AD A042746

ARO Report 77-2

PROCEEDINGS OF THE TWENTY-SECOND CONFERENCE ON THE DESIGN OF EXPERIMENTS IN ARMY RESEARCH DEVELOPMENT AND TESTING

DDC
RECEIVED
AUG 11 1977
C

Approved for public release; distribution unlimited. The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

AD NO.
DDC FILE COPY

Sponsored by

The Army Mathematics Steering Committee

on Behalf of

THE CHIEF OF RESEARCH, DEVELOPMENT AND ACQUISITION

U. S. Army Research Office

Report No. 77-2
June 1977

PROCEEDINGS OF THE TWENTY-SECOND CONFERENCE ON THE DESIGN OF EXPERIMENTS

Sponsored by the Army Mathematics Steering Committee

HOST

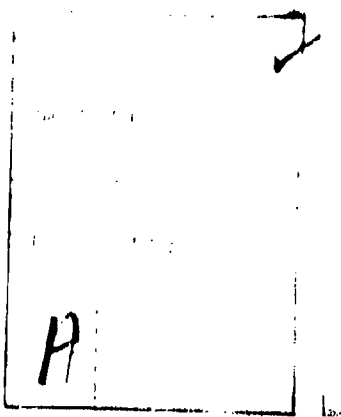
Harry Diamond Laboratories

Adelphi, Maryland

20-22 October 1976

Approved for public release; distribution unlimited.
The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, North Carolina



FOREWORD

The Harry Diamond Laboratories located in Adelphi, Maryland served as the site for the 22nd Conference on the Design of Experiments in Army Research, Development, and Testing held 20-22 October, 1976. This Army agency co-hosted the first three conferences in this series with the National Bureau of Standards when it was located there. It was a pleasure to meet in the new quarters of the Harry Diamond Laboratories and take advantage of their excellent facilities. Planning for these meetings requires much effort and attention to detail and we are indebted to Dr. Joseph Kirschner who served as Chairman for Local Arrangements and was ably assisted by Grace Frazier and Steven Kimmel. We are pleased that Colonel Thomas McGregor, Commanding Officer of the Laboratories opened the Conference and welcomed the participants. We look forward to meeting at the Laboratories again in the future.

It is traditional to have invited speakers give essentially expository talks on topics of current interest in statistics and probability. There is also an attempt to provide talks that are somewhat consistent with the theme of the mission of the Army installation at which the annual Conference is held. This confluence of purposes was achieved. The first talk was given by Professor J. Stuart Hunter of Princeton University on "The Measurement Process." The crux of this talk was measurement when data is available over time such as in air pollution studies and the speaker presented two different models by which this could be accomplished. Later in the first morning Professor Benjamin S. Blanchard of Virginia Polytechnic Institute and State University gave a talk on, "Management of Reliability." The reliability theme pervades many Army installations and this is so at the Harry Diamond Laboratories. On the afternoon of the second day there were two sessions for invited speakers and each was devoted to a very current topic in statistics where each topic has a fast developing literature. The first speaker was Dr. Carl N. Morris of the RAND Corporation who spoke on, "Stein's Estimator, Its Generalizations and Its Applications." This was followed by Professor Robert Hogg of the University of Iowa who spoke on, "Robust Statistical Procedures." The subject matter in both of these talks has wide ranging applications in a number of diverse activities of the Army. On the morning of the last day of the meetings Professor Nozer D. Singpurwalla of the George Washington University spoke on, "Accelerated Life Testing." This topic has a long history in Defense Department programs and is still a quite active subject for statistical investigations.

The audience consisted of a large number of participants from Army installations, other government agencies, and a number of investigators from universities. A major purpose of the conference is to bring together those engaged in scientific work in Army installations with other investigators. This interaction has been going on successfully since the inception of the program and it continued at this Conference. Statisticians and others in Army installations discuss their work at technical sessions and clinical sessions at each Annual Conference. For this Conference

PRECEDING PAGE NOT FILMED
BLANK

there were eight technical sessions comprising eighteen papers and four clinical sessions. At the clinical sessions a panel of experts responds to problems raised by those in Army installations who have usually given advance manuscript copies to the panelists. Besides the technical aspects, these sessions provide a source for initiating future collaboration between scientists in Army installations and those in university life.

On the evening of the first day of the Conference a banquet is held at which the Samuel S. Wilks Memorial Award of the American Statistical Association and the Department of the Army is presented. At this meeting the twelfth award was presented to Dr. Solomon Kullback, Profesor Emeritus of Statistics at the George Washington University. The award was made by Dr. Joan Rosenblatt, Chairman of the Wilks Award Committee. Professor Kullback was cited for substantive contributions to both the theory and the application of statistics, including his work on multidimensional contingency table analysis and cryptanalysis, and his outstanding contributions in the application of statistics in the service of the Nation.

The Army Mathematics Steering Committee sponsors these meetings on behalf of the Office of the Chief of Research and Development and Acquisition to bring new developments in statistics to Army scientists and engineers and to expose them to thinking that could be profitable to them in the execution of their missions. The Committee has asked that the Proceedings of the Conference be published and issued Army wide and to other scientific communities.

At the beginning of each calander year the Program Committee for these conferences is selected and meets in Washington, D.C. to suggest areas of interest, to outline a program, and to suggest speakers for the meeting to be held later that year. I would like to express my appreciation to Dr. Frank Grubbs, Program Chairman for this year's Committee and to Dr. Douglas Tang, Chairman of the Subcommittee on Probability and Statistics, Army Mathematics Steering Committee, for their efforts and great help. My thanks also go to other committee members involved in developing this year's program: Drs. Walter D. Foster, Bernard Harris, Joseph M. Kirschner, Badrig Kurkjian, Clifford J. Maloney, Robert J. Launer, Douglas B. Tang. Dr. Francis G. Dressel, Program Committee Secretary, as always was helpful in many ways in making sure the program was a success. Thus, many helped in guiding this Conference to a successful conclusion and this is very much appreciated.

Herbert Solomon
Conference Chairman

TABLE OF CONTENTS

Title	Page
Foreword	iii
Table of Contents	v
Agenda	vii
Management of Reliability, Availability, and Maintainability (RAM), Benjamin S. Blanchard	1
Problems in Analyzing Pharmacokinetic Data, Carl Peck and Alan Hopkins	11
Effects of Dietary Bran and Cellulose on Serum Lipids, Charlotte M. Heggie, Daniel H. Connor, Frank A. Franklin, Jr., Walter D. Foster, Samuel M. Wylde, Joe M. Blumberg	23
Analysis of an Error-time Response Performance, Michael Hacskaylo	31
Reliability Analysis of Airfield Lighting Systems, Frank Kuo and Edward S. Lindow	47
Simplified Method for Determining Approximate Lower Confidence Bounds of a System Whose Posterior Reliability Distribution is Described as a Beta, Louis M. Iannuzzelli and Robert Dostal	65
Evaluation of Gunner Errors Through Time Series Analysis, Latricha Greene and John Howerton	73
A Method for Determining Pairwise Contrasts from a Friedman Two-way Layout Based on a Theorem by Marascuilo, Jimmie C. DeLoach and Eugene F. Dutiot	91
Estimate of Reliability in the Stress-strength Model, Asit P. Basu	97
Underlying Probability Distribution of Gun Tube Fatigue Life, Ronald L. Racicot	111
Estimation and Effect of Noise Correlation on Variance Estimation from Moving Arc Smoothing, Paul H. Thrasher	125

Robust Outlier Detection in Trajectory Data Reduction
William S. Agee and Robert H. Turner 143

Table Look-up and Interpolation for a Normal Random Number Generator
William L. Shepherd and John N. Hynes 153

Eigenvectors Analysis of Empirical Data Versus Utilization of Standard Functions
Oskar M. Essenwanger 165

Induction on a Markov Chain
Richard M. Brugger 177

Markov and Path Dependent Processes Applied to Continuous Sampling Plans in Tandem
David L. Arp 187

Interval Estimation for Empirical Bayes Generalizations of Stein's Estimator
Carl Morris 219

Robust Statistical Procedures
Robert V. Hogg 247

Estimating Reliability from Small Samples
Donald W. Rankin 257

Estimation and Prediction of Confidenced Reliable Life from Small Sample Sizes
Eugene E. Coppola 297

Sequential Allocation of Observations in the Exponential Selection Problem
Robert M. Wharton and R. Srinivasan 311

Maximum Likelihood Estimation of 12D for Inoculated Packs
Edward W. Ross, Jr. 317

Confidence Bounds for the General Linear Model
J. Richard Moore and Malcolm S. Taylor 327

List of Attendees 339

AGENDA

THE TWENTY-SECOND CONFERENCE ON THE DESIGN OF EXPERIMENTS IN ARMY RESEARCH, DEVELOPMENT AND TESTING

20-22 October 1976

Harry Diamond Laboratories

***** Wednesday, 20 October *****

0815-0915 Registration -- Lobby of the Administration Building: Building 205

0915-1215 GENERAL SESSION I -- Auditorium of the Administration Building

CALLING OF CONFERENCE TO ORDER

Mr. Joseph Kirshner, Chairman on Local Arrangements, Harry Diamond Laboratories, Adelphi, Maryland

WELCOMING REMARKS

Colonel Thomas McGregor, Commanding Officer, Harry Diamond Laboratories, Adelphi, Maryland

CHAIRMAN OF SESSION I

Dr. Frank E. Grubbs, Program Committee Chairman, Aberdeen Proving Ground, Maryland

THE MEASUREMENT PROCESS

Professor J. Stuart Hunter, School of Engineering and Applied Science, Princeton University, Princeton, New Jersey

1030-1100 BREAK

1100-1215 GENERAL SESSION I (CONTINUED)

MANAGEMENT OF RELIABILITY

Professor Benjamin S. Blanchard, Jr., Engineering Extension Division, Virginia Polytechnic Institute and State University, Blacksburg, Virginia

1215-1315 LUNCH -- HDL Cafeteria

***** Wednesday *****

1315-1445 CLINICAL SESSION* A -- Auditorium of Building of 205

CHAIRMAN

Robert L. Launer, US Army Research Office, Research Triangle Park, North Carolina

PANELISTS

Seymour Geisser, School of Statistics, University of Minnesota, Minneapolis, Minnesota

Robert V. Hogg, The University of Iowa, Department of Statistics, Iowa City, Iowa

J. Stuart Hunter, School of Engineering and Applied Science, Princeton University, Princeton, New Jersey

Herbert Solomon, Department of Statistics, Stanford University, Stanford, California

PROBLEMS IN TESTING PHARMACOKINETIC MODELS

LTC Carl C. Peck and L. A. Hopkins, Blood Research Division, Department of Surgery, Letterman Army Institute of Research, Presidio of San Francisco, California

DIETARY BRAN AND CELLULOSE: EFFECTS ON SERUM LIPIDS

Walter D. Foster, Charlotte M. Heggi, Daniel H. Conner, Armed Forces Institute of Pathology; Frank A. Franklin, Jr., Walter Reed Army Medical Center; Samuel M. Wylde, Ener-G-Foods, Inc.; Joe M. Blumberg, Oscar B. Hunter Memorial Laboratory, Washington, DC

1315-1445 TECHNICAL SESSION 1 -- Room 2G016

CHAIRMAN

Langhorne P. Withers, US Army Operational Test and Evaluation Agency, Falls Church, Virginia

ANALYSIS OF AN ERROR-TIME RESPONSE PERFORMANCE

Michael Hacskaylo, Night Vision Laboratory, USA Electronics Command, Ft. Belvoir, Virginia

***** Wednesday *****

AN EXPERIMENTAL DESIGN TO DETERMINE THE FREQUENCY DISTRIBUTION OF LASER RADAR (LADAR) RETURN SIGNAL VOLTAGES

Jerry W. Vickers, Systems Evaluation, Aeroballistics Directorate, USA Missile R&D Command, Redstone Arsenal, Alabama

1445-1515 BREAK

1515-1645 CLINICAL SESSION B -- Auditorium of Building 205

CHAIRMAN

Joan R. Rosenblatt, Statistical Engineering Laboratory, National Bureau of Standards, Washington, DC

PANELISTS

A. Clifford Cohen, Institute of Statistics, University of Georgia, Athens, Georgia

Frank E. Grubbs, Aberdeen Proving Ground, Maryland

Bernard Harris, Mathematics Research Center, University of Wisconsin, Madison, Wisconsin

Nozer D. Singpurwalla, Department of Operations Research, George Washington University, Washington, DC

RELIABILITY ANALYSIS OF AIRFIELD LIGHTING SYSTEMS

Frank Kuo and Ed Lindow, Construction Engineering Research Laboratory, Champaign, Illinois

SIMPLIFIED METHOD FOR DETERMINING APPROXIMATE LOWER CONFIDENCE BOUNDS OF A SYSTEM WHOSE RELIABLITY FUNCTION IS DESCRIBED AS A BETA

Louis M. Iannuzzelli and R. Dostal, HQ, USA Armament Command, Rock Island, Illinois

1515-1645 TECHNICAL SESSION 2 -- Room 2G016

CHAIRMAN

Gertrude Weintraub, Picatinny Arsenal, Dover, New Jersey

EVALUATION OF GUNNER ERRORS THROUGH TIME SERIES ANALYSIS

Latricha Greene and John Howerton, Systems Evaluation, Aeroballistics Directorate, USA Missile R&D Command, Redstone Arsenal, Alabama

***** Wednesday *****

RANGE INSTRUMENTATION POSITION ACCURACY

F. L. Carter, Dugway Proving Ground, Dugway, Utah

1830 - SOCIAL HOUR AND BANQUET -- Hampshire Inn

PRESENTATION OF THE SAMUEL S. WILKS MEMORIAL AWARD

Dr. Frank E. Grubbs, Master of Ceremonies

***** Thursday, 21 October *****

0830-1010 CLINICAL SESSION C -- Auditorium of Building 205

CHAIRMAN

A. Clifford Cohen, Institute of Statistics, University of Georgia, Athens, Georgia

PANELISTS

Robert Bechhofer, Department of Operations Research, Cornell University, Ithaca, New York

Seymour Geisser, School of Statistics, University of Minnesota, Minneapolis, Minnesota

Robert V. Hogg, The University of Iowa, Department of Statistics, Iowa City, Iowa

J. Richard Moore, US Army Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

EXPERIMENTAL DESIGN FOR LABORATORY EVALUATION OF IMAGING SYSTEMS

R. Flaherty, J. Palmer and F. Shields, Night Vision Laboratory, USA Electronics Command, Ft. Belvoir, Virginia

A METHOD FOR DETERMINING PAIRWISE CONTRASTS FROM A FRIEDMAN TWO-WAY LAYOUT BASED ON A THEOREM BY MARASCUILO

Jimmie C. Deloach and Eugene Dutoit, USA Infantry Center, Ft. Benning, Georgia

0830-1010 TECHNICAL SESSION 3 -- Room 2G016

CHAIRMAN

J. Bart Wilburn, Jr., I&M Branch, US Army Electronic Proving Ground, Ft. Huachuca, Arizona

***** Thursday *****

ESTIMATE OF RELIABLITY IN THE STRESS-STRENGTH MODEL

Asit P. Basu, University of Missouri-Columbia, Department of Statistics, Columbia, Missouri

UNDERLYING PROBABILITY DISTRIBUTION OF GUN TUBE FATIGUE LIFE

Ronald L. Racicot, Watervliet Arsenal, Watervliet, New York

FAILURE PREDICTION OF FINITE FLAWED CERAMIC PLATES UNDER COMBINED STRESSES

Donald M. Neal, Army Materiel and Mechanics Research Center, Watertown, Massachusetts

1010-1040 BREAK

1040-1215 CLINICAL SESSION D -- Auditorium of Building 205

CHAIRMAN

Clifford J. Maloney, Bureau of Biologics, FDA, Bethesda, Maryland

PANELISTS

Robert Bechhofer, Department of Operations Research, Cornell University, Ithaca, New York

G.E.P. Box, Department of Statistics, University of Wisconsin, Madison, Wisconsin. Representing the Mathematics Research Center.

Bernard Harris, Mathematics Research Center, University of Wisconsin, Madison, Wisconsin

Herbert Solomon, Department of Statistics, Stanford University, Stanford, California

ESTIMATION AND EFFECT OF NOISE CORRELATION ON VARIANCE ESTIMATION FROM MOVING ARC SMOOTHING

Paul H. Thrasher, Quality Assurance Office, White Sands Missile Range, New Mexico

1040-1215 TECHNICAL SESSION 4 -- Room 2G016

CHAIRMAN

Walter D. Foster, Armed Forces Institute of Pathology, Washington, DC

***** Thursday *****

ROBUST OUTLIER DETECTION IN TRAJECTORY DATA REDUCTION

Robert H. Turner and William S. Agee, Analysis and Computation Division, National Range Operations Directorate, White Sands Missile Range, New Mexico

ON THE UPWARD CONTINUATION OF FIRST DERIVATIVES OF THE ANOMALOUS GRAVITY POTENTIAL UNDER CONSIDERATION OF A SUITABLE DATA BASE

H. Baussus von Luetzow, USA Engineer Topographic Laboratories, Ft. Belvoir, Virginia

COMPARISON OF ERROR RATES AND MISCLASSIFICATION PROBABILITIES USING BINOMIAL AND BAYESIAN MODELS FOR PERSONNEL CLASSIFICATION

Frederick H. Steinheiser, Jr. and Kenneth I. Epstein, USA Research Institute, Arlington, Virginia

1215-1315 LUNCH -- HDL Cafeteria

1315-1415 TECHNICAL SESSION 5 -- Auditorium of Building 205

CHAIRMAN

Joseph S. Tyler, Jr., Chemical Research Laboratory, Biophysics Laboratory, US Army Edgewood Arsenal, Edgewood, Maryland

TABLE LOOK UP AND INTERPOLATION FOR A NORMAL RANDOM NUMBER GENERATOR

William L. Shepherd and John N. Hynes, Systems Management Division, Instrumentation Directorate, White Sands Missile Range, New Mexico

EIGENVECTORS ANALYSIS OF EMPIRICAL DATA VERSUS UTILIZATION OF STANDARD FUNCTIONS

Oskar M. Essenwanger, Physical Sciences Directorate, USA Missile RD&E Laboratory, USA Missile Command, Redstone Arsenal, Alabama

1315-1415 TECHNICAL SESSION 6 -- Room 2G016

CHAIRMAN

Malcolm S. Taylor, US Army Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

***** Thursday *****

INDUCTION ON A MARKOV CHAIN

Richard M. Brugger, RAM Assessment Division, USA Armament Command, Rock Island, Illinois

MARKOV DEPENDENT PROCESSES AND CONTINUOUS SAMPLING PLANS IN TANDEM

David L. Arp, Naval Weapons Center, China Lake, California

1415-1645 GENERAL SESSION II -- Auditorium of Building 205

CHAIRMAN

Dr. Douglas B. Tang, Department of Biostatistics/Applied Mathematics, Division of Biometrics and Medical Information Processing, Walter Reed Army Institute of Research, Washington, DC

STEIN'S ESTIMATOR, ITS GENERALIZATIONS, AND ITS APPLICATIONS

Dr. Carl N. Morris, Rand Corporation, Santa Monica, California

1515-1545 BREAK

1545-1645 GENERAL SESSION II (CONTINUED)

ON ROBUST STATISTICAL PROCEDURES

Professor Robert V. Hogg, The University of Iowa, Department of Statistics, Iowa City, Iowa

***** Friday, 22 October 1976 *****

0830-1015 TECHNICAL SESSION 7 -- Room 2G016

CHAIRMAN

Gerald Andersen, Battlefield System Integration Directorate, Alexandria, Virginia

ESTIMATING RELIABILITY FROM SMALL SAMPLES

Donald W. Rankin, Army Materiel Test and Evaluation Directorate, White Sands Missile Range, New Mexico
(Jerry Short will present the paper)

***** Friday *****

ESTIMATION AND PREDICTION OF CONFIDENCED RELIABLE LIFE FROM SMALL SAMPLE SIZES

Eugene E. Coppola, Benet Weapons Laboratory, Watervliet Arsenal, Watervliet, New York

SEQUENTIAL ALLOCATION OF OBSERVATIONS IN THE EXPONENTIAL SELECTION PROBLEM

Robert M. Wharton and R. Srinivasan, Thomas Jefferson University and Temple University (respectively), Philadelphia, Pennsylvania

0830-1015 TECHNICAL SESSION 8 -- Auditorium of Building 205

CHAIRMAN

Bruce J. McDonald, Probability & Statistics Program, Office of Naval Research, Arlington, Virginia

MAXIMUM LIKELIHOOD ESTIMATION OF 12D FOR INOCULATED PACKS

Edward W. Ross, Jr., USA Natick R&D Command, Natick, Massachusetts

CONFIDENCE BOUNDS FOR THE GENERAL LINEAR MODEL

Malcolm S. Taylor and J. Richard Moore, USA Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

COST OF LIVING INDEX

K. S. Banerjee, Statistics and Computer Science Division, University of Delaware, Newark, Delaware

1015-1045 BREAK

1045-1215 GENERAL SESSION III -- Auditorium of Building 205

CHAIRMAN

Professor Herbert Solomon, Chairman of the Conference, Department of Statistics, Stanford University, Stanford, California

OPEN MEETING OF THE AMSC SUBCOMMITTEE ON PROBABILITY AND STATISTICS

Dr. Douglas B. Tang, Department of Biostatistics and Applied Mathematics Division, Biometrics and Medical Information Processing, Walter Reed Army Institute of Research, Washington, DC

ACCELERATED LIFE TESTING

Professor Nozer D. Singpurwalla, Department of Operations Research, George Washington University, Washington, DC

1215-1315 LUNCH

MANAGEMENT OF

RELIABILITY, AVAILABILITY, AND MAINTAINABILITY (RAM)

Benjamin S. Blanchard
College of Engineering
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

ABSTRACT. Our systems and equipment in the field have become more and more complex; are not operationally available a good percent of the time; require extensive maintenance and support; and are quite costly. One of the causes for this dilemma is the emphasis that has been placed on performance and advanced technology, while at the same time very little consideration has been given to reliability, availability, and maintainability (RAM). More recently, a concerted effort has been initiated to recognize RAM as necessary parameters of system/equipment design and development. Military specifications and standards have been generated and RAM requirements (to varying degrees) have been formally applied on many programs. Although this effort has forced the recognition of RAM to a considerable extent, many program implementation problems currently exist and our systems/equipment in the field are still experiencing significant difficulties.

In this paper the author has attempted to identify some of the problems associated with current RAM implementation practices, and to recommend some courses of action for improvement in the future. A significant challenge lies ahead if we intend to derive some of the benefits from RAM.

1. INTRODUCTION. Through the past few decades emphasis in the design and development of new systems and equipment has been placed primarily on performance factors, delivery schedules, and initial acquisition price. The pressures associated with increased performance has resulted in a dilemma where many items currently in government inventories are highly complex, inoperative a good percentage of the time, difficult to maintain, and in general too costly to justify. In other words, we have produced a large quantity of systems and equipment with low reliability and poor maintainability characteristics, and the level of support necessary to sustain them operationally is considerable! This in turn has:

a. Threatened the overall availability and operational effectiveness of systems and equipment in the field and hence, the defense of our country (either directly or indirectly).

b. Caused high maintenance work loads and increased logistics support resource requirements.

c. Increased life cycle costs for systems/equipment acquisition and utilization, particularly those costsassociated with system operation and support throughout the life cycle.

The current trends of rising system costs plus inflation, combined with some budgetary shifts from defense to other public sectors, are causing serious concerns relative to our future defense capability.

More recently, an attempt has been initiated to counter these trends through the recognition of some critical "cause and effect" relationships involving reliability, availability, maintainability, performance, logistics support, life cycle cost, etc. Experience has indicated that highly reliable and maintainable systems/equipment are a means for improving operational effectiveness while holding the line on life cycle costs. Reliability and maintainability are indeed characteristics which are inherent in system/equipment design, and the extent to which they are considered has a significant impact on logistics support requirements and life cycle cost. Further, the consideration of reliability and maintainability in the design process must commence at the conceptual phase of system development and extend through detailed full-scale engineering development, test and evaluation, and production. In essence, it has been recognized by many that the conditions noted below should be stressed in the future:

a. Reliability, availability, and maintainability should be considered in the system design and development process on an equivalent basis with performance and other related factors.

b. Logistics support should be considered in the design process and should be closely integrated with reliability, availability, maintainability, and performance considerations.

c. Life cycle cost should be considered as a design parameter (i.e., design to unit acquisition cost, design to unit operation and support cost, design to unit life cycle cost).

A primary objective is to provide the necessary management emphasis in all future system/equipment acquisitions, or modifications for improvement, to ensure that these considerations are addressed at the proper level.

With this objective in mind, it is now appropriate to review current practices, assess the pros and cons of such, and determine the steps necessary to further improve our systems and equipment. The author attempts to accomplish this in the paragraphs below, with the discussion basically focusing on the management of Reliability, Availability, and Maintainability (RAM).

2. CURRENT RAM PRACTICES. Although reliability, availability, and maintainability are recognized in many programs today, the implementation practices associated with these areas still require some improvement. A few characteristic problems as they currently exist in on-going programs are outlined below (not necessarily presented in any specific order).

a. Specification of System Technical Requirements

(1) In many instances, quantitative factors are included in requests for proposals (RFPs) and in contracts as "goals". Consequently,

these factors are indeed treated as goals and not as requirements, and are considered only lightly (if at all) in program implementation.

(2) Quantitative factors are not always specified in meaningful terms. Often, probabilistic values that can not be realistically demonstrated are specified instead of quantitative factors that can be understood, allocated, and verified. For instance, it is questionable that one can adequately verify a 0.9995 reliability requirement when limited quantities of equipment are procured and the test sample is small. Also, it is hard to explain a "0.9995 factor" to a design engineer in a meaningful manner, where a MTBF or MTBM value would be more appropriate. In essence, the mathematical "jargon" sometimes employed is difficult to relate directly to design and is often misunderstood by engineers and management personnel alike.

(3) The application of technical requirements is not always related to specific mission objectives. As a result, it is difficult to determine whether the requirements are too stringent or too loose relative to the ultimate mission need. In many cases mission requirements are not adequately defined early enough in the program, and one can not properly design equipment without a mission profile or scenario of some type; thus, RAM requirements result from a "best-guess" approach which is less than satisfactory.

b. RAM as Design Parameters

Reliability, availability, and maintainability are not recognized as design parameters. Past practices have promoted the concept of "design the system quickly, put it into a test program, and fix it if necessary". RAM have not been truly coupled into the design effort, but designated for measurement or demonstration at the conclusion of full-scale engineering development. This concept has been quite costly, particularly when extensive system/equipment modifications are necessary to meet RAM requirements at this late stage of engineering development.

c. Application of Specifications and Standards

(1) On a number of occasions the "panic" to release specifications for a procurement results in a fragmented document incorporating conflicts and contradictions. RAM is not properly integrated into the overall product. The specification is without doubt one of the most important aspects of a program, but is not always given the necessary level of attention because of the tight schedule requirement to publish something for immediate dissemination purposes. The consequences frequently result in problems occurring throughout the remainder of the program.

(2) Military specifications and standards (e.g., MIL-STDs-470,-471,-781, -785) are often arbitrarily applied to a contract in terms of "blanket coverage" without the tailoring of such to the specific program need. This can result in the application of meaningless requirements, untimely activities and information, too much data of little value,

and high consequential program costs. Specifications and standards should address real time task enforcement, product measurement and control, with less overall dependence on test and demonstration at the end of full-scale engineering development.

d. Structuring of Test Programs

(1) System/equipment testing is accomplished to different environmental profiles than what is actually experienced in the field. Hence, the test results are not necessarily a verification that the intended requirements have (or have not) been met. This relates to the initial inadequate definition of mission profiles or scenarios as discussed in Paragraph 2a(3) above.

(2) Many test and demonstration programs are accomplished at the end of full-scale engineering development after the commitment of production/operation funds. Testing at this stage can only measure the worth of a design configuration at a point when it is too costly and time consuming to make major changes to correct a problem for RAM.

e. Producer Accountability

Producer accountability is generally lacking! If the system/equipment fails in test and demonstration, the policy in some cases has been to discount the failures or to change the standards such that the system will pass. How often is the system/equipment actually rejected by the customer because of failure to pass RAM tests? In such cases, is the producer actually required to initiate the necessary corrective action at his own expense? Perhaps there are some cases where the producer is actually held accountable for his design for RAM; however, in numerous other instances the system/equipment is accepted regardless of the outcome of RAM verification testing.

The problems outlined above are representative of areas where current implementation practices concerning RAM need improvement. In all cases the type of problems indicated have been recognized, and some action is being taken (to varying degrees) in an attempt to improve future system/equipment acquisitions from the RAM standpoint. However, inspite of what is currently underway relative to RAM activities, a great deal of additional effort is required if the objectives of RAM in system/equipment design and development are to be truly realized.

3. CHALLENGES FOR THE FUTURE. At this point the major question is--Are We Serious About Reliability, Availability, And Maintainability? The author firmly believes that we are! However, every effort must be made to preclude or alleviate some of the problem areas mentioned above. It is felt that no new policies per se are necessary, but that a new approach to policy implementation is definitely required. Some key implementation factors and challenges for the future are noted.

a. More front end planning, programming, and budgeting is required relative to the inclusion of RAM factors. In other words, RAM considerations must be addressed in Decision Coordinating Papers (DCPs), Operational Capability Objective (OCOs), Letters Of Agreement (LOAs), Outline Development Plans (ODPs), Required Operational Capability (ROC) documentation, and so on.

Referring to Figure 1, which illustrates the classical program phases, RAM should be initially covered in the conceptual phase of system design and development. The intent is not to be overly stringent relative to the specification of RAM requirements at this stage, but to properly address the major issues pertaining to RAM. In addition, program budgets must reflect RAM decisions--i.e., increase the procurement dollars slightly to acquire the necessary RAM and reduce the support dollars to reflect the corresponding reduction in system support cost.

b. RAM must be treated more as a "discipline" throughout the system/equipment life cycle and in particular, the early design process.

Figure 2 illustrates the system life cycle process and addresses typical RAM activities in each phase of the process. Not only are RAM activities applicable in each major evolutionary stage of system development, but these activities must be closely interrelated throughout! Of particular significance are the decisions pertaining to RAM which are a part of the requirements depicted in Blocks 2 through 8 of Figure 2. Experience has indicated that ultimate system life cycle cost is significantly influenced by design decisions made during the conceptual and validation phases of a program. The overall impact of actions affecting life cycle cost is reflected by the "trend" curve in Figure 3. Further, experience has verified that life cycle cost is highly influenced by RAM, particularly those costs associated with system operation and support. Thus, RAM must be addressed early in the system life cycle if the end product is to be cost effective.

c. Program management for RAM must be significantly strengthened! More specifically:

(1) Realistic and meaningful requirements must be clearly specified early in the system life cycle.

(2) Specifications must be improved and more precisely "tailored" to meet the actual needs. User involvement in the initial preparation of specifications is recommended.

(3) Requests for proposals (RFPs) must leave no doubt that RAM and performance are "equals" in priority and importance.

(4) Program managers must be held technically accountable for RAM as well as for other factors.

(5) Program "checks and balances" must be provided for management control (and audit for compliance) relative to RAM requirements. Formal program reviews and technical design reviews must address major RAM issues.

(6) Integrated test planning is required. As the purpose of testing is to ensure that the system/equipment design meets all stated requirements (including RAM), it is essential that such testing be accomplished in the proper environment. If the test conditions duplicate or exceed the ultimate field environment, the test results

CLASSICAL PROGRAM PHASES

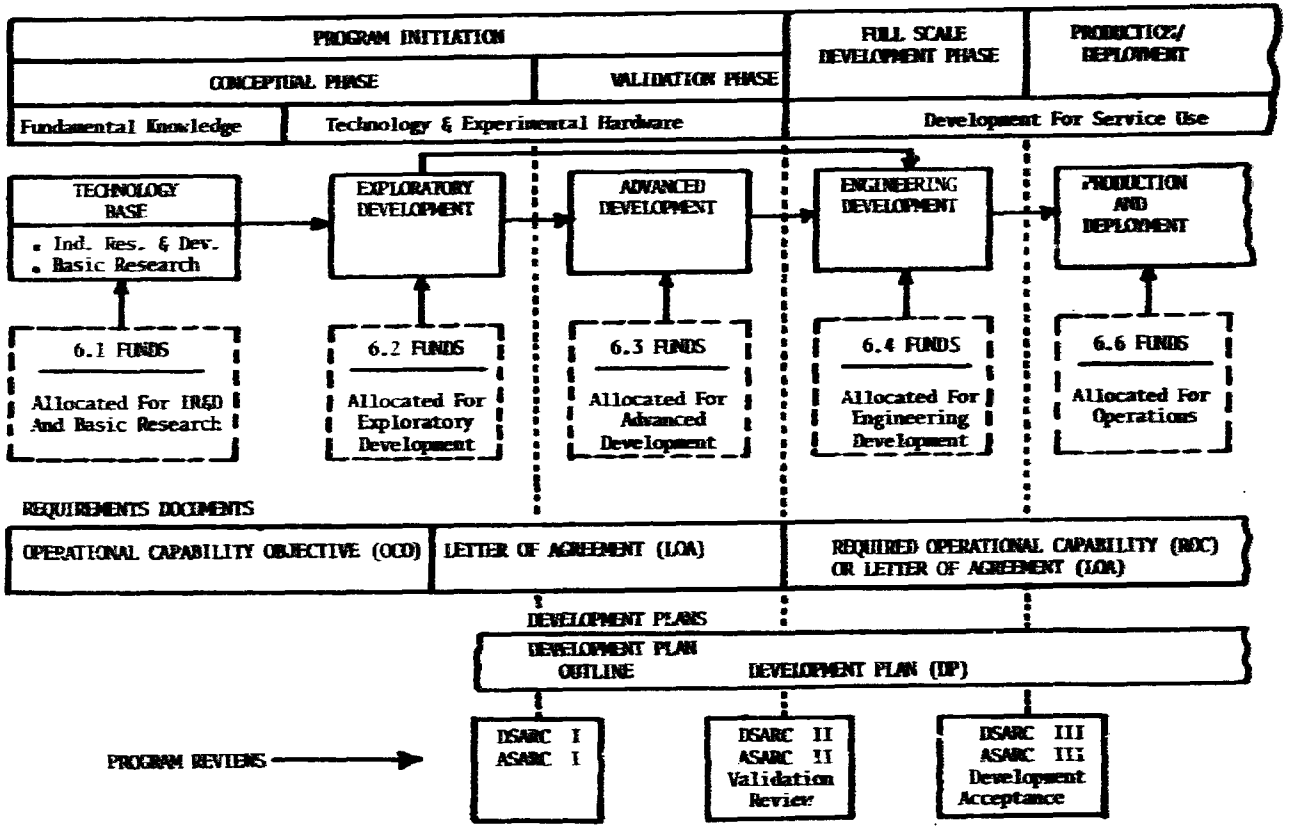


FIGURE 1

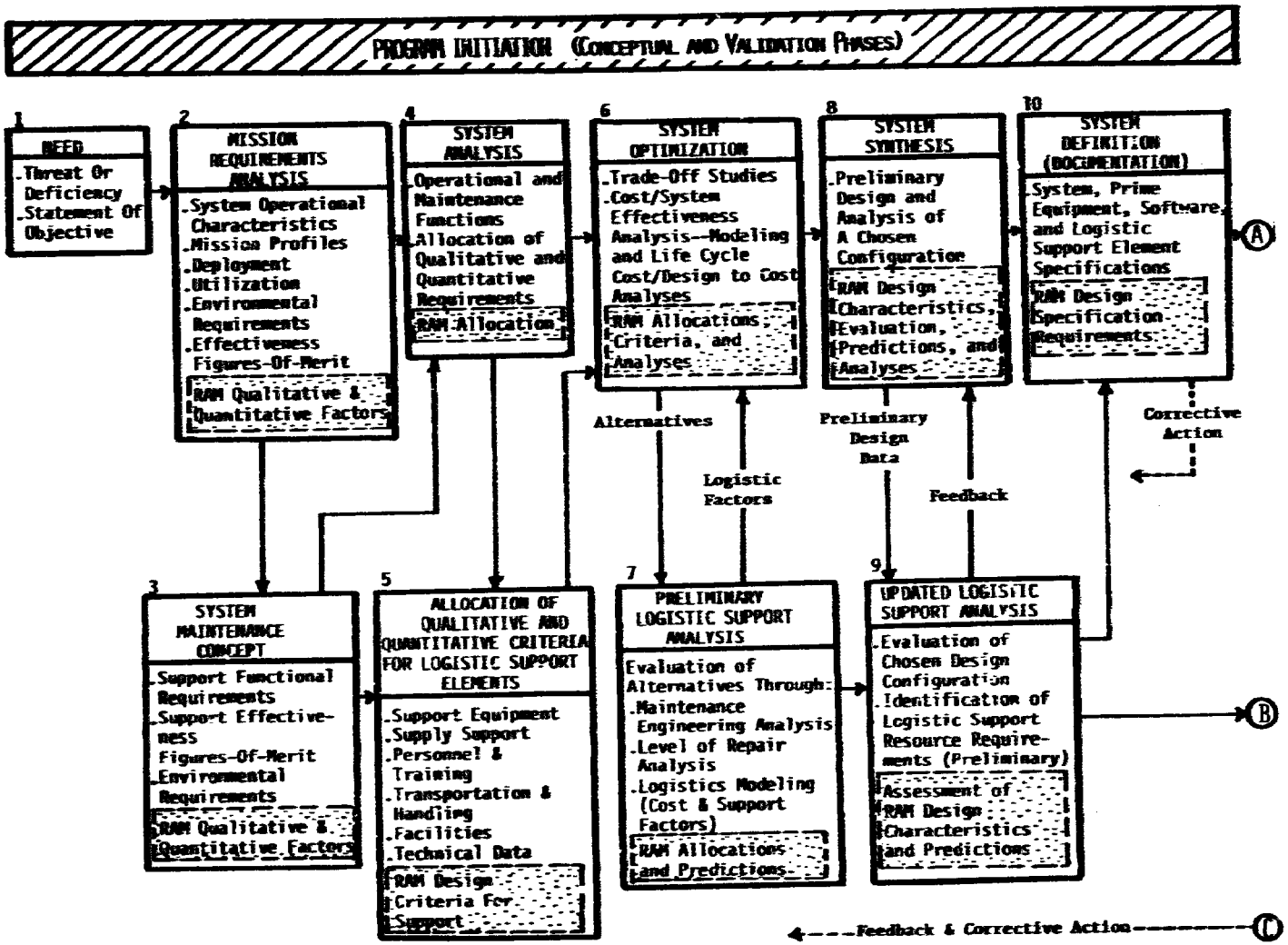


FIGURE 2 - SYSTEM LIFE CYCLE PROCESS (SHEET 1)

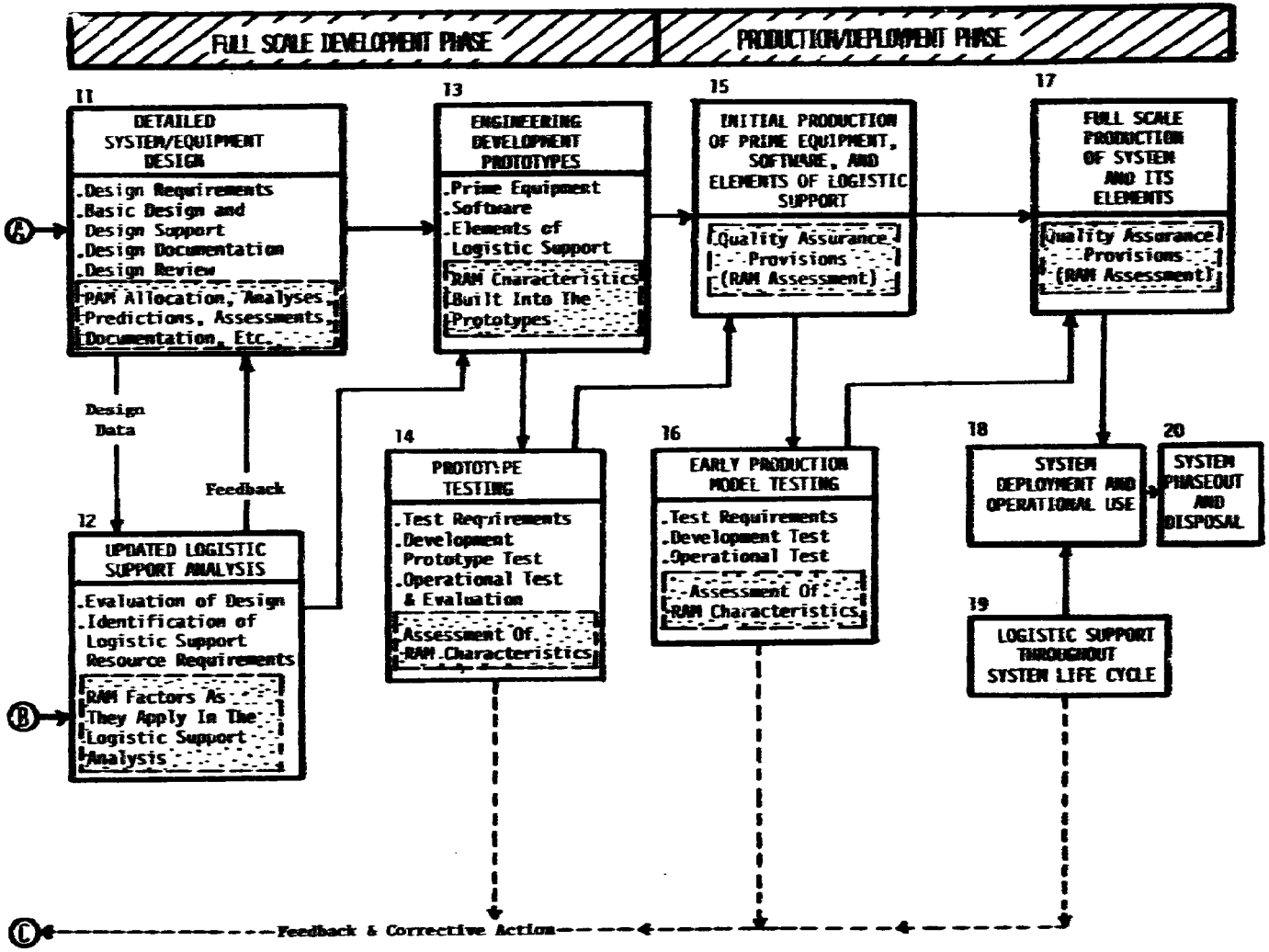


FIGURE 2 - SYSTEM LIFE CYCLE PROCESS (SHEET 2)

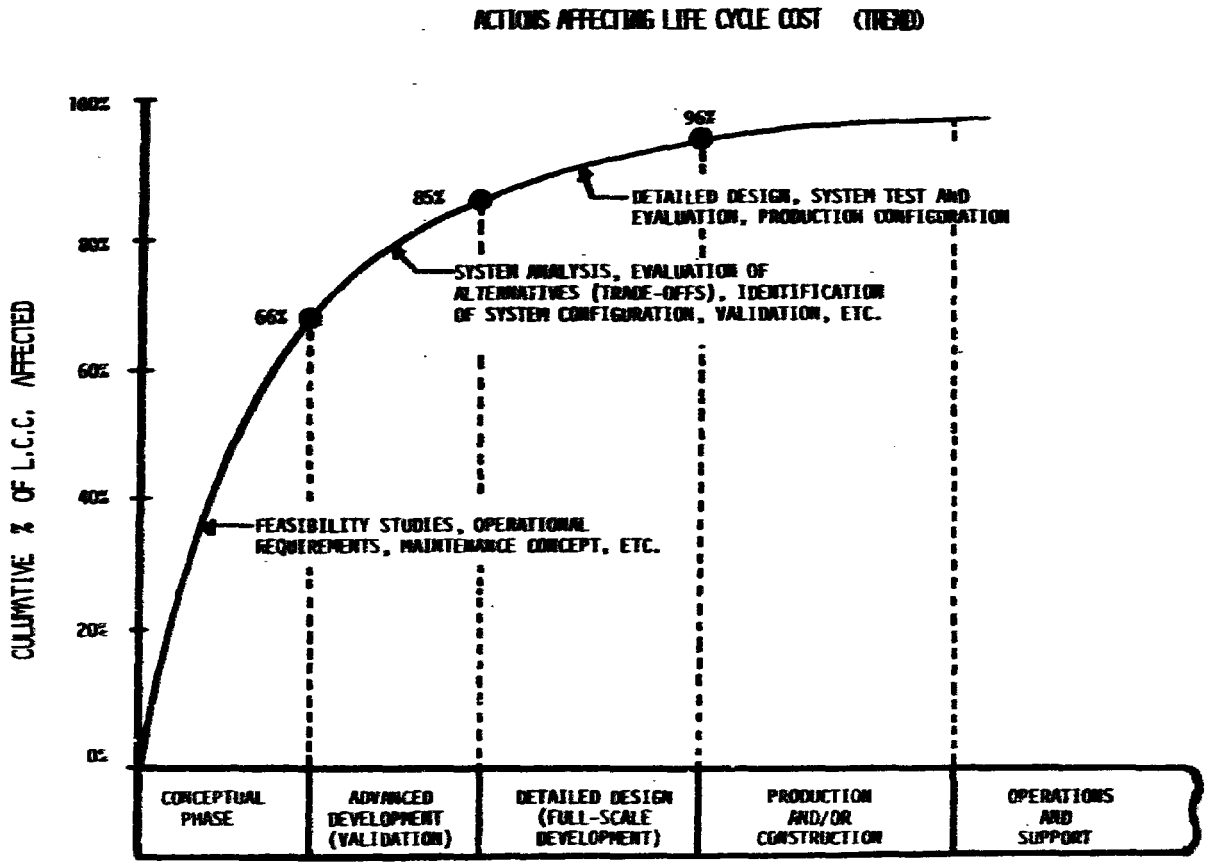

ACTIONS AFFECTING LIFE CYCLE COST (TREND)
CUMMITIVE % OF L.C.C. AFFECTED
100%
80%
60%
40%
20%
0%
96%
85%
66%
DETAILED DESIGN, SYSTEM TEST AND EVALUATION, PRODUCTION CONFIGURATION
SYSTEM ANALYSIS, EVALUATION OF ALTERNATIVES (TRADE-OFFS), IDENTIFICATION OF SYSTEM CONFIGURATION, VALIDATION, ETC.
FEASIBILITY STUDIES, OPERATIONAL REQUIREMENTS, MAINTENANCE CONCEPT, ETC.
CONCEPTUAL PHASE
ADVANCED DEVELOPMENT (VALIDATION)
DETAILED DESIGN (FULL-SCALE DEVELOPMENT)
PRODUCTION AND/OR CONSTRUCTION
OPERATIONS AND SUPPORT

FIGURE 3

will be effective in ensuring that RAM requirements will be met after system/equipment deployment. If not, testing will be ineffective.

Additionally, a number of different tests may be accomplished at different stages in the life cycle. All individual tests must be addressed on an integrated basis to ensure that the desired information is provided at the right time in the system life cycle. Too much testing too early is costly; accomplishing tests too late in the program could be costly; and redundancies in testing may also be costly.

(7) More producer involvement after the system is in operation is desirable. In many instances, the producer should be held responsible for correcting major field deficiencies.

(8) There should be more innovative approaches to better contracting for RAM. One should consider the appropriate use of: penalty/incentive provisions; penalty clauses to cover poor workmanship and design practices; warranties at the piece part level; and meaningful progress payment schedules. The application of the appropriate contractual provisions for RAM requirements should create the desired emphasis relative to RAM.

(9) Strict and timely enforcement of RAM program requirements is essential.

d. Managers and organizations must be educated relative to the benefits derived through the proper level and application of RAM. This is perhaps the greatest challenge, since it is felt that many of the problems experienced in the past could have been avoided had the benefits of RAM been adequately understood and accepted. In addition, with the proper education and understanding, many of the desired objectives mentioned above should be readily attainable.

4. CONCLUSION. The past few decades have led to many advances toward focusing attention on reliability, availability, and maintainability (RAM). The next decade is significant in terms of actual realization of the benefits derived through RAM. The proper levels and applications of RAM are indeed necessary to improve overall system/cost effectiveness at reduced life cycle cost. Addressing the issues outlined in Paragraph 3 is believed to be a step in the right direction and constitutes a major challenge for the future. With educational know-how, persistence, and dedicated effort, it is believed that this challenge can be met.

PROBLEMS IN ANALYZING PHARMACOKINETIC DATA

Carl Peck and
Alan Hopkins
Departments of Surgery and Information Sciences
Letterman Army Institute of Research
Presidio of San Francisco, California

ABSTRACT. Analyzing drug disposition data using pharmacokinetic modeling techniques is a commonly used approach to reducing such data to therapeutically useful facts. However, certain conceptual and statistical problems arise as a result of the data analyst's choice of (1) objectives of the analysis, (2) the class of models to fit the data, (3) the data fitting procedure, (4) the technique(s) for assessing goodness of fit, and (5) ultimately, the most acceptable model. These problems are introduced here along with some current techniques for overcoming them. The advice of the panelists is presented along with our consideration of their recommendations.

1. INTRODUCTION. Dosing decisions in medical therapeutics often involve deciding how much, how frequently, and how long to administer a given drug to a particular patient. Such decisions are rendered much less arbitrary if the therapist has some notion of the time course of drug distribution and elimination from the body, as well as a knowledge of the relationship of these quantitative features of drug disposition to pharmacologic effects. Surprising as it may seem, exacting knowledge of this sort is known for only a small proportion of substances currently used in medical therapeutics. In the main, dosing regimens have been developed on an empirical basis by a trial and error process.

Note. The presentation of this paper at the Conference included examples of problems encountered in analysis of pharmacokinetic data in our laboratory. In order to provide space for comments by the panelists (paraphrased by us) and subsequent discussions, numerical examples are omitted. The interested reader is referred to the paper of Boxenbaum et al.[1] for examples of pharmacokinetic data which typify the issues addressed in this paper.

In recent years a general approach to gathering and organizing drug disposition information has been developed and is frequently referred to as "pharmacokinetics." Pharmacokinetics has been defined by Gibaldi and Perrier as "the study of the time course of drug and metabolite levels in different fluids, tissues, and excreta of the body, and the mathematical relationships required to develop models to interpret such data."[2]

For the purpose of making quantitative therapeutic decisions, a pharmacokinetic analysis of drug data can contribute in several ways. First of all, a model which accurately describes the time course of the drug in the body as well as in particular pools can be quite helpful in choosing dosing size and dosing frequency. This presumes, of course, that the therapist has some notion of desirable upper and lower bounds for drug amounts in the body or pool of interest. The behavior of linear systems under single and multiple dose administration as well as oral ingestion and intravenous infusion is well worked out.[2,3] Certain "derived" parameters, such as "apparent distribution volume," "body clearance," "terminal elimination half-time," and "extent of bioavailability" can be operationally useful in making dosing decisions. Knowledge of the influence of pathologic states on these derived parameters can result in optimal dosing regimens in the face of disease-induced alterations in distribution and elimination.

Secondly, insights into drug-body interactions can be obtained from pharmacokinetic analyses. For example, a mathematically zero-order elimination process implies "saturation" of an elimination mechanism, perhaps a hepatic enzyme-system. Observation that the renal clearance and body clearances of a drug are identical suggests that the kidney is the major elimination organ. A renal clearance which is numerically in excess of glomerular filtration rate implies tubular secretion ± glomerular filtration as mechanisms of renal drug processing. Insights of this nature contribute to therapeutic decision-making by alerting the therapist to special precautions he must take in designing a therapeutic regimen for multiple dosing in a patient with a diseased liver or kidney.

In this paper we wish to summarize some current approaches to analyzing pharmacokinetic data by identifying some problem areas and presenting the responses of panelists to them.

2. PHARMACOKINETIC MODELS: MATHEMATICAL DESCRIPTIONS OF DRUG DISPOSITION. Conceptually, the pharmacokinetic model is usually viewed as a system of inter-connected pools or compartments (Figure 1). The arrows between pools represent drug transfer directions and the symbols "K_{ij}" are interpreted as transfer rates. The drug is considered to be introduced into one of the compartments and body fluid samples are taken from one or more of the pools. Mathematically, the model may be defined as a system of differential equations. Linear differential equations (first order) have been the most fully explored and frequently applied drug disposition models.[3] Although many drugs undergo apparent first order distribution and elimination processes, this is not always the case. Apparent zero order or combinations of zero and first order processes do occur in drug kinetics, which render models of the Michaelis-Menton type applicable.[4] However, for the purposes of this discussion we will confine our attention primarily to the class of linear models.

Integrated solutions to systems of linear differential equations assume a certain simplicity and order. An n-compartment open model (with bidirectional drug transfer between all adjacent pools) yields a linear combination of n-exponentials:

$$D_j = \sum_{i=1}^{n} A_i e^{-\lambda_i t} \qquad \text{Equation 1}$$

where D_j = drug amount or concentration in the j^{th} pool; n = number of compartments; A_i, λ_i = arbitrary parameters of the model which are various algebraic combinations of the original "micro"-rate constants (K_{ij}), volume scalars, and initial conditions.

3. METHODS OF PHARMACOKINETIC DATA ANALYSIS. Development of a pharmacokinetic analysis usually procedes as follows: (1) serial drug concentrations are measured in a body fluid following a dose administration, (2) some procedure is used to choose a class of probable models which are appropriate for the purpose of the analysis, (3) the data are fitted to the models by some procedure resulting in

SOME LINEAR PHARMACOKINETIC MODELS

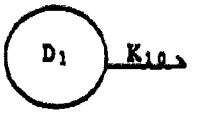
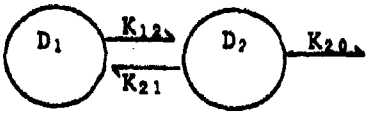
Conceptual Models	Mathematical Models	Generalized Solution
	$\dot{D}_1 = -K_{10}D_1$	$D_1(t) = A_1e^{-\lambda_1 t}$
	$\dot{D}_1 = K_{21}D_2 - K_{12}D_1$ $\dot{D}_2 = K_{12}D_1 - (K_{20}+K_{21})D_2$	$D_1(t) = A_1e^{-\lambda_1 t} + A_2e^{-\lambda_2 t}$

Figure 1

model parameter estimates, (4) an assessment is made of the goodness of fit of the model to the data, and (5) a "most acceptable" model is chosen. The remainder of the discussion is to focus on some problems encountered in steps (2), (3), (4), and (5).

Step 2. Specifying the Class of Probable Models. Choosing the class of models to be considered usually involves a preliminary study of the concentration/time-course data. If a cartesian plot reveals a predominatly linear decay profile, then a zero order model or Michaelis-Menton model is usually considered. Curvilinear decay curves are rendered segmentally linear on log-concentration/time plots if the data behaves as a poly-exponential. The number of straight-line segments can be used as the initial number of exponential terms to be included in the model. In addition, the slo and y-intercepts of these segments can be used as starting point.. .. iterative parameter estimation procedures. Although most pharmacokin. icists procede in this fashion using manual or partially automated graphical procedures, attempts have been made to fully automate this phase of the analysis.[5,6]

A decision must be made regarding the exact form of the mathematical model to use in the data fitting phase. While data may be fitted directly to systems of differential equations,[7,8] the usual practice is to use the integrated form of the model. In the case of linear pharmacokinetic models, this reduces to fitting data to a form of Equation 1. The analyst must also decide whether to parameterize the equation explicitly in the "micro"-rate constants (K_{ij}) or use the "macro"-rate constants (A_i, λ_i). This last issue was addressed by one of the panelists (G.B.) and is discussed below.

Step 3. Fitting the Model to Data. Fitting the model equations to pharmacokinetic data is usually done using an automated least-squares (LS) program such as SAAM[7] or NONLIN.[8] With two exceptions, currently employed pharmacokinetic models are nonlinear with respect to their parameters in their integrated forms and therefore require nonlinear LS data fitting procedures for estimating parameter values. The two exceptions are one-compartment open models with (a) purely zero order elimination or (b) first order elimination (which can be linearized by a $\log_e$ transformation of the entire model). Among problems encountered in this stage of the analysis

are (1) appropriateness of the LS criterion for minimization, especially as regards deviations of the system under study from assumptions inherent in the LS approach and the large influence that aberrant data values can have on the parameter estimates, and, (2) whether and how to "weight" data for the analysis. These two problems constitute part of the requirement for assessing adequacy of the model (addressed by panelist R.H.)

Step 4. Assessing Goodness of Fit. An evaluation of the goodness of fit of the model equation(s) to the data is a highly desirable procedure in pharmacokinetics. The exact form that this assessment takes will depend upon the overall objective of the pharmacokinetic analysis, models used, and the fitting procedure employed. For example, the analyst may be primarily interested in developing a descriptive equation employing "macro"-rate constants to use in computation of "derived" parameters, or his principal intent may be to estimate "macro"-rate constants of a specific compartmental model; these divergent objectives will determine the criteria as well as the technique employed in judging goodness of fit. If a LS data fitting procedure has been employed, use of residual plots and analysis of residuals for their distributional properties is appropriate.[1,9,10] We have employed these techniques to evaluate some of our pharmacokinetic data analyses and found them to be particularly useful. Plots of weighted standardized residuals against drug concentrations reveal patterns which at a glance allow assessment of adequacy of weighting (stabilizing the variance about the regression line), model specificity (search for systematic deviations of residuals from the regression line) and randomness of residual distribution. Further analysis of residuals alone for distributional properties (e.g. mean, median, variance, skewness, kurtosis, specific tests of normality) has been enlightening but not always useful. As pointed out by panelist Dr. R. Hogg, the use of normality tests may constitute too severe a criterion for use in an area where the validity of normal assumptions are in serious question from the outset. In this regard, it was suggested by one of the panelists that the Shapiro-Wilk[11] test for normality might be reasonable.

Step 5. Choosing the Most Acceptable Model. Ultimately, all data analyses must be terminated. This phase in pharmacokinetic data analysis can be a troublesome problem when no clear-cut model emerges more convincingly acceptable than others in the class of models

explored, or when attempts at weighting leave the analyst puzzled about adequacy of various weighting schemes. Analysis "termination criteria" do emerge, however, when the overall objective of the analysis is integrated with the other phases as is developed in the discussion below. It should be noted also that a satisfactory termination of data analysis is closely tied to the adequacy of the overall design of the pharmacokinetic experiment. Optimally, the experimentalist and the data analyst should communicate in the experimental design stage so that sampling times, number of replicates, etc. are designed to "optimize" the information gain from the effort. This translates into a pre-experimental consideration of models to be used in analysis of the data and design of the experimental details so that statistical estimates of model parameters are at minimum variance within the practical constraints of experimental technology and costs.

5. COMMENTS OF PANEL MEMBERS AND DISCUSSION.

Dr. G. E. P. Box: A central issue which is inherently important in each problem area cited above is the overall objective of the exercise. Clear recognition of the goal(s) of a particular pharmacokinetic experiment leads to clarity in the subsequent data analysis.

Discussion: On the surface, these remarks seem almost unnecessary, for the thoughtful data analyst should always have a clear idea of the goals of the exercise. However, Dr. Box correctly detected some vagueness in the objectives of analyzing pharmacokinetic data relevant to the ultimate use of the results. We accept Dr. Box's perspicacious comments and wish to cite some developments in recent pharmacokinetic literature which contribute to clarifying the objectives of pharmacokinetic analyses. While postulating a class of pharmacokinetic models in terms of compartmental schematics with specific inter-compartmental connections is intellectually attractive, the effort required to test, evaluate, and find an acceptable one may be far in excess of that necessary to fulfill the clinical goals of the experiment. If knowledge relevant to making dosing decisions is the principal purpose, then a data analytic approach which concentrates on estimation of macro-parameter models may be adequate. Wagner has recently published a series of articles which argue these points forcefully and which propose simplified data analytic techniques for computing useful pharmacokinetic parameters.[12,13,14]

If the analyst perceives that the objective of the analysis is to provide tools for prediction and for computing "derived" pharmacokinetic parameters, and not to test specific compartmental models which were derived from differential equations, then he is not restricted to exclusive use of the class of poly-exponentials. For example, Wold et al.[15] propose the use of cubic spline procedures for computing area under the curve and terminal drug decay half-time, and give a specific pharmacokinetic example to illustrate the method.

DR. R. V. Hogg: LS data fitting is not the only available option and "robust" statistical procedures should be considered. [In his formal presentation[16] "On Robust Statistical Procedures," Dr.Hogg outlined several possible alternatives to the LS approach to parameter estimation.]

Discussion: Use of robust statistical procedures indeed offers a potential contribution to analysis of pharmacokinetic data. Although these approaches pose computational difficulties, they are attractive both from the point of view of (a) relaxation of the more restrictive normal assumptions inherent in LS procedures and (b) minimization of the effects of "erratic" data (outliers). We have not yet applied any of these approaches to our own problems, although we are aware of one group which has. Frome and Yakata[17] used both LS and least-absolute-deviation criteria in obtaining parameter estimates from the fit of a one-compartment open first order model to a set of pharmacokinetic data.

Dr. S. Geisser: Consideration should be given to the use of the Cp statistic[18] and predictive sample reuse methods[19-21] for assessing goodness of fit and for developing data analysis termination strategies.

Discussion: The Cp statistic was originally derived for use in making decisions about the number of terms to include in linear models where normal assumptions hold. Therefore, use of this approach for deciding among several poly-exponential models must be viewed as an ad hoc procedure, the theoretical basis for which remains unexplored. Nevertheless, the technique is appealing. Given that the "total squared error" computed from a nonlinear regression bears some inexact but semiquantitative relationship to the "true" squared biases and squared random errors, then plotting Cp versus p for various pharmacokinetic models† may yield some

† here p might be considered the number of parameters of the model.

basis for choice.

Use of predictive sample reuse methodology for assessment of different predictive functions is apparently a rather recent development in statistics. The available papers on predictive sample reuse are not easy for the non-statistician to understand, therefore, a brief description of technique in the present context will be given. Like the Cp statistic, a number associated with a given fit of a specific model to data is desired which will allow discrimination between models such that a most reliably predictive model may be identified. This number, call it Γ_p, may be computed using the following "data reuse" approach. Model j is fit to all the data less the first datum by LS and the residual sum of squares is recorded (RSS_1). The procedure is repeated after replacing the first datum and omitting the second data point and the resulting RSS_2 is added to the first. This is repeated by replacing the i^{th} data point and removing the $(i+1)^{th}$ datum and so forth until $\Gamma_p = \sum_{i=1}^{n} RSS_i$. The entire procedure is replicated for each proposed model. Then all Γ_p may be compared and model p*, beyond which Γ_p does not get appreciably smaller, may be chosen as an acceptable model. We have no experience with this technique but it may be a useful data analysis termination strategy.

FINAL COMMENT. While following up on recommendations of the panel, we ran across two treatises generally covering the areas of goodness-of-fit and data analysis termination strategies which we feel are important to pass on to the reader. They are Daniel and Wood's book (see ref. 18) and a recent paper by Hocking.[22] These sources contain discussions of other techniques which may be applicable to the problems addressed in this paper.

ACKNOWLEDGEMENTS. We wish to express our sincere appreciation to the members of the panel for their thoughtful consideration of our problems in analysing pharmacokinetic data and for their thought-provoking comments. Any errors in paraphrasing the panelists' comments or misinterpretation of their advice remain the responsibility of the authors. In addition, assistance in preparing this manuscript of Dr. Lewis Sheiner of the University of California, San Francisco is gratefully acknowledged.

REFERENCES.

1. Boxenbaum HG, Riegelman S, Elashoff RM: Statistical estimations in pharmacokinetics. J Pharmacokin Biopharm 2:133-148, 1974.

2. Pharmacokinetics by M. Gibaldi and D. Perrier, p. V, Marcel Dekker, Inc, New York, NY, 1975.

3. Fundamentals of Clinical Pharmacokinetics by J. G. Wagner, Drug Intelligence Publications, Inc, Hamilton, Ill, 1975.

4. Wagner JG: Properties of the Michaelis-Manton equation and its integrated form which are useful in pharmacokinetics. J Pharmacokin Biopharm 1:103-121, 1973.

5. Sedman AJ, Wagner JG: CSTRIP, a FORTRAN IV computer program for obtaining initial polyexponential parameter estimates. J Pharm Sci 65:1006-1010, 1976.

6. AUTOAN, a Time-Sharing Digital Computer Program, available from Publication Distribution Service, 615 East University Avenue, Ann Arbor, Michigan 48106.

7. S. A. A. M. Manual by M. Berman and M. F. Weiss, Government Printing Office, Washington, DC.

8. A Users Manual for NONLIN and Associated Programs, Upjohn Co, Kalamazoo, Michigan, 1974.

9. Draper NR, Smith H: Chapter 3 "The Examination of Residuals" in Applied Regression Analysis, Wiley & Sons, New York, 1966.

10. Neter J, Wasserman W: Chapter 4 "Aptness of Model and Remedial Measures" in Applied Linear Statistical Models, R D Irwin, Inc, Homewood, Ill, 1974.

11. Shapiro SS, Wilk MB: An analysis of variance test for normality. Biometrika 52:591, 1965.

REFERENCES (continued).

12. Wagner JG: Do you need a pharmacokinetic model and, if so, which one? J Pharmacokin Biopharm 3:457-478, 1975.

13. Wagner JG: Linear pharmacokinetic models and vanishing exponential terms: Implications in pharmacokinetics. J Pharmacokin Biopharm 4:395-426, 1976.

14. Wagner JG: Linear pharmacokinetic equations allowing direct calculation of many needed pharmacokinetic parameters from the coefficients and exponents of polyexponential equations which have been fitted to the data. J Pharmacokin Biopharm 4:443-467, 1976.

15. Wold S: Spline functions in data analysis. Technometrics 16: 1-11, 1974.

16. Hogg RV: On Robust Statistical Procedures. Proceedings of the Twenty-second Conference on the Design of Experiments in Army Research, Developing and Testing, pp 1976.

17. Frome EL, Yakatan GJ: Statistical Estimation of the Pharmacokinetic Parameters in the One Compartment Open Model. Res Rep CSS 269, Center for Cybernetic Studies, University of Texas, Austin, Texas 78712, 1976.

18. Daniel C, Wood FS: pp 86-90 "Mallows graphical method of comparing fitted equations" in Fitting Equations to Data, Wiley-Interscience, J Wiley & Sons, Inc, New York, NY, 1971.

19. Geisser S: The predictive sample reuse method with applications. J Am Stat Assn 70:320-328, 1975.

20. Geisser S: A predictive approach to the random effect model Biometrika 61:101-107, 1974.

21. Stone M: Cross-validatory choice and assessment of statistical predictions. J Statist Soc B36:111-147, 1974.

22. Hocking RR: The analysis and selection of variables in linear regression. Biometrics 32:1-49, 1976.

EFFECTS OF DIETARY BRAN AND CELLULOSE ON SERUM LIPIDS

Charlotte M. Heggie (1); Daniel H. Connor (1);
Frank A. Franklin, Jr., (2); Walter D. Foster, (1);
Samuel M. Wylde (3); Joe M. Blumberg, (4).

(1) Armed Forces Institute of Pathology, Washington, D.C. 20306;
(2) The Preventive Medicine Office, Walter Reed Army Medical Center, Washington, D.C. 20306; (3) Ener-G-Foods, Inc., Seattle, WA 98134; and (4) Oscar B. Hunter Memorial Laboratory, Washington, D.C. 20005

ABSTRACT. Unprocessed bran (bran) and carboxy-methyl cellulose (CMC) were added to regular diets of overweight and normal weight volunteers to determine the effect on serum lipids. Downward mean trends of cholesterol and triglyceride levels were found in all groups taking bran and CMC after twelve weeks except the overweight men ingesting CMC. Downward mean trends for cholesterol ranged from 0.74 mg/100 to 1.65 mg/100 per week and for the triglyceride from 0.36 mg/100 to 4.78 mg/100 per week.

1. INTRODUCTION. Coronary atherosclerosis is the leading cause of death in the United States.[1,2] In spite of this atherosclerosis was rare in this country before 1900,[3] and remains almost unknown in some developing countries.

Dietary factors are under constant scrutiny, and a number of researchers have proposed that lack of dietary fiber may be an important causal factor,[8-11] because fiber is abundant in the diets of rural people in developing countries where atherosclerosis is rare and has decreased in the diets of westerners during the rise of fatal atherosclerosis.

Dietary fiber could lower serum lipids in various ways. It is hygroscopic and might absorb emulsified lipids taken with the diet. Dietary fiber would also absorb cholesterol secreted in the bile and thus reduce its reabsorption in the small intestine. Increased dietary fiber also reduces gastro-intestinal transit time and thismight also reduce absorptions of lipids.

In an attempt to determine whether dietary fiber reduces serum lipids, we performed the following study.

2. METHODS. Forty-four healthy men ages ranging from 23 to 65 years, volunteered for a 12-week study. All were on duty at the Armed Forces Institute of Pathology when the study began. Most were pathologists, and the remainder were trained in one of the medical specialties. All understood the purpose of the study and were "dedicated" volunteers. They continued their regular diets, did not alter their life styles, and maintained body weight.

PRECEDING PAGE NOT FILMED BLANK

The men were divided by height/weight ratio and age into three equivalent groups--control, bran and cellulose. Each member of the bran group added 56 gm of unprocessed bran to his daily diet--28 gm (1 ounce) with breakfast and 28 gm with his evening meal--a daily supplement of about 6 gm of nonnutritive fiber. Each member of the cellulose group added 6 gm of cellulose* to his daily diet--3 gm at breakfast and 3 gm with the evening meal. This is all nondigestable, so both groups ingested approximately 6 grams of nonnutritive hygroscopic substance. These supplements were ingested for 12 weeks.

Fasting blood samples were collected at intervals of two weeks; serum cholesterol determinations were done every two weeks; and serum triglyceride determinations, every four weeks. (The control group, however, had no triglyceride determinations on the fourth week.)

During the course of the study, 9 of the 44 men dropped out-- 4 were transferred, 4 could not tolerate the unprocessed bran, and 1 man substituted sweetened bran ("All Bran") for unprocessed bran. Of the remaining 35, 18 had a "normal" weight and 17 were overweight. Linear regression to estimate the trend of each man's serum lipids was calculated and the trends were averaged for each group. Because only slopes were averaged, the variation introduced by differences in lipid levels from subject to subject was removed--a valid approach since each subject acted as his own control in the trend analysis. A refinement of the analysis involved the recomputation of the average trends per group with each subject's degree of consistency of trend used as a weight in obtaining a weighted-average trend (where degree of consistency was measured as the reciprocal of the variance of the slope). The weighted-average, while conferring greater importance to consistent trends, also served to be selective, giving some subjects considerable prominence. Therefore, special care was taken in the interpretation of the weighted averages to ensure that they were also representative of the group.

The probabilities were obtained from Student's t-test on the average trends (weighted and unweighted) for each group under the null hypothesis of zero trend against the one-sided, alternative hypothesis of negative slope.

3. RESULTS. The triglyceride levels were sharply lowered in the normal-weight subjects eating bran and cellulose. The group of overweight subjects eating bran and cellulose and the control group did not show this striking trend. See Fig. 1. In addition mean cholesterol levels fell in the group of overweight men taking bran. The graphs in Fig. 1 are means of the individual trends so that the variation in lipid levels from subject to subject was removed.

---

*Purchased as sodium carboxy-methyl-cellulose tablets, 0.5 gm, from Interstate Drug Exchange Mfg. Co., Plainview, Long Island, New York 11803

Using a preliminary cutoff at $P \leq .10$, four of the seven negative slopes in the bran and cellulose groups were statistically significant. See Table 1. Expressed as a percentage of the initial levels, the reduction was 75% for the group taking CMC and 60% for the group taking bran. Three of the mean trends that failed the statistical cutoff were groups of overweight men. Because of the greater variability of serum lipid trends among the overweight men, a refined analysis was performed consisting of computing weighted-average trends using as weights the degree of consistency of each individual's trend. The weighted trends generally show a numerically steeper rate of reduction of serum lipids together with enhanced statistical probabilities. Thus, six of the eight trends for the bran and cellulose groups were statistically significant ($P \leq .05$) downward trends. The only non-downward trend was the overweight men ingesting CMC whose serum cholesterol unaccountably increased. This contrasts with the decreased triglyceride level for this same group.

Since each subject served as his own control--his pretreatment level was the initial point for his own trend--no reference thus far has been made to the actual control groups. They served to determine whether an unknown or subconscious factor influenced serum lipids during the study. The average trends for the control group revealed no such factor. See Table 1. One of the weighted-average trends--the triglyceride levels in the normal-weight control group--did fall with $P = .12$. To be conservative, therefore, this slope was subtracted from the slopes of the bran and cellulose groups for the normal-weight men, in computing the probability statements.

Laboratory variation, expressed as a ratio of the laboratory variance to the residual experimental variance, was 1/16, a negligible quantity as a possible factor affecting the analysis and interpretation of these data. The standard deviation for the laboratory, calculated over each two-month period, was found to be 5 mg/100 ml for serum cholesterol and 8 mg/100 ml for serum triglyceride.

4. COMMENT. A number of studies reporting the effects of whole or fractional grain products on serum lipids have produced varied results, but the majority support the view that whole grain and whole grain products tend to lower serum lipids.[19-22] In our study CMC lowered the average triglyceride levels by 75% in normal-weight subjects, and bran lowered the average serum triglyceride levels of normal-weight subjects by 60%. We do not know the mechanisms by which bran and CMC lowered serum lipids. Some possible mechanisms suggest that nonnutritive substance (1) increases the excretion of bile acids by increasing catabolism of cholesterol in the liber,[19] (2) shortens gastrointestinal transit time, thus allowing less time for lipids to be absorbed, and (3) absorbs water, bile salts and other solutes including lipids, thus reducing absorption of lipids. None of these hypotheses however, explains the fact that serum triglycerides in our normal-weight men dropped more quickly than serum lipids in our overweight men

ingesting bran and CMC. If nonnutritive substance lowers serum triglycerides more quickly in non-obese men, then other dietary factors probably play a role. One of these could be the ingestion of excessive amounts of refined carbohydrates by the overweight men. Sugar, for example, not only contributes to obesity but is an important cause of hyperlipidemia.[23]

Our study supports the opinion that nonnutritive substance (bran and CMC) lower serum lipids. And in particular, we found that the most striking lowering effect was on the serum triglycerides in men taking CMC who were not overweight.

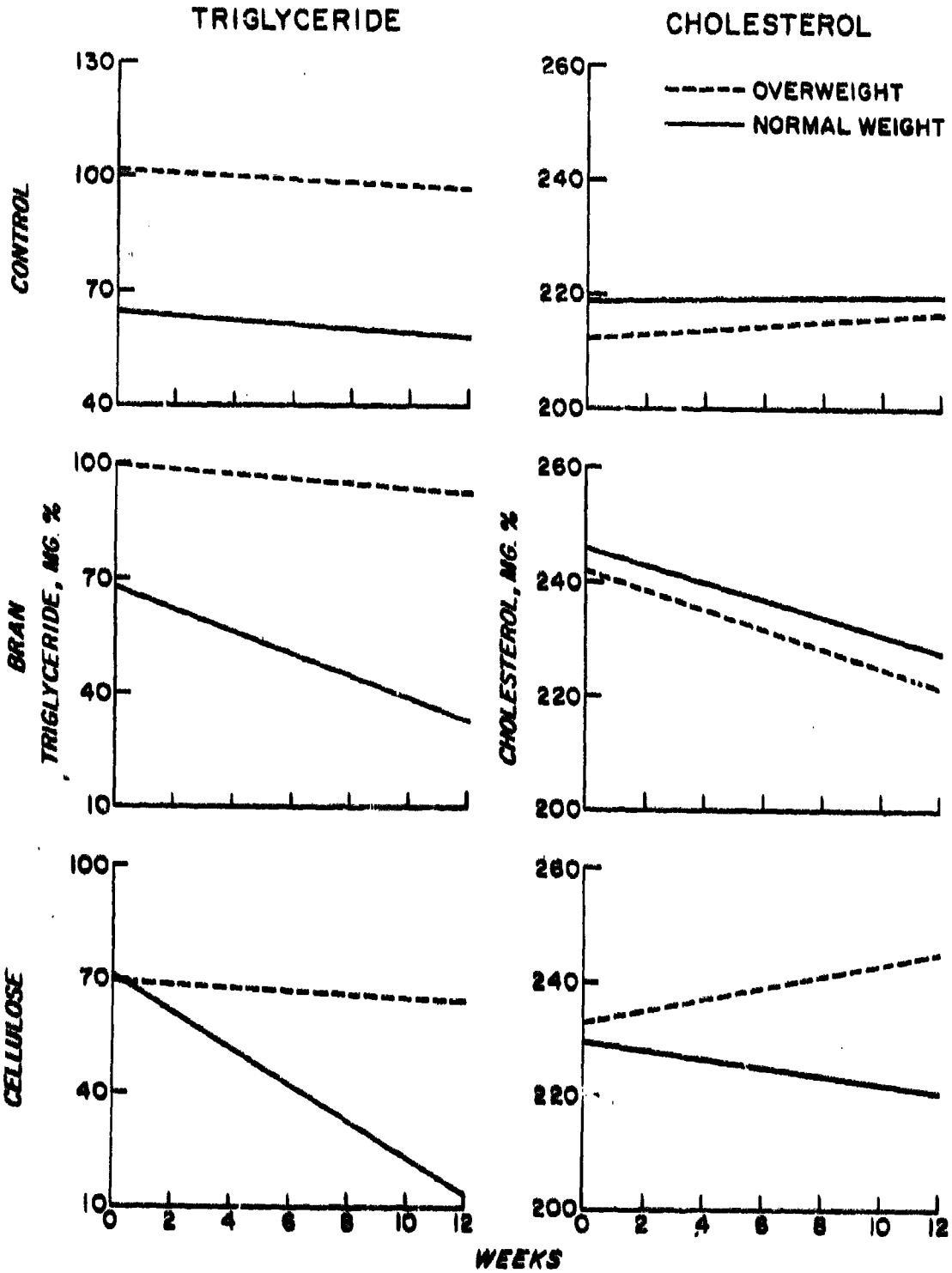


FIGURE 1.
MEAN TRENDS OF SERUM CHOLESTEROL AND TRIGLYCERIDE LEVELS FOR THREE GROUPS OF VOLUNTEERS--THE CONTROL GROUP, THE GROUP INGESTING BRAN, AND THE GROUP INGESTING CELLULOSE.

Table 1. Mean Tends ($\bar{b}$, mg % per week) of Serum Triglyceride and Serum Cholesterol Levels as Determined on 35 Volunteers for 12 Weeks.

	No. of Vol.	Averages $\bar{b}^*$	$s(\bar{b})$	Prob.	Weighted Averages $\bar{b}w^+$	$s(b\bar{w})$	Prob.
Serum Triglyceride							
Controls							
Normal Weight	9	-0.43	0.93	NS	-0.53	0.30	0.12
Overweight	5	-0.33	1.70	NS	0.80	1.36	NS
Bran							
Normal Weight	4	-2.88	1.77	0.08	-3.48	1.69	0.05
Overweight	6	- .60	1.86	NS	-2.06	0.71	0.01
Cellulose							
Normal Weight	5	-4.78	1.95	0.02	-4.12	0.90	0.001
Overweight	6	-0.36	1.72	NS	- .85	0.38	0.02
Serum Cholesterol							
Controls							
Normal Weight	9	0.22	0.70	NS	0.09	0.56	NS
Overweight	5	0.37	0.78	NS	-0.78	0.62	NS
Bran							
Normal Weight	4	-1.49	0.96	0.07	-3.42	0.59	0.001
Overweight	6	-1.65	1.11	0.08	-2.85	0.68	0.001
Cellulose							
Normal Weight	5	-0.74	0.85	NS	-0.98	0.75	0.10
Overweight	6	1.04	0.75	NS	1.52	0.80	NS

$$^{*}\bar{b} = \frac{\Sigma Sx_i}{\Sigma Sx^2}$$

$$V(\bar{b}) = \frac{\Sigma s^2_{y.x(D.F.)}}{\Sigma sx^2 \ (D.F.)}$$

$$^{+}\bar{b} = \frac{Sw_i b_i}{Sw_i}$$

$$w_i = 1/V(b_i)$$

$$V(\bar{b}_w) = 1/Sw_i$$

References

1. Stamler J, et al: Primary prevention of the atherosclerotic diseases. Circulation 42:A55-A95, 1970.

2. Burkitt DP: Some diseases characteristic of modern western civilization. Br Med J 1:274-278, 1973.

3. Burkitt DP, Walker ARP, Painter NS: Dietary fiber and disease. JAMA 229:1068-1074, 1974.

4. Antonis A, Bersohn I: The influence of diet on serum triglycerides in South African white and Bantu prisoners. Lancet 1:3-9, 1961.

5. Walker ARP: Biological and disease patterns in South African interracial populations as modified by rise in privilege. S Afr Med J 46:1127-1134, 1972.

6. Trowell H: Crude fibre, dietary fibre, and atherosclerosis. Atherosclerosis 16:138-140, 1972. (Letter to the Editor.)

7. Schaefer O: When the Eskimo comes to town. Nutr Today. Nov/Dec: 8-16, 1971.

8. Roughage in the diet. Med Wld News 35-42 Sept 6, 1974.

9. Trowell H: Fiber: A natural hypocholesteremic agent. Am J Clin Nutr 25:464-465, 1972.

10. Trowell H: Brown bread versus white. Br Med J 2:752-753, 1937. (Editorial.)

11. Trowell H: Ischemic heart disease and dietary fiber. Am J Clin Nutr 25:926-932, 1972.

12. Bruno M: Non-nutritive fiber can boost your health. Mag Better Health 129-147, March 1973.

13. Painter NS, Almeida AZ, Colebourne KW: Unprocessed bran in treatment of diverticular disease of the colon. Br Med J 1:137-140, 1972.

14. Burkitt DP, Walker ARP, Painter NS: Effect of dietary fibre on stools and transit-times, and its role in the causation of disease. Lancet 2:1408-1412, 1972.

15. Painter NS: Diverticulosis and diverticulitis. Practitioner 208:669-670, 1972.

16. Mayer J: Obesity: Diagnosis. Postgrad Med 25:469-475, (April) 1959.

17. Cummings JH: Progress report. Dietary fibre. Gut 14:69-81, 1973.

18. Eastwood MA, et al: The effects of dietary supplements of wheat bran and cellulose on faeces. Proc Nutr Soc 32:22A, 1973. (Abstract.)

19. Mathur KS, Khan MA, Sharma, RD: Hypocholesterolaemic effect of bengal gram: A long-term study in man. Br Med J 1:30-31, 1968.

20. De Groot AP, Luyken R, Pikaar NA: Cholesterol-lowering effect of rolled oats. Lancet 2:303-304, 1963.

21. Edwards CH, et al: Utilization of wheat by adult man: Nitrogen metabolism, plasma amino acids and lipids. Am J Clin Nutr 24:181-193, 1971.

22. Heaton KW, Pomare EW: Effect of bran on blood lipids and calcium. Lancet 1:49-50, 1971.

23. Ahrens E et al: Carbohydrate-Induced and Fat-Induced Lipemia. Trans Assoc Am Physicians 74:134-146, 1961.

ANALYSIS OF AN ERROR-TIME RESPONSE PERFORMANCE

Michael Hacskaylo
U.S. Army Electronics Command
Night Vision Laboratory
Fort Belvoir, Virginia

ABSTRACT. The analyses of the error-time response performances of groups of naive subjects permitted to make discrete right/wrong decisions are presented for three experimental display panels of increasing complexity. The panel designs were based on a circular representation of light bulbs, where the lights corresponded to the angles of a circle. The first panel design consisted of a ring of lights that portrayed one contiguous angular representation by the lights. A contiguous representation of light was defined as a domain. The complexities of the second and third designs were increased to two contiguous semicircular representations of the ring of lights, where for each design the semicircular representatives were defined as two domains. The function of the panels was to display the azimuthal angular source location of infrared lasers when detected by infrared detection systems.

The subjects were randomly selected from a large population having no prior knowledge (zero degree of learning) of the panels and separated into three groups of seven subjects each. Each subject evaluated two of the three panels in an ABBA manner for one and only one set of six trials per panel. Such a group of subjects, constrained to the same degree of learning of the panels and limited to the one set of trials, is defined as an eigengroup for this analysis.

The subjects were instructed to mark on a response panel as accurately and rapidly as possible the corresponding angular light of the stimulus panel, viz, the display panel. The response panel was a five inch circle drawn on a 8 inch by 10 inch paper.

The number of errors of the eigengroups was analyzed as a function of time for each of the experimental designs. It was found that for the experiment, the error-time response equation is $\log E = -2n \log T + K$, where E is the number of total errors per eigengroup, n is the number of domains of the stimulus panel, T is the mean time for the total number of trials for each eigengroup per system, and K is a constant. It was necessary to introduce new terms, i.e., domain and eigengroup to unambiguously define the stimulus panel and interpret the results consistent with the equation.

1. INTRODUCTION. The purpose of this paper is to present an error-time analyses of the designs of the informational display panels of infrared detection systems. The systems detected and displayed the azimuthal angular position of a laser source to a crew during a laser-tank

engagement as shown in figure 1. Since the error-time response performance of a well-trained crew would more apt reflect the selectivity and training of the crew, an evaluation procedure was required that would reflect the panel design rather than the personnel capability, training and cumulative learning process. To implement the procedure it was decided to employ naive subjects who had no knowledge of the systems, exposed only to instructional procedures (without preliminary learning trials) and constrained to make one and only one decision per trial. The decision would be considered right or wrong. The results should be different than the cumulative learning performance where error decisions were allowed until the correct decision was made, Gagné and Foster (1949). In other procedures, errors were treated as partially correct answers (Fitts and Seeger, 1953), and the error-time response data are statistically treated to determine the mean and standard deviation of the error-time parameters. These parameters are interpreted as how far from the correct value the errors are as a function of learning and response times. The determination of the number of discrete errors as a function of time for a group of subjects, who were not trained nor subjected to the cumulative learning process, is not apparent in literature.

Figure 1. Artist's conception of a laser-tank engagement

In this paper, the error-time response performance of groups of individuals subjected to only one set of trials resulted in a frequency-distribution curve which was different than a cumulative performance curve. The mathematical analysis of error-time response data of random groups of individuals subjected to the one-trial set method appears to be significantly new. To assure that the groups were not subjected to a cumulative learning process, each subject was instructed as to the procedure and then dismissed after evaluating the panels. In this manner, each group was considered to be of the same or identical state of conditioning or training for all sets of trials. Such groups are defined as eigengroups. (The word, eigen means proper, inherent, peculiar). In a fuller context, the error-time response performance of eigengroups is properly satisfied only when the groups are subjected to the one-set trial method.

2. METHOD. The error-time response performance data were obtained on panel designs similar to those of Fitts and Seeger (1953). Due to the similarity, the Fitts and Seeger experiment is briefly described. Their experiment, in essence, was to determine the learning skills of matched groups of individuals to a singlefold response. The stimulus panel had a ring of eight equally spaced light bulbs. The stimulus was a light flashing on. This action keyed the subject to associate the light with the angular position on the ring. The response panel had a stylus. The response was the action by the subject in moving the stylus to the corresponding position on the response panel as the interpreted position of the stimulus panel. (Two variations of the stimulus panel were geometrically configured with increasing complexity to simulate the ring design. The corresponding response panels were also increased in complexity. The S-R compatibility of those designs were also determined.)

The panel designs reported here were also based on a circular representation of equally spaced light bulbs. Since the physical entity is the light bulb embodying the stimulus, the physical entity (light bulb) is defined as the stimulant. The stimulant and the configurational display of the stimulant (ring of light bulbs) on the stimulus panel is defined in this paper as the significand. The significands were geometrically configured to increase the complexity of the stimulus panel for the singlefold response. The three designs are now described.

Panel A. The significand of the panel was a three and one-half inch diameter ring of 36 equally spaced light bulbs as portrayed in figure 2(a). The ring was positioned on the front surface of a box 4 inches wide, 8 inches long and 2 inches deep. The light bulbs were angularly marked in degrees from zero to 360 degrees in ten degree increments in a clockwise direction with zero at the top. The continuous clockwise direction of the marked light bulbs is considered as a domain of the significand, i.e., one contiguous representation of the stimulus panel as portrayed in figure 2(b). When a light came on it signified the angular position on the ring.

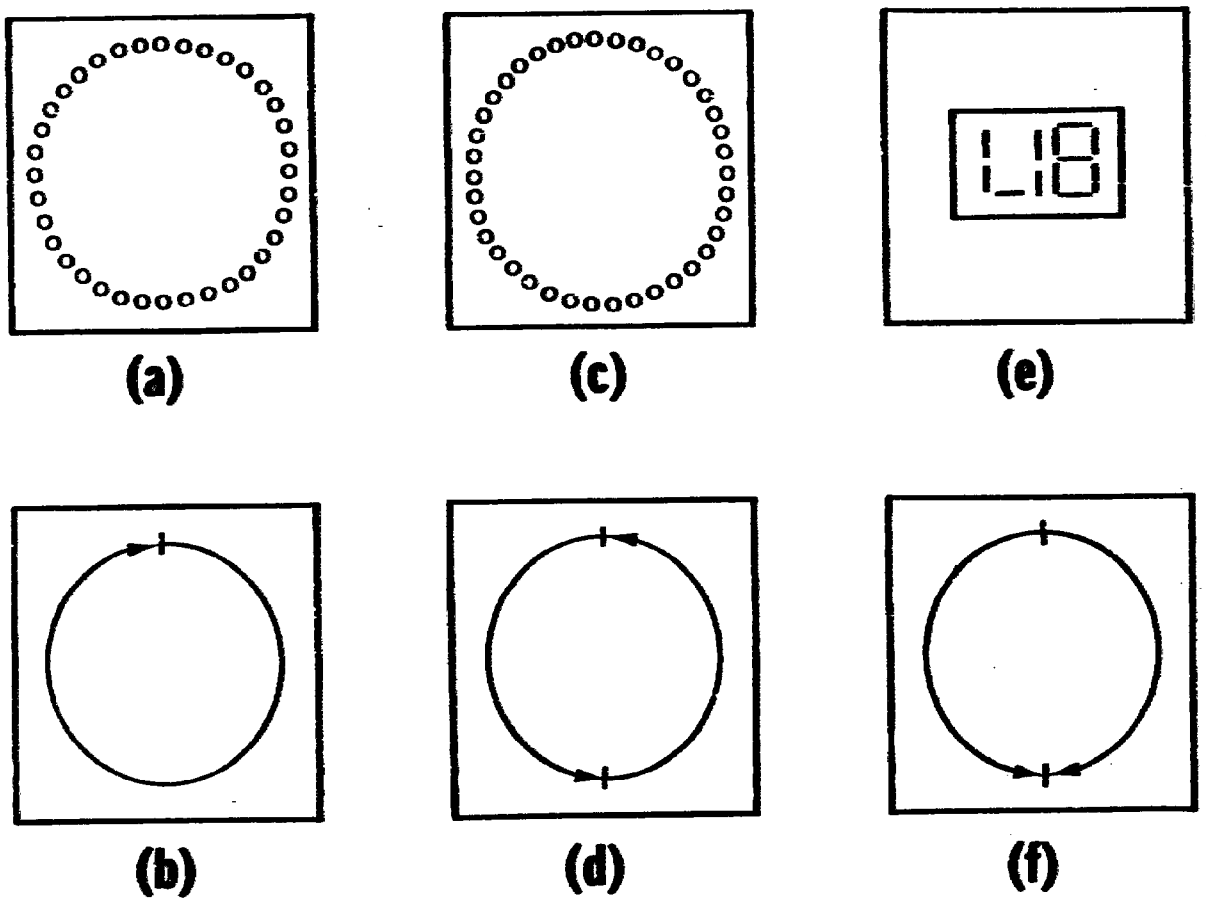


Figure 2. Schematic representation of display panels A, B, and C are shown as (a), (c) and (e) with the corresponding domains, (b), (d) and (f)

Panel B. The significand of the panel was a three and one-half inch diameter ring of 36 equally spaced light bulbs marked in angular mils as portrayed in figure 2(c). The ring was positioned on the front surface of a box identical in dimensions as in panel A. The light bulbs were angularly marked in mils from zero to 3200 in 177.78 mil increments in a counterclockwise direction with zero at the top for one-half of the circle. (There are 6400 mils per 360 degrees of a circle, therefore, each position corresponds to 177.78 mils as well as 10 degrees.) The angular marking started at zero again at the bottom, and continued in the counterclockwise direction to 3200 at the top. The two halves completed the circle. The two counterclockwise directional iterations of the marked light bulbs are considered as two domains of the significand, i.e., two contiguous representations of the panel as shown in figure 2(d). When a light came on, it signified the angular position on the ring.

Panel C. The significand of the panel was a stimulant in the form as an alphanumeric readout display as portrayed in figure 2(e). The display window was positioned on the front surface of a box of identical dimension as in panel A. The first of three characters was a letter, L or R, and the next two were digits ranging from 00 to 32. The letter R indicated a circular representation in a clockwise direction. The numerical values indicated the angular position in 100 mil increments (equivalent to 5.625 degrees) with zero at the top and increasing to 3200 mils for one-half of the circular representation. The letter L indicated a circular representation in a counterclockwise direction. The numerical values indicated the angular position in 100 mil increments with zero at the top and increased to 3200 mils for the completion of the circular representation. The one clockwise and one counterclockwise directional representations of the circle are considered as two domains of the significand, i.e., two semicircular representations of the stimulus panel as shown in figure 2(f). When an alphanumeric readout came on, it signified the angular position on the circular representation.

Response Panel. The response panel was identical for each panel. A five inch circle was drawn on a 8 x 10 inch sheet of plain paper. The circle was divided into quadrants and marked into degrees and mils as follows: Zero degrees (0°) and zero mils (0 mils) were marked at the top. In a clockwise direction, each quadrant was successively marked 90°, 1600 mils; 180°, 3200 mils; 270°, 4800 mils; and again at the top, 360°, 6400 mils. A pencil was used for marking angular positions with an "X" on the circle.

3. PROCEDURE. Twenty-one U.S. Army enlisted men of all ranks, who were not formally matched but had no prior knowledge of the experimental panels, were randomly selected and separated into three groups of seven subjects each. One at a time, each subject was thoroughly briefed on the operational procedures of two preselected display panels just prior to evaluation. The subject was instructed as follows: As quickly and as accurately as possible, read the angular representation of a light (or

digital readout) and the appropriate direction when the stimulus light came on, and mark with an "X" that angular position on the circle of the sheet of paper. The position of the intersection of the "X" was considered to be the angular position. For familiarization the subject was given two preliminary runs if so desired.

Each subject performed a series of six trials on the two preselected cases in an ABBA manner for a total number of 12 trials. On a preprogrammed schedule of randomness, each subject read the angular position and marked the circle as quickly and accurately as possible. There were three "X"'s per response (paper) panel since one panel was supplied for each A, B, B, A sequence. The time interval from when the light came on to when the subject marked the panel was measured to 0.001 second, however, the time for each trial was recorded to the nearest 0.01 second. It is assumed that the reaction-time error introduced by the investigative team for the time measurements was constant for the trials.

Upon completion of the set of trials, each subject was dismissed. Care was taken to insure that subsequent subjects for evaluation did not associate with any of the previously dismissed subjects.

The display panels were evaluated in a room that consisted of a 36 inch high bench, chair, and associated equipment required to activate the lights of the panels. The panels, two at a time, were positioned one on top of the other on the bench in the following sequence: for eigengroup 1, panel A on panel C; for eigengroup 2, panel B on panel C; and for eigengroup 3, panel B on panel A. It is to be noted that the sequence for eigengroup 2 was incorrect to maintain proper counterbalancing block order. However, this flaw did not appear to be evidenced in the analyses as described later. The only persons permitted in the room were the subject and the investigative personnel.

For each stimulus panel angle (light), three angular resolution ranges for determining the accuracy of the response panel angle "X" were considered to be a) $\pm$40 degrees; b) $\pm$20 degrees; and c) $\pm$10 degrees. The readout angle was considered as an error if the response angle was greater than the angular resolution for each stimulus angle, i.e., each response angle would be a right/wrong decision for three ranges.

The angular position marked on the response panels (from the preprogrammed readout angles) were measured in degrees. This was done by using a transparent template graduated to 0.5 degree which was superimposed on the marked response panels. The accuracy of the marked angle was measured to $\pm$0.5 degrees.

The number of errors that each subject made with respect to each of the three ranges for a set of six trials for each panel were counted.

The average times of the six trials for the panels that each subject evaluated were determined. The number of errors and associated time for each of the 252 trials were tabulated for data reduction.

4. RESULTS. The mean time and number of angular errors in each range for the set of six trials for the subjects are shown in Table 1. The table separates the subjects in their respective eigengroups for the panels evaluated. The average of the mean times as well as the total errors per range for the groups for each panel are also shown in Table 1. Note that the errors are considered as completed events and that the standard deviation of the angular errors have no significance in this analysis.

An accepted method for the portrayal of the frequency-distribution data of Table 1, is to plot the number of errors (per subject) as a function of the mean time (per subject). To illustrate the method, plots of the number of errors in the range $\pm 40°$ as a function of the mean time of the subjects of the groups for each of the panels are shown in figure 3. The data presented in such a fashion cannot be clearly interpreted. The only two significant observations that can be made are as follows: The first is that the error-mean time response performance curves of the two eigengroups for the same design exhibit some degree of similarity, and the second is that most of the errors occur between 3 and 5 second time interval.

However, if the data are plotted in a different fashion, a strikingly new set of parametric curves are generated. If, for the data of Table 1, the number of total errors, E in the range $\pm 40°$ per eigengroup is plotted as a function of the mean time on a log E vs log T scale, it can be seen that two distinct linear curves are generated as shown in figure 4. The eigengroup datum points for panel A fall on one line, and eigengroup datum points of panel B and panel C fall on a second line. The two curves are separated by at least one order of magnitude in the error count, and this separation indicates that there is an uniqueness between panel A and panels B and C.

The curve for panel A can be expressed as

$$\log E_{A,40} = -2 \log T_A + 1.16 \qquad (1)$$

where $E_{A,40}$ is the total number of errors $\pm 40°$ per eigengroup for panel A; T_A is the mean time per eigengroup of panel A; and 1.16 is a constant. The negative sign is interpreted to mean that as the amount of time is increased for reading the stimulus panel angle and marking the "X" on the response panel, the number of errors decrease.

HUMAN FACTORS DATA, N=252

Eigengroup 1

Subject	Panel A Mean Time (Sec)	Panel A Angular Errors (Number) ±40°	±20°	±10°	Panel C Mean Time (Sec)	Panel C Angular Errors (Number) ±40°	±20°	±10°
1	5.85	0	0	0	5.86	1	1	1
2	2.83	0	2	4	3.45	0	0	1
3	5.10	0	1	2	5.84	1	1	1
4	2.58	0	0	1	2.96	1	1	1
5	2.95	1	1	2	3.52	4	5	6
6	4.68	0	0	2	3.98	0	0	1
7	3.17	0	1	3	4.86	1	2	3
	3.87(Ave)	1	5	14	4.35(Ave)	8	10	14

Eigengroup 2

Subject	Panel B Mean Time (Sec)	Panel B Angular Errors (Number) ±40°	±20°	±10°	Panel C Mean Time (Sec)	Panel C Angular Errors (Number) ±40°	±20°	±10°
8	3.78	0	0	1	3.51	1	1	1
9	4.38	1	1	3	3.37	0	0	0
10	4.11	0	0	1	3.66	0	0	0
11	5.06	0	0	1	3.87	0	1	3
12	4.59	0	1	2	4.23	0	1	1
13	11.51	0	1	3	10.08	0	0	0
14	5.24	1	2	4	5.77	2	3	5
	5.52(Ave)	2	5	15	4.93(Ave)	3	6	10

Eigengroup 3

Subject	Panel B Mean Time (Sec)	Panel B Angular Errors (Number) ±40°	±20°	±10°	Panel A Mean Time (Sec)	Panel A Angular Errors (Number) ±40°	±20°	±10°
15	2.95	5	5	5	2.90	0	1	4
16	3.33	4	6	6	2.75	0	1	3
17	3.85	3	4	5	3.49	1	1	4
18	1.87	4	6	6	0.98	1	2	6
19	2.79	5	5	6	2.55	0	2	4
20	3.24	1	1	4	3.18	0	2	4
21	4.26	0	0	2	3.82	0	0	4
	3.26(Ave)	22	27	34	2.81(Ave)	2	9	29

Table 1. Mean time and number of errors for each subject per eigengroup tabulated for each angular resolution range per panel.

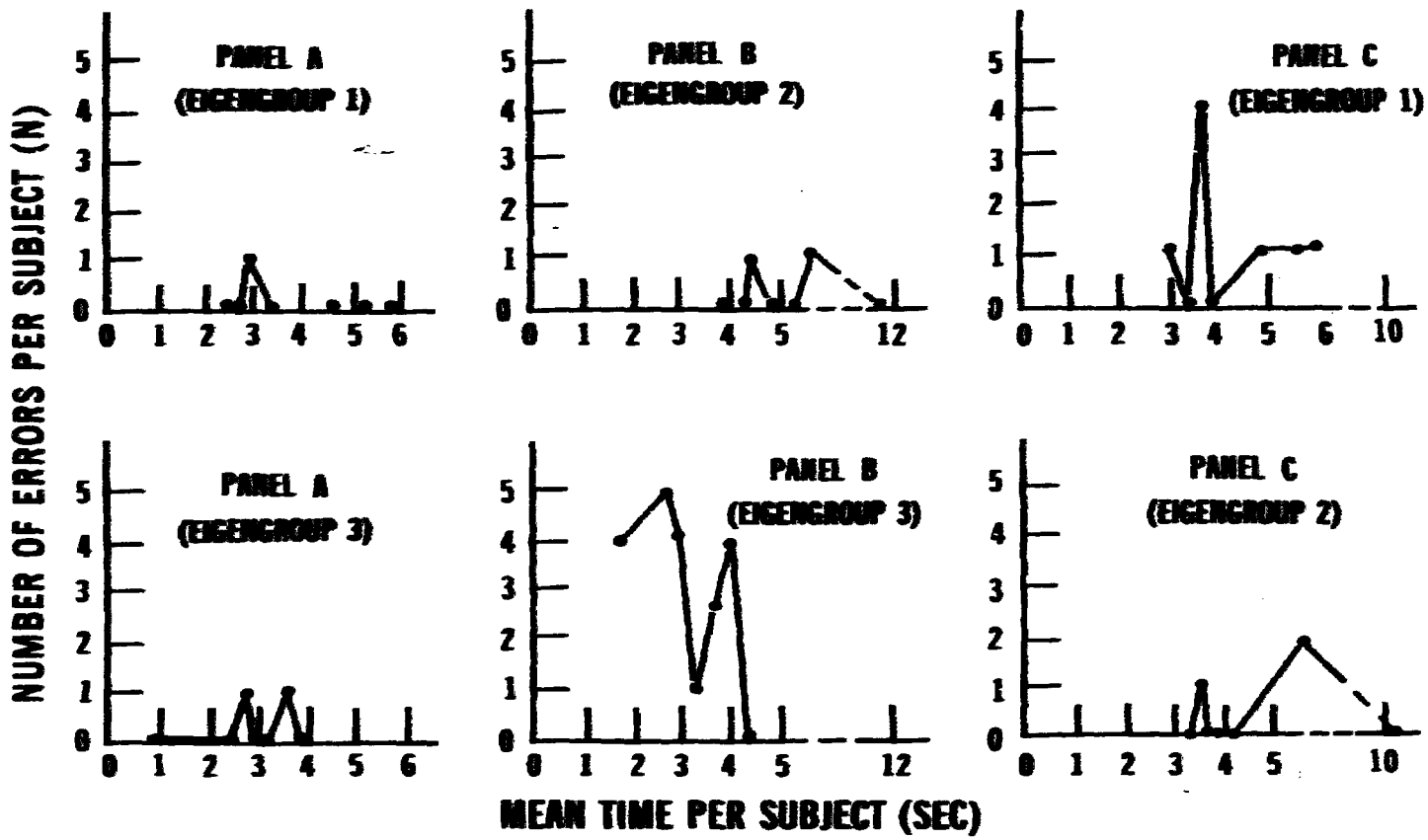


Figure 3. Number of errors and corresponding mean time for panels evaluated by subjects per eigengroup, ±40 degree resolution range

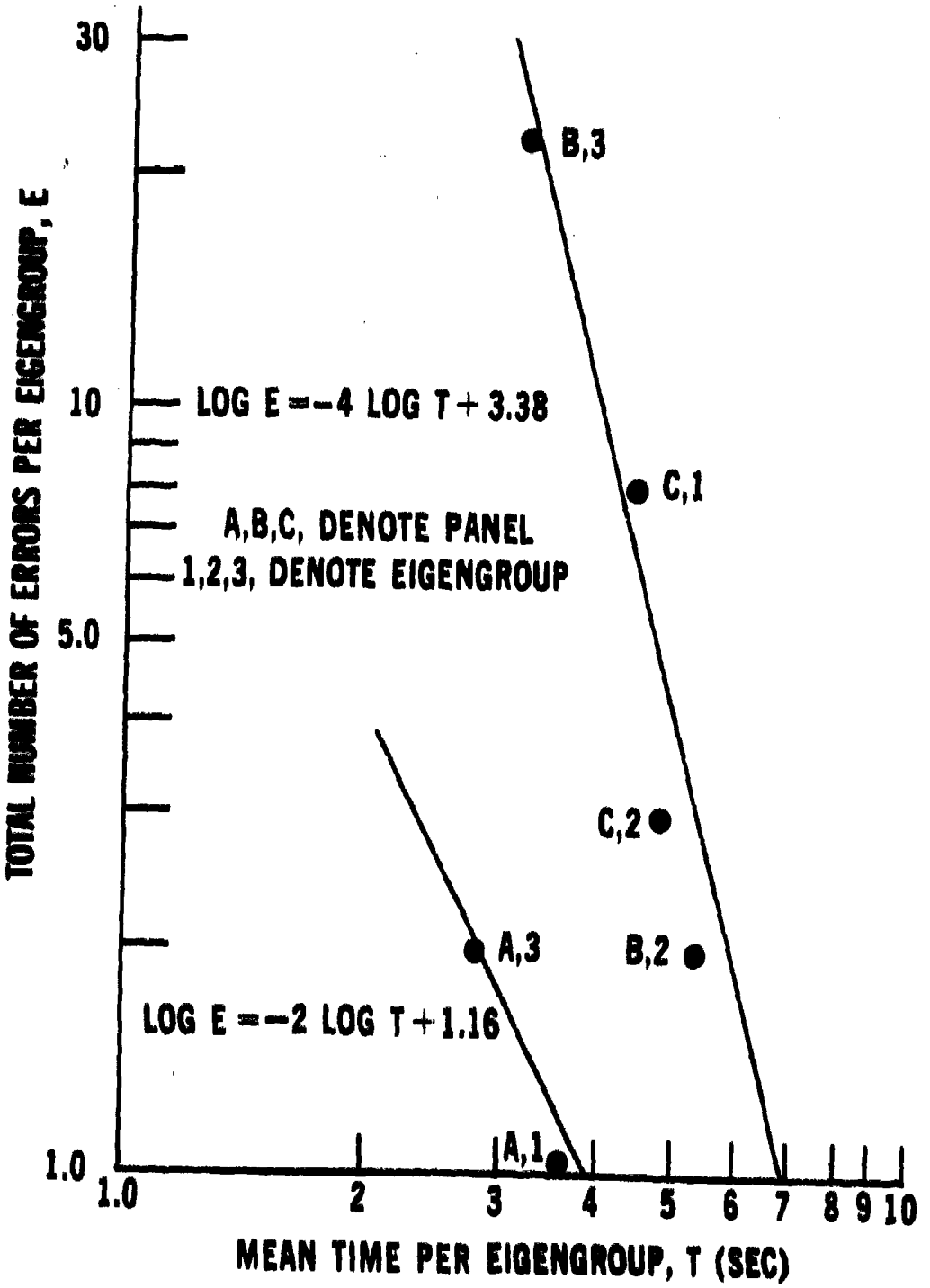


Figure 4. Error dependence on the mean time for panels evaluated per eigengroup, ±40 degree resolution range

The curve for panel B and panel C can be expressed as

$$\log E_{B,C,40} = -4 \log T_{B,C} + 3.38 \qquad (2)$$

where $E_{B,C,40}$ is the total number of errors $\pm 40°$, per eigengroup for panels B and C; $T_{B,C}$ is the mean time per eigengroup for panels B and C; and 3.38 is a constant.

The similar plots of the number of total errors $\pm 20°$ per eigengroup are shown in figure 5. The curve for panel A can be expressed as

$$\log E_{A,20} = -2 \log T_A + 1.86 \qquad (3)$$

where $E_{A,20}$ is the total number of errors $\pm 20°$ per eigengroup for panel A, T_A is the mean time and 1.86 is a constant. The curve for panels B and C can be expressed as

$$\log E_{B,C,20} = -4 \log T_{B,C} + 3.57 \qquad (4)$$

where the terms have the same comparable definitions as for Eq. (2).

The plots of the number of total errors $\pm 10°$ per eigengroup are shown in figure 6. The curve for panel A can be expressed as

$$\log E_{A,10} = -2 \log T_A + 2.35, \qquad (5)$$

and the curve for panels B and C can be expressed as

$$\log E_{B,C,10} = -4 \log T_{B,C} + 3.72 \qquad (6)$$

where the terms are defined similarly as those in Eqs. (1) and (2).

The general equation can be expressed as

$$\log E = -2n \log T + K \qquad (7)$$

where E is the number of total errors per eigengroup, n is the number of domains of the significand of the stimulus panel, T is the mean time for the total number of trials for each eigengroup per system, and K is a constant. The general equation and the definitions of the terms are limited to the results and discussions of the above analyses of the error-time response performance for a singlefold response.

5. DISCUSSION. The purpose of this experiment was to evaluate the human factors of three variations of a display panel by subjects with zero (minimal) bias. The mathematical analysis of the error-time response performances of groups of "unbiased" individuals resulted in a new

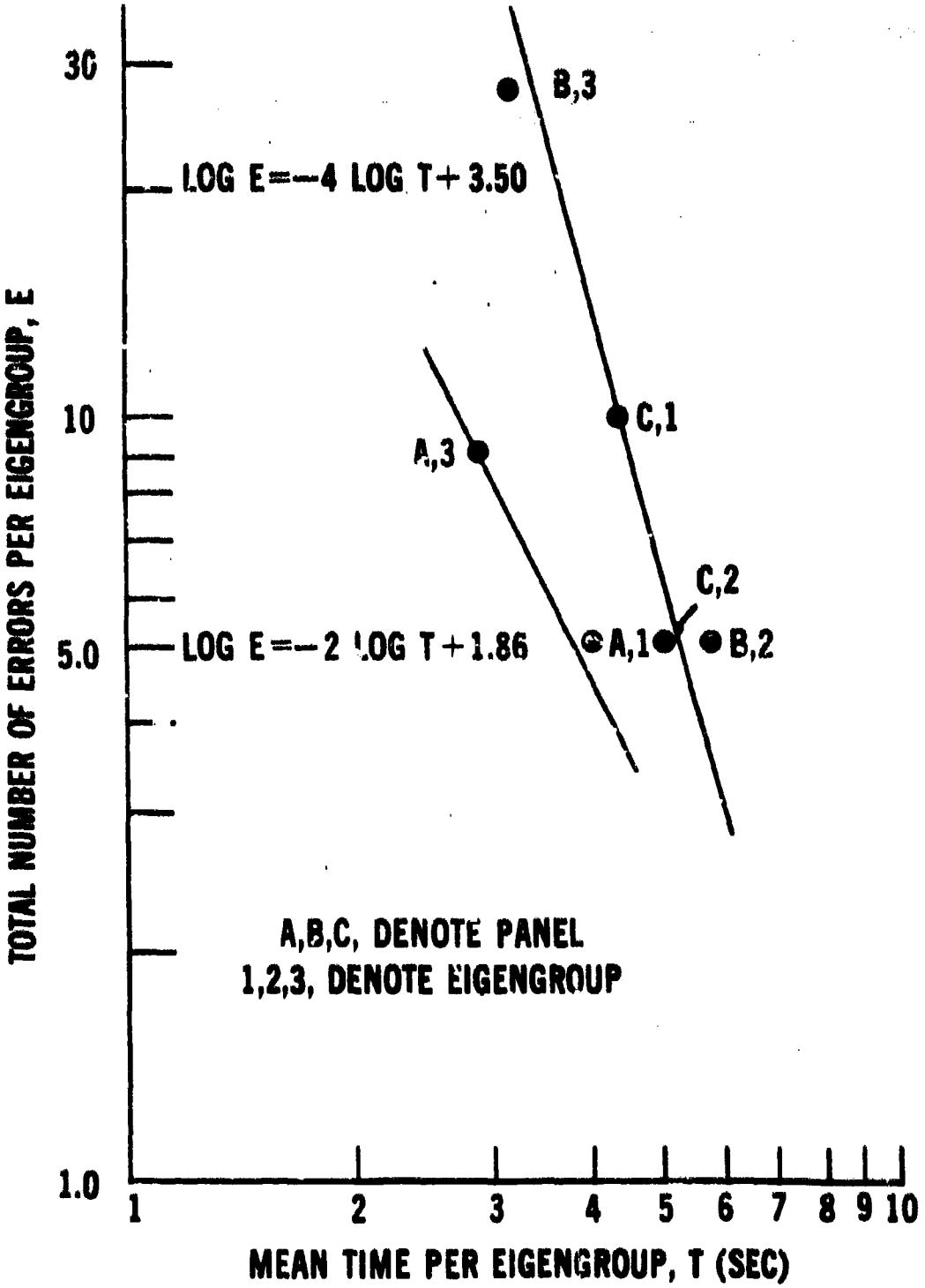


Figure 5. Error dependence on the mean time for panels evaluated per eigengroup, ±20 degree resolution range

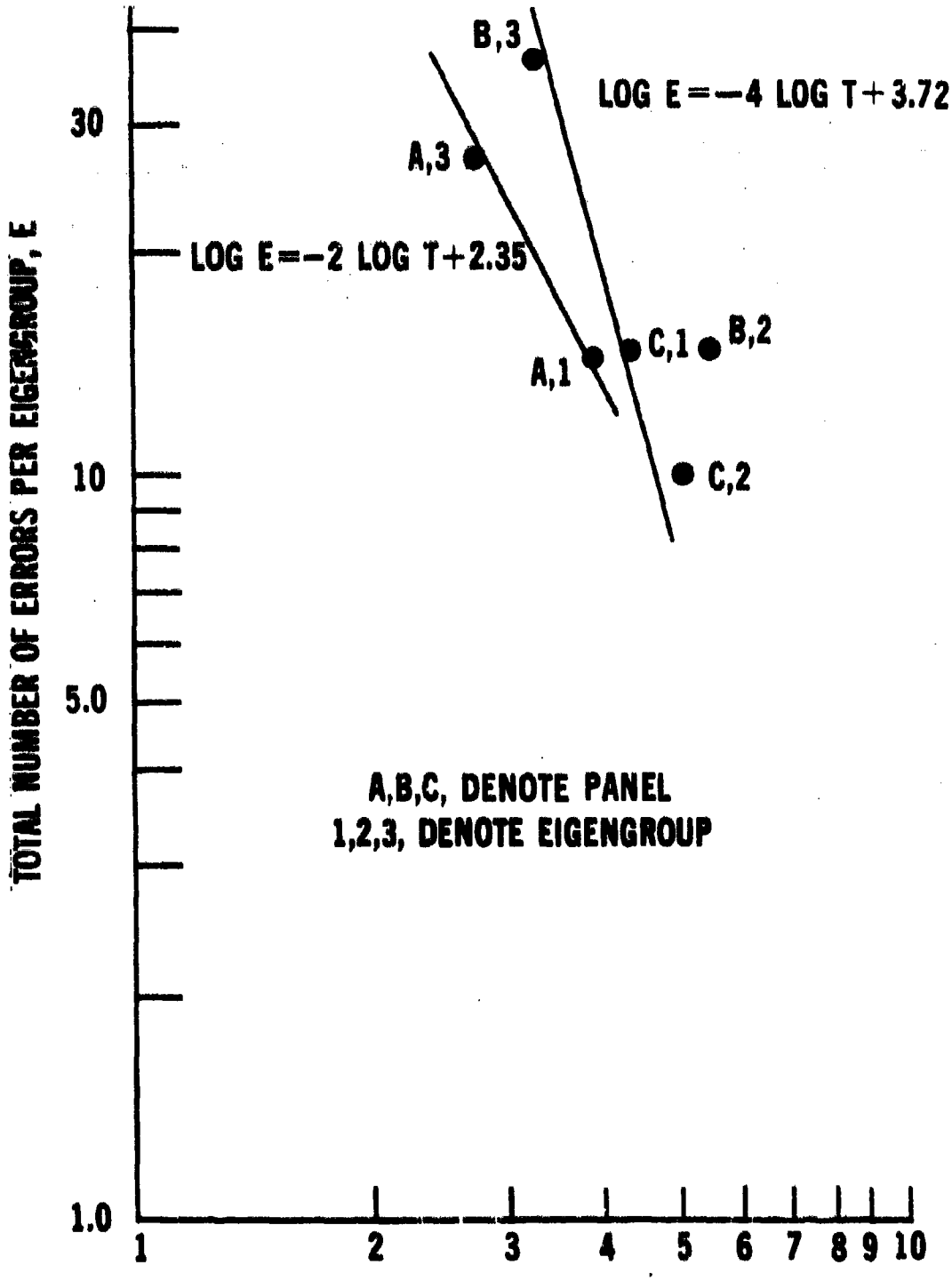


Figure 6. Error dependence on the mean time for panels evaluated per eigengroup, ±10 degree resolution range

frequency-distribution equation. In order to maintain clarity in describing the experimental designs and procedures, it was necessary to introduce and define new parameters which would be relevant to the analyses of the panels and interpretation of the equations. The three display panels were designed to perform the same response function, but the complexity and domains of the significands of the stimulus panels were increased. In particular, the methods of the angular readouts were changed from the circular display of degrees in one direction of one domain of panel A to the two circular sequential displays of mils in the same direction of two domains of panel B, and finally to the alphanumeric readout of mils generating two semicircular displays in opposite directions of two domains of panel C.

The subjects, selected at random for this experiment were considered to be identical, but not matched with respect to knowledge and training associated with the designs. (Random groups are comprised of subjects which would be considered representative of a large assembly of those subjects, whereas matched groups are defined as groups comprised of those subjects whose evaluated characteristics have been found to be similar within some norm of a criterion. Both groups can be considered as eigengroups if they are constrained to being nearly the same state of knowledge and training, and evaluated once and only once for one set of trials for each of the experimental designs.)

The curves for the three error ranges for the panel having one domain (panel A) can be portrayed by an empirical equation, $\log E = f(\log T)$ with each having the same slope of -2. The displacement constant increases from 1.20 to 1.86 to 2.35 with increasing angular readout resolution. The curves for the three ranges for the panels having two domains (panels B and C) can be portrayed by the same empirical equation as above with each having a slope of -4. The displacement constant increases from 3.38 to 3.57 to 3.72 with increasing angular readout resolution. Since the slope of the general equation is $-2n$, where n is the number of domains, and the constants, increasing with increasing angular resolution as well as number of domains, it appears that the general equation is an explicit function of both the number of domains of the panels and the resolution of the response data. This implies that the general equation is independent of the amount of training of the eigengroups. However, it is logical to expect that for a given number of trial sets, the total number of errors per eigengroup would decrease with increased level of training. Since it is not known how the training would effect the equation, if at all, it is assumed that the general equation is an implicit function of training. In order that the error-time response experiment to be meaningful, it is required for the number of trials sets be sufficiently large so that at least one error be committed per trial set for each of the experimental designs.

The increased complexity, i.e., changing the significand from a circular representation of two domains of panel B, to an alphanumeric representation of two domains of panel C had no apparent influence or deviation from the linearity of the curves representing those panels. The lack of deviation is not unexpected as a result of the Gagné and Foster (1949) studies.

It is realized that the analyses presented here are of a small sample evaluation of groups of individual subjects. However, the analysis of variance indicates that trials of the right/wrong decision-timed response performances on the three systems are valid ($F_{(3,65)} = 5.08$; $p<0.001$). The analysis of variance for the systems ($F_{(2,156)} = 8.84$; $p<0.001$) indicated that the systems were different, and that TrialXSystem ($F_{(10,156)} = 0.39$) was not significant. Some learning did occur for the subjects (six trials each), however, the learning did not interact with the systems, and all subjects learned equally to about the same degree.

From the above discussion, it is postulated that the general equation is valid for other error-time response experiments similar to those described in this paper. Efforts were made to apply the error-time response data of Gagné and Foster (1949) and Fitts and Seeger (1953) to the analysis. This was done for the purpose of subjecting Eq. (7) to experimental results of other investigators for corroboration. The error-time response equation could not be generated from the above sources due to the following reasons: (1) the mean time was measured only for the correct choice which included the wrong choices until the correct choice was made; (2) the total number of errors were determined as a function of preliminary and accumulated training; and (3) most importantly, the error-time measurements were not made on eigengroups, i.e., those groups having identical prior knowledge of the panels, and the same acquired learning for each set of trials for the entire series of trial sets. Further investigational work is required to subject the general equation to experimental verification.

6. SUMMARY. The experiment presented here is similar to those reported in literature, and the stimulus-response procedures are standard practices. It is known, in general, that as a subject takes less time to make decisions, considered to be right/wrong, the number of errors increases and the standard deviation becomes larger. However, the experiment here differs on two important aspects with respect to the control of the subjects and data analyses. The first is that the subjects were separated into groups of equally biased knowledge (no pretraining) concerning the panels and were not subjected to a cumulative learning process for the entire series of trial sets. The second aspect is that an error was considered as a discrete response of a right/wrong decision and the errors were analyzed as a function of the mean time of the total number of decisions per eigengroup. The analyses of the error-time

equations necessitated the introduction of new parameters in order to unambiguously define the stimulus panels and interpret the procedures and results consistent with the equations. The general equation is a mathematical expression which, for this experiment, describes the relationship between the number of errors of right/wrong decisions and the mean time in making the decisions.

7. ACKNOWLEDGMENT. The author wishes to thank Mr. Jeff Abraham and Specialist Fifth Class David Hardman, Electronic Test Command, Fort Huachuca, Arizona for assistance in carrying out the test program, and to Mr. Richard Flaherty, Night Vision Laboratory for the analysis of variance as well as valuable discussions.

REFERENCES

Fitts, P.M. and Seeger, C.M., S-R Compatibility: Spatial characteristics of stimulus and response codes, J. Experimental Psychology, 1953, Vol. 46, pp 199-210.

Gagné, R.M., Contributions of learning to human development, Psychological Review, 1968, Vol. 75, pp 177-191.

Gagné, R.M. and Foster, H., Transfer of training from practise on components in a motor skill, J. Experimental Psychology, Vol. 39, pp 47-68, 1949.

RELIABILITY ANALYSIS OF AIRFIELD LIGHTING SYSTEMS

Frank Kuo
Edward S. Lindow
US Army Corps of Engineers
Construction Engineering Research Laboratory
Champaign, Illinois

ABSTRACT. The reliability analysis of a system with multiple types of components under maintenance is a complex problem. This paper presents a model for such analysis with specific application to airport lighting systems. A set of consecutive coefficients is introduced to account for system failure criteria which includes random light outages, consecutive light outages, and consecutive light bar failures. Probability theory and simulation techniques are used along with the consecutive coefficients in determining system reliability. The computerized model has been used in a sensitivity analysis to determine the effect on system reliability of parameters such as unit reliability, system configuration, maintenance strategy, and unit performance characteristics.

1. INTRODUCTION. Visual guidance lighting systems for airports provide necessary information for aircraft operation during the approach, landing, takeoff, and ground movement (taxiing). In darkness, inclement weather or other periods of low visibility, the information provided by these systems is critical to safe and efficient air travel.

Although significant research has been devoted to improving component equipment in these lighting systems and to delineating the pilot's information requirements, little has been done to determine the operational reliability of the systems currently in use. Because these systems are critical to safe and efficient aircraft operations and because installation and maintenance costs for such systems are high, procedures to analyze the reliability of present airfield lighting systems are needed.

The purpose of the research summarized in this paper was to develop procedures for evaluating the functional reliability of airfield lighting systems.

2. AIRFIELD LIGHTING SYSTEM MODEL. There are numerous types of lighting systems involved in the visual guidance of aircraft traffic. The number and the configuration of lights in each system will depend on factors such as the information conveyance requirements, the area to be served, the category of operations, and the terrain.

Although individual systems are comprised of specialized equipment in configurations designed to satisfy specific information requirements, all lighting systems have the common elements of a power source, power circuitry and light transmission equipment. Because of these similarities, a general model can be used to define all visual guidance lighting systems.

The model developed for this purpose consists of 12 types of components: commercial power, auxillary power, control panel, control circuitry, control vault, regulator, primary cable, isolating transformer, secondary cable, fixture, lens, and lamp. Division of the model into these components considered function, maintenance, physical proximity and connection, and failure modes. Some of the components include several elements (e.g., the control vault includes power transformers, relays, switches, etc) while others are composed of a single element (e.g., the lamp).

Figure 1 illustrates the general lighting system model. Since the number of components of each type can be varied (or deleted if not applicable), this model provides the necessary flexibility to define all airfield lighting systems.

By defining the geometry of a system, the operating characteristics, and the failure criteria, any lighting system can be analyzed using this general model.

3. RELIABILITY MODEL. System reliability is typically defined as the probability that a system will perform its intended function in a specific environment for a specified period of time. However, systems which undergo constant maintenance, as is the case with airfield lighting, are composed of equipment of various ages and thus a time period can not be realistically analyzed. For such maintained systems, the steady-state reliability, which can be interpreted as the probability of the system being in a nonfailure state while under operation, is significant.

Figure 2 is a tree structure depicting the parameters which must be considered in analyzing the reliability of airfield lighting systems in the steady state. Essentially three steps are required.

a. Develop the component reliability function for each type of component.

b. Simulate the average light unit reliability.

c. Calculate the system reliability by applying the system failure criteria.

Thus, the reliability model includes both deterministic and stochastic parameters which must be combined by using analytic and simulation procedures. The following sections summarize the procedures employed in the three steps of the model.

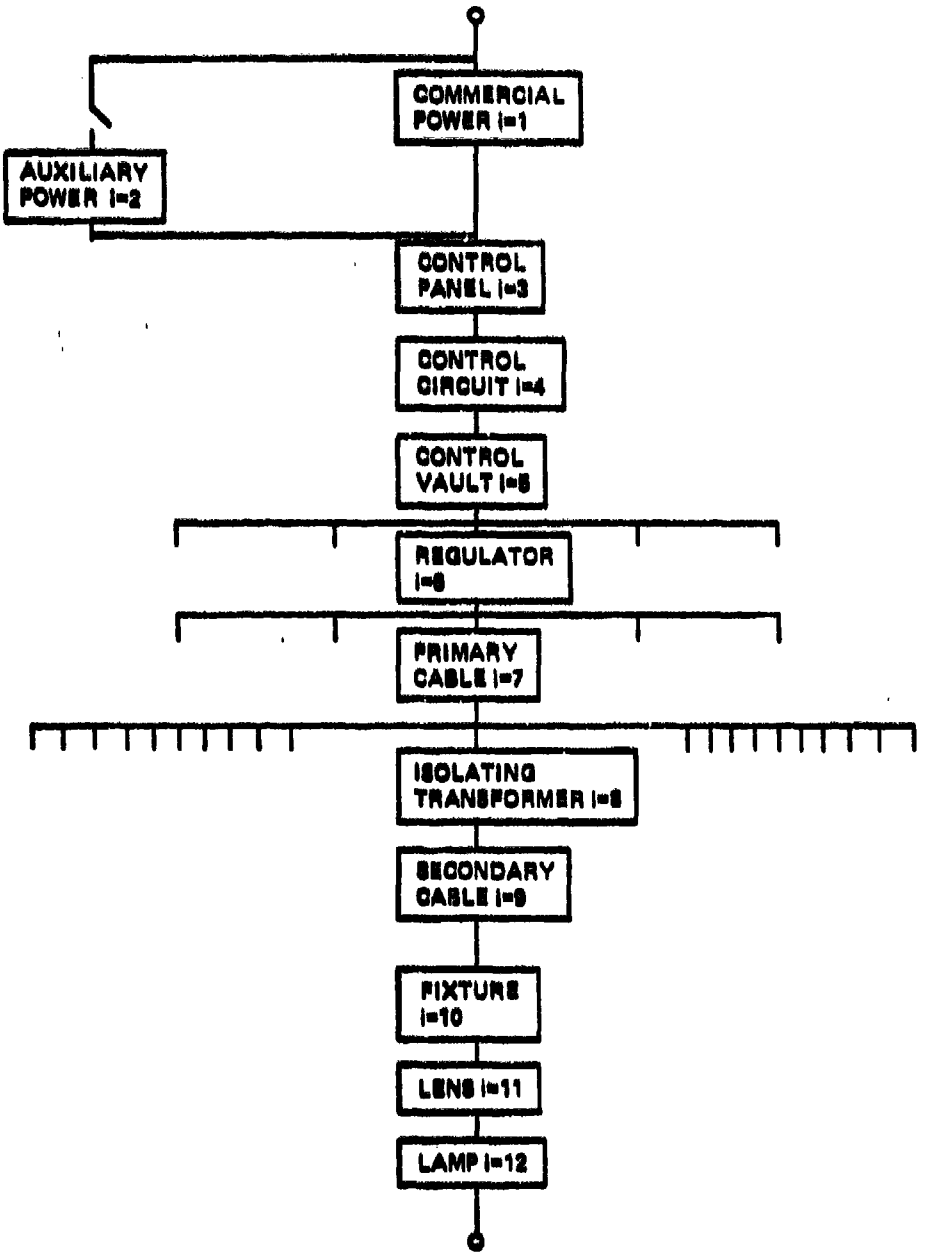


Figure 1
GENERAL LIGHTING SYSTEM MODEL

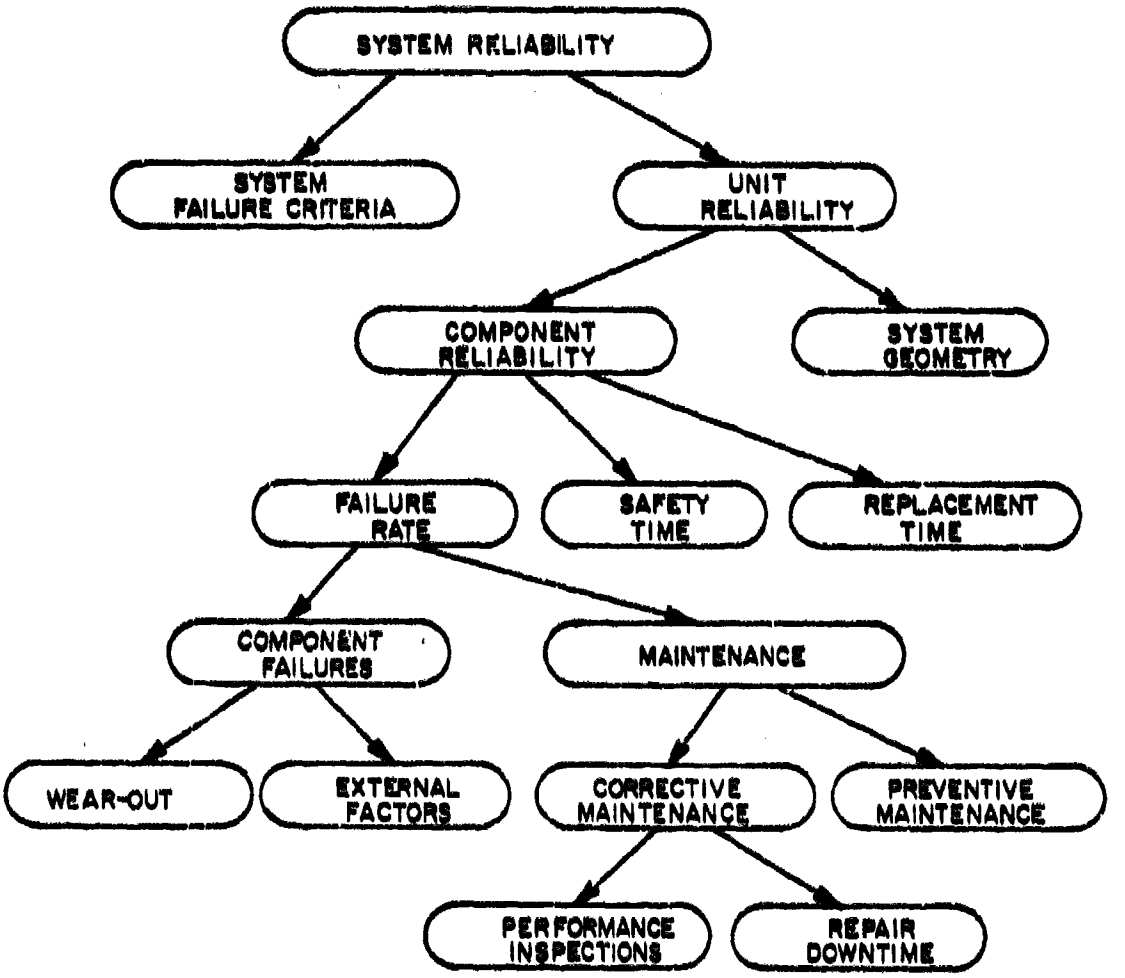


Figure 2

TREE STRUCTURE OF RELIABILITY ANALYSIS

4. COMPONENT RELIABILITY. The reliability of each component type in the general lighting system model can be approximated by an exponential distribution over the component's design life. This distribution is defined by Eq 1 and illustrated in Figure 3.

$$R(t) = e^{-\lambda t} \qquad \text{[Eq 1]}$$

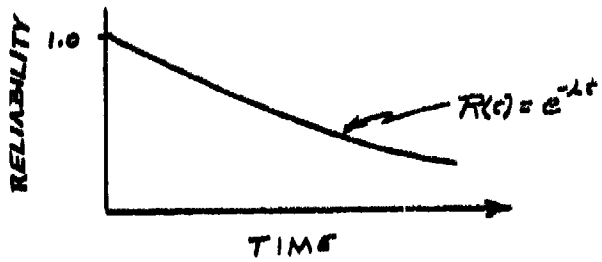


Figure 3. Component Reliability Distribution.

Determination of the failure rate (λ) for each component in the system is quite complex when maintenance and operation practices are considered. Full-scale testing of lighting systems to determine failure rates would be very expensive and time consuming, while accelerated testing of systems or individual components introduces inaccuracies. Thus, field data on system performance are the best source of information for determining a component's reliability function.

Considering the field data anticipated to be available, the reliability function for each component in the lighting system can be expressed by:

$$R(t) = e^{-\frac{(t-t_s)C_m}{C_f}} \qquad \text{[Eq 2]}$$

where C_f = the coefficient of failure

C_m = the coefficient of maintenance

t_s = the safety time (i.e., the period of time when the component is known to have no chance of failure).

The coefficient of failure (C_f) for each component is computed from Eq 3.

$$C_{fi} = \left(\frac{\text{total no. of component } i \text{ in the system}}{\text{no. of component } i \text{ failures per year}}\right) 50{,}000 \qquad \text{[Eq 3]}$$

The coefficient of maintenance (C_m) for each component is computed from Eq 4.

$$C_{mi} = \frac{\ln\left(-C_{fi} \ln R_a\right)}{\ln\left((t_L - t_s)/2\right)} \qquad \text{[Eq 4]}$$

where t_L = design life of component i

t_s = safety time for component i

R_a = average reliability for component i.

$$= 1 - \left(\frac{N_f}{N_o}\right)\left(\frac{T_d}{T}\right)$$

where N_f = failures per year for component i

N_o = total number of component i in the system

T_d = average downtime for component i (hours)

T = operation time per year (hours)

Utilizing these relationships, Eq 2 can then empirically account for preventive maintenance, corrective maintenance, and failure rate. Preventive maintenance (PM) considers the component's design life, replacement time (i.e., that period preceding the design life when group replacement is undertaken), and, indirectly, the level of PM activities (i.e., the more PM performed, the lower the failure rate). Corrective maintenance includes the time to detect a failure and the time required to perform repairs. The failure rate is the annual number of failures of that component type in a system due to all failure modes (e.g., wear-out, human error, etc.).

6. UNIT RELIABILITY. Once the individual component reliabilities have been determined, they can be combined to obtain a unit reliability using Eq 5. The unit reliability (Ru) is defined as the probability that a randomly chosen single unit in the system will be operational when called upon to perform. The unit is composed of one of each component type in the general lighting model as depicted in Figure 4.

$$R_u = \{1 - [1-R_1(t_1)][1-R_2(t_2)]\} \cdot R_3(t_3) \cdot R_4(t_4) \cdot R_5(t_5) \cdot R_6(t_6) \cdot R_7(t_7) \cdot R_8(t_8) \cdot R_9(t_9) \cdot R_{10}(t_{10}) \cdot R_{11}(t_{11}) \cdot R_{12}(t_{12}) \quad \text{[Eq 5]}$$

To determine the average unit reliability, a Monte Carlo simulation routine was developed to stochastically account for the time function and system geometry factors. That is, the component's reliability is actually a function of time and, in the steady-state, the component's reliability may be at any point of time on the function. In addition, the system geometry, or the number of each component in the system, will also influence the average unit reliability. The routine used is illustrated in Figure 5.

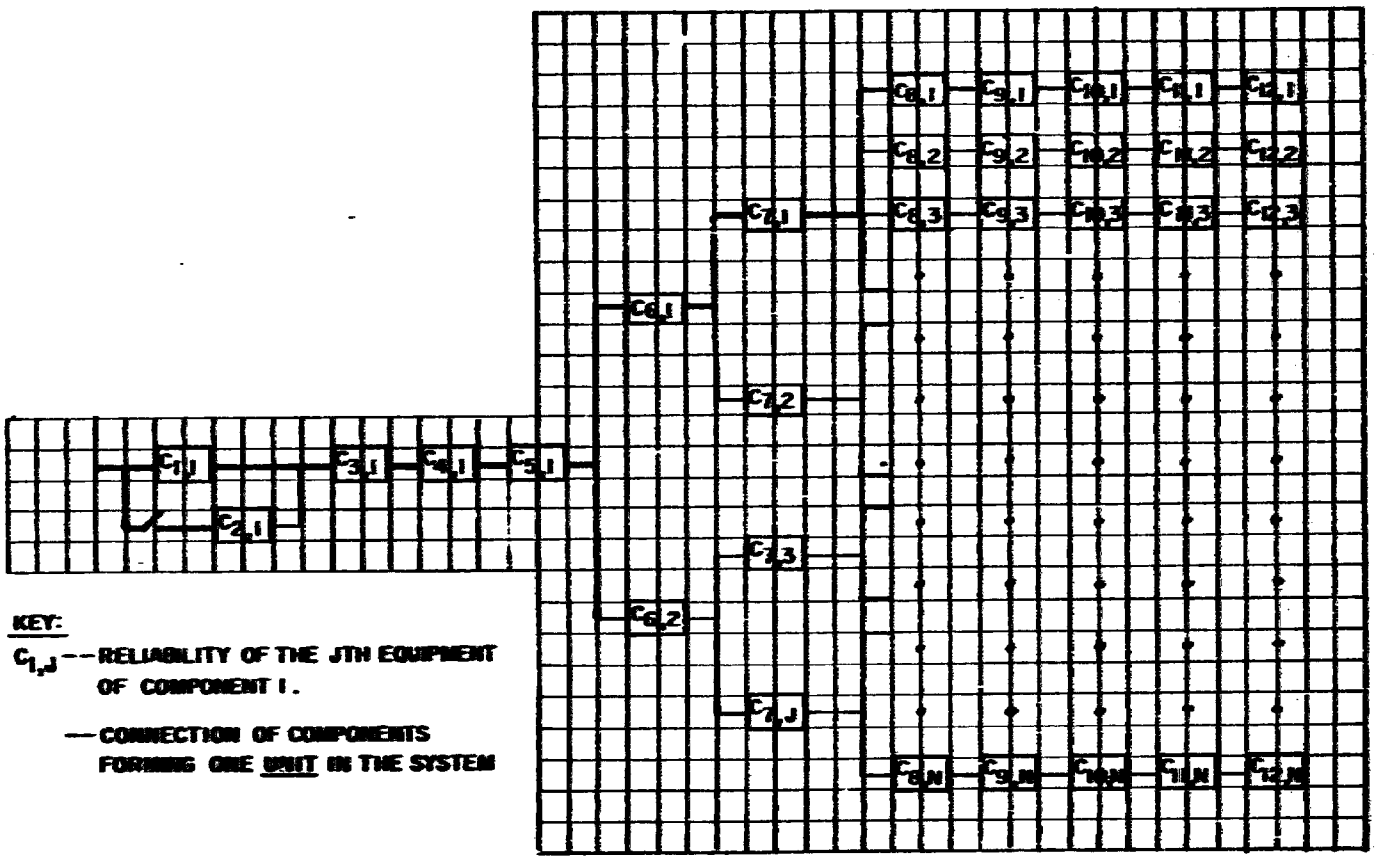


Figure 4

GENERAL RELIABILITY MODEL

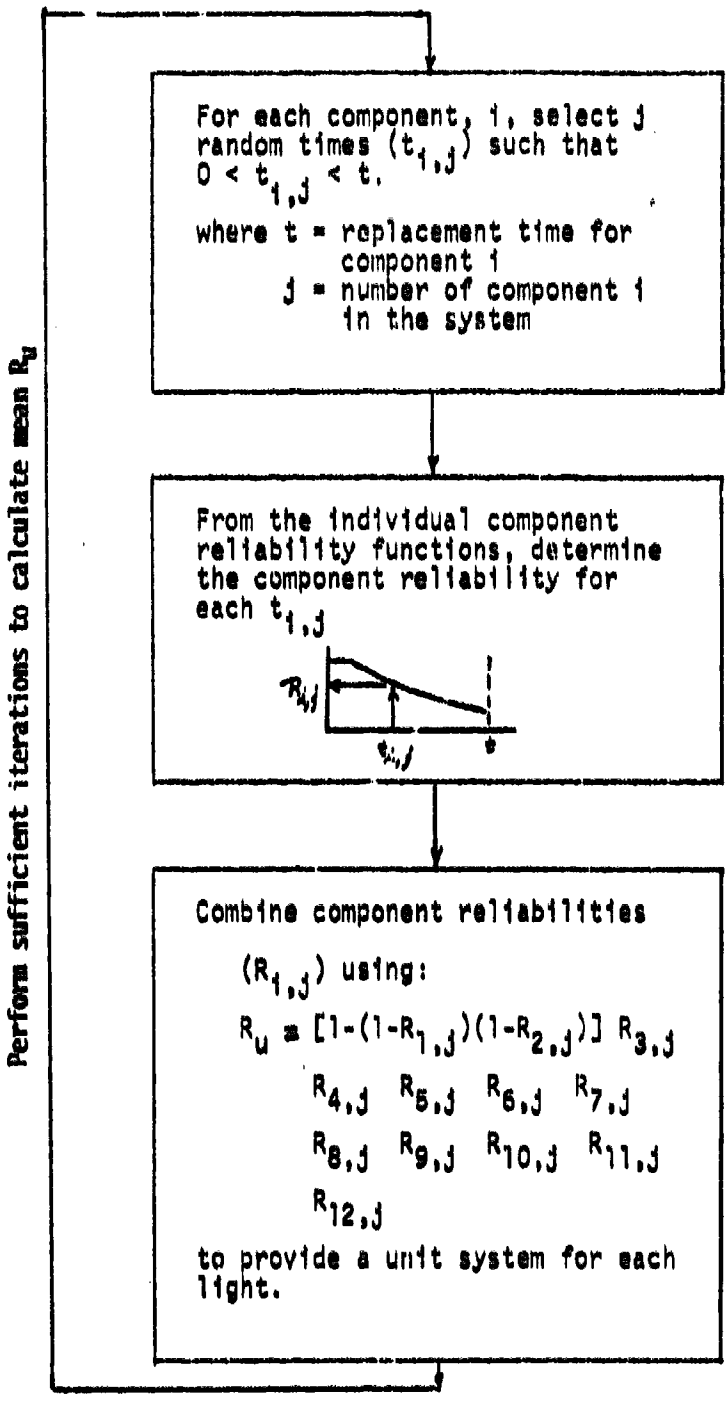


Figure 5. Simplified framework of average unit reliability simulation routine.

6. SYSTEM RELIABILITY. An airfield lighting system fails when it does not accurately transmit the information required by a pilot for safe operation of an aircraft. Since pilot perception is involved, system failure is subjective in nature. Through research studies, the FAA has established objective failure criteria which provides minimum operating standards for each type of lighting system.

In defining failure criteria, the airfield lighting systems have been catagorized as linear and bar systems. The linear system criteria stipulates the percent of random outages and the number of consecutive outages. The bar system criteria stipulates the percent of random outages, the number of outages in a bar creating bar failure, and the number of consecutive bar failures.

Using the appropriate failure criteria and the average unit reliability, the reliability of the lighting system can be determined from Eq 6 for both categories of systems.

$$R_s = \sum_{i=0}^{n} W_i \, R_u^{n-i} \, (1-R_u)^i \qquad \text{[Eq 6]}$$

where R_s = the system reliability

R_u = the unit reliability

n = the total number of lights in the system

i = the number of light failures in the system

W_i = the number of ways i failures can occur in a system of n total lights without the system reaching failure by either the random or the consecutive failure criteria.

The following example illustrates the application of this equation. The example involves finding the system reliability for a three-lamp system ($n=3$) with system failure defined as all three lamps out or two consecutive lamps out. The probability of a lamp being on is R_u. Table 1 shows the eight possible conditions in which this system can be; three are failures and five are successes.

The system reliability is

$$R_s = \sum_{i=0}^{n} W_i \, R_u^{n-1} \, (1-R_u)^1$$

$$R_s = 1R_u^{3-0} (1-R_u)^0 + 3 R_u^{3-1} (1-R_u)^1 + 1R_u^{3-2} (1-R_u)^2 + 0R_u^{3-3} (1-R_u)^3$$

TABLE 1

Possible Conditions for Example

	S	S	S	S	F	S	F	F	S = Success
Lamp 1	0	x	0	0	x	x	0	x	F = Failure
Lamp 2	0	0	x	0	x	0	x	x	0 = Light operating
Lamp 3	0	0	0	x	0	x	x	x	x = Light failed
	i=0		i=1			i=2		i=3	
	$W_i = 1$		$W_i = 3$			$W_i = 1$		$W_i = 0$	

In a linear lighting system, if the consecutive failure criterion is not considered, Eq 6 reduces to a binomial distribution or

$$R_s = \sum_{i=0}^{NR} \binom{n}{i} R_u^{n-i} \, (1-R_u)^i \qquad \text{[Eq 7]}$$

where NR = number of random failures allowed in the system.

To consider consecutiveness as well as random outages in the failure criteria, an analitical procedure has been developed to compute each W_i. Since W_i is a multivariate integer function of n, NC, and i (where NC = number of consecutive failures allowed and n and i as previously defined), there is a unique constant for each (n, NC, i) which is defined here as the consecutive coefficient, C(n, NC, i). This coefficient is the number of ways that i outages can be distributed in n total lights without having more than NC consecutive outages. Substituting the coefficient in Eq 6 produces:

$$R_s = \sum_{i=0}^{NR} W_i \, R_u^{n-i} \, (1-R_u)^i \qquad \text{[Eq 8]}$$

(Note that the summation is from i=0 to i=NR since W_i goes to zero when the number of outages, i, exceeds the allowable random outages, NR).

An automated procedure is used to compute the consecutive coefficients based on the following recursive function:

$$C(n, NC, i) = C(n-1, NC, i) + C(n-1, NC, i-1) - C(n-NC-2, NC, i-NC-1) \qquad \text{[Eq 9]}$$

The derivation and development of the program may be found elsewhere.[1]

The method for analyzing the bar lighting systems is similar. However, the determination of W_i is much more complex due to the nature of the bar system failure criteria. A detailed description of the bar system analitical technique is given in the project final report.[1]

---

[1] Lindow, E. S. and Yuo, F. "Reliability Analysis For Airfield Lighting Systems" Final Report for Contract DOT-FAGGWAI-118, CERL, September 1976.

7. THE RAALS PROGRAM. The reliability methodology summarized in the previous sections would be difficult to apply manually when considering the number of lights in a system, the stochastic properties of the component reliabilities, and the sophistication of the failure criteria. Thus, the procedures have been computerized in the RAALS (Reliability Analysis of Airfield Lighting System) program. This program is capable of efficiently estimating the functional reliability of any lighting system used in the visual guidance of aircraft. Flexibility is provided in the program to consider various system configurations and failure criteria as well as different component failure rates, design lives, and levels of maintenance.

Figure 6 is a simplified flow chart of the RAALS program. Figure 7 presents the input data listing, a typical component reliability function, and the system reliability output resulting from an example problem.

8. CONCLUSIONS. The automated procedure for analyzing reliability of airfield lighting systems (RAALS) is an implementable tool which can be used to:

a. Compare the reliability of similar systems,

b. Determine where a system should be improved to increase its reliability,

c. Form a basis for decisions on implementing changes to failure criteria, equipment, or maintenance policies,

d. Monitor the reliability of a system as it becomes older or as modifications are installed.

The RAALS program logic is based on traditional reliability theory. However, due to the number and complexity of lighting systems and the necessity to consider consecutiveness in the failure criteria, original analytical techniques were developed and interfaced with traditional theory. These techniques included:

a. Formulation of a general lighting system model capable of considering all of the diverse equipment and geometry encountered in airfield lighting

b. Adaptation of a Monte Carlo simulation routine to the analysis to account for the stochastic nature of the component reliabilities

c. Derivation of the consecutive coefficient to consider consecutiveness in the system failure criteria

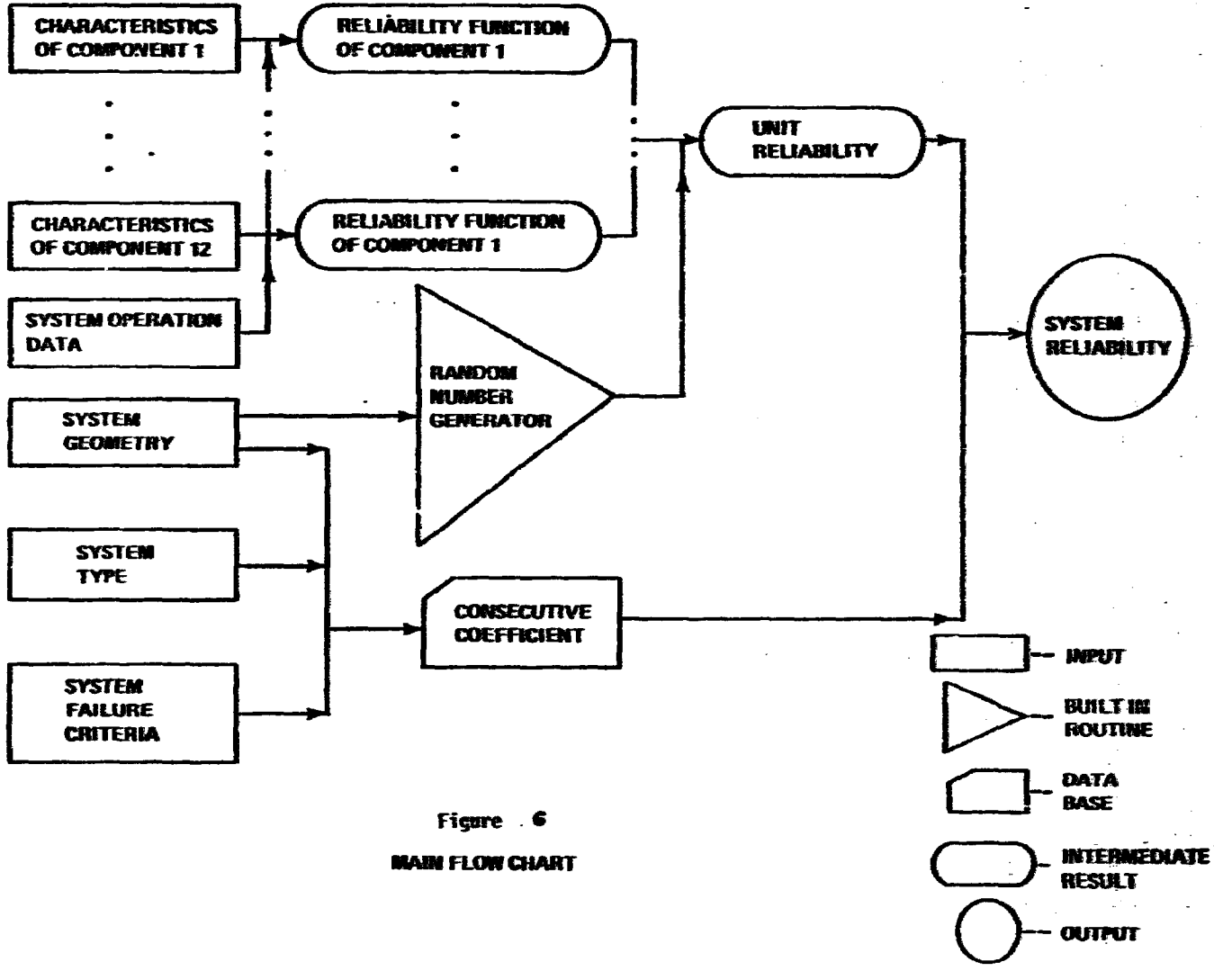


Figure 6

MAIN FLOW CHART

LOCATION: EXAMPLE PROBLEM

SYSTEM TYPE: ALSF-2 CENTERLINE

COMPONENT CHARACTERISTICS

COMPONENT TYPE	NUMBER	FAILURES PER YEAR	REPLACEMENT TIME (HRS.)	SAFETY TIME (HRS.)	DOWN TIME (HRS.)	DESIGN LIFE (HRS.)	MAINTENANCE COEFFICIENT	FAILURE COEFFICIENT	AVERAGE COMPONENT AVAILABILITY
COMMERCIAL POWER	1	1	175000.	0.	24.	175000.	.493504	50000.	.994521
AUXILIARY POWER	1	1	175000.	0.	36.	175000.	.529256	50000.	.991781
CONTROL PANEL	1	1	175000.	0.	24.	175000.	.493504	50000.	.994521
CABLE TO VAULT	1	1	100000.	0.	24.	100000.	.519029	50000.	.994521
CONTROL VAULT	1	3	100000.	0.	12.	100000.	.455089	16666.	.991781
REGULATOR	3	2	175000.	0.	24.	175000.	.493423	75000.	.994347
PRIMARY CABLE	3	3	100000.	0.	12.	100000.	.456439	50000.	.991764
TRANSFORMER	150	15	60000.	0.	24.	60000.	.544508	500000.	.994457
SECONDARY CABLE	150	40	60000.	0.	48.	60000.	.611968	187500.	.997878
FIXTURE	150	25	60000.	0.	36.	60000.	.583879	300000.	.995638
LENSES	150	75	40000.	0.	24.	40000.	.566911	100000.	.997760
LAMP	150	160	8000.	0.	24.	10000.	.659367	46875.	.994155

OPERATION TIME:

HOURS PER DAY = 12.0
DAYS PER YEAR = 365.0

FAILURE CRITERIA:

1. 20. % LIGHT FAILURES CONSTITUTES SYSTEM FAILURE.
2. 2 CONSECTIVE LIGHT BAR FAILURES CONSTITUTES SYSTEM FAILURE
3. 30% LIGHT FAILURES IN A BAR CONSTITUTES LIGHT BAR FAILURE.

Figure 7

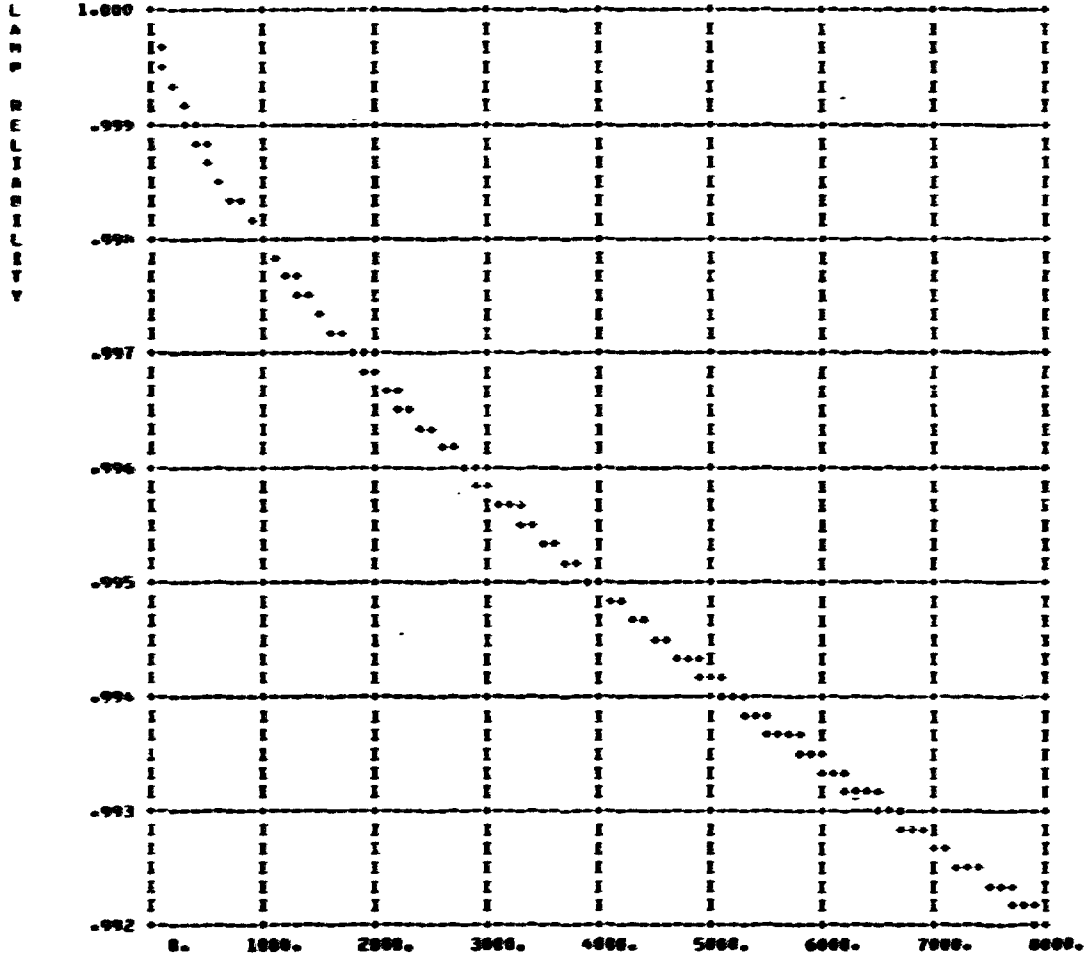
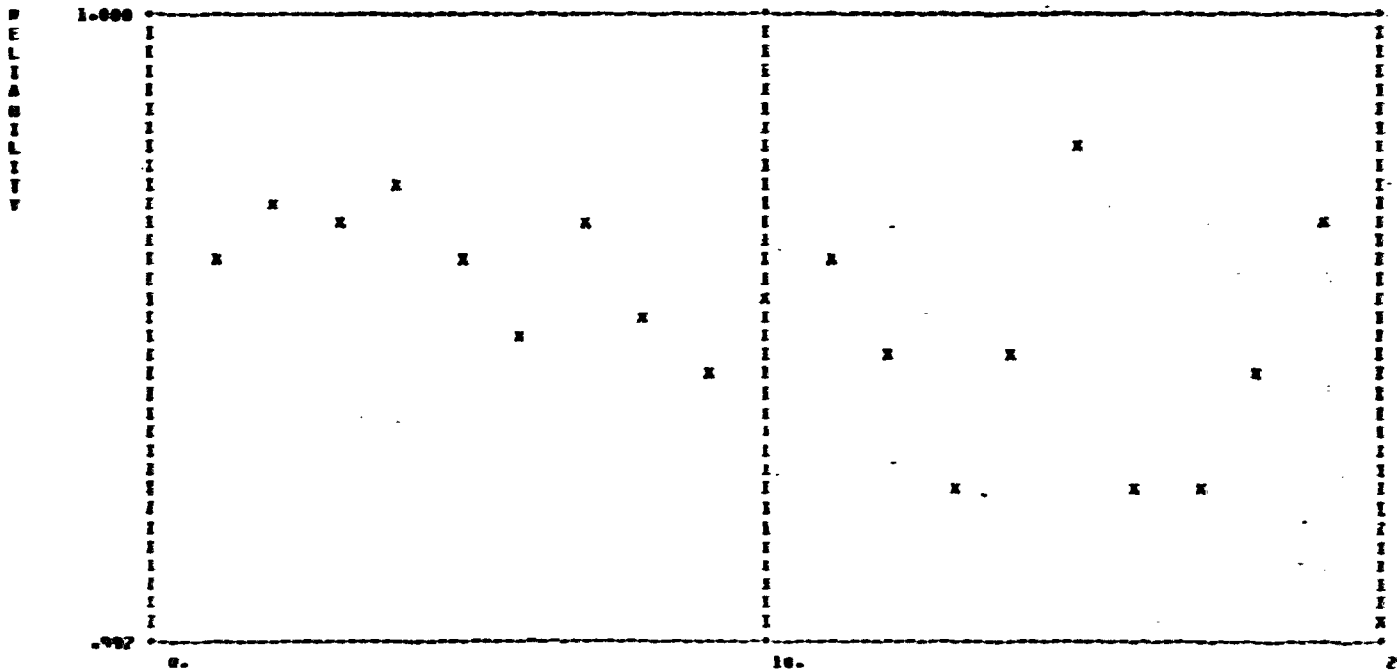


Figure 7 (cont'd)

ITERATION	MEAN UNIT REL	VAR UNIT REL	SYSTEM REL
1	.96436787	.00000719	.99679771
2	.96893124	.00000717	.99763923
3	.96776483	.00000656	.99727936
4	.96948571	.00000978	.99779753
5	.96633605	.00000555	.99671635
6	.96412021	.00000571	.99589091
7	.96773060	.00000609	.99727805
8	.96471664	.00000473	.99615103
9	.96208355	.00000581	.99532972
10	.96536044	.00000721	.99641749
11	.96642035	.00000724	.99681630
12	.96384437	.00000584	.99578305
13	.96815716	.00000542	.99738221
14	.96348752	.00000554	.99561336
15	.97101224	.00000771	.99819337
16	.96011214	.00000761	.99385598
17	.96040576	.00000666	.99402683
18	.96317831	.00000626	.99546975
19	.96764212	.00000736	.99723832
20	.95745191	.00000673	.99215189

Figure 7 (cont'd)

BY INPUT DATA SET # 6 THE MEAN SYSTEM RELIABILITY IS .996015988
THE VARIANCE OF IT IS .000002442



d. Development of an anlaytical procedure to determine system reliability which accounts for the operation, maintenance, and failure variables of each component

e. Automation of the combined procedures into a concise, efficient computer program.

Although this research effort was devoted to airfield lighting systems, the methodology developed is applicable to any system which can be similarly defined and for which failure criteria stipulate consecutive failures as well as random failures.

ACKNOWLEDGEMENT

This research was conducted for the Federal Aviation Administration, Visual Aids Section, under Inter-Agency Agreement No. DOT-FA66WAI-118. Mr. Walter C. Fisher, Chief, Visual Aids Section, was the Technical Monitor.

The work was performed by the Military and Base Engineering Branch, Facility Operations Division, of the U.S. Army Construction Engineering Research Laboratory, Champaign, Illinois.

The opinions expressed in this paper are those of the author and do not necessarily reflect the views of the U.S. Army Construction Engineering Research Laboratory or the Federal Aviation Administration.

More detailed information on this project and the RAALS program are contained in the Final Report for contract DOT-FAGGWAI-118.

SIMPLIFIED METHOD FOR DETERMINING APPROXIMATE LOWER CONFIDENCE BOUNDS OF A SYSTEM WHOSE POSTERIOR RELIABILITY DISTRIBUTION IS DESCRIBED AS A BETA

Louis M. Iannuzzelli
Product Assurance Directorate
US Army Armament Command

and

Robert Dostal
US Army Management Engineering Training Agency
Rock Island, IL 61201

ABSTRACT. This paper deals with a simplified method of determining the approximate lower confidence bounds on reliability of a system, given the system posterior reliability beta parameters A'_s and B'_s (integer or non-integer) and/or trials and failures observed and the interval desired. Prior to the development of this method, a computer was utilized to determine the lower bounds due to the fact that the beta parameters were, for the most part, non-integer. The method described in this paper was empirically developed and provides a method of determining approximate reliability bounds very simply with the use of a SR 51, HP 45 etc., hand calculator. The unsolved problem simply stated is "Why does the method work as well as it does?"

1. INTRODUCTION. A need arose in ARMCOM for a simplified method of determining approximate lower bounds on reliability, given subsystem data, a model and the confidence interval desired. As a result, a literature search was made of current available methods. These methods are referred to by comparison in our paper titled, "Confidence Limits for System Reliability When Testing Takes Place at the Component Level," dtd 31 Oct 75. Based on the review of the current available methods, it was decided to see if a more simplified method could be developed which would overcome some of the shortcomings of the current methods and still provide results which would satisfy our needs. A method was developed as described in reference paper; however, the mathematical expression derived empirically for calculating the lower bound is still, to this day, not fully understood.

2. THE LOWER BOUND ON RELIABILITY. The lower bound on reliability is determined as follows:

Given: A_s, B_s system posterior reliability parameters of a beta

$1 - \alpha$ = Confidence interval desired

$f_1 = 2B_s$

$f_2 = 2A_s$

$f_1 < f_2$

$$R_{(1-\alpha)} = \exp\left(\frac{-\chi^2_{\alpha;f_1}}{f_1 + f_2 - 2 - (f_1 - 2)\ .831\ \left\{1 - \left(\frac{1}{f_1 + 1} - \frac{1}{f_2 + 1}\right)\right\} Q^Q} \right)$$

where $Q = \frac{f_1}{f_1 + f_2}$

$f_1 \geq f_2$

$$\bar{R}_{(1-\alpha)} = \exp\left(\frac{-\chi^2_{(1-\alpha);f_2}}{f_1 + f_2 - 2 - (f_2 - 2)\ .831\ \left\{1 - \left(\frac{1}{f_2 + 1} - \frac{1}{f_1 + 1}\right)\right\} Q'^{Q'}} \right)$$

where $Q' = \frac{f_2}{f_1 + f_2}$

$$R_{(1-\alpha)} = 1 - \bar{R}_{(1-\alpha)}$$

3. EXPRESSION FOR BINOMIAL DATA.

Given: N = No. of trials

f = failures

$1 - \alpha$ = Confidence interval

$f + 1 < (N - f)$

$$R_{(1-\alpha)} = \exp\left(\frac{-\chi^2_{\alpha;2(f+1)}}{2\left\{N - f\ .831\ \left\{1 - \left(\frac{1}{2f + 3} - \frac{1}{2(N - f) + 1}\right)\right\} Q^Q\right\}} \right)$$

where $Q = \frac{f + 1}{N + 1}$

$f + 1 \geq (N - f)$

$$\bar{R}_{(1-\alpha)} = \exp\left(\frac{-\chi^2_{(1-\alpha);\,2(N-f)}}{2\{N-((N-f)-1).831\{1-(\frac{1}{2(N-f)+1}-\frac{1}{2f+3})\}Q'Q}\right)$$

where $Q' = \frac{N-f}{N+1}$

$$R_{(1-\alpha)} = 1 - \bar{R}_{(1-\alpha)}$$

A fractional chi-square table is required; however, linear interpolation can be utilized.

4. PROBLEM. An understanding of the expression:

$$R_{(1-\alpha)} = \exp\left(\frac{-\chi^2_{\alpha;f_1}}{f_1+f_2-2-(f_1-2).831\{1-(\frac{1}{f_1+1}-\frac{1}{f_2+1})\}Q^Q}\right)$$

where $Q = \frac{f_1}{f_1+f_2}$

is needed in order to provide an answer to the many inquiries concerning the mathematical validity of the above expression.

5. RESULTS. Many values of A_s and B_s, both integer and non-integer, were compared. The values shown are just a few of the comparisons made. Other comparisons at different confidence intervals were made as shown in Table 1 through 4.

For whatever help it may be, the relationship between the F distribution and the expression was found to be:

$$F_{\alpha;\,f_1,\,f_2} = \frac{f_2}{f_1}\left(\frac{1-p}{p}\right);\ f_1 < f_2$$

where $p = \exp(-K)$

$$F_{\alpha;\,f_1,\,f_2} = \frac{f_2}{f_1}\left(\frac{1}{\exp(-K)}-1\right)$$

$$F_{\alpha;\,f_1,\,f_2} = \frac{f_2}{f_1}(\exp(K)-1)$$

$$K = \frac{\chi^2_{\alpha;f_1}}{f_1+f_2-2-(f_1-2).831\{1-(\frac{1}{f_1+1}-\frac{1}{f_2+1})Q^Q}$$

$$Q = \frac{f_1}{f_1 + f_2}$$

From this expression, approximate values of the F distribution can be obtained for non-integer degrees of freedom.

A_β	B_β	CONFIDENCE	TRUE	CALCULATED	ERROR
4.18	8.00	99	.091	.091	0
		95	.144	.144	0
		90	.179	.179	0
102.34	28.44	99	.693	.692	+.001
		95	.721	.721	0
		90	.735	.735	0
7.01	6.04	99	.234	.237	-.003
		95	.314	.315	-.001
		90	.361	.361	0
7.49	6.09	99	.251	.254	-.003
		95	.332	.333	-.001
		90	.379	.379	0
18.71	3.99	99	.607	.606	+.001
		95	.681	.679	+.002
		90	.718	.717	+.001
10.08	5.00	99	.559	.557	+.002
		95	.632	.631	+.001
		90	.670	.668	+.002
17.84	4.99	99	.555	.554	+.001
		95	.629	.627	+.002
		90	.667	.666	+.001

TABLE 1

A_S	B_S	CONFIDENCE	TRUE	CALCULATED	ERROR
98.23	4.98	99	.891	.890	+.001
		95	.913	.912	+.001
		90	.924	.923	+.001
94.21	10.99	99	.816	.814	+.002
		95	.843	.841	+.002
		90	.856	.854	+.002
38.58	.75	99	.901	.901	0
		95	.937	.937	0
		90	.953	.953	0
6.12	.71	99	.521	.519	+.002
		95	.669	.668	+.001
		90	.743	.742	+.001
49.45	.52	99	.934	.934	0
		95	.961	.961	0
		90	.972	.972	0
647.45	.63	99	.99426	.99432	-.00006
		95	.99651	.99656	-.00005
		90	.99744	.99750	-.00006
122.23	.68	99	.96919	.96916	+.00003
		95	.98104	.98102	+.00002
		90	.98604	.98602	+.00002
49.86	.35	99	.945	.944	+.001
		95	.970	.970	0
		90	.980	.980	0

TABLE 2

LOWER BOUNDS OF SINGLE TAIL CONFIDENCE INTERVAL (BINOMIAL DATA)

N	f	.90 OPT	.90 CAL.	.90 ERROR	.95 OPT	.95 CAL.	.95 ERROR	.99 OPT	.99 CAL.	.99 ERROR
3	0	.464	.464	0	.368	.360	0	.215	.215	0
	1	.196	.198	-.002	.135	.137	-.002	.059	.060	-.001
	2	.035	.035	0	.017	.017	0	.003	.003	0
6	0	.681	.681	0	.607	.607	0	.464	.464	0
	1	.490	.492	-.002	.418	.421	-.003	.294	.298	-.004
	2	.333	.336	-.003	.271	.275	-.004	.173	.179	-.006
	3	.201	.202	-.001	.153	.154	-.001	.085	.085	0
	4	.093	.092	+.001	.063	.063	0	.027	.027	0
	5	.017	.017	0	.009	.009	0	.002	.002	0
10	0	.794	.794	0	.741	.741	0	.631	.631	0
	1	.663	.663	0	.606	.606	0	.496	.496	0
	2	.550	.551	-.001	.493	.494	-.001	.388	.390	-.002
	3	.448	.449	-.001	.393	.395	-.002	.297	.300	-.003
	4	.354	.354	0	.304	.305	-.001	.218	.222	-.004
	5	.267	.271	-.004	.222	.226	-.004	.150	.153	-.003
	6	.188	.189	-.001	.150	.151	-.001	.093	.094	-.001
	7	.116	.116	0	.087	.087	0	.048	.048	0
	8	.054	.054	0	.037	.037	0	.016	.016	0
	9	.009	.010	-.001	.005	.005	0	.001	.001	0
30	0	.926	.926	0	.905	.905	0	.858	.858	0
	1	.876	.876	0	.851	.851	0	.798	.798	0
	2	.832	.831	+.001	.805	.804	+.001	.748	.747	+.001
	3	.791	.789	+.002	.761	.760	+.001	.702	.701	+.001
	4	.751	.749	+.002	.720	.719	+.001	.660	.658	+.002
	5	.713	.711	+.002	.681	.679	+.002	.619	.617	+.002
	6	.675	.673	+.002	.643	.641	+.002	.580	.578	+.002
	7	.639	.636	+.003	.606	.604	+.002	.543	.541	+.002
	8	.603	.601	+.002	.570	.568	+.002	.507	.505	+.002
	9	.568	.566	+.002	.535	.533	+.002	.473	.471	+.002
	10	.534	.531	+.003	.501	.498	+.003	.439	.437	+.002
	11	.500	.497	+.003	.467	.465	+.002	.406	.405	+.001
	12	.467	.464	+.003	.434	.432	+.002	.374	.373	+.001
	13	.434	.431	+.003	.402	.399	+.003	.343	.342	+.001
	14	.401	.398	+.003	.370	.367	+.003	.313	.312	+.001
	15	.370	.376	-.006	.339	.345	-.006	.284	.290	-.006
	16	.338	.344	-.006	.308	.314	-.006	.256	.260	-.004
	17	.308	.312	-.004	.279	.283	-.004	.228	.232	-.004
	18	.277	.281	-.004	.250	.253	-.003	.201	.204	-.003
	19	.248	.250	-.002	.221	.224	-.003	.176	.178	-.002
	20	.218	.221	-.003	.193	.196	-.003	.151	.153	-.002
	21	.190	.192	-.002	.165	.168	-.003	.127	.129	-.002
	22	.162	.164	-.002	.140	.142	-.002	.104	.106	-.002
	23	.135	.136	-.001	.115	.116	-.001	.083	.084	-.001

TABLE 3

LOWER LIMITS OF 50% CONFIDENCE INTERVAL (BINOMIAL DATA)

N	f	OPT.	CAL.	ERROR	N	f	OPT.	CAL.	ERROR
2	0	.707	.707	0	20	16	.181	.182	-.001
	1	.293	.293	0		17	.131	.132	-.001
						18	.083	.083	0
3	0	.794	.794	0		19	.034	.034	0
	1	.500	.501	-.001					
	2	.206	.206	0	30	0	.977	.977	0
						1	.945	.945	0
6	0	.891	.891	0		2	.912	.911	+.001
	1	.736	.736	0		3	.879	.878	+.001
	2	.579	.578	+.001		4	.846	.845	+.001
	3	.421	.422	-.001		6	.780	.778	+.002
	4	.264	.264	0		8	.714	.712	+.002
	5	.109	.109	0		10	.648	.645	+.003
						12	.582	.578	+.004
10	0	.933	.933	0		14	.516	.511	+.005
	1	.838	.838	0		16	.451	.455	+.004
	2	.741	.741	0		18	.385	.388	-.003
	3	.645	.644	+.001		20	.319	.321	-.002
	4	.548	.546	+.002		22	.253	.255	-.002
	5	.452	.454	-.002		24	.187	.188	-.001
	6	.355	.356	-.001		26	.121	.122	-.001
	7	.259	.259	0		27	.088	.089	-.001
	8	.162	.162	0		28	.055	.055	0
	9	.067	.067	0		29	.023	.023	0
20	0	.966	.966	0					
	1	.917	.917	0					
	2	.869	.868	+.001					
	3	.819	.818	+.001					
	4	.770	.769	-.001					
	5	.721	.720	+.001					
	6	.672	.670	+.002					
	7	.623	.620	+.003					
	8	.574	.570	+.004					
	9	.525	.520	+.005					
	10	.475	.480	-.005					
	11	.426	.430	-.004					
	12	.377	.380	-.003					
	13	.328	.330	-.002					
	14	.279	.280	-.001					
	15	.230	.231	-.001					

TABLE 4

EVALUATION OF GUNNER ERRORS THROUGH TIME SERIES ANALYSIS

Latricha Greene and John Howerton
Systems Evaluation
Aeroballistics Directorate
US Army Missile Research and Development Command
Redstone Arsenal, Alabama 35809

ABSTRACT

This paper describes a procedure used at the Army Missile Command (primarily with command to line of sight systems) for modelling man in the loop. The model developed here with its parameters can be used to simulate data or to drive a total systems simulation.

The procedure outlined here was developed initially by L. Greene, J. Howerton, N. Rich, and M. Wise of the Army Missile Command in conjunction with M. Yang from the University of Florida for the optical mode of Air Defense Systems in which a man was used to track the target. Current plans call for using this same technique to evaluate tracking radars during an ECM environment.

The analysis of the original work as described here was concerned only with stationary data.

1. Introduction

Prediction of the amount of error due to gunner tracking of a moving target is an important phase in the development of weapon systems. Data of this type occur in the form of time series. The observations are dependent and the nature of this dependence is of utmost importance.

The purpose of this paper is to present a method for evaluating gunner error data described below, thereby defining a time series model. This model and its parameters can be used to simulate data for future problems of a similar nature or may be used as a subroutine to missile flight simulation.

2. Data Description

The initial tests to determine the gunner tracking error characteristics were conducted at Redstone Arsenal during the period 13 through 18 July 1972. The King Air, a twin engine Beechcraft, was the target utilized for these tests.

A 16mm film camera was attached to the monocular output of the tracker unit. This output presents the same view to the film camera as the binocular output presents to the gunner.

There were four gunners who participated in the tests. They were instructed to track the centroid to the target aircraft when details were not resolvable. When resolvable they were to track the intersection at the wing and fuselage. The amount of error was shown to be independent of individual gunner, that is, there was no statistical significance.

3. Model Building

This section discusses the time series model building for the gunners' error data. After examining all the data available, we conclude that the data forms a stationary time series except at the beginning where a transient occurs during acquisition, and at the end where a transient is introduced by the simulated missile in flight signal. Runs with too few data were eliminated. The total number of runs was then 143. A few nonstationary data can also be seen. They occupy 13.29 percent of the total.

When the data are recorded with equally spaced time intervals, we generally use a linear time series model to fit the data. A commonly used model for univariate time series can be written as

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \ldots + \phi_p(Y_{t-p} - \mu)$$

$$+ a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \qquad (1.1)$$

where

subscript t = time

Y_t = the value of the time series at time t

μ = the expected value of Y_t

a_t = a white noise process, i.e., a_t is independent, identically distributed $N(0, \sigma_a^2)$

p, q = two parameters depending on the properties of a particular time series.

Model (1.1) is called a mixed model with autoregressive and moving average components. It has been widely used in practice with fruitful results (see e.g., Box and Jenkins [1], Fuller and Tsokos [2], Cleveland [3, 4], and Box et. al [5]). The intuitive idea behind the model (1.1) is the assumption that the present value Y_t depends on the values of Y_t in the near past, i.e., $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$. This is the autoregressive component

$$(Y_t - \mu) = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \ldots + \phi_p(Y_{t-p} - \mu) \quad .$$

The moving average component $a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$ indicates that the present value Y_t depends not only on the present noise a_t, but also the previous noise $a_{t-1}, \ldots, a_{t-q}$. This is reasonable since the noise will not diminish very rapidly in real situations. The noise prolongs its influence on Y_t for a certain period.

In practice when time series data are given, a model of the form (1.1) can generally be built. The detailed procedure has been given in Box and Jenkins [1]. There are four main steps.

a. Model Identification

In this first step, autocorrelation coefficients, partial autocorrelation coefficients, and inverse correlation coefficients (e.g., Cleveland [3]) are used to determine the values of p and q in

model (1.1). The value p is called the order of the autoregressive component and the value q is called the order of the moving average component in a mixed model (1.1).

b. Parameter Estimation

After the values of p and q have been determined, there are $p + q + 2$ parameters: $\mu, \phi_1, \phi_2 \ldots, \phi_p, \theta_1, \ldots, \theta_q$ and the variance σ_a^2 of a_t to be determined. The method used to estimate ϕ's and θ's has been described in Box and Jenkins (Chapter 7, [1]), Clevenson [6], and Parzen [7]. The main technique is the maximum likelihood estimation. Generally, the calculation needs the help of spectral density estimation [7] or nonlinear least squares estimation [1].

c. Diagnostic Checking

The estimated values $\hat{\mu}$, $\hat{\phi}$, $\hat{\theta}$, and $\hat{\sigma}_a^2$ of the parameters μ, ϕ, θ, and σ_a^2, respectively, are not generally equal to the real value of these parameters. The model with estimated parameters

$$(Y_t - \hat{\mu}) = \hat{\phi}_1(Y_{t-1} - \hat{\mu}) + \ldots + \hat{\phi}_p(Y_{t-p} - \hat{\mu}) + a_t - \hat{\theta}_1 a_{t-1} - \ldots - \hat{\theta}_q a_{t-q} \qquad (1.2)$$

may not fit the original data well. Diagnostic checking determines whether our estimated model fits the data well. The residual process $[\hat{a}_t]$ is examined. If the $[\hat{a}_t]$ is close to a white noise process, the model is considered to be adequate and the whole model building procedure is over. Otherwise, we go to the next step.

d. Modification of the Model

If the model we built is found inadequate through the diagnostic checking, we will try to fit the data by a new modified model. Generally, the residual process $[\hat{a}_t]$ will reveal some information on how the model should be rebuilt. In most cases, a pair of new values of and q will be obtained. Using these new values of p and q, we undergo steps b., c., and d. for this new model building.

All the four steps have been carefully followed for building the gunners' error data model. For the (apparently) stationary time series, with azimuth and elevation both counted, the total number of realizations

was 248. Each time series of azimuth and elevation is run separately (Tables 1 and 2). Sixty-two percent of the stationary series can be fitted well by a third order autoregressive process [p = 3, q = 0 in model (1.1)], i.e.,

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \phi_3(Y_{t-3} - \mu) + a_t \quad . \tag{1.3}$$

A few data can not be fitted well by (1.3); they are fitted by a more complicated model. These models and their percentages of the total data are given in Table 1. Due to the biological and psychological differences among gunners, there are variations in these parameters. The means and variances of these parameters are also given in Table 1.

TABLE 1. GUNNER'S ERROR MODEL FOR AZIMUTH

General model (3rd order autoregressive process) (62.90%)

	μ	ϕ_1	ϕ_2	ϕ_3	σ_a^2
Mean	0.0393	0.4489	0.2362	0.1245	0.0128
Variance	0.0108	0.0170	0.0066	0.0087	0.0001

Special Model

		Mean	Variance
1) $\phi_4 \neq 0$	(17.74%)	0.1490	0.0050
2) $\phi_5 \neq 0$	(4.84%)	0.0962	0.0076
3) $\phi_6 \neq 0$	(5.65%)	0.0485	0.0189
4) $\phi_7 \neq 0$	(3.22%)	0.1218	0.0060
5) $\phi_8 \neq 0$	(0.81%)	0.1844	0.0000
6) $\phi_9 \neq 0$	(2.42%)	0.0196	0.0142
7) $\phi_{10} \neq 0$	(1.61%)	0.0568	0.0145
8) $\phi_{11} \neq 0$	(0.81%)	0.0970	0

TABLE 2. GUNNER'S ERROR MODEL FOR ELEVATION

General model (3rd order autoregressive process) (60.48%)

	μ	ϕ_1	ϕ_2	ϕ_3	σ_a^2
Mean	-0.0515	0.3692	0.2165	0.1448	0.0045
Variance	0.0124	0.0188	0.0051	0.0057	0.0001

Special Model

		Mean	Variance
1)	$\phi_4 \neq 0$ (15.32%)	0.1535	0.0014
2)	$\phi_5 \neq 0$ (8.87%)	0.1255	0.0088
3)	$\phi_6 \neq 0$ (7.25%)	0.1125	0.0039
4)	$\phi_7 \neq 0$ (4.84%)	0.1221	0.0087
5)	$\phi_8 \neq 0$ (0.81%)	0.1497	0
6)	$\phi_9 \neq 0$ (0.81%)	0.1294	0
7)	$\phi_{10} \neq 0$ (0.81%)	0.1178	0
8)	$\phi_{11} \neq 0$ (0.81%)	0.0573	0

A question arises whether the azimuth error and elevation error are dependent on each other during a gunner's aiming. The data show that we can consider the azimuth error and elevation error to be two independent processes. The following procedure is followed.

A general model describing the relation between two time series is a linear transfer function model. Let X_t be the time series of azimuth and Y_t be the time series of elevation. A linear transfer function model can be written as

$$(Y_t - \mu_y) = \alpha_1(Y_{t-1} - \mu_y) + \ldots + \alpha_m(Y_{t-m} - \mu_y)$$

$$+ \beta_1(X_{t-1} - \mu_x) + \ldots + \beta_n(X_{t-n} - \mu_x) + N_t \qquad (1.4)$$

where

$\mu_y = E(Y_t)$

$\mu_x = E(X_t)$

N_t = a noise process

m, n = the numbers of past values of X_t and Y_t on which the present Y_t depends.

Intutitively, model (1.4) indicates that the present azimuth value Y_t may depend on the previous values of both azimuth and elevation. This model has been used in many practical situations and gives good results (see e.g., Box and Jenkins [1]). Since we have already found a good model for Y_t in the previous model buildings, we may combine the Y model and (1.4) and have

$$a_t = \beta_1(X_{t-1} - \mu_x) + \ldots + \beta_n(X_{t-n} - \mu_x) + N_t \qquad (1.5)$$

where a_t is the noise process from the model of Y_t. Since a_t is a white noise process, the values β's can be easily estimated (Box and Jenkins [1] p. 380).

An attempt has been made to fit all the corresponding pairs of azimuth error data and elevation error data by model (1.4). Except for a few exceptions (1 percent of the total), the β values are very small (less than 0.05 for all $\beta_1, \beta_2, \ldots, \beta_{25}$). Hence, we consider that the error in elevation has no significant influence on that in azimuth. A similar model fitting by replacing X by Y and Y by X in (1.4) has also been run for all pairs of data. An independence relation is also obtained here. Hence, we conclude that there is no significant dependence between azimuth error and elevation error.

4. Simulation Procedure

In order to simulate the total performance of a guided missile system with a man in the loop, we may use the gunner's model described in the previous section. Considering the nonrepeatability of man's reactions, it must be realized that for any single simulation the error model will not give the same results as given by man. However, man's behavior on the average should agree with that of the error model.

Simulation of a gunner's behavior may be performed as follows:

a) Choose 2 random numbers γ_1 and γ_2 in [0, 1]. γ_1 is used to construct azimuth error, if

$\gamma_1 \in$ [0, 0.6290], a third order autoregressive model will be used,

$\gamma_1 \in$ [0.6291, 0.8064], a fourth order autoregressive model with $\phi_4 \neq 0$ will be used,

$\gamma_1 \in$ [0.8065, 0.8548], a fifth order autoregressive model with $\phi_5 \neq 0$ will be used,

$\gamma_1 \in$ [0.8549, 0.9113], a sixth order autoregressive model with $\phi_6 \neq 0$ will be used,

$\gamma_1 \in$ [0.9114, 0.9435], a seventh order autoregressive model with $\phi_7 \neq 0$ will be used,

$\gamma_1 \in$ [0.9436, 0.9516], an eighth order autoregressive model with $\phi_8 \neq 0$ will be used,

$\gamma_1 \in$ [0.9517, 0.9758], a ninth order autoregressive model with $\phi_9 \neq$ will be used,

$\gamma_1 \in$ [0.9759, 0.9919], a tenth order autoregressive model with $\phi_{10} \neq 0$ will be used,

$\gamma_1 \in$ [0.9920, 1.00], an eleventh order autoregressive model with $\phi_{11} \neq 0$ will be used.

Thus, we have chosen a model for azimuth error process. γ_2 is used to construct elevation error, if

$\gamma_2 \in$ [0, 0.6048], a third order autoregressive model will be used,

$\gamma_2 \in$ [0.6049, 0.7580], a fourth order autoregressive model with $\phi_4 \neq 0$ will be used,

$\gamma_2 \in$ [0.7581, 0.8467], a fifth order autoregressive model with $\phi_5 \neq 0$ will be used,

$\gamma_2 \in$ [0.8468, 0.9192], a sixth order autoregressive model with $\phi_6 \neq 0$ will be used,

$\gamma_2 \in$ [0.9193, 0.9676], a seventh order autoregressive model with $\phi_7 \neq 0$ will be used,

$\gamma_2 \in [0.9677, 0.9757]$, an eighth order autoregressive model with $\phi_8 \neq 0$ will be used,

$\gamma_2 \in [0.9758, 0.9838]$, a ninth order autoregressive model with $\phi_9 \neq 0$ will be used,

$\gamma_2 \in [0.9839, 0.9919]$, a tenth order autoregressive model with $\phi_{10} \neq 0$ will be used,

$\gamma_2 \in [0.9920, 1.0]$, an eleventh order autoregressive model with $\phi_{11} \neq 0$ will be used.

b) Use normal random number generator to generate the required parameters μ, ϕ's, and σ_a^2.

c) Using a polynomial root solver, check the roots of $X^p - \phi_1 X^{p-1} \ldots - \phi_p = 0$. If any of the roots is greater than or equal to 1, discard this set of ϕ's and select another group of parameters.

d) Let X_t denote the azimuth error process and Y_t denote the elevation error process. Then according to the models and parameters chosen by steps a) and b), we can simulate X_t and Y_t consecutively by generating normal random derivates a_t from $N(0, \sigma_a^2)$.

e) If the perfect aim of a gunner at time t is $(\underline{A}_t, \underline{E}_t)$, then our simulated coordinate of a gunner at time t is $(A_t + X_t, E_t + Y_t)$. A simulation example:

BEST AVAILABLE COPY.

INITIALLY SELECTED VALUES

	MEAN	AVAR	PHI1	PHI2	PHI3	JUMP	PHI(JUMP)
AZIMUTH	.2804	.1113	.0000	.0000	.0000		
ELEVATION	.1990	.1107	.0000	.0000	.0000		

TIME	AZIMUTH	ELEVATION	TIME	AZIMUTH	ELEVATION	TIME	AZIMUTH	ELEVATION	TIME	AZIMUTH	ELEVATION
1	.1799	.1884	90	.1713	.0000	135	.0959	.0968	180	.0437	.0000
2	.1337	.1884	91	.1563	.1171	136	.0874	.0499	181	.0933	.1051
3	.1863	.1938	92	.1414	.0000	137	.1000	.0000	182	.0000	.0000
4	.1603	.1348	93	.1263	.1419	138	.1038	.0499	183	.0833	.0948
5	.0863	.1938	94	.1114	.0000	139	.0854	.0000	184	.1168	.1668
6	.1844	.1320	95	.0963	.1761	140	.0419	.1019	185	.0395	.0487
7	.1844	.1908	96	.0815	.0000	141	.0988	.0000	186	.0814	.1774
8	.1811	.1777	97	.0663	.1916	142	.1333	.1494	187	.0744	.1419
9	.1899	.1918	98	.0514	.0000	143	.1488	.1984	188	.0396	.0000
10	.1768	.1438	99	.0364	.2060	144	.1264	.1926	189	.0046	.0716
11	.1918	.1908	100	.0214	.0000	145	.1094	.2037	190	.0906	.1496
12	.1686	.1773	101	.0061	.2179	146	.1778	.1748	191	.0921	.1919
13	.1697	.1908	102	.0093	.0000	147	.0966	.0961	192	.0816	.1108
14	.1879	.1761	103	.0244	.2411	148	.0883	.0648	193	.0633	.0882
15	.1807	.1896	104	.0393	.0000	149	.0804	.0401	194	.0681	.0889
16	.1768	.1617	105	.0547	.2643	150	.0778	.0335	195	.0418	.0831
17	.1738	.1817	106	.0698	.0000	151	.0966	.0388	196	.0446	.0844
18	.1838	.1971	107	.0848	.2877	152	.1277	.0629	197	.0992	.1312
19	.1811	.1896	108	.1000	.0000	153	.1473	.0911	198	.1009	.1484
20	.1843	.1884	109	.1151	.3111	154	.1273	.1016	199	.1309	.1868
21	.1872	.1896	110	.1302	.0000	155	.1107	.1184	200	.1408	.1908
22	.1838	.1773	111	.1451	.3345	156	.1179	.1128	201	.1828	.2357
23	.1807	.1896	112	.1599	.0000	157	.1276	.1343	202	.1808	.2404
24	.1794	.1872	113	.1747	.3342	158	.1273	.1343	203	.1788	.2723
25	.1814	.1908	114	.1896	.0000	159	.1494	.1494	204	.1808	.2708
26	.1838	.1947	115	.2038	.3078	160	.1632	.1607	205	.1807	.2808
27	.1838	.2044	116	.2188	.0000	161	.1607	.1607	206	.1808	.2678
28	.1872	.2261	117	.2337	.2849	162	.1768	.1607	207	.1808	.2678
29	.1881	.1896	118	.2486	.0000	163	.1699	.1804	208	.1808	.2687
30	.1903	.1908	119	.2635	.2832	164	.1959	.1949	209	.1804	.2687
31	.1940	.1908	120	.2784	.0000	165	.1804	.1949	210	.1808	.2687
32	.1940	.1908	121	.2933	.2778	166	.1818	.2116	211	.3018	.3219
33	.1940	.1908	122	.3082	.0000	167	.1968	.1949	212	.1808	.2687
34	.1940	.1908	123	.3231	.2778	168	.1813	.1897	213	.1808	.2687
35	.1940	.1908	124	.3380	.0000	169	.1808	.1876	214	.1808	.2687
36	.1940	.1908	125	.3529	.2778	170	.1953	.1949	215	.1808	.2687
37	.1940	.1908	126	.3678	.0000	171	.1968	.1968	216	.1808	.2687
38	.1940	.1908	127	.3827	.2778	172	.1968	.1968	217	.1808	.2687
39	.1940	.1908	128	.3976	.0000	173	.1968	.1968	218	.1808	.2687
40	.1940	.1908	129	.4125	.2778	174	.1968	.1968	219	.1808	.2687
41	.1940	.1908	130	.4274	.0000	175	.1968	.1968	220	.1808	.2687
42	.1940	.1908	131	.4423	.2778	176	.1968	.1968	221	.1808	.2687
43	.1940	.1908	132	.4572	.0000	177	.1968	.1968	222	.1808	.2687
44	.1940	.1908	133	.4721	.2778	178	.1968	.1968	223	.1808	.2687
45	.1940	.1908	134	.4870	.0000	179	.1968	.1968	224	.1808	.2687
46	.1940	.1908	135	.5019	.2778	180	.1968	.1968	225	.1808	.2687
47	.1940	.1908	136	.5168	.0000	181	.1968	.1968	226	.1808	.2687
48	.1940	.1908	137	.5317	.2778	182	.1968	.1968	227	.1808	.2687
49	.1940	.1908	138	.5466	.0000	183	.1968	.1968	228	.1808	.2687
50	.1940	.1908	139	.5615	.2778	184	.1968	.1968	229	.1808	.2687
51	.1940	.1908	140	.5764	.0000	185	.1968	.1968	230	.1808	.2687
52	.1940	.1908	141	.5913	.2778	186	.1968	.1968	231	.1808	.2687
53	.1940	.1908	142	.6062	.0000	187	.1968	.1968	232	.1808	.2687
54	.1940	.1908	143	.6211	.2778	188	.1968	.1968	233	.1808	.2687
55	.1940	.1908	144	.6360	.0000	189	.1968	.1968	234	.1808	.2687
56	.1940	.1908	145	.6509	.2778	190	.1968	.1968	235	.1808	.2687
57	.1940	.1908	146	.6658	.0000	191	.1968	.1968	236	.1808	.2687
58	.1940	.1908	147	.6807	.2778	192	.1968	.1968	237	.1808	.2687
59	.1940	.1908	148	.6956	.0000	193	.1968	.1968	238	.1808	.2687
60	.1940	.1908	149	.7105	.2778	194	.1968	.1968	239	.1808	.2687
61	.1940	.1908	150	.7254	.0000	195	.1968	.1968	240	.1808	.2687
62	.1940	.1908	151	.7403	.2778	196	.1968	.1968	241	.1808	.2687
63	.1940	.1908	152	.7552	.0000	197	.1968	.1968	242	.1808	.2687
64	.1940	.1908	153	.7701	.2778	198	.1968	.1968	243	.1808	.2687
65	.1940	.1908	154	.7850	.0000	199	.1968	.1968	244	.1808	.2687
66	.1940	.1908	155	.7999	.2778	200	.1968	.1968	245	.1808	.2687
67	.1940	.1908	156	.8148	.0000	201	.1968	.1968	246	.1808	.2687
68	.1940	.1908	157	.8297	.2778	202	.1968	.1968	247	.1808	.2687
69	.1940	.1908	158	.8446	.0000	203	.1968	.1968	248	.1808	.2687
70	.1940	.1908	159	.8595	.2778	204	.1968	.1968	249	.1808	.2687
71	.1940	.1908	160	.8744	.0000	205	.1968	.1968	250	.1808	.2687
72	.1940	.1908	161	.8893	.2778	206	.1968	.1968	251	.1808	.2687
73	.1940	.1908	162	.9042	.0000	207	.1968	.1968	252	.1808	.2687
74	.1940	.1908	163	.9191	.2778	208	.1968	.1968	253	.1808	.2687
75	.1940	.1908	164	.9340	.0000	209	.1968	.1968	254	.1808	.2687
76	.1940	.1908	165	.9489	.2778	210	.1968	.1968	255	.1808	.2687
77	.1940	.1908	166	.9638	.0000	211	.1968	.1968	256	.1808	.2687
78	.1940	.1908	167	.9787	.2778	212	.1968	.1968	257	.1808	.2687
79	.1940	.1908	168	.9936	.0000	213	.1968	.1968	258	.1808	.2687
80	.1940	.1908	169	.0085	.2778	214	.1968	.1968	259	.1808	.2687
81	.1940	.1908	170	.0234	.0000	215	.1968	.1968	260	.1808	.2687
82	.1940	.1908	171	.0383	.2778	216	.1968	.1968	261	.1808	.2687
83	.1940	.1908	172	.0532	.0000	217	.1968	.1968	262	.1808	.2687
84	.1940	.1908	173	.0681	.2778	218	.1968	.1968	263	.1808	.2687
85	.1940	.1908	174	.0830	.0000	219	.1968	.1968	264	.1808	.2687
86	.1940	.1908	175	.0979	.2778	220	.1968	.1968	265	.1808	.2687
87	.1940	.1908	176	.1128	.0000	221	.1968	.1968	266	.1808	.2687
88	.1940	.1908	177	.1277	.2778	222	.1968	.1968	267	.1808	.2687
89	.1940	.1908	178	.1426	.0000	223	.1968	.1968	268	.1808	.2687
90	.1940	.1908	179	.1575	.2778	224	.1968	.1968	269	.1808	.2687
91	.1940	.1908	180	.1724	.0000	225	.1968	.1968	270	.1808	.2687
92	.1940	.1908	181	.1873	.2778	226	.1968	.1968	271	.1808	.2687
93	.1940	.1908	182	.2022	.0000	227	.1968	.1968	272	.1808	.2687
94	.1940	.1908	183	.2171	.2778	228	.1968	.1968	273	.1808	.2687
95	.1940	.1908	184	.2320	.0000	229	.1968	.1968	274	.1808	.2687
96	.1940	.1908	185	.2469	.2778	230	.1968	.1968	275	.1808	.2687
97	.1940	.1908	186	.2618	.0000	231	.1968	.1968	276	.1808	.2687
98	.1940	.1908	187	.2767	.2778	232	.1968	.1968	277	.1808	.2687
99	.1940	.1908	188	.2916	.0000	233	.1968	.1968	278	.1808	.2687
100	.1940	.1908	189	.3065	.2778	234	.1968	.1968	279	.1808	.2687

5. Data Listing

BEST AVAILABLE COPY

ID	C	OO	END	N	GAP	MEAN	PHI(1)	PHI(2)	PHI(3)	PHI(4)	J	PHI(J)
71	A	200	1871	1672	.0119	.0144	.4843	.2255	.0507	.1050		
71	F	200	1871	1672	.0137	-.0475	.7749	.1886	.1747			
72	A	200	1944	1447	.0140	.0609	.7718	.2084	.1484	.1940		
72	F	200	1944	1447	.0112	-.0495	.7445	.2078	.1087			
73	A	200	1900	1701	.0040	.4498	.5830	.2446	.1238			
73	F	200	1900	1701	.0152	.0033	.2498	.7010	.1491	.1792		
74	A	200	1840	1641	.0078	.7835	.4834	.7704	.1410			
74	F	200	1840	1641	.0134	-.0572	.7250	.2309	.1009			
75	A	200	1949	1770	.0148	.1717	.4411	.2000	.0904	.7001		
75	F	200	1949	1770	.0044	.0096	.7507	.2438	.1067	.1418		
76	A	100	700	601	.0152	.2484	.4608	.2501	.0574			
76	F	100	700	601	.0084	-.0707	.7432	.2474	.1018		6	.1097
77	A	1	1000	1000	.0008	.0770	.6091	.1700	.2710			
77	F	1	1000	1000	.0020	-.1049	.1017	.1749	.1405	.1947		
78		NONSTATIONARY ERROR										
80	A	1	800	800	.0058	.0994	.1702	.1944	.1148			
80	F	1	800	800	.0030	-.0702	.2284	.1347	.1401	.1707		
82		NONSTATIONARY ERROR										
83	A	50	400	551	.0170	-.0010	.6704	.1411	.0804	.1414		
83	F	50	400	551	.0182	-.1807	.6010	.2172	.0898			
84	A	300	700	401	.0201	.7117	.5844	.1014	.0078			
84	F	300	700	401	.0201	.0740	.4187	.2114	.0918			
84	A	1	800	800	.0118	.0618	.2842	.2704	.2078		7	.1574
84	F	1	800	800	.0258	-.1498	.2001	.2507	.2067			
87	A	100	940	870	.0207	.1407	.5742	.7007	.0404			
87	F	100	940	870	.0461	-.2400	.2477	.2918	.2701			
88	A	100	450	551	.0104	.0275	.5144	.1816	.1847			
88	F	100	450	551	.0222	-.1410	.7048	.1108	.1521			
90	A	140	420	281	.0072	.0641	.5248	.2078	.1811		7	.0072
90	F	140	420	281	.0058	-.1017	.4402	.0700	.2242		9	.0484
91	A	50	410	370	.0787	-.0144	.5027	.7148	.0584			
91	F	50	410	370	.0781	-.1760	.7097	.2194	.1000			
92		NONSTATIONARY ERROR										
97	A	50	475	526	.0074	.0494	.4477	.7077	-.0782			
97	F	50	475	526	.0104	-.0817	.2748	.2114	.1013			
98	A	100	442	742	.0184	-.0038	.7707	.1408	.1114		9	.1404
98	F	100	442	742	.0089	-.0444	.7164	-.0287	.2530			
99	A	100	400	301	.0047	.0477	.6700	.7107	.1479			
99	F	100	400	301	.0284	-.1107	.0411	.1491	.1749		8	.1471
94	A	50	478	428	.0078	.0710	.7477	.1414	-.0484			
94	F	50	478	428	.0000	-.0271	.6178	.2714	.0758			
98	A	100	481	982	.0077	-.0097	.4408	.1474	.0870			
98	F	100	481	982	.0140	.0800	.7740	.1887	.1074	.1408		
98	A	50	444	747	.0174	-.0998	.2480	.1778	.7458			
98	F	50	444	747	.0066	-.0770	.7808	.2177	.2967			
100	A	40	780	741	.0098	.0141	.7187	.1440	.0142			
100	F	40	780	741	.0170	-.0727	.2740	.7981	.1481			
101	A	50	280	231	.0202	-.1701	.2144	.7872	.2270			
101	F	50	280	231	.0040	.0442	.4545	.2044	-.0284			
102	A	50	500	451	.0244	-.1771	.4110	.1744	-.0178	.2140		
102	F	50	577	401	.0074	.0178	.4478	.0748	.2777			
104	A	100	957	458	.0047	.0048	.4178	.7874	.0700			
104	F	100	957	458	.0241	-.1957	.1087	.1747	.0714	.1447		
105	A	100	1000	901	.0044	-.7411	.4528	.7484	.0440			
105	F	100	1000	901	.0047	.0000	.4410	.7841	.0789			

BEST AVAILABLE COPY

104	A	100	500	401	.0050	-.0160	.3821	.2955	.1596		5	.0604
105	F	100	500	401	.0352	-.2446	.0917	.2914	.1532		6	.1982
107	A	100	540	540	.0[illegible]50	-.1320	.3541	.1730	.2001			
107	F	100	540	540	.0[illegible]4	-.0874	.3435	.2648	.1481		6	.0889
108	A	50	547	498	.0013	.0715	.4978	.3109	.1838		9	-.1140
108	F	50	547	498	.0284	-.2015	.2893	.1421	.2104			
109			NONSTATIONARY ERRORS									
110	A	100	400	301	.0121	.0040	.3860	.1844	.2899			
110	F	100	400	301	.0148	-.1021	.1354	.1738	.0144		5	.2499
111	A	100	1300	1201	.0350	-.0277	.2706	.1624	.1902	.1479	5	.1407
111	F	100	1300	1201	.0276	-.1158	.1800	.2494	.2173			
112	A	100	1500	1401	.0117	-.0216	.3428	.2952	.0883	.1778		
112	F	100	1500	1401	.0142	-.0647	.2767	.1841	.0994	.1803		
113	A	100	1865	1766	.0059	.1911	.4102	.2917	.2474			
113	F	100	1865	1766	.0158	-.0590	.2720	.2171	.2946		7	.1522
114			NONSTATIONARY ERRORS									
115	A	100	1801	1702	.0069	.0497	.4333	.2710	.2077			
115	F	100	1801	1702	.0176	-.1812	.2378	.1540	.1471		5	.1657
118	A	200	1400	1201	.0217	-.0544	.4501	.1877	.0830		6	.1722
118	F	200	1400	1201	.0134	-.1277	.4769	.2734	.1630			
119			NONSTATIONARY ERRORS									
120	A	100	1842	1743	.0693	-.1709	.2288	.2018	.2572	.2196		
120	F	100	1842	1743	.0047	-.0300	.4493	.3072	.1324			
121	A	200	1895	1696	.0214	-.0311	.5367	.2134	.1149			
121	F	200	1895	1696	.0131	-.0900	.3927	.2074	.0991			
122	A	100	1900	1801	.0369	-.1522	.4019	.2378	.1770			
122	F	100	1900	1801	.0094	-.0312	.4026	.1900	.1988			
123	A	200	1887	1688	.0219	-.0334	.3915	.2248	.1776	.1341		
123	F	200	1887	1688	.0132	-.0708	.3842	.2544	.2030			
124	A	200	1400	1401	.0245	-.1127	.4491	.1291	.1740		10	.1772
124	F	200	1400	1401	.0132	-.0114	.4844	.3035	.0708			
125	A	1	1300	1300	.0025	.1230	.4558	.2471	.0760			
126	F	1	1300	1300	.0169	-.1089	.2814	.1412	.1816			
127	A	1000	1800	801	.0421	-.1025	.2016	.2487	.2784	.1782		
127	F	1000	1800	801	.0105	-.1500	.3954	.2123	.1483	.0598		
128	A	1000	1582	583	.0092	.0356	.5331	.2604	.0934			
128	F	1000	1582	583	.0129	-.0579	.4845	.2104	.1408			
129	A	200	1400	1201	.0519	-.2150	.3650	.2150	.2142			
129	F	200	1400	1201	.0104	-.0479	.3428	.2448	.1580			
130	A	200	1794	1595	.0106	.1741	.4884	.2714	.2202			
130	F	200	1794	1595	.0044	.0140	.4880	.2404	.1794			
132	A	200	1400	1201	.0217	-.0830	.3873	.3047	.1221			
132	F	200	1400	1201	.0134	-.0402	.5328	.1906	.1414			
133	A	200	1785	1586	.0271	-.1945	.3814	.2477	.1885		11	.0970
133	F	200	1785	1586	.0167	-.0701	.3827	.3045	.1737			
134			TOO FEW DATA									
135			TOO FEW DATA									
134	A	700	1000	301	.0170	-.0444	.4147	.2400	.1433			
134	F	700	1000	301	.0068	.0041	.7814	.1808	-.1004			
138	A	1	600	600	.0067	.0068	.2708	.3788	.1934	.1079		
138	F	1	600	600	.0078	-.0814	.3373	.3601	.1412			
139	A	1	600	600	.0013	.0558	.4478	.2608	.0025			
139	F	1	600	600	.0037	-.0429	.3054	.2348	.1887		8	.1497
140			NONSTATIONARY ERRORS									
141	A	50	450	401	.0064	-.0227	.5580	.2370	.0872			
141	F	50	450	401	.0070	-.0455	.5764	.2072	.1042			
142	A	50	600	551	.0078	-.0214	.4943	.2384	.0457			
142	F	50	600	551	.0061	-.0816	.4462	.2150	.1841			
144	A	20	700	681	.0074	.0261	.4710	.2212	.1200			
144	F	20	700	681	.0080	.0070	.4274	.2442	.2341			
145	A	50	500	451	.0036	.1239	.5018	.3745	-.0077			
145	F	50	500	451	.0044	.0015	.4275	.2415	.0011			
147	A	80	400	321	.0080	.0525	.3444	.0274	.0100			

BEST AVAILABLE COPY

147	F	80	400	321	.0062	-.0027	.2448	.1401	.1827		
148	A	100	400	501	.0066	-.0011	.5002	.2178	.1479		
148	F	100	400	501	.0109	-.0503	.2785	.2402	.1961	.1767	
149			TOO FEW DATA								
170	A	1	570	570	.0058	.0896	.5446	.7217	.0556		
170	F	1	570	570	.0183	-.1240	.3994	.1918	.0974		
171	A	1	1898	1898	.0075	.0724	.4225	.2449	.0520		
171	F	1	1898	1898	.0040	.0854	.5918	.2418	.1567		5 -.1359
172			NONSTATIONARY ERRORS								
173	A	40	390	281	.0094	-.0926	.4094	.3748	.0724		6 -.0501
173	F	40	390	281	.0091	-.0857	.5412	.1802	.1704		
174	A	1	600	600	.0038	.0202	.4119	.2442	.2328		
174	F	1	600	600	.0055	-.0002	.4752	.2882	-.0027	.1531	
175	A	100	500	401	.0213	-.1836	.2822	.2247	.1780		8 .1844
175	F	100	500	401	.0039	.0870	.8778	.1777	-.0891		
176	A	200	500	301	.0028	.2021	.7783	.2541	.0404		10 -.0437
176	F	200	500	301	.0094	-.1054	.3540	.2204	.2185		10 .1178
177	A	1	599	599	.0040	.3372	.4751	.2048	.0784		
177	F	1	599	599	.0103	-.0112	.3458	.2070	.1444		7 .2212
178	A	1	1100	1100	.0096	.0037	.3930	.7117	.2187		
178	F	1	1100	1100	.0144	-.0087	.3044	.2574	.2502		
179	A	150	450	301	.0028	.1744	.6491	.1243	.1330		
179	F	150	450	301	.0043	-.0040	.3331	.2782	.1441		7 .1448
180	A	50	250	201	.0018	.1214	.7548	.1150	.0508		
180	F	50	250	201	.0050	-.0504	.4701	.1844	.1287	.0541	
183	A	1	800	800	.0041	.0529	.4148	.3013	.2038		
183	F	1	800	800	.0044	-.0026	.4478	.2828	.1453		
187	A	1	500	500	.0042	-.0280	.3210	.1871	.1928		
187	F	1	500	500	.0052	.0555	.4090	.1876	.2244		
184	A	1	872	872	.0087	-.0019	.3832	.2414	.2418		
184	F	1	872	872	.0082	-.0233	.3590	.2145	.1441		8 .1444
185	A	1	600	600	.0100	-.0474	.6081	.2588	.1964		
185	F	1	600	600	.0049	.0087	.4598	.2370	.1978		
186	A	100	800	701	.0050	-.0059	.3448	.2877	.1498		
186	F	100	800	701	.0089	-.0521	.4041	.2707	.1482		
187	A	20	198	180	.0047	-.0101	.4441	.0980	.2277		
187	F	20	198	180	.0090	-.0148	.3018	.2593	.1934		
188	A	1	898	898	.0044	-.0049	.3811	.4139	.1418		
188	F	1	898	898	.0058	.0037	.3320	.2208	.1894		6 .1799
189	A	1	800	800	.0020	.0240	.4445	.3001	.1594		
189	F	1	800	800	.0029	.0048	.3021	.2407	.1548	.1644	
190	A	100	450	351	.0043	.1041	.4401	.2741	.0894	.1368	
190	F	100	450	351	.0048	.0552	.4257	.2704	.2091		
191	A	150	450	301	.0050	-.0087	.3077	.2441	.2877		
191	F	150	450	301	.0054	-.0816	.2767	.2411	.1778		
192			TOO FEW DATA								
197	A	80	701	652	.0127	-.1395	.4148	.2855	.1194		
197	F	80	701	652	.0078	-.0409	.4067	.2037	.1470		9 .1150
194	A	200	500	301	.0097	-.0274	.4494	.3024	-.0714		
194	F	200	500	301	.0058	-.0425	.4848	.1829	.2072		
197			NONSTATIONARY ERRORS								
198	A	1	1000	1000	.0041	.0441	.5476	.2508	.2351		4 -.1958
198	F	1	1000	1000	.0193	-.0444	.3442	.2204	.1450		
199	A	320	400	281	.0048	.0801	.4040	.2277	.0553		
199	F	320	400	281	.0045	.0514	.4907	.2285	.1748		
200	A	40	240	201	.0043	.1814	.5789	.2242	.1186		
200	F	40	240	201	.0433	-.1407	.3994	.2530	.0987	.1428	
202	A	100	516	417	.0120	.1440	.5747	.1840	.0470	.1144	
202	F	100	516	417	.0040	-.0584	.3994	.1982	.2079		7 -.0582
203	A	100	1100	1001	.0093	-.0238	.2472	.1955	.1284	.2052	

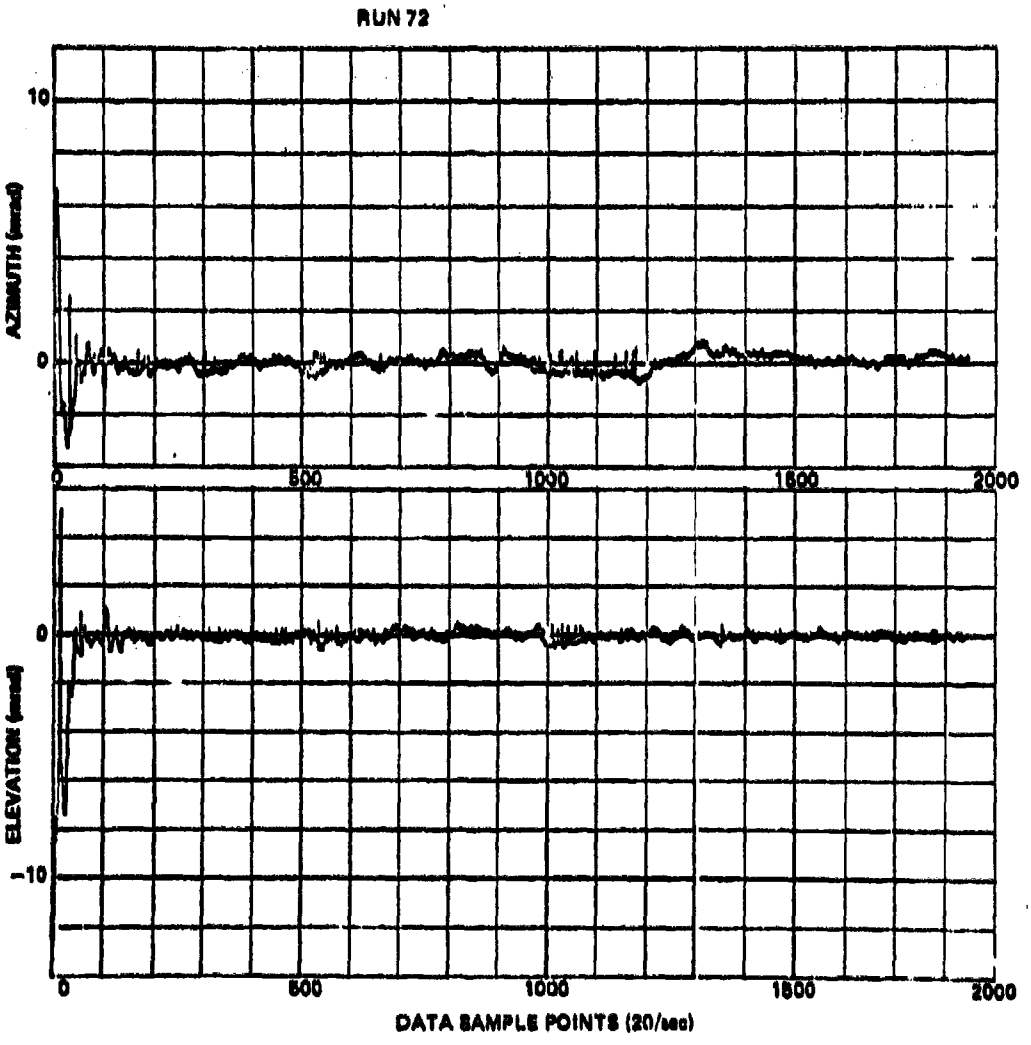
BEST AVAILABLE COPY

203	F	100	1100	1001	.0043	.0451	.4234	.1770	.1757		5	.1749
204	A	40	320	281	.0088	.2267	.7140	.1837	-.0267			
204	F	40	320	281	.0098	.1232	.8419	.0751	-.1867		9	.1904
205	A	60	360	301	.0038	.3490	.6280	.1415	.1349			
205	F	60	360	301	.0069	-.0442	.6280	.1849	.1521			
206			NONSTATIONARY ERRORS									
207	A	80	377	298	.0208	-.0324	.7117	.1995	.0940	.1707		
207	F	80	377	298	.0130	-.0493	.5976	.0189	.1347			
211			NONSTATIONARY ERRORS									
214			NONSTATIONARY ERRORS									
215	A	100	500	401	.0035	.0384	.5450	.2528	.1008		6	-.0538
215	F	100	500	401	.0122	-.0790	.2962	.2748	.1362			
217	A	50	500	451	.0309	.0854	.5270	.2847	.1073			
217	F	50	500	451	.0870	-.0727	.2136	.2414	.1211	.1944		
218	A	50	503	454	.0239	.0331	.3507	.1844	.3449	-.0940		
218	F	50	503	454	.0401	-.1041	.2235	.2509	.0482		5	.1242
219	A	100	604	505	.0025	.0171	.4240	.3383	.0871			
219	F	100	604	505	.0124	-.0736	.2455	.2244	.1811		7	.1944
220			TOO FEW DATA									
221	A	1	750	750	.0049	-.0178	.1926	.1114	.2641	.1347	5	.1657
221	F	1	750	750	.0061	-.0918	.2299	.3270	.2160			
222			TOO FEW DATA									
227	A	100	777	678	.0068	.0732	.4116	.2217	.0440			
227	F	100	777	678	.0080	-.1048	.2128	.2529	.1171	.1490		
227			TOO FEW DATA									
230	A	50	400	351	.0062	-.0126	.3807	.2251	.1437			
230	F	50	400	351	.0080	-.0667	.2035	.1904	.0410			
231			NONSTATIONARY ERRORS									
235	A	80	400	321	.0060	-.0176	.3314	.1237	.1483			
235	F	80	400	321	.0078	-.0267	.4128	.1800	.3000			
236	A	100	550	451	.0078	.0511	.4520	.1307	.1537			
236	F	100	550	451	.0078	-.0354	.3476	.1773	.2981			
3	A	60	300	241	.0033	-.0405	.6078	.0985	-.0434			
3	F	60	300	241	.0029	-.0953	.7443	.1345	.0487			
4			TOO FEW DATA									
155			NONSTATIONARY ERRORS									
156			NONSTATIONARY ERRORS									
158	A	60	360	301	.0175	.0070	.4184	.1585	-.0910			
158	F	60	360	301	.0154	-.1151	.4147	.1200	.0478			
159	A	100	1400	1301	.0128	-.0328	.5253	.1515	.0545	.1141		
159	F	100	1400	1301	.0078	-.0350	.4031	.2417	.0732			
160	A	100	1197	1098	.0088	.0946	.4450	.2358	-.0105			
160	F	100	1197	1098	.0101	-.0486	.4276	.2013	.1808			
161	A	50	500	451	.0173	.0192	.4155	.1339	.1244			
161	F	50	500	451	.0130	-.0387	.3884	.1832	.1884			
273	A	100	450	351	.0014	.0120	.7734	-.0775	.0787		5	-.0827
273	F	100	450	351	.0025	-.0491	.0717	.1309	.1040			
274	A	1	501	501	.0057	.0837	.4347	.2254	.0788	.1240		
274	F	1	501	501	.0049	.0520	.3490	.1481	.2144		5	.1500
275	A	1	424	424	.0030	.0714	.4330	.2571	.1980			
275	F	1	424	424	.0038	-.0254	.4414	.2543	.1478			
276	A	100	500	401	.0087	.0243	.4056	.2024	.2070			
276	F	100	500	401	.0134	-.1145	.1148	.2010	.1248	.2017		
278	A	100	500	401	.0074	.0122	.2981	.3321	.1420			
278	F	100	500	401	.0048	.0691	.2073	.3808	.2150			
279	A	150	524	377	.0041	.0435	.5474	.1207	-.1848	.2088		
279	F	150	524	377	.0061	-.0214	.4449	.2482	.1486			
280			TOO FEW DATA									
294	A	60	475	416	.0210	.1433	.2292	.1487	.1447		4	.2310
294	F	60	475	416	.0330	-.1198	.2493	.1748	.0954			

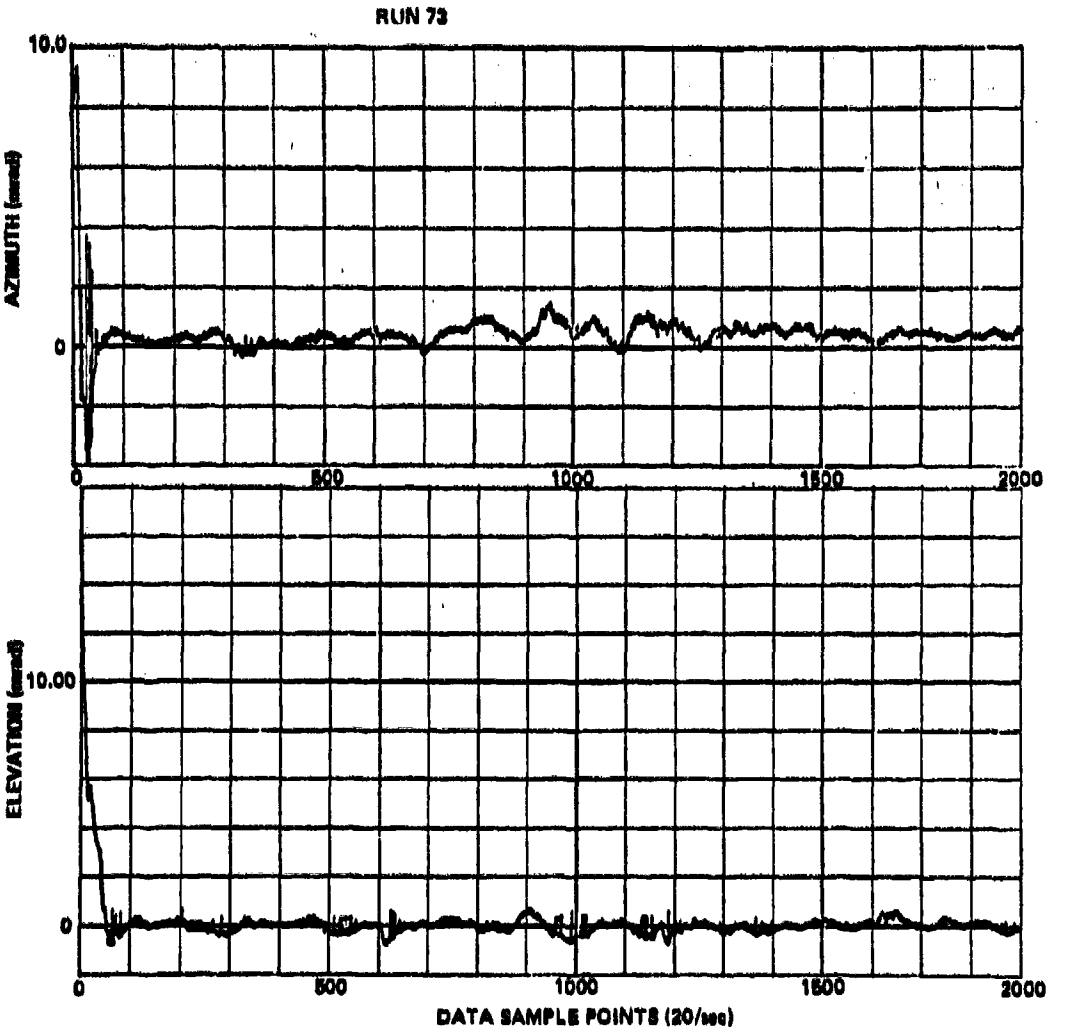
BEST AVAILABLE COPY

294	A	70	550	441	.0081	.1784	.4127	-.2478	.1094	.1209		
295	F	70	550	481	.0156	-.1140	.2491	-.0252	.1004	.1392		
296		NONSTATIONARY ERRORS										
297		TOO FEW DATA										
298	A	50	400	351	.0097	.0462	.1275	.1271	.1831		7	.2177
299	F	50	400	351	.0042	-.0090	.1776	.3094	.0113		7	.0787
300		NONSTATIONARY ERRORS										
301	A	100	424	327	.0095	.0415	.1084	.2234	.0752		9	.1755
301	F	100	424	327	.0075	-.0187	.3315	.1330	.1028		6	.1420
302	A	100	424	327	.0008	.0784	.2405	.2744	.2581			
302	F	100	424	327	.0008	.0457	.1504	.2488	.0878		6	.1021
303	A	1	450	450	.0079	.2411	.1008	.2531	.2379			
303	F	1	450	450	.0094	-.0578	.1478	.2444	.0861	.1792		
304		NONSTATIONARY ERRORS										
304		NONSTATIONARY ERRORS										
307	A	100	474	377	.0044	.1582	.1902	.1418	.2037		5	.1325
307	F	100	474	377	.0277	-.1740	.0900	.0809	.1051		5	.1182
308		TOO FEW DATA										
309	A	1	501	501	.0142	.1436	.2404	.1772	.1308		6	.1091
310	F	1	501	501	.0044	.0175	.4156	.1507	.1871		6	-.0317
311	A	100	400	501	.0058	.1495	.2518	.2818	.2417			
311	F	100	400	501	.0056	-.0444	.3140	.1270	.1808			
312	A	100	1050	940	.0010	.0387	.1731	.4868	.0037	.0277		
312	F	100	1050	940	.0034	-.0025	.5415	.2178	.0940		11	.0873
313	A	50	1200	1151	.0027	.0414	.4023	.4314	.0958		9	-.0024
313	F	50	1200	1151	.0095	-.0714	.3128	.1286	.1214		5	.1464
314	A	50	1144	1117	.0031	.0454	.5050	.3634	.0300			
314	F	50	1144	1117	.0043	-.0122	.6353	.2750	.1384			
315	A	100	1453	1536	.0032	.0746	.5541	.4114	-.0447			
315	F	100	1453	1536	.0074	-.0411	.3794	.2809	.1491			
316	A	200	1400	1201	.0044	.0035	.2894	.2784	.2144	.1485		
316	F	200	1400	1201	.0095	-.0897	.3209	.1710	.1300	.1795		
317	A	100	1500	1401	.0057	-.0238	.4104	.2352	.1497		7	.1044
317	F	100	1500	1401	.0118	-.0950	.3550	.2604	.2404			
318	A	150	700	551	.0037	.0084	.4045	.3131	.0978			
318	F	150	700	551	.0054	-.0443	.3288	.2745	.0808			
319	A	80	498	619	.0057	.0448	.4484	.2378	.1035		6	.1172
319	F	80	498	619	.0104	-.0484	.3130	.2574	.2078		6	.0851
320	A	50	1021	972	.0041	.0430	.3949	.2223	.1447	.1233		
320	F	50	1021	972	.0068	-.0440	.3730	.2447	.1997			

6. Sample of Data Plots



Sample of Data Plots



7. Future Use

Although the original work dealt mostly with a stationary set of data, there is no reason why this technique could not be used with non-stationary data simply by using the difference equations as outlined in Time Series Analysis, Forecasting, and Control, by Box and Jenkins.

A study is underway to evaluate the ROLAND Air Defense System during an ECM environment. This is a very critical area and one that so far has not been investigated with a systematic quantitative approach. The approach offered here would be valid regardless of the type of engagement (optical or radar). Simply stated: A series of target tracks are carried out and a time series model is built of the resulting radar errors as a function of ECM and other parameters.

The final output of this study would be a computer program (or subroutine integrated with the weapon system simulation) that could be used for predicting end game results as a function of different types of ECM throughout the ROLAND system engagement boundary.

The basic data needed to build the proposed model comes from a video camera bore-sighted to the track radar. An investigation of the advantages of putting a missile beacon on the target is being conducted at this time.

REFERENCES

1. Box, G. E. P. and Jenkins, G. M., Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco, 1970.

2. Fuller, F. C., Jr. and Tsokos, C. P., "Time Series Analysis of Water Pollution Data," Biometrics, 1971, pp. 1017-1034.

3. Cleveland, W. S., "Fitting Time Series Models for Prediction," Technometrics, 1971, pp. 713-723.

4. Cleveland, W. S., "The Inverse Autocorrelations of a Time Series and Their Applications," (with discussion by Parzen), Technometrics, 1972, pp. 277-298.

5. Box, G. E. P., Jenkins, G. M., and Bacon, D. W., "Models for Forecasting, Seasonal and Nonseasonal Time Series," Spectral Analysis of Time Series, ed by B. Harris, Wiley, N. Y., 1967.

6. Clevenson, M. L., "Asymptotically Efficient Estimates of the Parameters of a Moving Average Time Series," Stanford University Tech. Rep. #15, 1970.

7. Parzen, E., "Efficient Estimation of Stationary Time Series Mixed Schemes," SUNY/Buffalo Tech. Rep., 1972.

A METHOD FOR DETERMINING PAIRWISE CONTRASTS FROM A FRIEDMAN TWO-WAY LAYOUT BASED ON A THEOREM BY MARASCUILO

Jimmie C. DeLoach and Eugene F. Dutoit
United States Army Infantry Center
Fort Benning, Georgia 31905

1. INTRODUCTION.

The authors wish to express their appreciation to the US Army Research Office and the Clinical panelists at the Twenty-second conference in the Design of Experiments for their valuable comments about this problem.

In recent years there has been an increased effort to produce more and more non-parametric statistical tests. These tests have had broad based applications in education and psychological research and to some extent in military testing and evaluation of new products and training methods.

The value of such non-parametric tests is well known. Although it is not the purpose of this paper to demonstrate the usefulness of these tests, it is worthwhile to restate one of the more salient features of non-parametric tests and that is the fact that they do not depend upon sometimes unrealistic distribution assumptions, such as the normality of error distribution and that in many cases they are more readily comprehended and their test statistics more easily computed by a broader spectrum of statisticians and researchers.

Friedmen in 1937 introduced a test which is sometimes referred to as the two-way analysis of variance by ranks. The method is outlined in detail in Conover [ref 1, pp 264-274]; the test is considered to be the non-parametric version of the familiar parametric two-way analysis of variance (ANOVA). The parametric ANOVA is the usual way of testing the hypothesis of no treatment differences. For experiments of the randomized block design, and where there is one observation per block, the Friedman test is used as a non-parametric method to test this same hypothesis.

The subject of this paper is related to an extension of the Friedman test to the case of several observations per block, given in Conover [ref 1, p 273]. The example given in the next section will illustrate the use of this extension. The data come from unpublished lecture notes of reference 4.

2. EXAMPLE.

The hypothetical data of Table [1] represent scores on a reading test given to seventh grade students following one, three, or five weekly 20 minute training periods on an electric talking typewriter programed to teach reading skills. The study was conducted across four different schools, drawing from different social strata in the community and taught by four different sets of teachers in four different classroom environments.

Table [1] Scores on a Reading Test Following One, Three, or Five Weekly 20 Minute Training Periods on an Electric Talking Typewriter for Four Different Schools.

	Sessions per Week		
School	1	3	5
A	110	82	118
	87	84	96
	79	74	104
	102	70	126
B	41	93	111
	76	76	76
	43	91	91
	74	40	105
C	56	102	83
	50	40	72
	64	39	60
	61	62	105
D	67	68	126
	60	87	101
	50	69	126
	80	65	103

The data of Table 1 are ranked within each block. These rankings appears in Table 2. The sum of ranks R_j are also given.

Table 2 Observations Ranked Within Blocks and the Sum of Ranks.

	Sessions per Week		
School	1	3	5
A	10	4	11
	6	5	7
	3	2	9
	8	1	12
B	2	10	12
	6	6	6
	3	8.5	8.5
	4	1	11

Table continued on following page

School	Sessions per Week 1	3	5
C	4	11	10
	3	2	9
	8	1	5
	6	7	12
D	4	5	11
	2	8	9
	1	6	12
	7	3	10
	$R_1 = 77$	$R_2 = 80.5$	$R_3 = 154.5$

The expected value of R_j is given by:

$$E(R_j) = \frac{bm(mk+1)}{2} = \frac{(4)(4)[(4)(3)+1]}{2} \qquad (1)$$

$$= \frac{16(13)}{2}$$

$$= 104$$

where b = # blocks (schools)
k = # treatments (sessions per week)
m = # observations per cell

The Friedman test statistic is given by

$$T_4 = \frac{12}{bkm^2(mk+1)} \sum_{j=1}^{k} [R_j - E(R_j)]^2 \qquad (2)$$

$$= \frac{12}{(4)(3)(16)(13)} \sum_{j=1}^{3} [R_j - 104]^2$$

$$= \frac{12}{(12)(208)} [27^2 + 23.5^2 + 50.5^2]$$

$$= \frac{1}{208} [729 + 552.25 + 2550.25]$$

$$= \frac{3831.5}{208}$$

$$\approx 18.4$$

The distribution of T_4 can be approximated by the chi-square distribution with k-1 degrees of freedom. For this example k-1=2 and $\chi^2_{.95}(2) = 5.99$. Thus, we would reject a null hypothesis of no treatment differences.

3. PROBLEM.

In the preceding section, the results of the extension of the Friedman test to the case of several observations indicate that significant differences between the three treatments exist at the $\alpha = .05$ level. A natural question arises, i.e., which treatments differ significantly in a statistical sense? No post-hoc pairwise comparison procedures are given in Conover for this extension. Also, Hollander and Wolfe (1973) do not address this problem. A possible solution lies in extending a theorem given by Marascuilo and McSweeney (1967) which is given in the next section.

4. THEOREM (MARASCUILO - McSWEENEY). Let $\psi = a_1\theta_1 + a_2\theta_2 + \ldots + a_k\theta_k$, where $\sum_{i=1}^{k} a_i = 0$ is a linear contrast of unknown parameters. Consider the set of all possible linear contrasts of the form ψ. Let

$$\hat{\psi} = a_1\hat{\theta}_1 + a_2\hat{\theta}_2 + \ldots a_k\hat{\theta}_k \tag{3}$$

be an estimate of ψ with estimated variance given by

$$\text{Var}(\hat{\psi}) = \sum_{i=1}^{k} a_i^2 \,\text{Var}(\hat{\theta}_k) + 2\sum_{i<i'}^{k}\sum^{k} a_i a_{i'} \,\text{Cov}(\hat{\theta}_i, \hat{\theta}_{i'}) \tag{4}$$

Then in the limit the probability is $1-\alpha$ that simultaneously for all linear contrasts of the form ψ

$$\hat{\psi} - \sqrt{\chi^2_{1-\alpha}(k-1)}\ \sqrt{\text{Var}(\hat{\psi})} < \psi < \hat{\psi} + \sqrt{\chi^2_{1-\alpha}(k-1)}\ \sqrt{\text{Var}(\hat{\psi})}$$

The reader will note that this theorem is a chi-square analog to the more familiar Scheffe' theorem.

The proof of this theorem may be obtained from Marascuilo and McSweeney (reference 3) upon request.

5. APPLICATION OF THE THEOREM. Let R_i be the sum of the ranks as in section 2. Let

$$\psi = a_1\theta_1 + a_2\theta_2 + \ldots + a_k\theta_k \tag{5}$$

be a linear contrast with estimate

$$\hat{\psi} = a_1 R_1 + a_2 R_2 + \ldots + a_k R_k \qquad (6)$$

The variance of the contrast will be determined two ways; assuming independence between treatment observations [i.e., Cov $(\hat{\theta}_i, \hat{\theta}_{i'}) = 0$] and the case where the assumption of independence cannot be justified [i.e., Cov $(\hat{\theta}_i, \hat{\theta}_{i'}) \neq 0$].

a. If Cov $(\hat{\theta}_i, \hat{\theta}_{i'}) = 0$

$$\text{Var}(\hat{\psi}) = a_1^2 \text{ Var}(R_1) + a_2^2 \text{ Var}(R_2) + \ldots a_k^2 \text{ Var}(R_k)$$

$$= \left[\frac{bm^2 (m\,k+1)(k-1)}{12}\right] \Sigma a_i^2$$

Where Var (R_i) is given in Conover (p. 273).

$$\sqrt{\text{Var}(\hat{\psi})} = \sqrt{\left[\frac{bm^2 (m\,k+1)(k-1)}{12}\right] \Sigma a_i^2} \qquad (7)$$

b. If Cov $(\hat{\theta}_i, \hat{\theta}_{i'}) \neq 0$

$$\sqrt{\text{Var}(\hat{\psi})} = \sqrt{\left[\frac{bm(m\,k+1)(mk-m+1)}{12}\right] \Sigma a_i^2} \qquad (8)$$

Now

$$\hat{\psi}_1 = R_1 - R_2 = 77 - 80.5 = -3.5$$

$$\hat{\psi}_2 = R_1 - R_3 = 77 - 154.5 = -77.5$$

$$\hat{\psi}_3 = R_2 - R_3 = 80.5 - 154.5 = -74$$

are the possible pairwise comparisons and their estimated values from our original example. In order to test these values for significance, we apply the Marascuilo - McSweeney theorem and compute the critical differences.

$$CD = \sqrt{\chi^2_{1-\alpha(k-1)}} \quad \sqrt{Var\ (\hat{\psi})} \qquad (9)$$

a. If Cov $(\hat{\theta}_i, \hat{\theta}_{i'}) = 0$

$$CD = \sqrt{5.99} \qquad \sqrt{277.33}$$

$$= (2.45) \qquad (16.65)$$

$$= 40.79$$

b. If Cov $(\hat{\theta}_i, \hat{\theta}_{i'}) \neq 0$

$$CD = \sqrt{5.99} \qquad \sqrt{312}$$

$$= (2.45) \qquad (17.66)$$

$$= 43.28$$

Any contrast which has an absolute value greater than CD is a statistically significant contrast. Thus, at the $\alpha = .05$ level of significance, $\hat{\psi}_2$ and $\hat{\psi}_3$ are significant contrasts. Therefore, in relationship to our example, it would appear that five sessions per week are necessary to increase the test scores and improve reading skills. This conclusion is consistent with the findings of the example source (reference 4).

REFERENCES

1. Conover, W. J. Practical Nonparametric Statistics. New York: Wiley, 1971.

2. Hollander, M. and Wolfe, D. A. Nonparametric Statistical Methods. New York: Wiley, 1973.

3. Marascuilo, L. A. and McSweeney, M. "Nonparametric Post Hoc Comparisons For Trend", Psychological Bulletin, 1967, Vol 67, No. 6, 401-412.

4. Penfield, D. and McSweeney, M. "Powerful Statistical Techniques for Qualitative as well as Quantitative Variables". American Educational Research Association. Pre Session, 1970.

ESTIMATE OF RELIABILITY IN THE STRESS-STRENGTH MODEL

Asit P. Basu*
University of Missouri-Columbia

ABSTRACT

Suppose Y is the strength of a component which is subject to a stress X. Then the component fails whenever $X \geq Y$, and there is no failure when $X < Y$. In this paper the problem of estimating the reliability function

$$R = P(X < Y)$$

is considered. A survey of available results is presented and some new results are considered.

*Research supported by Army Research Office under Grant No. DAA 29-76-G-0301 and by the Air Force Office of Scientific Research under Grant No. AFOSR-75-2795B.

INTRODUCTION

Let X and Y be two random variables with cumulative distribution functions F(x) and G(y) respectively. Suppose Y is the strength of a component subject to a stress X. Then the component fails if at any moment the applied stress (or load) is greater than its strength or resistance. The stress is a function of the environment to which the component is subjected, and its value at any point of time is considered a random variable. The strength of a component is measured by the stress required to failure. Strength depends on material properties, manufacturing procedures and so on. If the components under question are mass produced and their selection in a given system is assumed to be made at random, then the strength should also be considered a random variable. The reliability of a component during a given period [0,T] is taken to be the probability that its strength exceeds the stress during the entire interval, that is, the reliability function R is given by

$$R = P(X < Y)$$

From practical considerations it is desirable to draw inference about the reliability function. The problem of estimating R has been considered by many using nonparametric, Bayesian and parametric approach. We shall present a survey of available results and consider some new results.

The above model was first considered by Birnbaum (1956) and has since found an increasing number of applications in many different

areas, especially in the structural and aircraft industries.

As an example, consider the following problem discussed by Lloyd and Lipow (1962). A solid propellant rocket engine is successfully fired provided the chamber pressure (X) generated by ignition stays below the burst pressure (Y) of the rocket chamber. If $X \geq Y$, the engine blows up and the operation is a failure.

Note the problem of inference about $R = P(X < Y)$ is similar to the problem of estimation of $P = P(X \geq Y)$, the probability of failure. So one can either talk of R, or of P.

2. Nonparametric approach

Let $(X_1, X_2, \ldots, X_m)$ and $(Y_1, Y_2, \ldots, Y_n)$ be two independent samples of measurements on X and Y respectively. Let

$$\phi(X_i, Y_j) = \begin{cases} 1 & \text{if } Y_j < X_i \\ 0, & \text{otherwise} \end{cases}$$

then

$$U = \sum_{i=1}^{m} \sum_{j=1}^{n} \phi(X_i, Y_j)$$

is the well known two sample Mann-Whitney statistic, that is

$$U = \text{number of pairs } (X_i, Y_j) \text{ such that } Y_j < X_i \ .$$

Birnbaum (1956) showed that the Mann-Whitney statistic U could be used to estimate 1 - R (Probability of failure), and hence R. In particular

$$\hat{P} = 1 - \hat{R} = U/mn \qquad (2.1)$$

was proposed as an estimator of $P = Pr$ (failure), and it was used to obtain one sided confidence interval for P for the cases F known, G unknown $(m \to \infty)$, and both F and G unknown. Birnbaum and McCarty (1958) considered a numerical procedure for computing the sample sizes needed for the confidence interval based on U/mn.

Owen, Craswell and Hanson (1964) showed that the assumption of continuity required in Birnbaum (1956) was not essential and produced some tables for use in computing sample sized and confidence intervals for the Birnbaum-McCarty procedure.

Govindarajulu (1968) also has explicitly derived one sided and two sided distribution free confidence bounds for P based on the asymptotic normality of $\hat{P} = U/mn$. This bounds are approximately one half of the corresponding bounds due to Birnbaum and McCarty (1958). In particular, Govinderajulu showed that for all F and G and large m or n, the solution ϵ of the equations

$$P(P \leq \hat{P} + \epsilon) = P(P \geq \hat{P} - \epsilon) \geq \gamma, \quad 0 < \gamma < 1$$

is given by

$$\epsilon \geq (4\nu)^{-1/2} \phi^{-1}(\gamma),$$

and the solution of the equation

$$P(|\hat{P} - P| \leq \epsilon) \geq \gamma, \quad 0 < \gamma < 1$$

is given by

$$c \geq (4\nu)^{-1/2}\, \phi^{-1}\left(\frac{1+\gamma}{2}\right).$$

Here

$$\phi(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X} e^{-u^2/2}\, du\,,$$

and $\phi^{-1}(\cdot)$ is the inverse function of $\phi(\cdot)$.

Recently Govindarajulu (1974) has also considered a sequential distribution-free procedure for obtaining fixed-width confidence limits for P (and hence for R). However, in the absence of additional numerical computation, it is not known how good is the performance of this sequential procedure.

3. Bayesian Approach

Not much has been done from the Bayesian point of view Enis and Geisser (1971) investigated Bayesian approach for estimating R assuming X and Y to be independently distributed and that X and Y are either exponentially distributed or normally distributed.

4. Parametric Approach

In many situations, the distribution of X or (of both X and Y) will be known, and it is desired to obtain parametric solutions. Thus, in case of missile flights, the stress may be expensive to sample, but the physical characteristics of the missile system, such as the propulsive force, angle of elevation, changes in atmospheric condition, and so on may all have known distributions; consequently,

the distribution of stresses may be calculated. In this section, we shall consider the problem of estimating R (or P) for specific parametric distributions.

4.1 Normal Distribution: Owen, Craswell and Hanson (1964) considered above problem and gave one sided confidence intervals for R when both stress and strength are (a) jointly bivariate normally distributed and observations are in pairs, or (b) when X and Y are independent normal with a common unknown variance. Note if X and Y follow a joint bivariate distribution

$$R = P(X < Y) = P(Y - X > 0)$$

$$= \Phi\left(\frac{\mu_Y - \mu_X}{\{\sigma_X^2 - 2\rho\sigma_X\sigma_Y + \sigma_Y^2\}^{1/2}}\right)$$

and $$\hat{R} = \Phi\{\bar{Y} - \bar{X} / (\sigma_X^2 - 2\rho\sigma_X + \sigma_Y^2)^{1/2}\}$$

if σ_X, σ_Y and ρ are known. Similarly if X and Y are independent

$$P(X < Y) = \int_{-\infty}^{\infty} F(x)\, dG(x).$$

Same problems have been considered by Govidarajulu (1976), who obtained two sided confidence intervals for R. Church and Harris (1970) have also considered the same problems under the assumption that X and Y are independent, normally distributed and the distribution of X is known. Assume, without any loss of generality, that $E(X)=0$ and $Var(X)=1$. In this case,

$$R = P\{X < Y\} = \Phi\left(\frac{\mu}{\sqrt{1+\sigma^2}}\right)$$

where $\mu = E(Y)$ and $\sigma^2 = E(Y - \mu)^2$. Church and Harris considered considered the estimator

$$R = \Phi\left(\frac{\bar{Y}}{\sqrt{1 + s^2}}\right) = \Phi(V), \text{ say,}$$

where $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ and $s^2 = \sum_{1}^{n} (Y_i - \bar{Y})^2/(n - 1)$, from which they obtained the following confidence interval for R.

$$P\{\Phi(V - \Phi^{-1}(1 - \tfrac{\gamma}{2})\ \hat{\sigma}_V) < R < \Phi\ (V + \Phi^{-1}(1 - \tfrac{\gamma}{2})\ \hat{\sigma}_V)\} \approx 1 - \gamma$$

Similarly, a one sided confidence interval is given by

$$P\{R > \Phi\ (V - \Phi^{-1}(1 - \gamma)\ \hat{\sigma}_V)\} \approx 1 - \gamma\ .$$

Here

$$\hat{\sigma}_V = \left[\frac{s^2}{1 + s^2}\left(\frac{1}{n} + \frac{\bar{Y}^2 s^2}{2(n - 1)(1 + s^2)^2}\right)\right]^{\frac{1}{2}}$$

The confidence interval obtained by Church and Harris compare favorably with that of Govindarajulu (1968). Their procedure, although empirically demonstrated to be superior to that of Govindarajulu is, however, inexact since it uses the asymptotic normal approximation of a given statistic and requires the substitution of the population mean and standard deviations by their observed sample values. In fact, all the parametric estimators suffer from same weakness as

they are based on maximum likelihood estimators. Mazumdar (1970), has considered the same problem of obtaining point and interval estimates of reliability and obtained mvue of reliability using interference theory. Minimum variance unbiased estimator of R in the normal case has also been considered by Downton (1973).

4.2 Gamma and Exponential distribution: Since in many physical situations, specially in reliability and life testing problems, exponential and gamma distributions provide more realistic models, it is desirable to obtain estimators of R in these cases.

Let X and Y be independently distributed with density functions

$$f(x) = \frac{1}{\Gamma(p)\alpha^p} e^{-x/\alpha} x^{p-1}, \quad x > 0, \; p > 0$$

$$g(y) = \frac{1}{\Gamma(q)\beta^q} e^{-y/\beta} y^{q-1}, \quad y > 0, \; q > 0$$

respectively. Then

$$R = P(X<Y) = \int_0^\infty [1-G(x)]\, dF(x)$$

$$= \int_0^\infty \left[\int_x^\infty \frac{1}{\Gamma(q)\beta^q} e^{-y/\beta} y^{q-1} dy\right] \frac{1}{\Gamma(p)\alpha^p} e^{-x/\alpha} x^{p-1} dx$$

$$= \sum_{k=0}^{q-1} \frac{\Gamma(p+k)}{\Gamma(p)\Gamma(k+1)} \frac{\alpha^k \beta^p}{(\alpha+\beta)^{p+k}} .$$

Here p and q are assumed to be known integers. If two independent random samples $(X_1, X, \ldots X_m)$ and $(Y_1, Y_2 \ldots, Y_n)$ from the two gamma populations are available mle of α and β are given by $\hat{\alpha}=\frac{\bar{X}}{p}$ and $\hat{\beta}=\frac{\bar{Y}}{q}$. Hence mle of R is

$$\hat{R} = \sum_{k=0}^{q-1} \frac{\Gamma(p+k)}{\Gamma(p)\Gamma(k+1)} \frac{\hat{\alpha}^k \hat{\beta}^p}{(\hat{\alpha}+\hat{\beta})^{p+k}} .$$

As special cases, if q=1, that is if X follows the gamma distribution and Y follows the exponential distribution

$$R = (\hat{\beta}/(\hat{\alpha}+\hat{\beta})^p .$$

Finally, if both p and q are equal to 1, we have the case of two independent exponential distributions and we have

$$R = \frac{\hat{\beta}}{\hat{\alpha}+\hat{\beta}} = \frac{\bar{Y}}{\bar{X}+\bar{Y}}$$

The distribution of $\hat{R}$, for large m and n, can be shown to be normal and hence asymptotic confidence interval for $\hat{R}$ can be obtained.

Tong (1974, 1975) has obtained mvube of R for gamma and exponential distributions. The variance of the mvbe of R, in the exponential case has been derived by Kelley et al (1976)

4.3 Weibull distribution: Let X and Y be independent random variables each following the Weibull distribution with common shape

parameter δ. That is let

$$F(x) = 1 - e^{-x^{\delta}/\alpha} \quad , \quad \alpha > 0 \ , \ x > 0$$

$$G(y) = 1 - e^{-y^{\delta}/\beta} \quad , \quad \beta > 0 \ , \ y > 0.$$

We can readily see

$$R = P(X^{\delta} < Y^{\delta}) = P(X < Y) = \frac{\beta}{\alpha + \beta}$$

Note above is independent of δ. Again, we can obtain the mle of R to be

$$R = \hat{\beta}/(\hat{\alpha} + \hat{\beta})$$

where $\hat{\alpha}$ and $\hat{\beta}$ are mle of α and β.

4.4 Bivariate exponential distribution: Since exponential distribution is considered a useful model in life testing problems, it is desirable to consider bivariate analogue of univariate exponential distributions which will have properties similar to the univariate exponential distribution. Marshall and Olkin (1967) have proposed a very important bivariate exponential distribution (BVE), which is given by

$$\bar{F}(x,y)=P(X>x,Y>y)=e^{-\lambda_1 x-\lambda_2 y-\lambda_{12}\max(x,y)},\ 0\le\lambda_1,\lambda_2,\lambda_{12}<\infty,\ \lambda_1+\lambda_{12}>0,\ \lambda_2+\lambda_{12}>0\ (x>0,y>0).$$

The BVE does arise in several natural ways and its properties appear to be fundamental. In particular, marginal distributions of BVE are exponential and BVE has the loss of memory property (LMP) given by

$$\bar{F}(x+t,y+t) = \bar{F}(x,y)\bar{F}(t,t) \quad \text{for} \quad s_1, s_2, t \geq 0$$

However, this distribution is not absolutely continuous and there are clearly situations when it can not be applied. Thus, from data, it is found that $X \neq Y$ for any pair (X,Y) the model is clearly not applicable. An alternative absolutely continuous distribution related to the BVE and having some of its properties would appear to be of interest. To this end, Block and Basu (1974) have proposed an absolutely continuous bivariate exponential extension (ACBVE), which turns out to be the absolutely continuous part of the BVE of Marshall and Olkin. ACBVE is also seen to be a variant of the distribution Freund (1961). The ACBVE is given by

$$\bar{F}(x,y) = \frac{\lambda}{\lambda_1+\lambda_2} \exp[-\lambda_1 x - \lambda_2 y - \lambda_{12}\max(x,y)]$$

$$- \frac{\lambda_{12}}{\lambda_1+\lambda_2} \exp[-\lambda \max(x,y)] \quad \text{for } x>0, y>0.$$

Here

$$\lambda = \lambda_1 + \lambda_2 + \lambda_{12} .$$

Estimates of R when the underlying distribution is BVE or ACBVE has been obtained by Basu (1976). These results will be communicated elsewhere.

5. Reliability of complex systems

The model described before can be extended to more complex systems. For example, a single component system of strength Y could be subjected to k different independent stresses $X_1, X_2, \ldots X_k$. Here reliability of the system is given by

$$R = P\{X_1 < Y, X_2 < Y, \ldots, X_k < Y\}$$

or

$$R = P\{\max(X_1, X_2, \ldots, X_k) < Y\}.$$

An example of interest is the case where a beam with strength Y is subjected to several stresses $X_1, X_2, \ldots, X_k$. Another similar problem of interest is to evaluate the reliability function R' of a k-component system of strengths $Y_1, Y_2, \ldots, Y_k$ respectively each of which is subject to a common stress X. Here

$$R = P\{X < Y_1, X < Y_2, \ldots, X < Y_k\}$$

$$= P\{X < \min(Y_1, \ldots, Y_k)\}.$$

As an example, the flow of a current X through an electronic component assembled from several subcomponents with abilities to accommodate currents $Y_1, Y_2, \ldots, Y_k$ would follow this pattern.

Chandra (1975) has considered the problem of estimating R and R' under the assumption that the X's and Y's are all independent random variables and (a) all follow normal distributions, (b) Y's are all exponential and X is normal with known variance.

Bhattacharyya and Johnson (1974) considered the problem of estimating reliability function R for a more complex m-out-of-k system. Here each of m components of a system of strengths Y_1, Y_2, ... ,Y_k is subjected to a stress X and the system survives if at least m out of the k components survive. Assuming $X, Y_1, \ldots, Y_k$ to be independent with distribution functions $F(x)$, $G_1(y_1)$, $G_2(y_2)$, ... ,$G_k(y_k)$. Bhattacharyya and Johnson considered the problem of estimating the reliability function $R=P_r$(at least m of the $Y_1, \ldots, Y_k$ exceed X), under the assumption $G_1=G_2=\ldots=G_k=G$, say, and that F and G are exponential distributions with known scale parameters. Here

$$R = \sum_{\alpha=m}^{k} \binom{k}{\alpha} \int_{-\infty}^{\infty} [1-G(x)]^{\alpha}[G(x)]^{k-\alpha} d(F(x)) .$$

Bhattacharyya and Johnson (1973) have also considered a nonparametric approach for the above problem.

The author is currently investigating additional problems in this area results of which will be communicated elsewhere.

REFERENCES

1. Basu, A. P. (1971). Bivariate failure rate. J. Amer. Stat. Assn. 66, 103-104.

2. Basu, A. P. (1976). Further results on reliability estimation. (unpublished).

3. Bhattacharyya, G. L. and Johnson, R. (1974). Estimation of reliability in a multicomponent stress-strength model. J. Amer. Stat. Assn, 69, 966-970.

4. Bhattacharyya, G. K. and Johnson R. (1973). Nonparametric estimation of reliability for an s-out-of-k stress-strength system. Technical Report 341, Department of Statistics, University of Wisconsin.

5. Bhattacharyya, G. K. and Johnson, R. (1975). Stress-Strength model for Reliability. Reliability and fault free analysis. Edited by Barlow et.al SIAM, 509-532.

6. Birnbaum, Z. W. (1956). On a use of the Mann-Whitney statistic Proc. 3rd Berkeley Symposium. Vol. 1, 13-17.

7. Birnbaum, A. W. and McCarty, R. C. (1958). A distribution free upper confidence bound for Pr(Y<X) based on independent sample of X and Y. Ann. Math. Statist, 29, 558-562.

8. Block, Henry and Basu, A. P. (1974). A continuous bivariate exponential extension. J. Amer. Stat. Assn. 69. 1031-1037.

9. Chandra, Satish (1975). On estimating the reliability of a component subject to several different stresses (strength). Naval Research Logistic Quarterly 22, 31-39.

10. Church, J. D. and Harris, Bernard (1970). The estimation of reliability from stress-strength relationships. Technometrics 12, 49-54.

11. Downton, F. (1973). The estimation of Pr(Y<X) in the normal case. Technometrics 15, 551-558.

12. Enis, Peter and Geisser, Seymour (1971). Estimation of the probability that Y<X. J. Amer. Stat. Assn. 66, 162-168.

13. Freund, J. E. (1967). A bivariate extension of the exponential distribution. J. Amer. Stat. Assn. 56, 971-977.

14. Govindarajulu, Z. (1967). Two sided confidence limits for P(X<Y) based on normal samples of S and Y. Sankhyā, Series B 29, 35-40.

15. Govindarajulu, Z. (1968). Distribution free confidence bounds for P(X<Y). Ann. Inst. Stat. Math. 20, 299-238.

16. Govindarajulu, Z. (1974). Fixed-width confidence intervals for P(X<Y). Reliability and Biometry, Edited by Proschan and Serfling, 747-757.

17. Kelley, G. D. Kelley, J. A. and Schuncany, W. R. (1976). Efficient estimation of P(Y<) in the exponential case. Technometrics 18, 359-360.

18. Marshall, A. W. and Olkin, I. (1967). A multivariate exponential case. J. Amer. Stat. Assn. 62, 30-44

19. Mazumdar, M. (1970). Some estimates of reliability using inference theory. Naval Research Logistic Quarterly 17, 159-165.

20. Owen, D. B. Craswell, R. J. and Hanson, D. L. (1964). Nonparametric upper confidence bounds for Pr(Y<X) and confidence limits for Pr(Y<X) when X and Y are normal. J. Amer. Stat. Assn. 59, 906-924.

21. Tong, H. (1974). A note on the estimation of Pr(Y<X) in the exponential case. Technometrics 16, 625.

22. Tong, H. (1975). Letters to the editor. Techometrics 17, 393.

UNDERLYING PROBABILITY DISTRIBUTION OF GUN TUBE FATIGUE LIFE

Ronald L. Racicot
Applied Mathematics and Mechanics Division
Benet Weapons Laboratory
Watervliet Arsenal
Watervliet, New York

ABSTRACT. The fracture mechanics studies of gun tube fatigue conducted thus far are essentially deterministic. That is, crack growth and failure are described exactly by assuming that all pertinent parameters are known. Much information has been gained by this approach in studying the important parameters that affect fatigue life. Fatigue life, however, is known to be a random variable. The probabilistic nature of fatigue life must, therefore, be taken into account in the development of gun tubes.

The development approach used at the present time is to schedule gun barrel replacement early enough to forestall failure during firing. Since fatigue life is a random variable, this is accomplished by statistically determining a "safe life" from fatigue test results on a small number of tubes.

In this paper, a probabilistic approach starting with existing theories of fracture mechanics is used to determine the best fit theoretical distribution of life. The main purpose is to improve the present statistical methods for determing safe life by providing a basis for choosing a distribution in analyzing small sample data. The approach used is to assume that the material properties and design parameters in crack growth and failure laws are random variables. Fatigue life is then given as a function of a number of random variables. The fatigue test results for the 105mm M137A1 and 175mm M113E1 tubes are used as bases to estimate means and variances of the model parameters. Monte Carol simulation studies are then conducted by assuming various probability distributions for the model parameters and computing the statistics of the distribution of fatigue lives. Results of the Monte Carlo studies indicate that the best-fit theoretical distributions of fatigue life are the 2- and 3-parameter log-normal.

1. INTRODUCTION. The general problem considered is the fatigue failure of gun tubes resulting from repetitive firing pressure cycles. Numerous studies have been performed at the Watervliet Arsenal and elsewhere on fatigue crack growth and failure of gun tubes [1-12]. These studies include both theoretical fracture mechanics which relate material properties and design parameters to crack growth and experimental measurement on actual gun tubes of crack depth versus number of cycles.

The fracture mechanics studies conducted thus far are essentially deterministic. That is, crack growth and failure are described exactly by assuming that all pertinent parameters are known. Empirical methods are used to estimate some of the model parameters. Much information has been gained by this approach in studying the important parameters that affect fatigue life [10-12]. Fatigue life, however, is known to be a random variable. The probabilistic nature of fatigue life must, therefore, be taken into account in the development of gun tubes.

The development approach used at the present time is to schedule gun barrel replacement early enough to forestall failure during firing. Since fatigue life is a random variable, this is accomplished by statistically determining a "safe life" from fatigue test results on a small number of tubes [4-6,13,14]. The safe life is a statistical tolerance limit [15] for fatigue life for which current specifications require at least a 0.999 probability that tubes will survive the specified safe life. This is determined by first assuming a theoretical distribution of fatigue life and then statistically computing the 0.999 tolerance limit at 90% confidence from a six tube test. The main drawback of this approach is the lack of justification for choosing the theoretical distribution. In the past the 3-parameter Weibull has been arbitrarily assumed [4-6,13].

In this paper, a probabilistic approach starting with existing theories of fracture mechanics is used to determine the best fit theoretical distribution of life. The main purpose is to improve the present statistical methods for determing safe life by providing a basis for choosing a distribution in analyzing small sample data.

The approach used here is to assume that the material properties and design parameters in crack growth and failure laws are random variables. Fatigue life is then given as a function of a number of random variables. The fatigue test results for the 105mm M137A1 and 175mm M113E1 tubes [4,5] are used as bases to estimate means and variances of the model parameters. Monte Carlo simulation studies are then conducted by assuming various probability distributions for the model parameters and computing the statistics of the distribution of fatigue lives [16, p. 124].

2. PROBABILISTIC MODEL BASED ON FRACTURE MECHANICS. There are essentially three phases in the fatigue failure of gun tubes: 1) initiation of cracks; 2) stable crack growth; and 3) failure through unstable crack growth or perforation of the tube surface. Initiation of cracks occurs very early in the life of a tube due primarily to the heat effects of firing the first few rounds [5,10]. The main phenomena in tube fatigue, therefore, are crack growth and failure.

The theories of fracture mechanics for fatigue of tubes are well covered in the literature and Army reports; so only the final results are summarized here (see [11] and references listed in this paper). The crack growth model used in this study is based on the Paris [17] expression for rate of crack growth and on analyses and experimental results of Throop [12], Throop and Miller [11], and others [1-10]. The rate of crack growth is approximated by the expression

$$\frac{db}{dN} = \frac{(\Delta K)^m}{M} \tag{1}$$

in which b = crack depth
N = number of cycles
ΔK = range of variation of stress intensity factor K for one cycle (see [18] for discussion of stress intensity factor)
m = empirical parameter dependent on material and stress intensity
M = empirical parameter dependent on material properties.

In the Throop model [12], a value of m equal to 3.0 gives an adequate overall fit to tube fatigue data although m is known to vary from specimen-to-specimen and for different tube designs. The variables ΔK and M in this model are given as

$$\Delta K = \alpha S\sqrt{\pi b} \tag{2}$$

$$M = \frac{E K_{IC} \sigma_y}{C} \tag{3}$$

in which S = maximum hoop stress at the bore of the tube, $= P(w^2+1)/(w^2-1)$; P = internal pressure, w = O.D./I.D.
α = empirical parameter which depends on crack shape and residual stresses. Compressive residual stresses at the bore of the tube are introduced using the autofrettage process [19,20].
E = Young's modulus
K_{IC} = fracture toughness for a crack in a tangential stress field. K_{IC} is the value of stress intensity K at which unstable crack growth begins.
σ_y = yield strength
C = empirical parameter which varies with m to maintain dimensional homogeneity and may be a function of other material properties.

Substituting (2) and (3) into (1) gives

$$\frac{db}{dN} = \frac{C}{E\sigma_y K_{IC}} (\alpha S\sqrt{\pi b})^m \tag{4}$$

In the probability model, the exponent m is allowed to be a random variable with the mean being determined empirically. The variables E, σ_y, K_{IC}, α and S are random variables.

All of the parameters in (4) can statistically vary from cycle-to-cycle, as a function of crack depth and for different cracks within a given tube. Depth measurements of the largest crack versus number of cycles as well as results of probabilistic studies indicate, however, that the greatest sources of fatigue life variability stem from tube-to-tube variability in the controlling crack growth parameters. Fatigue crack growth in a given tube, therefore, is essentially deterministic in comparison to tube-to-tube variability. The problem then reduces to integrating (4) assuming that material and tube parameters remain constant within a given tube:

$$N_f = N - N_i = \frac{2E\sigma_y K_{IC}}{C(\alpha S\sqrt{\pi})^m (m-2)} \left(b_i^{-\frac{1}{2}(m-2)} - b^{-\frac{1}{2}(m-2)}\right) \quad \text{for } m \neq 2 \tag{5}$$

$$= \frac{E\sigma_y K_{IC}}{C(\alpha S)^2 \pi} \ln(b/b_i) \quad \text{for } m = 2$$

in which b_i = initial crack depth which depends on the heat affected zone and residual stresses.
N_i = initial number of cycles yielding b_i.

In (5), N_i is relatively small and can be assumed zero. The initial crack depth b_i is assumed to be a random variable.

Failure occurs when the crack depth b is either equal to the tube wall thickness B or equal to the critical depth at which unstable growth begins. Unstable crack growth in tubes occurs when

$$b_c = \frac{A}{\pi} \left(\frac{K_{IC}}{\alpha S}\right)^2 \tag{6}$$

in which b_c = critical crack depth
A = empirical constant which accounts for differences in crack shape in the tube and in the specimens used to determine K_{IC}.

Finally, fatigue life N_f is equal to $(N-N_i)$ in (5) where $b = \min(B, b_c)$.

3. LEVELS OF VARIABILITY OF MATERIAL PROPERTIES AND DESIGN PARAMETERS. Equation (5) is a model of fatigue life given as a function of random material and design parameters. The theoretical forms of the distributions of the different model parameters are unknown. The normal, lognormal and Weibull distributions [21] were consequently assumed for the model parameters in studying the form of the distribution of N_f. For these distributions, the mean and variance of each parameter are sufficient to fully characterize the random variables.

Available test data for the 105mm M137A1 and 175mm M113E1 tubes were used as bases to estimate means and variances of the model parameters. Once the model parameters are characterized in a probabilistic sense, sensitivity studies can be performed to determine important factors that influence the statistics of N_f.

a. 105mm M137A1 Tube Data. Table I lists fatigue life and property data for nine 105mm tubes [4]. The fracture toughness was not measured for these tubes and had to be estimated from the yield strength and critical crack depth data using (6) and an empirical relationship for σ_y versus K_{IC} [22]. In addition to this data, crack depth versus number of cycles data were measured on these tubes. The model parameters m, α, and b_i were estimated from this data by fitting the model (5) to the data. Figure 1 shows a comparison of the model to the data for some of the tubes.

TABLE I: FATIGUE AND PROPERTY DATA FOR 105MM M137A1 TUBES

Tube No.	Fatigue Life, Rounds + Cycles	b_c, in	σ_y, ksi	K_{IC}, ksi$\sqrt{in}$ (1)	α (2)
59421	16798	0.80	196	90	.777
59071	12576	0.80	190	99	.851
58046	12469	1.07	171	116	.864
59906	12162	0.60	189	85	.841
62103	10971	0.85	192	107	.891
59895	10801	0.80	187	104	.892
59527	10397	1.05	204	121	.910
59239	9503	0.70	187	100	.921
59531	8882	0.75	207	106	.944

(1) Estimates using equation (6) and $\sigma_y = 334 - 1.39K_{IC}$ [22].
(2) Estimates from crack depth vs. cycles data.

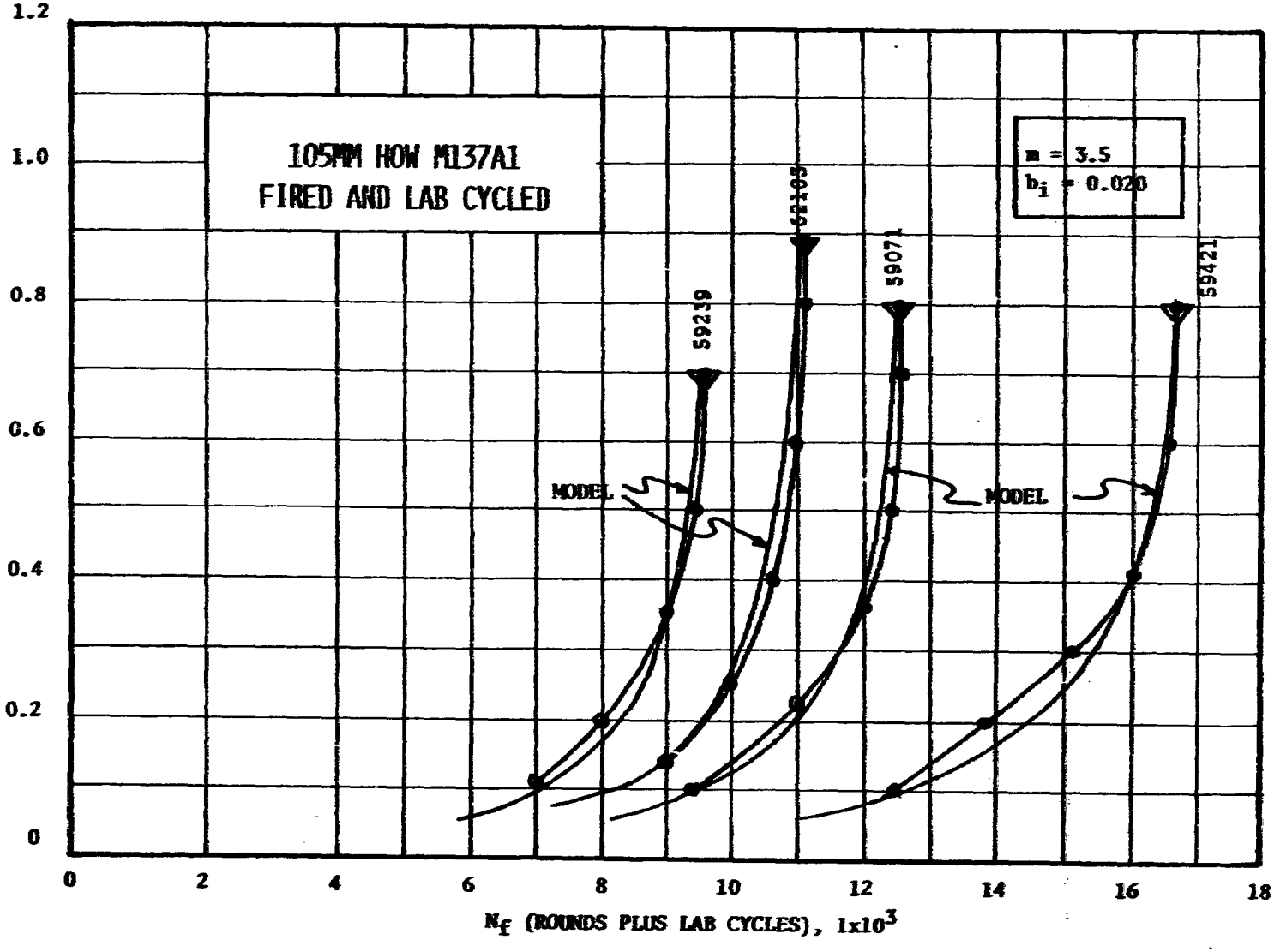


FIGURE 1: CRACK DEPTH VS ROUNDS PLUS CYCLES, 42 KSI, COMPARISON OF MODEL WITH MEASURED DATA

Table II is a summary of the means and standard deviations of the model parameters either estimated from the 105mm tube data or assumed if no data was available.

TABLE II: SUMMARY OF MEANS AND STANDARD DEVIATIONS OF MODEL PARAMETERS FOR 105MM TUBES

Parameter	Mean	Standard Deviation(1)	
D_o, Outside diam., in	6.85	0.0	A
D_i, Inside diam., in	4.21	0.0	A
P, Max. Pressure, ksi	42	0.0	A
α, Crack shape-residual stress parameter	0.877	0.050	E
K_{IC}, Fracture toughness, ksi$\sqrt{in}$	103	11.5	E
σ_y, Yield strength, ksi	191	$\sigma_y = 334 - 1.39K_{IC}$	
b_i, Initial crack depth, in	0.02	0.001	A
m, Rate exponent	3.5	0.1	A
E, Young's Modulus, ksi	30,000	300	A
A, Critical crack depth constant	1.604	0.0	A
C, Empirical constant	0.0333	0.0	A

(1) E ≡ Estimated; A ≡ Assumed

b. 175mm M113E1 Tube Data. Table III summarizes the fatigue and property data either measured or estimated from tests on four 175mm tubes [5]. Figure 2 is a comparison of the model to the crack depth versus cycles data for these tubes. The means and standard deviations estimated from data or assumed for the model parameters are summarized in Table IV.

TABLE III: FATIGUE AND PROPERTY DATA FOR 175MM M113E1 TUBES

Tube No.	Fatigue Life, Rounds + Cycles	b_c, in	σ_y, ksi	K_Q, ksi$\sqrt{in}$	K_{IC}(1) ksi$\sqrt{in}$	α (2)
4134	10974	$\geq$ 3.98	156	130	152	.900
4133	12313	2.40	169	115	115	.874
4127	15255	$\geq$ 3.98	151	124	139	.819
4130	16201	$\geq$ 3.98	153	135	136	.805

(1) K_{IC} was adjusted to account for b_c = 2.40 for tube 4133 by applying equation (6). K_Q is an estimate of K_{IC} using a nonstandard specimen.
(2) Estimates from crack depth versus cycles data.

TABLE IV: SUMMARY OF MEANS AND STANDARD DEVIATIONS OF MODEL PARAMETERS FOR 175MM TUBES

Parameter, (See Table II for Definitions)	Mean	Standard Deviation(1)	
D_o	15.0	0.0	A
D_i	7.04	0.0	A
P	46	0.0(2)	A
α	.8495	0.045	E
K_{IC}	135.5	15.3	E
σ_y(3)	146	$\sigma_y = 334 - 1.39K_{IC}$	
b_i	0.06	0.005	A
m	3.0	0.1	A
E	30000	300	A
A	2.26	0.0	A
C	0.2413	0.0	A

(1) E ≡ Estimated; A ≡ Assumed
(2) Tube-to-tube variation assumed zero; however, cycle-to-cycle standard deviation ≅ 0.90 from [23].
(3) σ_y was computed from the equation given. This resulted in a somewhat lower value than the measured values given in Table III. The computed σ_y is still within the required specifications of 140-160 ksi.

4. BEST FIT PROBABILITY DISTRIBUTION OF FATIGUE LIVES. In this section, the model expressed by equation (5) is used to generate probability distributional information for fatigue lives of tubes. This is accomplished by first assuming probability distributions for the model parameters and then using Monte Carlo simulation to generate the fatigue life distribution. The simulation trials were conducted as follows:

a. The general form of the distribution for the model parameters is fixed. A choice of one of three possible distributions is used; normal, lognormal or Weibull.

b. The mean and standard deviation for each parameter is fixed using the test results and assumptions given in Section 3 as bases. It should be noted that the 105mm and 175mm tube data are used only to provide a starting point for conducting the Monte Carlo trials.

c. A value for each of the random model parameters is generated using random numbers [16, p. 124].

d. The fatigue life for the given set of parameters is computed using (5) and (6).

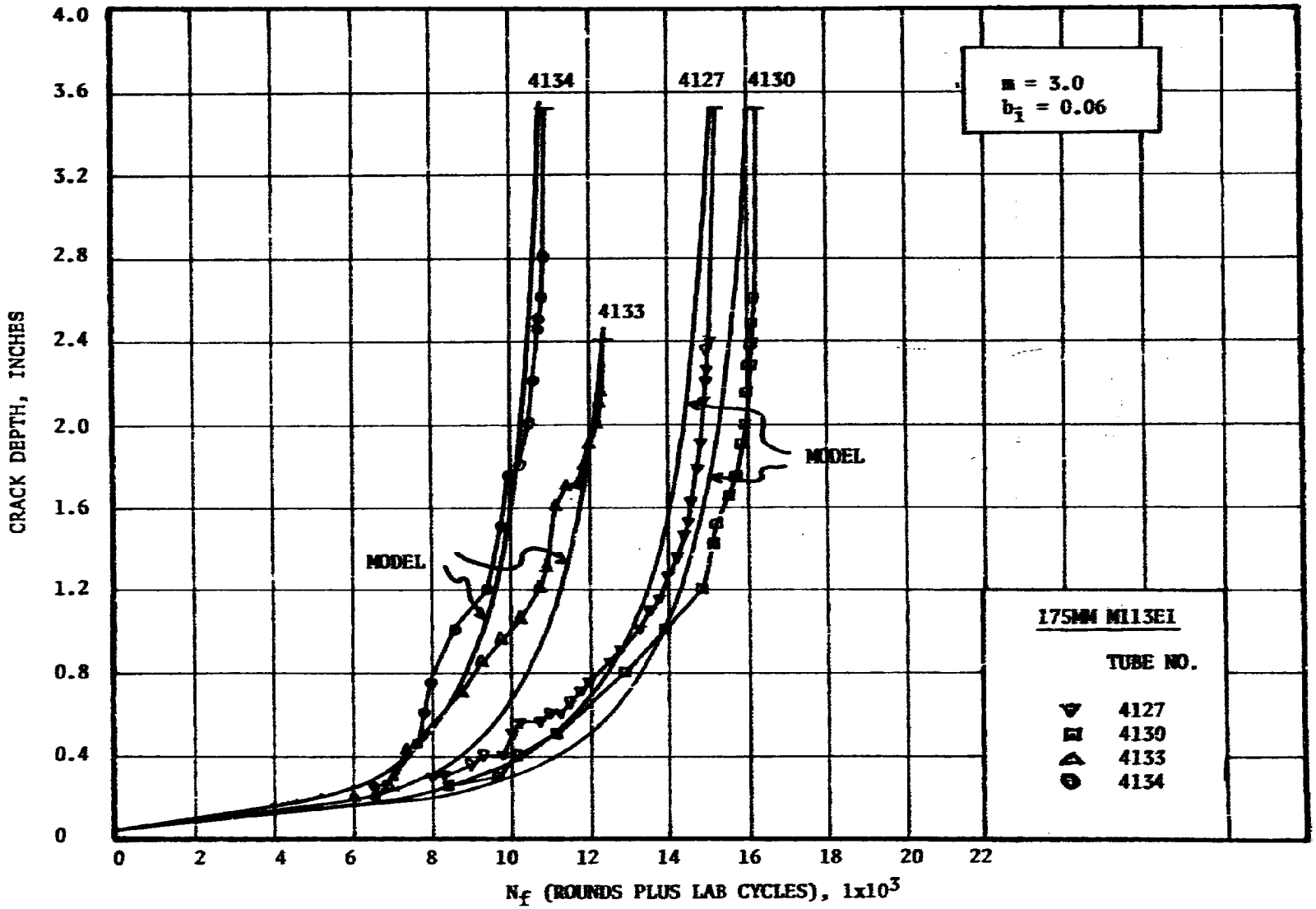


FIGURE 2: CRACK DEPTH VS ROUNDS PLUS CYCLES, ORIGIN OF RIFLING, COMPARISON OF MODEL WITH MEASURED DATA

e. Steps c) and d) are repeated J times (usually 1,000 to 10,000) yielding J different values of fatigue failure times.

f. Various distributional statistics are computed from the J failure times; eg. mean, variance, coefficients of skewness and kurtosis [16, p. 146], 99.0 and 99.9 lower percentiles, and the K-S (Kolmogorov-Smirnov) statistic [16, p. 466].

Steps a) through f) can be repeated for different model parameter distributions, different values of parameter means and standard deviations, different failure criteria, etc.

A number of candidate theoretical distributions were considered for fatigue life; normal, 2- and 3-parameter lognormal, 2- and 3-parameter Weibull and gamma [16,21]. A comparison was made of the various theoretical distributions to the Monte Carlo model distribution. This was done by first fitting the theoretical distribution to the model distribution by equating means and variances. The third parameter in the 3-parameter distributions were fixed by equating the 99.9 lower percentile of the theoretical and model distributions. The reason for this was to match as closely as possible the lower tails of the distributions for comparative purposes. Goodness of fit was then checked using the K-S statistic and by comparing the coefficients of skewness and kurtosis (third and fourth moments) and the 99.0 and 99.9 lower percentiles.

The K-S statistic is a measure of the maximum deviation of a theoretical cumulative distribution from a set of data; the lower the K-S statistic, the better the fit. The data in this case are the Monte Carlo failure times. Table V lists the K-S statistics for the various theoretical distributions as a function of parameter distribution and data bases.

TABLE V: K-S STATISTIC FOR COMPARING MODEL WITH VARIOUS THEORETICAL DISTRIBUTIONS

	K-S Statistic*					
	105mm M137A1 Tubes			175mm M113E1 Tubes		
Failure Time	Parameter Distribution			Parameter Distribution		
Distribution	Normal	Lognormal	Weibull	Normal	Lognormal	Weibull
Normal	.068	.061	.120	.050	.040	.109
2-p Weibull	.084	.078	.135	.082	.073	.136
3-p Weibull	.081	.073	.299	.143	.138	.330
2-p Lognormal	.029	.022	.075	.019	.010	.073
3-p Lognormal	.021	.023	.046	.014	.010	.036
Gamma	.041	.034	.090	.029	.019	.085

*Only 1,000 Monte Carlo trials were used in this case to reduce excessive computer time.

It should be noted that the distributions of the material and design parameters in equation (5) are not known. Different distributions were consequently assumed to indicate the importance of this factor, if any, on conclusions made about the failure time distribution. The K-S statistics given in Table V indicate that the 2- and 3-parameter lognormal provide the best overall fit to the model for the different parameter distributions considered.

An explanation is required for why the K-S statistic in Table V increased in some cases for the 3-parameter distribution in comparison to the 2-parameter distribution. Generally, one would expect a better fit when the number of distribution parameters is increased. This would be true if the 3rd parameter was chosen to minimize the K-S statistic. However, in the gun fatigue problem the main concern is estimating probabilities at the lower tails of the distributions. The third distribution parameter was consequently chosen by equating a given lower percentile. This resulted in a worse fit at the upper tail for some of the cases considered, particularly for the 3-parameter Weibull distribution, resulting in a higher K-S statistic.

In light of the above discussion, it is of interest to compare other goodness-of-fit statistics which would indicate behavior at the lower tails. Table VI lists the coefficients of skewness and kurtosis and the 99.0 and 99.9 lower percentiles for the model and theoretical distributions. The parameter distributions were assumed normal for these particular results with 10,000 Monte Carlo trials run for each case. Again, the lognormal, particularly the 3-parameter lognormal, yielded the best overall fit to the model statistics. Compare, for example, the 99.0 percentiles of the assumed failure time distributions to the model value.

TABLE VI: COMPARISON OF SIMULATED MODEL DISTRIBUTION WITH THEORETICAL DISTRIBUTIONS

Failure Time Dist.	Coefficients of Skewness		Coefficients of Kurtosis		Lower Percentile 99.0		Lower Percentile 99.9*	
	105mm Tubes	175mm Tubes	105mm Tubes	175mm Tubes	105mm Tubes	175mm Tubes	105mm Tubes	175mm Tubes
Normal	0.0	0.0	3.00	3.00	5589	7891	3598	5977
2-p Weibull	-0.27	-0.41	2.90	3.11	5171	7181	3298	5007
3-p Weibull	0.37	0.26	2.87	2.78	6688	8707	6050	7954
2-p Lognormal	0.68	0.55	3.84	3.55	6802	8857	5745	7712
3-p Lognormal	0.81	0.65	4.18	3.75	6992	9003	6050	7954
Gamma	0.45	0.37	3.30	3.20	6456	8571	5218	7256
Model	0.86	0.76	4.45	3.97	6996	9154	6050	7954

*The third parameter for the 3-p distributions was chosen such that the 99.9 percentile was equal to the model results.

There is theoretical justification for why the lognormal could be expected to provide a representation of the fatigue life distribution. The model (5) gives fatigue life as a product of random variables. The limiting distribution for the product of an infinite number of random variables is the lognormal regardless of the form of the distribution of the individual random variables [16, p. 262]. In practice the actual number of random variables required to give a lognormal depends on a number of factors including the form of the distribution of the individual random variables as well as accuracy required for the distribution which is to represent the product. For example, if each random variable in the product is itself lognormal then the product is always lognormal regardless of the number of random variables. It appears that even though equation (5) represents the product of at most seven random variables, this is apparently enough to give a trend toward lognormal as indicated by the results.

5. FUTURE RESEARCH EFFORTS. The results reported in this paper were based on the particular fracture mechanics model given by equation (5). As additional experimental results are obtained this model may be revised as well as the values of the model parameters and their variances. The effect on life distribution must be rechecked in this instance.

In any case, a number of interesting studies may be pursued using the developed probabilistic model:

a. determine the relative effects of variability in design and material parameters on the variability of fatigue life;

b. study possible methods of increasing safe life through control of statistical parameters;

c. study different methods of computing safe life; and

d. improve the initial design approach for new gun tubes.

REFERENCES

[1] J. F. Throop, "Fatigue Crack Growth in Thick-Walled Cylinders", Proc. National Conf. on Fluid Power, Vol. XXVI, September 1972, pp. 115-131.

[2] J. F. Throop, "Fatigue Crack Tolerance in Thick-Walled Cylinders", Technical Report No. WVT-7035, Watervliet Arsenal, Watervliet, New York, November 1970.

[3] B. A. Austin, A. N. Reiner, T. E. Davidson, "Low Cycle Fatigue Strength of Thick-Walled Pressure Vessels", Technical Report No. WVT-6819, Watervliet Arsenal, Watervliet, New York, June 1968.

[4] B. B. Brown, A. N. Reiner, T. E. Davidson, "The Fatigue Life Characteristics of the 105mm M137A1 Howitzer Barrel", Technical Report No. WVT-7202, Watervliet Arsenal, Watervliet, New York, January 1972.

[5] T. E. Davidson, et al, "The Fatigue Life Characteristics of the 175mm M113E1 Gun Tube", Technical Report No. WVT-6912, Watervliet Arsenal, Watervliet, New York, March 1969.

[6] T. E. Davidson, A. N. Reiner, "Fatigue Life Characteristics of the 155mm Howitzer M1A1 Tube", Technical Report No. WVT-7134, Watervliet Arsenal, Watervliet, New York, May 1971.

[7] T. E. Davidson, et al, "Fatigue and Fracture Analysis of the 175mm M113 Gun Tube", Technical Report No. WVT-6822, Watervliet Arsenal, Watervliet, New York, November 1968.

[8] D. P. Kendall, J. H. Underwood, D. C. Winters, "Fracture Toughness and Crack Growth Measurements with "C" Shaped Specimens", Technical Report No. R-WV-T-6-39-73, Watervliet Arsenal, Watervliet, New York, October 1973.

[9] J. J. Miller, D. C. Winters, R. R. Fujczak, "The Measurement and Analysis of Fatigue Crack Growth in Thick-Walled Cylinders", Technical Report No. WVT-TR-74040, Watervliet Arsenal, Watervliet, New York, September 1974.

[10] T. E. Davidson, J. F. Throop, A. N. Reiner, "The Role of Fracture Toughness and Residual Stresses in the Fatigue and Fracture Behavior of Large Thick-Walled Pressure Vessels", Proc. National Conf. on Fluid Power, Vol. XXVI, September 1972, pp. 102-114.

[11] J. F. Throop, G. A. Miller, "Optimum Fatigue Crack Resistance", ASTM Special Technical Publication 467, Philadelphia, PA, 1970, pp. 154-168.

[12] J. F. Throop, "Fatigue Crack Propagation and Fracture in High Pressure Cylinders", Report to be published, Watervliet Arsenal, Watervliet, New York, 1976.

[13] T. Moraczewski, R. W. Clarke, R. Soanes, "The Statistical Determination of Confidenced Safe Fatigue Life for the 175mm M113B1 Gun Tube", Technical Report No. WVT-6909, Watervliet Arsenal, Watervliet, New York, March 1969.

[14] R. W. Clarke, "A General Computational Algorithm for Bayesian Confidence Bounds", Technical Report No. WVT-6911, Watervliet Arsenal, Watervliet, New York, May 1969.

[15] N. L. Johnson, F. C. Leone, Statistics and Experimental Design, Volume I, New York: Wiley, 1964, p. 278.

[16] J. R. Benjamin, C. A. Cornell, Probability, Statistics, and Decision for Civil Engineers, New York: McGraw Hill, 1970.

[17] P. C. Paris, "The Fracture Mechanics Approach to Fatigue", Fatigue, an Interdisciplinary Approach, Syracuse University Press, 1964, pp. 107-132.

[18] C. C. Osgood, "A Basic Course in Fracture Mechanics", Machine Design, Lesson 1: What is Fracture Mechanics, July 22, 1971, pp. 58-64 and Lesson 2: Building on the Fundamentals, August 5, 1971, pp. 88-93.

[19] T. E. Davidson, et al, "The Autofrettage Principle as Applied to High Strength Light Weight Gun Tubes", Technical Report No. WVT-RI-5907, Watervliet Arsenal, Watervliet, New York, October 1959.

[20] T. E. Davidson, et al, "Analysis of the Effect of Autofrettage on the Fatigue Life Characteristics of the 175mm M113 Gun Tube", Technical Report No. WVT-6901, Watervliet Arsenal, Watervliet, New York, January 1969.

[21] O. B. Moan, "Application of Mathematics and Statistics to Reliability and Life Studies", Reliability Handbook, W. G. Ireson, Editor, New York: McGraw Hill, 1966, pp. 4-1 to 4-75.

[22] R. R. Fujczak, "Computer-aided Prediction of Remaining Fatigue Life of Internally Pressurized Cylinders", Technical Report No. WVT-76008, Watervliet Arsenal, Watervliet, New York, March, 1976.

[23] H. B. Anderson, "Effects of Environment on Safe Operating Limits for Artillery Weapon Systems", Report No. DPS-2005, Watervliet Arsenal, Watervliet, New York, June 1966.

ESTIMATION AND EFFECT OF NOISE CORRELATION ON VARIANCE ESTIMATION FROM MOVING ARC SMOOTHING

Paul H. Thrasher
Quality Assurance Office
US Army White Sands Missile Range
White Sands Missile Range, New Mexico

ABSTRACT. Correlation in the noise on Y, in measurements of Y versus X with X assumed exact, does not formally effect the moving arc least-squares estimate of Y. It does, however, effect the variance estimate of Y. Analysis has been done to find correction factors to the zero correlation estimates of (1) the moving arc smoothing factor and (2) the degrees of freedom in the relation

$$[\text{Variance Estimate}] = [\Sigma(Y_{\text{Smoothed}}-Y_{\text{Data}})^2] \frac{[\text{Smoothing Span Factor}]}{[\text{Degrees of Freedom}]} .$$

Both correction factors depend on the correlation matrix. An algorithm has been devised to estimate the correlation matrix by assuming First Order Markov correlation. Problems with the application of the theory are discussed and possible modifications are suggested.

1. INTRODUCTION. In many physical measurements of related quantities X and Y, two conditions exist. First, the independent variable X can be measured so much more accurately than the dependent variable Y that X can be assumed exact. Second, the man and/or machine system which measures Y introduces correlated noise. In one example, the tracking of missiles, X is time and Y is position.

The statistical analysis may be complicated by a lack of knowledge about the physical model describing the data. One approach to this dilemma is to do a least-squares fit of a polynomial to a smoothing span of N data points in order to find a "smoothed" value for the middle point i. To analyze the (i+1)th point, the smoothing span must be shifted one point forward in X and the least-squares analysis must be repeated. In the example of missile tracking, a quadratic polynomial fits a highly restricted physical situation. The quadratic description is rendered invalid by such factors as air resistence, changing rocket thrusts, and stage separation. Since the correct physical description is unknown, however, the quadratic polynomial is normally used.

The theory presented below is based on a polynomial model of degree n. Three sections are devoted to the theory.

First, an algebraic derivation yields values of (a) smoothed positions and corresponding derivatives $d^m Y_{S;i}/dX^m$, (b) estimates of variances of $d^m Y_{S;i}/dX^m$ when the noise correlation is not considered, and (c) correction factors to these variance estimates in order to take correlation into account.

These correction factors are functions of the correlation coefficients and the number of degrees of freedom in the smoothing span.

The second theory section is a matrix derivation which obtains (a) an alternate expression for the polynomial obtained used in the first section and (b) the relation between the number of degrees of freedom and the correlation matrix.

The third section estimates the correlation coefficient in the correlation matrix by using a First Order Markov approximation.

The fourth section reports on difficulties encountered in applying the theory to (a) the output of a white noise generator that has had First Order Markov correlation introduced into it and (b) actual missile tracking data. The basic problem is that the results appear to depend on analysis variables which have no physical influence on the correlation present.

A brief fifth section lists the primary cause of the difficulty and possible corrective procedures. This information was provided by the panel at the presentation of this problem to the Twenty-Second Conference on the Design of Experiments in Army Research, Development, and Testing.

2. ALGEBRAIC RELATION BETWEEN VARIANCE ESTIMATES FOR IGNORING AND CONSIDERING CORRELATION.

This section discusses the effect on the covariance of measurements,

$$\mathrm{COV}(Y_{i+j},Y_{i+j'}) = \rho(i,j,j',\omega,s)\ \mathrm{VAR}(Y_i) \quad , \tag{2.1}$$

on the variance of a least-squares polynomial. If the data's correlation is either non-existent or ignored, the correlation coefficient, $\rho(i,j,j',\omega,s)$, is set equal to $\delta_{jj'}$. In general, however, the measuring device's bandwidth and measurement interval, ω and s, result in $\rho \neq \delta_{jj'}$. The following equations trace the influence of the data correlation through the moving arc smoothing process.

The calculation of the smoothed dependent variable does not formally depend on the correlation in the data. An n^{th} degree polynomial, $Y_{S;i+j}$, is constructed through N data points. The i^{th} point is in the center of this smoothing span and j ranges from $-\alpha = -(N-1)/2$ to α to locate individual measurements. The polynomial is a summation over orthonormal function which are defined by

$$F_k(sj) = \sum_{\ell=0}^{k} C_{k\ell}(j)^{\ell} \quad ; \tag{2.2}$$

where orthonormality determines the $C_{k\ell}$'s; thus, the polynomial is

$$Y_{S;i+j} = \sum_{k=0}^{n} A_k(i)\, F_k(sj) = \sum_{k=0}^{n} \sum_{\ell=0}^{k} A_k(i)\, C_{k\ell}(j)^{\ell} \quad . \tag{2.3}$$

A least-squares calculation minimizes

$$\sum_{j=-\alpha}^{\alpha} e_j^2 = \sum_{j=-\alpha}^{\alpha} (Y_{S;i+j} - Y_{i+j})^2 \tag{2.4}$$

to determine the constants to be

$$A_k(i) = \sum_{j=-\alpha}^{\alpha} F_k(sj)\, Y_{i+j} \quad . \tag{2.5}$$

The m^{th} derivative of $Y_{S;i}$ is obtained by m differentiations of $Y_{S;i+j}$ with respect to $(X_i + sj)$ and then setting j equal to 0; this results in

$$d^m Y_{S;i}/dX^m = \left[(1/s^m)\, d^m Y_{S;i+j}/dj^m\right]_{j=0} = (m!/s^m) \sum_{k=m}^{n} A_k(k)\, C_{km}$$

$$= \sum_j a_{jm}\, Y_{i+j} \tag{2.6}$$

where

$$a_{jm} = \frac{m!}{s^m} \sum_{k=m}^{n} \sum_{\ell=0}^{k} C_{km}\, C_{k\ell}(j)^{\ell} \quad .$$

Since the $d^m Y_{S;i}/dX^m$ values are functions of the Y_{i+j} data through the $A_k(i)$ values, the errors in these derivatives are also dependent on the errors and correlation of the Y_{i+j} data. The variance of $d^m Y_{S;i}/dX^m$ is calculated from expectation relations to be

$$VAR(d^m Y_{S;i}/dX^m) = (m!/s^m)^2 \sum_{k=m}^{n} \sum_{k'=m}^{n} C_{km} C_{k'm}\, COV[A_k(i), A_{k'}(i)] \quad . \tag{2.7}$$

The covariance of $A_k(i)$ and $A_{k'}(i)$ is found to be

$$\mathrm{COV}[A_k(i),A_{k'}(i)] = \sum_{j=-\alpha}^{\alpha} \sum_{j'=-\alpha}^{\alpha} (F_{k'}(sj') \times$$

$$E[(Y_{i+j}-Y_{\mu;i+j})(Y_{i+j'}-Y_{\mu;i+j'})]) \qquad (2.8)$$

where $Y_{\mu;i+j}$ is the true mean of the i+j point. By using

$$\mathrm{COV}(Y_{i+j},Y_{i+j'}) = \delta_{jj'}\mathrm{VAR}(Y_i)$$

for

$$E[(Y_{i+j}-Y_{\mu;i+j})(Y_{i+j'}-Y_{\mu;i+j'})] \quad ,$$

the correlation-ignored result is found to be

$$[\mathrm{VAR}(d^m Y_{S;i}/dX^m)]_{\text{CORR-IG}} = (m!/s^m)^2\ \mathrm{VAR}(Y_i) \sum_{k=m}^{n} C_{km}^2 \quad . \qquad (2.9)$$

The degrees of freedom used in $\mathrm{VAR}(Y_i)$ is n+1 less than N. By using the general expression $\rho(i,j,j',\omega,s)\ \mathrm{VAR}(Y_i)$ for $\mathrm{COV}(Y_{i+j},Y_{i+j'})$, the correlation-considered result is found and the ratio of the correlation-considered estimate of variance to the correlation-ignored estimate is calculated to be the product of

$$R_{mm}(i) = \frac{\sum_{k=m}^{n} \sum_{k'=m}^{n} \sum_{\ell=0}^{k} \sum_{\ell'=0}^{k'} \sum_{j=-\alpha}^{\alpha} \sum_{j'=-\alpha}^{\alpha} C_{km}C_{k'm}C_{k\ell}C_{k'\ell'}(j)^{\ell}(j')^{\ell'}\rho(i,j,j',\omega,s)}{\sum_{k=m}^{n} C_{km}^2} \qquad (2.10)$$

where $(j)^{\ell} = 1$ for $j=\ell=0$ and $(j')^{\ell'} = 1$ for $j'=\ell'=0$, and

$$F_n = \frac{N-(n+1)}{N-T} \qquad (2.11)$$

where T is the "true" reduction in the degrees of freedom discussed in Section 3.

For segments of the data in which the correlation coefficient may be assumed constant in i and symmetric in j and j', the $R_{mm}(i)$ may be rewritten to expedite computer calculations. For j=j', the correlation coefficient must

be unity. This fact and the symmetry about the two diagonals in the array of possible j and j′ values are used to rewrite the numerator of R_{mn}. The definition and orthonormality of the $F_k(sj)$ functions simplify the sum over the main diagonal; the result is

$$R_{mn} = \left[\sum_{k=m}^{n} C_{km}^2\right]^{-1} \left[\sum_{k=m}^{n} C_{km}^2 + 2 \sum_{k=m}^{n} \sum_{k'=m}^{n} \sum_{\ell=0}^{k} \sum_{\ell'=0}^{k'} \sum_{j=1}^{(N-1)/2} C_{km}C_{k'm}C_{k\ell}C_{k'\ell'}(j)^{\ell}(-j)^{\ell'}\rho(j,-j,\omega,s) + 4 \sum_{k=m}^{n} \sum_{k'=m}^{n} \sum_{\ell=0}^{k} \sum_{\ell'=0}^{k'} \sum_{j=1}^{(N-1)/2} \sum_{j'=-(j-1)}^{(j-1)} C_{km}C_{k'm}C_{k\ell}C_{k'\ell'}(j)^{\ell}(j')^{\ell'}\rho(j,j',\omega,s)\right] \qquad (2.12)$$

where $(j')^{\ell'} = 1$ for $j'=\ell'=0$.

The constants in the orthonormal functions may be obtained by a bootstrap derivation by starting with C_{00}. The non-zero constants for subscripts less than or equal to 5 are

$$C_{00} = \sqrt{1/N}$$

$$C_{11} = \sqrt{12/N(N^2-1)}\ ,$$

$$C_{20} = \sqrt{5(N^2-1)/4N(N^2-4)}\ ,$$

$$C_{22} = \sqrt{180/N(N^2-1)(N^2-4)}\ ,$$

$$C_{31} = \sqrt{7(3N^2-7)^2/N(N^2-1)(N^2-4)(N^2-9)}\ ,$$

$$C_{33} = \sqrt{2800/N(N^2-1)(N^2-4)(N^2-9)}\ ,$$

$$C_{40} = \sqrt{81(N^2-1)(N^2-9)/64N(N^2-4)(N^2-16)}\ ,$$

$$C_{42} = \sqrt{225(3N^2-13)^2/N(N^2-1)(N^2-4)(N^2-9)(N^2-16)}\ ,$$

$$C_{44} = \sqrt{44100/N(N^2-1)(N^2-4)(N^2-9)(N^2-16)} \ ,$$

$$C_{61} = \sqrt{11(15N^4-230N^2+407)^2/16N(N^2-1)(N^2-4)(N^2-9)(N^2-16)(N^2-25)} \ ,$$

$$C_{63} = \sqrt{53900(N^2-7)^2/N(N^2-1)(N^2-4)(N^2-9)(N^2-16)(N^2-25)} \ , \text{ and}$$

$$C_{66} = \sqrt{898544/N(N^2-1)(N^2-4)(N^2-9)(N^2-16)(N^2-25)} \ . \qquad (2.13)$$

The systematic occurrence of zeros in the table of $C_{k\ell}$ values may be used to further expedite computer calculations. Since $C_{k\ell} = 0$ unless $k + \ell$ is even, each term in the sums of Equation (2.12) is identically zero unless $k + m$, $k' + m$, $k + \ell$, and $k' + \ell'$ are all even.

3. MATRIX DERIVATION OF NUMBER OF DEGREES OF FREEDOM.

The raw variance of data analyzed with a smoothing span of N points is given by

$$VAR = \frac{E(\underset{\sim}{E}_e^T \underset{\sim}{E}_e)}{N-T} \ . \qquad (3.1)$$

The numberator, $E(\underset{\sim}{E}_e^T \underset{\sim}{E}_e)$, is the expectation value for the sum of the squares of the differences between data values and corresponding smoothed or filtered values. The denominator, N-T, is called the number of degrees of freedom. The reduction in the degrees of freedom, T, is dependent on the correlation of the data in the smoothing span. For zero correlation, T is one more than the degree of the polynomial used for smoothing. This corresponds to the number of constants in the polynomial. For the total correlation, T is equal to N. In this case, the variance is undefined. The following derivation yields a description of the degrees of freedom for intermediate correlations.

The dependent variables, Y_r, may be arranged in N by 1 matrices. Each of these column matrices are related to the independent variables, polynomial coefficients, and random errors by

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{B} + \underset{\sim}{E}_c \ . \qquad (3.2)$$

The rth row of the random error column matrix, $\underset{\sim}{E}_c$, contains the error, ϵ_r, of the rth dependent variable, Y_r. The nth degree polynomial coefficients, B_n, B_{n-1}, ..., B_0, are in the n+1 by 1 column matrix $\underset{\sim}{B}$. The N by n+1 matrix $\underset{\sim}{X}$ may

be considered as a composite of n+1 column matrices, $\underset{\sim}{X}_n$, $\underset{\sim}{X}_{n-1}$, ..., $\underset{\sim}{X}_0$. The rth row of each $\underset{\sim}{X}_c$ contains the cth power of the independent variable X_r that corresponds to the dependent variable Y_r. The smoothed or filtered dependent variables $\hat{Y}_r$, are given by the independent variables, $\underset{\sim}{X}$, and estimates of the polynomial coefficients, $\hat{\underset{\sim}{B}}$, by

$$\hat{\underset{\sim}{Y}} = \underset{\sim}{X}\,\hat{\underset{\sim}{B}} \quad . \tag{3.3}$$

A least-squares calculation may be used to find $\hat{\underset{\sim}{B}}$. The summation of squares for the deviations,

$$e_r = Y_r - \hat{Y}_r \quad , \tag{3.4}$$

is given by

$$\sum_{r=1}^{N} e_r^2 = \underset{\sim}{E}_e^T \underset{\sim}{E}_e \tag{3.5}$$

where $\underset{\sim}{E}_e^T$ is the row matrix which is the transpose of the column matrix $\underset{\sim}{E}_e$ containing the e_r's. Substitution of $\underset{\sim}{E}_e = \underset{\sim}{Y} - \hat{\underset{\sim}{Y}} = \underset{\sim}{Y} - \underset{\sim}{X}\hat{\underset{\sim}{B}}$, differentiation of the sum of squares with respect to $\hat{B}_i$, and setting the result equal to zero yields

$$0 = [-\underset{\sim}{Y}^T \underset{\sim}{X} \underset{\sim}{I}_r + \hat{\underset{\sim}{B}}^T \underset{\sim}{X}^T \underset{\sim}{X} \underset{\sim}{I}_r] + [-\underset{\sim}{Y}^T \underset{\sim}{X} \underset{\sim}{I}_r + \hat{\underset{\sim}{B}}^T \underset{\sim}{X}^T \underset{\sim}{X} \underset{\sim}{I}_r]^T \tag{3.6}$$

where $\underset{\sim}{I}_r$ is a column matrix defined in terms of Kronicher delta functions, $\delta_{k,\ell} = 0$ if $k \neq \ell$ and $\delta_{k,k} = 1$, by

$$\underset{\sim}{I}_r = \begin{bmatrix} \delta_{n,1} \\ \delta_{n-1,1} \\ \vdots \\ \delta_{01} \end{bmatrix} \tag{3.7}$$

Since each of the two terms in Equation (3.6) are scalars (i.e., 1 by 1 matrices) and the second is the transpose of the first, the two terms are equal. Thus, Equation (3.6) simplifies to

$$0 = 2[-\underset{\sim}{Y}^T \underset{\sim}{X} \underset{\sim}{I}_r + \hat{\underset{\sim}{B}}^T \underset{\sim}{X}^T \underset{\sim}{X} \underset{\sim}{I}_r] \quad . \tag{3.8}$$

This further simplifies to

$$\hat{\underset{\sim}{B}} = (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{Y} \tag{3.9}$$

where the $^{-1}$ superscript is the standard notation for inverse. This result utilizes the raw data to estimate polynomial coefficients.

The raw variance of the data within the smoothing span is found by relating the expectation values of $\underset{\sim}{E}_e^T\underset{\sim}{E}_e$, $\underset{\sim}{E}_t^T\underset{\sim}{E}_t$, and $\underset{\sim}{E}_c\underset{\sim}{E}_c^T$. The first is estimated by the sum of the squares of Y_r-Y_r, the second is the product of N and the desired variance denoted by σ^2, and the third is the product of the correlation matrix, $\underset{\sim}{V}$, and σ^2. Substitution of Equations (3.3) (3.9) and (3.2) into the definition $\underset{\sim}{E}_e = \underset{\sim}{Y}-\underset{\sim}{Y}$ yields

$$\underset{\sim}{E}_e = [\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T]\,\underset{\sim}{E}_t \tag{3.10}$$

where $\underset{\sim}{I}$ is the standard unit matrix whose elements are defined by $I_{ij} = \delta_{ij}$, and $(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}$ is the inverse of $(\underset{\sim}{X}^T\underset{\sim}{X})$ defined such that

$$(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}(\underset{\sim}{X}^T\underset{\sim}{X}) = (\underset{\sim}{X}^T\underset{\sim}{X})(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} = \underset{\sim}{I} \quad .$$

Equation (3.10) leads immediately to

$$\underset{\sim}{E}_e^T\underset{\sim}{E}_e = \underset{\sim}{E}_t^T[\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T]\,\underset{\sim}{E}_t \quad . \tag{3.11}$$

Taking the expectation value of Equation (3.11) yields

$$E(\underset{\sim}{E}_e^T\underset{\sim}{E}_e) = E\{\underset{\sim}{E}_t^T\underset{\sim}{E}_t\} - E\{\underset{\sim}{E}_t^T\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{E}_t\} \quad .$$

The first term on the right is just $N\sigma^2$. The last term may be simplifed by noting that the quantity in braces is a 1 by 1 matrix, replacing this simple matrix by its trace, and using the identify

$$\text{Trace}\,(\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C}) = \text{Trace}\,(\underset{\sim}{B}\underset{\sim}{C}\underset{\sim}{A}) = \text{Trace}\,(\underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B}) \quad .$$

Further simplification is made by interchanging the order of expectation and trace operations and finally by making the usual assumption that the measurements in $\underset{\sim}{X}$ are exact so

$$E[f(\underset{\sim}{X})(\underset{\sim}{E}_c\underset{\sim}{E}_c^T)] = f(\underset{\sim}{X})E(\underset{\sim}{E}_c\underset{\sim}{E}_c^T) \quad .$$

The result is

$$E(\underset{\sim}{E}_e^T \underset{\sim}{E}_e) = N\sigma^2 - \text{Trace}\{\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T E(\underset{\sim}{E}_e\underset{\sim}{E}_e^T)\} \quad . \tag{3.12}$$

The use of $E(\underset{\sim}{E}_e\underset{\sim}{E}_e^T) = \underset{\sim}{V}\sigma^2$ yields

$$\sigma^2 = \frac{E(\underset{\sim}{E}_e^T \underset{\sim}{E}_e)}{N-\text{Trace}\{\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{V}\}} \quad . \tag{3.13}$$

Equation (3.13) is cumbersome because the trace is performed on an N by N matrix. Trace algebra converts the quantity inside the braces to a n+1 by n+1 matrix. Estimation of $E(\underset{\sim}{E}_e^T \underset{\sim}{E}_e)$ by $\underset{\sim}{E}_e^T \underset{\sim}{E}_e$ then yields an estimate of the raw variance to be

$$\hat{\sigma}^2 = \frac{\underset{\sim}{E}_e^T \underset{\sim}{E}_e}{N-\text{Trace}\{(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{V}\underset{\sim}{X}\}} \quad . \tag{3.14}$$

The evaluation of the effective degrees of freedom, i.e., the denominator of Equation (3.14) is dependent on the data through $\underset{\sim}{V}$ and the smoothing process through $\underset{\sim}{X}$ and N. The $\underset{\sim}{X}$ matrix is given in terms of $x_j = js + s_i$ where x_i is the mid-point of the smoothing span and s is the measurement interval. The general form is

$$\underset{\sim}{X} = \begin{bmatrix} x_{-\alpha}^n & x_{-\alpha}^{n-1} & \cdots\cdots & 1 \\ x_{-\alpha+1}^n & x_{-\alpha+1}^{n-1} & \cdots\cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_{\alpha}^n & x_{\alpha}^{n-1} & \cdots\cdots & 1 \end{bmatrix} \tag{3.15}$$

where α is defined by (N-1)/2. Although $\underset{\sim}{X}$ depends on s and x_i, the degrees of freedom do not. The independent variable's increments, s, has no effect because it does not effect either the variance or the sum of the squares of the deviations. The midpoint of the independent variable segment, x_i, has no effect under the necessary assumption that $\underset{\sim}{V}$ matrix describes the correlation in all segments considered. For computational ease, s and x_i may be set equal to 1 and 0 for

the degrees of freedom calculation. This simplifies x_j to $x_j = j$. The calculation of the n+1 by n+1 matrix, $\underset{\sim}{X}^T\underset{\sim}{\underset{\sim}{V}}\underset{\sim}{X}$, is further simplified if $\underset{\sim}{\underset{\sim}{V}}$ is described with a single Markov constant ρ by $V_{rc} = \rho^{|r-c|}$. The other n+1 by n+1 matrix, $(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}$, may be obtained by either analytically or computationally finding the inverse of

$$\underset{\sim}{X}^T\underset{\sim}{X} = \begin{bmatrix} \sum x_j^{2n} & \sum x_j^{2n-1} & \cdots & \sum x_j^{n} \\ \sum x_j^{2n-1} & \sum x_j^{2n-2} & \cdots & \sum x^{n-1} \\ \vdots & \vdots & & \vdots \\ \sum x_j^{n} & \sum x_j^{n-1} & \cdots & 2\alpha+1 \end{bmatrix} = \begin{bmatrix} \sum j^{2n} & \cdots & \sum j^{n} \\ \cdot & & \cdot \\ \vdots & & \vdots \\ \sum j^{n} & \cdots & 2\alpha+1 \end{bmatrix} \tag{3.16}$$

where all summations are over the range $-\alpha \le j \le \alpha$. The summations over powers of j may be found with either a computer or a mathematics handbook.

If one desires an explicit equation for the degrees of freedom, the procedure of the above paragraph can be done analytically. The results for n=0, n=1, and n=2 are, respectively:

$$DOF_0 = N - \frac{1}{(2\alpha+1)}\,\underset{\sim}{H}_0^T\underset{\sim}{\underset{\sim}{V}}\underset{\sim}{H}_0 \quad , \tag{3.17}$$

$$DOF_1 = N - \frac{1}{2\alpha+1}\left[\frac{3}{\alpha(\alpha+1)}\,\underset{\sim}{H}_1^T\underset{\sim}{\underset{\sim}{V}}\underset{\sim}{H}_1 + \underset{\sim}{H}_0^T\underset{\sim}{\underset{\sim}{V}}\underset{\sim}{H}_0\right] \quad , \text{ and} \tag{3.18}$$

$$DOF_2 = N - \frac{45}{\alpha(\alpha+1)(2\alpha+1)(2\alpha+3)(2\alpha-1)}\left[\underset{\sim}{H}_2^T\underset{\sim}{\underset{\sim}{V}}\underset{\sim}{H}_2 + \frac{(2\alpha-1)(2\alpha+3)}{15}\,\underset{\sim}{H}_1^T\underset{\sim}{\underset{\sim}{V}}\underset{\sim}{H}_1 - \frac{2\alpha(\alpha+1)}{3}\,\underset{\sim}{H}_2^T\underset{\sim}{\underset{\sim}{V}}\underset{\sim}{H}_0 + \frac{\alpha(\alpha+1)(3\alpha^2+3\alpha-1)}{15}\,\underset{\sim}{H}_0^T\underset{\sim}{\underset{\sim}{V}}\underset{\sim}{H}_0\right] \tag{3.19}$$

where $\underset{\sim}{H}_j$ is given by

$$\underset{\sim}{H}_j = \begin{bmatrix} (-\alpha)^j \\ (-\alpha+1)^j \\ \vdots \\ (\alpha)^j \end{bmatrix} \tag{3.20}$$

with the understanding that $[0]^0 = 1$. For each of these three equations, substitution of $\underset{\sim}{V} = \underset{\sim}{H}_0\underset{\sim}{H}_0^T$ yields zero. This simply states that totally correlated data has zero degrees of freedom.

If the identify matrix is used for $\underset{\sim}{V}$ in Equations (3.17), (3.18), and (3.19), the results are $DOF_0 = N-1$, $DOF_1 = N-2$, and $DOF_2 = N-3$. This checks with $DOF_n = N-(n+1)$, i.e., the number of degrees of freedom equals the number of points in the smoothing span minus the number of constants in the polynomial.

4. ESTIMATE OF CORRELATION COEFFICIENTS.

The covariance of two raw data points, Y_{i+j} and $Y_{i+j'}$, is related to their correlation coefficient and the variance of the points in the range $i-\alpha \leq (j \text{ or } j') \leq i+\alpha$ by

$$COV(Y_{i+j}, Y_{i+j'}) = \rho_{i+j,i+j'} \cdot VAR(Y_i) \quad . \tag{4.1}$$

The pseudo-deviations are defined by

$$e_{i+j} = Y_{i+j} - Y_{S;i+j} \tag{4.2}$$

where $Y_{S;i+j}$ is not the true mean which would yield the true deviations; instead, it is the kth degree polynomially smoothed value from the operation

$$Y_{S;i+j} = \sum_{p=-\alpha}^{\alpha} a_p \, Y_{i+j+p} \quad . \tag{4.3}$$

The a_p's are restricted by

$$\sum_{p=-\alpha}^{\alpha} a_p = 1 \quad ,$$

and are defined by Equation (2.6) with $j = p$ and $m = 0$.

By using two fast Fourier transforms and associated manipulations, $COV(e_{i+j}, e_{i+j'})$, may be obtained. The needed quantities, however, are either $COV(Y_{i+j}, Y_{i+j'})$ or $\rho_{i+j,i+j'}$ and $VAR(Y_i)$. Unfortunately, these cannot be obtained without applying constraints. Presented below is a method of determining $\rho_{i+j,i+j'}$ and $VAR(Y_i)$ assuming that Y_{i+j} is a kth degree polynomial with additive First Order Markov error.

By defining $b_p \equiv \delta_{p_0} - a_p$, the pseudo-deviations may be found from

$$e_{i+j} = \sum_{p=-\alpha}^{\alpha} b_p Y_{i+j+p} \quad . \tag{4.4}$$

By using

$$\sum_{p=-\alpha}^{\alpha} a_p = 1$$

and

$$\sum_{p=-\alpha}^{\alpha} b_p = 0 \quad ,$$

the expectation value of $(e_{i+j}, e_{i+j'})$ may be shown to equal both $COV(e_{i+j}, e_{i+j'})$ and

$$\sum_{p=-\alpha}^{\alpha} \sum_{q=-\alpha}^{\alpha} b_p b_q \, COV(Y_{i+j+p}, Y_{i+j'+q}) \quad .$$

These results and the use of Equation (4.1) leads to an expression,

$$COV(e_{i+j}, e_{i+j'}) = VAR(Y_i) \sum_{p=-\alpha}^{\alpha} \sum_{q=-\alpha}^{\alpha} b_p b_q \rho_{i+j+p,i+j'+q} \quad , \tag{4.5}$$

which relates the known pseudo-deviation covariances to the desired raw data correlation factors. This equation cannot be solved for $\rho_{i+j,i+j'}$, however, because the double summation is over $(2\alpha+1)^2$ terms. In order to circumvent this problem of having more unknowns than equations, it is convenient to mathematically model the correlation factor.

The First Order Markov error in the i+1 point, ε_{i+1}, is given in terms of a single Markov constant, ρ, the error of the i point, ε_i, and a random variable, η_{i+1}, by

$$\varepsilon_{i+1} = \rho \varepsilon_i + \eta_{i+1} \quad . \tag{4.6}$$

Relating expectation values of (, ε_{i+j}) for all values of j may be used to express the correlation coefficient $\rho_{j,j'}$ as

$$\rho_{j,j'} = \rho^{|j-j'|} \quad . \tag{4.7}$$

By using Equation (4.7) and defining an index $\phi = j' - j$, Equation (4.5) becomes

$$COV(e_{i+j}, e_{i+j'}) = VAR(Y_i) \sum_{p=-\alpha}^{\alpha} \sum_{q=-\alpha}^{\alpha} b_p b_q \rho^{|\phi+q-p|} , \qquad (4.8)$$

This set of equations has only two unknowns, ρ and $VAR(Y_i)$. The straightforward approach would be to define a deviation by

$$\Delta = COV(e_{i+j}, e_{i+j'}) - VAR(Y_i) \sum_{p=-\alpha}^{\alpha} \sum_{q=-\alpha}^{\alpha} b_p b_q \rho^{|\phi+q-p|} , \qquad (4.9)$$

calculate a sum of squares by

$$S = \sum_{\phi=-2\alpha}^{2\alpha} \Delta^2 \qquad (4.10)$$

and find the values of β and $\widehat{VAR}(Y_i)$ that simultaneously satisfy

$$0 = \left.\frac{\partial S}{\partial \rho}\right]_{\rho=\beta,\ VAR(Y_i) = \widehat{VAR}(Y_i)} \qquad (4.11)$$

and

$$0 = \left.\frac{\partial S}{\partial VAR(Y_i)}\right]_{\rho=\beta,\ VAR(Y_i) = \widehat{VAR}(Y_i)} \qquad (4.12)$$

Unfortunately, the direct procedure is algebraically intractable. An alternate approach is to first perform a calculation of Equation (4.12) and find the $\widehat{VAR}(Y_i)$ as a function of ρ to be

$$\widehat{VAR}(Y_i) = \frac{\sum_{\phi=-2\alpha}^{2\alpha} \left[COV(e_{i+j}, e_{i+j'}) \sum_{p} \sum_{q} b_p b_q \rho^{|\phi+q-p|} \right]}{\sum_{\phi=-2\alpha}^{2\alpha} \left(\sum_{p} \sum_{q} b_p b_q \rho^{|\phi+q-p|} \right)^2} ; \qquad (4.13)$$

and second, define a new deviation as a function of ρ only as

$$\Delta' \equiv \mathrm{COV}(e_{i+j}, e_{i+j-\phi}) - \widehat{\mathrm{VAR}}(Y_i) \sum_p \sum_q b_p b_q \rho^{|\phi+q-p|} \tag{4.14}$$

and graphically find $\hat{\rho}$ as the value of ρ which minimizes

$$S' \equiv \sum_\phi \Delta'^2 \tag{4.15}$$

where ϕ is still bounded by $-2\alpha \leq \phi \leq 2\alpha$. In this graphical procedure S' is of course a function of ρ. The saving restriction which makes the procedure tractable is that ρ is bounded by $-1 \leq \rho \leq +1$. The computation work is still considerable; however, so it is worthwhile to use the invariance under change in sign of ϕ of $\mathrm{COV}(e_{i+j}, e_{i+j+\phi})$, $\sum_p \sum_q b_p b_q \rho^{|\phi+q-p|}$, and $\Delta'(\phi)$.

5. NUMERICAL RESULTS. Two sets of numbers have been analyzed in order to determine the usefulness of the theory in the last three sections. The first set has been generated by using a random noise generator and introducing First Order Markov correlation of known ρ. The second set is from a missile versus drone test at White Sands Missile Range.

The generated numbers do not lead to completely desirable results from the analysis. Table I shows the input and one set of output of the computerized equations from the last three sections. For large values of ρ, the two resulting variance estimates agree with each other but diverge considerably from the input variance. The basic discrepancy occurs in the output ρ.

Comparison of the left and right columns of Table I shows deviations for all values of ρ. Table II shows sample output of ρ's for ranges of smoothing span N and polynomial degree n. Since the output average is 0.57 ± 0.07 when the input is 0.5, and 0.17 ± 0.05 when the input is 0.2, it appears that the problem is in the variability of the output.

Analyzing data from missile versus drone missions displays more variability of the output. Table III shows the results of varying smoothing span and/or polynomial degree on missile position data. The resulting output ρ varies in an unsystematic manner. A further lack of uniformity is shown in Table IV. The drone, which the missile of Table III was attacking, was airborne for sufficient time to analyze eight successive segments of 256 data points. The variation in output ρ between segments is evident; but again there is no evident system of variation. A final illustration of the non-uniformity of the output ρ is shown in Table V. The Cartesian coordinates of Table IV were calculated from azimuths and elevations measured with several cinetheodolites. Table V shows the averages and variance estimates of five elevation output ρ's from one cinetheodolite.

6. POSSIBLE CAUSES OF DIFFICULTIES. The panel at the Twenty-Second Conference on the Design of Experiments in Army Research, Development, and Testing made some comments on this problem.

First, the use of polynomials was seriously questioned. The fluxuation in calculated ρ should not occur if the mathematical model fits the physical situation. Since the form of the equation for missile trajectories is unknown except in idealized circumstances, a parameter free approach was suggested.

Second, if polynomials must be used to compare with current correlation-ignored results using quadratics, it was suggested that the sum of squares of deviations should not be minimized; instead of deviation, the deviation divided by the square root of a previous estimate of the variance should be used. This procedure, which would change both the position estimates and its variance estimates, should be iterated until the position estimates stablize.

Third, since the path of an object depends on previous position, velocity, and acceleration of the object and not on future values, it was suggested that estimates of position and variance should be determined from the forward time end of the smoothing span, instead of its midpoint, estimate position and variance.

7. ACKNOWLEDGEMENT. In addition to the members of the panel at the Twenty-Second Conference on the Design of Experiments, several personnel at White Sands Missile Range have contributed to this project. A special acknowledgement is due to Elton P. Avara, Atmospheric Sciences Laboratory, Meteorological Satellite Tech Area, WSMR.

TABLE I

INPUT			OUTPUT USING n=0 AND N=17		
			VARIANCE ESTIMATES		
MARKOV ρ OF GENERATOR	VARIANCE OF GENERATOR	RMS OF GENERATOR OUTPUT	MOVING ARC "RAW VARIANCE" CORRECTED FOR DEGREES OF FREEDOM	$\rho - \sigma^2$ ROUTINE	ρ ESTIMATE
.0	1	1.03	1.14	1.13	.06
.1	1	.97	.96	.99	.12
.2	1	.98	.91	.86	.11
.3	1	.94	.85	.86	.27
.4	1	1.02	.87	.90	.37
.5	1	1.07	1.12	1.15	.61
.6	1	.98	1.01	1.12	.66
.7	1	1.01	.92	1.05	.77
.8	1	.94	.67	.73	.70
.9	1	.84	.43	.47	.79

TABLE II

SMOOTHING SPAN	OUTPUT ρ n = 0 or 1	n = 2 or 3	n = 4 or 5	INPUT ρ
17	.61	.62	.51	.5
19	.64	.58	.46	.5
21	.67	.54	.50	.5
17	.11	.20	.25	.2
19	.11	.19	.18	.2
21	.13	.14	.21	.2

TABLE III
MISSILE ANALYSIS

SMOOTHING SPAN	OUTPUT ρ n = 0 or 1	n = 2 or 3	n = 4 or 5
7	.02	.20	.98
11	.00	.09	.30
15	.10	.04	.09
21	.35	.05	.07
31	.58	.10	.07

TABLE IV
DRONE ANALYSIS
OUTPUT FOR n = 2 or 3 AND SMOOTHING SPAN = 15

SEGMENT	X	Y	Z	AVERAGE
1	.38	.80	.59	.59
2	.46	.34	.48	.43
3	.57	.31	.59	.49
4	.89	.69	.50	.69
5	.86	.80	.68	.78
6	1.00	.43	.52	.65
7	1.00	1.00	1.00	1.00
8	-.05	-.06	.25	.05
AVERAGE	.64	.54	.58	.59

TABLE V

OUTPUT ρ FOR ELEVATION OF DRONE FROM
FIVE SUCCESSIVE SEGMENTS ON ONE FILM

n = 0 or 1	n = 2 or 3	n = 4 or 5
.88 ± .27	.34 ± .39	-.08 ± .82

ROBUST OUTLIER DETECTION IN TRAJECTORY DATA REDUCTION

William S. Agee and Robert H. Turner
Analysis and Computation Division
National Range Operations Directorate
US Army White Sands Missile Range
White Sands Missile Range, NM 88002

ABSTRACT. A data reduction program at White Sands Missile Range that often has an hour of flight time is called the Multiple Radar Tracking System (MRTS). Undetected outliers destroy automated data reduction causing a significant number of reruns with human detection of these outliers. The procedure described in this paper enables the MRTS to reduce large quantities of radar data with very little chance of being influenced or ruined by outliers.

Outliers are detected by examining residuals from a least squares estimation. Three robust methods of estimation which are insensitive to outliers are described. The masking effect is almost nonexistent in these methods.

1. INTRODUCTION. An entire trajectory of Cartesian position velocity and acceleration data is produced from radar (range, azimuth, and elevation) data by the Multiple Radar Trajectory System (MRTS). The MRTS consists of four distinct areas:

a. Data gathered from several sources are merged onto one file after being calibrated and time corrected.

b. A preprocessor eliminates outliers and computes initial observation variances and initial X, Y, Z positions. The robust outlier detector is in this stage.

c. A batch processor produces the entire trajectory simultaneously from all observations (except outliers).

d. A fixed lag optimal smoother then produces smoothed positions, velocities and accelerations.

The remainder of this paper is about the preprocessor stage. As the program is at present, whenever outliers are found they are discarded instead of being deweighted.

In order to detect outliers an examination of residuals should be made. But these residuals must not come from an estimation of the observation process that is influenced by the outliers. Three estimation schemes are described which are resistant to outliers. Two methods of examining the residuals for outlying observations are described. The use

of the outlier resistant estimation and residual examination make up the robust outlier detector used in the preprocessor stage of the MRTS.

2. OUTLIER RESISTANT ESTIMATION. The observation model is

$$x_i = a_0 + a_1 t_i + a_2 t_i^2 + e_i, \quad i = 1, n$$

The three methods described are called:

a. Least squares with robust weights,

b. Brown-Mood, and

c. Theil-Sen.

The first one is used in the MRTS.

Least Squares with Robust Weights. The median of the observations x^* and its respective time t^* are found. For each observation compute

$$d_i = \frac{(x_i - x^*)^2}{|t_i - t^*|}$$

Solve for the

$$\hat{A}^T = (\hat{a}_0, \hat{a}_1, \hat{a}_2)$$

by minimizing

$$\sum_{i=1}^{n} w_i \, (x_i - AT_i)^2$$

where

$$w_i = \frac{\sum_{j=1}^{n} d_j}{d_i}$$

Brown-Mood. The following steps show the iterative process for slope and curvature coefficients:

a. Initialize

$$\hat{a}_1^{(0)} = \hat{a}_2^{(0)} = 0$$

$$t^* = \text{med}\ (t_i)$$

$$t^+ = \text{med}\ (t_i > t^*)$$

$$t^- = \text{med}\ (t_i \leq t^*)$$

$$j = 0$$

b. Find median residual in each half

$$x^+ = \underset{t_i > t^*}{\text{med}}\ (x_i - \hat{a}_1^{(j)} t_i - \hat{a}_2^{(j)} t_i^2)$$

$$x^- = \underset{t_i \leq t^*}{\text{med}}\ (x_i - \hat{a}_1^{(j)} t_i - \hat{a}_2^{(j)} t_i^2)$$

c. Update coefficients

$$\hat{a}_1^{(j+1)} = \hat{a}_1^{(j)} + \frac{1}{2}\left[\frac{x^+ - x^-}{t^+ - t^-}\right]$$

$$\hat{a}_2^{(j+1)} = \hat{a}_2^{(j)} + \frac{1}{2}\left[\frac{x^+ - x^-}{(t^+)^2 - (t^-)^2}\right]$$

The relaxation factor of 1/2 seems to provide faster and more stable convergence.

d. Repeat steps two and three until convergence, then compute the intercept coefficient

$$\hat{a}_0 = \text{med}\ (x_i - \hat{a}_1^{(j+1)} t_i - \hat{a}_2^{(j+1)} t_i^2)$$

Theil-Sen. This method is not iterative but it does require many divided differences be taken. First, all the divided differences $d_{j,i}$ without duplication

$$d_{j,i} = \frac{x_j - x_i}{t_j - t_i} \quad j > i$$

To compute all possible divided differences of the $d_{j,i}$ would take too much time and space. Instead a smaller number of divided differences which represent the $d_{j,i}$ well is computed

$$e(i, i+\ell, i+2\ell) = \frac{d(i+2\ell, i+\ell) - d(i+\ell, i)}{t_{i+2\ell} - t_i}$$

for

$$i = 1,\ n-2\ell$$

$$\ell = 1,\ [n/3]$$

Let

$$\hat{a}_2 = \text{med}\ (e(i, i+\ell, i+2\ell))$$

Now

$$d_{j,i} = \frac{x_j - x_i}{t_j - t_i} = a_1 + a_2 (t_j + t_i)$$

since

$$x_i = a_0 + a_1 t_i + a_2 t_i^2$$

Let

$$\hat{a}_1 = \text{med}\ (d_{j,i} - a_2 (t_i + t_j))$$

and finally

$$\hat{a}_0 = \text{med}\,(x_i - \hat{a}_1 t_i - \hat{a}_2 t_i^2)$$

3. OUTLIER DETECTION. A Grubb's-type statistic proposed by Teitjen and Moore [2] is described. A modified version of this statistic is used in the MRTS.

Grubbs Type Statistic. All residuals are ordered by absolute values. We make a change of variable names so that the r's correspond to the observations

$$|z_1| \leq |z_2| \leq \cdots \leq |z_n|$$

$$z_1 = r_{j(1)}, \cdots, z_n = r_{j(n)}$$

After finding the largest gap

$$(|z_{n-k+1}| - |z_{n-k}|)$$

compute the test statistic

$$E_k(n) = \frac{\sum_{i=1}^{n-k} (z_i - \bar{z}_k)^2}{\sum_{i=1}^{n} (z_i - \bar{z})^2}$$

where

$$\bar{z}_k = \frac{\sum_{i=1}^{n-k} z_i}{n-k}$$

and

$$\bar{z} = \frac{\sum_{i=1}^{n} z_i}{n}$$

If $E_k(n)$ is smaller than the desired critical value, we conclude that these k most extreme residuals correspond to outlying observations.

Modified Grubbs Type Statistics. This is the same as previously described except for the denominator of the test statistic and the critical value selected. Instead of testing for k outliers in n samples, we test for one outlier in n-k+1 samples. We compute

$$E_1(n-k+1) = \frac{\sum_{i=1}^{n-k} (z_i - \bar{z}_k)^2}{\sum_{i=1}^{n-k+1} (z_i - \bar{z}_{k+1})^2}$$

where

$$\bar{z}_k = \frac{\sum_{i=1}^{n-k} z_k}{n-k}$$

and

$$\bar{z}_{k+1} = \frac{\sum_{i=1}^{n-k+1} z_k}{n-k+1}$$

If $E_1(n-k+1)$ is smaller than the desired critical value, we conclude that the k most extreme of the n residuals correspond to outlying observations.

3. EXAMPLES. The three previously described estimation procedures and an unweighted least squares were applied to four sets of real data. The original sets of data and residuals from each estimation are listed.

Example 1 - a set of 16 observations where the last two are outliers (residuals$\times 10^5$):

	OBSERVATIONS	LEAST SQUARES W/ROBUST WTS	BROWN-MOOD	THEIL-SEN	LEAST SQUARES
1.	-.0051	-658	-74	-11	-3829
2.	-.0048	-281	-47	-6	-6479
3.	-.0044	24	-22	0	-7532
4.	-.0041	257	0	6	-6991

	OBSERVATIONS	LEAST SQUARES W/ROBUST WTS	BROWN-MOOD	THEIL-SEN	LEAST SQUARES
5.	-.0037	418	20	11	-4854
6.	-.0033	557	87	67	-1071
7.	-.0033	524	51	22	4207
8.	-.0027	419	13	-22	11081
9.	-.0023	342	72	33	19650
10.	-.0021	-6	-71	-111	29615
11.	-.0017	-227	-17	-56	41375
12.	-.0013	-520	34	0	54731
13.	-.0010	-985	-17	-44	69583
14.	-.0006	-1421	29	11	86129
15.	-.9590	-960730	-958727	-958733	-854528
16.	.4451	442389	445014	445022	568910

Example 2 - a set of 15 observations where the third, fourth and fifth are outliers (residuals×10^6):

	OBSERVATIONS	LEAST SQUARES W/ROBUST WTS	BROWN-MOOD	THEIL-SEN	LEAST SQUARES
1.	.21709	-1611	-444	-135	-332222
2.	.21824	-1497	-313	-87	-314194
3.	.95519	734413	735591	735744	441640
4.	.94511	723287	724437	724529	452449
5.	.93499	712116	713215	713256	465224
6.	.22288	-1051	-24	-23	-221986
7.	.22405	-943	-11	-39	-193910
8.	.22530	-760	54	9	-163748
9.	.22652	-612	61	9	-131611
10.	.22770	-510	0	-47	-97508
11.	.22900	-293	32	0	-61280
12.	.23028	-101	15	10	-23066
13.	.23155	75	-39	7	17144
14.	.23286	286	-81	0	59399
15.	.23418	502	-140	0	103670

Example 3 - a set of 15 observations where the twelfth, thirteenth, and fifteenth are outliers (residuals×10^6):

	OBSERVATIONS	LEAST SQUARES W/ROBUST WTS	BROWN-MOOD	THEIL-SEN	LEAST SQUARES
1.	-1.70987	-3359	-599	9	-157774
2.	-1.70942	-867	387	0	-204
3.	-1.70893	991	225	12	105480

	OBSERVATIONS	LEAST SQUARES W/ROBUST WTS	BROWN-MOOD	THEIL-SEN	LEAST SQUARES
4.	-1.70845	2166	61	-5	159227
5.	-1.70793	2708	-54	0	161087
6.	-1.70741	2576	-159	-14	111021
7.	-1.70682	1841	-186	23	9099
8.	-1.70626	402	-233	12	-144780
9.	-1.70571	-1721	-282	-28	-350595
10.	-1.70510	-4456	-262	-28	-608277
11.	-1.70449	-7866	-232	-45	-917885
12.	1.43777	3129701	3141456	3141568	1862231
13.	1.44602	3132585	3149144	3149153	1456410
14.	-1.70257	-22044	0	-121	-2158177
15.	1.44667	3120482	3148695	3148416	473139

Example 4 - a set of 21 observations where the seventh, twentieth, and twenty-first are outliers. This example illustrates dropped sign bits and zeroed data (residuals$\times 10^6$):

	OBSERVATIONS	LEAST SQUARES W/ROBUST WTS	BROWN-MOOD	THEIL-SEN	LEAST SQUARES
1.	-.00988	-123	-423	-248	-4433
2.	-.00995	-88	-337	-178	-2839
3.	-.00976	212	9	154	-1203
4.	-.01017	-83	-243	-114	-386
5.	-.01016	47	-75	39	632
6.	-.01023	102	14	112	1351
7.	.0	10463	10404	10487	12152
8.	-.01047	129	95	162	2034
9.	-.01083	-90	-103	-52	1807
10.	-.01089	-4	0	35	1661
11.	-.01089	147	164	182	1356
12.	-.01121	-16	8	11	513
13.	-.01143	-75	-46	-60	-449
14.	-.01152	2	30	0	-1500
15.	-.01185	-156	-133	-179	-3010
16.	-.01200	-128	-114	-178	-4558
17.	-.01206	-6	-5	-85	-6236
18.	-.01241	-168	-185	-282	-8422
19.	-.01239	45	5	-108	10457
20.	.01215	24783	24717	24587	11810
21.	.01301	25846	25750	25603	10177

4. CONCLUSIONS

a. Least Squares with Robust Weights:

(1) Almost always can produce residuals which reveal up to half the sample to be outliers,

(2) Is the fastest of the three estimators described, and

(3) May be improved with other choices for weights and iteration.

b. Brown-Mood Estimator:

(1) Has unknown convergence properties and

(2) May not work if too many outliers are in one half.

c. Theil-Sen Estimator:

(1) Has robust coefficient estimates,

(2) Is slowest and simplest of the three estimators described, and

(3) May be made more efficient by taking advantage of equally spaced data and for other schemes of selecting divided differences.

d. Grubbs' Type Statistic:

(1) Has no masking effect,

(2) Is fast and easy to use, and

(3) Could use 2d difference criteria to determine which k residuals to be tested.

e. Modified Grubbs-Type Statistic:

(1) Simplifies table look-up and

(2) Detects same outliers as the Grubbs'-type statistic on all samples tried so far.

5. REFERENCES

[1] Mood, A. M., "Introduction to the Theory of Statistics", First Edition 1950, McGraw-Hill, pp 408-410.

[2] Theil, H., "A Rank Invariant Method of Linear and Polynomial Regression Analysis", Indag Math 1950, Vol 12, p 85-91, 173-177, 467-482.

[3] Tietjen, G. L. and R. H. Moore, "Some Grubbs-Type Statistics for the Detection of Several Outliers", Technometrics Aug 1972, Vol 14 Nr 3, pp 583-597.

TABLE LOOK-UP AND INTERPOLATION FOR A NORMAL RANDOM NUMBER GENERATOR

William L. Shepherd and John N. Hynes
Systems Management Division
Instrumentation Directorate
US Army White Sands Missile Range
White Sands Missile Range, New Mexico 88002

ABSTRACT. A normal random number generator using table look-up and interpolation for the inverse normal distribution function is presented and compared to one where the inverse function is computed from a commonly used formula.

1. INTRODUCTION. In Monte Carlo problems and in simulations of noisy measurements, the cost effectiveness of the required normal pseudo-random number generators is still of some economic importance. We present and compare two such generators. One of them is available on the Univac 1108 computer at White Sands Missile Range (WSMR); the other is the main subject of this report.

2. INVERSE DISTRIBUTION FUNCTION METHOD. Let

$$P(x) = \frac{1}{2} + \int_0^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \qquad (2.1)$$

and $\{y\}$ be the output (sequence) generated by a uniform random number generator with density function equal to 1 over the interval $[0, 1]$ and 0 elsewhere. Then $\{P^{-1}(y)\}$ can be thought of as the output of an $n(0, 1)$ random number generator [1, p. 950]. As mentioned in [1], the principal difficulty in using this principle is in the computation of $P^{-1}(y)$. In one of the normal random number generators in use at WSMR, $P^{-1}(y)$ is computed by the formulas

$$P^{-1}(y) = -P^{-1}(1 - y) \qquad \text{for } \frac{1}{2} < y < 1 \quad , \qquad (2.2)$$

$$= \eta - \frac{a_0 + a_1\eta + a_2\eta^2}{1 + b_1\eta + b_2\eta^2 + b_3\eta^3} \qquad \text{for } 0 < y \leq \frac{1}{2} \quad ,$$

where

$$\eta = \sqrt{-2\ln y}$$

$$a_0 = 2.515517 \qquad b_1 = 1.432788$$

$$a_1 = .802853 \qquad b_2 = .189269$$

$$a_2 = .010328 \qquad b_3 = .001038 \ ,$$

with error less than 4.5×10^{-4}. (This formula is also given in [1] and [2].) We refer to this generator as Generator A.

In the following sections, we describe another approximation to $P^{-1}(x)$, referred to as Generator B.

3. A SPLINE APPROXIMATION TO $P^{-1}(y)$. First, consider

$$g(y) = g(a) + g'(a)(y - a) + \beta(y - a)^2 \quad \text{for } a \le y \le \frac{a + b}{2}$$

$$= g(b) + g'(b)(y - b) + \gamma(y - b)^2 \quad \text{for } \frac{a + b}{2} < y \le b \ . \qquad (3.1)$$

Set $h = b - a$, $\bar{y} = \frac{a + b}{2}$,

$$\beta = \frac{2}{h^2}(g(b) - g(a)) - \frac{1}{2h}(3g'(a) + g'(b)) \ , \qquad (3.2)$$

$$\gamma = -\frac{2}{h^2}(g(b) - g(a)) + \frac{1}{2h}(3g'(b) + g'(a)) \ . \qquad (3.3)$$

With some laborious manipulation, it can be verified that

$$g(\bar{y}-) = g(\bar{y}+) \ , \quad g'(\bar{y}-) = g'(\bar{y}+) \ . \qquad (3.4)$$

$g(y)$ is a quadratic spline, with knots $\{a, \bar{y}, b\}$, on $[a, b]$, which interpolates locally between $(a, g(a))$ and $(b, g(b))$.

Now, consider the 2N+1 knots

$$t_0 = \frac{1}{2} < t_2 < t_4 \ldots < t_{2N} < 1 \quad ,$$

$$t_{2i+1} = \frac{1}{2}(t_{2i} + t_{2i+2}), \quad i = 0, 1, \ldots, N-1 \quad ,$$

and the splines

$$g_i(y) = P^{-1}(t_{2i}) + P^{-1,}(t_{2i})(y - t_{2i}) + \beta_i(y - t_{2i})^2 \quad , \quad t_{2i} \le y < t_{2i+1}$$

$$= P^{-1}(t_{2i+2}) + P^{-1,}(t_{2i+2})(y - t_{2i+2}) + \gamma_i(y - t_{2i+2})^2 \quad , \qquad (3.5)$$

$$t_{2i+1} \le y < t_{2i+2} \quad ,$$

where

$$\beta_i = \frac{2}{h_i^2}(P^{-1}(t_{2i+2}) - P^{-1}(t_{2i})) - \frac{1}{2h_i}(3P^{-1,}(t_{2i}) + P^{-1,}(t_{2i+2})) \quad , \qquad (3.6)$$

$$\gamma_i = -\frac{2}{h_i^2}(P^{-1}(t_{2i+2}) - P^{-1}(t_{2i})) + \frac{1}{2h_i}(3P^{-1,}(t_{2i+2}) + P^{-1,}(t_{2i})) \quad , \qquad (3.7)$$

$$h_i = t_{2i+2} - t_{2i} \quad . \qquad (3.8)$$

Define

$$g(t) = g_i(t) \qquad t_{2i} \le t < t_{2i+2} \quad , \quad i = 0, 1, \ldots, N-1 \quad . \qquad (3.9)$$

From (3.1), (3.2), and (3.3), with a, b, g(a), g(b), g'(a), and g'(b) replaced by t_{2i}, t_{2i+2}, $P^{-1}(t_{2i})$, $P^{-1}(t_{2i+2})$, $P^{-1,}(t_{2i})$, $P^{-1,}(t_{2i+2})$, g(t) is a spline agreeing, in function and first derivative, with $P^{-1}(t)$ on $\{t_{2i}\}_{i=0}^{N}$. (g(t) also has the knots $\{t_{2i+1}\}_{i=0}^{N-1}$.) g(t) interpolates the table

$$\{t_{2i}, P^{-1}(t_{2i}), P^{-1,}(t_{2i})\}_{i=0}^{N} \quad . \qquad (3.10)$$

In order to use (3.9), (3.10) for a normal random number generator, we need a suitable table and a means for computing $P(t)$ for $t_{2N} < t < 1$.

Define

$$||g - P^{-1}||_{(a,b)} = \max\{g(t) - P^{-1}(t) | t \in [a,b]\} \quad , \quad b < 1 \quad .$$

We need a t_{2N} close to 1 and a sequence $\{t_{2i}\}_{i=0}^{N}$ with a small N so that $||g - P^{-1}||_{(1/2,t_{2N})} < \varepsilon$ for a prescribed tolerance ε. We used numerical search not described in detail here. It took up much computer time and is not optimal. Essentially, we started with $t_0 = \frac{1}{2}$, computed t_2 so that

$$||g - P^{-1}||_{(t_0,t)} \leq \varepsilon \quad , \qquad t \leq t_2$$

$$||g - P^{-1}||_{(t_0,t)} > \varepsilon \quad , \qquad t > t_2 \quad ,$$

then recursively determined t_{2i+2} so that

$$||g - P^{-1}||_{(t_{2i},t)} \leq \varepsilon \quad , \quad t_{2i} \leq t < t_{2i+2}$$

$$> \varepsilon \quad , \quad t > t_{2i+2} \quad .$$

t_{2N} was determined empirically by stopping when $t_{2N} - t_{2N-2}$ was less than a prescribed tolerance δ.

For the computation of $P^{-1}(y)$ in the above, we used Newton's method for solving

$$P(y) - x = 0$$

for x, with

$$P(y) = \frac{1}{2}\,(1 + \mathrm{erf}(y/\sqrt{2})) \quad ,$$

$$\mathrm{erf}\, y = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{y^{n+1}}{n!(2n+1)}$$

from [1], and

$$P^{-1\prime}(y) = \frac{1}{P'(P^{-1}(y))} \quad ,$$

$$P'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad .$$

Our numerical experience indicated that near one the knot spacing needed to obtain the required accuracy is not feasible, as N is large. We now present two methods for computing $P^{-1}(t)$ for $t_{2N} < t < 1$.

The first method is to use the approximation (2.2) from Generator A.

The second method is to approximate $P^{-1}(x)$ by a quadratic spline with knots $\{t_{2N}, \frac{1}{2}(1 + t_{2N}), 1\}$ which has the same area under the curve as does $P^{-1}(t)$ over each of the intervals $[t_{2N}, \frac{1}{2}(t_{2N} + 1)]$, and $[\frac{1}{2}(t_{2N} + 1), 1]$. (A discussion is given in Appendix A.)

Table 2, 3 gives the requisite coefficients for N = 29, N = 89, respectively. LEFT INTERVAL, RIGHT INTERVAL refer to $[t_{2i}, t_{2i+1}]$, $[t_{2i+1}, t_{2i+2}]$, respectively. The last row shows 1 to be a knot. The entries in this row were obtained according to the equal area criterion and would not be used in a computer program where a rational approximation is used for $t_{2N} \leq t < 1$.

EXAMPLE FOR TABLE 2:

For i = 13, $.914502271 < t \leq .920905352$

and

$$g(t) = 1.41118763 + 6.78465741(t - .920905352) + 30.8378448(t - .920905352)^2 \quad .$$

4. NUMERICAL RESULTS. The Generator B was run under 4 separate conditions as indicated in Table 1. The interpolation tolerance for N = 29 is 10^{-4} and for N = 89 it is 10^{-6}. Results for Generator A are also includes in Table 1.

TABLE 1. AVERAGE RUN TIMES (CPU) IN SECONDS WITH IDENTICAL INPUT OF 10,000 POINTS

GENERATOR A	2.427272
GENERATOR B	
For N = 29 (with rational function approximation at the end)	1.307336
For N = 29 (with spline approximation at the end)	1.111576
For N = 89 (with rational function approximation at the end)	1.503344
For N = 89 (with spline approximation at the end)	1.276524

5. CONCLUSIONS AND ADDITIONAL RESEARCH NEEDED. From Table 1, the following empirical inferences can be made.

a. Generator B with N = 29 and with either end option is slightly more accurate, and about twice as fast as Generator A. It requires 186 stored constants.

b. As compared to Table 2, Table 3 provides for interpolation over a larger interval, is a little slower, provides six significant digit interpolation accuracy but requires 543 stored constants.

Additional research could be done in the approximations at the end. (2.2) is not necessarily optimal for $t_{2N} \leq t < 1$.

The constants for Generator B are believed to be of nine significant digit accuracy. It is possible that they do not have to be this accurate. Further research could address this problem.

Since computation for N = 89 is only a little slower than for N = 29, but interpolates much more accurately, we think more of the CPU time is used in the interpolation than in the table look-up logic. As higher order interpolation is slower than quadratic, there is not much advantage in using it in order to reduce the required number of knots.

REFERENCES

[1] Abramowitz, Milton and Irene A. Stegun, "Handbook of Mathematical Functions," National Bureau of Standards AMS 55, 1964.

[2] Hastings, C., Jr., "Approximations for Digital Computers," Princeton University Press, Princeton, N. J., 1955.

APPENDIX A. In this Appendix, we outline the procedure for extending $g(t)$ over the interval $[t_{2N}, 1)$. Principally, we want to exhibit (A.3) and (A.4), omitting most of the detail.

First, we want β so that, with

$$g_N(t) \triangleq P^{-1}(t_{2N}) + P^{-1\prime}(t_{2N})(t - t_{2N}) + \beta(t - t_{2N})^2 \quad , \tag{A.1}$$

for $t_{2N} \leq t \leq \frac{1}{2}(t_{2N} + 1) = \bar{t}$, we have

$$\int_{t_{2N}}^{\bar{t}} g_N(t)dt = \int_{t_{2N}}^{\bar{t}} P^{-1}(t)dt \quad . \tag{A.2}$$

By the change of variables $t = P(x)$, we have

$$\int_{t_{2N}}^{\bar{t}} P^{-1}(t)dt = \int_{P^{-1}(t_{2N})}^{P^{-1}(\bar{t})} x\, P'(x)dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{P^{-1}(t_{2N})}^{P^{-1}(\bar{t})} x \exp(-x^2/2)dx$$

$$= \frac{1}{\sqrt{2\pi}} \left(\exp[-(P^{-1}(t_{2N}))^2] - \exp[-(P^{-1}(\bar{t}))^2]\right) \quad . \tag{A.3}$$

Similarly,

$$\int_{\bar{t}}^{1} P^{-1}(t)dt = \lim_{n \to 1} \int_{\bar{t}}^{n} P^{-1}(t)dt = \frac{1}{\sqrt{2\pi}} \exp[-(P^{-1}(\bar{t}))^2] \quad . \tag{A.4}$$

(A.3), (A.1), and (A.2) yield

$$\beta = \frac{-3}{(\bar{t} - t_{2N})^3}(P^{-1}(t_{2N})(\bar{t} - t_{2N}) + \frac{1}{2} P^{-1\prime}(t_{2N})(\bar{t} - t_{2N})^2$$

$$- \frac{1}{\sqrt{2\pi}} \left(\exp[-(P^{-1}(t_{2N}))^2] - \exp[-(P^{-1}(\bar{t}))^2]\right) \quad .$$

For $\bar{t} \le t \le 1$, we proceed similarly to define

$$g_N(t) = g_N(\bar{t}) + g_N'(\bar{t})(t - \bar{t}) + \gamma(t - \bar{t})^2 \quad ,$$

$$\gamma = \frac{-3}{(1 - \bar{t})^3}\left(g_N(\bar{t})(1 - \bar{t}) + \frac{1}{2} g_N'(\bar{t})(1 - \bar{t})^2 - \frac{1}{\sqrt{2\pi}} \exp[-(P^{-1}(\bar{t}))^2]\right) \quad .$$

With β and γ so determined,

$$g(t) = g_N(t) \quad , \quad t_{2N} \le t \le 1 \quad .$$

It should be noted that $g_N(t)$ is not, strictly, an interpolation function.

TABLE 2

QUADRATIC SPLINE COEFFICIENTS FOR .0001 TOLERANCE

i	t_{2i}	t_{2i-1}	FUNCTION VALUE	1ST DER.	(2ND DER.)/2 LEFT INTERVAL	(2ND DER.)/2 RIGHT INTERVAL
0	.500000000	*********	.000000000	2.50662827	.075764656	.388034412
1	.558171995	.529085997	.146336174	2.53361129	.545723915	.880786896
2	.612423226	.585297611	.285640412	2.61100122	1.05778217	1.44656276
3	.661641926	.637032576	.416948403	2.73426188	1.65836525	2.13361930
4	.705556743	.683599334	.540450319	2.90078621	2.39739418	2.99693870
5	.744341619	.724949181	.656788745	3.11000477	3.33320045	4.10281944
6	.778349951	.761345785	.966632414	3.36289148	4.53770637	5.53949546
7	.808126397	.793238174	.871012723	3.66295469	6.10705566	7.41560363
8	.834035993	.821081195	.970237718	4.01332127	8.16024017	9.88627243
9	.856645040	.845340515	1.06536692	4.42133582	10.8696708	13.1501464
10	.876292996	.866469018	1.15665337	4.89327612	14.4512634	17.4731445
11	.893344581	.884818788	1.24451295	5.43763796	19.1979064	23.2031402
12	.908099189	.900721885	1.32914050	6.06324880	25.4950866	30.8378448
13	.920905352	.914502271	1.41118763	6.78465741	33.8953704	41.0198211
14	.931977972	.926441662	1.49068561	7.61416498	45.1018066	54.6292266
15	.941551745	.936764858	1.56793411	8.56896728	60.0864257	72.8101501
16	.949798711	.945675228	1.64290506	9.66496074	80.1147766	97.2113037
17	.956927851	.953363281	1.71609695	10.9291433	107.027099	129.989746
18	.963070221	.959999036	1.78748238	12.3849888	143.155273	173.870666
19	.968333371	.965701796	1.85684465	14.0535436	191.572021	233.241760
20	.972895302	.970614336	1.92515959	15.9915148	257.207153	313.360229
21	.976808459	.974851881	1.99189042	18.2242343	345.408691	419.432983
22	.980102517	.978455488	2.05587086	20.7436675	461.757202	559.967285
23	.982907608	.981505063	2.11788529	23.6096984	616.023193	745.834716
24	.985296324	.984101966	2.17798233	26.8627900	821.866210	1006.04174
25	.987457596	.986376960	2.24009420	30.8133965	1112.46679	1363.56250
26	.989309937	.988383766	2.30120359	35.3998477	1508.10644	1846.82568
27	.990880787	.990095362	2.36074176	40.6699449	2043.21679	2509.15136
28	.992236637	.991558712	2.41985432	46.8422700	2776.53808	3398.48632
29	.993372664	.992804651	2.47685236	53.8572662	4829.71895	37970.1705
30	1.00000000	.996686332	3.25464832	255.761911	**********	**********

TABLE 3

QUADRATIC SPLINE COEFFICIENTS FOR .000001 TOLERANCE

i	t_{2i}	t_{2i-1}	FUNCTION VALUE	1ST DER.	(2ND DER.)/2 LEFT INTERVAL	(2ND DER.)/2 RIGHT INTERVAL
0	.500000000	********	.000000000	2.50662827	.036046982	.181203842
1	.527511597	.513755798	.069016097	2.51260520	.253838621	.401294708
2	.554553339	.541032468	.137173924	2.53032278	.475940704	.628820419
3	.580875465	.567714402	.204133636	2.55940243	.707033157	.868577957
4	.606299114	.593587290	.269686078	2.59946024	.951988220	1.12551879
5	.630684651	.618491883	.333667207	2.65012136	1.21565246	1.40468025
6	.653944977	.642314814	.395993198	2.71107122	1.50334739	1.71118545
7	.676039376	.664992176	.456651928	2.78209440	1.82032775	2.05090713
8	.696952964	.686496170	.515656885	2.86305588	2.17249679	2.42998504
9	.716695715	.706824340	.573053353	2.95392149	2.56611633	2.85541153
10	.735394360	.725995038	.628904958	3.05475458	3.00863266	3.33451080
11	.752785078	.744039719	.683280136	3.16570065	3.50746154	3.87625885
12	.769208248	.760996663	.736241888	3.28696482	4.07207870	4.49017715
13	.784623388	.776915818	.787903379	3.41895313	4.71248626	5.18707656
14	.799060028	.791841708	.838268462	3.56186948	5.43972778	5.98056793
15	.812598409	.805829218	.887512233	3.71648198	6.26869201	6.88455963
16	.825273288	.818935848	.935649995	3.88319780	7.21306610	7.91645813
17	.837133552	.831203420	.982745125	4.06263803	8.29160308	9.09383392
18	.848223594	.842678573	1.02884437	4.25544315	9.52256012	10.4404678
19	.858593434	.853408514	1.07402131	4.46245657	10.9304428	11.9809799
20	.868287360	.863440397	1.11833187	4.68455825	12.5411071	13.7451629
21	.877344058	.872815709	1.16181189	4.92262519	14.3872680	15.7641296
22	.885802651	.881573354	1.20450436	5.17766358	16.5016784	18.0794219
23	.893701702	.889752176	1.24645719	5.45082132	18.9241790	20.7341918
24	.901068493	.897385098	1.28766382	5.74297625	21.7013854	23.7821044
25	.907950355	.904509424	1.32823862	6.05598746	24.8919067	27.2801818
26	.914368497	.911159426	1.36815678	6.39083559	28.5552368	31.2964172
27	.920354167	.917361332	1.40745786	6.74908770	32.7624511	35.9124016
28	.925936506	.923145336	1.44617905	7.13245471	37.5995178	41.2204589
29	.931142691	.928539598	1.48435556	7.5428056	43.1614074	47.3302917
30	.935996797	.933569744	1.52201068	7.98206208	49.5636901	54.3616333

TABLE 3 (Continued)

i	t_{2i}	t_{2i-1}	FUNCTION VALUE	1ST DER.	(2ND DER.)/2 LEFT INTERVAL	(2ND DER.)/2 RIGHT INTERVAL
31	.940520710	.938258754	1.55915972	8.45221094	56.9315795	62.4606323
32	.944735739	.942628224	1.59582208	8.95545274	65.4213256	71.7829589
33	.948660574	.946698156	1.63200301	9.49395652	75.1904907	82.5385131
34	.952317143	.950488856	1.66774824	10.0707042	86.4736769	94.9212036
35	.955721520	.954019331	1.70305936	10.6882405	99.4630126	109.226562
36	.958890795	.957306157	1.73795689	11.3496354	114.445007	125.739074
37	.961340393	.960165594	1.77245406	12.0580832	131.775878	114.763549
38	.964584757	.963212575	1.80656273	12.8170072	151.735717	166.786010
39	.967138001	.965861379	1.84030141	13.6302717	174.836181	192.179077
40	.969511720	.968324860	1.87366538	14.5014626	201.486328	221.552856
41	.971718537	.970615128	1.90667314	15.4350326	232.297607	255.501708
42	.973770186	.972744362	1.93934266	16.4358269	267.935791	294.767822
43	.975677580	.974723883	1.97169145	17.5091240	309.224609	340.171630
44	.977450859	.976564220	2.00373662	18.6606822	356.886230	392.791503
45	.979095992	.978273426	2.03542611	19.8940017	411.999755	453.737548
46	.980625451	.979860722	2.06684143	21.2181148	476.125976	524.165527
47	.982047370	.981336411	2.09799870	22.6404429	550.188476	606.019042
48	.983366533	.982706952	2.12884685	24.1656689	636.053222	700.946777
49	.984592943	.983979738	2.15946494	25.8053788	735.710937	810.318359
50	.985723539	.98515824I	2.18960465	27.5533131	850.266601	936.277343
51	.986765808	.986244673	2.21926965	29.4153730	982.491210	1083.87255
52	.987747769	.987256788	2.24912629	31.4444682	1137.92187	1254.44921
53	.988653015	.988200392	2.27854762	33.5101517	1316.79980	1451.35644
54	.989487538	.989079276	2.30753660	35.9202470	1524.25488	1681.52832
55	.990272348	.989879943	2.33669011	38.4361673	1766.84277	1949.18457
56	.990995846	.990634097	2.36544725	41.1246983	2046.94531	2258.28613
57	.991662820	.991329333	2.39381051	43.9961846	2371.78710	2620.29687
58	.992289085	.991975952	2.42231843	47.1225577	2753.17968	3040.56445
59	.992866423	.992577754	2.45046574	50.4675066	3195.23046	3525.90234
60	.993398657	.993132540	2.47825468	54.0447095	3705.70703	4087.43554

TABLE 3 (Continued)

i	t_{2i}	t_{2i-1}	FUNCTION VALUE	1ST DER.	(2ND DER.)/2 LEFT INTERVAL	(2ND DER.)/2 RIGHT INTERVAL
61	.993889309	.993643983	2.50568683	57.8684187	4292.57031	4754.28906
62	.994351932	.994120621	2.53340187	62.0537570	4995.87500	5520.42187
63	.994778412	.994565172	2.56079898	65.5387341	5805.76953	6414.12890
64	.995171573	.994974993	2.58788028	71.3431056	6740.19921	7453.58593
65	.995534019	.995352796	2.61464734	76.4876203	7840.09375	8665.03515
66	.995874708	.995704364	2.64163945	82.1105753	9123.69531	10091.2890
67	.996188782	.996031745	2.66835193	88.4155599	10615.1484	11748.8593
68	.996478318	.996333550	2.69478698	94.6207663	12352.6562	13669.1796
69	.996745234	.996611776	2.72094638	101.566437	14373.9453	15894.6093
70	.996991297	.996868266	2.74683144	109.014404	16719.8437	18522.6875
71	.997221982	.997106640	2.77289314	117.144323	19492.0156	21585.5312
72	.997434644	.997328313	2.79871048	125.879954	22720.3125	25148.9375
73	.997630692	.997532668	2.82428547	135.264584	26428.7500	29341.1718
74	.997811424	.997721058	2.84961976	145.344227	30868.4843	34182.0625
75	.997980860	.997896142	2.87515591	156.365979	35983.7187	39976.8437
76	.998137059	.998058959	2.90048281	168.231127	42014.5000	46322.4375
77	.998273733	.998205396	2.92428049	180.304436	48601.7187	53272.0625
78	.998393322	.998333527	2.94655473	192.487446	55635.2500	60854.7812
79	.998497963	.998445643	2.96732080	204.677361	63404.3125	68674.5000
80	.998589524	.998543744	2.98660368	216.770522	71431.3750	77046.1250
81	.998669640	.998629582	3.00443796	228.665964	80480.9375	85049.1250
82	.998739741	.998704691	3.02086735	240.268802	87860.0000	93840.6250
83	.998792317	.998766029	3.03374742	249.822392	94369.6250	101113.500
84	.998831749	.998812033	3.04374878	257.531449	103357.750	103587.000
85	.998861323	.998846536	3.05145475	263.651087	106801.000	109387.250
86	.998883504	.998872413	3.05735550	268.446026	107243.500	116271.500
87	.998900139	.998891821	3.06185198	272.164678	128647.000	100745.000
88	.998912616	.998906377	3.06526543	275.025721	128965.000	104880.000
89	.998921968	.998917292	3.06784780	277.212287	158981.193	127441.48
90	1.00000000	.999460984	3.73782594	1213.45302	**********	**********

EIGENVECTORS ANALYSIS OF EMPIRICAL DATA VERSUS UTILIZATION OF STANDARD FUNCTIONS

Oskar M. Essenwanger
Physical Sciences Directorate
Technology Laboratory
US Army Missile Research and Development Command
Redstone Arsenal, Alabama 35809

ABSTRACT. Parameterization of empirical data (e.g., the wind profiles from surface to 25 Km altitude) in many cases entails the approximation of data by mathematical functions. In general, several options which lead to solutions are available but the question of which is the most suitable form is sometimes difficult to answer.

Often a specific goal of approximating data by mathematical functions is the derivation of one characteristic parameter or variate. Theoretically, eigenvector analysis (or equivalently the development of empirical polynomials) should lead to maximum information by a single parameter.

A comparison between approximations by eigenvectors and standard (orthogonal) functions has been made. It is shown that in particular cases standard functions can achieve equivalent reductions of the variance and they may be simpler and more economical to compute than eigenvector functions.

1. INTRODUCTION. Parametrization of atmospheric data (such as the wind profile as function of the altitude) requires the derivation of suitable mathematical expressions. The availability of high speed electronic data processing tools has opened the door to a utilization of the most sophisticated mathematical tools even for the generally huge collectives of atmospheric data. For example, the calculation of empirical polynomials (or eigenvectors in mathematical terminology) is now possible without too much difficulty for the large dimensions of atmospheric data matrices. Consequently it is very tempting to "grind" huge data collections through the computers without considering how much benefit these highly sophisticated tools render compared with the application of standard functions or simple parameters.

In this article, some light is shed on the utilization of empirical polynomials in comparison with the use of standard functions exemplified by the wind profiles of certain altitude ranges. Under curtain conditions, standard functions can achieve an equivalent reduction of the variance to the one obtained by eigenvector analysis.

2. THE CALCULATION OF EIGENVECTORS. The problem under consideration is the development of proper functions for the wind speed profile V_h where the h is a subscript denoting the altitude. $\overline{V}_h$ designates a mean wind speed profile. The wind direction θ_h can be treated equivalently. We formulate the representation of the wind speed profile:

$$V_{h,i} - \bar{V}_h = B_{1,i}\,\phi_{1,h} + B_{2,i}\,\phi_{2,h} + \ldots B_{n,i}\,\phi_{n,h} \qquad (1)$$

where $i = 1,\ldots, N$, and $n \ll N$. In this equation the coefficients $B_{j,i}$ and the functions $\phi_{j,h}$ must be determined.

The development of optimized characteristic functions $\phi_{j,h}$ is a known problem of matrix analysis. A mathematical formulation is:

$$M_\phi^{-1}\, M_V\, M_\phi = M_\lambda \qquad (2)$$

where M_ϕ designates a matrix of eigenvectors (or polynomials), M_V the data matrix for the wind profile, and M_λ a (diagonal) matrix of eigenvalues. The elements of the (symmetric) data matrix are either the covariances:

$$V_{h,k} = \Sigma\, V_{h,i} \cdot V_{k,i}/N \qquad (3a)$$

the standardized covariances:

$$\bar{V}_{h,k} = \Sigma(V_{h,i} - \bar{V}_h)(V_{k,i} - \bar{V}_k)/N \qquad (3b)$$

or the correlations:

$$r_{h,k} = \bar{V}_{h,k}/\left(\sigma_{V_h} \cdot \sigma_{V_k}\right) . \qquad (3c)$$

A judgement of the effectiveness of the systems can be made by a calculation of the residual or left variance, or the percentage reduction, which can be readily obtained from the eigenvalues λ by:

$$PR_j = \lambda_j^2 \Big/ \sum_{j=1}^{n} \lambda_j^2 . \qquad (4)$$

More details on the mathematical background can be found in the author's text (1976). The covariance and the correlation system has been compared in a recent article by Essenwanger (1975), and will not be repeated here. In this article it is illustrated that the percentage reduction varies largely with the particular system which is selected but the residual variance (error) is of the same magnitude for the same number of terms irrespective of the percentage reduction of the individual system.

3. EIGENVECTORS OF THE WIND PROFILE. First it should be clarified that under the term "wind profile" the structure of the wind velocity in the first 10 m of the atmosphere is not meant. The nomenclature designates the wind speed or direction as a function of the altitude up to about 25 or 30 Km.

The first eigenvectors of the wind direction covariance matrix for the altitude range surface to 24 Km are depicted in Figure 1 for January and July at stations representative of four climatic zones. We learn from inspecting Figure 1 that it would be very difficult to find an adequate standard function to approximate that particular structure of the atmospheric direction profile.

In turn, as displayed in Figure 2, the first eigenvector of the wind speed from surface to 10 Km altitude range lends itself readily for replacement by a standard function. A linear curve fit would adequately replace the eigenvector for three stations, and the fitting of a second order curve may be a successful approximation for Albrook.

The examination of the eigenvectors for the surface to 24 Km wind speed system follows next. Figure 3 discloses that at least for some climatic regimes a standard function such as the Fourier series may be applicable. This fact is supported by scrutinizing Figure 4 which exhibits the wind speed profile for Montgomery. As it is displayed, the major eigenvector comprises over 80% of the variance and resembles a sine wave. Indeed, a Fourier analysis of the first three eigenvectors revealed that at least the first two eigenvectors provide largely one dominant Fourier term. A comparison of the eigenvector and Fourier system appears to be a worthwhile study.

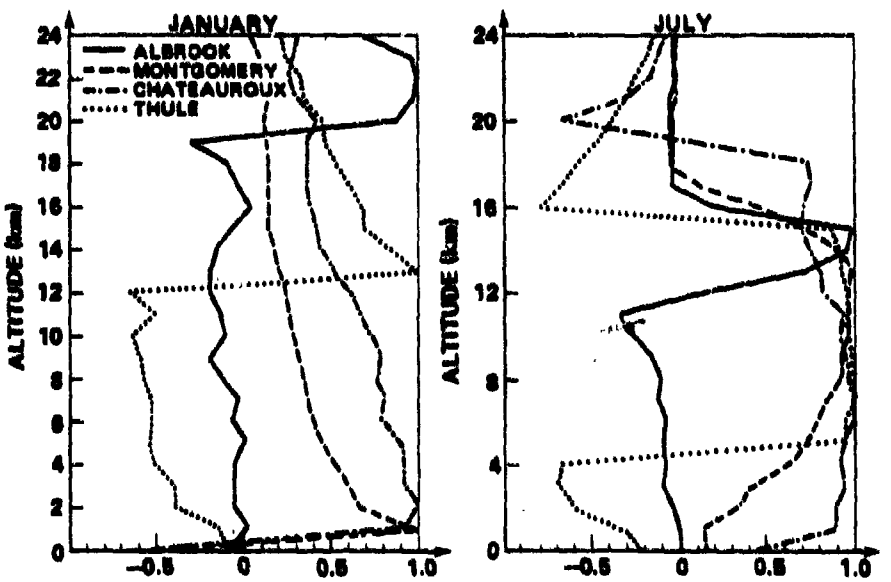


Figure 1. First Eigenvectors, Wind Direction, Covariance.

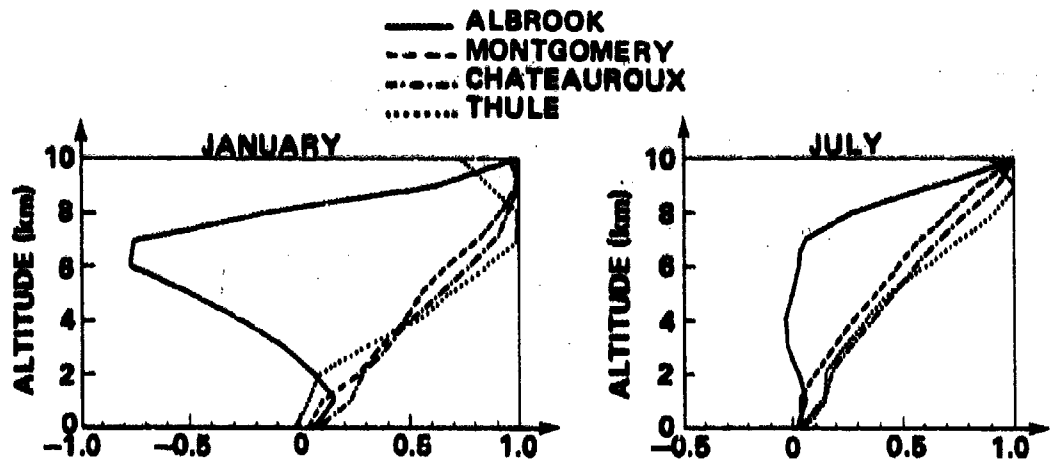


Figure 2. First Eigenvector Structure (Surface to 10 Km).

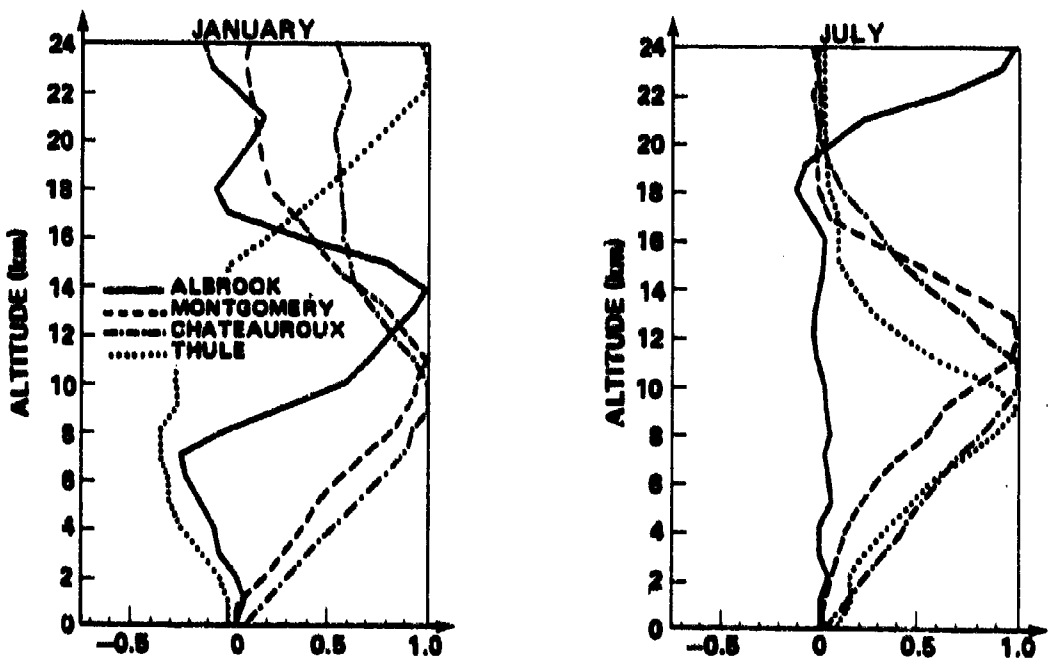


Figure 3. First Eigenvector Structure (Surface to 24 Km).

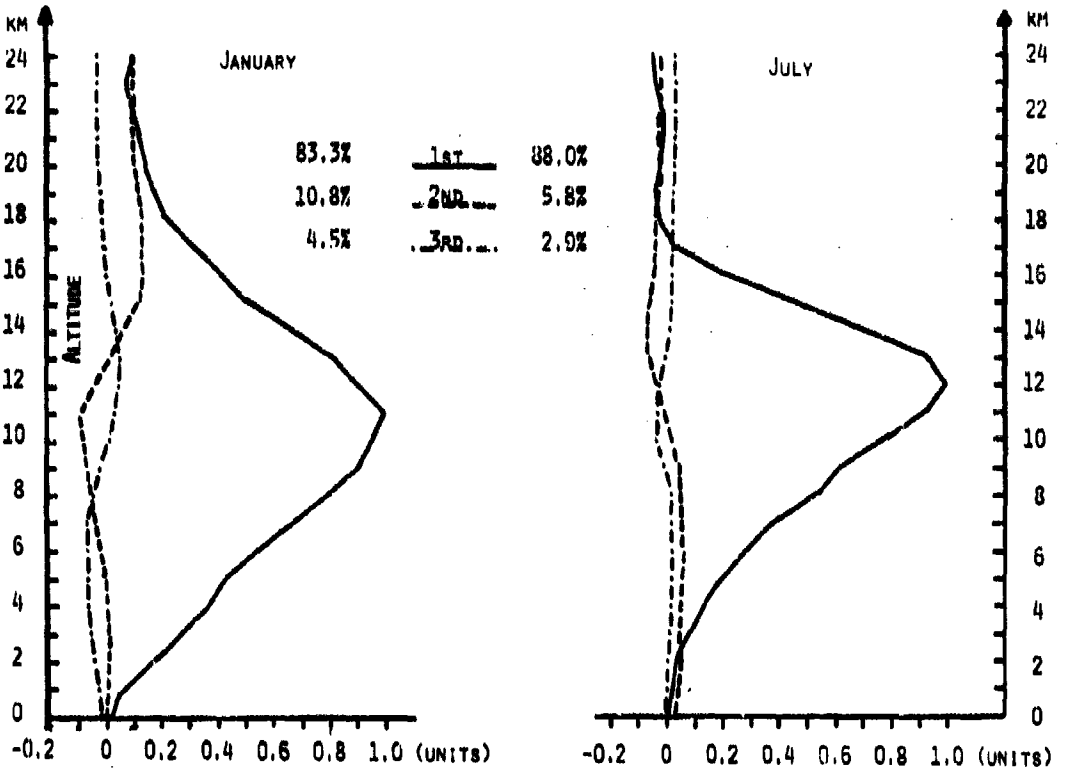


Figure 4. First Three Eigenvectors (Scaled)
Montgomery, Surface - 24 Km, $y = V_h$.

4. STANDARD FUNCTIONS FOR THE WIND PROFILE. While empirical polymomials provide an optimum of information in one single term, standard functions have other advantages. One of them is the homogeneous mathematical background for different collectives, e.g., data from different climatic regimes. This homogeneity is beneficial for a classification of the wind profile into categories (see Essenwanger, 1974). The differences of the percentage reductions between individual order terms at locations from typical climatic regimes are not partially or entirely caused by the diversity of this mathematical background. Because the present goal is the derivation of one characteristic parameter, the homogeneity of the background is of secondary importance here. Of interest, however, is the simplicity or the cost savings associated with the utilization of standard functions.

Table 1 serves as a basis for the examination of the reduction of the variance by individual order terms. Three systems are depicted for the surface to 25 Km wind profile at Montgomery (Alabama):

Table 1. Left Variance of Wind Profile (Surface to 25 Km)
Fourier and Eigenvector System
Montgomery, Alabama

	$V_h - \bar{V}_h$				V_h				Cov	Eigenvector	
	σ_1^2	1 Term	2 Terms	3 Terms	σ_2^2	1 Term	2 Terms	3 Terms	1 Term	2 Terms	3 Terms
Jan	56.41	24.34	15.05	10.65	234.5	31.9	17.7	12.7	44.8	28.9	18.6
Feb	61.26	23.09	13.80	9.52	255.8	29.0	15.4	10.4	38.5	24.7	17.7
Mar	50.77	23.42	15.12	11.64	267.3	31.0	16.6	12.6	35.1	24.3	19.0
Apr	52.37	17.55	11.00	7.96	180.7	23.9	12.2	8.4	26.1	16:9	12.2
May	28.44	11.38	6.51	4.59	66.6	15.1	7.4	4.9	10.5	10.5	7.8
Jun	23.38	9.56	5.38	3.65	33.8	15.2	7.8	4.6	12.0	8.7	6.6
Jul	14.45	7.25	4.41	3.06	25.4	16.8	8.9	5.8	9.8	7.4	5.7
Aug	17.53	8.10	4.75	3.30	30.5	19.9	9.0	5.9	9.6	7.6	5.8
Sep	25.08	10.06	5.49	3.67	41.1	16.8	7.3	4.5	12.5	9.2	7.2
Oct	32.65	11.91	6.69	4.75	93.6	19.0	7.8	5.1	17.0	11.6	8.3
Nov	48.81	18.17	10.89	7.88	151.9	25.6	12.7	9.1	29.6	20.5	13.6
Dec	58.50	21.48	13.86	9.67	182.3	30.6	18.0	12.4	36.7	23.8	16.6

$\text{cov} = (x - \bar{x})(y - \bar{y})/N$

Unit: $(m/sec)^2$

$$V_{h,i} - \bar{V}_h = A_{0,i} + A_{1,i} \sin(\alpha_h + \beta_{1,i}) + A_{2,i} \sin(2\alpha_h + \beta_{2,i}) + \ldots \quad (5a)$$

$$V_{h,i} = C_{0,i} + C_{1,i} \sin(\alpha_h + \varphi_{2,i}) + C_2 \sin(2\alpha_h + \varphi_{2,i}) + \ldots \quad (5b)$$

and the eigenvector system of equation (3b) which had emerged as the system with the smallest residual variance of the three eigenvector systems in a separate study.

We learn that one term of the eigenvector system with coefficient $B_{1,i}$ displays the lowest left variance. It should be noticed that A_0 or C_0 is the first coefficient of the Fourier system, which leads to:

$$\sigma_1^2 = \sum_N^i \sum_h (V_{h,i} - \bar{V}_h - A_{0,i})^2/(h \cdot N) \quad (6a)$$

or:

$$\sigma_2^2 = \sum_N^i \sum_h (V_{h,i} - C_{0,i})^2/(h \cdot N) \; . \quad (6b)$$

Consequently the column for one term of the eigenvector system must be compared with the columns σ_1^2 and σ_2^2. Attention should be called that an assumption:

$$V_{h,i} = \bar{V}_h = A_{0,i} \quad (5c)$$

leads to a residual variance which is quite comparable with the eigenvector system. Although the system requires that the mean wind speed profile $\bar{V}_h$ is known, the prerequisite is identical, however, with the one in the eigenvector system. It is self evident that the calculation of the average value $A_{0,i}$ is a trivial task.

A further reduction of the variance is gained by adding terms in the Fourier or eigenvector series. Because one term of the Fourier system has two parameters which can be fitted the columns should not be compared equivalently according to their headings. The left variance should be compared between one term of the Fourier series and three terms of the eigenvector system. Then the fact that the left variance is lowest for the empirical polynomials agrees with the expectation.

One additional fact deserves attention. If we are interested in a single-variate system, the eigenvector system can only be based on $B_{1,i}$ because the other coefficients $B_{j,1}$, $j \geq 2$, are independent of

$B_{1,i}$. Although the Fourier system is orthogonal the coefficients can be related (see Essenwanger, 1964).

5. SINGLE-PARAMETER FOURIER SYSTEM. Before a single-parameter system other than based on $A_{0,i}$ can be examined let us derive an analytical expression for the replacement of the current coefficients of the Fourier system by an approximation . We cast:

$$V_{h,i} - \bar{V}_h = \left(A_{0,i} + \epsilon_{A_{0,i}}\right) + \left(A_{1,i} + \epsilon_{A_{1,i}}\right) \sin\left(\alpha_h + \beta_{1,i} + \Delta\beta_{1,i}\right) + \ldots \quad (7)$$

By summation over h and omission of the terms which become zero we deduce the following expression for the left variance:

$$Var_L^2 = s_V^2 - A_0^2 + \epsilon_{A_0}^2 - A_1/2 + A_1\left(A_1 + \epsilon_{A_1}\right)(1 - \cos\Delta\beta_1) + \epsilon_{A_1}^2/2 - A_2^2/2 + A_2\left(A_2 + \epsilon_{A_2}\right)(1 - \cos\Delta\beta_2) + \epsilon_{A_2}^2/2 + \ldots \quad (8)$$

(The subscript i denoting the individual observation time has been omitted).

It is easily recognized that for $\epsilon = 0$ and $\Delta\beta = 0$ Eqn. (8) reduces to the well-known formula for the left variance (e.g. see Essenwanger, 1976) because the two terms after $A^2/2$ disappear. It may be reasonable that $A_j^2 > \epsilon_{A_j}^2$ for the dominant Fourier term. For the other term of the series it may not hold, and instead of a net decrease of the variance, an increase may result.

A critical contribution to the error variance is also made by $\Delta\beta$. It is obvious that for $|\Delta\beta| > \pi/2$ the cosine term becomes negative, and thus the error contribution of this term may become quite significant unless the amplitude is small. Inspection of Figure 5 reveals that β_1 for the system (5b) displays a distinct maximum for its frequency distribution, and a replacement of the individual $\beta_{1,i}$ by its mean $\bar{\beta}_1$ may suffice. However, β_1 for the system (5a) exhibits a bimodal distribution (Figure 6). Consequently we must find a characteristic parameter which provides a close approximation of β_1, A_0 and A_1. The investigation is still in progress but tentative results indicate that choosing a single characteristic such as:

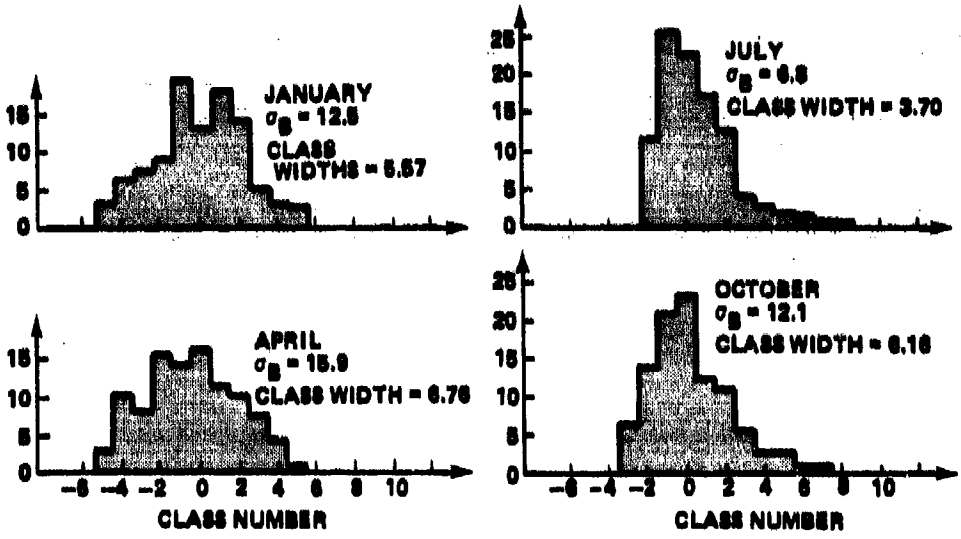


Figure 5. Frequency Distribution of β_1 Coefficient (Montgomery).

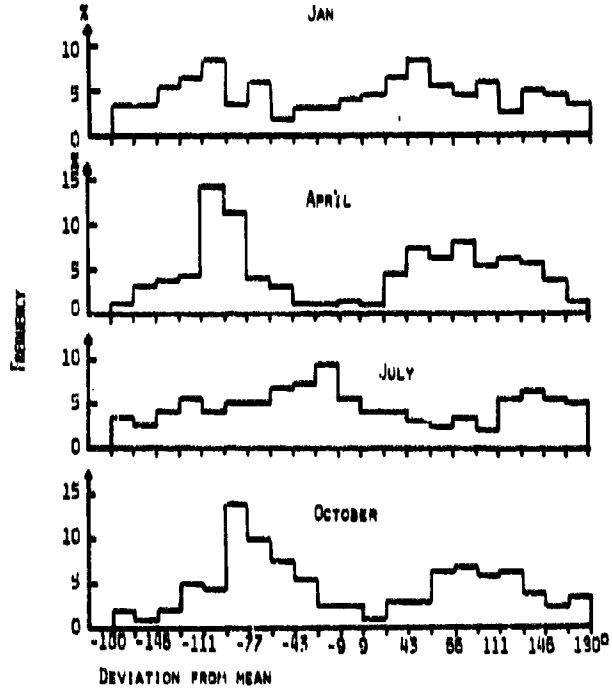


Figure 6. Frequency Distribution of β_1 Montgomery, Surface - 25 Km, $y = V_h - \bar{V}_h$.

$$y_{1,i} = \sum_k \omega_k V_{k,i} \tag{9a}$$

or:

$$y_{2,i} = \sum_r \omega_r (V_{r,i} - \bar{V}_r) \tag{9b}$$

may succeed. Then $\beta(y)$, $A_0(y)$ and $A_1(y)$ etc. k and r denote certain altitude levels, and ω stands for appropriate weights. Tentative results are depicted in Table 2 which had been obtained under favorable conditions. We learn that the surface to 25 Km single parameter system would be competitive with the eigenvector system. It should be considered that an increase of the variance of 25% is not significant at the 95% level of confidence for the F-test for N ~ 200.

It is emphasized that the replacement by standard functions cannot be generalized for the wind profile from all altitude ranges. For example, if our goal is the derivation of a single characteristic for the surface to 15 Km range, probably the eigenvector system is the best approach. The possibilities of a replacement by standard functions must be examined in every individual case.

6. CONCLUSIONS. A comparison was made between curve fitting systems based on empirical polynomials (i.e. eigenvectors) and standard functions. It was disclosed that the eigenvector system offers an optimum reduction of the variance with a minimum number of coefficients as expected from theory. It was illustrated, however, that under certain conditions standard functions may perform quite well, and these are simpler and more economical to compute than eigenvector functions.

7. REFERENCES CITED.

Essenwanger, Oskar M., 1964, Mathematical Characteristics of Individual Wind Profiles; Report RR-TR-64-12, US Army Missile Command, pp. 41.

Essenwanger, Oskar M., 1974, The Structure of the Wind Profile From Surface to 25 Km in Various Climatic Zones; Klimatologische Forschung (The Hermann Flohn 60th Anniversary Vol); Ferd, Dümmler, Bonn, p. 523-539.

Essenwanger, Oskar M., 1975, Eigenvector Representation of Wind Profiles, Preprints, Fourth Conference on Probability and Statistics in Atmospheric Sciences, p. 206-210; (Publ. American Meteorological Soc., Boston).

Essenwanger, Oskar M., 1976, Applied Statistics in Atmospheric Sciences, Vol A.; Elsevier, Amsterdam-New York, pp. 412.

8. ACKNOWLEDGMENT. The author wishes to express his gratitude to Dr. D. A. Stewart for her critical review of the manuscript.

Table 2. Left Variance

	Eigenvector			$V_h - \bar{V}_h$			V_h	
	1 Term	2 Terms	3 Terms	A_0	One Parameter	One F-Term	One Parameter	One F-Term
Jan	44.8	28.9	18.6	56.4	43.5	24.3	44.5	31.9
Feb	38.5	24.7	17.7	61.3	52.4	23.1	38.2	29.0
Mar	35.1	24.3	19.0	50.8	45.1	23.4	33.6	31.0
Apr	26.1	16.9	12.2	52.4	37.6	17.6	25.6	23.9
May	15.3	10.5	7.8	28.4	21.0	11.4	16.3	15.1
Jun	12.0	8.7	6.6	23.4	16.0	9.6	20.1	15.2
Jul	9.8	7.4	5.2	14.5	12.3	7.3	17.1	16.8
Aug	9.6	7.6	5.8	17.5	12.6	8.1	20.6	19.9
Sep	12.5	9.2	7.2	25.1	16.8	10.1	20.9	16.8
Oct	17.0	11.6	8.3	32.6	12.1	11.9	20.7	19.0
Nov	29.6	20.5	13.6	48.8	39.8	18.2	32.5	25.6
Dec	36.7	23.8	16.6	58.5	46.2	21.5	42.0	30.6

INDUCTION ON A MARKOV CHAIN

Richard M. Brugger

RAM Assessment Division
Product Assurance Directorate
U. S. Army Armament Command
Rock Island, Illinois

ABSTRACT. Through the use of Markov chain methods, expressions for Mean Rounds Between Failure (MRBF) were found for a class of weapon systems. The method led to an inductive determination of an expression for the general case.

Following the derivation of the general MRBF expression, expressions for reliability are obtained (but not a general expression).

1. INTRODUCTION. The problems treated in this paper relate to a shipboard weapon system of the following type. Some number (a variable) of gun mounts are connected in parallel. This parallel network is then connected in series with a fire control system. Each gun mount has the same number of guns (for simplicity, we will assume one gun per mount; the results are easily extended to some other number of guns per mount).

q_1 = Prob (given mount functions successfully) - (1)

q_2 = Prob (fire control functions successfully) (2)

$p_i = 1 - q_i$, $i = 1,2$. (3)

Once a mount fails, it is considered inoperative thereafter.

Note that we are assuming that each gun mount has the same success probability. This assumption simplifies the Markov chain work somewhat, but, as we will show later, even this simplifying assumption doesn't serve much purpose in the end.

In this particular application, the interest was focused only on the behavior of the gun mounts and fire control. We are therefore not concerned with failures of other parts of the system, such as the guns or the ammunition, and will, for convenience, assume that these function perfectly.

2. MEAN ROUNDS BETWEEN FAILURE (MRBF). In this application, MRBF will be defined as the expected number of rounds, successful and unsuccessful, attempted up to and including the first salvo where either none of the mounts function, the fire control does not function, or both events occur.

THIS PAGE NOT FILMED

For one mount, it is apparent that MRBF follows a geometric distribution. The probability of a salvo successfully occurring is $q_1 q_2$. By the properties of the geometric distribution, then

$$MRBF = 1/(1 - q_1 q_2) \tag{4}$$

For two mounts, a Markov chain was constructed with the following state definitions:

S0 = everything working

S1 = one mount out, fire control working

S2 = system not working

The transition matrix was as follows:

	S0	S1	S2
S0	$q_1^2 q_2$	$2p_1 q_1 q_2$	1 - Σ left elements
S1	—	$q_1 q_2$	1 - Σ left elements
S2	1	—	—

In the above matrix, the expression "left elements" refers to matrix elements in the same row but in columns to the left. One would ordinarily, and correctly, think that in row S2 the one should be in column S2 rather than S0, thereby reflecting the fact that state S2 is an absorbing state. This one is shifted to S0, however, to change the problem into one that can be treated as a first passage situation.

We will use column S2 to bring about degeneracy, so we are not concerned about what the actual values in this column turn out to be. Solving therefore, for the steady state probabilities in terms of the steady state probability for state S2 (denoted P(S2)), we have

$$P(S0) = \frac{1}{1 - q_1^2 q_2^2} P(S2) \tag{3}$$

$$P(S1) = \frac{2p_1q_1q_2}{(1-q_1^2q_2)(1-q_1q_2)}\,P(S2) \qquad (6)$$

$$P(S1) = P(S2) \qquad (7)$$

$$MRBF = \frac{2P(S0)}{P(S2)} + \frac{P(S1)}{P(S2)} \qquad (8)$$

$$= \frac{2}{1-q_1q_2}$$

For three mounts, the transition matrix becomes more complicated, so the simplified Markov chain method [1,2] was used. The states were defined in terms of situations, rather than on a salvo-by-salvo basis. The states were

S0 = system working

S1 = 1 mount failed in salvo of first failure

S2 = 2 mounts failed in salvo of first failure

S3 = system not working

The transition matrix becomes:

	S0	S1	S2	S3
S0	—	$\frac{3p_1q_1^2q_2}{1-q_1^3q_2}$	$\frac{3p_1^2q_1q_2}{1-q_1^3q_2}$	1 - Σ left elements
S1	1	—	—	—
S2	1	—	—	—
S3	1	—	—	—

Going through the steps required for solution, as described in 1,2 , we obtain

$$MRBF = \frac{3}{(1-q_1q_2)} \tag{9}$$

Using the simplified Markov chain method for 4,5 and 6 mounts, the pattern

$$MRBF_i = \frac{i}{(1-q_1q_2)} \tag{10}$$

continued, where i is the number of mounts.

Since only the top row of the transitional matrix in the simplified form has any new information as the number of mounts increase, and since the expected length of the various states (in the simplified Markov chain sense) was determined as the number of mounts increased, induction was considered.

By considering the result true for $k-1$ and considering what the structure of the top row of the transition matrix would be for k , it was seen that

$$MRBF = \frac{k}{1-q_1^k q_2} + q_2 \frac{\sum_{i=1}^{k-1} \binom{k}{i} p_1^i q_1^{k-i} (k-1)}{(1-q_1^k q_2)(1-q_1 q_2)} \tag{11}$$

$$= \frac{k}{1-q_1 q_2}$$

QED.

After obtaining this general solution, further relection was given to the problem. Because of the nature of the simplified Markov chain transition matrix for this proble (whereby all of the information of interest appeared along the top row) i was seen that a direct algebraic induction solution, without any use of Markov chains at all, was possible.

Finally, it was seen that the problem was actually much simpler; even algebraic induction was not necessary. This was determined as follows. Let us imagine that an observer is stationed by each mount, and that each observer will remain by his mount for an infinite length of time (or for an infinite number of trials from the situation "everything working" to "system down"). Each observer records how many rounds are fired from his mount until his system (the fire control and his mount) breaks down. His system is equivalent to a one mount system, as is each of the other observers, so, over the long run, the average number of rounds between failure for his system will be the same as the MRBF for one mount. For several mounts, then, the MRBF for the system is just equal to the sum of the MRBF's for individual mounts (whereby our earlier simplifying assumption that all mounts have the same MRBF is seen to be unnecessary). While the common fire control suggests dependency, the dependency exists only for each trial from "everything working" to "system down"; it does not exist for the system MRBF.

From the above, it is seen that almost no mathematics was necessary for solution. At the same time, the mathematics bears out the result obtained through the purely intuitive approach just described.

2. RELIABILITY. For this application, the reliability for an N round mission will be defined as the probability that a mission of N successful rounds will be accomplished.

For one mount, we have a simple geometric distribution, and the N round reliability can be expressed as

$$R_N = (q_1 q_2)^N \qquad (12)$$

For two mounts, let k_o be the number of salvos that would be required if the Nth successful round were fired in the k_o'th salvo and no breakdowns occurred in the first k_o-1 salvos. If N is even, $k_o = N/2$. If N is odd, $k_o = (N+1)/2$.

After some investigation into how the problem could best be algebraically treated, it was found that the best approach would be one whereby any necessary summations would be indexed by the number of successful fire control salvos. Thus, for two mounts, N even, we have

$$R_N = q_1^N q_2^{k_o} + \sum_{k=k_o+1}^{N} 2p_1 q_1^N q_2^k \tag{13}$$

which, for $q_2 < 1$, is found to be

$$q_1^N q_2^{k_o} + 2p_1 q_1^N \frac{q_2^{k_o+1} - q_2^{N+1}}{1 - q_2} \quad . \tag{14}$$

For two mounts, N odd, we have

$$R_N = q_1^{N+1} q_2^{k_o} + 2p_1 q_1^N q_2^{k_o} + \sum_{k=k_o+1}^{N} 2p_1 q_1^N q_2^k \tag{15}$$

which, for $q_2 < 1$, is found to be

$$q_1^{N+1} q_2^{k_o} + 2p_1 q_1^N q_2^{k_o} + 2p_1 q_1^N \frac{q_2^{k_o+1} - q_2^{N+1}}{1 - q_2} \quad . \tag{16}$$

For three mounts, the problem becomes slightly more complicated. Let us make the following definitions for k_o .

If

$$N = 0 \mod 3, \quad k_o = N/3 \tag{17}$$

$$N = 1 \mod 3, \quad k_o = (N+2)/3 \tag{18}$$

$$N = 2 \mod 3, \quad k_o = (N+1)/3 \quad . \tag{19}$$

The following probability of mutually exclusive events are defined.

1. P(0) is the probability that the N'th successful round occurs on the k_o'th salvo.

2. P(0A) is the probability that the N'th successful round occurs after the k_o'th salvo, but no mount failures occur in the first (k_o-1) salvos.

3. P(1) is the probability that one mount failure occurs in the first (k_o-1) salvos, no more failures occur, and the N'th successful round occurs after the k_o'th salvo.

4. P(2) is the probability that two mount failures occur, at least one before the k_0'th salvo, and the N'th successful round occurs after k_0'th salvo.

Then

$$R_N = P(0) + P(0A) + P(1) + P(2) \quad . \tag{20}$$

When N = 0 modulo 3

$$P(0) = q_1^N q_2^{N/3} \quad . \tag{21}$$

When N = 1 modulo 3

$$P(0) = q_1^{N-1} q_2^{(N+2)/3}(1 - p_1^3) \quad . \tag{22}$$

When N = 2 modulo 3

$$P(0) = q_1^{N-2} q_2^{(N+1)/3}(q_1^3 + 3p_1 q_1^2) \quad . \tag{23}$$

When N = 0 modulo 3

$$\begin{aligned} P(0A) &= q_1^{N-3} q_2^{(N-3)/3}\{(3p_1 q_1^2 q_2)(1 - p_1^2)q_2 \\ &\quad + (3p_1^2 q_1 q_2) q_1^2 q_1^2\} \qquad (24) \\ &= 3p_1 q_1^{N-1} q_2^{(N+3)/2}(1 - p_1^2 + p_1 q_1 q_2) \quad . \end{aligned}$$

When N = 1 modulo 3

$$P(0A) = 0 \quad . \tag{25}$$

When $N = 2$ modulo 3

$$P(OA) = q_1^{N-2} q_2^{(N-2)/3} [(3p_1^2 q_1 q_2) q_1 q_2]$$

$$= 3 q_1^N q_2^{(N+4)/3} p_1^2 \quad . \tag{26}$$

Let $S(N/2)$ be the smallest interger larger than or equal to $N/2$. Then

$$P(1) = \sum_{\substack{k=k_o+1 \\ k \leq N/2}}^{S(N/2)} 3p_1 q_1^{N-2k} q_1^{2k} q_2^k + \sum_{\substack{k=k_o+1 \\ k \leq (N+1)/2}}^{S(N/2)} 3p_1 q_1^{N-2k+1} q_1^{2k} q_2^k . \tag{27}$$

If $k_o+1 = N/2$ the above generalizes to

$$3p_1 (q_1^N + q_1^{N+1}) \frac{q_2^{k_o+1} - q_2^{S(N/2)+1}}{1 - q_2} \tag{28}$$

for $q_2 < 1$. If $(k_o+1) > N/2$, $P(1)$ does not axist (the case for N small).

$$P(2) = \sum_{k=k_o+1}^{N} 3p_1^2 C_k q_1^N q_2^k \tag{29}$$

where C_k indicates the number of ways two numbers can add up to N-k given the maximun of these two numbers is less than k.
For $S(N/2) \leq k \leq N$

$$C_k = N - k + 1 \quad . \tag{30}$$

For $k_o+1 \leq k \leq S(N/2) - 1$

$$C_k = (-N + 3k - 1) \tag{31}$$

Then

$$\sum_{k=S(N/2)}^{N} 3p_1^2(N - k + 1)q_1^N q_2^k$$

$$= 3p_1^2 q_1^N \{ \sum_{S(N/2)}^{N} (N - k - 1 + 2)q_2^k \} \tag{32}$$

$$= 3p_1^2 q_1^N \{ \sum_{S(N/2)}^{N} (N + 2)q_2^k - \sum_{S(N/2)}^{N} (k + 1)q_2^k \}.$$

Treating the right hand term within the brackets as the sum of derivatives (being equal to a derivative of a sum) the above becomes, for $q_2 < 1$

$$3p_1^2 q_1^N \{(N + 2) \frac{q_2^{S(N/2)} - q_2^{N+1}}{1 - q_2}$$

$$- [\sum_{0}^{N} (k + 1)q_2^k - \sum_{0}^{S(N/2)-1} (k + 1)q_2^k] \}$$

$$= 3p_1^2 q_1^N \{(N + 2) \frac{q_2^{S(N/2)} - q_2^{N+1}}{1 - q_2} \tag{33}$$

$$- [\frac{1 - (N + 2)q_2^{N+1} + (N + 1)q_2^{N+2}}{(1 - q_2)^2}$$

$$- \frac{1 - (S(N/2) + 1)q_2^{S(N/2)} + S(N/2)q_2^{S(N/2)+1}}{(1 - q_2)^2}] \},$$

In a similar manner, we find

$$\sum_{k_o+1}^{S(N/2)-1} (- N + 3k - 1)3p_1^2 q_1^N q_2^k$$

$$= 3p_1^2 q_1^N \Big(- (N + 4) \frac{q_2^{k_o+1} - q_2^{S(N/2)}}{1 - q_2}$$

$$+ 3\Big\{ \frac{1 - (S(N/2) + 1)q_2^{S(N/2)} + S(N/2)q_2^{S(N/2)+1}}{(1 - q_2)^2} \qquad (34)$$

$$- \frac{1 - (k_o + 2)q_2^{k_o+1} + (k_o + 1)q_2^{k_o+2}}{(1 - q_2)^2} \Big\} \Big) .$$

REFERENCES

1. Brugger, R.M., "A simplification of the Markov chain approach to continous sampling plan formulation", QEM 21-230-12, March 1972.

2. Brugger, R.M., "A simplification of skip-lot procedure formulation", Journal of Quality Technology, October 1975.

MARKOV AND PATH DEPENDENT PROCESSES APPLIED TO CONTINUOUS SAMPLING PLANS IN TANDEM

David L. Arp
Naval Weapons Center
China Lake, California

ABSTRACT. A continuous sampling scheme, consisting of two generic Continuous Sampling Plans (CSP) in series, is analyzed. This serial arrangement is used for the attribute sampling for two different independent characteristics of items in a given production run; the output from the first plan forms the input to the second. Using standard one dimensional Markov Chain (MC) models for the generic CSP's, the serial CSP model is shown to be equivalent to a two dimensional (or second order) MC wherein the state of the second component is directly dependent on that of the first.

The ergodic properties of the marginal distribution of the second component are analyzed by using 1) the ergodic theorem applied to matrix valued random variables, 2) a nonstationary MC approximation to a path dependent process, and 3) direct products of transition matrices constrained by the dependence mentioned above. In the latter two approches, the MC's are shown to be aperiodic and (strongly) ergodic; either one can be used to show convergence of the path dependent process. Taking the appropriate limits, as the production run becomes infinite, it is proven that the limiting probabilities for the second component are independent of those of the first.

Using direct products, the analysis is extended to the case of three or more CSP's in tandem. Under the additional assumption of a separable initial probability vector and for $n \geq 2$, the direct product MC, which is ergodic and stationary, is shown to be equivalent to a finite sequence of n MC's. In this sequence, the first MC is ergodic and stationary; the remaining MC's are (strongly) ergodic and nonstationary. Comparisons are also made with other naturally arising multicharacteristic sampling plans.

1.0 INTRODUCTION.

1.1 Continuous Sampling Plans. Given a production line of items, a (one characteristic) Continuous Sampling Plan (CSP) consists of two or more phases of attribute sampling for an item characteristic directly from the line. In at least one phase, the sampling frequency is zero with an exit occurring only after a fixed number of items are found to be consecutively nondefective (screening phase). The phases are always connected is such a way that each of them is "positive recurrent" for an (abstract)

infinite production run. Moreover, the number of phases is finite and an exit from any one of them takes place after a finite number of production units with probability one.

CSP's are modelled by Markov Chains (MC) which, because of the phase structuring, are finite, aperiodic, and irreducible. The plans and their MC models are discussed at length in References 6.2 and 6.5. The simplest of the CSP's, CSP-1, along with its usual MC model, is described in Chapter 2.

1.2 Origin of Tandem CSP's. In the past, CSP-1 has been used in a serial manner to sample for eight different characteristics per production unit. In practice, the characteristics were sampled for at successive stations along the production line. It is this type of sampling that is generalized and modelled in Chapter 2 and further analyzed in the succeeding chapters.

1.3 Contents of Paper. In Chapter 2, after describing CSP-1 and its MC model, Semi Markov Chains (SMC) are introduced and utilized to simplify the MC model in two ways: the "classical" way, driven by a particular functional, and a second way, motivated by the serial sampling plan and the idea of a controlled Markov Chain (MC). Such a SMC simplification of a MC is called SMC reduction (see Reference 6.2). The description of (2)-serial CSP-1 is then given followed by a second order MC ((2)-MC) model for it. The (2)-MC model is based on the assumption of independent characteristics.

In Chapter 3, the second SMC reduction is used in developing two similar approaches to the simplification of the (2)-MC model. The major connections between the resulting models are also brought out. The second, path dependent model is approximated by a strongly ergodic nonstationary MC. In Reference 6.2, it is erroneously stated that this approximation is equivalent to the (2)-MC. Thus, one of the major purposes of Chapter 3 is to clarify the assumptions made which make the nonstationary MC differ from the (2)-MC.

In Chapter 4, the longest of the chapters, a third method is given which utilizes the concept of the direct product of matrices. For $n > 2$, (n)-serial CSP-1 is also handled by the same techniques and the CSP-1 restriction is eventually dropped. For (n)-serial CSP-1, it is also shown that its direct product MC, which is stationary and ergodic, can be separated into n MC's. The first of these MC's is also stationary (and ergodic) in contrast to the remaining ones which are nonstationary (and strongly ergodic). Furthermore, the latter n-1 nonstationary MC's exhibit structures which are essentially different from the one exhibited by the nonstationary MC in Chapter 3. Of primary interest is the marginal Average Fraction Inspected (AFI) functional for the last plan in tandem. This functional is compared to the one which results from use of the plan by itself. The treatment of other reasonable nonserial multicharacteristic sampling plans concludes the chapter. Chapter 4 contains all the major results in the most satisfactory form.

Chapter 5 concludes the paper by summing up the major conclusions and theorems as well as suggesting some further possibilities for and modifications of multicharacteristic sampling plans.

1.4 Glossary. In References 6.1, 6.2, and 6.3, the clearance number, which characterizes the screening phase of CSP-1, is denoted by the capital letter I. However, in this paper, "I" might be confused with the identity matrix and thus small i will be used instead for the clearance number.

Henceforth, references will be denoted by numbers in brackets (e.g., "References 6.2 and 6.3" will be written as [6.2,6.3]. Common abbreviations and notations are given below.

rxm = r columns and m rows

[a.e.] = almost everywhere

pv = probability vector (non-negative entries with sum = 1)

CSP = Continuous Sampling Plan

FI(N) = Fraction Inspected out of N units

AFI(N) = Average of FI(N)

$AFI_n(\infty)$ = Marginal AFI(∞) for the nth plan in a (n)-serial CSP

MC = Markov Chain; SMC = Semi Markov Chain

M(·) = MC process; X(·) = SMC process

$A \otimes B$ = Direct product of the two matrices

A_{1n} = Transition matrix of (n)-serial CSP-1

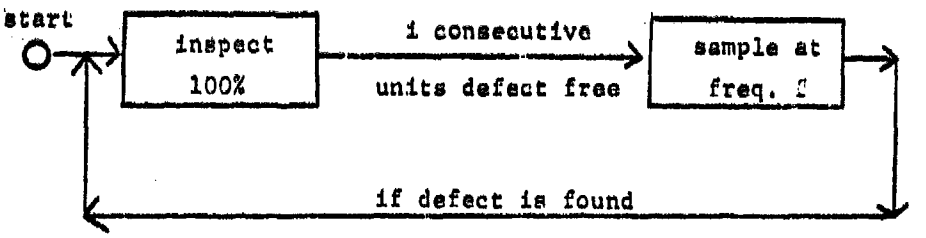
1.5 Acknowledgment. Mrs. Leah K. Jones deserves full credit for the excellent and expeditious typing of the paper as well as for the drafting of some complicated diagrams and the proper rendering of special technical symbols.

2.0 BACKGROUND.

2.1 CSP-1. This sampling plan, the simplest of its type, is characterized by one variable and two parameters. The variable, p, is the probability of finding a defective item (characteristic) under the assumption that the product flow forms a Bernoulli process. The two parameters are i, the clearance number required to exit from the screening phase (abbr. sc), and f, the sampling frequency to use during the unlimited sampling phase (abbr. uls). Thus, when necessary for clarity, a particular CSP-1 will be written explicitly as CSP-1[p; i,f]. The black box description of and the MC model for the plan appear in Figures 1 and 2, respectively.

Figure 1

Block Diagram of CSP-1[p; i,f]

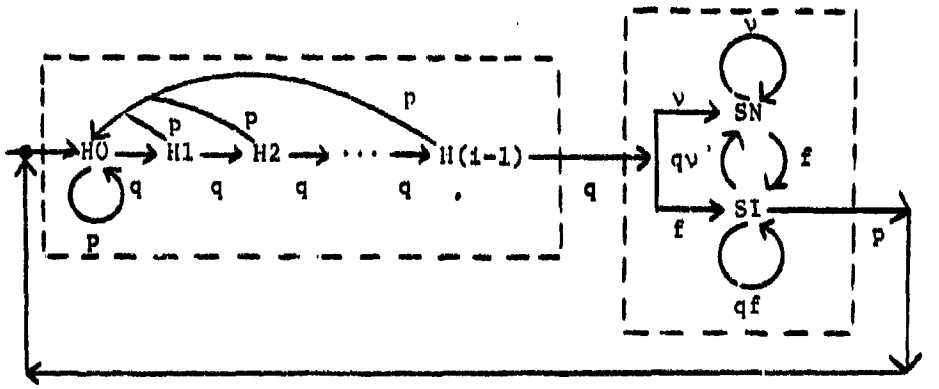


First box = screening phase (sc)

Second box = unlimited sampling phase (uls)

Figure 2

Markov Chain Model of CSP-1[p; i,f]



p = Probability of defective; q = 1-p

i = Clearance number

f = Sampling frequency; ν = 1-f

Hj = MC state of sc, $0 \le j \le i-1$

SN = Noninspection MC state of uls

SI = Inspection MC state of uls

2.2 Semi Markov Chains. Semi Markov Chains (SMC) can be used to simplify CSP's and, specifically, CSP-1. Below, a brief exposition of SMC's is given. For further details, see [6.2, Appendix of 6.3, or 6.6].

For discrete (and integral) $t \geq 0$, let $X(t)$ be a (discrete) stochastic process. Then we have

Definition 1. $X(t)$ is a finite Semi Markov Chain iff its state space is finite and and following relationship holds

$$\text{Prob}[Y(n), W(n) | Y(m), W(m);\ 0 \leq m \leq n-1]$$

$$= \text{Prob}[Y(n), W(n) | Y(n-1), W(n-1)]$$

where $Y(m+1) = X(t_{m+1})$, $T(m+1) = W(m+1) - W(m)$ is the time of sojourn in state $Y(m)$ from its entrance until its exit to state $Y(m+1)$, $t_{(m+1)}$ is a particular realization of the random variable $W(m+1)$ which in turn is the total time to $(m+1)$st transition, and $Y(m) \neq Y(m+1)$ for all m.

For further reference, we have

Definition 2. Let ℓ, k be in the state space of $X(\cdot)$. Then

a. The (defective) pdf of the time to transition from state ℓ to state k is

$$Q_{\ell,k}(t) = \text{Prob}[X(t) = k;\ X(t') = \ell,\ t > t' > 0 | X(0) = \ell],$$

for $\ell \neq k$ and is otherwise zero.

b. The probability of starting in state ℓ at time zero and being in state k at time t is given by .

$$P_{\ell,k}(t) = \text{Prob}[X(t) = k | X(0) = \ell].$$

In Definition 1, the process $Y(\cdot)$ is a MC called the embedded MC of the SMC. Letting H_0 be the Heaviside sequence, the transition matrix for this MC is

$$[H_0 * Q_{\ell,k}(\infty)]$$

where the asterisk denotes the operation of convolution. If in Definition 2, there should exist at least one state k such that $Q_{k,k}(\cdot)$ is not identically zero, then self transitions are possible without being recorded by the SMC apparatus. In this case, the concept of a Markov Renewal Process (MRP) must be used. Referring to Definition 1, the MRP would be the process $(X(\cdot), W(\cdot))$. In the rest of the paper, we will be dealing with aperiodic, irreducible, and stationary SMC's. The definitions of all these concepts parallel those for MC's. For further information on MRP's, types of SMC's and their relationships with their embedded MC's, see [6.2, 6.3, or 6.7].

We finish this section by stating two theorems needed later on.

Theorem 1. Given the SMC $X(\cdot)$, we have

$$P_{\ell,k}(t) = \sum_j Q_{\ell,j} * P_{j,k}(t) + (\delta_{\ell,k}) J_k(t)$$

where $P_{\ell,k}(\cdot)$ and $Q_{\ell,j}(\cdot)$ are defined in Definition 1, $\delta_{\ell,k}$ is the Kronecker delta, and

$$J_k(t) = H_0 * (\delta_0 - \sum_s Q_{k,s})(t)$$

for $\delta_0(t) = \delta_{0,t}$.

Proof. See [6.2, 6.6, or 6.7].

Theorem 2. Given the SMC $X(\cdot)$, the following limit holds.

$$\lim_{t\to\infty} P_{\ell k}(t) = \frac{e_k \mu_k}{\sum_s e_s \mu_s}$$

$$= a_k$$

where e = the unique eigenvector with eigenvalue 1 for the embedded MC and μ_k = the mean time of sojourn in state k.

Proof. See [6.2, 6.6, or 6.7].

2.3 Simplification of CSP-1. The first simplification is driven by the Fraction Inspection (FI) functional which is given in

Definition 3. For the model of CSP-1 appearing in Figure 2, the Fraction Inspected (FI) functional is

$$FI(N) = 1 - \frac{\nu}{N} \sum_{t=0}^{N} C_{(uls)}(t)$$

In the equation, N = the total number of units which have passed the inspection station in real time, $\nu = 1-f$, and

$$C_{(uls)}(t) = \begin{cases} 1, & \text{if } X(t) \text{ is in uls} \\ 0, & \text{otherwise} . \end{cases}$$

Taking the conditional average of FI(N) gives a function defined in

Definition 4. The Average Fraction Inspected (AFI), for the first N units and starting in either MC state HO or in any state under equilibrium conditions is

$$AFI(N) = E[FI(N)|M(0) = HO]$$

$$= E_{\underline{e}}[FI(N)], \text{ also}$$

where M(·) is the MC process, E[·] the expectation operator, and $\underline{e}$ the long run probability vector (pv).

Concerning the first simplification of CSP-1, we have

Theorem 3. Letting sc = 1 and uls =2 (see Figure 2), we can construct the following SMC whose states are defined in terms of the z transform [6.1, 6.2, or 6.11].

States: $(1, \hat{Q}_{12}(z))$ and $(2, \hat{Q}_{21}(z))$

where $\hat{Q}_{12}(z) = \dfrac{q^i(z-q)}{z^i(z-1)+\gamma}$, $\hat{Q}_{21}(z) = \dfrac{\delta}{z-\beta}$, $\gamma = pq^i$,

$\delta = fp$, and $\beta = 1-\delta$.

Proof. See [6.2].

Corollary 1. The unlimited sampling phase of CSP-1 can be reduced to a MC state Si with a geometric pdf.

Proof. From Theorem 3, the transform of the function $\hat{Q}_{21}(z)$ is a (nondefective) pdf which can be written (in the time domain) as

$$Q_{Si,HO}(t) = \delta\beta^{t-1}$$

In the above equation, HO is used since the application of this Corollary will be to the MC model.

Corollary 2. Starting in state 1 at time zero, the FI(N) functional in Definition 3 has a limit as N approaches infinity given by

$$\lim_{N\to\infty} FI(N) = 1-\nu\alpha_2 \qquad [a.e.]$$

$$= AFI(\infty)$$

Proof. The first equality follows from Theorem 2 applied to the SMC constructed in Theorem 3 and the ergodic theorem for functionals defined (or, in this case, definable) on SMC's. The second equality follows from

Definition 4 and the facts that "M(0) = H0" is equivalent to "X(0) = 1" and $E[C_2(t)|X(0) = 1] = P_{i2}(t)$.

The second simplification will be used in Chapter 3.

Theorem 4. From the MC model of CSP-1, a SMC can be constructed with the following states (again in terms of the z transform)

States: $(a, \hat{Q}_{ab}(z))$ and $(b, \hat{Q}_{ba}(z))$

where

$$\hat{Q}_{ab}(z) = \frac{q}{z-p} \quad \text{and} \quad \hat{Q}_{ba}(z) = \frac{\hat{Q}_{cd}(z)\cdot\hat{Q}_{da}(z)}{1-\hat{Q}_{ac}(z)\cdot\hat{Q}_{ca}(z)}$$

The transfer functions for the intermediate states c and d are

$$\hat{Q}_{cd}(z) = \left(\frac{q}{z}\right)^{i-1} \text{ and } \hat{Q}_{ca}(z) = \left(\frac{q}{z-p}\right)\left(1-\left(\frac{q}{z}\right)^{i-1}\right).$$

Proof. Let a = H0, $c = \{Hj\}$, for $1 \le j \le i-1$, and d = Si. Then a and d have geometric pdf's and are thus (trivial) SMC states. From [6.2], c is a SMC state with the given transform. Using a routine combinatorial argument, we have (dropping the argument z)

$$\hat{Q}_{ba} = \hat{Q}_{cd}\cdot\hat{Q}_{da}\left\{\sum_{j=0}^{\infty}(\hat{Q}_{ac}\cdot\hat{Q}_{ca})^j\right\}$$

which reduces to the given form by summation of a geometric series for $|z| > 1$.

2.4 MC Model for (2)-Serial CSP-1. We consider two (different) CSP-1's in tandem: CSP-1 $[p_k; i_k, f_k]$ with MC and SMC states $\{Hjk, Sik\}$ and $\{a_k, b_k\}$, respectively.

The (2)-MC model of (2)-serial CSP-1 is based on the assumption that the two item characteristics being sampled for are independent. Following the practical case discussed in Chapter 1, two item characteristics are sampled for at two successive stations along a production line, according to two (different) CSP-1 types. If an item is rejected because of a defective first characteristic, then the second characteristic is not sampled for. Thus a transition to H01 occurs in the first plan but no transition at all occurs in the second plan for the given operational time increment which the item represents. However, if the item passes muster for the first characteristic (i.e., the item is inspected and found to be nondefective in the 1st characteristic or, because of f_1, is not inspected), a transition takes place in the first plan to a state other than H01 (or a_1) and the item moves on to the second station. Thus, in this latter situation, a transition takes place in both MC's for the specific

operational time increment generated by the unit. We translate this viewpoint into the (2)-MC model given in Figure 3.

Figure 3

Second Order Markov Chain Model for (2)-Serial CSP-1

States: $\{(k1, i2), (k1, k2), (i1, k2),$ for $0 \leq kj \leq ij-1, j = 1,2\}$

Transitions: ((kj)+1 may be ij, j = 1,2)

State		State	Probability
(k1, k2)		((k1)+1, (k2)+1)	q_1q_2
(i1, k2)	$\Longrightarrow$	(i1, (k2)+1)	β_1q_2
(k1, i2)		((k1)+1, i2)	$q_1\beta_2$
(i1, i2)		(i1, i2)	$\beta_1\beta_2$
(k1, k2)		((k1)+1, 0)	q_1p_2
(i1, k2)	$\Longrightarrow$	(i1, 0)	β_1p_2
(k1, i2)		((k1)+1, 0)	$q_1\delta_2$
(i1, i2)		(i1, 0)	$\beta_1\delta_2$
(k1, x)	$\Longrightarrow$	(0, x)	p_1
(i1, x)		(0, x)	δ_1

(x = i2 or k2).

The result is a rather complicated 2 dimensional lattice. The remaining chapters reduce the study of this model and, more generally, similar models for (n)-serial CSP-1's and functionals defined on them to a manageable systematic analysis with various degrees of success. To help in this analysis, we fix some more ideas in two more definitions before leaving Chapter 2.

Definition 5. A (n)-serial plan is the same as a (n)-serial CSP and consists of CSP's arranged in tandem such that the output of the jth plan is the input to the (j+1)st plan, $1 \leq j \leq n-1$. For a given operational time increment given by the movement of a production unit through the sampling stations, a transition takes place in the (j+1)st plan only if no defects are found in the preceding j plans. Moreover, if a defect is found at the jth station, no transitions take place in the consecutive plans after j.

However, the interpretation of "virtual transition" for "no transition" will also be used when convenient to do so. If only a particular type of CSP is used, the serial plan will be called a (n)-serial CSP-"type". If the CSP's are mixed types, the plan will generally be written out: (CSP-type(1))— ... —(CSP-type(n)).

Definition 6. A multicharacteristic plan (MCP) will be used as a generic term while a non-CSP MCP will be called a variant MCP.

3.0 TWO APPROACHES TO (2)-SERIAL CSP-1. The two approaches are given in Sections 3.1 and 3.2. The connections between them are given in Section 3.3. In addition, a and b are the SMC states appearing in Theorem 4 for the first plan, A_2 is the usual transition matrix for the second plan used alone, and I_2 is the identity matrix of rank i_2.

3.1 Average Transition Matrix. Given the (2)-MC model for (2)-serial CSP-1, we first define the matrix valued characteristic functional in

Definition 7. Let 1) ω be a realization of the process $(X(t), M_2(t))$, where $X(\cdot)$ is the SMC variable for the first plan and $M_2(\cdot)$ is the MC variable for the second plan and 2) $Proj_t(\omega)$ be the projection to the first component at time t. Then the matrix valued characteristic functional is

$$C_t(\omega) = \begin{cases} A_2, \text{ if } Proj_t(\omega) = b \\ I_2, \text{ if } Proj_t(\omega) = a \end{cases}$$

Using the idea of a controlled MC (see [6.12]) and Definition 7, we can prove

Theorem 5. As N approaches infinity,

$$\frac{1}{N} \sum_{t=1}^{N} C_t(\omega) \to \alpha_a I_2 + \alpha_b A_2 \quad [a.e.]$$

Proof. We can break the matrix valued random variable up as

$$C_t(\omega) = a_t(\omega) I_2 + b_t(\omega) A_2$$

The functionals $a_t(\cdot)$ and $b_t(\cdot)$ have the obvious definitions: $a_t(\omega) = 0$ or 1 iff $Proj_t(\omega) = b$ or a, respectively and $b_t(\omega) = 1 - a_t(\omega)$. Then the above average sum can be similarly decomposed. The theorem then follows from the definition of the (2)-MC model given in Figure 3, the SMC reduction of CSP-1 in Theorem 4, and the ergodic theorem for functionals defined on SMC's.

Using Theorem 5, an average operator can be associated with the second plan in

Definition 8. Given the RHS of the limit in Theorem 5, the Average Transition Matrix for the second plan is

$$\bar{A}_2 = \alpha_a I_2 + \alpha_b A_2$$

Clearly, for the second plan, the expression for $\bar{A}_2$ can be looked upon as stating that, in the long run, I_2 is the (virtual) "transition matrix" $(100)\alpha_a\%$ of the time while A_2 is the appropriate matrix for the remaining $(100)\alpha_b\%$ of the time. To elaborate somewhat, I_2 can be interpreted as the "Stop" matrix. That is, when I_2 is employed, no transitions take place as far as production unit time is concerned. A possibly better interpretation is to consider (virtual) transitions as taking place according to the identity matrix but to define the relevant functionals only for transitions which occur according to A_2. With this latter viewpoint, we then have a path dependent nonstationary process (see Section 3.2).

Given $\bar{A}_2$, we have

Theorem 8.

$$\lim_{k\to\infty} \left(\bar{A}_2\right)^k = L_2$$

where L_2 is the usual long run matrix for the second plan. That is, the columns are all identically equal to the long run probability vector (pv) a_2.

Proof.

$$\left(\bar{A}_2\right)^k = \sum_{j}^{k} \binom{k}{j} \alpha_a^j \left(I_2 - A_2\right)^j A_2^{k-j} \qquad (1)$$

However,

$$\lim_{k\to\infty} A_2^k = L_2 \qquad (2)$$

Therefore, Eq. (2) and summability theory [6.9] imply that the limit exists for Eq. (1) and is L_2.

Theorem 8 shows that the use of the average matrix gives the same long run results that use of A_2 does. Thus, using this first approach results in a marginal AFI(∞) which is the same as that which would be obtained if the second plan were to be used by itself.

3.2 Path Dependent Model. The model is given by the matrices in

Definition 9. The path dependent matrices for (2)-serial CSP-1 is

$$A(t,\omega) = \prod_{k=1}^{t} \left(a_k(\omega) I_2 + b_k(\omega) A_2 \right)$$

where $a_k(\cdot)$ and $b_k(\cdot)$ are defined in the proof of Theorem 5 and the matrices are defined in Section 3.0.

Let the conditional expectations $E[\cdot | X_1(0) = a]$ and $E[\cdot | \underline{a1}]^*$ operate on the above matrices to yield matrices $A_a(t)$ and $A_\alpha(t)$, respectively. Also let $\underline{e}_1 = (1,0,\cdots,0)$, i_1 times, and $\underline{y}_2$ be an arbitrary pv with i_2 entries. Then, a little reflection shows that the (2)-MC model with initial pv = $[\underline{e}_1, \underline{y}_2]$ or $[\underline{e}_1, \underline{y}_2]$ is equivalent to using $A_a(t)$ or $A_\alpha(t)$, respectively, with initial pv = $\underline{y}_2$.

Using the equality "$b_k(\cdot) = 1-a_k(\cdot)$", we can rewrite the matrices in Definition 9 as

$$A(t;\omega) = \prod_{k=1}^{t} \left(a_k(\omega) \left(I_2 - A_2 \right) + A_2 \right) \qquad \text{(A3)}$$

Multiplying the RHS of Eq. (A3) out, we get

$$A(t;\omega) = (a_1 a_2 \cdots a_t) \left(I_2 - A_2 \right)^t$$

$$+ \cdots + \left(\sum{}' a_{j_1} a_{j_2} \cdots a_{j_s} \right) \left(I_2 - A_2 \right)^s A_2^{t-s}$$

$$+ \cdots + \left(\sum_{j=1}^{t} a_j \right) \left(I_2 - A_2 \right) A_2^{t-1} + A^t \qquad \text{(B3)}$$

In Eq. (B3), the arguments of the a_j's have been dropped for notational convenience and "$\sum'$" is the restricted summation obtained by requiring that $j_1 < j_2 < \cdots < j_s$.

From Eq. (B3), a recursive scheme can be developed. For a (complex) polynomial of degree n with roots r_j (j = 1 to n), let $S_k(r_1,\cdots,r_n)$ be the kth symmetric function of the roots (associated with the variable of power n-k). For simplification in using Eq. (B3), define

$$D_k^n(\omega) = S_k(a_1,\cdots,a_n).$$

---

*$\underline{a1}$ is SMC notation for $\underline{e}_1$.

Thus, for example,

$$D_0^n = 1,\ D_1^n = \sum_j a_j,\ \text{and}\ D_n^n = \prod_j a_j$$

Since $a_k(\omega) = 0$ or 1, we can consider the RHS of Eq. (B3) as a random matrix polynomial over the binary field. Using the symmetric function, we have

Theorem 9. With the special random symmetric functions defined above, we have a recursive relationship between the coefficients of $A(n+1;\omega)$ and $A(n;\omega)$ where we treat $(I_2 - A_2)$ as "1" and A_2 as the polynomial indeterminate.

Proof. The recursion is obtained by expressing $A(n+1;\omega)$ as $A(0,n;\omega) \cdot A(n,n+1;\omega)$ and equating coefficients. Explicitly, the recursion is given by

$$D_0^{n+1} = 1$$

$$D_{n+1}^{n+1} = a_{n+1} + D_n^n$$

$$D_k^{n+1} = D_k^n + a_{n+1} D_{k-1}^n,\ 1 \le k \le n$$

In particular,

$$D_{n-s}^n = \sum{}' a_{j_1} a_{j_2} \cdot \cdots \cdot a_{j_s}$$

Eq. (B3) is more useful for calculation of Eq. (A3) because of

Proposition 1.

$$E[a_{j_1} a_{j_2} \cdot \cdots \cdot a_{j_s} | X(0) = a]$$

$$= P_{aa}(j_1) P_{aa}(j_2 - j_1) \cdot \cdots \cdot P_{aa}(j_s - j_{s-1}).$$

Proof. Since state a has a geometric pdf,

$$P_{aa}(j) = \text{Prob}[M_1(j) = HO | M_1(0) = HO]$$

where $M_1(\cdot)$ is the MC process for plan 1.

Corollary 1. For the first plan, letting $E[\cdot | X(0) = a] = E_a[\cdot]$, we have

$$E_a[a_1 a_2 \cdot \cdots \cdot a_n] = p_1^n$$

$$E_a\left[\sum_j a_j\right] = \sum_j P_{aa}(j)$$

Proof. Proposition 1 and definitions.

Corollary 2. With the same conditions as Corollary 1,

$$E_a[D_0^{n+1}] = 1$$

$$E_a[D_{n+1}^{n+1}] = P_{aa}(n+1) + E_a[D_n^n]$$

$$E_a[D_k^{n+1}] = E_a[D_k^n] + E_a\left[D_k^n E_a[a_{n+1} | D_k^n]\right] \qquad \text{(C3)}$$

Proof. Theorem 9 and definitions.

Eq. (C3) is, in general, tedious to evaluate. As this equation stands, the probability of the union of k overlapping events would have to be evaluated. Thus we try an approximation such that Eq. (C3') holds:

$$E_a\left[D_k^n\, E_a[a_{n+1}|D_k^n]\right] = E_a[D_k^n]\, E_a[a_{n+1}] \qquad (C3')$$

However, Eq. (C3') is equivalent to the assumption that the random matrices $A(j,j+1;\omega)$ are independent. Proceeding with this simplifying assumption, we get the following nonstationary MC, $\{A'(k)\}$, where

$$A'(k) = \left(P_{aa}(k)I_2 + P_{ab}(k)A_2\right) \qquad (D3)$$

Concerning this MC, we have

Theorem 10. The nonstationary MC whose matrices are given by Eq. (D3) is strongly ergodic. Its limit is expressed by

$$\lim_n \prod_{k=1}^{n} A'(k) = L_2$$

where the strong convergence is in the sense of the norm supremum (or any norm equivalent to it in finite dimensional Euclidean space).

Proof. Each of the matrices has the unique eigenvector $\underline{e}_2$ with eigenvalue 1. From [6.10] and Theorem 8, the nonstationary MC is strongly ergodic with the above limit since

$$\lim_k A'(k) = \bar{A}_2,$$

where the limit is taken with respect to one of the above norms.

The nonstationary MC in Theorem 10 is the approximation which is erroneously stated to be equivalent to the (2)-MC with the expectation operator $E_a[\cdot]$. To get some idea of the relationship between $E_a[A(k,\omega)]$ and $A'(1,k)$, we prove

Proposition 2. $P_{aa}(j)$ is a monotonically non-decreasing function.

Proof. Recalling the stochastic sequence $W(\cdot)$ from Definition 1 and letting T_{ab} be a sojourn time in a until exit to b, we have

$$\begin{aligned}\Delta P_{aa}(n) &= P_{aa}(n+1) - P_{aa}(n) \\ &= -Q_{ab}(n) + \sum_{j=1}^{\infty} \Big\{P[W_j(a) = n] \\ &\qquad - P[W_j(a) + T_{ab} = n]\Big\} \qquad (1)\end{aligned}$$

But the expression inside the summation sign in Eq. (1) is

$$-P[W_j(a) + T_{ab} = n \text{ and } T_{ab} \neq 0] \leq 0 \qquad (2)$$

From (1) and (2), the Proposition follows.

Corollary 3. The coefficients of the nonstationary MC are all less than or equal to the corresponding ones of the expected value of the path dependent model.

Proof. Abbreviating $P_{aa}(\cdot)$ by $P(\cdot)$, Proposition 2 shows that

$$P(j_1)P(j_2)\cdots\cdots P(j_s) \leq P(j_1)P(j_2-j_1)\cdots\cdots P(j_s-j_{s-1})$$

Each side is a general term of the two models, the LHS coming from the nonstationary MC model and the RHS coming from the path dependent one.

3.3 Connections. The transition matrix $\bar{A}_2$ in Section 3.1 is clearly equal to

$$\text{strong-lim}_k A'(k).$$

The connection between the nonstationary MC and the average of the path dependent model has already been examined; the former is obtained from the latter upon assuming the independence of the one step random matrices. Nonstationary MC's also arise in Chapter 4 but they are more related to the SMC reduction in Theorem 1 than to the reduction given in Theorem 2.

4.0 DIRECT PRODUCTS AND MULTICHARACTERISTIC PLANS. In this chapter A_k will denote the transition matrix of the kth CSP in a serial plan. The plan variables and parameters will also be indexed in the same manner (e.g., p_k, q_k, f_k, and i_k for CSP-1). I_k will denote the identity matrix of rank i_k. We will use properties of direct products without detailed comment (see [6.8]).

4.1 (2)-Serial CSP-1. The direct product of two matrices is given by

Definition 10. Let A and B be nxm and rxs matrices, respectively. Then the direct product of A and B is the nrxms matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ & \cdots & & \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

(with some abuse of notation in using B rather than its entries). A direct product is sometimes referred to as a Kronecker product in the case of matrices and an (algebraic) tensor product when the factors are explicitly linear operators.

Given the (2)-MC model in Figure 3, Chapter 2, Definition 10 can be used to express its transition matrix in a compact form which is given by the third equation below. By construction, the (2)-MC matrix can be written as

$$A_{12} = \begin{bmatrix} p_1 I_2 & q_1 A_2 & 0 \cdots & 0 \\ p_1 I_2 & 0 & q_1 A_2 \cdots & 0 \\ & & \cdots & \\ p_1 I_2 & 0 & 0 \cdots & q_1 A_2 \\ \delta_1 I_2 & 0 & 0 \cdots & \beta_1 A_2 \end{bmatrix}$$

Using some simple properties of direct products, we can rewrite the above matrix as

$$A_{12} = \begin{bmatrix} p_1 I_2 & 0 \cdots 0 \\ p_1 I_2 & 0 \cdots 0 \\ \cdots & \\ p_1 I_2 & 0 \cdots 0 \\ \delta_1 I_2 & 0 \cdots 0 \end{bmatrix} - \begin{bmatrix} p_1 A_2 & 0 \cdots 0 \\ p_1 A_2 & 0 \cdots 0 \\ \cdots & \\ p_1 A_2 & 0 \cdots 0 \\ \delta_1 A_2 & 0 \cdots 0 \end{bmatrix}$$

$$+ \begin{bmatrix} p_1 A_2 & q_1 A_2 & 0 \cdots 0 \\ p_1 A_2 & 0 & q_1 A_2 \cdots 0 \\ & \cdots & \\ p_1 A_2 & 0 & 0 \cdots q_1 A_2 \\ \delta_1 A_2 & 0 & 0 \cdots \beta_1 A_2 \end{bmatrix}$$

$$= C_1 \otimes (I_2 - A_2) + A_1 \otimes A_2$$

where

$$C_1 = (p_1)\begin{bmatrix} 1\ 0\ \cdots\ 0 \\ 1\ 0\ \cdots\ 0 \\ \cdots \\ 1\ 0\ \cdots\ 0 \\ f_1\ 0\ \cdots\ 0 \end{bmatrix} \quad \text{(an } i_1 \times i_1 \text{ matrix)}$$

Concerning A_{12} we have

Theorem 11. A_{12} is aperiodic, irreducible, and finite. Moreover, if $\underline{e}_j$ is the long run probability vector (pv) of A_j, $j = 1,2$, then $\underline{e}_1 \otimes \underline{e}_2$ is the long run probability vector (pv) of A_{12}.

Proof. 1) Refer to Figure 3, Chapter 2. The state (0,0) is aperiodic. It is straightforward but tedious to verify that (0,0) can be reached from any state and that from (0,0) any state can be reached. Thus the matrix is irreducible. Being irreducible and having one aperiodic state (0,0) imply aperiodicity for the matrix. Finally, it is trivial that the matrix is finite since its direct product components are. 2) To prove the second part of Theorem 11, we use the fact that a finite, aperiodic, and irreducible MC matrix has an unique eigenvector with eigenvalue 1. By assumption,

$$\underline{e}_j A_j = \underline{e}_j,\ \|\underline{e}_j\|_2 = 1, \text{ and } \underline{e}_j \text{ is unique,}$$

where $\|\underline{v}\|_2 = \sum_s |v_s|$ and $j = 1,2$. Thus we have,

$$\begin{aligned}(\underline{e}_1 \otimes \underline{e}_2) A_{12} &= (\underline{e}_1 \otimes \underline{e}_2)\{C_1 \otimes (I_2 - A_2)\} \\ &\quad + (\underline{e}_1 \otimes \underline{e}_2)\{A_1 \otimes A_2\} \\ &= \underline{e}_1 C_1 \otimes \underline{e}_2 (I_2 - A_2) + \underline{e}_1 A_1 \otimes \underline{e}_2 A_2 \\ &= \underline{e}_1 C_1 \otimes 0 + \underline{e}_1 \otimes \underline{e}_2 \\ &= \underline{e}_1 \otimes \underline{e}_2\end{aligned}$$

Moreover the entries of $\underline{e}_1 \otimes \underline{e}_2$ are

positive and add to one by the definition of the direct (or tensor) product of two vectors (of course "$\underline{e}_1 \otimes \underline{e}_2$" requires, by Definition 10, that $\underline{e}_1$ be $i_1 \times 1$ and $\underline{e}_2$ be $1 \times i_2$). The uniqueness of a long run pv finishes the proof.

We now turn to the investigation of the marginal AFI(∞) for the second plan.

Theorem 12. For the second plan in (2)-serial CSP-1, the marginal AFI(∞) is given by

$$AFI_2(\infty) = 1 - \nu_2 \alpha_{b_1} e_{2i_2}$$

(α_{b_2} is used as shorthand since no SMC reduction is used in the proof.)

Proof.

$$\underline{e}_1 \otimes \underline{e}_2 = \left[(e_{1k})(e_{2j})\right]$$

by definition.

$$1\text{-}AFI_2(\infty) = \lim_N \frac{1}{N} \left\{ \sum_{k=1}^{N} \sum_{j=1}^{i_1} C_{(j,i_2)}(\omega_1, \omega_2; k) \right\}$$

[a.e.], by definition,

$$= \nu_2 \left(\sum_{j=1}^{i_1} e_{1j} \right) e_{2i_2} \qquad \text{[a.e.]}$$

$$= \nu_2 \alpha_{b_1} e_{2i_2}$$

Except for three comments, the theorem is finished.

If sampling begins with state (0,0) with probability one or with $\nu = \underline{e}_1 \otimes \underline{e}_2$, then operating on the characteristic functional,

$C_{(j,i_2)}(\omega_1, \omega_2; k)$, by $E[\cdot \mid S]$, where

S = "$M_{12}(0) = (0,0)$" or "$\underline{e}_1 \otimes \underline{e}_2$", allows the dropping of "[a.e.]." Secondly, $\omega_k = \text{Proj}_k(\omega)$, $k = 1,2$. Finally, the definition of the functional implies that we are considering the identity matrix as a legitimate, but virtual, transition matrix; this viewpoint has been mentioned in Definition 5 and after Definition 8, Chapter 3.

We see from Theorem 12 that the formula

$$1 - \nu_2 \alpha_{b_1} e_{2i_2}$$

is the average number of units which are actuall inspected for the second

characteristic. Thus, for the second plan in tandem, "not inspected" is not equivalent to "sampled" because of the control exerted by the first plan on the second (recall the two interpretations of the identity matrix in Definition 5). In other words, 1-AFI$_2(\infty)$ is the average fraction not inspected whereas 1-AFI(∞, p_2, i_2, f_2) is the average sampled (equal to $\nu_2 e_{2 i_2}$).

Before leaving this section, an alternate "proof" of Theorem 11 will be given which will, in addition, give some insight into the transient behavior of the (2)-serial CSP-1 model. If a (2)-MC pv can be expressed as the direct product of two pv's (one for each plan), then such a pv will be called separable. Given that the (2)-MC starting pv (initial pv) is separable, define the pv's $\underline{x}^0$ and $\underline{y}^0$ as the initial pv's for the first and second plans, respectively. Defining a vector as a unity vector iff each of its entries is unity, further define $\underline{a(k)}$ as the $1 \times i_k$ unity vector, $k = 1,2$, and $\underline{a(k)}^t$ as its transpose ($i_k \times 1$ unity vector). Let $\{\underline{u}^r\}$ and $\{\underline{v}^r\}$ be the vectors $\{\underline{x}^0 A_1^r\}$ and $\{\underline{y}^0 A_2^r\}$, respectively (if the meaning is clear, we won't use the notation $\underline{y}^t$ for a row (column) vector corresponding to the column (row) vector $\underline{y}$). Then*

$$\left(\underline{x}^1 \otimes \underline{y}^1\right)\underline{a(2)} = \left(\underline{x}^0 \otimes \underline{y}^0\right)\left[C_1 \otimes I_2 + \left(A_1 - C_1\right) \otimes A_2\right] \underline{a(2)}$$

$$= \begin{pmatrix} s_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} u_1^1 - s_0 \\ u_2^1 \\ \vdots \\ u_{i_1}^1 \end{pmatrix}$$

$$= \underline{u}^1$$

where $s_0 = 1 - \nu_1 x_{i_1}^0$, and, in the same way,

$$\underline{a(1)}^t(\underline{x}^1 \otimes \underline{y}^1) = s_0 \underline{y}^0 + (1 - s_0)\underline{y}^1$$

*Other ways of writing $\underline{x} \otimes \underline{y}$ are: $\underline{x} \cdot \underline{y}^t$, $\underline{x}$ and $\underline{y}$ column vectors, $|\underline{x}\rangle\langle\underline{y}|$, and just

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{i_1} \end{pmatrix}(y_1, \text{---}, y_{i_2}) = \begin{bmatrix} x_1 y_1 \text{---} x_1 y_{i_2} \\ \text{---} \\ x_{i_1} y_1 \text{---} x_{i_1} y_{i_2} \end{bmatrix}$$

In all four notations, the result is an outer product which is a matrix.

$$= \underline{y}^0 \left(g_0 \, I_2 + (1-g_0) \, A_2 \right).$$

However, $(\underline{x}^1 \otimes \underline{y}^1) \, \underline{a}(2) = \underline{x}^1$ and $\underline{a}(1)^t \, (\underline{x}^1 \otimes y^1) = \underline{y}^1$

which, combined with the above, gives

$$\underline{x}^1 = \underline{u}^1 \quad \text{and} \quad \underline{y}^1 = \underline{y}^0 \left(g_0 \, I_2 + (1-g_0) \, A_2 \right).$$

Repeating with $\underline{x}^2 \otimes \underline{y}^2$,

$$\underline{x}^2 = \underline{u}^2 \quad \text{and} \quad \underline{y}^2 = \underline{y}^1 \left(g_1 \, I_2 + (1-g_1) \, A_2 \right)$$

where $g_1 = 1-\nu_1 x^1_{i_1}$. In general, by induction, we have

$$\underline{x}^{r+1} = \underline{u}^{r+1} \text{ and } \underline{y}^{r+1} = \underline{y}^r \left(g_r \, I_2 + (1-g_r) \, A_2 \right)$$

where $g_r = 1-\nu_1 x^r_{i_1}$. Thus, for a separable pv, the first plan's pv propagates according to the plan's own individual structure. In contrast, the second plan's pv propagates as a nonlinear function (because of "$1-g_r$") of vectors (pv's) arising from both plans and dependent on ν_1. The relevance to Theorem 11 arises from the observation that for all practical purposes, any pv for the model can be considered separable even though theoretically, there are nonseparable pv's whose transfinite cardinal number is equal to that of the set of all linear functions from the unit interval to itself (loosely speaking, there are an infinite number of ways to factor a real number). In particular, $\underline{e}_1 \otimes \underline{e}_2$ is not only separable (by construction) but self replicating.

The connection with Theorem 11 will be completed by showing convergence of $\underline{x}^r$ and $\underline{y}^r$ to $\underline{e}_1$ and $\underline{e}_2$, respectively. In the process, we will see that the model can be decomposed into a stationary ergodic and a nonstationary strongly ergodic MC thereby providing a link to the results of Chapter 3. From Chapter 3 and [6.10], the matrix

$$\prod_{j=1}^{r} A_2(j) = \prod_{j=1}^{r} \left(g_j \, I_2 + (1-g_j) \, A_2 \right)$$

$$= A_2(1,r)$$

strongly converges to L_2 since

$$\lim_j A_2(j) = (1-\nu_1 a_{2i_1}) \, I_2 + (\nu_1 a_{2i_1}) \, A_2$$

$$= (AFI(1) I_2 + (1-AFI(1)) A_2)$$

$$= A''_2 \quad , \; AFI(1) = AFI(\infty, p_1; i_1, f_1),$$

and $(A''_2)^N$ (strongly) converges to L_2 by the usual summability arguments.

Thus $\underline{y}^r$ strongly converges to $\underline{e}_2$. The decomposition of A_{12}^r into A_1^r and $A_2(1,r)$ is not surprising since the first plan doesn't depend on the second while the nonstationary MC appears because we are restricting attention to the second plan which does depend on the first.

4.2 (n)-Serial CSP-1. These plans can also be easily handled by direct products. Before proving the next theorem, some new matrices must first be provided. By extension of direct products to three or more matrices, we define the needed matrices in

Definition 11. Given (n)-serial CSP-1, the (n)-serial transition matrices are

$$A_{1n} = \begin{bmatrix} p_1 I_{2n} & q_1 A_{2n} & 0 \cdots 0 \\ p_1 I_{2n} & 0 & q_1 A_{2n} \cdots 0 \\ & \cdots & \\ p_1 I_{2n} & 0 & 0 \cdots q_1 A_{2n} \\ \delta_1 I_{2n} & 0 & 0 \cdots \beta_1 A_{2n} \end{bmatrix}$$

where

$$A_{kn} = \begin{bmatrix} p_k I_{(k+1)n} & q_k A_{(k+1)n} & 0 \cdots 0 \\ p_k I_{(k+1)n} & 0 & q_k A_{(k+1)n} \cdots 0 \\ & \cdots & \\ p_k I_{(k+1)n} & 0 & 0 \cdots q_k A_{(k+1)n} \\ \delta_k I_{(k+1)n} & 0 & 0 \cdots \beta_k A_{(k+1)n} \end{bmatrix}$$

for $2 \le k \le n-1$ (and for $k = 1$). More explicitly,

A_{kn} = the transition matrix of the (n-k+1)-serial CSP-1, consisting of CSP-1's k through n from the original (n)-serial CSP-1.

For $k = n$, $A_{nn} = A_n$. Moreover,

$I_{(k+1)n}$ = the identity matrix of rank $(i_{k+1} \cdot \cdots \cdot i_n)$, for $1 \le k \le n-2$, and of rank i_n for $k = n-1$;

that is,

$$I_{(k+1)n} = I_{k+1} \otimes \cdots \otimes I_n, \quad 1 \le k \le n-2$$

$$= I_n \qquad , \quad k = n-1.$$

Some important relationships exist for these matrices in

Theorem 13. Given the matrices in Definition 11, we have

$$A_{1n} = C_1 \otimes \left(I_{2n} - A_{2n}\right) + A_1 \otimes A_{2n}.$$

More generally,

$$A_{kn} = C_k \otimes \left(I_{(k+1)n} - A_{(k+1)n}\right) + A_k \otimes A_{(k+1)n}, \qquad 1 \le k \le n-1$$

Proof. By definition,

$$\begin{aligned} A_{(n-1)n} &= C_{n-1} \otimes \left(I_{nn} - A_{nn}\right) + A_{n-1} \otimes A_{nn} \\ &= C_{n-1} \otimes \left(I_n - A_n\right) + A_{n-1} \otimes A_n \end{aligned}$$

since $A_{nn} = A_n$ and $I_{nn} = I_n$. Backward induction on k, the second equation in the statement of this theorem, and the decomposition of the transition matrices according to Definition 11 give the result for fixed n. Backward induction can be converted into forward induction by relabelling. Double induction can also be done by varying n, keeping k fixed, and then proceeding by induction on k, keeping n fixed.

To determine the long run pv of (n)-serial CSP-1, Theorem 13 will be used, in Theorem 14, along with

Proposition 3. If B_1 and B_2 are transition matrices for two finite, aperiodic, and irreducible MC's, then $B_1 \otimes B_2$ also has all three properties. Moreover, if $\underline{e}_1$ and $\underline{e}_2$ are the long run pv's for B_1 and B_2, respectively, then $\underline{e}_1 \otimes \underline{e}_2$ is the long run pv for the matrix direct product.

Proof. It is trivial that the direct product is finite. The other properties follow from the equation

$$p^n_{(j,k)(r,s)} = (p^n_{jr})^1 (p^n_{ks})^2 .$$

Theorem 14. Given (n)-serial CSP-1 together with the long run pv's, $\underline{e}_k$, for the constituent plans ($1 \le k \le n$), the long run pv for the serial plan (model) is

$$\underline{e}_1 \otimes \underline{e}_2 \otimes \cdots \otimes \underline{e}_n$$

Proof. Using backward induction on the index k, for fixed n (and more generally, double induction on k and n) as in Theorem 13, Theorem 12 shows that $\underline{e}_{n-1} \otimes \underline{e}_n$ is the long run pv for $A_{(n-1)n}$. The first equation in Theorem 13, the equation

"$\underline{e}_1 \otimes \underline{e}_2 \otimes \cdots \otimes \underline{e}_n = (\underline{e}_1) \otimes (\underline{e}_2 \otimes \cdots \otimes \underline{e}_n)$",

the fact that A_{1n} is a MC matrix, and Proposition 3 applied to $A_1 \otimes A_{2n}$ suffice to finish the proof.

As an example of Theorem 14, consider (3)-serial CSP-1. The "Go" probabilities for a transition in the third plan are:

q_1q_2, $q_1\beta_2$, β_1q_2, and $\beta_1\beta_2$.

The "Stop" probabilities for a (virtual) transition in the third plan are:

P_1, q_1P_2, $q_1\delta_2$, δ_1, β_1P_2, and $\beta_1\delta_2$.

The matrices are:

$$A_{13} = C_1 \otimes (I_{23} - A_{23}) + A_1 \otimes A_{23} \qquad \text{(A4)}$$

$$A_{23} = C_2 \otimes (I_3 - A_3) + A_2 \otimes A_3 \qquad \text{(B4)}$$

Rearranging Eqs. (A4) and (B4), expanding the "23" identity matrix, and substituting the rearranged Eq. (B4) into the altered Eq. (A4) yield

$$A_{13} = C_1 \otimes I_2 \otimes I_3 + (A_1 - C_1) \otimes C_2 \otimes I_3$$
$$+ (A_1 - C_1) \otimes (A_2 - C_2) \otimes A_3.$$

Looking at this last equation and the "Go" and "Stop" probabilities, the first term of the equation is the "Stop" matrix for transitions in the second and third plans together while the second term is the "Stop" matrix for transitions in the third plan alone. The third term is, of course, the "Go" matrix for all three plans together.

We investigate, at this point, an alternate "proof" for Theorem 14 analogous to the one given for Theorem 11. First of all, we derive the recursions and the decomposition which result from the assumption of an initial separable pv for (3)-serial CSP-1. The extension of the results to (n)-serial CSP-1 is then easily obtained.

Let $\underline{x}^0 \otimes \underline{y}^0 \otimes \underline{z}^0$ be an initial pv for the (3)-serial CSP-1 model. Furthermore, define the following three sequences of vectors:

$$\underline{u}^r = \underline{x}^0 A_1^r, \ \underline{v}^r = \underline{y}^0 A_2^r, \text{ and } \underline{w}^r = \underline{z}^0 A_3^r .$$

Also define $\underline{a(2,3)}$ to be the unity $1 \times (i_2 i_3)$ vector, and $\underline{a(k)}$ the unity $1 \times i_k$ vector, for $1 \le k \le 3$. Rewrite the equations for A_{13} and A_{23} as

$$A_{13} = C_1 \otimes I_{23} + (A_1 - C_1) \otimes A_{23} \qquad \text{(C4)}$$

and

$$A_{23} = C_2 \otimes I_3 + (A_2 - C_2) \otimes A_3 \qquad \text{(D4)}$$

From Eq. (C4), we have

$$\underline{x}^1 \otimes (\underline{y}^1 \otimes \underline{z}^1) = \underline{x}^0 C_1 \otimes (\underline{y}^0 \otimes \underline{z}^0) + (\underline{u}^1 - \begin{pmatrix} 1-\nu_1 x^0_{i_1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}) \otimes (\underline{y}^0 \otimes \underline{z}^0 A_{23})$$

Since the components of $\underline{y}^1 \otimes \underline{z}^1$ add to one, we have from the RHS of the last equation

$$(\underline{x}^1 \otimes (\underline{y}^1 \otimes \underline{z}^1))\, \underline{a(2,3)} = \begin{pmatrix} g_0(1) \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \underline{u}^1 - \begin{pmatrix} g_0(1) \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$= \underline{u}^1, \quad g_0(1) = 1-\nu_1 x^0_{i_1}$$

Thus, as before we have $\underline{x}^1 = \underline{u}^1$. Since the components of $\underline{x}^1$ add to one, we also have

$$\underline{a(1)}^t (\underline{x}^1 \otimes (\underline{y}^1 \otimes \underline{z}^1)) = g_0(1)\, (\underline{y}^0 \otimes \underline{z}^0) + (1-g_0(1)) \left(\underline{y}^0 \otimes \underline{z}^0 A_{23}\right)$$

$$= \underline{y}^1 \otimes \underline{z}^1$$

Using Eq. (D4) to evaluate the second factor of the second term of the RHS of the second to last equation gives the following string of manipulations.

$$(\underline{y}^0)^1 \otimes (\underline{z}^0)^1 = (\underline{y}^0 \otimes \underline{z}^0)\, A_{23}$$

$$((\underline{y}^0)^1 \otimes (\underline{z}^0)^1)\underline{a(2)} = (\underline{y}^0 \otimes \underline{z}^0)\, A_{23}\, \underline{a(2)} \quad \text{or}$$

$$(\underline{y}^0)^1 = \underline{y}^1 \text{ (as with the (2)-serial case)}$$

$$\underline{a(3)}^t((\underline{y}^0)^1 \otimes (\underline{z}^0)^1) = \underline{a(3)}^t \left((\underline{y}^0 \otimes \underline{z}^0)\, A_{23}\right) \quad \text{or}$$

$$(\underline{z}^0)^1 = g_0(2)\, \underline{z}^0 + (1-g_0(2))\underline{w}^1$$

$$= \underline{z}^0 \left(g_0(2)\, I_3 + (1-g_0(2))\, A_3\right)$$

where $g_0(2) = 1-\nu_2 y^0_{i_2}$. Finally,

$$(\underline{y}^1 \otimes \underline{z}^1)\, \underline{a(2)} = \underline{y}^0 \left(g_0(1)\, I_2 + (1-g_0(1))\, A_2\right)$$

$$= \underline{y}^1$$

and

$$\underline{a(3)}^t(\underline{y}^1 \otimes \underline{z}^1) = \underline{z}^1$$

$$= \underline{z}^0 \left(g_0(1,2)\, I_3 + (1-g_0(1,2))\, A_3\right)$$

where $1-g_0(1,2) = (1-g_0(1))(1-g_0(2))$. Proceeding by induction,

$$\underline{x}^{r+1} = \underline{u}^{r+1},\ \underline{y}^{r+1} = \underline{y}^r\, A_2''(r),\ \text{and}\ \underline{z}^{r+1} = \underline{z}^r\, A_3''(r)$$

where

$$A_2''(r) = g_r(1)\, I_2 + (1-g_r(1))\, A_2,\ g_r(1) = 1-\nu_1 x_{i_1}^r$$

$$A_3''(r) = g_r(1,2)\, I_3 + (1-g_r(1,2))\, A_3,\ 1-g_r(1,2) = (1-g_r(1))(1-g_r(2))$$

and $g_r(2) = 1-\nu_2 y_{i_2}^r$. Furthermore,

$$\underline{y}^r = \underline{y}^0 \prod_{j=1}^{r-1} A_2''(j) \text{ and } \underline{z}^r = \underline{z}^0 \prod_{j=1}^{r-1} A_3''(j)$$

Therefore once again $\underline{y}^r$ (strongly) converges to $\underline{e}_2$, $g_r(1)$ converges to $1-\nu_1\alpha_{2i_1}$, and $g_r(2)$ converges to $1-\nu_2\alpha_{2i_2}$. It follows then that

$1-g_r(1,2)$ converges to $(1-\nu_1\alpha_{2i_1})(1-\nu_2\alpha_{2i_2})$

and that

$A_3''(r)$ (strongly) converges to $(1-\text{AFI}(1)\cdot\text{AFI}(2))I_3 + \text{AFI}(1)\cdot\text{AFI}(2)\, A_3$.

From the last statement, the recursive relationships, summability theory, and [6.10], $\underline{z}^r$ converges to $\underline{e}_3$.

In summary, for an initial <u>separable</u> pv for (3)-serial CSP-1, all three components individually converge to their long run pv's which are independent of one another. Moreover, analogous to A_{12}, A_{13} decomposes into one ergodic stationary MC and two strongly ergodic nonstationary MC's, the third depending on the first two.

The vector approach can be generalized to (n)-serial CSP-1 as an alternate "proof" for Theorem 14. However, the major reason for the vector approach is to obtain recursions and the manner of convergence. By induction, one can now easily show the following recursions and decomposition for A_{1n} with an initial separable pv $= \underline{x}_1^0 \otimes \cdots \otimes \underline{x}_n^0$. An outline of the results is given below.

$$\underline{x}_s^{r+1} = \underline{x}_s^r\ A_s''(r),\ 1 \le s \le n$$

where

$$A_s''(r) = g_r(1,2,\cdots,s-1)I_s + (1-g_r(1,2,\cdots,s-1))\, A_s$$

and

$$1-g_r(1,2,\text{---},s-1) = \prod_{k=1}^{s-1} (1-\nu_k(\underline{x}_k^r)_{i_k})$$

(and $g_r(0) = 1$). Then taking limits, we have a "proof" for Theorem 14. In general then, A_{1n} decomposes into an ergodic stationary MC and (n-1) strongly ergodic nonstationary MC's of increasing dependence on the elements of all the preceding MC's.

We now deal with $AFI_n(\infty)$ in

Theorem 15. For (n)-serial CSP-1,

$$1-AFI_n(\infty) = \nu_n(\alpha_{b_1}\alpha_{b_2}\cdots\alpha_{b_{(n-1)}})e_{ni_n}$$

(again α_{b_j} is shorthand, $1 \leq j \leq n-1$).

Proof. From Theorem 14,

$$\underline{e}_1 \otimes \text{---} \otimes \underline{e}_n = \left[(e_{1j})(e_{2k})\cdots(e_{ns})\right]$$

The rest of the proof follows the logic of Theorem 12. For example, the functional is

$$\frac{1}{N}\sum_{k=1}^{N}\sum_{J} C_{(J,i_n)}(\omega_1(n-1),\ \omega_n;k)$$

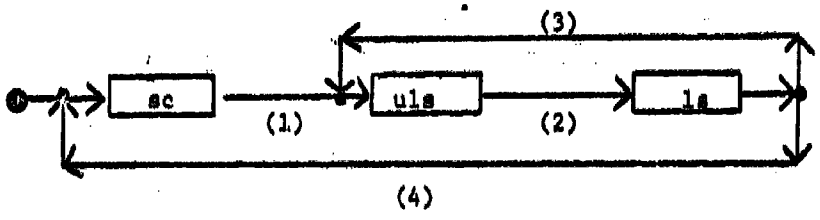
where J is the set of (n-1)-tuples of indices varying in a manner such that the rth index varies between 1 and i_r, $1 \leq r \leq n-1$.

The same comments made about $AFI_2(\infty)$ and $AFI(\infty,p_2;i_2,f_2)$ also apply to $AFI_n(\infty)$ and $AFI(\infty,p_n;i_n,f_n)$.

4.3 (n)-Serial CSP. An example of a CSP, different from CSP-1, is CSP-2 given in Figure 4; the limited sampling phase (abbr. ls) requires sampling at some frequency and, in addition, has a "clearance" number (for successive, but not consecutive, k nondefective inspected items). In a sense, the ls is a combination of the sc and uls phases.

Figure 4

Block Diagram of CSP-2



sc and uls = as in Figure 1

ls = limiting sampling phase

Arrows (1) and (2): As in Figure 1

Arrow (3): If k units are successively inspected and found to be defect free

Arrow (4): If the jth unit inspected is found to be defective, $1 \leq j \leq k$

The ls phase can be looked upon as consisting of k MC states. Further, each state, SLj, has transitions to HO and SL(j+1) (or to Si for j = k), given in the z transform mode, as follows (see [6.2] for further details)

SLj to SL(j+1) or Si given by "$\lambda/(z-v)$", $\lambda = fq$

SLj to HO given by "$\delta/(z-v)$", $\delta = fp$

As an example, consider the (2)-serial CSP given by a CSP-2 followed by a CSP-1 (the reverse order is easy since then the component matrix A_2 is just the transition matrix for CSP-2). The matrix for the total plan is

$$A_{12}(2,1) = C_1(2) \otimes \left(I_2 - A_2(1)\right) + A_1(2) \otimes A_2(1)$$

Dropping indices on the individual probabilities, those matrices used on the RHS above which come from use of CSP-2 are

$$A_1(2) = \begin{bmatrix} G_{ii} & G_{ki} \\ G_{ik} & G_{kk} \end{bmatrix}$$

where

$$G_{ii} = \begin{bmatrix} p & q & 0 & \cdots & 0 \\ p & 0 & q & \cdots & 0 \\ & \cdots & & & \\ p & 0 & 0 & \cdots & q \\ 0 & 0 & 0 & \cdots & \beta \end{bmatrix}, \quad G_{kk} = \begin{bmatrix} \nu & \lambda & 0 & \cdots & 0 \\ 0 & \nu & \lambda & & 0 \\ & & \cdots & & \\ 0 & 0 & 0 & \cdots & \nu \end{bmatrix}$$

$$G_{ik} = \begin{bmatrix} \delta & 0 & 0 & \cdots & 0 \\ \delta & 0 & 0 & \cdots & 0 \\ & & \cdots & & \\ \delta & 0 & 0 & & 0 \end{bmatrix}, \quad G_{ki} = \begin{bmatrix} & & 0 & & \\ & & & & \\ \delta & 0 & 0 & \cdots & 0 \end{bmatrix}$$

and G_{rs} is an rxs matrix. Also

$$C_1(2) = (p) \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ & & \cdots & \\ 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots f \cdots & 0 \\ f & 0 & \cdots & 0 \\ & & \cdots & \\ f & 0 & \cdots & 0 \end{bmatrix}$$

where the "f" not in column 1 is in (col, row) = (i+1,i). Formally, the analysis can proceed in a manner entirely analogous to that done in Section 4.2. For an initial separable pv, the decomposition of this (2)-serial CSP into a stationary MC and a nonstationary MC also holds. More generally, such decompositions, analogous to the one which holds for (n)-serial CSP-1, hold for any (n)-serial CSP.

4.4 Variant Multicharacteristic Plans. The first plan that would seem a natural variant is one whose MC matrix is given by

$$\tilde{A}_{12} = A_1 \otimes A_2$$

With this plan, the state determinations for each component are independent of one another. By Proposition 3, the above matrix is irreducible, finite, and aperiodic with long run pv $\underline{e}_1 \otimes \underline{e}_2$.

Another possible variant is given in Figure 5.

Figure 5

Variant Multicharacteristic Plan

States: Same as in Figure 3

Transitions ((kj)+1 may be ij for = 1,2):

State		State	Probability
(k1,k2)	⟹	((k1)+1,(k2)+1)	q_1q_2
(i1,k2)		(i1,(k2)+1)	β_1q_2
(k1,i2)		((k1)+1,i2)	$q_1\beta_2$
(i1,i2)		(i1,i2)	$\beta_1\beta_2$
Any of above +		(0,0)	1-Prob(state)

The transition rules in Figure 5 can be restated: transitions take place iff both characteristics are each either inspected and found nondefective or sampled. If we let $i_1 = i_2 = i$, this plan has one ergodic class given by the diagonal ordered pairs: $\{(j,j) \mid 0 \le j \le i\}$; all other states are transient. Moreover, if the inspection starts off with the state (0,0), we then have a plan equivalent to CSP 1 with $\bar{p} = 1-q_1q_2$ and $\bar{\beta} = \beta_1\beta_2$. However, with this plan, marginal AFI has no meaning because of the ambiguity expressed by $\bar{p}$ and $1-\bar{\beta}$. It is even doubtful whether the traditional AFI function would be a good measure of effectiveness for such a plan.

5.0 CONCLUSION. The motivation for this paper is Chapter 3 even though the main, workable results are contained in Chapter 4.

5.1 Chapter Three. The two models considered in Chapter 3 employ SMC reduction in an attempt to simplify the second order MC model at the end of Chapter 2 and highlight the difference between it and the (approximate) model given by the nonstationary MC. Any simplification of the (2)-MC model by using SMC reductions for both plans would probably not be worth the effort since superimposing two independent SMC's is quite a complex process in itself; here, of course, the SMC's are dependent!

If we are only interested in the long run case (ignoring the transient case which is hard to analyze anyway), SMC reduction of both plans can be used to yield a model consisting of the states $\{(a,1), (a,2), (b,1), (b,2)\}$ where the letters and numerals refer to the second and first SMC reduction, respectively, in Chapter 2. This model would replace the pdf's of states b and 1 by geometric pdf's. The conditions to be satisfied for this change are

$$\frac{1}{q_b'} = \mu_b \quad \text{and} \quad \frac{1}{q_1'} = \mu_1$$

The (q')'s are to be determined given the standard mean times μ_b and μ_1.

More results on products of random matrices may be found in [6,4] where various types of independence assumptions are invoked.

5.2 Chapter Four. One main result is Theorem 12 (and Theorem 14). As a consequence of the theorem, the expression "$\nu_2 e_{2} i_2$" has two interpretations: the average fraction sampled in the usual sense and the average fraction not inspected in the serial sense. The other main result, not formally stated in any theorem, is the decomposition of any (n)-serial CSP into a sequence of MC's, the first stationary, the remaining nonstationary.

The (2)-MC model assumes that the characteristics are independent. This condition can be relaxed if the ordered pairs remain independent but the two elements of any particular pair are allowed to be correlated. Let $Z_j = (X_j, Y_j)$ be the description of the jth unit. That is, $X_j(Y_j) = 0$ or 1 iff the first (second) characteristic is nondefective or defective, respectively. The relaxation is equivalent to the assumption that the Z_j form a Bernoulli process but that X_j and Y_j are not independent. Then, using the definitions of correlation coefficient and conditional probability ($\sigma_k^2 = p_k q_k$, $k = 1,2$) we have

$$P[Y_j = 1 | X_j = 0] = p_2' = \frac{p_2 + (r\sigma_1\sigma_2 + p_1 p_2)}{q_1}$$

and

$$P[Y_j = 0 | X_j = 0] = q_2' = 1 - p_2' .$$

Now $p_2' < p_2$ (or $> p_2$) iff $r\sigma_1\sigma_2 > 0$ (or < 0) and then iff the characteristics are positively (or negatively) correlated. In particular, if $r > 0$ (or < 0), then $AFI_2(=)$ will be smaller (or larger) than that obtained in Chapter 4. We finally note that for random variables X_j and Y_j, uncorrelated is equivalent to independence.

In the variant case, a (2)-characteristic plan is given where the very meaning of marginal AFI is nonexistent. Such a plan might be useful for cases of large positive correlation.

6.0 REFERENCES

6.1 Arp, D. L. "Expected values from a Markov Chain model of CSP-F and the z transform," Proc. 19th Conf. on Design of Experiments in Army R, D and T (ARO Rpt. 74-1), pp. 263-296.

6.2 Arp, D. L. "Semi Markov chains applied to Markov chain models of continuous sampling plans," Proc. 20th Conf. on Design of Experiments in Army R, D and T (ARO Rpt. 75-2, part 2), pp. 657-739.

6.3 Arp, D. L. "Semi Markov chains applied to Markov chain functionals partially dependent on random backward time shifts," Proc. 21st Conf. on Design of Experiments in Army R, D and T (ARO Rpt. 76-2), pp. 417-471.

6.4 Bharucha-Reid, A. T. Random Integral Equations, Academic Press, 1972.

6.5 Brugger, R. M. "Responsiveness properties of continuous sampling plans," Proc. 19th Conf. on Design of Experiments in Army R, D and T (ARO Rpt. 74-1), pp. 233-262.

6.6 Cinlar, E. "Markov renewal theory," Adv. Appl. Prob., Vol. 1 (1969), pp. 123-187.

6.7 Feller, W. "On semi Markov chains," Proc. of Nat'l. Acad. of Sci., Vol. 54 (1964), pp. 653-659.

6.8 Halmos, P. R. Finite Dimensional Vector Spaces, van Nostrand, 1958.

6.9 Hardy, G. H. Divergent Series, Oxford University Press, 1973.

6.10 Isaacson, D. L. and Madsen, R. W. Markov Chains Theory and Applications, John Wiley and Sons, 1976.

6.11 Jury, E. Theory and Application of the z-Transform Method, Krieger Publ. Co., 1973.

6.12 Rustagi, J. S. Variational Methods in Statistics, Vol. 121, Academic Press, 1976.

INTERVAL ESTIMATION FOR EMPIRICAL BAYES GENERALIZATIONS OF STEIN'S ESTIMATOR

Carl Morris*
Senior Statistician
The Rand Corporation
Santa Monica, California

ABSTRACT. The James-Stein estimator improves the expected mean square error of $k \geq 3$ independent sample means for all possible combinations of true means. In spite of this, it is not widely used in practical applications, partly because no confidence intervals accompany it. We derive interval estimates in this paper based on an uninformative prior distribution and illustrate the use and success of the method in an application. Not only is the estimator about three times as efficient as the sample mean vector in this example, but the intervals provided are 37 percent shorter while containing the true values with greater frequency than nominally claimed. The prior is used in the final section to extend the James-Stein estimator and to provide interval estimates for the case when the unknown parameters are exchangeable but the sample means have unequal variances.

1. INTRODUCTION. The James-Stein estimator (1961) of the means of $k \geq 3$ independent normal distributions is well-known for being uniformly and substantially better than the sample mean, on the basis of its expected sum of squared errors. The James-Stein estimator and its generalizations apply to many situations involving linear models, and offer mean squared error improvements over the classical estimators in many of the applications of statistics. Nevertheless, an informal poll of perhaps 150 statisticians at this conference revealed that only one (I would be a second) had ever used a Stein-like estimator in a real application.

Why? Polls of other groups of statisticians probably would yield similar results, although subjective Bayesians and ridge analysts may use related methods more frequently in actual data analysis. The reasons certainly include unfamiliarity on the part of many statisticians with the methods and the types of applications for which the James-Stein estimator in particular, and multiparameter estimation in general, is best suited. Long acceptance of the sample mean and its simplicity makes statisticians reluctant to reject it in favor of a more complicated and imperfectly understood method. Furthermore, the use of the James-Stein estimator requires making judgments about which problems to combine, which not to, and the choice of origin to shrink toward. If these judgments are not good, then the James-Stein estimator will improve on the total mean squared error of the sample mean insignificantly, and can be much worse for some coordinates. These reasons for the nonuse of the James-Stein estimator in applications are discussed more fully in Efron-Morris (1975, Secs. 1, 5).

*This work was partially supported by a grant from the U.S. Department of Health, Education, and Welfare, Washington, D.C.

Even those familiar with the James-Stein estimator often do not use it in its simplest form because the assumptions made for its derivation usually are not met. Instead, a generalization usually must be derived to estimate an appropriate origin, to account for nonnormal distributions, for unequal variances of the observations, for unknown variances of the observations, for regression situations, for multivariate data, or for another variation of the assumptions. Recent progress in providing these generalizations has not yet had much impact. Furthermore, the generalizations derived by different researchers are not always in agreement because they are not derived from any single principle. It seems to me, however, that data analysts probably will find the empirical Bayes viewpoint most useful both for identifying appropriate situations for using the James-Stein rule and its generalizations, and for deriving appropriate generalizations. For that reason the empirical Bayes viewpoint has been used in most of my papers with Professor Efron (March 1972, August 1972, March 1973, November 1973, 1975, March 1977, May 1977) on this topic.

Another deterrent to using the James-Stein estimator is that despite its ability to reduce mean squared error, no methods have been developed for estimating the precision of the estimates, or for determining confidence intervals. (Some attempts have been made by Stein (1962, 1975, 1974), but the results there are largely theoretical and asymptotic.)

The primary purpose of this paper is to provide a method for deriving interval estimates for the unknown parameters estimated in a matter similar to that of James-Stein and to illustrate the results on data. This is done in Section 2, using formal Bayesian ideas. The improper prior distribution used is not chosen subjectively, however, but is chosen because it yields an estimator similar to the James-Stein estimator, because the resulting estimator is minimax (uniformly dominating the vector of sample means) and admissible, because it should lead to conservative interval estimates, and because it results in easily computable statistics. It has been considered previously by several authors Baranchik (1964), Stein (1962), Leonard (1974).

The discussion in Section 2 is centered on the problem of estimating the true batting averages of eighteen baseball players. These data, which were used before in Efron-Morris (1975), are ideal for this work because the true values are available. The "confidence intervals" derived by the methods of Section 2 are about 37 percent shorter in this problem than those for the sample mean and they contain the true values with the proper probability. Since the true values were not chosen from the prior, the results encourage the idea that this method may be used generally. Such a recommendation must await further research.

The prior distribution also is used in Section 4 to derive a multiparameter estimator for parameters which have an exchangeable distribution, but whose sample means have markedly unequal variances. While the resulting estimates and interval estimates in this

illustration compare favorably to the sample mean, Section 4 is intended only to illustrate the use of this method. The resulting rule is known not to be minimax, however, and its properties await fuller investigation. Still, the method appears to be as good as any suggested to date for generalizing the James-Stein estimator to the case of unequal variances, and it does permit construction of interval estimates.

2. A WORKED EXAMPLE: EMPIRICAL BAYES INTERVAL ESTIMATES FOR THE BATTING AVERAGES OF EIGHTEEN BASEBALL PLAYERS. Let us consider the problem of estimating the true means $\{\theta_i\}$ of k normal distributions, having observed the independent sample means $X_1, X_2, \ldots, X_k$. Each X_i is assumed to have the same variance V which is known. Thus, given θ_i,

$$X_i \overset{ind}{\sim} N(\theta_i, V) \qquad i = 1, 2, \ldots, k. \tag{2.1}$$

The simplest version of the James-Stein estimator (1961) applies when $k \geq 3$ and requires making *a priori* guesses $\mu_1, \mu_2, \ldots, \mu_k$ at $\theta_1, \theta_2, \ldots, \theta_k$. Then θ_i is estimated by

$$\hat{\theta}_{i,JS} = \mu_i + (1 - \hat{B}_{JS})(X_i - \mu_i) \tag{2.2}$$

with

$$\hat{B}_{JS} = (k-2)V/\Sigma(X_i - \mu_i)^2. \tag{2.3}$$

The value in (2.3) determines how much X_i should be shrunk toward μ_i. Whenever $\hat{B}_{JS}$ exceeds unity, it should be replaced by 1 in (2.2)

The usual estimator of θ_i is X_i, being the best unbiased estimator, the best fully invariant estimator, the maximum likelihood, the least squares and the Gauss-Markov estimator. It is minimax with the expected sum of squared errors, the "risk," being

$$E_\theta \Sigma(X_i - \theta_i)^2/V = k. \tag{2.4}$$

The subscript θ on the expectation operator indicates that $\theta_1, \ldots, \theta_k$ are fixed and $X_1, \ldots, X_k$ vary according to (2.1). The James-Stein estimator is uniformly better by this criterion, having risk

$$E_\theta \Sigma(\hat{\theta}_{i,JS} - \theta_i)^2/V = k - (k-2)E_\theta \hat{B}_{JS}, \tag{2.5}$$

which is less than k, since $\hat{B}_{JS} > 0$ always. If $\theta_i = \mu_i$ for all i, then $E_\theta \hat{B}_{JS} = 1$ resulting in a risk equal to 2.

If the statistician prefers not to guess at the $\{\mu_i\}$, but believes $\mu_1 = \mu_2 = \ldots = \mu_k = \mu$ (say), he may estimate μ by $\bar{X} = \Sigma X_i/k$ and modify (2.2) to

$$\tilde{\theta}_{i,JS} = \bar{X} + (1 - \tilde{B}_{JS})(X_i - \bar{X}), \tag{2.6}$$

defining

$$\tilde{B}_{JS} = (k-3)/S, \qquad S = \Sigma(X_i - \bar{X})^2/V. \tag{2.7}$$

This version of the James-Stein estimator applies only if $k \geq 4$ (one degree of freedom is lost in estimating μ by $\bar{X}$), but it ordinarily would be preferred to (2.2) in applications to data. Its risk is

$$E_\theta \Sigma(\tilde{\theta}_{i,JS} - \theta_i)^2/V = k - (k-3)E_\theta \tilde{B}_{JS}, \tag{2.8}$$

dominating the risk (2.4) of the sample means. If $\theta_1 = \ldots = \theta_k$, it is easily checked from the chi-square distribution that $E\tilde{B}_{JS} = 1$ and hence that (2.8) is equal to 3. Otherwise (2.8) increases from 3 to k as $\Sigma(\theta_i - \bar{\theta})^2$ increases. Once again, it is better to modify (2.6) so that every θ_i is estimated by $\bar{X}$ in the event that $\tilde{B}_{JS} > 1$.

The estimator (2.6) was applied in Efron-Morris (1975) to the baseball data of Table 1. The observations X_i in the second column are the batting averages* of 18 batters in 1970 after 45 attempts. The variance of each X_i is known to be $V = (0.0659)^2$. The batting averages for these players during the remainder of the season, considered to be the true values" θ_i, will be presented later.

Instead of the James-Stein estimator (2.6), the one recommended in this paper for $k \geq 4$ uses

$$\hat{\theta}_i = \bar{X} + (1 - \hat{B})(X_i - \bar{X}), \tag{2.9}$$

as in (2.6) but replaces (2.7) by the smaller value

---

*Actually the values X_i in Table 1 are minor adjustments to the observed averages after 45 appearances given by $X_i = 0.4841 + 0.0659\sqrt{45} * \arcsin(2\hat{p}_i - 1)$, rounded to three significant figures. The observed average actually is $\hat{p}_i$; for example, $\hat{p}_1 = 18/45 = 0.400$ for player 1 (Roberto Clemente). The arcsin transformation stabilizes variances, as required for assumption (2.1), and the constants 0.4841 and 0.0659 are chosen so that the $\{X_i\}$ and the $\{\hat{p}_i\}$ have the same mean (0.26567) and standard deviation (0.0659). The same transformation $\theta_i = 0.4841 + 0.0659\sqrt{45}$ arcsin $(2p_i - 1)$ was made to the true values p_i, being the proportion of successes during the remainder of the season for batter i. The names of the players and other information about this problem are contained in Efron-Morris (1975).

Table 1

THE MAXIMUM LIKELIHOOD ESTIMATES (MLE), EMPIRICAL BAYES ESTIMATES (EBE), AND TWO ESTIMATES OF THE EBE RISK FOR EACH OF EIGHTEEN BASEBALL PLAYERS

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MLE	EBE				
i	X_i	$\hat{\theta}_i$	$\sigma_i^*(X)$	P_i	R_i^*	$\hat{R}_i$
1	0.395	0.308	0.046	0.203	0.491	1.738
2	0.375	0.301	0.044	0.145	0.454	1.163
3	0.355	0.295	0.043	0.097	0.424	0.685
4	0.334	0.288	0.042	0.057	0.398	0.287
5	0.313	0.281	0.041	0.027	0.379	-0.006
6	0.313	0.281	0.041	0.027	0.379	-0.006
7	0.291	0.274	0.040	0.008	0.367	-0.198
8	0.269	0.267	0.040	0.000	0.362	-0.274
9	0.247	0.260	0.040	0.004	0.365	-0.234
10	0.247	0.260	0.040	0.004	0.365	-0.234
11	0.224	0.252	0.040	0.021	0.376	-0.067
12	0.224	0.252	0.040	0.021	0.376	-0.067
13	0.224	0.252	0.040	0.021	0.376	-0.067
14	0.224	0.252	0.040	0.021	0.376	-0.067
15	0.224	0.252	0.040	0.021	0.376	-0.067
16	0.200	0.244	0.041	0.052	0.395	0.243
17	0.175	0.236	0.043	0.100	0.425	0.714
18	0.148	0.227	0.045	0.168	0.469	1.391
MEAN	0.266	0.266	0.042	0.056	0.397	0.274
STDEV	0.068	0.022	0.002	0.060	0.038	0.593

COMPUTATIONS: $k = 18$, $m = 7.5$, $V = (0.0659)^2$,

$\bar{X} = 0.26567$, $S = \Sigma(X_i - \bar{X})^2/V = 18.93244$,

$1 - \Phi(\sqrt{S}) = 6.76 \times 10^{-6}$, $e_{0.5}(S) = 3720.30214$, $e_{7.5}(S) = 6.77428$,

$\hat{B} = \frac{15}{S}(1 - 1/e_{7.5}(S)) = 0.79229 \times 0.85238 = 0.67534$.

$\hat{\theta}_i = \hat{B}\bar{x} + (1-\hat{B})x_i = 0.17941 + 0.32466\ X_i$, $P_i = (X_i-\bar{X})^2/SV = (X_i-\bar{X})^2/0.08222$,

$v = [2\hat{B} - 15(1-\hat{B})/e_{7.5}(S)]/S = 0.63178/S = 0.03337 = (0.1827)^2$,

$R_i^* = \frac{1}{18} + \frac{17}{18}(1-\hat{B}) + P_i vS = 0.36218 + 0.63178\ P_i$,

$\sigma_i^*(X) = (V * R_i^*)^{\frac{1}{2}} = 0.03966(1 + 1.7444\ P_i)^{\frac{1}{2}}$,

$\hat{R}_i = 1 - 2\frac{17}{18}\hat{B} + P_i\ S[2v + \hat{B}^2] = -0.27563 + 9.89823\ P_i$,

$\Sigma R_i^*/k = 0.39728$, $\Sigma \hat{R}_i/k = 0.27427$.

$$\hat{B} = \frac{k-3}{S}\left(1 - \frac{1}{e_m(S)}\right), \quad m = \frac{k-3}{2}, \tag{2.10}$$

where for $S = \Sigma(X_i - \bar{X})^2/V$ we have defined

$$e_m(S) = m \exp(S/2) \int_0^1 B^{m-1} \exp(-BS/2)\,dB. \tag{2.11}$$

The theory behind this estimator will be presented in Section 3. Here it will be described and its application illustrated. The function $e_m(S)$ increases with S from $e_m(0) = 1$ at $S = 0$ to infinity as $S \to \infty$. Thus $e_m(S) > 1$ always and therefore $\hat{B}$ in (2.10) shrinks X_i toward $\bar{X}$ less than the James-Stein estimator does. One can compute $e_m(S)$ by direct integration, or by using tables of the chi-square distribution, of the incomplete gamma function, or of the confluent hypergeometric function $M(a, b, z)$, Abramowitz-Stegun (1965, Chapter 13), since

$$e_m(S) = \Gamma(m+1)\left(\frac{2}{S}\right)^m \exp(S/2) \int_0^{S/2} \frac{g^{m-1} \exp(-g)\,dg}{\Gamma(m)}, \tag{2.12}$$

$$= M(1, m+1, S/2) = \Gamma(m+1) \sum_{j=0}^{\infty} (S/2)^j/\Gamma(m+1+j). \tag{2.13}$$

However, it usually is simplest to compute it recursively from

$$e_m(S) = \frac{2m}{S}\left(e_{m-1}(S) - 1\right) \tag{2.14}$$

using the initial values

$$\begin{aligned} e_1(S) &= (\exp(S/2) - 1)(2/S), \\ e_{.5}(S) &= \left(\frac{2\pi}{S}\right)^{\frac{1}{2}} \exp(S/2)\left[\Phi(\sqrt{S}) - .5\right], \end{aligned} \tag{2.15}$$

$\Phi(x)$ being the cumulative distribution function of a standard normal distribution. For large values of S, the approximation

$$1 - \Phi(\sqrt{S}) \doteq (2\pi S)^{-\frac{1}{2}} \exp(-S/2)(S+1)/(S+2) \tag{2.16}$$

may be used in (2.15), Abramowitz-Stegun (1965, p. 932). For small values of S, $e_m(S) \doteq 1 + S/2(m+1) + S^2/4(m+1)(m+2)$, from (2.13), so

$$\hat{B} \doteq \frac{m}{m+1}\left(1 - \frac{S/2}{(m+1)(m+2)}\right) \tag{2.17}$$

ignoring terms of order S^2. Hence $\hat{B}$ decreases monotonically from $m/(m+1) = (k-3)/(k-1)$ at $S = 0$ to 0 as $S \to \infty$. The reader is cautioned about the use of (2.14) for small values of S. It can be numerically unstable in such cases, and then (2.13) should be used instead.

Using the values of $\{X_i\}$ in Table 1, we calculate

$$\bar{X} = 0.26567, \quad S = \Sigma(X_i - \bar{X})^2/(0.0659)^2 = 18.93244, \quad m = (k-3)/2 = 7.5,$$

$1 - \Phi(\sqrt{S}) = 6.76 \times 10^{-6}$ from (2.16), $e_{.5}(S) = 3720.30214$ from (2.15),

$e_m(S) = e_{7.5}(S) = 6.77428$ by iteration of (2.14) seven times,

$\tilde{B}_{JS} = 15/S = 0.79229$, $\hat{B} = 0.79229 * 0.85238 = 0.67534$,

$$\hat{\theta}_i = 0.26567 + (1 - 0.67534)(X_i - 0.26567) = 0.17941 + 0.32466\, X_i.$$

In this case, (2.9) shrinks the MLE toward the grand mean only 85.238 percent as much as the James-Stein estimator (2.6) does. The values $\hat{\theta}_i$ are recorded as the empirical Bayes estimates in the third column of Table 1.

What precision should be attached to the estimates just derived? The error of estimate we will use is given in column 4 of Table 1 as $\sigma_i^*(X)$, computed as follows. Define

$$v = [2\hat{B} - (1-\hat{B})(k-3)/e_m(S)]/S \tag{2.18}$$

and

$$R_i^* = \frac{1}{k} + \frac{k-1}{k}(1-\hat{B}) + P_i v S \tag{2.19}$$

where

$$P_i = (X_i - \bar{X})^2/\Sigma(X_j - \bar{X})^2 = (X_i - \bar{X})^2/VS. \tag{2.20}$$

Then $\sigma_i^*(X)$ is defined to be

$$\sigma_i^*(X) = (VR_i^*)^{\frac{1}{2}}. \tag{2.21}$$

From the values already obtained, we compute

$$v = 0.63178/S = 0.03337, \quad R_i^* = 0.36218 + 0.63178\, P_i,$$

$$\sigma_i^*(X) = 0.0659\ (0.36218 + 0.63178\, P_i)^{\frac{1}{2}} = 0.03966(1 + 1.7444\, P_i)^{\frac{1}{2}}.$$

The values $\{P_i\}$, which are recorded in column 5 of Table 1, measure in relative terms the squared distances from the individual means to the grand mean. The precision (2.21) is better for those components i having X_i near the center $\bar{X}$ of the data. This fact is completely analogous to a similar result in linear regression, that prediction errors are smaller near the mean of the explanatory variables. Values of $\sigma_i^*(X)$ appear in

column 4 of Table 1. A player at the mean would have $\sigma_i^*(X) = 0.03966$, but player number 1 is farthest from the center with $P_1 = 0.203$, and therefore has $\sigma_1^*(X) = 0.046$, 16 percent larger. The typical value of $\sigma_i^*(X)$ is about 0.0415, or 37 percent less than the standard deviation 0.0659 of X_i. Thus, a considerable improvement in precision is claimed, equivalent to using the sample means of a sample 2.52 times as large.

Formula (2.19) is one of two estimates of the risk $(\hat{\theta}_i - \theta_i)^2/V$ of the empirical Bayes estimator (2.9). These values are given as R_i^* in column 6 of Table 1, and are less than the risk of the sample average $E_\theta(X_i - \theta_i)^2/V = 1$ for every player.

In column 7, the unbiased estimates $\hat{R}_i$ of the risks of $\hat{\theta}_i$ are given, computed from the formula

$$\hat{R}_i = 1 - 2\frac{k-1}{k}\hat{B} + P_i[2v + \hat{B}^2]S. \tag{2.22}$$

The estimator in (2.22) is the unique unbiased estimator of the squared error risk of the estimator (2.9). That is

$$E_\theta \hat{R}_i = E_\theta(\hat{\theta}_i - \theta_i)^2/V, \tag{2.23}$$

for all fixed $(\theta_1, \ldots, \theta_k)$. Summing the values of (2.23) over all k players, with (2.18) substituted in (2.22), we obtain

$$\hat{R}_+ = \Sigma \hat{R}_i = k - (k-3)[\hat{B} + (2-\hat{B})/e_m(S)]. \tag{2.24}$$

Since $\hat{R}_+ < k$ for all $(X_1, \ldots, X_k)$, and $\hat{R}_+$ is unbiased for the risk of (2.9), it follows that (2.9) is a minimax estimator of $(\theta_1, \ldots, \theta_k)$ for $k \geq 4$. That is

$$E_\theta \Sigma(\hat{\theta}_i - \theta_i)^2/V = k - (k-3)E[\hat{B} + (2-\hat{B})/e_m(S)] < k \tag{2.25}$$

for every set of values $(\theta_1, \ldots, \theta_k)$. The minimax character of (2.9) was proved by Baranchik (1964).

Clearly the values $\hat{R}_i$ in Table 1 are unreasonable, being negative estimates of a positive quantity in the central 11 of the 18 cases. With other data these estimates might look better, but they generally tend to be quite variable. The smoother values R_i^* provide more reasonable estimates of component risk, although as a group they tend to be conservative, for the following reasons. The sum of the values R_i^* can be written

$$R_+^* = \Sigma R_i^* = k - (k-3)[\hat{B} + (1-\hat{B})/e_m(S)], \qquad (2.26)$$

or using (2.24),

$$R_+^* = \hat{R}_+ + 2m/e_m(S). \qquad (2.27)$$

It follows from (2.27) that R_+^* overestimates the total risk of $(\hat{\theta}_1, \ldots, \hat{\theta}_k)$, since $\hat{R}_+$ is unbiased for this risk. For the data of Table 1, we calculate $\hat{R}_+ = 4.937$ from (2.24), $R_+^* = 7.151$ from (2.26), and $2m/e_m(S) = 2.214$. The amount $2m/e_m(S)$ that R_+^* overestimates the total risk decreases as S increases, and would tend to be smaller for most examples, where the true values are likely to be more dispersed.

How well does this analysis do? The true values* are given in Table 2, column 2. Column 3 presents the values $(\hat{\theta}_i - \theta_i)/\sigma_i^*(X)$, a distribution which ideally has zero mean and unit standard deviation. The mean of these values is -0.027, only about one-tenth of a standard deviation from that expected, the standard deviation is 0.862, meaning that the intervals are conservative. This is expected, since from (2.27),

$$E_\theta \Sigma\, \sigma_i^{*2}(X) = VE_\theta R_+^* > E_\theta \Sigma\, (\hat{\theta}_i - \theta_i)^2 \qquad (2.28)$$

and so the $\sigma_i^*(X)$ tend to be too large (by about 15 percent in this case). For comparison, the distribution of errors of X_i, relative to the standard deviation of X_i, is given in column 4 of Table 2. The mean and standard deviation of these numbers are almost exactly what is expected from a sample of 18 numbers from a N(0,1) distribution. Hence, the intervals for $\hat{\theta}_i$ in this example are *both* shorter and more conservative than those for X_i.

The signs of the MLE errors in column 4 are strongly correlated with the X_i values, because the true means θ_i have regressed markedly toward the mean, relative to the observed means X_i. Figure 1 shows this regression effect vividly, and how the $\{\hat{\theta}_i\}$ shrink the $\{X_i\}$ to produce better estimates. The dispersion of the $\{\hat{\theta}_i\}$ is even smaller than that of the true values $\{\theta_i\}$ since the ordering of the $\{\theta_i\}$ is not highly correlated with that of the $\{X_i\}$ (Spearman's rank correlation

*These really are only the batting averages for the remainder of the 1970 season, being independent estimates of the true values with standard deviation 0.0659 $(45/N_i)$, N_i given in column 7 of Table 2.

coefficient is only $\rho(\theta, X) = 0.218$ for these data). The regression to the mean effect also occurs for the EB estimates $\hat{\theta}_i$ in column 3, although it is much less pronounced. An even less conservative shrinking constant than $\hat{B}$ would be needed to eliminate the regression to the mean for these estimates and true values.

Table 2

TRUE VALUES, RELATIVE ERRORS, AND LOSSES FOR EMPIRICAL BAYES ESTIMATES (EBE) AND MAXIMUM LIKELIHOOD ESTIMATES (MLE)

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	TRUE VALUE	EBE RELATIVE ERROR	MLE RELATIVE ERROR	EBE LOSS	MLE LOSS	
i	θ_i	$\frac{\hat{\theta}_i - \theta_i}{\sigma_i^*(X)}$	$\frac{X_i - \theta_i}{\sqrt{V}}$	$\frac{(\theta_i - \hat{\theta}_i)^2}{V}$	$\frac{(X_i - \theta_i)^2}{V}$	N
1	0.346	-0.831	0.744	0.339	0.553	367
2	0.300	0.026	1.138	0.000	1.295	426
3	0.279	0.365	1.153	0.057	1.330	521
4	0.223	1.560	1.684	0.968	2.837	275
5	0.276	0.124	0.561	0.006	0.315	418
6	0.273	0.198	0.607	0.015	0.368	466
7	0.266	0.198	0.379	0.014	0.144	586
8	0.211	1.406	0.880	0.716	0.775	138
9	0.271	-0.286	-0.364	0.030	0.133	510
10	0.232	0.694	0.228	0.175	0.052	200
11	0.266	-0.343	-0.637	0.044	0.406	277
12	0.258	-0.145	-0.516	0.008	0.266	270
13	0.306	-1.334	-1.244	0.668	1.548	435
14	0.267	-0.368	-0.653	0.051	0.426	538
15	0.228	0.598	-0.061	0.134	0.004	186
16	0.288	-1.054	-1.335	0.439	1.783	558
17	0.318	-1.903	-2.170	1.540	4.709	408
18	0.200	0.609	-0.789	0.174	0.623	70
MEAN	0.267	-0.027	-0.022	0.299	0.976	369
STDEV	0.037	0.862	0.988	0.412	1.157	150

The observations of the preceding paragraph are expressed differently in Figure 2. The central dashed line is the maximum likelihood estimator (the 45 degree line); the other four dashed lines are the MLE plus or minus 1.00 and 1.96 standard deviations of X_i. These determine the classical 68 percent and 95 percent confidence intervals. Each player is plotted at his point $(X_i \quad \theta_i)$. The dashed confidence bands do very well: 12/18 of of the true values are located between the 16th and 84th percentiles; and

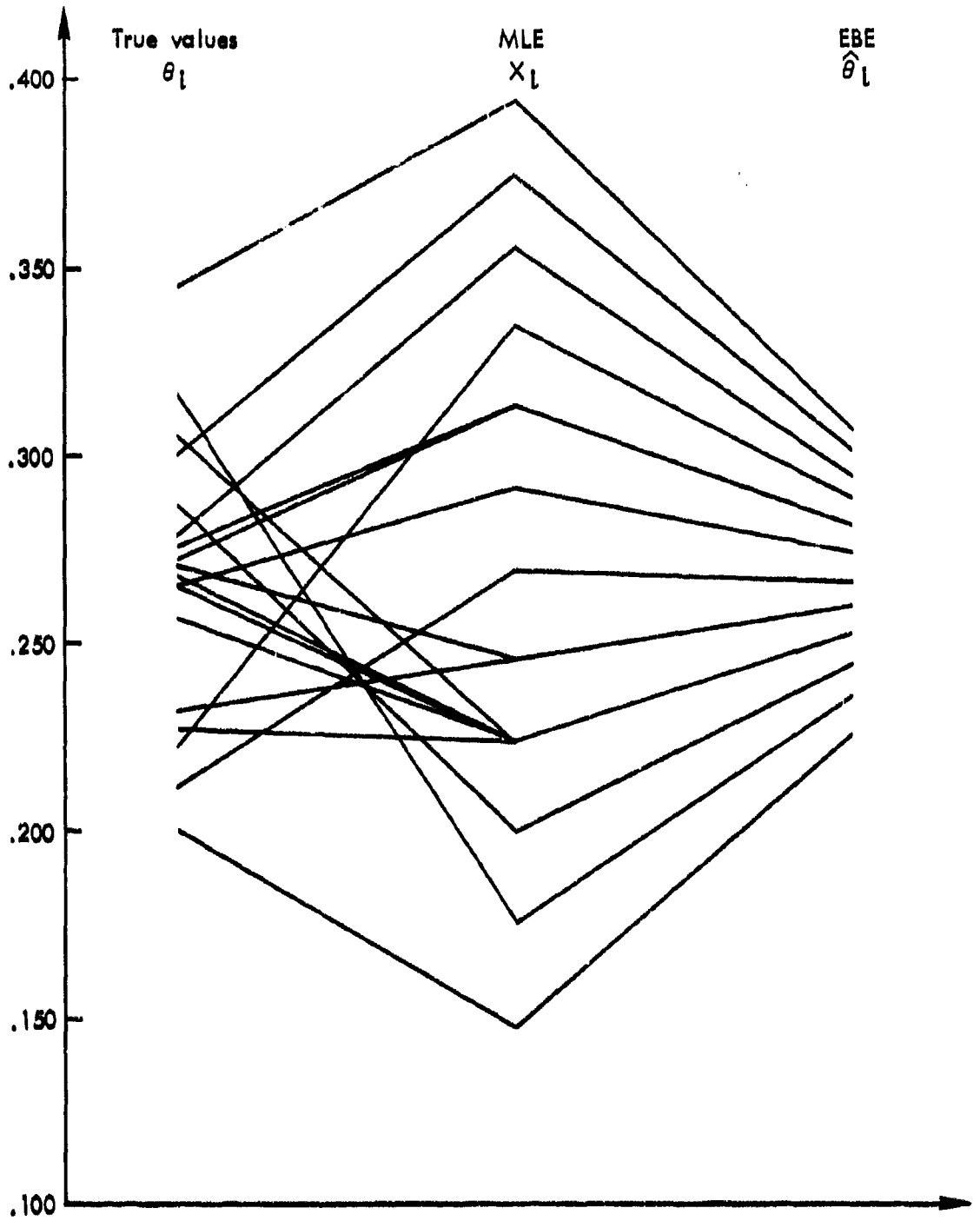


Fig. 1—True values, maximum likelihood estimates, and empirical Bayes estimates for each of 18 players illustrating regression to the mean of the true values and shrinkage of the empirical Bayes estimator

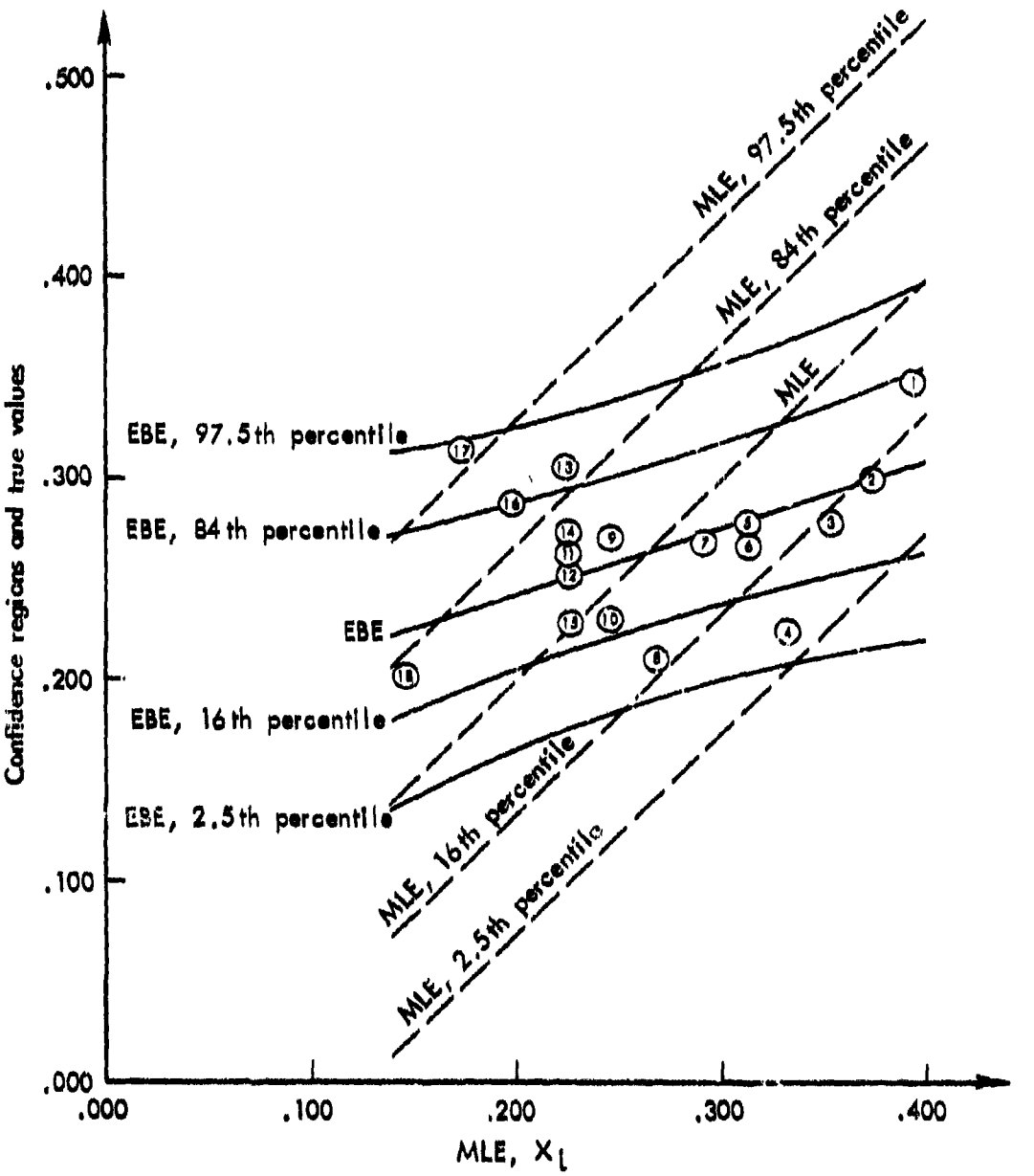


Fig. 2 — MLE X_ℓ, MLE $\pm V^{1/2}$, and MLE $\pm 1.96\ V^{1/2}$ (dashed lines) and EBE $= \hat{\theta}_\ell$, EBE $\pm \sigma_\ell^*(X)$, and EBE $\pm 1.96\ \sigma_\ell^*(X)$ (solid curves). Eighteen players plotted at (X_ℓ, θ_ℓ) using data of Tables 1, 2

17/18 are within the 95th confidence band. The middle solid line is the empirical Bayes estimator $\hat{\theta}_i = 0.17941 + 0.32466\ X_i$. This value $\pm\ \sigma_i^*(X)$ is intended to correspond approximately to 68 percent confidence, and $\pm\ 1.96\ \sigma_i^*(X)$ to 95 percent confidence. Notice that these solid line confidence bands curve to allow for greater errors at extreme values of X_i. The confidence bands are conservative in this application, in congruence with the theoretical statements made after (2.27); 13/18 = 0.722 of the true values are in the central 68 percent confidence region, and all 18 are in the 95 percent region.

An extremely interesting point raised by Figure 2 is that when the 95 percent confidence region is used to make a statistical test that the true value of a player is a specified value, then conflicting results can be obtained from the classical and empirical Bayes methods. Because it has shorter intervals, we expect the empirical Bayes methodology to reject certain true values when the MLE does not. For example, from Figure 2, a 0.500 season average cannot be rejected for player number 1 according to classical theory, but is out of the question from the empirical Bayes standpoint. (No one has ever approached such a value for a full season.) The astonishing fact is that the empirical Bayes method includes two small regions that are excluded by the classical methodology. To illustrate this, a true value of 0.318 is rejected at the 95 percent level for player number 17 (Thurmond Munson) by the classical test, but is not rejected at the same level using empirical Bayes intervals in Table 2. It turns out that 0.318 *was* Munson's true value. (And in 1976 he was voted the most valuable player in the American League!) We will not discuss this hypothesis testing problem further here, but obviously it is a worthy topic for further research.

Columns (5) and (6) show the losses incurred by the two estimators $\hat{\theta}_i$ and X_i. Only for the 10th and 15th players does $\hat{\theta}_i$ fail to improve on X_i, and in those cases the $\hat{\theta}_i$ loss is small. The empirical Bayes loss $\Sigma(\hat{\theta}_i - \theta_i)^2/V$ for the 18 players is 5.38. The sample means give 17.57, close to what is expected for 18 components, but worse by a multiple of 3.27 than 5.38. The values R_i^* and $R_i^\dagger$ from Table 1 estimate the expected value of entries in column (5) of Table 2. Since $\hat{R}_+ = 4.94$ and $R_+^* = 7.15$, $\hat{R}_+$ is closer to the combined loss $\Sigma(\hat{\theta}_i - \theta_i)^2/V = 5.38$. However the R_i^* values, being smoother estimates of $E(\hat{\theta}_i - \theta_i)^2/V$, are much closer to the individual losses $(\hat{\theta}_i - \theta_i)^2/V$ of the players than are the $\hat{R}_i$.

Do these results hold up for other samples $\{X_i\}$ from these true values $\{\theta_i\}$? A simulation was conducted to check this and to determine whether the intervals computed by $\hat{\theta}_i \pm \sigma_i^*(X)$ and $\hat{\theta}_i \pm 1.96\ \sigma_i^*(X)$ contain the true values at least 68 percent and 95 percent of the time. Using the same true values $\{\theta_i\}$ of Table 2 each time, new values of $(X_1, \ldots, X_{18})$ were randomly drawn from the normal distribution (2.1) one hundred times, with $\mathrm{Var}(X_i) = (0.0659)^2$ in all cases.

In the 1800 experiences, the true values were contained in their nominal 68 percent intervals (in $\hat{\theta}_i \pm \sigma_i^*(X)$) 74 percent of the time, and in the nominal 95 percent intervals 97.3 percent of the time. In one of the 100 cases three of the true values fell outside their nominal 95 percent intervals, in nine cases two true values fell outside, in 28 cases one fell outside, and in the remaining 62 cases all 18 of the true values were in the interval $\hat{\theta}_i \pm 1.96\ \sigma_i^*(X)$. The average shrinking value $\hat{B}$ was 0.608, and $\sigma_i^*(X)$ was typically 65 percent of $V\hat{B}$, so empirical Bayes confidence intervals were both 35 percent shorter and more conservative than those based on the sample mean.

The estimate $\{X_i\}$ had average loss 18.45 (0.75 of a standard deviation above that expected), while $\{\theta_i\}$ has 6.41, more efficient than the MLE by a factor of 2.88. In no case did $\{\hat{\theta}_i\}$ have combined loss exceeding 13.1, and its total loss never exceeded 60 percent of that of $\{X_i\}$ in any of the 100 cases.

Next consider the estimates of risk. The risk of $\{\hat{\theta}_i\}$ is close to 6.4, the average loss in the 100 simulations. The average value of $\hat{R}_+$ was 6.3, R_+^* averaged 8.3. However R_+^* was a better estimate of the total loss, which varied from case to case, than $\hat{R}_+$ in 59 of the 100 cases, and had root mean squared error 2.9 for estimating the total loss $L_+ = \Sigma(\hat{\theta}_i - \theta_i)^2/V$ (i.e., $\Sigma(R_+^* - L_+)^2/100 = (2.9)^2$ for the 100 cases), whereas $\hat{R}_+$ had an inferior root mean squared error of 3.7. The component estimates of risk R_i^* were much better than $\hat{R}_i$, as estimates of the loss $L_i = (\hat{\theta}_i - \theta_i)^2/V$. In the root mean square sense, $|R_i^* - L_i|$ averaged 0.51 while $|\hat{R}_i - L_i|$ typically was 0.78. The latter errors also were more variable from problem to problem. In only 9 of the 100 cases was the root mean square of the 18 $|\hat{R}_i - L_i|$ values smaller than the root-mean-square of the $|R_i^* - L_i|$ values.

The analysis presented in Tables 1 and 2 then is typical (although slightly on the favorable side) of what would be expected from a random draw of observed values $\{X_i\}$ from the true values $\{\theta_i\}$ of Table 2. The conclusion from the simulation for these $\{\theta_i\}$ is that in addition to substantial improvement in the risk of the sample means, the empirical Bayes estimates $\{\hat{\theta}_i\}$ of (2.9) provide much shorter confidence intervals than the classical estimator, with nominal values that are conservative.

We cannot make similar claims at this time for the confidence intervals generated by the empirical Bayes estimator for other combinations of true values, but there is reason to expect similar results if the statistician is careful to combine estimates from problems for which the true values are exchangeable (i.e., the distribution of the $\{\theta_i\}$ should

be invariant under permutations). For large $\Sigma(\theta_i-\bar{\theta})^2$ the rule (2.9) is nearly equal to the James-Stein estimator, which Stein (1962) has shown leads to approximately correct confidence sets when either S or k is large. Over all components, (2.9) is minimax, and conservative both because it shrinks less than the James-Stein rule, and because

$$\frac{E_\theta\Sigma(\hat{\theta}_i-\theta_i)^2}{E_\theta\Sigma\,\sigma_i^{*2}(X)} < 1, \tag{2.29}$$

which follows from (2.27). But the statistician who cares about each individual component really needs to know not that (2.29) holds, but that for every $i = 1, \ldots, k$,

$$E_\theta\,\frac{(\hat{\theta}_i-\theta_i)^2}{\sigma_i^{*2}(X)} \leq 1, \tag{2.30}$$

or nearly so. This can fail badly if the true values fall into distinct groups (so they could not have come from the exchangeable prior on which (2.9) is based). The most dramatic example of this failure occurs for large k when $\theta_2 = \theta_3 = \ldots = \theta_k$ and $\theta_1 = \theta_2 \pm \sqrt{kV}$. Then, although $\theta_1^{*2}(X) \doteq V$ for $i = 1$, $E_\theta(\hat{\theta}_1-\theta_1) \doteq Vk/4$. However the unbiased estimate of risk of θ_1, $\hat{R}_1$ will be close to the correct value $k/4$ and therefore is a much better estimate of risk than R_1^* in this instance. More generally, an upper bound for R_1^* is 1.5 for all k, X, achieved for $P_1 = 1$ near $S = 2k$. Thus $\sigma_1^*(X) \leq 1.225\sqrt{V}$ always, resulting in nonconservative intervals for components that are badly estimated. A limited translation modification of the estimator (2.9) would reduce this error significantly without substantially reducing the overall efficiency of the estimate (Efron-Morris, 1972). Obviously, considerable caution must be taken when applying $\hat{\theta}_i$ to components with large P_i or large $\hat{R}_i$ values. This example warns against too much reliance on the Bayesian interpretation of the estimator and illustrates why the statistician must consider the exchangeability assumption to be plausible before using either the James-Stein estimator or (2.9).

3. DERIVATION OF THE EMPIRICAL BAYES ESTIMATOR. The James-Stein rule (2.6) may be derived as an empirical Bayes estimator (see Efron-Morris, March 1973, and Efron-Morris, 1975) by assuming that the true values $\{\theta_i\}$ independently follow the same prior distribution with two unknown parameters $\mu = E\theta_i$, $A = \mathrm{Var}(\theta_i)$,

$$\theta_i \overset{\text{ind}}{\sim} N(\mu, A) \qquad i = 1, 2, \ldots, k. \tag{3.1}$$

Given $\{\theta_i\}$, the sample means have the normal distribution specified in (2.1). If μ and A are known, the Bayes estimator of θ_i for squared error loss is the posterior mean

$$E(\theta_i | X, A, \mu) = \mu + (1-B)(X_i-\mu), \tag{3.2}$$

defining

$$B = \frac{V}{V+A} . \tag{3.3}$$

The marginal distribution of $\{X_i\}$ given μ, A is obtained by integrating $\{\theta_i\}$ out of the conditional distribution (2.1) of $\{X_i\}$, obtaining

$$X_i \overset{ind}{\sim} N(\mu, V + A) \qquad i = 1, 2, \ldots, k. \tag{3.4}$$

Thus $\bar{X}$ is the usual estimator of μ, from (3.4), and can be used to replace the unknown μ in (3.2), while $S = \Sigma(X_i-\bar{X})^2/V$, being distributed as

$$S \sim \frac{1}{B} \chi^2_{k-1}, \tag{3.5}$$

because of (3.4), provides a basis for estimating B. The unbiased estimate of B is $\tilde{B}_{JS} = (k-3)/S$,

$$E\tilde{B}_{JS} = E\frac{k-3}{S} = B \tag{3.6}$$

from (3.5). Substitution of $\bar{X}$ and $\tilde{B}_{JS}$ for the unknown values μ and B in (3.2) yields the James-Stein estimator (2.6) of $\{\theta_i\}$ as an empirical Bayes estimator.

Instead of the unbiased estimate, we will derive a formal Bayes estimator of B by assuming A is uniformly distributed on $[0, \infty)$, that is, with probability element dA on $[0, \infty)$. A compelling reason for this choice is that the James-Stein estimator is the formal Bayes estimator resulting from distributing A uniformly on $[-V, \infty)$. Since it is known that A cannot be negative, being a variance, restricting it to $[0, \infty)$ leads to an estimator similar to but better than the James-Stein estimator. This prior has been studied before with μ known, by Stein (1962), by Baranchik (1964) who proved the resulting estimator is minimax, and again by Stein (1973) where he developed the unbiased estimator $\hat{R}_+$ of its risk and also observed that the rule is admissible because of a theorem of Brown (1971). Leonard (1974) discussed the prior in a Bayesian setting, and it is similar to, but not identical with, priors recommended by other Bayesians for this problem: Jeffreys (1948), Lindley-Smith (1972), Zellner-Vandaele (1975), and Good and Wallace (as interpreted by Stein (1962, p. 281). An appealing property of this prior is that it does not depend on the variance V. The estimators

of Strawderman (1971) do not share this property, which renders them inapplicable in the context of Section 4.

Using the density from (3.5) and $dA = -VdB/B^2$, the density of B given S is proportional to

$$f(B|S) = B^{m-1} \exp(-BS/2)dB \tag{3.7}$$

with $m = (k-3)/2$ on $0 \leq B \leq 1$. Therefore the formal Bayes estimate of B is

$$\hat{B} = EB|S = \frac{\int_0^1 B^m \exp(-BS/2)dB}{\int_0^1 B^{m-1} \exp(-BS/2)dB} . \tag{3.8}$$

The denominator of (3.8) is, up to a scaler multiple, the marginal density function of S (being an improper density). Integrating the numerator of (3.8) by parts once yields

$$-\frac{2}{S} \exp(S/2) + \frac{2m}{S} \int_0^1 B^{m-1} \exp(-BS/2)dB$$

and hence (3.8) simplifies to

$$\hat{B} = \frac{k-3}{S}(1 - \frac{1}{e_m(S)}) \tag{3.9}$$

with $e_m(S)$ defined in (2.10). Estimating μ by $\bar{X}$ and B by (3.9) in (3.2) yields the estimator (2.9) as an empirical Bayes estimator.

The variance v of B given S also can be obtained. We have

$$v = Var(B|S) = -2 \frac{d}{dS} EB|S = -2 \frac{d}{dS} \hat{B}. \tag{3.10}$$

Since

$$\frac{de_m(S)}{dS} = e_m(S)(1-\hat{B})/2, \tag{3.11}$$

it follows from (3.10), (3.9), and then (3.11) that

$$v = 2[\hat{B} - (1-\hat{B})m/e_m(S)]/S. \tag{3.12}$$

The unbiased estimate of component risk of any estimator of θ_i of the form $\hat{\theta}_i = \bar{X} + (1-B(S))(X_i-\bar{X})$ is, denoting $B'(S) = dB(S)/dS$ and $P_i = \Sigma(X_i-\bar{X})^2/SV$,

$$\hat{R}_i(S) = 1-2\frac{k-1}{k}B(S) + P_iS[B^2(S) - 4B'(S)] \qquad (3.13)$$

for any shrinking function B(S) which depends only on S. That is

$$E_\theta\hat{R}_i(S) = E_\theta(\hat{\theta}_i-\theta_i)^2/V. \qquad (3.14)$$

This follows from writing

$$E_\theta(\hat{\theta}_i-\theta_i)^2/V = E_\theta[(X_i-\theta_i) - B(S)(X_i-\bar{X})]^2/V$$
$$= 1 + E_\theta B^2(S)(X_i-\bar{X})^2/V - 2E_\theta\frac{X_i-\theta_i}{V}B(S)(X_i-\bar{X}),$$

and then applying Stein's formula, Stein (1973), Efron-Morris (1976), to obtain the identity

$$E_\theta\frac{X_i-\theta_i}{V}B(S)(X_i-\bar{X}) = E_\theta\frac{\partial}{\partial X_i}B(S)(X_i-\bar{X}). \qquad (3.15)$$

Formula (3.13) is obtained by computing

$$\frac{\partial}{\partial X_i}B(S)(X_i-\bar{X}) = B'(S)(X_i-\bar{X})\frac{\partial S}{\partial X_i} + B(S)(1-1/k),$$

noting that $\partial S/\partial X_i = 2(X_i-\bar{X})/V$, and collecting terms. The expression (2.22) for $\hat{R}_i$ follows from substituting (3.10) and (3.9) into (3.13).

It is interesting to note if B(S) in (3.13) is any Bayes estimator of B, computed as $B_\pi(S) = E_\pi B|S$ with π the prior density of A, then $\mathrm{Var}_\pi(B|S) = -2dB_\pi(S)/dS$ and the unbiased estimate (3.13) of risk becomes

$$\hat{R}_{i,\pi} = 1-2\frac{k-1}{k}B_\pi(S) + P_iS[B_\pi^2(S) + 2\,\mathrm{Var}_\pi(B|S)]. \qquad (3.16)$$

To compute the posterior distribution of $\{\theta_i\}$ given the data $\{X_i\}$, we need a prior distribution for μ in (3.4). This distribution is chosen to be Lebesgue (uniform) measure on $(-\infty, \infty)$, independent of the distribution on A, because it leads to the classical estimate $\bar{X}$ for μ. Assuming the normal distributions (2.1) and (3.1) for $\{X_i\}$ given $\{\theta_i\}$ and $\{\theta_i\}$ given μ and A, Bayes theorem gives

$$\mu|X, A \sim N(\bar{X}, \frac{A+V}{k}). \qquad (3.17)$$

To extend the result (3.2), denote $\theta = (\theta_1, \ldots, \theta_k)'$, $X = (X_1, \ldots, X_k)'$, $e = (1, 1, \ldots, 1)'$, and I the kxk identity matrix. The distribution of θ is

$$\theta|(X,\ \mu,\ A) \sim N_k(\mu e + (1-B)(X-\mu e),\ V(1-B)I). \tag{3.18}$$

Integrating the distribution of μ (3.17) out of (3.18) yields

$$\theta|(X,\ A) \sim N_k(\bar{X}e + (1-B)(X-\bar{X}e),\ V(1-B)I + B\frac{V}{k}ee'). \tag{3.19}$$

Finally, the distribution of $A = V(1-B)/B$ given X is given by (3.7). So integrating B out of (3.19), using (3.7), yields

$$E(\theta|X) = \bar{X}e + (1-\hat{B})(X-\bar{X}e), \tag{3.20}$$

with $\hat{B}$ given by (3.9), and the conditional covariance matrix as

$$\frac{1}{V}\,\mathrm{Cov}(\theta|X) = (1-\hat{B})I + \hat{B}ee'/k + v(X-\bar{X}e)(X-\bar{X}e)'/V \tag{3.21}$$

with $v = \mathrm{Var}(B|S)$ given by (3.12).

It is not precisely true that

$$\theta|X \sim N_k(E\theta|X,\ \mathrm{Cov}(\theta|X)) \tag{3.22}$$

has the normal distribution. But θ does have a normal distribution for every fixed B (3.19), and if either B is estimated by $\hat{B}$ without large variance, or if the normal distribution of X is considered as an added source of variation in (3.20), then the normal distribution should hold approximately in (3.22). We assumed this to produce interval estimates in Section 2. Formula (3.21) actually shows that the $\{\theta_i\}$ values are correlated, a fact not mentioned or used in Section 2. Thus (3.21) could be used to find posterior credibility ellipsoids for θ given X. Instead, Section 2 uses only the diagonal elements of (3.21)

$$\sigma_i^{*2}(X)/V = R_i^* = 1 - \frac{k-1}{k}\hat{B} + P_i vS, \tag{3.23}$$

and ignores the covariance.

In some problems the prior mean μ (3.1) may be known, and then (3.17) would be inappropriate. All the results given so far cover the case of known μ provided: $\bar{X}$ is replaced by μ throughout; k-1 and k-3 are changed to k and k-2 in (3.5), (3.6); m is changed to (k-2)/2 throughout (this is the reason for using the subscript m on e_m in (2.10)); (k-1)/k is replaced by 1 in the middle terms of (3.13), (3.16), and (3.23); (3.19) is ignored in favor of (3.18); and the ee'/k term in the middle of (3.21) is eliminated.

As stated before, the James-Stein rule is a formal Bayes estimator against the prior taking A uniform on $[-V, \infty)$. Therefore the risk estimates, interval estimates, posterior distributions, and all other quantities computed in this section can be computed for the James-Stein estimator. These results are obtained by replacing $e_m(S)$ by

infinity ($1/e_m(S) = 0$) in all formulas. Recall, for example, that $\hat{B} = \tilde{B}_{JS}(1-1/e_m(S))$, so setting $e_m(S) = \infty$ modifies $\hat{B}$ to $\tilde{B}_{JS}$. More generally, if the prior takes A uniform on $[\alpha, \infty)$, $\alpha \geq -V$, then the resulting value of $\hat{B}$ is

$$\hat{B} = EB|S = \frac{k-3}{S}\left(1 - \frac{1}{e_m(\beta S)}\right), \quad \beta = \frac{V}{V+\alpha} . \tag{3.24}$$

The James-Stein rule is obtained by letting $\beta \to \infty$ in (3.24), hence $e_m(\beta S) \to \infty$. The estimator of this paper is given by $\alpha = 0 (\beta = 1)$, while other more conservative estimators result from choices of $\alpha > 0$ $(\beta < 1)$.

4. FORMAL BAYES ESTIMATORS IN THE UNEQUAL VARIANCES CASE. Because of the success in previous sections of the formal Bayes estimator resulting from the prior

$$\theta_i|A \overset{\text{ind}}{\sim} N(\mu_i, A) \qquad i = 1, 2, \ldots, k \tag{4.1}$$

with the variance A distributed as

$$A \sim \text{Uniform } (0, \infty), \tag{4.2}$$

we use this prior again in the case where the variances of the sample means X_i are not necessarily equal. That is, (2.1) is generalized to

$$X_i|\theta_i \overset{\text{ind}}{\sim} N(\theta_i, V_i) \qquad i = 1, 2, \ldots, k \tag{4.3}$$

with the V_i known, but possibly unequal. This is the case that arises most frequently in applications. The equal variance situation rarely occurs, except in some designed experiments. We shall assume that the $\{\mu_i\}$ are known, because while they can be estimated, doing so causes the formulas of this section to become much more complicated without providing much additional insight. In most applications, however, estimating the $\{\mu_i\}$ would be worthwhile. Having assumed $\{\mu_i\}$ known, we take them to be zero without essential loss of generality, and replace (4.1) with

$$\theta_i|A \overset{\text{ind}}{\sim} N(0, A), \qquad i = 1, 2, \ldots, k. \tag{4.4}$$

By making use of Bayes' formula, and by obtaining the marginal distribution of $\{X_i\}$, (4.3) and (4.4) are equivalent to

$$\theta_i|(X_i, A) \overset{\text{ind}}{\sim} N((1-B_i)X_i, V_i(1-B_i)), \quad i = 1, 2, \ldots, k \tag{4.5}$$

$$X_i|A \overset{\text{ind}}{\sim} N(0, A + V_i), \quad i = 1, 2, \ldots, k \tag{4.6}$$

where we have defined

$$B_i = \frac{V_i}{V_i + A} . \tag{4.7}$$

Letting $S_i = X_i^2/V_i$, it follows from (4.6) that

$$S_i|A \overset{\text{ind}}{\sim} \frac{1}{B_i}\chi_1^2 \qquad i = 1, 2, \ldots, k. \tag{4.8}$$

The posterior distribution of A may be obtained from application of Bayes' formula to (4.2) and (4.6), or more simply to (4.2) and (4.8), to obtain the posterior probability element of A given $X_1, \ldots, X_k$

$$f_S(A)dA = \frac{\exp(-\frac{1}{2}\sum_{j=1}^{k}[B_jS_j - \log(B_j)])dA}{\int_0^\infty \exp(-\frac{1}{2}\sum_{j=1}^{k}[B_jS_j - \log(B_j)])dA}, \tag{4.9}$$

with each B_j a function of A given by (4.7).

Formulas (4.5) and (4.9) summarize all information relevant to a Bayesian analysis. In particular

$$\hat{\theta}_i = E\theta_i|X = (1-\hat{B}_i)X_i \tag{4.10}$$

and

$$R_i^m = \frac{1}{V_i}\mathrm{Var}(\theta_i|X) = 1-\hat{B}_i + S_iv_i, \tag{4.11}$$

where we have defined, with $B_i = B_i(A)$ given by (4.7),

$$\hat{B}_i = EB_i|S = \int_0^\infty B_i(A)f_S(A)dA \tag{4.12}$$

and

$$v_i = \mathrm{Var}(B_i|S) = \int_0^\infty B_i^2(A)f_S(A)dA - \hat{B}_i^2. \tag{4.13}$$

Although there are many methods and some tricks to help in computing the integrals (4.12), (4.13), none yield simple answers like those of the preceding sections. The simplest way to compute (4.12) and (4.13) we have found so far is to evaluate the numerator of (4.9) at a number of points (about 100, not equally spaced), then to divide these values by their sum to obtain $f_S(A)$ at those points, and finally compute the 2k integrals (4.12), (4.13) as finite sums. This is a minor task using a computer. More thought should be given to these computational issues if the method is used frequently.

The symbols (4.10) and (4.11) are the same as those used in Sections 2, 3, and retain their meanings (except that here μ is not estimated). So do

$\hat{B}_i$ and v_i, although they now vary with the component. The standard deviation of θ_i given $X = (X_1, \ldots, X_k)$ is, once again,

$$\sigma_i^*(X) = (V_i R_i^*)^{\frac{1}{2}}, \tag{4.14}$$

and this quantity again will be used to define interval estimates.

The unbiased estimates of risk are

$$\hat{R}_i = 1 - 2\hat{B}_i + S_i(\hat{B}_i^2 + 2v_i), \tag{4.15}$$

being derived by the same argument used to obtain (3.13) and (3.16). Then $\hat{R}_i$ satisfies

$$E_\theta \hat{R}_i = E_\theta(\hat{\theta}_i - \theta_i)^2/V_i \tag{4.16}$$

for every $(\theta_1, \ldots, \theta_k)$.

For illustration, these estimates are computed on the eight observations given in Table 3. The variances V_i have unit geometric mean, and nearly increase by a factor of two (actually 1.9921)) each time, leading to $\max(V_i)/\Sigma V_i = \frac{1}{8}$. The data X_i and true values θ_i are fictitious, but are carefully chosen functions of the square roots of the 16 expected squared N(0, 1) order statistics so that: (i) the $\{\theta_i\}$ look like a sample from N(0, 1)(hence A = 1); (ii) the values $(X_i-\theta_i)/V_i^{\frac{1}{2}}$ look like a N(0, 1) sample with $\Sigma(X_i-\theta_i)^2/V_i$ and $\Sigma(X_i-\theta_i)^2$ nearly equal to their expected values; (iii) $\Sigma(\theta_i-X_i/(1+V_i))^2$ and $\Sigma(\theta_i-X_i/(1+V_i))^2$ are nearly equal to their conditional expected values given X_i ($X_i/(1+V_i)$ is the Bayes estimator of θ_i if A=1 in (4.5)); and (iv) the three squared correlations between the pairs $(\theta_i, \log(V_i))$, $(\log(V_i), (X_i-\theta_i)/V_i^{\frac{1}{2}})$, and $(\theta_i, (X_i-\theta_i)/X_i^{\frac{1}{2}})$ have been controlled to be near their expected values, $1/(k-1) = 0.143$. The sample is called "surprise-free" for obvious reasons. Such a sample is desired because the purpose of this section is to illustrate the methods on only one data set, while we hope the results will typify more general experience.

The data and true values appear in column (1), (2), (3), (11) of Table 3. The amount of shrinking $\hat{B}_i$, column (4), increases sharply as V_i increases. The values $\hat{A}_i$, defined in comparison to (4.7) by

$$\hat{B}_i = \frac{V_i}{V_i + \hat{A}_i}, \tag{4.17}$$

also increase, nearly linearly in the standard deviation $(V_i)^{\frac{1}{2}}$.

Table 3

SURPRISE-FREE DATA WITH UNEQUAL VARIANCES, PARAMETER ESTIMATES, INTERVAL ESTIMATES, TRUE VALUES, AND ERRORS

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
V_i	$V_i^{\frac{1}{2}}$	X_i	$\hat{B}_i$	$\hat{A}_i$	$\hat{\theta}_i$	v_i	$\sigma_i^*(X)$	R_i^*	$\hat{R}_i$	θ_i	$\frac{\hat{\theta}_i-\theta_i}{\sigma_i^*(X)}$	$\frac{(\hat{\theta}_i-\theta_i)^2}{V_i}$
0.0896	0.2993	-2.0757	0.0243	3.60	-2.025	0.000	0.299	0.996	1.020	-1.7592	0.891	0.790
0.1786	0.4226	-1.0525	0.0466	3.65	-1.003	0.001	0.415	0.962	0.937	-0.4816	1.259	1.525
0.3557	0.5964	1.7510	0.0864	3.78	1.600	0.004	0.581	0.949	0.962	1.9836	0.661	0.414
0.7085	0.8417	0.6777	0.1527	3.92	0.574	0.010	0.778	0.854	0.723	-0.0267	-0.773	0.510
1.4113	1.1880	0.2406	0.2521	4.19	0.180	0.020	1.028	0.749	0.500	-0.5405	-0.701	0.368
2.8114	1.6767	-3.3058	0.3827	4.53	-2.041	0.030	1.435	0.733	1.035	-2.4542	-0.288	0.061
5.6002	2.3665	4.8068	0.5302	4.97	2.258	0.033	1.845	0.608	1.375	0.4345	-0.989	0.594
11.1554	3.3400	1.5676	0.6723	5.44	0.514	0.028	1.930	0.334	-0.233	0.5025	-0.006	0.000

The "true value" of A is 1.000, and $\Sigma\theta_i^2/k = 1.752$, so all these values of $\hat{A}_i$ are conservative, although for small V_i they are slightly less conservative because those components get higher relative weight when estimating A. The estimates $\hat{\theta}_i$ and their Bayesian standard errors, which increase with V_i, appear in columns (6) and (8). The $\hat{\theta}_i$ differ little from X_i for small V_i, but are shrunk considerably for large V_i. As usual in the unequal variances situation, the empirical Bayes estimates order the true means differently than the sample means do (the 4th and 8th components are reordered). This has important implications for the theory of ranking and selection.

The Bayesian estimate R_i^* of the ratio of the mean squared error of $\hat{\theta}_i$ relative to that of X_i is given in column (9). These values average 0.773, a quantity one cares about if the loss function is $\frac{1}{k}\Sigma(\hat{\theta}_i-\theta_i)^2/V_i$. The square root, 0.879, is the average ratio of confidence interval widths, although little improvement over the sample mean is possible for components with small V_i and much for large V_i. The unbiased estimates of risk, $\hat{R}_i$, appear in column (10), averaging 0.790, slightly higher than the R_i^* average. All the quantities in columns (1)-(10) can and should be computed when utilizing these estimates.

The "true values" appear in column (11). The relative errors of the estimate $\hat{\theta}_i$, given in column (12), have root mean square of 0.788, much less than the nominal value 1.000. Thus, confidence intervals based on $\sigma_i^*(X)$ would be conservative in this example. The weighted squared errors, whose expectations are estimated by R_i^* and $\hat{R}_i$ in columns (9), (10), appear in column (13). The sum of the values in column (13), corresponding to the loss function $\Sigma(\hat{\theta}_i-\theta_i)^2/V_i$ is 4.26, while if X_i is used, 8.00 (the expected loss) is obtained. For squared error loss, $\Sigma(\hat{\theta}_i-\theta_i)^2 = 4.87$ while $\Sigma(X_i-\theta_i)^2 = 22.56$ (the expectation of this last quantity is 22.32).

The values of the shrinking coefficients $\hat{B}_i$ (Table 3, column (4)) are plotted in Figure 3 against $\log(V_i)$, which is linear in i. The amount of shinking increases sharply as V_i increases, but not as much as the shrinking coefficient $B_i = V_i/(1+V_i)$ for the Bayes estimator $(1-B_i)X_i$ which would be used in (4.5) if A were known to be equal to 1.

The value $\hat{A}$ of A that maximizes $f_S(A)$ in (4.9) is $\hat{A} = 2.345$, being the maximum likelihood estimate of A based on the joint distribution (4.8) of $(S_1, S_2, \ldots, S_k)$. Use of this in the Bayes estimator (4.5) yields the empirical Bayes estimator, labeled "EBMLE" in Figure 3. As Figure 3 illustrates, this shrinking value $B_i = V_i/(2.345 + V_i)$ is less conservative than $\hat{B}_i$. For large values of k it should be nearly equal to $\hat{B}_i$.

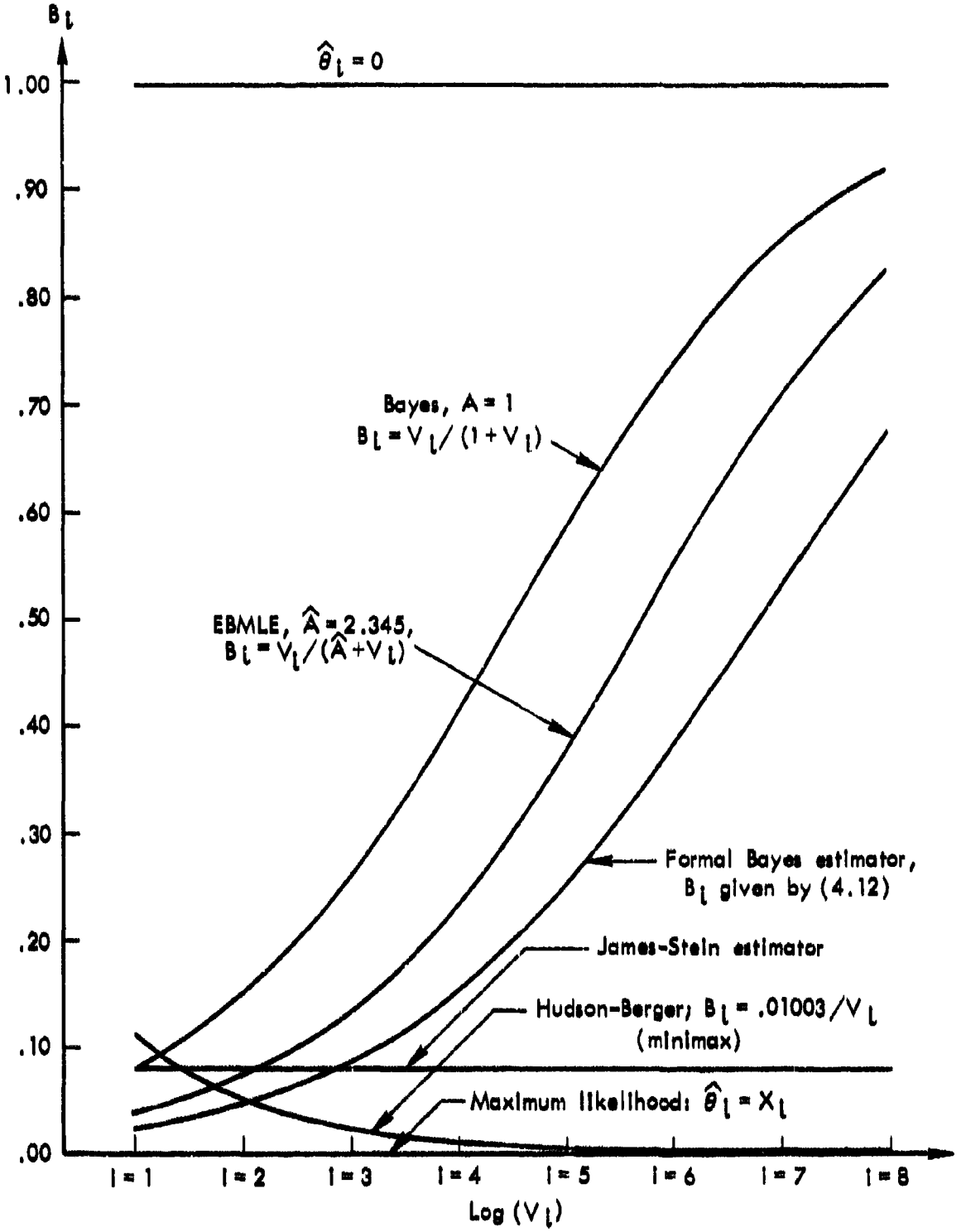


Fig. 3—Values of the shrinking coefficient $B_l = V_l/(\hat{A}_l + V_l)$ for several estimators of the form $(1-B_l)\,X_l$ plotted as a function of the logarithmic variance, for surprise-free data of Table 3

The two horizontal lines at $B_i = 0$ (no shrinkage) and $B_i = 1$ (full shrinkage to the prior mean) correspond respectively to the maximum likelihood estimator and to the estimator that ignores the data and estimates $\theta_i = 0$ in every case. The other estimators compromise between these extremes.

The James-Stein estimator, modified for the unequal variances situation, has constant shrinkage $B_i = 0.084$ for $i = 1, 2, \ldots, 8$. This estimator estimates θ_i by

$$(1 - \frac{k-2}{\Sigma X_j^2/V_j})X_i, \tag{4.18}$$

being minimax for the loss function $\Sigma(\hat{\theta}_i-\theta_i)^2/V_i$. It is derived by setting $\tilde{X}_i = X_i/V_i^{\frac{1}{2}}$, $\tilde{\theta}_i = \theta_i/V_i^{\frac{1}{2}}$, applying the James-Stein estimator (2.6) to $\{\tilde{X}_i\}$, and then transforming back. These transformations do not preserve the prior distribution (4.4), however, so the resulting estimator is unsatisfactory if the statistician thinks *a priori* that the $\{\theta_i\}$ are exchangeable. The result in Figure 3 slightly overshrinks the components $i = 1, 2$, which are well estimated by X_i and forfeits the big improvements possible for the components with large V_i.

The estimator of Hudson (1974) and Berger (1976), which estimates θ_i by

$$(1 - \frac{1}{V_i}\ \frac{k-2}{\Sigma X_j^2/V_j^2})X_i, \tag{4.19}$$

is minimax for the loss function $\Sigma(\hat{\theta}_i-\theta_i)^2$. But it shrinks less, not more, as the variances increase, and therefore can hardly shrink at all, see Figure 3.

This is the price one pays in order to use a minimax estimator in the case of unequal variances: almost no shrinkage will be allowed on those components that are not well estimated by X_i, although they are precisely the components where shrinkage is needed. Implicit in this statement is another assertion: estimators that are empirical Bayes against exchangeable prior cannot be minimax if the variance of some component is large relative to the others. A data analyst wishing to improve on the maximum likelihood estimator therefore must choose between two very different kinds of estimators. Since he probably is more able to recognize exchangeable prior distributions than to choose loss functions (and minimax estimators are highly sensitive to the weights L_i assumed in the loss function $\Sigma L_i(\hat{\theta}_i-\theta_i)^2$) he generally will be better off using empirical Bayes estimators. This approach also will permit him to identify many situations when he should stay with the maximum likelihood esimator.

The empirical Bayes approach, combined with formal Bayes theory, has one other advantage that is central to this paper. It provides a coherent method for computing interval estimates for the estimated parameters. For priors that yield estimators similar to the one of this paper, these intervals promise to contain the true means in most problems with the specified probability if the true means $\{\theta_i\}$ have any orthogonally invariant distribution, and perhaps will do so for most exchangeable prior distributions. If further research shows this, data analysts will be able to identify many situations for which powerful alternatives to the sample mean can be used.

REFERENCES

1. Abramowitz, M., and I. A. Stegun (eds.), *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc., New York, 1965.

2. Alam, Khursheed, "A Family of Admissible Minimax Estimators of the Mean of a Multivariate Normal Distribution," *Ann. Statist.*, Vol. 1, No. 3, May 1973, pp. 517-525.

3. Baranchik, A. J., "Multiple Regression and Estimation of the Mean of a Multivariate Normal Distribution," Technical Report No. 51, Stanford University, Department of Statistics, 1964.

4. Berger, J. O., "Admissible Minimax Estimation of a Multivariate Normal Mean with Arbitrary Quadratic Loss," *Ann. Statist.*, Vol. 4, No. 1, 1976, pp. 223-226.

5. Brown, L. D., "Admissible Estimators, Recurrent Diffusions, and Insoluble Boundary Value Problems," *Ann. Math. Statist.*, Vol. 42, 1971, pp. 855-903.

6. Efron, B., and C. Morris, "Limiting the Risk of Bayes and Empirical Bayes Estimators--Part II: The Empirical Bayes Case," *J. Amer. Statist. Assoc.*, Vol. 67, No. 337 (March 1972), pp. 130-139.

7. _____, and C. Morris, "Empirical Bayes on Vector Observations--An Extension of Stein's Method," *Biometrika*, Vol. 59, No. 2 (August 1972), pp. 335-347.

8. _____, and C. Morris, "Stein's Estimation Rule and Its Competitors--An Empirical Bayes Approach," *J. Amer. Statist. Assoc.*, Vol. 68, No. 341 (March 1973), pp. 117-130.

9. _____, and C. Morris, "Combining Possibly Related Estimation Problems," *J. Royal Statist. Soc.*, Series B, Vol. 35, No. 3 (November 1973; with discussion), pp. 379-421.

10. _____, and C. Morris, "Families of Minimax Estimators of the Mean of a Multivariate Normal Distribution," *Ann. Statist.*, Vol. 4, No. 1, 1976, pp. 11-21.

11. _____, and C. Morris, "Data Analysis Using Stein's Estimator and Its Generalizations," *J. Amer. Statist. Assoc.*, Vol. 70, No. 350, June 1975, pp. 311-319.

12. _____, and C. Morris, "Comment," *J. Amer. Statist. Assoc.*, Vol. 72, No. 357, March 1977, pp. 91-93.

13. _____, and C. Morris, "Stein's Paradox in Statistics," to appear in *Scientific American*, May 1977.

14. Hudson, H. M., "Empirical Bayes Estimation," Stanford University, Department of Statistics, Technical Report No. 58, May 31, 1974.

15. James, W., and C. Stein, "Estimation with Quadratic Loss," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, University of California Press, Berkeley, 1961, pp. 361-379.

16. Jeffreys, H., *The Theory of Probability*, 2nd Ed., Oxford University Press, 1948.

17. Leonard, T., "A Bayesian Method for the Simultaneous Estimation of Several Parameters," University of Warwick, Statistics Department, Coventry, England, April 1974.

18. Lindley, D. V., and A.F.M. Smith, "Bayes Estimates for the Linear Model (with discussion)," *J. Royal Statist. Soc.*, Series B, Vol. 34, 1972, pp. 1-42.

19. Stein, C., "Confidence Sets for the Mean of a Multivariate Normal Distribution," *J. Royal Statist. Soc.*, Series B, Vol. 24, No. 2 (1962), pp. 265-296.

20. _____, "Estimation of the Mean of a Multivariate Normal Distribution," *Proceedings of the Prague Symposium on Asymptotic Statistics*, 3-6 September 1973, pp. 345-381.

21. _____, "Estimation of the Parameters of a Multivariate Normal Distribution--I. Estimation of the Means," Stanford University, Department of Statistics, Technical Report No. 63, November 1974.

22. Strawderman, W. E., "Proper Bayes Minimax Estimators of the Multivariate Normal Mean," *Ann. Math. Statist.*, Vol. 42, No. 1, 1971, pp. 385-388.

23. Zellner, A., and W. Vandaele, "Bayes-Stein Estimators for k-Means, Regression and Simultaneous Equation Models," in S. E. Fienberg and A. Zellner (eds.), *Studies in Bayesian Econometrics and Statistics (in Honor of Leonard J. Savage)*, North-Holland Publishing Company, Amsterdam, 1975, pp. 627-653.

ROBUST STATISTICAL PROCEDURES

Robert V. Hogg
Department of Statistics
The University of Iowa
Iowa City, Iowa 52242

ABSTRACT. Two proposals are given that can be used to modify the method of least squares. The first replaces one of the factors in the squaring process by the rank of that factor. While some success has been achieved in applications with this procedure, the computations involved are not as easy as with the second method. In the latter, the square function is replaced by another function, say ρ. This ρ function can be convex, as in Huber's M-estimators, but it can also be non-convex, as in the descending M-estimators of Andrews and Hampel. The descending M-estimator scheme thus requires a better preliminary estimate so as not to find the "wrong" solution. Three examples using real data are considered.

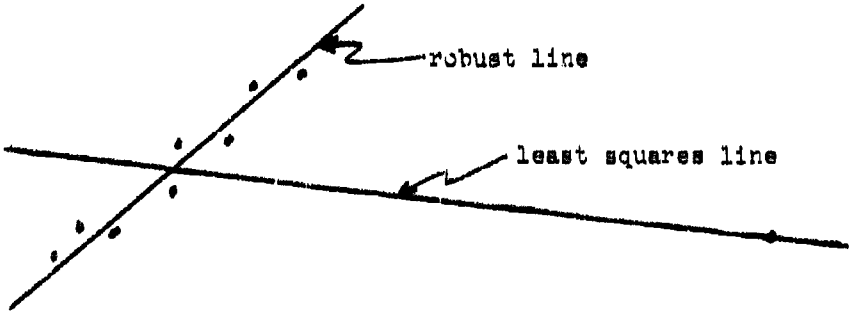
1. INTRODUCTION. The method of least squares, that is,

$$\text{minimizing} \sum_{i=1}^{n} (y_i - \sum_{j=1}^{p} \beta_j x_{ij})^2,$$

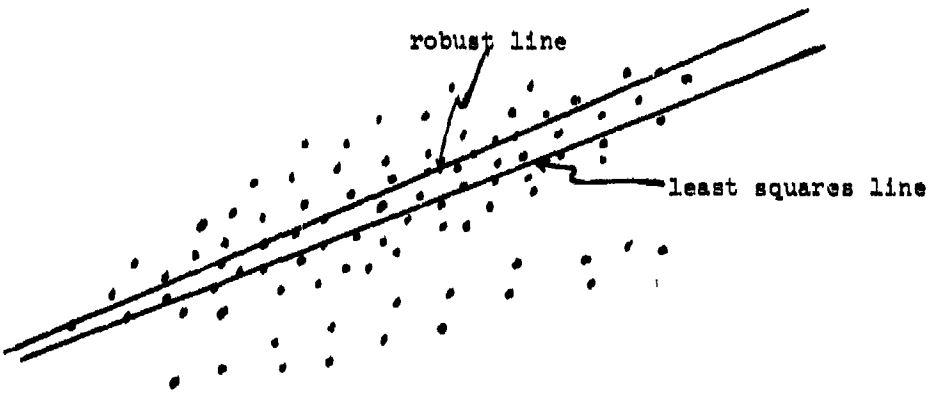
has served us well for many years! But there now is concern about the influence of "outliers" as they tend "to pull" the solution towards them too much. Consequently the residuals (if they are even considered) are distorted too much, and accordingly the outliers are difficult to detect. Of course, the situation is worse if the investigator blindly takes one of the many packaged programs and treats the answers as if they were the "truth" without checking assumptions, etc.

Two examples are:

(a)



(b)



While the first is one that I constructed, the second is <u>like</u> some lumber data that Boardman [3] considered. The investigator of that project at first fit the least square line, and <u>later</u> Boardman discovered that they were really dealing with <u>two</u> populations.

To see exactly what can be gained by robust methods, consider the example in Chapter 5 of the book by Daniel and Wood. This concerns the operation of a plant for the oxidation of Ammonia to Nitric Acid. There are 21 observations, in which the 3 independent variables are air flow, cooling water inlet temperature, and acid concentration while the stack loss is the dependent variable. The following table shows the "least squares" betas, the "least squares" betas with four <u>bad</u> points thrown out, and two sets of "robust" betas <u>based on all 21 observations</u>.

<u>ESTIMATES OF BETAS</u>

<u>METHOD</u>	<u>β_1</u>	<u>β_2</u>	<u>β_3</u>
Least squares	.72	1.30	-.15
Least squares (without outliers)	.80	.58	-.07
M-estimates (Andrews)	.82	.52	-.07
Nonparametric (median scores)	.83	.58	-.06

While some of the details of the latter two procedures will be explained later, please note that they give essentially the same answers using all 21 points as does least squares after 4 bad points have been removed; seemingly these robust schemes provide a BIG advantage in applications!

2. NONPARAMETRIC PROCEDURES. While I like "nonparametrics" myself, there are programming problems and hence we will not discuss that technique at length. The idea is this: Instead of minimizing

$$\sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} \beta_j x_{ij}\right)^2,$$

replace one of the factors $(y_i - \Sigma\beta_j x_{ij})$ by its rank, say R_i, and

$$\text{minimize} \quad \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} \beta_j x_{ij}\right) R_i.$$

Please note that R_i is a function of the $\beta_1, \beta_2, \cdots, \beta_p$ and hence an iterated process must be used (while there are a few short cuts, the ranking requires most of the computer time).

Of course, this nonparametric scheme can be generalized easily. Consider the "scores"

$$a(1) \leq a(2) \leq \cdots \leq a(n)$$

and then

$$\text{minimize} \quad \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} \beta_j x_{ij}\right) a(R_i).$$

Examples: (i) $a(i) = i$, then $a(R_i) = R_i$.

(ii) $a(i) = \begin{cases} -1, & i < (n+1)/2, \\ 1, & i > (n+1)/2, \end{cases}$ and if $n+1/2 = \text{integer}$,

then $a\left(\frac{n+1}{2}\right) = 0$.

The scoring in (ii) is often referred to as "median scores," and these scores were actually used in the nonparametric scheme associated with the Daniel and Wood example.

One final remark about these nonparametric procedures, if a constant is subtracted from the scores $a(i)$ so that the resulting a's are such that

$$\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a(i) = 0,$$

the minimization is equivalent to solving the p approximate equalities:

$$\sum_{i=1}^{n} x_{ij}a(R_i) \approx 0, \quad j = 1,2,\cdots,p.$$

While several persons have worked in this area, I believe that Hettmansperger and McKean [4] have developed the programs the most.

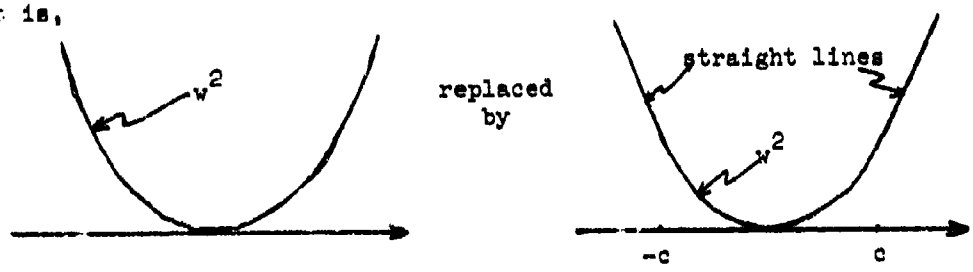
3. M-ESTIMATORS. Huber [6] first proposed these estimators. He suggested replacing, in least squares, the square function $\rho(w) = w^2$ by some other ρ function and

$$\text{minimizing} \quad \sum_{i=1}^{n} \rho\Big(y_i - \sum_{j=1}^{p} \beta_j x_{ij}\Big).$$

For some theoretical reasons, his first substitution was

$$\rho(w) = \begin{cases} w^2 & , \quad |w| \le c \\ 2c|w| - c^2, & |w| > c. \end{cases}$$

That is,



To clearly understand this substitution, let $y_1, y_2, \cdots, y_n$ be an observed random sample. Let us try to estimate the unknown middle θ by the method of least squares, noting the modification as we proceed.

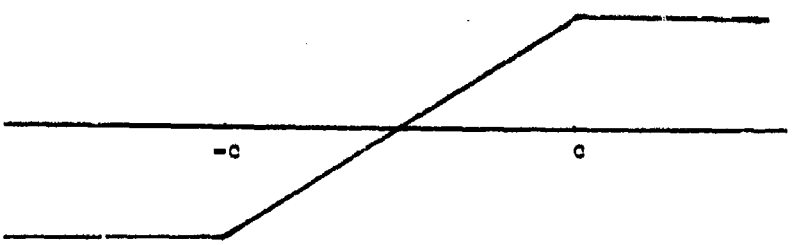
$$\rho(w) = w^2 \quad \text{and} \quad \min \sum_{i=1}^{n} (y_i - \theta)^2 = \min \sum_{i=1}^{n} \rho(y_i - \theta).$$

Take the derivative and equate to zero to obtain

$$\sum_{i=1}^{n} (-2)(y_i-\theta) = \sum_{i=1}^{n} -\psi(y_i-\theta) = 0,$$

where $\psi = \rho'$. In Huber's M-estimates (called this because if $\psi = -f'/f$, where f is the density, the resulting estimate is the Maximum likelihood of the location parameter), we have some difficulty because the formula changes at "c". Huber's $\psi = \rho'$ is

$$\psi(w) = \begin{cases} -2c, & w < -c, \\ 2w, & -c \leq w \leq c, \\ 2c, & c < w. \end{cases}$$



To make the equation,

$$\sum_{i=1}^{n} \psi(y_i-\theta) = 0,$$

have a <u>scale invariant</u> solution, we need to introduce a scale factor s in the following way:

$$\sum_{i=1}^{n} \psi\left(\frac{y_i-\theta}{s}\right) = 0.$$

A familiar s used by "robustniks" is given by

$$(.6745)s = \text{med}(|y_i - \text{med}(y_i)|) = \text{MAD},$$

the <u>m</u>edian of the <u>a</u>bsolute <u>d</u>eviations. The constant c should be selected so that if $y_1, y_2, \cdots, y_n$ actually arose from a normal population, most of

the numbers $|(y_i-\theta)/s| \le c$. Values of c around 1.5 or 2.0 are popular.

In the more general regression situation, we could take

$$s = \underset{\text{(of non-zero deviations)}}{\text{median}} \frac{|y_i - \text{prel. est. of middle}|}{(.6745)},$$

where preliminary estimate of middle should be fairly robust. While numerically difficult to determine, the β's that

$$\text{minimize} \sum \left| y_i - \sum_{j=1}^{p} \beta_j x_{ij} \right|$$

would provide robust estimates. However, while not real robust, many use least squares estimates, which is satisfactory with Huber's procedure.

The equations that we must solve in fitting $\sum_{j=1}^{p} \beta_j x_{ij}$ are, $j=1,2,\cdots,p$,

$$\sum_{i=1}^{n} \psi\left(\frac{y_i - \Sigma\beta_j x_{ij}}{s}\right) x_{ij} = 0.$$

Several iterations are usually required, and s would be recalculated on each (there are other suggestions for s in the literature [9] that are possibly easier to calculate).

Also note that if we wish to fit a non-linear function h of some parameters, say $\beta_1,\beta_2,\cdots,\beta_p$, we simply replace x_{ij} by $\frac{\partial h_i}{\partial \beta_j}$, where h_i is h with the i^{th} independent variables inserted (that is, those observations corresponding to y_i). Note in the special case h is linear, then

$$\frac{\partial h_i}{\partial \beta_j} = x_{ij}.$$

Hence, in the non-linear case, we solve (by iteration)

$$\sum_{i=1}^{n} \psi\left(\frac{\Delta_i}{s}\right) \frac{\partial h_i}{\partial \beta_j} = 0, \quad j = 1,2,\cdots,p,$$

where

$$\Delta_i = y_i - h_i.$$

One interesting way (there are others [9]) of handling this is by weighted non-linear least squares.

$$\sum_{i=1}^{n} \left[\frac{\psi(\Delta_i/s)}{(\Delta_i/s)}\right] \Delta_i \frac{\partial h_i}{\partial \beta_j} = 0,$$

where the weight

$$w_i = \psi(\Delta_i/s)/(\Delta_i/s) \quad \text{and} \quad \frac{\partial h_i}{\partial \beta_j}$$

are found from previous steps in the iteration (of course, recalculating s each time). Of course, ordinary non-linear least squares is

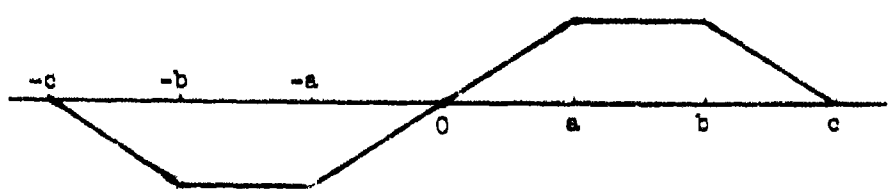
$$\min \sum_{i=1}^{n} \Delta_i^2 \quad \text{yields} \quad \sum_{i=1}^{n} \Delta_i \frac{\partial h_i}{\partial \beta_j} = 0.$$

Now we have

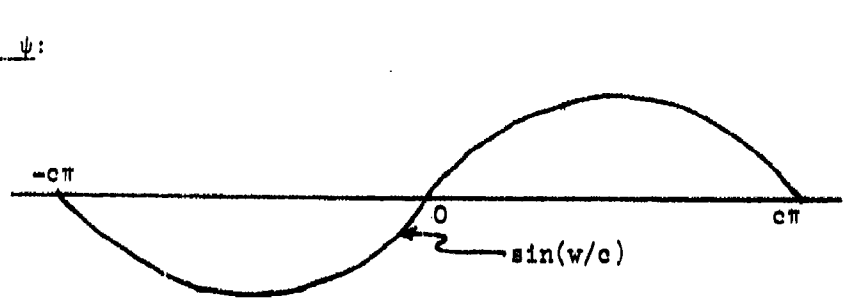
$$\sum_{i=1}^{n} w_i \Delta_i \frac{\partial h_i}{\partial \beta_j} = 0, \quad j = 1,2,\cdots,p.$$

4. DESCENDING M-ESTIMATORS. Several statisticians (Hampel, Andrews, etc. [1,2]) have modified Huber's ψ function (and, of course, the corresponding ρ) with functions that descend back to zero.

Hampel's ψ:



Andrew's ψ:



The problem, in the general regression situation, is still to solve

$$\sum_{i=1}^{n} \psi\left(\frac{\Delta_i}{s}\right) \frac{\partial h_i}{\partial \beta_j} = 0, \quad j = 1,2,\cdots,p.$$

Again, weighted (linear or non-linear as is the case) least squares is frequently used. However, since the corresponding ρ function is not convex, the solutions may not be unique. Thus, it is extremely important to start with a reasonably good preliminary estimate or else the iteration process could end up with the wrong solution. One way to avoid the wrong solution is through the use of Huber's ψ function on several iterations before using a descending ψ function.

It is also extremely interesting to study the weights associated with the various observations; they indicate the importance of the points. In particular, very low or zero weights (using Hampel's or Andrew's ψ) indicate that the corresponding points are probably outliers. To see how all of this fits together, let us consider two illustrations, both of which were obtained from the statisticians at the Los Alamos Sci. Lab. In each case, the Andrew's sine ψ function was used.

Ex. 1. Evaluating the lognormal assumption on bids for wildcat oil leases. There were 174 leases under consideration and in each case the number of bids ranged from 10 to 18. The logs of the bids were taken, and normality was tested using the Shapiro-Wilk W. In 64 cases out of the 174, normality was rejected. Hence it seemed that bids did not follow a lognormal assumption.

However, it was observed that there seemed to be some very low (noise) bids (oil firms trying to get a lease cheap). Hence, using Andrew's procedure, the middle of the values was estimated and the weights recorded with each observation. For illustration, here is a sample of $n = 14$ after 10 iterations (starting with $w_i = 1$).

$y_i = \log(i^{th})$	Δ_i	w_i (using c = 1.00)	
15.612	-.284	1.405	
15.080	-.816	1.133	
15.824	-.072	1.442	
15.872	-.024	1.444	
15.896	.000	1.445	weighted mean = 15.8962
14.916	-.980	1.009	
14.763	-1.133	.881	
16.148	.251	1.413	
16.246	.350	1.384	
16.727	.831	1.122	
17.289	1.392	.649	
13.529	-2.367	.000	
17.458	1.562	.495	outliers
10.463	-5.433	.000	

This was done for each of the 174 leases. The outliers (low, but noise bids) were eliminated from each. Then normality of the logs tested again. In this testing, only 5 of 174 cases were rejected. That is, about 3% were rejected, which is in good agreement with a 5% testing procedure. Thus it seems that bids do have an approximate lognormal distribution once the noise bids have been eliminated.

Ex. 2. Half-life of Plutonium-241. Six laboratories in the U.S. started a sample exchange program to follow the isotopic content of a Plutonium sample which had some of ^{238}Pu, ^{240}Pu, ^{241}Pu, ^{242}Pu, and ^{239}Pu, the latter of which was used as a base. That is, for example, values of the ratio of the contents of ^{241}Pu to ^{239}Pu were reported and denoted by R. Every 3 to 6 months, each of the six labs would report the value of this ratio giving a total of 78 points. They wished to fit the non-linear function $h(t) = R_0 e^{-kt}$. The data and print-out looked like this after 25 iterations.

item	w_i	mos	R_i	h_i	Δ_i
1.	7.62	0	.04471	.04470	.00001
2.	7.58	0	.04468	.04470	-.00002
	⋮		⋮		⋮
34.	4.20	16	.04168	.04191	-.00023
35.	0.00	16	.04271	.04191	.00080
	⋮ outlier		⋮		⋮

There were 6 points with zero weights (out of 78). The interesting thing is that upon checking these "bad" points it was discovered that all 6 were

from one lab, due to a technical difficulty. (Incidentally, the half-life seems to be about 14.4 ± .1 years. Without robust procedures, this was about 14.8 ± 1 year.)

While there are more improvements to be made using these robust procedures, they already provide substantial protection against outliers or bad data points and could be used in place of standard least squares procedures; for examples, regression, ANOVA, time series, and fitting by splines.

5. ACKNOWLEDGMENT. Robert V. Hogg's research on this topic was supported in part by NIH grant GM 22271-02.

REFERENCES

[1] D. F. Andrews, et al. (1972), Robust estimates of location: survey and advances. Princeton Univ. Press.

[2] D. F. Andrews (1974), A robust method for multiple linear regression. Technometrics 16, 523-531.

[3] T. J. Boardman, A review of the use of qualitative variables in regression analysis. Fall Technical Conference of ASQC and ASA, Denver, Colorado, October, 1975.

[4] T. P. Hettmansperger and J. W. McKean (1976), A robust alternative based on ranks to least squares in analyzing linear models, to appear in Technometrics.

[5] R. V. Hogg (1974), Adaptive robust procedures: a partial review and some suggestions for future applications and theory. J. Amer. Statis. Assoc. 69, 909-923.

[6] P. J. Huber (1964), Robust estimation of a location parameter. Ann. Math. Statist. 35, 73-101.

[7] P. J. Huber (1972), Robust statistics: a review. Ann. Math. Statist. 43, 1041-1067.

[8] P. J. Huber (1973), Robust regression: asymptotics, conjectures and Monte Carlo. Ann. Statist. 1, 799-821.

[9] P. J. Huber (1976), Robust methods of estimation of regression coefficients, Swiss Federal Inst. of Tech., Zürich.

ESTIMATING RELIABILITY FROM SMALL SAMPLES

Donald W. Rankin
Army Materiel Test and Evaluation Directorate
US Army White Sands Missile Range
White Sands Missile Range, New Mexico

ABSTRACT. Exact probability formulae are developed, with no restrictive assumptions, for use with tests which produce data of the go-no-go type. Although universally valid, the formulae are particularly apropos when small sample size is dictated. Since a programmable calculator greatly facilitates the solutions, programming suggestions are included.

1. INTRODUCTION. Often it is found that military, economic, or time limitations preclude the employment of any testing technique which requires that a large sample be taken.

Statistical treatment of small-sample data, always difficult enough, should not be degraded by requiring unnecessary postulates or by using formulae which yield only approximations. Consequently, the methods developed herein are based on no assumptions other than that of random sampling, and the formulae yield exact answers.

Increasing availability of programmable calculators with external program storage makes this approach completely feasible. With this in mind, programming suggestions are included where they seem to be indicated.

Since the formulae are exact, there is no theoretical limit to sample size. There is, however, a practical one, depending jointly upon the size and operating speed of the computer or calculator and upon the ingenuity of the programmer.

It may prove helpful to insert here a few remarks on notation and terminology, since there are to be found variations in the literature.

Factorials are variously indicated as

$$\lfloor n = n! = n^{(n)} = 1 \cdot 2 \cdot 3 \cdots n \quad .$$

The symbol $\lfloor n$ is chosen for use, since it acts as parentheses and thus reduces confusion when parentheses are used for another purpose within the same expression (e.g. Equation 38).

Generalized factorials are found as $n^{(m)}$ or $(n)_m$.

$$n^{(m)} = n(n-1)(n-2) \cdots (n-m+1) = \frac{\lfloor n}{\lfloor n-m}$$

will be used.

Binomial coefficients appear in many ways:

$$ {}_nC_k = C(n,k) = \binom{n}{k} = \frac{n^{(k)}}{\lfloor k} = \frac{\lfloor n}{\lfloor k \, \lfloor n-k} \quad . $$

The symbol C(n,k) is adopted, since it can be typed easily on a single line.

The indefinite summation symbol is taken to mean

$$ \Sigma\phi(x) = \phi(a) + \phi(a+1) + \cdots + \phi(x-1) \quad , $$

a series which consists of exactly x-a terms. The indefinite finite integral thus is

$$ \Delta^{-1}\phi(x) = \Sigma\phi(x) + C \quad . $$

Here, $\Delta^{-1}\phi(x)$ is analogous to $\int f(x)\,dx$ in the infinitesimal calculus.

The generalized notation used for a series is

$$ S = T_1 + T_2 + T_3 + \cdots + T_i + \cdots \quad . $$

If there exists some value of i such that $T_j = 0$ for all $j > i$, the series is finite.

Derivatives are shown by primes:

$$ \frac{d}{dr} f(r) = f'(r) \quad . $$

Level of confidence is denoted by L.

By "insignificant" is meant "insignificant to the computer." For example, if a series S is being summed and S_j represents the sum of the first j terms, T_{j+1} is insignificant if it is too small to affect the least significant digit of S_j as programmed in the calculator.

2. BINOMIAL PROBABILITY. Sometimes the testing technique permits sampling with replacement. Even when replacement is not possible, the same condition can be achieved (mathematically) by assuming* a population of infinite size. In other words:

The act of sampling does not alter the characteristics of the population.

*This does not belie the statement of Paragraph 1, since the opposite case -- when an infinite population cannot be assumed -- also is covered in Paragraph 4.

Given the above condition, let us specify that in a certain population, the probability of observing a success is given by r. Obviously, the probability of observing a failure is given by 1-r, which we shall call p.

$$p = 1 - r \quad .$$

It follows that $(p+r)^n = 1$. Thus, if we draw a sample of size n, the probability of observing exactly k failures is given by the appropriate term* of the binomial expansion

$$(p+r)^n = C(n,0)\, p^n + C(n,1)\, p^{n-1}\, r + C(n,2)\, p^{n-2}\, r^2 + \cdots + C(n,n)\, r^n \quad .$$

Since,

$$\sum_{k=0}^{n} C(n,k)\, p^{n-k}\, r^k = 1 \quad , \qquad (k = 0, 1, 2, \cdots, n) \quad , \qquad (1)$$

it may be said that

$$C(n,k)\, p^{n-k}\, r^k$$

defines a probability function in the discrete variable k.

Noting that

$$C(n,k) = \frac{\lfloor n}{\lfloor k \, \lfloor n-k} = C(n,\, n-k) \quad ,$$

we define

$$p(k) = C(n,k)\, p^k\, (1-p)^{n-k} = C(n,k)\, r^{n-k}\, (1-r)^k \qquad (2)$$

as the probability of observing exactly k defectives (failures) in n trials.

Unfortunately, the problem rarely is that simple. In most test designs, it is possible to control the value of n arbitrarily, and to observe the value of k exactly, but nothing is known about r. A probability function in r is required.

*The term containing p^k

Now r can take any value within the prescribed limits, $0 \leq r \leq 1$; i.e., it is a <u>continuous</u> variable and necessarily

$$\int_{r=0}^{1} f(r)\, dr = 1 \qquad (3)$$

describes f(r), whatever it may turn out to be, as the required probability function in r. Setting

$$g(r) = C(n,k)\, r^{n-k}\, (1-r)^{k} \quad , \qquad (4)$$

n and k being constant, we see from Equation (2) that g(r) is a density function in r. In order to discover a relationship between g(r) and f(r), we must evaluate

$$\int_{r=0}^{1} g(r)\, dr = C(n,k) \int_{0}^{1} r^{n-k}\, (1-r)^{k}\, dr \quad . \qquad (5)$$

To integrate*

$$\int x^{n-k}\, (1-x)^{k}\, dx$$

let

$$u = (1-x)^{k}$$

and

$$dv = x^{n-k}\, dx \quad .$$

Then,

$$du = -k(1-x)^{k-1}\, dx$$

and

$$v = (n-k+1)^{-1}\, x^{n-k+1} \quad .$$

---

*The value of $\int_{0}^{1} x^{n-k}\, (1-x)^{k}\, dx$ is found in many tables. But a program for computing $\int_{0}^{z} x^{n-k}(1-x)^{k}\, dx$, $0 < z < 1$ is required, hence it is considered desirable to show the complete process of integration. These two definite integrals are sometimes referred to as the complete and incomplete Beta-function.

Note that n and k are integers such that $0 < n \geq k \geq 0$.

$$\int x^{n-k}\,(1-x)^{k}\,dx = \frac{1}{n-k+1}\,x^{n-k+1}\,(1-x)^{k} + \frac{k}{n-k+1}\int x^{n-k+1}\,(1-x)^{k-1}\,dx \quad .$$

Another similar integration by parts is performed upon the last term, yielding

$$\begin{aligned}\int x^{n-k}\,(1-x)^{k}\,dx = {} & \frac{1}{n-k+1}\,x^{n-k+1}\,(1-x)^{k} \\ & + \frac{k}{n-k+1}\left\{\frac{1}{n-k+2}\,x^{n-k+2}\,(1-x)^{k-1}\right. \\ & \left. + \frac{k-1}{n-k+2}\int x^{n-k+2}\,(1-x)^{k-2}\,dx\right\} \quad .\end{aligned}$$

Iterating k times results in

$$\begin{aligned}\int x^{n-k}\,(1-x)^{k}\,dx = {} & \frac{1}{n-k+1}\,x^{n-k+1}\,(1-x)^{k} + \frac{k}{(n-k+1)(n-k+2)}\,x^{n-k+2}\,(1-x)^{k-1} \\ & + \cdots + \frac{k(k-1)\cdots(k-k+2)}{(n-k+1)(n-k+2)\cdots(n-k+k)}\,x^{n-k+k}\,(1-x)^{k-k+1} \\ & + \frac{k(k-1)\cdots(k-k+1)}{(n-k+1)(n-k+2)\cdots(n-k+k)}\int x^{n-k+k}\,(1-x)^{k-k}\,dx \quad . \qquad (6)\end{aligned}$$

But now the last term submits to integration. It can be rewritten

$$\frac{\lfloor k}{n^{(n-k)}}\int x^{n}\,dx = \frac{\lfloor k\ \lfloor n-k}{\lfloor n+1}\,x^{n+1} \quad . \qquad (7)$$

It now becomes very easy to evaluate the definite integral

$$\int_{x=0}^{1} x^{n-k}\,(1-x)^{k}\,dx \quad ,$$

since at the lower limit, all terms become zero, and at the upper limit (x=1), all terms <u>except the last</u> become zero. Hence

$$\int_{x=0}^{1} x^{n-k}\,(1-x)^{k}\,dx = \frac{\lfloor k\ \lfloor n-k}{\lfloor n+1} = \frac{1}{(n+1)C(n,k)} \quad . \qquad (8)$$

Substituting this expression into Equation (5), we arrive at the remarkable result

$$\int_{r=0}^{1} g(r)\,dr = \int_{0}^{1} C(n,k) r^{n-k} (1-r)^{k}\,dr = \frac{1}{n+1} \quad , \tag{9}$$

i.e., $\int_{0}^{1} g(r)\,dr$ depends upon sample size only! And thus, the desired probability function in r is

$$f(r) = (n+1)\,g(r) = \frac{\lfloor n+1}{\lfloor k \lfloor n-k} r^{n-k} (1-r)^{k} \quad . \tag{10}$$

That it be a useful probability function requires that other definite integrals can be computed. Substituting Equation (7) into Equation (6) and multiplying through by (n+1)C(n,k) enables us to write

$$\int \frac{\lfloor n+1}{\lfloor k \lfloor n-k} x^{n-k} (1-x)^{k}\,dx = C + \sum_{i=0}^{k} C(n+1,k-i) x^{n+1-k+i} (1-x)^{k-i} \quad . \tag{11}$$

Without loss of generality, we can choose the lower limit (of the definite integral) to be zero. The function there conveniently reduces to the constant of integration. Also, to avoid programming problems, we can restrict the upper limit to values less than unity. Thus, for an arbitrary value of z,

$$\int_{r=0}^{z} f(r)\,dr = z^{n-k+1} \sum_{i=0}^{k} C(n+1,k-i) z^{i} (1-z)^{k-i} \quad , (0 \leq z < 1) \tag{12}$$

expresses the probability that $r \leq z$. The case of z = 1 already has been covered by Equations (3) and (10), i.e.,

$$\int_{r=0}^{1} f(r)\,dr = \int_{0}^{1} (n+1) C(n,k) r^{n-k} (1-r)^{k}\,dr = 1 \quad .$$

The same formula (Equation 12) can be used to solve the inverse problem;* i.e., when the level of confidence is specified. Set

$$L = 1 - \int_{0}^{z} f(r)\,dr \quad , \tag{13}$$

then solve for z.

---

*See Paragraph 5C.

3. ESTIMATES OF THE RELIABILITY.

A. The Function f(r).

It is worthwhile to examine the probability function

$$f(r) = (n+1)C(n,k)r^{n-k}(1-r)^{k} \quad . \tag{10}$$

A typical graph (n = 7, k = 2) is shown in Figure 1. The area under the curve is divided into quarters by the ordinates at r = 0.567, 0.679, and 0.779. A maximum occurs when, exclusive of the end points, f'(r) = 0; i.e., when (n-k)(1-r) - kr = 0. We shall call this maximum the "maximum likelihood estimate" of the reliability and identify it with a circumflex (^). It computes easily to be

$$\hat{r} = 1 - \frac{k}{n} \quad . \tag{14}$$

When n-1 > k > 1, the curve exhibits two inflection points, equally spaced about the maximum. They occur at

$$r_{pi} = \hat{r} \pm \sqrt{\hat{r}^2 - \hat{r}\left(1 - \frac{k}{n-1}\right)} \quad . \tag{15}$$

As will be seen later, they are of interest to the programmer. Figure 1 shows inflection points at (0.530, 1.55) and (0.899, 1.01).

When k = 1, only one inflection point appears at

$$r_{p_1} = \hat{r} - \frac{k}{n} \quad ,$$

(see Figure 2).* Any program must take this fact into account.

B. Level of Confidence.

It is the nature of a function of a continuous variable that an area below the curve (i.e., a definite integral) cannot be described by a single point. A pair of points is required.

When the function under consideration is a probability function,** the ordinates erected at the selected pair of points enclose an area called the level of confidence. It is proper to think of a level of confidence as an area,

*Quartiles, r = 0.697, 0.799, and 0.870. r_{p_1} at (0.714, 2.125).

**i.e., when $\int_a^b f(r)\,dr = 1$.

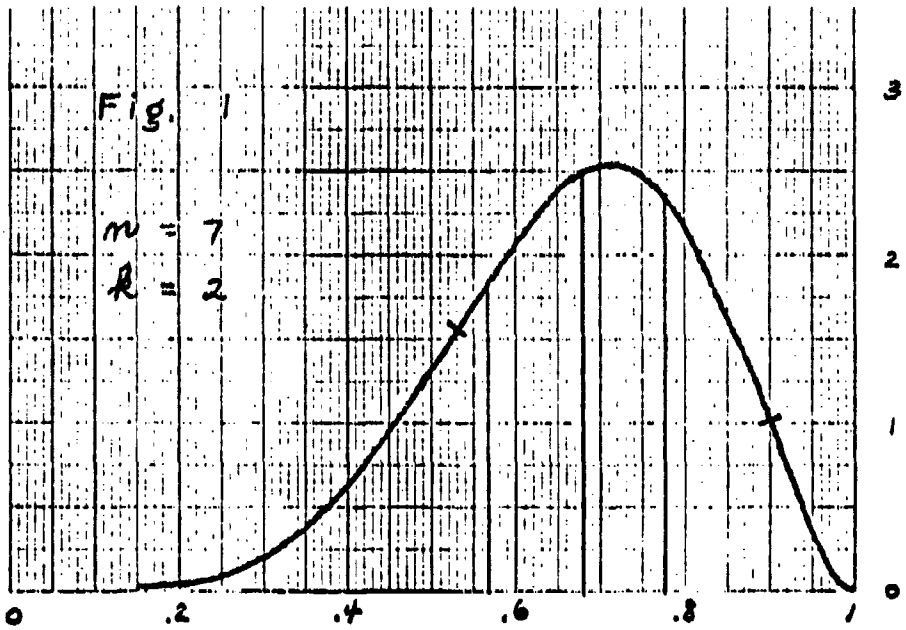


Fig. 1
n = 7
k = 2
0
.2
.4
.6
.8
1
0
1
2
3

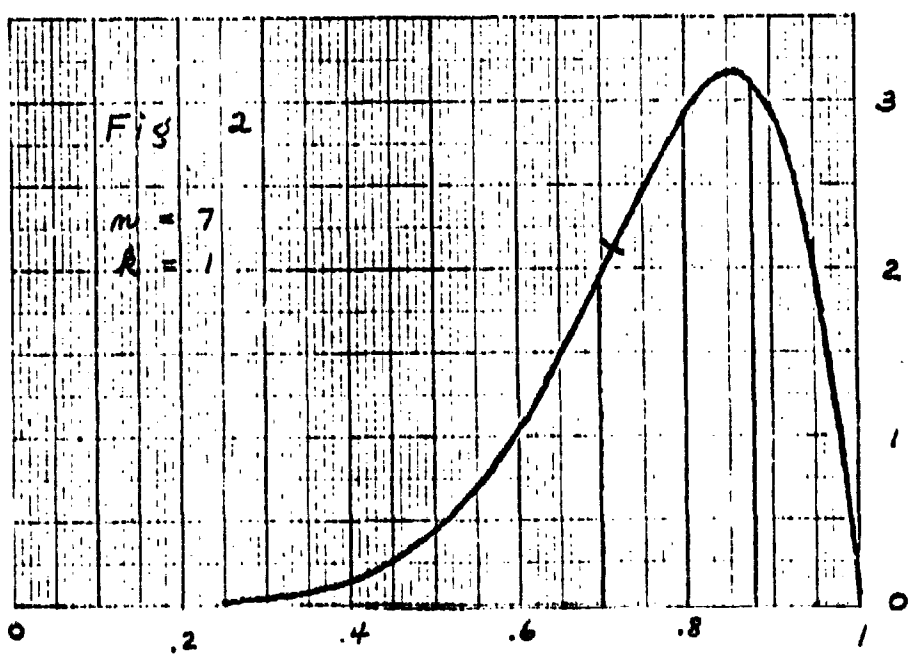


Fig 2
n = 7
k = 1
0
.2
.4
.6
.8
1
0
1
2
3

as a definite integral, or as a probability. Again referring to Figure 1, it can be stated "at the 50% level of confidence, $0.567 < r < 0.779$" or "at the 75% confidence level, $r > 0.567$." In the latter case, $r = 1$ is the second member of the pair.

Selection of a level of confidence may be, and often should be, quite arbitrary. However, deferment of this selection until after preliminary test results are in, in an effort to "improve" the data, usually can be regarded as a reprehensible practice.

When selecting a confidence level (in advance, of course) it sometimes helps in vizualizing it, to couch it in terms of ordinary gambler's odds, rather than the more commonly used decimal fraction. Thus, a confidence level of 0.96 gives odds of 24 to 1 against the analyst issuing erroneous advice. At 0.90, the odds drop to 9 to 1 and at 0.75 to an alarming 3 to 1.

However, there is another side to this coin. Consider what happens when a 100% level of confidence is chosen. Obviously, the pair of defining points is located at 0 and 1, regardless of the nature and shape of the probability function. Selecting too high a confidence level produces a strong masking effect by driving the defining points (limits of integration) far into the tails. A higher-than-necessary level of confidence may be a luxury the analyst can ill afford.

In summary, there are two approaches for handling the data. The first is to select (perhaps arbitrarily) two values of the argument, then compute the level of confidence (area) between them. The second* is to choose a confidence level, then compute two values of r which will bound it.

C. The Case of Zero Failures.

Specifically, when $k = 0$, the function degenerates to

$$f(r) = (n+1)r^n \quad . \tag{16}$$

Additionally, given $n > 1$ and $r > 0$,

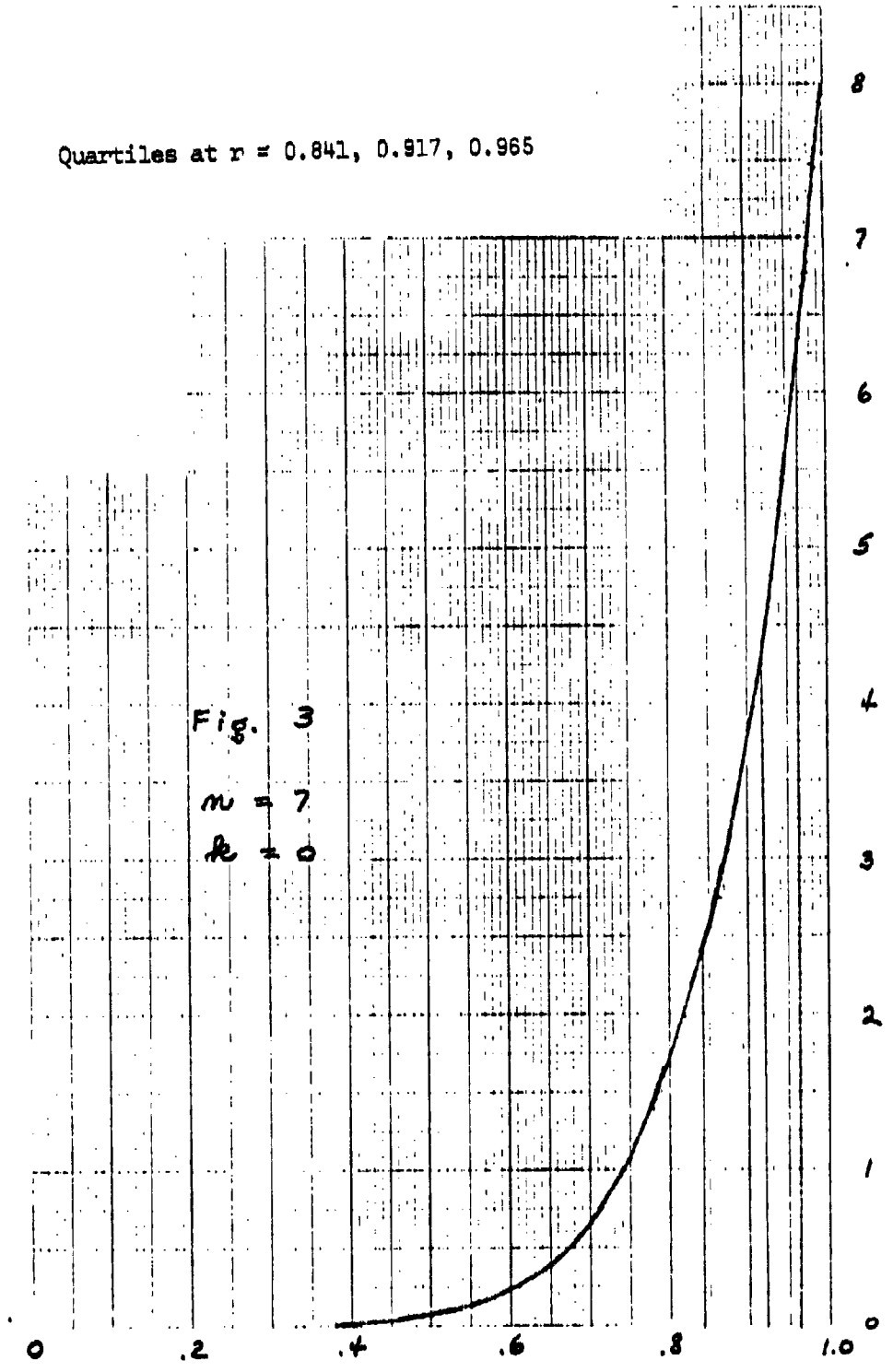
$$f'(r) \neq 0$$

$$f''(r) \neq 0$$

and the algorithms which will be developed will fail. The function for $n=7$, $k=0$ is shown in Figure 3. Note that there is no point of inflection and no maximum in the usual sense. However, we still can define

$$\hat{r} = 1 - \frac{k}{n} = 1 - \frac{0}{n} = 1 \quad .$$

*See Paragraph 3D and the opening remarks of 3E.


Quartiles at r = 0.841, 0.917, 0.965
Fig. 3
n = 7
k = 0
0
.2
.4
.6
.8
1.0
0
1
2
3
4
5
6
7
8

The solution of this case is very simple and can be effected with an ordinary table of logarithms, since

$$\int_0^z f(r)\,dr = z^{n+1} \quad . \tag{17}$$

The practical solution possibilities are limited to two.

(1) Choose r = z and r = 1 as the two values of the argument, z being arbitrary but less than 1. Then the level of confidence, L, is given by

$$L = 1 - z^{n+1} \tag{18}$$

(2) Choose L. Then set r = 1 as the upper bound. The lower bound, r = z, is given by

$$z = \left(1 - L\right)^{\frac{1}{n+1}} \quad . \tag{19}$$

If the programmer wishes to include the case of zero failures, he should write it as a separate sub-routine.

D. The Best Estimate of the Reliability ($0 < k < n$).

When both values of r are specified ($r = z_1$ and $r = z_2$), the problem is straightforward enough. Simply use Equation (12) twice to compute L.

$$L = \int_0^{z_2} f(r)\,dr - \int_0^{z_1} f(r)\,dr \quad . \tag{20}$$

If either $z_1 = 0$ or $z_2 = 1$, then Equation (12) need be employed only once.

But when L is specified, there are an infinite number of solution-pairs which satisfy the required condition. The usual way out of this dilemma is to set one of the limits to be 0 or 1, then solve Equation (12) for the other. Newton's method of successive approximations is well-suited to effect this solution. An algorithm will be given which converges quite rapidly upon the correct answer.

Sometimes a confidence level is specified which arbitrarily excludes equal areas from each end of the distribution. This is equivalent to two solutions with $z_1 = 0$.

But a programmable calculator makes practicable a more elegant solution. Let it be called "The Best Estimate of the Reliability." Briefly described, it is this: The level of confidence being specified, the best estimate of the reliability is given by the particular values of z_1 and z_2 which minimize the difference $z_2 - z_1$. We shall designate them with a tilde thus:

$$\tilde{z}_1, \tilde{z}_2 \quad \text{or} \quad \tilde{r}_1, \tilde{r}_2 \quad .$$

The best estimate of the reliability possesses several distinguishing properties:

(1) The solution is unique.

(2) $z_2 - z_1$ is a minimum, by definition.

(3) $f(\tilde{z}_1) = f(\tilde{z}_2)$. (21)

That this is true is evident from Figure 4. If either ordinate is displaced away from the maximum, the other must be displaced a smaller amount to conserve area; i.e., $z_2 - z_1$ increases. This important equality is made use of in the solution.

(4) $\tilde{z}_1$ and $\tilde{z}_2$ always lie on opposite sides of $\hat{r}$. Thus is avoided the absurdity of excluding $\hat{r}$ from the solution area. This property also is used in the solution.

(5) Any included value of r is more likely than every excluded value.

Note that when k = 0, the solution is degenerate.* This should not be surprising, since r = 1 yields an absolute extremal, not a relative one.

E. Comparison of Methods.

It is common practice to specify L, then set $z_2 = 1$ and compute z_1. Under these conditions, z_1 is a function of L. Although this does not invalidate the method, it indicates that due caution be exercised, lest the published value of z_1 reflect little more than the analyst's whim. The method can make only one kind of statement, viz. "At the 75% confidence level, r exceeds 0.567." No attempt is made to predict what r actually is (it may be far from 0.567) and nothing is said about the shape of the distribution, save that the right-hand "tail" surely is included. The method might be used by a manufacturer or user to test for compliance with a minimum standard.

*See Paragraph 3C.

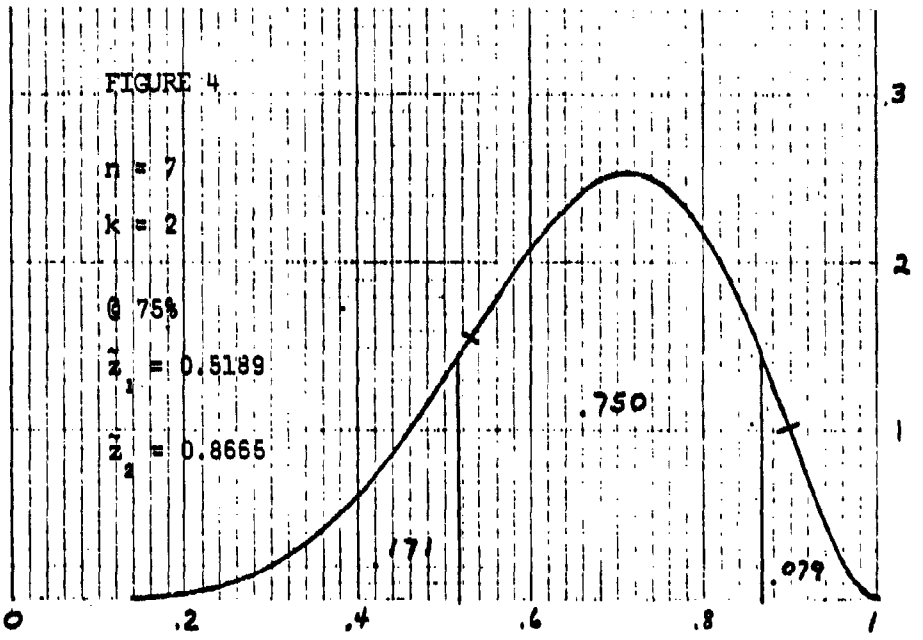


FIGURE 4
n = 7
k = 2
@ 75%
$\hat{z}_1$ = 0.5189
$\hat{z}_2$ = 0.8665
.750
.171
.079
0
.2
.4
.6
.8
1
2
3

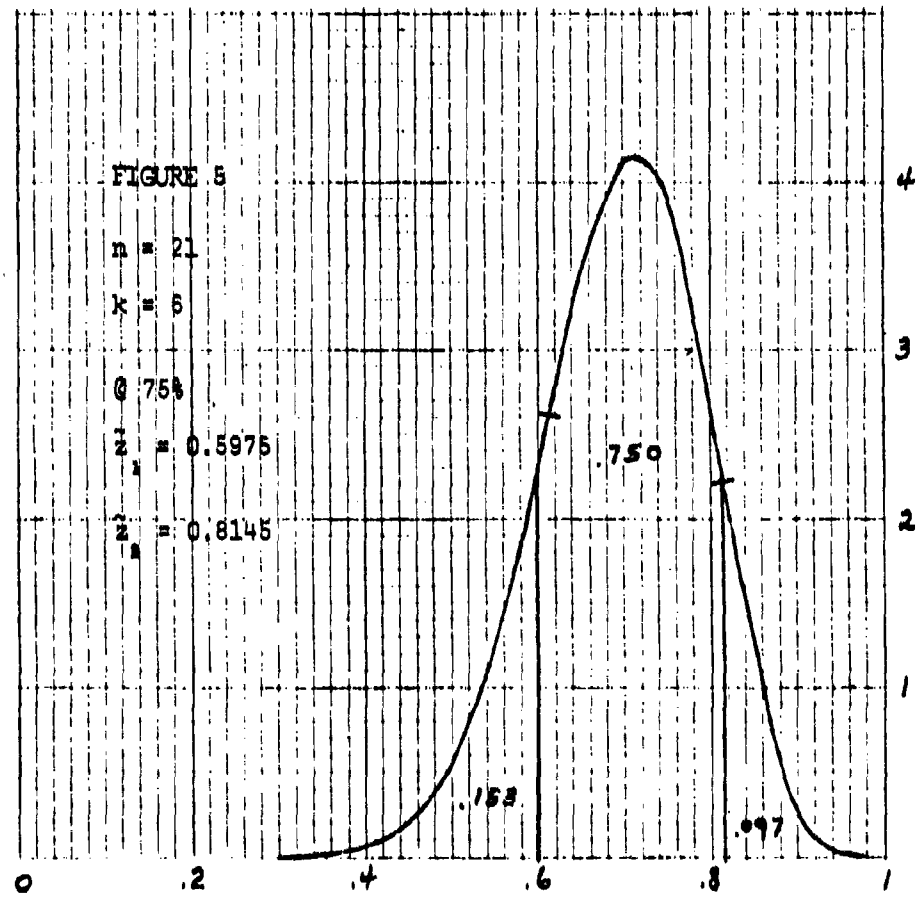


FIGURE 5
n = 21
k = 6
@ 75%
$\hat{z}_1$ = 0.5975
$\hat{z}_2$ = 0.8145
.750
.153
.097
0
.2
.4
.6
.8
1
2
3
4

On the other hand, the best estimate of the reliability states "The maximum likelihood estimate of r is 0.714 and, in any event, at the 75% level of confidence r lies between 0.519 and 0.867."

The values of $z_2 - z_1$ are 0.433 and 0.348, respectively.

The "best estimate" might be used to evaluate a new device or procedure, without reference to a pre-established criterion.

In a nut-shell, one method measures, the other tests for compliance. Before choosing between them, the analyst must decide what sort of question he is attempting to answer.

F. The Effect of Increasing Sample Size.

What happens when the same failure rate is observed in a larger sample? This is graphically illustrated in Figures 4 and 5. It is observed that $\hat{r}$ is unchanged, but $f(\hat{r})$ increases. Also, $\tilde{z}_1$ and $\tilde{z}_2$ both move inward toward $\hat{r}$; i.e., $\tilde{z}_2 - \tilde{z}_1$ decreases. It is clear that enlarging the sample size will increase the precision of the "best estimate." If n becomes great enough, the graph of the function virtually is reduced to a tall spike at $\hat{r}$.

4. HYPERGEOMETRIC PROBABILITY. When test conditions do not permit sampling with replacement, and when the population is known to be finite (and measureable!) in size,* the theory of Paragraph 2 is not applicable. We must perforce develop another method for dealing with sampling without replacement. To parallel our earlier statement, we say:

The act of sampling measurably alters some characteristic of the remaining population.

In this Paragraph, we shall not speak of the reliablity, nor shall we employ as a symbol the letter r. (As will be seen, the analogous quantity is $1 - x/N$.)

Given a population consisting of N items, x of which are defective, the probability that a sample of size n will contain exactly k defectives is

$$p(k) = p(N,x,n,k) = \frac{\lfloor N-n \rfloor n \lfloor N-x \rfloor x}{\lfloor N-n-x+k \rfloor n-k \lfloor x-k \rfloor N \lfloor k} \quad . \tag{22}$$

Notice that x and n are interchangeable in the formula, which, at our convenience, can be written in either of two ways:

$$p(k) = \frac{C(N-n,x-k)\cdot C(n,k)}{C(N,x)} = \frac{C(N-x,n-k)\cdot C(x,k)}{C(N,n)} \quad . \tag{22}$$

*e.g., test-firing guided missiles.

But in the usual case, N, n, and k are known and it is required to estimate x; i.e., x IS THE ONLY VARIABLE. What is needed is a probability function in x.* Now Equation (22) in any of its forms gives p(k) as a probability function in k, but not necessarily in x. It is observed that with n and k being held constant, p(k) serves as a density function in the discrete variable x. In attempting to disclose the relationship between p(k) and the desired probability function in x -- which we shall write as $p_x = p_x(N,x,n,k)$ -- we must, as the first step, evaluate the finite definite integral**

$$Q_x = Q_x(N,n,k) = \sum_{x=k}^{k+N-n} p(k) \quad , \tag{23}$$

a series consisting of N-n+1 terms. The limits of integration are obvious, since k defectives already have been observed, and N-n is the population remaining. Substituting Equation (22) in its first form for p(k) and factoring out the constants (which do not contain x) we find

$$Q_x = \frac{\lfloor N-n \rfloor n}{\lfloor n-k \rfloor N \rfloor k} \sum_{x=k}^{k+N-n} (N-x)^{(n-k)} \, x^{(k)} \tag{24}$$

where $(N-x)^{(n-k)}$ and $x^{(k)}$ denote generalized factorials.*** An expression for this integral is obtained as follows:

Let

$$u_x = (N-x)^{(n-k)}$$

and

$$\phi(x) = x^{(k)} \quad .$$

Then,

$$C + \Sigma u_x \phi(x) = \Delta^{-1} u_x \phi(x) = (EE' - 1)^{-1} \, u_x \phi(x) \tag{25}$$

---

*Cf the discussion following Equation (3), Paragraph 2.

**i.e., sum the finite series over all possible values of x.

***The basic reference for the following derivation is George Boole's "Calculus of Finite Differences." Boole's notation (third and later editions) is used throughout.

where, temporarily, E operates on u alone, E′ on ϕ alone. Continuing,

$$(EE'-1)^{-1}u_x\phi(x) = [(1+\Delta)E'-1]^{-1}u_x\phi(x)$$

$$= (\Delta' + \Delta E')^{-1}u_x\phi(x)$$

$$= \frac{1}{\Delta'}\left(1+\frac{\Delta E'}{\Delta'}\right)^{-1}u_x\phi(x)$$

$$= \frac{1}{\Delta'}\left\{1-\frac{\Delta E'}{\Delta'}+\left(\frac{\Delta E'}{\Delta'}\right)^2-\cdots\right\}u_x\phi(x) \quad . \qquad (26)$$

From Equation (26) we can write the desired expansion, dropping the primes as no longer necessary.

$$\Sigma u_x\phi(x) = -C + u_x\Sigma\phi(x) - \Delta u_x\Sigma^2\phi(x+1) + \Delta^2 u_x\Sigma^3\phi(x+2) - \cdots$$

$$\cdots + (-1)^j\Delta^j u_x\,\Sigma^{j+1}\phi(x+j) + \cdots \quad . \qquad (27)$$

The series of Equation (27) will terminate after n-k+1 terms, fewer by N+k-2n than that of Equation (23). It can be used to sum any number of terms of Equation (23) or Equation (24).

It will prove useful to list a breakdown of the terms in Equation (27). This is done below.

$$u_x = (N-x)^{(n-k)}$$

$$-\Delta u_x = (n-k)(N-x-1)^{(n-k-1)}$$

$$\Delta^2 u_x = (n-k)^{(2)}(N-x-2)^{(n-k-2)}$$

$$\vdots$$

$$(-1)^j\Delta^j u_x = (n-k)^{(j)}(N-x-j)^{(n-k-j)}$$

$$\vdots$$

$$(-1)^{n-k}\Delta^{n-k}u_x = \lfloor n-k$$

$$\Delta^{n-k+1}u_x = 0 \qquad (28)$$

$$\Sigma\phi(x) = \frac{x^{(k+1)}}{k+1}$$

$$\Sigma^2\phi(x+1) = \frac{(x+1)^{(k+2)}}{(k+2)^{(2)}}$$

$$\Sigma^3\phi(x+2) = \frac{(x+2)^{(k+3)}}{(k+3)^{(3)}}$$

$$\vdots$$

$$\Sigma^{j+1}\phi(x+j) = \frac{(x+j)^{(k+j+1)}}{(k+j+1)^{(j+1)}}$$

$$\vdots$$

$$\Sigma^{n-k+1}\phi(x+n-k) = \frac{(x+n-k)^{(n+1)}}{(n+1)^{(n-k+1)}} \qquad (29)$$

Noting that when $x < k$, $\phi(x) = x^{(k)} = 0$, we have

$$\sum_{x=0}^{k-1} u_x\phi(x) = 0$$

and

$$\sum_{x=0}^{k} u_x\phi(x) = (N-k)^{(n-k)} \lfloor k \quad .$$

From the definition of the operator Σ,* we have

$$\Sigma u_k\phi(k) = \sum_{x=0}^{k-1} u_x\phi(x) = 0 = -C + 0 + 0 + \cdots$$

since all $\Sigma^j\phi(x)$ vanish when $x = k$. Thus $C = 0$,** and

$$\Sigma u_x\phi(x) = u_x\Sigma\phi(x) - \Delta u_x\Sigma^2\phi(x+1) + \cdots \qquad (30)$$

holds for all admissible values of x, i.e., $k \leq x \leq k + N - n$.

*Paragraph 1.

**A more rigorous demonstration is given in Appendix A.

Again remembering the definition of the operator Σ, we can evaluate the expression in Equation (24) as follows:

$$\sum_{x=k}^{k+N-n} (N-x)^{(n-k)}\, x^{(k)} = \Sigma u_{k+N-n}\phi(k+N-n) + (n-k)^{(n-k)}(k+N-n)^{(k)} \quad . \quad (31)$$

Utilizing Equations (28) and (29) to write down the full expansion, we obtain

$$\begin{aligned}\sum_{x=k}^{k+N-n} (N-x)^{(n-k)}\, x^{(k)} = {} & (k+N-n)^{(k)}\, \underline{|n-k} \\ & + \underline{|n-k}\, \frac{(N-n+k)^{(k+1)}}{k+1} \\ & + (n-k)\, \underline{|n-k-1}\, \frac{(N-n+k+1)^{(k+2)}}{(k+2)^{(2)}} \\ & + (n-k)^{(2)}\, \underline{|n-k-2}\, \frac{(N-n+k+2)^{(k+3)}}{(k+3)^{(3)}} \\ & + \cdots + (n-k)^{(n-k)}\, \underline{|0}\, \frac{N^{(n+1)}}{(n+1)^{(n-k+1)}} \quad . \quad (32)\end{aligned}$$

Since $0^{(0)} = \underline{|0} = 1$, and since $\underline{|n-k}$ is a factor of every term, the last equation becomes

$$\begin{aligned}\sum_{x=k}^{k+N-n} (N-x)^{(n-k)}\, x^{(k)} = \underline{|n-k} \Bigg\{ & (N-n+k)^{(k)} + \frac{(N-n+k)^{(k+1)}}{k+1} \\ & + \frac{(N-n+k+1)^{(k+2)}}{(k+2)^{(2)}} + \frac{(N-n+k+2)^{(k+3)}}{(k+3)^{(3)}} \\ & + \cdots + \frac{N^{(n+1)}}{(n+1)^{(n-k+1)}} \Bigg\} \quad . \quad (33)\end{aligned}$$

To sum the series inside the braces, we return to the list of Equation (29) and notice that, except for the first term, we have exactly the values taken on when $x = N-n+k$. The first term, of course, is $\phi(N-n+k)$. Whence we can write symbolically

$$\sum_{x=k}^{k+N-n} (N-x)^{(n-k)}\, x^{(k)} = \lfloor n-k \left\{ 1 + \frac{1}{\Delta} + \frac{E}{\Delta^2} + \frac{E^2}{\Delta^3} + \cdots + \frac{E^{n-k}}{\Delta^{n-k+1}} \right\} \phi(N-n+k)$$

$$= \lfloor n-k \left\{ 1 + \frac{1}{\Delta} \left[\frac{1 - \left(\frac{E}{\Delta}\right)^{n-k+1}}{1 - \frac{E}{\Delta}} \right] \right\} \phi(N-n+k)$$

$$= \lfloor n-k \left\{ 1 - \left[1 - \left(\frac{E}{\Delta}\right)^{n-k+1} \right] \right\} \phi(N-n+k)$$

$$= \lfloor n-k \left\{ \left(\frac{E}{\Delta}\right)^{n-k+1} \right\} \phi(N-n+k) \qquad (34)$$

From the basic definitions of the operators,

$$\left\{ \left(\frac{E}{\Delta}\right)^{n-k+1} \right\} \psi(y) = \Sigma^{n-k+1}\, \psi(y+n-k+1) \quad , \quad \text{and}$$

$$\therefore \sum_{x=k}^{k+N-n} (N-x)^{(n-k)}\, x^{(k)} = \lfloor n-k \; \Sigma^{n-k+1}\, \phi(N+1) \quad . \qquad (35)$$

Again referring to Equation (29),

$$\sum_{x=k}^{k+N-n} (N-x)^{(n-k)}\, x^{(k)} = \lfloor n-k \, \frac{(N+1)^{(n+1)}}{(n+1)^{(n-k+1)}} = \frac{\lfloor n-k \; \lfloor N+1 \; \lfloor k}{\lfloor N-n \; \lfloor n+1} \quad . \qquad (36)$$

Substituting this value into Equation (24) yields the desired integral

$$Q_x = \sum_{x=k}^{n+N-m} p(N,x,n,k) = \frac{N+1}{n+1} \quad . \qquad (37)$$

Finally, the required probability function in x is

$$P_x = P_x\,(N,x,n,k) = \frac{n+1}{N+1}\, p(k) = \frac{\lfloor N-n \; \lfloor N-x \; \lfloor n+1 \; \lfloor x}{\lfloor (N-x) - (n-k) \; \lfloor n-k \; \lfloor x-k \; \lfloor N+1 \; \lfloor k}$$

$$= \frac{C(N-x,\, n-k) \cdot C(x,k)}{C(N+1,\, n+1)} \qquad (38)$$

Note that this probability function differs by only a constant multiplier from the original function, p(k), given in Equation (22). However, x and n no longer are interchangeable, due to the presence of the factor n+1.

The technique of Equations (27), (28), (29), and (30) can be used to sum any number of terms of the probability integral. Thus, provided only that m is some proper value of x, $(k \leq m \leq k + N - n)$,

$$\frac{\lfloor N-n \rfloor n+1}{\lfloor n-k \rfloor N+1 \lfloor k} \sum_{x=k}^{m-1} (N-x)^{(n-k)} x^{(k)} = \frac{\lfloor N-n \rfloor n+1}{\lfloor n-k \rfloor N+1 \lfloor k} \Sigma u_m \phi(m) \quad , \qquad (39)$$

gives the probability that FEWER than m defectives will be found in N. As previously noted, (Equation (27)), the right-hand side of Equation (39) will contain n-k+1 terms. An alternate expansion for $\Sigma u_m \phi(m)$ which sums in fewer terms whenever $m < n$ is given in Appendix B. This alternate expansion is preferable for programming.

The graph of the function is, of course, a histogram composed of rectangles of equal width but varying height. (Figure 6). For any arbitrary value of x, the area (integral) of the corresponding rectangle can be computed by Equation (38). The combined area of any number of consecutive rectangles can be computed by Equation (39) and interpreted as a level of confidence.

The inverse problem is not so clear-cut, however, since no attempt is made to attach meaning to "a portion of a rectangle." Thus, any assigned confidence level must include the phrase "greater than" or "less than." Repeated application of Equation (39) to successive values of x will reveal the correct answer. It may be useful to employ Equation (13) to obtain a fairly close first approximation.

Borrowing the terminology of Paragraph 3 and referring to Figure 6, we can make statements like:

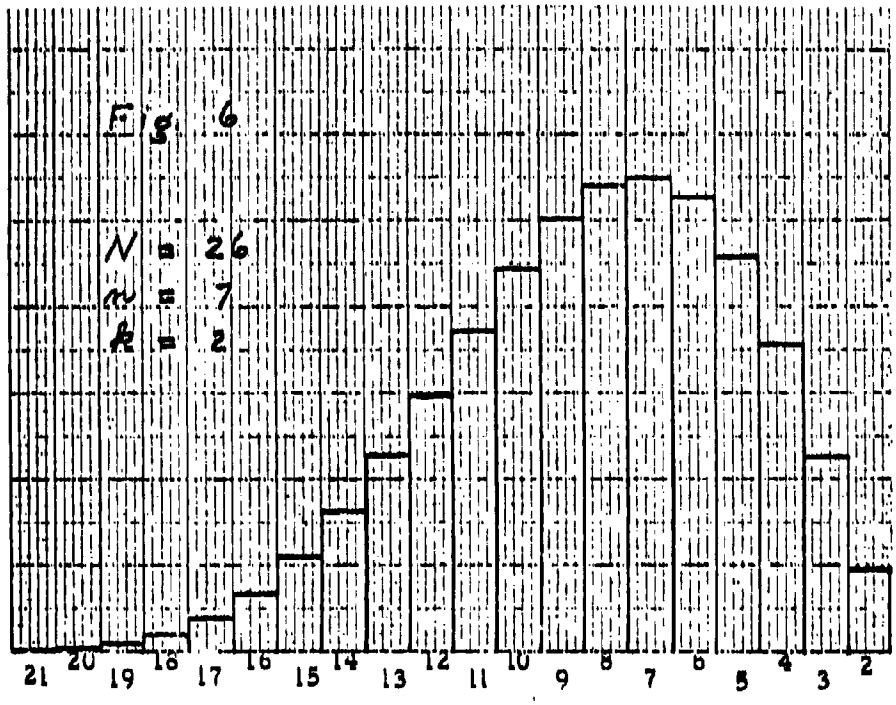
"That $x < 12$ exceeds the 80% confidence level," or

"Best estimate of x: $\hat{x} = 7$, and at the 0.74891 level of confidence, $\tilde{x}_1 = 11$ and $\tilde{x}_2 = 4$;" i.e., $4 \leq x \leq 11$.

5. COMPUTATIONAL PROCEDURES.

A. Significant Digits. The occurrence of large factorials really permits no alternative to computation by logarithms. Now two processes which are prodigal of significant digits are subtraction of nearly equal numbers and computing antilogarithms." We can be subjected to both hazards within the same algorithm. Therefore, it is suggested that computations be carried to 12 or 13 significant digits. For machines which do not compute logarithms accurately enough, the following is suggested:

"$\ln_e \lfloor 1000$ Is of the order of 5300. Four significant digits will be lost when subsequently passing to an antilogarithm.



X	P_x	X	P_x
2	0.01915	12	0.05952
3	0.04547	13	0.04522
4	0.07117	14	0.03246
5	0.09166	15	0.02185
6	0.10475	16	0.01362
7	0.10999	17	0.00772
8	0.10806	18	0.00386
9	0.10034	19	0.00162
10	0.08854	20	0.00051
11	0.07440	21	0.00009

Express the number in scientific notation thus: $378 = 3.78 \times 10^2$. If the resulting units digit is 1, proceed directly. If it is 2, 3, 4, or 5, divide the left-hand member by $e = 2.718\ 281\ 828\ 459$, intending to add $\ln_e e = 1$ to the result later. If the units digit is 6, 7, 8, or 9, divide by e^2 and add 2 later. Call the resulting number y. For our present example,

$$y = \frac{3.78}{e} = 1.39 \cdots$$

Now use the transformation

$$\xi = \frac{y-1}{y+1} \ .$$

The series,

$$\tfrac{1}{2}\ln_e y = \xi + \frac{\xi^3}{3} + \frac{\xi^5}{5} + \frac{\xi^7}{7} + \cdots \ , \tag{40}$$

will converge rapidly.* The exponent is recovered by adding or subtracting $\ln_e 10 = 2.302\ 585\ 092\ 994$, a suitable number of times.

B. Stirling's Formula for $\lfloor n$. Bernouilli's Numbers.

Stirling's formula for $\lfloor n$ is

$$\lfloor n = \frac{n^n \sqrt{2\pi n}}{e^{nS}} \tag{41}$$

where S is the asymptotic series

$$S = 1 - \frac{B_1 n^{-2}}{1 \cdot 2} + \frac{B_2 n^{-4}}{3 \cdot 4} - \frac{B_3 n^{-6}}{5 \cdot 6} + \cdots \ . \tag{42}$$

The B_j are Bernouilli's numbers, the first six of which are

$$B_1 = \frac{1}{6} \qquad B_7 = \frac{1}{30}$$

$$B_2 = \frac{1}{30} \qquad B_9 = \frac{5}{66}$$

$$B_3 = \frac{1}{42} \qquad B_{11} = \frac{691}{2730}$$

*To continue the example, $\xi = \frac{0.39\cdots}{2.39\cdots} = 0.163\cdots$ and the eighth term is 1.02×10^{-13}.

For thirteen-digit accuracy, $n > 11$ requires four terms of the series S, $n > 39$ but three terms. Thus,

$$\ln_e \lfloor n = \tfrac{1}{2}\ln_e(2\pi n) + n(\ln_e n - 1) + \frac{1}{12n}\left[1 - \frac{1}{30n^2}\left(1 - \frac{2}{7n^2}\right)\right], \quad n > 11$$

or

$$\ln_e \lfloor n = \tfrac{1}{2}\ln_e(2\pi n) + n(\ln_e n - 1) + \frac{1}{12n}\left(1 - \frac{1}{30n^2}\right), \quad n > 39 \quad . \qquad (43)$$

Logarithms of smaller factorials must be computed directly, of course.

C. Newton's Method. For the solution of otherwise-difficult inverses, Newton's method of successive approximations is indispensable. However, certain precautions must be taken by the programmer.

Ideally, the graph of the function is an ogive. But it serves the purpose equally well if two values of the argument can be found which surely bracket the desired solution and between which the function behaves like an ogive.*

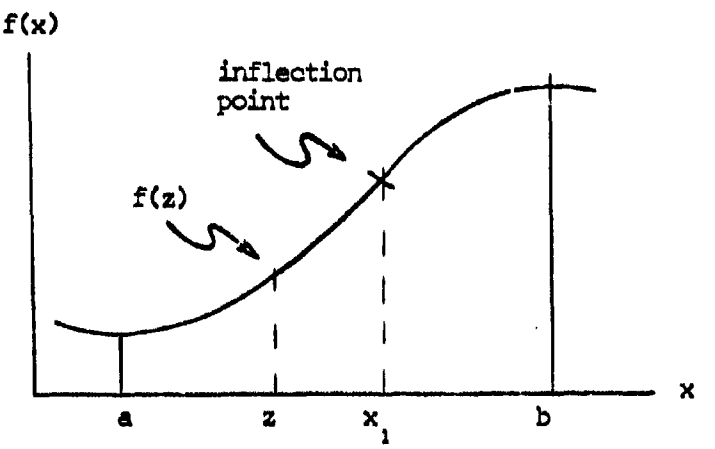


FIGURE 7

The basic operation, of course, is

$$x_{i+1} = x_i + \frac{f(z) - f(x_i)}{f'(x_i)} \quad , \qquad (44)$$

*i.e., bounded by a maximum and a minimum, with a single point of inflection between.

f(z) being given, and from which it is required to find z. Let the first approximation be taken at the inflection point. Since the slope is steepest there, it insures that the approximate solutions will not overshoot the true one. Thus the x_i's will remain within bounds, avoiding a spurious solution or runaway.

Were we to express the cumulative probability of Equation (12) as a function of z,

$$F(z) = \int_{r=0}^{z} f(r)\, dr \quad , \tag{45}$$

we would find that its graph is a true ogive, that its derivative is simply $F'(z) = f(r)$, and that the inflection point occurs at $\hat{r}$.

D. Summation of Series. Many of the formulae herein developed for use involve the summation of series. A convenient way of handling this type of computation in a programmable calculator is to discover and employ a term-to-term recurrence relationship.

Usually, infinite series offer no problem. For example, in Equation (40) we can choose to assign only odd subscripts to terms, whence

$$T_{i+2} = \frac{i\xi^2}{i+2}\, T_i \quad .$$

$\theta(i) = \frac{i\xi^2}{i+2}$ is known as the recurrence ratio. It is of most use to the programmer when it is a constant or a function of position only.

Finite series ostensibly offer a choice -- they can be summed from either end. Not really. When there are only a few terms, it probably makes no difference. But when there are many, the least term always should be left until last. There are three compelling reasons for this:

(1) The earlier the large terms are computed, the less accumulated round-off or truncation error they will contain.

(2) When employing a recurrence ratio, no term can contain more significant digits than the first term. In a fixed point machine, computing the least significant term first may result in complete disaster.

(3) If some terms are insignificant, it is unnecessary to waste computer time on them, provided the significant terms are computed first. In this case, the effect is quite similar to summing an infinite series.

E. Exiting a Loop. Many of the formulae developed can advantageously employ an iterative process in the computation.* A program must employ some device for terminating this process; i.e., exiting the loop. Basically, there are two cases which must be treated.

The first occurs when the number of iterations is known, or can be determined readily. The programmer merely finds a factor (or sets up a dummy index) which is known to reach zero eventually, and tests it.

The second (and more sensitive) obtains when the number of iterations depends upon the results of the calculations. It is a mistake to test the untreated single term, since it may become insignificant to the result, but yet not zero. It is tempting to test the difference between two successive solutions, but it is possible (particularly with Newton's method) to reach two alternating solutions which differ only in the least significant digit. A nearly foolproof procedure is to establish a maximum allowable error (call it δ), subtract it from the absolute value of the quantity in question, then test the sign of the difference. It may be necessary or desirable to choose a δ which squanders two or three (ostensibly) significant digits, in order to hasten the exit.

*Examples: Summation of Series, Newton's Method, Factorials.

APPENDIX A

EVALUATING A CONSTANT OF INTEGRATION

Equation (27) states

$$\Sigma u_x \phi(x) = -C + u_x \Sigma \phi(x) - \Delta u_x \Sigma^2 \phi(x+1) + \Delta^2 u_x \Sigma^3 \phi(x+2) - \cdots$$

$$\cdots + (-1)^j \Delta^j u_x \Sigma^{j+1} \phi(x+j) + \cdots \qquad (27)$$

and it is required to evaluate C, the constant of integration.

Now the admissible values of x are

$$k \leq x \leq k + N - n$$

and the fastidious may object to the development and inclusion of an expression like

$$\sum_{x=0}^{k-1} u_x \phi(x) = 0$$

So, let us increase the upper limit by unity. That the expression

$$\sum_{x=0}^{k} u_x \phi(x)$$

has a real sum, and that the sum is

$$\sum_{x=0}^{k} u_x \phi(x) = (N-k)^{(n-k)} \, \underline{|k}$$

there can be no doubt.

Continuing in the manner of Equation (30), we have

$$\Sigma u_{k+1} \phi(k+1) = \sum_{x=0}^{k} u_x \phi(x) = (N-k)^{(n-k)} \, \underline{|k}$$

$$= -C + (N-k-1)^{(n-k)} \frac{\underline{|k+1}}{k+1} + (n-k)(N-k-2)^{(n-k-1)} \frac{\underline{|k+2}}{(k+2)^{(2)}}$$

$$+ (n-k)^{(2)} (N-k-3)^{(n-k-2)} \frac{\underline{|k+3}}{(k+3)^{(3)}} + \cdots .$$

If we transpose -C, then

$$\frac{\lfloor k+h}{(k+h)^{(h)}} = \lfloor k$$

is a factor of the right-hand side, so that

$$C + (N-k)^{(n-k)}\ \lfloor k = \lfloor k \left\{ (N-k-1)^{(n-k)} + (n-k)(N-k-2)^{(n-k-1)} \right.$$
$$\left. + (n-k)^{(2)}(N-k-3)^{(n-k-2)} + \cdots + (n-k)^{(n-k)}(N-n-1)^{(0)} \right\} .$$

If we substitute $x = k+1$ into Equation (28), we obtain exactly the succession of terms exhibited within the braces above. This allows us to write symbolically

$$C + (N-k)^{(n-k)}\ \lfloor k = \lfloor k \left\{ 1 - \Delta + \Delta^2 - \Delta^3 + \cdots + (-1)^{n-k}\Delta^{n-k} \right\} u_{k+1}$$
$$= \lfloor k \left\{ \frac{1 + \Delta^{n-k+1}}{E} \right\} u_{k+1} = \lfloor k \left\{ \frac{1}{E} \right\} u_{k+1} \quad ,$$

since, by Equation (28), $\Delta^{n-k+1} u_x = 0$ for all x.

$$\therefore\ C + (N-k)^{(n-k)}\ \lfloor k = \lfloor k\ E^{-1} u_{k+1} = \lfloor k\ u_k = \lfloor k\ (N-k)^{(n-k)} \quad ;$$

i.e., $C = 0$. Q.E.D.

APPENDIX B

AN ALTERNATE EXPANSION OF EQUATION (25)

An alternate expansion of Equation (25) follows. As before, let

$$u_x = (N-x)^{(n-k)}$$

and

$$\phi(x) = x^{(k)}$$

Then,

$$C + \Sigma u_x \phi(x) = \Delta^{-1} u_x \phi(x) = (EE' - 1)^{-1} u_x \phi(x) \quad , \tag{25}$$

where, temporarily, E operates on u alone, E′ on ϕ alone. Continuing in a different manner,

$$(EE' - 1)^{-1} u_x \phi(x) = [E(1 + \Delta') - 1]^{-1} u_x \phi(x) = (\Delta + E\Delta')^{-1} u_x \phi(x)$$

$$= \frac{1}{E\Delta'} \left(1 + \frac{\Delta}{E\Delta'}\right)^{-1} u_x \phi(x)$$

$$= \frac{1}{E\Delta'} \left\{1 - \frac{\Delta}{E\Delta'} + \left(\frac{\Delta}{E\Delta'}\right)^2 - \left(\frac{\Delta}{E\Delta'}\right)^3 + \cdots\right\} u_x \phi(x) \tag{B-1}$$

From Equation (B-1), we can write the desired expansion, once more dropping the primes as no longer necessary.

$$\Sigma u_x \phi(x) = -C + u_{x-1} \Sigma \phi(x) - \Delta u_{x-2} \Sigma^2 \phi(x) + \Delta^2 u_{x-3} \Sigma^3 \phi(x)$$

$$- \cdots + (-1)^j \Delta^j u_{x-j-1} \Sigma^{j+1} \phi(x) + \cdots \tag{B-2}$$

Again it is useful to list a breakdown of terms.

$$
\begin{aligned}
u_x &= (N-x)^{(n-k)} \\
u_{x-1} &= (N-x+1)^{(n-k)} \\
-\Delta u_{x-2} &= (n-k)(N-x+1)^{(n-k-1)} \\
\Delta^2 u_{x-3} &= (n-k)^{(2)}(N-x+1)^{(n-k-2)} \\
&\vdots \\
(-1)^j \Delta^j u_{x-j-1} &= (n-k)^{(j)}(N-x+1)^{(n-k-j)} \\
&\vdots \\
(-1)^{n-k} \Delta^{n-k} u_{x-n+k-1} &= \underline{\lfloor n-k} \\
\Delta^{n-k+1} u_{x-n+k-2} &= 0
\end{aligned}
\tag{B-3}
$$

$$
\begin{aligned}
\phi(x) &= x^{(k)} \\
\Sigma\phi(x) &= \frac{x^{(k+1)}}{k+1} \\
\Sigma^2\phi(x) &= \frac{x^{(k+2)}}{(k+2)^{(2)}} \\
\Sigma^3\phi(x) &= \frac{x^{(k+3)}}{(k+3)^{(3)}} \\
&\vdots \\
\Sigma^{j+1}\phi(x) &= \frac{x^{(k+j+1)}}{(k+j+1)^{(j+1)}} \\
&\vdots \\
\Sigma^{n-k+1}\phi(x) &= \frac{x^{(n+1)}}{(n+1)^{(n-k+1)}}
\end{aligned}
\tag{B-4}
$$

Following the method of Appendix A,

$$\sum_{x=0}^{k} u_x \phi(x) = (N-k)^{(n-k)} \underline{|k} = \Sigma u_{k+1} \phi(k+1)$$

$$= -C + (N-k)^{(n-k)} \frac{(k+1)^{(k+1)}}{k+1} - 0 + 0 - \cdots \qquad \text{(B-5)}$$

since all $\Sigma^j \phi(x)$ vanish when $k + j > x$.

Simplifying,

$$(N-k)^{(n-k)} \underline{|k} = \Sigma u_{k+1} \phi(k+1) = -C + (N-k)^{(n-k)} \underline{|k}$$

and again $C = 0$. Thus, Equation (B-2) can be written

$$\Sigma u_x \phi(x) = u_{x-1} \Sigma \phi(x) - \Delta u_{x-2} \Sigma^2 \phi(x) + \cdots \quad , \qquad \text{(B-6)}$$

which holds for all admissible values of x. From Equations (B-3) and (B-4), it is apparent that when substituted into Equation (39), the expansion never will contain more than $n - k + 1$ terms, and will contain fewer whenever $m < n$.

Other expansions are possible, but usually prove to be more cumbersome than the two already developed.

APPENDIX C

PROGRAM PLANNING - BINOMIAL

1. INTRODUCTION. Reliability is expressed by z or by r, depending upon whether or not it is a limit of integration.

In general, loops will be exited by comparing the difference between two successive iterations with some standard, δ. (See Paragraph 5E.)

Nearly every formula of interest is greatly simplified if expressed as a function of f(r). Thus,

$$p(k) = \frac{1}{n+1} f(r)$$

$$f'(r) = \left[\frac{n-k}{r} - \frac{k}{1-r}\right] f(r) = \frac{1}{r}\left[n - \frac{k}{1-r}\right] f(r) .$$

$$\int_{r=0}^{z} f(r)\, dr = T_1 + T_2 + T_3 + \cdots$$

where

$$T_1 = \frac{zf(z)}{n+1-k}$$

and

$$T_{i+1} = \frac{h_i\, z\, T_i}{(n+2-h_i)(1-z)}$$

$$h_1 = k,\ h_2 = k-1,\ h_3 = k-2, \text{ etc.}$$

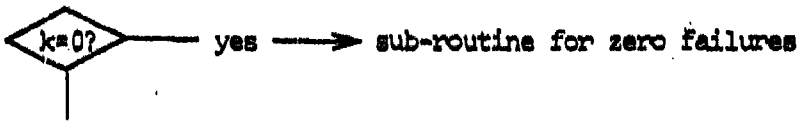
Note that if n and k do not change, there is no need to compute

$$\ln_e \lfloor n+1 - \ln_e \lfloor n-k - \ln_e \lfloor k$$

more than once.

2. COMPUTING L (z specified). Equations (12) and (13).

Enter data



no

Compute $\ln_e \frac{\lfloor n+1}{\lfloor k \lfloor n-k}$

Compute $(n-k) \ln_e z + k \ln_e (1-z)$

Add the above, yielding $\ln_e f(z)$

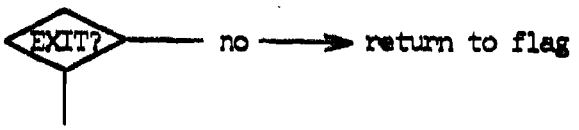
Compute and store T_1

Set $h_1 = k$

Flag

Compute T_{i+1} and add to partial sum

decremant h_i



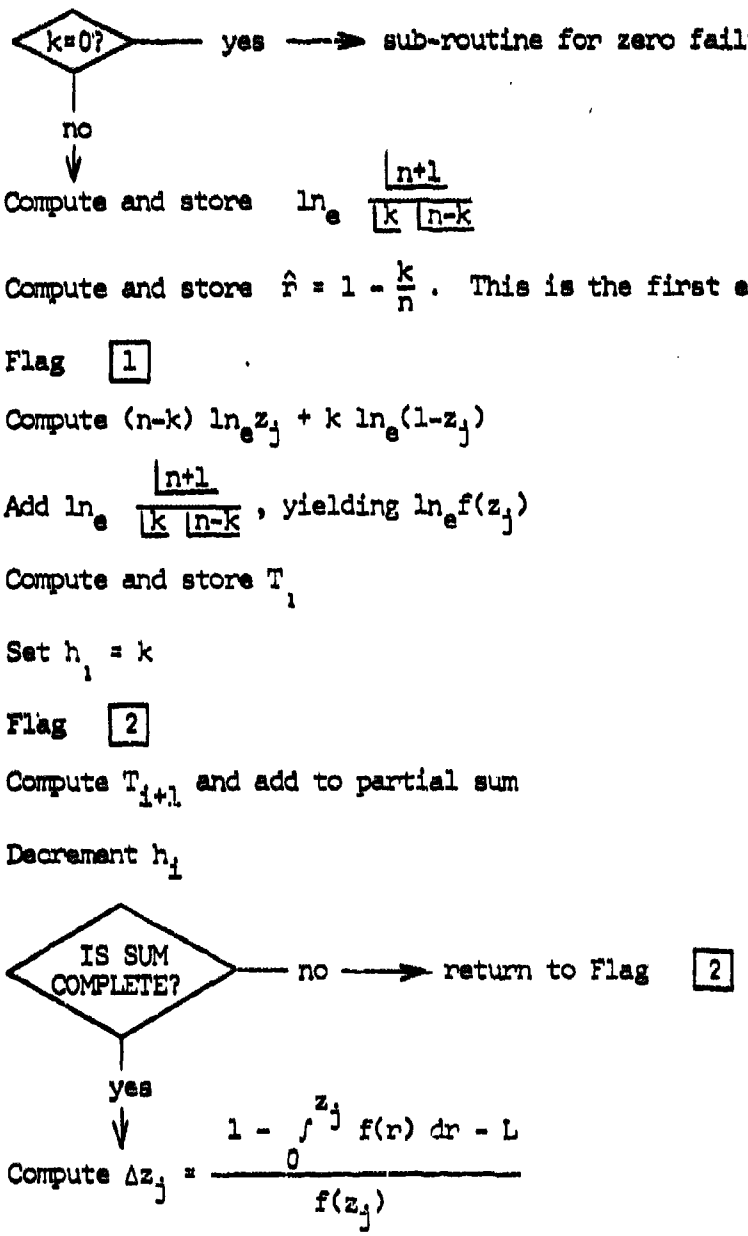
yes

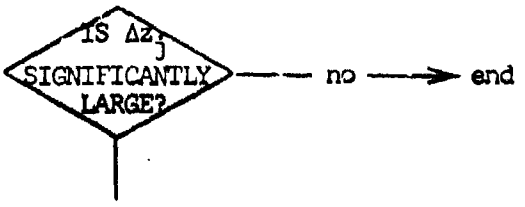
Subtract final sum from 1

End.

3. COMPUTING z (L specified). Equations (12), (13), (44), and (45).

Enter data



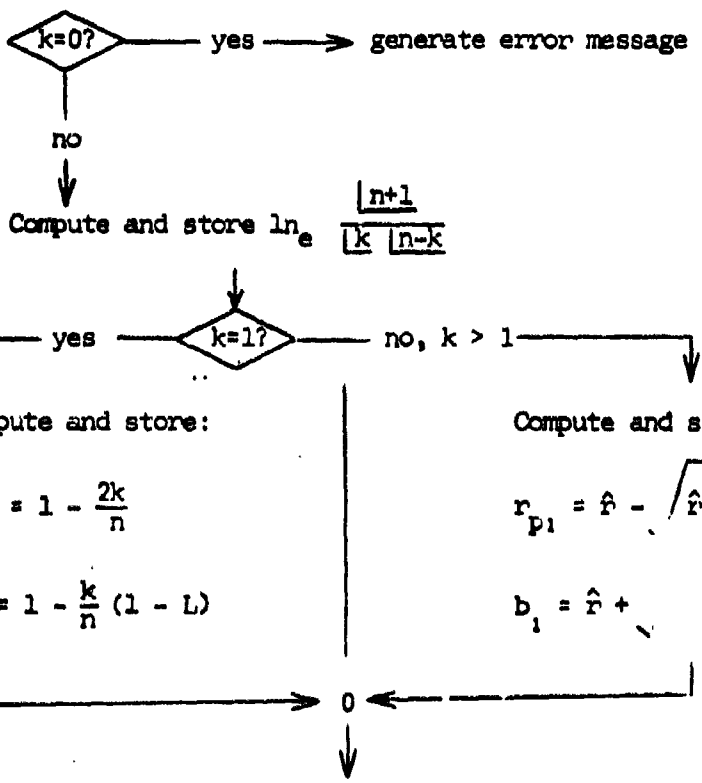


$z_{j+1} = z_j + \Delta z_j$. Store.

Return to Flag [1]

4. COMPUTING "BEST ESTIMATE OF THE RELIABILITY." (L specified). It might be said that the method employed is (Newton)^2. Therefore, it is mandatory that the program include realistic exit routines, in order to keep computer time within reason. Both $\tilde{z}_1$ and $\tilde{z}_2$ are computed. Equations (14), (15), (21), and Paragraphs 2 and 3 above are employed.

Enter data



The b_i are the successive estimates of $\tilde{z}_2$. The program will not run with $b_i = 1$, hence the above split. To continue:

Flag [1]

Compute $(n-k)\ \ln_e b_i + k\ \ln_e(1-b_i)$

Compute and store $f(b_i)$

To compute $\int_0^{b_i} f(r)\,dr$, call "integral" subroutine. Store.

Set $a_0 = r_{p_1}$

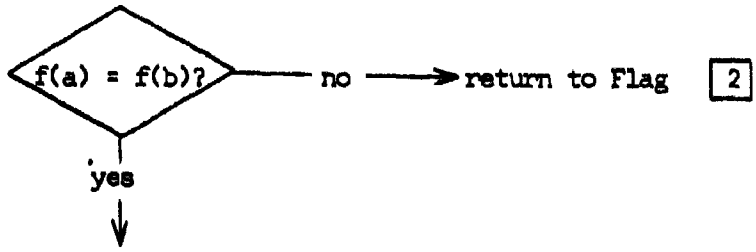
Flag [2]

Compute $(n-k)\ \ln_e a_j + k\ \ln_e(1-a_j)$

Compute $f(a_j)$

Compute $f'(a_j) = \frac{1}{a_j}\left[n - \frac{k}{1-a_j}\right] f(a_j)$

$$a_{j+1} = a_j + \frac{f(b_i) - f(a_j)}{f'(a_j)}$$



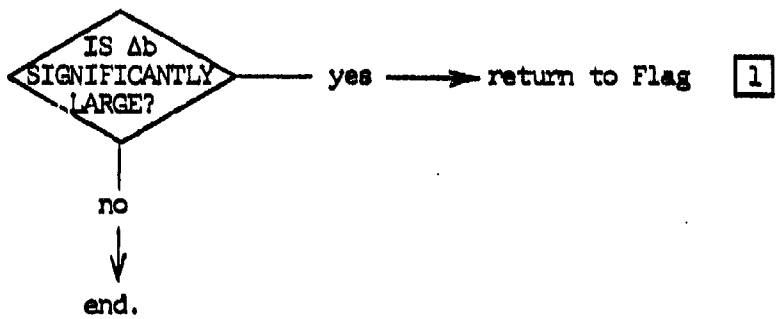
To compute $\int_0^{a_j} f(r)\,dr$, call "integral" subroutine

$$\Delta b_i = \left(\frac{L - \int_0^{b_i} f(r)\,dr + \int_0^{a_j} f(r)\,dr}{f(b_i)}\right) \cdot \left(\frac{f'(a_j)}{f'(a_j) - f'(b_i)}\right)$$

NOTE: Since at this point in the solution f(a) = f(b), the second fractional expression reduces to

$$\frac{n - \frac{k}{1-a}}{n - \frac{k}{1-a} - \frac{a}{b}\left(n - \frac{k}{1-b}\right)}$$

$b_{i+1} = b_i + \Delta b_i$



Subroutine "integral"

From x and f(x), compute and store

$$T_1 = (n+1-k)^{-1} \times f(x)$$

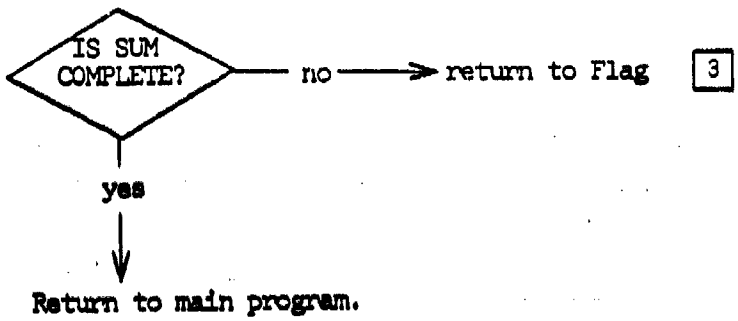
set $h_1 = k$

Flag 3

Compute and add to partial sum

$$T_{\ell+1} = \frac{h_\ell \times T_\ell}{(n+2-h_\ell)(1-x)}$$

Decrement h_ℓ


IS SUM COMPLETE?
no
return to Flag
3
yes
Return to main program.

APPENDIX D

PROGRAM PLANNING - HYPERGEOMETRIC

1. INTRODUCTION. The variable x is to be associated with the probability of a specific number of defectives. The variable m is to be associated with the cumulative probability that FEWER than the stated number of defectives exist.

In general, the series to be summed are all finite, but when both m and n are quite large, it will measureably hasten exiting the loop to compare the term with some arbitrarily small standard, δ, rather than zero.

The formulae of interest are conveniently expressed as functions of $p_x(N,x,n,k)$. Thus,

$$P_x = \frac{C(x,k)C(N-x,n-k)}{C(N+1,n+1)} \tag{D-1}$$

$$p(k) = \frac{N+1}{n+1} P_x \tag{D-2}$$

$$\sum_{x=k}^{m-1} P_x = T_1 + T_2 + T_3 + \cdots \tag{D-3}$$

where

$$T_1 = \frac{C(m,k+1)C(N-m+1,n-k)}{C(N+1,n+1)}$$

$$= \frac{m-k}{k+1} \cdot \frac{N-m+1}{(N-m-n)+(k+1)} \cdot P_m \tag{D-4}$$

and

$$T_{i+1} = \frac{m-(k+1)}{k+i+1} \cdot \frac{(n+1)-(k+i)}{(N-m-n)+(k+i+1)} \cdot T_i \tag{D-5}$$

Note that in implementing the above formula (D-5), the factors of the numerator should be computed before incrementing the index, the factors of the denominator after. In fact, under this scheme, the denominator of Equation (D-4) becomes equivalent to that of Equation (D-5), and the index can be set to k+1 before computing T_1.

2. COMPUTING p_x (x ARBITRARY). Equations (38) and (43)

Although the computation of nine logarithms is involved, there should be no difficulty encountered worthy of notice. It is preferable to compute and add (±) the largest logarithms last. ($\lfloor N+1$, $\lfloor N-m$, and $\lfloor N-x$ will be the largest numbers.)

3. COMPUTING $\sum_{x=k}^{m-1} p_x$ (m SPECIFIED). Equations (39) and (B-6)

The program plan is left to the reader. Sufficient suggestions should be found in Paragraph 2, Appendix C, and in Paragraphs 1 and 2, above.

4. COMPUTING m (L SPECIFIED). See Paragraph 4, "Hypergeometric Probability."

The problem is to find an integral value of m such that

$$\sum_{x=k}^{m-2} p_x < L < \sum_{x=k}^{m-1} p_x \quad . \tag{D-6}$$

A first approximation is obtained by using the method of Appendix C, Paragraph 3 to solve for z from the observed values of n and k, then applying the transformation

$$m = k - \frac{1}{2} + (1-z)(N-n+1) + \varepsilon \quad . \tag{D-7}$$

The quantity $0 \leq \varepsilon < 1$ is necessary to insure that m is an integer. A study of Figure 6 will reveal why Equation (D-7) is a suitable transformation.

In actual practice, ε need not even be determined. Instead, the estimate of m from Equation (D-7) is truncated at the decimal point, yielding m-1.

Next $\sum_{x=k}^{m-2} p_x$ is computed from the estimate of m-1, (see Paragraph 3, above). It is not necessary to compute the second integral of Equation (D-6), since

$$\sum_{x=k}^{m-1} p_x = p_{m-1} + \sum_{x=k}^{m-2} p_x \tag{D-8}$$

and both members of the right-hand side already are available.

If the inequality (D-6) holds, the problem is solved. If not, the estimate of m-1 is adjusted by unity and the last process repeated. (Only p_m or p_{m-2}, as the case may be, need be computed.)

5. COMPUTING THE "BEST ESTIMATE OF x, THE NUMBER OF DEFECTIVES." (L SPECIFIED).

The problem is similar to that discussed in Paragraph 4, above.

First approximations to $\tilde{x}_1$ and $\tilde{x}_2$ are obtained by using the method of Appendix C, Paragraph 4, to compute $\tilde{z}_1$ and $\tilde{z}_2$, then transforming the variables.

Several values of both the simple and cumulative probabilities are computed for arguments near the estimates of $\tilde{x}_1$ and $\tilde{x}_2$. The results are tabulated and inspected. The simple rectangles are discarded one at a time, beginning with the smallest in area. The process stops when one more step would reduce the remaining integral (area) to less than the value of L.

The "maximum likelihood estimate" is merely the value of x associated with the tallest rectangle.

ESTIMATION AND PREDICTION OF CONFIDENCED RELIABLE LIFE FROM SMALL SAMPLE SIZES

Eugene E. Coppola
Benet Weapons Laboratory
Watervliet Arsenal
Watervliet, New York

RELIABLE LIFE AND ITS LOWER CONFIDENCE BOUND. Reliable life is that time S during which a specified proportion R of a population of devices will operate continuously without failure. The proportion R is called the reliability. Reliable life is important for devices that fail catastrophically, that is, failure of the devices generally results in distruction of the devices and possibly destruction of surrounding equipment and injury or death to operating personnel. Cannon components such as tubes, breeches and chambers fall into this category.

Since such catastrophic failures must be avoided, it is important that the device be operated only during the time when the probability of failure is low. The reliable life for a new device, however, is not known and hence it must be estimated from test data. In addition, for gun components, a confidence requirement is added. Namely, it must be shown with a specified confidence level C that the actual reliable life exceeds a given value. In other words, what we want is a lower confidence bound $\tilde{S}$ at confidence level C for the reliable life S. The lower confidence bound $\tilde{S}$ will be called lower confidenced reliable life (LCRL). Note that when applied to cannon components, reliable life is usually called safe life.

The testing of cannon components is quite costly and time-consuming. Consequently economic and time considerations greatly limit the number of components that can be tested. This number, the sample size, is generally around six, although in some instances it has been as low as three and as high as 20. If the reliability were low, then this restriction of sample size would be relatively unimportant. However, the reliability for cannon components is generally specified to be 99.9 per cent. Further, the confidence C is generally specified to 90 per cent. On first glance, one might imagine that the smallness of the sample size would give highly undesirable results in calculating S. This is, however, not always the case, as we shall see below.

THE LOGNORMAL DISTRIBUTION. Because of the smallness of the sample size, non-parametric methods do not give satisfactory results. Consequently, it is necessary to assume that the failure times follow a distribution of known mathematical form. The lognormal and Weibull families of distributions are usually used for this purpose. In this paper, we will restrict ourselves to the lognormal family.

A real-valued random variable X is said to follow a lognormal distribution of X is positive with probability 1 and log X follows a normal distribution. The normal distribution of log X will depend the two usual parameters, μ and σ, defined by $\mu = E(\log X)$ and $\sigma = \text{Var}(\log X)$. These two parameters are also the parameters of the lognormal distribution of X.

In terms of the parameters, the reliable life S is given by:

$$S = \exp(\mu - \sigma z_R)$$

where z_R is the 100 R'th per cent point of the standard normal distribution.

Assume that m specimens have been tested to failure, with the failure times being $x_1, \ldots, x_m$. We further assume that the specimens were randomly selected from the population and that they are independent. Then the maximum likelihood estimates (MLE's) of μ, σ and S are given by:

$$\hat{\mu} = \frac{1}{m} \sum_{j=1}^{m} \log x_j \qquad \hat{\sigma}^2 = \frac{1}{m} \sum_{j=1}^{m} (\log x_j - \hat{\mu})^2$$

$$\hat{S} = \exp(\hat{\mu} - \hat{\sigma} z_R)$$

The LCRL $\tilde{S}_m$ is given by:

$$\tilde{S}_m = \exp(\hat{\mu} - \hat{\sigma} K_m)$$

Where K_m is a tolerance factor that depends on m, R and C. Values of K_m for various m, r and C have been tabulated and are readily available in the statistical literature. Note that since we are most interested in examining the sensitivity of the LCRL to the sample size m, we have added a subscript on to the LCRL notation $\tilde{S}m$ to emphasize that $\tilde{S}m$ is being calculated from a sample of size m.

STATISTICAL PROPERTIES OF $\tilde{S}m$. To eliminate the parameter μ, we consider $\tilde{S}m/S$ rather than $\tilde{S}m$. The distribution of $\tilde{S}m/S$, in fact, does not depend on the parameter μ; it does, however, depend on m, R and C and the parameter σ. Now, σ is generally not known. However, from past experience it appears that for cannon tubes and breeches, σ will be between 0 and 0.3 in the vast majority of cases, with an average value of about 0.2. The expected values of $\tilde{S}m/S$ and $(\tilde{S}m/S)^2$ are given by:

$$E\left(\frac{\tilde{S}_m}{S}\right) = \frac{2\exp[\sigma z_R + \sigma^2(\frac{1}{2m} + L_m^2)]}{\Gamma(\frac{m-1}{2})} \int_{\sigma L_m}^{\infty} (u-\sigma L_m)^2 e^{-u^2} du$$

$$E\left(\frac{\tilde{S}_m}{S}\right)^2 = \frac{2\exp[2\sigma z_R + 2\sigma^2(\frac{1}{m} + 2L_m^2)]}{} \int_{2\sigma L_m}^{\infty} (u-2\sigma L_m)^2 e^{-u^2} du$$

where $L_m = (2m)^{-1/2} K_m$

By evaluating these, the mean and variance of $\tilde{S}m/S$ may be determined.

Table 1 shows the expected value of $\tilde{S}m/S$ for R = 0.999, C = 0.9 and for various σ and m. Since we would like $\tilde{S}m$ to be close to S, values of $E(\tilde{S}m/S)$ close to 1 are most desirable. However, since $\tilde{S}m \leq S$ with probability C, we should have $E(\tilde{S}m/S) < 1$. As can be seen from Table 1, $E(\tilde{S}m/S)$ is much smaller than 1 for very small m. For example, for σ = 0.2 and m = 3, $E(\tilde{S}m/S)$ is approximately 50 per cent. This means that on the average, $\tilde{S}m$ will only be half as large as S. For the developer, this means that if a policy were adopted that LCRL were to be based exclusively on samples of 3, then the developer would have to insure that on the average the actual reliable life of the equipment be twice as large as the reliable life he desires to demonstrate. For this reason alone, a policy of basing LCRL on samples of 3 is highly undesirable.

It should be no surprise that m = 3 gives undesirable results. What is surprising is that for m not much larger than 3, the results are not too bad. The author finds it remarkable, for example, that for m = 10, one will obtain $\tilde{S}m$ on the average about 77 per cent of the actual S.

Table 2 shows the variance of $\tilde{S}m/S$. A variance near 0 is most desirable. As can be seen, the variances for very small m are relatively far from 0. For $\sigma = 0.2$, the variance is fairly small for $m \geq 6$ and changes relatively little with increasing m. Use of any of the standard inequalities such as Chebyshev's Inequality shows that even for relatively small m, $\tilde{S}m/S$ will tend to be fairly close to its expected value.

Hence, the restriction to small m, whole not ideal from a statistical viewpoint, is not especially damaging either, provided m is not too small. In fact when σ is close to 0, the LCRL will have quite good properties.

THE EFFECT OF INCREASING SAMPLE SIZES. Although in some cases small m may give acceptable results, in other cases small m may not be as desirable. Let us investigate the following question: m specimens have been tested. What would happen if we tested an additional k specimens and added them to the sample to give a sample size of m + k?

The question is of more than academic interest. For example, Table 3 shows the probability that by adding one more specimen to the sample, we can increase LCRL. As can be seen, it is likely that LCRL will increase. If the amount of increase is sufficiently large, it may be worthwhile to test one or two more specimens.

Assume that we have a sequence $X_1, X_2, X_3, \ldots$ of independent, randomly selected failure times. For each m, let

$$\mu_m = \frac{1}{m}\sum_{j=1}^{m} \log x_j \quad , \quad \theta_m^2 = \frac{1}{m}\sum_{j=1}^{m} (\log x_j - \mu_m)^2$$

$$\tilde{S}_m = \exp(\mu_m - \theta_m K_m)$$

$\hat{\mu}_m$, $\hat{\sigma}_m$ and $\tilde{S}_m$ are just the MLE's and LCRL that we would calculate from the first m failure times. To see the effect of adding k additional specimens to a sample of m, we want to study $\tilde{S}_m$ in relation to $\tilde{S}_{m+k}$.

Define

$$T_{m,k} = \frac{\log \tilde{S}_{m+k} - \log \tilde{S}_m}{\hat{\sigma}_m}$$

The distribution of Tm,k does not depend on either μ or σ, so that Tm,k can be used whatever the actual values of these parameters. A knowledge of the distribution of Tm,k is useful for the following reason: Once m specimens have been tested, we can calculate $\tilde{S}m$ and $\hat{\sigma}m$. The only unknown quantity in the definition of Tm,k is $\tilde{S}m+k$. Consequently, probability statements concerning Tm,k can be translated into probability statements concerning $\tilde{S}m+k$. In particular, we can construct prediction intervals for $\tilde{S}m+k$ in terms of $\tilde{S}m$ and m, as follows: Assume p ($0<p\leq 1$) is given and that we have determined two numbers t_1 and t_2 such that

$$Pr(t_1 \leq T_{m,k} \leq t_2) = p .$$

This last equation is equivalent to:

$$Pr(\tilde{S}_m \exp(\hat{\sigma}_m t_1) \leq \tilde{S}_{m+k} \leq \tilde{S}_{m+k} \exp(\hat{\sigma}_m t_2) = p$$

Consequently, $(\tilde{S}_m \exp(\hat{\sigma}_m t_1) , \tilde{S}_{m+k} \exp(\hat{\sigma}_m t_2))$ will be a prediction interval for $\tilde{S}m+k$ at level p.

An interesting fact about Tm,k is that with probability 1, Tm,k is bounded from above under a certain mild condition. In fact, T can be written in the following form:

$$\frac{C_1 z}{x^{1/2}} - C_2\left(1+\frac{y+z^2}{x}\right)^{1/2} + K_m \qquad (1)$$

where $C_1 = (\frac{k}{m+k})^{1/2}$

$C_2 = (\frac{m}{m+k})^{1/2} K_{m+k}$

z is standard normal, x is chi-squared with k-1 degrees of freedom, y is chi-squared with m-1 degrees of freedom and x, y, and z are independent. (Note: we assume here that $m \geq 2$. We allow $k = 1$ and interpret a chi-square variate with 0 degrees of freedom as a random variable which takes the value 0 with probability 1). The function given in(1)above will take a maximum value if $C_2 \geq C_1$, and in this case, the maximum value is $K_m - (C_2^2 - C_1^2)^{1/2}$ The condition $C_2 \geq C_1$ is equivalent to:

$$K_{m+k}^2 \geq \frac{k}{m} \qquad (2)$$

When this inequality is satisfied, the maximum value will be

$$H_{m,k} = K_m - (\frac{mK_{m+k}^{-k}}{m+k})^{1/2}$$

Consequently, when inequality 2 is satisfied,

$$Pr(T_{m,k} \leq H_{m,k}) = 1$$

This is equivalent to:

$\tilde{s}_{m+k} \leq \tilde{s}_m \exp(\hat{\sigma}_m H_{m,k})$ with probability 1. Inequality(2)will not be true for all m and k. (In fact, as $k \to \infty$, the left side of (2) approaches z_R^2 while the right side approaches ∞. Hence, inequality(2)will not be true for large k). However, (2) will be true for m and k of interest in this paper. For example, when $m = 6$, $R = 0.999$ and $C = 0.9$, inequality(2)will be true for $k \leq 50$.

The distribution of $T_{m,k}$ for $R = 0.999$, $C = 0.9$ and for some m and k have been determined by Monte-Carlo simulation. The cumulative distributions of $T_{m,k}$ for $m = 6$ and $k = 1, 2, 3$ and for $m = 10$ and $k = 1, 2,$

3 are shown. Note that for m = 10, the bulk of the distribution is concentrated near 0, so that adding up to 3 more samples to an already-existing sample of 10 will probably not produce much change in LCRL. For m = 6, the distribution is not so closely concentrated near 0. However, depending on the actual numbers involved, the prediction intervals may be fairly tight.

An Example

As an example, consider the following six failure times: 2596, 2536, 2811, 2141, 2416, 2839. We calculate from these:

$\hat{\mu}_6 = 7.841$

$\hat{\sigma}_6 = 0.09493$

$\tilde{S}_6 = 1427$, for $R = 0.999$, $C = 0.9$

Now suppose that the original test plan is to test 8 specimens, of which the first 6 gave the failure times above. Then prediction intervals for $\tilde{S}_7$ and $\tilde{S}_8$ at a level of 90% are

$\tilde{S}_7$ 1305, 1554

$\tilde{S}_8$ 1265, 1649

The figures on the right represent the upper bounds mentioned above. That is, with 100% confidence, $\tilde{S}_8 \leq 1649$.

Now, if the original aim of the test was to demonstrate a reliable life of 2000, it is clear that this will be impossible. For after all 8 specimens have been tested, the LCRL cannot be higher than 1649. Consequently, the testing can be halted after 6.

Suppose instead that only 1500 reliable life was desired. From the distribution of $T_{6,2}$, we can calculate that 1500 is a lower prediction

bound of $\tilde{S}_8$ at a level of approximately 70 per cent. One can therefore be fairly confident the test will show at least 1500 reliable life. On the other hand, if only 1250 reliable life were desired, then one can be about 90% confident that the final results will show a reliable life of at least 1250.

TABLE 1

Expected Value of $\tilde{S}_n/S_R$

$R = 0.999, \gamma = 0.9$

n \ σ	0.2	0.4	0.6	0.8	1.0
3	0.471	0.354	0.348	0.334	0.332
4	0.577	0.424	0.374	0.363	0.359
5	0.641	0.479	0.408	0.386	0.380
6	0.684	0.522	0.439	0.402	0.397
7	0.715	0.556	0.469	0.420	0.403
8	0.739	0.584	0.491	0.438	0.412
9	0.757	0.607	0.512	0.455	0.422
10	0.773	0.627	0.531	0.471	0.434
11	0.786	0.643	0.548	0.485	0.445
12	0.797	0.658	0.564	0.499	0.456
13	0.806	0.671	0.577	0.511	0.467
14	0.814	0.683	0.589	0.523	0.476
15	0.822	0.694	0.601	0.534	0.486
16	0.828	0.703	0.611	0.544	0.495
17	0.834	0.712	0.621	0.554	0.504
18	0.840	0.720	0.630	0.562	0.512
19	0.845	0.727	0.638	0.571	0.520
20	0.849	0.734	0.646	0.579	0.527
21	0.853	0.740	0.653	0.586	0.535
22	0.857	0.746	0.660	0.593	0.541
23	0.860	0.751	0.666	0.600	0.548
24	0.863	0.756	0.672	0.606	0.554
25	0.867	0.761	0.678	0.612	0.560

TABLE 2

Variance of $\tilde{S}_n/S_R$

$R = 0.999$, $\gamma = 0.9$

n \ σ	0.2	0.4	0.6	0.8	1.0
3	0.132	0.279	0.602	1.598	4.700
4	0.091	0.199	0.384	0.809	1.916
5	0.068	0.156	0.281	0.523	1.064
6	0.054	0.130	0.225	0.388	0.709
7	0.045	0.111	0.190	0.311	0.528
8	0.038	0.097	0.166	0.262	0.421
9	0.033	0.086	0.147	0.228	0.353
10	0.029	0.078	0.133	0.203	0.305
11	0.026	0.071	0.122	0.184	0.271
12	0.024	0.065	0.113	0.169	0.244
13	0.022	0.061	0.105	0.156	0.223
14	0.020	0.056	0.098	0.146	0.206
15	0.018	0.053	0.092	0.137	0.192
16	0.017	0.050	0.087	0.129	0.180
17	0.016	0.047	0.083	0.123	0.170
18	0.015	0.044	0.079	0.117	0.161
19	0.014	0.042	0.075	0.111	0.153
20	0.013	0.040	0.072	0.107	0.146
21	0.013	0.038	0.069	0.102	0.140
22	0.012	0.037	0.066	0.098	0.134
23	0.011	0.035	0.063	0.095	0.129
24	0.011	0.034	0.061	0.091	0.125
25	0.010	0.032	0.059	0.088	0.120

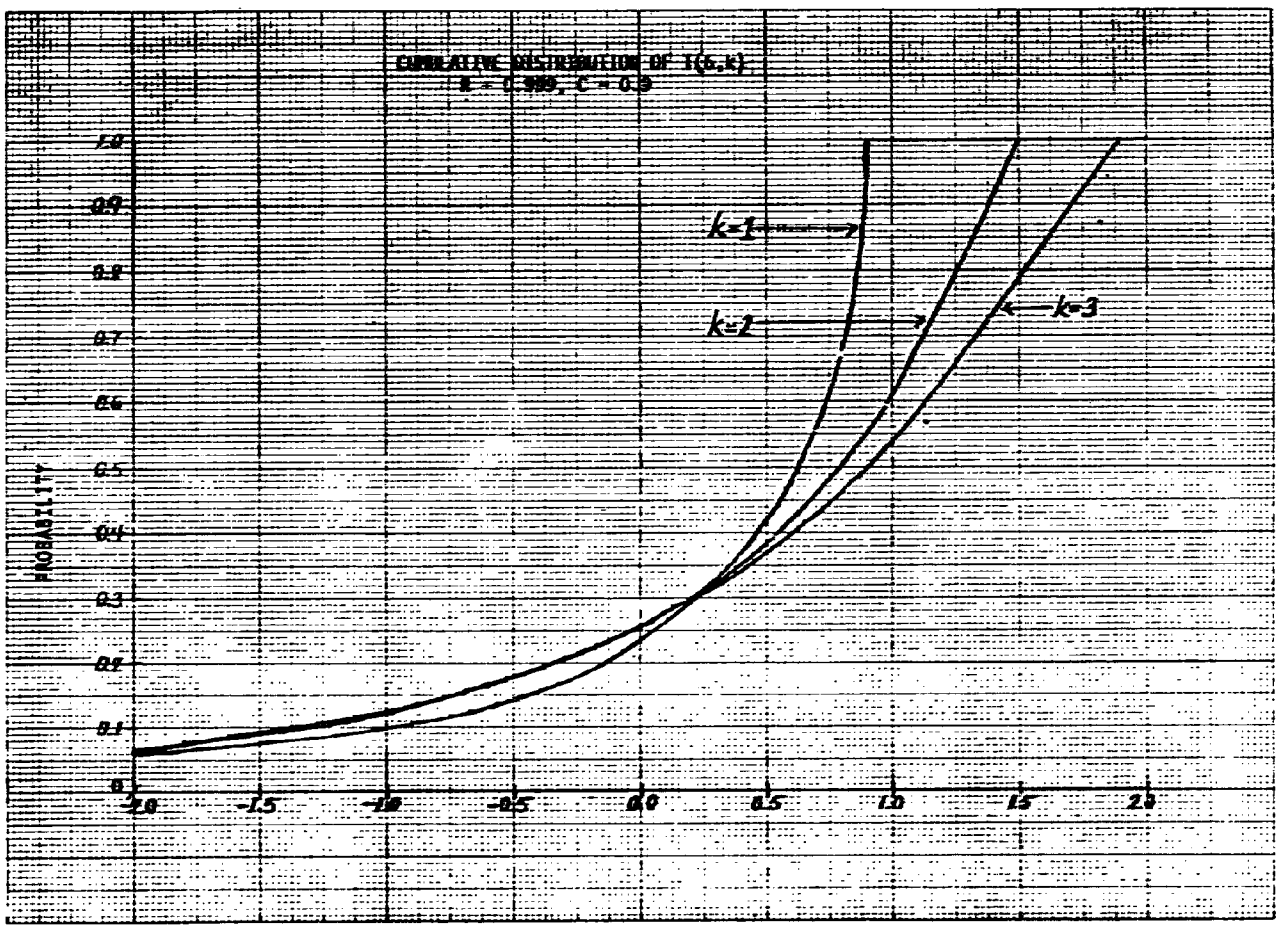
TABLE 3

$\Pr(\tilde{S}_{n+1} \geq \tilde{S}_n)$

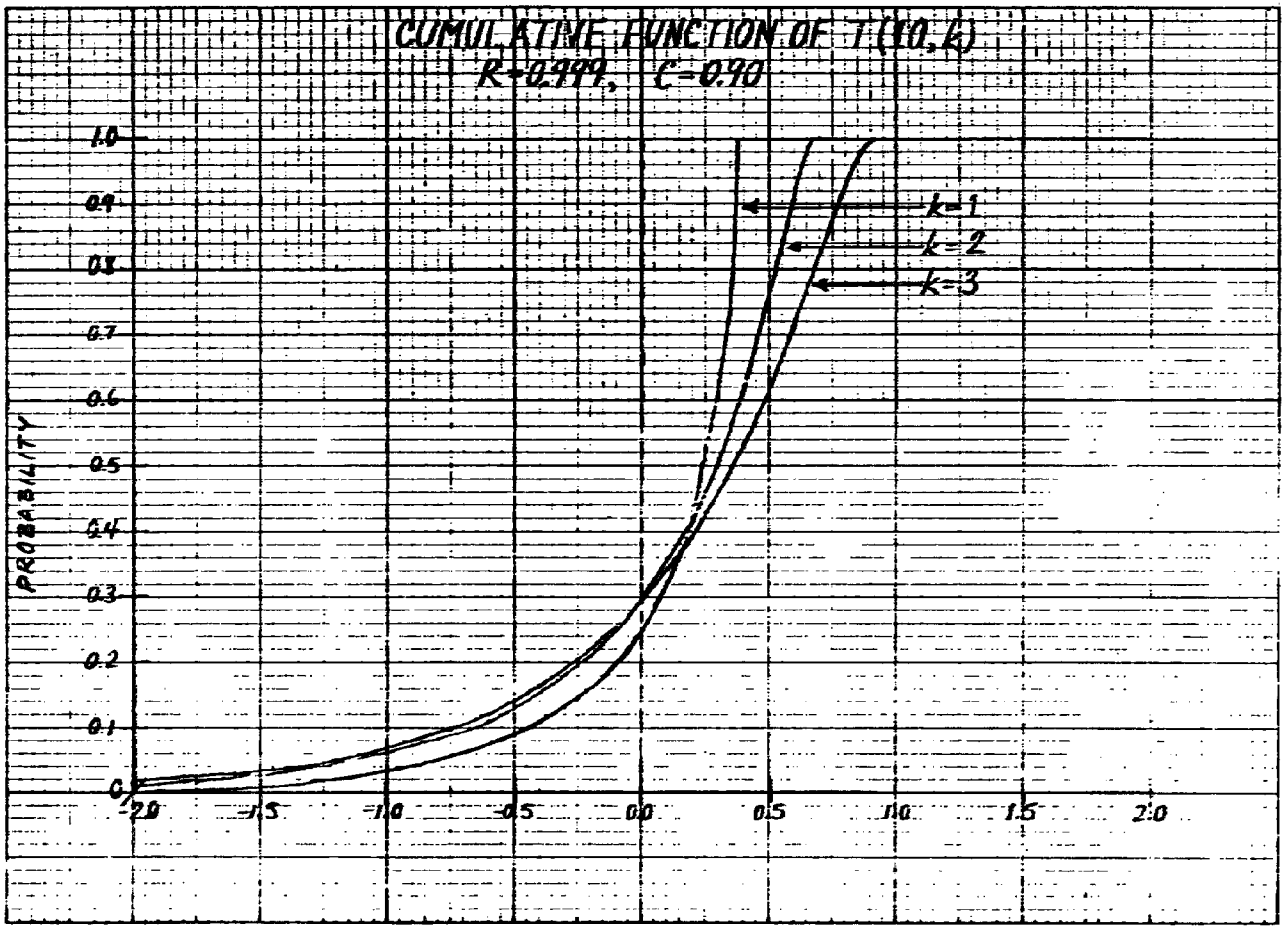
$\gamma = 0.9$

n \ R	0.999	0.99	0.95
2	.821	.819	.815
3	.797	.796	.793
4	.783	.782	.780
5	.774	.773	.771
6	.766	.765	.763
7	.760	.759	.758
8	.755	.755	.755
9	.752	.750	.749
10	.748	.747	.745
11	.746	.745	.744
12	.743	.742	.739
13	.741	.740	.736
14	.739	.737	.734
15	.736	.736	.733
16	.736	.736	.729
17	.733	.731	.728
18	.733	.730	.728
19	.730	.730	.728
20	.729	.729	.726
21	.728	.727	.725
22	.727	.727	.724
23	.726	.725	.723
24	.726	.725	.723

NO. 3404-20 KIETZKEN GRAPH PAPER
20 X 20 PER INCH
EUGENE DIETZGEN CO.
MADE IN U. S. A.



K+E 10 X 10 TO THE INCH 46 0780
7 X 10 INCHES
MADE IN U.S.A.



SEQUENTIAL ALLOCATION OF OBSERVATIONS IN THE EXPONENTIAL SELECTION PROBLEM

Robert M. Wharton, Ph.D.
Thomas Jefferson University
Philadelphia, Pennsylvania

R. Srinivasan, Ph.D.
Temple University
Philadelphia, Pennsylvania

ABSTRACT. Two sequential data-dependent allocation rules for assigning patients to clinical trials are explored in this paper. The objective of the designs is to test the null hypothesis that there is no difference in mean survival times associated with two treatments where survival time is assumed to be exponentially distributed and at the same time to minimize the number of patients assigned to the inferior treatment. Both the single patient and multiple patient entry cases are discussed.

1. INTRODUCTION. This paper is concerned with protocols for clinical trials in which we desire to determine whether there is any difference in the effects of two treatments. We will base our decision on some measurable response associated with a treatment (i.e. survival time, time to remission, etc.). We also assume that patients arrive for treatments sequentially in time either individually or in groups.

Most clinical trials addressing this question require approximately equal numbers of patients to be assigned to each treatment. Now suppose it becomes clear to the treating physician that one treatment is better than the other before sufficient patients have been accrued to reach a decision with the significance and power specified in the original trial design. He then faces an ethical problem. He can not continue to treat patients with an inferior treatment and yet by terminating the trial prematurely, he may lose information which would be invaluable in planning the treatment of many future patients.

To reduce this ethical problem, it would be useful to design the clinical trial using the data collected up to a given point to choose the treatment for a patient entering the trial at that point. The aim being a design which tends to assign the majority of the patients to the superior method of treatment, while meeting the classical statistical criteria of significance and power.

For the exponential selection problem in which we test the null hypothesis that there is no difference in mean survival times associated with the two treatments where survival times are assumed to be exponential with parameters (death rates) depending upon treatment, Flehinger and Louis (1971) have investigated a whole range of sequential data-dependent assignment rules ranging from strict alternation of treatment to assignment of treatment with lower estimated death rate. Clearly the most data-dependent allocation rule would be to assign the next patient to the treatment with the smallest expected death rate (maximum likelihood estimate). Unfortunately this rule would often have the

effect of allocating an overwhelming proportion of patients to one or the other treatment and thus extending the length of the trial indefinately (Armitage 1975). To reduce this difficulty Flehinger and Louis have proposed the following range of allocation rules:

Let D_{in} = the number of deaths of patients treated by method i by time n

T_{in} = the total time lived by patients treated by method i by time n

γ = be a constant between 0 and 1

then at time n

a) if $|D_{1n} - D_{2n}| > \gamma n$ and $D_{1n} < D_{2n}$ treatment 1 is used whereas if $D_{1n} > D_{2n}$ treatment 2 is used

b) if $|D_{1n} - D_{2n}| \leq \gamma n$ and $D_{1n}/T_{1n} \leq D_{2n}/T_{2n}$

Treatment 1 is used whereas if $D_{1n}/T_{1n} > D_{2n}/T_{2n}$ treatment 2 is used.

2. ALLOCATION RULES. We wish to examine two further allocation rules. The first treats the same situation as the Flehinger-Louis rules (i.e. exponential survival time, patients arriving sequentially over a period of time and being assigned immediately to a single treatment.) This allocation rule which we will refer to as R1 assigns treatment to the next three incoming patients based on accumulated data with two of the patients receiving treatment 1 if it has the smallest expected death rate and one patient then receiving treatment 2 and the reverse if treatment 2 is associated with the lowest expected death rate. If the two treatments have the same expected death rate then the treatment given to two of the next three patients is reversed from that of the previous triple of patients. This relatively simple rule overcomes the difficulty of an overwhelming proportion of patients going to any one treatment and is comparable to the Flehinger-Louis rules with respect to Average Sample Number (A.S.N.) and Inferior Treatment Number (I.T.N.).

It also has the advantage that it can be extended in a natural way to the case of multiple patient entry. We will only consider here the situation where three patients arrive for treatment every third day but the suggested approach can easily be extended to a more general setting. In the three patient entry case as in the above allocation rule we assign two of the next three patients

to the treatment which has the smallest expected death rate on the basis of accumulated data. We shall refer to this allocation rule as R3 . The protocols for clinical trials utilizing R1 and R3 are open sequential design which terminate when the likelihood ratio crosses a given boundry. This is also true of the Flehinger-Louis rules. Comparisons of R1 and R3 with two allocation rules denoted R2 and R4 involving strict alternation are presented in Section 3. These results have been obtained on the basis of computer simulation with 1000 replications for each entry.

3. DEFINITIONS AND SIMULATION RESULTS.

Hypothesis under consideration:

It is assumed that there are two treatments available. A patient is given one of these treatments at a point in time, after which his remaining life length has an exponential distribution, the death rates λ_1 and λ_2 depend upon the treatments. The clinical trial is intended to choose one of the following hypotheses:

$$H_0 : \lambda_1 = \lambda_2 \; ; \; H_1 : \lambda_2 = k\lambda_1 \; ; \; H_2 : \lambda_1 = k\lambda_2$$

where $k > 1$ is chosen in advance as a ratio which represents a medically significant difference.

Allocation Rules:

For any given time t , after the trial begins.

Let

X_{ijt} be the time lived since treatment if he is still alive for a patient given treatment i at time j.

Y_{ijt} be the time lived from treatment to death if he has died.

D_{it} be the number of deaths of patients treated by method i by time t.

T_{it} be the total time lived by patients treated by method i by time t.

We consider four allocation rules denoted R1 thru R4 .

R1 The patients arrive one per day but the treatment plan for the next three (3) days is defined every third day by randomly assigning one of the two treatments to two of the patients and the other treatment to the remaining patient in the triple. Which treatment is used twice is determined by the following rule:

If $D_{1t} / T_{1t} < D_{2t} / T_{2t}$ treatment 1 is used

if $D_{1t} / T_{1t} > D_{2t} / T_{2t}$ treatment 2 is used

if $D_{1t} / T_{1t} = D_{2t} / T_{2t}$ change from previous assignment.

R2 Is simple alternation of treatment one and two, with the treatment for the first patient randomly selected.

R3 Three patients arrive on the first day and every third day thereafter. The treatment received by the majority of the next triplet is determined by the same rule used in R1.

R4 The patients arrive as in R3. Treatment one or two is randomly selected and this treatment is randomly assigned to two of the 3 patients with the third patient receiving the other treatment. The treatment scheme is reversed for the next 3 patients.

Termination Rules

The termination rules considered utilize the liklihood ratio's

$$L_{1t} = k^{D_{2t}} \{ (T_{1t} + T_{2t}) / (T_{1t} + kT_{2t}) \}^{(D_{1t} + D_{2t})}$$

$$L_{2t} = k^{D_{1t}} \{ (T_{1t} + T_{2t}) / (kT_{1t} + T_{2t}) \}^{(D_{1t} + D_{2t})}$$

and are of the form: select two numbers A and B with $A < 1 < B$.

$\max (L_{1t}, L_{2t}) < A \rightarrow$ terminate and accept H_0.

$\max (L_{1t}, L_{2t}) > B \rightarrow$ Terminate and accept H_1 where 1 corresponds to the larger of L_{1t}, L_{2t}.

$A \leq \max (L_{1t}, L_{2t}) \leq B \rightarrow$ continue testing.

Flehinger and Louis (1971) showed that for $k = 2$, $A = .1$ and $B = 30$. give a significance level of .05 and a power of .95 . These values were used in the results that follow.

The authors are currently working on more extensive computer simulations of the schemes presented here for the exponential and similar results for the normal case. The implications of introducing further randomization and its effect on selection and trend bias are also being explored.

PERFORMANCE OF PROTOCOLS
A = .1, B = 30, k = 2

λ_1 = .10 R1

λ_2	O.C.	A.S.N.	I.T.N.
.03333	.999	47.8	18.3
.04	.993	56.2	21.3
.05	.953	76.2	28.6
.06667	.554	104.7	38.7
.075	.315	107.7	42.5
.10	.053	82.3	--
.15	.555	107.3	40.3
.20	.955	71.4	26.1
.25	.994	49.7	18.0
.30	1.0	40.1	14.7

*** *** *** ***

λ_1 = .10 R2

λ_2	O.C.	A.S.N.	I.T.N.
.03333	.999	44.0	22.0
.04	.995	53.0	26.5
.05	.944	70.0	35.0
.06667	.568	97.0	48.5
.075	.319	97.0	48.5
.10	.050	76.0	38.0
.15	.564	94.0	47.0
.20	.955	64.0	32.0
.25	.992	47.0	23.5
.30	.999	37.0	18.5

*** *** *** ***

λ_1 = .10 R3

λ_2	O.C.	A.S.N.	I.T.N.
.03333	1.0	47.0	17.7
.04	.998	52.3	19.7
.05	.968	67.6	25.3
.06667	.615	98.8	37.9
.075	.397	100.2	39.5
.10	.057	81.1	--
.15	.565	106.3	39.7
.20	.961	72.3	27.5
.25	.995	48.5	17.8
.30	1.0	38.5	14.0

$\lambda_1 = .10$ R4

λ_2	O.C.	A.S.N.	I.T.N.
.03333	1.0	43.5	21.7
.04	.997	52.1	26.1
.05	.965	66.2	33.1
.06667	.579	104.5	52.3
.075	.293	102.7	51.4
.10	.052	79.5	39.8
.15	.558	101.9	50.9
.20	.956	65.4	32.7
.25	.992	43.7	21.9
30	1.0	34.9	17.4

*** *** ***

References

1. Armitage, P. (1975) Sequential Medical Trials, Blackwell

2. Flehinger, B.J., Louis, T.A. (1971) Sequential Treatment Allocation in Clinical Trials, Biometrika 58, pp. 419-26.

Maximum Likelihood Estimation of 12D for Inoculated Packs

Edward W. Ross, Jr.
Office of the Technical Director
U.S. Army Natick R&D Command
Natick, MA 01760

Abstract. This paper describes a statistical procedure for estimating the 12D dose in the radiation-sterilization of canned food, using data from an inoculated pack experiment. The method assumes a two-parameter distribution, of which the shifted-exponential is taken as a prototype, and uses the maximum-likelihood principle to estimate the parameters and hence 12D. The procedure is embodied in a computer program which estimates 12D and provides confidence limits on both 12D and the kill at zero dose. The method is illustrated by an example.

1. Introduction. This paper is concerned with methods for assessing the effectiveness of ionizing radiation as a means of food-preservation. In particular, it deals with the problem of estimating the 12D dose, using the data obtained from an inoculated pack experiment. A number of papers have dealt wholly or partly with this question, including those of Anellis et al. (1968, 1969, 1975), Grecz et al. (1971) and Ross (1974, 1976). The general problem is one of determining a dose-response function and is discussed by Finney (1952).

The purpose of the present paper is to describe a method of data analysis, based on the maximum-likelihood (ML) criterion, for estimating 12D from inoculated pack data. An example will be presented, showing how the method works.

The ML method is a very widely-used procedure for deriving estimates of unknown parameters from experimental data and is described in most books on mathematical statistics, e.g., Hoel (1971). It seems not to have been applied in analyzing inoculated pack data, possibly because it leads to complicated formulas that can in practice only be solved with the aid of a high-speed computer. Despite this drawback the ML method is worth considering because it can extract more useful information than other procedures from the same data.

2. Theory. This section is divided into two parts, dealing with the inoculated pack experiment and a description of the ML method.

2a. The Inoculated Pack Experiment. The inoculated pack experiment consists of inoculating cans of the food substrate with a large number of the test-microorganisms. The cans are then vacuum-sealed, groups of them are exposed to different doses of radiation and then incubated. After incubation each can is examined to see whether it contains survivors. In the example described later, the test microorganisms were spores of ten strains of Clostridium botulinum, the incubation period was six months and the method of examination was the recovery of viable botulinal cells.

If we denote the different groups of cans by index, i, $i = 1,2,\ldots, M$, we define

x_i = dose which the i-th group received.

n_i = number of organisms per can in the i-th group.

N_i = number of cans in the i-th group.

K_i = number of sterilized cans (i.e., cans without survivors in the i-th group.

Usually the experiment is designed so that all n_i are approximately equal, and all N_i are the same. This simplifies the experiment and analysis; but there are advantages to be gained by varying n_i and N_i. In any case, the procedure described here applies to the situation where n_i and N_i may all be different.

The data consist of x_i, n_i, N_i and K_i for $i = 1,2,\ldots M$, where x_i and n_i are non-negative numbers and N_i and K_i are negative integers. The data-analysis must deduce an estimate of the 12D dose from this data.

2b. The ML Method. The method described here is based on the general probability theory for inoculated packs, see Ross (1974).

It is assumed that under the test conditions, the probability that an individual organism will be killed at dose x is given by the distribution function G(x), the survival probability being 1-G(x). The 12D-dose, which we denote x_c, satisfies

$$1-G(x_c) = 1 \times 10^{-12} \quad (1)$$

The probability that a can containing n organisms will be sterilized (i.e., all organisms will be killed) at dose x is denoted by $\Phi(x)$. $\Phi(x)$ and G(x) are related by means of

$$\Phi(x) = [G(x)]^n \quad (2)$$

or, approximately for n large and 1-G small,

$$\Phi(x) = e^{-n[1-G(x)]} \quad (3)$$

$$G(x) = 1 + n^{-1} \ln \Phi(x) \quad (4)$$

In the inoculated pack test at dose x_i, N_i cans are exposed each having $\Phi(x_i) = \Phi_i$ as the probability of sterilization. The probability that K_i cans are sterilized at dose x_i is given by the binomial distribution

$$P_i = \binom{N_i}{K_i} \Phi_i^{K_i} \left(1 - \Phi_i\right)^{N_i - K_i} \quad (5)$$

So far, nothing has been assumed about the form of the function G(x). We now assume that G(x) has a general form, $G(x; B_1, B_2)$, by which is meant that G depends not only on x (i.e., dose) but on two other quantities, B_1 and B_2, which are independent of dose. For example, assuming a shifted-exponential distribution for G(x) means

$$G(x) = 1 - \exp\{-B_1(x-B_2)\}$$

If B_1 and B_2 were known, we could immediately estimate 12D by solving Equation (1), which becomes in this case

$$1 - 1 \times 10^{-12} = 1 - \exp\{-B_1 (x_c - B_2)\}$$

or

$$x_c = B_2 + (12 \ln 10)/B_1 = B_2 + 27.63/B_1$$

Usually we do not know B_1 and B_2, and our problem is then to estimate them from the inoculated pack data.

The ML method tells us to do this by choosing B_1 and B_2 so that the probability (likelihood) of getting the observed experimental results is as large as possible. If B_1 and B_2 are given, then Φ_i is known and P_i is the likelihood of getting the observed outcome at dose x_i. Since the cans at different doses are tested independently, the joint likelihood of getting the observed experimental results for all the M doses is

$$P = P_1 P_2 \ldots P_M$$

The ML procedure directs us to find B_1 and B_2 so that P will be maximized. This is equivalent to maximizing

$$\Gamma = \ln P = \sum_{i=1}^{M} \{\ln C_i + K_i \ln \Phi_i + (N_i - K_i) \ln(1-\Phi_i)\} \quad (6)$$

where

$$C_i = \binom{N_i}{K_i} = \frac{N_i!}{K_i!(N_i-K_i)!}$$

$$\Phi_i = \exp\{-n_i (1-G_i)\}.$$

The usual procedure for finding B_1 and B_2 is to solve the equations

$$\partial\Gamma/\partial B_1 = 0$$

$$\partial\Gamma/\partial B_2 = 0.$$

The plausible forms for G(x) all lead to equations which are too complicated to solve by simple formulas. Usually one uses, instead, a successive approximation scheme, like the one written in matrix form as

$$\beta = B - (\Gamma'')^{-1}\Gamma' \tag{7}$$

where

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \quad \Gamma' = \begin{bmatrix} \partial\Gamma/\partial B_1 \\ \partial\Gamma/\partial B_2 \end{bmatrix}$$

$$\Gamma'' = \begin{bmatrix} \partial^2\Gamma/\partial B_1^2 & \partial^2\Gamma/\partial B_1 \partial B_2 \\ \partial^2\Gamma/\partial B_2 \partial B_1 & \partial^2\Gamma/\partial B_2^2 \end{bmatrix}$$

In using this, an initial guess is made for B_1 and B_2. The terms on the right of Equation (7) are evaluated for those values of B_1 and B_2 and the quantities β_1 and β_2 are calculated using Equation (7). These are then taken as the new values of B_1 and B_2 and the process is repeated. This continues until the β's and B's are equal to some desired accuracy.

The properties of this Gauss-Newton iteration scheme are reasonably well-known. It converges if the initial guess is good enough, and the Hessian matrix, Γ'', is positive-definite. If it converges, the inverse of the Hessian Matrix, $(\Gamma'')^{-1}$, gives the estimated variance-covariance matrix of B_1 and B_2. However, the method may occasionally fail to converge.

Given any assumed form for G(x), one can write explicit formulas for the quantities $\partial\Gamma/\partial B_1$, $\partial\Gamma/\partial B_2$, $\partial^2\Gamma/\partial B_1^2$, $\partial^2\Gamma/\partial B_1\partial B_2$ etc. as functions of B_1 and B_2. These formulas are necessary, but they are complicated and not especially informative, so we omit them.

The calculations involved in carrying out the ML method are obviously very tedious. However, the author has prepared a FORTRAN computer program which does the calculation of B_1 and B_2, then finds the 12D dose, x_c. The program also finds confidence limits on x_c and the logarithm of the survival probability at x=0. It does all of these calculations for each of the following five general forms of G(x):

$$G(x) = 1 - \exp\{B_1(x-B_2)\}$$ shifted exponential

$$G(x) = 1 - \exp\{-(B_2x)^{B_1}\}$$ Weibull

$$G(x) = F_g\{B_1(x-B_2)\}$$ normal

$$G(x) = F_g\{B_1 \ln(x/B_2)\}$$ lognormal

$$G(x) = 1 - \exp(-B_1x)$$ unshifted exponential

where

$$F_g(y) = \int_{-\infty}^{y} (2\pi)^{-1/2} e^{-t^2/2} dt$$

The program receives the inoculated pack data as input, including doses where all or none of the cans are sterilized as well as partial spoilage doses. It first carries out least-squares fitting of the data from partial-spoilage doses only, obtaining in this way initial estimates of B_1 and B_2 for all five forms of $G(x)$. These are used to start the ML method. Having found the optimizing (i.e., maximizing) values of B_1 and B_2, the program also finds for each form the quantity

$$\chi^2 = \sum_{1}^{M} \frac{(K_i - N_i\phi_i)^2}{N_i\phi_i(1 - \phi_i)}$$

This is distributed approximately as a χ^2-random variable with M-2 degrees of freedom and is an overall measure of how well that form can be made to fit the data.

3. Example. In this Section we describe an inoculated pack that was recently carried out at the U. S. Army Natick R&D Command and shows the results of using the proposed method of data analysis.

An inoculated pack was done at -30°C using C. botulinum spores in canned pork. The data (Anellis, unpublished, based on can-swelling) are shown in Table 1. The results from the preliminary least-squares (LS) fitting and the final, maximum likelihood (ML), estimates are shown in Table 2. Figure 1 is a graph of the data points and the four distributions fitted by the ML method.

Examining the ML results, we see that the normal distribution predicts that the entire 95% confidence interval of Z lies below -1, i.e. there is more than 90% kill at zero dose with 95% confidence. We therefore discard this distribution. For the exponential distribution ML predicts a very small shift (Z = -.23) whose 95% confidence limits straddle zero. There is, therefore, no reason to conclude that the shift is non-zero, which means that in this case the simple-exponential hypothesis is acceptable (i.e., it is not contradicted by the data). Similarly the Weibull shape parameter is very close to 1.0, B_1 = .9733, which also supports this hypothesis. These two distributions give almost the same 12-D dose, x_c = 3.89, and the 12D-dose of a simple-exponential is x_c = 3.83. The Schmidt-Nank formula yields x_c = 3.76. The lognormal leads to the estimate x_c = 4.11.

The theoretical value of χ^2 is $\chi^2_{M-2}(.95) = \chi^2_4(.95) = 9.49$, which exceeds the computed χ^2 for all four distributions, so we have no evidence against any of the four distributions on grounds of goodness-of-fit.

In this case we can adopt a procedurally conservative viewpoint and reason as follows. In the past, the simple exponential has always been used. The data does not refute its use here, so we may conclude that the distribution is exponential, the best 12D estimate is 3.89, and the 95% confidence limits are $3.62 \leq x_c \leq 4.32$. An alternative algorithm is to suspend judgment on the distribution but use the largest 12D-value given by any acceptable distribution. This leads to use of the lognormal estimates, $\hat{x}_c$ = 4.11 and $3.73 \leq x_c \leq 4.77$.

Either of these two viewpoints can be taken in this case and the two 12-D values obtained are not statistically different at the 95% confidence level. Also in practical terms the difference between 3.9 and 4.1 megarads for 12D is not very important.

4. Discussion. The ML method has the following advantages:

(i) It is a generally accepted statistical procedure.

(ii) It is a very flexible method that can be used with many different assumed distribution functions.

(iii) Because it uses the data at points where $K_i = 0$ or N_i, it comes closer than existing methods to using all the information that is in the data.

It has two drawbacks, namely, it is complicated and may occasionally fail to converge. The former is not a problem since a computer program already exists for it, and the latter happens very rarely in the writer's experience.

On balance, it appears that the ML method is promising and deserves further study.

REFERENCES

Anellis, A., and Werkowski, S. 1968. Estimation or radiation resistance values of microorganisms in food products. Appl. Microbiol. 16: 1300.

Anellis, A., Berkowitz, D., Jarboe, C., and El-Bisi, H.M. 1969. Radiation sterilization of prototype military foods. III. Pork loin. Appl. Microbiol. 18:604.

Anellis, A., Shattuck, E., Rowley, D.B., Ross, E.W., Whaley, D.N., and Dowell, V.R. 1975. Low-temperature irradiation of beef and methods for evaluation of a radappertization process. Appl. Microbiol. 30: 811.

Finney, D. J. 1952. "Probit Analysis", p. 8. Cambridge University Press, Cambridge, England.

Grecz, N., Walker, A.A., Anellis, A., and Berkowitz, D. 1971. Effect of irradiation temperature in the range -196 to 95C on the resistance of spores of Clostridium botulinum 33A in cooked beef. Can. J. Microbiol. 17:135.

Hoel, P. G. 1971. "Introduction to mathematical statistics", p. 196, John Wiley & Sons, New York.

Ross, E. W. 1974. Statistical estimation of 12D for radappertized foods. J. Food Sci. 1974:800.

Ross, E. W. 1976. On the statistical analysis of inoculated packs. J. Food Sci. 1976:578.

x_i	n_i	N_i	K_i
2.0	1.83×10^7	100	0
2.2	"	100	11
2.4	"	100	51
2.6	"	100	85
2.8	"	100	98
3.0	"	100	100

Table 1: Swelling data from inoculated pack for C. botulinum in irradiated canned pork.

	Exp		Weib		Norm.		Logn.	
	ML	LS	ML	LS	ML	LS	ML	LS
$\bar{x}$	2.5	-	2.5	-	2.5	-	2.5	-
Ψ	18.0	-	18.0	-	5.0	-	5.0	-
B_1	6.963	7.751	.9733	1.059	1.247	1.351	3.015	3.345
B_2	-.077	.171	7.771	6.151	-1.932	-1.601	2.507	2.104
95% x_{c1}	3.616	-	3.559	-	3.469	-	3.730	-
12D = x_c	3.891	3.736	3.895	3.738	3.710	3.606	4.113	3.897
95% x_{cu}	4.315	-	4.473	-	4.079	-	4.773	
95% Z_L	1.743	-	-	-	-2.912	-	-	-
Z	-.232	.576	-	-	-2.097	-1.815	-	-
95% Z_u	1.278	-	-	-	-1.433	-	-	-
Γ_2	-10.07	-11.11	-10.09	-10.73	-9.59	-10.17	-10.58	-11.59
χ^2	2.93	6.03	3.00	5.03	2.20	3.94	3.75	6.93

Table 2: Maximum likelihood (ML) and Least Squares (LS) estimates for four distribution-forms, based on data in Table 1. $Z = \log_{10}$ (survival probability at zero dose). LS simple exponential 12D = 3.829, Schmidt-Nank 12D = 3.755. The LS results are based on the data $2.2 \leq x \leq 2.8$, the ML results in $2.0 \leq x \leq 3.0$.

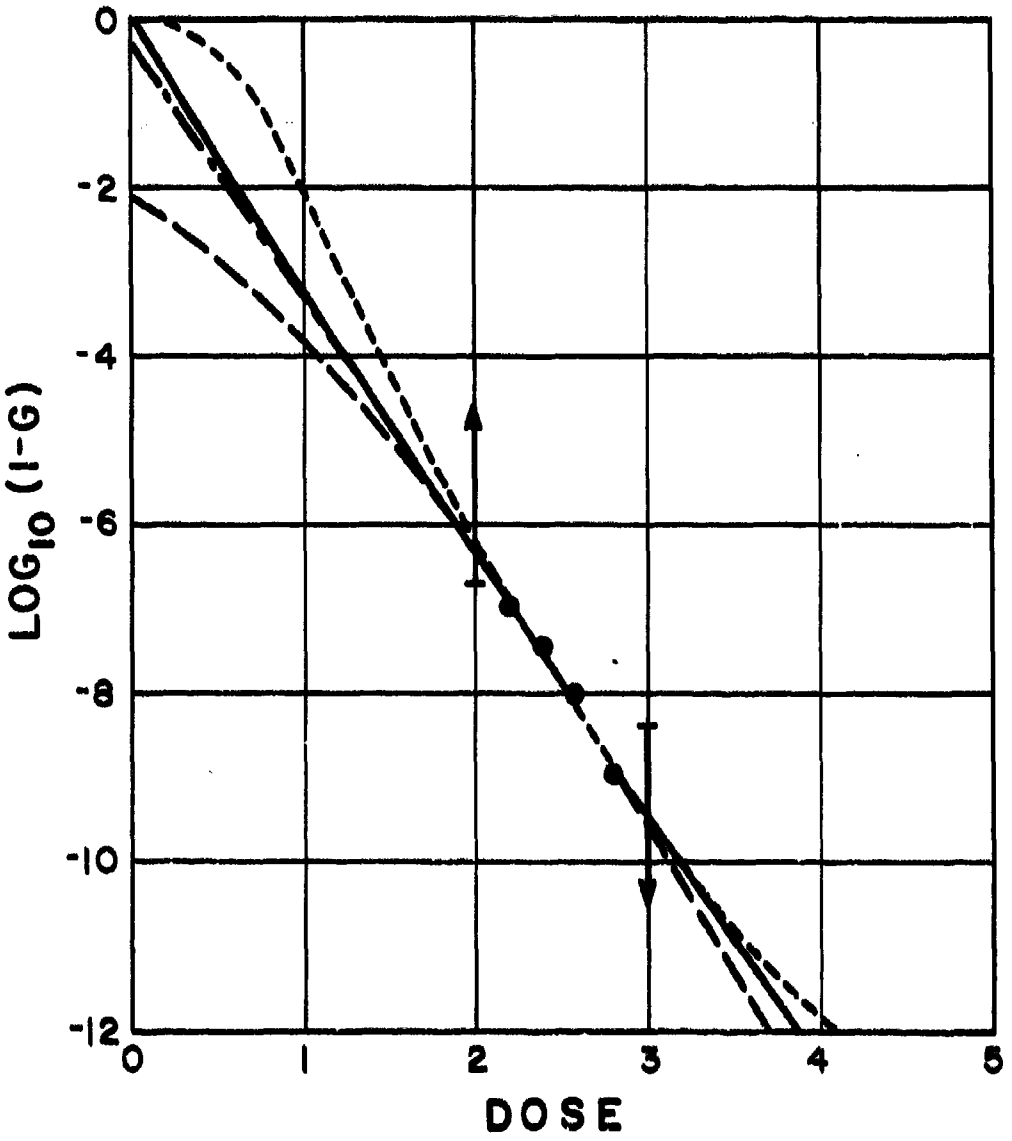


Figure 1: Graph of $\log_{10}$ (survival probability) as a function of dose in megarads for irradiated canned pork, based on viable cells. Data points are shown as •, and the four distributions fitted by ML are shown as lines: ——— Weibull, ——— - ——— exponential, — — — normal, ——— lognormal. ⊥ and ⊤ are the 95% confidence ranges for theoretical probabilities at doses where no cans or all cans are sterilized.

CONFIDENCE BOUNDS FOR THE GENERAL LINEAR MODEL

J. Richard Moore
Malcolm S. Taylor

USA Ballistic Research Laboratories
Aberdeen Proving Ground, Maryland 21005

In this paper, for the general linear model $Y = X\beta + e$, we consider the construction of confidence bounds about the entire regression line. To accomplish this we exploit a powerful theorem of Scheffé. A procedure often encountered is one in which a set of confidence intervals about $E(y|x)$ or prediction intervals for future observations are determined and then the end points are connected in such a fashion as to describe an envelope. The belief is that what has been accomplished is precisely what Scheffé's theorem allows one to do.

In addition, we present some extensions concerning confidence bounds about combinations of regression lines and suggest a useful application of these results. Specifically, we propose to use the confidence bounds about the difference of regression lines to make a quantitative assessment of when and where independent sets of data characterizing the same phenomena are in agreement or disagreement.

1. INTRODUCTION

It is appropriate at the onset that we devote a few paragraphs to the introduction of the general linear hypothesis model of full rank. We want to consider uncorrelated observations $y_1, y_2, \ldots, y_n$ that satisfy the relation

$$y_i = \sum_{j=1}^{p} x_{ij}\beta_j + e_i, \quad i = 1, 2, \ldots, n \qquad (1.1)$$

and are linear in the unknown parameters $\beta_1, \beta_2, \ldots, \beta_p$ with known coefficients x_{ij} and random term e_i satisfying

$$E(y_i) = \sum_{j=1}^{p} x_{ij}\beta_j,$$

and

$$\mathrm{Var}(y_i) = \sigma^2.$$

In other words, the random term e_i is a random variable with expected value $E(e_i)$ equal to zero and unknown variance $\mathrm{Var}(e_i)$ equal to σ^2. The problem, in its most general sense, involves determining point and interval estimates of several quantities of interest of the model and the testing of various statistical hypotheses.

For compactness of notation and ease of manipulation let

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & & & \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \quad e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix};$$

then we can write the system of relations (1.1) as

$$Y = X\beta + e$$

and proceed to define the general linear hypothesis model of full rank as follows:

Definition 1.1. The model $Y = X\beta + e$ where Y is a random observed vector, e is a random vector, X is an n x p matrix of known fixed quantities, and β is a vector of unknown parameters is called the general linear hypothesis model of full rank, provided the rank of X is equal to p where $p \leq n$.

In the present inquiry we restrict our consideration to the normal theory case, which means the random vector e, already satisfying $E(e) = 0$ and $\text{cov}(e) = \sigma_1^2$, will, in addition, be assumed to be normally distributed.

The problem most frequently addressed is that of estimating the unknown parameters β_j on the basis of the observations y_i. These estimates of β_j, denoted by $\hat{\beta}_j$, are functions of y_i; and, as such, are themselves random variables about which confidence intervals can be constructed. These ideas are fully developed in a number of textbooks.[1,2] A point not so widely expounded is that the usual frequency interpretation of a confidence interval based on a single sample $y_1, y_2, \ldots, y_n$ holds only for a single coefficient β_j; if the same data are used to determine confidence intervals for both β_i and β_j, $i \neq j$, the probability is not 1-α that the confidence intervals thus constructed will simultaneously contain β_i and β_j. The complexity is advanced by the fact that the interval estimates are not independent; so, in general, only a single confidence statement can be made from a single set of observations.

It is not our intent here to address this problem directly; such an inquiry falls into the general area of simultaneous confidence intervals. It is our intent, however, to consider a ramification of this problem: namely, the construction of a confidence envelope about the entire regression line. We will, in addition, provide some results concerning confidence envelopes about combinations of regression lines and implications of their use.

Toward this end consider the following definition due to Bose:[3]

---

1 Graybill, F. A., An Introduction to Linear Statistical Models, Volume I, McGraw-Hill Book Company, Inc., New York, 1961.

2 Rao, C. R., Linear Statistical Inference and Its Applications, John Wiley & Sons, Inc., New York, 1965.

3 Bose, R. C., "The Fundamental Theorem of Linear Estimation", Proceedings of the 31st Indian Science Congress, 1944, pp. 2-3.

Definition 1.2. A parametric function ψ is called an estimable function if it has an unbiased linear estimate, i.e., if there exists an n-vector a of constant coefficients such that $E(a'y) = \psi$.

If L is a p-dimensional space of estimable functions with basis $\{\psi_1, \psi_2, \ldots, \psi_p\}$ and $\hat{\psi}$ is the least squares estimate of $\psi \in L$, then we have the following theorem due to Scheffé[4].

Theorem 1.1. Under the general linear hypothesis model (normal case) the probability is $1 - \alpha$ that simultaneously for all $\psi \in L$

$$\hat{\psi} - S\hat{\sigma}_{\hat{\psi}} \leq \psi \leq \hat{\psi} + S\hat{\sigma}_{\hat{\psi}}$$

where the constant $S = \{pF_p(p,n-r)\}^{1/2}$ and rank $X = r$.

The implications of this theorem are far reaching; and in this article we will exploit a single facet, albeit an important and useful one. To facilitate this we need to be aware of the fact that since least squares estimates $\hat{\beta}$ are BLUE, the elements of the vector β of the general linear model of full rank form a basis of a space L of estimable functions which includes polynomials as a special case.

2. CONFIDENCE REGION FOR A POLYNOMIAL

To determine a confidence region for a polynomial with observational equations

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \ldots + \beta_{p-1} x_i^{p-1} + e_i, \quad i = 1, 2, \ldots, n$$

in the model $Y = X\beta + e$, the n x p matrix $X = (x_{ij})$ of known constant coefficients takes the form

---

4 Scheffé, H., The Analysis of Variance, John Wiley & Sons, Inc., New York, 1959.

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{p-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{p-1} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 1 & x_n & x_n^2 & \dots & x_n^{p-1} \end{pmatrix}$$

The least squares estimate of β is given by $\hat{\beta} = (X'X)^{-1}X'Y$. If we choose $\psi_i = \beta_i$, $i = 0, 1, \dots, p-1$, then $\{\psi_i\} = \{\beta_i\}$ is a set of p linearly independent estimable functions which forms a basis for the space L. For any value x_o denote $X_o' = (1, x_o, \dots, x_o^{p-1})$. Clearly, $E(y_o) = X_o'\beta \in L$ and has least squares estimate

$$X_o'\hat{\beta} = X_o'(X'X)^{-1}X'Y = \sum_{i=1}^{n} a_i y_i$$

where the coefficients a_i are the elements of the 1 x n vector $X_o'(X'X)^{-1}X'$. Thus

$$\sum_{i=1}^{n} a_i^2 = X_o'(X'X)^{-1}X'[X_o'(X'X)^{-1}X']'$$

$$= X_o'(X'X)^{-1}X'X(X'X)^{-1}X_o$$

$$= X_o'(X'X)^{-1}X_o$$

so that $\sigma^2_{\hat{\psi}} = \sigma^2 X_o'(X'X)^{-1}X_o$ with unbiased estimate $s^2 X_o'(X'X)^{-1}X_o$.

From Scheffé's theorem we can assert with probability $1-\alpha$ that simultaneously for all $\psi \in L$ and, in particular, $X_o'\beta \in L$

$$X_o'\hat{\beta} - S\hat{\sigma}_{\hat{\psi}} \leq X_o'\beta \leq X_o'\hat{\beta} + S\hat{\sigma}_{\hat{\psi}}$$

where $S = [pF_\alpha(p,n-p)]^{1/2}$.

As an illustration, suppose the paired data (1.20, 0.34), (1.37, 0.94), (1.38, 0.99), (1.65, 1.58), (1.71, 2.08), (1.82, 2.25) are characterized by the quadratic $y = -0.31x^2 + 3.97x - 3.95$ over the interval of interest, $1 \leq x \leq 2$. The 95% confidence region for the entire true line is given by

$$X_o'\hat{\beta} - (5.28)\hat{\sigma}_{\hat{\psi}} \leq X_o'\beta \leq X_o'\hat{\beta} + (5.28)\hat{\sigma}_{\hat{\psi}}$$

as shown in Figure 1.

Grubbs[5] showed that for the case $y = \beta_o + \beta_1 x$ the confidence bounds resulting from Scheffé's theorem are

$$\hat{\beta}_o + \hat{\beta}_1 x \pm [2F_\alpha(2,n-2)]^{1/2} S \left(\frac{1}{n} + \frac{n(x-\bar{x})^2}{A_{xx}} \right)^{1/2} \qquad (2.1)$$

where $S = \left(\frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{\beta}_o - \hat{\beta}_i x_i)^2 \right)^{1/2}$ and $A_{xx} = n\Sigma x_i^2 - [\Sigma x_i]^2$.

Note that the value x appearing in (2.1) is not limited to an x_i which appears in the observations (x_i, y_i), $i = 1, 2, \ldots, n$.

3. THE TWO-SAMPLE CASE

Suppose two independent sets of data have given rise to two characterizations of the same phenomenon so that we are now confronted with what is, in essence, two models:

$$Y_1 = X_1\beta_1 + e_1 \text{ , an } n_1 \times p_1 \text{ problem,}$$

and

$$Y_2 = X_2\beta_2 + e_2 \text{ , an } n_2 \times p_2 \text{ problem.}$$

We can still represent this situation as $Y = X\beta + e$ where now

---

5 Grubbs, F. E., Linear Statistical Regression and Functional Relations, BRL Report No. 1842, November 1975.

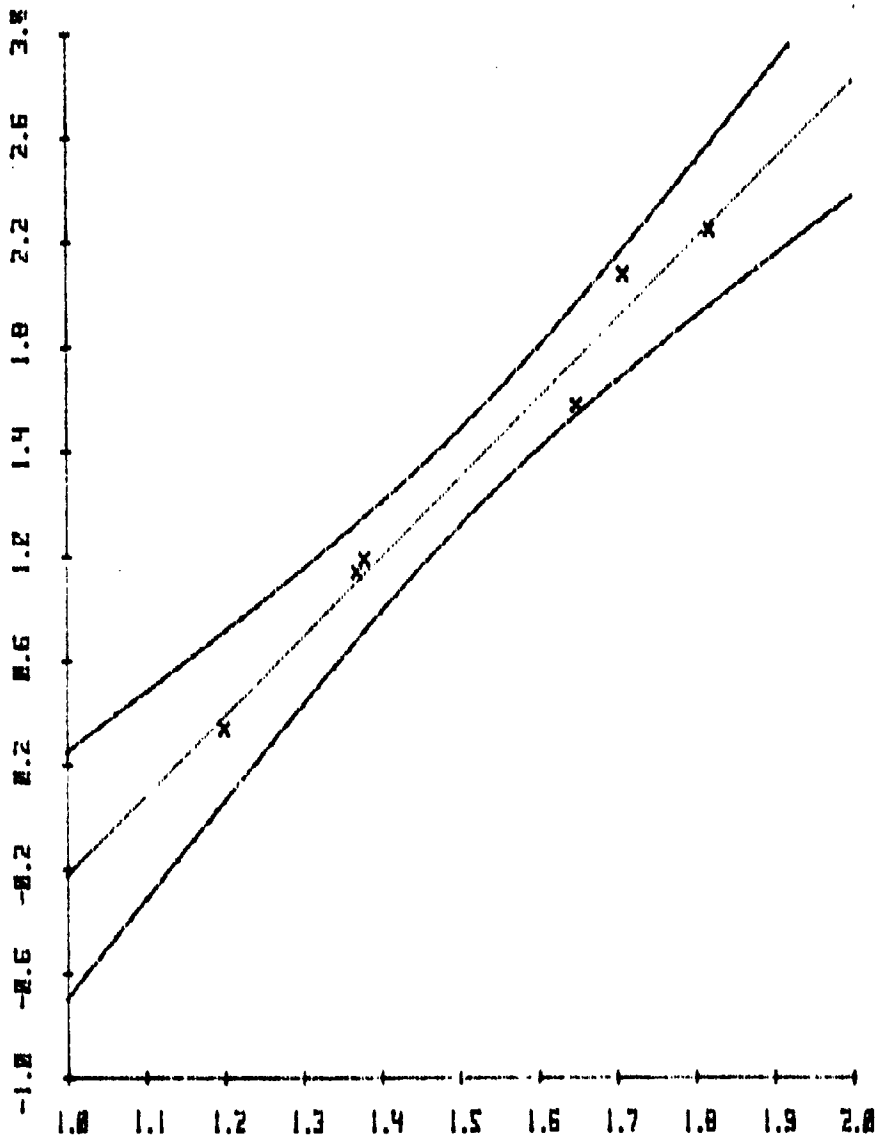


FIGURE 1

95 % CONFIDENCE LIMITS FOR A LINEAR FUNCTION

$$Y = \begin{pmatrix} Y_1 \\ \hline Y_2 \end{pmatrix} \quad , \quad X = \left(\begin{array}{c|c} X_1 & 0 \\ \hline 0 & X_2 \end{array}\right) , \quad \beta = \begin{pmatrix} \beta_1 \\ \hline \beta_2 \end{pmatrix} \quad , \quad e = \begin{pmatrix} e_1 \\ \hline e_2 \end{pmatrix} .$$

The least squares estimate is given, as before, by $\hat{\beta} = (X'X)^{-1}X'Y$.

Consider now the difference of two polynomials $y_1^* - y_2^* \in L$ with LS estimate $X_1^{*'}\hat{\beta}_1 - X_2^{*'}\hat{\beta}_2$ where $X_1^{*'} = (1, x_i, \ldots, x_i^{p_i-1})$. Rewriting,

$$X_1^{*'}\hat{\beta}_1 - X_2^{*'}\hat{\beta}_2 = \left(X_1^{*'} \mid - X_2^{*'}\right) \begin{pmatrix} \hat{\beta}_1 \\ \hline \hat{\beta}_2 \end{pmatrix} = \left(X_1^{*'} \mid - X_2^{*'}\right) (X'X)^{-1}X'Y$$

$$= \sum_{i=1}^{n_1+n_2} a_i y_i$$

where the coefficients a_i are the elements of the $1 \times (n_1+n_2)$ vector $X^{*'}(X'X)^{-1}X'$. Thus,

$$\Sigma a_i^2 = X^{*'}(X'X)^{-1}X'[X^{*'}(X'X)^{-1}X']' = X^{*'}(X'X)^{-1}X'X(X'X)^{-1}X^* = X^{*'}(X'X)^{-1}X^*.$$

Now $X^{*'} = \left(X_1^{*'} \mid - X_2^{*'}\right)$ and $(X'X)^{-1} = \left(\begin{array}{c|c} (X_1'X_1)^{-1} & 0 \\ \hline 0 & (X_2'X_2)^{-1} \end{array}\right)$

so $X^{*'}(X'X)^{-1}X^* = X_1^{*'}(X_1'X_1)^{-1}X_1^* + X_2^{*'}(X_2'X_2)^{-1}X_2^*$.

As before, $\mathrm{Var}(\hat{\psi})$ has the unbiased estimate $\hat{\sigma}^2_{\hat{\psi}} = s^2X^{*'}(X'X)^{-1}X^*$, where s^2 is now the pooled estimate of the variance, and with probability $1 - \alpha$ simultaneously for all $\psi \in L$

$$X_1^{*\prime}\hat{\beta}_1 - X_2^{*\prime}\hat{\beta}_2 - S\hat{\sigma}_{\hat{\psi}} \leq y_1^* - y_2^* \leq X_1^{*\prime}\hat{\beta}_1 - X_2^{*\prime}\hat{\beta}_2 + S\hat{\sigma}_{\hat{\psi}}$$

with $S = \left[(p_1 + p_2)F_\alpha[p_1 + p_2, n_1 + n_2 - (p_1 + p_2)]\right]^{1/2}$.

To illustrate one of the most useful potential applications of this result consider the situation where we are presented with two sets of data, $\{(x_{11},y_{11}), (x_{12},y_{12}),\ldots,(x_{1n},y_{1n})\}$, $\{(x_{21},y_{21}), (x_{22},y_{22}),\ldots,(x_{2m},y_{2m})\}$, collected from the same process; and we want to say something about the similarity or dissimilarity of the two descriptions. Suppose each set is fitted with a quadratic; and we construct the confidence bound about the difference $y_1^* - y_2^*$, as shown in Figure 2. Over the region $(1.65 \leq x \leq 2.05)$, where the confidence bounds cover the line $y = 0$, we will say the two descriptions are consistent, although the associated probability level cannot be attached without qualification and interpretation.

The extension of Grubbs result (2.1) to this case is direct; the bounds take the form

$$X_1^{*\prime}\hat{\beta}_1 - X_2^{*\prime}\hat{\beta}_2 \pm \left[(p_1 + p_2)F_\alpha(p_1 + p_2, n_1 + n_2 - p_1 - p_2)\right]^{1/2} \cdot S \cdot \cdot \left[\frac{1}{n_1} + \frac{1}{n_2} + \frac{n_1(X^* - \bar{X}_1)}{A_{xx}^1} + \frac{n_2(X^* - \bar{X}_2)}{A_{xx}^2}\right]^{1/2}$$

where S^2 is the pooled estimate of variance and A_{xx}^i is computed from the i-th data set.

4. THE k-SAMPLE CASE

The straightforward generalization to k sets of data proceeds as follows:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & X_k \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_k \end{bmatrix}$$

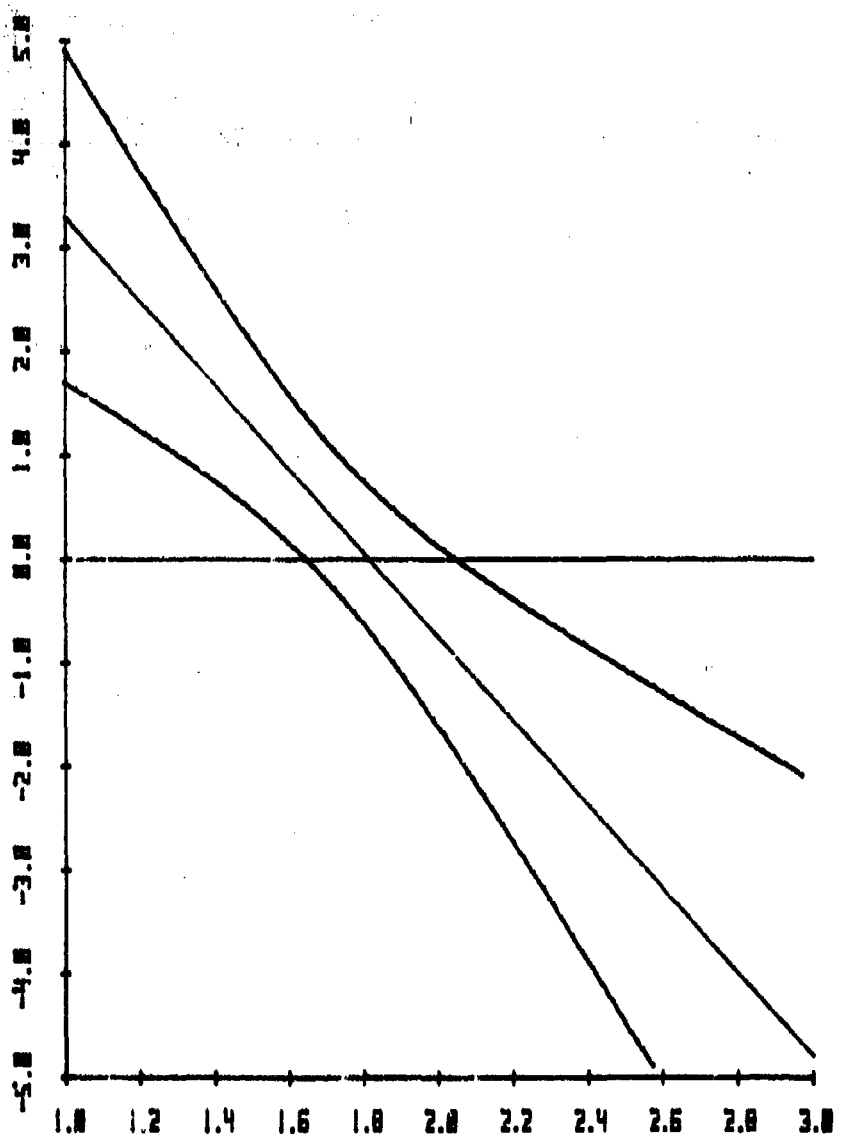


FIGURE 2

95 % CONFIDENCE LIMITS FOR THE DIFFERENCE IN TWO LINEAR FUNCTIONS

and

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hline \hat{\beta}_2 \\ \hline \vdots \\ \hline \hat{\beta}_k \end{pmatrix} = \begin{pmatrix} (X_1'X_1)^{-1}X_1'Y_1 \\ \hline (X_2'X_2)^{-1}X_2'Y_2 \\ \hline \vdots \\ \hline (X_k'X_k)^{-1}X_k'Y_k \end{pmatrix}.$$

Now for $X^{*'} = \left(X_1^{*'} \mid \ldots \mid X^{*'}_k\right)$ where $X_i^{*'} = \left(1, x, \ldots, x^{p_i-1}\right)$

we can write $\hat{y}_i^* = X_i^{*'}\hat{\beta}_i = X_i^{*'}(X_i'X_i)^{-1}X_i'Y_i$, $i = 1, 2, \ldots, k$.

Suppose we now consider the estimate of $\Sigma c_i y_i^*$:

$\Sigma c_i \hat{y}_i^* = \Sigma c_i X_i^{*'}\hat{\beta}_i = \Sigma c_i X_i^{*'}(X_i'X_i)^{-1}X_i'Y_i$. Rewriting,

$$\Sigma c_i X_i^{*'}\hat{\beta}_i = \left(c_1 X_1^{*'} \mid \ldots \mid c_k X_k^{*'}\right)(X'X)^{-1}X'Y = \sum_{i=1}^{n_1+\ldots+n_k} a_i y_i$$

where the coefficients a_i are the elements of the $1 \times \Sigma n_i$ vector $CX^{*'}(X'X)^{-1}X'$.

Thus,

$$\begin{aligned} \Sigma a_i^2 &= CX^{*'}(X'X)^{-1}X'[CX^{*'}(X'X)^{-1}X']' \\ &= CX^{*'}(X'X)^{-1}X'X(X'X)^{-1}CX^* \\ &= CX^{*'}(X'X)^{-1}CX^*. \end{aligned}$$

Now

$$CX^{*'} = \left[c_1X_1^{*'} \;\middle|\; \cdots \;\middle|\; c_kX_k^{*'}\right] \text{ and } (X'X)^{-1} = \begin{pmatrix} (X_1'X_1)^{-1} & & \\ & \ddots & \\ & & (X_k'X_k)^{-1} \end{pmatrix}$$

so $CX^{*'}(X'X)^{-1}CX^{*} = \sum_{i=1}^{k} c_i^2 X_i^{*'}(X_i'X_i)^{-1}X_i^{*}$. The confidence region now assumes the form

$$\Sigma c_i X_i^{*'}\hat{\beta}_i - S\left(sCX^{*'}(X'X)^{-1}CX^{*}\right)^{1/2} \leq \Sigma c_i y_i^{*} \leq \Sigma c_i X_i^{*'}\hat{\beta}_i + S\left(sCX^{*'}(X'X)^{-1}CX^{*}\right)^{1/2};$$

with $S = \left(\Sigma p_i \cdot F_{\alpha}(\Sigma p_i, \Sigma n_i - \Sigma p_i)\right)^{1/2}$.

For the linear case we obtain

$$CX^{*'}(X'X)^{-1}CX^{*} = \sum_{i=1}^{k} c_i^2 X_i^{*'}(X_i'X_i)^{-1}X_i^{*}$$

$$= \sum_{i=1}^{k} c_i^2\left(\frac{1}{n_i} + n_i(x_o - \bar{X}_i)^2\right),$$

and the two sample case (Section 3) is obtained by setting $c_1 = 1$ and $c_2 = -1$.

ATTENDEES

22nd CONFERENCE ON DESIGN OF EXPERIMENTS

20 Oct -- 22 Oct 1976

HDL Auditorium

NAME	ORGANIZATION
ANDERSEN, Gerald R.	DARCOM
ARP, David L.	Naval Weapons Center
ASCHER, Harold	Naval Weapons Laboratories
BARTYCZAK, Ronald J.	IRS
BASU, Prof. Asit P.	Department of Statistics
BATES, Carl B.	Concepts Analysis Agency
BECHHOFER, Robert E.	Cornell University
BERRY, 1LT Patrick L.	Edgewood Arsenal
BLANCHARD, Ben S.	College of Engineering
BOX, George	University of Wisconsin
BROOKS, James	AMSAA
BROWN, Mark	City College
BRUGGER, Richard M.	Rock Island
BURGE, J. Robert	Walter Reed
CARROLL, Dr. Robert M.	Army Research Institute
CHANDRA, Dr. Jagdish	Army Research Office
CHEEVER, Herbert L.	Human Engineering Lab
CLARK, Offie E.	Product Assurance Dir.
COHEN, A. Clifford, Jr.	University of Georgia
COHEN, Stanley A.	Internal Revenue Service
COPPOLA, Eugene E.	Watervliet Arsenal
CRIGLER, John R.	Naval Center

NAME	ORGANIZATION
CROARKIN, Mary C.	National Bureau of Standards
CULPEPPER, Gideon A.	White Sands Missile Range
DAVID, John	Harry Diamond Laboratories
DAY, Donna N.	Internal Revenue Service
DeARMON, I. A.	Frederick Cancer Res. Center
DEFRIES, Dennis L.	MILPERCEN
DELOACH, Jimmie	US Infantry Center
DEPRIEST, Douglas	ONR
DINGER, Donald B.	Ft. Belvoir
DOSTAL, Robert G.	Rock Island
DRESSEL, Dr. Francis G.	Army Research Office
EHRLICH, Karl	NSWC
ENGELMAN, George F.	Naval Sea Systems Command
EPSTEIN, Kenneth	Army Research Institute
ESSENWANGER, Dr. Oskar M.	Redstone Arsenal
FOSTER, Walter D.	AFIP
FRISHMAN, Fred	Internal Revenue Service
GARDNER, Dr. Leonard B.	Rock Island
GARRETT, Wilbert A.	Aberdeen Proving Ground
GEISSER, Seymour	University of Minnesota
GELFAND, Alan E.	Dept. of Statistics
GRUBBS, Dr. Frank E.	Ballistic Research Labs
HAAS, Barbara R.	Internal Revenue Service
HACSKAYLO, Michael	Night Vision Laboratory
HARDIN, Donald E.	Ft. Belvoir
HARRIS, Bernard	University of Wisconsin

NAME	ORGANIZATION
HERBACH, Dr.	(?)
HIGBEE, Glenn	Ft. Detrick
HIGHFILL, Jerry W.	Ft. Detrick
HIRES, Joyce A.	Aberdeen Proving Ground
HIRSHFELD, Dr. Stephen F.	Research Institute
HODGE, John W., Jr.	Ft. Detrick
HOFFMAN, Joe	ARMCOM
HOGG, Robert V.	University of Iowa
HOLCOMBE, Stephanie	Infantry Board
HORNER, Theodore	Theodore Horner Assoc.
HUGHES, Paul F.	System Planning Corp.
HUSTAD, Jeffrey P.	MILPERCEN
IANNUZZELLI, Louis M.	Rock Island
JACKSON, William E.	DNA
JOLLENSTEN, Dr. Ralph W.	Department of Defense
KAPLAN, Eves	Walter Reed
KASHPORENKO, Daniel	MILPERCEN
KU, Harry H.	National Bureau of Standards
LARSEN, MAJ. L. E.	Walter Reed
LAUNER, Robert	Army Research Office
LEONARD, Randall E.	Ft. Belvoir
LEVIN, Sheldon G.	DNA
LUETZOW, Dr. H. Baussus von	Army Engr. Topogr. Labs
MAAR, Dr. James R.	Department of Defense
MADRO, Wm.	AMSAA
MALONEY, Clifford J.	Bureau of Biologics

NAME	ORGANIZATION
MATTILE, Craig A.	MILPERCEN
McAFEE, Dr. Walter S.	Army Electronics Command
McCARTHY, Thomas G.	Picatinny Arsenal
McDONALD, Bruce	NRL
McCLOSKEY, Joseph P.	Department of Defense
MELLINGER, Dr. J. J.	Army Research Institute
MOORE, J. Richard	Aberdeen Proving Ground
MORRIS, Carl	The Rand Corporation
MUMFORD, Stanford	Edgewood Arsenal
NATRELLA, Mary G.	National Bureau of Standards
NEAL, Donald M.	Army Materials & Mech. Res. Center
NOBLE, Dr. Ben	University of Wisconsin
PALMER, Charlotte	MILPERCEN
PASINI, Harold C. Jr.	Aberdeen Proving Ground
PECK, Carl C., MD	Letterman Inst. of Research
PEIRCE, L. Douglas	Army Logistics Man. Center
PERLMAN, Sanford	Internal Revenue Service
PETERSON, David	NSWC
POWERS, CPT Thomas	Ft. Detrick
POWERS, Thomas J.	Res. Inst. of Infectious Diseases
RACICOT, Ronald L.	Watervliet Arsenal
REALE, Ronald P.	Concepts Analysis Agency
RICHARDSON, Ralph M.	Army MILPERCEN
RIGGS, Charles W.	Frederick Cancer Res. Cen.
ROBINSON, Kenneth W.	MILPERCEN
ROSE, Carol D.	DARCOM

NAME	ORGANIZATION
ROSENBLATT, Joan	National Bureau of Standards
ROSS, Edward W.	NARADCOM
RUBIN, Herman	University of Wisconsin
SABALOS, Dionisios	Ft. Huachuca
SEIDL, George	NSWC
SECREST, Lawrence M.	MERADCOM
SELIG, Seymour	NRL
SHADID, Raymond J.	IRS
SHEPERD, William L.	White Sands Missile Range
SHIMI, Dr. I. N.	Bolling AFB
SHOENFELD, Martin	INTEREX
SHORT, P. J.	ARMTE-TE-PZ
SINGPURWALLA, Nozer D.	George Washington University
SMITH, Larry A.	MILPERCEN
SMUDA, Daniel	MILPERCEN
SOLOMON, Herbert	Stanford University
SPINDLER, Dr. George B.	Canadian Embassy
SROKA, Dr. Stephen A.	Department of Defense
STEINHEISER, Frederick	Army Research Institute
STOKES, Nathaniel	MERADCOM Lab 4000
TANG, Douglas B.	Walter Reed
TAUB, Audrey E.	Naval Surface Weapons Center
TAYLOR, Malcom S.	Aberdeen Proving Ground
TAYLOR, Vincent K.	DARCOM
THOMAS, Jerry	Concept Analysis
THOMPSON, W. A., Jr.	Statistics

NAME	ORGANIZATION
THRASHER, Paul H.	White Sands Missile Range
TRAVERS, Gary D.	Department of Defense
TRYBUL, Dr. T. S.	DARCOM
TURNER, Robert H.	White Sands Missile Range
VISNAW, Vernon V.	Aberdeen Proving Ground
VITALE, Dr. Richard A.	Mathematics Research Center
WALPERT, Harold	Naval Surface Weapons Center
WEINTRAUB, Gertrude	Picatinny Arsenal
WHARTON, Dr. Robert	Thomas Jefferson University
WIDBY, James B.	Infantry Board
WILBURN, John Bart	Electronic Proving Ground
WILLARD, Daniel	Pentagon
WITHERS, Langhorne P.	OPN Test & Eval. Agency
WOLMAN, Agatha S.	Concepts Analysis Agency
WOLMAN, Dr. William W.	Highway Administration
WOMACK, Franklin	NSWC
WYKOFF, Norman L.	Jefferson Proving Ground
YOUNG, Patricia A.	Ft. Meade
ZARWYN, Dr. Berthold	Harry Diamond Laboratories

KIRSHNER, Joseph M. (Chairman on Local Arrangements) Harry Diamond Laboratories

*U.S.GOVERNMENT PRINTING OFFICE: 1977-742390/ REGION NO. 4

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ARO [redacted]-77-2
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): PROCEEDINGS OF THE [redacted] CONFERENCE ON THE DESIGN OF EXPERIMENTS IN ARMY RESEARCH, DEVELOPMENT, AND TESTING (22nd) Held at Harry Diamond Laboratories, Adelphi, Maryland on 20-22 October 1976.
5. TYPE OF REPORT & PERIOD COVERED: Interim Technical Report
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)
8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: (12) 358p.
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Army Mathematics Steering Committee on Behalf of the Chief of Research, Development and Acquisition
12. REPORT DATE: June [redacted] 1977
13. NUMBER OF PAGES: 344
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): U. S. Army Research Office, P. O. Box 12211, Research Triangle Park, NC 27709
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. The findings in this report are not to be construed as official Department of the Army position, unless so designated by other authorized documents.

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

18. SUPPLEMENTARY NOTES

This is a technical report resulting from the Twenty-second Conference on the Design of Experiments in Army Research, Development and Testing. It contains most of the papers presented at that meeting. These treat various Army statistical and design problems.

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

management of reliability
pharmacokinetic models
serum lipids
error-time analysis
reliability
confidence bounds
time series analysis
pairwise contrasts
stress-strength models
probability distributions
fatigue life
variance estimation
robust outlier detection
random number generator
eigenvectors analysis
Markov chains
continuous sampling plans
Stein's estimator
robust statistical procedures
small samples
estimation and prediction
exponential selection problems
maximum likelihood estimation
confidence bounds

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

Unclassified

SUPPLEMENTARY

INFORMATION

AD-A042746

Subject: Errata Sheet

TO: Recipients of Proceedings of the 22nd Conference on the Design of Experiments in Army Research, Development and Testing

The undersigned apoligizes for some errors in my paper "Induction on a Markov chain" appearing on pages 177-186 of the proceedings. Four pen and ink corrections will correct these errors:

a. The denominator in Equation (5) should be $1 - q_1^2 q_2$ rather than $1 - q_1^2 q_2^2$.

b. Equation (7) should be $P(S2) = P(S2)$ rather than $P(S1) = P(S2)$.

c. The numerator of the second term in Equation (11), following the summation sign, should be

$$\binom{k}{i} p_1^i q_1^{k-i} (k-i) \qquad \text{rather than}$$

$$\binom{k}{i} p_1^i q_1^{k-i} (k-1)$$

d. The phrase between Equation (27) and Equation (28) should be:

"If $k_o+1 \leq N/2$ the above generalizes to"

rather than

"If $k_o+1 = N/2$ the above generalizes to".

RICHARD M. BRUGGER
Mathematical Statistician
Quality Evaluation Division
Product Assurance Directorate
US Army Armament Materiel Readiness Command
Rock Island, IL 61201