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

Army Research Office
Durham, North Carolina

November 1974

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: ARO Report 74-1	2. GOVT ACCESSION NO.
	3. RECIPIENT'S CATALOG NUMBER: AD/A-002 564
4. TITLE (and Subtitle): PROCEEDINGS OF THE NINETEENTH CONFERENCE ON THE DESIGN OF EXPERIMENTS IN ARMY RESEARCH, DEVELOPMENT AND TESTING	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	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 and Development and Acquisition.	12. REPORT DATE: November 1974
	13. NUMBER OF PAGES: 597
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): Army Research Office, Box CM, Duke Station, Durham, North Carolina 27706 RDRD	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 an 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 Nineteenth Conference on the Design of Experiments in Army Research, Development and Testing. It contains most 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)

Baysian statistics
selection procedures
analysis of variance
vitamin A modeling
circular probable error
hit probability
the beta distribution
bimodality
prediction analysis
laser countermeasures
linear transformations
regression analysis
reliability
continuous sampling
system simulation
the Weibull distribution
multivariate analysis
sensitivity tests
life testing
tolerance techniques
normal order statistics
rank tests
jackknife techniques
multiple component systems

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

*U.S. GOVERNMENT PRINTING OFFICE: 1974-643-778/N.P.

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

U. S. Army Research Office

Report No. 74-1
November 1974

PROCEEDINGS OF THE NINETEENTH CONFERENCE ON THE DESIGN OF EXPERIMENTS

Sponsored by the Army Mathematics Steering Committee

HOSTS

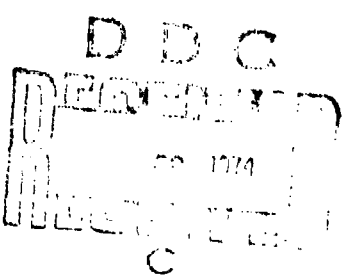
Headquarters U. S. Army Armament Command

and

U. S. Army Management Engineering Training Agency

24 - 26 October 1973

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



FOREWORD

The Nineteenth Conference on the Design of Experiments in Army Research, Development and Testing was held 24-26 October 1973 at Rock Island, Illinois. Like the first conference in this series, it had two hosts. They were the Headquarters U.S. Army Armament Command and the U.S. Army Management Engineering Training Agency. The last-named host furnished the excellent conference rooms for the meeting. In the planning phases of the meeting Mr. Raymond B. Loecke was their representative. Dr. Norman Coleman, a scientist at the U.S. Army Armament Command, served as the Local Chairman. Those in attendance at the conference are indebted to him and his assistants for issuing the host invitational letter, arranging for local accommodations and handling the many details needed for a successful conference. Dr. Coleman also helped in the initial plans for this meeting by bringing to the first meeting of the Program Committee a list of statistical areas that would be of interest to the scientific personnel of the hosts.

There were six addresses by invited speakers. The first of these was by Professor Jerome Cornfield who talked on "Bayesian Statistics." He first pointed out that the Bayesian approach uses prior information as well as objective information to avoid incoherence. He then gave several examples illustrating the inconsistent results that follow from attempting to judge statistical procedures by using only average properties over repeated samples. The second invited speaker was Professor Shanti S. Gupta. "Selection and Ranking Procedures for Multivariate Normal Populations" was the title of his address. After discussing such procedures, he surveyed known results and mentioned some unsolved problems in this area. In his address on "Generalized Jackknife Techniques" Professor H. L. Gray first carefully defined the jackknife statistic and several of its generalizations. He indicated how one can use these statistics as point and interval estimators for data from a random sample or a stochastic process with a continuous index. In his talk on "Reliability Growth" Professor Frank Proschan pointed out that in the development of a complex system it is useful to examine the importance of each component and to determine the optimum effort to allocate to each component to achieve a desired system reliability growth. One way of attacking this problem is through use of the Birnbaum measure of component importance. Proschan showed how this measure could be used to determine system reliability growth from component reliability growth. "Some Critical Remarks on Accelerated Life Testing" was the title of the address by Professor Sam C. Saunders. He stated that one of the current needs in reliability theory is for methods that permit the prediction of the life of a system from a few tests in which the process of wear has been accelerated. This paper made a strong plea for the statistician as well as the engineer to have a thorough understanding of each appropriate cumulative damage process in terms of its chemical or physical behavior before an analysis is undertaken. The invited speaker who spoke last was Professor William A. Thompson, Jr. In 1973 E. Brindley and he introduced a multivariate concept of monotone failure rate. Here the system fails when the first of its components fails so that not all component lifetimes are observable. In his address Dr. Thompson discussed mathematical models for problems arising in the biological and the engineering sciences. In the biological context one refers to the theory of competing risks, while in the engineering applications the model represents a non-repairable series system.

It is interesting to note that all the invited speakers, except Professor Gray, have appeared on the programs of earlier conferences in this series. We are grateful to these gentlemen for their valuable help and for their willingness to give of their time to bring new and useful developments in the fields of statistics and the design of experiments to the attention of Army scientists.

There is a large amount of scientific work being conducted in the many Army installations. These conferences offer an opportunity for individuals in these laboratories to present to an interested audience the work they are pursuing. At this meeting twenty-six technical papers and four clinical papers appeard on the agenda. These presentations were well received and both the members of the audience, as well as the speakers benefitted from the contents of the papers as well as the suggestions and questions raised in the discussions after each talk.

Following the banquet, held in the evening of the first day of the conference the ninth Samuel S. Wilks Memorial Award of the American Statistical Association and the Department of the Army was presented to Professor H. O. Hartley for his many outstanding contributions to the developments and teaching of statistics. Dr. Hartley is Professor of Economics and Distinguished Professor of Statistics at the Texas A and M University, at College Station, Texas. He has served many times as an invited speaker and also as a panelist at past Conferences on the Design of Experiments. His deep insight into statistical theory enabled him to offer valuable suggestions to many speakers at these meetings. We are pleased to be able to print in these proceedings the remarks Dr. Hartley made following the presentation of the Wilks Award.

The Army Mathematics Steering Committee, on behalf of the Office of the Chief of Research and Development and Acquisition, sponsors these meetings in order to expose Army scientists and engineers to various areas and new developments in statistics, thus upgrading the competence of in-house personnel. Members of this Committee have asked that these proceedings be published and issued Army-wide as well as to other scientific communities.

Near the first of each year the Program Committee for these conferences is selected and meets in Washington, D. C. to outline the program and suggest possible speakers for the next meeting. I would like to express my personal thanks to the members of this year's committee: A. Clifford Cohen, Norman P. Coleman, Francis Dressel, A. Ross Eckler, Walter Foster, Fred Frishman, Bernard Harris, Badrig Kurkjian, Clifford J. Maloney and Herb Solomon. Francis Dressel served as secretary of this committee until 1 September. From then on the duties of the secretary were carried by Fred Frishman. Their help in guiding this conference to a successful conclusion is duly appreciated.

FRANK E. GRUBBS
Conference Chairman

TABLE OF CONTENTS

A G E N D A

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

Wednesday, 24 October 1973

0830-0900 REGISTRATION - Room 16, AMETA Building 90

0900-1130 GENERAL SESSION I
Auditorium, Second Floor, Bldg. 90

CALLING OF CONFERENCE TO ORDER
Dr. Norman Coleman, Local Chairman

WELCOMING REMARKS
Dr. Colin Hudson, Chief Scientist, US Army Armament Command
Dr. A. Lynn Bryant, Director, US Army Management Engineering Training Agency

CHAIRMAN OF SESSION I
Professor Herbert Solomon, Department of Statistics, George Washington, University, Washington, DC

BAYESIAN STATISTICS
Professor Jerome Cornfield, Department of Statistics, George Washington University, Washington, DC

RANKING AND SELECTION PROCEDURES FOR MULTIVARIATE NORMAL POPULATIONS
Professor Shanti S. Gupta, Department of Statistics, Purdue University, West Lafayette, Indiana

1130-1245 LUNCH

1245-1430 CLINICAL SESSION A
Room 14, Bldg. 90

CHAIRMAN
Douglas Tang, Walter Reed Army Institute of Research

PANELISTS
Jerome Cornfield, Department of Statistics, George Washington University, Washington, DC
Walter Foster, Fort Detrick, Maryland
Gerald Lieberman, Department of Operstions Research, Stanford University, Palo Alto, California
William A. Thompson, Jr., Department of Statistics, University of Missouri, Columbia, Missouri

Wednesday

1245-1430 CLINICAL SESSION A - (Cont'd)

INTERPRETATION OF ANALYSIS OF VARIANCE EFFECTS IN DESIGNS YIELDING A SUBJECTS X TREATMENT INTERACTION
Ray T. Sterner, Richard S. Teplick and James T. Wheeler, Fitzsimons Army Medical Center, Denver, Colorado

VITAMIN A MODELING
R. S. Teplick, Fitzsimons Army Medical Center, Denver, Colorado

1245-1430 TECHNICAL SESSION 1
Auditorium, Bldg. 90

CHAIRMAN
William McIntosh, US Army Test and Evaluation Command

SAMPLE SIZE DETERMINATION AND THE OPERATING CHARACTERISTIC CURVE FOR THE CIRCULAR PROBABLE ERROR
G. A. Culpepper, White Sands Missile Range, New Mexico

DESIGN AND ANALYSIS OF A HIT PROBABILITY EXPERIMENT BASED ON A BIVARIATE NORMAL DISTRIBUTION
R. W. Mai, Yuma Proving Ground, Yuma, Arizona

MODEL FOR PROBABILITY HIT ANALYSIS OF 20MM PROJECTILES
Diana L. Frederick, Frankford Arsenal, Philadelphia, Pennsylvania

1430-1500 BREAK

1500-1630 CLINICAL SESSION B
Room 14, Bldg. 90

CHAIRMAN
Gerard Dobrindt, US Army Test and Evaluation Command

PANELISTS
A. S. Cohen, Institute of Statistics, University of Georgia, Athens, Georgia
S. S. Gupta, Department of Statistics, Purdue University, West Lafayette, Indiana
J. Richard Moore, Ballistic Research Laboratories, Aberdeen, Maryland
Herbert Solomon, Department of Statistics, George Washington University, Washington, DC

CONDITION/EFFECTIVENESS MODEL FOR FACILITY COMPONENTS
R. J. Colver and J. G. Kirby, Construction Engineering Research Laboratory, Champaign, Illinois

USING THE BETA DISTRIBUTION TO SUGGEST A DISTRIBUTION OF ACCESS TIME TO COMPUTER MEMORY
W. Foster, Ft. Detrick, Maryland

1500-1630 TECHNICAL SESSION 2
Room 15, Bldg. 90

Wednesday

1500-1630 TECHNICAL SESSION 2 - (Cont'd)

CHAIRMAN
Roger Brauer, US Army Construction Engineering Research Laboratory

ON BIMODALITY IN PARAMETERIZED ATMOSPHERIC MODELS AND SOME SOLUTIONS TO AVOID IT
O. E. Essenwanger, U.S. Army Missile Research, Development and Engineering Laboratory, Huntsville, Alabama

PREDICTION OF EFFECTS OF ULTRA SHORT LASER PULSES
E. W. Steubing, E. A. Lucia and F. D. Verderame, Frankford Arsenal, Philadelphia, PA

MODELLING SEE-THRU AEROSOLS FOR INFRARED COUNTERSURVEILLANCE AND LASER COUNTERMEASURES
E. W. Steubing, R. Doherty, J. J. Pinto and F. D. Verderame, Frankford Arsenal, Philadelphia, PA

1500-1630 TECHNICAL SESSION 3
Auditorium, Bldg. 90

CHAIRMAN
Frank Proschan, Florida State University, Tallahassee, Florida

TRANSFORMATIONS THROUGH A NON-EUCLIDEAN SPACE IN A LINEAR TRANSFORMATION CONTEXT
O. N. Dalton, White Sands Missile Range, New Mexico

REGRESSION ANALYSIS APPROACH TO INTERPOLATION ALGORITHMS
E. L. McDowell, Construction Engineering Research Laboratory, Champaign, Illinois

1830-1915 SOCIAL HOUR

1915- BANQUET

PRESENTATION OF THE SAMUEL S. WILKS MEMORIAL AWARD
Dr. Frank Grubbs, Ballistic Research Laboratories, Master of Ceremonies

****

Thursday, 25 October 1973

0830-0945 TECHNICAL SESSION 4
Auditorium, Bldg. 90

CHAIRMAN
Arlen Dillin, GEN Thomas J. Rodman Laboratory, Rock Island, Illinois

Thursday

0830-0945 TECHNICAL SESSION 4 - (Cont'd)

RELIABILITY GROWTH OF THE TF-30 ENGINE
B. W. Haines, F. L. Carter and M. F. Cummings, Naval Air Station, Norfolk, Virginia

RELIABILITY GROWTH ESTIMATION FROM FAILURE AND TIME TRUNCATED TESTING
L. H. Crow, Army Materiel Systems Analysis Agency, Aberdeen Proving Ground, Maryland

0830-0945 TECHNICAL SESSION 5
Room 14, Bldg. 90

CHAIRMAN
William Fulkerson, GEN Thomas J. Rodman Laboratory, Rock Island, Illinois

RESPONSIVENESS PROPERTIES OF CONTINUOUS SAMPLING PLANS
R. M. Brugger, U.S. Army Armament Command, Rock Island, Illinois

A COMPUTATIONALLY FEASIBLE ALGORITHM FOR THE AFI FUNCTION ARISING FROM SHORT-RUN SCP-1
D. L. Arp, Weapons Planning Group, China Lake, California

0830-0945 TECHNICAL SESSION 6
Room 15, Bldg. 90

CHAIRMAN
Richard D'Accardi, U.S. Army Electronics Command, Fort Monmouth, New Jersey

SYSTEM SIMULATION OF FIVE MODULE ARRAY PRODUCTION LINES CONTAINING INDEPENDENT COUPLED AND CROSS LINKED BUFFER CONFIGURATIONS
M. Roffman, Frankford Arsenal, Philadelphia, Pennsylvania

CONFIGURATION MANAGEMENT FOR RELIABILITY/MAINTENANCE TESTING
J. Fargher, Jr. Frankford Arsenal, Philadelphia, Pennsylvania

0945-1015 BREAK

1015-1130 TECHNICAL SESSION 7
Auditorium, Bldg. 90

CHAIRMAN
Michael H. Strub, U.S. Army Research Institute

BAYESIAN CONFIDENCE INTERVALS FOR INTERVAL RELIABILITY OF WEIBULL COMPONENTS
R. Racicot, Watervliet Arsenal, Watervliet, New York

RELIABILITY OF A SERIAL SYSTEM
A. E. Johnsrud, U.S. Army Concepts Analysis Agency, Bethesda, Maryland

Thursday

1015-1130 TECHNICAL SESSION 8
Room 14, Bldg. 90

CHAIRMAN
William S. Agee, White Sands Missile Range

THE ROLE OF INSTRUMENT PERFORMANCE STANDARDS IN THE CONTROL OF THE WHITE SANDS MISSILE RANGE TEST SUPPORT PROCESS
J. B. Gose and W. R. Jenkins, Jr., White Sands Missile Range, New Mexico

MULTIVARIATE ANALYSIS TECHNIQUES APPLIED TO EQUIPMENT TESTING
J. B. Wilburn, Jr., U.S. Army Electronic Proving Ground, Fort Huachuca, Arizona

1015-1130 TECHNICAL SESSION 9
Room 15, Bldg. 90

CHAIRMAN
Bernard F. Engebos, White Sands Missile Range

A ONE-SHOT SENSITIVITY TEST PROCEDURE AND MODEL FOR EXTREME PERCENTAGE POINTS
S. K. Einbinder, Picatinny Arsenal, Dover, New Jersey

ANALYSIS OF TECHNIQUES FOR FINDING A FORWARD OBSERVER
Charles McElwee, Frankford Arsenal, Philadelphia, Pennsylvania

1130-1245 LUNCH

1245-1345 TECHNICAL SESSION 10
Room 14, Bldg. 90

CHAIRMAN
Edward N. Fiske, Edgewood Arsenal

PREDICTIVE EQUATIONS FOR ACCELERATED LIFE TESTING OF IMPREGNATED CHARCOAL
W. S. Magee, Jr., J. A. Baker, E. J. Poziomek, John Boardway and Harold Ball, Edgewood Arsenal, Aberdeen Proving Ground, Maryland

A GAUSSIAN TOLERANCE TECHNIQUE FOR MULTILAYER OPTICAL COATINGS
J. J. Walls, Jr., R. McKyton and A. Kawalec, Frankford Arsenal, Philadelphia, Pennsylvania

1245-1345 TECHNICAL SESSION 11
Auditorium, Bldg. 90

CHAIRMAN
George Kalemkarian, GEN Thomas J. Rodman Laboratory, Rock Island, Illinois

Thursday

1245-1345 TECHNICAL SESSION 11 - (Cont'd)

A TABLE OF CUMULATIVE DISTRIBUTION FUNCTION VALUES OF EXPECTED VALUES OF NORMAL ORDER STATISTICS
G. L. Aasheim, U.S. Army Armament Command, Rock Island, Illinois

A STRONG JUSTIFICATION FOR THE USE OF RANK TESTS IN THE CASE OF NON-NORMALITY
J. Sethuraman, Florida State University, Tallahassee, Florida and D. H. Jones, Rutgers University, New Brunswick, New Jersey

1245-1345 TECHNICAL SESSION 12
Room 15, Bldg. 90

CHAIRMAN
Mrs. Genevieve Meyer, U.S. Army Mobility Equipment R&D Center

STATISTICAL ANALYSIS AND MODELING OF PATH LOSS DISTANCE DEPENDENCY INFORMATION
R. D'Accardi and D. Dense, Electronics Command, Ft. Monmouth, New Jersey, and C. Tsokos, University of South Florida, Tampa, Florida

AN APPROACH TO OCCUPANT EVALUATION OF ARMY FAMILY HOUSING INTERIORS
R. Brauer and R. Neathammer, Construction Engineering Research Laboratory, Champaign, Illinois

1345-1415 BREAK

1415-1630 GENERAL SESSION II
Auditorium, Second Floor, Bldg. 90

CHAIRMAN
Dr. Richard Moore, U.S. Army Armament Command, Rock Island, Illinois

GENERALIZED JACKKNIFE TECHNIQUES
Dr. H. L. Gray, Charles F. Frensley, Professor of Mathematical Sciences, Southern Methodist University, Dallas, Texas

RELIABILITY GROWTH
Professor Frank Proschan, Department of Statistics and Statistical Consulting Center, Florida State University, Tallahassee, Florida

****

Friday, 26 October 1973

0830-1000 GENERAL SESSION III
Auditorium, Second Floor, Bldg. 90

CHAIRMAN
Dr. Frank E. Grubbs, Ballistics Research Laboratories

Friday

0830-1000 GENERAL SESSION III - (Cont'd)

OPEN MEETING OF THE SC SUBCOMMITTEE ON PROBABILITY AND STATISTICS
Dr. Walter D. Fo[illegible]r, Fort Detrick, Frederick, Maryland

ACCELERATED LIFE TESTING
Professor Sam C. Saunders, Department of Pure and Applied Mathematics, Washington State University, Pullman, Washington

1000-1030 BREAK

1030-1130 GENERAL SESSION III - (Cont'd)

RELIABILITY OF MULTIPLE COMPONENT SYSTEMS
Professor William A. Thompson, Jr., Department of Statistics, University of Missouri, Columbia, Missouri

BAYESIAN STATISTICS

Jerome Cornfield
Department of Statistics
The George Washington University
Washington, D.C. 20006

ABSTRACT. In the standard frequency approach to statistics one is concerned with properties over the sample space in contrast to the Bayesian approach which is concerned with properties over the parameter space. Thus, given n independent trials at constant probability p the frequentist is concerned with the probability of such and such a number of successes, whereas the Bayesian asks, given whatever number of successes have been observed, for the probability that p falls within certain limits. The answer to the latter, but not the former, question involves the use of prior probabilities. It therefore seems natural to conclude that the Bayesian is distinguished from the rest of statistical mankind by some special wish to use some unidentified entity called "prior information" in addition to the "objective" information available to everyone else. This is an incomplete and unproductive introduction to Bayesian statistics and is more likely to lead to polemics than to understanding.

A more informative introduction can be achieved by criticizing some of the ideas underlying the frequentist viewpoint. I shall therefore start by giving several examples illustrating the anomalous results that can follow from attempting to judge statistical procedures by reference to their average properties over repeated samples. These anomalies are seen to have the common property of incoherence (defined below). We then go on to sketch the argument that Bayesian procedures are necessary and sufficient to avoid incoherence. From this point of view the Bayesian augument is a natural extension and correction of that of the frequentist, and the prior probabilities an inevitable consequence of the search for procedures which will, in full generality, avoid incoherence.

Ex. 1

In this example we exhibit two different tests of the same hypothesis. Both

have the same Type I error, but the one with the smaller type II error seems clearly absurd. Something more than Type I and Type II errors must therefore be involved in selecting tests of hypotheses.

We consider the following simple null and alternative hypotheses.

$$H_0: \quad x \overset{d}{=} N(0,\sigma^2)$$

$$H_1: \quad x \overset{d}{=} N(10,\sigma^2)$$

σ^2 is known, once an auxiliary random variable, y, is observed. When y = 0, σ^2 is known to be 1, and when y = 1, $\sigma^2$is known to be 100. Pr(y=1) =p, which in the numerical example that follows is taken to have value $.05-(3\times10^{-6})$

We can think of the experiment as involving two steps:

(a) choosing an instrument for measuring x, with the imprecise instrument having a probability, p, of being chosen

and (b) using the instrument chosen to measure x.

These considerations are summarized in the table below:

Instrument	σ^2	y	P(y)
Precise	1	0	1-p
Imprecise	100	1	p $(=.05-3\times10^{-6})$

We now consider two different tests fo H_0. For

Test A

Reject H_0 when x>1.645, given y = 0

x>(10)(1.645), given y = 1

Test B

Reject H_0 when x> 4.5, given y = 0

x>-∞ , given y = 1

Test A seems like a plausible test, since for the case where p = 0 or 1 it is the most powerful test, while test B seems unacceptable, since for y = 1, it rejects

the hypothesis that $\mu = 0$, even for an observed $x = 3$. Nevertheless, as summarized below, if the choice between A & B is to be based on standard frequency criteria, B is the test of choice, since it has essentially zero type II error.

Operating characteristics

Test	Type I error	Type II error
A	.05	.037
B	.05	$(1-p)\Phi(-5.5) \approx 0$

($\Phi(x)$ is the standard normal integral from $-\infty$ to x.)

Why do we prefer A to B, even though on standard frequentist criteria it is the inferior test? A concept with which Fisher familiarized statisticians, although the idea is older and goes back at least to Keynes, is that our criteria must take account not only of properties over the entire sample space but over any recognizable subset of it. For the present example two such recognizable subsets are defined by the cases $y = 0$ and $y = 1$. For test A the type I error is .05 for both subsets, but for set B it is not. Thus,

Recognizable subset	Conditional probability of rejection $\mid H_0$	
	Test A	Test B
$y = 0$	.05	3×10^{-6}
$y = 1$	.05	1

Although the unconditional probability of rejection over the entire sample space is the same for both tests, this is not true for the conditional probabilities. Given that the sample point is in the subset defined by $y = 1$ we know for sure that Test B will give the wrong answer and why should anyone use a test that is certain to be wrong in identifiable circumstances?

Ex. 2

The same point is made here but for a very common rather than contrived problem, confidence limits on the ratio of two independent normal variables. An argument leading to the most restrictive confidence set is as follows. Given x_i $(i=1,2)$ which is $N(\mu_i, 1)$ and $\theta = \mu_1/\mu_2$, confidence limits on θ can be obtained by noting

(1) $x_1 - \theta x_2$ is normal

(2) $E(x_1 - \theta x_2) = 0$

(3) $\mathrm{Var}(x_1 - \theta x_2) = 1 + \theta^2$,

so that if $\chi^2_{1-\alpha}$ is the upper $(1-\alpha)$ probability point of one degree of freedom chi-square distribution, the confidence set of θ with coefficient $1-\alpha$ is the set of θ satisfying

$$(1) \qquad \frac{(x_1 - \theta x_2)^2}{1 + \theta^2} \leq \chi^2_{1-\alpha}.$$

By multiplying out and collecting coefficients of the powers of θ this is seen to be equivalent to

$$(2) \qquad \theta^2(x_2^2 - \chi^2_{1-\alpha}) - 2\theta x_1 x_2 + x_1^2 - \chi^2_{1-\alpha} \leq 0.$$

When the discriminant of the quadratic on the left hand side is positive, the inequality will hold for all θ between the two roots of the quadratic obtained by setting the left hand side equal to zero. These two values of θ supply a standard confidence interval for points in the sample space leading to a positive discriminant. Consider now the subset of points in the sample space in which both the discriminant and the coefficient of θ^2 in (2) are negative.

For this case it is easy to see that the inequality (2) holds for all θ so that the confidence interval is $-\infty$ to $+\infty$ - asserted with confidence $1-\alpha$. Surely this is being over-cautious.

The anomalous feature of the result - which, from the point of view of standard frequentist criteria cannot be improved upon - can again be highlighted by considering the concept of recognizable subsets. Thus, for one such subset the confidence interval is certain to include θ and for the other it includes it with probability $< 1-\alpha$

Recognizable subset	Conditional probability of including θ
$x_2^2 < \chi^2_{1-\alpha}$, discriminant <0	unity
all other	$<1-\alpha$

Unconditionally, the probability of including θ is indeed $1-\alpha$, but this does not correspond to our confidence for either of the two identifiable subsets.

We now show that a fully general effort to avoid such anomalies leads to Bayes theorem. We start, with example 2 in mind, by asking a frequentist statistician who assigns a confidence coefficient $1-\alpha$ to an interval whether he would be willing to bet that the interval always contains the true value at odds of $1-\alpha$ to α, plus say a small premium for being obliging. He naturally agrees since his expected loss on the bets is zero no matter what the true value and he gains the premium. This is a consequence of the frequency properties of the confidence set defined over the entire sample space.

We now modify our question. As before we ask him to offer odds of $1-\alpha$ on the set to α against it, but to <u>permit us to decide whether we bet for or against the interval.</u> Thus, for sample points leading to an infinite interval on θ we bet for the interval and offer him odds of $1-\alpha$ to α. For this subset we always win, since the interval includes all θ. For the remaining subset we always bet against the interval, again accepting his odds of $1-\alpha$ to α, since

for this subset we shall have probability of winning greater than α. Our expected gain is thus positive no matter what the value of θ. (The premium is clearly a theoretical irrelevancy, and can always be eliminated by letting the size of the bet overpower the size of the stake.)

This experience of losing for all values of θ, despite the use of confidence limits with optimum properties causes the frequentist statistician to consider whether there is some alternative system of setting probabilities on parameter sets that would safeguard him against this certain loss. This sets him on the Bayesian path, since his probability assignments must be Bayesian in the sense of the theorem stated below. We here sketch out the ideas leading to the theorem. For the detailed argument see [1]. We consider a betting game with a cast and a set of rules. The cast consists of

A - an antagonist

B - a probability setter

C - a referee

The rules are that 1. We are given an m x n matrix in which each row corresponds to a possible sample point and each column to a possible state of nature, of which discrete parameter points are a special case. The entry in the i^{th} row and j^{th} column, p_{ij}, gives the probability of the i^{th} sample when the j^{th} state of nature obtains ($\sum_{i=1}^{m} p_{ij} = 1$ for $j = 1,2...n$). Both A and B know the p_{ij}.

2. A value of i is selected with probability p_{ij} and both A and B are informed as to the value of i, this is, which sample is drawn, but not the value of j, i.e. not the true state of nature. C however, knows j.

3. B's task is to announce a probability $P_i(I)$, where I is an interval containing some subset of the states of nature, i.e. a subset of the integers 1 to n. (Clearly there are 2^n possible I's.) 4. A selects a stake, S_i, which may be

positive or negative, and gives B amount $P_i(I)S_i$. 5. C determines whether I includes the true state of nature j. If it does, the interval is correct and A has won the bet, and B pays him amount S_i. If it does not, A has lost, and B pays nothing. (When S_i is negative the direction of money flow is reversed.)

It will be recognized that these rules are equivalent to B offering odds of

$P_i(I)$ to $1-P_i(I)$ that I is correct when $S_i>0$

and $1-P_i(I)$ to $P_i(I)$ that I is incorrect when $S_i<0$.

Now consider B's expected gain when the j^{th} state of nature obtains. B would clearly like to select the $P_i(I)$ in such a way that this is in some sense maximized. But clearly a minimum requirement would be that the gain not be negative for all states of nature, as in example 2, no matter how A chooses the S_i. We state this more precisely by denoting B's expected gain by $G_j(P,S)$ and then introducing the key notion of coherence [2].

Definition

If $G_j(P,S)\leq 0$ for all j for any S and the equality can be deleted for at least one j, B's assignment of the P is incoherent. Otherwise it is coherent.

With these concepts it is then possible to prove:

Theorem

A necessary and sufficient condition for coherence is that

$$P_i(I) = \sum_{j\in I} p_{ij}q_j \Big/ \sum_{j=1}^{n} p_{ij}q_j .$$

where the q_j are any positive quantitites.

$P_i(I)$ can then be interpreted, by Bayes' theorem, as the conditional, or posterior, probability that I includes j, given the realization of sample i. Similarly p_{ij} (or more strictly any quantity proportional to p_{ij} where the

proportionality constant does not depend on j) is the likelihood of state of nature j, given sample i. Finally $q_j / \sum_1^n q_j$ can be interpreted as the prior probability of state of nature j.

Note that aside from the restriction to finite sample and parameter spaces the formulation is general. Without the restriction to finite parameter spaces the theorem need not be true, since the proof, which depends on the interchange of order of summation, need not be true when the restriction to finite spaces is removed. Although the restriction to finite parameter spaces is very severe from a mathematical point of view, I, and many statisticians who are far better mathematicians than I, do not regard it as any real restriction at all.

Nothing in this formulation tells us how to select the q_j. So long as they are positive for all j, any q_j will lead to coherence. This has been regarded by many statisticians as the fatal flaw in the Bayesian argument, but to maintain this consistently one must reject coherence as an overriding criterion. Some statisticians have been willing to do this, rather than permit the q_j to enter the domain of statistics. It is not easy to discuss this point constructively, and so I simply record my own view that such rejection is premature, and that the introduction of the Bayesian outlook and of prior probabilities has already paid handsome dividends in new insights and results over the entire range of statistical theory.

Thus, unless one rejects coherence as a criterion, the theoretical inevitability of prior probabilities seems beyond dispute. What seems most needed now is experience in applying these ideas to real problems, rather than continued theoretical exegesis. In this spirit I summarize an application to the computer interpretation of the ECG described in detail elsewhere [3].

We start with a k-dimensional vector of ECG characteristics, $\underline{x}$, from a patient of unknown diagnostic status. (This statement glosses over what is in some respects the hardest part of the problem of computerizing ECG's, but one that is not relevant here.) We let j index the possible diagnostic classes ($j = 1,2...,n$) and estimate the conditional density of $\underline{x}$, given j, say $f(\underline{x}|j)$, from data on patients of known diagnostic status. Denoting, as before, the prior probability of being in the j^{th} diagnostic class by q_j, we compute the posterior probability of being in category j, given $\underline{x}$ as

$$P(j|\underline{x}) = f(\underline{x}|j)q_j / \sum_{i=1}^{n} f(\underline{x}|i)q_i .$$

These probabilities are printed out for each patient.

The prior probabilities are interpreted as the relative frequency with which patients in the different diagnostic classes present themselves. These quantities are often not known precisely and for that reason many attempts at medical diagnosis have dispensed with the concept entirely [4]. We, on the other hand, have gone ahead and estimated them as best as we could, separately for each institution using the program, as illustrated in Tables 1 and 2.

There are of course many problems involved in computer interpretation of the ECG. These are slowly being identified and handled. The imprecision with which the prior probabilities are estimated is one of the less pressing of these. A more important one, which is a consequence of the Bayesian formulation, but which exists for non-Bayesian approaches as well, is that a patient may not belong to any of the diagnostic classes included in the program, in which case the computed probabilities are misleading. (A preliminary test of the hypothesis that a patient comes from one of the n classes seems called for here even though this falls outside the realm of Bayesian ideas.)

There are of course many other applications of Bayesian ideas that might be mentioned, but the basic point, that their usefulness must be appraised in the context of applications, remains. I hope that some readers will be stimulated to try them out.

References

[1] Cornfield, J., "The Bayesian Outlook and its Application," Biometrics, 25, 617, 1969.

[2] DeFinetti,B.,"Foresight: Its logical Laws, Its Subjective Sources." In: Studies in Subjective Probability, ed. by H.E. Kyburg, Jr. and H.E. Smokler, New York, Wiley (reprinted from Annales de l'Institut Henri Poincaré, 71,(1937)).

[3] Cornfield, J., Dunn, R.A., Batchlor, C.D., and Pipberger, H.V.,"Multigroup Diagnosis of Electrocardiograms", Computers and Biomedical Research, 6,97,1973

[4] Collen, M.F., Rubin, L.,Neyman, J., Dantzig, G.B., Baer, R.M., Siegelaub, A.B., "Automated Multiphasic Screening & Diagnosis," Amer. Journal Pub. Health, 54, 741, 1964.

TABLE 1

MISCLASSIFICATION MATRIX USING PRIORS OF THE ADMITTING OFFICE OF VA HOSPITAL, WASHINGTON, D.C. (2336 PATIENTS)[a]

Clinical		Computer (1)	(2)	(3)	(4)	(5)	(6)	(7)	Prior probabilities
N	(1)	98.0[b]	0.3	0.7	0.0	0.5	0.0	0.5	0.60
AMI	(2)	14.5	60.2	3.5	2.6	12.5	0.6	6.1	0.07
PMI	(3)	11.6	4.0	77.3	0.9	3.1	0.7	2.4	0.10
LMI	(4)	8.2	14.1	2.4	69.4	2.4	0.0	3.5	0.02
LVH	(5)	34.8	9.1	5.6	0.8	44.7	2.0	3.0	0.13
RVH	(6)	48.8	4.1	2.5	0.0	4.1	28.9	11.6	0.02
PE	(7)	27.7	8.4	9.2	0.9	3.4	2.6	47.8	0.06

[a] Total percentage correctly classified = $98.0 \times 0.60 + 60.2 \times 0.07 + \cdots = 81.4$.

[b] Note that only 2% of the normals are misclassified with a concomitant increase in misclassifications of abnormal records.

N = normal

AMI = anterior myocardial infarct

PMI = posterior myocardial infarct

LMI = lateral myocardial infarct

LVH = left ventricular hypertrophy

RVH = right ventricular hypertrophy

PE = pulmonary embolism

TABLE 2

MISCLASSIFICATION MATRIX USING PRIORS OF THE CARDIOLOGY SERVICE OF THE VA HOSPITAL WEST ROXBURY, MA (2336 PATIENTS)[a]

Clinical		Computer (1)	(2)	(3)	(4)	(5)	(6)	(7)	Prior probabilities
N	(1)	76.3[b]	1.8	3.2	0.0	15.4	1.0	2.2	0.12
AMI	(2)	3.5	68.0	3.8	3.2	19.2	0.3	2.0	0.19
PMI	(3)	1.3	4.7	84.2	0.9	6.5	1.1	1.3	0.24
LMI	(4)	1.2	17.6	1.2	71.8	7.1	0.0	1.2	0.06
LVH	(5)	12.7	10.9	6.1	0.8	64.7	1.8	3.0	0.30
RVH	(6)	21.3	9.1	8.3	0.0	14.9	35.5	10.7	0.03
PE	(7)	9.8	15.0	12.1	0.9	11.0	3.5	47.8	0.06

[a] Total percentage correctly classified = $76.3 \times 0.12 + 68.0 \times 0.19 + \cdots = 69.9$.

[b] Note that 24% of the normals were misclassified with an increase in correct classifications of abnormals. Thus sensitivity and specificity of the method can be adjusted according to specific requirements.

Selection and Ranking Procedures for Multivariate Normal Populations*

Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, Indiana

ABSTRACT.

This paper deals with selection and ranking procedures for multivariate normal populations. Let π_i be $N(\underline{\mu}_i, \Sigma_i)$, $i = 1,\ldots,k$, be $k(\geq 2)$ multivariate normal problems. Procedures for selecting a subset containing the (unknown) population with the smallest generalized variance, the largest Mahalanobis distance function and the largest (smallest) multiple correlation coefficient are described. The paper also surveys other known results in ranking problems for these populations and mentions some unsolved problems.

1. INTRODUCTION. The classical tests of homogeneity are inadequate in many practical situations in which the experimenter has to make a decision regarding k populations. This inadequacy is not in the development of these tests but rather in the basic formulation itself which is not designed to answer many questions which are of real interest to the experimenter. The early attempts to find more meaningful formulations led to partial answers in the form of slippage tests and tests for outliers. Some important contributions have been made in this area, among others, by Mosteller (1948), Grubbs (1950), Karlin and Truax (1960), and Doornbos (1966). The initial contributions in the direction of multiple decision problems were notably made by Paulson (1949) and Bahadur (1950). The basic formulations of what is now commonly known as selection and ranking problems are due to Bechhofer (1954) and Gupta (1956).

Let $\pi_1,\ldots,\pi_k$ be k independent populations with underlying distribution functions F_{θ_i}, $i = 1,\ldots,k$. The θ_i are unknown values of a quality characteristic θ which is used as the criterion for ranking the populations. To be specific,

*This research was supported in part by the Office of Naval Research Contract N00014-67-A-0226-00014 at Purdue University. Reproduction in whole or in part is permitted for any purposes of the United States Government.

we define π_i to be better than π_j if $\theta_i \geq \theta_j$. Let the ordered θ_i be denoted by $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$. The experimenter is assumed to have no prior knowledge regarding the correct pairing of the ordered and the unordered θ_i.

The goal of the experimenter in its simple version under Bechhofer's formulation is to choose one of the populations and claim that it is the best (i.e. the one associated with $\theta_{[k]}$). A correct selection (CS) in this formulation is the selection of any one of the populations associated with $\theta_{[k]}$. Any selection rule R is required to have the associated probability of a correct selection denoted by $P(CS|R)$ at least as large as a pre-determined $P^*(\frac{1}{k} < P^* < 1)$ whenever the distance (suitably defined) between the best and the second best populations, denoted by $\delta(\theta_{[k]}, \theta_{[k-1]})$, is at least as large as a specified constant $\delta^* > 0$. Of course, the experimenter has to specify P^* and δ^* satisfactory to him. The problem is to determine the smallest sample size depending on δ^* and P^* and other parameters which will meet the basic probability requirement. This formulation known as the indifference zone formulation derives its name from the fact that the experimenter is 'indifferent' to distinguish between the best and the second best unless they are sufficiently apart.

The goal of the second formulation due to Gupta (1956) is to select a non-empty subset from the given populations so as to include the best population. In the case of more than one population associated with $\theta_{[k]}$, one of them is assumed to have been tagged as the best. A correct selection in this case will be the selection of any subset containing the best. Under this subset selection approach, it is required that for any rule R, $P(CS|R) \geq P^*$ whatever be the true configuration of the unknown θ_i. Here the subset selected is of random size and one wishes to select a rule which satisfies the probability required with as small an expected subset size as possible.

Many authors have contributed to the area of subset selection procedures and a general survey of the work in this area has been made by Gupta and Panchapakesan (1973). Recently some work has been done by Santner (1973), and Gupta and Santner (1973) connecting the indifference zone and subset selection approaches.

In the present paper, we are concerned with subset selection procedures for multivariate normal populations using different criteria for ranking such as the generalized variance, Mahalanobis distance function and multiple

correlation coefficient. As it can be seen from the subsequent sections, the statistics employed in the selection rule are scalar-valued and have univariate distributions. Let $\underline{x}_{ij}$, $j = 1,2,\ldots,n$, be a sample of size n of vector observations from π_i which is $N(\underline{\mu}_i, \Sigma_i)$, $i = 1,2,\ldots,k$. Let

$$S_i = \frac{1}{n-1} \sum_{\alpha=1}^{n} (\underline{x}_{i\alpha} - \bar{\underline{x}}_i)(\underline{x}_{i\alpha} - \bar{\underline{x}})' = \text{Sample Covariance Matrix.}$$

Now we discuss the selection procedures for the three problems mentioned above.

2. Selection in terms of generalized variance, $|\Sigma_i|$.

For a multivariate normal distribution, the natural measure of dispersion is the covariance matrix. However, it is necessary, for the purpose of selection, to define a meaningful univariate measure of dispersion. Various different measures have been considered in statistical literature, but none of these is uniformly best, in the sense of being a robust estimator of the dispersion. The generalized variance is quite frequently used as a measure of dispersion. So we discuss selection in terms of this generalized variance $|\Sigma_i|$ associated with π_i. Σ_i are assumed to be unknown. Assume also that $\underline{\mu}_i$ are unknown. For selecting a subset containing the smallest $|\Sigma_i|$, Gnanadesikan and Gupta (1970) studied the following rule R, based on the sample covariance matrices S_i, $i = 1,2,\ldots,k$.

R: Select π_i iff

$$|S_i| \leq \frac{1}{c} |S|_{min}$$

where $|S|_{min} = \min(|S_1|,\ldots,|S_k|)$ and $c = c(k,p,n,P^*)$ is the largest value to satisfy the basic probability requirement $\inf P(CS|R) \geq P^*$. Note $0 < c \leq 1$. It has been shown that

$$\inf_{\Omega} P(CS|R) = P(Y_1 \leq \frac{1}{c} Y_j,\ j = 2,\ldots,k) \text{ where } Y_1,\ldots,Y_k$$

are independent. identically distributed random variables, each being the product of p independent factors, the rth factor being distributed as a χ^2 variable with (n-r) d.f. The proof follows from the fact that $|S_i|$ is distributed as $|\Sigma_i|/(n-1)^p$ times the product of p ind. χ^2 with (n-1), (n-2),...,(n-p) d.f.

The exact distribution of Y_i is unknown except when p = 2. In this case, $\inf P(CS|R) = P(Z_1 \leq \frac{1}{\sqrt{c}} Z_j,\ j = 2,\ldots,k)$ where Z_i are i.i.d. chi square with 2(n-2) d.f. In this case $\sqrt{c}$ is the lower 100(1-P*) percentage point of $F_{min} = \min_{2 \leq j \leq k} \chi^2_{\nu,j}/\chi^2_{\nu,1}$, $\nu = 2n-4$. Tables of $\sqrt{c}$ are available from Gupta and Sobel (1962) who studied the problem of selection of variances of normal populations.

When p > 2, one can use Hoel's approximation for the distribution of Y_i or use the normal approximation for χ^2. Some studies of these approximations were made in Gnanadesikan and Gupta (1970). Hoel (1937) suggested approximating $Y_i^{1/p}$ by the distribution having the density function $g(y) = \lambda^{p(n-p)/2} y^{(p(n-p)/2)-1} e^{-\lambda y}/\Gamma(\frac{1}{2} p(n-p))$ when $\lambda = \frac{p}{2}\left(1 - \frac{(p-1)(p-2)}{2n}\right)^{1/p}$. For p = 1,2, this approximation is exact.

Performance of the procedure R was studied in terms of risk functions using three different loss functions. If the ordered generalized variances are denoted by $|\Sigma|_{[1]} \leq \ldots \leq |\Sigma|_{[k]}$, the loss functions considered for including in the subset the population whose generalized variance is Σ_i, were:

(i) $L_1(\Sigma_i) = |\Sigma_i|/|\Sigma|_{[1]} - 1$

(ii) $L_2(\Sigma_i) = (\text{Rank of pop. } \pi_i)/\frac{k(k+1)}{2}$, where the ranks increase with the generalized variance, and,

(iii) $L_3(\Sigma_i) = \frac{S}{k}$, where $S(1 \leq S \leq k)$ is the random number of populations included in the subset.

The computations of the risk functions associated with the above loss functions, for p = 2, k = 2(1)5, $|\Sigma|_{[i]}/|\Sigma|_{[1]} = a^{2i-2}$ where a = 1.2(.2)2.0(.5)3.0, n = 3(1)7 and P* = .75 indicate that $E(L_2)$ and $E(L_3)$ are sensitive to changes in the values of the parameters and are decreasing functions of a and n. In the case of L_1 it increases in the range of values of a considered when n = 3 and for other values of n it increases upto a certain point and then decreases as a increases. This lack of monotonicity of $E(L_1)$ as the best population moves further away from the other populations, and the difficulty of its interpretation

render L_1 less suitable than L_2 and L_3. Between L_2 and L_3, due to ease of interpretation, L_3 seems more appropriate as the criterion of performance of R.

Finally, R was shown to be monotone i.e. for all $i < j$,
$P[R \text{ selects the pop. with } |\Sigma|_{[i]}] \leq P[R \text{ selects the pop. with } |\Sigma|_{[j]}]$.

<u>On approximating the constant c</u> - <u>Approximation to the dist. of |S|</u>

Let η = product of p factors each ind. central χ^2, the rth factor having (n-r)d.f.

(i) Then using methods of probability plotting, it is found that Hoel's approximation to the distribution of $\eta^{1/p}$ decreases in accuracy as p increases.

(ii) The approximation of the distribution of $\log \chi^2_\nu$ by the normal distribution improves with d.f., ν, of the χ^2 variable. For $\nu \geq 25$, the maximum difference between the quantiles is $\leq .4$ while for $.05 \leq \alpha \leq .95$, the difference is less than .05. The maximum error in probability in using the percentage point of the standard normal distribution as those of standardized $\log \chi^2$ is less than .02 for $n \leq 25$ and decreases with n.

(iii) The normal approximation for log(generalized sample variance) improves with both p and n. Approximating the dist. of $\frac{1}{p} \log \eta$ by the normal

$\inf_{\Omega} P(CS|R) \approx \int_{-\infty}^{\infty} \phi^{k-1}(y-c_0)d\phi(y)$ where $c_0^2 \approx (\log c)^2 / \sum_{i=1}^{p} \lambda_2(n-i)$ where

$\lambda_2(n) = \mathrm{Var}(\log \chi^2_n) \approx \frac{2}{n-1}$

3. <u>Selection in terms of distance functions</u>

Gupta (1966) considered the problem of selection of a subset of k multivariate normal populations which would include the population located farthest away from the origin, where the distance from the origin of the population with mean vector, $\underline{\mu}_i$ and covariance matrix, Σ_i, is defined as $\underline{\mu}_i' \Sigma_i^{-1} \underline{\mu}_i$, the Mahalanobis distance function. Let $y_{ij} = \underline{x}_{ij}' \Sigma^{-1} \underline{x}_{ij}$ where we assume Σ_i are all equal to Σ and $\underline{x}_{ij}$, $i = 1,\ldots,k$; $j = 1,\ldots,n$ is the vector with p-components of observations on the ith population. Then $y_i = \sum_{j=1}^{n} y_{ij}$ has a non-central χ^2 distribution with np d.f. and non-centrality parameter $\lambda_i' = n\lambda_i$ where $\lambda_i = \underline{\mu}_i' \Sigma_i^{-1} \underline{\mu}_i$. For selecting a subset containing the largest λ_i, Gupta proposed the rule R.

R: Select π_i iff $y_i \geq c \max(y_1,\ldots,y_k)$ where $0 < c = c(k,n,p,P^*) \leq 1$ is determined to satisfy the P* probability condition. It was shown that

$\inf P(CS|R) = \inf_{\lambda' \geq 0} \int_0^{\infty} F_{\lambda'}^{k-1}(x/c)dF_{\lambda'}(x)$ where $F_{\lambda'}(x)$=cdf of a non-central λ^2 with np d.f.

For selecting the populated nearest to the origin, R' was defined

R': Select π_i iff $y_i \leq b \min(y_1,\ldots,y_k)$ where $b = b(k,n,p,P^*) > 1$ is again determined to satisfy the P* condition. For this rule R',

$$\inf P(CS|R') = \inf_{\lambda' \geq 0} \int_0^\infty [1-F_{\lambda'}(x/b)]^{k-1}\, dF_{\lambda'}(x).$$

Monotonicity of the above two integrals wrt λ' has been shown by Gupta (1966) and Gupta and Studden (1970). Both integrals are monotonically increasing in λ' so that the inf takes place when $\lambda' = 0$. Thus the problem reduces to selecting the gamma population with the largest (smallest) scale parameters, respectively, which is solved in Gupta (1963) and in Gupta and Sobel (1962) where appropriate constants are available.

To be precise, Gupta and Studden (1970) considered the case where Σ_i are not necessarily equal but known. In this case the procedure is modified by using $y_{ij} = \underline{x}_{ij}' \Sigma_i^{-1} \underline{x}_{ij}$. They also studied the case where Σ_i are different but all unknown. In this case, let $z_i = \bar{x}_i' S_i^{-1} \bar{x}_i$. Then the procedures R_1 and R_1' are:

R_1: Select π_i iff $cz_i \geq \max(z_1,\ldots,z_k)$, $c > 1$.

R_1': Select π_i iff $z_i \leq b \min(z_1,\ldots,z_k)$, $b > 1$.

Gupta and Studden (1970) obtained a sufficient condition for the monotone behavior of $I(\lambda) = \int_0^\infty F_\lambda^{k-1}(cx)dF_\lambda(x)$ and $J(\lambda) = \int_0^\infty [1-F_\lambda(\frac{x}{c})]^{k-1}\, dF_\lambda$ where where $f_\lambda(x) = \sum_{j=0}^\infty \frac{e^{-\lambda}\lambda^j}{j!} g_j(x)$, $x \geq 0$ and $F_\lambda(x)$ = cdf of $f_\lambda(x)$ and $g_j(x)$, $j = 0,1,\ldots$ is a sequence of density functions on $[0,\infty)$. This condition is a special case of a mere general condition obtained by Gupta and Panchapakesan (1972) for a class of more general rules. The sufficient condition is satisfied for R,R' and also for R_1 and R_1'. For R_1 and R_1' the constants c and b are given by

$$\int_0^\infty F_{p,n-p}^{k-1}(cx)dF_{p,n-p}(x) = P^* \text{ and } \int_0^\infty [1-F_{p,n-p}(x/b)]^{k-1}\, dF_{p,n-p}(x) = P^* \text{ where}$$

$F_{p,n-p}(\cdot)$ is the cdf of a central F with p and n-p d.f.

It should be pointed out that the procedures R and R' for Σ_i known case are not strictly analogous to the procedures R_1 and R_1' for the unequal and unknown Σ_i case. If we use $\bar{x}_i \Sigma_i^{-1} \bar{x}_i$ in R and R', the constants c and b do not

depend on n, a feature which is undesirable. Alam and Rizvi (1966) use this type of rule. Another procedure, say, for the case of $\Sigma_i = \Sigma$, for all i,

$$z_i \geq z_{max} - d, \quad z_i = \bar{x}_i' \Sigma^{-1} \bar{x}_i$$

where d is given by $\inf_{\lambda' \geq 0} \int_0^\infty F_{\lambda'}^{k-1}(x+d) f_{\lambda'}(x)dx = P^*$ where F and f refer to non-central χ^2 with p d.f. and non-centrality parameter $\lambda' = n\lambda$.

Procedures of the above type when $\Sigma_i = \Sigma$, known, or Σ_i not equal but known, have <u>not been determined</u> in the sense that one does not know the monotone behavior of the integral involving d (above).

Another unsolved problem is the case where $\Sigma_1 = \ldots = \Sigma_k = \Sigma$, unknown, and we use a pooled estimate of Σ.

4. Selection in terms of multiple correlation coefficient

The random vector $\underline{X}_i'$ has multivariate distribution $N(\underline{\mu}_i, \Sigma_i)$ where $\underline{\mu}_i$ and Σ_i are unknown. Let $\rho_i \equiv \rho_{1.2\ldots p}$ be the multiple correlation coefficient between the first variable and the rest in the population π_i defined by

$1 - \rho_i^2 = \dfrac{|\Sigma_i|}{\sigma_{i11}|\Sigma_{i(11)}|}$ where σ_{i11} is the leading element of Σ_i and $\Sigma_{i(11)}$ is the matrix obtained from Σ_i by deleting the first row and the first column. This ρ_i (taken to be the positive square root of ρ_i^2) is the maximum of the correlation between X_{i1} and a linear combination of $x_{i2},\ldots,x_{ip}$ over all possible linear combinations and as such is a measure of the dependence of X_{i1} on $X_{i2},\ldots,X_{ip}$.

Gupta and Panchapakesan (1969) investigated the problem of selecting a subset containing the population associated with $\rho_{[k]}(\rho_{[1]})$. Let $R_i \equiv R_{1.2\ldots p}^{(i)}$ denote the multiple correlation coefficient obtained from the sample $\underline{x}_{ij}$, $j = 1,\ldots,n$. Two cases:

(i) $x_{i2},\ldots,x_{ip}$ are fixed; the conditional case.

(ii) $x_{i2},\ldots,x_{ip}$ are random; the unconditional case.

The following rule $\mathscr{R}$ has been investigated by Gupta and Panchapakesan (1969) for the selection of $\rho_{[k]}$.

$\mathscr{R}$: Select π_i iff $R_i^{*2} \geq c \max(R_1^{*2},\ldots,R_k^{*2})$ where $R_i^{*2} = R_i^2/(1-R_i^2)$. In the above rule, we write $c(0 < c \leq 1)$ formally both for the conditional case and the unconditional.

Letting $\lambda_i = \rho_i^2$, we can write the density of R_i^{*2} as

$$u_\lambda(x) = \sum_{j=0}^{\infty} \frac{\Gamma(q+m+j)\lambda^j}{\Gamma(q+m)\,j!}(1-\lambda)^{q+m} f_{2(q+j),2m}(x), \text{ unconditional case}$$

$$u_\lambda(x) = \sum_{j=0}^{\infty} \frac{e^{-m\lambda}(m\lambda)^j}{j!} f_{2(q+j),2m}(x), \text{ conditional case,}$$

where $q = \frac{p-1}{2}$, $m = \frac{n-p}{2}$ and $f_{r,s}(x)$ = density function of central F with r,s d.f.

It is easy to show that $u_\lambda(x)$ has MLR in x and hence the distribution of R^{*2} is stochastically increasing in λ. Hence

$$\inf P(CS|\mathscr{R}) = \inf_{\lambda \geq 0} \int_0^\infty U_\lambda^{k-1}(x/c)\,dU_\lambda(x), \text{ when } U_\lambda(x) = \text{cdf corresponding to } u_\lambda(x).$$

A sufficient condition for the monotonicity of the above integral is obtained by Gupta and Panchapakesan and it is verified so that the inf takes place when $\lambda = 0$ for the unconditional case. For the conditional case the result follows from the paper of Gupta and Studden. In either case we obtain

$\inf P(CS) = \int_0^\infty F_{2q,2m}^{k-1}(x/c)\,dF_{2q,2m}(x) = P^*$, F and f refer to central F. Table 1 of Gupta and Panchapakesan (1969) gives values of the constants c, which are same for both cases.

For selecting the population with the smallest ρ, the rule is $\mathscr{R}'$

$\mathscr{R}'$: Select π_i iff $dR_i^{*2} \leq \min_{1 \leq j \leq k} R_j^{*2}$, $0 < d = d(k,n,p,P^*) \leq 1$.

In an analogous manner, it follows for both cases

$$\inf P(CS|\mathscr{R}') = \int_0^\infty [1-F_{2q,2m}(xd)]^{k-1}\,dF_{2q,2m}(x) = P^*.$$

Since $1-F_{2q,2m}(xd) = F_{2m,2q}(\frac{1}{xd})$, the constants d can be obtained from constants c by interchanging q and m.

Consider $\int_0^\infty F_{2q,2m}^{k-1}(\frac{x}{c})\,dF_{2q,2m}(x) = P^*$.

When q and m are integers i.e. p and n are odd, we can use series expansion for $F_{2q,2m}(x)$ and obtain formulae for computing c for specified values of q,m and P^*. The final result is:

$$P^* = \frac{\Gamma(q+m)}{\Gamma(q)\Gamma(m)(1-c)^m} \sum_{\alpha=0}^{qk-1} \sum_{j=0}^{(k-1)(m-1)} (-1)^{\alpha}\binom{qk-1}{\alpha} a(k-1,j)\left(\frac{c}{1-c}\right)^{\alpha+j} K(c,m,q,\alpha,j)$$

where a(r,j) and K(c,m,q,α,j) are given by certain recurrence relations.

References

(1) K. Alam and M. H. Rizvi, Selection from multivariate populations, Ann. Inst. Statist. Math. 18, 307-318, (1966).

(2) R. R. Bahadur, On the problem in the theory of k populations, Ann. Math. Statist. 21, 362-375, (1950).

(3) R. E. Bechhofer, A single-sample multiple-decision procedure for ranking means of normal populations with known variances, Ann. Math. Statist. 25, 16-39, (1954).

(4) R. Doornbos (1966). Slippage Tests. Mathematical Research Centre, Amsterdam.

(5) M. Gnanadesikan and S. S. Gupta, Selection procedures for multivariate normal distributions in terms of measures dispersion, Technometrics, 12, 103-117, (1970).

(6) F. E. Grubbs (1950). Sample criteria for testing outlying observations. Ann. Math. Statist. 21, 27-58.

(7) S. S. Gupta, On a decision rule for a problem in ranking means. Mimeograph Series No. 150, Institute of Statistics, University of North Carolina, Chapel Hill, N. C. (1956).

(8) S. S. Gupta, On a selection and ranking procedure for gamma populations, Ann. Inst. Statist. Math., 14, 199-216, (1963).

(9) S. S. Gupta, On some selection and ranking procedures for multivariate normal populations using distance functions. In Multivariate Analysis (ed. P. R. Krishnaiah), 457-475, Academic Press, N. Y. (1966 b).

(10) S. S. Gupta and S. Panchapakesan, Some selection and ranking procedures for multivariate normal populations. In Multivariate Analysis II (ed. P. R. Krishnaiah) 475-505, Academic Press, N. Y. (1969).

(11) S. S. Gupta and S. Panchapakesan (1972a), On multiple decision (subset selection) procedures J. Mathematical and Physical Sci. 6, 1-72.

(12) S. S. Gupta and S. Panchapakesan (1972b), On a class of subset selection procedures, Ann. Math. Statist. 43, 814-822.

(13) S. S. Gupta and T. J. Santner (1973), On selection and ranking procedures - a restricted subset selection rule. Proceedings of the Ninth Session of the International Statistical Institute, Contributed Papers I, pp. 409-417.

(14) S. S. Gupta and M. Sobel, On the smallest of several correlated F-statistics, Biometrika, 49, 509-523 (1962).

(15) S. S. Gupta and W. J. Studden, On some selection and ranking procedures with applications to multivariate populations. In Essays in Probability and Statistics (ed. R. C. Bose et al), 327-338, University of North Carolina Press, Chapel Hill, N. C. (1970)

(16) S. Karlin and D. Truax (1960), Slippage problems, Ann. Math. Statist. 31, 296-394.

(17) F. Mosteller, A K-Sample slippage test for an extreme population, Ann. Math. Statist. 19, 58-65. (1948).

(18) E. Paulson, A multiple decision procedure for certain problems in the analysis of variance, Ann. Math. Statist. 20, 95-98, (1949).

(19) M. H. Rizvi and H. Solomon (1973), Selection of largest multiple correlation coefficients: asymptotic case, J. Amer. Statist. Assoc. 68, 184-188.

(20) T. J. Santner (1973), A restricted subset selection approach to ranking and selection problems. Mimeograph Series No. 318, Department of Statistics, Purdue University.

INTERPRETATION OF ANALYSIS OF VARIANCE EFFECTS IN DESIGNS YIELDING A SUBJECTS X TREATMENT INTERACTION

Ray T. Sterner, Richard S. Teplick, and James T. Wheeler
Letterman Army Institute of Research
Fitzsimons Army Medical Center
Denver, Colorado 80240

ABSTRACT. This paper outlines a proposed alternative analysis for data commonly obtained under a one-way, within-subjects analysis of variance design. Briefly, for situations in which multiple replicates are obtained within treatment cells, it is argued that inclusion of Subjects as an indigenous factor yields important information regarding the nature of subjects' responsivity to experimental treatments, as well as the reliability of population inferences. Results of an empirical application of the analysis to a set of plasma Vitamin A observations obtained from 8 male volunteers during a successive, 50-week period of dietary Vitamin A restriction are used to illustrate the approach. Comments from discussants aimed at clarifying (a) legitimacy of the approach (b) potential limitations imposed by the occurrence of a significant Subjects X Treatment term upon Treatment inferences and (c) alternative analytical procedures or methods for dealing with cases yielding such a Subjects X Treatment interaction are sought.

1. INTRODUCTION. One-way, within-subjects (i.e., repeated measures) analysis of variance designs are often used in experiments where: (1) costs of running multiple, independent treatment groups are prohibitive and/or (2) estimates of experimental effects relative to prior control values offer sufficient test of proposed hypotheses. Typically, this analysis partitions total experimental variation into a Between-Subjects and Within-Subjects components, with a subsequent F-Ratio of Between-Treatment/Residual developed from the Within-Subjects component (see Winer, 1971). Although this design controls inter-subject differences in variability (i.e., variability due to differences in the average responsiveness of subjects is removed from the Residual), no direct test of subjects' responsivity across treatments is afforded by this analysis. In cases where multiple replicates are obtained within treatment cells, however, this same analysis can be viewed as involving both a Treatment and Subject variance component. That is, estimates of the Between-Subjects and Between-Subjects X Treatment variance can be obtained. It is felt these components provide useful information in the assessment of the (1) nature of subject responsivity (i.e., the homogeneous or heterogeneous nature in the way subjects respond to specific treatments) and (2) generalizability of an observed treatment effect to a population.

2. PROPOSED MODEL. Consider the mixed effects analysis of variance model (i.e., with Treatment a fixed and Subjects a random effect variable):

$$X_{ij} = M + a_i + b_j + ab_{ij} + e_{ij}, \qquad (1)$$

where, X_{ij} refers to an observation of $subject_i$ under $treatment_j$, M refers to the grand mean term, a_i refers to a constant associated with $subject_i$, b_j refers to a constant associated with $treatment_j$, ab_{ij} refers to a constant associated with the $subject_i$ by $treatment_j$ interaction, and e_{ij} refers to experimental error. This is the basic additive model which characterizes a straightforward two-way analysis of variance comprised of Subjects and Treatments (see Winer, 1971).

Extending the above model to the repeated measures case in which multiple replicates are obtained for each subject within treatment cells yields:

$$X_{ijk} = M + a_i + b_j + c_k(a_i) + ab_{ij} + bc_{jk}(a_i) + e_{ijk}, \qquad (2)$$

where, X_{ijk} refers to a replicate of $subject_i$ under $treatment_j$, M refers to the grand mean term, a_i refers to a constant associated with $subject_i$, b_j refers to a constant associated with $treatment_j$, $c_k(a_i)$ refers to a constant associated with $replicate_k$ nested within $subject_i$, ab_{ij} refers to a constant associated with the $subject_i$ by $treatment_j$ interaction, $bc_{jk}(a_i)$ refers to a constant associated with the $treatment_j$ X $replicate_k$ nested within $subject_i$, and e_{ijk} refers to error. Thus, the proposed model makes it feasible to discuss the potential testing of variances attributable to Treatment, Subjects, and Treatment X Subjects sources. Additionally, inclusion of Subjects as an indigenous variable would seem to yield important information about the uniformity of subjects' responsivity to selected experimental treatments, highly useful data in attempts to generalize observed findings (treatment effects) to a designated population.[1]

3. EMPIRICAL APPLICATION OF THE ANALYSIS. In a recent dietary Vitamin A restriction experiment, 50 weekly Vitamin A plasma estimates (i.e., mg Vitamin A per 100 ml Plasma) were obtained from 8 adult male volunteers

---

[1] In the Discussion Session which followed presentation of the current paper, Prof. Cornfield referred to certain similarities between the current issue and one alluded to by himself and J. W. Tukey in an earlier publication (see Cornfield & Tukey, 1956).

for purposes of determining the "characteristic Vitamin-A-depletion pattern of humans." Due to considerable variability present in the observed plasma estimates (particularly during the early part of the experiment), it was decided to treat restriction as a successive series of 25, 2-week treatment values (2 replicates/session) for each subject. Moreover, a Log transformation of the data was performed in order to (1) avoid the likelihood of a treatment cell effect yielding a negative Vitamin A value, and (2) invoke a multiplicative model of Vitamin A depletion, an approach which seemed more realistic for describing the expected decreases in subjects' plasma A. Thus, the design was set up as a 25 (Treatments) X 8 (Subjects) X 2 (Replicates) factorial, with Treatments and Subjects viewed as repeated measures and Replicates nested within Subjects. This final model is the same as that listed in Equation 2, with the exception that Log X_{ijk} is substituted for X_{ijk}.

Results of the analysis of variance are presented in Table 1. As shown, the Treatment X Subjects ($F = 2.65$; $df = 168/192$, $p<.05$), Treatment ($F = 6.58$; $df = 24/168$, $p<.05$), and Subjects terms ($F = 7.04$; $df = 7/8$, $p<.05$) were all found to be significant. Whereas interpretations of the Treatment and Subjects main effects are fairly straightforward (i.e., mean Plasma Vitamin A estimates differed significantly across the 25 treatment sessions as well as between the 8 subjects), interpretation of the Treatment X Subjects term is more difficult. Basically, this term indicates that the Plasma A values for subjects varied uniquely as a function of specific treatment sessions.

TABLE 1
Analysis of variance for Log mg Vitamin A/100 ml Plasma values of 8 volunteers obtained during 25, 2-week treatment sessions of dietary A restriction (i.e., 2 replicates/subject/treatment session)

Source	df	MSq	F
Treatment	24	.988	6.58*
Treatment X Subjects	168	.150	2.65*
Treatment X Replicates w. Subjects	192	.057	
Subjects	7	.777	7.04*
Replicates w. Subjects	8	.110	

* $p<.05$

In an attempt to clarify the Treatment X Subjects interaction, plots of both the raw Plasma A values and Log Plasma A values for each subject by treatment session are shown in Figures 1a and 1b, respectively. (Note that the insert of Figure 1a also shows the mean mg Vitamin A/100 ml Plasma by Sessions for the group of 8 subjects.) As can be seen in these graphs, whereas all 8 of the subjects displayed similar decreases in their Plasma A values during the first part of restriction, midway through the restrictive period, 5 subjects showed stabilization and 3 subjects continued to display large decreases in Plasma A. This indicates subjects were heterogeneous in their response to prolonged A restriction, a result which limits the straightforward inference that subjects displayed a characteristic treatment (depletion) pattern throughout restriction.

While such interaction is explainable in a number of ways (i.e., widely differing depletion rates among individuals, intake of foods containing Vitamin A in violation of dietary regimen, etc.), the significance of a Subjects X Treatment term would seem to argue against the generalizability of the Treatment main effect. That is, if it is concluded on the basis of the Treatment effect that 50 weeks exposure to dietary Vitamin A restriction produces significant depletion, any inference regarding the "characteristic depletion rate of humans" still remains suspect due to the apparent differences shown by certain of the subjects. Thus, whereas under the classical one-way, within-subjects analysis, one would be led to conclude that Vitamin A was significantly decreased during a 50-week period of restriction (i.e., a result also yielded by the current approach), the occurrence of Subject X Treatment interaction in the proposed analysis would suggest such a conclusion be qualified, if not ignored altogether. Regardless of the pragmatic conclusion, however, the current analysis would seem to provide additional information regarding interpretation and inferences about the Treatment effect which has utility to a variety of research situations.

4. DISCUSSANT QUESTIONS. In conclusion, several questions concerning application of the proposed analysis are noted.[2] Specifically, these are:

1. Are there any problems with legitimacy of the analysis?

2. What are the inference limitations imposed upon the Treatment effect by occurrence of the Subjects X Treatment interaction?

3. What would be an appropriate method to quantify the divergence of subjects?

4. What criteria can be used for determining the appropriateness of data transformation(s) in the current analysis?

---

[2] Discussants for the current paper were: Prof. J. Cornfield, Prof. G. Lieberman, and Prof. W. A. Thompson, Jr.

5. REFERENCES.

1. Cornfield, J., and J. W. Tukey. 1956. "Average values of mean squares in factorials," ANN. MATH. STAT., Vol. 27 (1956), pp. 907-949.

2. Winer, B. J. Statistical Principles in Experimental Design (2nd ed.) McGraw-Hill, New York, 1971, pp. 278-282.

6. DISCUSSANT COMMENTS.

Prof. Lieberman:

I commend the authors for an interesting paper with well stated mathematical assumptions. I have no qualms about the mathematical model and their treatment of it. However, I have sound reservations about the application of the physical problem to this math model. I explain as follows:

The term "treatment" is used, and there are 25 "treatments." However, the "treatments" are just bi-weekly observations, i.e., "treatments" are time. Thus, the term treatment is artificial. I believe that the authors are interested in long run trends, as opposed to minor fluctuations. The suggested analysis of variance approach doesn't answer this question. Furthermore, there is a question in my mind whether or not the homoscedasticity assumption, within a subject, is valid. I believe that one subject may be more variable than another, and this variation is important. I believe that these comments address the first concern mentioned in the paper.

The second concern can be answered by the usual answer for main effects in the presence of interaction. The interaction says that the vitamin A concentration decreases in some subjects and stays the same or perhaps increases in other subjects. This implies there are "treatment differences."

The third area of concern is the heart of the problem in that it requires some suggestion about how to handle the overall analysis. In the absence of more information, I would plot the vitamin A plasma as a function of time (per week) for each subject, and see what conclusions can be drawn from each subject.

The final area of concern is the "transformation." I approve of the log transformation because it handles the case where variability is proportional to the concentration level, and I believe this to be valid in this case.

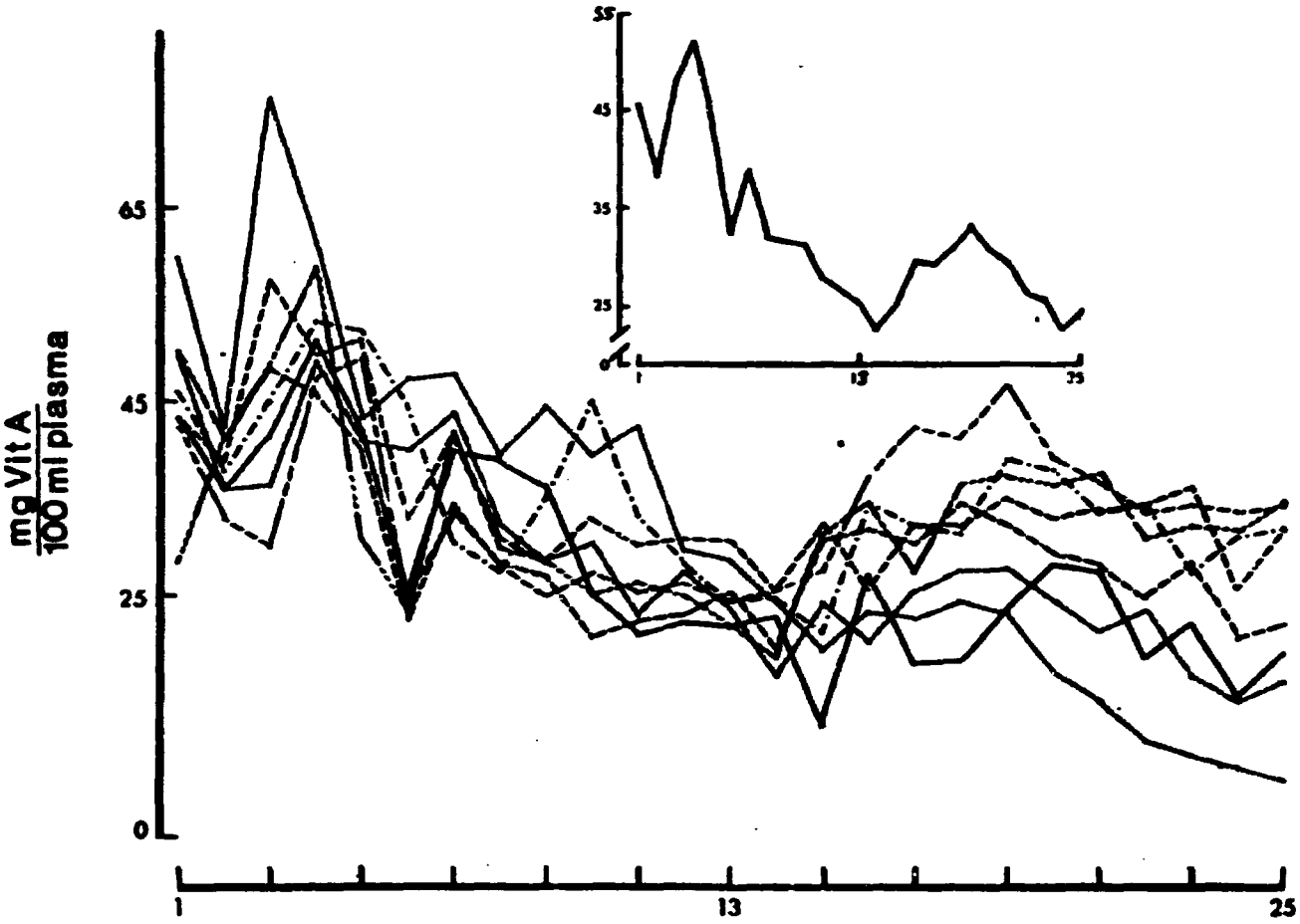
Prof. Thompson:

An alternative analysis of these data is suggested by Figure 1. It appears that a quadratic regression equation of Log concentrations (y) vs. time (x) may be fitted to the data for each volunteer. This leads to eight regression equations of the form

$$y = e^{\alpha + \beta x + \gamma x^2}, \qquad (3)$$

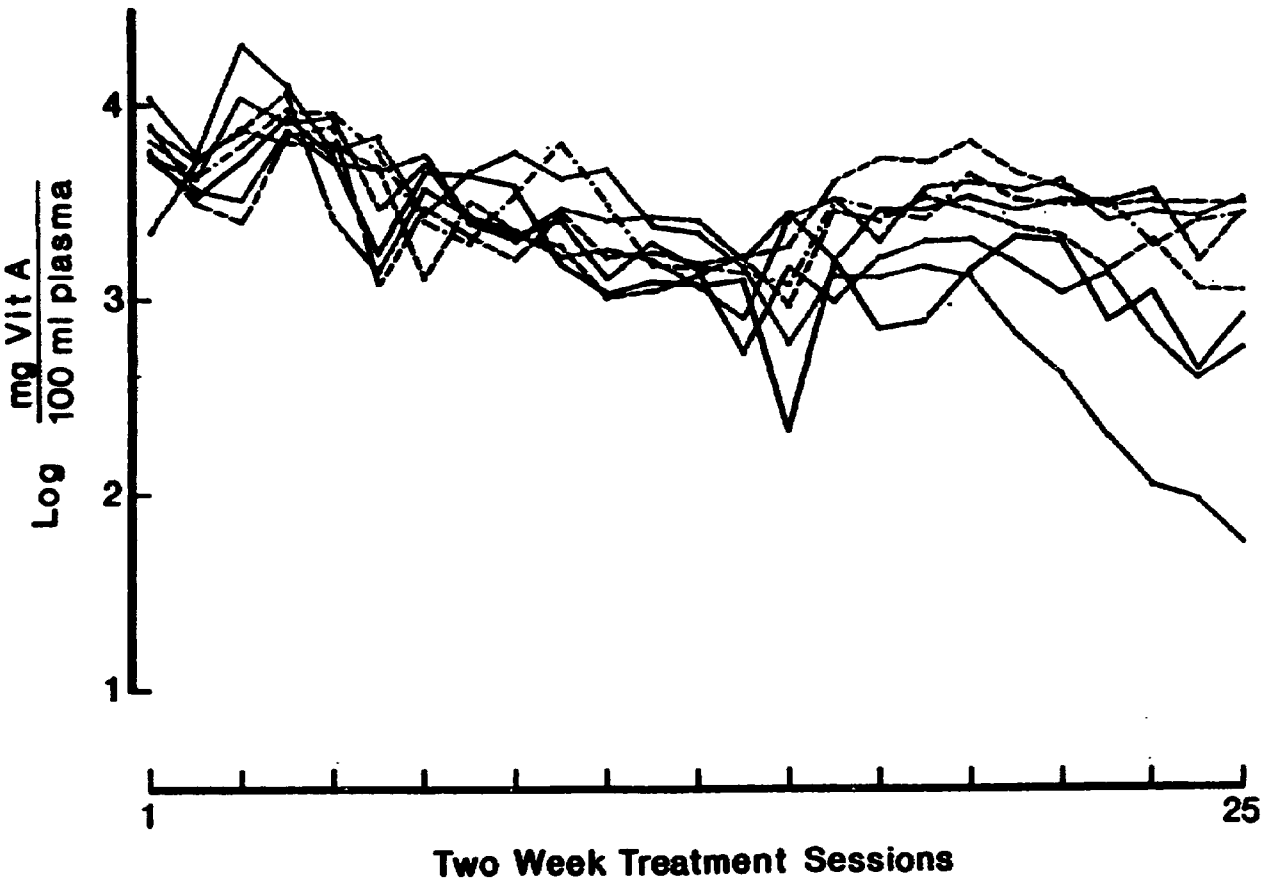
one for each volunteer. Regression equations of this form would be particularly informative if $\beta \leq 0$, $\gamma \leq 0$, for the e^{α} would be an asymptote as $x \to \infty$. It would also be interesting to test the hypothesis that all eight regression equations are in fact the same, for then the individuals of the population can be thought of as having the same depletion pattern.

FIGURE: 1a



Graph of the raw mg Vitamin A/100 ml plasma values for each subject across the 25, 2-week treatment sessions. (Mean values for the group of 8 subjects across treatment sessions are shown in the insert).

FIGURE: 1b



Graph of the Log mg Vitamin A/100 ml plasma values for each subject across the treatment sessions.

PROBLEMS WITH A COMPARTMENT MODEL FOR ASSESSING HUMAN VITAMIN A KINETICS

Richard Teplick
Captain, M.C.
Assistant Chief Department of Information Sciences
Letterman Army Institute of Research
Fitzsimons Army Medical Center
Denver, Colorado

ABSTRACT. To utilize tracer data to assess relative rates of Vitamin A utilization under dietary restriction a four-compartment model was developed. Ideally, the matrix of transition probabilities for this model could be obtained by finding its four eigenvalues and eigenvectors by fitting an equation with four exponential terms to the tracer data from each compartment. However, since data was obtained from only two compartments, two less direct techniques were utilized.

In one attempt to solve the model, the data from each of these two compartments was fit with such a multiexponential equation. The resulting two eigenvectors and four eigenvalues were used to try to solve twelve nonlinear simultaneous equations. This proved to be very complex and dependencies were eventually uncovered. A second attempt utilized the eigenvalues from the first fit to derive an equation describing the second set of data explicitly in terms of transition probabilities, volumes and two eigenvectors. The parameters of this equation were then adjusted to fit the data. This technique yielded a physically unrealizable solution. However the attempts to obtain solutions using these methods illustrate a number of problems common to this class of models.

1. INTRODUCTION. In the late 1960s eight adult males were studied for a two and one half year period to acquire information about human requirements for Vitamin A. This study included a 360 to 771 day depletion phase during which exogenous Vitamin A intake was to be eliminated. At the beginning of the depletion phase, seven of the subjects were given an intravenous tracer dose of C^{14} labeled retinol. At varying intervals the label present in plasma, plasma retinol, feces, urine and breath was determined. The original intent was to use specific activity of the tracer to calculate the body stores of Vitamin A. However, since this involves an open system the label will never equilibrate with the body stores. Thus specific activity should vary with time and cannot be used directly to calculate body stores of retinol. Therefore, an alternative method to estimate total retinol stores as a function of time was sought. The approach selected was to model the system.

2. THE MODEL. The simplest model which is in any way realistic is shown in Figure 1. The assumptions made for this model are as follows:

a. The loss or gain of retinol in any compartment follows first order kinetics. Therefore, there is no active transport.

b. Retinol exists in blood both bound to serum proteins and in a free form.

c. All compartments communicate directly with the free serum retinol compartment but not directly with each other.

d. The volumes of all compartments are fixed.

e. The bolus of labeled retinol was introduced directly into he free blood compartment.

The general form of the set of first order differential equations generated by this model is shown at the bottom of Figure 1.

These equations, written explicitly for this model in terms of concentrations rather than quantities, are seen in Figure 2A. The initial conditions are given in Figure 2B. Since the bolus of label is introduced only into the free-blood compartment (C_2); if rapid injection and mixing are assumed, the label in this compartment at zero time equals the initial dose. Additionally, at zero time, there can be no label in any of the other three compartments.

The known or measured parameters used in seeking solutions of the model are given in Figure 2C. The concentration of labeled retinol in blood was determined at varying intervals. Thus these values reflect both blood compartments. Fecal samples were collected and pooled for one week periods. Aliquots were counted and total weights measured giving a figure for label lost per week. This was converted to a one-day figure and assumed to reflect the loss at mid-week. The blood volume was determined by dilution techniques. The problem, then, is to derive the parameters of the model from the data.

3. SOLVING THE MODEL. The general form of the solution to the set of differential equations is shown in Figure 3A. This solution is derived by taking the Laplace Transform of the equations and solving them for the transform of the concentration in any compartment. This solution, shown on the right of the equal sign in Figure 3A, always has a fourth degree polynomial in S for a denominator. A second degree polynomial in S forms the numerator for all compartments except the free-blood compartment which is third degree in S. The inverse Laplace transform of this function, shown on the left in Figure 3A, is a weighted sum of four exponential terms where the a_j's are the same for all compartments and the A_{ij}'s differ.

Any attempt to solve for the defining parameters of the model will involve nonlinear curve fitting. The major differences among these approaches depend upon the available information. If data are available for the concentration or amount of label as a function of time in each of the four compartments the simplest technique is to fit the four exponential equation to the data for each compartment. Since the a_j are the same in each compartment, once they are determined for any compartment using nonlinear fitting techniques, the data for the remaining three compartments may be fit linearly. As illustrated in Figure 3B, the A_{ij} are components of the eigenvector corresponding to the eigenvalue a_j for the matrix of transition probabilities. Therefore, as illustrated, the transition probabilities can easily be calculated if (A_{ij}) is non-singular. This method could not be used for this model since a direct measure of labeled retinol was available only for the combined concentration of the two blood compartments.

It is readily shown that each A_{ij} may be written as an explicit nonlinear function of transition probabilities and volumes. The exact form can be derived from the equation in Figure 3C.

The α, β and δ may be determined by placing solutions for concentrations in Laplace space into the form given on the right in Figure 3A. The total blood concentration can be expressed as a weighted sum of four exponential terms since it is a linear combination of the two blood compartments. A similar form can be derived for fecal excretion. Therefore, the four eigenvalues, and the eigenvectors for the total blood concentration for the fecal excretion may be found by nonlinearly fitting one set of data and using the eigenvalues determined from this fit to linearly fit other set of data. The two eigenvectors and four eigenvalues thus determined can be used to derive twelve equations which are nonlinear functions of the eight transition probabilities and the four volumes. Eight of these equations are derived from the equation in Figure 3C. The other four are formed by writing the denominator of the solutions in Laplace space in the form of the denominator given on the right in Figure 3A. It appeared that there were twelve equations and twelve unknowns. However, after considerable algebraic manipulation, dependencies could be demonstrated between several of the equations. Thus it was apparent that a solution could not be derived in this fashion and this method was abandoned. Retrospectively, this is obvious as half of the eigenvectors and all of the eigenvalues are insufficient for reconstructing the original matrix.

A second approach was devised which uses nonlinear minimization techniques to find the unknown parameters using as few variables in the minimization as possible. Figure 4 illustrates this approach for the blood compartments. By using the equation relating C_3 to C_2, (Figure 4A) the general form of the solution (Figure 4B) and the

unknown constraints (Figure 4C) the total blood concentration can be written as a function of the four eigenvalues, and the bound blood eigenvectors, volume and associated transition probabiliites.

Figure 4E illustrates this equation. The eigenvalues were determined by fitting the fecal data using a nonlinear least squares approach. This data was chosen rather than labeled blood retinol data because more data points were available and because there was no blood data available for the first ten hours of the study. Initial values for the fit were estimated by peeling two exponentials from the tail of the fecal data and guessing the remaining values. It was also evident that all eigenvalues and at least one of the A_{ij} had to be negative.

The fitting routine incorporates provisions for random searches and several gradient and nongradient minimization techniques. Parameters may also be fixed while others are permitted to vary. A procedure is also available to attempt to subtract out a local minima and continue the search. All runs were on a CDC 7600. A positive definite Hessian matrix was eventually found although several parameters had to be altered manually to escape local minima. A positive definite Hessian was also obtained for the equations of Figure 4E using the eigenvalues derived from the fecal data. The derived parameters appeared reasonable. Subsequently a linear fit was performed on the total blood compartment. Using this fit and the calculated bound-blood equation, the equation for the concentration of retinol in the free blood compartment was calculated. This produced one negative component in the eigenvector which drove the calculated labeled retinol negative for the first five hours of the study and caused a maximum to be reached at about 26 hours. Because these results are physically impossible, further efforts were abandoned. However, had this not occurred, a similar technique would have been used to determine the parameters associated with the liver compartment in conjunction with the measured fecal excretion and calculated free-blood pool. The parameters for the last compartment (other tissues) would have been determined using the same technique along with the calculated label remaining in the body (assuming the percent metabolites to be insignificant) minus the quantity calculated to be in the three determined pools. The resultant eigenvectors and eigenvalues were to be used to recalculate the volumes and transition probabilities as in Figure 3B. These new values would be used as starting points for new solutions to be calculated in the same manner as the original solutions. It was hoped that eventually the fitted parameters would correspond to those calculated from the eigenvalues and eigenvectors. The reason this technique was to be employed was in the hope of circumventing the difficulties inherent in the use and interpretation of nonlinear minimization techniques.

4. PROBLEMS WITH THE TECHNIQUE. Regardless of the approach taken, some nonlinear curve fitting must be performed. If measurements are available for all compartments, only one nonlinear fit adjusting seven parameters is needed. This method was utilized to fit the fecal data. In the second approach used to solve the model this fit was used as the basis for a second nonlinear fit adjusting six additional parameters. An alternative would have been one fit for 10 parameters which would have also determined the eigenvalues. It is evident that when fewer parameters are fit more nonlinear fits are needed. This is counterbalanced by the mounting difficulties in fitting more parameters simultaneously. All fits, regardless of technique, are dependent upon the initial values chosen. In this study, only crude guesses could be made for initial values. If an improper set of initial values is used, the nonlinear fit becomes trapped in a local minimum. This was amply clear from futile attempts to better the fit using random searches; successively fixing parameters; and subtracting out local minima. Yet when a different set of starting values was selected a better fit resulted.

It is obvious from the form of the fecal data, which is initially zero; reaches a peak in about 24 hours and then monotonically decreases; that at least one of the A_{ij} had to be negative and that all of the eigenvalues must be negative. If a physical interpretation could be affixed to the eigenvectors, reasonable starting points and constraints might be derived. However, I am unable to find any such interpretion. Is it valid to assume that a fit with a positive definite Hessian matrix is the best fit? If so, should starting values be altered until a positive definite Hessian can be found. If, as in this case, the model fails to yield a reasonable solution, how can one decide if it is the model, the techniques, the data, or a combination of these factors causing the difficulty. Additionally, if a reasonable solution is obtained without finding a positive definite Hessian is it valid to use this solution and to compare the solution for one subject to another. Finally, can any form of confidence intervals be placed around the calculated parameters?

5. COMMENTS BY ONE OF THE PANELISTS. The author is grateful to Professor Jerome Cornfield for the following remarks on a possible methods for treating the troublesome problems:

I have been stimulated by this very interesting paper to re-examine [1] the application of H.O. Hartley's method of internal least squares [2] to the analysis of multi-compartment models. The major reason for investigating this method is that certain parameters of the model can be made to appear linearly, so that the computations become greatly simplified. A major consequence of this simplicity is that the issue of whether the existence of four compartments can be inferred from measurements on only two, can be investigated in a very direct way. My own working hypothesis is that only direct measurements of concentrations in the compartments in question, such as those reported by Duncan et al [3], will provide satisfactory answers. Such direct measurement is not possible when the subjects are humans, so that it seems essential to have a statistical basis for determining whether a given body of data will permit reliable indirect estimation of the parameters of unobserved compartments from direct measurement on observed ones.

<u>One-Compartment Model</u>

Starting with a one-compartment model, the differential equation, using quantities of label throughout is

$$(1) \qquad \frac{dQ}{dt} = -\lambda Q$$

Integrating both sides from time 0 to t, we have

$$(2) \qquad Q = I - \lambda \int_0^t Q\,dt \ ,$$

where I is the bolus injected.

If observations q_i are made at time t_i (i=1,2...r), estimate

$$\int_{t_{i-1}}^{t_i} Q\,dt \qquad \text{as}$$

$$(3) \qquad y_i = \frac{(t_i - t_{i-1})(q_{i-1} - q_i)}{\log_e q_{i-1} - \log_e q_i}$$

where $t_0 = 0$ and $q_0 = I$.

((3) is obtained by passing an exponential through the points at t_{i-1} and t_i and integrating.) Set

$$(4) \qquad x_i = \sum_{j=1}^{i} y_j,$$

and estimate λ as the constant in the linear model.

$$(5) \qquad q_i = I - \lambda x_i + \epsilon\ ,$$

i.e. by minimising $\sum_{i=1}^{r} (q_i - I + \lambda x_i)^2 / V_i$ (where $\mathrm{Var}(q_i) = V_i$) or some other convenient procedure. A standard error, confidence interval or other measure of uncertainty can be similarly obtained. This formulation, which is essentially due to Hartley, linearises the problem. The question of the efficiency of the estimate seems of secondary importance in this context and we shall not pursue it.

Two-compartment model

The model is defined by the equations

$$(6) \qquad \begin{aligned} \frac{dQ_1}{dt} &= -(\lambda_{12} + \lambda_{10})Q_1 + \lambda_{21}Q_2 \\ \frac{dQ_2}{dt} &= \lambda_{12}Q_1 - (\lambda_{21} + \lambda_{20})Q_2\ . \end{aligned}$$

Compartment 1 corresponds to the blood-free compartment, compartment 2 to a homogeneous all-other compartment, while λ_{10} and λ_{20} are the constants governing

the two forms of excretion, Integrating as before

$$Q_1 = I - (\lambda_{12} + \lambda_{10})\int_0^t Q_1 dt + \lambda_{21}\int_0^t Q_2 dt$$

(7)

$$Q_2 = \lambda_{12}\int_0^t Q_1 dt - (\lambda_{21} + \lambda_{20})\int_0^t Q_2 dt .$$

Observations q_{1i} of amounts in compartment 1, are made at time t_{1i} and of u_{ji}, the amount excreted from compartment j between $t_{j(i-1)}$ and t_{ji} (j=1,2). No observations are made on compartment 2. We define x_{1i} as in the one compartment model[/1] and estimate $\int_0^{t_i} Q_1 dt$ in equations (7) as x_{1i}. To estimate $\int_0^{t_i} Q_2 dt$ we let

(8) $q_{2i} = I - q_{1i}$ --cumulative amt. excreted,

where the last term is estimated from the u_{ji}. We pass over the details of this estimation as straightforward and of no intrinsic interest. We now treat the q_{2i} in the same fashion as the q_i of the one-compartment model to obtain x_{2i} as estimates of $\int_0^{t_i} Q_2 dt$. We then estimate $(\lambda_{12} + \lambda_{10})$ and λ_{21} from the linear model

(9) $$q_{1i} = I - (\lambda_{12} + \lambda_{10})x_{1i} + \lambda_{21}x_{2i} + \epsilon$$

by least squares or some other convenient procedure. λ_{10} and λ_{20} can be estimated from the linear models

(10) $$u_{ji} = \lambda_{j0}y_{ji} + \epsilon_j \qquad j = 1,2.$$

[/1] For the 2 compartment model Q_i is a linear combination of 2 exponentials and alternatively one might consider passing such a curve through the 4 points q_{i-2}, q_{i-1}, q_i and q_{i+1} and integrating between q_{i-1} and q_{-i}.

Estimates of the uncertainty of the estimated constants can be obtained, as for the one compartment case, from standard linear regression theory.

Three compartment model

The foregoing is all preliminary to a consideration of whether the observations q_{1i} and u_{ji} will support the separation of the assumed homogeneous all-other compartment into two or more. The differential equations are

$$\frac{dQ_1}{dt} = -(\lambda_{12}+\lambda_{13}+\lambda_{10})Q_1 + \lambda_{21}Q_2 + \lambda_{31}Q_3$$

$$\frac{dQ_2}{dt} = \lambda_{12}Q_1 - (\lambda_{21}+\lambda_{20})Q_2 \qquad (11)$$

$$\frac{dQ_3}{dt} = \lambda_{13}Q_1 - (\lambda_{31}+\lambda_{30})Q_3,$$

which on integration become

$$Q_1 = I-(\lambda_{12}+\lambda_{13}+\lambda_{10})\int_0^t Q_1\,dt + \lambda_{21}\int_0^t Q_2\,dt + \lambda_{31}\int_0^t Q_3\,dt$$

$$Q_2 = \lambda_{12}\int_0^t Q_1\,dt - (\lambda_{21}+\lambda_{20})\int_0^t Q_2\,dt \qquad (12)$$

$$Q_3 = \lambda_{13}\int_0^t Q_1\,dt - (\lambda_{31}+\lambda_{30})\int_0^t Q_3\,dt$$

Letting $Q_2^* = Q_2 + Q_3$ we have from the sum of the last two equations

$$Q_2^* = (\lambda_{12}+\lambda_{13})\int_0^t Q_1\,dt - (\lambda_{21}^*+\lambda_{20}^*)\int_0^t \theta_2^*\,dt, \qquad (13)$$

$$\text{where } \lambda_{21}^*+\lambda_{20}^* = \frac{(\lambda_{21}+\lambda_{20})\int_0^t Q_2\,dt + (\lambda_{31}+\lambda_{30})\int_0^t Q_3\,dt}{\int_0^t Q_2\,dt + \int_0^t Q_3\,dt}$$

Equation (13) is of the same form as the second of equations (7), except that the quantity $\lambda_{21}+\lambda_{20}$, which is constant over time in that equation

is replaced by the weighted average $\lambda^*_{21} + \lambda^*_{20}$. This quantity will change over time if $\int_0^t Q_2\,dt / \int_0^t Q_3 dt$ is not constant for all t. The ability to break the lumped all-other compartment into two or more separate compartments in the absence of direct measurements is thus entirely dependent on the non-constancy of $\lambda^*_{21} + \lambda^*_{20}$ and the inferences that can be drawn from this non-constancy.

Behavior of $\lambda^*_{21} + \lambda^*_{20}$

The behavior of the function $\lambda^*_{21} + \lambda^*_{20}$ can now be directly investigated by first fitting a two compartment model from equations (9) and (10). We then substitute observed quantities for the second of equations (7) to obtain

$$(14) \qquad q_{2i} = \lambda_{12} x_{1i} - (\lambda_{21} + \lambda_{20}) x_{2i}.$$

Using the value of λ_{12} estimated from the two compartment assumption one can compute for each t_i the quantity

$$(15) \qquad \frac{q_{2i} - \lambda_{12} x_{1i}}{x_{2i}}$$

and observe its behavior over time. If it is constant, within observational error, then there is no basis in the data for breaking down the lumped all-other compartment. This step of the investigation would benefit, however, by knowing the behavior of $\lambda^*_{21} + \lambda^*_{20}$ when the three compartment model holds. One way of proceeding would be to substitute for $\int_0^t Q_2 dt$ and $\int_0^t Q_3 dt$ in $\lambda^*_{21} + \lambda^*_{20}$ the linear combination of three exponentials obtained by the solution of equations (12). All linear properties are now lost, but there may be no simpler procedure. Under some circumstances $\lambda^*_{21} + \lambda^*_{20}$

may be monotone in t, although I have been unable to prove that this is generally true when the three compartment model holds. Although more can and should be done to investigate the behavior of $\lambda^*_{e1} + \lambda^*_{e0}$, the point of this note is that direct investigation of the behavior over time of expression (15), which depends only on simply computed observables, may provide a more perspicuous analysis than is achievable by more standard methods.

Comment

The foregoing is best interpreted as an informal memo to a colleague on a possible new line of investigation that may bypass some difficulties, but overlooking many details which must eventually be filled in. Until this investigation is undertaken, I would claim no more than that the general approach seems promising.

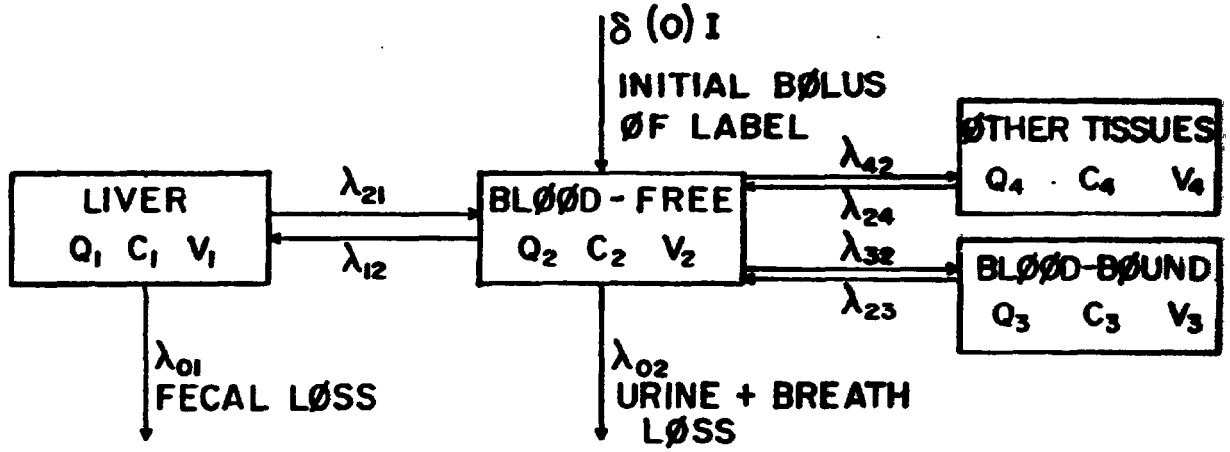
References

[1] Cornfield, J., Steinfeld, J. and Greenhouse, S.W., "Models for the Interpretation of Experiments Using Tracer Compounds", Biometrics, 16:212, 1960

[2] Hartley, H.O., The Estimation of Non-Linear Parameters by Internal Least Squares, Biometrika 35:32, 1948

[3] Duncan, L.E. Jr., Cornfield, J. and Buck, K.,"Circulation of Iodinated Albumin through Aortic and other Connective Tissues of the Rabbit", Circulation Research, VI:244, 1958

PASSIVE ØPEN FØUR CØMPARTMENT MØDEL FØR VITAMIN A



GENERAL FØRM ØF EQUATIØNS DICTATED BY MØDEL

$$\frac{d\,Q_i(t)}{dt} = \sum_{j=1}^{4} \frac{\lambda_{ij}}{V_j}\, Q_j(t) \qquad \lambda_{ii} = -\sum_{\substack{k=0 \\ k\neq i}}^{4} \lambda_{ik}$$

$$i \in [1, 4]$$

Q_i = QUANTITY ØF LABEL IN CØMPARTMENT I AS A FUNCTIØN ØF TIME.
C_i = CØNCENTRATIØN ØF LABEL IN CØMPARTMENT I AS A FUNCTIØN ØF TIME.
V_i = FIXED VØLUME ØF CØMPARTMENT i.
λ_{ij} = FIXED TRANSFER CØNSTANT FRØM CØMPARTMENT j TØ CØMPARTMENT I.

FIGURE: 1

EXPLICIT EQUATIØNS FØR VITAMIN A MØDEL

(A) EQUATIØNS GENERATED BY MØDEL

1) $V_1 \frac{dC_1}{dt} = \lambda_{12} C_2 - (\lambda_{21} + \lambda_{01}) C_1$

2) $V_2 \frac{dC_2}{dt} = \lambda_{21} C_1 + \lambda_{23} C_3 + \lambda_{24} C_4 - (\lambda_{02} + \lambda_{12} + \lambda_{32} + \lambda_{42}) C_2$

3) $V_3 \frac{dC_3}{dt} = \lambda_{32} C_2 - \lambda_{23} C_3$

4) $V_4 \frac{dC_4}{dt} = \lambda_{42} C_2 - \lambda_{24} C_4$

(B) INITIAL CØNDITIØNS

1) $Q_2(0) = I$

2) $Q_1(0) = Q_3(0) = Q_4(0) = 0$

(C) MEASURABLE QUANTITIES

1) $\frac{Q_2(t) + Q_3(t)}{V_2 + V_3}$

2) $\lambda_{01} C_1(t)$

3) $V_2 + V_3$

Ci - CØNCENTRATIØN ØF LABEL IN CØMPARTMENT I AS A FUNCTIØN ØF TIME
I - INITIAL INTRAVENØUS DØSE ØF LABEL

FIGURE: 2

FIGURE: 3

SØLUTIØNS TØ CØMPARTMENT MØDEL

A GENERAL FØRM ØF SØLUTIØNS

$$C_i(t) = \sum_{j=1}^{4} A_{ij} e^{a_j t} = \mathcal{L}^{-1}\left\{\frac{KS^3 + \alpha_i S^2 + \beta_i S + \delta_i}{\prod_{j=1}^{4} (S - a_j)}\right\}$$

B IF A_{ij} KNØWN FØR ALL i

$$[\lambda_{ij}]\,[A_{ij}] = [A_{ij}]\,[a_j\ \delta_{ij}]$$

$$[\lambda_{ij}] = [A_{ij}]\ [a_j\,\delta_{ij}]\ [A_{ij}]^{-1}$$

C IF EXPLICIT SØLUTIØNS ARE DESIRED

$$A_{ij} = \frac{Ka_j^3 + \alpha_i a_j^2 + \beta_i a_j + \delta_i}{\prod_{\substack{j=1 \\ j\neq i}}^{4} (a_i - a_j)}$$

K = 1 IF i = 2

K = 0 IF i ≠ 2

$\mathcal{L}^{-1}$ = INVERSE LAPLACE TRANSFØRM

S = CØØRDINATES IN LAPLACE SPACE

A_{ij}, a_j ARE CØNSTANTS FRØM FITTING DATA

δ_{ij} IS KRØRECKER DELTA

[ ] ARE SQUARE MATRICES

A SOLUTION FOR BLOOD COMPARTMENTS

A GIVEN THE EQUATION RELATING $C_3(t)$ AND $C_2(t)$

(1) $$V_3 \frac{dC_3(t)}{dt} = \lambda_{32} C_2(t) - \lambda_{23} C_3(t)$$

B WHERE

(1) $$C_K(t) = \sum_{j=1}^{4} A_{Kj} e^{\alpha_j t}$$

C WITH CONSTRAINTS

(1) $$V_T = V_2 + V_3$$

(2) $$C_T(t) = \frac{C_2(t) V_2 + C_3(t) V_3}{V_2 + V_3}$$

(3) $$\sum_{j=1}^{4} A_{3j} = 0$$

D AND DETERMINED OR MEASURABLE QUANTITIES

(1) V_T

(2) $C_T(t)$

E THEN $C_T(t)$ CAN BE EXPRESSED IN $V_3, \lambda_{32}, \lambda_{23}, A_{32}, A_{33}, A_{34}$

$$C_T(t) = \left(\frac{V_2 - V_3}{V_t}\right)\left\{\sum_{j=2}^{4}\left[\left(\alpha_j \frac{V_3}{\lambda_{32}} + \frac{\lambda_{23}}{\lambda_{32}} + \frac{V_3}{V_t - V_3}\right) A_{3j} e^{\alpha_j t}\right] - \left(\alpha_1 \frac{V_3}{\lambda_{32}} + \frac{\lambda_{23}}{\lambda_{32}} + \frac{V_3}{V_t - V_3}\right) e^{\alpha_1 t} \cdot \sum_{j=2}^{4} A_{3j}\right\}$$

FIGURE: 4

SAMPLE SIZE DETERMINATION AND THE OPERATING CHARACTERISTIC CURVE FOR THE CIRCULAR PROBABLE ERROR

Gideon A. Culpepper

Reliability, Availability, and Maintainability Division
Army Missile Test and Evaluation Directorate
US Army White Sands Missile Range, New Mexico 88002

ABSTRACT. An equation is developed containing the sample size required in order to meet a Circular Probable Error requirement expressed in terms of a desired and essential value and the corresponding producer and user risks. From this equation a table is produced giving the required sample size for the risk factors involved and the value of the Circular Probable Error ratio. Also, the necessary condition is given that the CPE estimate must fulfill in order to meet the criteria.

An Operating Characteristic Curve is also constructed for an example. The OC Curve is based upon the chi-square distribution with 2N-2 degrees of freedom and the ratio of the essential CPE to the true CPE.

The Circular Probable Error (CPE) is often used as a measure of accuracy in evaluating the performance of surface-to-surface missiles and rockets. Various formulas are used in estimating the CPE depending upon the characteristics of the miss distance data collected from the firing tests. The estimator for the CPE used in this paper is $\widehat{CPE} = 1.1774 s_P$, where $s_P^2 = \frac{1}{2}\left[s_X^2 + s_Y^2\right]$. The sample variance s_P^2 is the pooled or average variance of the component miss distance data.

The problem considered here is to determine first, the sample size needed in order to meet the CPE criteria given in requirements documents and second, to find what the condition is that the estimated CPE must fulfill to satisfy these requirements. The CPE criteria are assumed to be given in the following form:

CPE_E = the essential or minimum acceptable CPE

CPE_D = the desired or specified CPE

$CPE_E > CPE_D$

The following risk factors are either given or assumed for the purpose of test planning:

β = the user's probability of the missile passing the accuracy test if the true CPE of the missile is greater than CPE_E.

α = the producer's probability of the missile failing this test if the true CPE of the missile is equal to CPE_D. $100(1-\beta)\%$ is interpreted as the confidence that the true CPE $\leq CPE_E$ if $\widehat{CPE}$ satisfies a certain condition (see paragraph on next page). The parameters

$\{CPE_E, \beta; CPE_D, \alpha\}$ determine the sample size (N) of missiles to be fired in an accuracy demonstration test.

After the miss distance data have been collected, the CPE is estimated by the formula $\widehat{CPE} = 1.1774 s_p$. If $\widehat{CPE} \leq K(CPE_E)$, where $K = \left[\chi^2_\beta(2N-2)/2N-2\right]^{\frac{1}{2}}$, then the missile has met the accuracy criteria, that is, there is a demonstrated $100(1-\beta)\%$ confidence that the true $CPE \leq CPE_E$ and if the true $CPE = CPE_D$, there is a probability of $100(1-\alpha)\%$ that the missile will be accepted.

The partial table on the next page lists the α and β risks, the ratio CPE_E/CPE_D, N, and the value of K. In forming this table N, α, and β were selected first and the ratio CPE_E/CPE_D was computed last. The equation

$$\frac{\chi^2_\alpha(2N-2)}{\chi^2_{1-\beta}(2N-2)} = \left(\frac{CPE_E}{CPE_D}\right)^2$$

connects the quantities. The upper percentage points of chi-square are used with 2N-2 degrees of freedom.

Numerical Example.

$$CPE_E = 300 \text{ meters}, \quad CPE_D = 250 \text{ meters}$$

$$\beta = 0.20 \qquad \alpha = 0.50$$

$$\frac{CPE_E}{CPE_D} = \frac{300}{250} = 1.20. \quad N = 7 \text{ and } \widehat{CPE} \leq 0.8067(CPE_E)$$

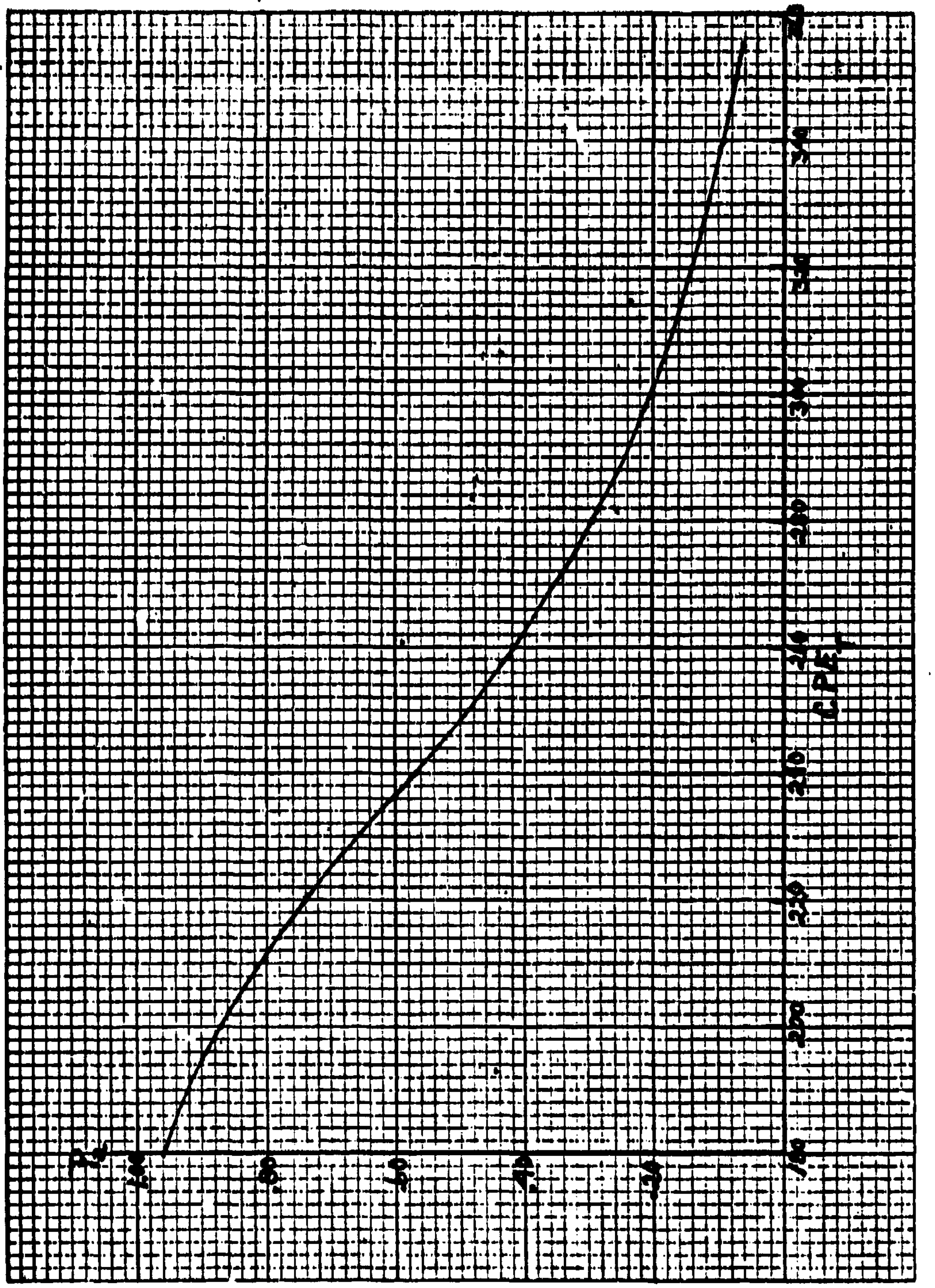
	$\frac{CPE_E}{CPE_D}$	N	K such that if $CPE \leq K(CPE_E)$ the accuracy criteria have been met.
$\alpha = .50$	1.451	5	.6605
	1.385	6	.6975
$\beta = .10$	1.342	7	.7248
	1.309	8	.7459
	1.284	9	.7629
	1.262	10	.7769
	1.247	11	.7888
	1.225	12	.7987
$\alpha = .50$	1.264	5	.7575
	1.229	6	.7861
$\beta = .20$	1.205	7	.8067
	1.187	8	.8224
	1.173	9	.8367
	1.162	10	.8466
$\alpha = .25$	1.365	6	.8208
	1.326	7	.8386
$\beta = .25$	1.298	8	.8521
	1.275	9	.8628
	1.257	10	.8716
	1.242	11	.8790
	1.229	12	.8853
	1.218	13	.8906
	1.208	14	.8953
	1.200	15	.8995
	1.192	16	.9033

If an OC Curve is desired for this design, then the probability that the missile will pass this accuracy test (be accepted by the test) is given by

$$P_a = \text{Prob}\left[\chi^2(2N-2) \leq \chi^2_\beta(2N-2)\left(\frac{CPE_E}{CPE_T}\right)^2\right]$$

where CPE_T = the true CPE. If N = 7, β = 0.20, the coordinates of the OC Curve are given in the table below. $\chi^2_{.20}(12) = 7.81$. The graphed OC Curve follows.

CPE_T	$(7.81) \times \left(\frac{CPE_E}{CPE_T}\right)^2$	$P_a \simeq$
180	21.79	.96
200	17.57	.87
220	14.52	.74
240	12.20	.57
260	10.40	.42
280	8.96	.29
300	7.81	.20
320	6.86	.14
340	6.08	.09
350	5.73	.07



DESIGN AND ANALYSIS OF A HIT PROBABILITY EXPERIMENT BASED ON A BIVARIATE NORMAL DISTRIBUTION

Robert W. Mai, Data Reduction Section, Technical Test Support Division, Materiel Test Directorate, Yuma Proving Ground, Yuma, Arizona

ABSTRACT

The paper describes a method for analyzing the results of a hit probability experiment in such a way that the producer's risk (α risk) of rejecting a population of items that does in fact meet a specified criterion is controlled. The procedure described is an adaptation of the methods ordinarily used within AMC when a binomial distribution or exponential distribution is the model.

The method starts with a population known to have a hit probability equal to the specification. A procedure is defined for getting a point estimate of hit probability from a sample. A rejection point for the point estimate from a given sample is then established at a level that will force the α risk to be within a given interval.

This method will provide information useful in evaluating the acceptability of a test item's hit probability when a bivariate normal distribution is indicated. It will also help to determine a sample size which will afford desired protection against making a Type I error. Investigation is underway to determine interval estimates for the risk of a Type II error (i.e. to provide quantitative control of the β risk using similar techniques.)

INTRODUCTION

This paper is concerned with the estimation of producer's risk α for tests of hit probability where a bivariate normal distribution is indicated and where the miss distances in the horizontal (x) and vertical (y) directions are independent. The α risk had to be quantified so that the project engineer could design the test and choose an analytical plan that would follow latest Army guidance for this phase of testing. Guidance has been: As a result of reliability growth during the test cycle, it is imperative that agencies insure against rejecting the test item early in the test cycle. This can be accomplished by insuring a high progability of accepting a population of test items if the true hit probability, P_H, is greater than or equal to the specified hit probability, P_{SPEC}.

APPROACH TO THE PROBLEM

The approach taken in this investigation was:

(1) Find a hit probability estimator, $\hat{P}_H$, that could be easily measured and/or computed during the test.

(2) A rejection point, P_{RP}, must be computed and $\hat{P}_H$ compared to this P_{RP}.

(3) If we set up the hypothesis that

$$\hat{P}_H \leq P_{RP} \text{ implies } P_H < P_{SPEC}$$

$$\hat{P}_H > P_{RP} \text{ implies } P_H \geq P_{SPEC}$$

then we must be able to choose P_{RP} such that the risk of a Type I error is less than or equal to α.

INVESTIGATION PROCEDURES AND RESULTS

The physical set up of the test hints strongly at the choice for $\hat{P}_H$ (the hit probability estimator). The target was to be rectangular (or square) with height H and length L. Previous test results had shown miss distances in x and y directions to be independent and normally distributed. Also the system could be boresighted to assure

$$\mu_x = 0 = \mu_y$$

For this set up we know that the hit probability could be computed as the product of the probability of hitting in x and the probability of hitting in y. Also we know that the probability of hitting in the x or y direction can be computed based on the normal distribution.

An obvious choice for the hit probability estimator $\hat{P}_H$ is the one described by the following equations:

$$\hat{P}_H = \hat{P}_x \hat{P}_y$$

$$\hat{P}_x = A(Z_{x,1}) + A(Z_{x,2})$$

$$\hat{P}_y = A(Z_{y,1}) + A(Z_{y,2})$$

$$Z_{x,1} = \frac{\bar{x} - x_1}{\hat{\sigma}_x}$$

$$Z_{x,2} = \frac{x_2 - \bar{x}}{\hat{\sigma}_x}$$

$$Z_{y,1} = \frac{\bar{y} - y_1}{\hat{\sigma}_y}$$

$$Z_{y,2} = \frac{y_2 - \bar{y}}{\hat{\sigma}_y}$$

where x_1: x coordinate of left edge of target

x_2: x coordinate of right edge of target

y_1: y coordinate of lower edge of target

y_2: y coordinate of upper edge of target

$\bar{x}$: average x coordinate of sample impacts

$\bar{y}$: average y coordinate of sample impacts

$\hat{\sigma}_x$: sample standard deviation in x direction

$\hat{\sigma}_y$: sample standard deviation in y direction

Z : standard normal variable with mean 0 and variance 1

A(Z): Probability that the absolute value of a standard normal variable will not exceed Z

$\hat{P}_x$: Sample estimate of the probability that the x coordinate of an impact will be within the bounds of the target

$\hat{P}_y$: Sample estimate of the probability that the y coordinate of an impact will be within the bounds of the target

The next step in the investigation is to find rejection points P_{RP} for $\hat{P}_H$ in such a way that if

$$P_H = P_{SPEC}$$

then in a sample of size N

$$(*) \qquad P(\hat{P}_H \leq P_{RP}) = \alpha \ .$$

We see from the definition of P_H that it is computed from the sample parameters $\bar{x}$, $\bar{y}$, $\hat{\sigma}_x$, and $\hat{\sigma}_y$. Thus to solve equation (*) for P_{RP}, we would need to study the distributions of $\bar{x}$, $\bar{y}$, $\hat{\sigma}_x$, and $\hat{\sigma}_y$.

The sample parameters $\bar{x}$ and $\bar{y}$ are normally distributed with means μ_x and μ_y respectively, and standard deviations $\sigma_x / \sqrt{N}$ and $\sigma_y / \sqrt{N}$, respectively [1]. The parameters $\hat{\sigma}_x$ and $\hat{\sigma}_y$ are each related to chi-squared with N-1 degrees of freedom [1]:

$$\chi^2(N-1) = \frac{(N-1)\,\hat{\sigma}_x^2}{\sigma_x^2}$$

$$\chi^2(N-1) = \frac{(N-1)\,\hat{\sigma}_y^2}{\sigma_y^2}$$

To investigate P_{RP}, we need to know how $\bar{x}$, $\bar{y}$, $\hat{\sigma}_x$ and $\hat{\sigma}_y$ behave, which in turn depends on μ_x, μ_y, σ_x, σ_y and the normal and chi-square distributions. We know how to approximate the distributions and we know that

$$\mu_x = 0 = \mu_y \ .$$

The remaining unknowns are σ_x and σ_y. Recall that we are trying to solve equation (*) under the assumption that

$$P_H = P_{SPEC} \ .$$

We also know that

$$P_H = P_x \, P_y$$

$$P_x = A(Z_{x,1}) + A(Z_{x,2})$$

$$P_y = A(Z_{y,1}) + A(Z_{y,2})$$

$$Z_{x,1} = \frac{-L/2 - \mu_x}{\sigma_x}$$

$$Z_{x,2} = \frac{L/2 - \mu_x}{\sigma_x}$$

$$Z_{y,1} = \frac{-H/2 - \mu_y}{\sigma_y}$$

$$Z_{y,2} = \frac{H/2 - \mu_y}{\sigma_y}$$

But $\mu_x = 0$ implies

$$A(Z_{x,1}) = A(Z_{x,2})$$

and

$$P_x = 2A(Z_{x,1}) \quad .$$

Similarly,

$$P_y = 2\ A(Z_{y,1}) \ .$$

In this test the size of target was such that

$$\sigma_x = a\ \sigma_y \quad \text{where}$$

$$L = a\ H \ .$$

With this simplification we have

$$P_x = P_y$$

and

$$P_H = P_x^2 = P_y^2 = P_{SPEC} \ .$$

Thus if $P_{SPEC} = .90$ we have

$$P_H = P_x^2 = P_y^2 = .90$$

$$P_x = .94868 = P_y$$

$$A(Z_x) = .47434 = A(Z_y)$$

$$\sigma_x = 1.92426 = \sigma_y.$$

If the assumption that $\sigma_x = a\,\sigma_y$ when $L = a\,H$ is not valid then iterative numerical techniques are required to solve for σ_x and σ_y given a value for $P_{SPEC} = P_H$.

We now have all the information required to solve equation (*). We let

$$S = \{ (a,b,c,d) \,/\, P(\hat{\sigma}_x \geq a,\ \hat{\sigma}_y \geq b,\ |\bar{x}| \geq c,\ |\bar{y}| \geq d) = \alpha \}$$

We then compute

$$\hat{P}_H(a,b,c,d),$$

for every $(a,b,c,d) \in S$ by

substituting $a = \hat{\sigma}_x$

$b = \hat{\sigma}_y$

$c = \bar{x}$

$d = \bar{y}$

into the equations for $\hat{P}_H$.

Then we let

$$P_{RP} = \text{Min}\{ \hat{P}_H(a,b,c,d) \,/\, (a,b,c,d) \in S \}.$$

If we do the above, then

$$P_H = P_{SPEC}$$

implies

$$P(\hat{P}_H \leq P_{RP}) \leq \alpha .$$

To find the set S, we observe first due to independence of x and y that

$$P(\hat{\sigma}_x \geq a,\ \hat{\sigma}_y \geq b,\ |\bar{x}| \geq c,\ |\bar{y}| \geq d)$$

$$= P(\hat{\sigma}_x \geq a)\ P(\hat{\sigma}_y \geq b)\ P(|\bar{x}| \geq c)\ P(|\bar{y}| \geq d).$$

We now compute values for (a,b,c,d) associated with the probabilities

$$a_{1,x} = P(\hat{\sigma}_x \geq a)$$

$$a_{1,y} = P(\hat{\sigma}_y \geq b)$$

$$a_{2,x} = P(|\bar{x}| \geq c)$$

$$a_{2,y} = P(|\bar{y}| \geq d) .$$

For ease of notation, we let

$$a_1 = a_{1,x}\ a_{1,y}$$

$$a_2 = a_{2,x}\ a_{2,y} ,$$

so that we are looking for (a,b,c,d) such that

$$\alpha = a_1\ a_2 .$$

An example of the computation required is given in table 1 for $P_{SPEC} = .90 = P_H$, $\alpha = .10$, $N = 38$.

The observations made during computation were:

(1) For a given a_1, the minimum $\hat{P}_H(a,b,c,d)$ occurs where $a_{1,x} = a_{1,y} = \sqrt{a_1}$.

(2) For a given value of a_2, the minimum value for $\hat{P}_H(a,b,c,d)$ occurs when $a_{2,x} = a_2$ or $a_{2,y} = a_2$ depending on whether $a_{1,y} > a_{1,x}$ or $a_{1,y} < a_{1,x}$, respectively.

(3) For a given value of α, the minimum $\hat{P}_H(a,b,c,d)$ occurs where $a_2 = 1.00$ and $a_1 = \alpha$ and (from (1)) where $a_{1,x} = \sqrt{a_1} = a_{1,y}$.

Thus P_{RP} can be computed for a given N, α, P_{SPEC} by making one set of computations for $\hat{P}_H(a,b,c,d)$ for the (a,b,c,d) associated with the probabilities

$$a_{1,x} = \alpha = a_{1,y}$$

$$a_{2,y} = 1.0 = a_{2,x} .$$

Examples of various rejection points and associated α risks are shown in figures 1 - 4. In figure 1, for example, where the specification hit probability was .90 and the sample size was 38, one can choose any desired α risk for a test by simply finding the rejection point P_{RP} associated with α. Then compute $\hat{P}_H$ for the sample and reject it if $\hat{P}_H \leq P_{RP}$ and accepting it if $\hat{P}_H > P_{RP}$.

ACKNOWLEDGEMENTS:

I wish to express my gratitude to Don Nelson for his very conscientious efforts in the seemingly endless calculations required to discover and substantiate the methods described herein.

It has been brought to my attention that Professor Lieberman (Stanford University) has done extensive work in this area and arrived at results more comprehensive than those in this investigation.

REFERENCES

[1] Ostle, Bernard, Statistics in Research, The Iowa State University Press, Ames Iowa, 1963.

[2] Huntsberger, David V., Elements of Statistical Inference, Allyn and Bacon, Boston, Massachusetts, 1967.

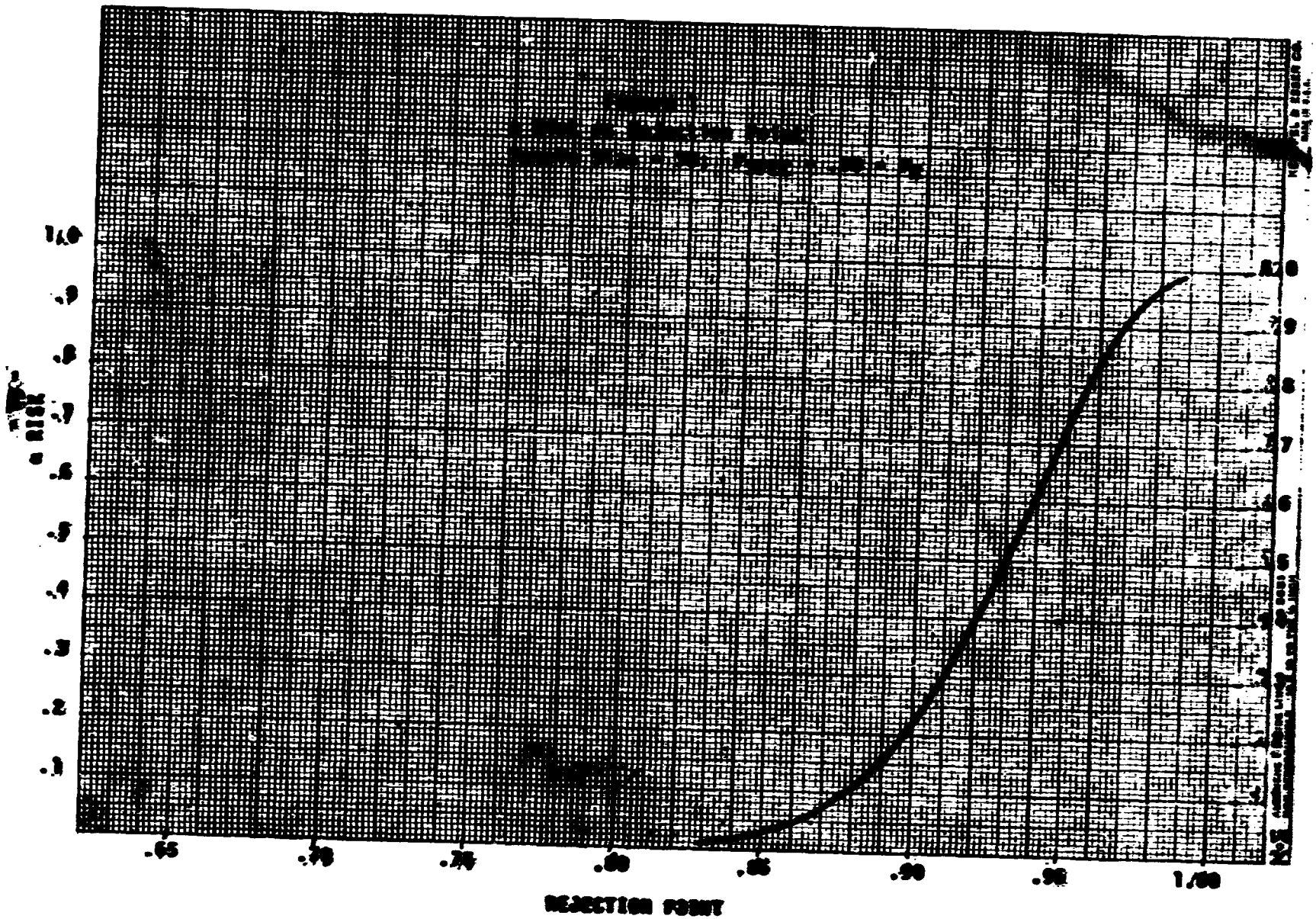
TABLE 1

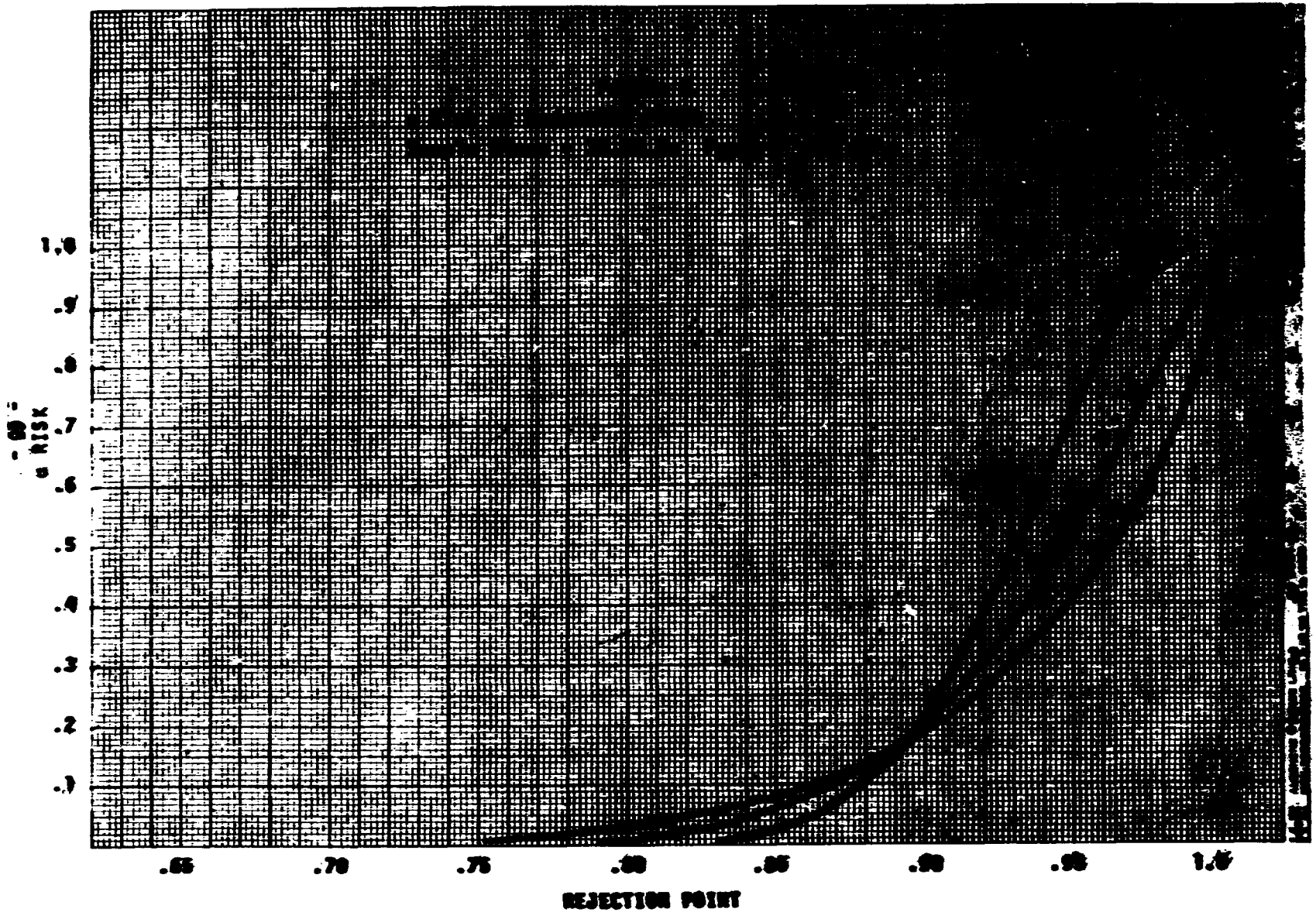
$\alpha = .10$ $P_{SPEC} = .90 = P_H$ SAMPLE SIZE = 38

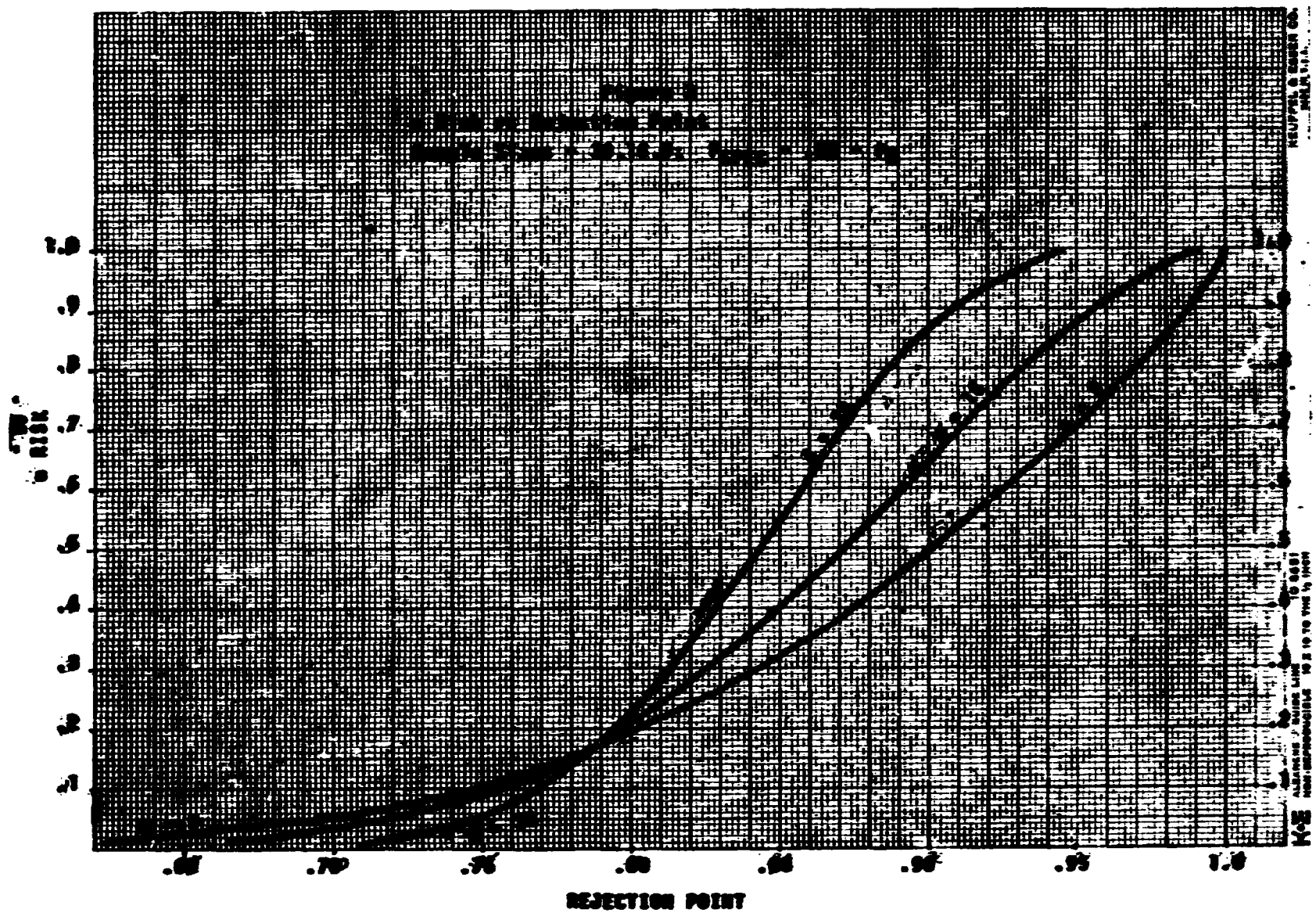
$a_{1,x}$	$a_{1,y}$	$a_{2,x}$	$a_{2,y}$	$\hat{P}_H$
$a_1 = a_{1,x}\ a_{1,y} = .10$,		$a_2 = a_{2,x}\ a_{2,y} = 1.00$		
.1000	1.0000	1.0000	1.0000	.9118
.2000	.5000	"	"	.8806
.3000	.3333	"	"	.8786
.3162	.3162	"	"	.8785
.3500	.2857	"	"	.8787
.4000	.2500	"	"	.8791
.5000	.2000	"	"	.8806
.6000	.1667	"	"	.8828
1.0000	.1000	1.0000	1.0000	.9118

$a_{1,x}$	$a_{1,y}$	$a_{2,x}$	$a_{2,y}$	$\hat{P}_H$
$a_1 = a_{1,x}\, a_{1,y} = .20$		$a_2 = a_{2,x}\, a_{2,y} = .50$		
.2000	1.0000	.5000	1.0000	.9245
"	"	.6000	.8333	.9253
"	"	.7071	.7071	.9257
"	"	.8333	.6000	.9258
.2000	1.0000	1.0000	.5000	.9262
.3000	.6667	.5000	1.0000	.8987
"	"	.6000	.8333	.8992
"	"	.7071	.7071	.8992
"	"	.8333	.6000	.8991
.3000	.6667	1.0000	.5000	.8988
.4162	.4162	.5000	1.0000	.8917
"	"	.6000	.8333	.8922
"	"	.7071	.7071	.8923
"	"	.8333	.6000	.8922
.4162	.4162	1.0000	.5000	.8917
.5000	.4000	.5000	1.0000	.8960
"	"	.6000	.8333	.8965
"	"	.7071	.7071	.8966
"	"	.8333	.6000	.8965
.5000	.4000	1.0000	.5000	.8960
.6667	.3000	.5000	1.0000	.8988
"	"	.6000	.8333	.8991
"	"	.7071	.7071	.8992
"	"	.8333	.6000	.8992
.6667	.3000	1.0000	.5000	.8987
1.0000	.2000	.5000	1.0000	.9262
"	"	.6000	.8333	.9258
"	"	.7071	.7071	.9257
"	"	.8333	.6000	.9253
1.0000	.2000	1.0000	.5000	.9245

$a_{1,x}$	$a_{1,y}$	$a_{2,x}$	$a_{2,y}$	$\hat{P}_H$
$a_1 = a_{1,x}$	$a_{1,y} = .3162$	$a_2 = a_{2,x}$	$a_{2,y} = .3162$	
.3162	1.0000	.3162	1.0000	.9342
.4000	.7905	.3162	1.0000	.9124
.5000	.6324	.3162	1.0000	.9092
.5623	.5623	.3162	1.0000	.9088
.6324	.5000	1.0000	.3162	.9092
.7905	.4000	1.0000	.3162	.9124
1.0000	.3162	1.0000	.3162	.9342
$a_1 = a_{1,x}$	$a_{1,y} = .4000$	$a_2 = a_{2,x}$	$a_{2,y} = .2500$	
.6325	.6325	.2500	1.0000	.9167
$a_1 = a_{1,x}$	$a_{1,y} = .5000$	$a_2 = a_{2,x}$	$a_{2,y} = .2000$	
.7071	.7071	.2000	1.0000	.9251
$a_1 = a_{1,x}$	$a_{1,y} = .7000$	$a_2 = a_{2,x}$	$a_{2,y} = .1429$	
.8367	.8367	.1429	1.0000	.9419
$a = a_{,x}$	$a_{,y} = 1.0000$	$a = a_{,x}$	$a_{,y} = .1000$	
1.0000	1.0000	.1000	1.0000	1.0000


REJECTION POINT


1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
.65
.70
.75
.80
.85
.90
.95
1.0
REJECTION POINT


1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
.70
.80
.90
1.0
REJECTION POINT

MODEL FOR PROBABILITY HIT ANALYSIS OF 20MM PROJECTILES

Diana L. Frederick
Frankford Arsenal
Philadelphia, Pennsylvania

ABSTRACT. An analytical determination of probability of hit (PH) for ammunition fired from a helicopter or fixed wing aircraft was calculated by a model utilizing Hastings Error Function. The model can incorporate errors for various slant ranges and target sizes. Primarily, the model provided trajectory comparisons and probability of hit data for a proposed 20mm 1950 grain HE projectile with an M505 fuse nose contour and the recently developed 20mm PGU 3/B Semi-Armor Piercing Incendiary (SAPI) projectile for mixed firings at 10 x 10 and 100 x 100 feet targets.

1. INTRODUCTION. This analytical model provides a means of comparing mismatched ammunition when fired in the same burst and is particularly useful; for example, in comparing kinetic projectiles with high explosive projectiles. However, specific data on these projectiles for a fixed set of conditions were not immediately available; therefore, the model used existing data for ammunition under development. The model represents an input to the Systems Analysis task for ACTS.

The purpose of this model was to provide a tool for determining and comparing probability of hit data for a proposed 20mm 1950 grain HE projectile, Figure 1, and the newly developed Armor Piercing Incendiary (AP) ammunition, Figure 2, when used in combination in the M61 gun. Since the firing system allows only one ammunition type to be entered in the computer system of the weapon and since the weapon could fire mismatched ammunition in the same burst, the question arose: How does the limitation on the weapon system and the variation in projectile design effect the probability of hit?

A summary of projectile characteristics follows in Table I.

TABLE I

Ammunition Characteristics

Projectile	Length (in)	Weight (grs)	Muzzle Velocity (fps)	Nose Length (in)
HE	3.600	1950	2950	1.440
AP	3.125	2100	2850	1.462

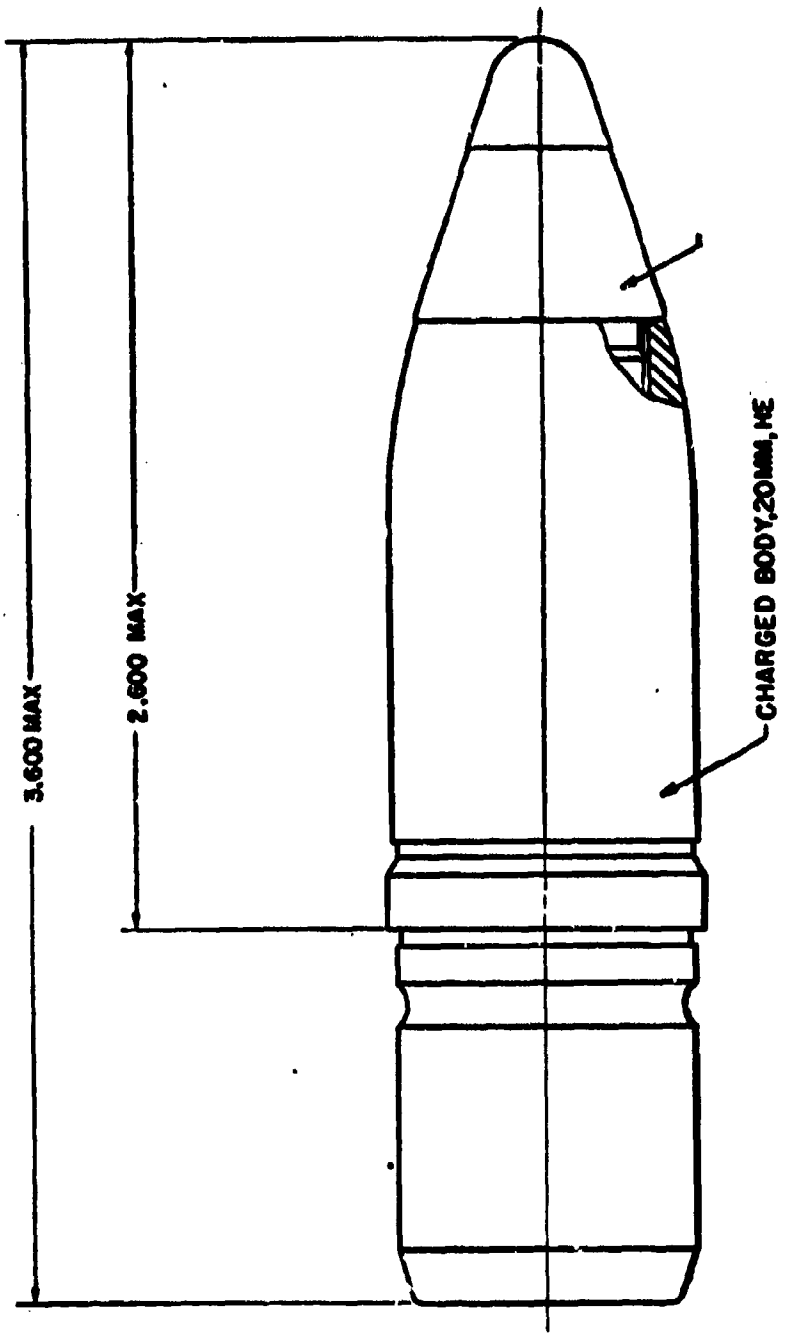
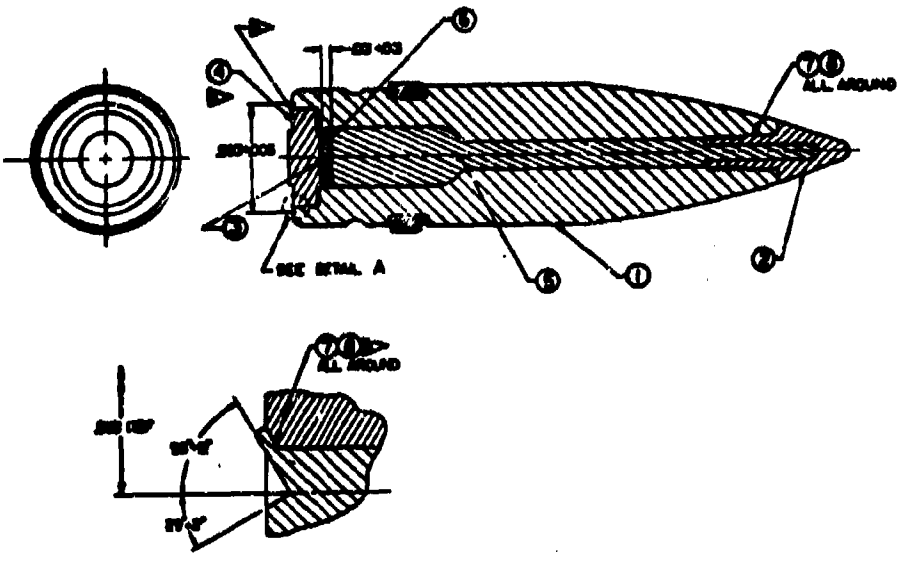
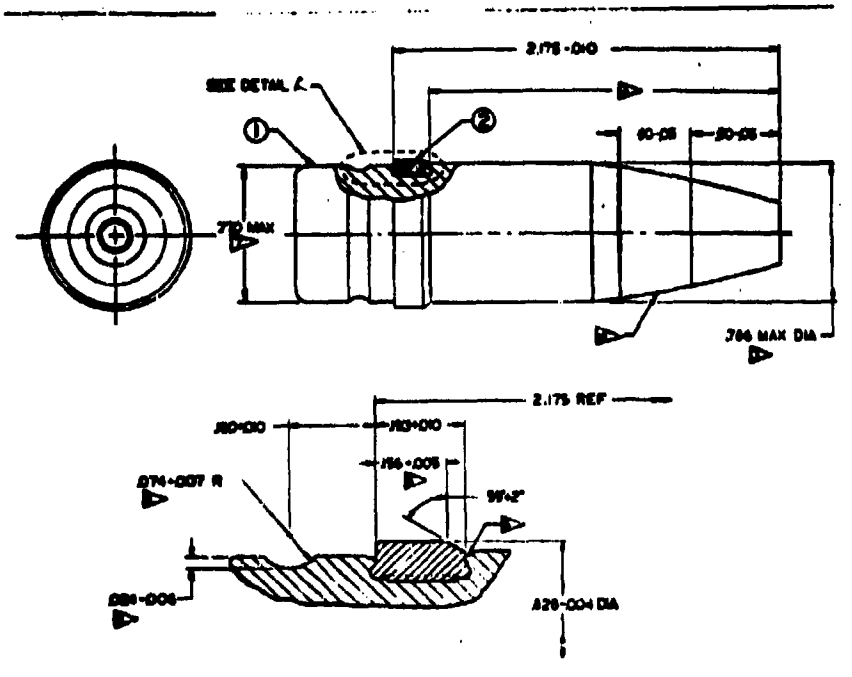


Figure 1. 20 mm Mod 1 HE Projectile



a. Projectile Assembly



b. Body and Rotating Band Assembly

Figure 2. 20 mm AP Projectile

The probability of hit calculations require estimates for the following error values:

1. Trajectory dispersion caused by variation in ammunition parameters such as drag weight and muzzle velocity,

2. Projectile precision,

3. System errors caused by the weapon mount type and movement, functional mechanism, and ammunition interaction.

2. PROCEDURE. The computer program was set up for use through card data input or data input on a teletype. The output from DIANA PH (name of program) is the probability of hit for a specified slant range, and target size for single and multiple rounds. Data inputs to the program are target size, slant range and the error estimates. All distributions considered were normal.

$$P_h = Px\ Py$$

where

P_h = probability of hit
Px = target length probability
Py = target width probability

$$Px = \text{Prob}\left(\frac{-Hx - \bar{X}}{\sqrt{2}\,\hat{\sigma}} < x < \frac{Hx - \bar{X}}{\sqrt{2}\,\hat{\sigma}}\right)$$

$$Py = \text{Prob}\left(\frac{-Hy - \bar{Y}}{\sqrt{2}\,\hat{\sigma}} < y < \frac{Hy - \bar{Y}}{\sqrt{2}\,\hat{\sigma}}\right)$$

where

X = continuous random variable of dispersion in x direction

$\bar{X}$ = mean dispersion in x direction

Hx = 1/2 target width

Y = continuous random variable of dispersion in y direction

$\bar{Y}$ = mean dispersion in y direction

Hy = 1/2 target height

SR = slant range

Ai = all errors, mils

$$\bar{\sigma} = \sqrt{\sum Ai^2}$$

$$\hat{\sigma} = \bar{\sigma}(SR/1000)$$

Px and Py were evaluated from application of Hastings Error Approximation, that is if

$$Px = \frac{2}{\sqrt{\pi}} \int_0^x e^{-tx^2} dtx$$

where $$tx = \frac{Hx}{\sqrt{2}\ \hat{\sigma}}$$

and $$Py = \frac{2}{\sqrt{\pi}} \int_0^y e^{-ty^2} dty$$

where $$ty = \frac{Hy}{\sqrt{2}\ \hat{\sigma}}$$

then the approximations are:

$$Px \approx 1 - \frac{1}{\left[1 + a_1 \bar{x} + a_2 \bar{x}^2 + a_3 \bar{x}^3 + a_4 \bar{x}^4 + a_5 \bar{x}^5 + a_6 \bar{x}^6\right]^{16}}$$

and

$$P\bar{y} \approx 1 - \frac{1}{\left[1 + a_1 \bar{y} + a_2 \bar{y}^2 + a_3 \bar{y}^3 + a_4 \bar{y}^4 + a_5 \bar{y}^5 + a_6 \bar{y}^6\right]^{16}}.$$

with:

$a_1 = .0705,2307,84$
$a_2 = .0422,8201,23$
$a_3 = .0092,7052,72$
$a_4 = .0001,5201,43$
$a_5 = .0002,7656,72$
$a_6 = .0000,4306,38$

Also for multiple rounds

$$Pn = 1 - (1 - P)^n$$

Symbols used in the computer program are as follows:

SR	-	Slant range
BML	-	Ballistic error in mils
AML	-	Ammunition mismatch error in mils
REML	-	Range estimate error in mils
SHAPE	-	Projectile design name
Hx	-	One-half target length
Hy	-	One-half target width
IND	-	Index
Px	-	Target length probability
Py	-	Target width probability
P	-	Probability of hit
Pn	-	Probability of hit of n rounds

The program is described in the flow chart in Figure 3. It is set up to do multiple runs when IND is equal to zero. The program will perform probability of hit calculations for various errors using the same target size and projectile shape. If IND is not equal to zero the program can be terminated, the target size can be changed, or the projectile can be changed.

TABLE II

Data Inputs

Card Type	Columns	Format	Explanation
1	2 - 24	4A6	Projectile shape
2	1 - 10	F10.2	One-half target width
	11 - 20	F10.2	One-half target length
3	1 - 10	F10.2	BML
3	11 - 20	F10.2	AML
	21 - 30	F10.2	REML
	31 - 40	F10.2	SR
	41 - 45	I5	IND

The fortran program, sample data input and sample data output follow in the Appendices.

DISCUSSION

An example of an application of DIANA PH is the probability of hit analysis performed on two 20mm projectiles using trajectory data generated

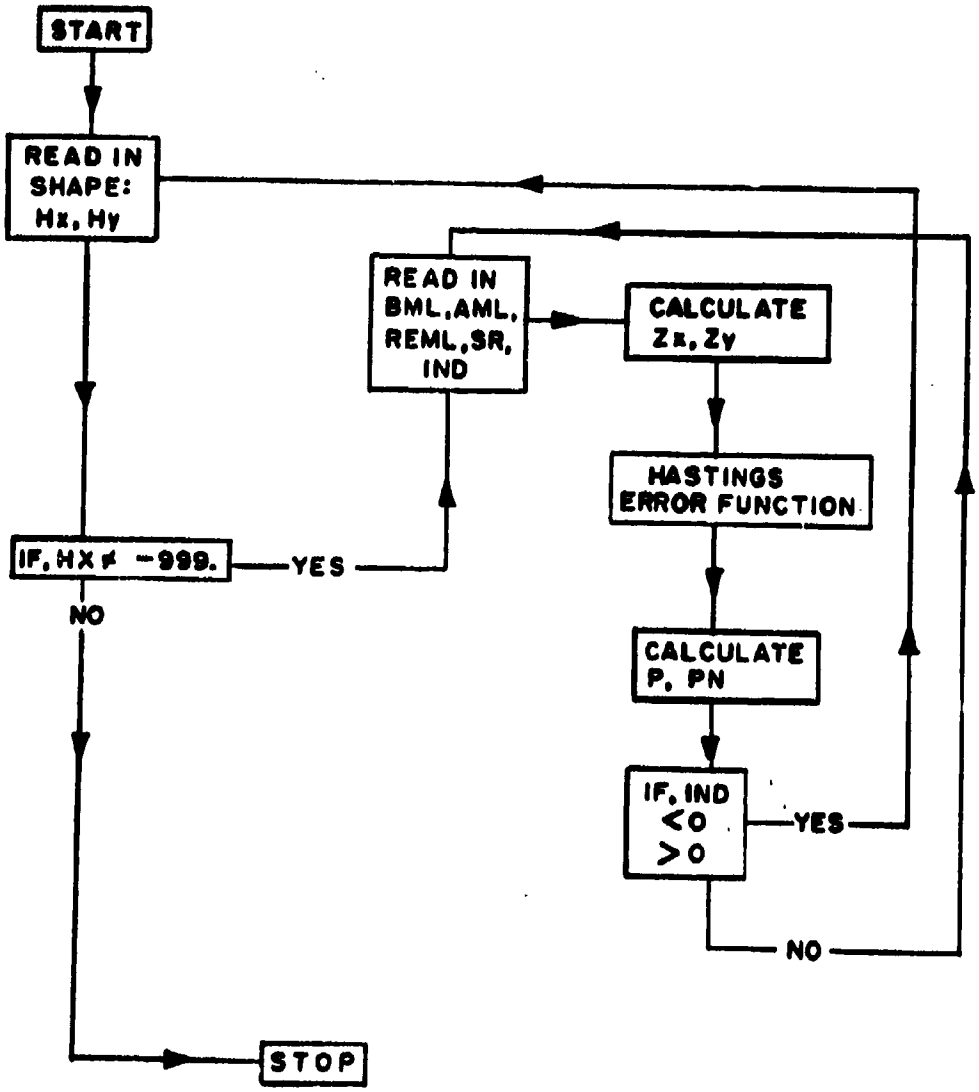


Figure 3. Flow Chart for P_{11} Program

at Frankford Arsenal. Figure 4 contains the drag curves for these two projectiles: the HE and AP. The lower drag of the AP projectile results from its longer and more slender nose shape. Thus, the AP projectile travels further with a shorter time of flight, see Figure 5.

Mismatch is a function of range, muzzle velocity, altitude upon firing, and aircraft dive angle and speed. Mismatch was calculated from the difference between the trajectories of the two projectiles. This study concerned variation of only two errors: system dispersion, BML, and trajectory dispersion, AML. Range estimate dispersion, REML, remained zero. The study covered the following four cases:

TABLE III

Variation of AML and BML Errors (Probability of Hit)

Case I - BML = 1 for both projectiles,
AML = 0 for AP
AML = X(i) for HE
X(i)= mil dispersion between trajectories of the two projectiles at the specified slant range. In this case the value is taken to be 3.47 at a range of 2309 feet.

Case II - Same as Case I except range was 6928 feet and X(i) was 13.7.

Case III- Same as Case II except BML = 4 for both projectiles.

Case IV - BML = 1 for HE
BML = 2 for AP
X(i)= 18.3 at a range of 10,381 feet

Table IV contains tabulations of the probability of hit comparisons between the HE and AP projectiles for single and multiple round bursts. The values came from exercise of the computer program with the input data as shown in the table plus that shown in Table III. The program has incorporated the drag curve for the AP and HE projectiles. From the resultant trajectory mismatch between the two projectiles, the program computes the probability of hit for both types of projectiles but with the AP being the reference. The trajectory mismatch is shown in Figure 6.

The firing system is programmed for one ammunition system only. This report assumes the projectile drag curve programmed is the AP. Thus the AML error is zero for this round and some value Xi for HE.

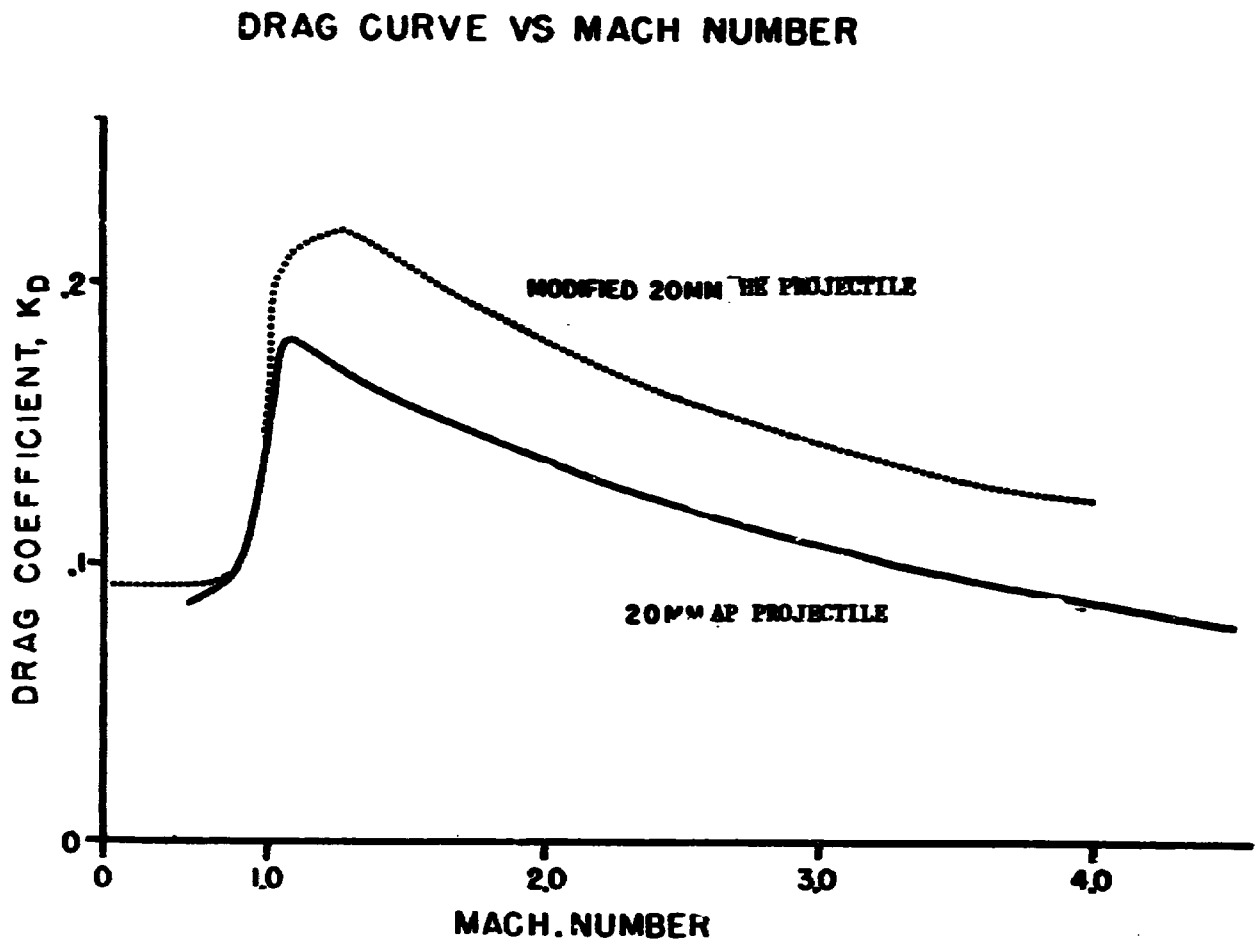


Figure 4. Drag Curves for PGU 3/B and Mod M56

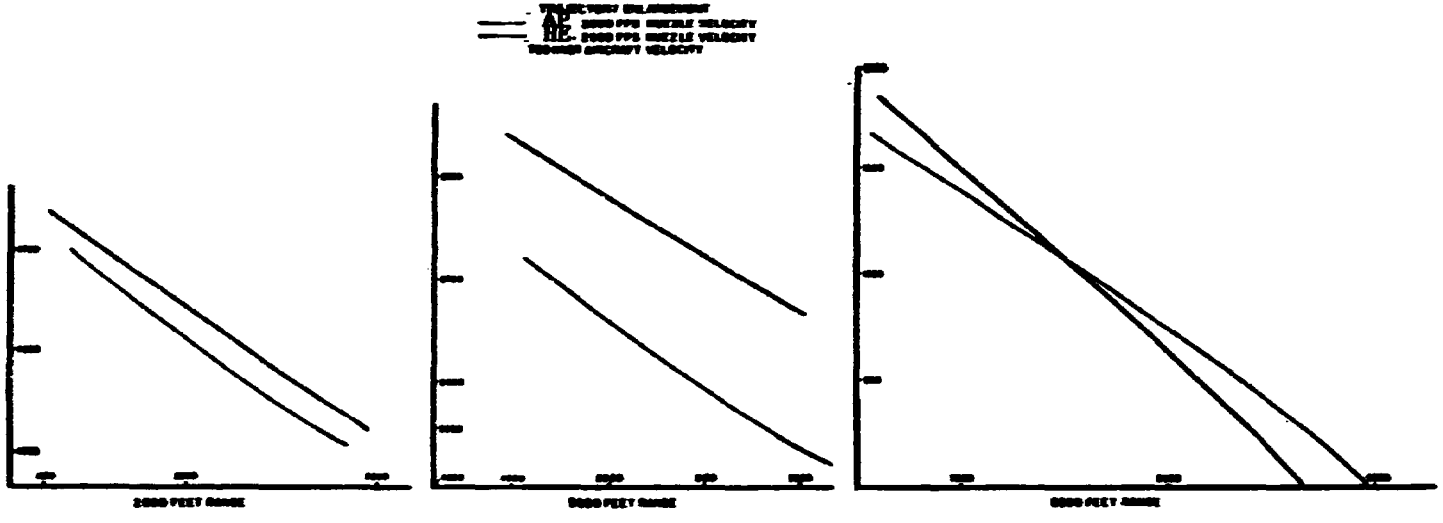


Figure 5. Trajectory Enlargement

TABLE IV

Probability of Hit (Target 10 x 10 feet)

Case Study	Round Type	Wt (grs)	Mux Vel (fps)	Slant Range (feet)	BML (mils)	AML (mils)	REML (mils)	Probability of Hit of N Rounds					
								1	5	10	15	30	76
I	HE	1950	2950	2309.4	1	3.47	0	.2036	.6796	.8973	.9671	.9998	1.0
	AP	2100	2850	2309.4	1	0	0	.9402	1.0	1.0	1.0	1.0	1.0
II	HE	1950	2950	6928.0	1	13.7	0	.0017	.0087	.0174	.0260	.0646	.1250
	AP	2100	2850	6928.0	1	0	0	.2804	.8071	.9627	.9928	.9999	1
III	HE	1950	2950	6928.0	4	13.7	0	.0016	.0081	.0161	.0241	.0599	.1163
	AP	2100	2850	6928.0	4	0	0	.0205	.0984	.1871	.2671	.5409	.7929
IV	HE	1950	2950	10381.0	1	18.3	0	.0004	.0021	.0043	.0065	.0165	.0328
	AP	2100	2850	10381.0	2	0	0	.0362	.1684	.3085	.4250	.7539	.9394

NOTES:

Altitude - 6000 feet
Aircraft Velocity - 790 knots (1267 fps)
Dive Angle - 30° from horizontal

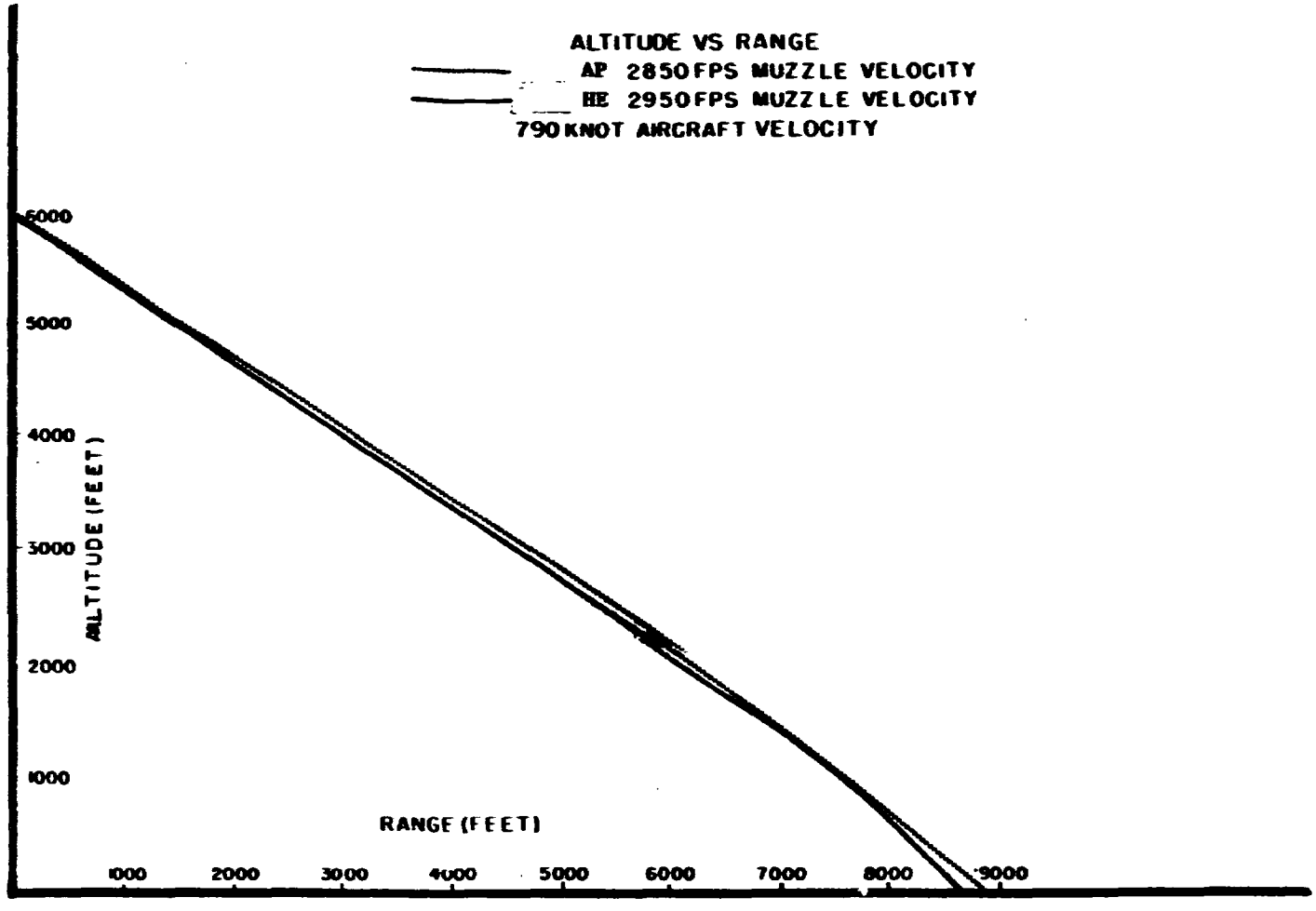


Figure 6. Trajectory Mismatch

Figure 7 shows the probability of hit versus number of rounds. The data graphed is taken from Table IV.

Table V compares probability of hit between a small and large target at short and long ranges with data obtained from exercise of the program as a second example.

The data sets for these exercises are contained in the Appendices. These data are graphically presented in Figures 8 and 9 for two different target sizes: a small target 10 by 10 feet and a large target 100 by 100 feet.

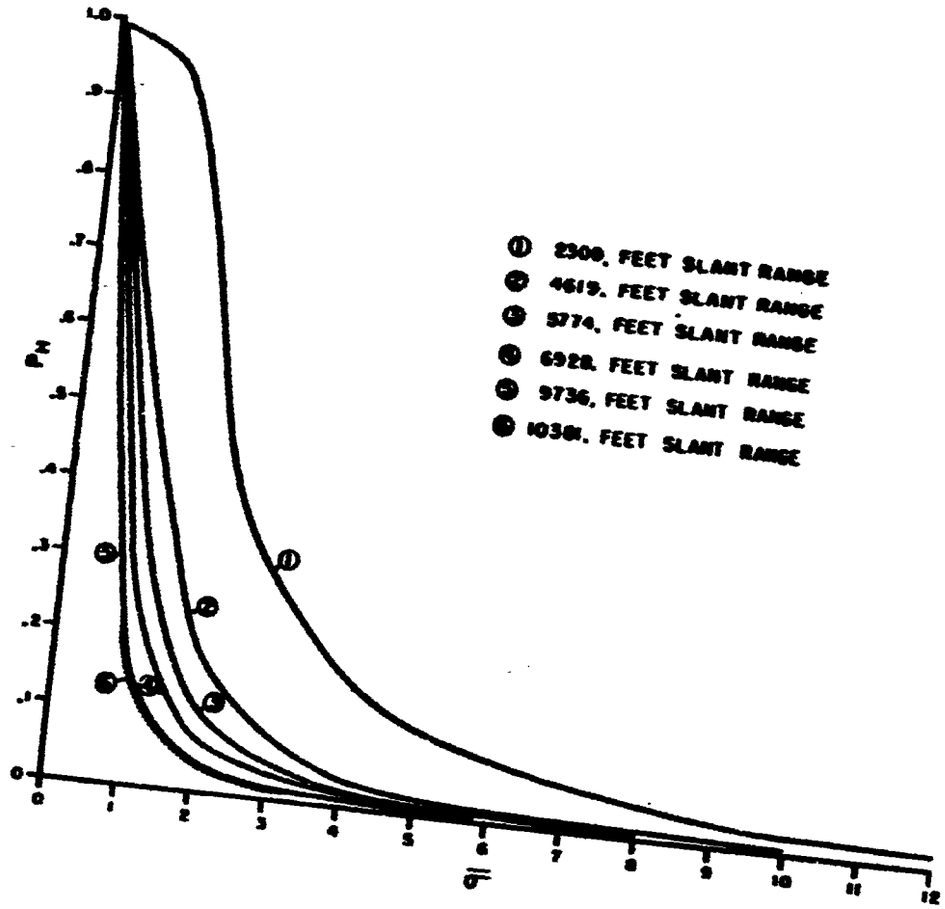


Figure 7. Probability of Hit versus Errors, Mils 10 by 10 Feet Target

TABLE V

Probability of Hit

Slant Range (feet)	Target - 10 x 10 feet BML (mils)	AML (mils)	REML (mils)	P_H	Target - 100 x 100 feet BML (mils)	AML (mils)	REML (mils)	P_H
2309	1	0	0	.9402	1	0	0	1.
	4	0	0	.1695	4	0	0	1.
	0	3.47	0	.2184	0	3.47	0	1.
	4	3.47	0	.1007	4	3.47	0	0.999
					4	3.47	18.3	0.554
					4	3.47	50.0	0.1111
					4	3.47	75.0	0.0514
10381	1	0	0	.1368	1	0	0	0.999
	4	0	0	.0091	4	0	0	0.5952
	0	18.3	0	.0004	0	18.3	0	0.0431
	4	18.3	0	.0004	4	18.3	0	0.0411

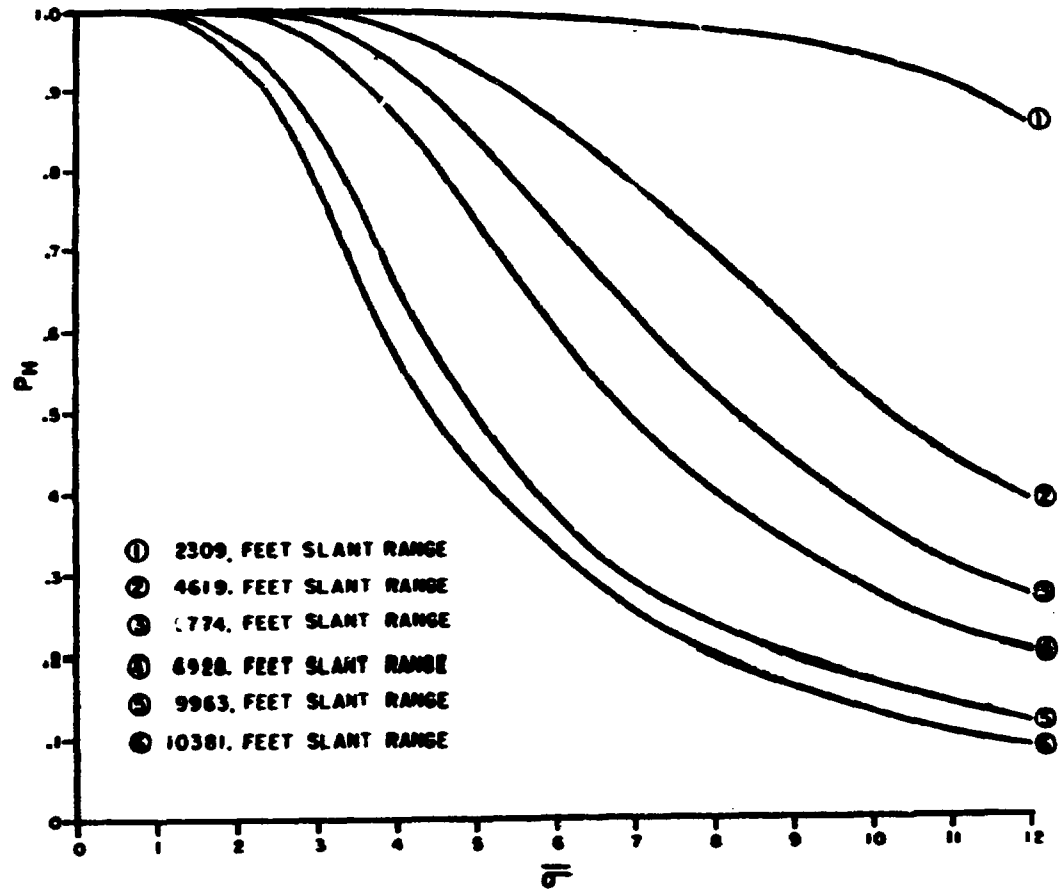


Figure 8. Probability of Hit versus Errors, Mils 100 by 100 Feet Target

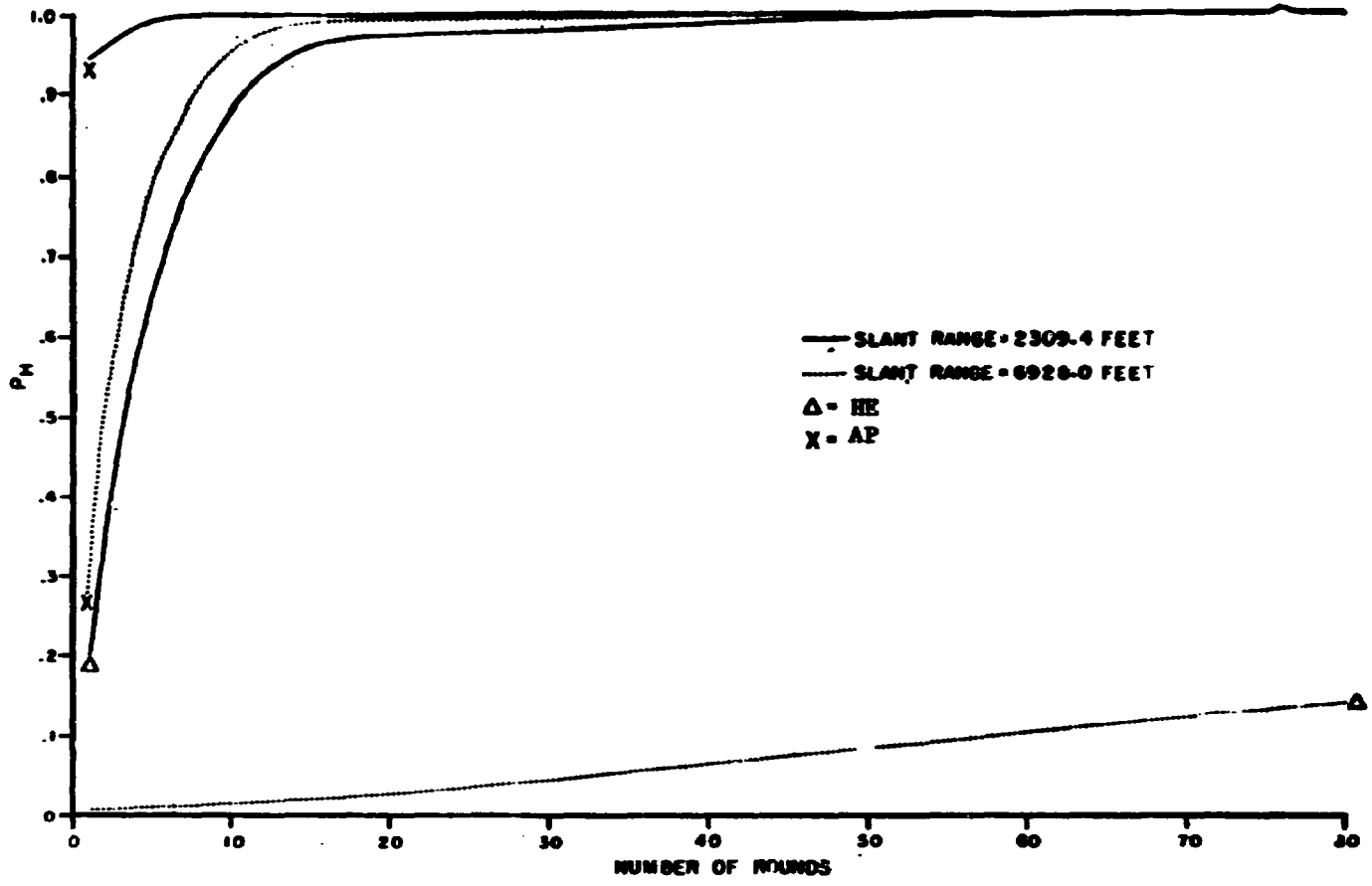


Figure 9. Probability of Hit versus Number of Rounds
10 feet x 10 feet target

BIBLIOGRPAHY

1. Bernard Harris, Theory of Probability, Addison-Wesley, Reading, MA, 1966.

2. Paul Doel, Introduction to Mathematical Statistics, Wiley Publication, NY, 1954.

3. Cecil Hastings, Jr., Approximations for Digital Computers, Princeton University Press, 1955.

4. US Army Weapons Command, "Some Mathematical Models and Computer Programs for Small Arms Analysis," SY-TN 10-80, Ad Hoc Small Arms Systems Analysis Working Group, 1971.

5. U. S. Army Management Engineering Training Agency, "Elements of Reliability and Maintainability," 1967.

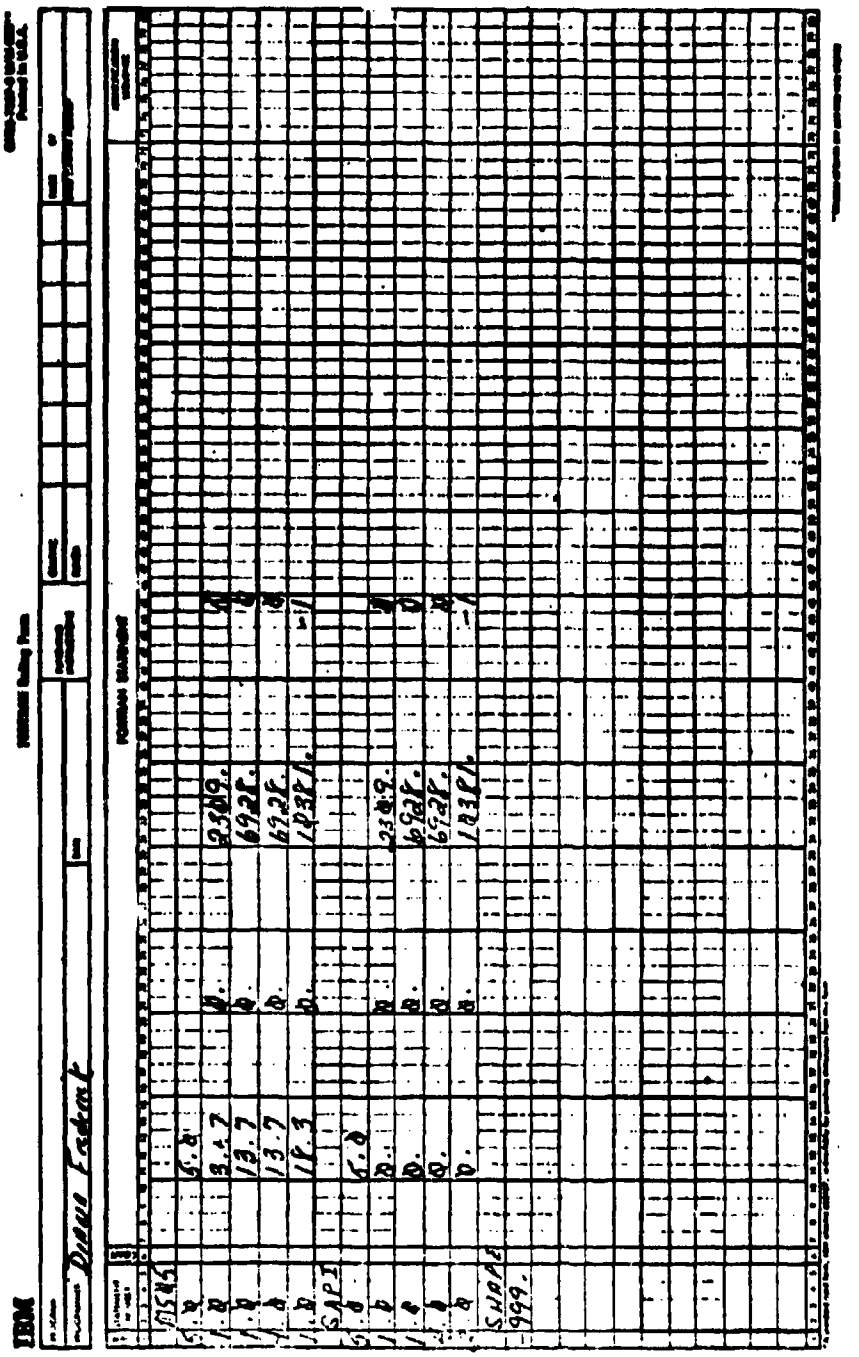
6. J. E. Miller, C. Snyder, H. Greveris, "A Malfunction Prediction Method for Small Arms Weapon Systems," Frankford Arsenal Report R-2059, November 1972.

7. J. E. Miller, H. Greveris, H. Jaslow, "Diagnostic Testing of Small Arms Weapons," Report R-2086, July 1973.

8. Diana Frederick, "Graphical Trajectory Compendium of Caliber .50 and 20mm Projectiles in an Air-to-Ground and Ground-to-Ground Role," Frankford Arsenal TN-1181, October 1973.

APPENDIX A

SAMPLE DATA INPUT



APPENDIX B

FORTRAN PROGRAM

```
      PROGRAM BIANAPN(INPUT,OUTPUT,TAPE1=INPUT,TAPE3=OUTPUT)
      DIMENSION SHAPE(6)
  100 FORMAT(4F10.2,I5)
  101 FORMAT( 30X, 20H PROBABILITY OF HIT , ////)
  102 FORMAT( 25X, 31H AIRCRAFT MUNITIONS LABORATORY ,///)
  300 FORMAT(3F10.2,7F8.4)
  310 FORMAT(30H        SR        BML        AML,32H   REML   P    , P38
     1    P76      , 20H  P5       P10    P15    / )
  320 FORMAT(/17H TARGET WIDTH IS , E15.8)
  330 FORMAT(/18H TARGET LENGTH IS , E15.8 / )
  340 FORMAT(/19H SR IS SLANT RANGE )
  350 FORMAT(/32H BML IS BALLISTIC ERROR IN MILS , 29H AML IS AIMING ERR
     2OR IN MILS , 38H REML IS RANGE ESTIMATE ERROR IN MILS )
  360 FORMAT(/29H P IS THE PROBABILITY OF HIT , 33H PN IS THE PROBABILIT
     2Y OF N HITS )
  370 FORMAT( 1H1)
  380 FORMAT(6A4)
  390 FORMAT(/  21H PROJECTILE DESIGN IS,6A4)
      IN=1
      IO=3
      WRITE(IO,370)
      WRITE(IO,101)
      WRITE(IO,102)
      WRITE(IO,350)
      WRITE(IO,340)
    4 CONTINUE
      READ(IN,380) SHAPE
      WRITE(IO,390) SHAPE
    2 READ(IN,100)RX,RY
      IF(RX-999.)3,40,3
    3 CONTINUE
      R=2.*RX
      WRITE(IO,320)R
      R=2.*RY
      WRITE(IO,330)R
      WRITE(IO,310)
    1 READ(IN,100)BML,AML,REML,SR,IND
      SIG=BML*BML+AML*AML+REML*REML
      SIG=SQRT(SIG)
      SIGM=SIG*SR/1000.
      ZX=RX/SIGM
      ZY=RY/SIGM
      Z=ZX/1.414
      PX=ERR(Z)
      Z=ZY/1.414
      PY=ERR(Z)
      P=PX*PY
      P38=1.-(1.-P)**38.
      P76=1.-(1.-P)**76.
      P5=1. - (1.-P)**5
      P10=1.-(1.-P)**10
      P15=1.-(1.-P)**15
      WRITE(IO,300)SR,BML,AML,REML,P,P38,P76,P5,P10,P15
      IF(IND)4,1,2
   40 STOP
```

```
      FUNCTION ERR(A)
C     ERROR FUNCTION SUBROUTINE
      A1=0.0705230784
      A2=0.0422820123
      A3=0.0092705272
      A4=.0001520143
      A5=.0002765672
      A6=.0000430638
      A=A
      C=1.0
      C=SIGN(C,A)
      Z=ABS(A)
      IF(Z-0.0000001)1,2,2
    2 IF(Z-4.0)3,3,4
    3 ERR= ((((((A6*Z+A5)*Z+A4)*Z+A3)*Z+A2)*Z+A1)*Z+1.0)**16.0
      ERR= ABS(1.0-(1.0/ERR))
      ERR= C*ERR
      RETURN
    1 ERR= 0.0
      RETURN
    4 ERR=C
      RETURN
      END
```

PROBABILITY OF HIT
AIRCRAFT MUNITIONS LABORATORY

BAL IS BALLISTIC ERROR IN MILS AML IS AIMING ERROR IN MILS AEML IS RANGE ESTIMATE ERROR IN MILS

SW IS SLANT RANGE

PROJECTILE VELOCITY IS 4000

TARGET WIDTH IS .1000000E+02

TARGET LENGTH IS .1000000E+02

SW	BAL	AML	AEML	P	P30	P70	P95	P10	P15
2300.00	1.00	0.00	0.0000	.2830	.9990	1.0000	.0707	.0976	.9471
4020.00	1.00	12.70	0.0000	.0810	.8840	.1251	.0088	.0174	.0280
6020.00	0.00	12.70	0.0000	.0810	.0880	.1104	.0081	.0162	.0241
10341.00	1.00	10.30	0.0000	.0804	.0160	.0329	.0022	.0044	.0068

PROJECTILE VELOCITY IS 5000

TARGET WIDTH IS .1000000E+02

TARGET LENGTH IS .1000000E+02

SW	BAL	AML	AEML	P	P30	P70	P95	P10	P15
2300.00	1.00	0.00	0.0000	.9902	1.0000	1.0000	1.0000	1.0000	1.0000
4020.00	1.00	0.00	0.0000	.8805	1.0000	1.0000	.0871	.9828	.9926
6020.00	0.00	0.00	0.0000	.8705	.8450	.7929	.0994	.1871	.2671
10341.00	2.00	0.00	0.0000	.8302	.7200	.4349	.1009	.3000	.4251

BML IS BALLISTIC ERROR IN MILS AML IS AIMING ERROR IN MILS REML IS RANGE ESTIMATE ERROR IN MILS

SR IS SLANT RANGE

PROJECTILE DESIGN IS M505

TARGET WIDTH IS .10000000E+02

TARGET LENGTH IS .10000000E+02

SR	BML	AML	REML	P	P30	P70	P90	P10	P15
2304.40	1.00	3.47	-0.0000	.2036	.9998	1.0000	.6797	.8974	.9671
2304.40	2.00	3.47	-0.0000	.1691	.9991	1.0000	.6046	.8432	.9379
2304.40	3.00	3.47	-0.0000	.1319	.9954	1.0000	.5069	.7568	.8801
2304.40	4.00	3.47	-0.0000	.1007	.9823	.9997	.4119	.6542	.7967
2304.40	1.00	4.85	-0.0000	.0300	.6854	.9010	.1411	.2624	.3665
2304.40	2.00	4.85	-0.0000	.0291	.6744	.8940	.1373	.2557	.3576
2304.40	3.00	4.85	-0.0000	.0277	.6567	.8822	.1312	.2452	.3443
2304.40	4.00	4.85	-0.0000	.0260	.6332	.8655	.1236	.2320	.3269
2304.40	1.00	13.96	-0.0000	.0151	.4395	.6858	.0733	.1413	.2043
2304.40	2.00	13.96	-0.0000	.0149	.4346	.6803	.0723	.1393	.2015
2304.40	3.00	13.96	-0.0000	.0145	.4266	.6712	.0706	.1362	.1971
2304.40	4.00	13.96	-0.0000	.0141	.4159	.6588	.0683	.1319	.1912
2304.40	1.00	15.60	-0.0000	.0121	.3713	.6047	.0592	.1150	.1674
2304.40	2.00	15.60	-0.0000	.0120	.3677	.6002	.0585	.1136	.1655
2304.40	3.00	15.60	-0.0000	.0118	.3620	.5929	.0574	.1115	.1625
2304.40	4.00	15.60	-0.0000	.0114	.3542	.5829	.0559	.1087	.1585
4419.00	1.00	21.30	-0.0000	.0016	.0605	.1173	.0082	.0163	.0243
4419.00	2.00	21.30	-0.0000	.0016	.0601	.1165	.0081	.0162	.0242
4419.00	3.00	21.30	-0.0000	.0016	.0594	.1153	.0080	.0160	.0239
4419.00	4.00	21.30	-0.0000	.0016	.0586	.1137	.0079	.0158	.0235
5774.00	1.00	24.60	-0.0000	.0008	.0295	.0581	.0039	.0078	.0117
5774.00	2.00	24.60	-0.0000	.0008	.0293	.0578	.0039	.0078	.0117
5774.00	3.00	24.60	-0.0000	.0008	.0291	.0574	.0039	.0077	.0116
5774.00	4.00	24.60	-0.0000	.0008	.0288	.0568	.0038	.0077	.0115
6928.20	1.00	13.70	-0.0000	.0018	.0646	.1250	.0087	.0174	.0260
6928.20	2.00	13.70	-0.0000	.0017	.0636	.1232	.0086	.0172	.0256
6928.20	3.00	13.70	-0.0000	.0017	.0621	.1203	.0084	.0167	.0250
6928.20	4.00	13.70	-0.0000	.0016	.0600	.1164	.0081	.0162	.0241
6928.20	1.00	23.80	-0.0000	.0006	.0220	.0434	.0029	.0058	.0087
6928.20	2.00	23.80	-0.0000	.0006	.0219	.0432	.0029	.0058	.0087
6928.20	3.00	23.80	-0.0000	.0006	.0217	.0429	.0029	.0057	.0086
6928.20	4.00	23.80	-0.0000	.0006	.0214	.0424	.0028	.0057	.0085
6928.20	1.00	26.50	-0.0000	.0005	.0178	.0352	.0024	.0047	.0070
6928.20	2.00	26.50	-0.0000	.0005	.0177	.0351	.0023	.0047	.0070
6928.20	3.00	26.50	-0.0000	.0005	.0176	.0348	.0023	.0047	.0070
6928.20	4.00	26.50	-0.0000	.0005	.0174	.0345	.0023	.0046	.0069
7490.00	1.00	30.00	-0.0000	.0003	.0113	.0225	.0015	.0030	.0045
7490.00	2.00	30.00	-0.0000	.0003	.0113	.0224	.0015	.0030	.0045
7490.00	3.00	30.00	-0.0000	.0003	.0112	.0223	.0015	.0030	.0044
7490.00	4.00	30.00	-0.0000	.0003	.0111	.0221	.0015	.0029	.0044
8083.00	1.00	22.10	-0.0000	.0005	.0187	.0371	.0025	.0050	.0074
8083.00	2.00	22.10	-0.0000	.0005	.0186	.0369	.0025	.0049	.0074
8083.00	3.00	22.10	-0.0000	.0005	.0184	.0365	.0024	.0049	.0073
8083.00	4.00	22.10	-0.0000	.0005	.0182	.0360	.0024	.0048	.0072
9237.60	1.00	6.56	-0.0000	.0042	.1487	.2753	.0210	.0415	.0616
9237.60	2.00	6.56	-0.0000	.0040	.1399	.2602	.0196	.0389	.0578
9237.60	3.00	6.56	-0.0000	.0036	.1274	.2385	.0178	.0352	.0524
9237.60	4.00	6.56	-0.0000	.0032	.1132	.2135	.0157	.0311	.0463
9963.30	1.00	30.40	-0.0000	.0002	.0066	.0131	.0009	.0017	.0026
9963.30	2.00	30.40	-0.0000	.0002	.0065	.0130	.0009	.0017	.0026
9963.30	3.00	30.40	-0.0000	.0002	.0065	.0130	.0009	.0017	.0026
9963.30	4.00	30.40	-0.0000	.0002	.0065	.0129	.0009	.0017	.0026
10381.00	1.00	18.30	0.0000	.0004	.0166	.0329	.0022	.0044	.0066
10381.00	2.00	18.30	0.0000	.0004	.0164	.0326	.0022	.0043	.0065
10381.00	3.00	18.30	0.0000	.0004	.0162	.0321	.0021	.0043	.0064
10381.00	4.00	18.30	0.0000	.0004	.0159	.0315	.0021	.0042	.0063
10797.00	1.00	24.60	-0.0000	.0002	.0085	.0170	.0011	.0023	.0034
10797.00	2.00	24.60	-0.0000	.0002	.0085	.0169	.0011	.0022	.0034
10797.00	3.00	24.60	-0.0000	.0002	.0084	.0168	.0011	.0022	.0033
10797.00	4.00	24.60	-0.0000	.0002	.0083	.0166	.0011	.0022	.0033
11365.00	1.00	33.33	-0.0000	.0001	.0042	.0084	.0006	.0011	.0017
11365.00	2.00	33.33	-0.0000	.0001	.0042	.0084	.0006	.0011	.0017
11365.00	3.00	33.33	-0.0000	.0001	.0042	.0083	.0006	.0011	.0016
11365.00	4.00	33.33	-0.0000	.0001	.0041	.0083	.0005	.0011	.0016

PROJECTILE DESIGN IS SAM1

TARGET WIDTH IS .10000000E+02

TARGET LENGTH IS .10000000E+02

SM	RML	AML	REML	P	P3A	P76	P5	P10	P15
2389.40	1.00	-0.00	-0.0000	.9402	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	2.00	-0.00	-0.0000	.5199	1.0000	1.0000	.9745	.9993	1.0000
2389.40	3.00	-0.00	-0.0000	.2805	1.0000	1.0000	.8071	.9628	.9928
2389.40	4.00	-0.00	-0.0000	.1695	.9991	1.0000	.6050	.8439	.9384
2389.40	5.00	-0.00	-0.0000	.1123	.9892	.9999	.4486	.6963	.8324
2389.40	6.00	-0.00	-0.0000	.0794	.9569	.9981	.3389	.5629	.7110
2389.40	7.00	-0.00	-0.0000	.0590	.9009	.9902	.2623	.4557	.5985
2389.40	8.00	-0.00	-0.0000	.0455	.8297	.9710	.2078	.3724	.5028
2389.40	9.00	-0.00	-0.0000	.0362	.7532	.9391	.1682	.3080	.4244
2389.40	10.00	-0.00	-0.0000	.0294	.6781	.8964	.1386	.2579	.3607
2389.40	11.00	-0.00	-0.0000	.0244	.6082	.8465	.1160	.2185	.3092
2389.40	12.00	-0.00	-0.0000	.0205	.5449	.7929	.0984	.1871	.2671
4619.80	1.00	-0.00	-0.0000	.5199	1.0000	1.0000	.9745	.9993	1.0000
4619.80	2.00	-0.00	-0.0000	.1695	.9991	1.0000	.6049	.8439	.9383
4619.80	3.00	-0.00	-0.0000	.0794	.9569	.9981	.3388	.5629	.7110
4619.80	4.00	-0.00	-0.0000	.0455	.8297	.9710	.2078	.3724	.5028
4619.80	5.00	-0.00	-0.0000	.0294	.6781	.8964	.1385	.2579	.3607
4619.80	6.00	-0.00	-0.0000	.0205	.5449	.7929	.0984	.1871	.2671
4619.80	7.00	-0.00	-0.0000	.0151	.4392	.6856	.0733	.1412	.2042
4619.80	8.00	-0.00	-0.0000	.0116	.3578	.5876	.0566	.1100	.1604
4619.80	9.00	-0.00	-0.0000	.0092	.2953	.5034	.0450	.0880	.1290
4619.80	10.00	-0.00	-0.0000	.0074	.2469	.4328	.0366	.0719	.1059
4619.80	11.00	-0.00	-0.0000	.0061	.2089	.3741	.0304	.0598	.0882
4619.80	12.00	-0.00	-0.0000	.0052	.1787	.3255	.0256	.0505	.0748
5774.80	1.00	-0.00	-0.0000	.3764	1.0000	1.0000	.9057	.9911	.9992
5774.80	2.00	-0.00	-0.0000	.1122	.9892	.9999	.4486	.6959	.8323
5774.80	3.00	-0.00	-0.0000	.0516	.8665	.9822	.2328	.4114	.5484
5774.80	4.00	-0.00	-0.0000	.0294	.6780	.8963	.1385	.2579	.3607
5774.80	5.00	-0.00	-0.0000	.0189	.5159	.7657	.0910	.1738	.2490
5774.80	6.00	-0.00	-0.0000	.0132	.3958	.6350	.0642	.1242	.1804
5774.80	7.00	-0.00	-0.0000	.0097	.3094	.5231	.0475	.0928	.1360
5774.80	8.00	-0.00	-0.0000	.0074	.2468	.4328	.0366	.0719	.1059
5774.80	9.00	-0.00	-0.0000	.0059	.2007	.3611	.0290	.0572	.0846
5774.80	10.00	-0.00	-0.0000	.0048	.1659	.3043	.0236	.0466	.0691
5774.80	11.00	-0.00	-0.0000	.0039	.1393	.2591	.0195	.0387	.0575
5774.80	12.00	-0.00	-0.0000	.0033	.1184	.2228	.0164	.0326	.0485
6928.20	1.00	-0.00	-0.0000	.2805	1.0000	1.0000	.8071	.9628	.9928
6928.20	2.00	-0.00	-0.0000	.0794	.9569	.9981	.3389	.5629	.7110
6928.20	3.00	-0.00	-0.0000	.0362	.7532	.9391	.1682	.3080	.4244
6928.20	4.00	-0.00	-0.0000	.0205	.5449	.7929	.0984	.1871	.2671
6928.20	5.00	-0.00	-0.0000	.0132	.3959	.6350	.0642	.1242	.1804
6928.20	6.00	-0.00	-0.0000	.0092	.2953	.5034	.0450	.0880	.1290
6928.20	7.00	-0.00	-0.0000	.0067	.2268	.4021	.0333	.0654	.0965
6928.20	8.00	-0.00	-0.0000	.0052	.1787	.3255	.0256	.0505	.0748
6928.20	9.00	-0.00	-0.0000	.0041	.1441	.2674	.0203	.0401	.0596
6928.20	10.00	-0.00	-0.0000	.0033	.1184	.2228	.0164	.0326	.0485
6928.20	11.00	-0.00	-0.0000	.0027	.0989	.1880	.0136	.0270	.0403
6928.20	12.00	-0.00	-0.0000	.0023	.0838	.1606	.0114	.0228	.0340
7480.00	1.00	-0.00	-0.0000	.2462	1.0000	1.0000	.7567	.9408	.9856
7480.00	2.00	-0.00	-0.0000	.0686	.9327	.9955	.2989	.5084	.6554
7480.00	3.00	-0.00	-0.0000	.0311	.6990	.9094	.1461	.2709	.3774
7480.00	4.00	-0.00	-0.0000	.0176	.4911	.7410	.0850	.1629	.2341
7690.00	1.00	-0.00	-0.0000	.2347	1.0000	1.0000	.7375	.9311	.9819
7490.00	2.00	-0.00	-0.0000	.0650	.9222	.9939	.2854	.4893	.6350
7690.00	3.00	-0.00	-0.0000	.0294	.6789	.8969	.1388	.2584	.3613
7490.00	4.00	-0.00	-0.0000	.0167	.4722	.7215	.0807	.1548	.2230
7940.00	1.00	-0.00	-0.0000	.2220	.9999	1.0000	.7150	.9188	.9768
7940.00	2.00	-0.00	-0.0000	.0611	.9089	.9917	.2704	.4676	.6115
7940.00	3.00	-0.00	-0.0000	.0277	.6555	.8813	.1308	.2445	.3433
7940.00	4.00	-0.00	-0.0000	.0157	.4509	.6985	.0759	.1460	.2107
8083.00	1.00	-0.00	-0.0000	.2152	.9999	1.0000	.7022	.9113	.9736
8083.00	2.00	-0.00	-0.0000	.0590	.9009	.9902	.2623	.4557	.5985
8083.00	3.00	-0.00	-0.0000	.0267	.6423	.8721	.1265	.2371	.3336
8083.00	4.00	-0.00	-0.0000	.0151	.4393	.6856	.0733	.1412	.2042

PROJECTILE DESIGN IS SAP1 (cont'd)
TARGET WIDTH IS .10000000E+02
TARGET LENGTH IS .10000000E+02

SR	RML	AML	R&ML	P	P38	P76	P5	P10	P15
8244.00	1.00	-0.00	-0.0000	.2068	.9998	1.0000	.6859	.9014	.9690
8244.00	2.00	-0.00	-0.0000	.0565	.8902	.9880	.2523	.4409	.5819
8244.00	3.00	-0.00	-0.0000	.0255	.6257	.8599	.1213	.2279	.3215
8244.00	4.00	-0.00	-0.0000	.0144	.4247	.6691	.0702	.1354	.1961
8432.00	1.00	-0.00	-0.0000	.1997	.9998	1.0000	.6717	.8922	.9646
8432.00	2.00	-0.00	-0.0000	.0544	.8805	.9857	.2439	.4283	.5677
8432.00	3.00	-0.00	-0.0000	.0246	.6113	.8489	.1169	.2201	.3113
8432.00	4.00	-0.00	-0.0000	.0139	.4124	.6547	.0676	.1306	.1893
9237.60	1.00	-0.00	-0.0000	.1695	.9991	1.0000	.6050	.8439	.9384
9237.60	2.00	-0.00	-0.0000	.0455	.8297	.9710	.2078	.3724	.5028
9237.60	3.00	-0.00	-0.0000	.0205	.5449	.7929	.0984	.1871	.2671
9237.60	4.00	-0.00	-0.0000	.0116	.3579	.5877	.0566	.1100	.1604
9736.40	1.00	-0.00	-0.0000	.1540	.9983	1.0000	.5667	.8123	.9187
9736.40	2.00	-0.00	-0.0000	.0411	.7969	.9587	.1842	.3426	.4670
9736.40	3.00	-0.00	-0.0000	.0185	.5077	.7577	.0890	.1702	.2440
9736.40	4.00	-0.00	-0.0000	.0104	.3338	.5496	.0511	.0996	.1456
9736.00	5.00	-0.00	-0.0000	.0067	.2353	.3998	.0330	.0650	.0958
9736.00	6.00	-0.00	-0.0000	.0047	.1624	.2985	.0231	.0456	.0676
9736.00	7.00	-0.00	-0.0000	.0034	.1221	.2293	.0170	.0337	.0501
9736.00	8.00	-0.00	-0.0000	.0026	.0949	.1808	.0130	.0259	.0386
9736.00	9.00	-0.00	-0.0000	.0021	.0758	.1458	.0103	.0205	.0306
9736.00	10.00	-0.00	-0.0000	.0017	.0618	.1198	.0084	.0167	.0249
9736.00	11.00	-0.00	-0.0000	.0014	.0514	.1001	.0069	.0138	.0206
9736.00	12.00	-0.00	-0.0000	.0012	.0434	.0848	.0058	.0116	.0173
9963.30	1.00	-0.00	-0.0000	.1477	.9977	1.0000	.5502	.7977	.9090
9963.30	2.00	-0.00	-0.0000	.0393	.7818	.9574	.1815	.3301	.4517
9963.30	3.00	-0.00	-0.0000	.0177	.4918	.7417	.0852	.1632	.2345
9963.30	4.00	-0.00	-0.0000	.0100	.3167	.5331	.0489	.0954	.1396
10341.00	1.00	-0.00	-0.0000	.1369	.9963	1.0000	.5210	.7706	.8901
10341.00	2.00	-0.00	-0.0000	.0362	.7540	.9395	.1685	.3086	.4251
10341.00	3.00	-0.00	-0.0000	.0163	.4639	.7126	.0788	.1513	.2182
10341.00	4.00	-0.00	-0.0000	.0092	.2959	.5042	.0451	.0882	.1293
10381.00	5.00	-0.00	-0.0000	.0059	.2011	.3618	.0291	.0574	.0848
10381.00	6.00	-0.00	-0.0000	.0041	.1444	.2679	.0203	.0402	.0597
10341.00	7.00	-0.00	-0.0000	.0030	.1082	.2048	.0150	.0297	.0442
10341.00	8.00	-0.00	-0.0000	.0023	.0840	.1689	.0115	.0228	.0340
10301.00	9.00	-0.00	-0.0000	.0019	.0670	.1294	.0091	.0181	.0270
10381.00	10.00	-0.00	-0.0000	.0015	.0546	.1062	.0074	.0147	.0219
10341.00	11.00	-0.00	-0.0000	.0012	.0453	.0886	.0061	.0121	.0181
10381.00	12.00	-0.00	-0.0000	.0010	.0382	.0750	.0051	.0102	.0153
10797.00	1.00	-0.00	-0.0000	.1273	.9943	1.0000	.4937	.7437	.8702
10797.00	2.00	-0.00	-0.0000	.0335	.7265	.9252	.1568	.2990	.4005
10797.00	3.00	-0.00	-0.0000	.0151	.4381	.6843	.0730	.1407	.2035
10797.00	4.00	-0.00	-0.0000	.0085	.2770	.4772	.0418	.0818	.1202
11365.00	1.00	-0.00	-0.0000	.1156	.9906	.9999	.4591	.7074	.8417
11365.00	2.00	-0.00	-0.0000	.0303	.6897	.9037	.1427	.2650	.3699
11365.00	3.00	-0.00	-0.0000	.0138	.4056	.6467	.0662	.1279	.1856
11365.00	4.00	-0.00	-0.0000	.0077	.2537	.4431	.0378	.0741	.1091

BML IS BALLISTIC ERROR IN MILS. AML IS AIMING ERROR IN MILS REML IS RANGE ESTIMATE ERROR IN MILS

SR IS SLANT RANGE

PROJECTILE DESIGN IS M505

TARGET WIDTH IS .10000000E+03

TARGET LENGTH IS .10000000E+03

SR	BML	AML	REML	P	P38	P76	P5	P10	P15
2309.40	1.00	3.47	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2309.40	2.00	3.47	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2309.40	3.00	3.47	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2309.40	4.00	3.47	-0.0000	.9999	1.0000	1.0000	1.0000	1.0000	1.0000
2309.40	1.00	9.85	-0.0000	.9434	1.0000	1.0000	1.0000	1.0000	1.0000
2309.40	2.00	9.85	-0.0000	.9386	1.0000	1.0000	1.0000	1.0000	1.0000
2309.40	3.00	9.85	-0.0000	.9303	1.0000	1.0000	1.0000	1.0000	1.0000
2309.40	4.00	9.85	-0.0000	.9184	1.0000	1.0000	1.0000	1.0000	1.0000
2309.40	1.00	13.96	-0.0000	.7712	1.0000	1.0000	.9994	1.0000	1.0000
2309.40	2.00	13.96	-0.0000	.7662	1.0000	1.0000	.9993	1.0000	1.0000
2309.40	3.00	13.96	-0.0000	.7580	1.0000	1.0000	.9992	1.0000	1.0000
2309.40	4.00	13.96	-0.0000	.7466	1.0000	1.0000	.9990	1.0000	1.0000
2309.40	1.00	15.60	-0.0000	.6956	1.0000	1.0000	.9974	1.0000	1.0000
2309.40	2.00	15.60	-0.0000	.6913	1.0000	1.0000	.9972	1.0000	1.0000
2309.40	3.00	15.60	-0.0000	.6842	1.0000	1.0000	.9969	1.0000	1.0000
2309.40	4.00	15.60	-0.0000	.6744	1.0000	1.0000	.9963	1.0000	1.0000
4619.00	1.00	21.30	-0.0000	.1508	.9980	1.0000	.5584	.8050	.9139
4619.00	2.00	21.30	-0.0000	.1499	.9979	1.0000	.5561	.8029	.9125
4619.00	3.00	21.30	-0.0000	.1484	.9978	1.0000	.5522	.7994	.9102
4619.00	4.00	21.30	-0.0000	.1464	.9976	1.0000	.5468	.7946	.9069
5774.00	1.00	24.60	-0.0000	.0756	.9496	.9975	.3251	.5445	.6926
5774.00	2.00	24.60	-0.0000	.0753	.9489	.9974	.3238	.5427	.6908
5774.00	3.00	24.60	-0.0000	.0747	.9476	.9973	.3216	.5398	.6878
5774.00	4.00	24.60	-0.0000	.0739	.9459	.9971	.3187	.5358	.6837
6928.20	1.00	13.70	-0.0000	.1606	.9987	1.0000	.5833	.8263	.9276
6928.20	2.00	13.70	-0.0000	.1583	.9986	1.0000	.5775	.8215	.9246
6928.20	3.00	13.70	-0.0000	.1546	.9983	1.0000	.5682	.8136	.9195
6928.20	4.00	13.70	-0.0000	.1497	.9979	1.0000	.5556	.8025	.9122
6928.20	1.00	23.80	-0.0000	.0567	.8912	.9882	.2531	.4472	.5834
6928.20	2.00	23.80	-0.0000	.0564	.8899	.9879	.2520	.4405	.5815
6928.20	3.00	23.80	-0.0000	.0559	.8878	.9874	.2501	.4376	.5783
6928.20	4.00	23.80	-0.0000	.0553	.8848	.9867	.2475	.4337	.5739
6928.20	1.00	26.50	-0.0000	.0460	.8331	.9721	.2099	.3757	.5067
6928.20	2.00	26.50	-0.0000	.0458	.8318	.9717	.2091	.3744	.5052
6928.20	3.00	26.50	-0.0000	.0455	.8297	.9710	.2078	.3724	.5028
6928.20	4.00	26.50	-0.0000	.0451	.8267	.9700	.2060	.3695	.4994
7690.80	1.00	30.00	-0.0000	.0294	.6785	.8966	.1387	.2581	.3610
7690.80	2.00	30.00	-0.0000	.0293	.6772	.8958	.1383	.2574	.3601
7690.80	3.00	30.00	-0.0000	.0292	.6752	.8945	.1376	.2562	.3585
7690.80	4.00	30.00	-0.0000	.0289	.6724	.8927	.1366	.2545	.3563
8083.80	1.00	22.10	-0.0000	.0485	.8489	.9772	.2202	.3918	.5257
8083.80	2.00	22.10	-0.0000	.0482	.8471	.9766	.2190	.3900	.5236
8083.80	3.00	22.10	-0.0000	.0478	.8442	.9757	.2170	.3870	.5200
8083.80	4.00	22.10	-0.0000	.0471	.8402	.9745	.2144	.3828	.5151
9237.60	1.00	6.54	-0.0000	.3427	1.0000	1.0000	.8773	.9849	.9982
9237.60	2.00	6.56	-0.0000	.3250	1.0000	1.0000	.8599	.9804	.9972
9237.60	3.00	6.56	-0.0000	.2992	1.0000	1.0000	.8310	.9714	.9952
9237.60	4.00	6.56	-0.0000	.2693	1.0000	1.0000	.7917	.9566	.9910
9963.30	1.00	30.40	-0.0000	.0172	.4824	.7320	.0830	.1591	.2289
9963.30	2.00	30.40	-0.0000	.0171	.4813	.7309	.0827	.1586	.2282
9963.30	3.00	30.40	-0.0000	.0170	.4794	.7290	.0823	.1578	.2272
9963.30	4.00	30.40	-0.0000	.0169	.4769	.7264	.0817	.1568	.2257
10391.00	1.00	18.30	0.0000	.0430	.8117	.9645	.1972	.3556	.4827
10391.00	2.00	18.30	0.0000	.0426	.8089	.9635	.1957	.3530	.4796
10391.00	3.00	18.30	0.0000	.0420	.8042	.9617	.1931	.3489	.4747
10391.00	4.00	18.30	0.0000	.0412	.7978	.9591	.1897	.3433	.4679
10797.00	1.00	24.60	-0.0000	.0223	.5750	.8194	.1065	.2016	.2866
10797.00	2.00	24.60	-0.0000	.0222	.5732	.8179	.1060	.2007	.2855
10797.00	3.00	24.60	-0.0000	.0220	.5702	.8153	.1052	.1993	.2835
10797.00	4.00	24.60	-0.0000	.0217	.5661	.8118	.1041	.1973	.2808
11365.00	1.00	33.33	-0.0000	.0110	.3437	.5693	.0539	.1049	.1532
11365.00	2.00	33.33	-0.0000	.0110	.3430	.5683	.0538	.1046	.1528
11365.00	3.00	33.33	-0.0000	.0109	.3417	.5667	.0535	.1042	.1521
11365.00	4.00	33.33	-0.0000	.0109	.3400	.5644	.0532	.1036	.1513

PROJECTILE DESIGN IS SAPI

TARGET WIDTH IS .10000000E+03

TARGET LENGTH IS .10000000E+03

SN	BML	AML	REML	P	P38	P76	P5	P10	P15
2389.40	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	2.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	3.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	4.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	5.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	6.00	-0.00	-0.0000	.9994	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	7.00	-0.00	-0.0000	.9960	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	8.00	-0.00	-0.0000	.9865	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	9.00	-0.00	-0.0000	.9680	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	10.00	-0.00	-0.0000	.9402	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	11.00	-0.00	-0.0000	.9044	1.0000	1.0000	1.0000	1.0000	1.0000
2389.40	12.00	-0.00	-0.0000	.8628	1.0000	1.0000	1.0000	1.0000	1.0000
4619.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4619.00	2.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4619.00	3.00	-0.00	-0.0000	.9994	1.0000	1.0000	1.0000	1.0000	1.0000
4619.00	4.00	-0.00	-0.0000	.9965	1.0000	1.0000	1.0000	1.0000	1.0000
4619.00	5.00	-0.00	-0.0000	.9402	1.0000	1.0000	1.0000	1.0000	1.0000
4619.00	6.00	-0.00	-0.0000	.8627	1.0000	1.0000	1.0000	1.0000	1.0000
4619.00	7.00	-0.00	-0.0000	.7710	1.0000	1.0000	.9994	1.0000	1.0000
4619.00	8.00	-0.00	-0.0000	.6790	1.0000	1.0000	.9966	1.0000	1.0000
4619.00	9.00	-0.00	-0.0000	.5944	1.0000	1.0000	.9990	.9999	1.0000
4619.00	10.00	-0.00	-0.0000	.5199	1.0000	1.0000	.9745	.9993	1.0000
4619.00	11.00	-0.00	-0.0000	.4555	1.0000	1.0000	.9522	.9977	.9999
4619.00	12.00	-0.00	-0.0000	.4008	1.0000	1.0000	.9227	.9940	.9995
5774.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5774.00	2.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5774.00	3.00	-0.00	-0.0000	.9922	1.0000	1.0000	1.0000	1.0000	1.0000
5774.00	4.00	-0.00	-0.0000	.9402	1.0000	1.0000	1.0000	1.0000	1.0000
5774.00	5.00	-0.00	-0.0000	.8404	1.0000	1.0000	.9999	1.0000	1.0000
5774.00	6.00	-0.00	-0.0000	.7244	1.0000	1.0000	.9984	1.0000	1.0000
5774.00	7.00	-0.00	-0.0000	.6147	1.0000	1.0000	.9915	.9999	1.0000
5774.00	8.00	-0.00	-0.0000	.5199	1.0000	1.0000	.9745	.9993	1.0000
5774.00	9.00	-0.00	-0.0000	.4410	1.0000	1.0000	.9454	.9970	.9998
5774.00	10.00	-0.00	-0.0000	.3764	1.0000	1.0000	.9057	.9911	.9992
5774.00	11.00	-0.00	-0.0000	.3237	1.0000	1.0000	.8585	.9800	.9972
5774.00	12.00	-0.00	-0.0000	.2804	1.0000	1.0000	.8071	.9628	.9928
6928.20	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6928.20	2.00	-0.00	-0.0000	.9994	1.0000	1.0000	1.0000	1.0000	1.0000
6928.20	3.00	-0.00	-0.0000	.9680	1.0000	1.0000	1.0000	1.0000	1.0000
6928.20	4.00	-0.00	-0.0000	.8628	1.0000	1.0000	1.0000	1.0000	1.0000
6928.20	5.00	-0.00	-0.0000	.7245	1.0000	1.0000	.9984	1.0000	1.0000
6928.20	6.00	-0.00	-0.0000	.5945	1.0000	1.0000	.9890	.9999	1.0000
6928.20	7.00	-0.00	-0.0000	.4865	1.0000	1.0000	.9643	.9987	1.0000
6928.20	8.00	-0.00	-0.0000	.4008	1.0000	1.0000	.9227	.9940	.9995
6928.20	9.00	-0.00	-0.0000	.3334	1.0000	1.0000	.8684	.9827	.9977
6928.20	10.00	-0.00	-0.0000	.2805	1.0000	1.0000	.8071	.9628	.9928
6928.20	11.00	-0.00	-0.0000	.2384	1.0000	1.0000	.7438	.9344	.9832
6928.20	12.00	-0.00	-0.0000	.2047	.9998	1.0000	.6819	.8988	.9678
7480.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7480.00	2.00	-0.00	-0.0000	.9983	1.0000	1.0000	1.0000	1.0000	1.0000
7480.00	3.00	-0.00	-0.0000	.9490	1.0000	1.0000	1.0000	1.0000	1.0000
7480.00	4.00	-0.00	-0.0000	.8197	1.0000	1.0000	.9998	1.0000	1.0000
7690.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7690.00	2.00	-0.00	-0.0000	.9977	1.0000	1.0000	1.0000	1.0000	1.0000
7690.00	3.00	-0.00	-0.0000	.9405	1.0000	1.0000	1.0000	1.0000	1.0000
7690.00	4.00	-0.00	-0.0000	.8028	1.0000	1.0000	.9997	1.0000	1.0000
7940.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7940.00	2.00	-0.00	-0.0000	.9967	1.0000	1.0000	1.0000	1.0000	1.0000
7940.00	3.00	-0.00	-0.0000	.9297	1.0000	1.0000	1.0000	1.0000	1.0000
7940.00	4.00	-0.00	-0.0000	.7826	1.0000	1.0000	.9995	1.0000	1.0000
8883.80	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8883.80	2.00	-0.00	-0.0000	.9960	1.0000	1.0000	1.0000	1.0000	1.0000
8883.80	3.00	-0.00	-0.0000	.9232	1.0000	1.0000	1.0000	1.0000	1.0000
8883.80	4.00	-0.00	-0.0000	.7710	1.0000	1.0000	.9994	1.0000	1.0000

PROJECTILE DESIGN IS SAPI (cont'd)
TARGET WIDTH IS .10000000E+03
TARGET LENGTH IS .10000000E+03

SR	BML	AML	REML	P	P38	P76	P5	P10	P15
8268.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8268.00	2.00	-0.00	-0.0000	.9950	1.0000	1.0000	1.0000	1.0000	1.0000
8268.00	3.00	-0.00	-0.0000	.9143	1.0000	1.0000	1.0000	1.0000	1.0000
8268.00	4.00	-0.00	-0.0000	.7560	1.0000	1.0000	.9991	1.0000	1.0000
8432.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8432.00	2.00	-0.00	-0.0000	.9940	1.0000	1.0000	1.0000	1.0000	1.0000
8432.00	3.00	-0.00	-0.0000	.9062	1.0000	1.0000	1.0000	1.0000	1.0000
8432.00	4.00	-0.00	-0.0000	.7428	1.0000	1.0000	.9949	1.0000	1.0000
9237.60	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9237.60	2.00	-0.00	-0.0000	.9865	1.0000	1.0000	1.0000	1.0000	1.0000
9237.60	3.00	-0.00	-0.0000	.8628	1.0000	1.0000	1.0000	1.0000	1.0000
9237.60	4.00	-0.00	-0.0000	.6741	1.0000	1.0000	.9966	1.0000	1.0000
9736.40	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9736.40	2.00	-0.00	-0.0000	.9797	1.0000	1.0000	1.0000	1.0000	1.0000
9736.40	3.00	-0.00	-0.0000	.8338	1.0000	1.0000	.9999	1.0000	1.0000
9736.40	4.00	-0.00	-0.0000	.6414	1.0000	1.0000	.9941	1.0000	1.0000
9736.00	5.00	-0.00	-0.0000	.4840	1.0000	1.0000	.9634	.9987	1.0000
9736.00	6.00	-0.00	-0.0000	.3697	1.0000	1.0000	.9005	.9901	.9990
9736.00	7.00	-0.00	-0.0000	.2883	1.0000	1.0000	.8174	.9666	.9939
9736.00	8.00	-0.00	-0.0000	.2296	1.0000	1.0000	.7286	.9263	.9800
9736.00	9.00	-0.00	-0.0000	.1865	.9996	1.0000	.6436	.8730	.9547
9736.00	10.00	-0.00	-0.0000	.1541	.9913	1.0000	.5668	.8123	.9187
9736.00	11.00	-0.00	-0.0000	.1292	.9948	1.0000	.4993	.7493	.8745
9736.00	12.00	-0.00	-0.0000	.1098	.9830	.9999	.4410	.6875	.8253
9963.30	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9963.30	2.00	-0.00	-0.0000	.9760	1.0000	1.0000	1.0000	1.0000	1.0000
9963.30	3.00	-0.00	-0.0000	.8203	1.0000	1.0000	.9998	1.0000	1.0000
9963.30	4.00	-0.00	-0.0000	.6248	1.0000	1.0000	.9926	.9999	1.0000
10381.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10381.00	2.00	-0.00	-0.0000	.9682	1.0000	1.0000	1.0000	1.0000	1.0000
10381.00	3.00	-0.00	-0.0000	.7951	1.0000	1.0000	.9996	1.0000	1.0000
10381.00	4.00	-0.00	-0.0000	.5953	1.0000	1.0000	.9891	.9999	1.0000
10381.00	5.00	-0.00	-0.0000	.4418	1.0000	1.0000	.9458	.9971	.9998
10381.00	6.00	-0.00	-0.0000	.3340	1.0000	1.0000	.8690	.9828	.9978
10381.00	7.00	-0.00	-0.0000	.2587	1.0000	1.0000	.7762	.9499	.9888
10381.00	8.00	-0.00	-0.0000	.2051	.9998	1.0000	.6827	.8993	.9681
10381.00	9.00	-0.00	-0.0000	.1661	.9990	1.0000	.5967	.8373	.9344
10381.00	10.00	-0.00	-0.0000	.1369	.9963	1.0000	.5210	.7706	.8901
10381.00	11.00	-0.00	-0.0000	.1146	.9902	.9999	.4560	.7040	.8390
10381.00	12.00	-0.00	-0.0000	.0973	.9795	.9996	.4005	.6406	.7846
10797.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10797.00	2.00	-0.00	-0.0000	.9593	1.0000	1.0000	1.0000	1.0000	1.0000
10797.00	3.00	-0.00	-0.0000	.7698	1.0000	1.0000	.9994	1.0000	1.0000
10797.00	4.00	-0.00	-0.0000	.5672	1.0000	1.0000	.9848	.9998	1.0000
11365.00	1.00	-0.00	-0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11365.00	2.00	-0.00	-0.0000	.9452	1.0000	1.0000	1.0000	1.0000	1.0000
11365.00	3.00	-0.00	-0.0000	.7354	1.0000	1.0000	.9987	1.0000	1.0000
11365.00	4.00	-0.00	-0.0000	.5310	1.0000	1.0000	.9773	.9995	1.0000

"CONDITION/EFFECTIVENESS MODEL FOR FACILITY COMPONENTS"

R.J. Colver, J.G. Kirby, P.V. Kauffold, R.E. DeVor and M.J. Kraitsik

U.S. Army Construction Engineering Research Laboratory
Champaign, Illinois

INTRODUCTION

The quantification of facility deterioration is a problem currently facing the U.S. Army Corps of Engineers. Quantification is immediately necessary for determination of maintenance costs over the life of the facility and for projection of facility maintenance requirements for use in programming and budgeting maintenance resources.

The analysis of maintenance cost over the life of the facility is inherent in the process of life cycle costing alternate facility construction systems, e.g., brick and tile vs. frame and shingle. Life cycle costing is required, by regulation, on all new construction. The projection of maintenance requirements is made necessary by the scarcity of maintenance resources and the need to start programming process some two to three years prior to the budget year. A major problem with both efforts is the quantification of the condition of the facility over time in terms that are common to all facilities and are easily convertible to dollar costs or some other unit of measure necessary for policy formulation.

The Office of the Chief of Engineers (OCE) and the Office of the Deputy Chief of Staff for Logistics (DCSLOG) are currently developing an Army facilities information system, the Integrated Facilities System (IFS), which provides a scheme for breaking facilities into common components (See Figure 1) and evaluating each component in terms of its condition, i.e., good, marginal, poor. Condition is also reported in terms of dollars required to return component to good condition. Condition data is collected through direct inspection of all components. The IFS is primarily designed as a reporting system for actual maintenance requirements. When implemented, however, it may provide data useful for analyzing and projecting maintenance costs over time. It was this conjecture which led to the development of the proposed model.

In the course of developing this model, it became apparent that facilities could not be easily fit into traditional models of failure such as the exponential chance failure model or wearout failure model. A deterioration phenomenon, however, seemed to be common to most facility component wearout. In particular, there is seldom a point where one can identify a complete component, or facility, failure. Instead, the component characteristics describing its performance change gradually over time, making failure difficult to define in meaningful terms. The deterioration phenomenon, or parameter drift, is more difficult to deal with when applied to complex systems; the traditional model is not easily applied to situations where data gathering is automated and large amounts of data are encountered. In light of the present configuration of the IFS, from which input data will be taken, a discrete-state Markov model has been formulated, making use of existing data formats and compatible with the general scheme of IFS reporting.

NOMENCLATURE

$p_{ij}(k)$	one-step transition probability, the conditional probability that component k will go from state i to state j
$P(k)$	one-step transition matrix for component k
$\pi_j(k)$	steady-state probability that component k will be in state j in the long run
$[\pi_j(k)]$	steady-state probability vector for component k
$\mu_j(k)$	average recurrence time for component k to return to state j
$X_{hij}(k)$	the occurrence of a transition of component k from state i to state j, for the hth facility observed
$[o(k)]_i$	matrix of observations of transitions at data point i for component k
T_i	transition data matrix at data point i
$\hat{P}_{ij}(k)$	theoretical transition probabilities calculated from a polynomial model
$\hat{P}(k)$	theoretical transition matrix
β_i	parameter of polynomial equation
b_i	least squares estimate of parameter of polynomial equation
V_i	ith variable describing a component
$d_{ij}(k)$	the extent that the kth component in the ith state degrades the jth mission of a facility, expressed as a decimal fraction
$D(k)$	mission effects matrix for component k
$D_j(k)$	column vector of mission effects for mission j, component k
$F(k)$	ineffectiveness vector for component k

$f_j(k)$	ineffectiveness of component k on facility mission j, expressed as a decimal fraction
$e_j(k)$	effectiveness of co ent k on facility mission j, expressed as a decim. fraction
Q	matrix portion for absorbing Markov chain which does not contain absorption probabilities
I	identity matrix
N	$(I - Q)^{-1}$, or fundamental matrix of absorbing chain
R	absorption state vector portion of absorbing state transition matrix
$\underline{1}$	column vector of all 1's
NR	vector of mean absorption probabilities from any state
M	mean number of steps for absorption from any state

APPLICATION OF THE MARKOV MODEL

The typical facility as a system exhibits a continuous process of deterioration over time, a complex process in which a large number of interrelated performance characteristics change to produce overall reduced system effectiveness. The relatively complex nature of the facility system does not lend itself to the discrete failure model in which there are only two possible states, "failed" and "not failed," it being difficult even to define the term "failure."

Chance failure and wearout failure are treated primarily as discrete, attribute events; the process is one of counting rather than measuring. Failure rates, the parameter of the chance failure models, are defined as failures per unit time, or percent failures per unit time; wearout failure parameters are the mean wearout time and variance of the wearout time.

Deterioration, or degradation, is a different phenomenon in that the performance parameter of a component is measured over time, with the change in the parameter representing the deterioration of the component. Presumably, the parameter being measured contributes to the ability of the component to perform its function. If the parameter value falls outside some design limit, a deterioration failure occurs. In a deterioration failure model, the rate at which the parameter changes over time is a measure of the component reliability.

A stochastic deterioration model would define the performance characteristics of the system as random variables, where the parameters of the distributions of the performance characteristics are functions of time; for example, let q equal the component performance measure and let q have a normal distribution at any given time with mean μ_q and variance σ_q^2. Let these parameters

be a function of time,

$$\mu_q(t) = f(t) \tag{9}$$

$$\sigma_q^2(t) = s(t) \tag{10}$$

Let q_o = the design limit of the component.

The hazard function for the component would take the form

$$g(q) = \frac{1}{2\sqrt{s(t)}} e^{\frac{-[q - f(t)]^2}{2\sqrt{s(t)}}} \tag{11}$$

By integration, the failure time distribution function would be:

$$F(t) = \int_{t=0}^{t} \frac{1}{2\sqrt{s(\phi)}} \int_{q=0}^{q_o} e^{\frac{-[q - f(\phi)]^2}{2\sqrt{s(\phi)}}} dq\, d\phi \tag{12}$$

The deterioration model is more descriptive of the deterioration in a facility system, but it does not make easier the definition of the failed condition of the component. The deterioration model still permits only two distinct states and makes it difficult to adequately include the intermediate steps in the deterioration process, which are also of interest in arriving at policy decisions.

The Markov model falls somewhere between the chance failure model and the continuous deterioration model by providing for any number of discrete states into which the system may fall during the process of deterioration. The Markov model, unlike either of the others, looks at the *change* in state of the system and is in this sence a more dynamic model. This approach is particularly suited to the facility situation in providing very relevant management information in a form that is directly useful.

Use of the discrete state model for the deterioration phenomenon requires that the component performance continuum be divided into a number of mutually exclusive, but all-inclusive, states, and for practical purposes it is desired that these states be minimum in number while adequately representing the range of deterioration and in sufficient detail to avoid ambiguity in the definition of the states. It is assumed that this approximation to the continuous nature of the deterioration model will not seriously effect the results obtained with the model.

No attempt is made here to define all components and the corresponding states. General guidelines for component definitions are suggested as follows:

a. Component breakdown should be on a functional basis as each component relates to the facility as a whole, rather than according to shop responsibilities, size, capacity, etc. The general function for the component defined should be the same for all facilities regardless of type.

b. All components should be defined as the same level of detail; in this respect, a hierarchical or tree structure may be helpful.

c. Components should be primarily divided as building components and non-building components. The number of components which may be applied to both broad categories will probably be small.

d. Components should be mutually exclusive.

The definition of the component states is similar to the condition rating scheme used in the early versions of IFS [2]. This scheme defined four states into which the condition of all components would fall. Benchmarks for evaluation of component condition related the component condition classification with the effectiveness of the facility as a whole and did not

differentiate the component condition from the relative effect of any particular component on a particular facility mission.

The definition of component states suggested for the Markov model is performed independently of the consideration of the effect that a component condition may have on a facility as a whole and represents a significant improvement on the present condition rating scheme with the greatest improvement being that each component condition is given in terms of that component only. General guidelines for definition of the component states are suggested as follows:

a. Component states should be clearly defined in terms of measurable quantities such as capacity, or in clearly defined attribute terms.

b. Each state should be defined in terms of the particular component without regard to other components or facility functions.

c. States should be defined in sufficient detail to avoid ambiguity; unnecessary detail should be avoided to minimize complexity of the resulting state model.

It may be reasonably assumed that the state transition probabilities at any given time will depend upon a set of variables which determine the way the system will behave. Added dynamic dimension is achieved in the model by formulating the transition probabilities as functions of this set of variables, as

$$p_{ij} = f(\{V\}, \{\beta\}) \qquad (13)$$

where $\{V\}$ = the set of variables and

$\{\beta\}$ = a set of parameters of a polynomial model.

The polynomial model allows the inclusion of all relevant and significant factors which, logically, will have an impact on the deterioration of a facility.

It should be noted that the variables are defined for the facility as a whole, and the set of values of the variables will be roughly the same for all components; in practice, the values of the variables may differ for different components within the same facility, or may only differ from facility to facility, or may vary only from installation to installation. In all cases the set of variables arrived at will be the same for all installations, all facilities, and all components. It is, therefore, through the definitions of the components, the component states, and the significant variables, that the model achieves its generality, and through the application of the particular values of the variables and mission analysis (covered later) that the model is applied to specific circumstances.

Estimates of the parameters of the polynomial model are obtained by the method of least squares from the observation of samples of facilities over a "transition step," normally a change of state from one year to the next.

APPLICATION OF ABSORBING STATE MODEL TO FACILITIES

The absorbing state model has a somewhat different development. All though the basic states are the same as those listed above, there is one more state that must be added. This state, the absorbing state, will represent the point in time where the component of the facility has degraded to such a point where it is disposed of.

It should be noted that a particular component's disposal may or may not mean that the facility itself is disposed of. For example, a gas heating system may be replaced by an electrical heating system with no effect to the facility. However, if a roof is disposed of, the building itself will also be disposed of unless the roof is only being replaced; in this case, the component goes back to the best condition.

The probabilities for the absorbing chain are computed the same way the probabilities for the regular chain, with the exception that for the kth absorbing state, $p_{kk} = 1$ and p_{kj}, for all j is 0.

EXPERIMENTAL DESIGN

Experience with facility data has suggested that the variables which may determine the change in state of components may be divided into three broad categories with corresponding important variables in each category as follows:

1. Design Variables--the variables whose values are determined at the time the component is designed and constructed and remain unchanged throughout the life of the component, such as:
   (a) Type of construction--permanent or temporary, masonry or wood, etc.,
   (b) Size or capacity--square feet, Btu/hour, persons, etc., and
   (c) Year constructed--the age of the component.
2. Environmental Variables--describe the actual conditions under which the component is operated, such as:
   (a) Component loading--actual load being placed on the component as compared with designed capacity, and
   (b) Climactic conditions--heating degree-days, annual rainfall, etc.
3. Policy Variables--variables describing the operational and maintenance policies of the organization responsible for the operation and/or maintenance of the component, such as:
   (a) Dollars spent on preventive maintenance of the component and
   (b) Dollars spent on corrective maintenance of the component.

This list of variables, although reasonably complete without the benefit of extensive data analysis, should not be considered as

exhaustive by any means. All reasonably possible variables of cause should be included in the data gathering process so that all possible sources of variation may be accounted for. The larger the list, however, the more data points will be required in order to get independent estimates of the least squares parameters, and for this reason the variables included in the analysis should be chosen with care.

The data for the analysis should be collected according to a two-level factorial design scheme; in the case of the seven variables listed, a 2^{7-3} fractional factorial design could be used, and by confounding the main effects of three of the variables with third-order and fourth-order interactions, it is possible to gain independent estimates of main effects and all second-order interactions. An assumption required for the model is that all third- and higher-order interactions will be statistically insignificant, a reasonable assumption under the circumstances. The design matrix is shown on the following page. High and low levels of the variables should be chosen so that they lie at or near the extreme values to be reasonably encountered in practice.

The fractional factorial design will permit independent estimates of the least squares parameters of first- and second-order terms; use of the factorial also permits the inclusion of attribute-valued variables which have two levels in the model [3].

The response of the component to be observed is whether or not the component changed state from one condition evaluation to the next. Obviously, this requires that the same set of facilities and components be observed under the same or nearly same conditions for two years consecutively. The observed data would be recorded as a matrix as:

Design Matrix--2^{7-3} Fractional Factorial Design

Test	1	2	3	4	12	13	14	23	24	34	123	124	7=234	6=134	5=1234
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2	-	+	+	+	-	-	-	+	+	+	-	-	+	-	-
3	+	-	+	+	-	+	+	-	-	+	-	-	-	-	-
4	-	-	+	+	+	-	-	-	-	+	+	+	-	-	+
5	+	+	-	+	+	-	+	-	+	-	-	+	-	+	-
6	-	+	-	+	-	+	-	-	+	-	+	-	-	+	+
7	+	-	-	+	-	-	+	+	-	-	+	-	+	-	+
8	-	-	-	+	+	+	-	+	-	-	-	+	+	+	-
9	+	+	+	-	+	+	-	+	-	-	+	-	-	-	-
10	-	+	+	-	-	-	+	+	-	-	-	+	-	+	+
11	+	-	+	-	-	+	-	-	+	-	-	+	+	-	+
12	-	-	+	-	+	-	+	-	+	-	+	-	+	+	-
13	+	+	-	-	+	-	-	-	-	+	-	-	+	+	+
14	-	+	-	-	-	+	+	-	-	+	+	+	+	-	-
15	+	-	-	-	-	-	-	+	+	+	+	+	-	+	-
16	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+

$$[o(k)]_i = \begin{array}{c} \\ s_1 \\ s_2 \\ \vdots \\ s_m \end{array} \begin{array}{c} \begin{array}{cccc} s_1 & s_2 & \cdots & s_m \end{array} \\ \begin{bmatrix} \sum_{h=1}^{n} x_{h11}(k) & \sum_{h=1}^{n} x_{h12}(k) & \cdots & \sum_{h=1}^{n} x_{h1m}(k) \\ \sum_{h=1}^{n} x_{h21}(k) & \sum_{h=1}^{n} x_{h22}(k) & \cdots & \sum_{h=1}^{n} x_{h2m}(k) \\ \vdots & & & \\ \sum_{h=1}^{n} x_{hm1}(k) & \sum_{h=1}^{n} x_{hm2}(k) & \cdots & \sum_{h=1}^{n} x_{hmm}(k) \end{bmatrix} \end{array}$$

where n is the size of the total sample taken at data point i

$$x_{hij} = \begin{cases} 1 & \text{if component goes from state i to state j} \\ 0 & \text{otherwise} \end{cases}$$

It should be noted that in effect there will be three samples taken at each data point, inasmuch as the transition matrix which will result from the observation matrix is a series of conditional probabilities, and thus, each row of the transition matrix must add to 1.0.

From the observation matrix, taken at each data point of the factorial design, a transition matrix is calculated:

$$p_{ij}(k) = \frac{\sum_{h=1}^{n} x_{hij}(k)}{\sum_{j=1}^{m} \sum_{h=1}^{n} x_{hij}(k)} \qquad (14)$$

The resulting transition matrix is referred to as the transition data

matrix, and there will be a transition data matrix for each data point; in the case of the 2^{7-3} design, there will be 16 such data matrices:

T_i = transition data matrix at data point i

$$[p_{ij}(k)]_i = T_i \tag{15}$$

Because the data being collected will be of an attribute nature, a rather large sample will be required in order to maintain the desired precision of the estimated transition probabilities. The estimate $p_{ih}(k)$ will have a variance

$$p_{ij}(k) = \frac{p(1 - p)}{n} \tag{16}$$

where n = the total number of components which were initially in state i. The variance of the estimate p will be a maximum when $p_{ij}(k) = 0.50$; therefore, in order to have a precision of ±1 percent,

$$n \geq \frac{0.50^2}{0.10} \qquad n \geq 25 \tag{17}$$

This is not to suggest that a sample of at least 25 components in state 1, 25 in state 2, etc., be taken for purposes of precision; the total sample of components taken at a data point should be as random as possible in order to have some idea of the distribution of the initial states. The samples should be sufficiently large that there are about 25 components in each of the initial states; barring this, the resulting precision of the estimate should be kept in mind.

LEAST SQUARES MODEL

The data derived from the observations is used to estimate the parameters of a second-order least squares model as follows:

$$p_{ij}(k) = b_o + b_1 V_1 + b_2 V_2 + \ldots + b_n V_n + b_{12} V_1 V_2 + \ldots + b_{1n} V_1 V_n + \ldots b_{n-1,n} V_{n-1} V_n + \text{Residual.} \quad (18)$$

Estimates of parameters are obtained by the method of least squares in matrix notation:

$$\underline{b(k)} = (V'V)^{-1} Y_{ij}(k) \quad (19)$$

where $\underline{b(k)}$ = vector of parameter estimates

V = the design matrix

$Y_{ij}(k)$ = the vector of the $p_{ij}(k)$ estimates from all data points, where $y_h(k) = p_{ij}(k)$ at data point h

Analysis of the data by least squares will result in a set of polynomial models, one for each possible state transition:

N^2 = the number of state transitions, and least-squares polynomial models required, where N = the number of states.

This provides an additional reason for keeping the number of states at the minimum necessary.

With the parameter estimates, it is possible to calculate a transition matrix from a set of values for the independent variables. Because of the lack of fit and residual error inherent in the least-squares process, it may be necessary to adjust the derived transition probabilities so that

$$\sum_{j=1}^{m} p_{ij}(k) = 1.0 \quad (20)$$

To do this, a dummy term is introduced so that

$$\sum_{j=1}^{m} p_{ij}(k) - C = 1.0 \tag{21}$$

$$\sum_{j=1}^{m} p_{ij}(k) = 1.0 + C \tag{22}$$

$$\frac{\sum_{j=1}^{m} p_{ij}(k)}{1 + C} = \frac{1 + C}{1 + C} = 1.0 \tag{23}$$

Thus, dividing each element in a row of the derived transition matrix by the sum of the probabilities in that row will result in a corrected row of transition probabilities whose sum will be equal to 1.0.

MISSION ANALYSIS

Each component of a facility can be expected to effect the mission of the facility depending on the state that the component is in; the states of the component can be expected to have different impacts on the different missions for which a facility may be used. As an independent part of the analysis, and as a part of the generalized aspect of the model, it is necessary to assess in quantitative terms the impacts that the component states will have on the different missions. This multiple mission concept will apply most to the components which will apply to the buildings rather than to the non-building facilities. Non-building facilities will more likely have a sinlge mission or function.

The process for estimating these effects is, for the most part, subjective; the accuracy and reliability of the effectiveness measure obtained will largely be a function of the skill and experience of the person or persons making the estimates. The procedure and resulting mission effects matrix is easily understood and lends itself to easy correction and arrival at a consensus where there may be differences of opinion. In many cases, the degradation of the facility mission will be very closely correlated with the effectiveness of the facility as a whole, e.g., the electrical power distribution component in a missile launch facility, while in other cases there may not be a clear-cut relationship (e.g., with the roof component of a warehouse facility). It is hoped, however, that by making the mission analysis an independent undertaking from the condition of the component, that a maximum of objective and reliable mission effectiveness estimates will result.

At the onset, all facility missions should be listed to include all types of facilities which may contain the component in question. Next, these facility missions should be grouped according to any commonality of the specific function that the component is to perform (see numerical example). This will better relate the component to the mission of the facility and cut down the complexity of the estimating process.

When all missions of the facility have been listed, a fraction is assigned to each state of the component for each facility mission that may be encountered, the fraction indicating the degree to which the facility mission is degraded when the component is in a particular state, the resulting mission effects matrix D(k) will appear:

$$
\begin{array}{c} \\ \\ s_i = \\ \\ \\ \\ \end{array}
\begin{array}{c}
\text{mission } j = \\
\begin{array}{cccc} 1 & 2 & \cdots & a \end{array} \\
\begin{bmatrix}
d_{11}(k) & d_{12}(12) & \cdots & d_{1a}(k) \\
d_{21}(k) & d_{22}(k) & \cdots & d_{2a}(k) \\
\vdots & & & \\
d_{m1}(k) & d_{m2}(k) & \cdots & d_{ma}(k)
\end{bmatrix}
\end{array}
\begin{array}{l} \\ s_1 \\ s_2 \\ \\ \\ s_m \end{array}
$$

Mission Effects Matrix

The estimates $d_{ij}(k)$ are required to be in the range:

$$0 \leq d_{ij}(k) \leq 1.0 \qquad (24)$$

TOTAL COMPONENT EFFECTIVENESS

The effectivneness of a component may be computed from two determining inputs: the set of design, environmental, and policy variables which describe the component, and the mission of the facility or facilities which contain the component by way of the mission effects matrix.

The variable values are applied to the polynomial equations to arrive at a theoretical one-step transition matrix. After any necessary adjustments, the matrix $\hat{P}(k)$ is taken to a power, usually 5 or 6, sufficient to arrive at steady-state probabilities $[\pi(k)]$. These represent the long-run probabilities that the component, under the conditions as set by the values of the input variables, will be in a given state.

Multiplying the steady-state probability vector by the mission effects matrix results in an ineffectiveness vector, showing the long-run degradation of the component on the missions of facilities containing the component, as:

$$[\pi(k)] \cdot D(k) = F(k) \tag{25}$$

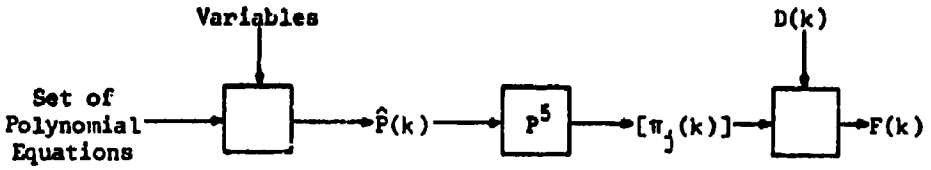
The component ineffectiveness is expressed as a decimal fraction. Multiplying the steady-state probability vector by a single column vector from D(k) will result in the element from the ineffectiveness vector for a particular mission in question:

$$[\pi(k)] \cdot D_j(k) = f_j(k) \text{ for the } j\underline{\text{th}} \text{ mission} \tag{26}$$

The effectiveness of the component, as applied to a facility with mission j, is:

$$e_j(k) = 1.0 - f_j(k) \tag{27}$$

$$0 \leq e_j(k) \leq 1.0 \tag{28}$$



Component Effectiveness Model

USING THE BETA DISTRIBUTION TO SUGGEST THE DISTRIBUTION OF COMPUTER ACCESS TIMES TO STORAGE

Walter D. Foster
Office of The Surgeon General
Information Systems Office

ABSTRACT

In comparing two or more candidate storage devices for all or parts of a computer configuration, problems arise in which it would be helpful if a distribution of access times to storage were available. For the case at hand, three values were listed by the manufacturer: maximum, minimum, and mean. The problem consisted of contriving estimates of several percentiles of a distribution purporting to represent the unknown distribution of access times.

Because the Beta distribution is bounded, is defined by mean and standard deviation as parameters (thru m and n), can take many degrees of skewness either positive or negative, and is extensively tabulated, it was selected as a basis for suggesting the requested percentiles. Its characteristics are presented along with helpful ideas for plotting.

What is lacking is any idea of how good the percentile estimates might be. Because actual data on which to base statistical estimates generally do not exist, the question of accuracy cannot be answered. An approach is offered by postulating that the true distribution is Beta in form with known mean but with poorly estimated standard deviation. The effect of error in estimating the standard deviation on the percentiles is given in tabular form. A second approach to the effect of error constitutes the question for the panel.

I. BACKGROUND AND OBJECTIVES:

In preparing to simulate the use of 1) Digital Equipment Corp's RPO-3 disk and 2) IBM's 3330 disk with the PDP-11 computer in GPSS on the IBM 360-30, a specific problem arose concerning the distribution of access times to disk memory for various storage locations on the disk and for a variety of information items to be retrieved.

PDP-11 literature gave the minimum access time as 7.5 msecs, the maximum as 55, and the average as 42 msec. With only this amount of information, the problem was to select or contrive a probability distribution which would serve as a first approximation to the (unknown) true distribution of access times and to obtain the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles of this distribution.

II. ANALYSIS

Because of the obvious skewness of the true distribution as indicated by the three given values of access time, any symmetrical distribution such as the normal or rectangular would not be appropriate. However, the Beta distribution,

$$B(m, n) = \int_0^1 k\, x^{m-1} (1 - x)^{n-1}\, dx \quad \text{- - - - - - - - - - - - - - - (1)}$$

which is a bounded, unimodal, continuous distribution with domain zero to unity and parameters m and n, is capable of exhibiting skewness to the right ($m < n$), left ($m > n$), or symmetry ($m = n$). It has mean,

$$\mu = m/(m + n) \quad \text{- (2)}$$

mode,

$$\text{mode} = (m - 1)/(m + n - 2) \quad \text{- (3)}$$

and variance,

$$\sigma^2 = mn/(m + n)^2 (m + n + 1) \quad \text{- - - - - - - - - - - - - - - - - - - (4)}$$

This distribution was deemed able to provide a first approximation to the actual distribution of the access times under the assumptions that the true distribution is unimodal, continuous, bounded, and reasonably regular in its characteristics. The following mathematical development shows the basis for the computation of the desired percentiles.

The first step was to transform the scale of the original variable, t (for access time in milliseconds), to that of X with upper and lower bounds of zero and unity as required by the Beta distribution. Such a transformation is given by

$$X = (t - a)/(b - a) \quad \text{- (5)}$$

where

a = lower bound in t-scale, and
b = upper bound in t-scale, such that when

t = a, X = 0, and when
t = b, X = 1, as required.

The opportunity to estimate the Beta distribution parameters, m and n, lay in solving equations (2) and (4) simultaneously under the assumptions that the mean, μ, and variance, σ^2, were known and the distribution was known to be the Beta in form. The mean was one of the given values, 42 in the t-scale and .7263 in the X-scale according to (5). Knowledge of the variance was contrived according to the following development. It was assumed that the essential domain of both the actual and the fitted Beta distributions was equivalent to six standard deviations. A precedent for this assumption in using the Beta distribution exists in the statistical theory underlying the development of the PERT charting. Similarly, the domain of any normal distribution is often taken to be essentially equivalent to six standard deviations. Therefore, we set the range of the equivalent Beta distribution equal to six standard deviations:

$6\sigma = 1 - 0$ from which

$$\sigma = 1/6.$$

Simultaneous solution of equations (2) and (4) yielded explicit estimates of the two parameters as follows:

$$m = (\mu/\sigma)^2 (1 - \mu) - \mu \quad \text{- (6)}$$

$$n = m(1/\mu - 1) \quad \text{- (7)}$$

with numerical estimates

$m \doteq 4.713$ and

$n \doteq 1.6847.$

With the fitting of the equivalent beta distribution, it remained to

find the desired percentiles. These were available in Table III. 10, Percentage Points of the Beta Distribution 1/ . As required, we set

$v_1 = 2n = 3.37$ and

$v_2 = 2m = 8.94$

and employed double interpolation to create the following table of X for the requested percentiles:

Percentile	5	10	25	50	75	90	95
X	.4145	.491	.621	.753	.859	.922	.951
t	27.2	30.8	37.0	43.3	48.3	51.3	52.7

Conversion of the X values to the t-scale was accomplished by the inverse solution of (5) giving

$$t = (b - a)X + a \quad \text{- (8)}$$

For purposes of illustration, the ordinates of the fitted Beta distribution were computed from the first derivative of (1) for arbitrary increments of X. Figure 1 shows this distribution with both the X and t-scales.

III. HOW ACCURATE ARE THE PERCENTILES?

What is lacking is any idea of how good the percentile estimates might be. Because actual data on which to base statistical estimates did not exist, the question of accuracy cannot be answered. An approach is offered by postulating that the true distribution is Beta in form with known mean but with poorly estimated standard deviation.

Two approaches have been followed in assessing the effect on the percentiles of a poorly estimated standard deviation.

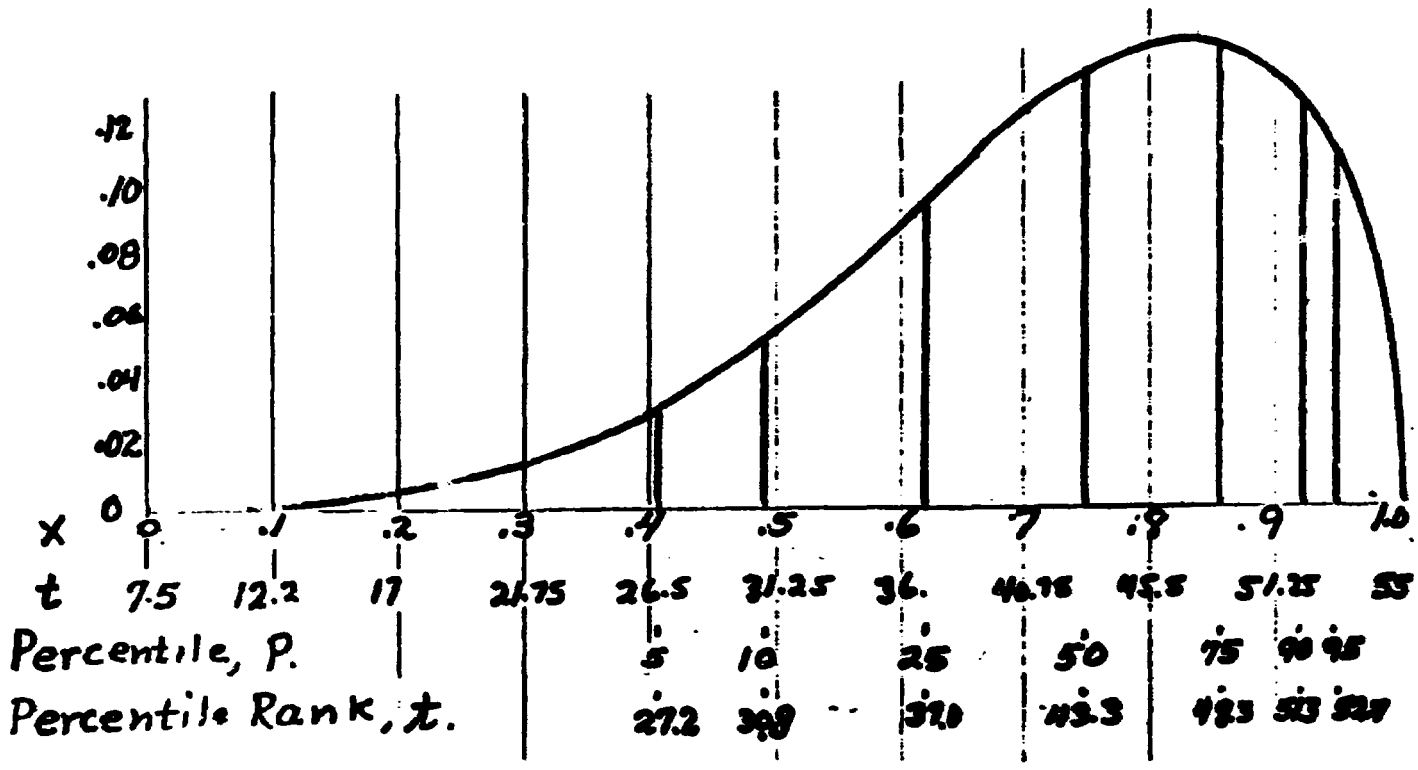
The first consists of a simple sensitivity analysis in which the 25-th percentile has been computed for a two-way table of mean and standard deviation. This is given in the following table where it is easy to perceive the change in the 25-th percentile for changes in either the mean or the standard deviation or both.

TABLE I. VALUES OF P_{25} FOR VARIOUS μ and σ

		Mean				
		.70	.75	.80	.85	.90
σ	.05	.666	.710	.754	.804	.868
	.10	.634	.686	.739	.795	.858
	.15	.601	.657	.716	.784	
	.20	.571	.630	.700		
	.25	.543	.597			

Suppose, for example, that the true distribution is Beta in form with mean .75 and standard deviation .25. However, it is assumed that the mean is .75 but the standard deviation is .15. Then from the table, it can be seen that P_{25} is thought to be .657 when it really is .597. Similar effects can be obtained for a change in the mean or a change in both mean and standard deviation.

FIGURE I. FITTED BETA DISTRIBUTION WITH PARAMETERS m = 4.47 and n = 1.68.



Percentiles are given in the 3rd row of abscissa values; corresponding t-scale values are given in the 4th row.

X-scale in units for Beta Distribution.
t-scale in milliseconds for original data.

The second approach has the same objective, what is the effect of a change in standard deviation on the percentile. It consists of applying the delta process to the Beta distribution. The question to the panel consists of both the need for this approach and its appropriateness.

Let us denote the change in X by ΔX as effected by a change in the standard deviation thru the Beta parameters m and n. By requiring the original Beta integral to have value equal to the desired percentile, e.g. .25, as well as the augmented Beta integral, we have

$$k\int_0^X t^{m-1}(1-t)^{n-1}dt = K\int_0^{X+\Delta X} t^{m+\Delta m-1}(1-t)^{n+\Delta n-1}dt = P$$

where $k = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)}$ and

$K = \frac{\Gamma(m+n+\Delta m+\Delta n)}{\Gamma(m+\Delta m)\Gamma(n+\Delta n)}$ and

P = specified percentile

Differentiation with respect to X gives

$$k(X)^{m-1}(1-X)^{n-1} = K(X+\Delta X)^{m+\Delta m-1}(1-X-\Delta X)^{n+\Delta n-1}$$

and

$$\frac{k}{K} = \left(\frac{X+\Delta X}{X}\right)^{m-1}(X+\Delta X)^{\Delta m}\left(\frac{1-X-\Delta X}{1-X}\right)^{n-1}(1-X-\Delta X)^{\Delta n}$$

and

$$\frac{k}{K} = \left(1+\frac{\Delta X}{X}\right)^{m-1}\left[\left(1+\frac{\Delta X}{X}\right)^{\Delta m}X^{\Delta m}\right]\left(1-\frac{\Delta X}{1-X}\right)^{n-1}\left[\left(1-\frac{\Delta X}{1-X}\right)^{\Delta n}(1-X)^{\Delta n}\right]$$

and

$$\frac{k}{K} = \left(1+\frac{\Delta X}{X}\right)^{m+\Delta m-1}X^{\Delta m}\left(1-\frac{\Delta X}{1-X}\right)^{n+\Delta n-1}(1-X)^{\Delta n}$$

Then $\ell n \frac{k}{R} = (m+\Delta m-1)\ell n(1+\frac{\Delta X}{X})+\Delta m\ \ell nX + (n+\Delta n-1)\ell n(1-\frac{\Delta X}{1-X}) + \Delta n\ \ell n(1-X)$

which becomes by taking the first term in the series expansion for $\ell n\ w$:

$$\ell n \frac{k}{R} = (m+\Delta m-1)\ \frac{\Delta X}{X} + \ell n\ X^{\Delta m} + (n+\Delta n-1)(\frac{-\Delta X}{1-X}) + \ell n(1-X)^{\Delta n}$$

whence

$$\Delta X = \frac{X(1-X)\ell n\left[\frac{k}{R} X^{-\Delta m}(1-X)^{-\Delta n}\right]}{\left[(1-X)(m+\Delta m-1)-X(n+\Delta n-1)\right]} \qquad \ldots\ldots\ldots\ldots (9)$$

Evaluation of Δm and Δn in (9) was accomplished by augmenting σ in equation (6):

$$m+\Delta m = (\frac{\mu}{\sigma+\Delta\sigma})^2\ (1-\mu)-\mu \qquad \text{whence}$$

$$\Delta m = (1-\mu)\left[(\frac{\mu}{\sigma+\Delta\sigma})^2 - (\frac{\mu}{\sigma})^2\right] \qquad \ldots\ldots\ldots\ldots (10)$$

Similarly from equation (7), we have

$$\Delta n = \Delta m\ (\frac{1}{\mu} - 1) \qquad \ldots\ldots\ldots\ldots (11)$$

The evaluation of equation (9) caused considerable consternation to my little desk-top Olivetti 101 and despite its valiant efforts, I think I managed to gum it up by errors of my own. At least I haven't been able to duplicate any effect in Table 1 as yet.

I had hoped that a formulation like that in equation (9) would provide some insight on the effect of $\Delta\sigma$ on ΔX and possibly a not-too-nasty computation of it. Perhaps your answer to my question is the same as mine... Uggghhhh!

1/ Chemical Rubber Company Handbook of Tables for Probability and Statistics.

ON BIMODALITY IN PARAMETERIZED ATMOSPHERIC MODELS AND SOME SOLUTIONS TO AVOID IT

Oskar Essenwanger
Physical Sciences Directorate
US Army Missile Research, Development and Engineering Laboratory
US Army Missile Command
Redstone Arsenal, Alabama 35809

ABSTRACT. The bimodality in the frequency distributions of atmospheric data samples is examined. If the bimodality is generated by a mixture of populations based on differences in physical behavior, it may only sometimes be possible to separate the individual collectives a' priori.

A second class of data exists where bimodality in the frequency distributions is also caused by physical behavior, but is produced by either parameterization or design requirements. This bimodality can usually be avoided, but the methods of solution may not always be trivial.

Some bimodalities of frequency distributions for atmospheric elements fall into this second category. It is illustrated how to eliminate the bimodality in case of an angular variate profile such as in the parameterization of the wind direction profile as a function of altitude. Another example is taken from engineering requirements where the principle of stationarity of time series or homogeneity of climate is violated.

1. INTRODUCTION. Ordinarily it is expected that sampling from one population should lead to unimodal frequency density distributions. Bi- or multimodality in the sample frequency is therefore attributed to random effects of data sampling and can generally be neglected by the statistician.

The problem becomes more serious when the physical background for data sampling is heterogeneous, and consideration must be given to the fact that the sample may comprise more than one individual population. This mixture of populations could lead to multimodality in the frequency distribution. In this case, the individual peaks in the frequency density are statistically significant, and the distribution can no longer be treated by standard statistical methods.

Although advanced techniques have been developed for treating mixed distributions (e.g. Pearson, 1894; Doetsch, 1928, 1936; Essenwanger, 1955a and others) every bimodal or multimodal distribution should not automatically be considered as in the group of mixtures. The history of the data sample should be analyzed, and the question investigated whether the physical background or behavior of the observed property justifies the classification as a mixture of populations. Even though multimodality may be produced by differences of physical behavior unimodality can sometimes be restored by appropriate procedures. These procedures for the elimination of mixtures in populations may sometimes be very intricate.

The following discussions illustrate the case of a bimodality generated for an angular variate and some techniques of circumventing this bimodality.

A second example is presented which frequently occurs in the establishment of design criteria for systems analysis. Input requirements for atmospheric data may have been requested which ignore the stationarity or homogeneity such as in atmospheric time series or in the uniformity of the climatic regime. Multimodality in the frequency density will be the result of this neglect. In this case, efforts should be made by the climatologist to achieve modification of the requirements.

The outcome may be beneficial to both the engineer and climatologist. Usually the result only needs a simple treatment by standard statistical procedures with a meaningful interpretation of the statistical parameters with respect to physical properties.

2. EXAMPLES OF BI- AND MULTIMODALITY. In the following examples the wind as an atmospheric element has been selected. Although cases of bi- or multimodality are not restricted to this atmospheric parameter, quantities of multidimensional proportion are more likely to develop a heterogeneous background. A complex treatment may enhance the chances that physical conditions for one of the parameters involved become dissimilar as is the case for the mathematical description of the wind direction profile presented below.

Complexity of description alone, however, does not necessarily lead to the generation of bimodality. E.g. simple models of surface temperature disclose multiple peaks as can be found in numerous references, among them this writer's articles in 1955b.

Even in the case where bimodality is being introduced by the mathematical or statistical tool under consideration, it should be pointed out that the effect on the frequency distribution of the data sample cannot always be predicted a' priori. In final analysis, it is the physical behavior which causes these multiple peaks. The bimodality by the left or right turn of the wind direction profile as a function of altitude or the non-stationarity of meteorological time series is a physical phenomena.

2.a. The Wind Direction Profile. Before the actual bimodality of the frequency distribution of the characteristic parameter for the wind direction profile is illustrated, a few comments on parameterization may be appropriate.

While wind speed, temperature and density can be treated with straight-forward methods, the analysis of the wind direction requires some definitions. It is self-evident that the wind velocity could be separated into zonal and meridional components whose profile analysis needs little further explanation. These profiles would be, however, a function of the coordinate system as they differ when new references are established. Nonetheless, zonal and meridional components are coordinates of special interest to the meteorologist. The rare separation into wind speed and direction can also be attributed, however, to the difficulties arising in statistical treatment of angular variates.

The first problem is the assignment of a direction for calms. Otherwise they would appear as an angle of zero degrees according to the reported code. Without transformation the calms may create excessive outliers in profile analysis. It was decided to perform a linear vertical inter- or extrapolation (end points) of the profile. Although a higher order curve fitting was studied the outcome displayed no significant difference from the linear system and could be neglected for practical purposes.

The second problem is the elimination of the discontinuity at 0 or 360 degrees. This goal is achieved by restricting the maximum magnitude of the angular differences between two altitude levels to 180° and adjusting the subsequent directional values by adding or subtracting 360 degrees. (E.g. the sequence 350, 10 would be converted to 350, 370 but 10, 350 would be changed to 10, -10, etc.)

These two conditions eliminate most of the difficulties in directional profile analysis. One remaining problem is illustrated in Figure 1. Two directional profiles are depicted at Chateauroux in July, where the wind velocity changes from a westerly regime in the troposphere and lower stratosphere to easterlies above 20 km. No complication would arise below 20 km. The important set of surface to 25 km profiles exceeds this altitude threshold, however. This bimodality is also absent in the winter months.

While the profile of 21 July veers with height to east, the 19 July profile turns counterclockwise into easterlies. This difference in the directional shifts creates a bimodality in the distribution of the slope of a polynomial representation (Fig. 2). The mean slope resulting from the bimodality in the distribution of the linear coefficient would provide an analytical profile with no turn to the east in the layers above 20 km. Although this "mathematical compromise" is correct from a formalistic viewpoint of putting a mathematical formula into practice, the physical behavior would be completely suppressed.

A study of the average wind direction by individual altitude levels (see Essenwanger 1961, 1964) reveals that "backing" of the profile is supported by the turn of the mean direction (Fig. 1). A procedure was designed to relax the second condition (limitation of $|\Delta\theta|$ to 180°) for one point of the profile. This is accomplished by the establishment of a consistent profile, first with the surface direction and then with the angle of the top altitude (i.e. 25 km) as the initial reference. The two profiles coincide or are $n \cdot 360$ degrees apart. Then the two directional profiles are compared with the mean profile. Using a point by point selection at every altitude level each point with the minimum departure from the mean directional profile is selected. In the case of conformity of the two constructed profiles (from the surface or from the top altitude) no change or switch will be necessary. If the profiles are not congruent a switch from one profile to the other will take place at one point. This point will represent the minimum deviation from condition two ($|\Delta\theta| \leq 180^{\circ}$) in other words the least relaxation (see Fig. 3).

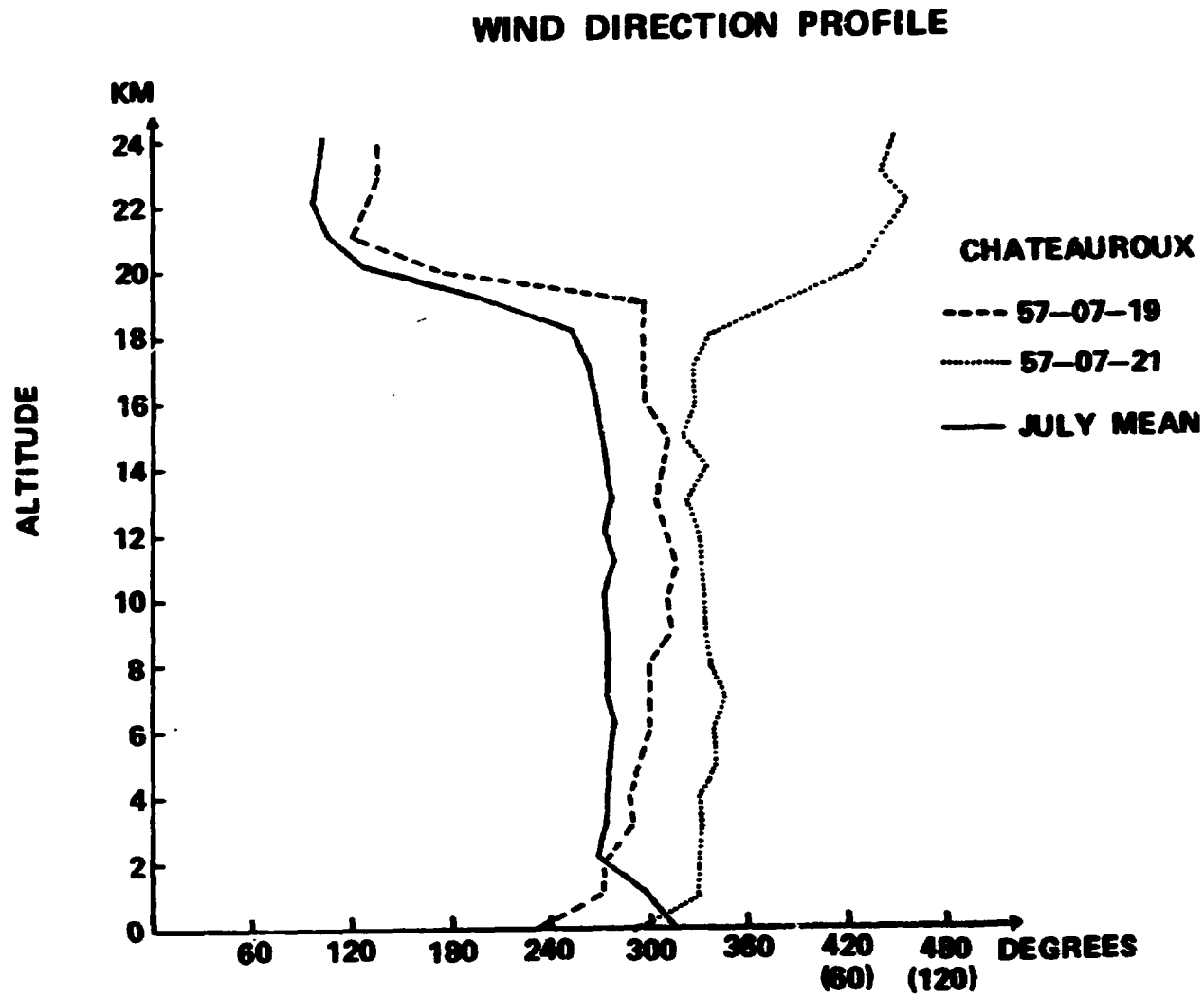

WIND DIRECTION PROFILE
KM
ALTITUDE
24
22
20
18
16
14
12
10
8
6
4
2
0
60
120
180
240
300
360
420
(60)
480
(120)
DEGREES
CHATEAUROUX
57–07–19
57–07–21
JULY MEAN

FIGURE 1

After this step was introduced the bimodality disappeared (see Fig. 2), and the new mean slope reflects the shift of the wind direction from west to east. Now the profile analysis can proceed in a manner similar to procedures outlined for the wind speed or other atmospheric elements (see Essenwanger, 1972).

First a mathematical description is obtained by

$$\theta(h) = D_0 + D_1\phi_1(h) + D_2\phi_2(h) + \ldots + D_n\phi_n(h) \qquad (1)$$

The θ denotes the wind direction as a function of the altitude as described above, and the ϕ_i terms are Tchebychef orthogonal polynomials. The D_0 as the mean direction may exceed 360 degrees from the conversion of the directional profile into a consistent form, but can be converted without any effect into the boundaries of 0 through 360 degrees.

The final model assumes the form

$$\theta(h) = (D_1 - \bar{D}_1)[d_0 + \phi_1(h) + d_2\phi_2(h) + \ldots + d_n\phi_n(h)] + \bar{\theta}(h) \qquad (2)$$

which again displays only one variate whose frequency distribution can be found. After the bimodality was eliminated, these distributions exhibited one central peak, and could be well fitted by a Weibull model.

In some instances

$$\bar{\theta}(h) \sim \bar{D}_1[d_0 + \phi_1(h) + d_2\phi_2(h) + \ldots + d_n\phi_n(h)] \qquad (3)$$

and the mean term cancels out in equation (2). Then

$$\theta(h) = D_1[d_0 + \phi_1(h) + d_2\phi_2(h) + \ldots + d_n\phi_n(h)] \qquad (2a)$$

As it has been demonstrated, the bimodality introduced by backing or veering of the wind with altitude can be streamlined to provide a unimodal frequency density. The reconstructed analytical direction profile now reflects the proper change from the westerlies in the lower troposphere to the easterlies in the 20-25 km altitude layer during July at Chateauroux. Thus the introduced change in the profile analysis procedures has not only eliminated the bimodality, it has also led to a more appropriate and simpler representation of the physical background in general.

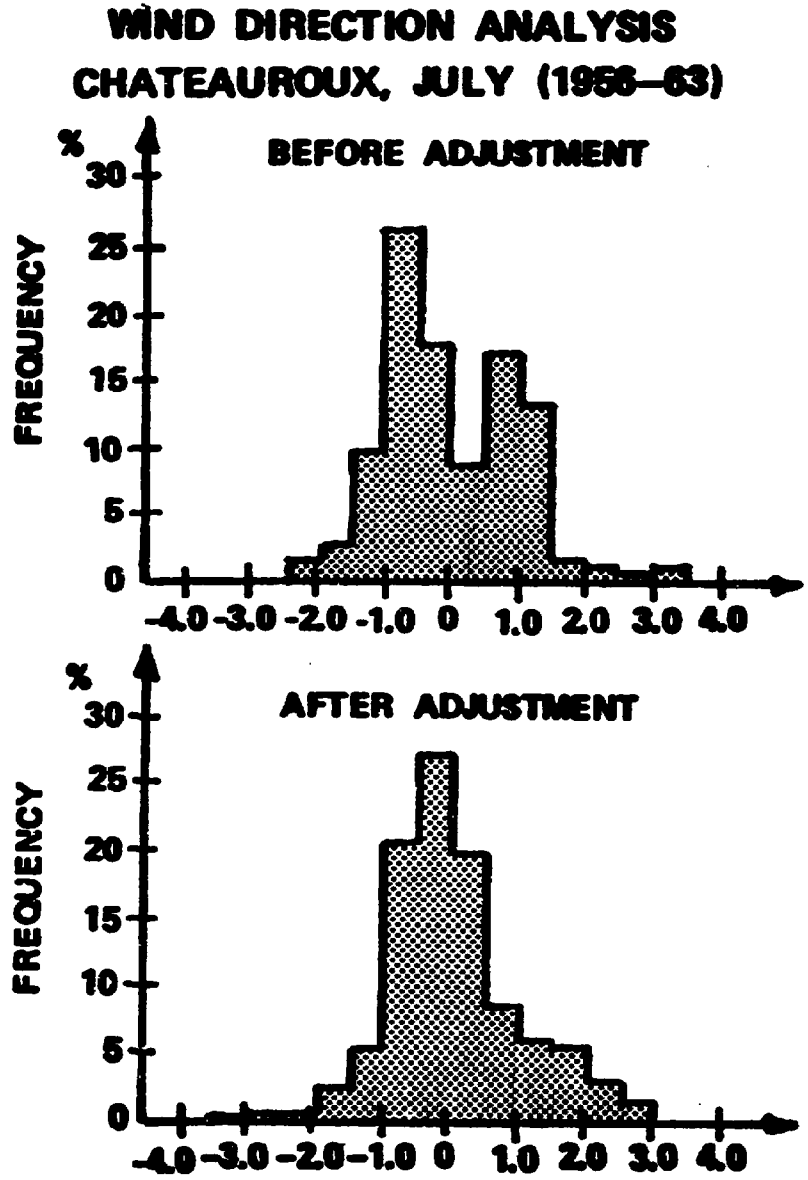

WIND DIRECTION ANALYSIS
CHATEAUROUX, JULY (1956–63)
BEFORE ADJUSTMENT
%
FREQUENCY
30
25
20
15
10
5
0
-4.0 -3.0 -2.0 -1.0 0 1.0 2.0 3.0 4.0
AFTER ADJUSTMENT
%
FREQUENCY
30
25
20
15
10
5
0
-4.0 -3.0 -2.0 -1.0 0 1.0 2.0 3.0 4.0

FIGURE 2

WIND DIRECTION PROFILE, 57–07–21

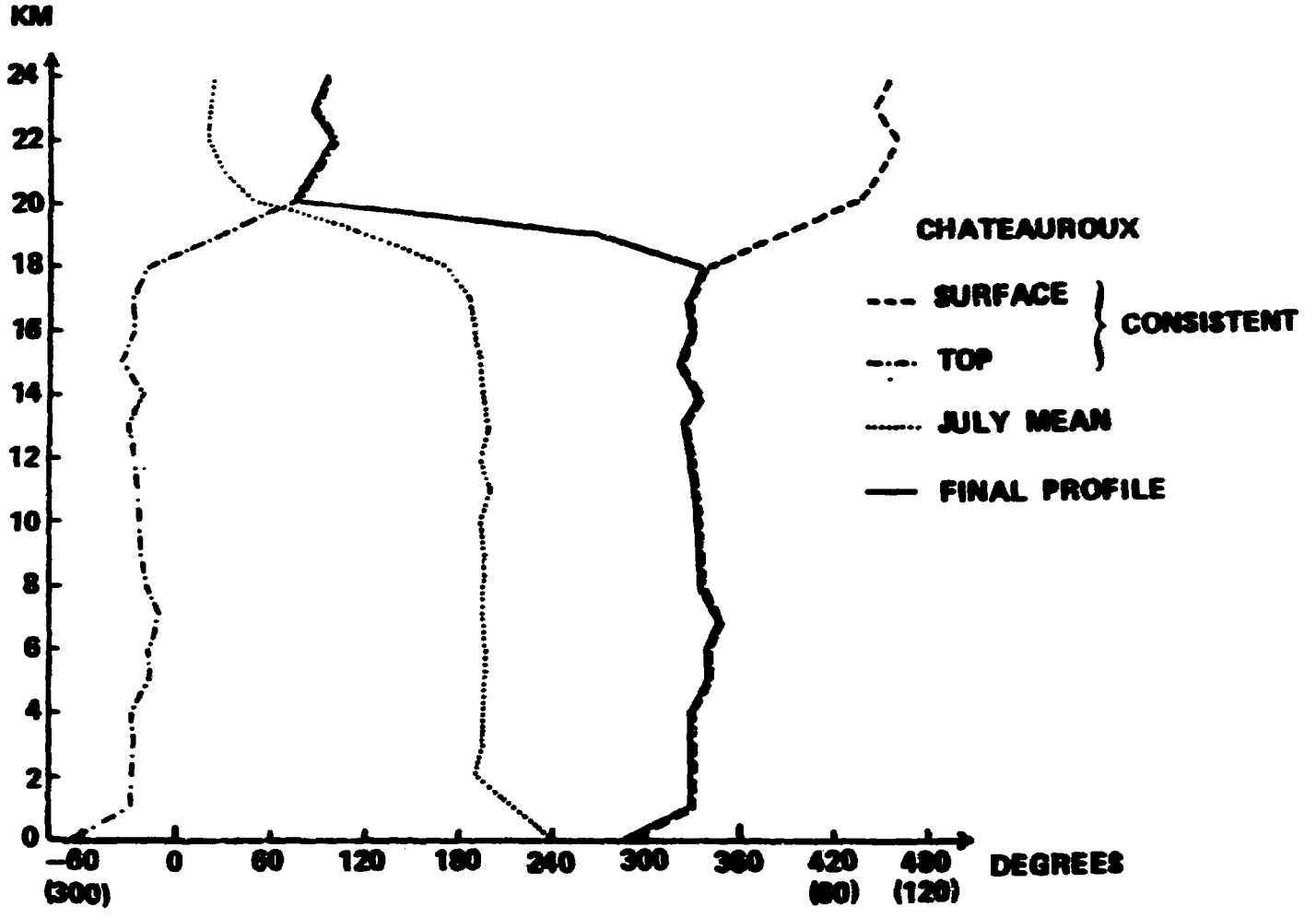


FIGURE 3

It should be emphasized that the adopted additional condition has not eliminated veering or backing of the wind with altitude. First, examination of the adjusted frequency in Figure 2 reveals that deviations from the mean are positive and negative. This fact indicates deviations from the mean conditions which still remain. The major change took place, however, in the alignment with the mean conditions (Table 1). The seasonal shift from westerlies in winter to easterlies in summer at 20-25 km altitude at Chateauroux, which is of interest here, is now reflected in the seasonal variation of the mean coefficients. This result can be much more easily interpreted and is meaningful because the mean conditions are most frequently encountered.

Table 1. Mean Coefficients for Wind Direction Profile at Chateauroux.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$\bar{D}_o$	329	303	311	305	316	240	258	245	273	290	276	318 degrees
$\bar{D}_1$	-7.3	-.1	3.0	4.1	9.5	-16.0	-15.5	-12.6	1.1	6.0	7.2	-8.0 degrees

$(\bar{\theta}_h = \bar{D}_o + \bar{D}_1 \phi_1)$

2.b. Stationarity of Meteorological Time Series. This second point of caution, the non-stationary behavior of meteorological elements, is well known and has also been treated by other authors such as Godske (1962,1966a,1966b) etc. It is especially applicable to the derivation of characteristic variates for meteorological elements (see Essenwanger, 1972). The goal of the reduction is the establishment of a variate whose frequency distribution is simple or at least unimodal.

The elimination of one kind of bimodality was discussed in section 1. In that case the bimodality was caused by the peculiarity of an angular variate, such as the wind direction, and the physical behavior in terms of the general circulation.

The frequency distribution of the characteristic variates may also disclose bimodality caused by the non-stationary behavior of atmospheric parameters. Let us assume that we were to combine the observational data of the entire year to derive an "annual" average, as it is favorably looked upon mostly by engineers.

Formalistically the techniques described by the author (1972) can be applied to the observational data, and schematically a variate with annual constants k_i (see eqn. 4 below) can be obtained. Unless this combination is justified by the climatic homogeneity or stationarity within the selected time frame for the meteorological elements under consideration the resulting frequency distribution of the characteristic variate may not be unimodal or may consist of a mixture of frequency distributions whose statistical treatment is as difficult as the tool of multivariate analysis. We would have gained very little in comparison with a multivariate system.

An example is now illustrated in Figure 4. The surface to 25 km wind profile has been represented by

$$V_n = A_o[1 + k_1 \sin(\alpha_n + \beta_1) + k_2 \sin(2\alpha_n + \beta_2) + \ldots] \qquad (4)$$

where the k_i denotes constants derived over a "stationary" time period such as a month or season.

The frequency distribution of the A_o coefficient for Montgomery, Alabama, is given in Figure 4, where the middle and lower diagram display the winter and summer season, respectively. Combination of the frequencies for an annual representation inevitably leads to a bimodal frequency density as exhibited in the top diagram of Figure 4.

It is evident that this bimodality must be attributed to the violation of the principle of stationarity. In fact, a closer examination of the middle diagram (Fig. 4) from the winter season reveals that even the winter season may comprise an inhomogeneous time period. The question of whether this is the case cannot be answered from this diagram alone. It would be necessary to break the frequency down into the individual months. Another check is the fitting of the frequency density by a statistical model (such as the Weibull frequency) and the examining of the statistical significance of the departure from the model by using established statistical test procedures. The latter procedure was chosen. In this particular case it was possible to conclude that the deviations were within the margins of non-significance. Since statistical tests are a function of the number of observations, however, re-examination with a longer period of record may not confirm this result.

FREQUENCY DISTRIBUTION OF A_0 COEFFICIENT,
SURFACE TO 25 KM WIND SPEED PROFILE
MONTGOMERY (ALABAMA), 1956–64

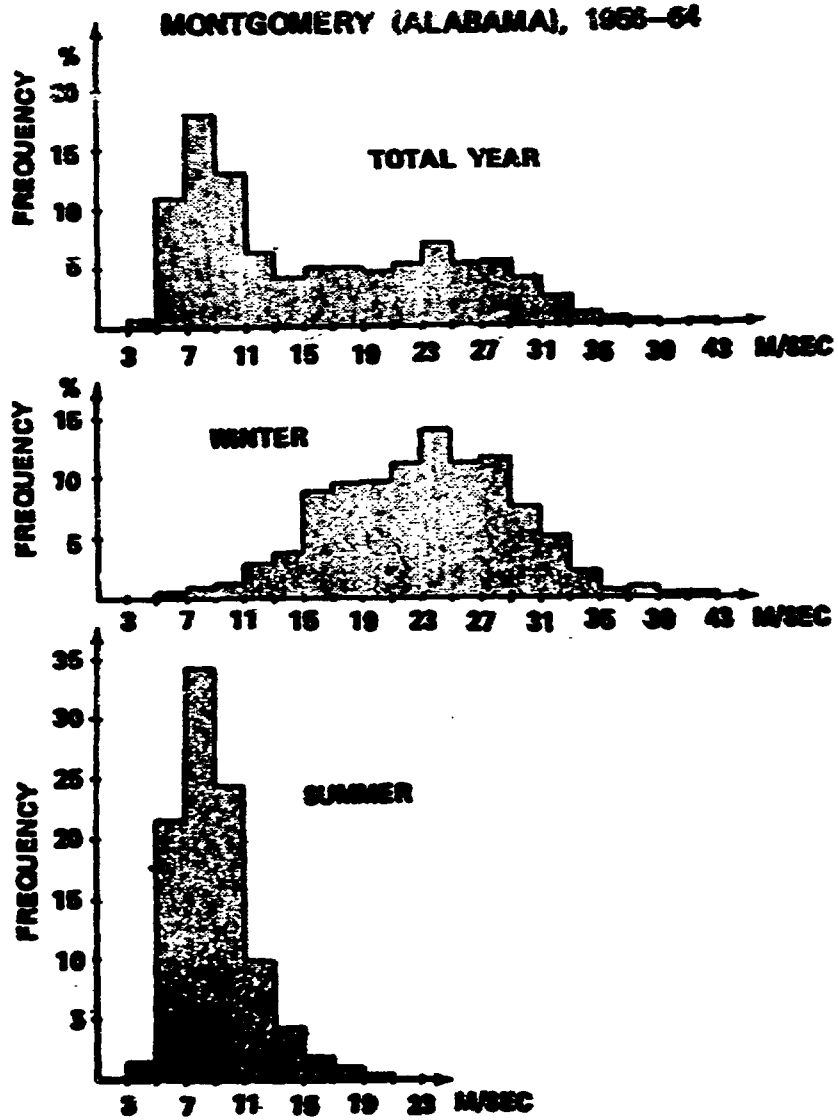


FIGURE 4

The displayed bimodality in the frequency density of the annual summary for the characteristic coefficient of the surface to 25 km profile was easy to recognise by the two peaks. A further example is added from the analysis of the characteristic coefficient for the surface to 16 km profile. In this particular case the multimodality is not striking at the first glance (see Fig. 5). Although the expanded peak region as illustrated by the histogram of the frequency density for the summer season at Chateauroux may create suspicion, only the detailed analysis with the individual months and the statistical test can reveal the heterogeneity. As discerned by the three diagrams June, July and August, only the two summer months July and August indicate some kind of a uniform stationary regime. During June apparently a different pattern of the general circulation exists which generates a different set of wind profiles (surface to 16 km) than in the subsequent months. Again, final judgment must be reserved until the study is repeated with a longer period of record.

The statistical modelling by eqn. (4) demonstrates the advantage of a representation by a simple meaningful mathematical characteristic for a multivariate system. Non-stationary behavior can be readily discovered by the elementary analysis of the frequency distribution of this characteristic parameter while complex matrix tabulations of correlations or vector notations usually would not permit insight into stationarity.

Preservation of stationarity and homogeneity of the meteorological time series may at first sight appear to the engineer to be an inconvenience rather than a necessity because it leads to more than one frequency model. It should be realised that the trade-off is the gain of a meaningful statistical representation by a unimodal frequency. Even though more than one monthly or seasonal frequency needs to be studied in systems analysis or design, the benefit of an interpretation of the statistical characteristics in terms of a unimodal distribution may by far exceed the lower computer costs for a meaningless statistical characteristic from one multipeaked model. Erroneous conclusions drawn from non-representative statistics may lead to extra costs which more than offset the savings intended by utilisation of an annual average, not to mention the possibility of improper functioning of a system. The exact penalty being paid for the malfunction or the inability to achieve the stated requirements under true atmospheric conditions can sometimes not be determined.

FREQUENCY DISTRIBUTION OF A_0 COEFFICIENT, SURFACE TO 16 KM WIND SPEED PROFILE CHATEAUROUX (FRANCE), 1966–84

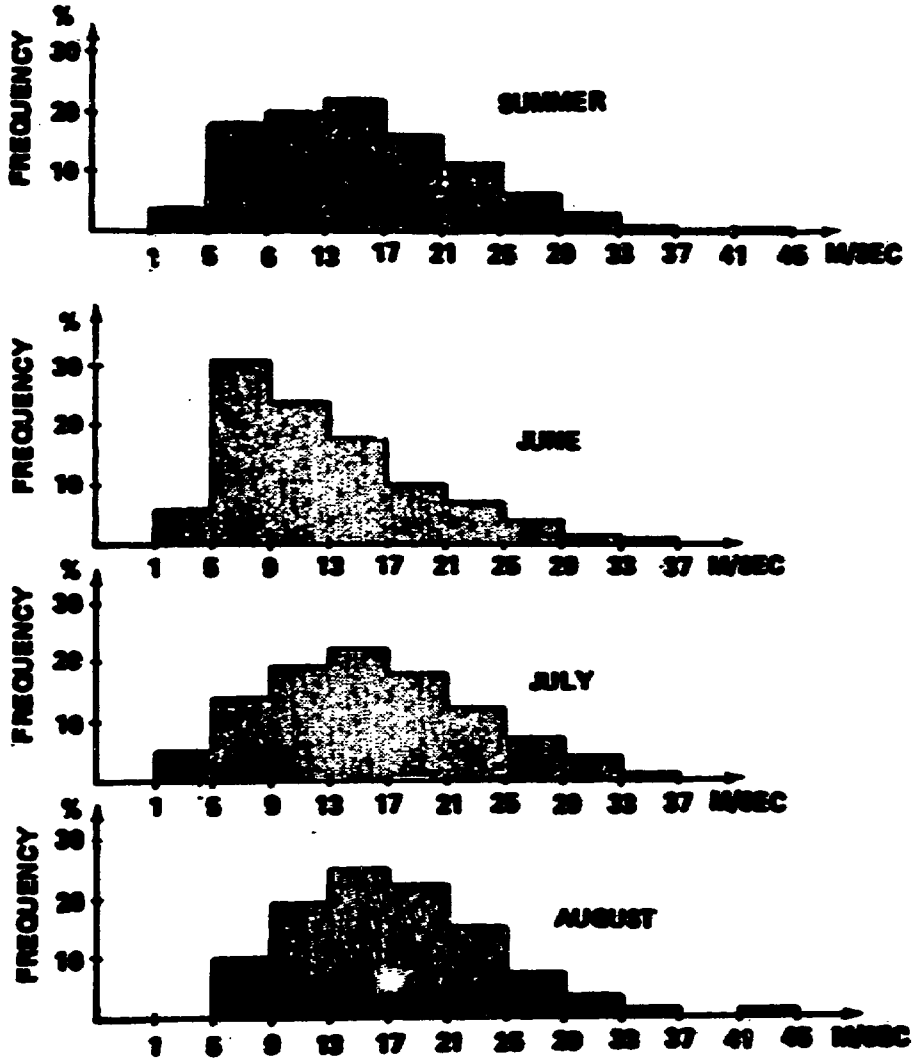


FIGURE 5

3. CONCLUSIONS. The bi- or multimodality was investigated for some frequency density distributions from data samples of atmospheric parameters for system analysis and design. Since multimodality may lead to misinterpretation of statistical characteristics the reason for this multiple peakedness is very important.

The author (e.g. 1955) has discussed in some articles that the heterogeneous physical background may lead to a mixture of populations and consequently to multiple peaks in the frequency density. In this case the only procedure for better statistical representation than a meaningless approximation by a unimodal frequency is usually the decomposing into partial components.

A second cause for bimodality may be the parameterisation or the stationarity. As illustrated in the preceding sections, this type of bimodality can be avoided although the method of solution may not always be trivial.

In the first example the bimodality is introduced by the method of establishing a characteristic coefficient for mathematical representation of an angular variate profile (surface to 25 km). The technique to avoid the bimodality was successful in streamlining the wind profile spiral to turn most frequently in the direction of a mean profile defined from individual altitude level statistics.

In the second example it is shown that violation of the principle of stationarity by engineering requirements leads inevitably to bimodality which can only be avoided by modification of the requirement to account for the reality of atmospheric behavior. Then unimodal frequency densities will be restored. This correction may sometimes be accepted very reluctantly because it may lead to more computer work in system analysis or expanded studies for design. The benefits should be obvious, however. The calculations of statistical characteristics peculiar to a unimodal frequency may lead in the case of multiple frequency peaks to a meaningless interpretation and consequently to improper functioning of the designed system. Thus atmospheric conditions should always be examined for their agreement with reality, and some pitfalls can be avoided by proper consideration of the physical behavior of atmospheric variables.

REFERENCES CITED:

DOETSCH, G., 1928. Die Elimination des Doppler effekts bei spektroskopischen Feinstrukturen und exakte Bestimmung der Komponenten.
Zeitschr.f.Physik, Vol. 49, p. 705-730.

DOETSCH, G., 1936. Zerlegung einer Funktion in Gauss'sche Fehlerkurven und zeitliche Zurückverfolgung eines Temperaturzustandes.
Math. Zeitschr. Vol. 41, p. 283-318.

ESSENWANGER, O.M., 1955a. Neue Methode der Zerlegung von Häufigkeitsverteilungen in Gauss'sche Normalkurven und ihre Anwendung in der Meteorologie.
Ber. Dt. Wetterd., No. 10, pp. 11.

ESSENWANGER, O.M., 1955b. Zur Haufigkeitsanalyse von Grosswetterlagen.
Meteor. Rundschau, Vol. 8, p. 55-56.

ESSENWANGER, O.M., 1972. Characterisation of Global Atmospheric Conditions by Parameterized Daily Profiles.
Preprint Vol. of the International Conf. on Aerospace and Aeronautical Meteorology, Washington, D.C., p. 25-31
(Publ. by Am. Meteor. Soc., Boston).

GODSKE, C.L., 1962. Contributions to Statistical Meteorology.
Geofysiske Publikasjoner geophysica Norvegica, Vol. XXIV, p. 161-210.

GODSKE, C.L., 1966a. A Statistical Approach to Climatology.
Arch. Met. Geoph Bioklim, Ser B, Bd 14, H3-4, p. 269-279.

GODSKE, C.L., 1966b. On the Time Dependence of Smoothed Variables.
Tellus XVIII, p. 714-721.

PEARSON, K., 1894. Contributions to the Mathematical Theory of Evolution.
Phil. Trans. Roy. Soc., A. Vol. 184, p. 71-110.

TOWARD THE PREDICTION OF EFFECTS OF ULTRA-SHORT LASER PULSES

E. W. Stuebing, E. A. Lucia, and F. D. Verderame
Pitman-Dunn Laboratory, Frankford Arsenal
US Army Armament Command
Philadelphia, PA 19137

ABSTRACT. The recently developed mode-locking technology for producing ultrashort (picosecond) laser pulses is being explored for military applications. Understanding the behavior of absorbing media toward laser pulses of picosecond duration requires detailed modeling of intramolecular energy transfer processes because the time scale of such processes may be considerably longer than the entire duration of the pulse. Systems of coupled differential equations in two dimensions are formulated for candidate energy transfer schemes and their solutions compared to original experimental evidence in four time regimes. The equations are solved numerically by finite difference methods and alternatively by Monte Carlo techniques. A comparison of the two approaches is given.

1. INTRODUCTION. The tris (p-dialkylaminophenyl) aminium hexafluoroantimonates[1] are a series of dyes developed by the American Cyanamid Corporation, Bound Brook, NJ under contracts with Frankford Arsenal. One compound of this class, tris (p-diethylaminophenyl) aminium hexafluoroantimonate (TEAAF) has been the key ingredient in formulations for laser countermeasures, goggles and visors effective against ruby and neodymium lasers. When dissolved in vinyl, cellulostic, or acrylic plastics, TEAAF strongly absorbs ruby, Nd, 2 x Ruby, GaAs, Neon and argon ion laser radiation, yet transmits a large portion of the visible spectrum (See Figure 1). With suitable additives, this dye is very stable toward heat and light and seems not to be irreversibly affected by very high powered Q-switched ruby and neodymium/glass laser pulses. Plastic formulations of TEAAF are now commonplace for providing adequate and reliable protection especially against pulsed 1.06μm laser radiation by absorbing the preponderance of the intense Q-switched light to which they are exposed while maintaining their chemical and physical properties even after repeated exposure. In the near future, it is anticipated that new liquid formulations utilizing this compound for protective and masking tactics will be developed and used. The

[1] U.S. Patent 3,440,257 4/22/69

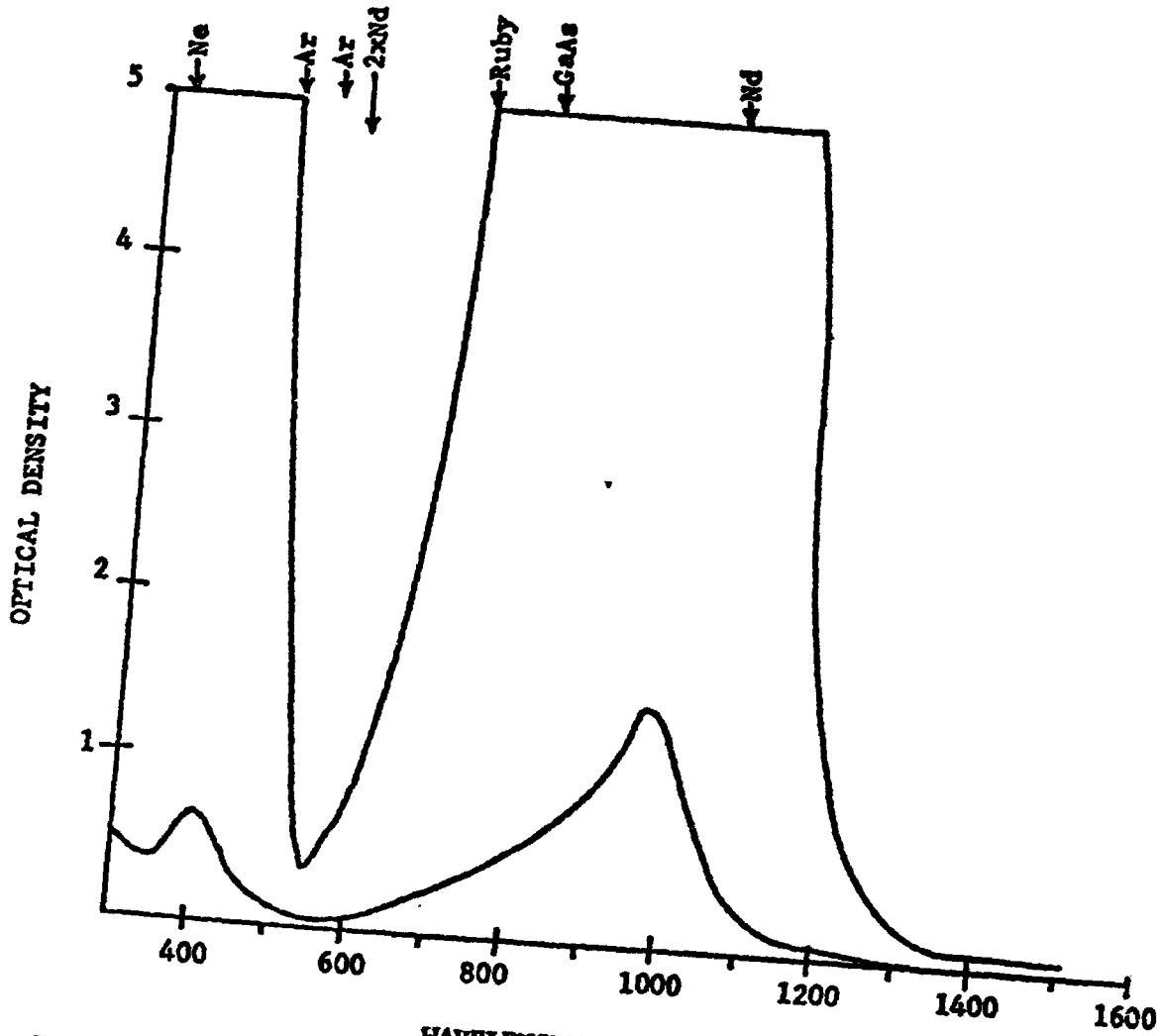


FIGURE 1. ABSORPTION SPECTRUM OF TEAAF IN POLYVINYL PLASTIC UNDER LIGHT OF ORDINARY INTENSITY (CARY 14)

objective of this study was to discover the effects of ultrashort pulses of neodymium laser radiation on several TEAAF plastic formulations now in use and to understand the intramolecular and intermolecular energy transfer processes which affect absorption of radiation as a function of peak power and pulse widths in the picosecond/gigawatt regime.

Data for this study was gathered at $\lambda = 1.06\mu m$ in four time regimes using the Cary 14 spectrophotometer (CW), a Korad Model K-2 Nanosecond Laser System (20 nsec), and a Korad Model K-1500 Mode-locked Laser System (18 psec), all in our laboratory, and laser equipment available to us at Naval Research Laboratory with 250 picosecond (FWHM) duration pulses. The last two lasers will be referred to as picosecond systems. Reference to Figure 1 (lower curve) shows that $1.06\mu m$ is not exactly at the absorption maximum but rather on the low frequency side of the broad, infrared absorption maximum of TEAAF centered at about 990 nm. With the picosecond systems at $1.06\mu m$ new and unexpected effects have been observed with TEAAF dissolved at several concentrations in methyl methacrylate monomer, clear plastic polymethylmethacrylate and in polycellulose propionate matrix. Namely, we have observed a markedly enhanced optical density at $1.06\mu m$ at irradiance levels (power per unit area) comparable to those of Q-switched pulses; a reversible bleaching at higher peak powers in the several gigawatt regimes, and a strong fluorescence in many samples especially during exposure at higher powers. The results are correlated with theory based upon rates of energy transfer among the available TEAAF energy levels. In these energy level schemes the rate of vibrational-vibrational relaxations cannot be neglected and therefore would give rise to complicated rate expressions.

In view of conjectures before these experiments were conducted about the kinds of effects to be expected, it is equally important to stress what was not observed. No cracking, crazing, spallation, burn spots, or permanent damage of any kind could be observed under microscopic examination. This kind of damage is more likely energy dependent rather than power dependent, and only a small total energy (~ 1 joule) was deposited even at the highest peak powers.

2. EXPERIMENTAL RESULTS. Typical results for samples of TEAAF in polycellulose propionate (PCP, Lexan) and in polymethylmethacrylate (PMMA, Plexiglass) are shown in Figures 2-4. Sample thicknesses were adjusted so that the optical density toward light of ordinary intensities (as measured on the Cary 14) would be approximately one. The Cary 14 OD of each sample is shown by a dotted line on each figure, where OD is measured by

$$OD = - \log_{10} (I/I_o) \qquad (1)$$

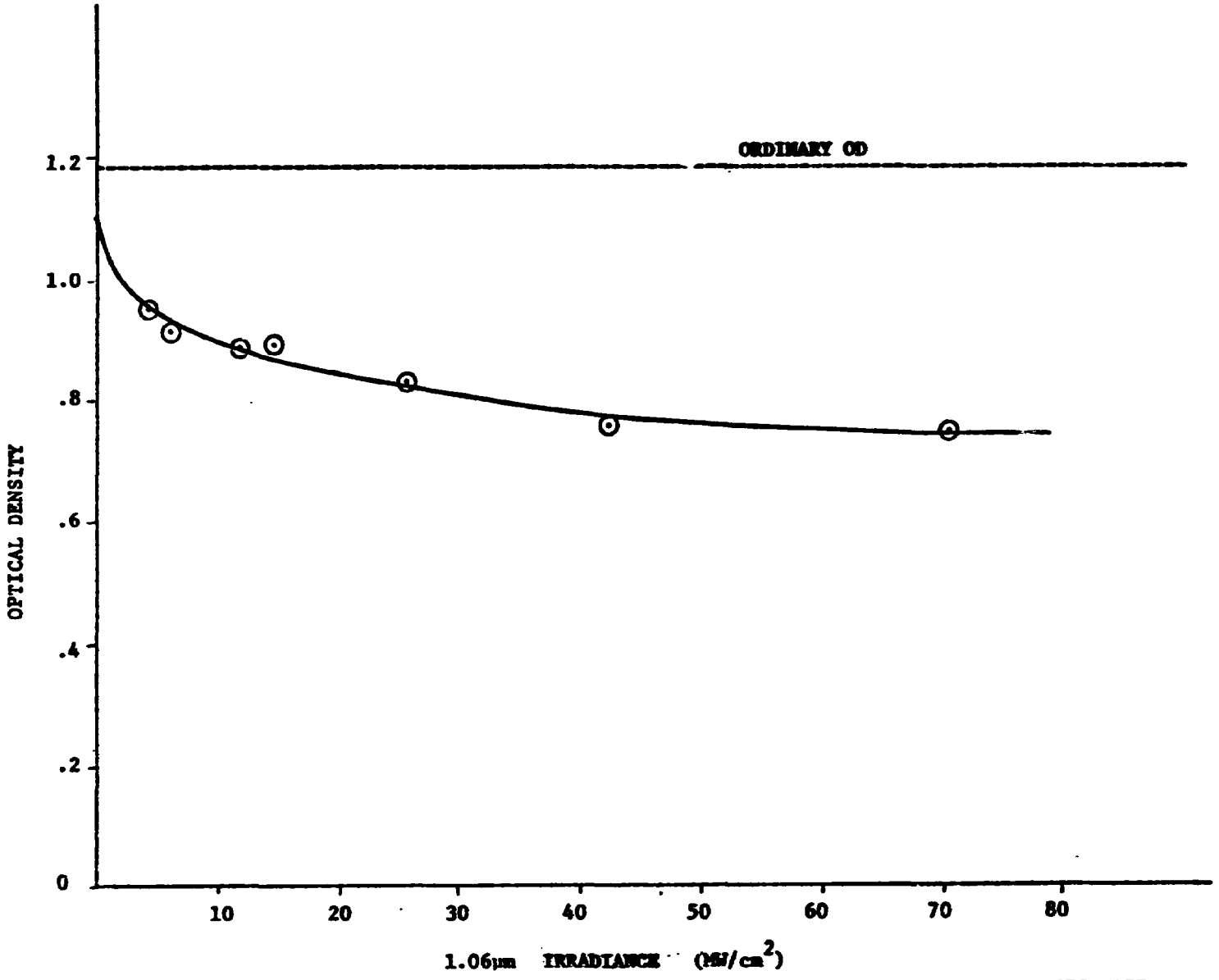


FIGURE 2. PEAK POWER IRRADIANCE DEPENDENCE OF TEAAF ABSORPTION IN POLYCELLULOSE PROPIONATE FOR 20 NSEC LASER PULSES.

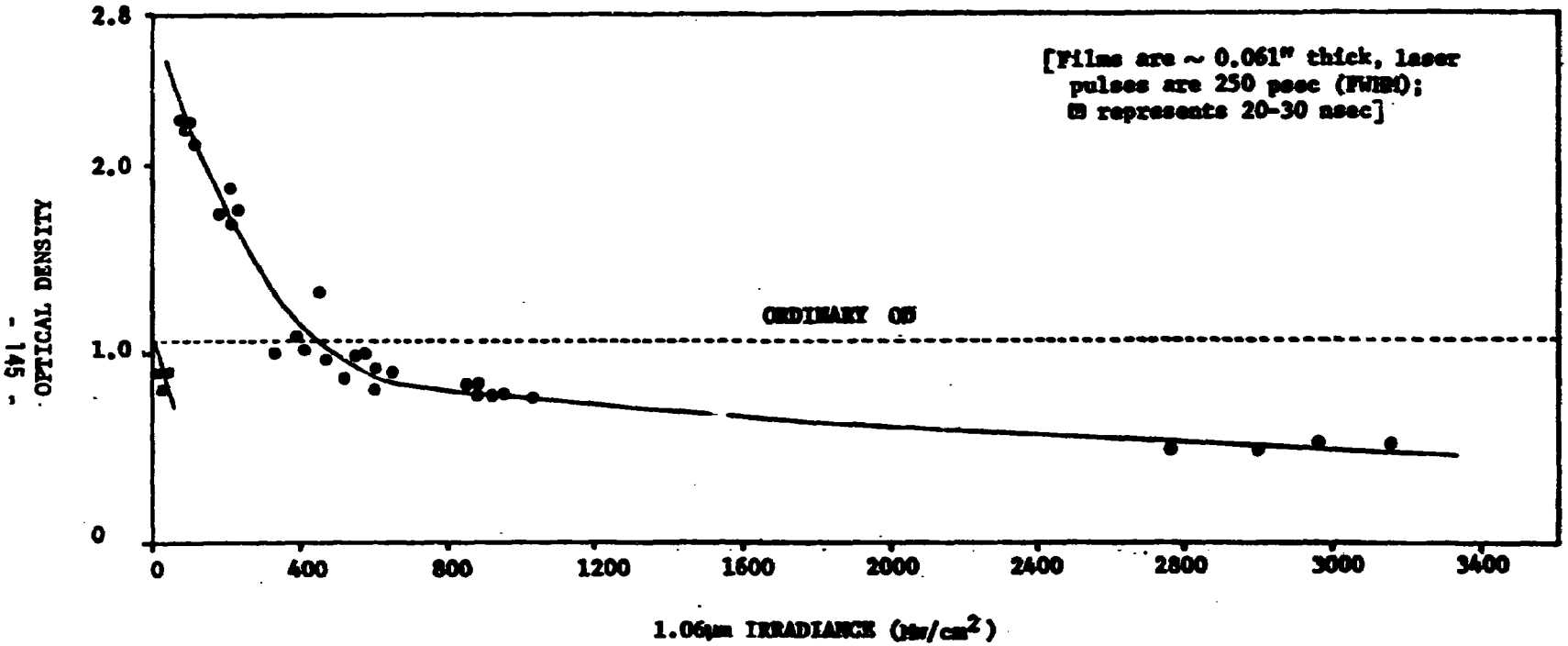


FIGURE 3. OPTICAL DENSITY VS. PEAK POWER IRRADIANCE FOR TRAAF IN PMMA

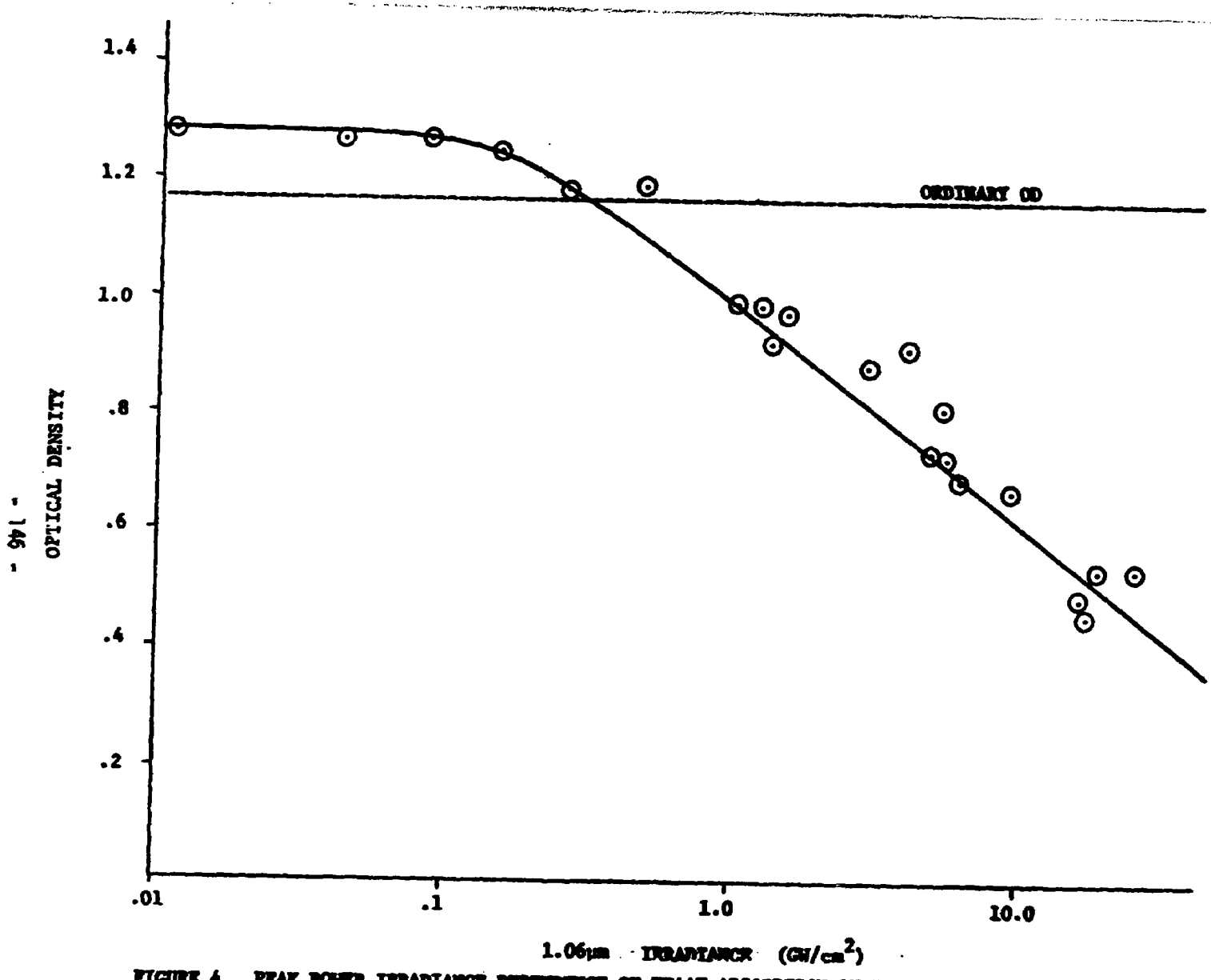


FIGURE 4. PEAK POWER IRRADIANCE DEPENDENCE OF TRAAF ABSORPTION IN POLYCELLULOSE PROPIONATE FOR 18 PSEC LASER PULSES

for I_o being the incident beam intensity, and I being the intensity transmitted by the sample. Figure 2 shows that the dye bleaches under exposure to 20 nsec pulses over irradiance levels of 5 to 70 megawatts/cm^2; however, the bleaching is not catastrophic (virtually complete loss of OD) but levels off after a reduction to about 2/3 of the original OD and then appears to be fairly impervious to further increases in power. Furthermore, the bleaching is reversible; a sample exposed to 70 MW/cm^2 shows complete recovery of its absorption properties when subsequently exposed to 5 MW/cm^2 or Cary 14 measurements at the same site.

The results with 250 psec pulses, Figure 3, show a new and unexpected effect. For irradiance levels in the range 100-400 MW/cm^2, the sample showed a marked enhanced absorption rather than bleaching. At higher powers the sample bleaches in about the same way (noncatastrophically) as with 20 nsec pulses, with the crossover from enhancement to bleaching occuring at roughly 400 MW/cm^2.

For 18 psec pulses, Figure 4, enhancement again occurs in the range 100-250 MW/cm^2, converting to bleaching at higher powers. The abcissa on Figure 4 is plotted on a log scale due to the wide power range of the instrument. When this data is plotted on a linear scale comparable to Figures 2 and 3, then the data of Figure 4 falls on a curve whose shape resembles the two previous figures: peaked sharply to the left with the bleaching flattening out to a long, horizontal line at about OD = 0.6.

3. THEORETICAL ANALYSIS. The lowest lying energy states of a typical aromatic molecule are schematically illustrated in Figure 5, where solid arrows indicate radiative transitions undergone by absorbing or emitting photons, and dotted arrows indicate radiationless transitions. The ordinary optical density of an absorbing molecule is due to transitions from the ground state to the first excited singlet state from which it may relax to the ground state by radiative (fluorescence) or radiationless (quenching) paths. In light beams of ordinary intensity, these relaxation processes are sufficiently fast to keep the ground state essentially fully populated. In laser beams of high irradiance, however, this relaxation rate may be overwhelmed by the arrival rate of photons so that many, and perhaps virtually all, of the absorption sites are occupied by molecules in the first excited singlet state (or the lowest triplet state on the right side of the diagram, if the pulse length is sufficiently long compared to the rate of intersystem crossing). If there is no available absorption from this excited state (e.g., K_{12} in the figure) at the proper wavelength, then the material bleaches catastrophically. If such an absorption exists but is weaker than the ground state absorption (K_{01}) then the OD of the sample decreases from its

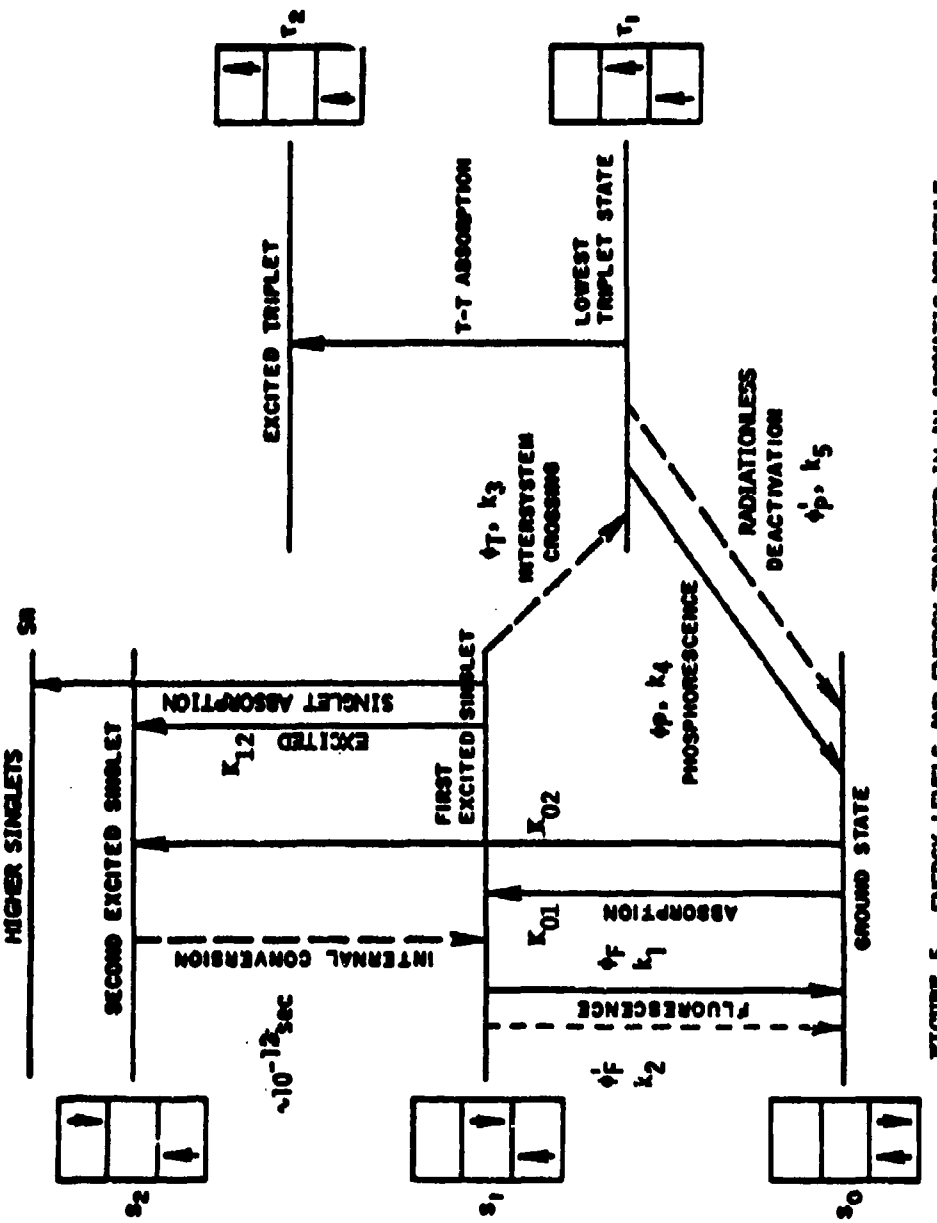


FIGURE 5. ENERGY LEVELS AND ENERGY TRANSFER IN AN AROMATIC MOLECULE.

initial value toward that corresponding to K_{12}. Because radiationless relaxations among excited states (internal conversion) are generally exceedingly fast compared to relaxations from the lowest excited state, the K_{12} absorption will not depopulate S_1 at power levels which do depopulate S_0, so that the bleaching levels off over at least some range of powers at the OD corresponding to K_{12}. This appears to be the behavior of TRAAP toward 20 nsec pulses and toward picosecond pulses at their higher powers. It does not explain the enhancement.

Previous discussions of enhancement have focused on the existence of an excited absorption (K_{12}) stronger than K_{01}. However, this would lead to an enhancement that switched on at higher powers and leveled off as they increased, an effect we have observed weakly in our laboratory at the ruby wavelength (0.6943μm). It cannot explain the low-power enhancement which steadily bleaches away which is observed here. We therefore offer the conjecture that this effect is due to a concerted two-photon absorption directly from the ground state to the second singlet (represented by K_{02} in Figure 5). Such absorptions are very improbable compared to K_{01}; however, they depend on the square of the photon intensity, therefore switching on at higher powers. When the power reaches a level such that this absorption appears, it occurs in addition to K_{01} which continues as usual so that the net effect is an addition of the OD's due to the two absorptions and an enhancement of the sample OD above that due to K_{01} alone. Increasing the power beyond this point should increase the rate of the two-photon absorption; however, this additional absorption arises only by additionally depleting the ground state and results in even more rapid population of the first excited singlet (following the rapid internal conversion). Hence, the overall effect may be somewhat additionally enhanced absorption during the leading edge of the pulse accompanied by equally more rapid bleaching of the ground state and lower OD toward the bulk of the pulse, resulting in net bleaching from the initial enhanced value for pulses with additional power beyond the onset of two-photon absorption. As detectors are not yet capable of following the detailed pulse shape of picosecond pulses, this cannot now be tested experimentally.

To examine the feasibility of such an explanation in terms of the intramolecular energy transfer rates (K's) required to produce such an overall result, the differential equations governing the above processes have been formulated for numerical solution. This was originally done by Monte-Carlo methods which proved too time consuming. Subsequent finite difference solutions have proved much more economical and provide detailed profiles of the incident and transmitted pulses from which peak powers and integrated energy content can be determined. These two parameters give some measure of the degree of distortion of the pulse shape by the

sample and can be compared with results of power-sensitive and energy-sensitive detectors which are available. In addition, molecular orbital calculations are being carried out on the TEAAF molecule using the semi-empirical CNDO method followed by configuration interaction to generate the excited states. These results will be examined to determine whether excited states of the TEAAF molecule exist with the proper energy differences and relative transition intensities to fit the requirements predicted by the rate equations.

4. A PROBLEM IN EXPERIMENTAL DESIGN. A diagram of the laser apparatus is shown in Figure 6. Sample optical densities are determined by measuring the relative signals at the sample detector and at the reference detector with reliable results. However, determining irradiance levels is dependent on absolute calibration of the reference detector which is quite difficult and a major source of experimental error, most especially with the 18 psec system. For example, even though it would be valuable to determine whether the enhancement we observe is dependent on irradiance power level alone or is also modified by pulse width, due to the uncertainty in the ordinate values we cannot say whether or not the 250 psec data crosses the ordinary OD line at the same irradiance level as with 18 psec pulses. In the same way it is difficult to sort out any dependence of the enhancement on the square of the intensity. Present practice calls for reference detector calibration runs at the beginning of each day which are both inadequate and too time consuming. These lead to a linear fit of reference detector response versus calibrated laser power delivered. Unfortunately the time required for enough shots to establish tight confidence limits is so great as to consume virtually the entire working day. This is confounded by a lack of control by the operator over the power to be delivered by the laser on a given shot. For a particular setting of instrument controls, a series of shots will result in roughly a normal distribution of powers about some mean with a range of perhaps 50% of the mean value. It is hoped that with experience the detector calibration will be found to be sufficiently stable that a very highly characterized calibration curve can be established and the daily calibrations replaced with periodic samplings of small size but adequate to establish with the high confidence required that we are still sampling the same calibration curve.

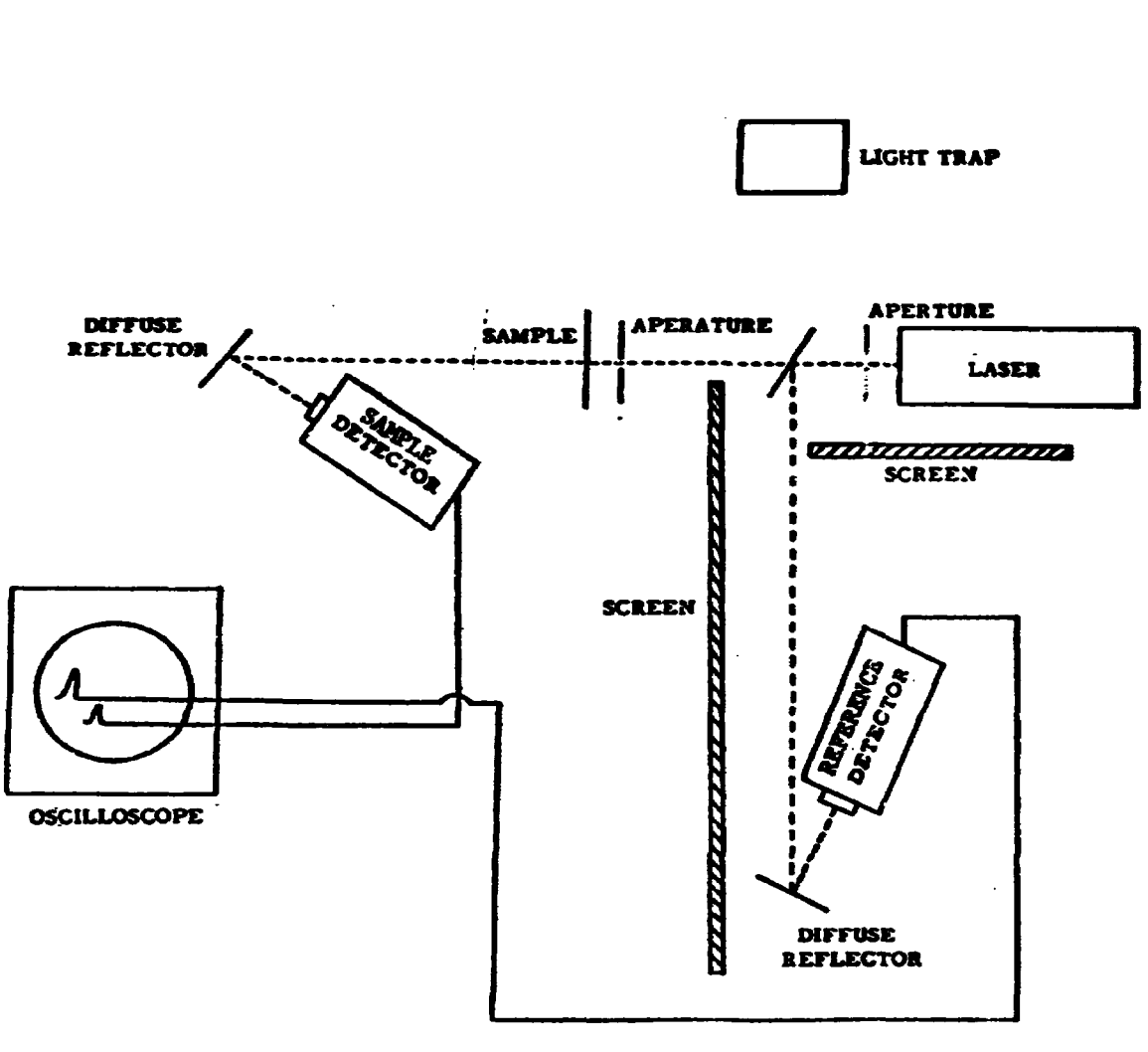


FIGURE 6. LASER IRRADIANCE APPARATUS

MODELING SEE-THROUGH AEROSOLS FOR INFRARED COUNTERSURVEILLANCE AND LASER COUNTERMEASURES

R. Doherty, J. J. Pinto, E. W. Stuebing, and F. D. Verderame
Pitman-Dunn Laboratory, Frankford Arsenal
US Army Armament Command
Philadelphia, PA 19137

ABSTRACT. All targets emit and reflect radiation which can be sensed by infrared detectors and seekers. To reduce the infrared radiation from the target, particulate clouds containing infrared absorbers can be employed. Mie theory analysis of a cloud of liquid microemulsion with known composition and known complex refractive index generates the luminous transmittance and optical density (at $\lambda = 1.06\mu m$) of the cloud as functions of number density, particle size, and path length. These calculations are used to define the particle size and number density requirements to produce a cloud with optical density of one and a luminous transmittance of 25%. A pneumatic spray nozzle is used to produce a cloud of the microemulsion for which the number density, particle size, and optical density at $\lambda = 1.06\mu m$ are determined and correlated with calculated Mie theory results.

1. INTRODUCTION. At the Applied Science Laboratory of Frankford Arsenal, we are interested in the development of particulate clouds as effective countermeasures against infrared guidance and tracking systems. To establish the physical characteristics of a particulate cloud necessary to achieve an effective countermeasure, a Mie scattering program was acquired from the Atmospheric Sciences Laboratory, White Sands Missile Range, New Mexico. The program was modified to use the scattering cross section and absorption cross section of representative cloud particles to predict the luminous transmittance and optical density at 1.06μm over various path lengths within the cloud.

In this manner the effect of particle size, number density and the complex refractive index of various materials were studied. This report presents the results of these calculations for a particular microemulsion containing the molecular absorber IR-99 (absorption maximum at 980 nm) where a luminous transmittance of 25% and an optical density of one at 1.06μm have been arbitrarily established as desired cloud properties. In addition, an experimental procedure worked out in cooperation with Edgewood Arsenal and used to verify the computational figures is also presented.

2. THEORETICAL INVESTIGATIONS. The program PGAUS was acquired from the US Army Atmospheric Sciences Laboratory, White Sands Missile Range, New Mexico. The program performs single scattering calculations from first principles according to the Mie theory, resulting in the scattering cross section and absorption cross section for an average scatterer in a cloud. In addition to calculating these values at 1.06μm, the program was modified to calculate them at various wavelengths throughout the visible region. Tables of 1.06μm extinction and luminous transmittance through a dispersed microemulsion were then generated for number densities in the range of 10^{9}-10^{16} at particle diameters ranging from 0.01μm to 2.0μm. Additional phase function interpolation and integration routines were included to calculate the amount of radiation scattered in small angles about the forward and backward directions. The modified program has been renamed "PGAUS-LT" and is described in detail along with operating instructions in a report presently under preparation.[1]

Mie calculations in "PGAUS" are limited to single scattering calculations, hence the possibility of light having once been scattered being rescattered one or more times before leaving the cloud is not considered. Single scattering calculations are less costly to perform and are likely to provide valid results for those aerosols in which extinction is due almost entirely to absorption. They also identify those cases for which the number density and particle size require multiple scattering calculations and are useful as a survey tool to eliminate cases where the cloud is too transparent to the 1.06μm wavelength.

An oil/alcohol/soap/dye microemulsion in water[2] has been formulated to extinguish the 1.06μm wavelength. The oil drops in the microemulsion as determined by low angle x-ray diffraction are at least as small as 0.01μm. However in our computer calculations we consider larger drops as well because dispersion of the microemulsion leads to large drops containing many dye bearing oil drops which aggregate on evaporation of the water. The exact structure of the microemulsion droplet is not known but is suspected to be similar to two concentric spheres with oil in the center and an oil/alcohol/soap/dye layer around the outside. In order to perform Mie calculations using "PGAUS", the oil droplet was treated as a homogeneous sphere with the real portion of its complex refractive index, $M = n(1 - ik)$, due to the oil and alcohol. Neither material absorbs in the region 0.4μm - 1.06μm. Hence, large anomalies are not expected and the real index is taken to be constant at $n = 1.44$ over this region. The imaginary component, k, is taken entirely from the dye as no other component of the microemulsion absorbs in the 0.4μm - 1.06μm regions and is given by

$$k = \frac{\varepsilon C \lambda}{4\pi} \qquad (1)$$

where ε is the molar extinction coefficient of the dye at wavelength, λ, and C is the molar concentration of the dye in the drop. The concentration of dye in bulk microemulsion is 2.6×10^{-2} molar. Values for k based on this concentration and on ε's taken in methylmethacrylate (MMA) solution using a Cary 14 Spectrophotometer are given in Table I.

Optical density [$OD = -\log_{10}(I/I_o)$] is a linear function of path length and number density, so for example, at fixed number density, doubling the path length will double the OD. On the other hand, luminous transmittance is a complicated function of the product of number density times path length* (through an exponent under an integral) and therefore must be tabulated as a function of one or the other. For our purposes we chose a ten meter path length and tabulated results as a function of number density given in orders of magnitude. Manipulation of the results to other path lengths is straightforward.

Table II plots the results of luminous transmittance and optical density calculations based on our model where the albedo at $\lambda = 1.06\mu m$ for each drop size gives the ratio of scattered IR energy to total IR extinction.

We see from Table II that for oil drops of $0.5\mu m$ in diameter and larger, almost all of the $1.06\mu m$ light lost is due to scattering. It turns out that particle diameters larger than $0.5\mu m$ act as Mie scatterers with phase functions strongly peaked in the forward direction so that a certain amount of energy will be scattered forward out of the cloud and may be useful as a surveillance or homing signature. Thus, although a number density of $10^5 - 10^6$ for the $0.5\mu m$ diameter particles appears to be close to satisfying our hypothetical case (OD = 1 @ $1.06\mu m$ with LT = 25%), it is uncertain whether these particles would acceptably screen a target. However, at a number density of $10^8 - 10^9$ multiple scattering effects would be important and would tend to distribute energy away from the forward direction. The problem of forward scattered (or backward scattered) light is eliminated by choosing particle diameters less than or equal to $0.1\mu m$. Such particles strongly absorb the $1.06\mu m$ light and, provided adequate number densities (or longer path lengths) can be achieved, they would provide an effective screen for a target.

Figure 1 shows the percent luminous transmittance through these aerosols when they present an OD of 1.0 to $1.06\mu m$ light. These results

* See Appendix A.

TABLE I. - IMAGINARY COMPONENT INDEX OF REFRACTION FOR IR-99

AT A CONCENTRATION OF 2.6×10^{-2} M

Wavelength	ϵ^a	k^b
.40	5.65×10^4	4.66×10^{-3}
.45	2.14×10^4	1.99×10^{-3}
.50	0.39×10^4	0.40×10^{-3}
.55	0.24×10^4	0.27×10^{-3}
.60	0.33×10^4	0.41×10^{-3}
.65	0.62×10^4	0.83×10^{-3}
.70	1.44×10^4	2.06×10^{-3}
1.06	3.92×10^4	8.55×10^{-3}

a. Determined in methylmethacrylate solution

b. From eq. (1)

TABLE II.— OPTICAL DENSITY AT λ = 1.06μm AND LUMINOUS TRANSMITTANCE OVER A PATH LENGTH OF TEN METERS FOR AEROSOLS OF OIL DROPS CONTAINING IR-99 DYE[a] AT A CONCENTRATION OF 2.6×10^{-2} M

	Drop Diameter (μm)											
	2.0		1.0		0.5		0.1		0.05		0.01	
log (N)[d]	LT[b]	OD[c]	LT	OD	LT	OD	LT	OD	LT	OD	LT	OD
0												
1	100											
2	99		100		100							
3	92	0.05	97		99							
4	44	0.47	75	0.10	95							
5	0	4.74	6	0.97	60	0.05						
6			0	9.67	1	0.52	100					
7					0	5.20	98					
8							85	0.03	100			
9							22	0.32	96	0.03		
10							0	3.19	67	0.34	100	
11									3	3.35	99	0.03
12									0		88	0.26
13											32	2.59
14											0	
Albedo at λ = 1.06	0.90		0.94		0.90		0.15		0.02		1.9×10^{-4}	

a. Table I
b. Percent luminous transmittance
c. Optical Density at λ = 1.06μm
d. N = number density (cm^{-3})

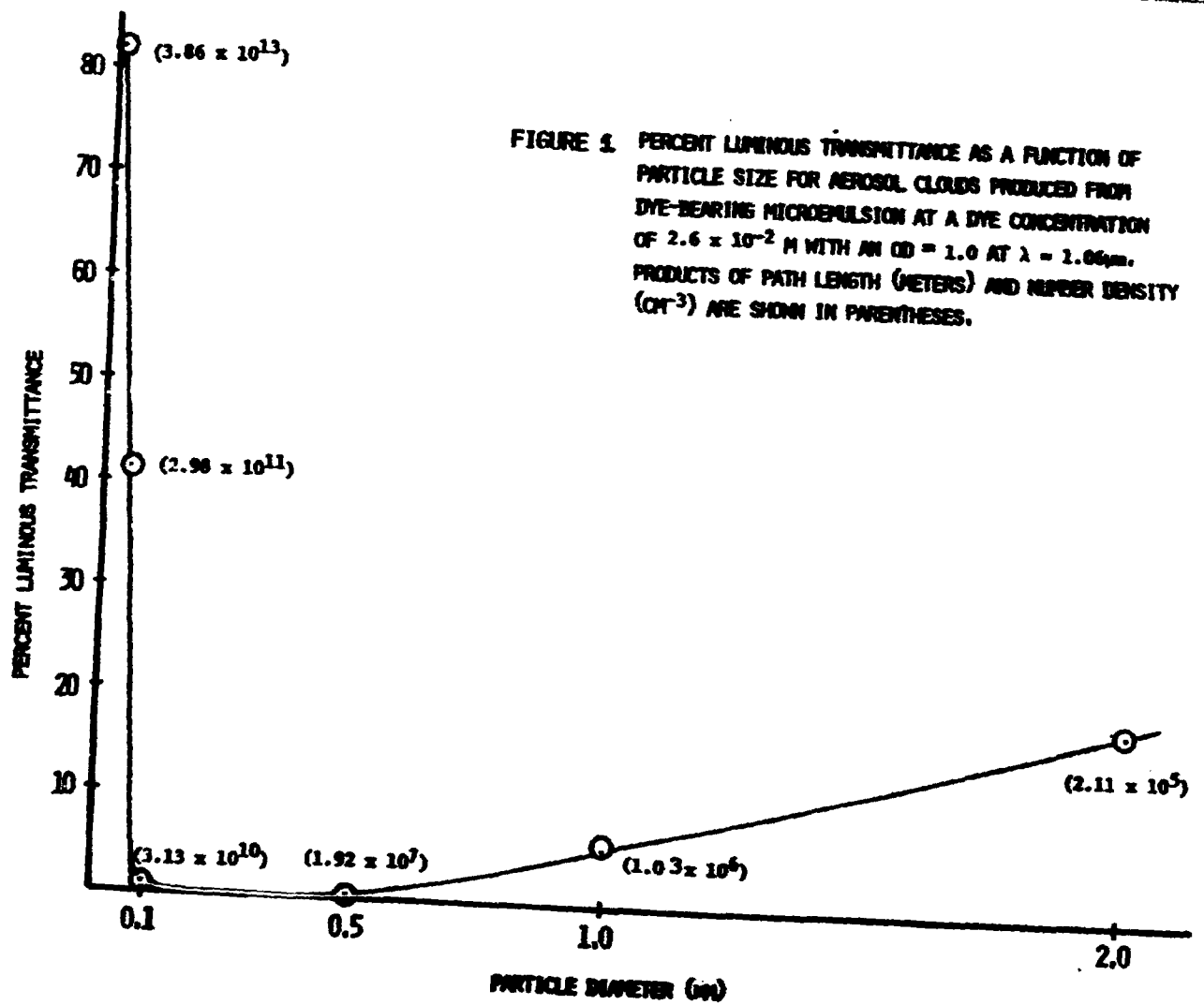


FIGURE 1. PERCENT LUMINOUS TRANSMITTANCE AS A FUNCTION OF PARTICLE SIZE FOR AEROSOL CLOUDS PRODUCED FROM DYE-BEARING MICROEMULSION AT A DYE CONCENTRATION OF 2.6×10^{-2} M WITH AN OD = 1.0 AT λ = 1.06μm. PRODUCTS OF PATH LENGTH (METERS) AND NUMBER DENSITY (CM^{-3}) ARE SHOWN IN PARENTHESES.

are independent of path length, i.e., whenever the product of path length times number density is such that an OD of 1.0 is achieved at 1.06μm, then the cloud will have the given luminous transmittance. The value of this product in ($M \times cm^{-3}$) is given at each point. The minimum in this curve at 0.5μm is due to the strong scattering of light at the wavelength of maximum contribution to the luminous transmittance integral (~ 0.55μm) by particles of the same size. For larger particles, the luminous transmittance increases because, although the majority of the energy lost is due to scattering, the scattering no longer strongly prefers 0.5μm to 1.0μm as the particle size moves closer to, and beyond, 1.0μm. The increased luminous transmittance for particles of 0.1μm and smaller is due to the increasing predominance of absorption over scattering as sizes diminish. Then the relative extinction of 1.06μm to the visible more and more closely resembles the absorption spectrum of the dye which is strongly selective for the IR. Consequently, the most advantageous way to achieve and OD = 1 and a LT of 25% will be with particle sizes below 0.1μm. Furthermore, such particles tend not to settle out of air and hence provide a very persistent aerosol. Since monodisperse aerosols of these liquid materials are unlikely, estimates of the upper limits of larger drops that can be tolerated in the mixture without degrading luminous transmittance can be made from Table II (i.e. we can tolerate densities of $10^5 cm^{-3}$ of 0.5μm drops and $10^4 cm^{-3}$ of 1.0 to 2.0μm drops). Moreover, the presence of water drops in the dispersed aerosol may cause scattering of so much visible light that 25% luminous transmittance cannot be achieved. Table III shows calculated luminous transmittance through ten meter clouds of water drops as functions of drop diameter and number density. Thus, tolerable limits on water drops can be added to the dye bearing oil drops to complete the characterization of an adequate aerosol.

Thus far the results reported are for a near saturation concentration of the dye, i.e., 18 mg/mℓ or 2.6×10^{-2} M. Should a better solvent for the dye be discovered, it is of interest to examine the effect of the resulting change in dye concentration. For this purpose we consider the 0.1μm oil drops with 0.03 M concentration (corresponding to the bulk microemulsion saturation concentration of 21 mg/mℓ) and with 10-fold and 100-fold increases in concentration. The results are presented in Table IV. The extinction increase with concentration almost follows Beer's Law so that each order of magnitude increase in concentration buys very nearly an order of magnitude decrease in the number density (or number density x path length) required. For smaller particles with less scattering, this increase even more closely follows Beer's Law. For the larger particles with large albedos (scattering rather than absorption dominates), increases in concentration buy little relaxation in density times path length requirements.

TABLE III. - PERCENT LUMINOUS TRANSMITTANCE THROUGH TEN METER CLOUDS OF WATER DROPS AS A FUNCTION OF NUMBER DENSITY N (DROPS PER cc) AND DROP DIAMETER (μm)

log (N)	Drop Diameter (μm)					
	2	1	0.5	0.1	0.05	0.01
0	100	100	100	100	100	100
1	100	100	100	100	100	100
2	99	100	100	100	100	100
3	92	97	100	100	100	100
4	44	75	95	100	100	100
5	0	6	60	100	100	100
6		0	1	100	100	100
7			0	98	100	100
8				85	100	100
9				21	96	100
10				0	67	100
11					3	99
12					0	88
13						32
14						0

TABLE IV. - OPTICAL DENSITY AT λ = 1.06μm AND LUMINOUS TRANSMITTANCE OVER A PATH LENGTH OF TEN METERS IN AEROSOLS OF 0.1μm DIAMETER OIL DROPS CONTAINING H-99 DYE AT VARIOUS CONCENTRATIONS

	Concentration Factor[a]							
	1.0		2.0		10.0		100.0	
log (N)[d]	LT[b]	OD[c]	LT	OD	LT	OD	LT	OD
0								
1								
2								
3								
4								
5							100	
6	100		100		100		99	.02
7	98		98		98	.03	92	.26
8	85	.04	85	.07	80	.32	44	2.64
9	22	.36	21	.68	12	3.21	0	
10	0	3.65	0	6.81	0			
11								
12								
13								
14								
Albedo at λ = 1.06	0.13		0.07		0.02		0.02	

a. To be multiplied by 3.0×10^{-2} M
b. Percent luminous transmittance
c. Optical density at λ = 1.06μm
d. N = number density (cm^{-3})

In summary, only particles of 0.1μm diameter or smaller take advantage of the selective absorbing properties of the dye. Beer's Law improvements in the absorption of the 1.06μm light of up to two orders of magnitude can be achieved with the 0.1μm diameter particles by increasing the dye concentration by up to two orders of magnitude. Furthermore, since the attenuation of the 1.06μm radiation is accomplished through the mechanism of absorption rather than scattering, the problem of target signature enhancement through forward (or backward) scattered light is eliminated. As shown in Table II, a number density of 10^{10} - 10^{11} with particle diameters of 0.05μm is needed to achieve a luminous transmittance of 25% and an optical density of one at 1.06μm in a 10 meter cloud where the OD is due to absorption. Such small particles and high number densities will at best be extremely difficult to form in a laboratory experiment. However, the relevance of our theoretical calculations using a homogeneous drop model can be determined by the characterization of experimentally produced clouds of larger particles and lower number densities.

3. EXPERIMENTAL INVESTIGATIONS. The microemulsion to be disseminated is siphoned into a pneumatic spray nozzle operating at 150 psig of N_2. The resulting spray is directed against the wall of a jar which contains the nozzle and microemulsion. In this manner the larger spray particles impact on the inside surface of the jar and only the smaller particles egress through a cut-out in the top of the top of the jar into a seven-foot diameter holding chamber. The holding chamber as shown in Figure 2 is equipped with window ports, sampling valves, and interfaced with a 1.06μm Neodymium laser. Pulse energy of the beam is 0.05 joules with a pulse width of 50 nanoseconds.

Particle size distribution for the cloud is obtained by drawing samples through a six-stage cascade type impactor which has a particle diameter range in graduated steps from 0.25μm to 8+μm. Coincident with the impactor samples, a separate filter sample of the cloud is drawn at a known rate. Colorimetric analysis of this filter yields the amount of dye in a unit volume of the cloud. Multiplying this value by the tank volume, the total mass of dye contained in the cloud can be determined. A sample plot of this data is shown in Figure 3 as a function of time after dispersal. Knowing the dye content in a unit volume of cloud, the original concentration of the dye in the microemulsion and the mean particle size, the number density of the cloud is established. Typical values for the pneumatically sprayed microemulsion after several minutes are allowed for the cloud to reach equilibrium are a number density of 10^6 and a particle size of 0.9μm.

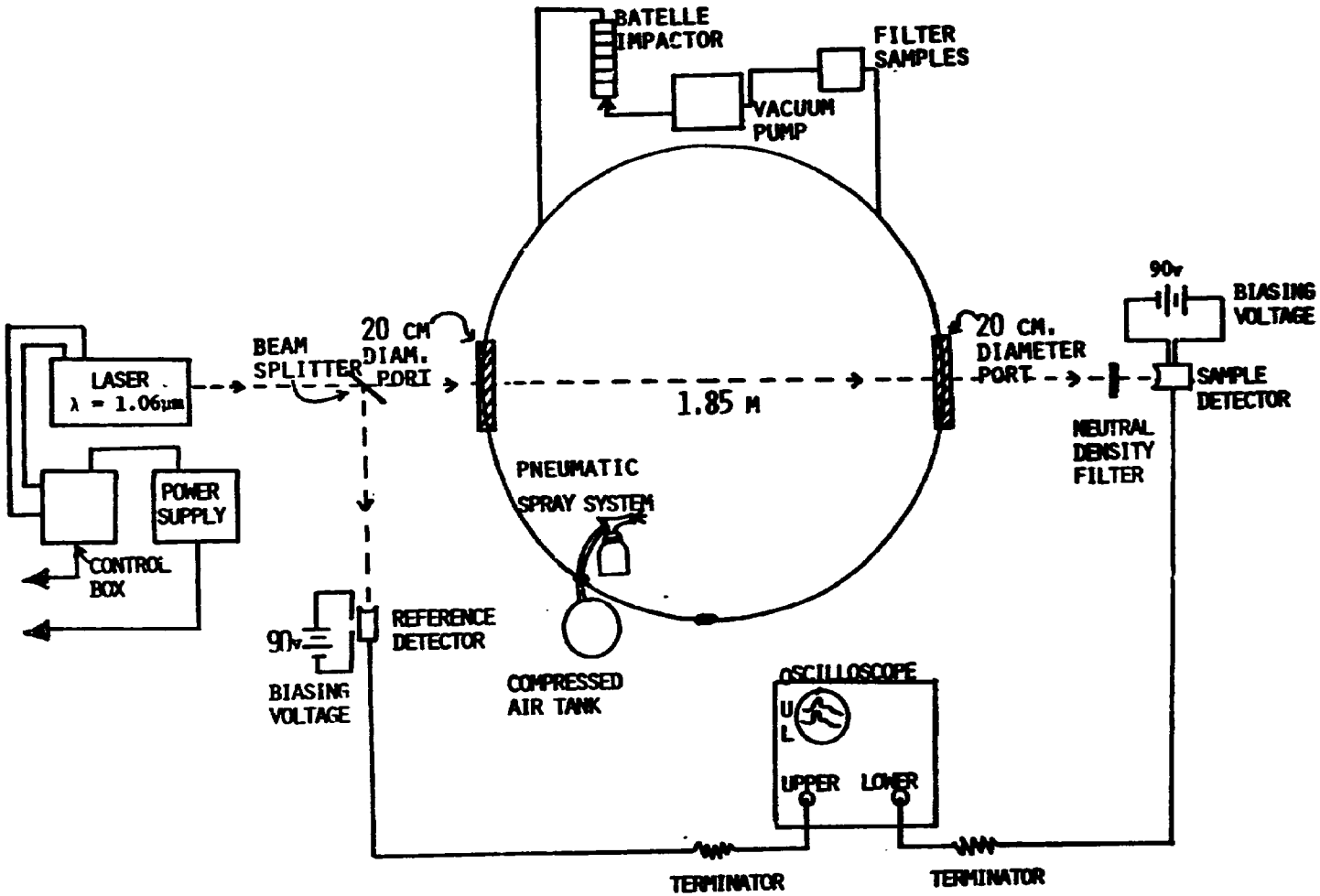
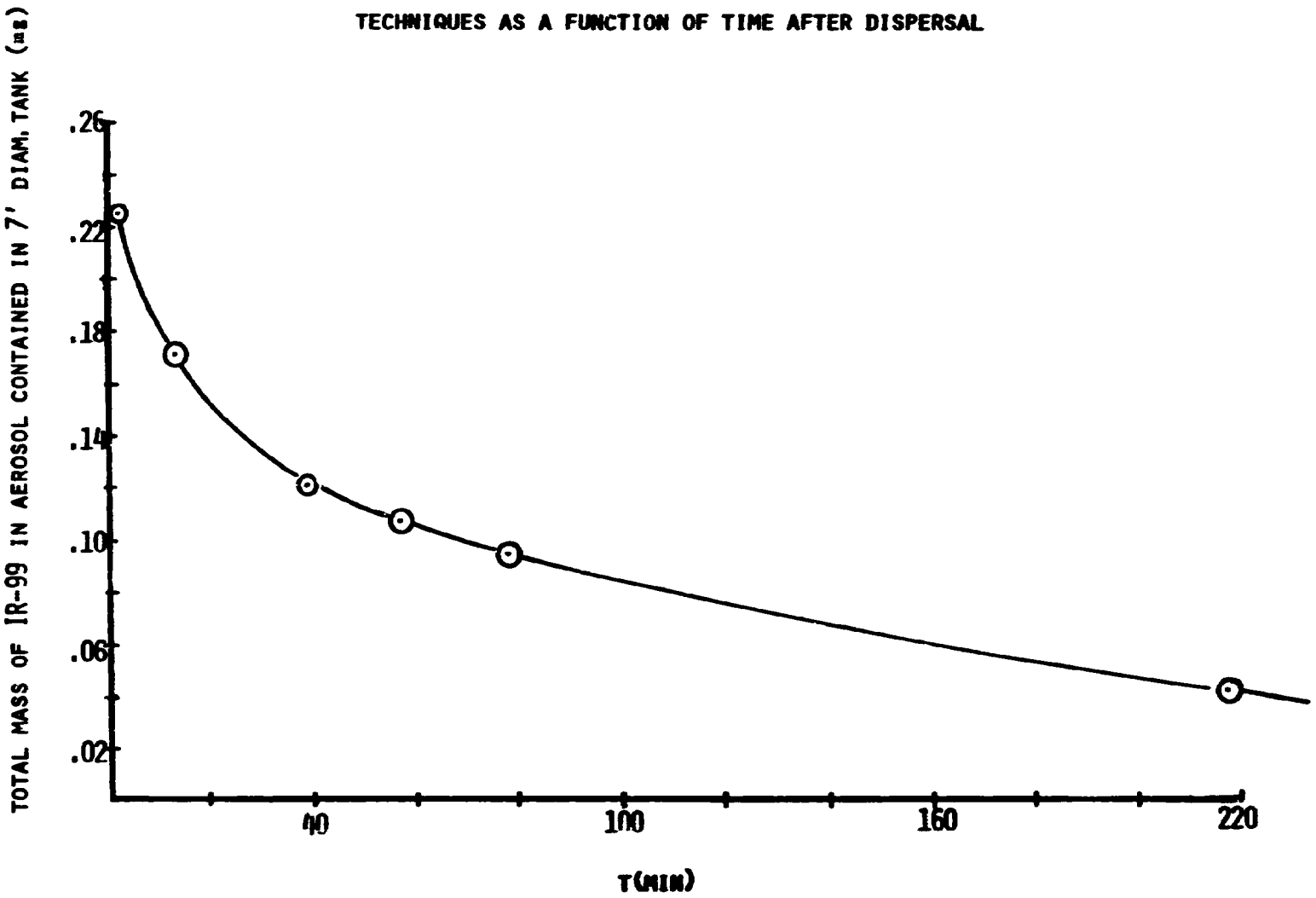


FIGURE 2 EXPERIMENTAL EQUIPMENT USED TO CHARACTERIZE PARTICLE SIZE, NUMBER DENSITY, AND OPTICAL DENSITY OF IR-99 DYED MICROEMULSION.

FIGURE 3 - MASS OF DYE, IR-99 IN AEROSOL DISPERSED BY THE SPRAY NOZZLE TECHNIQUES AS A FUNCTION OF TIME AFTER DISPERSAL



Evaluation of the optical density at 1.06μm for the cloud is accomplished by means of two photodiode detectors in combination with the 1.06μm Nd laser. A beam splitter inserted in the path of the light source causes a portion of the beam to enter one of the detectors (the reference detector). The remaining portion of the beam is allowed to pass through the chamber after which it encounters the second detector (the sample detector). Recording of the pulse heights for the two detectors is done on a dual beam oscilloscope equipped with a camera. Variations in source output, efficiency factors for the detectors, and reflective losses are eliminated by comparing the pulse heights for sprayed aerosols with those obtained when only air was inside the chamber by means of the following equation:

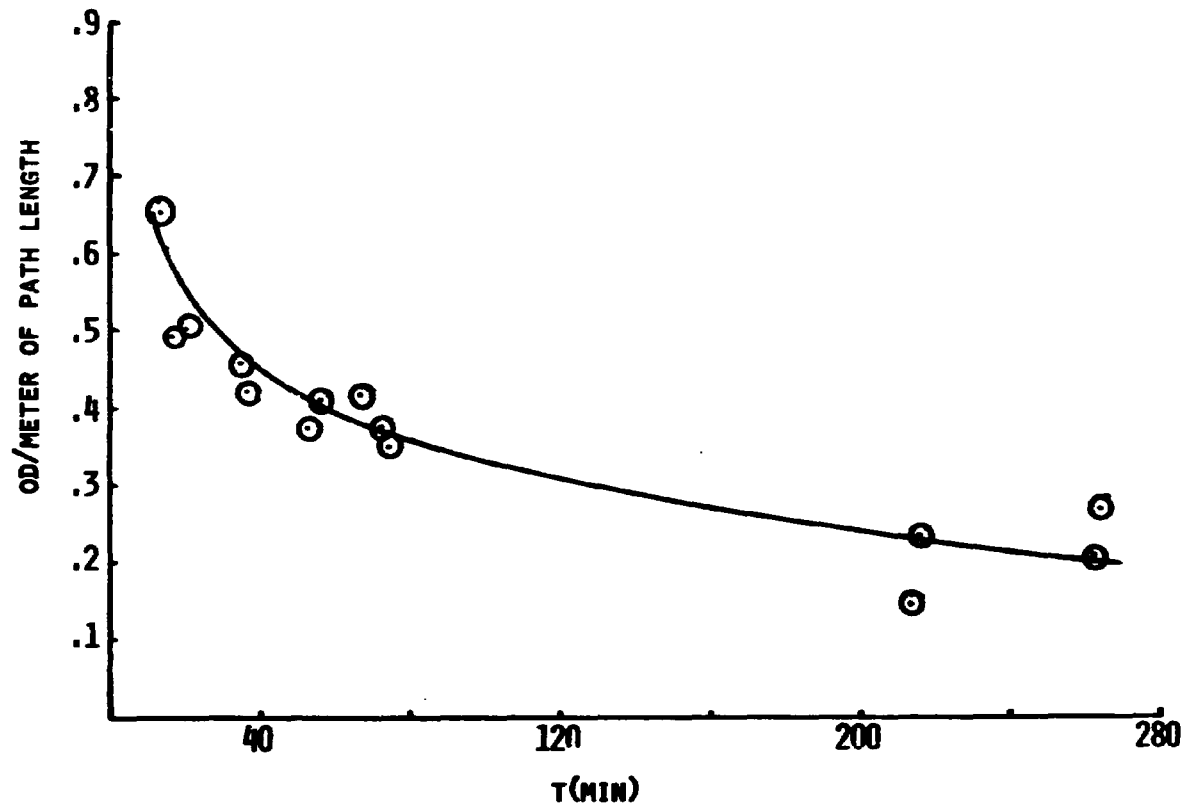
$$OD = - \log I/I_o = \log (V_S/V_R)_{air} + \log (V_R/V_S)_{aerosol} \qquad (2)$$

where V_S is the oscilloscope voltage corresponding to the sample detector and V_R is the voltage corresponding to the reference detector. In Figure 4 the results of the measurements reported as OD/meter of path length versus time is shown.

4. CONCLUSIONS. The optical density for the dye-bearing microemulsion clouds as a function of particle size and number density agrees fairly well with the results of the Mie calculations of the PGAUS program. For example, it can be seen that at 40 minutes after dissemination the OD/meter for the cloud is 0.45, the measured number density recorded at that time is 1.25×10^6 with an average particle size of 0.9μm. Theoretical calculations from the Mie program predict an OD/meter of 0.66. The difference between observed (0.45) and calculated (0.66) may be partially accounted for by scattered light in the forward direction which is gathered by the detector and recorded, resulting in a lower OD than was predicted on the basis of transmitted light. (Scattered light in the forward direction is counted as light not transmitted in the PGAUS program.)

Future experiments on the generation of finer particle sizes will include the use of explosive disseminators and propellant actuated devices. In addition, methods to increase the concentration of the dye in the present microemulsion as well as the use of such materials as dye-impregnated plastics will be investigated. Modifications to the computer program will include the use of multiple scattering routines and the separate treatment of forward and backward scattered light.

FIGURE 4 - OPTICAL DENSITY OF IR-99 AEROSOL/METER OF PATH LENGTH AS A FUNCTION OF TIME AFTER DISPERSAL BY SPRAY NOZZLE



REFERENCES.

1. E. W. Stuebing and J. Pinto,"PGAUS-LT--A Program for Computing Optical Properties of Single Scattering Aerosol Clouds of Homogeneous Particles", Frankford Arsenal report in preparation.

2. R. Mackay, "A Study of Microemulsions for Laser Absorbers," Drexel University, report in preparation, classified.

APPENDIX A: LUMINOUS TRANSMITTANCE. The output from a PGAUS-LT calculation provides the phase function as a function of μ (cosine of the scattering angle) along with scattering and absorption cross sections (C's) and efficiencies (Q's) for each of the wavelengths, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, and 1.06μm. These cross sections are combined to give the cross section for total extinction at each wavelength, which is used to calculate and tabulate the optical density (OD) per meter as a function of number density at each wavelength. Finally, these OD's for the wavelength 0.40-0.70μm are used to calculate the luminous transmittance (LT) through the aerosol over various path lengths as a function of number density from the formula,

$$LT: \frac{\sum_{\lambda} S(\lambda)\ 10^{-OD(N,\lambda)\cdot L}\ P(\lambda)}{\sum_{\lambda} S(\lambda)\ P(\lambda)} \qquad (A1)$$

where $S(\lambda)$ is the relative energy at wavelength λ of the CIE Standard Illuminant (Artificial Daylight), $OD(N,\lambda)$ is the optical density per meter at wavelength λ of the aerosol with number density N (cm^{-3}), L is the path length in meters, and $P(\lambda)$ is the relative luminosity coefficient at wavelength λ for the CIE Standard Photopic Observer.

TRANSFORMATIONS THROUGH A NON-EUCLIDEAN SPACE IN A LINEAR TRANSFORMATION CONTEXT

Application of First-Degree-Affine Transformation To Probability Density Functions in the α-Log Space

OREN N. DALTON
Mathematical Services Branch
Analysis & Computation Division
White Sands Missile Range, White Sands, New Mexico 88002

ABSTRACT

This paper discusses the technique of transforming curves through non-Euclidean spaces, which in vector/matrix notation appear virtually identical to linear (or affine) transformations. The method is called first-degree-affine. Any continuous function of n class C^1 variables can be considered as a definition of the variables in non-Euclidean n-space with a non-constant metric. An additional space is defined orthogonal to this n-space characterized by the parameter t, and the transformation is defined relative to this non-Euclidean n+1 space in a certain manner. The Dirac Delta function, $\delta(t-1)$, is introduced and it is shown that the symbolic integral is nothing more than evaluating the transformed function on a cross section in the t-dimension; on the hyperplane, t=1. In this paper the technique is applied, in particular, to a space called the "α-log" space, and it is shown that such functions as the α-log, Chi-squared, Maxwell (and Rayleigh), Gamma and Normal probability density functions are members of an equivalence class to the limit of a defining parameter, under the aegis of the first-degree-affine transform. It is pointed out that there are only about three non-commensurate spaces which encompass almost all standard continuous probability density functions.

The research for this study was supported in part by the Army In-House Laboratory Independent Research Program (ILIR), Agency Accession Number DA OK 1512, Work Unit Number IT 061101A91A 00 055.

TRANSFORMATIONS THROUGH A NON-EUCLIDEAN SPACE IN A LINEAR TRANSFORMATION CONTEXT

Application of First-Degree-Affine Transformations to Probability Density Functions in a Particular Space

I. INTRODUCTION. The term "first-degree-affine" [1] was coined by the author, and the transformation to which the technique refers, as far as can be determined, was invented by the author. The reason for this choice of a name should become clear later.

Essentially, [2] the technique allows one to transform curves through non-Euclidean spaces in the same manner a linear transformations are used to transform curves through Euclidean spaces. (In fact, the appearance of a first-degree-affine transformation is identical to that of a linear transformation.) Ignoring the additive constant implied by the integration, an affine transformation has the following property:

Let $x_1, x_2, \ldots, x_n$ be variables of class C^1 [3] (that is, they have continuous first partial derivatives over their domain) and let y be a continuous function of the x_i, i=1, 2, ..., n

$$y = y(x_1, x_2, \cdots x_n)$$

so that

$$dy = \sum_{i=1}^{n} \frac{dy}{dx_i}\, dx_i$$

If y is affine, we can write

$$y = \sum_{i=1}^{n} \frac{\partial y}{\partial x_i} x_i$$

Then, [3] y is affine if and only if the partials are constants (or, at most, functions of a parameter.) We note, using the symbols, "<" and ">", the Dirac "bra" and "ket", for row and column vectors, respectively, that y can be written as

$$y = \langle J_1 X \rangle$$

where

$$\langle J_1 = \left(\frac{\partial y}{\partial x_1} \quad \frac{\partial y}{\partial x_2} \cdots \frac{\partial y}{\partial x_n} \right)$$

is the Jacobean matrix, and

$$X\rangle = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

As was shown [1, 2] there exists certain classes of functions in which the partials are functions of the x_i's but, nevertheless, under the stated restrictions, can be transformed in the same manner as above. For example, let

$$y = \frac{x^2}{Z}$$

then

$$\langle J_1 = \left(\frac{2x}{z} \quad -\frac{x^2}{z^2}\right)$$

and

$$\langle J_1 X\rangle = \left(\frac{2x}{z} \quad -\frac{x^2}{z^2}\right)\begin{bmatrix} x \\ z \end{bmatrix} = \frac{2x^2}{z} - \frac{x^2}{z} = \frac{x^2}{z}$$
$$= y$$

Functions (which can be) written in this manner are called first-degree-affine transformations on the variables.

The bulk of [1] explores methods for converting any continuous function of C^1 variables to first-degree-affine form and then recovering the correct final function. This with the theorems in [2] show that any continuous vector function or functional can be respresented in first-degree-affine form, and a first-degree-affine transformation can always be found which transforms any continuous function into another function in a manner whose appearance is identical to that of a linear transform. In addition, it is readily apparent that the first-degree-affine transformation is to non-Euclidean spaces as the linear transformation is to Euclidean spaces.

To illustrate these ideas, observe the following simple function

$$y = x^2$$

This equation is interpreted as a graph in two dimensional Euclidean space, and in such a space describes a parabola. But this is one of two

interpretations. It is equally correct to view this equation (or, perhaps, just the term "x^2") as the definition of a variable, called "x", with a metric 2x. That is, x is defined in a one-dimensional non-Euclidean space having the metric 2x.

Obviously

$$y \neq \frac{dy}{dx} x$$

In order to convert this equationto a form which can be written as a matrix transformation of a vector [1] we introduce the parameter "t" which, by definition, defines a space orthogonal to the non-Euclidean space of x.

Define the Euclidean graph

$$y^* = \frac{x^2}{t}$$

Then if

$$X\rangle = \begin{bmatrix} x \\ t \end{bmatrix}$$

we know from above that

$$y^* = \langle J_1 X \rangle$$

where $\vec{J}_1$ is the Jacobean matrix. We now introduce the concept of a symbolic integral. Let $\delta(t - \xi)$ be the Dirac delta function. Then [4]

if $f(x,t)$ is any function of x and t we have that

$$\int_{-\infty}^{\infty} f(x,t)\,\delta(t-\tau)\,dt = f(x,\tau)$$

In discussions of the symbolic integral and, in particular, the symbolic function $\delta(t-\tau)$, it seems to have been overlooked [1, 2] that the symbolic integral using the Dirac delta function is no more than a statement that $f(x,\tau)$ is an evaluation of $f(x,t)$ on the cross-sectional cut in the hyperplane: $t=\tau$, scaled by the value of τ.

That is, in the problem above

$$\int_{-\infty}^{\infty} y^* \delta(t-1)\,dt = \int_{-\infty}^{\infty} \frac{x^2}{t}\delta(t-1)\,dt = \left.\frac{x^2}{t}\right|_{t=1} = x^2$$

$$= y$$

To underline this point of view, y^* is graphed in Figure (1) in three-dimensional Euclidean space. Note that as $t\to 0$, the (cutaway) surface narrows and steepens like a box canyon and pinches together toward the y-axis, and as $t\to\infty$ the curve flattens toward a line parallel to the x-axis in the x-t plane. The surface has very much the appearance of an alluvial fan. Any cross-sectional cut by a plane $t=t_1$, contains a parabola scaled by $1/t_1$, but in a cross section of the plane, $t=1$, the embedded curve is the parabola x^2.

Briefly, the method for converting a function to first-degree-affine form [1, 2] is as follows:

1. Assume that n distinct variables are defined in a Euclidean n-space, E^n. Increase this space to n+1 by defining a space orthogonal to E^n with the parameter t.

2. The terms in the function are multiplied by appropriate values of t so that the algebraic sums of the powers of the variables equal unity.

3. The arguments of transcendental functions and variables of variable power are "zeroed out" and the term multiplied by t. As examples, the following functions are first-degree-affine:

(1) $$x^n t^{1-n}$$

(2) $$\frac{x^4 y^3}{z^4} t^{-2}$$

(3) $$t \cos\left(\frac{w^2 x}{z} t^{-2}\right)$$

(4) $$\frac{t^4}{w^3} e^{-[(x/t)^2 + (w/t)^2]}$$

(5) $$t\left(\frac{y}{t}\right)^{x/t} = t^{1-x/t} y^{x/t}$$

(6)

$$t\left[\cos\left(\frac{w^2}{t^2}\right)\right]^{x/t}$$

and may be represented in the form

$$\langle J, X \rangle$$

where $\vec{X}$ includes all the variables in the function, including "t".

In a sense, the first-degree-affine transformation is a formalization of the process of substitution which has long been used. The same statement can be made about the linear transformation. However, either type of transformation provides insights not always apparent from the basic equation; the power of either technique as a computational tool can only be appreciated after using it.

The rest of this paper is devoted to a discussion of only one space which is generated by certain probability density functions and named, herein, the α-log space. The first-degree-affine transform illustrates groups of spaces in which the functions, under certain restrictions (in this case transformations using constant parameters), belong to an equivalence class. In [2] it is shown that three or four spaces or equivalence classes encompass virtually all continuous probability density functions (PDF's) normally in use. The following table is a listing of some of them. (The "Parimensic" space refers to functions of two independent α-log variables, a Quotient space which includes the Beta, F, Student-t and Cauchy PDF's and a Product space which includes a modified Bessel function of the second kind [5]):

SPACE

	UNIFORM	α-LOG	PARIMENSIC
PDF	Uniform	α-log	Beta
	(Cauchy)	Chi-squared	F
	Tangent	Gamma	Student-t
	Logistic	Maxwell	Bessel (Mod of 2nd kind)
	Exponential	Rayleigh	
		Normal	(Cauchy)
		(Exponential)	
	Ratio		
	Power		

We note that the intersections of the spaces are not empty (Exponential, Cauchy) but, in general, functions which define the spaces cannot be transformed from one space to another. Such PDF's as the Cauchy, Exponential and Normal are defined by explicit values for the defining parameters. The Cauchy PDF, for example, arises in the Uniform space as one member of the Tangent sequence, and in the Parimensic space as one member of the F-sequence. (The word "sequence" in this context implies a succession of functions based on rational or integral values of the defining parameters; for certain functions it is synonymous with the definition "degrees-of-freedom".) Generalized functions define each space, the Tangent, Gamma, Beta, for example. These functions can be transformed into others in the same space of which a member may be given a particular name , but, in general, a specific member cannot generate the space. The meaning of this paragraph will, hopefully, become clearer in the text.

NOTATION. Certain terms will be used in a standard manner:

I is a unit matrix,

I_n is a unit n×n Matrix,

$\langle I$ or $\underline{1}\rangle$ is the summing vector, a vector of ones,

$\langle I_n$ or $I_n\rangle$ is a summing vector having n components,

$\langle e_i$ or $e_i\rangle$ is the i^{th} orthonormal vector. I.e., a vector with zero's in every position except the i^{th} position which has a one,

$\delta(t-\tau)$ is the Dirac Delta function which has meaning [4] only in an integral of the form

$$\int_{-\infty}^{\infty} f(\vec{x}, t)\delta(t-\tau)\,dt = f(\vec{x}, \tau)$$

and is tantamount to the evaluation of $f(\vec{x}, t)$ on the cross-section of the hyperplane, $t=\tau$.

t is the parameter used to describe a space orthogonal to the space of any other variables involved in a particular problem, and $-\infty < t < \infty$.

$|J_x|$, J_x a matrix, will <u>always</u> imply the absolute value of the determinant.

$\langle \phi$ or $\underline{0}\rangle$ are vectors of zeros.

THE α-LOG SPACE. This name was chosen because the α-log density, described in this space, generates some of the more important standard PDF's. If $f(v)$, the PDF for the variable v, is the α-log density, then . . .

$$f(v) = \frac{1}{\Gamma(\alpha)} \ln^{\alpha-1}(1/v), \quad \alpha > 0, \quad \alpha \neq 1, \quad 0 < v < 1$$

or letting

$$K \triangleq \frac{\delta(t-1)}{\Gamma(\alpha)} ,$$

$$f(\vec{v}) = K \ln^{\alpha-1}\langle e_1 V\rangle^{-1} ,$$

where

$$V\rangle = \begin{bmatrix} v \\ t \end{bmatrix}$$

The requirement that α be greater than zero is necessary because of the term $\Gamma(\alpha)$, and if $\alpha = 1$, $f(v) = 1$, $0 < v < 1$, the Uniform PDF. Since a direct transformation, between the Uniform space and the α-log space is probably impossible, $\alpha \neq 1$ is excluded.

The function

$$\frac{1}{\Gamma(\alpha)} \ln^{\alpha-1}(1/v) , \quad 0 < v < 1$$

describes a variable in a one dimensional non-Euclidean space with metric

$$\frac{(1-\alpha)}{\Gamma(\alpha)} \cdot \frac{1}{v} \ln^{\alpha-2}(1/v)$$

This function does not appear to describe a spherical surface but there exists a transform which warps this space into a function (the Chi-squared PDF) which is proportional to a spherical surface. The space defined by $f(v)$ has a curvature of $\Gamma^2(\alpha)$, since

$$\int_0^1 \ln^{\alpha-1}(1/v)\, dv = \Gamma(\alpha)$$

The α-log space is defined (using a single variable) by transformations involving functions with real parameters. The intent of this statement is to exclude such transformations as y^y, $\sin^y y$, etc.[1] Since the α-log space forms an equivalence class, i.e.: if J,K are first-degree-affine transformations, then

$\vec{x} = J\vec{x}$, where $J \neq I$, the unit matrix

If $\vec{y} = J\vec{x} \Rightarrow \vec{x} = J^{-1}\vec{y}$

If $\vec{y} = J\vec{x}$ and $\vec{z} = K\vec{y}$ then $\vec{z} = KJ\vec{x}$,

[1] As a matter of fact such transformations result in functions which are logs of logs or logs of trigonometric functions. Although these functions are legitimate members of the α-log space, they lack the interest of the other functions and would lead the discussion too far afield.

and the properties of a vector space hold, the resultant functions from any (general) first-degree-affine transformation on $f(\vec{v})$ can be used to describe the properties of any other form of the variable. The term "general transformation" implies a transform using parameters (excluding the types of functions mentioned above) which are allowed to assume any real value within a given domain. For this paper, the parameters employed will be called α and β where $\alpha,\beta>0$ and $\alpha\neq 1$. In particular, we will examine the exponential form of the functions since most functions of interest are in that form.

As above, we represent the α-log PDF as

$$f(\vec{v}) = K \ln^{\alpha-1}\langle e, V\rangle^{-1}, \quad 0<v<1, \; -\infty<t<\infty, \; \alpha>0, \; \alpha\neq 1$$

In this section some of the details of the transformation will be presented; these will be, in themain, dispensed with thereafter.

Let

$$X\rangle \triangleq \begin{bmatrix} x \\ t \end{bmatrix}$$

where

$$X\rangle = \begin{bmatrix} x \\ t \end{bmatrix} = \begin{bmatrix} t\ln(t/v) \\ t \end{bmatrix}$$

This relationship is first-degree-affine so we can write:

$$|x\rangle = \begin{bmatrix} \frac{\partial x}{\partial v} & \frac{\partial x}{\partial t} \\ \frac{\partial t}{\partial v} & \frac{\partial t}{\partial t} \end{bmatrix} |v\rangle = \begin{bmatrix} -\frac{t}{v} & 1+\ln\frac{v}{t} \\ 0 & 1 \end{bmatrix} |v\rangle = J_x |v\rangle$$

We have

$$J_x = \begin{bmatrix} -e^{x/t} & 1+\frac{x}{t} \\ 0 & 1 \end{bmatrix}$$

and

$$J_x^{-1} = \begin{bmatrix} -e^{-x/t} & e^{-x/t}+\frac{x}{t}e^{-x/t} \\ 0 & 1 \end{bmatrix}$$

$$|J_x^{-1}| = e^{-x/t}$$

Thus

$$f(\vec{x}) = K|J_x^{-1}| \ln^{\alpha-1}\langle e_1 J_x^{-1} x\rangle^{-1}$$

and

$$J_x^{-1} X\rangle = \begin{bmatrix} t\, e^{-x/t} \\ t \end{bmatrix}$$

from which

$$\langle e_1 J_x^{-1} X\rangle^{-1} = \frac{1}{t} e^{x/t}$$

Substituting

$$f(\vec{x}) = \frac{\delta(t-1)}{\Gamma(\alpha)} e^{-x/t} \left[\ln\left(\frac{1}{t} e^{x/t}\right)\right]^{\alpha-1}$$
$$= \frac{\delta(t-1) e^{-x/t}}{\Gamma(\alpha)} \left[x/t - \ln t\right]^{\alpha-1}$$

Then

$$f(x) = \int_{-\infty}^{\infty} f(\vec{x})\, dt = \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} ,$$

$$0 < x < \infty$$

which is the Gamma PDF for a variable y=2x.

If z=2x and $\alpha = k/2 \Rightarrow dx = 2\, dx$

then

$$f(z) = \frac{z^{k/2-1}\, e^{-z/2}}{2^{k/2}\, \Gamma(k/2)}, \quad 0 < z < \infty$$

the Chi-squared PDF.

There are two levels of generality which might be introduced for transformations of one variable and constant parameters. The most general form can be written as

$$\vec{X} = \begin{bmatrix} x \\ t \end{bmatrix} = \begin{bmatrix} t \ln^{q/p}(t/v)^{s} \\ t \end{bmatrix} = J_x \vec{V}$$

since this transformation is first-degree-affine, and

$$J_x = \begin{bmatrix} -\frac{\alpha s}{p} \frac{t}{v} \ln^{q/p-1}(t/v)^{s} & \frac{x}{t} + \frac{\alpha s}{p} \ln^{q/p-1}(t/v)^{s} \\ 0 & 1 \end{bmatrix}$$

except that the power of the log could vary as, say $-\frac{q}{p}$, $\pm\frac{1}{\alpha p}$, $\pm\frac{1}{p}$ and the argument of the log could be v/t. However, any functions involving a transformation using the two parameters, p and q, is equivalent to any other irrespective of the values of p and q. The effect of this generality will be seen when Pearson's Incomplete Gamma Function tables are used to find the distribution of the sum of the squares of n independent normal variables to the $\frac{1}{m}$ power -- defined, herein, as the Maxwell distribution.

In the above problem, since the transformation is first-degree-affine and since

$$f(\vec{v}) = K \ln^{\alpha-1}\langle e, V\rangle^{-1} ,$$

we have that

$$f(\vec{x}) = K \left|J_x^{-1}\right| \ln^{\alpha-1}\langle e, J_x^{-1} X\rangle^{-1} ,$$

$$J_x^{-1} = \begin{bmatrix} \frac{-p}{\alpha q} e^{-\frac{1}{q}(x/t)^{p/\alpha}} \left(\frac{x}{t}\right)^{\frac{p}{\alpha}-1} & e^{-\frac{1}{q}(x/t)^{p/\alpha}} \left[1 + \frac{p}{\alpha q}\left(\frac{x}{t}\right)^{p/\alpha}\right] \\ 0 & 1 \end{bmatrix},$$

$$\left|J_x^{-1}\right| = \frac{p}{\alpha q} e^{-\frac{1}{q}(x/t)^{p/\alpha}} \left(\frac{x}{t}\right)^{\frac{p}{\alpha}-1} , \quad J_x^{-1} X\rangle = \begin{bmatrix} t\, e^{-\frac{1}{q}(x/t)^{p/\alpha}} \\ t \end{bmatrix}$$

and

$$f(\vec{x}) = \frac{p}{\alpha q} \frac{\delta(t-1)}{\Gamma(\alpha)} e^{-\frac{1}{q}(x/t)^{p/\alpha}} \left(\frac{x}{t}\right)^{\frac{p}{\alpha}-1} \left[\frac{1}{q}\left(\frac{x}{t}\right)^{p/\alpha} - \ln t\right]^{\alpha-1}$$

from which

$$f(x) = \int_{-\infty}^{\infty} f(\vec{x})\, dt = \frac{p}{\alpha q^{\alpha} \Gamma(\alpha)} x^{p-1} e^{-\frac{1}{q}x^{p/\alpha}} , \quad 0 < x < \infty$$

(We note that $\int_0^\infty f(x)dx = \frac{\alpha q^\alpha \Gamma(\alpha)}{p}$.) In the special case when $q=1$, we have

$$f(x; q=1) = \frac{p}{\alpha\Gamma(\alpha)} x^{p-1} e^{-x^{p/\alpha}}$$

and we see that the parameter q, acts as no more than a scaling constant.

If $p=1, \alpha=\frac{1}{2}$ and $q=2$, then

$$f(x; p=1, \alpha=\tfrac{1}{2}, q=2) = \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad 0<x<\infty$$

the normal PDF over the right half-plane. We might also note that if $p=\alpha$, then

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

the Gamma PDF.

It turns out that the left half-plane, $-\infty<x<0$, is defined by a transformation involving v/t rather than t/v. For a similar transformation as above, let

$$f(\vec{v}) = K \ln^{\alpha-1}\langle e, V\rangle^{-1}, \quad \begin{matrix} 0<v<1, -\infty<t<\infty, \\ \alpha>0, \alpha\neq 1 \end{matrix}$$

and define

$$\vec{X} = \begin{bmatrix} x \\ t \end{bmatrix} = \begin{bmatrix} t \ln^{\alpha/p}(v/t)^{q} \\ t \end{bmatrix} = J_x \vec{V}$$

from which we know that

$$f(\vec{x}) = K|J_x^{-1}| \ln^{\alpha-1}\langle e_1, J_x^{-1} X\rangle^{-1}, \quad -\infty < x < 0$$

We have that

$$J_x = \begin{bmatrix} \frac{\alpha q}{p} \frac{t}{v} \ln^{\frac{q}{p}-1}(v/t)^{q} & \frac{x}{t} - \frac{\alpha q}{p} \ln^{q/p-1}(v/t)^{q} \\ 0 & t \end{bmatrix},$$

$$J_x^{-1} = \begin{bmatrix} \frac{p}{\alpha q} e^{\frac{1}{q}(x/t)^{p/\alpha}} \left(\frac{x}{t}\right)^{\frac{p}{\alpha}-1} & e^{\frac{1}{q}(x/t)^{p/\alpha}} \left[1 - \frac{p}{\alpha q}\left(\frac{x}{t}\right)^{p/\alpha}\right] \\ 0 & t \end{bmatrix},$$

$$J_x^{-1} X\rangle = \begin{bmatrix} t\, e^{\frac{1}{q}(x/t)^{p/\alpha}} \\ t \end{bmatrix}, \quad |J_x^{-1}| = \frac{p}{\alpha q}\left(\frac{x}{t}\right)^{\frac{p}{\alpha}-1} e^{\frac{1}{q}(x/t)^{p/\alpha}}$$

and we have

$$f(x) = \int_{-\infty}^{\infty} f(\vec{x})\, dt$$

$$= \frac{p}{\alpha q \Gamma(\alpha)} \int_{-\infty}^{\infty} \left(\frac{x}{t}\right)^{\frac{p}{\alpha}-1} e^{\frac{1}{q}(x/t)^{p/\alpha}} \left[\frac{1}{q}\left(\frac{x}{t}\right)^{p/\alpha} - \ln t\right]^{\alpha-1} \delta(t-1)\, dt$$

We note that this transformation is the same as the previous transformation had we used $-q$ instand of q (all except the $\frac{1}{q}$ scale factor which is positive since $|J_x^{-1}|$ is always positive.) Thus, this can be written as

$$f(x) = \frac{p}{\alpha q^{\alpha} \Gamma(\alpha)} |x|^{p-1} e^{-\frac{1}{q}|x|^{p/\alpha}}, \quad -\infty < x < 0$$

As before, when $p = 1$, $\alpha = 1/2$, $q = 2$:

$$f(x; p=1, \alpha=1/2, q=2) = \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad -\infty < x < 0$$

Putting the two half-planes together and averaging (since the probability that $x=0$ is 0) we have

$$f(x; p=1, \alpha=1/2, q=2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad -\infty < x < \infty$$

Also when $p = \alpha$:

$$f(x;\ p=\alpha) = \frac{1}{\theta^{\alpha}\Gamma(\alpha)} |x|^{\alpha-1} e^{-\frac{1}{\theta}|x|} \quad , -\infty < x < 0$$

If we wished to extend the Gamma function to both half-planes we would have

$$f(x;p=\alpha) = \frac{1}{2\theta^{\alpha}\Gamma(\alpha)} |x|^{\alpha-1} e^{-\frac{1}{\theta}|x|} \quad , -\infty < x < \infty$$

although this extended range does not appear to have much practical value.

II. INTEGRATION IN THE α-LOG SPACE USING PEARSON'S INCOMPLETE GAMMA FUNCTION TABLES. It was proved [2] for transformations involving a single variable, that[1]

$$\int_0^{x'} f(x)\,dx = \int_0^{y'} f(y)\,dy$$

where $\vec{x} = J_x \vec{y}$ is a first-degree-affine transform and

$$f(x) = \int_{-\infty}^{\infty} |J_x^{-1}|\, f(J_x^{-1}\vec{x})\, \delta(t-1)\,dt$$

---

[1] NOTE: Since the differentials are treated as components in a Euclidean space, then $|\vec{dx}|$ is tantamount to $\left|\int_0^{dx_1} \cdots \int_0^{dx_n}\right|$, and for any transform $\vec{dx} = J\,\vec{dy} \Rightarrow |\vec{dx}| = |J| \cdot |\vec{dy}| = |J|\,dy_1 \cdots dy_n$. For first-degree-affine (or linear) transforms, the differentials can be replaced by the vector itself.

if there are no discontinuities in the range of y over the domain $0<y<y'$. This is based on the following observations. Assume that $\vec{Y}$ is continuous over the interval: $\vec{Y}_0 < \vec{Y} < \vec{Y}_f$. Then we can write: $\vec{Y}_0 < J_x^{-1}\vec{X}, J_x^{-1}\vec{X} < \vec{Y}_f$. If there are no discontinuties due to the transformation, the two inequalities can be combined so that $\vec{Y}_0 < J_x^{-1}\vec{X} < \vec{Y}_f$. We can solve this system to produce the inequalities: $\vec{X}_0 < \vec{X} < \vec{X}_f$, if the variables are at all separable. (We note that $\vec{X}_0$ and $\vec{X}_f$ may be vector functions of the components of $\vec{X}$.)

Without computing details, observe the following. Let

$$f(\vec{v}) = K \ln^{\alpha-1}\langle e, V\rangle^{-1} ,$$

$$0 < v < 1, \quad -\infty < t < \infty, \quad \alpha > 0, \quad \alpha \neq 1,$$

as above, and define two first-degree-affine transformations

$$Y\rangle = \begin{bmatrix} t\ln(t/v) \\ t \end{bmatrix} = J_Y V\rangle , \quad 0 < y < \infty, \quad -\infty < t < \infty,$$

$$X\rangle = \begin{bmatrix} \sqrt{\alpha}\, t \ln^{1/2}(t/v) \\ t \end{bmatrix} = J_X V\rangle , \quad 0 < x < \infty, \quad -\infty < t < \infty,$$

$$\int_0^{v'} K \ln^{\alpha-1}\langle e, V\rangle^{-1} dv = \int_0^{y'} K \left|J_Y^{-1}\right| \ln^{\alpha-1}\langle e, J_Y^{-1} Y\rangle^{-1} dy$$

$$= \int_0^{x'} K \left|J_x^{-1}\right| \ln^{\alpha-1}\langle e, J_x^{-1} X\rangle^{-1} dx$$

where

$$y' = \langle e_1, J_y V' \rangle$$

$$x' = \langle e_1, J_x V' \rangle$$

Since

$$\vec{Y} = J_y \vec{V}\,, \quad \vec{X} = J_x \vec{V} \Rightarrow \vec{Y} = J_y J_x^{-1} \vec{X}$$

and computing we would have

$$J_y J_x^{-1} X\rangle = \begin{bmatrix} -t/v & 1 + \ln(t/v) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -xv/t^2 & \frac{v}{t} + \frac{x^2 v}{t^3} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix}$$

$$= \begin{bmatrix} \frac{x}{t} & -\frac{1}{2}\frac{x^2}{t^2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} = \begin{bmatrix} \frac{x^2}{2t} \\ t \end{bmatrix}$$

Thus, after integrating over t we have that

$$y' = \frac{x'^2}{2}$$

and that

$$\frac{1}{\Gamma(\alpha)} \int_0^{y'} y^{\alpha-1} e^{-y}\, dy = \frac{2^{2-\alpha}}{\Gamma(\alpha)} \int_0^{x'} x^{2\alpha-1} e^{-\frac{1}{2}x^2}\, dx$$

$0 < y' < \infty$ $\qquad$ $0 < x' < \infty$

In particular, if $\alpha = 1/2$ we can use the Γ-tables [6] to compute the normal distribution and cross compare, i.e.

$$\frac{1}{\sqrt{\pi}} \int_0^{y'} y^{-1/2} e^{-y} dy = 2 \frac{1}{\sqrt{2\pi}} \int_0^{x'} e^{-\frac{1}{2}x^2} dx$$

$0 < y' < \infty$ $\qquad$ $0 < x' < \infty$

The accuracy of the comparison, perforce, depends on fineness of the mesh of the tabulated values.

The tables [6] are published in terms of I(u,p) (Pearson's notation) where

$$I(u,p) = \frac{1}{\Gamma(p+1)} \int_0^{\sqrt{p+1}\,u} v^p e^{-v} dv \quad , \quad 0 < u < \infty$$

Thus, for $p = \alpha - 1, \Rightarrow p = -\frac{1}{2}$ for $\alpha = \frac{1}{2}$, and since $y' = \sqrt{p+1}\,u$, $y' = u/\sqrt{2}$, $x' = (\sqrt{2}\,u)^{1/2}$. The tables in [7] are published as

$$\frac{1}{\sqrt{2\pi}} \int_0^{x'} e^{-\frac{1}{2}x^2} dx \quad , \quad 0 < x' < \infty$$

If we call this latter integral $J(x')$, we have that

$$I(u, p = -1/2) = 2\,J(x')$$

The following table compares the two integrals for a few values of u which were chosen from Pearson's tables (with seven-place accuracy) to the limit of a linear interpolation from the tables of [7] (with four-place acuracy).

u	$I(u, -\frac{1}{2})$	x' (app)	$J(x')$	$2J(x')$
.5	.5995940	.841	.29927	.59954
2	.9073910	1.682	.4537	.9074
4	.9826127	2.378	.4913	.9826
6	.9964197	2.912	.4982	.9964
10	.9998305	3.760	.4999	.9998
12	.9999620	4.120	.5000	1.0000

It was shown in [2], that if $\{\chi_i, i = 1, 2, .., n\}$ is a set of n independent standard normal variables, then if

$$w = \left[\frac{1}{2}\sum_{i=1}^{n}\chi_i^2\right]^{\frac{1}{m}}$$

that the PDF for w (called, there, the PDF for the Maxwell distribution) $f(w)$ is

$$f(w) = \frac{m}{\Gamma(\alpha)} w^{m\alpha-1} e^{-w^m}, \quad \alpha > 0, m > 0, 0 < w < \infty$$

If we replace α by $n/2$, then

$$F(w') = \int_0^{w'} \frac{m}{\Gamma(n/2)} w^{\frac{mn}{2}-1} e^{-w^m} dw \;, \quad 0 < w' < \infty$$

Proceeding by transformations similar to that done above it can be shown that in terms of the Gamma PDF, we have

$$\frac{1}{\Gamma(n/2)} \int_0^{y'} y^{n/2-1} e^{-y} dy = \frac{m}{\Gamma(n/2)} \int_0^{w'} w^{\frac{mn}{2}-1} e^{-w^m} dw$$

$$0 < y' < \infty \qquad\qquad 0 < w' < \infty$$

where $y' = w'^m$. Converting to Pearson's notation as was done above, $p = n/2 - 1$, and since $y' = \sqrt{p+1}\, u$, we have that

$$y' = \sqrt{n/2}\, u = w'^m \;,$$

or

$$u = \sqrt{\frac{2}{n}}\, w'^m$$

and m is a free index.

The following table gives the integral for three values each of m (the m^{th} root of $\frac{1}{2}\sum_{i=1}^{n} \chi_i^2$), n (the number of variables), and u (the value of the Γ-variable):

m	n	p	$\sqrt{n/2}$	u	w^{-m}	w^{-} (app)	F (w^{-})
2	3	½	1.225	.5	.612350	.7825	.2529242
				4	4.898800	2.2130	.9796360
				12	14.696400	3.8330	.9999981
	6	2	1.732	.5	.866000	.9310	.0573798
				4	6.929200	2.6820	.9687180
				12	20.785000	4.5570	.9999998
	8	3	2.000	.5	1.000000	1.0000	.0189882
				4	8.000000	2.8280	.9687180
				12	24.000000	4.8990	.9999999
½	3	½	1.225	.5	.612350	.3750	.2529242
				4	4.898800	24.0100	.9796360
				12	14.696400	216.0800	.9999981
	6	2	1.732	.5	.866000	.7500	.0573798
				4	6.928200	48.0250	.9687180
				12	20.785000	431.3900	.9999998
	8	3	2.000	.5	1.000000	1.0000	.0189882
				4	8.000000	64.0000	.9687180
				12	24.000000	576.0000	.9999999
3	3	½	1.225	.5	.612350	.8492	.2529242
				4	4.898800	1.6984	.9796360
				12	14.696400	2.4495	.9999981
	6	2	1.732	.5	.866000	.9532	.0573798
				4	6.928200	1.9064	.9687180
				12	20.785000	2.7495	.9999998
	8	3	2.000	.5	1.000000	1.0000	.0189882
				4	8.000000	2.0000	.9687180
				12	24.000000	2.8845	.9999999

In the above table, values for u were chosen from entries in Pearson's tables of incomplete Γ-functions to avoid interpolation. The values for w' were derived from them. Since the integral of the Γ-function is independent of m, every value of m produces the same distribution function; like its logarithmic ancestory, the only requirement is that w'^{m} be a fixed value for a given n and u. Although these statements are obvious in the above context, it is not so obvious that the distribution for any root m (or power) of the sums of squared normal variables is independent of the root.

I would like to conclude with a remark concerning transformations between the α-log space and the so-called Uniform space epitomized by

$$f(\vec{v}) = \Gamma(t-1)\langle e, V\rangle, \quad 0<v<1, \quad -\infty<t<\infty,$$

0 otherwise. Any first-degree-affine transformation in this space has an anti-derivative, i.e.

$$\int_0^{v'} dv = v' = \int_0^{w'} dw \int_{-\infty}^{\infty} |J_{\vec{w}}^{-1}| t\, dt$$

where $\vec{W} = \begin{bmatrix} w \\ t \end{bmatrix}$ is a first-degree-affine transformation of $\vec{V}$. If a transformation existed to the α-log space, and, in particular, to the normal PDF, say, we could write

$$v' = \int_0^{w'} K e^{-\frac{1}{2}w^2}\, dw$$

a closed form solution.

There is another argument against this possibility. The Uniform PDF is related to the α-log PDF only when $\alpha = 1$. Thus, we could expect that any transformation in the α-log space in which the resultant function is determined for $\alpha = 1$ would be derivable from the Uniform PDF. This, indeed, proves to be the case. This is theprincipal reason that the value $\alpha = 1$ was forbidden. Although it has been stated that the Uniform PDF could be warped into any other continuous PDF, this statement, on the basis the material in [2] and this paper, should be qualified, that such is true only if it is modified in the process by an "external" source. The Uniform PDF and an α-log variable are incommensurable for essentially the same reason a plane and a sphere are incommensurable.

III. FIRST-DEGREE-AFFINE TRANSFORMATIONS WHICH TRANSFORMS THE SPACE OF N VARIABLES INTO A SPACE OF M VARIABLES, M<N. It was shown in [2] that if n variables defined in a function appear as a quadratic form, and a transformation is desired to an m-dimensional subspace, $m < n$, then the affine (or linear) and the first-degree-affine transformation can be used (in part) as an $m \times n$ matrix. This is true, for example, for the Maxwell distribution defined in the last section, and in particular when the root, m, is equal to one. In this case the sums of squares of independent standard normal variables results in the Chi-squared PDF, a particular member of the Maxwell PDF.

On the other hand, the Jacobean must be square; the resultant function involves n variables and $n-m$ variables must be integrated out. Thus the Maxwell or Chi-squared PDF's are marginal PDF's. (In this sense, all PDF's which result using the first-degree-affine transformation are

marginal since t must be integrated out.) Usually the extra $n-m$ variables can be transformed unchanged.

These ideas will be illustrated for the sums of the squares of standard normal variables.

Define

$$|X_1\rangle = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad |X\rangle = \begin{bmatrix} X_1\rangle \\ t \end{bmatrix}, \quad g(t) = e^{-\frac{1}{2}(t^2-1)} \delta(t-1)$$

in which

$$x_i \sim N(0,1), \forall x_i \in \vec{X}_1$$

then

$$f(\vec{x}) = g(t) f(\vec{x}_1) = K e^{-\frac{1}{2}\langle XX\rangle}$$

where

$$K = \frac{e^{\frac{1}{2}} \delta(t-1)}{(2\pi)^{n/2}}$$

is the PDF for $\vec{X}$.

Now, let

$$W\rangle = \begin{bmatrix} u \\ t \end{bmatrix} = \begin{bmatrix} \frac{1}{2t}\langle X_1 X_1\rangle \\ t \end{bmatrix} = \begin{bmatrix} x_1/t & x_2/t & \cdots & x_n/t & \frac{-u}{t} \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \vec{X}$$

$$= J_w X\rangle$$

the first-degree-affine form which projects the n standard normal variables from the n-dimensional curved space spanned by $\vec{X}$ to the two-dimension curved space spanned by u and t. Since

$$u = \frac{\langle X_1 X_1\rangle}{2t} \Rightarrow 0 < u < \infty \quad \text{for} \quad -\infty < x_i < \infty,\ x_i \in \vec{X}_1$$

for any i, the resultant function must be multiplied by two.

If we define

$$W_1\rangle = \begin{bmatrix} u \\ x_2 \\ x_3 \\ \vdots \\ x_n \\ t \end{bmatrix}$$

and

$$J_{1w} = \begin{bmatrix} x_1/t & x_2/t & \cdots & x_n/t & -u/t \\ 0 & 1 & \cdots & 0 & 0 \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$

then $J_1 w$ is nonsingular, and $f(\vec{w}_1)$ can be written as

$$f(\vec{w}_i) = K \left| J_{iw}^{-1} \right| e^{-\frac{1}{2}\langle w (J_w J_w^T)^{-1} w \rangle}$$

$$\left| J_{iw}^{-1} \right| = \frac{t}{x_i} = \frac{t}{\sqrt{2u - x_2^2 - \ldots - x_n^2}}$$

in which the reduced $2 \times n+1$ matrix J_w appears in the quadratic form in the exponent and the Jacobean maintains the relationship of the new variables.

We have

$$J_w J_w^T = \begin{bmatrix} \frac{u^2}{t^2} + \frac{1}{t^2}\langle X_i X_i \rangle & -\frac{u}{t} \\ -\frac{u}{t} & 1 \end{bmatrix} = \begin{bmatrix} \frac{2u}{t} + \frac{u^2}{t^2} & -\frac{u}{t} \\ -\frac{u}{t} & 1 \end{bmatrix}$$

$$\left(J_w J_w^T\right)^{-1} = \begin{bmatrix} \frac{t}{2u} & \frac{1}{2} \\ \frac{1}{2} & 1 + \frac{u}{2t} \end{bmatrix}$$

so that

$$\langle w\,(J_w J_w^T)^{-1} w \rangle = 2ut + t^2$$

and

$$f(u, \chi_2, \chi_3, \cdots, \chi_n) = \int_{-\infty}^{\infty} f(\vec{w}) \, dt = \frac{2e^{-u}}{(2\pi)^{n/2} \sqrt{A_3^2 - \chi_2^2}}$$

$$= K(u)(A_3^2 - \chi_2^2)^{-1/2}$$

where

$$A_3^2 = 2u - \chi_3^2 - \chi_4^2 - \cdots - \chi_n^2$$

Integrating the terms $\chi_2, \ldots, \chi_n$, successively, and noting that for $n = 1$ f has the following form

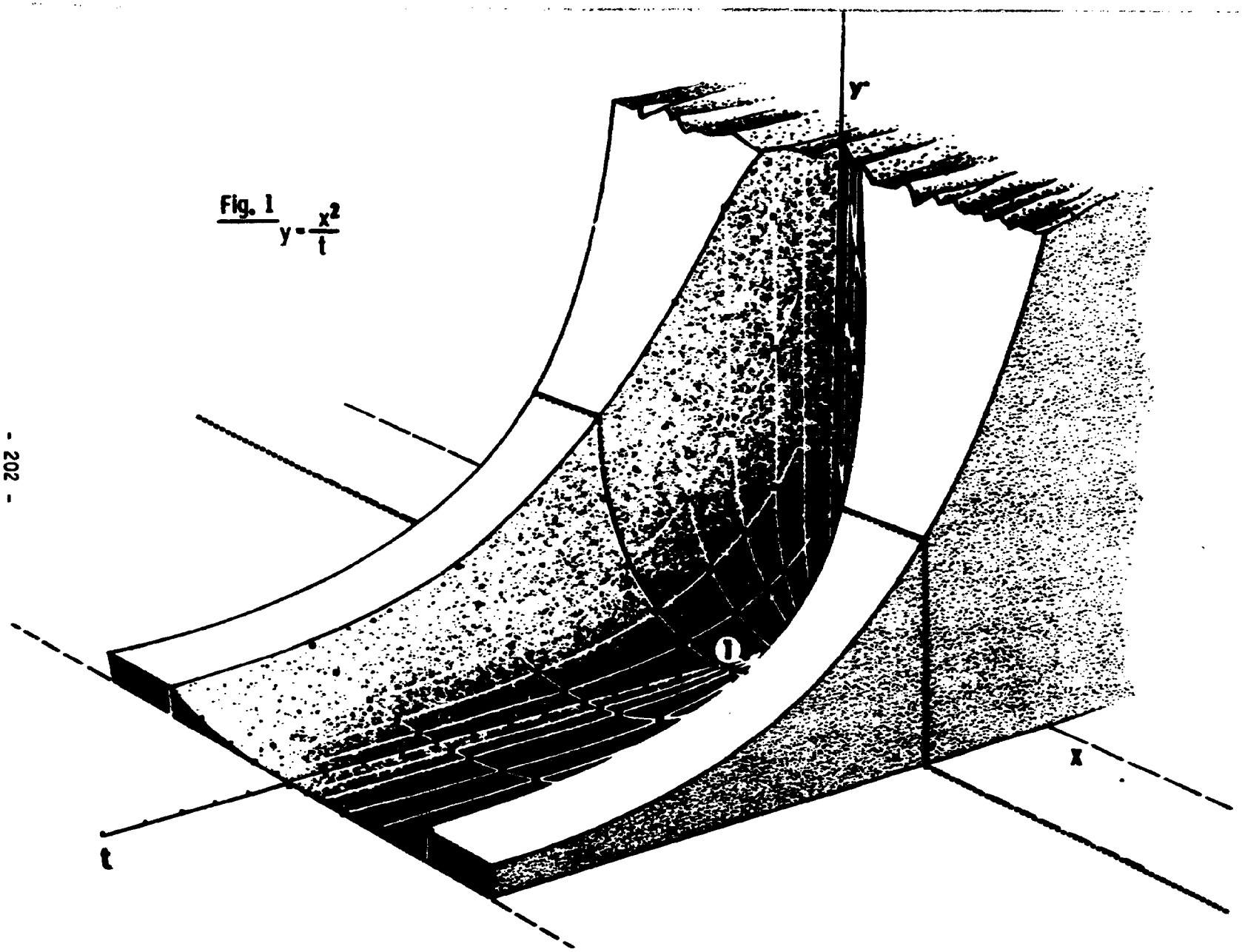
$$\frac{e^{-u}}{\sqrt{\pi u}} = \frac{u^{-1/2} e^{-u}}{\Gamma(1/2)}$$

we have, for all values for $n \geq 1$:

$$f(u) = \frac{u^{\frac{n}{2}-1} e^{-u}}{\Gamma(n/2)} ,$$

$n \in \mathbb{N}$ (the natural numbers), $0 < u < \infty$, which is the Chi-squared PDF for parameters n and $\nu = 2u$.

Fig. 1 $y = \frac{x^2}{t}$



REFERENCES

1. Dalton, Oren N., "First-Degree-Affine Formulation of Functions, Non-linear Transformations in a Linear Transformation Context", Technical Report No. 32, NR-AM-I, WSMR, NM, Jan '72.

2. Dalton, Oren N., "Transformations Through Non-Euclidean Spaces, in a Linear Transformation Context, Applied to Continuous Probability Density Functions", Technical Report, STEWS-NR-AM, WSMR, NM, Apr '73.

3. Sokolnikoff, I. S., "Tensor Analysis, Theory and Applications", John Wiley and Sons, Inc., NY, NY, 1951.

4. Friedman, B., "Principals and Techniques of Applied Mathematics", John Wiley and Sons, Inc., NY, NY, 2nd printing, 1957.

5. Hendrick, J. M., "Exact Distribution of the Product of Independent Generalized Gamma Variables with the Same Shape Parameter", The Annals of Mathematical Statistics, 39(5), 1951-2, 1968.

6. Pearson, Karl, Editor, "Tables of the Incomplete Γ-Functions", Cambridge University Press, London, England, Reissue, 1946.

7. Chemical Rubber Pub. Co., "Mathematical Table from Handbook of Chemistry and Physics", Cleveland, Ohio, 6th Printing, Sep '67.

PRESENTATION OF THE SAMUEL S. WILKS MEMORIAL MEDAL

After, the banquet on the evening of 24 October 1973, Chairman Grubbs made a few remarks about the Wilks Medalist. Professor J. S. Hunter of Princeton University headed up the 1973 committee of the American Statistical Association to make the selection. As a result, we are very pleased to announce that Professor H. O. Hartley was selected to receive the 1973 Samuel S. Wilks Memorial Medal. Professor Hartley's remarks follow.

ACCEPTANCE REMARKS OF H. O HARTLEY ON RECEIVING THE SAMUEL S. WILKS MEMORIAL MEDAL FOR 1972

It is of course a great honor for me to receive this award. In spite of your generous comments, Frank, I have a certain feeling of trepidation to join the august body of those who have received the award before me. My sincere appreciation of receiving this honor is two-fold.

First and foremost, I realize that this award is in memory of the late Samuel S. Wilks, and it is a tremendous inspiration to me to be associated with this great man. Sam Wilks is undoubtedly a pioneer, if not the pioneer of statistics in the U.S.A.. As a brilliant young scholar, he was sent to England to 'sample' both the Pearsonian School at London and R. A. Fisher's activities at Rothampstead Experiment Station. Then he was appointed to an important chair at Princeton University to develop a statistics program there. Practically all initial efforts in the science of statistics in the U.S.A. can be traced back to an input by Sam Wilks.

But Sam Wilks did more than that. He created the cooperative spirit in the statistical community in the U.S.A.. Right from the outset, he represented the blending of sound statistical theory with a strong relevance to applications of statistics in the various subject matter areas to which it is applied. This blending of theory and applications is of the utmost importance. We can appreciate this if we observe the unnecessary controversies between theoretician and practitioner that nowadays often arise:- The pure theoretician often regards applied statistics as 'substituting numbers into his equations'. He apparently does not realize that the effective application of statistics requires a sound knowledge of the subject matter area in which the data have arisen. On the other side of the fence, the hard boiled practitioner often regards much of the statistical theory as 'useless'. His definition of 'useless' often covers anything he is incapable of understanding and, unfortunately on this definition, much of the theory is labeled as 'useless'.

Sam Wilks has done a tremendous job in bridging this conflict. He has stressed again and again that we statisticians can not be expert in all subject areas to which statistics must be applied, and so we must cooperate with the subject matter area specialist. The first prerequisite for this is proper communication with the expert. The right attitude to take is to tell him "look, I am not a chemist (if you are talking to a chemist), or I am not an engineer, or I am not a clinician..., so you must allow me to ask stupid questions about your subject, and I allow you to ask stupid questions about statistics". This attitude sets the background for the team work needed in applying statistics. It was this spirit of cooperative team work that made Sam Wilks popular with practitioners and theoreticians alike, and if we can learn this lesson from him, if we can live up to this principle, then I am sure we shall do justice to his memory.

And now I would like to reminisce on my personal memories of Sam Wilks. My first contact with him was in 1949, when I had an assignment to spend the summer at Princeton. I got to know him as a most congenial Director of the Statistical Unit at Fine Hall. There was a relaxed atmosphere and yet an efficiently administered unit. He always had time for you. He was a perfect host and I was invited to his house to meet his charming wife and his son, Stanley.

John Tukey, the first recipient of this award, was then Sam's right hand man at Princeton and it was quite an experience to meet him for the first time. There are lots of stories about John that I could tell you, but let me tell you just one. We used to have lunch together at a diner on U.S.A. #1 to which he used to drive me, and immediately we arrived there, he would pull out his famous yellow pad and we got down to talking shop. We were notorious at the diner for this and hardly left time for lunch. One day Sam Wilks joined us for lunch. There were the three of us, Sam, the tall slender Texan, John, who was then very heavy and over-powering, and me. As we were entering the diner, the manager could not help blurting out "They come in all shapes and sizes in the Statistics Department!"

I did not know at the time that Sam was a Texan, and I certainly did not know that I would end up in Texas. Yes, it is true, you are looking now at what is clearly a rather imperfect simulation of a tall Texan!

An appreciation of Sam Wilks would not be complete without mentioning this 'Conference on the Design of Experiments in Army Research, Development and Testing', and this brings me to my second point of appreciation. The creation of this important series of conferences resulted from Sam Wilks' successful cooperation with the Army Research Office. This Office, indeed, the Army in general, has carried the burden of organizing this splendid annual event. The conferences have been conducted in the spirit of Sam Wilks, in the spirit of the blending of advanced methodology with practical applications. This is clearly reflected in the program of these conferences and indeed in today's proceedings.

In these days of pressure for 'mission oriented research', the Army Research Office must be congratulated for taking a responsible view of the problem. We cannot today afford unrestricted basic research no matter whether it will ever be of relevance to applications. On the other hand, it would be narrow minded, indeed reckless, to support only such research which has immediate applications to problems in DOD. Such a short sighted attitude may prevent the basic research for important new developments from being implemented (remember good old Sputnik!). The right attitude is to evaluate the relevance of basic research with regard to its application potentials. The assessment of such research potentials is sometimes extremely difficult, but the Army Research Office is to be congratulate for at least facing up to this responsible task.

Now let me finally come to my pleasant assignment of accepting this award, and this is clearly a very personal matter, and so I will end on a very personal note. At our Institute of Statistics at Texas A&M University, my wife runs a "Statistics Wives' Group" to which the wives of both faculty and students belong. When a student gets his degree, this group awards a degree, also, to his wife. She is awarded the P.H.T. (this stands for put hubby through). This degree to her is awarded for her unselfish sacrifice of letting her husband get on with his job and take the responsibility for all (and I mean all) family affairs. And so in accepting this certificate, I would like to also accept it as a P.H.T. award for my wife.

It remains, then, for me to thank the A.S.A., the original donors, and the Army Research Office for this magnificent award.

RELIABILITY GROWTH OF THE TF-30 ENGINE

Bertram W. Haines
Statistics and Mathematics Department
Naval Safety Center
Norfolk, Virginia

Frederick L. Carter
Chemical System Division
Dugway Proving Ground, Utah

Martha R. Cummings
Statistics and Mathematics Department
Naval Safety Center
Norfolk, Virginia

ABSTRACT. The introduction of any new complex system into its operational environment can be expected to be followed by a period during which improvements, designed to increase the system's reliability, will be incorporated. An aircraft engine is such a system. Changes and improvements continue to be made on aircraft engines after they are put into production and introduced operationally to the fleet. As a result of these changes and improvements, the system's reliability is expected to increase. Several methods of measuring the resulting reliability growth have been proposed. At least one of these, the system proposed by Barlow and Scheuer,* provides for both the grouping of the system development program into stages and the discrimination between two distinct causes of failure. It, therefore, seemed to be an appropriate model to describe the reliability growth of the TF-30 engine as it is used in some of the Navy's A-7 Corsair II single jet engine attack aircraft.

The Barlow-Scheuer Trinomial Model (BSTM) requires the system development period to be divided into stages that are initially defined in terms of either time periods or a given number of trials. In this study, the first of these definitions was used. The model further requires the definition of three possible outcomes for each trial. These are: 1. inherent failure, 2. assignable cause failure,

---

*Barlow, R.E., and Scheuer, E.M., "Reliability Growth During a Development Testing Program", Technometrics 8-1, February 1966.

and 3. success. Inherent failures are defined as any failures caused by factors not attributable to the system. Assignable cause failures can be assigned to one or more components of the system. Any trial that does not result in an inherent failure or an assignable cause failure is called a success.

Once the reliability growth is observed and measured using the BSTM, the relation between stages of relatively rapid reliability growth and the incorporation of power plant changes (PPC) was examined. It was found that some PPC's could not be associated with relatively rapid reliability growth, whereas, others could. This comparison was made on a small sample of PPC's that had been incorporated during the entire system development period covered by the study. The failure history of some specific components associated with two of the sampled PPC's was examined to evaluate the hypothesis that some PPC's were the cause of observed relatively rapid reliability growth.

It was concluded that it was feasible to use the BSTM as an index of the efficacy of some power plant changes toward increasing the reliability of aircraft engines.

1. INTRODUCTION.

1.1. Background. The introduction of any new complex system into its operational environment can be expected to be followed by a period during which improvements, designed to increase the system's reliability, will be incorporated. An aircraft engine is such a system, for improvements continue to be made after the system has been put into production and introduced operationally to the fleet. The expanded use of the system often reveals weaknesses that were previously not apparent. As weaknesses are discovered, corrective changes are made in existing systems and incorporated into new systems. Hence, during some period following development and initial production, the system is improved and its reliability is increased.

1.2. Objectives. There are many techniques available to evaluate improvements in the reliability of a system. The Barlow-Scheuer Trinomial Model (BSTM) should be applicable to the particular case of aircraft engine development. The objective of this paper is to report the results of applying Navy operational data to the BSTM in order to evaluate the feasibility of using this technique on available data concerning other Navy operated systems. The specific

data used in this study were derived from operation of the TF-30 turbo-fan jet engine built by Pratt and Whitney. The engine is used in the A-7 Corsair II single engine attack aircraft operated by the Navy. If application of available data on Navy aircraft/engine operations to the BSTM proves technically feasible, the model can become a powerful tool in providing an objective measure of the effects of engineering efforts designed to improve the safety or reliability of the system. It is manifest that an improvement in the reliability of an aircraft engine will result in an increase in the safety of that aircraft. If, therefore, improvements in reliability as measured by the BSTM can be associated with specific events such as power plant changes (PPC), then it can be inferred that those PPC's were responsible for an increase in safety. Thus, this method promises to provide an objective measure of the efficacy of one of the Navy's accident prevention programs.

1.3. Overview of the BSTM. The BSTM requires the system development program to be divided into stages. The stages can be initially defined in terms of either time periods or a given number of trials performed on the system under development. Within each stage, however defined, the number of trials of the system is noted. Each trial must result in exactly one of three possible outcomes: 1. inherent failure, 2. assignable cause failure, or 3. success. An inherent failure is defined as any failure caused by factors that are not attributable to the system. Assignable cause failures are those for which the cause can be directly assigned to one or more components of the system. A success is any trial during which the system did not fail so that the system is as ready to operate at the end of the trial as it was at the beginning of the trial. The assumption is made that all trials within a single stage are performed on essentially similar systems.

2. ANALYSIS.

2.1. Definition of Parameters. The BSTM provides for both the grouping of the system development program into stages of development and the discrimination between two distinct causes of failure. Stages of development in this study were initially defined to include all trials performed during a given time period. The assumption also was made that during any calendar month there exists a sufficient degree of homogeneity among the aircraft engines in use in the fleet so that all the trials occurring within that month could constitute a single stage of development. A trial of an aircraft engine system is made each time the system is used. Hence,

the number of flights made or attempted with the TF-30 engine during a calendar month constitutes the total number of trials of that system in one stage. Each trial, in each stage, results in exactly one of three outcomes: inherent failure, assignable cause failure, or success. A flight is considered a success when the aircraft returns without incident, i.e., no maintenance action is required as a result of malfunction. An inherent failure is one that can be eliminated only as the state of the art of flying is advanced and human error or acts of God are eliminated. Bird strikes, weather, and pilot errors are typical examples of inherent failure. Assignable cause failures are those that can be corrected by some operational or equipment modification. The malfunction of number six bearing or a crack in the combustion chamber are examples of assignable cause failures. By the BSTM, the number of stages of development and the number of trials in the ith stage may be treated as random variables. They were so treated in this study.

2.2. Data Sources. The Standard Naval Maintenance and Material Management File (3M) was used to ascertain the occurrence and nature of TF-30 engine malfunctions. All maintenance action reports on the TF-30 engine between 1 October 1968 and 30 June 1971 were included. The resulting data are known to exclude some malfunctions because some engines were striken from the inventory following severe damage even though no maintenance action report was filed. The number of such instances is small and will not significantly affect the conclusions of this study.

The total number of trials was obtained from three different sources. Flights of the A-7A's and A-7B's, both of which are powered by the TF-30 engine, were obtained from OPNAV 50-104, "Flight Activity of Navy Aircraft" which is published monthly by CNO. A count of all A-7A/B flights that were aborted before take off was obtained from the 3M data system. Flights of the TF-30 equipped A-7E (now A-7C) were obtained by aircraft bureau number from the Individual Flight Activity Reporting System maintained by the Naval Safety Center.

2.3. Variable of Analysis. The reliability, r, or an engine is an expression of the probability of a successful operation on any given trial. (Figure 1) Let q_0 represent the probability of an inherent failure. It is assumed that the state of the art will not fluctuate significantly from month to month; therefore, q_0 is held constant through all stages. Let q_i represent the probability of an assignable cause failure in the ith stage of development. This probability is assumed to be a stagewise non-increasing variable.

From this, it is seen that the reliability of the system at the ith stage of development is as shown in equation (1) in Figure 1. When the q_i have decreasing values with respect to time, the r_i are seen to be stagewise increasing. This has been termed "reliability growth".

2.4. Analytic Technique. Assume there are n_i trails (flights) in the ith stage ($i = 1,2...,k$). It has been noted previously that each trial must result in one of three outcomes; inherent failure, assignable cause failure, or success. Thus at the end of the ith stage, the n_i trails in that stage will have resulted in a_i inherent failures, b_i assignable cause failures, and c_i successes. It is necessary that each trial results in exactly one of the three possible outcomes, as indicated by equation (2) in Figure 1. The probability, in any stage, of observing exactly a_i inherent failures, b_i assignable cause failures, and c_i successes is written as equation (3) in Figure 1 where q_0 is the probability of observing an inherent failure in any trial, and q_i is the probability of observing an assignable cause failure in the ith stage. The assumption that each of these probabilities remains constant throughout each stage is necessary for the validity of this equation.

Estimators of the probabilities of inherent failure in all stages and assignable cause failure in the ith stage were required for the computation of engine reliability. The estimating procedure is maximum likelihood. Upon differentiating the logrithm of the likelihood function with respect to q_0 and q_i setting the derivatives equal to zero, we find the maximum likelihood estimates as shown in Figure 2.

It is easily seen that $\hat{q}_0$ is the total number of inherent failures occurring in all k stages divided by the total number of trials. The $\hat{q}_i$ are the maximum likelihood estimates of the q_i's in general. Finally, the maximum likelihood estimate for system reliability at the ith stage is as shown.

2.5. Calculations. Data were collected to cover the period from 1 October 1968 through 30 June 1971. Each of the 33 months in this period was considered as a stage of development. Thus, in this study, $k = 33$. The number of inherent failures (a_i), assignable cause failures (b_i), successes (c_i), and trials (n_i) for each stage were recorded. The value of $\hat{q}_0$ was then calculated as shown in Figure 3.

Once $\hat{q}_0$ had been obtained, stagewise estimates of the probabilities of assignable cause failures followed easily. The calculated values of the $\hat{q}_i$ in the first four stages of development are shown in Figure 3.

On inspection of these four estimates, it is seen that the $\hat{q}_i$ are not sequentially non-increasing. It is, thus, apparent that the arbitrarily designated stages of development do not reflect the reliability growth of the system. It was, therefore, required by the BSTM to alter the time periods designating stages of development until the desired non-increasing values for $\hat{q}_i$ were observed.

Beginning with the first stage of development, the next sequential stage observed with an increasing $\hat{q}_i$ must be included with its immediate predecessor. In this instance, $\hat{q}_2$ is greater than $\hat{q}_1$, so stage 2 must be included with stage 1. To facilitate the discussion and for notational ease, as the analysis proceeds, the $\hat{q}_i$ will be renamed. In the second iteration in this example, the new $\hat{q}_1$ will be a combination of the above $\hat{q}_1$ and $\hat{q}_2$. Hence, we have these new values as shown in Figure 4. Again, there is an increasing sequence, and the adjustment of stages must be made. The result is as shown in Figure 5. The requirement of non-increasing values of $\hat{q}_i$ has now been met and two true stages of development have been defined. This process was followed for the complete 33 months of data.

Having defined $\hat{q}_0$ and $\hat{q}_i$ for each of nine true stages of development, and recalling the estimate of system reliability at the end of the ith stage, we see that the reliability growth of the TF-30 engine is quantified as shown in Figure 6. For each of the nine end stage reliabilities ($\hat{r}_i$), a 95% lower confidence bound has been calculated. Barlow and Scheuer describe a conservative lower confidence bound for $\hat{r}_k$ considered as the binomial parameter. In this study, since n_k is very large, normal theory was used. The calculated 95% lower confidence bound for each of the $\hat{r}_i$ is shown in Figure 6.

2.6. Interpretation. Now that the reliability growth has been quantified and the lower bounds calculated, it is desirable to relate the stages of that growth to events that were designed to cause it. Ideally one would expect every Power Plant Change to result in an associated and relatively rapid increase in the reliability of the system. In fact, this does not happen. This is not to say that those PPC's for which no associated reliability increase can be observed are not worthwhile. It is only a recognition of the fact that some PPC's produce more dramatic results than others. Between December 1968, the first month of data used in this study,

and June 1971, the last month of data used in this study, slightly more than 100 PPC's were developed for the TF-30 engine and incorporated into these engines. From these 100, six were chosen as illustrations and examples for this paper. A list and brief description of these six PPC's is shown in Figure 7.

Three dates were provided for each PPC by the Planning Division of the Naval Air Rework Facility (NARF), Norfolk. These dates are: 1) the month the PPC was first installed, labeled "Introduction", 2) the month when about 40-60 percent of the engines in the inventory had the PPC installed, labeled "Midway", and 3) the month when from 80-90 percent of the engines in the inventory had the PPC installed, labeled "Completion".

The installation of the PPC was considered essentially complete after being installed into only 80-90 percent of the engines because at that time the remaining 10-20 percent of the engines were either in transit to the NARF or out of service for some other reason. The midway point of incorporation was considered to approximate the earliest time when the effect of the PPC could be readily observed. Thus, if any PPC caused a dramatic increase in the reliability of the engine, one would expect to be able to detect this effect by the midway month of incorporation.

The reliability growth of the TF-30 engine from the end of October 1968 through June 1971 is shown graphically in Figure 8. In this graph, the horizontal axis is time in monthly increments. The vertical axis is system reliability. Note that the vertical scale does not start at zero so that the character of the reliability growth will be emphasized. In addition to the end stage reliabilities and their 95% lower confidence bounds, the six chosen PPC's are depicted on the graph. Each PPC is shown as a heavy horizontal bar. The left end of the bar concides with the month of introduction. The right end of the bar concides with the month of completion. The midway month is shown as an arrow underneath each bar. It should be noted that the end stage reliabilities are measured and graphed as of the end of the month, whereas, the three dates associated with each PPC are shown as the middle of the month. The number identifying each PPC is written on top of its bar.

Power Plant Changes 167 and 180 were both introduced in August 1968. The midway month for each was January 1969. From the graph, it is seen that there was a dramatic increase in the rate of reliability growth during the month of January 1969. It, thus,

appears that these two PPC's were at least partially responsible for this segment of reliability growth.

PPC 199 was introduced in December 1968. Its midway month was March 1969. Installation of this PPC was essentially complete by July 1969. During the period, over which the installation of this PPC could logically have been expected to have an observable effect on system reliability, there was no unusually rapid increase observed. From the graph, it is seen that the reliability of the TF-30 did improve significantly over this period, but relatively the improvement was small compared to January 1969 or the period from September through December 1969. Thus, it cannot be said that PPC 199 was the principal cause of any relatively rapid improvement in reliability. This does not in itself mean that PPC 199 was of no value. It merely means that any improvements realized as a result of installing PPC 199 were not large enough, in relation to the effects of other changes that were occurring during the same period, to result in relatively rapid reliability growth.

PPC 212 was introduced in March 1969. Its midway month was September 1969. October through December 1969 was a period over which relatively rapid reliability growth occurred. It is reasonable to assume that PPC 212 made some substantial contribution to this growth.

Of the four PPC's just discussed, it was assumed that the coincidence of the midway months with a period of relatively rapid reliability growth represented a causative relation. In order to verify this assumption, it would be necessary to examine failure records before, during, and after the installation period of the PPC. Specifically, the number of failures of all components associated with the PPC would have to be determined and compared among the three time periods. It would also have to be verified that any differences were not the result of different levels of utilization (trials).

Fortunately, most of the data needed for such a comparison were available from the computerized data banks at the Naval Safety Center. In order to retrieve them, however, it was first necessary to identify all components that are associated with the PPC under investigation. This in itself is an arduous task requiring the effort of engineers experienced with the TF-30 engine. Unfortunately, the time required to identify all components associated with each of the six PPC's was not available. However, it was possible to identify

one component associated with PPC's 220 and 267. That component is the tower shaft bearing.

From Figure 8, it appears that PPC 220 had an observable effect on reliability growth early in its history but that effect ended shortly after the midway month. The graph also indicates that PPC 267 had a very marked effect on reliability growth.

Figure 9 shows the number of failures of the tower shaft bearing for each month from October 1968 through June 1971. Also shown are the introduction, midway, and completion months of PPC 220 and the introduction and midway months of PPC 267. The data on this figure tend to verify the conclusions drawn from the previous figure. From these data, it is seen that the tower shaft bearing associated with PPC 220 showed no failures prior to the introduction of the change. By the midway month, only one failure was observed. By the completion of the change, 17 failures had occurred. If PPC 267 had not been introduced, there undoubtedly would have been many more failures of the new tower shaft roller bearing. However, by the midway month of PPC 267, the tower shaft bearing failures had been effectively stopped resulting in a period of rapid and significant reliability growth.

Thus when it is realized that components other than the tower shaft bearing were affected by PPC 220, it is seen that this PPC contributed substantially to the reliability growth of the TF-30 during the months of November and December 1969. Following that, the problems created by PPC 220 overcame the problems solved by the change and the reliability growth was substantially slowed. The introduction of PPC 267 to correct the problem created by PPC 220 allowed the benefits of PPC 220 to again be seen.

3. CONCLUSION.

This report clearly indicates the feasibility of using the Barlow-Scheuer Trinomial Model for evaluating the reliability of the TF-30 engine. The growth of reliability was observed and the degree of reliability was stated. Further, three specific stages of development during which reliability increased dramatically were discovered. Each of those three periods was associated with a specific PPC and a rationale supporting a causal relationship was developed.

It is reasonable to assume that if this method is applicable to the TF-30 system, it could be similarly applied to other systems

such as airframes, engines, tanks, or individual components. The only limitation would be the availability of data concerning a given system. The feasibility of using the BSTM as an index of the efficacy of some engineering efforts in accident prevention has been demonstrated. In effect, this has made available another tool to demonstrate objectively and unequivocally the effects of some of these efforts on Naval aviation.

FIGURE 1

$$(1) \quad r_i = 1 - q_0 - q_i \qquad (i = 1,2,\ldots, k)$$

$$(2) \quad a_i + b_i + c_i = n_i$$

$$(3) \quad P(a_i, b_i, c_i) = \frac{n_i!}{a_i!\, b_i!\, c_i!}\, q_0^{a_i}\, q_i^{b_i}\, (1 - q_0 - q_i)^{c_i}$$

FIGURE 2

$$\left.\begin{aligned} \hat{q}_0 &= \frac{\sum_{i=1}^{k} a_i}{\sum_{i=1}^{k} n_i} \\ \hat{q}_i &= \frac{(1 - \hat{q}_0)\, b_i}{b_i + c_i} \\ \hat{r}_i &= 1 - \hat{q}_0 - \hat{q}_i \end{aligned}\right\} \qquad i = 1,2,\ldots, k$$

FIGURE 3

$\hat{q}_0 = 1.629603 \times 10^{-3}$

$\hat{q}_1 = 2.533787 \times 10^{-2}$

$\hat{q}_2 = 2.890294 \times 10^{-2}$

$\hat{q}_3 = 3.103555 \times 10^{-2}$

$\hat{q}_4 = 2.578506 \times 10^{-2}$

FIGURE 4

$\hat{q}_0 = 1.629603 \times 10^{-3}$

$\hat{q}_1 = 2.696155 \times 10^{-2}$

$\hat{q}_2 = 3.103555 \times 10^{-2}$

$\hat{q}_3 = 2.578506 \times 10^{-2}$

FIGURE 5

$\hat{q}_0 = 1.629603 \times 10^{-3}$

$\hat{q}_1 = 2.820535 \times 10^{-2}$

$\hat{q}_2 = 2.578506 \times 10^{-2}$

FIGURE 6

Stage	Calendar Period	$\hat{q}_i$	$\hat{r}_i$	$\hat{r}_i$(LB)
1	Oct-Dec 1968	2.820535×10^{-2}	.970165	.968088
2	Jan 1969	2.578506×10^{-2}	.972585	.970855
3	Feb-Sep 1969	2.431031×10^{-2}	.974060	.973128
4	Oct 1969	2.329376×10^{-2}	.975077	.974198
5	Nov-Dec 1969	2.155256×10^{-2}	.976818	.976028
6	Jan-Aug 1970	2.043030×10^{-2}	.977940	.977308
7	Sep 1970	1.576172×10^{-2}	.982609	.982057
8	Oct-Dec 1970	1.272812×10^{-2}	.985642	.985163
9	Jan-Jun 1971	1.221275×10^{-2}	.986158	.985730

FIGURE 7

SELECTED POWER PLANT CHANGES

PPC

167	Introduction	August 1968
	Midway	January 1969
	Completion	August 1969

Purpose: Reduce wear in the vane slots of the shrouds by increasing the bearing area of the vanes.

180	Introduction	August 1968
	Midway	January 1969
	Completion	June 1969

Purpose: Improve durability of the bleed air duct assembly by incorporating a baffle configuration to funnel cooling air over the duct diaphragm.

199	Introduction	December 1968
	Midway	March 1969
	Completion	July 1969

Purpose: Provide more durable #6 bearing support.

212	Introduction	March 1969
	Midway	September 1969
	Completion	April 1970

Purpose: Provide a new fuel nozzle nut assembly and a new, more durable combustion chamber assembly featuring a revised hole pattern and a deflector baffle. This new configuration is designed to reduce exhaust smoke density.

220	Introduction	June 1969
	Midway	December 1969
	Completion	July 1970

Purpose: Provide a more durable bearing configuration in the tower shaft assembly by replacing the ball bearing assembly with a roller bearing assembly.

267	Introduction	March 1970
	Midway	October 1970
	Completion	March 1971

Purpose: To prvent the roller bearing outer race rotation thereby eliminating bearing housing wear and possible misalignment of the bearing inner race on the shaft.

FIGURE 8

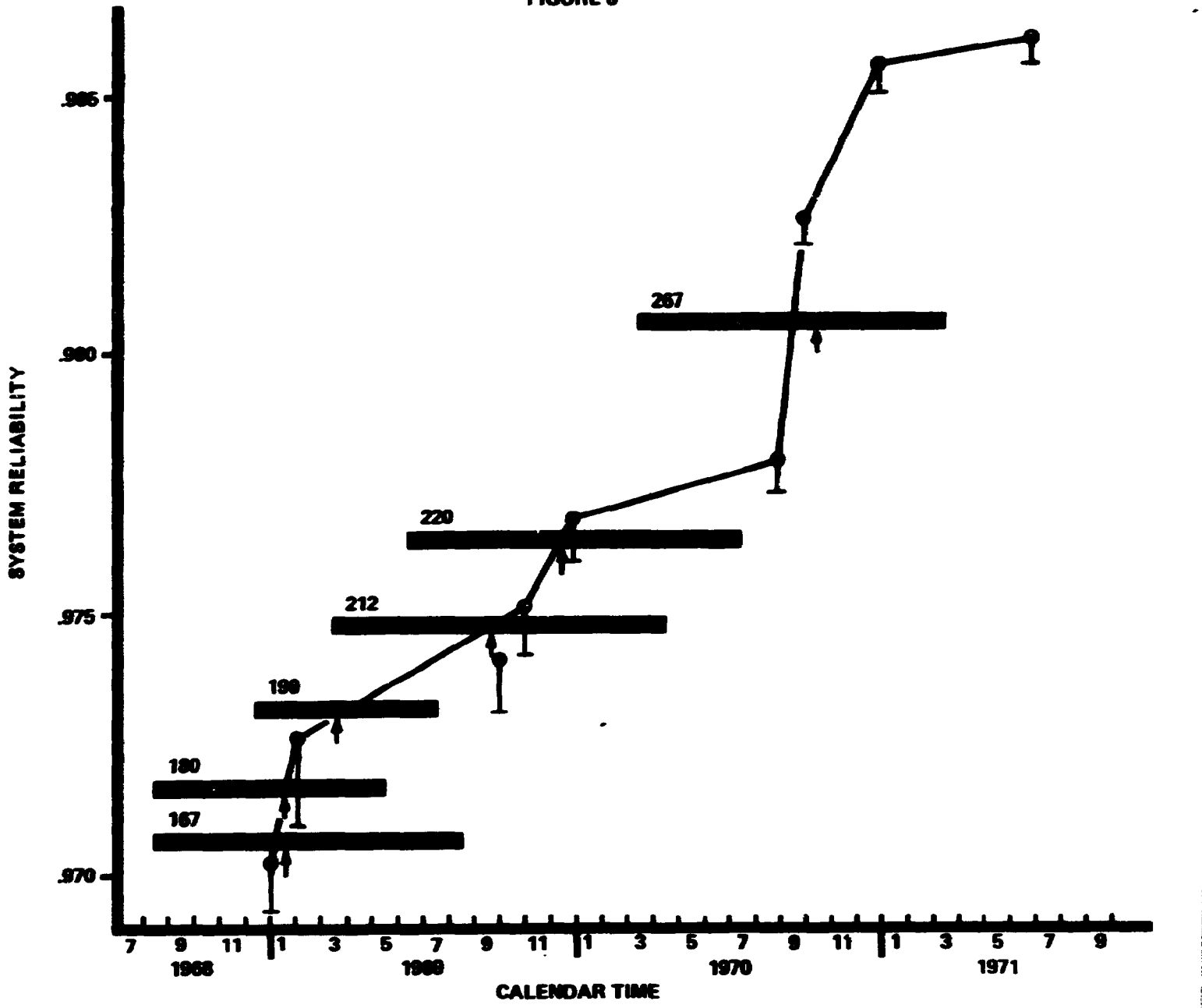

.985
.980
.975
.970
SYSTEM RELIABILITY
267
220
212
198
180
167
7 9 11 1 3 5 7 9 11 1 3 5 7 9 11 1 3 5 7 9
1968
1969
1970
1971
CALENDAR TIME

FIGURE 9

INCIDENCE OF TOWER SHAFT BEARING FAILURE ASSOCIATED WITH POWER PLANT CHANGES 220 AND 267

MONTH	TRIALS	NO. OF FAILURES	REMARKS
OCT 68	6,867	0	
NOV	5,745	0	
DEC	5,544	0	
JAN 69	5,939	0	
FEB	6,196	0	
MAR	6,973	0	
APR	8,855	0	
MAY	6,754	0	
JUN	6,696	0	PPC 220 Introduced
JUL	6,868	0	
AUG	6,642	1	
SEP	5,643	0	
OCT	6,442	0	
NOV	6,766	0	PPC 220 Midway
DEC 69	6,176	0	
JAN 70	5,797	3	
FEB	7,203	1	
MAR	7,743	6	PPC 267 Introduced
APR	7,499	4	
MAY	6,656	2	
JUN	4,847	0	PPC 220 Complete
JUL	4,313	0	
AUG	4,194	1	
SEP	5,204	0	
OCT	5,141	0	PPC 267 Midway
NOV	5,299	0	
DEC	4,198	0	
JAN 71	5,437	0	
FEB	5,588	0	
MAR	6,544	0	
APR	5,770	0	
MAY	5,344	0	
JUN 71	5,793	0	

RELIABILITY GROWTH ESTIMATION FROM FAILURE AND TIME TRUNCATED TESTING

Larry H. Crow
Reliability, Availability and Maintainability Division
U. S. Army Materiel Systems Analysis Agency
Aberdeen Proving Ground, Maryland

1. INTRODUCTION. In 1962, J. T. Duane of General Electric Company's Motor and Generator Department [see Duane (5)] published a report in which he presents his observations on failure data for five divergent types of systems during their development programs at G. E. These systems included complex hydromechanical devices, complex types of aircraft generators and an aircraft jet engine. The study of the failure data was conducted in an effort to determine if any systematic changes in reliability occurred during the development programs for these systems. His analysis revealed that for these systems, the observed cumulative failure rate versus cumulative operating hours fell close to a straight line when plotted on log-log paper. Similar plots have been noted in industry for other types of systems, and by the U. S. Army for various military weapon systems during development [see Crow (3)].

From a mathematical interpretation of these straight line plots on log-log papers, we will show that the reliability growth of these systems during development was governed by a certain Weibull process. From this result a Weibull reliability growth model is formulated which can be used with test data for monitoring and projecting system reliability during a development testing program. This paper gives appropriate estimation procedure for this tracking of system reliability for various combinations of time and failure truncated data.

2. THE WEIBULL RELIABILITY GROWTH MODEL. Let $N(t)$ denote the number of system failures by time t, $t > 0$. The observed cumulative failure rate $C(t)$ at time t is, therefore, equal to $C(t) = N(t)/t$. The plots on log-log paper imply that $\log C(t)$ is approximately a straight line. That is, $\log C(t) \doteq \delta + \gamma \log t$. Equating $C(t)$ to its expected value and assuming an exact linear relationship, we have $\log (E[C(t)]) = \delta + \gamma \log t$. Taking exponentials gives $E[C(t)] = \lambda t^{\gamma}$, $\lambda = e^{\delta}$. Hence, $E[N(t)] = \lambda t^{\beta}$, for $\beta = \gamma + 1$, since $E[C(t)] = E[N(t)]/t$. Thus, the expected number of system failures by time t is λt^{β}.

The instantaneous failure rate, $r(t)$, of the system is the change per unit time of $E[N(t)]$. Thus, $r(t) = \frac{d}{dt} E[N(t)] = \lambda \beta t^{\beta-1}$, which is recognized as being the Weibull failure rate function. It is important to note that since the system configuration is changing, the data are not homogeneous and, therefore, the usual theory for a Weibull distribution

will not apply. In fact, it has been shown by the author [see Crow (2)] that when the configuraiton of the system is changing during a development program and failures are governed by the failure rate $r(t) = \lambda\beta t^{\beta-1}$, then the system failure times follow a nonhomogeneous Poisson process with Weibull intensity function r(t).

At time t_o the Weibull failure rate is $r(t_o) = \lambda\beta t_o^{\beta-1}$. If no further system improvements are made after time t_o, then it is reasonable to assume that the failure rate would remain constant at the value $r(t_o)$ if testing were continued. In particular, if the system were put into production with the configuration fixed as it was at time t_o, then the life distribution of the systems produced would be exponential with mean time between failure (MTBF) $M(t_o) = [r(t_o)]^{-1} = t_o^{1-\beta}/\lambda\beta$. Hence, for $0 < \beta < 1$, the MTBF M(t) increases as the development testing time t increases, and is proportional to $t^{1-\beta}$. Thus, β is a growth parameter reflecting the rate at which reliability, or MTBF, increases with development testing time.

If the successive times of failures are being recorded for a system undergoing development testing, then a statistical goodness of fit test developed by the author in (2) can be performed to determine if the Weibull reliability growth model may be used to track system reliability during development testing. To track this reliability improvement would, of course, require estimating from test data the two unknown parameters λ and β by say $\hat{\lambda}$, $\hat{\beta}$. One would then estimate the failure rate function by $\hat{r}(t) = \hat{\lambda}\hat{\beta}t^{\hat{\beta}-1}$ and the MTBF function by $\hat{M}(t) = [\hat{r}(t)]^{-1} = t^{1-\hat{\beta}}/\hat{\lambda}\hat{\beta}$. If the system were tested to time T, say, then $\hat{M}(T)$ would estimate the current MTBF and $\hat{M}(t)$, $t > T$ would project estimates of system MTBF into the future.

3. ESTIMATION PROCEDURES. Suppose that a system has experienced N failures during development testing. Let X_i be the age (time on test) of the system at the i-th failure, i=1,...,N. If testing were stopped at the N-th failure, the data are said to be failure truncated.

$$0 \qquad X_1 \qquad X_2 \qquad X_3 \qquad \cdots \qquad X_N$$

The maximum likelihood (ML) estimate of β, the growth parameter, is

$$\hat{\beta} = \frac{N}{\sum_{i=1}^{N-1} \log \frac{X_N}{X_i}}, \tag{1}$$

and the ML estimate of λ is

$$\hat{\lambda} = \frac{N}{X_N^{\hat{\beta}}}. \qquad (2)$$

Thus, calculating $\hat{\lambda}$, $\hat{\beta}$ one may estimate the failure rate function $r(t) = \lambda\beta t^{\beta-1}$ by $\hat{r}(t) = \hat{\lambda}\hat{\beta}t^{\hat{\beta}-1}$. The MTBF function $M(t) = [r(t)]^{-1}$ is similarly estimated by $\hat{M}(t) = [\hat{r}(t)]^{-1}$. In particular the current estimate of the MTBF is $\hat{M}(X_N) = X_N/N\hat{\beta}$, and $\hat{M}(t)$, $t > X_N$, projects expected future growth of system MTBF.

Example

Suppose that a system undergoing development testing recorded the following 40 successive failure times; .7, 3.7, 13.2, 17.6, 54.5, 99.2, 112.2, 120.9, 151.0, 163.0, 174.5, 191.6, 282.8, 355.2, 486.3, 490.5, 513.3, 558.4, 678.1, 688.0, 785.9, 887.0, 1010.7, 1029.1, 1034.4, 1136.1, 1178.9, 1259.7, 1297.9, 1419.7, 1571.7, 1629.6, 1702.3, 1928.9, 2072.3, 2525.2, 2928.5, 3016.4, 3181.0, 3256.3. That is, the system was of age .7 when the first failure occurred, of age 3.7 when the second failure occurred, etc. At age 3256.3 the system had the 40-th failure. From these data, and equations (1) and (2) we find that $\hat{\lambda} = 0.761$, $\hat{\beta} = 0.490$.

Using $\hat{\lambda}$, $\hat{\beta}$, the failure rate function is estimated by $\hat{r}(t) = \hat{\lambda}\hat{\beta}t^{\hat{\beta}-1}$ and the MTBF function is estimated by $\hat{M}(t) = [\hat{r}(t)]^{-1}$. The current failure rate $r(3256.3)$ is estimated to be $\hat{r}(3256.3) = 0.006$, and the estimate of current MTBF is $[.006]^{-1} = 166.7$. Further, $\hat{r}(4000) = 0.005$ and $\hat{M}(4000) = [.005]^{-1} = 200.0$ are projections of failure rate and MTBF if development testing were continued to $T = 4000$.

Suppose that $K \geq 1$ systems have been simultaneously tested to time T, where T is not a failure time. In this case, the data are said to be time truncated. If design and engineering modifications are made on all K systems at the same time, then at any time during the testing the systems will have basically the same configuration. In this situation we may combine the failure data on these K systems to obtain estimates of λ and β.

$$0 \quad X_{11} \quad X_{21} \quad X_{31} \quad X_{41} \quad . \quad . \quad . \quad X_{N_1,1} \quad T$$

$$0 \quad X_{12} \quad X_{22} \quad X_{32} \quad X_{42} \quad X_{52} \quad . \quad . \quad . \quad X_{N_2,2} \quad T$$

$$0 \quad X_{1K} \quad X_{2K} \quad X_{3K} \quad . \quad . \quad . \quad X_{N_K,K} \quad T$$

Let N_r be the random number of failures observed for the r-th system, r=1,...,K. Let X_{ir} be the age of the r-th system at the i-th failure. The ML estimate of β is

$$\hat{\beta} = \frac{N}{\sum_{r=1}^{K} \sum_{i=1}^{N_r} \log \frac{T}{X_{ir}}} \tag{3}$$

where $N = \sum_{r=1}^{K} N_r$ is the total number of failures experienced by the K systems. The ML estimate of λ is

$$\hat{\lambda} = \frac{N}{KT^{\hat{\beta}}}. \tag{4}$$

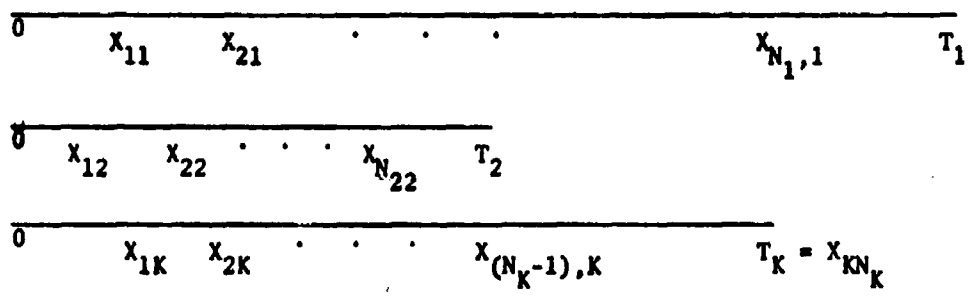
Using these estimates of λ and β, the ML estimates of the failure rate function and MTBF function are $\hat{r}(t) = \hat{\lambda}\hat{\beta}t^{\hat{\beta}-1}$ and $\hat{M}(t) = [\hat{r}(t)]^{-1}$, $t > 0$, respectively. The ML estimate of current system MTBF is $\hat{M}(T)$, and $\hat{M}(t)$, $t > T$, projects expected future growth of system MTBF with increased testing time. Note that the total test time for these K systems is KT. However, the calendar time or system configuration age is T.

Example

Suppose K = 3 systems were each tested for time T = 200 with the successive failure times given in Table 1. From these data, and using equations (3) and (4), the ML estimate of β is $\hat{\beta} = 0.615$ and the ML estimates of λ is $\hat{\lambda} = 0.461$. From $\hat{\lambda}$, $\hat{\beta}$ we estimate M(200), the current MTBF, by $\hat{M}(200) = 27.12$.

If development testing were stopped at say T = 300, the model states that future times between failures will follow the exponential distribution with mean M(300). Based on test data to T = 200 the projection of the MTBF at T = 300 is $\hat{M}(300) = 31.70$. That is, with 100 more units of test time we estimate an increased in MTBF from 27.12 to 31.70.

Again, suppose that $K \geq 1$ systems begin development testing at the same time (time 0), but the test times for each of the systems are not necessarily the same. If T_r, the test time for the r-th system, is also a failure time, then the data on this system are said to be failure truncated. Otherwise the data are said to be time truncated.



In the above illustration, the data on systems 1 and 2 are time truncated while the data on system K are failure truncated. For this general testing situation the ML estimates of λ and β are not in closed mathematical form as they were for the two special cases previously discussed. In this situation the ML estimates of λ and β are values $\hat{\lambda}$ and $\hat{\beta}$ satisfying the equations

$$\hat{\lambda} = \frac{N}{\sum_{r=1}^{K} T_r^{\hat{\beta}}} \qquad (5)$$

$$\hat{\beta} = \frac{N}{\hat{\lambda}\sum_{r=1}^{K} T_r^{\hat{\beta}} \log T_r - \sum_{r=1}^{K}\sum_{i=1}^{N_r} \log X_{ir}} \qquad (6)$$

In general, these equations cannot be solved explicity for $\hat{\lambda}$ and $\hat{\beta}$, but must be solved by iterative procedures. As before, the ML estimates of the failure rate and MTBF function are $\hat{r}(t) = \hat{\lambda}\hat{\beta}t^{\hat{\beta}-1}$ and $\hat{M}(t) = [\hat{r}(t)]^{-1}$, $t > 0$, respectively.

A computer program has been written at AMSAA for calculating $\hat{\lambda}$ and $\hat{\beta}$ from the general expressions (5) and (6). This computer program is reported in Belbot (1).

Example

Consider again the data in Table 1. Suppose that it is assumed a priori that systems 1, 2 and 3 are tested only to the 10th, 15th and 11th failure respectively. In this case the data are failure truncated on the three systems. Thus, the ML estimates of λ and β are calculated from expression (5) and (6) using iterative procedures. From these data $T_1 = 197.2$, $T_2 = 190.8$, $T_3 = 195.8$, and the ML estimates of λ and β are $\hat{\lambda} = 0.443$, $\hat{\beta} = 0.626$.

Even though the ML estimates of λ and β are not in closed mathematical form for this general testing situation, it can be shown that an unbiased estimate of β can be written in closed form. To calculate this estimate let

$$M_r = \begin{cases} N_r & \text{if data on the r-th system are time truncated} \\ N_r - 1 & \text{if data on the r-th system are failure truncated} \end{cases}$$

$r=1,\ldots,K$. Then an unbiased estimate of β is

$$\bar{\beta} = \frac{M-1}{\sum_{r=1}^{K}\sum_{i=1}^{M_r} \log \frac{T_r}{X_{ir}}}$$

where $M = \sum_{r=1}^{K} M_r$.

Other statistical results and actual Army applications of the Weibull reliability growth model may be found in Crow (2), (3), (4).

REFERENCES

1. Belbot, Edward F., A Computer Program For Estimation of Parameters of the Weibull System Failure Rate Function, AMSAA Technical Report (In Preparation), 1974.

2. Crow, L. H., Reliability Analysis For Complex, Repairable Systems, AMSAA Reliability, Availability and Maintainability Division Interim Note. No. 26, January 1974.

3. Crow, L. H., On Reliability Growth Modeling, AMSAA Reliability, Availability and Maintainability Division Interim Note No. 27, January 1974.

4. Crow, L. H., Tracking Reliability Growth, AMSAA Reliability, Availability and Maintainability Division Interim Note No. R-30, April 1974.

5. Duane, J. T., Learning Curve Approach to Reliability Monitoring, IEEE Trans., Aerospace, Vol. 2, 1964.

TABLE 1

Simulated Data for K=3 Systems Operated for Time T=200 when $\lambda = 0.6$ and $\beta = 0.5$

System 1	System 2	System 3
X_{i1}	X_{i2}	X_{i3}
4.3	0.1	8.4
4.4	5.6	32.5
10.2	18.6	44.7
23.5	19.5	48.4
23.8	24.2	50.6
26.4	26.7	73.6
74.0	45.1	98.7
77.1	45.8	112.2
92.1	75.7	129.8
197.2	79.7	136.0
	98.6	195.8
	120.1	
	161.8	
	180.6	
	190.8	
$N_1 = 10$	$N_2 = 15$	$N_3 = 11$
$\sum_{i=1}^{10} \log\left(\frac{T}{X_{i1}}\right)$	$\sum_{i=1}^{15} \log\left(\frac{T}{X_{i2}}\right)$	$\sum_{i=1}^{11} \log\left(\frac{T}{X_{i3}}\right)$
=19.661	=26.434	=12.398

$$N = N_1 + N_2 + N_3 = 36$$

$$\sum_{r=1}^{3} \sum_{i=1}^{N_r} \log\left(\frac{T}{X_{ir}}\right) = 58.493$$

$$\hat{\beta} = \frac{N}{\sum_{r=1}^{3} \sum_{i=1}^{Nr} \log\left(\frac{T}{X_{ir}}\right)} = 0.615 \qquad \bar{\beta} = \frac{N-1}{N} \hat{\beta} = 0.598$$

$$\hat{\lambda} = \frac{N}{KT^{\hat{\beta}}} = 0.461$$

RESPONSIVENESS PROPERTIES OF CONTINUOUS SAMPLING PLANS

RICHARD M. BRUGGER

US ARMY ARMAMENT COMMAND
PRODUCT ASSURANCE DIRECTORATE
RAM ASSESSMENT DIVISION

ABSTRACT:

The usual methods of describing the functional properties of a continuous sampling plan (such as a display of Average Fraction Inspected (AFI) and Average Outgoing Quality (AOQ) curves) are founded on the assumption that the probability that a unit is defective is constant over all units in a continuing flow of manufactured product. While the use of this assumption greatly simplifies much of the mathematics involved in describing the properties of the plans, it is, at the same time, not a realistic assumption. In designing a continuous sampling plan, the statistician is concerned with, among other things, the plan's responsiveness, that is, how the plan reacts when the probability that a unit is defective suddenly increases. This paper describes some Markov chain methods that have been developed to describe responsiveness properties of plans complicated in structure.

Various terms will be used in this paper, and will be defined at appropriate times.

VUGRAPH 1

Cases occur in practice where it is not practical to group units of material into lots prior to sampling for inspection purposes; for example, product may be manufactured and moved in a continuous flow for efficiency or safety reasons. As an alternative, we have continuous sampling, introduced by Harold Dodge in a paper in the Annals of Mathematical Statistics in 1943.

VUGRAPH 2

This is Dodge's first continuous sampling plan, called CSP-1. At the start of inspection, a screening crew inspects 100% of the units. When a pre-specified number, i, of consecutive units are found to be free of the defects concerned, that is, the kinds of defects being inspected for, the screening crew is released from 100% inspection, and the sampling inspector inspects a pre-specified fraction, f, of the units, where the units sampled are selected in a random manner. If, during sampling, the inspector finds one of the defects concerned, the screening crew returns to 100% inspection, and from that point on, as you can see, the plan consists of alternating periods of 100% inspection and sampling inspection. The clearance number, i, might be some value like 100, and the sampling frequency, f, might be 1/100.

VUGRAPH 3

Some of the functional properties of a continuous sampling plan that are of interest are the Average Fraction Inspected, or AFI, which is the expected value of the fraction of material that will be inspected over an indefinitely long period of time when each unit has probability p of being defective: The Average Outgoing Quality or AOQ, which is the expected fraction of material that is defective in accepted material over an indefinitely long period of time when each unit has probability p of being defective, and the Average Outgoing Quality Limit, or AOQL, which is the maximum value of the AOQ that can be obtained for a particular sampling plan.

VUGRAPH 4

Here is an AOQ curve. The abscissa represents the percent defective of product submitted to the point of inspection. The ordinate represents the average outgoing quality. Starting at the origin, you can see that as the quality of submitted product becomes poorer, so does the quality of outgoing material, until it reaches some maximum, the AOQL, after which the quality of outgoing material becomes better as the quality of incoming material becomes poorer. There is really no paradox here, since poorer quality product will cause the 100% inspection phase of the plan to be in effect more and more, and at some point this will be reflected by the AOQ curve bending downward.

Remember that the AOQ function is defined for an infinitely long sequence of units for which the probability, p, that a unit is defective, remains unchanged for all units ------forever.

By using a direct algebraic modeling approach, Dodge derived formulas for AFI and AOQ for CSP-1.

VUGRAPH 5

In 1951, a paper by Dodge and Torrey appeared in Industrial Quality Control, and introduced two more continuous sampling plans, designated CSP-2 and CSP-3, portrayed here is CSP-3. As you can see, the plan is a little more complicated that CSP-1. Instead of returning immediately to 100% inspection upon finding a defect during sampling inspection, the sampling inspector inspects the next four units. If a unit is defective, 100% inspection is resumed. Otherwise the sampling inspector resumes sampling, but keeps count of the number of sample units inspected.

If a defect is found during sampling, 100% inspection is resumed. If k consecutive sample units are free of the defects concerned, sampling is continued without counting.

The only reason I am showing this vugraph is to touch on some points regarding the construction of the plan. Remember that CSP-1 was describable as a unique continuing cycle of events, that is, the plan alternated between 100% inspection and sampling inspection. CSP-3, on the other hand, does not at first look like it could be described by a unique cycle, since some boxes have two outputs, which as the arrows indicate, go to different places. On the other hand, if we inclose all boxes except 100% inspection in one big black box, we see that once again we have a unique cycle of events, that is, we alternate between 100% inspection and something else. Mathematically, then, our only problem is to define what goes on inside the big black box, if we can. This, in effect, was the approach taken by Dodge and Torrey to derive the mathematical properties of CSP-3.

VUGRAPH 6

In short, then, Dodge's method is an algebraic method describing the properties of plans, and can be used only when the properties of each black box can be determined by algebraic means.

The limitations of this first method for complicated black boxes were overcome with the application of Markov chain theory to continuous sampling theory. This was begun by Lieberman and Solomon, and described in their paper appearing in 1955 in the Annals of Mathematical Statistics.

After using the Markov chain approach for some time, a simplifying algorithm suggested itself to us, and we developed a proof of its validity. The same algorithm was apparently developed about the same time by Sackrowitz, who mentioned it and used it in a paper in Technometrics in 1972. Since the Sackrowitz paper does not show the derivation of the algorithm, our technical paper on the subject is provided as a handout.

VUGRAPH 7

The effect of the simplified Markov chain approach is shown in this illustration. Under former Markov chain methods, each unit of product, or, in some cases, each inspected unit of product, would be indexed, with the index corresponding to one of a finite number of states of a Markov chain. We would then be interested in such things as what proportion of units reach the inspector during a 100% inspection phase, or what proportion of units reach the inspector during a sampling phase.

In the simplified approach, we temporarily forget the units themselves, and pretend that these units are contained in boxes as they reach the inspector. Each box is labeled with the name of a phase of inspection, such as 100% or sampling. The labels on the box become indices that we relate to the states of a Markov chain. Immediately, we see that the number of indices, and hence, the complexity of the Markov chain, is greatly reduced. We now find what proportion of boxes reach the inspector labeled 100%, what proportion are labeled sampling, and so on. After determining these proportions, we can easily determine proportions of units rather than boxes involved in the various phases of inspection by finding the expected number of units in each kind of box and carrying out appropriate algebraic operations. By introducing the units only at the end of the mathematical gyrations, however, all of the mathematics along the way will be simpler. We will illustrate this with an example.

VUGRAPH 8

Recall the CSP-3 plan which we just discussed. Here is the transitional probability matrix for the Markov chain describing the plan. The labels for the rows indicate the states we are leaving, the labels for the columns indicate the states to which we are going. The label 100 denotes the 100% inspection phase, next unlimited sampling, checking, and limited sampling. The simplicity of the matrix yields extremely simple equations to work with, as you can see. The symbol P_{100} denotes the probability of a box labeled 100% inspection reaching the inspector at some specified time indefinitely long after the start of inspection, or equivalently, the proportion of an indefinitely long stream of boxes that are labeled 100%. The other P's are similarly defined. We have further abbreviated the words unlimited sampling, checking, and limited sampling, as you can see.

In the lower right of the chart, we have solved each of the unknowns in terms of some arbitrary one of them, in this case, P_{us}. Our interest now focuses on the coefficients of P_{us}.

VUGRAPH 9

We have carried these coefficients over into column one of this chart, which is our working table for finding phase occupancy probabilities. Column two is for eliminating any factors that may be common to all expressions in column one, or for clearing denominators if there are any. Any simplifying operation carried out for one expression must simultaneously be carried out for each of the other expressions. In this case, there is nothing in column one that can be simplified, so column two is the same as column one. Column three provides the expression for the expected number of units for each of the phases. Column four is for simplifying column three. In this case, we have cleared all denominators by multiplying each expression by fpq^i. Column five serves as our device for getting the units out of the boxes, so to speak, and is the product of columns two and four.

VUGRAPH 10

We can now use the expressions in column five to determine several properties of the plan. For example, by letting the sum of the terms in column five be denoted by D, we obtain the values for $P_a(U)$, where the U signifies that we now have probabilities expressed in proportions of units.

VUGRAPH 11

Let us now return to the subject of our paper -- responsiveness, by which we mean the speed, following a deterioration in product quality, with which a 100% inspection phase will be initiated, thereby permitting the discovery of a need for corrective action.

For example, let's consider the situation shown on the chart. Units of product are manufactured and placed on a conveyor. The sign on the wall indicates past quality history was great, so we are very likely in a sampling phase of the plan we are using.

Suddenly we have trouble with our manufacturing process. Defective units are being manufactured at a higher rate than usual and are being fed into our stream of product. While we are sampling at some frequency we will probably not catch all of the defective product. We feel that if 100% inspection were resumed, we would discover the problem and take corrective action. How long will it take before 100% inspection is resumed?

VUGRAPH 12

We became interested in this problem partly because of suspicions raised by Dodge about the plans contained in H106, which are the same as the plans called CSP-M in MIL-STD-1235.

This is an example of a three level plan. MIL-STD-1235 and H-106 permit the use of CSP-M plans, subject to certain conditions, ranging from one to five levels. A one level plan would be the same as a CSP-3 plan.

In a review presented in Technometrics in 1960, Dodge expressed an intuitive concern about the behavior of a CSP-M plan if quality should suddenly deteriorate when the plan is, say, in the fifth sampling level of a five level plan. Dodge feared that 100% inspection might be resumed only after many units had been accepted.

Work in the area of CSP responsiveness had previously been done by Guthrie and Johns and described in a Stanford University Technical Report on the subject in 1958, and by Hiller in a Technometrics paper in 1964.

Their methods are readily adaptable to plans which are relatively simple in construction, but do not yield practical computational algorithms for plans as complicated as CSP-M. Let's examine briefly the work of Guthrie and Johns, since, conceptually, our work bears some resemblance.

VUGRAPH 13

This is the type of plan to which Guthrie and Johns applied their method. It is known as MLP-T, which means, "Multi-Level Plan - Tightened", and was formulated by Derman, Littauer and Solomon and described in a 1957 paper in the Annals of Mathematical Statistics. There are variations of this plan; this is one example.

If, anytime during the operation of the plan, a defective unit is found, 100% inspection is initiated and continued until i consecutive good units are found. The sampling frequency is reduced after i consecutive units are defect free. In this example, the frequency is reduced from f to f/2 to f/4.

The Guthrie and Johns method involved finding the expected number of units produced following deterioration until 100% inspection could be initiated.

VUGRAPH 14

This involved, for each step in the plan, for example, the fifth sampled unit of the second sampling phase, finding the probability of being at that step at some specific time indefinitely long after the start of production, and multiplying this by the expected number of units passed from this point until a defective unit is found. Summing over all steps in the plan yields the expected number of units produced following deterioration until 100% inspection is resumed. Here p_0 denotes the probability before deterioration that a given unit is defective, while p_1 denotes the probability after deterioration that a given unit is defective. Note that for any value of p_1, E(N) would be maximum over p_0 when $p_0 = 0$. This is because the probability becomes one that we are on a step in the most liberal sampling level. This was the assumption that we made in our method.

VUGRAPH 12 (REPEAT)

The real problem now presented itself. How do we find the expected number of units to return to 100% inspection from the most liberal sampling level? For a one level plan, simple algebraic methods could be used. For a two level plan, somewhat more complicated, but still algebraic methods could be used. For three levels and higher, algebraic methods could no longer be used, since the infinite number of ways to return to 100% inspection could not be nicely ordered so that something like converging sums could be obtained. This was the problem. The solution eventually suggested itself.

VUGRAPH 15

This chart shows the restructuring of our CSP-M flow diagram to permit solution of the problem. We said that we wanted to know the expected number of units produced from time of deterioration, while we are at the most liberal sampling phase of the plan, until time of return to 100% inspection. An expected number would correspond to an average over an infinite number of trials, so we thought of using a Markov chain approach, which in effect would consider an infinite number of trials. We would be interested in the expected number of units from entrance to the most liberal sampling phase until entrance to the 100% inspection phase. It is sufficient to consider only entrance to the most liberal sampling phase, since, given that we are in the phase, at some point, regardless of how long we have been in the phase, the expected number of units from that point to return to 100% inspection is the same as from entrance to the phase.

Solution of the Markov chain will yield proportions of time spent in the different phases, but it won't yet give us our expected number of units to return to 100% inspection, so we still need something else. By including the 100% inspection phase, which will occur once and only once per cycle, we can get what else we need.

It is for this reason that we have shown the arrow going from the 100% inspection phase to the most liberal sampling phase. The first sampling level has disappeared, because for all plans except a one level plan, it is not a part of the path from the most liberal sampling level to a resumption of 100% inspection.

VUGRAPH 16

The mathematical model of our flow diagram model can now be described.

Upon solving the Markov chain, let P_{100} denote the expression for the proportion of the time the plan is in the 100% inspection phase. Then, by reason of the way our flow diagram was constructed, the rest of the time we were on our return path. We have an expression for finding the expected length of the 100% inspection phase.

The only unknown is E(N), where N denotes the expected number of units to return to 100% inspection after first entrance (per cycle) in the most liberal sampling phase. We can therefore find an expression for E(N) as shown on the bottom of the chart.

VUGRAPH 17

We now had to derive expressions for the known quantities for plans of each of the levels one through five.

As mentioned previously, we also could obtain solutions for levels one and two by using direct algebraic methods, and this served as a partial check on our work. Nothing would be served now by going through each of the derivations step by step, so instead I merely am showing you the resulting equations; you can see how the complexity increases as the number of levels goes up.

This is the result for a one level plan.

VUGRAPH 18

This is the result for a two level plan.

VUGRAPH 19

This is the result for a three level plan.

VUGRAPH 20

This is the result for a four level plan.

VUGRAPH 21

This is the result for a five level plan.

VUGRAPH 22

Now that we had our equations derived, we were ready to carry out computations. We computed not only the expected number of units produced between time of deterioration and resumption of 100% inspection, but we broke this down into expected number of inspected units and expected number of uninspected units. We also computed the expected number of defective units passing during this period. These computations were carried out for several values of p for each plan.

With numerical results, we were ready to evaluate their significance. This became somewhat subjective, since we had no pre-specified criteria. Comparison of CSP-M with other plans was carried out. Even this approach had limitations, since a certain amount of subjectivity is introduced in trying to determine the appropriate plan parameters that should be matched when a comparison of responsiveness is made.

Suffice it to say that, in comparison with other plans, CSP-M performed about as Dodge had suggested; that is, plans of three levels or less performed adequately, while plans of more than three levels were somewhat unresponsive. Again, this conclusion is reached through a certain amount of subjectivity.

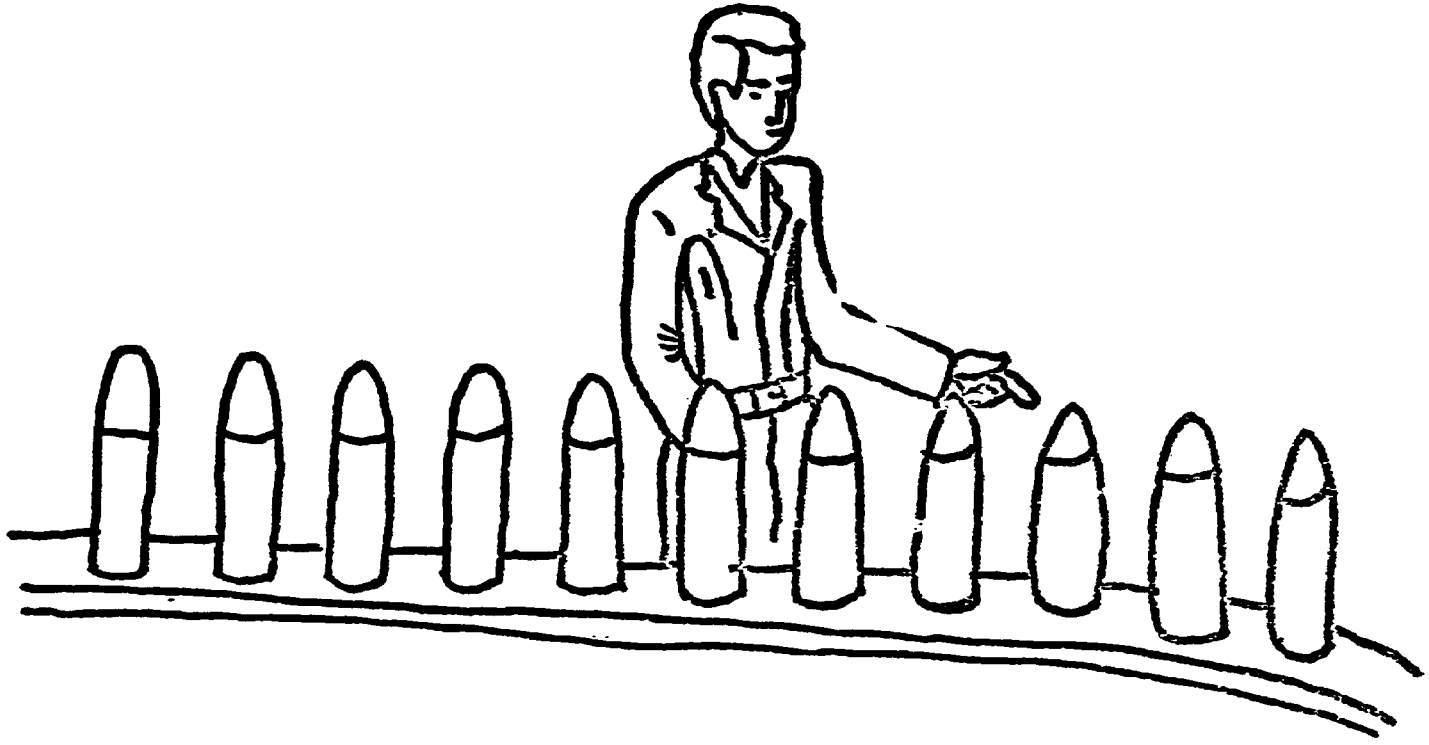
VUGRAPH 23

It might be well to point out here that the method we used could be extended to the case of an assumed value of p other than zero prior to deterioration, in other words, we could use the more general Guthrie and Johns assumptions. This could be done by considering entrance to each step of the plan directly from the 100% inspection phase. It is apparent, though, that the resulting summations would be so complicated that the computational nightmare that would result might not be worth the effort, considering the amount of additional information it would provide.

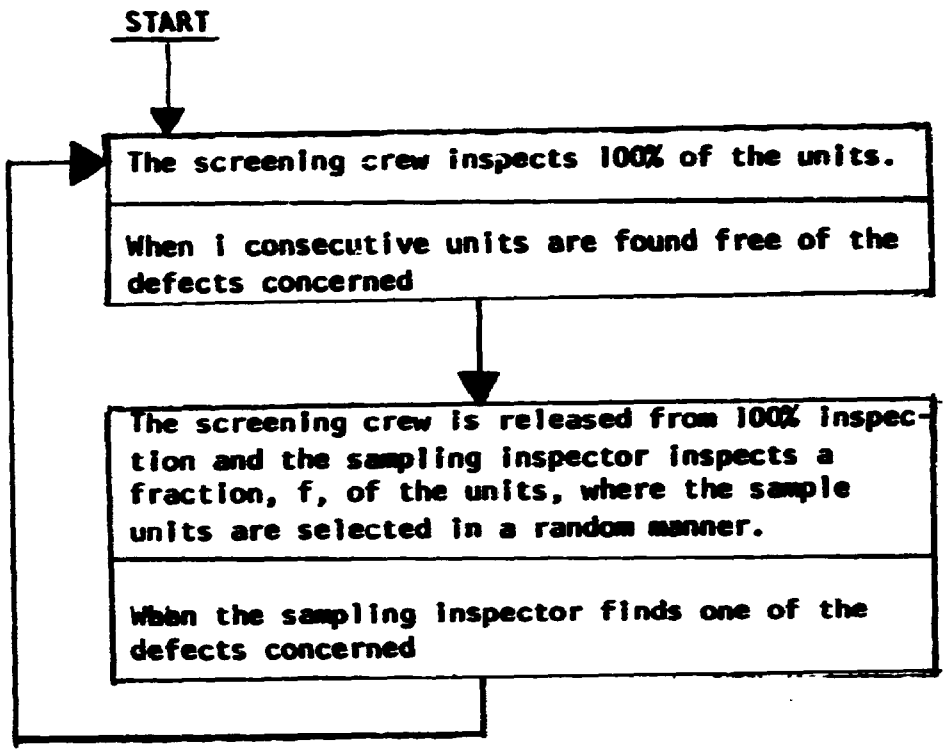
VUGRAPH OFF

I would like to acknowledge the assistance of three of my former employees: Mr. Gary Aasheim and Mr. David Heidt assisted in the review of the works of Hillier and of Guthrie and Johns, while Mr. Heidt and Mr. James Fordice assisted in deriving the so-called known expressions for the responsiveness formulas, and coordinating the programming work.

Inspecter



VUGRAPH 1


START
The screening crew inspects 100% of the units.
When i consecutive units are found free of the defects concerned
The screening crew is released from 100% inspection and the sampling inspector inspects a fraction, f, of the units, where the sample units are selected in a random manner.
When the sampling inspector finds one of the defects concerned

VUGRAPH 2

SOME FUNCTIONAL PROPERTIES

AVERAGE FRACTION INSPECTED (AFI)

THE EXPECTED VALUE OF THE FRACTION OF MATERIAL THAT WILL BE INSPECTED OVER AN INDEFINITELY LONG PERIOD OF TIME WHEN EACH UNIT HAS PROBABILITY p OF BEING DEFECTIVE.

AVERAGE OUTGOING QUALITY (AOQ)

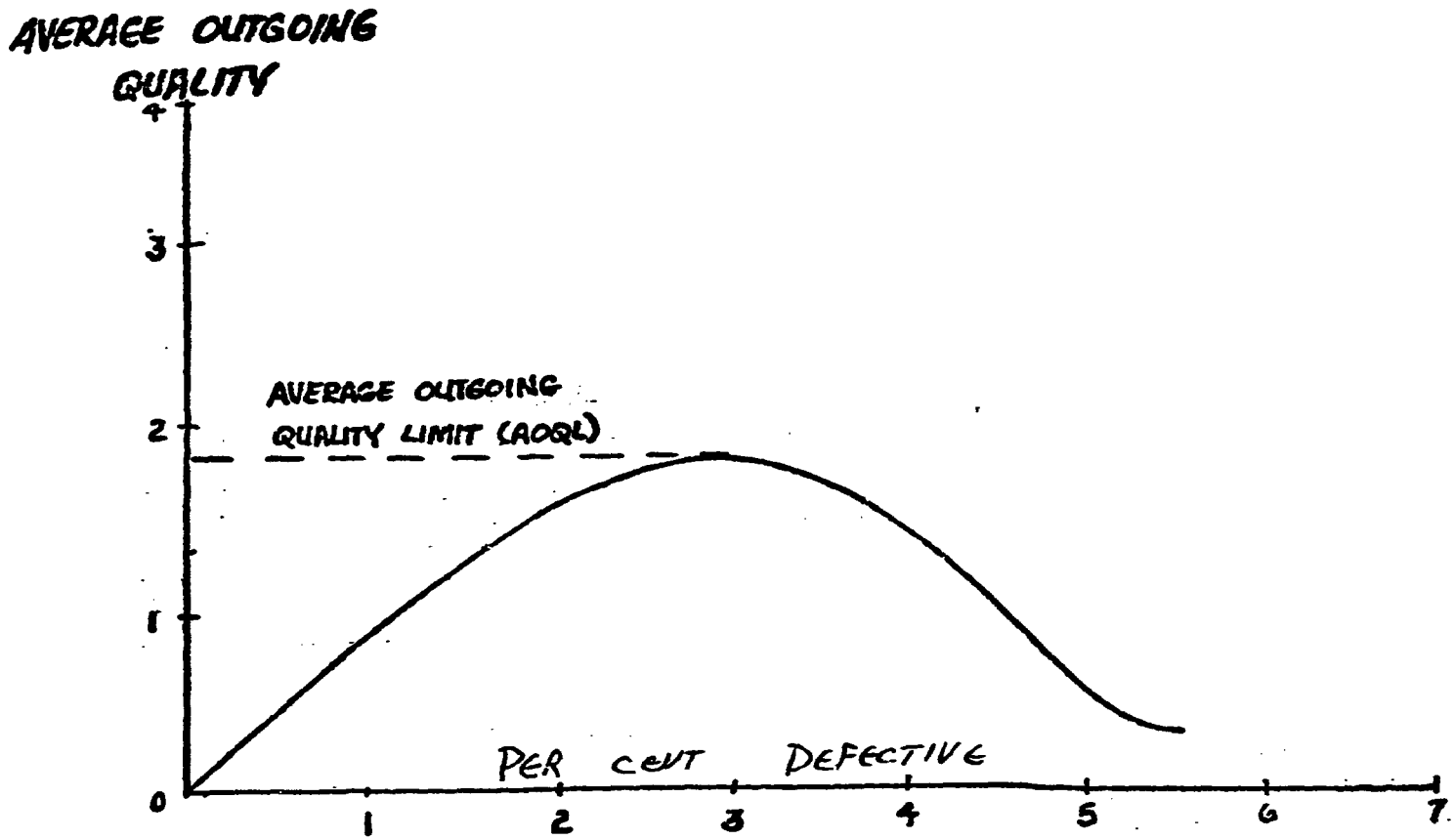
THE EXPECTED FRACTION OF MATERIAL THAT IS DEFECTIVE IN ACCEPTED MATERIAL OVER AN INDEFINITELY LONG PERIOD OF TIME WHEN EACH UNIT HAS PROBABILITY p OF BEING DEFECTIVE.

AVERAGE OUTGOING QUALITY LIMIT (AOQL)

THE MAXIMUM VALUE OF AOQ

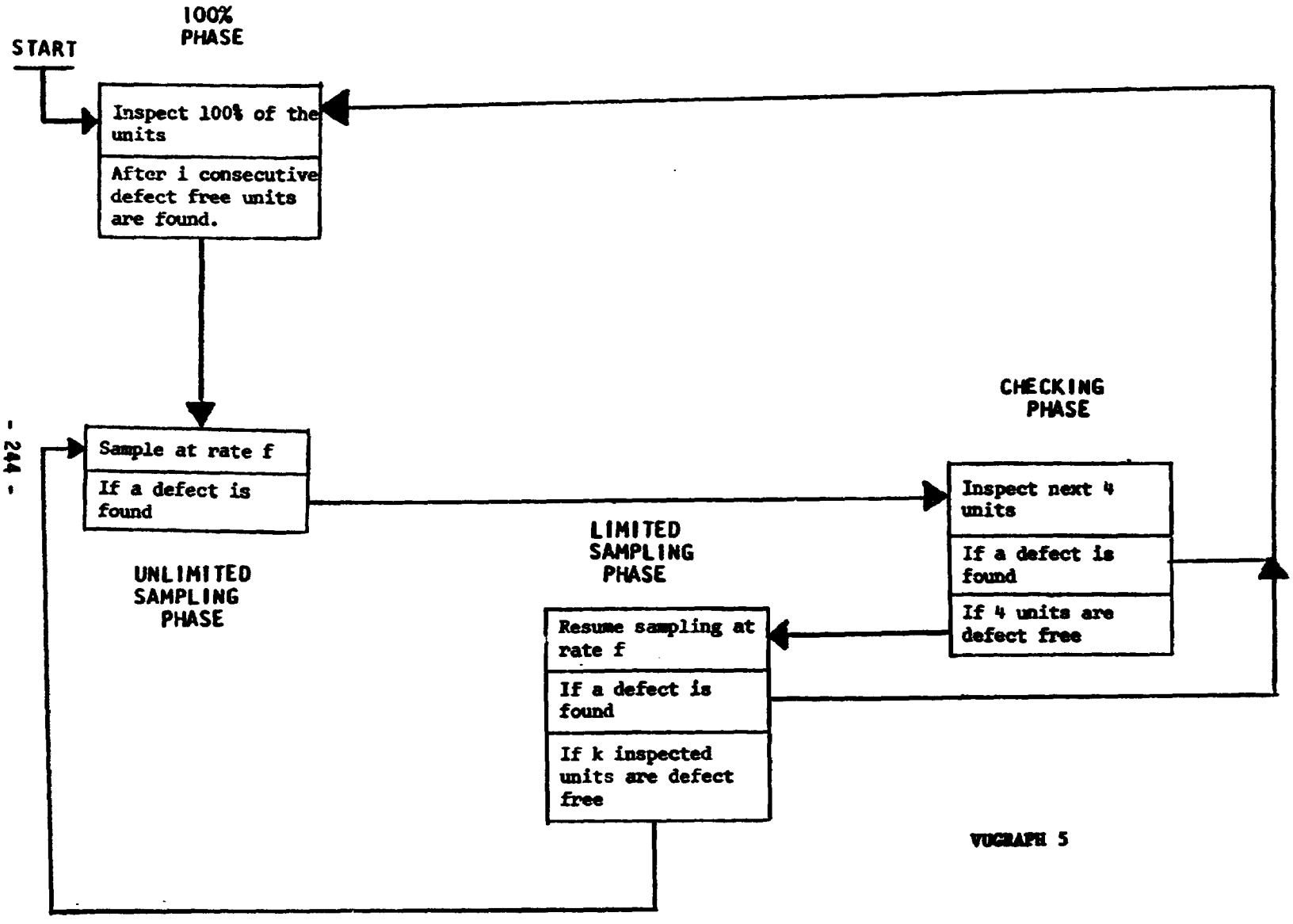
VUGRAPH 3

THE AVERAGE OUTGOING QUALITY (AOQ) CURVE



VUGRAPH 4

PROCEDURE FOR CSP-3 PLANS



VUGRAPH 5

APPROACHES TO FORMULATING MATHEMATICAL PROPERTIES OF CONTINUOUS SAMPLING PLANS

1. AN ALGEBRAIC METHOD DESCRIBING THE PROPERTIES OF THE PLAN - USABLE WHEN THE PLAN CAN BE DESCRIBED AS A UNIQUE CONTINUING CYCLE OF EVENTS, AND THE PROPERTIES OF EACH BLACK BOX CAN BE DETERMINED BY ALGEBRAIC MEANS.

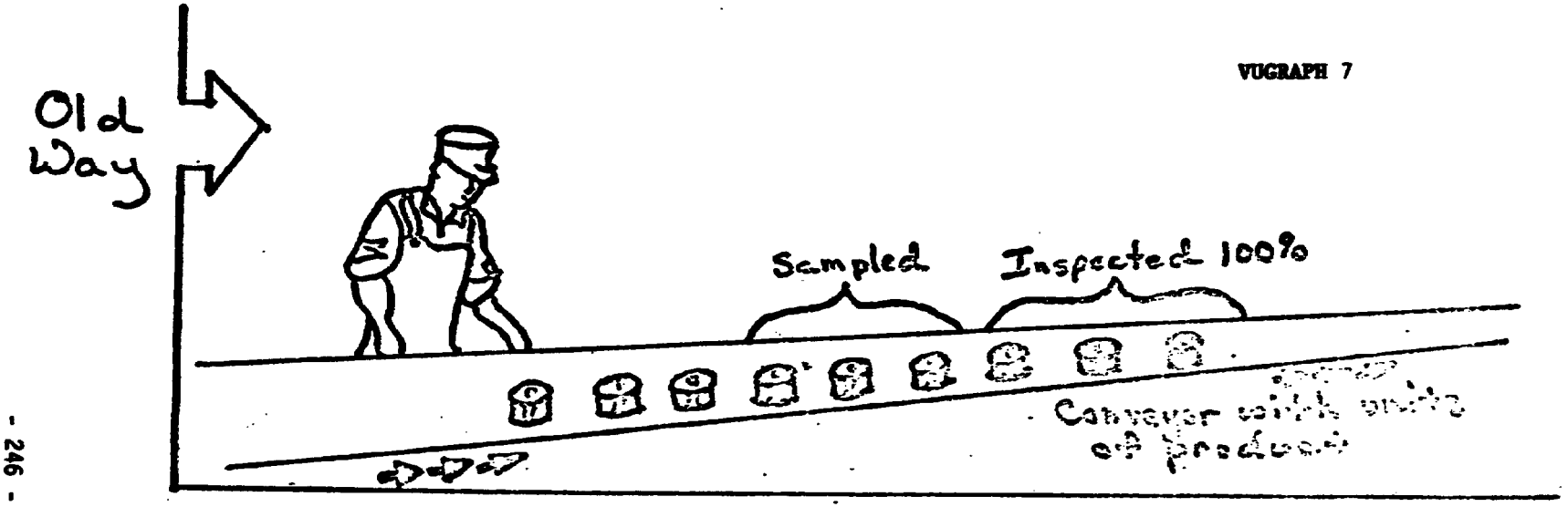
2. THE MARKOV CHAIN APPROACH - USABLE EVEN WHEN THE PLAN CANNOT BE DESCRIBED AS A UNIQUE CONTINUING CYCLE OF EVENTS WITH THE PROPERTIES OF EACH BLACK BOX BEING DETERMINABLE BY ALGEBRAIC MEANS.

3. THE SIMPLIFIED MARKOV CHAIN APPROACH - MATHEMATICALLY EQUIVALENT TO THE MARKOV CHAIN APPROACH, BUT MUCH EASIER TO USE.

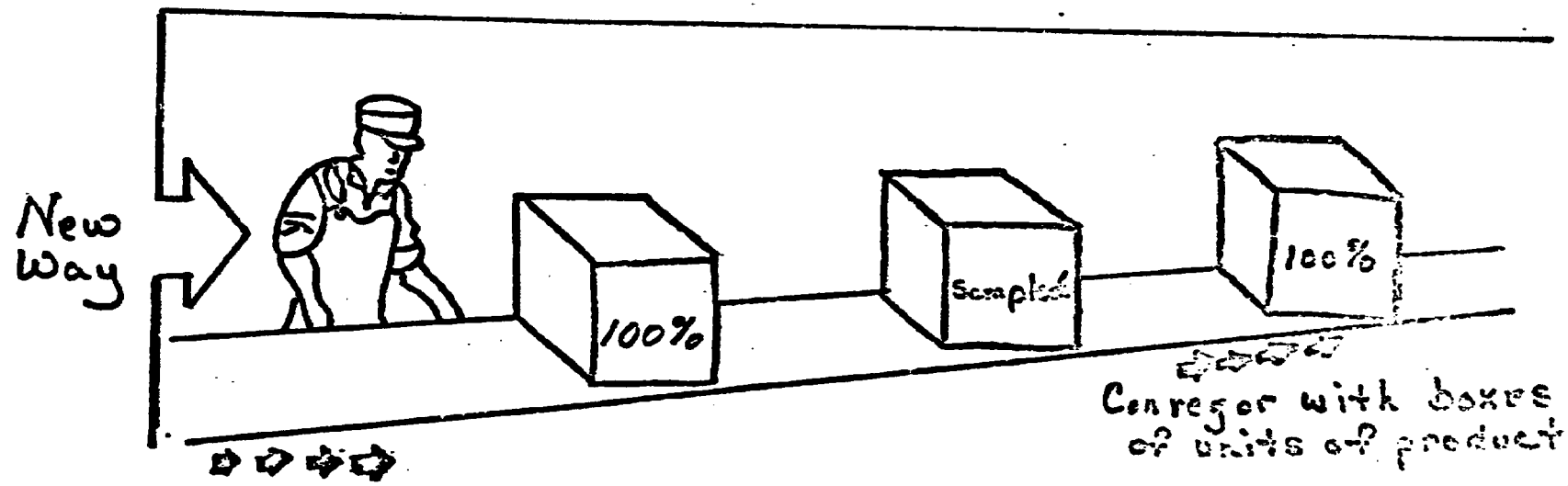
VUGRAPH 6

The Simplified Markov Chain Method

VUGRAPH 7



How The Mathematician Would View the Situation



TRANSITIONAL PROBABILITY MATRIX FOR SIMPLIFIED MARKOV CHAIN DESCRIBING CSP-3

		TO			
		100	UNL SAMP	CH	LIM SAMP
FROM	100	0	1*	0	0
	UNL SAMP	0	0	1**	0
	CH	$1-q^4$	0	0	q^4
	LIM SAMP	$1-q^k$	q^k	0	0

*FOR $p<1$ **FOR $p>0$

RESULTING EQUATIONS	SOLUTION IN TERMS OF P_{us}
$P_{100} = P_{ch}(1-q^4) + P_{ls}(1-q^k)$	$P_{100} = P_{us}(1-q^{k+4})$
$P_{us} = P_{100} + P_{ls}q^k$	$P_{us} = P_{us}$
$P_{ch} = P_{us}$	$P_{ch} = P_{us}$
$P_{ls} = P_{ch}q^4$	$P_{ls} = P_{us}q^4$

VUGRAPH 8

WORKING TABLE FOR FINDING PHASE OCCUPANCY PROBABILITIES

	1	2	3	4	5
PHASE	COEFFICIENT	SIMPLIFICATION OF COLUMN 1	EXPECTED NUMBER OF UNITS IN PHASE	SIMPLIFICATION OF COLUMN 3	PRODUCT OF COLUMNS 2 & 4
100	$1-q^{k+4}$	$1-q^{k+4}$	$\frac{1-q^{i}}{pq^{i}}$	$f(1-q^{i})$	$f(1-q^{i})(1-q^{k+4})$
us	1	1	$\frac{1}{fp}$	q^{i}	q^{i}
ch	1	1	4	$4fpq^{i}$	$4fpq^{i}$
ls	q^{4}	q^{4}	$\frac{1-q^{k}}{fp}$	$q^{i}(1-q^{k})$	$q^{i}q^{4}(1-q^{k})$

VUGRAPH 9

USE OF RESULTS OF WORKING TABLE

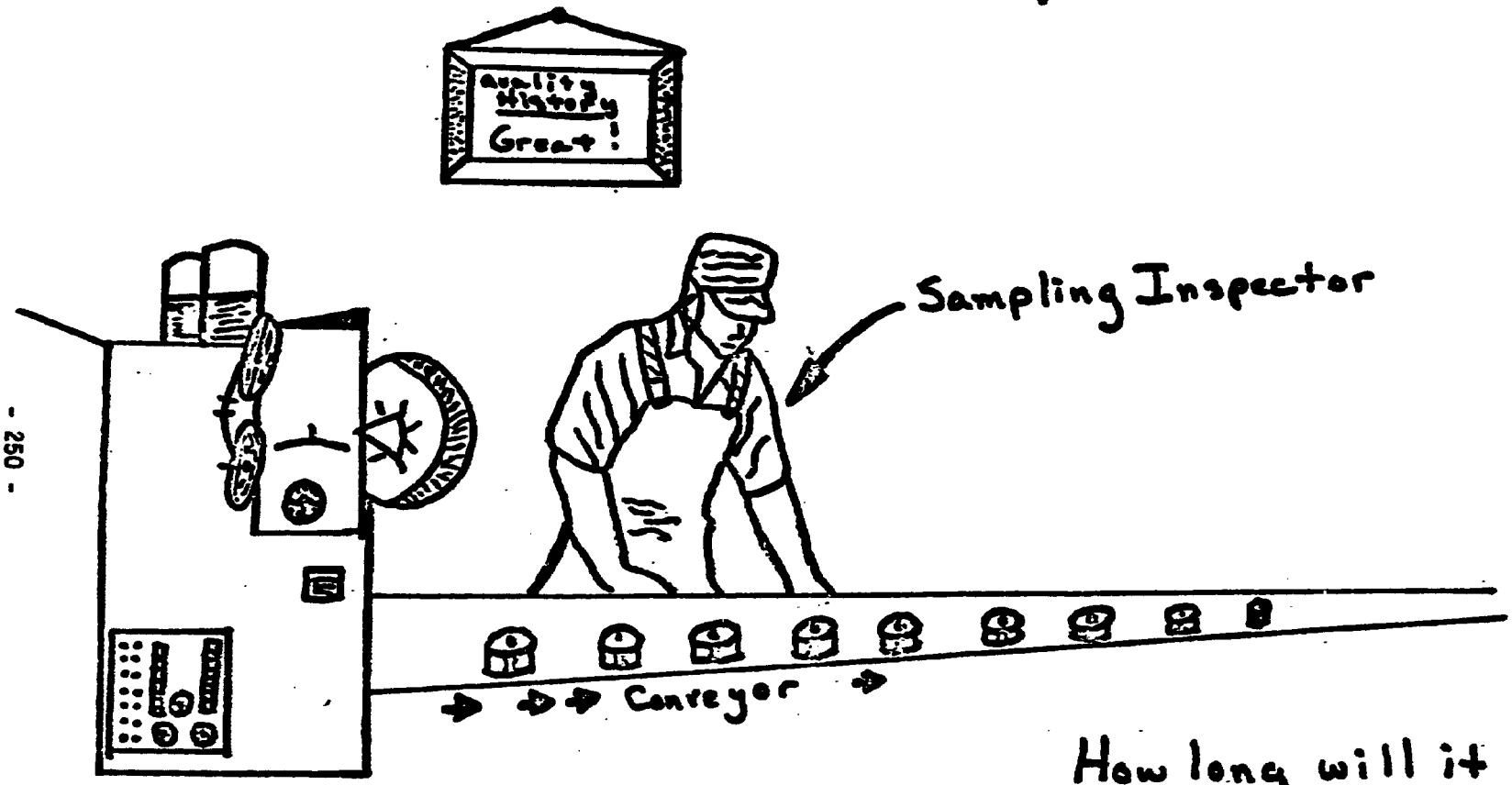
LET SUM OF EXPRESSIONS IN COLUMN 5 = D

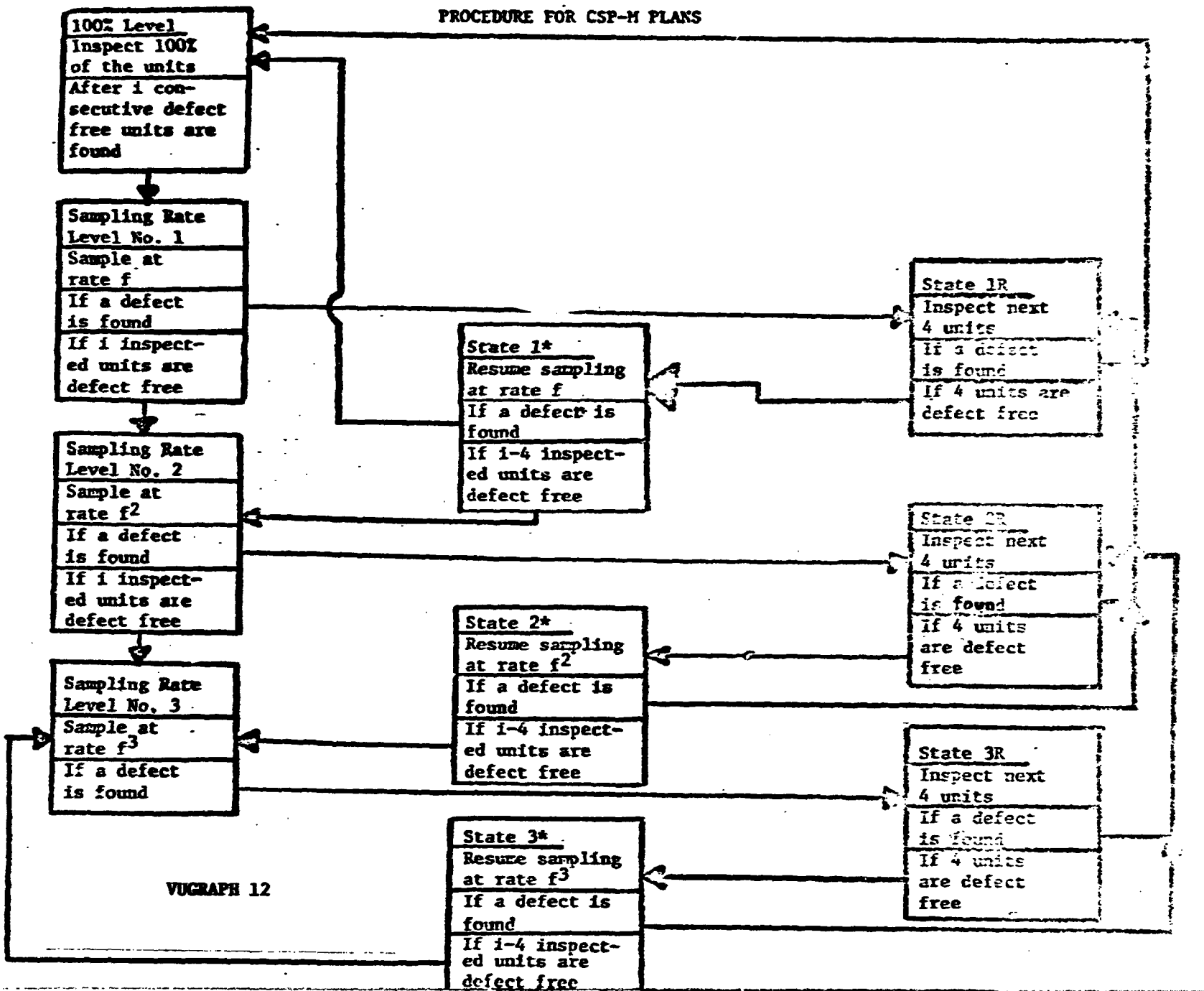
THEN

	COLUMN 5			
100	$f(1-q^l)(1-q^{k+4})$	$P_{100}(U)$	=	$f(1-q^l)(1-q^{k+4})$
us	q^l	$P_{us}(U)$	=	$\frac{q^l}{D}$
ch	$4fpq^l$	$P_{ch}(U)$	=	$\frac{4fpq^l}{D}$
ls	$q^l q^4(1-q^k)$	$P_{ls}(U)$	=	$\frac{q^l q^4(1-q^k)}{D}$

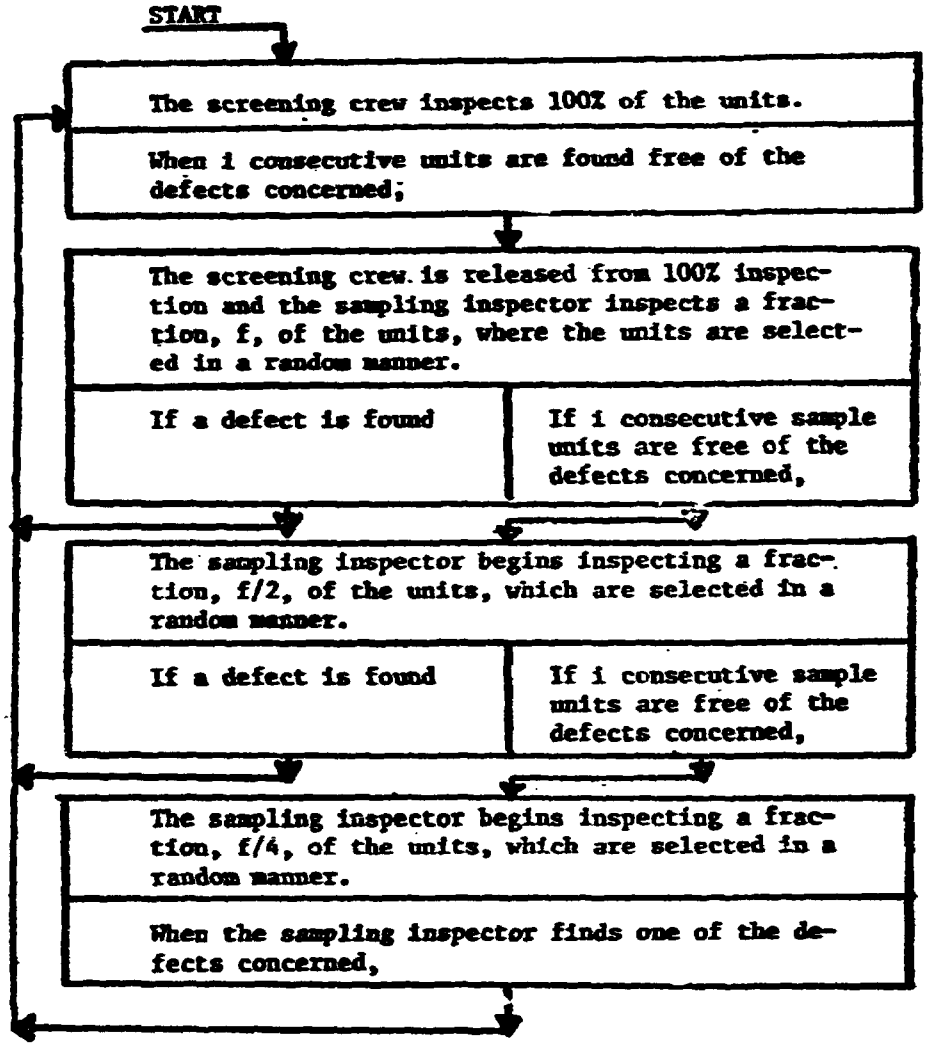
VUGRAPH 19

The Problem of Responsiveness




PROCEDURE FOR CSP-M PLANS
100% Level
Inspect 100% of the units
After i consecutive defect free units are found
Sampling Rate Level No. 1
Sample at rate f
If a defect is found
If i inspected units are defect free
Sampling Rate Level No. 2
Sample at rate f2
If a defect is found
If i inspected units are defect free
Sampling Rate Level No. 3
Sample at rate f3
If a defect is found
State 1*
Resume sampling at rate f
If a defect is found
If i-4 inspected units are defect free
State 2*
Resume sampling at rate f2
If a defect is found
If i-4 inspected units are defect free
State 3*
Resume sampling at rate f3
If a defect is found
If i-4 inspected units are defect free
State 1R
Inspect next 4 units
If a defect is found
If 4 units are defect free
State 2R
Inspect next 4 units
If a defect is found
If 4 units are defect free
State 3R
Inspect next 4 units
If a defect is found
If 4 units are defect free

VUGRAPH 12


START
The screening crew inspects 100% of the units.
When i consecutive units are found free of the defects concerned,
The screening crew is released from 100% inspection and the sampling inspector inspects a fraction, f, of the units, where the units are selected in a random manner.
If a defect is found
If i consecutive sample units are free of the defects concerned,
The sampling inspector begins inspecting a fraction, f/2, of the units, which are selected in a random manner.
If a defect is found
If i consecutive sample units are free of the defects concerned,
The sampling inspector begins inspecting a fraction, f/4, of the units, which are selected in a random manner.
When the sampling inspector finds one of the defects concerned,

VUGRAPH 13

GUTHRIE AND JOHNS METHOD

$$E(N) = \sum_{\substack{\text{all steps} \\ \text{in plan}}} P(\text{being at step } j \text{ and finding defect } \mid p = p_o)(\text{Expected number of units to return to 100\% inspection from step } j \mid p = p_1),$$

where p_o = probability before deterioration that a given unit is defective.

p_1 = probability after deterioration that a given unit is defective.

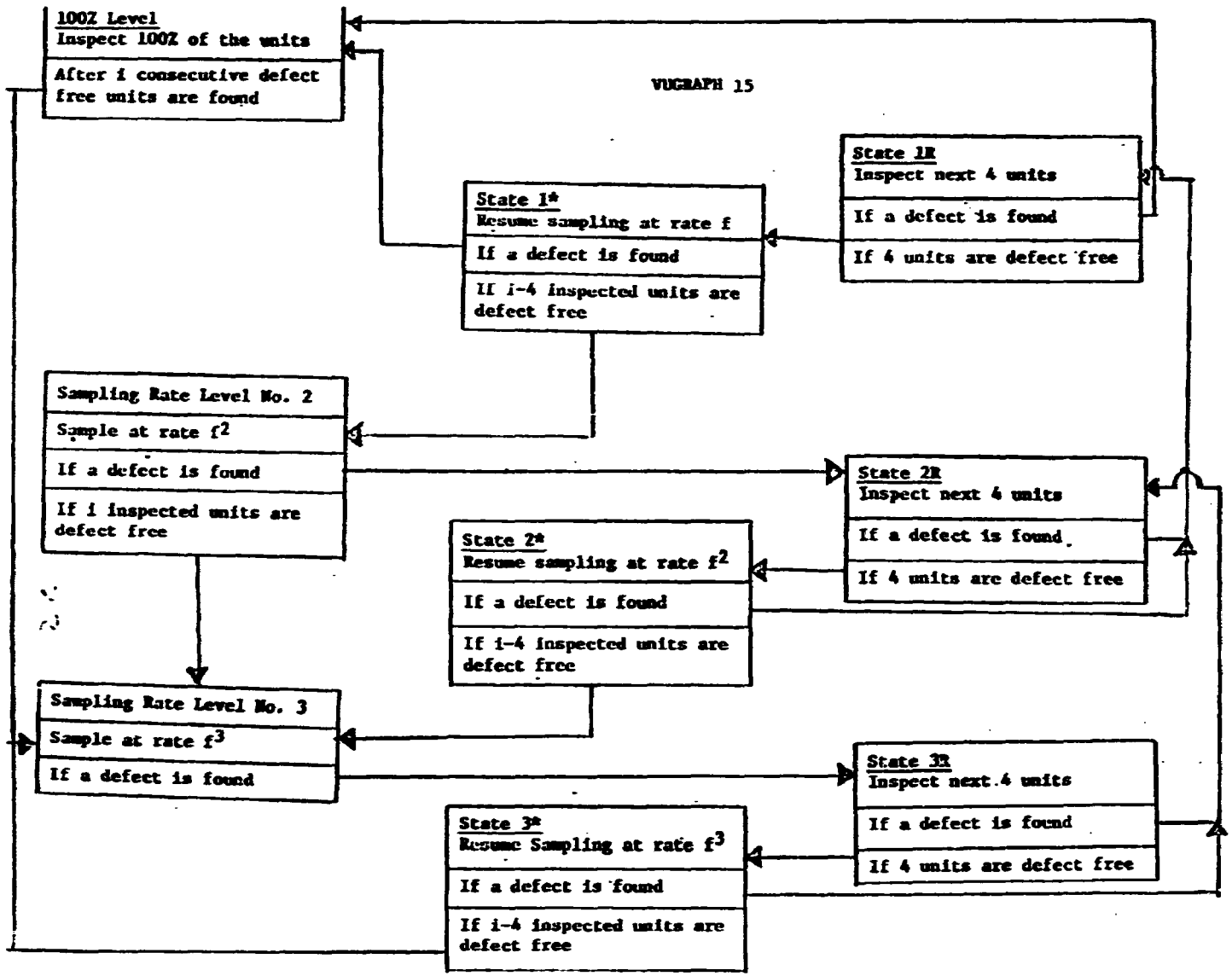
OUR METHOD

$$E(N) = (\text{Expected number of units to return to 100\% inspection from most liberal sampling step} \mid p = p_1),$$

in effect, assuming $p_o = 0$.

VUGRAPH 14

VUGRAPH 15


100Z Level
Inspect 100Z of the units
After i consecutive defect free units are found
State 1R
Inspect next 4 units
If a defect is found
If 4 units are defect free
State 1*
Resume sampling at rate f
If a defect is found
If i-4 inspected units are defect free
Sampling Rate Level No. 2
Sample at rate f^2
If a defect is found
If i inspected units are defect free
State 2R
Inspect next 4 units
If a defect is found
If 4 units are defect free
State 2*
Resume sampling at rate f^2
If a defect is found
If i-4 inspected units are defect free
Sampling Rate Level No. 3
Sample at rate f^3
If a defect is found
State 3R
Inspect next 4 units
If a defect is found
If 4 units are defect free
State 3*
Resume Sampling at rate f^3
If a defect is found
If i-4 inspected units are defect free

$$\frac{P_{100}}{P_{return}} = \frac{E(100\% \text{ phase length})}{E(N)}$$

where N denotes number of units required to return to 100% inspection after first entrance (per cycle) in the most liberal sampling phase.

Therefore, $E(N) = \frac{(P_{return})E(100\% \text{ phase length})}{P_{100}}$

VUGRAPH 16

ONE LEVEL PLAN

$$E(N) = \frac{1+4fp+q^4-q^1}{fp(1-q^1)}$$

VUGRAPH 17

TWO LEVEL PLAN

$$E(N) = \frac{4}{1-q^4} + \frac{q^4-q^1}{fp(1-q^4)} + \frac{1}{f^2p(1-q^4)^2} + \frac{4}{(1-q^4)^2} + \frac{q^4-q^1}{f^2p(1-q^4)^2}$$

VUGRAPH 18

THREE LEVEL PLAN

$$E(N) = \frac{4}{1-q^1} + \frac{q^4+q^1}{fp(1-q^1)} + \frac{q^1}{f^2p} + \frac{4}{(1-q^1)^2} + \frac{q^4-q^1}{f^2p(1-q^1)^2} + \frac{1-q^1(1-q^1)^2}{f^3p(1-q^1)^3}$$

$$+ \frac{4[1-q^1(1-q^1)^2]}{(1-q^1)^3} + \frac{(q^4-q^1)[1-q^1(1-q^1)^2]}{f^3p(1-q^1)^3}$$

VUGRAPH 19

FOUR LEVEL PLAN

$$E(N) = \frac{4}{1-q^i} + \frac{q^4-q^i}{fp(1-q^i)} + \frac{q^i}{f^2p} + \frac{4}{(1-q^i)^2} + \frac{q^4-q^i}{f^2p(1-q^i)^2} + \frac{q^i[1+q^i(1-q^i)]}{f^3p(1-q^i)} +$$

$$\frac{4[1-q^i(1-q^i)^2]}{(1-q^i)^3} + \frac{(q^4-q^i)[1-q^i(1-q^i)^2]}{f^3p(1-q^i)^3} + \frac{(1-q^i)^2+q^{2i}(2+q^i-3q^{2i}+q^{3i})}{f^4p(1-q^i)^4} +$$

$$\frac{4[(1-q^i)^2+q^{2i}(2+q^i-3q^{2i}+q^{3i})]}{(1-q^i)^4} + \frac{(q^4-q^i)[(1-q^i)^2+q^{2i}(2+q^i-3q^{2i}+q^{3i})]}{f^4p(1-q^i)^4}$$

VUGRAPH 20

FIVE LEVEL PLAN

$$E(N) = \frac{4}{1-q^1} + \frac{q^4-q^1}{fp(1-q^1)} + \frac{q^1}{f^2p} + \frac{4}{(1-q^1)^2} + \frac{q^1-q^4}{f^2p(1-q^1)^2} + \frac{q^1[1+q^1(1-q^1)]}{f^3p(1-q^1)} +$$

$$\frac{4[1-q^1(1-q^1)2]}{(1-q^1)^3} + \frac{(q^4-q^1)[1-q^1(1-q^1)2]}{f^3p(1-q^1)^3} + \frac{q^1[1+q^{21}(1-q^1)\ (2-q^1)]}{f^4p(1-q^1)^2} +$$

$$\frac{4[1-q^1(1-q^1)^2(2+q^1-q^{21})]}{(1-q^1)^4} + \frac{(q^4-q^1)[1-q^1(1-q^1)^2(2+q^1-q^{21})]}{f^5p(1-q^1)^5} +$$

$$\frac{4\{(1-q^1)^4+q^1[q^1(1-q^1)^2(-q^{31}+3q^{21}-q^1-2)+q^1(1-q^1)+1]\}}{(1-q^1)^5} +$$

$$\frac{(q^4-q^1)\{(1-q^1)^4+q^1[(1-q^1)^2(-q^{31}+3q^{21}-q^1-2)+q^1(1-q^1)+1]\}}{f^5p(1-q^1)^5}$$

VUGRAPH 21

QUANTITIES COMPUTED

EXPECTED NUMBER OF UNITS PRODUCED ● ● ●

EXPECTED NUMBER OF UNITS INSPECTED ● ● ●

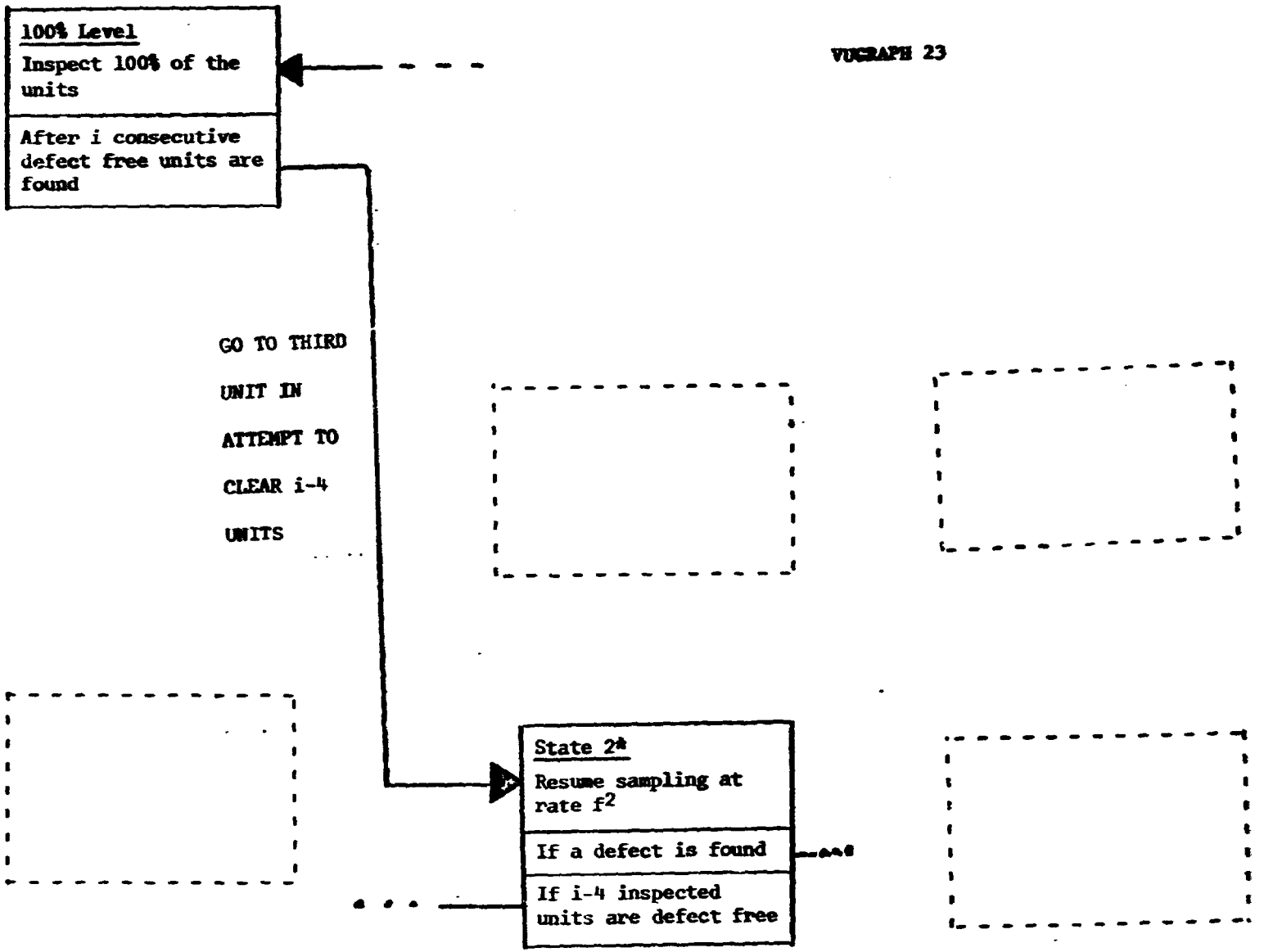
EXPECTED NUMBER OF UNITS NOT INSPECTED ● ● ●

EXPECTED NUMBER OF DEFECTIVE UNITS PASSING ● ● ●

● ● ● BETWEEN TIME OF DETERIORATION AND RESUMPTION OF 100% INSPECTION

VUGRAPH 22

VUGRAPH 23


100% Level
Inspect 100% of the units
After i consecutive defect free units are found
GO TO THIRD UNIT IN ATTEMPT TO CLEAR i-4 UNITS
State 2*
Resume sampling at rate f2
If a defect is found
If i-4 inspected units are defect free

EXPECTED VALUES FROM A MARKOV CHAIN MODEL OF CSP-F AND THE Z-TRANSFORM

David L. Arp
Weapons Planning Group
Naval Weapons Center
China Lake, California
93555

ABSTRACT. The main topic of this paper deals with the development of a computationally feasible expansion of the AFI function for the short-run case of a continuous sampling plan (CSP-F) which, in turn, is treated as a finite ergodic Markov chain.

A time-lapsed difference equation for the probability of being in the sampling stage at a given step is derived; for a finite run of length N, the expected value of the random variable "fraction of total (operational) time in inspection states" (AFI(N)) is then expressed as a linear combination of the above probabilities. Unfortunately, employing the above difference equation as a basis for recursion results in an algorithm which rapidly becomes too complex and time-consuming for any computer now in use--including, *a fortiori*, the smaller programmable calculators. Mr. Richard Brugger attempted to solve this problem by deriving an exact, computable expression for AFI(N) based on probabilistic reasoning [7.1, pp. 929-948]. Unfortunately, it turned out that, although his approach did yield a workable expression which converged to the long-run AFI function, it did not provide an exact formula, nor, for that matter, any concisely expressed error terms.

My approach to this problem is to derive a true asymptotic expansion of AFI(N) by employing the Z-transform method to solve the above mentioned difference equation (yielding a function of a complex variable, regular outside of the unit circle); contour integration, series expansions, and the residue calculus to "project out" the desired terms; and finally, the binomial and LaGrange expansions to yield practical approximations with implicit error terms. It is further shown how to apply this technique to obtain a similar expansion for the variance of the aforementioned random variable. Utilizing this latter expansion along with Martingale Theory, bounds are obtained on certain probability statements. Finally, the method is applied to other quantities of interest arising not only from CSP(N) but also from more complex continuous sampling plans.

1.0 INTRODUCTION. This paper is the result of work done by the author on finite run CSP-1 while working for the now defunct Army Munitions Command (MuCom). The main topic, that of deriving compact formulas for the expected value and variance of the random variable "fraction inspected in a finite run of length N," was done at the request of Mr. Richard M. Brugger, mathematician and former chief of a Concepts Branch located in a command subordinate to MuCom. I wish to thank him for his encouragement and his efforts to provide a stimulating work-research environment.

A more than passing acquaintance with the theory of functions of one complex variable is assumed throughout this paper. In addition to the principal topic briefly mentioned above, another one, appearing in chapter 6, has been added.

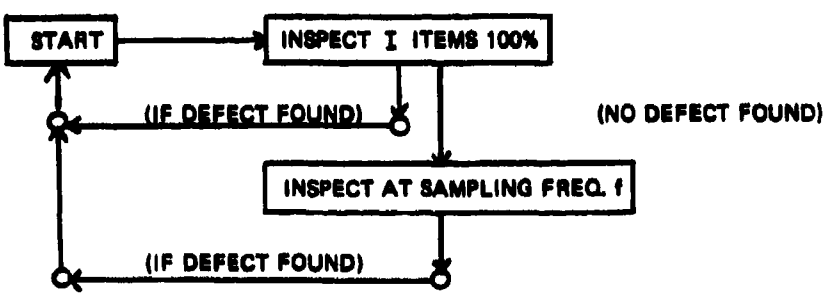
Before tackling the body of the paper, some background discussion and preliminary theoretical work are necessary.

2.0 BACKGROUND

2.1 CSP(N) and CSP(∞). In this paper CSP-1 (CSP-F) is denoted by CSP(∞) (CSP(N), N = total number of units in the finite run). The "box" scheme for both plans is given in Figure 1.

FIGURE 1

CSP(∞) AND CSP(N)



FOR CSP(N), STOP AFTER UNIT N

The plan parameters, with practical ranges, are

(a) I = CLEARANCE NO., $10 \leq I \leq 2 \times 10^4$.

(b) f = FREQUENCY NO., $5 \times 10^{-3} \leq f \leq 5 \times 10^{-1}$.

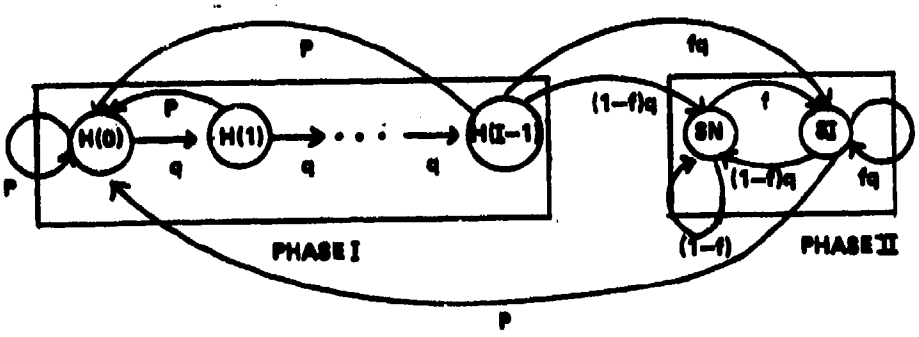
(c) p = PROBABILITY OF A DEFECTIVE UNIT, $p \sim 10^{-1}$.

(d) N = RUN NO. FOR CSP(N), $10^3 \leq N \leq 10^6$.

In order to analyse the effectiveness of such a plan, and, in particular, the values of the parameters to achieve a desired level of quality control, it is necessary to formulate a reasonable and mathematically tractable model which in turn can be studied. One such possible model, in general use, is to treat CSP as a finite, ergodic, and homogeneous Markov chain. Homogeneity necessarily implies that the probability of a defective unit is constant over the entire production run. To learn more of the history and justifications of this particular model, one can consult reference 7.1 where a lucid discussion of and further references to these topics are given. For the purposes of this presentation, it suffices to note that such a model is known to give upper bounds for the quantities of interest. A layout of this model appears in Figure 2.

FIGURE 2

STANDARD MARKOV-CHAIN APPROACH TO CSP



(1) PHASE I = INSPECTION PHASE
PHASE II = SAMPLING PHASE

(2) H(J) = Jth STATE IN PHASE I

SN = NON-INSPECTION STATE

SI = INSPECTION STATE

(3) q = 1 − p

In Figure 2 we have (I+2) possible states with the non-zero transitional probabilities between them given by the quantities below the connecting arrows. A tedious but routine analysis of this model shows it to be in fact ergodic with an easily constructed finite (I+2) × (I+2) constant transitional matrix [7.1, p. 938].

Henceforth, use of CSP, without any further qualification, will always mean both types of plans. Furthermore, the use of "plans" relative to one or both of these types will mean "for various selections of the parameters"; thus, for example, CSP(N,I,f) and CSP(N',I', f') are considered to be two different kinds of one type of plan.

Practically speaking, there are only two types of starting conditions (initial probability vectors) that need to be considered for CSP: JOB-SHOP and ARBITRARY-ENTRY.

In the JOB-SHOP case, the sampling begins in state H(0) (or the "start" box in Figure 1). The corresponding initial probability vector, denoted by ι (iota), and its conditional expected-value operator are given, respectively, by

$$\iota = (1, 0, \ldots 0) \text{ of dimension } I+2, \text{ and}$$

$$E_{\iota}[\cdot].$$

In contrast to this case, where realistically sampling begins at the start of production, we have the ARBITRARY-ENTRY case. In this latter type, sampling has been going on long enough for statistical equilibrium to have been achieved. Therefore, at any arbitrary time, sampling can be assumed to be at a given state with its long-run (ergodic) probability. In this situation, the initial probability vector and conditional expected-value operator are given, respectively, by

$$\Pi = (\Pi_1, \Pi_2, \ldots, \Pi_{I+2}) \text{ and } E_{\Pi}[\cdot],$$

where

where Π_j is the long run probability of being in (or starting in) state j; I+1 (I+2) means SN (SI), and 1 (I) refers to H(0) (H(I-1)). For future use, we define

$$\Pi_{II} = \Pi_{I+1} + \Pi_{I+2}.$$

We can now define the AFI(N) function. Let ω be an ordered infinite sequence of states whose order is compatible with the Markov chain structure (if we agree to concentrate the probability measure induced by the Markov chain on such sequences, we can consider any ordered sequence--compatible or not). Then we have the following definitions:

Definition 1. Given ω as above,

$$PROJ_k(\omega) = \text{The (Markov) state in the } k^{th} \text{ slot of } \omega.$$

Furthermore, the outcome space of all such ω's with the probability measure, P, induced by the Markov chain, is denoted by (Ω, P) [see 7.9, for a clear, concise description of σ-algebras and measures induced on product spaces].

Two characteristic random variables are defined next, then a random variable is defined in terms of either one of them. This dual situation will occur throughout the rest of this paper and arises because of alternate ways of looking at AFI(N).

Definition 2. (Characteristic random variables)

$$Z_k(\omega) = \begin{cases} 1, & \text{if } PROJ_k(\omega) \in PHASE\ II \\ 0, & \text{otherwise.} \end{cases}$$

$$X_{SN}(k)(\omega) = \begin{cases} 1, & \text{if } PROJ_k(\omega) \in SN \\ 0, & \text{otherwise.} \end{cases}$$

Definition 3. For $\nu = 1 - f$,

$$FI(N)(\omega) = 1 - \frac{\nu}{N} \sum_{j=1}^{N} Z_{j-1}(\omega)$$

$$= 1 - \frac{1}{N} \sum_{j=1}^{N} X_{SN}(j)(\omega).$$

In terms of the random variable FI(N), we can finally define AFI(N) for the two "initial-condition" cases:

Definition 4. For the JOB-SHOP case (case I),

AVERAGE FRACTION INSPECTED = E_t [FI(N)].

For the ARBITRARY-ENTRY case (case II):

$$AFI(N) = E_{\Pi}[FI(N)].$$

Concerning these two functions and their relationships with AFI(∞), we have

Proposition 1. For case I,

$$\lim_{N \to \infty} AFI(N) = AFI(\infty) = 1 - \Pi_{SN} = 1 - \nu\alpha(\infty).$$

For case II,

$$AFI(N) = 1 - \Pi_{SN} = AFI(\infty).$$

Proof. (case I)

$$\lim_{N\to\infty} AFI(N) = E_t\left[\lim_{N\to\infty} FI(N)\right], \text{ by the}$$

Lebesque Dominated Conv. Thm.,

$$= E_t\,[FI(\infty)], \text{ by the}$$

Individual Ergodic Thm.,

$$= AFI(\infty), \text{ by definition.}$$

Case II follows trivially from the definitions. Q.E.D.

To round out the relationship between AFI(N) and AFI(∞), we state, without proof [see the proposition, chp. 4], a final equation:

$$\lim_{N\to\infty} VAR(FI(N)) = 0 \quad \text{(both cases)}.$$

2.2 The difference equation. We now turn our attention to the principal difference equation, which will eventually be used to derive an expansion for AFI(N) (JOB-SHOP case).

Letting $\alpha(k) = E_t\,[Z_{k-1}]$ $(K \geq 1)$, we clearly have, by definition,

$$AFI(N;p,f,I) = 1 - \frac{\nu}{N}\sum_{j=1}^{N} \alpha(j).$$

Whether reasoning directly from the definition of $\alpha(j)$ or using the Chapman-Kolmogorov equation [7.5, p. 194] and the relation between Z_j and $X_{SN}(j)$, we can derive the following difference equation for $\alpha(\cdot)$:

$$(A)\quad \alpha(k) = \begin{cases} 0, & 1 \leq k \leq I \\ q^I, & k = I+1 \\ \beta\alpha(k-1) + \gamma - \theta\alpha(k-1-I), & k \geq I+2. \end{cases}$$

Here, $\beta = 1 - fp$, $\gamma = pq^I$, $\theta = \nu\gamma$, and $q = 1 - p$. Two courses of action are now open to us. Either solve the given difference equation theoretically or use the equation itself as a basis for a recursive method for computing AFI(N) on a (digital) computer. The trouble with this latter route is the step or (operational) time lag, I, which is blocked out and emphasized by an arrow in (A).

2.3 Computational difficulties. We will demonstrate in this section that a recursion algorithm based on the lag in (A) is impractical for many realistic applications of the AFI(N) function.

With minicomputers a step lag of size I for the practical values given in section 1 would require the continual dynamic storage of up to 2×10^4 quantities, which is too large (and at least ungainly for the larger machines) for them. Moreover, for the larger machines this lag would lead to a large computation time for many of the standard calculations which are done in continuous sampling work. Short explanations of three of the more common types of these calculations will give a clearer idea of the difficulties involved for both types of computers. In the discussion below, variables appear to the left of the semicolon, parameters to the right.

The first kind of calculation arises in the graphical approach to the study of CSP(N), which can involve plotting some function derived from AFI(N). Thus the problem confronted by the mathematician (or lesser beings) could be of the form "plot FN(AFI(p;N,f,I)) as p varies over the unit interval and do a sensitivity analysis on the parameters." A second common type of calculation encountered could be of the form: find (constrained) I such that $\max_p$[FN(AFI(p,I;N,f))] is equal to some fixed number. If, for instance, the function were the AOQ, the problem would then read: for a given run number N and sampling frequency f, find the clearance number I to insure that the fraction of defective items passed will not exceed, on the average, a given number. The last frequently met type of problem is to plot FN(AFI(N);p,f,I) in order to study convergence properties as N becomes large. Recalling the practical ranges of the parameters from the first section and the step lag from the second, we see that the kind of computational routine for any of the above examples would require a great deal of processing time.

In concluding this section, it is hoped that the need for the former course of action suggested in section 2 has now been demonstrated. The remainder of this paper is devoted to the development of such an expansion as well as certain ancillary results.

3.0 DERIVATION OF THE AFI(N) EXPANSION

3.1 The Z-transform method. The basic idea of this method is to (1) consider a sequence $\{\sigma_j\}$ $j = 0$ to ∞, as a function from the natural numbers, NN, to the real line (i.e., σ: NN$\mapsto$R); (2) to employ the Z-transform to map such a function to one of a complex variable, $\sigma(z)$, regular in a neighborhood of ∞; and (3) to project back to the original function by contour integration. As expected with a transform method, it is found that this particular method transfers the analytical problem of solving a linear difference equation to the much simpler algebraic one of solving for an unknown (complex function) which is implicitly given by an algebraic equation.

To carry out this program, we must make precise definitions.

Definition 1. Given a sequence, $\sigma(j)$, considered as a function $\sigma: \mathbb{N} \to \mathbb{R}$, its Z-transform is

$$\hat{\sigma}(z) = \sum_{j=0}^{\infty} \frac{\sigma(j)}{z^j} \quad \text{(Laurent expansion)}.$$

$\hat{\sigma}(z)$ is a function which is clearly either regular in a neighborhood of ∞ (i.e., for all z such that $|z| > R(\sigma)$) or only at ∞ itself; in the latter situation, the Z-transform method still works. It should also be clear that the size of $R(\sigma)$ depends on the growth properties of $\sigma(j)$ as $j \to \infty$. We next define two standard sequences and an operation between any two arbitrary ones.

Definition 2. The Dirac sequence at k, $\delta_k: \mathbb{N} \to \{0,1\}$, is defined via

$$\delta_k(j) = \begin{cases} 1, & \text{for } k = j \\ 0, & \text{otherwise.} \end{cases}$$

From definition 1, its Z-transform is clearly given by

$$\hat{\delta}_k(z) = \frac{1}{z^k}.$$

Definition 3. The Heaviside sequence at k, $H_k: \mathbb{N} \to \mathbb{N}$, is defined as

$$H_k(j) = \sum_{h=k}^{\infty} \delta_h(j).$$

Again, from definition 1, we have its Z-transform given by

$$\hat{H}_k(z) = \left(\frac{1}{z^{k-1}}\right)\left(\frac{1}{z-1}\right).$$

Definition 4. Given two sequences, $\sigma_i (i = 1, 2)$, their convolution, $\sigma_1 * \sigma_2$, is a new sequence given by

$$(\sigma_1 * \sigma_2)(k) = \sum_{h=0}^{k} \sigma_1(k-h)\, \sigma_2(h).$$

Letting $RM = \text{Max}\,(R(\sigma_1), R(\sigma_2))$ and recalling the Cauchy product for the multiplication of two power series, we have the following relation between "$*$" and "$\wedge$":

$$\widehat{(\sigma_1 * \sigma_2)}(z) = \hat{\sigma}_1(z)\,\hat{\sigma}_2(z), \text{ for } |z| > RM.$$

3.2 Difference equation revisited. Given the definitions in the last section, we can now express the difference equation given in chapter 2 more concisely--including the initial conditions as well.

Proposition 1. The difference equation for α: $NN \to [0, 1]$ is given by

$$\text{(A)} \quad \{\alpha * (\delta_0 - \beta\delta_1 + \theta\delta_{I+1})\}(k) = q^I \delta_{I+1}(k) + \gamma H_{I+2}(k).$$

Its Z-transform is given by

$$\text{(B)} \quad \hat{a}(z) = \frac{q^I}{p(z)} + \frac{\gamma}{p(z)(z-1)}, \quad p(z) = z^I(z - \beta) + \theta.$$

Proof. (A) follows from the relevant definitions; (B) also follows from these definitions and collecting terms. Q.E.D.

The next proposition will be useful both as a check and a tool for formula derivations.

Proposition 2. (End-point property of $\hat{\sigma}(z)$)

$$\lim_{z \to 1} \left(\frac{z-1}{z}\right)\hat{\sigma}(z) = \lim_{N \to \infty} \sigma(N) \text{ if the R.H.S.}$$

exists; moreover, $\hat{\sigma}(z)$ is then regular (at least) outside of the unit disk.

Proof. [see 7.4, pp. 257-258.] Q.E.D.

Corollary. $\hat{a}(z)$ satisfies the above proposition 2 and is therefore regular outside of the unit disk.

Proof. From proposition 1,

$$\lim_{z \to 1} \left(\frac{z-1}{z}\right)\hat{a}(z) = \frac{\gamma}{1-(\beta-\theta)} = \Pi_{II}.$$

Since the Markov chain is ergodic,

$$\lim_{N \to \infty} a(N) = \Pi_{II}. \qquad \text{Q.E.D.}$$

3.3 AFI(M) formula. We now show that contour integration is logically equivalent to the inverse Z-transform for the complex functions considered here. In the rest of this paper, we use the following abbreviation:

$$\frac{1}{2\pi i} \oint_{\Gamma(R)} \triangleq \int',$$

where $\Gamma(R) = \{z/|z| = R\}$ with R changing depending on the context.

Theorem 1. $\int' (\cdot) z^{j-1} dz$ operates on $\hat{a}(z)$ as a projection operator for $R > 1$.

Proof. Recall the simple fact,

$$\text{(SF)} \qquad \int' \frac{dz}{z^r} = \delta_1(r),$$

which is easily proved upon letting $z = Re^{i\phi}$ and integrating.

By definition 1,

$$\hat{a}(z) = \sum_{k=0}^{\infty} \frac{a(k)}{z^k},$$

which converges uniformly for $R > 1$ by the corollary to proposition 2. Since a uniformly convergent series can be integrated term by term and since (SF) holds, we finally have

$$\int' \hat{a}(z)\, z^{j-1}\, dz = \sum_{k=1}^{\infty} a(k) \int' \frac{dz}{z^{(k-j)+1}}$$

$$= \sum_{k=1}^{\infty} a(k)\, \delta_k(j)$$

$$= a(j) \qquad \text{Q.E.D.}$$

Theorem 2. For any R > 1,

$$AFI(N) = 1 - \frac{\nu}{N} \int' \hat{a}(z) \frac{z^N - 1}{z - 1} \, dz.$$

Proof. From chapter 2, we have

$$AFI(N) = 1 - \frac{\nu}{N} \sum_{j=1}^{N} a(j).$$

From theorem 1 and summing a geometric series, we get (for R > 1)

$$AFI(N) = 1 - \frac{\nu}{N} \sum_{j=1}^{N} \int' \hat{a}(z) \, z^{j-1} dz$$

$$= 1 - \frac{\nu}{N} \int' \hat{a}(z) \left(\sum_{j=1}^{N} z^{j-1} \right) dz$$

$$= 1 - \frac{\nu}{N} \int' \hat{a}(z) \left(\frac{z^N - 1}{z - 1} \right) dz. \qquad \text{Q.E.D.}$$

Since $|z| > 1$ in theorem 2, the geometric sum $(z^N - 1)/(z - 1)$ diverges as $N \to \infty$. We must therefore first evaluate the compact expression in theorem 2 (using series expansions, the residue calculus, the Cauchy-Gousat theorem, and the Cauchy integral theorem) and then take the limit of the result to get AFI(∞). In other words,

$$\lim_{N \to \infty} \int' (\cdot) \neq \int' \lim_{N \to \infty} (\cdot).$$

Consequently, our first task is to evaluate the expression in theorem 2.

Theorem 3. (The Main Theorem)

$$\sum_{j=1}^{N} a(j) = \text{Part I} + \text{Part II}.$$

Part I, the principal part, is a compact expression that approaches $\nu \Pi_{II}$ as $N \to \infty$. Part II, the residue, is a finite alternating series whose terms are "filtered" binomial expressions and that moreover, approaches zero as $N \to \infty$.

Proof. Decomposing $\hat{a}(z)$ according to proposition 1, breaking up the geometric sum, and using theorem 2, we have

$$(1)\quad \sum_{j=1}^{N} a(j) = q^I \int' \frac{z^N}{p(z)(z-1)}\,dz + \gamma \int' \frac{z^N}{p(z)(z-1)^2}\,dz$$

$$- \left[q^I \int' \frac{dz}{p(z)(z-1)} + \gamma \int' \frac{dz}{p(z)(z-1)^2}\right].$$

However, for $s = 1$ and 2,

$$\left| \int' \frac{dz}{p(z)(z-1)^s} \right| \leq \frac{R}{(R^I(R-1)-\theta)(R-1)^s}$$

$$\sim \frac{1}{R^{I+s}} \qquad \text{(large R)}$$

$$\to 0, \text{ as } R \to \infty.$$

Therefore, the last two terms in D are zero since R can be taken arbitrarily large. Hence, we now have left in (1):

$$\text{SUM} = q^I\ (q^I - \text{term}) + \gamma\ (\gamma - \text{term}).$$

Our task is thus reduced to evaluating this sum.

For sufficiently large R,

$$\left| \frac{\beta z^I - \theta}{z^{I+1}} \right| < 1 \quad \text{(see A.1, Appendix)}$$

and we can therefore expand $1/p(z)$ by expanding $1 - ((\beta z^I - \theta)/z^{I+1})$ in a geometric series yielding

$$\frac{1}{p(z)} = \sum_{m=0}^{\infty} \frac{(\beta z^I - \theta)^m}{z^{(I+1)(m+1)}} = \sum_{m=0}^{\infty} E_m(z).$$

Using the above expansion in the remaining terms of SUM and letting $J_{1m} = (q^I - \text{term})_m$ and $J_{2m} = (\gamma - \text{term})_m$, we have for $k = 1, 2$

$$\left|J_{km}\right| = \left|\int' E_m(z)\,\frac{z^N}{(z-1)^k}\,dz\right|$$

$$\leq \frac{R^{N+1}\,|\beta R^I - \theta^m|}{R^{(I+1)(m+1)}\,(R-1)}$$

$$\sim \frac{R^{N+1+mI}}{R^{k+(m+1)(I+1)}} \qquad \text{(large R)}$$

$$\to 0, \text{ as } R\to\infty \text{ for}$$

$m > N - (I + k)$. Since a uniformly convergent series can be integrated term by term, we conclude that

J_{km} is nonzero only for $\boxed{m \leq N - (I + k)}$.

Since zero and one are the only poles in the new integrands appearing after the expansion of $1/p(z)$ was made, we can use the Cauchy-Gousat theorem to reduce the contour in the integrals above from $|z| = R > 1$ to $|z| = \varepsilon$ and $|z - 1| = \varepsilon$, thereby getting

$$q^I\text{ - term} = \int'_{|z|=\varepsilon}{}^{(1)} + \int'_{|z-1|=\varepsilon}{}^{(1)} = L_1 + G_1$$

$$\gamma\text{ - term} = \int'_{|z|=\varepsilon}{}^{(2)} + \int'_{|z-1|=\varepsilon}{}^{(2)} = L_2 + G_2 .$$

$L_1 + L_2$ will yield Part II while $G_1 + G_2$ will give Part I. We turn our attention first to evaluating the easier Part I.

Letting $x = (\beta - \theta)$ and integrating the finite series in Part I for both G_1, we get

$$G_1 = \frac{1 - x^{(N-I)}}{1 - x}$$

$$G_2 = \sum_{m=0}^{N-(I+2)} D_z^{(1)}\left[\frac{z^N(\beta z^I - \theta)^m}{z^{(I+1)(m+1)}}\right]\Big|_{z=1}$$

$$= (N - I - 1)\left(\frac{1 - x^{(N-I+1)}}{1 - x}\right) + C, \text{ where}$$

$$C = [I\beta - (I + 1)x]\left[\frac{(1 - x^{(N-I-1)}) - (N - I - 1)(1 - x)x^{(N-I-2)}}{(1 - x)^2}\right].$$

These final results, for Part I, follow directly from summing geometric series (after the integration has been done) with the appearance of the first derivative, $D_z^{(1)}(\cdot)$, following from Cauchy's integral theorem.

Turning now to Part II, we get, after expanding $(\beta z^I - \theta)^m$ and $1/(z-1)^s$ in a binomial series and upon setting $L_k = f'(k)$ with contour "$|z| = \varepsilon$,"

$$L_k = (-1)^k \sum_{s=0}^{\infty} (\delta_2(k)s + 1) \sum_{m=0}^{N-(I+k)} \sum_{r=0}^{m} Q(m,r,s)$$

where

$$Q(m,r,s) = \binom{m}{r} (-\theta)^r \beta^{(m-r)} \int'_{|z|=\varepsilon} z^{ex} dz,$$

where

$$ex = (m - r)I + N - (m + 1)(I + 1) + s.$$

(-1) and (s+1) appear from the geometric series expansion of $1/(z-1)^k$ for k = 1 and 2, respectively; the binomial terms appear as a result of the binomial expansion of $(\beta z^I - \theta)^m$. Once again we have used the theorem on the term-by-term integration of a uniformly convergent series.

To evaluate $f' z^{(ex)} dz$ (contour: $|z| = \varepsilon$), we recall (SF) in the first equation of the proof of theorem 1. Thus, we must now consider triples of integers, (m,r,s), such that ex = -1. Simplifying this equation, we get the following indicial equation to solve; the terms in Part II are nonzero iff

$$N - (m + I)s = rI, \text{ for all triples } (m,r,s)$$

subject to inherent natural constraints elucidated below. Because r appears in the binomial coefficient dominated by m, we use the indicial equation to get for fixed but arbitrary s:

$$r(s) \leq m \Leftrightarrow \boxed{\left\lceil \frac{N - (I - s)}{(I + 1)} - 1 \right\rceil < \operatorname{Max}_m (r(s))}$$

$$\boxed{\leq \left\lceil \frac{N - (I - s)}{I + 1} \right\rceil}.$$

Again using the indicial equation and the bounds already found for m, we further get for the L_k's (k = 1,2):

$$r \leq m \leq (N - (I + k)) \Leftrightarrow \boxed{s \leq (N - I)I - k(I + 1)}$$

$$N - (r(s,k) + 1)\, I + s \leq N - (I + k) \Leftrightarrow \boxed{\operatorname{Min}_m (r(s,k)) \geq \frac{s + k}{I}}.$$

Consequently, we can conclude, in particular, that the "filtered" binomial coefficient is of the form

$$\binom{N - (r(s) + 1)\, I + s}{r(s)} .$$

Noting the form for the L_k and the results on the bounds for the integral variables r, s, and m, we can write the final results for Part II.

$$q^I L_1 = - q^I \left[\sum_{s=0}^{s_1} \sum_{r=r_{11}(s)}^{r_2(s)} T(s) \right]$$

$$\gamma\, L_2 = \gamma \left[\sum_{s=0}^{s_2} (s + 1) \sum_{r=r_{12}(s)}^{r_2(s)} T(s) \right] ,$$

where $s_k = (N - I)I - k\,(I + 1)$, $r_{1k}(s) = \min_m(r(s,k))$, $r_2(s) = \max_m(r(s))$, and

$$T(s) = \binom{N - (r(s) + 1)I + s}{r(s)} (-\theta)^{r(s)} \beta^{N-(r(s)+1)I+s} . \qquad \text{Q.E.D.}$$

In summary, we have the following synopsis of theorem 3 ($x = \beta - \theta$):

$$AFI(N) = 1 - \frac{\nu}{N} \left\{ q^I (G_1 + L_1) + \gamma (G_2 + L_2) \right\}$$

$$= 1 - \frac{\nu}{N} \left\{ R + \text{Part II} \right\} ,$$

$$R = q^I \left(\frac{1 - x^{(N-I)}}{1 - x} \right) + \gamma \left[(N - I - 1) \left(\frac{1 - x^{(N-I-1)}}{(1 - x)} \right) + C \right]$$

$$C = [I\beta - (I + 1)x] \left[\frac{(1 - x^{(N-I-1)}) - (N - I - 1)(1 - x)x^{(N-I-2)}}{(1 - x)^2} \right]$$

In conclusion to this, the main section, examination of the summary to theorem 3 shows that the Z-transform method applied to AFI(N) is logically equivalent to an extensive rearrangement of the finite sum $\Sigma a(j)$. Hence this paper could have been entitled in part: A Rearrangement Theorem for AFI(N)...."

3.4 Convergence properties and approximations. We have seen that the AFI(N) expansion is made up of essentially two different parts: a compact (easily programmable) Part I and a messier Part II consisting of a fairly large finite double sum of binomial terms each of which being "filtered" or selected from its own binomial expansion by the solution triples to the indicial equation. However, the difficulty of handling such an iterated sum for computational use is only apparent for the practical magnitudes of the parameters involved. Stated briefly, for N large relative to I in the sense that $N \sim (10)(I)$ or larger, and, in particular, for $N >> I$, only the first few terms in Part II need to be retained. The essential reason for this is that Part II is, in reality, an asymptotic "factorial" series when considered as a function of N. Turning now to a computational routine on a digital computer, it suffices to calculate the terms $(T(2s) + T(2s + 1)) = \hat{T}(s)$ sequentially at the maximum value of $\beta(q;f, I)$, which is easily found, by the derivative test, to be

$$q(\text{Critical}) = I/(I + 1).$$

For this value of q, a cut-off routine can then be inserted to truncate the double sum as soon as $\hat{T}(s)$ is less than some predetermined small number.

Keeping the above discussion concerning the computational feasibility of Part II in mind, we now prove some theoretical results on the asymptotic behavior of Part II, as well as the whole AFI(N) formula, as $N \to \infty$.

Proposition 2. For $k = 1, 2$, we have

$$(s_k + 1) \sum T(s_k) \leq [(\delta_2(k)\, N + 1)I + O\left(\tfrac{1}{N}\right)].$$

$$\left[\sum_{r=0}^{N+O(\frac{1}{N})} \left(\binom{N(I+1)}{r} + O\left(\tfrac{1}{N}\right)\right)(-1)^{N(I+1)-r}\, \theta^r\, \beta^{N(I+1)} + O\left(\tfrac{1}{N}\right)\right]$$

as $N \to \infty$ through a sequence of values such that $N(I + 1)$ is even (with a similar expression for $\Sigma T(0)$).

Proof. Factor out the leading terms appearing above from the expressions occurring in the summary to Main Theorem. The proof is finished by recalling the definition of $O(\cdot)$ [7.8, chp. 1]. Q.E.D.

Directly from this proposition, we obtain (k = 1, 2)

$$(D)\quad \frac{1}{N}\sum_{s=0}^{s_k}(\delta_2(k)s+1)\sum T(s)$$

$$\leq (\delta_2(k)NI+1)(I+O(\tfrac{1}{N^2}))\,\beta^{N(I+1)+O(\frac{1}{N})}(1-\theta)^{N+O(\frac{1}{N})}$$

which approaches zero as $N\to\infty$. The analysis that resulted in (D) can be refined in a number of ways (e.g., use of the LaGrange expansion; see A.2, Appendix) to obtain approximations to Part II for $N \gg I$.

Combining (D) with Part I from Main Theorem, we can prove the theorem below.

Theorem 4. As $N\to\infty$, $AFI(N) \to AFI(\infty)$.

Proof. In the summary to theorem 3, the expression blocked out and emphasized with an arrow in Part I is the only portion of the whole formula which does not vanish as $N\to\infty$; indeed, examination will reveal that it approaches $1 - \nu\alpha(\infty) = 1 - \Pi_{sN} = AFI(\infty)$. The proof is finished by (D). Q.E.D.

A more direct proof of convergence to $AFI(\infty)$ can be obtained for the (unfortunately infrequent) case that $\theta + \beta < 1$. First we prove

Lemma. If $\theta + \beta < 1$, then there exists an $R < 1$ such that $R_0^I(R_0 - \beta) > \theta$ for all $R_0 > R$. Furthermore, $\beta < R = \theta^{1/I+s}$ for some positive integer s implicitly defined by

$$\frac{Ln(a-\beta)}{Ln(a)} > s,\quad a = \theta^{1/I+s}.$$

Proof. Since $\theta^{1/n} = e^{1/n(Ln\theta)}$ and $0 < \theta < 1$, we have

$$(1)\ \lim_{n\to\infty}\theta^{1/n} = 1,\quad (2)\ \frac{1}{n}Ln\theta < \frac{1}{n+1}Ln\theta,\ \text{and}\ (3)\ \lim_{s\to\infty}\theta^{s/(I+s)} = \theta.$$

Assume now that $R = \theta^{1/(I+s)}$ for some s. Then,

$$R^I(R - \beta) > \theta \ddagger \theta^{I/(I+s)} (\theta^{1/(I+s)} - \beta) > \theta$$

$$\ddagger (\theta^{1/(I+s)} - \beta) > \theta^{s/(I+s)}.$$

Taking Ln of both sides of the last inequality yields the defining inequality for s. Finally assume that such an s does not exist. Again, from the last inequality, we get

$$(4)\ s \text{ does not exist } \ddagger (\theta^{1/(I+s)} - \beta) \leq \theta^{s/(I+s)} \text{ for all } s.$$

Using (1), (2), and (3) in the R.H.S. of (4) and taking the limits give us $1 - \beta \leq \theta$, which is contrary to assumption. Q.E.D.

Theorem 5. If $\theta + \beta < 1$, then there exists an $R < 1$ such that

$$|AFI(\infty) - AFI(N)| \leq \frac{1}{N} \text{ constant } (R^{N+1}).$$

Proof. That such an R exists follows from the lemma. Taking the sum part of AFI(N) and an initial contour of radius $R_1 > 1$, we can immediately reduce the integrands as in (1) of theorem 3. Furthermore, because the existence of R implies that the roots of p(z) lie within $z = R$, we can also reduce $|z| = R_1$ to $|z| = R$ and $|z - 1| = \varepsilon$ $(1 - \varepsilon > R)$ since the poles of the integrands are one and the roots of p(z). Thus we have

$\sum_{j=1}^{N} a(j) = q^I [J_1 + J_3] + \gamma[J_2 + J_4]$, where the contour of $J_1(J_2)$ is $|z| = R$ and that of $J_3(J_4)$ is $|z - 1| = \varepsilon$. Now,

$$|J_k| \leq \frac{R^{N+1}}{|R^I|R - \beta| - \theta|(1 - R)^k} \quad (k = 1,2),$$

$$J_1 = \frac{1}{p(1)}, \text{ and } J_2 = \frac{N}{p(1)} - \frac{p'(1)}{p(1)}.$$

The remainder of the proof follows from trivial manipulation. Q.E.D.

In conclusion, it is hoped that the need for an expansion of AFI(N), demonstrated in chapter 2, has now been met.

4.0 VARIANCE. The variance of AFI(N) is not zero (see below). In order to utilize CSP(N) to effect savings over that of CSP(∞), it is necessary now to investigate this subject.

4.1 Calculation of $P^{k}_{HO,SN}$. From the Chapman-Kolmogorov equation [7.5, p. 194] or directly in terms of $\alpha(k)$ [7.1, p. 950], we obtain the following equations for this transitional probability:

$$P^{k}_{HO,SN} = \nu\alpha(k+1) = \omega(k+1) = \omega_1(k)$$

$$(A)\quad \hat{\omega}(z) = \nu\hat{\alpha}(z) = \frac{\hat{\omega}_1(z)}{z}\,.$$

4.2 Calculation of $P^{m}_{SN,SN}$. Since there is not any direct, simple relationship with $\alpha(\cdot)$ for this case [7.1, pp. 951-952], another application of the Chapman-Kolmogorov equation yields, for this expression, the following two simultaneous equations:

$$P^{m}_{SN,SN} = P_{SN,SN}\,P^{m-1}_{SN,SN} + P_{SN,SI}\,P^{m-1}_{SI,SN}$$

$$P^{k}_{SI,SN} = P_{SI,SN}\,P^{k-1}_{SN,SN} + P_{SI,SI}\,P^{k-1}_{SI,SN} + P_{SI,HO}\,P^{k-1}_{HO,SN}.$$

Letting $\phi(m) = P^{m}_{SN,SN}$ and $\psi(k) = P^{k}_{SI,SN}$, the corresponding difference equations, with the initial conditions written explicitly, would be

$$\phi(m) = \nu(\phi * \delta_1)(m) + f\,(\psi * \delta_1)(m) + \nu\delta_1(m)$$

and

$$\psi(k) = \nu q(\phi * \delta_1)(k) + fq\,(\psi * \delta_1)(k) + p\nu\omega(k) + \nu q\,\delta_1(k).$$

Applying the Z-transform to the above two equations yields

$$\hat{\phi}(z) = \nu\,\frac{\hat{\phi}(z)}{z} + f\,\frac{\hat{\psi}(z)}{z} + \frac{\nu}{z}$$

$$\hat{\psi}(z) = q\left[\nu\hat{\phi}\,\frac{(z)}{z} + f\,\frac{\hat{\psi}(z)}{z}\right] + p\nu\hat{\omega}(z) + \frac{\nu q}{z}.$$

Since $\hat{\omega}(z)$ is known from 4.1, we therefore have two simultaneous equations in two unknowns of the form

$$(a-1)\,x + (b)y = -c$$

$$(qa)\,x + (qb-1)\,y = -(qe+d)$$

whose determinant is (-1). Thus, for the unique solution, $\hat{\phi}(z)$, we finally obtain

$$(B) \quad \hat{\phi}(z) = \nu(1 - \beta) \frac{\hat{\alpha}(z)}{(z - \beta)} + \frac{\nu}{(z - \beta)} .$$

As a simple check on this derivation, we calculate the limit below (remembering that $\hat{\alpha}(z) = \nu\hat{a}(z)$).

$$\lim_{z \to 1} \left(\frac{z - 1}{z}\right) \hat{\phi}(z) = \nu a(\infty) = \Pi_{SN} = \lim_{m \to \infty} \phi(m)$$

which is correct since the Markov chain is ergodic.

As a corollary to calculating the variance, we can now calculate $P^k_{SN,SN}$ for any k (and, for that matter, $P^k_{SI,SN}$ from $\psi(k)$), which is done by the sequence of equations below.

$$P^k_{SN,SN} = \int^{1} \hat{\phi}(z) z^{k-1} dz$$

$$= \nu(1 - \beta) \sum_{j=0}^{k-(I+2)} \beta^j \int^{1} \hat{a}(z)\, z^{(k-(j+1))-1} dz$$

$$+ \nu \int^{1} \frac{z^{k-1}}{(z - \beta)} \, dz$$

$$= \nu(1 - \beta) \sum_{j=0}^{k-(I+2)} \beta^j a(k - (j + 1)) + \nu\beta^{k-1}$$

$$(C) \qquad = (1 - \beta) \sum_{j=0}^{k-(I+2)} \beta^j P^{k-(j+2)}_{HO,SN} + \nu\beta^{k-1}$$

The above calculation illustrates the generality of the Z-transform method as well as being useful for the variance calculations to be given in the next section.

4.3 Calculation of variances. From the simple identity VAR (1 - cX) $= c^2 VAR(X)$, we easily derive the following equations:

$$(D)\quad VAR_{\iota}(FI(N)) = \frac{1}{N}\left(\sum \frac{a_j}{N}\right) - \left(\sum \frac{a_j}{N}\right)^2 + T(N)\ (j = 1,N),$$

$$= R_1 - S_1 + T(N),$$

and

$$(E)\quad VAR_{\Pi}(FI(N)) = \frac{1}{N}(\Pi_{SN}) - (\Pi_{SN})^2 + T(N)$$

$$= R_2 - S_2 + T(N)$$

where

$$T(N) = 2\,\frac{\Sigma(a * b)(j)}{N^2}\quad (j = 0,\ N - 1),$$

$$a_j = P^j_{HO,SN},\ \text{and}\ b_k = P^k_{SN,SN}.$$

Using (A) and (B), we can now calculate the above variances by evaluating T(N) directly. However, since, as $N\to\infty$, both variances approach zero (see proposition below); R_1, $R_2 \to 0$; and $S_1 \to \Pi^2_{SN}$; we must have: $T(N)\to\Pi^2_{SN}$. Hence, we can approximate this last term, T(N), by

$$\left(\sum \frac{a_j}{N}\right)\left(\sum \frac{b_j}{N}\right)\quad (j = 0,\ N - 1).$$

Proposition. As $N\to\infty$, $VAR_x(IF(N))$ approaches zero for $x = \iota$ and Π, respectively.

Proof. Letting $E[\cdot]$ stand for either $\hat{E}_{\iota}[\cdot]$ or $E_{\Pi}[\cdot]$, the ergodic theorem implies

$$E\left[\,|FI(N) - E[FI(N)]|\,\right]\to 0,\ \text{as}\ N\to\infty.$$

Since $|FI(N) - E[FI(N)]| \le 1$ and the chain induced probability measure, P, is necessarily finite, we have

$$E\left[\,|FI(N) - E(FI(N))|^2\,\right] \le E\left[\,|FI(N) - E[FI(N)]|\,\right]$$

where we recall the definition of $E[\cdot]$ in terms of Ω and P as being

$$E[X] = \int_{\omega\in\Omega} X(\omega)\,dP(\omega),\ \Omega = \text{outcome space.}\qquad \text{Q.E.D.}$$

The practical problem of estimating the variances given in (D) and (E) has now been reduced to the calculation of the expression $\Sigma b_j/N$--a calculation similar to that already done in (C). However, a straight-forward approximation can also be obtained by the following considerations. From (C), noting that $a(j) = 0$ for $j < I$, we have

$$\frac{\Sigma\, b_k}{N} = c_1 \frac{\Sigma\,(\alpha * \beta)(k)}{N} + c_2 \frac{\Sigma\,(\beta * \delta_1)(k)}{N}$$

where $\beta(k) = \beta^k$ and $k = 0, (N-1)$. For $\underline{x} = (x_1, x_2, \ldots x_N) \in R^N$, define the $\ell^1(N)$ norm of $\underline{x}$ by $||\underline{x}|| = \Sigma|x_i|$ $(i = 1,N)$. Using this definition, the equation preceding it, and multiplying through by N, we conclude

$$||\underline{b}|| \leq c_1 ||\,(\underline{\alpha} * \underline{\beta})|| + c_2\, ||(\underline{\beta} * \underline{\delta})||.$$

Recalling the $\ell^1(N)$ inequality for the convolution, $||\underline{x} * \underline{y}|| \leq ||\underline{x}||\,||\underline{y}||$, we therefore have

$$||\underline{b}|| \leq c_1\, ||\underline{\alpha}||\,||\underline{\beta}|| + c_2\, ||\underline{\beta}||\,||\underline{\delta}||.$$

Dividing the above inequality through by N and defining $a(\cdot)$ in terms of $\alpha(\cdot)$, we finally obtain

$$(F)\qquad \left(\frac{\sum_{k=1}^{N-1} b_k}{N}\right) \leq \nu(1-\beta)\left(\frac{\sum_{k=1}^{N} a_k}{N}\right)\left(\frac{1-\beta^{N+1}}{1-\beta}\right) + \left(\frac{\nu}{N}\right)\left(\frac{1-\beta^{N+1}}{1-\beta}\right)(1).$$

Rewriting (F) in terms of AFI(N), we get

$$(G)\qquad \left(\frac{\sum_{k=1}^{N-1} P^k_{SN,SN}}{N}\right) \leq (1 - AFI(N))(1-\beta^{N+1}) + \left(\frac{\nu}{N}\right)\left(\frac{1-\beta^{N+1}}{1-\beta}\right).$$

Using (G), the approximation to T(N), (A), (B), and rewriting all other expressions in terms of AFI(N), we obtain the final expression for upper bounds to (A) and (B).

$$\begin{aligned} VAR_L(FI(N)) &\leq \tfrac{1}{N}\left(1 + \tfrac{1}{N}\right)(1 - AFI(N+1)) \\ &\quad - \left(1 + \tfrac{1}{N}\right)^2 (1 - AFI(N+1))^2 \\ &\quad + (1 - (AFI(N)))\left[1 - AFI(N)(1-\beta^{N+1}) + \frac{\nu}{N}\left(\frac{1-\beta^{N+1}}{1-\beta}\right)\right] \end{aligned}$$

and

$$\mathrm{VAR}_{\Pi}(\mathrm{FI}(N)) \leq \frac{1}{N}\Pi_{SN} - \Pi_{SN}^{2} + (1 - \mathrm{AFI}(N))^{2}\,(1 - \beta^{N+1})$$

$$+ \frac{\nu}{N}\,(1 - \mathrm{AFI}(N))\,\left(\frac{1 - \beta^{N+1}}{1 - \beta}\right).$$

In concluding this section, it should be noticed that both the upper bounds above approach zero as $N \to \infty$.

5.0 PROBABILITY BOUNDS. In this section, use will be made of some elementary Martingale theory [7.6, pp. 209-215; 7.7, 7.9] to prove the following two statements:

(A) $$P\left[(\mathrm{FI}(N) - E_{\ell}[\mathrm{FI}(N)]) > a\right] \leq \frac{1}{a^2}\,\mathrm{VAR}_{\ell}\,(\mathrm{FI}(N))$$

(B) $$P\left[|\mathrm{FI}(N) - E_{\Pi}[\mathrm{FI}(N)]| > a\right] \leq \frac{1}{a^2}\,\mathrm{VAR}_{\Pi}(\mathrm{FI}(N)).$$

5.1 Martingales. We start by giving the following definition for a sequence of random variables.

Definition. A sequence of random variables, $\{V_N\}$, is a (super ..., sub)-Martingale iff

$$E[V_N | V_1, \ldots, V_{N-k}]\ (\leq, =, \geq)\ V_{N-k}\ [a.e],\ \text{for } 1 \leq k \leq N - 1.$$

Now let $V_k = X_{SN}(k)$ and $U_N = \Sigma\, V_k/N$, $k = 1, N$. Then we have

Proposition. For $E[\,\cdot\,] = E_{\Pi}[\,\cdot\,]$, $\{U_N\}$ is a Martingale

Proof. By the Markov property, we have

$$E[U_k | U_1, \ldots, U_{\ell}] = E[U_k | U_{\ell}].$$

Splitting U_k up, we obtain

$$(C)\quad E[U_k | U_{\ell}] = \frac{\ell}{k}\left(\frac{1}{\ell}\sum_{j=1}^{\ell} V_j\right) + E\left[\frac{1}{k}\sum_{\ell+1}^{k} V_j \,\middle|\, U_{\ell}\right]$$

$$= \frac{\ell}{k}\,U_{\ell} + E[\sigma | U_{\ell}].$$

Letting $\alpha = E[V_j]$, for any j (since we are dealing with the arbitrary-entry case) and using the definition of conditional expectation, we have

$$E\left[E[\sigma|U_\ell]\right] = \frac{1}{k} E\left[\sum_{\ell+1}^{k} V_j\right] = \left(\frac{k-\ell}{k}\right) \alpha.$$

On the other hand, we also have

$$E\left[\left(\frac{k-\ell}{k} U_\ell\right)\right] = \left(\frac{k-\ell}{k}\right) E[U_\ell] = \frac{k-\ell}{k} \alpha.$$

Comparing the R.H.S.'s of the last two equations, we conclude that the L.H.S.'s are equal:

$$E\left[E[\sigma|U_\ell]\right] = E\left[\left(\frac{k-\ell}{k} U_\ell\right)\right].$$

Since both of the integrands in the last equation are non-negative almost everywhere ([a.e.]), we can conclude that

$$E[\sigma|U_\ell] = \left(\frac{k-\ell}{k}\right) U_\ell \quad [a.e.].$$

Comparing this equation with (C), we finally get

$$E[U_k|U_1,\ldots,U_\ell] = \frac{\ell}{k} U_\ell + \frac{k-\ell}{k} U_\ell \quad [a.e.],$$

$$= U_\ell \quad [a.e.].$$

From this proposition, we can derive two corollaries.

Corollary 1. $\{|U_k - \alpha|\}$ is a Submartingale.

Proof. From the proposition above, we clearly have that $\{(U_k - \alpha)\}$ is a Martingale. Using Jensen's inequality, we have

$$E[|U_k - \alpha| \,|\, U_\ell] \geq \left|E[(U_k - \alpha) \,|\, U_\ell]\right|$$

$$= |U_\ell - \alpha|.$$

The equality follows from the first statement. Q.E.D.

Corollary 2. $\{|FI(N) - E(FI(N))|\}$ is a Submartingale, where again, $E[\,\cdot\,] = E_\Pi[\,\cdot\,]$.

Proof. Let $\alpha = E[(FI(N))]$, then

$$FI(N) - \alpha = (1 - U_N) - (1 - \alpha) = U_N - \alpha).$$

We know that $\{|U_N - \alpha|\}$ is a Submartingale from corollary 1. Q.E.D.

In a similar manner, using the fact that $E\left[E[U_k|U_\ell]\right] \geq E[U_\ell]$ for $E = E_\ell$, we can show that $\{(FI(N) - E_\ell[FI(N)])\}$ is a Submartingale; here, however, there are no absolute value signs!

5.2 Kolmogorov's inequality. The next result needed is Kolmogorov's inequality for Submartingales [see 7.6, pp. 241-246, for an excellent proof], which is given here as a theorem.

Theorem. (Kolmogorov's Inequality) given that $\{W_k\}$ is a Submartingale and $\phi(x)$ a convex, monotonically increasing function, we have

$$p[\underset{1\leq k\leq N}{\text{Max}} \{W_k\} > a] \leq \frac{1}{\phi(a)} E[\phi(W_N)]. \qquad \text{Q.E.D.}$$

The demonstrations of the inequalities (A) and (B) now follow from corollary 2, the comment following its proof, Kolmogorov's inequality on setting $\phi(x) = x^2$, and the set inclusion

$$\{\omega|W_N(\omega) > a\} \subset \{\omega|\underset{1\leq k\leq N}{\text{Max}} \{W_k(\omega) > a\},$$

where we can abbreviate such sets as

$$[W_N > a] \subset [\text{Max } \{W_k\} > a].$$

6.0 FURTHER APPLICATIONS OF THE Z-TRANSFORM. Two examples, indirectly related to AFI(N), and a discussion of the transform method applied to more complex sampling plans make up this concluding chapter.

6.1 Examples. The first example, part of which is to be used later, deals with the derivation of the mean time of sojourn in the inspection phase from entrance to exit (to phase II). It suffices to consider a modified Markov chain with all of phase II as a new absorption state---called A. Then, the problem becomes: calculate $E_\ell[\hat{I}]$ where the "^" symbol makes note of the modified chain. From the C-K equation, we can derive the following equation for calculating $P^k_{HO,A} = a_2(k)$.

$$a_2(k) = \sum_{j=0}^{I-1} pq^j(a_2 * \delta_{j+1})(k) + q^I H_I(k).$$

Using the Z-transform yields

$$\hat{a}_2(k) = \left(\frac{z}{z-1}\right) \frac{q^I(z-q)}{p(z)} \quad , \; f = 0 \text{ in } p(z).$$

Performing the simple end-point test,

$$\lim_{z \to 1} \left(\frac{z-1}{z}\right) \hat{a}_2(z) = 1 = a_2(\infty),$$

as expected. Employing the method of first entrance [7.5, pp. 214-215] produces the following equations;

$$P^k_{HO,A} = (f_{HO,A} * P_{A,A})\ (k) \quad \text{or}$$

$$a_2(k) = (\gamma_1 * \gamma_2)\ (k), \text{ where}$$

$\gamma_1(k)$ is the probability of first entrance to A on step k. Using the Z-transform on the last equation yields

$$\hat{\gamma}_1(z) = \frac{\hat{a}_2(z)}{\hat{\gamma}_2(z)} = \frac{q^I(z-q)}{p(z)} .$$

By definition of first entrance probabilities,

$$E_L[\hat{I}] = \sum_{k=1}^{\infty} k\, \gamma_1(k).$$

Using the formula for $\gamma_1(z)$, we finally have

$$E_L[\hat{I}] = \lim_{N \to \infty} \int \hat{\gamma}_1(z)\, D_z \left(\frac{z^N - 1}{z - 1}\right) dz$$

which can be evaluated as in chapter 3.

The second example, similar to the one above, concerns the calculation of $E_{SN}[\hat{II}]$ where again a (new) modified Markov chain is used with phase I now serving the role of the absorption state--called B. For this example, the C-K equation gives

$$P^{k}_{SN,B} = P_{SN,SN}\, P^{k-1}_{SN,B} + P_{SN,SI}\, P^{k-1}_{SI,B} \quad \text{and}$$

$$P^{k}_{SI,B} = P_{SI,SI}\, P^{k-1}_{SI,B} + P_{SI,SN}\, P^{k-1}_{SN,B} + fp\, H_1(k).$$

Letting $\alpha_3(k) = P^{k}_{SN,B}$, using the Z-transform on the above simultaneous equations, solving for $\hat{\alpha}_3(z)$, and transforming back yields

$$\hat{\alpha}_3(z) = \frac{(fp)z}{z(z-1)(z-\beta)} \quad \text{and}$$

$$\left((1 - \frac{\nu}{1-fq}\ \delta_1) * \alpha_3\right)(k) = \frac{f^2 p}{1-fq}\ H_1(k).$$

End-point analysis again gives $\alpha_3(\infty) = 1$. One could then proceed as in the previous example.

We now use parts of the examples in the last topic to be treated in this chapter.

6.2 Higher level sampling plans. We will present below a possible methodology for analyzing higher level plans through a systematic use of "operator-box" analysis. First, it is convenient to give a short outline of operator-box notations (also called block diagrams in control theory) which appears in Figure 3.

FIGURE 3.

OPERATOR NOTATION

(1) SUM

$L_1 + L_2$

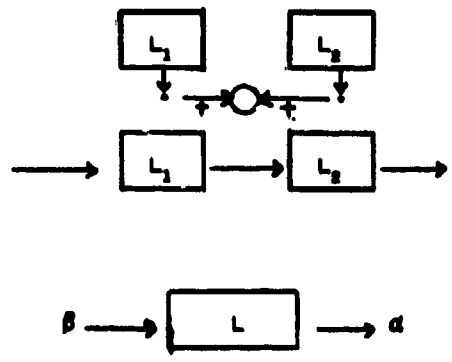
(2) COMPOSITION

$L_2 \circ L_1$

(3) INPUT—OUTPUT

$L(\alpha) = \beta$

"OPERATOR—BOX" SYMBOLS



With reference to diagram 3 in Figure 3, we can apply the Z-transform to the operator and both sequences to get

$$\hat{L}(z)\,\hat{\alpha}(z) = \hat{\beta}(z) \text{ or } \hat{\alpha}(z) = \left(\frac{1}{\hat{L}(z)}\right)\hat{\beta}(z) = \hat{f}(z)\,\hat{\beta}(z).$$

This is precisely the formalized version of what was previously done in this paper. Going one step further, remembering the relationship between $\wedge$ and *, we can write

$$\hat{\alpha}(z) = \hat{f}(z)\,\hat{\beta}(z) \overset{*}{\div} \alpha(k) = (f * \beta)(k) \quad (\hat{f} = \hat{f}).$$

We can thus write the diagram in the more convenient way:

$$\beta \longrightarrow \boxed{\quad f \quad} \longrightarrow \alpha.$$

With these preliminaries out of the way, we can now put forth our tentative procedure for handling any type of sampling plan. A starting point for research would be CSP(N). Concerning phases I and II, we have the following information.

(1) From chapter 3, we have for CSP(N)

$$L_1\,\alpha_1 = \beta_1 = \nu\,(q^I\delta_I + \gamma\,H_{I+1}).$$

(2) From example 1, we have for phase I alone

$$L_2\,\alpha_2 = \beta_2 = q^I\,H_I.$$

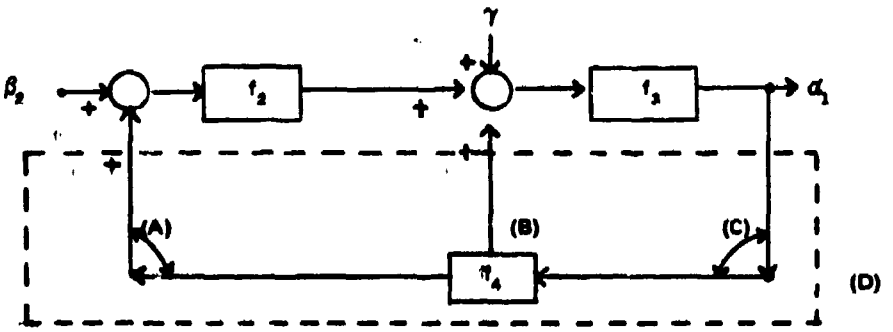
(3) From example 2, we have for phase II alone

$$L_3\,\alpha_3 = \beta_3 = \left(\frac{f^2 p}{1 - fq}\right) H_1.$$

In the above, $L_1 = 1 - \beta\delta_1 + \theta\delta_{I+1}$, $L_2 = \sum_{0}^{I-1} pq^j\,\delta_{j+1}$, and $L_3 = 1 - c\delta_1$, $c = \nu/(1 - fq)$. Given α_1 as known, we now try to find new inputs and/or new operators that can be combined with portions (2) and (3) above to yield a diagram equivalent to that of (1). With this procedure in mind, the following three types of arrangements strongly suggest themselves for investigation (using the more manageable f * notation); refer to Figure 4 for the symbols used.

FIGURE 4.

PHASE I AND PHASE II $\overset{+}{+}$ CSP(N)



(1) $\gamma = 0$,

(a) $f_3 * (f_2 * (f_4 * \alpha_1 + \beta_2) + f_4 * \alpha_1) = \alpha_1$

(b) $\hat{f}_4 = (\hat{\alpha}_1 - \hat{f}_2 \hat{f}_3 \hat{\beta}_2)/(1 + \hat{f}_2) \hat{f}_3 \hat{\alpha}_1$

(c) $\overset{\bullet\bullet}{f}_4(z) = 1/\hat{L}_4(z)$.

(2) $\gamma = \beta_3$, leaving out (B).

(a) $f_3 * (f_2 * (\beta_2 + f_4 * \alpha_1) + \beta_3) = \alpha_1$

(b) $\hat{f}_4 = \hat{\alpha}_1 - \hat{f}_3 (\hat{f}_2\hat{\beta}_2 + \hat{\beta}_3)/\hat{f}_2\hat{f}_3\hat{d}_1$

(c) Same as (1).

(3) γ = unknown input, leaving out (D).

(a) $f_3 * (f_2 * \beta_2 + \gamma) = \alpha_1$

(b) $\hat{\gamma} = (\hat{\alpha} - \hat{f}_2\hat{f}_3\hat{\beta}_2)/\hat{f}_3$

(c) $\gamma(k) = \int' \hat{\gamma}(z)\, z^{k-1}\, dz$.

The results gleaned from the foregoing investigations, concerning the forms that f_4 and γ take, could then be used in the more complex plans to solve for the final unknown output.

Thus for a sampling plan with one clearance stage, (r - 1) limited sampling stages, one unlimited sampling stage, and (r) checking phases, the procedure envisaged would be (1) analyze each phase (or stage) by itself and, taking account of its place in the plan, derive a formula for either $P^{k}_{HO,SN1}$, $P^{k}_{SNj,A}$, or $P_{CK\ell,A}$(A = all possible entry phases from the phase under consideration); (2) the outcome of this analysis should result in an f* for each phase; and (3) put the "operator-box" symbols together with the appropriate additions and solve for the unknown output.

In conclusion, use of operator diagrams further suggests two additional areas of inquiry. The first is the possible mechanization by electronic circuits of a given plan since all operators involved can eventually be represented in terms of the fundamental δ_j's. The second area is the possible development of a relationship between the responsiveness of a plan and the stability of a discrete linear control system.

APPENDIX

A.1 Expansion of $1/p(z)$. In the proof of theorem 3, we expanded $1/p(z)$ by using the binomial series for $1/(1 - w)$, $w = (\beta z^I - \theta)/z^{I+1}$. However, two other expansions are also possible but result in integrals that are hard to evaluate. The results appear in (1) and (2) below.

$$(1)\quad \sum_{m=0}^{\infty} \frac{(\beta z^I)^m}{(z^{I+1} + \theta)^{m+1}} \longrightarrow \text{EVAL: } \int' \frac{(\beta z^I)^m \, dz}{\prod_k (z - z_k)(z - 1)^t}$$

for $t = 1, 2$; k varying from 1 to $I + 1$; and z_k = $(I + 1)$st root of θ.

$$(2)\quad \sum_{m=0}^{\infty} \frac{(-\theta)^m}{(z^I(z - \beta))^{m+1}} \longrightarrow \text{EVAL: } \int' \frac{dz}{(z^I(z - \beta))^{m+1} (z - 1)^t}$$

for $t = 1, 2$. Both evaluations require higher order derivatives by Cauchy's integral theorem. Nevertheless, a connection does exist between the above expansions and the one used in theorem 3.3. This connection is provided by Newton's formula for the derivatives of the product of two functions:

$$D^{(k)} (u(z)v(z)) = \sum_{j=0}^{k} \binom{k}{j} u^{(k-j)}(z) \, v^{(j)}(z).$$

A.2 The LaGrange Expansion. [7.10, p. 6] In this section, we give an alternative to the upper bound for $\Sigma T(s)$ in chapter 3 by using the LaGrange Expansion Theorem. First of all, it clearly follows that

$$(A)\quad \sum_{r=r_1}^{r_2} T(s) \le \sum_{r=0}^{\infty} T(s).$$

By regrouping the factors on the R.H.S. of (A) and using the elementary fact that

$$\frac{1}{k!} D_y^{(k)} \left(y^{N-(k+1)I+s} \right) \Big|_{y=1} = \binom{N-(k+1)I+s}{k},$$

we can rewrite the R.H.S. of (A) as

$$(B)\quad 1 + \sum_{k=1}^{\infty} \frac{(-\theta)^k}{k!} D_y^{(k)} \left[(\beta y)^{N-I+s} \left((\beta y)^{-I} \right)^k \right] \Big|_{y=1}.$$

In order to use the LaGrange Expansion to obtain a closed expression for (B), we must first consider

$$(C)\quad D_y \left(f(y) + g(y) \right), \text{ where}$$

$$f(y) = y - \beta^{N-I+s} h(y),$$

$$g(y) = \beta^{N-I+s} h(y) + \sum_{k=1}^{\infty} \frac{(-\theta)^k}{k!} D_y^{(k-1)} \left[\beta^{N-1+s} D_y h(y) \left((\beta y)^{-I} \right)^k \right]$$

$$\text{and} \qquad h(y) = \frac{y^{(N-I+1+s)}}{(N-I+1+s)} .$$

However, (C) = (B') where the prime denotes "no evaluation of 1." We now rewrite (C) as Dy(g(y)) + Dy(f(y)); from LaGrange's Theorem,

$$D_y(f(y)) = D_y \left[\frac{(r(y))^{N-I+s+1}}{N - I + s + 1} \right]$$

where r(y) is the unique solution to

$$(D)\quad r(y) = y - \frac{\theta}{(\beta r(y))^I}$$

provided that, by Rouche's Theorem,

$$(D')\quad \left| \frac{\theta}{(\omega(y)\beta)^I} \right| < |\omega(y) - y|$$

for any ω on some contour $|z - y| = b$, $b > 0$. Once more applying the Expansion Theorem, this time to (D), we get

$$(E)\quad r(y) = y - \frac{\theta}{(\beta y)^I} \sum_{k=0}^{\infty} \binom{-I}{k} \left(\frac{-\theta}{y} \right)^k$$

$$= y - \frac{\theta}{(\beta y)^I} \left(\frac{1}{(1 - (\theta/y)^I} \right) \text{for } \theta/y < 1.$$

From (E), we consequently have

$$(F)\quad D_y\,(r(y)) = 1 + \frac{\theta I}{(\beta y)^I}\left[\frac{1}{(1-(\theta/y))^I}\left[\frac{1}{y} + \frac{\theta}{y^2(1-(\theta/y))}\right]\right].$$

Thus we have,

$$(C')\quad D_y g(y) + D_y f(y) = 1 - (\beta y)^{N-I+s} + r(y)^{N-I+s}\, D_y(r(y)).$$

Evaluating (C') at 1, we finally have from (A)

$$\boxed{\sum_{r=r_1}^{r_2} T(s) \le 1 - \beta^{N-I+s} + (r(1))^{N-I+s}\, r'(1)}$$

where r(1) and r'(1) are given by (E) and (F), respectively.

To conclude this section, we notice the trivial fact that $r(1) < 1$ from (E). Thus we have another proof that Part II, in Theorem 3.2, converges to zero provided condition (D') holds for the case $y = 1$. However, letting $\omega = 1 + \varepsilon e^{i\phi}$, the validity of (D') follows from

$$(1-\varepsilon)\beta > (\varepsilon/\theta)^I$$

from sufficiently small $\varepsilon < \theta$.

7.0 REFERENCES

7.1 Brugger, R. M., "Functional Properties of CSP-1 Applied to a Finite Length Production Run," Proceedings of the 17th Conference on the Design of Experiments in Army Research Development and Testing, Part 2, pp. 929-953.

7.2 Knopp, K. *Theory of Functions*, Part I, Dover Publications, 1947.

7.3 Kyrala, A. *Applied Functions of a Complex Variable*, Wiley-Interscience, 1972,

7.4 de la Barrière, P. *Optimal Control Theory*, W. B. Saunders Co., 1967.

7.5 Parzen, E. *Stochastic Processes*, Holden-Day, 1967.

7.6 Feller, W. *Introduction to Probability Theory and its Applications*, Vol. II, Wiley, 1971.

7.7 McKean, H. P. *Stochastic Integrals*, Academic Press, 1969.

7.8 Sirovich, L. *Techniques of Asymptotic Analysis*, Springer-Verlag, 1971.

7.9 Loéve, M. *Probability Theory*, 3rd edition, Van Nostrand.

7.10 Bellman, R. *Perturbation Techniques in Mathematics, Physics, and Engineering*, Holt, Rinehart and Winston, Inc.

AUTOMATED PRODUCTION LINE MODELING USING GERTS IIIQ

Martin M. Roffman
Manufacturing Technology Directorate
Frankford Arsenal
Philadelphia, Pennsylvania
19137

and

. Lawrence P. McNamee
Computer Science Department
University of California
Los Angeles, California
90024

ABSTRACT. Applications of discrete event, digital simulation techniques to production line modeling are herein discussed. Six network models are diagrammed for GERTS IIIQ, a language specially designed to resolve stochastic networks such as those exhibited in production line problems. The applications are then generalized to a wider class of stochastic networks characterized by nonnegative, unidirectional branches in which renewal processes (cycles) may occur. GERTS IIIQ symbolic diagrams accompany each application, followed by an explanation of the network and its functional utility. Where necessary, analytical expressions for certain statistical results are derived from conditional probability theory and presented for verification purposes.

INTRODUCTION. During recent systematic analysis of a high speed, automated production facility, several modeling techniques were developed to realistically portray common manufacturing operations (1). Each technique was designed to simplify the task of simulating a production line, component failure, buffer storage, and component transfer activities within a continuous type production/inventory environment. Although it was not readily apparent, these methods can be extended to resolve analogous problems in network applications far removed from the production line. The purpose of this paper is to describe basic procedures which can provide insight into a class of networks containing repetitive or renewal (cyclic) production operations, positive flow capacities, and unidirectionality of flow along each arc.

All of the procedures to be described are modeled using GERTS IIIQ, a discrete event, digital simulation language. GERTS IIIQ, an acronym meaning Graphical Evaluation and Review Technique Queueing Simulation language, is particularly well suited to resolve the discrete scheduling activities that typically arise in production and inventory systems. Stochastic process events, such as the occurrence of a tool failure, for example, can be conveniently controlled by adjusting the parameter settings

on GERTS decision nodes. Activity times for various network events may be selected from any of nine different probability distributions and appropriately scaled to meet simulation requirements. GERTS is also suitable for modeling production systems which include buffer storage. Buffers (queues) are commonly installed between succeeding productive processes to smooth component transfer irregularities arising from machine failure. Statistics can then be computed on average, minimum, maximum, and standard deviation of such buffer characteristics as utilization level, busy time, and number of components balked from full buffers. This feature allows relatively complex queueing problems to be evaluated and tested for parameter sensitivity within a reasonable period of time.

A brief description of GERTS IIIQ is contained in the appendix. For more detailed information, the interested reader is directed to the references on GERTS and its antecedent, GASP (2-6).

NETWORK MODELS. Six basic network configurations are described in operational detail.

Network 1. - Individual Manufacturing Process

An individual manufacturing process is modeled in figure 1. by means of a decision node and two deterministic nodes on which statistics of production and failure events are maintained. A single starting event emanating from source node 2 initiates the manufacturing process and from that time, the source exerts no further influence on simulation statistics. Several theoretical results may be verified on network 1. First of all, availability is a quantity related to operative and repair times that characterize a given process. Let

MTBF = Mean time between failures

MTTR = Mean time to repair .

Using these definitions,

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}} .$$

Productive and failure cycle times are symbolically represented by T and MTTR while probabilities of production and failure are P and 1-P, each respectively. From these values, the expected operative and repair times are PT and (1-P)MTTR so that process availability can be expressed as follows:

$$\text{Availability} = \frac{PT}{PT + (1-P)\text{MTTR}} .$$

One particular choice of P equates these two availability expressions and forces the simulation model to demonstrate desired operative, failure, and

cyclic time relationships, that being

$$P = \frac{1}{1 + T/MTBF} .$$

For example, assume MTBF = 100, MTTR = 10, and T = 1. Then availability = .9091, and the required probability setting would be P = .9901.

All of the parameters necessary to test network one using GERTS IIIQ have now been defined. Output statistics should display an availability of 90.91%, production cycle time of unity, repair time of 10 units, and branch realization probabilities of 99% and 1%, for production and repair events, respectively.

A second theoretical result of this network arises from consideration of completion time statistics. The expected completion time, E(T), and its variance, V(T), for a single unit of production are derived from condition probability arguments, as illustrated by Pritsker and Whitehouse (7).

$$E(T) = T_1 + (1-P)T_2/P$$

$$V(T) = T_1 + T_2(T_2-1)(1-P)/P + (T_2(1-P)/P)^2.$$

Since figure 1 calls for completion of 100 units of production, observed values for E(T) and V(T) will be 100 times greater than predicted here. Statistical convergence of simulation data to these theoretical predictions is easily demonstrated with GERTS IIIQ by plotting both quantities against the total number of units produced.

Network 2. - Serially Coupled Processes

Figure 2 illustrates serially coupled processes configured such that the second stage (node 5) does not function until it receives output from the first stage (node 3). In addition, each item manufactured by the first stage must be accepted downline before further production on the first stage can resume. Both criteria are satisfied by demanding receipt of two completed activities at node 4; one arriving from node 3 and a second from node 5. These branches signal the operability status of both stages and serve to regulate further production. Due to the interactive nature of this network, its output cycle time, i.e., the time between finished items observed at node 6, is constrained by that stage containing the longest cycle time. Thus,

$$\text{Network Cycle Time} = \text{Max}(T_1+NT_2,\ T_3+MT_4)$$

where N = number of failure loops occurring between nodes 3 and 8,

M = number of failure loops occurring between nodes 5 and 9.

The probability that exactly N and M failure loops will precede a production event is just

$$P_1P_2(1-P_1)^N(1-P_2)^M.$$

Expected network cycle time, E(T), can be obtained by summing the product of network cycle time with the failure loop probability for all values of N and M between zero and infinity.

$$E(T) = P_1P_2 \sum_{N=0}^{\infty} \sum_{M=0}^{\infty} (1-P_1)^N(1-P_2)^M \text{Max}(T_1+NT_2, T_3+MT_4).$$

Since figure 2 calls for completion of 100 production units, the observed value of E(T) will be 100 times greater than predicted here.

Network 3. - Buffered Serial Processes

Figure 3 demonstrates two processes connected in series but physically separated by a buffer activity, queue node 5. The buffer absorbs productive output from node 4 and directs it to the second process via node 8, given that the second process is currently operable and awaiting such input. Violation of either condition causes the buffer to augment its inventory level until the second process resumes operability status. Augmentation ceases, however, when the inventory level attempts to exceed maximum buffer capacity since the GERTS program automatically balks further input away from a full buffer. The balking unit then triggers network modification three whereby production on the first process terminates pending return of the second process to operability status. In essence, modification feature three severs the feedback loop from node 4 through 12 to 3. Since no feedback pulse can propogate between nodes 3 and 13, further production on the first process immediately halts until the second process resumes its operative status and signals that fact by sending an impulse between nodes 10 and 8. At this point, a second impulse passes from node 8 to 16, 17, and finally to node 3 so that the first process may resume production. Small time phasing factors of magnitude .0001 have been inserted between certain critical nodes to guarantee proper functioning of this network.

Individual process availabilities are determined by proper selection of the various T_i and P_i, and explained for network one. Network availability in this type of serial configuration displays a strong dependence upon the buffer characteristics of maximum capacity and initial inventory level. For a zero capacity buffer,

$$\text{Network Availability} = \prod_{i=1}^{2} \text{Availability}_i$$

This condition reflects the interdependency attributable to lack of storage capacity between processes. Essentially, the entire network is operative only during those periods when both processes are simultaneously operative. In contrast, specification of an infinite buffer capacity facilitates maximum independence between processes so that

$$\text{Network Availability} = \min_i (\text{Availability}_i)$$

Network availability under the infinite buffer situation is effectively throttled by that process containing the lowest availability.

For values of buffer capacity between zero and infinity, network availability varies monotonically within the limits just discussed. The exact form of this dependence is strongly influenced by the totality of system parameters and cycle time distributions employed and it cannot in general be conveniently predicted. However, statistical data relating to the dependence are easily generated from the GERTS model by graphing observed network availability as a function of maximum buffer capacity. It should be emphasized that observed availabilities may differ from theoretical values assigned to the model. Monte Carlo solutions generally consume copious quantities of computer time before numerical predictions assume sufficient levels of accuracy to become useful. Divergence between observed and theoretical availabilities can be significant for short computing runs or even moderate length runs utilizing probability parameter settings near zero and unity since these settings are particularly sensitive to discretization error associated with the decision process. A more detailed discussion of decision node discretization error is presented in reference 7.

Network 4. - Distributor with Input Rate Dependence

Figure 4 utilizes the GERTS network modification feature to sequentially distribute incoming production between nodes 4 and 5 on a first-in-first-out (FIFO) policy. Sequential distribution exists as long as the interarrival time, T_1, at node 4 exceeds the sum of the distributor cycle times, T_2+T_3. Violation of this requirement results in a production flow bias favoring the loop containing the longest cycle time. Network 4 distribution characteristics are sensitive to input rate, a condition whose significance varies according to the context in which the distributor must fuction. Observe, however, that the distributor can accommodate additional nodes by extending the principle of network modification as often as is necessary. Figure 5 illustrates such an extension to a three node distributor.

Network 5. - Distributor with Input Rate Independence

For those situations demanding strict FIFO distribution, the distributor should be reconfigured as in figure 6. The basic difference between networks 4 and 5 resides in the additional feedback branches connecting nodes 6 to 3 and 7 to 3. Their presence limits the netowrk input rate to values which guarantee sequential properties in the distribution by effectively providing a lower bound on network cycle time:

$$\text{Cycle Time Limit} = \text{Max}\ (T_1, T_2, T_3)$$

Network 6. - Multiplexer Unit

Figure 7 illustrates a multiplexer network designed to accept and transfer input from several channels to a single output channel. Basic

characteristics of this multiplexing operation include a physical separation of input from output, periodic input channel sampling to determine whether units are actually awaiting transfer, and rate control on input cycle time. As units enter the various input channels, they first engage network modification activities and then enqueue in storage areas for further processing. Given a clear output channel, transfer is conveniently accomplished during the next sampling interval by means of a detection signal that circulates in the output section and initiates appropriate switching mechanisms. Simultaneous transference of more than one unit is precluded by network logic since only one set of switching relays activates at any given instant. All other units excepting that actually undergoing transfer are held in abeyance in the input section. Input unit sampling is periodically performed every T_3 time intervals. The output channel cycling time, T_4, is composed of component transfer plus switching times lumped together as a single quantity.

Input rate control is a logical function that has been implemented as follows. Once a unit enters a particular channel for processing, the entrance to that channel closes until the unit successfully transfers. Completion of unit transference initiates an unblocking operation that opens the input channel. This on-off logic can also be extended further back through the input channel to control the rate of unit production before such production enters the multiplexer.

SUMMARY. Six network designs were documented to foster greater interest in the use of GERTS IIIQ, illustrate problems arising in verification of model behavior, and avoid duplication of developmental effort in discrete event production line modeling. The operation of each network was discussed with theoretical predictions being presented for comparison against observed statistical results. Although the techniques were primarily conceived to model production line activities, their applicability can easily be generalized to any network flow situation involving renewal operations, positive flow capacities, and unidirectionality of flow along each arc.

BIBLIOGRAPHY

1. Roffman, Martin M., System Simulation of Parallel Array Production Lines. Frankford Arsenal Research Report R-2077, Philadelphia, Pennsylvania, 19137. September, 1973.

2. Pritsker, A. Alan B. and Philip J. Kiviat, Simulation with GASP II, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

3. Pritsker, A. Alan B. and W. William Happ, "GERT: Graphical Evaluation and Review Technique, Part 1: Fundamentals," The Journal of Industrial Engineering, 17(5): 267-274, May, 1966.

4. Pritsker, A. Alan B. and Gary E. Whitehouse, "GERT: Graphical Evaluation and Review Technique, Part II: Probabilistic and Industrial Engineering Applications," The Journal of Industrial Engineering, 17(6): 293-301, June 1966.

5. Pritsker, A. Alan B. and Roland R. Burgess, The GERT Simulation Programs: GERTS III, GERTS IIID, GERTS IIIC, and GERTS IIIR. Department of Industrial Engineering, Virginia Polytechnic Institute, Blacksburg, Virginia, 24061.

6. Whitehouse, Gary E. and A. Alan B. Pritsker, "GERT: Part III- Further Statistical Results: Counters, Renewal Times, and Correlations," AIIE Transactions, I(1): 45-50, March, 1969.

7. Roffman, Martin M., Digital Filtering of Correlation and Discretization Noise Generated During Stochastic Network Simulation. Ph. D. Dissertation, Engineering Systems Department, University of California, Los Angeles, California, 90024. February, 1973.

APPENDIX

GERTS IIIQ networks consist of nodes and directed branches in which each node represents completion of an event. The basic GERTS program performs a simulation of the network by advancing time from event to event. The start nodes are realized at time zero and all activities are scheduled from them. Events are sequentially realized thereafter until a specified number of sink nodes is realized, this number representing the completion of a simulation run. The procedure repeats itself until all the desired runs are performed, at which time, the output statistics are printed.

A node is deterministic if the output branches (activities) are all simultaneously scheduled, i.e., the probability is unity that all branches emanating from the node are taken. A node is probabilistic if only one emanating branch is taken, where the selection probability is proportional to a predetermined parameter setting. For event and sink nodes, table A.1 demonstrates that the A parameter specifies the number of times activities incident to a node must be completed before the node is realized. The B parameter specifies the number of times the activities incident to the node must be completed after the node is realized the first time to realize the node again and to reschedule activities. The N parameter is a number assigned to each node; all nodes must be assigned number and all numbers must be unique. The C parameter of a queue node specifies the intial number of items in the queue; the D parameter the maximum number of items allowed in the queue. Coding format identifies deterministic nodes by semicircles on the output side and probabilistic nodes by triangles.

Branches of GERTS IIIQ networks represent activities and/or information transfer. Associated with each branch are the following characteristics.

(1) The probability that the activity will be initiated, given that the node from which it emanates is realized;

(2) The time to perform the activity once it is started. The time function can be distributed according to one of the available types: Constant, Normal, Uniform, Erlang, Lognormal, Poisson, Beta, Gamma, and Beta fitted to three parameters as in PERT;

(3) A counter that records the number of times a group of activities is exercised before a node is realized;

(4) An activity number which permits network modifications.

The activity emanating from a queue node represents a service activity. With the present version of the program, only one service activity can be associated with a queue node.

Counters can be assigned to a branch or a group of branches. The number of counts recorded is always referenced to a particular node.

Network modifications are possible by assigning an activity number to a branch, and when the branch is activated, replacement of a node by another node occurs. The input to the replacement node is the same. Only the output branches are changed by whatever branches emanate from the replaced node.

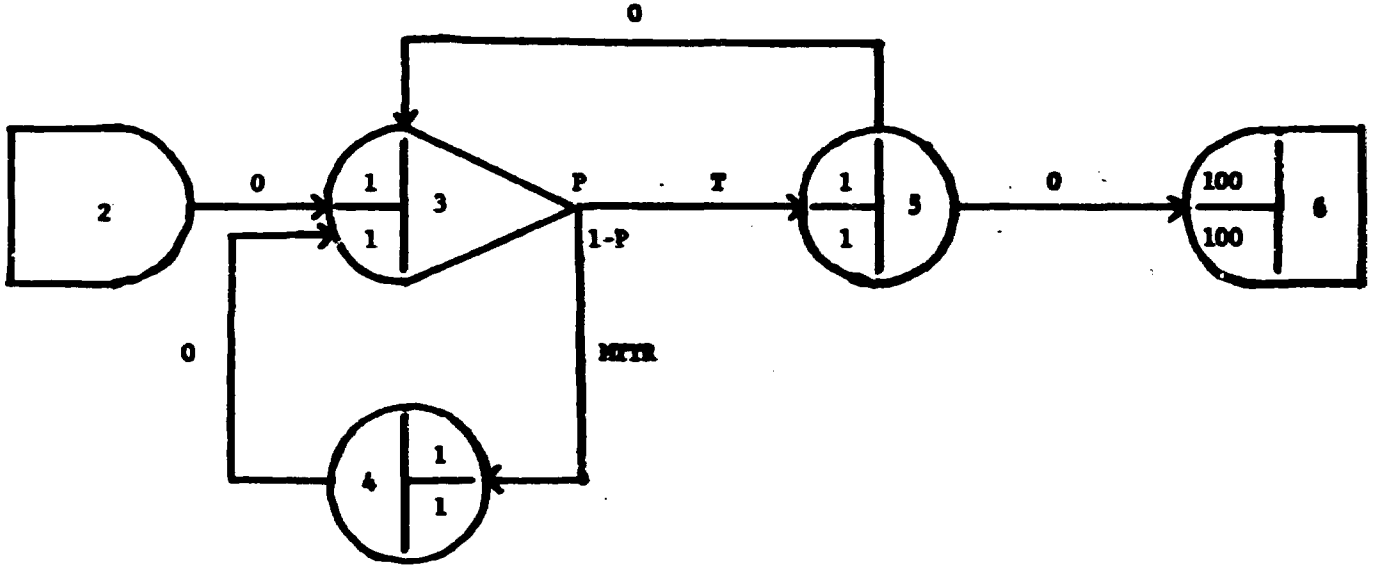


FIGURE 1. - SINGLE MANUFACTURING PROCESS

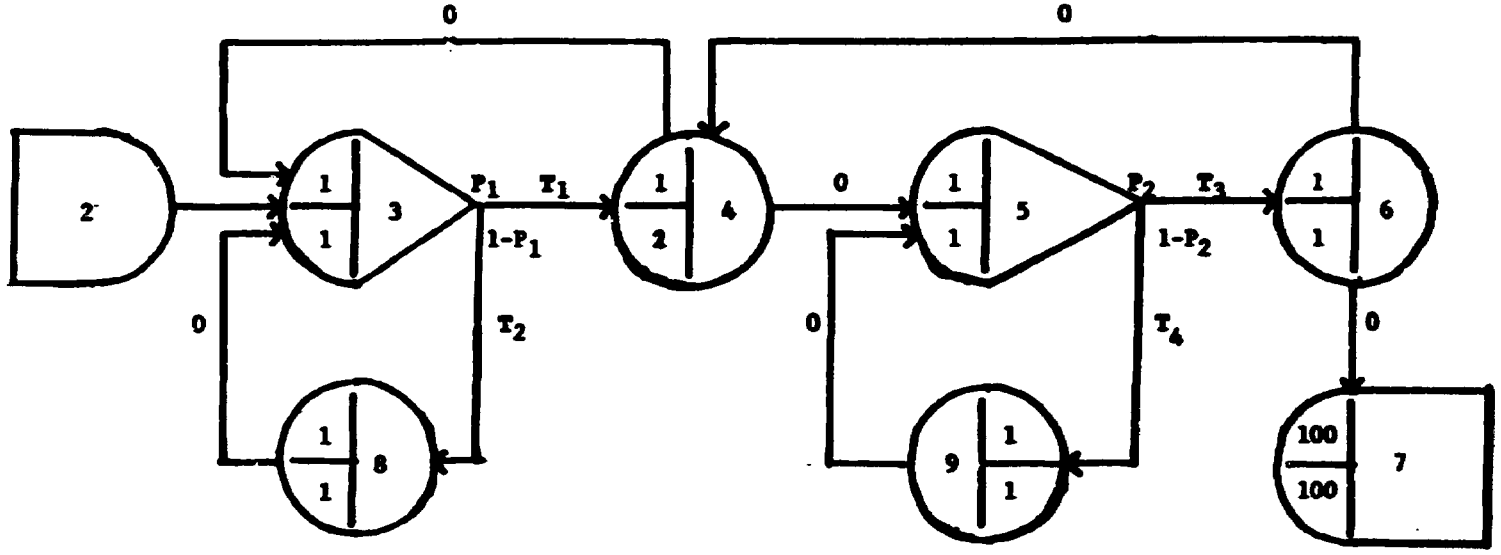


FIGURE 2. - SERIALLY COUPLED PROCESSES

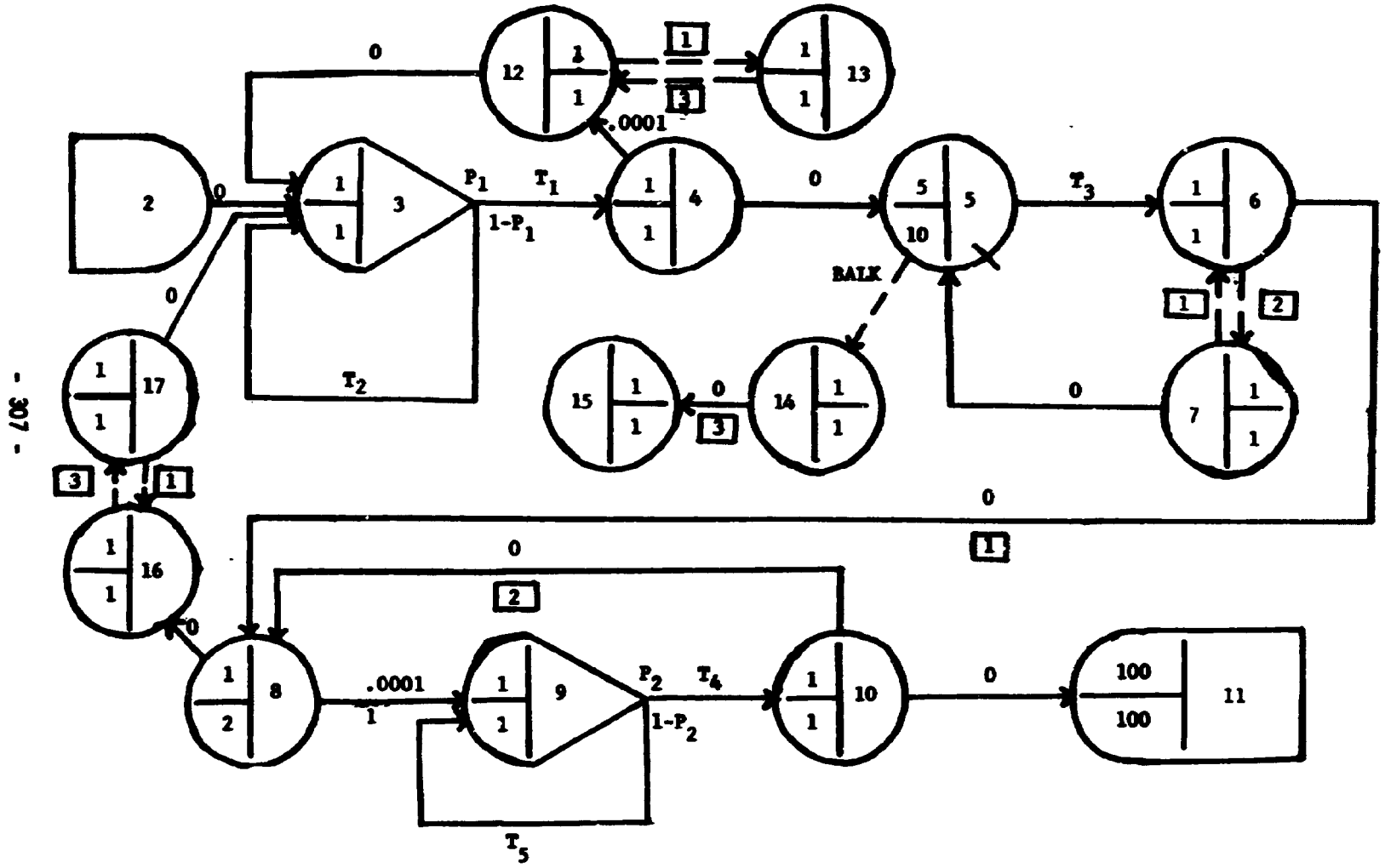


FIGURE 3. - BUFFERED SERIAL PROCESSES

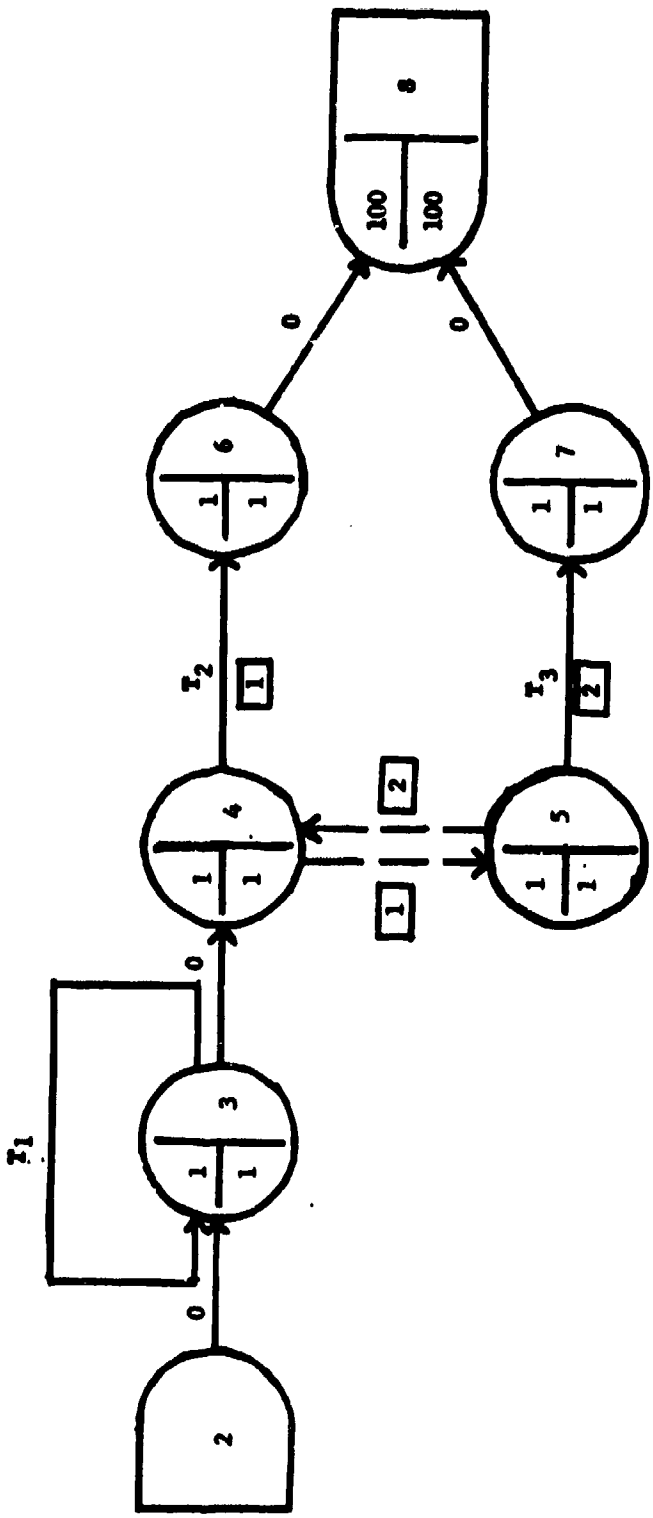


FIGURE 4. - DISTRIBUTOR WITH INPUT RATE REFERENCE

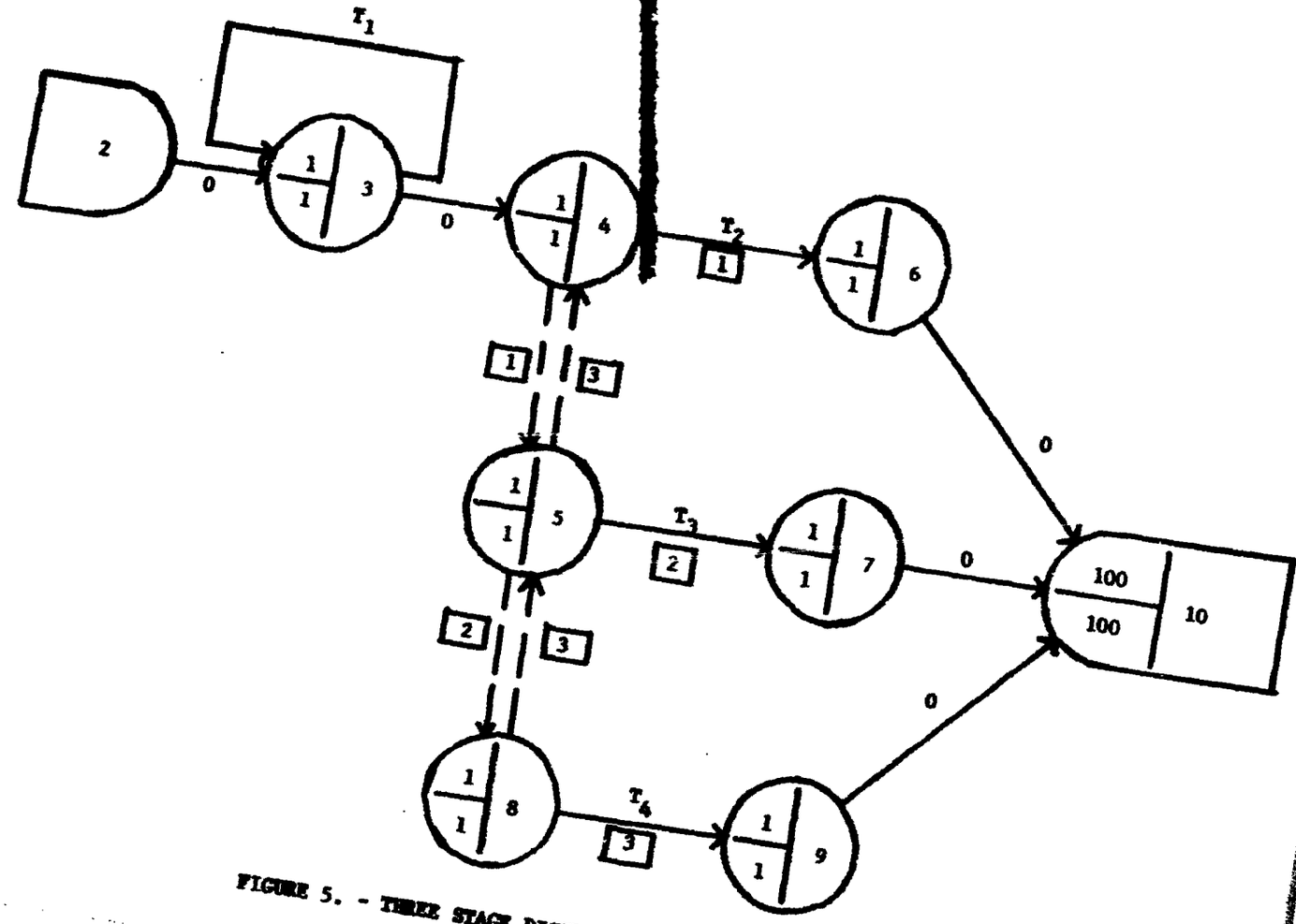


FIGURE 5. - THREE STAGE DISTRIBUTOR WITH INPUT RATE DEPENDENCE

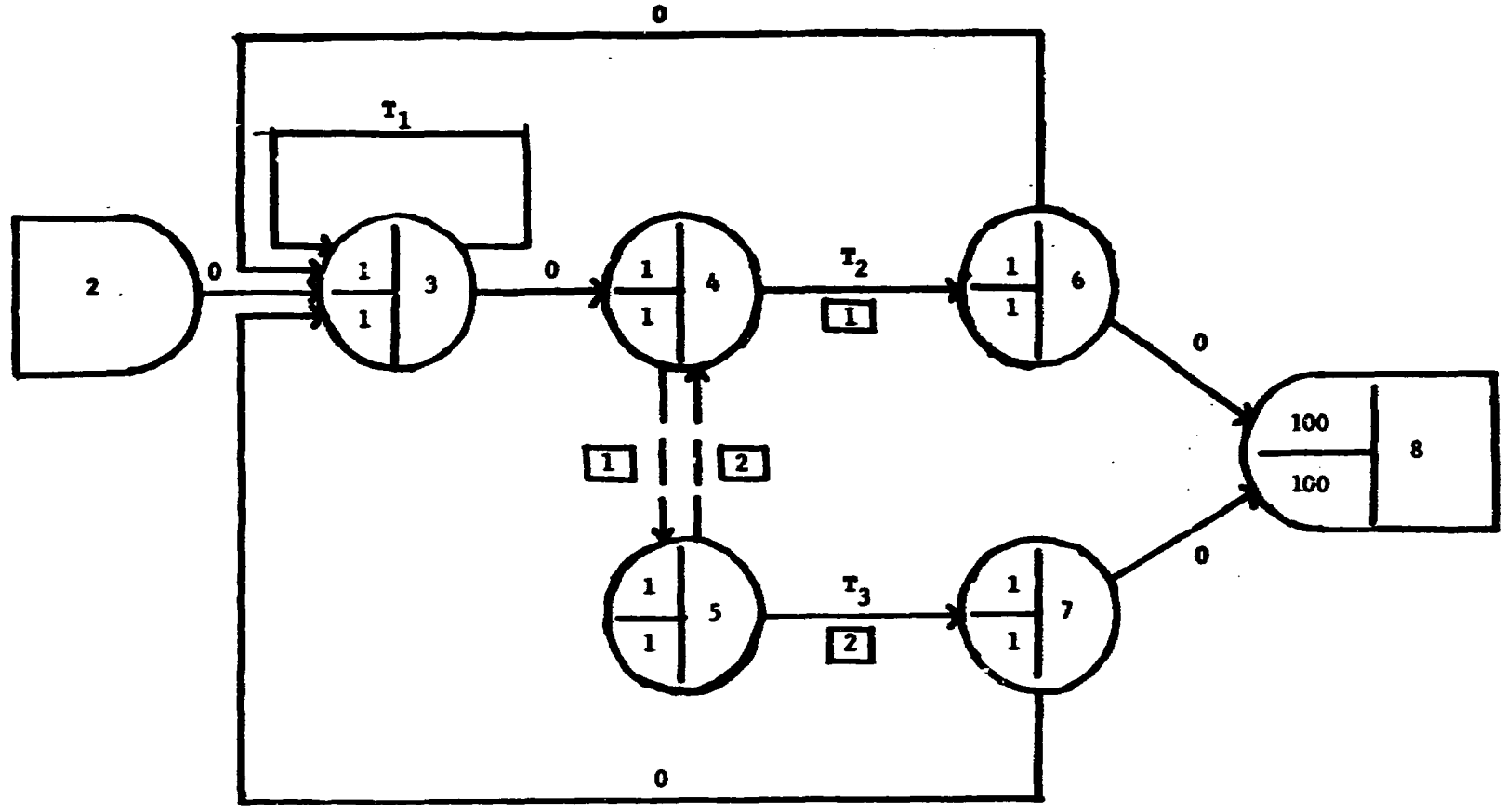


FIGURE 6. - DISTRIBUTOR WITH INPUT RATE INDEPENDENCE

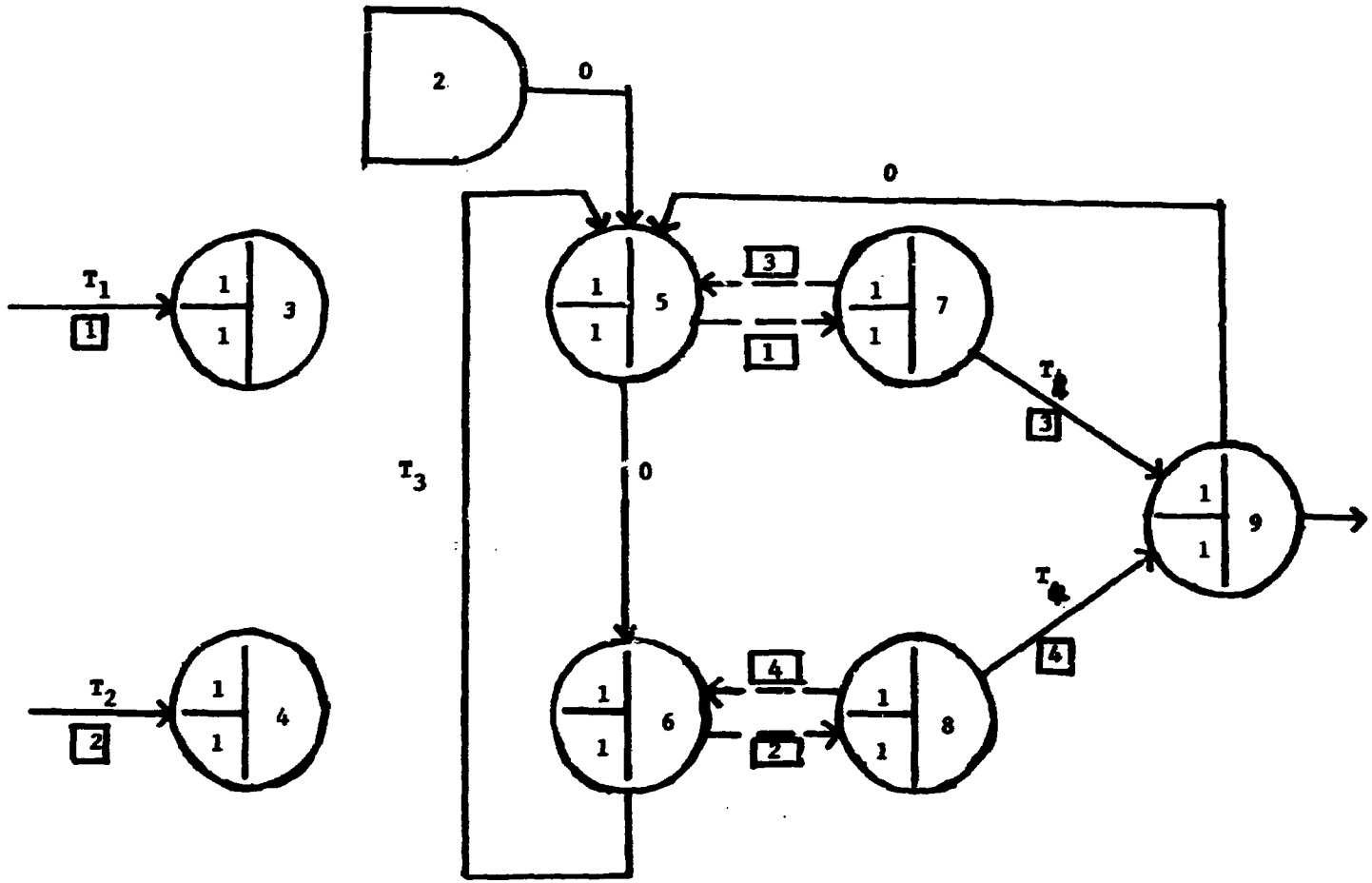
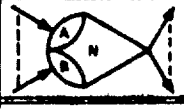
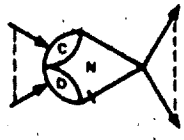


FIGURE 7. - MULTIPLEX NETWORK

PROGRAMS	NODE FUNCTION	DETERMINISTIC SYMBOL	PROBABILISTIC SYMBOL	PARAMETERS
Basic to Gert III	Source			A = No. of times activities incident to the node must be realized for node to be realized B = No. of times activities incident to the node must be realized after the first time the node is realized to repeat N = Node number
	Sink[1]		Not Applicable	
	Event			
Added to Form Gert IIIQ	Queue			C = Initial number in the queue D = Maximum number allowed in queue N = Node number

Note: [1] The sink node can not really be classified in this manner since it can be an event node with a sink node assigned to it. The symbol shown is the most common form used and is included here as a matter of convenience.

TABLE A-1
GERTS IIIQ NODE CAPABILITIES

CONFIGURATION MANAGEMENT FOR RELIABILITY/MAINTENANCE TESTING

John Fargher
Production Planning and Control
Ordnance Department
Naval Ammunition Depot
Hawthorne, Nevada 89415

ABSTRACT. This paper presents the methods employed by the Small Caliber Ammunition Modernization Program (SCAMP) to automate the reporting of maintenance actions on Module A1 of their new generation production equipment (NGPE) for manufacturing 5.56mm ammunition. The reporting procedure utilizes the process controller (minicomputer) to store failure times and causes of failure. Item control is obtained by use of serial numbers on assemblies. Sub-assemblies are not serial numbered but are included in the failure cause code to the controller.

Checks on new items, reliability degradation, and on out-of-control upstream operations can be obtained because of the configuration of tooling (24 separate, identified lines). Statistical tests being implemented to test for reliability include the use of non-parametric statistics before sufficient historical data has been gathered to generate a probability density function. After probability density function can be assumed, continuous control charts are illustrated by use of computer graphics for a management information system.

Any production plant, from a job shop through a dedicated facility requires a certain sequence of operations. These can be described in a broad sense as:

1. Quality inspection of raw materials and materials used inprocess (parts, capital equipment, oils, greases, utilities, etc.).
2. Processing the raw materials into a finished product.
3. In-process inspection for quality and process control.
4. Final verification inspection of quality prior to shipment or stowage.

Quality inspection of raw materials is vital to eliminate quality and production (reliability of production equipment) problems. Final verification inspection is required to assure that all processes are in control. In-process inspection is utilized to segregate "out-of-tolerance" items from subsequent operations. In-process inspections, if properly placed and planned, can be essential in reducing downstream downtime and discovering upstream quality and reliability problems so that corrective maintenance can be instituted. Control can be maintained by manual or automated methods, but this paper will deal in only automated, in-process reporting of inspection and reliability deficiencies. Figure I illustrates the process and information flow.

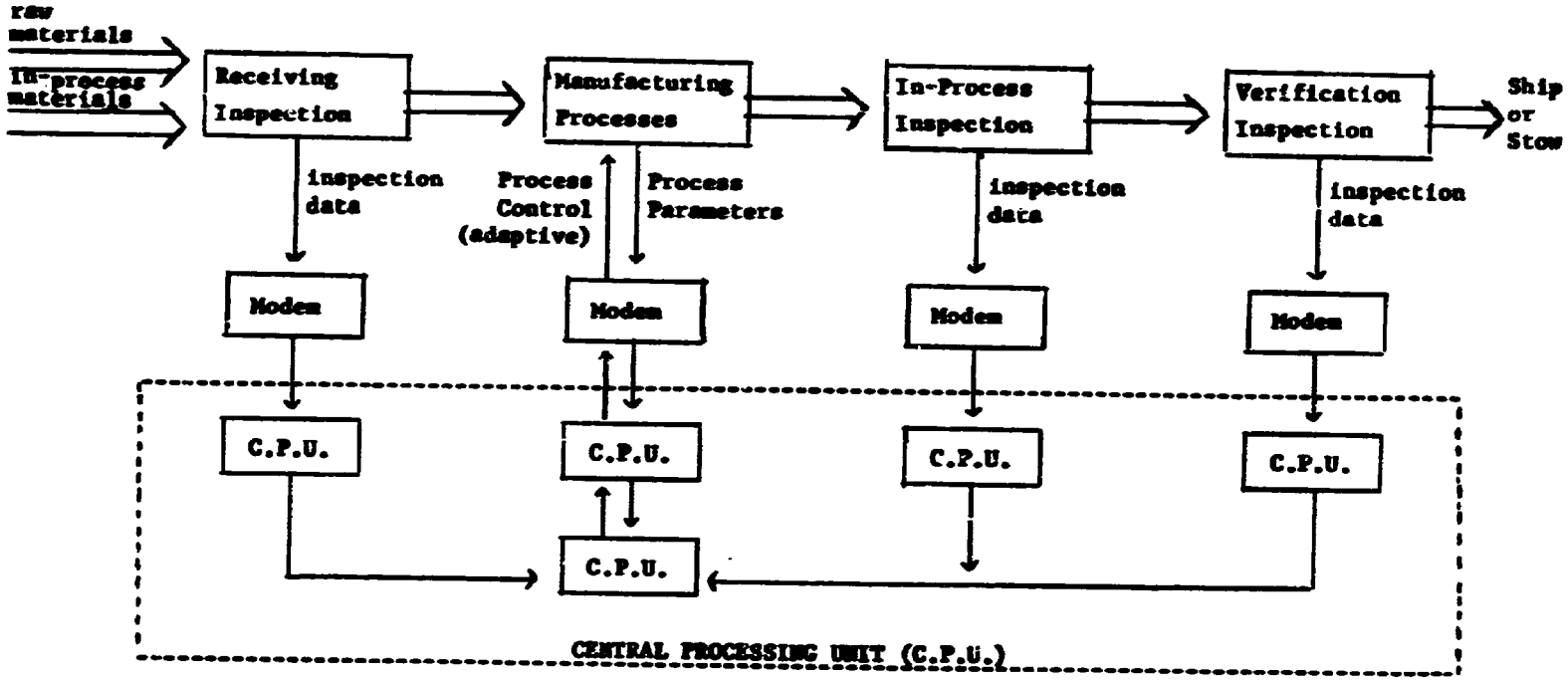


Figure I Process Control System Utilizing A Computer Network.

Optimal control of any process can be attained by monitoring the process parameters. Conditions creating "out-of-tolerance" conditions can be detected before "out-of-tolerance" items are produced. But, because of the difficulty of measuring these parameters (and even determining what parameters to measure) and reliability problems within the equipment measuring these parameters, in-process inspections are utilized as "the" control parameters or as a back-up system for verification and control.

Digital process control computers are emerging as an important part of process and quality control. The process control computer is a reliable and accurate tool for solving complex manufacturing problems as well as simple logical control of machines. An important new development is that of the minicomputer, with their low cost and ease of programming changes. Inspection parameters can be readily changed and input parameters added or changed with only minor programming.

Inspection reporting can be accomplished through digital signals such as go-no go, visual defects entered on a keyboard, and photodiode circuits and analog signals such as contracting automatic gaging and non-contacting air, light or laser gaging (See Figure 2).

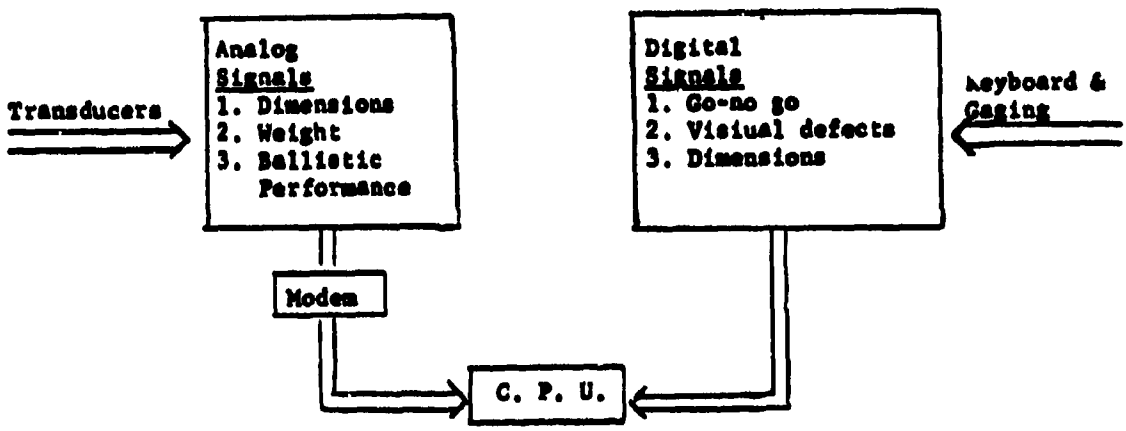


Figure 2. In-Process Inspection Reporting

Reliability, the probability that an item will perform properly for a certain specified time, is an especially important parameter in the operation of a system. Maintenance downtime is the other essential parameter determining production system availability. Inventory control of spare parts, production planning and scheduling, maintenance action scheduling (both preventive and corrective) and maintenance manpower requirements are all functions of reliability, availability and maintainability (RAM). The purpose of this paper is to tie the reporting of maintenance actions with (1) computerized data processing and (2) proven statistical analysis techniques to

1. Record tool and equipment life data,
2. Generate process control notices when out-of-control condition is noted,
3. Relate failures to possible causes,
4. Prepare standard maintenance reports and
5. Prepare special reports upon request.

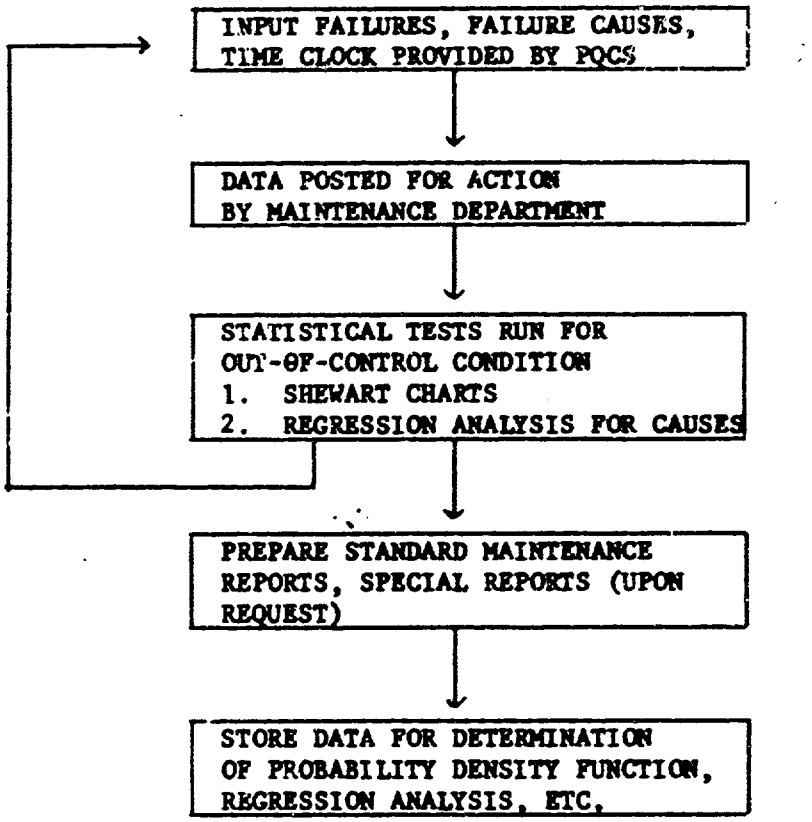


Figure 3. Flow of Equipment Failure Data

An example of a reporting system being implemented is Module A1 of their new generation production equipment (NGPE) by the Small Caliber Ammunition Modernization Management Office (SCAMMO) at Frankford Arsenal. Module A1 manufactures 5.56mm ammunition. The reporting procedure utilizes a process controller (minicomputer network) to store failure times and causes of failures. The process controller, called a Process Quality Control System (PQCS), is responsible for monitoring quality and production rates. From these, the PCQS controls production rates and issues management reports on the operation of Module A1. The inputs to the PQCS are (1) Inspection data (quality assurance), (2) Tools lost

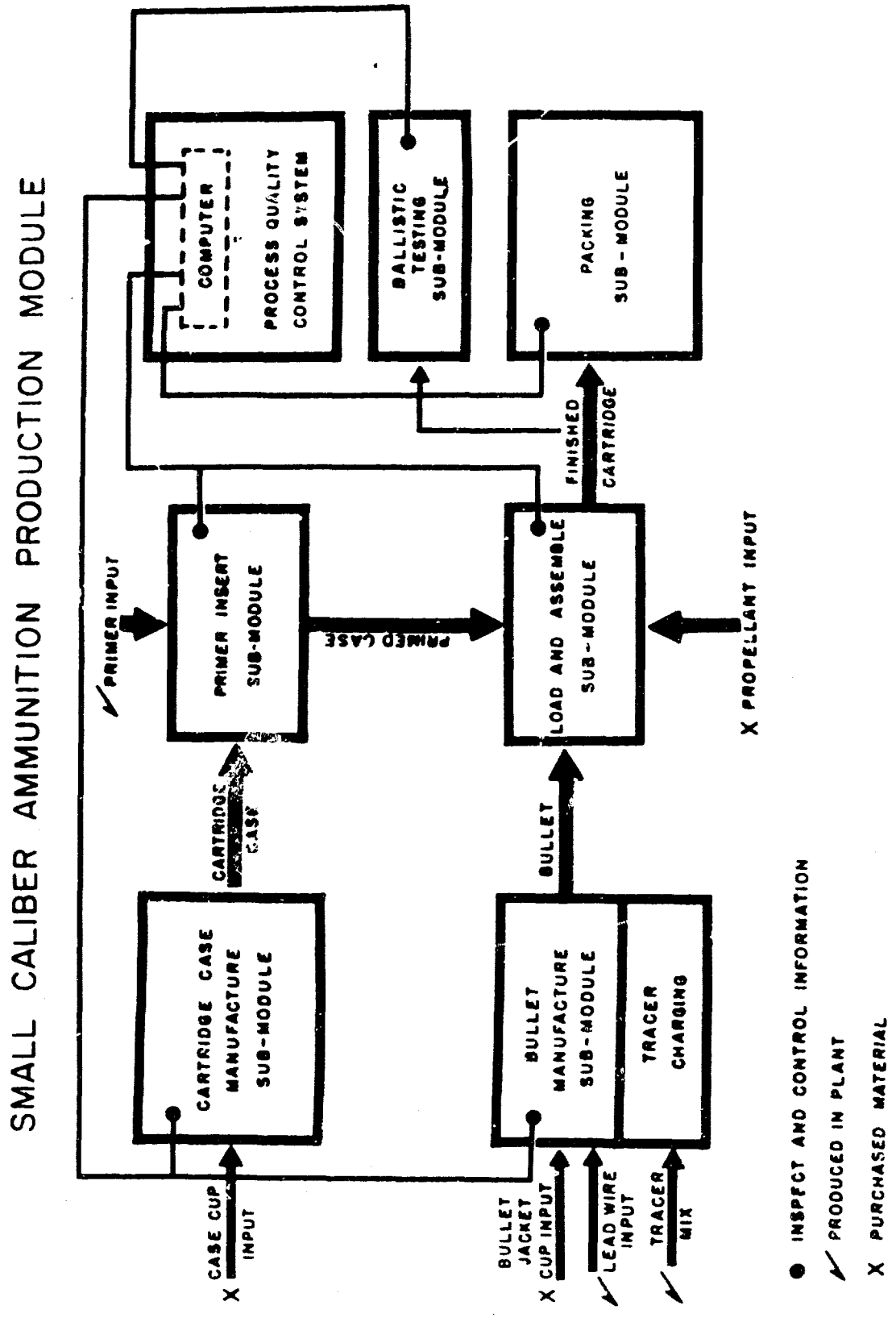


Figure 4. SCAMP Module A1 System Configuration

from production (FQCS provides time clock), and (3) Coded information from the operating and maintenance personnel for equipment identification and failure cause. The coded information is enterred through teletype-writers. Figure 4 illustrates the system configuration with inputs to FQCS.

Configuration Management. In order to identify tooling as to submodule and turret, the following convention will be utilized for providing each tool module with a serial number (8 digit number).

Type of Equipment	Machine	Consecutively Numbered
01	02	0 184

Figure 5. Serial Number for Tooling.

This S/N will allow data to be generated for each tool and by tooling type (by turret).

Additionally, whenever a tool is brought into service (i.e., a serial number is assigned it) a file is reserved in the FQCS for data and as a check for actually having that tool in service. Whenever data is fed into the FQCS about tooling, a search is performed to make sure that that file is in existance, i.e., the tool S/N exists. As tooling is replaced on the machinery, a data input is prepared showing the new tooling S/N, the old tooling S/N, and the time and date (time and date will be supplied by FQCS itself), and tool station number (1 to 24). An example of the data input is:

New Tool S/N	Old Tool S/N	Tool Station No.
01 020184	01 020207	005

Figure 6. Report of a Tool Change.

By replacing a tool module S/N, the FQCS will not allow the old tooling to return to the production system until the tool room performs repairs. A remarks card relating maintenance action would be

S/N	Action	Repairs Required
01 020207	0	07

Figure 7. Tool Room Report.

The "action" column would signify tool module disposition, 0 for return to queue and 1 for retirement of that tool module. As the type of repairs required are documented, separate codes will be generated for the "Repairs Required" column.

Advantages of Using S/N for Configuration Management. Maintenance actions can be broken down by tool module, turret station and line, and tooling type (also by module when follow-ons are installed). By considering only tool modules, "bad" tool modules can be identified as "out-of-control" and not allowed to re-enter the production system. Tool stations and upstream operations can be checked for out-of-tolerance performance by checking by turret station and line. Particular tooling design problems can be checked using data by tooling type.

Non-Serial Numbered Items. Where S/N identification is not required (because of simplicity of design) the following convention will be used to report repairs:

Submodule	Turret	Tool Station	Item	Repairs Required
01	07	00	09	109

Figure 8. Report of a Repair to a Non-Serial Numbered Item

This data input will be supplied to the FQCS as the item is replaced showing submodule number, turret number in that submodule, tool station, item number, a code to indicate what repairs were required and probably cause of failure. Whenever a tool station is not involved, such as a chain, a 00 code will be used. Whenever verification of failure mode is required, the item should be tagged (physically by a red tag requesting analysis and in the FQCS by a flag). The flag is to be removed only by inputing an updated failure mode.

Statistical Tests to be Performed on the Data. As a check for process control, analysis of the data will be required for tool and item life. Results of the statistical tests will provide a means of verifying the process control check itself with a time lag.

Before sufficient historical data is available for generating a probability density function, p.d.f., a non-parametric reliability estimation procedure will be used. Whenever sufficient data becomes available to estimate the parameters of the p.d.f., the proper density function may be found for reliability estimation.

REFERENCES

1. P. C. Chen, Ulrich Remfold & Jeremy S. Weinstein, "On-Line Computerized Testing", Institute of Electronics and Electrical Engineers Transactions on Industrial Applications, Volume IA-9, No. 1, Jan/Feb 1973, pp. 69-73.

2. Donald E. Henry, "Programmable Controllers: What They Really Involve", Factory, June 1973, pp. 44-46.

3. R. A. Radford & J. E. Stuehler, "The Role of Process Data in Controlling & Optimizing Manufacturing Processes", Institute of Electronics & Electrical Engineers Transactions on Manufacturing Technology, Volume MFT-1, No. 1, June 1972, pp. 5-8.

4. J. E. Stuehler, "An Intergrated Manufacturing Process Control System: Implementation in IBM Manufacturing", IBM Journal of Research and Development, Nov 1970.

5. Lawrence E. Seaton & Carl P. Green, Jr., "Automated Process Audit and Certification Program", Quality Management & Engineering, March 1973, pp. 24-28.

6. David T. Symmes, "Programmable Controllers, What They Are and What They Can Do For You", Assembly Engineering, August 1973, pp. 32-35.

BAYESIAN INFERENCES FOR INTERVAL RELIABILITY OF WEIBULL COMPONENTS

Ronald L. Racicot
Benét Weapons Laboratory
Watervliet Arsenal
Watervliet, New York 12189

ABSTRACT. Bayesian intervals have been generated for the interval reliability of a component within a system given component failure data. The reliability index considered is the average interval or mission reliability of a component over system life under the assumption of ideal repair. Under this assumption a component is replaced or renewed with a new component whenever failure occurs during system operation yielding a renewal process.

The distribution assumed for component interarrival failure times is the two-parameter Weibull with both parameters unknown. Since classical confidence limits cannot generally be found for this problem, a Bayesian approach is used to render the problem at least numerically tractable. Uniform priors were assumed for the Weibull parameters to simplify derivations and computations and to approximately represent complete prior ignorance. Since this is not necessarily true, a number of Monté Carlo trials were conducted to study the exactness of the Bayesian intervals from the classical frequency viewpoint. Results indicate that the Bayesian intervals generated yield close to exact frequency intervals depending on sample size, Weibull shape parameter and the true reliability.

1. INTRODUCTION. The general problem considered is the determination of inference information for the interval reliability of a component within a system from failure data. The reliability index considered is the average interval or mission reliability of a component over system life under the assumption of ideal repair [1]. Under this assumption a component is replaced or renewed with a new component whenever failure of the component occurs during system operation. In this instance component failures and subsequent replacements form a renewal process. The distribution assumed for component interarrival failure times in this study is the two-parameter Weibull with both parameters unknown.

The problem is basically one of determining inferencing information on a relatively complicated function of more than one population distribution parameter. Since classical confidence limits cannot generally be found for this problem, a Bayesian approach is used to render the problem at least numerically tractable. Numerical techniques were developed for generating Bayesian limits for the average interval reliability of a Weibull component over system life. An approximation to the interval reliability valid for high reliability components was also made in the numerical work to improve computational efficiency.

In the Bayesian formulation, uniform priors were assumed for the Weibull parameters to simplify derivations and computations and to approximately represent complete prior ignorance. Since this is not necessarily true, a number of Monte Carlo trials were conducted to study the exactness of the Bayesian intervals from a classical frequency viewpoint. Results indicate that the Bayesian intervals generated yield close to exact frequency intervals depending on sample size, Weibull shape parameter and the true reliability.

2. STATEMENT OF THE PROBLEM. The problem of specific interest is the determination of the reliability of a component within a system for a mission time interval $(t,t+\tau)$ where t is the system age and τ is the mission length. Prior to time t the component could have failed and been replaced one or more times. For the non-constant failure rate component, the interval reliability is a transient function of system time. The general formulation of this problem is well covered in the literature with only the final results being summarized here [1-5]. The reliability in this case is given as

$$R(t,\tau) = 1-F(t+\tau) + \int_0^t [1-F(t+\tau-x)]h(x)dx \qquad (1)$$

in which $R(t,\tau)$ = Interval reliability at system time t for a mission length τ.

$F(t)$ = Distribution of the interarrival times or first failure times of the component as a function of component age.

$h(x)dx$ = Renewal rate or density which describes the unconditional probability of failure and subsequent renewal at time x.

Another form of equation (1) which is more suitable for numerical computation is given by the relation [6]

$$R(t,\tau) = 1 - \int_t^{t+\tau} [1-F(t+\tau-x)]h(x)dx. \qquad (2)$$

In equation (2), integration is required only over the interval $(t,t+\tau)$ whereas in equation (1) integration over the entire interval (o,t) is required for each different value of t. In addition, a simplifying assumption can generally be made in equation (2) for high reliability components (i.e. τ<< Mean time between failure). In this case $F(\tau) \cong 0$ and equation (2) reduces to

$$R(t,\tau) \cong 1 - \int_t^{t+\tau} h(x)dx. \qquad (3)$$

The renewal rate h(x) in the above equations is a function of the underlying failure distribution of a component given by the following equation:

$$h(t) = f(t) + \int_0^t h(x)f(t-x)dx \qquad (4)$$

in which $f(t) = dF(t)/dt$.

Solution of equation (4) for given f(t) can be accomplished using numerical methods [6-8] or, for the Weibull distribution, through the use of the power series solution given in reference [9].

In practical applications, the transient reliability in itself is not entirely useful, for example, in making general comparisons of component reliabilities or in specification of required reliability levels for components. This difficulty is often overcome by considering either the worst reliability or the average reliability over some specified system life. In this particular study, the average reliability is considered and is defined by the relation

$$R_a(\tau) = n^{-1} \sum_{i=1}^{n} R_i(\tau) \qquad (5)$$

in which $R_a(\tau)$ = Average interval reliability
$R_i(\tau)$ = Reliability for the ith mission.
n = Expected number of missions over system life.
τ = Mission length.

Using equation (2) for the reliability of the ith mission with system time t being set equal to $(i-1)\tau$, equation (5) for average interval reliability reduces to

$$R_a(\tau) = 1 - n^{-1} \int_0^T G(x)h(x)dx \qquad (6)$$

in which T = The specified system life
= τn
G(x) = Function $[1-F(\tau-y(x))]$ where y(x) is equal to x on the initial interval (o,τ) and is then periodic with period τ for $x>\tau$.

For the approximate reliability given by equation (3) average reliability becomes

$$R_a(\tau) \cong 1 - n^{-1} \int_0^T h(x)dx. \qquad (7)$$

The failure distribution F(t) assumed in this work is the two-parameter Weibull given by the expression

$$F(t;\alpha,\beta) = 1 - \exp(-\alpha t^{\beta})$$
$$f(t;\alpha,\beta) = \alpha\beta t^{\beta-1} \exp(-\alpha t^{\beta}) \qquad (8)$$

in which β = Shape parameter

$\alpha = 1/\eta^{\beta}$ where η is the scale parameter or Weibull characteristic life.

The parameter α is used here instead of the usual characteristic life to facilitate derivations of the Bayesian intervals to be presented later.

Finally, it is assumed that component failure and suspension data has been generated where x_{fi} for $i=1,\ldots,n_f$ represents n_f failures and x_{sj} for $j=1,\ldots,n_s$ represents n_s suspension or censored times. Inference information in the form of confidence limits is then required for the average system reliability given by equation (6) or (7).

3. DETERMINATION OF BAYESIAN INTERVALS. In this study Bayesian intervals are derived through the use of the likelihood function and Bayes' theorem. Details of this approach are well known and can be found in the literature [10].

The likelihood function for the Weibull distribution is given as [11]

$$L(X;\alpha,\beta) = \alpha^{n_f}\exp(-\alpha\sum_{j}^{T.S.} x_j^{\beta})\beta^{n_f}\prod_{i=1}^{n_f}(x_{fi})^{\beta-1}$$

$$= a(\beta)\alpha^{n_f}\exp(-\alpha b(\beta)). \qquad (9)$$

in which X = Sample outcome x_{fi} for $i=1,\ldots,n_f$ and x_{sj} for $j=1,\ldots,n_s$.

T.S. = Summation over total sample including both failure and suspension times.

$$a(\beta) = \beta^{n_f}\prod_{i=1}^{n_f}(x_{fi})^{\beta-1}$$

$$b(\beta) = \sum_{j}^{T.S.} x_j^{\beta}$$

A simple assumption for prior distribution of the parameters α and β is the uniform where α and β are assumed independent. This simplifies numerical computation and derivation of posteriori information about the parameters and, more important, functions of the parameters. Assuming uniform priors does not necessarily mean assuming maximum ignorance of the parameters but, as will be shown shortly, the Bayesian intervals prove to be nearly exact in a classical frequency sense. Using Bayes' theorem [10] the posteriori distribution for α and β can now be written directly as

$$f(\alpha,\beta;X) = C\, a(\beta)\alpha^{n_f} \exp(-\alpha b(\beta)) \tag{10}$$

in which C = Normalizing constant such that area under $f(\alpha,\beta;X)$ is unity.

The function $f(\alpha,\beta;X)$ is defined as the conditional density distribution of the parameters α and β given the sample outcome X. The constant C in equation (10) can be determined from the relation

$$C = 1.0/\int_0^\infty \int_0^\infty a(\beta)\alpha^{n_f} \exp(-\alpha b(\beta))\,d\alpha\,d\beta. \tag{11}$$

Letting $y = \alpha b(\beta)$ in equation (11) yields

$$C = 1.0/\int_0^\infty a(\beta)(b(\beta))^{-n_f-1} \int_0^\infty y^{n_f} e^{-y}\,dy\,d\beta$$

$$= 1.0/[n_f! \int_0^\infty a(\beta)(b(\beta))^{-n_f-1}\,d\beta]. \tag{12}$$

Consideration is given next to determination of the posteriori distribution of average interval reliability $R_a(\tau;\alpha,\beta)$ which is essentially a function of the parameters α and β. From probability theory [12] the cumulative distribution for R_a can be determined from $f(\alpha,\beta;X)$ using the relation

$$F_{Ra}(z) = \int_{D_{Ra}} \int f(\alpha,\beta;X)\,d\alpha\,d\beta \tag{13}$$

in which D_{Ra} = Domain of R_a and represents all values of α and β such that $R_a(\tau;\alpha,\beta) \leq z$.

Assume next that for given β and R_a, equation (6) or (7) can be solved for α. That is

$$\alpha = R_a^{-1}(R_a,\beta).$$

Letting $\alpha=\alpha^*$ whenever $R_a=z$ gives for this expression

$$\alpha^* = R_a^{-1}(z,\beta). \tag{14}$$

Numerical solution of equation (14) for α^* is considered in the next section. Also, it is shown in the Appendix that for reliability given by equation (7)

$$R_a(\alpha,\beta) \leq R_a(\alpha^*,\beta) \tag{15}$$

whenever $\alpha \geq \alpha^*$.

The domain of R_a, D_{Ra}, can now be inferred directly from equation (15). The posteriori distribution for R_a given by equation (13) can now be written as

$$F_{Ra}(z) = \int_0^{\infty}\int_{\alpha^*}^{\infty} f(\alpha,\beta;X)d\alpha d\beta \qquad (16)$$

in which $\alpha^* = \alpha^*(z,\beta)$.

Substituting equation (10) into (16) yields the following expression:

$$F_{Ra}(z) = \int_0^{\infty}\int_{\alpha^*}^{\infty} C\ a(\beta)\alpha^{n_f} \exp(-\alpha b(\beta))d\alpha d\beta. \qquad (17)$$

Letting $y=\alpha b(\beta)$ and using equation (12) for the constant C finally yields after some manipulation

$$F_{Ra}(z) = 1.0 - K^{-1}\int_0^{\infty} a(\beta)(b(\beta))^{-n_f-1} P(n_f+1, w(\beta))d\beta \qquad (18)$$

in which $K = \int_0^{\infty} a(\beta)(b(\beta))^{-n_f-1} d\beta$

$$a(\beta) = \beta^{n_f}\prod_{i=1}^{n_f} x_{fi}^{\beta-1}$$

$$b(\beta) = \sum_{f}^{T.S.} x_f^{\beta}$$

$$w(\beta) = \alpha^*(z,\beta)b(\beta)$$

$P(n,x)$ = Incomplete Gamma function

$= 1 - (1+x+x^2/2!+\ldots+x^{n-1}/(n-1)!)e^{-x}$ for interer n.

Solution of equation (18) for given z and sample outcome X is accomplished using numerical quadrature.

Bayesian intervals on R_a can now be constructed directly from $F_{Ra}(z)$. For example, a lower Bayesian limit z_ℓ for confidence or probability level $(1-\phi)$ is given as

$$z_\ell = F_{Ra}^{-1}(\phi). \qquad (19)$$

4. COMPUTATION OF $\alpha = R_a^{-1}(z,\beta)$.

In the solution for the Bayesian intervals, the inverse of the reliability function given by equation (6) or (7) is required in determining D_{R_a} in equation (13). For the approximate interval reliability equation (7), this was accomplished numerically where a particular property of $h(t;\alpha,\beta)$, the renewal rate for the Weibull process, was employed to improve numerical efficiency. Namely, it is shown in the Appendix that

$$h(t;\eta,\beta)dt = h(t/\eta;\beta)d(t/\eta) \tag{20}$$

in which $\eta = \alpha^{-1/\beta}$.

The average interval reliability therefore has the property

$$\begin{aligned} R(T;\eta,\beta) &= 1-n^{-1}\int_0^T h(t;\eta,\beta)dt \\ &= 1-n^{-1}\int_0^{T/\eta} h(t/\eta;\beta)d(t/\eta) \\ &= R(T/\eta;\beta). \end{aligned} \tag{21}$$

Given a solution for equation (21) it is clear that for a different value of η, say η_1, that a value T_1 can be found such that

$$\begin{aligned} R(T;\eta,\beta) &= 1-n^{-1}\int_0^{T_1/\eta_1} h(t/\eta_1;\beta)d(t/\eta_1) \\ &= R(T_1;\eta_1,\beta) \end{aligned} \tag{22}$$

in which $T_1 = (\eta_1/\eta)T$.

Computational speed can consequently be greatly improved if the function $h(t;\eta,\beta)$ is generated separately for a finite number of β values within a predetermined range of significance or applicability and for some fixed η and T. That is, generate $h(t/\eta;\beta)$ as a function of t/η for various values of β. What in essence is actually generated upon integration of h is $n(1-z)$ as a function of T/η and β. Numerical methods were used to compute $h(t/\eta;\beta)$ [6] for the Weibull distribution as a function of β with the results being generated in tabular form. It is a simple matter then for given z, β and T to determine the corresponding η by table lookup with interpolation. Given η one can determine $\alpha = 1/\eta^\beta$ and hence solve equation (14):

$$\alpha = R_a^{-1}(z,\beta).$$

A similar procedure could be used to solve for the inverse of equation (6) where the approximation $F(\tau) \doteq 0$ is not used although more complicated numerical procedures would be required.

5. FREQUENCY INTERPRETATION OF THE BAYESIAN INTERVALS. In the application of Bayesian inferencing, difficulty arises in the interpretation of the derived intervals. This stems primarily from difficulties in objectively establishing the prior distribution to be used or its interpretation, particularly when little or no prior information exists. Although a number of papers have been written dealing with priors representing complete ignorance for single parameter problems (for example references [13-15]), little has been reported on choosing priors for multiparameter situations in which a function of the parameter is of interest.

A number of Monte Carlo computer trials were consequently conducted to determine frequency interpretation of the derived Bayesian intervals, if any. In these trials various sample sizes of failure times were generated from the Weibull distribution using different values of the true R_a, designated as R_{at}, and shape parameter β. The approximate R_a, equation (7), was used in this study. Suspension times were not generated in these particular trials. System life was fixed at unity with the number of missions over system life being fixed at 150. This number of missions was chosen to yield mission times τ much less than the mean time between failure required for the approximation used and to represent an actual system testing problem of current interest to this writer. The parameter values used in this study are summarized in Table I.

TABLE I: PARAMETER VALUES USED IN MONTÉ CARLO STUDY OF BAYESIAN INTERVALS.

R_{at} (True Reliability)	β (Shape Parameter)	Mean Time Between Failures
0.99	1.0	0.661
	3.0	0.512
	6.0	0.494
0.95	1.0	0.113
	3.0	0.126
	6.0	0.125
0.90	1.0	0.067
	3.0	0.065
	6.0	0.065

Number of Monte Carlo Trials Per Case = 1000
System Life = 1.0
Number of Missions = 150
Mission Time = 0.007
Number of Failures = 5, 10, 20, 30

In checking for exactness, 1000 samples each of size n_f were generated and, for each sample, the value of F(z) given by equation (18) was evaluated at z equal to R_{at}, the true value of R_a. If the posteriori distribution is exact in a frequency sense, then the distribution of $F(R_{at})$ for the 1000 generated samples should be uniform on the interval (0,1) [16].

Two statistical tests were used to check exactness. The first was the K-S test [17] in which the distribution of $F(R_{at})$ was hypothesized to be uniform. The second was a binomial test [17] for the lower and upper 90% probability limits. In the binomial test, for example, if the lower 90% limit is exact then an average of 10% of the Monté Carlo samples should yield values of $F(R_{at})$ less than 0.1. Significance levels of 1 to 20% were used in both of these hypothesis tests to draw the conclusions made.

Results of the various trials conducted indicate that the generated Bayesian intervals generally are not exact. To investigate degree of exactness (or inexactness) the function $F(R_{at}-\varepsilon)$ was investigated for exactness. In this instance if ε is small relative to R_{at} then the intervals can be judged to be nearly exact for practical purposes.

Table II summarizes the results of the Monté Carlo trials performed. As can be seen from the results given in this table, the Bayesian intervals are nearly exact at all confidence levels for sample sizes of 20 or greater. The lower and upper confidence limits are nearly exact for all of the cases considered except for the case of sample size equal to 5. In this instance, the lower 90% limits are nearly exact.

6. CONCLUSIONS. The results indicate the feasibility of computing Bayesian intervals for average interval reliability which are nearly exact in a frequency sense. It should be noted, however, that a limited number of cases were studied and this general conclusion cannot be made for all conditions not considered. Intuitively, the approach used is a sound one in that frequency interpretation is considered. It seems reasonable that Bayesian intervals can be found for many similar problems which can at least be made conservative through frequency studies similar to the one used for interval reliability.

Other methods for studying Bayesian inferencing include study of priors which yield frequency intervals and questions of optimizing derived intervals. Some difficulty is encountered in these approaches if classical confidence intervals are not available. Monté Carlo simulation can be used for such a study which is a subject for future efforts.

TABLE II: Results of Monté Carlo Trials to Study Frequency Exactness of Bayesian Intervals for Interval Reliability.

SAMPLE SIZE, n_f	SHAPE PARAMETER, β	TRUE RELIABILITY, R_{at}	ϵ	K-S*	p(0,.1)**	p(.9,1.0)**
5	3.0	0.99	0.0	0.103	0.138	0.196
			-0.0005	0.154	0.104	0.236
			0.001	0.136	0.225	0.115
		0.95	0.0	0.105	0.123	0.199
			-0.001	0.114	0.101	0.226
			0.003	0.120	0.204	0.124
		0.90	0.0	0.170	0.123	0.195
			-0.002	0.149	0.100	0.228
			0.005	0.107	0.190	0.137
10	3.0	0.99	0.0	0.059	0.109	0.137
			-0.0002	0.096	0.096	0.164
			0.0005	0.093	0.168	0.081
		0.95	0.0	0.081	0.100	0.145
			0.001	0.051	0.134	0.112
		0.90	0.0	0.084	0.101	0.144
			0.002	0.050	0.138	0.110
20	1.0	0.99	0.0	0.072	0.082	0.140
			0.0005	0.023	0.112	0.102
		0.95	0.0	0.109	0.075	0.164
			0.002	0.044	0.110	0.125
			0.004	0.065	0.142	0.097
		0.90	0.0	0.110	0.073	0.167
			0.002	0.077	0.088	0.149
			0.005	0.032	0.118	0.119
	3.0	0.99	0.0	0.056	0.088	0.129
			0.0001	0.029	0.094	0.112
		0.95	0.0	0.070	0.084	0.137
			0.0005	0.031	0.103	0.109
		0.90	0.0	0.070	0.082	0.138
			0.001	0.029	0.104	0.110
	6.0	0.99	0.0	0.075	0.084	0.140
			0.0001	0.024	0.099	0.106
		0.95	0.0	0.063	0.087	0.130
			0.0002	0.031	0.095	0.113
		0.90	0.0	0.069	0.087	0.134
			0.0005	0.027	0.098	0.108
30	3.0	0.99	0.0	0.040	0.098	0.109
			0.0001	0.029	0.115	0.100
		0.95	0.0	0.054	0.090	0.121
			0.0003	0.026	0.111	0.106
		0.90	0.0	0.056	0.090	0.122
			0.0005	0.030	0.109	0.108

TABLE II: Results of Monté Carlo Trials to Study Frequency Exactness of Bayesian Intervals for Interval Reliability.

(Continued)

*K-S = Kolomogorov-Smirnov statistic

Reject hypothesis that distribution of $F(R_{at}-\epsilon)$ is uniform if K-S > theoretical value where

THEORETICAL VALUE	SIGNIFICANCE LEVEL
0.034	0.20
0.039	0.10
0.052	0.01

**p(0,.1) and p(.9,1.0) = Proportion of Monté Carlo trials for which $F(R_{at}-\epsilon) \leq 0.1$ and ≥ 0.9 respectively

Reject hypothesis that lower or upper 90% probability levels for $F(R_{at}-\epsilon)$ are exact if p(0,.1) or p(.9,1.0) lies outside theoretical interval where

THEORETICAL INTERVAL	SIGNIFICANCE LEVEL
0.092-0.108	0.20
0.088-0.112	.10
0.078-0.122	.01

7. REFERENCES.

[1] E. Pieruschka, Principles of Reliability, Englewood Cliffs, New Jersey: Prentice-Hall, 1963, Chap. 7.

[2] D. R. Cox, Renewal Theory, New York: Wiley, 1963.

[3] E. Parzen, Stochastic Processes, San Francisco: Holden-Day, 1962, Chap. 5.

[4] R. E. Barlow and F. Proschan, Mathematical Theory of Reliability, New York: Wiley, 1965, Chap. 3.

[5] B. V. Gnedenko, Yu. K. Belyayev and A. D. Solovyev, Mathematical Methods of Reliability Theory, New York: Academic Press, 1969, Chap. 2.

[6] R. L. Racicot, Numerical Solutions of Renewal and Interval Reliability Functions, submitted for publication to IEEE Transactions on Reliability, June 1973.

[7] G. Schlenker, Numerical Methods in Renewal Theory, OR-68-2, Weapons Operations Research Office, U.S. Army Weapons Command, Rock Island, Illinois, February 1968, available from Defense Documentation Center, Cameron Station, Alexandria, Virginia 22314.

[8] R. E. Bellman, A New Approach to the Numerical Solution of a Class of Linear and Nonlinear Integral Equations of Fredholm Type, RM-4762-PR, The RAND Corp., Santa Monica, California, October 1966, AD640687, available from NTIS, Springfield, Virginia 22151.

[9] W. L. Smith and M. R. Leadbetter, "On the Renewal Function for the Weibull Distribution," Technometrics, 5, 1963, pp. 393-396.

[10] D. V. Lindley, "Introduction to Probability and Statistics from a Bayesian Viewpoint," Part 1: Probability and Part 2: Inference, Cambridge, Mass.: Cambridge University Press, 1965.

[11] A. Clifford Cohen, "Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and On Censored Samples," Technometrics, Vol. 7, No. 4, November, 1965, pp. 579-588.

[12] A. Papoulis, Probability, Random Variables, and Stochastic Processes," New York: McGraw-Hill, 1965.

[13] E. T. Jaynes, "Prior Probabilities," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, No. 3, Sept. 1968, pp.227-241.

[14] D. J. Bartholomew, "A Comparison of Some Bayesian and Frequentist Inferences," Biometrika, Vol. 52, pp. 19-35.

[15] A. R. Thatcher, "Relationships Between Bayesian and Confidence Limits for Predictions," Journal of the Royal Statistical Society, Vol. 26, No. 2, 1964, pp. 174-192.

[16] E. Parzen, Modern Probability Theory and Its Application, New York: Wiley, 1960, p. 313.

[17] S. Siegel, Nonparametric Statistics for the Behavioral Sciences, New York: McGraw-Hill, 1956, Chap. 4.

7. APPENDIX.

In this appendix, the following two properties required for development of the Bayesian probability limits on interval reliability are derived for the Weibull distribution:

(I) $h(t;\eta,\beta)dt = h(t/\eta;\beta)d(t/\eta)$

(II) $R_a(\alpha,\beta) \leq R_a(\alpha^*,\beta)$

whenever $\alpha \geq \alpha^*$ where $\alpha = 1/\eta^\beta$.

The renewal rate $h(t;\eta,\beta)$ satisfies the relation

$$h(t;\eta,\beta)dt = f(t;\eta,\beta)dt + \int_0^t h(x;\eta,\beta)f(t-x;\eta,\beta)dxdt \qquad \text{(A-1)}$$

in which $f(t;\eta,\beta)$ is the Weibull distribution given by

$$f(t;\eta,\beta)dt = (\beta/\eta)(t/\eta)^{\beta-1} \exp(-(t/\eta)^\beta)dt. \qquad \text{(A-2)}$$

Since η is a scale parameter

$$f(t;\eta,\beta)dt = f(t/\eta;\beta)d(t/\eta) \qquad \text{(A-3)}$$

and

$$F(t;\eta,\beta) = F(t/\eta;\beta). \qquad \text{(A-4)}$$

It follows directly then from the relation

$$h(t/\eta;\beta)d(t/\eta) = f(t/\eta;\beta)d(t/\eta) + \int_0^{t/\eta} h(x/\eta;\beta)f((t-x)/\eta;\beta)\, d(x/\eta)\, d(t/\eta) \qquad \text{(A-5)}$$

that (I) holds by virtue of equation (A-3) since solution of the renewal equation in this instance is unique.

The relation (II) can be shown for the approximate reliability given by equation (7) using the property (I). The average reliability can be written as

$$\begin{aligned} R_a(T;\eta,\beta) &= 1-\eta^{-1}\int_0^T h(t;\eta,\beta)dt \\ &= 1-\eta^{-1}\int_0^{T/\eta} h(t/\eta;\beta)d(t/\eta) \\ &= 1-\eta^{-1}\int_0^{T/\eta} h(y;\beta)dy \end{aligned} \qquad \text{(A-6)}$$

Since h is a positive function, the integral in equation (A-6) is monotonically increasing with decreasing η where T and β are held fixed. Hence $R_a(\alpha,\beta)$ is monotonically decreasing with increasing $\alpha = \eta^{-\beta}$, thus establishing property (II).

RELIABILITY OF A SERIAL SYSTEM

A. E. Johnsrud
U.S. Army Concepts Analysis Agnecy
Bethesda, Maryland

Introduction

A serial system is an assembly of subsystems of such form that the failure of any subsystem will constitute a failure of the entire system (e.g. artillery round and fuze). The reliability, as used in this paper, is one minus the probability of failure in one trial. The question to be answered concerns the confidence level that is to be associated with a given reliability for the system when the reliability-confidence relationship has been established for each of the subsystems through testing. A rigorous derivation is presented showing how the confidence level can be calculated exactly for any number of subsystems when the subsystem test results (trials and failures) are completely general; i.e., any number of trials and failures may be specified for any subsystem.

Confidence Level for a Single System

A single system (which may later be a subsystem of a serial system) is tested to determine its reliability. The test is a set of Bernoulli trials (pass or fail on each trial) and the result is m failures out of n trials. The confidence level for a particular reliability (R) is called C. It is defined to be the probability that the reliability exceeds R. How its value depends upon R, m, and n is shown by the following argument:

Suppose that the reliability is precisely r. Then the probability of obtaining m failures in n trials is

$$P_{m,n} = \frac{n!}{m!(n-m)!}\, r^{n-m}\,(1-r)^m = \binom{n}{m}\, r^{n-m}\,(1-r)^m$$

This can be converted to a probability distribution function by multiplying by the probability that the reliability is between r and r+dr. Since no prior knowledge about the reliability exists, all probabilities between 0 and 1 are equally likely and the probability that the reliability is between r and r+dr is simply dr. Thus the probability distribution function for the system is

$$dP_{m,n}/dr = \binom{n}{m} r^{n-m} (1-r)^m (n+1).$$

The factor (n+1) is for normalizing, to make the integral over $0 \leq r \leq 1$ equal to unity.

The confidence level for a given reliability, R, is the integral of the probability distribution function from R to 1. Although the general expression can be integrated, the treatment is simpler if one proceeds stepwise, letting the values of m increase from zero.

For m = 0 (no failures),

$$C = (n+1)(1)\int_R^1 r^n \, dr = (n+1) \cdot (1-R^{n+1})/(n+1)$$

$$= 1 - R^{n+1}$$

For m = 1,

$$C = (n+1)(n)\int_R^1 r^{n-1}(1-r) \, dr = (n+1)(n)\int_R^1 (r^{n-1} - r^n) \, dr$$

$$= 1 - R^{n+1} + R^{n+1} - (n+1)R^n + nR^{n+1}$$

$$= 1 - R^{n+1} - (n+1)R^n(1-R)$$

For m = 2,

$$C = (n+1)(n)(n-1)(\tfrac{1}{2})\int_R^1 r^{n-2}(1-r)^2\,dr$$

$$= (n+1)(n)(n-1)(\tfrac{1}{2})\left[\frac{r^{n-1}}{n-1} - \frac{2r^n}{n} + \frac{r^{n+1}}{n+1}\right]_R^1$$

$$= 1 - (\tfrac{1}{2})(n)(n+1)R^{n-1} + (n^2-1)R^n - (\tfrac{1}{2})(n)(n-1)R^{n+1}$$

$$= 1 - R^{n+1} - (n+1)R^n(1-R) + R^{n-1}\left[R^2 + (n+1)R(1-R)\right.$$
$$\left. - (\tfrac{1}{2})(n)(n+1) + (n^2-1)R - (\tfrac{1}{2})(n)(n-1)R^2\right]$$

$$= 1 - R^{n+1} - (n+1)R^n(1-R) - (n+1)(n/2)R^{n-1}(1-R)^2.$$

(Many algebraic steps have been left out.)

A general expression for C is

$$C_{m,n} = 1 - \sum_{j=0}^{m}\binom{n+1}{j}R^{n+1-j}(1-R)^j \quad .$$

This says that, given m failures in n trials, the confidence that the reliability exceeds R is one minus the probability that there are <u>at most</u> m failures in <u>n+1</u> trials when the reliability is precisely R. A check on the validity of the expression can be made by calculating the confidence level for an R of 0.5 when half the trials result in misses. Symmetry requires that the confidence be 0.5 also. If n = 2m, we have

$$C_{m,2m} = 1 - (\tfrac{1}{2})^{2m+1}\sum_{j=0}^{m}\binom{2m+1}{j} = 1 - (\tfrac{1}{2})^{2m+1}(\tfrac{1}{2})(2)^{2m+1}$$

$$= \tfrac{1}{2} \quad ,$$

which is the desired result. It is pointed out here that the literature on this subject uses a different equation for the confidence level, one in which n is used instead of (n+1). Tables calculated from that equation show the 50% confidence level for R = 0.5 occurring at n = 2m+1, so that the absence of symmetry is evident.

The expression for $C_{m,n}$ is valid for any n, including zero, and for m going from 0 to n. It is symmetric for m values on either side of n/2.

For applications in which it may be desirable to make use of a single reliability-related figure, instead of the reliability-confidence pair, the concept of the <u>average reliability</u> is introduced. By average reliability is meant the average of the reliability over all confidence intervals. It is very simply expressed in terms of m and n, being

$$\bar{R} = \frac{n+1-m}{n+2}$$

The derivation follows from

$$\bar{R} = \int_{C=0}^{1} R\, dC = \int_{C=0}^{1} [d(RC) - C\, dR]$$

$$= RC \Big|_{C=0}^{1} - \int_{R=1}^{0} C\, dR = C + \int_{R=0}^{1} C\, dR$$

The function $C_{m,n}$ is known from above and one can choose some small values of m and n to demonstrate the resultant $\bar{R}$. The average reliability can probably be used everywhere that the "point" estimate, (n-m)/n, is now used; it agrees with the latter for reliabilities near ½ and for n large. It gets

away from the difficulties that arise when there are zero or n failures. $\bar{R}$ calculated for a serial system is the product of the $\bar{R}$'s for the subsystems, as will be demonstrated later.

The Two-Subsystem System

The confidence for reliability exceeding R in a system which is composed of subsystem A and subsystem B is determined through calculations in three dimensional space. Consider an x, y, z coordinate system in which x and y are the reliabilities of subsystems A and B (R_A and R_B) and the z coordinate, the dependent variable, is the product of the probability distribution functions for R_A and R_B. (These distribution functions arise from the testing of A and B; they are given in the previous section.) The z coordinate now represents the joint probability distribution function for the two subsystems.

The reliability for the composite system is defined as $R_A R_B$. A change in variables can now be made in such a way that this product $R_A R_B$ becomes one of the new coordinates. Curves designating constant values of $R_A R_B$ compose a family of hyperbolas in the space of R_A, R_B. Orthogonal to that set is another set of hyperbolas and these two sets make up the new coordinate system. The new coordinates are

$$u = R_A R_B$$

$$v = (\tfrac{1}{2})(R_A^2 - R_B^2) .$$

The confidence level associated with a given composite system reliability R_C is found by integrating over the plane from $u = R_C$ to $u = 1$. A diagram of the R_A, R_B plane over which the

integration is to be done is shown on Figure 1.

Associated with the transformation is a Jacobian (J) whose purpose is to adjust for the stretching or shrinking of the size of the area element ($dR_A dR_B$). This Jacobian has the value

$$J = \tfrac{1}{2}(u^2 + v^2)^{-\frac{1}{2}}$$

The product (f) of the distribution functions is a function of R_A and R_B, which can be changed over to a function of u and v through the relations

$$R_A = \left[v + (u^2 + v^2)^{\frac{1}{2}}\right]^{\frac{1}{2}}$$

$$R_B = u\left[v + (u^2 + v^2)^{\frac{1}{2}}\right]^{-\frac{1}{2}}$$

so that $f(R_A, R_B) = f(u,v)$.

This leads to the following expression for the confidence that the reliability of the composite system exceeds R_C:

$$C = \int_{u=R_C}^{1} \int_{v=-\frac{1}{2}(1-u^2)}^{\frac{1}{2}(1-u^2)} f(u,v)(\tfrac{1}{2})(u^2 + v^2)^{-\frac{1}{2}}\, dv\, du$$

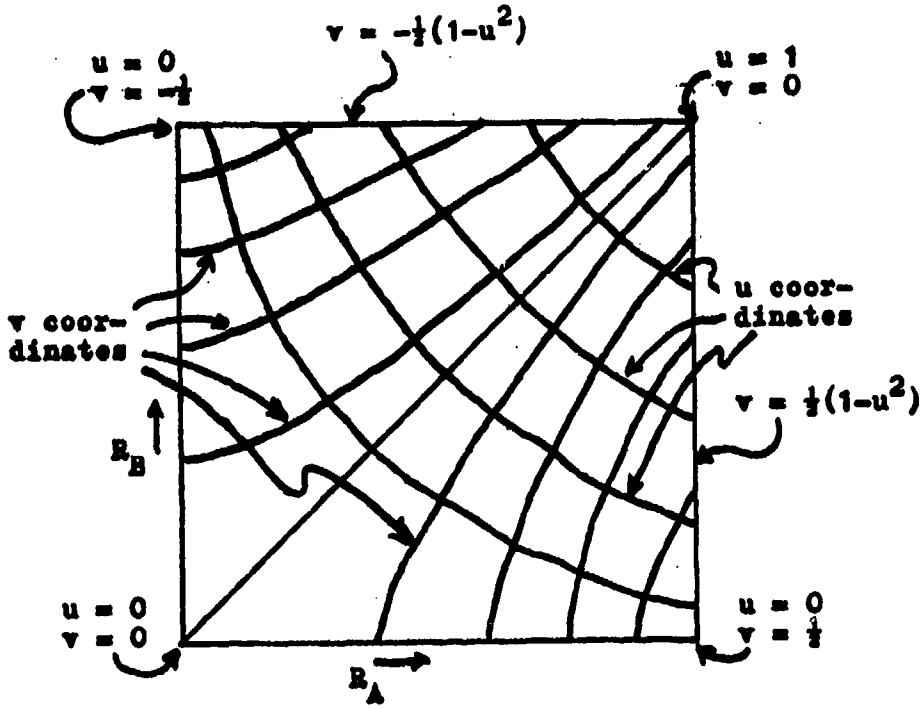


FIGURE 1. Region of Integration

It is assumed here that the distribution function is always normalized. Verifying that C approaches unity as R_C approaches zero is a good check for mistakes in the calculation.

The extension to any number of subsystems proceeds step-by-step, with a pair of subsystems being done first to yield C as a function of R_C for the pair. Differentiating this function will produce a probability distribution function which is to be paired with the next subsystem. The process continues until all subsystems are used.

Illustrative Example

Three subsystems produce respective test results of 1 failure out of 5 trials, 2 failures out of 6 trials, and 0 failures out of 10 trials. Labelling the first two of these subsystems A and B, for the moment, we have a joint probability distribution function

$$f = (6)(7)(5)R_A^4(1-R_A)(6)(5)(\tfrac{1}{2})R_B^4(1-R_B)^2$$

$$= (42)(75)\, u^4\, (1-R_A)(1-2R_B+R_B^2)$$

At this point it would be helpful to list some evaluated forms for the inner integral, forms which occur frequently. These are the following:

$$\int_{-\frac{1}{2}(1-u^2)}^{\frac{1}{2}(1-u^2)} \tfrac{1}{2}(u^2+v^2)^{-\frac{1}{2}}\, dv = -\ln u$$

$$\int_{-\frac{1}{2}(1-u^2)}^{\frac{1}{2}(1-u^2)} \tfrac{1}{2}(u^2+v^2)^{-\frac{1}{2}}\, R_A^n\, dv = (1/n)(1-u^n) \quad \text{for } n \neq 0$$

(and similarly for R_B)

$$\int_{-\frac{1}{2}(1-u^2)}^{\frac{1}{2}(1-u^2)} \tfrac{1}{2}(u^2+v^2)^{-\frac{1}{2}} R_A^{\,n} \ln R_A \, dv = (-1/n^2)(1-u^n+nu^n \ln u) \quad \text{for } n \neq 0 \text{ (and similarly for } R_B)$$

$$\int_{-\frac{1}{2}(1-u^2)}^{\frac{1}{2}(1-u^2)} \tfrac{1}{2}(u^2+v^2)^{-\frac{1}{2}} \ln R_A \, dv = -(\ln u)^2 \quad \text{(and similarly for } R_B).$$

Returning now to f, we can write it as

$$f = (42)(75)\, u^4\, (1-2R_B+R_B^{\,2}-R_A+2u-uR_B)$$

which we can rewrite for integration purposes (after noting that R_A and R_B behave the same under the integration) as follows:

$$f = (42)(75)\, u^4\, (1-3R-uR+R^2+2u)$$

The integral over v, after f is multiplied by the Jacobian, then becomes

$$(42)(75)\left[u^4(1+2u)(-\ln u) - (5/2)\, u^4 + 2u^5 + \tfrac{1}{2}u^6\right].$$

Integrating this over u, from R_C to 1, gives

$$C = (42)(75)\left\{ -\ln u \left[(u^5/5 + (u^6/3)\right] - (23/50)u^5 + (7/18)u^6 + (1/14)u^7 \Big|_{R_C}^{1} \right\}$$

$$= 1 + (42)(15)R_C^{\,5} \ln R_C + (42)(25)R_C^{\,6} \ln R_C + (21)(69)R_C^{\,5} - (49)(25)R_C^{\,6} - 3(75)R_C^{\,7}$$

[The rest of this paper will finish this example. The final expression for the confidence level of the three-subsystem composite becomes

$$C = 1 + (15)(77)(R^5 + 2R^6) \ln R + (7/4)R^{11} - (77)(33)\cdot$$

$$(1/4)R^7 - (77)(29)R^6 + (77)(37)R^5 .$$

The E calculated by integrating this expression is

$$(25)(11)/(56)(12) ,$$

the same as one obtains by multiplying together the three subsystem E's.]

THE ROLE OF INSTRUMENT PERFORMANCE STANDARDS IN THE CONTROL OF THE WHITE SANDS MISSILE RANGE TEST SUPPORT PROCESS

J. B. GOSE
W. R. JENKINS
QUALITY ASSURANCE OFFICE
US ARMY WHITE SANDS MISSILE RANGE
NEW MEXICO

1. INTRODUCTION. The WSMR Plant Process has control inputs of data requirements, instrument assignments, inspection results, instrument errors, and final data errors. Prior to feedback of the performance quality data to management, the data are compared with appropriate standards and control limits. The percent out of control are calculated and the out of control processes are tagged for management action. This process has as plant control variables:

a. Number of stations in solution - n

b. Geometry - $(H^tH)^{-1}$, from the Error Equation $\sigma_z^2 = (H^tH)^{-1}\sigma_x^2$

c. Sample Rate - s

d. Points in Filter - N

2. STANDARDS. The instrument performance standards are the cumulative frequency curves of observed performance for each System. Examples are shown in Figures 1 through 8. The control limits (indicated by arrows on these figures) are established through consideration of such factors as: the resources available for follow-up actions, the effect of error values and the geographical location of instruments in the solution process, and the frequency of use of the instruments. The upper control limit for measurement error is used for support planning purposes, and the lower control limit for data loss is used for support planning. The WSMR conventions for error type are included on the figures, these are defined as follows:

"σ" is used to denote error estimators computed by hybrid (two or more different Instrumentation Systems) solutions.

$$\sigma_{x_i}^2 = \frac{\Sigma(x_i-\mu)^2}{df} = \frac{\Sigma(\Delta_{SYS} + \Delta_{INST} + \epsilon_i)^2}{df} = \xi[(x_i-\mu)^2]$$

"S" is used to denote error estimators computed by an N-Station Single-Instrument System solution.

$$S_{x_i}^2 = \frac{\Sigma(\Delta_{INST} + \epsilon_i)^2}{df} = \xi[(x_i - \mu - \Delta_{SYS})^2]$$

"E" is used to denote error estimators computed by variate difference or moving arc curve fitting techniques.

$$E_{x_i}^2 = \frac{\Sigma\epsilon_i^2}{df} = \xi[(x_i - \mu - \Delta_{SYS} - \Delta_{INST})^2]$$

3. EQUATIONS OF THE PLANT PROCESS. The performance of the plant process is measured by the number of requirements met for each test of each project. In terms of the probabilities involved, this performance may be defined as

$$P(RM) = P(T \geq T_R)\ P(\alpha \leq \alpha_R) < P(t \geq t_S)\ P(\beta \leq \beta_S) \qquad \text{Eq 1}$$

where P(RM) is the probability that a requirement is met

T is data coverage time

α is data error (σ, S, or E)

t is coverage from instrument records

β is instrument measurement error (σ, S, or E)

Subscript R is the requirement value

Subscript S is the standard or control limit value.

To make Eq 1 an equality, it is proper to introduce a factor "g". Thus Eq 1 becomes

$$P(RM) = P(T \geq T_R)\ P(\alpha \leq \alpha_R) = P(t \geq t_S)\ P(\beta \leq \beta_S)g \qquad \text{Eq 2}$$

where "g" is the probability that the planning is done correctly and is comprised of the plant variables previously cited i.e.,

$$g = P\{f(n, (H^tH)^{-1}, s, N, \ldots)\}$$

Consideration of the realities of operation at WSMR allows us to make the further simplifications that

$$P(\alpha \leq \alpha_R) = P(\beta \leq \beta_S)g \quad * \qquad \text{Eq 3}$$

which is to say that the error is independent of the time frame when measurements are made.

4. FORECASTING (SUPPORT PLANNING). The standards are used with Eq 2 to predict process performance. An approximate value of g is obtained by evaluating n, $(H^tH)^{-1}$, N, s for the instruments performance standards and process configuration planned. This approximation (g_o) is used in Eq 2.

$$P(t \geq t_S)\ P(\beta \leq \beta_S)g = P(T \geq T_R/2)\ P(\alpha \geq \alpha_R) \geq 0.68$$

Comparison of g_o to g (evaluated as in Paragraph 4) indicates the degree of operational conformance to the process support plan.

5. A CASE OF PLANT PROCESS PERFORMANCE. In order to evaluate equations for an actual case, the entire output from WSMR's plant process was analyzed for three months of tests. The probabilities found were:

$$P(\alpha \leq \alpha_R) = 0.89$$

$$P(\beta \leq \beta_S) = 0.95$$

The values of P(RM) were found for $P(T \geq k\ T_R)$ for an array of k from 0 through 100 percent. These values are shown on Figure 9. WSMR has a performance standard for P(RM) at 0.68, the standard is seen to occur at k = 50%. The proper expression for $P(T \geq T_R)$ is $P(T \geq T_R/2)$ which is in fact the expression in use.

*The coverages may be considered independently from Eq 2 using

$$P(T \geq T_R) = P(t \geq t_S)h$$

wherein h is the probability that a data record is not lost.

From the above data, we may evaluate the remaining probabilities in Equation 2. From Eq 3, we find

$$g = 0.68.$$

From Eq 2, we find

$$P(T \geq T_R/2) = 0.76$$

and

$$P(t \geq t_S) = 0.81.$$

CUMULATIVE DISTRIBUTION OF RADAR DATA LOSS

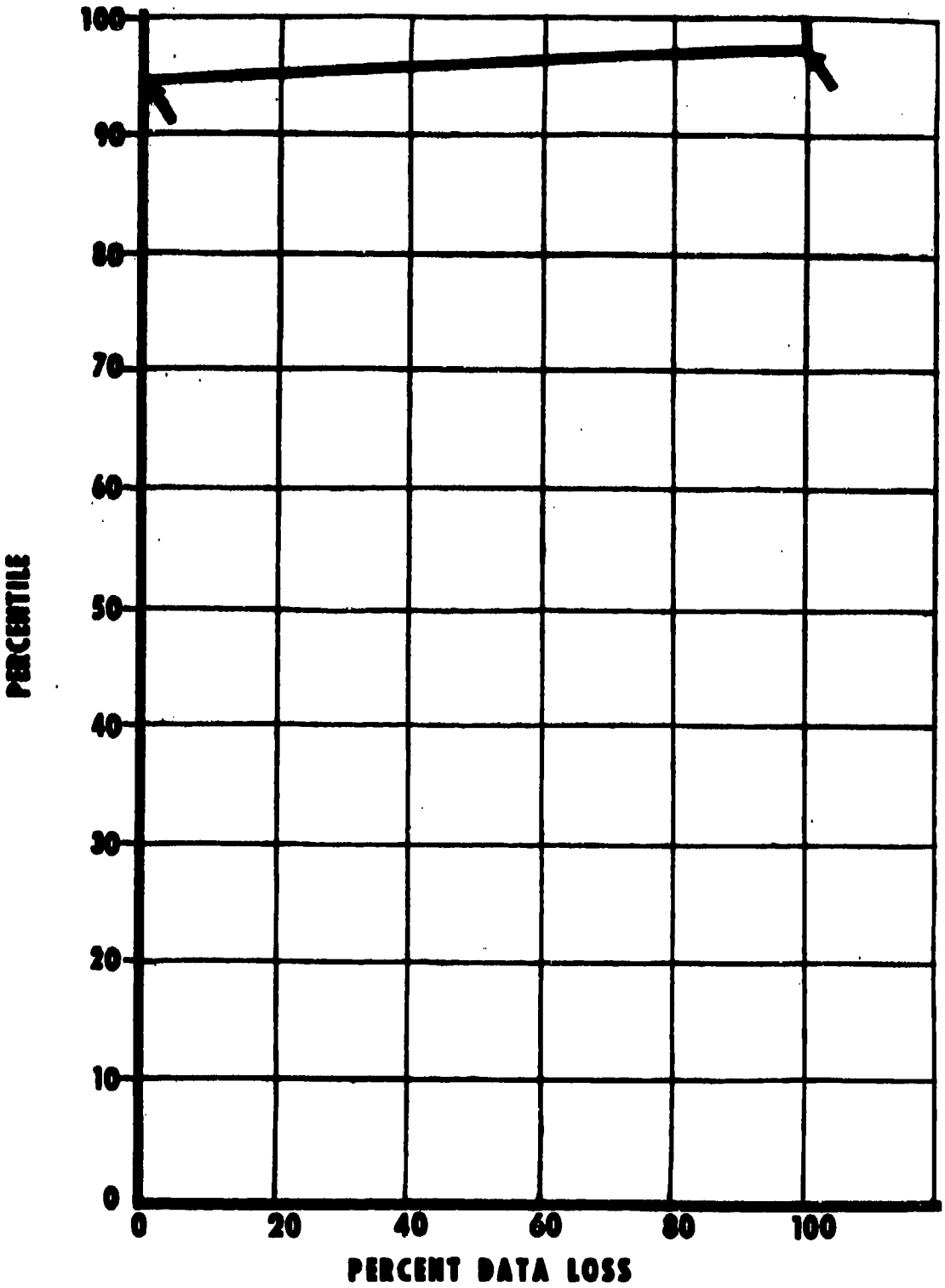


FIGURE 1

(TYPE E ESTIMATOR)
CUMULATIVE DISTRIBUTION OF RADAR AZIMUTH MEASUREMENT ERRORS

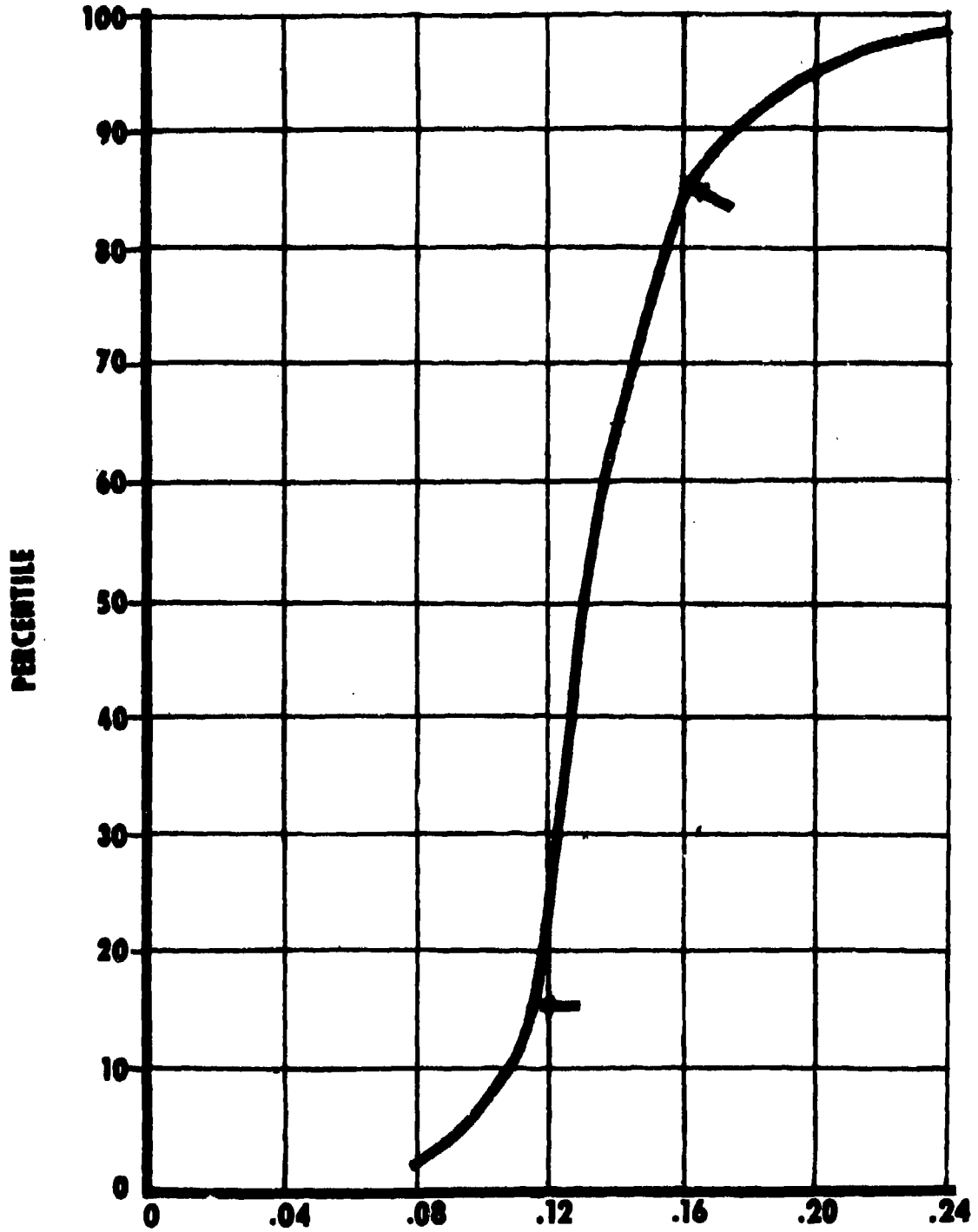


FIGURE 2

(TYPE E ESTIMATOR)
CUMULATIVE DISTRIBUTION OF RADAR ELEVATION MEASUREMENT ERRORS

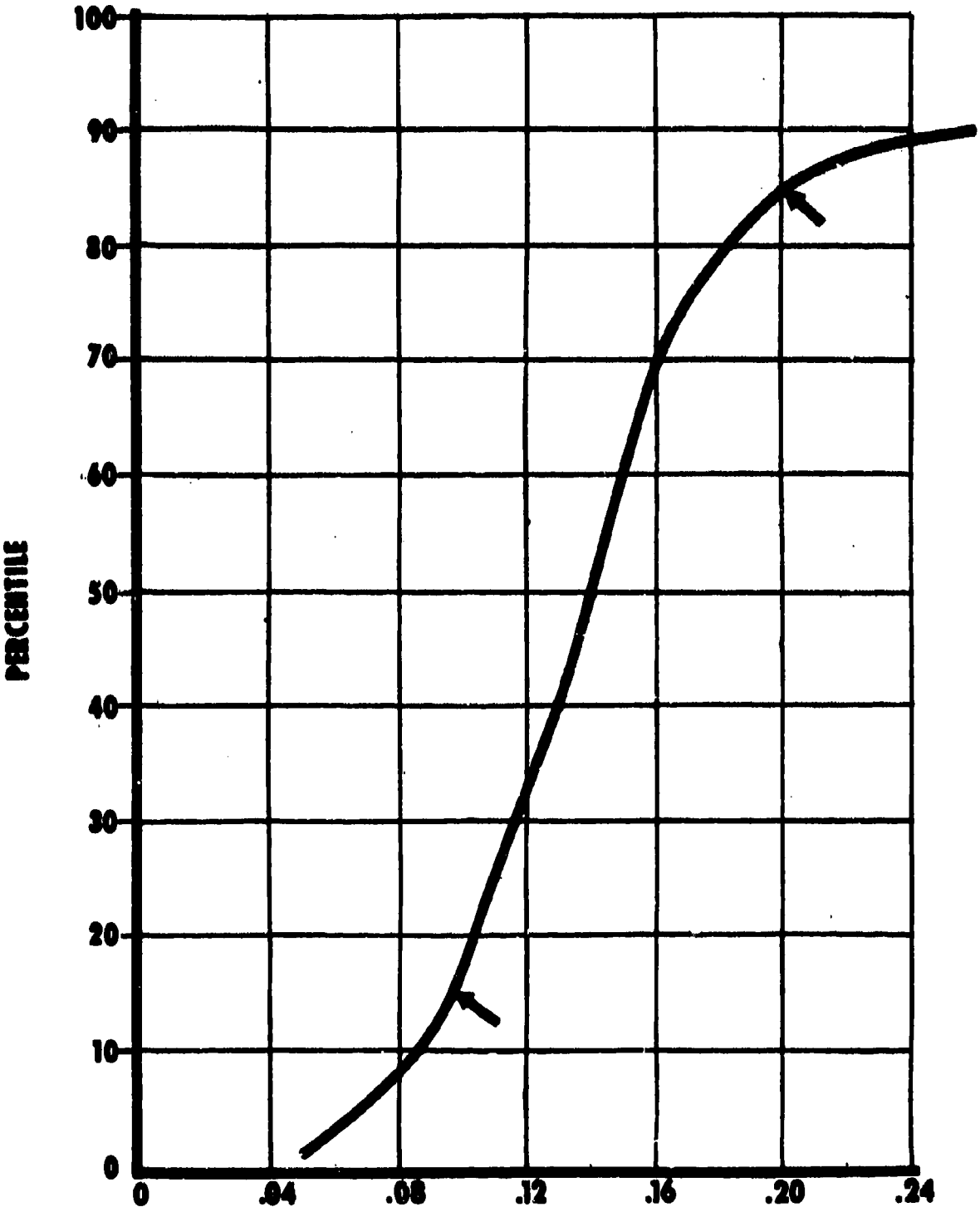


FIGURE 3

(TYPE E ESTIMATOR)
CUMULATIVE DISTRIBUTION OF RADAR RANGE MEASUREMENT ERRORS

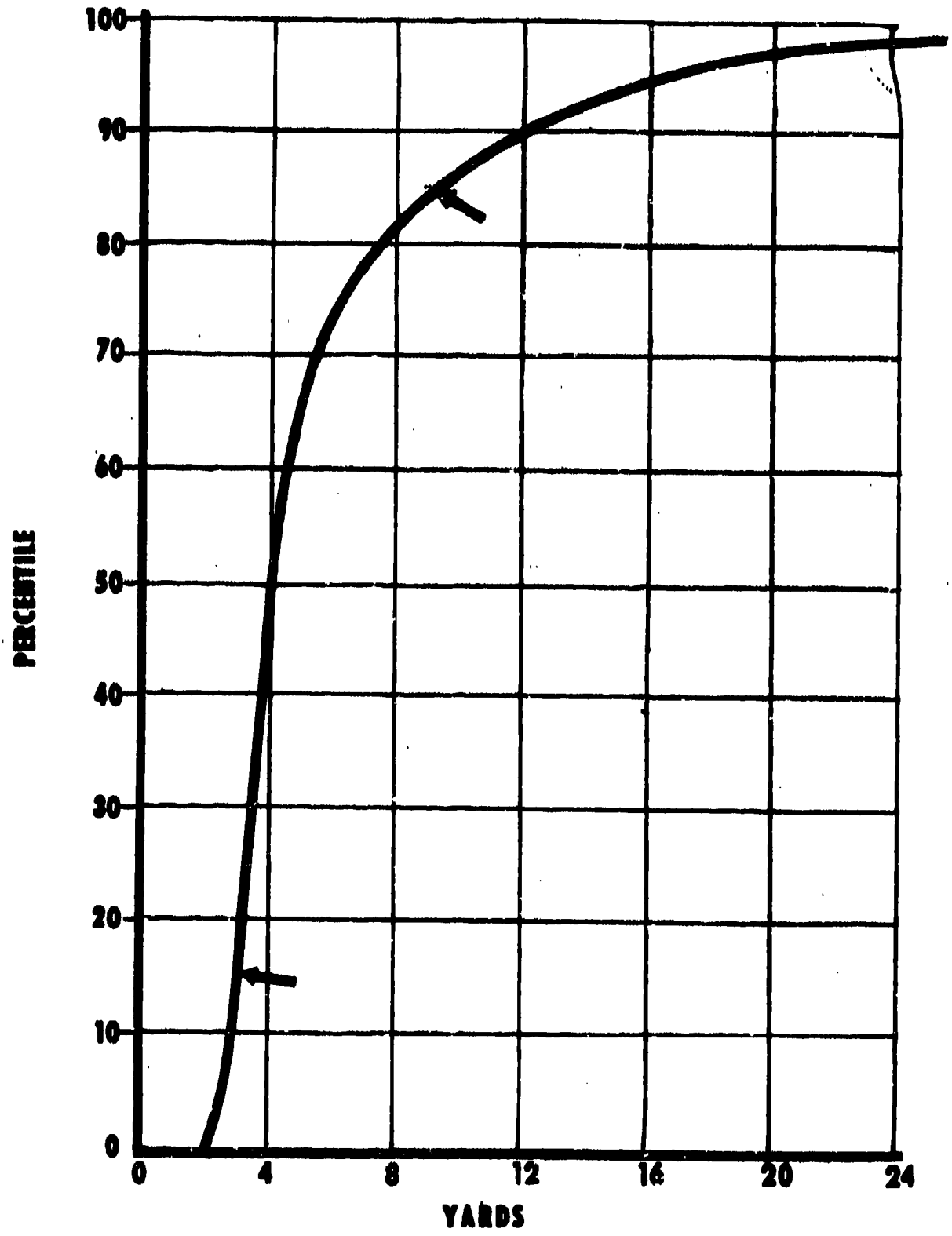


FIGURE 4

CUMULATIVE DISTRIBUTION OF ASKANIA DATA LOSS

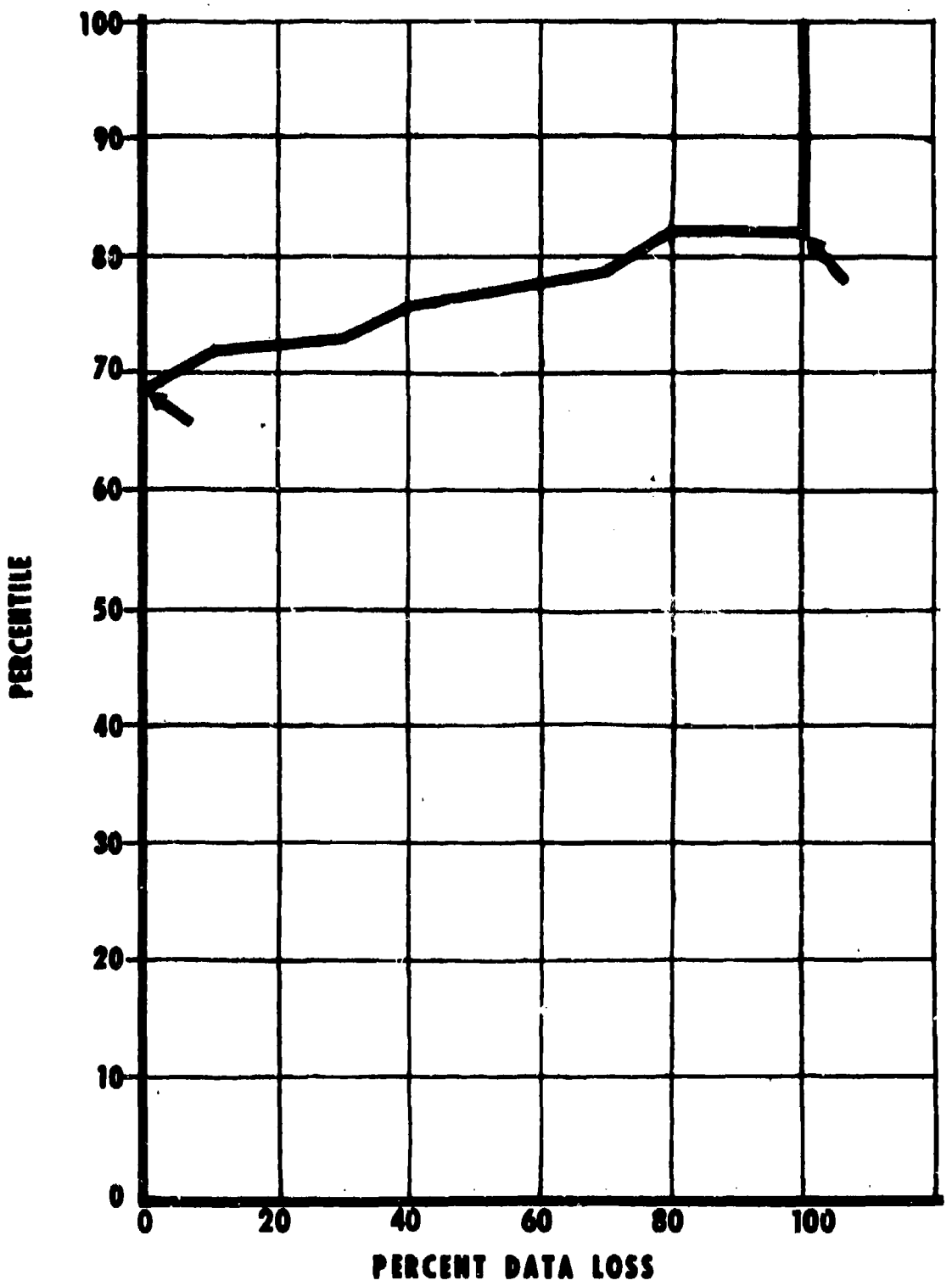


FIGURE 5

(TYPE S ESTIMATOR)
CUMULATIVE DISTRIBUTION OF ASKANIA AZIMUTH MEASUREMENT ERRORS

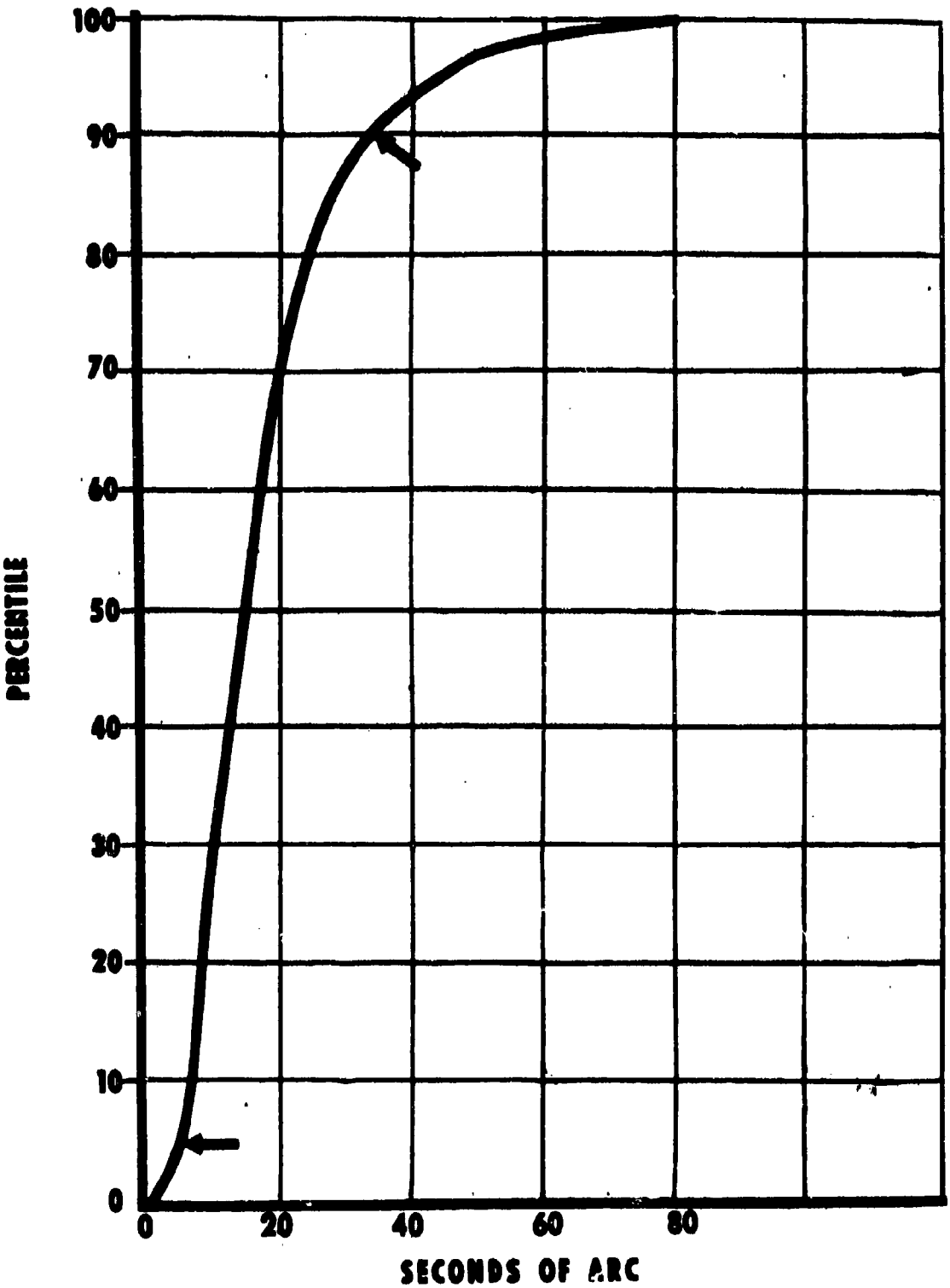


FIGURE 6

(TYPE S ESTIMATOR)
CUMULATIVE DISTRIBUTION OF ASKANIA ELEVATION MEASUREMENT ERRORS

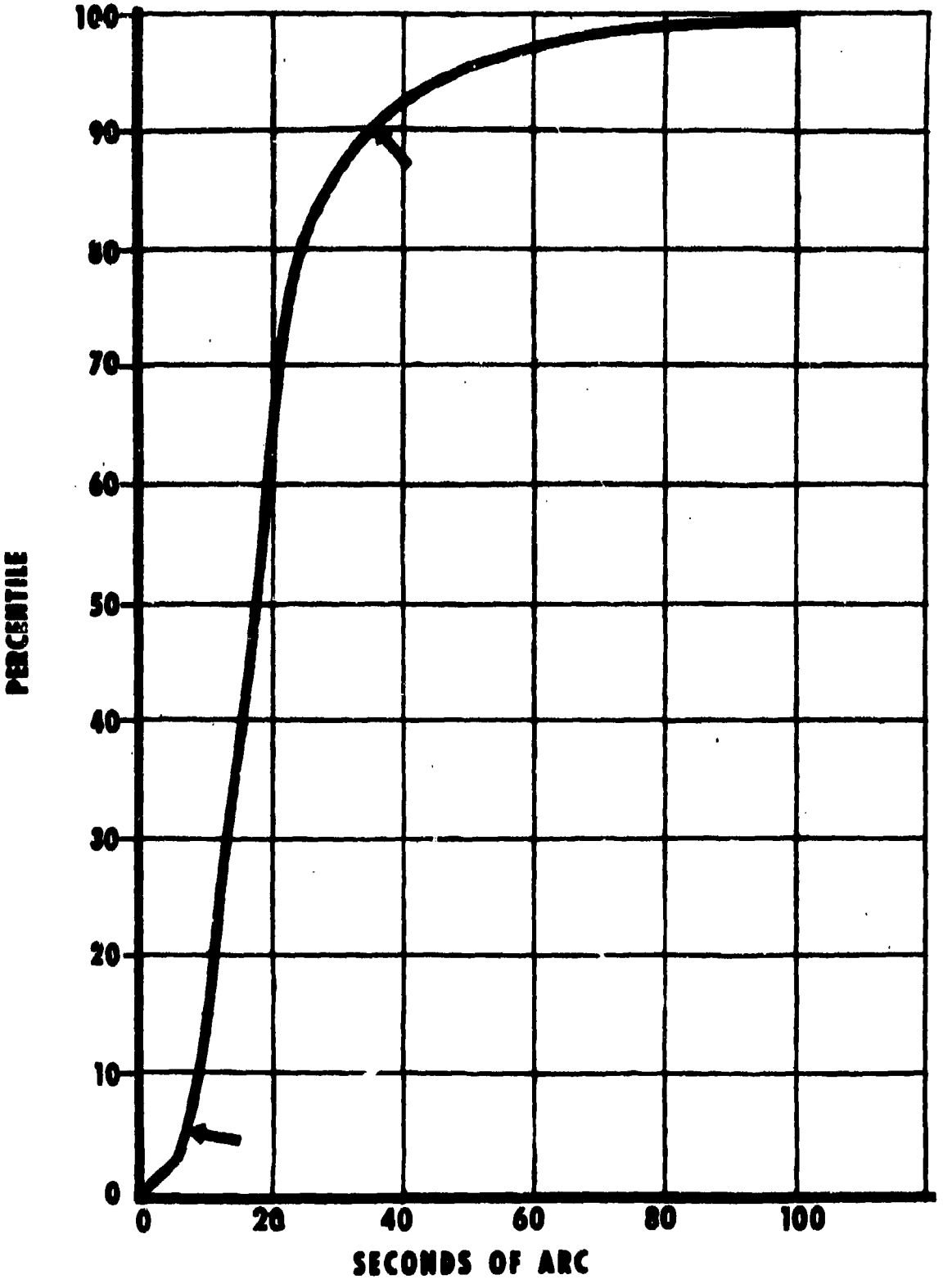


FIGURE 7

(TYPE σ ESTIMATOR)
CUMULATIVE DISTRIBUTION OF ASKANIA
V-ANGLE MEASUREMENT ERRORS

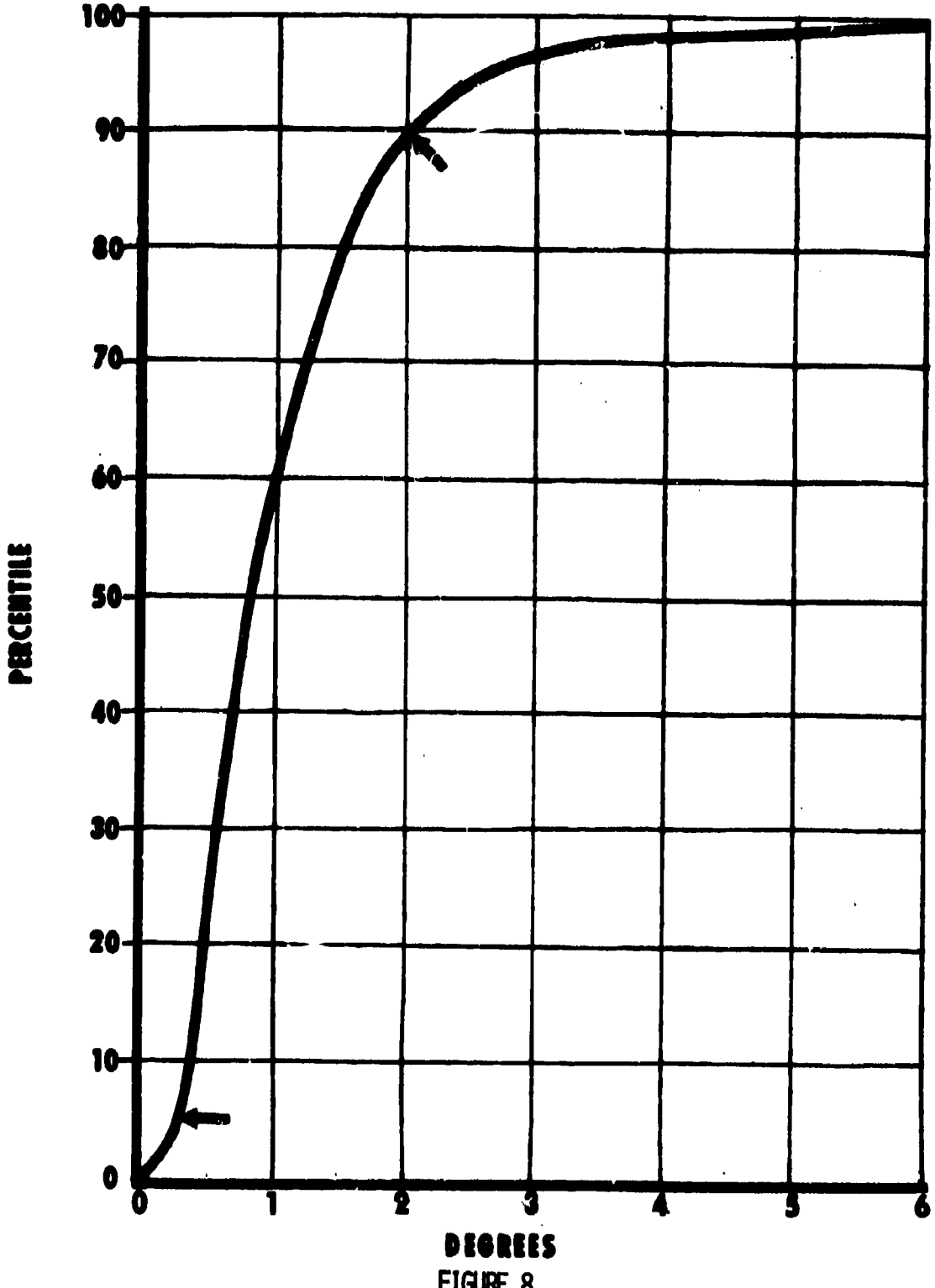


FIGURE 8

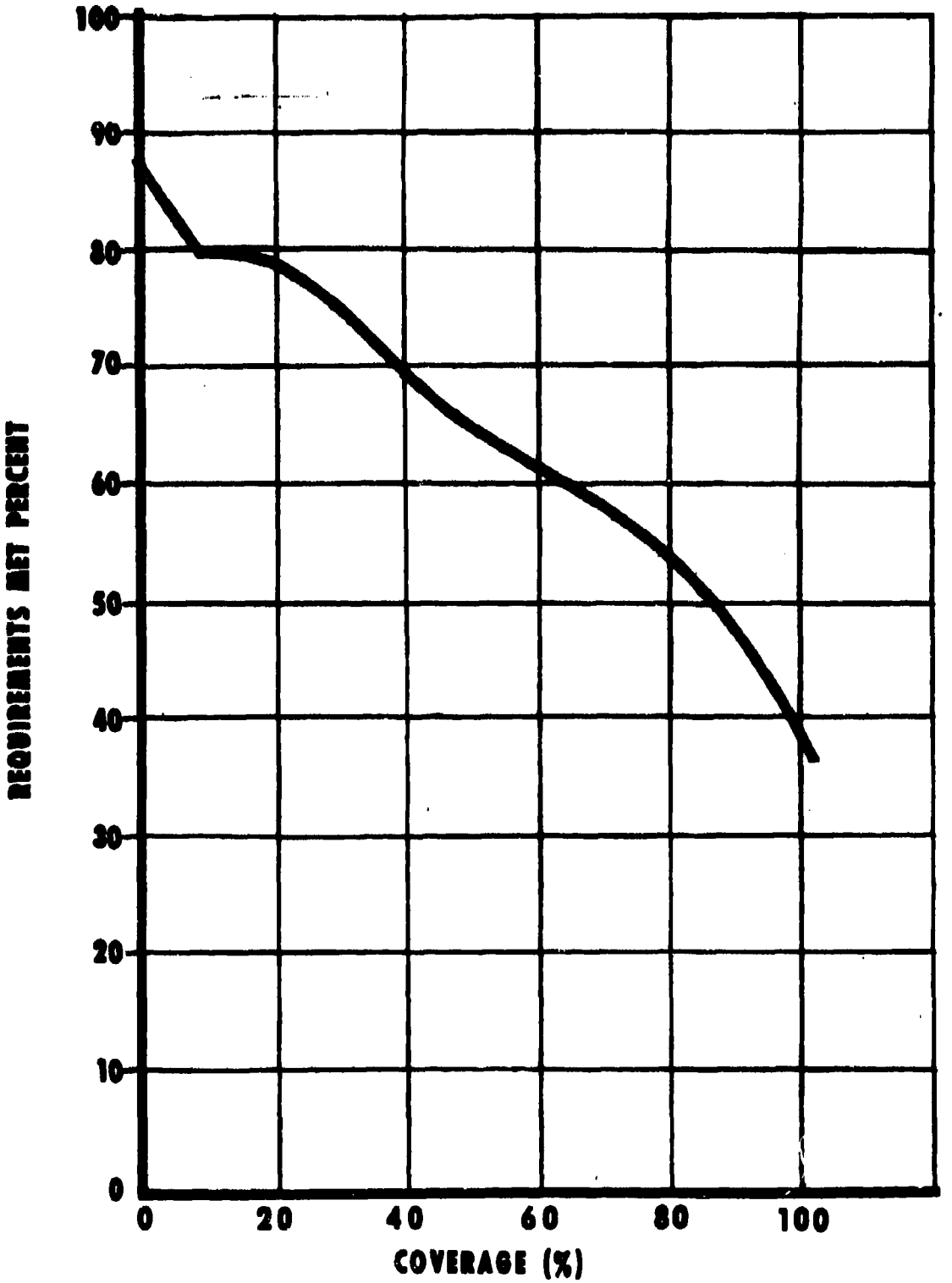

REQUIREMENTS MET PERCENT
100
90
80
70
60
50
40
30
20
10
0
0
20
40
60
80
100
COVERAGE (%)

FIGURE 9

MULTIVARIATE ANALYSIS TECHNIQUES APPLIED TO EQUIPMENT TESTING

John Bart Wilburn, Jr.
US Army Electronic Proving Ground
Fort Huachuca, Arizona 85613

ABSTRACT

This report discusses the general theory of multivariate eigenvector analysis and describes an application to equipment testing for developing a model specifying equipment performance and for constructing a climate model from the vantage point of equipment performance.

The purpose of this report is to outline a method of analyzing complex phenomena in an operating environment. The report discusses the general theory of the method of multivariate analysis and describes an application to equipment testing for developing a model specifying equipment performance and for constructing a climate model from the vantage point of equipment performance.

Method: The Method of analysis has been known by several names since its earliest forms of about 40 years ago described by H. Hotelling (1) as principal factor analysis and the more modern treatment of T. W. Anderson (2) as principal component analysis. Current terminology has been Principal Component Eigenvector Analysis, Empirical Orthogonal Functions or simply Multivariate Analysis.

It is of paramount importance that one point be made clear from the outset - that is, that this method of analysis does not imply any causal relationships. The relationships are purely coincidental with

an estimable determinancy and confidence. Causal relationships are to be explained in subsequent investigations of the coincidental relationships revealed by this analysis.

We can start with an array of observations of variables D^{l}_{nm} where l indexes the class of variables, e.g., location or type and the nxm array of n variables observed m times, i.e., m is the measure of time. For convenience, the l superscript will be dropped; furthermore, it is desirable that the n variables are of the same dimensions and similar variance. This latter constraint is assured if the array D_{nm} is a normalized array such that when the co-variance matrix is computed, it is the correlation matrix

$$C_{nn} = (\tfrac{1}{n}) D_{nm} D^{\dagger}_{mn} \qquad \text{Eqn 1}$$

with the correlation matrix, a routine eigenvector analysis is performed resulting in:

$$C_{nn} E_{nn} = E_{nn} V_{n} \qquad \text{Eqn 2}$$

where the E_{nn} are the orthogonal eigenvectors and the V_n are the eigenvalues.

The eigenvector set constitutes an orthogonal basis set of the correlation matrix.

To analyze the data, some workers have referred to these eigenvectors as the principal component eigenvectors as they are used to compute the principal components or amplitudes as some workers have called them.

The computation is:

$$X_{nm} = E_{nn}^{-1} D_{nm} \qquad \text{Eqn 3}$$

The meaning of this equation and subsequently the significance of the eigenvector can best be illustrated if the first element of the principal components is written in long form:

$$\begin{aligned} X_{11} &= \begin{bmatrix} e_{11} & e_{21} & e_{31} & e_{41} \cdots & e_{n1} \end{bmatrix} \begin{bmatrix} d_{11} \\ d_{21} \\ \vdots \\ d_{n1} \end{bmatrix} \\ X_{12} &= \begin{bmatrix} e_{12} & e_{22} & e_{32} & e_{42} \cdots & e_{n2} \end{bmatrix} \\ &\vdots \\ X_{n1} &= \begin{bmatrix} e_{1n} & e_{2n} & e_{3n} & e_{4n} \cdots & e_{nn} \end{bmatrix} \end{aligned} \qquad \text{Eqn 4}$$

Thus: $$X_{11} = (e_{11} d_{11} + e_{21} d_{21} + e_{31} d_{31} \cdots e_{n1} d_{n1}) \qquad \text{Eqn 5}$$

As one can see, the X_{nm} is a mapping of the normalized data onto the orthogonal coordinate system defined by the set E_{nn}. The set X_{nm} is a set of row vectors (X_m^n) of n new variables. Each row vector is orthogonal to every other row vector. Each row vector X_m^n in X_{nm} can be thought of as a vector of a state variable of a system defined by the original set of n variables in D_{nm} with each j^{th} element of X_m^n ; $j = 1,m$ denoting the variance with time of the n^{th} system state variable.

Considered in this way, the components of the eigenvectors are the coefficients of the n original variables constituting the j^{th} component of X_m^n. As with the correlation coefficient, the amount of variance in X_j^n produced by the i^{th} variable in the j^{th} column of

D_{nm} is denoted by the square of its appropriate eigenvector component. Thus:

$$\sum_{k=1}^{n} (e_{kn})^2 = 1.0 + \theta \qquad \text{Eqn 6}$$

where θ is the computational round off error.

To illustrate the orthogonality of the X_{nm}, note that:

$$X_{nm} X^{-1}_{mn} = E^{-1}_{nn} D_{nm} D^{-1}_{mn} E_{nn} \qquad \text{Eqn 7}$$

and

$$(\tfrac{1}{n}) E^{-1}_{nn} D_{nm} D^{-1}_{mn} E_{nn} = E^{-1}_{nn} C_{nn} E_{nn} = IV_n \qquad \text{Eqn 8}$$

Thus they are orthogonal. The singularity plane is defined by the eigenvectors whose eigenvalue is zero, within system noise and round-off error. These zero value eigenvectors reveal the linear dependence of the data. The eigenvalues not close to zero provide a measure of the deviation from that plane in orthogonal hyperspace of their associated eigenvectors. It can be shown (2) that the relative magnitude of an eigenvalue denotes the amount of variance in the original data set explained by the combinations of the original data variables dictated by the eigenvector associated with that eigenvalue.

Consequently, the amount of variance of the u^{th} variable in D_{nm} explained by the p^{th} eigenvector is denoted by:

$$\text{explained variance} = \sum_{i=1}^{n} \frac{V_p}{V_i} (e_{up})^2 \qquad \text{Eqn 9}$$

Combining equation 6 and 9, and summing over n, the amount of variance in D_{nm} accounted for by the p^{th} eigenvector is computed by:

Eqn 10

$$\text{explained variance} = \sum_{u=1}^{n} \sum_{i=1}^{n} \frac{V_p}{V_i} (e_{up})^2 = \sum_{i=1}^{n} \frac{V_p}{V_i}$$

Using this information, we can select from E_{nn} a set E_{np} which account for a desired amount of the variance. With E_{np} we can compute

Eqn 11

$$X_{pm} = E^{-1}_{pn} D_{nm}$$

where the set of row vectors X_{pm} are orthogonal, uncorrelated, are time varying and explained an amount of variance in D_{nm} given by:

Eqn 12

$$\text{explained variance} = \sum_{i=1}^{n} \sum_{q=a}^{p} \frac{V_q}{V_i}$$

Note that the $V_q ; q = a, p$ are not in general sequential as the first p of V_n. With these properties, the X_{pm} are ideally suited for regression analysis onto other measured quantities in the form

Eqn 13

$$Y_{lm} = A_{lp} X_{pm}$$

where A_{lp} is a row vector of regression coefficients. In this way, the data D_{nm} can be transformed to Y_{lm} by the transfer function $T_{ln} = A_{lp} E^{-1}_{pn}$. The precision of the estimate of Y_{lm} by equation 13 is given by:

Eqn 14

$$\text{explained variance} = \sum_{i=1}^{n} \sum_{q=a}^{p} s_e^2 \frac{V_q}{V_i}$$

where s_e^2 is the coefficient of determination of A_{lp} . Confidence limits $(1-\alpha)$ for each element of the transfer function T_{ln} can be obtained from the respective standard errors of A_{lp} . The computation is performed by (see Draper & Smith, #7, and Fritt et al #4).

$$S_{nn} = t\left\{m-p-1, 1-\tfrac{1}{2}\alpha\right\} \cdot \left(E_{np}\, U_{pp}\, U_{pp}\, E_{pn}^{-1} \right)^{\frac{1}{2}} \qquad \text{Eqn 15}$$

where U_{pp} is a diagonal matrix of standard errors of the elements of A_{lp} .

APPLICATION:

The application of these multivariate techniques to equipment modeling can proceed as follows. First, we must realize that the scope of analysis can be varied from a detailed analysis of an individual equipment piece functioning in a micro-climate to a general analysis of a group of identical equipment pieces functioning in a regional climate. The distinction will become clearer later.

For the micro-climate analysis, we start with a set of measured parameters such as frequencies of internal oscillators, various currents, pulse widths, accuracy of output signals compared to a standard or whatever else is considered to be important to the functioning of the equipment. These variables are incorporated into the normalized set D'_{nm}, l = location or equipment type for n variables and m measurements. Then follow the method outlined in the previous section so as to select the p most important eigenvectors E_{np} and compute the equipment principal components set X'_{pm} .

Now, perform a similar operation with the meteorological data F'_{rm} of r meteorological variables analogous to D'_{nm} and derive the climate principal components Y'_{qm} from the q most important eigenvectors comprising the set G_{rq}. With these two sets of principal components, calculate a matrix of row vectors of regression coefficients A'_{pq} such that

$$X'_{pm} = A'_{pq} Y'_{qm} \qquad \text{Eqn 16}$$

Additional information can be gotten from the analysis. For example, the correlation matrix C_{nn} contained all correlations and interlocking correlations of the parameters. However, since the original data in D_{nm} is not in general, orthogonal and uncorrelated, the inter-locking correlations can be deceptive. An analysis of the eigenvectors is more promising. Also, the following operations can be performed which amounts to a mapping of the Y_{qm} and X_{pm} back to the "real world" space of normalized data. It is:

$$\hat{D}_{nm} = E_{np} \; A'_{pq} \; G^{-1}_{qr} \; F'_{rm} \qquad \text{Eqn 17}$$

where $\hat{D}$ denotes a regression-based estimate of D. The transfer function

$$T'_{nr} = E_{np} \; A'_{pq} \; G^{-1}_{qr} \qquad \text{Eqn 18}$$

is identified which is a set of n row vectors

$$T'_{nr} = (t^n_r)' \qquad \text{Eqn 19}$$

so that

$$\hat{D}'_{nm} = T'_{nr} \; F'_{rm} \qquad \text{Eqn 20}$$

This function can then be used with validity for transforming additional, normalized data F'_{rs} to $\hat{D}'_{ns}$ constituting a prediction estimate, if careful attention is paid toward maintaining the original test conditions present in the derivation of T'_{nr} . The form then is

$$\hat{D}'_{ns} = T'_{nr} \; F'_{rs} \qquad \text{Eqn 21}$$

The amount of variance in the n^{th} equipment variable explained by all r climate variables is estimable by the method outlined by equation 14 and the confidence of the analysis given by equation 15.

Clearly, this technique of analysis could be applied to a general testing condition of a complex system just as easily as to a piece of equipment. The set X_{pm} could be regressed on both climate and on some other measurable "effectiveness" parameter. Thus giving a two-fold set of performance explanations.

On the regional scale, the application appears as follows: The mean-subtracted data array D^e_{nm} consists of: e : equipment parameter, n variable locations and m measurements. This then leads to E^e_{np} and X^e_{pm} in the same manner as before. Similarly the climate data is composed in the same way F^v_{nm} ; v: meteorological parameter, e.g., precipitation, temperature, or pressure, n locations of measurement which should be at or near the same location as the equipment and m measurements. The same analysis is followed as before except that in

this case D^{c}_{nm} and F^{v}_{nm} are of only one parameter and as such the correlation matrix of the micro-climate model may be replaced by the covariance matrix. This has the possible advantage in transforming back to the real world through T^{cv}_{nn} where

$$\hat{D}^{c}_{nm} = T^{cv}_{nn} \; F^{v}_{nm} \qquad \text{Eqn 22}$$

where

$$T^{cv}_{nn} = E^{c}_{np} \; A_{pq} \; G^{v}_{qn} \qquad \text{Eqn 23}$$

and $F^{v}_{nm} = (f_{nm} + \bar{f}_{n})^{v}$ and $D^{c}_{nm} = (d_{nm} + \bar{d}_{n})^{c}$ instead of being expressed in standard deviation units.

As before, an estimate, or prediction, of $\hat{D}^{c}_{ns}$ from F^{v}_{ns} by T^{cv}_{nn} can be made with the associated precision and confidence computed as before.

There is an advantage in using the regional analysis in that the climate construction is on a regional scale and as such is more easily related to current technology in meteorological science since the Y^{v}_{qm} varies only in time for the entire region and G^{v}_{nq} only in location for all time. Similarly, the X^{c}_{pm} vary only in time and the E^{c}_{np} vary only in location. Thus, plotting equation 9 for each u^{th} location can reveal some very interesting results concerning the sensitivity of locations from the vantage point of the equipment to regional climate conditions.

BIBLIOGRAPHY

1. Hotelling, H. - Analysis of a Complex of Statistical Variable into Principal Components; J. E. P., 1933.

2. Anderson, T. W. - An Introduction to Multivariate Statistical Analysis; John Wily, 1958.

3. Hoel, P. G. - Introduction to Mathematical Statistics; Wiley, 1947.

4. Fritts, H. C., et al - Multivariate Techniques for Specifying Tree Growth and Climate Relationships and for Reconstructing Anomalies in Paleoclimate, Jnl. App. Met., Vol. 10, No. 5., Oct 1971.

5. Sellers, W. D. - Climatology, of Monthly Precipitation Patterns in Western United States, 1931-1966, Mon. Wea. Rev, Vol. 96, No. 9, Sep 68.

6. Box, G. E. P., et al - Some Problems Associated with The Analysis of Multiresponse Data, Technometrics, Vol 15, No. 1, Feb 73.

7. Draper, N. R. and Smith, H. - Applied Regression Analysis, Wiley, New York, 1966.

ONE SHOT SENSITIVITY TEST FOR EXTREME PERCENTAGE POINTS

Seymour K. Einbinder
Concepts and Effectiveness Division
Ammunition Development and Engineering Directorate
Picatinny Arsenal
Dover, New Jersey

ABSTRACT. Sensitivity testing deals with continuous variables which cannot be measured directly. Several well known statistical techniques are available for the treatment of experimental responses which are quantal, i.e., all-or nothing. In this paper, a sequential sensitivity test strategy and estimation methodology is proposed. This procedure appears to be more efficient than the methods that are in common use for sensitivity testing to determine extreme percentage points of a response function. No loss of efficiency or accuracy results in the estimation of central percentage points such as $L_{.50}$ or $LD_{.50}$. The method is robust to many forms of the underlying response distribution, doesn't require a possibly limiting assumption of normality and stimulus step size, and is insensitive to the choice of initial stimulus level. The test procedure is discussed in three parts: (1) the test strategy, (2) the response model and (3) test data analysis and estimation.

1. INTRODUCTION. The basic ingredients of a sensitivity test are: a stimulus, a test specimen and a quantal response. Associated with each object is a critical stimulus or strength such that if the stimulus exceeds the strength, the object responds, and vice versa. The distribution of strengths in a population of objects is called the response distribution or the response function. Some typical responses may be failure, survival, functions properly, explodes, etc. The response must be chosen such that an increase in the stress level results in an increase in the response function or the probability of the occurrence of the response.

The basic problem in sensitivity testing is to estimate the response function either completely or locally over some region. Also required are estimates of percentage points of the response distribution and probabilities of response at specified levels of stress.

Some simple, basic rules for the design and analysis of a sensitivity experiment underlie the test procedure that was developed. In the first place extrapolation of results beyond the region of the test data should be minimized. Tests should be conducted as close to the region of interest as possible. The use of continuously variable stress levels applied sequentially is preferable to discrete ones. Finally, the response model should be robust to the form of the usually unknown underlying response distribution and be a good approximation in the local region of primary interest.

2. REVIEW OF SEQUENTIAL SENSITIVITY TESTS. A brief review of some well known sequential sensitivity test methods will be given next to facilitate the presentation of the new test procedure.

2.1 Up and Down Test. Probably the most well known sequential sensitivity test procedure is the Dixon and Mood Up and Down Test (1) which is illustrated in Figure 1. One item is tested at a time starting at the best initial estimate of the 50% response point. The test level is moved up one step after each negative response and down after a positive one. The step size is fixed and must be determined in advance of the test. The recommended step size is equal to the standard deviation of the response distribution. This method of testing tends to concentrate the observations near the mean. As a result, the procedure is quite good in estimating the mean or 50% point of a symmetric distribution but does not do too well with extreme percentage points. The method of analysis assumes a normal distribution for the response curve.

2.2 Langlie One Shot Test. Langlie (2) developed a sequential test strategy that overcame certain difficulties with the Up and Down test. This strategy makes use of continuously variable stress levels and is insensitive to the starting level and the apriori choice of a step size. The analysis is based upon a normal response distribution and has been shown to be more efficient than the Up and Down method in estimating the mean or 50% point as well as the standard deviation.

The Langlie Method is best illustrated by means of an example from his report (See Figure 2). The objective was to determine the high temperature performance of thermal batteries. The procedure consists of first choosing upper and lower limits such that all items function satifactorily at the lower limit and all fail at the upper limit. Failure is defined as a

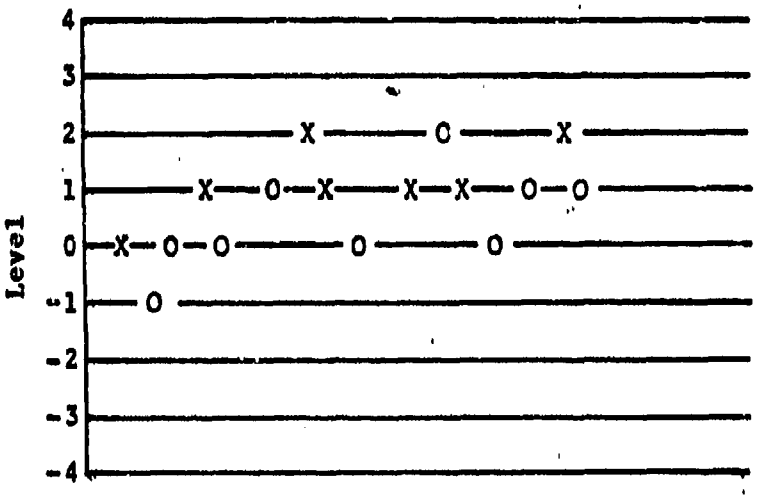
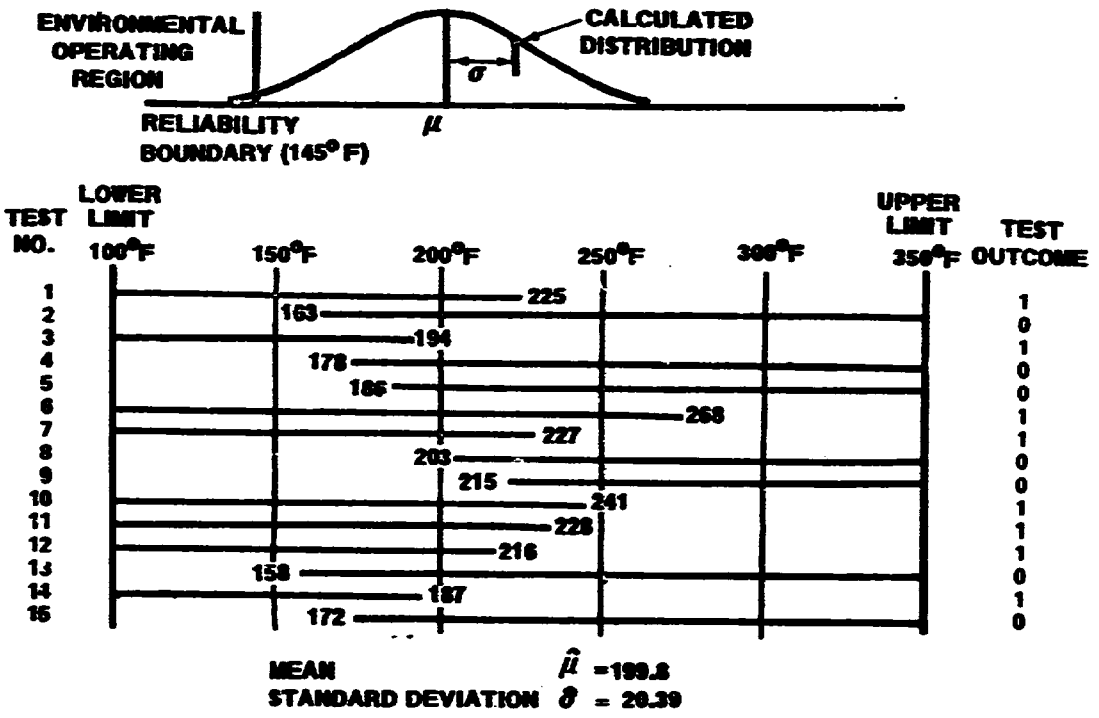


FIGURE 1. A TYPICAL UP AND DOWN EXPERIMENT

FIGURE 2. **SAMPLE "ONE-SHOT" TEST**



TEST-TO-FAILURE OF THERMAL BATTERIES IN TEMPERATURE

positive response and is denoted by 1 and successful battery performance is denoted by 0. The first stress level is selected at the midpoint of the interval. Since a failure resulted, the second trial is at the average of the first stress level and the lower limit. The general rule for the next stress level is to average the last stress level with an earlier stress level such that the number of 1's equals the number of 0's in going from one to the other. If no previous level exists that satisfies that condition then the last stress is averaged with either the upper or lower limit depending upon whether an increase or decrease in stress is required.

The accuracy and efficiency of the Langlie strategy are not sensitive to starting level, and no apriori knowledge of the standard deviation is required. Analysis of data is by the method of maximum likelihood and is based upon the assumption of a normal distribution of strengths.

2.3 Sensitivity Testing for Extreme Percentage Points. The test methods described thus far tend to concentrate the observations near the central part of the response distribution and are efficient for estimating its properties there. Often, however, the experimenter is interested in the nature of the extreme parts of the distribution. This requires appropriate test data from the region of interest. Wetherill (3) published the results of an investigation of sequential test methods for the estimation of general percentage points of a quantal response function. He found that most of the available procedures such as the Up and Down and the Robbins - Monro were not suitable for estimation of extreme percentage points. A rule for transforming the response in an Up and Down test was proposed by Wetherill which tends to concentrate the observations in the tail areas of the response curve. The question of stopping rules was also addressed by Wetherill. For more detail refer to Reference 3.

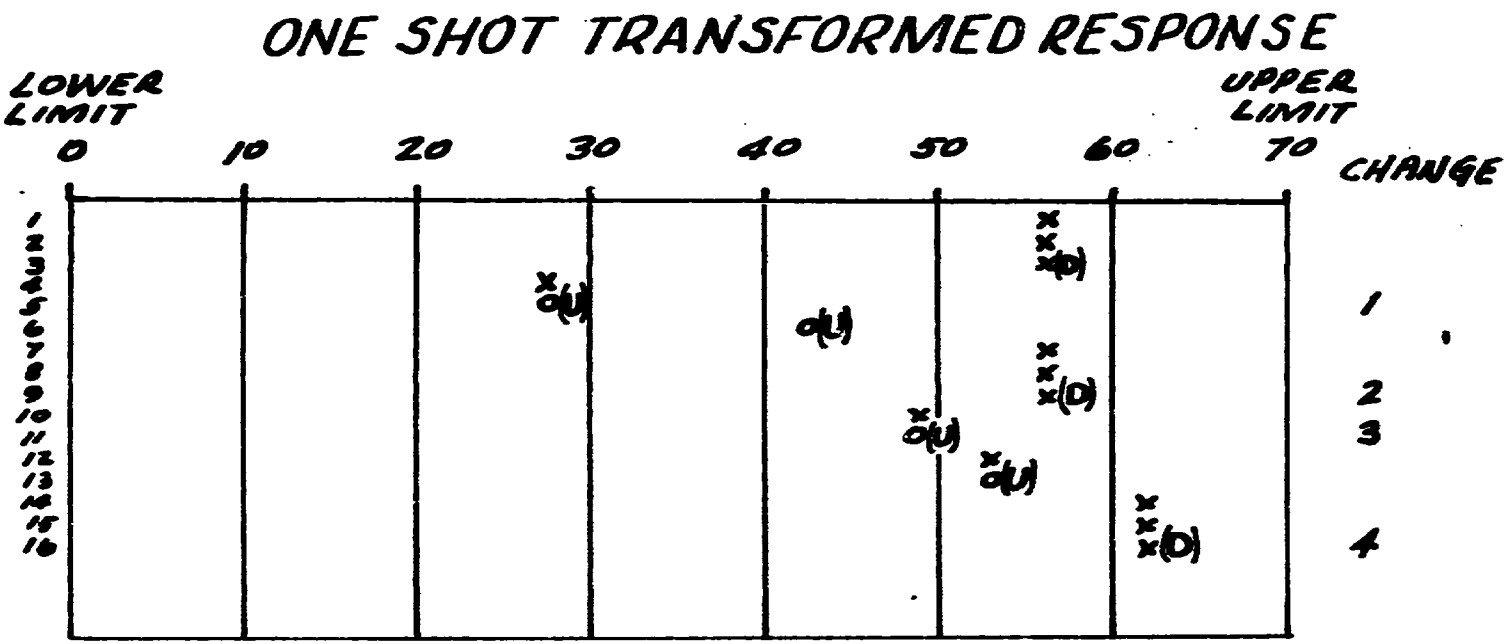
3. ONE SHOT TRANSFORMED RESPONSE TEST. In this test procedure, the Wetherill response transformation is applied to the Langlie One-Shot test algorithm. In this way, advantage is taken of the best features of both techniques which results in an efficient and effective new method of sensitivity testing. This test procedure is called the One-Shot Transformed Response (OSTR) Rule.

3.1 One-Shot Transformed Response Test Strategy. The OSTR strategy is best illustrated by an example which is protrayed graphically in Figure 3. The objective in the example is to estimate the upper tail performance of the response distribution. An OSTR test strategy with $N_o=3$ was used. The transformation is defined by the value of N_o which determines the response quantile around which the test levels tend to concentrate. This quantile is called the transformed median percentage (TMP) point for reasons which will be apparent later. For $N_o=3$, the TMP = 79%. The response transformation is designed to make an increase in stress easier than a decrease. The greater the difficulty in decreasing the stress level, the greater will be the transformed median percentage. A positive response is denoted by an X or 1 and a negative response by 0. A type D response, which requires a reduction in stress level, is allowed to occur after N_o confirmations of a positive response. In this case a type D response consists of a (111) outcome and a type U response of the set of outcomes (0), (10) or (110). After choosing lower and upper stress limits, as in the Langlie procedure, the first test level is chosen at the percentage point of the test interval corresponding to the approximate TMP of the transformation. The sequence of tests at the first stress level resulted in three positive responses or a type D. The process proceeds by applying the Langlie strategy to the U's and D's. Thus, the second stress level is at the average of the lower limit and the first stress.

3.2 Stopping Rules and the $\bar{W}$ Estimator. The same test sequence and results are shown in tabular form in Table 1. A change of response type is said to occur when an alternation of response is obtained. In this experiment the fourth change of response occurred at the 16th trial. Wetherill proposed a stopping rule based upon a specified number of changes of response type rather than a fixed number of trials. Occasionally, peculiar sequences of outcomes occur which provide little or no information about the response distribution. This condition is minimized by using the change of response stopping rule rather than a fixed sample size.

The number of observations required in an experiment is a random variable with this stopping rule. The expected sample size with a particular number of changes or response increases with N_o or the farther out in the tails of the response curve in which the testing takes place. For each sequence of trials on the transformed scale that represents a change of response, a reasonable estimate of the 50 percentile is the midpoint of

FIGURE 3.



$n_0 = 3$ TMP = .7937

x(1) = .8(70) = 56

RESPONSE TYPE: D = XXX
U = O, XO, XXO

x ↔ 1

TABLE 1

ONE SHOT TR TEST FOR n_o = 3, TMP = .7937

A = 0, B = 70, x(1) = .8(70) = 56

Trial I	Stress X(I)	Response Y(I)	Response Type*	Change Number	w
1	56.	1			
2	56.	1			
3	56.	1	D		
4	28.	1			
5	28.	0	U	1	42.
6	42.	0	U		
7	56.	1			
8	56.	1			
9	56.	1	D	2	49.
10	49.	1			
11	49.	0	U	3	52.5
12	52.5	1			
13	52.5	0	U		
14	61.25	1			
15	61.25	1			
16	61.25	1	D	4	56.875
					200.375

$$\bar{w} = \frac{1}{4}(200.375) = 50.09$$

* D: 111
U: 0, 10, 110

the stress interval where the change took place. These estimates are denoted by w after Wetherill who proposed this estimator for the Up and Down routine which uses equally spaced stress levels. The Wetherill approach was also found to apply to the Langlie strategy and appears to be efficient and accurate. Each change of response results in a separate estimate of the transformed 50% point and the overall point estimate $\bar{w}$ is the average of all of the individual estimates.

In this example, $\bar{w}$ is an estimate of $L_{.79}$. An iterative method of calculating maximum likelihood estimates of the response curve parameters is used which requires initial estimates of the parameters. These are obtained by a method of matching percentage points. The $\bar{w}$ estimator is useful for estimating percentage points to be used for maximum likelihood estimation.

3.3 Transformed Response Strategies. Transformed response strategies for several values of N_o are illustrated in Table 2. Also shown are the upper and lower tail percentage points that are estimated and around which the one shot test levels tend to concentrate. This table can be easily extended for additional values of N_o. For P>.5, X denotes a positive response, and 0 designates a negative one. For lower tail strategies, the responses are redefined. Zero represents a positive response, and the U and D designations are interchanged. The TMP for a given N_o is found in the following manner. Call P(x) the probability of a positive response at a test level x. Then the probability that the outcome at this level will result in a downward change of level is $[P(x)]^{N_o}$. The transformed response curve, in which the responses are classified as U or D, is then

$$F(x) = P[(x)]^{N_o}$$

The OSTR strategy is actually the Langlie routine applied to a transformed response curve. The standard Langlie procedure, which is obtained when $N_o=1$, may be used to estimate the 50% point of the transformed response curve F(x). By solving for the value of P(x) corresponding to F(x) = .5, we obtain the probability value of the original response function corresponding to the 50% point of the transformed response. This is the transformed median percentage. As noted before, for $N_o=3$ this value is .7937 in the upper tail and .2063 in the lower tail.

TABLE 2

CHARACTERISTICS OF SOME TRANSFORMED RESPONSE STRATEGIES

	Response Type*		Trans formation,	Percentage Point Estimated	
	D if p>.5,	U if p>.5,			
n_o	U if p<.5	D if p<.5	F =	p<.5	p>.5
2	XX	X0, 0	p^2	.2929	.7071
3	XXX	XX0, X0, X	p^3	.2063	.7937
3	XXX, XX0X	XX00, X0, 0	$p^3(2-p)$	.2664	.7336
4	XXXX	XXX0, XX0, X0, 0	p^4	.1591	.8409
4	XXXX, XXX0X	XXX00, XX0, X0, 0	$p^4(2-p)$	.1959	.8041
5	XXXXX	XXXX0, XXX0, XX0 X0, 0	p^5	.12945	.87055
5	XXXXX, XXXX0X	XXXX00, XXXX0 XX0, X0, 0	$p^5(2-p)$	.1540	.8460
6	XXXXXX	XXXXX0, etc.	p^6	.1092	.8908
7	XXXXXXX	XXXXXX0, etc.	p^7	.0944	.9056
8	XXXXXXXX	XXXXXXX0, etc.	p^8	.0829	.9171
9	XXXXXXXXX	XXXXXXXX0, etc.	p^9	.0740	.9260
10	XXXXXXXXXX	XXXXXXXXX0, etc.	p^{10}	.0670	.9330
14	XXXXXXXXXXXXXX	XXXXXXXXXXXXX0, etc.	p^{14}	.0484	.9516

*For p>.5, X→response and 0→ non-response.
For p<.5, X→non-response and 0→ response.

4. THE RESPONSE FUNCTION. Next we consider the sensitivity model or the form of the response function. The well known Weibull distribution was selected as a general response model for several reasons.

The Weibull distribution is general and very useful as a response function because it exhibits many different shapes. Consequently, it is robust to the form of the true response distribution. Since skew shapes may also be approximated, the need for normalizing transformations that are required by most conventional methods is minimized or eliminated.

The main difficulty with the Weibull is the fact that it is a 3 parameter distribution and the location parameter is difficult to estimate. Fortunately, however, the estimation of responses and percentage points in the stress regions of interest appear to be insensitive to variations in the location parameter.

Figure 4 illustrates the variety of shapes of the Weibull density for a number of values of the shape parameter γ. An exponential distribution results for $\gamma=1$. The distributions are skewed to the right for $\gamma<3.6$ and to the left for $\gamma>3.6$. Normal densities are approximated by $\gamma = 3.6$.

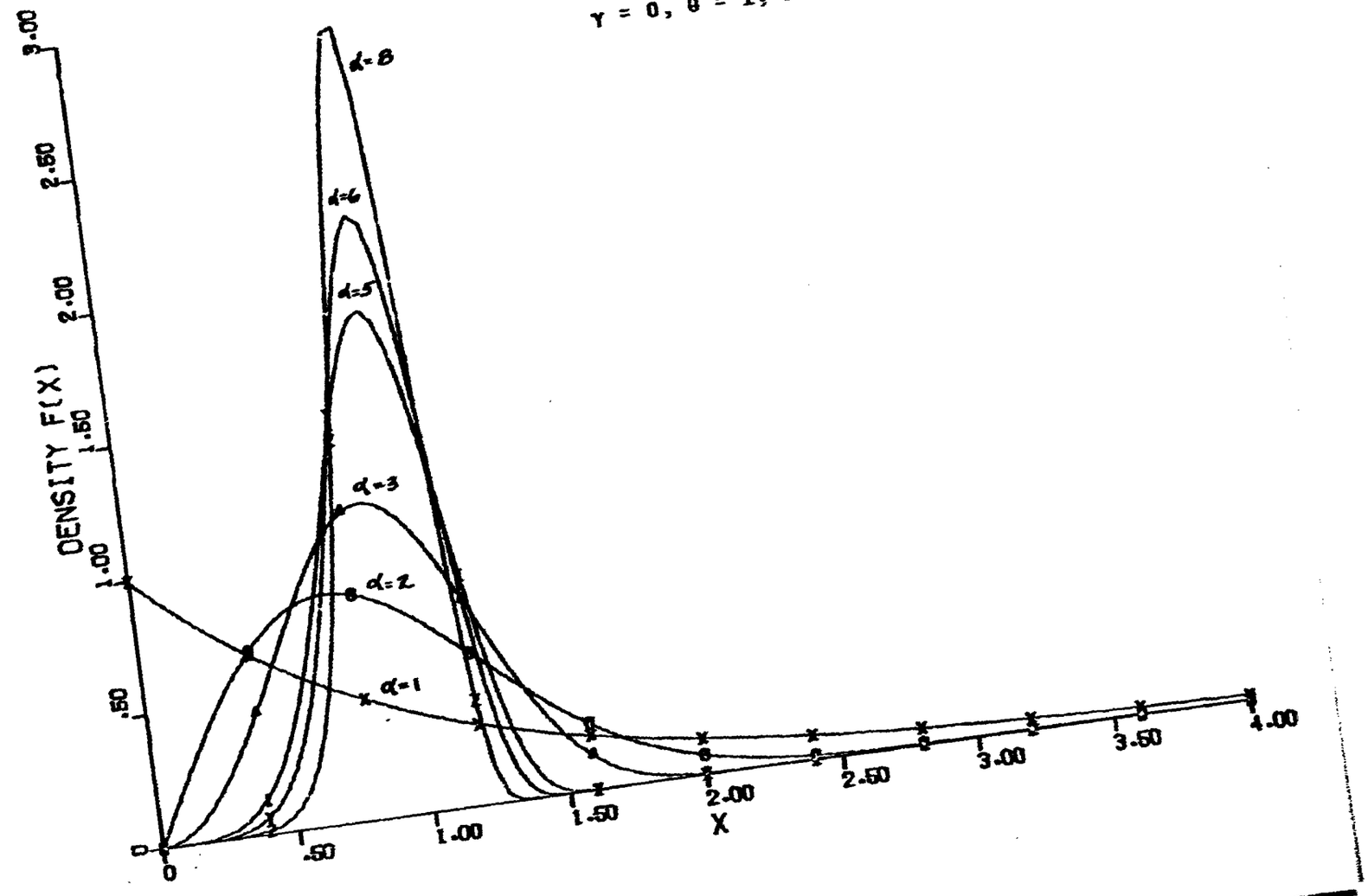
To provide additional flexibility and capability in fitting the best response functions to experimental outcomes in some local stress region, a reflected version of the Weibull distribution was also introduced. This allows the tail of the Weibull that best fits the data to be used for estimation purposes and minimizes the dependence of the estimates on the location parameter γ.

5. ESTIMATION. Given the results of a sensitivity experiment, maximum likelihood theory was applied to estimate the best Weibull response distribution. Let the stresses be $x_1,\ldots,x_n$ and the outcomes $y(x_1),\ldots,y(x_n)$ for a series of n trials. An outcome $y(x)=1$ denotes a positive response, and $y(x)=0$ indicates no response. To find the maximum likelihood estimate of $\underline{\theta} = (\theta_1,\ldots,\theta_p)'$, for the response distribution function $F(x;\underline{\theta})$, we require the likelihood equation:

$$L(\underline{y};\underline{\theta}) = \prod_{i=1}^{n} P_i^{y_i} Q_i^{1-y_i}$$

FIGURE 4. WEIBULL DISTRIBUTION CURVES

$\gamma = 0, \theta = 1; \alpha = 1, 2, 3, 5, 6, 8.$



where $P_i = P[x_i] = P[Y(x_i)=1] = \int_{-\infty}^{x_i} dF(x;\underline{\theta})$.

Differentiating log $L(y;\underline{\theta})$ with respect to each parameter and equating each derivative to zero results in a set of p equations whose solution set gives the maximum likelihood estimate for $\underline{\theta}$.

$$g_j(\underline{\theta}) \triangleq \frac{\partial \log L}{\partial \theta_j} = \sum_i \frac{y_i - P_i}{P_i Q_i} \frac{\partial P_i}{\partial \theta_j} = 0 \quad (j=1,\ldots,p)$$

Solution of these equations is accomplished by linearizing them by means of a Taylor series expension about $\underline{\theta}_o$. Thus,

$$g_j(\underline{\theta}_o+\Delta\underline{\theta}) = g_j(\underline{\theta}_o) + \sum_{k=1}^{p} \Delta\theta_k \left.\frac{\partial g_j(\underline{\theta})}{\partial \theta_k}\right|_{\underline{\theta}=\underline{\theta}_o} = 0$$

The linearized equations are then solved using a Newton-Raphson iterative procedure. The problem of an initial estimate $\underline{\theta}_o$ for the first iteration of the solution was solved by matching percentage points of the response curve with particular stimulus levels.

The covariance matrix for the estimators is the inverse of the Fisher information matrix B which is defined as

$$B(jk) = \left[-E\left(\frac{\partial^2 \log L}{\partial \theta_j \partial \theta_k}\right)\right]^{p\times p}$$

It then follows that the covariance matrix for $\hat{\underline{\theta}}$ is

$$\Sigma_{\underline{\theta}} = B^{-1}$$

5.1 M.L. Estimates of Weibull Parameters. The form of the Weibull distribution function found to be most practical and useful for fitting to sensitivity data is

$$F(x) = 1-\exp-[(x-\gamma)/\theta]^{\alpha} \qquad x \geq \gamma$$

$$= 0 \qquad x < \gamma$$

Maximum likelihood estimates of the Weibull parameters $\underline{\theta} = (\theta,\alpha,\gamma)'$ are obtained as described in the previous section. Starting values for the iterative solution are found by matching two percentage points for a fixed value of the location parameter γ. The $\bar{w}$ estimator is useful for finding these percentage points. The M.L.E. of γ is found by searching over the domain of γ. Convergence problems were encountered in solving the non-linear equations. A transformation of the data into an exponential form based upon the critical estimates of the Weibull parameters was found to stabilize and speed convergence to a solution.

5.2 Confidence Region for Weibull Parameters. A joint confidence region for the Weibull parameter vector $\underline{\theta}$ is obtained by making use of the asymptotic normality of the maximum likelihood estimator $\hat{\underline{\theta}}$ for $\underline{\theta}$.

$$\hat{\underline{\theta}} \sim N[\underline{\theta}, \Sigma_{\hat{\underline{\theta}}}]$$

The quadratic form in the parameters, $\nu = (\hat{\underline{\theta}}-\underline{\theta})' \Sigma_{\hat{\underline{\theta}}}^{-1} (\hat{\underline{\theta}}-\underline{\theta})$, is approximately chi square with 3 degrees of freedom. Thus, a $(1-\alpha)$ confidence region for $\underline{\theta}$ is given by

$$(\underline{\theta}-\hat{\underline{\theta}})' \Sigma_{\hat{\underline{\theta}}}^{-1} (\underline{\theta}-\hat{\underline{\theta}}) \leq \chi^2_{1-\alpha;q}.$$

$\chi^2_{1-\alpha;q}$ is the value of χ^2 for q degrees of freedom that satisfies

$$P[\nu \leq \chi^2_{1-\alpha;q}] = 1-\alpha$$

where q is the number of parameters estimated.

5.3 Estimation of Reliability and Percentage Points. The reliability at a given stress level x_c is defined as the probability that the strength exceeds this stress level. For the Weibull response function or strength distribution the reliability is

$$R(x_c;\underline{\theta}) = 1-\exp-[(x-\gamma)/\theta]^{\alpha},\ x\geq\gamma$$

An M.L. estimate of R is obtained by substituting $\hat{\underline{\theta}}$ for $\underline{\theta} = (\theta,\alpha,\gamma)'$.

$$\hat{R}(x_c) = R(x_c;\hat{\underline{\theta}}).$$

By the theorem on the asymptotic distribution of functions of consistent statistics, $\hat{R}(x_c)$ is asymptotically $N[R(x_c;\underline{\theta}), V_{\infty}(\hat{R})]$

where

$$V(R) = \sum_i \sum_j \frac{\partial R}{\partial \hat{\theta}_i} \frac{\partial R}{\partial \hat{\theta}_j} \operatorname{cov}(\hat{\theta}_i\hat{\theta}_j)$$

Thus, a lower (1-α) confidence limit for $R(x_c;\underline{\theta})$ is

$$\hat{R}(x_c) = z_{\alpha}\sqrt{V_{\infty}(\hat{R})}$$

where z_{α} is the 100α% point of the N(0,1) distribution such that

$$P[Z > z_{\alpha}] = \alpha$$

Percentage points L_p are estimated by solving the Weibull distribution function for L_p. Thus, we have

$$P = 1-\exp-[(L_p-\gamma)/\theta]^{\alpha}$$

which gives us

$$L_p = \theta(-\log Q)^{1/\alpha} + \gamma$$

where $Q = 1-P$.

Using the asymptotic distribution theorem as before, we get

$$\hat{L}_p = L_p(\underline{\hat{\theta}})$$

$$V_\infty (\hat{L}_p) = \sum_i \sum_j \frac{\partial L_p}{\partial \theta_i} \frac{\partial L_p}{\partial \theta_j} \operatorname{cov}(\hat{\theta}_i \hat{\theta}_j)$$

where

$$\frac{\partial L_p}{\partial \theta} = (-\log Q)^{1/\alpha}$$

$$\frac{\partial L_p}{\partial \alpha} = -\frac{\theta}{\alpha^2}(-\log Q)^{1/\alpha}\log(-\log Q)$$

$$\frac{\partial L_p}{\partial \gamma} = 1$$

Thus, $\hat{L}_p \sim N(L_p, V_\infty(\hat{L}_p))$. This permits the simple computation of asymptotic confidence intervals for percentage points of the response function.

5.4 Reflected Weibull Estimation. The reflected Weibull density and distribution functions are

$$f(x;\theta,\alpha,\gamma_R) = \alpha[(\gamma_R-x)/\theta]^{\alpha-1}\exp-[(\gamma_R-x)/\theta]^{\alpha} \quad x\leq\gamma_R$$
$$= 0 \quad x>\gamma_R$$

$$F(x;\theta,\alpha,\gamma_R) = \exp-[(\gamma_R-x)/\theta]^{\alpha} \quad x\leq\gamma_R$$
$$= 1 \quad x>\gamma_R$$

Fitting of a reflected Weibull distribution to a set of experimental data $(x_i, y(x_i))$, $i=1,\ldots,n$ was accomplished by reflecting the stress levels and outcomes about an arbitrary point A. The data transformed to the standard Weibull form is

$$x_s = 2A - x_i$$

$$y_s = 1 - y(x_i)$$

where x_s, y_s are the transformed stress and response which are used to obtain M.L. estimates for the parameters of a standard Weibull distribution, $\underline{\theta}_s = (\theta_s,\alpha_s,\gamma_s)'$. Since the scale and shape parameters (θ,α) are invariant under this transformation, the M.L. estimates of the reflected Weibull parameters are

$$\hat{\theta} = \hat{\theta}_s$$

$$\hat{\alpha} = \hat{\alpha}_s$$

$$\hat{\gamma}_R = 2A - \hat{\gamma}_s$$

where $\hat{\theta}_s$, $\hat{\alpha}_s$, $\hat{\gamma}_s$ are the M.L. estimates for the standard Weibull distribution based upon reflection of the original data about A.

The reflection transformation allows all of the statistics and estimators developed for the standard Weibull to be applied to the reflected Weibull.

6. SUMMARY AND CONCLUSIONS. In summary a new OSTR strategy has been developed that is relatively easy to apply. It has been found to be more efficient than existing methods for estimating extreme percentage points on a quantal response curve.

A general response model has been incorporated based upon the Weibull distribution which is robust and minimizes the need for normalizing transformations. In addition, the ability to fit response curves has been expanded by reflecting the Weibull distribution.

Maximum likelihood theory has been applied to derive, for both the standard and reflected forms of the Weibull distribution, point and confidence estimates of the parameters, the Fisher information matrix, the asymptotic covariances matrix and point and interval estimates of the reliability and percentage points.

The Wetherill $\bar{w}$ estimator was successfully applied to the Langlie routine and provides a simple and efficient estimate of certain percentage points.

Finally, a Fortran IV computer program was written that analyzes quantal data and calculates all of these statistics and estimates.

REFERENCES

1. Dixon, W.J. and Mood, A.M., "A Method for Obtaining and Analyzing Sensitivity Data", J. Amer. Statist. Assoc., Vol. 43, pp. 109-126, 1948.

2. Langlie, H.J., "Reliability Test Method for 'One-Shot' Items", Proceedings of the Eighth Conference on the Design of Experiments in Army Research Development and Testing, ARO-D Report 63-2, pp. 145-165, 1963.

3. Wetherill, G.B., "Sequential Estimation of Quantal Response Curves", J. Royal Statist. Soc., B, Vol. 25, pp. 1-48, 1963.

ANALYSIS OF TECHNIQUES FOR FINDING A FORWARD OBSERVER

Charles McElwee
Frankford Arsenal, Philadelphia, Pennsylvania

ABSTRACT

This paper is the result of work done in conjunction with projects designed to produce a point target first round hit capability. The initial portion of this task was to consider a new technique of forward observer location possible only since the development of precise Laser Rangefinders and optimize it. In the process of optimization and statistical comparison with existing techniques, the feasibility of a still newer idea came to light.

DEVELOPMENT OF TWO-HIT-RANGE-ONLY METHOD OF LOCATING A FORWARD OBSERVER

PROBLEM: Locate a Forward Observer by Firing Two Rounds from a Mortar and Having the FO Supply Range Information to the Hit Points

By using vector equations a computer sub-routine which is easily utilized was developed.

Input data required to use the sub-routine is range and directional information between the firing point and each of the hit points. The solution set contains two sets of range and direction calculations. The correct set may be ascertained by having the FO note whether the angle (as he observed) subtended by the first and second hits was clockwise or counterclockwise. The output ranges from the sub-routine will show one positive and one negative. The sign establishes the clockwise-counterclockwise sensing and should be ignored when plotting the FO's position once the proper solution is determined.

To test reliability of the technique, fifty-five (55) selected cases were tested and four (4) probability outlines including 100 percent were plotted accounting for the following error sources:

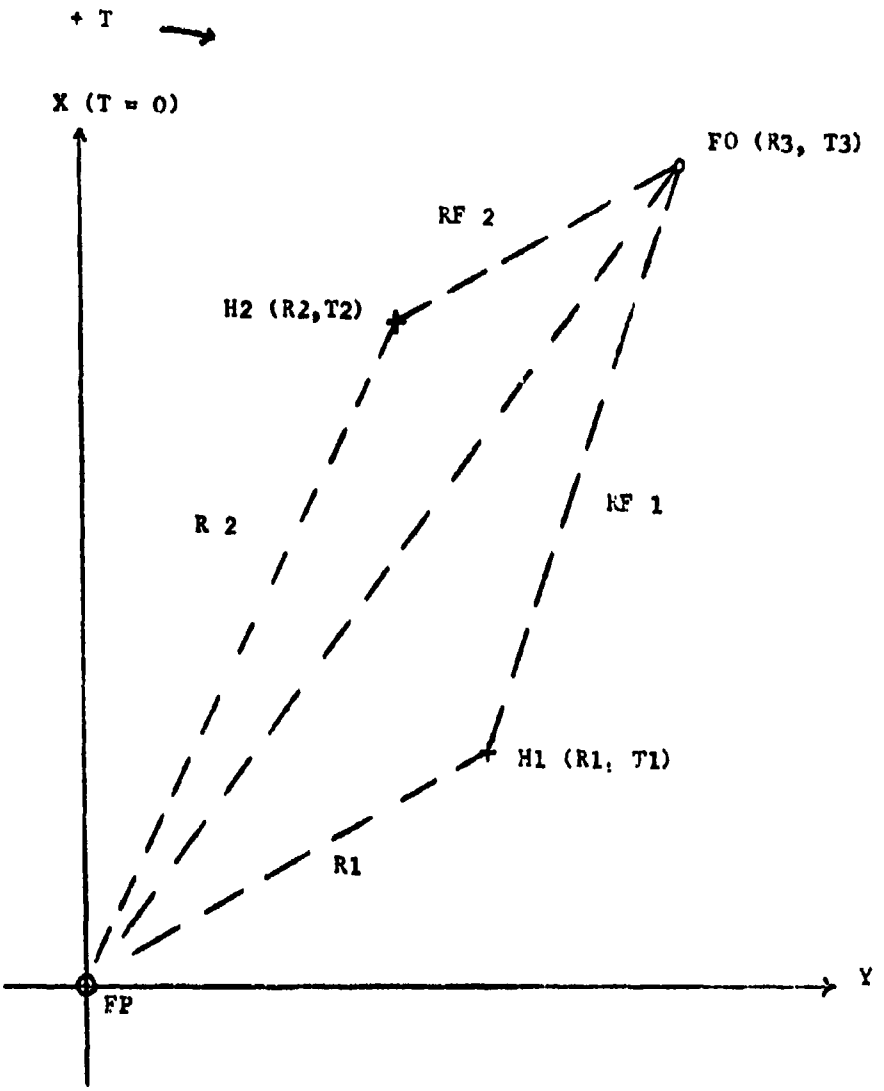
1. Elevation error in mortar (firing table)
2. Deflection error in mortar (firing table)
3. Rangefinder error.

Depending on the choice of data, the 100% error radius ranged from 65 to over 900 meters at extreme mortar ranges - reasonable through unreasonable. Investigation revealed that the error could be minimized by proper selection of the hit points with respect to the FO. There is, however, the requirement of some prior knowledge about the FO's position.

ASSUMPTIONS FOR INVESTIGATION

1. Round dispersion (Gaussian, according to firing tables)
2. Flat, level ground
3. ± 10 meter Rangefinder (accuracy plus precision, uniform)
4. ± 8 degrees compass error (gaussian, due to local magnetic disturbances).

The Mortar is used throughout this discussion but the results should in no way be construed as useful only in Mortar Systems. All aimed hit locations, FO actual locations, and "errors" incorporated in calculation were generated randomly internal to the program. Sample sizes in each sample set were also generated randomly, restricted to the region 30 to 170. Uniform or Gaussian random numbers were used in the program as appropriate.



FP - Firing Point

H1 - First Hit

H2 - Second Hit

FO - Forward Observer

R - Range

T - Angle

Problem Equations

$$RF2^2 = (F - G\,Y31 - X1)^2 + (Y31 - Y1)^2 = (F - G\,Y32 - X1)^2 + (Y32 - Y1)$$

These Are Second Degree Equations In Y31, Y32
Both Equations Are Identical With Y31, Y32 As Roots

$$(G^2+1)\,Y3^2 + (2X2G - 2H\,DY/DX^2 - 2Y2)\,Y3 + [(X/DX)^2 - 2X2/DX + R2^2 - RF2^2$$

A B C

$$D = -B/2A$$

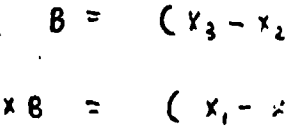
$$E = C/A$$

$$SR = \sqrt{\left(\frac{B}{2A}\right)^2 - \frac{C}{A}} = \sqrt{D^2 - E}$$

$$Y31 = D + SR$$

$$Y32 = D - SR$$

SENSING



$A \times B = |A||B| \sin\theta$

$\therefore A \times B$ Has Same Sign As θ

$$A = (X_3 - X_1)\,\hat{i} + (Y_3 - Y_1)\,\hat{j}$$

$$B = (X_3 - X_2)\,\hat{i} + (Y_3 - Y_2)\,\hat{j}$$

$$A \times B = (X_1 - X_3)(Y_2 - Y_3) - (Y_1 - Y_3)(X_2 - X_3)$$

$$= Y_3 Y_2 + Y_1 Y_3 - X_2 Y_1 - X_3 Y_2 - X_1 Y_3 + X_1 Y_2$$

$S = A \times B$ S1 (X31, Y31) S2 (X32, Y32)

The Sign of S Determines Observed Rotation From F0

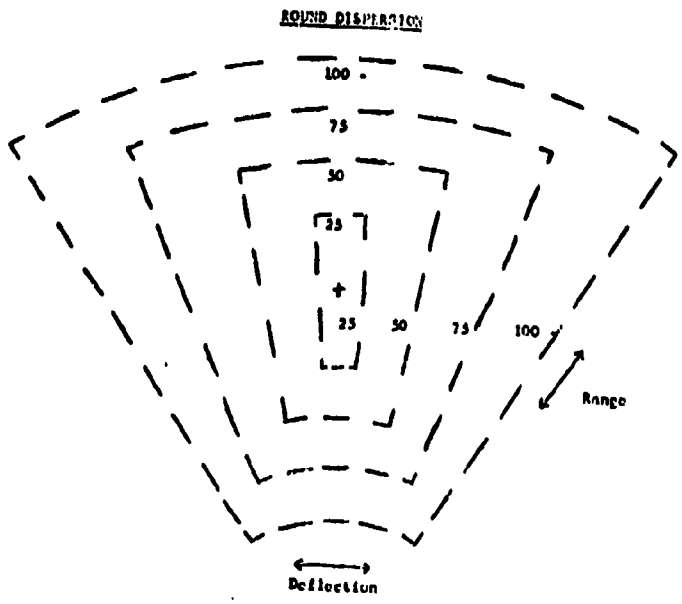
COMPUTER PLOTS

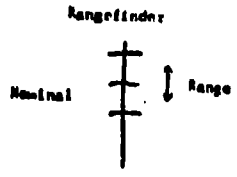
The following computer plots represent outlines resulting from the use of 25, 50, 75 and 100 percent boundaries of each error source. The asterisk near the center of the outline is the location of the observer if no errors were encountered. The scale is in meters/inch, with the maximum error appearing 4.166 inches from the center. The scale was established differently on each plot to obtain maximum resolution. Other information on the plot was used for researching purposes only. The non-continuous nature of the outlines is due to discrete point selection on the error outlines ("walking the boundary").

The plots for the FO- are in the minimum error configuration which has been found to exist. This configuration is discussed in the conclusion for this part.

ERROR OUTLINES

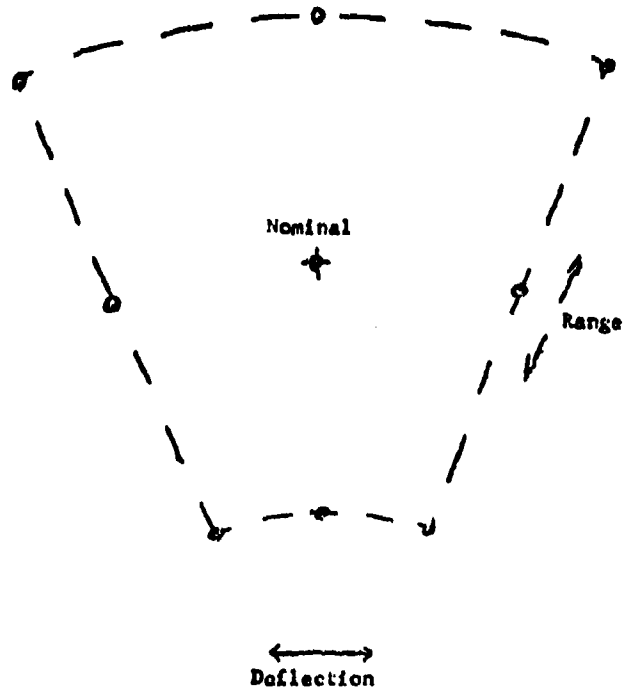
The error outlines below are indicative of those used for this discussion. Note that two hits and two ranges were used for each calculation.

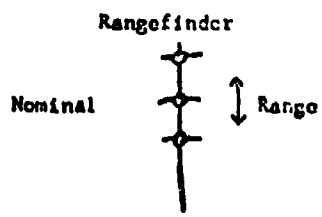




A typical "walkaround" is depicted below. There are nine (9) points per reference round and three (3) points per ranging for a total of 729 points.

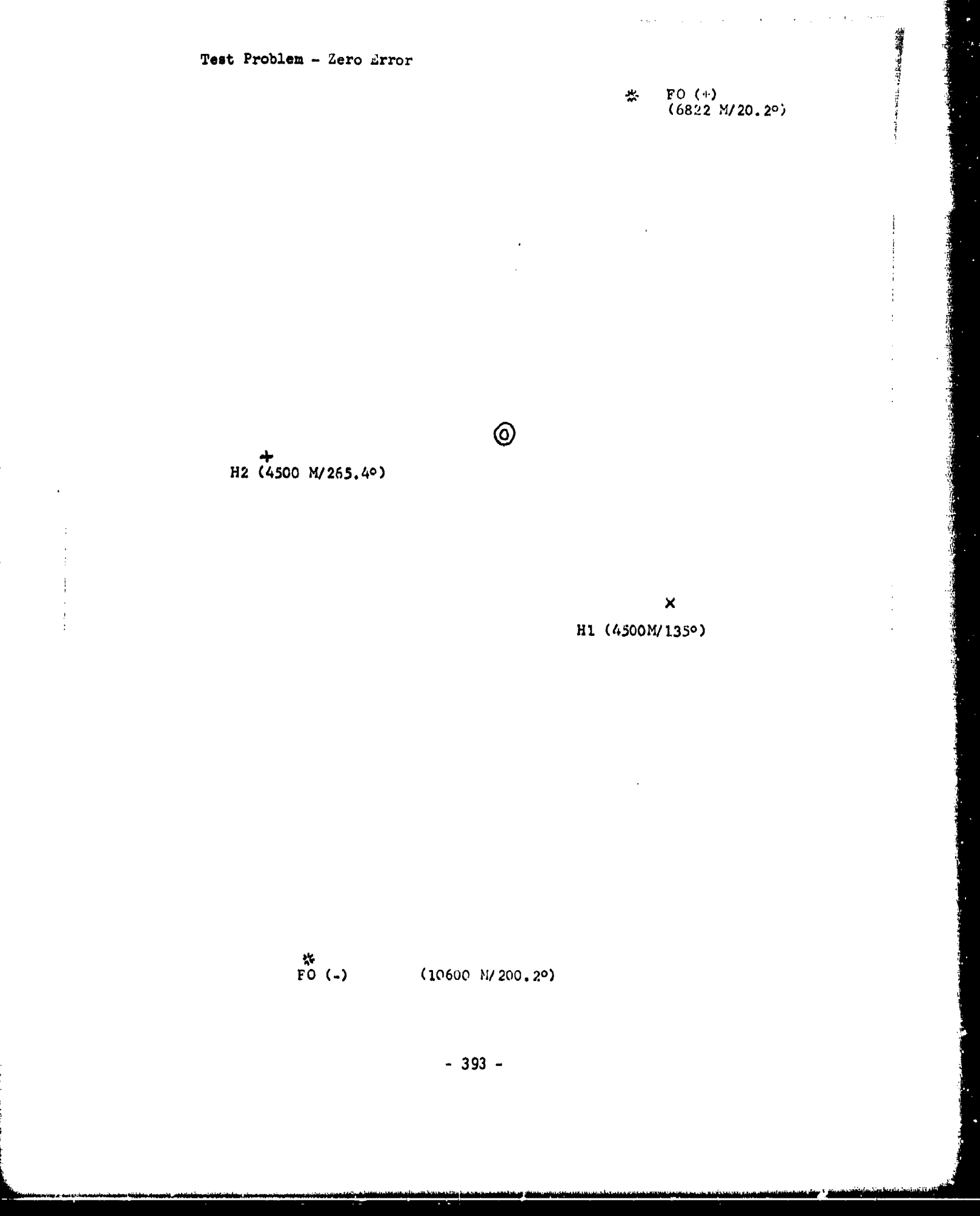
ROUND DISPERSION





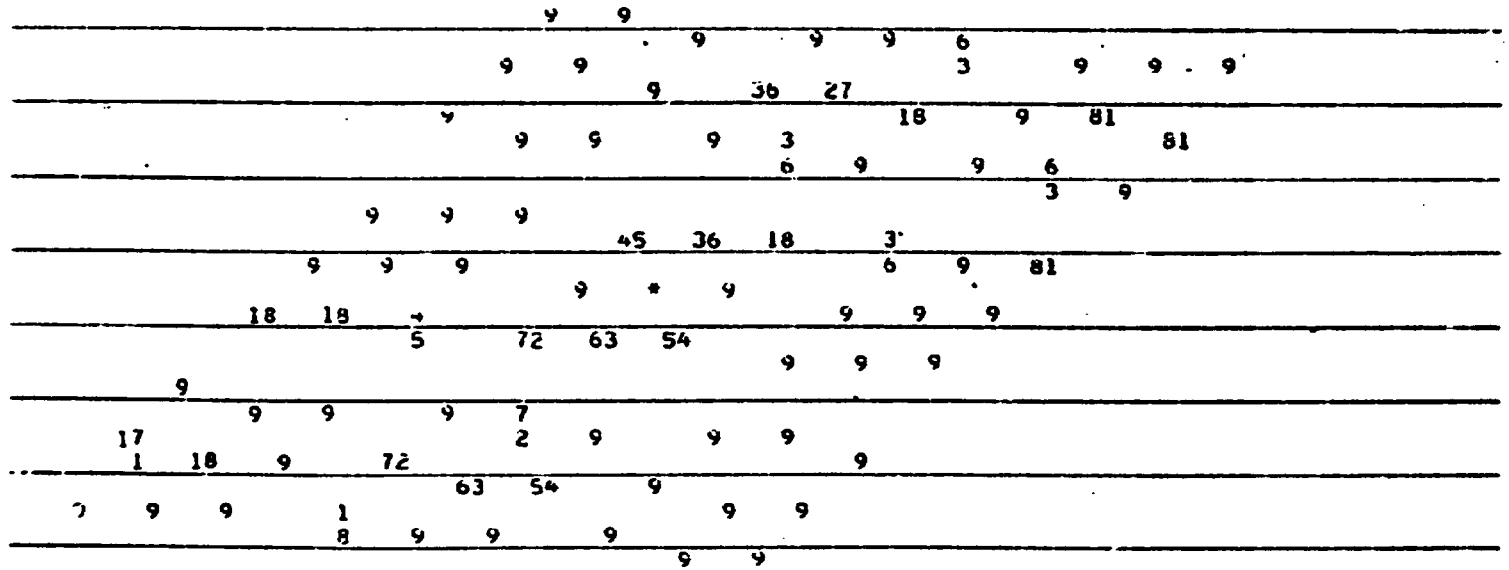
SLIDE 8

Test Problem - Zero Error



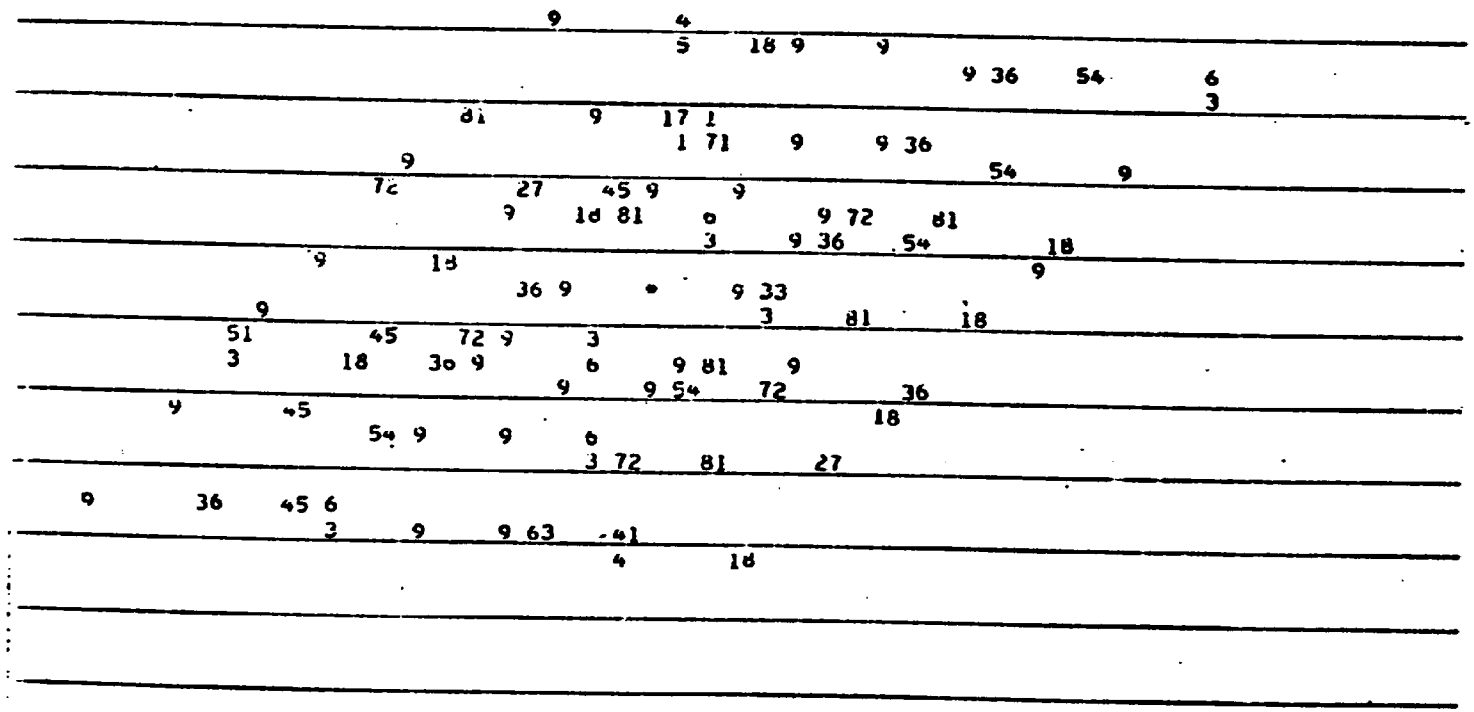
SCALE = 7.2 INCHES/50 UNITS = 0.87 9 POINTS = 729

FO- .259



SCALE = 9.0 INCHES/50 UNITS = 5.53 9 POINTS = 229

FO- 50%



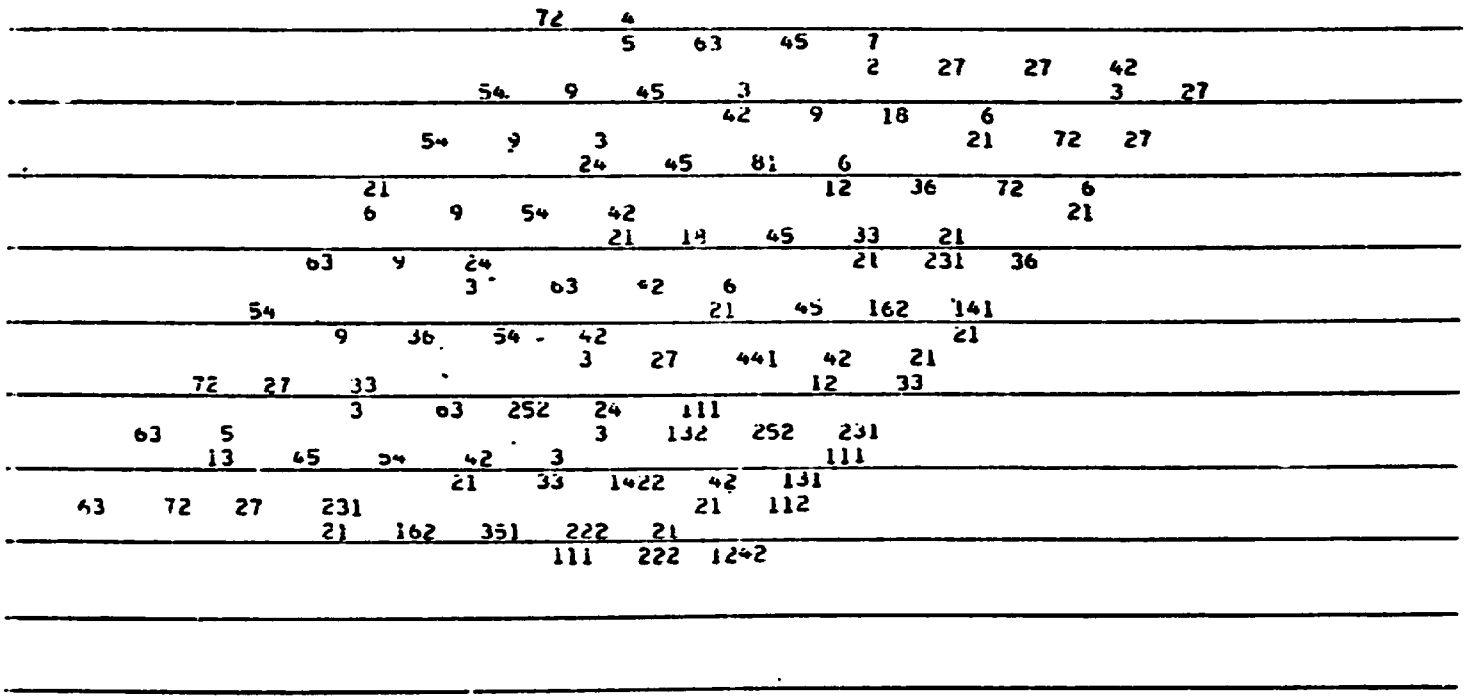
SCALE = 11.5 INCHES/50 UNITS = 0.36 76 POINTS = 729

FO - 75%.

9
2G Q1
87 5
13
G2 5
50 •3
PB 1E
21
M5 7•
74 1•4 33
YE 6L
E4 8P
1•3 F8
82 4E
72 6C
102 C6
36

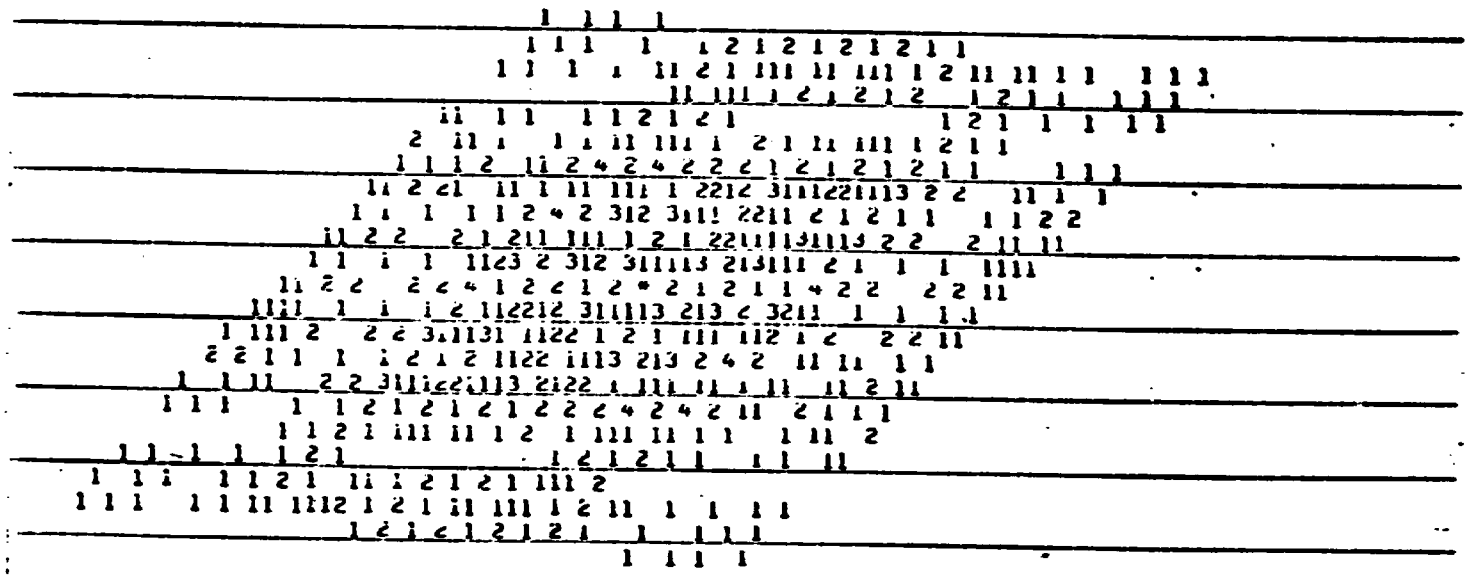
SCALE = 23.4 INCHES/50 UNITS = 2.14 7 POINTS = 729

FO- 100%



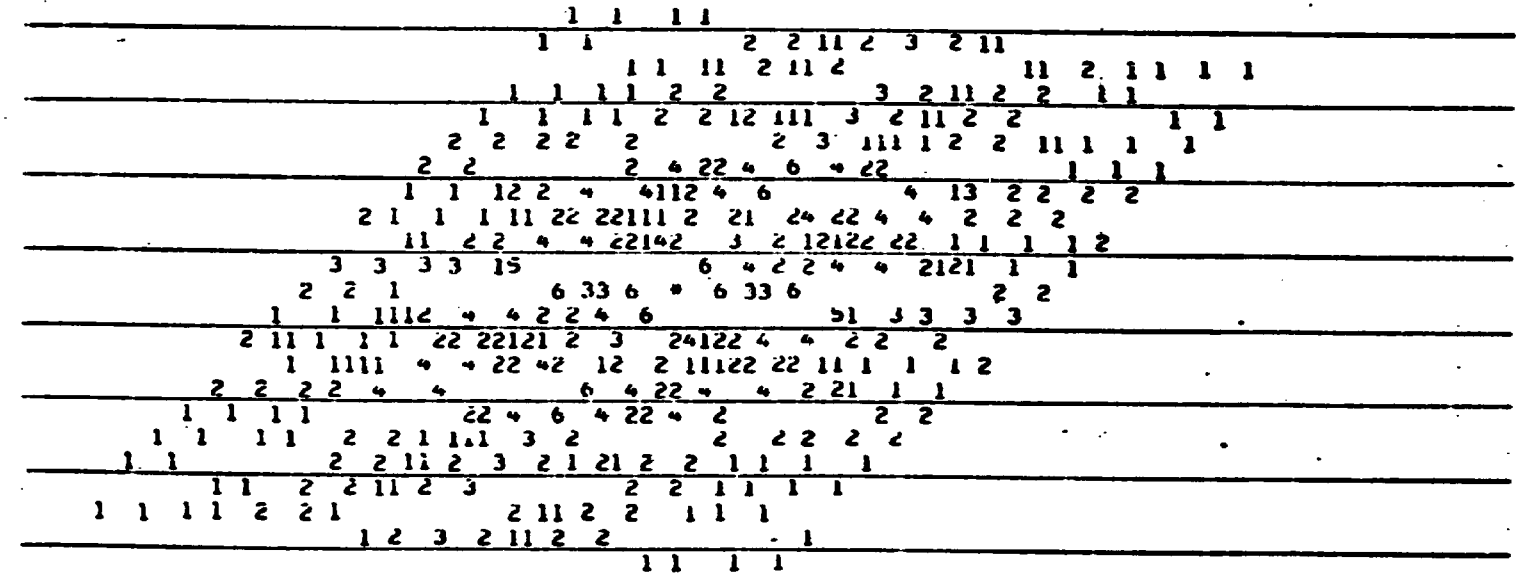
SCALE = 10.4 INCHES/50 UNITS = 4.80 1 POINTS = 729

Fo+ 25%



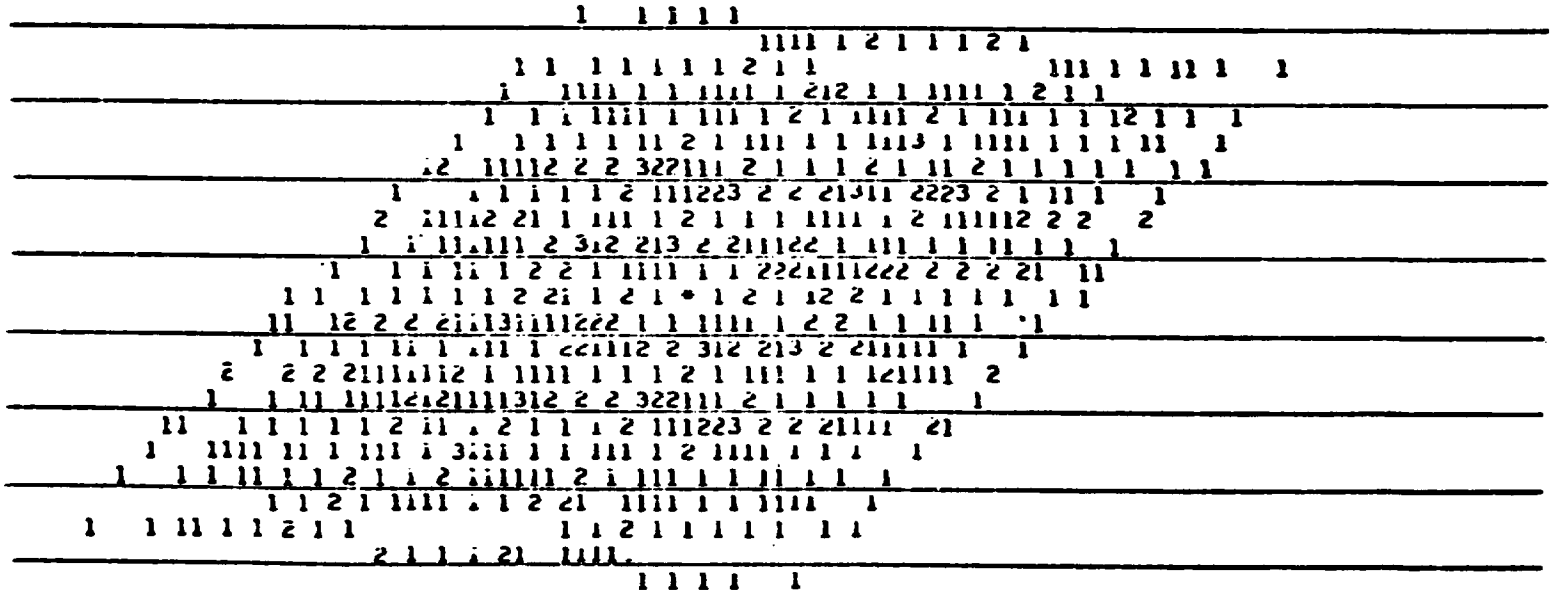
SCALE = 15.6 INCHES/50 UNITS = 3.20 9 POINTS = 729

F0+ 50%



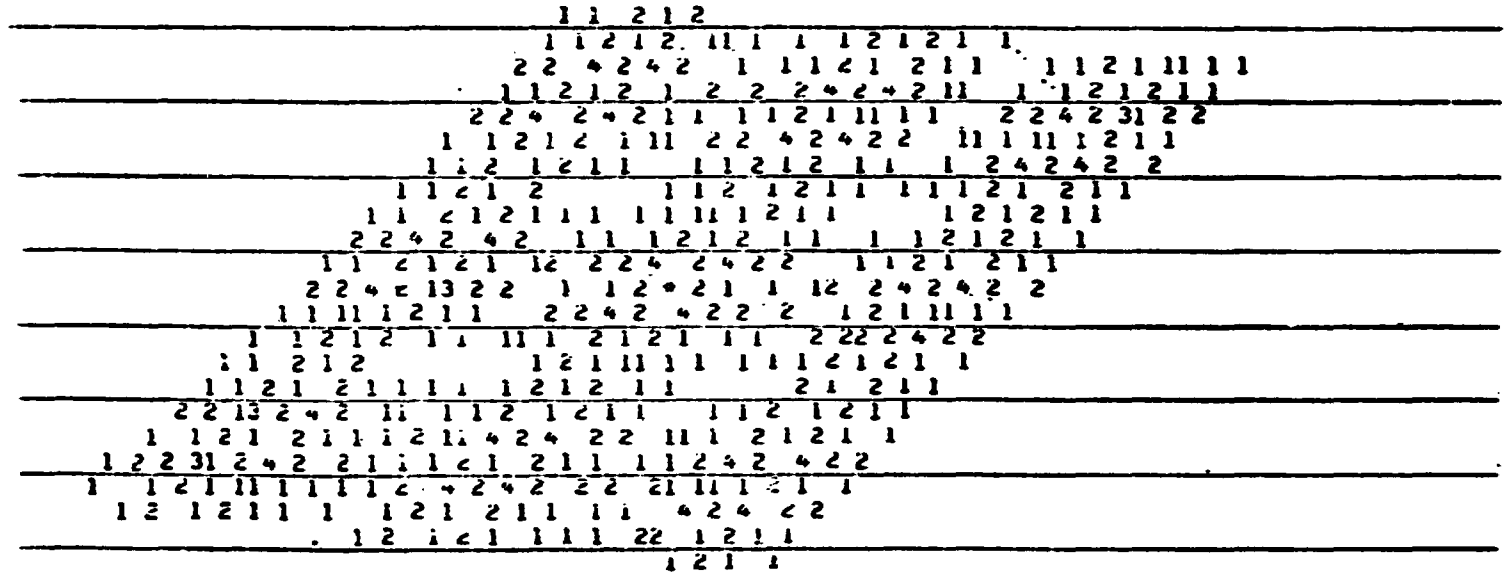
SCALE = 22.6 INCHES/50 UNITS = 2.19 1 POINTS = 729

Fo+ 75%



SCALE = 57.7 INCHES/50 UNITS = .87 1 POINTS = 729

F0 + 100%



CONCLUSION

The plots attached for FO- are an example of the minimum error radius configuration discovered during the research for the project. Please note that this plot represents extreme mortar ranges (4500 meters) and near extreme rangefinder ranges (9620 meters). The resulting 100% probability outline for FO- is contained within a circumscribed circle of radius 100 meters. The same outline for FO+ (not in minima configuration) requires a 240 meter radius. The error is larger even though the FO is closer to the firing point.

The following diagram illustrates the minima criterion:

The minima criterion is that θ_2 and θ_1 both be identically 90°. To perform this feat requires actual knowledge of the FO's position. It was found, however, that error sensitivity to the actual valves of θ_1 and θ_2 is low, therefore, it is necessary to know only an approximate location for the FO and establish H_1 and H_2 to produce a minimal error condition.

Enlarged View Of Minima Criteria

QUALITATIVE COMPARISON

Seven techniques of FO location were employed in the statistical analysis included in this portion of the report

Technique	Description (on basis of required information)
A	Range and Bearing to Hit One
B	Range and Bearing to Hit Two
C	Ranges only to Hits One and Two
D	Bearings only to Hits One and Two
E	Average Locations Found in A,B
F	Average Locations Found in C,D
G	Average Locations Found in A,B,C

Number of times (out of 36) technique showed least value of error statistic.

Technique	Mean	Standard Deviation	Standard Error of Mean	SD + SEM
A	0	0	0	0
B	0	0	1	1
C	2	0	0	2
D	0	0	0	0
E	10	34	33	18
F	0	0	0	0
G	24	2	2	13

Additionally, Technique E and G turn out to be statistically similar. This result is logical as no different information is incorporated when utilizing G. Techniques A,B,C have similar values also, as should be expected. Ranking of these techniques according to statistical accuracy would be:

Best	E,G
Fair	A,B,C
Worst	D,F

The added criterion of simplicity causes E to be chosen as the best over-all technique.

STATISTICAL COMPARISON OF LOCATING TECHNIQUES

The following pages contain Computer Outputs comparing the seven techniques tested as listed under "Qualitative Analysis". The first set of 36 samples assumes the $\pm$ 8 degree compass error. The second set assumes $\pm$ 2 degrees. Column headings are:

MEAN (Expected Value)
SD (Standard Deviation)
SEM (Standard Error of the MEAN)
SKEW (Skewness)
KURT (Kurtosis)

COMPUTER STATISTICAL OUTPUT:

±8 DEGREES COMPASS ERROR

MEAN	SD	SEM	SKEW	KURT
SAMPLE SET NUMBER 7. SAMPLE SIZE IS 47				
213.0	180.7	27.2	66.1	233.2
171.8	144.4	21.1	74.3	253.3
103.5	210.8	30.7	230.6	1434.4
679.1	917.1	133.8	156.7	723.2
151.7	130.4	19.0	97.0	416.5
345.0	461.1	67.3	152.6	692.8
110.4	97.1	14.2	88.6	218.9
SAMPLE SET NUMBER 8. SAMPLE SIZE IS 146				
191.4	176.7	14.6	260.9	971.2
199.4	191.1	15.8	306.7	1883.1
141.5	317.6	26.3	735.5	4329.7
1174.9	3008.0	248.8	824.1	5487.8
154.5	116.8	9.8	207.9	783.4
600.0	1512.4	125.2	827.8	5557.6
131.3	127.8	10.6	307.8	1834.6
SAMPLE SET NUMBER 9. SAMPLE SIZE IS 66				
204.5	176.0	21.3	73.1	229.5
215.9	217.4	26.4	135.7	443.7
200.9	533.8	64.7	272.8	1243.5
937.7	1368.0	165.9	193.4	768.2
163.9	118.8	14.4	79.0	244.0
524.2	763.8	92.6	202.1	850.3
153.3	161.4	19.8	193.9	835.3
SAMPLE SET NUMBER 10. SAMPLE SIZE IS 96				
204.5	101.8	21.8	64.3	251.3
193.5	207.2	27.7	90.2	277.9
206.7	826.3	84.0	254.5	1320.2
1338.5	4720.9	630.9	300.3	2546.3
162.7	134.3	17.9	85.9	278.3
717.5	2373.3	317.1	350.8	2454.5
150.9	204.0	27.1	179.3	785.7
SAMPLE SET NUMBER 11. SAMPLE SIZE IS 62				
212.9	174.3	22.1	71.7	242.4
186.3	142.8	18.1	51.6	164.9
240.8	600.9	83.9	209.7	1340.7
1750.7	5937.3	754.0	427.5	3212.3
140.3	104.8	13.3	61.7	207.9
947.1	3009.1	381.6	411.2	3032.9
152.4	217.2	27.6	245.8	1200.1
SAMPLE SET NUMBER 12. SAMPLE SIZE IS 150				
243.3	227.8	18.6	256.8	895.5
208.3	163.6	13.4	210.5	742.5
220.8	779.8	63.7	837.4	5277.8
1718.5	5069.3	413.9	1213.2	11703.5
163.7	127.3	10.4	207.9	762.4
878.4	2499.1	204.1	1195.8	11475.9
162.9	255.4	21.7	745.3	4532.8

MEAN	SD	SEM	SKEW	KURT
SAMPLE SET NUMBER 13, SAMPLE SIZE IS 155				
220.1	199.0	16.0	276.4	1118.5
242.9	205.3	16.5	243.5	831.4
165.1	770.1	62.3	1226.4	10311.6
1180.4	1899.8	152.6	767.1	5722.6
124.4	124.4	10.0	206.5	786.9
622.3	1011.5	81.2	696.0	4663.6
145.9	257.8	20.7	1058.9	8512.9
SAMPLE SET NUMBER 14, SAMPLE SIZE IS 134				
247.3	234.2	20.2	254.7	945.6
205.3	195.9	16.9	223.7	732.2
539.7	851.8	73.6	955.3	7358.1
2317.5	13127.0	1134.0	1340.1	15339.4
173.2	130.6	11.3	163.2	554.4
1211.8	6616.4	571.7	1361.1	14827.8
162.7	274.1	23.7	850.5	6329.3
SAMPLE SET NUMBER 15, SAMPLE SIZE IS 165				
227.3	185.0	14.4	165.2	647.2
223.6	179.7	14.0	223.9	840.0
956.9	956.5	74.5	827.0	4894.7
3337.2	16494.0	1284.1	1424.1	13010.5
160.3	120.8	9.4	253.2	996.2
1714.9	8242.8	641.7	1423.4	13008.8
184.1	313.0	24.4	781.7	6678.0
SAMPLE SET NUMBER 16, SAMPLE SIZE IS 58				
190.0	150.8	19.8	52.7	216.6
183.9	183.7	24.1	96.6	208.5
178.1	507.8	66.7	315.3	2008.3
49734.7	*0199.1	48609.5	419.1	3137.4
141.5	108.2	14.2	69.1	219.9
24900.9	*5094.9	24304.2	419.1	3137.4
130.2	178.2	23.4	240.7	1446.6
SAMPLE SET NUMBER 17, SAMPLE SIZE IS 106				
205.8	176.8	17.2	144.6	459.3
220.4	224.7	21.8	210.5	783.4
283.2	885.0	86.0	505.7	2738.5
11782.5	*0483.1	9759.8	1051.7	10693.8
162.5	137.4	13.3	229.3	1005.6
5916.3	50230.5	4879.4	1051.8	10654.9
180.5	281.1	27.3	446.2	2324.7
SAMPLE SET NUMBER 18, SAMPLE SIZE IS 37				
284.1	287.7	47.3	90.6	350.5
193.8	102.7	26.7	32.2	95.5
270.6	877.9	144.3	139.3	579.9
1483.6	3624.2	595.8	180.2	1062.1
174.6	162.9	26.8	82.5	308.7
616.6	2012.7	330.9	165.3	1056.3
172.2	307.1	50.5	113.1	440.4

MEAN	SD	SEM	SKEW	KURT
SAMPLE SET NUMBER 19, SAMPLE SIZE IS 71				
193.2	200.3	22.3	177.1	634.1
205.1	219.3	24.4	202.1	872.2
222.4	642.3	71.4	419.2	2547.7
642.4	1160.1	128.9	232.8	991.1
166.7	143.1	15.9	120.0	340.3
479.5	631.5	70.2	169.2	711.5
163.1	220.0	24.4	292.0	1561.4
SAMPLE SET NUMBER 20, SAMPLE SIZE IS 67				
179.9	153.0	18.8	93.3	335.6
217.5	179.4	21.4	62.4	204.3
237.7	666.3	81.7	324.5	1786.2
1440.3	2907.4	350.3	284.0	1456.9
141.4	91.0	11.1	24.7	147.8
703.5	1441.6	176.1	276.3	1408.6
142.6	204.8	25.0	279.4	1454.0
SAMPLE SET NUMBER 21, SAMPLE SIZE IS 74				
186.7	144.1	16.7	98.6	330.9
213.6	164.7	19.1	80.1	260.2
406.3	998.8	116.0	264.3	1177.4
2401.4	6433.4	747.9	375.9	2382.3
128.7	96.9	10.6	106.4	392.7
1379.3	3331.8	387.3	302.0	2201.8
187.1	311.1	36.2	259.4	1153.2
SAMPLE SET NUMBER 22, SAMPLE SIZE IS 129				
206.7	173.8	15.3	148.0	467.6
219.9	176.6	15.5	115.1	381.0
253.8	747.9	65.9	568.0	2984.9
9842.4	99279.5	8741.1	1430.7	16123.3
136.3	118.5	10.4	172.1	580.2
4895.8	49634.7	4370.1	1430.6	16121.8
188.3	237.2	20.9	488.8	2485.8
SAMPLE SET NUMBER 23, SAMPLE SIZE IS 134				
209.3	176.5	15.2	208.9	721.9
217.6	195.9	16.9	203.1	705.7
148.7	315.2	27.2	737.6	5226.4
1273.3	3569.4	310.0	1014.9	9078.7
161.1	107.3	9.3	174.4	725.2
631.3	1781.5	153.9	1011.5	9137.2
127.3	113.3	9.8	387.2	2223.1
SAMPLE SET NUMBER 24, SAMPLE SIZE IS 30				
214.9	246.0	44.9	43.5	122.8
193.5	164.6	30.1	36.8	108.5
131.5	261.7	47.8	114.1	547.0
976.4	1761.9	321.7	105.0	479.0
149.0	112.0	20.4	26.0	82.8
494.5	903.7	165.0	106.7	490.9
114.7	121.7	22.2	77.4	328.6

MEAN	SD	SEM	SKW	KURT
SAMPLE SET NUMBER 25, SAMPLE SIZE IS 105				
215.1	201.6	19.7	167.1	703.5
197.0	169.0	16.5	119.2	381.9
134.7	255.7	25.0	431.4	2279.0
3210.1	15450.3	1556.6	745.8	5788.9
159.8	116.2	11.5	132.9	452.8
1623.2	7973.3	778.1	745.4	5784.9
127.6	116.5	11.4	193.6	703.8
SAMPLE SET NUMBER 26, SAMPLE SIZE IS 83				
220.6	207.4	22.8	134.1	443.5
194.1	184.4	20.2	129.4	454.4
160.4	490.5	53.8	500.9	4438.7
2378.9	9607.3	1054.5	491.1	3142.2
157.0	110.0	12.1	71.8	254.1
1232.2	4807.3	527.7	488.9	3122.8
138.2	172.7	19.0	426.7	3047.9
SAMPLE SET NUMBER 27, SAMPLE SIZE IS 143				
217.1	195.5	16.4	282.4	1252.6
222.9	208.1	17.4	287.2	1144.7
271.3	822.0	68.7	606.3	3602.5
3774.8	24600.2	2480.3	1651.4	19514.5
151.7	115.1	9.6	202.7	1078.4
1959.5	14830.7	1240.2	1649.8	19477.2
167.1	259.0	21.6	622.0	3423.4
SAMPLE SET NUMBER 28, SAMPLE SIZE IS 30				
202.9	184.6	33.7	85.9	349.9
173.5	178.4	32.6	47.8	137.4
152.1	322.9	59.0	107.9	485.7
1036.1	1424.8	260.1	74.7	261.8
148.9	116.9	21.3	39.8	123.2
536.0	738.7	134.9	77.4	275.4
114.7	123.2	22.5	66.9	261.3
SAMPLE SET NUMBER 29, SAMPLE SIZE IS 58				
192.1	168.5	22.1	97.3	191.0
252.8	204.3	26.8	80.5	314.1
249.8	708.6	92.8	257.1	1349.1
1494.6	2404.5	343.6	242.1	1276.9
130.2	102.6	13.5	54.3	207.3
721.2	1056.6	138.2	173.8	785.6
157.8	220.5	28.9	230.6	1199.0
SAMPLE SET NUMBER 30, SAMPLE SIZE IS 60				
202.1	164.5	21.2	87.3	342.3
202.7	193.6	25.3	94.1	314.3
213.7	716.3	92.3	339.5	2161.8
1267.7	3208.7	414.2	360.7	2536.5
144.3	119.5	15.0	138.8	615.7
675.2	1670.1	215.6	329.0	2161.7
140.9	240.6	31.8	280.5	1020.6

MEAN	SD	SEM	SKEW	KURT
SAMPLE SET NUMBER 31 SAMPLE SIZE IS 104				
208.1	100.5	16.3	155.0	545.2
233.4	223.2	21.9	244.6	1190.1
114.5	326.3	32.0	749.2	6143.6
6264.5	26419.2	2590.6	902.1	8431.8
101.8	122.3	12.0	183.0	801.9
148.6	13210.2	1295.4	902.0	8429.9
123.4	424.7	12.2	387.1	2392.7
SAMPLE SET NUMBER 32 SAMPLE SIZE IS 166				
140.3	172.2	13.4	217.4	723.3
208.1	173.5	13.5	187.6	585.9
187.4	592.0	46.0	1314.5	12352.2
1578.7	6313.2	490.0	1305.1	13055.1
140.7	110.8	8.6	177.4	667.2
830.0	3170.8	246.1	1340.6	12734.0
135.9	201.9	15.7	1130.1	10257.7
SAMPLE SET NUMBER 33 SAMPLE SIZE IS 113				
218.8	184.8	17.4	210.6	890.7
217.2	200.1	18.8	170.0	534.1
243.9	748.0	74.1	571.8	3145.1
23518.9	*1347.9	21763.4	1108.9	12314.7
160.1	135.7	12.6	154.8	541.2
11777.8	*5871.1	10881.4	1108.9	12314.5
155.7	280.4	26.5	520.0	3017.2
SAMPLE SET NUMBER 34 SAMPLE SIZE IS 48				
260.3	205.9	29.7	48.0	156.3
291.1	255.5	36.9	74.5	277.6
235.2	800.4	116.4	269.3	1703.9
1752.5	4635.2	669.0	292.0	1533.8
200.8	159.1	23.0	89.3	265.9
940.3	2378.1	343.3	234.4	1370.6
184.3	207.2	38.6	233.7	1431.9
SAMPLE SET NUMBER 35 SAMPLE SIZE IS 106				
225.0	179.6	17.4	125.1	389.4
230.6	210.8	20.5	139.3	454.1
242.8	701.2	68.1	517.7	2979.3
2908.7	10195.7	990.3	681.6	5182.7
177.7	124.4	12.1	123.8	515.5
1465.7	5104.2	495.8	678.8	5152.6
173.2	238.8	23.2	494.2	2559.6
SAMPLE SET NUMBER 36 SAMPLE SIZE IS 78				
174.1	159.7	18.1	122.3	423.1
181.9	169.4	19.2	93.5	298.6
443.5	1272.4	144.1	271.9	1123.2
2937.8	18462.4	2090.5	652.1	5653.0
150.6	117.7	13.3	113.0	426.5
1904.2	9403.5	1064.7	646.3	5594.8
217.1	395.5	44.8	204.9	1132.7

COMPUTER STATISTICAL OUTPUT:

±2 DEGREES COMPASS ERROR

MEAN	SD	SEM	SKEW	KURT
SAMPLE SET NUMBER 7, SAMPLE SIZE IS 47				
60.9	43.7	6.4	50.5	186.9
54.3	39.4	5.7	82.0	307.0
103.5	210.8	30.7	236.6	1434.4
202.2	390.6	57.0	228.2	1343.9
41.6	32.4	4.7	85.9	326.7
123.3	212.5	31.0	185.6	942.5
50.3	68.4	10.0	207.2	1195.2
SAMPLE SET NUMBER 8, SAMPLE SIZE IS 146				
63.8	44.7	3.7	220.3	876.8
60.0	46.2	4.0	302.9	1626.7
141.5	317.6	26.3	735.5	4329.7
1330.5	12699.3	1051.0	1722.8	20654.2
48.2	30.5	2.5	167.7	650.7
703.6	6418.6	531.2	1721.1	20627.5
69.8	105.1	8.7	671.2	3819.8
SAMPLE SET NUMBER 9, SAMPLE SIZE IS 68				
62.9	43.9	5.3	68.0	239.0
67.1	56.3	6.8	122.0	424.8
200.9	533.8	64.7	272.6	1243.5
361.5	1236.4	149.9	435.5	3183.8
49.5	31.2	3.8	68.3	263.1
237.0	631.9	76.6	398.2	2809.2
89.4	169.8	20.6	271.0	1250.3
SAMPLE SET NUMBER 10, SAMPLE SIZE IS 56				
65.3	41.9	5.6	58.3	212.7
61.4	52.8	7.1	74.0	224.8
208.7	626.3	84.0	254.5	1320.2
441.9	1736.1	232.0	373.6	2704.8
49.2	36.9	4.9	70.9	243.2
283.5	925.7	123.7	324.8	2173.5
89.1	204.1	27.3	243.1	1228.9
SAMPLE SET NUMBER 11, SAMPLE SIZE IS 62				
66.0	43.9	5.6	55.8	218.7
57.3	36.9	4.7	36.1	156.0
240.6	660.8	83.9	269.7	1340.7
840.1	3994.1	507.3	435.6	3276.3
45.1	27.0	3.4	35.5	150.8
450.9	1974.8	250.8	442.3	3364.7
97.4	215.7	27.4	273.0	1364.3
SAMPLE SET NUMBER 12, SAMPLE SIZE IS 150				
72.9	56.5	4.6	230.4	783.8
63.2	40.6	3.3	187.0	720.0
226.6	774.8	63.7	837.9	5277.8
640.4	2526.7	206.5	1334.4	13767.1
48.4	32.6	2.7	177.7	683.3
360.6	1276.2	104.4	1204.3	12915.6
96.4	259.1	21.2	837.9	5278.3

MEAN	SD	SEM	SKEW	KURT
SAMPLE SET NUMBER 13, SAMPLE SIZE IS 155				
66.3	48.2	3.9	244.0	960.9
72.4	52.1	4.2	204.4	708.5
185.1	776.1	62.3	1228.4	10311.6
855.2	4261.3	342.3	1490.6	16163.5
48.3	32.0	2.6	186.1	733.8
479.5	2161.8	173.6	1423.0	15126.7
79.7	256.1	20.6	1217.9	10198.8
SAMPLE SET NUMBER 14, SAMPLE SIZE IS 134				
73.9	57.3	4.9	260.9	1010.6
64.0	47.6	4.1	183.9	616.8
216.7	851.8	73.6	955.3	7368.1
491.5	2192.5	189.4	1285.5	13587.7
50.8	32.1	2.8	126.0	498.5
312.1	1341.9	115.9	1272.8	13466.7
92.3	276.7	24.1	954.0	7365.4
SAMPLE SET NUMBER 15, SAMPLE SIZE IS 165				
68.2	44.3	3.4	174.7	698.5
68.4	47.7	3.7	221.6	832.6
306.9	956.5	74.5	827.0	4894.7
525.2	1701.1	132.4	1365.1	14035.9
49.0	31.7	2.5	230.6	952.3
355.5	1020.6	79.5	1088.8	9502.5
121.1	313.6	24.4	827.9	4935.1
SAMPLE SET NUMBER 16, SAMPLE SIZE IS 58				
62.5	40.1	5.3	42.5	193.3
57.8	45.4	6.0	82.3	256.1
178.1	507.8	66.7	315.3	2008.3
246.9	446.9	58.7	194.2	800.5
43.2	29.3	3.8	54.7	190.3
176.2	316.8	41.6	177.1	732.2
77.2	168.5	22.1	307.5	1935.6
SAMPLE SET NUMBER 17, SAMPLE SIZE IS 106				
66.9	43.2	4.2	130.2	489.7
67.4	57.9	5.6	213.6	791.2
285.2	885.0	86.0	505.7	2738.5
453.3	1070.8	104.0	581.7	4134.2
49.0	35.6	3.5	206.8	842.9
307.9	643.9	62.5	396.8	2066.2
114.8	288.5	28.0	504.7	2735.7
SAMPLE SET NUMBER 18, SAMPLE SIZE IS 37				
74.7	69.9	11.5	83.5	325.6
61.5	41.1	6.8	27.2	89.9
276.6	877.9	144.3	139.3	579.9
279.1	412.6	67.9	157.1	855.1
51.1	41.0	6.7	72.3	274.6
207.7	391.6	64.4	135.2	568.6
118.0	291.6	47.9	136.9	506.1

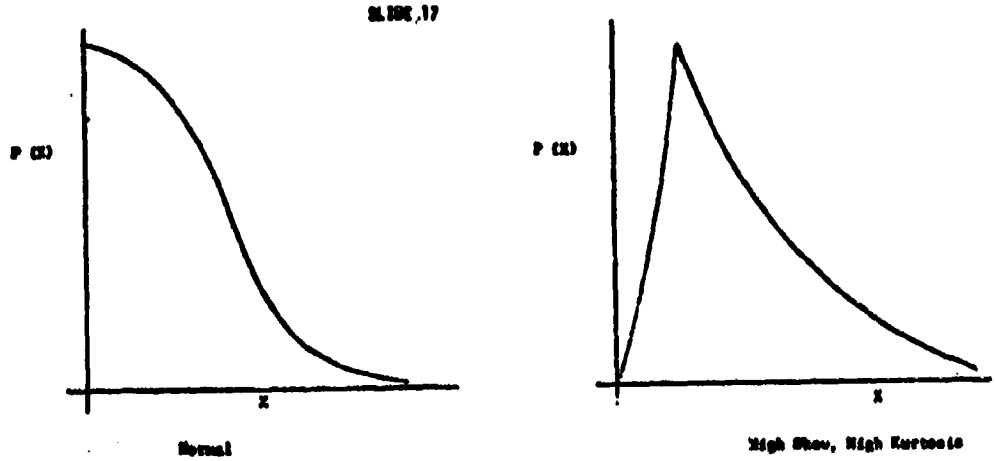
MEAN	SD	SEM	SKEW	KURT
SAMPLE SET NUMBER 19, SAMPLE SIZE IS 61				
64.1	51.1	5.7	140.2	498.1
65.8	54.7	6.1	182.3	767.3
222.4	642.3	71.4	419.2	2597.7
223.9	327.4	36.4	242.3	1039.0
53.1	37.0	4.1	114.1	374.0
189.2	354.7	39.4	319.8	1647.7
97.0	210.2	23.4	401.4	2438.1
SAMPLE SET NUMBER 20, SAMPLE SIZE IS 67				
60.0	39.5	4.8	77.4	299.7
61.3	43.5	5.3	86.9	175.1
237.7	668.5	81.7	324.5	1786.2
307.8	437.5	53.5	170.5	598.8
45.6	25.8	3.2	29.9	145.7
231.3	350.0	43.5	190.9	776.4
97.8	214.9	26.3	322.7	1773.9
SAMPLE SET NUMBER 21, SAMPLE SIZE IS 74				
62.6	36.8	4.3	88.4	373.8
64.1	43.2	5.0	65.1	217.7
406.3	998.0	116.0	264.3	1177.4
655.9	1599.9	186.0	366.5	2369.9
42.0	25.7	3.0	67.6	280.3
426.5	902.1	104.9	281.5	1524.9
149.3	325.7	37.9	265.5	1184.4
SAMPLE SET NUMBER 22, SAMPLE SIZE IS 129				
61.0	43.5	3.8	117.4	402.8
58.9	42.5	3.7	79.3	328.1
253.8	747.9	65.9	500.0	2989.9
337.1	1060.3	93.4	1055.8	9910.3
47.8	29.8	2.5	135.0	555.0
234.1	544.2	47.9	581.0	3321.8
105.3	243.6	21.4	567.5	3007.8
SAMPLE SET NUMBER 23, SAMPLE SIZE IS 134				
63.8	44.9	3.9	219.0	894.3
68.1	48.3	4.2	170.8	678.4
146.7	315.2	27.2	737.6	5226.4
312.9	620.3	54.1	578.9	3154.2
46.8	27.1	2.3	162.0	702.8
190.0	359.8	31.1	570.6	3206.1
69.4	99.4	8.6	724.8	5238.6
SAMPLE SET NUMBER 24, SAMPLE SIZE IS 30				
65.7	64.1	11.7	35.0	91.9
62.5	37.9	6.9	30.7	104.3
131.5	261.7	47.8	114.1	547.0
261.5	456.7	83.4	120.3	507.4
45.8	31.2	5.7	22.6	83.6
148.3	246.8	45.1	118.6	587.1
63.9	89.4	16.3	119.1	593.0

MEAN	SD	SEM	SKEW	KURT
SAMPLE SET NUMBER 25, SAMPLE SIZE IS 105				
64.7	49.6	4.8	153.6	544.2
61.3	44.9	4.4	92.2	319.3
134.7	255.7	25.0	431.4	2279.0
394.1	1292.3	126.1	691.3	5137.1
46.9	31.7	3.1	126.3	448.3
229.6	676.4	66.0	660.6	4676.9
67.0	88.1	8.6	377.8	1883.1
SAMPLE SET NUMBER 26, SAMPLE SIZE IS 83				
70.1	52.8	5.8	123.8	425.0
61.8	47.3	5.2	107.1	389.7
160.4	490.5	53.8	566.9	4438.7
515.2	1796.5	197.2	522.8	3809.6
46.9	30.1	3.3	67.8	253.1
303.8	920.9	101.7	470.2	3224.5
75.8	164.7	18.1	563.4	4427.6
SAMPLE SET NUMBER 27, SAMPLE SIZE IS 143				
68.2	49.5	4.1	249.5	1108.2
66.6	53.9	4.5	260.1	983.3
271.3	822.0	68.7	606.3	3602.5
1854.4	19113.5	1598.4	1673.6	19872.3
48.1	30.2	2.5	225.0	1009.1
1004.6	9564.7	799.8	1670.0	19815.1
111.1	267.8	22.4	670.2	3676.2
SAMPLE SET NUMBER 28, SAMPLE SIZE IS 30				
63.9	50.8	9.3	71.9	320.2
49.1	45.1	8.2	53.6	166.6
152.1	322.9	59.0	107.9	485.7
487.6	1030.7	188.2	99.1	391.1
42.6	29.3	5.4	42.7	160.4
263.7	518.9	94.7	97.0	378.4
63.0	105.0	19.2	113.0	533.5
SAMPLE SET NUMBER 29, SAMPLE SIZE IS 58				
59.5	43.3	5.7	84.5	341.4
70.5	50.7	6.7	87.6	338.9
249.8	708.6	92.8	257.1	1349.1
345.3	476.4	62.8	132.0	438.9
43.6	27.2	3.6	37.0	170.5
243.4	369.5	48.5	138.1	476.2
100.5	228.9	30.1	255.8	1340.5
SAMPLE SET NUMBER 30, SAMPLE SIZE IS 60				
62.0	43.2	5.6	71.9	290.6
63.6	50.0	6.5	89.8	274.1
213.7	715.3	92.3	339.5	2131.0
275.4	417.4	53.9	197.0	938.5
44.8	29.1	3.8	118.8	537.7
194.5	479.5	61.9	343.5	2332.4
87.4	238.3	30.8	333.5	2097.5

MEAN	SD	SEM	SKEW	KURT
SAMPLE SET NUMBER 31, SAMPLE SIZE IS 104				
66.6	45.7	4.5	124.6	462.5
72.2	58.0	5.7	231.5	1067.8
114.5	326.3	32.0	749.2	6143.6
428.2	1293.2	126.8	647.2	5650.2
49.9	34.3	3.4	133.7	514.5
235.1	651.9	63.9	676.1	5425.0
60.9	106.7	10.5	732.0	6010.1
SAMPLE SET NUMBER 32, SAMPLE SIZE IS 166				
63.3	44.6	3.5	208.6	726.0
65.4	43.0	3.3	164.0	593.9
167.4	592.0	46.0	1314.5	12352.2
274.5	612.0	47.5	957.3	7344.4
47.0	28.7	2.2	152.4	659.3
187.9	444.3	34.5	829.0	5151.2
75.9	197.4	15.3	1319.8	12483.5
SAMPLE SET NUMBER 33, SAMPLE SIZE IS 113				
69.0	47.8	4.5	181.1	768.2
67.6	49.2	4.6	151.6	500.7
243.9	788.0	74.1	571.8	3185.1
422.4	1323.1	124.7	978.1	9611.4
51.2	35.9	3.4	141.2	509.8
275.8	908.0	85.4	938.5	8949.0
101.0	259.7	24.4	581.4	3350.7
SAMPLE SET NUMBER 34, SAMPLE SIZE IS 48				
77.0	50.2	7.2	37.4	132.6
79.6	67.4	9.7	72.8	270.4
230.2	808.4	116.4	269.3	1703.9
389.5	711.7	102.7	168.9	803.5
57.0	40.5	5.8	73.3	302.6
272.5	582.0	84.0	192.9	997.7
101.8	259.3	37.4	265.9	1675.9
SAMPLE SET NUMBER 35, SAMPLE SIZE IS 106				
68.0	47.5	4.6	108.1	371.1
71.5	52.2	5.1	129.1	433.9
242.8	701.2	68.1	517.7	2979.3
523.9	1162.1	112.9	507.8	3314.2
53.2	35.2	3.4	122.7	560.8
321.9	760.1	74.4	473.5	2629.1
103.5	230.2	22.4	518.6	3017.2
SAMPLE SET NUMBER 36, SAMPLE SIZE IS 78				
58.4	43.5	4.9	106.5	401.3
60.6	43.4	4.9	99.9	402.9
443.5	1272.4	144.1	271.9	1123.2
4500.4	37342.9	4228.3	662.3	5774.7
49.0	31.7	3.6	117.4	422.6
2415.3	18785.0	2127.0	661.3	5763.3
167.6	412.9	46.8	274.0	1149.0

QUANTITATIVE COMPARISON

Generally, certain two-hit locating systems offer a reduction of order 1/3 in the mean error magnitude of locating the FO over the one-hit locating system. There is also a significant reduction in standard deviation and increase in kurtosis. All curves have a skewed characteristic.



Normal

High Skew, High Kurtosis

The extremely high skewness of the data indicates that simply, most of the time the error in location will be less than the mean (or expected error) but that occasionally outliers will occur. An increase in skewness coupled with a decrease in standard deviation indicates that the outliers will occur less frequently. The choice of the FO coordinate sets indicated appears from investigation to be the best compromise between time, accuracy and simplicity available at the state of technology assumed.

ASSUMPTIONS FOR INVESTIGATION

1. Flat, level ground for all points with no obstructions.
2. Round dispersions used for calculations were Gaussian according to firing tables (81mm Mortar was basis).
3. A rangefinder is available to the forward observer which has a $\pm$ 10 meter accuracy plus precision uniform error.
4. The compass techniques employed by the forward observer have as a dominant error $\pm$ 8 degrees Gaussian (3.5 standard deviation) resulting from local magnetic disturbances.

The Mortar is used throughout this discussion but the results should in no way be construed as useful only in Mortar Systems. All aimed hit locations, FO actual locations, and "errors" incorporated in calculation were generated randomly internal to the program. Sample sizes in each sample set were also generated randomly, restricted to the region 30 to 170. Uniform or Gaussian random numbers were used in the program as appropriate.

CONCLUSION

The use of the average of locations predicted by using two ranges with bearings between the Forward Observer and two hits has shown to be the best overall technique examined when consideration is given to time, accuracy and simplicity. Errors imposed include aiming, terrain, ranging, compass and round-to-round variations due to physical properties of ammunition and movement of the weapon.

ERRORS REDUCTION

Round-to-Round Variations: Ammunition properties are state-of-the-art and only highly improved technology can reduce this error.

Terrain: Terrain factors are uncontrollable.

Ranging: The Rangefinders implied are also near state-of-the-art with little improvement imminent.

Aiming and Movement: Aiming errors are individual and subject to training and experience. Weapon movement is minimized by re-aiming.

Compass: This source of error offers the most promise for improvement. $\pm$ 8 degree errors can be expected when using a magnetic device due to local disturbance. If a non-magnetic for determining bearing was produced which could hold within $\pm$ 2 degrees of the grid established by the fire direction center, the error in locating a Forward Observer can be reduced by a factor of three. Statistical comparison is shown on the two sets of computer listings.

PREDICTIVE EQUATIONS FOR ACCELERATED LIFE TESTING OF IMPREGNATED CHARCOAL

W.S. Magee, Jr., J.A. Baker, E.J. Poziomek,
J. Boardway and H. Ball
Development and Engineering Directorate
Physical Protection Branch
Edgewood Arsenal
Aberdeen Proving Ground, MD 21010

ABSTRACT. This paper presents a mathematical model and ensuing predictive equations for normal and accelerated life testing of impregnated charcoals under various conditions of temperature and water content. Predicted results are compared with data for charcoal in actual storage as well as laboratory scale accelerated storage. Analysis of the results gives information on the mechanism of degradation during storage.

INTRODUCTION. The general adsorbent used in respiratory protective systems is activated charcoal. For removal of toxic vapors from air, such systems prove more than adequate as measured by service times. However, for removal of toxic gases, such systems, although adequate, do not match the performance with vapors.

Fortunately, impregnation of activated charcoal with various substances markedly improves the performance with respect to removal of gases. A typical case is the enhanced service life for removal of cyanogen chloride by charcoals impregnated with an ammoniacal solution containing compounds of copper, silver and chromium. This case was extensively studied[1] during World War II.

Sealed storage at elevated temperatures and in the presence of moisture of respiratory protective devices containing such impregnated charcoal results in a deterioration of the improved service life eventually back to that for the unimpregnated material. Storage of dried impregnated material at virtually any temperature shows no deterioration. Although such sealed ,dry storage yields a virtually unlimited shelf life, devices used intermittently pick up moisture and degrade.

This study is concerned with the prediction of residual service lives of items stored sealed with known water pick-up and at a given temperature. The results are presented in terms of a mathematical model and compared with data for experimental charcoals.

2. DEVELOPMENT OF THE MODEL. Jonas and Rehrmann[2] modified an equation developed by Wheeler and Robell[3] and applied the result to studies on the service lives of fixed beds of charcoal exposed to a variety of gases and vapors. The modified equation is the form

$$t = W_e\left\{\frac{W}{C_0 Q} - \frac{\rho \ln(C_0/C)}{C_0 k}\right\} \qquad (1)$$

where t is the service life (or breakthrough time) of the bed as defined by a concentration C of the toxic material being attained in the effluent air stream, W_e is the dynamic saturation capacity of the bed in terms of weight of toxic picked up per unit weight of charcoal in the bed, W is the weight of charcoal in the bed, C_0 is the influent concentration of the toxic, Q is the volumetric flow rate, ρ is the bulk density of charcoal in the bed and k is a quasi first order rate constant for removal of the toxic molecules by sites on the charcoal. Similar equations have been reported[4,5] in the literature, their forms differing slightly depending upon which process in the bed is rate determining. A more complete discussion[6] of this will be published separately.

Equation (1) is applicable to systems in which the instantaneous effluent concentration C can be determined. Experimentally, cyanogen chloride service lives are determined by the collection of a specified cumulative amount passing through the bed. Calling this amount m, we can write the relation

$$m = \int_0^{t_s} CQ\,dt \qquad (2)$$

where t_s is the service life defined by the cumulative amount m of cyanogen chloride passed through the bed and C is the instantaneous effluent concentration given by solution of equation (1). Performing the integration at constant Q and collecting terms, we have the relation

$$t_s = \frac{W_e \rho}{C_0 k} \ln\left\{1 + \frac{mk}{QW_e\rho} \exp\left(\frac{Wk}{Q\rho}\right)\right\}. \qquad (3)$$

From experimental data on aged and unaged samples, the various quantities are found to have values which permit equation (3) to be well approximated by the form

$$t_s \approx W_e \left(\frac{\rho}{C_o k} \ln\left[\frac{mk}{W_e \rho Q}\right] + \frac{W}{C_o Q} \right). \qquad (4)$$

Details on the derivation of (4) from (3) are given in the Appendix.

The quantity W_e is a direct measure of the number of active sites for removal of cyanogen chloride per unit weight of charcoal. Unequivocal experimental evidence as to the precise chemical nature of these sites is difficult to obtain, so no clear-cut mechanism for their disappearance is available. Aging data indicate that there are a number of different types of sites. There are physical adsorptive sites which do not deteriorate with aging (these are the only sites on unimpregnated charcoal) and a number of types of chemisorptive sites, each type deteriorating with its own characteristic pseudo-first order rate constant. Thus, W_e can be written in the form

$$W_e = W_{ep}^{\circ} + \sum_{j=1}^{N} W_{ej}^{\circ} \exp\left[-k_j t_a\right] \qquad (5)$$

where W_{ep}° is a measure of the (non-deteriorating) physi-sorptive sites, W_{ej}° is a measure of the j^{th} type of chemi-sorptive sites at zero aging time, k_j is the quasi-first order rate constant for deterioration, t_a is the time of aging and N is the total number of types of chemisorptive sites.

Because of the numerical values of the quantities in equation (4), the first term in parenthesis is very small compared to the second. In addition, for our purposes the mild logarithmic functionality of t_s on W_e is negligible compared to the strong linear dependence. By these two observations, we can rewrite equation (4) after insertion of (5) to give

$$t_s = M W_{ep}^{\circ} + \sum_{j=1}^{N} M W_{ej}^{\circ} \exp\left[-k_j t_a\right] \qquad (6)$$

where M is given by definition as

$$M \equiv \left(\frac{\rho}{C_o k} \ln\left[\frac{mk}{W_e \rho Q}\right] + \frac{W}{C_o Q} \right). \qquad (7)$$

As mentioned, for all practical purposes M is virtually independent of aging time and is treated as such in the remainder of this paper.

Analysis of aging data with equation (6) under conditions of several moisture loadings and temperatures yields the following relation for k_j:

$$k_j = \frac{A}{10^{j-1}} [H_2O\%]^2 \exp\left[-E^*/RT\right] \qquad (8)$$

where A is an apparent Arrhenius pre-exponential constant, $[H_2O\%]$ is the weight percentage of water on the charcoal, E^* is an apparent Arrhenius activation energy for the degradation, R is the gas constant and T is the absolute temperature of storage. In this, the word apparent is used because k_j is most likely the product of an equilibrium and rate constant.

Combining equations (5), (6) and (8) and rearranging we have the form

$$r = r_0 + \sum_{j=1}^{N} r_j \exp\left\{-\frac{t_a A [H_2O\%]^2}{10^{j-1}} \exp\left[-E^*/RT\right]\right\} \qquad (9)$$

where r is the ratio t_s (at t_a) / t_s (at $t_a = 0$), r_0 is the ratio MW^o_{ep}/t_s (at $t_a = 0$), and each r_j is the ratio MW^o_{ej}/t_s (at $t_a = 0$). It is this equation we use to predict residual service lives of items stored under specified conditions.

DISCUSSION

Figures (1) and (2) and (3) compare the predictions of equation (9) with actual data on an experimental gas-aerosol filter material containing impregnated charcoals. In the time range of the storage experiment the following values were found for the quantities in equations (9):

$N = 2 \quad r_0 = 0.185 \quad r_1 = 0.300 \quad r_2 = 0.515$

$A = 2.9 \times 10^8 \text{ wks}^{-1} \quad E^* = 16,159.898 \text{ cal mole}^{-1} \text{ deg}^{-1}$

Thus, a predictive equation with fairly good reliability was derived without knowledge of the species involved or the process of degradation. A main use of a predictive equation like this is in accelerated storage studies providing that the more severe conditions do not alter the process of degradation. Comparison of actual storage data with predictions using parameters from accelerated testing confirms that in this case the processes do not change.

APPENDIX

Equation (3) has the form

$$t_s = X \ln\{1 + Y\exp(ZW)\} . \qquad (A1)$$

For large W (large beds) t_s is found to be linear in W. The slope of t_s vs W curves is given by

$$\frac{dt_s}{dW} = \frac{XZYe^{ZW}}{1 + Ye^{ZW}} \qquad (A2)$$

Empirically Ye^{ZW} is much larger than unity for large W so that equation (A2) becomes

$$\frac{dt_s}{dW} \cong XZ . \qquad (A3)$$

Therefore, from (A3) and (A1) using $Ye^{ZW} + 1 \sim Ye^{ZW}$ we have

$$t_s \cong X \ln Y + XZW \qquad (A4)$$

which is a rearrangement of equation (4).

References

(1) NDRC Div 10 Report (146).

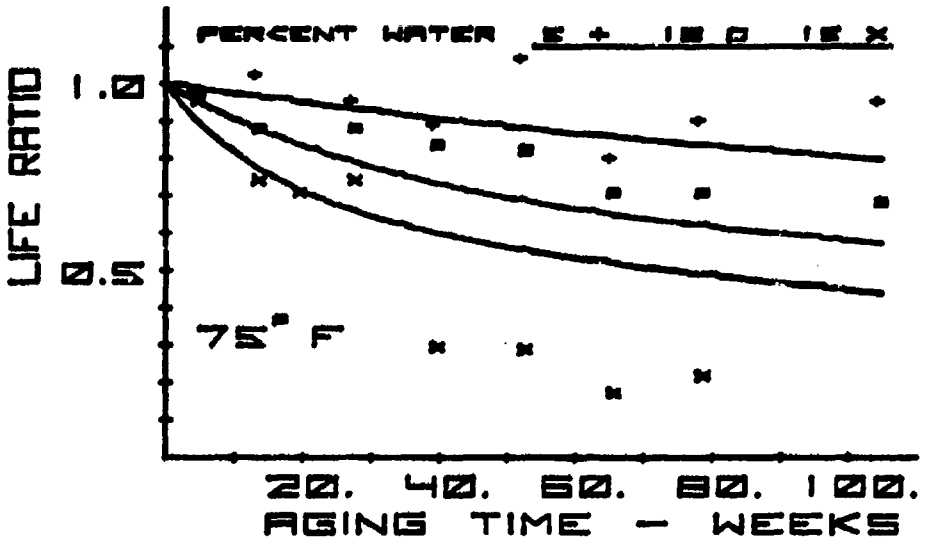
(2) Jonas, L.A., and Rehrmann, J., Carbon 10, 657 (1972).

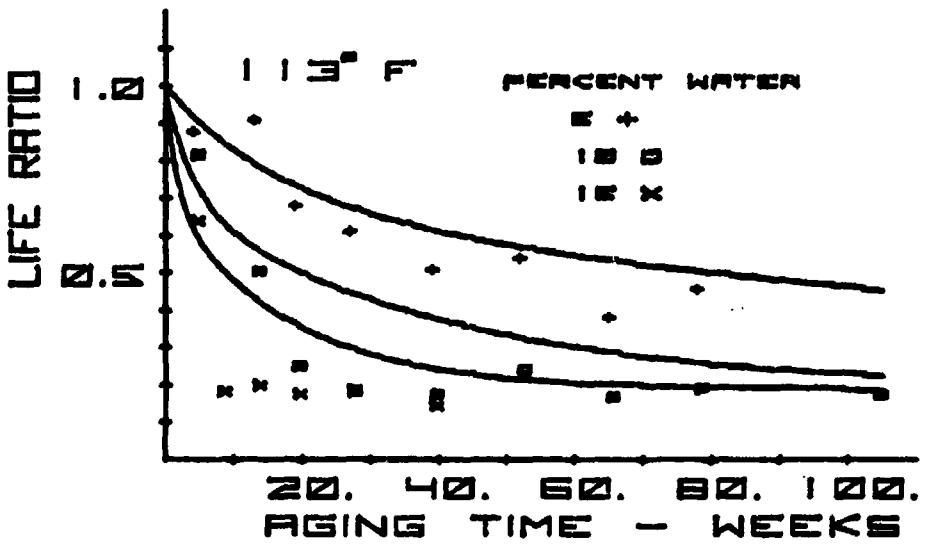
(3) Wheeler, A. and Robell, A., J. Catal. 13, 299 (1969).

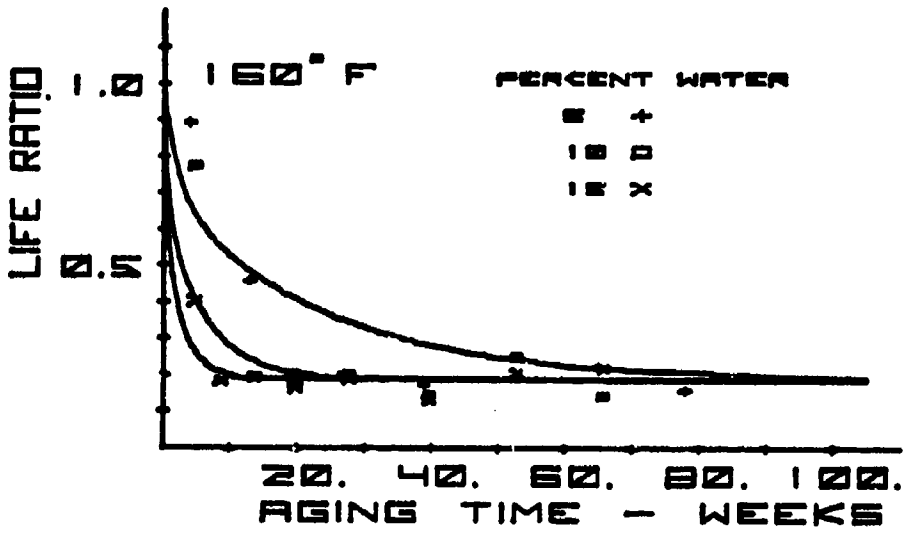
(4) Dole, M. and Klotz, I., Ind & Eng Chem. 38, 1289 (1946).

(5) Hutchins, R., Chem. Eng., p. 133, Aug 20, 1973.

(6) Magee, W. S. (to be published).


PERCENT WATER 5 + 10 □ 15 X
LIFE RATIO
1.0
0.5
75° F
20. 40. 60. 80. 100.
AGING TIME — WEEKS


113° F
PERCENT WATER
5 +
10 □
15 X
LIFE RATIO
1.0
0.5
20. 40. 60. 80. 100.
AGING TIME — WEEKS



DATA UNCORRECTED FOR ACTUAL WATER AND CHARCOAL LOADINGS. SOLID LINES ARE THEORETICAL, NOT BEST FITS TO THE DATA.

A GAUSSIAN TOLERANCE TECHNIQUE FOR MULTILAYER OPTICAL COATINGS

J. J. Walls, Jr., R. A. McKyton, and A. Kwalec
Fire Control Development & Engineering Directorate, Optical Branch
Frankford Arsenal, Philadelphia, Penna.

ABSTRACT. Discussed is a computer program that mathematically simulates production manufacturing tolerances for multilayer optical coatings. This technique operates on an ideal theoretical multilayer design by assigning random Gaussian film thickness errors for each layer within the film stack. "N" iterations of this process are performed to obtain a realistic effect of the film thickness variations on the spectral reflection and transmission characteristics of the coating. Thus this program enables the identification of those optical coating designs that have sensitive thickness requirements and also allows for the generation of practical thickness tolerances for the many optical coatings used in fire control instrumentation.

1. INTRODUCTION. Frankford Arsenal is engaged in the research, development, and manufacture of fire control systems. Items, such as optical sights, rangefinders, telescopes, and other optical instruments form the fire control support for tanks, helicopters, mortars, and many other weapons. This paper will describe a technique for predicting realistic tolerances for the numerous types of multilayer optical coatings used in fire control systems.

Through use of large scale computers both analysis and synthesis design techniques can be used to generate specific ideal coating designs for any given spectral input requirement; that is, given a spectral reflection or transmission specification - theoretical design parameters, such as the number of layers, their order, film thicknesses, and associated film refractive indices can be computed for the multilayer coating to ideally fit the spectral input specification. The problem that arises is that this ideal computed design is errorless - that is, it does not include the effects of the fabricational tolerances that are required to produce the design. Hence, before any realistic spectral performance of this design can be obtained it is necessary to include the effects of these fabricational errors in order to predict the practical spectral performance of the coating.

In order to form a basis for the understanding of the specifics of this problem it would be appropriate at this time to briefly describe what a multilayer optical coating is and how it is fabricated. A multilayer optical coating can be described as a combination of thin films that are deposited in a given order to predetermined thicknesses (t_k) and refractive indices (n_k) - see figure 1. The thicknesses of these films are thin enough to cause interference effects, and thus change the spectral properties of the coated component. The spectral performance of this coating is completely specified once the order of the films is specified and the individual thicknesses and refractive indices of each film within the stack are given. Changing any of these design parameters can change the spectral performance of the coating. The absorption of the film's within the multilayer stack is assumed zero. This is a valid assumption for most of the films used within a given spectral range. These multilayer coatings are fabricated by a number of techniques - most employ the vacuum vaporization technique. That is, the items to be coated will be placed within a chamber under vacuum. The materials to be deposited will be heated to their corresponding vaporization temperatures and the vapors then condensed onto the substrates under vacuum. Since each film within the multilayer stack must be deposited to a given thickness ($\sim 10^{-6}$in) - an optical thickness monitoring technique shown in figure 2 is the conventional method used to accomplish this task. Basically, the reflectance or transmittance of each film is monitored during film deposition and this optical parameter is related to the optical thickness of each film. In addition to thickness monitoring errors, other major factors such as evaporant distribution and substrate temperature contribute to the total error of the system.

With that discussion as a background - let us now return to the basic question: How do we simulate these system thickness errors in order to predict the realistic spectral performance of a multilayer coating? The answer to this question is the use of a computer program that statistically simulates system thickness errors and computes their complicated effects on the spectral performance of the multilayer optical coating.

2. PROGRAM LOGIC, INPUT, OUTPUT. The basis for the simulation of manufacturing tolerances is the generation of expected thickness errors and the analysis of their effects on the spectral properties of an ideal multilayer design. In order to determine the distribution of thickness errors which would be valid for the coating process a statistical analysis was performed on a set of 80 individual film thicknesses (40 points for ZnS films and 40 for MgF_2 films.) Analysis of the data showed

a slightly skewed distribution for both materials. However, due to the relatively small sample size from only one coating facility and because of the need to have the simulation of thickness errors as general as possible, it was decided to employ the Gaussian distribution as a first effort.

In performing the spectral computations for the coatings each film of the multilayer is assigned a matrix (M_k) based on its optical and physical properties.

$$M_k = \begin{bmatrix} \cos \delta_k & j\, n_k^{-1} \sin \delta_k \\ j\, n_k \sin \delta_k & \cos \delta_k \end{bmatrix} \qquad (1)$$

where δ_k = phase retardation of the k^{th} layer.

n_k = refractive index of the k^{th} layer.

A resultant product matrix (M) defined

$$M = M_1 * M_2 * \ldots\ldots\ldots * M_{1-m} * M_m = \begin{bmatrix} A & jB \\ jC & D \end{bmatrix} \qquad (2)$$

is computed at each wavelength of interest. This final matrix is then used to define the reflection or transmission of the initial design.

$$R = (1-T) = \frac{(X - U)^2 + (Y - V)^2}{(X + U)^2 + (Y + V)^2} \qquad (3)$$

where

$$\begin{aligned} X &= NoA & U &= N_s D \\ Y &= NoN_s B & V &= C \end{aligned} \qquad (4)$$

Figure 3 shows the basic computational operation performed for simulating the effects of film thickness tolerance on the spectral performance of a coating. The first operation is performed by taking the input design parameters (order, film thicknesses and film refractive indices) of the ideal design and computing the spectral reflection or transmission of this coating (M = 1 operation). When M = 2 or more the program creates new designs with induced errors by the following procedure: A three-digit integer or seed is inputed into a CDC library function, RANF, which creates

random values (X_1) from a uniform distribution in the interval (0,1). By virtue of the Central limit theorem, summing over every seven values returned from RANF results in a distribution of sums (A) which approaches a normal distribution as the number of samples increases. These values are used to generate a thickness error which is Gaussian such that 95% of these generated errors will be within the maximum tolerance or ERROR which the designer inputs. The new film thickness is created from the following equation:

$$t_k' = t_k + (A - 3.5)\ \frac{(E \cdot t_k)}{1.96} \qquad (5)$$

where

$$A = \sum_{i=1}^{7} X_1 \qquad \text{(generates normal distribution)}$$

t_k' = new film thickness (induced error)

t_k = original film thickness

E = Error (Thickness Tolerance)

Each layer of the initial design is successively modified by this induced error in thickness. The new multilayer design is then analyzed by the matrix technique and its spectral properties are computed. This error iteration process is performed as many times as necessary to determine the maximum changes in R or T as a function of a given thickness tolerance or ERROR.

Initial design input required for the tolerance program is shown in figure number 4. It includes angle of incidence, index of substrate and medium, spectral range of interest, a seed for RANF function and the index and physical thickness of each layer of the ideal film design. Also required as an input is the maximum allowed thickness change or ERROR.

The program has numerous output options also shown in Figure (4) which the designer can call upon to assist in the best presentation of data. The program permits any number of error designs to be generated. Each design's film index, thickness, percent thickness error may be outputed. However, when performing many iterations on an ideal design such information becomes excessive. The program allows for the deletion of this data unless imperative for determining which design created the largest change in R or T. For critical film systems the exact values of R & T vs λ for each design can be listed but for most investigations a plot of R or T versus wavelength is sufficiently informative to determine

maximum allowable tolerances. The plot routine of this program is limited to plotting eleven individual R or T curves, since it would be impossible to resolve anymore due to printing constraints. Since most film system require more than ten iterations to determine the greatest changes in spectral performance the program can present maximum and minimum valued R or T envelopes from any specified number of iterations. In addition, the output always includes a table giving the thickness error distribution within the specified tolerance range.

In order to establish that the spectral computation part of this program includes those critical parameters that completely define the spectral performance of any coating in the real world it was necessary to experimentally verify this point. To substantiate this, an eight layer design using quarter-wave films of zinc sulfide and magnesium fluoride designed for 45 degree incidence was chosen as a test design. The vacuum vaporization technique was used to deposit this multilayer stack. In addition to the multilayer coating deposit a set of the individual films making up the multilayer coating was obtained during the same coating run. The thicknesses of each film within the stack were measured and this along with the corresponding refractive indices was inputed into the thin film computer analysis program. The multilayer coating was measured with a Cary 14 spectrophotometer. A comparison of both the computational and experimental results obtained is shown in figure 5. As indicated, good agreement was obtained. The slight difference in the blue spectral region is attributed to film absorption.

3. ANALYSIS OF M-16 REFLEX SIGHT BEAMSPLITTER COATING. To determine the validity and usefulness of the tolerance program an eight-layer dichroic type beamsplitter coating that is presently being used in the prototype reflex sight for the M-16 rifle was analyzed. The design and spectral characteristics of this multilayer coating are shown in figure 6. The coating is designed to reflect the red spectral region and transmit the blue-green spectrum for a 45° degree incident angle.

To establish the practical validity of the tolerance program it is necessary to compare the spectral data from several controlled experimental evaporation runs with the spectral tolerance data predicted by the tolerance program for this specific coating design. With this end in mind ten separate evaporation runs were performed to obtain ten multilayer coatings and 10 sets of individual films for the multilayer coatings. Shown in figure 7 are a few spectral transmission curves obtained and also the maximum and minimum valued spectral envelopes formed from these ten coatings. The thicknesses of all the individual

films for these 10 multilayer coatings were determined by a spectral technique. The errors were then determined with respect to the ideal film thickness value. It was determined that 95% of film thickness errors fell within ±10%. This error of 10% and the ideal design parameters were inputed into the tolerance program. Iterations of 100, 600, and 1200 were performed in order to determine the maximum change in transmission for this thickness error. After 600 iterations the transmission envelope did not expand any further for the 10% thickness error. This indicated that the maximum expected spectral transmission tolerance envelope has been achieved for this given tolerance. Shown in figure 8 are the spectral tolerance envelopes generated for 100 error designs and 600 error designs for the 10% tolerance. Figure 9 shows that the envelope generated by the tolerance program encompasses all of the experimentally generated envelope values for the 10% tolerance.

If the film thickness tolerances were made tighter one would expect that the spectral tolerance band would become tighter. To demonstrate this effect, the input tolerance was changed to 5% and iterated until the maximum spectral tolerance band was obtained. Shown in figure 10 is a comparison of a 10% and 5% band. Not only is the 5% band significantly tighter, but its shape approaches that of the spectral transmission curve of the ideal design. This demonstrates that for this specific design a 5% tolerance would be preferred because color and brightness characteristics of this filter would be assured.

4. CONCLUSIONS. Based on the theoretical and experimental results presented, the following conclusions can be drawn:

a. If a spectral computational program is defined such that it generates realistic spectral output for a given set of design parameters and if an error distribution that describes the system fabricational errors is interacted with the spectral program for the essential number of iterations then the generated tolerance envelope will be a real and practical representation of the spectral tolerance for the multilayer coating.

b. As indicated from the experimental results the Gaussian error distribution is a good approximation for system fabricational errors.

c. One would expect as the film thickness tolerance is made tighter, the spectral tolerance envelope would become tighter and its shape would approach that of the ideal design - this is exactly what the tolerance program indicated.

5. SUMMARY. Some of the advantages and use of this program include:

a. The prediction of critical multilayer components. That is, for small film thickness errors, large spectral changes are observed.

b. It affords the optical designed a practical tool for assigning realistic spectral tolerances for optical components.

c. The government will have its own means of determining production tolerances independent of industry.

d. Several design attempts can be made for a given spectral requirement and a determination made on which design is least sensitive to error. This would definitely have cost implications.

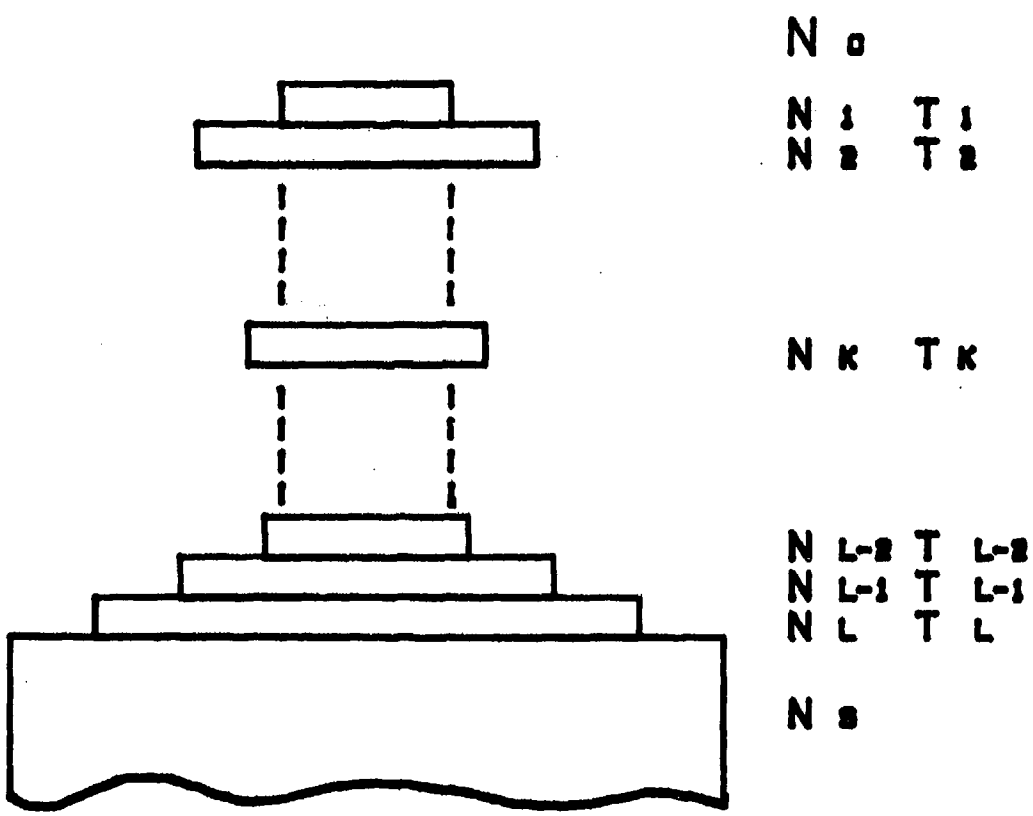
In summary, whereas before it has been impossible for the thin film coating designer to predict the complicated error interaction effects on the spectral performance of an optical coating, the statistical approach presented provides the designer with a method that will give him a high degree of confidence in predicting production tolerance effects.

Bibliography

Hicks, C.R. Fundamental Concepts in the Design of Experiments. New York: Holt, Rinehart and Winston, Inc., 1964.

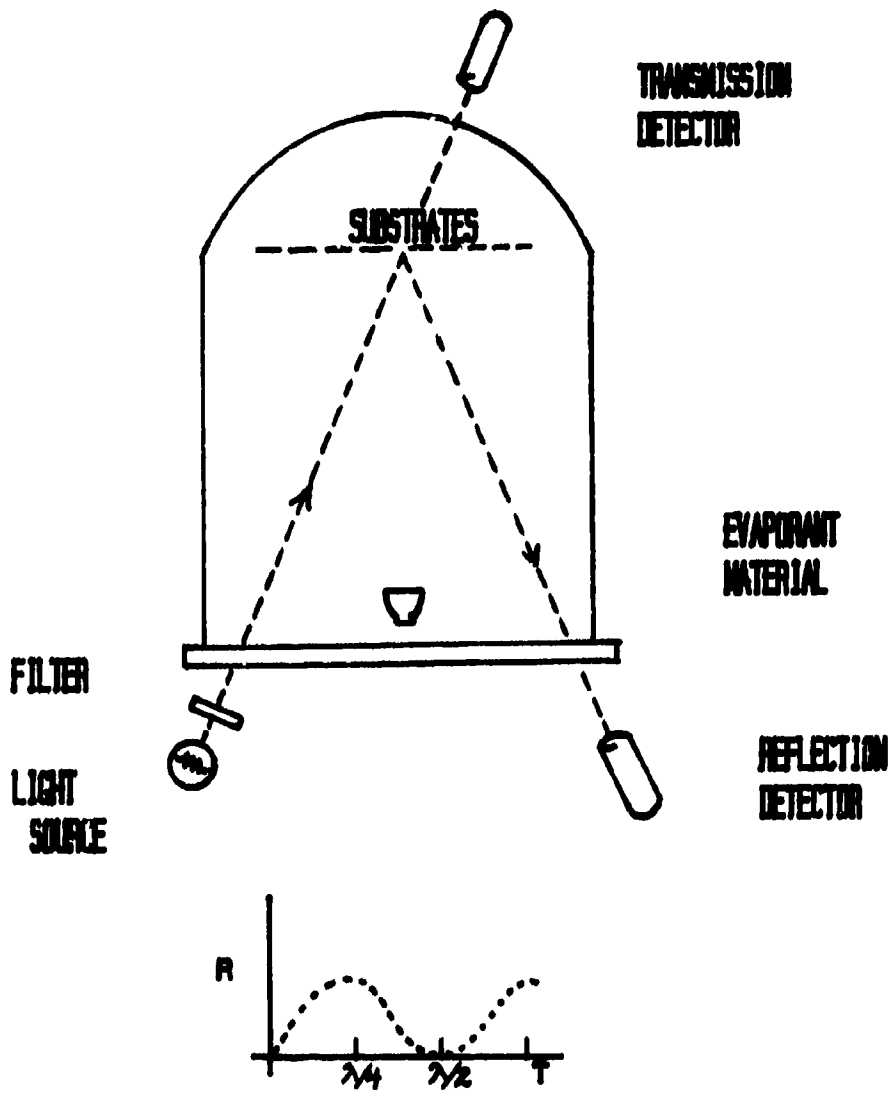
Hass, G. Physics of Thin Films. Vol. 1 New York, Academic Press, 1963.

Kingslake, R. Applied Optics and Optical Engineering, Vol. 1. New York, Academic Press, 1965.



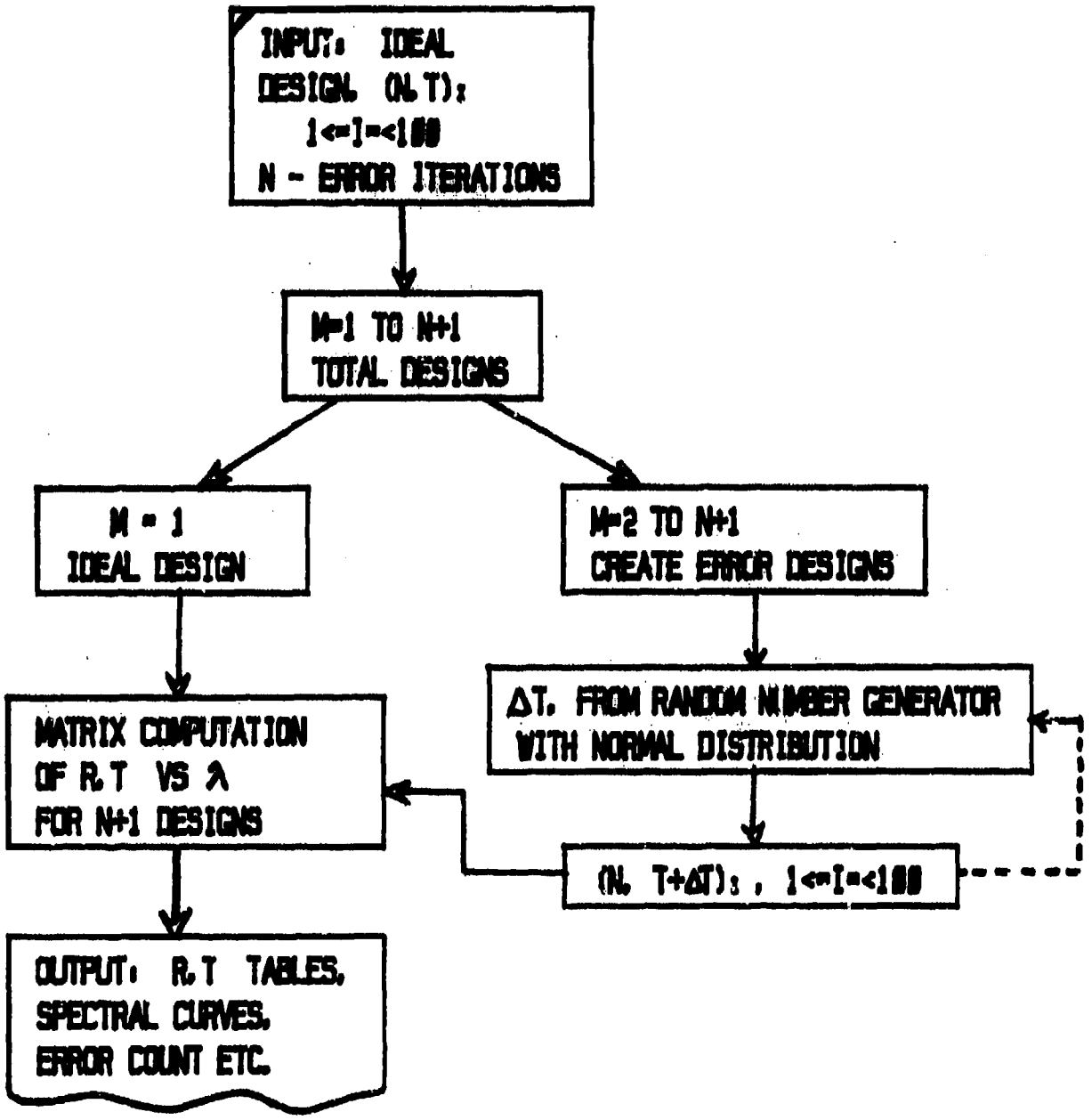
MULTILAYER OPTICAL COATING

FIGURE NO. 1



FILM THICKNESS MONITOR

FIGURE NO. 2



PROGRAM LOGIC

FIGURE NO. 3

INPUT

Φ - ANGLE OF INCIDENCE

N_S - SUBSTRATE INDEX

N_o - MEDIUM INDEX

$\lambda_i, \lambda_f, \Delta\lambda$ - INITIAL. FINAL. INCREMENT WAVELENGTHS

N - NUMBER OF ERROR ITERATIONS

SEED - 3 DIGIT INTEGER

N_i . T_i - INDEX. PHYSICAL THICKNESS FOR EACH LAYER

ERROR - MAXIMUM THICKNESS TOLERANCE

OUTPUT OPTIONS

1. LIST ALL DESIGNS OR ONLY IDEAL
2. LIST R & T VALUES OR PRESENT SPECTRAL CURVES
3. PLOT R OR T CURVES
4. PLOT 11 INDIVIDUAL CURVES OR MAX/MIN ENVELOPES FOR N - ITERATIONS

FIGURE NO. 4

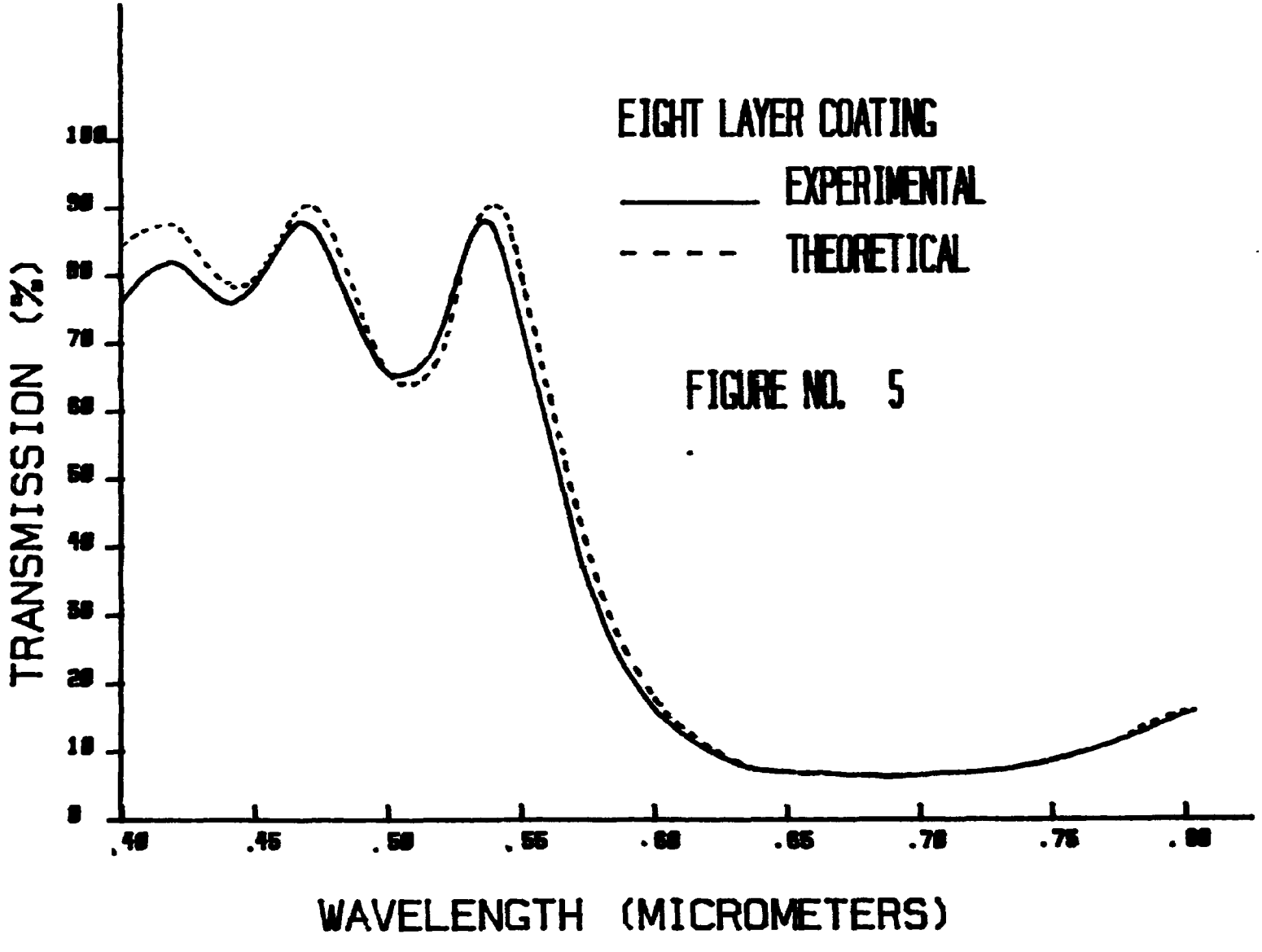

EIGHT LAYER COATING
EXPERIMENTAL
THEORETICAL
TRANSMISSION (%)
WAVELENGTH (MICROMETERS)

FIGURE NO. 5

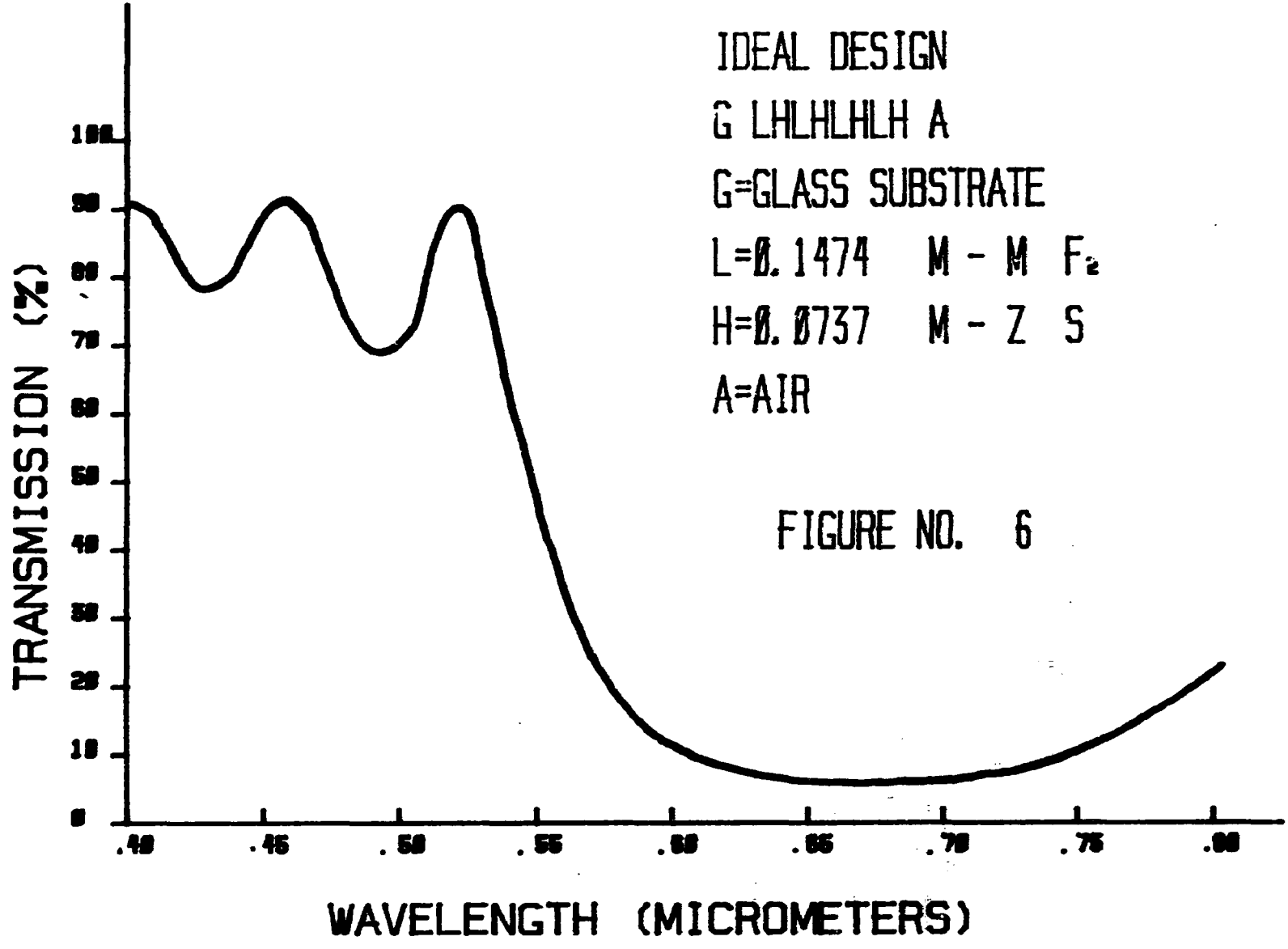

IDEAL DESIGN
G LHLHLHLH A
G=GLASS SUBSTRATE
L=Ø.1474 M - M F₂
H=Ø.Ø737 M - Z S
A=AIR
TRANSMISSION (%)
WAVELENGTH (MICROMETERS)

FIGURE NO. 6

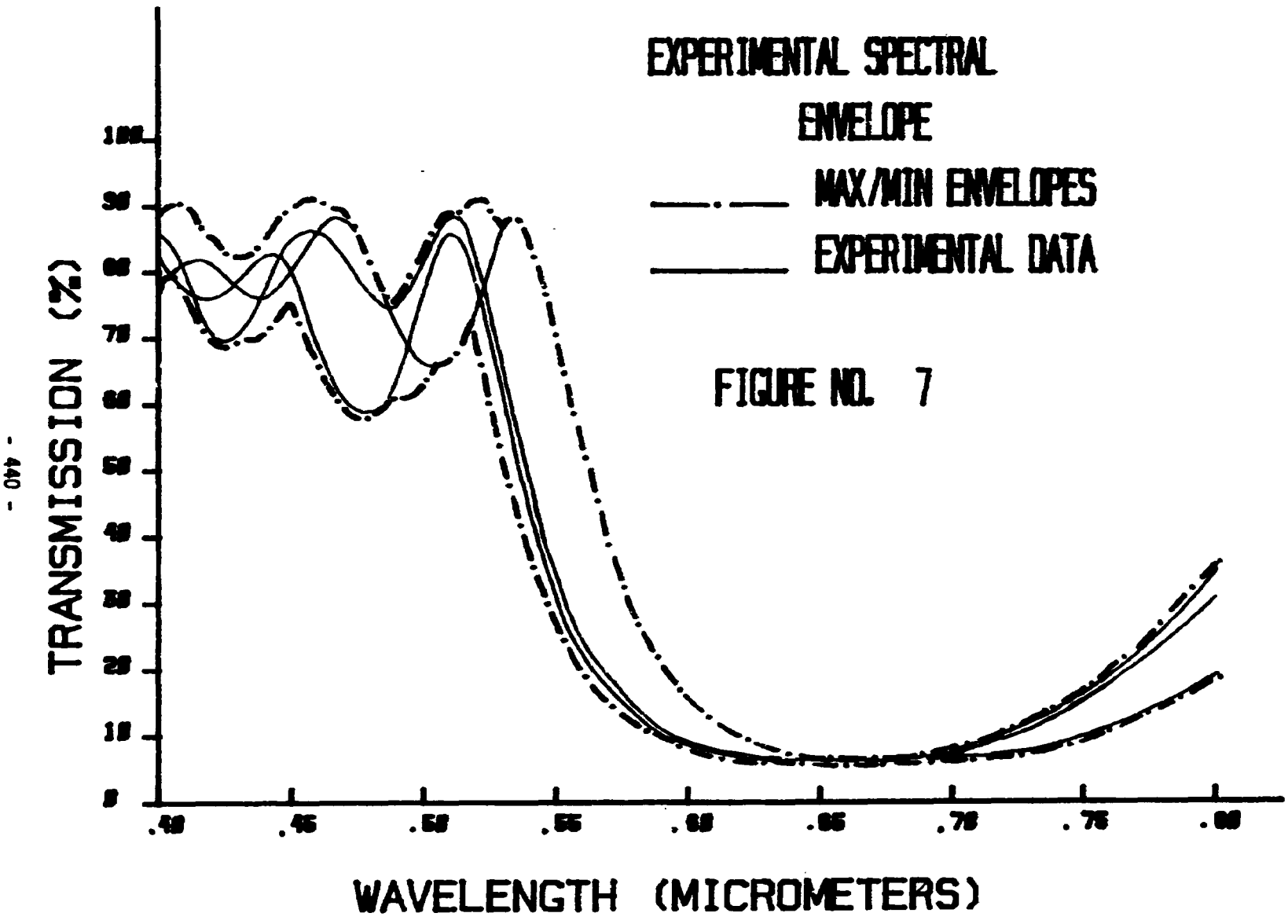

EXPERIMENTAL SPECTRAL ENVELOPE
MAX/MIN ENVELOPES
EXPERIMENTAL DATA
TRANSMISSION (%)
WAVELENGTH (MICROMETERS)

FIGURE NO. 7

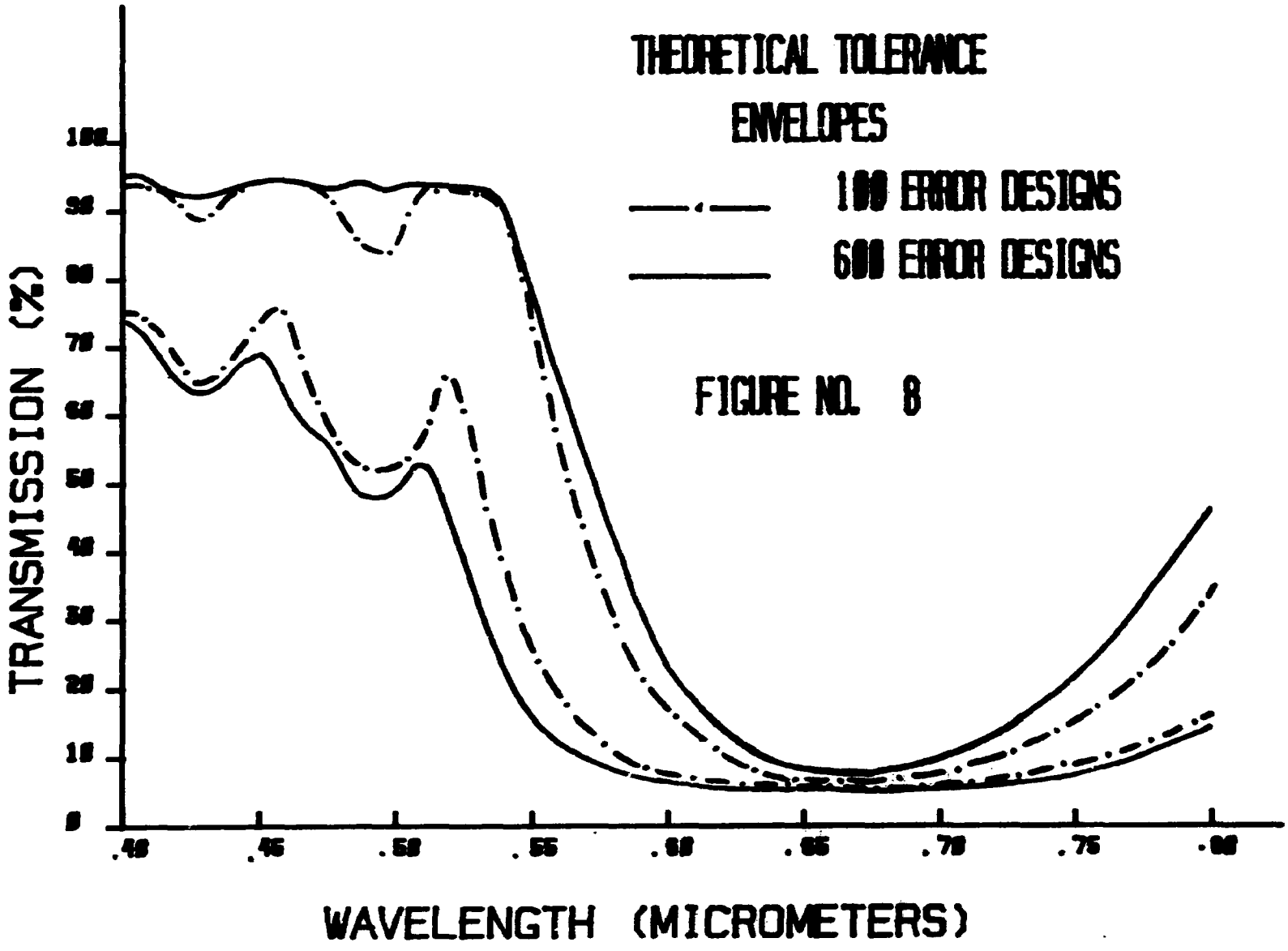

THEORETICAL TOLERANCE
ENVELOPES
100 ERROR DESIGNS
600 ERROR DESIGNS
TRANSMISSION (%)
WAVELENGTH (MICROMETERS)

FIGURE NO. 8

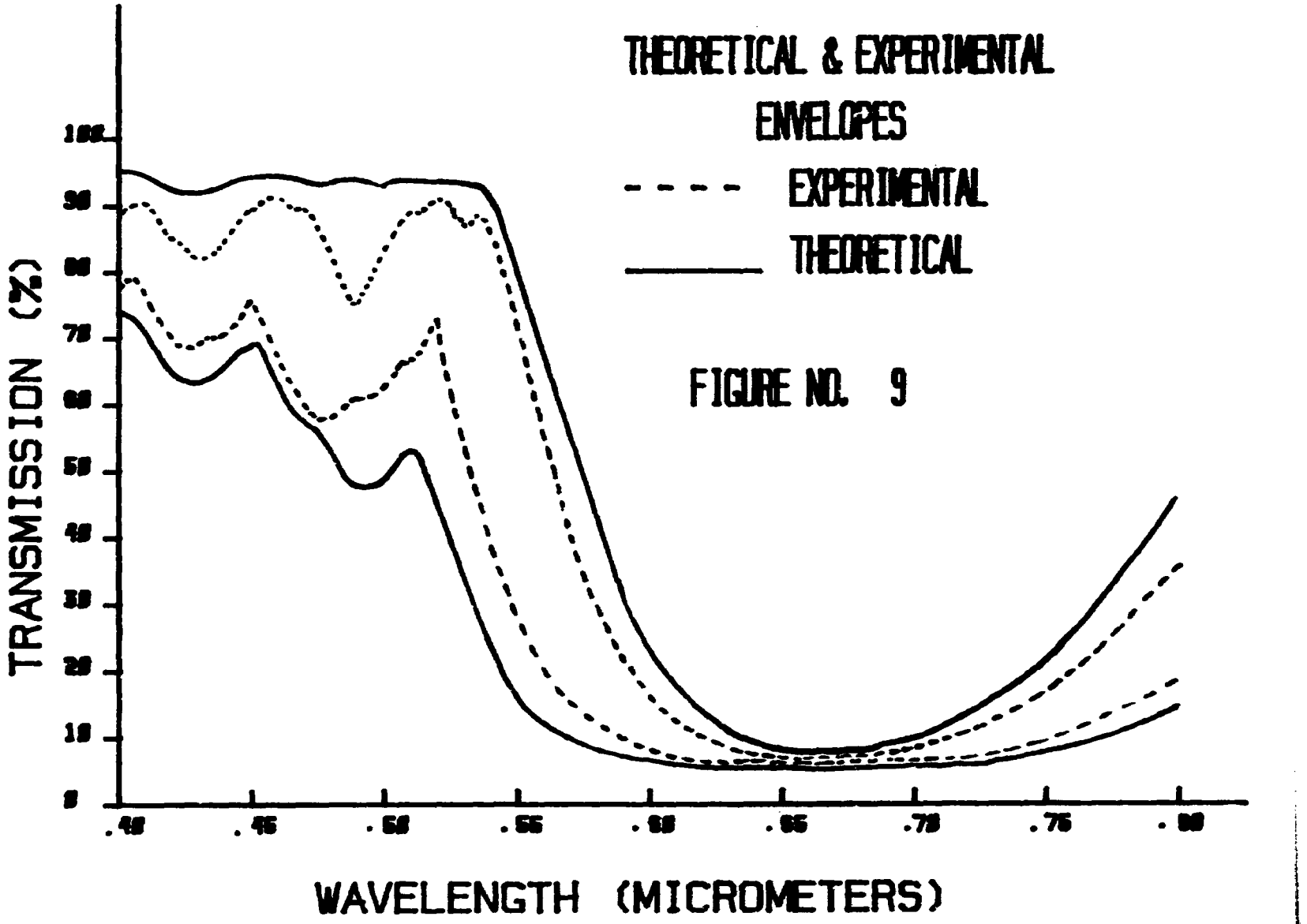

THEORETICAL & EXPERIMENTAL
ENVELOPES
EXPERIMENTAL
THEORETICAL
TRANSMISSION (%)
WAVELENGTH (MICROMETERS)

FIGURE NO. 9

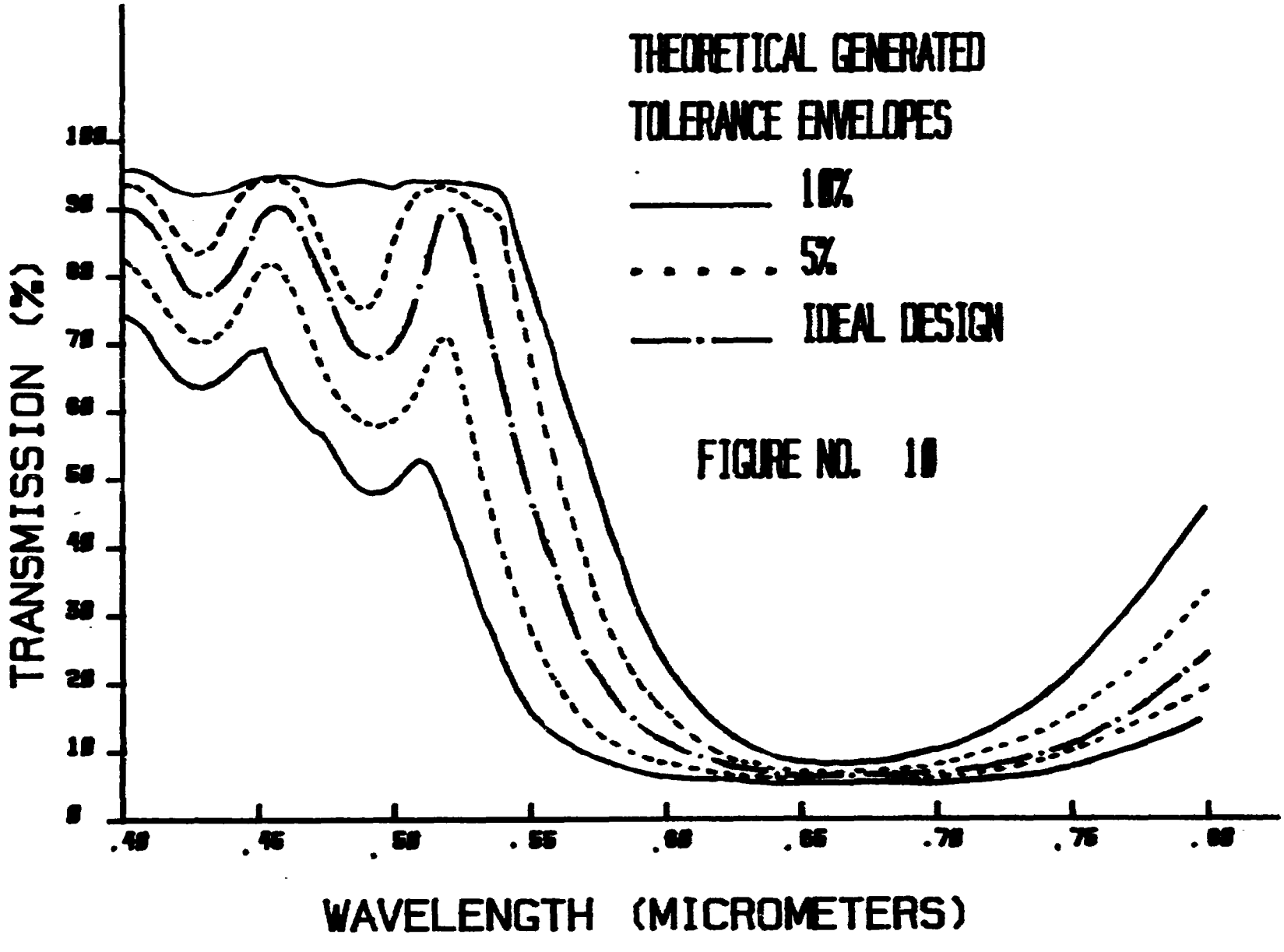

THEORETICAL GENERATED
TOLERANCE ENVELOPES
10%
5%
IDEAL DESIGN
TRANSMISSION (%)
WAVELENGTH (MICROMETERS)

FIGURE NO. 10

A TABLE OF CUMULATIVE DISTRIBUTION FUNCTION VALUES OF EXPECTED VALUES OF NORMAL ORDER STATISTICS

Gary L. Aasheim
Product Assurance Directorate
US Army Armament Command
Rock Island, Illinois

ABSTRACT

This article presents a table of cumulative distribution function values of expected values of normal order statistics. An example is given in which the table is used to test a population for normality and a brief discussion is made of the method used to construct the table. The latter part of the paper deals with origins of commonly used expressions for approximating values given in the table.

1.0 INTRODUCTION

In this paper we present a table (Table 2) of cumulative distribution function values of expected values of normal order statistics, an extension of H. L. Harter's table of expected values of normal order statistics (6.2). The project was undertaken at the suggestion of Mr. Richard Brugger and as a result of the US Army's participation in the Quadripartite Standardization Program and the Army Ammunition Procurement and Supply Agency's role at that time as the Army's representative in a project which deals with sampling plans for inspection by variables. As Army representatives on the project, we reviewed a proposed British Defence Standard that deals with that same type of sampling plan. The British publication presents an abridged version of Harter's table as an aid in probability plotting. The table presented here complements the abridged British version of Harter's work rather than Harter's full table and is intended as a further aid to probability plotting.

2.0 AN APPLICATION

The problem of determining whether it is reasonable to assume that a particular population follows a normal distribution can be dealt with by any one of several methods. Probability plotting is one of the most easily used of these methods and can be applied in either one of two equivalent forms.

One form requires the use of ordinary graph paper on which are plotted the ranked data points, x_i, versus the expected value of the corresponding normal order statistics, $E(x_i)$. Harter's table of expected values of normal order statistics can be used in this form of the test.

Table 2 can be used to implement the second form which requires the use of normal probability paper. The ranked data points, x_i, of a sample of size n are plotted against $\Phi[E(x_i)]$, where Φ is the standardized normal cumulative distribution function, or against some approximation to $\Phi[E(x_i)]$ such as $(i - 1/2)/n$. The method of this second form is now detailed followed by an example in which the method is illustrated with the use of Table 2.

2.1 PROBABILITY PLOTTING

The method of probability plotting using probability paper is as follows:

1. Rank the data in order of magnitude from smallest to largest. A set of size n of ordered observations is thus obtained: $x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n$.

2. From Table 2, obtain $\Phi[E(x_i)]$.

3. Plot the ranked data, x_i, vs. $\Phi[E(x_i)]$.

4. Draw the straight line which seems to best fit the data points.

5. If the points appear to cluster around the straight line, proceed on the assumption that the population is normally distributed.

2.2 AN EXAMPLE

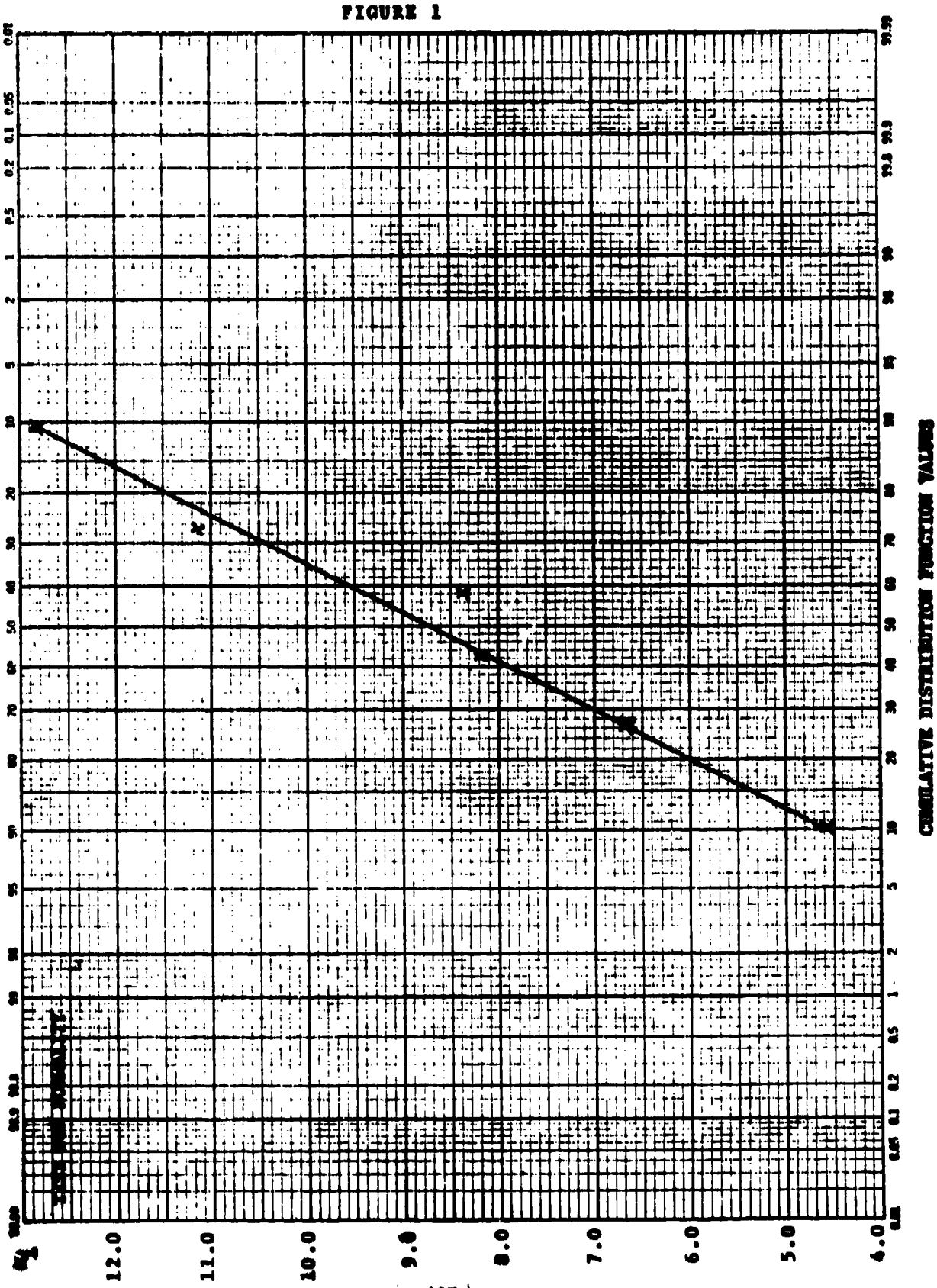
Consider the following simulated sample of size n = 6 from some population: 8.36, 12.81, 4.58, 6.66, 11.14, 8.19. If we rank this data in Table 1 and immediately below it place values of $\Phi[E(x_i)]$ obtained from Table 2 for a sample size 6, we get:

TABLE 1

i	1	2	3	4	5	6
x	4.58	6.66	8.19	8.36	11.14	12.81
$\Phi[E(x_i)]$	.10254	.26051	.42013	.57987	.73049	.89746

The probability plot of this data, Figure 1, shows the data points clustered closely enough around the fitted straight line that the user can reasonably conclude that he is dealing with a normally distributed population.

FIGURE 1



CUMULATIVE DISTRIBUTION FUNCTION VALUES

3.0 CONSTRUCTION OF THE TABLE

The data for Table 2 was taken from (6.3) and (6.4). Interpolation was required to obtain Table 2 from tables in (6.4). The interpolation formula (6.3) is based upon the first 2 terms of the Taylor series expansion of the normal cumulative distribution function. This interpolation gives results which are accurate to at least 7 decimal places, more than adequate for the 5 decimal place accuracy of Table 2.

The interpolation formula:

$$\phi[E(x_i)] = [2[E(x_i)-x_{i0}]Q_0 + P_0] / 2 + .5 ,$$

$i = 1, 2, \ldots n/2$ (n even)

$i = 1, 2, \ldots (n+1)/2$ (n odd)

$$\phi[E(x_i)] = 1 - \phi[E(x_{n-i+1})],$$

$i = [n/2]+1, [n/2]+2, \ldots n$ (n even)

$i = [(n+1)/2]+1, \ldots n$ (n odd)

where:

$E(x_i)$ is the expected value of the ith normal order statistic,

x_{i0} is the tabulated normal deviate nearest to $E(x_i)$,

$$Q_0 = Q(x_{i0}) = (1/\sqrt{2\pi}) \exp(-x_{i0}^2/2)$$

$$P_0 = P(x_{i0}) = \int_{-x_{i0}}^{x_{i0}} Q_0 \, dt .$$

Q_0 and P_0 are tabulated in (6.3).

4.0 BACKGROUND

Blom (6.1) gives the development of $(i - 1/2)/n$ and $i/(n + 1)$ as estimates of $\phi[E(x_i)]$. We will make a short review here of certain of his results. Consider n random points of the form u_{in} where $0 \leq u_{1n} \leq u_{2n} \leq \ldots \leq u_{nn} \leq 1$. u_{in} is the ith order statistic of a sample of size n from a uniform distribution. Let $x = G(u)$ be a Borel-measurable function defined over the range of u with $G^{-1}(x) = F(x)$. The random variable u_{in} follows a beta distribution, and, hence, the $x_{in} = G(u_{in})$ are known as transformed beta variables.

The problem considered in Blom's text which is of concern here is that of determining the expected value of transformed beta (TRB) variables. For the case in which the TRB-variable, x_{in}, follows a symmetric distribution, such as the normal, the expression for the expected value is:

$$(1) \qquad E(x_{in}) = G[(i - \alpha_i)/(n - 2\alpha_i + 1)] + R_i$$

where α_{in} is a correction factor which improves an earlier expression derived for $E(x_i)$ and which has a unique value for each pair (i,n). For the remainder term, R_i, we have $\lim_{n\to\infty} nR_i$ finitely bounded. A good approximation of $E(x_i)$ can then be had by replacing α_{in} with a single properly selected α for all (i,n) and by dropping the remainder term. We then have

$$(2) \qquad E(x_i) \doteq G[(i - \alpha)/(n - 2\alpha + 1)]$$

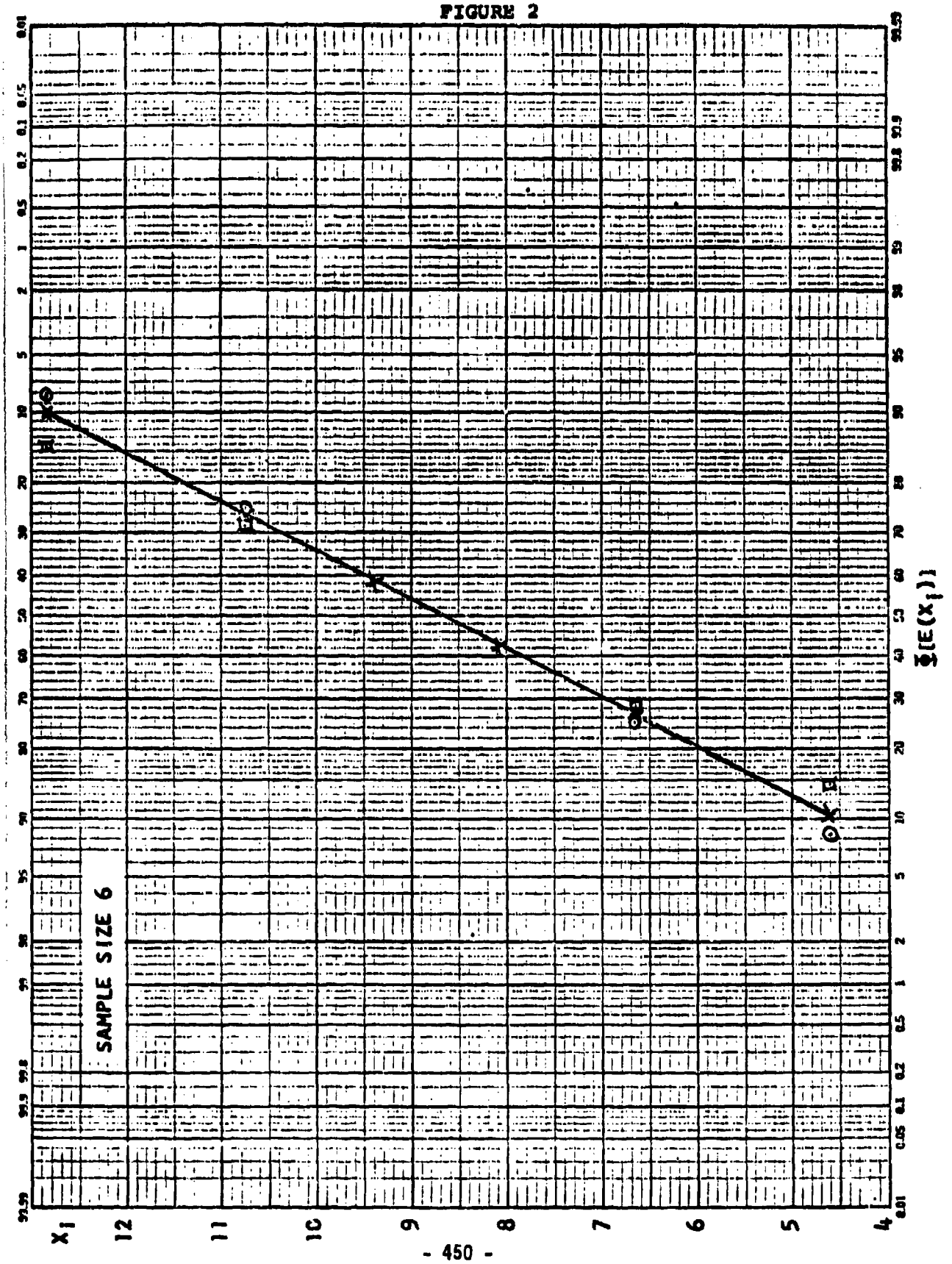
Consider the case of special interest in which $G = \phi^{-1}$. Assigning values to α gives the following expressions for $G^{-1}[E(x_i)] = \phi[E(x_i)]$ from (2):

α	$\phi[E(x_i)]$
0	$i/(n + 1)$
1/2	$(i - 1/2)/n$
3/8	$(i - 3/8)/(n + 1/4)$

The first two expressions give good approximations to $\phi[E(x_i)$ and are the ones most commonly used in probability plotting. Blom suggests that using the third expression above with $\alpha = 3/8$ gives a still better approximation to $\phi[E(x_i)]$ and is a good compromise of all α values for approximating $\phi[E(x_i)]$. However, this third expression is relatively untidy and only slightly better approximates $\phi[E(x_i)]$ than do the first two expressions. Probably for these reasons it is little used in probability plotting.

Figure 2 illustrates the deviations of $i/(n + 1)$ and $(i - 1/2)/n$ from $\phi[E(x_i)]$ for a sample of size 6. The comparison is made by first constructing a set of data that will fall on a straight line when plotted against $\phi[E(x_i)]$. The same data set is then plotted against $(i - 1/2)/n$ and then $i/(n + 1)$. In the figure, the points circled to the lower left and upper right are for $(i - 1/2)/n$; those enclosed by squares to the upper left and lower right are for $i/(n + 1)$. The deviation of the two approximation lines from the ϕ line is greatest at the end points and decreases toward the middle of the plot where the approximation curves cross the ϕ curve. Where the abscissa values are so close toward the middle, the approximation points have not been indicated separately.

FIGURE 2



Numerically, the deviations grow smaller with increased sample size. Due to the nature of the changing scale of the abscissa on the probability paper, however, the decreasing deviations are not apparent on a plot. For a given n, points on the $i/(n + 1)$ curve deviate more from the corresponding points on the ϕ curve than do points on the $(i - 1/2)/n$ line. A straight edge lying on the end points will show that the points for each of the two approximations lie nearly on a straight line, but are actually bent very slightly into an "S" shape, the bend being more pronounced for the $i/(n + 1)$ points than for $(i - 1/2)/n$ points. Since a test for normality using a probability plot is generally a strictly visual test, it appears doubtful that the relatively small errors introduced using either of the approximations given here would influence a test conclusion.

5.0 CONCLUSION

We undertook the project in order to provide a convenient method for plotting on normal probability paper which would be equivalent to methods for plotting on standard graph paper using Harter's table of expected values of normal order statistics. The table does not, as we originally thought it might, provide for significantly improved probability plotting with respect to improved abscissa values for the plot.

6.0 REFERENCES

6.1 Blom, Gunnar (1958), "Statistical Estimates and Transformed Beta Variables", John Wiley and Sons, New York.

6.2 Harter, H. L. (1961), "Expected Values of Normal Order Statistics, Biometrika 48, 151-165.

6.3 National Bureau of Standards (1953), "Tables of Normal Probability Functions", Applied Mathematics Series 23, US Government Printing Office, Washington, D. C.

6.4 SS Branch (1972), MP 191/1, "Sampling Procedures and Charts for Inspection by Variables", Appendix II, Quality Assurance Directorate, (Weapons), Woolwich, England SE18.

TABLE 2

Cumulative Distribution Function Values of Expected Values of Normal Order Statistics

k \ n	2	3	4	5	6	7	8
1	.28631	.19870	.15165	.12242	.10254	.08816	.07728
2	.71369	.50000	.38323	.31029	.26051	.22441	.19705
3		.80130	.61677	.50000	.42013	.36215	.31817
4			.84835	.68971	.57987	.50000	.43939
5				.87758	.73949	.63785	.56061
6					.89746	.77559	.68183
7						.91184	.80295
8							.92272

k \ n	9	10	11	12	13	14	15
1	.06877	.06193	.05632	.05163	.04766	.04425	.04129
2	.17559	.15833	.14414	.13227	.12220	.11354	.10603
3	.28367	.25589	.23305	.21394	.19771	.18376	.17164
4	.39568	.35355	.32205	.29569	.27330	.25406	.23734
5	.50000	.45118	.41103	.37743	.34889	.32435	.30303
6	.60432	.54882	.50000	.45914	.42445	.39462	.36870
7	.71633	.64645	.58897	.54086	.50000	.46487	.43435
8	.82441	.74411	.67795	.62257	.57555	.53513	.50000
9	.93123	.84167	.76695	.70431	.65111	.60538	.56565
10		.93807	.85586	.78606	.72670	.67565	.63130
11			.94368	.86773	.80229	.74594	.69697
12				.94837	.87780	.81624	.76266
13					.95234	.88646	.82836
14						.95575	.89397
15							.95871

TABLE 2 *Continued*

k \ n	16	17	18	19	20	21	22
1	.03878	.03641	.03438	.03256	.03092	.02943	.02809
2	.09962	.09362	.08844	.08380	.07962	.07584	.07240
3	.16102	.15163	.14327	.13579	.12904	.12293	.11737
4	.22268	.20972	.19819	.18785	.17853	.17009	.16241
5	.28434	.26781	.25309	.23990	.22802	.21725	.20746
6	.34597	.32588	.30798	.29195	.27749	.26441	.25249
7	.40759	.38392	.36285	.34397	.32696	.31154	.29751
8	.46920	.44196	.41772	.39599	.37641	.35866	.34252
9	.53080	.50000	.47257	.44800	.42585	.40578	.38752
10	.59241	.55804	.52743	.50000	.47528	.45289	.43252
11	.65403	.61608	.58228	.55200	.52472	.50000	.47750
12	.71566	.67412	.63715	.60401	.57415	.54711	.52250
13	.77732	.73219	.69202	.65603	.62359	.59422	.56748
14	.83898	.79028	.74691	.70805	.67304	.64134	.61248
15	.90038	.84837	.80181	.76010	.72251	.68846	.65748
16	.96122	.90638	.85673	.81215	.77198	.73559	.70249
17		.96359	.91156	.86421	.82147	.78275	.74751
18			.96562	.91620	.87096	.82991	.79254
19				.96744	.92038	.87707	.83759
20					.96908	.92416	.88263
21						.97057	.92760
22							.97191

TABLE 2 *Continued*

n / k	23	24	25	26	27	28	29
1	.02686	.02573	.02469	.02373	.02284	.02202	.02125
2	.06925	.06624	.06372	.06127	.05900	.05689	.05492
3	.11229	.10763	.10334	.09938	.09571	.09230	.08912
4	.15540	.14896	.14303	.13756	.13249	.12778	.12339
5	.19851	.19029	.18273	.17574	.16927	.16326	.15766
6	.24160	.23161	.22242	.21392	.20605	.19873	.19192
7	.28469	.27279	.26210	.25209	.24282	.23420	.22618
8	.32777	.31422	.30176	.29024	.27957	.26966	.26042
9	.37083	.35553	.34142	.32839	.31632	.30511	.29466
10	.41389	.39680	.38107	.36653	.35306	.34055	.32889
11	.45695	.43808	.42072	.40467	.38980	.37599	.36312
12	.50000	.47936	.46036	.44280	.42654	.41142	.39734
13	.54305	.52064	.50000	.48093	.46327	.44686	.43156
14	.58611	.56192	.53964	.51907	.50000	.48228	.46578
15	.62917	.60320	.57928	.55720	.53673	.51772	.50000
16	.67223	.64448	.61893	.59533	.57346	.55314	.53422
17	.71531	.68578	.65858	.63347	.61020	.58858	.56844
18	.75840	.72721	.69824	.67161	.64694	.62401	.60266
19	.80149	.76839	.73791	.70976	.68368	.65945	.63688
20	.84460	.80971	.77758	.74791	.72043	.69489	.67111
21	.88771	.85104	.81727	.78608	.75718	.73034	.70534
22	.93075	.89237	.85697	.82426	.79395	.76580	.73958
23	.97314	.93376	.89666	.86244	.83073	.80127	.77382
24		.97427	.93628	.90062	.86751	.83674	.80808
25			.97531	.93873	.90429	.87222	.84234
26				.97627	.94100	.90770	.87661
27					.97716	.94311	.91088
28						.97798	.94508
29							.97875

TABLE 2 *Continued*

n / k	30	31	32	33	34	35	36
1	.02054	.01987	.01924	.01865	.01810	.01758	.01708
2	.05309	.05138	.04977	.04826	.04683	.04549	.04423
3	.08616	.08338	.08078	.07833	.07603	.07386	.07181
4	.11929	.11545	.11186	.10847	.10529	.10229	.09946
5	.15242	.14753	.14294	.13862	.13456	.13073	.12711
6	.18556	.17960	.17402	.16877	.16382	.15916	.15476
7	.21868	.21166	.20509	.19890	.19308	.18759	.18240
8	.25180	.24372	.23615	.22903	.22233	.21601	.21004
9	.28490	.27577	.26720	.25915	.25158	.24442	.23767
10	.31800	.30781	.29825	.28927	.28081	.27283	.26530
11	.35110	.33985	.32930	.31938	.31005	.30124	.29292
12	.38419	.37189	.36034	.34949	.33928	.32964	.32054
13	.41728	.40392	.39138	.37960	.36850	.35804	.34815
14	.45037	.43594	.42241	.40970	.39773	.38643	.37576
15	.48346	.46797	.45345	.43980	.42695	.41483	.40337
16	.51654	.50000	.48449	.46990	.45617	.44323	.43098
17	.54963	.53203	.51551	.50000	.48539	.47161	.45879
18	.58272	.56406	.54655	.53010	.51461	.50000	.48620
19	.61581	.59608	.57759	.56020	.54383	.52839	.51380
20	.64890	.62811	.60862	.59030	.57305	.55678	.54121
21	.68200	.66015	.63966	.62040	.60227	.58517	.56902
22	.71510	.69219	.67070	.65051	.63150	.61357	.59663
23	.74820	.72423	.70175	.68062	.66072	.64196	.62424
24	.78132	.75628	.73280	.71073	.68995	.67036	.65185
25	.81444	.78834	.76385	.74085	.71919	.69876	.67946
26	.84758	.82040	.79491	.77097	.74842	.72717	.70708
27	.88071	.85247	.82598	.80110	.77767	.75558	.73470
28	.91384	.88455	.85706	.83123	.80692	.78399	.76233
29	.94691	.91662	.88814	.86138	.83618	.81241	.78996
30	.97946	.94862	.91922	.89153	.86544	.84084	.81760
31		.98013	.95023	.92167	.89471	.86927	.84524
32			.98076	.95174	.92397	.89771	.87289
33				.98135	.95317	.92614	.90054
34					.98190	.95451	.92819
35						.98242	.95577
36							.98292

TABLE 2 *Continued*

n / k	37	38	39	40	41	42	43
1	.01662	.01617	.01575	.01536	.01498	.01462	.01427
2	.04303	.04189	.04081	.03979	.03882	.03789	.03717
3	.06987	.06803	.06629	.06463	.06305	.06155	.06012
4	.09677	.09423	.09182	.08953	.08735	.08527	.08329
5	.12368	.12044	.11736	.11443	.11165	.10899	.10646
6	.15059	.14664	.14289	.13933	.13595	.13272	.12964
7	.17749	.17251	.16842	.16423	.16024	.15643	.15281
8	.20439	.19903	.19395	.18912	.18453	.18015	.17597
9	.23128	.22522	.21947	.21401	.20881	.20386	.19844
10	.25816	.25140	.24499	.23889	.23309	.22756	.22229
11	.28504	.27758	.27050	.26377	.25737	.25126	.24545
12	.31192	.30376	.29601	.28864	.28164	.27496	.26859
13	.33879	.32993	.32151	.31352	.30591	.29866	.29174
14	.36566	.35609	.34701	.33838	.33017	.32235	.31489
15	.39253	.38226	.37251	.36325	.35444	.34604	.33803
16	.41940	.40843	.39801	.38812	.37870	.36973	.36117
17	.44627	.43459	.42351	.41298	.40296	.39341	.38431
18	.47314	.46076	.44901	.43784	.42722	.41710	.40745
19	.50000	.48692	.47451	.46271	.45148	.44079	.43059
20	.52686	.51308	.50000	.48757	.47574	.46447	.45373
21	.55373	.53924	.52549	.51243	.50000	.48816	.47686
22	.58060	.56541	.55099	.53729	.52426	.51184	.50000
23	.60747	.59157	.57649	.56216	.54852	.53553	.52314
24	.63434	.61774	.60199	.58702	.57278	.55921	.54627
25	.66121	.64391	.62749	.61188	.59704	.58290	.56941
26	.68808	.67007	.65299	.63675	.62130	.60659	.59255
27	.71496	.69624	.67849	.66162	.64556	.63027	.61569
28	.74184	.72242	.70399	.68648	.66983	.65396	.63883
29	.76872	.74860	.72950	.71136	.69409	.67765	.66197
30	.79561	.77478	.75501	.73623	.71836	.70134	.68511
31	.82251	.80097	.78053	.76111	.74263	.72504	.70826
32	.84941	.82749	.80605	.78599	.76691	.74874	.73141
33	.87632	.85336	.83158	.81088	.79119	.77244	.75455
34	.90323	.87956	.85711	.83577	.81547	.79614	.77771
35	.93013	.90577	.88264	.86067	.83976	.81985	.80156
36	.95697	.93197	.90818	.88557	.86405	.84357	.82403
37	.98338	.95811	.93371	.91047	.88835	.86728	.84719
38		.98383	.95919	.93537	.91265	.89101	.87036
39			.98425	.96021	.93695	.91473	.89354
40				.98464	.96118	.93845	.91671
41					.98502	.96211	.93988
42						.98538	.96283
43							.98573

TABLE 2 *Continued*

n / k	44	45	46	47	48	49	50
1	.01395	.01363	.01333	.01305	.01277	.01251	.01225
2	.03616	.03536	.03458	.03385	.03314	.03246	.03181
3	.05875	.05745	.05620	.05500	.05385	.05275	.05170
4	.08140	.07959	.07786	.07620	.07462	.07310	.07163
5	.10405	.10174	.09953	.09742	.09539	.09345	.09158
6	.12670	.12389	.12120	.11863	.11616	.11380	.11152
7	.14935	.14604	.14287	.13984	.13693	.13414	.13147
8	.17199	.16818	.16453	.16104	.15770	.15449	.15141
9	.19463	.19032	.18619	.18225	.17846	.17483	.17134
10	.21726	.21245	.20785	.20344	.19922	.19516	.19127
11	.23989	.23458	.22950	.22464	.21997	.21550	.21120
12	.26252	.25671	.25115	.24583	.24073	.23583	.23113
13	.28514	.27883	.27280	.26702	.26147	.25616	.25106
14	.30776	.30095	.29444	.28820	.28222	.27649	.27098
15	.33039	.32307	.31608	.30939	.30297	.29598	.29090
16	.35300	.34519	.33773	.33057	.32371	.31713	.31082
17	.37562	.36731	.35936	.35175	.34445	.33745	.33073
18	.39823	.38943	.38100	.37293	.36520	.35778	.35065
19	.42085	.41154	.40264	.39411	.38594	.37809	.37056
20	.44346	.43366	.42427	.41529	.40668	.39841	.39048
21	.46608	.45577	.44591	.43647	.42741	.41873	.41039
22	.48869	.47789	.46755	.45764	.44815	.43905	.43031
23	.51131	.50000	.48918	.47882	.46889	.45937	.45022
24	.53392	.52211	.51082	.50000	.48963	.47969	.47013
25	.55654	.54423	.53245	.52118	.51037	.50000	.49004
26	.57915	.56634	.55409	.54236	.53111	.52032	.50996
27	.60177	.58846	.57573	.56353	.55185	.54063	.52987
28	.62438	.61057	.59736	.58471	.57259	.56095	.54978
29	.64700	.63269	.61900	.60589	.59332	.58127	.56969
30	.66961	.65481	.64064	.62707	.61406	.60159	.58961
31	.69224	.67693	.66228	.64825	.63480	.62191	.60952
32	.71486	.69905	.68392	.66943	.65555	.64222	.62944
33	.73748	.72117	.70556	.69061	.67629	.66255	.64935
34	.76011	.74329	.72720	.71180	.69703	.68287	.66927
35	.78274	.76542	.74885	.73298	.71778	.70402	.68918
36	.80537	.78755	.77050	.75417	.73853	.72351	.70910
37	.82801	.80968	.79215	.77536	.75927	.74384	.72902
38	.85065	.83182	.81381	.79656	.78003	.76417	.74894
39	.87330	.85396	.83547	.81775	.80078	.78450	.76887
40	.89595	.87611	.85713	.83896	.82154	.80484	.78880

TABLE 2 *Continued*

n / k	44	45	46	47	48	49	50
41	.91860	.89826	.87880	.86016	.84230	.82517	.80873
42	.94125	.92041	.90047	.88137	.86307	.84551	.82866
43	.96384	.94255	.92214	.90258	.88384	.86586	.84859
44	.98605	.96464	.94380	.92380	.90461	.88620	.86853
45		.98637	.96542	.94500	.92538	.90655	.88848
46			.98667	.96615	.94615	.92690	.90842
47				.98695	.96686	.94725	.92837
48					.98723	.96754	.94830
49						.98749	.96819
50							.98775

TABLE 2 *Continued*

n / k	51	52	53	54	55	56	57
1	.01201	.01178	.01155	.01134	.01113	.01093	.01073
2	.03118	.03058	.03000	.02944	.02890	.02839	.02789
3	.05068	.04970	.04876	.04786	.04699	.04615	.04534
4	.07023	.06888	.06758	.06633	.06512	.06396	.06284
5	.08979	.08806	.08640	.08480	.08326	.08178	.08034
6	.10934	.10724	.10522	.10328	.10140	.09959	.09785
7	.12890	.12642	.12404	.12175	.11954	.11741	.11535
8	.14845	.14560	.14286	.14022	.13767	.13522	.13285
9	.16799	.16477	.16167	.15868	.15580	.15303	.15035
10	.18754	.18394	.18048	.17715	.17393	.17084	.16785
11	.20708	.20311	.19929	.19561	.19206	.18864	.18534
12	.22662	.22227	.21809	.21406	.21018	.20644	.20283
13	.24615	.24143	.23689	.23252	.22831	.22424	.22032
14	.26568	.26059	.25569	.25097	.24643	.24204	.23781
15	.28521	.27975	.27449	.26943	.26454	.29600	.25529
16	.30475	.29891	.29330	.28788	.28266	.27763	.27278
17	.32427	.31806	.31208	.30633	.30078	.29542	.29026
18	.34380	.33722	.33088	.32477	.31889	.31322	.30774
19	.36333	.35637	.34967	.34322	.33701	.33101	.32522
20	.38286	.37552	.36847	.36167	.35512	.34880	.34270
21	.40238	.39467	.38726	.38011	.37323	.36659	.36018
22	.42191	.41382	.40605	.39856	.39134	.38438	.37766
23	.44127	.43298	.42484	.41700	.40945	.40217	.39514
24	.46095	.45213	.44363	.43545	.42756	.41995	.41262
25	.48048	.47128	.46242	.45389	.44567	.43774	.43009
26	.50000	.49043	.48121	.47234	.46378	.45553	.44757
27	.51952	.50957	.50000	.49078	.48189	.47332	.46509
28	.53905	.52872	.51879	.50922	.50000	.49111	.48252
29	.55873	.54787	.53758	.52766	.51811	.50889	.50000
30	.57809	.56702	.55637	.54611	.53622	.52668	.51748
31	.59762	.58618	.57516	.56455	.55433	.54447	.53495
32	.61714	.60533	.59395	.58300	.57244	.56226	.55243
33	.63667	.62448	.61274	.60144	.59055	.58005	.56991
34	.65620	.64363	.63153	.61989	.60866	.59783	.58738
35	.67573	.66278	.65033	.63833	.62677	.61562	.60486
36	.69525	.68194	.66912	.65678	.64488	.63341	.62234
37	.71479	.70109	.68792	.67523	.66299	.65120	.63982
38	.73432	.72025	.70671	.69367	.68111	.66899	.65730
39	.75385	.73941	.72551	.71212	.69922	.68678	.67478
40	.77338	.75857	.74431	.73057	.71734	.70458	.69226

TABLE 2 *Continued*

n / k	51	52	53	54	55	56	57
41	.79292	.77773	.76311	.74903	.73546	.72237	.70974
42	.81246	.79689	.78191	.76748	.75357	.70400	.72722
43	.83201	.81606	.80071	.78594	.77169	.75796	.74471
44	.85155	.83523	.81952	.80439	.78982	.77576	.76219
45	.87110	.85440	.83833	.82285	.80794	.79356	.77968
46	.89066	.87358	.85714	.84132	.82607	.81136	.79717
47	.91021	.89276	.87596	.85978	.84420	.82916	.81466
48	.92977	.91194	.89478	.87825	.86233	.84697	.83215
49	.94932	.93112	.91360	.89672	.88046	.86478	.84965
50	.96882	.95030	.93242	.91520	.89860	.88259	.86715
51	.98799	.96942	.95124	.93367	.91674	.90041	.88465
52		.98822	.97000	.95214	.93488	.91822	.90215
53			.98845	.97056	.95301	.93604	.91966
54				.98866	.97110	.95385	.93716
55					.98887	.97161	.95466
56						.98907	.97211
57							.98927

TABLE 2 *Continued*

k \ n	58	59	60	65	70	75	80
1	.01055	.01036	.01019	.00940	.00872	.00813	.00762
2	.02740	.02694	.02649	.02444	.02269	.02117	.01984
3	.04455	.04380	.04307	.03975	.03690	.03444	.03228
4	.06175	.06070	.05969	.05510	.05116	.04774	.04476
5	.07896	.07762	.07633	.07046	.06542	.06106	.05724
6	.09658	.09453	.09296	.08582	.07969	.07438	.06973
7	.11337	.11145	.10959	.10117	.09395	.08769	.08222
8	.13057	.12836	.12622	.11653	.10821	.10101	.09470
9	.14776	.14527	.14285	.13188	.12248	.11432	.10718
10	.16496	.16217	.15947	.14723	.13673	.12763	.11967
11	.18215	.17907	.17610	.16258	.15099	.14094	.13214
12	.19934	.19597	.19272	.17793	.16524	.15425	.14462
13	.21653	.21287	.20934	.19327	.17950	.16756	.15710
14	.23372	.22977	.22595	.20862	.19375	.18086	.16958
15	.25091	.24667	.24257	.22396	.20800	.19416	.18205
16	.26809	.26356	.25918	.23930	.22225	.20747	.19453
17	.28527	.28045	.27579	.25464	.23650	.22077	.20700
18	.30245	.29735	.29241	.26998	.25075	.23407	.21947
19	.31963	.31424	.30902	.28532	.26499	.24737	.23194
20	.33682	.33112	.32563	.30065	.27924	.26067	.24442
21	.35399	.34802	.34224	.31599	.29399	.27397	.25689
22	.37117	.36490	.35884	.33133	.30773	.28727	.26936
23	.38835	.38179	.37545	.34666	.32197	.30057	.28183
24	.40553	.39868	.39206	.36200	.33622	.31386	.29430
25	.42270	.41557	.40867	.37733	.35046	.32716	.30677
26	.43988	.43245	.42527	.39267	.36471	.34046	.31923
27	.45706	.44934	.44188	.40800	.37895	.35375	.33170
28	.47423	.46623	.45849	.42333	.39319	.36705	.34417
29	.49141	.48311	.47509	.43867	.40743	.38035	.35663
30	.50859	.50000	.49170	.45400	.42167	.39364	.36910
31	.52577	.51689	.50830	.46934	.43591	.40694	.38157
32	.54294	.53377	.52491	.48467	.45015	.42023	.39404
33	.56012	.55066	.54151	.50000	.46440	.43353	.40650
34	.57730	.56755	.55812	.51533	.47864	.44682	.41897
35	.59447	.58443	.57473	.53066	.49288	.46012	.43144
36	.61165	.60132	.59133	.54600	.50712	.47341	.44390
37	.62883	.61821	.60794	.56133	.52136	.48671	.45637
38	.64601	.63510	.62455	.57667	.53560	.50000	.46883
39	.66318	.65198	.64116	.59200	.54985	.51329	.48130
40	.68037	.66888	.65776	.60733	.56409	.52659	.49377

TABLE 2 *Continued*

n / k	58	59	60	65	70	75	80
41	.69755	.68576	.67437	.62267	.57833	.53988	.50623
42	.71473	.70265	.69098	.63800	.59257	.55318	.51870
43	.73191	.71955	.70759	.65334	.60681	.56647	.53117
44	.74909	.73644	.72421	.66867	.62105	.57977	.54363
45	.76628	.75333	.74082	.68401	.63529	.59306	.55610
46	.78347	.77023	.75743	.69935	.64954	.60636	.56856
47	.80066	.78713	.77405	.71468	.66378	.61965	.58103
48	.81785	.80403	.79066	.73002	.67803	.63295	.59350
49	.83504	.82093	.80728	.74536	.69227	.64625	.60596
50	.85224	.83783	.82390	.76070	.70601	.65954	.61843
51	.86943	.85473	.84053	.77604	.72076	.67284	.63090
52	.88663	.87164	.85715	.79138	.73501	.68614	.64337
53	.90342	.88855	.87378	.80673	.74925	.69943	.65583
54	.92104	.90547	.89041	.82207	.76350	.71273	.66830
55	.93825	.92238	.90704	.83742	.77775	.72603	.68077
56	.95545	.93930	.92367	.85277	.79200	.73933	.69323
57	.97260	.95620	.94031	.86812	.80625	.75263	.70570
58	.98945	.97306	.95693	.88347	.82050	.76593	.71817
59		.98964	.97351	.89883	.83476	.77923	.73064
60			.98981	.91418	.84901	.79253	.74311
61				.92954	.86327	.80584	.75558
62				.94490	.87752	.81914	.76806
63				.96025	.89179	.83244	.78053
64				.97556	.90605	.84575	.79300
65				.99060	.92031	.85906	.80547
66					.93458	.87237	.81795
67					.94884	.88568	.83042
68					.96310	.89899	.84290
69					.97731	.91231	.85538
70					.99128	.92562	.86786
71						.93894	.88033
72						.95226	.89282
73						.96556	.90530
74						.97883	.91778
75						.99187	.93027
76							.94276
77							.95524
78							.96772
79							.98016
80							.99238

TABLE 2 *Continued*

n / k	85	90	95	100	125	150	175
1	.00716	.00676	.00640	.00608	.00485	.00403	.00345
2	.01866	.01762	.01669	.01585	.01267	.01054	.00903
3	.03037	.02868	.02717	.02581	.02063	.01718	.01472
4	.04212	.03978	.03768	.03579	.02862	.02384	.02043
5	.05387	.05088	.04820	.04578	.03662	.03051	.02614
6	.06563	.06198	.05872	.05578	.04462	.03717	.03186
7	.07738	.07308	.06924	.06577	.05261	.04384	.03757
8	.08913	.08418	.07975	.07577	.06061	.05051	.04329
9	.10088	.09528	.09027	.08576	.06861	.05717	.04910
10	.11263	.10638	.10079	.09575	.07661	.06384	.05472
11	.12438	.11748	.11130	.10574	.08460	.07050	.06043
12	.13613	.12857	.12182	.11573	.09260	.07717	.06614
13	.14787	.13967	.13233	.12572	.10059	.08383	.07186
14	.15962	.15076	.14284	.13571	.10858	.09050	.07757
15	.17136	.16186	.15335	.14569	.11658	.09716	.08328
16	.18310	.17295	.16386	.15568	.12457	.10382	.08899
17	.19485	.18404	.17437	.16566	.13256	.11048	.09471
18	.20659	.19513	.18488	.17565	.14055	.11715	.10042
19	.21833	.20622	.19539	.18563	.14855	.12381	.10613
20	.23007	.21731	.20589	19561	.15654	.13047	.11184
21	.24181	.22840	.21640	.20560	.16453	.13713	.11755
22	.25355	.23949	.22691	.21558	.17252	.14379	.12326
23	.26528	.25058	.23742	.22557	.18051	.15045	.12897
24	.27702	.26167	.24792	.23555	.18850	.15711	.13468
25	.28876	.27275	.25842	.24553	.19649	.16377	.14039
26	.30050	.28384	.26893	.25551	.20448	.17043	.14610
27	.31224	.29493	.27944	.26549	.21247	.17709	.15182
28	.32397	.30601	.28994	.27547	.22046	.18375	.15752
29	.33571	.31710	.30045	.28545	.22844	.19041	.16323
30	.34744	.32819	.31095	.29543	.23643	.19707	.16894
31	.35918	.33927	.32145	.30541	.24442	.20373	.17465
32	.37092	.35036	.33196	.31539	.25241	.21039	.18036
33	.38265	.36144	.34246	.32537	.26040	.21705	.18607
34	.39439	.37253	.35296	.33535	.26839	.22371	.19178
35	.40613	.38361	.36347	.34533	.27637	.23037	.19749
36	.41786	.39470	.37397	.35531	.28436	.23703	.20320
37	.42959	.40578	.38447	.36529	.29235	.24368	.20891
38	.44133	.41687	.39498	.37527	.30034	.25034	.21462
39	.45306	.42795	.40548	.38525	.30832	.25700	.22032
40	.46476	.43904	.41598	.39523	.31631	.26366	.22603

TABLE 2 *Continued*

n / k	85	90	95	100	125	150	175
41	.47653	.45012	.42649	.40520	.32430	.27032	.23174
42	.48826	.46120	.43699	.41518	.33229	.27698	.23745
43	.50000	.47229	.44749	.42516	.34027	.28363	.24316
44	.51174	.48337	.45799	.43514	.34826	.29029	.24887
45	.52347	.49446	.46849	.44512	.35625	.29695	.25458
46	.53524	.50554	.47900	.45510	.36423	.30361	.26028
47	.54694	.51663	.48950	.46508	.37222	.31027	.26599
48	.55867	.52771	.50000	.47505	.38021	.31692	.27170
49	.57041	.53880	.51050	.48503	.38819	.32358	.27741
50	.58214	.54988	.52100	.49501	.39618	.33024	.28312
51	.59387	.56096	.53151	.50499	.40417	.33690	.28883
52	.60561	.57205	.54201	.51497	.41215	.34356	.29453
53	.61735	.58313	.55251	.52495	.42014	.35021	.30024
54	.62908	.59422	.56301	.53492	.42812	.35687	.30595
55	.64082	.60530	.57351	.54490	.43611	.36353	.31166
56	.65256	.61639	.58402	.55488	.44410	.37019	.31737
57	.66429	.62747	.59452	.56486	.45208	.37684	.32307
58	.67603	.63856	.60502	.57484	.46007	.38350	.32878
59	.68776	.64964	.61553	.58482	.46806	.39016	.33449
60	.69950	.66073	.62603	.59480	.47604	.39681	.34020
61	.71124	.67181	.63653	.60477	.48403	.40347	.34590
62	.72298	.68290	.64704	.61475	.49201	.41013	.35161
63	.73472	.69399	.65754	.62473	.50000	.41679	.35732
64	.74645	.70507	.66804	.63471	.50799	.42344	.36302
65	.75819	.71616	.67855	.64469	.51597	.43010	.36873
66	.76993	.72725	.68905	.65467	.52396	.43676	.37444
67	.78167	.73833	.69955	.66465	.53195	.44341	.38015
68	.79341	.74942	.71006	.67463	.53993	.45007	.38585
69	.80515	.76051	.72056	.68461	.54792	.45673	.39156
70	.81690	.77160	.73107	.69459	.55590	.46338	.39727
71	.82864	.78269	.74158	.70457	.56389	.47004	.40298
72	.84038	.79378	.75208	.71455	.57188	.47670	.40868
73	.85213	.80487	.76258	.72453	.57986	.48336	.41439
74	.86387	.81596	.77309	.73451	.58785	.49002	.42010
75	.87562	.82705	.78360	.74449	.59583	.49667	.42581
76	.88737	.83814	.79411	.75447	.60382	.50333	.43151
77	.89912	.84924	.80461	.76445	.61181	.50998	.43722
78	.91087	.86033	.81512	.77443	.61979	.51664	.44293
79	.92262	.87143	.82563	.78442	.62778	.52330	.44864
80	.93437	.88252	.83614	.79440	.63577	.52996	.45434

TABLE 2 *Continued*

k \ n	85	90	95	100	125	150	175
81	.94613	.89362	.84665	.80439	.64375	.53662	.46005
82	.95788	.90472	.85716	.81437	.65174	.54327	.46576
83	.96963	.91582	.86767	.82435	.65973	.54993	.47146
84	.98134	.92692	.87818	.83434	.66771	.55659	.47717
85	.99284	.93802	.88870	.84432	.67570	.56324	.48288
86		.94912	.89921	.85431	.68369	.56990	.48858
87		.96022	.90973	.86429	.69168	.57656	.49429
88		.97132	.92025	.87428	.69966	.58321	.50000
89		.98238	.93076	.88427	.70765	.58987	.50571
90		.99324	.94128	.89426	.71564	.59653	.51142
91			.95180	.90425	.72363	.60319	.51712
92			.96232	.91424	.73161	.60984	.52283
93			.97283	.92423	.73960	.61650	.52854
94			.98331	.93423	.74759	.62316	.53424
95			.99360	.94422	.75558	.62981	.53995
96				.95422	.76357	.63647	.54566
97				.96421	.77156	.64313	.55136
98				.97419	.77954	.64979	.55707
99				.98415	.78753	.65644	.56278
100				.99392	.79552	.66310	.56849
101					.80351	.66976	.57419
102					.81150	.67642	.57990
103					.81949	.68308	.58561
104					.82748	.68973	.59132
105					.83547	.69639	.59702
106					.84346	.70305	.60273
107					.85145	.70971	.60844
108					.85945	.71637	.61415
109					.86744	.72302	.61986
110					.87543	.72968	.62556
111					.88342	.73634	.63127
112					.89142	.74300	.63698
113					.89941	.74966	.64268
114					.90740	.75632	.64839
115					.91540	.76297	.65410
116					.92339	.76963	.65980
117					.93139	.77629	.66551
118					.93939	.78295	.67122
119					.94739	.78961	.67693
120					.95538	.79627	.68264

TABLE 2 *Continued*

n / k	125	150	175	n / k	125	150	175
121	.96338	.80293	.68834	161			.91672
122	.97138	.80959	.69405	162			.92243
123	.97937	.81625	.69976	163			.92814
124	.98733	.82291	.70547	164			.93386
125	.99515	.82957	.71117	165			.93957
126		.83623	.71688	166			.94528
127		.84289	.72259	167			.95090
128		.84955	.72830	168			.95671
129		.85621	.73401	169			.96243
130		.86287	.73972	170			.96814
131		.86953	.74542	171			.97386
132		.87619	.75113	172			.97957
133		.88285	.75684	173			.98528
134		.88952	.76255	174			.99097
135		.89618	.76826	175			.99655
136		.90284	.77397				
137		.90950	.77968				
138		.91617	.78539				
139		.92283	.79109				
140		.92950	.79680				
141		.93616	.80251				
142		.94283	.81822				
143		.94949	.81393				
144		.95616	.81964				
145		.96283	.82535				
146		.96949	.83106				
147		.97616	.83677				
148		.98282	.84248				
149		.98946	.84819				
150		.99597	.85390				
151			.85961				
152			.86532				
153			.87103				
154			.87674				
155			.88245				
156			88816				
157			.89387				
158			.89956				
159			.90529				
160			.91101				

TABLE 2 *Continued*

n / k	200	225	250	n / k	200	225	250
1	.00302	.00268	.00241	41	.20280	.18028	.16226
2	.00790	.00702	.00631	42	.20779	.18472	.16626
3	.01287	.01144	.01029	43	.21279	.18916	.17026
4	.01787	.01588	.01429	44	.21779	.19360	.17426
5	.02287	.02032	.01829	45	.22278	.19805	.17825
6	.02787	.02477	.02229	46	.22778	.20249	.18225
7	.03287	.02921	.02629	47	.23277	.20693	.18625
8	.03787	.03366	.03029	48	.23777	.21137	.19025
9	.04287	.03811	.03429	49	.24276	.21581	.19424
10	.04787	.04255	.03829	50	.24776	.22025	.19824
11	.05287	.04700	.04229	51	.25275	.22469	.20224
12	.05787	.05144	.04630	52	.25775	.22923	.20624
13	.06287	.05589	.05030	53	.26275	.23358	.21023
14	.06787	.06033	.05430	54	.26774	.23802	.21423
15	.07287	.06477	.05830	55	.27274	.24246	.21823
16	.07787	.06922	.06230	56	.27773	.24690	.22223
17	.08287	.07366	.06630	57	.28273	.25134	.22622
18	.08787	.07811	.07030	58	.28772	.25578	.23022
19	.09287	.08255	.07430	59	.29272	.26022	.23422
20	.09787	.08699	.07830	60	.29771	.26466	.23821
21	.10286	.09144	.08230	61	.30271	.27240	.24221
22	.10786	.09588	.08629	62	.30770	.27354	.24621
23	.11286	.10032	.09029	63	.31270	.27798	.25021
24	.11786	.10477	.09429	64	.31769	.28242	.25420
25	.12285	.10921	.09829	65	.32269	.28686	.25820
26	.12785	.11365	.10229	66	.32768	.29130	.26220
27	.13285	.11809	.10629	67	.33268	.29574	.26619
28	.13785	.12254	.11029	68	.33767	.30019	.27019
29	.14284	.12698	.11429	69	.34267	.30463	.27419
30	.14784	.13142	.11828	70	.34766	.30907	.27819
31	.15284	.13586	.12228	71	.35266	.31351	.28218
32	.15783	.14031	.12628	72	.35765	.31795	.28618
33	.16283	.14475	.13028	73	.36265	.32239	.29018
34	.16783	.14919	.13428	74	.36764	.32683	.29417
35	.17282	.15363	.13828	75	.37264	.33127	.29817
36	.17782	.15807	.14227	76	.37763	.33571	.30217
37	.18281	.16251	.14627	77	.38263	.34015	.30616
38	.18781	.16696	.15027	78	.38762	.34459	.31016
39	.19281	.17140	.15427	79	.39262	.34903	.31416
40	.19780	.17584	.15827	80	.39761	.35347	.31815

TABLE 2 *Continued*

n / k	200	225	250	n / k	200	225	250
81	.40261	.35791	.32215	121	.60239	.53552	.48201
82	.40760	.36235	.32615	122	.60739	.53996	.48601
83	.41259	.36679	.33014	123	.61238	.54440	.49001
84	.41759	.37123	.33414	124	.61737	.54884	.49400
85	.42258	.37567	.33814	125	.62237	.55328	.49800
86	.42758	.38011	.34213	126	.62736	.55772	.50200
87	.53257	.38455	.34613	127	.63236	.56216	.50600
88	.43757	.38899	.35013	128	.63735	.56660	.50999
89	.44256	.39343	.35412	129	.64235	.57104	.51399
90	.44756	.39788	.35812	130	.64734	.57548	.51799
91	.45255	.40231	.36212	131	.65234	.57992	.52198
92	.45755	.40676	.36612	132	.65733	.58436	.52598
93	.46254	.41120	.37011	133	.66233	.58880	.52997
94	.46753	.41564	.37411	134	.66732	.59324	.53397
95	.47249	.42008	.37811	135	.67232	.59769	.53797
96	.47752	.42452	.38210	136	.67731	.60212	.54196
97	.48252	.42896	.38610	137	.68231	.60657	.54596
98	.48752	.43340	.39010	138	.68730	.61101	.54996
99	.49251	.43784	.39409	139	.69230	.61545	.55396
100	.49750	.44228	.39809	140	.69729	.61989	.55795
101	.50250	.44672	.40208	141	.70229	.62433	.56195
102	.50749	.45116	.40608	142	.70728	.62877	.56595
103	.51248	.45560	.41008	143	.71228	.63321	.56994
104	.51748	.46004	.41407	144	.71727	.63765	.57394
105	.52248	.46448	.41807	145	.72227	.64209	.57793
106	.52751	.46892	.42207	146	.72726	.64653	.58193
107	.53247	.47336	.42606	147	.73226	.65097	.58593
108	.53746	.47780	.43006	148	.73725	.65541	.58992
109	.54245	.48224	.43405	149	.74225	.65985	.59392
110	.54745	.48668	.43805	150	.74725	.66429	.59792
111	.55244	.49112	.44205	151	.75224	.66873	.60191
112	.55744	.49556	.44605	152	.75724	.67317	.60591
113	.56243	.50000	.45004	153	.76223	.67761	.60990
114	.56743	.50444	.45404	154	.76723	.68205	.61390
115	.57242	.50888	.45804	155	.77222	.68649	.61790
116	.57742	.51332	.46203	156	.77722	.69093	.62189
117	.58241	.51776	.46603	157	.78221	.69537	.62589
118	.58741	.52220	.47003	158	.78721	.69981	.62989
119	.59240	.52664	.47402	159	.79221	.70426	.63388
120	.59740	.53108	.47802	160	.79720	.70870	.63788

TABLE 2 *Continued*

n / k	200	225	250	n / k	200	225	250
161	.80220	.71314	.64188	201		.89079	.80176
162	.80719	.71758	.64588	202		.89523	.80576
163	.81219	.72202	.64987	203		.89968	.80975
164	.81719	.72646	.65387	204		.90412	.81375
165	.82218	.72760	.65787	205		.90856	.81775
166	.82718	.73534	.66186	206		.91301	.82175
167	.83217	.73978	.66586	207		.91745	.82574
168	.83717	.74422	.66986	208		.92189	.82974
169	.84217	.74866	.67385	209		.92634	.83374
170	.84716	.75310	.67785	210		.93078	.83774
171	.85216	.75754	.68185	211		.93523	.84173
172	.85716	.76198	.68584	212		.93967	.84573
173	.86215	.76642	.68984	213		.94411	.84973
174	.86715	.77077	.69384	214		.94856	.85373
175	.87215	.77531	.69783	215		.95300	.85773
176	.87715	.77975	.70183	216		.95745	.86172
177	.88214	.78419	.70583	217		.96189	.86572
178	.88714	.78863	.70982	218		.96634	.86972
179	.89214	.79307	.71382	219		.97079	.87372
180	.89714	.79751	.71782	220		.97523	.87772
181	.90213	.80195	.72181	221		.97968	.88172
182	.90713	.80640	.72581	222		.98412	.88571
183	.91213	.81084	.72981	223		.98856	.88971
184	.91713	.81528	.73381	224		.99298	.89371
185	.92213	.81972	.73780	225		.99732	.89771
186	.92713	.82416	.74180	226			.90171
187	.93213	.82860	.74580	227			.90571
188	.93713	.83304	.74979	228			.90971
189	.94213	.83749	.75379	229			.91371
190	.94713	.84193	.75779	230			.91771
191	.95213	.84637	.76179	231			.92170
192	.95713	.85081	.76578	232			.92570
193	.96213	.85525	.76978	233			.92970
194	.96713	.85970	.77378	234			.93370
195	.97213	.86414	.77777	235			.93770
196	.97713	.86858	.78177	236			.94170
197	.98213	.87302	.78577	237			.94570
198	.98713	.87746	.78977	238			.94970
199	.99210	.88191	.79376	239			.95370
200	.99698	.88635	.79776	240			.95771

TABLE 2 *Continued*

n / k	200	225	250
241			.96171
242			.96571
243			.96971
244			.97371
245			.97771
246			.98171
247			.98571
248			.98971
249			.99369
250			.99759

A STRONG JUSTIFICATION FOR THE USE OF RANK TESTS IN THE CASE OF NON-NORMALITY

J. Sethuraman and D. H. Jones
The Florida State University and Rutgers University

This paper is to give a new and strong justification for rank tests in the case of non-normality.

To keep the talk simple we will consider only two problems. It is expected that the same situation will hold in other problems.

Problem I. To test that the median of a symmetric distribution is a specified number, 0, say, versus that the median is not 0.

Problem II. To test that two distributions are identical versus that they are not the same.

1. Problem I. Description of the problem and the basic tests.

Let us consider Problem I in detail. A random variable X has a distribution function F. Independent observations $X_1, X_2, \ldots, X_n$ are available on X. We wish to test the hypothesis

H : $F(x) = G(x)$ where G is an unknown symmetric continuous distribution function versus the alternative

A : $F(x) = G(x - \theta)$ where the unknown $\theta \neq 0$ and G is an unknown symmetric continuous distribution function

In the presence of normality, we can restrict the null-hypothesis as follows:

H_Φ : $F(x) = \Phi(\frac{x}{\sigma})$ where σ is unknown and $\Phi(\frac{x}{\sigma})$ stands for a normal distribution function with mean 0 and variance σ^2.

The preferred parametric test for this hypothesis is the one-sample t-test which is to reject for large values of $|t_n|$ where

$$t_n = \sqrt{\frac{n-1}{n}} \sum X_i \Big/ \sqrt{\frac{1}{n}\sum X_i^2 - (\frac{1}{n}\sum X_i)^2}.$$

A large sample approximation states that, with $a > 0$,

$$P(|t_n| > a) \to 2(1 - \phi(a))$$

where $\phi(x)$ is the standard normal distribution function.

The rank tests one usually uses to test H are called one-sample signed rank tests. They may be briefly described as follows. Arrange $|X_1|, |X_2|, \ldots$ in increasing order of magnitude and let the rank of $|X_i|$ be R_i. Define $\psi_i = 1$ if $X_i \geq 0$, $= 0$ if $X_i < 0$. Let $E_{n1} \ldots E_{nn}$ be some numbers to be called scores. Define

$$T_n = \sum_{i=1}^{n} E_{nR_i}\psi_i$$

to be the one-sample signed rank statistic based on the scores $E_{n1}, \ldots, E_{nn}$. One rejects H if

$$|T_n - \sum_{1}^{n} E_{ni}/2| \text{ is large.}$$

Two such rank tests are commonly used. When $E_{ni} = i$, the scores are called the Wilcoxon scores, and in this case T_n, to be called $T_n(W)$, is known as the one-sample Wilcoxon statistic. When $E_{ni} = E(W_{(i)})$ where $W_{(1)} \ldots W_{(n)}$ are the order statistics of a sample of size n from a half-normal distribution (also called the square root X_1^2 distribution), T_n, to be called $T_n(\phi)$, is known as the one-sample half-normal scores statistic or the Fraser statistic. These two rank tests have special local optimal properties and we will not detail them here.

The following large sample approximations are known. Let

$$T_n^*(W) = \left(T_n(W) - \frac{n(n+1)}{4}\right) \Big/ \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

and

$$T_n^*(\phi) = \left(T_n(\phi) - \frac{n}{\sqrt{2\pi}}\right) \Big/ \sqrt{\frac{n}{4}}$$

Then, under H and with $a > 0$,

$$P(|T_n^*(W)| > a) \to 2(1 - \Phi(a))$$

and

$$P(|T_n^*(\phi)| > a) \to 2(1 - \Phi(a)).$$

2. Illustrative example.

Here is an example taken from the recent book 'Non-parametric Statistical Methods' by Hollander and Wolfe (John Wiley and Sons, 1973, p. 53 Problem 20.) It is expected that the percentage of chromium in the samples of steel used in a certain process will be 18%. In 12 samples, the percentage of chromium differed from 18% by the following figures

-0.6, -0.1, -0.4, 0.1, -0.4, 0.9, -1.1, -0.5, -0.2, -0.6, 6.6, and 8.0.

By direct calculation we find

$t_n = 1.1229$

$T_n^*(W) = -0.4315$

$T_n^*(\phi) = -0.0628$

By using the large sample approximations we should accept the null hypothesis at the 5% level.

3. A heuristic reasoning and a discussion of efficiencies.

Here is the heuristic argument that rank tests should be more efficient than the t-test for testing the hypothesis H. The t-test is designed for the problem of testing the more limited hypothesis H_ϕ and so should fall flat on its face when trying to test H.

The pioneering work of Hodges and Lehmann, and Chernoff and Savage show that if the alternatives are normal and local then the t-test can hold its fort against rank statistics. However, if the alternatives are non-normal and local, the t-test can become less efficient. It is interesting that these conclusions are the same whether the hypothesis H or H_ϕ is being tested.

When the alternatives are not local, say the alternative states that $F(x) = \phi(x - \theta)$ where $\theta \neq 0$ is known, i.e. even if the alternative is normal, the t-test is inferior to the rank tests when testing the hypothesis H. (Of course, when testing for H_ϕ, the t-test is superior.) We present a graph below to show this. The criterion of efficiency we use is Bahadur efficiency, which measures the rate of convergence to zero of the type I error when the type II error is fixed.

To summarize, if the null-hypothesis to be tested includes non-normal distributions, then the rank tests fare better than the t-test even when the alternative hypothesis is normal (of course, also when the alternative hypothesis is non-normal). The older work which studied local alternatives said that the rank tests fare no worse than the t-test. We have shown that they actually fare better.

4. Problem II. Description of the problem and basic tests.

Let $X_1, X_2, \ldots X_m, Y_1, \ldots Y_n$ be two independent samples from two continuous distribution functions F, G, respectively. We write $N = m + n$. We wish to test the hypothesis

$$H : F = G$$

versus the alternative

$$A : F \neq G.$$

In the case of normality, the null hypothesis is restricted to

$$H_\phi : F(x) = G(x) = \phi\left(\frac{x - \theta}{\sigma}\right)$$

for some unknown θ and σ.

In this case, one uses the two-sample t-statistic, t_N, defined by

$$t_N = \frac{\sum X_i/m - \sum Y_i/n}{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right) \cdot \frac{1}{m+n-2}\left(\sum X_i^2 - \frac{1}{m}\left(\sum X_i\right)^2 + \sum Y_i^2 - \frac{1}{n}\left(\sum Y_i\right)^2\right)}}$$

and rejects for large values of $|t_N|$. One can use the large sample approximation that holds under the null hypothesis when $a > 0$:

$$P(|t_N| > a) \to 2(1 - \Phi(a)).$$

To define the rank statistics in this problem arrange the combined sample, $W_1, \ldots W_N$, say, in increasing order and put $Z_{Ni} = 1$ if the i^{th} largest combined observation is an X and $= 0$ if it is a Y. Let $E_{N1}, \ldots E_{NN}$, to be called scores, be some numbers. A simple linear rank statistic, T_N, is defined as follows:

$$T_N = \sum_{i=1}^{N} E_{Ni} Z_{Ni}.$$

One rejects the null hypothesis if $|T_N - n \sum E_{Ni}/N|$ is large.

When $E_{Ni} = i$, the scores are called Wilcoxon scores and T_N, to be called $T_N(W)$, is known as the two-sample Wilcoxon rank statistic. When $E_{Ni} = E(L_{(i)})$, when $L_{(1)}, \ldots L_{(N)}$ are the order statistics from a sample of size N from the standard normal, T_N, to be called $T_N(\phi)$, is known as the normal scores rank statistic.

In large samples we can use the following approximations, which hold under the null hypothesis. Let

$$T_N^*(W) = \left(T_N(W) - \frac{n(N+1)}{2}\right) \Big/ \sqrt{\frac{mn(N+1)}{12}}$$

and

$$T_N^*(\phi) = T_N(\phi) \Big/ \sqrt{\frac{mn}{N}}$$

Under the null hypothesis, with $a > 0$,

$$P(|T_N^*(W)| > a) \to 2(1 - \phi(a))$$

and

$$P(|T_N^*(\phi)| > a) \to 2(1 - \phi(a)).$$

5. Illustrative numerical example.

The following numerical example is taken from 'Non-parametric Statistical Methods' by Hollander and Wolfe (John Wiley and Sons, 1973, Page 74, Problem 1.)

"The data in Table 2 are a subset of the data obtained by Thomas and Simmons (1969), who investigated the relation of sputum histamine levels to inhaled irritants or allergens. The histamine content was reported in micrograms per gram dry weight of sputum. The subjects for this portion of the study consisted of 22 smokers; nine of them were allergics and the remaining 13 were asymptomatic (nonallergic) individuals. Care was taken to avoid people who carried out part of their daily work in an atmosphere of noxious gases or other respiratory toxicants. Table 2 gives the ordered sputum histamine levels for the 22 individuals in the study."

Table 2.

Sputum histamine levels (μg/g dry weight sputum)

Allergics	Nonallergics
1651.0	48.1
1112.0	48.0
102.4	45.5
100.0	41.7
67.6	35.4
65.9	34.3
64.7	32.4
39.6	29.1
31.0	27.3
	18.9
	6.6
	5.2
	4.7

Source. H. V. Thomas and B. Simmons (1969).

From the data we can easily compute that

$t_N = 2.0219$

$T_N^*(W) = 3.1719$

$T_N^*(\phi) = 3.0207$

Using the large sample approximations we conclude that all three tests reject the null hypothesis at the 5% level.

6. Discussion of efficiencies.

Earlier studies, like those in the one-sample case, have shown that even though the t-statistic is designed to test the hypothesis H_0, it is as efficient as the normal scores test for normal local alternatives. We have shown that, even for a fixed alternative of the form $F(x) = \phi(x)$, $G(x) = \phi(x - \theta)$, $\theta \neq 0$, the t-test is less efficient than the Wilcoxon and normal scores test. The same table as in the one-sample case applies. As before, our measure of efficiency is Bahadur efficiency.

To summarise; the t-test is well suited to test null-hypothesis that specify normality. If non-normality is present in the null hypothesis, there will be loss of efficiency if the t-test is used even when the alternative hypothesis is normal.

7. Miscellanea.

Extensions to one-sided alternatives and multi-sample problems along the same lines as above are obvious.

Details regarding the Bahadur efficiency of the t-test are available in a technical report currently under preparation. This new work rectifies an error in earlier work by Woodworth (Annals of Mathematical Statistics, 1970) and others who reported that the t-test had Bahadur efficiency more than unity relative to rank tests.

One needs the special scores mentioned in sections 1 and 4 to compute $T_1^*(\phi)$ and $T_2^*(\phi)$. These scores may be found in Govindarajulu and Eisenstat (Rep. Stat. Appl. Res. JUSE, 1965, 12, 149-164) and Harter, Order Statistics and their use in testing and estimation, Vol. 2, 1969, U.S. Govt. Printing Office.

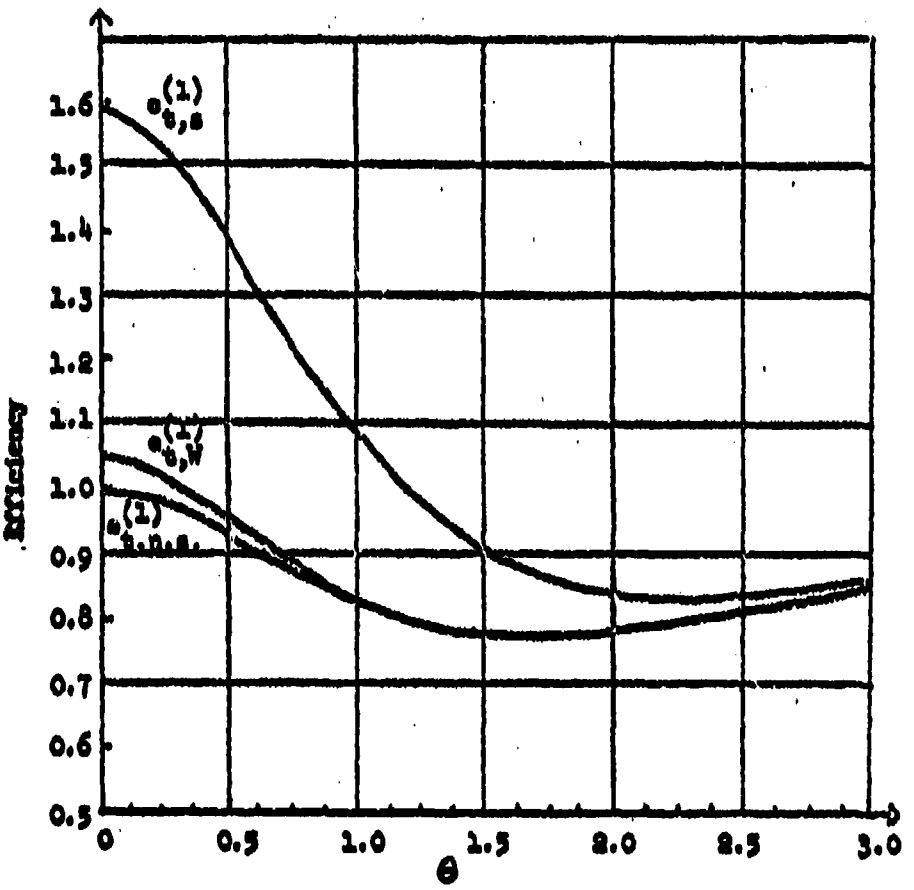


Fig. 1. Efficiencies for normal alternatives, one-sample case

t : t-test
s : sign test
W : Wilcoxon scores test
n.s. : half normal scores test

STATISTICAL ANALYSIS AND MODELING OF PATH LOSS DISTANCE DEPENDENCY INFORMATION

R. D'Accardi and D. Dence, U. S. Army Electronics Command, Fort Monmouth, New Jersey. C. Tsokos, University of South Florida, Tampa, Florida.

ABSTRACT

In recent years extensive investigations have been made in Southeast Asia and the United States to improve communications-electronics performance in heavily forested environments. The investigation involved making extensive propagation measurements at various locations in Thailand. Experimental design included collection of path loss measurements for a frequency range of 100 KHz to 10 GHz encompassing a wide range of antenna heights, distances, polarizations, and seasonal variations.

The initial analysis of the path loss information was basically a descriptive presentation with little statistical analysis and interpretation of the physical characteristics of the experiment. One of the most basic and important aspects of the analysis is the manner by which propagation data should be normalized to a common distance. Further, a careful inspection of the information, the physical aspects of the design of the experiment, and the topological characteristics of the environment reveals that one should not view the path loss data from a deterministic point of view, but rather as a stochnastic realization.

This presentation deals with one aspect of these measurements, namely the statistical analysis and modeling of selected propagation information in the 2-400 MHz range. Several statistical models are proposed for normalizing path loss information which treat the data as stochastic variables. A complete analysis is presented which includes the best estimates of the true path loss distance dependency, α, the formulation of the standard errors involved, and the confidence intervals for the true states of nature.

****

1. INTRODUCTION:

Historical

In recent years extensive investigations and measurements have been made in Southeast Asia to determine the communication conditions that exist in heavily forested environments. Such studies were initiated in 1962, sponsored by the Advanced Research Projects Agency and performed under the direction of the U.S. Army Electronics Command as part of the Southeast Asia Communications Research (SEACORE) Program. The overall aim was to help overcome severe radio communications problems occurring in Southeast Asia. Similar investigations were conducted in the United States over various terrain which included mountains, hills, and forests.

The path loss data in Thailand was obtained by Jansky and Baily, a Division of Atlantic Research Corporation, one of the prime contractors engaged in the SEACORE project. Experiments were designed to collect path loss measurements at various locations in Thailand for a frequency range of

100 KHz to 10 GHz. They encompassed a wide range of antenna heights (transmitting and receiving), locations, polarizations, and seasonal rainfall variations. For the complete chronological progress see references (3), (4), (5) and (1).

Initial Analysis of Path Loss Measurements

A descriptive presentation of the path loss data was presented in references (3) and (4) with minimal statistical analysis and interpretation of the physical characteristics of the experiments. One of the most basic and important aspects of the path loss data during the initial stage of the analysis, is the manner by which the propagation loss measurements should be normalized with reference to a common distance.

Jansky and Baily, (3) and (4), utilized a deterministic model given by the analytical expression:

$$Y_i = X_i - 40 \log d_i \qquad [1]$$

where:

d_i is the path length in miles, $i = 1,2,\ldots,n$,

X_i is the measured path loss over distance d_i, and

Y_i represents the normalized path loss data.

This expression gives a deterministic description of the path loss data as a function of distance. That is, when one theoretically views the jungle as a deterministic environment.

Theoretical studies, (2) and (3), have shown that the main mechanism which guides the energy from the transmitter to the receiver in a jungle environment is a lateral wave, which exhibits a distance power loss of 40 log d in the frequency range of 2 MHz to 100 MHz. In obtaining this analytical relationship one characterizes the electrical characteristics of the jungle environment as a deterministic phenomenon. However, after a careful inspection of the path loss data, the physical aspects of the design of the experiment, and the topological characteristics of the environment, one should not view the path loss information from a deterministic point of view. Preliminary findings show strong evidence that one should normalize the path loss data stochastically.

In the subsequent section we shall mention some relevant findings. A more detailed discussion of the problem from a statistical setting along with other preliminary findings is given in "Statistical Modeling Of Propagation Loss Data", by M. Acker, R. D'Accardi, D. Dence and C. Tsokos, (1).

Statistical Characterization of the Path Loss Data

It was previously mentioned that one of the initial and most important factors which must be considered prior to performing a rigorous statistical

analysis of the SEACORE data is the manner by which one should normalize the data to a common distance. Preliminary findings indicate that the theoretical distance behavior of 40 log d, which generally applies when one considers the electromagnetic environment to be deterministic in nature, is not a realistic relationship, especially above 100 MHz.

A statistical model was proposed, (1), for normalizing the path loss data to a common distance and is given by

$$Z_i = X_i - \alpha_i \log d_i + \epsilon$$

where

ϵ is the random error involved in the experiment which is assumed to be standard normal,

d_i is the path length in miles,

α_i is the parameter to be estimated which describes the path length and propagation loss dependence,

X_i is the path loss over distance d_i,

and

Z_i represents the normalized path loss.

The best statistical estimate in terms of possessing minimum variance of the parameter α_i in the above model, was found to be

$$\hat{\alpha}_i = \frac{\sum_{i=1}^{N} X_i \log(d_i)}{\sum_{i=1}^{N} \log^2(d_i)}$$

Selected data were used (4), to obtain estimates of α_i for the following frequencies:

2, 6, 12, 25.5, 50, 100, 250, 400, MHz,

and for specific combinations of transmitter and receiver antenna heights at distances from 0.2 to 2.0 miles. It was observed that the estimate of α_i fluctuated from 13.3 to 52.5 dB for the configurations analyzed, (1). This fluctuation in the estimate of α_i gives a clear indication that the parameter α should not be treated as a constant, but rather as a stochastic

variable for a more realistic description of the propagation loss distance relationship.

In the preliminary investigation of the SEACORE data (1) an attempt was made to determine whether or not the propagation loss varied from a wet season to a dry season. The information was classified as wet or dry by two criteria:

(i) rainfall greater or less than 3 inches per month,

(ii) rainfall greater or less than 6 inches per month.

For the dry climate a mean estimate of α_1 was found to be 35.7 dB, with a standard error of 6 dB; and for the wet classification a sample mean of 34.8 dB was obtained for α_1 with a standard error of 9.0 dB. Finally, parametric and non-parametric confidence intervals for the mean propagation loss were obtained for the specific configurations investigated. The analysis was performed for the wet and dry classifications as well as for some cumulative seasonal SEACORE data. For additional details of the analysis refer to (1).

Aims of the present study.

The initial information obtained in the experiment is given in dB form. That is, the propagation loss measurements were logarithmically transformed prior to being normalized to a common distance, (3), (4). In the preliminary modeling the path loss data was in dB form and the effects of the logarithmic transformation on the statistics of the information were not investigated. Therefore, it is the partial aim of this study to investigate the fundamental question:

> To what extent did the logarithmic transformation alter the statistics of the data?

An approach to answering this question is to determine whether or not the parameter α_1 varied significantly with a corresponding antilogarithmic model. Such a model is given by the following analytical expression:

$$10^{X_i} = \left(\frac{d_i}{.2}\right)^{\alpha_2}$$

where

d_i is path length referenced at 0.2 miles, $i = .,2,\ldots,n$

X_i is the measured path loss over distance d_i, and

α_2 is the parameter which relates the propagation loss and distance dependence.

Thus, the aim is to obtain the best estimate of the true state of nature, α_2. An exact estimate of α_2 is not possible, because in attempting to obtain a least square estimate for the parameter, a transcendental equation resulted which is not analytically tracticable and therefore, one must rely on approximation techniques for a solution. This expression will be discussed in a subsequent section.

In addition to investigating the fundamental question concerning the analysis of the path loss data using logarithmic vs. anti-logarithmic information, the following will be pursued:

A comparison of two statistical models will be given to determine which one of them will best exhibit the path loss distance dependency for normalizing purposes;

A complete analysis of the data to include all configurations of the SEACORE experiments;

Determination of the mathematical relationship that exists between the propagation loss and distance;

Standard errors involved for the logarithmic model will be formulated;

Upper and lower confidence bounds for the true state of nature will be calculated for 95% confidence.

2. FORMULATION OF THE PROPOSED MODELS

As we mentioned above, one of the primary concerns of the present study is to investigate the behavior of the logarithmic and anti-logarithmic path loss with respect to modeling, analysis, and interpretation of the propagation loss data. To clarify this objective we shall begin the presentation by reviewing the basic linear model that was previously developed (1) and attempt to relate the analytic results to those of the newly proposed model.

A Statistical Model for α_1 :

In the previous analysis of SEACORE information, (1), consideration was given to a linear model of the form

$$Z_i = X_i - \alpha_1 \log \left(\frac{d_i}{.2}\right) + \epsilon \qquad [2]$$

where

X_i = observed path loss data in dB;

α_i = path loss distance dependency parameter

d_i = distance over which path loss was measured,

ε = normally distributed random error due to measurement.

A maximum likelihood estimate for α_i is given by;

$$\hat{\alpha}_i = \frac{\sum_{i=1}^{N} X_i \log\left(\frac{d_i}{.2}\right)}{\sum_{i=1}^{N} \log^2\left(\frac{d_i}{.2}\right)}, \qquad [3]$$

with the data being referenced to a distance of 0.2 miles. This estimate of α_i is unbiased with variance and standard error given by:

$$\text{Var}(\hat{\alpha}_i) = \frac{\frac{1}{n-2}\sum_{i=1}^{n} (X_i - \hat{X}_i)^2}{\sum_{i=1}^{n} \left(\log\left(\frac{d_i}{.2}\right) - \bar{d}\right)^2} \qquad [4]$$

and

$$\text{Standard error } (\hat{\alpha}_i) = \sqrt{\text{var } (\hat{\alpha}_i)} \qquad [5]$$

respectively, where

$$\hat{X}_i = \hat{\alpha}_i \log\left(\frac{d_i}{.2}\right)$$

and the other variables are as previously defined.

Furthermore, the correlation coefficient between the distance, d_i, and the estimate of the propagation path loss behavior, $\hat{X}_i$, will be estimated by:

$$r_{\alpha_i} = \frac{n \sum_{i=1}^{n} \hat{X}_i \log\frac{d_i}{.2} - \left(\sum_{i=1}^{n} \hat{X}_i\right)\left(\sum_{i=1}^{n} \log\frac{d_i}{.2}\right)}{\sqrt{\left\{ n \sum_{i=1}^{n} \hat{X}_i^2 - \left(\sum_{i=1}^{n} \hat{X}_i\right)^2 \right\}\left\{ n \sum_{i=1}^{n} \left(\log\frac{d_i}{.2}\right)^2 - \left(\sum_{i=1}^{n} \log\frac{d_i}{.2}\right)^2 \right\}}} \qquad [6]$$

The purpose of obtaining this estimate of the true correlation, ρ, is to provide an indication as to the type of relationship that exists between

the distance and propagation loss. In addition, interval estimates were obtained for α_1 to answer the question: "How accurate is the MLE estimate of the true state of nature, α_1?" If one has 95% confidence intervals, then this question is adequately answered. The upper and lower confidence intervals for α_1 are given by

$$UB_1 = \alpha_1 + \frac{s\, t_{\alpha/2;\,(n-2)}}{\sqrt{\sum_{i=1}^{n} \left(\log \frac{d_i}{.2} - \bar{d}\right)^2}} \qquad [7]$$

$$LB_1 = \alpha_1 - \frac{s\, t_{\alpha/2;\,(n-2)}}{\sqrt{\sum_{i=1}^{n} \left(\log \frac{d_i}{.2} - \bar{d}\right)^2}} \quad , \qquad [8]$$

respectively, where

$$s = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n} (X_i - \hat{X}_i)^2} \quad , \qquad [9]$$

$$\bar{d} = \frac{1}{n} \sum_{i=1}^{n} \log \frac{d_i}{.2} \quad , \qquad [10]$$

and $t_{\alpha/2;\,(n-2)}$ is the appropriate t statistic for either $\alpha = .05$, or $\alpha = .01$, with n - 2 degrees of freedom, and the other variables are as previously defined.

It is to be noted that if these confidence intervals are small, our estimate of α_1, which gives a measure of the distance and propagation loss dependence, will be fairly close to the true state of nature. Of course, if Var (α_1) is small, one should expect the confidence intervals to be small.

The main reason for including this model, which was previously studied for a small number of configurations, is to obtain a much more complete analysis of the behavior of α_1 by considering a much larger data base. Secondly, exact estimates of the standard errors involved in estimating α_1 are possible; and, thirdly, the question of the primary differences between this stochastic model and the deterministic version (i. e., considering an estimate of α_1 being identically equal to 40) should be answered.

The previous findings utilizing SEACORE data (1) have indicated that for frequencies up to 100 MHz the deterministic and statistical manner of normalizing the data do not seem to bear significant differences. However, beyond 100 MHz there seems to be a significant difference between the two approaches. Thus, by considering a larger class of configurations one should be able to reach a specific decision with regard to this stochastic model versus the deterministic model. It should be mentioned that "not significantly different" with respect to the two models simply means that the dB difference between α_1 and 40 seems to be within ± 5-6 dB.

A Statistical Model For α_2

The most important model in terms of providing the most realistic approach to the question posed, namely, "is there a significant difference between modeling propagation loss information using logarithmic or anti-logarithmic data"?, is given by the expression

$$10^{X_i} = \left(\frac{d_i}{.2}\right)^{\alpha_2} \qquad [11]$$

The X_i again respresent propagation path loss data and the d_i are the corresponding distances involved which are normalized to 0.2 mile. Our aim is to obtain an estimate of the parameter α_2 which represents a nonlinear dependence of X_i and d_i. In attempting ot obtain an estimate of this true state of nature, the least squares approach was utilized which guarantees a minimum variance of the parameter estimated. The implementation of this method yields a transcendental equation which is of the form

$$\sum_{i=1}^{n} \left\{ 10^{X_i} \left(\frac{d_i}{.2}\right)^{\alpha_2} \ln \left(\frac{d_i}{.2}\right) \right\} = \sum_{i=1}^{n} \left\{ \left(\frac{d_i}{.2}\right)^{2\alpha_2} \ln \left(\frac{d_i}{.2}\right) \right\} \qquad [12]$$

Clearly, there is no exact solution to this equation. Thus one must rely primarily on computer techniques to obtain a workable estimate of α_2. This equation was programmed and preliminary results (see section 3) have yielded some significant findings with respect to the model which was discussed in section 2(a), namely, the model for α_1 which utilizes logarithmic data.

Since an exact expression for α_2 cannot be obtained, it is impossible to obtain exact estimates of the variance and standard error of $\hat{\alpha}_2$.

Thus, one must rely on approximate estimates of these statistics, although they may be extremely difficult to obtain. However, in the final analysis, (which will be published at a later date) acceptable formulations for approximating variance and standard error of α_1 will be derived. In addition, it is anticipated that estimates for the confidence bounds on the true state of nature, α_2, will be obtained.

3. SUMMARY OF RESULTS:

In Section 1, we gave a brief historical discussion of the SEACORE path loss experiments; and the fact that previous analysis considered the path-loss distance dependency, $\alpha = 40$, from a purely deterministic point of view,(3),(4). Utilizing the models of Section 2, which treat path loss as a stochastic realization, it was found that for most of the configurations of the experiment, the deterministic approach previously used to normalize the data to a common distance (see equation 1) is not suitable in relating the path loss to a common distance. Tables Ia and Ib show, for three representative frequencies, the minimun variance estimates using the logarithmic data (see equation 3) and the anti-logarithmic data (see equation 12). From these estimates, one can obtain the following information:

a. At lower frequencies (6.0 MHz) for a transmitter antenna height of 40 feet and vertical polarization $\hat{\alpha}_1$ was approximately 2dB less than the deterministic factor (40 dB), while the non-linear descriptor, $\hat{\alpha}_2$, was generally about 11 dB greater. For the same transmitter antenna height, however, the horizontal polarization exhibited a fairly close agreement for $\hat{\alpha}_1$ with respect to its deterministic counterpart. The estimates of α_2 were generally 5 dB greater than the 40 dB factor.

b. An increase in the transmitter antenna height to 80 feet at 6.0 MHz shows a fairly close agreement between $\hat{\alpha}_1$ and $\hat{\alpha}_2$ for smaller receiver antenna heights, however, as the receiver antenna heights increase above 53 feet, both $\hat{\alpha}_1$ and $\hat{\alpha}_2$ vary significantly from 40 dB. $\hat{\alpha}_2$ deviates quite rapidly to 11 dB above the deterministic value for a receiver antenna height of 79 feet. The behavior of both $\hat{\alpha}_1$ and $\hat{\alpha}_2$ is approximately the same as the frequency is increased to 100 MHz, with deviations for horizontal polarization being more pronounced.

c. At 100 MHz for a transmitter antenna height of 40 feet, the α_1 estimates were approximately 3 to 5 dB less than the deterministic case for vertical polarization, while the $\hat{\alpha}_2$'s averaged approximately 6 dB above the "40" factor. For the horizontal polarization, the $\hat{\alpha}_1$'s were fairly consistent with the vertical behavior, however, the $\hat{\alpha}_2$'s deviated more rapidly at lower receiver antenna heights (up to 22 dB for a receive height of 31 feet).

d. For a transmit antenna height of 80 feet at 100 MHz, the α_1 estimates were fairly close (within 3 dB) to the deterministic case for vertical polarization. However, the $\hat{\alpha}_2$'s at receiver antenna heights up to 53 feet, were greater than the deterministic case by as much at 14 dB. For receive heights above 53 feet, the estimates were less by 3 to 8 dB. The $\hat{\alpha}_1$'s deviated from nominal "40" by as much as 4 dB for horizontal polarization, while for higher receiver antenna heights, the α_2 estimates were as much as 16 dB less than the 40 dB factor.

e. As the transmitting frequency is increased to 400 MHz, for all transmitter antenna heights and vertical polarization, $\hat{\alpha}_1$ and $\hat{\alpha}_2$ deviated significantly from the deterministic factor. This deviation was as large as 26 dB for $\hat{\alpha}_1$ and 35 dB for $\hat{\alpha}_2$.

f. The behavior of the α_1 and α_2 estimates for horizontal polarization at 400 MHz were generally the same, however, the $\hat{\alpha}_2$'s deviated from 3 dB to 38 dB from the "40" factor. Not enough data was available at lower receiver heights to be conclusive at this frequency.

g. Table II shows the overall means and standard errors of the path loss parameter estimates. In all cases, small standard errors were exhibited and, in general, as the frequency increased, the means of the α_1 and α_2 estimates decreased significantly. It was observed that at a lower transmitting frequency (6.0 MHz) and transmitting antenna heights of 40 feet and 80 feet (and also for a mid-frequency of 100 MHz and transmitting height of 80 feet) the mean values of $\hat{\alpha}_1$, for both horizontal and vertical polarization, were close to the deterministic case. However, the mean values of $\hat{\alpha}_2$ were generally 5 to 10 dB above the deterministic factor. When the frequency is increased to 100 MHz, more significant variations were observed for horizontal polarization. For higher frequencies (400 MHz), there were very significant deviations for the parameter estimates. Therefore, for higher frequencies, normalizing propagation data with the deterministic factor, 40 dB, can result in a misleading description of the path loss behavior.

h. Table III gives a direct comparison between the means of $\hat{\alpha}_1$ and the means of $\hat{\alpha}_2$ by polarization. In both cases, the deviation from 40 dB is significant.

i. Tables IVa, IVb, IVc IVd and IVe show the selected parameter estimates, correlation, and confidence bounds for the α_1 estimates. The confidence bounds are generally small indicating that the α_1 model yields good estimates. No confidence bounds were obtained for $\hat{\alpha}_2$ because of the limited mathematical tracticability of the model. However, approximate estimates can be obtained.

4. CONCLUSIONS:

In Section 3, a representative data analysis was presented covering a typical low, medium, and high transmit frequency. The conclusions reached for this selected presentation were consistent with those obtained for other frequencies as stated in Section 1. (Other frequencies were not presented because of the magnitude of the results.)

From the techniques proposed in Section 2 and from the results of Section 3, utilizing stochastic models, the following questions were answered:

(a) Is there a significant difference between the results ($\hat{\alpha}_1$ and $\hat{\alpha}_2$) obtained in analysing the path loss when one used either the dB data or the anti-log data?

(b) If a significant variation exists, is it uniform at all frequencies or uniform for different ranges of frequencies?

TABLE IA: PATH LOSS COEFFICIENTS FOR SELECTED CONFIGURATIONS

Frequency (MHz)	Transmitter Antenna Ht.	Polarization	Coefficient	Receiver Antenna Heights-Ft. 11	20	26	31	37	42	48	53	59	64	69	73	79
6.0	40	V	ALPHA 1	37.4	37.7	38.4	38.2	38.8	39.2	38.0	38.4	38.4	37.2	38.1	37.8	38.8
			ALPHA 2	51.	51.0	51.	51.	48.	51.	48.	51.	51.	48.	51.	51.	51.
		H	ALPHA 1	40.1	41.1	40.3	40.7	40.8	41.1	40.8	41.3	41.4	41.0	41.3	40.5	40.8
			ALPHA 2	45.	48.	45.	48.	48.	48.	45.	45.	45.	42.0	45.	42.	45.
	80	H	ALPHA 1	--	--	38.2	39.3	38.9	39.4	39.0	37.8	37.0	37.5	36.9	36.4	36.1
			ALPHA 2	--	--	34.	37.	40.	40.	40.	37.	32.	34.	32.	32.	29.
100	13	V	ALPHA 1	--	--	--	--	32.2	33.2	34.1	34.2	33.4	35.5	36.3	36.6	38.1
			ALPHA 2	--	--	--	--	37.	43.	45.	40.	35.	37.	40.	43.	48.
	40	V	ALPHA 1	31.6	35.5	34.1	34.9	35.9	36.2	37.1	37.9	37.6	37.0	37.8	36.5	36.6
			ALPHA 2	48.	48.	37.	45.	56.	53.	48.	48.	45.	42.	40.	36.	39.
	80	V	ALPHA 1	--	--	38.1	36.9	39.0	40.2	40.0	40.2	40.9	40.9	41.1	40.8	40.5
			ALPHA 2	--	--	48.	43.	54.	54.	45.	45.	37.	35.	35.	35.	32.
100	13	H	ALPHA 1	--	37.8	38.5	38.4	39.6	39.7	39.1	39.2	39.8	38.9	38.8	38.7	39.5
			ALPHA 2	--	56.	48.	51.	57.	56.	45.	42.	45.	42.	42.	37.	39.
	40	H	ALPHA 1	36.0	37.3	35.9	37.6	37.0	33.5	33.5	35.3	34.3	33.1	34.2	34.3	34.8
			ALPHA 2	53.	56.	50.	62.	59.	45.	44.	42.	42.	36.	39.	42.	42.

TABLE IB: PATH LOSS COEFFICIENTS FOR SELECTED CONFIGURATIONS

Frequency (MHz)	Transmitter Antenna Ht. (FT.)	Polarization	Coefficient	Receiver Antenna Heights-Ft.												
				11	20	26	31	37	42	48	53	59	65	69	73	79
100	80	H	ALPHA 1	38.7	39.7	37.5	40.0	38.2	35.8	37.4	37.2	37.8	37.8	38.1	37.6	37.5
			ALPHA 2	48.	42.	32.	45.	39.	29.	31.	29.	29.	26.	26.	26.	24.
400	13	V	ALPHA 1	10.7	6.9	22.7	20.6	20.6	16.3	16.8	25.3	30.0	33.0	27.9	22.8	28.3
			ALPHA 2	9.	4.	21.	18.	18.	14.	12.	21.	24.	29.	22.	15.	20.
	40	V	ALPHA 1	14.8	13.1	17.0	13.3	11.8	17.4	19.1	17.4	21.3	26.3	26.1	28.5	35.5
			ALPHA 2	12.	12.	16.	9.	6.	8.	7.	4.	9.	16.	22.	26.	32.
	80	V	ALPHA 1	---	---	20.5	24.6	25.5	25.1	24.0	24.8	13.5	20.0	30.8	31.8	32.9
			ALPHA 2	---	---	23.	35.	21.	19.	12.	16.	4.	10.	26.	25.	26.
	13	H	ALPHA 1	17.1	16.2	15.4	25.2	12.5	7.4	8.3	20.6	9.5	12.5	29.9	34.2	29.9
			ALPHA 2	20.	18.	16.	27.	8.	2.	---	13.	4.	8.	31.	37.	31.
	40	H	ALPHA 1	---	---	---	---	---	---	---	---	---	97.3	---	97.7	34.4
			ALPHA 2	---	---	---	---	---	---	---	---	---	45.	76.	45.	43.
	80	H	ALPHA 1	---	---	---	---	22.9	30.9	31.6	33.1	31.3	38.9	35.7	35.5	38.7
			ALPHA 2	---	---	---	---	16.	24.	22.	24.	26.	42.	33.	31.	32.

TABLE II: SIGNIFICANCE BETWEEN ALPHA 1 AND ALPHA 2

Frequency (MHz)	Transmitter Antenna Ht. Ft.	Polarization	Overall Mean & Std. Deviation Alpha 1 $\bar{\alpha}_1$	S	Alpha 2 $\bar{\alpha}_2$	S	Test of Significance $\mu_{\alpha_1} = \mu_{\alpha_2}$
6.0	40	V	38.2	0.6	50.3	1.3	R
	40	H	40.9	0.4	45.5	2.1	R
	80	H	37.9	1.2	35.2	3.8	A
100.0	13	V	30.9	1.90	40.9	4.2	R
	40	V	36.1	1.8	45.0	6.0	R
	80	V	40.2	0.90	41.8	7.7	A
	13	H	39.0	0.60	47.7	6.9	R
	40	H	35.2	1.5	46.5	8.3	R
	80	H	38.0	1.1	32.7	8.0	R
400.0	13	V	21.7	7.6	17.5	6.6	A
	40	V	20.1	7.1	13.8	8.4	A
	80	V	24.9	5.7	19.7	8.8	A
	13	H	18.4	8.9	17.9	11.5	A
	40	INDETERMINATE					
	80	H	33.2	4.9	27.7	7.5	A

TABLE III. COMPARISON BETWEEN MEAN ALPHA 1 AND MEAN ALPHA 2 BY POLARIZATION

Frequency (MHz)	Transmitter Antenna Ht. Ft.	Polarization	Overall Mean Alpha 1	Overall Mean Alpha 2	Standard Error Alpha 1	Standard Error Alpha 2
6.0	40	H	40.9	45.5	0.4	2.1
		V	38.2	50.3	0.57	1.3
100	13	H	39.0	47.7	0.6	6.9
		V	30.8	40.9	1.9	4.2
	40	H	35.2	46.6	1.5	8.3
		V	36.1	45.0	1.9	6.0
	80	H	38.0	32.8	1.1	8.0
		V	40.2	41.8	0.9	7.7
400	13	H	18.4	17.9	8.9	11.5
		V	21.7	17.5	7.6	6.6
	40	H	INDETERMINATE			
		V	20.1	13.8	7.1	8.4
	80	H	33.2	27.7	4.9	.5
		V	24.9	19.7	5.7	8.8

TABLE IVA. SELECTED PARAMETER ESTIMATES & CONFIDENCE INTERVALS FOR ALPHA 1

Frequency (MHz)	Transmitter Antenna Ht. Ft.	Polarization	Receiver Antenna Ht. Ft.	Lower 95% Confidence Bound	Alpha 1	Upper 95% Confidence Bound	Standard Error	Correlation (Path Loss Distance)
6	80	H	11	--	--	--	--	--
			20	--	--	--	--	--
			26	33.7	38.2	42.7	2.4	.99
			31	34.3	39.3	44.2	2.6	.99
			37	34.5	38.9	43.4	2.4	.99
			42	34.5	39.4	44.3	2.6	.99
			48	33.8	39.0	44.2	2.7	.99
			53	33.2	37.8	42.5	2.4	.99
			59	31.5	37.0	42.5	2.9	.99
			64	32.7	37.5	42.4	2.5	.99
			69	31.2	36.9	42.5	3.0	.99
			73	30.3	36.4	42.5	3.2	.98
			79	30.2	36.1	41.9	3.1	.99

TABLE IVB. SELECTED PARAMETER ESTIMATES & CONFIDENCE INTERVALS FOR ALPHA 1

Frequency (MHz)	Transmitter Antenna Ht. Ft.	Polarization	Receiver Antenna Ht. Ft.	Lower 95% Confidence Bound	Alpha 1	Upper 95% Confidence Bound	Standard Error	Correlation (Path Loss Distance)
100	80	V	11	--	--	--	--	--
			20	--	--	--	--	--
			26	29.6	38.1	46.6	4.7	.95
			31	28.4	36.9	45.5	4.7	.95
			37	29.8	39.0	48.2	5.0	.95
			42	32.0	40.2	48.6	4.5	.96
			48	31.6	40.0	48.3	4.6	.96
			53	31.5	40.2	48.9	4.8	.96
			59	32.4	40.9	49.4	4.7	.96
			64	32.7	40.9	49.1	4.5	.97
			69	33.1	41.1	49.1	4.4	.97
			73	32.8	40.8	48.9	4.4	.97
			69	32.5	40.5	48.6	4.4	.97

TABLE IVC. SELECTED PARAMETER ESTIMATES & CONFIDENCE INTERVALS FOR ALPHA 1

Frequency (MHz)	Transmitter Antenna Ht. Ft.	Polarization	Receiver Antenna Ht. Ft.	Lower 95% Confidence Bound	Alpha 1	Upper 95% Confidence Bound	Standard Error	Correlation (Path Loss Distance)
100	80	H	11	29.7	38.7	47.8	4.9	.96
			20	31.5	39.7	47.9	4.4	.97
			26	28.7	37.5	46.3	4.7	.97
			31	31.0	40.0	48.9	4.8	.96
			37	27.8	38.2	48.7	5.6	.95
			42	23.2	35.8	46.4	6.8	.94
			48	25.5	37.4	49.3	6.4	.95
			53	26.0	37.1	48.3	6.0	.95
			59	26.6	37.8	48.9	6.0	.95
			64	25.6	37.8	49.9	6.5	.95
			69	26.1	38.1	50.1	6.5	.95
			73	25.2	37.6	50.0	6.7	.95
			79	24.6	37.5	50.5	7.0	.95

TABLE IVD. SELECTED PARAMETER ESTIMATES AND CONFIDENCE INTERVALS FOR ALPHA 1

Frequency (MHz)	Transmitter Antenna Ht. Ft.	Polarization	Receiver Antenna Ht. Ft.	Lower 95% Confidence Bound	Alpha 1	Upper 95% Confidence Bound	Standard Error	Correlation (Path Loss Distance)
400	80	V	11	--	--	--	--	--
			20	--	--	--	--	--
			26	11.3	20.5	29.8	4.6	.92
			31	16.8	24.6	32.4	3.9	.97
			37	20.6	25.5	30.3	2.4	.99
			42	19.4	25.1	30.9	2.9	.99
			48	12.5	24.0	35.5	5.7	.96
			53	10.6	24.8	39.0	7.0	.94
			59	-8.05	13.5	35.12	10.7	.79
			64	-0.21	20.2	40.3	10.0	.86
			69	12.0	30.8	49.7	9.4	.92
			73	13.3	31.8	50.4	9.2	.93
			79	14.2	32.9	51.6	9.3	.93

TABLE IVE. SELECTED PARAMETER ESTIMATES AND CONFIDENCE INTERVALS FOR ALPHA 1

Frequency (MHz)	Transmitter Antenna Ht. Ft.	Polarization	Receiver Antenna Ht. Ft.	Lower 95% Confidence Bound	Alpha 1	Upper 95% Confidence Bound	Standard Error	Correlation (Path Loss Distance)
400	80	H	11	--	--	--	--	--
			20	--	--	--	--	--
			26	--	--	--	--	--
			31	--	--	--	--	--
			37	13.4	22.9	32.4	4.7	0.96
			42	22.7	30.9	39.1	4.1	.98
			48	24.9	31.6	38.3	3.3	.99
			53	26.5	33.1	39.7	3.3	.99
			39	22.7	31.3	39.9	4.3	.98
			64	28.3	38.9	49.4	5.2	.98
			69	23.3	35.7	48.0	6.1	.97
			73	22.9	35.5	48.0	6.2	.97
			79	22.2	38.7	55.2	8.2	.96

Tables IA, IB, II and III show that $\hat{\alpha}_2$ was generally within 13 dB of the corresponding $\hat{\alpha}_1$. Table II shows the decisions resulting from the hypothesis that $\mu_{\alpha_1} = \mu_{\alpha_2}$, with regard to the overall means. In one-half the configurations tested, this hypothesis was rejected. These decisions were consistent using both parametric and non-parametric (distribution free) approaches. The differences between $\hat{\alpha}_1$, and $\hat{\alpha}_2$ are not uniform for the different frequencies involved, and vary considerably for each receiver antenna height measured.

The results show conclusively, that the deterministic view of the jungle environment is not reasonable, especially at higher frequencies. The wide variation of $\hat{\alpha}_1$, and $\hat{\alpha}_2$ over all configurations tested shows that one should view path loss as a stochastic realization.

Both models depicted yield reasonable results, but there are indications that for larger distances and higher frequencies, the α_2 model (anti-log) is a better approach. A logical conclusion, however, is that the model for α_1, is reasonable for lower frequencies and shorter distances considering the small confidence intervals over all receiver antenna heights at the lower frequencies. Intuitively, α_2, since it does not rely on logarithmic transformation, should be more effective especially at higher frequencies and longer distances. A decision as to which model, α_1, or α_2, is best will rely on future work in the subject area.

****

REFERENCES

(1) Acker, M., R.D'Accardi, D.Dence, and C.P. Tsokos, "Statistical Modeling of Propagation Path Loss Data", Proc. 18th Conference on the Design of Experiments in Army Research Development and Testing, Oct. 1972.

(2) Dence, D., and T. Tamir, "Radio Loss of Lateral Waves in Forest Environments", Radio Science, 4 (1969), 307-318.

(3) Hicks,Et al, "Tropical Propagation Research Final Report, Vol. II", Jansky and Bailey Engineering Department, Atlantic Research Corporation, Contract DA36-039, SC-90889, (1969).

(4) Jansky and Bailey Engineering Department, Atlantic Research Corporation. "Tropical Propagation Research Final Report, Vol. I", Contract DA36-039, SC-90889, (1966).

(5) Robertson, R.G., J.J. Hicks, C.B. Sykes, P.A. Anti, "Tropical Propagation Research Final Report Vol. III", Jansky and Bailey Engineering Department, Atlantic Research Corporation, Contract DA36-039, SC-90889, (1971).

(6) Tamir, T., "On Radio Wave Propagation in Forest Environments", IEEE Trans. AP-15, (1967), 806-817.

AN APPROACH TO OCCUPANT EVALUATION OF ARMY FAMILY HOUSING INTERIORS

Roger Brauer and Robert Neathammer
U. S. Army Construction Engineering Research Laboratory
Champaign, Illinois 61820

INTRODUCTION. In architectural research it is very difficult to achieve an experimental arrangement where the behavioral responses of building users can be related to architectural variables. The difficulty lies with both the control of architectural variables and the measurement of behavioral responses by means of observation methods or attitude assessment.

Architectural variables are difficult to control for, simply because people find it difficult to justify the expense of constructing buildings for experimental reasons. As a result, differences in building features must be identified in buildings which already exist. The limitation here is that differences may not extend over the range desired. New features may simply not yet exist in available buildings. Furthermore, features of interest may be confounded by the presence or absence of other features or by varying rules in management of the spaces.

Behavioral responses to architectural variables could be obtained through simulation or models. This would work reasonably well for appearance and visual characteristics. However, responses to functional considerations gain validity through "live-in" experience.

In general architectual psychology research is interested in relating attitudes and behaviors to building design characteristics. The attitudes and behaviors are not simply functions of design features, but are also influenced by other factors. A schematic representation of this concept is presented in Figure 1. The degree of satisfaction an individual expresses about the characteristics of his dwelling or a part of the dwelling, such as a bedroom, are directly influenced by many factors as well as mediated by intervening factors.

Therefore, experiments relating responses of people to built environments must be designed to control for some of the extraneous variables as well as those of specific interest.

FAMILY HOUSING

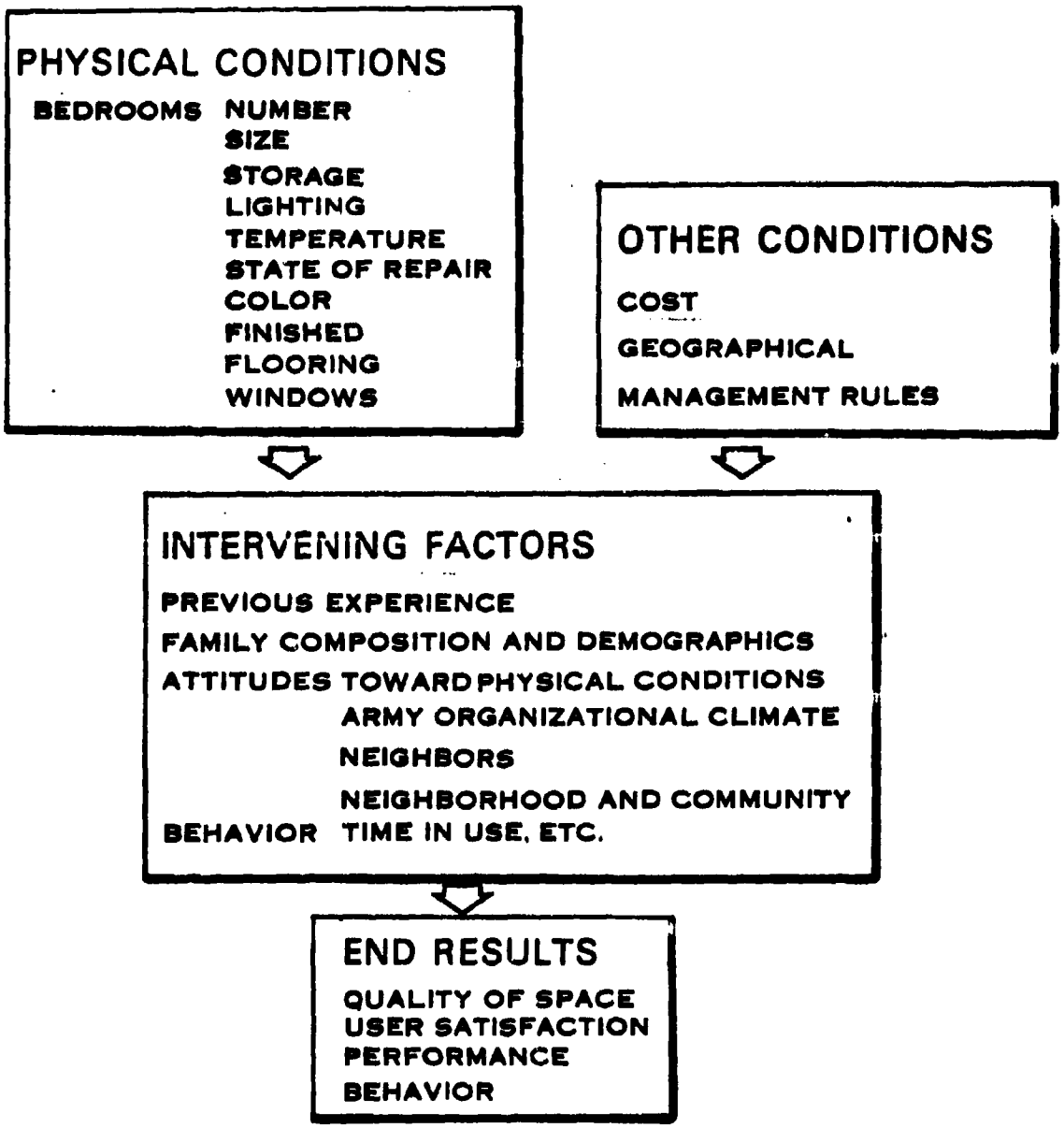


Figure 1. Factors influencing quality and occupant satisfaction.

The Experimental Problem

It is a well known fact that the quality of military family housing is an irritant to the soldier. Recently, Senator Stennis, who is Chairman of the Senate Armed Services Committee, reported in a news release that his interviews with military personnel revealed housing to be the greatest deterrent to reenlistment. The task which CERL was given was to evaluate Army housing from the occupant's point of view, with major emphasis on interiors. The main goal was to relate the attitudes and preferences of occupants to architectural variables. The study was to focus on housing that had been occupied for the first time in the last three years, but older housing was to be considered as well.

How the Experiment was Designed

As noted earlier, it was known that the attitudes and opinions would be influenced by many things. Since the new homes were to be found at twelve different installations, it was apparent that installations could be a source of variation (from differences in geographical location and from other factors).. Furthermore, the quality of construction workmanship was known to be a source of variation, even when homes were identical or very similar in plans and specifications. Another obvious source of variation was differences in floor plans and finish materials. In addition, variation could also result from differences in family composition, rank, age, possessions, motivation for providing the information requested, and countless other things. Because of time and funding constraints, responses of occupants had to be obtained through the use of survey methods. However, a rather precise questionnaire was developed for documenting four groups of information: demographic information, number of major possessions, preferences about design features, and attitudes toward present quarters.

In order to isolate some of the variation in the architectural characteristics, a nested design was selected for the experiment and for establishing a sampling procedure. As seen in Figure 2, the first level in the design was installation. Installation was selected as the first level to isolate much of the variation due to geographic and level differences.

The second level selected was construction project, identified by a year of appropriation. Thirty-five construction projects were included at this level. The number of homes in each housing project, appropriated from 1955 through 1971, ranged widely from 40 units to 1,000 units. At this level, it was anticipated that much of the variation due to the age of the quarters as well as the quality of construction could be isolated.

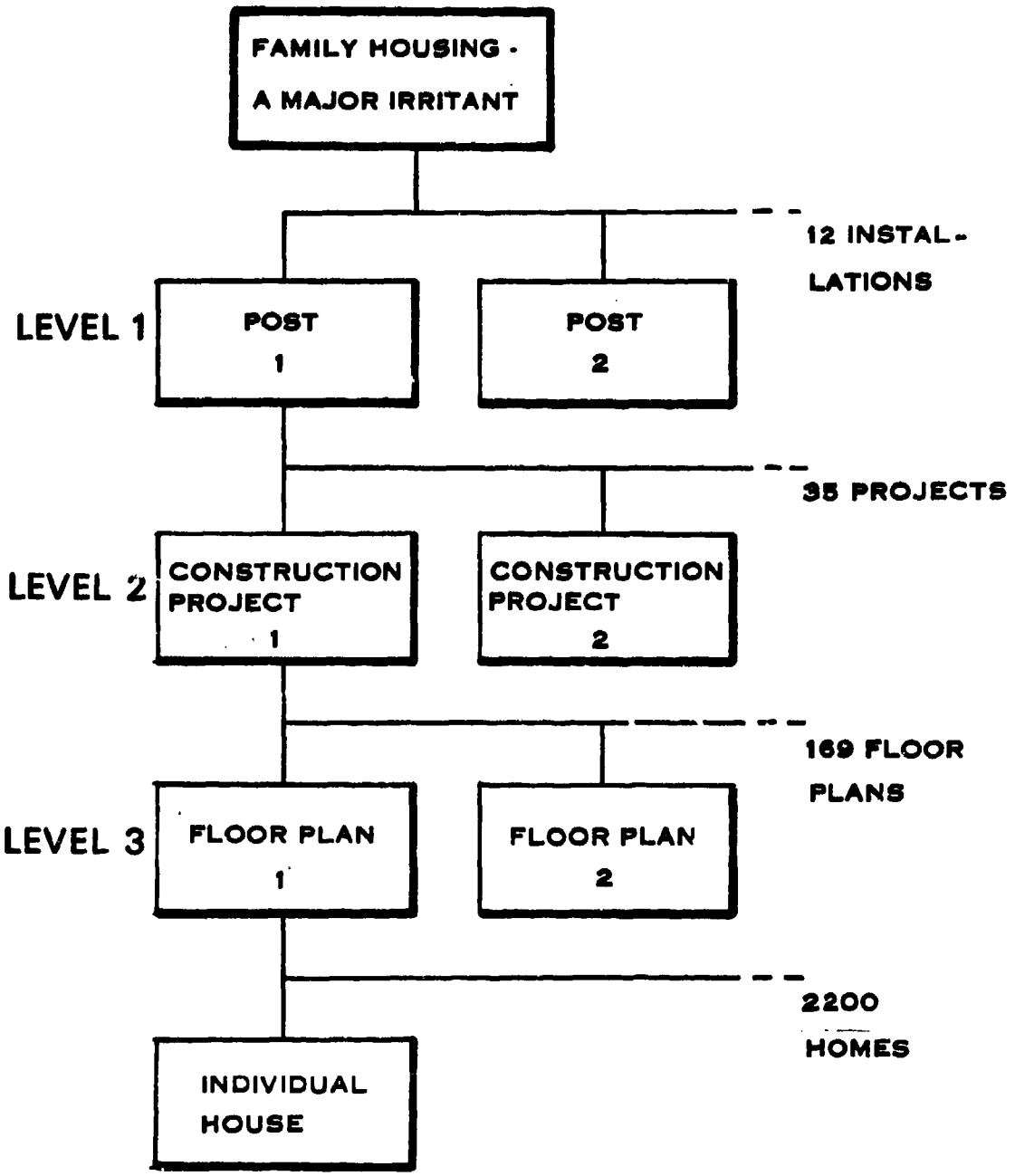


Figure 2. Schematic of nested design for house study.

The third and last level in the nested design was the floor plan. About 170 different floor plans were included in the study in an attempt to isolate much of the variation resulting from bedroom count, total living space, spatial characteristics and other features.

Being limited by the amount of funds available for the project, sample sizes were established at the second level, that of construction project. Sample sizes were determined so that 95% confidence bands on a sample percentage of 50% would be ≠ 10% in width. The sample was distributed at the floor plan level according to the percent of housing units of each floor plan type within the construction project. As a result, a total of about 2,300 homes were included in the study, representing a population of about 9,000 homes.

From previous studies, it was known that many housing occupants would not be motivated to participate in the study. The introduction of a high degree of personal contact was used to ensure cooperation. Questionnaires were delivered in person and picked up in person. When questionnaires were picked up, time was spent with respondents who wished to discuss good and bad features of their housing. With voluntary cooperation of respondents, field workers made precise photographic records of general conditions, as well as records of good and bad design features. Through the use of personal contact, a 95% participation rate was achieved and over 3,000 photographs in support of statistical evidence were collected.

In addition to occupant response, a detailed record was made of the size, materials of construction, and other characteristics of each floor plan, so that relationships between occupant responses and architectural variables could be made effectively.

Additional analyses can be achieved from the nested design for the study. Since demographics are not independent of housing size, bedroom count and possibly other characteristics of floor plans, the data can be rearranged according to rank, family composition, housing density and similar factors to achieve other results about responses. Furthermore, the nested design will permit us to relate housing characteristics to other dependant variables of a sociological nature in future analysis, such as theft rate, mental illness, divorce, and reenlistment.

Major project issues which this project was designed to address and the analyses planned to assist in resolving these issues are presented in Table 1.

There have been several previous surveys conducted to obtain attitudes, opinions, and preferences of occupants of Army family housing. This study

was designed not only to collect such responses but, more importantly, to relate them to housing characteristics, family characteristics and possessions, the things to which people are responding. This step forward in the planning of such studies should provide more useful information for housing designers and planners.

Table 1

Major Project Issues and Corresponding Analyses

PROJECT ISSUE	ANALYSES PLANNED
What are the main opinions, preferences, demographics and possessions of housing occupants?	1 way tabulations, 2 and 3 way cross tabulations
How do family and housing characteristics contribute to preferences?	Cross tabulations, correlations, and multivariate methods.
How can housing requirements (bedroom count) be predicted from demographics?	Cross tabulations, regression
How are space requirements in family housing related to family characteristics and attitudes?	Cross tabulations, regression
Can overall satisfaction with housing be predicted from satisfaction with housing elements and housing characteristics?	Multiple regression and AID (Automatic Interaction Detection)

GENERALIZED JACKKNIFE TECHNIQUES

H. L. Gray
Southern Methodist University
Dallas, Texas

ABSTRACT. In this paper the jackknife method is reviewed and some of its applications are considered. In particular its use as a point and interval estimator for data from a random sample or a stochastic process with continuous index set is discussed. Finally some results, implied by the generalized jackknife statistic, concerning unique minimum unbiased estimators are demonstrated.

I. INTRODUCTION. The jackknife statistic, $J(\hat{\theta})$, was first introduced by Quenouille (1956) in [6]. It is obtained as follows. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n and let $\hat{\theta}_n$ be an estimator for θ. Now partition the sample into N subsets of size M and define $\hat{\theta}^i$ as the estimator $\hat{\theta}$ restricted to a subsample obtained by deleting the i-th subset from the data (we of course assume $\hat{\theta}^i$ is definable on such a subsample). Then letting

$$J_i(\hat{\theta}) = N\hat{\theta} - (N-1)\hat{\theta}^i, \quad i = 1, \ldots, N, \tag{1}$$

the jackknife estimator is defined by

$$J(\hat{\theta}) = \frac{1}{N}\sum_{i=1}^{N} J_i(\hat{\theta})$$

$$= N\hat{\theta} - (N-1)\overline{\hat{\theta}^i} \quad . \tag{2}$$

Although it is not necessary to let M = 1, when possible it is usually desirable to do so. Thus we will henceforth make that assumption, i.e.

$$J_i(\hat{\theta}) = n\hat{\theta} - (n-1)\hat{\theta}^i, \quad i = 1, \ldots, n \tag{3}$$

and

$$J(\hat{\theta}) = n\hat{\theta} - (n-1)\overline{\hat{\theta}^i}. \tag{4}$$

The statistic $J(\hat{\theta})$ was introduced by Quenouille because of its bias reduction property. Namely, if

$$E[\hat{\theta}] = \theta + \frac{c_1(\theta)}{n} + \frac{c_2(\theta)}{n^2} + \ldots \quad . \tag{5}$$

then

$$E[J(\hat{\theta})] = \theta + \frac{d_2(\theta)}{n^2} + \frac{d_3(\theta)}{n^3} + \ldots \quad . \tag{6}$$

However, in 1958 J. Tukey [9] suggested the statistics in (3), which were later called pseudo values, could be treated as a random sample and hence furnish a general method for obtaining an approximate confidence interval for θ. This suggestion has evoked much research and the method has been found to be a useful tool in such problems.

In 1971 Schucany, Gray and Owen [8] significantly expanded the jackknife method by introducing the generalized jackknife and higher order generalized jackknife. Their definition was as follows.

Definition 1.

Let $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{k+1}$ be $k+1$ estimators for θ defined on the random sample $X_1, X_2, \ldots, X_n$. Further let a_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, k+1$ be real numbers and

$$D(\hat{\theta}_i; a_{ij}) = \begin{vmatrix} \hat{\theta}_1 & \hat{\theta}_2 \cdots & \hat{\theta}_{k+1} \\ a_{11} & a_{12} \cdots & a_{1,k+1} \\ \vdots & & \vdots \\ a_{k1} & a_{k2} \cdots & a_{k,k+1} \end{vmatrix} \tag{7}$$

Then the generalized jackknife, $G(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{k+1})$, is defined by

$$G(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{k+1}) = \frac{D(\hat{\theta}_i; a_{ij})}{D(1; a_{ij})}, \tag{8}$$

where it is assumed $D(1; a_{ij}) \neq 0$.

In general there is no restriction on the manner in which $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{k+1}$ in (7) can be selected, however they are usually chosen in the same manner as indicated in (2). That is for a given estimator $\hat{\theta}$ and a random sample of size n

$$\left.\begin{aligned} \hat{\theta}_1 &= \hat{\theta}(X_1, \ldots, X_n) \\ \hat{\theta}_2 &= \frac{1}{n} \sum_{i=1}^{n} \hat{\theta}^i(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n) = \overline{\hat{\theta}^i} \\ \hat{\theta}_3 &= \frac{1}{\binom{n}{2}} \sum_{i,j} \hat{\theta}^{i,j}(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n) = \overline{\hat{\theta}^{ij}}, \end{aligned}\right\} \tag{9}$$

etc.

The following theorem holds regardless of the method of selecting the $\hat{\theta}_i$. This result was first shown in [8].

Theorem 1. If

$$E[\hat{\theta}_j] = \theta + \sum_{i=1}^{k} f_{ij}(n)\, b_i(\theta), \quad j = 1, 2, \ldots, k+1$$

and if $a_{ij} = f_{ij}(n)$, then

$$E[G(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{k+1})] = \theta. \tag{10}$$

If the generalized jackknife is to be utilized as an estimator, the random variable $G(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{k+1})$ must not depend on θ. When

$$E[\hat{\theta}_1] = \theta + \sum_{i=1}^{\infty} f_i(n)\, b_i(\theta) \tag{11}$$

and $\hat{\theta}_2, \ldots \hat{\theta}_{k+1}$ are selected by the procedure described in (9), letting $a_{ij} = f_i(n-j+1)$ satisfies that requirement as well as Theorem 1 when the series in (11) is finite. For this reason the a_{ij} and the $\hat{\theta}_i$ in Definition 1 are usually chosen accordingly. When this is the case we write $G^{(k)}(\hat{\theta})$ in place of (8), i.e.

$$G^{(k)}(\hat{\theta}) = G(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{k+1}), \tag{12}$$

where the $\hat{\theta}_i$ are defined in (9) and $a_{ij} = f_i(n-j+1)$. One should note that

$$G^{(1)}(\hat{\theta}) \equiv G(\hat{\theta}) = \frac{\begin{vmatrix} \hat{\theta} & \overline{\hat{\theta}^i} \\ a_{11} & a_{12} \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ a_{11} & a_{12} \end{vmatrix}} = \frac{\hat{\theta} - R(n)\,\overline{\hat{\theta}^i}}{1 - R(n)}, \tag{13}$$

where $R(n) = a_{11}/a_{12}$ and hence $J(\hat{\theta}) = G(\hat{\theta})$ when $a_{11} = n^{-1}$, $a_{12} = (n-1)^{-1}$. The form (13) suggests an appropriate definition for the pseudo values, $G_i(\hat{\theta})$, associated with $G(\hat{\theta})$. That is,

$$G_i(\hat{\theta}) = \frac{\hat{\theta} - R(n)\hat{\theta}^i}{1 - R(n)}. \tag{14}$$

The following theorem which was first shown in [3] (1972) is the basis for the jackknife method as a technique for approximate confidence intervals.

Theorem 2. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with mean μ and finite variance σ^2. Let

$$\begin{aligned} \theta &= f(\mu) \\ \bar{X}_n &= \frac{1}{n} \sum_{i=1}^{n} X_i \\ \hat{\theta} &= f(\bar{X}_n), \end{aligned} \tag{15}$$

where f is a real-valued function, defined on the real line, which possesses a bounded second derivative in a neighborhood of μ. Further

suppose

$$a = \lim_{n \to \infty} \frac{R(n)}{(1-R(n))\ (n-1)} \neq \pm \infty,\ 0. \tag{16}$$

Then the random variable

$$\frac{a(G(\hat{\theta})-\theta)\ \sqrt{n(n-1)}}{\sqrt{\sum_{i=1}^{n} (G_i(\hat{\theta}) - G(\hat{\theta}))^2}} \to N(0,1) \tag{17}$$

as $n \to \infty$. That is, the random variable in (17) is asymptotically distributed as a normal random variable with zero mean and unit variance. As was pointed out in [3] this theorem can be extended to the case where $\hat{\theta}$ is a function of a U-statistic.

A final modification of the jackknife should be mentioned before proceeding to the application, namely its extension to stochastic processes. By considering an appropriate limiting case of (12), Gray, Watkins and Adams in [5] (1972) extended the notions above for k = 1,2 to stochastic processes. We mention briefly the nature of those results. The following definition is the counterpart of (4) in that development.

Definition 2.

Let $\{X(t) \mid t \in S\}$ be a stochastic process defined over an index set S containing the interval [a,b]. Let $\{I_X(t) \mid t \in [a,b]\}$ be a piecewise continuous stochastic process which is completely determined by the process $\{X(t) \mid t \in [a,b]\}$. Let N_γ be the random variable defined by the number of discontinuities of I_X of size γ, observed on the interval [a,b], and let

$$\hat{\theta} = \hat{\theta}(a,b) = \frac{I_X(b) - I_X(a)}{T}, \quad T = b - a.$$

Then for any real valued function f, differentiable over the range of $\hat{\theta}$, we define the J_∞ estimator by the following equation:

$$J_\infty[f(\hat{\theta})] = f(\hat{\theta}) - \sum_{\gamma \in \Gamma} N_\gamma [f(\hat{\theta} - \frac{\gamma}{T}) - f(\hat{\theta}) + \frac{\gamma}{T} f'(\hat{\theta})],$$

where Γ is the set of all possible values of γ,

$$f'(\hat{\theta}) = \left.\frac{df(\theta)}{d\theta}\right|_{\theta = \hat{\theta}}$$

and by the notation $\sum_{\gamma \in \Gamma}$ we mean that for any realization x, the sum is to be taken over all γ for which N_γ has an observed value different from zero on [a,b].

By making use of Definition 2 it is possible to get an extension corresponding to (13). Moreover theorems corresponding to Theorems 1 and

2 can also be obtained so that a normal r.v.like (17) can be used to obtain approximate confidence intervals for $f(\theta)$. The details of this procedure and further results can be found in [3] (1972). We state the following theorem for later use. It is the result for stochastic processes which corresponds to Theorem 2.

Theorem 3. Let $\{I_x(t) \mid t \in [a,\infty)\}$ be a stochastic processes with stationary independent increments such that $E[\hat{\theta}(a,t)] = \theta$ for each $t \in [a,\infty)$ and suppose that f has bounded second derivative in a neighborhood of θ. Then under mild regularity conditions

$$\frac{J_\infty[f(\hat{\theta})] - f(\theta)}{\left\{ \sum_{\gamma \in \Gamma} N_\gamma [f(\hat{\theta}-\frac{\gamma}{T}) - f(\hat{\theta})]^2 \right\}^{1/2}} \to N(0,1) \qquad (18)$$

as $T \to \infty$, i.e., the random variable in (18) is asymptotically normal with zero mean and unit variance.

II. APPLICATIONS. In this section we consider some of the standard applications of the jackknife which are suggested by the above theorems. The first of these was initially considered by Mosteller and Tukey in [7] (1968) and in more detail by Gray and Schucany in [3] (1972).

Example 1. Suppose the following is a random sample of size eleven from an unknown distribution, $x_1 = 0.1$, $x_2 = 1.0$, $x_3 = 1.3$, $x_4 = 1.9$, $x_5 = 0.4$, $x_6 = 0.1$, $x_7 = 4.7$, $x_8 = 0.5$, $x_9 = 1.9$, $x_{10} = 1.1$, $x_{11} = 0.1$, and we desire a confidence interval for σ. Moreover let us assume that the statistic we decide to use is

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2} = \sqrt{S^2} = f(S^2) \; . \qquad (19)$$

Since S^2 is an unbiased U-statistics we can use the extension of Theorem 2 to U-statistics which was mentioned earlier. Since we assume no specific knowledge of the bias we make the mild assumption that the bias in S is an analytic function of n^{-1}. This leads us to define $R(n) = (n-1)n^{-1}$ so that $G(\hat{\theta}) = J(\hat{\theta}) = J[f(S^2)] = J(S)$ and $\alpha = 1$. We then obtain from (17) (essentially) the approximate confidence interval (a,b), where

$$a = J(S) - t_{\frac{\alpha}{2}}(n-1) \frac{S_J}{\sqrt{n}} \; ,$$

$$b = J(S) + t_{\frac{\alpha}{2}}(n-1) \frac{S_J}{\sqrt{n}} \; , \qquad (20)$$

$$S_J = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (J_i(S) - J(S))^2} \; .$$

and $t_{\frac{\alpha}{2}}(n-1)$ is the $100(1-\frac{\alpha}{2})$ percentile point from the Student t-distribution with n-1 degrees of freedom. Thus for these data

$$J_1(S) = J_6(S) = J_{11}(S) = 1.1400$$
$$J_2(S) = .6325$$
$$J_3(S) = .6219$$
$$J_4(S) = J_9(S) = .8355$$
$$J_5(S) = .8894$$
$$J_7(S) = 7.7040$$
$$J_8(S) = .8243$$
$$J_{10}(S) = .6204$$

and

$$J(S). = \frac{1}{11} \sum_{i=1}^{11} J_i(S) = 1.4894 \qquad (21)$$

$$S_J/\sqrt{11} = 0.6244. \qquad (22)$$

From (21) and (22), we obtain

TABLE 1

Confidence Level 1-α	Interval
2/3	(0.85, 2.13)
0.95	(0.10, 2.88)

The extrapolation from Theorem 2 to the modification suggested by (20) is employed for the following reason. Experience has indicated that in most cases the lack of normality of (17) for finite samples usually leads to confidence intervals which undercover. Since the Student-t is also asymptotically normal, Theorem 2 is also a justification for assuming the random variable in (17) is approximately t for large n. When this assumption is made the net effect is simply to slightly enlarge the confidence interval which in most instances will improve the approximation.

The question naturally arises as to how good are such approximate confidence intervals as those of Table 1 and what additional modifications might improve them. This question was considered in some detail for this and other examples in [3]. Several modifications are suggested there. We will consider two of them. The first is to jackknife log S rather than S and obtain a confidence interval for log σ which is then converted to a confidence interval for σ.

The basis for this approach is the assumption that log S possesses a more symmetrical distribution than S and hence the convergence to normality of J(log S) should be more rapid than that of J(S). A second modification is to use $G(\hat{e})$ with

$$R(n) = \left(\frac{n-1}{n}\right)^{1/2} \tag{23}$$

The reasoning behind such an alteration as this would of course be the belief that the bias was closer to order $n^{-1/2}$ than n^{-1}. Table 2 below shows the results of a Monte Carlo study when the unknown sample is exponential with $\sigma = 1$. The purpose is to compare these three methods and the normal method, i.e., the method of obtaining a confidence interval by assuming $(n-1)\ S^2/\sigma^2$ is $\chi^2(n-1)$. The results are accurate to within 3% of the true confidence level associated with each of the statistics.

TABLE 2

Approximate 90% Confidence Intervals, n = 11

Source of Interval	Estimated True Confidence Level
$(n-1)S^2/\sigma^2$	67.6
J(S)	72.0
G(S)	73.6
J(log S)	80.6

In this case it is clear that the lack of normality of any of our statistics was certainly a major source of error in our approximation and that jackknifing log S seemed to improve that situation. This of course will not always be the case but another example of where it was beneficial is the following which was first given by Arvensen and Schmitz in [2] (1970).

Example 2. Consider now the following elementary variance components model:

$$Y_{ij} = \mu + a_i + e_{ij},\ i = 1,\dots I, j = 1, \dots M,$$

where μ is constant; $\{a_i\}$ are independent zero mean, σ_A^2 random variables; and the $\{e_{ij}\}$ are independent, mean zero, variance σ_e^2 random variables. The test of hypothesis to be considered is

$$H_o: \eta = \sigma_A^2 / \sigma_e^2 \leq \eta_o$$

vs.

$$H_1: \eta > \eta_o. \tag{24}$$

The usual procedure here is to assume normality and utilize the F-statistic to construct the required test. However it is well known that the F-test is quite sensitive to departures from normality and therefore other tests have been proposed. To employ the jackknife here we need a bit more theory than we have thus far suggested. That is, theorems, similar to those we have stated, for the jackknife of a function of several U-statistics are needed. Such results have been obtained and can be found in [1] (1969). Thus we may assume both $J(\hat{\theta})$ and $J(\log\ \hat{\theta})$ are approximately normal, where

$$\hat{\theta} = \frac{MSA}{MSE} \tag{25}$$

and MSA and MSE denote the between and within group mean squares respectively. The jackknife is then obtained by grouping the n= IM observations $\{Y_{ij}\}$ into natural groups of size M, i.e. by successively deleting the I

groups of observations $\{(Y_{i1},Y_{i2},\ldots,Y_{iM}),\ i = 1,2,\ldots,I\}$ in the computation of $J_i(\hat{\theta})$ and $J_i(\log \hat{\theta})$. The approximate t-statistic (under H_o),

$$\frac{J(\hat{\theta}) - \eta_o}{s_{J(\hat{\theta})}}\sqrt{n} \text{ or } \frac{J(\log \hat{\theta}) - \log \eta_o}{s_{J(\log \hat{\theta})}}\sqrt{n}\ , \qquad (26)$$

depending on whether or not we have decided to use $\hat{\theta}$ or $\log \hat{\theta}$ as our statistic to be jackknifed, is then used to test the appropriate hypothesis. This test was compared to the F-test by Arvensen in [2] (1970). In that study $\mu = 0$, $\eta_o = 1$, $I = 15$ and $M = 3$. The results are shown in Table 3 below for significance level $\alpha = .10$.

It is interesting to note in this table how well the method employing $J(\log \hat{\theta})$ agrees with the F-test on normal data where the F is of course exact. Also the nonrobust nature of the F with respect to the significance level is well demonstrated in the case of the double expenential.

TABLE 3

Probability of Rejecting $\eta \leq 1$

	$\eta = \sigma_A^2/\sigma_e^2$	1	1.5	2.5	4	6
Distribution						
Normal	F test (theoretical)	0.10	0.29	0.65	0.90	0.98
	F test (empirical)	0.098	0.269	0.655	0.900	0.979
	Jackknife $J(\log \hat{\theta})$	0.090	0.257	0.609	0.881	0.972
	Jackknife $J(\hat{\theta})$	0.022	0.080	0.263	0.469	0.700
Double Exponential	F test	0.183	0.334	0.567	0.764	0.887
	Jackknife $J(\log \hat{\theta})$	0.096	0.218	0.420	0.616	0.767
Uniform	F test	0.085	0.253	0.690	0.951	0.994
	Jackknife $J(\log \hat{\theta})$	0.093	0.308	0.733	0.947	0.993

Example 3. Use Theorem 3 to construct an approximate confidence interval, based on a fixed time test, for the reliability function

$$r(\theta,x) = e^{-\theta x} \qquad (27)$$

associated with the Poisson process $\{N(t) \mid t \in [o,\infty)\}$. In this case let

$$\hat{\theta} = \frac{N}{T} \qquad (28)$$

and

$$f(\hat{\theta}) = e^{-\hat{\theta}x} = r(\hat{\theta},x),$$

where N is the number of failures occurring in time T. In this event $\Gamma = \{1\}$ and we have

$$J_\infty(f(\hat{\theta})) = f(\hat{\theta}) - \hat{\theta} f(\hat{\theta}) - N[f(\hat{\theta}-\tfrac{1}{T}) - f(\hat{\theta})]$$

$$= e^{-\hat{\theta}x}\left\{1-N(e^{\frac{x}{T}}-1+\tfrac{1}{T})\right\}. \tag{29}$$

It can easily be shown (see [3]) that in this case $J_\infty(f(\hat{\theta}))$ has a lower order bias, as a function of T, than $f(\hat{\theta})$. Moreover Theorem 3 can be applied here with

$$\sum_{\gamma\in\Gamma} N_\gamma \left[f(\hat{\theta}-\tfrac{\gamma}{T}) - f(\hat{\theta})\right]^2 = N e^{-2\hat{\theta}x}(e^{x/T}-1)^2 \triangleq S_T. \tag{30}$$

Thus from the same reasoning as our previous examples an approximate confidence interval for $r(\theta,x)$ is

$$\left[J_\infty(r(\hat{\theta},x)) - S_T\, t_{\alpha/2},\ J_\infty r(\hat{c},x) + S_T\, t_{\alpha/2}\right] \tag{31}$$

where $t_{\alpha/2}$ is the $100(1-\alpha/2)$% point from normal distribution.

In order to exemplify this result the following Monte Carlo studies were made. Random Poisson numbers with known parameters were generated and the corresponding statistics substituted into (31). In Table 4, the columns labeled P_1, P_2 and P_3 give the percent of 1000 samples generated for which $r(\theta,x)$ was contained in the confidence interval obtained from (31) for each of the three α-levels shown.

TABLE 4

θ	x	T	P_1 $1-\alpha = 0.50$	P_2 $1-\alpha = 0.70$	P_3 $1-\alpha = 0.95$
2	0.05	0.5	35.9	54.5	61.0
2	0.05	2.5	44.6	59.1	86.3
2	0.05	5.0	49.0	65.0	92.0
5	0.02	0.2	36.1	55.6	63.4
5	0.02	1.0	44.2	63.8	86.8
5	0.02	2.0	47.8	63.8	91.4
10	0.01	0.1	37.9	54.4	62.1
10	0.01	0.5	50.8	62.7	86.5
10	0.01	1.0	46.3	64.7	91.5

Inspection of the above table suggests that use of the percentile points from the Student-t distribution as in our previous example would improve the approximation even though, when T is sufficiently large for

the expected number of failures, θT, to be near 10, the approximation is quite good.

In the above examples we have found no great value in $G(\hat{\theta})$ over $J(\hat{\theta})$. We therefore give one final example which clearly demonstrates the wider utility of $G(\hat{\theta})$. Since this example is rather lengthy and has only appeared this month in the literature we set it apart as a separate section.

III. UMVU ESTIMATORS FOR $f(\mu,\sigma^2)$ IN THE NORMAL CASE. Before we proceed to the primary result of this final section let us consider the following examples.

Example 4. Consider the differential equation

$$y'' + k^2 y = 0.$$

Then

$$y = C_1 \operatorname{Sin} kt + C_2 \operatorname{Cos} kt \tag{32}$$

Now in many physical problems k is not precisely known and in fact it might be better to refer to it as $\bar{K}$, where $\bar{K}$ is a random variable. In this sense $\bar{k}$ is an approximation to $\mu = E[\bar{K}]$ and $y = C_1 \operatorname{Sin} \bar{k}t + C_2 \operatorname{Cos} \bar{k}t$ is an approximation to $y_\mu = C_1 \operatorname{Sin} \mu t + C_2 \operatorname{Cos} \mu t$. If this latter solution is the solution of interest and a sample $k_1, k_2, \ldots, k_n$ of values of K are obtained which can reasonably be assumed to be from a normal population with unknown mean and variance, then the problem is one of estimating a function of μ (i.e. y_μ) from normal samples.

Example 5. Let $x \sim N(\mu,\sigma^2)$. Estimate

$$P[X \leq k] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\frac{k-\mu}{\sigma}} e^{-1/2 t^2} dt = \phi\left(\frac{k-\mu}{\sigma}\right) \tag{33}$$

from a random sample of size n. The problem is again one of estimating a function of μ and σ^2 from normal data.

Example 6. Suppose x is a random variable such that $\ln x \sim N(\mu,\sigma^2)$, i.e. $x \sim$ lognormal.

Then

$$E[x] = e^{\mu + \frac{\sigma^2}{2}}$$

$$\operatorname{Var} x = e^{2\mu+\sigma^2}(e^{\sigma^2}-1). \tag{34}$$

Then estimating the mean and variance of a lognormal distribution can be formulated as a problem of estimating a function of μ and σ^2 from a normal population.

Example 7. Suppose $x \sim B(n,p)$ (Binomial). Then for large n the variance of arc sin $\sqrt{x/n}$ is essentially independent of p. A transformation, T, defined on a random variable x such that the variance of T(x) is independent of the parameters in the x distribution is called a variance stabili-

zing transformation. In many cases even when x is not binomial such transformed data are approximately normal, and it is often assumed that $T(x) \sim N(\mu,\sigma^2)$. In the final analysis one is however interested in estimating E[x] and Var x as in Example 6 but in this case the distribution of x is unknown. Suppose $Y \sim N(\mu,\sigma^2)$ and $Y = \text{arc Sin } \sqrt{x}$. Then

$$E[x] = \frac{1}{2}\,[1-\text{Cos } 2\mu]\; e^{-2\sigma^2} \tag{35}$$

and

$$\text{Var } x = \frac{1}{8}\left(e^{-8\sigma^2}-1\right)\text{ Cos } 4\mu - \frac{1}{4}\left(e^{-4\sigma^2}-1\right)\text{Cos}^2\, 2\mu.$$

Hence again the problem is one of estimating a function of μ and σ^2 with normal data.

Each of the above examples lead to a problem of estimating a function of the parameters of a normal distribution. Each of these problems have been solved in various places in the literature. However in a recent article (1973) [4] Gray, Watkins, and Schucany have made use of the generalized jackknife to obtain a general solution to this class of problems. We will now state their main result and show how it can be utilized in the above examples. To see how the jackknife plays a fundamental role in obtaining this result see the above mentioned article.

Theorem 4. Under suitable regularity conditions the unique minimum variance unbiased estimator for:

(i) $f(\mu)$, f analytic over the reals, is $\hat{f}(\mu)$, where

$$\hat{f}(\mu) = f\,(\bar{X}) + \sum_{m=1}^{\infty} \frac{(-1)^m\, \Gamma(\frac{n-1}{2})\, f^{(2m)}(\bar{X})}{\Gamma(\frac{n-1}{2}+m)\; m!}\left(\frac{s^2}{4}\right)^m \tag{36}$$

where σ^2 is assumed unknown (σ^2 known, replace $\left[\Gamma(\frac{n-1}{2})\Big/\Gamma(\frac{n-1}{2}+m)\right]\left(\frac{s^2}{4}\right)^m$ by $\left(\frac{\sigma^2}{2n}\right)^m$, see [4]).

(ii) $f(\mu,\sigma^2)$, f analytic on $(-\infty,\infty)$ X $[0,\infty)$, is $\hat{f}\,\mu,\sigma^2)$, where

$$\hat{f}(\mu,\sigma^2) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^i\, (s^2)^{i+j}\, n^j\, \Gamma(\frac{n-1}{2})}{i!\; j!\; 2^{2i+j}\; \Gamma(\frac{2i+2j+n-1}{2})}\, f^{(2i,j)}(\bar{X},0), \tag{37}$$

where

$$f^{(2i,j)}\,(\bar{X},0) = \frac{\partial^{2i}\,\partial^{j}\, f(\mu,\sigma^2)}{\partial\,\mu^{2i}\,\partial(\sigma^2)^j}\,\Bigg|_{\bar{X},0}.$$

In (i) and (ii) above, $s^2 = n^{-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$.

We can of course now write down the solutions to examples 4,5,6 and 7 at once. That is, if:

(a) $f(\mu) = C_1 \text{ Sin } \mu\, t + C_2 \text{ Cos } \mu\, t$, then

$$\hat{f}(\mu) = C_1 \sin \bar{K}\, t + C_2 \cos \bar{K}\, t + \sum_{m=1}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}\right)\left[C_1 \sin \bar{K}t + C_2 \cos \bar{K}t\right]\left(\frac{t^2 s^2}{4}\right)^m}{\Gamma\left(\frac{n-1}{2}+m\right) m!}$$

$$= f(\bar{K})\, \Gamma\left(\frac{n-1}{2}\right)\left(\frac{tS}{2}\right)^{-\frac{n-3}{2}} I_{\frac{n-3}{2}}(tS), \qquad (38)$$

where $I_{\frac{n-3}{2}}$ is the modified Bessel function of index $(n-3)/2$.

(b) $f(\mu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{k-\mu} e^{-\frac{1}{2}t^2}\, dt$, then

$$\hat{f}(\mu) = f(\bar{X}) - \frac{e^{-\frac{1}{2}(k-\bar{X})^2}}{\sqrt{2\pi}} \sum_{i=1}^{\infty} \frac{(-1)^i H_{2i-1}(k-\bar{X})}{i!\,(2n)^i}$$

$$= f\left(\frac{\sqrt{n}\,(k-\bar{X})}{\sqrt{n-1}}\right), \qquad (39)$$

which is the well known UMVU.

(c) $f(\mu,\sigma)^2 = e^{\mu + \frac{1}{2}\sigma^2}$ = E[lognormal random variable].

Therefore

$$\hat{f}(\mu,\sigma^2) = e^{\bar{x}} \sum_{m=0}^{\infty} \frac{\left[\frac{(n-1)s^2}{4}\right]^m}{\left(\frac{n-1}{2}\right)_m m!}, \text{ where } \left(\frac{n-1}{2}\right)_m = \frac{\Gamma\left(\frac{n-1}{2}+m\right)}{\Gamma\left(\frac{n-1}{2}\right)}.$$

Also as stated in Example 6, the variance of a lognormal is given by

$$\text{Var } Y = e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right).$$

Simple differentiation yields

$$\widehat{\text{Var } Y} = e^{2\bar{x}} \sum_{m=0}^{\infty} \frac{(s^2)^m}{\left(\frac{n-1}{2}\right)_m m!}\left[(n-1)^m - \left(\frac{n-2}{2}\right)^m\right].$$

(d) $f(\mu,\sigma^2) = \frac{1}{2}\left[1-\cos 2\mu\right] e^{-2\sigma^2}$.

In this case

$$f^{(2i,j)}(\bar{x},0) = \frac{(-2)^j}{2}[1-\cos 2\bar{x}], \quad i = 0 \qquad i = 0, 1, \ldots$$

$$= (-1)^{i+j+1}\, 2^{2i+j-1} \cos 2\bar{x}, \quad i = 1,\ldots \qquad j = 0, 1, \ldots$$

and hence

$$\hat{f}(\mu,\sigma^2) = \frac{\alpha!}{2}\left[w_1^{-\alpha} J_\alpha (2w_1) - w_2^{-\alpha} J_\alpha(2w_2) \cos 2\bar{x}\right],$$

$$\alpha = \frac{n-3}{2}, \quad w_1 = \sqrt{n}\, s = \sqrt{\sum (x_i-\bar{x})^2}, \quad w_2 = \sqrt{\frac{n}{n-1}}\, w_1,$$

where J_α is the Bessel function of index α.

ACKNOWLEDGMENTS

Most of the research reported on in this paper was supported under ONR Contracts N00014-68-A-0515 and N0014-72-A-0296.

REFERENCES

[1] Arvensen, J. [1969], "Jackknifing U-Statistics," Ann. Math. Statist. 40, 2076-2100.

[2] Arvensen, J., and Schmitz, T. [1970], "Robust Procedures for Variance Component Problems Using the Jackknife," Biometrics 26, 677-686.

[3] Gray, H. L. and Schucany, W. R. [1972], The Generalized Jackknife Statistic, Marcel Dekker, Inc. New York.

[4] Gray, H. L., Watkins, T. A., and Schucany, W. R. [1973], "On the Jackknife Statistic and Its Relation to UMVU Estimators in the Normal Case," Communications in Statistics, Vol. 2, No.4.

[5] Gray, H. L., Watkins, T. A., and Adams, J. E. [1972], "On the Generalized Jackknife, and Its Extensions, and Relation to e_n-transformations," Ann. Math. Statist., Vol. 43, No. 1, 1-30.

[6] Quenouille, M. [1956], "Notes on Bias in Estimation," Biometrika 43, 353-360.

[7] Mosteller, F., and Tukey, J. W. [1968], "Data Analysis, Including Statistics," Handbook of Social Psychology, G. Lindzey and E. Aronson, Eds., Addison-Wesley, Reading, Mass.

[8] Schucany, W. R., Gray, H. L., and Owen, D. B. [1971], "On Bias Reduction in Estimation," JASA 66, 524-533.

[9] Tukey, J. W. [1958], "Bias and Confidence in Not Quite Large Samples," (abstract), Ann. Math. Statist. 29, 614.

RELIABILITY GROWTH[1]

Frank Proschan

Department of Statistics
The Florida State University
Tallahassee, Florida

ABSTRACT

This is an expository paper prepared for presentation at the Nineteenth Conference on the Design of Experiments in Army Research, Development and Testing, Rock Island, Illinois, October 24-26, 1973.

In this paper, we discuss, interpret, and illustrate the Birnbaum measure of the importance of each component in a coherent system. (See Z. W. Birnbaum, On the importance of different components in a multicomponent system, in Multivariate Analysis - II, ed. by P. R. Krishnaiah, Academic Press, 1969.) We show how this measure of component importance can be used to determine system reliability growth from component reliability growth. Finally, we present and illustrate an algorithm to calculate the optimum amount of effort to expend on improving individual component reliabilities so as to achieve a desired system reliability growth at minimum total cost for reliability growth.

1 Research sponsored by the Air Force Office of Scientific Research, AFSC,USAF, under Grant No. AFOSR 74-2581. The United States Government is authorized to reproduce and distribute reprints for government purposes not withstanding any copyright notation hereon.

1. Introduction. In the development of a complex system, it is very useful to (a) determine the importance of each component in contributing to system reliability, and (b) determine the optimum effort to allocate to each component to achieve a desired system reliability growth, taking into account each component importance and each component cost.

In this paper, we discuss, interpret, and illustrate the Birnbaum (1969) measure of component importance. We show how this measure of component importance can be used to determine system reliability growth from component reliability growth. Finally, we present an algorithm to calculate the optimum amount of effort to expend on improving individual component reliabilities so as to achieve a desired system reliability growth.

2. Preliminaries. We shall consider the large, natural class of systems known as coherent systems. To define a coherent system, first we indicate the state of the system by ϕ, where

$$\phi = \begin{cases} 1 & \text{if the system is functioning,} \\ 0 & \text{if the system is failed.} \end{cases}$$

Next we indicate the state of the i^{th} component in the system by x_i, where

$$x_i = \begin{cases} 1 & \text{if the } i^{th} \text{ component is functioning,} \\ 0 & \text{if the } i^{th} \text{ component is failed,} \end{cases}$$

$i = 1, \ldots, n$. We assume that system state ϕ is a deterministic function of the component states, $x_1, \ldots, x_n$:

$$\phi = \phi(\underline{x}),$$

where $\underline{x} = (x_1, \ldots, x_n)$. We call ϕ the structure function of the system; for brevity, we refer to "structure ϕ."

2.1. Definition. A structure function ϕ is coherent if

(a) $\phi(\underline{0}) = 0$,

(b) $\phi(\underline{x}) \leq \phi(\underline{y})$ for $x_1 \leq y_1, \ldots, x_n \leq y_n$;

(c) $\phi(\underline{1}) = 1$, and

(d) $\phi(\underline{x}|x_i = 1) \neq \phi(\underline{x}|x_i = 0)$ for fixed $i = 1, \ldots, n$,

where $\underline{0} \overset{\text{dgf}}{=} (0, \ldots, 0)$ and $\underline{1} \overset{\text{dgf}}{=} (1, \ldots, 1)$.

(a) states that the system is in the failed state if each component is in the failed state. (b) states that repairing failed components can only improve system performance. (c) states that the system is in the functioning state if each component is functioning. (d) states that each component is relevant in defining system state as a function of component states.

Birnbaum, Esary, and Saunders (1961), Barlow and Proschan (1965), and Barlow and Proschan (1974) discuss the properties of coherent structures and their applications in reliability theory.

Throughout the paper we consider only coherent structures, since these

are the structures generally encountered in applications.

In many practical applications, the state X_i of component i is a Bernouilli random variable; i.e.,

$$X_i = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } q_i \overset{\text{dgf}}{=} 1 - p_i. \end{cases}$$

p_i represents the reliability of component i. We shall assume throughout the paper that component states $X_1, \ldots, X_n$ are stochastically independent.

Corresponding to these random component states $X_1, \ldots, X_n$, the system random state $\phi(\underline{X})$ is also a Bernouilli random variable with $P[\phi(\underline{X}) = 1] \overset{\text{dgf}}{=} h(\underline{p})$, where $\underline{p} = (p_1, \ldots, p_n)$. $h(\underline{p})$ yields system reliability as a function of the individual component reliabilities.

In the model treating optimum allocation of effort for component reliability growth (Sections 4 and 5), we shall need the following concepts:

2.2. Definition. A component with reliability p is said to have hazard $r = -\log p$.

2.3. Definition. Let $h(\underline{p})$ be the reliability function of a coherent system. Let

$$\eta(\underline{r}) = -\log h(e^{-r_1}, \ldots, e^{-r_n}) \qquad (2.1)$$

yield system hazard as a function of component hazards. The function $\eta(\underline{r})$ is called the hazard transform of the system. See Esary, Marshall, and Proschan (1970) for a discussion of properties and applications of the hazard transform.

3. Measuring Component Importance. Birnbaum (1969) has proposed the

following reasonable measure of the importance of each component in a coherent system.

3.1. Definition. For structure ϕ with corresponding reliability function $h(\underline{p})$, the importance $I_\phi(i; \underline{p})$ of component i is given by

$$I_\phi(i; \underline{p}) = \frac{\partial h(\underline{p})}{\partial p_i}, \quad i = 1, \ldots, n. \tag{3.1}$$

We often suppress ϕ and $\underline{p}$, and simply write $I(i)$.

The importance $I(i)$ of component i can thus be interpreted as the rate at which system reliability grows as the reliability of component i grows.

Equivalent forms of (3.1) are often useful and provide additional insight concerning component importance. To develop these, we start with the easily verified identity:

$$h(\underline{p}) = p_i h(\underline{p}|p_i = 1) + q_i h(\underline{p}|p_i = 0) \tag{3.2}$$

for $i = 1, \ldots, n$. Differentiating, we immediately obtain an equivalent definition of component importance:

$$I(i) = h(\underline{p}|p_i = 1) - h(\underline{p}|p_i = 0) \quad \text{for} \quad i = 1, \ldots, n. \tag{3.1a}$$

Eq. (3.1a) also provides a strong intuitive motivation for the definition of component importance. It states that the importance of component i may be measured by the growth in system reliability upon repair of failed component i.

From (3.1a) we may obtain a third formula for component importance.

From the definition of $h(p)$, and the fact that a Bernouilli random variable X with parameter p satisfies $EX = p$, we obtain:

$$I(i) = E[\phi(\underline{X}|X_i = 1) - \phi(\underline{X}|X_i = 0)] \tag{3.1b}$$

for $i = 1, \ldots, n$. Eq. (3.1b) tells us that the importance of component i may be measured by the average improvement in the state of the system as component i improves from the failed state to the functioning state.

3.2. Remarks.

(1) From (3.1a) it is obvious that, in general:

$$0 \leq I(i) \leq 1 \quad \text{for} \quad i = 1, \ldots, n. \tag{3.3}$$

Further analysis easily shows that if $0 < p_i < 1$, $i = 1, \ldots, n$, then

$$0 < I(i) < 1 \quad \text{for} \quad i = 1, \ldots, n. \tag{3.4}$$

(2) From (3.1a) it is clear that the importance of component i does not depend on p_i, but, of course, does depend on the remaining p_j, $j \neq i$.

System Reliability Growth as a Function of Component Reliability Growth

The measure of component importance defined by (3.1) (or equivalently by (3.1a) or (3.1b)) may be used to evaluate system reliability growth resulting from the reliability growth of individual components. By the chain rule for differentiation, we obtain

$$\frac{dh}{dt} = \sum_{i=1}^{n} \frac{\partial h}{\partial p_i} \frac{dp_i}{dt} ,$$

where t is a common parameter, say the time elapsed since system development began or the money expended on system development. Using Definition (3.1) for component importance, we have

$$\frac{dh}{dt} = \sum_{i=1}^{n} I(i) \frac{dp_i}{dt} . \qquad (3.5)$$

Thus the rate at which system reliability grows is a weighted combination of the rates at which component reliabilities grow, where the weights are the measures of component importance. Formula (3.5) gives the system designer quantitative information as to the payoff accruing from effort expended on achieving individual component reliability growth.

A closely related formula with similar practical applications is obtained from (3.5):

$$\Delta h = \sum_{i=1}^{n} I(i) \Delta p_i , \qquad (3.6)$$

where Δh represents the growth in system reliability corresponding to growths Δp_i in component reliabilities. As in (3.5), the component measures of importance enter as weights. Thus small growths Δp_i in component reliabilities lead to a corresponding growth Δh in system reliability

in accordance with (3.6).

Some Standard Systems. For convenience, we label components so that $p_1 \leq \ldots \leq p_n$.

3.3. Series System. For a series system, $h(\underline{p}) = \prod_1^n p_i$, so that $I(i) = \frac{\partial h}{\partial p_i} = \prod_{j \neq i} p_j = \frac{1}{p_i} h(\underline{p})$. It follows that component importance satisfies:

$$I(1) \geq \ldots \geq I(n). \tag{3.7}$$

Inequality (3.7) states that for a series system, the smaller the component reliability, the greater the component importance. In particular, the component with the lowest reliability is the most important to the system. This reflects the well known principle that "a chain is as strong as its weakest link".

3.4. Parallel system. For a parallel system, $h(\underline{p}) = 1 - \prod_1^n q_i$, so that $I(i) = \frac{\partial h}{\partial p_i} = \prod_{j \neq i} q_j = \frac{1}{q_i} [1 - h(\underline{p})]$. Thus component importance satisfies:

$$I(i) \leq \ldots \leq I(n). \tag{3.8}$$

For a parallel system the larger the component reliability, the greater the component importance. In particular, the component with highest reliability is the most important to the system. This, too, is intuitively reasonable, since if just one component functions, the system functions.

3.5. 2-out-of-3 system. A 2-out-of-3 system functions if and only if at least 2 out of the 3 components in the system function. The reliability function $h(\underline{p}) = p_1p_2q_3 + p_1q_2p_3 + q_1p_2p_3 + p_1p_2p_3$. Thus component importance is given by

$$I(1) = \frac{\partial h}{\partial p_1} = p_2q_3 + q_2p_3,$$

$$I(2) = p_1q_3 + q_1p_3,$$

$$I(3) = p_1q_2 + p_2q_1.$$

It follows that $I(2) - I(1) = (p_2 - p_1)(p_3 - q_3) \geq 0$ for $p_3 \geq 1/2$. Similarly $I(3) - I(2) \geq 0$ for $p_1 \geq 1/2$. Thus for a 2-out-of-3 system with $1/2 \leq p_1 \leq p_2 \leq p_3$, $I(1) \leq I(2) \leq I(3)$; i.e., component importance is in the same order as component reliability. If however $p_1 \leq p_2 \leq p_3 \leq 1/2$, then component importance is in the reverse order.

Note that in this case, component importance ranks depend on component reliability values.

4. Optimum Component Reliability Growth.

4.1. Model. Assume a series system of n stochastically independent components. The present reliability of component i is $p_i^\circ \overset{\text{dgf}}{=} e^{-r_i^\circ}$, where r_i° is the corresponding hazard. To increase the component reliability to $p_i \overset{\text{dgf}}{=} e^{-r_i}$ costs $c_i(r_i^\circ - r_i)$ (cost may be measured in dollars, man-hours, time, etc., or some weighted combination of factors). We suppress the dependence of c_i on r_i°, showing only its dependence on $r_i^\circ - r_i$, the decrease in component hazard. c_i satisfies:

(a) $c_i(0) = 0$ (no improvement costs nothing),

(b) $c_i(\cdot)$ is a strictly convex increasing function (the cost of achieving successive fixed-size reductions in hazard increases), with continuous first derivative.

We wish to achieve a growth in system reliability to the level $p^* > \prod_1^n p_i^\circ$, the present level, with minimum total expenditure on individual component reliability growth. Specifically, we wish to determine component reliability level $p_i^* \geq p_i^\circ$ (or equivalently, component hazard $r_i^* \leq r_i^\circ$), $i = 1, \ldots, n$, satisfying

$$\prod_1^n p_i^* = p^*, \tag{4.1}$$

at minimum total cost $\sum_1^n c_i(r_i^\circ - r_i^*)$.

Solution. Taking logarithms in (4.1), our problem is equivalent to finding $r_i^* \leq r_i^\circ$, $i = 1, \ldots, n$, such that

$$\sum_1^n r_i^* = r^* \tag{4.1a}$$

with $\sum_1^n c_i(r_i^\circ - r_i^*)$ a minimum. Note that we are minimizing a sum of convex functions, each depending on a separate decision variable, subject to a linear constraint on the decision variables.

The problem is a standard problem of optimization subject to constraints. The intuitive idea underlying the solution is simple. We spend continuously towards the growth of various components until we just attain the desired system reliability p^*. At every stage, the rate of expenditure

per unit decrease in system (or equivalently, component) hazard is the least possible at that stage, or equivalently, the rate of decrease in system hazard is the most possible per unit of expenditure. The expenditure continues until the desired system reliability is attained. It is intuitively clear that when the desired system reliability is attained, the rate of expenditure per unit decrease in system hazard is the same (say, λ) for all component types, insofar as this is possible. (For some component types, achieving the rate λ may be impossible because the rate of expenditure is always greater than λ, while for other component types, achieving the rate λ may be impossible because the rate of expenditure is always smaller than λ.) When the solution is achieved, the rate of expenditure is the same for all component types as far as possible, since if it were not, we could always improve the outcome by shifting expenditures from one component to another. The solution may be summarised as follows.

(1) For $\lambda \geq 0$, $i = 1, \ldots, n$, define

$$r_i(\lambda) = \begin{cases} r_i^{\circ} & \text{if } c_i'(0) > \lambda \quad (4.2a) \\ 0 & \text{if } c_i'(r_i^{\circ}) < \lambda \quad (4.2b) \\ x \text{ such that } c_i'(r_i^{\circ} - x) = \lambda, & \text{otherwise.} \quad (4.2c) \end{cases}$$

λ may be thought of as the final rate of expenditure upon attainment of some, as yet unspecified, system reliability. Thus (4.2c) defines the corresponding component hazard required. (4.2b) states that if the maxi-

mum rate of expenditure possible for the i^{th} component is below λ, then spend enough to reduce component i hazard to 0. (4.2a) states that if the minimum rate of expenditure for the i^{th} component is $> \lambda$, then retain the present hazard r_i^o (i.e., spend nothing for reliability growth of component i).

Note that $r_i(\lambda)$ is uniquely determined, since c_i' is continuous and strictly increasing.

(2) For $c_i'(0) \leq \lambda \leq c_i'(r_i^o)$, $i = 1, \ldots, n$, $r_i(\lambda)$ is a continuous decreasing function of λ. This follows from the fact that c_i' is an increasing continuous function.

(3) For $\lambda \geq 0$, define

$$r(\lambda) = \sum_{i=1}^{n} r_i(\lambda) \qquad (4.3)$$

$r(\lambda)$ represents the system hazard resulting from the choice of λ. $r(\lambda)$ is a continuous decreasing function of λ for $\lambda \geq 0$, since $r(\lambda)$ is the sum of continuous decreasing functions. $r(\lambda)$ decreases from r^o to 0 as λ increases from 0 to ∞.

(4) There exists at least one value of $\lambda \geq 0$ such that $r(\lambda) = r$ for $0 < r < r^o$. This follows immediately from (3).

Using these facts, we may now state the solution to the problem posed in Model 4.1 in

4.2. Theorem. Let λ^* be the minimum value of λ such that $r(\lambda) = r^*$. Then hazard $r_i^* = r_i(\lambda^*)$, $i = 1, \ldots, n$, minimizes total cost $\sum_{i=1}^{n} c_i(r_i^o - r_i^*)$ subject to $0 \leq r_i^* \leq r_i^o$, $i = 1, \ldots, n$, and $\sum_{1}^{n} r_i^* = r^*$.

Proof. Consider any other choice $(r_1, \ldots, r_n)$ satisfying the con-

straints $0 \le r_i \le r_i^0$, $i = 1, \ldots, n$, and $\sum_1^n r_i = r^*$. Let $r_i < r_i^*$ for i in I_1, and $r_i > r_i^*$ for i in I_2, where I_1 and I_2 are disjoint subsets of $(1, \ldots, n)$, one of which may be empty, but not both. For i in I_1, $c_i'(r_i^0 - r_i) \ge c_i'(r_i^0 - r_i^*)$; for i in I_2, $c_i'(r_i^0 - r_i^*) \le c_i'(r_i^0 - r_i^*)$. Hence

$$\sum_1^n c_i(r_i^0 - r_i) - \sum_1^n c_i(r_i^0 - r_i^*) = \sum_{i \text{ in } I_1} [c_i(r_i^0 - r_i) - c_i(r_i^0 - r_i^*)]$$

$$- \sum_{i \text{ in } I_2} [c_i(r_i^0 - r_i^*) - c_i(r_i^0 - r_i) = \sum_{i \text{ in } I_1} \int_{r_i}^{r_i^*} c_i'(r_i^0 - x)dx$$

$$- \sum_{i \text{ in } I_2} \int_{r_i^*}^{r_i} c_i'(r_i^0 - x)dx \ge \lambda^* \left[\sum_{i \text{ in } I_1} (r_i^* - r_i) - \sum_{i \text{ in } I_2} (r_i - r_i^*) \right]$$

$$= \lambda^* \left[\sum_1^n r_i^* - \sum_1^n r_i \right] \ge \lambda^*(r^* - r^*) = 0.$$

Thus $(r_1^*, \ldots, r_n^*)$ minimizes $\sum_1^n c_i(r_i^0 - r_i)$ subject to $\sum_1^n r_i = r^*$. ||

4.3. Remark. The corresponding optimum expenditure for the reliability growth of component i is $c_i(r_i^0 - r_i^*)$, $i = 1, \ldots, n$; the corresponding total expenditure on component reliability growth of $\sum_1^n c_i(r_i^0 - r_i^*)$ is the minimum expenditure that will achieve a system reliability growth to the desired level of p^*.

5. Illustrations of Optimum Reliability Growth. We illustrate the calculations required by working out an example. The example is based on the Carhart-Herd (1957) model in which the cost of developing component reliability is assumed to be proportional to a positive power of the mean life achieved. The model assumes further that component lifelength is governed by an exponential distribution. From these two assumptions, it is

easy to see that $c_i(\cdot)$ is of the form:

$$c_i(r_i^\circ - r_i) = b_i(r_i^{-a_i} - r_i^{\circ -a_i}) \quad \text{for} \quad 0 < r_i \leq r_i^\circ, \quad (5.1)$$

where $b_i > 0$ and $a_i > 0$ for $i = 1, \ldots, n$. It follows, of course, that $c_i''(r_i^\circ - r_i) = b_i a r_i^{-a-1}$. Thus $c_i(\cdot)$ is an increasing, strictly convex function for $0 < r_i \leq r_i^\circ$.

Example. To illustrate the solution of Section 4 in the simplest fashion, we consider a two component series system. Presently, component 1 has reliability $p_1^\circ = .75$ or hazard $r_1^\circ = -\log .75$, while component 2 has reliability $p_2^\circ = .60$ or hazard $r_2^\circ = -\log .60$; thus the present system reliability is $p^\circ = p_1^\circ p_2^\circ = .45$.

To decrease the hazard of component 1 from $r_1^\circ = -\log .75$ to $r_1 \leq -\log .75$ requires an expenditure of

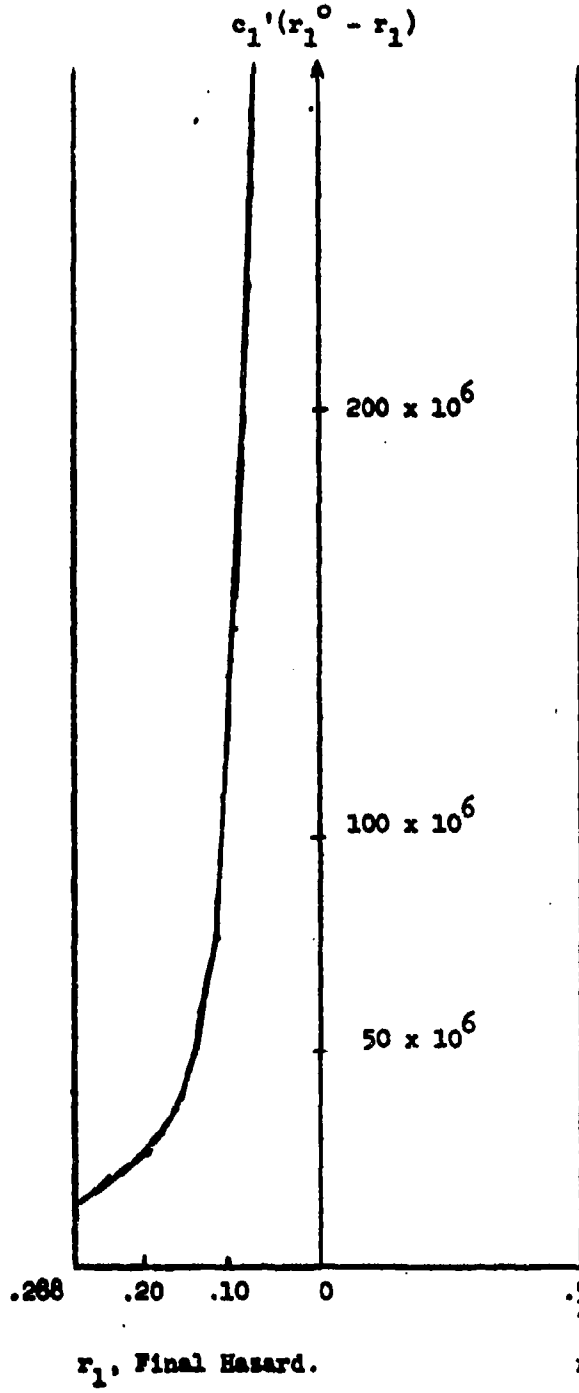
$$c_1(-\log .75 - r_1) = 10^6[r_1^{-1} - (-\log .75)^{-1}].$$

Similarly, to decrease the hazard of component 2 from $r_2^\circ = -\log .60$ to $r_2 \leq -\log .60$ requires an expenditure of

$$c_2(-\log .60 - r_2) = 2 \times 10^5[r_2^{-2} - (-\log .60)^{-2}].$$

Subsystem 1.

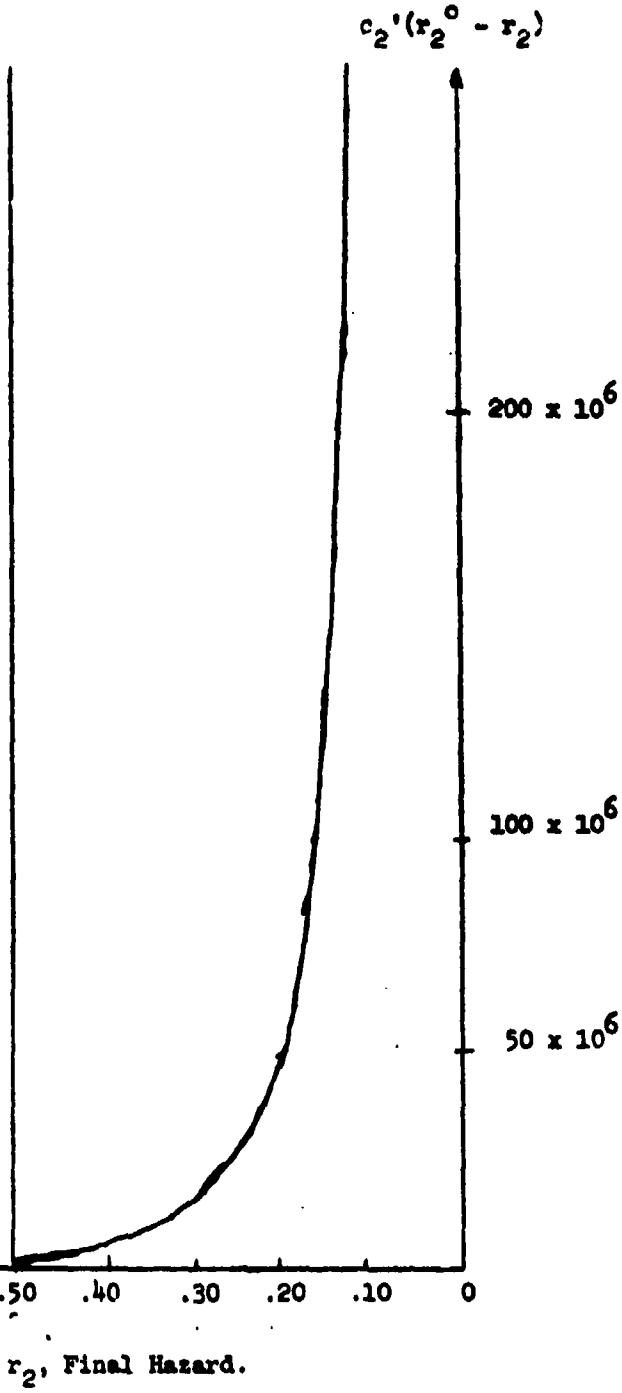
$c_1'(r_1^\circ - r_1) = 10^6 r_1^{-2}$



r_1, Final Hazard.

Subsystem 2.

$c_2'(r_2^\circ - r_2) = .4 \times 10^6 r_2^{-3}$



r_2, Final Hazard.

The problem is to determine the required amount of reliability growth for component 1 to a new value $p_1^* \geq .75$, and for component 2 to a new value $p_2^* \geq .60$, so as to achieve system reliability growth to the desired value $p^* = p_1^* p_2^* > .45$, at minimum total cost.

Differentiating, we obtain

$$c_1'(-\log .75 - r_1) = 10^6 r_1^{-2} ,$$

and

$$c_2'(-\log .60 - r_2) = 4 \times 10^5 r_2^{-3} .$$

By a second differentiation, we verify that $c_1(\cdot)$ and $c_2(\cdot)$ are strictly convex. The functions $c_1'(\cdot)$ and $c_2'(\cdot)$ are plotted in Figure 5.1.

Let us obtain solutions corresponding to each of a family of values of p^*. We may accomplish this most readily by specifying various values of λ^* and obtaining the corresponding sets r_1^*, r_2^* (or equivalently, p_1^*, p_2^*). In a sense, we will be solving a series of different problems but until we have solved the problem we will not know which problem we have solved (i.e., if we specify λ^*, we will not know what value $p^* = p_1^* p_2^*$ will be until we have obtained the individual values p_1^*, p_2^*.

Suppose then we pick $\lambda^* = 3 \times 10^6$, 5×10^6, 6×10^6, etc., as listed in column (1) of Table 5.1. Consider first $\lambda^* = 3 \times 10^6$. Since the curve of $c_1'(\cdot)$ (Fig. 5.1) lies entirely above $\lambda^* = 3 \times 10^6$, we obtain from 4.2(a), $r_1^* = r_1(3 \times 10^6) = -\ln .75$. Next, from the curve of $c_2'(\cdot)$ (Fig. 5.1), we read off $r_2^* = r_2(3 \times 10^6) = .50$, the abscissa corresponding to an ordinate of 3×10^6 (in accordance with (4.2c)). Taking antilogs, we obtain $p_1^* = .75$ (column (4) of Table 5.1) and $p_2^* = .61$

Table 5.1. Computation of Solution

(1)† λ^*	(2) $r_1(\lambda^*)$	(3) $r_2(\lambda^*)$	(4) p_1^*	(5) p_2^*	(6) System reliability $p = p_1^* p_2^*$	(7)† $10^6[r_1^{*-1} - (-\ln .75)^{-1}]$	(8)† $.2 \times 10^6[r_2^{*-1} - (-\ln .60)]^{-2}$	(9)† Minimum total expenditure = (7) + (8)
3	.28 (= ln.75)	.50	.75	.61	.46	0	.03	.03
5	.28	.45	.75	.64	.48	0	.2	.2
6	.28	.40	.75	.67	.51	0	.5	.5
9	.28	.35	.75	.71	.53	0	.9	.9
15	.23	.30	.80	.74	.59	.8	1.4	2.2
25	.20	.26	.82	.77	.63	1.5	2.2	3.7
45	.15	.21	.86	.81	.70	3.2	3.8	7.0
50	.13	.20	.88	.82	.72	4.2	4.2	8.4
100	.10	.17	.90	.84	.76	6.5	6.0	12.5
400	.05	.10	.95	.90	.86	16.5	19.2	35.7

† The entries in columns 1, 7, 8, and 9 are in multiples of 10^6.

(column (5) of Table 5.1). In column (6), we list $p^* = p_1^* p_2^* = .46$. Since $p_1^* = .75$, the original reliability of subsystem 1, the corresponding expenditure is 0 (column (7)); to attain $p_2^* = .61$, the required expenditure is $.03 \times 10^6$ (column (8)). Finally, in column (9) is shown the total expenditure of $.03 \times 10^6$ obtained by summing the corresponding entries in columns (7) and (8).

In a similar fashion, we compute the rows in Table 5.1 corresponding to $\lambda^* = 5 \times 10^6$, 6×10^6, etc. The resulting relationship between the minimum total cost required as a function of system reliability attained is shown in Figure 5.2. We can now state that to attain a system reliability of $p^* = .70$ say, we should aim to improve the reliability of subsystem 1 to .86 and of subsystem 2 to .81. The resulting total expenditure of 7.0×10^6 is minimum among allocations of funds which would yield a system reliability of .70. A similar statement may be made about each of the points plotted in Fig. 5.2.

If we wish to compute the optimal set p_1^*, p_2^* corresponding to a point not explicitly plotted in Fig. 5.2, such as corresponding to $p^* = .65$, say, we proceed as follows. We plot a curve showing λ^* as a function of p^* (Fig. 5.3). From the curve, we read off the value of $\lambda^* = 26 \times 10^6$ corresponding to $p^* = .65$. We then compute $p_1^* = .84$, $p_2^* = .78$ corresponding to $\lambda^* = 26 \times 10^6$. The resulting system reliability $p^* = p_1^* p_2^* = .65$ and the expenditure required is 2.4×10^6. Note that the resulting $p^* = p_1^* p_2^*$ might <u>not</u> have been the desired value .65; suppose it had been $< .6.5$. In this case, if the discrepancy had been large enough, a new computation would be made using a slightly larger value of λ^*

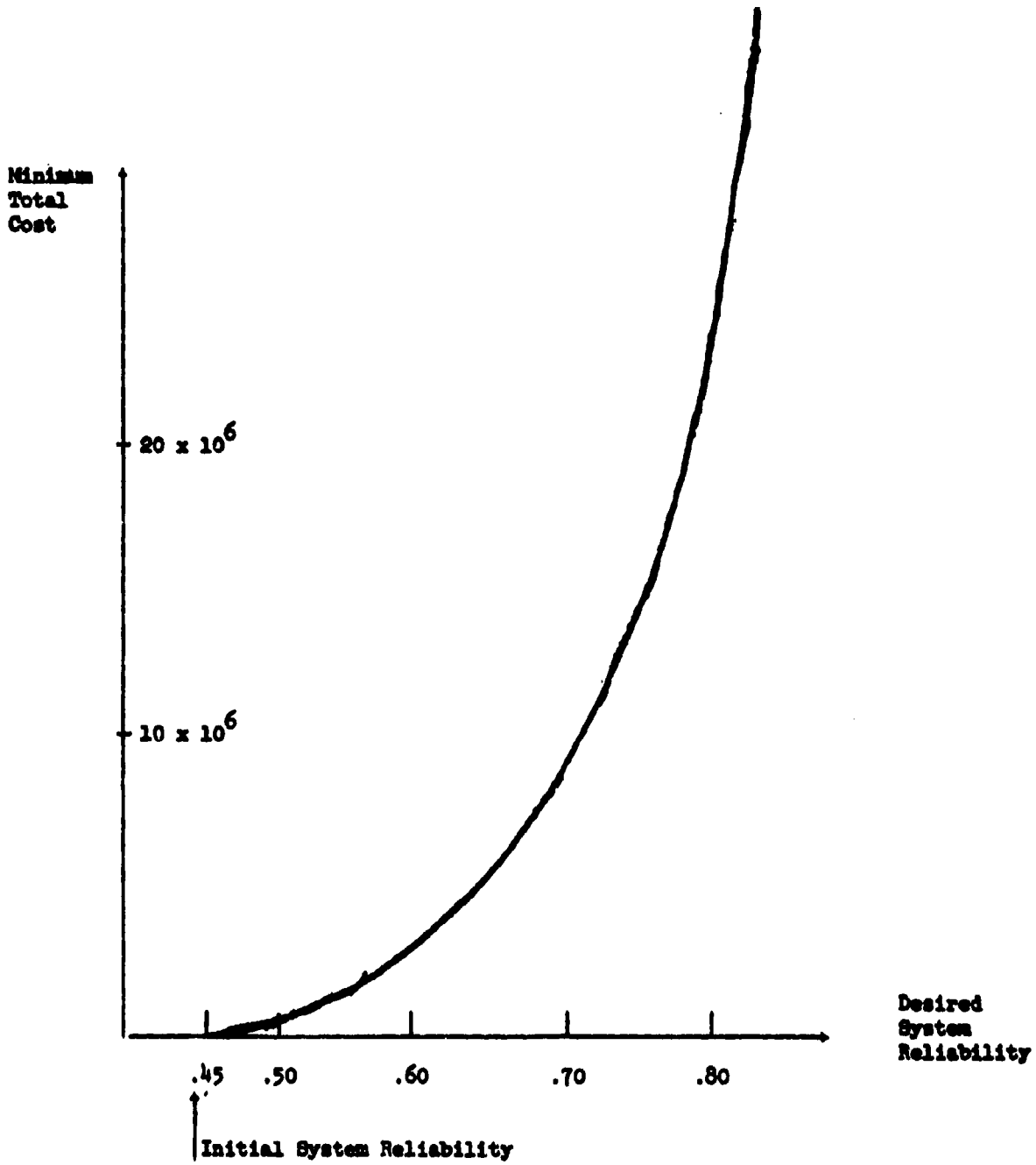


Figure 5.2. Cost Vs. Reliability Under Optimal Allocation.

than previously. By successive trials, we can reduce the error to as small a value as we please.

An interesting point is brought out by this graphical example. In actual applications of this model, it may be that the relationships $c_i(\cdot)$ will be available in the form of graphical curves based on observed data, rather than in the form of explicit mathematical relationships. For example, in the course of developing a subsystem, a record of subsystem reliability attained corresponding to various expenditures of funds may have been accumulated. Note that the graphical procedure for computing the optimal $p_1^*, \ldots, p_n^*$ may be used very nicely, without any need for an explicit mathematical formula for the $c_i(\cdot)$.

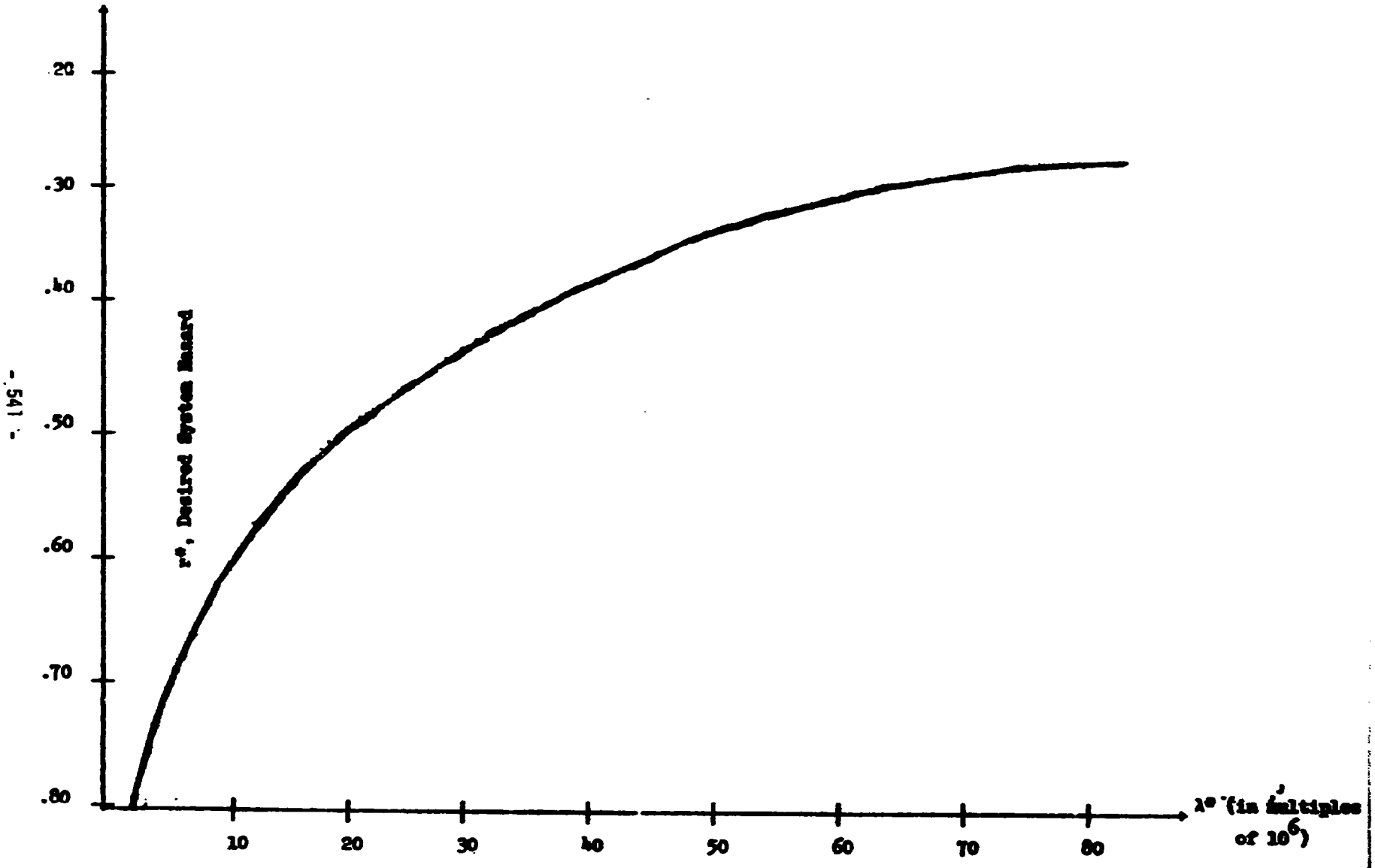


Figure 5.3. Desired System Hazard Vs. Resulting Rate of Expenditure, λ^*

REFERENCES.

[1] Albert, A. and F. Proschan (1959) Increased reliability with minimum effort. Technical Report No. 30, Stanford University Statistics Department.

[2] Barlow, R. E. and F. Proschan (1965) Mathematical Theory of Reliability. Wiley, New York.

[3] Barlow, R. E. and F. Proschan (1974) Statistical Theory of Reliability Vol. I. Holt, Rinehart, and Winston, New York - in press.

[4] Birnbaum, Z. W. (1969) On the importance of different components in a multicomponent system. Multivariate Analysis - II, ed. by P. R. Krishnaiah. Academic Press, New York.

[5] Birnbaum, Z. W., J. D. Esary, and S. C. Saunders (1961) Multicomponent systems and structures and their reliabilities. Technometrics 3 55-77.

[6] Carhart, R. R. and G. R. Herd (1957) A simple cost model for optimizing reliability, in Reliability of Military Electronic Equipment by the Advisory Group on Reliability of Electronic Equipment, 64-73.

[7] Esary, J. D., A. W. Marshall, and F. Proschan (1970) Some reliability applications of the hazard transform. SIAM J. Appl. Math. 18 849-860.

SOME CRITICAL REMARKS ON ACCELERATED LIFE TESTING

Sam C. Saunders
Department of Pure and Applied Mathematics
Washington State University
Pullman, Washington

SUMMARY. The history of the development of mathematical statistics has been that of successive efforts by its practitioners at increasing the generality and applicability of statistical inference. This has often resulted in the gradual weakening of the dependence of statistical models upon specific physical, chemical or engineering concepts which make the model appropriate to a particular problem. One of the current needs in reliability theory is for methods which allow the prediction of life from a few tests in which the process of wear has been accelerated. This paper is an effort to point out the necessity of a thorough understanding of each appropriate cumulative damage process in terms of its chemical or physical behavior before a sufficiently useful analysis can be made. Some remarks are also made on the futility of only a mathematical approach to the problem of accelerated life testing. The reinforcement possible between statistical theory and engineering reality is illustrated through the discussion of the application of certain aspects of renewal theory, reliability theory, fracture mechanics, strength of materials and physics of solids in the mathematical calculation of metallic fatigue in airframe design. This problem represents an extreme in accelerated life testing because of the paucity of data from which long term predictions are made. Some of the difficulties of such an interdisciplinary approach are discussed and the advisability of a cooperative attitude of statisticians toward the contribution from other disciplines is mentioned. Some consequences of utilizing the assumption of constant failure rate, when it is unjustified, are given. The attitudes toward the causes of failure resulting from this constant failure rate syndrome may prove to be equally destructive to the design of a valid accelerated life test.

One such attitude is that certain "isolated failures" may be ignored in the calculation of life expectancy from accelerated life tests. The contrasting reaction is to determine the cause of each failure and "fix" all units presently in service and/or all units to be subsequently manufactured. It is maintained that the statisticians proper role in the analysis of such failure data from accelerated life tests must be more than providing universally applicable statistical techniques which "crunch numbers". Instead his role must involve a synthesis of scientific knowledge from relevant areas for the proper interpretation of the failure and its cause and cure.

INTRODUCTION. The idea of shortening the period of testing which is necessary to determine the life distribution of a system is certainly an appealing one. It offers the hope of decreasing the lead time between design concept and prototype introduction with the consequent advantages of both cost reduction and possible performance superiority. Moreover, it fits in with the aspirations of mankind, expressed throughout history, in the beneficient result of scientific endeavor and it also reflects that ubiquitous hope of getting something for nothing. An instance of close parallel in an earlier age was the research program seeking the transmutation of lead into gold.

After expressing this mild skepticism about the universal applicability of accelerated life testing, let me begin by presenting an archetypical situation in which our fondest hopes for such a program may be realized.

If $T>0$ is a random variable, interpreted as life length, and

$$U(t) = P[T>t] \qquad \text{for } t>0 \tag{1}$$

is the probability at birth of a life exceeding t, then

$$P[T>s+t \mid T>s] = \frac{U(t+s)}{U(s)} \tag{2}$$

is the probability that given an age s the residual life length exceeds t. Now this residual lifetime after s coincides with the life distribution at birth iff

$$U(t+s) = U(s)U(t). \tag{3}$$

This equation has a unique solution; namely,

$$(4) \qquad U(t) = e^{-\lambda t} \qquad \text{for some } \lambda > 0.$$

If T has a distribution given by U, it is said to be exponentially distributed. A component with an exponential life does not wear out! Such a component may fail, but if it has not failed at a given time, it is as good as new.

In what situations can one expect the exponential life to be a suitable description of reality? I now consider two:

(i) the life length of a jeweled bearing in a watch,

(ii) tensile strength of a steel rod.

In (i) the bearing receives such little wear during the watch's operation that it fails only as a result of severe shock, such as being dropped, which can occur at any time for reasons not related to the usual utilization. Of course, the analogy of the jeweled bearing in a watch with the electrical properties of transistors and integrated circuits is so seductive that widespread adoption of such a model for electronic components has been observed.

Suppose that a component has an exponential distribution, given by equation (4), and from a group of n such components which are on life test we wait until the first r≥1 failures occur, say

$$X_{(1)} < X_{(2)} < \cdots < X_{(r)}$$

are the times of failure. We have the famous result of Epstein-Sobel:

Theorem: If the failure note λ is unknown, the best unbiased estimatator of λ^{-1} is, in this situation, the total life statistic divided by the number of failures namely,

$$T_r = \frac{1}{r}\left[\sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)}\right]$$

in which case $\text{var}(T_r) = \frac{\lambda^{-2}}{r}$ and $2r\lambda T_r$ has a χ^2- distribution with 2r degrees of freedom.

As a consequence of this result, it is believed that estimators of, and confidence bounds for the mean life (or the reliability at any given time) can be calculated from counts made only of the total number of hours of observed life and the number of observed failures. This result I call "Pandora's Box" because of its similarity with the mythological account of an instance whereby great mischief was inadvertantly done. (the mischief was accomplished by the release of naive hope which usually proves vain.) "Pandora's Box" has resulted in some dire-practical consequences; namely, the wholesale, but hidden, assumption of exponential life.

Further, there has been a relaxation of the felt obligation to determine the cause of failure instead of just noting the service times at which they occurred. The prospect of great monetary gain resulting from the possible utilization of one-hour tests on each of one-hundred new components to demonstrate a reliability of one-hundred hours of life for each one has great attraction for persons who have but a little knowledge of statistics. (Las Vegas provides the same attraction for the probabilistically ingenuous and with, I believe, the same results.)

Any doubt about the universal applicability of the exponential assumption was met with the rejoinder: "For any continuous life length random variable there exists a transformation which will make the distribution exponential; therefore, one needs merely find that transformation and then use the results for the exponential law."

Of course, blaiming mathematics for one's financial losses at chance has a long history going back at least as far as Cheval De Mere and one cannot justifiably blame the exponential distribution for the idiocies which have been committed in its name. However, one does decry the attitude that accelerated life testing is only a subset of statistical theory and as such should be pursued by mathematical statisticians using just pencil and paper.

Let us make a definition: Accelerated life testing is the design of tests on a component's (or system's) reliability which are to take place under increased stress (or a more stringent environment, or both) in order to predict, as soon as possible, the reliability under the proposed service conditions.

We now state one such problem formulated mathematically:

Given test conditions θ and life lengths $X_1(\theta),\ldots,X_n(\theta)$, estimate the reliability under service conditions ϕ, i.e. estimate

$$\bar{F}(t;\phi) = P[X(\phi)>t] \quad \text{for } t>0,$$

and then place bounds on this estimate.

Almost any person with scientific training will immediately say that this problem is ill-defined and virtually insoluble because there are no stated assumptions relating the parametric values θ and ϕ with the distribution. And that is exactly the point. Any useful model for the solution of this problem must supply information about the parameter space and the ostensibly definable relationships between parameter values and environmental conditions and, we maintain, this problem is more than a problem in mathematical statistics.

As an illustration of a mathematical formulation of a class of problems, we present the following:

Assume that under two different, but fixed, sets of conditions, we have

$$X \sim F \quad \text{and} \quad Y \sim G,$$

where

$$G(x) = h[F(x)] \quad (\text{or } G(x) = F[a(x)])$$

We can further assume the function h, or a, known (or $a(x)\geq x$) and then estimate G, F or h. We might also wish to find confidence bounds. One also might simplify the problem by assuming that G or F is the exponential distribution. See [1] and the references given there.

All of these variations are very interesting mathematical problems which doubtless have some utility. My contention is that any statistical theory which applies equally well to the width of stripes on the back of a zebra and to the life of transistors, ignoring the physical, chemical and biological differences, can't be specific enough for either situation to make any but the grossest of predictions. What is needed in accelerated life testing is less generality and more specificity.

Further, it is contended that it is necessary to understand the physical process governing the influence of increased stress and/or more stringent environment on the distribution of life in each specific instance.

To make this difference of opinion clear, let us consider the simple problem of predicting the wearout of government-issued socks under severe service conditions, given data on the wearout of socks under normal conditions. We can estimate the mean and variance, then guess the percentage of increased wear under the severe conditions and take that percentage reduction of the mean service life and then calculate wearout using an appropriate distribution. Alternatively, we could perform impact and abrasion tests on the socks with appropriate amounts of chemical additives simulating both sweat and dirt until the parameters of the life distribution are known functions of the impact, abrasion and chemical influences on the constituent material in the socks. Then the distributi . of life is calculated under the altered properties of resistance under the severe service regime. In many instances accelerated life testing is at least as much an engineering, physics or chemistry problem as it is a statistical one and in fact I sometimes believe the number of parameters fitted statistically may be proportional to the amount of ignorance of the underlying mechanisms of failure.

A constant failure rate is said to generate "random failures" which is interpreted to mean failures for which the cause will never be duplicated and so such "isolated failures" need not be analyzed. It has been the experience of many system managers that there is no such thing as a "random" or "isolated" failure. In fact, I heard one say "An isolated failure is just the first failure in what will become a flood of failures unless system modifications are made as soon as possible."

The failure times of integrated circuits are often regarded as random variables with decreasing failure rates (indeed, sound arguments can be made for this contention.) However, accelerated testing depends upon whether the electrical malfunctions are, for example, caused by such divergent means as:

(1) the unfortunate redeposition of dust particles due to the introduction of a zero-gravity state,

(2) the fatigue and rupture of electrical circuit welds due to sonic or ultrasonic vibration, or

(3) the degradation of electrical response due to continual exposure to high levels of gamma radiation.

All of these are known to occur and it is certainly important to know which.

We believe that it is not possible to accelerate the life testing of a system unless you know in general

(i) the distribution of wear or damage which a specified service will impose,

(ii) the mechanisms (or modes) of failure which are important to a specified service and need to be studied.

(iii) the alteration of stress behavior caused by accelerated testing.

There is not likely to be a theory based on statistical considerations alone which will be of much utility, just as a deterministic theory alone based only upon physical considerations may not be worth much in determining an accelerated testing procedure.

To illustrate the type of cooperation between physical theory and statistical analysis which I feel must be accomplished to obtain realistic accelerated life tests, I want to discuss the current state of affairs, vis a vis accelerated life testing, in the determination of airplane reliability.

First of all, there are several desiderata in airframe design: strength, lightness and fatigue resistance. (Actually, designers have known for some time how to construct an airplane which would not fatigue; the reason they have not done so is that all such designs will not fly, but only taxi.) This parenthetical comment is intended to indicate the mutually, oppositional nature of these requirements. Consequently, any design must be a compromise effected by some balance between these competitive factors. To construct an optimal design requires an acceptable definition of optimality. For commercial design, one such definition might be profitability.

The potential profitability of a particular airframe design necessitates a determination of both the strength requirements (for increased payload) and the maintainability requirements (for increased longevity.). Each of these is, of course, subject to stochastic variation. The design motto "get the weight out," which is intended to increase the payload of the plane, usually results in "putting the fatigue in" and as a consequence, the service life of the model is stochastically shortened, which in turn reduces the expected lifetime profitability. There is an added complexity. The weight of a structure made of aluminum can be reduced by half by substituting titanium but at an increase of approximately ten times the cost. Thus, not only may certain designs be much more expensive, but they may have a shorter life. I need not mention any examples of current models. A successful design is one that steers between the Scylla and Charibdis of these alternatives. Moreover, since the designed life is 60,000 hours (30 years) it is clear that all testing of its durability must be of an accelerated nature.

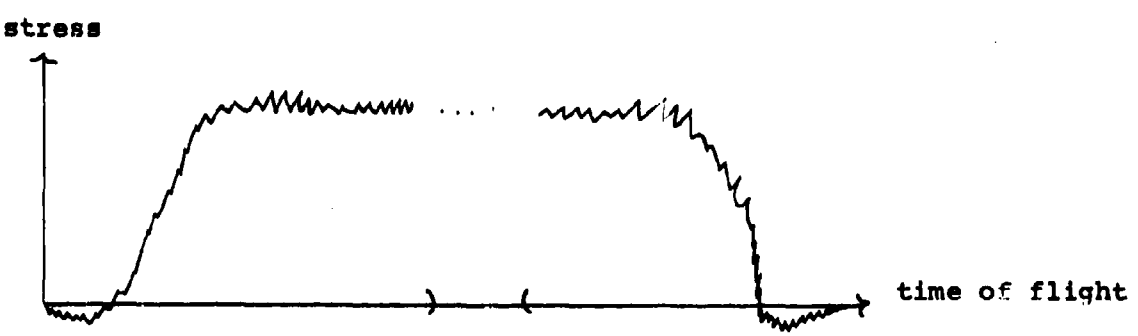
Our first task is to delineate the mechanisms of failure which will be of importance under the specified service regime. Clearly, we must have adequate theories to calculate both the yield strength of materials and the time until the occurrence of metallic fatigue. The former is available in textbooks,(see [2]), it is the latter theory about which there is yet some disagreement especially as regards the interaction with environmental effects. (see [3] and [4]).

The strength requirements and the environmental stress usually dictate the choice of material as well as its fabrication. For example, the brittleness or ductility of the metal and whether it should be extended, cast or forged should be determined. Whether it will periodically be subjected to a corrosive environment and/or elevated to 400F° during each ground-air-ground cycle of the flight must be taken into account for each component. It is beyond the scope of this talk to do any more than sketch the outlines of the general procedure.

Of course, the concept of a fixed service regime is a simplification since, in fact, one of the major sources of variability of the service life is the stochastic nature of the loading spectrum. The identification of this distribution is of fundamental importance since the design must sustain the extreme loads encountered as a consequence of both maneuver and random gust loading but during most of its service life it will encounter stress at a somewhat lower level.

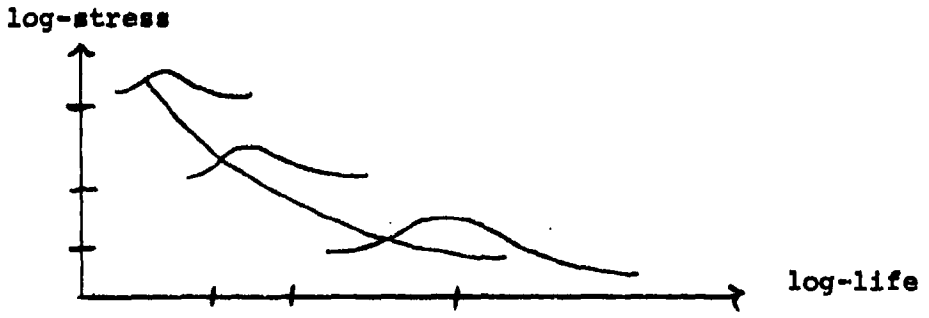
The loads encountered in service are a result of the gross weight and the gusts encountered in the atmosphere. It is necessary that a summary of the statistical characteristics of the atmosphere be available. We presume that $F(\omega;h)$, the distributions of gusts of magnitude ω at a given altitude h, is available, at least in tabular form. The time between the occurrence of gusts is assumed to be exponential with the parameter a function of the airspeed. The airframe is a filter which transforms these gust and maneuver loads as a function of the gross weight and the location of the station on the fuselage or the wing.

For example, if we select a typical lower wing station and graph the ground-air-ground work cycle, we might have:



In the modeling of this G-A-G cycle, we consider two random variables. The time of flight and the gross take-off weight. The distribution of these quantities must be determined from the intended usage of the plane and the route structure of the carrier. This information provides us at each station (of which there are several hundred on each wing) with a distribution of (S,N) which are respectively the loading spectrum and the length of the cycle. Since the stress as a function of time at each station will be transmitted throughout the structure and be born proportionally by each metallic component, it is necessary that something be known about the fatigue behavior of that metal.

Many constant amplitude fatigue life tests are made with notched metallic coupons of a particular alloy in order to obtain an equation between stress and life. This relation is what is called the Wöhler equation and is identical to what statisticians would call a regression equation between log-stress and log-life:



From this information, which will be different for each type of metal and its method of fabrication, is calculated the distribution of time until a given loading spectrum will initiate a crack. A formula which is used for this purpose is an adaptation of an empirical result called "Miner's Rule." After both empirical corroberation and some belated theoretical justification (see [5] and [6]), it is now used to predict the median time until crack initiation. The actual time itself is regarded as a random variable.

The distribution of gust loads during a time interval of given length determines the distribution of the random loading spectrum. The time until crack initiation of a given material for a given distribution of the loading spectrum is assumed to have an extreme value distribution with two parameters α,β say where α is determined as a material constant and β is the median life calculated using Miner's rule and the Wöhler diagrams.

After the crack has been initiated, i.e. is of the length of roughly .01 in, it enters a phase of stable growth. During this time, it will propagate as a function, not only of the maximum and minimum stress of each oscillation but of the order of the loads within the spectrum as well. To assume you that this is a significant influence, let me remark that there are known instances where simply the reversal of the lead order of a spectrum, therefore with exactly the same maximum and minimum stress and oscillation frequencies, results in an alteration of the average rate of crack propagation by a factor of 3 (see [7]).

Moreover, there have been observed certain other non-intuitive effects of load order. These effects, called crack arrest and crack jump, are also significant. A quantified theoretical explanation of such a complex phenomenon, which has universal acceptance has not been attained. However, several partial attempts have been recently made ([3]).

As an additional complication, two identical loading spectra imposed on what are nominally identical specimens can result in lines differing by as much as a factor of ten because of environmental variation causing stress corrosion. This is because fatigue is a chemical, as well as physical, phenomena of the surface and the interaction between fatigue and corrosion plays an important role. In fact, anything that affects the chemistry of the surface such as the polish of the metal, the finish, or intermittant exposure to such agents as salt air or jet fuel vapors, must be taken into account.

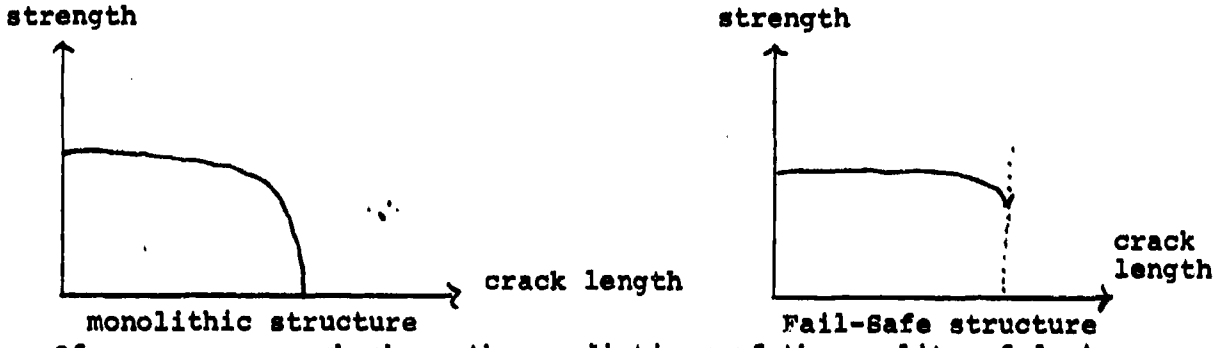
Thus, any accelerated life testing at elevated stress levels in dry laboratory air must consider the ambient conditions of the metallic surface during its service life in order to make valid predictions.

An alternative theory for calculating the crack length during the stable phase of crack growth following crack initiation is to use a

differential equation with two random parameters governing it behavior. These parameters are determined for each specimen by metalurgical measurement for one and the other from stress intensity factors calculated by Fracture Mechanics, geometry of the specimen and the type of metal. (see [ ]).

From the solution of this equation, we obtain the stochastic process, say $S(t-T_o)$ for $t>0$, which is the crack length at a time t with the initiation occurring at a time T.

The residual strength of a component which contains a crack of length $s>0$ is calculated theoretically from its design and gross area stresses. In a monolithic structure, the crack can attain a length for which the structural strength is zero. In a fail-safe structural design there are "crack stoppers" which arrest the crack at some size before the strength is reduced too far. We illustrate with two schematic plots of strength versus crack length:



Of course, as a check on the predictions of the quality of design concerning both the ultimate and fatigue strength of structural members, many critical components are fabricated and tested until failure occurs in one or the other mode.

I have mentioned all of these ancillary subjects, which are currently important in the accelerated life testing of an airframe, to illustrate my contention that in many instances in modern technology where accelerated life testing can be applied for the prediction of reliability that the statistical techniques for accelerated testing when they are properly integrated with the appropriate scientific knowledge from physics, chemistry and engineering provides a most powerful tool.

Conclusion

Statisticians working on reliability in industry during the past few years have been used like "hired guns" or "counselors". On retainer they were expected to collect data and advance statistical argument supporting the position that their patron advocated. The attempts to "prove" the reliability of certain products using statistical arguments has resulted in many statements such as "With $100(1-\varepsilon)\%$ confidence this component meets or exceeds its reliability goal of $1-\delta$" where ε and δ are both small.

Unfortunately in many cases the first n trials (at hundreds of thousands of dollars per trial) resulted in total failure. (Here n is large). In many instances the cause was subsequently determined to be an oversight in the design caused by disregarding or misinterpreting the results of testing. Of course, statisticians as a group were not more culpable than persons from other fields of knowledge but I think they were blaimed more frequently.

Clearly, not all of the difficulties of designing and constructing reliable systems can be cured by statisticians learnfng more engineering and engineers learning more statistics, but I believe nothing will be lost by such efforts. Furthermore, it is my personal, perhaps chauvinistic, belief that statisticians will finish this assignment first and that the profitable development of accelerated life testing procedures based on sound engineering knowledge with a thorough understanding of the physics of materials will be accomplished largely by their efforts. In any case, it is virtually certain that the development of such procedures constitutes a major need for statistical applications and presents a challenging task to scientific endeavor in the years ahead.

References

[1] Steck, G. P., and Zimmer, W. J. Estimation of Acceleration and Ageing Functions, 1972 NATO Conference Proceedings, Reliability Testing and Reliability Evaluation. Ernest M. Scherer - Editor, VII-B, Bureau of Business Services and Research. California State University - Northridge, 1972.

[2] Palakowski, N. H., and Ripling, E. J. Strength and Structure of Engineering Materials, Prentice Hall, Inc. Englewood Cliffs, New Jersey, 1966.

[3] Effects of Environment and Complex Load History on Fatigue Life, ASTM - STP- 462, American Society for Testing and Materials, 1910 Race Street, Philadelphia, PA 19103, 1970.

[4] Probabilistic Aspects of Fatigue, R. A. Heller - Editor, ASTM-STP - 511, American Society for Testing and Materials, 1910 Race Street, Philadelphia, PA 19103, 1972.

[5] Birnbaum, Z. W. and Saunders, Sam C., " A Probabilistic Interpretation of Miner's Rule", Siam J. Appl. Math 16, p. 637-652, 1968.

[6] Saunders, Sam C., " A Probabilistic Interpretation of Miner's Rule II". Siam J. Appl. Math 19, p. 251-265, 1970.

[7] McMillan, J. C., and Pelloux, R. M. M., "Fatigue Crack Propagation under Program and Random Loads". Fatigue Crack Propagation, ASTM - STP - 415, American Society for Testing and Materials, p. 505-532, 1907.

RELIABILITY OF MULTIPLE COMPONENT SYSTEMS

Larry Lee and W. A. Thompson, Jr.*
University of Missouri - Columbia

Abstract. Brindley and Thompson (1973) have introduced a multivariate concept of monotone failure rate.

Here we treat a mathematical model which has arisen in biological and engineering applications. In a biological context one refers to the theory of competing risks; in the engineering applications the model represents a non repairable series system. The system fails when the first of its components fails so that not all component lifetimes are observable.

Initially, we assume the components of the system to be independent. Then we introduce dependence in terms of sets of minima of independent random variables. The resulting multivariate distribution of component lifetimes generalizes Marshall & Olkin's multivariate exponential distribution but allows for the possibility of monotone failure rates.

The above dependence distribution is then derived through a "fatal shock" model where the shocks arrive according to a time dependent Poisson process. The failure rates of the component life times are determined by the intensity functions of the processes.

*This research was supported in part by Office of Naval Research Contract N00014-67-A-0287-004.

1. INTRODUCTION

Let $X_1, \ldots, X_k$ be theoretical failure times of k components of a series system; that is, if the system were to continue, the ith component would fail at time X_i. The system fails when the first component fails so that failure times of the other components become unobservable. Only the system failure time, $U = \min(X_1, \ldots, X_k)$, and the component or components which caused the failure are observable. The configuration of components causing the failure is called the failure pattern, we describe it in more detail later.

Problems of this type have arisen in two diverse applications. First, in the context of actuarial science, Cornfield (1957), Kimball (1958), Chiang (1968), and Berkson and Elveback (1960) use the "competing risk model" in the preparation of life tables for biological populations; Moeschberger and David (1971) discuss applications of the competing risk model and consider the problem of estimating parameters of the underlying life distributions.

Second, problems having the same mathematical structure occur in connection with the reliability and safety of engineering systems. Marshall and Olkin (1967), Arnold (1968), and Bemis, Higgins and Bain (1972) are papers which appear to be motivated by engineering applications.

1.1 Survival Times and Functions

Let X denote survival time, i.e., the length of time until a particular functioning object fails to function properly. Once the object fails it stays in that state, we are not considering it to be repairable. Except for the intuitive background, in this article one may think of survival time as meaning simply a non-negative random variable (r.v.). The survival function of X,

$$\bar{F}(x) = P[X > x]; \quad x \geq 0,$$

is the probability that the object survives at least x units of time.

As a consequence of the frequency intepretation of probability, $\bar{F}(x)$ is also the proportion of a large population which will survive till age x. Thus, as Grubbs and Shuford (1973) have done in constructing a probabilistic theory of combat, if interactions between the strengths of the two armies are ignored, then the proportions of combatants on each side surviving at time x can be estimated by $\bar{F}(x)$.

This article treats two or more survival times jointly, particularly when they are dependent. If X and Y are survival times, then

$$\bar{F}(x,y) = P[X > x, Y > y]; \quad x,y \geq 0,$$

is their joint survival function. Joint survival functions for more than two objects are defined in an analogous manner.

Possible applications of joint survival functions are suggested by the following examples. First, denoting the life times of husband and wife by X and Y, respectively, an insurance company selling an annuity will be interested in the bivariate survival function. Second, the two engines of a twin-engine airplane can fail separately or simultaneously; the joint survival function is

important in safety considerations. Third, for traffic congestion studies, one is interested in the time gaps between cars on a two lane or multilane highway.

The exponential distribution

$$F(x) = 0, \quad x < 0, \quad F(x) = 1 - e^{-\lambda x}, \quad x \geq 0 \tag{1.1}$$

has proved useful as a model for life testing, see Epstein and Sobel (1953), but it has a "no aging property" which is peculiar in this context. If the r.v. X is exponential then

$$P[X > x + \Delta \mid X > x] = P[X > \Delta]$$

for all $x \geq 0$, $\Delta > 0$. That is, in a probability sense, residual life is independent of age.

Obviously many objects age, i.e., become more prone to failure, as they become older. Some actually strengthen as they get older, e.g., some electronic circuits and many new mechanical devices.

The concept of failure rate plays a role at this point. Let X be a non-negative random variable with density $f(x)$, distribution function (d.f.) $F(x)$, and survival function $\bar{F}(x) = 1 - F(x)$. The <u>failure rate</u> is

$$r(x) = \frac{f(x)}{\bar{F}(x)} = -\frac{d}{dx}(\log \bar{F}(x)). \tag{1.2}$$

Alternatively we may write

$$\bar{F}(x) = \exp\left[-\int_0^x r(t)dt\right]. \tag{1.3}$$

The failure rate is useful and has a meaningful interpretation, for $r(x)\Delta x$ represents approximately the probability that an object of age x will fail in the interval $[x, x + \Delta x]$.

Barlow and Proschan (1965) introduce monotone failure rates as follows.

Definition. A nondiscrete univariate distribution $F(x)$ is IFR (DFR) if

$$P(X > x + \Delta \mid X > x) = \frac{\bar{F}(x+\Delta)}{\bar{F}(x)}$$

is decreasing (increasing) in x for every fixed $\Delta > 0$, $x \geq 0$ such that $\bar{F}(x) > 0$.

If $F(x)$ has a density and $F(0-) = 0$, then $F(x)$ being IFR (DFR) is equivalent to the failure rate $r(x)$ of (1.2) being increasing (decreasing).

Some distributions which have been important in life studies are i) the exponential with constant failure rate ii) the Weibull, with $r(x) = \rho\beta x^{\alpha-1}$ and iii) the Gompertz with $r(x) = B \exp(Cx)$; $B,C > 0$. Makeham's formula, $r(x) = A + B \exp(Cx)$; $B,C > 0$, has been important in the theory of life insurance.

1.2 Multivariate Exponential Distributions

Since the exponential distribution plays a crucial role in many univariate lifetime problems, we are concerned with multivariate extensions of it.

The simplest multivariate distribution with exponential marginals is composed of independent exponential distributions. With the multivariate hazard rate defined as $r(x_1, \ldots, x_k) = f(x_1, \ldots, x_k)/\bar{F}(x_1, \ldots, x_k)$, the hazard rate of independent exponentials is obviously constant. Basu (1971) shows that the only absolutely continuous bivariate distribution with exponential marginals and constant bivariate hazard rate is that of two independent exponentials.

Gumbel (1960) presents a bivariate distribution with exponential marginals and joint survival function

$$\bar{G}(x,y) = e^{-x-y-\delta xy}; \quad 0 \leq \delta \leq 1; \quad x,y \geq 0.$$

The coefficient of correlation for this bivariate distribution is either negative or zero.

Freund (1961) studies the following model. Suppose that two exponential lifetimes, with parameters α and β, function independently until the first failure. At failure the remaining lifetime becomes exponential with a new parameter, either α' replacing α or β' replacing β. This may realistically represent a situation where two components perform the same function, and the failure of one component puts additional responsibility on the remaining one.

Freund's distribution has the "no aging property"

$$\bar{F}(x + \Delta, y + \Delta) = \bar{F}(x,y) \cdot \bar{F}(\Delta,\Delta); \quad \Delta,x,y \geq 0.$$

But $F_1(x)$, the marginal distribution of X, is IFR (DFR) if and only if $\alpha < \alpha'(\alpha > \alpha')$ and similarly $F_2(y)$ is IFR (DFR) if and only if $\beta < \beta'(\beta > \beta')$. Since the hazard rate $r(x)$ of an exponential distribution is equal to its parameter, this result is intuitive. Note that the failure rates of the marginals can be increasing, decreasing, or even monotone in opposite directions.

Marshall and Olkin (1967) derive from three different models a bivariate distribution which has exponential marginals and joint survival function

$$\bar{F}(x,y) = \exp\{-\lambda_1 x - \lambda_2 y - \lambda_{12} \max(x,y)\};$$

$x,y \geq 0$; $\lambda_1,\lambda_2,\lambda_{12} \geq 0$. We call this class of distributions the bivariate exponential distribution, BVE, and its extension to n variables the multivariate exponential distribution, MVE.

Marshall and Olkin derive the BVE through a "fatal shock" model, a "non-fatal shock" model, and a "no aging" model. In the "fatal shock" model three independant Poisson processes, with parameters λ_1,λ_2, and λ_{12}, govern the respective occurrences of failures of component one, component two, or both components in a two component system. Their "no aging" model shows that, analogous to the univariate exponential distribution,

$$\bar{F}(x + \Delta, y + \Delta) = \bar{F}(x,y)\bar{F}(\Delta,\Delta); \quad \Delta > 0; \; x,y \geq 0;$$

i.e., $P[X > x + \Delta, Y > y + \Delta | X > x, Y > y] = P[X > \Delta, Y > \Delta]$, with exponential marginals, if and only if $F(x,y)$ is BVE. Allowing the Δ's to differ, they show that $\bar{F}(x + \Delta_1, y + \Delta_2) = \bar{F}(x,y)\bar{F}(\Delta_1, \Delta_2)$ for all positive Δ_1 and Δ_2 if and only if X and Y are independent exponential r.v.'s.

Marshall and Olkin find the d.f. (which has a line of singularity along the main diagonal of the first quadrant), the moment generating function, moments, and several characteristics of the BVE. For example they show that (X, Y) is BVE if and only if there exist independent exponential r.v.'s U, V and W such that $X = \min(U, W)$ and $Y = \min(V, W)$. Also if (X, Y) is BVE, then $\min(X, Y)$ is exponential.

Marshall and Olkin also have a complete discussion of the MVE, with survival function given by

$$\bar{F}(x_1, \ldots, x_k) = \exp(-\sum_{i=1}^{k} \lambda_i x_i - \sum_{i<j} \lambda_{ij} \max(x_i, x_j)$$

$$- \sum_{i<j<k} \lambda_{ijk} \max(x_i, x_j, x_l) - \ldots$$

$$- \lambda_{1\ldots n} \max(x_1, \ldots, x_k))$$

where the λ's are non-negative and not all zero.

2. DEPENDENCE AND AGING ASPECTS OF MULTIVARIATE SURVIVAL

The theory of monotone failure rate has proved useful as a probabilistic model of univariate survival time, particularly in reliability theory. The exponential distribution is important in this theory as the boundary between IFR and DFR distributions. In searching for multivariate extensions of the monotone failure rate idea, the "no aging" property makes it appealing to require that the boundary between multivariate IFR and DFR should be the class of MVE distributions. Brindley and Thompson (1973) obtain this result for the following generalization of the monotone failure rate concept. A multivariate d.f. $F(x_1, \ldots, x_k)$ defined on the positive orthant is IFR (DFR) if

$$\frac{P(X_1 > x_1 + \Delta, \ldots, X_k > x_k + \Delta)}{P(X_1 > x_1, \ldots, X_k > x_k)} = \frac{\bar{F}(x_1 + \Delta, \ldots, x_k + \Delta)}{\bar{F}(x_1, \ldots, x_k)}$$

is decreasing (increasing) in $x_1, \ldots, x_k$ for each $\Delta > 0$, and all $x_1, \ldots, x_k \geq 0$ such that $\bar{F}(x_1, \ldots, x_k) > 0$. The failure times (non-negative r.v.'s) $X_1, \ldots, X_k$ are jointly IFR (DFR) if the d.f. of each subset of them is IFR (DFR).

The point here is that it is possible for $\bar{F}(x_1, \ldots, x_k)$ to be increasing in each variable and yet some subset of $X_1, \ldots, X_k$

may have a marginal distribution which is not increasing in each variable. For example, Freund's bivariate exponential distribution has the no aging property and hence is IFR but, if $\alpha > \alpha'$, $F_1(x)$ will be DFR.

In the definition of jointly IFR, the requirement that each subset of the variables have a property is reminiscent of the definition of independent events

Harris (1970) defines a d.f. $F(x_1, \ldots, x_k)$ to be multivariate IHR if i. $F(x_1, \ldots, x_k)$ is IFR in the sense of the previous paragraph and ii. the variables $X_1, \ldots, X_k$ possess a positive dependence property called right corner set increasing (RCSI).

In the bivariate case, RCSI is the requirement that

$$P(X > x', Y > y' | X > x, Y > y)$$

be increasing in x and y. The RCSI property implies the series bound

$$\bar{F}(x,y) \geq \bar{F}_1(x)\ \bar{F}_2(y).$$

Harris obtains several results for IHR variables including the property that subsets of IHR r.v.s are IHR. This shows that multivariate IHR r.v.s are multivariate IFR. Gumbel's distribution is an example of IFR r.v.s which are not IHR; the series bound need not hold.

Positive dependence properties, like RCSI, will be reasonable for studying the life times of components all subjected to the same environment. But we may wish to study life times subject to different environments and there are several other types of positive dependence which imply the series bound and are as intuitively appealing as RCSI. For example, positive likelihood ratio

dependence and positive regression dependence, see Lehmann (1966) and Dykstra, Hewett, and Thompson (1973) also imply the series bound. Further the parallel definition of DHR is disappointing in that the boundary between IHR and DHR consists of independent exponential distributions

Finally, there is no reason why aging and positive dependence need go together. If X and Y are r.v.s uniformly distributed on the triangle with vertices (0,0), (0,1), and (1,0), then X and Y are jointly IFR but they exhibit a negative dependence property which we may call right corner set decreasing. Dependence and aging are in fact orthogonal properties.

Since dependence logically need not accompany monotone failure, such concepts need not be included in multivariate extensions of univariate monotone failure rate. Multivariate IFR and DFR as defined by Brindley and Thompson (1973) are strictly aging concepts which lead to a symmetric theory, and the MVE distributions form the boundary between them. Sets of minimums of IFR lifetimes are IFR, and Harris' IHR distributions form a substantial subclass of the IFR distributions.

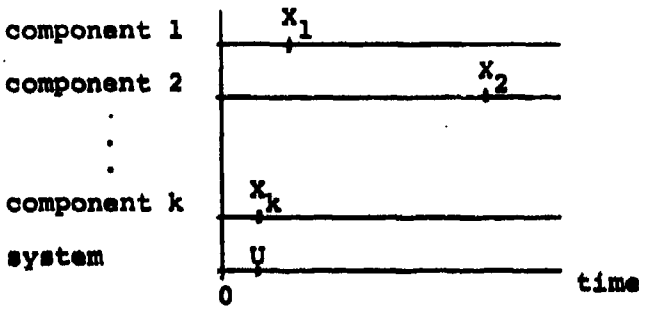
3. INDEPENDENT COMPONENTS

3.1 Independence Model

The initial systems to be considered are those consisting of independent components. The model is as indicated in Figure 1.

Figure 1

Theoretical Failure Times for Independent Components



Denoting the survival functions of U and X_i by $\bar{G}$ and $\bar{F}_i$ (i=1, ..., k), we have $U = \min(X_1, \ldots, X_k)$ and

$$\bar{G}(x) = \prod_{i=1}^{k} \bar{F}_i(x). \tag{3.1}$$

It is well known that, for independent components, system failure rate is the sum of component failure rates. In fact, from (1.2) and (3.1),

$$r_U(x) = -\frac{d}{dx}(\log \bar{G}(x)) = \sum_{i=1}^{k} -\frac{d}{dx}(\log \bar{F}_i(x)).$$

$$= \sum_{i=1}^{k} r_i(x). \tag{3.2}$$

The probability of tied values is zero so that the failure pattern is simply which one of the components causes the system to fail.

The joint probabilities of failure time and failure pattern are, for i=1, ..., k:

$$P(U > u, X_i = U) = P(u < X_i < \min_{j \neq i}(X_j)) = \int_u^\infty \bar{G}(x) r_i(x) dx.$$

The probability that the ith component causes the system to fail is $P(X_i = U) = \pi_i$, say.

$$\pi_i = \int_0^\infty \bar{G}(x) r_i(x) dx.$$

An example of the utility of these ideas appears in Vesely, Waite, and Keller (1971). They are concerned with the design of a safety system which will shut down an atomic reactor should it begin to go out of control. They consider a manual as well as an automatic system and for each, they estimate reliabilities from theoretical considerations. Estimated component reliabilities for the manual system appear in Table 1. From this Table they conclude that, effort to improve reliability of the manual system should center on relays and console switches; improvement of reliability of terminals and connectors, and wires does not pay off in improved system reliability.

Table 1 - Manual Control System

Component	Π's
Relays (8)	.6477
Console Switches (2)	.3076
Terminals and Connectors (27)	.0262
Wires (76)	.0185

The conditional survival function of system life given that the ith component caused failure is

$$\bar{G}(u|X_i = U) = \pi_i^{-1} \int_u^\infty \bar{G}(x) r_i(x) dx,$$

and the conditional density is

$$g(u|X_i = U) = \pi_i^{-1} \bar{G}(u) r_i(u). \qquad (3.3)$$

This is equation (2.5) of Moeschberger and David (1971). From equations (1.3) and (3.3) we obtain

$$\bar{F}_i(x) = \exp\left[-\pi_i \int_0^x \frac{g(u|X_i=U)}{\bar{G}(u)}\, du\right], \quad i=1, \ldots, k. \tag{3.4}$$

Thus, as Berman (1963) has observed, the distribution of failure time and failure pattern uniquely determines that of the component lifetimes.

3.2 Proportional Failure Rates

For two series systems of independent and identical components, consisting of k_1 and k_2 components respectively, then $r_1(x) = \frac{k_1}{k_2} r_2(x)$. In general we say that X and Y have proportional failure rates if there exists a constant $\theta > 0$ such that

$$r_X(x) = \theta r_Y(x) \tag{3.5}$$

for all $x \geq 0$.

The assumption of proportional failure rates for the component lifetimes of a series system has occurred several places in the literature. See Allen (1963), David (1970), Sethuraman (1965) and Nádas (1970). We may summarize the results concerning proportional failure rates as follows.

Theorem 1. For continuous and independent $X_1, \ldots, X_k$, the following are equivalent:

i) $X_1, \ldots, X_k$ have proportional failure rates

ii) $r_i(x) = \Pi_i \cdot r_U(x); \quad i=1, \ldots, k$ (3.6)

iii) $\bar{F}_i(x) = [\bar{G}_U(x)]^{\Pi_i}; \quad i=1, \ldots, k$ (3.7)

iv) failure time is independent of failure pattern and

v) there is a common transformation h so that $h(X_1), \ldots, h(X_K)$ are independent exponential r.v.s.

Proof. Clearly ii) implies i), but also i) implies ii). For if $r_i(x) = \theta_{ij} r_j(x)$ with $\theta_{ij} > 0$ for $i \neq j$ then $r_U(x) = r_j(x) \cdot \theta_{\cdot j}$ where $\theta_{\cdot j} = \sum_i \theta_{ij}$ and

$$\Pi_j = \int_0^\infty \bar{G}(x) r_j(x) dx = \frac{1}{\theta_{\cdot j}} .$$

The equivalence of ii) and iii) is a result of (1.2) and (1.3).

The equivalence of ii) and iv) follows from (3.3). We have

$$r_i(u) = \Pi_i \frac{g(u|X_i = U)}{\bar{G}(u)} .$$

Hence $r_i(u) = \Pi_i \cdot r_U(u)$ if and only, for i=1, ..., k, $g(u|X_i = U) = g(u)$, the density function of U.

Now iii) implies v) where the transformation h is given by

$$h(x) = \int_0^x r_U(t) dt.$$

Note that h is continuous and non-decreasing.
From (1.3),

$$\bar{F}_i(x) = [\bar{G}(x)]^{\Pi_i} = \exp[-\Pi_i h(x)].$$

Let $Y_i = h(X_i)$ and $h^{-1}(z) = \inf_x \{x : h(x) \geq z\}$.

$$\bar{F}_{Y_i}(y) = P(Y_i > y) = P[h(X_i) > y]$$

$$= P[X_i > h^{-1}(y)] = \bar{F}_i[h^{-1}(y)]$$

$$= \exp\{-\Pi_i h[h^{-1}(y)]\} = \exp(-\Pi_i y),$$

which is the survival function of an exponential r.v.

Finally v) implies i), since

$$\bar{F}_i(x) = P(X_i>x) = P[h(X_i)>h(x)]$$

$$= \bar{F}_{Y_i}(h(x)) = \exp\{-\theta_i h(x)\}.$$

where h is the assumed transformation. Thus

$$r_i(x) = -\frac{d}{dx}[\log \bar{F}_i(x)] = \theta_i \frac{dh}{dx}.$$

In the case of proportional failure rates we have

$$\bar{G} = \bar{F}_i^{\frac{1}{\pi_i}}, \quad i=1, \ldots, k.$$

If some π_i is small, then these equations make it appealing to assume that the distribution of U can be well approximated by a limiting extreme value distribution. For, if $\pi_i^{-1}=n$ then $\bar{G}=\bar{F}_i^n$ is the survival function of the minimum of n independent r.v.s. all having d.f. F_i.

The possibilities appear in Table 10.2 of Thompson (1969). The Cauchy type limit assigns no probability to positive values and hence is unacceptable as a distribution of failure time. The exponential type would imply that lifetimes could be negative as well as positive. Clearly the limited type with lower limit zero is the most appropriate choice of distribution. The limited type with that limit is the Weibull which has density.

$$w(u) = \rho\,\alpha\,u^{\alpha-1}e^{-\rho u^{\alpha}}, \quad \rho,\alpha > 0.$$

4. DEPENDENT COMPONENTS

4.1 Dependent Component Model

Marshall and Olkin characterize their MVE in terms of sets of minima of exponential r.v.s. We may use this idea as one way to introduce dependence among component lifetimes. The components

causing the system to fail can be indicated by a random vector $V = (V_1, \ldots, V_k)$ where V_i equals 1 or 0 according as the ith component is or is not failed. The sample space S of values which V can assume contains 2^k-1 elements since the zero vector is excluded. If V=s where $s \in S$ then we say that the system has exhibited failure pattern s. We assume a collection of independent and continuous r.v.s $\{Z_s : s \in S\}$ where Z_s is the theoretical time of occurence of failure pattern s.

Now, the theoretical failure time of the ith component is

$$X_i = \min_{\{s:s_i=1\}} (Z_s); \qquad i=1, \ldots, k \tag{4.1}$$

and system failure time is

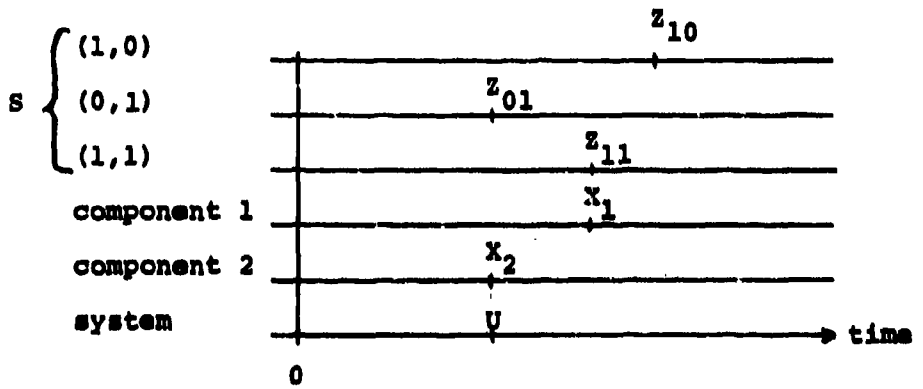
$$U = \min(X_1, \ldots, X_k) = \min_{s \in S}(Z_s). \tag{4.2}$$

We wish to observe that David (1973) has also suggested the model (4.1).

For the bivariate case, the model is indicated in Figure 2.

Figure 2

Theoretical Failure Times for Bivariate Dependent Components



Let $\bar{G}$ and $\bar{F}_s$ denote the survival functions of U and Z_s respectively, and let π_s be the probability of failure pattern s;

$$\bar{G}(z) = \prod_{s \in S} \bar{F}_s(z) \tag{4.3}$$

and

$$\pi_s = P(V = s) = P(Z_s = U) \tag{4.4}$$

Let r_U, r_i, and r_s be the failure rate functions of U, X_i, and Z_s respectively. We have

$$r_i(x) = \sum_{\{s:s_i=1\}} r_s(x) \quad \text{and} \quad r_U(x) = \sum_{S} r_s(x)$$

but there will be no general expression of r_U in terms of $\{r_i\}$.

A special case of an observation of Harris is that

$$\{X_1 > x_1, \ldots, X_k > x_k\} = \bigcap_{s \in S} \{Z_s > y_s\}$$

where $y_s = \max(x_1 s_1, \ldots, x_k s_k)$, $s \in S$. Hence

$$\bar{F}(x_1, \ldots, x_k) = \prod_{s \in S} \bar{F}_s(y_s). \tag{4.5}$$

Note that $X_1, \ldots, X_k$ have multivariate d.f.(4.5) if and only if there exists a collection of independent r.v.s $\{Z_s: s \in S\}$ such that $X_i = \min_{\{s:s_i=1\}}(Z_s)$. Hence, equation (4.5) is an alternative way of representing the dependent component model of this section.

The marginal distributions of (4.5) have the same form as the parent distribution. In fact

$$\bar{F}(x_1, \ldots, x_m, 0, \ldots, 0) = \prod_{s \in S} \bar{F}_s[\max(x_1 s_1, \ldots, x_m s_m, 0, \ldots, 0)]$$

$$= \prod_{s_1, \ldots, s_m} \bar{F}_{s_1, \ldots, s_m}[\max(x_1 s_1, \ldots, x_m s_m)]$$

where $\bar{F}_{s_1, \ldots, s_m}(x) = \prod_{s_{m+1}, \ldots, s_k} \bar{F}_s(x)$.

However, consider generating a bivariate Weibull distribution by taking

$$\bar{F}_s(t) = \exp[-\rho_s t^{\alpha_s}]$$

for $s = (0,1), (1,0), (1,1)$. We obtain

$$\bar{F}(x_1,x_2) = \exp\left\{-\rho_{10}x_1^{\alpha_{10}} - \rho_{01}x_2^{\alpha_{01}} - \rho_{11}[\max(x_1,x_2)]^{\alpha_{11}}\right\}. \qquad (4.6)$$

Note that (4.6) differs from the bivariate Weibull mentioned in Marshall and Olkin (1967) and discussed in Moeschberger (1974); the marginals are not Weibull.

For the conditional density of system failure time given failure pattern s, using (3.3), we obtain the expression

$$g(u|V=s) = \pi_s^{-1}\bar{G}(u) \cdot r_s(u). \qquad (4.7)$$

Using (4.7) and (1.3) we may write (4.5) in the alternative form

$$\bar{F}(x_1, \ldots, x_k) = \exp\left[-\sum_{s \in S} \pi_s \int_0^{y_s} \frac{g(u|V=s)}{\bar{G}(u)} du\right]$$

where $y_s = \max(x_1 s_1, \ldots, x_k s_k)$. Again, the distribution of failure time and failure pattern uniquely determines that of the component survival times.

Brindley and Thompson observe that sets of minimums of multivariate IFR(DFR) failure times are multivariate IFR(DFR). Hence if $\{Z_s: s \in S\}$ are univariate IFR(DFR) then $X_1, \ldots, X_k$ are multivariate IFR(DFR). For example, component failure times having the bivariate Weibull of (4.6) will be multivariate IFR if α_{10}, α_{01}, and α_{11} are all greater than 1.

4.2 Proportional Failure Rates

The assumption of proportional failure rates in the model (4.5) amounts to

$$r_s(x) = \pi_s \cdot r_U(x), \quad s \in S. \tag{4.8}$$

That is, the constants of proportionality are the probabilities of the failure patterns. Now from (3.2) and (4.1),

$$r_i(x) = \pi_i \cdot r_U(x) \tag{4.9}$$

where

$$\pi_i = \sum_{\{s:s_i=1\}} \pi_s = P(X_i = U).$$

By summing (4.9) we obtain, as a generalization of (3.2),

$$r_U(x) = \sum_{i=1}^{k} r_i(x) / \sum_{i=1}^{k} \pi_i.$$

With the additional assumption of (4.8), the dependent component model (4.5) becomes

$$\bar{F}(x_1, \ldots, x_k) = \prod_{s \in S} [\bar{G}(y_s)]^{\pi_s} \tag{4.10}$$

where $y_s = \max(x_1 s_1, \ldots, x_k s_k)$. This is a joint survival distribution, for the k components of the system, which is similar to that of Marshall and Olkin's MVE. In fact, if we take $\bar{G}(t) = \exp(-\lambda t)$ we obtain their MVE survival function:

$$\bar{F}(x_1, \ldots, x_k) = \exp[-\lambda \sum_{s \in S} \pi_s \max(x_1 s_1, \ldots, x_k s_k)].$$

The marginal distributions of (4.10) again satisfy (4.10) only in fewer variables. The π's have the same significance and even the d.f.G is the same. For example

$$\bar{F}(x_1, \ldots, x_m, 0, \ldots, 0) = \prod_{s \in S} \{\bar{G}[\max(x_1 s_1, \ldots, x_m s_m)]\}^{\pi_s}$$

$$= \prod_{s_1, \ldots, s_m} \{\bar{G}[\max(x_1 s_1, \ldots, x_m s_m)]\}^{\pi'_{s_1 \ldots s_m}}$$

where $\pi'_{s_1 \ldots s_m} = \sum_{s_{m+1}, \ldots, s_k} \pi_{s_1 \ldots s_m s_{m+1} \ldots s_k}$, the marginal probability of failure pattern $(s_1, \ldots, s_m)$ among the first m components.

Theorem 1 carries over directly to the dependent component model.

Theorem 2. For continuous and independent $\{Z_s : s \in S\}$, the following are equivalent:

i) $\{Z_s : s \in S\}$ have proportional failure rates

ii) $r_s(x) = \Pi_s \cdot r_U(x)$, $s \in S$ (4.8)

iii) $\bar{F}_s(x) = [\bar{G}_U(x)]^{\Pi_s}$, $s \in S$

iv) failure time is independent of failure pattern

and

v) there is a common transformation h so that $h(Z_s)$, $s \in S$, are independent exponential r.v.s.

Since $U = \min_{s \in S}(Z_s)$, and the events $\{V=s\}$ and $\{U=Z_s\}$ are equivalent, the proof is exactly as in Theorem 1 except that (4.7) is used instead of (3.3).

As an example of Theorem 2, consider the bivariate Weibull, (4.6). By calculating the failure rates, we see that failure time and pattern will be independent if and only if $\alpha_{10}=\alpha_{01}=\alpha_{11}$. Then from (4.8),

$\pi_{10}=\rho_{10}/(\rho_{10}+\rho_{01}+\rho_{11})$, $\pi_{01}=\rho_{01}/(\rho_{01}+\rho_{10}+\rho_{11})$, and $\pi_{11}=\rho_{11}/(\rho_{01}+\rho_{10}+\rho_{11})$.

As in the case of independence, since one of the π's will be small, the equations

$$\bar{G} = \bar{F}_s^{\pi_s^{-1}}, \quad s \in S$$

make it appealing to assume that the distribution of failure time is well approximated by a limiting extreme value distribution. The Weibull seems to be the most satisfactory of these.

In a sample of size N from (4.10), let U_j and $V^{(j)}$ denote the time of failure and the failure pattern for the jth observation, j=1, ..., N. These are the only observable quantities, and from them we might wish to make inferences about the distribution of component lifetimes. In such a sample, let N_s be the number of occurences of pattern s and let n_s be an observed value of N_s.

$$\sum_{s \in S} N_s = N.$$

Because of the independence between failure time and failure pattern, the joint density of the observations U_j and $V^{(j)}$, j=1, ..., N is

$$\prod_{j=1}^{N} g(u_j) \cdot \prod_{s \in S} \pi_s^{n_s}.$$

Incorporating the Weibull choice of G into this equation, the joint density of the observations becomes

$$(\rho\alpha)^N \prod_{j=1}^{N} u_j^{\alpha-1} e^{-\rho \sum_{j=1}^{N} u_j^{\alpha}} \cdot \prod_{s \in S} \pi_s^{n_s}, \tag{4.11}$$

where $\rho, \alpha > 0$ are the parameters of the Weibull.

Thus, with proportional failure rates, we may assume that system failure time has a Weibull distribution independent of failure pattern; and we may take observed failure patterns to have the multinominal distribution and be independent of the observed failure times.

If $\alpha=1$, then (4.5) yields the MVE distribution studied by Arnold (1968), and Bemis, Higgins and Bain (1972). Then $\sum_{i=1}^{N} U_i$ and $\{N_s\}$ are complete sufficient statistics and $\sum_{i=1}^{N} U_i$ has a gamma distribution.

5. TIME DEPENDENT FATAL SHOCK MODELS

Marshall and Olkin derive their MVE distribution from three points of view, including the "fatal shock" model. In the univariate case this model would hypothesize that shocks arrive according to a Poisson process and that the first shock destroys the object. Survival time would have the exponential distribution (1.1). But the exponential has the "no aging" property which is non-intuitive for many applications. How might we alter the fatal shock model to allow age dependent reliability behavior? For example, to model an IFR lifetime? One way is to allow the process controlling the arrival of the shocks to be a non-homogeneous or time dependent Poisson process.

A description of the time dependent Poisson process can be found, for example, in Parzen (1962). Let N(t) be the random number of shocks to the object in time t. The times at which the shocks occur are $\tau_1, \tau_2, \ldots$ where $0 < \tau_1 < \tau_2 < \ldots$. The inter

arrival times between shocks are

$$T_1 = \tau_1, \quad T_2 = \tau_2 - \tau_1, \ \ldots, \ T_n = \tau_n - \tau_{n-1}, \ \ldots .$$

Axioms of the Poisson process are:

Axiom 0. $N(0) = 0$.

Axiom 1. Independent increments: for all choice of indices $t_0 < t_1 < \ldots < t_n$ the random variables

$$N(t_1) - N(t_0), \ N(t_2) - N(t_1), \ \ldots \ , \ N(t_n) - N(t_{n-1})$$

are independent.

Axiom 2. For any $t > 0$, $0 < P\{N(t) > 0\} < 1$.

Axiom 3. For any $t \leq 0$

$$\lim_{h \to 0} \frac{P\{N(t+h) - N(t) \geq 2\}}{P\{N(t+h) - N(t) = 1\}} = 0.$$

Axiom 4. For some function $\nu(t)$, called the intensity function,

$$\lim_{h \to 0} \frac{1 - P\{N(t+h) - N(t) = 0\}}{h} = \nu(t).$$

These axioms imply that $N(t)$ has generating function

$$\Psi(z,t) = \exp[\lambda(t)(z-1)]$$

where

$$\lambda(t) = \int_0^t \nu(x)\,dx.$$

Since

$$P(T_1 > t) = P(N(t) = 0) = e^{-\lambda(t)}$$

we see that there is a one-one correspondence between the process and the distribution of T_1.

The function $\lambda(t)$ has the interpretation $\lambda(t) = E[N(t)]$. If $\nu(t)$ is constant we obtain the ordinary or homogeneous Poisson process where $\lambda(t) = \nu \cdot t$.

The time dependent Poisson process can be transformed into a homogeneous Poisson process. In fact, the process $\{M(u), u \geq 0\}$ defined by

$$M(u) = N(\lambda^{-1}(u)), \quad u \geq 0$$

is an ordinary Poisson process with

$$E(M(u)] = E[N(\lambda^{-1}(u))] = \lambda(\lambda^{-1}(u)) = u.$$

If shocks to an object are fatal then the survival time of the object is the arrival time of the first shock; $X = T_1$. If shocks arrive according to a time dependent Poisson process then the survival function for the object is

$$\bar{F}(x) = P(T_1 > x) = \exp(-\lambda(x)).$$

The failure rate of X is the intensity function of the process:

$$r(x) = -\frac{d}{dx}(\log \bar{F}(x)) = \nu(x).$$

Hence by specializing the intensity function, the fatal shock model yields all of the usual univariate life distributions and failure properties as special cases.*

The neatness of this result was not inevitable; it is a consequence of the particular model assumed. Renewal theory is another common way of modeling reliability problems. But as

* Conversations with Larry Crow and Lee Bain led to this observation.

Gnedenko, Belyayev, and Solovyev (1969, p. 105) point out one must be careful to distinguish the failure rate of X from the renewal density of the process.

Proceeding now to the multivariate case, we see that the distribution (4.5) can be derived from a fatal shock model. We consider that 2^k-1 independent random shock processes are operating and that a shock occurring in the process labeled $(s_1, \ldots, s_k)$ destroys those components i, for which $s_i=1$. The r.v. Z_s of Section 4.1 is the time of occurence of the first shock in the process labeled s. As before, with X_i defined by (4.1), the joint survival function (4.5) results.

If shocks arrive according to time dependent Poisson processes then (4.5) becomes

$$\bar{F}(x_1, \ldots, x_k) = \exp[- \sum_{s \in S} \int_0^{y_s} \nu_s(t)dt] \qquad (5.1)$$

where $y_s = \max (x_1 s_1, \ldots, x_k s_k)$, $s \in S$, and $\nu_s(t)$ is the intensity function of the process labeled s. Also $\nu_s(t) = r_s(t)$, the failure rate of Z_s.

We have already seen that the time dependent Poisson process can be made homogeneous by transformation. In the fatal shock model, when can we perform a single time transformation so that all shock processes are homogeneous? Answer: in the case of proportional failure rates.

To see this, first suppose that $r_s(x) = \Pi_s \; r_U(x)$, $s \in S$. Let $\lambda(t) = \int_0^t r_U(x)dx$ and $M_s(u) = N_s(\lambda^{-1}(u))$, $s \in S$. We have $E[M_s(u)] = E[N_s(\lambda^{-1}(u)] = \int_0^{\lambda^{-1}(u)} \nu_s(x)dx = \Pi_s \lambda (\lambda^{-1}(u)) = \Pi_s u$; and the $M_s(u)$, $s \in S$ are all homogeneous Poisson processes.

Conversely, suppose $M_s(u) = N_s(h(u))$, $s \in S$ are independent homogeneous Poisson processes.

$$P[h^{-1}(Z_s) > u] = P[Z_s > h(u)]$$

$$= P[N_s(h(u)) = 0] = P(M_s(u) = 0)$$

$$= \exp(-\theta_s u).$$

Theorem 2 then states that $\{Z_s : s \in S\}$ have proportional failure rates.

REFERENCES

Allen, W. (1963). A note on conditional probabilities when hazards are proportional. Opns. Res. 11 658-659.

Arnold, B. (1968). Parameter estimation for a multivariate exponential distribution. J. Amer. Stat. Assoc. 63 848-852.

Barlow, R. and Proschan F. (1965). Mathematical Theory of Reliability. John Wiley and Sons Inc., New York.

Basu, A. (1971). Bivariate failure rate. J. Amer. Stat. Assoc. 66 103-4.

Bemis, B.; Higgins, J. and Bain, L. (1972). Estimation and hypothesis testing for the parameters of a bivariate exponential distribution. J. Amer. Stat. Assoc. 67 927-929.

Berkson, J. and Elveback, L. (1960). Competing exponential risks with particular reference to the study of smoking and lung cancer. J. Amer. Stat. Assoc. 55 415-428.

Berman, S. (1963). Note on extreme values, competing risks and semi-markov processes. Ann. Math. Statist. 34 1104-1106.

Brindley, E.; and Thompson, W. (1973). Dependence and aging aspects of multivariate survival. J. Amer. Stat. Assoc. 67 822-830.

Chiang, C. (1968). Introduction to Stochastic Processes in Biostatistics. John Wiley & Sons, Inc., New York.

Cornfield, J. (1957). The estimation of the probability of developing a disease in the presence of competing risks Amer. J. Publ. Hlth. 47 601-607.

David, H. (1970). On Chiang's proportionality assumption in the theory of competing risks. Biometrics 26 336-339.

_____ (1973). Parametric approaches to the theory of competing risks. Proceedings of the 1973 Conference on Reliability and Biometry. Florida State University, Tallahasse, Florida.

Dykstra, R.; Hewett, J.; and Thompson, W. (1973). Events which are almost independent. Ann. Statist. 1 674-681.

Epstein, B. and Sobel, M. (1953) Life testing. J. Amer. Stat. Assoc. 48 486-502.

Freund, J. (1961). A bivariate extension of the exponential distribution. J. Amer. Stat. Assoc. 56 971-7.

Gnedenko, B.; Belyayev, Y.; and Solovyev, A. (1969). Mathematical Methods of Reliability Theory. Academic Press, New York.

Grubbs, F. and Shuford, J. (1973). A new formulation of Lanchester combat theory. Operations Research 21 926-941.

Gumbel, E. (1960). Bivariate exponential distributions. J. Amer. Stat. Assoc. 55 698-707.

Harris, R. (1970). A multivariate definition for increasing hazard rate distribution functions. Ann. Math. Statist. 41 713-717.

Kimball, A. (1958). Disease incidence estimation in populations subject to multiple causes of death. Bull. Inst. Stat. 36 193-204.

Lehmann, E. (1966). Some concepts of dependence. Ann. Math. Statist. 37 1137-53.

Marshall, A. and Olkin, I. (1967). A multivariate exponential distribution. J. Amer. Stat. Assoc. 62 30-44.

Moeschberger, M. (1974). Life tests under dependent competing causes of failure. Technometrics to be published.

Moeschberger, M. and David, H. (1971). Life tests under competing causes of failure and the theory of competing risks. Biometrics 27 909-933.

Nádas, A. (1970). On proportional hazard functions. Technometrics 12 413-416.

Parzen, E. (1962). Stochastic Processes. Holden-Day Inc., San Francisco.

Sethuraman, J. (1965). On a characterization of the three limiting types of the extreme. Sankhya A27 357-364.

Thompson, W. (1969). Applied Probability. Holt, Rinehart and Winston, Inc. New York.

Vesely, W.; Waite, E.; and Keller, R. (1971). Fault tree applications within the safety program of Idaho nuclear. Idaho nuclear corporation, Idaho Falls, Idaho.