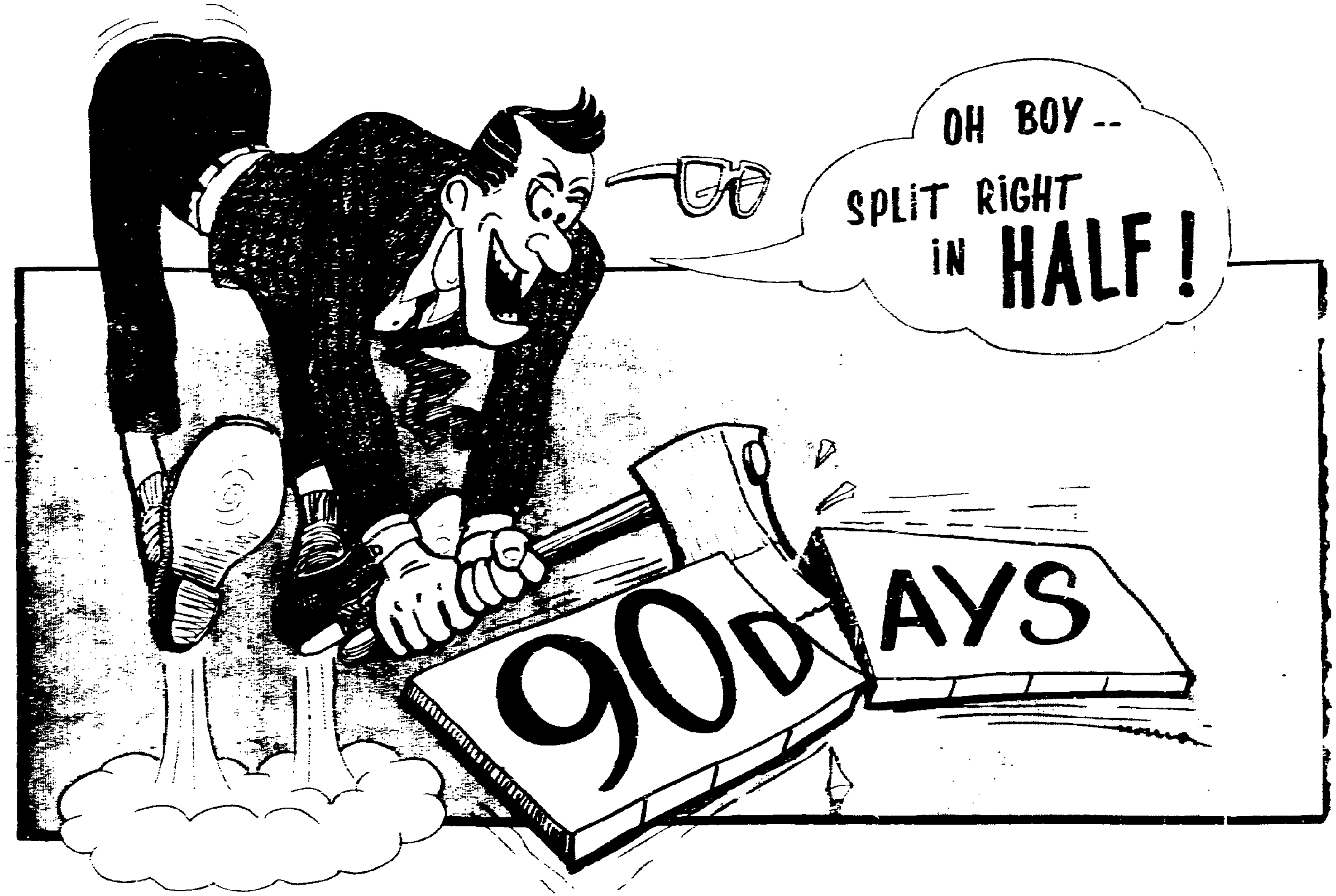
SAMPLES OF OCCURRENCES OF PDS ON VARIOUS ROUTES

ROUTES & COLORS	MISSILES (% OCCURRENCE)				TOTALS	AVERAGE		ROUTES
	P	S	H	K				
B1 W	22	23	20	15	80	20.00	Route 1	1
B2 W	19	7	23	25	74	18.50		1a
B3 W	15	24	11	8	58	14.50	Route 2	2
B4 W	6	12	10	26	54	13.50		2a
B5 G	2	4	2	10	18	4.50		2
B6 G	2	0	4	4	10	2.50		2a
[illegible]	13	3	4	0	20	5.00	Route 3	3S
[illegible]	1	1	1	0	3	.75		3Sa
B9 W	3	1	4	0	8	2.00		3L
B10 W	0	0	4	5	9	2.25		3La
[illegible]	5	0	3	2	10	2.50		3S
[illegible]	0	0	4	1	5	1.25		3Sa
B13 Y	9	20	7	3	39	9.75	Route 4	4
B14 Y	3	5	3	1	12	3.00		4a
Totals	100	100	100	100	400	100.00		
SUMMARY BY COLORS								
White PDS	79	71	77	79	306	76.50		
Green PDS	9	4	13	17	43	10.75		
Yellow PDS	[illegible]	25	10	4	51	12.75		
Totals	100	100	100	100	400	100.00		

SLIDE 11. 400 PDS sample. Current procedures provided a maximum of 90 days before a PDS became delinquent. With certain exceptions for some PDS's on a few routes, no interim time limits were specified. This analysis indicated that, on the average (95% of the time), all PDS's can be processed through each station in 2.2 days, with a range from 0.8 days to 6.3 days. A recommendation for the establishment of time limits for each station would preclude lengthy delays.

Unfortunately, S&M cannot predetermine which route a PDS will flow (with certain exceptions) but on the average, the longest route length (without broadcast) was approximately 30 days. The final step obviously was the determination of the frequency of occurrence of this longest route.

In accordance with a 400 PDS sample (100 from each missile system), this route (30 day length) occurred approximately 13% of the time. This slide illustrates the results of this sample.



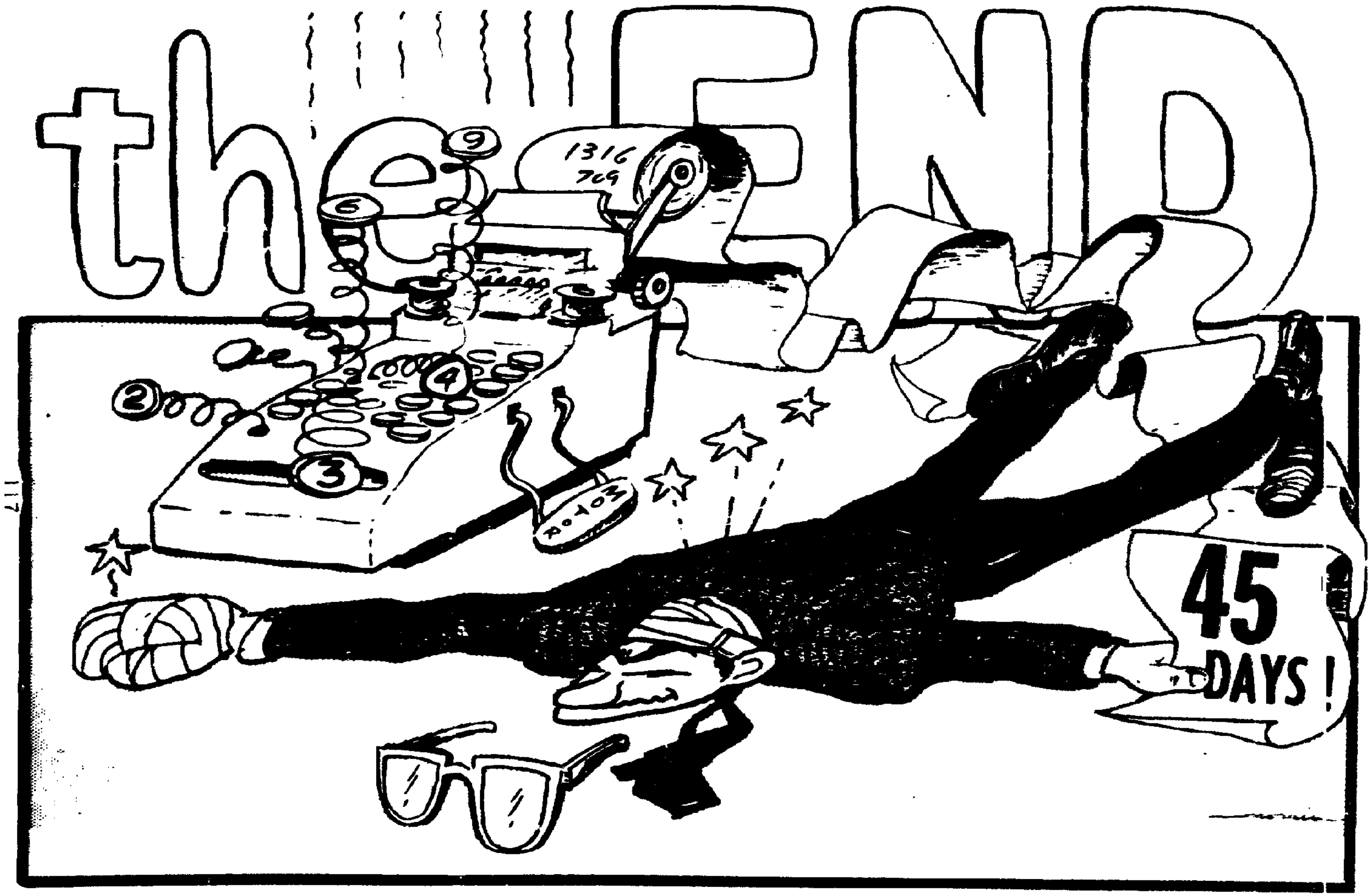
(12)

<u>SLIDE 12. Axe head cutting time in half.</u> Since a PDS did not become delinquent until the expiration of 90 days and, on the average, our longest route length was approximately 30 days plus 5 days for broadcast plus nearly 8 additional days for two standard deviations, or a total of approximately 43 days, 45 days (for the sake of a nice round number) was selected for the maximum time limit. To support the recommendation of reducing the 90-day time limit to a 45-day time limit, the probability that this 30-day route would exceed 45 days had to be calculated. This probability, using the t-distribution, was calculated to be 0.04. Thus the probability that this route length would exceed 45 days is 13% times the probability (.04) which is five one-thousandths or only 5 times out of 1000.

Consequently, a recommendation for a 45-day limit was tendered in my final analysis.

Although this now concludes my presentation, I would like to provide you with a very short follow-up. The 45-day limitation was not accepted because the powers in control felt that this cut in time was too drastic. Instead, a sixty-day time limit was substituted. However, through improved flow procedures and the subsequent recommended elimination of one station, this 60-day time limit was recently cut down to 45 days. Additional studies (non-statistical) are currently being performed to effect a further reduction in time.

I thank you!



13

OPTIMAL ECONOMY IN PLANNING EXPERIMENTS*

Regina C. Elandt-Johnson
Biostatistics Department
University of North Carolina at Chapel Hill,
North Carolina

ABSTRACT. Suppose that a cost, y, (which is a random variable) is a non-linear function of some controlled variable x, and in a general case, is expressed as a polynomial of k-th degree in x. Let

(1) $$y = c(x) + \sum_{t=1}^{k} a_t x^t \quad ,$$

and

(2) $$Y = E(y \mid x) = \sum_{t=1}^{k} A_t x^t$$

be the estimated and expected ("true") cost functions respectively. Let $\hat{x}_o$ and x_o be the values of x at which the estimated and expected cost functions attain minima respectively. Further, let $\hat{Y}_o = E(y \mid \hat{x}_o)$ be the actual expected cost when $\hat{x}_o$ is substituted for (unknown) x_o, and $Y_o = E(y \mid x_o)$ the 'true' minimum cost. We define the 'allowance' cost as

(3) $$E(\hat{Y}_o - Y_o) = \sum_{t=1}^{k} A_t \left[E(\hat{x}_o^t) - x_o^t\right] \quad .$$

If c(x) estimates C(x) closely, then (3) will usually be small.

To evaluate (3) we have to find the distribution of $\hat{x}_o$ which is a function of regression coefficients $a_1, a_2, \ldots, a_k$,

(4) $$\hat{x}_o = g(a_1, a_2, \ldots, a_k) \quad .$$

In the general case this may be complicated, but for sufficiently large sample size, n, we can find an approximate distribution using the Central Limit Theorem and a Taylor series expansion of the multivariable function

*This paper has been accepted for publication in "Operations Research".

(4). Application of orthogonal polynomials appears to be relevant to this situation.

It is easy to see from (3) that $E(\hat{Y}_o - Y_o)$ depends on the shape of the true cost function, $C(x)$, even if the fitted regression function, $c(x)$, is of the right order. Incorrect choice of the degree of $c(x)$ might affect the 'allowance' cost more severely.

ON A CLASS OF NONPARAMETRIC TESTS FOR MANOVA IN TWO WAY LAYOUTS*

PRANAB KUMAR SEN

University of North Carolina, Chapel Hill,

and University of Calcutta.

SUMMARY. The object of the present investigation is to propose and study a class of nonparametric tests for the multivariate analysis of variance (MANOVA) problem relating to complete two way layouts. In this context, the concept of rank-permutations for multidimensional interchangeability is developed, and the same is incorporated in the formulation of a class of genuinely distribution-free rank order tests. Asymptotic properties of the class of proposed tests are studied and compared with those of the standard parametric ones.

1. INTRODUCTION

Let us consider a complete two way layout comprising of n complete blocks (replicates), each block containing $r(\geq 2)$ plots where r different treatments are applied. The yield (response) is a p variate quantitative (stochastic) vector, and we denote by $X_{ij}^{(k)}$ the k-th response for the jth treatment placed in the ith block for $i = 1, \ldots, n$, $j = 1, \ldots, r$, $k = 1, \ldots, p$. In the sequel, it will be assumed that $n, r, p \geq 2$. Let then

$$\underset{\sim}{X}'_{ij} = (X_{ij}^{(1)}, \ldots, X_{ij}^{(p)}),\ i = 1, \ldots, n,\ j = 1, \ldots, r; \tag{1.1}$$

$$\underset{\sim}{\mu}' = (\mu^{(1)}, \ldots, \mu^{(p)}); \tag{1.2}$$

---

*Work supported by the Army Research Office, Durham, Grant DA-31-124-ARO-D-G432.
This article was reproduced photographically.

$$\underset{\sim}{\alpha}_i' = (\alpha_i^{(1)}, \ldots, \alpha_i^{(p)}), \ i = 1, \ldots, n; \tag{1.3}$$

$$\underset{\sim}{\tau}_j' = (\tau_j^{(1)}, \ldots, \tau_j^{(p)}), \ j = 1, \ldots, r; \tag{1.4}$$

and
$$\underset{\sim}{e}_{ij}' = (e_{ij}^{(1)}, \ldots, e_{ij}^{(p)}), \ j = 1, \ldots, r, \ i = 1, \ldots, n. \tag{1.5}$$

We adopt the usual linear model as

$$\underset{\sim}{X}_{ij} = \underset{\sim}{\mu} + \underset{\sim}{\alpha}_i + \underset{\sim}{\tau}_j + \underset{\sim}{e}_{ij}, \ i = 1, \ldots, n, \ j = 1, \ldots, r, \tag{1.6}$$

where $\underset{\sim}{\mu}$ is the vector of <u>mean effects</u>, $\underset{\sim}{\alpha}_i$ the <u>block effects</u> ($i = 1, \ldots, n$), $\underset{\sim}{\tau}_j$ the <u>treatment effects</u> ($j = 1, \ldots, r$), and $\underset{\sim}{e}_{ij}$ the residual error vectors ($i = 1, \ldots, n, \ j = 1, \ldots, r$). These component vectors are assumed to be mutually independent. Our problem is to have a comprehensive test for the hypothesis of no treatment effects i.e.,

$$H_0: \ \underset{\sim}{\tau}_1 = \ldots = \underset{\sim}{\tau}_r. \tag{1.7}$$

In the parametric case, it is usually assumed that $\underset{\sim}{e}_{ij}$ ($i = 1, \ldots, n, \ j = 1, \ldots, r$) are $N(= nr)$ independent and identically distributed stochastic vectors distributed according to a multinormal distribution with a null mean vector and a dispersion matrix (positive definite) $\underset{\sim}{\Sigma} = ((\sigma_{kq}))$, where σ_{kq} is the covariance of $(e_{ij}^{(k)}, e_{ij}^{(q)})$, for $k, q = 1, \ldots, p$. The parametric MANOVA tests are either based on the likelihood ratio criterion or on the characteristic roots of some determinantal equations. The likelihood ratio criterion reduces to the ratio of two generalized variances and can be expressed as the product of several (p) independent beta variables (cf. Anderson (1958, Chapter 8)). Alternatively, one may work with the smallest characteristic root of the determinental equation involving the same generalized variances. Occasionally, some symmetric function of the roots are also used. For details, the reader is referred to Rao (1965, chapter 8). The parametric tests thus appear to be deterministic, but they are not very simple, especially for $p > 2$. Further, in this procedure the assumptions of independence and multinormality of the error vectors play an indispensible role. Unlike the univariate case, very little has been investigated about the effects of departure from these two basic assumptions on the performance characteristics of the parametric MANOVA tests. On the otherhand, the assumption of multinormality of the error vectors is often found to be dubious, especially in many biometric problems. Further, in many problems, there appears to be sufficient evidence on the stochastic dependence of the error vectors within the same block. For example, in agricultural experiments, the presence of spatial correlation may distort the stochastic independnece of the error vectors within the same block. Similar dependence may be due to genetic effects in many animal feeding experiments. The object of the present investigation is to relax both the assumptions of multinormality as well as independence of the error components. In fact, for the tests proposed here, we require only that

(i) the joint distribution function $F(\underset{\sim}{e}_{i1}, \ldots, \underset{\sim}{e}_{ir})$ of $\underset{\sim}{e}_{i1}, \ldots, \underset{\sim}{e}_{ir}$ is continuous and independent of $i = 1, \ldots, n$, and

(ii) $F(\underset{\sim}{e}_{i1}, \ldots, \underset{\sim}{e}_{ir})$ is a symmetric function of its r arguments (vectors) $\underset{\sim}{e}_{i1}, \ldots, \underset{\sim}{e}_{ir}$ i.e., F remains invariant under any permutation of the r vectors among themselves, or in other words, $\underset{\sim}{e}_{i1}, \ldots, \underset{\sim}{e}_{ir}$ are symmetric dependent stochastic vectors.

Evidently, both the assumptions (i) and (ii) are much less restrictive than the usual assumptions of independence and multinormality. Thus, the proposed method appears to have a comparatively wider scope of applicability.

In the nonparametric case, practically no work has been done on this line. For completely randomized layouts, very recently some nonparametric MANOVA tests have been offered by Chatterjee and Sen (1964, 1966), Sen (1965, 1966a), Puri and Sen (1966), and Anderson (1965), among few others. Bhapkar (1965) has also presented some asymptotically distribution-free test for the same problem. The present author (1966 b) has considered some rank methods for combination of independent experiments in MANOVA. The same procedure is applicable in our situation here, but it fails to be suitable in some respects. This problem may also be regarded as the multivariate generalization of the nonparametric ANOVA tests relating to two way layouts. Such ANOVA tests have been considered by Friedman (1937), Durbin (1951), Brown and Mood (1951), Benard and Elteren (1953), and others. These are all based on intra-block rankings, and the same method can be generalized to the MANOVA problem. The present author (1966 c) has considered a modified approach to nonparametric ANOVA tests for two way layouts. Extending an idea of Hodges and Lehmann (1962), he has considered the rankings after alignment, and under a suitable permutation model, has obtained a class of genuinely distribution-free tests based on these modified rankings. This results, in most of the cases, in an increased (at least asymptotically) efficiency of the proposed test. The object of this paper is to generalize the method of rankings after alignment to the MANOVA problem and to offer some suitable non-parametric tests for the same. For this purpose, the concept of <u>multidimensional interchangeability</u> is developed and certain rank permutational ideas are formulated. With the aid of this a class of properly distribution-free rank order tests for the hypothesis in (1.7) is developed. Further, the celebrated Chernoff-Savage (1958) theorem on the asymptotic normality and power-efficiency of a class of univariate nonparametric test-statistics, as extended to the multivariate case by Puri and Sen (1966) and to the problem of compound symmetry of multivariate distributions by Sen (1966 c), is extended further to take care of the problem of multidimensional interchangeability, to be considered here. With the aid of this, the asymptotic power and power-efficiency of the proposed class of tests are studied.

2. SOME PRELIMINARY NOTIONS.

Let us define a set of r^2 real quantities by

$$c_{\ell j} = \delta_{\ell j} - 1/r \text{ for } j, \ell = 1, \ldots, r, \tag{2.1}$$

where $\delta_{\ell j}$ is the usual Kronecker delta. Thus, $\Sigma_{j=1}^{r} c_{\ell j} = 0$ for all $\ell = 1, \ldots, r$. Let us then consider the r intra-block contrasts

$$\underset{\sim}{Y}_{i\ell} = \Sigma_{j=1}^{r} c_{\ell j} \underset{\sim}{X}_{ij}, \ \ell = 1, \ldots, r. \tag{2.2}$$

From (1.6) and (2.2), we have

$$\underset{\sim}{Y}_{i\ell} = (\underset{\sim}{\tau}_{\ell} - \frac{1}{r} \Sigma_{j=1}^{r} \underset{\sim}{\tau}_{j}) + (\underset{\sim}{e}_{i\ell} - \frac{1}{r} \Sigma_{j=1}^{r} \underset{\sim}{e}_{ij}), \tag{2.3}$$

where the first factor on the right hand side of (2.3) vanishes when H_o in (1.7) holds. Further, by assumption (ii) of section 1, we get with some simple reasonings that the joint distribution of $\left[(\underset{\sim}{e}_{i\ell} - \frac{1}{r} \Sigma_{j=1}^{r} \underset{\sim}{e}_{ij}), \ \ell = 1, \ldots r\right]$ is a symmetric function of the r (vector) arguments. Consequently, from (2.3), we get that under H_o in (1.7), the joint distribution of $(\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir})$ will be a symmetric function of the r vectors $\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir}$. On the otherhand, if H_o in (1.7) does not hold, the joint distribution of $(\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir})$ will be a symmetric function of its (vector) arguments only when each one of them is adjusted by appropriate location vectors. Thus, if instead of the observed responses $\underset{\sim}{X}_{ij}$'s, we work with the block-adjusted yields $\underset{\sim}{Y}_{ij}$'s, our problem of testing H_o in (1.7) reduces to that of testing the hypothesis of interchangeability of the vectors $\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir}$ (for all $i = 1, \ldots, n$), against translation type of alternatives. This is termed the problem of multidimensional interchangeability, and a formulation of an appropriate rank permutation model for the same, will be considered in the next section. The necessary rank order statistics will be defined now.

Let us pool the $N(= nr)$ observations $\{Y_{ij}^{(k)}, j = 1, \ldots, r, i = 1, \ldots, n\}$ into a combined set and denote the ordered observations by

$$Y_{(1)}^{(k)} < \ldots < Y_{(N)}^{(k)}, \tag{2.4}$$

where by virtue of the assumed continuity of the distribution of the error vectors, the possibility of ties in (2.4) may be neglected, in probability. Let then $C(u)$ be the usual sign-function viz.,

$$c(u) = \begin{cases} 1, & \text{if } u > 0 \\ 0, & \text{if } u \leq 0, \end{cases} \tag{2.5}$$

and let

$$R_{ij}^{(k)} = 1 + \sum_{\alpha=1}^{N} c(Y_{ij}^{(k)} - Y_{(\alpha)}^{(k)}), \tag{2.6}$$

for $i = 1, \ldots, n$, $j = 1, \ldots, r$.

Thus $R_{ij}^{(k)}$ stands for the rank of $Y_{ij}^{(k)}$ within the set (2.4). This ranking procedure is employed separately for each $k = 1, \ldots, p$. Consequently, any vector $\underset{\sim}{Y}_{ij}$ having p elements is made to correspond to a rank p-vector

$$\underset{\sim}{R}'_{ij} = (R_{ij}^{(1)}, \ldots, R_{ij}^{(p)}), \tag{2.7}$$

for $i = 1, \ldots, n$, $j = 1, \ldots, r$. The composite collection is a $p \times N$ matrix

$$\underset{\sim}{R}_N^{p\times N} = (\underset{\sim}{R}_{11}, \ldots, \underset{\sim}{R}_{1r}, \ldots, \underset{\sim}{R}_{n1}, \ldots, \underset{\sim}{R}_{nr}). \tag{2.8}$$

$\underset{\sim}{R}_N$ will be termed a <u>collection (rank) matrix.</u> Each row of $\underset{\sim}{R}_N$ is a permutation of the numbers $1, \ldots, N$. For any positive integer $N(= nr, n = 1, 2, \ldots)$ we define p sequences of real numbers by

$$\underset{\sim}{E}_N^{(k)} = (E_{N,1}^{(k)}, \ldots, E_{N,N}^{(k)}), k = 1, \ldots, p. \tag{2.9}$$

$E_{N,\alpha}^{(k)}$'s are all real quantities and are explicit functions of $(\frac{\alpha}{N+1})$. We adopt the coventional Chernoff-Savage (1958) form and write

$$E_{N,\alpha}^{(k)} = J_N^{(k)}(\frac{\alpha}{N+1}), \alpha = 1, \ldots, N, k = 1, \ldots, p, \tag{2.10}$$

where the function $J_N^{(k)}$ need be defined only at $\frac{\alpha}{N+1}$, $\alpha = 1, \ldots, N$. However, we shall find it more convenient to extend its domain of definition to (0, 1) according to the Chernoff-Savage convention. Also, we define rp requences of indicator functions $\{Z_{N,\alpha}^{(j,k)}, \alpha = 1, \ldots, N\}$, for $j = 1, \ldots, r$, $k = 1, \ldots, p$ by

$$Z_{N,\alpha}^{(j,k)} = \begin{cases} 1, \text{ of } Y_{(\alpha)}^{(k)} \text{ is some } Y_{ij}^{(k)} (i = 1, \ldots, n), \\ 0, \text{ otherwise,} \end{cases} \tag{2.11}$$

for $\alpha = 1, \ldots, N$. Then we define rp rnak order statistics

$$T_{N,j}^{(k)} = \frac{1}{n} \sum_{\alpha=1}^{N} E_{N,\alpha}^{(k)} Z_{N,\alpha}^{(j,k)}, j = 1, \ldots, r, k = 1, \ldots, p. \tag{2.12}$$

It may be noted that

$$\frac{1}{r} \sum_{j=1}^{r} T_{N,j}^{(k)} = \frac{1}{N} \sum_{\alpha=1}^{N} E_{N,\alpha}^{(k)} = \bar{E}_N^{(k)} \text{(say,)}, k = 1, \ldots, p; \tag{2.13}$$

where $\bar{E}_N^{(1)}, \ldots, \bar{E}_N^{(p)}$ are all known constants (depending on N). Thus, at most $(r - 1)p$ of ths rp variables in (2.12) are linearly independent. Our proposed test is based on the set of random variables in (2.12). To develop strictly distribution-free tests for the hypothesis (1.7), we shall consider in the next section some permutation model. But, before that it may be worth writing a point of clarification. The class of statistics in (2.12) has some similarity with that of a similar class of statistics considered by Puri and Sen (1966). However, in the later case, we have a one way classification with N independent p-variate observations, while in this case, we have a two way classification with n independent pr-variate observations. This makes the situation somewhat more complicated, and requires a more specialized attention for both the permutation as well as asymptotic test theory.

3. RANK PERMUTATIONS FOR MULTIDIMENSIONAL INTERCHANGEABILITY.

The collection matrix $\underset{\sim}{R}_N^{p\times N}$, given by (2.8), is now expressed in terms of n submatrices $\underset{\sim}{R}_1^{p\times r}, \ldots, \underset{\sim}{R}_n^{p\times r}$, where $\underset{\sim}{R}_i^{p\times r}$ is the matrix of the r rank p-tuplets corresponding to $(\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir})$, for $i = 1, \ldots, n$. Thus, we have

$$\underset{\sim}{R}_N^{p\times N} = (\underset{\sim}{R}_1^{p\times r}, \ldots, \underset{\sim}{R}_n^{p\times r}). \tag{3.1}$$

Now under the null hypothesis (1.7), the joint distribution function $G(\underset{\sim}{Y}_{i1},\ldots,\underset{\sim}{Y}_{ir})$ is a symmetric function of $\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir}$, and hence, the same remains invariant under any permutation of the r vectors in the r positions of G. Since, there are r! possible permutations of the r vectors among themselves, the permutational probability (i.e., conditional probability) mass associated with each of the r! possible permutations is equal to $(r!)^{-1}$, (under H_o in (1.7),) for all $i = 1, \ldots, n$. Since, $(\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir})$ is distributed (jointly) independently of $(\underset{\sim}{Y}_{i'1}, \ldots, \underset{\sim}{Y}_{i'r})$ for all $i \neq i' = 1, \ldots, n$, the joint distribution of

$$Y_N = (Y_{11}, \ldots, Y_{1r}, \ldots, Y_{n1}, \ldots, Y_{nr}) \tag{3.2}$$

remains invariant under the following finite group $\mathcal{G}_n$ of transformations $\{g_n\}$ which maps the sample space of $\underset{\sim}{Y}_N$ onto itself. The number of elements of $\mathcal{G}_n$ is equal to $(r!)^n$, and typically a transformation g_n is such that

$$g_n \underset{\sim}{Y}_N = \underset{\sim}{Y}_N^* = (\underset{\sim}{Y}_{11}^*, \ldots, \underset{\sim}{Y}_{1r}^*, \ldots, \underset{\sim}{Y}_{n1}^*, \ldots, \underset{\sim}{Y}_{nr}^*), \tag{3.3}$$

where $(\underset{\sim}{Y}_{i1}^*, \ldots, \underset{\sim}{Y}_{ir}^*)$ is any permutation of $(\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir})$, $i = 1, \ldots, n$. Let $\mathcal{Y}_N$ be the Np-dimensional sample space of $\underset{\sim}{Y}_N$, (and we take it to be the Np-dimensional Euclidean space). Evidently, the sample space of $\underset{\sim}{Y}_N^*$ is the same as that of $\underset{\sim}{Y}_N$, and moreover, under H_0 in (1.7), the joint distribution of $\underset{\sim}{Y}_N$ remains invariant under the group of transformations $\mathcal{G}_n$. Let now $S(\underset{\sim}{Y}_N)$ be a (real or vector valued) function on $\underset{\sim}{Y}_N$. Then, for any $\underset{\sim}{Y}_N \in \underset{\sim}{\mathcal{Y}}_N$, we will have a set of $(r!)^n$ values of $S(\underset{\sim}{Y}_N)$, obtained under the group of transformations $\mathcal{G}_n$, and this set is denoted by $\Sigma(\underset{\sim}{Y}_N)$. Then, under the null hypothesis (1.7), the conditional distribution of $S(\underset{\sim}{Y}_N)$ over the set $\Sigma(\underset{\sim}{Y}_N)$ will be uniform. Let us define $T_{N,j}^{(k)}$ as in (2.2), and let

$$\underset{\sim}{T}_N^{rxp} = ((T_{N,j}^{(k)}))_{j} = 1, \ldots, r, \; k = 1, \ldots, p \tag{3.4}$$

Then, it follows that $\underset{\sim}{T}_N$ is a stochastic matrix, which under the group of transformations $\mathcal{G}_n$ can have only $(r!)^n$ possible realizations. Since $\underset{\sim}{T}_N$ is an explicit function of the N rank p-tuplets $\underset{\sim}{R}_{ij}$, $i = 1, \ldots, n$, $j = 1, \ldots, r$, it will be more convenient for us to review the above invariance argument in terms of the following rank-invariance argument.

The way in which we have defined $\underset{\sim}{R}_N$ in (2.8) and (3.1), it follows that for any $\underset{\sim}{Y}_N \in \underset{\sim}{\mathcal{Y}}_N$ there will be a corresponding collection matrix $\underset{\sim}{R}_N$. On examining the group of transformations $\mathcal{G}_n$, it will be clear that the transformation g_n on $\underset{\sim}{Y}_N$, given by (3.3), gives rise to another collection matrix $\underset{\sim}{R}^*_N$, which is obtained by applying the same transformation g_n on the original collection matrix $\underset{\sim}{R}_N$. Thus, under the group of transformations $\mathcal{G}_n$ of $\{g_n\}$, the rank collection matrix $\underset{\sim}{R}_N$ (corresponding to $\underset{\sim}{Y}_N \varepsilon \mathcal{Y}_N$) gives rise to a set of $(r!)^n$ rank collection matrices (obtained by applying the same transformations $\{g_n\}$,) and this set is denoted by $\Sigma(\underset{\sim}{R}_N)$. If $\underset{\sim}{R}^*_N$ is any member of $\Sigma(\underset{\sim}{R}_N)$, we note that $\underset{\sim}{R}^*_N$ is really derived from $\underset{\sim}{R}_N$ by a finite number of inversions of the columns of the later. Thus we may write

$$\underset{\sim}{R}^*_N \sim \underset{\sim}{R}_N \ (\mathrm{mod}\ \mathcal{G}_n) \text{ for all } \underset{\sim}{R}^*_N \ \varepsilon \ \Sigma(\underset{\sim}{R}_N). \tag{3.5}$$

Hence, the set $\Sigma(\underset{\sim}{R}_N)$ contains $(r!)^n$ rank-matrices which are permutationally (under inversions of intra-block columns) equivalent (under $\mathcal{G}_n$) to $\underset{\sim}{R}_N$. Thus, we term $\Sigma(\underset{\sim}{R}_N)$ as the <u>permutation set</u> (mod $\mathcal{G}_n$) of $\underset{\sim}{R}_N$. $\underset{\sim}{R}_N$ like $\underset{\sim}{Y}_N$ is a stochastic variable, and each row of $\underset{\sim}{R}_N$ is a permutation of 1, ..., N. Thus, $\underset{\sim}{R}_N$ can have $(N!)^p$ possible realizations, and this set of all possible realizations of $\underset{\sim}{R}_N$ is denoted by $\mathcal{R}_N$, so that

$$\underset{\sim}{R}_N \ \varepsilon \ \Sigma(\underset{\sim}{R}_N) \subset \mathcal{R}_N. \tag{3.6}$$

The probability distribution of $\underset{\sim}{R}_N$ on $\mathcal{R}_N$ (defined on an additive class of subsets A_N of $\mathcal{R}_N$,) will depend on the unknown joint distributions $G(\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir})$, $i = 1, \ldots, n$, even under H_0 in (1.7). Thus, unlike the case of univariate one way classified data, the use of the unconditional distribution of $\underset{\sim}{R}_N$ will fail to provide a distribution-free test. However, from what has been discussed before, it follows that

$$P\{\underset{\sim}{R}_N = \underset{\sim}{R}^*_N \mid \Sigma(\underset{\sim}{R}_N), H_0\} = (r!)^{-n}, \tag{3.7}$$

for all $\underset{\sim}{R}^*_N \ \varepsilon \ \Sigma(\underset{\sim}{R}_N)$, independently of $G(\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir})$, $i = 1, \ldots, n$. Now, the way in which $\underset{\sim}{E}_N^{(k)}$, k = 1, ..., p, are defined by (2.9), (2.10), it follows that $\underset{\sim}{T}_N$ in [(2.12), (3.4)] is an explicit function of $\underset{\sim}{R}_N$. Thus, the set $\Sigma(\underset{\sim}{R}_N)$ will give rise to a set of $(n!)^n$ realizations of $\underset{\sim}{T}_N$, and this set is denoted by $\Sigma(\underset{\sim}{T}_N)$. Hence, under the permutational probability measure (3.7), we will have a completely specified permutational distribution of $\underset{\sim}{T}_N$, and the corresponding permutational probability measure is denoted by $\mathcal{P}_n$. Let us then consider a test function $\phi(\underset{\sim}{Y}_N)(0 \leq \phi \leq 1)$, which to each $\underset{\sim}{Y}_N \in \underset{\sim}{Y}_N$ associates a probability of rejecting H_0 in (1.7), with the aid of $\mathcal{P}_n$. It follows that we can always select $\phi(\underset{\sim}{Y}_N)$ in such a manner that

$$\sum_{Y_N^* \in \Sigma(Y_N)} \phi \; (\underset{\sim}{Y}_N^*) = (r!)^n \cdot \varepsilon, \tag{3.8}$$

where $\varepsilon(0 < \varepsilon < 1)$ is the preassigned level of significance of the test. Consequently, $\phi(\underset{\sim}{Y}_N)$ has the S(ε) - structure of tests [cf. Lehmann and Stein (1949)], and is a similar size ε test for the null hypothesis (1.7).

Now, in actual practice, we prefer to use some single-valued function of $\underset{\sim}{T}_N$ as a test-statistic. There seems to be no definite suggestions regarding the structure of this test-statistics, and an optimum choice naturally may depend appreciably on the particular class of alternatives we have in mind. However, it may be suitable (though not necessarily optimum) to consider the following test-statistic which is the quadratic-form associated with the asymptotic permutation distribution of $\underset{\sim}{T}_N$. For this, let us consider first the permutational moments of $\underset{\sim}{T}_N$. It readily follows that

$$E\,(T_{N,j}^{(k)}\,|\mathcal{P}_n\,) = \bar{E}_N^{(k)}, \text{ for } k = 1, \ldots, p,\ j = 1, \ldots, r. \tag{3.9}$$

Let us define

$$\bar{E}^{(k)}_{NR_{i\cdot}}(k) = \frac{1}{r}\,\Sigma^{r}_{j=1}\, E^{(k)}_{N,R_{ij}}(k)\,,\ i = 1, \ldots, n,\ k = 1, \ldots, p, \qquad (3.10)$$

as the intra-block averages. Also let

$$v_{kq}(\underset{\sim}{R}_N) = \frac{1}{n(r-1)} \sum_{i=1}^{n} \sum_{j=1}^{r} \left\{ E^{(k)}_{N,R_{ij}}(k) - \bar{E}^{(k)}_{N,R_{i.}}(k) \right\} \left\{ \bar{E}^{(q)}_{N,R_{i}}(q) - \bar{E}^{(q)}_{N,R_{i.}}(q) \right\}, \qquad (3.11)$$

for k, q = 1, ..., p;

$$\underset{\sim}{V}_N(\underset{\sim}{R}_N) = ((v_{kq}(\underset{\sim}{R}_N)))_{k,\ q = 1, \ldots, p} \qquad (3.12)$$

It is then easy to verify that

$$\text{Cov}\ \{T^{(k)}_{N,j}, T^{(q)}_{N,j'} \mid {}_n\} = \frac{1}{nr}\,(\delta_{jj'}r - 1)v_{kq}(\underset{\sim}{R}_N), \qquad (3.13)$$

for k, q = 1, ..., p, j, j' = 1, ..., r, where $\delta_{jj'}$ is the usual Kronecker delta. For the time being, let us assume that $\underset{\sim}{V}_N(\underset{\sim}{R}_N)$, given by (3.12), is positive definite, and denote its reciprocal matrix by

$$\underset{\sim}{V}^{-1}_N(\underset{\sim}{R}_N) = ((v^{kq}(\underset{\sim}{R}_N)))_{k,\ q = 1, \ldots, p} \qquad (3.14)$$

Our proposed test-statistic S_N can then be expressed as

$$S_N = n \sum_{k=1}^{p} \sum_{q=1}^{p} v^{kq}(\underset{\sim}{R}_N) \sum_{j=1}^{r} [T^{(k)}_{N,j} - \bar{E}^{(k)}_N][T^{(q)}_{N,j} - \bar{E}^{(q)}_N]\,, \qquad (3.15)$$

and it may be noted that S_N is essentially a non-negative stochastic variable. We shall see later on that under certain regularity conditions on $G(\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir})$, $\underset{\sim}{V}_N(\underset{\sim}{R}_N)$ is positive definite with a very high probability, (precisely, in probability). However, if $\underset{\sim}{V}_N(\underset{\sim}{R}_N)$ fails to be non-singular, we may work with the highest order principal minor of $\underset{\sim}{V}_N(\underset{\sim}{R}_N)$ which is positive definite, and proceed similarly only with the responses pertaining to this minor. Thus, for convenience, we may assume $\underset{\sim}{V}_N(\underset{\sim}{R}_N)$ to be positive definite. Now,

$$E(S_N \mid \mathcal{P}_n) = p(r - 1), \tag{3.16}$$

and S_N measures the distance of $\underset{\sim}{T}_N$, in (3.4), from the permutational centre of gravity of the same. If H_0 in (1.7) does not hold, it can be shown that for at least one $k = 1, \ldots, p$ and one $j = 1, \ldots, r$, $T_{N,j}^{(k)}$ will converge to a point (stochastically) other than $\bar{E}_N^{(k)}$, and hence, by (3.15), S_N will be stochastically larger. Thus, we may propose the following test function:

$$\phi(\underset{\sim}{Y}_N) = \begin{cases} 1, \text{ if } S_N > S_{N,\epsilon}(\underset{\sim}{R}_N), \\ \gamma(\underset{\sim}{R}_N), \text{ if } S_N = S_{N,\epsilon}(\underset{\sim}{R}_N), \\ 0, \text{ if } S_N < S_{N,\epsilon}(\underset{\sim}{R}_N), \end{cases} \tag{3.17}$$

where the constants $S_{N,\epsilon}(\underset{\sim}{R}_N)$ and $\gamma(\underset{\sim}{R}_N)$ may usually depend on $\underset{\sim}{R}_N$ and are so chosen that

$$E\{\phi(\underset{\sim}{Y}_N) \mid \mathcal{P}_n\} = \epsilon; \quad 0 < \epsilon < 1. \tag{3.18}$$

(3.18) implies that $E\{\phi(\underset{\sim}{Y}_N) \mid H_0\} = \epsilon$. For small values of n(and r), one may venture to evaluate the exact values of $S_{N,\epsilon}(\underset{\sim}{R}_N)$ and $\gamma(\underset{\sim}{R}_N)$ with the aid of (3.7). However, the labor of this process of evaluation increases considerably with the increase in n(or r), and hence, as in other permutation tests, we are faced with the problem of finding out the asymptotic form of the permutation distribution of S_N. This is done in the next section.

4. ASYMPTOTIC PERMUTATION DISTRIBUTION OF S_N.

We shall impose certain regularity conditions on the p sequences $\{\underset{\sim}{E}_N^{(k)}\}$, $k = 1, \ldots, p$, defined by (2.9) and (2.10), as well as on the joint distribution function $G(\underset{\sim}{Y}_{i1}, \ldots, \underset{\sim}{Y}_{ir})$. Let us define

$$F_{N[j]}^{(k)}(x) = \frac{1}{n}\left[\text{Number of } Y_{ij}^{(k)} \le x\right], \; k = 1, \ldots, p, \; j = 1, \ldots, r; \tag{4.1}$$

$$H_N^{(k)}(x) = \frac{1}{n}\, \Sigma_{j=1}^{r} F_{N[j]}^{(k)}(x), \; k = 1, \ldots, p; \tag{4.2}$$

$$F_{N[j,\ell]}^{(k,q)}(x, y) = \frac{1}{n}\left[\text{Number of } (Y_{ij}^{(k)}, Y_{i\ell}^{(q)}) \le (x, y)\right], \tag{4.3}$$

for $k, q = 1, \ldots, p$, $j, \ell = 1, \ldots, r$ with either $j \ne \ell$ or $k \ne q$ or both. Now, corresponding to the joint cdf G, let us denote the marginal cdf of $Y_{ij}^{(k)}$ and of $(Y_{ij}^{(k)}, Y_{i\ell}^{(q)})$ by $F_{[j]}^{(k)}(x)$ and $F_{[j,\ell]}^{(k,q)}(x, y)$, respectively, for $j, \ell = 1, \ldots, r$, $k, q = 1, \ldots, p$, with at least one of $j \ne \ell$, $k \ne q$ being true, and let

$$H^{(k)}(x) = \frac{1}{r}\, \Sigma_{j=1}^{r} F_{[j]}^{(k)}(x), \text{ for } k = 1, \ldots, p. \tag{4.4}$$

With the definition of $E_{N,\alpha}^{(k)}$'s as in (2.10), we make the following assumptions concerning $J_N^{(k)}$'s.

<u>ASSUMPTION 1.</u> $\lim_{n\to\infty} J_N^{(k)}(H) = J^{(k)}(H)$ <u>exists for all</u> $0 < H < 1$ <u>and is not a constant.</u> Since, we shall be interested here in translation type of alternatives, we shall further assume that

$$J^{(k)}(H) \text{ is } \uparrow \text{ in } H\text{: } 0 < H < 1 \text{ for all } k = 1, \ldots, p. \tag{4.5}$$

ASSUMPTION 2. $\frac{1}{N} \Sigma_{\alpha=1}^{N} \left| J_N^{(k)} \left(\frac{\alpha}{N+1}\right) - J^{(k)}\left(\frac{\alpha}{N+1}\right) \right| = o(N^{-\frac{1}{2}})$, (4.6)

for $k = 1, \ldots, p$, and

$$\int_{-\infty}^{\infty} [J_N^{(k)}(\frac{N}{N+1} H_N^{(k)}(x)) - J^{(k)}(\frac{N}{N+1} H_N^{(k)}(x))] \, dF_{N[j]}^{(k)}(x) = o_p(N^{-\frac{1}{2}}), \tag{4.7}$$

for all $k = 1, \ldots, p$, $j = 1, \ldots, r$.

ASSUMPTION 3. $J^{(k)}(H)$ is absolutely continuous in H: $0 < H < 1$, and

$$\left[\frac{d^r}{dH^r} J^{(k)}(H) \right] \leq K \, [H(1 - H)]^{-r-\frac{1}{2}+\delta}, \tag{4.8}$$

for $r = 0, 1$, and some $\delta > 0$, where K is a finite positive constant.

Also for the positive definiteness and asymptotic convergence of the covariance matrix $\underset{\sim}{V}_N(\underset{\sim}{R}_N)$, given by (3.12), we require two more mild regularity conditions.

ASSUMPTION 4. $\frac{1}{N} \Sigma_{\alpha=1}^{N} \left| \{J_N^{(k)}(\frac{\alpha}{N+1})\}^2 - \{J^{(k)}(\frac{\alpha}{N+1})\}^2 \right| = o(1)$, (4.9)

for $k = 1, \ldots, p$, and

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big[J_N^{(k)}(\frac{N}{N+1} H_N^{(k)}(x)) J_N^{(q)}(\frac{N}{N+1} H_N^{(q)}(y)) - J^{(k)}(\frac{N}{N+1} H_N^{(k)}(x)) J^{(q)}(\frac{N}{N+1} H_N^{(q)}(y))\Big] dF_{N[j,\ell]}^{(k,q)}(x,y)$$

$$= o_p(1) \text{ for all } j, \ell = 1, \ldots, p, \; k, q = 1, \ldots, p, \tag{4.10}$$

where either $k \neq q$ or $j \neq \ell$ or both. Let us also define

$$Z_{ij}^{(k)} = J^{(k)}(H^{(k)}(Y_{ij}^{(k)})), \; k = 1, \ldots, p, \; j = 1, \ldots, r; \tag{4.11}$$

$$\underset{\sim}{Z}_{ij} = (Z_{ij}^{(1)}, \ldots, Z_{ij}^{(p)}), \quad j = 1, \ldots, r; \tag{4.12}$$

$$a_{kq \cdot j\ell} = E\{Z_{ij}^{(k)} \cdot Z_{i\ell}^{(q)}\} \text{ for } k,\ q = 1, \ldots, p,\ j,\ \ell = 1, \ldots, r; \tag{4.13}$$

$$A_{j\ell} = ((a_{kq \cdot j\ell}))_{k,q = 1, \ldots, p},\ j,\ \ell = 1, \ldots, r; \tag{4.14}$$

$$\nu_{kq} = \frac{1}{r} \Sigma_{j=1}^{r} a_{kq \cdot jj} - \frac{1}{r^2} \sum_{j=1}^{r} \sum_{\ell=1}^{r} a_{kq \cdot j\ell}, \text{ for } k,\ q=1, \ldots, p \tag{4.15}$$

$$\underset{\sim}{\nu} = ((\nu_{kq}))_{k,\ q=1, \ldots, p}. \tag{4.16}$$

<u>**ASSUMPTION 5.**</u> <u>$\underset{\sim}{\nu}$ is positive definite</u> (4.17)

Before we present the main theorems of this section, let us consider the conditions under which assumption 5 holds. Using (4.14), let us define

$$\underset{\sim}{A}_{(j,\ell)} = [\underset{\sim}{A}_{jj} + \underset{\sim}{A}_{\ell\ell} - 2\, \underset{\sim}{A}_{j\ell}], \quad j \neq \ell = 1, \ldots, r. \tag{4.18}$$

<u>**THEOREM 4.1**</u> <u>Assumption 5 holds if</u>

$$\max_{j \neq \ell = 1, \ldots, r} [\text{Rank of } \underset{\sim}{A}_{(j,\ell)}] = p \tag{4.19}$$

<u>**PROOF**</u>. Let $\underset{\sim}{\ell}' = (\ell_1, \ldots, \ell_p)$ be any real and <u>non-null</u> p-vector, and let

$$t_j = \underset{\sim}{\ell}' \underset{\sim}{Z}_{ij}, \quad j=1, \ldots, r, \qquad t. = \frac{1}{r} \Sigma_{j=1}^{r} t_j, \tag{4.20}$$

where $\underset{\sim}{Z}_{ij}$'s are defined by (4.12). It is then easily seen that

$$\underset{\sim}{\ell}' \underset{\sim}{\nu} \underset{\sim}{\ell} = \frac{1}{r} \Sigma_{j=1}^{r} E(t_j^2) - E(t_.^2) \geq 0. \tag{4.21}$$

Thus, we require only to show that for any non-null $\underset{\sim}{\ell}$, (4.21) is strictly positive. Using essentially the proof of lemma 4.1 of Sen (1966), it can be shown that $\frac{1}{r} \Sigma_{j=1}^{r} E(t_j^2) - E(t_.^2)$ will be strictly positive unless

$$E(t_j t_\ell) = E(t_j^2) = \text{constant, for all } j, \ell=1, \ldots, r. \tag{4.22}$$

Now, using (4.18) and (4.19), we get that

$$E(t_j-t_\ell)^2 = \underset{\sim}{\ell}' \underset{\sim}{A}_{(j,\ell)}\underset{\sim}{\ell} > 0, \tag{4.23}$$

for at least one pair (j,ℓ), $j\neq\ell=1,\ldots,r$. As $E(t_j-t_\ell)^2 \leq 2[E(t_j^2)+E(t_\ell^2)]$, (4.23) implies that $E(t_j^2) > 0$ for at least one $j=1,\ldots,r$. Again, for the specific (j,ℓ) for which (4.23) holds, we may assume without any loss of generality that $E(t_\ell^2) \leq E(t_j^2)$, $E(t_j^2) > 0$, and thus, we require only to show that $E(t_jt_\ell) < E(t_j^2)$. If $E(t_\ell^2) = 0$, the proof is evident, while, if $E(t_\ell^2) > 0$, we have from (4.23) $2E(t_jt_\ell) < E(t_j^2) + E(t_\ell^2) < 2E(t_j^2)$. Hence, (4.22) can not hold for all $j,\ell=1,\ldots r$, if (4.19) holds. Consequently, (4.21) is strictly positive.

Hence, the theorem.

It may be noted that (4.19) really implies that the vector $(\underset{\sim}{Z}_{ij} - \underset{\sim}{Z}_{i\ell})$ is of full rank for at least one $j\neq\ell=1,\ldots,r$.

THEOREM 4.2. Under the assumptions 1 to 5, $\underset{\sim}{V}_N(\underset{\sim}{R}_N)$, defined by (3.12), converges in probability to $\underset{\sim}{\nu}$, defined by (4.16), and hence, is positive definite, in probability.

PROOF. The proof of this theorem follows as a more or less straightforward generalization of theorem 4.2 of Puri and Sen (1966) and of theorem 4.2 of Sen (1966c). Hence, for the intended brevity of the paper, it is not considered in detail.

THEOREM 4.3. Under the assumptions 1 to 5, the permutation distribution of the statistic S_N, defined by (3.15), converges asymptotically, in probability, to a chi square distribution with p(r-1) degrees of freedom (d.f.).

PROOF. We shall first prove that under the permutation model considered in Section 3, $[n^{\frac{1}{2}}(T_{N,j}^{(k)} - \bar{E}_N^{(k)})$, $j=1,\ldots,r-1$, $k=1,\ldots,p]$ has asymptotically a p(r-1) multinormal distribution. This would be done by proving that any arbitrary linear function of these p(r-1) statistics has asymptotically a normal distribution under the permutation model of section 3. Such a linear compound can be equivalently written as (by virtue of (2.13),)

$$W_n = n^{\frac{1}{2}} \Sigma_{j=1}^{r} \Sigma_{k=1}^{p} d_{jk} T_{N,j}^{(k)} \quad \text{where } \Sigma_{j=1}^{r} d_{jk} = 0,\ k=1,\ldots,p. \tag{4.24}$$

Under assumption 2, (4.24) can be rewritten as

$$n^{\frac{1}{2}} \sum_{i=1}^{n} \{ \sum_{j=1}^{r} \sum_{k=1}^{p} d_{jk} J^{(k)} (\frac{R_{ij}^{(k)}}{N+1}) \} + o_p(1). \tag{4.25}$$

Let us then write

$$U_{N,i}(\underset{\sim}{R}_N) = \sum_{j=1}^{r} \sum_{k=1}^{p} d_{jk} J^{(k)}(\frac{R_{ij}^{(k)}}{N+1}), \quad i=1,2,\ldots,n. \tag{4.26}$$

The random variable $U_{N,i}(\underset{\sim}{R}_N)$ can have only $r!$ possible equally likely values under our permutation model. These values are obtained by permuting the r vectors $\underset{\sim}{R}_{ij}$, $j=1,\ldots,r$ (defined by (2.7),) among themselves. Thus,

$$E\{U_{N,i}(\underset{\sim}{R}_N)|\mathcal{P}_n\} = \sum_{k=1}^{p} \{\frac{1}{r} \sum_{j=1}^{r} J^{(k)}(\frac{R_{ij}^{(k)}}{N+1})\} \sum_{j=1}^{r} d_{jk} = 0, \tag{4.27}$$

for $i=1,\ldots,n$. Similarly,

$$E(U_{N,i}^2(\underset{\sim}{R}_N)|\mathcal{P}_n) = \sum_{k=1}^{p} \sum_{q=1}^{p} \sum_{j=1}^{r} \sum_{\ell=1}^{r} d_{jk} d_{\ell q} E\{J^{(k)}(\frac{R_{ij}^{(k)}}{N+1}) J^{(q)}(\frac{R_{i\ell}^{(q)}}{N+1}) | \mathcal{P}_n\}$$

$$= \sum_{k=1}^{p} \sum_{q=1}^{p} (\sum_{j=1}^{r} d_{jk} d_{jq}) \{\frac{1}{r-1} \sum_{j=1}^{r} [J^{(k)}(\frac{R_{ij}^{(k)}}{N+1}) J^{(q)}(\frac{R_{ij}^{(q)}}{N+1}) - \frac{1}{r^2}(\sum_{j=1}^{r} J^{(k)}(\frac{R_{ij}^{(k)}}{N+1})) \tag{4.28}$$

$$(\sum_{j=1}^{r} J^{(q)}(\frac{R_{ij}^{(q)}}{N+1}))]\} \qquad \text{for } i=1,\ldots,n.$$

Since the permutations of the rank-vectors within the ith block is independent of the permutations within the i'th block for $i \neq i' = 1, \ldots, n$, under our permutation model, $\{U_{N,i}(\underset{\sim}{R}_N),\ i = 1, \ldots, n\}$ are mutually independent. Hence, to prove the desired result, we may use the Berry-Esseen theorem [cf. Loeve (1962, p. 288)], according to which it is sufficient to show that

$$\lim_{n\to\infty} \frac{\Sigma_{i=1}^{n} E\{|U_{N,i}(\underset{\sim}{R}_N)|^3 | \mathcal{P}_n\}}{[\Sigma_{i=1}^{n} E(U_{N,i}^2(\underset{\sim}{R}_N) | \mathcal{P}_n)]^{3/2}} = 0. \tag{4.29}$$

From (3.11), (3.12) and (4.28), we get that

$$\frac{1}{n}\Sigma_{i=1}^{n} E(U_{N,i}^2(\underset{\sim}{R}_N) | \mathcal{P}_n) = \sum_{j=1}^{r}\sum_{k=1}^{p}\sum_{q=1}^{p} d_{jk} d_{jq} v_{kq}(\underset{\sim}{R}_N)$$

$$= \sum_{j=1}^{r} \{\sum_{k=1}^{p}\sum_{q=1}^{p} d_{jk} d_{jq} v_{kq}(\underset{\sim}{R}_N)\}$$

$$\xrightarrow{P} \sum_{j=1}^{r} \{\sum_{k=1}^{p}\sum_{q=1}^{p} d_{jk} d_{jq} \nu_{kq}\}, \tag{4.30}$$

whereby theorem 4.1 and assumption 5, the right hand side of (4.30) is a (non-zero) positive constant, for any given $(d_{jk},\ j = 1, \ldots, r,\ k = 1, \ldots, p)$. Thus, it is sufficient to show that the numerator of the left hand side of (4.29) is $o_p(N^{3/2})$, and this readily follows from assumption 3 and (4.26). Hence, under our permutation model, the first term of (4.25) has asymptotically, in probability, a normal distribution. Once this is established, we consider the quadratic form associated with the asymptotic multinormal distribution of $\{n^{\frac{1}{2}}(T_{N,j}^{(k)} - \bar{E}_N^{(k)}),\ j = 1, \ldots, r-1,\ k = 1, \ldots, p\}$, and using some well-known results on the limiting distribution of continuous functions of random variables [cf. Sverdrup (1952)], it is easily seen that under our permutation model, the statistic S_N, given by (3.15), has asymptotically, in probability, a chi square distribution with $p(r-1)$ d.f.

Hence, the theorem.

It may be noted that the permutation distribution of S_N being essentially a conditional distribution, the convergence in theorem 4.3 holds, in probability, i.e., for almost all $\underset{\sim}{Y}_N$. If we now denote by $\chi^2_{t,\epsilon}$ the upper $100\epsilon\%$ point of the chi square distribution with t d.f., then from (3.17) and theorem 4.3, we arrive at the following.

THEOREM 4.4. *$S_{N,\epsilon}(\underset{\sim}{R}_N)$ and $\gamma(\underset{\sim}{R}_N)$, defined by (3.17), converge, in probability to $\chi^2_{p(r-1),\epsilon}$ and 0, respectively.*

By virtue of theorem 4.4, the *exact permutation test*, considered in (3.17), reduces asymptotically to

$$\phi(Y_N) = \begin{cases} 1, & \text{if } S_N \geq \chi^2_{p(r-1),\epsilon} \\ 0, & \text{otherwise;} \end{cases} \tag{4.31}$$

and (4.31) will be termed henceforth the *asymptotic permutation test*.

5. ASYMPTOTIC POWER OF THE PROPOSED TESTS.

In this section we shall study the asymptotic power and power-efficiency of our proposed class of tests. This requires first of all the study of the asymptotic (unconditional) distribution of S_N, when the null hypothesis (1.7) is not necessarily true. For this study, we also adopt the same notations as in section 4, and write

$$T^{(k)}_{N,j} = \int_{-\infty}^{\infty} J^{(k)}_N \left(\frac{N}{N+1} H^{(k)}_N(x)\right) dF^{(k)}_{N[j]}(x), \tag{5.1}$$

for $j=1,\dots,r$, $k=1,\dots,p$. The statistics in (5.1) has some analogy with a class of similar statistics considered by Puri and Sen (1966). However, in this case of two way layout we are faced with n independent pr-variate observations, while in the earlier case, Puri and Sen were faced with the oneway layout involving $N(=nr)$ p-variate observations. This makes the situation somewhat more complicated in our case, and the necessary modifications will be studied here. Let us define

$$\mu_j^{(k)} = \int_{-\infty}^{\infty} J^{(k)}(H^{(k)}(x))\, dF_{[j]}^{(k)}(x), \tag{5.2}$$

for $j=1,\dots,r$, $k=1,\dots,p$. Also let

$$\beta_{jj'\cdot\ell\ell'}^{(k,q)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [F_{[j,j']}^{(k,q)}(x,y) - F_{[j]}^{(k)}(x)F_{[j']}^{(q)}(y)]\, J'^{(k)}(H^{(k)}(x))\, J'^{(q)}(H^{(q)}(y))\cdot$$

$$dF_{[\ell]}^{(k)}(x)\, dF_{[\ell']}^{(q)}(y), \tag{5.3}$$

for $j,j',\ell,\ell' = 1,\dots,r$, $k, q=1,\dots,p$, with either $j\neq j'$ or $k\neq q$ or both, while

$$\beta_{jj\cdot\ell\ell'}^{(k,k)} = \iint_{-\infty<x<y<\infty} F_{[j]}^{(k)}(x)[1-F_{[j]}^{(k)}(y)]\, J'^{(k)}(H^{(k)}(x))\, J'^{(k)}(H^{(k)}(y))\, dF_{[\ell]}^{(k)}(x)dF_{[\ell']}^{(k)}(y)$$

$$+ \iint_{-\infty<x<y<\infty} F_{[j]}^{(k)}(x)[1-F_{[j]}^{(k)}(y)]\, J'^{(k)}(H^{(k)}(x))\, J'^{(k)}(H^{(k)}(y))\, dF_{[\ell']}^{(k)}(x)dF_{[\ell]}^{(k)}(y),$$

$$\text{for } j=1,\dots,r,\ k=1,\dots,p,\ \ell,\ell' = 1,\dots,r. \tag{5.4}$$

Finally, let

$$\beta_{\cdot jj'\cdot}^{(k,q)} = \frac{1}{r}\left\{\sum_{\ell=1}^{r}\sum_{\ell'=1}^{r}\left[\beta_{jj'\cdot\ell\ell'}^{(k,q)} + \beta_{\ell\ell'\cdot jj'}^{(k,q)} - \beta_{\ell j'\cdot j\ell}^{(k,q)} - \beta_{j\ell'\cdot\ell j'}^{(k,q)}\right]\right\} \tag{5.5}$$

for $k,q=1,\dots,p$; $j,j' = 1,\dots,r$.

<u>THEOREM 5.1.</u> <u>If the assumptions 1,2 and 3 of section 4 hold, then for arbitrarily</u>

continuous $G(Y_{i1},\ldots,Y_{ir})$, the random variables $[N^{\frac{1}{2}}(T_{N,j}^{(k)} - \mu_j^{(k)}),\ j=1,\ldots,r,$ $k=1,\ldots,p]$ has asymptotically a multinormal distribution with a null mean vector and a dispersion matrix with elements $\rho_{jj'}^{(k,q)}$, defined by (5.5).

(It may be noted that by virtue of (2.13), (4.4) and (5.2), the above multinormal distribution will be essentially singular with a rank less than or equal to $p(r-1)$.)

PROOF. We shall present only a brief sketch of the proof, as the same will follow precisely on similar lines as in theorem 5.1 of Puri and Sen (1966) and theorem 5.1 of Sen (1966c). Proceeding precisely on the same line as in the proofs of these two theorems it can be easily shown that

$$N^{\frac{1}{2}}|(T_{N,j}^{(k)} - \mu_j^{(k)}) - (B_{j,1N}^{(k)} + B_{j,2N}^{(k)})| = o_p(1), \tag{5.6}$$

for all $j=1,\ldots,r$, $k=1,\ldots,p$, where

$$B_{j,1N}^{(k)} + B_{j,2N}^{(k)} = \frac{1}{r}\sum_{j'=1}^{r}\left\{\frac{1}{n}\sum_{i=1}^{n}\left[B_{j:j'}^{(k)}(Y_{ij}^{(k)}) - B_{j':j}^{(k)}(Y_{ij'}^{(k)})\right]\right\} \quad ; \tag{5.7}$$

$$B_{j:\ell}^{(k)}(Y_{ij}^{(k)}) = \int_{-\infty}^{\infty}\left[F_{[j](i)}^{(k)}(x) - F_{[j]}^{(k)}(x)\right] J'^{(k)}(H^{(k)}(x))\, dF_{[\ell]}^{(k)}(x) \quad ; \tag{5.8}$$

$$F_{[j](i)}^{(k)}(x) = \begin{cases} 0, & \text{if } x < Y_{ij}^{(k)} \\ 1, & \text{if } x \geq Y_{ij}^{(k)}, \end{cases} \tag{5.9}$$

for $i=1,\ldots,n$, $j,\ \ell=1,\ldots,r$, $k=1,\ldots,p$. It is therefore sufficient to show that for any arbitrary non-null $\underset{\sim}{\delta} = (\delta_{11},\ldots,\delta_{pr})$, $N^{\frac{1}{2}}\sum_{j=1}^{r}\sum_{k=1}^{p}\delta_{jk}(B_{j,1N}^{(k)} + B_{j,2N}^{(k)})$ has asymptotically a normal distribution. By virtue of (5.7), the same can be written as $n^{-\frac{1}{2}}\sum_{i=1}^{n} B(\underset{\sim}{Y}_{i1},\ldots,\underset{\sim}{Y}_{ir})$, where

$$B(\underset{\sim}{Y}_{i1},\ldots,\underset{\sim}{Y}_{ir}) = r^{-\frac{1}{2}} \sum_{j=1}^{r} \sum_{\ell=1}^{r} \sum_{k=1}^{p} \delta_{jk}[B_{j:\ell}^{(k)}(Y_{ij}^{(k)}) - B_{\ell:j}^{(k)}(Y_{i\ell}^{(k)})],\ i=1,\ldots,n. \tag{5.10}$$

Since, the random variables in (5.10) are independent and identically distributed, in order to make use of the central limit theorem under the Lindeberg's condition, it is sufficient to show that these have finite second order moments. Using (5.8), it is easily seen that $E\{B(\underset{\sim}{Y}_{i1},\ldots,\underset{\sim}{Y}_{ir})\} = 0$ for all $i=1,\ldots,n$, and by virtue of (5.10), it appears to be sufficient to show that $E(|B_{j:\ell}^{(k)}(Y_{ij}^{(k)}|^2) < \infty$ for all $j,\ell=1,\ldots,r$, $k=1,\ldots,p$, $i=1,\ldots,n$. Now, under the assumption 3 of section 4, it is easily seen that for any η: $0 < \eta < \delta$ (defined by (4.8),)

$$E(|B_{j:\ell}^{(k)}(Y_{ij}^{(k)})|^{2+\eta}) < \infty, \tag{5.11}$$

uniformly in $j,\ell=1,\ldots,r$, $k=1,\ldots,p$. Hence, the desired asymptotic normality follows readily. Again, by (5.7), (5.8) and (5.9), we have

$$E(B_{j:\ell}^{(k)}(Y_{ij}^{(k)})\, B_{j':\ell'}^{(q)}(Y_{i'j'}^{(q)})) = \delta_{ii'}\, \beta_{jj':\ell\ell'}^{(k,q)}, \tag{5.12}$$

where $\delta_{i\ell'}$ is the usual Kronecker delta and $\beta_{jj':\ell\ell'}^{(k,q)}$'s are defined by (5.3) and (5.4), for $j,j',\ell,\ell'=1,\ldots,r$, $k,q=1,\ldots,p$. Hence, it is easily seen that

$$N\, E((B_{j,1N}^{(k)} + B_{j,2N}^{(k)})(B_{\ell,1N}^{(q)} + B_{\ell,2N}^{(q)})) = \beta_{j\ell}^{(k,q)}\ , \tag{5.13}$$

which is defined by (5.5), for $k,q=1,\ldots,p$, $j,\ell=1,\ldots,r$. Consequently, by (5.6), we may conclude that the dispersion matrix of the asymptotic normal distribution has elements $\beta_{j\ell}^{(k,q)}$, defined by (5.5).

Hence, the theorem.

We have already noted that the asymptotic normal distribution of theorem 5.1 is singular and of rank at most equal to $p(r-1)$. If the null hypothesis in (1.7) is true, $G(Y_{i1},\ldots,Y_{ir})$ will be a symmetric function of the r vectors, and hence it is easily seen that (i) the marginal cdf of $Y_{ij}^{(k)}$ will be the same for all $j=1,\ldots,r$, $i=1,\ldots,n$, and is denoted by $H^{(k)}(x)$ for $k=1,\ldots,p$; (ii) the marginal cdf of $(Y_{ij}^{(k)}, Y_{ij}^{(q)})$ $(k\neq q)$ will not depend on j, and is denoted by $H_1^{(k,q)}(x,y)$ for $k\neq q=1,\ldots,p$, and (iii) the marginal cdf of $(Y_{ij}^{(k)}, Y_{i\ell}^{(q)})$ $(j\neq\ell)$ will not depend on $(j\neq\ell)$, and is denoted by $H_2^{(k,q)}(x,y)$ for $j\neq\ell=1,\ldots,r$, $k, q=1,\ldots,p$. Thus, it follows from (5.3), (5.4), (4.11) through (4.14) that in this case

$$\begin{aligned}\beta_{jj':\ell\ell'}^{(k,q)} = a_{kq\cdot jj'} &= a_{kq}^{(1)}, \quad \text{if} \quad j=j'=1,\ldots,r,\\ &= a_{kq}^{(2)} \quad \text{if} \quad j\neq j'=1,\ldots,r,\end{aligned} \tag{5.14}$$

where $a_{kq}^{(1)}$ depends only on $H_1^{(k,q)}(x,y)$ and $a_{kq}^{(2)}$ on $H_2^{(k,q)}(x,y)$, respectively. Thus, from (4.15) and (5.14), we get that in this case ν_{kq}, defined by (4.15), reduces to

$$\nu_{kq} = [(r-1)/r](a_{kq}^{(1)} - a_{kq}^{(2)}), \; k,q=1,\ldots,p, \tag{5.15}$$

and

$$\beta_{j\ell}^{(k,q)} = (\delta_{j\ell}r-1)\,\nu_{kq}, \; j,\ell=1,\ldots,r, \; k,q=1,\ldots,p, \tag{5.16}$$

where $\delta_{j\ell}$ is the usual Kronecker delta. Consequently, it is easily seen that under H_o in (1.7),

$$S_N^* = n \sum_{k=1}^{p} \sum_{q=1}^{p} \nu^{kq} \sum_{j=1}^{r} (T_{N,j}^{(k)} - \mu_{.}^{(k)})(T_{N,j}^{(q)} - \mu_{.}^{(q)}) \tag{5.17}$$

(where $((\nu^{kq}))$ is the reciprocal of $((\nu_{kq}))$, and

$$\mu^{(k)} = \int_0^1 J^{(k)}(u)du, \; k=1,\ldots,p,)$$

has asymptotically a chi square distribution with p(r-1) d.f. Now, under assumption 2 of section 4

$$|N^{\frac{1}{2}}(\bar{E}_N^{(k)} - \mu^{(k)})| = o(1), \text{ for } k=1,\ldots,p, \tag{5.18}$$

and by theorem 4.2, we have under assumption 5 that

$$\underset{\sim}{V}_N(\underset{\sim}{R}_N) \xrightarrow{P} \underset{\sim}{\nu} \text{ i.e., } \underset{\sim}{V}_N^{-1}(\underset{\sim}{R}_N) \xrightarrow{P} \underset{\sim}{\nu}^{-1}. \tag{5.19}$$

Hence, from (3.15), (5.17), (5.18) and (5.19), we get that under H_o in (1.7)

$$S_N \overset{P}{\sim} S_N^*. \tag{5.20}$$

Hence, we arrive at the following.

<u>THEOREM 5.2</u>. <u>Under H_o in (1.7) and assumptions 1 to 5 of section 4, the statistic S_N in (3.15) has asymptotically a chi square distribution with p(r-1) d.f.</u>

Let now $\hat{\underset{\sim}{\nu}}$ be any consistent estimator of $\underset{\sim}{\nu}$, defined by (4.15) and (5.15). If $\hat{\underset{\sim}{\nu}}$ is positive definite and we denote its reciprocal by $\hat{\underset{\sim}{\nu}}^{-1} = ((\hat{\nu}^{kq}))$, then we can have an <u>asymptotically distribution-free test</u> based on

$$\hat{S}_N = n \sum_{k=1}^{p} \sum_{q=1}^{p} \hat{\nu}^{kq} \sum_{j=1}^{r} (T_{N,j}^{(k)} - \bar{E}_N^{(k)})(T_{N,j}^{(q)} - \bar{E}_N^{(q)}). \tag{5.21}$$

Since, $\hat{S}_N$ can be shown to have the chi square distribution with p(r-1) d.f., when H_o in (1.7) holds, the test function may be proposed as

$$\hat{\phi}(\underset{\sim}{Y}_N) = \begin{cases} 1, & \text{if } \hat{S}_N > \chi^2_{p(r-1),\epsilon} \\ 0, & \text{otherwise.} \end{cases} \tag{5.22}$$

We shall now consider the power properties of the permutation test in (3.17) and (4.31) and the large sample test in (5.22). We shall obtain certain power-equivalence relations among these tests, and compare them with the parametric tests referred to in Section one.

By virtue of theorem 5.1, it can be shown that if the linear model (1.5) holds but the null hypothesis (1.7) is not true, then $(\mu_j^{(k)} - \bar{E}_N^{(k)})$, $j=1,\ldots,r$, $k=1,\ldots,p$, can not all converge to zero as $N\to\infty$, and hence, S_N, defined by (3.15), will be stochastically indefinitely large, as N increases. Consequently, the tests considered will be all consistent. Thus, for any given $(\tau_1,\ldots,\tau_r)$ in (1.6), (not all null), the power of the test (3.17) or (4.31) or (5.22) will be asymptotically equal to unity. Hence, forthe study of the asymptotic power properties of the tests, we shall consider a sequence of alternative hypotheses for which the power asymptotically lies in the open interval $(\epsilon,1)$. This we specify as

$$H_N: \tau_j = N^{-\frac{1}{2}} \lambda_j, \; j=1,\ldots,r, \tag{5.23}$$

where λ_j, $j=1,\ldots,r$ are all real p-vectors, not all equal (or null). Further, for simplification of the asymptotic power function, we shall assume that the cdf $F_{[j]}^{(k)}(x)$, $F_{[j,j]}^{(k,q)}(x,y)$ and $F_{[j,\ell]}^{(k,q)}(x,y)$ are all absolutely continuous and have continuous density functions. Under $\{H_N\}$ in (5.23), we will thus have sequences of cdf's $\{F_{[j],N}^{(k)}(x)\}$ etc, defined for each N, and it is easy to verify that

$$\lim_{N=\infty} F_{[j],N}^{(k)}(x) = H^{(k)}(x) \quad \text{for all } j=1,\ldots,r, \tag{5.24}$$

$$\lim_{N=\infty} F_{[j,j]}^{(k,q)}(x,y) = H_1^{(k,q)}(x,y) \text{ for all } j=1,\ldots,r, \; k\neq q=1,\ldots,p \tag{5.25}$$

$$\lim_{N=\infty} F_{[j,\ell]}^{(k,q)}(x,y) = H_2^{(k,q)}(x,y) \quad \text{for } j\neq\ell=1,\ldots,r, \; k,q=1,\ldots,p. \tag{5.26}$$

Hence, in this case also (5.16) holds in the limit as $N\to\infty$. Also, if we define

$$\zeta_k = \int_{-\infty}^{\infty} \frac{d}{dx} J^{(k)}(H^{(x)}(x))\ dF^{(k)}(x),\ k=1,\ldots,p, \tag{5.27}$$

then, it is easy to show that

$$\lim_{N=\infty} E(N^{\frac{1}{2}}(T_{N,j}^{(k)} - \mu_{.}^{(k)})|H_N) = \lambda_j^{(k)}\zeta_k, \tag{5.28}$$

for all $j=1,\ldots,r$, $k=1,\ldots,p$. Hence, from the results of theorem 5.1 it follows that under $\{H_N\}$, S_N^* has asymptotically a noncentral chi square distribution with $p(r-1)$ d.f. and the noncentrality parameter

$$\Delta_S = \sum_{k=1}^{p}\sum_{q=1}^{p} \nu^{kq}\ \zeta_k\zeta_p\ [\frac{1}{r}\sum_{j=1}^{r}(\lambda_j^{(k)}-\bar{\lambda}^{(k)})(\lambda_j^{(q)}-\bar{\lambda}^{(q)})], \tag{5.29}$$

where

$$\bar{\lambda}^{(q)} = \Sigma_{j=1}^{r}\ \lambda_j^{(q)}/r,\ \text{for}\ q=1,\ldots,p.$$

Now, from theorem 4.2, (5.24), (5.25), (5.26) and the discussion following it, it follows that under $\{H_N\}$ also $S_N \overset{P}{\sim} S_N^*$, and hence, we have the following.

THEOREM 5.3. Under the sequence of alternatives $\{H_N\}$ in (5.23), S_N, defined by (3.15), has asymptotically a non-central chi square distribution with $p(r-1)$ d.f. and the non-centrality parameter Δ_S, defined by (5.29), provided the conditions of theorem 5.1 hold, and in addition, the marginal cdf's corresponding to the joint cdf $G(\underset{\sim}{Y}_{i1},\ldots,\underset{\sim}{Y}_{ir})$ are all absolutely continuous and have continuous density functions.

If we consider the large sample test, defined by (5.22), then it can be shown similarly that $\hat{S}_N \overset{P}{\sim} S_N^*$, under $\{H_N\}$, and hence, the conclusions of theorem 5.3 also applies to $\hat{S}_N$. Thus, the permutation test considered in sections 3 and 4 and the large sample test considered in (5.29), are asymptotically power equivalent for the sequence of alternatives $\{H_N\}$, in (5.23). As we have seen that the permutation tests are easy to define for small samples, we are now in a position to recommend the use of the same, for all sample sizes.

In the parametric case, the limiting distributions of various test-statistics for this problem have been studied by various workers, and the reader may be referred to Anderson (1958, Ch. 8·10), Rao [(1952, Ch. 7), (1965, Ch. 8)], and James (1960), among others. Most of the results relate to the null case, while it may be considerably difficult to formulate a general theory for the non-null cases, though some work has also been done on this line. For the likelihood ratio test, however, the asymptotic non-null distribution may be found without much difficulty, and for the sequence of alternatives in (5.23), this statistic can be shown to have asymptotically a non-central chi square distribution with p(r-1) d.f. and the non-centrality parameter

$$\Delta_U = \sum_{k=1}^{p} \sum_{q=1}^{p} \sigma^{kq} \{\frac{1}{r} \sum_{j=1}^{r} (\lambda_j^{(k)} - \bar{\lambda}^{(k)})(\lambda_j^{(q)} - \bar{\lambda}^{(q)})\}, \tag{5.30}$$

where $\lambda_j^{(k)}$ and $\bar{\lambda}^{(k)}$ are defined by (5.23) and (5.29), respectively, and $\underset{\sim}{\Sigma}^{-1} = ((\sigma^{kq})) = ((\sigma_{kq}))^{-1}$ is the reciprocal of the common dispersion matrix $\underset{\sim}{\Sigma}$. The comparison of Δ_S and Δ_U (for the purpose of studying asymptotic relative efficiency) poses the same problem as has been studied in some detail by Puri and Sen (1966). For intended brevity, this is therefore not reproduced again. The only remark that may be made here is that if we work with $\underset{\sim}{E}_N^{(k)}$'s (defined by (2.9), (2.10),) as the expected values of the order statistics in a sample of size N drawn from a standardized normal distribution and term the resulting test as Normal score MANOVA test for the two way lay out, then it is easily seen that for normal alternatives, this test is asymptotically power equivalent to the likelihood ratio test. In actual practice, the use of rank sums (i.e., $E_{N,\alpha}^{(k)} = \alpha/(N+1)$, $\alpha=1,\ldots,N$, $k=1,\ldots,p$) often results in a quite simplified procedure and at the same time does not involve any serious loss of efficiency. For details of these points, the reader may be referred to Puri and Sen (1966), the same argument being true in the two way layout case.

REFERENCES

Anderson, T. W. (1958): An Introduction to Multivariate Statistical Analysis. John Wiley and Sons Inc., New York.

Anderson, T. W. (1965): Some multivariate nonparametric procedures based on statistically equivalent blocks. Proc. Inter. Symp. Multivariate Analysis. (in press).

Benard, A. and Van Elteren, P. H. (1953): A generalization of the method of M-rankings. Proc. Kon. Ned. Ak. Van Wet. Ser A. 56, 358-369.

Bhapkar, V. P. (1965): Some nonparametric tests for the multivariate several sample location problem. Proc. Inter. Sym. Multivariate Analysis (in press).

Brown, G. W. and Mood, A. M. (1951): On median tests for linear hypotheses. Proc. Second Berkeley Symp. Math. Stat. Prob. II, 159-166.

Chatterjee, S. K. and Sen, P. K. (1964): Nonparametric tests for the bivariate two sample location problem. Calcutta Stat. Asso. Bull. 13, 18-58.

Chatterjee, S. K. and Sen, P. K. (1966): Nonparametric tests for the multivariate multisample location problem. S. N. Roy Memorial Volume (in press).

Chernoff, H. and Savage, I. R. (1958): Asymptotic normality and efficiency of certain nonparametric test statistics. Ann. Math. Stat. 29, 972-996.

Durbin, J. (1951): In complete blocks in ranking experiments. Bri. J. Psychol. 4, 85-90.

Friedman, M. (1937): The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J. Amer. Stat. Assoc. 32, 675-99.

Hodges, J. L., Jr. and Lehmann, E. L. (1962): Rank methods for combination of independent experiments in analysis of variance. Ann. Math. Stat. 33, 482-497.

Lehmann, E. L. and Stein, C. (1949): On the theory of some nonparametric hypotheses. Ann. Math. Stat. 20, 28-45.

Loeve, M. (1962). Probability Theory. D. Van Nostrand Co. Princeton, New Jersey.

Puri, M. L. and Sen, P. K. (1966): On a class of multivariate multi-sample rank order tests. Sankhyā (Submitted).

Rao, C. R. (1952). Advanced Statistical Methods in Biometric Research. John Wiley and Sons Inc., New York.

Rao, C. R. (1965). Linear Statistical Inference and Its Applications. John Wiley and Sons Inc., New York.

Sen, P. K. (1965): On a class of bivariate two sample nonparametric tests. Proc. 5th Berkeley Symp. Math. Stat. Prob. (in press).

Sen, P. K. (1966 a): On a class of multisample multivariate non-parametric tests. Ann. Inst. Stat. Math. (Submitted).

Sen, P. K. (1966 b): Rank methods for combination of independent experiments in MANOVA. Part one: two treatment multi-response case. S. N. Roy Memorial Volume. (in press).

Sen, P. K. (1966 c): On some non-parametric generalizations of Wilks' tests for H_M, H_{VC} and H_{MVC}, I. Inst. Stat. Univ. North Carolina, mimeo series report 468.

Sverdrup, E. (1952). On the limit distribution of a continuous function of random variables. Skand. Actdskft. 35, 1-10.

TESTS FOR OUTLIERS*

H. A. David
University of North Carolina at Chapel Hill
North Carolina

1. INTRODUCTION. The proper treatment of outliers has long been a subject for study. It is an active area now and likely to remain so for some time to come. The reason for this is easy to see: there seems no limit to the multitude of different situations in which outliers are important. Excellent recent surveys of the subject have been given by Dixon (1962) and Ferguson (1961a).

It is useful to distinguish three aims of procedures designed to deal with outliers:

(a) to screen data in routine fashion preparatory to analysis (this includes but is more general than the old problem of 'rejection of outliers');

(b) to sound an alarm that outliers are present, thus indicating the need for closer study of the data-generating process;

(c) to pinpoint observations which may be of special interest just because they are extreme.

Numerous test-statistics have been devised, mostly from intuitive considerations, and their percentage points tabulated on the assumption of a common normal parent population. However, much more needs to be known about the performance of the various statistics in use for the non-null situation when outliers are in fact present. We will here be concerned primarily with cases (b) and (c). This is not in any way to belittle the importance of case (a), and I will just mention a recent proposal by Anscombe (1966). If the primary aim of screening data is the estimation of parameters, Anscombe suggests a two-stage procedure: (1) Apply the appropriate test for outliers at a very stringent level of significance, so stringent that good observations will very seldom be rejected. The purpose of this is to get rid only of wild observations very far removed from the main stream. (2) Apply the same outlier test again to the reduced data but now at quite a moderate level of significance. This time, unlike the preceding stage, observations found to be outlying will not be rejected

*Research supported by the Army Research Office, Durham.

but rather given reduced weight in the estimation of parameters. This second stage process is commonly termed Winsorization.

This kind of approach promises to be fruitful for the situation of case (a) although its properties are by no means easy to investigate.

We shall begin with a discussion of several measures of performance, including the power function, of some well-known test statistics relevant to cases (b) and (c). We assume that the underlying variation is normal and consider in some detail the case where a single true outlier is present which differs from the remaining observations in mean only. Some limited results will also be given for the case when two observations are from a common outlying or contaminating distribution. The situation of an unknown number of outliers is briefly treated. Some of the statistics we use, and others, have been studied under these assumptions by experimental sampling.

2. <u>MEASURES OF PERFORMANCE</u>. Let x_i $(i = 1, 2, \ldots, n)$ be independent normal variates, x_i having mean μ_i and variance σ^2. On the null hypothesis of homogeneity, H_0, the μ_i are all equal to some unspecified value μ. We shall consider alternatives H_a representing a shift or slippage to the right in one or a small fraction of the μ_i. A suitable class of statistics for testing H_0 against H_a is of the form

$$v = \max_i d_i , \tag{1}$$

where d_i is the difference, $x_i - \bar{x}$, appropriately divided. Of particular interest are the following special cases of v corresponding to various degrees of information on σ:

(i) standardized extreme deviate (from the sample mean)

$$v_1 = \max(x_i - \bar{x}) / \sigma = (x_{max} - \bar{x}) / \sigma \;;$$

(ii) internally studentized extreme deviate

$$v_2 = (x_{max} - \bar{x}) / s, \quad s^2 = \Sigma (x_i - \bar{x})^2 / (n - 1);$$

(iii) externally studentized extreme deviate

$$v_3 = (x_{max} - \bar{x}) / s_\nu \;,$$

where s_ν is a root-mean-square estimate of σ based on ν degrees of freedom and independent of the x_i;

(iv) internally and externally studentized extreme deviate

$$v_4 = (x_{max} - \bar{x})/s_p ,$$

where $$s_p^2 = [\Sigma(x_i - \bar{x})^2 + \nu s_\nu^2]/(n - 1 + \nu) .$$

v_1 is appropriate when σ is known, v_2 in the absence of any knowledge of σ. The use of v_3 and v_4 requires an independent estimate of σ. In v_4 such external information is combined with internal information by means of a pooled estimate of σ^2. Formally v_1 and v_2 may be regarded as the special cases, $\nu = \infty$ and $\nu = 0$, of v_4.

If v_α is the upper α significance point of the null distribution of v, then H_o is rejected for $v > v_\alpha$, and the warning required in case (b) of the Introduction is thereby given. For (c) this must be followed up by declaring one or more of the x_i to be outliers, for example, those x_i for which d_i exceeds v.

Because of the difficulty of dealing with more general alternatives we shall first suppose that just one of the observations — we do not know which — is a true outlier and has mean $\mu + \lambda(\lambda > 0)$. In the formulation of slippage tests we may say that H_a consists of n mutually exclusive hypotheses of which the ith, H_i, specifies that

$$\mu_i = \mu + \lambda, \ \mu_j = \mu \qquad (j = 1, 2, \ldots, i - 1, i + 1, \ldots, n) .$$

It is known (e.g. Kudo, 1956) that in this situation v_4 (and hence v_1, v_2 when applicable) has the desirable optimal property of maximizing the probability of rejecting a true outlier in the class of all level α tests which are invariant under the transformation $x_i' = ax_i + b(a > 0)$ applied to each x_i.

It is clear that a reasonable measure of the performance of any of the v-statistics can depend only on the sample size n and the ratio λ/σ. In particular, the measure must be independent of which of the H_i holds. For convenience we therefore take $i = 1$, and also $\sigma = 1$. The following measures come to mind:

1. Power function $P_1 = Pr(v > v_\alpha \mid H_1)$.

2. Probability that the observation x_1 from the slipped population is significantly large

$$P_2 = Pr(d_1 > v_\alpha \mid H_1).$$

3. Probability that x_1 is significantly large and the largest in the sample

$$P_3 = Pr(d_1 > v_\alpha,\ x_1 > x_2,\ x_3,\ \ldots,\ x_n \mid H_1).$$

4. Probability that only x_1 is significant

$$P_4 = Pr(d_1 > v_\alpha,\ d_2,\ d_3,\ \ldots,\ d_n < v_\alpha \mid H_1).$$

5. (Dixon, 1950) Probability that x_1 is significantly large given that it is the largest in the sample

$$P_5 = Pr(d_1 > v_\alpha \mid x_1 > x_2,\ \ldots,\ x_n;\ H_1).$$

We see that

(2) $$P_1 \geq P_2 \geq P_3 \geq P_4$$

and also that

(3) $$P_5 = P_3/Pr(x_1 > x_2,\ x_3,\ \ldots,\ x_n)\ ,$$

where the probability in the denominator has been tabulated by Teichroew (1955) for $n \leq 10$.

It can be shown that

$$P_2 \leq P_1 \leq P_2 + \alpha\ ,$$

provided $n < 2/\alpha$; in fact, a somewhat stronger general inequality holds (David and Paulson, 1965). Also for v_2 one has $P_2 = P_3 = P_4$. We therefore confine attention to P_2 as the most convenient measure.

The graphs of Figure 1 show inter alia just how much is added to the value of P_2 by the use of v_4 rather than v_3 in the present case of a single true outlier. Of course, the gain is highest when the internal information on σ^2 is large compared to the external information, i.e. when $n - 1$ is large compared to ν. However, there are indications that internal degrees of freedom are less valuable than external ones. Thus for $n = 6$ the solid curve $\nu = 5$ lies well above the dotted curve $\nu = 0$ although in both cases there is a total of 5 D.F.

3. A SEQUENTIAL PROCEDURE. It will be a rare occasion when we actually know the number of outliers for which to test. Ideally we might wish to proceed sequentially as follows:

> Apply a certain test-statistic to the sample of n. If significance is obtained eliminate the most extreme observation and apply the same test-statistic to the reduced sample of n-1, adjusting the significance point to the new sample size. If significance holds again, repeat the procedure until the test-statistic ceases to be significantly large.

We consider now such a sequential procedure for v_1, the case where σ is known and may be taken equal to unity. To this end note the following easily proved algebraic results:

(a) $x_i - \bar{x} = \frac{n-1}{n}(x_i - \bar{x}_i)$, where $\bar{x}_i = \sum_{\substack{t=1 \\ t \neq i}}^{n} x_t/(n-1)$.

(b) $x_i - \bar{x}_j \gtrless x_j - \bar{x}_i$ according as $x_i - \bar{x} \gtrless x_j - \bar{x}$.

(c) $x_i - \bar{x} + \frac{1}{n-1}(x_j - \bar{x}) = x_i - \bar{x}_j$.

Also when the x_i's are normally distributed,

(d) $x_i - \bar{x}$ and $x_j - \bar{x}_i$ are statistically independent.

(e) $v_{1,\alpha}^{(n)}$, the upper α significance point of v_1 in samples of n, is an increasing function of n.

From (c) and (d) we see that the joint occurrence of

$$(4) \qquad x_i - \bar{x} > v_{1,\alpha}^{(n)} \ , \ x_j - \bar{x} > v_{1,\alpha}^{(n)}$$

implies

$$(f) \quad x_i - \bar{x}_j > v_{1,\alpha}^{(n)} > v_{1,\alpha}^{(n-1)} \ ,$$

and by symmetry that $x_j - \bar{x}_i > v_{1,\alpha}^{(n-1)}$.

This result means that we do not have to take the above procedure too literally: if (4) holds we can immediately declare both x_i and x_j to be outliers, and next apply our test-statistic to the remaining sample of n-2, etc.

To evaluate the performance of this procedure we consider a rather special case: two observations x_i and x_j (i, j unknown) are from a contaminating $N(\mu + \lambda, 1)$ $(\lambda > 0)$ population, the remaining $n - 2$ are from $N(\mu, 1)$. This is a reasonable model for the situation when a common source is responsible for the shift in the two observations. Any acceptable measure of performance will not depend on i and j which we take to be 1 and 2. We consider the following measures:

1. Probability that at least one of x_1, x_2 is significantly large:

$$\Pi_1 = \Pr\{\max(x_1 - \bar{x}, x_2 - \bar{x}) > v_{1,\alpha}^{(n)}\} .$$

2. Probability that both x_1, x_2 are significant in a 2-stage procedure:

$$\Pi_2 = \Pr\{\max(x_1 - \bar{x}, x_2 - \bar{x}) > v_{1,\alpha}^{(n)} , \min(x_1 - \bar{x}_2, x_2 - \bar{x}_1) > v_{1-\alpha}^{(n-1)}\}.$$

3. Probability that both x_1, x_2 are significant at the first stage:

$$\Pi_3 = \Pr\{x_1 - \bar{x} > v_{1,\alpha}^{(n)} , x_2 - \bar{x} > v_{1,\alpha}^{(n)}\} .$$

(In these measures we are not concerned with the possibility that good observations may also be declared outliers.) It is clear that $\Pi_1 > \Pi_2 > \Pi_3$. Π_2 may be found from

$$\Pi_2 = \Pr\{x_1 - \bar{x} > v_{1,\alpha}^{(n)},\ x_2 - \bar{x}_1 > v_{1,\alpha}^{(n-1)}\}$$

$$+ \Pr\{x_2 - \bar{x} > v_{1,\alpha}^{(n)},\ x_1 - \bar{x}_2 > v_{1,\alpha}^{(n-1)}\}$$

$$- \Pr\{x_1 - \bar{x} > v_{1,\alpha}^{(n)},\ x_2 - \bar{x}_1 > v_{1,\alpha}^{(n-1)},\ x_2 - \bar{x} > v_{1,\alpha}^{(n)},\ x_1 - \bar{x}_2 > v_{1,\alpha}^{(n-1)}\}$$

$$= 2\Pr\{x_1 - \bar{x} > v_{1,\alpha}^{(n)}\}\ \Pr\{x_2 - \bar{x}_1 > v_{1,\alpha}^{(n-1)}\}$$

$$- \Pr\{x_1 - \bar{x} > v_{1,\alpha}^{(n)},\ x_2 - \bar{x} > v_{1,\alpha}^{(n)}\} \text{ by (d) and (f) .}$$

Hence Π_2 as well as Π_1 and Π_3 can be evaluated from tables of the univariate and bivariate normal distribution function. Figure 2 gives some numerical results comparison being also made with the earlier probability (P_2 for $\nu = \infty$) of detecting a single outlier when only one is present. The difference between Π_2 and Π_3 is seen to become less marked as n increases.

Some extensions of these results to the above cases of σ unknown are planned. It must not be supposed that the results will all be much the same. When σ has to be estimated from the sample at hand the presence of a second outlier tends to "mask" (Murphy, 1951) the effect of the first. In fact, for $\alpha = .05$ and $n \leq 14$, the probability of detecting any outliers by the use of v_2 tends to zero as $\lambda \to \infty$. (cf. Ferguson, 1961b). For finite λ the probability of detection may be quite unsatisfactorily low and the sequential process has little chance of ever getting started. The masking effect applies also to other statistics such as Dixon's $r_{10} = \dfrac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}}$ which might be used sequentially in this case. Ferguson (1961b) recommends Karl Pearson's

$$b_2 = n\Sigma(x_i - \bar{x})^4/[\Sigma(x_i - \bar{x})^2]^2$$

as a general statistic appropriate for both one and two-sided tests.

It should also be noted that in the artificial case where the above model of exactly two outliers is known to be the right alternative to H_0

the optimal procedure consists (Murphy) in rejecting the largest two observations when $(x_{(n)} + x_{(n-1)} - 2\bar{x})/s$ is too large. Percentage points are not known but are available for

$$\sum_{i=1}^{n-2} (x_{(i)} - \bar{x}_{n-1,n}) / \sum_{i=1}^{n} (x_i - \bar{x})^2 \qquad \text{Grubbs (1950).}$$

Dixon (1951, 1962) gives percentage points for several of his r-statistics, e.g. for

$$r_{20} = \frac{x_{(n)} - x_{(n-2)}}{x_{(n)} - x_{(1)}}$$

designed as a test "for $x_{(n)}$ avoiding $x_{(n-1)}$".

Although only a fraction of the many questions of interest have been considered in this paper I hope that the need for much more detailed knowledge of the performance of tests for outliers has been demonstrated. Of course, it must never be forgotten that the problem of outliers is only partly statistical.

Section 2 of this paper is based on David and Paulson (1965) where further details are given. I am indebted to R. G. McMillan for Figure 2.

REFERENCES

1. Anscombe, F. J. (1966). Treatment of outliers in samples of size three. J. Res. Nat. Bur. Standards, 70B, 141-147.

2. David, H. A. and Paulson, A. S. (1965). The performance of several tests for outliers. Biometrika, 52, 429-436.

3. Dixon, W. J. (1950). Analysis of extreme values. Ann. Math. Statist., 21, 488-506.

4. Dixon, W. J. (1962). Rejection of observations. Contributions to Order Statistics, A. E. Sarhan and B. G. Greenberg (Eds.), 299-342. New York: John Wiley and Sons.

5. Ferguson, T. S. (1961a). Rules for rejection of outliers. Revue Inst. Int. de Stat., 29, 29-43.

6. Ferguson, T. S. (1961b). On the rejection of outliers. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1, 253-87.

7. Grubbs, F. E. (1950). Sample criteria for testing outlying observations. Ann. Math. Statist., 21, 27-58.

8. Kudo, A. (1956). On the testing of outlying observations. Sankhyā, 17, 67-76.

9. Murphy, R. B. (1951). On tests for outlying observations. Ph.D. thesis, Princeton University.

10. Teichroew, D. (1955). Probabilities Associated with Order Statistics in Samples from two Normal Populations with Equal Variance. Army Chemical Center, Maryland: Chemical Corps Engineering Agency.

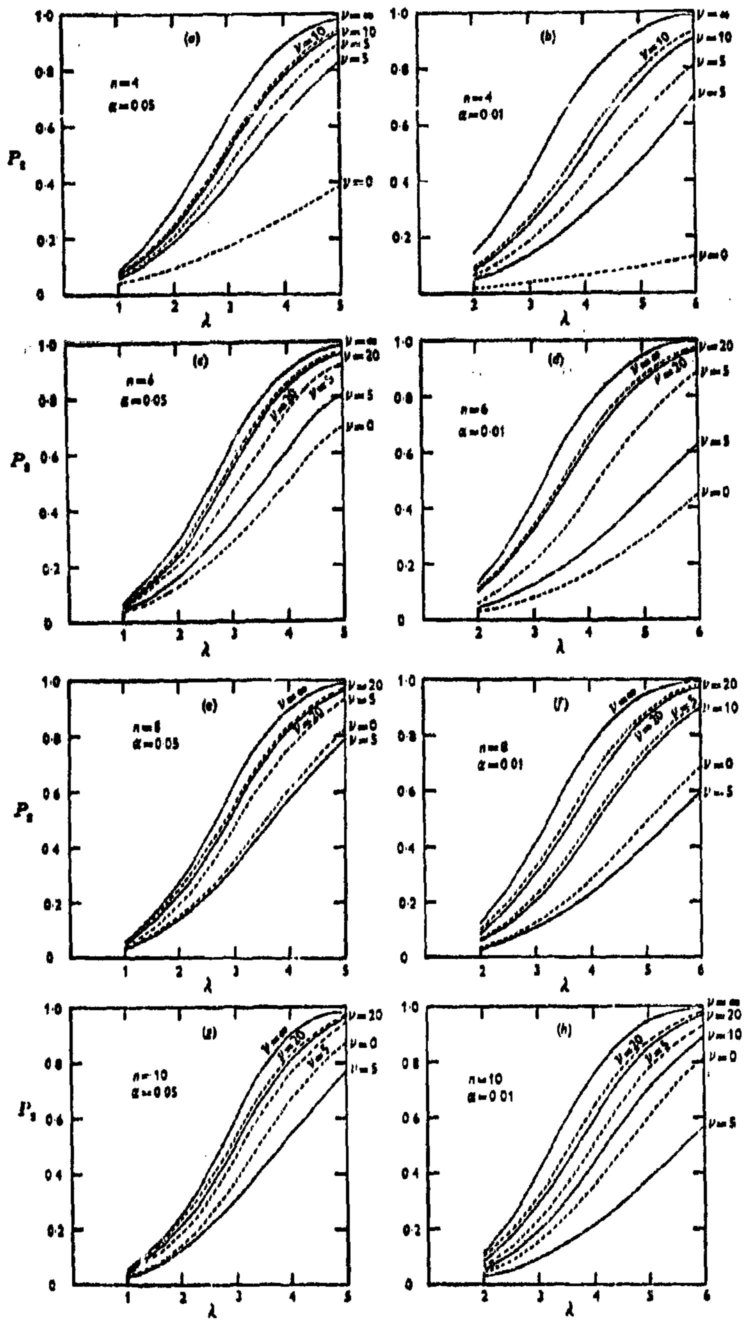


Fig. 1. Probability P_2 that the outlier x_1 is detected when v_1 to v_4 are used at level α and H_1 holds.

$\nu=\infty$: $\Pr\{(x_1 - \bar{x})/\sigma > v_{1,\alpha}\}$ $\quad$ $\nu=0$: $\Pr\{(x_1 - \bar{x})/s > v_{2,\alpha}\}$

$\cdots\cdots\nu=5,10,20$: $\Pr\{(x_1 - \bar{x})/s_\nu > v_{3,\alpha}\}$ $\quad$ $---\nu=5,10,20$: $\Pr\{(x_1 - \bar{x})/s_p > v_{4,\alpha}\}$

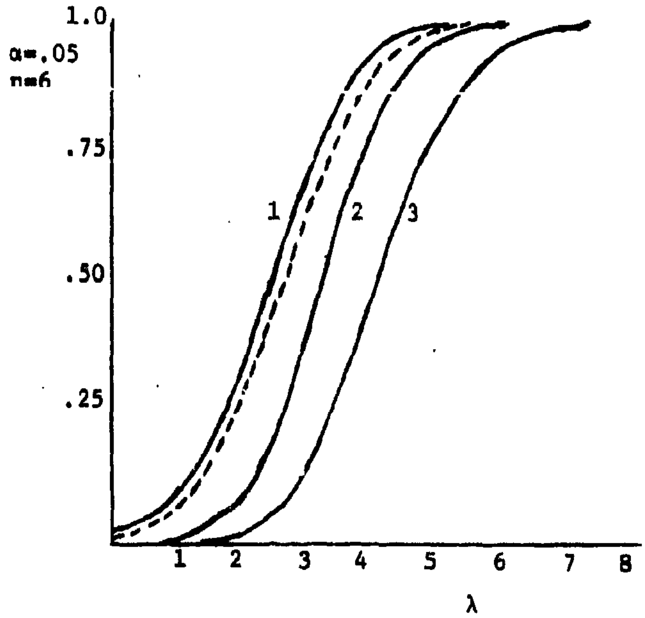
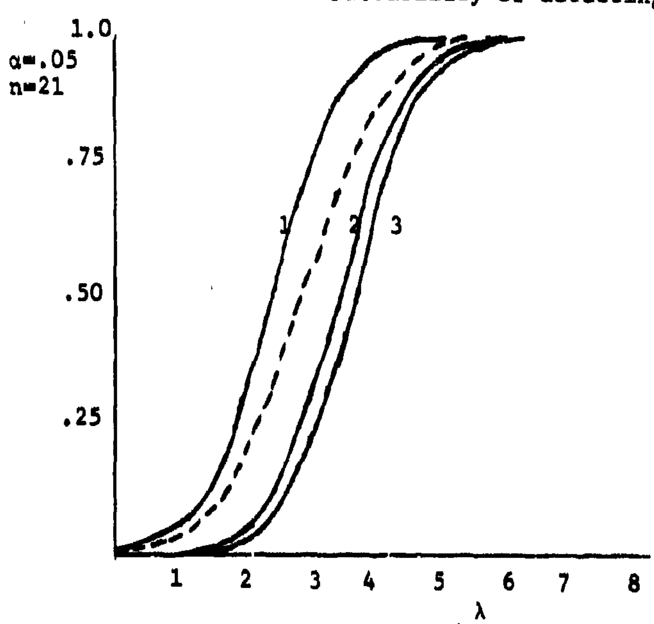


Fig. 2. Probabilities Π of detection at level α when $H_{1,2}$ holds.

1 : Π_1 2 : Π_2 3 : Π_3

Dotted line: Probability of detecting a single outlier (P_2 for $\nu=\infty$)



THE PROBABILITY OF SURVIVAL OF A SUBTERRANEAN TARGET UNDER INTENSIVE ATTACK

Bernard Harris, Herman F. Karreman
and J. Barkley Rosser
Mathematics Research Center, U. S. Army
The University of Wisconsin, Madison, Wisconsin

ABSTRACT. This report deals with the analysis of a model for studying the probability of survival of a subterranean target under an intensive attack. Most of the analysis is based on the assumption that the explosions are circularly distributed about the target and that the number of explosions is known. In the last two sections it is shown what effect a relaxation of these assumptions has on the probability of survival of the target.

The rest of this artical was reproduced photographically. It is MRC Technical Summary Report No. 653, December 1966.

1. Introduction

This report deals with the analysis of a model for studying the probability of survival of a subterranean target under an intensive attack.

The target is located below the surface at a distance d from the surface. The projection of the target on the surface will be identified as the origin in ordinary two-dimensional rectangular coordinates. K explosions occur at points $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_K$, where $\tilde{X}_i$ are independent identically distributed random vectors, $\tilde{X}_i = (X_{i1}, X_{i2})$. They will be assumed to have the bivariate normal distribution centered at the origin with zero correlation coefficient, i.e.,

$$(1) \qquad f(x_1, x_2) = (2\pi\sigma_1\sigma_2)^{-1} \exp\left(-\frac{1}{2}\left\{\frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2}\right\}\right) \quad (-\infty < x_1, x_2 < \infty ; \sigma_1, \sigma_2 > 0) .$$

The energy directly applied into the ground will be denoted by E and the seismic velocity of the rock will be denoted by c. The distances R_i, $i = 1, 2, \ldots, K$, of the explosions from the origin are consequently independent identically distributed random variables, and from (1), their common probability density function is given by

Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No.: DA-31-124-ARO-D-462.

$$(2) \qquad f_1(r) = \frac{r}{2\pi\sigma_1\sigma_2}\int_0^{2\pi} \exp\left(-\frac{r^2}{2}\left\{\frac{\cos^2\theta}{\sigma_1^2} + \frac{\sin^2\theta}{\sigma_2^2}\right\}\right) d\theta \qquad (0 \le r < \infty) \ .$$

In particular, if $\sigma_1 = \sigma_2 = \sigma$, then

$$(3) \qquad f_1(r) = r\sigma^{-2} e^{-r^2/2\sigma^2} \qquad (0 \le r < \infty) \ .$$

It will be assumed that the free field stresses P_i, $i = 1, 2, \ldots, K$, due to the explosions are given by

$$(4) \qquad P_i = \lambda c^{\alpha} E^{\beta} (R_i^2 + d^2)^{-\gamma} \qquad i = 1, 2, \ldots, K$$

where λ, α, β, and γ are positive parameters. Therefore $P_1, P_2, \ldots, P_K$ are independent and identically distributed random variables.

The following assumptions will be made about the survivability of the target.

(1) If $\max_{1 \le i \le K} P_i \ge M$, the target will fail. That is to say, M is the maximum loading from a single burst which the target can withstand without failure.

(2) If $P_i \le p_0$, no damage to the target takes place from the i^{th} burst. Further, if $P_i > p_0$, some permanent damage is done to the target, in an amount proportional to $P_i - p_0 \cdot p_0$ is the elastic limit of the target structure. Thus we define

$$(5) \qquad D_i = \begin{cases} P_i - p_0 & \text{if } P_i > p_0 \\ 0 & \text{otherwise} \ . \end{cases} \qquad (i = 1, 2, \ldots, K)$$

and D_i is known as the degradation due to the i^{th} burst. The target will also fail to survive the K explosions whenever

$$(6) \qquad \sum_{i=1}^{K} D_i \ge D^* \ .$$

Here D^* is called the maximum allowable cumulative degradation. It is assumed that the accumulation of permanent damage is additive and has no effect on the amount of damage produced by any subsequent explosion, or on M, the maximum loading from a single burst which the target can sustain.

Thus we have that the target will survive K explosions whenever

$$(7) \qquad \sum_{i=1}^{K} D_i < D^* \text{ and } \max_{1 \le i \le K} P_i < M \ .$$

The following relations between p_0, M, and D^* will be assumed to hold,

$$p_0 < M < D^* + p_0 \ .$$

Minor modifications in the analysis that follows would be needed, if this were not the case. However, it is clear that these are consistency requirements which should reasonably be satisfied by the three parameters given above.

In Section 2, we obtain the probability density function of the free field stress due to a single explosion, when $\sigma_1^2 = \sigma_2^2$. This will be exploited in Section 5, by exhibiting a number of examples to show how a straightforward examination of this probability density function may be employed in estimating the probability of survival.

Section 3 contains a discussion of techniques for estimating the probability of survival when K is fixed (i. e., not a random variable), and when $\sigma_1^2 = \sigma_2^2$ (the circular bivariate normal distribution). The approximation methods used here have been employed as the basis for a computer program.

In Section 4, some comments concerning the suitability of the model are given.

Section 6 discusses some methods which may be employed if $\sigma_1^2 \neq \sigma_2^2$ (the elliptic case). These are compared with results obtained in Section 3 for $\sigma_1^2 = \sigma_2^2$.

Finally, Section 7 provides a brief discussion of the extension of the previous results, if K is a random variable, rather than a fixed quantity.

2. The Probability Distribution of the Free Field Stress of a Single Explosion

A substantial amount of useful information may be obtained by a careful examination of the probability density function of P, the free field stress. We will derive this function in this section, and note some of its properties. These will be exploited in Section 5 of this report.

It will be convenient to define

(8) $$\theta = \lambda_C{}^{\alpha} E^{\beta} .$$

Thus, from (4), we have

(9) $$P = \theta(R^2 + d^2)^{-\gamma} \qquad (0 \leq R < \infty) .$$

$P = P(R)$ is a mapping from $[0,\infty)$ to $(0, \theta d^{-2\gamma}]$. On $[0,\infty)$, P is a monotonic decreasing function of R, and thus the inverse mapping $P^{-1}(p)$ is uniquely defined for every p, $0 < p \leq \theta d^{-2\gamma}$, and is a positive monotonic decreasing function of p. Indeed

(10) $$P^{-1}(p) = [(\theta/p)^{1/\gamma} - d^2]^{\frac{1}{2}} \qquad (0 < p \leq \theta d^{-2\gamma}) .$$

Hence

(11) $$\Pr\{P \leq p\} = \Pr\{R \geq P^{-1}(p)\} = \int_{P^{-1}(p)}^{\infty} f_1(r)\,dr \qquad (0 < p \leq \theta d^{-2\gamma})$$

where $f_1(r)$ is given by (2) or (3) and $P^{-1}(p)$ is given by (10).

We will restrict ourselves to the case $\sigma_1 = \sigma_2$ until Section 6. With this restriction, $f_1(r)$ is given by (3), and integration of (11) yields

$$(12) \qquad \Pr\{P \leq p\} \equiv G(p) = \begin{cases} 0 & (p < 0) \\ e^{-\{(\theta/p)^{1/\gamma} - d^2\}/2\sigma^2} & (0 \leq p \leq \theta d^{-2\gamma}) \\ 1 & (p > \theta d^{-2\gamma}) \end{cases} .$$

Then, the probability density function of P is given by

$$(13) \qquad g(p) = (2\gamma)^{-1} \sigma^{-2} \theta^{1/\gamma} p^{-(\gamma+1)/\gamma} \exp\{-[(\theta/p)^{1/\gamma} - d^2]/2\sigma^2\} \quad (0 \leq p \leq \theta d^{-2\gamma}) .$$

We now proceed to investigate some of the characteristics of $g(p)$.

The median M_p of $g(p)$ is readily obtained by solving

$$(14) \qquad G(M_p) = e^{-\{(\theta/M_p)^{1/\gamma} - d^2\}/2\sigma^2} = \frac{1}{2}$$

or

$$\{(\frac{\theta}{M_p})^{1/\gamma} - d^2\}/2\sigma^2 = \log 2 .$$

Hence

$$(15) \qquad M_p = \theta(2\sigma^2 \log 2 + d^2)^{-\gamma} .$$

Here $\sigma\sqrt{2 \log 2}$ is frequently referred to as the CEP (circular error probability), so that we may also write

$$M_p = \theta[(CEP)^2 + d^2]^{-\gamma} .$$

Similarly, the νth percentile of P may be obtained by setting the right hand side of (14) equal to $\nu/100$.

We can find the mode of $g(p)$, which we denote by m_p, by solving

(16) $$\frac{d \log g(m_p)}{d\, m_p} = -\frac{\gamma+1}{\gamma m_p} + \frac{\theta^{1/\gamma}}{2\gamma\sigma^2\, m_p^{(\gamma+1)/\gamma}} = 0$$

or

(17) $$m_p = \theta[2(\gamma+1)]^{-\gamma}\,\sigma^{-2\gamma} .$$

Thus, since

$$g(0) = 0,\ g(\theta d^{-2\gamma}) = (2\gamma)^{-1}\,\sigma^{-2}\,\theta^{-1}\,d^{2(\gamma+1)} > 0$$

$g(p)$ has a unique mode given by (17), whenever

$$m_p \le \theta d^{-2\gamma}$$

or equivalently,

(18) $$d \le [2(\gamma+1)]^{1/2}\,\sigma .$$

If, on the other hand, $d \ge [2(\gamma+1)]^{1/2}\sigma$, then $g(p)$ is monotone increasing, and the maximum of $g(p)$ occurs at $\theta d^{-2\gamma}$.

We conclude the characterization of $g(p)$ by evaluating the moments (both conditional and unconditional). Let A be any measurable set on $(-\infty, \infty)$; then the conditional k^{th} moment of $g(p)$, $\mu_{k,A}$ is given by

(19) $$\begin{aligned}\mu_{k,A} &= E\{P^k \mid P \in A\}\\ &= E\{\theta^k(R^2+d^2)^{-\gamma k} \mid R \in P^{-1}(A)\}\\ &= \frac{\theta^k\sigma^{-2}}{\Pr\{R \in P^{-1}(A)\}}\int_{P^{-1}(A)} (r^2+d^2)^{-\gamma k}\, e^{-r^2/2\sigma^2}\, r\,dr\end{aligned}$$

which is obtained using (3) and (9).

Two particular cases of (19) merit explicit statement.

(1) If A is an interval (p_1, p_2) with $0 \le p_1 < p_2 \le \theta d^{-2\gamma}$, then

(20) $$\mu_{k,A} = E\{P^k | p_1 \le P \le p_2\}$$

$$= \frac{\theta^k \sigma^{-2}}{e^{-r_1^2/2\sigma^2} - e^{-r_2^2/2\sigma^2}} \int_{r_1}^{r_2} (r^2+d^2)^{-\gamma k} \; e^{-r^2/2\sigma^2} \, rdr$$

where

(21) $$r_1 = [(\theta/p_2)^{1/\gamma} - d^2]^{\frac{1}{2}}, \quad r_2 = [(\theta/p_1)^{1/\gamma} - d^2]^{\frac{1}{2}} .$$

We can write (20) in terms of a tabulated function, the incomplete gamma-function, as follows. In (20), make the substitution

$$r = (2\sigma^2 y - d^2)^{\frac{1}{2}}$$

and hence

$$rdr = \sigma^2 \, dy \; .$$

Thus

$$\mu_{k,A} = \frac{\theta^k (2\sigma^2)^{-\gamma k}}{e^{-y_1} - e^{-y_2}} \int_{y_1}^{y_2} y^{-\gamma k} \; e^{-y} \, dy$$

where

(22) $$y_1 = (\frac{\theta}{p_2})^{1/\gamma} (2\sigma^2)^{-1}, \quad y_2 = (\frac{\theta}{p_1})^{1/\gamma} (2\sigma^2)^{-1} .$$

Accordingly, we now have

(23) $$E\{P^k | p_1 < P < p_2\} = \frac{\theta^k \{\Gamma(1-\gamma k\, ,\ y_1) - \Gamma(1-\gamma k\, ,\ y_2)\}}{(2\sigma^2)^{\gamma k} \; (e^{-y_1} - e^{-y_2})}$$

where

$$\Gamma(a,x) = \int_x^{\infty} e^{-t} t^{a-1} \, dt$$

is the incomplete gamma-function.

(2) If $p_1 = 0$, $p_2 = \theta d^{-2\gamma}$, then $r_1 = 0$, $r_2 = \infty$, and we obtain the unconditional k^{th} moment

$$(24) \qquad E\{P^k\} = \mu_k = \frac{e^{d^2/2\sigma^2}\,\theta^k\{\Gamma(1-\gamma k,\; d^2/2\sigma^2)\}}{(2\sigma^2)^{\gamma k}} .$$

3. Estimating the Probability of Survival

We will provide two formulas for estimating the probability of survival. The first (27) is more accurate, but substantially more difficult to compute. The second (30) should nevertheless provide a good approximation for large K. Both approximations employ the central limit theorem of probability theory.

Let T be the event described by (7). Then $Pr\{T\}$ is the probability of survival. We may write the event T as follows:

$$T = \bigcup_{m=0}^{K} \{m \text{ of } P_1, P_2, \ldots, P_K > p_0, \quad \sum_{P_i > p_0} P_i \le D^* + mp_0, \quad \max_{1\le i\le K} P_i \le M\} .$$

Thus

$$Pr\{T\} = \sum_{m=0}^{K} \binom{K}{m} Pr\{P_1, P_2, \ldots, P_m > p_0,\ P_{m+1}, P_{m+2}, \ldots, P_K < p_0,\ \sum_{i=1}^{m} P_i \le D^* + m p_0,$$

$$\max_{1\le i\le K} P_i \le M\}$$

$$= \sum_{m=0}^{K} \binom{K}{m} Pr\{\sum_{i=1}^{m} P_i \le D^* + m p_0 \mid p_0 < P_1, P_2, \ldots P_m < M\}\, e^{-(K-m)[(\theta/p_0)^{1/\gamma} - d^2]/2\sigma^2}$$

$$\times Pr\{p_0 < P_1, P_2, \ldots, P_m < M\} .$$

The last factor is evaluated as follows:

(25) $\Pr\{p_0 < P_1, P_2, \ldots, P_m < M\} =$

$$=\begin{cases} \{e^{-[(\theta/M)^{1/\gamma}-d^2]/2\sigma^2} - e^{-[(\theta/p_0)^{1/\gamma}-d^2]/2\sigma^2}\}^m & (M \le \theta d^{-2\gamma}) \\ \{1 - e^{-[(\theta/p_0)^{1/\gamma}-d^2]/2\sigma^2}\}^m & (p_0 \le \theta d^{-2\gamma} \le M) \\ 0 & (\theta d^{-2\gamma} \le p_0) \end{cases} .$$

To complete the approximation, we estimate

(26) $$\Pr\{\sum_{i=1}^{m} P_i \le D^* + m\mu_0 | p_0 < P_1, P_2, \ldots, P_m < M\}$$

by means of the central limit theorem. Noting that if $M > \theta d^{-2\gamma}$, M plays no role in conditioning, we replace (26) by

$$\Pr\{\sum_{i=1}^{m} P_i \le D^* + m p_0 | p_0 < P_1, P_2, \ldots, P_m < \min(M, \theta d^{-2\gamma})\} .$$

Then, setting $p_1 = p_0$, $p_2 = \min(M, \theta d^{-2\gamma})$ and $A = (p_1, p_2)$, we can obtain $\mu_{1,A}$ and $\mu_{2,A}$ from (20) and (23). Finally, the central limit approximation to (26) is given by

$$\Pr\{\sum_{i=1}^{m} P_i \le D^* + m p_0 | p_0 < P_1, P_2, \ldots, P_m < p_2\} = \Phi\left(\frac{D^* + m p_0 - m\mu_{1,A}}{\sqrt{m(\mu_{2,A} - (\mu_{1,A})^2)}}\right),$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt .$$

Thus we have

$$(27) \qquad \Pr\{T\} = e^{Kd^2/2\sigma^2} \sum_{m=0}^{K} \binom{K}{m} \Phi\left(\frac{D^* + m p_0 - m\mu_{1,A}}{\sqrt{m(\mu_{2,A} - (\mu_{1,A})^2)}} \right)$$

$$\times e^{(m-K)y_2} \{e^{-y_1} - e^{-y_2}\}^m .$$

where y_1 and y_2 are given by (22) with $p_1 = p_0$ and $p_2 = \min(M, \theta d^{-2\gamma})$.

The second and more tractable approximation is given by applying the central limit theorem directly to the random variables D_i, $i = 1, 2, \ldots, K$. In order to do this, we need to evaluate the first two moments of D_i.

In general, we have, for $k = 1, 2, \ldots$

$$\begin{aligned} E\{D^k\} &= E\{D^k | P < p_0\} \Pr\{P < p_0\} + E\{D^k | P \geq p_0\} \Pr\{P \geq p_0\} \\ &= E\{(P - p_0)^k | P \geq p_0\} \Pr\{P \geq p_0\} \\ &= \int_0^{r_2} (p - p_0)^k r\sigma^{-2} e^{-r^2/2\sigma^2} dr \end{aligned}$$

where r_2 is given by (21) with $p_1 = p_0$. Thus

$$\begin{aligned} E\{D^k\} &= \int_0^{r_2} (\theta(r^2+d^2)^{-\gamma} - p_0)^k r\sigma^{-2} e^{-r^2/2\sigma^2} dr \\ &= \int_0^{r_2} \sum_{j=0}^{k} \binom{k}{j} \theta^j (r^2+d^2)^{-\gamma j} (-1)^{k-j} p_0^{k-j} r\sigma^{-2} e^{-r^2/2\sigma^2} dr . \end{aligned}$$

Hence, as in the derivation of (23),

$$(28) \quad E\{D^k\} = \sum_{j=0}^{k} \frac{\binom{k}{j} \theta^j e^{d^2/2\sigma^2}}{(2\sigma^2)^{\gamma j}} (-1)^{k-j} p_0^{k-j} \{\Gamma(1-\gamma j, y_1) - \Gamma(1-\gamma j, y_2)\}$$

where y_1 and y_2 are defined as in (22) with $p_1 = p_0$ and $p_2 = \theta d^{-2\gamma}$.

We now extend (28) to obtain the conditional moments of D conditioned on $\{P_i < M,\ i = 1, 2, \ldots, K;\ M > p_0\}$. Clearly,

$$E\{D^k | P < M\} = E\{D^k | P < \min(M, \theta d^{-2\gamma})\} .$$

Let $p_1 = p_0$, $p_2 = \min(M, \theta d^{-2\gamma})$ and define r_1 and r_2 by (21) and y_1 and y_2 by (22) . Then,

$$(29) \qquad E\{D^k | P > p_0\}$$

$$= e^{r_1/2\sigma_2} \int_{r_1}^{r_2} (p - p_0)^k r\sigma^{-2} e^{-r^2/2\sigma^2} dr$$

$$= \sum_{j=0}^{k} \binom{k}{j} \frac{\theta^j (-1)^{k-j} p_0^{k-j}}{(2\sigma^2)^{\gamma j} e^{-y_1}} \{\Gamma(1 - \gamma j, y_1) - \Gamma(1 - \gamma j, y_2)\} .$$

In particular, if we denote $E\{D | P < p_2\}$ by ν_1 and $E\{D^2 | P < p_2\}$ by ν_2 , then we have

$$Pr\{T\} = Pr\{\sum_{i=1}^{K} D_i < D^* | \max_{1 \le i \le K} P_i \le M\} Pr\{\max_{1 \le i \le K} P_i < M\}$$

and

$$(30) \qquad Pr\{T\} \sim \Phi\left(\frac{D^* - K\nu_1}{\sqrt{K(\nu_2 - (\nu_1)^2)}}\right) e^{-K(y_1 - d^2/2\sigma^2)} .$$

4. A Discussion of the Model

At this point, we digress briefly to note certain aspects of the assumptions which have been made.

In Sections 2 and 3, we have assumed that the number of explosions was a fixed quantity. However, it may appear more reasonable to suppose it to be a random variable. We can see this as follows. If N missiles are fired at the target then some may not explode by virtue of defects and some may be intercepted by defenses. Hence, for any given target, it may be reasonable to assume that the number of missiles which explode is a random variable whose probability distribution depends on the number of missiles fired at the target, the reliability of the missile system, and the nature and extent of the defenses of the target. In Section 7, we provide a brief analysis of this problem. The results of Sections 2 and 3 will nevertheless provide reasonable approximations to this more complicated model in a large variety of situations. In order to use these results in this manner, "K" in Sections 2 and 3 should be interpreted as the expected value of the random variable. This is accomplished by permitting K in (30) to assume arbitrary real positive values, despite the fact that in the derivation of (30), K has been presumed to be an integer.

Then we note that the assumption of the circular normal distribution, i. e. $\sigma_1 = \sigma_2$, which has been employed throughout Sections 2 and 3, may not be completely justified. The usual nature of ballistics problems would suggest that the two parameter family of probability density functions given by (1) should be more appropriate, since there seems to be no reason to assume that the two error components, distance and lateral errors, should have the same variance. This assumption is relaxed in Section 6, in which we give a brief discussion of some suggestions for treating the more general problem.

In addition, it may be noted that the model is quite sensitive to the choice of the coefficients and exponents in (4); such as $\lambda c^{\alpha} E^{\beta}$, which we have denoted by θ, for instance.

In Section 5, an example is provided, which shows that two moderately different choices of θ can produce drastically different results.

We also note that M remains constant during the entire bombardment. It would seem more reasonable to assume that if $\sum_{i=1}^{m} D_i > 0$, $m < K$, then the vulnerability to a single shock should be reduced for later shocks, since the target has already suffered some damage.

Moreover, the basic formula (4), used in computing the free field stress, appears to have certain defects. We point out in particular one defect.

If the target is located on the surface (i.e. $d = 0$), then the free field stress for a direct hit is infinite, regardless of the magnitude of E. There are many other plausible choices which might be used in place of (4) and would still approximate (4) for $d > 0$ without the defect at $d = 0$. However, we have proceeded under the assumption that (4) will give satisfactory results for those values of R, d, c, λ, α, β, γ and E which are in regions of interest to potential users of the results cited in this paper.

Whatever assumption we use in place of (4), there is still the following concern. Since (4), or its replacement, will be obtained from empirical data, we must assume that it is only approximately valid, but not exactly valid. Then let

$$\rho = \Pr\{P < M\}$$

be the exact, but unknown probability. The answer given by (4), may be denoted by $\rho + \delta$. Then, for K explosions,

$$\Pr\{P_1 < M,\ P_2 < M,\ \ldots,\ P_K < M\} = \rho^K$$

which we estimate by $(\rho + \delta)^K$. If, we compare these two quantities, we have that approximately

$$(\rho + \delta)^K \sim \rho^K e^{\delta K/\rho}$$

for δ small compared to ρ. Thus, for K large, very substantial errors may be produced. We exhibit the magnitude for one simple example. Let $\rho = .93$ $\delta = .04$ and $K = 12$. Then,

$$\rho^K = .4186 \qquad (\rho + \delta)^K = .8704$$

and

$$[(\rho + \delta)/\rho]^K = 2.079 \ .$$

Hence, even if (4) is nearly correct, so that ρ is approximated fairly well by use of (12), raising to a large power will introduce very big errors.

5. Some Illustrations

We now show how Sections 2 and 3 may be employed to analyze the problem for several ranges of parameter values, using a variety of rough approximation methods.

Example 1. For a certain subset of parameter values, $q_M = \Pr\{\max_{1 \le i \le K} P_i \ge M\}$ may be close to unity. If this is the case, it is immediately apparent that since $\Pr\{T\} \le 1 - q_M$, $\Pr\{T\}$ is close to zero.

From (12), we have, for $M \le \theta d^{-2\gamma}$

$$q_M = \Pr\{\max_{1 \le i \le K} P_i \ge M\} = \left(1 - e^{-\{(\theta/M)^{1/\gamma} - d^2\}/2\sigma^2}\right)^K \qquad (31)$$

Thus, for $0 < \epsilon < 1$, $q_M > 1 - \epsilon$ is equivalent to

$$1 - (1-\epsilon)^{1/K} \, e^{-\{(\theta/\widetilde{M})^{1/\gamma} - d^2\}/2\sigma^2} .$$

Hence, $q_M > 1-\epsilon$, whenever,

$$d^2 - 2\sigma^2 \log(1-(1-\epsilon)^{1/K}) < \left(\frac{\theta}{M}\right)^{1/\gamma} \tag{32}$$

Therefore, whenever (32) holds for sufficiently small $\epsilon > 0$, it is apparent that the probability of survival is negligible. In general, evaluating q_M (31) provides an easily computable upper bound for $\Pr\{T\}$.

Example 2. We now assume that $M \geq \theta d^{-2\gamma} = p^*$, so that M plays no role in the computation of the survival probability. If in addition, p_0 exceeds m_P then, since $g(p)$ is monotone decreasing for $p_0 \leq p \leq p^*$, we may be able to replace $g(p)$ by a simpler function, such as a linear function or an exponential function in that region. We will now briefly discuss the approximations obtained in this manner.

From (12),

$$\Pr\{P \leq p_0\} = e^{-\{(\theta/p_0)^{1/\gamma} - d^2\}/2\sigma^2} = G(p_0) . \tag{33}$$

Thus, out of K explosions, on the average, $(1-G(p_0))K$ will have a free field stress exceeding p_0.

Let

$$\hat{p} = \frac{1}{2}(p_0 + p^*) . \tag{34}$$

Then, if we expand $g(p)$ in a Taylor series about $\hat{p}$, we obtain

$$g(p) = g(\hat{p}) + (p-\hat{p})\, g'(\hat{p}) + R(p), \quad p_0 \leq p \leq p^* , \tag{35}$$

where

$$R(p) = \frac{(p-\hat{p})^2}{2} g''(\hat{\beta})$$

for some $\hat{\hat{p}}$ between $\hat{p}$ and p .

If $R(p)$ is sufficiently small, we can replace $g(p)$ by

$$\hat{g}(p) = g(\hat{\beta}) + (p-\hat{p})\, g'(p), \qquad p_0 \leq p \leq p^* . \tag{36}$$

Defining $h(p)$ by

$$h(p) = \hat{g}(p)/(1-G(p_0)), \qquad p_0 \leq p \leq p^* , \tag{37}$$

we see that $h(p)$ is approximately a probability density function and the conditional moments of P are approximately the unconditional moments of $h(p)$. That is,

$$E\{P^k \mid p_0 \leq P \leq p^*\} \sim \int_{p_0}^{p^*} p^k\, h(p)\, dp . \tag{38}$$

To evaluate the integral in (38), it is convenient to write

$$h(p) = h_1(p) + h_2(p) ,$$

where, since $h(p)$ is trapezoidal by (36) and (37), we can write

$$h_1(p) = h(p^*), \qquad p_0 \leq p \leq p^*$$

and

$$h_2(p) = h(p) - h(p^*), \qquad p_0 \leq p \leq p^* ,$$

is a linear function with $h_2(p^*) = 0$.

The following elementary results of probability theory can now be employed:

1) The moments of the rectangular distribution on $(0, b)$ are given by:

$$\mu_k = b^{-1} \int_0^b x^k\, dx = \frac{b^k}{k+1}$$

2) The moments of the triangular distribution on $(0, b)$ defined by $f(x) = 2b^{-2}(b-x)$, $0 < x < b$ are given by

$$\nu_k = \int_0^b 2x^k b^{-2}(b-x)\,dx = 2b^k/(k+2)(k+1) \ .$$

It is convenient now to identify p_0 with 0 and $p^* - p_0$ with b . That is, we define

$$\hat{h}(p - p_0) = h(p) \qquad p_0 \leq p \leq p^*$$

and

$$\hat{h}_i(p - p_0) = h_i(p) \qquad i = 1, 2;\ p_0 \leq p \leq p^* \ .$$

Then, it is clear that there is a constant ζ, $0 < \zeta < 1$, namely

$$\zeta = \int_{p_0}^{p^*} h_1(p)\,dp = h(p^*)(p^* - p_0)$$

such that $\zeta^{-1}\hat{h}_1(q)$ is the rectangular distribution on $(0, p^* - p_0)$ and $(1-\zeta)^{-1}\hat{h}_2(q)$ is approximately the triangular distribution on $(0, p^* - p_0)$. Hence, the moments of $\hat{h}(q)$ are approximately given by

$$\alpha_k = \frac{\zeta(p^* - p_0)^k}{k+1} + \frac{(1-\zeta)\,2(p^* - p_0)^k}{(k+1)(k+2)} = \frac{(p^* - p_0)^k}{(k+1)}\left\{\frac{\zeta k+2}{k+2}\right\} \ .$$

Thus, for $k = 1$, we have

$$E\{P \mid p_0 < P < p^*\} = p_0 + \alpha_1 .$$

We can interpret the above calculations as follows. About $(1-G(p_0))K$ explosions will exceed p_0, and of these, the average force will be about $p_0 + \alpha_1$. Consequently the average degradation per explosion exceeding p_0 will be about α_1. Hence, the probability of survival $Pr(T)$ will approximately satisfy

$$Pr\{T\} \geq .5$$

if

$$D^* \geq \alpha_1 (1 - G(p_0))K$$

and less than $.5$ otherwise.

A more refined estimate of the probability of survival can be obtained by computing the variance of P conditioned on $p_0 < P < p^*$. Since the variance is translation invariant, i. e. $\sigma_x^2 = \sigma_{x-a}^2$ for all real numbers a, the variance is given by $\alpha_2 - \alpha_1^2$, and hence

$$\sigma^2_{P \mid p_0 < P < p^*} = \frac{(p^* - p_0)^2}{3} \left\{ \frac{2\zeta + 2}{4} \right\} - \alpha_1^2 .$$

We can now apply the central limit theorem, obtaining

$$Pr\{T\} \sim \Phi \left(\frac{D^* - \alpha_1 (1 - G(p_0))K}{\sqrt{(1 - G(p_0))K \sigma^2_{P \mid p_0 < P < p^*}}} \right) .$$

An alternative procedure which leads to estimating $g(p)$ by an exponential function can be constructed as follows.

Expand $\dfrac{d \log g(p)}{dp}$ given in (16) in a Taylor series about β obtaining,

$$\frac{d \log g(p)}{dp} \sim \tau + \omega(p - \hat{p}) .$$

Solving the indicated differential equation, we have

$$g(p) \sim e^{\kappa + \tau p + (\omega/2)(p - \hat{p})^2} .$$

If we can assume that the second degree term in the exponent can be ignored, then $g(p)$ has an exponential approximation. The conditional moments can now be readily computed and the central limit theorem can be applied precisely in the same manner as above. The specific details are omitted.

Example 3. This example is introduced to give some indication of the sensitivity of the probability of survival to changes in the parameters θ and p_0 .

It is apparent from (31) that we can choose θ so that $M < p^*$ and $q_M > 1-\epsilon$ for any $\epsilon > 0$, so that the probability of survival will not exceed ϵ . Now reduce θ so that $M = p^*$ and hence M plays no role in the analysis. Then, the damage per explosion exceeding p_0 is bounded by $p^* - p_0$ and the proportion of explosions that exceed p_0 is given by $1 - G(p_0)$. Thus, the average total cumulative degradation can not exceed $K(p^* - p_0)(1 - G(p_0))$ which for suitable choice of p_0, can be made less than D^*; and hence $\Pr\{T\}$ can be made arbitrarily close to unity.

6. Estimating the Probability of Survival when $\sigma_1^2 \neq \sigma_2^2$

From (2), we note that the marginal distribution of r, $f_1(r)$ cannot be obtained in closed form in this case, and consequently, the marginal distribution of P can not be written in closed form either. Hence, in this case, we must resort to numerical integration. This section will therefore be devoted to a brief discussion of our ideas in this direction, and to the manner in which they may be exploited to obtain estimates of the probability of survival.

Consider the integrand on the right of (2), and note that for $\sigma_0^2 = \max(\sigma_1^2, \sigma_2^2)$, we have

$$\frac{r^2}{2}\left(\frac{\cos^2\theta}{\sigma_1^2} + \frac{\sin^2\theta}{\sigma_2^2}\right) \geq \frac{r^2}{2}\left(\frac{\cos^2\theta + \sin^2\theta}{\sigma_0^2}\right) = \frac{r^2}{2\sigma_0^2} \quad .$$

Thus for r^2/σ_0^2 sufficiently large, the integrand on the right of (2) does not provide any appreciable contribution to $f_1(r)$, and for purposes of integration, we can replace $f_1(r)$ by zero.

In brief, for any bounded function $g(r)$, there is a real number S, such that we can replace

$$\int_0^\infty g(r) f_1(r)\, dr \quad \text{by} \quad \int_0^S g(r) f_1(r)\, dr \ .$$

We will therefore evaluate $f_1(r)$ numerically for a sufficiently dense set of r values, $0 < r < S$, so that integrations of the type denoted above can be evaluated by numerical methods (Simpson's rule, for example) , with sufficient accuracy for our purposes.

Since the integrand in (2) depends on θ only through $\sin^2\theta$ and $\cos^2\theta$ for each fixed r we can choose a uniformly spaced and sufficiently dense set of θ values in $0 \leq \theta \leq \pi/2$ to evaluate $f_1(r)$ numerically.

Then, using (11), we compute

(39)
$$\Pr\{P \leq p_0\} = \int_{r_2}^{S} f_1(r)\,dr$$

and

(40)
$$\Pr\{P \leq M\} = \int_{r_1}^{S} f_1(r)\,dr$$

where

(41)
$$r_1 = \begin{cases} [(\theta/M)^{1/\gamma} - d^2]^{1/2}, & M \leq \theta d^{-2\gamma} \\ 0, & M \geq \theta d^{-2\gamma} \end{cases}$$

and

(42)
$$r_2 = \begin{cases} [(\theta/p_0)^{1/\gamma} - d^2]^{1/2} & 0 \leq p_0 \leq \min(M, \theta d^{-2\gamma}) \\ 0 & \text{otherwise} \end{cases}.$$

We now proceed, much in the same manner as in section 3, by applying the central limit theorem to the random variables $D_1, D_2, \ldots, D_K$, and therefore obtaining the analogue of (30). In order to do so, it is necessary to compute the conditional first two moments of the degradations D_i, given that $P < M$. Hence, we readily have that

$$E\{D^k | P < M\} = E\{D^k | R > r_1\}$$

$$= \left\{ \int_{r_1}^{r_2} (p-p_0)^k f_1(r)\,dr \right\} \left\{ \int_{r_1}^{S} f_1(r)\,dr \right\}^{-1} .$$

Hence

$$(43) \quad E\{D^k | P < M\} \sim \left\{ \int_{r_1}^{r_2} [\theta(r^2+d^2)^{-\gamma} - p_0]^k f_1(r)\,dr \right\} \left\{ \int_{r_1}^{S} f_1(r)\,dr \right\}^{-1} .$$

Designating $E\{D | P < M\}$ by β_1 and $E\{D^2 | P < M\}$ by β_2, we can now write the analog of (30), i.e.,

$$(44) \quad Pr\{T\} \sim \Phi\left(\frac{D^* - K\beta_1}{\sqrt{K(\beta_2 - (\beta_1)^2)}} \right) (Pr\{P < M\})^K .$$

Two suggestions for applying the methods of sections 2 and 3 have been considered. In one of these, we compute

$$E(R) = \int_0^{\infty} r f_1(r)\,dr$$

and equate this to $\frac{1}{2}\sqrt{2\pi}\,\sigma$. Then, the solution for σ,

$$(45) \quad \sigma = 2E(R)/\sqrt{2\pi} \sim 0.85776\ E(R)$$

can be used in (3) to obtain an approximation to $f_1(r)$ which avoids the complications of this section.

Alternatively, one may consider finding R^* such that

$$\int_0^{R^*} f_1(r)\,dr = \frac{1}{2}$$

and equate R^* to the median of $f_1(r)$ in (3), i.e. set

$$R^{*2} = 2(\log 2)\,\sigma^2 \tag{46}$$

and use the value of σ^2 thus obtained in (3). Some numerical comparisons have been made between the results of (44) and those obtained by using (45) and (46) to simplify the problem.

It was noted that the discrepancies are substantial, suggesting that the two proposed approximations are not very good. A careful examination of the discrepancies shows that the estimation of $\Pr(P < M)$ is fairly good for a single explosion, but the exponentiation for K explosions produces large errors; this phenomenon was previously noted in section 4.

7. Estimating the Probability of Survival When K is a Random Variable

In the preceding sections, it was tacitly assumed that K, the number of missiles that penetrate the defenses and explode in the neighborhood of the target is a fixed quantity. The purpose of this section is to give some idea of the extent to which the probability of survival of the target is affected by allowing K to be a random variable rather than a fixed quantity. To this extent a number of computations have been performed in which K is a random variable with a probability distribution $p_N(k)$, where

$$\Pr\{k \text{ missiles explode} \mid N \text{ sent}\} = \Pr\{K = k\} = p_N(k) \; .$$

We consider two possible models which lead to the following different choices of $p_N(k)$.

(i) the binomial distribution

(ii) a mixture of two binomial distributions

(i) Assume the probability that each missile explodes remains the same for all missiles, and that whether a given missile explodes or not is independent of the performance of any other missile that is sent. We denote the constant probability that a missile explodes by r and hence we have

$$P\{K = k\} = p_N(k) = \binom{N}{k} r^k (1-r)^{N-k} \; , \tag{47}$$

Then, the probability of survival is given by

$$\Pr\{T\} = \sum_{k=0}^{N} p_N(k) \Pr\{\text{Survival} \mid K = k\} \tag{48}$$

Some numerical computations have been made, for $N = 17$, $r = .7$, and in which (30) has been used to estimate $Pr\{Survival \mid K = (17)(.7) = 11.9\}$, which is the expected value of K. These comparisons were made for 6 selected choices of θ and two choices of D^*, leaving σ, M, and p_0 fixed throughout (12 comparisons in all). Over the set of comparisons, it was noted that the maximum difference between the probability of survival computed using (48) and the probability of survival computed using (30) was .015, suggesting that the approximation using (30) may be quite good for a fairly large range of parameter values.

It has also been noted that the approximation tends to improve as D^* increases. This is fairly natural, since the central limit approximation employed in (30) will tend to become more accurate as D^* increases.

(ii) It is natural also to envision circumstances in which the probability of a missile exploding may change as the circumstances governing the defense of the target change. Suppose that if the defenders have been warned (for instance, with respect to the direction of approach of the missiles by the DEW line) then the defenses can eliminate about 15 out of a flight of 17 missiles on the average, but if they are not warned they can eliminate only about 2 of them on the average. Suppose further that the chance of getting such a warning is approximately 25%.

In general, from this point of view, we will get a mixture of two binomial distributions, i. e.

$$(49) \qquad p_N(k) = \zeta \binom{N}{k} r_1^k (1-r_1)^{N-k} + (1-\zeta)\binom{N}{k} r_2^k (1-r_2)^{N-k}$$

where $0 < \zeta < 1$. For the above set of circumstances, we would have $\zeta = .25$,

4

and for $N = 17$, as in part (i) we could take $r_1 \sim .12$ and $r_2 \sim .88$. (48) applies with $p_N(k)$ as given in (49).

To approximate in this case using (30), we can evaluate (30) numerically for $K = Nr_1$ and for $K = Nr_2$, and then average these two results with weights ζ and $1-\zeta$ respectively. The numerical comparisons which have been made suggest that this recommendation should have wide applicability.

SIMON AWARDED 1966 WILKS MEMORIAL MEDAL

Major General Leslie E. Simon (Ret.) received the 1966 (second) Samuel S. Wilks Memorial Medal during the Twelfth Annual Conference on Design of Experiments in Army Research, Development and Testing, which was held at the National Bureau of Standards, Gaithersburg, Maryland, 19-21 October 1966. General Simon has long been recognized both on a national and international basis for his outstanding contributions to Army statistics, reliability, quality control and promotion of statistical activities generally. General Simon was a long-standing friend of Sam Wilks and conferred with Sam on many statistical problems and activities.

The Wilks Award is given each year to a statistician and is based primarily on his contributions, either recent or past, to the advancement of scientific or technical knowledge in Army statistics, ingenious application of such knowledge, or successful activity in the fostering of cooperative scientific matters which coincidentally benefit the Army, the DOD, and the Government, as did Samuel S. Wilks himself.

Dr. Frank E. Grubbs received the initial Wilks Medal in November 1964, and Dr. John W. Tukey of Princeton University received the first Wilks Memorial Medal in October 1965 at the Eleventh Design of Experiments Conference.

The Award consists of a medal, with a profile of Professor Wilks and the name of the Award on one side, and the seal of the American Statistical Association and the name of the recipient on the other side; an honorarium related to the magnitude of the award funds donated by Mr. Rust; and a citation.

With the approval of President Frederick F. Stephan of the American Statistical Association, the Wilks Award Committee for 1966 consisted of:

Professor Robert E. Bechhofer, Cornell University

Dr. Francis G. Dressel, Duke University and the Army Research Office-Durham

Dr. Churchill Eisenhart, National Bureau of Standards

Professor Oscar Kempthorne, Iowa State University

Dr. Alexander M. Mood, U. S. Office of Education

Dr. Frank E. Grubbs, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland -- Chairman

The citation for the second Wilks medalist reads as follows:

> "To Major General Leslie E. Simon for his pioneering contributions to Quality Control, Sampling Inspection, Reliability and Army Design of Experiments, and for his timely promotion of statistical activities which have benefited not only the Army but our government and country as well."

General Simon received the second Wilks Memorial Medal at the banquet of the Twelfth Design of Experiments Conference, the presentation being made by President Frederick F. Stephan. General Simon replied as follows:

"President Stephan, Chairman Grubbs, ladies and gentlemen: I am most grateful for the honor that our Association has seen fit to bestow upon me. However, I am primarily a professional soldier; and secondarily a statistician. Thus, I had difficulty in rationalizing the bases on which my colleagues came to the conclusion that one of my statistical attainments should be so honored.

"While considering this matter, I happened to read a letter from Alfred S. Romer, President of the American Association for the Advancement of Science, that was published in the September issue of the Bulletin of that association. Two paragraphs of that letter, I believe, not only explain the place of the AAAS in the whole regime of the scientific community, but by analogy apply equally to the very important role of the American Statistical Society in the large and diverse field of statistics. Additionally, these paragraphs may be of some application to individuals. I would like to read to you these two paragraphs.

> 'When the Association was founded, well over a century ago, all American scientists could -- and did -- meet in one small hall; in those days specialization had not advanced far in any field, so that an astronomer, a chemist, a botanist could all talk more or less understandably to one another. But the number of American scientists grew constantly and specialization increased, creating a babel of often mutually unintelligible scientific tongues. In consequence, a centrifugal process set in: special societies in various fields were established; and with the continual increase in number of scientists, it long ago became impossible for any city in the country to accommodate at one time all the members of all scientific groups.

'Although many major societies now meet separately, the annual meetings of the AAAS still include technical sessions in nearly every area. Most important is the fact that the Association is the one organization appropriate for symposia and conferences in interdisciplinary areas. Still further, there are many subjects of common interest to scientists of every sort (government relations to science, for example), and the AAAS is the appropriate forum for discussion of such problems.'

"The centrifugal process described by Dr. Romer surely took place in the science of statistics quite as much as in any field of science. Furthermore, the American Statistical Association is the one organization that binds together the common interests of all the specialized statistical organizations. About twenty years ago, I had the honor of being a member of an ad hoc committee appointed to consider the future of ASA and it recommended that rendering this service should be a goal of ASA.

"The implication of Romer's remarks to individuals is one additional step in logic. The number of statisticians has increased enormously, during the last quarter century, along with concomitant gains in powerful statistical tools and increased recognition of the importance of Statistics. One who enters a field while it is in a rapidly expanding stage naturally has more opportunities for identifiable achievement than one who enters after it has become mature and more densely populated. In a mature activity, one exchanges some of the challenges of pioneering for the important, but less conspicuous satisfactions of pleasurable cooperative work with colleagues, the enjoyment of more sophisticated techniques and pride in the perfection of one's work.

"Timing one's entry into a field is only slightly more practicable than making a judicious selection of one's ancestors. as sometimes recommended by the medical profession. I made no choice. However, due to the need for better methods for solving Army problems, I happened to begin work in Statistics relatively early and under favorable circumstances for ready identification; and I cannot escape the belief that the perspective of my work is enhanced by a rather chance sequence of events similar to that described by Dr. Romer. Thus, I am doubly grateful; first, for the pleasure and satisfaction of working in a most engaging and rewarding field, and second, for the generous recognition awarded me by my colleagues."

SINGLE DEGREE OF FREEDOM ORTHOGONAL COMPONENTS OF A FACTOR AT 2^K LEVELS IN TERMS OF LINEAR COMBINATIONS OF THE 2^K CONTRASTS OF K FACTORS AT 2 LEVELS

Joseph Weinstein
Electronic Components Laboratory
U. S. Army Electronics Command
Fort Monmouth, New Jersey

INTRODUCTION. F. Yates (1937) algorithm for resolving a set of 2^K observations from a factorial experiment on K 2-level factors has many desirable properties for the data analyst. Briefly, it is easily learned, readily programmed on a computing machine, requires only the simplest arithmetic operations which limit the possibility of "blunders", and it gives estimates of the effects of all K experimental factors singly and their joint effects 2 or more at a time.

To illustrate the algorithm for those unfamiliar with it we consider the set of numbers

$$(1) \qquad X_{ij} \qquad \begin{matrix} i = 1, 2, 3, \ldots, 2^K \\ j = 0, 1, 2, \ldots, K \end{matrix}$$

where X_{io} represents the input data obtained from the factorial experiment and is structured in the standard order.

Then one computes

$$(2) \qquad X_{i,j+1} = \begin{cases} X_{(2i-1),j} + X_{2i,j} : & i=1,2,\ldots 2^{(K-1)} \\ -X_{(2i-1-2^K),j} + X_{(2i-2^K),j} : & i=2^{(K-1)}+1,\ 2^{(K-1)}+2, \ldots, 2^K \end{cases}$$

iterating until $j = K-1$ so that the contrast vector X_{ik} has been computed.

For $K = 2$ and the treatment factors A and B the input data, X_{io} can be represented by the treatment combinations: (1), a, b, ab; where the non-appearance of the small letter (a) implies that one of the levels of that factor (A) was included in the conditions giving rise to that observation and the appearance of the small letter (a) implies that the other level of the factor (A) was present. Then application of the algorithm gives

X_{io}	X_{i1}	X_{i2}
(1)	(1) + a	(1) + a + b + ab
a	b + ab	-(1) + a - b + ab
b	-(1) + a	-(1) - a + b + ab
ab	-b + ab	(1) - a - b + ab

The entries of the last column will be observed to be respectively the

$$(3) \qquad \begin{bmatrix} \text{sum of the input observations} \\ \text{the contrast of the A factor} \\ \text{the contrast of the B factor} \\ \text{the interaction contrast of A and B} \end{bmatrix} .$$

Where an experimental design involves one or more factors which are varied over a number of levels which is a power of 2 we shall show that the computational advantages of Yates Algorithm can still be retained in the data analysis by relating linear combinations of the results obtained to the desired factor effects. For a factor at 2^K levels these desired factor effects are of course

$$(4) \qquad \begin{bmatrix} \text{the sum of the observations} & & & & \\ \text{Linear effect of the factor} & & & & \\ \text{Quadratic} & '' & '' & '' & '' \\ \text{Cubic} & '' & '' & '' & '' \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (2^K-1) & '' & '' & '' & '' \end{bmatrix} .$$

Thus for a factor X which is to be varied over 4 levels K = 2 and our interest is in the first four rows above and we shall examine how they relate to the four rows of (3) which involve 2 factors A and B each at 2 levels.

Practical Solutions and General Solutions. Given that we will represent the four levels of an experimental factor W by the four treatment combinations available from 2 dummy factors A and B which range over 2 levels, the question of assignments arises (for this can be made in

4! ways). These 24 possible assignments are detailed in Figure I - "O Matrices" which relates the levels of W (w_1, w_2, w_3, w_4) to the treatment combinations of A and B by the appearance of a 1 in the row and column and 0 elsewhere.

For example

	(1)	a	b	ab
w_1	0	1	0	0
w_2	1	0	0	0
w_3	0	0	1	0
w_4	0	0	0	1

assigns w_1 to a, w_2 to (1), w_3 to b and w_4 to ab.

It is immediately apparent that practical limitations will constrain a catalog of relationships to the case of a factor at 4 levels. For the next step would involve K = 3, or a factor at 8 levels, and 8! (=40,320) possible ways of assigning the factor levels to the eight treatment combinations (1), a, b, ab, c, ac, bc, abc. However, the constraint is certainly not severe in the case of experiments with physical factors since the investigator is rarely concerned with effects higher than cubic. Furthermore the difficulty arises from the desire to catalog all cases. But, if an arbitrary assignment is made between the factor at 2^K levels and the treatment combination of K dummy factors at 2 levels each, then the following argument applies in general for all values of K for that assignment.

<u>The General Solution</u>. Given a factor W at 2^K levels and K dummy factors F_1, F_2, ---, F_k each at 2 levels which can represent 2^K treatment combinations of a full factorial experiment we may represent the assignment of treatment combinations to the levels of W as a matrix equation.

$$(5) \qquad W = OX$$

where W is the vector $(w_1, w_2, \ldots, w_{2^K})^T$.

O is a 2^K by 2^K matrix of 1's and 0's such that only one 1 can appear in any row or column

X is a vector of treatment combinations in standard order for the K 2-level factors

$$\left((1),\ f_1,\ f_2,\ f_1f_2,\ f_3,\ f_1f_3,\ldots,f_k,\ \ldots,f_1f_2f_3\cdots f_K\right)^T .$$

We represent the computations of the Yates Algorithm by a matrix operator N which is also 2^K by 2^K and define the results of operating on the input observations X_o by N as Y (the contrast vector for the sum of all the observations, the K factor effects F_1, F_2, ... F_k and the 2^K -K -1 joint effects of two or more factors). Thus,

$$(6) \qquad Y = NX$$

since N^{-1} Y = X equation (5) yields

$$(7) \qquad W = ON^{-1} Y \quad .$$

Consider that a direct method of operating on W by some matrix operator M which would yield the desired vector of factor effects (such as (4)) could be represented by Ω. Then

$$(8) \qquad \Omega = MW = MON^{-1}Y \quad \text{(From (7))} \quad .$$

Such direct operators M do exist and are in fact the contrast coefficient vectors to be found in tables of orthogonal polynomials (usually limited to components of 5th degree or less).

Since M and N exist and O can be cataloged for K = 2 (or assigned arbitrarily for K > 2) it is possible to evaluate MON^{-1} for all values of O and thus define the linear combinations of components of Y which correspond to the desired components of the real factor W.

The procedure is illustrated for K = 2 and particular values of O selected from the catalog of 24 possible values of O given in Figure I.

For K = 2 M and N are given by

$$(9) \qquad M = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \\ 1 & -1 & -1 & 1 \\ -1 & 3 & -3 & 1 \end{pmatrix} \qquad N = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

then N^{-1} can be shown to be $(1/4)\,N^{T}$ and the special cases for 0 labelled 1, and 7 will be used to evaluate MON^{-1}

$$\text{for } O_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad MO_1N^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 2 & -1 & 0 \end{bmatrix} .$$

Thus $\Omega = MO_1N^{-1}\,Y$ has components

- W_T: sum of the input date = sum of the input data
- W_L: linear contrast for W = dummy contrast A plus twice dummy contrast B
- W_Q: quadratic contrast for W = interaction contrast for dummy factors AB
- W_C: cubic contrast for W = twice dummy contrast A minus dummy contrast B .

$$\text{Similarly for } O_7 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad MO_7N^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

and the components of Ω are defined as

W_T: sum of the input data = sum of the input data

W_L: linear contrast for W = twice the dummy contrast B plus the interaction contrast of dummy factors AB

W_Q: quadratic contrast for W = dummy contrast A

W_C: cubic contrast for W = negative of dummy contrast B plus twice the interaction contrast for dummy factors AB.

The complete catalog of such relationships for the 24 possible assignment matrices O is given in Figure II "W components of MON^{-1} Y."

Finally, the equivalence of results obtained by this Extended Yates procedure and by conventional procedures for obtaining single degree of freedom contrasts for experimental data is detailed in Figures III, IV and V for the case K = 2 and assignment matrix O. Here we consider a factorial experiment in 16 runs where a factor W is at 4 levels and two factors C and D are each at 2 levels. First consider the conventional procedure. Figure III in column heading "real" lists the standard order of the real treatment combinations (in practice this would be the column of observations obtained from the experiment). The 16 columns at the right list the coefficients for multiplying the input observation on the same row such that the sum of products of the input observation by its coefficient estimates the factor contrast named at the head of that column.

Similarly, Figure IV also develops single degree of freedom contrasts from an experiment on 16 runs assumed to be a factorial experiment on four 2-level factors A, B, C, and D. Here the 16 columns at the right list the coefficients required to obtain the contrasts named at the head of the respective columns.

<u>Extended Yates Procedures</u>. Figure V is the result of combining the columns of Figure IV according to the rules of row 1 of Figure II (since O_1 was used to assign W to dummy factors A and B). It is observed that these 16 columns are exactly equivalent to those in Figure III.

O - matrices
available for assignment of A and B
treatment combinations to levels of W

1	2	3	4	5	6
1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0
0 1 0 0	0 1 0 0	0 0 0 1	0 0 0 1	0 0 1 0	0 0 1 0
0 0 1 0	0 0 0 1	0 0 1 0	0 1 0 0	0 1 0 0	0 0 0 1
0 0 0 1	0 0 1 0	0 1 0 0	0 0 1 0	0 0 0 1	0 1 0 0

7	8	9	10	11	12
0 1 0 0	0 1 0 0	0 1 0 0	0 1 0 0	0 1 0 0	0 1 0 0
1 0 0 0	1 0 0 0	0 0 0 1	0 0 0 1	0 0 1 0	0 0 1 0
0 0 1 0	0 0 0 1	0 0 1 0	1 0 0 0	1 0 0 0	0 0 0 1
0 0 0 1	0 0 1 0	1 0 0 0	0 0 1 0	0 0 0 1	1 0 0 0

13	14	15	16	17	18
0 0 1 0	0 0 1 0	0 0 1 0	0 0 1 0	0 0 1 0	0 0 1 0
0 1 0 0	0 1 0 0	0 0 0 1	0 0 0 1	1 0 0 0	1 0 0 0
1 0 0 0	0 0 0 1	1 0 0 0	0 1 0 0	0 1 0 0	0 0 0 1
0 0 0 1	1 0 0 0	0 1 0 0	1 0 0 0	0 0 0 1	0 1 0 0

19	20	21	22	23	24
0 0 0 1	0 0 0 1	0 0 0 1	0 0 0 1	0 0 0 1	0 0 0 1
0 1 0 0	0 1 0 0	1 0 0 0	1 0 0 0	0 0 1 0	0 0 1 0
0 0 1 0	1 0 0 0	0 0 1 0	0 1 0 0	0 1 0 0	1 0 0 0
1 0 0 0	0 0 1 0	0 1 0 0	0 0 1 0	1 0 0 0	0 1 0 0

Figure 1

W Components of $MON^{-1}Y$

$\underset{\sim}{\Omega}$	Total				W_L				W_Q				W_C			
	T	A	B	AB	T	A	B	AB	T	A	B	AB	T	A	B	AB
1	1					1	2					1		2	-1	
2	1						2	-1		-1					-1	-2
3	1					1		-2			-1			2		1
4	1						1	-2		-1					2	1
5	1					2	1					1		-1	2	
6	1					2		-1			-1			-1		-2
7	1						2	1		1					-1	2
8	1					-1	2					-1		-2	-1	
9	1					-2		1			-1			1		2
10	1					-2	1					-1		1	2	
11	1						1	2		1					2	-1
12	1					-1		2			-1			-2		-1
13	1					1		2			1			2		-1
14	1						-1	2		-1					-2	-1
15	1					1	-2					-1		2	1	
16	1						-2	1		-1					1	2
17	1					2		1			1			-1		2
18	1					2	-1					-1		-1	-2	
19	1					-2	-1					1		1	-2	
20	1					-2		-1			1			1		-2
21	1						-1	-2		1					-2	1
22	1					-1		-2			1			-2		1
23	1					-1	-2					1		-2	1	
24	1						-2	-1		1					1	-2

Figure II

$2^2 \times 4$ in 16 runs

Treatment st'd ordering: real	Treatment st'd ordering: dummy	Total	WL	WQ	WC	C	CWL	CWQ	CWC	D	DWL	DWQ	DWC	CD	CDWL	CDWQ	CDWC
w_1	(1)	1	-3	1	-1	-1	3	-1	1	-1	3	-1	1	1	-3	1	-1
w_2	a	1	-1	-1	3	-1	1	1	-3	-1	1	1	-3	1	-1	-1	3
w_3	b	1	1	-1	-3	-1	-1	1	3	-1	-1	1	3	1	1	-1	-3
w_4	ab	1	3	1	1	-1	-3	-1	-1	-1	-3	-1	-1	1	3	1	1
cw_1	c	1	-3	1	-1	1	-3	1	-1	-1	3	-1	1	-1	3	-1	1
cw_2	ac	1	-1	-1	3	1	-1	-1	3	-1	1	1	-3	-1	1	1	-3
cw_3	bc	1	1	-1	-3	1	1	-1	-3	-1	-1	1	3	-1	-1	1	3
cw_4	abc	1	3	1	1	1	3	1	1	-1	-3	-1	-1	-1	-3	-1	-1
dw_1	d	1	-3	1	-1	-1	3	-1	1	1	-3	1	-1	-1	3	-1	1
dw_2	ad	1	-1	-1	3	-1	1	1	-3	1	-1	-1	3	-1	1	1	-3
dw_3	bd	1	1	-1	-3	-1	-1	1	3	1	1	-1	-3	-1	-1	1	3
dw_4	abd	1	3	1	1	-1	-3	-1	-1	1	3	1	1	-1	-3	-1	-1
cdw_1	cd	1	-3	1	-1	1	-3	1	-1	1	-3	1	-1	1	-3	1	-1
cdw_2	acd	1	-1	-1	3	1	-1	-1	3	1	-1	-1	3	1	-1	-1	3
cdw_3	bcd	1	1	-1	-3	1	1	-1	-3	1	1	-1	-3	1	1	-1	-3
cdw_4	abcd	1	3	1	1	1	3	1	1	1	3	1	1	1	3	1	1

The column headings Total through CDWC are printed under the spanning heading "Single Degree of Freedom Contrast Vectors".

Assignment Matrix

<u>B</u> \ <u>A</u>	(1)	a
(1)	w_1	w_2
b	w_3	w_4

FIGURE III. Conventional Procedure for Analysis of 2^4 Factorial

Treatment st'd ordering		Linear Combinations of 1-d.f. Contrast Vectors															
real	dummy	Total	A+2B=WL	AB=WQ	2A-B=WC	C	AC+2BC=CWL	ABC=CWQ	2AC-BC=CWC	D	AD+2BD=DWL	ABD=DWQ	2AD-BD=DWC	CD	ACD+2BCD=CDWL	ABCD=CDWQ	2ACD-BCD=CDWC
		1	-3	1	-1	-1	3	-1	1	-1	3	-1	1	1	-3	1	-1
		1	-1	-1	3	-1	1	1	-3	-1	1	1	-3	1	-1	-1	3
		1	1	-1	-3	-1	-1	1	3	-1	-1	1	3	1	1	-1	-3
		1	3	1	1	-1	-3	-1	-1	-1	-3	-1	-1	1	3	1	1
		1	-3	1	-1	1	-3	1	-1	-1	3	-1	1	-1	3	-1	1
		1	-1	-1	3	1	-1	-1	3	-1	1	1	-3	-1	1	1	-3
		1	1	-1	-3	1	1	-1	-3	-1	-1	1	3	-1	-1	1	3
		1	3	1	1	1	3	1	1	-1	-3	-1	-1	-1	-3	-1	-1
		1	-3	1	-1	-1	3	-1	1	1	-3	1	-1	-1	3	-1	1
		1	-1	-1	3	-1	1	1	-3	1	-1	-1	3	-1	1	1	-3
		1	1	-1	-3	-1	-1	1	3	1	1	-1	-3	-1	-1	1	3
		1	3	1	1	-1	-3	-1	-1	1	3	1	1	-1	-3	-1	-1
		1	-3	1	-1	1	-3	1	-1	1	-3	1	-1	1	-3	1	-1
		1	-1	-1	3	1	-1	-1	3	1	-1	-1	3	1	-1	-1	3
		1	1	-1	-3	1	1	-1	-3	1	1	-1	-3	1	1	-1	-3
		1	3	1	1	1	3	1	1	1	3	1	1	1	3	1	1

FIGURE V. Extended Yates Procedure for Analysis of $2^2 \times 4$ Factorial.

CONDITIONAL EFFECTS AND INTERACTIONS IN SYMMETRICAL FACTORIAL CONFOUNDING WITH APPLICATION TO BIOLOGY

N. R. Bohidar
Biomathematics Division, Fort Detrick, Frederick, Maryland
C-E-I-R, Inc., Los Angeles, California

INTRODUCTION. The conditional effects and interactions (CE&I's) associated with a factorial experiment have the property of establishing direct and reciprocal relationship among the various main effects and thereby improving the interpretative information of such effects. They have also the virtue of alleviating, to some extent, the broad problems of interpretation of higher order interactions such as four or more factor interaction, contrasts of the type linear x quadratic x linear x cubic or AB^3C^2, etc. by assigning appropriate interpretation to their respective conditional entities. In this treatise the concept of conditional effects and interactions is introduced in consistence with the general theory and modulo notation associated with symmetrical factorial experiments. The treatment consists of algebraic definitions, determination of conditional effects and interactions for a given situation and orthogonal partition of sums of squares in general anova procedures. The problem of estimability of the CE&I's under classical confounding have been considered. Simple rules have been developed for rapid examination of the estimability of the CE&I's under confounding conditions by the application of elementary operations of theory of sets. The interpretation of CE&I's is explained by a numerical example from a biological experiment. The topics such as conditional confounding and its impact on fractional replication, fractional factorial, asymmetrical factorials, etc., are not presented in this treatise. The theorems and the proofs are heavily based on the properties of finite geometries derived from Galois fields and finite projective and Euclidean Geometry and combinatorial theorems. No proofs will be given here. Only the definitions and the salient properties will be described. The notation will be consistent with the general factorial notation.

DEFINITION OF CE&I'S AND PROBLEM OF ESTIMABILITY. The CE&I's are generated by decomposing the total dimension of the factor space into interpretable dimensions and it is expected that the problem of estimability of such effects and interactions becomes an immediate concern. The mathematical theory of factorial experimental design follows directly from the theory of linear models based on the Gauss-Markoff theorem which states that Y, the response vector with n components if expressed in terms of the following linear model,

$$Y = X\beta + e$$

where β is the column vector of p unknown parameters, X is the design matrix of dimension n x p and e is the error column vector with n components, then the best linear unbiased estimator of β is,

$$\hat{\beta} = (X'X)^{-1}X'Y$$

obtained from the solution of the following normal equations

$$X'X\hat{\beta} = X'Y$$

where $X'X$ is non-singular and there is a unique inverse associated with $X'X$. But the factorial experimental design matrix is not always of full rank and so one is interested in investigating the conditions of estimability of a linear function of the β's such as $\lambda'\beta$ where λ is a column vector with p components. Now let $\lambda'\beta$ be estimated by $a'y$. One proceeds to minimize the variance of $a'y = a'a\sigma^2$ under the condition of unbiasedness, $a'X = \lambda'$ with p constrains by the use of the Lagrangian Multiplier ρ. Now solving the equations $\partial Q/\partial a_i$, $i = 1,2\ldots n$ and where Q is the expression to be minimized we have $a = X\rho$ or $X'X\rho = \lambda$. This equation provides one with a condition of estimability which states that, if there exists a ρ such that $X'X\rho = \lambda$, the coefficient of the linear function of the β's, then $\lambda'\beta$ is estimable. In defining CE and I's, the conditions of estimability are appropriately incorporated into the definition, and if one follows the definitions and the constrains associated with the definition, the problem of estimability does not arise.

In consistence with the normal factorial notation, p denotes the level of the factors where p is a prime number, n denotes the number of the factor and the treatment combinations are denoted by $x_1x_2\ldots x_n$, x_i, $i=1,2\ldots n$, being the level of the i^{th} factor where x takes the value from 0 to (p-1). There are p^n treatment combinations, there are p^n-1 degrees of freedom, there are $(p^n-1)/(p-1)$ contrasts each with (p-1) degrees of freedom and each contrast with (p-1) degrees of freedom is associated with p-sets of p^{n-1} treatment combinations. All numbers are expressed as reduced modulo p. Confounding for p^n in blocks of p^s requires n-s independent effects or interactions to be confounded along with all generalized effects and interactions with a total of $(p^{n-s}-1)/(p-1)$ effects and interactions confounded. By considering the modulo definitions for the (n-s) independent effect or interaction confounded, one can generate the p^{n-s} blocks. The total degrees of freedom confounded is $(p^{n-s}-1)$ and the number of effects and interactions each with (p-1) degrees of freedom have $(p^{n-s}-1)/(p-s)$ effects and interactions confounded. The total number of systems of confounding for a p^n experiment in blocks of p^s is equal to

$$[(p^n-1)(p^n-p)\ldots(p^n-p^{n-s-1})] \,/\, [(p^{n-s}-1)(p^{n-s}-p)\ldots(p^{n-s}-p^{n-s-1})] \;.$$

Definition I. The symbolical representation of conditional effect and interaction with one condition is

$$(1) \qquad A^{a_1}B^{a_2}\ldots N^{a_n}/X_j$$

where a_i's, $i = 1,2,\ldots n$ take integral values between 0 and p-1 and by convention first $a_i \neq 0 = 1$, the X_j is any factor letters A, B, ... N for which its corresponding a_k $(k = 1,2\ldots n) = 0$ and $j = 0,1,2\ldots(p-1)$.

The expression in (1) defines contrasts among p-sets of treatment combinations satisfying one of the p following equations,

$$\sum_{i=1}^{n} a_i x_i = 0 \bmod p \,/\, x_k = q$$

$$\sum_{i=1}^{n} a_i x_i = 1 \bmod p \,/\, x_k = q$$

$$\vdots$$

$$\sum_{i=1}^{n} a_i x_i = (p-1) \bmod p \,/\, x_k = q$$

where a_i's take integral values between 0 and p-1, $k = 1,2,\ldots n$, q takes integral values between 0 and p-1, k can take any value $1,2,\ldots n$ for which $a_k = 0$ and it refers to the k^{th} coordinate in the n-dimensional space.

If $p = 3$ and $n = 4$, then one is dealing with 3^4-case. AB^2/C_0 is estimable since $a_1 = 1$, $a_2 = 2$, $a_3 = 0$ $a_4 = 0$, $X = C$ and $j = 0$ with congruential equation $[x_1 + 2x_2 = 0,1,2 \bmod 3/x_3 = 0]$.

Definition II. An effect or interaction conditioned on more than one effect is a conditional effect or interaction with multiple conditions

(2) $$A^{a_1}B^{a_2}C^{a_3}\ldots N^{a_n}/W_j/Y_k/\ldots/Z_m$$

where a_i's take integral values between 0 and (p-1)

$$j = 0,1,\ldots,p-1$$
$$k = 0,1,\ldots,p-1$$
$$m = 0,1,\ldots,p-1$$

W takes the factor letters A, B, ... N for which $a_w = 0$

Y takes the factor letters A, B, ... N for which $a_y = 0$ and for which $W \neq Y$

Z takes the factor letters A, B, ... N for which $a_z = 0$ and for which $W \neq Y \neq Z$.

The expression in (2) is defined by the contrasts among p-sets of treatment combinations satisfying one of the p following equations,

$$\sum_{i=1}^{n} a_i x_i = 0 \bmod p \;/\; x_j = q_1 \;/\; x_k = q_2 \;/\ldots/\; x_m = q_m$$

$$\sum_{i=1}^{n} a_i x_i = 1 \bmod p \;/\; x_j = q_1 \;/\; x_k = q_2 \;/\ldots/\; x_m = q_m$$

$$\vdots$$

$$\sum_{i=1}^{n} a_i x_i = (p-1) \bmod p \;/\; x_j = q_1 \;/\; x_k = q_2 \;/\ldots/\; x_m = q_m$$

and $j = 1,2,\ldots n$ and $0 \leq q_1 \leq p-1$

$k = 1,2,\ldots n$ and $0 \leq q_2 \leq p-1$

$\vdots$

$m = 1,2,\ldots n$ and $0 \leq q_m \leq p-1$.

In the case of 3^4 the conditional effect $A/B_0/C_1/D_2$ is estimable since $\alpha_1 = 1$, $\alpha_2 = \alpha_3 = \alpha_4 = 0$ and $W \neq Y \neq Z$. For this case the congruential equation is: $x_1 = 0, 1, 2 \bmod 3 \,/\, x_2 = 0 \,/\, x_3 = 1 \,/\, x_4 = 2$.

Definition III. For $p \geq 3$, one is interested in interpreting effects and interactions in terms of their polynomial effects such as linear effect, quadratic effect, cubic effect, etc. The linear effect is defined as

$$A' = \sum_{i=1}^{p-1} \left(i - \frac{p-1}{2}\right) A_i \Big/ \frac{p(p^2-1)}{12}$$

and the quadratic effect is defined as

$$A'' = \sum_{i=1}^{p-1} A_i \left[\left(i - \frac{p-1}{2}\right) - \left(\frac{p^2-1}{12}\right)\right] \Big/ \frac{p(p^2-1)(p^2-4)}{180}$$

for p equally spaced levels and A_i represents the i^{th} set of the p sets defining the contrast. For unequally spaced levels $q_0, q_1, \ldots q_{p-1}$, the linear effect is defined as

$$A' = \sum_{i=0}^{p-1} (q_i - \bar{q}) A_i \Big/ \sum_{i=1}^{p-1} (q_i - \bar{q})^2$$

and $$A'' = \sum_{i=1}^{p-1} C_i A_i$$

where $\Sigma C_i = 0$ and $\Sigma C_i (q_i - \bar{q}) = 0$.

The same line of argument is followed for the other higher order polynomial effects.

Definition IV. The expression in (1) and (2) will be called conditional effects and interactions (CE&I) and the unconditioned effects and interactions of the type $A^{\alpha_1} B^{\alpha_2} \ldots N^{\alpha_n}$ will be called classical effects and interactions.

PROPERTIES OF CE AND I'S. The following are the combinatorial and statistical properties of the conditional effects and interactions.

Property No. 1: The total number of conditional effects and interactions and the classical effects and interactions for a given p and n is

$$N(n,p) = [(2p)^n - (p+1)^n]$$

where $n \geq 2$ and p is a positive integer. The exact number of CE and I's are

$$N'(n,p) = (2p)^n - (p+1)^n - (p^n - 1) \quad .$$

A Table of $N(n,p)$ and $N'(n,p)$ have been presented in the Appendix.

Property No. 2: Consider an effect or an interaction, denoted by $A^{\alpha_1}B^{\alpha_2}\ldots N^{\alpha_n}/W_j/Y_k/\ldots/Z_m$ defined by the following equations

$$\sum_{i=1}^{n} \alpha_i x_i = 0 \bmod p/x_k = q_1/\ldots/x_m = q_m$$

$$\vdots$$

$$\sum_{i=1}^{n} \alpha_i x_i = p-1 \bmod p/x_k = q_1/\ldots/x_m = q_m \quad .$$

With other conditions satisfying, each equation satisfies

$$[p^{n-1-k}]$$

treatment combinations where k = number of conditions associated with the conditional effects and interactions.

Property No. 3: Consider a conditional effect and interaction $A^{\alpha_1}B^{\alpha_2}\ldots N^{\alpha_n}/W_j$, where W is any factor letter A, B, ... N for which $\alpha_w = 0$ and $j = 0, 1, 2, \ldots (p-1)$. If W is kept fixed and the congruential equations associated with the effects and interactions are solved for each value of $j = 0, 1, 2, \ldots (p-1)$, then p contrasts are generated and they are mutually orthogonal contrasts.

Property No. 4: Consider a conditional effect of interaction denoted by $A^{a_1}B^{a_2}\ldots N^{a_n}/W_j/Y_k/\ldots/Z_m$ with m conditions, where W, Y, ..., Z and j, k, ... m all satisfy conditions proposed in (2). If W, Y, ... Z are kept fixed and the congruential equations for the effect or interaction are solved for each combination of the values of j, k, ... m, then p^m contrasts are generated. They are all mutually orthogonal contrasts.

Property No. 5: Let the symbol $A^{a_1}B^{a_2}\ldots N^{a_n}$ denote the numerical totals of the effect or interaction under consideration, then we have

$$\sum_{j=0}^{p-1} (A^{a_1}A^{a_2}\ldots N^{a_n}/W_j) = A^{a_1}B^{a_2}\ldots N^{a_n} \quad .$$

This can be extended to the case with multiple conditions

$$\sum_{j=0}^{p-1}\sum_{k=0}^{p-1}\ldots\sum_{m=0}^{p-1} (A^{a_1}B^{a_2}\ldots N^{a_n}/W_j/Y_k/\ldots/Z_m)$$

$$= A^{a_1}B^{a_2}\ldots N^{a_n} \quad .$$

Property No. 6: For $p \geq 3$, property 5 can be extended to the polynomial effects, such as

$$\sum_{j=0}^{p-1} [(A^{a_1}B^{a_2}\ldots N^{a_n})'/W_j] = (A^{a_1}B^{a_2}\ldots N^{a_n})'$$

$$\sum_{j=0}^{p-1} [(A^{a_1}B^{a_2}\ldots N^{a_n})''/W_j] = (A^{a_1}B^{a_2}\ldots N^{a_n})'' \quad .$$

So also for the multiple condition case

$$\sum_{j=0}^{p-1}\sum_{k=0}^{p-1}\ldots\sum_{m-0}^{p-1} [(A^{a_1}B^{a_2}\ldots N^{a_n})'/W_j/Y_k/\ldots/Z_m]$$

summed over p^m terms $= (A^{a_1}B^{a_2}\ldots N^{a_n})' \quad .$

Property No. 7: Let $A^{\alpha_1}B^{\alpha_2}\ldots N^{\alpha_n}$ be the total of the effect or interaction under consideration, then the sum of squares for the conditional effects and interactions (SS(CE&I)) can be expressed explicitly as follows

$$\text{(3)} \qquad SS(CE\&I) = (A^{\alpha_1}A^{\alpha_2}\ldots N^{\alpha_n}/X_j/Y_k/\ldots/Z_m)^2\,(r\Sigma\lambda_i^2)^{-1}$$

where r is the number of replicates and λ's are the coefficients such that $\Sigma\lambda_i = 0$. If X, Y, ... Z are kept fixed and for each combination of the values assumed by j, k, ... m, a sum of squares is calculated then this p^m-set of sum of squares forms an orthogonal set for the analysis of variance tests. The conditional sum of squares can be expressed in terms of the combination of classical sum of squares as follows

$$\sum_{j=0}^{p-1}\sum_{k=0}^{p-1}\cdots\sum_{m=0}^{p-1}(A^{\alpha_1}B^{\alpha_2}\ldots N^{\alpha_n}/X_j/Y_k/\ldots/Z_m)^2/(r\Sigma\lambda_i^2)$$

$$= SS(A^{\alpha_1}B^{\alpha_2}\ldots N^{\alpha_n}) + SS(A^{\alpha_1}B^{\alpha_2}\ldots N^{\alpha_n}X) + \ldots + SS(A^{\alpha_1}B^{\alpha_2}\ldots N^{\alpha_n}Z)$$

where the definition of X, Y, ... Z and j, k, ... m are the same as given in (2). It is also noted that the expression in (3) generates single degrees of freedom contrast sum of squares.

CLASSICAL CONFOUNDING AND ESTIMABILITY OF CE&I'S. The problem here is the following:

Let a classical effect or interaction $A^{\alpha_1}B^{\alpha_2}\ldots N^{\alpha_n}$ defined by

$$\sum_{i-1}^{n} \alpha_i x_i = 0, 1, 2, \ldots (p-1) \bmod p$$

be completely confounded with blocks, then what are the conditions of estimability of the following CE&I

$$A^{\alpha_1}B^{\alpha_2}\ldots N^{\alpha_n}/W_j/Y_k/\ldots Z_m$$

defined by $\sum_{i=1}^{n} \alpha_i x_i = 0, 1, 2\ldots(p-1) \bmod p/x_i = q_1/x_k = q_2/\ldots/x_m = q_m$

where the a_i's, $W, Y, \ldots Z$ and $j, k, \ldots m$ meet specifications given in (2). The problem can be extended to cases in which two or more classical effects or interactions and their generalized interactions are completely confounded. The problem reduces to the fact that there are p^{n-s} blocks each containing p^s treatment combinations and one is interested in finding a contrast among p^{n-i-1} treatment combinations such that the contrast is orthogonal to p^{n-s} blocks.

The approach to the problem here will be to develop rules for rapid examination of the estimability of a given conditional effect or interaction under classical confounding based on simple mathematical manipulation. The theorems and proofs of the results are completely omitted.

In this problem we have three types of effects or interactions:

(i) Confounded effects or interactions and their generalized interactions,
(ii) Conditioned effect or interaction,
(iii) "Conditions" (effects used as conditions).

Each of the effects can be represented by their respective coordinates of the factor space.

Let Γ be a finite set of coordinates in the n-dimensional factor space,

$$\Gamma = \{x_1, x_2, \ldots x_n\} .$$

Let β be a subset of Γ containing the coordinates associated with the confounded effects or interactions

$$\beta = \{x_j, x_k, \ldots x_m\}$$

where $j, k, m = 1, 2, \ldots n$ and

$$\beta = \beta_1 \cup \beta_2 \cup \ldots \cup \beta_r$$

where β_i's are subsets containing coordinates of the i^{th} confounded effect or interaction out of r such confoundeded effects or interactions. β is then the union of the coordinates of r confounded effects or interactions and their generalized interactions.

Let γ be a subset of Γ containing the coordinates associated with the conditioned effects.

$$\gamma = \{x_i, x_j, \ldots x_k\} \quad .$$

Let Π_1 be a subset of Γ containing the coordinates associated with the "conditions"

$$\Pi_1 = \{x_i, x_j, \ldots x_k\}$$

where $\Pi_1 = \Pi_{11} \cup \Pi_{12} \ldots \cup \Pi_{1r}$

if there are r conditions associated with the conditional effect or interaction.

Now by the application of simple rules of set operation we derive the following new quantities, in three steps:

Step 1. $\delta = \beta \cup \gamma$

Step 2. $\Pi_2 = \beta \cup \gamma$

Step 3. $\delta_0 = (\Pi_1 \cup \Pi_2) \cap \delta$

The conditions of orthogonality to p^{n-s} blocks are as follows:

(i) If $\delta_0 = \delta$

then the conditional effect or interaction is not orthogonal to blocks

(ii) If $\delta_0 \neq \delta$

then the conditional effects and interactions under consideration are orthogonal to blocks and consequently estimable.

Examples: -

E1. Consider 2^4 case and confound ABC, BCD and their generalized interaction AD. Then the conditional effect A/C_0 is orthogonal to blocks, since $\Gamma = \{x_1, x_2, x_3, x_4\}$, $\beta_1 = \{x_1, x_2, x_3\}$, $\beta_2 = \{x_2, x_3, x_4\}$, $\beta_3 = \{x_1, x_4\}$, $\beta = \beta_1 \cup \beta_2 \cup \beta_3 = \{x_1, x_2, x_3, x_4\}$, $\gamma = \{x_1\}$, $\Pi_1 = \{x_3\}$, $\delta = \{x_1, x_2, x_3, x_4\}$, $\Pi_2 = \{x_1\}$, $\delta_0 = (\Pi_1 \cup \Pi_2) \cap \delta = \{x_1, x_3\} \neq \delta$.

E2. Consider the same case as in E1, the conditional interaction $AB/C_1/D_0$ is not orthogonal to blocks and consequently not estimable because $\beta = \{x_1, x_2, x_3, x_4\}$, $\gamma = \{x_1, x_2\}$, $\Pi_1 = \{x_3, x_4\}$, $\delta = \{x_1, x_2, x_3, x_4\}$, $\Pi_2 = \{x_1, x_2\}$, $\delta_0 = (\Pi_1 \cup \Pi_2) \cap \delta = \{x_1, x_2, x_3, x_4\} \cap \{x_1, x_2, x_3, x_4\}$ and by the application of the idempotency law, $\delta_0 = \delta$.

E3. Consider the case 3^5 and confound the four factor interaction ABCD, then the polynomial conditional interaction $A'B''/C_1/E_0$ is orthogonal to blocks because $\beta = \{x_1, x_2, x_3, x_4\}$, $\gamma = \{x_1, x_2\}$, $\Pi_1 = \{x_1, x_2, x_3, x_4\}$, $\Pi_2 = \{x_1, x_2\}$, $\delta_0 = (\Pi_1 \cup \Pi_2) \cap \delta = \{x_1, x_2, x_3\} \neq \delta$.

The construction phase of the factorial experimental designs involving conditional effects and interactions will be presented separately. The impact of confounding of conditional effect and interaction (conditional confounding) on the structure of fractional factorials, fractional replication and asymmetrical factorials with balanced and partially balanced configuration will also be presented separately.

INTERPRETATION OF FACTORIAL EXPERIMENTS. When an effect is conditioned on another effect, a conditional effect is generated. One of the examples of such effect is the well-known nested effect in the hierarchical classification situation. The conditional effects and interactions discussed in this study differ from the nested effects in that the nested effects do not permit consideration of reciprocal relationship between the conditional effect and its conditions, whereas the conditional effects do permit establishment of reciprocal relationship between the conditional effects and its conditions and do yield to meaningful interpretation when expressed in its reciprocal form. Consider a nested effect Farm/Counties, the reciprocal nested effect Counties/Farm is not defined, whereas a conditional effect A/B has a reciprocal nested effect B/A which is well defined. It possesses the property of commutativity with respect to the conditional operator "/". The conditional effect and interaction not only establish direct relationship between two or more effects but it also yields information on the reciprocal effects. An effect or interaction is usually defined orthogonal to other effects and interactions. By establishing direct and reciprocal relationship among the main effects, the conditional effects and interaction yield very meaningful and unambiguous interpretation. By reducing the higher dimensions to lower interpretable dimensions the higher order interaction does yield informative information with meaningful interpretation. All possible situations cannot be listed in this note.

Nowlet us consider a small experiment in which

$E(4A) = -68$; $E(2A/B_0) = 44$; $E(2A/B_1) = -122$

$E(4B) = 224$; $E(2B/A_0) = 190$; $E(2B/A_1) = 34$

$E(4AB) = -156$, where E stands for effect of, and the numberical value stands for the magnitude of the yield associated with the effects. By examining the numerical values of the effects and the conditional effects one can immediately appreciate the virtue of the information given by the conditional effects. The E(4A) yields the information that there is loss in the yield as one increases the level of A, but the two conditional effects following exactly tells us where is the loss and where is the gain, meaning that the loss associated with A is not a total loss. The E(4AB) give information on the loss associated with increasing levels of A or B or both, but the conditional effects associated with B clearly define where are the gains and their exact magnitudes. This is given here purely from the standpoint of appreciation of the usefulness of the conditional effects. The true use of conditional effects is appreciated in systems where p is large and n is large. An analysis of variance on conditional effects of this experiment is given in the Appendix, Table 2. The Appendix also contains a table (Table 3) of effects and sum of squares for a 3^2-experiment, where the polynomial effects have been isolated. It is interesting to note that the quadratic effect of B(B'') yields a gain of 56 units in the presence of higher dose of factor A. The interpretation of the other situations are self-explanatory.

REFERENCE

Bose, R. C. 1947. Mathematical theory of the symmetrical factorial design. Sankbya 8: 107-166.

Carmichael, R. D. 1937. Introduction to the theory of groups of finite order. Ginn, Boston.

Cochran, W. G., and Cox, G. M. 1962. Experimental Design 2d Ed., John Wiley and Sons, Inc. New York

Fisher, R. A. 1935. The design of experiment. Oliver and Boyd, Ltd. Edinburgh and London

Kempthorne, O. 1952. The design and analysis of experiments. John Wiley and Sons, Inc. New York

Yates, F. 1933. Complex Experiments, Supplement to the Journal of the Royal Statistical Society 2: 181-247.

APPENDIX

Table 1

N(n,p) and N'(n,p)

p/n	2		3		4	
2	7,	4	37,	30	175,	160
3	20,	12	152,	126	1040,	960
4	39,	24	387,	324	3471,	3216

Table 2

ANOVA 2^2 With Two Replications

Effect	d.f.	S.S.	M.S.	F
A	1	578	578	16.5*
A/B_0	1	484	484	13.8*
A/B_1	1	3136	3136	89.6**
B	1	6272	6272	179.2**
B/A_0	1	9025	9025	257.9**
B/A_1	1	289	289	8.3*
AB	1	3042	3042	86.9**
Error*(Total -3)=4		140	35	
Total $(2^3 r - 1) = 7$		10032		

* Replication pooled with error.

**Significant at .01 level of probability of Type I error.

Table 3

Sum of Squares for 3^2 Factorial

Contrasts	Effects	$\sum_{i=1}^{s} \lambda_i^2$	Sum of Squares
A'	-270	6	12,150.0
A'/B_o	-123	2	7,564.5
A'/B_1	-118	2	6,962.0
A'/B_2	- 29	2	420.5
B'	-257	6	11,008.2
B'/A_o	- 14	2	98.0
B'/A_1	-135	2	9,112.5
B'/A_2	-108	2	5,832.0
A''	30	18	50.0
A''/B_o	- 51	6	416.7
A''/B_1	- 16	6	42.7
A''/B_2	- 97	6	1,568.2
B''	3	18	.5
B''/A_o	- 28	6	130.7
B''/A_1	- 25	6	104.2
B''/A_2	+ 56	6	522.7
$A'B'$	94	4	2,209.0
$A'B''$	84	12	588.0
$A''B'$	148	12	1,825.3
$A''B''$	78	36	169.0

THE NEGATIVE BINOMIAL DISTRIBUTION APPLIED TO ATMOSPHERIC PARAMETERS

Oskar Essenwanger
Aerophysics Branch, Physical Sciences Laboratory
Research & Development Directorate
U. S. Army Missile Command, Redstone Arsenal, Alabama

ABSTRACT. The negative binomial distribution can be very helpful for determining wind speed and wind shear frequency distributions. The derivation of frequency distributions of vector wind shear data for small shear increments (100 and 50m) from existing common radiosonde data with 1 km altitude intervals is explained. The frequency distributions for smaller shear increments differ drastically from "scaled down" distributions. Considerable error for engineering evaluation would be introduced if the shape change of the negative binomial distribution with the shear increment through the change of the mean and sigma were neglected.

Finally computations of the cumulative 90, 95 and 99% exceedance thresholds for wind speed and wind vector shear by use of the negative binomial and the bivariate normal distribution are compared with the observations. The analytical values for employing the negative binomial prove best.

I. INTRODUCTION. Although the negative binomial distribution (NBD) has been known to statisticians for a long time, applications in atmospheric physics are not very wide spread. This can be explained for the following reasons.

After early discussion by Pascal and Fermat [see Todhunter, 1] one can find largely two versions of interpretation. Greenwood and Yule [2] assume that the events are mutually independent, but the intensity varies from individual to individual event. Polya [3] and Eggenberger [4] interpret that the events are statistically dependent, i.e. the occurrence of one event increases the probability that further events will occur.

In the latter sense applications have been attempted mainly for distribution of precipitations or runs of days with or without precipitation [see Wanner, 5, 6] . As has been pointed out by the author [7] , applications to the continuous frequency distribution of precipitation prove to be a problem. Therefore Thom [8] has suggested the use of the incomplete gamma function. Recently the model of the Markov chain [see Caskey, 9 and Weiss, 10] has been more successfully applied. Thus utilization of the negative binomial for the field of precipitation appears to be very limited.

In the sense of Greenwood and Yule's interpretation the NBD may apply for wind and wind shear. One would expect from other theoretical background, however, that the wind vector follows a bivariate normal distribution (the components being normally distributed). Then the non central chi-square distribution should adequately describe the distribution of the scalar wind speed.

The non central chi-square distribution, however, does not fit extreme values very well, especially for wind shear distributions of smaller increments [Essenwanger, 11]. Thus one may attempt to fit the empirical distribution with the NBD, as is later demonstrated. An earlier application has been made by Wanner [12], who concludes that the frequency distribution of the wind speed follows the shape of the NBD the closer the higher the altitude of his sampling (mountain observations). Since the present discussion is mostly concerned with upper atmospheric observations, the employing of the NBD with wind data may be investigated.

II. FREQUENCY DISTRIBUTIONS OF WIND AND WIND SHEAR. As previously mentioned the NBD is employed to describe the observed frequency distributions of wind speed and wind shear values. It is therefore of vital interest to ascertain how close is the agreement between observed and analytical distribution. Further, since the NBD is a discontinuous distribution, testing has to proceed to determine whether the given class division of the continuous wind distribution can be adequately adjusted to provide fair resemblance with the NBD. This adjustment is difficult for precipitation [see Essenwanger, 7].

The problem is discussed by the author in detail in a recent report [13]. Figure 1 serves as an example to summarize results for wind shear distributions. The figure displays a typical wind shear distribution for 1 km shear intervals (histogram at top of figure 1). The other 3 diagrams exhibit the deviations from the observed frequency for 3 types of fitted curves, the NBD, the incomplete gamma function with maximum likelihood fit and with moments fit. Statistical tests showed no significant difference between these 3 fitted curve types and the observation.

The NBD was selected for its convenience of computation. Since there are no observed data on the frequency distributions of smaller shear increments and the recommended distribution is predicted, data for a maximum likelihood fit of the gamma function are not available. Thus both analytical distributions rely on the moments fit. It is immaterial to select the negative binomial rather than the incomplete gamma function. It will be further seen that observed data in the range of extreme values such as threshold exceeded by 10, 5 and 1% of the data fit the observations quite well with the NBD.

III. WIND SHEAR INTERVALS OF SMALL INCREMENTS. In a basic article [14] the author has derived that a relationship between the mean shear $\bar{\nu}$ and the shear interval Δh exists as follows

$$(1) \qquad \bar{\nu}_{(\Delta h)} = a_o (\Delta h)^{a_1}$$

where a_o and a_1 are constants depending on climatological conditions. The Δh denotes the difference of the altitudes (shear interval), from which the vector shear ν as the residual of two wind vectors is computed.

It has further been deduced that a similar relationship holds for the standard deviation

$$(2) \qquad \sigma_{(\Delta h)} = A_o + B_1 (\Delta h)^{a_1}$$

where A_o and B_1 are again constants depending on climatic conditions. The constant A_o can be determined from

$$(3) \qquad \sigma_{(\Delta h)} = A_o + A_1 \bar{\nu}_{(\Delta h)} \quad .$$

Equation (1) has further been confirmed by Armendariz and Rider [15] and Belmont and Shen [16]. Although Armendariz and Rider [15] derived a similar equation to (1) for the standard deviation, which means A_o in (2) would equal zero, it is presently open whether A_o approximates zero in the ground layers, from which Armendariz' and Rider's data are derived, while the author included data up to 50 km altitude [see Reisig 17]. The absence of A_o may be further an effect of the terrain, as Armendariz and Rider work with data from White Sands, New Mexico, while the author's data were obtained at Cape Kennedy, Florida.

One has now two parameters, the mean and standard deviation, which can be utilized to compute the expected frequency distribution of shear values. Figure 2 demonstrates the agreement between observed and analytical distributions, employing the NBD for computing the analytical model. Five layers from various conditions of upper atmospheric shear distributions have been selected. The first 3 layers show excellent agreement between analytical and observed data. Some discrepancies are noted for 15-20 and 20-25 km. Although the deviations are statistically not significant, the problem of a distribution with a better fit or some adjustment to the fitting procedure is still open.

One point must be stressed, however. The present method, employing the NBD for describing the analytical distribution is far superior to the generally practiced technique of "scaling down" frequency distributions of vector wind shears. A typical example is given in Table 1. For 3 atmospheric layers a comparison was made between analytically derived and scaled down distributions. The "scaling down" technique assumes that the same distribution for smaller shear intervals as for larger intervals exists, e.g. the regular available shear distribution of 1 km intervals (easily obtained from the present radiosonde network) would be divided by 10 to obtain the distribution of 100 m shear interval. Table 1 demonstrates clearly that this technique is out of place as it does not take into consideration any shape and scale change. The real distribution produced such changes, which are quite adequately expressed by the NBD. It is important to include these changes into the derived frequency distribution. As becomes quite obvious from Table 1, considerable error for engineering evaluation would be introduced if the shape changes of the negative binomial distribution with the shear increment through the change of the mean and sigma were neglected.

More details can be found in pertinent articles as cited under 13 and 14.

IV. <u>COMPUTATION AND COMPARISON OF 90, 95 AND 99 PERCENT PROBABILITY THRESHOLDS.</u> One of the important criteria in missile application are percentile values such as the thresholds of shear values exceeded in 10 percent of the cases or similar tolerance values. Thus it is quite reasonable to demand that the employed distribution must be successful in describing said thresholds. The 90, 95 and 99% observational values were selected for this purpose.

Three types of distribution were tested, the bivariate distribution (BD), the negative binomial and the incomplete gamma function (IGF). The results for the NBD and the IGF were, however, similar and showed no statistically significant or obvious differences. Thus the comparison between NBD and IGF may be omitted here.

a. <u>Computation of the Threshold for the Bivariate Distribution.</u>

The computation of the 90, 95 or 99% value for the bivariate distribution is quite cumbersome. One has to solve the following type of integral

$$P_{(L)} = \int_{0}^{v_L} v \, f_{(v)} dv \tag{4}$$

Table 1

Comparison of Derived 100-Meter Interval Vector Wind Shear Frequency Distributions From 1000-Meter Intervals for January, Washington, D. C., (1956 through 1962)

Class Interval m/sec per 100 m	5 to 6 km		10 to 11 km		16 to 17 km	
	Scaled 100 m	Analytical Method	Scaled 100 m	Analytical Method	Scaled 100 m	Analytical Method
0.00 - 0.99	630	236.9	443	149.4	382	167.6
1.00 - 1.99	129	246.0	102	163.2	41	142.9
2.00 - 2.99	12	156.0	24	117.4	1	72.9
3.00 - 3.99	0	77.9	2	69.8	2	28.8
4.00 - 4.99		33.7	0	37.2	0	9.7
5.00 - 5.99		13.2		18.4		3.0
6.00 - 6.99		4.8		8.7		0.8
7.00 - 7.99		1.7		3.9		0.2
8.00 - 8.99		0.5		1.7		0.1
9.00 - 9.99		0.2		0.7		0.0
10.00 - 10.99		0.1		0.3		
11.00 - 11.99		0.0		0.2		
12.00 - 12.99				0.1		
13.00 - 13.99				0.0		
	771	771	571	571	426	426

where $P_{(L)}$ denotes the probability level of the threshold, v_L the threshold value of the wind speed, v the wind speed and $f_{(v)}$ the distribution function, in this case the bivariate normal distribution. A similar equation exists for the wind shear, replacing v by the pertinent parameter for the wind shear.

The solution is complicated, but can be approximated by the cumulative distribution of the non central χ^2 distribution or by determining an offset circle of the bivariate distribution (see 18). In the present application, one has

$$v_L = R \cdot \sigma \tag{5}$$

and $1 < R < 5$. Thus the equation for approximation is transformed for obtaining R in explicit form:

$$R^2 = a \left[1 - \frac{2}{9}\left(\frac{1+b}{a}\right) + \frac{2}{3}\left(\frac{1+b}{a}\right)^{\frac{1}{2}} C \right]^3 \tag{6}$$

where

$$C = \left[-(\ln P + \ln \sqrt{2\pi}) \right]^{\frac{1}{2}} \tag{7}$$

$$a = 2 + r^2 \tag{8}$$

$$b = \frac{r^2}{a} \tag{9}$$

$$r = \frac{v_r}{\sigma} \tag{10}$$

v_r denotes the resultant wind vector or the equivalent for the shear. The solution presents no problem when employing a high speed electronic computer.

A simplified approach, provides the same correlation coefficient, although in the winter months the average threshold is slightly higher than the observed value. This can be based upon the following assumption.

The mean wind speed $\bar{v}$ can be computed from

$$\bar{v} = \int_0^\infty v \, f_{(v)} \, dv \tag{11}$$

If the mean components have the values $\bar{x} = \bar{y} = 0$, and $\sigma_x = \sigma_y$ then the solution is

$$\bar{v} = C_o \sigma_v , \tag{12}$$

where

$$2\sigma_v^2 = \sigma_x^2 + \sigma_y^2 \text{ and } C_o = 1.2533. \tag{13}$$

If $\sigma_x \neq \sigma_y$, but not $\sigma_x \ll \sigma_y$ or $\sigma_y \ll \sigma_x$, equation (12) is a good approximation.

It has been shown by the author [11] that for $\bar{x} \neq 0$ and $\bar{y} \neq 0$, the following type of solution can be found

$$\bar{v} = C_o \sigma_v e^{A \frac{v_r}{\sigma_v}} . \tag{14}$$

This checks out well as demonstrated in Table 2 by the high correlation between $\ell n \frac{\bar{v}}{C_o \sigma}$ and $\frac{v_r}{\sigma_v}$.

If one assumes a similar form for the solution of v_L, namely

$$v_L = C_L \sigma_v e^{A \frac{v_r}{\sigma_v}} , \tag{15}$$

then v_L becomes simply

$$v_L = \frac{C_L}{C_o} \bar{v} . \tag{16}$$

If one considers that the C_o is taken from the circular normal distribution, where $\sigma_x = \sigma_y = \sigma_v$, then the high correlation between observed and computed values of the 90, 95 and even 99% as later shown is remarkable.

TABLE 2

Linear Correlation Coefficient for Check of Formula

$$\bar{v} = C_o \, \sigma_v \, e^{A \frac{v_r}{\sigma_v}}$$

Month	Mean Wind Speed		Mean Wind Shear
	El Paso	Chateauroux	Montgomery
Jan	.986	.987	.954
Feb	.995	.906	.915
Apr	.992	.888	.946
May	.991	.973	.971
Jul	.969	.998	.933
Aug	.998	.975	.944
Oct	.985	.953	.928
Nov	.997	.994	.853
Average	.989	.959	.931

b. Computation of the Threshold for the Negative Binomial Distribution.

The computation of the threshold value for the negative binomial is also based on a solution of the integral (3) as before but this time the $f_{(v)}$ is the negative binomial distribution and the integral is one-dimensional. Similar explicit formulae as for the bivariate distribution are presently not available. One can convert the cumulative NBD, however, into the incomplete beta function. This was pointed out by Pearson and Fieller [19], or rediscovered by Patil [20] and was recently discussed by Bartko [21, 22]. The 90, 95 and 99% values can then be obtained from the tables of the incomplete beta function [23]. The procedure is somewhat elaborate, but does not involve computations of the cumulative distribution by electronic computer. It was performed to obtain the necessary analytical values for comparison with the observed threshold.

Although the maximum likelihood fit could have been utilized by employing the frequency distributions and finding solutions to the maximum likelihood equation (Haldane, [24] and cited by Bartko, [22]), the moments method for parameter estimation was employed for the following reasons. One of the goals is the derivation of distributions for small shear intervals, for which the frequency distribution is not known. Thus the information necessary for the maximum likelihood fit is not available, while the parameters for the moments fit can be computed. If the NBD with moments fit would therefore give a poor result for computation of the threshold values, the NBD could not be used without first developing a prediction scheme for the information needed for maximum likelihood fit. Thus the question of maximum likelihood fit is of secondary importance for this particular problem.

When using the tables of the incomplete beta functions [23], the parameters p, q and the scale parameter b must be known. They have been obtained from

$$p = 2\,\frac{\left(\dfrac{\bar{x}^2}{\sigma^2} + \dfrac{\bar{x}^3}{\mu_3} - \dfrac{\bar{x}\sigma^2}{\mu_3}\right)}{\left(4\dfrac{\bar{x}\sigma^2}{\mu_3} + 1 - \dfrac{\bar{x}^2}{\sigma^2}\right)} \tag{17}$$

$$q = \frac{p\,(p+1)}{\dfrac{\bar{x}^2}{\sigma^2} - p} \tag{18}$$

(19) $$b = \mu_3 \frac{\left(\frac{\bar{x}^2}{\sigma^2} - 1\right) - 4\,\bar{x}\sigma^2}{\frac{\mu_3 \bar{x}}{\sigma^2} - 2\,\sigma^2}$$

where μ_3 is the third moment with reference to $\bar{x}$, the mean and σ^2 the variance*. The pertinent parameters for wind and wind shear have to be introduced into equations (17) thru (19).

The threshold value then becomes

(20a) $$v_L = b\,(1 - x_L) \text{ or}$$

(20b) $$v_L = b\,x_L$$

depending on whether $q > p$ (then 20a) or $q < p$ (then 20b).

c. Comparison of the Computed Thresholds with the Observed Values.

The threshold values of 90, 95 and 99% were computed for wind speed and wind shear for several stations and compared with the respective observed values. The latter were obtained from a computer program, listing certain thresholds of the cumulative distributions as begun in the Climatological Ringbook [25].

The differences between the computed and observed thresholds could have been checked with the Chi-square test for statistical significance of the deviations. Since the computed values were close to the observed thresholds, another tool of comparison has been employed. It was obvious from randomly selected samples that the chi-square test would not render statistical significance for most of the deviations of the computed threshold from the observed values. Thus the correlation coefficient was

*Footnote: The μ_3 for the negative binomial distribution is known, when the $\bar{x}$ and the σ^2 are known: $\mu_3 = \bar{x}\,(1 + 3d + 2d^2)$, where $d + 1 = \frac{\sigma^2}{\bar{x}}$.

utilized, which cannot only give information about the agreement between theory and observation, but can also delineate a systematic bias, if the means of the analytical and observed threshold differ.

The correlation coefficients are contained in Tables 3 - 5 for the data of Montgomery, Alabama as a typical example of the results. It is evident from the tables that the correlation is very high and therefore the analytical values are very close. However, a detailed inspection of the coefficients shows that there are some differences. First one notices that the coefficients display a slight tendency to decrease towards the 99% threshold. Thus the analytical values appear to fit less towards the extreme values. Further, this tendency to decrease is more pronounced for the bivariate fit than for the negative binomial and more for the wind shear than for the wind speed. This result is not unexpected. The tendency of deviations from the bivariate distribution, especially for wind shears, has been pointed out by the author in an earlier article [11]. Further, the analytical method for the bivariate distribution approximates the thresholds by either using mean wind speed only as in equation (16) or basing it on the circular distribution for equation (6). The method employing the negative binomial avoids these problems. Besides the mean, the variance of the distribution is needed, and in our particular case the variance of the wind speed and wind shear. By fitting the incomplete beta function, even the third moment μ_3 could be included, which is a 3 parameter fit. Thus the basis for analytically determining the thresholds comprises more or better parameters for the negative binomial approach. The result confirms this. The analytical thresholds agree better with the observed ones for the negative binomial method.

Whether there is a bias between the computed and observed thresholds can be answered from Table 6, where typical examples for the wind speed thresholds are displayed. The results for the negative binomial distribution look generally good, although there is a slight tendency towards a lower average than the observed. But the result may be considered within the tolerance limits of errors. The scatter for the analytical values v_L around the average $\bar{v}_L$ expressed by the standard deviation σ_a is the same as for the observed values, denoted by σ_o. This confirms the closeness of the computed results in addition to the high correlation coefficient.

The averages for the bivariate distribution also agree very well, thus no systematic large bias is visible. It is noticed, however, that the scatter represented by σ_a is higher than the scatter for the observed data. This indicates that not all of the computed values have good agreement, a conclusion already stated above in the consideration of the correlation coefficient.

Table 3

LINEAR CORRELATION COEFFICIENT FOR COMPARISON OF OBSERVED AND ANALYTICALLY DERIVED 90% LEVEL

Mon gomery

Month	Wind Speed		Wind Shear	
	Bivariate	Neg. Binomial	Bivariate	Neg. Binomial
Jan	.987	.994	.971	.915
Feb	.991	.985	.957	.960
Apr	.996	.996	.966	.950
May	.996	.994	.985	.974
Jul	.945	.875	.984	.979
Aug	.882	.867	.983	.968
Oct	.996	.994	.986	.976
Nov	.993	.995	.977	.973
Average	.973	.963	.976	.962

Table 4

LINEAR CORRELATION COEFFICIENT FOR COMPARISON OF OBSERVED AND ANALYTICALLY DERIVED 95% LEVEL

Montgomery

	Wind Speed		Wind Shear	
Month	Bivariate	Neg. Binomial	Bivariate	Neg. Binomial
Jan	.980	.996	.940	.938
Feb	.981	.992	.863	.940
Apr	.993	.997	.950	.905
May	.991	.997	.972	.978
Jul	.883	.906	.986	.985
Aug	.809	.898	.966	.959
Oct	.994	.997	.972	.958
Nov	.990	.997	.931	.950
Average	.953	.973	.948	.952

Table 5

LINEAR CORRELATION COEFFICIENT FOR COMPARISON OF OBSERVED AND ANALYTICALLY DERIVED 99% LEVEL

Montgomery

	Wind Speed		Wind Shear	
Month	Bivariate	Neg. Binomial	Bivariate	Neg. Binomial
Jan	.980	.995	.868	.936
Feb	.975	.990	.724	.925
Apr	.986	.996	.785	.888
May	.978	.996	.950	.975
Jul	.752	.919	.932	.980
Aug	.680	.929	.854	.958
Oct	.976	.992	.860	.902
Nov	.980	.997	.728	.943
Average	.913	.977	.838	.938

Table 6

Comparison of Computed and Observed Thresholds (Summary)

(a) bivariate

Station	Threshold	N	$\bar{v}_L$	$\bar{V}_L$	σ_a	σ_o	r
El Paso	90%	224	24.4	23.8	14.1	12.8	.984
	95	224	27.9	27.2	16.1	14.2	.970
	99	224	34.6	33.9	20.0	16.4	.929
Chateauroux	90	248	24.1	25.5	11.9	12.2	.970
	95	248	27.6	29.5	13.5	14.2	.967
	99	248	34.2	37.8	16.8	17.2	.941
Montgomery	90	372	26.6	25.6	18.5	15.3	.984
	95	372	30.3	29.2	21.1	16.8	.970
	99	372	37.6	36.2	26.1	19.5	.951

(b) negative binomial

Montgomery	90	372	24.9	25.6	15.3	15.3	.992
	95	372	28.2	29.2	16.8	16.8	.993
	99	372	34.4	36.2	19.5	19.5	.992

$\bar{v}_L$ mean wind speed of computed threshold (m/sec)

$\bar{V}_L$ mean wind speed of observed threshold (m/sec)

σ_a standard deviation of analytical values

σ_o standard deviation of observed values

r correlation coefficient

One may think about other distribution functions as being more suitable for deriving analytical values such as the Weibull distribution [26]. The negative binomial distribution, however, describes the thresholds already satisfactorily and preliminary computations with the Weibull distribution did not render better results rather than thresholds in the middle between the bivariate and negative binomial method. Besides, it is very difficult to objectively determine the location parameter for the Weibull distribution, and thus the negative binomial distribution offers an advantage in the estimation of parameters. Under these circumstances the question of determining the thresholds based on the Weibull distribution is not further pursued for this report.

V. SUMMARY AND CONCLUSIONS. It has been demonstrated that the negative binomial distribution has its place in problems of atmospheric physics, especially in missile climatology for wind speed and wind shear distributions. For this purpose the NBD serves largely as a practical and convenient tool for describing the frequency distribution. Especially the application to derive realistic frequency distributions of wind shear for small increments is important. This technique is far superior to the general practice of scaling down wind shear distribution for larger intervals which are commonly available. The utilization of the NBD, however, can accommodate the change of shape of the distribution with the shear interval, a property, which the scaling down neglects. Considerable error for engineering application may arise if this shape change is overlooked.

It has further been discussed in detail that the NBD can also be useful in deriving threshold values for the cumulative 90, 95 and 99% levels, if mean and variance for the distribution are known. Comparison between analytically derived and observed thresholds displayed excellent agreement without bias. The method proved superior to the application of the bivariate normal distribution for the same purpose. The only advantage for the latter practice could be the possibility of establishing a relationship between the threshold value and the mean, as expressed in equation (16). In this relationship one parameter, the mean only, needs to be known. This simplifies the computation of statistical parameters and increases the use of numerous data collection, in which the mean only is given.

The conversion of the NBD for the use of the tables of the incomplete beta function [23] to obtain the pertinent threshold values has been described. The need for knowing the third moment μ_3 does not introduce a new condition, as the μ_3 for the NBD is known with given mean and variance. Making use of three parameters, however, points to the possibility of utilizing the incomplete beta function for the curve fitting, although the third moment μ_3 then must be computed from the observations to offer some advantage. Utilization of two parameters, mean and variance, is sufficient only for the NBD.

Another 3 parameter fit would be the Weibull distribution. Preliminary computations did not produce better results, however, and therefore no detailed discussion and comparison were included in this report.

The NBD has therefore a definite place among the statistical distributions useful for application to atmospheric parameters.

REFERENCES

1. Todhunter, I., A history of the mathematical theory of probability from the time of Pascal to that of Laplace. MacMillan & Co. 1865.

2. Greenwood M. and Yule, G. U.: An inquiry into the nature of frequency distributions representative of multiple happenings. Journ. Roy. Stat. Soc., 83, 1920 p. 255 - 279.

3. Polya, G.: Sur quelques points de la theorie des propabilites. Annales de l' Institut Henri Poincare, 1, 1930 p. 117 - 161.

4. Eggenberger, F. and Polya, G.: Über die Statistik verketteter Vorgänge. Zeitschr. Ang. Math. Mech. 1, 1923, S. 279 - 289.

5. Wanner, E.: Über die Frequenz der täglichen Niederschläge. Verh Schweiz Naturf. Gesellsch. 1939, p. 27 ff.

6. Wanner, E.: Niederschlagsfrequenzkurven. Meteor. Zeitschr. 59, H. 3, 1942, p. 92 ff.

7. Essenwanger, O.: Wahrscheinlichkeitsansteckung und Erhaltungsneigung beim Niederschlag. Meteor. Rundschau, 9, 1956 S. 13 - 25.

8. Thom, H. C. S.: A statistical method of evaluating augmentation of precipitation by cloud seeding. Advisory Committee on Weather Control, Vol II, 1957, p. 2 - 25.

9. Caskey, J. E.: A Markov chain model for the probability of precipitation occurrence in intervals of various length. MWR Vol. 91, 1963, p. 298 - 301.

10. Weiss, L. L.: Sequences of wet or dry days described by a Markov chain probability model. MWR Vol. 92, 1964, p. 169 - 176.

11. Essenwanger, O.: Statistical parameters and percentile values for vector wind shear distribution of small increments. Archiv Meteor., Geoph., Bioklim Ser. A, Bd 15, 1965, p. 50 - 61.

12. Wanner, E.: Über die Frequenz der Windstärke auf dem'Santis 1932 - 1939. Ann. Schweiz. Meteor. Zentralanst. Bd 76, 1939, Anteung Nr. 7.

13. Essenwanger, O. and Billions, N.: On wind distributions for smaller shear intervals. U. S. Army Missile Command Report No. RR-TR-65-4, 1965.

14. Essenwanger, O.: On the derivation of frequency distributions of vector wind shear values for small shear intervals. Geofisica pura e applicata, Vol. 56, 1963, p. 216 - 224.

15. Armendariz, M. and Rider,: Wind shear for small thickness layers. U. S. Army Electronics Command Report ECOM-5040, 1966, 22 pp. Manuscript 66-352 of Sixth Conference on Applied Meteorology, Los Angeles, California, 1966.

16. Belmont, A. D. and Shen, W. C. C.: Comparison of jimsphere and rawinsonde wind shear. Control Data Corporation Final Report Contract N00014-66-C0127-1966.

17. Reisig, G.: Instantaneous and continuous wind measurements up to the higher stratosphere. Jour. Meteor., Vol. 13, 1956, p. 443 - 455.

18. Handbook of Mathematical Functions, 26. U. S. Department of Commerce, National Bureau of Standards, Fourth Printing, December 1965.

19. Pearson, K. and Fieller, E. C.: On the Application of the double Bessel function to statistical problems. Biometrica, Vol. 25, 1933, p. 158 - 178.

20. Patil, G. P.: On evaluation of negative binomial distribution function. Am. Math. Stat., Vol. 31, 1960, p. 527.

21. Bartko, J. J.: A note on the negative binomial distribution. Technometrics, Vol. 4, 1962, p. 609 - 610.

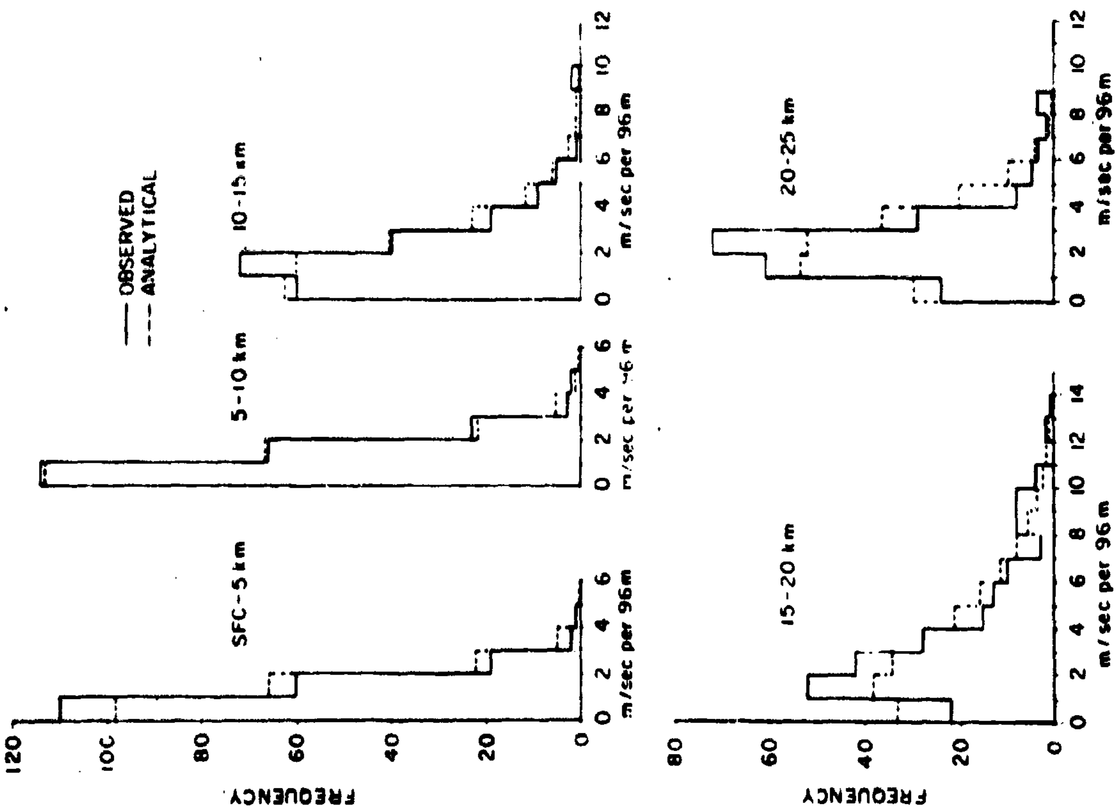
22. Bartko, J. J.: The negative binomial distribution: A review of properties and application. The Virginia Journal of Science, Vol. 12, 1961, p. 18 - 37.

23. Pearson, K.: Tables of incomplete beta function. Cambridge University Press, ed. 1956.

24. Haldane, J. S. B.: The fitting of binomial distributions. Annals Eugenics, Vol. 11, 1941, p. 179 - 181.

25. ABMA Climatological Ringbook. Army Ballistic Missile Agency Reports, 1961.

26. Weibull, W.: A Statistical Distribution Function of Wide Application. Jour. Applied Mechanics, Vol. 18, No. 1, 1951, p. 293 - 297.


OBSERVED
ANALYTICAL
SFC-5 km
5-10 km
10-15 km
15-20 km
20-25 km
FREQUENCY
m/sec per 96m

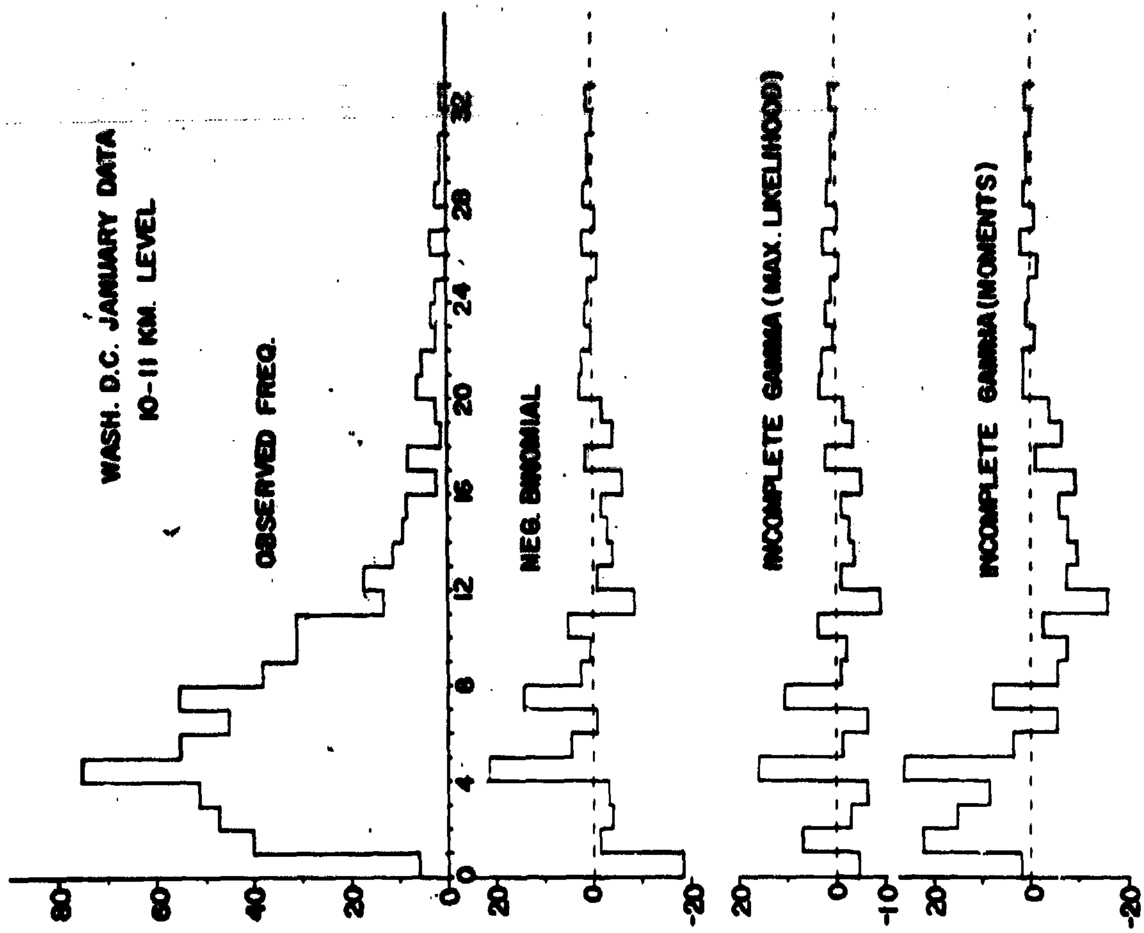

WASH. D.C. JANUARY DATA
10-11 KM. LEVEL
OBSERVED FREQ.
NEG. BINOMIAL
INCOMPLETE GAMMA (MAX. LIKELIHOOD)
INCOMPLETE GAMMA (MOMENTS)
0
4
8
12
16
20
24
28
32
80
60
40
20
0
20
0
-20
20
0
-10
20
0
-20

Figure 1

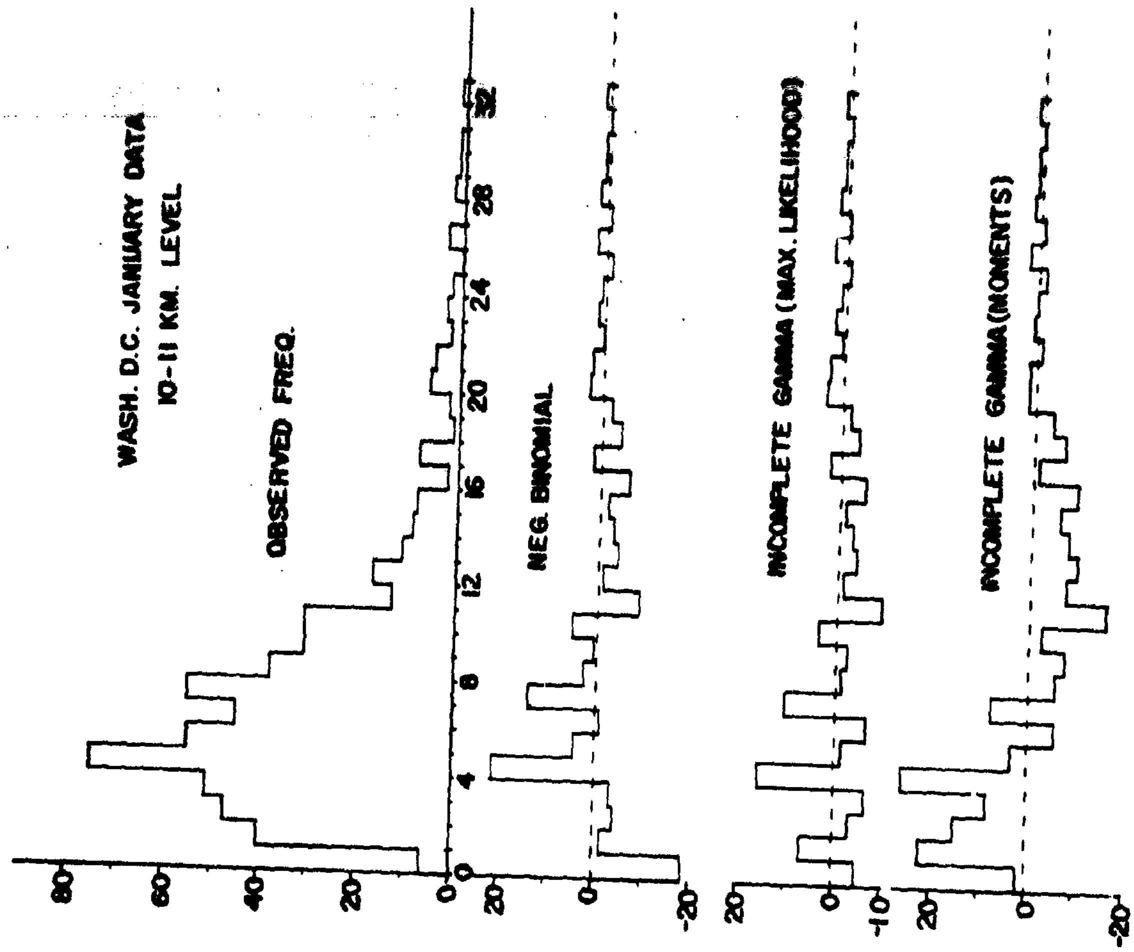

WASH. D.C. JANUARY DATA
10-11 KM. LEVEL
OBSERVED FREQ.
NEG. BINOMIAL
INCOMPLETE GAMMA (MAX. LIKELIHOOD)
INCOMPLETE GAMMA (MOMENTS)
80
60
40
20
0
20
0
-20
20
0
-10
20
0
-20
0
4
8
12
16
20
24
28
32

Figure 2

TRIAL VARIABILITY INTERPRETED AS DIFFERENCES IN TRANSLATION OR ROTATION IN FUNCTION ANALYSIS OF VARIANCE

Walter D. Foster
U. S. Army Biological Laboratories
Fort Detrick, Frederick, Maryland

ABSTRACT. Referee experimentation connotes in general a set of participants performing the same experiment under nearly identical circumstances. Variance analysis of results often takes the form of Between Stations and Within Stations. As a device for interpreting the magnitude of the mean square for repeated trials at a station, the mean square is converted to a corresponding vertical change in centroid (translation) or to a change in slope (rotation). The variable of analysis is a multiple-parameter function representing decay.

The concept and practice of the Analysis of Variance when the response variable is a function rather than a single value was given by Foster [1] in 1962. Comparison of this technique to the multivariate analysis of variance was given by Foster [2] in 1963. Brownlee [3] showed how to make simultaneous tests of slopes and centroids if the response is a linear function with two parameters. Church [4] gave the partition of variance for a factorial experiment for each parameter of a curvilinear model when used as the response variable.

The development in this paper is described in terms of its application: referee experimentation. Referee (or collaborative or standardization) experiments consist basically of several independent laboratories performing the same experiment in nearly identical circumstances. The simplest case compares laboratories (or stations as they are referred to here) using repeated trials at each station as the criterion -- the standard Between and Within analysis of variance. A more sophisticated design would introduce a range of treatments in order to estimate a Station X Treatment effect. Thus, the two major objectives of a referee experiment are the comparison of stations, treatment means and the estimation of reproducibility at each station. When each trial is a biological aerosol produced in a closed chamber and allowed to settle, the response is the decay function which describes the loss of biological activity with time;

$$C = C_0(t + 1)^{-b}e^{-kt}$$

is the decay model chosen for this analysis. The comparison of stations, treatments and S x T was given by Foster [1]. It is the purpose of this paper to examine the mean square for repeated trials at a station which was used as a measure of reproducibility and to translate this variance whose magnitude is generally meaningless to the experimenter into a familiar scale to facilitate subjective appraisal and evaluation of reproducibility.

Using the techniques of multiple regression, the data for a single trial of n points can be represented by

$$\ln C = \ln C_o - b \ln(t + 1) - kt$$

where $\ln C_o$, b, and k are estimated by least squares. Partition of the variation in the analysis of variance format is given in Table I, using Snedecor's [5] notation.

TABLE I. A. V. for a Single Trial

Line	Source	df		SS
1	Function	3		
2	C_o		1	$(SY)^2/n$
3	b, k		2	$bSx_1y + kSx_2y$
4	Deviations	n-3		$Sy^2 - bSx_1y - kSx_2y$
5	TOTAL	n		SY^2

Note: $X_1 = \ln(t + 1)$; $X_2 = t$; $Y = \ln C$

$x_1 = X_1 - \overline{X}_1$; $x_2 = X_2 - \overline{X}_2$; $y = Y - \overline{Y}$

For t trials at a station, the analysis of variance of the decay curve, showing partition and corresponding sums of squares is given in Table II.

TABLE II. A. V. for t Trials

Line	Source	df	SS
6	Mean	3	$(S_1^{tn} Y)^2/tn + b S_1^{tn} x_1 y + k S_1^{tn} x_2 y$
7	Among trials	3(t - 1)	Σ_1^t Line 1_i - Line 6
8	Deviations	t(n - 3)	Σ_1^t Line 4
9	TOTAL	tn	Σ_1^t Line 5

When the mean decay function for a station is compared to those of other stations, the comparison is both visual and objective -- visual because the functions can be graphed and their parameters tabled; objective because a test of significance is available [1] , but not given here. Thus, the comparison of means is complete and in a scale meaningful to the participants. Comparisons of trial M S for the various stations can also be done statistically, but the mean square itself has little meaning to the experimenter.

Two strategems involving translation and rotation in the original scale are presented as a method of interpreting the magnitude of the trial mean square. Since most aerobiologists are thoroughly familiar with the simple exponential function,

$$C = C_o e^{-kt}$$

as a decay model, the trial mean square has been scaled into translations of C_o and into rotation of k. The technique is simple.

Let the experimenter visualize the trial variability as being expressed by two parallel lines, the plot of

$$\ln C = \ln C_o - kt$$

whose vertical separation or translation is equivalent to the trial variability. Obviously, the greater the variability, the greater the distance between the two parallel lines. He thus may consider his trial variability as if he had run only two aerosols with equal decay rates but displaced starting points (intercepts).

Algebraically, the displacement or translation is derived by considering the same partition of the trial decay functions in Table II with only two trials. This is shown in Table III. The notation has the form of

$$Y = \bar{Y} - b\ (X - \bar{x}) .$$

TABLE III. Development of Trial Variability as Translation

Line	Identification	SS
10	SS Function 1:	$n\ \bar{Y}_1^2 + bSxy$
11	SS Function 2:	$n\ \bar{Y}_2^2 + b\ Sxy$
12	Line 10 + 11:	$n(\bar{Y}_1^2 + \bar{Y}_2^2) + 2b\ S\ xy$
13	Mean Function:	$2n[\ (\bar{Y}_1 + \bar{Y}_2)/2]^2 + 2bSxy$
14	Line 12 - Line 13:	$n(\bar{Y}_1 - \bar{Y}_2)^2/2$

The Mean Square corresponding to the sum of squares in Line 14 is simply

$$n(\bar{Y}_1 - \bar{Y}_2)^2/4 .$$

Upon equating the observed trial mean square to the derived translation and solving for the translation, we have

$$\bar{Y}_1 - \bar{Y}_2 = \sqrt{4\ \text{MS Trials}/n} ,$$

which as a distance applies to the intercepts, ln C_o, as well as to the centroids because of the assumed parallelism.

The following example of six trials at a station illustrates the use of this technique.

Function Analysis of Variance

Line	Source	df	MS
15	Mean	3	2427.4869
16	Among Trials	15	.1975
17	Deviations	30	.0878

$\bar{Y}_1 - \bar{Y}_2 = \sqrt{4(.1975)/8} = .314$ in ln C scale,

or a 1.37 fold (antiln .314) effect.

Had the trial mean square been .960 the translation would have been

$$\bar{Y}_1 - \bar{Y}_2 = \sqrt{4(.96)/8} = .693 \text{ or a } 2.0 \text{ fold effect.}$$

Interpreted to the aerobiologist, trial variation of this magnitude (MS = .1975) implies that his ability to reproduce an aerosol is no better than 1.37 fold. It should be noted in passing that the translation concept is applicable to any decay function under the requirement of parallelism.

The second approach to relate trial variability to experience is by rotation, i.e., a change in the slope of the linear decay function; in this case it refers to a change in the parameter k. The centroids for each of two lines are required to be identical; the MS for trial variability is equated to change in slope. This approach is more subtle since changes in k effected through equivalent size of the mean square depend upon the domain of the independent variable and the particular design. For a large domain the change in k will be small; for a narrow interval, the change will be large (because the variance of slope is proportional to $1/S\ x^2$). The development is given in the following table. The notation again has the form of

$$Y = \bar{Y} + b(X - \bar{X}) \quad .$$

TABLE IV. Development of Trial Variability as Rotation

Line	Identification	SS
15	SS Function 1:	$n\bar{Y}^2 + b_1 Sxy_1$
16	SS Function 2:	$n\bar{Y}^2 + b_2\ Sxy_2$
17	Line 15 & 16:	$2n\ \bar{Y}^2 + b_1\ Sxy_1 + b_2\ Sxy_2$
18	Mean Function:	$2n\ \bar{Y}^2 + (\frac{b_1 + b_2}{2})(Sxy_1 + S\ xy_2)$
19	Line 17-18:	$(b_1 - b_2)^2\ S\ x^2/2$

The Mean Square corresponding to the Sum of Squares in Line 19 is simply

$$(b_1 - b_2)^2\ \ S\ x^2/4\ \ .$$

As before, this quantity is equated to the observed trial mean square and the amount of rotation is

$$b_1 - b_2 = \sqrt{4(.1975)/1.66 \times 10^3}$$

$$= .022\ \ .$$

Note that the apparently small change in slope is due to the extremely large domain of t, 1300 minutes. While the concept of translation was applicable to any decay function, the rotation approach required a linear model for its straight-forward interpretation as a change in a single parameter.

REFERENCES

1. Foster, Walter D. (1962) Analysis of a Function in Collaborative Experimentation. Proceedings of the Eighth Conference on the Design of Experiments in Army Research, Development and Testing, ARO-D Report 63-2.

2. Foster Walter D. (1963) FAV versus MAV in a Referee Experiment. Presented at Biometrics - IMS Joint Meeting at Harvard University, 6 May 63.

3. Brownlee, K. A. Statistical Theory and Methodology in Science and Engineering. Wiley 1960, pp. 329-342.

4. Church, Alonzo, Jr. (1966) Analysis of Data when the Response is a Curve. Technometrics, Vol 8 No. 2, May 1966, pp 229-246.

5. Snedecor, George W. Statistical Methods, Fourth Edition. Iowa State College Press, 1946.

A METHOD FOR ADJUSTING FOR PARTICLE SIZE AND MATRIX EFFECTS IN THE X-RAY FLUORESCENCE ANALYSIS PROCEDURE

Raymond H. Myers
Department of Statistics,
Virginia Polytechnic Institute, Blacksburg, Virginia

Donald Womeldorph
Phillips Petroleum Company
Bartlettsville, Oklahoma

X-Ray fluorescence methods have been widely used in the analysis of multicomponent mixtures. The advantage is due, of course, to the high speed and precision of the method. It is unfortunate, however, that one is not always able to attain accurate analyses in practice because of the existence of sample matrix effects and particle size effects.

Existence of matrix effects implies that the intensity of fluorescent radiation from the analytical element is a function of the concentration of the matrix elements as well as its own concentration. This phase of our problem has been discussed by several workers. Mitchell [7] describes the problem in elaborate detail. In a recent paper, Alley and Myers [1] discuss ways of using inverse estimation in linear regression to account for these effects. Also, Campbell and Brown [3] have reviewed mathematical and empirical methods.

The consideration of particle size of the components is extremely important in X-Ray fluorescence analysis for the case of granular materials. In fact, variations in particle size of the materials can, in some cases, have a greater effect on the X-Ray intensity than variations in concentration. The fluorescent X-Ray intensity is affected by both the fluorescent and matrix component particle sizes and their relative concentrations in the sample. Claisse and Sampson [4], and Bernstein [2] discuss the particle size-intensity relationship.

This paper describes the use of a procedure involving estimation in a statistical functional relationship to approximate the structural form that exists between the X-Ray intensity of each component and the concentration of all of the components in the mixture. The non-linear functional relationships, which include the effects of measurement errors, permit the estimation of component concentrations in unknowns over wide ranges at constant particle size by using data obtained from the analyses of a series of calibration mixtures having the same particle size. Methods are also shown for estimating the concentrations of components in mixtures at any other combination of particle sizes by analyzing only one additional calibration mixture having the new particle size combination. This substantially reduces in comparison with conventional procedures the amount of experimental work required to recalibrate when one or more of the component particle sizes varies upon changing lots or batches of material.

Special attention is given with numerical results, to "Tichloral" igniter mixtures manufactured by the U. S. Army Missile Command, Redstone Arsenal, Alabama. These mixtures are comprised of potassium perchlorate, titanium, and aluminum powders, and sometimes a small percentage of a binder such as polyisobutylene. The estimation procedure is presented and "check samples" of known concentration (with particle size differing from that of the calibration data) were analyzed by the procedure.

The method described here differs considerably from the usual multiple regression technique.

THEORETICAL DEVELOPMENT OF PROCEDURE. Lucas Tooth and Pyne [6] developed a theoretical concentration - intensity model accounting for matrix effects. It is this model that serves as the basis for our development (other models such as a "complete quadratic" polynomial can perhaps be used as well). This model can be expressed as:

$$(1) \qquad V_n = \alpha_0^{(n)} + \sum_{1 \leq j \leq q} \alpha_j^{(n)} I_j + I_n \sum_{1<j<q} \alpha_{n,j}^{(n)} I_j$$

where V_n is the percentage of component n in the mixture; I_j is the X-Ray intensity for component j; the $\alpha^{(n)}$'s are constant parameters related to mass absorption coefficients [3]. $\alpha_0^{(n)}$ includes background intensity when peak intensity measurements are made. The subscript (n) implies that the parameters are characteristic of the n^{th} component, i.e., the α's describe enhancement or absorption of the other components with the n^{th} component. For example, for a three component mixture, we can write the percentage of component 1 in the mixture as:

$$(2) \quad V_1 = \alpha_0^{(1)} + \alpha_1^{(1)} I_1 + \alpha_2^{(1)} I_2 + \alpha_3^{(1)} I_3 + \alpha_{11}^{(1)} I_1^2 + \alpha_{12}^{(1)} I_1 I_2 + \alpha_{13}^{(1)} I_1 I_3 .$$

Often terms beyond those describing a linear equation can be deleted without serious consequence.

One might expect that a classical least squares procedure for estimating the coefficients in equation (2) would be appropriate. Actually, the papers [8], [5], and [1] rely heavily on this procedure. In the latter paper, the authors use a linear relationship in which the concentrations are on the right hand side of the equation, while intensity appears on the left. The coefficients are estimated by least squares and the equations (One for each component) are inverted for the analysis of an unknown. However, the particle sizes of the solid components were held constant in the experimental work. If the particle size effect is assumed to vary from batch to batch of raw materials that are used, then the coefficients in (2) would be dependent on particle size and thus it would be necessary to develop a different relationship involving different coefficients for each batch of raw materials.

Experimental methods are presented here for which the experimenter can use concentration - intensity data under one particle size condition, to determine the percentages of components in unknown samples under a second particle size condition.

Assumptions Concerning Equation (2). Suppose we consider the model of equation (2). We shall drop the subscript on the coefficients and thus refer to the relationship for component 1.

$$(2a) \qquad V_1 = \alpha_0 + \alpha_1 I_1 + \alpha_a I_2 + \alpha_3 I_3 + \alpha_{11} I_1^2 + \alpha_{12} I_1 I_2 + \alpha_{13} I_1 I_3 \ .$$

We could, of course, write a similar expression for components 2 and 3.

Suppose we consider two particle size levels, say 1 and 2. Suppose we have intensity - concentration data at particle size 2, but we wish to estimate the coefficients in (2.a) when the raw materials are from a batch at particle size 1. It must be emphasized here that we do not need to know what these particle sizes are; we simply know that two different conditions exist. We will assume that the measured intensity of component 1 at some concentration level (V_1, V_2, V_3) and at particle size 2 can be written:

$$(3) \qquad X_1 = I_1 + d_1 + f_1$$

and similarly for components 2 and 3.

I_1 is the "effective" or true X-Ray intensity of component 1 for the mixture at concentration (V_1, V_2, V_3), and at particle size condition 1.

d_1 is the particle size correction, i.e., the constant which represents the affect on the intensity of the particle size difference (between level 1 and level 2).

X_1 is the measured X-Ray intensity of component 1 when the mixture is composed of raw materials at particle size 2.

f_1 is a random measurement error effect on the intensity. It represents the affect of counting and other instrumental errors.

Further discussion of d_1 and f_1 are in order here. f_1 is considered to be a statistical "random" error, owing to inaccuracy in measuring the intensity. The measurement error as defined here includes components such as the counting error, and errors in the preparation of samples and pellets from the same calibration mixture. d_1 is not considered to be a random error but rather a constant value (plus or minus) which describes the affect on the intensity of particle size 2 over and above particle size 1. It is assumed for our purposes that the particle size within a batch is reasonably homogeneous. Otherwise one might consider d_1 to be a mean or average particle size affect. It must also be emphasized here that the d_1 represents an affect on intensity of ingredient 1 of the overall particle size of the mixture and not merely the particle size of any one ingredient. Finally, for our purposes, it is assumed that d_1, d_2, and d_3 (for the case of a three component system) are independent of the concentration level (V_1, V_2, V_3). This does not appear to be an unreasonable assumption if the concentration spread of interest is not excessive.

We shall now proceed to incorporate the model of equation (3) with that of (2.a) into a procedure for estimating the coefficients of (2.a). Suppose we have concentration - intensity experimental data for which the basic materials are at particle size level 2. We would like to be able to use this data to estimate (2.a) for materials at any particle size level. Suppose we consider (2.a) in which the materials are at particle size level 1. Substituting the actual intensities at particle 1 into (2.a) yields:

$$V_1 = \alpha_0 + \sum_{j=1}^{3} \alpha_j (X_j - d_j - f_j) + \sum_{j=1}^{3} \alpha_{1j} (X_1 - d_1 - f_1)(X_j - d_j - f_j) + \epsilon_1. \quad (4)$$

We have added the usual ϵ_1 (error term) as a random term to account for inaccuracies in equation (3) since this equation is certainly not completely deterministic in its derivation.

Estimation of Coefficients in Equation (3). Suppose the chemist were to prepare samples at preselected concentration levels and intensity readings are taken, the ingredients being from a batch at particle size level 2. We wish to use this information to obtain an estimate of the concentration - intensity relationship for particle size 1 (or for the ingredients from a batch at any other particle size).

Equation (4) can be written as:

$$(5) \qquad V_1 = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + a_{11} X_1^2 + a_{12} X_1 X_2 + a_{13} X_1 X_3 + Z_1$$

where:

$$\begin{aligned}
Z_1 = &-\left[\sum_{j=1}^{3} a_j f_j + \sum_{j=1}^{3} a_j d_j \right] + a_{11} \left[d_1^2 + f_1^2 \right] - 2 a_{11} X_1 \left[d_1 + f_1 \right] \\
&+ 2a_{11} d_1 f_1 + a_{12} \left[d_1 d_2 + f_1 f_2 \right] + a_{12} \left[d_1 f_2 + f_1 d_2 - f_2 X_1 - f_1 X_2 \right. \\
&\left. - d_2 X_1 - d_1 X_2 \right] + a_{13} \left[d_1 d_3 + f_1 f_3 + d_1 f_3 + f_1 d_3 - f_3 X_1 - f_1 X_3 - d_3 X_1 - d_1 X_3 \right] \\
&+ \epsilon_1 \, .
\end{aligned} \qquad (6)$$

Thus the "error" associated with the least squares model of equation (5) is given by equation (6). Note the terms that are translated to Z_1 through measurement errors and through the important particle size effects. The X's in equation (5) are the measured values of the intensitites and thus are random variables. One notices that if the usual least squares procedure is used, i.e., by minimizing the sum of squares of the errors in determining the estimates of the coefficients, that the error, Z_1, is correlated with the X's, since both involve the f's. This, of course, invalidates the usual regression assumption [9] that the residual error and the X's are independent. Of course, the errors in measuring the intensities may well be negligible, in which case we need only consider the effects translated by particle size. We shall discuss this situation in a later part of the paper.

It is not unreasonable to assume that these errors are independently distributed with zero mean and variance $\sigma_{f_j}^2$. Suppose we make n observations of the type $(X_{1i}, X_{2i}, X_{3i}, V_i)$. If we sum both sides of equation (5) over these n values, we obtain:

$$\sum_{i=1}^{n} V_{1i} = n\alpha_0 + \alpha_1 \Sigma X_{1i} + \alpha_2 \Sigma X_{2i} + \alpha_3 \Sigma X_{3i} + \alpha_{11} \Sigma X_{1i}^2 + \alpha_{12} \Sigma X_{1i}X_{2i}$$

(7)

$$+ \alpha_{13} \Sigma X_{1i}X_{3i} + \Sigma Z_{1i} .$$

All terms in equation (7) are known except ΣZ_{1i}. The latter contains sample quantities which certainly are unknown. For example, if we were to expand ΣZ_{1i}, such terms as $-\alpha_1 \sum_i f_{1i}$, $\alpha_{11} \sum_i f_{1i}^2$, $-2\alpha_{11} \sum_i f_{1i} X_{1i}$, etc. would appear, and since we have no knowledge as to the measurement error on any given sample, these quantities are unknown. However, we can replace these quantities by parameters that represent their "expected" or "average" values, the latter which we can estimate by a separate experimental procedure. If we assume the measurement error variable $f_j (j=1, 2, 3)$ has mean 0 and variance $\sigma_{f_j}^2$, then

$$E[\sum_{i=1}^{n} f_{1i}] = E[\Sigma f_{2i}] = E[\Sigma f_{3i}] = 0 .$$

Here the "E" notation refers to expectation. For $\alpha_1 \Sigma f_{1i}^2$, we can write

$$E(f_{1i})^2 = \sigma_{f_1}^2$$

and, if we further assume that the measurement errors are independent,

$$E[\alpha_{11} \Sigma f_{1i}^2] = n\alpha_{11}\sigma_{f_1}^2 .$$

After performing these operations, we can then write

$$\Sigma V_{1i} = n\hat{a}_0 + \hat{a}_1 (\Sigma x_{1i} - nd_1) + \hat{a}_2(\Sigma x_{2i} - nd_2) + \hat{a}_3(\Sigma x_{3i} - nd_3)$$

(8)

$$+ \hat{a}_{11}(\Sigma x_{1i}^2 + nd_1^2 - 2d_1 \Sigma x_{1i} - n\sigma_{f_1}^2) + \hat{a}_{12}(\Sigma x_{1i}x_{2i} + nd_1 d_2$$

$$- d_2 \Sigma x_{1i} - d_1 \Sigma x_{2i}) + \hat{a}_{13}(\Sigma x_{1i}x_{3i} - d_3 \Sigma x_{1i} + nd_1 d_3 - d_1 \Sigma x_{3i}).$$

Equation (8) is unbiased in the sense that both sides have the same expectation. We have inserted "hats" on the α terms to imply that they will be estimated by equations of this type.

For the next estimating equation, we can multiply both sides of (5) by x_{1i} and sum over the n observations as before.

$$\Sigma V_{1i}x_{1i} = \alpha_0\Sigma x_{1i} + \alpha_1\Sigma x_{1i}^2 + \alpha_2\Sigma x_{1i}x_{2i} + \alpha_3\Sigma x_{1i}x_{3i} + \alpha_{11}\Sigma x_{1i}^3$$

$$+ \alpha_{12}\Sigma x_{1i}^2 x_{2i} + \alpha_{13}\Sigma x_{1i}^2 x_{3i} + \Sigma x_{1i}Z_{1i} .$$

$\Sigma x_{1i}X_{1i}$ will contain unknown sample quantities which we shall once again replace by their expectation. The term Σf_{1i}^3 is replaced by $n\sigma_{f_1}^3$, which we are defining as $E(f_1)^3$, the third moment of the distribution of f_1. In doing this, we arrive at the following equation:

$$\Sigma x_{1i}V_{1i} = \hat{\alpha}_0\Sigma x_{1i} + \hat{\alpha}_1 (\Sigma x_{1i}^2 - n\sigma_{f_1}^2 - d_1\Sigma x_{1i}) + \hat{\alpha}_2 (\Sigma x_{1i}x_{2i} - d_2\Sigma x_{1i})$$

$$+ \hat{\alpha}_3(\Sigma x_{1i}x_{3i} - d_3\Sigma x_{1i}) + \hat{\alpha}_{11}(\Sigma x_{1i}^3 + 2nd_1\sigma_{f_1}^2 - n\sigma_{f_1}^3$$

$$- 3\sigma_{f_1}^3\Sigma x_{1i} + d_1^2\Sigma x_{1i} - 2d_1\Sigma x_{1i}^2) + \hat{\alpha}_{12}(\Sigma x_{1i}^2 x_{2i} + nd_2\sigma_{f_1}^2 \tag{9}$$

$$- \sigma_{f_1}^2\Sigma x_{2i} + d_1d_2\Sigma x_{1i} - d_2\Sigma x_{1i}^2 - d_1\Sigma x_{1i}x_{2i}) + \hat{\alpha}_{13} (\Sigma x_{1i}^2 x_{3i}$$

$$+ nd_3\sigma_{f_1}^2 - \sigma_{f_1}^2\Sigma x_{3i} + d_1d_3\Sigma x_{1i} - d_3\Sigma x_{1i}^2 - d_1\Sigma x_{1i}x_{3i}) .$$

This equation also has unbiased property as does equation (8).

At this stage we have two estimating equations. We can proceed to derive five more for estimating the seven coefficients in model (2.a). We obtain these equations by multiplying both sides of equation (5) by x_{2i}, x_{3i}, x_{1i}^2, $x_{1i}x_{2i}$, and $x_{1i}x_{3i}$ and performing the necessary operations, as described here for the first two equations, on $\Sigma x_{2i}Z_{1i}$, $\Sigma x_{3i}Z_{1i}$, etc.

Estimation of the d's, $\sigma^2_{f_j}$, etc. The quantities d_j and $\sigma^2_{f_j}$ (j=1, 2, 3) which appear in the estimating equations are, of course, unknown and must be estimated before we can use the equations in estimating the a's.

Obtaining an estimate of $\sigma^2_{f_1}$ is quite easily accomplished by preparing several samples and obtaining intensity measurements $x_{11}, x_{12}, \ldots, x_{1N}$ (independent of the samples used in section (b)) at some concentration and computing $\hat{\sigma}^2_{f_1} = \Sigma(x_{1i}-\bar{x}_1)^2/N$, the sample variance. One can then compute estimates for $\sigma^2_{f_2}$, $\sigma^2_{f_3}$ by obtaining similar sample variances for the intensities of components 2 and 3. We can, of course, estimate the third and fourth moments in a similar manner.

To obtain estimates of the d's, the experimenter needs first to obtain replicated analyses (on component 1 for the case of d_1) for a sample of raw materials from particle size 2. One must then obtain similar readings for the materials from the batch of interest, in our case this refers to the batch at particle size 1. It is important that the two sets of readings be taken at the same concentration. One can then obtain the averages $\bar{x}_1^{(1)}$ and $\bar{x}_1^{(2)}$, where the superscript denotes the particle size condition. The unbiased estimate of d_1 is then $\bar{x}_1^{(2)} - \bar{x}_1^{(1)}$. The same procedure is used to obtain estimates of d_2 and d_3.

APPLICATION TO IGNITER MIXTURES. Samples of the igniter mixture were prepared at various concentrations of $KClO_4$, Ti, and Al. The intensity for each component was measured for each sample. The data is shown in Table 1. The overall particle size effect on each intensity was assumed to be the same for these samples, and the materials in this batch were relatively "coarse" for all three ingredients. Thus we shall refer to the particle size as c-c-c. This is particle size 2 in our theoretical development.

Experimental.

Instrumentation-Analyses were made with a universal vacuum X-Ray spectrometer marketed by Philips Electronic Instruments. Spectrometer components consisted of an FA-60 tungsten target X-Ray tube, a 4-inch by 0.020-inch entrance collimator, an ethylenediamine D-tartrate (EDDT) analyzing crystal, and a gas flow detector flushed with P-10 gas. The X-Ray tube was operated at 45KV - constant potential, and 40 ma. Pulse height discrimination was used for the analysis of aluminum.

TABLE 1

Concentration-X-Ray Intensity Data for Igniter Mixtures

Potassium Perchlorate		Titanium		Aluminum	
Weight %	Counts/sec.	Weight %	Counts/sec.	Weight %	Counts/sec.
31.0	11,609.	31.0	7,279.	34.0	4,917.
31.0	11,582.	31.0	7,135.	34.0	5,116.
5.0	2,113.	30.0	8,302.	34.0	3,967.
5.0	2,146.	30.0	8,194.	34.0	3,831.
34.0	11,775.	7.0	1,691.	35.0	3,825.
34.0	12,003.	7.0	1,735.	35.0	3,857.
31.0	15,180.	30.0	8,660.	6.0	690.
31.0	15,345.	30.0	8,638.	6.0	712.
8.0	3,266.	6.0	1,746.	35.0	2,986.
8.0	3,294.	6.0	1,776.	35.0	3,180.
34.0	15,345.	6.0	1,907.	8.0	730.
34.0	15,250.	6.0	1,884.	8.0	737.
4.0	2,509.	29.0	11,507.	6.0	551.
4.0	2,486.	29.0	11,326.	6.0	551.
7.0	3,968.	5.0	2,240.	8.0	531.
7.0	3,938.	5.0	2,198.	8.0	534.
19.0	8,198.	18.0	5,159.	21.0	2,077.
19.0	8,574.	18.0	5,279.	21.0	2,160.

Procedure-Calibration and "check" mixtures were prepared for analysis as follows: 10 g. of each mixture including a variable amount of a cellulose binder was weighed into a 1-inch by 2-inch stainless steel vial and a 3/8 inch diameter plexiglas ball was added to the mixture. The mixture was then blended on a pica blender mill for 10 minutes. The ball facilitated blending without reducing the particle sizes of the powders. Two 5 g. pellets of each mixture were made in a 1 1/4-inch diameter pellet die under a pressure of 30,000 psi. The surface of each pellet that was against the die plunger was subsequently analyzed.

Pellet Samples were completely randomized among the calibration mixtures and analyzed in pairs in conjunction with a stable reference pellet containing the same analytical elements as the mixtures. The reference standard was used to correct X-Ray intensities for short and long term instrumental fluctuations. Peak intensity measurements were made by a fixed count technique and recorded as corrected counts per second. Specific analytical parameters are given in Table 2.

TABLE 2

Analytical Parameters for the Analysis of Igniter Mixtures

Component	Emission line	*Angle, °2θ	Fixed Counts
Potassium Perchlorate	KKα	22.23	200,000
Titanium	TiKαII	49.25	100,000
Aluminum	AlKα	114.77	50,000

*EDDT crystal advanced approximately 30° 2θ

ESTIMATION OF CONCENTRATION-INTENSITY MODEL AT SECOND PARTICLE SIZE CONDITION. A second batch of material was considered, one which contained relatively coarse particles of $KClO_4$ and fine particles of Ti and Al. Suppose one wished to estimate equation (2.a) for the c-f-f(particle size 1) lot using, however, the available concentration-intensity data for particle size 2 namely that in Table 1.

For the purposes of estimating the d_i, a sample from the c-f-f batch was prepared at 19, 18, and 21 per cent $KClO_4$, Ti, and Al respectively. Duplicates were taken and the intensities in counts per second obtained were:

$KClO_4$	Ti	Al
5,787.	3.676.	2,453.
5,770.	3,646.	2,461.

Point 9 of Table 1, with ingredients also at 19, 18, and 21 per cent concentration was used as the appropriate sample for the c-c-c batch. Subtraction of the average intensities was performed as indicated previously. Solution of equations (8) through (14), using the data of Table 1 was then accomplished on an IBM 7040 computer for each of the three ingredients. These coefficients are listed in Table 3. The coefficients can now be used for analysis of mixtures for the materials from the c-f-f lot.

TABLE 3

Estimates of the Coefficients for Coarse-Fine-Fine Lot

$KClO_4$(component 1)	Ti (component 2)	Al(component 3)
$\hat{a}_0 = 1.80189$	$\hat{a}_0 = -7.99781$	$\hat{a}_0 = -4.12577$
$\hat{a}_1 = 2.28358\text{x}10^{-3}$	$\hat{a}_1 = -4.34502\text{x}10^{-5}$	$\hat{a}_1 = -2.65274\text{x}10^{-5}$
$\hat{a}_2 = -3.71806\text{x}10^{-7}$	$\hat{a}_2 = 3.68540\text{x}10^{-3}$	$\hat{a}_2 = 2.87296\text{x}10^{-5}$
$\hat{a}_3 = 5.22015\text{x}10^{-4}$	$\hat{a}_3 = 1.51577\text{x}10^{-4}$	$\hat{a}_3 = 1.43834\text{x}10^{-2}$
$\hat{a}_{11} = 4.99141\text{x}10^{-9}$	$\hat{a}_{22} = -7.89669\text{x}10^{-8}$	$\hat{a}_{33} = -6.38742\text{x}10^{-7}$
$\hat{a}_{12} = -4.46347\text{x}10^{-8}$	$\hat{a}_{12} = 4.91218\text{x}10^{-8}$	$\hat{a}_{13} = -1.48698\text{x}10^{-7}$
$\hat{a}_{13} = 1.47985\text{x}10^{-7}$	$\hat{a}_{23} = 1.96862\text{x}10^{-7}$	$\hat{a}_{23} = -2.88265\text{x}10^{-7}$

Analysis of Check Sample. More samples were prepared using materials from the c-f-f lot in order that the analytical equation for $KClO_4$ and Ti could be checked. Notice that it is only necessary in this case to analyze for two components. The third can be obtained by difference because the α-cellulose binder is added by the analyst and is always known. The per cent of Al for the "check samples" was computed by difference. The results were compared with the known concentrations in order that the quality of the estimating equations could be evaluated. In order to illustrate the improvement obtained by the method over that of ordinary least squares without the particle size correction, the results for the check samples were compared with those obtained by estimating the intensity-concentration relationship of equation (2.a) by ordinary least squares.

The first sample contained the known concentration; 25% $KClO_4$, 25% Ti, 25% Al, and 25% α-cellulose binder. The intensities in counts per second were observed in duplicate. The results are:

$KClO_4$	Ti	Al
8,453.	8,107.	3,353.
8,332.	8,129.	3,379.

Using these intensities from the duplicates, the average calculated percentage compositions (Using the coefficients in Table 2) are below:

$KClO_4$	Ti	Al
24.71	25.59	24.70

This indicates the agreement between actual and estimated concentrations. One would, of course, expect even better agreement if the range of concentration of the original data in Table 1 were more narrow. The estimated concentration, using conventional least squares, neglecting particle size and measurement errors are:

$KClO_4$	Ti	Al
20.17	30.40	27.20

The difference between these values and the ones for our proposed procedure is primarily due to the introduction of the d's into the method.

Additional samples from the c-f-f lot were prepared and the estimates of concentration were obtained, using both conventional least squares, and our procedure. The results are shown in Table 4.

TABLE 4

	Actual Concentrations			Predicted Concentrations			Least Squares		
Sample	%K	%Ti	%Al	%K	%Ti	%Al	%K	%Ti	%Al
1	21	21	21	20.8	21.65	21.55	16.88	26.56	19.56
2	21	25	19	20.53	26.3	18.17	16.72	30.6	17.68
3	18	20	24	18.09	20.4	23.5	17.33	25.23	19.45

Note the improvement in the procedure over the least squares results.

Use of a Linear Model. In many cases of quality control analysis the materials to be analyzed will vary over small concentration ranges and the procedure of estimating concentrations at a given particle size and compensating for recognized particle size changes can be simplified by using a linear model such as:

$$I_i = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \ldots$$

The same procedures apply to this model and the estimating equations are considerably more simple than those for the second order model discussed in detail here.

Discussion of Sources of Error. The d_j and the moments of the f's are based only on sample estimates. This is obviously a source of error in the procedure. For the igniter system presented here, the d_j are based on only two observations. We would expect better results on the check samples if we had used more observations.

In many practical situations where the X-Ray fluorescence technique is used, the range of interest in concentration would be more narrow than what we used here (Table 2). In practice one might wish to narrow the range of experimentation to insure the truth of the assumption that the d_j are truly constant and do not depend on concentration.

When determining $\sigma^2_{f_j}$ one must be sure to include all source of error which cause X_j to differ from the true intensity I_j. As pointed out earlier this involves more than just making repeated measurements on the same sample which gives primarily the counting error. The error of blending mixtures and preparing pellets as well as uncompensated instrumental mechanical, and electronic variations must also be accounted for. A good estimate of the measurement error can be easily obtained, however.

The composition selected for determining both d_j and $\sigma^2_{f_j}$ should lie close to the center of the calibration compositions. Also, the calibration compositions should be selected according to a statistically designed experiment to insure accurate estimates of the coefficients in equation (2. a).

In a controlled process the normal variation of particle sizes among lots of materials will be smaller than the variation shown here for sizes 1 and 2. These variations were purposely made large to illustrate the suitability of the method.

The α-cellulose binder of the igniter mixtures was considered as a variable component in this work. Although it could not be analyzed directly by X-Ray spectrometry; the binder was allowed to vary to simulate production igniter mixtures which may contain a binder subject to production variations in the same manner as the other components. The binder, of course, also results in the formation of stronger pellets, and thereby allows a wider range of composition to be analyzed. The binder would normally be added to the mixture in a constant amount by the analyst. Results of analyses with constant binder would probably be more accurate than results with variable binder.

LITERATURE CITED

1. Alley, B. J., Myers, R. H., Anal. Chem. 37, 1685 (1965).

2. Bernstein, F., Advances in X-Ray Analysis. 6, 436 (1963).

3. Campbell, W. J., Brown, J. D., Anal. Chem. 36, 312R (1964).

4. Claisse, F. and Samson, C., Advances in X-Ray Analysis. 5, 335 (1962).

5. Lamborn, R. E., Sorenson, F. J., Advances in X-Ray Analysis. 6, 422 (1963).

6. Lucus-Tooth J., Pyne, C., Ibid. 7, 523 (1964).

7. Mitchell, B. J., "Encyclopedia of X-Rays and Gamma Rays", G. L. Clark ed., p. 531, Reinhold New York (1963).

8. Sugimoto M., Burseki-Kaguku 11, 1168 (1962); C. A. 58, 3076g (1963).

9. Williams, E. J., "Regression Analysis," p. 162, Wiley, New York, 1951.

DETERMINING THE CONFIDENCE LIMITS FOR SOME TIME INDEPENDENT SYSTEM RELIABILITY ESTIMATES WHEN ATTRIBUTE DATA FOR THE INDEPENDENT SUB-COMPONENTS ARE GIVEN. (A Proposed Solution and Approximating Formula)

Eugene Dutoit
Picatinny Arsenal, Dover, New Jersey

STATEMENT OF PROBLEM: A problem that arises often in ammunition engineering is estimating the reliability of some "one shot" weapon systems. This clinical problem is concerned with the situation where the only data available are attribute (the fraction: number of successful functionings/total number of items tested) and pertain to the components of the system. The ammunition or reliability engineer arranges the independent system components in some logical configuration (called the reliability block diagram) and he constructs a mathematical model of the overall system reliability. Established procedures do exist for determining the reliability of each separate component at any appropriate confidence level, but this problem of interest is to establish some techniques for combining these component data so that some reliability estimate can be made about the system (note: -no "system" data are available) at any desired confidence level. In essence, this problem is hopefully designed to:

(1) Raise some interest and thought for this problem which appears to have been treated too lightly considering the frequency with which it arises. Perhaps someone who might be writing or considering to write a textbook on reliability might develop a computational procedure for publication and reference. The use of computer/simulation studies have already been proposed. These methods may be applicable when a computer is available and time is not a crucial factor, but we are seeking a solution that would give a quick but good approximation to some rigorous and lengthy solution.

(2) Encourage the examination of data indicating the distribution of failures for conventional weapon systems to determine if some characteristic distribution can be used to describe some types of items. This paragraph reflects similar statements made by Lt. Colonel M. S. Hochmuth during the "opening remarks" of this conference.

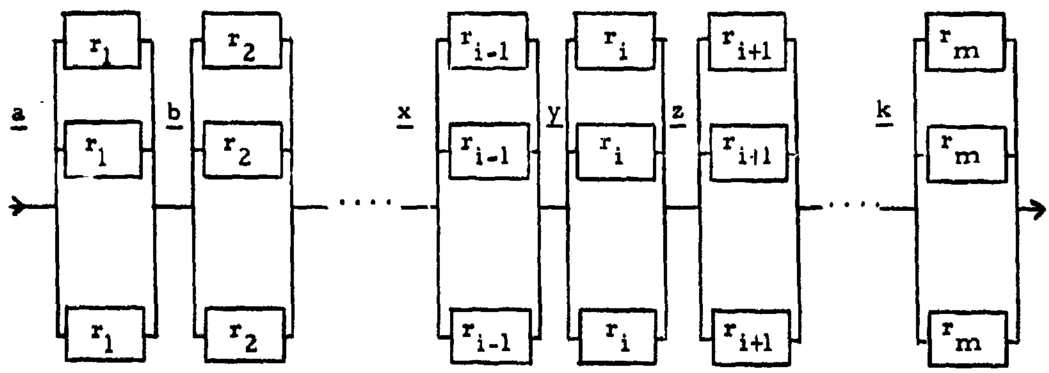
ACKNOWLEDGMENTS: Before continuing with some proposed solution and approximating technique I would like to express my appreciation to the Army Mathematics Steering Committee for giving me the opportunity to present this clinical problem at the "Twelfth Conference", I am also appreciative to all the panel members (Dr. F. Frishman, Chairman; Mr. O. Bruno, Professor A. Cohen, Jr., Professor B. Harshbarger, Dr. J. Rosenblatt and Professor H. Solomon) who offered constructive suggestions/comments either at the meeting or by writing.

I also wish to thank Mr. Stuart Ritter who developed the computer program and charts used in this work.

A PROPOSED SOLUTION: The author of this report has independently arrived at a "similar" solution to the problem as Mr. H. DeCicco [1]* and Messrs. Lloyd and Lipow [4], therefore the derivations presented here shall be "quick and dirty". The interested reader should consult these references, and the other sources cited in this report, in order to become more familiar with the problem. Those readers who are interested in researching the problem might compare this enclosed solution or some other possible solutions with each other to determine if some extra degree of accuracy obtained by a more rigorous/analytic method is worth the extra effort. DeCicco mentions in his paper [1] that it is "unrealistic to expect serious support for assurance to more than two significant digits". This criterion might be used to determine significant differences between all proposed solutions to this problem. This proposed solution will be reduced to some approximation and graphic procedure which will hopefully simplify the computation for non-mathematically oriented personnel.

SERIES-PARALLEL CASE (GENERAL): Consider the following configuration:

Figure 1



Notation: Number of "y" components in parallel in set i, where the reliability of each item at the 50% confidence level is r_i.

*Numbers in brackets indicate references at the end of paper.

The derivation of the "error propagation" formula is well known and need not be discussed here (see Bowker and Lieberman; Engineering Statistics, Prentice Hall, Inc., Englewood Cliffs, N. J., 1959, page 62). Given a function of m variables $f(r_1, r_2, \ldots, r_m)$ with expected values $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_m$, the expected value of the function is approximated by:

$E[f(r_1, r_2, \ldots, r_m)] \cong f(\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_m)$ with approximate variance:

$$VAR\,[f(r_1, r_2, \ldots, r_m)] \cong VAR\,(r_1) \left[\frac{\partial f}{\partial r_1} \bigg| \hat{r}_1, \hat{r}_2, \ldots, \hat{r}_m \right]^2 + \cdots$$

$$\cdots + VAR\,(r_m) \left[\frac{\partial f}{\partial r_m} \bigg| \hat{r}_1, \hat{r}_2, \ldots, \hat{r}_m \right]^2 .$$

Considering the general series-parallel configuration, the equation for the reliability of this system is:

$$(1) \quad \begin{cases} \hat{R} = [1-(1-\hat{r}_1)^a]\ [1-(1-\hat{r}_a)^b] \cdots\cdots \cdot [1-(1-\hat{r}_m)^k] \\ \text{or} \\ \hat{R} = \prod_{i=1}^{m} [1-(1-\hat{r}_i)^y] \end{cases} .$$

Equation (1) corresponds to the expected value of the function. The variance of equation (1) is

$$(2) \qquad \sigma^2_{\hat{R}} = \left(\frac{\partial \hat{R}}{\partial \hat{r}_1}\right)^2 \sigma^2_{\hat{r}_1} + \cdots + \left(\frac{\partial \hat{R}}{\partial \hat{r}_m}\right)^2 \sigma^2_{\hat{r}_m} .$$

Consider that $\hat{r}_i$ = number of successful functionings/total number fired or tested, where $\hat{r}_i$ is a best estimate of a proportion describing a population where a proportion r_i of the individuals have a certain characteristic and a proportion $1-r_i$ of the individuals do not have it. If $\hat{r}_i$ is the best estimate of some binomial parameter r_i, then the variance of $\hat{r}_i$ is:

$$(3) \qquad \sigma^2_{\hat{r}_i} = \frac{\hat{r}_i(1-\hat{r}_i)}{n_i} \quad .$$

Where n_i = total number tested of item i. The general term for equation (2) is:

$$(4) \qquad \left(\frac{\partial\hat{R}}{\partial\hat{r}_i}\right)^2 \sigma^2_{\hat{r}_i} = \{[1-(1-\hat{r}_1)^a] \cdots [1-(1-\hat{r}_{i-1})^x]\,[1-(1-\hat{r}_{i+1})^z] \cdots$$
$$\cdots [1-(1-\hat{r}_m)^k]\,[y(1-\hat{r}_i)^{y-1}]\}^2 \; \frac{\hat{r}_i(1-\hat{r}_i)}{n_i} \quad .$$

The total variance of the system reliability estimate is:

$$(5) \qquad \sigma^2_{\hat{R}} = \sum_{i=1}^{m} \left(\frac{\partial\hat{R}}{\partial\hat{r}_i}\right)^2 \sigma^2_{\hat{r}_i} \quad .$$

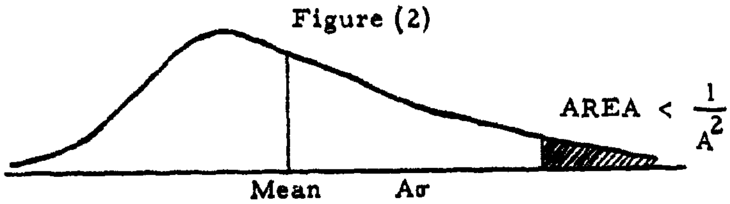
Equation (1) describes the nominal value of the true system reliability R, namely $\hat{R}$ and equations (4) and (5) give the variance of the system estimates. In the area of convential ammunation reliability, we are interested in computing the lower 90% confidence limit of R. This is done in the usual way:

$$(6) \qquad 90\%\ \text{C.L.}\ \ R = \hat{R} - A\,\sigma_{\hat{R}}$$

where "A" depends on the distribution of R.

Since we have no data for the overall system performance (reference second part of Statement of Problem in this report) it was decided to use distribution-free methods - see reference [1]. Chebyshev's inequality states that the amount of area under any distribution which is farther away from the mean than "A" standard deviation units is less than $\frac{1}{A^2}$. This is described in figure (2) below:

Figure (2)



"A" is determined so that at least 90% of the distribution is explained; i.e., the shaded area must be no larger than 10%. Applying the above theorem:

$$.10 = \frac{1}{A^2}$$

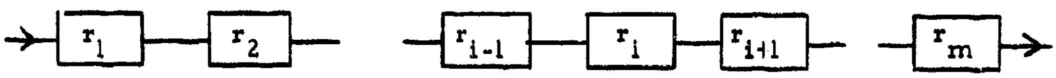
$$A = 3.16$$

therefore equation (6) becomes:

$$(7) \qquad 90\%\ \text{C L. } R \geq \hat{R} - 3.16\,\sigma_{\hat{R}} \quad .$$

SERIES SYSTEMS: The most common case of conventional ammunition reliability assessments have been on systems without replicated components. Referencing figure (1) and letting a=b=...=k=1 we have the following condition:

Figure (2)



Equation (1) becomes:

$$(8) \qquad \hat{R} = \hat{r}_1 \cdot \hat{r}_2 \cdots \hat{r}_m = \prod_{i=1}^{m} \hat{r}_i \quad .$$

Equation (4) becomes:

$$(9) \qquad \left(\frac{\partial \hat{R}}{\partial \hat{r}_i}\right)^2 \sigma^2_{\hat{r}_i} = (\hat{r}_1 \cdot \hat{r}_2 \cdots \hat{r}_{i-1} \cdot \hat{r}_{i+1} \cdots \hat{r}_m)^2 \; \frac{\hat{r}_i (1-\hat{r}_i)}{n_i} \quad .$$

So that equation (5) is:

$$(10) \qquad \sigma^2_{\hat{R}} = \sum_{i=1}^{m} (\hat{r}_1 \cdot \hat{r}_2 \cdots \hat{r}_{i-1} \hat{r}_{i+1} \cdots \hat{r}_m)^2 \; \frac{\hat{r}_i (1-\hat{r}_i)}{n_i} \quad .$$

The values obtained by equations (8) and (10) are then "stuffed into" equation (7) to obtain the lower 90% confidence limit on the system reliability.

Example: Consider two (2) elements in series:

Applying equations (8) and (10):

$$\hat{R} = \hat{r}_1 \cdot \hat{r}_2$$

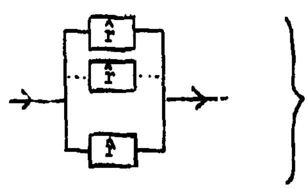
$$\sigma^2_{\hat{R}} = \frac{\hat{r}_2^{\,2}\, \hat{r}_1 (1-\hat{r}_1)}{n_1} + \frac{\hat{r}_1^{\,2}\, \hat{r}_2 (1-\hat{r}_2)}{n_2}$$

or

$$\sigma^2_{\hat{R}} = \hat{r}_1 \cdot \hat{r}_2 \left[\frac{\hat{r}_2 (1-\hat{r}_1)}{n_1} + \frac{\hat{r}_1 (1-\hat{r}_2)}{n_2} \right] .$$

Figure (3) on the next page gives the estimates of $\hat{R}$ and $\sigma^2_{\hat{R}}$ for 2 through 5 components connected in series.

PARALLEL SYSTEMS: If S components are arranged in a parallel configuration, each component with reliability $\hat{r}$ and all S components must fail for the system to fail, then by applying equations (1), (5) and (6):



"S" components

$$\hat{R} = 1-(1-\hat{r})^s$$

and $\sigma^2_{\hat{r}} = \frac{\hat{r}(1-\hat{r})}{n}$

$$\left(\frac{\partial \hat{R}}{\partial \hat{r}}\right)^2 = s^2 \left[(1-\hat{r})^{s-1}\right]^2 .$$

Since:

$$\sigma^2_{\hat{R}} = \left(\frac{\partial \hat{R}}{\partial \hat{r}}\right)^2 \sigma^2_{\hat{r}}$$

FIGURE 3

SUMMARY OF SERIES SYSTEM FORMULAE

No. of Components in System	Estimate of Reliability ($\hat{R}_s$)	Estimate of Reliability Variance ($\sigma^2_{\hat{R}_s}$) [ALL $r \equiv \hat{r}$]
2	$\prod_{i=1}^{2} \hat{r}_i$	$\prod_{i=1}^{2} r_i \left[\frac{r_2(1-r_1)}{n_1} + \frac{r_1(1-r_2)}{n_2} \right]$
3	$\prod_{i=1}^{3} \hat{r}_i$	$\prod_{i=1}^{3} r_i \left[\frac{r_2 r_3 (1-r_1)}{n_1} + \frac{r_1 r_3 (1-r_2)}{n_2} + \frac{r_1 r_2 (1-r_3)}{n_3} \right]$
4	$\prod_{i=1}^{4} \hat{r}_i$	$\prod_{i=1}^{4} r_i \left[\frac{r_2 r_3 r_4 (1-r_1)}{n_1} + \frac{r_1 r_3 r_4 (1-r_2)}{n_2} + \frac{r_1 r_2 r_4 (1-r_3)}{n_3} + \frac{r_1 r_2 r_3 (1-r_4)}{n_4} \right]$
5	$\prod_{i=1}^{5} \hat{r}_i$	$\prod_{i=1}^{5} r_i \left[\frac{r_2 r_3 r_4 r_5 (1-r_1)}{n_1} + \frac{r_1 r_3 r_4 r_5 (1-r_2)}{n_2} + \frac{r_1 r_2 r_4 r_5 (1-r_3)}{n_3} + \frac{r_1 r_2 r_3 r_5 (1-r_4)}{n_4} + \frac{r_1 r_2 r_3 r_4 (1-r_5)}{n_5} \right]$
		IN ALL SERIES CASES: 90% C.L. $R = \hat{R}_{sys} - 3.16\, \sigma_{\hat{R}_s}$

then

$$\sigma_{\hat{R}} = S(1-\hat{r})^{s-1}\left(\frac{\hat{r}(1-\hat{r})}{n}\right)^{\frac{1}{2}}$$

therefore:

(11) $$90\%\ \text{CL.}\ R = 1-(1-\hat{r})^{s}-(3.16)(S)(1-\hat{r})^{s-1}\left(\frac{\hat{r}(1-\hat{r})}{n}\right)^{\frac{1}{2}} .$$

APPROXIMATION FORMULAE. The aforementioned equations can be cumbersome to work with and (as mentioned earlier in the paper) it might be useful if some approximation technique could be used in its place.

Series Case. Consider the estimate of the system reliability $\hat{R}$ where -

(8) $$\hat{R} = \prod_{i=1}^{m} \hat{r}_i .$$

Suppose we were to "assume" that $\hat{R}$ was on estimate of some binomial parameter R. The estimated variance $[\hat{\sigma}^2_{\hat{R}}]$ of σ^2_R would be:

(12) $$\hat{\sigma}^2_{\hat{R}} = \frac{\hat{R}(1-\hat{R})}{n} .$$

The value of n would be chosen so that $\sigma^2_{\hat{R}}$ would be a conservative maximum.

If the sample sizes n_i are the same, then this common sample size should be used in the denominator; if $n_i \neq n_j$, then the minimum value of n should be used in order to maximize $\hat{\sigma}^2_{\hat{R}}$. Equation (12) can be re-written as:

(12)' $$\hat{\sigma}^2_{\hat{R}} = \frac{\hat{R}(1-\hat{R})}{n(\min)} .$$

Equation (7) now becomes:

(13) $$90\%\ \text{C.L.}\ R \geq \hat{R} - 3.16\,\hat{\sigma}_{\hat{R}} .$$

The difference between the proposed solution (equation 7) and its approximation (equation 13) was examined in a general fashion. From this cursory investigation, the differences between these methods can range from about 0.00 (two significant decimal places - see DeCicco) to roughly 0.05 at extreme conditions. It appears that this approximation becomes more effective for $\hat{r}_i \to 1.0$ or large values of n_i or both.

Parallel Case: The estimate of the system variance (equation 12) could be applied to parallel system configurations. The range of differences has not been investigated but it is expected to be in close agreement with the series situation.

GRAPHIC PROCEDURE: As stated earlier, it would be useful to reduce the computations of both approaches to the problem.

Proposed Solution - Series Case - Equation (4) can be substituted into equation (2) to express the total variance of the system reliability estimate (expanded form of equation 5):

(14)

$$\sigma^2_{\hat{R}} = \sum_{i=1}^{m} \{[1-(1-\hat{r}_1)^a] \ldots [1-(1-\hat{r}_{i-1})^x] [1-(1-\hat{r}_{i+1})^z] \ldots [1-(1-\hat{r}_m)^k] [y(1-\hat{r}_i)^{y-1}]\}^2 \times \frac{\hat{r}_i(1-\hat{r}_i)}{n_i} .$$

In the series case a=x=y=z=k=...=1, and by factoring out $\prod_{i=1}^{m} \hat{r}_i$ we obtain:

$$(15) \qquad \sigma^2_{\hat{R}} = \hat{R} \sum_{i=1}^{m} \frac{\hat{r}_1 \cdot \hat{r}_2 \cdots \hat{r}_{i-1} \hat{r}_{i+1} \cdots \hat{r}_m (1-\hat{r}_i)}{n_i} .$$

By letting

$$(16) \qquad \phi_i = \hat{r}_1 \hat{r}_2 \cdots \hat{r}_{i-1} \hat{r}_{i+1} \cdots \hat{r}_m (1-\hat{r}_i) .$$

Equation (15) becomes

$$(17) \qquad \frac{\sigma^2_{\hat{R}}}{\hat{R}} = \sum_{i=1}^{m} \frac{\phi_i}{n_i} .$$

For specific paired values of ϕ_i and n_i (which are computed from sample data) we can set up a graph of the form:

$$\Phi_i = \frac{\phi_i}{n_i}$$

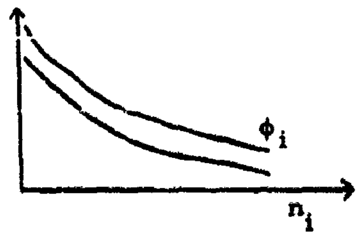


Figure (4)

Defining $\frac{\phi_i}{n_i} = \Phi_i$, equation (17) becomes

$$\frac{\sigma^2_{\hat{R}}}{\hat{R}} = \sum_{i=1}^{m} \Phi_i \qquad \text{or} \qquad \sigma_{\hat{R}} = \left[\left(\sum_{i=1}^{m} \Phi_i\right)(\hat{R})\right]^{\frac{1}{2}}$$

so that equation (7) can be written as:

$$90\%\ \text{C L. } R \geq \hat{R} - 3.16 \left[\left(\sum_{i=1}^{m} \Phi_i\right)(\hat{R})\right]^{\frac{1}{2}}$$

which is defined by $\Sigma \Phi_i$ and $\hat{R}$. A graph can be set up -

90% C.L. $R \geq$

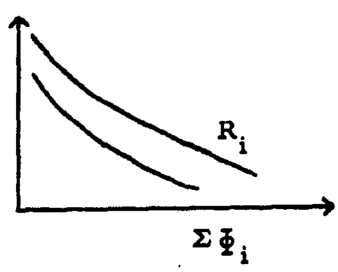


Figure (5)

to give the proposed solution. The range of values for n_i, m_i and $\hat{R}$ were considered to fall in the following intervals:

$$25 \le n_i \le 200$$

$$.85 \le \hat{R} \le .99$$

$$2 \le m_i \le 5 \quad .$$

From these initial boundary intervals the range of values for ϕ_i and $\Sigma \Phi_i$ were determined to be:

$$.005 \le \phi_i \le .150$$

$$50 \times 10^{-6} \le \Sigma \Phi_i \le 30,000 \times 10^{-6} \quad .$$

Figures 4 and 5 have been worked out per the above ranges of values and are presented in the appendix as Figures 4'a, b and 5'a, b.

Approximate Solution - Series Case: Equation (13) is:

$$90\% \text{ C.L. } R \ge \hat{R} - 3.16 \left[\frac{\hat{R}(1-\hat{R})}{n(\text{Min})}\right]^{\frac{1}{2}}$$

which is defined by $\hat{R}$ and n(Min). These parameters will be assumed to have the following range of values-

$$.85 \le \hat{R} \le .99$$

$$25 \le n(\min) \le 200$$

- so that the following graph can be determined:

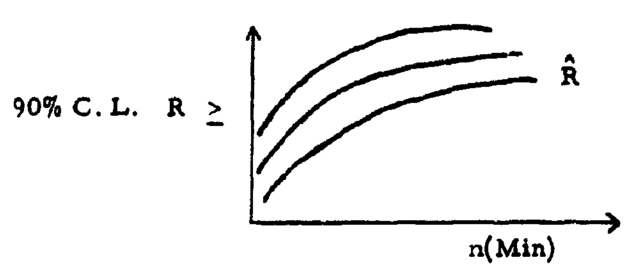


Figure (6)

The details for Figure (6) are given in the appendix as Figure 6'.

This method is certainly much easier to use than any of the previous methods.

Proposed Solution - parallel case - Consider equation (11)

90% C. L. $R = 1-(1-\hat{r})^{s} - (3.16)(S)(1-\hat{r})^{s-1}\left[\frac{\hat{r}(1-\hat{r})}{n}\right]^{\frac{1}{2}}$ which is defined by s number of components in parallel, $\hat{r}$=reliability of each component in the parallel network and n=the same size used to compute $\hat{r}$. For practical purposes let s = 2 and 3 components in parallel.

The following graphs can be constructed:

90% C.L. R $\geq$

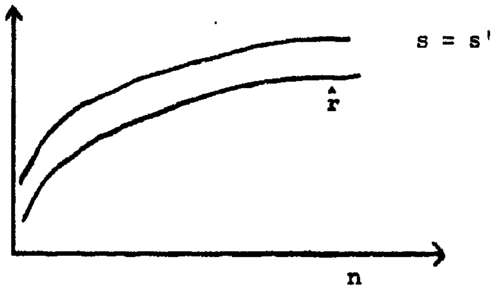


Figure (7)

The details are given in the appendix as figure 7'a,b.

CONCLUDING REMARKS: The above procedure is a proposed "type" of answer examplifying the kind of solution requested. Any solutions to this problem that can be published/circulated as a standard reference would be appreciated.

REFERENCES

Some solutions to this problem can be found/derived from the information contained in the following references:

1. DeCicco, H., The Reliability of Weapon Systems Estimated From Component Test Data Alone; Reliability Branch ORDSW-DR, 1 December 1959.

2. Dowling, J., Computer Simulations in Reliability; 8th Conference on the Design of Experiments.

3. Kniss, J., Reliability Estimation for Multi-Component Systems; 9th Conference on the Design of Experiments.

4. Lloyd, D. K., Lipow, M., Reliability, Management and Mathematics; Prentice Hall, Inc., Englewood Cliffs, N. J., 1962.

5. Orkand, D., A Monte Carlo Method for Determining Lower Confidence Limits for System Reliability on the Basis of Sample Component Data: Picatinny Arsenal report, June 1960.

For those readers who wish to persue this problem further I recommend the following additional references:

6. Cohn, A., Reliability In Complex Systems, paper given at the 12th Conference of the Design of Experiments, 1966.

7. Rosenblatt, J., "Confidence Limits for the Reliability of Complex Systems" - section 4 (zero-one components) printed in Statistical Theory of Reliability, Marvin Zelen (Editor), The University of Wisconsin Press, 1962.

APPENDIX

FIGURE 4'a

PHI-FACTOR RATIOS

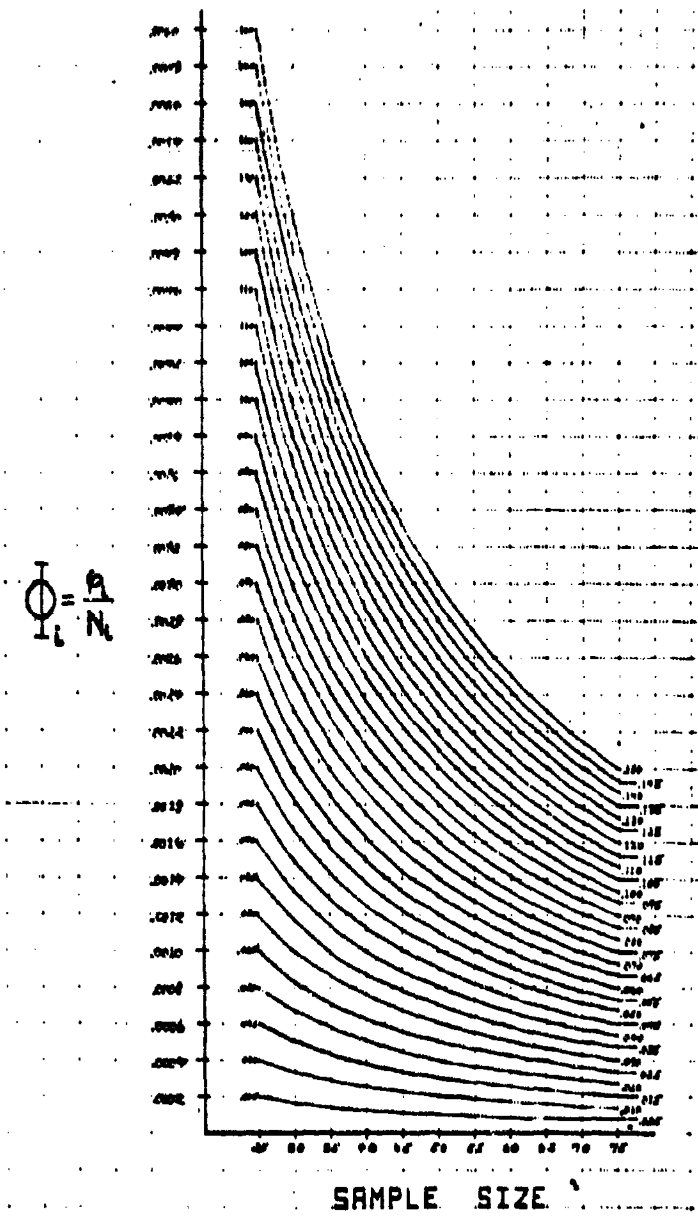


FIGURE 4'b

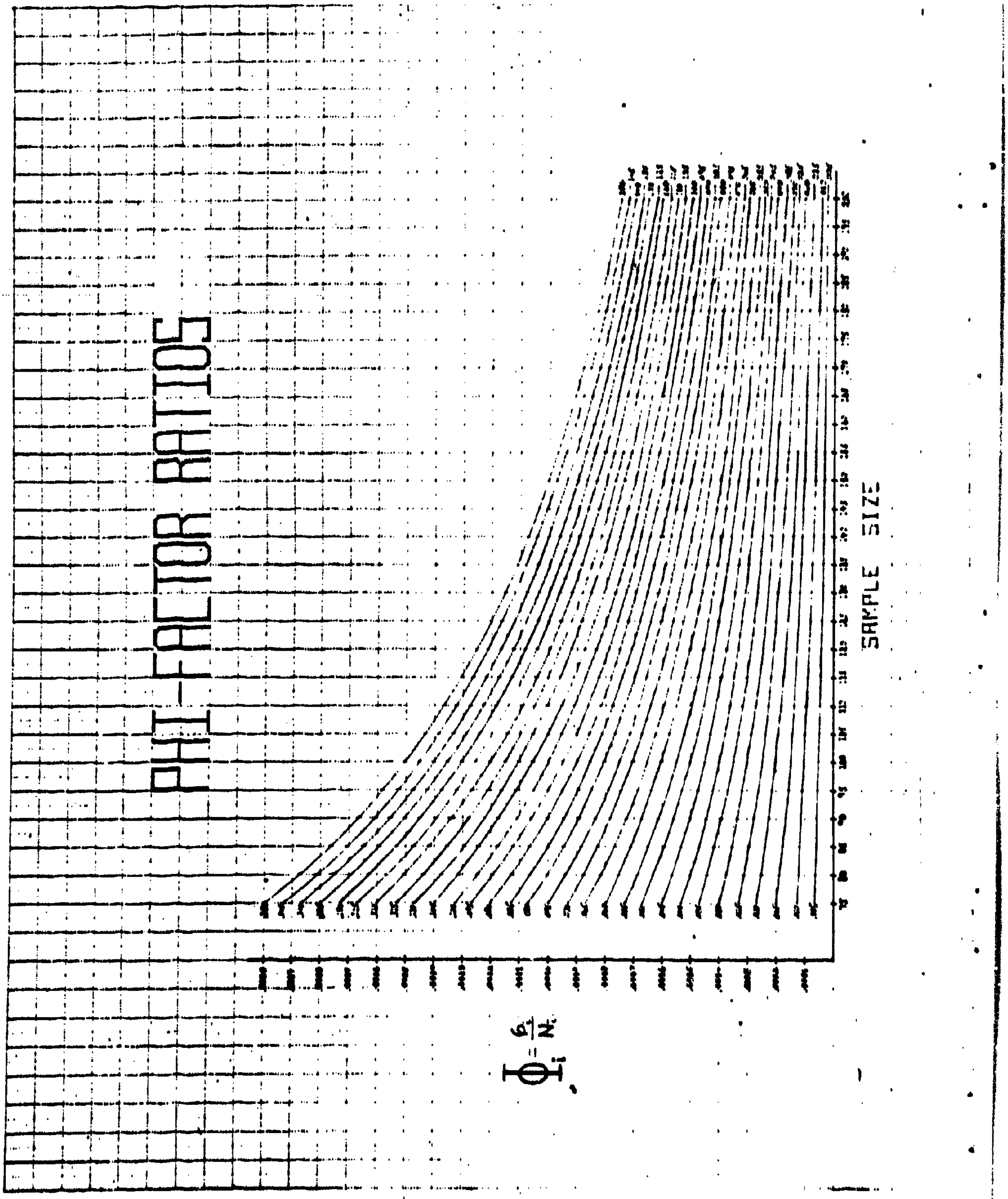


FIGURE 5'a

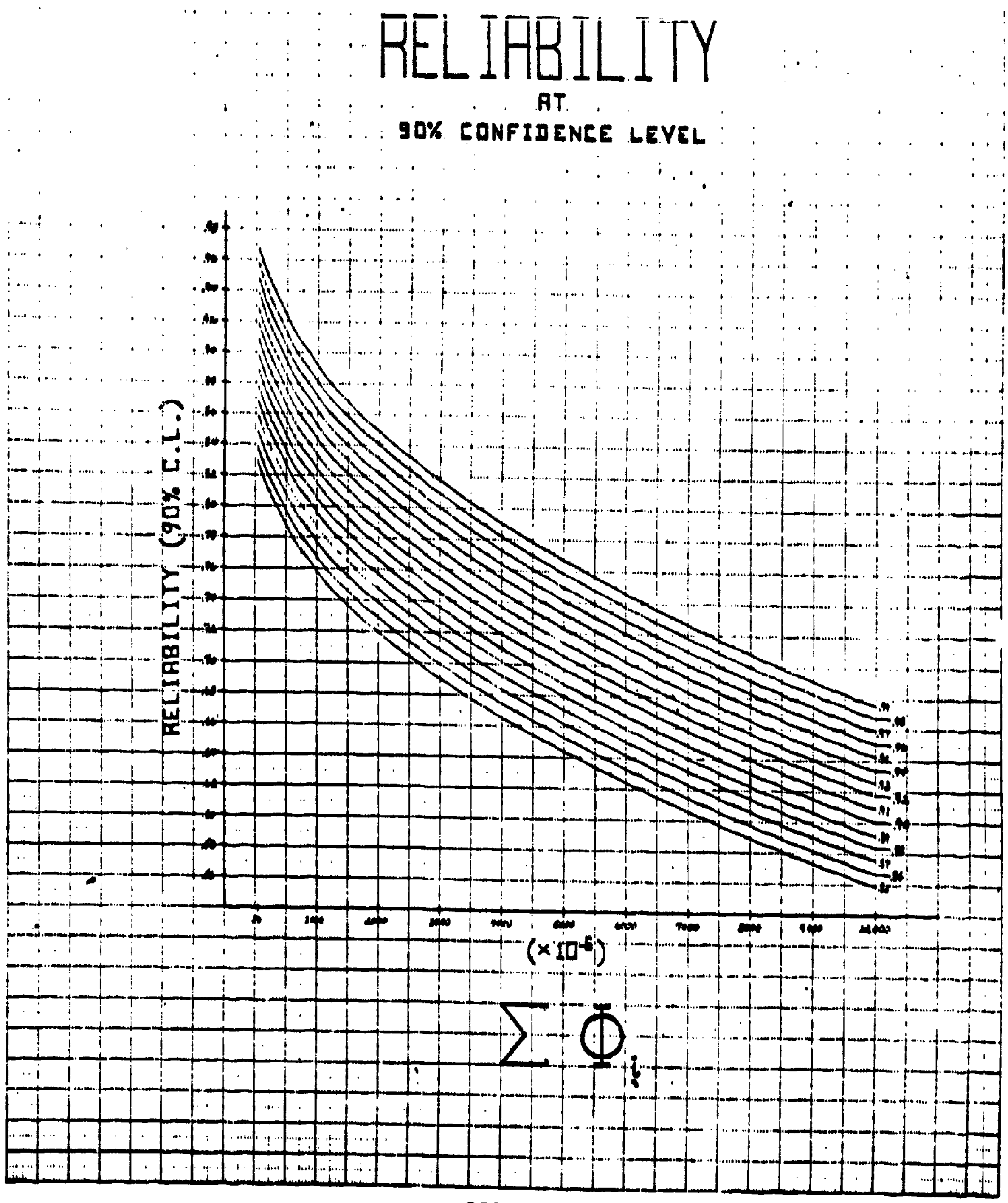
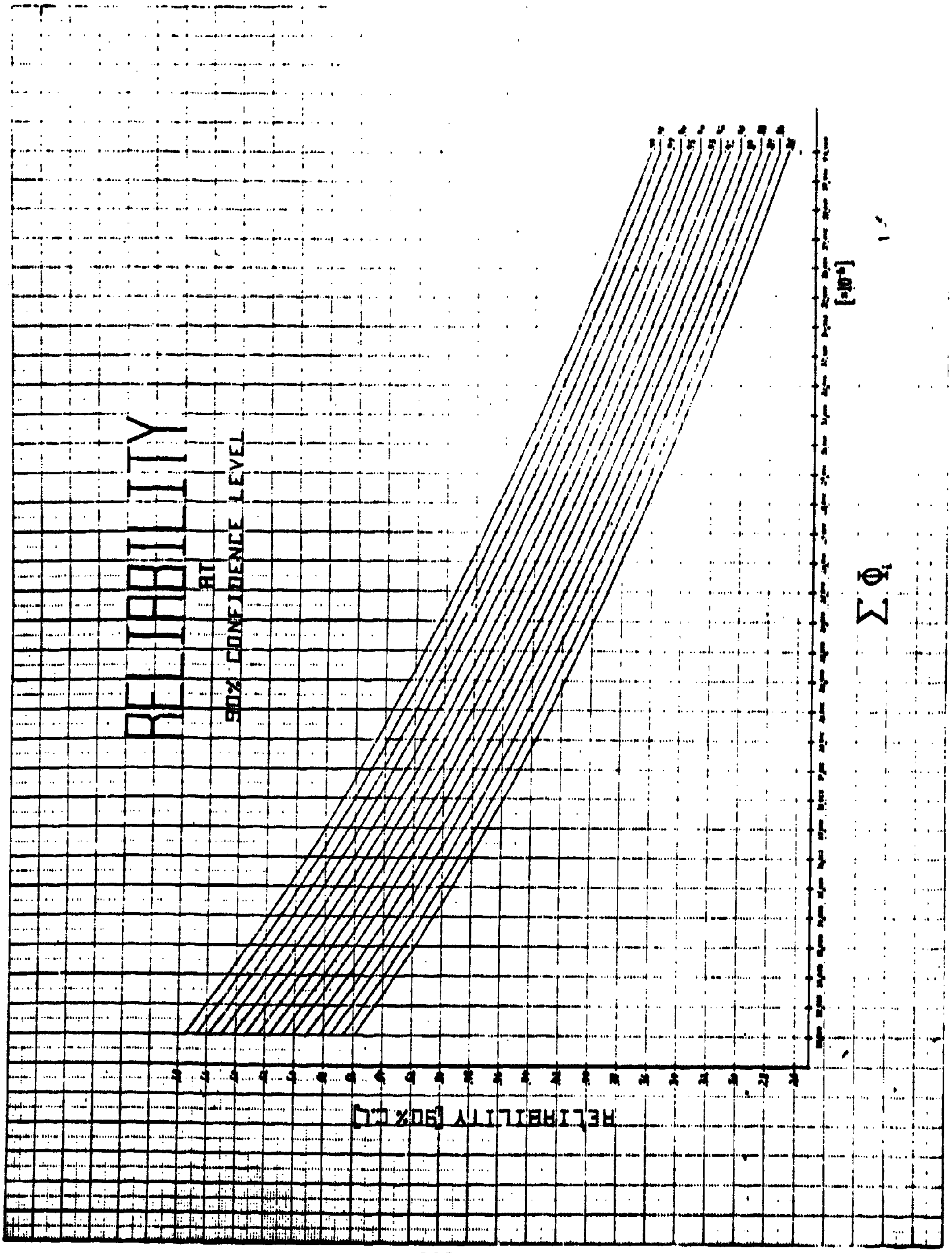

RELIABILITY
AT
90% CONFIDENCE LEVEL
RELIABILITY (90% C.L.)
(×10⁻⁶)
ΣΦ

FIGURE 5'b



RELIABILITY
AT
90% CONFIDENCE LEVEL
RELIABILITY [90% C.L]
$\sum \Phi_i$

FIGURE 6

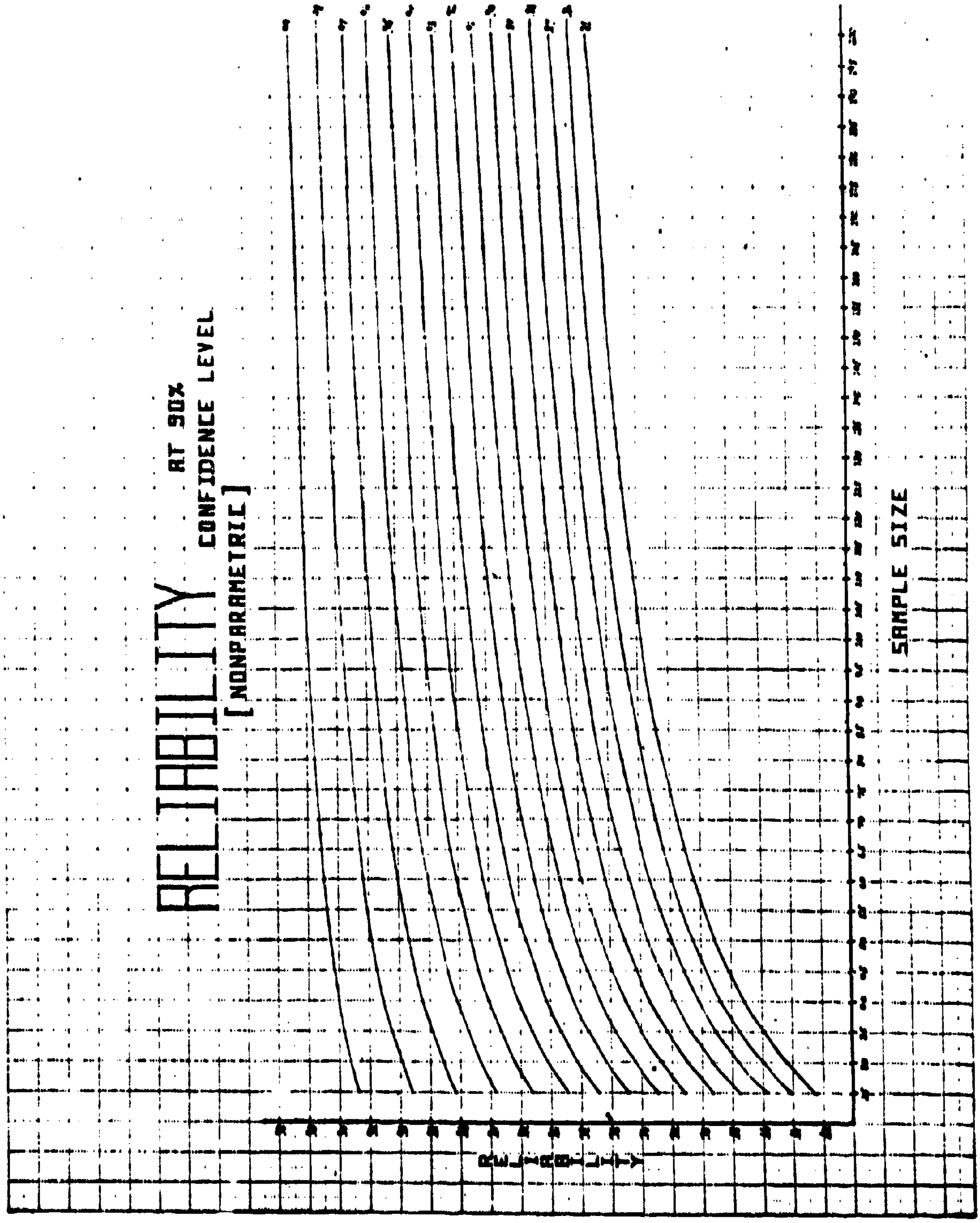

RELIABILITY
AT 90%
CONFIDENCE LEVEL
[NONPARAMETRIC]
SAMPLE SIZE
RELIABILITY

FIGURE 7'a

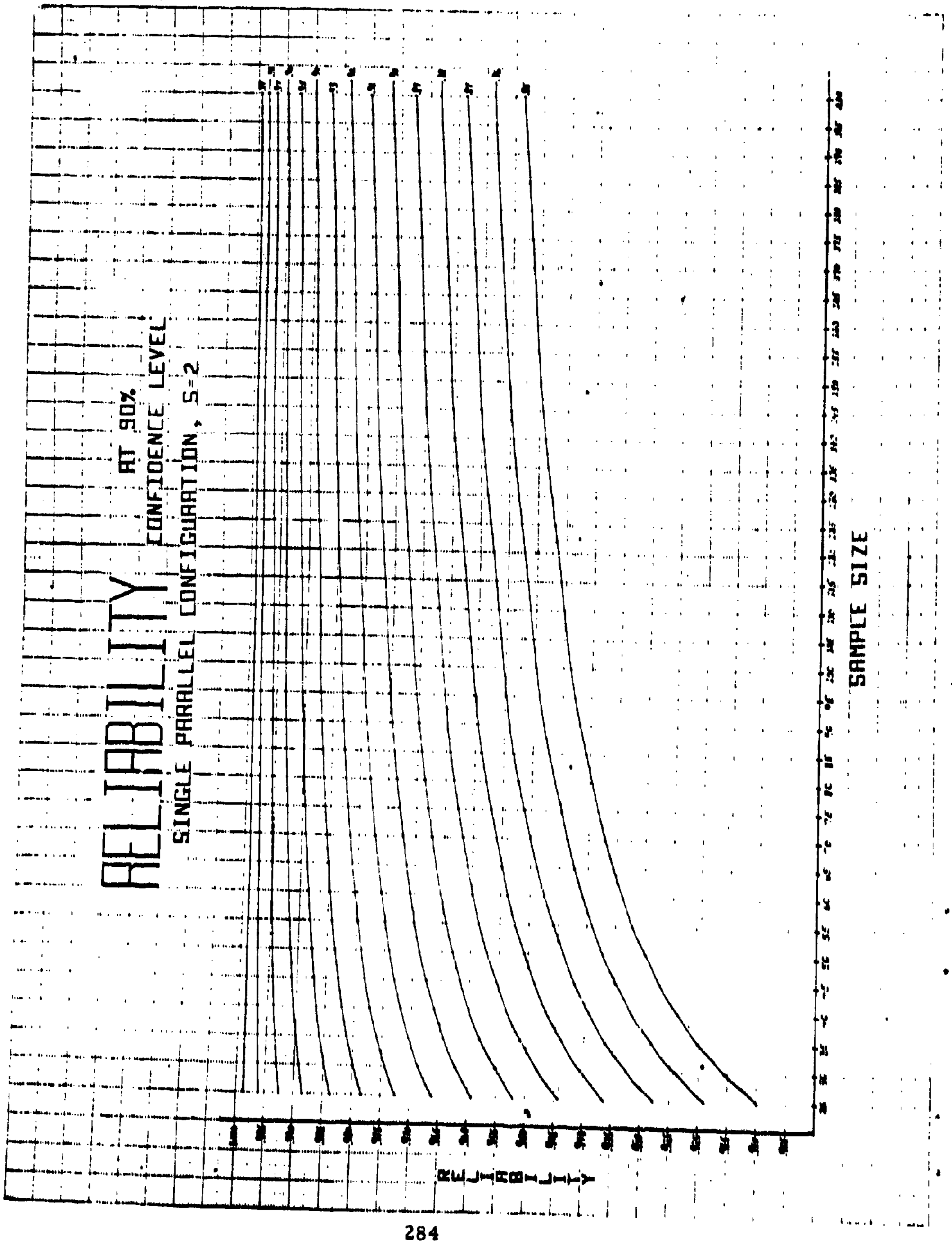
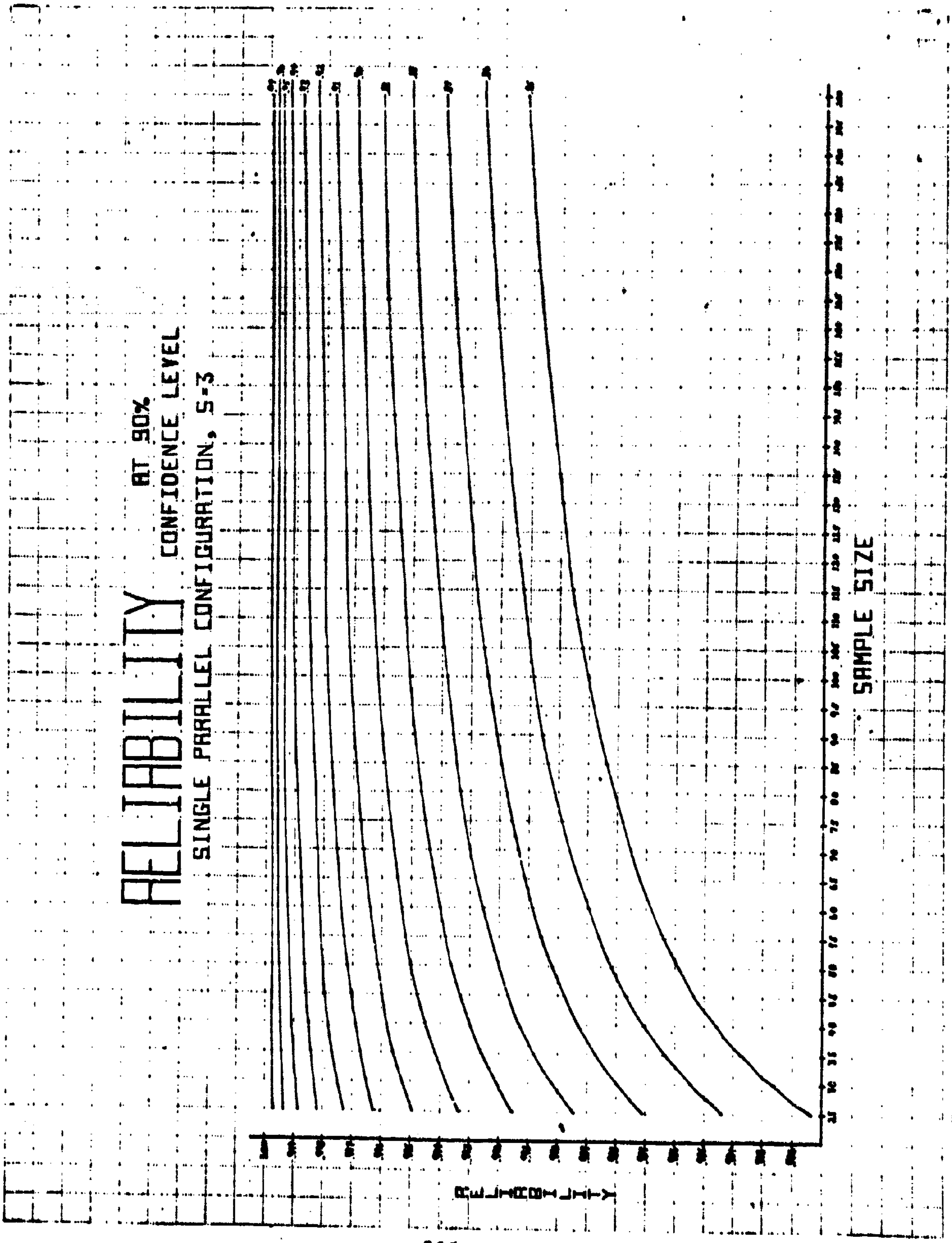

RELIABILITY
AT 90%
CONFIDENCE LEVEL
SINGLE PARALLEL CONFIGURATION, S=2
SAMPLE SIZE
RELIABILITY

FIGURE 7'b


RELIABILITY
AT 90%
CONFIDENCE LEVEL
SINGLE PARALLEL CONFIGURATION, S=3
SAMPLE SIZE
RELIABILITY

STATISTICS, PROBABILITY, AND DETERMINISM IN A RELIABILITY IMPROVEMENT PROGRAM

Woodie R. Jenkins, Jr.
National Range Operations
White Sands Missile Range, New Mexico

The Data Collection Directorate of White Sands Missile Range (WSMR) is presently engaged in the task of increasing the probability of obtaining usable data from several data gathering systems. These systems are used on various projects to collect vehicle performance data. The projects are tests of weapon systems. The data gathering systems are optical cameras and electronic instruments used to measure the position, velocity, attitude, events, and internal status of test vehicles. The probability of obtaining usable data is the "Reliability" that is referred to in this paper. Data records are obtained by instruments of the optical and electronic systems, and the records are assessed "Usable in Reduction" or "Unusable in Reduction" by the WSMR Data Reduction personnel.

It is the policy of the Data Collection Directorate to allow a data gathering system to exhibit a total fraction of unusable records, over a given time period, that does not exceed P_o. In other words, if U = the number of unusable records and I = the number of attempts to obtain data, then the fraction of unusable data obtained by a system over a given period of time is

$$P = \frac{U}{I} \quad (1)$$

[Note that U/I is a measure of the unreliability of the system, and one minus the unreliability is the reliability of the system.] And, in order for the process of obtaining usable data to perform in an acceptable manner,

$$P \text{ Must be } \leq P_o .$$

When, over a specified period of time, $P > P_o$, then the Directorate must take action to improve its data gathering reliability.

It is the $P > P_o$ problem that we address ourselves to in this paper.

The question to be answered is "What action must the Directorate take in order to ensure that P will be $\leq P_o$ for the next equal sampling period?" It is my hypothesis that "The areas that <u>should</u> be controlled can be found by determining the most significant differences between the deterministic relationships that existed at the time the <u>unusable</u> records were obtained by specific instruments and the relationships that existed at the time the usable records were obtained. This requires that the same instruments at the same locations be operated by the same personnel on the same projects in both cases. Moreover, hypotheses about how to control physically the appropriate deterministic relationships can be formulated, tested, and verified with satisfactory results."

[If other hypotheses are made available, I will certainly consider them.]

Once the relationships or parameters that must be controlled are known, then statement(2) of my hypothesis can be performed.

The following example illustrates how statement (1) of my hypothesis can be accomplished.

Let us say that we must assure ourselves that the $P > P_o$ condition for sample (1) will be a $P \leq P_o$ condition for sample (2) for the tracking camera system (cinetheodolites). Sample (1) is the original data for which $P > P_o$. Sample (2) is the necessary and sufficient amount of data needed to make a decision about whether the controlled process yields $P \leq P_o$. The following observation was obtained from all of the sample (1) data.

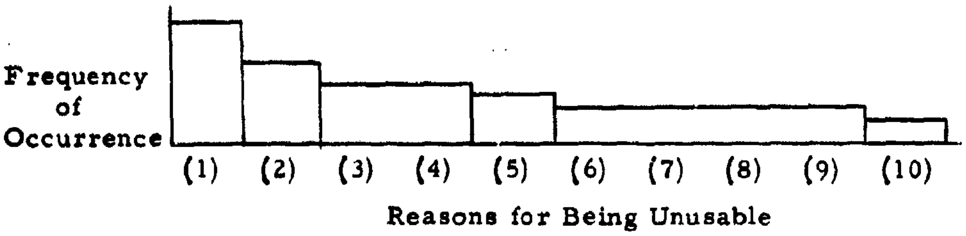


Figure I

From the definition of P [Eq. (1)] it can be seen that

$$(2) \qquad P = \frac{\sum_{\text{Reasons}} \text{Frequencies}}{I} .$$

If there is no reason to expect that the P for sample (2) will be significantly different from the P for sample (1) if the process were left unchanged and if

$$(3) \qquad \frac{\sum_{\text{Reasons}} \text{Frequencies} - \text{Frequency of Reason (1)}}{I} \leq P_o ,$$

then each occurrence of reason (1) should be analyzed for the deterministic conditions or relationships that existed at the time that thé data records were obtained.

If reason (1) is identified as: "Insufficient Coverage", then the equation describing the probability of obtaining "Sufficient Coverage" by a camera is derived as follows.

Sufficient coverage is defined as the required number of consecutive frames of data, M_o, for any optical system. If a cinetheodolite is assigned to operate on a project from time t'_a to time t'_n at a data gathering rate of $\hat{r}$ frames per unit time, then the expected total number of frames of data is

(6) $$\hat{r}\ (t'_n - t'_a)\ .$$

If r = the obtained frame rate and $[t_a, t_n]$ is the time interval over which the camera operated, then the total number of frames of data obtained is

(7) $$r\ (t_n - t_a)\ .$$

Also, if $\theta_{o,t}$ and $\Phi_{o,t}$ = azimuth and elevation angles respectively of the optical axis of the camera at time t, if θ_t and Φ_t = the azimuth and elevation angles respectively of the aerial target to be tracked at time t, and if β_θ and β_Φ are the angular sized of the camera' field of view in the horizontal and vertical planes of the camera respectively, then it can be shown that the aerial target is contained in the camera's field of view it and only if

$$\theta_{o,t} + \frac{57.3}{2}\beta_\theta > \theta_t\ \theta_{o,t} - \frac{57.3}{2}\ \beta_\theta$$

(8) and

$$\Phi_{o,t} + \frac{57.3}{2}\beta_\Phi > \Phi_t > \Phi_{o,t} - \frac{57.3}{2}\beta_\Phi$$

are satisfied simultaneously. Note that since θ and Φ are in degrees and β_θ and β_Φ are in radians, 57.3 coverts β_θ and β_Φ into degrees. If we call A_t the probability of the camera acquiring the aerial target at the instant t and use the concept of Delta Functions, then

$$(9)\quad A_t = \begin{cases} 1, & \theta_{o,t} + \frac{57.3}{2}\beta_\theta > \theta_t > \theta_{o,t} - \frac{57.3}{2}\beta_\theta \text{ and} \\ & \Phi_{o,t} + \frac{57.3}{2}\beta_\Phi > \Phi_t > \Phi_{o,t} - \frac{57.3}{2}\beta_\Phi \\ 0, & \text{Otherwise} \end{cases}$$

Since we must have at least M_o number of consecutive frames of $A_t = 1$ over $[t_a, t_n]$ in order to have sufficient coverage, then a concise mathematical statement of the required condition is defined as follows.

The camera operates at a frame rate of r frames per second. The time required to obtain one frame of film is Δt, where

$$(10)\quad r = \frac{1}{\Delta t}\ ; \qquad \Delta t = \frac{1}{r}\ .$$

Moreover, if it takes Δt units of time to obtain one frame of film, then it takes $M_o \Delta t$ units of time to obtain M_o consecutive frames of film. Therefore, sufficient coverage is obtained if and only if

$$(11)\quad \sum_{t_1 = t_a}^{t_n - M_o \Delta t} \ \prod_{t = t_1 + \Delta t}^{t_1 + M_o \Delta t} A_t \geq 1\ ,$$

where

$$(12)\quad t_1 = \{t_a,\ t_a + M_o\Delta t,\ t_a + 2M_o\Delta t,\ t_a + 3M_o\Delta t, \ldots, t_n - 2M_o\Delta t,\ t_n - M_o\Delta t\}.$$

If there is any sampled instant in M_o consecutively sampled instants for which $A_t = 0$, then the product term of equation (11) is zero for that series of frames. If all such series of frames yield product terms of zero, then the film record will surely be assessed unusable due to insufficient coverage. Again relying on the Delta Function concept, the probability of having obtained sufficient coverage is given by

$$(13) \qquad C_{[t_a, t_n]} = \begin{cases} 1, & \displaystyle\sum_{t_1 = t_a}^{t_n - M_o \Delta t} \; \prod_{t = t_1 + \Delta t}^{t_i + M_o \Delta t} A_t \geq 1 \\ 0, & \text{Otherwise} \end{cases}$$

Equations (9) and (13) provide a means for attempting to find a physical cause for each occurrence of unusable records due to insufficient coverage. For example, the following relationships can be compared by using both usable and unusable data for each station (camera) that obtained unusable records due to insufficient coverage.

After analyzing Figure II, we will be in a position to formulate hypotheses about how to control physically the relationship(s) exhibiting the most significant differences between the usable and unusable data for a given camera on a specific project.

The above discussion has illustrated my approach to solving the $P > P_o$ problem. Since the proposed method has not been tried as yet, I am seeking an evaluation of the method along with alternate approaches to solving the problem. I will now entertain questions and/or comments about this problem.

STATION (CAMERA)

Relationships	Usable Data	Insufficient Coverage Data	Usable Data Minus Insufficient Coverage Data
$r - \hat{r}$			
$(t_n - t_a) - (t'_n - t'_a)$			
$\sum_{t=t_a}^{t_n} A_t - \sum_{t=t'_a}^{t'_n} A_t$			
$\prod_{t=t_a}^{t_n} A_t - \prod_{t=t'_a}^{t'_n} A_t$			
$[\Delta\theta_{o,t}]_{t_a}^{t_n} - [\Delta\theta_t]_{t_a}^{t_n}$			
$[\Delta\Phi_{o,t}]_{t_a}^{t_n} - [\Delta\Phi_t]_{t_a}^{t_n}$			
$[\Delta\theta_{o,t}]_{\text{Max Possible}} - [\Delta\theta_{o,t}]_{t_a}^{t_n}$			
$[\Delta\Phi_{o,t}]_{\text{Max Possible}} - [\Delta\Phi_{o,t}]_{t_a}^{t_n}$			
$[\Delta\theta_{o,t}]_{\text{Max Possible}} - [\Delta\theta_t]_{t_a}^{t_n}$			
$[\Delta\Phi_{o,t}]_{\text{Max Possible}} - [\Delta\Phi_t]_{t_a}^{t_n}$			

FIGURE II

A COMPUTERIZED PROCEDURE FOR WRITING MATHEMATICAL MODELS FOR SYSTEMS RELIABILITY

John G. Mardo and Anthony J. Ricciardi
Mathematics and Statistics Branch, Nuclear Reliability Division,
Quality Assurance Directorate, Picatinny Arsenal
Dover, New Jersey

I. ABSTRACT. A method for the determination of mathematical models for the reliability of missile adaption kit (AK) systems is presented. The method consists of a computer program, the input of which is a Boolean expression of the system configuration. The program constructs a success-failure tree from the Boolean expression resulting in all possible success paths for the system. The union of these success paths is the reliability model for the system. The number of components and not the complexity with which they are combined limits the use of the present procedure.

II. ACKNOWLEDGEMENT. The assistance of Bruce Barnett of the Data Processing Systems Office of Picatinny Arsenal in the development of the techniques used in this report is acknowledged. In particular, his contributions to the underlying theoretical aspects are appreciated.

The remainder of this article has been reproduced photographically.

III Introduction

The purpose of mathematical estimations of reliability in any stage of the life cycle of a system is to determine the expected probability of successful functioning in use. These estimations should anticipate potential reliability problems and reveal system configurations that have greatest probabilities of failure in use.

An important tool in evaluating these estimates is the reliability equation. This equation is a mathematical model of the system under consideration, relating the reliability of the system to the reliability of the components which comprise it. For complex electrical systems, these equations are difficult to obtain. The difficulties encountered are dependent upon the number of components in the system and the degree of complexity of the configuration.

The dependent operations of the various components require the use of conditional probabilities in developing the mathematical models. The determination of these conditional probabilities is difficult; as a result, reliability equations for complex systems are usually approximations based on the assumption of independence. Simplified models of the system are presently used which ignore the less likely modes of operation.

Although numerical estimates obtained from models which represent the system exactly do not differ markedly from those that would result from approximate methods, there are a number of advantages in using the more exact method. Arguments as to the validity of the model used for the analysis are largely eliminated because the "model", in this case, is the most complete mathematical representation of system operation possible. The ability to handle large numbers of components permits breaking down the system into very small elements and the reliabilities of these small elements can be established with greater confidence and can be established by testing at less expense. Finally, the equations arising from this analysis permit component effect studies on a more realistic level, since a more exact representation of a component's role in the operation of the system is given by the resulting equation.

An automated procedure will be presented for analyzing systems. This procedure results in a reliability equation which is a mathematical model representing the system. The primary purpose of this automated procedure is the determination of success models in the shortest length of time by the most economical means. Complicated networks require months of manual effort to determine reliability models even with the previously discussed approximations. Using the computer procedure to be discussed, it is necessary to understand the logical functioning of the system. With this understanding, it will reqire only a few days rather than months to derive the final algebraic equation using computer techniques.

IV Foundations of the Computerized Procedure

Consider the system composed of the components $C_1, C_2, \ldots, C_n$ arranged in a configuration which makes ordinary parallel-series reliability analysis of the network difficult. Interdependency of component operation causes such a situation. It follows from Baye's Theorem that

$$P(f) = P(f|C_i)P(C_i)+P(f|\bar{C}_i)P(\bar{C}_i) \tag{1}$$

where: $P(f)$ = probability of the system functioning

$P(f|C_i)$ = the probability of the system functioning given that component C_i operates correctly

$P(C_i)$ = probability of component C_i operating correctly

$P(f|\bar{C}_i)$ and $P(\bar{C}_i)$ are defined similarly where $\bar{C}_i$ represents the event where component C_i fails to operate correctly

$P(C_i)$ is the reliability of the component C_i and $P(\bar{C}_i) = 1 - P(C_i)$

Equation (1) would be the desired expression of the system reliability in terms of the component reliabilities if the conditional probabilities $P(f|C_i)$ and $P(f|\bar{C}_i)$ were evaluated either numerically or as functions of the component reliabilities. The computer program listed in Appendix A performs these evaluations of the conditional probabilities by using the Boolean algebraic expression which represents the logic of the system operation. This expression is a function $B(C_1^*, C_2^*, \ldots, C_n^*)$ which takes on the values 1 or 0 representing system success or failure, respectively, where C_i^* is a variable which takes on the value 1 or 0 depending on whether component C_i operates or fails to operate, respectively.

It is possible using this function B to evaluate the conditional probabilities $P(f|C_i)$ and $P(f|\bar{C}_i)$. If, when C_i^* is given the truth value 1 in the Boolean function B, and all other C_j^*, where $j \neq i$, are given truth values 0, the value of B is 1, then $P(f|C_i) = 1.0$. However, if $B = 0$ then $P(f|C_i)$ cannot be determined directly and $P(f|C_i)$ must be further expanded as follows:

$$P(f|C_i) = P(f|C_iC_j)P(C_j)+P(f|C_i\bar{C}_j)P(\bar{C}_j) \text{ for any } j \neq i$$

Similarly, when C_i^* is given the value 0 in B and all other C_j^*, where $j \neq i$, are given values 1, B = 0, then $P(f|\bar{C}_i) = 0.0$. However, if B = 1, then $P(f|\bar{C}_i)$ can not be determined directly and thus may also require further expansion as follows:

$$P(f|\bar{C}_i) = P(f|\bar{C}_iC_j)P(C_j)+P(f|\bar{C}_i\bar{C}_j)P(\bar{C}_j) \text{ for any } j \neq i$$

At this point, an attempt is again made to evaluate the conditional probabilities using the function B. The procedure is continued until all the conditional probabilities have been eliminated by substitution of either their numerical equivalents or these conditional probabilities expressed as combinations of the individual component reliabilities. When this point is reached, the P(f) has been expressed algebraically as a combination of the individual component reliabilities and the program is terminated.

Applying this procedure to the following circuit:

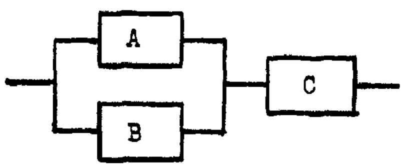


FIG. 1

the Boolean expression for the circuit is

$$\text{SYSTEM} = B(A,B,C) = (A+B)\cdot C$$

Expanding as described above using Baye's Theorem:

1. $P(f) = P(f|A)\cdot P(A)+P(f|\bar{A})\cdot P(\bar{A})$

2. $P(f|A) = P(f|AB)\cdot P(B)+P(f|A\bar{B})\cdot P(\bar{B})$

3. $P(f|\bar{A}) = P(f|\bar{A}B)\cdot P(B)+P(f|\bar{A}\bar{B})\cdot P(\bar{B})$

4. $P(f|AB) = P(f|ABC)\cdot P(C)+P(f|AB\bar{C})\cdot P(\bar{C})$

5. $P(f|A\bar{B}) = P(f|A\bar{B}C)\cdot P(C)+P(f|A\bar{B}\bar{C})\cdot P(\bar{C})$

6. $P(f|\bar{A}B) = P(f|\bar{A}BC)\cdot P(C)+P(f|\bar{A}B\bar{C})\cdot P(\bar{C})$

From the Boolean Expression it follows that:

$P(f|\overline{AB}) = P(f|AB\bar{C}) = P(f|A\overline{BC}) = P(f|\bar{A}B\bar{C}) = 0$ and $P(f|ABC) =$

$P(f|A\bar{B}C) = P(f|\bar{A}BC) = 1$

Hence:

$$P(f) = P(C)\cdot P(B)\cdot P(A)+P(A)\cdot P(C)\cdot P(\bar{B})+P(C)\cdot P(B)\cdot P(\bar{A})$$

$$P(f) = [P(B)\cdot P(A)+P(A)\cdot(1-P(B))+P(B)\cdot(1-P(A))]\cdot P(C)$$

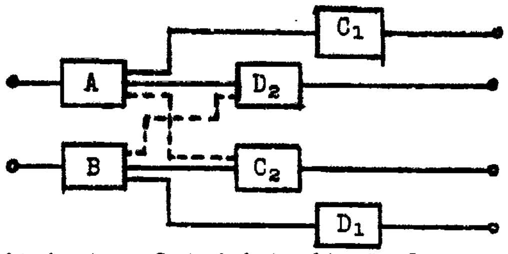
$$P(f) = [P(A)\cdot P(B)+P(A)-P(A)\cdot P(B)+P(B)-P(A)\cdot P(B)]\cdot P(C)$$

$$P(f) = (P(A)+P(B)-P(A)\cdot P(B))\cdot P(C)$$

This result is the algebraic reliability equation for the circuit shown in FIG. 1.

V Application of Computer Procedure

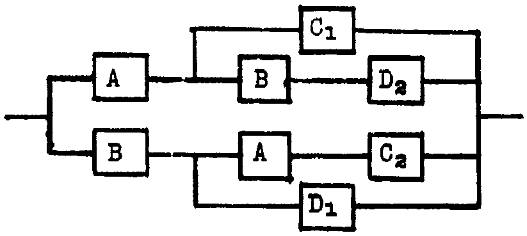
The following example demonstrates the computer method of handling the procedure on a simple circuit. Consider the circuit with components A, B, C_1, D_1, C_2, D_2 in the figure below:



FUNCTIONAL DIAGRAM

FIG. 2

The above circuit is translated into its Boolean or logic diagram:



BOOLEAN OR LOGIC DIAGRAM

FIG. 3

The Boolean logic diagram is converted into a Boolean expression using Boolean algebraic techniques. The resulting expression for the above diagram is as follows:

(1) $\text{SYSTEM} = A(C_1+BD_2)+B(D_1+AC_2)$

This expression is then programmed using whatever means are available in the programming language being used.

The next step is to set up an "order of consideration" of the components. This will be A, B, C_1, D_2, D_1, C_2.

The steps that follow are handled by the computer as follows:

Using the Boolean expression, a "tree" is generated within the computer. Such a tree will now be generated for the circuit under discussion.

Symbol (A) is used to represent the success of component (A). Symbol ($\bar{A}$) is used to represent the failure of component (A). Similar notations are used for components B, C, and D. Starting with component A (since A is the first component under the order of consideration) the associated success-failure symbols (A and $\bar{A}$) are used as the first two branches of the tree. To determine how far to continue a branch, each branch is tested or examined using the following rules:

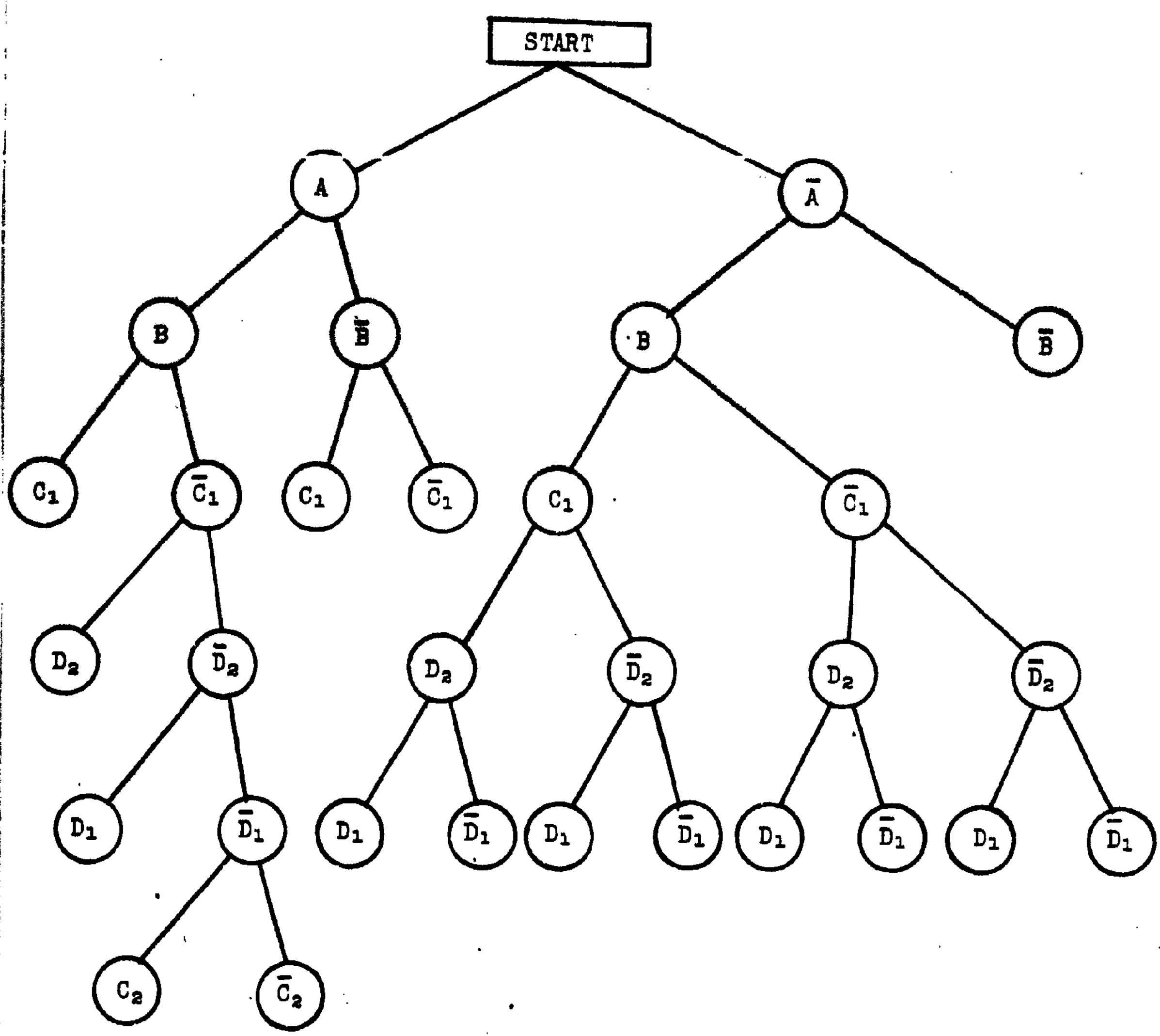
1. Starting with the A branch, assign the value 1 to component (A) and the value 0 to all the remaining components in the system. Substitute these truth values into the Boolean expression for the system and determine whether this combination of values causes a system success or a system failure. If the result is a system success, end the A branch of the tree. If the result is a failure, plan to continue the A branch by adding the two branches (B and $\bar{B}$) of component B.

2. Using the $\bar{A}$ branch, assign the value 0 to component (A) and the value 1 to all remaining components in the system. Substitute these truth values into the Boolean expression for the system and determine whether this combination of values causes a system success or a system failure. If the result is a system failure, end the $\bar{A}$ branch of the tree. If the result is a system success, plan to continue the $\bar{A}$ branch by adding the two branches (B and $\bar{B}$) of component B.

3. Continue to generate the tree diagram by adding components and testing each branch of each component for termination or continuation as described above. The expression that describes the success path to the last component in the branch can be used to develop the algebraic equation for the system.

When all success paths have been generated, the program creates an algebraic success model which can be used for the generation of reliability point estimates for the overall network described by the Boolean expression.

The tree diagram for the circuit of FIG. 2 is shown in FIG. 4 along with the resulting success paths.



SUCCESS PATHS

ABC_1

$AB\bar{C}_1D_2$

$AB\bar{C}_1\bar{D}_2D_1$

$AB\bar{C}_1\bar{D}_2\bar{D}_1C_2$

$A\bar{B}C_1$

$\bar{A}BC_1D_2D_1$

$\bar{A}BC_1\bar{D}_2D_1$

$\bar{A}B\bar{C}_1D_2D_1$

$\bar{A}B\bar{C}_1\bar{D}_2D_1$

FIG. 4

The computer program then saves these susscess paths, substituting for the $\bar{A}$, $\bar{B}$, etc., (1-A), (1-B), etc., respectively as follows:

$$
\begin{aligned}
R\ (\text{SYSTEM}) = {} & ABC_1+ABD_2(1-C_1)+ABD_1(1-C_1)(1-D_2)+ABC_2(1-C_1)(1-D_2)(1-D_1) \\
& +AC_1(1-B)+BC_1D_2D_1(1-A)+BC_1D_1(1-A)(1-D_2)+BD_2D_1(1-A)(1-C_1) \\
& +BD_1(1-A)(1-C_1)(1-D_2)
\end{aligned}
$$

The equation for R (SYSTEM) is then stored in computer memory as follows (see Appendix; RELIABILITY MODEL):

$$
\begin{aligned}
R\ (\text{SYSTEM}) = {} & ABC_1+ABD_2-ABC_1D_2+ABD_1-ABC_1D_1-ABD_1D_2+ABC_1D_1D_2+ABC_2-ABC_1C_2 \\
& -ABC_2D_2+ABC_1C_2D_2-ABC_2D_1+ABC_1C_2D_1+ABC_2D_1D_2-ABC_1C_2D_1D_2+AC_1 \\
& -ABC_1+BC_1D_1D_2-ABC_1D_1D_2+BC_1D_1-ABC_1D_1-BC_1D_1D_2+ABC_1D_1D_2+BD_1D_2 \\
& -ABD_1D_2-BC_1D_1D_2+ABC_1D_1D_2+BD_1-ABD_1-BC_1D_1+ABC_1D_1-BD_1D_2+ABD_1D_2 \\
& +BC_1D_1D_2-ABC_1D_1D_2
\end{aligned}
$$

Noting that A and B are similar components which will always have the same function and the same reliability as will C_1, C_2, D_1, and D_2 the above equation will be reduced to the following by the computer:

$$
\begin{aligned}
R\ (\text{SYSTEM}) = {} & A^2C+A^2C-A^2C^2+A^2C-A^2C^2-A^2C^2+A^2C^3+A^2C-A^2C^2-A^2C^2+A^2C^3-A^2C^2+A^2C^3 \\
& +A^2C^3-A^2C^4+AC-A^2C+AC^3-A^2C^3+AC^2-A^2C^2-AC^3+A^2C^3+A^2C^2-AC^3+A^2C^3 \\
& +AC-A^2C-AC^2+A^2C^2-AC^2+A^2C^2+AC^3-A^2C^3
\end{aligned}
$$

The "combine terms" routine is then applied to obtain the final result, the algebraic success model for the circuit in FIG. 2 or any circuit represented by the logic diagram of FIG. 3.

$$R\ (\text{SYSTEM}) = 2A^2C-6A^2C^2+4A^2C^3-A^2C^4+2AC$$

VI Conclusion

The program to carry out the procedure described above has been developed and tested on many hypothetical systems. (See Appendix) Results of this testing brought to light a few drawbacks to the method. These weaknesses will now be discussed.

On a system consisting of N distinct components, the number of branches which may be considered is 2^N. This number may be reduced greatly if a proper order of consideration of components is used. The procedure is extremely sensitive to this order and efforts are now being made to develop decision mechanisms within the program to construct non-redundant success paths which result from improper order of consideration. An illustration of this redundancy can be shown on the demonstration circuit used above. Because of the order A, B, C_1, etc. used two success paths which result are ABC_1 and $A\bar{B}C_1$. The same contribution that these paths make to the final reliability equation would have resulted had the order been slightly altered; i.e., A, C_1, B, etc. The only success path resulting from this order would have been AC_1 and $P(AC_1) = P(A)P(C_1)$. Notice that $P(ABC_1)+P(A\bar{B}C_1) = P(A)P(B)P(C_1)+P(A)P(\bar{B})P(C_1) = P(A)P(B)P(C_1)+P(A)P(C_1)-P(A)P(B)P(C_1) = P(A)P(C_1)$. Hence, this change in order eliminates the use of two branches to come up with the same contribution to the final algebraic success model. The presence of each causes unreasonable amounts of computer time to be used even when only point estimates rather than the algebraic equations are being computed.

A second problem is caused by the need for large amounts of computer storage. This need arises only when the algebraic equation is being sought, since each success path must be stored in some manner so that final refinement of all success paths as a whole can be made to determine the final model in a well organized form. When only a numerical point estimate is sought, there is, in general, no need to be concerned about memory size.

The third and final problem is a minor one. It results from the cumulative round-off error that is present when many accumulative multiplications are performed with very small numbers while generating numerical estimates of reliability. This problem, however, has largely been overcome due to the availability on most present day computers of the double precision variable.

The present stages of development of the procedure are concerned primarily with overcoming these difficulties. When the flaws are eliminated, the computer program will provide to the engineers a means of predicting and estimating the reliability of their systems. It will provide engineering with the efficiency and accuracy of the computer in determining the reliability success models it requires, saving a good deal of time and money. Reliability equations that previously require months to derive manually, can now be solved in a matter of days.

VII References

1. Moskowitz, F. , "The Nature and Behavior of Redundancy Networks and Functions," Rome Air Development Center, Griffiss AFB, New York, RADC-TR-60-4, May 1960

2. Moskowitz, F., "Properties of Redundancy Networks with Equal Element Reliability Values," RADC-TN-58-146, June 1958

3. Moskowitz, F., "The Analysis of Redundancy Networks," Communications and Electronics AIEE, pp. 627-632, November 1958

4. Weinstock, George D., "Topological Analysis of Non-Series-Parallel Redundant Networks," RCA Surface Communications Systems Laboratories, May 1962

Appendix A

Included in this appendix are a listing of the computer program discussed in the body of this report and an example of the program output. This output resulted from the application of the computer procedure on the simple circuit discussed in Section V of this report. Note that a reliability point estimate was generated for a given set of component reliability values, as well as, the final reliability model.

PROGRAM LISTING

```
      DIMENSION  ICOMP(30),IBRNCH(30),TRUTH(30),ELEMNT(30),IKOUNT(2000),
     1ITERMS(300,30),SIGN(300),AFACTO(30,30),L(30)
      DOUBLE PRECISION B(30),PROB,PROB1,PROB2
      DATA Q000HL/4H    /
      BLANK= Q000HL
      DATA Q001HL/4HN   /
      NEGATE= Q001HL
      DATA Q002HL/4H +  /
      PLUS= Q002HL
      DATA Q003HL/4H -  /
      TINUS= Q003HL
      WRITE (6,601)
  601 FORMAT(1H1)
  100 READ (5,500) N,ICODE,ISECT
  500 FORMAT(3(I3,2X))
C ICODE =+1  CALCULATE POINT ESTIMATE ONLY
C ICODE = 0   EQUATION ONLY
C ICODE =-1  BOTH
      IF(ICODE)89,89,91
   89 READ (5,501) (ELEMNT(LIST),             LIST=1,N)
  501 FORMAT(18A4)
      DO 90 ILIST=1,N
   90 WRITE(6,598)  ELEMNT(ILIST),ILIST
  598 FORMAT(1H0,39X,10HCOMPONENT ,A4,35H IS REPRESENTED AS VARIABLE NUM
     1BER ,I2)
   91 IF(ICODE)92,93,92
   92 READ(5,502)(B(KM),KM=1,N)
      WRITE(6,599)N,(B(KN),KN=1,N)
  599 FORMAT (1H0//54X,22HNUMBER OF COMPONENTS =,I2,/54X,24H-----------
     1------------//(1H0,4(9X,D18.12))/)
C THE ABOVE STATEMENTS HANDLE THE TABULATING OF COMPONENT SYMBOL REPRE-
C SENTATION, RELIABILITIES OR BOTH.
   93 ITERM = 0
      ICOMP(1)= 1
      K=1
    6 IBRNCH(K)=1
    5 ICOMP(K+1)=1
      M=K+2
      DO 10 I=M,N
   10 ICOMP(I)=0
      DO 700 I9=1,N
  700 L(I9)=ICOMP(I9)
      IBOOLE=ICOMP(1)*(ICOMP(2)*ICOMP(4)+ICOMP(3))+ICOMP(2)*(ICOMP(1)*
     1ICOMP(6)+ICOMP(5))
      IF(IBOOLE)71,70,71
   70 IF(K-N+1)13,14,14
   14 WRITE(6,600)
  600 FORMAT(15X,46HNO SUCCESS PATH EXISTS FOR CIRCUIT CONSTRUCTED)
      GO TO 100
```

```
 13 K=K+1
    GO TO 6
 71 ITERM = ITERM + 1
    DO 40 IM=1,N
    IF(ICOMP(IM))42,41,42
 42 L1=IM
    TRUTH(IM)=BLANK
    GO TO 40
 41 TRUTH(IM)=NEGATE
 40 CONTINUE
    IF(ICODE)73,74,73
 73 PROB2 = 1.0
    DO 80 IT2=1,L1
    IF(ICOMP(IT2))81,82,81
 81 PROB1=B(IT2)
    GO TO 80
 82 PROB1=1.0 - B(IT2)
 80 PROB2=PROB2*PROB1
    PROB = PROB + PROB2
    IF(ICODE)74, 2, 2
 74 DO 201 INDEX=1,L1
201 ITERMS(1,INDEX)=INDEX
    KOUNT=0
    K2=1
199 IF(TRUTH(K2) - NEGATE)203,202,203
202 KOUNT = KOUNT + 1
    IKOUNT(KOUNT)= K2
198 K2 = K2 + 1
    GO TO 199
203 IF(K2-L1)198,204,198
204 IF(KOUNT)197,205,197
205 NOTERM= 1
    NUM=1
    SIGN(1)=PLUS
    GO TO 194
197 NOTERM = 2**KOUNT
    NOSIGN = 2*NOTERM
    DO 206 INDICE = 1,NOSIGN
206 SIGN(INDICE)= PLUS
    NUM=1
    KOUNT1=0
    ISTAGE =1
172 MOUNT =KOUNT1+1
    KOUNT1 =KOUNT1 + 2**(ISTAGE - 1)
    DO 180 INUM=MOUNT,KOUNT1
    NEXT = IKOUNT(ISTAGE)
    NUM = NUM +1
    IF(NEXT-1)181,182,181
181 NEXT1=NEXT-1
    DO 150 INDEX1=1,NEXT1
```

```
150 ITERMS(NUM,INDEX1)=ITERMS(INUM,INDEX1)
182 ITERMS(NUM,NEXT)= 0
    NEXT2 = NEXT + 1
    DO 151 INDEX2 = NEXT2,L1
151 ITERMS(NUM,INDEX2) = ITERMS(INUM,INDEX2)
    SIGN(NUM)=SIGN(INUM)
    NUM=NUM + 1
    IF(SIGN(INUM)-TINUS)152,162,152
152 SIGN(NUM) = TINUS
    GO TO 170
162 SIGN(NUM)=PLUS
170 IF(NEXT-1)191,192,191
191 DO 153 INDEX3=1,NEXT1
153 ITERMS(NUM,INDEX3)=ITERMS(INUM,INDEX3)
192 ITERMS(NUM,NEXT) = ITERMS(INUM,NEXT)
    DO 154 INDEX4=NEXT2,L1
154 ITERMS(NUM,INDEX4)=ITERMS(INUM,INDEX4)
    IF(NUM+1-2**(KOUNT+1))180,194,194
180 CONTINUE
    GO TO 200
194 DO 955 J= NOTERM,NUM
    DO 111 IBSO=1,L1
    IF(ITERMS(J,IBSO))112,113,112
112 AFACTO(J,IBSO)=ELEMNT(IBSO)
    GO TO 111
113 AFACTO(J,IBSO)=BLANK
111 CONTINUE
    KNOUT=KOUNT+1
    DO 144 KOUN=1,KNOUT
    DO 114 I=1,LI
    IF(AFACTO(KOUN,I)-BLANK ) 114,331,114
331 TEMP=AFACTO(KOUN,I)
    AFACTO(KOUN,I)= AFACTO(KOUN,I+1)
    AFACTO(KOUN,I+1)=TEMP
114 CONTINUE
144 CONTINUE
155 IF(ITERM-1)955,954,955
954 WRITE(6,593)
955 WRITE(6,605) SIGN(J),(AFACTO(J,I),I=1,L1)
605 FORMAT(1H ,20(A4,1X))
593 FORMAT (1H1,12X,17HRELIABILITY MODEL/1H ,11X,19H-----------------
   1-/)
    GO TO 2
200 ISTAGE = ISTAGE +1
    GO TO 172
  2 IBRNCH(K)=2
    ICOMP(K+1)=0
    M1=K+2
    DO 20 J=M1,N
 20 ICOMP(J)=1
```

```
      DO 701 I8=1,N
  701 L(I8)=ICOMP(I8)
      IBOOLE=ICOMP(1)*(ICOMP(2)*ICOMP(4)+ICOMP(3))+ICOMP(2)*(ICOMP(1)*
     1ICOMP(6)+ICOMP(5))
      IF(IBOOLE)13,22,13
   22 K=K-1
      IF(IBRNCH(K)-1)100,2,15
   15 IF(K-1)22,16,22
   16 IF(ICOMP(1))23,101,23
   23 ICOMP(1)= 0
      IBRNCH(1)=2
      DO 30  IN= 2,N
   30 IBRNCH(IN)=0
      GO TO 5
  101 IF(ICODE)102,100,102
  102 WRITE(6,610)PROB
  610 FORMAT(1H ///42X,18H POINT ESTIMATE = , D18.12///)
      GO TO 100
  502 FORMAT(4D18.12)
      END
```

PROGRAM OUTPUT

```
COMPONENT A    IS REPRESENTED AS VARIABLE NUMBER  1
COMPONENT B    IS REPRESENTED AS VARIABLE NUMBER  2
COMPONENT C1   IS REPRESENTED AS VARIABLE NUMBER  3
COMPONENT D2   IS REPRESENTED AS VARIABLE NUMBER  4
COMPONENT D1   IS REPRESENTED AS VARIABLE NUMBER  5
COMPONENT C2   IS REPRESENTED AS VARIABLE NUMBER  6
```

NUMBER OF COMPONENTS = 6

```
0.9938010000000D 00   0.9907970000000D 00   0.9957060000000D 00   0.9938010000000D 00

0.1000000000000D 01   0.9938010000000D 00
```

RELIABILITY MODEL

```
+   A . B . C1
+   A . B   .   D2
-   A . B . C1 . D2
+   A . B   .         D1
-   A . B   .   D2 . D1
-   A . B . C1  .     D1
+   A . B . C1 . D2 . D1
+   A . B         .        C2
-   A . B     .       D1 . C2
-   A . B   .   D2    .    C2
+   A . B   .   D2 . D1 . C2
-   A . B . C1       .     C2
+   A . B . C1  .     D1 . C2
+   A . B . C1 . D2    .   C2
-   A . B . C1 . D2 . D1 . C2
+   A   .   C1
-   A . B . C1
+   B   .   C1 . D2 . D1
-   A . B . C1 . D2 . D1
+       B . C1  .     D1
-       B . C1 . D2 . D1
-   A . B . C1  .     D1
+   A . B . C1 . D2 . D1
+       B   .   D2 . D1
-       B . C1 . D2 . D1
-   A . B   .   D2 . D1
+   A . B . C1 . D2 . D1
+       B      .      D1
-       B   .   D2 . D1
-       B . C1  .     D1
+       B . C1 . D2 . D1
-   A . B      .      D1
+   A . B   .   D2 . D1
+   A . B . C1  .     D1
-   A . B . C1 . D2 . D1
```

POINT ESTIMATE = 0.99990367759ID 00

BEST FITTING LINEAR VARIETIES

Robert M. Thrall
Willow Run Laboratories
University of Michigan, Ann Arbor, Michigan

1. INTRODUCTION. We consider a generalization of the classical problem of finding the best fitting linear function for a set of data. The results obtained are stated in the language of eigenvalues and principal components and take a form which is not explicit in the usual textbook treatments of principal components. In 1901 Karl Pearson in his paper "On Lines and Planes of Closest Fit to Systems of Points in Space" (London Philosophical Magazine, Sixth Series, Vol. 2, 1901, pp. 559-572) stated and solved the problem for ordinary three space. The texts M. J. Kendall, A Course in Multivariate Analysis, and T. W. Anderson, Multivariate Statistical Analysis, treat the standard principal component theory and give useful numerical examples. R. Bellman (Introduction to Matrix Analysis, McGraw Hill, New York 1960, pp. 113-115) develops the same topic from a slightly different point of view using the Courant-Fischer min-max Theorem.

2. SOME ALGEBRAIC BACKGROUND. Let $V = V_k$ be the space of column vectors of degree k over the real field. A sequence $C_1, \ldots, C_r$ of vectors in V is said to be orthonormal if

$$C_i^T C_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \qquad (i, j = 1, \ldots, n).$$

(We use the superscript T to denote matrix transposition.) A matrix $C = [C_1 \ldots C_r]$ is said to be orthonormal if its columns constitute an orthonormal sequence or equivalently if $C^T C = I_r$.

A subset W of V is said to be a subspace if it is closed under addition and multiplication by scalars, a subset M of V is said to be a linear manifold if it has the form $M = W + X_o = \{X_o + X \mid X \in W\}$ for some subspace W; i.e., a linear manifold is just the parallel displacement of a vector space. For any matrix C we denote by $L^*(C)$ the set (subspace) of all solutions of the equation $CX = 0$ and denote by $L(C)$ the subspace consisting of all linear combinations of columns of C. For any subspace W there exist matrices A and B such that $W = L^*(A) = L(B)$. If $\dim W = r$ we may assume that A has shape $(k-r)$-by-k and that B has shape k-by-r ; we may also assume that A^T and B are both orthonormal.

A square orthonormal matrix P is said to be orthogonal ; for any non-square orthonormal matrix C there is a second orthonormal matrix D such that $P = [CD]$ is orthogonal and for which $L(C) = L^*(D^T)$.

A k-by-k matrix A is said to be positive semi-definite if $X^T AX \geqq 0$ for all X in V; if also $X^T AX = 0$ implies $X = 0$, A is said to be positive definite. If A is positive semi-definite there exists an orthogonal matrix P such that

$$\Lambda = P^T AP = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{bmatrix}$$

is a diagonal matrix whose diagonal entries satisfy the condition

$$\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_k \geqq 0 .$$

The numbers $\lambda_1, \ldots, \lambda_k$ are called eigenvalues of A ; a positive semi-definite matrix is positive definite if and only if $\lambda_k > 0$. The j-th column P_j of P satisfies the condition $AP_j = \lambda_j P_j$ and is called an eigenvector for A belonging to the eigenvalue $\lambda_j (j = 1, \ldots, k)$.

If A is any k-by-k matrix we define the trace of A by

$$\text{tr } A = a_{11} + a_{22} + \ldots + a_{kk} .$$

If a product BC of two matrices B and C is square then so is CB and

$$\text{tr } BC = \text{tr } CB .$$

A matrix G is said to be a projection if

$$GG^T = G .$$

If C is orthonormal then CC^T is a projection.

If $W = L(C)$ where C is a k-by-r orthonormal matrix the <u>projection</u> of any vector X on W is the vector

$$X_o = \Sigma_{i=1}^{r} (X^T C_i) C_i ;$$

then $X - X_o$ is perpendicular to X_o, $(X - X_o)^T X_o = 0)$ and the squared distance $d(X, W)^2$ from X to W is given by

$$d(X, W)^2 = (X - X_o)^T (X - X_o) = X^T X - X_o^T X_o$$

$$= X^T X - \Sigma_{i=1}^{r} \Sigma_{j=1}^{r} (X^T C_i)(X^T C_j) C_i^T C_j$$

$$= X^T X - \Sigma_{i=1}^{r} (X^T C_i)^2$$

$$= X^T X - X^T CC^T X$$

$$= X^T (I - CC^T) X .$$

Let $[CD]$ be an orthogonal matrix where D has $s = k - r$ columns. Then

$$I = [CD] [CD]^T = CC^T + DD^T$$

so that

$$d(X, W)^2 = X^T DD^T X$$

and, moreover, $W = L^*(D)$.

Next, let $B = [B_1 \ldots B_n]$ be a k-by-n matrix and let

$$d(B,\ W)^2 = \Sigma_{i=1}^{n} d(B_i,\ W)^2$$

$$= \Sigma_{i=1}^{n} B_i^T\ DD^T B_i$$

$$= \mathrm{tr}\ B^T\ DD^T B$$

since

$$(B^TD)(D^TB) = \begin{bmatrix} B_1^TD \\ \vdots \\ B_n^TD \end{bmatrix} \cdot [D^TB_1 \dots D^TB_n] = \begin{array}{cc} & j \\ i & \begin{bmatrix} B_i^TDD^TB_j \end{bmatrix} \end{array} .$$

Now, since tr is a symmetric function, we have

$$d(B,\ W)^2 = \mathrm{tr}\ B^T\ DD^TB = \mathrm{tr}\ D^T\ BB^TD\ .$$

Next, let M be an orthogonal matrix for which

$$M^T(BB^T)M = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} = \Lambda$$

where $\lambda_1 \geqq \lambda_2 \geqq \dots \geqq \lambda_k \geqq 0$.

Then

$$d(B,\ W)^2 = \mathrm{tr}\ (M^TD)^T(M^T\ BB^TM)\ (M^TD)$$

$$= \mathrm{tr}\ D'^T \Lambda D' = d(\Lambda,\ W')^2$$

where $D' = M^TD$ is also orthonormal and $W' = L^*(D'^T)$.

3. <u>THE BEST FITTING LINEAR SPACE</u>. We now state our problem. Given B, find the space W of dim r which minimizes $d(B,W)^2$, the sum of the squared distances from the columns B to W.

We see that

$$\min \{d(B,W)^2 \mid \dim W = r\} = \min\{\operatorname{tr} D^T BB^T D \mid D \text{ orthonormal of rank } s = k - r\}$$

$$= \min\{\operatorname{tr} D'^T \Lambda D' \mid D' \text{ orthonormal of rank } s = k - r\} .$$

We now show that

(1) $$D_0' = \begin{bmatrix} 0 \\ I_s \end{bmatrix} \text{ minimizes } \operatorname{tr} D'^T \Lambda D ,$$

(2) $$\min \{d(B,W)^2 \mid \dim W = r\} = \lambda_{r+1} + \dots + \lambda_k$$

and for the minimizing space W_0 we have

(3) $$W_0 = L^*(D_0^T) = L(C_o)$$

where $M = [C_0 D_0]$ is the partitioning of M into its first r and last s columns.

Thus, we conclude that W_0 is the space spanned by the r eigenvectors with largest eigenvalues.

Since (2) and (3) follow at once from (1), we need only establish (1): Now

$$D'^T \Lambda D' = [F_1^T \dots F_n^T] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \begin{bmatrix} F_1 \\ \vdots \\ F_k \end{bmatrix}$$

$$= \Sigma_{i=1}^k \lambda_i F_i^T F_i$$

where F_i is the i-th row of $D'(i=1,\dots,k)$. Next, let $y_i = F_i^T F_i (i=1,\dots,k)$. Then, since D' is orthonormal, $0 \leqq y_i \leqq 1$. Moreover,

$$\Sigma_{i=1}^k y_i = \operatorname{tr} D'^T D' = \operatorname{tr} D' D'^T = \operatorname{tr} I_s = s .$$

Thus,

$$\min_{D'} \operatorname{tr} D'^T D' \geq \min \{\Sigma \lambda_i y_i \mid 0 \leq y_i \leq 1, y_1 + \\ + \ldots + y_k = s\} .$$

This simple linear programming problem has the solution $y_1 = \ldots = y_r = 0$, $y_{r+1} = \ldots = y_k = 1$ and since these y_i are realized by D'_0, (1) is established.

4. BEST FITTING LINEAR MANIFOLD. Any linear manifold M of dimension r can be written in the form

$$M = L(C) + h\, C_o$$

where $[C_o C]$ is an orthonormal matrix and $W = L(C)$ is the related linear space. To calculate the distance from any vector X to M we first find the unique vector Y in M for which $X - Y$ is orthogonal to M or, equivalently to W. Let X_o be the projection of X on W as defined in Section 2 above. Then $X - X_o$ is orthogonal to W and since C_o is also orthogonal to W the vector $X - X_o - hC_o$ is also orthogonal to W. But $X_o + hC_o$ is in M; hence $Y = X_o + hC_o$. Thus we have

$$\begin{aligned} d(X, M)^2 &= (X - (X_o + hC_o))^T (X - (X_o + hC_o)) \\ &= (X - X_o)^T (X - X_o) - 2h(X - X_o)^T C_o + \\ &\quad + h^2 C_o^T C_o . \end{aligned}$$

Referring to the notations and calculations in Section 2 and using the fact that $[C_o C]$ is orthonormal this can be written as

$$d(X, M)^2 = X^T DD^T X - 2hX^T C_o + h^2 .$$

Now, let $A = [A_1, \ldots, A_n]$ be any k-by-n matrix. Then the sum of the squared distances of the columns of A to M is given by (cf Section 3)

$$d(A,M)^2 = \mathrm{tr}\, A^T DD^T A - 2nkA_o^T C_o + nk^2$$

$$= \mathrm{tr}\, A^T DD^T A + n(h - A_o^T C_o)^2 - n(A_o^T C_o)^2$$

where $nA_o = \Sigma A_i$; i.e. A_o is the mean of the vectors $A_1, \ldots, A_n$.

From this formula it is clear that for any choice of the matrix C, $d(A,M)^2$ is then minimized by taking $h = A_o^T C_o$ and choosing C_o orthogonal to W and so as to maximize $A_o^T C_o$. This is clearly acheived by taking C_o as the unit vector in the direction of the projection of A_o on $L(D)$ (the orthogonal complement of W). Then the projection of A_o on $L(D)$ will be $(A_o^T C_o)C_o = hC_o$ and we conclude that the <u>minimizing linear manifold contains the mean</u> A_o <u>of the columns of</u> A. We use this fact to reduce the best fitting manifold problem to the best fitting vector space problem which we have already solved.

Clearly, for any vector Z,

$$d(X,M)^2 = d(X - Z, M - Z)^2 .$$

In particular for $Z = A_o$ we have

$$d(X,M)^2 = d(X - A_o, W)^2$$

and hence

$$d(A,M)^2 = d(B,W)$$

where

$$B = [(A_1 - A_o) \ldots (A_n - A_o)]$$

is obtained by subtracting A_o from each column of A. Hence, if W is the best fitting linear space for B then $M = W + A_o$ is the best fitting linear manifold for A.

5. SUMMARY. The results of the preceding two sections can be summarized as follows. Let $A_1, \ldots, A_n$ be any n vectors in k-space, let A be the matrix whose columns are these vectors, let A_o be the mean of the n vectors, and let B be the matrix whose columns are $A_1 - A_o, \ldots, A_n - A_o$. Then the best fitting linear space $W_r(A)$ of dimension r for $A_1, \ldots, A_n$ has a basis the eigenvectors corresponding to the r largest eigenvalues of AA^T and the sum of the squared distances of the vectors to this space is the sum of the k - r smallest eigenvalues of AA^T. (In the case of equal eigenvalues the generating eigenvectors must be independent but this is guaranteed if they are selected as columns of an orthonormal matrix as above.)

The best fitting linear manifold $M_r(A)$ of dimension r for these vectors is then $W_r(B) + A_o$ and the sum of the squares of the distance is the sum of the k - r smallest eigenvalues of BB^T.

If one wishes the average squared distance from the vectors to $M_r(A)$ the number above is divided by n. This can be acheived alternatively by using the matrix $G = (1/\sqrt{n})B$. The result is that $M_r(A) = W_r(G) + A_o$ and the average squared distance is the sum of the k - r smallest eigenvalues of the covariance matrix of $GG^T = (1/n)BB^T$. Suppose that $M = [M_1, \ldots, M_h]$ is an orthogonal matrix for which

$$M^T G G^T M = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix}$$

where $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_k \geqq 0$. Then the columns $M_1, \ldots, M_k$ are called the principal components of the distribution $A_1, \ldots, A_n$, and the first r principal components constitute a basis for $W_r(G)$.

What is usually stated in statistical texts is that the first principal component gives the best fitting line; that the second principal component gives the best fitting line orthogonal to the first; and, in general, that the r-th principal component gives the best fitting line orthogonal to the space generated by the first (r - 1) principal components. It is not stated explicitly that the first r principal components give the best fitting space of dimension r.

PLANNING AND ANALYSIS OF NON-EXPERIMENTAL STUDIES*

W. G. Cochran
Harvard University
Cambridge, Massachusetts

1. INTRODUCTION. During the past 20 years a marked increase in statistical studies of human populations has taken place. Several reasons for this can be suggested. Successful applications of operations research during World War II led to an expanded use of this technique in business and marketing after the war. Public opinion polls, which proved interesting and informative as news media, stimulated the growth of agencies equipped to take sample surveys for clients. The provision of increased amounts of money for field research in the social sciences also contributed.

In many of these studies, the objective is primarily descriptive--to get the basic facts about some problem. Examples are the monthly estimates of numbers of employed and unemployed, or a survey undertaken in a city to estimate the amount of delinquency among teenage boys according to some definition of this term.

In other investigations, interest focuses on the study of relationships. For my purposes, I should like to distinguish two classes within this type, although they shade into one another. The first class consists of broad analytical surveys in which a number of variables are being investigated simultaneously by multiple classification or multiple regression, or by setting up models involving systems of equations, as in econometrics. For instance, in a recent study organized by the U. S. Office of Education [1], standard tests were given to school children in grades 1, 3, 6, 9 and 12. By multiple regression methods, estimates were obtained of the contribution made to the child's performance by various characteristics of the school attended, by the home environment and parental attributes, and by the child's aspirations and self-concept.

When these studies are exploratory, the discovery of the relationships that are present suggests the question: Why?, leading the investigator to set up plausible hypotheses about the causal forces at work. In other studies, causal hypotheses may already have been proposed, the purpose of the study being to verify whether the predictions about relationships made from a casual model are consistent with the results.

*This work was facilitated by a grant from the National Science Foundation (GS-341).

My second class of analytical surveys is narrower in scope and more intimately bound up with the idea of cause and effect. The investigator concentrates on a specific presumed causal agent and tries to measure certain aspects of its effects. Examples are the effects of wearing lap seat belts on the amount and types of injury sustained in auto accidents the effects of air pollution on illness associated with the respiratory organs, the effect of a new contraceptive device on the birth rate during the next five years, and, to cite a World War II study, the effect of bombing on the morale of the bombed people.

These studies resemble controlled experiments, because we set out to measure the effects of certain 'treatments'--the causal agents. However, in the 'non-experimental' studies with which I am concerned, the investigator is unable, for practical or ethical reasons, to use the two chief weapons of controlled experimentation. He cannot select the subjects who are to receive the causal agent and the subjects from whom it is to be withheld. If the agent is one that may be present in greater or less amount, as with air pollution or bombing, he has no control over these amounts, but must take them as he finds them.

The design and analysis of controlled experiments has become fairly well categorized and standardized. Most university courses on the subject discuss completely randomized, randomized blocks, and latin square plans (sometimes under different names) and go on to factorial experimentation and to techniques for estimating response surfaces. This standardization brings with it the usual benefit of economy of effort: once learned, the techniques of planning and analysis can be applied, often with only minor variations, in widely different areas of research.

With non-experimental studies much less standardization of this type has occurred. There is less cumulative experience with the various types of study plan. In the principal fields in which these plans are used--sociology, psychology, education, market research, and public health--workers have only recently begun to learn from one another. Statisticians have shown limited interest in the logical structures of the plans.

While non-experimental studies present many issues that merit discussion, this paper will be confined to three topics, as follows.

Some preliminary aspects of planning.
Simple types of study plan.
Techniques for increasing precision and
eliminating bias.

2. PRELIMINARY ASPECTS OF PLANNING. Being unable to apply the causal agent in which he is interested, the investigator in a non-experimental study must first find some locale in which the agent is operating or will operate under conditions suitable for measuring its effects. In this search the following questions must be kept in mind, all of them matters of judgment rather than of black and white.

1. Is the cause operating in sufficient strength? Sometimes, for reasons of convenience or expense, the investigator chooses an environment in which the causal force operates too weakly to allow its effect to be measured in the size of sample that is feasible. For instance, airline pilots might be considered a convenient source from which to study predictors of heart disease, since they receive repeated and thorough medical examinations of which records are kept. On the other hand, one of the criteria by which they are selected is that they are the kind of men who are unlikely to develop heart disease.

2. What other important variables are present whose effects may be confounded with those of the causal variable? How will they be handled? In planning a study of the effects of air pollution, an investigator might look for three residential areas in the same city, one heavily polluted, one moderately, and one relatively free from pollution. But it is likely that the residents of these areas will show a sizeable gradient in socioeconomic levels, which might account for any differences found in respiratory illness. If the investigator confines himself to areas closely similar in socioeconomic level, he may find that the differences in amounts of air pollution are quite small, thus becoming involved in the difficulty mentioned in point 1. Methods for handling confounded variables are discussed later in this paper. If, however, an important variable is too highly correlated with the causal variable, as might be the case in the air pollution example, there may be no way to disentangle their effects.

3. What measurements are to be taken? What is known about the precision and accuracy of the measurements? Many aspects of human life and behavior present formidable problems of measurement; e.g., how does one measure morale? In large studies, the measurement process may be restricted, for reasons of expense, to responses on a printed questionnaire. Substantial biases in measurement can, of course, produce badly misleading results. "Random" errors of measurement of the effects decrease the precision of the results. "Random" errors in measuring the strength of the causal variable (e.g. number of cigarettes smoked per day) will produce an underestimate of the size of the effect. Similarly, "random" errors in measuring a confounded variable decrease the effectiveness of the standard statistical methods for removing the disturbing effects of this variable.

4. If the study is to be made from records already collected by someone else, have the records been checked as to completeness, accuracy, and accessibility? It is always worth considering whether a study can be made from existing records, not only because of cost but because this may be the only way to obtain results in a reasonably short time. Sometimes, investigators construct plans and engage staff for a study on the basis of someone's assurance about the quality of the records that turns out to be greatly over-optimistic, particularly when the records are kept for some legal or administrative purpose but rarely used or examined. A careful pilot survey of the records, designed to reveal any weaknesses for the purpose at hand, is essential before commitments are made.

5. How will the sample size or sizes be determined? In controlled experimentation there are formulas that provide guidance about sample size by calculating the size needed to estimate the effect with a prescribed width of 95% confidence interval, or the size for which some basic test of significance will have a prescribed power. It is advisable to try to use these formulas in non-experimental studies also. However, in order to obtain useful numerical answers from these formulas one must have an estimate of (i) the standard deviation per observation and (ii) the likely size of the effect that is being estimated. In exploratory studies these estimates may be lacking, and the investigator may have to use simply the largest sample size that can be afforded, having speculated that this size is more likely to be too small than too large.

6. If non-response or later melting-away of the sample is anticipated, what are the plans for coping with it? This is a common problem, especially when participation in the study is somewhat of an imposition on the subjects, or when the study extends for several years. Investigators tend to be lax about non-response. The standard call-back or follow-up questionnaire procedures developed in sample surveys are often surprisingly helpful. Sometimes it is feasible to follow people who move within the same metropolitan area even if it is too costly to follow those who leave the area. Sometimes background information about non-respondents is available, or can be obtained by mail, that assists a judgement about the extent to which they bias the conclusions. Speculations about the extent to which non-respondents might bias the results can always be made much more comfortably with a 10% than with a 30% non-response rate.

7. What are the comparisons from which the size of the presumed causal effect will be estimated? Numerous points arise here. In some studies the 'cause present' group is clearly defined, but it is less clear what can be used as a 'cause absent' group for comparable purposes. Often it is important to estimate the causal effect separately in different subgroups of the population (e.g. for people of different ages, for men and women).

The types of adjustment to be made for handling confounded variables are also relevant.

8. Is the environment a 'typical' one from the viewpoint of generalizability of results? Sometimes an ingenious investigator finds a group of people (for instance a special religious sect) among whom the causal force is operating with no important confounded variables. But he may reluctantly decide not to attempt the study in this group, because they seem atypical in so many respects that any generalization of results would appear hazardous.

With some problems of great interest and importance, investigators have to search for a long time before a suitable environment is found. Sometimes none is found: in other cases we are restricted to the type of study that can be done rather than the type we would like to do. Consider the problem of investigating in human subjects the effects of exposure to atomic radiation on illness and death rates. Ideally, the answer would take the form of a dosage-response curve, the rate being expressed as a function of the exposure history (amount and duration).

As pointed out by Seltser and Sartwell [2], the principal opportunities for investigations in human subjects are confined to the following: (a) the Japanese survivors of the atomic bombs in Hiroshima and Nagasaki, involving a single exposure, (b) groups occupationally exposed to radiation at times when the possible danger from this source was not realized--radiologists, dentists, and makers of watches with luminous dials, (c) persons who received medical radiation, as in the treatment of some forms of cancer, or infants exposed <u>in utero</u> through pelvic X-rays of the mother in the late stages of pregnancy, and (d) areas of the earth in which natural radioactivity is unusually high.

None of these sources provides more than limited material for constructing a dosage-response curve. To illustrate the types of study that have been undertaken, long-term studies in Hiroshima and Nagasaki were initiated in 1950. In Hiroshima the sample contains about 12,000 people, divided into 4 groups of about 3,000 each, according to their distances from the point of impact of the bomb. The subjects receive regular health examinations, with particular attention to any symptom that might be an after-effect of radiation exposure.

A study of this type is expensive and administratively difficult. Fortunately, the health data also permit many useful investigations of general health questions. From the viewpoint of the dosage-response curve, a weakness is that the dose to which any person was exposed is not known, but has had to be estimated roughly from memory of a person's location and local shielding by buildings at the time when the bomb fell.

Also, the group furthest from the epicenter, who serve as the non-exposed group, differ in some important characteristics from the three exposed groups, and have proved unsatisfactory as a 'control' [3].

The study by Seltser and Sartwell [2] of the mortality of radiologists is an excellent example of the possibilities from groups occupationally or medically exposed. They chose male members of the Radiological Society of North America. For each member they obtained by a painstaking search the status (dead or alive) as of December 31, 1958, with cause of death and any available information on other factors such as age that might influence duration of life. Research of this type always raises the question: with what are the exposed group to be compared? Ideally, we seek a non-exposed group which is similar to the exposed group with regard to any other variable that is known or suspected to have a material effect on duration of life. (In this example an obviously relevant variable is age.) In an observational study the extent to which this goal can be met is of course dependent on our ability to measure such variables and to find a group that has similar distributions with respect to them.

The authors chose two comparison groups. As the nearest to a non-exposed group they used the American Academy of Ophthalmology and Otolaryngology, whose members rarely have occasion to employ X-radiation. As an intermediate group they also included the American College of Physicians, since some of these members use X-rays, for example, in heart examinations. In such studies the inclusion of a middle group is advantageous in either adding confirmation to the results given by the two extreme groups or in casting doubt upon them. This study, however, again has the weakness that no measures of the doses of radiation experienced by the subjects are available, except as a rough guess for the group as a whole.

3. SIMPLE TYPES OF STUDY PLAN. This section introduces some simpler types of plan, with a brief discussion of their strengths and weaknesses and of the statistical analysis.

3.1 A single group, measured before and after the action of the causal agent. This type is common when the causal agent is of short duration. For example, after complaints about the time taken to go through a cafeteria line, a change in the service is proposed that it is claimed will remove the bottleneck. Before this change is made, the times taken to go through are recorded for a random sample of the users, the same being done after the change is made. In other situations, the causal agent might be DDT spraying of 10 villages, an estimate of the misquito population being made before and after spraying, or a radio and TV appeal which the stations in an area agree to give on a certain day, urging mothers to bring their children into the clinics in a city for immunizations, the number of children appearing for immunization being counted in each clinic during the week before and the week after this appeal.

Unlike the radiologists example, such studies have no comparison group, usually because all members of the population of interest are exposed (at least potentially) to the causal agent. Sometimes, as in the DDT example, a comparison group of unsprayed villages might have been chosen, but is excluded for administrative or financial reasons. Often, a single-group study is the only feasible approach in attempting to learn something about the effects of new governmental programs or laws that apply to everyone.

The absence of a comparison group is, of course, the major weakness. Any other event that produces a change in the level of the variable during the Before-After period has its effects inevitably confounded with those of the causal agent. Campbell and Stanley [4] give a detailed catalogue of these sources of bias in educational research. If the investigator is aware of such other influences he can sometimes ask questions about the reasons for people's change in behavior that help him to judge whether these influences have been important. Knowledge that a change is coming may influence people's behavior immediately before the change, so that the After-Before difference is misleading.

Although the conclusions from studies of this type involve a substantial element of judgment, the studies are, as Campbell and Stanley put it, "worth doing when nothing better can be done". I might express it a little more positively. With new public programs, plans to estimate their effects are often not initiated until some time after the program has been running. By this time it is difficult to get good 'Before' measurements and too late to take precautions or gather supplementary information that might have helped in judging the effects. The question: How can we study the effects of this program? should be raised some time before the program begins.

The statistical analysis usually involves examining the difference between two paired or independent samples. The samples may be subclassified by another variable, e.g., age of subject, in order to reveal any variation in effect with age.

Sometimes there is reason to expect that the Before measurement will itself influence the subject's behavior. A plan that has been proposed is to have two groups, both exposed to the cause. Whenever feasible, these can be random halves of an initially chosen group. Group 1 is measured 'Before' and 'After', group 2 is measured 'After' only. The idea is that by comparing the two 'After' sets of results, we can test whether the 'Before' measurement influenced the level of the 'After' responses in group 1.

The best method of estimating the size of the causal effect presents a problem involving the pooling of data after performing a test of significance. If the subscripts a and b denote 'After' and 'Before', the difference $(\bar{Y}_{2a}-\bar{Y}_{1b})$ is an unbiased estimate of the causal effect. Assuming a constant variance σ^2 per subject, this difference has variance $2\sigma^2/n$. The difference $(\bar{Y}_{1a}-\bar{Y}_{1b})$ has variance $2\sigma^2(1-\rho)/n$, where ρ is the correlation between the 'Before' and 'After' measurements for the same subject, but is unbiased only if the 'Before' measurement did not affect the level of the 'After' measurement. The estimates $(\bar{Y}_{2a}-\bar{Y}_{1b})$ and $(\bar{Y}_{1a}-\bar{Y}_{1b})$ are themselves correlated, since $\bar{Y}_{1b}$ appears in both. One approach is to seek a weighted mean of these estimates, with weights determined from the results of the preliminary test of significance of $(\bar{Y}_{1a}-\bar{Y}_{2a})$, that has minimum mean square error subject to a condition that the bias be kept small.

The preceding discussion has been confined to studies in which it is satisfactory to measure the causal effect at a single time after the causal event. In many situations, the causal event may have prolonged effects, or if its effect is likely to die away, the investigator wants to measure this decay curve. For these purposes we need, at a minimum, a series of measurements at intervals of time before the event, followed by a series at intervals after the event. The problem of the model to be used for the analysis of results of this type raises some interesting questions which have been illustrated by Campbell and Stanley [4]. Model-fitting and interpretation are easiest when the 'Before' measurements appear to fluctuate about a constant level; the difficulty increases when the 'Before' and 'After' measurements display trends, particularly those with curvature. The question of serial correlations must also be considered.

3.2 <u>'Cause present' and 'Cause absent' groups. Y measured 'After' only.</u> This is a very common type. The Hiroshima and radiologist studies, investigations of the effectiveness of seat belts in preventing injury in automobile accidents, and the large studies of the death rates of non-smokers and cigarette, cigar, and pipe smokers are examples. As we have seen, there may be several 'cause present' groups, representing different strengths or variations inthe causal agent, and more than one 'cause absent' group, particularly where the selection of a control group presents difficulty.

At its simplest, the analysis follow the usual methods for the analysis of one-way classifications or of two-way classifications if pairing or blocking has been employed in forming the groups. Often, however, the analysis of a multiple classification is involved, other variables being introduced in order to diminish the risk of bias, as discussed in section 4, or because the investigator wants to examine interactions of the causal effects with these variables.

An important variant of this method, often called the retrospective method, is much used in epidemiological research. In this, we find a group in which the effect is present and one from which it is absent, and compare the frequency with which the presumed causal agent is found in the two groups. This approach is natural when a group of people show symptoms of food poisoning at a picnic and the cause is being sought. As another example, numerous investigators have selected a group of lung cancer patients and another group of patients in the same hospitals who do not have this disease, comparing the proportions of cigarette smokers in the two groups. With this approach, it is often hard to select the 'effect absent' group and to obtain measurements of high quality. Further, erroneous results may be obtained when there are several causal agents and attention is focussed on one. But with an effect that is rare, this approach may be the only practicable one, and it is often the quickest way of obtaining a preliminary indication for or against a postulated relationship. For a discussion, see [5].

3.3 'Cause present' and 'cause absent' groups. Y measured Before and After. This plan has been used, for example, in studies of the effects of new public housing, as against slum housing, on health and social behavior. When it became known which group of applicants were to move into a new public housing development, a control group of families who would in general remain in slum housing were selected. The basic questionnaries on health and social behavior were obtained both before the move took place and at several times after the successful applicants had moved. In a study of the effects of fluoridation of town water on children's teeth, usually done by a plan of type 3.1, a nearby control town which did not plan to fluoridate could be included if the resources permitted. The state of dental health of a sample of children from both towns would be measured before and some time after the fluoridation in the first town.

With this plan the investigator is in a better position to guard against bias than with plan 3.2. Ideally, the initial distribution of the response variable Y should be the same in the 'cause present' and 'cause absent' groups. Since he has the initial measurements, he can verify whether this seems to be the case. Even if the distributions are somewhat different, it is still possible to compare the amount of change in the two groups during the 'Before-After' period.

A general estimate of the size of the causal effect is

$$(\overline{Y}_{1a}-\overline{Y}_{2a}) - \beta(\overline{Y}_{1b}-\overline{Y}_{2b}) , \tag{3.1}$$

where the value of β is to be chosen. Suppose that the model is as follows.

$$\text{Before:} \quad y_{1bj} = \mu_1 + e_{1bj}; \qquad y_{2bj} = \mu_2 + e_{2bj}$$

$$\text{After:} \quad y_{1aj} = \mu_1 + \delta + \tau_1 + e_{1aj}; \quad y_{2aj} = \mu_2 + \tau_2 + e_{2aj} \quad .$$

Here, δ represents the causal effect to be estimated; τ_1 and τ_2 represent other time-changes that affect the two groups; and the e's are random variables with means zero. From this model we see that

$$E\{ (\overline{Y}_{1a}-\overline{Y}_{2a}) - \beta(\overline{Y}_{1b}-\overline{Y}_{2b}) \} = \delta + (\tau_1-\tau_2) + (\mu_1-\mu_2)(1-\beta) \quad .$$

Hence, (i) if $\tau_1 \neq \tau_2$, the plan provides no unbiased estimate of δ: this is, of course, obvious, (ii) if $\tau_1 = \tau_2$ but $\mu_1 \neq \mu_2$ (i.e., the initial levels of the two groups differ), the only unbiased estimate of δ is given by taking $\beta = 1$. (iii) if $\tau_1 = \tau_2$ and $\mu_1 = \mu_2$, any value of β gives an unbiased estimate. Assuming that the e's all have the same variance σ^2, the estimate (3.1) has variance

$$2\sigma^2 (1-2\beta\rho+\beta^2)/n$$

where ρ is the correlation coefficient between an 'After' and a 'Before' measurement. If the 'After' and 'Before' samples are independent, so that $\rho = 0$, we take $\beta = 0$. If these measurements are paired, the minimum variance is given by $\beta = \rho$. In practice, β is estimated in this case by an analysis of covariance of the 'After' on the 'Before' measurements.

4. TECHNIQUES FOR INCREASING PRECISION AND ELIMINATING BIAS. In controlled experiments the investigator relies on randomization, plus other precautions such as 'blindness' in the measurement process, to ensure that biases are kept to a negligible level. As means of increasing precision, blocking and adjustments made by the analysis of covariance are two of the principal weapons.

Devices analogous to blocking and covariance are commonly used in non-experimental studies also. However, since randomization is not available, these devices must perform the double function of eliminating bias and of increasing precision. In fact, since bias is regarded as the chief source of erroneous conclusions, control of bias becomes their principal function.

Suppose that Y is the response or effect variable, and that there is a 'cause present' and a 'cause absent' group. If X is any variable that is related to Y, a bias may arise in $(\overline{Y}_1 - \overline{Y}_2)$, the estimated difference between the means of the two groups, if the distribution of X differs in the two groups. For instance, if the regression of Y on X is linear,

$$Y_{ij} = \mu_i + \beta X_{ij} + e_{ij},$$

where i = 1, 2 denotes the group, and the e_{ij} are residuals with mean zero, then

$$E(\overline{Y}_1 - \overline{Y}_2 | X) = \mu_1 - \mu_2 + \beta(\overline{X}_1 - \overline{X}_2) \quad . \tag{4.1}$$

The term $\beta(\overline{X}_1 - \overline{X}_2)$ is the bias.

In handling these variables the investigator makes a list of the X variables known or thought to be related to Y. These variables are placed in one of the following classes.

(I) Important variables whose effects the investigator will try to remove, either because there seems a danger of bias or because removal will bring a worthwhile increase in precision.

(II) Variables for which the investigator will check whether their distribution is similar in the 'cause present' and 'cause absent' groups. No adjustment will be made for these variables unless the distributions appear sufficiently different so that there seems a danger of bias. This method is employed for variables whose correlation with Y is modest. If Y and X are linearly related, with correlation ρ, the fractional reduction in the variance of Y due to elimination of the effect of X cannot exceed ρ^2. If $|\rho| \leq 0.3$, this reduction is less than 9%: the potential increase in precision is small.

In practice, verification that the distribution of X is similar in the two groups of subjects is often done by forming the frequency distribution of X in each group, with, say, k classes, and making the χ^2 test for a 2 x k contingency table. A low value of χ^2 is taken as assurance that the distributions of X are similar and that there is little risk of bias from the relation between Y and X. This χ^2 test may not be the best procedure. If the regression of Y on X is linear, equation (4.1) shows that comparison of the mean values of X in the two groups is more relevant, since the bias in $(\overline{Y}_1 - \overline{Y}_2)$ comes from the term $(\overline{X}_1 - \overline{X}_2)$. Similarly, if the relation between Y and X is curved and can be approximated by a quadratic regression, comparison of the first two moments of X in the two groups is relevant.

(III) Variables about which nothing will be done, because their relation to Y is judged too tenuous to create trouble. This class also contains X variables which it is not feasible to measure and those of which the investigator is ignorant.

A natural question at this point is: Why not put all the X variables in class I, or at least do so whenever there is any doubt? I don't know the full answer to this, but a partial answer is that the techniques (matching and adjustment) by which we attempt to remove the effects of these X variables become steadily more cumbersome to apply and to interpret as the number of X variables increases. These techniques may be described as follows.

'Ideal' matching. Each member of group 1 has a partner in group 2 who has, within narrow limits, the same value for any X variable for which adjustment is being made. By taking the difference between partners, the effects of these X variables are eliminated, provided that the regression of Y on these X variables is the same in both groups. Clearly, this matching is effective whether the regression is linear or curved.

In practice, the construction of matched pairs often presents difficulty, particularly if matching has to be done on several X variables. Usually, it is necessary to have a large reservoir of subjects for at least one of the two groups; otherwise, it will not be possible to locate partners who agree closely on the values of all the desired X variables. A common experience is that the construction of partners takes much longer than anticipated, that the rules set up about the closeness of the match have to be continually relaxed, and that some subjects have to be omitted because no match is found.

Stratified or frequency matching. This is a looser form of matching which facilitates the construction of partners. The range of each X variable is divided into a number of classes, commonly from 2 to 5 or 6. Thus the X variables create a multiple classification: for instance, with 3 X's and 4 classes per variable there are 64 cells. For a member of the 'cause present' group, any member of the 'cause absent' group who falls in the same cell is an acceptable partner. In the end, what this method amounts to is that in any cell of the multiple classification the two groups have an equal number of subjects. Often, there is no specific designation of partners, since this seems rather pointless.

Stratified matching is the only kind that is feasible for an X that is an ordered classification, such as "mild", "moderate", "severe" or is qualitative, e.g., religious affiliation or urban, suburban, rural.

Adjustment by subclassification. This method is very similar to stratified matching. When selecting the 'cause present' and the 'cause absent' groups we do not attempt any matching. Adjustment for differences in the X distributions in the two groups is accomplished by forming the multiple classification used in stratified matching and making adjustments by a least squares or analysis of variance model.

To illustrate the relation between the two methods, suppose there are X variables and that only 100 subjects are available for the 'cause present' group. To see how the land lies, we classify these subjects, plus 100 from the 'cause absent' group, into 9 cells, assuming that each X variable has 3 cells. In table 1, the numbers of subjects found in each cell are shown, P and A denoting the two groups, Both the P and A sets add to 100.

TABLE 1

Subclassification on two X variables.

		X_1		
		< 20	21 - 50	Over 50
	Mild	P 8	P 10	P 19
		A 23	A 7	A 4
X_2	Moderate	P 8	P 8	P 16
		A 26	A 9	A 3
	Severe	P 5	P 11	P 15
		A 19	A 6	A 3

If we are using stratified matching, we select 8 at random out of the 23 A's in the top left cell, discarding the rest. In both the other cells in the top row, we need more A's to reach the desired numbers 10 and 19. Looking the table over, it appears that a reservoir of perhaps 700 or more subjects suitable for the 'cause absent' group would be necessary to build up all the cells to the desired numbers in the P group.

In adjustment by subclassification, as I am using this term, we either accept the A sample as it stands or attempt only to build up cells in which the A sample is very small. The decision depends on the size of the reservoir for the $\underline{A}$ group, the time and trouble involved in any build up, and the investigator's opinion as to whether the effort is worthwhile.

From the viewpoint of estimation of effects we face a 2 x 3 x 3 table with either stratified matching or adjustment by subclassification. It is assumed that Columns (X_1) and Rows (X_2) both show real effects, and possibly an interaction, since otherwise there would be no need to match or adjust for these X variables.

The simplest situation is that in which there is no interaction of the (P-A) difference with either X_1 or X_2. In this event the 9 differences $(\overline{P}_{ij}-\overline{A}_{ij})$ are all estimates of the same quantity. It follows that with stratified matching, the difference between the overall sample means $(\overline{P}-\overline{A})$ is free from any confounding with the levels of X_1 or X_2. The estimate $(\overline{P}-\overline{A})$ has variance $\sigma^2/50$, where σ^2 is the within-cell variance (assumed constant from cell to cell). If the A sample is accepted as it stands, the corresponding estimate for adjustment by subclassification is a weighted mean of the differences $(\overline{P}_{ij}-\overline{A}_{ij})$, weighting each inversely as its variance. The weights are $n_{1ij}n_{2ij}/(n_{1ij}+n_{2ij})$, where the $\underline{n}$'s are the sample sizes in the (i, j) cell. For table 1 the variance of this weighted mean difference turns out to be $\sigma^2/36.6$, about 35% larger than with stratified matching. In this situation stratified matching provides a simpler estimate that is more precise.

We may, however, wish to examine whether the (P-A) difference changes with the level of X_1 and X_2. As Billewicz [6] has pointed out, the ability to examine these interactions is an advantage which these methods hold over 'Ideal' matching. If interactions are found, estimation of the overall difference may become of little interest. The technique needed here is the analysis of multiple classifications with unequal numbers in the cells. While the general least squares theory covering this technique is not new, much remains to be learned about the practical handling and interpretation of such analyses, particularly for investigators who are not expert in statistical methods. The recent paper by Federer and Zelen [7] is a useful contribution.

Adjustment by covariance. Conceptually, this is the same approach as adjustment by subclassification for the case in which the $\underline{X}$'s are continuous. Covariance may have an advantage and a disadvantage. The grouping of continuous X's into classes in adjustment by subclassification loses some information: covariance avoids this loss. On the other hand, adjustment by subclassification does not involve any assumption that the relation between Y and X is linear. If the investigator follows the common practice of adjusting in covariance only for linear effects of X, covariance is at a disadvantage if the true regression has substantial non-linearity. Of course, this loss can be avoided by adopting a more accurate model in the covariance analysis.

How effective are these techniques? The following comments are based on results quoted in [8] and on some unpublished work. As already mentioned, 'ideal' matching removes bias due to $X_1 \ldots X_k$ under any regression

$$Y_{ij} = \mu_i + \phi(X_{1ij}, \ldots X_{kij}) + e_{ij}, \quad (i = 1, 2)$$

if the regression function ϕ is the same in both groups. The variance of $(\overline{Y}_1 - \overline{Y}_2)$ is reduced by the matching to a fraction $(1-\rho^2)$ of its original value, where ρ is the correlation coefficient between Y and ϕ. In practice, 'ideal' matching is likely to be at its best when the X's are quantitative and one of the groups has a large reservoir in which matches may be sought, while the other group is small. In this situation, matching should not prove too difficult. Moreover, the other disadvantage of matching--that one cannot examine effectively the interactions of the causal variable with the X variables--scarcely applies when one group is small, since the sample size would probably preclude any precise estimates of interactions.

Covariance adjustment should have about the same effects on bias and precision, with the qualifications that the correct form of the regression equation must be fitted, and that there is some loss of precision from sampling errors in the estimated regression coefficients. If the regression is linear and there happens to be no bias due to the X's, the fraction to which $V(\overline{Y}_1 - \overline{Y}_2)$ is reduced by the covariance adjustment is roughly

$$(1-\rho^2)\left\{1 + \frac{k}{(2n-k-3)}\right\} \tag{4.2}$$

where n is the size of sample in each group, so that the regression coefficients are estimated from $2(n-1)$ degrees of freedom. The term in curly brackets will be close to 1 if k is small relative to 2n. However, if there are substantial biases in some of the X's, (4.2) no longer applies, and the corresponding term in curly brackets can be much larger. The performance of this covariance adjustment when the fitted model is of the wrong form deserves further study. Linear covariance adjustments seem to perform surprisingly well when the true regression has a moderate degree of curvature.

The preceding remarks about matching and covariance assume that the X's are measured without appreciable error. Suppose that for an X variable the recorded measurement is $x = X + d$, where d is a random error of measurement with mean zero, independent of X and of e, the deviation of Y from its regression on X. The effects of these errors of measurement are roughly as follows, where $f = \sigma_d^2/\sigma_x^2 = \sigma_d^2/(\sigma_X^2 + \sigma_d^2)$.

(i) Matching and covariance remove only a fraction $(1-f)$ of the bias in Y due to X.

(ii) $V(\bar{Y}_1 - \bar{Y}_2)$ is reduced to the fraction $\{1 - (1-f)\rho^2\}$ of its original value.

While imprecise measurement weakens the performance of these techniques, it is easy to form an exaggerated notion of the size of this effect if some check calculations are not made. For instance, suppose that $\sigma_X = 25$, nearly all the correct values of X lying between 0 and 125. If we are told that half the observed measurements are wrong by more than 5 units, this seems rather poor quality of measurement. However, a probable error of 5 corresponds to $\sigma_d = 7.4$, $\sigma_d^2 = 55$, $\sigma_x^2 = 680$, and $f = 0.08$. Thus, 92% of the bias is still removed.

Now consider stratified matching and adjustment by subclassification as applied to quantitative X's. From the viewpoint of errors of measurement of X, these methods appear crude, since the quantitative scale of an X variable is replaced by a classified variable that takes only the number of distinct values that the number of classes allow. With stratified matching the values of $(1-f)$ are 0.64, 0.79, 0.86, 0.90, and 0.92 for 2, 3, 4, 5, and 6 classes, respectively. Strictly, these values hold only if the regression of Y on X is linear, X is normally distributed, and the classes are of equal size. However, they appear accurate enough as guides to practice when the regression of Y on X is nonlinear, when X has some skewness and kurtosis and when the class sizes depart moderately from equality. The results indicate that at least five or six classes should be used for any X variable which is thought to be a source of substantial bias.

With adjustment by subclassification the preceding $(1-f)$ values apply so far as the removal of bias due to X is concerned. This method suffers an additional loss of precision, as illustrated previously, because of inequalities in the sample sizes of the two groups in the individual cells of the multiple classification.

The situation when $\underline{X}$ is an ordered classification is not so clear. If an ordered classification can be regarded as essentially a grouping of an underlying quantitative $\underline{X}$, the preceding values of $(1-f)$ should be applicable. In practice, however, ordered classifications are often used because no more precise method of measurement is known. If we envisage some accurate measurement X, not yet discovered, it seems reasonable that the ordered classification will contain errors of misclassification as well as grouping errors. These additional errors presumably reduce the values of $(1-f)$, to an extent that does not seem to have been investigated.

Finally, none of the methods can guarantee to remove bias due to an $\underline{X}$ variable that has been omitted from the matching or adjustments. The situation with regard to such omitted variables is interesting. If they happen to have a high correlation with the included $\underline{X}$'s--in other words, if we are lucky--most of their bias will also be removed by the matching or adjustments. This explains, I think, why linear covariance often works well when Y has a quadratic regression on X, since X and X^2 have a high correlation in many bodies of data. But one can also meet the opposite situation in which the bias due to omitted X's is inflated by the adjustments. Thus in non-experimental studies there always remains an element of uncertainty in our claims about the size and reality of a presumed causal effect.

REFERENCES

1. J. S. Coleman et al. Equality of Educational Opportunity. U. S. Office of Education. 1966.

2. R. Seltser and P. E. Sartwell. The influence of occupational exposure to radiation on the mortality of American Radiologists and other medical specialists. Amer. J. Epidemiology, 81, 2-22, 1965.

3. S. Jablon, M. Ishida, and G. W. Beebe. Studies of the Mortality of A-bomb survivors. 2. Radiation Research 21, 423-445, 1964.

4. D. T. Campbell and J. C. Stanley. Experimental and Quasi-Experimental Designs for Research. Rand McNally & Co., Chicago, 1966.

5. B. MacMahon, T. F. Pugh, and J. Ipsen. Epidemiologic Methods. Little, Brown & Co., Boston, 1960.

6. W. Z. Billewicz. The efficiency of matched samples: an empirical investigation. Biometrics, 21, 623-644, 1965.

7. W. T. Federer and M. Zelen. Analysis of multifactor classifications with unequal numbers of observations. Biometrics, 22, 525-551, 1966.

8. W. G. Cochran. The planning of observational studies of human populations. J. Roy. Statist. Soc., A, 128, 234-365, 1965.

A MODERATELY DISTRIBUTION FREE APPROACH TO RELIABILITY ESTIMATION BASED ON THE FIRST ORDER STATISTIC*

Michael G. Billings
C-E-I-R, Inc., Dugway Field Operation
Dugway, Utah

ABSTRACT. This paper describes a small sample reliability test design and evaluation technique based on properties of the first order statistic. The technique is "moderately distribution free" in that it is applicable to any problem which satisfies the following conditions: 1) The random variable X involved is continuous; 2) X can take on only non-negative real values; 3) the "mission" of the system under investigation is a set of real numbers of the form $[T, \infty)$, where $T \geq 0$; 4) there exists a set of real numbers $M \subset [1, \infty)$ such that $m \in M \Rightarrow F_X(mT) \geq mF_X(T)$, where F_X is the distribution function of X. Some sufficient conditions are given which define classes of distributions to which the technique is applicable. Also, it is shown that the technique is a highly accurate approximate procedure for reliability evaluation when in fact the random variable X involved has an exponential distribution, so that Condition 4 is not satisfied. Finally, a brief consideration of the Weibull distribution is presented.

1. INTRODUCTION. The purpose of this paper is to derive and demonstrate a small sample reliability test design and evaluation technique which appears to have applicability over a wide class of distributional forms. The technique derived, referred to as the Modified Distribution Free (MDF) technique, is based upon certain properties of order statistics and is conceptually similar to the strictly distribution free binomial approach to reliability evaluation. The MDF technique introduces certain fairly nonrestrictive assumptions in order to achieve a trade off between sample size and system performance. Before proceeding it will be useful to introduce and interpret the concepts and symbols which will be encountered in the ensuing discussion.

Technically, the term reliability is always used relative to some system, conceptual or real, the primary purpose of which has been determined to be the accomplishment of a specific objective called the system mission. The reliability of the system is defined to be the probability that

*This article prepared for U. S. Army Test and Evaluation Command under Contract No. DA-42-007-AMC-141(R).

the system will accomplish its designated mission. In order to meaningfully discuss system mission reliability it is necessary to establish a method for measuring system performance relative to the particular mission. For this purpose it is convenient to assume the existence of a population $\mathcal{S}$ of systems of the type under consideration. On this population a random variable X is defined in such a way that the mission can be characterized as a subset T* of the probability space $\mathcal{S}_X$ induced from $\mathcal{S}$ by X. If the probability measure on $\mathcal{S}_X$ is P and the associated distribution function of X on $\mathcal{S}_X$ is F_X, then the definition of system mission reliability becomes

$$P\{x \in \mathcal{S}_X \mid x \in T^*\} = Pr(X \in T^*) = \int_{T^*} dF_X . \text{1/}$$

In practice it is desired to obtain an interval estimate of the system reliability to which we are able to attach a measure of assurance that the interval contains the true reliability. Conventionally, this has been done as follows:

1) One obtains for $\int_{T^*} dF_X$ a confidence interval estimator which depends on an estimator $\hat{F}$ of F_X;

2) A value of $\hat{F}$ is then observed, and the corresponding confidence interval estimate for $\int_{T^*} dF_X$ is calculated;

3) The confidence coefficient associated with the interval estimator for $\int_{T^*} dF_X$ is taken to be the measure of assurance (confidence) that the calculated interval estimate contains the true value of $\int_{T^*} dF_X$.

The result of this procedure is a statement of the form "with γ-confidence the reliability is at least α", hereafter abbreviated $r(\alpha, \gamma)$, where α is the lower bound of the interval estimate obtained for $\int_{T^*} dF_X$, and γ is the associated confidence coefficient.

1/ In accordance with convention, if X is a continuous random variable with density function $f_X = (dF_X)/(dx)$, then $\int_{T^*} dF_X = \int_{T^*} f_X dx$. If X is discrete, then the integral $\int_{T^*} dF_X$ is a sum over the set T*.

Under certain assumptions on the random variable involved, it is possible to equivalently formulate the reliability evaluation problem within a hypothesis test framework. The Modified Distribution Free technique described in Section 2 utilizes this approach.

2. THE MODIFIED DISTRIBUTION FREE (MDF) TECHNIQUE.

The MDF approach to reliability estimation presupposes that the following conditions are satisfied by the particular problem involved:

1) The random variable X under consideration is continuous.

2) $\mathcal{S}_X = \{x \mid x \geq 0\}$.

3) The mission T^* can be described by $T^* = \{x \in \mathcal{S}_X \mid x \geq T,\ T \in \mathcal{S}_X\}$.

4) There exists a set of real numbers $M \subset [1, \infty)$ such that

$$m \in M \implies F(mT) \geq mF(T),$$

where T is given in Condition 3, and $F(x) = Pr(X \leq x)$.

The particular hypothesis test structure employed in the MDF approach is described as follows: Suppose it can be assumed that $m \in M$ (see Condition 4 above), and that it is desired to either conclude or fail to conclude the reliability statement $r(\alpha, \gamma)$ on the basis of a sample of size n from $\mathcal{S}_X$. Let β be determined so that $(1-\beta)/m = 1-\alpha$, and let the sample size n be such that $1-\gamma = \beta^n$. Let the null hypothesis be given by

$$H_0: F(mT) > 1-\beta;$$

and the alternate hypothesis be given by

$$H_1: F(mT) \leq 1-\beta .$$

The test statistic to be used is $X_{(1)}$, the first order statistic, and H_0 will be rejected if $X_{(1)} \geq mT$.[2]

It is clear that $Pr(\text{Reject } H_0 \mid H_0 \text{ is true}) \leq 1-\gamma$, since

[2] Note that this means testing can be truncated once each sample system has operated for mT units.

$$\Pr(X_{(1)} \geq mT \,|\, H_0 \text{ is true}) = [1 - F(mT)]^n \leq \beta^n$$
$$= 1-\gamma \ .$$

Thus, if H_0 is rejected, it is concluded that $F(mT) \leq 1-\beta$, and the significance level of the test does not exceed $1-\gamma$. By virtue of the initial assumption that $F(mT) \geq mF(T)$, rejection of H_0 implies that $F(T) \leq (1-\beta)/m = 1-\alpha$. The probability that H_1 is accepted erroneously does not exceed $1-\gamma$; thus, if H_0 is rejected the conclusion is $r(\alpha,\gamma)$.

The usefulness of the MDF approach as a design tool when the appropriate conditions and assumptions are satisfied is evident from Proposition 1:

Proposition 1. Suppose for a given reliability estimation problem that the Conditions 1-4 above are met. Suppose further that $m \epsilon M$ (i.e., $F(mT) \geq mF(T)$). If

$$n = \frac{\log(1-\gamma)}{\log(1-m(1-\alpha))},$$

and if the null hypothesis H_0 of the MDF hypothesis test is rejected (i.e., if $X_{(1)} \geq mT$), then $r(\alpha,\gamma)$.

Proof: This follows immediately from the relations $(1-\beta)/m = 1-\alpha$; and $\beta^n = 1-\gamma$.

Making use of this result, it is possible to construct tables which give sample sizes from which rejection of H_0 will lead to conclusion $r(\alpha,\gamma)$ for various values of m and reliability-confidence level combinations (α,γ). An abbreviated table is presented below. The entries are the minimum sample sizes necessary for rejection of H_0, i.e., when $X_{(1)} \geq mT$, to lead to the conclusion $r(\alpha,\gamma)$ under the assumption that $F(mT) \geq mF(T)$ for the value of m shown.

TABLE 2.1

	(α, γ)		
m	(.99, .90)	(.95, .90)	(.90, .90)
1	230	45	22
2	114	22	11
2.5	91	18	9
3	76	15	7
3.75	61	12	5
4	57	11	5
5	45	8	4
6	38	7	3

Example. If the mission is T = 50 hours, and if it can be assumed that $F(4T) \geq 4F(t)$, then to be able to conclude $r(.95, .90)$ a minimum of 11 systems would have to operate successfully for at least $4 \cdot T = 200$ hours. (See Footnote 2.)

The MDF-hypothesis test can also be used to provide descriptive reliability statements, i.e., statements of the form $r(\alpha, \gamma)$ based on the actually observed value of $X_{(1)}$. Suppose, for example, that for a sample of n systems we observe $X_{(1)} = m^*T$, and it can be assumed that $m^* \epsilon M$. It is then possible to determine the strongest statement $r(\alpha, \gamma)$, which can be made on the basis of the test, for a fixed upper bound $1-\gamma$ on the significance level by solving the equation

$$n = \frac{\log(1-\gamma)}{\log(1-m(1-\alpha))}$$

for α. It is also possible to determine the highest confidence which can be associated with a given reliability level α on the basis of the test.

Example. Suppose T = 50 hours and a sample of 17 systems yields a value of $X_{(1)} = 178$ hours. Further, suppose it is possible to assume that $178/50 \epsilon M$. Then the strongest statement of the form

$r(\alpha, .90)$ which can be concluded on the basis of the MDF-hypothesis is $r(.9645, .90)$, obtained by solving the equation

$$17 = \frac{\log(.10)}{\log(1-3.56(1-\alpha))}$$

for α.

Example. On the basis of the performance described in the previous example, the statement $r(.95, .964)$ could also be concluded.

3. APPLICABILITY OF THE MDF APPROACH. Whether the MDF-approach can be applied to a particular problem depends on the extent to which the experimenter can justify the necessary assumptions regarding the problem and the distribution function involved. The purpose of this section is to discuss certain fairly nonrestrictive conditions which define classes of distributions to which the MDF-approach is applicable. It will also be shown that the MDF-approach is a highly accurate approximate procedure for reliability evaluation when the distribution involved is exponential, and thus does not satisfy Condition 4.

Proposition 2 establishes that the MDF-approach is applicable to a fairly commonly occurring class of distributions.

Proposition 2. Let X be a continuous random variable with $\mathscr{S}_X = \{x \mid x \geq 0\}$, and let $T \epsilon \mathscr{S}_X$. If the density function $f(x)$ is monotone nondecreasing on $[0, mT]$, where $m \geq 1$, then $F(mT) \geq mF(T)$.

Proof: Let $x \epsilon [T, mT]$. By hypothesis, $f(x)$ is monotone nondecreasing on $[0, x]$, so that $x \cdot f(x) \geq F(x)$. Thus $f(x)/F(x) \geq 1/x$. Since this is true for every $x \epsilon [T, mT]$, it follows that

$$\int_T^{mT} \frac{f(x)}{F(x)}\, dx \geq \int_T^{mT} \frac{dx}{x} .$$

This is equivalent to saying that

$$\log \frac{F(mT)}{F(T)} \geq \log \left(\frac{mT}{T}\right) = \log m,$$

or that $F(mT) \geq mF(T)$. Q.E.D.

From Proposition 2, it is immediate that if $f(x)$ is monotone nondecreasing on $[0, \xi_p]$, where $F(\xi_p) = p$, then the MDF technique is valid for all values of m such that $m \in [1, \xi_p/T]$.

It is appropriate here to point out that the hypothesis

$$H_0: F(mT) > 1-\beta$$

is logically equivalent to the statement

$$mT > \xi_{1-\beta} .$$

Hence, H_0 could be written in the more illuminating, if redundant, form

$$H_0: F(mT) > 1-\beta \text{ and } mT > \xi_{1-\beta} .$$

Therefore, to reject H_0 is to conclude that

$$F(mT) \leq 1-\beta$$

and

$$mT \leq \xi_{1-\beta} .$$

Thus, for example, if one can assume that $f(x)$ is monotone nondecreasing on $[0, \xi_p]$, and if $1-\beta \leq p$, then acceptance of H_1 implies the simultaneous validity of the relations $F(mT) \geq mF(T)$ and $1-\beta \geq F(mT)$. Thus acceptance of H_1 implies that

$$F(T) \leq \frac{1-\beta}{m} ,$$

so that

$$r\left(1 - \left[\frac{1-\beta}{m}\right], \gamma\right)$$

is concluded.

Example. Suppose it can be assumed that the density function involved is monotone nondecreasing on $[0, \xi_{.50}]$. Two problems are considered: 1) Design a test which will determine whether the conclusion $r(.95, .90)$ is valid on the basis of a sample of size 21; 2) given a sample of size 9, and $X_{(1)} \geq 4.3T$, what is the strongest statement of the form $r(\alpha, .90)$ that can be concluded?

Solution to Problem 1: The MDF-hypothesis test here can be expressed as follows: We are given $n = 21$, $1-\gamma = .10$, $\alpha = .95$. Thus $(1-\beta)/m = .05$, and $\beta^{21} = 1-\gamma = .10$. Thus, $\beta = .896$, so that $.104/.05 = m = 2.08$. Hence, the hypotheses are

$$H_0: F(2.08T) > .104$$

and

$$H_1: F(2.08T) \leq .104.$$

H_0 is rejected if $X_{(1)} \geq 2.08T$. If $X_{(1)} \geq 2.08T$, then it is accepted that $F(2.08T) \leq .104$ and $2.08T \leq \xi_{.104} < \xi_{.50}$, so that $.104 \geq F(2.08T) \geq 2.08\, F(T)$, with at least γ-confidence; i.e., if $X_{(1)} \geq 2.08T$, $r(.95, .90)$ is concluded.

Solution to Problem 2: Here $n = 9$, $m = 4.3$ and $1-\gamma = .10$. Hence, $(1-\beta)/4.3 = 1-\alpha$ and $\beta^9 = .10$, so that $\beta = .7745$, and $\alpha = .0524$. Hence, suppose

$$H_0: F(4.3T) > .2355$$

and

$$H_1: F(4.3T) \leq .2355.$$

Then, $X_{(1)} \geq 4.3T$ results in the conclusion $r(.9476, .90)$.

We now compare the results of applying the MDF-technique to a situation in which the random variable X involved actually has the exponential distribution with $f(x) = \lambda e^{-\lambda x}$, $\lambda > 0$, and thus does not satisfy Condition 4 for any value of $m > 1$. Table 3.1 provides comparisons of MDF-approach results with lower .90 confidence bounds $(1-\beta^n = .90)$ for the reliability obtained under the assumption that X actually has the exponential distribution and $X_{(1)} \geq mT$ for the sample sizes and values of m shown. The hypothesis test involved in obtaining the bounds under the exponential assumption is nearly identical to the MDF-hypothesis test, with the exception that the MDF assumption $F(mT) \geq mF(T)$ is omitted and the bound for F(T) is obtained from the fact that when X has the exponential distribution

$$F(mT) \leq 1-\beta \Rightarrow F(T) \leq 1-\beta^{1/m} .$$

The validity of this implication is seen as follows:

$$F(mT) = 1-e^{-\lambda mT} \leq 1-\beta \Leftrightarrow e^{-\lambda mT} \geq \beta \Leftrightarrow$$

$$e^{-\lambda T} \geq \beta^{1/m} \Leftrightarrow 1-e^{-\lambda T} \leq 1-\beta^{1/m}$$

$$\Leftrightarrow F(T) \leq 1-\beta^{1/m} .$$

TABLE 3.1

MDF-Conclusion	r(.99,.90)		r(.95,.90)		r(.90,.90)	
m	n	$\beta^{1/m}$	n	$\beta^{1/m}$	n	$\beta^{1/m}$
1	230	.990	45	.950	22	.900
2	114	.990	22	.949	11	.895
2.5	91	.990	18	.948	9	.892
3	76	.989	15	.947	7	.888
3.75	61	.989	12	.946	5	.881
4	57	.988	11	.946	5	.881
5	45	.988	8	.944	4	.871
6	38	.987	7	.943	3	.858

For example, the MDF-approach conclusion based on a sample of 76 systems and $X_{(1)} \geq 3T$ is r(.99,.90). The corresponding conclusion based on the assumption that X has an exponential distribution is r(.989,.90).

It is possible to analytically explain the lack of sensitivity of the MDF approach to certain types of departures from Condition 4. In particular, the condition $F(mT) \geq mF(T)$ for every $m \in [1, \theta]$ is equivalent to saying that for $x \in [T, \theta T]$ the distribution function $F(x)$ dominates the function

$$L(x) = \frac{F(T)}{T} x .$$

(See Figure 1.)

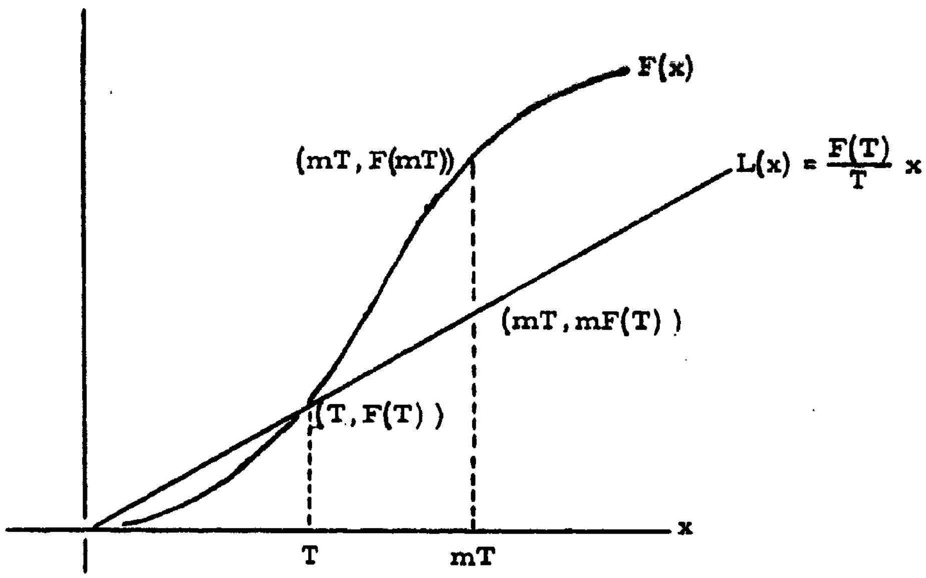


FIGURE 1

For $x \in [0, T/F(T)]$, $L(x)$ can be thought of as the distribution function of a random variable which is uniformly distributed on $[0, T/F(T)]$. Thus, if in reality X has a distribution function with the property that $F(mT) < mF(T)$ for $m \geq 1$, then $F(x)$ will be dominated by $L(x)$, the slope of which $\frac{F(T)}{T}$ shall be "small" when $T/F(T)$ is large.

What happens when X has the exponential distribution, with $F(x) = 1 - e^{-\lambda x}$, is this (see Figure 2): If λx is small, i.e., if $1/\lambda$ is large relative to x, then $1 - e^{-\lambda x} \doteq \lambda x$. That is, $F(x)$ is closely approximated by the distribution function of a random variable which is uniformly distributed on $[0, 1/\lambda]$. Since $\lambda \doteq F(x)/x$ for small values of λx, if λT is small $F(T)/T \doteq \lambda$, so that $F(x) \doteq xF(T)/T = L(x)$, which accounts for the relatively small error in the MDF conclusions for small values of m.

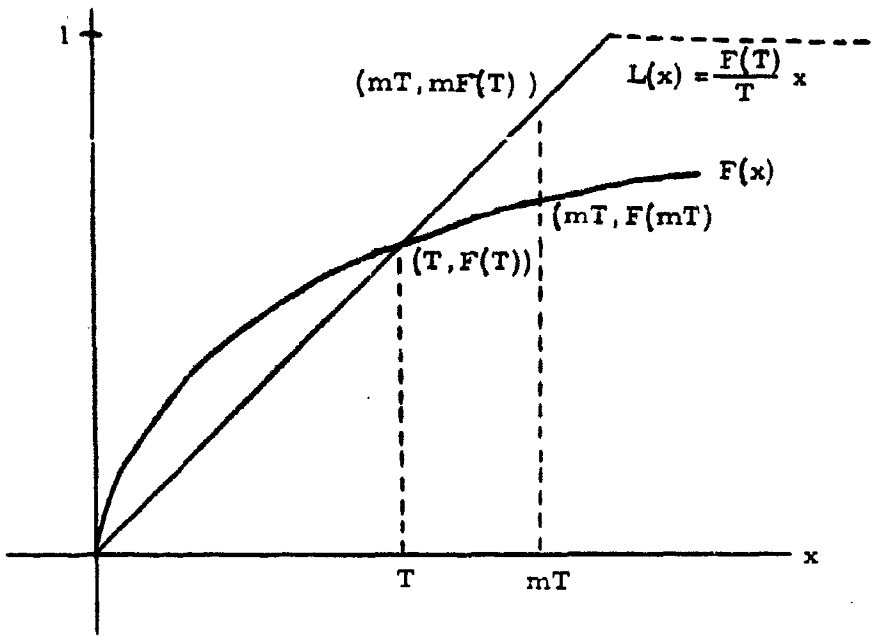


FIGURE 2

4. A CONSIDERATION OF THE WEIBULL DISTRIBUTION. The Weibull distribution occupies an important position in the theory of reliability. Thus, it is useful to compare MDF results with those obtained under the assumption that the random variable under consideration has the Weibull distribution. For these comparisons, it is assumed that the density function of X is given by $f(x) = \lambda x^{\theta - 1} e^{-\lambda x^{\theta}}$, where $x \geq 0$, $\theta \geq 1$ and $\lambda > 0$. The distribution function of X is $F(x) = 1 - e^{-\lambda x^{\theta}}$. Statements of the form $r(\alpha, \gamma)$ can be obtained for this case using the MDF hypothesis test structure with the implication $F(mT) \leq 1 - \beta(m)$ $\Rightarrow F(T) \leq \frac{1 - \beta(m)}{m}$ replaced by the implication $F(mT) \leq 1 - \beta(m) \Rightarrow F(T) \leq 1 - [\beta(m)]^{\frac{1}{m}}$. Note that this is the same substitution which was made in Section 3 when statements $r(\alpha, \gamma)$ were obtained for the exponential distribution using the MDF hypothesis test structure. That the implication $F(mT) \leq 1 - \beta(m) \Rightarrow F(T) \leq 1 - [\beta(m)]^{\frac{1}{m}}$ is valid when X has the Weibull distribution with $\theta \geq 1$ is seen as follows:

$$F(mT) \leq 1 - \beta(m) \iff 1 - e^{-\lambda(mT)^{\theta}} \leq 1 - \beta(m)$$

$$\iff e^{-\lambda T^{\theta} m^{\theta}} \geq \beta(m) \iff e^{-\lambda T \theta} \geq [\beta(m)]^{\frac{1}{m^{\theta}}}$$

$$\iff 1 - e^{-\lambda T \theta} \leq 1 - [\beta(m)]^{\frac{1}{m^{\theta}}} \iff F(T)$$

$$\leq 1 - [\beta(m)]^{\frac{1}{m^{\theta}}} .$$

But $1 - [\beta(m)]^{\frac{1}{m^{\theta}}} \leq 1 - [\beta(m)]^{\frac{1}{m}}$ since $\beta \epsilon [0, 1]$, $m \leq 1$ and $\theta \geq 1$. Thus, if $X_{(1)} \geq mT$, where $Pr(X_{(1)} \geq mT \mid H_0 \text{ is true}) \leq 1 - \gamma$, the statement $r([\beta(m)]^{\frac{1}{m}}, \gamma)$ may be concluded.

Example. In [2], Lieberman and Johns have presented a method for estimating reliability when the random variable involved has a Weibull distribution. Section 6 of [2] presents an illustrative example in which the reliability of a system for a mission of T = 40 hours is estimated, with γ = 90 confidence, on the basis of the following observations on the first five order statistics: $X_{(1)} = 50$, $X_{(2)} = 75$, $X_{(3)} = 125$, $X_{(4)} = 250$ and $X_{(5)} = 300$. The sample size used is 10. Using the estimation method they derived, the authors conclude r(.796,.90). Had the authors simply used binomial reliability tables [1], they would have concluded r(.794.,90), since no mission failures occurred in 10 trials. By way of comparison, if one employed the MDF technique under the assumption that f(x) is monotone increasing on $[0, \xi_p)$ for any $p \leq .20$, the conclusion would be r(.835,.90), while if one were to utilize the MDF type technique adapted, as described, to the assumption that X has the Weibull distribution with $\theta \geq 1$, the conclusion would be r(.832,.90). It should be noted that the estimating procedure of Lieberman and Johns does not involve any assumptions on the values of θ, which at least partially explains the relatively small difference between their estimate and the binomial estimate. The MDF technique becomes increasingly less accurate as an approximate method as θ approaches 0. Therefore, caution should be used in applying the technique to a Weibull situation if it is suspected that θ is actually less than 1.

REFERENCES

1. Confidence Limits for Attribute Data, second revision, Lockheed Missile and Space Company, Sunnyvale, California, 1966, pp. 1-87.

2. M. V. Johns, Jr., and G. J. Lieberman, "An Exact Asymptotically Efficient Confidence Bound for Reliability in the Case of Weibull Distribution", Technometrics, v. 8, 1966, pp. 135-175.

RELIABILITY IN COMPLEX SYSTEMS*

A. Clifford Cohen
The University of Georgia

1. INTRODUCTION. It is well established that the reliability of complex systems varies with time. Following a break-in and adjustment period during which minor deficiencies are corrected, a system is placed in service with an initial reliability R_o. Thereafter the reliability either increases as further system deficiencies are corrected or it decreases as components deteriorate with age. In the life of some systems, there is an early period during which reliability increases, and a subsequent period of constant or decreasing reliability. Our attention in this paper is limited to models of monotone increasing and monotone decreasing reliability.

2. EXPONENTIAL MODELS.

Increasing Reliability. With R(t) designating reliability at time t, a simple exponential model for increasing reliability may be expressed as

$$R(t) = 1 - (1 - R_o)e^{-\alpha t}, \quad \alpha \geq 0, \quad t \geq 0, \tag{1}$$

where R_o and α are parameters to be estimated from sample data.

Decreasing Reliability. When reliability decreases with time, we consider the following relationship

$$R(t) = R_o e^{-\alpha t}, \quad \alpha \geq 0, \quad t \geq 0, \tag{2}$$

where again R_o and α are parameters to be estimated from sample data.

3. MAXIMUM LIKELIHOOD ESTIMATION. Let n_i specimens be tested at time t_i and let x_i designate the number of successes achieved $(i = 0, 1, \ldots k)$. Sample data resulting from a sequence of such tests then consist of the triples $(t_o, n_o, x_o), (t_1, n_1, x_1), \ldots (t_k, n_k, x_k)$. From these data, we must determine which model is appropriate (i. e.

*Reprinted with permission of the American Society for Quality Control from Transactions of the Twentieth Annual A. S. Q. C. Technical Conference, June 1966.

increasing or decreasing) and then estimate the parameters. Before examining the problem of choosing between models, we will first consider estimation in each model separately.

With the reliability $R(t_i)$ at time t_i abbreviated to R_i, the likelihood function for our sample may be expressed as

$$(3)\qquad L[(n_o, x_o)\ldots(n_k, x_k)] = \prod_{i=0}^{k} \binom{n_i}{x_i} R_i^{x_i}(1 - R_i)^{n_i - x_i}\,.$$

Estimation in the Increasing Model. When R_i is given by equation (1), we make this substitution in (3) to obtain

$$(4)\qquad L_I(R_o, a) = \prod_{i=0}^{k} \binom{n_i}{x_i}[1 - (1 - R_o)e^{-a\,t_i}]^{x_i}[(1 - R_o)e^{-a\,t_i}]^{n_i - x_i},$$

where the subscript $_{(I)}$ indicates employment of the increasing model. On taking logarithms of (4) differentiating with respect to R_o and a in turn, we obtain

$$(5)\qquad \begin{cases} \dfrac{\partial \ln L_I}{\partial R_o} = -\dfrac{\sum_0^k (n_i - x_i)}{1 - R_o} + \sum_0^k \dfrac{x_i e^{-a\,t_i}}{1 - (1 - R_o)e^{-a\,t_i}}, \\[2ex] \dfrac{\partial \ln L_I}{\partial a} = -\sum_0^k t_i(n_i - x_i) + (1 - R_o)\sum_0^k \dfrac{t_i x_i e^{-a\,t_i}}{1 - (1 - R_o)e^{-a\,t_i}}\,. \end{cases}$$

On setting these equations equal to zero and simplifying, estimating equations in the case of increasing reliability become

$$(6)\quad \begin{cases} (1 - R_o) \sum_0^k \dfrac{x_i e^{-\alpha t_i}}{1-(1-R_o)e^{-\alpha t_i}} = \sum_0^k (n_i - x_i)\,, \\[2ex] (1 - R_o) \sum_0^k \dfrac{t_i x_i e^{-\alpha t_i}}{1-(1-R_o)e^{-\alpha t_i}} = \sum_0^k t_i(n_i - x_i)\,. \end{cases}$$

When (1) is the appropriate model, the required estimates $\hat{R}_o$ and $\hat{\alpha}$ can be found by simultaneously solving (6) using standard iterative techniques. Should the value $\hat{\alpha}$ thereby obtained from some given sample turn out to be negative, this suggests that the increasing model is inappropriate and that we should either set $\hat{\alpha} = 0$ or investigate the decreasing model of equation (2).

Estimation in the Decreasing Model. When R_i is given by equation (2), we make this substitution in (3) and thereby obtain

$$(7)\qquad L_D(R_o,\alpha) = \prod_{i=0}^{k} \binom{n_i}{x_i} (R_o e^{-\alpha t_i})^{x_i} (1 - R_o e^{-\alpha t_i})^{n_i - x_i},$$

where the subscript (D) indicates employment of the decreasing model. On taking logarithms and differentiating, we have

$$(8)\quad \begin{cases} \dfrac{\partial \ln L_D}{\partial R_o} = \dfrac{\sum_0^k x_i}{R_o} - \sum_0^k \dfrac{(n_i - x_i)e^{-\alpha t_i}}{1 - R_o e^{-\alpha t_i}}\,, \\[2ex] \dfrac{\partial \ln L_D}{\partial \alpha} = -\sum_0^k x_i t_i + R_o \sum_0^k \dfrac{t_i(n_i - x_i)e^{-\alpha t_i}}{1 - R_o e^{-\alpha t_i}}\,. \end{cases}$$

On equating the above partials to zero, the estimating equations become

$$(9)\quad \begin{cases} R_o \sum_0^k \dfrac{(n_i - x_i)e^{-\alpha t_i}}{1 - R_o e^{-\alpha t_i}} = \sum_0^k x_i \,, \\[2ex] R_o \sum_0^k \dfrac{t_i(n_i - x_i)e^{-\alpha t_i}}{1 - R_o e^{-\alpha t_i}} = \sum_0^k t_i x_i \,. \end{cases}$$

When (2) is the appropriate model, the required estimates $\hat{R}_o$ and $\hat{\alpha}$ are found by simultaneously solving the two equations of (9). In this case, should the value $\hat{\alpha}$ thereby obtained, turn out to be negative (an unacceptable result) this suggests that either we should set $\hat{\alpha} = 0$ or that the increasing model of (1) should be employed.

4. CHOOSING THE MODEL. In many applications, a' priori considerations dictate which of the models considered here is appropriate. In others, the sample data will clearly indicate which model is to be preferred. In perhaps the majority of applications, the choice of the model will involve a more careful analysis of sample data, and the following procedure is suggested for choosing between the increasing reliability model of (1) and the decreasing reliability model of (2).

1. Solve equations (6) for tentative estimates of R_o and α in the increasing reliability model. If the tentative estimate of α thus obtained is positive, accept both tentative estimates and designate them as $\hat{R}_{oI}$ and $\hat{\alpha}_I$.

   If the estimate of α obtained from (6) is negative then accept as estimates $\hat{\alpha}_I = 0$ and $R_{oI} = \sum_{i=0}^{k} x_i/n$ where $n = \sum_{i=0}^{k} n_i$.

2. Solve equations (9) for tentative estimates of R_o and α in the decreasing reliability model. If the tentative estimate of α from these equations is positive, then accept both tentative estimates from (9) and designate them as $\hat{R}_{oD}$ and $\hat{\alpha}_D$. If the estimate of α obtained from (9) is negative, accept as estimates $\hat{\alpha}_D = 0$ and $\hat{R}_{oD} = \sum_{i=0}^{k} x_i/n$.

3. Calculate $\hat{L}_I = L_I(\hat{R}_{oI}, \hat{a}_I)$ and $\hat{L}_D = L_D(\hat{R}_{oD}, \hat{a}_D)$ using equations (4) and (7) respectively.

4. If $\hat{L}_I > \hat{L}_D$, choose the increasing reliability model; if $\hat{L}_D > \hat{L}_I$, choose the decreasing reliability model. Otherwise (if $\hat{L}_I = \hat{L}_D$), we employ the constant reliability model $R_i = R_o$ with $a = 0$, and with $\hat{R}_o = \sum_{i=1}^{k} x_i/n$.

5. **ASYMPTOTIC VARIANCES AND COVARIANCES OF ESTIMATES.** The asymptotic variance-covariance matrix of the maximum likelihood estimates $\hat{R}_o$ and $\hat{a}$ is given as

$$(10)\quad \begin{bmatrix} -E\left(\frac{\partial^2 \ln L}{\partial R_0^2}\right) & -E\left(\frac{\partial^2 \ln L}{\partial R_o \partial a}\right) \\ -E\left(\frac{\partial^2 \ln L}{\partial a \partial R_o}\right) & -E\left(\frac{\partial^2 \ln L}{\partial a^2}\right) \end{bmatrix}^{-1} = \begin{bmatrix} V(\hat{R}_o) & Cov(\hat{R}_o, \hat{a}) \\ Cov(\hat{R}_o, \hat{a}) & V(\hat{a}) \end{bmatrix},$$

where E symbolizes expected values. In practice, satisfactory approximations can be obtained by replacing expected values of the partials with their actual values calculated using $R_o = \hat{R}_o$ and $a = \hat{a}$. The required second partials follow from further differentiation of (5) and (8) in turn. These results are given below.

<u>For Increasing Reliability.</u>

$$(11)\quad \left\{ \begin{aligned} \frac{\partial^2 \ln L_I}{\partial R_0^2} &= -\sum_0^k \frac{(n_i - x_i)}{(1-R_o)^2} - \sum_0^k \frac{x_i e^{-2at_i}}{[1-(1-R_o)e^{-at_i}]^2}, \\ \frac{\partial^2 \ln L_I}{\partial R_o \partial a} &= \frac{\partial^2 \ln L_I}{\partial a \partial R_o} = -\sum_0^k \frac{x_i t_i e^{-a t_i}}{[1-(1-R_o)e^{-a t_i}]^2}, \\ \frac{\partial^2 \ln L_I}{\partial a^2} &= -(1-R_o)\sum_0^k \frac{t_i^2 x_i e^{-a t_i}}{[1-(1-R_o)e^{-a t_i}]^2}. \end{aligned} \right.$$

For Decreasing Reliability.

$$(12)\quad \begin{cases} \dfrac{\partial^2 \ln L_D}{\partial R_o^2} = -\sum_0^k x_i/R_o^2 - \sum_0^k \dfrac{(n_i - x_i)e^{-2\alpha t_i}}{(1 - R_o e^{-\alpha t_i})^2}, \\ \dfrac{\partial^2 \ln L_D}{\partial R_o \partial \alpha} = \dfrac{\partial^2 \ln L_D}{\partial \alpha\, \partial R_o} = \sum_0^k \dfrac{(n_i - x_i)t_i e^{-\alpha t_i}}{(1 - R_o e^{-\alpha t_i})^2}, \\ \dfrac{\partial^2 \ln L_D}{\partial \alpha^2} = R_o \sum_0^k \dfrac{t_i^2 (n_i - x_i)e^{-\alpha t_i}}{(1 - R_o e^{-\alpha t_i})^2}. \end{cases}$$

Although asymptotic variances and covariances might be misleading for small samples, they should closely approximate the true variances and covariances for moderate size samples; i.e. for $n = \sum_1^k n_i$ in excess of say 20.

The variance of $\hat{R}_i$ which is of course a function of $\hat{R}_o$ and $\hat{a}$, can be approximated by employing a theorem of Cramér [1] which enables us to write

$$(13)\quad V(\hat{R}_i) \doteq \left(\frac{\partial R_i}{\partial R_o}\right)^2 V(\hat{R}_o) + 2\left(\frac{\partial R_i}{\partial R_o}\right)\left(\frac{\partial R_i}{\partial \alpha}\right)\mathrm{Cov}(\hat{R}_o, \hat{a}) + \left(\frac{\partial R_i}{\partial \alpha}\right)^2 V(\hat{a}) .$$

For the increasing reliability model, it follows from equation (1) that

$$\frac{\partial R_i}{\partial R_o} = e^{-\alpha t_i} \text{ and } \frac{\partial R_i}{\partial \alpha} = t_i(1 - R_o)e^{-\alpha t_i} .$$

Accordingly, in this case, we have

$$(14)\quad V(\hat{R}_i) \doteq e^{-2\hat{a} t_i}[V(\hat{R}_o) + 2(1 - \hat{R}_o)t_i \mathrm{Cov}(\hat{R}_o, \hat{a}) + (1 - \hat{R}_o)^2 t_i^2 V(\hat{a})] .$$

For the decreasing reliability model, it follows from equation (2) that

$$\frac{\partial R_i}{\partial R_o} = e^{-\alpha t_i}, \text{ and } \frac{\partial R_i}{\partial \alpha} = -t_i R_o e^{-\alpha t_i} .$$

In this case $V(R_i)$ becomes

(15) $$V(\hat{R}_i) \doteq e^{-2\hat{\alpha} t_i}[V(\hat{R}_o) - 2\hat{R}_o t_i \mathrm{Cov}(\hat{R}_o \hat{\alpha}) + \hat{R}_o^2 t_i^2 V(\hat{\alpha})] .$$

6. **ILLUSTRATIVE EXAMPLES.** In order to illustrate the practical application of results of this investigation, let us consider simulated test data on two complex systems, one with increasing reliability and the other with decreasing reliability.

Example 1. Increasing Reliability. Following are results of the initial and four subsequent tests conducted on this system.

t_i	(time periods)	0	1	2	3	4
n_i	(number tested)	20	10	5	5	5
x_i	(number successes)	13	8	5	4	5
x_i/n_i	(success ratio)	0.65	0.80	1.00	0.80	1.00

Summarizing, we have $n = \sum_0^4 n_i = 45$, $\sum_0^4 x_i = 35$, $\sum_0^4 t_i x_i = 50$, $\sum_0^4 (n_i - x_i) = 10$, and $\sum_0^4 t_i(n_i - x_i) = 5$. Our problem now is to substitute these values into (6) and solve for the required estimates $\hat{R}_o$ and $\hat{\alpha}$. Any standard iterative method might be employed for this purpose, but the following procedure seems relatively straightforward and should be generally satisfactory.

As an initial approximation, $R_o^{(o)}$, we select the initial success ratio $x_o/n_o = 0.65$, and as an initial approximation to R_1, the success ratio $x_1/n_1 = 0.80$. We substitute these two values in (1) with $t_1 = 1$ to obtain

$$0.80 = 1 - (1 - 0.65)e^{-\alpha} , \text{ and it follows that}$$

$$e^{-\alpha} = \frac{1 - 0.80}{0.35} = 0.57143.$$

Reading from a table of exponential functions, we have as an initial approximation, $\alpha^{(o)} = 0.56$. The superscripts serve to indicate the order of the approximations. We subsequently try letting $\alpha = 0.50$ and 0.55 respectively in the two equations of (6) and solve these in turn for R_o. We of course are seeking a value of α such that the two values of R_o thus obtained are identical. Following is a summary of these results including interpolation to obtain new approximations $\alpha^{(1)}$ and $R_o^{(1)}$.

α	R_o from 1st. Eq. of (6)	R_o from 2nd Eq. of (6)	Difference
0.500	0.662	0.685	-0.023
0.543	0.656	0.656	0
0.550	0.655	0.651	+0.004

As new approximations, we have $\alpha^{(1)} = 0.543$ and $R_o^{(1)} = 0.656$.

We now elect to seek further improvement through Newton's method which is based on Taylor series expansions of the estimating equations about a point in the vicinity of their simultaneous solution. Let h and k designate corrections to be determined by the iteration process so that $\hat{R}_o = R_o^{(1)} + h$ and $\hat{\alpha} = \alpha^{(1)} + k$. Using Taylor's theorem and neglecting terms containing powers of h and k above the first, we have as correction equations

$$h \frac{\partial^2 \ln L_I}{\partial R_o^2} + k \frac{\partial^2 \ln L_I}{\partial R_o \partial \alpha} = - \frac{\partial \ln L_I}{\partial R_o},$$

$$h \frac{\partial^2 \ln L_I}{\partial R_o \partial \alpha} + k \frac{\partial^2 \ln L_I}{\partial \alpha^2} = - \frac{\partial \ln L_I}{\partial \alpha},$$

which are to be solved simultaneously for h and k.

Using (5) and (11) we evaluate the partials in these equations at the point $R_o = 0.656$, $\alpha = 0.543$, and the correction equations become

$$-119.9088\, h - 16.7561\, k = 0.0998,$$

$$-16.7561\, h - 11.6602\, k = 0.0038.$$

Solving, we have $h = -0.00098$ and $k = 0.00109$. Thus the final estimates become

$$\hat{R}_o = 0.656 - 0.00098 = 0.6550,$$

$$\hat{\alpha} = 0.543 + 0.00109 = 0.5441 .$$

As verification of the accuracy of these final estimates, they were substituted into the first partials of (5) with results as follows:

$$\left.\frac{\partial \ln L_I}{\partial R_o}\right|_{\substack{R_o=0.6550 \\ \alpha\ =0.5441}} = 0.001, \qquad \left.\frac{\partial \ln L_I}{\partial \alpha}\right|_{\substack{R_o=0.6550 \\ \alpha\ =0.5441}} = 0.001.$$

Values of zero would have indicated perfect agreement. The small values obtained here are considered satisfactory and no further iterations are deemed necessary.

Rather than employing the intermediate interpolative procedure, we might have moved directly from the initial approximations to the Newton method. In that case, of course, one or more additional cycles of the Newton iteration might have been required to reach the same final results as those obtained here.

Using values of the second partials employed as coefficients in the correction equations, the variance-covariance matrix of (10) is approximated as

$$\begin{bmatrix} 119.9088 & 16.7561 \\ 16.7561 & 11.6602 \end{bmatrix}^{-1} = \begin{bmatrix} 0.0104 & -0.0150 \\ -0.0150 & 0.1073 \end{bmatrix} .$$

Accordingly we have

$$V(\hat{R}_o) \doteq 0.0104, \quad V(\hat{\alpha}) \doteq 0.1073, \quad Cov(\hat{R}_o, \hat{\alpha}) \doteq -0.0150.$$

Using these values in equation (14) we calculate $V(\hat{R}_i)$ at times $t_i = 0, 1, 2, 3$ and 4. We also calculate the predicted values of R_i (i.e. $\hat{R}_i$) at these times using equation (1) with $\alpha = \hat{\alpha} = 0.5441$ and $R_o = \hat{R}_o = 0.6550$. These results are displayed below along with actual success ratios for comparison.

t_i	0	1	2	3	4
x_i/n_i	0.65	0.80	1.00	0.80	1.00
$\hat{R}_i$	0.6550	0.7998	0.8838	0.9326	0.9609
$V(\hat{R}_i)$	0.0104	0.0043	0.0046	0.0036	0.0022

An attempt to fit the decreasing model of (2) to these data resulted in a value $\alpha < 0$ as a solution of (9). We were thus led to estimates $\hat{\alpha}_D = 0$ and $\hat{R}_{oD} = \sum_{i=0}^{4} x_i/n = 0.7778$. Using these estimates in (7), we calculate $\hat{L}_D = 0.0008$, whereas using the estimates $\hat{\alpha}_I = 0.1472$ and $\hat{R}_{oI} = 0.6074$ in equation (4), which applies when the increasing model of (1) is employed, we calculate $\hat{L}_I = 0.005$. Since $\hat{L}_I > \hat{L}_D$, our choice of the increasing reliability model of (1) in this instance is verified as being correct.

Illustrative Example 2. Decreasing Reliability. Following are test data on a system in which reliability is decreasing with time.

t_i	(Time periods)	0	1	2	3	4
n_i	(Number Tested)	20	5	5	5	3
x_i	(Number Successes)	12	3	2	2	1
x_i/n_i	(Success Ratio)	60	.60	.40	.40	.33

Summarizing, we have $n = \Sigma_1^4 n_i = 38$, $\Sigma_1^4 x_i = 20$, $\Sigma_1^4 t_i x_i = 17$, $\Sigma_1^4 (n_i - x_i) = 18$, and $\Sigma_1^4 t_i(n - x_i) = 25$. Proceeding to solve equations (9) using these data, we again select as an initial approximation to R_o, the initial success ratio. Thus we have $R_o^{(0)} = 0.60$. The initial approximation to α comes from a free-hand curve through the points on a plot of the success ratios versus time as $\alpha^{(0)} = 0.12$.

This time, we skip the intermediate approximations as used in the first illustration and proceed immediately to the Newton method. At the end of one cycle, we have as improved approximations

$$R_0^{(1)} = 0.605 \text{ and } \alpha^{(1)} = 0.145.$$

With the partials of (8) and (11) evaluated for $R_o = 0.605$ and $\alpha = 0.145$, the correction equations become

$$-122.9348\,h + 48.5840\,k = -0.1886,$$

$$48.5840\,h - 76.0959\,k = -0.0496.$$

On solving, we find

$$h = 0.0024 \text{ and } k = 0.0022,$$

and as final estimates (or new approximations) we have

$$\hat{R}_o = 0.6050 + .0024 = 0.6074$$

$$\hat{\alpha} = 0.1450 + .0022 = 0.1472.$$

These values are substituted into the first partials of (8) with the following results

$$\left.\frac{\partial \ln L_D}{\partial R_o}\right|_{\substack{R_o=0.6074 \\ \alpha=0.1472}} = -0.0001, \quad \left.\frac{\partial \ln L_D}{\partial \alpha}\right|_{\substack{R_o=0.6074 \\ \alpha=0.1472}} = -0.0012.$$

These values are considered to be sufficiently close to zero to justify acceptance of $\hat{R}_o = 0.6074$ and $\hat{\alpha} = 0.1472$ as final estimates, and no further iterations were made.

As in illustration 1, the variance-covariance matrix of $\hat{R}_o$ and $\hat{\alpha}$ is approximated using coefficients of the correction equations. Thus we have

$$\begin{bmatrix} 122.9348 & -48.5840 \\ -48.5840 & 76.0959 \end{bmatrix}^{-1} = \begin{bmatrix} 0.0109 & 0.0069 \\ 0.0069 & 0.0176 \end{bmatrix}.$$

Accordingly for this example, we have $V(\hat{R}_o) \doteq 0.0109$, $V(\hat{\alpha}) \doteq 0.0176$, and $Cov(\hat{R}_o, \hat{\alpha}) \doteq 0.0069$. The variance of $\hat{R}_i$ at $t_i = 0, 1, 2, 3$ and 4 is computed from (15) and the predicted (estimated) values of R_i

(designated $\hat{R}_i$) for these same time values are computed from (2). These results along with the success ratios are displayed below.

t_i	0	1	2	3	4
x_i/n_i	0.60	0.60	0.40	0.40	0.33
$\hat{R}_i$	0.6024	0.5246	0.4525	0.3906	0.3371
$V(\hat{R}_i)$	0.0109	0.0069	0.0112	0.0183	0.0250

An attempt to fit the increasing model of (1) to the data for this example resulted in a value $\alpha < 0$ as a solution of (6) and we were thus led to $\hat{\alpha}_I = 0$ and $\hat{R}_{oI} = \sum_{i=0}^{4} x_i/n = 0.5263$. Using these estimates in (4), we calculate $\hat{L}_I = 0.001$, whereas using the estimates $\hat{\alpha}_D = 0.1472$ and $\hat{R}_{oD} = 0.6072$ in (7), we calculate $\hat{L}_D = 0.002$. Thus with $\hat{L}_D > \hat{L}_I$ for these data, the decreasing model of (2) is the proper choice.

7. SOME CONCLUDING REMARKS. Although questions relating to how many tests should be conducted and when they should be scheduled, have not been formally examined here, they are not to be dismissed as being unimportant. When tests are destructive and the cost is great, there is considerable pressure to limit their number. Considerations having little to do with statistics or probability often dictate that a rather large proportion of available test specimens be expended in the initial tests. Such allocation, of course, limits the number available for subsequent testing. Further studies in this area to determine optimum test designs are still in progress.

When this investigation was begun, it was intended to consider not only the exponential models, but also the hyperbolic model

$$R(t) = R_\infty + \frac{R_o - R_\infty}{\alpha t + 1}; \quad t \geq 0, \quad \text{where } 0 \leq R_o \leq 1,$$

$0 \leq R_\infty \leq 1$, and $\alpha > 0$. As in the exponential models, R_o is the initial probability at time $t = 0$. R_∞ is the final or ultimately attainable reliability; i.e. $\lim_{t \to \infty} R(t) = R_\infty$. In this model, reliability is increasing or decreasing with time depending on whether $R_\infty > R_o$ or $R_\infty < R_o$.

A special case of the hyperbolic model with $\alpha = 1$ and $R_\infty > R_o$ has been considered by Lloyd and Lipow [2] .

Procedures similar to those employed in this paper can be used to estimate parameters α , R_o, and R_∞ in the general case, but in view of the length that the present paper has already attained, further consideration of this model is being temporarily deferred.

REFERENCES

[1] Cramér, Harold, 1946. "Mathematical Methods of Statistics," page 353, Princeton University Press.

[2] Lloyd, David K. and Lipow, Myron, 1962. Reliability: Management, Methods, and Mathematics, Prentice-Hall, Inc. , Englewood Cliffs, New Jersey.

ESTIMATION OF TIME FUZE CHARACTERISTICS BY NON-LINEAR REGRESSION METHODS

Weldon F. Willoughby
U. S. Army Ballistic Research Laboratories
Aberdeen Proving Ground, Maryland

INTRODUCTION. Ballistic tests of mechanical time fuzes provided data which indicated that the biases in functioning time (i.e., the differences in the running time and the set time) for a given time setting were relatively large and widely dispersed when the firings were conducted at low temperatures. For the firings conducted at higher temperatures, the biases decreased in magnitude and became more uniform as the temperature increased. Since, in the past, the bias in functioning time of mechanical time fuzes assembled to artillery projectiles had been expressed implicitly in the firing tables as a function of set time alone, an investigation was conducted to determine the dependence of fuze bias on temperature as well as set time. In addition, it was desired to find an equation expressing the relationship between fuze bias, temperature and set time which could be programmed for use on the Field Artillery Digital Automatic Computer (FADAC).

Plots of the bias in fuze functioning time versus set time for constant temperatures indicated that the two variables were linearly related. On the other hand, plots of fuze bias versus temperature for constant time settings resembled single branches of rectangular hyperbolas, indicating a nonlinear relationship between bias and temperature.

From these indications, and after trying several models, a candidate model equation containing two linear parameters and one nonlinear parameter was assumed to adequately describe the relationship among fuze bias, the dependent variable, and temperature and set time, the two independent variables. In the model, it was assumed that only the biases were affected by errors of measurement.

As is well known (see [2] and [8], etc.), the method of least squares, which is the method most often used in regression problems, may be used to estimate the parameters of functional relationships among sets of experimental data whenever it can be assumed that:

(a) the dependent variable, Y, is related to known levels of a set of independent variables, $X_1, X_2, \ldots X_k$, by a relationship of the form

(1) $$Y = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \epsilon$$

where the β_i (i = 1, 2, . . . , k) are unknown parameters and ϵ is the error in the observed value of the dependent variable, and

(b) the errors in the observed values of the dependent variable are independent and randomly distributed with zero mean and a common variance. (In addition, if valid statistical tests of significance are to be made, it is also necessary to assume that the errors are normally distributed.)

However, when the functional relationship among the variables cannot be expressed as a linear combination of the unknown parameters as in (1), the usual procedures for estimation by the method of least squares are not directly applicable. Several procedures are available (see [3], [4], [6], and [7]) for estimating the parameters of nonlinear functions. These procedures generally employ a transformation of the function into a linear form either by a change of variables or by an approximation based on a Taylor's series expansion under the assumption that the function is locally linear. In connection with the latter, the approximating procedures require iterative processes to converge to solutions and the advent of high speed computers has greatly facilitated the solution of nonlinear regression problems by these methods.

For this problem, the model equation was assumed to be of the form

$$Y_{ijk} = \frac{\beta_1 + \beta_2 X_{2j}}{X_{2j} + \beta_3} X_{1i} + \epsilon_{ijk}$$

(2)

$$= \frac{\beta_1 X_{1i}}{X_{2j} + \beta_3} + \frac{\beta_2 X_{1i} X_{2j}}{X_{2j} + \beta_3} + \epsilon_{ijk}$$

where Y_{ijk} is the observed fuze bias at time setting X_{1i} and temperature X_{2j} and ϵ_{ijk} is a random error with zero expectation. Assuming this model equation, the regression function to be fitted is

$$\mu = E(Y_{ijk}) = \frac{\beta_1 + \beta_2 X_{2j}}{X_{2j} + \beta_3} X_{1i}$$

(3)

$$= \frac{\beta_1 X_{1i}}{X_{2j} + \beta_3} + \frac{\beta_2 X_{1i} X_{2j}}{X_{2j} + \beta_3}$$

It can be seen by inspection of the first form of (3) that, for a constant temperature (X_{2j} is constant), the regression function represents a straight line passing through the origin (zero set time and zero bias) and for a constant time setting (X_{1i} is constant), the regression function represents a rectangular hyperbola with vertical and horizontal asymptotes.

The function given in (3) was fitted to sets of data obtained from ballistic functioning tests of the mechanical time fuze. Least squares estimates of the three parameters were determined first by an iterative process (after linearizing the function) which exploited the facility and speed of computation of the Ballistic Research Laboratories Electronic Scientific Computer (BRLESC) in scanning the parameter space. Then, as a check on the results obtained by this procedure, least squares estimates were also obtained by the Hartley [5] modification of the Gauss-Newton iteration which in theory has the highly desirable property of guaranteed convergence to estimates yielding the absolute minimum sum of squares of residuals provided the initial estimates of the parameters are in the neighborhood of the final values.

In order to obtain approximate confidence intervals about the individual parameters, as estimate of the variance-covariance matrix of the least squares estimates was obtained using the Fisher information matrix described by Rao [9]. The confidence intervals were constructed by the procedure described by Stone in his discussions on the paper by Beale in [1].

THE SCANNING PROCESS. To determine estimates of the unknown parameters by the scanning process, the regression function was linearized by substituting an initial estimate of the nonlinear parameter β_3. The two linear parameters, β_1 and β_2, were then estimated by the usual least squares procedure. The sum of squares of residuals was computed using the three estimates of the parameters. In the next iteration, the initial estimate of β_3 was changed by some small increment and new estimates of β_1 and β_2 were determined as before. Again, the sum of squares of residuals, using the new estimates, were computed. The process was repeated until a minimum sum of squares of residuals over a rather large range of estimates of β_3 was obtained. The estimates of the parameters which gave the minimum were considered to be the least squares estimates.

If β_3^*, the value of β_3 which gives the minimum sum of squares of residuals, is substituted in (3), the denominator of each of the terms could be considered to be of the form

(4) $$X_{2j}^* = X_{2j} + \beta_3^* \quad .$$

Now, let

$$(5) \qquad t_{1i} = \frac{X_{1i}}{X_{2j}^{*}} \; ; \quad t_{2j} = X_{2j} \quad .$$

Then (3) can be written as

$$(6) \qquad \mu = E(Y_{ijk}) = \beta_1 t_{1i} + \beta_2 t_{1i} t_{2j}$$

which is linear in β_1 and β_2. Least squares estimates of β_1 and β_2, for the given value of β_3^*, may be obtained by solution of the normal equations resulting from minimizing the sum of squares of residuals

$$(7) \qquad Q(\beta) = \sum_{ijk} \{ Y_{ijk} - (\beta_1 t_{1i} + \beta_2 t_{1i} t_{2j}) \}^2$$

with respect to β_1 and β_2, when the errors in the Y_{ijk} are independent and distributed with mean zero and constant variance σ^2. On the other hand, when the errors in the Y_{ijk} have different precision, i.e., the variance of errors in Y_{ijk} is not constant, the sum of squares of residuals to be minimized is of the form

$$(8) \qquad Q(\beta) = \sum_{ijk} \omega_{ij} \{ Y_{ijk} - (\beta_1 t_{1i} + \beta_2 t_{1i} t_{2j}) \}^2$$

where the ω_{ij} are relative weights of the Y_{ijk} which make the quantities

$$(9) \qquad Y_{ijk}^{*} = \sqrt{\omega_{ij}} \; Y_{ijk}$$

have a common variance. (In the case of homogeneous variances, the relative weights, $\omega_{ij} = 1$.) Thus, a predicted value of Y may be determined from the equation

$$\hat{Y}_{ij} = \hat{\beta}_1 t_{1i} + \hat{\beta}_2 t_{1i} t_{2j}$$

(10)
$$= \frac{\hat{\beta}_1 X_{1i}}{X_{2j} + \beta_3^*} + \frac{\hat{\beta}_2 X_{1i} X_{2j}}{X_{2j} + \beta_3^*}$$

The standard error of estimate is given by the expression

(11)
$$\hat{\sigma} = \sqrt{\frac{1}{n-3} \sum_{ijk} \omega_{ij} \{Y_{ijk} - (\hat{\beta}_1 t_{1i} + \hat{\beta}_2 t_{1i} t_{2j})\}^2} \ .$$

In the process of determining the prediction, it was noted in the examination of the data that the dispersions of the observed biases varied considerably from temperature to temperature and to some degree from time setting to time setting. As previously stated, a necessary assumption for the application of the least squares method is that the variances of the errors in the Y_{ijk} be constant. Therefore, it was necessary to perform a transformation of the biases to remove the effect of heterogenous variances at the various temperatures and time settings. A suitable transformation found in [8] is to let

(12)
$$Y'_{ijk} = \frac{Y_{ijk}}{\sigma_{ij}}$$

where Y_{ijk} is the k^{th} observed bias and σ_{ij} is the standard deviation of the biases at the i^{th} set time and j^{th} temperature. The transformed variates, Y'_{ijk}, have the property that their variances equal one. That is,

$$\mathrm{Var}\,(Y'_{ijk}) = E\left[Y'_{ijk} - E\,(Y'_{ijk})\right]^2 = E\left[\frac{Y_{ijk}}{\sigma_{ij}} - E\left(\frac{Y_{ijk}}{\sigma_{ij}}\right)\right]^2$$

(13)
$$= \frac{1}{\sigma_{ij}^{\ 2}} E\left[Y_{ijk} - E\,(Y_{ijk})\right]^2 = \frac{\sigma_{ij}^{\ 2}}{\sigma_{ij}^{\ 2}} = 1 \ .$$

Since the true variances of the biases were not known, the reciprocal of the sample variances were used as the relative weighting factors, ω_{ij}.

In order to cover the range of feasible values of β_3, the estimates of this parameter used in the determination ranged from -10,000 to +10,000. This range was scanned first at intervals of 100, 10, 1, 0.1, and 0.01 until the value of β_3 was found which gave the smallest sum of squares of residuals. In each iteration of the process, least squares estimates of β_1 and β_2 corresponding to the estimate of β_3 were computed. Plots of the error root mean squares (in the subrange indicating a minimum sum of squares of residuals) obtained in each iteration versus the estimated values of β_3 are given in Figure 1, for the three zones of fire. (A smooth curve has been drawn through the points.) Table I gives the least squares estimates of the regression parameters and the sum of squares of residuals for the three zones of fire.

THE HARTLEY MODIFICATION OF THE GAUSS-NEWTON ITERATION. The Hartley modification of the Gauss-Newton iteration initially requires the expansion of the regression function in a multiple first order Taylor's series about initial estimates of the parameters, $\tilde{\beta}_1$, $\tilde{\beta}_2$ and $\tilde{\beta}_3$, obtaining an expression of the form

$$
\begin{aligned}
Y &= f(X_1, X_2;\ \tilde{\beta}_1 + \Delta\beta_1, \tilde{\beta}_2 + \Delta\beta_2, \tilde{\beta}_3 + \Delta\beta_3) \\
&\approx f(X_1, X_2;\ \tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3) + \sum_{i=1}^{3} \frac{\partial f}{\partial \beta_i} \Delta\beta_i
\end{aligned}
\tag{14}
$$

where the partial derivatives are evaluated at $\beta_i = \tilde{\beta}_i$ ($i = 1, 2, 3$) and the $\Delta\beta_i$ are corrections to the $\tilde{\beta}_i$ to be computed. This step is based on the assumption that the regression function is linear in the neighborhood of the estimated values of the parameters. For convenience, (14) may be written as

$$
Y = f \approx f_o + \sum_{i=1}^{3} f_i \Delta\beta_i \quad . \tag{15}
$$

This expression is linear in the unknown corrections, $\Delta\beta_i$, and therefore, under the appropriate assumptions, the method of least squares may be employed to estimate the corrections to the initial estimates of the β_i. The normal equations are obtained by minimizing the sum of squares of residuals given by

(16) $$Q(\tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3) = \sum_{ijk} \omega_{ij} (Y_{ijk} - f_o - \sum_{i=1}^{3} f_i \Delta\beta_i)^2$$

assuming that the f_i are continuous over the range of values of the independent variables, X_1 and X_2.

Then, instead of applying the entire correction to the $\tilde{\beta}_i$ as is done in the Gauss-Newton iteration, a fraction v of the correction is applied, where v is determined as follows.

Consider the sum of squares of residuals to be a function of v by defining it as

(17) $$Q(v) = Q(\tilde{\beta}_1 + v\Delta\beta_1, \tilde{\beta}_2 + v\Delta\beta_2, \tilde{\beta}_3 + \Delta\beta_3) \quad o \leq v \leq 1.$$

(The $\tilde{\beta}_i$, the initial estimates of the parameters and the $\Delta\beta_i$, the corrections to the estimates are known values, leaving only v unknown.) The value of v giving a minimum of $Q(v)$ is found by an approximate method by determining the level of v at which the parabola passing through $Q(0)$, $Q(\frac{1}{2})$, and $Q(1)$ has a minimum. Using the Lagrange method, the parabola passing through these points can be written as

(18) $$\Phi(v) = [2Q(0) - 4Q(\tfrac{1}{2}) + 2Q(1)]\, v^2 - [3Q(0) + 4Q(\tfrac{1}{2}) - Q(1)]\, v + Q(0) .$$

After differentiating $\Phi(v)$ with respect to v and setting the results equal to zero, the value of v giving a minimum of $\Phi(v)$ is found to be

(19) $$\text{min} = \frac{3Q(0) - 4Q(\frac{1}{2}) + Q(1)}{4[Q(0) - 2Q(\frac{1}{2}) + Q(1)]} = \frac{Q(0) - Q(1)}{4[Q(0) - 2Q(\frac{1}{2}) + Q(1)]} + \frac{1}{2} .$$

Using this value of v, the new estimates of the β_i to be used in computing the sum of squares of residuals and in the next iteration is computed from the expression

(20) $$\tilde{\beta}^*_i = \tilde{\beta}_i + v_{min}\, \Delta\beta_i .$$

The above procedure is repeated until the estimates yielding the minimum sum of squares of residuals is obtained.

As indicated in [5], in the event that the value of v_{min} does not give a reduction in the sum of squares of residuals from one iteration to the next, the value of v yielding a minimum of Φ (v) in an interval of half the length should be used to compute the new estimates.

Because the Hartley modification requires initial estimates that are in the neighborhood of those yielding the absolute minimum, and since this procedure was to be used as a check, initial estimates of the parameters were selected in the neighborhood of the final estimates obtained in the scanning process. Table II gives the least squares estimates of the parameters of (3) and the sum of squares of residuals obtained in the final iteration of the Hartley modification.

CONSTRUCTION OF CONFIDENCE INTERVALS. Another procedure, presented in [1], yields least squares estimates of the parameters as well as an estimate of the variance-covariance matrix of the least squares estimates, which can be used to construct approximate confidence intervals about the individual parameters. This procedure is based on the Fisher information matrix as described in [9].

Corrections to initial estimates, $\tilde{\beta}_i$, are derived by expanding the normal equations in Taylor's series about the initial estimates, obtaining expressions of the form

$$(21) \qquad \frac{\partial Q(\beta)}{\partial \beta_i} = 0 \simeq \frac{\partial Q(\beta)}{\partial \tilde{\beta}_i} + \sum_{i=1}^{3} \Delta\beta_i \frac{\partial^2 Q(\beta)}{\partial \tilde{\beta}_i \, \partial \tilde{\beta}_j}, \quad (i = 1, 2, 3)$$

where $\Delta\beta_i = \beta_i - \tilde{\beta}_i$ and $Q(\beta)$ is the sum of squares of residuals. From this, the set of normal equations can be written in matrix notation as

$$(22) \qquad V \, \Delta\beta = G$$

where V is the Fisher information matrix with elements

$$(23) \qquad I^{ij} = - E \frac{\partial^2 Q(\beta)}{\partial \beta_i \, \partial \beta_j} = - \frac{\partial^2 Q(\beta)}{\partial \beta_i \, \partial \beta_j}, \quad (i, j = 1, 2, 3) .$$

$\Delta\beta$ is a column vector with components $\Delta\beta_i$, and 2G is a column vector with components $\frac{\partial Q(\beta)}{\partial \tilde{\beta}_i}$. Solution of (22) yields the corrections which are to be applied to the initial estimates to obtain estimates to be used in the next iteration. That is,

$$\Delta\beta = V^{-1} G \tag{24}$$

and

$$\beta^1 = \tilde{\beta} + \Delta\beta = \tilde{\beta} + V^{-1} G \tag{25}$$

where β^1 is the vector of estimates to be used in the next iteration and $\tilde{\beta}$ is the vector of initial estimates.

When the process has converged to the least squares estimates, the matrix $V^{-1}\hat{\sigma}^2$ is an approximation to the variance-covariance matrix of the least squares estimates. Using this approximation, it is possible to construct confidence intervals about the individual parameters such that

$$P\left\{\begin{array}{l} \hat{\beta}_1 - \hat{\sigma}\sqrt{3v^{11} F_\alpha\ (3,\ n-3)} \le \beta_1 \le \hat{\beta}_1 + \hat{\sigma}\sqrt{3v^{11} F_\alpha\ (3,\ n-3)} \\ \hat{\beta}_2 - \hat{\sigma}\sqrt{3v^{22} F_\alpha\ (3,\ n-3)} \le \beta_2 \le \hat{\beta}_2 + \hat{\sigma}\sqrt{3v^{22} F_\alpha\ (3,\ n-3)} \\ \hat{\beta}_3 - \hat{\sigma}\sqrt{3v^{33} F_\alpha\ (3,\ n-3)} \le \beta_3 \le \hat{\beta}_3 + \hat{\sigma}\sqrt{3v^{33} F_\alpha\ (3,\ n-3)} \end{array}\right\} \ge 1 - \alpha \tag{26}$$

where V^{ii} (i = 1, 2, 3) is the diagonal element of the i^{th} row of V^{-1}, F_α (3, n-3) is the α percentile of the F distribution with 3 and n-3 degrees of freedom and

$$\hat{\sigma} = \sqrt{\frac{Q(\hat{\beta})}{n-3}} \quad . \tag{27}$$

Ninety percent confidence interval estimates based on the estimates obtained in the scanning process and those obtained by the Hartley modification are given in Table III. A combination of the Hartley modification and the procedure described in this section yielded estimates that gave a slightly smaller sum of squares of residuals in each of the three zones.

Point estimates and 90% confidence interval estimates based on the combined procedure are given in Table IV.

DISCUSSION OF RESULTS. It can be seen by inspection ot Tables I and II, which give the estimates obtained in fitting the regression function by the scanning procedure and by the Hartley modification of the Gauss-Newton int eration, that the results of the two procedures do not differ to any great degree. In general, the estimates of β_3 obtained by the two processes differ more than the estimates of the other two parameters, especially in Zone II. However, it is pointed out that the error root mean squares in the neighborhood of the apparent minimum are less sensitive to small (positive) changes in this estimate than in those for the other two parameters. This can be seen from Figure 1. In addition, in examination of the error root mean squares in the various iterations of the Hartley modification, it was noted that a difference of as much as four in the estimates of β_3 in the neighborhood of the minimum caused a change of only 0.01 in the error root mean squares.

Further examination of the estimates presented in Tables I and II reveals that for each zone, the estimates of β_2 are relatively small in comparison to the estimates of β_1 and β_3. This may lead one to think that this parameter does not contribute significantly to the model and may be eliminated from consideration. In fact, tests of hypotheses based on the assumption that the statistic

$$(28) \qquad t = \frac{\hat{\beta}_1 - 0}{\hat{\sigma}_{\hat{\beta}_i}} = \frac{\hat{\beta}_i}{\hat{\sigma}^2 V^{ii}}$$

is distributed as "Student's" t distribution, indicated that the hypothesis that $\beta_2 = 0$ would be accepted in Zones I and IIω and the hypothesis that $\beta_1 = 0$ would be accepted in Zone III at the .05 level of significance. On the basis of these results, the model equations for the various zones could be as indicated below.

	Zone	Model Equation
(29)	I & II	$Y_{ijk} = \dfrac{\gamma_1 X_{1i}}{X_{2j} + \gamma_2} + \epsilon_{ijk}$
(30)	III	$Y_{ijk} = \dfrac{\gamma_1 X_{1i} X_{2j}}{X_{2j} + \gamma_2} + \epsilon_{ijk}$

Each of the two model equations above have properties that are similar to those of the original model in that the regression functions determined from these models represent straight lines passing through the origin when X_{2i} is constant and rectangular hyperbolas with horizontal and vertical asymptotes when X_{1i} is constant. These equations would be much more suitable for use on FADAC than the original model equations.

Figures 2, 3, and 4 give perspective sketches of the general shapes of the surfaces represented by the estimated regression functions. Sketches of the constant temperature and constant set time contour lines are given in Figures 5, 6, and 7. To indicate how well the curves fit the data, Figures 8, 9 and 10 give sketches of the constant set time curves (for selected time settings) with the data points plotted. Similar sketches for selected constant temperatures are given in Figures 11, 12, and 13.

It can be seen from the last two sets of sketches that the variability of the observed biases was relatively large at low temperatures for the giventime settings; however, there was little difference in the variability at the various time settings for a given temperature. It is also easy to see that the assumption that the effect of temperature on bias is relatively constant is not a bad assumption for temperatures slightly above zero degree Fahrenheit.

On the basis of the amount of information obtainable from the procedures discussed in the foregone sections, it appears that the best procedure is a combination of the Hartley modification and the procedure utilizing the Fisher Information matrix. Point estimates of the regression parameters and 90% confidence intervals obtained by this combined procedure are given in Table III. The combined procedure gave sums of squares of residuals that were slightly less than those obtained by either of the other two procedures, although, due to slightly larger estimates of the $\sigma_{\hat{\beta}_i}$, the confidence intervals obtained for this method were generally longer than those for the other two procedures.

ACKNOWLEDGMENTS. The author wishes to express his gratitude to Messrs. James R. Kniss and Robert M. Eissner for their suggestions and critical review of this report. Immeasurable assistance was rendered by the members of the Nuclear Ammunition Section of the Surveillance and Reliability Laboratory. Messrs. Jerry Thomas and Warren S. Wenger adapted the procedures discussed in Section I and II to computer programs. Mr. Sidney Gerard also assisted in programming various aspects of the problem for the computer. Mr. Emerson Jackson drew the sketches presented in the report. The model equation was arrived at after deliberations with Mr. Norbert Digiacomo of the Firing Tables Branch of the Computing Laboratory.

REFERENCES

1. Beale, E. M. L., Confidence Regions in Non-Linear Estimation, J. Roy. Stat. Soc., B-22, 1960.

2. Davies, O. L., The Design and Analysis of Industrial Experiments, Oliver and Boyd, Ltd., London and Edinburgh, 1956.

3. Draper, N. R. and Smith, H., Applied Regression Analysis, John Wiley and Sons, Inc., New York, 1966.

4. Guest, P. G., Numerical Methods of Curve Fitting, Cambridge at the University Press, 1961.

5. Hartley, H. O., The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares, Technometrics, Vol. 3, No. 2, May 1961.

6. Levenberg, K., A Method for the Solution of Certain Non-Linear Problems in Least Squares, Quart. Appl. Math., 2, 1944.

7. Marquardt, D. W., An Algorithm for Least-Squares Estimation of Non-Linear Parameters, J. Soc. Indust. Appl. Math. Vol. II, No. 2, June 1963.

8. Natrella, Mary G., Experimental Statistics, Section I, Basic Concepts and Analysis of Measured Data, ORDP 20-110, Ordnance Engineering Handbook Officer, Durham, N. C., 1962.

9. Rao, C. R., Advanced Methods in Biometric Research, John Wiley and Sons, Inc., New York, 1952.

TABLE I

Estimates Using the Scanning Procedure

Zone	No. Obs. Considered	Estimates of Regression Parameters $\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	Sum of Squares of Residuals
I	171	0.5604	0.0023	54.62	15.0688
II	96	0.6392	-0.0021	64.83	9.3593
III	217	0.0245	-0.0054	51.46	38.0889

TABLE II

Estimates Using the Hartley Modification

Zone	No. Obs. Considered	Estimates of Regression Parameters			Sum of Squares of Residuals
		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	
I	171	0.5585	0.0023	54.31	15.0676
II	96	0.6644	-0.0023	67.16	9.3068
III	217	0.0284	-0.0056	50.67	37.8275

TABLE III

Ninety Percent Confidence Intervals about Regression Parameters

Method of Estimation	Zone	No. of Observations Considered	90% Confidence Limits on -					
			β_1		β_2		β_3	
			Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit
Scanning Process	I	171	0.3746	0.7462	-0.0013	0.0059	45.46	63.78
	II	96	0.4029	0.8755	-0.0073	0.0031	50.74	78.92
	III	217	-0.0476	0.0966	-0.0083	-0.0025	43.15	59.77
Hartley Modification	I	171	0.3792	0.7377	-0.0012	0.0058	45.75	62.87
	II	96	0.3944	0.9344	-0.0078	0.0032	50.04	84.28
	III	217	-0.0349	0.0917	-0.0078	-0.0034	46.39	54.95

TABLE IV

Point Estimates and 90% Confidence Interval Estimates*

Zone	No. of Observations Considered	Sum of Squares of Residuals	β_1			β_2			β_3		
			Point Estimate	Lower Limit	Upper Limit	Point Estimate	Lower Limit	Upper Limit	Point Estimate	Lower Limit	Upper Limit
I	171	15.0165	0.6239	0.4035	0.8443	0.0021	-0.0017	0.0059	55.67	46.19	67.15
II	96	9.2150	0.6824	0.4001	0.9647	-0.0010	-0.0063	0.0043	66.03	48.64	83.42
III	217	36.9716	0.0510	-0.0134	0.1154	-0.0050	-0.0072	-0.0027	50.47	46.17	54.78

* Estimates presented in this Table were obtained by a combination of the procedures described in Sections II and III

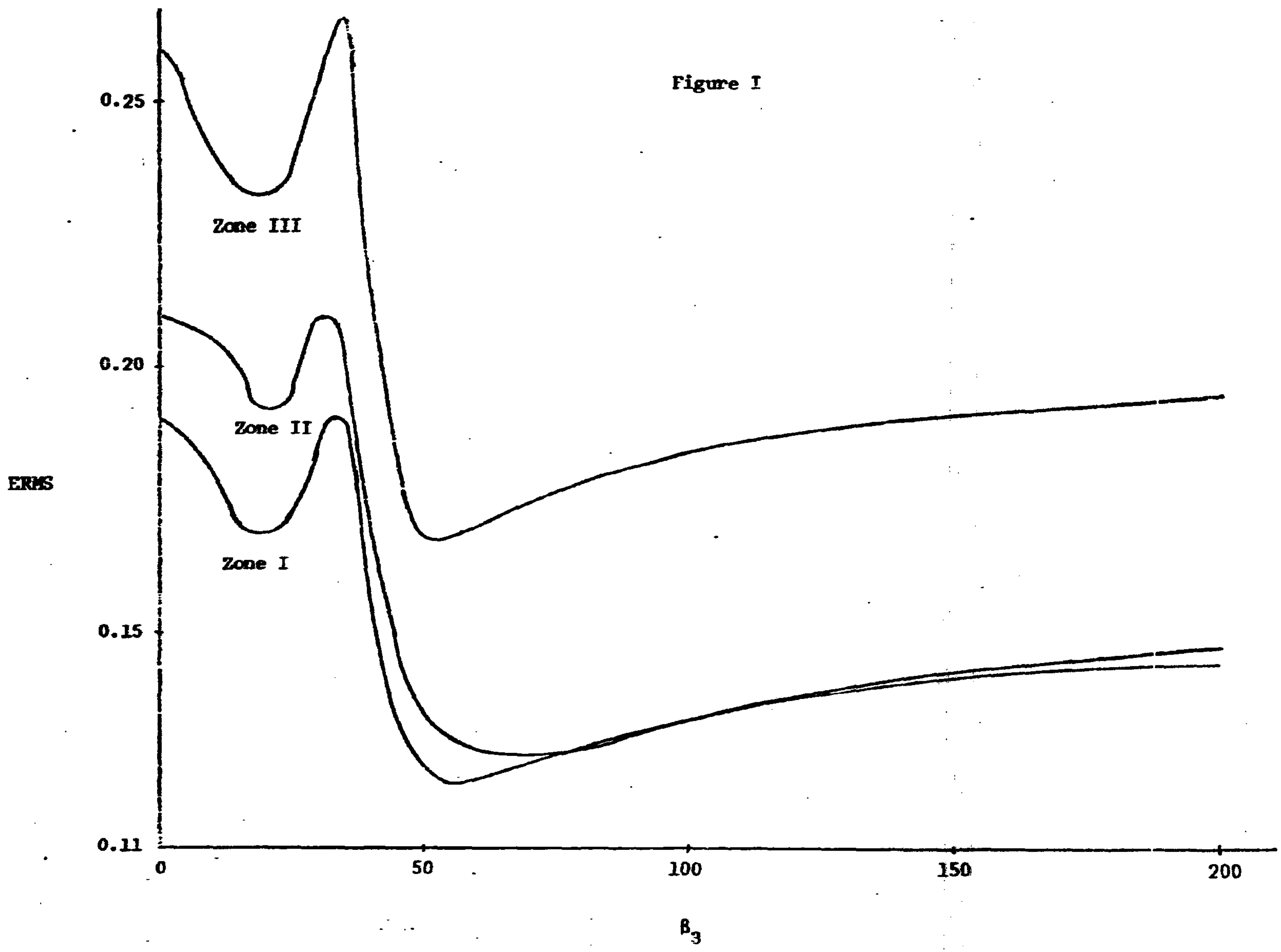

0.25
0.20
0.15
0.11
ERMS
Zone III
Zone II
Zone I
0
50
100
150
200
β_3

Figure I

Figure 2

Zone I

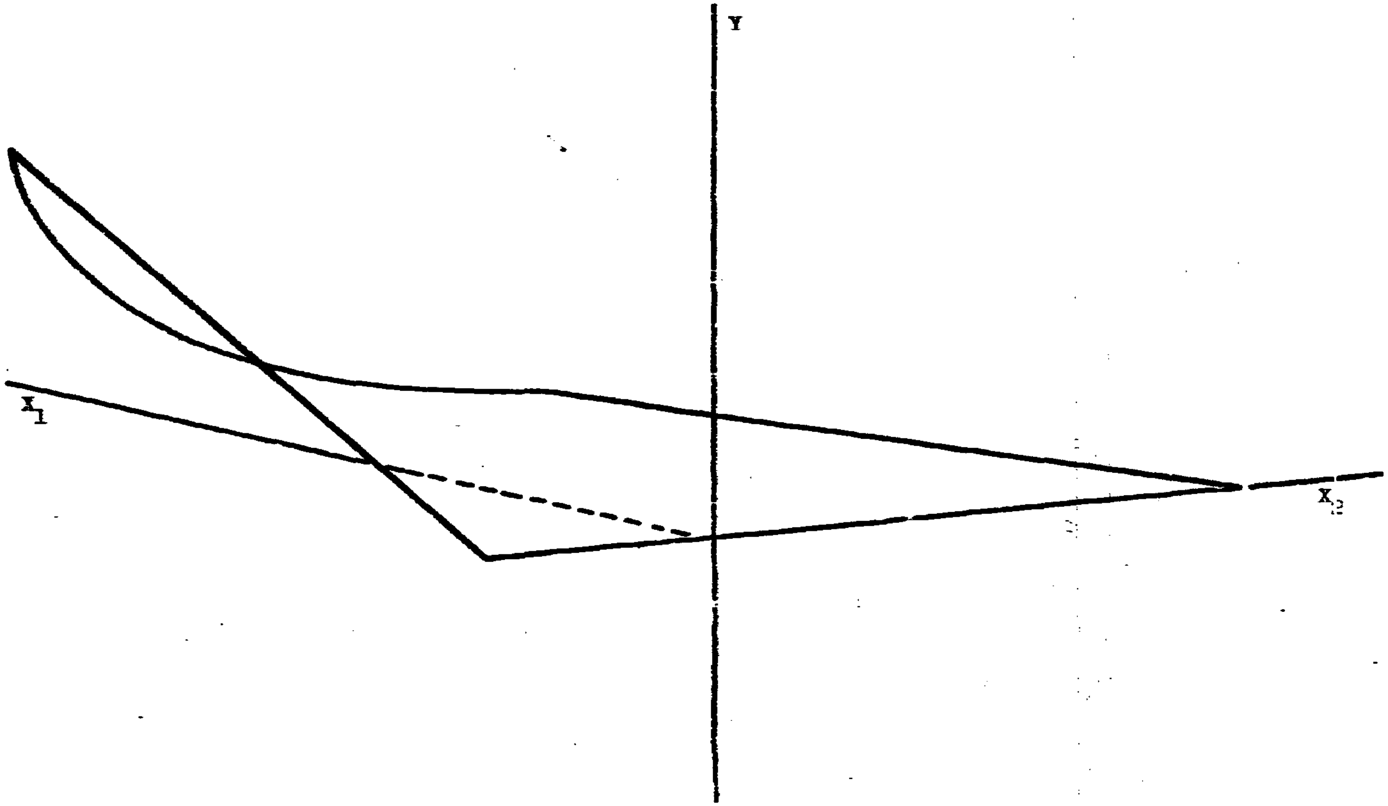


Figure 3

Zone II

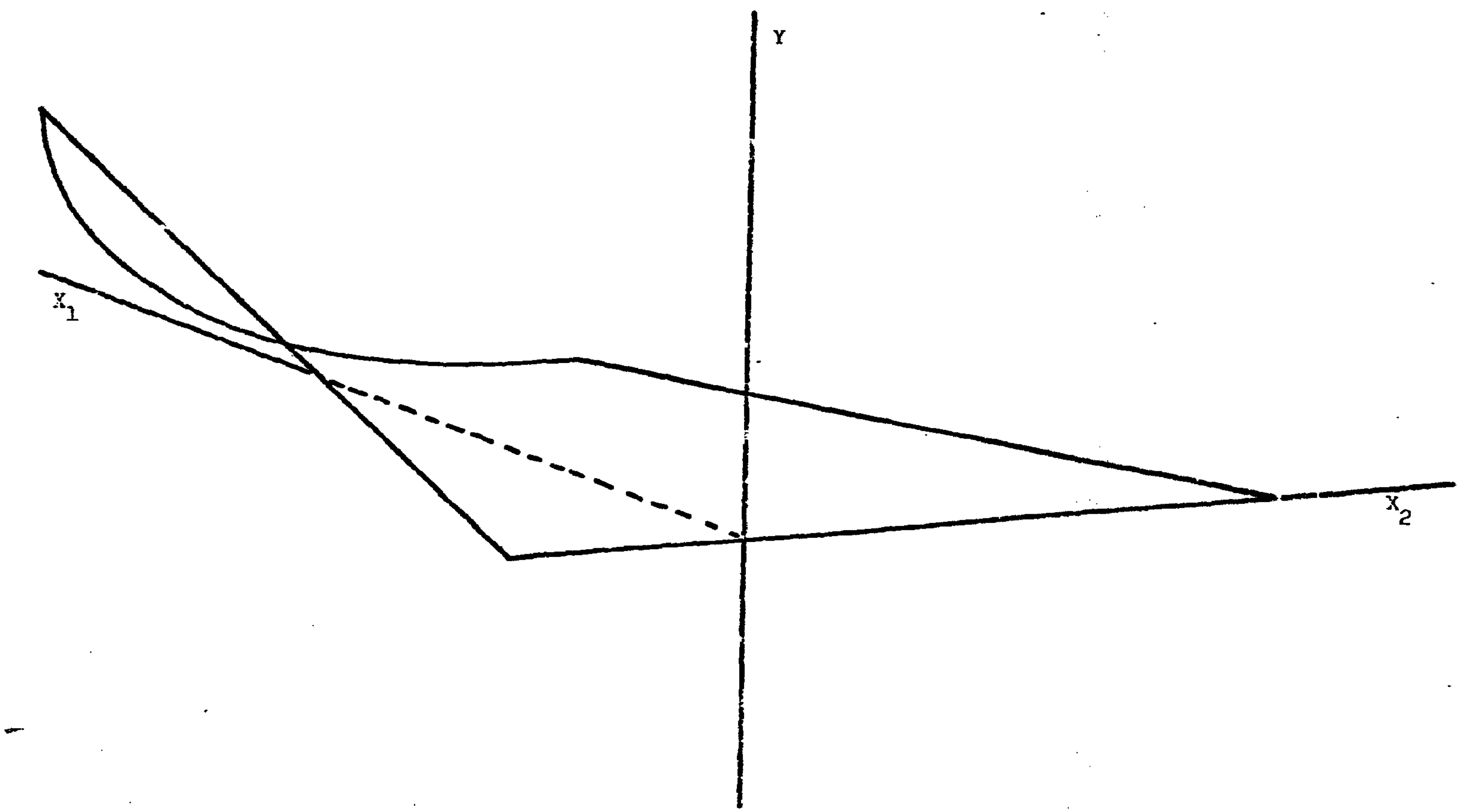


Figure 4

Zone III

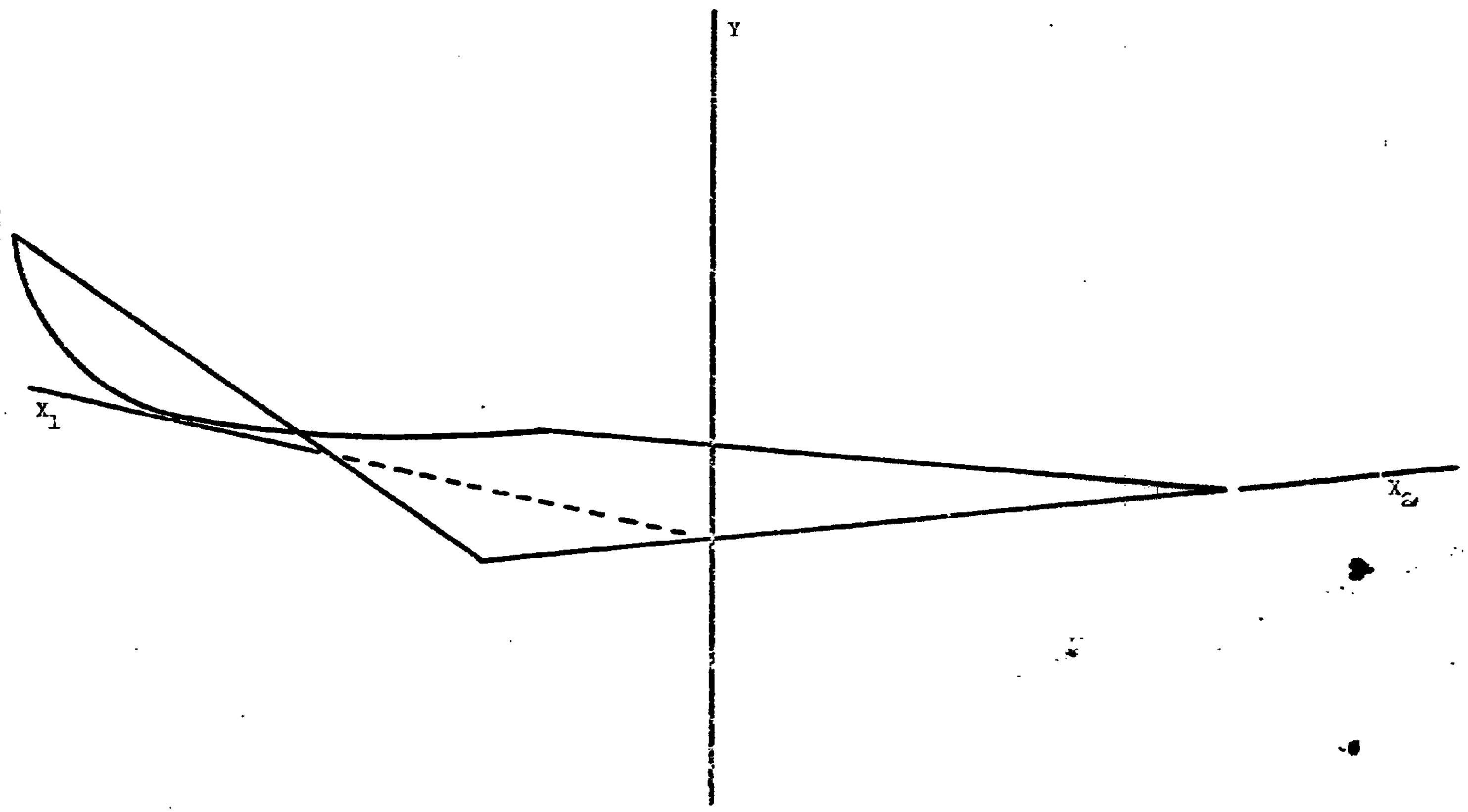


Figure 5

Zone I

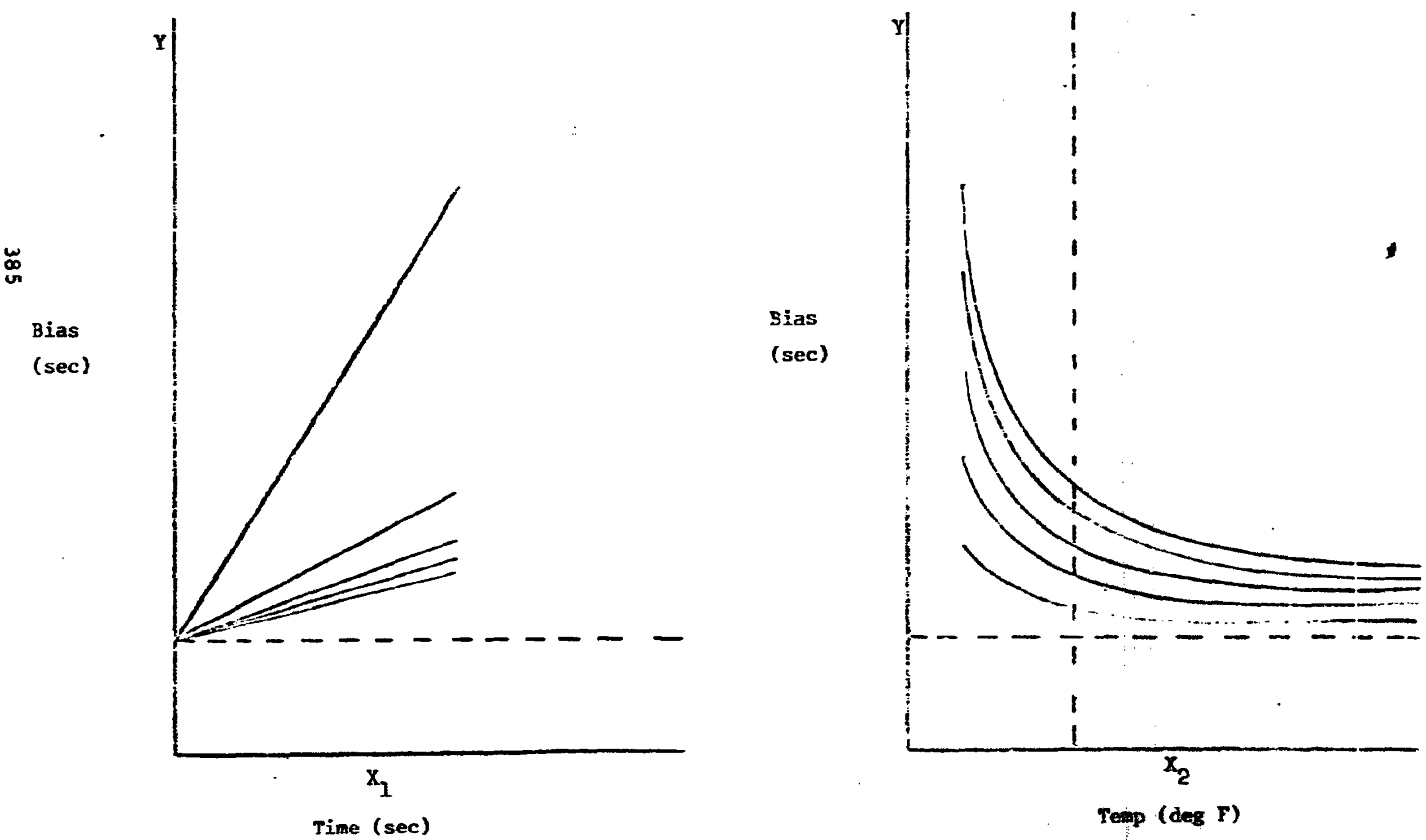


Figure 6

Zone II

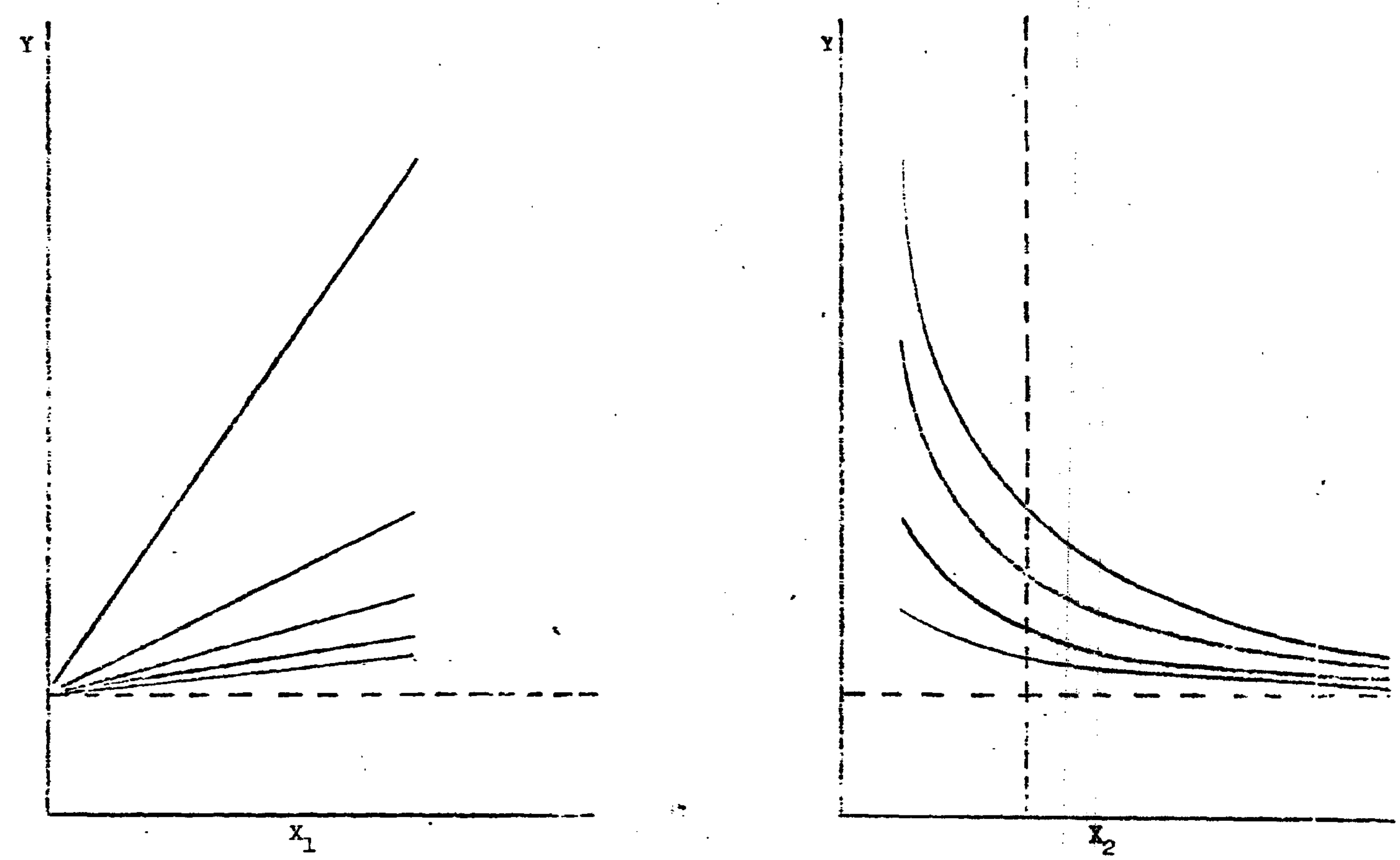


Figure 7

Zone III

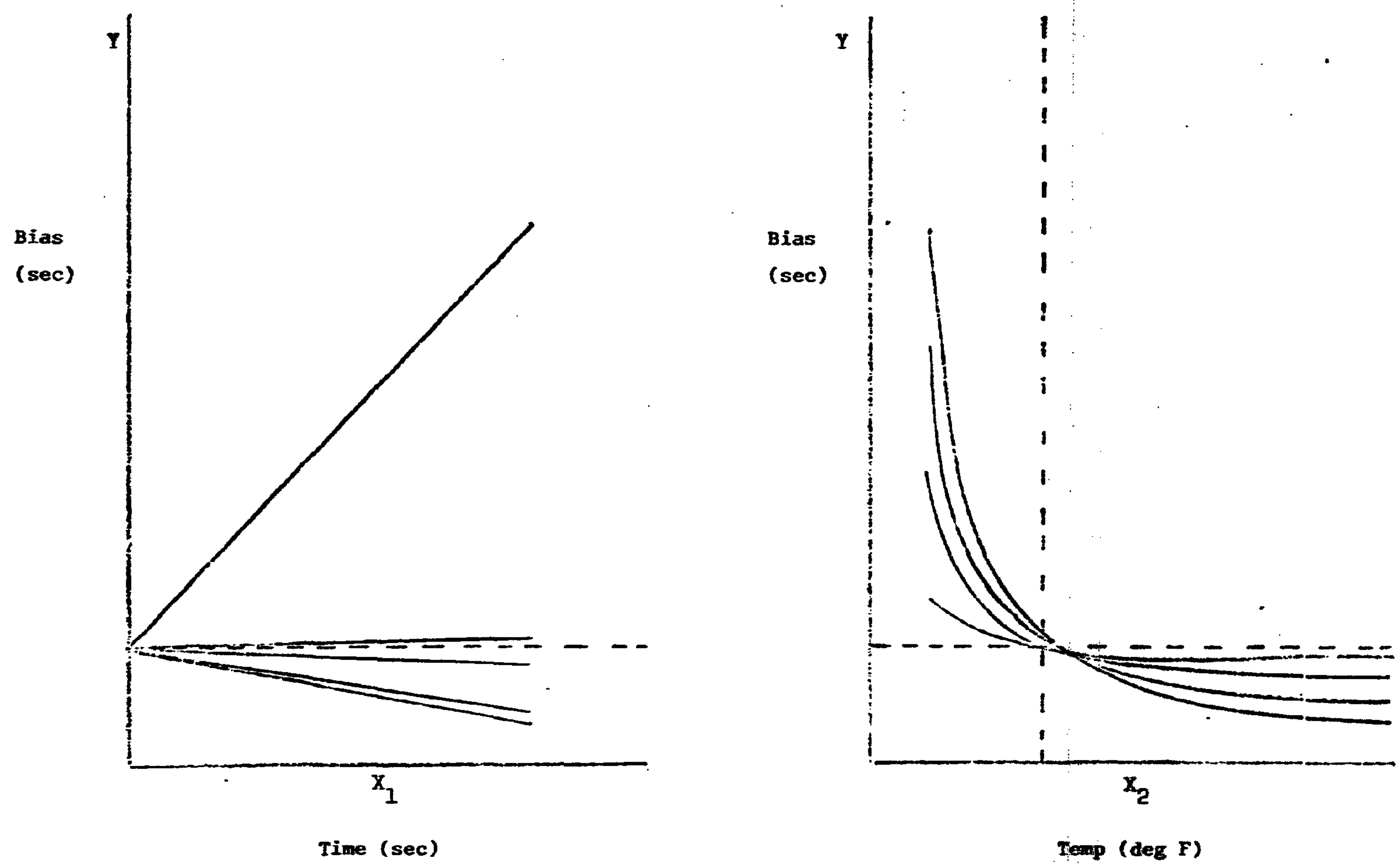


Figure 8

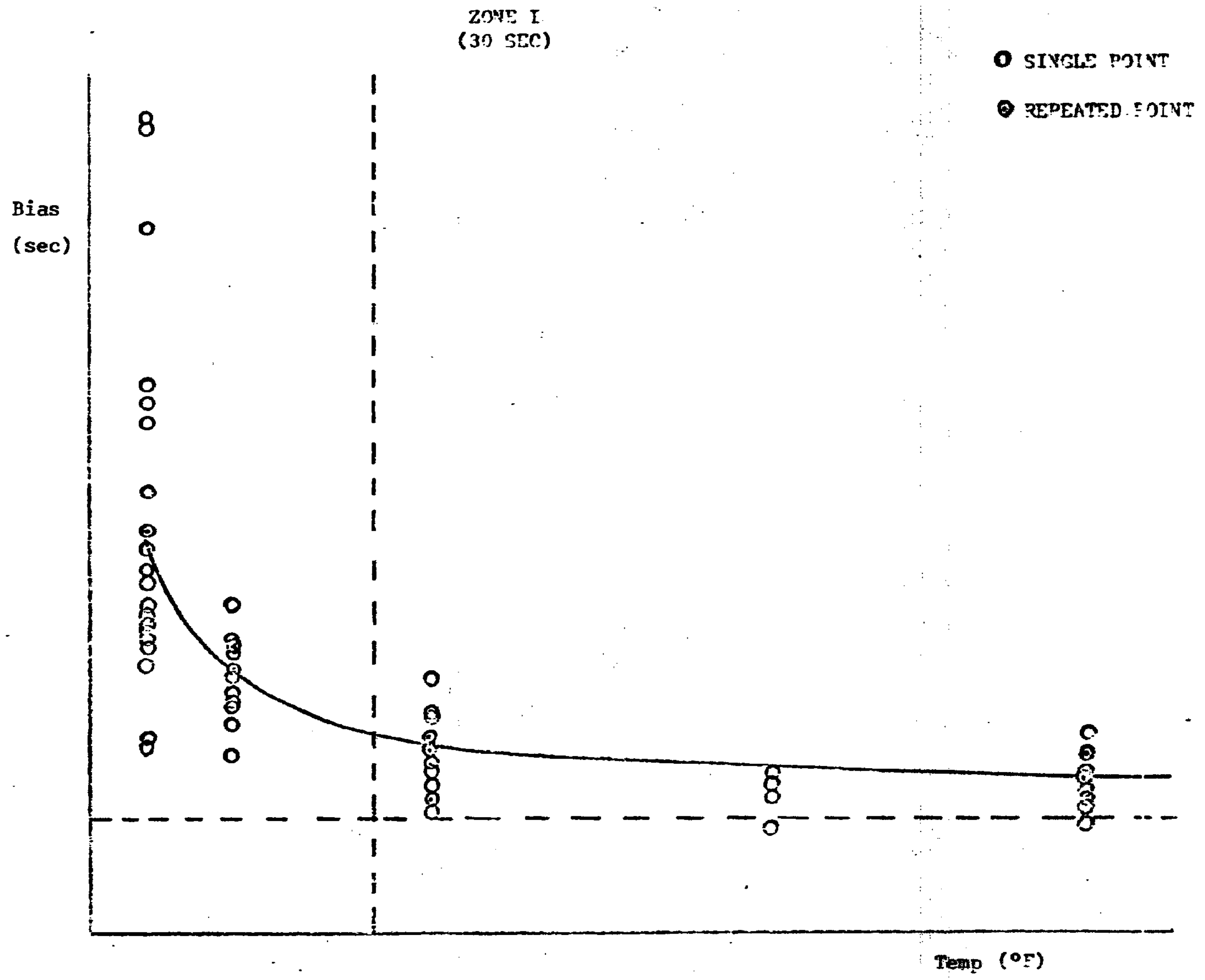


Figure 9

ZONE II
(55 SEC)

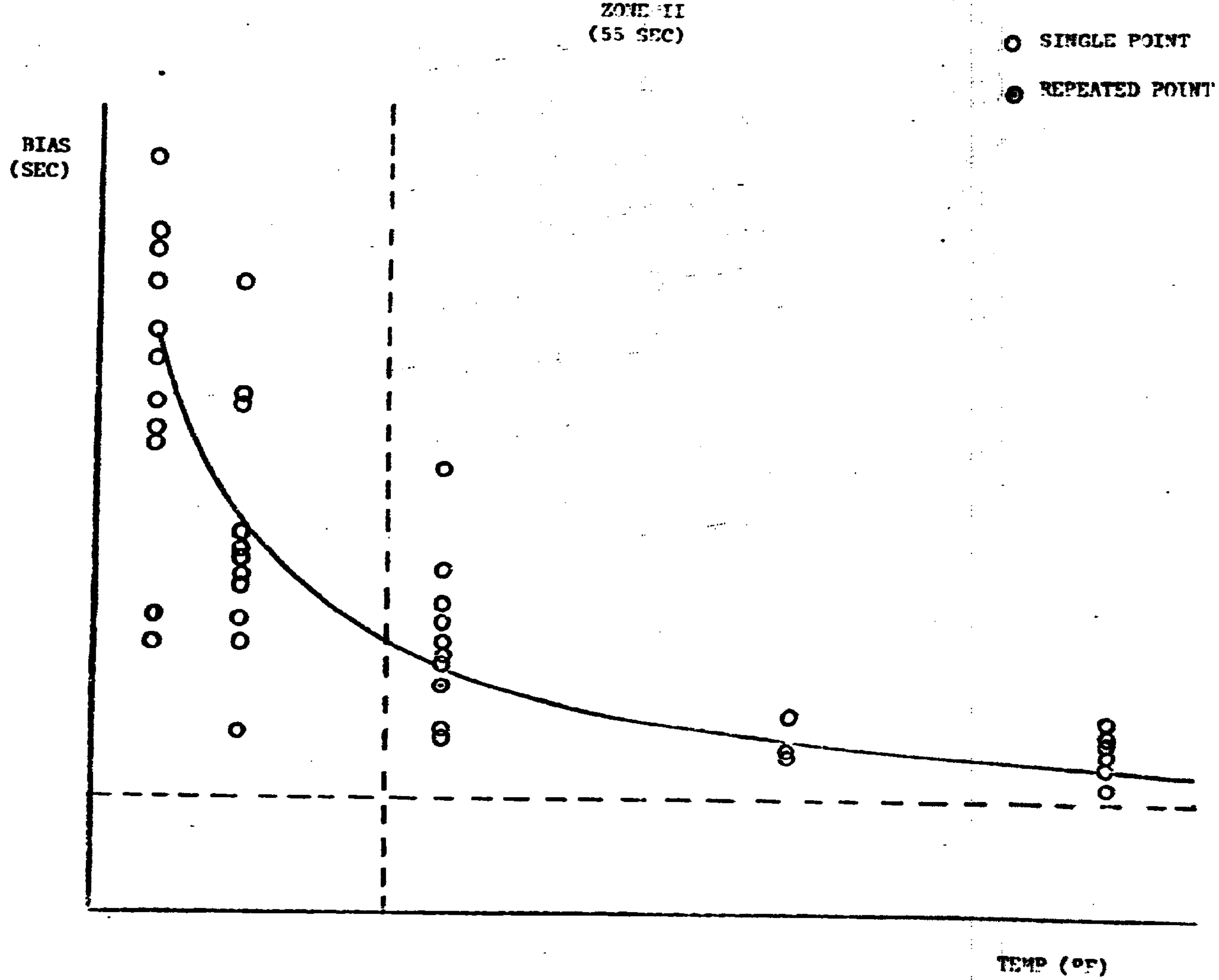


Figure 10

ZONE III
(60 SEC)

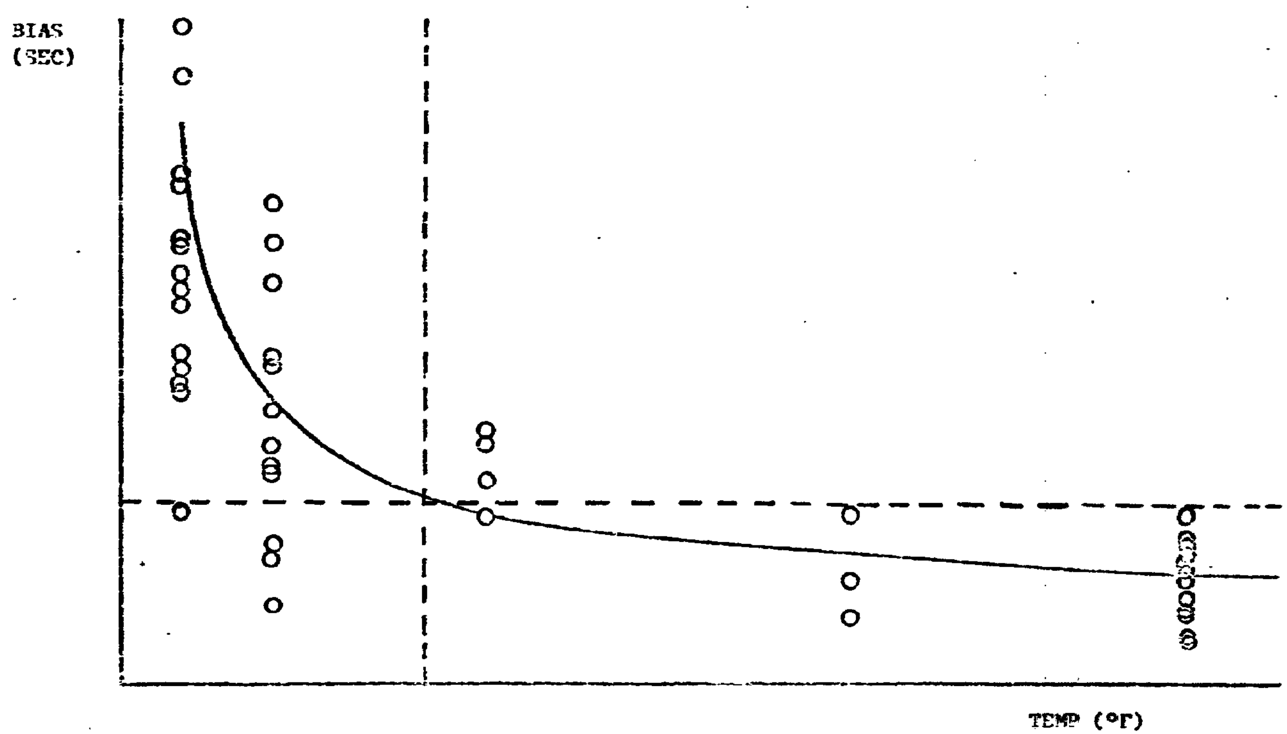


Figure 11

ZONE I
(10°F)

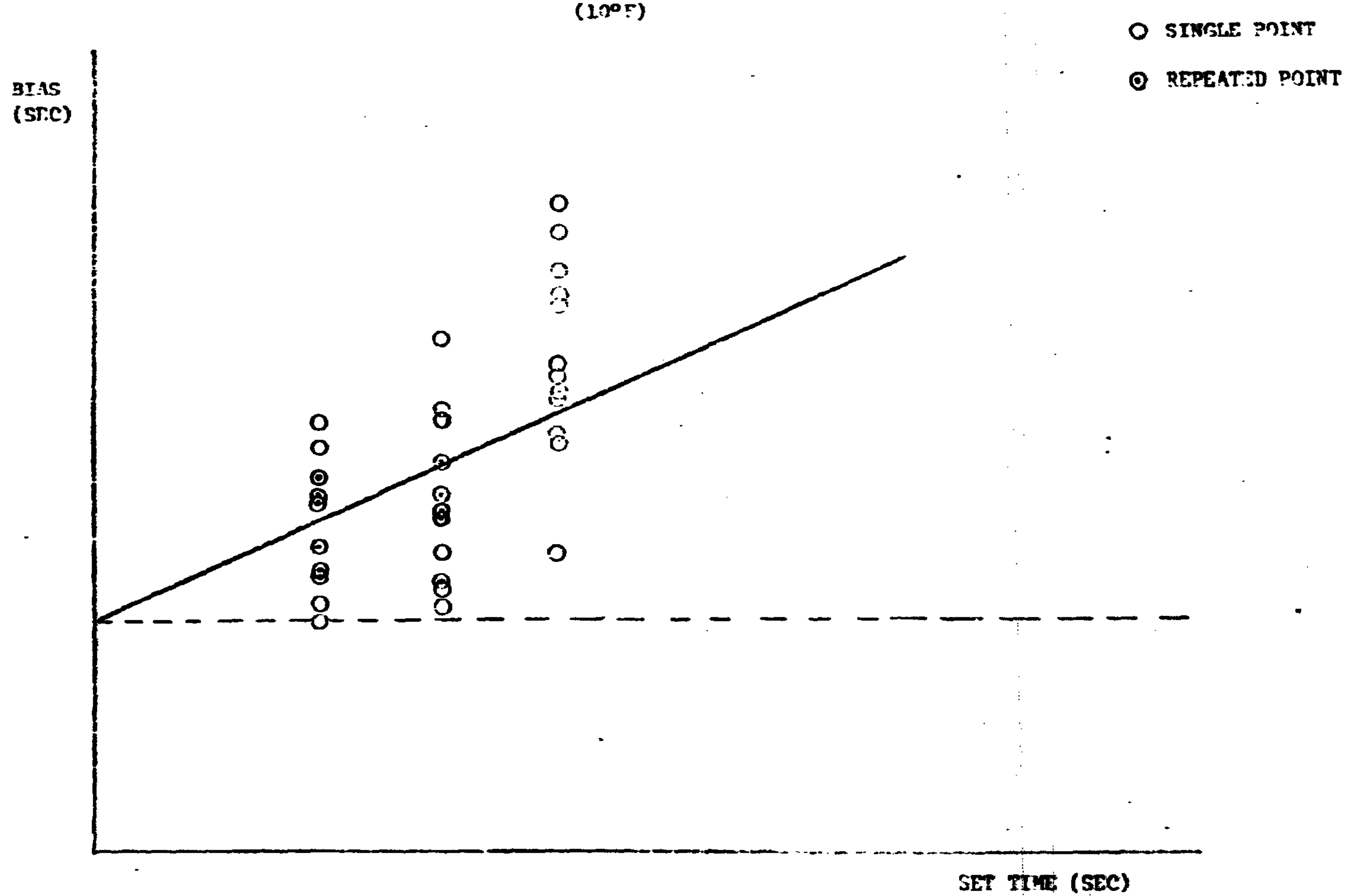


Figure 12

ZONE II
(10°F)

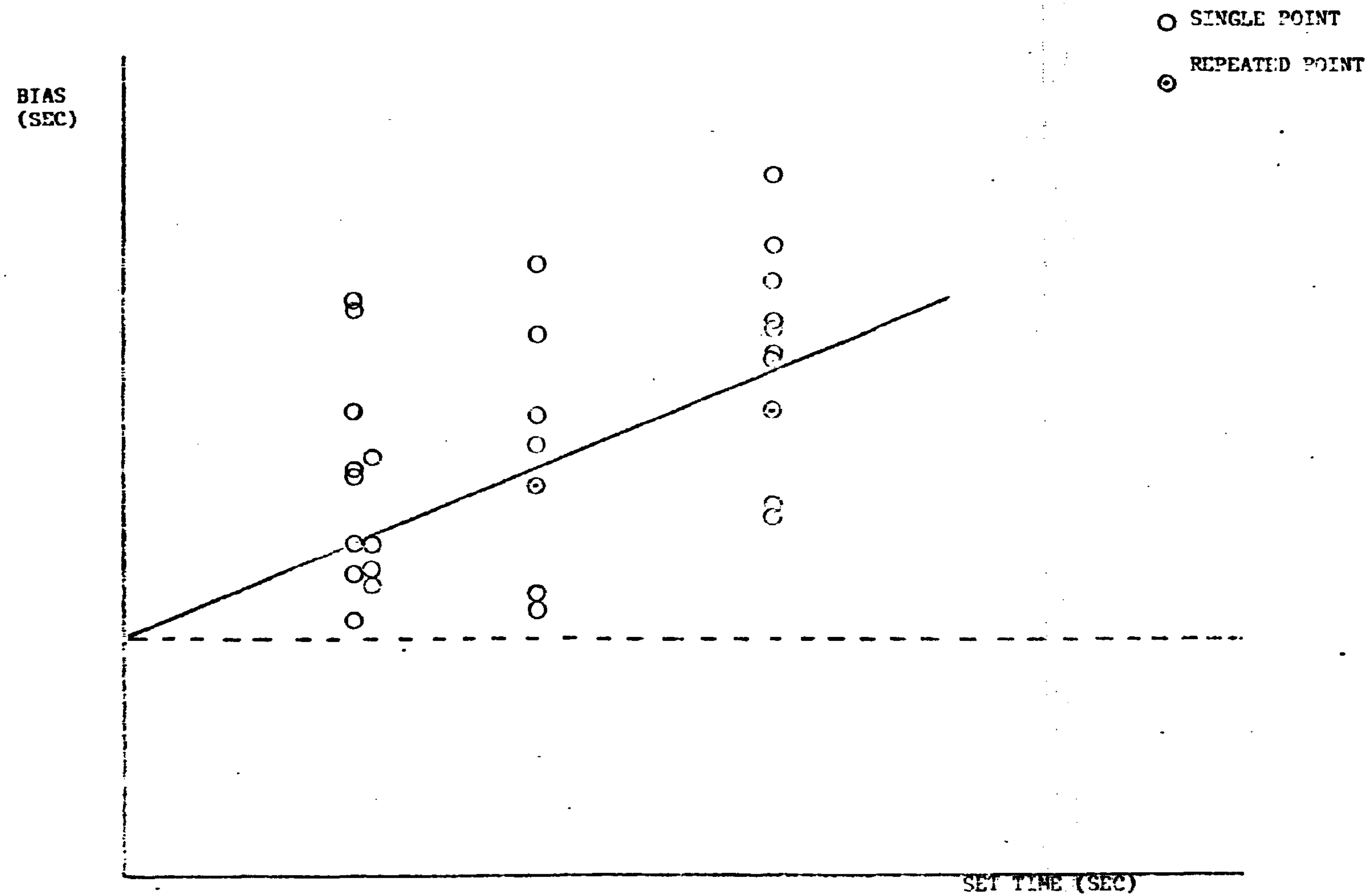
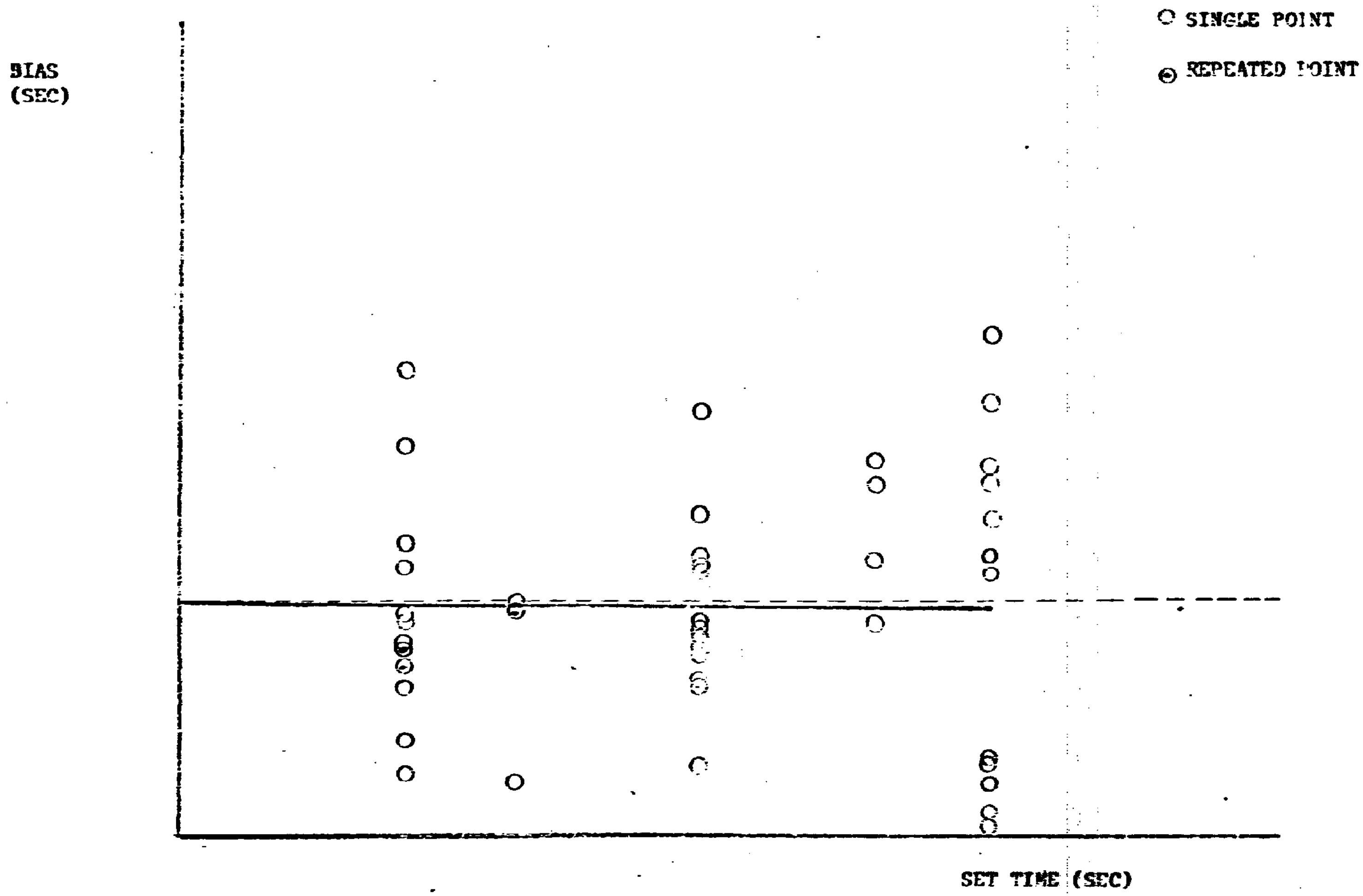


Figure 13

ZONE III
(10°F)



OBSERVATIONS ON THE SELECTION OF PREDICTORS

H. L. Lucas and A. C. Linnerud
North Carolina State University at Raleigh

1. INTRODUCTION AND SUMMARY. Most work on the selection of predictors has been done in the context of the general linear model,

$$(1) \qquad \underline{y} = [\underline{x}_1\, \underline{x}_2 \cdots \underline{x}_p\, \underline{\epsilon}]\ [\beta_1\, \beta_2 \cdots \beta_p\ 1]' ,$$

where $\underline{y}$ (observations on the predictand), $\underline{x}_i$ (observations on the ith predictor) and $\underline{\epsilon}$ (random residuals) are all $n \times 1$ matrices, the β_i (regression coefficients) are scalars, and the prime means transpose. For one class of practical situations it can be assumed that observations have been made on all predictors that are relevant (corresponding β's non-zero) and possibly on some that are not (corresponding β's zero). Given a set of data, the problem is to decide which one or more of the 2^p subsets of predictors is or are likely to be the correct one. In the present study, attention has been confined mainly to selection of the single best candidate.

Three criteria of selection, namely, the residual mean square (MS) the Mallows C-statistic (C) and a modification of the C-statistic (MC), have been studied to date. It was assumed that a set of data is a random sample from a population characterized by certain values of the β_i and by the form and the parameters of the joint distribution of the $\underline{x}_i$ and $\underline{\epsilon}$. The $\underline{x}_i$ were assumed to be measured without error. Performance was studied in terms of the probability that a criterion leads to selection of the correct subset of predictors.

Since the mathematics has appeared to be intractable, a highspeed computer has been used to study the problem empirically. So far, data have been obtained only on cases with $p = 3$, the $\underline{x}_i$ and $\underline{\epsilon}$ jointly normally distributed and

$$(2) \qquad E\begin{bmatrix}\underline{x}_1\\ \underline{x}_2\\ \underline{x}_3\\ \underline{\epsilon}\end{bmatrix} = \begin{bmatrix}\underline{1}\\ \underline{0}\\ \underline{0}\\ \underline{0}\end{bmatrix}, \qquad V\begin{bmatrix}\underline{x}_1\\ \underline{x}_2\\ \underline{x}_3\\ \underline{\epsilon}\end{bmatrix} = \begin{bmatrix}\underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0}\\ \underset{\sim}{0} & I & \rho I & \underset{\sim}{0}\\ \underset{\sim}{0} & \rho I & I & \underset{\sim}{0}\\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & I\end{bmatrix},$$

where $\underline{1}$ $(n \times 1)$ has elements all 1, $\underline{0}$ $(n \times 1)$ and $\underset{\sim}{0}$ $(n \times n)$ have elements all zero, I $(n \times n)$ is an identity matrix and ρ (1×1) is the correlation between $\underline{x}_2$ and $\underline{x}_3$.

Using sample size $n = 20$, 100-200 samples were drawn for each of several combinations of ρ (0.00, 0.95) and β-sets having different numbers (0, 1, 2, 3) and magnitudes (1, 2, 5) of non-zero elements. Although the results exhibited many qualitative and some quantitative features which were not unanticipated, the quantitative features were in general pleasantly surprising to the author. All three criteria were better for selecting the correct subset of predictors than was expected on the basis of some approximate and apparently naive preliminary considerations. This was particularly true for the cases, all β's zero or the non-zero β's small. For most situations studied, C was superior to MC and MC superior to MS. Exceptions occurred particularly when the magnitudes of β_1 and β_2 were small but both non-zero, and the correlation between $\underline{x}_2$ and $\underline{x}_3$ was .95. Exclusive of these exceptions, criterion C resulted in 57 - 100% selection of the correct subset of predictors, and MC and MS resulted in 30 - 100% correct choice. For the exceptions noted, however, good practical performance of the criteria was still obtained. The subset of predictors selected as best simply alternated between including $\underline{x}_2$ and $\underline{x}_3$ rather than both.

2. BACKGROUND.

2.1. General orientation:

There is a point which needs to be emphasized before focusing on the immediate setting of the results to be presented. It is this. Given a vector of observations on a predictand $\underline{u}$ and potential predictors $\underline{z}_j$ ($j = 1, 2, \ldots, \pi$), one should consider any theory and reasonable supposition regarding the nature of relationship of $\underline{u}$ to the $\underline{z}_j$. It is often not sensible to assume that $\underline{u} = \underline{y}$ of (1) and that $z_j = \underline{x}_i$ of (1), although this is too often done. Rather, it may be proper to let $\underline{y} = \underline{\eta}(\underline{u})$ and $\underline{x}_i = \underline{\phi}_i(\underline{z}_i, \underline{z}_2, \ldots, \underline{z}_\pi)$. Consideration here will be confined to situations such that transformations of the observations on predictand and predictors permit expression of the relationship as in (1).

It will be assumed that a practical situation can be characterized by certain values of the β_i and by the form and the parameters of the joint distribution of the $\underline{x}_i$ and $\underline{\epsilon}$. More specifically, it will be assumed to start that

(3)
$$E\begin{bmatrix}\underline{x}_1\\ \underline{x}_2\\ \underline{x}_3\\ \cdot\\ \cdot\\ \cdot\\ \underline{x}_p\\ \underline{\epsilon}\end{bmatrix} = \begin{bmatrix}\underline{1}\\ \underline{0}\\ \underline{0}\\ \cdot\\ \cdot\\ \cdot\\ \underline{0}\\ \underline{0}\end{bmatrix}, \quad V\begin{bmatrix}\underline{x}_1\\ \underline{x}_2\\ \underline{x}_3\\ \cdot\\ \cdot\\ \cdot\\ \underline{x}_p\\ \underline{\epsilon}\end{bmatrix} = \begin{bmatrix}\underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} & \underset{\sim}{0}\\ \underset{\sim}{0} & I & \rho_{23}I & \cdots & \rho_{2p}I & \underset{\sim}{0}\\ \underset{\sim}{0} & \rho_{23}I & I & \cdots & \rho_{3p}I & \underset{\sim}{0}\\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot\\ \underset{\sim}{0} & \rho_{2p}I & \rho_{3p}I & \cdots & I & \underset{\sim}{0}\\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} & I\end{bmatrix}.$$

A given set of data then represents a sample from the $\underline{x}_i$, $\underline{\epsilon}$ population which, with the β_i, implies $\underline{y}$.

The problem is to examine the performance of the aforementioned criteria for deciding on which of the 2^p subsets of the $\underline{x}_i$ is the relevant subset of predictors (i.e., which of the 2^p subsets of the β_i consists of only and all the non-zero β_i).

2.2. The selection criteria:

Let $v = 1, 2, 3, \ldots, 2^p$ index the subsets of the $\underline{x}_i$. Then upon rearrangement of the columns of $[\underline{x}_1\ \underline{x}_2 \ldots \underline{x}_p\ \underline{\epsilon}]$ of (1) and the elements of $[\beta_1\ \beta_2 \ldots \beta_p\ 1]'$, we can write

(4) $$\underline{y} = [Z_v\ X_v\ \underline{\epsilon}]\ [\underline{\theta}_v\ \underline{\beta}_v'\ 1]'$$

where the columns of Z_v $(n \times q_v)$ consist of the vth subset of the $\underline{x}_i$, those of X_v $(n \times \overline{p-q_v})$ the remaining $\underline{x}_i$, $\underline{\theta}_v$ and $\underline{\beta}_v$ consist of the correspondingly rearranged β_i, and $q_v = 0, 1, 2, \ldots, p$. Then, assuming Z_v, X_v to be of full rank, the total sum of squares $T = \underline{y}'\,\underline{y}$ can be partitioned into $S_v = (Z_v' Z_v)^{-1} Z_v'\underline{y}$ and $R_v = T - S_v$, and R_v carries $n - q_v$ degrees of freedom. It is useful to note that

(5) $$E(R_v) = \underline{\beta}_v' A_v \underline{\beta}_v + (n - q_v),$$

where $A_v = X_v' [I - Z_v (Z_v' Z_v)^{-1} Z_v'] X_v$.

The criteria compared were

$$MS_v = R_v/(n - q_v)$$

$$C_v = R_v/s^2 - n + 2q_v$$

$$MC_v = R_v/s^2 - (n-p)(n-q_v-2)/(n-p-2)$$

where s^2 is $R_v/(n-q_v)$ for the case in which Z_v contains all the $\underline{x}_i$. When analyzing data, the correct subset of predictors is chosen to be the one among the 2^p subsets which has the smallest value for the criterion being considered.

Under the assumption $\underline{\beta}_v = \underline{0}$, and $\underline{\epsilon}$ distributed with mean $\underline{0}$ and variance I.

$$\begin{aligned} & E(MS_v) = 1, \\ (6) \qquad & E(C_v) = (p-q_v)(n-p)/(n-p-2) + 2q_v - p \\ & \qquad \simeq q_v \text{ if n is large,} \\ & E(MC_v) = 0. \end{aligned}$$

These expectations are of some interest when studying the results in the next section.

3. RESULTS. As indicated in the introduction, the mathematics appeared to be intractable, so performance was studied empirically. The scope of this work has been restricted by the computational capacity available to date, but programming for a much faster computer is now in process and it will be possible to study more predictors than three. The current results, some of which are shown in the following tables, may provide some helpful insight toward obtaining at least an approximate analytic solution. They also may aid in constructing a sharper criterion for selecting predictors.

Table 1. Success in choosing correct predictors by smallest residual mean square.

Values of			No. of runs	Number of times each subset of predictors chosen															
				ρ = .00								ρ = .95							
β_1	β_2	β_3	k	-	1	2	3	12	13	23	123	-	1	2	3	12	13	23	123
0	0	0	100	*31	15	17	19	5	4	4	5	*31	24	8	10	6	6	12	3
0	0	0	100	*29	17	8	19	4	9	10	4	*33	18	12	4	5	4	16	8
1	0	0	100	0	*43	0	0	25	18	0	14	0	*52	0	0	14	12	0	22
0	1	0	100	0	0	*39	0	23	0	26	12	0	0	*41	12	19	7	10	11
1	1	0	100	0	0	0	0	*66	0	0	34	0	0	0	0	*64	21	0	15
0	1	1	100	0	0	0	1	0	0	*70	29	0	0	23	23	11	13	*16	14
1	1	1	100	0	0	0	0	0	1	0	*99	0	0	0	0	38	36	0	*26
2	0	0	100	0	*46	0	0	22	22	0	10	0	*46	0	0	15	15	0	24
0	2	0	100	0	0	*47	0	16	0	21	16	0	0	*46	6	20	5	16	7
2	2	0	100	0	0	0	0	*65	0	0	35	0	0	0	0	*70	5	0	25
2	2	2	100	0	0	0	0	0	0	0	*100	0	0	0	0	3	9	0	*88
0	5	0	100	0	0	*41	0	26	0	23	10	0	0	*46	0	26	0	24	4
0	5	0	100	0	0	*52	0	19	0	21	8	0	0	*45	0	29	0	16	10

*Correct choice

Table 2. Success in choosing correct predictors by smallest C-statistic.

Values of			No. of runs	Number of times each subset of predictors chosen															
				ρ = .00								ρ = .95							
β_1	β_2	β_3	k	-	1	2	3	12	13	23	123	-	1	2	3	12	13	23	123
0	0	0	100	*61	9	12	13	1	0	2	2	*69	9	5	6	0	5	5	1
0	0	0	100	*53	15	10	13	2	4	2	1	*63	15	8	3	2	3	5	1
1	0	0	100	0	*66	0	0	15	13	0	6	0	*76	0	0	7	16	0	11
0	1	0	100	0	0	*63	0	14	0	16	7	0	0	*57	15	11	4	11	2
1	1	0	100	0	0	0	0	*86	0	0	14	0	0	0	0	*68	21	0	11
0	1	1	100	0	0	0	1	0	0	*85	14	0	0	37	38	7	8	*7	3
1	1	1	100	0	0	0	0	0	1	1	*98	0	0	0	0	49	44	1	*6
2	0	0	100	0	*69	0	0	14	15	0	2	0	*76	0	0	6	9	0	9
0	2	0	100	0	0	*72	0	7	0	14	7	0	0	*66	10	12	4	6	2
2	2	0	100	0	0	0	0	*80	0	0	20	0	0	0	0	*83	7	0	10
2	2	2	100	0	0	0	0	0	0	0	*100	0	0	0	0	7	12	0	*81
0	5	0	100	0	0	*66	0	16	0	15	3	0	0	*71	0	10	0	18	1
0	5	0	100	0	0	*71	0	11	0	15	3	0	0	*64	0	19	0	11	6

*Correct choice

Table 3. Success in choosing correct predictors by smallest modified C-statistic.

Values of			No. of runs	Number of times each subset of predictors chosen															
				ρ = .00								ρ = .95							
β_1	β_2	β_3	k	-	1	2	2	12	13	23	123	-	1	2	3	12	13	23	123
0	0	0	100	*39	11	16	17	5	4	3	5	*40	19	9	8	4	6	11	3
0	0	0	100	*32	17	8	17	4	9	9	4	*41	16	10	5	4	5	15	4
1	0	0	100	0	*49	0	0	21	18	0	12	0	*56	0	0	13	11	0	20
0	1	0	100	0	0	*45	0	20	0	23	12	0	0	*44	13	17	5	12	9
1	1	0	100	0	0	0	0	*71	0	0	29	0	0	0	0	*64	21	0	15
0	1	1	100	0	0	0	1	0	1	*73	25	0	0	25	25	10	13	*[illegible]1	10
1	1	1	100	0	0	0	0	0	1	0	*99	0	0	0	0	39	38	0	*23
2	0	0	100	0	*55	0	0	18	21	0	6	0	*51	0	0	13	12	0	24
0	2	0	100	0	0	*53	0	14	0	20	13	0	0	*50	6	18	4	15	7
2	2	0	100	0	0	0	0	*70	0	0	30	0	0	0	0	*71	5	0	24
2	2	2	100	0	0	0	0	0	0	0	*100	0	0	0	0	3	9	0	*88
0	5	0	100	0	0	*47	0	25	0	20	8	0	0	*50	0	24	0	23	3
0	5	0	100	0	0	*54	0	17	0	23	6	0	0	*49	0	26	0	15	10

*Correct choice

Table 4. Comparison of choice by residual mean square (MS), C-statistic (C) and modified C-statistic (MC).

Values of			No. of runs	Number of correct choices					
				$\rho = .00$			$\rho = .95$		
β_1	β_2	β_3	k	MS	C	MC	MS	C	MC
0	0	0	100	31	61	39	31	69	40
0	0	0	100	29	53	32	33	63	41
1	0	0	100	43	66	49	52	76	56
0	1	0	100	39	63	45	41	57	44
1	1	0	100	66	86	71	64	68	64
0	1	1	100	70	85	73	16	7	11
1	1	1	100	99	98	99	26	6	23
2	0	0	100	46	69	55	46	76	51
0	2	0	100	47	72	53	46	66	50
2	2	0	100	65	80	70	70	83	71
2	2	2	100	100	100	100	88	81	88
0	5	0	100	41	66	47	46	71	50
0	5	0	100	52	71	54	45	64	49

SAMPLE CENSORING*

N. L. Johnson
Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

1. INTRODUCTION. There are currently available a number of methods designed to reduce the possible effects of "wild" ("maverick") observations on the analysis of sample values. Among these may be mentioned "trimming" and "Winsorisation". These methods involve the possible or sometimes automatic exclusion of extreme values among those observed. Apart from these methods, for which appropriate statistical analyses, taking proper account of the omission of sample values, are available, samples may be incomplete owing to inadequate recording, or, unfortunately, biassed selection of values which accord best with some preconceived ideas or desires.

While, under properly regulated conditions, information on any censoring of sample values should accompany the records of the values themselves, this is not always the case. Indeed, with the last situation described with the preceding paragraph, such information is not to be expected; but also, even in more respectable cases, information may be omitted by negligence.

The problems to be considered in this paper are those arising when it is suspected that there has been some form of censoring of the original sample. Complete, and reasonably tidy solutions are obtained only on the assumption that the population distribution of an observed character is known. However, study of this situation does give some clue as to what can be done when knowledge of the population distribution is incomplete.

Problems of a similar kind have been discussed in an earlier paper [1]. They were of a rather simple nature in that there was usually a direct choice between two possible sample sizes.

2. FORMAL STATEMENT OF PROBLEM. It will be supposed that there are available r observations of a character (X) which may be regarded as observed values of random variables $x_1', x_2', \ldots, x_r'$. These are a sub-set of the n ($\geq$ r) variables $x_1', x_2', \ldots, x_n'$ corresponding to a complete random sample of (unknown) size n. If r = n, then the 'sub-set' is identical with the complete sample. We will be interested in testing whether this is, in fact, the case. Various kinds of alternatives, specifying different kinds

*The work was supported in part by Army Research Office Contract AROD-4, and in part by Air Force Contract AF-AFOSR-760-65.

of censoring which might be applied to the complete sample, can be considered. Certain special kinds of censoring have been discussed in earlier papers [2] [3], and the results of these investigations will be summarized in Sections 3 and 4. Then, in Section 5, we will consider problems associated with general types of censoring. Certain practical problems arising in application of the tests described in Sections 3, 4, and 5 will be discussed in Section 6.

Discussion will be restricted to situations in which $x_1', x_2', \ldots, x_n'$ can be regarded as n independent continuous random variables, with a known common probability density function, represented by $f(x)$.

3. SYMMETRICAL CENSORING OF EXTREMES. We will suppose that if censoring occurs it takes the form of exclusion of the s greatest and s least among the original n sample values. Then $x_1', x_2' \ldots, x_r'$ are the r central values among an original set of $n(=r+2s)$ values. Denoting this hypothesis by $H_{s,s}$ the joint probability density function of the r ordered variables $x_1 \leq x_2 \leq \ldots \leq x_r$ (these being a rearrangement of $x_1', x_2', \ldots, x_r'$ in increasing order of magnitude) is:

$$p(x_1, x_2, \ldots, x_r \mid H_{s,s}) = \frac{(r+2s)!}{(s!)^2} [F(x_1)]^s [1-F(x_r)]^s \prod_{j=1}^{r} F(x_j) \tag{1}$$

$$(x_1 \leq x_2 \leq \ldots \leq x_r)$$

where $$F(x) = \int_{-\infty}^{x} f(x)\, dx .$$

The hypothesis that there has been no censoring and therefore that the complete sample is available is, in the notation already introduced, $H_{o,o}$. For brevity this will be denoted by H_o.

The most powerful test of H_o against the alternative $H_{s,s}$ has a critical (rejection) region of the form.

$$p(x_1, \ldots, x_r \mid H_{s,s}) \geq Cp(x_1, \ldots, x_r \mid H_o) \tag{2}$$

where C is a constant. Whatever be the value of s, inequality (2) can be written in the form.

$$F(x_1)\ [1 - F(x_r)] \geq K \tag{3}$$

where K is a constant. Inequality (3) does not depend on s, so the test defined by this inequality is uniformly most powerful with respect to $H_{s,s}$ for all $s > 0$; i.e. with respect to any symmetrical censoring of the extremes of the sample values. The value of K must be chosen to give a required level of significance, α say, when H_0 is true. This value depends on α and r, and may be denoted by $K(\alpha, r)$. Then

$$\text{(4)} \qquad \Pr[F(x_1)[1-F(x_r)] \geq K(\alpha, r) \mid H_0] = \alpha \quad .$$

Table 1 gives a few values of $K(\alpha, r)$. For

$r \geq 10$ the approximations

$K(0.10, r) \doteqdot 2.65(r+1.5)^{-2}$

$K(0.05, r) \doteqdot 4.1(r+2)^{-2}$

$K(0.01, r) \doteqdot 9.2(r+3.5)^{-2}$

give useful results. Mathematical analysis connected with the determination of $K(\alpha, r)$ is contained in Appendix I.

A discussion of the evaluation of the power of this test is contained in Appendix II.

TABLE 1

Upper 100 α% Significance Limits of $F(x_1)[1 - F(x_r)]$

r \ α	0.05	0.01
2	0.207	0.235
3	0.150	0.195
4	0.109	0.156
5	0.0822	0.125
6	0.0633	0.101
7	0.0503	0.0830
8	0.0408	0.0692
9	0.0338	0.0585

4. GENERAL CENSORING OF EXTREMES. If the requirement of symmetry is dropped we need to consider hypotheses of form H_{s_o, s_r} corresponding to exclusion of the s_o smallest and s_r largest individuals in the original sample, with s_o and s_r not necessarily equal. In this case there is no longer a uniformly most powerful test of H_0. There is a uniformly most powerful test of H_0 with respect to the subclass $H_{\theta s_r, s_r}$ in which s_o/s_r (= θ) is constant.

It has a critical region of form

$$[F(x_1)]^{\theta} [1-F(x_r)] \geq K(\alpha, r, \theta) \tag{5}$$

[If $s_r = 0$, we take $\theta = \infty$ and replace (5) by $F(x_1) \geq$ constant]

To obtain a significance level equal to α , the value of $K(\alpha, r, \theta)$, given H_0 is valid (i.e. there is no censoring), must make the probability that inequality (5) is satisfied equal to α . In [3] a heuristic method proposed by S. N. Roy [4] is applied to suggest a possible test of H_0 with respect to all alternative hypotheses of type H_{s_o, s_r} (for any values of s_o and s_r).

This calls for construction of the union of regions like (5) with $\alpha \simeq \alpha'$, over all values of θ. Points $(F(x_1), F(x_r))$ on the boundary of the critical region must satisfy the equations.

$$[F(x_1)]^{\theta} [1-F(x_r)] = K(\alpha', r, \theta) \tag{6.1}$$

$$\frac{\partial}{\partial\theta} \{ [F(x_1)]^{\theta} [1-F(x_r)] \} = \partial K(\alpha', r, \theta)/\partial\theta . \tag{6.2}$$

From (6.1) and (6.2) it follows that

$$\log F(x_1) = \partial \text{ long } K(\alpha', r, \theta)/\partial\theta . \tag{6.3}$$

If $K(\alpha', r, \theta)$ is known, $F(x_1)$ can be found from (6.3) and then $F(x_1)$ is determined by (6.1). However explicit evaluation of $K(\alpha', r, \theta)$ is troublesome, and approximate methods were used in [3] leading to the simple (through approximate) formula:

(7) $$F(x_1) + [1 - F(x_r)] \geq K_1(\alpha, r)$$

for the union of critical regions. Here $K_1(\alpha, r)$ represents a constant which can be chosen to give a required value, α say, for the significance level. (Note that α' appears only in the construction of (7); it is not the significance level of the resultant test.)

Although an approximate argument, applying a heuristic principle has been used in reaching (7), the critical region so obtained has a natural appeal; and seems worthy of further consideration.

The distribution theory associated with the critical region (7) is very simple. If H_{s_0, s_r} is valid then $F(x_1) + [1 - F(x_r)]$ is distributed as

$\chi^2_{2(s_0+s_r+2)} / (\chi^2_{2(s_0+s_r+2)} + \chi^2_{2(r-1)})$ where $\chi^2_{2(s_0+s_r+2)}$ and $\chi^2_{2(r-1)}$ are mutually independent. (Equivalently, the distribution is a beta distribution with parameters $(s_0 + s_r + 2)$, $(r-1)$.) It follows (putting $s_0 = s_r = 0$) that

(8) $K_1(\alpha, r)$ = upper 100 α% point of beta distribution with parameters 2, (r-1). These values can be obtained from Table 16 of [6].

The power of the test with respect to a specified alternative hypothesis H_{s_0, s_r} is also easily calculated. In fact

(9) $$\Pr[F(x_1) + (1-F(x_r)) \geq K_1 \mid H_{s_0, s_r}] = 1 - I_{K_1}(s_0 + s_r + 2, r-1)$$
$$= I_{1-K_1}(r-1, s_0 + s_r + 2)$$

where $I_p(M, N) = [B(M, N)]^{-1} \int_0^p t^{M-1}(1-t)^{N-1}\, dt$ is the incomplete beta function ratio.

For given s_0 and s_r, as r tends to infinity the power tends to

(10) $$\Pr[\chi^2_{2(s_0+s_r+2)} \geq \chi^2_{4, 1-\alpha}]$$

(where $\chi^2_{\nu, 1-\alpha}$ denotes the upper 100 α % point of the distribution of χ^2 with ν degrees of freedom).

A few values of the power are shown in Table 2. It appears that the asymptotic ($r \longrightarrow \infty$) values give a good indication of true value for $r > 30$.

TABLE 2

Power β_{s_o, s_r} of the general purpose test ($\alpha = 0.05$)

$s_o + s_r =$	2	6	10	14	18
$r = 4$	0.167	0.470	0.716	0.862	0.938
$r = 30$	0.281	0.845	0.989	-	-
$r = \infty$	0.303	0.892	0.996	-	-

A special case of some interest arises when censoring at one extreme only is suspected (i.e. $s_o = o$ or $s_r = o$). In this case the uniformly most powerful test has the critical region

$$y_r < \alpha^{1/r} \quad (\text{if } s_o = o)$$

or

$$y_1 > 1-\alpha^{1/r} \quad (\text{if } s_r = 0) \ .$$

The power of the test with critical region $y_r < \alpha^{1/r}$ with respect to the alternative H_{o, s_r} is

$$\beta(H_{o, s_r}) = \frac{(r+s_r)!}{(r-1)!\, s_r!} \int_o^{\alpha^{1/r}} y^{r-1}(1-y)^{s_r}\, dy$$

$$= I_{\alpha^{1/r}}(r, s_r + 1)$$

(where I denotes the incomplete beta function ratio).

5. GENERAL CENSORING. We first introduce the notation $H_{s_o, s_1, \ldots, s_r}$ to denote the hypothesis that s_j observations have been removed between x_{j-1} and x_j for $j=1, 2, \ldots, (r+1)$ with $x_o = -\infty$, $x_{r+1} = +\infty$.

In this notation the H_{s_o, s_r} considered in Sections 3 and 4 would be $H_{s_o, o, o, \ldots, o, s_r}$. Also, for convenience we will write

$$(11) \qquad y_j = F(x_j) \qquad (j = 1, \ldots r) \quad .$$

$$y_o = 0; \quad y_{r+1} = 1$$

Then the best critical region for testing the hypothesis of no censoring ($H_{o, o, \ldots, o, o}$) against the alternative $H_{s_o, s_1, \ldots, s_r}$ is of form

$$(12) \qquad \prod_{j=o}^{r} (y_{j+1} - y_j)^{s_j} \geq K(\alpha, r, s_o, s_1, \ldots, s_r) \quad .$$

It is clear that there is a uniformly most powerful test with respect to any set of alternatives $H_{s_o, s_1, \ldots, s_r}$ for which the ratios $s_o : s_1 : \ldots : s_r$ are constant, but not with respect to any other sets of alternatives. While one could attempt to apply Roy's heuristic principle, as in [5], to construct a general purpose critical region for the whole set of alternatives $H_{s_o, s_1, \ldots, s_r}$ the effect of approximations might well be much more important in the more general case, and is certainly more difficult to gauge. We therefore consider more or less arbitrarily chosen criteria which, however, do have some relation to criteria suggested from theoretical considerations.

We first consider a test with critical region

$$(13) \qquad g = \prod_{j=o}^{r} (y_{j+1} - y_j) > K_2(\alpha, r) = K_2 \quad .$$

It is quite likely that this criterion may be felt to have some practical drawbacks. These will be discussed in Section 6, but for the present we will just consider how to evaluate K_2 in (13), at any rate approximately.

It will be convenient to approximate to the distribution of log g, rather than of g itself. The moment generating function of log g, when

$H_{s_o, s_1, \ldots, s_r}$ is valid, is

$$(14) \qquad E\,[g^t \mid H_{s_o, s_1, \ldots, s_r}] = \frac{(r + \sum_{o}^{r} s_j)!}{\prod_{j=o}^{r} s_j} \times$$

$$\times \int \int \cdots \int \prod_{j=o}^{r} (y_{j+1}-y_j)^{s_j+t} dy_1 \ldots dy_r \quad .$$

[The region of integration is $0 \leq y_1 \leq y_2 \leq \cdots \leq y_r \leq 1$. Remember that $y_o = 0$ and $y_{r+1} = 1$.]

Since the joint probability density function of $y_1, \ldots, y_r$ is

$$p(y_1, \ldots y_r \mid H_{s_o, s_1, \ldots, s_r}) = \frac{\Gamma(r+1+\sum_{o}^{r} s_j)}{\prod_{j=o}^{r} \Gamma(s_j+1)} \prod_{j=o}^{r} (y_{j+1}-y_j)^{s_j}$$

it follows that

$$(15) \qquad \int \int \cdots \int \prod_{j=o}^{r} (y_{j+1}-y_j)^{s_j} dy_1 \ldots dy_r = \frac{\prod_{j=o}^{r} \Gamma(s_j+1)}{\Gamma(r+1+\sum_{o}^{r} s_j)}$$

and hence from (14) and (15)

$$(16) \qquad E\,[g^t \mid H_{s_o, s_1, \ldots, s_r}] = \frac{(r+\sum_{o}^{r} s_j)!}{\prod_{j=o}^{r} s_j!} \; \frac{\prod_{j=o}^{r} \Gamma(t+s_j+1)}{\Gamma(\,(r+1)(t+1) + \sum_{o}^{r} s_j)} \; .$$

Taking logarithms and differentiating, the following expression for the mth cumulant of log g is obtained:

(17) $$\kappa_m(\log g \mid H_{s_o, s_1, \ldots, s_r})$$

$$= \sum_{j=o}^{r} \Psi^{(m-1)}(s_j+1) - (r+1)^m \Psi^{(m-1)}(r+1+\sum_{j=o}^{r} s_j) \quad .$$

In particular when the null hypothesis H_0 ($\equiv H_{o,o,\ldots,o}$) is valid

(18) $$\kappa_m(\log g \mid H_o) = (r+1)\Psi^{(m-1)}(1) - (r+1)^m \Psi^{(m-1)}(r+1).$$

The polygamma functions have the values

$$\Psi(1) = -\gamma = -0.5772$$

and $\Psi^{(m-1)}(1) = (-1)^m (m-1)!\ S_m \quad (m \geq 2)$

where $S_m = 1 + 2^{-m} + 3^{-m} + \ldots$

Hence

(19.1) $\kappa_1(-\log g \mid H_0) = (r+1)(\gamma + \Psi(r+1))$

(19.2) $\kappa_m(-\log g \mid H_0) = (r+1)\left[(m-1)!\ S_m + (-1)^{m-1}(r+1)^{m-1}\Psi^{(m-1)}(r+1)\right]$

$(m \geq 2) \quad .$

For z not too small, we have, to a good approximation

(20.1) $\Psi(z) \doteq \log(z - 1/2)$

(20.2) $\Psi^{(m)}(z) \doteq (-1)^{m-1}(m-1)!\ (z - 1/2)^{-m} \ (m \geq 1)$

whence

(21.1) $\kappa_1(-\log g \mid H_0) \doteq (r+1)(0.57722 + \log(r+1/2))$

(21.2) $\kappa_m(-\log g \mid H_0) \doteq (r+1)(m-1)!\,[S_m - (m-1)^{-1}\{(r+1)/(r+1/2)\}^{m-1}]$.

Noting that

(i) the least possible value of $(-\log g)$ is $(r+1)\log(r+1)$, corresponding to $y_j = j/(r+1)$ for $j=1, 2, \ldots, r$

(ii) $$\frac{[\kappa_3(-\log g \mid H_0)]^2}{[\kappa_2(-\log g \mid H_0)]^3} \doteq \frac{(2S_3-1)^2}{(r+1)(S_2-1)^3} = \frac{7.35}{r+1}$$

and

$$\frac{\kappa_4(-\log g \mid H_0)}{\kappa_2(-\log g \mid H_0)} \doteq \frac{6S_4-2}{(r+1)(S_2-1)^2} = \frac{10.80}{r+1}$$

(while for χ^2 with $(r+1)$ degrees of freedom, $\kappa_3^2/\kappa_2^3 = 8/(r+1)$ and $\kappa_4/\kappa_2^2 = 12/(r+1)$)

(iii) $\mathrm{var}(-\log g \mid H_0) = 0.645(r+1)$

while $\mathrm{var}(0.57722\,\chi^2_{r+1}) = 0.666(r+1)$

it appears that we might take, as an approximation,

(22) $-\log g - (r+1)\log(r+1)$ to be distributed as $0.57722 \times$ (χ^2 with $(r+1)$ degrees of freedom) or, equivalently

(22)' $1.732\,[-\log g - (r+1)\log(r+1)]$ to be distributed as χ^2 with $(r+1)$ degrees of freedom. This implies

$$K_2 \doteq \frac{\exp[-\chi^2_{r+1,\alpha}/1.732]}{(r+1)^{r+1}}$$

where

$\chi^2_{r+1,\alpha}$ is the lower α% point of the distribution of χ^2 with $(r+1)$ degrees of freedom.

(If $-\log g - (r+1)\log(r+1)$ is approximated by $0.5587\,\chi^2_{1.0332(r+1)}$, then means and variances agree while the values of κ_3^2/κ_2^3 and κ_4/κ_2^2 for the approximating distribution are $7.74(r+1)^{-1}$ and $11.61(r+1)^{-1}$.)

The approximations cannot be expected to be good unless r is fairly large. In the extreme case $r = 1$ with $g = y_1(1-y_1)$ we have exactly

$$\text{(23.1)} \qquad \Pr[g > G \mid H_0] = (1-4G)^{1/2} \quad (0 \le G \le 1/4)$$

while (22) gives

$$\text{(23.2)} \qquad \Pr[g > G \mid H_0] \doteq 1 - (4G)^{0.866} .$$

The approximation (23.2) is substantially less than the true value (23.1) though it does have the correct limits (1 and 0) as G tends to 0 or 1/4. In order to assess the power of this test we return to equation (17). This gives the cumulants of log g when a general alternative hypothesis $H_{s_0, s_1, \ldots, s_r}$ is valid. It would seem reasonable to fit the distribution of $[-\log g - (r+1)\log(r+1)]$ by that of a multiple of χ^2, so that first and second moments agree. It may be that better approximations to upper percentage points of -log g would be obtained by fitting the first three moments (instead of the initial point and first two moments - see [4]). This method might therefore be employed when the power is, say, above 0.75.

6. MODIFIED TESTS. The test criteria described above are all based on the probability integral transformation

$$\text{(24)} \qquad y = \int_{-\infty}^{x} f(x)\,dx .$$

They explicitly assume that f(x) is known exactly (in practice to a close approximation) and that there are no errors in observation of x. This last condition is never satisfied when x is a continuous variable. There is always some kind of grouping error occasioned by the finiteness of the number of digits used in recording the observations. This is particularly important in relation to test functions like g of (13) in Section 5. If it so happens that any two of the y's are equal the value of g is zero and the null hypothesis will be accepted. Clearly, if this happens because of the use of too coarse a grouping interval, the test is likely to be very insensitive. Furthermore, the larger r is, the more likely it is that at least two x's (and so two y's) will be equal, thus giving rise to a zero value for g. We are thus led to consider modified tests, less sensitive to this kind of effect. A simple way of effecting this is to use only a selected number of the transformed order statistics $y_1, y_2, \dots, y_r$ -say $y_{a_1}, y_{a_2}, \dots, y_{a_k}$ (with the values $1 \le a_1 < a_2 < \dots < a_k \le r$ fixed before analyzing the data, of course) and to apply a test with critical region

$$(25) \qquad g_a = \prod_{j=0}^{k+1} (y_{a_j} - y_{a_{j-1}}) > K_3 ,$$

with $y_{a_{k+1}} = 1$, $y_{a_0} = 0$. (A natural choice would be to take the a's at equal intervals apart.)

The value of K_3 depends on the required significance level, α , and also on the selected a'_js, as well as on r. In fact the distribution of g_a, when H_0 is valid, is the same as that of g, with r replaced by k, when $H_{s_0, s_1, \dots, s_k}$ is valid and with $s_j = a_{j+1} - a_j - 1$ $(j=0, 1, 2, \dots, k)$ hence, the same calculations as those needed to evaluate the power of the test using g are required in calculating the value K_3 in (25). Also, of course, calculation of the power of the test with critical region (25) will follow the same lines.

A similar kind of modification can be applied to tests of symmetrical censoring of extremes (Section 3). In this case it would be natural to ignore the least and greatest m observations, and use only $y_{m+1}, \dots, y_{r-m}$. The uniformly most powerful test of H_0 against symmetrical alternatives $H_{s,s}$ has a critical region form similar to (3), viz:

$$(26) \qquad y_{m+1} (1 - y_{r-m}) \ge K_4 .$$

Determination of K_4 is, however, more troublesome than for K. The equation

$$(27)\quad \frac{r!}{(m!)^2(r-2m-2)!}\int\int y_{m+1}^{m}(y_{r-m}-y_{m+j})^{r-2m-2}(1-y_{r-m})^{m}dy_{m+1}dy_{r-m}$$

(where the region of integration is $y_{m+1}(1-y_{r-m}) \geq K_4$;

$0 \leq y_{m+1} \leq y_{r-m} \leq 1$)

has to be satisfied.

Evaluation of the integral of the left hand side, with K_4 replaced by K, gives the power of the test with critical region (3) with respect to the alternative hypothesis $H_{m,m}$. The notes in Appendix II are therefore relevant to this problem.

7. CONDITIONS OF APPLICABILITY. It may be felt that the condition stated at the beginning of Section 6, namely that the true probability density function f(x) must be known, is unlikely to be satisfied in practice. While this is so, in the strict sense that it is very rarely the case that a theoretically formulated model gives an exact representation of reality, it will sometimes be the case that there is sufficiently massive evidence to establish f(x), from observed relative frequencies, with adequate accuracy. Slight variations in form of f(x) can be tolerated without serious effect, particularly if a modified test of the type described in Section 6 is used. It may be noted that it is not essential that f(x) have a simple, or indeed any explicit, mathematical form - a graphical representation can suffice.

It would, however, be interesting, but beyond the scope of the present investigation, to inquire into the robustness of these tests with respect to variation in f(x). (i.e. to use of an incorrect function, $f_1(x)$ say, in (24)).

REFERENCES

1. Johnson, N. L. (1962) "Estimation of Sample Size", Technometrics, 4, 59-67.

2. Johnson, N. L. (1966) "Tests of Sample Censoring" Proc. 20th Tech. Conf. ASQC, 699-703.

3. Johnson, N. L. (1966) "A General Purpose Test of Censoring of Extreme Sample Values", Submitted to S. N. Roy Memorial Volume.

4. Pearson, E. S. (1959) "Note on an Approximation to the Distribution of Non-Central X^2", Biometrika, 46, 364.

5. Roy, S. N. (1957) Some Aspects of Multivariate Analysis, Wiley. New York: Indian Statistical Institute, Calcutta.

6. Biometrika Tables for Statisticians (Ed. E.S. Pearson and H.O. Hartley) Cambridge University Press 1958.

Appendix I

We have to consider the evaluation of $K(\alpha, r)$ from equation (4). Puting $y_j = F(x_j)$ (as in (11)), the joint probability density of y_i and y_r, given H_0, is

$$(A.1) \qquad p(y_1, y_r \mid H_0) = r(r-1)(y_r - y_1)^{r-2} \quad (0 \le y_1 \le y_r \le 1) \ .$$

Hence $K(\alpha, r)$ (now written as K for convenience) satisfies the equation

$$(A.2) \qquad r(r-1) \iint\limits_{y_1(1-y) \ge K} (y_r - y_1)^{r-2} \, dy_1 dy_r = \alpha \ .$$

The region $y_1(1-y_r) \ge K$ can be defined by the inequalities $y_1 \le y_r \le 1-K/y_1$ and these imply also $1-y_1 - K/y_1 \ge 0$ or $Y_- \le y_1 \le Y_+$ where $Y_{\pm} = [1 \pm \sqrt{1-4K}]/2$.

Hence from (A.2)

$$(A.3) \qquad r \int_{Y-}^{Y+} (1-Ky^{-1}-y)^{r-1} \, dy = \alpha \ .$$

Expanding the integrand and integrating term by term leads to the equation

$$(A.4) \qquad r \sum_{j=o}^{r-1} \binom{r-1}{j} (-1)^j K^j \sum_{i=o}^{r-j-1} \binom{r-j-1}{i} (-1)^i h_{i-j+1}(\sqrt{1-4K})$$

where $h_o(z) = \log\left(\frac{1+z}{1-z}\right)$

$$h_m(z) = 2^{-m} m^{-1} [(1+z)^m - (1-z)^m] \ .$$

Note that for $m > 0$,

(A.5) $$h_m(\sqrt{1-4K})$$

$$= [m^{-1} \sum_{\substack{o \le j \\ \le (m-1)/2}} (-1)^j \binom{m-1-j}{j} K^j] \sqrt{1-4K}$$

$$= K^m h_{-m} (\sqrt{1-4K}) \ .$$

For $r = 2(1)\,9$, the left hand side of (A.4) is shown in Table A.1 below.

TABLE A.1

$r =$	$\sqrt{1-4K}$ x	$-\log \frac{1+\sqrt{1-4K}}{1-\sqrt{1-4K}}$ x
2	1	2K
3	1 + 8K	6K
4	1 + 26K	12K (1 + K)
5	$1 + \frac{166}{3} K + \frac{128}{3} K^2$	$20K(1 + 3K)$
6	$1 + 97K + 226K^2$	$30K(1 + 6K + 2K^2)$
7	$1 + \frac{759}{5} K + \frac{3558}{5} K^2 + \frac{1024}{5} K^3$	$42K(1 + 10K + 10K^2)$
8	$1 + \frac{1102}{5} K + \frac{8654}{5} K^2 + \frac{7492}{5} K^3$	$56K(1 + 15K + 30K^2 + 5K^3)$
9	$1 + \frac{10618}{35} K + \frac{125634}{35} K^2 + \frac{218044}{35} K^3 + \frac{32768}{35} K^4$	$72K(1+21K+70K^2 + 35K^3)$

(For example for $r = 3$, $1+8K) \sqrt{1-4K} - 6K \log \frac{1+\sqrt{1-4K}}{1-\sqrt{1-4K}} = \alpha$.)

The calculations rapidly become more complicated as r increases. It therefore is desirable to search for some approximation to $K(\alpha, r)$ which will give useful results for r large (and preferably for $r \geq 10$).

Some empirical formulae have been given in Section 3. Here we use an analytical approach, starting from equation (A. 3). We first make a succession of transformations, aimed at obtaining an integrand for which useful bounds can be set.

Firstly, putting $y = z\sqrt{K}$

$$\int_{Y_-}^{Y+} (1-Ky^{-1}-y)^{r-1}\, dy = \sqrt{K} \int_{1/A(K)}^{A(K)} \{1-\sqrt{K}(z^{-1}+z)\}^{r-1}\, dz$$

where $A(K) = (1/2)(1+\sqrt{1-4K})/\sqrt{K}$.

Next making the transformation $z = e^t$ the integral becomes

$$\sqrt{K} \int_{-\log A(K)}^{\log A(K)} e^t \{1-\sqrt{K}(e^t + e^{-t})\}^{r-1}\, dt$$

$$= \sqrt{K} \int_{-\log A(K)}^{\log A(K)} e^{-t} \{1-\sqrt{K}(e^t + e^{-t})\}^{r-1}\, dt$$

$$\text{(A. 6)} \quad = 2\sqrt{K} \int_0^{\log A(K)} (1-2\sqrt{K}\cosh t)^{r-1} \cosh t\, dt.$$

Now making the transformation $v = 2\sqrt{K}\cosh t$, we obtain

$$\text{(A. 7)} \qquad \int_{2\sqrt{K}}^{1} (1-v)^{r-1} (v^2 - 4K)^{-1/2} v\, dv \quad .$$

Integrating by parts, this is equal to

$$\text{(A. 8)} \qquad (r-1) \int_{2\sqrt{K}}^{1} (1-v)^{r-2} (v^2 - 4K)^{1/2}\, dv.$$

Thus equation (A. 3) can be written

$$r(r-1) \int_{2\sqrt{K}}^{1} (v^2 - 4K)^{1/2} (1-v)^{r-2}\, dv = \alpha \quad .$$

Making the final transformation $v = 2\sqrt{K} + (1-2\sqrt{K})u$ we obtain

$$\text{(A.9)} \qquad r(r-1)(1-2\sqrt{K})^{r-1/2} \int_0^1 \{(1-2\sqrt{K})u^2 + 2\sqrt{K}\,u\}^{1/2} \cdot (1-u)^{r-2}\,du = \alpha \quad .$$

Since

$$\sqrt{2\sqrt{K}}\,\sqrt{u} \le \{(1-2\sqrt{K})u^2 + 2\sqrt{K}\,u\}^{\frac{1}{2}} \le \sqrt{1-2\sqrt{K}}\,u + \sqrt{2\sqrt{K}}\,\sqrt{u}$$

it follows that

$$\text{(A.10)} \qquad (1-2\sqrt{K})^{r-1/2}\sqrt{2\sqrt{K}} \;\; \frac{\frac{1}{2}\sqrt{\pi}\;\Gamma(r+1)}{\Gamma(r+1/2)} \le \alpha$$

$$\le (1-2\sqrt{K})^{r-1/2}\left[\sqrt{1-2\sqrt{K}} + 2\sqrt{K}\;\frac{\frac{1}{2}\sqrt{\pi}\;\Gamma(r+1)}{\Gamma(r+1/2)}\right] \quad .$$

As can be deduced by direct analysis, $K \to o$ as $r \to \infty$, but since $(1-2\sqrt{K})^r \le \alpha$

it follows that $K \ge 1/4\,(1-\alpha^{1/r})^2$ and hence Kr^2 cannot tend to zero.

If we put $K = Cr^{-2}$ (where C is, of course a function of r and α) then, approximately

$$\text{(A.11)} \qquad e^{-2\sqrt{C}}\,\sqrt{\pi/2}\;C^{1/4} \le \alpha \le e^{-2\sqrt{C}}(1 + \sqrt{\pi/2}\,C^{1/4}).$$

This implies that C lies between fixed limits, and suggests that, for large r, K is of the form $C\,r^{-2}$. (The form of function - $C_1(r+D_1)^{-2}$ - used as an approximation to K in Section 3 was suggested by this analysis.)

An alternative, heuristic approach is as follows:

If H_0 be valid, $y_1(1-y_r)$ is distributed as $uv/(u+v+w)^2$ where u, v and w are independent χ^2 random variables with 2, 2, 2(r-1) degrees of freedom respectively.

If r is large,

$$\Pr\left[\frac{uv}{(u+v+w)^2} > \frac{C}{r^2}\right] \doteq \Pr[uv > 4C]$$

(since $w/[2(r-1)] \sim 1$ as $r \to \infty$). Hence we have $K \sim Cr^{-2}$ where C satisfies the equation

$$\frac{1}{2}\int_0^\infty \exp\left(-\tfrac{1}{2}u - 2C/u\right) du = \alpha \ .$$

APPENDIX II

The joint probability density function of y_1 and y_r, when H_{s_o, s_r} is valid, is

$$\text{(A.12)} \qquad p(y_1, y_r) = \frac{(r+s_o+s_r)!}{s_o!(r-2)!\,s_r!} \, y_1^{\,s_o}(1-y_r)^{s_r}(y_r-y_1)^{r-2}$$

$$(o \le y_1 \le y_r \le 1) \, .$$

Hence

$$\Pr[y_1(1-y_r) \ge K \,|\, H_{s_o, s_r}]$$

$$\text{(A.13)} \qquad = \frac{(r+s_o+s_r)!}{s_o!(r-2)!\,s_r!} \int_{Y_-}^{Y_+} y_1^{\,s_o} \int_{y_1}^{1-K/y_1} (1-y_r)^{s_r}(y_r-y_1)^{r-2}\, dy_r\, dy_1$$

(where $Y_{\pm} = (1/2)\,[1 \pm \sqrt{1-4K}]$ as in (A.3)).

Noting that $(1-y_r)^{s_r} = \{(1-y_1)-(y_r-y_1)\}^{s_r}$ we see that the integral in (A.13) is equal to

$$\int_{Y_-}^{Y_+} y_1^{\,s_o} \sum_{j=o}^{s_r} \binom{s_r}{j} (-1)^j (1-y_1)^{s_r-j} (r+j-1)^{-1} (1-K/y_1-y_1)^{r-2}\, dy_1$$

$$= \int_{Y_-}^{Y_+} y_1^{\,s_o} \sum_{j=o}^{s_r} \binom{s_r}{j} (-1)^j (1-y_1)^{s_r-j} (r+j-1)^{-1} \sum_{i=o}^{r-2} \binom{r-2}{i} (-1)^i K^i y_1^{-i} (1-y_1)^{r-2-i}\, dy_1$$

$$\text{(A.14)} \qquad = \sum_{j=o}^{s_r} \sum_{i=o}^{r-2} (-1)^{j+i} \binom{s_r}{j} \binom{r-2}{i} (i+j-1)^{-1} K^i \int_{Y_-}^{Y_+} y_1^{\,s-i} (1-y_1)^{s_r+r-2-i-j}\, dy_1 .$$

Using (A.5) this can be expressed explicitly in terms of K. The resulting formula is rather cumbersome, and does not give much insight into the dependence of power on s_o and s_r. The following alternative approach, although it depends on some quite rough approximations, should give a reasonably accurate idea of the nature of this dependence, when r is large compared with s_o and s_r.

From (17) it follows that

$$\text{(A.15)} \qquad \kappa_m(-\log\{y_1(1-y_r)\})$$
$$= (-1)^m [\Psi^{(m-1)}(s_o+1) + \Psi^{(m-1)}(r) + \Psi^{(m-1)}(s_r+1)$$
$$- 3^m \Psi^{(m-1)}(s_o+s_r+r+2)] \quad .$$

Using the approximate formula (20) we obtain

$$\text{(A.16.1)} \qquad \kappa_1(-\log\{y_1(1-y_r)\}) \doteq 2\gamma - \sum_{j=1}^{s_o} j^{-1} - \sum_{j=1}^{s_r} j^{-1} - \log(r-1/2)$$
$$+ 3\log(s_o+s_r+r+3/2)$$

and, for $m \geq 2$

$$\text{(A.16.2)} \qquad \kappa_m(-\log\{y_1(1-y_r)\}) \doteq (m-1)! \left[\sum_{j=s_o+1}^{\infty} j^{-m} + \sum_{j=s_r+1}^{\infty} j^{-m} \right.$$
$$+ \{(m-1)(r-1/2)^{m-1}\}^{-1}$$
$$\left. - 3^m \{(m-1)(s_o+s_r+r+3/2)^{m-1}\}^{-1} \right] .$$

If r is large, then for the smaller values of $m(\geq 2)$

$$\text{(A.16.3)} \qquad \kappa_m(-\log\{y_1(1-y_r)\}) \doteq (m-1)! \left[\sum_{j=s_o+1}^{\infty} j^{-m} + \sum_{j=s_r+1}^{\infty} j^{-m} \right] \quad .$$

Note that r does not appear in this approximation.

In particular, taking m = 2

$$\text{(A.17)} \qquad \operatorname{var}(-\log\{y_1(1-y_r)\}) \doteq \sum_{j=s_o+1}^{\infty} j^{-2} + \sum_{j=s_r+1}^{\infty} j^{-2} \quad .$$

The variance decreases as s_o and /or s_r increases. The expected value (κ_1) also decreases.

A further approximation to (A. 16. 3) gives

$$\text{(A. 18. 1)} \quad \kappa_m(-\log\{y_1(1-y_r)\}) \doteq (m-2)!\,[(s_o + 1/2)^{-(m-1)} + (s_r + 1/2)^{-(m-1)}]$$

and in particular

$$\text{(A. 18. 2)} \qquad \kappa_2(-\log\{y_1(1-y_r)\}) \doteq (s_o + 1/2)^{-1} + (s_r + 1/2)^{-1} \quad .$$

If $s_o = s_r = s$, formula (A. 18. 1) becomes

$$\text{(A. 19. 1)} \qquad \kappa_m(-\log\{y_1(1-y_r)\}) \doteq 2(m-2)!\,(s+1/2)^{-(m-1)}$$

while (A. 16. 1) becomes

$$\text{(A. 19. 2)} \quad \kappa_1(-\log\{y_1(1-y_r)\}) \doteq 2\gamma - 2\sum_{j=1}^{s} j^{-1} - \log(r-1/2) + 3\log(2s+r+3/2).$$

If r is large this last equation may be replaced by

$$\text{(A. 19. 3)} \quad \kappa_1(-\log\{y_1(1-y_r)\} \doteq 2\gamma - 2\sum_{j=1}^{s} j^{-1} + 2\log(r+3s+5/2) \quad .$$

If s increases to s+1, κ_1 decreases by approximately $2(s+1)^{-1} - 6(r+3s+5/2)^{-1}$. It is not suggested that it will always be appropriate to use these approximations, particularly those appearing later, which depend heavily on r being large compared with s_o and s_r. The approximations are exhibited because they bring out rather clearly the way the distribution of $-\log\{y_1(1-y_r)\}$ depends on s_o and s_r.

LIST OF ATTENDEES

Banash, Robert C.	Rock Island Arsenal
Bartko, John J.	NIH, Bethesda, Maryland
Benton, Alan	BRL, Aberdeen Proving Ground
Billings, Michael G.	CEIR, Inc., Dugway, Utah
Biser, Erwin	Fort Monmouth
Bohidar, N. R.	Fort Detrick
Breaux, H. J.	BRL, Aberdeen Proving Ground
Bruno, O. P.	BRL, Aberdeen Proving Ground
Bulfinch, Alonzo	Picatinny Arsenal
Cameron, J. M.	NBS, Gaithersburg, Maryland
Campbell, Lloyd W.	BRL, Aberdeen Proving Ground
Cochran, W. G.	Harvard University
Cohen, A. C.	University of Georgia
Cox, E. L.	Agr. Res. Service, Beltsville, Md.
Currie, Lloyd	NBS, Gaithersburg, Md.
Danish, Michael B.	BRL, Aberdeen Proving Ground
Davis, James A.	STAG, Bethesda, Md.
David, H. A.	University of North Carolina
Dihm, H. A., Jr.	Redstone Arsenal
Dressel, F. G.	Army Research Office-Durham
Duff, J. B.	Fort Belvoir
Dutoit, E. F.	Picatinny Arsenal
Egner, D. O.	Limited War Lab., Aberdeen Proving Ground, Maryland
Eisenhart, C.	NBS, Gaithersburg, Md.
Eissner, R. M.	BRL, Aberdeen Proving Ground
Ellner, H.	Edgewood Arsenal
Embody, D. R.	Stat. Design of Expts, Section, Navy Bldg. Washington, D. C.
Endelman, Anna	FDA, Washington, D. C.
Essenwanger, O. M.	Redstone Arsenal
Foster, W. D.	Fort Detrick
Frishman, Fred	OCRD, Washington, D. C.
Gallivan, James P.	Army Mat's. Res Agcy, Watertown, Mass.
Gart, John J.	NIH, Bethesda, Md.
Gehan, Edmund	NIH, Bethesda, Md.
Gracia, David M.	Natick Laboratories
Greenwood, Joseph A.	FDA, Arlington, Va.
Grimes, James P.	Naval Res Lab., Washington, D. C.
Greenberg, Bernard G.	University of North Carolina Chapel Hill, North Carolina
Greenhouse, Samuel W.	NIH, Bethesda, Md.

LIST OF ATTENDEES (cont'd.)

Grubbs, Frank E.	BRL, Aberdeen Proving Ground
Harris, Bernard	University of Wisconsin
Harrison, Joseph O., Jr.	RAC, McLean, Va.
Harshbarger, Boyd	VPI, Blacksburg, Va.
Hershner, Ivan R., Jr.	OCRD, Washington, D. C.
Hogben, David	NBS, Gaithersburg, Md.
Horner, Theodore W.	Booz-Allen Applied Res., Bethesda, Md.
Isaac, Gerhard J.	Fitzsimons General Hospital Denver, Colorado
Jebe, Emil H.	USATAC, Ann Arbor, Michigan
Jenkins, W. R., Jr.	WSMR, New Mexico
Jessup, Gordon L., Jr.	Fort Detrick
Johnson, Norman L.	University of North Carolina Chapel Hill, North Carolina
Johnson, Regina C.	University of North Carolina Chapel Hill, North Carolina
Joiner, Brian	NBS, Gaithersburg, Md.
Jones, Marian W.	Fort Detrick
Karp, Abraham E.	NBS, Washington, D. C.
Kniss, James R.	BRL, Aberdeen Proving Ground
Kopec, Casimir S.	NBS, Washington, D. C.
Kruskal, William	University of Chicago
Ku, Harry H.	NBS, Gaithersburg, Md.
Kurkjian, Badrig M.	Harry Diamond Laboratories
Lachenbruch, P. A.	University of North Carolina Chapel Hill, North Carolina
Lancaster, Leslie E.	Center for Naval Analyses, Arlington, Va.
Lavin, George I.	Aberdeen Proving Ground, Md.
LeMaster, Vaughn	Ammunition Proc. & Supply Agency Joliet, Ill.
Levy, Bert	Harry Diamond Laboratories
Lucas, H. L., Jr.	North Carolina State at Raleigh, N. C.
Lum, Harry	Fort Detrick
Lum, Mary D.	Wright-Patterson AFB, Ohio
Lundegard, Robert J.	ONR, Washington, D. C.
Mah, V. W.	Radford College, Radford, Va.
Maloney, Clifford J.	NIH, Bethesda, Md.
Mandel, John	NBS, Washington, D. C.
Mann, Henry B.	University of Wisconsin
Martin, Cyrus A.	Fort Belvoir
McKeague, Robert I., Jr.	Army Munitions Command, Dover, N. J.
Mellinger, J. J.	Army Personnel Res Office Washington, D. C.

LIST OF ATTENDEES (cont'd.)

Mizuki, Mikiso	CMRD, Eou Gollie, Florida
Mohler, Lynn S.	BRL, Aberdeen Proving Ground
Myers, R. H.	VPI, Blacksburg, Va.
Natrella, Mary G.	NBS, Gaithersburg, Md.
Nickens, Paul B.	BRL, Aberdeen Proving Ground
Niehl, Betsy	Army Personnel Res. Office Washington, D. C.
Orleans, Beatrice S.	Naval Ship Systems Command Washington, D. C.
Overholt, John L.	Center for Naval Analysis, Arlington, Va.
Owen, William B.	MCRD-DPSR-ODCSPER DA, Alexandria, Virginia
Pabst, William	Naval Ordnance Systems Command Washington, D. C.
Parks, Albert B.	Harry Diamond Laboratories
Pettigrew, Karen D.	NIH, Bethesda, Md.
Pettigrew, H. M.	NIH, Bethesda, Md.
Provost, Robert G.	Redstone Arsenal
Raybold, Robert C.	NBS, Gaithersburg, Md.
Ricciardi, Anthony J.	Picatinny Arsenal
Robbins, B. L.	STAG, Bethesda, Md.
Roberson, J. D.	Bell Aerosystems Co., Tucson, Arizona
Roseberry, Thomas D.	CEIR, Dugway, Utah
Rosenblatt, Joan R.	NBS, Gaithersburg, Md.
Santone, Louis C.	NBS, Gaithersburg, Md.
Schnell, Raymond B.	Edgewood Arsenal
Schultz, E. Fred, Jr.	Agr. Res. Serv., USDA, Beltsville, Md.
Selig, S. M.	Office of Naval Research, Washington, DC
Sen, P. K.	University of North Carolina Chapel Hill, North Carolina
Sigman, M.	NBS, Washington, D. C.
Sloane, Harry S.	Dugway Proving Ground
Snowden, Shirley E.	Fort Detrick
Solomon, Herbert	George Washington University
Speckman, Janace A.	NBS, Gaithersburg, Md.
Starr, Selig	OCRD-ARO, Washington, D. C.
Stephanides, Agatha S.	STAG, Bethesda, Md.
Tarnove, Theodore L.	Edgewood Arsenal
Thomas, Donald G.	NIH, Bethesda, Md.
Thrall, Robert M.	University of Michigan
Tsutakawa, Robert K.	University of Chicago
Ullman, M. B.	Bureau of the Budget, Washington, D. C.

LIST OF ATTENDEES (cont'd.)

Varley, Thomas C.	ONR, Washington, D. C.
Wampler, Roy H.	NBS, Gaithersburg, Md.
Weinstein, Joseph	Fort Monmouth
West, Larry L.	Fort Lee, Virginia
Wiley, S. R.,	Army Map Service, Bethesda, Md.
Williams, Arthur L.,	STAG, Bethesda, Md.
Willoughby, Weldon F.	BRL, Aberdeen Proving Ground
Wolman, William	Commerce Department, Public Roads, Washington, D.C.
Youden, W. J.	Consultant, Washington, D. C.
Zelen, Marvin	NIH, Bethesda, Md.
Zirkel, Paul R.	BRL, Aberdeen Proving Ground

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
U. S. Army Research Office-Durham
Box CM. Duke Station
Durham, N. C. 27706

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP
NA

3. REPORT TITLE
PROCEEDINGS OF THE TWELFTH CONFERENCE ON THE DESIGN OF EXPERIMENTS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Interim Technical Report

5. AUTHOR(S) (Last name, first name, initial)

6. REPORT DATE
May 1967

7a. TOTAL NO. OF PAGES
428

7b. NO. OF REFS

8a. CONTRACT OR GRANT NO.

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
ARO-D Report 67-2

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES
This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of the U. S. Army Research Office-Durham.

11. SUPPLEMENTARY NOTES
NONE

12. SPONSORING MILITARY ACTIVITY
Army Mathematics Steering Committee on behalf of Office of the Chief of Research and Development

13. ABSTRACT
This is the technical report resulting from the 12th Conference on the Design of Experiments in Army Research, Development and Testing. It contains 24 papers presented at that meeting. These treat various Army statistical and design problems.

14. KEY WORDS
operations research
estimation of error rates
data analysis
statistical analysis of rubber components
probability of survival of subterranean targets
factor analysis
confidence limits
reliability
biological systems
analysis of non-experimental studies
regression analysis
predictors
sample censoring

DD FORM 1 JAN 64 1473

Unclassified

Security Classification