104
77

THE METHOD OF SEQUENTIAL SYSTEMATIC SAMPLING IN

DIGITAL SIMULATION

by

ChinFu Ho

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Industrial Engineering and Operations Research

APPROVED:

J. W. Schmidt, Chairman

H. D. Sherali

M. S. Jones

R. H. Myers

R. V. Foutz

August, 1986

Blacksburg, Virginia

THE METHOD OF SEQUENTIAL SYSTEMATIC SAMPLING IN

DIGITAL SIMULATION

by

ChinFu Ho

Committee Chairman: J. William Schmidt
Department of Industrial Engineering and Operations Research

(ABSTRACT)

This dissertation presents a methodology for the statistical analysis of simulation output data. The analysis deals with the predictability of statistical inferential procedures for means and variances when the data are realizations of correlated and nonnormally distributed random variables. The purpose of the methodology is to improve the predictability of an inferential procedure with respect to the level of confidence in confidence interval analysis, or the power function in hypothesis testing.

Conventional methods of statistical analysis for means lead to poor performance in their predictability if the sample observations are subject to strong autocorrelation. In addition, the predictability problem with respect to inferential procedures for variances is compounded by violation of the normality assumption.

The methodology presented in this dissertation sets forth a sampling procedure to collect sequences of essentially uncorrelated observations. With these observations at hand, the statistical formulation presented leads to an estimator of the variance of the sample mean, thus yielding inferential procedures for means through the classical techniques. The formulation also leads to an estimator of the variance of the population and inferential procedures for variances are developed with an improved property of robustness. The bias in each estimator is greatly reduced due to the sampling procedure employed. Finally the research includes an algorithm for testing the lag correlation such that the sampling procedure can be actually implemented.

The methods for means and variances developed in this research have been compared with corresponding conventional procedures. The comparison is based upon the predictability of the inferential procedure applied to the sample observations generated from autoregressive, simple moving average and M/M/1 queueing models. From the computational and simulation results reported in this research, the methods for means and variances suggested by this research have led to an improvement in the predictability of the analysis.

ACKNOWLEDGEMENTS

I offer my warm thanks to many generous people who contributed to the completion of this dissertation.

For three years in the course of research, Dr. J. W. Schmidt has provided step-by-step advice, insight to identifying and tackling problems, consistence of adhering to his academic standards, and tips in verbal and oral communications. Without his assistance this dissertation would not have been completed. The thanks to him is beyond description.

I am also particularly thankful for the help, guidance, and encouragement by the other members of the graduate committee, namely, Dr. Hanif D. Sherali, Dr. Robert V. Foutz, Dr. Raymond H. Myers, and Dr. Marilyn S. Jones. Their contribution to this dissertation is invaluable.

Special thanks should also be given to my parents who supported me morally and financially during the years of my graduate study. And, for for her work on typing, no praise is enough.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

TABLE

TABLE

LIST OF FIGURES

FIGURE

FIGURE

CHAPTER I

INTRODUCTION

1.1 THE SUBJECT AND PURPOSE OF THE RESEARCH

This research deals with the statistical analysis of simulation output data. Specifically, the research is focused on statistical inferential procedures for means and variances when the data available are correlated and may not be normally distributed.

The inferential procedures concerned in this research are confidence interval analysis and hypothesis tests. The predictability of the inferential procedures may be affected by data correlation, nonnormality, or both. The predictability of an inferential procedure refers to the precision with which the analyst may predict the level of confidence in confidence interval analysis or the power of a hypothesis test. The purpose of the research is to develop methods to reduce the deleterious impact of these phenomena on the predictability of the inferential procedures concerned.

1.2 IMPORTANCE OF THE STUDY

Digital simulation as a technique for modeling complex systems has become increasingly popular since the advent of the digital computer. The popularity of simulation may be explained by its advantages expressed in their relation to other alternatives to modeling. The most frequently employed alternative is mathematical modeling. The principal advantage of simulation over mathematical modeling lies in its simplicity. First, the system may be complicated enough to challenge a complete mathematical description while being amenable to representation by a simulation model. Second, for complex systems, the level of mathematical sophistication and training required for mathematical modeling are generally greater than that for simulation modeling. Thus, even if the system is amenable to mathematical analysis, the analyst may not possess the background necessary for the development of a mathematical model while he may be able to develop a valid simulation model. In summary, the relative advantages of simulation are versatility and simplicity.

A simulation model usually provides results which include estimates of the system parameters of interest. Typically these estimates are based upon output data from a single simulation run. Estimators such as the sample mean, variance, proportion, correlation coefficient, and median are commonly used. Since these estimators are random variables, if the results of a simulation study are to be meaningful, appropriate statistical inferential procedures must be used to analyze the output data. The most popular procedures usually assume that the sample data are independent and normally distributed. However, the results of a simulation experiment often yield realizations of random variables which are correlated and nonnormally distributed, leading to unpredictable performance of classical inferential procedures. This in turn may lead to a substantial increase or decrease in the likelihood that erroneous conclusions will be drawn.

The correlation of output data from a single simulation run is a well-known phenomenon. As stated by Law and Kelton [1982], "It is our belief that the output data from virtually all simulations are correlated". For example a major source of error arises in the estimation of the variance of the sample mean if the data are correlated. This error affects the precision in predicting the level of confidence for confidence interval analysis and the power of hypothesis tests if classical techniques are applied indiscriminately.

While nonnormally distributed sample data is also common, its impact on the predictability of classical procedures for means is usually less significant than in the case of correlated data. For example, if the sample size is reasonably large, violation of the assumption of normality alone has little effect on the accuracy of classical inferential procedures for means. However, for procedures for variances, violation of the assumption of normality may have considerable impact.

As a result of the inadequacy of classical statistics, extensive research has led to the development of parametric and nonparametric estimation and hypothesis testing procedures. However, most of the efforts have focused on inferential procedures for means. In the case of inferential procedures for means, several interval estimation techniques have been proposed. These include the method of replication, batch means, regeneration cycles, parametric modeling, spectral analysis, and standarized time series. Among these methods, batch means seems to be the most popular because it is relatively easy to use. All of these procedures rely on central limit theorems to describe the asymptotic behavior of the estimator. A detailed discussion of each of these methods will be presented in Chapter 2.

As for inferential procedures for variances, little research has been conducted for the case of correlated output. This may be due in part to the impact of nonnormality on the predictability of the inferential procedures for variance thus compounding the problem of dealing with correlated data. Law (1983) suggested a confidence interval procedure for variances, which applies an interval estimator used in the case for means to construct a confidence interval for the variance. However, this procedure only provides a lower bound for the confidence level being sought.

1.3 SCOPE AND LIMITATIONS OF THE RESEARCH

The following assumptions are made about the nature of realizations of sample observations in this research.

1. Observations in the process have finite first through fourth moments.
2. The process is strictly stationary, i.e. the entire probability structure of the process depends only on time differences (change of time origin has no effect on properties of the process). Therefore the covariance $Cov(X_i, X_{i+j})$ between two observations X_i and X_{i+j} depends only on the separation j and not on the initial ith point observed.

1.4 APPROACH OF THIS RESEARCH

The existing methods for means attempt to find an estimate for the variance of the sample mean or a substitute for such an estimate. The validty of these methods in terms of the resulting predictability depends on the asymptotic behavior of the estimate constructed. Based upon the empirical results reported, the predictability yielded by the existing methods are plagued by strong autocorrelation among the sample observations and small sample size. This phenomenon suggested the importance of estimating the variance of the sample mean for correlated data. In the case of variances, the predictability of any procedure depends in part on the error in estimating the variance of the population. Moreover, the distribution of a variance estimator is often complicated when the data are nonnormally distributed and/or correlated. Hence defining the sampling distribution for the variance estimator involved may be an arduous task.

Before applying or improving any of these methods or procedures, one needs to know the impact of the data correlation and nonnormality on the predictability of the inferential procedure considered. If the impact is considered insignificant, there would be no compelling reason to improve existing methods. Therefore, this research first establishes the significance of violation of the assumptions of normality and independence. The significance is measured by the magnitude of induced error in the probability statements for means and variances when the assumption/assumptions concerned is/are violated. Specifically, the error is represented by the difference between the desired and actual level of confidence in confidence interval analysis, or the difference between the desired and actual power of the test in hypothesis testing when the supporting assumption/assumptions is/are ignored. If the significance is established, the next step is to find a way to overcome the difficulties in estimating the variance of the sample mean and the variance of the population.

The correlation among sample observations is often a source of error in estimating the variance of the sample mean and the variance of the population. This research suggests a sampling procedure to collect sequences of uncorrelated or nearly uncorrelated observations from one realization of correlated data. By using these uncorrelated observations in a meaningful manner, the bias in estimating the variance of the sample mean and the variance of the population will be significantly reduced or eliminated.

The method of sampling relies on the identification of the point where the lag correlation of the process "dies out". If the lag correlation dies out at lag k, then the lag (k-1) is referred to as the order of serial dependence for the process. Hence observations sampled at intervals of k or greater may be considered uncorrelated. An algorithm for testing the lag correlation and then determining the order of serial dependence will be developed in this research.

The sampling method can be used to provide an improved technique in estimating the variance of the sample mean. By applying classical techniques, the estimator developed will lead to an improvement of the predictability of the inferential procedures concerned.

The method of batch means is the most widely used among the existing methods for means. The method of batch means is not only easy to use but also can improve the predictability of the statistical analysis when the data dependence is ignored. This research will compare the method for means developed in this research with batch means in improving the predictability of the analysis. The comparison is presented in the case of AR(1) and simple moving average models. To conclude the performance of the method for means developed in this research, it will be compared with other methods of means in the case of M/M/1 queueing models.

The sampling method developed in this research is also used to provide an estimator of the variance of the population in the case of correlated data. However, based upon this estimator, one can not directly apply classical techniques to construct inferential procedures for variances. This is because such construction requires the sampling distribution of the estimator, which is difficult to obtain in the case of correlated and nonnormally distributed data. Hence, based upon the variance estimator, associated inferential procedures for variances will be developed. This in turn leads to an improvement of the predictability of the procedure.

To illustrate the predictability of the inferential procedures for variances developed, they are applied to correlated data generated by a simple moving average process. The simple moving average process will first generate normal and then nonnormal data. For both normally and nonnormally distributed data, the classical inferential procedures for variances are applied for comparison. To summarize the performance of the methods for variances developed in this research, they will be compared with the classical procedures for variances with respect to the predictability of the inferential procedure applied in the case of M/M/1 queues.

To summarize, the methodology suggested in the approach of the research is delineated in Figure 1.

1.5 THE OBJECTIVES OF THIS RESEARCH

The objectives to be achieved in this research include the following:

1. Develop a sampling procedure to collect sequences of

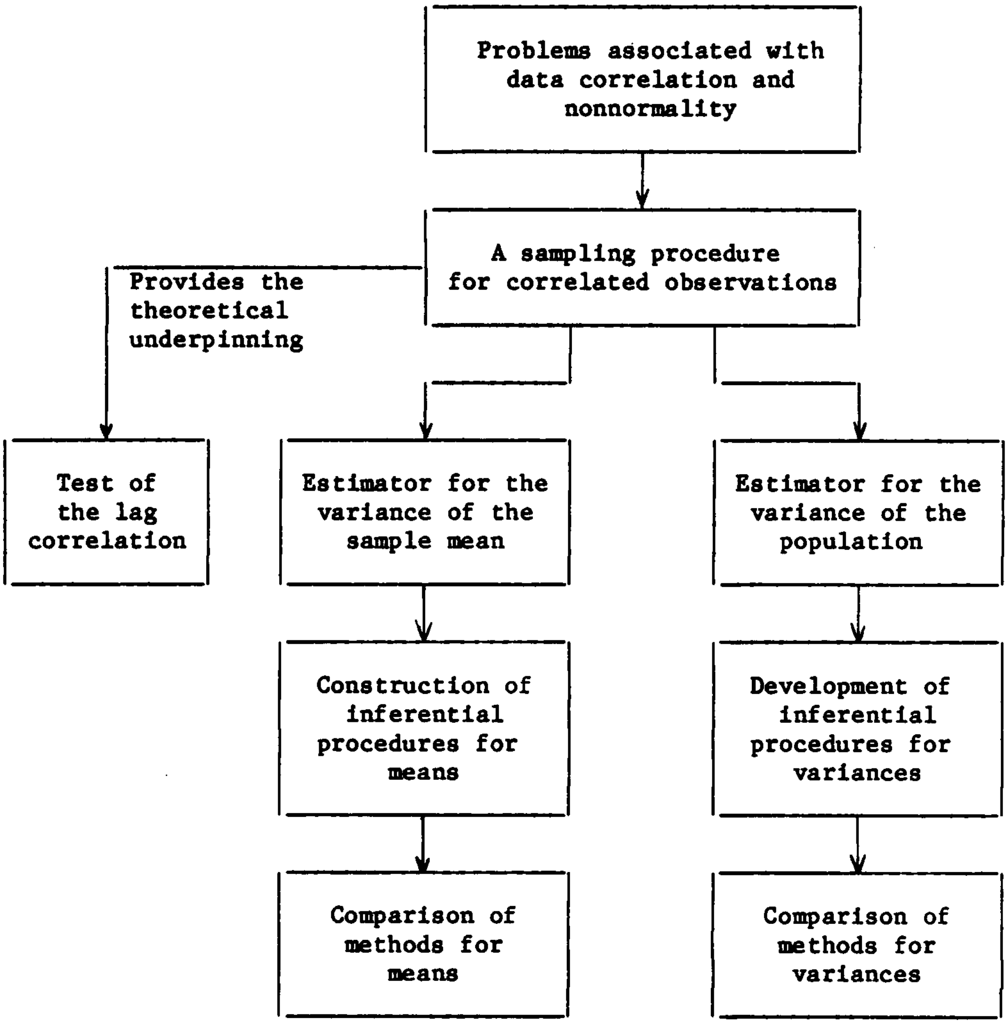

Figure 1. The Approach of Research

uncorrelated or nearly uncorrelated observations based upon an assumed order of serial dependence.

2. Develop an algorithm to determine the assumed order of serial dependence in the sampling procedure.
3. Develop an estimator for the variance of the sample mean based upon the sampling procedure in 1. Use this estimator to construct inferential procedures for means.
4. Compare the method for means developed in the research with other methods for means in the case of AR(1), simple moving average, and M/M/1 queueing models.
5. Develop an estimator of the population variance based upon the sampling procedure in 1. Then develop associated inferential procedures for variances.
6. Apply the inferential procedures for variances developed in this research to correlated data in cases where the data are normally and nonnormally distributed. In each case the classical inferential procedures for variances are applied for comparison.
7. Compare the methods for variances suggested in this research with the classical methods for variances in the case of simple moving average and M/M/1 queueing models.

1.6 RESEARCH CONTRIBUTIONS

The most important contributions made by this research are the following:

1. It provides a sampling procedure to collect observations which are essentially uncorrelated under the assumption of the order

of serial dependence.

2. Based upon the sampling method mentioned in 1, it provides an estimator of the variance of the sample mean. The estimator can be used to improve the predictability of inferential procedures for means.
3. It demonstrates the performance of the method for means developed in the research through the comparison with the method of batch means, the classical method. The comparison is presented in three cases: AR(1), simple moving average, and M/M/1 queueing models.
4. Based upon the sampling method developed in this research, it provides an estimator of the variance of the population. Employing this estimator it provides inferential procedures for variances.
5. It empirically tests the performance of the method for variances developed in this research by comparing the classical inferential procedures for variances in two cases: simple moving average and M/M/1 queueing models.
6. It provides an algorithm for the determination of the order of serial dependence. Through implementation of the algorithm, sequences of observations which are considered as uncorrelated can be obtained from a sequence of observations.

1.7 ORGANIZATION OF THE DISSERTATION

In presenting the results of this research, the following order of presentation is employed. Chapter II treats past research related to this research in the three areas of inferential procedures for means, variances, and determining the order of serial dependence. Chapter III treats measurement of potential significance of the problems associated with violation of the assumptions of independence and normality. The impact of violation of each of the two assumptions on the predictability of the inferential procedure concerned is discussed.

Chapter IV presents the sampling procedure mentioned in Section (1.4) as an improved technique for estimating both the variance of the population and the variance of the sample mean. Chapter V treats the comparison of the method for means developed in Chapter IV, the method of batch means, and the classical method where the data correlation is ignored. The comparison is presented in two cases where sample observations are generated from AR(1) and simple moving average models respectively.

Chapter VI presents parametric inferential procedures for variances based upon the variance estimator developed in Chapter IV. Sample observations are assumed to be correlated and normally distributed in this chapter. To evaluate the performance of the inferential procedures developed, they are applied to correlated and normally distributed data in the form of simple moving average model to predict the level of confidence and the power of hypothesis tests. Chapter VII presents revision of the inferential procedures developed in Chapter VI when the data are correlated and nonnormally distributed. To illustrate the performance of the revised procedures, they are applied to the sample observations which are correlated and nonnormally distributed and generated in the form of simple moving average model to predict the level of confidence and the power of hypothesis tests. In both Chapters VII and VIII, the suggested inferential procedures for variances are compared with the classical inferential procedures for variances.

Chapter VIII presents the algorithm for determining the order of serial dependence and thus provides the justification for the sampling procedure developed in Chapter IV. The sample observations at lag greater than this order can be considered uncorrelated. To demonstrate the performance of the algorithm, it is applied to the waiting time observations of M/M/1 queueing models to determine the order of serial dependence.

Chapter IX presents a comparison of the methods for means and variances developed in this research with existing procedures. The comparison is presented in the case of M/M/1 queues.

Chapter X summarizes the important features of the research, presents conclusions arising from the research and recommends areas for further research.

CHAPTER II

LITERATURE REVIEW

2.1 GENERAL INTRODUCTION

Three topics are covered in this literature review and include inferential procedures for means, inferential procedures for variances, and determination of the order of serial dependence.

One important problem often encountered in a real-world simulation study is the construction of a confidence interval for the true mean μ of a stochastic process. Information continaed in such a confidence interval provides the decision maker with a measure of how precisely μ is known. However, constructing the confidence interval is difficult because one realization of random variables from a simulation usually leads to a sequence of correlated and nonnormally distributed data. In the presence of data correlation classical statistical procedures should be applied with caution.

The purpose here is to review several well-known methods for means and discussion their shortcomings. Empirical studies of each method will also be discussed.

In addition to constructing inferential procedures for means, another important problem in real-world simulation study is that of constructing a confidence interval for the population variance σ^2 of one or several stochastic processes. Information contained in such a confidence interval provides the decision maker with a measure of how precisely σ^2 is known. However, few procedures for variances are reported in the literature. This phenomenon may arise from the following two considerations:

1. The sampling distribution of the test statistic considered is not known because the sample data are correlated.
2. The test procedures for variances are known to be sensitive to the departure from normality in the case where sample observations are independent. Similar conclusion is anticipated in the case where sample observations are correlated.

The purpose of literature review on inferential procedures for variances is to review the available procedures for variances.

As presented in Chapter I, a problem which is related to the statistical analysis of correlated data is that of determining the order of serial dependence for a realization of correlated random variables. Assuming that the serial correlation decreases as the separation between two observations increases, the serial correlation can be deemed to have died out at a certain point of separation. The order of serial dependence can be determined from the value of the separation. The process of constructing an inferential procedure for means or variances is assisted if information about the order of serial dependence is provided. The purpose of literature review on determination of the order of serial dependence is to review methods that can be used to determine the order of serial dependence.

2.2 INFERENTIAL PROCEDURES FOR MEANS

2.2.1 Introduction

Let $Y_1, Y_2, \ldots$ be an output process obtained from a simulation run. Given the assumption of strict stationarity for the process $\{Y_t,\ t > 1\}$, let μ be the common mean of the process, σ^2 the common variance, and Σ the covariance matrix given by

$$\Sigma = \begin{bmatrix} \sigma^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma 12 & \sigma^2 & \cdots & \sigma_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma^2 \end{bmatrix}.$$

The variance of the sample mean ($\bar{Y}$) is then given by

$$\sigma^2_{\bar{Y}} = \frac{1}{n} [\sigma^2 + \frac{2}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{ij}].$$

Simulation practitioners are often interested in seeking a 100(1-α)% confidence interval for a single mean or the difference of two means or in testing hypotheses about one or two means.

By way of example, if a 100(1-α)% confidence interval for μ is to be constructed, the confidence limits suggested in the past research are often presented in the following form:

$$\bar{Y} \pm t_{f,1-\alpha/2}\hat{\sigma}_{\bar{Y}}$$

where $\hat{\sigma}^2_{\bar{Y}}$ is an estimate of $\sigma^2_{\bar{Y}}$, f is an appropriate degrees of freedom, and $t_{f,1-\alpha/2}$ is the upper 1-α/2 critical value for a t distribution with f degrees of freedom.

In the presence of data correlation, the sample mean $\bar{Y}$ is still an unbiased estimate of μ. Thus precision of the confidence interval for μ would depend on the estimate of $\sigma^2_{\bar{Y}}$, $\hat{\sigma}^2_{\bar{Y}}$, since the half length of the confidence interval is given by ($t_{f,1-\alpha/2}\hat{\sigma}_{\bar{Y}}$). However, the data correlation leads to difficulties in estimating $\sigma^2_{\bar{Y}}$.

On the other hand, the nonnormality of the process $\{Y_t\}$ may not yeild significant impact as in the case of data correlation. The following statements from Box (1953) illustrate some history about the robustness of the inferential procedures for means in the case of independent data. He said:

> "thanks to the work of ...., Gayen (1950a,b), David and Johnson (1951a,b), there is abundent evidence that these comparative tests on means are remarkably insensitive to general nonnormality of the parent population."

Srivastava (1958) studied the effects of nonnormality on the power function of the single sample t-test. For samples of size 10 and 5% significance level his results indicate only moderate errors in estimating the true power of the test by that which would be predicted under the assumption of normality.

While robustness of procedures for means is confirmed in the case of independent data, empirical studies (Law (1977)) of the past research also suggested similar conclusion for the correlated output. This point will be presented in the discussion of methods for means suggested in the past research.

Six methods for means have been suggested in the literature. Law [1983] categorized these six methods into the following four basic types:

1. Those that seek independent observations: replication and batch means.
2. Those that seek to estimate correlation in output variables: parametric modeling and spectral analysis.

3. Those that exploit special structure of the underlying process: regeneration cycles.
4. Those that are based upon standardized time series.

Methods of the first type (replication and batch means) attempt to group simulation observations into batches which yield uncorrelated or nearly uncorrelated and identically distributed random variables with mean μ and variances σ^2. If this can be achieved then classical statistical procedures can be applied. Methods of the second type (parametric modeling and spectral analysis) do not try to obtain uncorrelated observations, but recognize the correlation structure of the process and use correlation estimates to form a confidence interval. Methods of the third type (regeneration cycles) also seek independent observations but exploit a particular known probabilistic structure of the process to form a confidence interval. Methods of the fourth type (standardized time series) use the weak convergence of functions of the process (see Billingsley (1968, Section 21)) to provide a general framework for developing a confidence interval.

The performance of each method in the confidence interval analysis may be extended to the case of hypothesis testing. However, empirical results for those methods are often reported for confidence interval analysis. Typically these method are applied to an M/M/1 queueing system to form a confidence interval for the mean waiting time or a system characteristic of interest. Usually the criterion on the performance of a method is based upon the predictability of the confidence interval procedure applied. The predictability is measured by the discrepency between the actual and the desired coverage with respect to the confidence interval constructed.

2.2.2 Procedures That Seek Independent Observations

The first of such procedures is the method of replication. This method employs separate replications with each replication being called a batch. One attempts to obtain the sample mean from each batch and treat each sample mean as an independent observation to estimate the variance of the sample mean.

Suppose the analyst chooses m independent replications of length k observations each, and uses independent random number streams for each replication. If $\bar{Y}_j$ is the sample mean of k observations in the jth replication, then $\bar{Y}_j$'s are independent and identically distributed (i.i.d.) random variables with mean μ. As a point estimtae of μ one uses

$$\bar{\bar{Y}} = \frac{1}{m} \sum_{j=1}^{m} \bar{Y}_j$$

and estimates $\mathrm{Var}(\bar{Y})$ by

$$S^2_{\bar{\bar{Y}}} = \frac{1}{m(m-1)} \sum_{j=1}^{m} (\bar{Y}_j - \bar{\bar{Y}})^2.$$

A $100(1-\alpha)\%$ confidence interval for μ is given by

$$\bar{\bar{Y}} \pm t_{m-1,1-\alpha/2} S_{\bar{\bar{Y}}}$$

where $t_{m-1,1-\alpha/2}$ is the upper $1-\alpha/2$ critical value for a t distribution with m-1 degrees of freedom. While the method of replication is easy to apply, two disadvantages may be identified:

1. Only one observation is obtained for a single run. Hence the cost of computer time may be quite high if a large number of observations are to be collected.

2. Some information about the output variables would be lost. For example, the underlying correlation and the population variances cannot be estimated.

An alternative to the method of replication is the method of batch means. The method of batch means is perhaps the most widely used method where inferential methods for means must be applied in the presence of correlated data.

As with replication, the method of batch means attempts to take independent observations from the process, but only a single simulation run of length n is made. The analyst divides the run into m batches of k consective observations each (m=mk). Let $\bar{Y}_j$ to be the sample mean of the k observations in the jth batch. $\bar{Y}_j$ is called the jth batch mean. $\bar{\bar{Y}}$ and $S^2_{\bar{Y}}$ are then defined as in the method of replication.

The method of batch means might be motivated by the following theoretical considerations. Assume $R_s = \mathrm{Cov}(Y_t, Y_{t+s})$ being the lag s covariance of the original process and $R_s(k) = \mathrm{Cov}(\bar{Y}_j, \bar{Y}_{j+s})$ being the lag s covariance among the batches. Then it can shown (see Law and Carson (1979)), if $0 < \sum_{s=-\infty}^{\infty} R_s < \infty$, then the correlation

$$\frac{R_s(k)}{R_0(k)} \to 0 \text{ as } k \to \infty \text{ for } s = 1,2,\ldots,m-1.$$

Thus, the absolute magnitude of correlation among batch means decreases asymptotically. Schmeiser and Kang (1981) have shown that the batch means of ARMA(1,1) processes under the assumption of normality are also ARMA(1,1). As a result, if the process $\{Y_t\}$ is MA(1), the process $\{\bar{Y}_j\}$ is also MA(1); if $\{Y_t\}$ is AR(1), then $\{\bar{Y}_j\}$ is ARMA(1,1). Therefore the correlation between batch means may be an important source of error if the batch means are used to form a confidence interval for μ. This viewpoint is supported by empirical results in Law and Kelton (1979) for M/M/1 queueing models, where the correlation between the $\bar{Y}_j$'s is the most serious source of error. For example, for the M/M/1 queue with the traffic density (ρ) of .8, the actual coverage of the confidence interval produced by this method is 70% for $m = 20$ and $k = 64$ while the desired is 90%.

Due to the problems associated with the correlation between batch means, procedures which seek to find a suitable batch size k are needed. One approach has been suggested (Gross and Harris (1974, pp. 426)) is to fix the number of batcvhes (m) and then increases the batch size k until the estimated correlation of adjacent batches means is less than a small number (0.05 for example). The difficulty with this approach, as quoted from Law (1983), is "correlation estimators are generally biased and for small n are highly variable". Fishman (1975) developed a method to determine the batch size that relies upon the von Neumann ratio. The method is applied in the case of M/M/1 queue. Fishman concluded that his method performs well with large sample sizes if the process is not too positively autocorrelated. Two examples of his results are cited here and in each example the desired leve of confidence is 95%. Based upon 60 replications for each sample size, for the M/M/1 queue with $\rho = .8$, the actual coverage of the confidence interval produced by method of batch means is 84.4% for $n = 8192$, where the batch size k is determined by Fishman's method in each replication. For the M/M/1 queue with $\rho = .9$, the actual coverage is 83.3% for $n = 8192$.

Law (1977) conducted a comparison of replication and batch means. He concluded that batch means is empirically superior to replication, but neither method works well if the sample size is too small. Two examples of his results in the case of M/M/1 queue with $\rho = .9$ are given here. For $m = 10$ and $k = 64$, Law reported that the actual coverage of the confidence interval is 62% for replication and 76.75% for batch means. For $m = 20$ and $k = 320$, the actual coverage is 27% for replication and 72% for batch means. In each example the desired coverage is 90%.

Based upon the empirical study of the M/M/1 queue with $\rho = .9$, Law (1977) reported that if the number of batch exceeds 20 violation of the normality assumption has no major impact on the confidence interval procedures on the basis of batch means and replication.

2.2.3 Procedures That Seek To Estimate The Correlation Structure

The first of such procedures is the method of parametric modeling. This method employs estimates of the correlation structure to obtain an estimate of the variance of the sample mean and was first developed by Fishman [1971,1973,1978]. Assume that the observations $Y_1, Y_2, Y_3, \ldots, Y_n$ from a single replication of the simulation with mean μ can be represented by an AR(p) model,

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + a_t$$

where $\{a_t\}$ is a sequence of uncorrelated random variables with mean zero and variance σ_a^2.

Let $\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$ be the estimator for μ. Let R_s represent the lag s covariance as previously defined. The variance of $\bar{Y}$ is given by

$$\mathrm{Var}(\bar{Y}) = \frac{1}{n}\,[\sigma^2 + \frac{2}{n}\sum_{s=1}^{n-1} (n-s)R_s]. \tag{2.2.3.1}$$

It can be shown (Pritsker and Pegden (1979), pp. 472)

$$\lim_{n\to\infty} n\,\mathrm{Var}(\bar{Y}) = \sum_{h=-\infty}^{\infty} R_h = m.$$

Given that $\sum_{h=-\infty}^{\infty} |R_h| < \infty$, it can also be shown

$$m = \frac{\sigma_a^2}{(1 - \sum_{j=1}^{p} \phi_j)^2}.$$

Based upon estimating the covariances R_s from the observations $Y_1, Y_2, \ldots, Y_n$, Fishman (1973) gives a procedure for determining the order p and obtaining estimates $\hat{\phi}_j$ $(j = 1,2,\ldots,\hat{p})$ and $\hat{\sigma}_a^2$, where $\hat{p}$ is the estimated order. For large n, an estimate of $\mathrm{Var}(\bar{Y})$ and an approximate $100(1-\alpha)\%$ confidence interval for μ are given by

$$\mathrm{Var}(\bar{Y}) = \frac{\hat{\sigma}_a^2}{n(1 - \sum_{j=1}^{\hat{p}} \phi_j)^2}$$

$$\bar{Y} \pm t_{\hat{f},1-\alpha/2}\sqrt{\mathrm{Var}(\bar{Y})}.$$

The expression for the estimated degrees of freedom, $\hat{f}$, is given in Fishman [1978]. The method of parametric modeling was empirically tested by Law and Kelton (1979). Their results indicated that the actual coverage of the confidence interval produced may yield coverage below that desired, if the sample size chosen is small. In comparison with the method of batch means for the case of M/M/1 queue with $\rho = .8$, Law and Kelton reported that the actual coverage of the confidence interval formed is 75% for n = 1280 while the desired is 90%.

Andrew and Schriber [1982] generalize Fishman's model by assuming that $Y_1, Y_2, Y_3, \ldots, Y_n$ can be represented by an ARMA(p,q) model. They used the G-K-M method (a model identification technique developed by Gray, Kelley and McIntire [1978] to determine the order p and q) to identify the process under investigation. Empirical results from two queueing systems indicated the coverage of the confidence interval is less than satisfactory and also not consistent with increasing sample size in the case of M/M/1 queues although the maximum sample size was only 400. For that sample size in the M/M/1 queue with $\rho = .8$, the actual coverage of the confidence intervals formed is 79% while the desired is 95%.

There are two disadvantages associated with parametric modeling:

1. The model identification procedure developed by Box and Jenkins [1976] is not very comprehensive - it is most effective for the AR(p) or MA(q) models. The G-K-M method is more general; it covers all ARMA(p,q) models, based on so called R- and S- arrays. However, because of the complexity of the R- and S- arrays, the recognition of the orders of the R- and S- arrays may be difficult when the sample covariance estimates are employed to obtain those arrays.
2. The estimator of $\mathrm{Var}(\bar{Y})$ is obtained through an asymptotic procedure and may not yield reliable estimates if the sample size is not large. As quoted from Pritsker and Pegden (1979), "In our experience, parametric modeling of the time series obtained from a simulation model has not produced reliable estimates of the variance of the sample mean".

Like parametric modeling, the method of spectral analysis also employs estimates of the correlation structure to obtain an estimate of the variance of the sample mean. The method of spectral analysis computes $\bar{Y}$ from the process $Y_1, Y_2, Y_3, \ldots, Y_n$. Given the lag h covariance, R_h, and the assumption, $\sum_{h=-\infty}^{\infty} |R_h|^2 < \infty$, the spectral density function $f(\lambda)$ is defined as

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} R_h e^{-i\lambda h}, \; -\pi < \lambda < \pi.$$

Then, as in the case of parametric modeling,

$$m = \sum_{h=-\infty}^{\infty} R_h = 2\pi f(0).$$

Therefore, an estimator of $\text{Var}(\bar{Y})$ is given by

$$\text{Var}(\bar{Y}) = \frac{\hat{m}}{n} .$$

In other words, estimating the spectral density function at zero frequency is equivalent to estimating $\text{Var}(\bar{Y})$. Estimating the spectral density function requries the determination of the number of covariances to be included in the computation, and the weighting function (lag window) to be applied to the estimated covariance obtained from finite observations. There are several types of weighting functions. These are Barlett, Tukey, Parzen, rectangular and "variance" (see Duket and Pritsker (1978)).

An estimator for $\text{Var}(\bar{Y})$ that is immediately available can be obtained by replacing R_s in equation (2.2.3.1) with an estimate C_s. Since C_s is quite unreliable for s near n, Fishman (1969) has suggested the following estimator

$$\mathrm{Var}(\bar{Y}) = \frac{1}{n} [C_o + \frac{2}{n} \sum_{s=1}^{q-1} w_q(s) C_s] \tag{2.2.3.2}$$

where $q > s$, and $w_q(s)$ is known as the weighting function. An approximate $100(1-\alpha)\%$ confidence interval for μ is given by

$$\bar{Y} \pm t_{f,1-\alpha/2} \sqrt{\mathrm{Var}(\bar{Y})}$$

where f depends on n, q, and the choice of weighting function (see Fishman (1969,1973)).

Duket and Pritsker (1978) applied the lag windows mentioned above to obtain estimates of the spectral density function at zero frequency in a simulation study of the M/M/1 queue with $\rho = .8$. They found that all estimates of the variance of the sample mean have extremely high variability. They concluded that due to such variability these estimates are not useful for statistical statements concerning the sample mean of the system. Law and Kelton (1979) also empirically tested the method of spectral analysis using the estimator of $\mathrm{Var}(\bar{Y})$ in equation (2.2.3.2) with the Tukey window. They found that the actual coverage of the confidence interval was smaller than that desired if the sample size chosen is small. For example, in comparison with the result of the method of batch means reported in the last section, for the M/M/1 queue with $\rho = .8$, the actual coverage is 77% for $n = 1280$ while the desired is 90%.

The method of spectral analysis is a fairly complicated technique, requiring a fairly sophiscated statistical background on the part of the analyst. It is also expensive to apply due to the large number of covariance estimates which must be computed. Moreover, there are no definitive rules in which the analyst can choose the number of covariances to be included in the computation.

2.2.4 Procedures That Exploits The Special Structure Of The Underlying Process

The method of regeneration cycles takes an altogether different approach in estimating the variance of the sample mean. This method divides a simulation run into independent cycles by defining states where the model starts anew. The estimate of the variance of the sample mean is based on observed values in the independent cycles. The method is discussed by Crane and Iglehart (1974,a,g;1975) and Fishman (1973b,1974).

For the output process $\{Y_i,\ i \geq 1\}$, assume that there exists a sequence of random times $1 \leq B_1 < B_2 < \ldots$, called regeneration points or regeneration times. At each of these times the process starts over with the same probabilistic structure that dictates its behavior at time B_1. Hence, the portion of the process between any two consecutive regeneration times, $\{Y_i,\ B_n \leq i < B_{n+1}\}$, is an independent and identically distributed replica. The portion of the process between two successive B_j's is called a cycle. For $N_j = B_{j+1} - B_j$, it is assumed that $E(N_j) < \infty$. If $Z_j = \sum_{i=B_j}^{B_{j+1}-1} Y_i$, the random vectors $U_j = (Z_j,N_j)'$ are i.i.d.. The mean of the output process μ can be expressed as $\mu = E(Z)/E(N)$.

Suppose a simulation of the process $\{Y_i,\ i \geq 1\}$ for n' cycles leads to the following data:

$$Z_1,Z_2,\ldots,Z_{n'}.$$

$$N_1,N_2,\ldots,N_{n'}.$$

Each of the two sequences comprises i.i.d. observations but Z_j and N_j are not independent in general.

A point estimator for μ can be given by $\hat{\mu} = \bar{Z}(n')/\bar{N}(n')$, where n' is the number of cycles. $\bar{Z}(n')$ and $\bar{N}(n')$ are unbiased estimators for E(Z) and E(N) respectively, but $\hat{\mu}$ is not an unbiased estimator for μ. An alternative is to obtain the sequence of i.i.d. observations $V_j = Z_j - \mu N_j$. Thus an approximate 100(1-α)% confidence interval for μ can be constructed by applying the classical central limit theorem to this sequence (Iglehart (1975)). The coverage of the confidence interval will approach 1-α as the number of cycles n' approaches infinity.

The significant advantage of the method of regeneration cycles is that it provides i.i.d. observations. However, the assumption of regeneration points requires the simulation practitioner to show that his model does have this probabilistic structure, which may be difficult or impossible for complex systems often encountered in real-world simulation studies. For example, as pointed out by Law and Carson (1979), "a queueing system with two or more non-Poisson arrival streams" does not have regeneration points. Even if the regeneration points can be identified, it might happen that $E(N_j)$ may be so large that only a few cycles can be simulated. This may lead to the actual coverage of the confidence interval constructed smaller than the desired coverage. Law and Kelton (1979) showed that if the number of cycles is small, then the resulting coverage of the confidence interval constructed may be disappointing. For example, in the M/M/1 queue with $\rho = .8$ and with the same sample size as in the case of batch means

($n = 1280$), the actual coverage of the confidence interval is 70% against the desired 90%.

2.2.5 A Procedure Based Upon Standardized Time Series

This method refers to a methodology which uses the weak convergence of functions of stochastic processes (see Billingsley (1969)) to construct procedures for means. General references for this method are Goldsman and Schruben (1982), Schruben (1982,1983), and Schruben et al. (1983).

The method of standardized time series assumes that the process $\{Y_i,\ i \geq 1\}$ is strictly stationary with mean μ for all i and also phi-mixing. Phi-mixing refers to a process for which Y_i and Y_{i+j} become essentially independent as j becomes large (see Billingsley (1969) for definition). This method divides the observations $Y_1, Y_2, \ldots, Y_n$ into m batches of size k where $n = mk$. Let $\bar{Y}_i$ be the sample mean of the ith batch, and $\bar{Y}_{i,j}$ be the sample mean of the first j observations of the ith batch and given by

$$\bar{Y}_{i,j} = \left(\frac{1}{j}\right) \sum_{p=1}^{j} Y_{p+(i-1)k} \qquad \text{where } i = 1,2,\ldots,m \quad j = 1,2,\ldots,k.$$

If n is large, the grand sample mean $\bar{\bar{Y}}$ willbe approximately normally distributed,

$$\bar{\bar{Y}} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

where $\sigma^2 = \lim_{n \to \infty} n(\mathrm{Var}(\bar{\bar{Y}}))$.

The term standardized time series for this method refers to the following sequences:

$$T_i(j/k) = \frac{\bar{Y}_i - \bar{Y}_{i,j}}{\frac{\sqrt{k}\sigma}{j}}$$

where $i = 1,2,\ldots,n$, and $j = 1,2,\ldots,k$. Define the statistic H as

$$H = \frac{12}{k^3-k} \sum_{i=1}^{m} [\sigma\sqrt{k} \sum_{j=1}^{k} T_i(j/k)]^2.$$

It can be shown that

$$H \sim \sigma^2 \chi_m^2$$

where χ_m^2 represents the chi-square statistic with m degrees of freedom. H is also asymptotically independent of $\bar{\bar{Y}}$. Therefore

$$\frac{(\bar{\bar{Y}} - \mu)/\sqrt{\sigma^2/n}}{\sqrt{H/m\sigma^2}} = \frac{\bar{\bar{Y}} - \mu}{\sqrt{H/mn}} \sim t(m)$$

where t(m) represents the t statistic with m degrees of freedom. An approximate $100(1-\alpha)\%$ confidence interval for μ is given by

$$\bar{\bar{Y}} \pm t_{m,1-\alpha/2}\sqrt{H/mn}.$$

Empirical results repoted in Schruben [1983] indicated that, for the case of M/M/1 queues, the coverage of the confidence interval performs well if k is large. However, like the method of batch means, the method of standardized time series relies on a large sample size for good results. In the M/M/1 queue with $\rho = .8$, with the sample size 60,000 ($m = 20$, $k = 3{,}000$) the actual coverage of the confidence interval is 85% while the desired coerage is 90%. In the M/M/1 queue with $\rho = .2$, for the sample size of 20,000 ($m = 20$, $k = 1{,}000$), the actual coverage is 93% while the desired is 90%. Additional experiments are needed to investigate the performance of the method for small values of k.

2.2.6 Conclusion And Impact On The Research

For the purpose of reference, empirical results on M/M/1 queueing systems, reported by Law (1977,1979), Fishman (1975) and Schruben (1983), are given in Table 1, 2, 3, and 4. As these tables illustrate, the performance of all the methods for the M/M/1 queueing systems with respect to the predictability of the confidence interval procedure applied is quite disappointing. The disappointing performance may arise in either or both of the following two situations:

1. The sample size is small.
2. The process is very much positively autocorrelated.

Based upon the empirical results reported, one conclusion can be drawn to summarize available confidence interval procedures for means. In the presence of positive correlation among data, a small sample size may yield the actual coverages that are significantly lower than desired since these procedures are only asymptotically correct. This research will develop a method (in Chapter IV) which will yield less discrepency between the desired and actual coverage than in the case of batch means if the data are positively correlated. The method developed will be compared with other methods for means for the M/M/1 queueing systems.

2.3 INFERENTIAL PROCEDURES FOR VARIANCES

2.3.1 Introduction

Assume the stochastic process $\{Y_t, t > 1\}$ with mean μ being simulated has variance

Table 1

Confidence Level in Law and Kelton (1979) for the M/M/1 Queueing Model with $\rho = .8$, $R = 400$ Replications, $\alpha = 10\%$.

	Batch Means				Parametric	Spectral Analysis				Regeneration
	m				Modeling	d.f. + 1				Cycles
n	5	10	20	40		5	10	20	40	
320	.6900	.5975	.4900	.3650	.6875	.7125	.6250	.5375	.4175	.6875
640	.7225	.7050	.5875	.4625	.7225	.7600	.7350	.6450	.5275	.7225
1280	.7800	.7400	.7050	.5850	.7525	.7825	.7700	.7450	.6525	.7525
2560	.7975	.8025	.7525	.6700	.7550	.8325	.8075	.7725	.7225	.7550

Table 2

Confidence Level in Law (1977) for the M/M/1 Queueing Model
with $\rho = .9$, R = 400 Replications, $\alpha = 10\%$.

	Replication				Batch Means			
n \ m	5	10	20	40	5	10	20	40
1600	0.5275	0.1825	0	0	0.6325	0.5750	0.4750	0.3650
3200	0.6700	0.4175	0.0450	0	0.7675	0.6825	0.6050	0.5225
6400	0.7525	0.6200	0.2700	0	0.8050	0.7675	0.7200	0.6125
12800	0.8325	0.7375	0.5275	0.0900	0.8650	0.8125	0.7975	0.7400

Table 3

Confidence Level in Fishman (1975) for M/M/1 Queueing Models with R = 60 Replications, α = 5%.

ρ \ n	Batch Means 2048	4096	8192	16384
0.5	0.934	0.967	0.900	0.950
0.8	0.783	0.884	0.884	0.932
0.9	0.734	0.867	0.833	0.900

Table 4

Confidence Level in Schruben (1983) for M/M/1 Queueing Models with R = 100 Replications, α = 10%.

ρ	k	m	Standardized Time Series
.2	20000	1	0.86
	10000	2	0.91
	4000	5	0.94
	2000	10	0.93
	1000	20	0.93
.5	40000	1	0.86
	20000	2	0.90
	8000	5	0.86
	4000	10	0.86
	2000	20	0.86
.8	60000	1	0.92
	30000	2	0.90
	12000	5	0.86
	6000	10	0.84
	3000	20	0.85

$$\sigma^2 = E(Y_t - \mu)^2.$$

Simulation practitioners may interested in seeking a 100(1-α)% confidence interval for σ^2 and conducting hypothesis tests for σ^2.

Unlike inferential procedures for means, inferential procedures for variances are quite sensitive to departures from normality. Box (1953) points out "this test is particularly sensitive to changes in γ_2 (coefficinet of kurtosis)," and "it is shown the sensitivity is even greater when the number of variances to be compared exceeds two."

In addition to nonnormality, inferential procedures for variances considered in this research are further influenced by the correlation of sample observations. Typically violation of the normality assumption alone does not cause bias in estimating σ^2, if the variance estimate of σ^2 is given by

$$s^2 = \sum_{i=1}^{n} (Y_i - \bar{Y})^2/(n-1).$$

However, correlated data often leads to bias in estimating σ^2. Also the sampling distribution of the variance estimator s^2 in the classical statistics may not be appropriate for either nonnormal or correlated data. Furthermore, developing the sampling distribution of the variance estimator is often an arduous task in the presence of data correlation or nonnormality. Thus classical inferential procedures for variances which are based upon the assumptions of independence and normality are not suitable for correlated and nonnormally distributed data.

Beside classical inferential procedures for variances, few inferential procedures for variances are reported in the literature.

This may be due in part to those characteristics such as nonnormality and data correlation, and in part to the interest of the academic community. Law (1983) suggested a confidence interval procedure for variances, which will be discussed in the next section. Also there is an indirect test procedure which can be used for testing variances and it will be discussed later in this chapter.

The literature view in this section will include the following:

1. The confidence interval procedure suggested by Law (1983).
2. Indirect testing procedure for variances.

2.3.2 Law's Confidence Interval Procedure For Variances

Law (1983) obtained a lower bound of the confidence level in his procedure through a simple algebraic formulation of an interval estimator for means. For an output process $\{Y_i,\ i \geq 1\}$ with mean μ, the variance of Y_i is given by

$$\mathrm{Var}(Y) = E(Y^2) - \mu^2.$$

In applying the procedures for means to the sample observations $Y_1^2, \ldots, Y_n^2$, a 100(1−α)% confidence interval may be constructed for $E(Y^2)$,

$$\Pr\{\underline{\theta} < E(Y^2) < \bar{\theta}\} = 1 - \alpha$$

where $\underline{\theta}$ and $\bar{\theta}$ are lower and upper confidence limits.

Let $\underline{\mu}$ and $\bar{\mu}$ be lower and upper confidence limits for the parameter μ for a 100(1−α)% confidence interval. Crane and Iglehart (1974) derived the following confidence interval for Var(Y)

$$\Pr\{\underline{\theta} - (\bar{\mu})^2 < \mathrm{Var}(Y) < \bar{\theta} - (\underline{\mu})^2\} \geq 1 - \alpha.$$

One shortcoming about this method is that the method can only obtain the lower bound of the confidence level being desired.

2.3.3 Indirect Testing Procedure For Variances

If an output process can be represented by an ARMA model, an approach to the inferential procedures for variances is to obtain the variance estimate through an identified ARMA model for the output process. Since the variance estimate is a function of the estimated model parameter (see Box and Jenkins (1976), Chapter 3), an inferential procedure for variances may be translated into an inferential procedure for the estimated model parameters in an indirect manner.

Consider the process $\{Y_t\}$ which would be illustrated by an AR(p) model.

$$Y_t = \phi_1 Y_{t1} + \ldots + \phi_p Y_{t-p} + a_t$$

where $\{a_t\}$ is a sequence of i.i.d. $N(0, \sigma_a^2)$ random variables. Based upon the maximum likelihood estimation procedure in Box and Jenkins, let the maximum likelihood estimates of $(\phi_1, \phi_2, \ldots \phi_p)$ be denoted by $(\hat{\phi}_1, \hat{\phi}_2, \ldots \hat{\phi}_p)$. The estimated variance of Y can be derived as a function of estimated model parameters $(\hat{\phi}_1, \hat{\phi}_2, \ldots \hat{\phi}_p)$ (see Box and Jenkins (1976, pp. 56)). That is

$$\hat{\sigma}^2 = \hat{\phi}_1 \hat{R}_1 + \ldots + \hat{\phi}_p \hat{R}_p + \hat{\sigma}_a^2$$

where $\hat{R}_p$ is the autocovariance estimate of Y at lag p. The precision of the estimates $(\hat{\phi}_1, \hat{\phi}_2, \ldots \hat{\phi}_p)$ can be evaluated from a simultaneous confidence interval procedure developed by Box and Jenkins (1976, pp. 227).

The drawback of this approach is that it can only indirectly construct an inferential procedure for variances.

2.3.4 Conclusion And Impact On The Research

So far the inferential procedures for variances are plagued by the presence of data correlation and the nonnormality of parent populations. The data correlation will hinder the direct application of classical statistics because the independence assumption is violated. The nonnormality of parent population would cause the test procedure developed under normality assumption to be quite non-robust if that assumption is removed. A plausible approach to solve the problem of data correlation and nonnormality may be proceeded as follows. In dealing with the problem of data correlation, past research indicated that an interval estimator for mean may be helpful. One may obtain an unbiased estimator of the population variance from the interval estimator and then approximate this variance estimator with a proper distribution form to resolve the problem of data correlation. In dealing with the problem of nonnormality, past research indicated that test procedures for variances are quite sensitive to the magnitude of kurtosis. To solve the problem of nonnormality, one should try to reduce the impact of kurtosis in the formulation of inferential procedures for variances.

The treatment of the impact of data correlation and nonnormality on inferential procedures for variances is presented in Chapter III. The method suggested to solve the problem of data correlation is presented in Chapter VI. The method suggested to solve the problem of nonnormality is presented in Chapter VII.

2.4 DETERMINATION OF THE ORDER OF SERIAL DEPENDENCE

2.4.1 Introduction

Several methods for means, including replication, batch means, and regeneration cycles, attempts to acquire uncorrelated observations to assist construction of an interval estimate for means. This research also intends to obtain uncorrelated observations. Unlike those methods for means, this research employs the order of serial dependence to obtain uncorrelated observations.

The terminology, the order of serial dependence, comes from Anderson (1971), and is explained as follows:

Given a sequence of observations $X_1, X_2, \ldots, X_n$ from a process with mean μ and variance σ^2, the definition of autocovariance, R_k, and autocorrelation, ρ_k, at lag k are

$$R_k = E(X_t - \mu)(X_{t+k} - \mu)$$

$$\rho_k = \frac{R_k}{R_0}$$

The most frequently used estimators for R_k are given by

$$\hat{R}_k = \frac{1}{n} \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})$$

or

$$\hat{R}'_k = \frac{1}{n-k} \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})$$

Thus an estimator of ρ_k is given by

$$\hat{\rho}_k = \frac{\hat{R}_k}{\hat{R}_0}$$

or

$$\hat{\rho}'_k = \frac{\hat{R}'_k}{\hat{R}_0}$$

This sample autocorrelation measures the serial dependence between observations k time units apart. The process $\{X_i\}$ would be said to exhibit only (k-1) order of serial dependence if the autocorrelations with lag k or greater are considered as zero's. Hence if the order of serial dependence is determined as (k-1), then observations at lag k or greater can be considered as uncorrelated.

The order of serial dependence (k-1) might be determined as follows: For an integer variable k ranging from $1 < k < M$ where M is a large number, sequentially test whether the lag correlation ρ_k dies out at lag k.

The hypothesis tests of such types often involve the distribution theory of sample autocovariances and autocorrelations. However, for finite n, the distribution theory of the sample autocovariances and autocorrelations is complicated. This is particularly true for sample autocorrelations; for each k, $\hat{\rho}_k$ is a ratio of two quadratic forms in the $\{X_i\}$ and therefore the exact result for its sampling distribution is virtually not available due to the complexity of the formulation of $\hat{\rho}_k$. In fact, most of the past research on finite sample distributions has been based upon the assumption of circular correlation to facilitate the derivation of the distributions of sample autocorrelations. The assumption of circular correlation refers to $X_{n+t} = X_t$ for $t = 0,1,2...n$. In case $\mu = 0$, the sample circular autocorrelation at lag s is given by

$$r_s = \sum_{t=1}^{n} X_t X_{t+|s|} / \sum_{t=1}^{n} X_t^2 .$$

The formulation of r_s in general when $\mu \neq 0$ is given in Hannan (1970).

Most of the exact results which have been obtained refer only to the distribution of r_1 for a simple model of $\{X_i\}$. References can be found in R. L. Anderson (1942). Koopmans (1942), Dixon (1944), and Madow (1945). Various approximations have been derived for the distribution of higher autocorrelation (see T. W. Anderson (1948), Hannan (1955), Leipnik (1947), and Hannan (1970)). However, these approximations are usually too complicated for practical use. Anderson (1971) provided normal approximations to the distribution of r_1 for $n = 5,10,...75$.

However, as far as the asymptotic theory is concerned, a plausible approach is that of determining limiting sampling distribution of autocorrelations. Using the limiting distribution the order of serial dependence can be determined.

An alternative approach is that of determining whether a sequence of random variables could have been generated by a given model, and then estimate the serial correlation based upon parameters of the model. Through the model fitting of sample observations, an analyst can obtain estimates of the theoretical autocorrelations based upon estimates of the model parameters (see section (2.4.3) for reference). Using the estimates of the theoretical autocorrelations he can decide where the autocorrelation of the data is deemed to have died out and then determine the order of the serial dependence (see section (8.2) of

Chapter VIII). However, this approach requires a goodness of fit test to examine the adequacy of the model fitted.

Based upon these two approaches, the literature search of tests for serial correlation will be directed to the following area:

1. Asymptotic normality of sample autocorrelations.
2. Goodness of fit test based upon the sample autocorrelations of residuals.

2.4.2 Asymptotic Normality of Sample Autocorrelations

The approach based upon asymptotic normality of sample autocorrelations is to obtain a limiting distribution (usually normal) for sample autocorrelations. The determination of the order of serial dependence can be based upon the limiting distribution for sample autocorrelations.

Anderson (1971, pp. 478) gives the following results of the asymptotic distribution of the sample autocorrelations:

Let $\{X_i\}$ be a general linear process of the form

$$X_t = \mu + \sum_{s=-\infty}^{\infty} g_s \varepsilon_{t-s}$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) < \infty$, $E(\varepsilon_t^4) < \infty$, $\sum_{s=-\infty}^{\infty} |g_s| < \infty$, and with $\sum_{s=-\infty}^{\infty} |s| g_s^2 < \infty$, the limiting joint distribution of $(\hat{\rho}_1 - \rho_1)$, $(\hat{\rho}_2 - \rho_2) \ldots (\hat{\rho}_k - \rho_k)$ (as $n \to \infty$) is multivariate normal with zero means and variance and covariance given by

$$\text{Cov}[\hat{\rho}_r,\hat{\rho}_{r+u}] \sim \frac{1}{n}\sum_{m=-\infty}^{\infty}[\rho_m\rho_{m+u} + \rho_{m+r+u}\rho_{m-r} + 2\rho_r\rho_{r+u}\rho_m^2 - 2\rho_r\rho_m\rho_{m-r-u} - 2\rho_{r+u}\rho_m\rho_{m-r}] \qquad (2.4.2.1)$$

$$\text{Var}(\hat{\rho}_r) \sim \frac{1}{n}\sum_{m=-\infty}^{\infty}[\rho_m^2 + \rho_{m+r}\rho_{m-r} + 2\rho_r^2\rho_m^2 - 4\rho_r\rho_m\rho_{m-r}] \qquad (2.4.2.2)$$

Priestley (1981, pp. 269) suggested that the asymptotic normality of the sample autocorrelation at lag r may be approximated by

$$\hat{\rho}_r \sim N(\rho_r,\text{Var}(\hat{\rho}_r)) \qquad (2.4.2.3)$$

If $\rho_s \to 0$ ans $|s| \to \infty$, and if r is sufficiently large so that

$$\rho_r \sim 0, \ |s| > r$$

then (2.4.2.1) and (2.4.2.2) are reduced to

$$\text{Cov}[\hat{\rho}_r,\hat{\rho}_{r+u}] \sim \frac{1}{n}\sum_{m=-\infty}^{\infty}[\rho_m\rho_{m+u}] \qquad (2.4.2.4)$$

$$\text{Var}(\hat{\rho}_r) \sim \frac{1}{n}\sum_{m=-\infty}^{\infty}\rho_m^2 \qquad (2.4.2.5)$$

Box and Jenkins [1976] employed (2.4.2.5) in deciding where the true autocorrelation has "died out". The null hypothesis is that the true autocorrelation dies out at lag k,

$$H_0 : \rho_k = 0.$$

Under the null hypothesis, for the variance of sample autocorrelation,

$$\text{Var}(\hat{\rho}_k) = \frac{1}{n}\{1 + 2\sum_{u=1}^{q}\hat{\rho}_u^2\}, \ k > q.$$

The null hypothesis is accepted if

$$-Z_{1-\alpha/2}\sqrt{\text{Var}(\hat{\rho}_k)} < \hat{\rho}_k < Z_{1-\alpha/2}\sqrt{\text{Var}(\hat{\rho}_k)}. \qquad (2.4.2.6)$$

Additional references on asymptotic normality of sample autocorrelations can be found in Barlett (1946), Fuller (1974), and

Anderson (1971).

There are two drawbacks in this approach. First, the limiting normal distribution in (2.4.2.3) applies only asymptotically. Second, the estimator of the sample autocovariances at lag h, $\hat{R}_h$, is not an unbiased estimator at that lag. According to Priestly (1981, pp. 323) the bias of $\hat{R}_h$ is given by

$$E(\hat{R}_h) - R_h = \frac{|h|}{n} - \frac{2\pi R_0(n-|h|)}{n^2} f(0) \qquad (2.4.2.7)$$

where f(0) is the spectral density function at zero frequency. Notice that the absolute value of bias increases as the lag h increases. Similar conclusions apply to the case of correlation estimators. Based upon 100 waiting time observations for the M/M/1 queueing model with .9 traffic density, Law and Kelton (1982b) estimated the correlation of the waiting time process. The results are disappointing and are illustrated in Figure 2. Hence the asymptotic normality for sample autocorrelations should be applied with caution.

2.4.3 Goodness of Fit Test

The approach which employs a goodness of fit test obtains an estimate of the theoretical autocorrelation based upon an ARMA(p,q) model of the sample observations and a test to justify the adequacy of the theoretical autocorrelation estimates. For example, if a sequence of correlated data can be represented by an AR(p) model,

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = a_t \qquad (2.4.3.1)$$

where $\{a_t\}$ is a sequence of i.i.d. $N(0,\sigma^2)$ random variables. The autocovariance at lag k, R_k, can be obtained from the assumed model

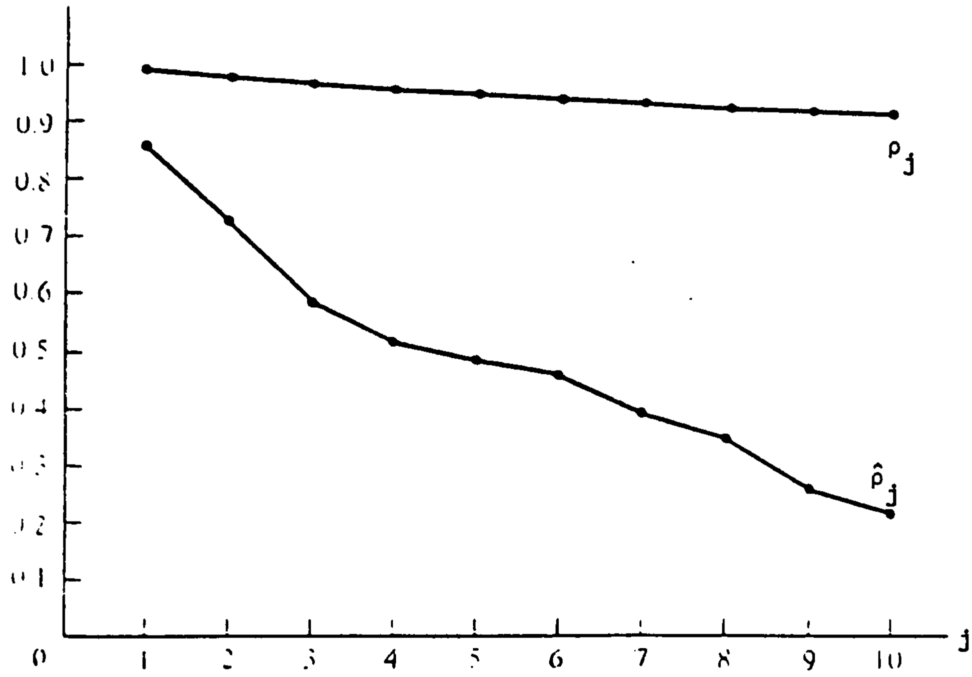

Figure 2. True and Estimated Correlations of the Waiting Time Process for the M/M/1 Queue with $\rho = 0.9$.

Source: Law and Kelton (1982).

$$R_k = \phi_1 R_{k-1} + \phi_2 R_{k-2} + \cdots + \phi_p R_{k-p}.$$

Then the autocorrelation at lag k, ρ_k, is

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}.$$

Based upon an assumed model for the data, the autocovariance and autocorrelation acquired in this manner are referred to as the theoretical autocovariance and autocorrelation in this proposal. Since most of the time the true model for a sequence of data is unknown, model fitting procedures are usually applied to find a suitable model. A very popular estimation procedure of model parameters is the maximum likelihood estimation procedure in Box and Jenkins (1976). Applying this procedure one can obtain the maximum likelihood estimation of $(\phi_1, \phi_2, \ldots, \phi_p)$. With these estimates the estimated theoretical autocorrelation can be derived using the above formulation of ρ_k (see Box and Jenkins (1976, pp. 55)).

Since the theoretical autocorrelation depends on the model parameters, it is essential to have a reliable goodness of fit test to examine the adequacy of the fitted model. There are four goodness of fit tests available. They include Whittle, Box-Pierce, Ljung-Box, and Godolphin test.

For example, all these test procedures attempt to test for the fit of the AR(p) process given by equation (2.4.3.1). Assuming that a realization $X_1, X_2, \ldots, X_n$ of the process (2.4.3.1) is available, the objective shared by these procedures is to test departures from the assumption that $\{a_t\}$ comprises a sequence of i.i.d. $N(0, \sigma^2)$ random variables. To this end, it is required to compute the residuals

$$e_t = X_t - \hat{\phi}_1 X_{t-1} - \cdots - \hat{\phi}_p X_{t-p}$$

where $\hat{\phi}_1,\hat{\phi}_2,\ldots,\hat{\phi}_p$, are estimates of $\phi_1,\phi_2,\ldots,\phi_p$, respectively. These estimates can be obtained from a model fitting procedure. Based upon the residuals, each test derives a general test statistic for the adequacy of $(\hat{\phi}_1,\hat{\phi}_2,\ldots,\hat{\phi}_p)$.

Whittle's procedure (see Shittle [1952]) is based on a variance ratio criterion that is asymptotically equivalent to the likelihood-ratio statistics. Box-Pierce test is based upon the distribution theory developed by Box and Pierce (1970). Ljung-Box test (1978) is a modification, in fact an improvement, of the Box-Pierce test. Godolphin (1980) extends the Box-Pierce theory by deriving a procedure which employs an uncorrelated residual transformation under the null hypothesis. Among these tests, the Ljung-Box test is the most widely-known. This may be due in part to that the test is simple to use.

From the empirical study by Davies and Newbold (1979), the Ljung-Box test was found to have considerable power in discriminating alternative models. Clarke and Godolphin (1982) conducted an empirical study investigating powers of the Shittle, Ljung-Box and Godolphin test. Their results indicated that the Godolphin test appears to be the most powerful with the Ljung-Box test ranked as the second. However, the residual transformation of the Godolphin test makes it the most computation-demanding procedure among the available. The maximum sample size employed in the two empirical studies conducted is 200. Such a sample size may not be large enough since all the test procedures are asymptotically developed.

2.4.4 Conclusion and Impact on the Research

Determination of the order of serial dependence involves tests of the lag correlation. There are two major difficulties in tests of the lag correlation. First, the exact distributions of the sample autocorrelations are difficult to obtain. Second, satisfactory estimates or autocorrelations can only be obtained if the sample observations are independent but they are assumed correlated for the purpose of this research. While the asymptotic normality of the sample autocorrelations may be established, the estimate of autocorrelations may yield considerable bias if the sample size is small and the lag is large.

The approach utilizing ARMA modeling may avoid the two difficulties cited above in determining the order of serial dependence. However, this approach requires a goodness of fit test to examine the adequacy of the model accepted. From empirical studies of Davies and Newbold (1979), Clarke and Godolphin (1982), the Ljung-Box test appears to be powerful and relatively easy to use. This test may serve well as a procedure for determining the adequacy of a fitted ARMA(p,q) model.

CHAPTER III

THE VIOLATIONS OF INDEPENDENCE AND NORMALITY ASSUMPTION

3.1 INTRODUCTION

Parametric inferential statistical procedures for means and variances typically assume that the sample data are normally and independently distributed. Two important questions are of interest. What consequences may be anticipated if those procedures are applied when either or both of these assumptions are violated? Second, if the consequences are considered serious with respect to the application at hand, what steps may be taken to negate or reduce the deleterious impact of violation of these assumptions? This chapter will treat the subject of the first question. The subject related to the second question will be treated from Chapter IV and Chapter VII.

The impact of violation of any assumption may be measured in a variety of ways. In the treatment here it will be measured in one dimension. The impact on the probability of error for the inferential procedure considered will be analyzed.

In the case of means, this chapter is focused on the impact of violation of the independence assumption. This is because violation of the normality assumption has little impact on the inferential procedures for means if a sufficiently large sample is employed (see Section (2.4.1) of Chapter II).

In the case of procedures for variances, this chapter addresses the effect on the predictability of the analysis if the assumptions of independence and normality are violated.

3.2 THE ASSUMPTION OF INDEPENDENCE AND NORMALITY

Throughout the discussion of confidence intervals and hypothesis for means the observations within each sample were assumed to be correlated. In the case of two samples (two unknown means) independence was assumed between samples.

From Section (3.3) to (3.6) in which the statistical analysis for means are presented, the sample observations will be assumed to be normally distributed so that the effect of violation of the independence assumption will be analyzed. The analysis of such violation is conducted by applying statistical inferential procedures that are based upon the assumption of independence and ignore the existence of data dependence.

The probable computational forms of confidence intervals and tests of hypothesis for variance are difficult to obtain if the data are either correlated or nonnormally distributed. For this reason, the discussion of violation of the assumptions of independence and normality is led by the simulation experiments which investigate the effect of such violation. In Section (3.7) where the analysis of the simulation experiments for variances is presented, the assumptions of sample observations in an experiment are in succession of the following:

1. correlated and normally distributed,
2. uncorrelated and normally distributed,
3. correlated and nonnormally distributed.

This sequence of assumptions is made to analyze the effect of violation of independence assumption, normality assumption, and both.

The simulation experiments based upon the three sets of assumptions are conducted for both single variance and two variances. In the case of two samples (two variances) independence was assumed between samples.

The effect of violation of the assumptions of independence and normality will be measured in terms of the resulting induced error in the level of confidence for confidence intervals and the power function for hypothesis tests. Let $1 - \alpha_D$ be assumed or design level of confidence and $1 - \beta_D(\phi)$ the assumed or design power function of the test where ϕ is the parameter tested. If, because of violation of the assumption of independence, the actual level of confidence is $1 - \alpha_T$ and the actual power function $1 - \beta_T(\phi)$ then the error induced in the level of confidence is given by $\Delta_{1-\alpha}$ where

$$\begin{aligned}\Delta_{1-\alpha} &= (1 - \alpha_T) - (1 - \alpha_D)\\ &= \alpha_D - \alpha_T.\end{aligned}$$

In a similar manner the induced error for the power of the test is $\Delta_{1-\beta(\phi)}$ where

$$\begin{aligned}\Delta_{1-\beta(\phi)} &= (1 - \beta_T(\phi)) - (1 - \beta_D(\phi))\\ &= \beta_D - \beta_T(\phi).\end{aligned}$$

The conclusion is drawn as to the magnitude of induced error on inferential characteristics such as the level of confidence and the power function.

3.3 BIAS CAUSED BY THE DATA CORRELATION

Assume that the sample observations $X_1, X_2, \cdots, X_n$ are normally distributed with mean μ and variance σ^2. If the purpose of the analysis is simply to provide a point estimate for μ, the sample mean, $\bar{X}$, proves to be entirely satisfactory whether the sample observations are independent or not since $E(\bar{X}) = \mu$ in either case.

Since $\bar{X}$ is a random variable, the discrepency between $\bar{X}$ and μ can be described only in probabilistic terms. This is the intent of a confidence interval for μ. However, to define a confidence interval for μ the analyst must know something about the variance of $\bar{X}$, $\sigma^2_{\bar{x}}$. It is at this point the effect of violation of the assumption of independence becomes evident.

When dealing with normally distributed random variables the degree of dependence between two observations is usually measured by the coefficient of correlation, ρ_{ij}. That is if X_i and X_j are the ith and jth observations in a sequence of observations then

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$$

where σ_i and σ_j are the standard deviations of X_i and X_j respectively, and σ_{ij} is the covariance of X_i and X_j.

In the presence of data correlation, the variance of $\bar{X}$ can be shown as

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} + 2\,\frac{\sigma^2}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \rho_{ij}$$

$$= \frac{\sigma^2}{n}\,(1 + 2R_\rho(n)) \qquad (3.3.1)$$

where

$$R_\rho(n) = \frac{1}{n} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \rho_{ij}.$$

Thus the assumption that the variance of $\bar{X}$ is σ^2/n carries the implication that $R_\rho(n) = 0$. The bias in σ^2/n as an estimate of $\sigma^2_{\bar{X}}$, $B_{\frac{\sigma^2}{n}, \sigma^2_{\bar{X}}}$, is given by

$$
\begin{aligned}
B_{\frac{\sigma^2}{n}, \sigma^2_{\bar{X}}} &= \frac{\sigma^2}{n} - \sigma^2_{\bar{X}} \\
&= \frac{\sigma^2}{n} - \frac{\sigma^2}{n}(1 + 2R_\rho(n)) \\
&= -2\frac{\sigma^2}{n} R_\rho(n). \qquad (3.3.2)
\end{aligned}
$$

One would expect the magnitude of the error in probability statements concerning $\bar{X}$, when the variance of $\bar{X}$ is erroneously taken to be σ^2/n, to depend upon the magnitude and sign of $R_\rho(n)$.

Typically the population variance (σ^2) is unknown. In such cases σ^2 is estimated by s^2 and $\sigma^2_{\bar{X}}$ by s^2/n. The expected value of s^2 is

$$
\begin{aligned}
E(s^2) &= E[\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2] \\
&= \sigma^2[1 - \frac{2}{n-1} R_\rho(n)]. \qquad (3.3.3)
\end{aligned}
$$

Hence

$$B_{s^2, \sigma^2} = -\frac{2\sigma^2}{n-1} R_\rho(n). \qquad (3.3.4)$$

If the variance of $\bar{X}$ is estimated by s^2/n in the case of correlated sample observations the resulting bias is

$$B_{\frac{s^2}{n},\sigma^2_{\bar{X}}} = E(\frac{s^2}{n}) - \sigma^2_{\bar{X}}$$

$$= -2\,\frac{\sigma^2}{n-1}\,R_\rho(n). \qquad (3.3.5)$$

Equations (3.3.2) and (3.3.5) indicate that the bias in the estimate of the variance of $\bar{X}$ when independence is erroneously assumed is nearly the same whether the population variance is unknown or not, at least for large n. Hence, as in the case of a known population variance, the magnitude of the error in probability statements concerning $\bar{X}$ will depend upon the magnitude and sign of $R_\rho(n)$. In the next two sections the magnitude of this error will be evaluated for confidence intervals and tests of hypothesis.

3.4 CONFIDENCE INTERVALS FOR MEANS

If the analyst assumes independence of the sample observations when $\rho_{ij} = 0$ for at least some $i \neq j$, one would anticipate that the actual level of confidence achieved would differ from that desired since the estimate of $\sigma^2_{\bar{x}}$ used to define the confidence limits is biased. Suppose that the analyst recognizes that the sample observations are not independent and he uses the estimator $s^2_{\bar{x}}$ which has the expected value $\hat{\sigma}^2_{\bar{x}}$. It should be noted that $s^2_{\bar{x}} \neq s^2/n$ in general. Hopefully, if $s^2_{\bar{x}}$ is not an unbiased estimator of $\sigma^2_{\bar{x}}$, it is less biased than σ^2/n or s^2/n. Unless $s^2_{\bar{x}}$ is an unbiased estimate of $\sigma^2_{\bar{x}}$, the level of confidence actually achieved, $1 - \alpha_T$, on the basis of this estimator will differ from the desired, $1 - \alpha_D$. However, the discrepency between $1 - \alpha_T$ and $1 - \alpha_D$ should be less using $s^2_{\bar{x}}$ than would be the case were independence assumed.

For the case of a single population, the $100(1 - \alpha_D)\%$ confidence limits for μ based upon $s_{\bar{X}}^2$ as an estimate of $\sigma_{\bar{X}}^2$ are given by

$$L,U = \bar{X} \pm s_{\bar{X}}\, t_{1 - \frac{\alpha_D}{2}, f} \tag{3.4.1}$$

where f is the degrees of freedom. The propriety of application of equation (3.4.1) depends upon whether or not $(\bar{X} - \mu)/s_{\bar{X}}$ is t distributed with f degrees of freedom. Let t_{exp} be defined as

$$t_{exp} = \frac{(\bar{X} - \mu)}{s_{\bar{X}}}$$

$$= \frac{(\bar{X} - \mu)/\sigma_{\bar{X}}}{s_{\bar{X}}/\sigma_{\bar{X}}}\,.$$

Now, assuming that the sample observations are normally and independently distributed, $(\bar{X} - \mu)/\sigma_{\bar{X}} \sim N(0,1)$. If, in addition, $fs_{\bar{X}}^2/\sigma_{\bar{X}}^2 \sim X^2(f)$, then, by definition of the t random variable, $t_{exp} \sim t(f)$ provided $\bar{X}$ and $s_{\bar{X}}^2$ are independent random variables. In the case of uncorrelated data, $\sigma_{\bar{X}}^2 = \sigma^2/n$ and $s_{\bar{X}}^2 = s^2/n$ and because of the normality assumption $s_{\bar{X}}^2$ and $\bar{X}$ are independent random variables, $s_{\bar{X}}^2/\sigma_{\bar{X}}^2 = s^2/\sigma^2$ and $fs^2/\sigma^2 \sim \chi^2(f)$ where $f = n - 1$. However, for the case of correlated data $s_{\bar{X}}^2$ is not an unbiased estimator of $\sigma_{\bar{X}}^2$, and $s_{\bar{X}}^2$ and $\bar{X}$ may not be independent. Hence $fs_{\bar{X}}^2/\sigma_{\bar{X}}^2$ may not be χ^2 distributed with f degrees of freedom and in turn t_{exp} may not be t distributed with f degrees of freedom.

In order to describe the behavior of t statistic in the case of correlated data, the approximate t statistic is defined as follows:

$$t' = \frac{(\bar{X} - \mu)/\sigma_{\bar{x}}}{s_{\bar{x}}/\sigma_{\bar{x}}}$$

$$= \frac{\sigma_{\bar{x}}}{\sqrt{E(s_{\bar{x}}^2)}} \frac{(\bar{X} - \mu)/\sigma_{\bar{x}}}{s_{\bar{x}}/\sqrt{E(s_{\bar{x}}^2)}}$$

$$= mt \sim t'(f). \tag{3.4.2}$$

where

$$m = \frac{\sigma_{\bar{x}}}{\sqrt{E(s_{\bar{x}}^2)}}, \text{ and } t = \frac{(\bar{X} - \mu)/\sigma_{\bar{x}}}{s_{\bar{x}}/\sqrt{E(s_{\bar{x}}^2)}}.$$

Assume

$$f \frac{s_{\bar{x}}^2}{E(s_{\bar{x}}^2)} \sim \chi^2(f).$$

The degrees of freedom, f, is obtained from solving the following two equations:

$$E[f \frac{s_{\bar{x}}^2}{E(s_{\bar{x}}^2)}] = f \tag{3.4.3}$$

$$Var[f \frac{s_{\bar{x}}^2}{E(s_{\bar{x}}^2)}] = 2f. \tag{3.4.4}$$

The procedure of obtaining the approximate degrees of freedom through solving the equations (3.4.3) and (3.4.4) simultaneously is originated from Satterthwaite (1946). Thus

$$f = \frac{2E^2(s_{\bar{x}}^2)}{Var(s_{\bar{x}}^2)}. \tag{3.4.5}$$

The mathematical expression developed for f, for the case that $s_{\bar{x}}^2$ is given by s^2/n, is given in Appendix A. In (3.4.2) the scalar m accounts for the impact attributed to the estimator bias of $\sigma_{\bar{x}}^2$, and the t statistic in (3.4.2) accounts for the impact attributed to the distortion of the degrees of freedom. Both the estimator bias and the distortion of the degrees of freedom are the results of data correlation.

Suppose the approximate degrees of freedom f_1 as computed from equation (3.4.5) is known and therefore the confidence limits are given by

$$L,U = \bar{\bar{X}} \pm s_{\bar{\bar{x}}}\, t_{1-\frac{\alpha_D}{2},f_1} . \tag{3.4.6}$$

The actual level of confidence, $1 - \alpha_T$ is

$$\begin{aligned}
1 - \alpha_T &= \Pr[L < \mu < U] \\
&= \Pr[\bar{\bar{X}} - s_{\bar{\bar{x}}}\, t_{1-\frac{\alpha_D}{2},f_1} < \mu < \bar{\bar{X}} + s_{\bar{\bar{x}}}\, t_{1-\frac{\alpha_D}{2},f_1}] \\
&= \Pr[\mu - s_{\bar{\bar{x}}}\, t_{1-\frac{\alpha_D}{2},f_1} < \bar{\bar{X}} < \mu + s_{\bar{\bar{x}}}\, t_{1-\frac{\alpha_D}{2},f_1}] \\
&= \Pr[-t_{1-\frac{\alpha_D}{2},f_1} < t'(f_1) < t_{1-\frac{\alpha_D}{2},f_1}] \\
&= \Pr[-\frac{1}{m}\, t_{1-\frac{\alpha_D}{2},f_1} < t < \frac{1}{m}\, t_{1-\frac{\alpha_D}{2},f_1}] \\
&= \Pr[-\frac{\sqrt{E(s_{\bar{\bar{x}}}^2)}}{\sigma_{\bar{\bar{x}}}}\, t_{1-\frac{\alpha_D}{2},f_1} < t < \frac{\sqrt{E(s_{\bar{\bar{x}}}^2)}}{\sigma_{\bar{\bar{x}}}}\, t_{1-\frac{\alpha_D}{2},f_1}]
\end{aligned}$$

$$= \Pr[-\frac{\hat{\sigma}_{\bar{\bar{X}}}}{\sigma_{\bar{\bar{X}}}} t_{1-\frac{\alpha_D}{2}, f_1} < t < \frac{\hat{\sigma}_{\bar{\bar{X}}}}{\sigma_{\bar{\bar{X}}}} t_{1-\frac{\alpha_D}{2}, f_1}] \tag{3.4.7}$$

where

$$\hat{\sigma}^2_{\bar{\bar{X}}} = E(s^2_{\bar{\bar{X}}}).$$

In the case of large samples, the critical value of t distribution in equation (3.4.6) may be replaced by that of normal distribution. Hence the approximation of the degrees of freedom may not be needed. Also the estimator of $\sigma^2_{\bar{\bar{X}}}$ may be replaced by its expectation in the case of large samples. That is,

$$t_{1-\frac{\alpha_D}{2}, f_1} \longrightarrow Z_{1-\frac{\alpha_D}{2}} \text{ and } s^2_{\bar{\bar{X}}} \longrightarrow \hat{\sigma}^2_{\bar{\bar{X}}}$$

for large samples. Hence the confidence limits in equation (3.4.6) become

$$L, U = \bar{\bar{X}} \pm \hat{\sigma}_{\bar{\bar{X}}} Z_{1-\frac{\alpha_D}{2}} . \tag{3.4.8}$$

Using the confidence limits in equation (3.4.8), one can obtain the actual level of confidence by

$$1 - \alpha_T \simeq \Pr(\bar{\bar{X}} - \hat{\sigma}_{\bar{\bar{X}}} Z_{1-\frac{\alpha_D}{2}} < \mu < \bar{\bar{X}} + \hat{\sigma}_{\bar{\bar{X}}} Z_{1-\frac{\alpha_D}{2}})$$

$$\simeq \Pr(\mu - \hat{\sigma}_{\bar{x}} Z_{1-\frac{\alpha_D}{2}} < \bar{\bar{X}} < \mu + \hat{\sigma}_{\bar{\bar{X}}} Z_{1-\frac{\alpha_D}{2}})$$

$$\simeq \Pr(-\frac{\hat{\sigma}_{\bar{\bar{X}}}}{\sigma_{\bar{\bar{X}}}} Z_{1-\frac{\alpha_D}{2}} < Z < \frac{\hat{\sigma}_{\bar{\bar{X}}}}{\sigma_{\bar{\bar{X}}}} Z_{1-\frac{\alpha_D}{2}}) \tag{3.4.9}$$

where

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \sim N(0,1).$$

Notice that one can obtain equation (3.4.9) from equation (3.4.7) by replacing $t_{1-\frac{\alpha_D}{2},f_1}$ with $Z_{1-\frac{\alpha_D}{2}}$.

To illustrate the effect of bias in estimating $\sigma^2_{\bar{X}}$, consider the case where $\sigma^2_{\bar{X}}$ is estimated by σ^2/n. That is $s^2_{\bar{X}} = \sigma^2/n$. From equation (3.3.1)

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}(1 + 2R_\rho(n))$$

and

$$\hat{\sigma}_{\bar{X}}/\sigma_{\bar{X}} = (1 + 2R_\rho(n))^{-1/2}.$$

Thus $1 - \alpha_T$ is given by

$$1 - \alpha_T = \Pr\left(-\frac{Z_{1-\frac{\alpha_D}{2}}}{\sqrt{1 + 2R_\rho(n)}} < Z < \frac{Z_{1-\frac{\alpha_D}{2}}}{\sqrt{1 + 2R_\rho(n)}}\right). \qquad (3.4.10)$$

In the case of confidence intervals for two population means, μ_1 and μ_2, the results are much the same. While the observations within the samples from each population are not independent, independence of observations between the samples is assumed. Let σ_1^2 and σ_2^2 be the respective variances of the observations for the two populations and let $\rho_{ij}^{(1)}$ and $\rho_{hk}^{(2)}$ the coefficients of correlation between observations i and j for population 1 and observations h and k for population 2. Assume that $\sigma_1^2 \neq \sigma_2^2$ and $\rho_{ij}^{(1)} \neq \rho_{hk}^{(2)}$ for at least some $i \neq h$ and $j \neq k$.

If $s^2_{\bar{x}_1}$ and $s^2_{\bar{x}_2}$ are the estimates of the variances of the sample means for the two populations, $\sigma^2_{\bar{x}_1}$ and $\sigma^2_{\bar{x}_2}$, then the $(1 - \alpha_D)100\%$ confidence limits for $\mu_1 - \mu_2$ are given by

$$L,U = (\bar{X}_1 - \bar{X}_2) \pm \sqrt{s^2_{\bar{x}_1} + s^2_{\bar{x}_2}}\; Z_{1 - \frac{\alpha_D}{2}}$$

for large samples, n_1 and n_2, such that $t_{1 - \frac{\alpha_D}{2}, f} \longrightarrow Z_{1 - \frac{\alpha_D}{2}}$ where f is the appropriate degrees of freedom, and $s^2_{\bar{x}_i} \longrightarrow \hat{\sigma}^2_{\bar{x}_i}$ where $i = 1,2$.

By arguments similar to those applied in the case of a single population mean it can be shown

$$1 - \alpha_T = \Pr\Big(\mu_1 - \mu_2 - \sqrt{\hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}}}\; Z_{1 - \frac{\alpha_D}{2}} < \bar{X}_1 - \bar{X}_2$$

$$< \mu_1 - \mu_2 + \sqrt{\hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}_2}}\; Z_{1 - \frac{\alpha_D}{2}}\Big)$$

Since

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}}}} \sim N(0,1)$$

$$1 - \alpha_T = \Pr\left(-\sqrt{\frac{\hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}_2}}{\sigma^2_{\bar{x}} + \sigma^2_{\bar{x}_2}}}\; Z_{1 - \frac{\alpha_D}{2}} < Z < \sqrt{\frac{\hat{\sigma}^2_{\bar{x}} + \hat{\sigma}^2_{\bar{x}_2}}{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}}}\; Z_{1 - \frac{\alpha_D}{2}}\right).$$

3.5 HYPOTHESIS TESTS FOR MEANS

The effect of bias in estimating the variance of the sample mean on the tests of hypotheses for means will be evaluated by examining the effect of bias on the power function for the test. Specifically, for the case of a single mean the power function for the test of H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$ will be examined and by inference will lead to conclusions concerning the alternative one-tail tests. The same approach will be adopted in the case of two population means.

The limits for the test of H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$ are given by

$$L,U = \mu_0 \pm s_{\bar{x}}\, t_{1-\frac{\alpha_D}{2},f}$$

where a large sample size is assumed such that $t_{1-\frac{\alpha_D}{2},f} \longrightarrow Z_{1-\frac{\alpha_D}{2}}$, and $s_{\bar{x}}^2 \longrightarrow \hat{\sigma}_{\bar{x}}^2$. The actual power of the test is given by

$$1 - \beta_T(\mu) = 1 - \Pr(\mu_0 - \hat{\sigma}_{\bar{x}}\, Z_{1-\frac{\alpha_D}{2}} < \bar{X} < \mu_0 + \hat{\sigma}_{\bar{x}}\, Z_{1-\frac{\alpha_D}{2}} / \mu).$$

Since $\bar{X} \sim N(\mu, \sigma_{\bar{x}}^2)$, letting $Z = (\bar{X} - \mu)/\sigma_{\bar{x}}$ yields

$$1 - \beta_T(\mu) = 1 - \Pr\left(\frac{\mu_0 - \mu}{\sigma_{\bar{x}}} - \frac{\hat{\sigma}_{\bar{x}}}{\sigma_{\bar{x}}} Z_{1-\frac{\alpha_D}{2}} < Z < \frac{\mu_0 - \mu}{\sigma_{\bar{x}}} + \frac{\hat{\sigma}_{\bar{x}}}{\sigma_{\bar{x}}} Z_{1-\frac{\alpha_D}{2}}\right). \tag{3.5.1}$$

The desired power of the test is given by

$$1 - \beta_D(\mu) = 1 - \Pr\left(\frac{\mu_0 - \mu}{\sigma_{\bar{x}}} - Z_{1-\frac{\alpha_D}{2}} < Z < \frac{\mu_0 - \mu}{\sigma_{\bar{x}}} + Z_{1-\frac{\alpha_D}{2}}\right). \tag{3.5.2}$$

Let $\delta = \mu - \mu_0$. The magnitude of the discrepency between the actual power achieved, $1 - \beta_T(\delta)$, and that desired, $1 - \beta_D(\delta)$, can be describe described by

$$\Delta_\beta = [1 - \beta_T(\delta)] - [1 - \beta_D(\delta)]$$

$$= \int_{-\frac{\delta}{\sigma_{\bar{x}}} - \frac{\hat{\sigma}_{\bar{x}}}{\sigma_{\bar{x}}} Z_{1-\frac{\alpha_D}{2}}}^{-\frac{\delta}{\sigma_{\bar{x}}} + \frac{\hat{\sigma}_{\bar{x}}}{\sigma_{\bar{x}}} Z_{1-\frac{\alpha_D}{2}}} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2}\, dZ$$

$$- \int_{-\frac{\delta}{\sigma_{\bar{x}}} - Z_{1-\frac{\alpha_D}{2}}}^{-\frac{\delta}{\sigma_{\bar{x}}} + Z_{1-\frac{\alpha_D}{2}}} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2}\, dZ.$$

For the tests of hypotheses concerning two population means the resulting error from bias may be developed following arguments similar to those applied in the case of a single population mean. Let μ_1 and μ_2 be the respective population means, σ_1^2 and σ_2^2 the population variances and $\rho_{ij}^{(1)}$ and $\rho_{k\ell}^{(2)}$ the coefficients of correlation for observations i and j from population 1 and for observations k and ℓ from population 2.

For the test of H_0: $\mu_1 = \mu_2$ against H_1: $\mu_1 \neq \mu_2$ the test limits, for large sample size, are given by

$$L,U = \pm \sqrt{\hat{\sigma}_{\bar{x}_1}^2 + \hat{\sigma}_{\bar{x}_2}^2}\; Z_{1-\frac{\alpha_D}{2}}$$

where α_D is the design or desired level of significance. Then the true power of the test, $1 - \beta_T(\delta)$, for $\delta = \mu_1 - \mu_2$ is

$$1 - \beta_T(\delta) = 1 - \Pr\left(-\sqrt{\hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}_2}}\, Z_{1-\frac{\alpha_D}{2}} < \bar{X}_1 - \bar{X}_2 < \sqrt{\hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}_2}}\, Z_{1-\frac{\alpha_D}{2}}\right)$$

$$= 1 - \Pr\left(-\frac{\delta}{\sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}}} - \sqrt{\frac{\hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}_2}}{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}}}\, Z_{1-\frac{\alpha_D}{2}} < Z < \frac{-\delta}{\sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}}} + \sqrt{\frac{\hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}_2}}{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}}}\, Z_{1-\frac{\alpha_D}{2}}\right).$$

Since

$$\bar{X}_1 - \bar{X}_2 \sim N(\delta, \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}),$$

let

$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2} = \sqrt{\hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}_2}}, \text{ and } \sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}}.$$

Then

$$1 - \beta_T(\delta) = 1 - \Pr\left(\frac{-\delta}{\sigma_{\bar{x}_1-\bar{x}_2}} - \frac{\hat{\sigma}_{\bar{x}_1-\bar{x}_2}}{\sigma_{\bar{x}_1-\bar{x}_2}}\, Z_{1-\frac{\alpha_D}{2}} < Z < \frac{-\delta}{\sigma_{\bar{x}_1-\bar{x}_2}} + \frac{\hat{\sigma}_{\bar{x}_1-\bar{x}_2}}{\sigma_{\bar{x}_1-\bar{x}_2}}\, Z_{1-\frac{\alpha_D}{2}}\right).$$

The desired power of the test, $1 - \beta_D(\delta)$, is given by

$$1 - \beta_D(\delta) = 1 - \Pr\left(\frac{-\delta}{\sigma_{\bar{x}_1-\bar{x}_2}} - Z_{1-\frac{\alpha_D}{2}} < Z < \frac{-\delta}{\sigma_{\bar{x}_1-\bar{x}_2}} - Z_{1-\frac{\alpha_D}{2}}\right).$$

3.6 TWO EXAMPLES

To illustrate the effect of bias on the confidence intervals and hypothesis tests for means, two examples are considered in this section.

Example A

The sample observations $X_1, X_2, \cdots, X_n$ are normally distributed with common unknown mean μ and common known variance $\sigma^2 = 10$. A 80% confidence interval for μ is to be constructed when the sample observations are erroneously assumed to be independent. If $n = 100$, determine the true level of confidence achieved if

$$\rho_{ij} = \rho^{|i-j|}$$

where $-1 < \rho < 1$.

The actual level of confidence in this case can be computed from equation (3.4.10) where $R_\rho(n)$ is given by

$$\begin{aligned} R_\rho(n) &= \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \rho_{ij} \\ &= \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \rho^{j-i} \\ &= \frac{1}{n}\sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \rho^{k} \\ &= \frac{\rho}{n(1-\rho)^2}\,[n(1-\rho) + \rho^n - 1]. \end{aligned}$$

The results are summarized in Table 5. As Table 5 illustrates, the discrepency between the actual level of confidence and the desired level of confidence (80% in this example) is affected by the magnitude

Table 5

Difference Between the True Level of Confidence $(1 - \alpha_T)$ and Desired Level of Confidence $(1 - \alpha_D)$ for Means for Correlated Normal Random Variables When the Traditional Procedure for Unknown Means is Applied

$(E_{1-\alpha} = (1 - \alpha_T) - (1 - \alpha_D))$, $\alpha_D = 20\%$.

ρ	$1 - \alpha_T$	$E_{1-\alpha}$
0.9	0.2424	-0.5576
0.8	0.3375	-0.4625
0.7	0.4144	-0.3856
0.6	0.4848	-0.3182
0.4	0.6000	-0.2000
0.2	0.7050	-0.095
0.1	0.7578	-0.0422
0	0.80	0.0
-0.1	0.8471	0.0471
-0.2	0.8823	0.0823
-0.4	1.0	0.2000

of ρ.

Example B

As in the case of example A, the sample observations $X_1, \cdots, X_n$ are normally distributed with common unknown mean μ and common known variance $\sigma^2 = 10$. To illustrate the effect of correlated data in the case of hypothesis tests, the following hypothesis is to be conducted:

$$H_0: \ \mu \leq \mu_0$$

$$H_1: \ \mu > \mu_0.$$

If n = 100, and the desired significance level is 20%, determine the actual and desired power of the test when sample observations are erroneously assumed to be independent. Assume

$$\rho_{ij} = \rho^{|i-j|}$$

where $\rho = 0.9, -0.2$.

Given the assumption about the correlation ρ_{ij}, the actual power of the test can be obtained from equation (3.5.1) and the desired power of the test can be obtained from equation (3.5.2). The results are presented in Table 6 and shown graphically in Figure 3. As the results indicate, the violation of independence assumption could lead to a significant loss in the power of the test considered.

3.7 PARAMETRIC INFERENTIAL PROCEDURES FOR VARIANCES

In the chapter of literature review, the general conclusion reached in the discussion about the parametric inferential procedures for means was that the effect of violation of the assumption of normality could be overcome by choosing sufficiently large samples. The same conclusion can not be drawn with respect to parametric

Table 6

True and Desired Powers for Means for Correlated Normal Random Variables When the Conventional t Test is Applied.

$d(\lvert\mu_1 - \mu_0\rvert)$	$1 - \beta_T$ (ρ = −0.2)	$1 - \beta_T$ (ρ = 0.9)	$1 - \beta_D$
0	0.1177	0.7576	0.2000
0.4	0.4936	0.7683	0.2200
0.8	0.9367	0.7970	0.2800
1.2	0.9989	0.8384	0.3710
1.6	1.0	0.8820	0.4810
2.0	1.0	0.9210	0.5980
2.4	1.0	0.9520	0.7090
2.8	1.0	0.9730	0.8030
3.2	1.0	0.9860	0.8770
3.6	1.0	0.9940	0.9280
4.0	1.0	0.9970	0.961

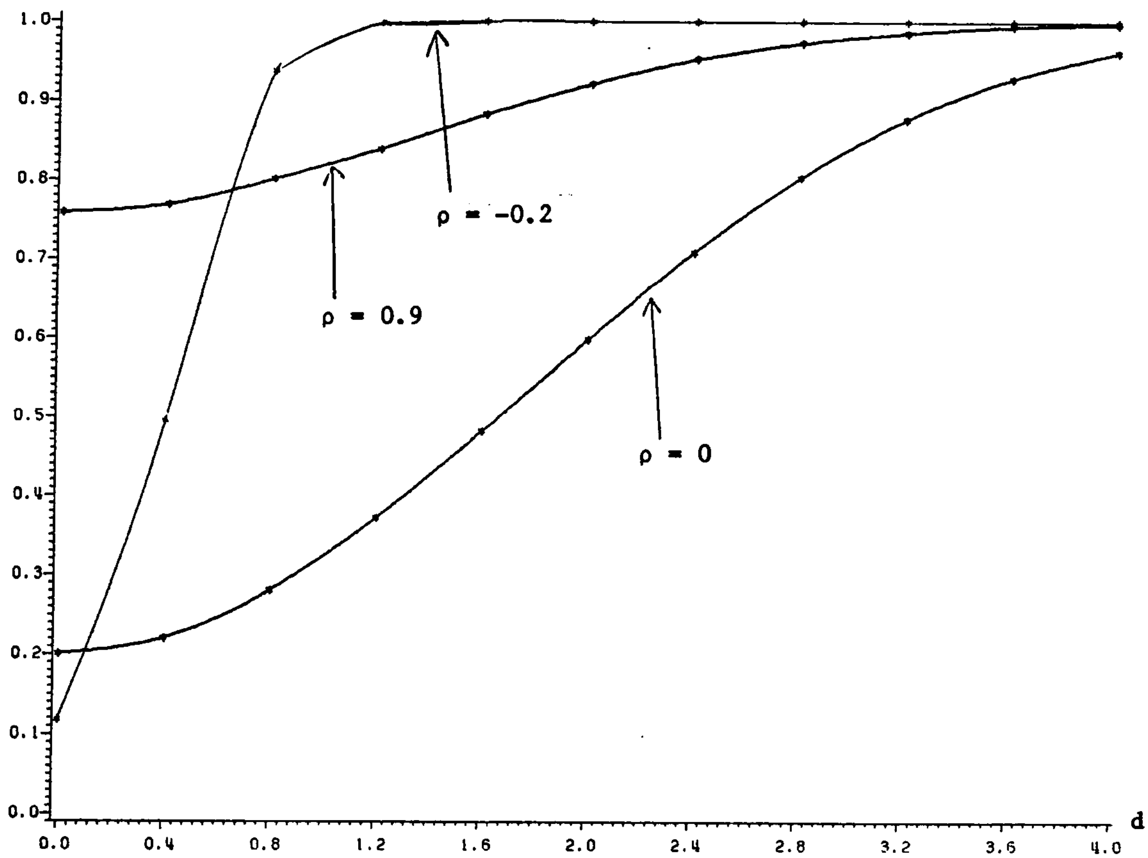

Figure 3. True and Desired Powers for Means for Correlated Data When the Conventional t Test is Applied

procedures for variances. Furthermore, the problem may be compounded by violation of the independence assumption. To illustrate the nature of the problem, consider a test of H_0: $\sigma^2 < \sigma_0^2$ against H_1: $\sigma^2 > \sigma_0^2$. For a sequence of observations $X_1, X_2, \cdots, X_n$, the test statistic usually applied in this case is

$$\chi^2_{exp} = \frac{(n-1)S^2}{\sigma_0^2}$$

where

$$s^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2.$$

The statistic χ^2_{exp} is chi-square distributed with n-1 degrees of freedom if the observations X_i are independent and normally distributed. However, if the sample observations are independent but nonnormally distributed then the test statistic χ^2_{exp} may not be chi-square distributed with n-1 degrees of freedom. A similar conclusion applies if the sample observations are correlated but normally distributed.

For the case of testing the variances of two populations, the classical technique employs the F test, which is subject to the same criticism encountered in the case of chi-square test.

Six examples in this section illustrate the potential consequences of violation of the assumption of normality and independence when the classical procedures for variances are conducted. The first two examples treat violation of the assumption of independence. The third and fourth treat violation of the assumption of normality. The last two examples treat violation of both assumptions. In each case, the effect of violation of an assumption is measured by the discrepency between the desired and actual power of the test.

To determine the precision with which any approximate test procedure predicts the true power of the test one needs to know something about the true power function. One might attempt to derive the exact distribution of the test statistic involved, but this is an arduous task indeed in the case of variance tests. Instead one can simulate the test procedure assuming a specified normal or nonnormal distribution with the known σ^2, recording the frequency with which the null hypothesis is rejected for a given $R_\sigma(\frac{\sigma_1^2}{\sigma_0^2})$. If r is the number of rejections of H_0 out of a total of s replications of the simulation experiment for a given R_σ then the estimated power, $1 - \beta_E$, is given by

$$1 - \beta_E = \frac{r}{s} .$$

If $1 - \beta_D$ is the desired power for the test procedure given R_σ, then the procedure may be considered to be reasonably reliable if the discrepency between $1 - \beta_E$ and $1 - \beta_D$ is not significant.

In the analysis used to evaluate a proposed test procedure for a single variance, 500 replications of the simulation experiment will be conducted. For each case for which the proposed test procedures is evaluated one must define the distribution of sample observations X_i, the number of replications, s. The specific distribution of X_i chosen will determine the value of R_σ since one must not only define the family of disributions for X_i but also the parametric values of the distribution. With respect to the test procedure let T_{exp} be the test statistic and T_L and T_U the lower and upper critical limits for the test associated with a level of significance α. For single-tail tests either T_L or T_U is defined but not both. If the test to be evaluated is the two-tail chi-square test then

$$T_{exp} = \frac{(n-1)s^2}{\sigma_0^2} \tag{3.7.1}$$

$$T_L = \chi^2_{\alpha/2}(n-1) \tag{3.7.2}$$

$$T_U = \chi^2_{1-\alpha/2}(n-1). \tag{3.7.3}$$

The desired power of the test, $1 - \beta_D$, is based upon the chi-square test and can be obtained from using equations (3.7.1), (3.7.2) and (3.7.3).

The simulation experiment with respect to the predictability of the classical F test can be conducted by inference.

Example C

The true and desired power of the chi-square test of

$$H_0: \sigma^2 > \sigma_0^2$$

against

$$H_1: \sigma^2 < \sigma_0^2$$

are to determine when $\sigma_0^2 = 2.263$ and the correlation structure of the sample observations are given by

$$\rho_{ij} = \rho^{|i-j|}$$

where $\rho = 0.9$. A lower-tail test is called for

$$T_L = \chi^2_\alpha \, (n-1)$$

and T_{exp} is given by equation (3.7.1). Let $n = 101$, $\alpha = 20\%$ and $s = 500$. The estimated power, $1 - \beta_E$, is to be computed for $R_\sigma = 1.00 + 0.1\,k$ where $k = 0,1,2,\ldots,13$.

Let σ_1^2 be the true variance of the population analyzed. Due to the correlation structure given above,

$$\sigma_1^2 = \frac{1}{1 - \rho^2} \, .$$

Hence

$$R_\sigma = \frac{\sigma_1^2}{\sigma_0^2} = \sigma_0^2 \, (1 - \rho^2).$$

Therefore

$$\rho = \sqrt{1 - \frac{R_\sigma}{\sigma_0^2}} \, .$$

The results of the analysis are presented in Table 7 and shown graphically in Figure 4. For the values of R_σ considered in the experiment, the estimated error in approximating the power of the chi-square test reaches a maximum absolute value of 0.5420 at $R_\sigma = 1.0$. The results suggest that violation of the independence assumption would lead to a significant loss of predictability if the chi-square test is applied indiscriminately.

Example D

Similar to the preceding Example C, this example examines the impact of violation of the independence assumption for two variances.

The true and desired power of the F test of

Table 7

Estimated True Power ($1 - \hat{\beta}_T$) and Desired Power ($1 - \beta_D$) for Single Variance for Correlated Normal Random Variables When the Classical Chi-Square Test is Applied.

(L,U = 99% confidence limits, $E_{1-\beta} = [(1 - \hat{\beta}_T) - (1 - \beta_D)]$)

$\frac{1}{R_\sigma}\left(\frac{\sigma_1^2}{\sigma_0^2}\right)$	L	$1 - \hat{\beta}_T$	U	$1 - \beta_D$	$E_{1-\beta}$
1.0	0.6920	0.7420	0.7878	0.20	0.5420
1.1	0.7047	0.7500	0.7953	0.2828	0.4672
1.2	0.7324	0.7760	0.8196	0.4868	0.2892
1.3	0.7885	0.8280	0.8675	0.7009	0.1271
1.4	0.8237	0.8600	0.8963	0.8567	0.0033
1.5	0.8550	0.8880	0.9210	0.9430	-0.055
1.6	0.9057	0.9320	0.9583	0.9809	-0.0489
1.7	0.9395	0.9600	0.9805	0.9946	-0.0346
1.8	0.9547	0.9720	0.9893	0.9987	-0.0267
1.9	0.9653	0.9800	0.9947	0.9997	-0.0197
2.0	0.9670	0.9880	0.9990	0.9999	-0.0119
2.1	0.9780	0.9960	0.9990	1.0	-0.004
2.2	0.9800	0.9980	0.9990	1.0	-0.002
2.3	1.0	1.0	1.0	1.0	0.

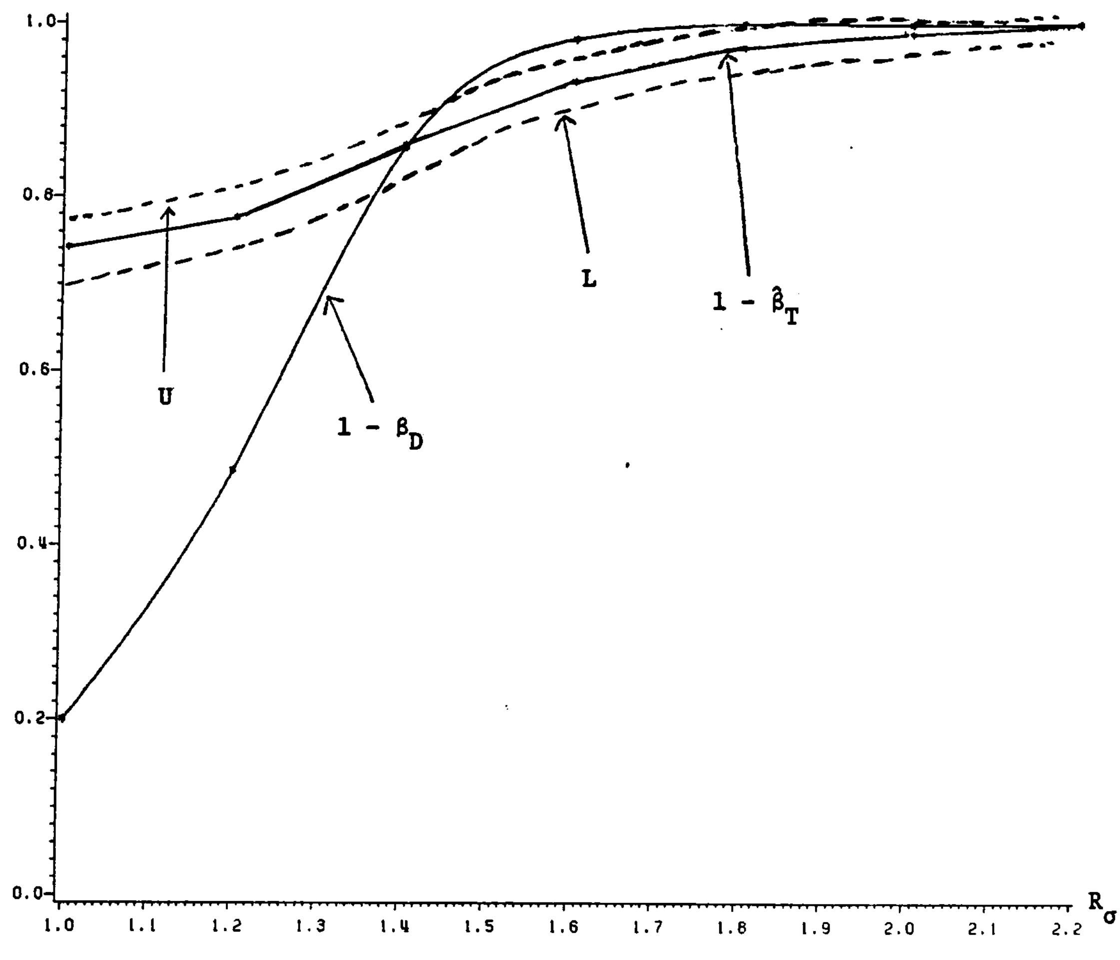

Figure 4. Lower (L) and Upper (U) 99% Confidence Limits for the Estimated True Power ($1 - \hat{\beta}_T$) and the Desired Power ($1 - \beta_D$) of the Chi-Square Test When the Sample Observations are Correlated and Normally Distributed.

$$H_0: \sigma_1^2 > \sigma_2^2$$

against

$$H_1: \sigma_1^2 < \sigma_1^2$$

are to be determined when $\sigma_1^2 = 2.263$ and the correlation structure in each population is given by

$$\rho_{ij} = \rho^{|i-j|}$$

where $\rho = 0.9$.

The lower-tail F test is called for $n_1 = n_2 = 51$, $\alpha = 20\%$.

The results of the analysis are presented in Table 8 and shown graphically in Figure 5. In comparison with the results reported in Table 7, the power of the test would be reduced if two variances are to be compared.

Example E

The true and desired power of the chi-square test of

$$H_0: \sigma^2 < \sigma_0^2$$

against

$$H_1: \sigma^2 > \sigma_0^2$$

are to be determined when the sample observations are exponentially distributed and $\sigma_0^2 = 1$. An upper-tail test is called for

$$T_U = \chi^2_{1-\alpha}(n-1)$$

and T_{exp} is given by equation (3.7.1). Let $n = 101$, $\alpha = 10\%$ and $s = 500$. The estimated power, $1 - \beta_E$, is to be computed for

Table 8

Estimated True Power $(1 - \hat{\beta}_T)$ and Desired Power $(1 - \beta_D)$ for Two Variances for Correlated Normal Random Variables When the Classical F Test is Applied.

(L,U = 99% confidence limits, $E_{1-\beta} = [(1 - \hat{\beta}_T) - (1 - \beta_D)]$)

$R_\sigma(\frac{\sigma_2^2}{\sigma_1^2})$	L	$1 - \hat{\beta}_T$	U	$1 - \beta_D$	$E_{1-\beta}$
1.0	0.5304	0.5820	0.6336	0.20	0.3820
1.4	0.5590	0.6100	0.6610	0.4669	0.1431
1.8	0.5795	0.6300	0.6805	0.7823	-0.1533
2.2	0.5980	0.6480	0.6980	0.9310	-0.2830
2.6	0.6646	0.7120	0.7594	0.9804	-0.2684
3.0	0.7196	0.7640	0.8084	1.0	-0.2486
3.4	0.7744	0.8152	0.8556	1.0	-.0236
3.8	0.8326	0.8678	0.9034	1.0	-0.1848
4.2	0.8847	0.9140	0.9433	1.0	-0.1322
4.6	0.9420	0.9620	0.9820	1.0	-0.038
5.0	1.0	1.0	1.0	1.0	0

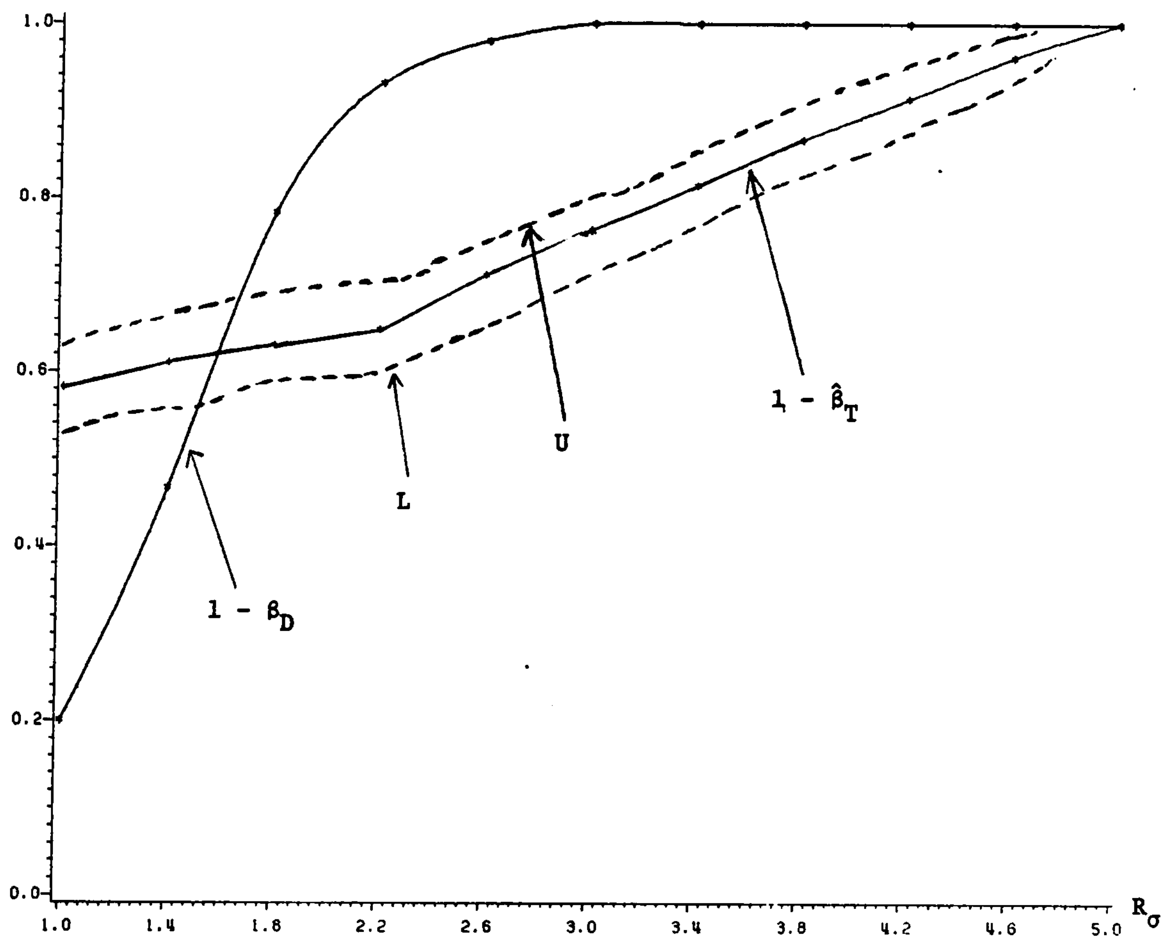

Figure 5. Lower (L) and Upper (U) 99% Confidence Limits for the Estimated True Power ($1 - \hat{\beta}_T$) and the Desired Power ($1 - \beta_D$) of the F Test When the Sample Observations are Correlated and Normally Distributed.

$R_\sigma = 1.00 + 0.1k$ where $k = 0,1,2,\cdots,11$.

Let σ_1^2 be the true variance of the population analyzed. Since the sample observations are exponentially distributed

$$\sigma_1^2 = \frac{1}{\lambda^2}$$

where λ is the parameter of the exponential distribution. Hence

$$\lambda = \frac{1}{\sigma_0^2\sqrt{R_\sigma}}$$

$$= \frac{1}{\sqrt{R_\sigma}}$$

since $\sigma_0^2 = 1$.

The results of the analysis are summarized in Table 9 and shown graphically in Figure 6. For the values of R_σ considered in the experiment, the estimated error in approximating the power of the chi-square test reaches a maximum absolute value of 0.176 at $R_\sigma = 1.40$ and the error in estimating the level of confidence is 0.158 ($R_\sigma = 1.00$). These results would suggest that the analyst should be skeptical about the use of the chi-square test for nonnormal data.

Example F

Similar to the preceding Example D, this example generates uncorrelated exponential random variables in the case of two populations with 51 observations for each population. The test of hypothesis considered is H_0: $\sigma_1^2 < \sigma_2^2$, against H_1: $\sigma_1^2 > \sigma_2^2$, where the estimated and desired powers of the F test are to be determined.

The estimated power is obtained from a simulation experiment of 500 replications. The desired power is derived directly from the F

Table 9

Estimated True Power ($1 - \hat{\beta}_T$) and Desired Power ($1 - \beta_D$) for Single Variance for Uncorrelated Exponential Random Variables When the Classical Chi-Square Test is Applied.

(L,U = 99% confidence limits, $E_{1-\beta} = [(1 - \hat{\beta}_T) - (1 - \beta_D)]$)

$R_\sigma(\frac{\sigma_1^2}{\sigma_0^2})$	L	$1 - \hat{\beta}_T$	U	$1 - \beta_D$	$E_{1-\beta}$
1.0	0.2047	0.250	0.2953	0.100	0.158
1.1	0.2790	0.326	0.3760	0.281	0.045
1.2	0.4078	0.460	0.5122	0.517	-0.057
1.3	0.5280	0.580	0.630	0.725	-0.125
1.4	0.6395	0.688	0.7365	0.864	-0.176
1.5	0.7230	0.770	0.8110	1.940	-0.170
1.6	0.8030	0.846	0.8810	1.976	-.0130
1.7	0.8450	0.884	0.9150	0.991	-0.107
1.8	0.8755	0.906	0.9365	0.997	-0.091
1.9	0.9160	0.946	0.9670	0.999	-0.053
2.0	0.9470	0.966	0.9850	1.000	-0.034
2.1	0.9470	0.972	0.9870	1.000	-0.028
2.1	0.9796	0.990	1.0	1.000	-0.010
2.2	1.0	1.0	1.0	1.000	0.

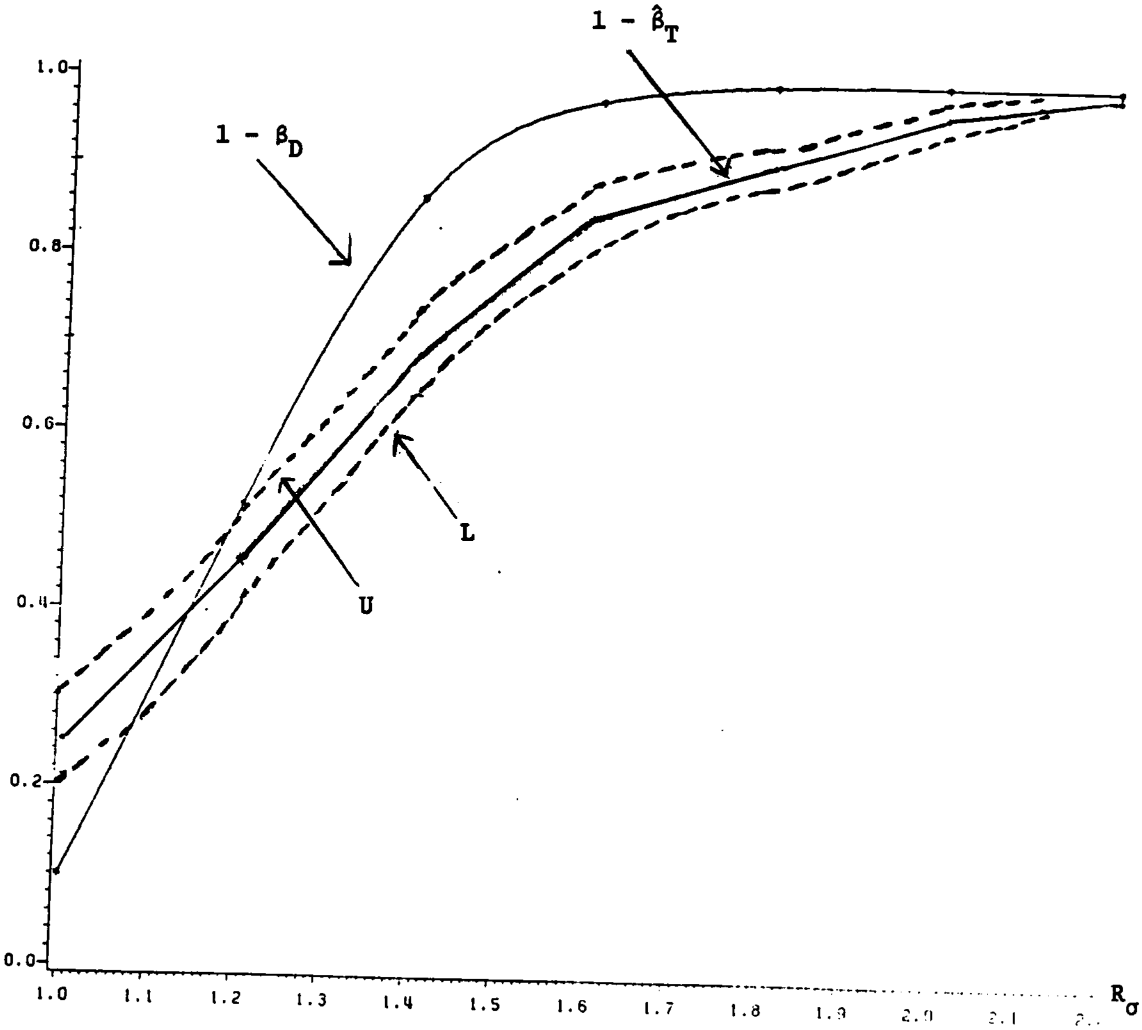

Figure 6. Lower (L) and Upper (U) 99% Confidence Limits for the Estimated True Power ($1 - \hat{\beta}_T$) and the Desired Power ($1 - \beta_D$) of the Chi-Square Test When the Sample Observations are Uncorrelated and Exponentially Distributed.

test. The results are presented in Table 10 and shown graphically in Figure 7. As Table 10 indicated, violation of the normality assumption induces a greater loss of predictability if two variances are to be compared.

<u>Example G</u>

The true and desired powers of the chi-square test of

$$H_0: \ \sigma^2 < \sigma_0^2$$

against

$$H_1: \ \sigma^2 > \sigma_0^2$$

are to be determined when the sample observations are generated by the simple moving average scheme given by

$$X_i = \sum_{t=1}^{k} Z_{t+i-1}$$

where $i = 1,2,\ldots,n$,

$$Z_\ell \sim N(\mu_Z, \sigma_Z^2),$$

and Z_ℓ and Z_n are independent for $\ell \neq h$. The random variable X_i then has mean μ given by

$$\mu = k \ \mu_Z$$

variance σ^2 given by

$$\sigma^2 = k \ \sigma_Z^2.$$

The coefficient of correlation between X_i and X_j ρ_{ij} is given by

$$\rho_{ij} = \begin{cases} 1 - \dfrac{|i-j|}{K}, & 0 < |i-j| < k \\ 0 & |i-j| > k \end{cases} .$$

Table 10

Estimated True Power ($1 - \hat{\beta}_T$) and Desired Power ($1 - \beta_D$) for Two Variances for Uncorrelated Exponential Random Variables When the Classical F Test is Applied.

(L,U = 99% confidence limits, $E_{1-\beta} = [(1 - \hat{\beta}_T) - (1 - \beta_D)]$)

$R_\sigma(\frac{\sigma_2^2}{\sigma_1^2})$	L	$1 - \hat{\beta}_T$	U	$1 - \beta_D$	$E_{1-\beta}$
1.0	0.3170	0.3675	0.4180	0.1000	0.2675
1.2	0.3684	0.4200	0.4716	0.1680	0.2520
1.4	0.4277	0.4800	0.5323	0.3223	0.1577
1.6	0.4828	0.5350	0.5872	0.5011	0.0339
1.8	0.5538	0.6050	0.6562	0.6612	-0.0562
2.0	0.6001	0.6500	0.6999	1.7839	-0.1339
2.2	0.6862	0.7325	0.7788	1.8685	-.01360
2.4	0.7313	0.7750	0.8187	0.9226	-0.1262
2.6	0.7989	0.8375	0.8761	0.9555	-0.1060
2.8	0.8376	0.8725	0.9074	0.9748	-0.1265
3.0	0.8516	0.8850	0.9184	0.9859	-0.1012
3.2	0.8829	0.9125	0.9421	0.9920	-0.0820
3.4	0.9549	0.9721	0.9893	1.0	-0.0795
3.6	0.9779	0.9889	0.999	1.0	-0.0280
3.8	1.0	1.0	1.0	1.0	0.

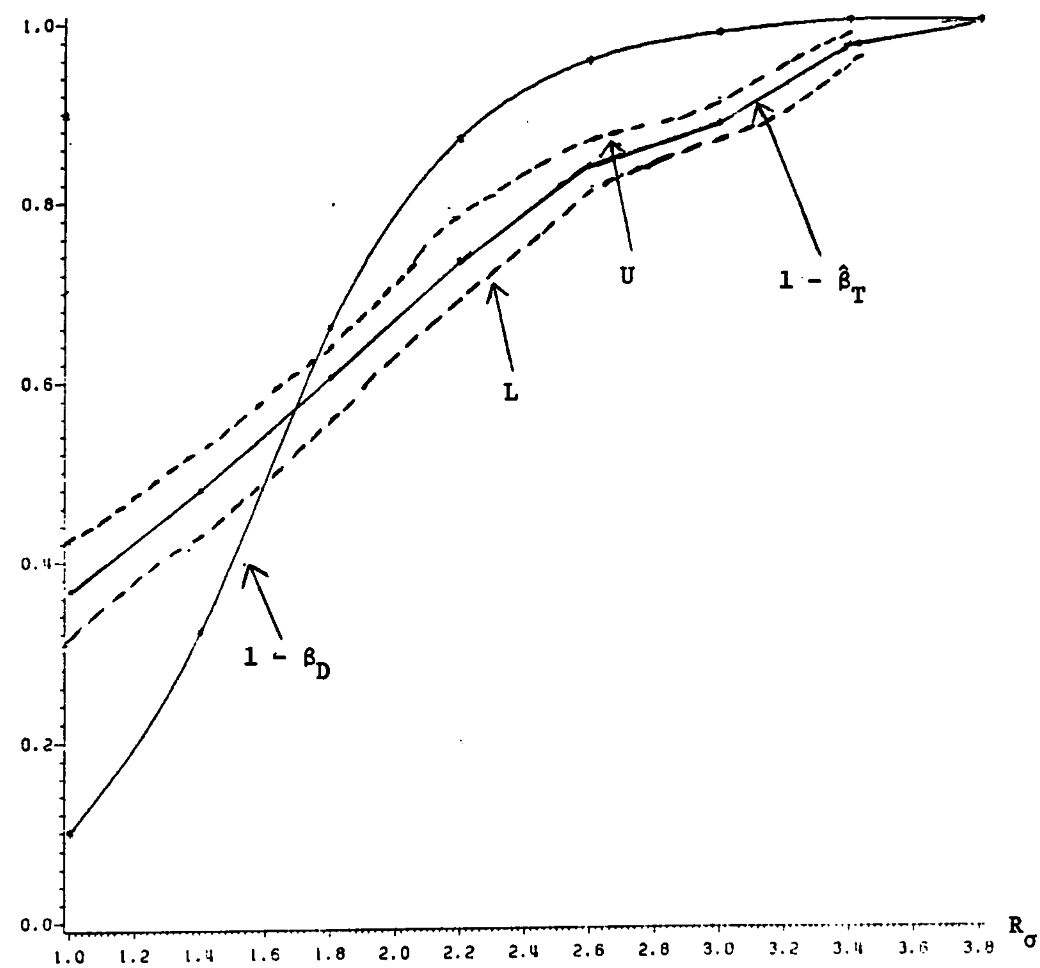

Figure 7. Lower (L) and Upper (U) 99% Confidence Limits for the Estimated True Power ($1 - \hat{\beta}_T$) and the Desired Power ($1 - \beta_D$) of the F Test When the Sample Observations are Uncorrelated and Exponentially Distributed.

If Z_ℓ is gamma distributed such that

$$Z_\ell \sim G(\lambda, \frac{1}{k}),$$

then the distribution of X_i is given by

$$X_i \sim G(\lambda, 1).$$

That is an exponential distribution with parameter λ and the variance of X_i, σ^2, is given by

$$\sigma^2 = \frac{1}{\lambda^2}.$$

Now consider $\sigma_0^2 = 25$, $k = 4$, $n = 360$ and $R_\sigma = 1.0 + 0.1I$ where $I = 0,1,2,\ldots,11$. The results of the analysis are presented in Table 11 and shown graphically in Figure 8. As the results indicate, the error in approximating the power of the chi-square test reaches a maximum value of 0.5920 at $R_\sigma = 1.0$. This suggests that data correlation together with departure from normality can constitute a major source of error in applying chi-square test procedure for variances.

Example H

Similar to the preceding example G, this example generates correlated and exponentially distributed random variables in the case of two populations with $k = 4$ and $n = 180$ for each population. The test of hypothesis considered is H_0: $\sigma_1^2 < \sigma_2^2$, against H_1: $\sigma_1^2 > \sigma_2^2$, where the estimated and desired powers of the F test are to be determined.

As in the case of single variances, the estimated power is

Table 11

Estimated True Power $(1 - \hat{\beta}_T)$ and Desired Power $(1 - \beta_D)$ for Single Variance for Correlated Exponential Random Variables When the Classical Chi-Square Test is Applied.

(L,U = 99% confidence limits, $E_{1-\beta} = [(1 - \hat{\beta}_T) - (1 - \beta_D)]$)

$R_\sigma(\frac{\sigma_1^2}{\sigma_0^2})$	L	$1 - \hat{\beta}_T$	U	$1 - \beta_D$	$E_{1-\beta}$
1.0	0.7026	0.7480	0.7934	0.2000	0.5480
1.1	0.7047	0.7500	0.7953	0.6703	0.0797
1.2	0.7196	0.7640	0.8084	0.9425	-0.1785
1.3	0.7281	0.7720	0.8159	0.9952	-0.2232
1.4	0.7366	0.7800	0.8233	0.9998	-0.2198
1.5	0.7733	0.8140	0.8547	1.0	-0.1860
1.6	0.8393	0.8740	0.9087	1.0	-.01260
1.7	0.8709	0.9020	0.9331	1.0	-0.0980
1.8	0.9321	0.9540	0.9759	1.0	-0.0460
1.9	0.9472	0.9661	0.9850	1.0	-0.0390
2.0	0.9681	0.9820	0.9959	1.0	-0.0180
2.1	1.0	1.0	1.0	1.0	0.

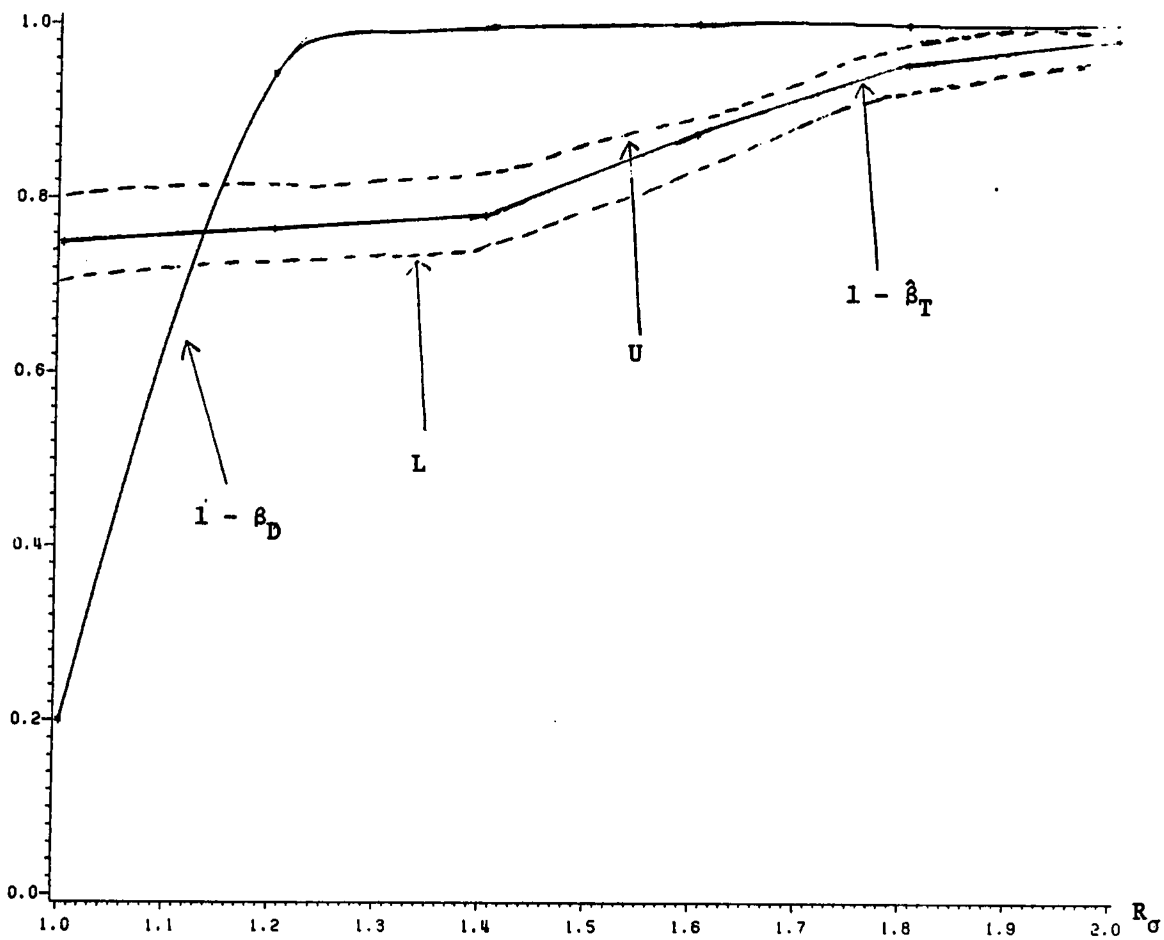

Figure 8. Lower (L) and Upper (U) 99% Confidence Limits for the Estimated True Power ($1 - \hat{\beta}_T$) and the Desired Power ($1 - \beta_D$) of the Chi-Square Test When the Sample Observations are Correlated and Exponentially Distributed.

obtained from a simulation of 500 replications. The desired power is derived directly from the F test. The results are presented in Table 12 and shown graphically in Figure 9. In comparison with the results reported in Example G, the resulting predictability in this example performs worse than the case of single variance.

3.8 CONCLUDING REMARKS

In the case of means, the predictability of the inferential procedures for means is affected by the bias in estimating the variance of the sample mean. The estimator bias is a result of data correlation when the convential statistical inferential procedures for means are applied with ignorance of the data correlation. The magnitude of the estimator bias determines the precision with which the inferential procedures predict the level of confidence and the power of hypothesis tests.

In the case of variances, the predictability of the classical procedures are affected by both the data correlation and the nonnormality. In the situation that two variances are concerned, the resulting predictability is worse than the case of single variance if the assumption of normality is violated.

Table 12

Estimated True Power $(1 - \hat{\beta}_T)$ and Desired Power $(1 - \beta_D)$ for Two Variances for Correlated Exponential Random Variables When the Classical F Test is Applied.

(L,U = 99% confidence limits, $E_{1-\beta} = [(1 - \hat{\beta}_T) - (1 - \beta_D)]$)

$R_\sigma(\frac{\sigma_2^2}{\sigma_1^2})$	L	$1 - \hat{\beta}_T$	U	$1 - \beta_D$	$E_{1-\beta}$
1.0	0.7495	0.7920	0.8345	0.2000	0.5920
1.1	0.7733	0.8140	0.8547	0.5242	0.3080
1.2	0.7831	0.8230	0.8629	0.8118	0.0112
1.3	0.8319	0.8674	0.9029	0.9496	-0.0822
1.4	0.8573	0.8900	0.9227	0.9904	-0.1004
1.5	0.8824	0.9120	0.9416	1.9986	-0.0866
1.6	0.8916	0.9204	0.9484	1.9998	-0.0794
1.7	0.9033	0.9300	0.9567	1.0	-0.0700
1.8	0.9163	0.9405	0.9657	1.0	-0.0595
1.9	0.9296	0.9520	0.9744	1.0	-0.0480
2.0	0.9395	0.9600	0.9805	1.0	-0.0400
2.1	0.9695	0.9830	0.9965	1.0	-0.0170
2.2	1.0	1.0	1.0	1.0	0.

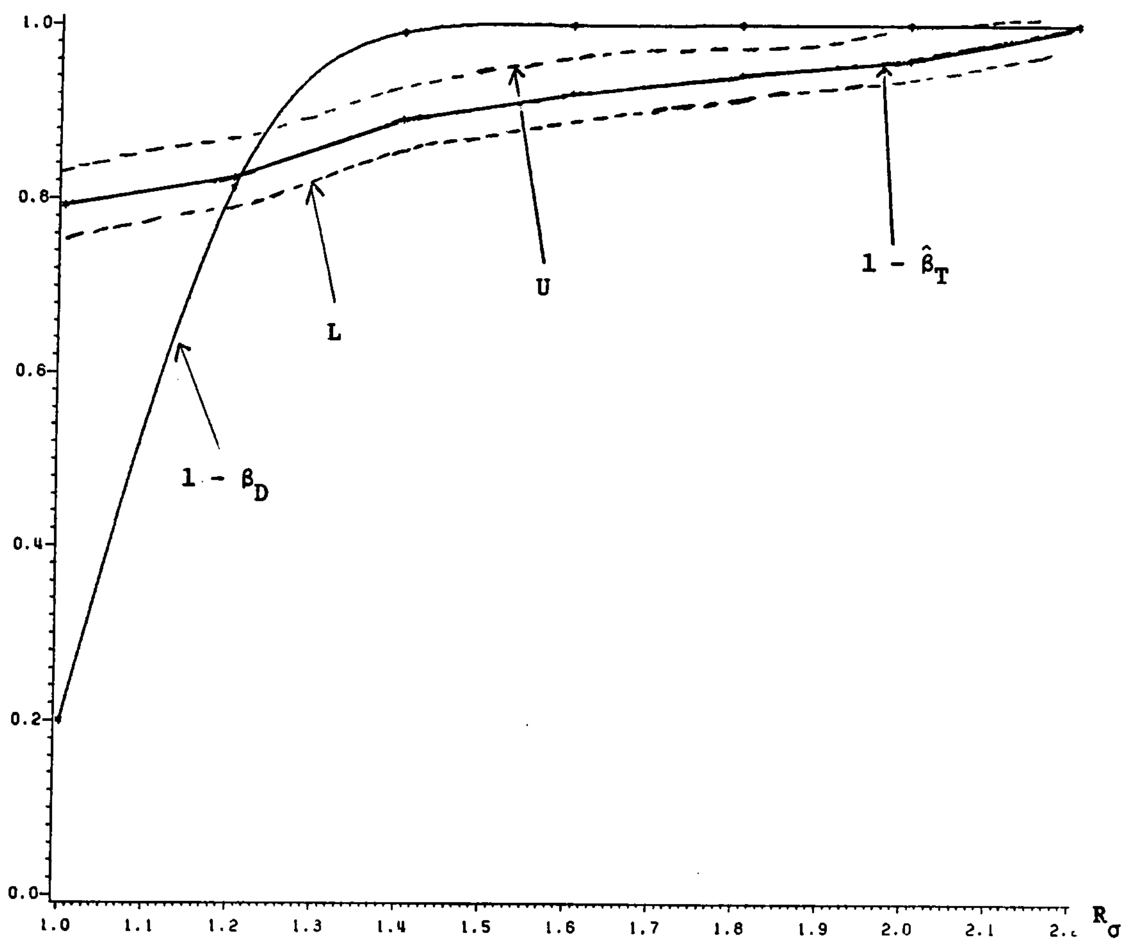

Figure 9. Lower (L) and Upper (U) 99% Confidence Limits for the Estimated True Power ($1 - \hat{\beta}_T$) and the Desired Power ($1 - \beta_D$) of the F Test When the Sample Observations are Correlated and Exponentially Distributed.

CHAPTER IV

THE METHOD OF SEQUENTIAL SYSTEMATIC SAMPLING

4.1 INTRODUCTION

The method of sequential systematic sampling is presented in this chapter. The method is intended to prove the predictability of the statistical analysis when the data are correlated and normally distributed. The impact of violating the assumption of normality will be demonstrated later in this research.

For inferences concerning means, the method of sequential systematic sampling is an alternative to the method of batch means. Both methods attempt to reduce bias in estimating $\sigma_{\bar{X}}^2$ thereby improving the predictability of the analysis of which they are a part. The method of sequential systematic sampling also allows direct estimation of σ^2. This will lead to associated inferential procedures for variances while the application of the method of batch means could only provide a lower bound of the desired level of confidence.

Both methods approach the problem of estimating $\sigma_{\bar{X}}^2$ by grouping sample observations. In the method of sequential systematic sampling, the method of regrouping may be considered as the converse of that used in the method of batch means. This usually leads to a more substantial reduction of bias in estimating $\sigma_{\bar{X}}^2$ and σ^2 therefore yielding more efficient techniques of analysis.

In applying the method of sequential systematic sampling, the sample observations $X_1, X_2, \cdots, X_n$ are sampled at intervals of length k. That is the first systematically drawn sample consists of observations $X_1, X_{k+1}, \cdots, X_{(m-1)k+1}$, the second consists of $X_2, X_{k+2}, \cdots, X_{(m-1)k+2}$, and so forth where $n = mk$. The intent of the method is to use the observations arranged in this sequence of systematically drawn samples in a manner which provides an estimator of $\sigma_{\bar{x}}^2$ and σ^2 with insignificant bias.

Note that

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} + \frac{2}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{i,j}. \tag{4.1.1}$$

The method of sequential systematic sampling attempts to estimate $\sigma_{\bar{x}}^2$ by estimating σ^2 and $\sum_i \sum_j \sigma_{ij}$, where it is assumed that all X_i are identically distributed with common mean μ, and variance σ^2, and σ_{ij} is the covariance of X_i and X_j. This method assumes that k is chosen sufficiently large that the terms $\sigma_{i,i+h}$ may be ignored for $h > k$. This value of $(k - 1)$ is referred to as the order of serial dependence of the process $\{X_i\}$ and observations of the process at lag k or greater can be considered as uncorrelated. While the true order of serial dependence for a stochastic process will usually be unknown, the problem of determining the proper value of k is left until Chapter VIII.

4.2 THE ESTIMATOR OF THE VARIANCE OF THE POPULATION

The estimator for σ^2 is obtained by pooling the conventional estimators of variance for the k systematic samples. That is, if s_i^2 is the estimate of variance for i^{th} systematic sample then

$$s_i^2 = \frac{1}{m-1} \sum_{\ell=1}^{m} [X_{(\ell-1)k+i} - \bar{X}_i]^2, \quad i = 1,2,\cdots,k$$

where $\bar{X}_i$ is the sample mean for the ith systematic sample and

$$\bar{X}_i = \frac{1}{m} \sum_{\ell=1}^{m} X_{(\ell-1)k+i}, \quad i = 1,2,\cdots,k.$$

The expected value of s_i^2 is

$$E(s_i^2) = \sigma^2 - \frac{2}{m(m-1)} \sum_{\ell=1}^{m-1} \sum_{h=\ell+1}^{m} \sigma_{(\ell-1)k+i,(h-1)k+i}, \quad i = 1,2,\cdots,k$$

and the estimate of σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{k} \sum_{i=1}^{k} s_i^2$$

with expectation given by

$$E(\hat{\sigma}^2) = \sigma^2 - \frac{2}{km(m-1)} \sum_{i=1}^{k} \sum_{\ell=1}^{m-1} \sum_{h=\ell+1}^{m} \sigma_{(\ell-1)k+i,(h-1)k+i}. \quad (4.2.2)$$

To simplify the expression for $E(\hat{\sigma}^2)$, note that there are (mk−jk) covariance terms at lag jk in equation (4.2.2). Let $\bar{\sigma}_{jk}$ be the mean covariance at lag jk. Then

$$\bar{\sigma}_{jk} = \frac{1}{k(m-1)} \sum_{i=1}^{k} \sum_{\ell=1}^{m-j} \sigma_{(\ell-1)k+i,(\ell+j-1)k+i}, \quad j = 1,2,\cdots,m-1$$

and

$$E(\hat{\sigma}^2) = \sigma^2 - \frac{2}{m(m-1)} \sum_{j=1}^{m-1} (m-j)\bar{\sigma}_{jk}.$$

The bias, $B_{\hat{\sigma}^2,\sigma^2}$, of $\hat{\sigma}^2$ as an estimate of σ^2 is given by

$$B_{\hat{\sigma}^2,\sigma^2} = \frac{2}{m(m-1)} \sum_{j=1}^{m-1} (m-j)\bar{\sigma}_{jk}. \quad (4.2.3)$$

The summation in equation (4.2.3) includes (m−1) mean covariance terms at lag k, (m−2) mean covariance terms at lag 2k,$\cdots$, and one mean covariance term for observations at lag (m−1)k.

4.3 THE ESTIMATOR FOR SUMMATION OF COVARIANCE

To estimate $\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{ij}$, one can express the double summation in equation (4.1.1) in terms of the mean covariance at lag h, $\bar{\sigma}_h$. That is,

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{i,j} = \sum_{j=1}^{n-1} \sigma_{j,i+1} + \sum_{j=1}^{n-2} \sigma_{j,j+1} + \cdots + \sigma_{1,n}$$
$$= \sum_{h=1}^{n-1} (n-h)\bar{\sigma}_h \tag{4.3.1}$$

where, as before,

$$\bar{\sigma}_h = \frac{1}{n-h} \sum_{j=1}^{n-h} \sigma_{j,j+h}.$$

Given the assumption that the mean covariance $\bar{\sigma}_h$ may be ignored for $h > k$, the equation (4.3.1) can be further reduced to the following:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{ij} = \sum_{h=1}^{k-1} (n-h)\bar{\sigma}_h. \tag{4.3.2}$$

Typically one might estimate $\bar{\sigma}_h$ by the conventional estimator

$$\hat{\sigma}_h = \frac{1}{n} \sum_{j=1}^{n-h} (X_j - \bar{X})(X_{j+h} - \bar{X}).$$

But this estimator may be significantly biased (see equation (2.3.2.6) of Chapter II) for correlated sample observations.

An alternative to estimate $\bar{\sigma}_h$ is to break the original sequence of sample observations into m successive sequences each of size k. If Y_{ij} is the j^{th} observation in the i^{th} sequence, then

$$Y_{i,j} = X_{(i-1)k+j}.$$

If observations at lag k or greater may be consider uncorrelated, then $Y_{i,j}$ and $Y_{k,\ell}$ are uncorrelated for $k > i+2$ for all j and ℓ. In terms of the observations $Y_{i,j}$, $\bar{\sigma}_h$ may be expressed (the derivation is given in Appendix B) as

$$\bar{\sigma}_h = \frac{1}{mk-h}\left[\sum_{i=1}^{m}\sum_{j=1}^{k-h} \mathrm{Cov}(Y_{i,j}\ Y_{i,j+h}) + \sum_{i=1}^{m-1}\sum_{j=1}^{h} \mathrm{Cov}(Y_{i,k-h+j},\ Y_{i+1,j})\right] \tag{4.3.3}$$

where the first term in equation (4.3.3) accounts for the covariance at lag h within each sequence and the second accounts for the covariance at lag k between successive sequences.

Let σ_{Wh} be the sum of covariance terms at lag h within the m sequences and σ_{Bh} be the sum of covariance terms at lag h between sequences.

Then

$$\begin{aligned}\sigma_{Wh} &= \sum_{i=1}^{m}\sum_{j=1}^{k-h} \sigma_{(i-1)k+j,(i-1)k+j+h} \\ &= \sum_{i=1}^{m/2}\sum_{j=1}^{k-h} [\mathrm{Cov}(Y_{2i-1,j},Y_{2i-1,j+h}) \\ &\quad + \mathrm{Cov}(Y_{2i,j},Y_{2i,j+h})]\end{aligned}$$

for even m and

$$\begin{aligned}\sigma_{Wh} &= \sum_{i=1}^{(m+1)/2}\sum_{j=1}^{k-h} \mathrm{Cov}(Y_{2i-1,j},Y_{2i-1,j+h}) \\ &\quad + \sum_{i=1}^{(m-1)/2}\sum_{j=1}^{k-h} \mathrm{Cov}(Y_{2i,j},Y_{2i,j+h})\end{aligned}$$

for odd m. Now let the sequence of sample means $\bar{Y}_{\cdot\ell,0}$ and $\bar{Y}_{\cdot\ell,e}$ be given by

$$\bar{Y}_{\cdot\ell,0} = \begin{cases} \dfrac{2}{m}\displaystyle\sum_{i=1}^{m/2} Y_{2i-1,\ell}, & \text{even } m \\[2ex] \dfrac{2}{m+1}\displaystyle\sum_{i=1}^{(m+1)/2} Y_{2i-1,\ell}, & \text{odd } m \end{cases}$$

$$\bar{Y}_{\cdot \ell, e} = \begin{cases} \frac{2}{m} \sum_{i=1}^{m/2} Y_{2i,\ell}, & \text{even } m \\ \frac{2}{m+1} \sum_{i=1}^{(m-1)/2} Y_{2i,\ell}, & \text{odd } m. \end{cases}$$

The estimate of σ_{Wh} is chosen as

$$\hat{\sigma}_{Wh} = \begin{cases} \frac{m}{m-2} \sum_{j=1}^{k-h} \sum_{i=1}^{m/2} [(Y_{2i-1,j} - \bar{Y}_{\cdot j,0})(Y_{2i-1,j+h} - \bar{Y}_{\cdot j+h,0}) \\ \quad + (Y_{2i,j} - \bar{Y}_{\cdot j,e})(Y_{2i,j+h} - \bar{Y}_{\cdot j+h,e})], \text{ even } m, \\ \sum_{j=1}^{k-h} [\frac{m+1}{m-1} \sum_{i=1}^{(m+1)/2} (Y_{2i-1,j} - \bar{Y}_{\cdot j,0})(Y_{2i-1,j+h} - \bar{Y}_{\cdot j+h,0}) \\ \quad + \frac{m-1}{m-3} \sum_{i=1}^{(m-1)/2} (Y_{2i,j} - \bar{Y}_{\cdot j,e})(Y_{2i,j+h} - \bar{Y}_{\cdot j+h,e})], \text{ odd } m. \end{cases} \tag{4.3.4}$$

The reason that odd and even sequences are treated separately is related to the assumption that observations at lag k or greater may be treated as uncorrelated. The observations comprising $\bar{Y}_{\cdot \ell,0}$ have minimum lag 2k as is the case for $\bar{Y}_{\cdot \ell,e}$. Thus the crossproduct terms in equation (4.3.4) such as $(Y_{2i-1,j}\ \bar{Y}_{\cdot j+h,0})$ are affected only by correlation at lag h. This is evident by taking the expected value of equation (4.3.4) which is given by

$$E(\hat{\sigma}_{Wh}) = \sum_{i=1}^{m} \sum_{j=1}^{k-h} \sigma_{(i-1)k+j,(i-1)k+j+h} + \varepsilon_W(k,h), \tag{4.3.5}$$

where $\varepsilon_W(k,h)$ includes covariance terms at lag k of greater for even m and is given by

$$\varepsilon_W(k,h) = -\frac{4}{m-2} \sum_{j=1}^{k-h} \sum_{i=1}^{m/2} \sum_{\substack{\ell=1 \\ \ell \neq i}}^{m/2} \sigma_{2(i-1)k+j+h,2(\ell-1)k+j} \tag{4.3.6}$$

and by

$$\varepsilon_W(k,h) = -2\sum_{j=1}^{k-h}\frac{1}{m-1}\sum_{i=1}^{(m+1)/2}\sum_{\substack{\ell=1\\ \ell\neq i}}^{(m+1)/2}\sigma_{2(i-1)k+j+h,2(\ell-1)k+j}$$

$$+\frac{1}{m-3}\sum_{i=1}^{(m-1)/2}\sum_{\substack{\ell=1\\ \ell\neq i}}^{(m-1)/2}\sigma_{(2i-1)k+j+h,(2\ell-1)k+j} \qquad (4.3.7)$$

for odd m.

One can express equations (4.3.6) and (4.3.7) in terms of the lag covariance. Then equation (4.3.6) becomes

$$\varepsilon_W(k,h) = \frac{-4(k-h)}{m-2}\sum_{i=1}^{\frac{m}{2}-1}\left(\frac{m}{2}-i\right)(\sigma_{2ik+h}+\sigma_{2ik-h}) \qquad (4.3.8)$$

and equation (4.3.7) becomes

$$\varepsilon_W(k,h) = \begin{array}{l} -2(k-h)\left[\frac{1}{m-1}\sum_{i=1}^{\frac{m+1}{2}-1}\left(\frac{m+1}{2}-i\right)(\sigma_{2ik+h}+\sigma_{2ik-h})\right. \\ \left.+\frac{1}{m-3}\sum_{i=1}^{\frac{m-1}{2}-1}\left(\frac{m-1}{2}-i\right)(\sigma_{2ik+h}+\sigma_{2ik-h})\right]. \end{array} \qquad (4.3.9)$$

Remember that σ_{Bh} is given by

$$\sigma_{Bh} = \sum_{i=1}^{m-1}\sum_{j=1}^{h}\sigma_{ik+j-h,ik+j}$$

$$= \sum_{j=1}^{h}\sum_{i=1}^{m/2}Cov(Y_{2i-1,k+j-h},Y_{2i,j})$$

$$+\sum_{i=1}^{(m-2)/2}Cov(Y_{2i,k+j-h},Y_{2i+1,j})$$

for even m and

$$\sigma_{Bh} = \sum_{j=1}^{h}\sum_{i=1}^{(m-1)/2}Cov(Y_{2i-1,k+j-h},Y_{2i,j})$$

$$+ Cov(Y_{2i,k+j-h},Y_{2i+1,j})$$

for odd m. The estimate of σ_{Bh} is given by

$$\sigma_{Bh} = \begin{cases} \sum_{j=1}^{h} [\frac{m}{m-2} \sum_{i=1}^{m/2} (Y_{2i-1,k+j-h} - \bar{Y}'_{\cdot k+j-h,0})(Y_{2i,j} - \bar{Y}'_{\cdot j,e}) \\ \quad + \frac{m-2}{m-4} \sum_{i=1}^{(m-2)/2} (Y_{2i,k+h-h} - \bar{Y}'_{\cdot k+j-h,e})(Y_{2i+1,j} - \bar{Y}'_{\cdot j,0})], & \text{even } m \\ \frac{m-1}{m-3} \sum_{j=1}^{h} \sum_{i=1}^{(m-1)/2} [(Y_{2i-1,k+j-h} - \bar{Y}'_{\cdot k+h-h,0})(Y_{2i,j} - \bar{Y}'_{\cdot j,e})] \\ \quad + (Y_{2i,k+j-h} - \bar{Y}'_{\cdot k+j-h,e})(Y_{2i+1,j} - \bar{Y}'_{\cdot j,0})], & \text{odd } m \end{cases}$$

where

$$\bar{Y}'_{\cdot k+j-h,0} = \begin{cases} \frac{2}{m} \sum_{i=1}^{m/2} Y_{2i-1,k+j-h}, & \text{even } m \\ \frac{2}{m-1} \sum_{i=1}^{(m-1)/2} Y_{2i-1,k+j-h}, & \text{odd } m \end{cases}$$

$$\bar{Y}'_{\cdot j,e} = \begin{cases} \frac{2}{m} \sum_{i=1}^{m/2} Y_{2i,j}, & \text{even } m \\ \frac{2}{m-1} \sum_{i=1}^{(m-1)/2} Y_{2i,j}, & \text{odd } m \end{cases}$$

$$\bar{Y}'_{\cdot k+j-h,e} = \begin{cases} \frac{2}{m-2} \sum_{i=1}^{(m-2)/2} Y_{2i,k+j-h}, & \text{even } m \\ \frac{2}{m-1} \sum_{i=1}^{(m-1)/2} Y_{2i,k+j-h}, & \text{odd } m \end{cases}$$

$$\bar{Y}'_{\cdot j,0} = \begin{cases} \frac{2}{m-2} \sum_{i=1}^{(m-2)/2} Y_{2i+1,j}, & \text{even } m \\ \frac{2}{m-1} \sum_{i=1}^{(m-1)/2} Y_{2i+1,j}, & \text{odd } m. \end{cases}$$

The expected value of $\hat{\sigma}_{Bh}$ is given by

$$E(\hat{\sigma}_{Bh}) = \sum_{i=1}^{m-1} \sum_{j=1}^{h} \sigma_{ik+j-h,ik+j} + \varepsilon_B(k,h)$$

where $\varepsilon_B(k,h)$ includes covariances at lag k or greater. For even m $\varepsilon_B(k,h)$ is given by

$$\varepsilon_B(k,h) = -2\sum_{j=1}^{h}\Big[\frac{1}{m-2}\sum_{i=1}^{m/2}\sum_{\substack{\ell=1\\ \ell\neq i}}^{m/2}\sigma_{(2i-1)k+j-h,(2\ell-1)k+j}$$

$$+\frac{1}{m-4}\sum_{i=1}^{(m-2)/2}\sum_{\substack{\ell=1\\ \ell\neq i}}^{(m-2)/2}\sigma_{2ik+j,2\ell k+j-h}\Big] \qquad (4.3.10)$$

and by

$$\varepsilon_B(k,h) = -\frac{2}{m-3}\sum_{j=1}^{h}\sum_{i=1}^{(m-1)/2}\sum_{\substack{\ell=1\\ \ell\neq i}}^{(m-1)/2}[\sigma_{(2i-1)k+j-h,(2\ell-1)k+j}$$

$$+\sigma_{2ik+j,2\ell k+j-h}] \qquad (4.3.11)$$

for odd m.

The equations (4.3.10) and (4.3.11) can also be simplified if all covariance terms σ_{ij} are expressed by the lag h covariances σ_h where $h = |i-j|$. From equation (4.3.10), one obtains

$$\varepsilon_B(k,h) = -2h\Big[\frac{1}{m-2}\sum_{i=1}^{\frac{m}{2}-1}(\frac{m}{2}-i)(\sigma_{2ik+h}+\sigma_{2ik-h})$$

$$+\frac{1}{m-4}\sum_{i=1}^{\frac{m}{2}-1}(\frac{m-2}{2}-i)(\sigma_{2ik+h}+\sigma_{2ik-h})\Big]. \qquad (4.3.12)$$

From equation (4.3.11), one obtains

$$\varepsilon_B(k,h) = \frac{-4h}{m-3}\sum_{i=1}^{\frac{m-1}{2}-1}(\frac{m-1}{2}-i)(\sigma_{2ik+h}+\sigma_{2ik-h}). \qquad (4.3.13)$$

Finally the estimate of $\bar{\sigma}_h$, $\hat{\bar{\sigma}}_h$ is given by

$$\hat{\bar{\sigma}}_h = \frac{1}{mk-h}(\hat{\sigma}_{Wh}+\hat{\sigma}_{Bh}). \qquad (4.3.14)$$

Since covariance terms at lag k or greater are assumed to be negligable, the left hand side of equation (4.3.2) can now be estimated by $\sum_{h=1}^{k-1}(n-h)\hat{\bar{\sigma}}_h$.

4.4 THE ESTIMATOR FOR THE VARIANCE OF THE SAMPLE MEAN

Recall that the variance of the sample mean is

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} + \frac{2}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{i,j}.$$

Given the estimates for σ^2 and $\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{ij}$, the estimate for $\sigma_{\bar{x}}^2$ can be expressed as

$$\hat{\sigma}_{\bar{x}_k}^2 = \frac{1}{mk} [\hat{\sigma}^2 + \frac{2}{mk} \sum_{h=1}^{k-1} (\hat{\sigma}_{Wh} + \hat{\sigma}_{Bh})]. \tag{4.4.1}$$

The expected value of $\hat{\sigma}_{\bar{x}_k}^2$ can be shown to be

$$E(\hat{\sigma}_{\bar{x}_k}^2) = \frac{1}{mk} [\sigma^2 + \frac{2}{mk} \sum_{h=1}^{k-1} (\sigma_{Wh} + \sigma_{Bh}) - \frac{2}{m(m-1)} \sum_{i=1}^{m-1} (m-i)\bar{\sigma}_{ik} + \frac{2}{mk} \sum_{h=1}^{k-1} (\varepsilon_W(k,h) + \varepsilon_B(k,h))]. \tag{4.4.2}$$

The bias in $\hat{\sigma}_{\bar{x}_k}^2$ as an estimate of $\sigma_{\bar{x}}^2$, $B_{\hat{\sigma}_{\bar{x}_k}^2, \sigma_{\bar{x}}^2}$, is given by

$$B_{\hat{\sigma}_{\bar{x}_k}^2, \sigma_{\bar{x}}^2} = -\frac{2}{(mk)^2} [\frac{k}{m-1} \sum_{i=1}^{m-1} (m-i)\bar{\sigma}_{ik} - \sum_{h=1}^{k-1} (\varepsilon_W(k,h) + \varepsilon_B(k,h)) + \sum_{h=k}^{mk-1} (mk-h)\bar{\sigma}_h]. \tag{4.4.3}$$

Replacing the covariance terms σ_{ij} in equation (4.4.3) by the lag covariances and then dividing by σ^2 lead to the following:

$$
\begin{aligned}
B_{\hat{\sigma}^2_{\bar{x}_k}, \sigma^2_{\bar{x}}} = - \frac{-2\sigma^2}{(mk)^2} \Big\{ & \frac{k}{m-1} \sum_{i=1}^{m-1} (m-i)\rho_{ik} \\
& - \sum_{h=1}^{k-1} \Big[\frac{-4k+2h}{m-2} \sum_{i=1}^{\frac{m}{2}-1} \left(\frac{m}{2} - i\right)(\rho_{2ik+h} + \rho_{2ik-h}) \\
& + \frac{-2h}{m-4} \sum_{i=1}^{\frac{m-2}{2}-1} \left(\frac{m-2}{2} - i\right)(\rho_{2ik+h} + \rho_{2ik-h}) \Big] \\
& + \sum_{h=k}^{mk-1} (mk-h)\rho_h \Big\}, \text{ for even } m
\end{aligned}
\tag{4.4.4}
$$

and

$$
\begin{aligned}
B_{\hat{\sigma}^2_{\bar{x}_k}, \sigma^2_{\bar{x}}} = - \frac{-2\sigma^2}{(mk)^2} \Big\{ & \frac{k}{m-1} \sum_{i=1}^{m-1} (m-i)\rho_{ik} \\
& - \sum_{h=1}^{k-1} \Big[\frac{-2k+2h}{m-1} \sum_{i=1}^{\frac{m+1}{2}-1} \left(\frac{m+1}{2} - i\right)(\sigma_{2ik+h} + \sigma_{2ik-h}) \\
& + \frac{-2h}{m-3} \sum_{i=1}^{\frac{m-1}{2}-1} \left(\frac{m-1}{2} - i\right)(\sigma_{2ik+h} + \sigma_{2ik-h}) \Big] \\
& + \sum_{h=k}^{mk-1} (mk-h)\rho_h \Big\}, \text{ for odd } m\,.
\end{aligned}
\tag{4.4.5}
$$

The bias in either equation (4.4.4) or equation (4.4.5) consists of covariance terms with lags no less than k and these terms would be zero if the autocorrelation dies out at lag k.

4.5 CONSIDERATION OF THE ORDER OF SERIAL DEPENDENCE

Application of the method of sequential systematic sampling requires definition of the order of serial dependence as well as the number of systematic samples. The magnitudes of k and m, of course, depend upon the ultimate objective of the analysis to which the method of sequential systematic sampling is applied. Hence, proper definition of k and m requires that the analyst trace the effect of k and m to possible outcomes of the experiment and the effect of those outcomes on appropriate measures of performance.

The presence of data correlation, if ignored, may lead to significant bias in estimating $\sigma_{\bar{x}}^2$ and σ^2. In turn, the resulting bias for each case leads to error in predicting either the level of confidence in confidence interval analysis or the behavior of the power function in hypothesis testing. If there is good reason for specifying a particular level of confidence or, in the case of a hypothesis test, a particular level of significance and power associated with a case of the alternative hypothesis, there is equally good reason to make sure that the procedure applied will meet those critera, at least in an approximate sense.

It is evident from equation (4.2.3) and (4.4.3) that the bias in $\hat{\sigma}^2$ and $\hat{\sigma}_{\bar{x}_k}^2$ as an estimator of σ^2 and $\sigma_{\bar{x}}^2$ are a function of both k and m respectively. Thus by increasing either the order of serial dependence, k, the number of systematic samples, m, or both, the error in predicting the probabilistic performance of an inferential procedure applied may be reduced thus increasing the predictability of the procedure.

In a practical setting the purpose in applying the method of sequential systematic sampling will usually be to improve the predictability of the procedure ultimately applied. The level of predictability required, or conversely the tolerable level of uncertainty in predicting the performance of the procedure, is application dependent. The outcome of decisions which may be reached or influenced as a consequence of interpretation of the results coming from the application of the method of sequential systematic sampling defines the level of uncertainty which may be tolerated.

4.6 CONCLUDING REMARK

The method of sequential systematic sampling attempts to estimate σ^2 and $\sigma_{\bar{x}}^2$ by assuming the recognition of the order of serial dependence, k-1. With this assumption this method can collect essentially uncorrelated observations to form the estimators of σ^2 and $\sigma_{\bar{x}}^2$. The estimator bias in each case consists of covariances with lag greater than or equal to k. The bias would be eliminated if the correlation dies out at lag k.

CHAPTER V

EVALUATION OF THE METHOD OF SEQUENTIAL SYSTEMATIC SAMPLING

This chapter compares the method of batch means, sequential systematic sampling, and the classical method. The purpose of the comparison is to examine how the method of sequential systematic sampling performs relative to the method of batch means and the classical method in overcoming the problems associated with correlated data when inferential procedures for means are concerned. The classical method refers to case where s^2/n is used as the estimator of $\sigma^2_{\bar{x}}$. The performance of each method is first evaluated based on the bias of estimator of $\sigma^2_{\bar{x}}$ and then the predictability of the inferential procedure applied.

5.1 COMPARISON OF ESTIMATOR BIAS

In each of the three methods to be compared in this chapter, the bias in estimating $\sigma^2_{\bar{x}}$ is developed for observations generated from AR(1) and simple moving average models.

Consider an AR(1) model

$$X_t - \phi X_{t-1} = a_t$$

where $\{a_t\}$ is a sequence of i.i.d. N(0,1) random variables. The variance of X_t, σ^2, is given by

$$\sigma^2 = \frac{1}{1-\phi^2} .$$

The autocorrelation function ρ_k is given by

$$\rho_k = \phi^k, \text{ where } k > 0.$$

The general formula of the resulting bias in estimating $\sigma_{\bar{x}}^2$ through the method of batch means is presented in equation (C.2.10) of Appendix C. Using equation (C.2.10), the bias, $B_{\frac{s_y^2}{m},\sigma_{\bar{x}}^2}$, is

$$B_{\frac{s_y^2}{m},\sigma_{\bar{x}}^2} = \frac{-2\sigma^2}{(m-1)mk^2} \sum_{j=1}^{m-1} (m-j)[k\phi^{jk} + \sum_{i=1}^{k-1} (k-i)(\phi^{jk-i} + \phi^{jk+i})]. \qquad (5.1.2)$$

Then equation (5.1.2) leads to

$$B_{\frac{s_y^2}{m},\sigma_{\bar{x}}^2} = \frac{-2\sigma^2}{(m-1)mk^2}\left[\frac{m\phi^k[((\phi^{k1})^{n-1})-1]}{\phi^k-1} - \frac{(m-1)(\phi^k)^{m+1} - m(\phi^k)^m + \phi^k}{(\phi^k-1)^2}\right]$$

$$\cdot \left[k + \frac{k\phi(\phi^{k-1}-1)}{\phi-1} - (k-1)\phi^{k+1} - \frac{k\phi^k+\phi}{(\phi-1)} + \frac{(-k+1)\phi k^{-k-1} + k\phi^{-k} - \phi^{-1}}{(\phi^{-1}-1)^2}\right]. \qquad (5.1.3)$$

The general formula of the bias in estimating $\sigma_{\bar{x}}^2$ by the estimator s^2/n is presented in equation (C.2.11) of Appendix C. Using equation (C.2.11), the bias, $B_{\frac{s^2}{n},\sigma_{\bar{x}}^2}$, is given by

$$B_{\frac{s^2}{n},\sigma_{\bar{x}}^2} = \frac{-2\sigma^2}{(mk-1)k}\left[\frac{k\phi(\phi^{k-1}-1)}{\phi-1} - \frac{(k-1)\phi^{k+1}-k\phi^k+\phi}{\phi-1}\right] + V \qquad (5.1.4)$$

where V is the right hand side of equation (5.1.3). From equation (4.4.4), assuming that m is even, one obtains the bias for sequential systematic sampling given by

$$
\begin{aligned}
B_{\hat{\sigma}^2_{\bar{X}_k},\,\sigma^2_{\bar{X}}} = \frac{-2\sigma^2}{(mk)^2}\Big\{\frac{k}{m-1}\sum_{i=1}^{m-1}(m-i)\phi^{ik} \\
&- \sum_{h=1}^{k}\Big[\frac{-4k+2h}{m-2}\sum_{i=1}^{\frac{m}{2}-1}\left(\frac{m}{2}-i\right)(\phi^{2ik+h}+\phi^{2ik-h}) \\
&+ \frac{-2h}{m-4}\sum_{i=1}^{\frac{m}{2}-1}\left(\frac{m-2}{2}-i\right)(\phi^{2ik+h}+\phi^{2ik-h}) \\
&+ \sum_{h=k}^{mk-1}(mk-h)\phi^{h}\Big\}. \qquad (5.1.5)
\end{aligned}
$$

From equation (5.1.5), it follows that

$$
\begin{aligned}
B_{\hat{\sigma}^2_{\bar{X}_k},\,\sigma^2_{\bar{X}}} &= -\frac{2\sigma^2}{(mk)^2}\Big\{\frac{k}{m-1}\Big[\frac{m\phi^k[(\phi^k)^{n-1}-1]}{\phi^k-1} - \frac{(m-1)(\phi^k)^{m+1}-m(\phi^k)^m+\phi^k}{(\phi^k-1)^2}\Big] \\
&- \frac{1}{m-2}\Big\{\Big[(-4k)\frac{\phi^k-\phi}{\phi-1} + (2)\frac{(k-1)\phi^{k+1}-k\phi^k+\phi}{(\phi-1)^2} + (-4k)\frac{\phi^{-k}-\phi^{-1}}{\phi^{-1}-1} \\
&+ (-2)\frac{(1-k)\phi^{-k-1}+k\phi^{-k}-\phi^{-1}}{(\phi^{-1}-1)^2}\Big] \cdot \Big[\left(\frac{m}{2}\right)\frac{\phi^{mk}-\phi^{2k}}{\phi^{2k}-1} - \frac{\left(\frac{m}{2}-1\right)(\phi^{2k})^{\frac{m}{2}+1} - \frac{m}{2}(\phi^{2k})^{\frac{m}{2}}+\phi^{2k}}{(\phi^{2k}-1)^2}\Big] \\
&+ \frac{1}{m-4}\Big[(-2)\frac{(k-1)\phi^{k+1}-k\phi^k+\phi}{(\phi-1)^2} + (2)\frac{(1-k)\phi^{-k-1}+k\phi^{-k}-\phi^{-1}}{(\phi^{-1}-1)^2}\Big] \\
&\cdot \Big[\left(\frac{m-2}{2}\right)\frac{\phi^{(m-2)k}-\phi^{2k}}{\phi^{2k}-1} - \frac{\left(\frac{m-2}{2}-1\right)(\phi^{2k})^{\frac{m-2}{2}+1} - \left(\frac{m-2}{2}\right)(\phi^{2k})^{\frac{m-2}{2}} + \phi^{2k}}{(\phi^{2k}-1)^2}\Big]\Big\} \\
&+ \Big[(mk)\frac{\phi^{mk}-\phi}{\phi-1} - \frac{(mk-1)\phi^{mk+1}-mk\phi^{mk}+\phi}{(\phi-1)^2} \\
&- (mk)\frac{\phi^k-\phi}{\phi-1} + \frac{(k-1)\phi^{k+1}-k\phi^k+\phi}{(\phi-1)^2}\Big]\Big\}. \qquad (5.1.6)
\end{aligned}
$$

Given equations (5.1.3), (5.1.4 and (5.1.6), the bias in estimating $\sigma_{\bar{X}}^2$ by each of the three methods were presented in rather complicated forms. To compare the estimator bias for each method, a computer program is developed for the computation of equations (5.1.3), (5.1.4) and (5.1.6). The results are summarized in Table 13.

Three conclusions based upon the results are stated as follows:

1. Both the estimators of $\sigma_{\bar{X}}^2$ yielded by the method of batch mean and sequential systematic sampling show considerable improvement over the classical estimator of $\sigma_{\bar{X}}^2$ where the data correlation is ignored.
2. For an AR(1) model with the positive AR(1) parameter, the estimator of the method of sequential systematic sampling yields the least bias for $k_S = k_B$, $m_S = m_B$, where the subscripts S and B denote the method of sequential systematic sampling and the method of batch means respectively. The bias from each method decreases as k_S or k_B increase.
3. For an AR(1) model with the negative AR(1) parameter, the method of sequential systematic sampling may not yield the least bias but the difference between bias in each method may not be significant in consideration of its absolute magnitude.

Now, consider a simple moving average model

$$X_t = \sum_{i=1}^{k_t} a_{t+i-1} \tag{5.1.7}$$

where $\{a_t\}$ is a sequence of i.i.d. N(0,1) random variables. Then the mean and variance of X_t are given by

$$\mu = 0$$

Table 13

Estimator bias of $\sigma_{\bar{x}}^2$ in AR(1) Models with Parameter ϕ.

n = mk, Scale of bias: σ^2.

ϕ	k	m	Assumed Independence	Batch Means	Sequential Systematic Sampling
.5	10	20	−0.0099	−0.0020	0.0000
.5	20	10	−0.0099	−0.0010	0.0000
.8	10	20	−0.0392	−0.0177	−0.0052
.8	20	10	−0.0392	−0.0099	−0.0006
.9	10	20	−0.0859	−0.0570	−0.0323
.9	20	10	−0.0859	−0.0389	−0.0112
−.5	10	20	0.0033	0.0002	0.0000
−.5	20	10	0.0033	0.0001	0.0000
−.8	10	20	0.0045	0.0002	−0.0006
−.8	20	10	0.0045	0.0001	−0.0001
−.9	10	20	0.0047	0.0002	−0.0019
−.9	20	10	0.0047	0.0001	−0.0007

$$\sigma^2 = k_t.$$

The covariance σ_{ij} is given by

$$\sigma_{i,j} = \begin{cases} k_t - |i-j|, & 0 < |i-j| < k_t \\ 0, & |i-j| > k_t \end{cases}.$$

The correlation coefficient ρ_{ij} can be obtained by

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma^2}$$

$$= \begin{cases} 1 - \frac{|i-j|}{k_t}, & 0 < |i-j| < k_t \\ 0 & |i-j| > k_t \end{cases}. \qquad (5.1.8)$$

In applying the method of batch means, if the batch size k is less than k_t, the development of resulting bias $B_{\frac{s_y^2}{m}, \sigma_{\bar{x}}^2}$ is a time-consuming job in the case of simple moving average models. For example, to determine the domains of i and j simultaneously for the term ρ_{jk+i} in equation (C.2.10), the following inequalities have to be solved,

$$jk - i < k_t - 1$$

$$1 < i < k - 1$$

$$1 < j < m - 1$$

where i and j are integers. Given i one can determine the range of j but that would be a formidable task if a large domain for i alone is involved. However, for $k > k_t$, it is not difficult to develop an algebraic expression for the bias. Since $\rho_{jk} = 0$, $\rho_{jk+i} = 0$, the only nonzero term left in equation (C.2.10) is ρ_{jk-i} where $j = 1$. This is obvious because $\rho_{2k-i} = 0$ as a result of $2k-i > k > k_t$ for $i = 1,\cdots,k-1$. Hence the domain of j is 1 and the domain of i is $k - k_t + 1 < i < k-1$. Thus, for $k > k_t$,

$$
\begin{aligned}
B_{\frac{s_y^2}{m},\sigma_{\bar{x}}^2} &= \frac{-2\sigma^2}{(m-1)mk^2}(m-1)\sum_{i=k-k_t+1}^{k-1}(k-i)\rho_{k-i} \\
&= \frac{-2\sigma^2}{mk^2}\sum_i (k-i)\frac{1-(k-i)}{k_t} \\
&= \frac{-2\sigma^2}{mk^2}\Big\{\frac{2k(k_t-1)-(k-1)k-(k-k_t)(k-k_t+1)}{2} \\
&\quad - \frac{1}{k_t}\Big[kk_t-2k^2+k+\frac{(k-1)(2k-1)-(k-k_t)(k-k_t+1)(2k-2k_t+1)}{6}\Big]\Big\}. \qquad (5.1.9)
\end{aligned}
$$

In the case of applying the method of sequential systematic sampling, the bias $B_{\hat{\sigma}^2_{\bar{x}_k},\sigma^2_{\bar{x}}} = 0$ if $k > k_t$. However, it would be very time-consuming to develop an algebraic expression for $B_{\hat{\sigma}_{\bar{x}_k},\sigma^2_{\bar{x}}}$ when $k < k_t$ following similar reasons in the case of batch means. Hence the comparison of bias between the method of batch means, sequential systematic sampling and the classical method in the case of simple moving average models will be conducted in a simulation study in the next section.

5.2 COMPARISON OF PREDICTABILITY

5.2.1 Introduction

The objective shared by both the methods of batch means and sequential systematic sampling is to provide the analyst with an improved technique for estimating $\sigma^2_{\bar{x}}$ when he must deal with correlated data. The adequacy of the resulting estimator in each method was measured in terms of bias and the improvement can be measured relative to the bias in s^2/n. However, the purpose in seeking an improved estimator of is to reduce the error in probability statements of the means which arises if s^2/n is used as an estimator of $\sigma_{\bar{x}}^2$. Such statements include the quoted level of confidence for a confidence interval and the power of a hypothesis test.

A comparative study for the inferential procedures concerning the application of the classical method, the method of batch means and sequential systematic sampling is given in this section while the classical method refers to the case where s^2/n is used as an estimator for $\sigma_{\bar{x}}^2$. In the last section, the estimator of $\sigma_{\bar{x}}^2$ yielded by the method of sequential systematic sampling, in the case of positive autocorrelation, is less biased than that of the method of batch means in the two cases when sample observations were generated from AR(1) models and simple moving average models where $k > k_t$. Therefore the method of sequential systematic sampling would lead to a more substantial improvement of the predictability of the inferential procedures applied in these two cases. In this section, in addition to the comparison of bias, the purpose here is to compare improvement of the predictability yielded from each method while applying inferential procedures for means. The comparison is presented in the case of simple moving average models. The predictability of each method is measured by the discrepency between the desired and actual levels of confidence, and between the desired and actual powers of hypothesis test.

5.2.2 Methodology

If $\hat{\sigma}^2_{\bar{X}}$ is an estimate of $\sigma^2_{\bar{X}}$ then the confidence limits for a $100(1 - \alpha_D)\%$ confidence interval for the case of single population mean, μ, are given by

$$L,U = \bar{X} \pm \hat{\sigma}_{\bar{X}}\, t_{1 - \frac{\alpha_D}{2}, f} \tag{5.2.2.1}$$

where f is an appropriate degree of freedom. The propriety of application of equation (5.2.2.1) depends upon whether or not $(\bar{X}-\mu)/\hat{\sigma}_{\bar{X}}$ is t distributed with f degrees of freedom. Let t_{exp} be defined as

$$t_{exp} = \frac{\bar{X} - \mu}{\hat{\sigma}_{\bar{X}}}$$

$$= \frac{(\bar{X} - \mu)/\sigma_{\bar{X}}}{\hat{\sigma}_{\bar{X}}/\sigma_{\bar{X}}} .$$

Now, assuming that the sample observations are normally distributed, $(\bar{X} - \mu)/\sigma_{\bar{X}} \sim N(0,1)$. If, in addition $f\ \hat{\sigma}^2_{\bar{X}}/\sigma^2_{\bar{X}} \sim \chi^2(f)$, then, by definition of the t_{exp} random variable, $t_{exp} \sim t(f)$ provided $\bar{X}$ and $\hat{\sigma}^2_{\bar{X}}$ are independent random variables. In the case of uncorrelated data, $\sigma^2_{\bar{X}} = \sigma^2/n$ and $\hat{\sigma}^2_{\bar{X}} = s^2/n$ and because of the normality assumption s^2 and $\bar{X}$ are independent random variables, $\hat{\sigma}^2_{\bar{X}}/\sigma^2_{\bar{X}} = s^2/\sigma^2$ and $(fs^2)/\sigma^2 \sim \chi^2(f)$. However, for the case of correlated data $\hat{\sigma}^2_{\bar{X}}$ is not an unbiased estimator of $\sigma^2_{\bar{X}}$, and $\hat{\sigma}^2_{\bar{X}}$ and $\bar{X}$ may not be independent. Hence $f\hat{\sigma}^2_{\bar{X}}/\sigma^2_{\bar{X}}$ may not be χ^2 distributed with f degrees of freedom and in turn t_{exp} may not be t distributed with f degrees of freedom.

To determine the effectiveness of the method of batch means and sequential systematic sampling in reducing the error in probability statements the problem will be examined empirically. That is, application of each method will be simulated in the case where the sample observations are correlated and normally distributed.

The specific question addressed in the analysis will be the error between the desired level of confidence, $1 - \alpha_D$, and the true level of confidence, $1 - \alpha_T$, and a comparison of that error with the asymptotic error to be expected. In the case of hypothesis testing, the question addressed will be the error between the desired power of the test, $1 - \beta_P$, and the true power, $1 - \beta_T$, and a comparison of that error with the asymptotic error to be expected. A benchmark for comparison is provided by conducting a similar analysis for the case where $\sigma^2_{\bar{x}}$ is estimated by s^2/n.

Consider a sequence of correlated and normally distributed observations $X_1, X_2, \cdots, X_n$. Let L_I, U_I be $1 - \alpha_D$ confidence limits for the case where $\hat{\sigma}^2_{\bar{x}} = s^2/n$. Let L_B, U_B be the associated confidence limits based upon the method of batch means, and let L_S, U_S be the associated confidence limits based upon the method of sequential systematic sampling. Then

$$L_I, U_I = \bar{\bar{X}} \pm \frac{s}{\sqrt{n}}\, t_{1 - \frac{\alpha_D}{2}, n-1} \qquad (5.2.2.2)$$

$$L_B, U_B = \bar{\bar{X}} \pm \frac{s_y}{\sqrt{m}}\, t_{1 - \frac{\alpha_D}{2}, m-1} \qquad (5.2.2.3)$$

$$L_S, U_S = \bar{\bar{X}} \pm \hat{\sigma}_{\bar{x}_k}, t_{1 - \frac{\alpha_D}{2}, n-1} \qquad (5.2.2.4)$$

where n, m, and k are as previously defined. Neither s^2/n nor s_y^2/n is an unbiased estimate of $\sigma^2_{\bar{x}}$. Moreover, $\hat{\sigma}^2_{\bar{x}_k}$ is not an unbiased estimate of $\sigma^2_{\bar{x}}$ unless the order of serial dependence can be truely identified. These facts leads to the conclusion that none of the equations (5.2.2.2), (5.2.2.3) and (5.2.2.4) may yield an exact $100(1 - \alpha_D)\%$ confidence interval.

For sufficiently large n, m, and k, each estimator of $\sigma^2_{\bar{x}}$ may be replaced by its expected value and $t_{1-\frac{\alpha_D}{2},f} \longrightarrow Z_{1-\frac{\alpha_D}{2}}$. Hence the following confidence limits would be obtained,

$$L_I, U_I = \bar{X} \pm \sqrt{E(\frac{s^2}{n})}\, Z_{1-\frac{\alpha_D}{2}} \qquad (5.2.2.5)$$

$$L_B, U_B = \bar{X} \pm \sqrt{E(\frac{s^2_{\bar{y}}}{m})}\, Z_{1-\frac{\alpha_D}{2}} \qquad (5.2.2.6)$$

$$L_S, U_S = \bar{X} \pm \sqrt{E(\hat{\sigma}^2_{\bar{x}})}\, Z_{1-\frac{\alpha_D}{2}}. \qquad (5.2.2.7)$$

If $\hat{\sigma}^2_{\bar{x}}$ is sample estimate of $\sigma^2_{\bar{x}}$ then the true level of confidence may be approximated by

$$1 - \alpha_T \simeq 2\Pr[Z < \frac{\sqrt{E(\hat{\sigma}^2_{\bar{x}})}}{\sigma_{\bar{x}}} Z_{1-\frac{\alpha_D}{2}}] - 1.$$

Let $1 - \alpha_{TI}$, $1 - \alpha_{TB}$ and $1 - \alpha_{TS}$ be the corresponding approximations to the true levels of confidence for the cases where the correlations are ignored and where the methods of batch means and sequential systematic sampling are applied respectively,

$$1 - \alpha_{TI} \simeq 2\Pr[Z < \frac{\sqrt{E(\frac{s^2}{n})}}{\sigma_{\bar{x}}} Z_{1 - \frac{\alpha_D}{2}}] - 1 \qquad (5.2.2.8)$$

$$1 - \alpha_{TB} \simeq 2\Pr[Z < \frac{\sqrt{E(\frac{s_y^2}{m})}}{\sigma_{\bar{x}}} Z_{1 - \frac{\alpha_D}{2}}] - 1 \qquad (5.2.2.9)$$

$$1 - \alpha_{TS} \simeq 2\Pr[Z < \frac{\sqrt{E(\hat{\sigma}^2_{\bar{x}_k})}}{\sigma_{\bar{x}}} Z_{1 - \frac{\alpha_D}{2}}] - 1. \qquad (5.2.2.10)$$

In hypothesis tests, consider the test H_0: $\mu < \mu_0$, against H_1: $\mu > \mu_0$. Applying each method with large samples, the resulting test limits are given by

$$U_I = \mu_0 + \sqrt{E(\frac{s^2}{n})}\, Z_{1 - \frac{\alpha_D}{2}} \qquad (5.2.2.11)$$

$$U_B = \mu_0 + \sqrt{E(\frac{s_y^2}{m})}\, Z_{1 - \frac{\alpha_D}{2}} \qquad (5.2.2.12)$$

$$U_S = \mu_0 + \sqrt{E(\hat{\sigma}^2_{\bar{x}})}\, Z_{1 - \frac{\alpha_D}{2}}. \qquad (5.2.2.13)$$

Let $1 - \beta_I(\mu)$, $1 - \beta_B(\mu)$ and $1 - \beta_S(\mu)$ be the corresponding approximations to the true power of the test in applying each method, and they can be described by

$$1 - \beta_I(\mu) \simeq 1 - \Pr[Z < \frac{\mu_0 - \mu}{\sigma_{\bar{x}}} + \frac{\sqrt{E(\frac{s^2}{n})}}{\sigma_{\bar{x}}} Z_{1 - \frac{\alpha_D}{2}}] \qquad (5.2.2.14)$$

$$1 - \beta_B(\mu) \simeq 1 - \Pr[Z < \frac{\mu_0 - \mu}{\sigma_{\bar{x}}} + \frac{\sqrt{E(\frac{s_y^2}{m})}}{\sigma_{\bar{x}}} Z_{1-\frac{\alpha_D}{2}}] \tag{5.2.2.15}$$

$$1 - \beta_S(\mu) \simeq 1 - \Pr[Z < \frac{\mu_0 - \mu}{\sigma_{\bar{x}}} + \frac{\sqrt{E(\hat{\sigma}^2_{\bar{x}_k})}}{\sigma_{\bar{x}}} Z_{1-\frac{\alpha_D}{2}}]. \tag{5.2.2.16}$$

5.2.3 Simulation Experiment for Single Mean

Let X_i, $i = 1,2,\cdots,n$, be a sequence of normally distributed random variables, each with mean μ_1, variance σ^2 and covariance matrix Σ given by

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{12} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n} & \rho_{2n} & \cdots & 1 \end{bmatrix}$$

where ρ_{ij} is the coefficient of correlation between X_i and X_j for $i \neq j$. A sequence of correlated normal random variables can be generated by the moving average scheme given by

$$X_i = \sum_{t=1}^{k_t} a_{t+i-1} \tag{5.2.3.1}$$

where $\{a_t\}$ is a sequence of i.i.d. N(0,1) random variables. Then the mean and variance of X_i are given by

$$\mu = 0$$

$$\sigma^2 = k_t.$$

The covariance σ_{ij} is given by

$$\sigma_{ij} = \begin{cases} k_t - |i-j|, & 0 < |i-j| < k_t \\ 0, & |i-j| > k_t \end{cases}.$$

The correlation coefficient ρ_{ij} can be obtained by

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma^2}$$

$$= \begin{cases} 1 - \frac{|i-j|}{k_t}, & |i-j| < k_t \\ 0, & |i-j| > k_t \end{cases}.$$

Given the serial correlation coefficient in equation (5.2.3.2), equations (5.2.2.8), (5.2.2.9) and (5.2.2.10) can now be evaluated. The asymptotic true levels of confidence are evaluated from equation (5.2.2.8) where the dependence of the sample observations is ignored, from equation (5.2.2.9) where the method of batch means is applied, and from equation (5.2.2.10) where the method of sequential systematic sampling is applied. The asymptotic true power of the test in equation (5.2.2.14) and (5.2.2.15) and (5.2.2.16) can also be evaluated for each method.

A total of five hundred replications of the simulation experiment are conducted in estimating the true levels of confidence and the true power of the test for each method applied. In each simulation experiment, a simple moving average process with known mean μ is simulated for a given sample size n where $n = mk$. Given one simulated process, take note whether the confidence interval formed by a given method covers μ. Then independently replicate the entire process R

(R = 500) times. From these R replications, for each method, one can compute the proportion s of the R confidence intervals which contain μ. The proportion s is refered as the estimated level of confidence.

In the case of hypothesis testing, a simple moving average process with known mean μ_1 is simulated for a given sample size n where n = mk. Given one simulated process, take note whether the test limits yielded by a given method do not cover μ_0, i.e. the rejection of null hypothesis. Then independently replicate the entire process R times. From these R replications, for each method, one can compute the ratio t as the proportion of rejections of H_0. The proportion t is refered as the estimated power of hypothesis test.

The results of comparing the methods applied are summarized in Table 14, 15, and 16. The results in Table 15 are also shown graphically in Figures 10 and 11. Table 14 summarizes the results with respect to bias reduction and confidence interval analysis for $k_t = 10$ and $20 < k < 4$. Table 15 summarizes the results of confidence interval analysis for $k_t = 5, 10, 20$, and $2 < k < 20$. Table 16 summarizes the results with respect to hypothesis testing.

As Table 14 illustrates, when dependence is ignored and the sample data are treated as a sequence of independent observations, the resulting bias and the level of confidence are disappointing. The method of batch means leads to a true level of confidence which approaches that desired, 80%, as the batch size k increases; however, the method of sequential systematic sampling approaches the desired level of confidence faster than the method of batch means does as the value k increases.

Table 14

Coverage of 80% Confidence Intervals for Single Mean in Simple Moving Average Models with Order k_t.

$m = 50$, $n = mk$, $k_t = 10$, $\alpha_D = .20$, Replication = 500

	k	Assumed Independence	Batch Means	Sequential Systematic Sampling
$\sigma^2_{\bar{x}}$	4	0.4917	0.4917	0.4917
	6	0.3297	0.3297	0.3297
	8	0.2479	0.2479	0.2479
	10	0.1987	0.1987	0.1987
	14	0.1422	0.1422	0.1422
	16	0.1245	0.1245	0.1245
	20	0.0997	0.0997	0.0997
$B_{\hat{\sigma}^2_{\bar{x}}, \sigma^2_{\bar{x}}}$	4	-0.4440	-0.3232	-0.2108
	6	-0.2973	-0.1720	-0.0672
	8	-0.2235	-0.1025	-0.0151
	10	-0.1790	-0.0660	0.
	14	-0.1281	-0.0337	0.
	16	-0.1121	-0.0258	0.
	20	-0.0898	-0.0898	0.
$1 - \alpha_T$	4	0.3109	0.5471	0.6670
	6	0.3124	0.6244	0.7469
	8	0.3131	0.6734	0.7856
(True level	10	0.3135	0.7047	0.80
of	14	0.3140	0.7368	0.80
Confidence)	16	0.3141	0.7459	0.80
	20	0.3144	0.7580	0.80
$1 - \alpha_E$	4	0.3000	0.5140	0.6520
	6	0.3160	0.6000	0.7280
	8	0.3160	0.6880	0.7800
(Estimated	10	0.3200	0.6860	0.7920
level of	14	0.3240	0.7380	0.80
Confidence)	16	0.3320	0.7600	0.80
	20	0.3180	0.7460	0.80

Table 15

Lower and Upper 99% Confidence Limits for the True Level of Confidence for Single Mean in Simple Moving Average Model with Order k_t.

		Assumed Independence			Batch Means			Sequential Systematic Sampling		
k_t	k	$1-\alpha_{TI}$	L_I	U_I	$1-\alpha_{TB}$	L_B	U_B	$1-\alpha_{TS}$	L_S	U_S
5	2	0.4293	0.3589	0.4384	0.5537	0.4823	0.5634	0.6408	0.5788	0.6576
	3	0.4309	0.4031	0.4837	0.6294	0.5981	0.6761	0.7347	0.6924	0.7645
	4	0.4316	0.4188	0.4997	0.6776	0.6563	0.7311	0.7807	0.7644	0.8296
	5	0.4321	0.4001	0.4807	0.7084	0.6893	0.7617	0.8000	0.7803	0.8436
	10	0.4330	0.3932	0.4737	0.7594	0.7330	0.8015	0.8000	0.7666	0.8315
	20	0.4335	0.3805	0.4606	0.7808	0.7445	0.8112	0.8000	0.7602	0.8259
10	2	0.3066	0.2679	0.3426	0.4140	0.3658	0.4455	0.4922	0.4505	0.5316
	3	0.3095	0.2670	0.3416	0.4908	0.4535	0.5346	0.5985	0.5656	0.6450
	4	0.3109	0.2651	0.3396	0.5471	0.4793	0.5604	0.6670	0.6042	0.6819
	5	0.3118	0.2545	0.3282	0.5903	0.5354	0.6156	0.7138	0.6635	0.7378
	6	0.3124	0.2795	0.3550	0.6244	0.5535	0.6332	0.7469	0.6810	0.7541
	7	0.3128	0.2498	0.3231	0.6516	0.6032	0.6810	0.7699	0.7372	0.8052
	8	0.3131	0.2727	0.3478	0.6734	0.6409	0.7167	0.7856	0.7309	0.7996
	9	0.3133	0.2641	0.3385	0.6909	0.6450	0.7205	0.7953	0.7550	0.8212
	10	0.3135	0.2450	0.3179	0.7047	0.6348	0.7109	0.7999	0.7434	0.8109
	14	0.3140	0.2785	0.3539	0.7368	0.6966	0.7684	0.8000	0.7644	0.8296
	16	0.3141	0.2795	0.3550	0.7459	0.7101	0.7807	0.8000	0.7550	0.8212
	20	0.3144	0.2785	0.3539	0.7580	0.7007	0.7722	0.8000	0.7445	0.8118
20	2	0.2120	0.1723	0.2377	0.2937	0.2526	0.3262	0.3548	0.2978	0.3744
	4	0.2189	0.1882	0.2555	0.4113	0.3707	0.4505	0.5209	0.4733	0.5545
	6	0.2211	0.1882	0.2555	0.4887	0.4465	0.5277	0.6165	0.5737	0.6528
	8	0.2223	0.1797	0.2461	0.5454	0.4753	0.5565	0.6792	0.6174	0.6945
	10	0.2230	0.1919	0.2597	0.5889	0.5324	0.6127	0.7223	0.6749	0.7483
	12	0.2234	0.1695	0.2345	0.6231	0.5667	0.6459	0.7527	0.7049	0.7780
	14	0.2237	0.1662	0.2313	0.6505	0.6032	0.6810	0.7738	0.7372	0.8052
	16	0.2240	0.1891	0.2566	0.6724	0.6195	0.6964	0.7880	0.7497	0.8165
	18	0.2242	0.1704	0.2356	0.6899	0.6656	0.7397	0.7964	0.7634	0.8267
	20	0.2243	0.1769	0.2429	0.7038	0.6491	0.7244	0.8000	0.7550	0.8212

Table 16

Comparison of Power Function for Simple Moving Average Models with Order k_t.

$m = 50$, $n = mk$, $k_t = 10$, $\alpha_D = .20$, $d = \mu_1 - \mu_0$, Replication = 500

		Independence Assumed		Batch Means		Sequential Systematic Sampling		Desired Power of the Test
k	d	$1 - \beta_E$	$1 - \beta_T$	$1 - \beta_E$	$1 - \beta_T$	$1 - \beta_E$	$1 - \beta_T$	$1 - \beta_D$
4	0.1	0.7000	0.5064	0.4829	0.4041	0.3485	0.2964	0.2420
	0.2	0.7020	0.5400	0.5049	0.4421	0.3792	0.3321	0.2877
	0.3	0.7080	0.6102	0.5379	0.4981	0.4100	0.3807	0.3372
	0.4	0.7102	0.6500	0.5659	0.5423	0.4396	0.4128	0.3936
	0.5	0.7440	0.6736	0.5889	0.5832	0.4588	0.4590	0.4483
8	0.1	0.6840	0.5240	0.3746	0.3224	0.2267	0.2676	0.2565
	0.2	0.7040	0.5860	0.4000	0.3920	0.2864	0.3540	0.3264
	0.3	0.7230	0.7200	0.4647	0.4762	0.3706	0.4348	0.4048
	0.4	0.7560	0.7462	0.5760	0.5800	0.4864	0.5220	0.5000
	0.5	0.7862	0.7642	0.6340	0.6368	0.5760	0.5714	0.5596
14	0.1	0.6760	0.5500	0.3120	0.3156	0.2736	0.2716	0.2776
	0.2	0.7020	0.6024	0.4310	0.4233	0.3810	0.3900	0.3745
	0.3	0.7430	0.6840	0.5200	0.5199	0.4754	0.4865	0.4761
	0.4	0.7721	0.7403	0.6144	0.6042	0.5920	0.5840	0.5793
	0.5	0.8202	0.8508	0.7180	0.7190	0.6900	0.6880	0.6808

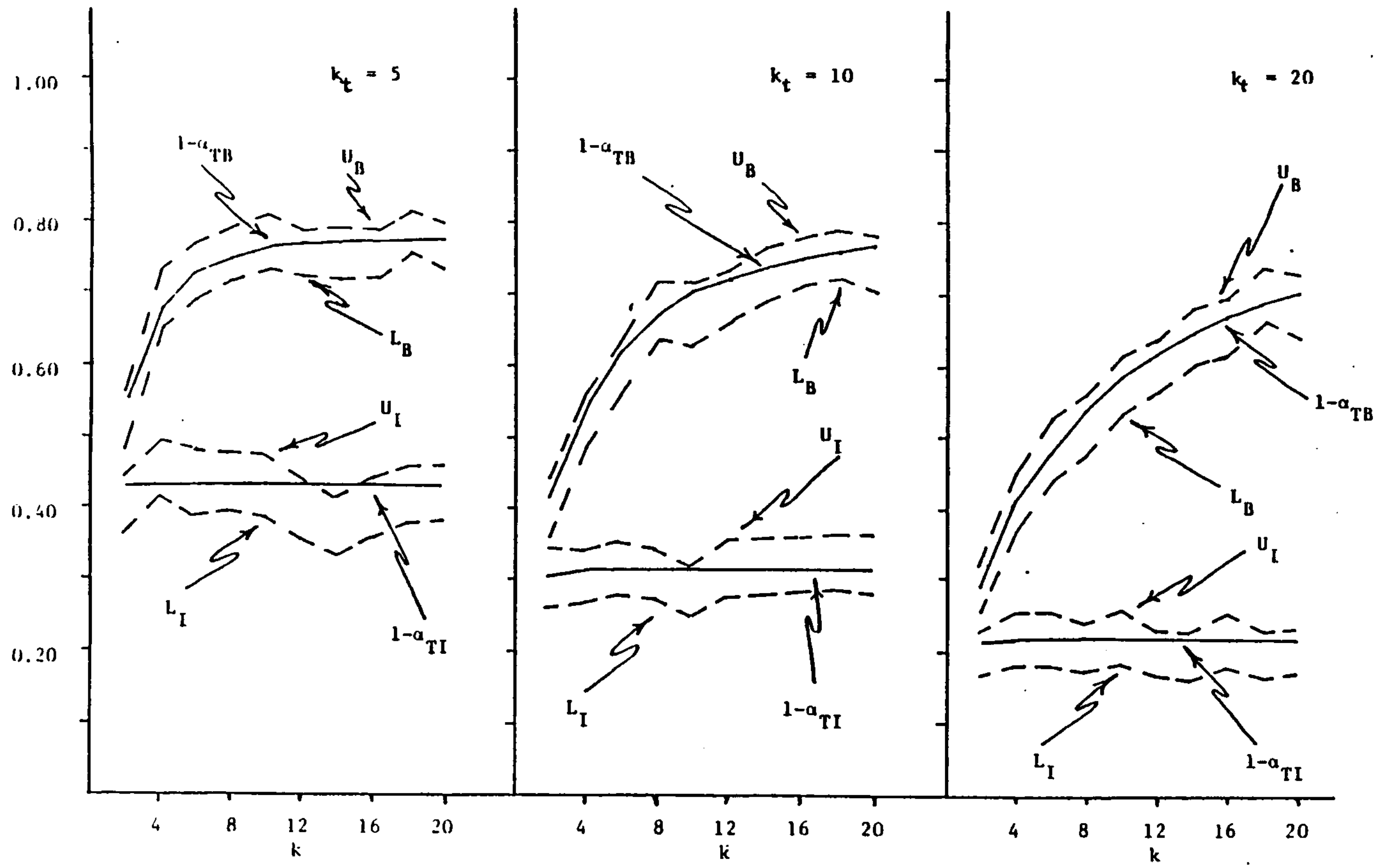

Figure 10

Results of the Analysis for Batch Means Where L_B, U_B = Lower and Upper 99% Confidence Limits for $1 - \alpha_{TB}$, L_I, U_I = Lower and Upper 99% Confidence Limits for $1 - \alpha_{TI}$, k = Batch Size for the Method of Batch Means, m = Number of Batches (50).

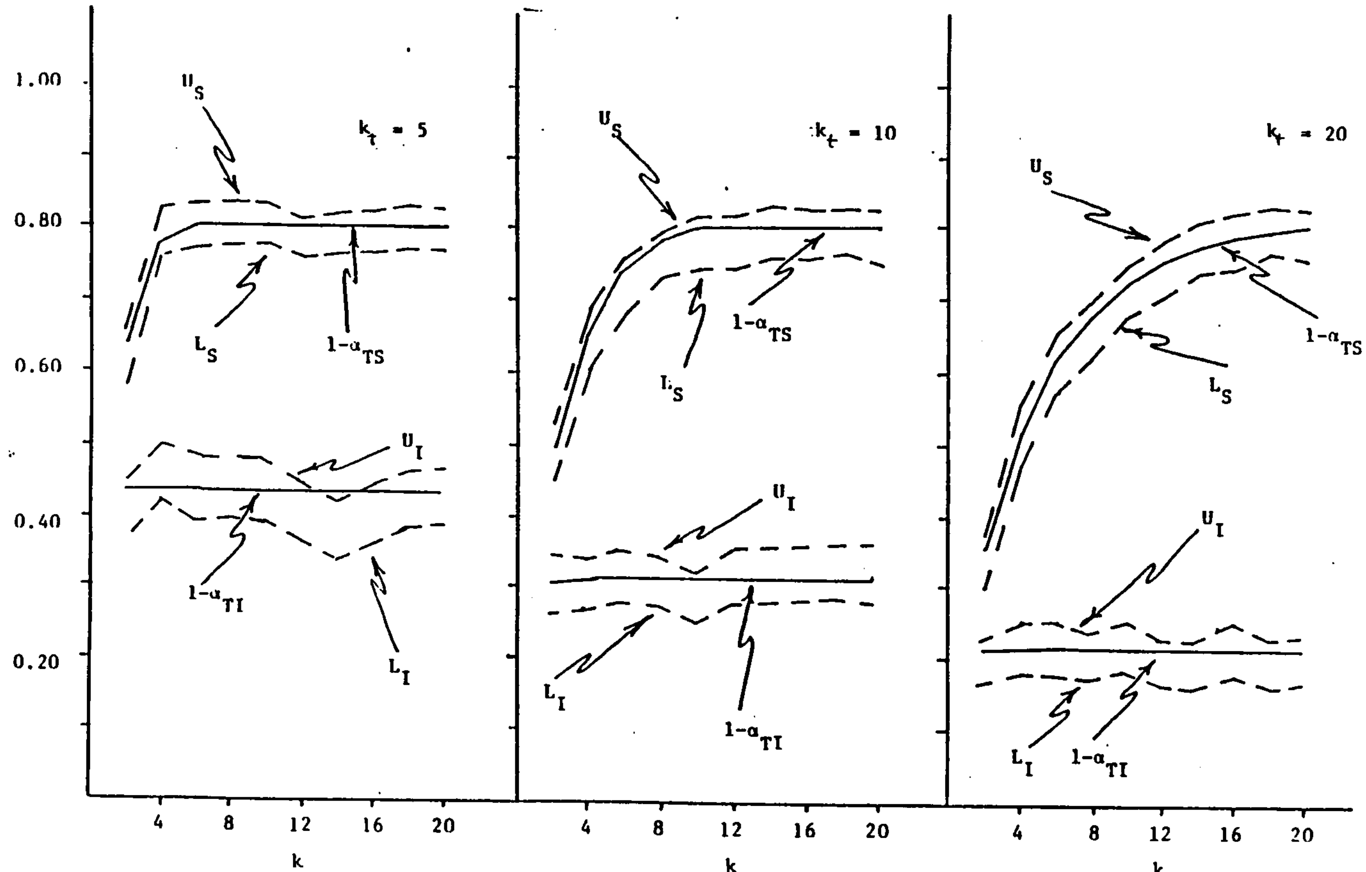

Figure 11

Results of the Analysis for Sequential Systematic Sampling Where L_S, U_S = Lower and Upper 99% Confidence Limits for $1 - \alpha_{TS}$, L_I, U_I = Lower and Upper 99% Confidence Limits for $1 - \alpha_{TI}$, k = Systematic Sampling Interval, m = Number of Observations per Systematic Sample (50).

The results in Table 15 indicate that the expressions for the predicted true level of confidence in equations (5.2.2.8), (5.2.2.9) and (5.2.2.10) adequately describe the performance of those methods in constructing confidence intervals for means.

In Figure 10 the predicted levels of confidence for batch means, $1 - \alpha_{TB}$, and those resulting if dependence is ignored, are computed by equations (5.2.2.9) and (5.2.2.8) respectively and are represented by the solid line segments. For batch size of 2,4,6,...,20 and 50 batches in each case, 500 sampling experiments were conducted where an 80% confidence interval for μ is computed in each case, first applying the method of batch means with k and m = 50 and second ignoring dependence where the sample size, n, is km = 50k. Based upon the simulation results of the experiments for each method a 99% confidence interval for the true level of confidence is computed. The 99% confidence limits in both cases are represented by the dashed line segments in Figure 10 where L_B, U_B are the 99% confidence limits for the true level of confidence for batch means and L_I, U_I are the 99% confidence limits for the true level of confidence achieved when the sample observations are erroneously assumed to be independent. In parallel with Figure 10, the results in comparing $1 - \alpha_{TS}$ and $1 - \alpha_{TI}$ are presented in Figure 11.

The results presented in Figure 10 and 11 suggested that both batch means and sequential systematic sampling lead to improvements on confidence interval analysis if the data dependence is ignored. However, by comparison the improvement yielded by sequential systematic sampling is more satisfactory than that contributed by batch means.

As for the performance of each method in the case of hypothesis test, the conclusions stated for the case of confidence interval apply to the results in Table 16.

5.3 NONNORMALITY AND CONFIDENCE INTERVALS FOR MEANS

The computational and simulation results presented in the last section have indicated the superior predictability of the method of sequential systematic sampling with respect to inferential procedures for means. Since the sample observations concerned in this research are correlated and may not be normally distributed, the robustness of confidence interval procedures considered should be investigated. Consider the simple moving average model in the following form:

$$X_i = \sum_{j=1}^{k_t} a_{j+i-1} \qquad (5.3.1)$$

where $\{a_i\}$ is a sequence of i.i.d. $G(\lambda,1)$ (exponential) random variables. The distribution of X_i is given by

$$X_i \sim G(\lambda,k_t). \qquad (5.3.2)$$

Then, by letting $m = 50$ and $\lambda = 0.2$, estimated level of confidence for means can be obtained by the methods of independence assumed, batch means, and sequential systematic sampling from 500 sampling experiments where $k_t = 5$, 10, and $2 < k < 20$. The results are summarized in Table 17. By comparison with the results presented in Table 15, the performance of the method of sequential systematic sampling was not affected in the presence of nonnormal populations, and neither were the cases of batch means and independence assumed. Hence the robustness of confidence interval procedurs for means is indicated.

Table 17

Coverage of 80% Confidence Intervals for Single Mean for Gamma Distributed Data in Simple Moving Average Model with Order k_t.

m = 50, Replication = 500.

k_t	k	Assumed Independence	Batch Means	Sequential Systematic Sampling
5	2	0.4080	0.5300	0.6380
	3	0.4500	0.6480	0.7360
	4	0.4560	0.6320	0.8020
	5	0.4340	0.7040	0.7904
	10	0.4500	0.7560	0.7860
	20	0.4500	0.8000	0.8100
10	2	0.2960	0.4120	0.4880
	3	0.3060	0.4800	0.5800
	4	0.3180	0.5620	0.6900
	5	0.3120	0.6140	0.7260
	7	0.2900	0.6800	0.8000
	8	0.3060	0.6720	0.7800
	9	0.3260	0.6980	0.7840
	10	0.3160	0.7260	0.8160
	14	0.3160	0.7480	0.8060
	16	0.3100	0.7540	0.7960
	20	0.3140	0.7780	0.8140
20	2	0.1880	0.2460	0.3040
	4	0.2200	0.4220	0.5200
	6	0.2580	0.4980	0.6380
	8	0.2240	0.5660	0.6920
	10	0.2440	0.5760	0.7000
	12	0.2060	0.6180	0.7580
	14	0.1820	0.6640	0.8060
	16	0.2180	0.6760	0.8220
	18	0.2360	0.7100	0.8220
	20	0.2360	0.7140	0.8020

5.4 CONFIDENCE INTERVAL PROCEDURES FOR TWO MEANS

The simulation experiments conducted in the case of single mean can be easily extended to the case of two means. Let X_{1i}, $i = 1,2,\cdots,n$, and X_{2j}, $j = 1,2,\cdots,n$, be two independent sequences of normally distributed random variables with mean μ_1 and μ_2 respectively. Assume that both X_{1i} and X_{2j} have the same covariance matrix Σ indicated in section (5.2.3), and they are also both generated by the same simple moving average process in equation (5.2.3.1). The 80% confidence interval procedures for two means based upon independence assumed, batch assumed, and sequential systematic sampling can be constructed and applied in the manner resembling the case of single mean (Appendix F and G contain the computational forms for these procedures). The results of the estimated level of confidence are summarized in Table 18, and, by comparison, they are in accordance with conclusions rendered for the case of single mean.

5.5 CONCLUDING REMARKS

With the ignorance of data correlation the classical method had the highest estimator bias and worst predictability in the analysis presented in this chapter. The method of batch means provides an improved technique in estimating and thereby improves the predictability of inferential procedures for means. However, this method does not perform relatively better than the method of sequential systematic sampling. In the case of AR(1) models, the method of sequential systematic sampling provides an estimator of $\sigma^2_{\bar{x}}$ with smaller bias than that of batch means. In the case of simple moving average models, the method of sequential systematic sampling also leads to less

Table 18

Coverage of 80% Confidence Intervals for Two Means in Simple Moving Average Models with Order k_t.

$m = 26$, $n = mk$, $k_t = 10$, $\alpha_D = .20$, Replication = 500

k	Assumed Independence	Batch Means	Sequential Systematic Sampling
4	0.3040	0.5160	0.6420
6	0.2960	0.6100	0.7120
8	0.3380	0.6900	0.7800
9	0.2980	0.6560	0.7740
10	0.3260	0.6740	0.7800
11	0.2860	0.6740	0.7820
12	0.2940	0.6820	0.7880
13	0.2920	0.6980	0.7620
14	0.2920	0.7120	0.7620
15	0.3100	0.7440	0.7980
16	0.3120	0.7400	0.7860

bias in estimating $\sigma_{\bar{X}}^2$ and in turn yields a greater improvement of the predictability of the inferential procedures applied. In general, one can anticipate less bias in estimating the variance of the sample mean when using the method of sequential systematic sampling than the method of batch means. This is particularly true as the sample size increases.

CHAPTER VI

THE PARAMETRIC INFERENTIAL PROCEDURES FOR VARIANCES

6.1 INTRODUCTION

Chapter III establishes the potential significance of the problems associated with the statistical anlaysis of nonnormal and correlated sample observations in the case where the population variance is concerned. The bias in estimating the population variance when the data are correlated was presented in Section (3.2), where the magnitude of the bias is similar to that in estimating the variance of the sample mean. The impacts of violation of the assumptions of independence and normality on the inferential procedures for variances are discussed in Section (3.7), where the predictability of the classical procedures is not reliable in the examples presented.

One important objective of this research is to reduce the deleterious impacts of such violation in the case of variances. To achieve this objective the approach in this research first attempts to reduce the bias in estimating the population variance and then improves the predictability of the inferential procedure considered. In the case of variances, the predicability is affected by both data correlation and nonnormality. The approach will first consider the case where the assumption of independence is violated and then the case where both the independence assumption and the normality assumption are violated.

This chapter treats methods which may be applied in attempting to reduce bias in estimating the population variance and thus can be applied to improve the predictability of inferential procedures for variances when the data are correlated and normally distributed. Chapter VII treats methods with the same objectives as in Chapter VI except that the data are correlated and nonnormally distributed. In each case the criterion for evaluating the performance of a proposed method is based upon the predictability of the inferential procedure applied.

6.2 METHODOLOGY

To construct a statistical inferential procedure for variances in the presence of data correlation, typically a statistician will first find an estimator for the population variance, and then he will define the sampling distribution of the variance estimator. As presented in Chapter IV, the development of the method of sequential systematic sampling leads to the estimator of the population variance and that of the variance of the sample mean. In the case of means, the distribution of the sample mean is normal for both independent and correlated data. Therefore, once the estimator of the variance of the sample mean is available, construction of inferential procedures for means follows. However, a similar situation may not be expected in the case of variances. The sampling distribution of a variance estimator is often difficult to define in the case of correlated data. To reduce the impact of data correlation the analyst may use uncorrelated observations obtained by the method of sequential systematic sampling to assist development of the sampling distribution of the variance estimator. Based upon this sampling distribution, inferential procedures for variances can then be developed. This in turn may lead to an improvement of the predictability of the analysis.

It is well-known that nonnormality affects the predictability of inferential procedures for variances in the case of independent data. Thus, for procedures for variances, the property of nonrobustness may also be anticipated in the case of correlated data. However, the robustness of the inferential procedures developed in this chapter will not be considered until the next chapter. Hence, throughout this chapter the sample observations are assumed to be normally distributed.

6.3 A PROCEDURE OF CHI-SQUARE APPROXIMATION

As indicated in Chapter IV, the method of sequential systematic sampling leads to the estimator of population variance given by

$$\hat{\sigma}^2 = \frac{1}{k} \sum_{i=1}^{k} s_i^2$$

where s_i^2 is the traditional variance estimator for the ith systematic sample and given by

$$s_i^2 = \frac{1}{(m-1)} \sum_{\ell=1}^{m} (X_{(\ell-1)k+i} - \bar{X}_i)^2$$

and

$$\bar{X}_i = \frac{1}{m} \sum_{\ell=1}^{m} X_{(\ell-1)k+i}, \quad i = 1,2,\ldots,k.$$

If the assumed order of serial dependence (k-1) is so large that $X_{(\ell-1)k+i}$ and $X_{(h-1)k+i}$ are independent for a given i, with $\ell,h = 1,2,\ldots,m$ and $\ell \neq h$, then observations in each systematic sample are uncorrelated. As a result,

$$E(s_i^2) = \sigma^2.$$

Therefore,

$$E(\hat{\sigma}^2) = \sigma^2.$$

Thus, given the assumption of the order of serial dependence, the method of sequential systematic sampling leads to an unbiased estimator of the population variance. However, while observations in the same systematic sample may be considered as independent, observations in different systematic samples may be correlated. Hence, the s_i^2 are not independent but each $(m-1)s_i^2/\sigma^2$ is chi-square distributed under the assumed order of serial dependence. Therefore the analyst may approximate the sampling distribution of the variance estimator $\hat{\sigma}^2$ by the following:

$$f \frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(f)$$

where

$$f = \frac{2E^2(\hat{\sigma}^2)}{Var(\hat{\sigma}^2)} .$$

For the numerator,

$$E^2(\hat{\sigma}^2) = (\sigma^2)^2 = \sigma^4.$$

For the denominator, the variance of $\hat{\sigma}^2$ can be expressed as

$$Var(\hat{\sigma}^2) = \frac{1}{k^2} [\sum_{i=1}^{k} Var(s_i^2) + 2 \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} Cov(s_i^2, s_j^2)]. \quad (6.3.1)$$

Notice that the variance of s_i^2 can be shown as

$$Var(s_i^2) = \sigma^4(\frac{\gamma_2}{m} + \frac{2}{m-1})$$

where γ_2 is the coefficient of kurtosis, which is a measure of flattening of a frequency curve near its peak. The value of γ_2 will be zero if the sample observations are normally distributed. The covariance term $\mathrm{Cov}(s_i^2, s_j^2)$ is defined by

$$\mathrm{Cov}(s_i^2, s_j^2) = E[(s_i^2 - \sigma^2)(s_j^2 - \sigma^2)]$$
$$= E(s_i^2 s_j^2) - \sigma^4.$$

If the variables X_i have common mean μ, let $\varepsilon_{(\ell-1)k+i}$ and $\bar{\varepsilon}_i$ denote the following:

$$\varepsilon_{(\ell-1)k+i} = X_{(\ell-1)k+i} - \mu \tag{6.3.2}$$

$$\bar{\varepsilon}_i = \frac{1}{m} \sum_{\ell=1}^{m} (X_{(\ell-1)k+i} - \mu). \tag{6.3.3}$$

Then

$$E(\varepsilon_{(\ell-1)k+i}) = 0, \text{ and } X_{(\ell-1)k+i} - \bar{X}_i = \varepsilon_{(\ell-1)k+i} - \bar{\varepsilon}_i.$$

Now,

$$E(s_i^2, s_j^2) = E\{\frac{1}{(m-1)^2} [\sum_{\ell=1}^{m} (X_{(\ell-1)k+i} - \bar{X}_i)]^2 [\sum_{h=1}^{m} (X_{(h-1)k+j} - \bar{X}_j)]^2\}.$$

Using equation (6.3.1) and (6.3.2), one obtains

$$E(s_i^2, s_j^2) = \frac{1}{(m-1)^2} E\{[\sum_{\ell=1}^{m} (\varepsilon_{(\ell-1)k+i} - \bar{\varepsilon}_i)^2]$$
$$[\sum_{h=1}^{m} (\varepsilon_{(h-1)k+j} - \bar{\varepsilon}_j)^2].$$

Following the formula of the product moment of fourth order for a multivariate normal distribution given in equation (A.1.4) of Appendix A, $E(s_i^2, s_j^2)$ and subsequently $\mathrm{Var}(\hat{\sigma}^2)$ is given in Appendix D as equation (D.1). The estimate of $\mathrm{Var}(\hat{\sigma}^2)$, $\widehat{\mathrm{Var}}(\hat{\sigma}^2)$, is also provided in Appendix D. Now the estimated degrees of freedom $\hat{f}$ can be expressed as follows:

$$\hat{f} = \frac{2\hat{\sigma}^4}{\widehat{Var}(\hat{\sigma}^2)} \tag{6.3.4}$$

where $\widehat{Var}(\hat{\sigma}^2)$ is given in Appendix D.

Based upon the degrees of freedom as computed by equation (6.3.4), the analyst can apply classical techniques to conduct the confidence interval estimation and hypothesis tests for variances through the use of approximate chi-square distribution. To construct an approximate $(1 - \alpha)100\%$ confidence interval for σ^2, the confidence limits are given by

$$L,U = \chi^2_{\alpha/2}(\hat{f}),\ \chi^2_{1-\alpha/2}(\hat{f}). \tag{6.3.5}$$

The sample chi-square statistic is given by $\chi^2_{exp} = \frac{\hat{f}\hat{\sigma}^2}{\sigma^2}$. Now, consider the test of H_0: $\sigma^2 = \sigma_0^2$, against H_1: $\sigma^2 \neq \sigma_0^2$, the test limits are given by

$$L,U = \chi^2_{\alpha/2}(\hat{f}),\ \chi^2_{1-\alpha/2}(\hat{f}). \tag{6.3.6}$$

Under the null hypothesis, the chi-square test statistic is given by $\chi^2_{exp} = \frac{\hat{f}\hat{\sigma}^2}{\sigma_0^2}$. The power of the test is

$$\begin{aligned} 1 - \beta(\sigma_1^2) &= 1 - \Pr[\chi^2_{\alpha/2}(\hat{f}) < \frac{f\hat{\sigma}^2}{\sigma_0^2} < \chi^2_{1-\alpha/2}(\hat{f})/\sigma_1^2] \\ &= 1 - \Pr[\frac{\sigma_0^2}{\sigma_1^2}\,\chi^2_{\alpha/2}(\hat{f}) < \chi^2(\hat{f}) < \frac{\sigma_0^2}{\sigma_1^2}\,\chi^2_{1-\alpha/2}(\hat{f})] \\ &= 1 - \Pr[\frac{1}{R_\sigma}\,\chi^2_{\alpha/2}(\hat{f}) < \chi^2(\hat{f}) < \frac{1}{R_\sigma}\,\chi^2_{1-\alpha/2}(\hat{f})] \end{aligned} \tag{6.3.7}$$

where $R_\sigma = \sigma_1^2/\sigma_0^2$.

Conclusions on single-tail tests can be obtained by inference.

In the case of classical statistics, the variance estimator $\hat{\sigma}^2$ is replaced by s^2, and the degrees of freedom is given by (n-1). Hence equations (6.3.5), (6.3.6), and (6.3.7) are given by the following:

$$L,U = \chi^2_{\alpha/2}(n-1),\ \chi^2_{1-\alpha/2}(n-1) \qquad (6.3.8)$$

$$L,U = \chi^2_{\alpha/2}(n-1),\ \chi^2_{1-\alpha/2}(n-1) \qquad (6.3.9)$$

$$1 - \beta(\sigma_1^2) = 1 - \Pr[\frac{1}{R_\sigma}\chi^2_{\alpha/2}(n-1) < \chi^2(n-1) < \frac{1}{R_\sigma}\chi^2_{1-\alpha/2}(n-1)]. \qquad (6.3.10)$$

6.4 AN APPROXIMATE NORMAL PROCEDURE

The inferential procedures for variances developed in the last section are based upon a chi-square approximation. Under the assumed order of serial dependence, inferential procedures for variances can also be developed by a normal approximation.

With the assumption of an appropriate order of serial dependence, the statistic $(m-1)s_i^2/\sigma^2$ has the chi-square distribution with (m-1) degrees of freedom. As a result,

$$E[\frac{(m-1)s_i^2}{\sigma^2}] = m-1$$

$$Var[\frac{(m-1)s_i^2}{\sigma^2}] = 2(m-1)$$

where $i = 1,2,\ldots,k$. For large m, it can be assumed that

$$\frac{(m-1)s_i^2}{\sigma^2} \sim N((m-1),\ 2(m-1))$$

for $i = 1,2,\ldots,k$. So,

$$s_i^2 \sim N(\sigma^2, \frac{2}{m-1}\sigma^4).$$

Let $S = [s_1^2, s_2^2, \ldots, s_k^2]$. Then

$$S \sim N(\sigma^2 I, \Sigma) \tag{6.4.1}$$

where I is the unit vector of order k x 1 and

$$\Sigma = \begin{bmatrix} \frac{2}{(m-1)}\sigma^4 & & \mathrm{Cov}(s_i^2, s_j^2) \\ & \ddots & \\ \mathrm{Cov}(s_i^2, s_j^2) & & \frac{2}{(m-1)}\sigma^4 \end{bmatrix}$$

Since

$$\hat{\sigma}^2 = \frac{1}{k}\sum_{i=1}^{k} s_i^2.$$

Given the distribution form of s_i^2 in (6.4.1), the following distribution of $\hat{\sigma}^2$ can be established,

$$\hat{\sigma}^2 \sim N(\sigma^2, \mathrm{Var}(\hat{\sigma}^2))$$

where

$$\mathrm{Var}(\hat{\sigma}^2) = \frac{1}{k^2}\left[\frac{2k}{m-1}\sigma^4 + 2\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\mathrm{Cov}(s_i^2, s_j^2)\right]$$

which is same as $\mathrm{Var}(\hat{\sigma}^2)$ in equation (6.3.1) in the procedure of chi-square approximation. Hence it would be reasonable to use the estimate, $\widehat{\mathrm{Var}}(\hat{\sigma}^2)$, suggested in Appendix D for $\mathrm{Var}(\hat{\sigma}^2)$. Given $\widehat{\mathrm{Var}}(\hat{\sigma}^2)$, one can apply classical techniques to conduct the confidence interval estimation and hypothesis tests for variances. To construct an approximate $(1-\alpha)100\%$ confidence interval for σ^2, the confidence limits are given by

$$L,U = \hat{\sigma}^2 \pm \sqrt{\widehat{Var}(\hat{\sigma}^2)}\, Z_{1-\alpha/2}. \qquad (6.4.2)$$

The true value of the variance may be used in equation (6.4.2) if it becomes known. Now, consider the test of H_0: $\sigma^2 = \sigma_0^2$, against H_1: $\sigma^2 \neq \sigma_0^2$, the test limits are given by

$$L,U = \hat{\sigma}_0^2 \pm \sqrt{\widehat{Var}(\hat{\sigma}^2)}\, Z_{1-\alpha/2}. \qquad (6.4.3)$$

The power of the test is

$$1 - \beta(\sigma_1^2) = 1 - Pr(L < \hat{\sigma}^2 < U/\sigma_1^2)$$

$$= 1 - Pr\left[\frac{\sigma_0^2-\sigma_1^2}{\sqrt{Var(\hat{\sigma}^2)}} - \frac{1}{R_\sigma} Z_{1-\alpha/2} < Z < \frac{\sigma_0^2-\sigma_1^2}{\sqrt{Var(\hat{\sigma}^2)}} + \frac{1}{R_\sigma} Z_{1-\alpha/2}\right]$$

$$= 1 - Pr\left[\frac{-d}{\sqrt{Var(\hat{\sigma}^2)}} - \frac{1}{R_\sigma} Z_{1-\alpha/2} < Z < \frac{-d}{\sqrt{Var(\hat{\sigma}^2)}} + \frac{1}{R_\sigma} Z_{1-\alpha/2}\right]$$

(6.4.4)

where $d = \sigma_1^2 - \sigma_0^2$, $R_\sigma = \dfrac{\sigma_1^2}{\sigma_0^2}$.

6.5 PERFORMANCE CRITERIA FOR PROCEDURES FOR VARIANCES

To examine the predictability of the inferential procedures developed for a single variance, simulation experiments are conducted. In the case of confidence interval analysis, the simulation experiment examines the precision with which the inferential procedure applied predicts the level of confidence. In the case of hypothesis testing, the simulation experiment examines the precision with which the inferential procedure applied predicts the power of hypothesis tests. The tests of hypotheses to be conducted are either the two-tail test, given by H_0: $\sigma^2 = \sigma_0^2$ against H_1: $\sigma^2 \neq \sigma_0^2$, or single-tail tests whose formulations can be inferred by comparison.

In the case of confidence interval estimation, a simple moving average process with known variance σ^2 is simulated for a given sample size n where n = mk with m and k as previously defined. Given one simulated process, take note whether the confidence interval formed by a given procedure cover σ^2. Then independently replicate the entire process R (R = 500) times. From these R replications, for each procedure, one can compute the proportion s of the R confidence intervals which contain σ^2. The proportion s is referred to as the estimated level of confidence in this chapter.

In the case of hypothesis testing, a simple moving average process with known variance σ_1^2 is simulated for a given sample size n where n = mk. Given one simulated process, take note whether the test limits formed by a given procedure do not cover σ_0^2, i.e. the rejection of null hypothsis. Then independently replicate the entire process R times. From these R replications, for each procedure, one can compute the ratio t as the proportion of rejections of H_0. The proportion t is referred to as the estimated power of hypothesis tests in this chapter.

If the inferential procedures applied are performing adequately, s should be near $1 - \alpha$, and t should be near $1 - \beta$, where $1 - \alpha$ and $1 - \beta$ are the predicted level of confidence and the predicted power of the test respectively. Both the predicted level of confidence and power of the test can be obtained directly from the application of classical techniques to the inferential procedure concerned.

Similar criteria are applied to the case of two variances.

While the criteria of performance thus described may illustrate the predictability of the analysis, comparing the inferential procedures developed with the classical inferential procedures for variances may indicate how much improvement on the predictability can be made in this research. Therefore the comparison will be presented in the case of simple moving average models.

6.6 SIMULATION EXPERIMENT UNDER NORMALITY ASSUMPTION

As in the case of means, correlated normal random variables in a simple moving average form of equation (5.1.8) are generated. Then equations (6.3.5), (6.3.8), and (6.4.2) are applied to obtain the estimated levels of confidence; equations (6.3.6), (6.3.9), and (6.4.3) are applied to obtain the estimated power functions of hypothesis test. Table 19 summarizes results in estimating and predicting the level of confidence. Table 20 presents similar results in the case of hypothesis test while the normal approximation is used in applying the method of sequential systematic sampling. Table 21 presents the results in hypothesis testing by the chi-square approximation.

Tables 19, 20 and 21 indicate that the predictability of the classical inferential procedures is disappointing. By comparison, these two tables illustrate the propiety of applying the method of sequential systematic sampling to improve the predictability of the analysis.

6.7 CONCLUDING REMARKS

In this chapter chi-square and normal distribution are suggested as the approximate distribution forms for the estimator of population variance $\hat{\sigma}^2$ which is acquired by applying the method of sequential

Table 19

Coverage of 80% Confidence Interval Procedures for Single Variance Moving Average Models with Normal Input.

$m = 40$, $\alpha_D = .20$, Replication = 500

k_t	Sequential Systematic Sampling		Independence Assumed
	Chi-square Approximation	Normal Approximation	
4	0.8042	0.8010	0.5900
5	0.8080	0.7990	0.5400
8	0.7842	0.8100	0.4160

Table 20

Power Function Obtained by Normal Approximation for Simple Moving Average Models with Normal Input.

$m = 40$, $\alpha_D = .20$, Replication = 500

k_t	Sequential Systematic Sampling			Independence and Normality Assumed		
	d	Simulation	Prediction	R_σ	Simulation	Prediction
4	0.10	0.1800	0.1524	1.04	0.4160	0.3143
	0.20	0.2120	0.2149	1.08	0.4720	0.4429
	0.50	0.4580	0.4312	1.20	0.5940	0.7832
	1.0	0.7500	0.7320	1.40	0.8360	0.9813
5	0.10	0.2019	0.1652	1.05	0.4860	0.3647
	0.20	0.2432	0.2460	1.10	0.5120	0.5475
	0.50	0.5100	0.5228	1.25	0.6760	0.9133
	1.0	0.8341	0.8447	1.50	0.9040	0.9988
8	0.10	0.2142	0.2076	1.08	0.5720	0.5551
	0.20	0.3600	0.3513	1.16	0.6300	0.8477
	0.50	0.7500	0.7716	1.40	0.8760	0.9980
	1.0	1.0	0.9858	1.80	0.9880	1.0

Table 21

Power Function Obtained by Chi-Square Approximation for Simple Moving Average Models with Normal Input.

$m = 40$, $\alpha_D = .20$, Replication = 500

k_t	R_σ	Simulation	Prediction
4	1.04	0.2940	0.2671
	1.08	0.3540	0.3397
	1.20	0.5810	0.5613
	1.40	0.8340	0.8290
5	1.05	0.2900	0.2851
	1.10	0.4000	0.3790
	1.25	0.6720	0.6553
	1.50	0.9500	0.9180
8	1.08	0.3521	0.3419
	1.16	0.5241	0.5021
	1.40	0.8800	0.8708
	1.80	1.0000	0.9966

systematic sampling. Based upon these two approximate distributrions, one can develop inferential procedures to predict the level of confidence and the power of hypothesis tests. The results of the simulation experiments under the normality assumption demonstrate the adequacy of the procedures developed with respect to the predictability of the inferential procedure applied.

CHAPTER VII

NONNORMALITY AND PROCEDURES FOR VARIANCES

7.1 INTRODUCTION

Since the parametric inferential procedures for variances are quite sensitive to departure from normality in the case of independent data, a similar conclusion may be applied to the case of correlated data. Hence the necessity of examining the property of robustness arises for the inferential procedures developed in the last chapter.

If the procedures do not have the property of robustness, development of inferential procedures for variances for correlated and nonnormally distributed data will be undertaken in this chapter. The inferential procedures developed will be evaluated through simulation experiments with the same performance criteria as in the case of normal data. Hence, correlated and nonnormal random variables in the simple moving average form will be generated in the experiments.

7.2 THE ROBUSTNESS OF THE PROCEDURES FOR VARIANCES

Concerning the impact of nonnormality, correlated and nonnormal random variables are generated to test the robustness of the parametric inferential procedures developed under the assumption of normality. Remember that the correlated normal random variables in the form of simple moving average model (equation (5.1.8) of Chapter V) are given by

$$X_i = \sum_{t=1}^{k} a_{t+i-1}$$

where $\{a_t\}$ is a sequence of i.i.d. N(0,1) random variables. In order to generate correlated and nonnormal random variables, one needs to change the input distribution N(0,1) to a nonnormal distribution such that the output variables X_i are nonnormally distributed. To this end, the input distribution is changed to $G(\lambda,1/k)$, where $G(\lambda,1/k)$ is a gamma distribution with parameter λ and $1/k$, and $(k-1)$ is the assumed order of serial dependence. The probability density function (p.d.f.) of the input variable $\{a_t\}$ is given by

$$f(a_t) = \frac{\lambda^{\frac{1}{k}}}{\Gamma(\frac{1}{k})} a_t^{\frac{1}{k}-1} e^{-\lambda a_t}, \quad 0 < a_t < \infty.$$

If $\{a_t\}$ is a sequence of i.i.d. $G(\lambda,1/k)$ random variables, then the output variables X_i will have the following distribution,

$$X_i \sim G(\lambda, \sum_{t=1}^{k} 1/k)$$

$$\sim G(\lambda,1).$$

Since $G(\lambda,1)$ refers to the exponential distribution with parameter λ, $E(\lambda)$,

$$X_i \sim E(\lambda).$$

Thus the output distribution is $E(\lambda)$ if the input distribution is $G(\lambda,1/k)$. The p.d.f. of the output variable X is given by

$$f(X_i) = \lambda e^{-\lambda X_i}, \quad X_i > 0.$$

The variance of the exponentially distributed output variable X_i is given by

$$\sigma^2 = \frac{1}{\lambda^2}.$$

Now, consider the hypothesis test of H_0: $\sigma^2 \leq \sigma_0^2$ against

H_1: $\sigma^2 > \sigma_0^2$ for the simulation experiment where nonnormal data is generated. The value of λ is given as 0.2 to obtain σ^2 in confidence interval analysis and ot obtain σ_0^2 for the test above. Both the estimated and predicted powers of the test are a function of λ. Thus, by changing the value of λ, the estimated and predicted powers of the test can be acquired.

Table 22 summarizes the results in estimating and predicting the level of confidence when the input distribution is given by $G(\lambda, 1/k)$. As Table 22 indicates, the predictability of the inferential procedures developed under the assumption of normality is disappointing and the predictability deteriorates in the case of the classical procedure. Therefore the simulation study has suggested that the inferential procedures developed in Chapter VI are not robust when the normality assumption is violated.

7.3 THE REVISED PROCEDURES FOR VARIANCES

If the sample obserations X_i are nonnormally distributed, the variance of s_i^2 is given by

$$\mathrm{Var}(s_i^2) = \sigma^4\left(\frac{\gamma_2}{m} + \frac{2}{m-1}\right).$$

Another impact of violating the assumption of normality, in the development of the procedures for variances, is that the formula developed for $\mathrm{Var}(\hat{\sigma}^2)$ in equation (D.1) is not correct. The derivation of product moment of fourth order for $\mathrm{Var}(\hat{\sigma}^2)$ is based on the equation (A.1.4) of Appendix A, which relies upon the assumption of normality. Therefore the expression for $\mathrm{Var}(\hat{\sigma}^2)$ in equation (D.1) is no longer valid when the assumption of normality is violated.

Table 22

80% Confidence Interval Procedures for Single Variance and Exponential Output.

$m = 40$, $\alpha_D = 0.20$, Replication = 500

INPUT: $G(\lambda, \frac{1}{k_t})$, OUTPUT: $E(\lambda)$

k_t	Estimated Level of Confidence		
	Sequential Systematic Sampling		Independence and Normality Assumed
	Chi-square Approximation	Normal Approximation	
4	0.4340	0.4580	0.2520
5	0.4120	0.4049	0.2140
8	0.3840	0.3745	0.1944

Anderson [1971] gives the formula for the fourth moment for stationary processes in general. For a stationary process $\{X_t\}$ with mean μ, assume that $\{X_t\}$ has the finite fourth moment as

$$\begin{aligned} & E(X_t - \mu)(X_{t+h} - \mu)(X_{t+g} - \mu)(X_{t+k} - \mu) \\ & = E(X_0 - \mu)(X_h - \mu)(X_g - \mu)(X_k - \mu) \\ & = V(h,g,k) \end{aligned}$$

where $t = \ldots,-1,0,1,\ldots$. The process $\{X_t\}$ can also be represented by the stationary moving average process

$$X_t - \mu = \sum_{S=-\infty}^{\infty} \theta_S \, Z_{t-S}, \quad t = ,,,.-1,0,1,\ldots$$

where $\{Z_t\}$ is a sequence of i.i.d. random variables with mean zero and variance σ_Z^2. Then, it can be shown that

$$V(h,g,k) = K(h,g,k) + \sigma_h\sigma_{g-k} + \sigma_g\sigma_{h-k} + \sigma_k\sigma_{h-g} \qquad (7.3.1)$$

where K(h,g,k) can be considered as the nonnormality factor for the process $\{X_t\}$, and σ_h is the lag h covariance.

Notice that if the variables X_t are normally distributed,

$$K(h,g,k) = 0.$$

As a result, equation (7.3.1) is equivalent to equation (A.1.4) in Appendix A. K(h,g,k) can be expressed in terms of θ_S and the fourth-order cumulant of z_t, k_4,

$$K(h,g,k) = k_4 \sum_{t=-\infty}^{\infty} \theta_t\theta_{t+h}\theta_{t+g}\theta_{t+k}.$$

Considering the nonnormality properties in V(h,g,k) and s_1^2, the variance of $\hat{\sigma}^2$ can be developed and is given in Appendix E as equation (E.1). The estimate of $Var(\hat{\sigma}^2)$, $\widehat{Var}(\hat{\sigma}^2)$, is also provided in Appendix

E.

Using the revised $\widehat{\mathrm{Var}}(\hat{\sigma}^2)$ given in Appendix E, the analyst can construct inferential procedures for variances as in the case of normal data. Since the inferential procedures based upon the chi-square or normal approximation are similar, the ones based upon the normal approximate procedure will be used in this chapter.

7.4 TWO CASE STUDIES

The revised confidence interval procedure for variances is first applied to the exponential data for which the unrevised confidence interval procedure for variances is applied. The results are summarized in Table 23. As Table 23 illustrates, the revised confidence procedure leads to an improvement on the predictability of the analysis.

Simulation experiments are conducted in two cases to examine the effectiveness of the revised inferential procedures for variances in predicting the level of confidence and the power of hypothesis tests. In each case correlated random variables are generated in a simple moving average form of equation (5.1.8). In the first case, the input distribution is given by $G(\lambda,1/k)$ as in Section (7.2). In the second case, to obtain a more skewed output distribution than that of the first case, the input distribution is given by $G(\lambda,1/(2k))$. Hence the distribution of the output variable X_i is given by

$$X_i \sim G(\lambda, \sum_{t=1}^{k} \frac{1}{2k})$$

$$\sim G(\lambda,\frac{1}{2}).$$

Table 23

Revised 80% Confidence Interval Procedures for Single Variance and Exponential Output.

$m = 40$, $\alpha_D = 0.20$, Replication $= 500$

INPUT: $G(\lambda, \frac{1}{k_t})$, OUTPUT: $E(\lambda)$

k_t	Estimated Level of Confidence	
	Sequential Systematic Sampling	Independence and Normality Assumed
4	0.8400	0.2520
5	0.8320	0.2140
8	0.8201	0.1944

The variance of the gamma distributed output variable X_1 is given by

$$\sigma^2 = \frac{1}{2\lambda^2} .$$

The results of simulation experiments for the two cases where the input distributions given by $G(\lambda,1/k)$ and $G(\lambda,1/(2k))$ respectively are summarized in Tables 24 and 25. As these two tables illustrate, the predictability of the procedure applied seems to be influenced by the magnitude of kurtosis. However, in comparison with the results reported in Table 22, the revised inferential procedures have led to an improvement on the predictability of the analysis when the assumption of normality is violated.

7.5 A METHOD OF ANALYSIS FOR GENERAL INPUT DISTRIBUTION

The simulation experiments so far conducted in this chapter were limited to input distributions which are gamma. The reason was in part due to the additive property of the gamma input variables which renders convenience for deriving the distribution of the output variables. Although the knowledge of the true value of kurtosis of the output process is not required to implement the revised inferential procedures for variances, such knowledge would provide a measure of departure from normality for the sample observations.

As indicated in the chapter of literature review, test procedures for variances are affected by the magnitude of kurtosis in the case of independent data. Hence it would be important to examine the predictability of the revised inferential procedures in the case where the magnitude of kurtosis for the output variable is large. To achieve a large value of kurtosis for the output variable, the input variable

Table 24

Confidence Level and Power Function for Single Variance and Exponential Output.

$m = 90$, $k_t = 4$, $\alpha_D = .20$, Replication $= 500$

INPUT: $G(\lambda, \frac{1}{K_t})$, OUTPUT: $E(\lambda)$

<Confidence Level>

γ_2	Sequential Systematic Sampling	Independence and Normality Assumed
6	0.8200	0.2520

<Power Function>

Sequential Systematic Sampling			Independence and Normality Assumed		
d	Simulation	Prediction	R_σ	Simulation	Prediction
0.17	0.2940	0.3084	1.11	0.4900	0.7036
0.37	0.4060	0.4300	1.23	0.6460	0.970
1.22	0.7440	0.7549	1.78	0.9540	1.0

Table 25

Confidence Level and Power Function for Single Variance and Gamma Output.

$m = 200$, $k_t = 4$, $\alpha_D = .20$, Replication = 500

INPUT: $G(\lambda, \frac{1}{2k_t})$, OUTPUT: $G(\lambda, \frac{1}{2})$

<Confidence Level>

γ_2	Sequential Systematic Sampling	Independence and Normality Assumed
12	0.8400	0.2040

<Power Function>

Sequential Systematic Sampling			Independence and Normality Assumed		
d	Simulation	Prediction	R_σ	Simulation	Prediction
0.08	0.2500	0.3212	1.11	0.5560	0.8848
0.18	0.3780	0.4577	1.23	0.7020	0.9995
1.61	0.8140	0.8140	1.78	0.9800	1.0

may not be gamma distributed. Therefore, the following method is used for deriving the true value of kurtosis for the output process from a known input distribution.

Now, consider again the following simple moving average model,

$$X_i = \frac{1}{k} \sum_{t=1}^{k} a_{t+i-1} \tag{7.5.1}$$

where $\{a_t\}$ is a sequence of i.i.d. random variables with finite fourth moment.

One may attempt to derive the distribution of X_i and then determines the kurtosis of that distribution to measure the departure from normality. However, this is an arduous task in the case of correlated and nonnormally distributed data. Instead, one can derive the true value of kurtosis of X_i from that of input variables. The variance of X_i, σ_x^2, is given by

$$\sigma_x^2 = \frac{1}{k^2} [\sum_{i=1}^{k} \sigma_a^2 + 2 \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \mathrm{Cov}(a_i, a_j)]$$

$$= \frac{\sigma_a^2}{k} . \tag{7.5.2}$$

The kurtosis of X_i^2 is defined as

$$\gamma_{2,x} = \frac{\mu_{4,x}}{\sigma_x^4} - 3 \tag{7.5.3}$$

where $\mu_{4,x}$ is the fourth central moment of X_i. Now, according to the definition of $\mu_{4,x}$

$$\mu_{4,x} = E(X_i - \frac{1}{k} \sum_{i=1}^{k} \mu_a)^4$$

where $\mu_a = E(a_t)$. It can be shown that

$$\mu_{4,x} = \frac{\sigma_a^4}{k^3}\,[\gamma_{2,a} + 3k]. \tag{7.5.4}$$

Substituting (7.5.4) into (7.5.3) yields

$$\gamma_{2,x} = \frac{\gamma_{2,a}}{k}\,. \tag{7.5.5}$$

As shown in equation (7.5.2) and (7.5.3), the variance and the value of kurtosis of the output variable can be derived from the corresponding parameters of the input distribution.

To evaluate the predictability of the revised inferential procedures for variances when the data involves a large value of kurtosis, simulation experiments are conducted in the case where the input distribution is lognormal. That is $a_t \sim LN(\mu,\sigma^2)$. The variance of a_t is given by

$$\sigma_a^2 = \omega(\omega - 1)e^{2\mu}$$

where $\omega = \exp(\sigma^2)$. Hence the variance of output variable in confidence interval analysis is a function of σ. The results of the experiments are summarized in Tables 26.

It is known that the lognormal distribution approaches normal for a small value of σ (see Johnson and Kotz [1972]). As a result, the distribution of the output variable will approach normal as σ decreases. On the other hand, as the magnitude of σ increases, the value of kurtosis of the output variable increases. This in turn leads to departure from normality for the output variable. From Tables 26, it is observed that the predictability of the revised procedures depends on the magnitude of the kurtosis $\gamma_{2,x}$. With relatively small

Table 26

Confidence Level for Single Variance and Lognormal Input Distributions.

$m = 200$, $K_t = 4$, $\alpha_D = .20$, Replication = 500

INPUT: $LN(\mu,\sigma^2)$, $\mu = 0$

σ	γ_2	Sequential Systematic Sampling	Independence and Normality Assumed
0.2	0.1696	0.8040	0.5340
0.5	1.4746	0.8080	0.4140
1.0	27.7341	0.9300	0.2180
2.0	2305137.0	0.9980	0.0220

values of $\gamma_{2,x}$, .1696 and 1.4746, the revised procedures perform well. However, with a larger value of kurtosis, 27.7341, the performance of the revised procedures is disappointing.

7.6 CONFIDENCE INTERVAL PROCEDURES FOR TWO VARIANCES

The simulation experiments conducted in the case of single variance can be readily modified for the case of two variances. Let X_{1i}, $i = 1,2,\ldots,n$, and X_{2j}, $j = 1,2,\ldots,n$, be two independent sequences of correlated random variables with the same mean μ and variances σ_1^2 and σ_2^2 respectively. Both X_{1i} and X_{2j} are generated by the same moving average process in equation (5.2.3.1) with variances σ_1^2 and σ_2^2 respectively. Following the case of single variance, the 80% confidence interval procedures for two variances can be constructed by applying the method of sequential systematic sampling and the classical F test. The computational forms for such confidence interval analysis can be found in Appendices F and G. The sequences of random variables generated in the simulation experiment follow the same order in deviating from normality as in the case of single variance. The empirical results of the estimated level of confidence are summarized in Table 27. As Table 27 illustrates, the confidence interval procedures for two variances are more sensitive to departures from normalty than those for single variance.

7.7 CONCLUDING REMARKS

The parametric inferential procedures for variances developed under the normality assumption are quite sensitive to departure from normality. The departure is measured by the value of kurtosis. Improvement of the performance of inferential procedures for variances

Table 27

Coverage of 80% Confidence Intervals for Two Variances in Simple Moving Average Models with Order k_t.

$m = 40, 90, 200, \alpha_D = .20$, Replication = 500

			Estimated Confidence Level	
k_t	γ_2	Input Distribution	Sequential Systematic Sampling	Independence and Normality Assumed
4	0	Normal	0.7800	0.5500
5	0	Normal	0.8120	0.5100
8	0	Normal	0.7980	0.3922
4	6	Gamma	0.7940	0.2520
4	12	Gamma	0.7740	0.1740
4	0.1696	Lognormal	0.6620	0.4140
4	1.4746	Lognormal	0.7080	0.3300
4	27.7341	Lognormal	0.8580	0.1720
4	2305137.0	Lognormal	0.9980	0.0700

is gained by revising the formulation of Var($\hat{\sigma}^2$). However, the degree of improvement depends on the magnitude of kurtosis. As the value of kurtosis increases, the performance of the revised inferential procedures for variances becomes disappointing.

CHAPTER VIII

AN ALGORITHM FOR DETERMINING THE ORDER OF SERIAL DEPENDENCE

8.1 INTRODUCTION

It is evident that the method of sequential systematic sampling depends upon identification of the order of serial dependence. In applying the method of sequential systematic sampling the precision with which the analyst predicts the level of confidence or the power of hypothesis tests depends in great part on a suitable choice of the order of serial dependence to reduce the error in estimating $\sigma^2_{\bar{x}}$ and σ^2. Hence, determination of the order of serial dependence is an important part to the implementation of the method of sequential systematic sampling.

Determining the order of serial dependence for the method of sequential systematic sampling is similar to determining the batch size for the method of batch means. In both of the two methods, the purpose of making such determination is to obtain uncorrelated or nearly uncorrelated observations. However, the difference between these two methods lies in the methodology by which the correlation among sample means can be reduced.

For the reported procedures for determining the batch size, the resulting batch means processes may still be subject to strong autocorrelation. This point is also presented in the comparative study in Chapter V. For simple moving average models, sequential systematic sampling is capable of eliminating correlation among observations provided that $k > k_t$, but batch means could not. For AR(1) models, batch means often leads to higher estimator bias than sequential systematic sampling.

If the assumed value of k is sufficiently large, it is likely to collect essentially uncorrelated observations through the method of sequential systematic sampling. This is possible because in most situations the absolute magnitude of the correlation between two observations decreases as the lag between them increases.

As discussed in the literature review on possible approaches to the determination of the order of serial dependence, the one which employs the ARMA modeling will be used to develop an algorithm in this chapter. The purpose of the algorithm is to collect uncorrelated or nearly uncorrelated observations from a sequence of correlated data. The performance of this algorithm will be evaluated in an empirical study for the case of M/M/1 queues.

8.2 SIGNIFICANCE OF TEST OF THE LAG CORRELATION

The adequacy of the method of sequential systematic sampling depends upon the validity of the assumed value of the order of serial dependence. For a sequence of random variables $X_1, X_2, \ldots, X_n$ with mean μ and common variance σ^2 the variance of the sample mean is given by

$$\sigma_{\bar{x}}^2 = \frac{1}{n} [\sigma^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{ij}].$$

In the application of the method of sequential systematic sampling, if the analyst has reason to believe that the order or serial dependence is (k-1), the expected value of the estimator of $\hat{\sigma}_{\bar{x}_k}^2$ is

$$E(\hat{\sigma}_{\bar{x}_k}^2) = \frac{1}{mk} [\sigma^2 + \frac{1}{mk} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \sigma_{ij}] + \frac{T}{mk}$$

where T is given by

$$T = \frac{-2}{m(m-1)} \sum_{j=1}^{m-1} (m-j)\bar{\sigma}_{jk} + \frac{2}{mk} \sum_{h=1}^{k-1} (\varepsilon_\omega(k,h) + \varepsilon_B(k,h))$$

where $\bar{\sigma}_{jk}$, $\varepsilon_\omega(k,h)$ and $\varepsilon_B(k,h)$ are as defined previously. Here T comprises covariance terms with lag greater than or equal to k.

Since one purpose of the method of sequential systematic sampling is to provide an interval estimator for conducting hypothesis tests for means, let H_0: $\mu < \mu_0$ against H_1: $\mu > \mu_0$, and let $1 - \beta_T$ and $1 - \beta_D$ be the actual and design power of the test respectively. Let E be the discrepancy between $1 - \beta_T$ and $1 - \beta_D$. For large sample, E can be shown as

$$E = \int_{Z_{1-\alpha}}^{Z_{1-\alpha}\left(\frac{\sigma_{\bar{x}}}{\sqrt{E(\hat{\sigma}^2_{\bar{x}})}}\right)} e^{-Z^2/2}\, dz. \tag{8.2.1}$$

The magnitude of E will depend on the ratio R given by

$$R = \frac{\sigma_{\bar{x}}}{\sqrt{E(\hat{\sigma}^2_{\bar{x}_k})}} = \frac{\sqrt{\frac{1}{mk}\left[\sigma^2 + \frac{1}{mk} \sum_{i=1}^{mk-1} \sum_{j=i+1}^{mk} \sigma_{ij}\right.}}{\sqrt{\frac{1}{mk}\left[\sigma^2 + \frac{1}{mk} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \sigma_{ij}\right.}}.$$

Now, in the numerator of R,

$$\begin{aligned}\sum_{i=1}^{mk-1} \sum_{j=i+1}^{mk} \sigma_{ij} &= \sum_{j=1}^{mk-1} (mk-j)\sigma_j \\ &= \sum_{j=1}^{k-1} (mk-j)\sigma_j + \sum_{j=k}^{mk-1} (mk-j)\sigma_j.\end{aligned} \tag{8.2.2}$$

For sufficiently large k, it is reasonable to assume that $\sigma_j < \sigma_k$ for $j > k$ under the assumption of strict stationarity. Hence the second summation in equation (8.2.2) is bound by

$$\sum_{j=k}^{mk-1} (mk-j)\sigma_j < \sigma_k \sum_{j=k}^{mk-1} (mk-j)$$

$$= \frac{\sigma_k}{2} [(mk)^2 - mk(2k-1) + k(k-1)]$$

Now, approximate the ratio R by

$$R \approx \frac{\sqrt{\frac{1}{mk} [\sigma^2 + \frac{1}{mk} \{\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \sigma_{ij} + \frac{\sigma_k}{2} [(mk)^2 - mk(2k-1) + k(k-1)]\}}}{\sqrt{\frac{1}{mk} [\sigma^2 + \frac{1}{mk} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \sigma_{ij}] + T}}$$

Observe that the error E in equation (8.2.1) is affected by $(\sigma_k/2)[(mk)^2 - mk(2k-1) + k(k-1))]$ and T, and these two factors go to zero if $\sigma_k = 0$. Therefore it is important to conduct hypothesis tests for testing $\sigma_k = 0$. Another way of reducing the error is to test that beyond the lag k the summation of covariance terms is small. That is

$$\sum_{j=k}^{mk-1} (mk-j)\sigma_j < c$$

where c is a small number.

8.3 MOTIVATION AND METHODOLOGY

Testing $\sigma_k = 0$ or $\sum_{j=1}^{mk}(mk-j)\sigma_j < c$ where c is a small number is the problem of the determination of the order of serial dependence for a sequence of correlated data. As indicated in the literature review there are two different approaches to this problem, and the one based upon the ARMA modeling is suggested. A realization of correlated data can often be specified as an ARMA model. However, more often than not the true parameters of the model are unknown and usually an estimation procedure is used in model fitting to provide estimates of the model parameters. Maximum likelihood (M.L.) estimation is often employed to obtain such estimates. The analyst can use these estimates from model fitting to obtain estimates of the theoretical autocovariances (see Section (2.3.4) for discussion). Hence a theoretical autocovariance estimate is then the function of M.L. estimates and therefore possesses the large sample properties of M.L. estimator.

The estimates of the theoretical autocovariances acquired can be used to determine the order of serial dependence with which sequences of essentially uncorrelated observations can be obtained. The reason is stated as follows: If a fitted ARMA model does represent the underlying random process, then the theoretical autocovariance estimates based upon the model should sufficiently describe the correlation structure of the process. Thus it would be legitimate in applying autocovariance estimates to determine the order of serial dependence for the process.

Given an order of serial dependence determined in a manner described above, two questions arise for the precision of the order of serial dependence. First, does the fitted model adequately describe the behavior of sample data such that the theoretical autocovariance estimates can represent the correlation of the data? Second, could one really obtain a sequence of uncorrelated data from the order of serial dependence determined by the use of the theoretical autocovariance estimates? To answer the first question, the Ljung-Box test is used to test the adequacy of the fitted model as recommended in the literature review. For the second question, the von Neumann ratio test is to be used to test whether a sequence of random variables are uncorrelated or not. The test was used in Fishman (1975) to determine the batch size. If the analyst has determined the order of serial dependence as (k-1), he can then use the von Neumann ratio test to determine whether observations at lag k or greater are uncorrelated.

Based upon the discussion above a methodology is suggested as follows:

1. Identify the correlation structure of sample observations and relate it to an ARMA(p,q) model.
2. Obtain parameter estimates of the assumed model.
3. Check the adequacy of the model by conducting a goodness of fit test to determine whether the assumed model should be accepted or not.
4. Obtain the theoretical autocovariance estimate from the accepted model parameters.
5. Employ the theoretical autocovariance estimate to determine the value of k such that

$$\sigma_k = 0$$

or

$$\sum_{j=k}^{mk-1} (mk-j)\sigma_j < c$$

where σ_j is the autocovariance estimate obtained in step 4.

6. Use the von Neumann ratio test to insure that observations at lag k or greater can be treated as uncorrelated.

A formal presentation of the methodology follows in the next section.

8.4 AN ALGORITHM FOR CORRELATED OBSERVATIONS

Consider a sequence of correlated data $Z_1, Z_2, \ldots, Z_n$ which can be represented by an ARMA(p,q) model. That is

$$Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \qquad (8.4.1)$$

where $\{a_t\}$ is a sequence of i.i.d. random variables with mean zero and variance σ^2.

To determine the order of serial dependence, this research suggests the following algorithm:

1. Generate a sequence of observations of size 1000.
2. Apply the corner method proposed by Beguin and Gourieoux [1979] to the data to determine the order of p and q of a simple ARMA(p,q) model.
3. Apply the maximum likelihood estimation procedure in Box and Jenkins (1976) to obtain the model parameter estimates ($\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p$ and $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_q$).
4. Given a significance level, apply the Box-Ljung test to check the adequacy of the model determined in step 2. If the model is accepted, go to step 5; otherwise return to step 2 and refit the model.
5. Based on the estimated model parameters, apply the algorithm developed by McLeod (1976) to compute the theoretical correlation estimates.
6. The lag k with which the correlation can be deemed as died out is determined if the lag k correlation estimate is less than or equal to a small number c.
7. Take another replication with size mk (m = 10).
8. Sequentially draw observations at intervals of k from the process and obtain a sample of size m.
9. Apply the von Neumann ratio test to the sample of size m to

determine whether the observations in this sample are uncorrelated or not.

10. If the independence of the data is not accepted by the test, increase the value k and go to step 7.

The corner method proposed by Beguin and Gourieroux (1979), gives a simple and general criterion for the determination of the order of p and q for an ARMA(p,q) model. Based upon a prespecified number of sample autocorrelations, this method computes the Δ-array (Δ(i,j)). A realization of the ARMA process is said to have order (p,q) if the Δ-array dies out at i = p and j = q.

With the order of p and q determined, one can obtain the M.L. estimates of the model parameters by the maximum likelihood estimation procedure in Box and Jenkins [1976], which minimizes the sum of squared residuals of the assumed model.

After fitting the model, the Ljung-Box test is used for testing the adequacy of the fitted model.

Given an assumed model for the sample observations, the algorithm by McLeod (1975) can be used to compute the theoretical autocorrelation of the model. Although the theoretical autocorrelation can be derived in the manner described by Box and Jenkins (1976, pp. 75), McLeod's algorithm is exact, easily coded, and can be used for ARMA models of all orders.

The algorithm suggested in this research initially determines the value of k if the lag k theoretical correlation is less than or equal to a user-defined value c (say c = 0.05). This approach is similar to that of Gross and Harris (1974) for determining the batch size.

However, a criticism of their approach is that "correlation estimators are generally biased and for small n are highly variable" (Law (1983)). For this reason, the conventional correlation estimators are not employed to determine the value of k. Instead, correlation estimates are obtained through ARMA modeling and such estimates may be close to the true correlations if the assumed model does describe the behavior of the sample data.

The purpose of this algorithm is to provide sequences of essentially uncorrelated data from a realization of correlated data. The von Neumann ratio test is suited for determining whether a sequence of observations are uncorrelated of not. Hence the test is applied to a systematic sample $Z_i, Z_{i+k}, \ldots, Z_{i+(m-1)k}$ where i can be chosen from 1 to k, to determine whether these observations are uncorrelated. If the null hypothesis of independence is rejected, increase the value k and apply the test to another systematic sample formed by a larger value of k. The rationale for this procedure is the following: As the lag k increases one would expect the correlation among observations in a systematic sample to decrease. Eventually the test fails to detect a departure from independence for a certain value of k. The manner of application of the von Neumann ratio test here is similar to that of Fishman (1975). However, this algorithm uses the lag correlation estimate to assist determination of a proper value of k. Therefore the way of applying the von Neumann ratio test to detect independence of the data in this algorithm may work better than Fishman's procedure.

8.5 APPLICATION OF THE SUGGESTED ALGORITHM

A computer program is developed to implement the algorithm suggested. The algorithm is applied to M/M/1 queueing systems. A simulation model is developed for the systems and generates observations of the waiting time. To avoid the startup problem the simulation study follows Fishman's approach (see Fishman (1975)): the first 459 observations of waiting time were deleted from the simulation output. The algorithm is then applied to the remaining observations to obtain the theoretical autocorrelation estimates for the determination of the order of serial dependence.

In examining the performance of the algorithm, a key question is how well the theoretical autocorrelation estimates approximate the corresponding true autocorrelations such that the order of serial dependence thereby acquired can render a sequence of uncorrelated observations. Resolving this problem may require the true autocorrelation of the model to assess precision of the autocorrelation estimates. Fortunately, Daley (1968) provided analytical results of serial correlation coefficients for the waiting time for M/M/1 queueing systems.

Performance of the algoriothm developed is measured in terms of the discrepency between the estimated and the true autocorrelations, and in terms of the efficiency of the algorithm in obtaining a sequence of uncorrelated data. If the magnitude of the discrepency is considered significant, the accuracy of the order of serial dependence might be affected. Furthermore, in most situations the true autocorrelation of the model is unknown. Therefore, it is very unlikely that the analyst can conduct the comparison of autocorrelation for every simulation model to be dealt with. An alternative to examining the accuracy of the order of serial dependence is to consider the efficiency in applying the von Neumann ratio test. If many iterations of the von Neumann ratio test are required to obtain the order of serial dependence or the test fails to determine the order of serial dependence, then the autocorrelation estimates acquired through the algorithm may be significantly biased.

Now, define ρ_n as the serial correlation coefficient of the waiting time in an M/M/1 queueing system with lag n. The formula of ρ_n is given in Daley (1968) by

$$\rho_n = \frac{(1-\rho)^3(1+\rho)}{2\pi 7^3(2-\rho)} \int_0^a \frac{t^n t^{1/2}(a-t)^{1/2}}{(1-t)^3}\, dt$$

where

$$\rho = \text{traffic density} = \frac{\text{Mean Service Time}}{\text{Mean Interarrival Time}}$$

$$a = \frac{4\rho}{(1+\rho)^2}, \quad 0 < a < 1.$$

The results of the estimated and the true autocorrelations of the waiting time observations are summarized in Tables 29, 30, and 31.

Table 28 gives the comparison between the estimated and the true autocorrelation with $\rho = 0.5$. The waiting time process generated from the M/M/1 queueing model is identified as the AR(1) model, with $\hat{\phi}_1 = 0.7516$, also $\chi_p^2 = 0.2762$, where $\chi_p^2 = \Pr\{\chi^2 < \chi_{exp}^2\}$, with χ_{exp}^2 as the sample chi-square statistic. From Table 29, one can observe that the estimated correlation coefficients approximate the corresponding true correlation coefficients very well. Because the traffic density is relatively low for a value of 0.5, both the estimated and the true correlation coefficients die out quickly at about $n = 27$. There is

Table 28

Determination of the Lag Correlation in the M/M/1 Queue with $\rho = 0.5$ and $c = 0.06$.

$\hat{\phi}_1 = .7156$, $\hat{\phi}_2 = .0713$, $\chi_p^2 = .2762$

n	$\rho(n)$	$\hat{\rho}(n)$
1	0.77595	0.78215
2	0.61698	0.62083
3	0.49629	0.50081
4	0.40322	0.41416
5	0.33010	0.33542
6	0.27193	0.28031
7	0.22518	0.22909
8	0.18728	0.18657
9	0.15634	0.15194
10	0.13095	0.12373
11	0.11000	0.10077
12	0.09265	0.08206
13	0.07821	0.06683
14	0.06616	0.05442
15	0.05608	0.04432
16	0.04761	0.03609
17	0.04049	0.02939
18	0.03449	0.02394
19	0.02941	0.01950
20	0.02512	0.01588
21	0.02148	0.01293
22	0.01838	0.01053
23	0.01575	0.00857
24	0.01351	0.00698
25	0.01160	0.00569
26	0.0100	0.00463
27	0.0086	0.00377

$\sum_{n=0}^{10} \rho_n = 4.59794$ $\qquad$ $\sum_{n=0}^{10} \hat{\rho}_n = 4.6241$

$\sum_{n=0}^{\infty} \rho_n = 5.33333$ $\qquad$ $\sum_{n=0}^{\infty} \hat{\rho}_n = 5.22782$

Table 29

Determination of the Lag Correlation in the M/M/1 Queue with $\rho = 0.8$ and $c = 0.0009$.

$\hat{\phi}_1 = .97045,\ \chi_p^2 = .3047$

n	Theoretical $\rho(n)$	Estimated $\hat{\rho}(n)$
1	0.96266	0.97045
2	0.92883	0.93055
3	0.89749	0.90529
4	0.86821	0.87821
5	0.84068	0.85616
6	0.81470	0.82297
7	0.79007	0.80293
8	0.76667	0.78845
9	0.74438	0.75284
10	0.72310	0.72804
11	0.70726	0.71239
12	0.68327	0.70421
13	0.66458	0.67821
14	0.64664	0.65053
15	0.62938	0.63412
16	0.61279	0.62345
17	0.59680	0.60981
18	0.58140	0.59010
19	0.56654	0.58000
20	0.55219	0.55694
21	0.53834	0.53821
22	0.52495	0.52241
23	0.51200	0.51001
24	0.49948	0.48048
25	0.48735	0.47035
26	0.47560	0.46990
27	0.46423	0.46717
28	0.45319	0.45447
29	0.44250	0.44049
30	0.43213	0.42003
31	0.42206	0.41828
32	0.41229	0.41682
33	0.40280	0.40566
34	0.39358	0.40479
35	0.38462	0.40419
36	0.37592	0.39386
37	0.36475	0.38380
38	0.35923	0.37399
39	0.35122	0.36444
40	0.34343	0.35513

Table 29 (Continued)

n	Theoretical $\rho(n)$	Estimated $\hat{\rho}(n)$
176	0.02662	0.01051
177	0.02617	0.01024
178	0.02573	0.00998
179	0.02530	0.00973
180	0.02488	0.00948
181	0.02446	0.00924
182	0.02405	0.00900
183	0.02365	0.00877
184	0.02326	0.00855
185	0.02287	0.00833
186	0.02249	0.00811
187	0.02212	0.00791
188	0.02175	0.00771
189	0.02139	0.00751
190	0.02103	0.00732
191	0.02068	0.00713
192	0.02034	0.00695
193	0.02000	0.00677
194	0.01967	0.00660
195	0.01935	0.00643
196	0.01903	0.00626
197	0.01872	0.00610
198	0.01841	0.00595
199	0.01810	0.00580
200	0.01781	0.00565
201	0.01751	0.00550
202	0.01723	0.00536
203	0.01694	0.00523
204	0.01667	0.00509

$$\sum_{n=0}^{10} \rho_n = 9.33957 \qquad \sum_{n=1}^{10} \hat{\rho}_n = 9.44589$$

$$\sum_{n=0}^{\infty} \rho_n = 41.8566 \qquad \sum_{n=0}^{\infty} \hat{\rho}_n = 40.13521$$

Table 30

Determination of the Lag Correlation in the M/M/1 Queue with $\rho = 0.9$ and $c = 0.026$.

$\hat{\phi}_1 = .99450$, $\chi_p^2 = .7088$

n	Theoretical $\rho(n)$	Estimated $\hat{\rho}(n)$
1	0.99016	0.99450
2	0.98098	0.98903
3	0.97210	0.98359
4	0.96350	0.97608
5	0.95512	0.97080
6	0.94695	0.96050
7	0.93898	0.95994
8	0.93118	0.95000
9	0.92354	0.94981
10	0.91605	0.94634
11	0.90871	0.94114
12	0.90150	0.93596
13	0.89442	0.93081
14	0.88747	0.92569
15	0.88062	0.92060
16	0.87389	0.91554
17	0.86727	0.91050
18	0.86074	0.90549
19	0.85431	0.90051
20	0.84798	0.89556
21	0.84174	0.89064
22	0.83559	0.88574
23	0.82952	0.88087
24	0.82354	0.87602
25	0.81763	0.87120
26	0.81180	0.86641
27	0.80605	0.86165
28	0.80037	0.85691
29	0.79476	0.85219
30	0.78992	0.84751
31	0.78375	0.84285
32	0.77835	0.83821
33	0.77300	0.83360
34	0.76772	0.82901
35	0.76250	0.82445
36	0.75734	0.81992
37	0.75224	0.81541
38	0.74720	0.81093
39	0.74221	0.80647
40	0.73727	0.80203

Table 30 (Continued)

n	Theoretical $\rho(n)$	Estimated $\hat{\rho}(n)$
356	0.15388	0.14037
357	0.15324	0.13960
358	0.15259	0.13883
359	0.15195	0.13807
360	0.15131	0.13731
361	0.15068	0.13656
362	0.14900	0.13580
661	0.04262	0.02610
662	0.04245	0.02596
663	0.04227	0.02582
664	0.04209	0.02568
665	0.04191	0.02554
666	0.04174	0.02539
667	0.04156	0.02526
668	0.04139	0.02512
669	0.04121	0.02498
670	0.04104	0.02484
671	0.04087	0.02470
672	0.04070	0.02457
673	0.04052	0.02443
674	0.04035	0.02430
675	0.04019	0.02417
676	0.04002	0.02403
677	0.03985	0.02390
678	0.03968	0.02377
679	0.03951	0.02364

$$\sum_{n=0}^{10} \rho_n = 10.52122 \qquad \sum_{n=0}^{10} \hat{\rho}_n = 10.6786$$

$$\sum_{n=0}^{\infty} \rho_n = 181.877 \qquad \sum_{n=0}^{\infty} \hat{\rho}_n = 181.81818$$

little difference in sums of those correlation coefficient terms between the estimated and the true autocorrelations. For the range of n from 0 to 10, the absolute value of difference is 0.02616. And for the range of n from 0 to ∞, the absolute value of difference is 0.10551. Given c = 0.06, the value k was chosen as 25. With a sample size 250 (m = 10, k = 25) the von Neumann ratio test accepted that observations sampled at intervals of 25 are uncorrelated.

Table 29 does the comparison between the estimated and the true autocorrelations with $\rho = 0.8$. The waiting time process generated is identified as the AR(1) model with $\hat{\phi}_1 = 0.97045$ and $\chi_p^2 = 0.3045$. As indicated in Table 30, the estimated correlation coefficients approximate their corresponding true values very well, while the values of both the estimated and the true correlation coefficients are persistently large. Unlike the queueing system with $\rho = 0.5$, this time both the estimated and the true correlation coefficients die out at about n = 200. The absolute value of difference for n from 0 to 10 is 0.10632. And the absolute value of difference from n from 0 to ∞ is 1.53146. In dealing with such a congested queueing system, an analyst may need to choose the order of serial dependence high enough to compensate the consistently large correlation coefficients. Given c = 0.009, the value k was chosen as 182. With the sample size 1820 (m = 10, k = 182) the von Neumann ratio test accepted that observations sampled at intervals of 182 are uncorrelated.

Table 30 carries out the comparison between the estimated and the true autocorrelations with $\rho = 0.9$. The waiting time process generated is identified as the AR(1) model with $\hat{\phi}_1 = 0.99450$ and $\chi_p^2 = 0.7088$.

For $\hat{\phi}_1 = 0.99450$, the AR(1) model is on the margin of stationarity assumption. The estimated correlation coefficients still trace the true correlation coefficients very well. Both the estimated and the true correlation coefficients die out at about n = 679. The absolute value of difference for n from 0 up to 10 is 0.15738, and 0.059 for n from 0 to ∞. Given c = 0.026, the value k was chosen as 662. With a sample size 6620 (m = 10, k = 662) the von Neumann ratio test accepted that observations sampled at intervals of 662 are uncorrelated.

In addition to the suggested algorithm, the true correlation ρ_j in the cases of M/M/1 queue is estimated by the conventional correlation estimate $\hat{\gamma}_j$ which is given by

$$\hat{\gamma}_j = \frac{1}{n-j} \sum_{i=1}^{n-j} (X_i - \bar{X})(X_{i+j} - \bar{X}).$$

Coupled with the correlation estimates yielded by the suggested algorithm, the results in using the conventional correlation estimates are shown graphically in Figures 12, 13, and 14. As the results in Figure 12 indicate, the conventional correlation estimates perform almost equally as those of the suggested algorithm. However, based upon the results presented in Figures 13 and 14, the conventional correlation estimates are subject to significant bias as the lag increases.

8.6 CONCLUDING REMARKS

Two conclusions are drawn for the performance of the algorithm developed in this chapter. First, the theoretical autocorrelation estimate acquired by the algorithm provides a good approximation to the corresponding true autocorrelation in the case of M?M/1 queues.

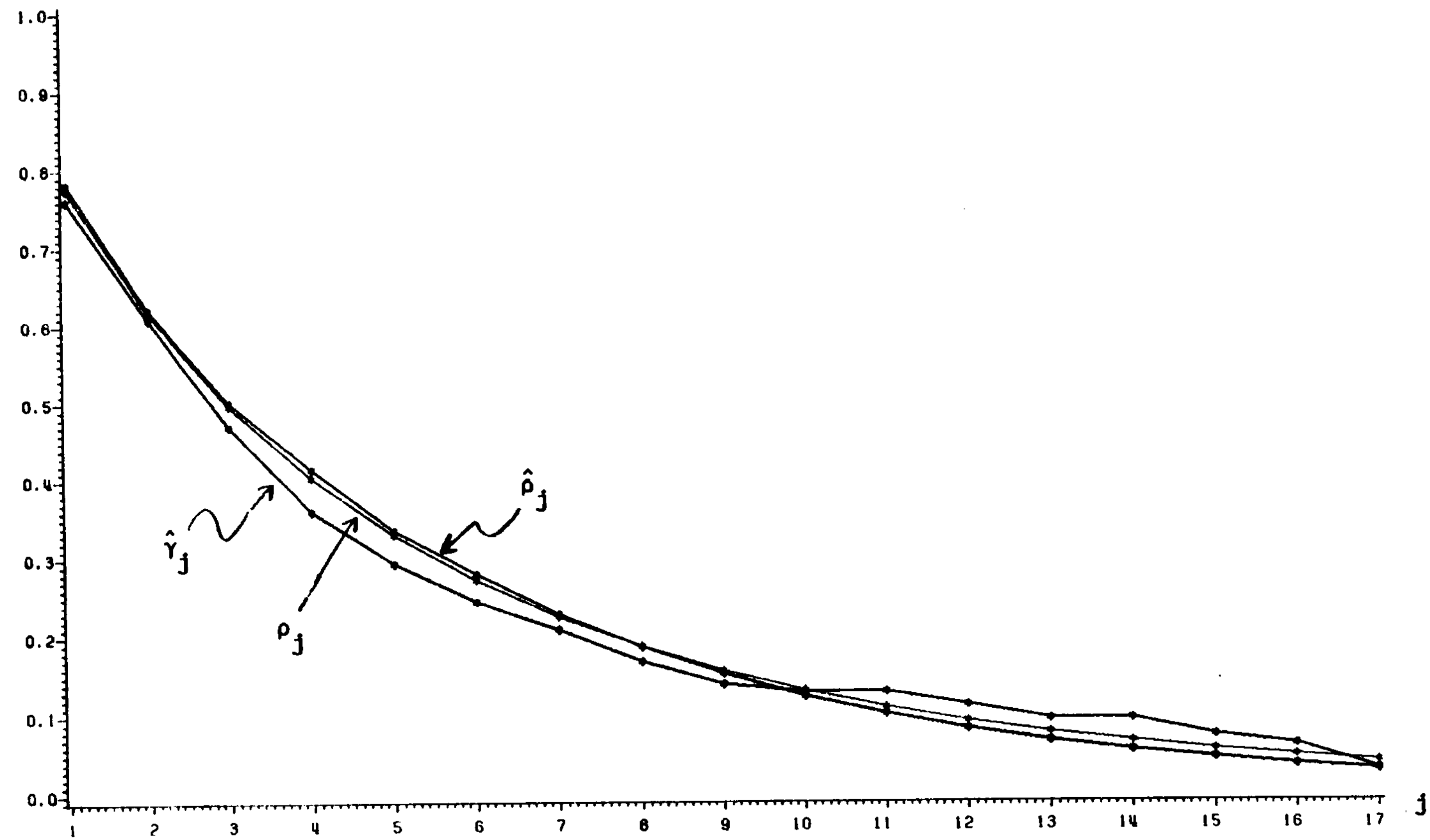

Figure 12

Results of the Suggested Algorithm and the Conventional Procedure in Estimating Correlations of the M/M/1 Queue with $\rho = 0.5$, $\rho_j \equiv$ True Correlation at Lag j, $\hat{\rho}_j \equiv$ Estimated Correlation by the Suggested Algorithm, $\hat{\gamma}_j \equiv$ Estimated Correlation by the Conventional Procedure.

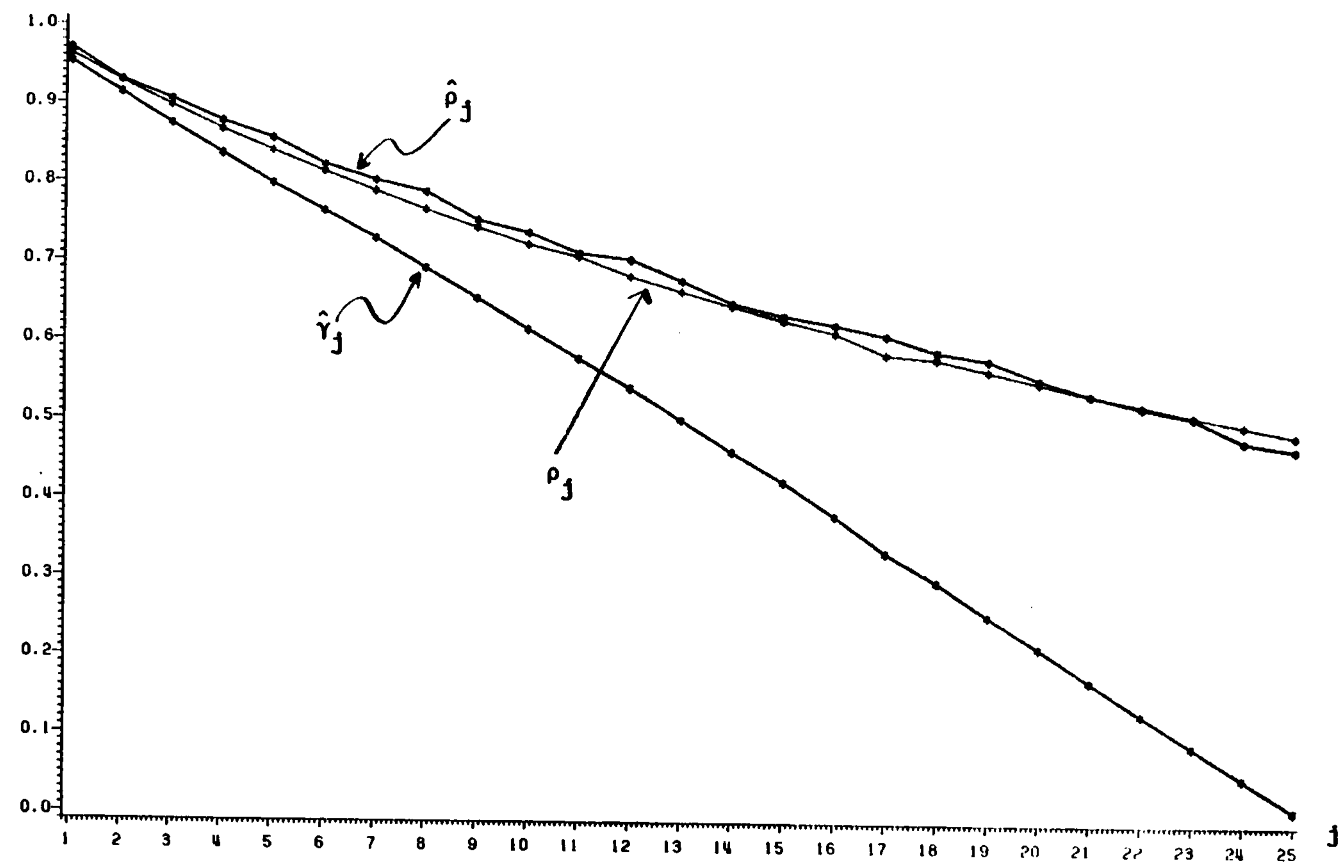

Figure 13

Results of the Suggested Algorithm and the Conventional Procedure in Estimating Correlations of the M/M/1 Queue with $\rho = 0.8$, $\rho_j \equiv$ True Correlation at Lag j, $\hat{\rho}_j \equiv$ Estimated Correlation by the Suggested Algorithm, $\hat{\gamma}_j \equiv$ Estimated Correlation by the Conventional Procedure.

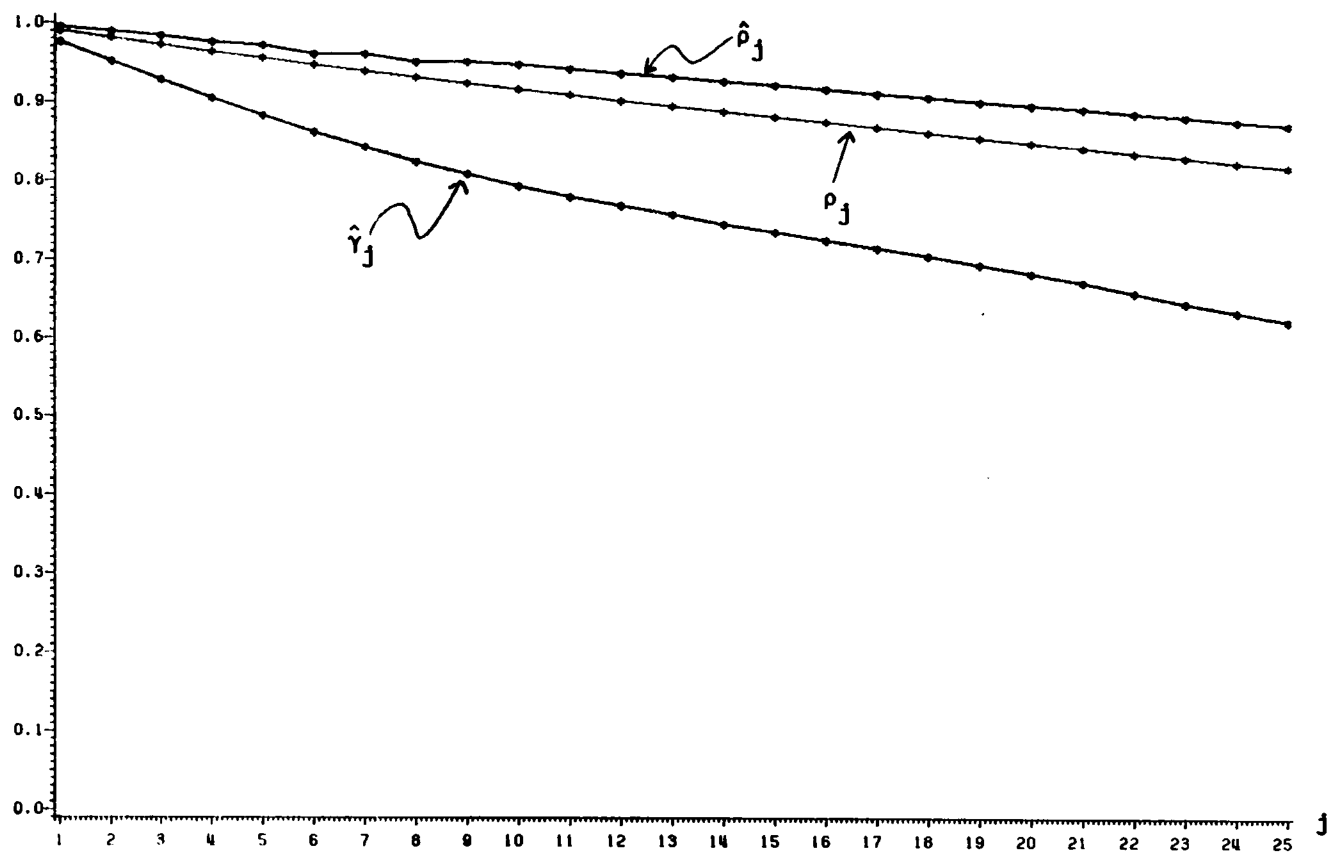

Figure 14

Results of the Suggested Algorithm and the Conventional Procedure in Estimating Correlation of the M/M/1 Queue with $\rho = 0.9$, $\rho_j \equiv$ True Correlation at Lag j, $\hat{\rho}_j \equiv$ Estimated Correlation by the Suggested Algorithm, $\hat{\gamma}_j \equiv$ Estimated Correlation by the Conventional Procedure.

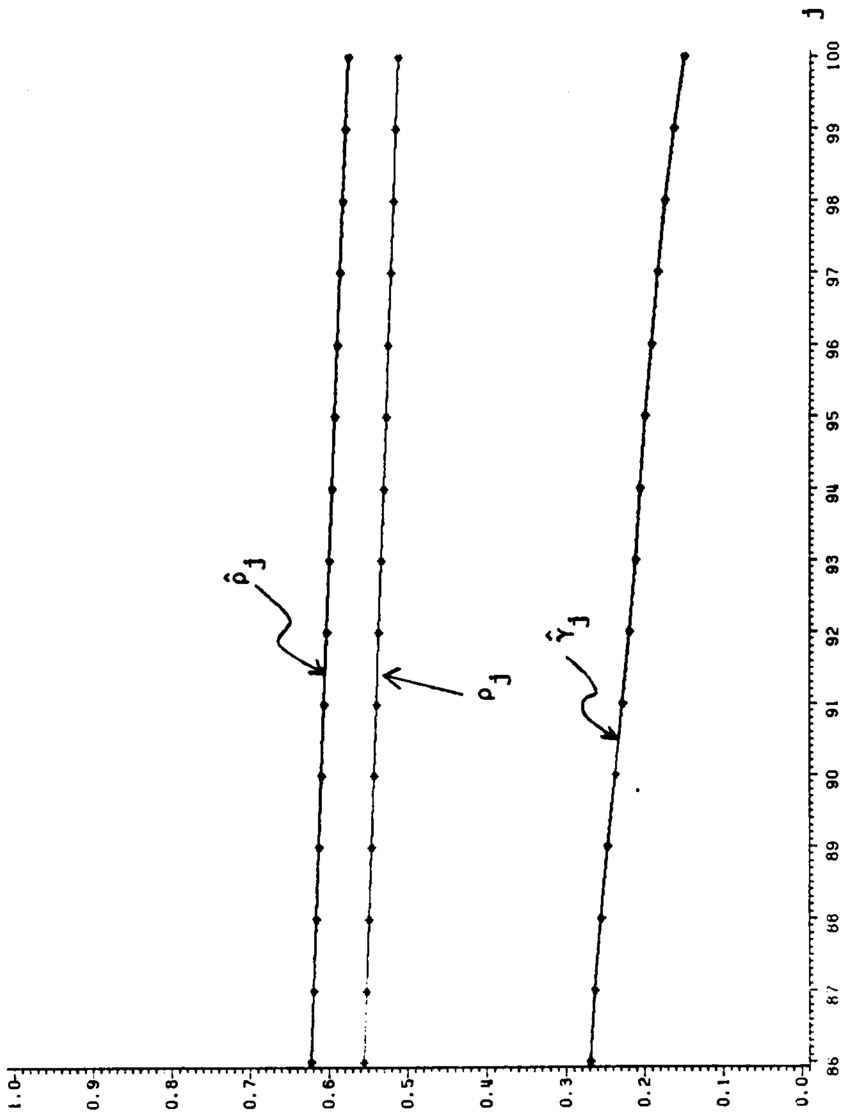

Figure 14 Continued

Second, as a result of the good approximation, the order of serial dependence can be determined in an efficient manner.

As indicated in the literature review, Fishman (1975) reported severe underestimation of the variance of the sample mean and failure to determine batch size in applying the von Neumann ratio test to observations generated from the M/M/1 queueing models with $\rho = 0.8$ and 0.9. The major source of error comes from the correlation among batch means. For heavily correlated output processes, the von Neumann ratio test has difficulty in determining an appropriate batch size since the resulting batch means processes are often highly correlated. To avoid the correlation among observations, the algorithm developed in this research employs the lag correlation of the data to assist the sampling of uncorrelated observations. Thus, in applying the sampling procedure suggested in the algorithm, the systematic data sequence often comprises essentially uncorrelated observations and therefore accepted by the von Neumann ratio test.

CHAPTER IX

COMPARISON OF METHODS FOR MEANS AND VARIANCES IN M/M/1 QUEUEING MODELS

9.1 INTRODUCTION

The ultimate purpose of devising statistical inferential procedures is to provide information about the precision of the parameters concerned (means or variances) of the simulated system. By improving the predictability of the inferential procedure applied, the information provided by the analysis will lead to an increase of the likelihood that the correct conclusion can be drawn.

In the broad sense, all existing methods suggested for the remedy of the classical procedures for means and variances are asymptotically correct, including the ones proposed in this research. However, in typical fashion of conducting a simulation study, excessively large sample sizes may be either prohibited by cost considerations or restricted by the nature of the system under study. It is for such reasons that the simulation analyst should consider the relative performance among methods available in terms of predictability when he must deal with a limited sample size. One way to evaluate the relative performance of the methods concerned is to apply them on a common ground to the data and compare their resulting predictability. In the past, various empirical studies for M/M/1 queueing systems have been conducted and were discussed in the literature review. This research will also consider the case of M/M/1 queues for the comparison of the following methods for means:

1. The classical method.
2. The method of batch means.

3. The method of sequential systematic sampling.

The comparison is pursued in two steps:

1. Compare the predictability of the inferential procedure applied for the three methods for a fixed sample size where batch means and sequential systematic sampling share the same values of m and k.
2. For a fixed sample size where the optimal batch size is determined, compare the method of batch means with the method of sequential systematic sampling for the same sample size where the order of serial dependence is identified by the proposed algorithm.

Considering again the case of M/M/1 queues, this research will compare the following two methods for variances:

1. The classical chi-square test.
2. The method of sequential systematic sampling.

Similar to the case of means, the comparison is presented for fixed sample sizes and optimal sample sizes.

9.2 EMPIRICAL RESULTS

The random process simulated is that of the system time for the M/M/1 queue with arrival rate λ and service rate μ, which give the traffic density $\rho = \lambda/\mu$. The system time in this case is exponentially distributed in steady state. Based upon the given values of λ and μ, the mean and variance of the system time can be determined (see Gross and Harris (1974)). For each M/M/1 queueing model specified by the traffic density, an $100(1 - \alpha)\%$ confidence interval is constructed for the true mean system time as well as for the variance of the system time. The sample size n employed is given by $n = mk$ where m and k are as previously defined. The value of k will be first determined in an arbitrary manner and then determined by the suggested algorithm. As suggested by Fishman (1975) to avoid the problem of the initial transient, the simulated process deletes the first 459 observations. The empirical results obtained in the experiment are summarized in Tables 31, 32, 33, 34, 35, and 36.

Table 31 presents the results of the comparative study when the fixed-sample-size procedures for means are applied. As expected, the method of independence assumed posesses the poorest performance in all cases of traffic density. The table also indicated that, for a given value of k, the difference of coverage between batch means and sequential systematic sampling increases as the traffic density multiplies, which is shown in the cases of $\rho = 0.8$, 0.9, and 0.95. This leads to the conclusion that the method of sequential systematic sampling achieves a greater predictability among the three methods applied with fixed sample sizes.

Table 32 presents the results in comparing the two methods for variances in the case of fixed sample sizes. As the table suggested, sequential systematic sampling has a overwhelmingly predominance over the classical chi-square test. In addition, for each value of k, the actual coverages of the methods deviated further from that desired as the traffic density increases. The derivaiton is considered far more severe than the case of means. This suggests that the method for variances developed in this research is not robust when the assumption or normality is violated.

Table 31

Coverage of 90% Confidence Intervals Using Fixed Sample Sizes for the Mean System Time in the M/M/1 Queueing Models with Traffic Density ρ.

$m = 20$, Replication = 400, I ≡ Independence Assumed,
B ≡ Batch Means, S ≡ Sequential Systematic Sampling

	$\rho = 0.2$			$\rho = 0.5$			$\rho = 0.8$		
k	S	B	I	S	B	I	S	B	I
10	0.8600	0.8725	0.7200	0.5815	0.7694	0.3133	0.4600	0.3875	0.1425
20	0.8593	0.9095	0.7286	0.8125	0.8050	0.4025	0.6225	0.5425	0.1425
30	0.8850	0.9200	0.7425	0.8650	0.8350	0.4350	0.6850	0.6705	0.1550
40	0.8700	0.9100	0.7125	0.8125	0.8300	0.4175	0.7425	0.6700	0.1250
50	0.8725	0.9025	0.7250	0.8500	0.8600	0.4225	0.7525	0.7075	0.1450
60	0.8650	0.9420	0.5422	0.8225	0.8450	0.4150	0.7300	0.6875	0.1250
70	0.8667	0.9550	0.5513	0.8659	0.9106	0.3166	0.7850	0.7375	0.1650

Table 31 (Continued)

	$\rho = 0.9$			$\rho = 0.95$		
k	S	B	I	S	B	I
10	0.3150	0.2500	0.0800	0.1525	0.1200	0.0375
20	0.3600	0.2750	0.0675	0.1850	0.1525	0.0300
30	0.5275	0.4200	0.0925	0.2225	0.1600	0.0200
40	0.5275	0.4275	0.0575	0.2600	0.1900	0.0150
50	0.5675	0.4675	0.0875	0.3025	0.2200	0.0275
60	0.5739	0.4837	0.0777	0.2925	0.2125	0.0325
70	0.5850	0.4925	0.0600	0.3325	0.2550	0.0275

Table 32

Coverage of 95% Confidence Intervals Using Fixed Sample Sizes for the Variance of the System Time in the M/M/1 Queueing Model with Traffic Density ρ.

$m = 10$, Replication = 150, I ≡ Independence and Normality Assumed, S ≡ Sequential Systematic Sampling

	$\rho = 0.2$		$\rho = 0.3$		$\rho = 0.4$		$\rho = 0.5$		$\rho = 0.8$		$\rho = 0.9$	
k	I	S	I	S	I	S	I	S	I	S	I	S
10	0.480	0.8333	0.4010	0.8000	0.3510	0.7833	0.2750	0.7000	0.1467	0.4800	0.020	0.14
20	0.4467	0.8500	0.3920	0.8833	0.3411	0.8167	0.2125	0.8000	0.1267	0.4933	0.053	0.2400
30	0.4533	0.8333	0.3800	0.8667	0.3000	0.7833	0.1750	0.7500	0.1133	0.5467	0.0200	0.2800

Table 33

Coverage of 80% Confidence Interval for the Method of Batch Means in M/M/1 Queueing Models.

Replication = 60

ρ \ n	2048	4096	8192	16384
0.5	0.8333	0.8136	0.9333	0.9333
0.8	0.6949	0.7627	0.7333	0.8667
0.9	0.2174	0.3966	0.6167	0.6610

Table 34

Proportion of Runs that Failed to Determine a Batch Size for the Method of Batch Means in M/M/1 Queueing Models.

ρ \ n	2048	4096	8192	16384
0.5	0.03	0.0	0.0	0.0
0.8	0.15	0.10	0.08	0.03
0.9	0.35	0.40	0.27	0.17

Table 35

Coverage of 80% Confidence Interval for the Method of Sequential Systematic Sampling in M/M/1 Queueing Models.

Replication = 60

ρ \ n	2048	4096	8192	16384
0.5	0.8000	0.7996	0.8305	0.8305
0.8	0.7797	0.7458	0.8136	0.8136
0.9	0.5652	0.6500	0.7667	0.8049

Table 36

Coverage of 80% Confidence Interval for Variance Yielded by the Method of Sequential Systematic Sampling in M/M/1 Queueing Models.

Replication = 60

ρ \ n	100	200	250
0.2	0.6520	0.7641	0.6940
0.3	0.7241	0.7500	0.6420
0.5	0.5021	0.6149	0.5500

Table 33 presents the results in applying the method of batch means when the optimal batch size is determined by Fishman's procedure in each of the sixty replicates generated. As this table suggested, the resulting predictability yielded by Fishman's procedure is not satisfactory. In addition, Table 34 presents the proportion of simulation runs where the von Neuman ratio test used by Fishman can't determine the optimal batch size. In contrast to the empirical finding due to Fishman's suggested procedure, Table 35 presents the results in applying the method of sequential systematic sampling where the proposed algorithm for determining the order of serial dependence to the same sample observations for which Fishman's procedure is conducted. In this application, the value of k is determined in each of the sixty replicates and the value of m is given by n/k where n is the total sample size. The user-defined value c in testing the lag correlation is taken as 0.001 universally in the empirical study. The results presented in this table suggested that the algorithm proposed in this research leads to an improvement on the predictability of the confidence interval procedures for means.

Table 36 presents the results in applying the method for variances developed in this research to the sample observations while in each replication the value of k is determined by the proposed algorithm. The results demonstrate the effectiveness of the method for variances developed in this research.

9.3 CONCLUDING REMARKS

The inferential procedures for means and variances developed in this research are applied in comparison with other procedures for means and variances for the case of M/M/1 queues. The comparison indicated that the method of sequential systematic sampling leads to a greater predictability of the analysis.

CHAPTER X

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS FOR FUTURE RESEARCH

10.1 INTRODUCTION

The purpose of this research was to present the method of sequential systematic sampling and its applications to the development of inferential procedures for means and variances. The inferential procedures thus developed lead to an improvement on the predictability of the statistical analysis for simulation output. A simulation output process in this research is featured with the data correlation and nonnormality. Observations of the process are assumed to be strict-stationary and to have finite fourth moment. The method of sequential systematic sampling uses the order of serial dependence of the process to provide uncorrelated observations for developing inferential procedures for means and variances. The algorithm for determining the order of serial dependence provides the theoretical underpinning for the method of sequential systematic sampling.

The significance of the problems associated with violation of the assumptions of independence and normality was presented in Chapter III.

The formal presentation of the method of sequential systematic sampling was presented in Chapter IV. To demonstrate the performance of the method developed for means, a comparative study for method of sequential systematic sampling, batch means, and the classical method in the case of AR(1) and simple moving average models was conducted in Chapter V.

The formal presentation of the parametric inferential procedures for variances for the case of correlated data was presented in Chapter

VI. To examine the predictability of the procedures, they were applied to correlated and normally distributed data in the case of simple moving models. A revision of the procedures for variances was presented in Chapter VII to deal with the case where the normality assumption is also violated. To evaluate the predictability of the revised procedures, they were applied to correlated and nonnormally distributed data in the case of simple moving average models.

The algorithm for determining the order of serial dependence was presented in Chapter VIII, and the algorithm was applied to the waiting time observations generated from M/M/1 queueing models to examine the adequacy of the algorithm.

The comparison of the methods for means and variances in the case of M/M/1 queues was presented in Chapter IX. The comparison covers both the cases of fixed and optimal sample sizes.

The purpose of this chapter is to summarize the results and conclusions scattered throughout the dissertation, and then recommend directions for further research.

10.2 RESULTS AND CONCLUSIONS

The five major results of the research effort presented in this dissertation are the following:

1. The method of sequential systematic sampling leads to the development of the estimator of the variance of the sample mean and the variance of the population. The estimator bias in each case will be significantly reduced or eliminated if the assumed order of serial dependence is sufficiently large.
2. The proposed algorithm for determining the order of serial

dependence would lead to the sampling of essentially uncorrelated observations for the method of sequential systematic sampling.

3. The method of sequential systematic sampling leads to a greater reduction of bias in estimating the variance of the sample mean and an improvement on the predictability of the analysis in comparison with the method of batch means and the classical method. The comparison was presented in the case of AR(1), simple moving average, and M/M/1 queueing models.
4. The parametric inferential procedures for variances developed in this research were proven to be quite predictable under the assumption of normality.
5. The magnitude of kurtosis of sample observations has significant impact on the predictability of the inferential procedures for variances.

The first major result is a conclusion reached during the analysis for the method of sequential systematic sampling in Chapter V.

An important purpose of estimating the variance of the sample mean is to construct an interval estimate for means. In constructing such an interval estimate, the methods of sequential systematic sampling devised the estimator through the use of uncorrelated observations. In the past research, the methods of batch means and regeneration cycles also attempt to obtain uncorrelated observations to assist construction of an interval estimate for means. The method of regeneration cycles exploits the existence of a random grouping of observations that leads to independent identically distributed blocks in the course of simulation. This grouping allows one to obtain an interval estimate for means. Details of criticism for this method can be seen in Chapter II. By comparison, the method of batch means and sequential systematic sampling acquire uncorrelated observations in an empirical manner—correlation among observations is considered essentially insignificant by statistical testing procedures and rules of thumb. However, as Schruben (1982) indicated, correlation in the batch means process may never die out. On the contrary, by applying the method of sequential systematic sampling, uncorrelated observations can be drawn from an output process if the value of k is sufficiently large.

In estimating the population variance, both regeneration cycles and batch means do not directly yield a variance estimator. A variance estimator may be obtained from the procedure suggested by Law (1983). However, the precision of his variance estimate has not been empirically tested.

The second major result is a conclusion reached during the analysis of the application of the algorithm for determining the order of serial dependence in Section (8.5) of Chapter VIII.

In determining the order of serial dependence, Priestley (1982), Box and Jenkins (1976) suggested the approach of asymptotic normality for sample autocorrelations. The determination of the order of serial dependence is also similar to the way that an "acceptable" batch size is determined for applying the method of batch means. For determining an appropriate batch size, Cross and Harris (1974) suggested that a batch size may be acceptable if the lag 1 correlation of the batch means process is less than a small number. These two approaches for determining the order of serial dependence may be crippled by the fact that correlation estimators are generally biased and for small samples are highly variable. Another procedure for determining batch size, suggested by Fishman (1975), is to use von Neumann's test for detecting the presence of correlation among batch means. However, this procedure has not been successful for sample observations which are subject to strong positive autocorrelation in the case of M/M/1 queues.

As developed in this research, the algorithm for determining the order of serial dependence obtains correlation estimates through the ARMA modeling of the data sequence. This algorithm also applies the von Neumann's test to vindicate the independence of a systematic sample drawn from an output process. For heavily congested queueing systems, the empirical study conducted in this research indicated that the correlation estimates acquired by the algorithm are near the true correlations. Therefore, the observations in the systematic sample are often found to be uncorrelated by the von Neumann ratio test.

The third major result is supported by the comparative studies conducted in Chapters V and IX. The fourth result follows from application of the inferential procedures for variances to correlated and normally distributed data in the case of simple moving average models in Section (6.4) of Chapter VI.

Few inferential procedures for variances are reported in the literature. Law (1983) suggested a confidence interval procedure for variances but it only obtains the lower bound of the desired level of confidence. The inferential procedures developed in this research are more complete in the sense that precise predictions can be made with respect to the level of confidence and the power of hypothesis tests.

The fifth result is a consequence of the application of the revised inferential procedures for variances to correlated and nonnormally distributed data in Sections (7.4) and (7.5) of Chapter VII and in Section (9.2) of Chapter IX. This result agrees with the results reported in the past research for the case of independent data.

10.3 RECOMMENDATIONS FOR FURTHER RESEARCH

The method of sequential systematic sampling provides a sampling procedure which contributes to the improvement of the predictability in applying statistical inferential procedures to correlated and nonnormally distributed data. For the research led by this sampling procedure and presented in this dissertation, there are five extensions. First, a research effort is needed to resolve the problem of the initial transient. The importance of this problem can be assessed from the following two perspectives:

1. An inferential procedure often assumes the existence of stationarity for the data and may not be robust if the assumption is violated.
2. Typically the assumption of steady state is used to derive the system parameters (for example: mean and variance of the system time) and the distribution of a system variable. Thus, the empirical study which assumes the knowledge based upon the steady-state assumption may not be able to produce the results which could represent the actual performance of the method applied should the assumption was violated.

Law (1983) summarizes several approaches to the problem of the initial transient. The research concerning this problem may start from these approaches and select one or devise a new approach toward the solution of the problem.

To further examine the effectiveness of the method of sequential systematic sampling, the method should be applied to inventory models. This will help to illustrate the capability of the method of sequential systematic sampling in dealing with real-world systems.

Third, the method for variances developed in this research is not robust when the assumption of normality is violated. In addition, the performance of the method deteriorates from single variance to two variances. Thus, further reserach ought to construct inferential procedures for mutiple variances, with special attention paid to the robustness of the procedures constructed. This is indeed an arduous task to be performed. The relevant research completed in this dissertation may be considered as a stepping stone for such task.

Fourth, the method for variances suggested in this research will be assisted if the value of kurtosis of the simulated process is known. This may be a luxury in real-world simulation studies. In this respect, further research should pursue a multivariate chi-square test applicable to a sequence of correlated data to determine the distribution form of the underlying process.

Fifth, the method of variances proposed in this research suffers from computation inefficiency. The estimation of the variance of the population variance demands a considerable amount of computer time, thus imposing a constraint if the computer resources are scarse.

Research is needed to reduce the execution time for the method, perhaps by approximating the variance estimator of the variance.

Sixth, testing the order of serial dependence may be conducted by Anderson's procedure (Chapter 6, Anderson (1971)). However, his procedure requires knowledge of the distribution of serial correlation.

REFERENCES

Anderson, R. L., "Distribution of the Serial Correlation Coefficient", Annals of Math. Stat., Vol. 13, 1942, pp. 1-13.

Anderson, T. W., "On the Theory of Testing Serial Correlation", Skand. Aktuarientidsakr., Vol. 31, 1948, pp. 88-116.

Anderson, T. W., An Introduction to Multivariate Statistical Analysis, New York: Wiley, 1958.

Anderson, T. W., The Statistical Analysis of Time Series, New York: Wiley, 1971.

Andrews, R. W. and T. J. Schriber, "Two ARMA-Based Confidence Interval Procedures for the Analysis of Simulation Output", Working Paper, Graduate School of Business Administration, University of Michigan.

Barlette, M. S., "On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series", Jour. Royal Stat. Soc., B8, Vol. 28, 1946.

Jean-Marc Beguin, Christian Gourieroux, Alain Monfort, "Identification of a Mixed Autoregressive-Moving Average Process: The Corner Method", Time Series, ed. O. D. Anderson, 1979, pp. 423-36, North-Holland, Amsterdam.

Bhattacharya, R. N., "On Errors of Normal Approximation", Annals of Prob., Vol. 3, No. 3, 1975, pp. 815-828.

Billingsley, P., Convergence of Probability Measures, New York: Wiley, 1968.

Box, G. E. P., "Non-normality and Test on Variances", Biometrika, Vol. 40, 1953, pp. 318-335.

Box, G. E. P. and D. A. Pierce, "Distribution of Residual Autocorrelation Integrated Moving Time Series Models", Jour. Amer. Stat. Assoc., Vol. 64, 1970.

Box, G. E. P. and G. M. Jenkins, Time Series Analysis: Forecasting and Control, Revised ED., San Francisco: Holden-Day, 1976.

Clarke, G. M., "Confidence Intervals Using Autocovariances for Simulation Performance Measures", Dept. of Industrial Engineering, Ohio State University.

Clarke, B. R. and E. J. Godolphin, "Comparative Power Studies for Goodness of Fit Tests of Time Series Models", Journal of Time Series Analysis, Vol. 23, 1975, pp. 141-151.

Cramer, H., Mathematical Methods of Statistics, Princeton University Press, 1946.

Crane, M. A. and D. L. Iglehart, "Simulation Stable Stochastic Systems; 1, General Multiserver Queues", J. Assoc. Comput., Vol. 21, 1974a, pp. 103-113.

Crane, M. A. and D. L. Iglehart, "Simulating Stable Stochastic Systems; 2, Markov Chains", J. Assoc. Comput., Vol. 21, 1974b, pp. 114-123.

Crane, M. A. and D. L. Iglehart, "Simulating Stable Stochastic Systems; 3, Regenerative Processes and Discrete Event Simulation", Opns. Res., Vol. 23, 1975, pp. 33-45.

Daley, D. J., "The Theoretical Correlation Coefficients of Waiting Times in a Stationary Single Queue", Journal of the Australlin Mathematical Society, Vol. 8, 1968, pp. 683-699.

Daniels, H. E., "The Approximate Distribution of Serial Correlation Coefficients", Biometrika, Vol. 43, 1956, pp. 169-185.

Daves, N., C. M. Triggs and P. Newbold, "Significance Levels of the Box-Pierce Portmanteau Statistics in Finite Samples", Biometrika, Vol. 64, NO. 3, 1977, pp. 517-522.

Daves, N. and P. Newbold, "Some Power Studies of a Portmanteua Test of Time Series Model Specification", Biometrika, Vol. 66, No. 1, 1979, pp. 153-155.

Dixon, W. J., "Further Contributions to the Problem of Serial Correlation", Annals of Math. Stat., Vol. 15, 1944, pp. 119-144.

Ducket, S. D. and A. A. B. Pritsker, "Examination of Simulation Output Using Spectral Methods", Math. Comput. Simulation, Vol. 20, 1978, pp. 53-60.

Fisher, R. A., "On the Probable Error of a Coefficient of Correlation Deduced form a Small Sample", Metron, Vol. 1, No. 4, 1921, pp. 3-32.

Fishman, G. S., Spectral Methods in Econometrics, Harvard University Press, Cambridge, Mass. 1969.

Fishman, G. S., "Estimating Sample Size in Computer Simulation Experiments", Mgnt. Sci., Vol. 18, 1971, pp. 21-38.

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

Fishman, G. S., "Batches Means in Digital Simulation", Technical Report No. 75-7, August 1975.

Fishman, G. S., Principles of Discrete Event Simualtion, New York: Wiley, 1978.

Fraser, D. A. S., Nonparametric Methods in Statistics, Wiley, New York, 1957.

Fuller, W. A., Introduction to Statistical Time Series, New York: Wiley, 1976.

Gayen, A. K., "The Frequency Distribution of the Product-Moment Correlation Coefficient in Random Samples of Any Size Drawn from Non-Normal Universes", Biometrika, Vol. 37, 1950, pp. 236-255.

Glynn, P. W., "Asymptotic Theory for Nonparametric Confidence Intervals", Technical Report 63, 1982, Dept. of Operations Research, Stanford University.

Godolphin, E. J., "A Method for Testing the Order of an Autoregressive-Moving Average Process", Biometrika, Voo. 67, No. 3, 1980, pp. 699-703.

Godolphin, E. J. and M. De Tullio, "Invariance Properties of Uncorrelated Residual Transformation", J. R. Statist. Soc. B, Vol. 40, NO. 3, 1978, pp. 313-321.

Goldsman, D. and L. W. Schruben, "Asymptotic Properties of Some Confidence Interval Estimators", Technical Report 544, School of Operations Research and Industrial Engineering, Cornell University.

Gray, H. L., G. D. Kelley and D. McIntire, "A New Approach to ARMA Modeling", Communications in Statistics, B7, 1978, pp. 1-177.

Gray, H. L. and W. A. Woodard, "Application of S-Arrays to Seasonal Data", Applied Time Series Analysis II, New York: Academic Press, 1981.

Gross, D., and C. M. Harris, Fundamentals of Queueing Theory, New York: Wiley, 1974.

Hannan, E. J., "Exact Test for Serial Correlation", Biometrika, Vol. 42, 1955, pp. 135-142.

Hannan, E. J., Multiple Time Series, New York: Wiley, 1970.

Hiedelberger, P. and P. P. Welch, "Adaptive Spectral Methods for Simulation Output Analysis", IBM J. Res. Develop., Vol. 25, 1981a, pp. 860-876.

Heidelberger, P. and P. D. Welch, "A Spectral Method for Confidence Interval Generation and Run Length Control in Simulations", Comm. Assoc. Comput., Vol. 24, 1981b, pp. 233-245.

Hiedelberger, P. and P. D. Welch, "Simulation Run Length Control in the Presence of Initial Transient", Opns. Res., Vol. 31, 1983, pp. 1109-1144.

Iglehart, D. L., "Simulating Stable Stochastic System 5: Comparison of Ratio Estimators", Nav. Res. Logist. Q., Vol. 22, 1975, pp. 553-565.

Johnson, N. L. and S. Kotz, Continuous Univariate Distributions - 1, New York: Wiley, 1972.

Johnson, N. J., "Modified t Tests and Confidence Interval for Asymmetrical Populations", J. Am. Statist. Ass., Vol. 73, NO. 363, 1978, pp. 536-544.

Jow, Y. L., "An Autoregressive Method for Simulation Output Analysis", Technical Report 21, Dept. of Operations Research, Stanford University.

Kendall, M. G. and A. Stuart, "The Advanced Theory of Statistics", Vol. 1, 3rd ED., New York: Hafner, 1969.

Koopmans, L. H., "The Spectral Analysis of Time Series", New York: Academic Press, 1974.

Koopmans, T., "Serial Correlation and Quadratic Forms in Normal Variables", Annals of Math. Stat., 1942, pp. 14-33.

Law, A. M., 1975, "A Comparison of Two Techniques for Determining the Accuracy of Simulation Output", Technical Report 75-11, Dept. of Industrial Engineering, University of Wisconsin.

Law, A. M., "Confidence Intervals in Discrete Event Simulation: A Comparison of Replication and Batch Means", Naval. Res. Logist. Quart., Vol. 24, 1977, pp. 667-678.

Law, A. M., "Statistical Analysis of Simulation Output Data", Opns. Res., Vol. 31, 1984, pp. 983-1029.

Law, A. M. and J. S. Carson, "A Sequential Procedure for Determining the Length of a Steady-State Simulation", Opns. Res., Vol. 27, 1979, pp. 1011-1025.

Law, A. M. and W. D. Kelton, "Confidence Intervals for Steady-State Simulations; 1. A Survey of Fixed Sample Size Procedures", Technical Report 78-5, Dept. of Industrial Engineering, University of Wisconsin.

Law, A. M. and W. D. Kelton, "Confidence Intervals for Steady-Sttae Simulations; 2. A Survey of Sequential Procedures", Mgmt. Sci., Vol. 28, 1982a), pp. 550-562.

Law, A. M. and W. D. Kelton, Simulation Modeling and Analysis, New York: McGraw Hill, 1982b.

Leipnik, R. B., "Distribution of the Serial Correlation Coefficient in a Circularly Correlated Universe", Ann. Math. Statist., Vol. 18, 1947, pp. 80-87.

Ljung, G. M. and G. E. P. Box, "On a Measure of Lack of Fit in Time Series Models", Biometrika, Vol. 65, 1978, pp. 297-303.

Madow, W. G., "Note on the Distribution of the Serial Correlation Coefficient", Ann. Math. Statist., Vol. 18, 1945, pp. 308-310.

McLeod, I., "Derivation of the Theoretical Autocovariance Function of Autoregressive-Moving Average Time Series", Appl. Statist., Vol. 24, No. 2, 1975, pp. 255-257.

Meketon, M. S. and P. Heidelberger, "A Renewal Theoretic Approach to Bias Reduction in Regenerative Simulations", Mgmt. Sci., Vol. 28, 1982, pp. 173-181.

von Neumann, J., "Distribution of the Ratio of the Mean Square Difference to the Variance", Annals of Math. Stat., Vol. 12, 1941, pp. 367-395.

Priesley, M. B., Spectral Analysis and Time Series, Vol. 1, New York: Academic Press, 1981.

Pritsker, A. A. B. and C. D. Pegden, Introduction to Simulation and SLAM., New York: Wiley, 1979.

Quenouille, M. H., "The Joint Distribution of Serial Correlation Coefficients", Annals Math. Stat., Vol. 20, 1949, pp. 561-571.

Roy, J. and M. L. Tiku, "A Laguerre Series Approximation to the Sampling Distribution of the Variance", Sankhya, SER. A, Vol. 24, 1962, pp. 181-184.

Satterthwaite, F. E., "An Approximate Distribution of Estimates of Variance Components", Biometrics Bulletin, Vol. 2, 1946.

Schruben, L. W., "Coverage Function for Interval Estimators of Simulation Response", Mgmt. Sci., Vol. 26, 1980, pp. 18-27.

Schruben, L. W., "Detecting Initialization Bias in Simulation Output", Opns. Res., Vol. 31, 1982, pp. 569-590.

Schruben, L. W., "Confidence Interval Estimation Using Standardized Time Series", Opns. Res., Vol. 31, 1983, pp. 1090-1108.

Schmeiser, B. W., "Batch Size Effects in the Analysis of Simulation Output", Opns. Res., Vol. 30, 1982, pp. 556-568.

Schmeiser, B. and K. Kang, "Properties of Batch Means from Stationary ARMA(1,1) Time Series", Technical Report 81-3, School of Industrial Engineering, Purdue University.

Solomon, H. and M. A. Stephens, "An Approximation to the Distribution of the Sample Variance", Cana. J. Stat., Vol. 11, No. 2, 1983, pp. 149-154.

Srivastave, A. B. L., "The Effect of Non-Normality on the Power Function of t-Test", Biometrika, Vol. 7, 1909.

Tan, W. Y. and S. P. Wang, "On the Roy-Tiku Approximation to the Distribution of Sample Variances from Nonnormal Universes", J. Amer. Statist. Assoc., Vol. 72, 1977, pp. 875-880.

Walker, A. M., "Large-Sample Estimation of Parameters for Moving Average Models", Biometrika, Vol. 48, 1961, pp. 343-357.

Watson, G. S., "On the Joint Distribution of the Circular Correlation Coefficients', Biometrika, Vol. 43, 1956, pp. 161-168.

Whittle, P., "Tests of Fit in Time Series", Biometrika, Vol. 39, 1952, pp. 309-318.

APPENDIX A

THE DEVELOPMENT OF THE APPROXIMATE DEGREES OF FREEDOM

In equation (3.4.5), if the estimator $s_{\bar{x}}^2$ is given by s^2/n the degrees of freedom f is given by

$$f = \frac{2E^2(s_{\bar{x}}^2)}{Var(s_{\bar{x}}^2)} = \frac{2E^2(s^2/n)}{Var(s^2/n)}$$

$$= \frac{2E^2(s^2)}{Var(s^2)} . \qquad (A.1.1)$$

For the numerator,

$$E(s^2) = E[\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2]$$

$$= \frac{1}{n-1} [\sum_i E(X_i^2) - nE(\bar{X}^2)] .$$

It can be shown that

$$E(X_i^2) = \mu^2 + \sigma^2$$

$$E(\bar{X}^2) = \mu^2 + \sigma^2/n + 2/n(R_\rho(n))$$

where

$$R_\rho(n) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \rho_{ij} .$$

Hence

$$E(s^2) = \sigma^2(1 - \frac{2}{n-1} R_\rho(n)) .$$

So

$$E^2(\frac{s^2}{n}) = [\frac{\sigma^2}{n} (1 - \frac{2}{n-1} R_\rho(n))]^2 .$$

For the denominator,

$$\mathrm{Var}(s^2) = E[s^2 - E(s^2)]^2$$
$$= E(s^4) - E^2(s^2).$$

Let

$$\varepsilon_i = X_i - \mu, \text{ and } \bar{\varepsilon} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu).$$

So

$$E(s^4) = \frac{1}{(n-1)^2} E[\sum_i (X_i - \bar{X})^2]^2$$

$$= \frac{1}{(n-1)^2} E[\sum_i (\varepsilon_i - \bar{\varepsilon})^2]^2$$

$$= \frac{1}{(n-1)^2} E[\sum_{i=1}^{n} \sum_{j=1}^{n} \varepsilon_i^2 \varepsilon_j^2 - 2n\bar{\varepsilon}^2 \sum_{i=1}^{n} \varepsilon_i^2 + n^2\bar{\varepsilon}^4].$$

Now

$$E(\sum_{i=1}^{n} \varepsilon_i^2) = n\mu_4 + \sum_{i=1}^{n} \sum_{j=1}^{n} E(\varepsilon_i^2 \varepsilon_j^2)$$

where μ_4 is the central moment of fourth order,

$$E(\bar{\varepsilon}^2 \sum_{i=1}^{n} \varepsilon_i^2)$$
$$= \frac{1}{n^2} [n\mu_4 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[\varepsilon_i^2 \varepsilon_j^2] + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{k=1}^{n} E(\varepsilon_i \varepsilon_j \varepsilon_k^2)],$$

$$E(\bar{\varepsilon}^4) = \frac{1}{n^4} E[\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{\ell=1}^{n} \varepsilon_i \varepsilon_j \varepsilon_k \varepsilon_\ell].$$

Based upon the characteristic function of multivariate normal random variables with zero means and common variance, the formulas of product moment of fourth order are the following:

$$E(\varepsilon_i^2 \varepsilon_j^2) = (1 + \rho_{ij}^2)\sigma^4 \qquad \text{(A.1.2)}$$

$$E(\varepsilon_i \varepsilon_j \varepsilon_k) = (\rho_{ij} + 2\rho_{ij}\rho_{jk})\sigma^4 \qquad \text{(A.1.3)}$$

$$E(\varepsilon_i\varepsilon_j\varepsilon_k\varepsilon_\ell) = (\rho_{ij}\rho_{k\ell} + \rho_{ik}\rho_{j\ell} + \rho_{i\ell}\rho_{jk})\sigma^4. \qquad (A.1.4)$$

Note that

$$E(s^2) = \sigma^2[1 - \frac{2}{n-1} R_\rho(n)].$$

Hence

$$\begin{aligned}\mathrm{Var}(s^2) = \frac{1}{(n-1)^2} &[(n-2)\mu_4 + (1 - \frac{2}{n})\sum_{i=1}^{n}\sum_{j=1}^{n} (1 + 2\rho_{ij}^2)\sigma^4 \\ &- \frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} (\rho_{ij} + \rho_{ik}\rho_{j\ell})\sigma^4 \\ &+ \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{\ell=1}^{n} (\rho_{ij}\rho_{k\ell} + \rho_{ik}\rho_{j\ell} + \rho_{i\ell}\rho_{jk})\sigma^4] \\ &- \sigma^4(1 - \frac{2}{n-1} R_\rho(n))^2. \qquad (A.1.5)\end{aligned}$$

Using equation (A.1.1), one obtains

$$f = \frac{\sigma^4(1 - \frac{2}{n-1} R_\rho(n))^2}{\mathrm{Var}(s^2)} \qquad (A.1.6)$$

where

$$R_\rho(n) = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \rho_{ij}$$

and $\mathrm{Var}(s^2)$ is given by equation (A.1.5).

APPENDIX B

DERIVATION OF EQUATION (4.3.3) OF CHAPTER IV

From the right hand side of equation (4.3.2),

$$(mk - h)\bar{\sigma}_h$$

$$= \sum_{i=1}^{mk-h} \sigma_{i,i+h}, \quad h = 1,2,\ldots,k-1.$$

The arrangement of sample observations by the method of sequential systematic sampling can be illustrated by Figure 15. Thus,

$$\sum_{i=1}^{mk-h} \sigma_{i,i+h}$$

$$= (\sigma_{1,h+1} + \sigma_{2,h+2} + \cdots + \sigma_{k-h,k})$$

covariance terms within row 1

$$+ (\sigma_{k-h+1,k+1} + \sigma_{k+h-2,k} + \cdots + \sigma_{k,k+h})$$

covariance terms between row 1 and row 2

$$+ (\sigma_{k+1,k+h+1} + \sigma_{k+2,k+h+2} + \cdots + \sigma_{2k-h,2k})$$

covariance terms within row 2

$$\vdots$$

$$+ (\sigma_{(m-1)k+1,(m-1)k+h+1} + \cdots + \sigma_{mk-h,mk})$$

covariance terms within row m

$$= \sum_{i=1}^{m} \sum_{j=1}^{k-h} \sigma_{(i-1)k+j,(i-1)k+j+h} + \sum_{i=1}^{m-1} \sum_{j=1}^{h} \sigma_{ik+j-h,ik+j}.$$

Hence

$$\bar{\sigma}_h = \frac{1}{mk-h} \left[\sum_{i=1}^{m} \sum_{j=1}^{k-h} \sigma_{(i-1)k+j,(i-1)k+j+h} + \sum_{i=1}^{m-1} \sum_{j=1}^{h} \sigma_{ik+j-h,ik+j} \right]$$

$$= \frac{1}{mk-h} \left[\sum_{i=1}^{m} \sum_{j=1}^{k-h} \mathrm{Cov}(Y_{ij}, Y_{i,j+h}) + \sum_{i=1}^{m-1} \sum_{j=1}^{h} \mathrm{Cov}(Y_{i,k-h+j}, Y_{i+1,j}) \right].$$

Row \ Column	1	2	3	. . .	k
1	X_1	X_2	X_3	. . .	X_k
2	X_{k+1}	X_{k+2}	X_{k+3}	. . .	X_{2k}
3	X_{2k+1}	X_{2k+2}	X_{2k+3}	. . .	X_{3k}
.	.	.	.		.
.	.	.	.		.
.	.	.	.		.
m	$X_{(m-1)k+1}$	$X_{(m-1)k+2}$	$X_{(m-1)k+3}$	. . .	X_{mk}

Figure 15

Arrangement of Sample Observations in the Method of Sequential Systematic Sampling.

APPENDIX C

THE METHOD OF BATCH MEANS

C.1 INTRODUCTION

The purpose of the method of batch means is to overcome the problems associated with correlated data in applying inferential produres for means. While this method does not totally eliminate the problems of correlated data, it can significantly reduce the magnitude of the error in probability statements concerning the level of confidence and power of the test which typically arise when correlation is ignored. The method of batch means is most effective when the absolute magnitude of correlation between two observations, X_i and X_j, decreases as $|i-j|$ increases.

In applying the method of batch means a sample of size n composed of the observations $X_1, X_2, \cdots, X_n$ with common mean μ is broken into m batches each of size k, where $n = mk$. If Y_ℓ denotes the sample mean for the ℓth batch,

$$Y_1 = \frac{1}{k} \sum_{i=1}^{k} X_i$$

$$\vdots$$

$$Y_\ell = \frac{1}{k} \sum_{i=1}^{k} X_{(\ell-1)k+i}$$

$$\vdots$$

$$Y_m = \frac{1}{k} \sum_{i=1}^{k} X_{(m-1)k+i} .$$

If the X_i have common mean, μ, then $E(Y_\ell) = \mu$, for $\ell = 1,2,\cdots,m$. Concerning confidence intervals or tests of hypotheses for μ, one will require an estimator for $\sigma^2_{\bar{x}}$. Based upon the discussion in Chapter III,

s_x^2/n may be a poor estimator for correlated data, where

$$s_x^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2.$$

This is due to the potential magnitude of the bias in s_x^2/n as an estimator of $\sigma_{\bar{X}}^2$.

Given the conjecture that $|\rho_{ij}|$ decreases with increasing $|i-j|$, one may reason that there is less overall correlation among the Y_ℓ's than there is among the X_i's. Therefore the Y_ℓ's are treated as if they were individual observations, and the estimator of $\sigma_{\bar{X}}^2$ is given by s_y^2/m where

$$s_y^2 = \frac{1}{(m-1)} \sum_{i=1}^{m} (Y_i - \bar{Y})^2.$$

Since

$$\begin{aligned}\bar{Y} &= \frac{1}{m} \sum_{\ell=1}^{m} Y_\ell \\ &= \frac{1}{mk} \sum_{\ell=1}^{m} \sum_{i=1}^{k} X_{(\ell-1)k+i} \\ &= \bar{X}\end{aligned}$$

$$s_y^2 = \frac{1}{m-1} \sum_{i=1}^{m} (Y_i - \bar{X})^2.$$

C.2 THE BIAS FOR THE ESTIMATOR OF THE VARIANCE OF SAMPLE MEAN

The error in probability statements concerning the mean of correlated data is due in great part to the bias in the estimator of $\sigma_{\bar{X}}^2$. It would be necessary to examine the bias in s_y^2/m as an estimator of $\sigma_{\bar{X}}^2$ to determine whether the situation is actually improved by batching observations. The bias in s_y^2/m as an estimator of $\sigma_{\bar{X}}^2$, $B_{\frac{s_y^2}{m}, \sigma_{\bar{X}}^2}$, is defined by

$$B_{\frac{s_y^2}{m},\sigma_{\bar{X}}^2} = E(\frac{s_y^2}{m}) - \sigma_{\bar{X}}^2 .$$

Now

$$E(s_y^2) = E[\frac{s_y^2}{m}] - \sigma_{\bar{X}}^2$$

$$= \frac{1}{m-1} [\sum_{\ell=1}^{m} E(Y_\ell^2) - mE(\bar{X}^2)]. \qquad \text{(C.2.1)}$$

Notice that

$$E(Y_\ell^2) = E[(\frac{1}{k} \sum_{i=1}^{k} X_{(\ell-1)k+i})^2]$$

$$= \frac{1}{k^2} [\sum_{i=1}^{k} E(X_{(\ell-1)k+i}^2) + 2 \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} E(X_{(\ell-1)k+i}X_{(\ell-1)k+j})] .$$

After some algebraic operations,

$$E(Y_\ell^2) = \frac{\sigma^2}{k} + \mu^2 + \frac{2\sigma^2}{k} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \rho_{(\ell-1)k+i,(\ell-1)k+j} . \qquad \text{(C.2.2)}$$

For $E(\bar{X}^2)$,

$$E(\bar{X}^2) = \frac{1}{m^2} E[\sum_{\ell=1}^{m} Y_\ell]^2$$

$$= \frac{1}{m^2} [\sum_{\ell=1}^{m} E(Y_\ell^2) + 2 \sum_{\ell=1}^{m-1} \sum_{t=\ell+1}^{m} E(Y_\ell Y_t)].$$

After some algebraic operations,

$$E(Y_\ell^2) = \frac{\sigma^2}{k} + \mu^2 + \frac{2\sigma^2}{(mk)^2} [\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \rho_{(\ell-1)k+i,(\ell-1)k+j}$$

$$+ \sum_{\ell=1}^{m-1} \sum_{t=\ell+1}^{m} \sum_{i=1}^{k} \sum_{j=1}^{k} \rho_{(\ell-1)k+i,(t-1)k+j}] . \qquad \text{(C.2.3)}$$

Substituting equations (C.2.2) and (C.2.3) into equation (C.2.1) yields

$$E(s_y^2) = \frac{\sigma^2}{k} + \frac{2\sigma^2}{mk} \sum_{\ell=1}^{m} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \rho_{(\ell-1)k+i,(\ell-1)k+j}$$

$$- \frac{2\sigma^2}{(m-1)mk^2} \sum_{\ell=1}^{m-1} \sum_{t=\ell+1}^{m} \sum_{i=1}^{k} \sum_{j=1}^{k} \rho_{(\ell-1)k+i,(t-1)k+j}.$$

Define

$$R_{w_\rho} = \frac{1}{mk} \sum_{\ell=1}^{m} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \rho_{(\ell-1)k+i,(\ell-1)k+j} \tag{C.2.4}$$

$$R_{B_\rho} = \frac{1}{mk} \sum_{\ell=1}^{m-1} \sum_{t=\ell+1}^{m} \sum_{i=1}^{k} \sum_{j=1}^{k} \rho_{(\ell-1)k+i,(\ell-1)k+j}. \tag{C.2.5}$$

Then

$$E(s_y^2) = \frac{\sigma^2}{k} + \frac{2\sigma^2}{k} R_{w_\rho} - \frac{2\sigma^2}{(m-1)k} R_{B_\rho}$$

where R_{w_ρ} accounts for the correlation within batches and R_{B_ρ} accounts for that between batches. Finally

$$E\left(\frac{s_y^2}{m}\right) = \frac{\sigma^2}{mk} [1 + 2R_{w_\rho} - \frac{2}{m(m-1)} R_{B_\rho}].$$

Recalling that $\sigma_{\bar{x}}^2 = \sigma^2(1 + 2R_\rho(n))/n$, $R_\ell(n)$ may be expressed by $R_\rho(n) = (R_{w_\rho} + R_{B_\rho})$. Then

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{mk} [1 + 2(R_{w_\rho} + R_{B_\rho})].$$

Hence the bias in s_y^2/m as an estimator of $\sigma_{\bar{x}}^2$ is

$$B_{\frac{s_y^2}{m},\sigma_{\bar{x}}^2} = - \frac{2\sigma^2}{(m-1)k} R_{B_\rho}. \tag{C.2.6}$$

Notice that the bias in s_x^2/n as an estimate of $\sigma_{\bar{x}}^2$ is given by

$$B_{\frac{s_x^2}{n},\sigma_{\bar{x}}^2} = \frac{-2\sigma^2}{(mk-1)}(R_{w_\rho} + R_{B_\rho}). \qquad (C.2.7)$$

Equations (C.2.6) and (C.2.7) indicate that for positively correlated data where $R_{w_\rho} > 0$ and $R_{B_\rho} > 0$, both s_x^2/n and s_y^2/m tend to underestimate $\sigma_{\bar{x}}^2$. On the other hand for $R_{w_\rho} < 0$ and $R_{B_\rho} < 0$ both estimator tend to overestimate $\sigma_{\bar{x}}^2$. However, examining the ratio of the bias in s_y^2/m to the bias in s_x^2/n an analyst can find that for m large relative to k, s_y^2/m has less bias than s_x^2/n in either case. That is

$$B_{\frac{s_y^2}{m},\sigma_{\bar{x}}^2} / B_{\frac{s_x^2}{n},\sigma_{\bar{x}}^2} = \frac{mk-1}{(m-1)k} \frac{R_{B_\rho}}{R_{w_\rho}+R_{B_\rho}}$$

$$\approx \frac{R_{B_\rho}}{R_{w_\rho}+R_{B_\rho}}$$

$$< 1.$$

The expressions of R_{w_ρ} and R_{B_ρ} in equations (C.2.4) and (C.2.5) can be simplified if they were expressed in terms of the lag covariance. That is to replace ρ_{ij} with ρ_h where $h = |i-j|$. Then R_{w_ρ} becomes

$$R_{w_\rho} = \frac{1}{mk} \sum_{\ell=1}^{m} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \rho_{ij}$$

$$= \frac{1}{k} \sum_{i=1}^{k-1} (k-i)\rho_i. \qquad (C.2.8)$$

From equation (C.2.5), it follows that

$$R_{B_\rho} = \frac{1}{mk} \sum_{\ell=1}^{m-1} \sum_{t=\ell+1}^{m} \sum_{i=1}^{k} \sum_{j=1}^{k} \rho_{(\ell-1)k+i,(t-1)k+j}$$

$$= \frac{1}{mk} \sum_{j=1}^{m-1} (m-j)[k\rho_{jk} + \sum_{i=1}^{k-1} (k-i)(\rho_{jk-i} + \rho_{jk+i})]. \qquad (C.2.9)$$

Substituting equation (C.2.9) into equation (C.2.6) yields

$$B_{\frac{s_y^2}{m},\sigma_{\bar{x}}^2} = \frac{-2\sigma^2}{(m-1)mk} \sum_{j=1}^{m-1} (m-j)[k\rho_{jk} + \sum_{i=1}^{k-1} (k-i)(\rho_{jk-i} + \rho_{jk+i})]. \qquad (C.2.10)$$

Using equation (C.2.8) and (C.2.9), one obtains

$$B_{\frac{s_x^2}{n},\sigma_{\bar{x}}^2} = \frac{-2\sigma^2}{(mk-1)} \{\frac{1}{k} \sum_{i=1}^{k-1} (k-i)\rho_i + \frac{1}{mk} \sum_{j=1}^{m-1} (m-j)[k\rho_{jk}$$

$$+ \sum_{i=1}^{k-1} (k-i)(\rho_{jk-i} + \rho_{jk+i})]\}. \qquad (C.2.11)$$

APPENDIX D

VARIANCE OF THE VARIANCCE ESTIMATOR, $\mathrm{Var}(\hat{\sigma}^2)$, FOR NORMAL DATA AND ITS ESTIMATE, $\widehat{\mathrm{Var}}(\hat{\sigma}^2)$

The variance of the estimator $\hat{\sigma}^2$ under the assumption of normality is given by

$$\begin{aligned}
\mathrm{Var}(\hat{\sigma}^2) &\\
&= \frac{1}{k^2}\{\sigma^4[k(\frac{\gamma_2}{m} + \frac{2}{m-1}) + k(k-1)(m-2)]\\
&+ 2\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\frac{1}{(m-1)^2}\sum_{\ell=1}^{m}\sum_{h=1}^{m}(2\sigma^2_{(\ell-1)k+i,(h-1)k+j}\\
&- \frac{4(m-1)}{m^2}\sum_{\ell_1=1}^{m-1}\sum_{\ell_2=\ell_1+1}^{m}\sum_{h=1}^{m}(2\sigma_{(h-1)k+j,(\ell_1-1)k+i}\sigma_{(h-1)k+j,(\ell_2-1)k+i})\\
&+ \frac{4}{m^2}\sum_{\ell_1=1}^{m-1}\sum_{\ell_2=\ell_1+1}^{m}\sum_{h_1=1}^{m-1}\sum_{h_2=h_1+1}^{m}(\sigma_{(\ell_1-1)k+i,(h_1-1)k+j}\sigma_{(\ell_2-1)k+i,(h_2-1)k+j}\\
&+ \sigma_{(\ell_1-1)k+i,(h_2-1)k+j}\sigma_{(\ell_2-1)k+i,(h_1-1)k+j})]\}. \qquad \text{(D.1)}
\end{aligned}$$

An estimate of $\mathrm{Var}(\hat{\sigma}^2)$ could be obtained by estimating σ^4 and each covariance term in equation (D.1). The estimate of σ^4 is chosen as $(\hat{\sigma}^2)^2$ where $\hat{\sigma}^2$ is the population variance estimate suggested by the method of sequential systematic sampling. To estimate the covariance terms such as $\sigma^2_{(\ell-1)k+i,(h-1)k+j}$ and the like, the lag covariance estimate suggested by the method of sequential systematic sampling in equation (4.3.14) of Chapter IV is used.

APPENDIX E

VARIANCE OF THE VARIANCCE ESTIMATOR, $Var(\hat{\sigma}^2)$, FOR NONNORMAL DATA AND ITS ESTIMATE, $\widehat{Var}(\hat{\sigma}^2)$

Incorporating K(h,g,k) into the fourth moment V(h,g,k), the variance of the variance estimator can be expressed in the following form:

$Var(\hat{\sigma}^2)$

$$= \frac{2k_4}{k^2(m-1)^2} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \{[\sum_{\ell=1}^{m} \sum_{h=1}^{m} \sum_{t=0}^{k} \theta_t^2\theta_{t+(h-\ell)k+(j-i)}^2]$$

$$- \frac{2}{m} [(\sum_{\ell_1=1}^{m-1} \sum_{\ell_2=\ell_1+1}^{m} \sum_{h=1}^{m} \sum_{t=0}^{k} \theta_t^2\theta_{t+(\ell_1-h)k+(i-j)}\theta_{t+(\ell_2-h)k+(i-j)})$$

$$+ (\sum_{\ell=1}^{m} \sum_{h=1}^{m} \sum_{t=0}^{k} \theta_t^2\theta_{t+(h-\ell)k+(j-i)}^2)]$$

$$+ \frac{1}{m^2} [(\sum_{\ell_1=1}^{m} \sum_{\ell_2=1}^{m} \sum_{t=0}^{k} \theta_t^2\theta_{t+(\ell_2-\ell_1)k+(j-i)}^2) + 2(\sum_{\ell_2=1}^{m} \sum_{h=1}^{m-1} \sum_{g=h+1}^{m} \sum_{t=0}^{k}$$

$$\theta_t^2\theta_{\ell+(h-\ell_2)k+(i-j)}\theta_{\ell+(g-\ell_2)k+(i-j)})$$

$$+ 2(\sum_{\ell_1=1}^{m} \sum_{\ell_2=1}^{m-1} \sum_{\ell_3=\ell_1+1}^{m} \theta_t^2\theta_{t+(\ell_2-\ell_1)k+j-i}\theta_{t+(\ell_2-\ell_1)k+(j-i)})]\}$$

$$+ \frac{1}{k^2} \{k\sigma^4(\frac{\gamma_2}{m} + \frac{2}{m-1}) + 2 \sum_{i=1}^{k-1} \sum_{j=i+1}^{k}$$

$$\{\frac{1}{(m-1)^2} [\frac{(m-1)^2}{m^2} \sum_{\ell=1}^{m} \sum_{h=1}^{m} (\sigma^4 + 2\sigma^2_{(\ell-1)k+i(h-1)k+j})$$

$$- \frac{4(m-1)}{m^2} \sum_{\ell_1=1}^{m-1} \sum_{\ell_2=\ell_1+1}^{m} \sum_{h=1}^{m} (\sigma^2\sigma_{(\ell_1-1)k+i}$$

$$+ 2\sigma_{(h-1)k+j,(\ell_1-1)k+i}\sigma_{(h-1)k+j,(\ell_2-1)k+i})$$

$$+ \frac{4}{m^2} \sum_{\ell_1=1}^{m-1} \sum_{\ell_2=\ell_1+1}^{m} \sum_{h_1=1}^{m-1} \sum_{h_2=h_1+1}^{m} (\sigma_{(\ell_1-1)k+i,(\ell_2-1)k+i}\sigma_{(h_1-1)k+j,(h_2-1)k+j}$$

$$+ \sigma_{(\ell_1-1)k+i,(h_1-1)k+j}\sigma_{(\ell_2-1)k+i,(h_2-1)k+j}$$

$$+ \sigma_{(\ell_1-1)k+i,(h_2-1)k+j}\sigma_{(\ell_2-1)k+i,(h_1-1)k+j})]\}$$

$$- \sigma^4\}. \tag{E.1}$$

In addition to invoking the estimate of $\text{Var}(\hat{\sigma}^2)$ in the case of normal data, the estimate of $\text{Var}(\hat{\sigma}^2)$ in the case of nonnormal data requires the estimates of θ_t's, γ_2, and k_4.

First, the estimates of the moving average parameters, $\hat{\theta}_t$'s, can be obtained by the maximum likelihood estimation procedure in Box and Jenkins (1976). A detailed discussion about such estimation is given in Chapter VIII.

Second, the kurtosis of the output process, γ_2, is defined as

$$\begin{aligned} \gamma_2 &= \frac{\mu_{4,x}}{\sigma^4} - 3 \\ &= \frac{\mu_{4,x} - 3\sigma^4}{\sigma^4} \\ &= \frac{k_{4,x}}{\sigma^4} \end{aligned} \tag{E.2}$$

where $\mu_{4,x}$ and $k_{4,x}$ are the fourth central moment and cumulant of the output process $\{X_t\}$. Then, to estimate γ_2, the following estimate is suggested,

$$\hat{\gamma}_2 = \frac{\hat{k}_{4,x}}{(\hat{\sigma}^2)^2} \tag{E.3}$$

where $\hat{\sigma}^2$ is the population variance estimate suggested by the method of sequential systematic sampling and $\hat{k}_{4,x}$ is an estimate of the fourth cumulant originated from Cramer (1946). Now, $\hat{k}_{4,x}$ is given by

$$\hat{k}_{4,x} = \frac{n}{(n-1)(n-2)(n-3)} [(n+1)m_4 - 3(n-1)(\hat{\sigma}^2)^2] \qquad (E.4)$$

where

$$n = mk$$

$$m_4 = \frac{1}{n} \sum_{i=1}^{k} \sum_{\ell=1}^{n} (X_{(\ell-1)k+i} - \bar{X}_i)^4 \qquad (E.5)$$

$$\bar{X}_i = \frac{1}{m} \sum_{\ell=1}^{m} X_{(\ell-1)k+i}.$$

To avoid the variability in estimating γ_2, the final estimate of γ_2, $\hat{\bar{\gamma}}_2$, is the sample mean of the γ_2 estimates from all replicates.

Third, by definition, k_4 (the fourth cumulant of the input process) is given by

$$k_4 = \gamma_{2,Z}\sigma_Z^4 \qquad (E.6)$$

where $\gamma_{2,Z}$ and σ_Z^2 are the kurtosis and variance of the input process $\{Z_t\}$. Hence, an estimate of k_4 can be obtained by multiplying estimates of $\gamma_{2,Z}$ and σ_Z^4 together. If the correlation of the output process $\{X_t\}$ dies out at k, then $\{X_t\}$ can be given by

$$X_t = \sum_{i=1}^{k} \theta_i Z_{t+i-1} \qquad (E.7)$$

where $\{Z_t\}$ is a sequence of i.i.d. random variables with mean μ and variance σ_Z^2. The variance of X_t, σ^2, is given by

$$\sigma^2 = \sigma_Z^2 \sum_{i=1}^{k} \theta_i^2. \qquad (E.8)$$

Thus an estimate of σ_Z^2, $\hat{\sigma}_Z^2$, is chosen as

$$\hat{\sigma}_Z^2 = \frac{\hat{\sigma}^2}{\sum_{i=1}^{k} \hat{\theta}_i^2} \tag{E.9}$$

where $\hat{\sigma}^2$ and $\hat{\theta}_i$'s are as previously defined. The kurtosis of the input process, $\gamma_{2,Z}$, can be expressed in terms of the kurtosis of the output process, γ_2, as

$$\gamma_{2,Z} = \frac{(\sum_{t=1}^{k} \theta_t^2)^2}{\sum_{t=1}^{k} \theta_t^4} \gamma_2. \tag{E.10}$$

Hence an estimate of $\gamma_{2,Z}$ is given by

$$\hat{\gamma}_{2,Z} = \frac{(\sum_{t=1}^{k} \hat{\theta}_t^2)^2}{\sum_{t=1}^{k} \hat{\theta}_t^4} \hat{\hat{\gamma}}_2. \tag{E.11}$$

where $\hat{\theta}_t$'s and $\hat{\hat{\gamma}}_2$ are as previously defined. Now the estimator of k_4 is given by

$$\hat{k}_4 = \hat{\gamma}_{2,Z}(\hat{\sigma}_Z^2)^2. \tag{E.12}$$

APPENDIX F

THE COMPUTATIONAL FORMS FOR TRADITIONAL METHODS FOR MEANS AND VARIANCES

Parameter	Confidence Interval
μ	$\bar{x} - \hat{\sigma}_{\bar{x}}\, t_{1-\alpha/2,f} < \mu < \bar{x} + \hat{\sigma}_{\bar{x}}\, t_{1-\alpha/2,f}$
$\mu_1 - \mu_2$	$(\bar{x}_1-\bar{x}_2) - \surd\ \hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}_2}\, t_{1-\alpha/2,f} < \mu_1-\mu_2 < (\bar{x}_1-\bar{x}_2) + \surd\ \hat{\sigma}^2_{\bar{x}_1} + \hat{\sigma}^2_{\bar{x}_2}\, t_{1-\alpha/2,f}$
σ^2	$\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{\alpha/2}}$
$\frac{\sigma_1^2}{\sigma_2^2}$	$\frac{S_1^2}{S_2^2} F_{1-\alpha/2;n_1-1,n_2-1} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} F_{\alpha/2;n_1-1,n_2-1}$

Hypothesis Test	Power Function
H_0: $\mu = \mu_0$, H_1: $\mu = \mu_1 \neq \mu_0$	$1 - Pr[\frac{\mu_0-\mu_1}{\sigma_{\bar{x}}} - \frac{\hat{\sigma}_{\bar{x}}}{\sigma_{\bar{x}}} Z_{1-\alpha/2} < Z < \frac{\mu_0-\mu_1}{\sigma_{\bar{x}}} + \frac{\hat{\sigma}_{\bar{x}}}{\sigma_{\bar{x}}} Z_{1-\alpha/2}]$
H_0: $\mu_1 - \mu_2 = 0$, H_1: $\mu_1 - \mu_2 = d$	$1 - Pr[\frac{d}{\sqrt{\sigma_{\bar{x}_1}^2+\sigma_{\bar{x}_2}^2}} - \frac{\sqrt{\hat{\sigma}_{\bar{x}_1}^2+\hat{\sigma}_{\bar{x}_2}^2}}{\sqrt{\sigma_{\bar{x}_1}^2+\sigma_{\bar{x}_2}^2}} Z_{1-\alpha/2} < Z < \frac{d}{\sqrt{\sigma_{\bar{x}_1}^2+\sigma_{\bar{x}_2}^2}} + \frac{\sqrt{\hat{\sigma}_{\bar{x}_1}^2+\hat{\sigma}_{\bar{x}_2}^2}}{\sqrt{\sigma_{\bar{x}_1}^2+\sigma_{\bar{x}_2}^2}} Z_{1-\alpha/2}]$
H_0: $\sigma^2 = \sigma_0^2$, H_1: $\sigma^2 = \sigma_1^2 \neq \sigma_0^2$	$1 - Pr[\frac{\sigma_0^2}{\sigma_1^2} \chi^2_{\alpha/2,n-1} < \chi^2_{n-1} < \frac{\sigma_0^2}{\sigma_1^2} \chi^2_{1-\alpha/2,n-1}]$
H_0: $\sigma_1^2 = \sigma_2^2$, H_1: $\sigma_1^2 \neq \sigma_2^2$	$1 - Pr[\frac{\sigma_2^2}{\sigma_1^2} F_{\alpha/2;n_1-1,n_2-1} < F_{n_1-1,n_2-1} < F_{n_1-1,n_2-1} < F_{1-\alpha/2;n_1-1,n_2-1}]$

n_1 = Number of observations taken from population 1.

n_2 = Number of observations taken from population 2.

APPENDIX G

THE COMPUTATIONAL FORMS SUGGESTED BY THIS RESERACH FOR MEANS AND VARIANCES

Parameter	Confidence Interval
μ	$\bar{x} - \hat{\sigma}_{\bar{x}_k} t_{1-\alpha/2,n-1} < \mu < \bar{x} + \hat{\sigma}_{\bar{x}_k} t_{1-\alpha/2,n-1}$
$\mu_1 - \mu_2$	$(\bar{x}_1-\bar{x}_2) - \sqrt{\hat{\sigma}^2_{\bar{x}_{k_1}} + \hat{\sigma}^2_{\bar{x}_{k_2}}}\, t_{1-\alpha/2,n_1+n_2-2} < \mu_1-\mu_2 < (\bar{x}_1-\bar{x}_2) + \sqrt{\hat{\sigma}^2_{\bar{x}_{k_1}} + \hat{\sigma}^2_{\bar{x}_{k_2}}}\, t_{1-\alpha/2,n_1+n_2-2}$
σ^2	$\hat{\sigma}^2 - \sqrt{\hat{Var}(\hat{\sigma}^2)}\, Z_{1-\alpha/2} < \sigma^2 < \hat{\sigma}^2 + \sqrt{\hat{Var}(\hat{\sigma}^2)}\, Z_{1-\alpha/2}$
$\sigma_1^2 - \sigma_2^2$	$(\hat{\sigma}_1^2-\hat{\sigma}_2^2) - \sqrt{\hat{Var}(\hat{\sigma}_1^2)+\hat{Var}(\hat{\sigma}^2)}\, Z_{1-\alpha/2} < \sigma_1^2 - \sigma_2^2 < (\hat{\sigma}_1^2-\hat{\sigma}_2^2) + \sqrt{\hat{Var}(\hat{\sigma}_1^2)+\hat{Var}(\hat{\sigma}^2)}\, Z_{1-\alpha/2}$

Hypothesis Test	Power Function
$H_0: \mu = \mu_0, H_1: \mu = \mu_1 \neq \mu_0$	$1 - \Pr[\frac{\mu_0-\mu_1}{\sigma_{\bar{x}}} - \frac{\hat{\sigma}_{\bar{x}_k}}{\sigma_{\bar{x}}} Z_{1-\alpha/2} < Z < \frac{\mu_0-\mu_1}{\sigma_{\bar{x}}} + \frac{\hat{\sigma}_{\bar{x}_k}}{\sigma_{\bar{x}}} Z_{1-\alpha/2}]$
$H_0: \mu_1 - \mu_2 = 0, H_1: \mu_1 - \mu_2 = d$	$1 - \Pr[\frac{d}{\sqrt{\sigma^2_{\bar{x}_1}+\sigma^2_{\bar{x}_2}}} - \frac{\sqrt{\hat{\sigma}^2_{\bar{x}_{k_1}}+\hat{\sigma}^2_{\bar{x}_{k_2}}}}{\sqrt{\sigma^2_{\bar{x}_1}+\sigma^2_{\bar{x}_2}}} Z_{1-\alpha/2} < Z < \frac{d}{\sqrt{\sigma^2_{\bar{x}_1}+\sigma^2_{\bar{x}_2}}} + \frac{\sqrt{\hat{\sigma}^2_{\bar{x}_{k_1}}+\hat{\sigma}^2_{\bar{x}_{k_2}}}}{\sqrt{\sigma^2_{\bar{x}_1}+\sigma^2_{\bar{x}_2}}} Z_{1-\alpha/2}]$
$H_0: \sigma^2 = \sigma^2_0, H_1: \sigma^2 = \sigma^2_1 \neq \sigma^2_0$	$1 - \Pr[\frac{\sigma^2_0-\sigma^2_1}{\sqrt{\mathrm{Var}(\hat{\sigma}^2)}} \frac{1}{R_\sigma} Z_{1-\alpha/2} < Z < \frac{\sigma^2_0-\sigma^2_1}{\sqrt{\mathrm{Var}(\hat{\sigma}^2)}} + \frac{1}{R_\sigma} Z_{1-\alpha/2}]$
$H_0: \sigma^2_1 = \sigma^2_2, H_1: \sigma^2_1 - \sigma^2_2 = d$	$1 - \Pr[\frac{d}{\sqrt{\mathrm{Var}(\hat{\sigma}^2_1)+\mathrm{Var}(\hat{\sigma}^2_2)}} - \frac{1}{R_\sigma} Z_{1-\alpha/2} < Z < \frac{d}{\sqrt{\mathrm{Var}(\hat{\sigma}^2_1)+\mathrm{Var}(\hat{\sigma}^2_2)}} + \frac{1}{R_\sigma} Z_{1-\alpha/2}]$

The vita has been removed from the scanned document