

AD-A124 077

ROBUST MINIMUM DISTANCE ESTIMATION OF THE
FOUR-PARAMETER GENERALIZED GAMMA.. (U) AIR FORCE INST OF
TECH WRIGHT-PATTERSON AFB OH SCHOOL OF SYST..
K F SHUMAKER SEP 82 AFIT-LSSR-27-82

1/1

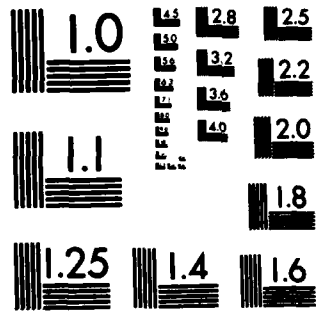
UNCLASSIFIED

F/G 12/1

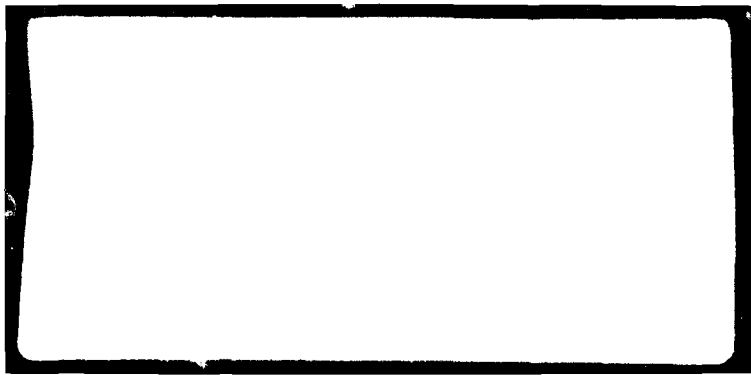
NL



END
SER.
FILMED
283
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



ROBUST MINIMUM DISTANCE ESTIMATION
OF THE FOUR-PARAMETER GENERALIZED
GAMMA DISTRIBUTION

Keith F. Shumaker, Captain, USAF

LSSR 27-82

DTIC
ELECTE
FEB 3 1983
S A D

This document has been approved
for public release and sale; its
distribution is unlimited.

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information are contained therein. Furthermore, the views expressed in the document are those of the author(s) and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the Air Training Command, the United States Air Force, or the Department of Defense.

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFIT thesis research. Please return completed questionnaires to: AFIT/LSH, Wright-Patterson AFB, Ohio 45433.

1. Did this research contribute to a current Air Force project?

- a. Yes b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not researched it?

- a. Yes b. No

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency received by virtue of AFIT performing the research. Can you estimate what this research would have cost if it had been accomplished under contract or if it had been done in-house in terms of manpower and/or dollars?

a. Man-years _____ \$ _____ (Contract).

b. Man-years _____ \$ _____ (In-house).

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3 above), what is your estimate of its significance?

- a. Highly Significant b. Significant c. Slightly Significant d. Of No Significance

5. Comments:

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification _____	
By _____	
Distribution/ _____	
Availability Codes	
Dist	Avail and/or Special
A	



Name and Grade

Position

Organization

Location

FOLD DOWN ON OUTSIDE - SEAL WITH TAPE

AFTT/LSB
WRIGHT-PATTERSON AFB OH 45433

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, \$300



NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST CLASS PERMIT NO. 7326 WASHINGTON D.C.

POSTAGE WILL BE PAID BY ADDRESSEE

AFTT/DAA

Wright-Patterson AFB OH 45433



FOLD IN

UNCLASSIFIED


SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER LSSR 27-82	2. GOVT ACCESSION NO. AD 4124 097	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ROBUST MINIMUM DISTANCE ESTIMATION OF THE FOUR-PARAMETER GENERALIZED GAMMA DISTRIBUTION		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Keith F. Shumaker, Captain, USAF		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Systems and Logistics Air Force Institute of Technology WPAFB OH 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Department of Communication and Humanities AFIT/LSH, WPAFB OH 45433		12. REPORT DATE September 1982
		13. NUMBER OF PAGES 63
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Approved for public release; LTR 278 1885. <i>John E. Wolaver</i> John E. WOLAYER Spec. for Research and Professional Development Air Force Institute of Technology (AFIT) Wright-Patterson AFB OH 45433 16 DEC 1982		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Reliability Analysis Minimum Distance Estimation Maximum Likelihood Estimation Life Distribution Generalized Gamma		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Thesis Chairman: Albert H Moore, Ph.D.		

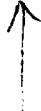
DD FORM 1473 1 JAN 73

EDITION OF 1 NOV 68 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)



A robust estimation technique (MLDE) is developed which uses minimum distance estimation in conjunction with maximum likelihood estimation (MLE). This technique is then applied to the four-parameter generalized Gamma distribution to obtain location, scale, shape, and power parameter estimates. A Monte Carlo analysis is conducted on three members of the four-parameter generalized Gamma distribution with sample sizes of 12, 16, 20, and 24 for a total of twelve cases. For each of these twelve cases, one thousand samples are generated for the analysis. Initial estimates of the location, scale, shape, and power parameters are found using a maximum likelihood estimator. Minimum distance estimation using the Anderson-Darling statistic is then employed to obtain a new estimate of the location parameter. Finally, this new improved location parameter estimate is used to refine the scale, shape, and power parameter estimates through maximum likelihood estimation. The performance of the MLDE technique is determined through use of mean square error and relative efficiency measures.



LSSR 27-82

ROBUST MINIMUM DISTANCE ESTIMATION OF THE
FOUR-PARAMETER GENERALIZED GAMMA DISTRIBUTION

A Thesis

Presented to the Faculty of the School of Systems and Logistics
of the Air Force Institute of Technology
Air University

In Partial Fulfillment of the Requirement for the
Degree of Master of Science in Systems Management

By

Keith F. Shumaker, BS
Captain, USAF

September 1982

Approved for public release;
distribution unlimited

This thesis, written by

Captain Keith F. Shumaker

has been accepted by the undersigned on behalf of the faculty
of the School of Systems and Logistics in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS MANAGEMENT

DATE: 29 September 1982


COMMITTEE CHAIRMAN

ACKNOWLEDGEMENTS

No work of this magnitude is ever accomplished through the efforts of just the author. I wish to express my gratitude to those persons who willingly gave of their time and expertise. Sincere thanks go to Dr. Albert H. Moore for his assistance in the completion of this thesis. A statistics course being taught by him stimulated my interest in estimation techniques and established the foundation on which this thesis was built.

A special note of thanks goes to my loving wife Susan whose patient and skilled editing helped me focus my thoughts. Susan's assistance in the typing of this thesis allowed me more time for research and resulted in a high quality product. The support and encouragement that she has given me has been helpful beyond words.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
 Chapter	
I. INTRODUCTION	1
Background	1
Objective	2
II. FOUR-PARAMETER GENERALIZED GAMMA DISTRIBUTION.	4
Generalized Gamma Function	4
History	4
Four-Parameter Generalized Gamma Cumulative Distribution	6
III. ESTIMATION BY MAXIMUM LIKELIHOOD METHODS	9
Estimators	9
Maximum Likelihood Estimators	9
IV. ROBUST MINIMUM DISTANCE ESTIMATION	13
Robust Estimation	13
Minimum Distance Estimation	13
Anderson-Darling Statistic	14
V. MONTE CARLO ANALYSIS	16
Overview	16
Generation of Sample Data	16
Estimation Technique (MLDE)	18

Chapter	Page
Estimation Technique (MLE)	19
Comparison of Estimation Techniques	19
VI. RESULTS AND CONCLUSIONS	21
Results	21
Conclusions	22
Recommendations	23
APPENDICES	25
A. RELATIVE EFFICIENCIES OF MLDE ESTIMATES AND ANDERSON-DARLING STATISTICS	26
B. AVERAGE MEAN SQUARE ERRORS	39
SELECTED BIBLIOGRAPHY	52
A. REFERENCES CITED	53
B. RELATED SOURCES	55

LIST OF TABLES

Table	Page
2-1. Special Cases	7
5-1. Combinations And Respective Parameter Values	17
A-1 thru A-12. Relative Efficiencies of MLDE Estimates and Anderson-Darling Statistics .	27-38
B-1 thru B-12. Average Mean Square Errors	40-51

LIST OF FIGURES

Figure	Page
2-1. Typical Graphs For $f(x;c,a,b,p)$	8

CHAPTER I

INTRODUCTION

Background

In recent years, statistical estimation has played an ever-increasing role in the reliability analysis of weapon systems. Of primary concern is the study of life distributions from which predictions of system and component failure can be made. Once failure test data from a particular weapon system is collected, the parameters of the appropriate probability model can be estimated and the system's probable life span determined. Unfortunately, the test data in such reliability analyses is often incompatible with the Exponential, Weibull, and other familiar probability models (19:349). As time and money for weapon systems development are placed under tighter constraints, the statistical community is faced with a challenge to find the best probability models for particular sets of test data. It is also tasked with finding the most efficient techniques of parameter estimation for specific models (4:1).

This thesis will focus on the generalized Gamma distribution which was first presented by Stacy in 1962 (18:1187). This distribution has since been recognized and used as a valuable life-testing model (6:1601; 8:159). With the modern

electronic computational capabilities now available, statisticians can expand and modify the classical methods of estimation such as method of moments and maximum likelihood. It is also possible to develop iterative techniques, as well as, utilize more recent methods such as minimum distance estimation.

Objective

The objective of this thesis is to develop a robust estimation procedure which uses minimum distance estimation in conjunction with a modified version of the maximum likelihood estimation (MLE) routine developed by Harter and Moore (8:159-165). This is an attempt to find a faster and more accurate iterative technique for parameter estimation of the four-parameter generalized Gamma distribution. Throughout the thesis, this new robust estimation procedure will be referred to as the MLDE technique. Estimates from the MLDE technique will be compared with those obtained from the original version of Harter's MLE technique to evaluate the degree to which the objective has been achieved. If the objective is successfully accomplished, the resulting estimation technique (MLDE) could also yield improved estimates for the guaranteed life of other distributions being used as life-testing models. The MLDE technique could then be used to reduce the data-gathering and computational costs of weapon system life-testing studies. This thesis is an extension of

the work by Drs. Albert H. Moore and H. Leon Harter and their
thesis students in the field of parameter estimation (2; 4;
7; 8; 10; 14).

CHAPTER II

FOUR-PARAMETER GENERALIZED GAMMA DISTRIBUTION

Generalized Gamma Function

The four-parameter generalized Gamma probability density function is defined as

$$f(x;c,a,b,p) = \frac{p(x-c)^{bp-1} \exp\{-[(x-c)/a]^p\}}{a^{bp} \Gamma(b)} \quad (2.1)$$

where $a, b, p \geq 0$ and $x \geq c \geq 0$

The location parameter c , also known as guaranteed life, is the value of x where the distribution begins to have a non-zero value. The relative scale of the distribution along the x -axis is determined by a , the scale parameter. The shape of the distribution is determined by the shape/power parameter b and the power parameter p .

History

In 1962, Stacy presented the generalized Gamma distribution and studied the properties of the three-parameter case (18:1187). Later, Stacy and Mihram made a further generalization by including cases in which the power parameter p was negative (19:351). For the purpose of this thesis and most life distribution studies, the restrictions of equation

2.1 are imposed since negative values of parameters p and c are not applicable when modeling life distributions. Stacy and Mihram also studied some of the basic properties of the density and considered parameter estimation by method of moments, maximum likelihood, and minimum variance with primary emphasis on the scale parameter a (19:352-355). Bain and Weeks, using one parameter at a time as being unknown, developed one-sided tolerance limits and confidence limits (1:1142). Parr and Webster obtained expressions for the maximum likelihood estimators of the shape/power parameter b and the shape parameter d , where $d=bp$ (16:2-3). They assumed the generalized Gamma distribution to be the correct life-testing model and provided tests for rejecting densities such as the Weibull and Exponential where they are not adequate models (6:1601). Harter added a fourth parameter, the location parameter c , which the previous authors had assumed to be zero. With the addition of this fourth parameter, he hoped to enhance the usefulness of the generalized Gamma population in the study of life distributions (8:160). Harter formulated an iterative procedure for maximum likelihood estimation of the four-parameter generalized Gamma population using methods previously used for the three-parameter Gamma population. He also gave mathematical expressions and tables for the asymptotic variances and covariances of the maximum likelihood estimators (7:Section 5).

Four-Parameter Generalized Gamma Cumulative Distribution

The four-parameter generalized Gamma cumulative distribution is given by

$$\begin{aligned} F(x) &= \int_0^x f(x;c,a,b,p) dx \\ &= \Gamma_{\omega}(b)/\Gamma(b) \end{aligned} \quad (2.2)$$

where $\omega = [(x-c)/a]^p$ and $p > 0$

The name "generalized" Gamma is suggested by the fact that the cumulative distribution function is an incomplete Gamma-function ratio.

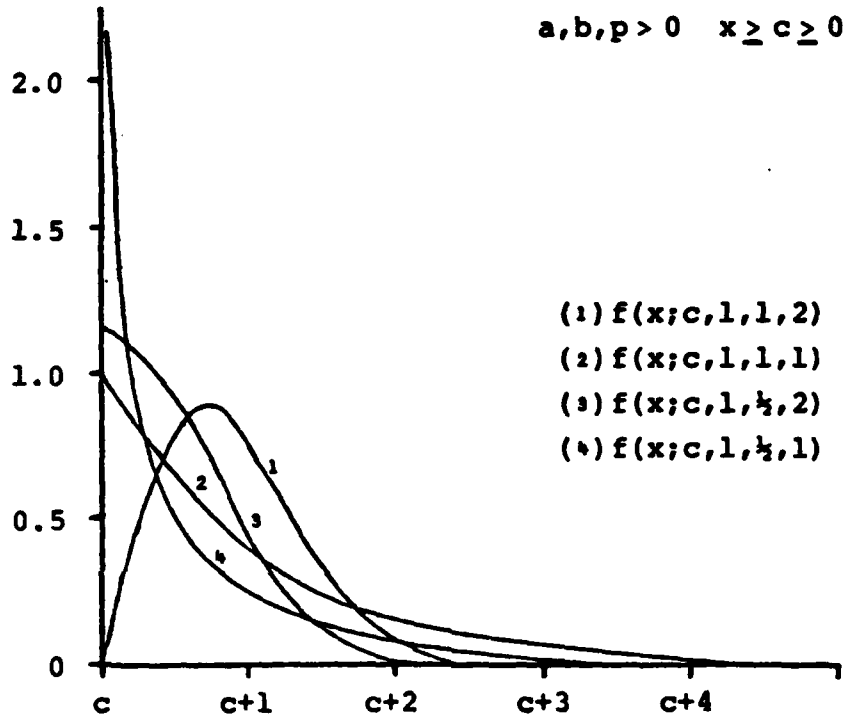
Special cases of the four-parameter generalized Gamma function include the two-parameter Exponential, the three-parameter Weibull and a variety of other well known probability functions. Table 2-1 lists some of these with the corresponding notation for their frequency functions. Letting the location parameter c take on the value of zero, the graphs of $f(x;c,a,b,p)$ can take a variety of shapes where $p > 0$. This can be seen from the graphs in Figure 2-1 where changes in shape are dramatically effected by varying the values of the product pb (19:352).

TABLE 2-1
Special Cases (19:351)

TYPE	FREQUENCY FUNCTION	RESTRICTIONS
Exponential	$f(x;c,a,1,1)$	$x \geq c \geq 0 \quad a > 0$
Gamma	$f(x;c,a,b,1)$	$x \geq c \geq 0 \quad a,b > 0$
Weibull	$f(x;c,a,1,p)$	$x \geq c \geq 0 \quad a,p > 0$
Half-Normal	$f(x;c,\sqrt{2},1/2,2)$	$x \geq c \geq 0$
Chi Squared	$f(x;c,2,n/2,1)$	n degrees of freedom, $c = 0$

FIGURE 2-1

Typical Graphs For $f(x;c,a,b,p)$ (19:352)



CHAPTER III

ESTIMATION BY MAXIMUM LIKELIHOOD METHODS

Estimators

Parameter estimation is the process of finding approximations of the true values of the parameters of a probability distribution. These approximations, also called estimates, are found by using rules called estimators. These rules tell "how to calculate an estimate based on the measurements contained in a sample [12:292]." A sample from a four-parameter generalized Gamma population will be generated and calculations will be performed on the data according to the rules of the particular estimator being used. The results of these calculations will be the estimates of the four parameters of the four-parameter generalized Gamma distribution. The estimators used in this research are a modified maximum likelihood estimator (MLE) and a minimum distance estimator (MLDE).

Maximum Likelihood Estimators

In the method of maximum likelihood, the parameter values which maximize the likelihood or joint density of the sample are used as estimates (12:262). Since the sample is selected randomly, the joint density of the sample is the

product of the probability distribution evaluated at each of the sample points (12:228). For the four-parameter generalized Gamma density of n observations, the likelihood is defined by the formula

$$\begin{aligned}
 L &= f(x_1; c, a, b, p) f(x_2; c, a, b, p) \cdots f(x_n; c, a, b, p) \\
 &= \prod_{i=1}^n f(x_i; c, a, b, p) \\
 &= \prod_{i=1}^n \frac{p(x_i - c)^{bp-1} \exp\{-[(x_i - c)/a]^p\}}{a^{bp} \Gamma(b)} \quad (3.1)
 \end{aligned}$$

To find the parameters of the distribution which maximize the likelihood function, the first partial derivatives of L with respect to each of the four parameters are set equal to zero. Usually the natural logarithm of L is taken before maximizing as this transforms the product into a sum which is easier to differentiate. The natural logarithm of L is a monotonically increasing function of L thus both $\ln L$ and L will be maximized for the same values (12:363). The natural logarithm of the likelihood function, the lowest r and the highest $(n-m)$ sample values having been censored, is given by

$$\begin{aligned}
 L = L_{r+1, m} &= \ln(n)! - \ln(r)! - \ln(n-m)! + (m-r) [\ln(p) - \ln(a)] \\
 &+ (bp-1) \sum_{i=r+1}^m [\ln(z_i)] - \sum_{i=r+1}^m [z_i^p] - (n) \ln \Gamma(b) \\
 &+ (r) \ln \Gamma(b; z_{r+1}^p) + (n-m) \ln [\Gamma(b) - \Gamma(b; z_m^p)] \quad (3.2)
 \end{aligned}$$

where $z_i = (x_i - c)/a$

If there is no censoring from above, the last term

$$(n-m) \ln[\Gamma(b) - \Gamma(b; z_m^p)] \quad (3.3)$$

drops out. If there is no censoring from below, the next to the last term

$$(r) \ln \Gamma(b; z_{r+1}^p) \quad (3.4)$$

drops out.

The first partial derivatives of $L = L_{r+1, m}$ with respect to a , b , p , and c are given by

$$\begin{aligned} \partial L / \partial a = & a^{-1} \{ -bp(m-r) + p \Sigma [z_i^p] - rz_{r+1} f(z_{r+1}) / F(z_{r+1}) \\ & + (n-m) z_m f(z_m) / [1 - F(z_m)] \} \end{aligned} \quad (3.5)$$

$$\begin{aligned} \partial L / \partial b = & p \Sigma [\ln(z_i)] - n \Gamma'(b) / \Gamma(b) + r \Gamma'(b; z_{r+1}^p) / \Gamma(b) F(z_{r+1}) \\ & + (n-m) [\Gamma'(b) - \Gamma'(b; z_m^p)] / \Gamma(b) [1 - F(z_m)] \end{aligned} \quad (3.6)$$

$$\begin{aligned} \partial L / \partial p = & (m-r) / p + b \Sigma [\ln(z_i)] - \Sigma [z_i^p \ln(z_i)] + r f(z_{r+1}) \ln(z_{r+1}) / \\ & p F(z_{r+1}) - (n-m) f(z_m) \ln(z_m) / p [1 - F(z_m)] \end{aligned} \quad (3.7)$$

$$\begin{aligned} \partial L / \partial c = & a^{-1} \{ (1-bp) \Sigma [z_i^{-1}] + p \Sigma [z_i^{p-1}] - r f(z_{r+1}) / F(z_{r+1}) \\ & + (n-m) f(z_m) / [1 - F(z_m)] \} \end{aligned} \quad (3.8)$$

where the primes in equation (3.6) indicate differentiation with respect to parameter b and where

$$f(z_i) = pz_i^{bp-1} \exp(-z_i^p) / \Gamma(b), \quad F(z_i) = \int_0^{z_i} f(t) dt = \Gamma(b; z_i^p) / \Gamma(b)$$

By setting the first partial derivatives of the likelihood

function equal to zero, the maximum likelihood estimates of the four parameters a , b , p , and c are obtained. In this case, the estimator could not be solved in closed form (19:352). Solutions can be found by iteration; and the iterative technique developed by Harter will be used to solve these equations for the maximum likelihood estimates (8:Section 6). Throughout this thesis, any references to maximum likelihood estimation (MLE) will denote the use of Harter's technique.

CHAPTER IV

ROBUST MINIMUM DISTANCE ESTIMATION

Robust Estimation

Estimation techniques which are reasonably insensitive to underlying assumptions were labeled as having the quality of "robustness" by Box (3:318). Much of the research in the area of robust estimation was initiated by the concern among the statistical community that the classical assumptions, such as normality, were sometimes not totally valid (5:30). One of the areas of robust analysis is the study of procedures for finding robust parameter estimates. A robust estimator can adapt to deviations in the underlying model and remain efficient (17:3). This thesis will deal with methods using estimators which utilize the information provided by the sample data to obtain estimates for the suspected underlying model (14:2-3).

Minimum Distance Estimation

Parr and Schucany expanded the concept of robust techniques into minimum distance estimation (17). Most of the initial work in minimum distance estimation was done by Wolfowitz in the early 1950's (22:75; 21:9). He outlined the minimum distance method and showed it to be consistent in a

wide variety of cases even when classical methods failed to give consistent estimates (21:9). Further research was done by Knüsel in 1961, and by Parr and Schucany in 1979 (11:Ch.1; 17). In 1978, Parr published an accumulative minimum distance bibliography covering the works of most researchers in the field of minimum distance estimation (15:Ch.1).

Minimum distance estimation requires that a family of distribution functions $F(x;\theta)$ be specified. It also requires a rule for obtaining the empirical distribution function (EDF), denoted $S_n(x)$, which is merely the distribution function of a particular sample. Finally, it requires a measure of the distance between $F(x;\theta)$ and $S_n(x)$. Minimum distance estimation then takes those values for the parameter θ which minimize the distance between $F(x;\theta)$ and $S_n(x)$. In this thesis, the scale, shape, and power parameter estimates, previously found by MLE, will be used with minimum distance estimation employing the Anderson-Darling statistic to estimate θ where θ is the location parameter c .

Anderson-Darling Statistic

The Anderson-Darling, A^2 , statistic is a special case of the Weighted Cramer-von Mises estimator and takes the form

$$A_n^2(S_n, F_\theta) = \int_{-\infty}^{\infty} [S_n(x) - F_\theta(x)]^2 [F_\theta(x)(1 - F_\theta(x))]^{-1} dF_\theta(x) \quad (4.1)$$

(17:7-8)

The A^2 statistic causes more emphasis to be placed on the

tails of the distribution and is noted as one of the more efficient for estimating the location parameter (20:735).

Stephen's analytical equation for A^2 is

$$A^2 = - \left[\frac{\sum_{i=1}^n (2i-1) [\ln(z_i) + \ln(1-z_{n+1-i})]}{n} \right] - n \quad (4.2)$$

where $z_i = F(x_i; \theta) = F_{\theta}(x_i)$ (20:731)

In the above equation, z_i is the standardized order statistics and n is the number of order statistics.

CHAPTER V

MONTE CARLO ANALYSIS

Overview

The performance of the MLE and MLDE technique is evaluated by use of a Monte Carlo analysis. The Control Data Corporation (CDC) computer systems, located at the Aeronautical Systems Division, Wright-Patterson Air Force Base, Ohio, was used in performing this analysis. The analysis of each technique involves three distinct steps. First, random samples from the selected underlying distribution are generated. Second, estimates of the parameters of this distribution are found by using an estimation technique on each sample. Third, these parameter estimates are compared to the true parameters of the distribution.

Generation of Sample Data

The comparisons in this thesis are made for a number of sample sizes and several different combinations of parameter values. Sample sizes of 12, 16, 20, and 24 are generated; and three combinations of parameters are used with each of these sample sizes. Table 5-1 is a table of these three combinations and their respective parameter values. For each of the twelve cases (four sample sizes times three combina-

TABLE 5-1

Combinations And Respective Parameter Values

PARAMETER	COMBINATION 1	COMBINATION 2	COMBINATION 3
Location	20	30	10
Scale	50	100	100
Shape	3	2	1
Power	1	1	2

tions), 1000 samples are generated.

The rejection method is used to accomplish the actual generation of the samples. This method makes use of the probability density function rather than the cumulative distribution function (13:118). To use this method, two subroutines from the International Mathematical Statistics Library (IMSL) are employed. The subroutine GGUBS is used to generate the needed random numbers. Once the data is generated by the rejection method and placed in an array, the subroutine VSRTA is used to sort the array in ascending order. Specific information about all IMSL subroutines can be found in the IMSL manual (9:Vol.1).

Estimation Technique (MLDE)

The MLDE technique is applied to every case for each ordered sample size. This technique consists of the following steps:

1. Initial estimates of the location, shape, scale, and power parameters are found using a modified version of Harter's MLE technique assuming all parameters are unknown. This modified version relaxes the tolerance limits of Harter's original version from $1D+07$ to $1D+03$.

2. With the estimates for the location, shape, scale, and power parameters found in step 1, minimum distance estimation using the A^2 statistic is employed to obtain a new estimate of the location parameter.

3. This new improved location parameter estimate from step 2 is used with the shape, scale, and power parameter estimates from step 1 in the modified MLE technique in step 1 to get new shape, scale, and power parameter estimates.

Estimation Technique (MLE)

The MLE technique is simply the original version of the iterative technique developed by Harter to find maximum likelihood estimates as described in Chapter III of this thesis (11:Section 6).

Comparison of Estimation Techniques

The performance of the MLE and the MLDE techniques will be determined through the use of mean square error (MSE) and relative efficiency (REFF) measures. The mean square error is defined as

$$MSE = \left[\sum_{i=1}^n (\theta_i - \theta)^2 \right] / n \quad (5.1)$$

where θ is the true value of the parameter, θ_i is the i -th estimate, and n is the number of times the estimation technique is applied. In this analysis, n is equal to 1000. When comparing MSE's for different parameter values, it should be noted that MSEs are not scale invariant. The same size MSE may be highly significant for small value estimates, but insignificant for larger ones (2:32).

The relative efficiency is defined as

$$REFF = MSE_i / MSE_j \quad (5.2)$$

where MSE_i is the mean square error of the MLE technique which is the base estimator being used for comparison and MSE_j is the mean square error of the MLDE technique which is the estimator being tested. Values greater than one indicate that the MLDE technique is more efficient than the MLE technique (10: 23,25).

CHAPTER VI

RESULTS AND CONCLUSIONS

Results

The numerical results of the Monte Carlo analysis are displayed in the appendices. There are twelve tables in each appendix, one for each case of sample sizes and parameter values. Appendix A contains tables of relative efficiencies and values of the Anderson-Darling statistics for the two estimation techniques. Appendix B displays the average mean square errors for the four parameters. The most apparent result is that for the sample sizes analyzed in this thesis, MLDE yields better estimates than MLE with the exception of the shape parameter estimates. Use of the Anderson-Darling statistics resulted in consistent improvement of the location parameter estimate over that of the MLE technique. Using this improved location parameter estimate and assuming all other parameters unknown, the MLDE technique also produced improved scale and power parameters.

The mean Anderson-Darling statistic was used as a means of comparison since it is an overall measure of how well the estimator performed. As seen in Appendix A, the MLE's Anderson-Darling statistic improves as the sample size of the particular probability density function increases. This

is expected and is primarily due to the fact that MLE is a strong asymptotic estimation procedure which performs better with increasing information. As seen from the MLDE's Anderson-Darling statistics, MLDE out performed MLE for all cases. The MLDE technique yielded an average computer time savings of sixteen percent over that of Harter's original MLE technique. Even with this time savings, the computer runs which accomplished the Monte Carlo analysis in this thesis involved a total of 180,000 seconds of processing time. In order to get such a large amount of computer resources, it was necessary to make successive runs during a time period of almost eleven months.

Conclusions

The objective of this thesis which was to find a faster and more accurate iterative technique for parameter estimation of the four-parameter generalized Gamma distribution has been successfully accomplished. For the cases analyzed in this thesis, MLDE provides estimates which closer fit the true distribution. MLDE is not meant to replace maximum likelihood techniques, but to be used as a refinement of those parameter estimates found by such techniques. MLDE requires less time than the MLE technique developed by Harter; and this savings of time, and thus computer resources, could play an important part in promoting recognition of minimum distance estimation methods. The average relative efficiencies fur-

ther indicates that over the entire range of sample sizes investigated in this thesis, the minimum distance estimation technique performed very well with the exception of the shape parameter estimate. It should be noted that the shape parameter estimate may appear distorted, however, sometimes this allows the overall estimated distribution to better fit the true distribution as indicated by the Anderson-Darling statistics. MLDE worked well considering the sizes of the samples and the fact that four parameters were estimated simultaneously. Despite the difficulty in estimating the shape and power parameters simultaneously due to a high negative correlation between them, minimum distance estimation did yield improved parameter estimates when applied to the four-parameter generalized Gamma distribution as compared to those found using the classical MLE estimator (8:163).

Recommendations

Due to the apparent lack of improvement in the estimates of the shape parameter when using the MLDE technique, a different technique should be studied to obtain an improvement. One approach would be the use of minimum distance estimation to obtain an improved shape parameter estimate as was done with the location parameter estimate in this thesis. This thesis dealt with sample sizes of 12, 16, 20, and 24. Further analysis should examine larger sample sizes to see if

MLDE continues to perform as well once the MLE technique starts to yield better estimates due to increased information from the larger samples.

APPENDICES

APPENDIX A

RELATIVE EFFICIENCIES OF MLDE ESTIMATES
AND ANDERSON-DARLING STATISTICS

TABLE A-1

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	12
Location	20
Scale	50
Shape	3
Power	1

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.1589	1.5975	.9783	1.0798

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
2.5576	1.3724

TABLE A-2

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	16
Location	20
Scale	50
Shape	3
Power	1

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.2176	1.0375	1.3148	1.1358

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
1.7192	.4526

TABLE A-3

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	20
Location	20
Scale	50
Shape	3
Power	1

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.1068	1.0873	1.0030	1.0897

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
.8026	.2687

TABLE A-4

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	24
Location	20
Scale	50
Shape	3
Power	1

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.1165	1.1046	.9770	1.1167

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
.7044	.2636

TABLE A-5

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	12
Location	30
Scale	100
Shape	2
Power	1

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.2786	1.0006	.7145	1.1594

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
2.5996	1.2705

TABLE A-6

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	16
Location	30
Scale	100
Shape	2
Power	1

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.1999	1.0012	.8902	1.1549

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
1.0608	.3792

TABLE A-7

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	20
Location	30
Scale	100
Shape	2
Power	1

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.3411	1.0043	.8071	1.1657

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
.9830	.3218

TABLE A-8

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	24
Location	30
Scale	100
Shape	2
Power	1

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.1390	1.0785	.9687	1.0984

ANDERSON-DARLING STATISTIC

<u>MLE</u>	<u>MLDE</u>
.6119	.2781

TABLE A-9

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	12
Location	10
Scale	100
Shape	1
Power	2

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.1455	1.6659	.3404	1.2893

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
2.5944	1.4522

TABLE A-10

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	16
Location	10
Scale	100
Shape	1
Power	2

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.1122	.9437	.9148	1.0913

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
.8855	.3111

TABLE A-11

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	20
Location	10
Scale	100
Shape	1
Power	2

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.0978	1.0001	.9945	1.0954

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
.5592	.2869

TABLE A-12

RELATIVE EFFICIENCIES
OF MLDE ESTIMATES

Monte Carlo	1000
Sample Size	24
Location	10
Scale	100
Shape	1
Power	2

<u>LOCATION</u>	<u>SCALE</u>	<u>SHAPE</u>	<u>POWER</u>
1.0915	1.0459	1.3357	1.1125

ANDERSON-DARLING STATISTICS

<u>MLE</u>	<u>MLDE</u>
.5407	.2711

APPENDIX B
AVERAGE MEAN SQUARE ERRORS

TABLE B-1

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	12
Location	20
Scale	50
Shape	3
Power	1

PARAMETER	MLE	MLDE
LOCATION	2054.55	1772.79
SCALE	7945010	4973260
SHAPE	24.5988	25.1438
POWER	73.1364	67.7283

TABLE B-2

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	16
Location	20
Scale	50
Shape	3
Power	1

PARAMETER	MLE	MLDE
LOCATION	1035.16	850.170
SCALE	8827.35	8508.45
SHAPE	3.28562	2.49892
POWER	.345935	.304579

TABLE B-3

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	20
Location	20
Scale	50
Shape	3
Power	1

PARAMETER	MLE	MLDE
LOCATION	1335.52	1206.65
SCALE	26389.0	24268.5
SHAPE	23.1567	23.0870
POWER	10.0422	9.21519

TABLE B-4

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	24
Location	20
Scale	50
Shape	3
Power	1

PARAMETER	MLE	MLDE
LOCATION	1083.81	970.706
SCALE	25589.4	23167.1
SHAPE	17.1962	17.6006
POWER	8.61239	7.71263

TABLE B-5

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	12
Location	30
Scale	100
Shape	2
Power	1

PARAMETER	MLE	MLDE
LOCATION	1736.82	1358.38
SCALE	6896990	6893010
SHAPE	4.42252	6.18980
POWER	6.82063	5.88276

TABLE B-6

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	16
Location	30
Scale	100
Shape	2
Power	1

PARAMETER	MLE	MLDE
LOCATION	1339.57	1116.44
SCALE	2980210	2976620
SHAPE	9.59494	10.7790
POWER	8.41545	7.28639

TABLE B-7

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	20
Location	30
Scale	100
Shape	2
Power	1

PARAMETER	MLE	MLDE
LOCATION	1130.46	842.914
SCALE	1020410	1016040
SHAPE	2.63146	3.26058
POWER	6.06143	5.19982

TABLE B-8

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	24
Location	30
Scale	100
Shape	2
Power	1

PARAMETER	MLE	MLDE
LOCATION	942.403	827.419
SCALE	48180.3	44672.5
SHAPE	3.57009	3.68541
POWER	11.2675	10.2577

TABLE B-9

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	12
Location	10
Scale	100
Shape	1
Power	2

PARAMETER	MLE	MLDE
LOCATION	621.546	542.614
SCALE	4903050	2943140
SHAPE	2.88021	8.46044
POWER	7.69565	5.96874

TABLE B-10

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	16
Location	10
Scale	100
Shape	1
Power	2

PARAMETER	MLE	MLDE
LOCATION	509.463	458.052
SCALE	2321.95	2460.50
SHAPE	7.91291	8.64999
POWER	8.80151	8.06528

TABLE B-11

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	20
Location	10
Scale	100
Shape	1
Power	2

PARAMETER	MLE	MLDE
LOCATION	286.557	261.030
SCALE	1741.65	1741.40
SHAPE	1.45520	1.46331
POWER	12.4612	11.3761

TABLE B-12

AVERAGE MEAN SQUARE ERRORS

Monte Carlo	1000
Sample Size	24
Location	10
Scale	100
Shape	1
Power	2

PARAMETER	MLE	MLDE
LOCATION	300.712	275.501
SCALE	2108.66	2016.02
SHAPE	110.632	82.8294
POWER	12.6957	11.4117

SELECTED BIBLIOGRAPHY

A. REFERENCES CITED

1. Bain, L. J., and D. L. Weeks. "Tolerance Limits for the Generalized Gamma Distribution," Journal of the American Statistical Association, LX (December 1965), pp. 1142-1152.
2. Bertrand, 2nd Lieutenant David E., USAF. "Comparison of Estimation Techniques for the Four-Parameter Beta Distribution." Unpublished master's thesis. AFIT/GOR/MA/81D-1, AFIT/EN, Wright-Patterson AFB OH, December 1981, AD A115562.
3. Box, G. E. P. "Non-Normality and Tests on Variances," Biometrika, XL (June 1953), pp. 318-335.
4. Daniels, Captain Tony G., USAF. "Robust Estimation of the Generalized t Distribution Using Minimum Distance Estimation." Unpublished master's thesis. AFIT/GOR/MA/80D-2, AFIT/EN, Wright-Patterson AFB OH, December 1980.
5. Geary, R. C. "Testing for Normality," Biometrika, XXXIV (January 1947), pp. 209-242.
6. Hager, H. W., and L. J. Bain. "Inferential Procedures for the Generalized Gamma Distribution," Journal of the American Statistical Association, LXV (December 1970), pp. 1601-1609.
7. Harter, H. L. "Asymptotic Variances and Covariances of Maximum Likelihood Estimators, from Censored Samples, of the Parameters of a Four-Parameter Generalized Gamma Population," ARL-66-0158, Aerospace Research Laboratories, Wright-Patterson AFB OH, (1966), AD 648045.
8. _____. "Maximum Likelihood Estimation of the Parameters of a Four-Parameter Generalized Gamma Population from Complete and Censored Samples," Technometrics, IX (February 1967), pp. 159-165.
9. IMSL Library Reference Manual. Houston: IMSL, Inc., 1980.

10. James, Captain William L., USAF. "Robust Minimum Distance Estimation Based on a Family of Three-Parameter Gamma Distributions." Unpublished master's thesis. AFIT/GOR/MA/80D-4, AFIT/EN, Wright-Patterson AFB OH, December 1980.
11. Knüsel, L. F. "Über Minimum-Distanz-Schätzungen." Unpublished doctoral dissertation, Swiss Federal Institute of Technology, 1969.
12. Mendenhall, William, Richard L. Scheaffer, and Dennis D. Wackerly. Mathematical Statistics with Applications. 2nd ed. Boston: Duxbury Press, 1982.
13. Mihram, G. A. Simulation: Statistical Foundations and Methodology. San Francisco: Academic Press, 1972.
14. Miller, Captain Robert M., USAF. "Robust Minimum Distance Estimation of the Three-Parameter Weibull Distribution." Unpublished master's thesis. AFIT/GOR/MA/80D-7, AFIT/EN, Wright-Patterson AFB OH, December 1980.
15. Parr, Van B., and J. T. Webster. "A Method for Discriminating Between Failure Density Functions Used in Reliability Predictions," Technometrics, VII (February 1965), pp. 1-10.
16. Parr, William C. "Minimum Distance and Robust Estimation." Unpublished doctoral dissertation, Department of Statistics, Southern Methodist University, 1978.
17. _____ and William R. Schucany. "Minimum Distance and Robust Estimation," Journal of the American Statistical Association, LXXV (September 1980), pp. 616-624.
18. Stacy, E. W. "A Generalization of the Gamma Distribution," Annals of Mathematical Statistics, XXXIII (September 1962), pp. 1187-1192.
19. _____ and G. A. Mihram. "Parameter Estimation for a Generalized Gamma Distribution," Technometrics, VII (August 1965), pp. 349-358.
20. Stephens, M. A. "EDF Statistics for Goodness of Fit and Some Comparisons," Journal of the American Statistical Association, LXIX (September 1974), pp. 730-737.
21. Wolfowitz, J. "Estimation by the Minimum Distance Method," Annals of the Institute of Statistical Mathematics, V (1953), pp. 9-23.

22. _____ . "The Minimum Distance Method," Annals of Mathematical Statistics, XXVIII (March 1957), pp. 75-88.

B. RELATED SOURCES

- Andrews, D. F., and others. Robust Estimation of Location. Princeton NJ: Princeton University Press, 1972.
- Blackman, J. "On the Approximation of a Distribution Function by an Empirical Distribution," Annals of Mathematical Statistics, XXVI (1955), pp. 256-267.
- Box, G. E. P., and George C. Tiao. "A Further Look at Robustness Via Baye's Theorem," Biometrika, IL (1962), pp. 419-431.
- Beyer, William H., ed. CRC Handbook of Tables for Probability and Statistics. (Second Edition), Cleveland: The Chemical Rubber Company, 1968.
- Deming, W. E. The Gamma and Beta Function. Ann Arbor MI: Edwards Brothers, Inc., 1946.
- Deutsch, Ralph. Estimation Theory. Englewood Cliffs NJ: Prentice-Hall Inc., 1965.
- Diggle, Peter J. "Robust Density Estimation Using Distance Methods," Biometrika, LXII (1975), pp. 39-48.
- Easterling, Robert G. "Goodness of Fit and Parameter Estimation," Technometrics, XVIII (February 1976), pp. 1-9.
- Hogg, Robert V. "Adaptive Robust Procedures: A Partial Review and Some Suggestions for Future Applications and Theory," Journal of the American Statistical Association, LXIX (December 1974), pp. 909-927.
- Kaper, K. C., and L. R. Lamberson. Reliability in Engineering Design. New York: John Wiley & Sons, Inc., 1977.
- Matusita, Kameo. "On the Estimation by the Minimum Distance Method," Annals of the Institute of Statistical Mathematics, V (1954) pp. 59-65.
- McNeese, Captain Larry B., USAF. "Adaptive Minimum Distance Estimation Techniques Based on a Family of Generalized Exponential Power Distribution." Unpublished master's thesis. AFIT/GOR/MA/80D, AFIT/EN, Wright-Patterson AFB OH, December 1980.

Pearson, E. S. and N. W. Please. "Relation Between the Shape of Population Distributions and the Robustness of Four Sample Test Statistics," Biometrika, LXII (1975), pp. 223-242.

Rey, William J. Robust Statistical Methods. New York: Springer-Verlag, 1978.

Sähler, W. "Estimation by Minimum-Discrepancy Methods," Metrika, XVI (1970), pp. 85-106.

Stigler, Stephen M. "Simon Newcomb, Percy Daniell, and the History of Robust Estimation, 1885-1920," Journal of the American Statistical Association, LXVIII (December 1973), pp. 872-879.

Walpole, Ronald E. and R. H. Myers. Probability and Statistics for Engineers and Scientists. (Second Edition), New York: MacMillan Publishing Co., Inc., 1978.