# DETERMINATION OF SAMPLE SIZE FOR SPECIFIED WIDTH CONFIDENCE INTERVALS IN DIGITAL SIMULATION EXPERIMENTS

By

BELVA JONES COOLEY

Bachelor of Science
McNeese State University
Lake Charles, Louisiana
1969

Master of Business Administration
Texas Tech State University
Lubbock, Texas
1970

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1976

# DETERMINATION OF SAMPLE SIZE FOR SPECIFIED WIDTH CONFIDENCE INTERVALS IN DIGITAL SIMULATION EXPERIMENTS

Thesis Approved:

Donald R Williams
Thesis Adviser

Wm Mendenhall

P. Larry Claypool

Gerald M Lage

N N Durham
Dean of Graduate College

# PREFACE

This study shows the development of a sample size determination method which may be used in experiments for which the data are exponentially or geometrically distributed. In addition, the procedure is shown to be valid for experimental data which are autocorrelated. The procedure can be applied to either physical or simulation experiments. Finally, the procedure is shown to be a practical one in that it may be easily calculated or easily incorporated into an on-going simulation experiment through the use of a subroutine.

I wish to thank my major advisor, Dr. Donald R. Williams, for his guidance and encouragement throughout this project. Appreciation is also expressed to the other committee members, Dr. Wayne Meinhart, Dr. Gerald Lage, and Dr. Larry Claypool, for their assistance in the study. Thanks is extended to Ms. Patty Barry and Mrs. Kathy Snyder for their assistance in typing the drafts and final copy of the manuscript.

Finally, special gratitude is expressed to my husband, John, and to our parents, Mr. and Mrs. Frank B. Jones and Mrs. Richard J. Cooley, for the support and encouragement that only a family can give.

TABLE OF CONTENTS


| Chapter | | Page |
|---|---|---|
| I. | INTRODUCTION | 1 |
| II. | REVIEW OF LITERATURE | 7 |
| III. | A SAMPLE SIZE METHOD FOR EXPONENTIAL DATA | 33 |
| IV. | COMPARISON OF SAMPLE SIZE METHOD FOR NORMAL DATA | 66 |
| V. | SUMMARY, CONCLUSIONS AND EXTENSIONS | 81 |
| A SELECTED BIBLIOGRAPHY | | 89 |
| APPENDIXES | | 93 |
| | APPENDIX A - THE PROPOSED METHOD FOR EXPONENTIAL DATA: ALGORITHM AND FORTRAN SUBROUTINE | 94 |
| | APPENDIX B - A PROGRAM FOR TESTING THE PROPOSED METHOD | 100 |
| | APPENDIX C - A PROGRAM FOR TESTING FISHMAN'S METHOD FOR DETERMINING SAMPLE SIZE | 105 |
| | APPENDIX D - STATISTICAL VALIDITY | 113 |
| | APPENDIX E - RESULTS OF EXPERIMENTS USING HIGHER AUTOREGRESSIVE ORDERS | 116 |

LIST OF TABLES


| Table | | Page |
|---|---|---|
| I. | Results Of Sample Size Experiments For Exponential Data With Autoregressive Order Of Zero. . . . . | 47 |
| II. | Results Of Sample Size Experiments For Exponential Data With Autoregressive Order Of One . . . . . | 48 |
| III. | Results Of Sample Size Experiments For Exponential Data With Autoregressive Order Of Two . . . . . | 49 |
| IV. | Comparison Of Time Requirements And Core Requirements For Sample Size Experiments Using Exponential Data . . . . . | 57 |
| V. | Evaluation Of Probability Statements . . . . | 58 |
| VI. | Results Of Sample Size Experiments For Geometric Data With Autoregressive Order Of Zero. . . . . | 61 |
| VII. | Results Of Sample Size Experiments For Geometric Data With Autoregressive Order Of One . . . . . | 62 |
| VIII. | Results Of Sample Size Experiments For Geometric Data With Autoregressive Order Of Two . . . . . | 63 |
| IX. | Results Of Sample Size Experiments For Normal Data With Autoregressive Order Of Zero. . . . . | 70 |
| X. | Results Of Sample Size Experiments For Normal Data With Autoregressive Order Of One . . . . . | 71 |
| XI. | Results Of Sample Size Experiments For Normal Data With Autoregressive Order Of Two . . . . . | 72 |

| Table | | Page |
|---|---|---|
| XII. | Evaluation Of Probability Statements: Normal Data. . . . . | 77 |
| XIII. | Selected Distributions Of Sample Size: Validation Of Sample Size And Distribution . . . . . | 115 |
| XIV. | Results Of Sample Size Experiments For Exponential Data With Autoregressive Order Of Three . . . . . | 117 |
| XV. | Results Of Sample Size Experiments For Exponential Data With Autoregressive Order of Five. . . . . | 118 |
| XVI. | Results Of Sample Size Experiments For Exponential Data With Autoregressive Order of Ten . . . . . | 119 |
| XVII. | Results Of Sample Size Experiments For Geometric Data With Autoregressive Order of Three . . . . . | 120 |
| XVIII. | Results Of Sample Size Experiments For Geometric Data With Autoregressive Order Of Five. . . . . | 121 |
| XIX. | Results Of Sample Size Experiments For Geometric Data With Autoregressive Order Of Ten . . . . . | 122 |

LIST OF FIGURES


| Figure | | Page |
|---|---|---|
| 1. | Histogram Of Sample Sizes Obtained With Fishman's Method And an Autoregressive Order Of Two. . . . . | 51 |
| 2. | Histogram Of Sample Sizes Obtained With The New Method: $\beta^2$ of .90 And Autoregressive Order Of Two . . . . . | 52 |
| 3. | Histogram Of Sample Sizes Obtained With The New Method: $\beta^2$ Of .95 And Autogressive Order Of Two . . . . . | 53 |
| 4. | Histogram Of Sample Sizes Obtained With The New Method: $\beta^2$ of .99 And Autoregressive Order Of Two . . . . . | 54 |
| 5. | Average Sample Sizes For Exponential Data Using The New Method. . . . . | 55 |
| 6. | Flowchart Of Sample Size Method For Exponential Data. . . . . | 97 |
| 7. | FORTRAN Subroutine For Computing Sample Sizes For Exponential Data. . . . . | 98 |
| 8. | FORTRAN Program For Testing The Sample Size Method Developed For Exponential Data . . . . | 103 |
| 9. | FORTRAN Program For Testing Fishman's Method For Computing Sample Size . . . . . | 109 |

# CHAPTER 1

# INTRODUCTION

The use of computer simulation as a tool for performing industrial and business-oriented experiments is fast becoming an important part of operations analysis. There are many problems, intractable using conventional analytical methods, that are particularly suited for analysis through the vehicle of computer simulation (e.g., complex multi-channel queuing problems, some reliability analysis, and a variety of inventory problems). It is expedient, therefore, that research be effected that is addressed to the subject of computer simulation experimentation and its potential as a decision-oriented tool.

Of particular importance to the researcher who uses computer simulation is the analysis of data generated by simulations. Because of characteristics inherent in digital simulations (e.g., non-normality, autocorrelation, and use of random numbers), conventional methods for experimental design and analysis are inappropriate for experimental data generated by computer simulations.

## Definition of Terms

Simulation. Simulation is most often construed as a

process of performing experiments on a model of reality rather than on an actual system. The simulation process, therefore, enables the analyst to gain information about systems as they would perform over time.

Monte Carlo Techniques. The term, Monte Carlo Techniques, is sometimes used as a synonym for simulation. In this study, the term is used more specifically to indicate experimentation on random numbers.

Sample Size Determination. The phrase, sample size determination, as used in this study, means that a procedure must be executed which indicates how many observations from an universe should be examined in order to adequately estimate a parameter of the population with a specified level of statistical precision.

Autocorrelation. Most conventional statistical techniques require that, in a data series, the value of any one observation is uninfluenced by the values for other observations. The term, autocorrelation, means that the values of the observations of a series of data are not independent of one another.

## Problem Definition

Determination of sample size requirements for digital simulation experiments is difficult because most data generated by simulations tend to be autocorrelated and tend to fit distributions other than the normal distribution. Several adequate sample size methods exist for

experiments in which the data are normal and independent. However, few methods have been devised which overcome the characteristics of non-normality and autocorrelation.

Simulation practitioners have approached the sample size problem in various ways. Many practitioners have simply ignored the lack of independence and proceeded as though the observations were not correlated. Others have suggested making many runs with different random number sequences and appealing to the Central Limit Theorem in order to view the sample means as normal and independent. This dependency upon the Central Limit Theorem, however, makes necessary the gathering of an unusually large number of observations for an adequate analysis of the data. Since digital simulation experiments are ordinarily quite large and expensive to execute, gathering a large number of observations becomes prohibitive when cost is of importance.

The purpose of this study is to develop, describe and demonstrate a valid method of determining sample sizes. As developed, the method will be efficient, where efficiency is defined in terms of the number of observations required for a given simulation experiment. Specifically, the method is particularly appropriate for data which are autocorrelated and which fit distributions frequently encountered in simulation experiments.

## Scope Of The Study

This study is limited to the development of a sample size determination procedure which is based on a method developed by Franklin A. Graybill. In his paper, Graybill outlines the conditions which must be met in order to develop a sample size determination method for distributions other than normal. His examples, however, pertain only to the normal distribution and there is no evidence in the literature that others have expanded his work to distributions other than normal. In addition, his procedure is based on an assumption of independence.

This study will expand Graybill's work by deriving a sample size determination procedure for the exponential and geometric distributions. The study will also show that the procedure may be used for autocorrelated data. In addition, experiments will be performed to validate the procedure and to compare the procedure to one developed by George S. Fishman.

Finally Graybill's method of normal data will be augmented with the procedure used by Fishman to adjust for autocorrelation. Tests will be run to validate the resulting procedure when used on autocorrelated data.

## Methodology

Validation of the method developed for computing sample size requirements for simulation experiments

generating exponentially or geometrically distributed data will employ a Monte Carlo approach. Sample size requirements will be computed for a data set which will be randomly generated to fit a specific distribution and to exhibit a specific degree of autocorrelation. The required number of observations will be taken as a sample and an interval estimate will be computed of the population mean. This procedure will be repeated a number of times and information will be gathered concerning: (1) the proportion of trials in which the confidence interval actually contained the true value of the mean; and (2) the proportion of trials in which the width of the confidence interval is less than or equal to the desired width.

The same procedure will be employed for Fishman's method using both exponential, geometric and normal data. The testing procedure will also be used for Graybill's method as it will be revised for autocorrelated, normal data.

## Chapter Organization

Chapter II describes the most relevant problems encountered when conducting an experiment using digital simulation techniques. The problem of sample size determination is shown to be a significant problem. Existing sample size determination techniques are described and the strengths and weaknesses of each are discussed.

Chapter III provides a detailed description of a pro-

cedure by which appropriate sample sizes for digital simulation can be determined, given that the data is exponentially or geometrically distributed. The technique is also shown to be successfully applied to autocorrelated data. In addition, this chapter explains the conceptual considerations which must be met if an appropriate sample size is to be determined with this procedure. The results of experiments designed to compare this procedure to one introduced by Fishman are also included in this chapter. Finally, a subroutine of the procedure is included so that the determination of sample size can become an integral part of a simulation experiment written in FORTRAN or GASP.

Chapter IV shows how the procedure introduced by Graybill can be augmented with a process designed to adjust for autocorrelation. Graybill's procedure can then be used on normally distributed data which is autocorrelated. Results of comparisons of Graybill's method with Fishman's method are reported in this chapter.

Chapter V reiterates the conclusions which may be made based on the data generated in the study. Possible extensions of the study and other areas of research related to the study are also discussed in this chapter.

# CHAPTER II

## REVIEW OF LITERATURE

Many definitions have been proposed for the term "simulation". Burdick and Naylor have defined it as "a numerical technique for conducting experiments on certain types of mathematical and logical models describing the behavior of a system (or some component thereof) on a digital computer over extended periods of real time."[1] Ackoff says simply that simulation is "to duplicate the essence of a system of activity without actually attaining reality itself."[2]

Within these two definitions are contained both the salient points of a definition for digital computer simulation and the attributes which make simulation an important tool for analysis. First, simulation requires that a model be built that adequately describes the operation of the system under study. In this way, systems which are too complex for mathematical analysis may be examined. Conway, et al., discuss the various aspects of model building that should be considered.[3] Secondly, experiments may be run to investigate the operation of the system for given configurations of components. As such, simulation is not an optimizing device but a search

procedure in which the configuration of components is varied in a number of experiments in order to find a maximum for those particular configurations. Finally, the added dimension of time as a variable enables the practitioner to incorporate a dynamic characteristic which is missing in most analytical techniques.

Because of the above attributes, the concept of computer simulation for discrete systems has become an important investigative tool for management scientists. Practitioners and thoreticians alike are viewing simulation as an acceptable method for approaching large scale, complex problems for which no analytical solutions are available.

## Problems in Simulation Experiments

As an experiment, computer simulation usually consists of the following nine steps which have been outlined by Burdick and Naylor:

1. Formulation of the problem.
2. Collection and processing of real world data.
3. Formulation of a mathematical model.
4. Estimation of the parameters of the operating characteristics of the model.
5. Evaluation of the model and parameter estimates.
6. Formulation of a computer program.
7. Validation.
8. Experimental design.
9. Analysis of simulation data.[4]

Thus a simulation experiment is quite similar to a physical experiment and, as such, is subject to all the problems of

a physical experiment. In addition, a simulation model exhibits problems brought about by the inherent characteristics of simulation. These difficulties can be categorized as problems of validation, experimental design, and sample size. Research is being actively conducted on all three fronts. The literature concerning validation, and experimental design will be discussed in this section. Sample size will be discussed in the following section.

## Validation

Validation of any experiment may be an onerous task in that inference to a population must be made on the basis of experiments on an environment which is abstracted from reality. For example, a plot of ground treated with various types of commercial fertilizer may be the environment of an agricultural experiment. On the basis of results from this experiment, the analyst draws inference to a broader spectrum of soils and conditions. The analyst using a simulation experiment also makes inference to the real world. However, his results may be viewed as more suspect since he is drawing inference from a mathematical model. Van Horn lists some of the primary reasons for this suspicion.

1. Simulations are often large and complex.
2. Including many processes, simulations allow these parts to interact in non-linear, non-stationary ways.
3. The assumptions may not be easily found.
4. For the uninitiated, simulations may appear as very close representations of

reality.[5]

Validating the model is viewed as a two-step process. The first step is to determine whether the model is internally correct in a logical and programming sense. Next, the analyst must determine whether it represents the phenomena it is supposed to represent. Meier, Newell and Pazer suggest the following tests as possible methods for uncovering defects in the model:

1. Run the model for a short time period on a small number of transactions so that results can be compared with hand calculations.
2. Run separate segments of complicated models alone so that results can be verified.
3. Eliminate random elements from stochastic models and run them as deterministic models.
4. Replace complex distributions with elementary ones so that results are more easily verified.
5. Construct simple test situations that test as many combinations of circumstances in the model as is feasible.[6]

A somewhat different approach is proposed by Schmidt and Taylor.[7] They suggest analytically checking results against theoretical values. Obviously simulation experiments would not be carried out if analytical techniques were readily available. In many cases, however, the model is relatively straight forward but the underlying empircal data fit unusual distributions. If this be the case, then a more appropriate distribution may be substituted for purposes of verification and the results checked against theoretical values. Similarly, authors have attempted verification of models through goodness-of-fit tests on

the resulting data.[8]

The analyst must be able to show that his model is a reasonably valid representation of the real system. One method that has been suggested is to show that the results of the simulation are consistent with results from the real system for a specific set of conditions.[9] The obvious shortcoming for this approach is the analyst's inability to check simulation results for every set of conditions against empirical data. Thus, the analyst would rarely, if ever, prove that a simulation model is a correct or "true" model of the actual system in question. Burdick and Naylor concluded statements on validation in this fashion:

> ...We consider the problem of validating computer models to be the most difficult unresolved problem facing individuals concerned with computer simulation experiments today.

In a later article, Naylor and Finger suggest a three-stage approach for validation:

1. Construct a set of hypotheses and postulates for the process using all available information - observations, general knowledge, relevant theory and intuition.
2. Attempt to verify the assumptions of the model by subjecting them to empirical testing.
3. Compare the input-output transformations generated by the real world.[11]

Van Horn approaches verification of the model from the standpoint of cost/value trade-offs.[12] He sees validation as a process of selecting a set of verification methods which is appropriate for a particular problem or

model. The selection criterion he proposes is the balancing of the cost of each action against the value of increased information about the validity of the results. He suggests eight possible methods:

1. Find models with high face validity.
2. Make use of existing research, experience, observation and other available knowledge to supplement models.
3. Conduct simple empirical tests of means, variances and distributions using available data.
4. Run tests of a Turing type. A Turing test, in this context, indicates the ability of an operationally experienced individual to discriminate between actual data and simulated data.
5. Apply complex statistical tests on available data.
6. Engage in special data collection.
7. Run prototype and field tests.
8. Implement the results with little or no validation.[13]

Thus, many suggestions have been made concerning the validation problem. To date, however, there seems to be very little concensus in the literature concerning a "proper" validation method.

## Experimental Design

Another important consideration in any discussion of the problems involved in computer simulation is that of the design of experiments. Several authors have expressed concern that too little is being done to solve design difficulties inherent in simulation experiments. Mize and Cox state:

> Literature dealing with the statistical design of simulation experiments is relatively scarce.

> ...The development of practical procedures is urgently needed in this area.[14]

They follow these statements with a discussion of the ways in which simulation experimentation differs from physical experimentation. These differences include the assumption of independence, the definition of experimental error, and the concept of randomization.

In physical experiments, great care is taken to insure that results from one combination of factor levels be independent of other results. This is not ordinarily the case in simulation experiments since analysts are able to control the amount of variability. An identical sequence of events can be reproduced by reproducing the same sequence of pseudo-random numbers for each alternative simulated. Conway, et al., sees this procedure as:

> the limiting case of the blocking concept commonly employed in experimental designs - blocks as homogenous as possible are selected to reduce the variability of the results.[15]

This reduction in variation not only sharpens the contrast between alternatives, but also allows a reduction in sample size.[16] However, reproducing the same stream of random numbers does not yield statistically independent results and the usual analysis of variance procedures are not applicable.[17] If only two alternatives are being considered, the results can be paired and the differences between pairs become the relevant sample observations. However, if more than two alternatives are compared, there is no satisfactory method available to make the

comparisons. Hillier and Lieberman do not see this as an over-riding limitation in that they view the procedure as:

> preferable to using different random numbers for each alternative and thereby obtaining statistically independent samples with relatively large variances.[18]

Another difference between physical and simulation experiments involves experimental error. In physical experiments, the residual variation is variation which is unexplained or beyond the control of the experimenter. In a simulation experiment, the stochastic variation, like every other feature of the model, is deliberately placed there by the constructor. The variation is introduced by generating pseudo-random numbers. As mentioned above, sequences of random numbers are usually reproduced to reduce the variance.[19]

Finally, the concept of randomness is unusual in simulation experiments. In physical experiments, randomness refers to the order in which tests are to be conducted and is imposed to average out the effects of uncontrolled variables. There are no "uncontrolled" variables in simulation. Therefore, since the programmer imposes each variable, the order of testing is unimportant.[20]

In 1966, Burdick and Naylor presented a comprehensive survey of experimental design methods which were applicable to simulation experiments. They included analysis of variance techniques, multiple ranking procedures,

sequential sampling and spectral analysis.[21] Later, in a book by Naylor, et al., designs including full factorial, fractional factorial, rotable, and response surface were discussed.[22] An extension of that work is found in an article by Ignall.[23] Multiple ranking procedures are more fully discussed in an article by Kleijnen, Naylor, and Seaks.[24] Williams and Weeks introduced a design allowing a sequential study of a factorial experiment.[25] Few studies other than these attack the problem of experimental design for simulations.

## The Problem of Sample Size

Burdick and Naylor classify the sample size problem as one of the major problems remaining in simulation experiments. The problem may be broken down into two subclassifications: 1) when to begin measurement; and 2) how many observations to measure.[26] The former is commonly called the problem of steady state and occurs when a process which operates on a substantially continuous basis is being studied. The determination of how many observations to measure becomes difficult when the data lack independence and normality.

### Steady State

For a continuously-running process or system, the question of when to begin measurement of observations hinges on a determination of when the system has reached

equilibrium. Starting the process in simulation introduces some mis-leading observations at the very beginning of the run. For instances, the first entities entering a queuing system would experience a very small waiting time while, in equilibrium, the waiting time might be an average of thirty minutes or more. Obviously inclusion of these observations would bias the results. Therefore, measurement should begin only after some preliminary running time during which the transient conditions are allowed to decay.[27] Equilibrium, however, implies only that the long range mean be stationary. The term does not require that the sample be normally distributed nor that it be void of runs or cycles.[28]

The question that lingers, though, is when does steady state occur. Schmidt and Taylor define the steady state for a parameter as the condition that occurs when "the mean and variance of a particular parameter stabilize to essentially constant values."[29] Over-all steady state is reached only when the mean and variance of each parameter stabilize. Conway, however, cautions against using cumulative statistics to determine steady state on the basis that these statistics lag behind the current state.[30] Their use results in discarding information unnecessarily. In addition, these statistics will "settle down even for systems which do not have a stationary state probability distribution."[31] As yet, no satisfactory method for determining steady state has been found.

## Stochastic Convergence

The second phase of the sample size problem could be called the problem of stochastic convergence and refers to the convergence of sample averages to population averages for large sample sizes. A measure of precision is the standard error of the mean, $\sigma/\sqrt{n}$ (the ratio of the population standard deviation, $\sigma$, to the square root of the sample size, n). To insure that the estimate is twice as precise, the sample size must be increased by four times. Thus, stochastic convergence is relatively slow and an extremely large sample size may be necessary to produce a reasonable level of precision.

This problem is compounded by the realization that conventional sample size techniques may not be appropriate for most simulation experiments in that assumptions of normality and independence are required. Simulation data, however, many times fit distributions other than normal.[32] In addition simulation data are usually not independent for observations which coexist in a system during a particular unit of simulated time. In fact, Conway states,

> in every investigation with which I have been concerned the correlation has been found to be appreciable and has had to be considered in the assessment of precision.[33]

In the past many practitioners have simply ignored the problem of lack of independence and proceeded as though the observations were not correlated. Others have suggested that a sample of large size can be used without

significantly under-estimating the parameter. Theoreticians have suggested making many runs with different random number sequences.[34] With an appeal to the Central Limit Theorem, the distribution of the means of these independent runs could be viewed as normal and independent. This method also requires an unusually large number of observations for analysis of the data. In addition, restarting to get independent observations means that steady state conditions must be reached each time.[35] A similar method which circumvents the steady state problem is making one continuous run and sub-dividing the observations into sets with intervening observations, the intervening observations quaranteeing independence.[36]

In a 1964 article, Murray A. Geisler suggested a sample size method for inventory simulations which allows the analyst to specify precision and confidence.[37] Geisler recognized the autocorrelation problem and proposed the values of $p^{th}$-order lag correlations be found and used in formulas for computing the variance of the mean of an autocorrelated series, as given by Moran.[38] From this the sample size could be found. Giesler, however, did not give an analytical method for calculating the lag correlations. Instead, a Monte Carlo approach was taken to estimate these values.

In 1967, another method was proposed by George S. Fishman.[39] Fishman showed that the variance of the sample mean for autocorrelated data is inversely proportional to

a fraction of the number of observations. Using this relationship, he defined the number of equivalent independent observations contained in an autocorrelated series.

## Fishman's Method for Computing Sample Size

In 1971, Fishman provided the theoretical basis for writing a subroutine for computing sample size for autocorrelated data which could be used in simulation programs.[40] Fishman shows that this procedure overcomes the assumption of independence and that it can be used in simulation experiments for which autocorrelation is a problem. This method also guarantees that the confidence interval on the estimated parameter will fall within the specified or desired width of the interval 100 percent of the time.

Fishman's method for computing sample size depends on a transformation of the correlated observations to an autoregressive form. A linear combination of deviations of the observations from their mean is found which produces new observations which are independent random variables. The sample variance, adjusted for autocorrelation, can then be estimated as a function of the estimates of the sample residual variations and the estimated coefficients of the autoregressive scheme. Fishman's development of the method can be explained as follows:

If an experiment should result in the observation of a sequence of events defined as $X_t$ where t is an integer

and $-\infty < t < \infty$, the mean of that sequence can be expressed as

$$\mu = E(X_t). \tag{2.1}$$

The autocovariance function of the sequence can be expressed as:

$$R_{t-s} = R_{s-t} = E\left[(X_s - \mu_s)\ (X_t - \mu_t)\right], \tag{2.2}$$

and the sample mean for N observations as

$$\overline{X} = (1/N)\ \Sigma_{t=1}^{N}\ X_t. \tag{2.3}$$

The variance can then be written as

$$\begin{aligned} \mathrm{Var}(\overline{X}_N) &= (1/N^2)\ \Sigma_{s,t=1}^{N}\ R_{s-t} \\ &= (1/N)\ \ \Sigma_{s=1-N}^{N-1}\ (1 - |s|/N)\ R_s. \end{aligned} \tag{2.4}$$

This approach assumes that the covariance between events in the sequence vanishes as the number of intervening events increases; that is,

$$\lim_{s\to\infty} R_s = 0. \tag{2.5}$$

With these assumptions, M can be defined as the sum of an infinite stream of autocovariances,

$$M = \lim_{N\to\infty}\ \Sigma_{s=1-N}^{N-1}\ (1 - |s|\ /N)\ \ R_s\ ; \tag{2.6}$$

this is equivalent to

$$M = \Sigma_{s=-\infty}^{\infty}\ R_s.$$

Thus, for large values of N the variance of the mean can be expressed as

$$\mathrm{Var}\ (\overline{X}_N) \simeq V_N = M/N. \tag{2.7}$$

The inconvenience of viewing M as a sum of an infinite sequence can be alleviated by expressing the sequence of events $X_t$ as an autoregressive sequence:

$$X_t = \mu + \Sigma_{s=0}^{\infty} a_s Y_{t-s} , \qquad (2.8)$$

where $a_s$ is a sequence of real constants such that

$$\Sigma_{s=0}^{\infty} |a_s| < \infty$$

and $Y_t$ is a sequence of uncorrelated, identically distributed random variables with mean zero and variance $\sigma^2$. Using this definition of $X_t$, the autocovariance of the $Y_t$ sequence can be written as

$$R_s = \sigma^2 \Sigma_{t=0}^{\infty} a_t \; a_{s+t} . \qquad (2.9)$$

The sum of the infinite stream of autocovariances, M, can then be written as

$$M = \Sigma_{s=-\infty}^{\infty} R_s$$

$$= \sigma^2 \; (\Sigma_{s=0}^{\infty} \; a_s). \qquad (2.10)$$

This result can be used to show that X has an autoregressive representation:

$$\Sigma_{s=0}^{\infty} \; b_s \; X'_{t-s} \; = Y_t , \qquad (2.11)$$

where

$$X'_t \; = X_t - \mu,$$

with

$$\Sigma_{s=0}^{\infty} \; a_s \; = 1/(\Sigma_{s=0}^{\infty} \; b_s),$$

$$M = \sigma^2/(\Sigma_{s=0}^{\infty} \; b_s)^2. \qquad (2.12)$$

Fishman also shows that for a finite autoregressive representation of order p, only p parameters need to be

estimated and

$$M = \sigma^2/(\Sigma_{s=0}^{p} b_s)^2 \text{ , where } b_o = 1. \qquad (2.13)$$

The estimation process which Fishman suggests involves the determination of the autoregressive order which is applicable for the given set of data and the subsequent application of equation (2.13). The general estimation process proceeds as follows:

The sample autocovariances are computed for several possible autoregressive orders, r, which range from values of zero to R':

$$C_{N,r} = (1/N)\Sigma_{t=1}^{N-r} X'_t X'_{t-r} \qquad (2.14)$$

where

$$X'_t = X_t - \overline{X}.$$

Then for a scheme of order, r+1, the $s^{th}$ coefficient is $\hat{b}_{r+1,s}$ and is determined by a recursive formula. If $\hat{b}_{r+1,0} = 0$ for all values of r, then

$$\hat{b}_{r+1,r+1} = \frac{-\Sigma_{s=0}^{r}\left[\hat{b}_{r,s} \quad C_{N,r-s+1}\right]}{\Sigma_{s=0}^{r}\left[\hat{b}_{r,s} \; C_{N,s}\right]} \qquad (2.15)$$

and

$$\hat{b}_{r+1,s} = \hat{b}_{r,s} + \hat{b}_{r+1,r+1}\hat{b}_{r,r-s+1}, \qquad (2.16)$$

where

$$s = 1, \ldots, r.$$

The sample residual variances can be computed as

$$\sigma^2_{r+1} = (n-r-1)^{-1} \Sigma_{t=r+2}^{N} (\Sigma_{s=0}^{r+1} b_{r+1,s} X'_{t-s})^2, \qquad (2.17)$$

where

$$r = 0,\ldots,R'-1.$$

Fishman shows that a confidence interval can be placed on the estimated coefficients with the interval

$$b_{r,r} \pm Z_{1-\alpha/2}[(1-\hat{b}^2_{r,r})/N]^{\frac{1}{2}} \qquad (2.18)$$

where $Z_{1-\alpha/2}$ is the point of the normal curve corresponding to a significance level $\alpha$ in

$$(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{Z_{1-\alpha/2}} e^{-x^2/2}\, dx = 1 - \alpha/2. \qquad (2.19)$$

The coefficient is not significantly different from zero if the confidence interval in equation (2.18) contains zero. The order of the autoregressive scheme, designated as "p", is the largest "r" for which $b_r$ is significant. Thus, the variance, M, of the sequence of original random numbers can then be estimated as

$$\hat{M} = \sigma_p^2/(\Sigma_{s=0}^{p} b_{p,s})^2. \qquad (2.20)$$

A confidence interval can be constructed on the mean by use of the probability statement

$$\Pr\left[|\bar{X}_N - \mu| < Z_{1-\phi/2}(\hat{M}/N)^{\frac{1}{2}}\right] \simeq 1-\phi, \qquad (2.21)$$

where $Z_{1-\phi/2}$ is the normal point corresponding to

$$(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{Z_{1-\phi/2}} e^{-x^2/2}\, dx = 1-\phi/2.$$

If a confidence interval of width 2C be preferred, then

$$C = Z_{1-\phi/2}\ (\hat{M}/N)^{\frac{1}{2}}. \qquad (2.22)$$

The necessary sample size for a confidence interval of 2C can be then computed by

$$N^* = \hat{M}(Z_{1-\alpha/2}/C)^2. \tag{2.23}$$

Fishman also shows that the use of an estimate for M in equation (2.23) requires that a t distribution be used in the construction of the confidence interval with L - 1 degrees of freedom where

$$L = N\,(C_{N,0} / \hat{M}). \tag{2.24}$$

Rather than read values from a table, the value of t is estimated by the following series of equations:

$$t_{1-\phi/2} \simeq Z_{1-\phi/2} + \frac{g_1}{(L-1)} \quad \frac{g_2}{(L-1)^2} \quad \frac{g_3}{(L-1)^3} + \ldots \tag{2.25}$$

where

$$g_1 = 1/4\,(Z_{1-\phi/2}{}^3 + Z_{1-\phi/2})$$

$$g_2 = 1/96\,(5Z_{1-\phi/2}{}^5 + 16Z_{1-\phi/2}{}^3 + 3_{1-\phi/2})$$

$$g_3 = 1/384\,(3Z_{1-\phi/2}{}^7 = 19Z_{1-\phi/2}{}^5 + 17_{1-\phi/2}{}^3 - 15_{1-\phi/2})$$

$$g_4 = 1/92160\,(79Z_{1-\phi/2}{}^9 - 776Z_{1-\phi/2}{}^7 + 14827_{1-\phi/2}{}^5 - 1920Z_{1-\phi/2}{}^3 - 945Z_{1-\phi/2}).^{41}$$

The estimate for sample size is, then,

$$N^* = \hat{M}(t_{1-\phi/2}/C)^2.$$

If the assumption that the sample means are normal is viewed to be inappropriate, Fishman shows how the process may be modified to accomodate an unimodal assumption. The probability statement on the mean can be written as:

$$PR\left[\frac{|\bar{X}_N - \mu|}{(\hat{M}/N)^{\frac{1}{2}}} \leq k\right] > (5k^2 - 3) / [3(1+k^2)], \text{ where}$$

$$0 \leq k \leq (5/3)^{\frac{1}{2}}$$

or

$$> (1-9k^2) / [9(1+k^2)], \text{ where}$$

$$k \geq (5/3)^{\frac{1}{2}}. \qquad (2.27)$$

If a confidence interval of 90% is desired, for instance, k can be found by:

$$(1 + 9k^2) / [9(1 + k^2)] = .90 \qquad (2.28)$$

or

$$k^2 = 7.89. \qquad (2.29)$$

An estimate for the sample size, N*, can then be computed by:

$$N^* = k^2\hat{M}/C^2, \qquad (2.30)$$

where 2C is the desired interval width.

Graybill's Method For Determining Sample Size

It is useful to study a conventional sample size method which proves to be applicable to simulation studies. The basic theory for sample size determination for distributions other than normal was introduced by F.A. Graybill for physical experiments with independent observations.[42] In addition, a recent article by Narula and Li discusses a sample size method for exponential life testing data which allows the experimenter to specify the level of significance and the probability of Type II error.[43] Although other methods may be superior under certain

conditions for normal data, Graybill's method is the only one purporting to handle many other known distributions.[44]

As in other methods, the analyst specifies the confidence coefficient $1-\alpha/2$, and the desired confidence interval width, d. In addition, he specifies the probability, $\beta^2$, that the actual width of the resulting confidence interval is less than or equal to the desired width, d.

Suppose the analyst wishes to determine the sample size, n, necessary to form a confidence interval of desired width, d, on some unknown parameter, $\mu$. He must first determine the actual width of the confidence interval, w, as a function of the sample size and some unknown parameter, $\theta$, which may be equal to $\mu$. Suppose there exists a function of the confidence interval width, sample size and $\theta$, $Y = g(w;\theta,n)$, such that Y is monotonically increasing in w for every $\theta$ and n and the distribution of Y depends only on the sample size. Then a function of the sample size, f(n), may be found so that

$$P[Y < f(n)] = \beta \text{ for } 0 < \beta < 1. \qquad (2.31)$$

If the equation, $f(n) = g(w;\theta,n)$, is solved for w, then the confidence interval width is expressed as a function of $\theta$ and the sample size, $h(\theta,N)$. The function, $h(\theta,n)$ be monotonically increasing for every n and monotonically decreasing in n for every $\theta$.

This method is a two-staged procedure in that a random variable, z, must be obtained in a preliminary sample of size, m. A function of z, f(z), must be defined so that

it does not depend on any unknown parameters or on the sample size, n, and

$$P\,[t(z) > \theta] = \beta \text{ for } 0 < \beta < 1. \tag{2.32}$$

Given that d is the desired confidence interval width, if a value of n is found such that

$$h\,[t(z),n] \leq d, \tag{2.33}$$

then the actual confidence interval width will be less than the desired width with a probability of $\beta$ or, stated mathematically,

$$P(w \leq d) \leq \beta^2. \tag{2.34}$$

Graybill used these conditions to develop a sample size formula to estimate the mean of a normal population. An initial sample, $u_1, u_2, \ldots u_m$, is taken of size m. The sample variance, z, is found and N* is the smallest integral value of the following:

$$\frac{2 \cdot t_{\alpha/2} \cdot \sqrt{z} \cdot \sqrt{f(N^*-1)}}{\sqrt{f(m-1)} \cdot \sqrt{N^*(N^* - 1)}} \leq 2C, \tag{2.35}$$

where:

1) $t_{\alpha/2}$ is a variate of Student's distribution with N*-1 degress of freedom or $\int_{t_{\alpha/2}}^{\infty} U(t,N^*-1)dt = \alpha/2$;

2) f(N*-1) is a chi-square variate with N*-1 degrees of freedom or $\int_0^{f(N^*-1)} W(\chi^2;\ N^*-1)\ d\chi^2 = \beta$ ;

3) f(m-1) is a chi-square variate with m-1 degrees of freedom or $\int_{f(m-1)}^{\infty} W(\chi^2;\ m-1)d\chi^2 = \beta$.

The estimate of the normal population is computed, then, by

taking a new sample of size N*, ($v_1$, $v_2$,...$v_{N*}$), and forming a confidence interval about the sample mean.

## Summary

This chapter discusses the literature available concerning three basic problems in simulation experiments: 1) validation; 2) experimental design; and 3) sample size. Each category of problems is broken into subproblems and examined. The problem of sample size is a ripe area for research given the dearth of practical solution techniques for this problem. Methods currently in use, requiring large blocks of computer core and time, tend to limit the use of simulation as a reliable tool of analysis. For these reasons, the section concerning sample size is more detailed and includes complete discussions of two methods, Fishman's and Graybill's, which are viewed as particularly relevant.

## FOOTNOTES

[1] Donald S. Burdick and Thomas H. Naylor, "Design of Computer Simulation Experiments For Industrial Systems," Communications of the ACM, IX (1966), p. 329.

[2] Russell L. Ackoff, Progress in Operations Research (New York, 1961), p. 30.

[3] R.W. Conway, B.M. Johnson and W.L. Maxwell, "Some Problems in Digital Systems Simulation," Management Science, VI (1959), p. 93.

[4] Burdick and Naylor, p. 329.

[5] Richard L. Van Horn, "Validation of Simulation Results," Management Science, XVII (1971), p. 248.

[6] Robert C. Meier, William T. Newell and Harold L. Pazer, Simulation In Business and Economics (New Jersey, 1969), p. 294.

[7] J.W. Schmidt and R.E. Taylor, Simulation and Analysis of Industrial Systems (Illinois, 1970), p. 502.

[8] K.J. Cohen and Richard M. Cyert, "Computer Models in Dynamic Economies," The Quarterly Journal of Economics, LXXV (1961), p. 112.

[9] Conway, Johnson and Maxwell, p. 104.

[10] Burdick and Naylor, p. 330.

[11] T.H. Naylor and J.M. Finger, "Verification of Computer Simulation Models," Management Science, XIV (1967), pp. B-95 - B-96.

[12]Van Horn, p. 257.

[13]Ibid.

[14]Joe H. Mize and J. Grady Cox, Essentials of Simulation (New Jersey, 1968), p. 152.

[15]Conway, Johnson and Maxwell, p. 105.

[16]Frederick S. Hillier and Gerald J. Lieberman, Introduction to Operations Research (San Francisco, 1967), p. 452.

[17]Conway, Johnson and Maxwell, p. 106.

[18]Hillier and Lieberman, p. 453.

[19]Mize and Cox, p. 154.

[20]Ibid.

[21]Burdick and Naylor.

[22]Thomas H. Naylor, Joseph L. Balintfy, et al., Computer Simulation Techniques (New York, 1966).

[23]Edward J. Ignall, "On Experimental Design For Computer Simulation Experiments," Management Science XVIII (1972), p. 384.

[24]Jack P.C. Kleijnen, Thomas H. Naylor and Terry G. Seaks, "The Use of Multiple Ranking Procedures to Analyze Simulations of Management Systems: A Tutorial," Management Science, XVIII (1972), p. B-245.

[25]Donald R. Williams and David L. Weeks, "A Technique For Designing and Augmenting Simulation Experiments," Management Science, XX (1974), p. 1385.

[26]Burdick and Naylor, p. 335.

[27]Conway, Johnson and Maxwell, p. 107.

[28]Ibid.

[29]Schmidt and Taylor, p. 346.

[30]R.W. Conway, "Some Tactical Problems In Digital Simulation," Management Science, X (1963), p. 49.

[31]Ibid.

[32]Phillip M. Morse, "Stochastic Properties of Waiting Lines," Operations Research, XIII (1955).

[33]Conway, p. 57.

[34]Meier, Newell and Pazer, p. 304.

[35]Conway, Johnson and Maxwell, p. 109.

[36]Ibid.

[37]Murray A. Geisler, "The Sizes of Simulation Samples Required to Compute Certain Inventory Characteristics With Stated Precision and confidence," Management Science, X (1964), p. 261.

[38]P.A.P. Moran, The Theory of Storage (New York, 1959).

[39]George S. Fishman, "Problems in the Statistical Analysis of Simulation Experiments: The Comparison of Means and the Length of Sample Records," Communications of the ACM, X (1957), p. 94.

[40]George S. Fishman, "Estimating Sample Size in Computing Simulation Experiments," Management Science, XVIII (1971), p. 21.

[41]Milton Abramowitz and Irene A. Stegun, (ed.), Handbook of Mathematical Functions (Washington, D.C., 1964), p. 949.

[42]Franklin A. Graybill, "Determining Sample Size For a Specified Width Confidence Interval," Annals of Mathematical Statistics, XXIX (1958), p. 282.

[43]Subhash C. Narula and Franz S. Li, "Sample Size Calculations in Exponential Life Testing," Technometrics, XVII (1975), p. 229.

[44]Aaron S. Goldman and R.K. Zeigler, "Comparisons of Some Two Stage Sampling Methods," Annals of Mathematical Statistics, LXVI (1966), p. 891.

# CHAPTER III

# A SAMPLE SIZE METHOD FOR EXPONENTIAL DATA

## Introduction

As explained in Chapter II, the issues involved in the sample size problem have largely been skirted in the literature. George S. Fishman has developed the only algorithmic approach to handling the complications of non-normality and autocorrelation of the data. The purpose of this chapter is to develop a new approach to the problem by revising a sample size method introduced by Franklin A. Graybill. Graybill's approach, as outlined in Chapter II, is designed for distributions of data other than normal in which the data points are independent. The method proposed here is for exponential or geometric distributions with either independent or autocorrelated data.

Also included in this chapter are the results of tests comparing the proposed method with Fishman's method. The tests included comparisons of the relative abilities of the two models to produce samples with a specified confidence interval width and with a stated confidence coefficient. The average sample sizes which are produced are compared

for the two methods using the same data - generating models and the same data. In addition, the computer times necessary to generate samples of the required sizes are compared. Finally, tests were performed to validate the confidence coefficients used in Fishman's method.

## The Proposed Method Revised For Exponential Distributions

### Use of The Exponential Distribution In Simulation

The data generated by many simulation experiments fit expoential distributions. In queuing experiments, for instance, one may be interested in the distribution and mean for the total amount of time spent in the system by the entities.[1] Other simulations may generate time-to-failure data for reliability estimates. For these and other simulations generating exponentially distributed data, the knowledge that the data fit that distribution can be used to facilitate sample size calculations.

### Development of the Method

The theory for developing a sample size method for distributions other than normal was developed by Franklin A. Graybill and is discussed in Chapter II. In this section a method for estimating sample size for exponential populations will be developed using the conditions prescri-

bed by Graybill and summarized in the previous chapter.

Graybill's conditions can be utilized to formulate a model for exponential data if a basic sampling property of the exponential distribution is recognized. A relationship exists whereby the ratio of twice the sum of sample observations ($v_1, v_2, \ldots, v_n$) and the true mean of the population exhibits a chi-square distribution with 2n degrees of freedom:

$$\frac{2n\hat{\theta}}{\theta} \sim \chi^2 (2n), \text{ where } \hat{\theta} = \frac{1}{n} \cdot \sum_{i=1}^{n} v_i^2 \qquad (3.1)$$

Let the parameter, $\theta$, be: 1) the true mean of the population for which a confidence interval is formed; and 2) the unknown parameter as specified in the second chapter.

Given equation (3.1), the following statement may be made:

$$P\left[\chi^2_2 < \frac{2n\hat{\theta}}{\theta} < \chi^2_1\right] = 1 - \alpha, \qquad (3.2)$$

where $\chi^2_1$ and $\chi^2_2$ are defined as

$$\int_0^{\chi^2_2} W(\chi^2;2n)\, d\chi^2 = \frac{\alpha}{2}$$

and

$$\int_{\chi^2_1}^{\infty} W(\chi^2;2n)\, d\chi^2 = \frac{\alpha}{2} \quad .$$

Then a confidence interval on the man can be derived as follows:

$$P\left[\frac{\chi^2_2}{2n\hat{\theta}} < \frac{1}{\theta} < \frac{\chi^2_1}{2n\hat{\theta}}\right] = 1 - \alpha$$

$$P\left[\frac{2n\hat{\theta}}{\chi^2_1} < \theta < \frac{2n\hat{\theta}}{\chi^2_2}\right] = 1 - \alpha. \qquad (3.3)$$

Therefore, the width of the confidence interval w, is:

$$w = 2n\hat{\theta}\left[\frac{1}{\chi^2_2} - \frac{1}{\chi^2_1}\right].$$

If the variable, $D_n$, is defined as

$$D_n = \frac{1}{\chi^2_2} - \frac{1}{\chi^2_1} \qquad (3.4)$$

then

$$w = 2n\hat{\theta}D_n; \qquad (3.5)$$

Recalling equation (3.1), let

$$Y = \frac{2n\hat{\theta}}{\theta} \qquad (3.6)$$

be a chi-square distribution with 2n degrees of freedom and, thus, dependent only on n. Solving equation (3.5) for $2n\theta$ and substituting into equation (3.6) results in

$$Y = g(w;\theta,n) = \frac{w}{\theta D_n}. \qquad (3.7)$$

If f(n) is defined as

$$\int_0^{f(n)} W(\chi^2;\ 2n)\ d\chi^2 = \beta, \qquad (3.8)$$

then condition (2.31) is satisfied. If $f(n) = g(w;\theta,n)$ then

$$f(n) = \frac{w}{\theta D_n}$$

and

$$w = f(n) \cdot \theta \cdot D_n,$$

or, equivalently,

$$h(\theta,n) = f(n) \cdot \theta \cdot D_n \ . \qquad (3.9)$$

In the first stage of the sampling procedure, m items $(u_1, u_2,\ldots,u_m)$ are summed so that

$$z = 2 \cdot \sum_{i=1}^{m} u_i \ , \qquad (3.10)$$

or

$$z = 2m\hat{\theta}, \qquad (3.11)$$

From the relationship in equation (3.1),

$$\frac{2m\hat{\theta}}{\theta} \sim \chi^2 \ (2m), \qquad (3.12)$$

a statement may be made that

$$P\left[\frac{z}{\theta} > \chi^2_3\right] = \beta, \qquad (3.13)$$

where $\chi^2_3$ is defined as

$$\int_{\chi^2_3}^{\infty} W(\chi^2;\ 2m)\ d\chi^2 = \beta. \qquad (3.14)$$

Therefore, condition (2.32) is satisfied by

$$P\left[\frac{z}{\chi^2_3} > \theta\right] = \beta, \qquad (3.15)$$

where

$$t(z) = \frac{z}{\chi^2_3} \ . \qquad (3.16)$$

Substituting equation (3.16) into equation (3.9) gives

$$h\left[t(z),\ n\right] = \frac{f(n) \cdot z \cdot D_n}{\chi^2_3}. \quad (3.17)$$

As stated above, any value of n satisfying the requirement,

$$h\left[t(z),n\right] < d,$$

insures that

$$P(w \leq d) \geq \beta^2.$$

Therefore, the necessary sample size, N*, is the smallest integral value of n satisfying

$$\frac{f(n) \cdot z \cdot D_n}{\chi^2_3} \leq d. \quad (3.18)$$

The procedure, then, for determining the necessary sample size for a specified width confidence interval and constructing that confidence interval is: 1) select a random sample of m observations; 2) based on the mean of the m observations, find the smallest integral value of n satisfying equation (3.18); 3) select a second random sample of size N*; and 4) compute a confidence interval using equation (3.3), and the sample mean from the N* number of observations.

The development given above provides a sample size method for exponential data. However, the assumption of independence still remains. The hypothesis was made that autocorrelation of data would not be a complicating factor

for this particular method. The presence of autocorrelation affects the estimate of the variance and, therefore, changes the relationship between the mean and variance. However, since the necessary sample size for this method depends only on the mean of the data and not on the variance, the only requirement is that the estimate of the mean be correct. Therefore, the proposed method is sufficient for exponentially distributed data which is either independent or autocorrelated.

A FORTRAN subroutine for this algorithm has been developed for use in FORTRAN or GASP simulations. A detailed explanation, flowchart, and program listing may be found in Appendix A.

## Comparisons of Sample Size Methods For Exponential Data

In order to validate the new method a number of tests were devised. This section includes a discussion of the measures of effectiveness which were considered important and the results of tests to estimate these measures.

Each test was also applied to Fishman's method to provide a comparison between the new method and Fishman's method for computing sample size. After some initial testing, it became apparent that when using Fishman's method for exponential data it is necessary to use the unimodal assumption. This procedure is consistent with results

reported by Fishman for the same kind of data.[3]

Measures of Effectiveness

The primary concern in testing the two methods was to determine if the method actually attained a confidence interval with the required level of confidence. Therefore, a sample size, N*, was computed for a given method and that number of observations was collected. Based on these observations, confidence limits were computed. Next, the true value of the mean of the distribution was compared to the confidence limits in order to determine if it were contained in the confidence interval. Upon completion of many trials, statistics were computed for the proportion of times the confidence interval actually contained the true value of the mean, $1 - \hat{\alpha}$. This proportion was compared to the confidence coefficient, $1-\alpha$, for the interval. In order to be satisfactory, the sample size method was required to produce a statistic, $1-\hat{\alpha}$, that was at least as great as the desired confidence coefficient $1 - \alpha$, or:

$$1 - \hat{\alpha} \geq 1 - \alpha, \qquad (3.19)$$

The second criterion of importance was the proportion of times, $\beta^2$, the method produced a confidence interval width which was less than or equal to the desired width. Therefore,

$$\hat{\beta}^2 \geq \beta^2, \qquad (3.20)$$

where $\beta^2$ is the stated probability that the computed confidence interval width is less than or equal to the desired width. Stated mathematically,

$$P(w \leq d) \geq \beta^2, \qquad (3.21)$$

where d is the desired confidence interval width and w is the computed width. Accumulating the estimate, $\hat{\beta}^2$, entailed finding the difference between the confidence interval limits for the sample and comparing this difference to the desired width, d.

Comparisons were made of the estimated sample sizes produced by each method. The objective of this test was to determine which method was more efficient in terms of the number of data points needed to satisfy the restrictions on the quantities, $1 - \alpha$ and $\beta^2$.

In addition, comparisons were made in regard to the computer time required to compute the value for the estimated sample size and to collect that many observations. The criterion of time was seen as an important measure of effectiveness in the context of simulation experiments. As stated previously, many computer simulations are quite large and complex. A sample size method which is to be incorporated in a simulation, then, should be efficient in terms of time to avoid unnecessarily increasing the time required for the total simulation. Average compilation times and average execution times were computed for each method.

The methods were also compared on the basis of computer core required to complete the required computations. The amount of core required for most simulations is quite large. Therefore, a sample size routine which is incorporated in the simulation should use as little core as possible. This test should indicate which method is more efficient in terms of core requirements.

Finally, tests were made concerning the probability statements inherent in Fishman's method.

## Data-Generating Models

Autoregressive formulas were used to achieve several levels of autocorrelation. Therefore, an observation, $X_t$, might depend on several previous values of X, $X_{t-1}$, and $X_{t-2}$, and the corresponding independent variable, $Y_t$. The models used to generate data are given below.

1) Independent data: Autoregressive order = 0.0

$$X_t = .5 + Y_t, \qquad (3.22)$$

2) Autocorrelated data: Autoregressive order = 1.0

$$X_t = .5X_{t-1} + .5 + Y_t, \qquad (3.23)$$

3) Autocorrelated data: Autoregressive order = 2.0

$$X_t = .5X_{t-1} + .25X_{t-2} + .5 + Y_t, \qquad (3.24)$$

4) Autocorrelated data: Autoregressive order = 3.0

$$X_t = .5X_{t-1} + .25X_{t-2} + .05X_{t-3} + .5 + Y_t, \qquad (3.25)$$

5) Autocorrelated data: Autoregressive order = 5.0

$$X_t = .5X_{t-1} + .2X_{t-2} + .05X_{t-3} + .03X_{t-4} + .02X_{t-5} + .5 + Y_t. \quad (3.26)$$

6) Autocorrelated data: Autoregressive order = 10.0

$$X_t = .5X_{t-1} + .2X_{t-2} + .05X_{t-3} + .03X_{t-4} + .01X_{t-5} + .005X_{t-6} + .003X_{t-7} + .001X_{t-8} + .0006X_{t-9} + .004X_{t-10} + .5 + Y_t. \quad (3.27)$$

For each of the above models, different distributions can be achieved by varying the distribution of Y . If $Y_t$ is distributed normally with a mean of 0.0 and a variance of 1.0, then, using model (1), $X_t$ is distributed normally with a mean of 0.5:

$$E(X_t) = E(.5) + E(Y_t) \quad (3.28)$$

and

$$E(X_t) = .5 + 0.0 = .5. \quad (3.29)$$

Using model (3), X is distributed normally with a mean of 2.0. The mean is computed as follows:

$$E(X_t) = .5E(X_{t-1}) + .25\ E(X_{t-2}) + E(.5) + E(Y_5) \quad (3.30)$$

and

$$.25E(X_t) = .5 + E(Y_t) \quad (3.31)$$

and

$$E(X_t) = 2.0.$$

Extending this idea, exponential and geometric data can be generated for various autoregressive orders and the true means for the distributions can be computed.

## Test Procedures

Each test described in this chapter was performed using a Monte Carlo procedure. Data were generated for a specified distribution and autoregressive order. An initial sample of the data population was randomly selected to provide input for the two sample size methods. An estimated sample size was computed using the proposed method for exponential data and Fishman's method augmented with an assumption of unimodally distributed means. Then a sample of that size was taken from the population. Based upon this sample, confidence intervals were calculated. This procedure was repeated a number of times for each data-generating model and estimates were computed for:

1) Mean sample size for each method;
2) Variance of the sample size;
3) Probability that the true mean is contained in the confidence interval constructed about the sample mean;
4) Probability that the confidence interval width is less than or equal to the desired width.

Through the use of the data-generating models

described in the previous section, many different kinds of distributions may be generated. A series of programs were written to incorporate these models and to produce the data necessary to make comparisons of the two methods. Examples of these programs may be found in Appendixes B and C. Notice that random numbers were produced using a random number generator called LLRANDOM.[4] This generator has been found to be a better generator than other random number generators which are commonly used. In addition, the generator will produce exponentially distributed, as well as normally distributed, numbers. For these tests exponential numbers with $\mu = 2$ were generated and transformed to autoregressive data using the data-generating models.

For each model, two hundred (200) trials, or repetitions of the experiment, were taken. After analyzing initial data it was found that two hundred sample size estimates were sufficient to produce a confidence interval with a confidence coefficient of .9 and a confidence interval width which at most is fifteen percent of the mean sample size. Calculations for selected experiments are found in Appendix D.

## Test Results

Confidence Coefficient. One of the primary purposes of these tests was to verify that the proposed method for

exponential data did actually produce the desired confidence coefficient, $1 - \alpha$. The results of these tests may be found in Tables I - III. A confidence coefficient of .90 was used in each experiment. After two hundred trials of each model, at least 90 percent of the computed confidence intervals contained the true mean using the new method and Fishman's method. In fact, the proposed method appears to be quite conservative in some cases. The experiments using autoregressive orders one and two produced results showing $1 - \hat{\alpha}$ to be greater than .96 for every value of $\beta^2$ used.

Confidence Interval Width. The probability that the computed confidence interval width is within limits set by the analyst has been defined as $\beta^2$. For these tests six values for $\beta^2$ have been chosen for the new method. Fishman, of course, implied a $\beta^2$ of 1.0 for his method. In Tables I - III the desired level of $\beta^2$ is given at the top of the table. The simulated estimate of $\beta^2$, $\hat{\beta}^2$, is given at the bottom of the table. Again, both methods performed as expected by producing values well above the stated value for $\beta^2$. Using this criterion, the proposed method was again very conservative.

Size of Sample. Test results for estimated sample size are also given in Tables I - III. For independent data and for data with an autoregressive order of two, the proposed method was superior to Fishman's method. For independent data, Fishman's method required an average of

TABLE I

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR EXPONENTIAL DATA WITH AUTOREGRESSIVE ORDER OF ZERO

| Variable | Fishman's Method | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|---|
| | $\beta^2=1.00$ | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 631 | 318 | 333 | 369 | 430 | 485 | 613 |
| Standard Deviation of Distribution of Sample Sizes | 91.9 | 78.7 | 82.0 | 90.1 | 103.6 | 116.0 | 145.2 |
| Estimated Mean of Data ($\bar{X}$) | 9.01 | 9.02 | 9.03 | 9.01 | 9.00 | 8.99 | 8.99 |
| True Mean of Data ($\mu$) | 9.00 | 9.00 | 9.00 | 9.00 | 9.00 | 9.00 | 9.00 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\bar{X}$ | 1.00 | .94 | .94 | .93 | .92 | .90 | .90 |
| Proportion of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | 1.00 | .87 | .91 | .96 | .99 | 1.00 | 1.00 |

TABLE II

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR EXPONENTIAL DATA WITH AUTOREGRESSIVE ORDER OF ONE

| Variable | Fishman's Method | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|---|
| | $\beta^2=1.00$ | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 79 | 102 | 108 | 121 | 143 | 163 | 209 |
| Standard Deviation of Distribution of Sample Sizes | 57.2 | 19.7 | 30.6 | 22.8 | 26.6 | 30.3 | 37.9 |
| Estimated Mean of Data ($\overline{X}$) | 4.78 | 5.00 | 5.00 | 4.99 | 5.00 | 5.00 | 5.01 |
| True Mean of Data ($\mu$) | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\overline{X}$ | .87 | .97 | .96 | .96 | .96 | .97 | .98 |
| Proportion of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | 1.00 | .93 | .97 | .99 | 1.00 | 1.00 | 1.00 |

TABLE III

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR EXPONENTIAL DATA WITH AUTOREGRESSIVE ORDER OF TWO

| Variable | Fishman's Method | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|---|
| | $\beta^2=1.00$ | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 401 | 330 | 345 | 383 | 445 | 502 | 635 |
| Standard Deviation of Distribution of Sample Sizes | 175.2 | 67.6 | 70.3 | 77.2 | 88.8 | 99.3 | 124.2 |
| Estimated Mean of Data ($\bar{X}$) | 9.87 | 10.02 | 10.02 | 10.00 | 10.00 | 9.99 | 9.99 |
| True Mean of Data ($\mu$) | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 |
| Proportion of Trials ($1-\alpha$) In Which $\mu$ Is Contained In Confidence Interval About $\bar{X}$ | .94 | .99 | .98 | .98 | .97 | .97 | .99 |
| Proportion of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | 1.00 | .77 | .83 | .92 | .99 | 1.00 | 1.00 |

631 items. The proposed method indicated that only 613 were necessary for a $\beta^2$ of .99. In practice, $\beta^2$ would usually be set at .90 or less. For a more useful level of $\beta^2$, then, the proposed method indicated a sample size of at least 200 items fewer than Fishman's method. For a second order autoregressive equation, Fishman's method indicated a sample size of 401. This sample size would fall between the sample sizes necessary for $\beta^2 = .80$ and $\beta^2 = .90$ for the new method; however, the $1 - \hat{\alpha}$ value is less for Fishman's method than for the new method for this experiment. Both values are still acceptable. This is not the case for the first order autoregressive experiments. The $1 - \hat{\alpha}$ level of .87 is insufficient for Fishman's method. Therefore the smaller sample size must be viewed as suspect. Histograms of the sample sizes for selected experiments are shown in Figures 1-4. Figure 5 shows the distribution of average sample size estimates using the new method for various values of $\beta^2$ and autoregressive orders.

Time Requirements. The time requirements for each method are shown in Table IV. The proposed method was superior in terms of compilation time and execution time in every instance. For independent data, the new method was at least 64 times faster than Fishman's method for execution time. For autoregressive order two, the new method was 19 times faster than Fishman's. However, the new method was only 6.5 times faster than Fishman's

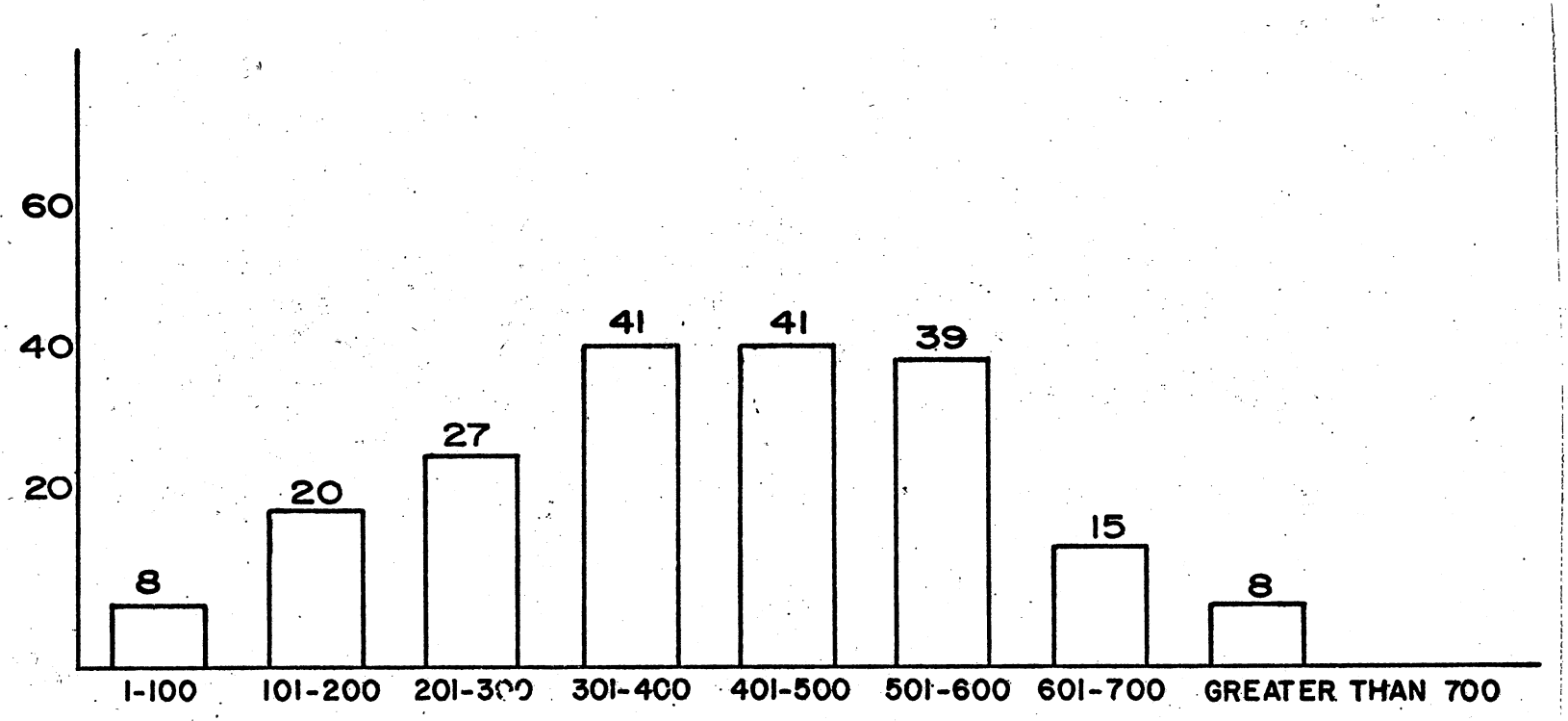


Figure 1. Histogram Of Sample Sizes Obtained With Fishman's Method And An Autoregressive Order of Two

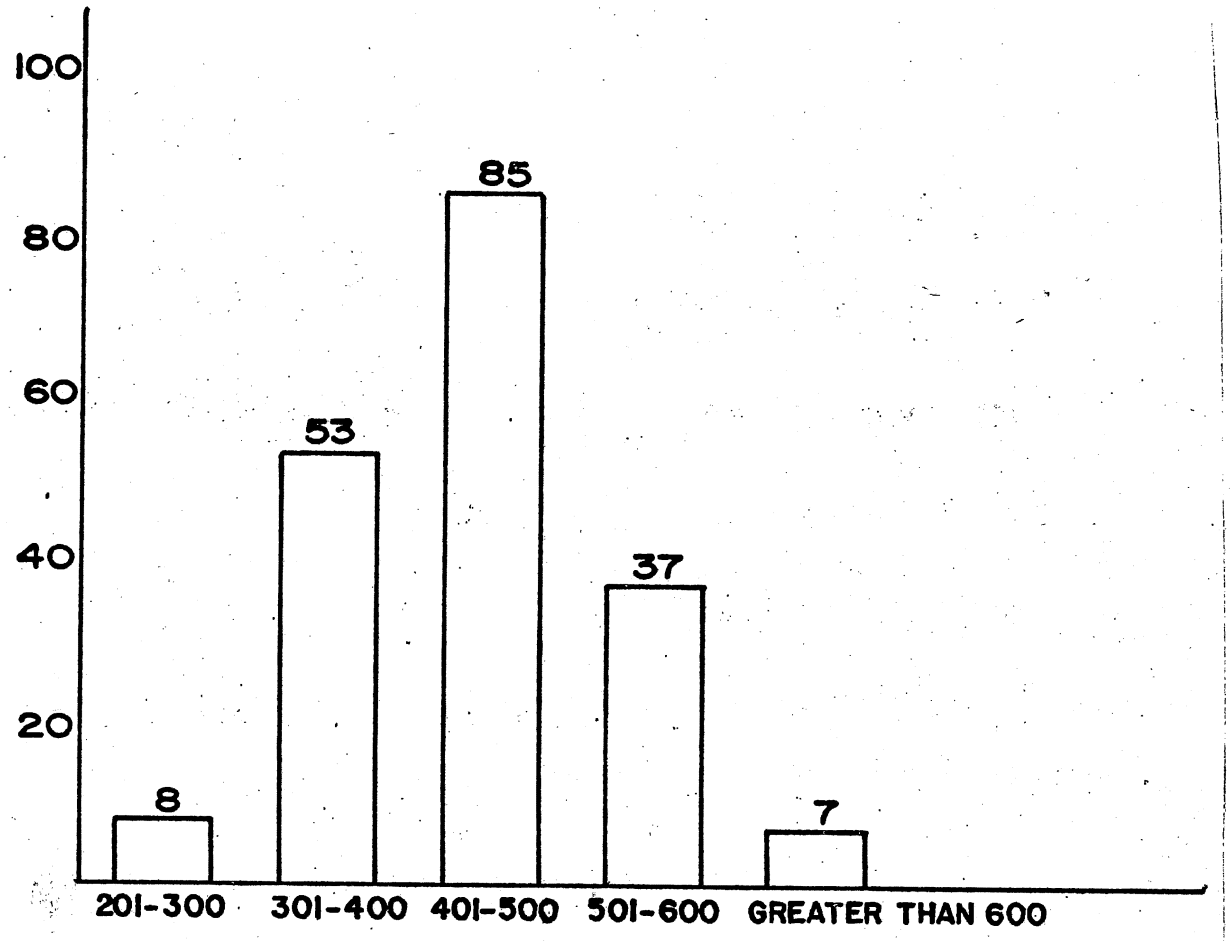


Figure 2. Histogram Of Sample Sizes Obtained With The New Method: $\beta^2$ of .90 And Autoregressive Order of Two

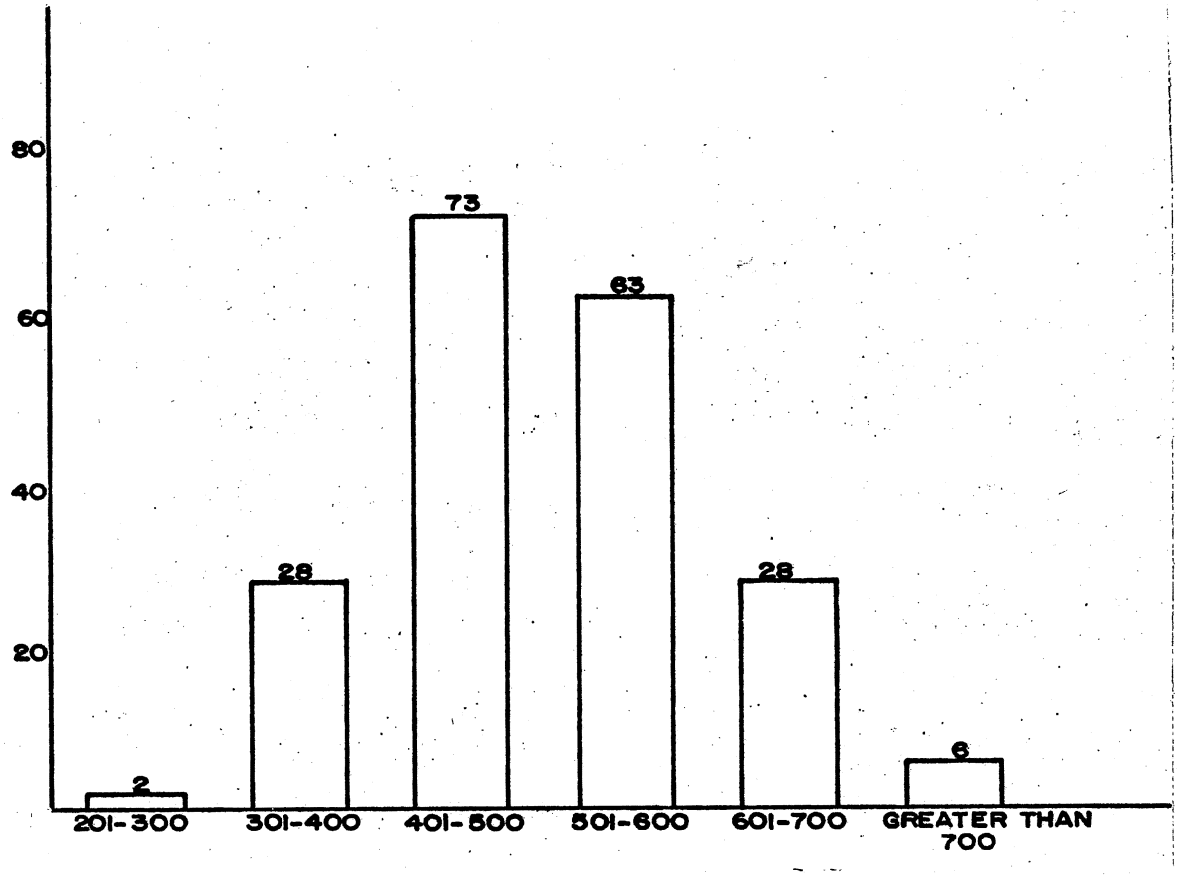


Figure 3. Histogram Of Sample Sizes Obtained With The New Method: $\beta^2$ Of .95 And Autoregressive Order Of Two

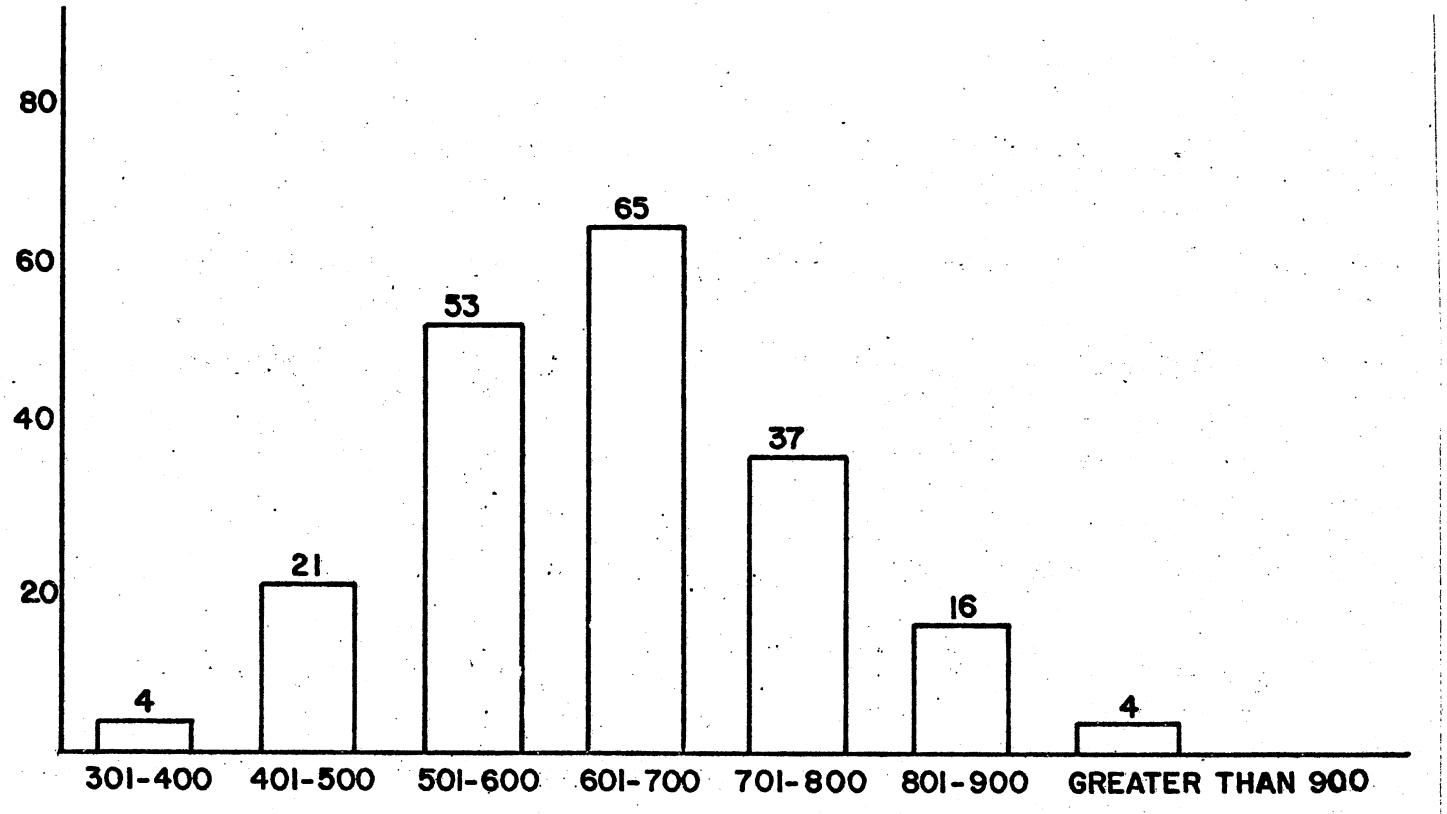


Figure 4. Histogram Of Sample Sizes Obtained With The New Method: $\beta^2$ of .99 And Autoregressive Order of Two

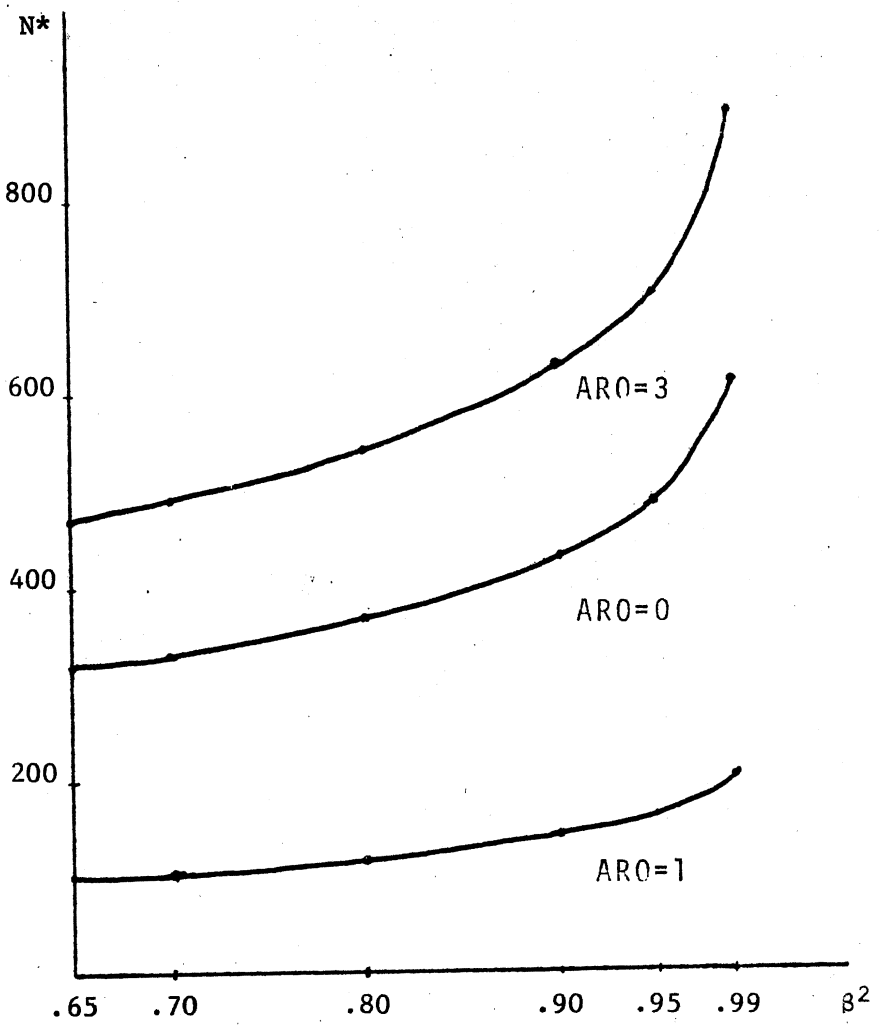


Figure 5. Average Sample Sizes For Exponential Data Using The New Method

method for autoregressive order one. This last comparison was a result of Fishman's smaller sample size for an autoregressive order of one.

Core Requirements. The proposed method is also superior based on the criterion of core requirements. Only 46 K is required for the new method while 64 K is necessary for Fishman's, as shown in Table IV.

Probability Statements. In analyzing the results of the previous tests, a doubt arose as to the interpretation of Fishman's implcation that $\beta^2$ is 1.00 using his method. Obviously, the confidence interval width must be less than the desired width in Fishman's method because it is forced to be less. Fishman ends his procedure only when the standard error of the mean is small enough to satisfy the confidence interval width requirement. The question, then, was whether this value could be interpreted in the same way as the $\beta^2$ value is interpreted in the proposed method. The interpretation of the value in the new method is that $\beta^2$ is the proportion of times that the confidence interval width will be less than a desired width if that size sample is collected a large number of times. This same definition was employed using data generated by Fishman's method. A repeated number of samples were collected of the size indicated by the average sample size for each autoregressive order. The results of this experiment are shown in Table V.

TABLE IV

COMPARISON OF TIME REQUIREMENTS AND CORE REQUIREMENTS FOR SAMPLE SIZE EXPERIMENTS USING EXPONENTIAL DATA

| Variable | Fishman's Method | The Proposed Method | | |
|---|---|---|---|---|
| | | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| ARO = 0 | | | | |
| Compilation Time (seconds) | 6.09 | 4.22 | 3.99 | 3.80 |
| Average Execution Time (seconds) | 8.43 | .13 | .13 | .13 |
| ARO = 1 | | | | |
| Compilation Time (seconds) | 6.45 | 4.37 | 4.40 | 4.42 |
| Average Execution Time (seconds) | .98 | .15 | .15 | .15 |
| ARO = 2 | | | | |
| Compilation Time (seconds) | 6.42 | 4.59 | 4.80 | 5.47 |
| Average Execution Time (seconds) | 3.70 | .19 | .19 | .21 |
| Core Requirements | 64K | 46K | 46K | 46K |

TABLE V

EVALUATION OF PROBABILITY STATEMENTS

| | ARO = 0 | ARO = 1 | ARO = 2 |
|---|---|---|---|
| Fishman's Method: | | | |
| Average Estimated Sample Size Using Fishman's Method | 631 | 79 | 401 |
| Estimate of $\beta_1^2$, $\hat{\beta}_1^2$, By Generating Repeated Samples Of Size N* | .45 | .29 | .35 |
| The Proposed Method: | | | |
| Estimate of Sample Size Using The $\hat{\beta}_1^2$ Value From Fishman's Method as $\beta_2^2$ | 268 | 71 | 257 |
| Resulting $\hat{\beta}_2^2$ | .71 | .55 | .39 |

The computation of estimated sample sizes is explained in a previous section and the results are shown in Tables I through III. For independent data, Fishman's method indicated that a sample of 631 items should be taken for $1 - \alpha = .9$ and with $\beta^2 = 1$ implied. Two hundred samples of 631 observations were collected and confidence intervals were computed for each. In each trial the confidence interval width was compared to the desired width. The computed width was less than or equal to the desired width in only 45% of the trials. The next step was to test the new method with a $\beta^2$ value of .45. As shown in Table V, the new method produced an average estimated sample size of 268 when $\beta^2 = .45$. Of the 200 samples taken, 71% had confidence intervals with widths less than or equal to the desired width. The experiment was repeated for autoregressive orders one and two with similar results.

This experiment would indicate that a probability statement concerning the confidence interval width cannot be made for Fishman's method. The data are not actually random data in that the sample is taken in sequential iterations and the sampling process ends when the specific data points selected have a standard error of the mean small enough to satisfy the confidence interval width criterion. Thus, a $\beta^2$ of one is forced and cannot be said to have taken place as a result of random sampling.

## Comparisons of Sample Size Methods for Geometric Data

Just as many simulations produce exponential data, they also produce geometrically distributed data. For instance, in queuing experiments the number of entities in the system takes on a geometric distribution.[5] Since the exponential distribution is the continuous analog of the geometric, it was felt that the proposed method for exponential data would also be satisfactory for geometric data.[6]

Geometric data were generated by calling a LLRANDOM subroutine for normal data.[7] The normal data were then transformed to geometric data using a method which is reported by Schmidt and Taylor.[8] Different autoregressive orders are obtained in the same manner as for exponential data.

The tests used to verify the use of the proposed method on geometric data are the same as were reported for exponential data. The results of these tests are shown in Tables VI - VIII. Notice that in each case the estimate for the confidence coefficient, $1 - \hat{\alpha}$, is larger than the lowest acceptable value of .90. The computed confidence interval widths were also acceptable. The value of $\hat{\beta}^2$ was at least as large as the desired level, $\beta^2$, in every case. Tests of models with higher autoregressive orders were also run. The results of these tests are shown in

TABLE VI

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR GEOMETRIC DATA WITH AUTOREGRESSIVE ORDER OF ZERO

| Variable | Fishman's Method | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|---|
| | $\beta^2=1.00$ | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Average Required Sample Size | 138 | 105 | 111 | 125 | 147 | 168 | 215 |
| Estimated Mean ($\overline{X}$) of Data | 4.89 | 5.00 | 5.01 | 5.01 | 5.01 | 5.01 | 5.01 |
| True Mean ($\mu$) of Data | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\overline{X}$ | .95 | .95 | .93 | .93 | .93 | .93 | .92 |
| Proportion of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | 1.00 | .90 | .91 | .97 | 1.00 | 1.00 | 1.00 |

TABLE VII

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR GEOMETRIC DATA WITH AUTOREGRESSIVE ORDER OF ONE

| Variable | Fishman's Method | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|---|
| | $\beta^2=1.00$ | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Average Required Sample Size | 600 | 445 | 465 | 515 | 597 | 671 | 846 |
| Estimated Mean ($\overline{X}$) of Data | 10.95 | 11.02 | 11.02 | 11.04 | 11.03 | 11.02 | 10.98 |
| True Mean ($\mu$) of Data | 11.00 | 11.00 | 11.00 | 11.00 | 11.00 | 11.00 | 11.00 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\overline{X}$ | .99 | .97 | .97 | .96 | .96 | .97 | .95 |
| Proportion of Trials ($\hat{\beta}^2$) In Which The Confidence Interval Width Is Within Desired Limits | 1.00 | .86 | .89 | .95 | 1.00 | 1.00 | 1.00 |

TABLE VIII

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR GEOMETRIC DATA WITH AUTOREGRESSIVE ORDER OF TWO

| Variable | Fishman's Method | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|---|
| | $\beta^2=1.00$ | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Average Required Sample Size | 523 | 395 | 413 | 458 | 531 | 598 | 755 |
| Estimated Mean ($\overline{X}$) of Data | 21.55 | 22.03 | 22.03 | 22.05 | 22.07 | 22.05 | 22.03 |
| True Mean ($\mu$) of Data | 22.00 | 22.00 | 22.00 | 22.00 | 22.00 | 22.00 | 22.00 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\overline{X}$ | .92 | .97 | .97 | .97 | .96 | .95 | .95 |
| Proportion of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | 1.00 | .71 | .79 | .88 | .97 | 1.00 | 1.00 |

Appendix E.

Fishman's method with an assumption of unimodality was also used on the same data. The sample sizes indicated by this method fall between the sizes indicated by the new method for $\beta^2 = .8$ and $\beta^2 = .9$. Since these levels of $\beta^2$ would be acceptable in a practical application, the two methods can be said to produce comparable results.

## Summary

Although Graybill stated that his sample size method could be used for distributions other than normal, there is no indication in the literature that the method has ever been revised and tested for exponential data. This chapter shows the development of the sample size method for exponential or geometric data which is independent or autocorrelated. The tests which were run and reported show that the method does work even for high orders of autoregression.

Finally, the method was compared to Fishman's method for estimating sample size. The proposed method is more straight forward conceptually and computationally. It also proved more efficient than Fishman's method in terms of size of sample, of time required for compilation and execution, and of core requirements. In addition, the new method appears to be less erratic from one autoregressive order to another.

## FOOTNOTES

[1]J.W. Schmidt and R.E. Taylor, Simulation and Analysis of Industrial Systems (Illinois, 1970), p. 67.

[2]Norman L. Johnson and Samuel Kotz, Continuous Univariate Distributions - 1 (Boston, 1970), p. 216

[3]George S. Fishman, "Estimating Sample Size in Computing Simulation Experiments," Management Science, XVIII (1971), p. 35.

[4]Joe Gray, "LLRANDOM: A Psuedo-Random Number Package" (unpub. paper, Oklahoma State University, 1974).

[5]Schmidt and Taylor, p. 66.

[6]Johnson and Kotz, p. 209.

[7]Gray.

[8]Schmidt and Taylor, p. 275.

# CHAPTER IV

# COMPARISON OF SAMPLE SIZE METHOD FOR NORMAL DATA

The purpose of this chapter is to record the results of experiments designed to investigate the most appropriate method of obtaining sample size estimates when data are normally distributed and autocorrelated. The initial hypothesis was that Graybill's method could be augmented with Fishman's method for correcting for autocorrelation and, thus, could be used for autocorrelated data. Since Graybill's method allows the analyst to make a probability statement concerning the width of a confidence interval computed about the parameter of interest, the hypothesis was that smaller sample sizes would be required using Graybill's method than would be required using Fishman's method. Fishman states that the width of a confidence interval computed using the sample size generated by his method will be within the desired width every time or with a probability of 1.0.[1]

## Test Procedures

The measures of effectiveness which were identified for tests on normal data are essentially the same as were

used in the tests for exponentially and geometrically distributed data. A comparison of the mean and variance of the average sample size which was estimated by each method was of primary importance. The tests were also designed to verify that the methods were operating in accordance with the probability statements which were made concerning each method. Finally, checks were made on the actual confidence coefficients in use for Fishman's method. The tests for these criteria were Monte Carlo simulations and were performed using the program in Appendix C.

Normally distributed data points, $Y_t$, were generated through the use of the random number generator, RANDU, and a process generator described by Schmidt and Taylor.[2] The data were then transformed from independent data to autocorrelated data using one of the following data-generating models:

1) Independent data

$$X_t = .5 + Y_t \ ;$$

2) Autocorrelated data with autoregressive order of 1.0

$$X_t = .5\ X_{t-1} + .5 + Y_t \ ;$$

3) Autocorrelated data with autoregressive order of 2.0

$$X_t = .5\ X_{t-1} - .25\ X_{t-2} + .5 + Y_t .$$

Since RANDU is not completely reliable in producing sequential blocks of data in which each block is uniformly distributed, a subroutine was employed to insure that the

data used for each trial in the experiment were actually normally distributed. Any data set which failed to meet the normality requirement was discarded.

For each trial, then, a sample of 50 was initially taken and the statistics from this sample were used to estimate sample size. For Fishman's method the statistics from the initial sample were used to obtain a first estimate of the sample size, N, necessary to meet the conditions imposed by the probability statements. Instead of collecting the N - 50 new data points necessary to obtain the required sample size, a portion, $\gamma(N-50)$, of the data points were gathered. The estimating procedure was then repeated until a sufficient number of data points, N*, were obtained to satisfy the conditions imposed by the probability statements. A confidence interval was constructed about the mean of the sample of N* points. Information was then collected concerning the width of the interval and the proportion of trials in which the true mean of the population was contained in the interval.

Next, Graybill's method was used with the same 50 data points taken as the initial sample. An estimate for N* was computed using the variance of the sample (adjusted for autocorrelation) and N* new data points were collected. A confidence interval was computed and the information necessary to investigate the test criteria was collected. Graybill's procedure was repeated for five additional values of $\beta^2$.

The number of trials repeated for each method was limited by the large size of the program and the excessive time required by the program. However, after observing the variances of the estimated sample sizes for some initial runs, it was found that one hundred trials would be sufficient to insure a confidence interval about the mean sample size with a width of no more than 15 percent of the true value of the mean. (See Appendix D for the calculations which were used to verify that a sample size of one hundred would be adequate.)

## Test Results

### Comparison of Sample Size

The primary criteria for comparison of Graybill's method with Fishman's were a comparison of the average sample sizes produced by each method and a comparison of the variance of the sample sizes produced. The results of the tests are shown in Tables IX, X, and XI. For each of the three data-generating models used, Graybill's method indicated that a larger sample size would be necessary than that indicated by Fishman. Fishman's method produced a smaller average sample size than was produced by Graybill's method with even the smallest value for $\beta^2$, .65. In addition, the variance of the sample sizes was greater for Graybill's method than for Fishman's.

TABLE IX

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR NORMAL DATA WITH AUTOREGRESSIVE ORDER OF ZERO

| Variable | Fishman's Method | Graybill's Method | | | | | |
|---|---|---|---|---|---|---|---|
| | $\beta^2=1.00$ | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 250 | 339 | 350 | 378 | 423 | 463 | 557 |
| Standard Deviation of Distribution of Sample Sizes | 53.3 | 107.7 | 111.2 | 119.5 | 133.0 | 144.9 | 173.1 |
| Estimated Mean of Data ($\bar{X}$) | .506 | .506 | .505 | .505 | .506 | .505 | .507 |
| True Mean of Data ($\mu$) | .5 | .5 | .5 | .5 | .5 | .5 | .5 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\bar{X}$ | .93 | .93 | .94 | .93 | .97 | .96 | .96 |
| Proportion of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | 1.00 | .78 | .79 | .85 | .89 | .94 | .96 |

TABLE X

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR NORMAL DATA WITH AUTOREGRESSIVE ORDER OF ONE

| Variable | Fishman's Method | Graybill's Method | | | | | |
|---|---|---|---|---|---|---|---|
| | $\beta^2=1.00$ | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 233 | 325 | 335 | 361 | 404 | 442 | 531 |
| Standard Deviation of Distribution of Sample Sizes | 74.5 | 177.4 | 183.3 | 197.7 | 220.6 | 240.4 | 286.6 |
| Estimated Mean of Data ($\overline{X}$) | 1.04 | 1.03 | 1.03 | 1.03 | 1.02 | 1.01 | 1.01 |
| True Mean of Data ($\mu$) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which Is Contained In Confidence Interval About $\overline{X}$ | .90 | .93 | .94 | .94 | .95 | .95 | .94 |
| Proportion of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | 1.00 | .54 | .55 | .58 | .69 | .72 | .81 |

TABLE XI

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR NORMAL DATA WITH AUTOREGRESSIVE ORDER OF TWO

| Variable | Fishman's Method | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|---|
| | $\beta^2=1.00$ | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 212 | 237 | 350 | 378 | 422 | 463 | 556 |
| Standard Deviation of Distribution of Sample Sizes | 58.1 | 199.6 | 205.6 | 220.3 | 244.4 | 265.7 | 316.3 |
| Estimated Mean of Data ($\overline{X}$) | .6658 | .6681 | .6661 | .6653 | .6633 | .6615 | .6611 |
| True Mean of Data ($\mu$) | .6667 | .6667 | .6667 | .6667 | .6667 | .6667 | .6667 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\overline{X}$ | .93 | .94 | .96 | .96 | .96 | .96 | .94 |
| Proportion of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | 1.00 | .66 | .70 | .74 | .79 | .81 | .92 |

## Verification Of Confidence Coefficients

One of the primary purposes of these tests was to verify that Fishman's method and Graybill's method did actually produce the desired confidence coefficient, $1 - \alpha$. The results of the tests may be found in Tables IX, X, and XI. Notice that in each case the value for $1 - \hat{\alpha}$ is greater than or equal to the specified level of $1 - \alpha$, .90. Based on these tests, then, each method appears to produce the confidence coefficient which is specified. As in the case of the exponential method, Graybill's method appears to be more conservative. The values of $1 - \hat{\alpha}$ generated by Graybill's method are larger than those generated by Fishman's and would indicate that larger samples had been taken than were necessary.

## Verification of $\beta^2$

A fourth purpose of the experiment was to verify that Graybill's method produced a value for $\hat{\beta}^2$ (an estimate of probability that the computed confidence interval width was within a desired width) which was consistent with the stated value of $\beta^2$. The results of the experiment may be found in Tables IX, X, and XI. Notice that the value for $\hat{\beta}^2$ was greater than or equal to the stated value for $\beta^2$ in only five of eighteen cases. These estimates indicate that Graybill's method did not produce the sample size estimates necessary to satisfy the confidence interval

width criterion.

The method proposed by Graybill is based on the assumptions of normality and independence. Clearly the data used in these tests were not independent. However, the transformation to an autoregressive scheme should have adjusted for the autocorrelation and should have made possible the use of Graybill's method just as it made possible Fishman's method. At the point where the sample size is estimated, both methods require independence.

The problem occurred in that a departure to Graybill's method was made after the first fifty observations were taken. It was found that a correct estimate of the autoregressive order could not consistently be obtained after only fifty data points. The estimate of autoregressive order was used to correct the computed variance of the sample for autocorrelation. Fishman's method requires several iterations before a sample size can be estimated. After more data points are collected the true autoregressive order can be found more consistently. Therefore the variance of the sample was more often correctly adjusted for the presence of autocorrelation using Fishman's method. Graybill's method, however, depended on the correction for autocorrelation which was based only on the first fifty observations. Since the variance used was frequently in error, the value of $\hat{\beta}^2$ was not consistent with the stated value for $\beta^2$. Further study could be done using a larger initial sample to ascertain whether

Graybill's method can successfully be used in conjunction with Fishman's for normally distributed, autocorrelated data.

## Implications of the Probability Statements

As indicated in Chapter III, the method developed by Fishman forces the confidence interval width to be equal to or less than the desired confidence interval width in every case. The iterative process terminates only when the standard error of the mean is small enough to guarantee a confidence interval of the desired width. The implication is that the probability, $\beta^2$, is equal to 1.0.

An experiment was performed to see if Fishman's procedure could actually be said to meet a requirement that $\beta^2 = 1.0$ in terms of an ordinary interpretation of the probability statement. The interpretation given Graybill's probability statement concerning confidence interval width is that if a large number of samples were collected from a population, the width of a confidence interval constructed about the sample mean would be less than or equal to the desired width in $\beta^2$ proportion of the trials.

The experiment constructed to test this interpretation involved the use of the average size of the sample, N*, required when using Fishman's method. A number of samples of size, N*, were collected and confidence intervals were constructed about the sample means. Each interval was

checked to ascertain whether or not the width was within the specified limits. The results of this test are shown in Table XII. Notice that for each autoregressive order used, the values for $\hat{\beta}^2$ (.21, .32, .49) are much smaller than the value of 1.0 which is implied throughout Fishman's article. The result would indicate that, although the confidence interval width is forced to be within limits, a probability statement concerning $\beta^2$ cannot be interpreted in the same way as it would be interpreted using Graybill's method. A probability statement of the type made by Graybill is based on an assumption of a random sample. As shown in this test, a sample taken according to Fishman's method is not a random sample. Therefore, other methods of analysis based on an assumption of a random sample could also lead to inappropriate conclusions.

The second phase of this experiment was to test Graybill's method using a stated value of $\beta^2$ equal to the estimated value, $\hat{\beta}^2$, obtained from Fishman's method. Again the results are shown in Table XII. For each autoregressive order, the resulting sample size and the resulting value of $\hat{\beta}^2$ is given. Notice that the sample sizes are still greater than the average size calculated by Fishman's method.

A final observation in conjunction with the probability statements revolves around a special problem of simulation experiments. The problem of steady state, as defined in Chapter II, is that the estimates of the values of the para-

TABLE XII

EVALUATION OF PROBABILITY STATEMENTS:
NORMAL DATA

| | ARO = 0 | ARO = 1 | ARO = 2 |
|---|---|---|---|
| Fishman's Method | | | |
| Average Estimated Sample Size Using Fishman's Method, N* | 250 | 233 | 217 |
| Estimate of $\beta_1^2$, $\beta_1^2$; Calculated By Generating Repeated Samples Of Size N* | .21 | .32 | .49 |
| Graybill's Method: | | | |
| Estimate of Sample Size Using The $\hat{\beta}_1^2$ Value From Fishman's Method as $\beta_2^2$ | 292 | 309 | 225 |
| Resulting $\hat{\beta}_2^2$ | .54 | .48 | .63 |

meters of a system fluctuate wildly for the first observations. Therefore, an estimate for a parameter from the first fifty observations might appear to be estimating an entirely different value than the same estimate made on the basis of the next fifty observations. This problem would have the same effect on the interpretation of probability statements as the situation described by Graybill where the true value of the parameter changes in the intervening time between samples. Graybill states that if $\mu_1 \neq \mu_2$ and if $\sigma_1 \neq \sigma_2$ then his procedure will produce a confidence interval on $\mu$ which has a known confidence coefficient but which has an unknown probability of a specified width.[3] However, Fishman's method would produce a confidence interval on $\mu$ in which the confidence coefficient is not known. This situation may be responsible for the low values of $1-\alpha$ obtained by Fishman in his tests on data generated by simulation experiments.[4]

## Summary

This chapter describes the way in which Graybill's sample size procedure for normal and independent data was augmented to adjust for autocorrelation. The chapter discusses the test procedures used to validate the resulting procedure and to compare the procedure to the method developed by Fishman. The tests indicate that the adjustment for autocorrelation is imperfect. The confidence intervals which result from the use of Graybill's method

with this adjustment may be wider than desired. In addition, the necessary sample sizes indicated by Graybill's method are somewhat larger than would be required using Fishman's method.

The chapter also includes an analysis of the interpretations which can be given to the probability statements used as criteria for each method. Tests indicate that the value for $\beta^2$ used in Fishman's method cannot be interpreted as it would be using Graybill's method. The sample as taken according to the Fishman procedure is not a random sample and, therefore, ordinary interpretations of probability statements do not apply.

## FOOTNOTES

[1]George S. Fishman, "Estimating Sample Size In Computing Simulation Experiments," Management Science, XVIII (1971), p. 28.

[2]J.W. Schmidt and R.E. Taylor, Simulation And Analysis Of Industrial Systems (Homewood, Illinois, 1970), p. 265.

[3]Franklin A. Graybill, "Determining Sample Size For A Specified Width Confidence Interval," Annals Of Mathematical Statistics, XXIX (1958), p. 287.

[4]Fishman, p. 35.

# CHAPTER V

# SUMMARY, CONCLUSIONS AND EXTENSIONS

## Summary

Sample size determination is recognized by simulation theorists and practioners to be one of the remaining problems in the area of digital simulation theory. The facts that simulation output data are often autocorrelated and often fit distributions other than normal make the sample size problem one that cannot be adequately attacked using conventional statistical techniques. Therefore, the purposes of this study were as follows:

1) To develop a sample size determination technique for autocorrelated data fitting exponential or geometric data based on a method by Graybill.
2) To empirically test the method for exponential or geometric data in order to ascertain whether or not the method performs correctly.
3) To compare the method for exponential or geometric data with a method developed by Fishman.
4) To adjust Graybill's method for normal data for autocorrelation.

5) To compare Graybill's method for normal data with Fishman's method.

## Sample Size Determination For Exponential or Geometric Data

The method for determining the necessary size of a sample for simulation data fitting an exponential or a geometric distribution was based on the method for independent data proposed by Graybill in 1958. Graybill's method enables the analyst to select a value for the probability that the true mean of the population is contained in the confidence interval constructed about the sample mean. In addition, the analyst may select a value for the probability that the width of the confidence interval constructed about the sample mean is less than or equal to a desired confidence interval width. Graybill stated that similar procedures may be used to determine sample sizes for distributions other than normal. However, before this study no one had developed the method for other distributions.

This study has extended Graybill's work by showing the development of a sample size method for exponential or geometric data. The theory involved in the development is discussed in Chapter III. In order to make the procedure accessible to a practitioner of simulation who does not have a sophisticated mathematical or statistical background, an algorithm is given for the procedure in

Appendix A. Appendix A also includes a flowchart of the procedure and the documentation and program listing for a FORTRAN subroutine. This FORTRAN subroutine is designed to enable the practioner to include the process of sample size determination as an integral part of his simulation. Using this subroutine, his simulation program can access the sample size subroutine, determine the proper sample size, return to the main program to collect more data points, and continue with the simulation experiment.

## Correcting Graybill's Procedure (Normal Data) For Autocorrelation

Graybill's procedure for determining sample size requirements for normal data is based on an assumption of independence. This assumption of independence would prohibit the use of Graybill's method using simulation data which is often autocorrelated. This study shows how Graybill's procedure may be augmented by the procedure for correcting for autocorrelation as discussed by Fishman. In addition to the discussion in Chapter IV, the linkage of Graybill's procedure with Fishman's correction for autocorrelation is demonstrated in the program listed in Appendix C.

# Conclusions

## Validity of the Sample Size Method For Exponential or Geometric Data

The procedure which was developed for determining sample size requirements for exponential or geometric data was empirically tested to determine whether or not the probability statements were accurate. The probability that the true mean was contained in a confidence interval constructed about the sample mean was stated to be $1 - \alpha$. A series of samples were taken using the sample size method developed in this study. For each sample, a confidence interval was constructed about the sample mean. Each interval was checked to determine whether or not it contained the true value of the mean. In each series of tests (both exponential and geometric), the empirical value $1-\hat{\alpha}$, was at least as great as the theoretical probability, $1-\alpha$. In addition, the widths of each confidence interval were compared to a desired width. In each series of tests, the empirically developed estimate, $\hat{\beta}$, (the probability that the width of the confidence interval be within a specified limit) was at least as great as the theoretical value, $\beta^2$. These tests show that the procedure for determining sample size requirements for exponential or geometric data produces results which are consistent with the stated probability criteria.

## Comparison Of The Proposed Method With Fishman's Method

The proposed method for determining sample size, as developed in this work, was compared to Fishman's method using data generated by the program in Appendix C. It was found that an assumption of unimodally distributed sample means (rather than normal) should be used when the data are distributed exponentially or geometrically. In most cases, the exponential method produced a sample size requirement which was comparable to or less than that required by Fishman's method. In addition, the variances of the distributions of sample sizes were often less for the method developed in this study than for Fishman's. The method described in this paper required less computer time for compilation and execution and also required less core for execution. The proposed method appeared to be less eratic from one autoregressive order to another. Finally, the method developed in this study is relatively easy to understand and to compute. This one attribute alone should be extremely important from a practitioner's point of view.

## Comparison Of Graybill's Method With Fishman's Method

Graybill's method, as augmented with a correction for autocorrelation, did not compare favorably with

Fishman's method. Graybill's method appeared to perform correctly for the criterion that the probability of the true mean being contained in the confidence interval about the sample mean be at least $1-\alpha$. However, the empirically developed values of $\hat{\beta}^2$ fell well below the stated values for $\beta^2$ (the probability that the width of the confidence interval is within desired limits). An explanation of why Graybill's method did not perform well using the correction for autocorrelation may be found in Chapter IV. In addition, the sample size requirements produced by Graybill's method were consistently higher than the estimates produced by Fishman's method.

Additional Findings

The conclusions reported in the preceding section would seem to imply that Fishman's method is superior to Graybill's method for normally distributed, autocorrelated data. However, Fishman's method was also found to be suspect when one considers the method in which his sample is taken. The sampling process he prescribes is an iterative process where successive blocks of data are collected until the standard error of the mean is small enough to guarantee the desired confidence interval width. In this sense, the sample is not a random sample. Therefore, one could not take a new sample of size, $N^*$, (estimated with Fishman's iterative procedure) and expect the resulting confidence interval width to be within the

desired limits. Confirming tests are found in Chapter IV for the normal distribution and Chapter III for the exponential distribution. This finding should make the procedure developed for exponential data even more desirable than Fishman's.

Finally, in simulation experiments the steady state problem may be responsible for a change in the values of the parameters of the distribution ($\mu$ and $\sigma^2$) from one phase of sampling to another. If this be the case, Fishman's method would produce a confidence interval with a known probability of a specified width but the confidence coefficient would not be known. Although the probability of a specified width would be unknown using Graybill's procedure, the confidence coefficient would be known. Therefore, Graybill's procedure or a method based on Graybill's criteria would appear to be more appropriate for simulation data.

## Extensions

Several other studies are suggested by the results of this work. One problem encountered was the failure to successfully augment Graybill's method with the correction for autocorrelation used in Fishman's method. A form of sensitivity analysis could be used to discover whether the size of the initial sample has an effect on the correct determination of autoregressive order. A larger initial sample size might make possible a more accurate estimate

of the variance and, hence, a better value for $\hat{\beta}^2$.

Secondly, a larger initial sample when using Fishman's method might eliminate the pre-mature termination of the iterative procedure as evidenced in Chapter III. The results for Fishman's procedure appeared to be suspect for exponential data with an autoregressive order of two. A second area for research, then, would be to determine if the size of the initial sample could be responsible for unrealistically small sample size requirements.

A third area for research would be the development of Graybill's procedure for other relevant distributions. FInally, a more general area of research would be to interface the sample size problem with the related area of steady state.

# A SELECTED BIBLIOGRAPHY


Abromowitz, Milton and Irene A. Stegun. Handbook of Mathematical Functions. Washington, D.C.: National Bureau of Standards, 1964.

Ackoff, Russell L. Progress in Operations Research. New York: John Wiley and Sons, 1961.

Aigner, D.J. "Note On Verification of Computer Simulation Models." Management Science, XVIII (July, 1972), 615-619.

Burdick, Donald S. and Thomas H. Naylor. "Design of Computer Simulation Experiments For Industrial Systems." Communications of the ACM, IX (May, 1966), 329-339.

Cohen, K.J. and Richard M. Cyert. "Computer Models in Dynamic Economics" The Quarterly Journal of Economics, LXXV (February, 1961), 112-127.

Conway, R.W., B.M. Johnson and W.L. Maxwell. "Some Problems in Digital Systems Simulation." Management Science, VI (October, 1959), 92-110.

Conway, R.W. "Some Tactical Problems In Digital Simulation." Management Science, X (October, 1963), 47-61.

Ehrenfeld, S. and S. Ben-Tuvia. "The Efficiency of Statistical Simulation Procedures." Technometrics, IV (May, 1962), 257-275.

Fishman, George S. "Problems in the Statistical Analysis of Simulation Experiments: The Comparison of Means and the Length of Sample Records." Communications of The ACM, X (February, 1967), 94-99.

Fishman, G.S. "Bias Considerations in Simulation Experiments," Operations Research, XX (July, 1971), 785-790.

Fishman, George S. "Estimating Sample Size In Computing Simulation Experiments," Management Science, XVIII (September, 1971), 21-38.

Fishman, George S. Concepts and Methods in Discrete Event Digital Simulation. New York: John Wiley and Sons, 1973.

Fishman, George S. "Statistical Analysis for Queuing Simulations." Management Science, XX (November, 1973), 363-369.

Gafarian, A.V. and C.J. Anaker. "Mean Value Estimation From Digital Computer Simulation." Operations Research, XIV (January - February, 1966), 25-44.

Geisler, Murray A. "The Sizes of Simulation Samples Required to Compute Certain Inventory Characteristics With Stated Precision and Confidence." Management Science, X (January, 1964), 261-286.

Goldman, Aaron S. and R.K. Zeigler. "Comparisons of Some Two Stage Sampling Methods." Annals of Mathematical Statistics, XXXVII (August, 1966), 891-897.

Gray, Joe. "LLRANDOM: A Pseudo-Random Number Package." (Unpublished paper, Oklahoma State University, 1974).

Graybill, Franklin A. "Determining Sample Size For A Specified Width Confidence Interval." Annals of Mathematical Statistics, XXIX (March, 1958), 282-287.

Hammersley, J.M. and D.C. Handscomb. Monte Carlo Methods. New York: John Wiley and Sons, 1964.

Hillier, Frederick S. and Gerald J. Lieberman. Introduction to Operations Research. San Francisco: Holden-Inc., 1967.

Hunter, J.S. and T.H. Naylor, "Experimental Designs For Computer Simulation Experiments." Management Science, XVI (March, 1970), 422-434.

Ignall, Edward J. "On Experimental Design For Computer Simulation Experiments." Management Science, XVIII (March, 1972), 384-388.

Johnson, Norman L. and Samuel Kotz. Continuous Univariate Distributions-1. Boston: Houghton Mifflin 1970.

Kleijnen, Jack P.C., Thomas H. Naylor and Terry G. Seaks. "The Use of Multiple Ranking Procedures to Analyze Simulations of Management Systems: A Tutorial." Management Science, XVIII (February, 1972), B-245-B-257.

McMillan, Claude and Richard F. Gonzales. Systems Analysis. Homewood, Illinois: Richard D. Irwin, Inc., 1965.

Meier, Robert C., William T. Newell and Harold L. Pazer. Simulation In Business and Economics. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1969.

Miller, H.D. The Theory of Stochastic Processes. New York: John Wiley and Sons, 1965.

Mize, Joe H. and J. Grady Cox. Essentials of Simulation. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

Moran, P.A.P. The Theory of Storage. New York: John Wiley and Sons, 1959.

Morse, Phillip M. "Stochastic Properties of Waiting Lines." Operations Research, III (August, 1955), 255-261.

Narula, Subhash C. and Franz S. Li. "Sample Size Calculations in Exponential Life Testing." Technometrics, XVII (May, 1975), 229-231.

Naylor, Thomas H., Joseph L. Balintfy, et al. Computer Simulation Techniques. New York: John Wiley and Sons, 1966.

Naylor, T.H. and J.M. Finger. "Verification of Computer Simulation Models." Management Science, XIV (October, 1967), B-92-B-101.

Naylor, Thomas H., Kenneth Wertz and Thomas Wonnacott. "Methods For Analyzing Data From Computer Simulation Experiments." Communications of the ACM, X (November, 1967), 703-710.

Schmidt, J.W. and R.E. Taylor. Simulation and Analysis of Industrial Systems. Homewood, Illinois: Richard D. Irwin, Inc., 1970.

Shannon, Robert E. Systems Simulation: The Art and Science. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1975.

Stein, Charles. "A Two Sample Test For A Linear Hypothesis Whose Power Is Independent of the Variance." Annals of Mathematical Statistics, XVI (September, 1945), 243-258.

Tocher, K.D. The Art of Simulation. Princeton, New Jersey: D. Van Nostrand Co., 1963.

Van Horn, Richard L. "Validation of Simulation Results." Management Science, XVII (January, 1971), 247-257.

Williams, Donald R. and David L. Weeks. "A Technique For Designing and Augmenting Simulation Experiments." Management Science, XX (June, 1974), 1385-1392.

# APPENDIXES

# APPENDIX A

## THE PROPOSED METHOD FOR EXPONENTIAL DATA: ALGORITHM AND FORTRAN SUBROUTINE

The proposed method, as it was revised for exponential or geometric data, can be most explicitly stated in algorithm form. The following is a detailed explanation of the steps which must be completed in order to implement this method.

1. Set upper and lower bounds on the sample size of zero: SSL = SSR = 0.
2. Collect a random sample of M observations.
3. Compute a value, Z, which is twice the sum of M observations: $Z = 2 \cdot \Sigma v_j$, where $v_j$ is the value of the $j^{th}$ observation in the sample.
4. Let GN = M.
5. Select a value for D, the desired width of the confidence interval computed about the sample mean.
6. Select values for $\beta^2$, the probability that the confidence interval width will be less than or equal to the desired width, and $1-\alpha$, the probability that the confidence interval about

the sample mean will contain the true value of the mean.

7. Find the standard normal deviates, $Z_{\beta}$ and $Z_{1-\alpha/2}$, for the chosen values of $\beta^2$ and $1 - \alpha$.

8. Approximate the values of $f(n)$, $\chi_1^2$, $\chi_2^2$ and $\chi_3^2$ using the following formulas:[1]

   a) $f(n) = (Z_{\beta} + \sqrt{4(GN) - 1})^2/2$ (A.1)

   b) $\chi_1^2 = (Z_{1-\alpha/2} + \sqrt{4(GN) - 1})^2/2$ (A.2)

   c) $\chi_2^2 = (-Z_{1-\alpha/2} + \sqrt{4(GN) - 1})^2/2$ (A.3)

   d) $\chi_3^2 = (-Z_{\beta} + \sqrt{4M - 1})^2/2.$ (A.4)

9. Calculate CN according to the following formula:

   $CN = 1/\chi_2^2 - 1/\chi_1^2$ (A.5)

10. Calculate a trial value for D, D* as follows:

    $D^* = [f(n) \cdot CN \cdot Z] / \chi_3^2$ (A.6)

11. Calculate the absolute difference between the values for D* and D:

    $DIFF = |D^* - D|$ (A.7)

12. If the value calculated for DIFF is less than or equal to .001, use the value selected for GN as the necessary sample size, N, and go to step 15. If not, continue.

13. If D is less than D*, set the new lower bound on the sample size, SSL, equal to GN. If the upper bound, SSR, is equal to zero, the new

value of GN is twice the lower limit, SSL. If SSR is not equal to zero, then the new value of GN is half-way between the upper and lower bounds:

$$GN = SSL + .5(SSR - SSL) \qquad (A.8)$$

Return to Step 8.

14. If D is greater than D*, set a new upper bound on the sample size equal to GN. Select a new value for GN which is halfway between the upper and lower bounds on the sample size:

$$GN = SSR - .5(SSR - SSL). \qquad (A.9)$$

Return to Step 8.

15. Select a new sample of data points, $u_j$, of size M.

16. Compute the lower confidence limit on the mean as follows:

$$CL_L = 2\Sigma u_j / x_1^2. \qquad (A.10)$$

17. Compute the upper confidence limit on the mean as follows:

$$CL = 2\Sigma u_j / x_2^2 \qquad (A.11)$$

A desirable property of a sample size method for simulation experiments is the capability of incorporating the method in an on-going simulation.[2] The remainder of this appendix describes a FORTRAN subroutine for the exponential method which can be used in simulations written in FORTRAN or GASP.

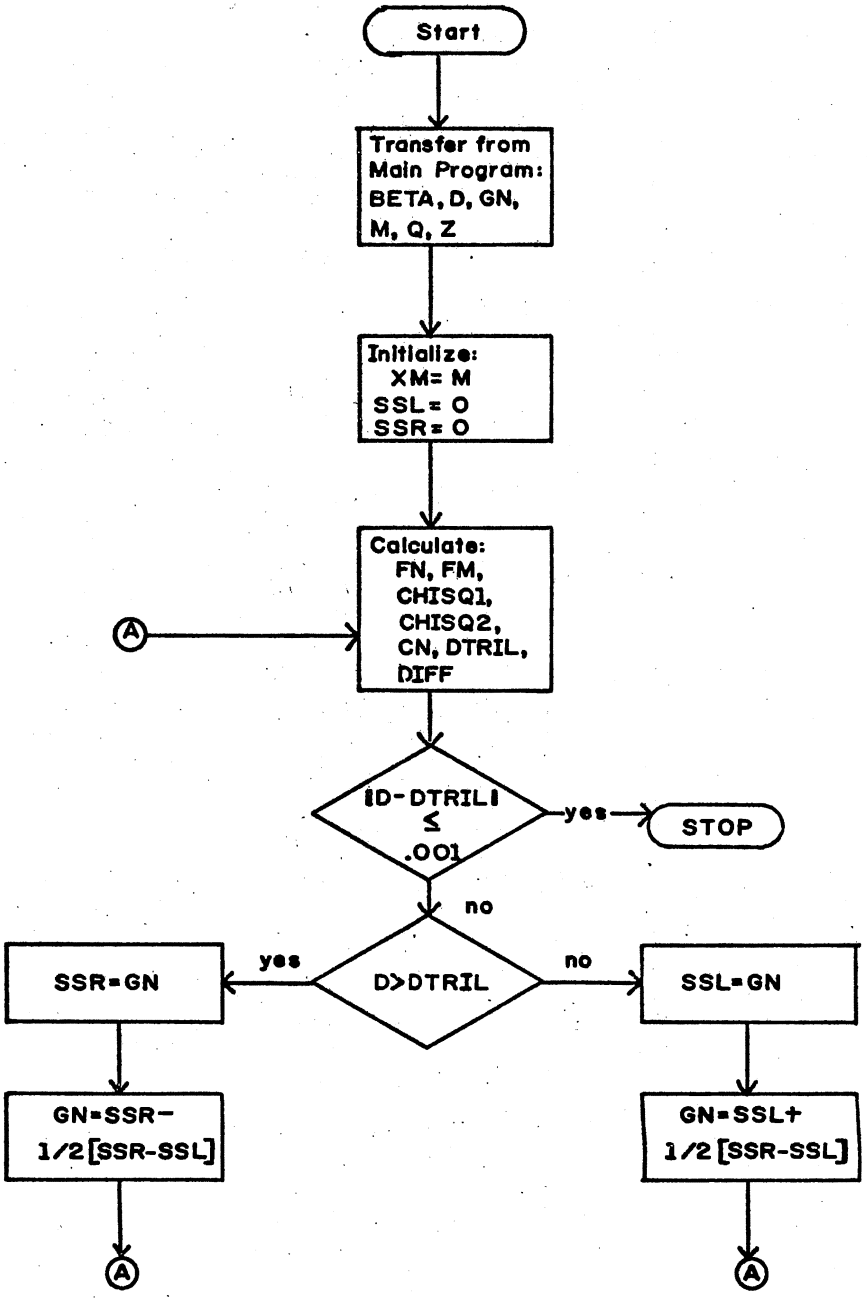


Figure 6. Flowchart Of Sample Size Method For Exponential Data

```
0001              SUBROUTINE EXPSZ (M,GN,Q,Z,D,BETA,CHISQ1,CHISQ2)
0002              XM = M
CCC3              SSL = 0.0
CC04              SSR = 0.0
0005            1 FN = ((BETA + SQRT(4.* GN - 1.))**2.) / 2.
C006              FM = ((-1. * BETA + SQRT(4. * XM - 1.))**2.) / 2.
CC07              CHISQ1 = ((Q + SQRT(4. * GN - 1.0))**2.) / 2.
0008              CHISQ2 = ((-1.0 * Q + SQRT(4.*GN - 1.0))**2.) / 2.
0009              CN = 1. / CHISQ2 - 1. / CHISQ1
C010              DTRIL = (FN*CN*Z) / FM
0011              DIFF = ABS(D - DTRIL)
0012              IF (DIFF.LE.0.001) GO TO 4
0013              IF (D.GT.DTRIL) GO TO 3
CC14              SSL = GN
0015              IF (SSR.EQ.0.0) GO TO 2
0016              GN = SSL + .5*(SSR-SSL)
CC17              GO TO 1
0018            2 GN = 2.*SSL
0019              GO TO 1
CC2C            3 SSR = GN
0021              GN = SSR - .5*(SSR - SSL)
0022              GO TO 1
CC23            4 CONTINUE
0024              RETURN
0025              END
```

Figure 7. FORTRAN Subroutine For Computing Sample Sizes For Exponential Data

## FOOTNOTES

[1]J.W. Schmidt and R.E. Taylor, Simulation And Analysis of Industrial Systems (Homewood, Illinois, 1970), p. 618.

[2]George S. Fishman, "Estimating Sample Size In Computing Simulation Experiments," Management Science, XVIII (September, 1971), p. 21.

# APPENDIX B

## A PROGRAM FOR TESTING THE PROPOSED METHOD

As explained in Chapter III, a program was written to allow testing of the new method for various data-generating models. These autoregressive models, specified by the programmer, operate on exponential data with $\mu = 2$ to produce exponential distributions with new means. In this way data can be generated with a variety of means and autoregressive orders.

To use this program, the analyst must specify six values of $Z_\beta$. The value, $\beta^2$ is the desired probability that the confidence interval width computed for a sample will be no wider than a width specified by the programmer. Thus, the value, $Z_\beta$, is a normal variate with probability, $\beta$. For each iteration, then, this program will produce six different sample size estimates which correspond to the six different values for $\beta^2$.

Within the program, the analyst may change the values for Q, D, GN, M or ITER. The definitions for these variable names are given below:

The subroutine is based on the concept of binary search to find the smallest integral value of n satisfying equation (3.18). A flow chart of the subroutine is shown in Figure 6 and a print-out of the subroutine is given in Figure 7.

Variable names used in the subroutine are as follows:

| | |
|---|---|
| BETA | Z or a standard normal variate for the probability, $\beta$ |
| CHISQ1 | $\chi^2_1$ |
| CHISQ2 | $\chi^2_2$ |
| D | Desired confidence interval width |
| DTRIL | Trial confidence interval width based on a trial sample size |
| DIFF | Absolute difference between the desired confidence interval width and the trial confidence interval width |
| FN | f(n) |
| FM | $\chi^2_3$ |
| GN | Trial sample size |
| M | Size of first sample |
| Q | $Z_{1-\alpha/2}$ or a standard normal variate with probability, $1-\alpha/2$ |
| SSL | Lower limit for the estimated sample size |
| SSR | Upper limit for the estimated sample size |

| | |
|---|---|
| Z | Twice the sum of observations in the first sample |

The following variables should be read in to the main program: BETA, D, GN, M and Q. The variable, Z, will be calculated by the main program. CHISQ1, CHISQ2, and GN must be transferred back to the main program so that a sample of size GN may be collected and a confidence interval computed about the mean of that sample.

```
C     THIS PROGRAM CALCULATES .90 CONFIDENCE INTERVALS AND COMPUTES
C     THE SAMPLE SIZE NECESSARY FOR OBTAINING A SPECIFIED WIDTH INTERVAL
C     FOR EXPONENTIAL DATA ACCORDING TO GRAYBILLS METHOD.
0001        DIMENSION X(1200)
0002        DIMENSION BETA1(6)
0003        DIMENSION Y(1200)
CCC4        DIMENSION N(6)
0005        DIMENSION AMEAN(6)
0006        DIMENSION SUMCM(6)
CC07        DIMENSION SUMCW(6)
C008        DIMENSION SUMMN(6)
0009        DIMENSION SUMN(6)
CC10        PRINT 201
0011      201 FORMAT (41X,'COMPUTING SAMPLE SIZE WITH A SPECIFIED',//,47X,'WIDTH
           $ CONFIDENCE INTERVAL',///)
0012        PRINT 202
C013      202 FORMAT (' THE DATA ARE EXPONENTIALLY DISTRIBUTED WITH A MEAN OF 9.
           $0.',/,' A 90 PERCENT CONFIDENCE INTERVAL WILL BE CALCULATED WITH A
           $ MAXIMUM WIDTH OF 2.0.',///)
0014        PRINT 203
C015      203 FORMAT (42X,'RESULTS FOR 100 INDEPENDENT TRIALS',/)
0016        PRINT 204
0017      204 FORMAT (' TRIAL',5X,'VARIABLE',21X,'.65',11X,'.70',11X,'.80',11X,'
           $.90',11X,'.95',11X,'.99',//)
C018        READ 101,(BETA1(I),I=1,6)
0019      101 FORMAT (6(F5.4,5X))
0020        CHISQ1 = 0.0
0021        CHISQ2 = 0.0
0022        Q = 1.645
0023        D = 2.
C024        ITER = 5
0025        DO 1 I = 1,6
0026        SUMCM(I) = 0.0
0027        SUMCW(I) = 0.0
CC28        SUMMN(I) = 0.0
0029        SUMN(I) = 0.0
C030      1 CONTINUE
CC31        IX = 7459183
0032        DO 12 IT = 1,ITER
C033        CALL OVFLOW
0034        CALL EXPON (IX,Y,1200)
0035        DO 2 I=1,1200
C036        Y(I) = Y(I) * 2.0
0037        IF(I.NE.1) GO TO 20
0038        X(1) = .5 + Y(1)
0039        GO TO 2
C04C       20 IF(I.NE.2) GO TO 21
0041        I1 = I - 1
0042        X(I) = .5 * X(I1) + .5 + Y(I)
C043        GO TO 2
0044       21 I1 = I - 1
0045        I2 = I - 2
CC46        X(I) = .5 * X(I1) + .25 * X(I2) + .5 + Y(I)
C047      2 CONTINUE
0048        DO 11 J = 1,6
0049        GN = 50.
C05C        M = 50
0051        L = 1
C052        Z1= 0.0
C053      3 DO 4 I = L,M
0054        Z1= Z1+ X(I)
0055      4 CONTINUE
CC56        Z = 2.*Z1
0057        IF (M.NE.50) GO TO 5
0058        BETA = BETA1(J)
CC59        CALL EXPSZ (M,GN,Q,Z,D,BETA,CHISQ1,CHISQ2)
0060        N(J) = GN
0061        L = M + 1
0062        M = N(J) + M
CC63        Z1 = 0.
0064        GO TO 3
```

Figure 8. FORTRAN Program For Testing The Sample Size Method Developed For Exponential Data

```
0065            5 GN = N(J)
0066              AMEAN(J) = Z1/GN
0067              CONFR = Z/CHISQ2
0068              CONFL = Z/CHISQ1
0069              IF(CONFR - 10.0) 8,7,6
0070            6 IF(CONFL - 10.0) 7,7,8
0071            7 SUMCM(J) = SUMCM(J) + 1.
0072            8 WIDTH = CONFR - CONFL
0073              IF (WIDTH - D) 9,9,10
0074            9 SUMCW(J) = SUMCW(J) + 1.
0075           10 SUMMN(J) = SUMMN(J) + AMEAN(J)
0076              SUMN(J) = SUMN(J) + N(J)
0077           11 CONTINUE
0078              PRINT 205, IT  ,N(1),N(2),N(3),N(4),N(5),N(6)
0079          205 FORMAT (1X,I3,5X,'SAMPLE SIZE',8X,6(10X,I4))
0080              PRINT 206, AMEAN(1),AMEAN(2),AMEAN(3),AMEAN(4),AMEAN(5),AMEAN(6)
0081          206 FORMAT (9X,'MEAN',16X,6(7X,F7.3),//)
0082           12 CONTINUE
0083              DO 13 J = 1,6
0084              N(J) = SUMN(J) / ITER
0085              SUMMN(J) = SUMMN(J) / ITER
0086              SUMCM(J) = SUMCM(J) / ITER
0087              SUMCW(J) = SUMCW(J) / ITER
0088           13 CONTINUE
0089              PRINT 207
0090          207 FORMAT (1H1,/////,51X,'SUMMARY OF RESULTS',///)
0091              PRINT 208
0092          208 FORMAT (48X,'.65',10X,'.70',10X,'.80',10X,'.90',10X,'.95',10X,'.99
                 $',//)
0093              PRINT 209, N(1),N(2),N(3),N(4),N(5),N(6)
0094          209 FORMAT (8X,'SAMPLE SIZE',19X,6( 9X,I4),/)
0095              PRINT 210,SUMMN(1),SUMMN(2),SUMMN(3),SUMMN(4),SUMMN(5),SUMMN(6)
0096          210 FORMAT (8X,'MEAN',27X,6(6X,F7.4),/)
0097              PRINT 211,SUMCM(1),SUMCM(2),SUMCM(3),SUMCM(4),SUMCM(5),SUMCM(6)
0098          211 FORMAT (8X,'PROPORTION OF TRIALS',/,8X,'TRUE MEAN CONTAINED',/,8X,
                 $'IN CONFIDENCE INTERVAL',9X,6(6X,F7.4),/)
0099              PRINT 212,SUMCW(1),SUMCW(2),SUMCW(3),SUMCW(4),SUMCW(5),SUMCW(6)
0100          212 FORMAT (8X,'PROPORTION OF TRIALS',/,8X,'CONFIDENCE INTERVAL',/,8X,
                 $'WIDTH IS WITHIN LIMITS',9X,6(6X,F7.4))
0101              STOP
0102              END


0001              SUBROUTINE EXPSZ (M,GN,Q,Z,D,BETA,CHISQ1,CHISQ2)
0002              XM = M
0003              SSL = 0.0
0004              SSR = 0.0
0005            1 FN = ((BETA + SQRT(4.* GN - 1.))**2.) / 2.
0006              FM = ((-1. * BETA + SQRT(4. * XM - 1.))**2.) / 2.
0007              CHISQ1 = ((Q + SQRT(4. * GN - 1.0))**2.) / 2.
0008              CHISQ2 = ((-1.0 * Q + SQRT(4.*GN - 1.0))**2.) / 2.
0009              CN = 1. / CHISQ2 - 1. / CHISQ1
0010              DTRIL = (FN*CN*Z) / FM
0011              DIFF = ABS(D - DTRIL)
0012              IF (DIFF.LE.0.001) GO TO 4
0013              IF (D.GT.DTRIL) GO TO 3
0014              SSL = GN
0015              IF (SSR.EQ.0.0) GO TO 2
0016              GN = SSL + .5*(SSR-SSL)
0017              GO TO 1
0018            2 GN = 2.*SSL
0019              GO TO 1
0020            3 SSR = GN
0021              GN = SSR - .5*(SSR - SSL)
0022              GO TO 1
0023            4 CONTINUE
0024              RETURN
0025              END
```

Figure 8. (Continued)

# APPENDIX C

## A PROGRAM FOR TESTING FISHMAN'S METHOD FOR DETERMINING SAMPLE SIZE

This program is designed to produce comparisons of sample size estimates using Fishman's method and Graybill's method on normal, autocorrelated data. Although Graybill's method assumes independence, the variance adjusted for autocorrelation is used to circumvent this assumption.

The program will generate the following data:

1) The estimated sample size averaged over the number of iterations.
2) The mean averaged over the samples taken.
3) The proportion of iterations in which the true autoregressive order was actually found.
4) The proportion of iterations in which the computed confidence interval actually contained the true mean.
5) The proportion of iterations in which the computed confidence interval width was less than or equal to the desired interval width.

The input data necessary are as follows:

| | |
|---|---|
| Q | A standard normal variate, $Z_{1-\alpha/2}$, with probability, $1-\alpha/2$. |
| D | Desired confidence interval width |
| GN | The initial "guess" for the correct estimate of the necessary sample size |
| M | The size of the first sample of each iteration |
| ITER | The number of iterations desired of the program. |

In addition, different autoregressive models may be used to produce a desired level of autocorrelation. For example in the accompanying print-out, statements 37 through 46 produce data with autoregressive order two. Once the model has been selected, the true value for the mean of that model should be placed in statements 69 and 70.

The program will generate the following data:

1) The estimated sample size averaged over the number of iterations.
2) The mean averaged over the samples taken.
3) The proportion of iterations in which the computed confidence interval actually contained the true mean.
4) The proportion of iterations in which the computed confidence interval width was less than or equal to the desired confidence interval width.

Card 1

| | |
|---|---|
| columns 1-5 | A five digit integer to be used as a random number seed |

Card 2

| | |
|---|---|
| columns 1-5 | A right-justified integer giving the size of the initial sample of each iteration |
| columns 11-15 | A standard normal variate with probability $1-\alpha/2$ |
| columns 21-25 | A standard normal variate used in computing confidence intervals around the coefficients of the autoregressive equations |
| columns 31-35 | A number between 0 and 1 which indicates the proportion, G, of the observations remaining to be collected which are taken on the subsequent sample |
| columns 41-46 | The specified width, D, of the confidence interval |
| columns 51-55 | A right-justified integer giving the largest autoregressive order to be tested |
| columns 61-65 | A right-justified integer giving the number of methods to be compared (one plus the number of $\beta$ values used) |

columns 71-75 — The number of iterations to be completed

Card 3

columns 1-5 — The true autoregressive order of the data-generating model

columns 11-15 — The true mean of the data

Card 4

columns 1-5 — The sample size to be used as a first "guess" in Graybill's method

Card 5

columns 1-5, 11-15, ..., 51-55 — The standard normal variates with probability, $\beta$, corresponding to the various $\beta^2$ values to be used for Graybill's method.

The programmer may also specify a data-generating model which produces the desired order of autocorrelation. For instance, in the following program, the model is located at statement 76.

```
00010       PROGRAM SAMPSIZ(INPUT,OUTPUT)
00050       DIMENSION XM1(8)
00060       DIMENSION SUMSQ1(8)
00070       DIMENSION IEXES(8)
00080       DIMENSION BETA(8)
00090       DIMENSION XBAR(8)
00100       DIMENSION XVAR(8)
00110       DIMENSION SUMCM(8)
00120       DIMENSION SUMCW(8)
00130       DIMENSION SUMP(8)
00140       DIMENSION SUMMN(8)
00150       DIMENSION SUMVR(8)
00160       DIMENSION SUMNR(8)
00170       DIMENSION NXN(8)
00180       DIMENSION PMCI(8)
00190       DIMENSION PCWC(8)
00200       DIMENSION PAROC(8)
00210       DIMENSION X(1500),Y(1500),X1(1500)
00220       DIMENSION W(10)
00230       DIMENSION VAR(10)
00240       DIMENSION B(10,10)
00250       DIMENSION C(10)
00260       READ 2000,IX
00270  2000 FORMAT (I5)
00280       READ 100,MSTAR,Q,P,G,D,LR,NMETH,ITERL
00290   100 FORMAT (I5,4(5XF5.0),3(5XI5))
00300       READ 2501, TARO,TMEAN
00310  2501 FORMAT (2F10.0)
00320       READ 707,BN
00330   707 FORMAT (F10.0)
00340       PRINT 2502
00350  2502 FORMAT (41X,38HCOMPUTING SAMPLE SIZE WITH A SPECIFIED,//,47X,25HWI
00360      1DTH CONFIDENCE INTERVAL,///)
00370       PRINT 2503
00380  2503 FORMAT (77H THE MODEL TO BE USED IS X(T) = .5 + Y(T) , WHERE Y IS
00390      1DISTRIBUTED AS N(0,1).)
00400       PRINT 2504, TARO
00410  2504 FORMAT (34H THE TRUE AUTOREGRESSIVE ORDER IS ,F2.0,2H .)
00420       PRINT 2505, TMEAN
00430  2505 FORMAT (31H THE TRUE VALUE OF THE MEAN IS ,F10.4,2H .)
00440       PRINT 2506 , D
00450  2506 FORMAT (23H A CONFIDENCE WIDTH OF , F10.2,13HIS SPECIFIED.,///)
00460       PRINT 2507,ITERL
00470  2507 FORMAT (    42X,11HRESULTS FOR ,I5,20H INDEPENDENT TRIALS.)
00480       PRINT 2508
00490  2508 FORMAT (6H TRIAL,5X,8HESTIMATE,5X,7HFISHMAN,41X,8HGRAYBILL)
00500       PRINT 2509
00510  2509 FORMAT (5H  NO.,34X,3H.65,11X,3H.70,11X,3H.80,11X,3H.90,11X,3H.95
00520      1,11X,3H.99,//)
00530       DO 95 I = 1,NMETH
00540       SUMMN(I) = 0.
00550       SUMVR(I) = 0.
00560       SUMNR(I) = 0.
00570       BETA(I) = 0.
00580       SUMCM(I) = 0.
00590       SUMCW(I) = 0.
00600       SUMP(I)  = 0.
00610       SUMNR(I) = 0.
00620       SUMMN(I) = 0.
00630    95 SUMVR(I) = 0.
00640       READ 2500,(BETA(I),I=2,NMETH)
00650  2500 FORMAT (6(F5.4,5X))
00670       DO 99 ITER = 1,ITERL
00680    98 DO 2006 J = 1,1500
00690       CALL RANDU(IX,IY,RN)
00700       IX = IY
00710       IF(RN-.5) 2002,2001,2001
00720  2001 W1 = 1. - RN
00730       GO TO 2003
00740  2002 W1 = RN
```

Figure 9. FORTRAN Program For Testing Fishman's Method For Computing Sample Size

```
00750  2003 Z = 1. / (W1**2.)
00760       Z1 = ALOG(Z)
00770       V = SQRT(Z1)
00780       GW = V - ((2.515517 + .802853 * V + .010328 * (V**2.)) / (1. +
00790      11.432788 * V + .189269 * (V**2.) + .001308 * (V**3.)))
00800       IF(RN-.5) 2004,2005,2005
00810  2004 Y(J) = -1. * GW
00820       GO TO 2009
00830  2005 Y(J) = GW
00840  2009 CONTINUE
00860       X(J) = Y(J) + .5
00950  2006 CONTINUE
00960    94 DO 97 METH = 1,NMETH
00970       M =MSTAR
00980       N = M
00990       LR1 = LR + 1
01000       SUMX = 0.
01010       K = 0
01020       KDC = K + 1
01030 C
01040 C     FIND THE MEAN OF ALL DATA POINTS
01050 C
01060     1 SUMX = 0.
01070       DO 2 J = KDC,N
01080     2 SUMX = SUMX + X(J)
01090       GN = BN
01100       ZN = N - K
01110       XBAR1= SUMX / ZN
01120 C
01130 C     SUBTRACT MEAN FROM EACH DATA POINT
01140 C
01150       DO 3 J = KDC,N
01160     3 X1(J) = X(J) - XBAR1
01170 C
01180 C     CALCULATE THE VECTOR OF COVARIANCES
01190 C
01200       DO 5  I = 1,LR1
01210       K3 = N - I + 1
01220       K2 = I + K
01230       SUM = 0
01240       DO 4 J = KDC,K3
01250       SUM = SUM + (X1(J) * X1(K2))
01260     4 K2 = K2 + 1
01270     5 C(I) = SUM / ZN
01280 C
01290 C     COMPUTE THE COEFFICIENTS OF THE AUTOREGRESSIVE EQUATIONS
01300 C
01310       B(1,1) = 1.
01320       J = 2
01330       DO 71 I=1,LR
01340       SUM1 = 0.
01350       SUM2 = 0.
01360       B(J,1) = 1.
01370       DO 6 KS = 1,I
01380       K2 = I - KS + 2
01390       SUM1 = SUM1 + (B(I,KS) * C(K2))
01400       SUM2 = SUM2 + B(I,KS) * C(KS)
01410       IF(I.EQ.KS) GO TO 72
01420     6 CONTINUE
01430    72 B(J,J) = -1. * (SUM1 / SUM2)
01440       IF(J.LT.3) GO TO 70
01450       DO 7 KS = 2,I
01460       K2 = I - KS + 2
01470     7 B(J,KS) = B(I,KS) + (B(J,J) * B(I,K2))
01480    70 J = J + 1
01490    71 CONTINUE
01500 C
01510 C     CALCULATE SAMPLE RESIDUAL VARIANCES
01520 C
01530       DO 10  I=2,LR1
01540       SUM2 = 0.
01550       K2 = I + 1 + K
01560       NPLUS = N + 1
```

Figure 9. (Continued)

```
01570       DO 9 J=K2,NPLUS
01580       SUM1 = 0.
01590       DO 8  L1 =1,I
01600       K3 = J - L1
01610     8 SUM1 = SUM1 + (B(I,L1) * X1(K3))
01620     9 SUM2 = SUM2 + SUM1**2
01630       AN = ZN - I + 1
01640    10 VAR(I) = SUM2 / AN
01650 C
01660 C     CALCULATE CONFIDENCE LIMITS AND DETERMINE AUTOREGRESSIVE ORDER
01670 C
01680       I = 1
01690       KP = 1.
01700    11 I = I + 1
01710       W(I) = 1. - (B(I,I)**2)
01720       CONFR = B(I,I) + P * ((W(I) / ZN)**.5)
01730       CONFL = B(I,I) - P * ((W(I) / ZN)**.5)
01740       IF(CONFR) 13,14,12
01750    12 IF(CONFL) 14,14,13
01760    13 KP = I
01770    14 IF (I.NE.LR1) GO TO 11
01780 C
01790 C     CALCULATE THE NECESSARY SAMPLE SIZE
01800 C
01810    15 IF (KP.NE.1) GO TO 16
01820       XM = C(1)
01830       K = 0
01840       GO TO 18
01850    16 SUM1 = 0.
01860       DO 17  I = 1,KP
01870    17 SUM1 = SUM1 + B(KP,I)
01880       XM = VAR(KP) / (SUM1**2)
01890       K = (XM / C(1)) - 1.
01900    18 G1 = (Q**3. + Q) / 4.
01910       G2 = (5. * (Q**5.) + 16. * (Q**3.) + 3. * Q) / 96.
01920       G3 = (3. * (Q**7.) + 19. * (Q**5.) + 17. * (Q**3.) - 15. * Q) /
01930      1384.
01940       G4 = (79. * (Q**9.) + 776. * (Q**7.) + 1482 * (Q**5.) - 1920. *
01950      1 (Q**3.) - 945. * Q) / 92160.
01960       DF = ((ZN * C(1)) / XM) - 1.
01970       QE = Q + (G1 / DF) + (G2 / (DF**2.)) + (G3 / (DF**3.)) +
01980      1(G4 / (DF**4.))
01990       IF(METH.NE.1) GO TO 23
02000       CW = D/2.
02010       XN = XM * ((QE/CW)**2)
02020       GO TO 24
02030    23 IF(N.NE.MSTAR) GO TO 19
02040       SUMSQ = XM * ZN
02050       FN = ((BETA(METH) + SQRT(2. * GN - 3.)) **2.) / 2.
02060       FM = ((-1.*BETA(METH) + SQRT(2. * ZN - 3.)) **2.) / 2.
02070       TO = QE
02080       DTRIL = (2.* TO * SQRT(SUMSQ) * SQRT(FN)) / (SQRT(FM) * SQRT(GN *
02090      1(GN - 1.)))
02100       IF (DTRIL.LE.D) GO TO 22
02110       GN1 = (2. * TO * SQRT(SUMSQ) * SQRT(FN)) / (SQRT(FM) * D)
02120       NGN1 = GN1
02130       NGN = GN
02140       IF(NGN1.EQ.NGN) GO TO 21
02150       GN = GN1
02160       GO TO 23
02170    21 GN = GN1 + 1.
02180       GO TO 23
02190    22 XN = GN
02200       M = GN
02210       K = N
02220       KDC = K +1
02230       N = N + M
02240       GO TO 1
02250    24 KDC = K + 1
02260       ZN = N - K
02270       IF (XN.LE.ZN) GO TO 19
02280       M = G * (XN - ZN)
02290       N = N + M
```

Figure 9. (Continued)

```
02300       IF(M.GE.5) GO TO 26
02310       NXN1= XN
02320       NZN = ZN
02330       M = NXN1- NZN  +1
02340       N = N + M
02350    26 GO TO 1
02360    19 SUM = 0.
02370       DO 20 I = KDC,N
02380    20 SUM = SUM + X(I)
02390       XBAR(METH) = SUM / ZN
02400       XVAR(METH) = XM / ZN
02410       CONFR = XBAR(METH) + (SQRT(XVAR(METH))*QE)
02420       CONFL = XBAR(METH) - (SQRT(XVAR(METH))*QE)
02430       IF (CONFR - TMEAN) 63,62,61
02440    61 IF (CONFL - TMEAN) 62,62,63
02450    62 SUMCM(METH) = SUMCM(METH) + 1.
02460    63 WIDTH = CONFR - CONFL
02470       IF (WIDTH - D) 64,64,65
02480    64 SUMCW(METH) = SUMCW(METH) + 1.
02490    65 KP = KP - 1
02500       IF (KP.NE.TARO) GO TO 66
02510       SUMP(METH) = SUMP(METH) + 1.
02520    66 SUMMN(METH) = SUMMN(METH) + XBAR(METH)
02530       SUMVR(METH) = SUMVR(METH) + XVAR(METH)
02540       SUMNR(METH) = SUMNR(METH) + XN
02550       SUMSQ1(METH) = SUMSQ
02560       XM1(METH) = XM
02570       NXN(METH) = XN
02580    97 CONTINUE
02590       PRINT 2511, ITER,NXN(1),NXN(2),NXN(3),NXN(4),NXN(5),NXN(6),NXN(7)
02600  2511 FORMAT (1X,I4,4X,2HN*,13X,I4,6(10X,I4))
02610       PRINT 2560, XBAR(1),XBAR(2),XBAR(3),XBAR(4),XBAR(5),XBAR(6),XBAR(7
02620      1)
02630  2560 FORMAT (9X,4HMEAN, 9X,F6.4,6(8X,F6.4))
02640       PRINT 2562,XVAR(1),XVAR(2),XVAR(3),XVAR(4),XVAR(5),XVAR(6),XVAR(7)
02650  2562 FORMAT(9X,8HVARIANCE,/,11X,7HOF MEAN,4X,F6.4,6(8X,F6.4))
02660       PRINT 2561,XM1(1),XM1(2),XM1(3),XM1(4),XM1(5),XM1(6),XM1(7)
02670       PRINT 2559,SUMSQ1(2),SUMSQ1(3),SUMSQ1(4),SUMSQ1(5),SUMSQ1(6),
02680      $SUMSQ1(7)
02690  2559 FORMAT(9X,6HSUM OF,/,11X,7HSQUARES,10X,6(5X,F9.4))
02700    99 CONTINUE
02710  2561 FORMAT (9X,8HVARIANCE,5X,F6.4,6(8X,F6.4),//)
02720       PRINT 2512
02730  2512 FORMAT (1H1,/////,51X,18HSUMMARY OF RESULTS,///)
02740       PRINT 2513
02750  2513 FORMAT (33X,7HFISHMAN,36X,8HGRAYBILL)
02760       PRINT 2514
02770  2514 FORMAT (48X,3H.65,10X,3H.70,10X,3H.80,10X,3H.90,10X,3H.95,10X,3H.9
02780      19,//)
02790       DO 67 I = 1,NMETH
02800       XBAR(I) = SUMMN(I) / ITERL
02810       XVAR(I) = SUMVR(I) / ITERL
02820       NXN(I) = SUMNR(I) / ITERL
02830       PMCI(I) = SUMCM(I) / ITERL
02840       PCWC(I) = SUMCW(I) / ITERL
02850    67 PAROC(I) = SUMP(I) / ITERL
02860       PRINT 2515, NXN(1), NXN(2), NXN(3), NXN(4), NXN(5), NXN(6), NXN(7)
02870  2515 FORMAT (8X,2HN*,25X,I3,6(10X,I3),/)
02880       PRINT 2516,XBAR(1),XBAR(2),XBAR(3),XBAR(4),XBAR(5),XBAR(6),XBAR(7)
02890  2516 FORMAT (8X,4HMEAN,21X,F6.4,6(7X,F6.4),/)
02900       PRINT 2517,PAROC(1),PAROC(2),PAROC(3),PAROC(4),PAROC(5),PAROC(6),
02910      1PAROC(7)
02920  2517 FORMAT (8X,20HPROPORTION OF TRIALS,/,8X,20HTRUE AUTOREGRESSIVE,/,8
02930      1X,14HORDER IS FOUND,10X,F7.4,6(6X,F7.4),/)
02940       PRINT 2518,PMCI(1),PMCI(2),PMCI(3),PMCI(4),PMCI(5),PMCI(6),PMCI(7)
02950  2518 FORMAT (8X,20HPROPORTION OF TRIALS,/,8X,19HTRUE MEAN CONTAINED,/,8
02960      1X,22HIN CONFIDENCE INTERVAL,2X,F7.4,6(6X,F7.4),/)
02970       PRINT 2519,PCWC(1),PCWC(2),PCWC(3),PCWC(4),PCWC(5),PCWC(6),PCWC(7)
02980  2519 FORMAT (8X,20HPROPORTION OF TRIALS,/,8X,19HCONFIDENCE INTERVAL,/,8
02990      1X,22HWIDTH IS WITHIN LIMITS,2X,F7.4,6(6X,F7.4))
03000       STOP
03010       END
```

Figure 9. (Continued)

# APPENDIX D

## STATISTICAL VALIDITY

The purpose of this appendix is to substantiate the techniques used in performing the experiments described in Chapter III and IV. The experiments were based on two assumptions: 1) that the distributions of sample sizes were normally distributed; and 2) that sufficient trials had been run to adequately estimate a mean sample size for the particular data and sample size determination method in question. The calculations reported in this appendix show that the sample size distributions can be treated as being normally distributed. Also, sufficient trials were taken so that a confidence interval computed about the estimated mean sample size would have a width no more than fifteen percent of the estimated mean sample size.

Several distributions of sample sizes were selected for testing. Normal distributions were assumed and confidence intervals ($\alpha = .05$) were computed about the estimated mean of each distribution. The widths of the resulting intervals are reported in the fourth column of Table XIII. The maximum allowable width was chosen to be fifteen percent of the estimated mean sample size for each distribution. The maximum allowable widths are reported

in the fifth column of Table XIII. Notice that, for each distribution selected, the interval width is well within the desired limits.

The confidence intervals which were calculated were based on the assumption that the distributions of sample sizes were normal. For each selected distribution, a Lilliefors test was performed to check the goodness-of-fit for a normal distribution. The sixth column lists the test statistic, D, for each set of sample sizes. Column seven gives the .90 quantile for the Lilliefors test. The null hypothesis (i.e., the data fit a normal distribution) should be rejected if the test statistic, D, exceeds the .90 quantile. Based on the test statistics, the null hypothesis cannot be rejected for any of the selected distributions.

TABLE XIII

SELECTED DISTRIBUTIONS OF SAMPLE SIZE:
VALIDATION OF SAMPLE SIZE
AND DISTRIBUTION

| Distribution of Data | Autoregressive Order | $\beta^2$ | Computed Width Of Confidence Interval About Mean Sample Size | Maximum Allowable Interval Width | D | $W_{.90}$ |
|---|---|---|---|---|---|---|
| Normal | 1 | 1.00 | 29.56 | 34.95 | .0440 | .0805 |
| Normal | 1 | 1.00 | 23.05 | 31.8 | .0654 | .0805 |
| Exponential | 0 | 1.00 | 25.634 | 94.65 | .0357 | .0569 |
| Exponential | 0 | .95 | 32.356 | 72.75 | .0525 | .0569 |
| Exponential | 1 | .95 | 8.368 | 24.45 | .0469 | .0569 |
| Exponential | 2 | .95 | 27.698 | 75.3 | .0422 | .0569 |

# APPENDIX E

## RESULTS OF EXPERIMENTS USING HIGHER AUTOREGRESSIVE ORDERS

This supplementary material gives the results of experiments for higher autoregressive orders. The proposed method was tested on both exponential and geometric data for autoregressive orders three, five, and ten. The results shown in Tables XIV through XIX indicate that the procedure performs properly under the conditions of higher autoregressive orders.

TABLE XIV

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR EXPONENTIAL DATA WITH AUTOREGRESSIVE ORDER OF THREE

| Variable | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|
| | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 469 | 390 | 541 | 627 | 706 | 889 |
| Estimated Mean Of Data ($\bar{X}$) | 12.49 | 12.48 | 12.49 | 12.49 | 12.48 | 12.41 |
| True Mean Of Data ($\mu$) | 12.50 | 12.50 | 12.50 | 12.50 | 12.50 | 12.50 |
| Proportion of Trials ($1-\alpha$) In Which $\mu$ Is Contained In Confidence Interval About $\bar{X}$ | .970 | .955 | .965 | .980 | .955 | .935 |
| Proportion Of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | .665 | .710 | .845 | .940 | .985 | 1.00 |

TABLE XV

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR EXPONENTIAL DATA WITH AUTOREGRESSIVE ORDER OF FIVE

| Variable | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|
| | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 212 | 222 | 247 | 289 | 327 | 416 |
| Estimated Mean Of Data ($\overline{X}$) | 12.52 | 12.52 | 12.52 | 12.54 | 12.52 | 12.49 |
| True Mean Of Data ($\mu$) | 12.50 | 12.50 | 12.50 | 12.50 | 12.50 | 12.50 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\overline{X}$ | .970 | .970 | .975 | .970 | .975 | .975 |
| Proprtion Of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | .665 | .710 | .845 | .965 | .985 | 1.00 |

TABLE XVI

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR EXPONENTIAL DATA WITH AUTOREGRESSIVE ORDER OF TEN

| Variable | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|
| | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 211 | 221 | 246 | 288 | 326 | 414 |
| Estimated Mean Of Data ($\bar{X}$) | 12.52 | 12.52 | 12.52 | 12.54 | 12.52 | 12.49 |
| True Mean Of Data ($\mu$) | 12.50 | 12.50 | 12.50 | 12.50 | 12.50 | 12.50 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\bar{X}$ | .975 | .975 | .975 | .975 | .975 | .975 |
| Proportion Of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | .660 | .700 | .845 | .965 | .985 | 1.00 |

TABLE XVII

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR GEOMETRIC DATA WITH AUTOREGRESSIVE ORDER OF THREE

| Variable | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|
| | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 260 | 272 | 303 | 353 | 399 | 506 |
| Estimated Mean Of Data($\bar{X}$) | 27.48 | 27.48 | 27.46 | 27.53 | 27.54 | 27.58 |
| True Mean Of Data ($\mu$) | 27.50 | 27.50 | 27.50 | 27.50 | 27.50 | 27.50 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\bar{X}$ | .950 | .945 | .97 | .955 | .965 | .965 |
| Proportion Of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | .655 | .720 | .865 | .940 | .996 | 1.00 |

TABLE XVIII

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR GEOMETRIC DATA WITH AUTOREGRESSIVE ORDER OF FIVE

| Variable | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|
| | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 253 | 266 | 295 | 344 | 389 | 494 |
| Estimated Mean Of Data ($\bar{X}$) | 27.48 | 27.48 | 27.46 | 27.52 | 27.53 | 27.56 |
| True Mean of Data ($\mu$) | 27.5 | 27.5 | 27.5 | 27.5 | 27.5 | 27.5 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\bar{X}$ | .950 | .943 | .965 | .960 | .950 | .970 |
| Proportion of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | .635 | .705 | .835 | .93 | .99 | 1.00 |

TABLE XIX

RESULTS OF SAMPLE SIZE EXPERIMENTS FOR GEOMETRIC DATA WITH AUTOREGRESSIVE ORDER OF TEN

| Variable | The Proposed Method | | | | | |
|---|---|---|---|---|---|---|
| | $\beta^2=.65$ | $\beta^2=.70$ | $\beta^2=.80$ | $\beta^2=.90$ | $\beta^2=.95$ | $\beta^2=.99$ |
| Mean Required Sample Size | 252 | 265 | 294 | 343 | 388 | 492 |
| Estimated Mean of Data ($\overline{X}$) | 27.49 | 27.48 | 27.46 | 27.52 | 27.54 | 27.56 |
| True Mean Of Data ($\mu$) | 27.50 | 27.50 | 27.50 | 27.50 | 27.50 | 27.50 |
| Proportion of Trials ($1-\hat{\alpha}$) In Which $\mu$ Is Contained In Confidence Interval About $\overline{X}$ | .950 | .945 | .965 | .960 | .950 | .970 |
| Proportion Of Trials ($\hat{\beta}^2$) In Which Confidence Interval Width Is Within Desired Limits | .625 | .69 | .835 | .93 | .99 | 1.00 |

VITA[2]

Belva Jones Cooley

Candidate for the Degree of

Doctor of Philosophy

Thesis: DETERMINATION OF SAMPLE SIZE FOR SPECIFIED WIDTH CONFIDENCE INTERVALS IN DIGITAL SIMULATION EXPERIMENTS

Major Field: Business Administration

Biographical:

Personal Data: Born in Lake Charles, Louisiana, September 18, 1947, the daughter of Mr. and Mrs. Frank B. Jones

Education: Graduated from Westlake High School, Westlake, Louisiana, in May, 1965; received Bachelor of Science degree in Management from McNeese State University in 1969; received Master of Business Administration degree from Texas Tech State University in 1970; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in May, 1976.

Professional Experience: Management trainee, First National Bank, Lake Charles, Louisiana, 1968-69; Research Assistant, Texas Tech State University, 1969-70; Instructor, Stephan F. Austin State University, 1970-72; Graduate teaching assistant, Oklahoma State University, 1972-75; Assistant Professor, University of Nebraska at Omaha, 1975-76.