ECONOMIC DESIGN OF CONTROL CHARTS

FOR

CORRELATED, MULTIVARIATE OBSERVATIONS

A THESIS

Presented to

The Faculty of the Division of

Graduate Studies

by

Francis Bernard Alt

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

in the School of Industrial and Systems Engineering

Georgia Institute of Technology

October, 1977

ECONOMIC DESIGN OF CONTROL CHARTS FOR CORRELATED MULTIVARIATE OBSERVATIONS

Approved:

Stuart J. Deutsch, Chairman

Stanley A. Mulaik

Russell G. Heikes

James J. Goode

Harrison M. Wadsworth

Date approved by Chairman: October 21, 1977

ACKNOWLEDGMENTS

It is a pleasure to express my appreciation and gratitude to my thesis advisor, Dr. S. J. Deutsch. His deep insight, guidance, and friendship were most instrumental in the completion of this dissertation. Without his help, this research would not have been possible.

My sincerest appreciation also goes to three other members of my dissertation committee: Dr. R. G. Heikes, Dr. S. A. Mulaik, and Dr. H. M. Wadsworth. Their conscientious and helpful attitudes and willingness to help me meet my deadline far exceeded the requirements of committee members.

The fourth member of my reading committee, Dr. J. J. Goode, deserves special thanks for unselfishly extending his outstanding mathematical ability whenever the need arose ,which was frequent. His friendship and guidance throughout my graduate program will long be remembered.

Dr. J. W. Walker also deserves special thanks for the direct contributions he has made via his outstanding mathematical ability in enabling me to complete this dissertation. However, my debt to him goes far deeper. How do I express my gratitude to someone who has provided deep inspiration and warm friendship throughout my years of graduate studies?

Two fellow graduate students, V. Venkata Rao and Suleyman Tufecki, unbegrudgingly supplied their time and computer programming competence needed for the preparation of most of the tables and Appendices. Their assistance is much appreciated.

Many former and current graduate students deserve my gratitude for their friendship and assistance during my graduate program. Among them are Dr. M. Bazaraa, Dr. D. Sipper, Dr. R. Rardin, Dr. R. Bulfin, Dr. B. Schmeiser, Dr. H. Vaish, Dr. G. Zalmai P. Pfeifer and F. Cullen. The list goes on.

Special thanks also goes to Ms. Aileen Arnold, Ms. Joyce Alexander, Ms. Judith Alt and Ms. Post for their excellent typing under an ever pressing deadline. Most of the typing was done by Ms. Arnold and Ms. Alexander.

I gratefully acknowledge the encouragement received from my family. My wife, Judith, and my daughter, Beth, sacrificed much attention and time that they deserved. Quite frequently, Judith was called on to be both mother and father to Beth. To her I express my special thanks. My parents deserve much gratitude for their constant encouragement throughout my life.

I would like to express my appreciation to the National Institute of Law Enforcement and Criminal Justice for providing contractual support for this research. A formal acknowledgment follows.

> This work was performed under Grant No. 75 NI-990091 from the National Institute of Law Enforcement and Criminal Justice. Points of view or opinions stated in this document are those of the author and do not necessarily represent the official position or policies of the United States Department of Justice.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

The scenario for the interrupted time series quasi experiment (ITSQE) is a set of $n = n_1 + n_2$ observations recorded at equispaced epochs of time, with an intervention or treatment introduced after the n_1^{th} observation. Since the observations are correlated, autoregressive-moving average models have been used to describe the behavior of observations obtained from the ITSQE. However, in order to take into account that the intervention has the potential to affect the post-intervention level of the time series, an additive shift parameter is included in the post-intervention model. In this dissertation, the models for the ITSQE were made even more flexible by taking into account that the intervention also has the potential to affect the variability-covariability of the process. These models were designated the multi-consequence intervention models.

Two methods of parameter estimation were investigated for the multi-consequence intervention model with particular emphasis directed towards the first and second order multi-consequence intervention model. The first method was designated "iterative conditional least squares estimation," and the basic idea is to transform the n original observations to another set of observations amenable to statistical linear model analysis. A search is conducted over the permissible parameter space of the moving average parameters until those values are found which minimize the sum of squared residuals of the transformed observations. The method of maximum likelihood was the second method.

While closed form expressions were obtained for the level and shift parameters, no such expressions could be obtained for the moving average parameters. However, an algorithm was presented for efficiently calculating the likelihood function. One advantage of using the maximum likelihood method is that an asymptotic likelihood ratio test can be employed to test whether the pre-intervention moving average parameters are equal to the post-intervention moving average parameters. The Appendices contain computer programs for both methods of estimation.

The detection of a shift in the level of an underlying process is also a problem of utmost importance in the area of quality control. Since the quality control scenario involves repeated samples of size n, the monitoring of the process level is usually recorded on a control chart. Whether there be one or multiple quality characteristics, the control chart scenario had previously assumed independent observations. This research has extended that to include correlated observations. Furthermore, the properties of the statistics used to monitor the process were also investigated.

For the quality control scenario, this research has also determined the economic parameters of sample size and control chart constant by using the scheme of minimizing the average run length of an out of control process for a large fixed value of the average run length of an in control process. This was done for two cases: multiple (2 and 3) quality characteristics for independent obse-vations; and, one quality characteristic for first-order serially correlated observations.

Finally, the concept of a multivariate, multiconsequence intervention model was introduced, and its properties were presented.

CHAPTER I

INTRODUCTION

In recent years, there has been increasing demand by consumers for quality products, and there is no sign of abatement. In order to be responsive to this demand, manufacturers have increasingly adopted various techniques of statistical quality control. One technique that has been very successful in monitoring a process is the control chart.

Shewhart [59] is generally credited with the development of the control chart in 1924. The basic idea behind the control chart is that there are two sources of variation in the quality of a product: chance causes and assignable causes. While the chance cause variation cannot be controlled, it is assumed that this variation follows a certain statistical pattern such as the normal distribution. When the variations do not conform to this assumption, a search is undertaken for one or more assignable causes such as a difference among raw materials. Additional discussion of this can be found in Duncan [24] and Grant and Leavenworth [30].

There are two distinct phases of control chart practice. The distinction being that in Phase I the control chart is used for analyzing past data for a lack of control and to assist in establishing control charts when no standards are given while in Phase II the chart is used to detect any departure of the underlying process from standard values. This dissertation is primarily concerned with Phase II.

Furthermore, although control charts are used to monitor both the process mean and variability, this dissertation concentrates on those used for the mean. Thus, primary attention is directed towards Phase II control charts for the mean.

In order to implement a Shewhart chart, a sample of n independent observations is obtained from the process at time t, and the value of a statistic (w_t) is calculated and plotted on a control chart. The chart usually has a central line and upper and lower control limits, which are taken to be $E(W) \pm 3\sqrt{Var(W)}$. If the value of the statistic falls within $E(W) \pm 3\sqrt{Var(W)}$, the decision maker can safely conclude that the process is under control. If w_t falls outside the control limits, assignable causes of variation are sought. With respect to monitoring the process mean, W is usually taken to be $\bar{X}$ (the sample mean). Furthermore, if it can be assumed that the process is normally distributed with the nominal values of the process mean and standard deviation denoted by μ_0 and σ_0, respectively, then upper and lower control limits are given by $\mu_0 \pm 3(\sigma_0/\sqrt{n})$.

Two underlying assumptions in the methodological development of control charts are that the process is normally distributed and the observations within a sample are independent as well as the between sample values of $\bar{X}$. Quite frequently, these assumptions are not warranted. The normality assumption is frequently justified by the central limit theorem. Moreover, a recent paper by Schilling and Nelson [57] provides tables which show the rate of approach of the distribution of sample means to normality for various underlying distributions and sample sizes. One of their findings is that this rate of approach is particularly slow for exponential and contaminated distributions.

Although a failure to satisfy the normality assumption is a serious error, a failure to satisfy the independence assumption is by far the more serious type of error. This has been demonstrated by Daniel [16], Scheffé [56], and Padia [47]. Actually, their investigations were performed for significance tests for the mean of a normal population. However, most of their results are applicable to control charts for the mean because of the one-to-one correspondence between control charts and significance tests. Walsh [67] has investigated the effect of intraclass correlation (the correlation between each two sample values is the same) on the significance level of the test for a single mean of a univariate normal population, and this has been extended by Basu, Odell, and Lewis [12] to samples drawn from a multivariate normal population. However, intraclass correlation appears to be a rare phenomenon in quality control for one is far more likely to encounter serial correlation such as was investigated by Scheffé and Padia. Although research into the effect of serial correlation on specific quality control techniques appears to be scarce, Johnson and Bagshaw [39][40] have investigated its effect on the CUSUM chart. Their primary conclusion is that the cusum chart is not robust to departures from independence.

Chapter II of this dissertation is an attempt to partially fill the existing gap concerning the effect of correlated observations on Shewhart control charts for the mean. The adaptation and development of control charts for the mean in the presence of correlated sample values will be investigated for both univariate and multivariate characteristics. In the latter case, the quality of each item is dependent upon several characteristics. Thus, there is correlation within each vector of measurements as well as across the vectors of measurements for a given sample. However, there is no correlation among the vectors of different samples. The risk properties of the statistics developed for correlated observations will be explored, as well as the power of the control chart. Examples will be provided.

While Chapter II develops the statistics to be used in the presence of correlated observations, it leaves unanswered the questions of how large a sample to select and at what value should the control chart constant be set. The answers to these questions are investigated in Chapter IV in accordance with the scheme of minimizing the average run length of an out of control process for a large fixed value of the average run length of an in control process. This scheme was originally used by Page [48] for one quality characteristic and uncorrelated observations. By comparing our newly developed results with those of Page, the effect of correlated observations and multiple quality characteristics can be determined. Although the extension of Page's scheme does not answer the question of when to sample, it is felt that his scheme is the most easily understood and implemented.

A very useful representation of correlated observations is provided by the autoregressive, moving average models of order (p,q): where p represents the order of the autoregressive component and q that of the moving average component. Although these models have existed for quite some time, it is only within the last few years that they been widely adopted to model various temporal occurrences. One reason for this popularity is the publication of Box and Jenkins [13]. They have increased the flexibility of these models to allow for processes which exhibit a nonstationary level and a seasonality component. For this reason, the most general of these models are denoted as Box-Jenkins multiplicative empirical stochastic models of order (p,d,q) x $(P,D,Q)_S$ where d denotes the degree of differencing needed to achieve stationarity and the upper case letters refer to the order of the seasonality component. The popularity and success of Box-Jenkins models is evident by the increasing number of textbooks and journal articles in diverse areas devoted to this subject.

Some of the more recent textbooks are those of Nelson [46], T. W. Anderson [9], Fuller [26], and O. D. Anderson [8]. Although the number of journal articles is too exhaustive too list, the following represents a few of the varied applications: Saboia [54] improved present methods of forecasting births by using these models, Leuthold et al [45] used these models to forecast daily hog prices; Thompson and Tiao [63] analyzed telephone data with these models, and the list goes on. One of the most successful modeling applications has been by Deutsch [19], [20] and Deutsch and Rardin [22], [23], who employed these models in describing monthly crime occurrences. They have shown that each of the seven index crimes across ten different cities was represented by the same form of model.

In discussing the estimation of the model parameters, Box and Jenkins [13] given primary emphasis to the estimation of the autoregressive and moving average parameters with only a passing interest in estimating the level of the series. The need for a reversal of this emphasis arose with the introduction of the interrupted time-series quasi-experiemnt (ITSQE) by Campbell and Stanley [15] in 1963. In the ITSQE, n_1 equally spaced observations are available prior to the implementation or occurrence of some treatment. After the intervention occurs, a set of n_2 observations becomes available. For example, the observations might be the monthly occurrences of homicide for the city of Boston and it is suspected that a change in the level could occur because of the introduction of a gun control law.

Inferential statistical methods for the ITSQE were first developed by Box and Tiao [14] for the integrated first-order moving average process. Although their results were applicable to only this model, it enabled improved data analyses to be performed for many diverse areas. For example, Glass [27] used the Box and Tiao results to analyze the Connecticut speeding crackdown, while Deutsch and Alt [21] used it to investigate Massachusetts' gun control law. Glass, Wilson, and Gottman [28] extended the Box and Tiao results to include other types of models. However, their model formulations assume that the autoregressive, moving average parameters (which describe the process covariance) before the intervention are the same as those afterwards. In Chapter III, these models are made more flexible to allow for the consequences of the intervention affecting these parameters as well as the process level. For this reason, the extended models are called multi-consequence intervention models.

After formulating the multi-consequence intervention model, Chapter III then considers the estimation of the model parameters via least squares and maximum likelihood. The least squares procedure consists of transforming the original $n_1 + n_2$ observations to another set of variables amenable to statistical linear model analysis. In the maximum likelihood estimation procedure, explicit expressions can be obtained for the estimates of the level and shift parameters for fixed values of the moving average parameters. While such closed form expressions do not exist for the maximum likelihood estimates of the moving average and autoregressive parameters, an algorithm is presented for the numerical computation of these estimates. Chapter III also demonstrates how the least squares estimates differ from the maximum likelihood estimates.

In investigating the effect of an intervention on a temporal sequence of occurrences, it quite frequently occurs that the intervention has also affected another temporal sequence of occurrences. For example, when a municipality introduces a gun control law, it may not only affect the level of its own monthly occurrences of homicide, but also the levels of surrounding municipalities. In order to study the simultaneous effect of an intervention on two or more temporal sequences of occurrences, Chapter V introduces the multivariate multi-consequence intervention model. Chapter V also considers the least squares estimation of the model parameters.

Lastly, Chapter VI discusses conclusions and directions for future research.

CHAPTER II

CONTROL CHARTS FOR CORRELATED OBSERVATIONS

In Chapter I, a brief introduction to the concept of statistical quality control was presented. The first section of this chapter elaborates upon that introduction by reviewing the statistical basis of the traditional control chart used to maintain surveillance over the process mean when there is only one quality characteristic and the sample is random. It is shown that the statistic used to monitor this control has favorable risk properties both in the traditional sense as well as in the Bayesian and minimax interpretations. Section 2.1.2 extends the work of the first section by allowing the sample elements to possess any type of known autocorrelative structure. The statistic used here to monitor control also enjoys favorable risk characteristics, and this statistic reduces to that used in the first section when there is no autocorrelation. Section 2.2.1 continues to elaborate upon the first section by assuming that the quality of process output is governed by several characteristics. Thus, each sample element is a vector of correlated observations. However, as in Section 2.2.1, it is assumed that the sample elements are uncorrelated. Again, we explore the risk properties of the statistic used. Section 2.2.2 treats the most general problem: the quality of process output is governed by several characteristics and there is correlation across the vectors of observations. After exploring the risk properties of the statistic used, we formulate a decision rule for maintaining control over the process mean vector. To illustrate the foregoing concepts, several examples are presented. Selected portions of this chapter appear in a paper by Alt, Deutsch, and Walker [5].

2.1 One Quality Characteristic

2.1.1 Independent Observations

When there is only one characteristic determining the quality of output from a process, measurements denoted by $x_1, x_2, \ldots, x_n$ are obtained from a sample output of size n, and these measurements are used to make inferences about the process quality. If the process has been refined to the extent that assignable causes are not affecting the variation in the measurements, then any remaining variation can be attributed solely to chance causes, which are inherent in the process. Thus, in the absence of assignable causes of variation, the measurements should behave as a sample coming from a probability distribution which has a certain mean and variance. If this is indeed the case, then the process is said to be in a state of statistical control. To maintain surveillance over the state of control, successive samples of size n are obtained, a summary statistic is calculated from each sample's measurements, and this statistic is plotted on a control chart. If this statistic falls within certain limits on the control chart, the process is judged to be in control.

This chapter will consider the use of control charts for watching over the <u>mean</u> of a process, when the process is already in an existing state of control. That is, the process has been refined and has evolved to an in-control state where the underlying probability distribution is completely known, and the parameters of the distribution have stabilized to fixed values with interest centered on the mean of the probability distribution. Guttman, Wilks, and Hunter [35] state that control charts used for this purpose are essentially devices for detecting important departures from an existing known state of statistical control and are called "theoretical control charts." Duncan [24] refers to these charts as "charts for attaining current control." Quite frequently, they are also called "charts based on standard values" because parameter values are specified at which the process can hopefully be controlled.

Let X be the random variable associated with the underlying probability distribution of the measurements. It is not unreasonable to assume that X is normally distributed, denoted $X \sim N$. Let μ_0 and σ_I^2 denote the standard or nominal values of the process mean and variance, respectively. The values of μ_0 and σ_I^2 may be derived from past data (where the data base is sufficiently large so that μ_0 and σ_I^2 may be treated as parameter values and not their estimates), determined from experience with similar past processes, or selected to attain certain objectives.

In a single sample of size n from X, let $X_1, X_2, \ldots, X_n$ denote the elements of the sample and assume that $X_1, X_2, \ldots, X_n$ constitute a random sample. That is, $X_1, X_2, \ldots, X_n$ are <u>independent</u> random variables such that

$$f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_X(x_i) \qquad (1)$$

where

$$f_X(x_i) = (2\pi\sigma_I^{\,2})^{-1/2} \exp\{-(x_i - \mu)^2/(2\sigma_I^{\,2})\} \tag{2}$$

Using equations (1) and (2), it is easily shown (see, for example, Hoel, Port, and Stone [37]) that the maximum likelihood estimator of μ is given by $\bar{X}$ and that $\bar{X}$ is a sufficient, unbiased estimator of μ with variance given by $\sigma_I^{\,2}/n$. Although $\bar{X}$ has very many other desirable statistical properties, only those relating to its risk will be stated. Let $\mathcal{D}_0$ denote the class of all unbiased estimators of μ. Then, in $\mathcal{D}_0$,

(i) $\bar{X}$ is the uniformly minimum variance unbiased estimator of μ, frequently denoted UMVUE;

(ii) $\bar{X}$ is a Bayesian estimator with respect to every prior when the loss function is quadratic; and,

(iii) $\bar{X}$ is a minimax estimator when the loss function is quadratic.

It will be shown that (i) implies (ii) and (iii).

Property (i) is a direct result of the Cramér-Rao inequality and is demonstrated in numerous textbooks (see [37], [65]). If $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)^t$ and $d(\underset{\sim}{X})$ is any other estimator belongton to $\mathcal{D}_0$, then property (i) can be stated as $\sigma_{\bar{X}}^{\,2} = E_{\underset{\sim}{X}}(\bar{X} - \mu)^2 \leq E_{\underset{\sim}{X}}(d(\underset{\sim}{X}) - \mu)^2 = \sigma_d^{\,2}$ for all μ where the subscript $\underset{\sim}{X}$ on the expected value operator indicates that the expectation is over the sample result space of the X_i's. Property (i) can also be restated in terms of loss and risk. Let $L(\mu, d(\underset{\sim}{X}))$ be the loss associated with using the estimator d when μ is the true process mean. Only quadratic loss functions will be considered whereby $L(\mu, d(\underset{\sim}{X})) = (d(\underset{\sim}{X}) - \mu)^2$. Let $R(\mu, d)$ denote the risk incurred when μ is the true mean and the estimator $d(\underset{\sim}{X})$ is used. Intuitively, it seems reasonable to minimize the average quadratic loss and this is generally defined to be the risk. Specifically,

$$\mathcal{R}(\mu,\ d) = E_{\underset{\sim}{X}}\ \mathcal{L}(\mu,\ d(\underset{\sim}{X})) = E_{\underset{\sim}{X}}(d(\underset{\sim}{X}) - \mu)^2 = \sigma_d{}^2.$$

Thus the mean square error reduces to the variance in $\mathcal{D}_0$. Note that the risk is usually a function of μ. Thus, to say that $\bar{X}$ is a minimum variance estimator in $\mathcal{D}_0$ is equivalent to saying that $\bar{X}$ is a minimum risk estimator for all μ when quadratic loss is used, and then $\mathcal{R}(\mu,\ \bar{X}) = \sigma_I{}^2/n$.

In order to demonstrate property (ii), recall that the mean risk denoted by $r(\pi,\ d(\underset{\sim}{X}))$, for a given prior distribution $\pi(\mu)$ and estimator $d(\underset{\sim}{X})$ is defined to be

$$r(\pi,\ d(\underset{\sim}{X})) = E_\mu\ \mathcal{R}(\mu,\ d) \quad = \quad E_\mu[E_{\underset{\sim}{X}|\mu}(d(\underset{\sim}{X}) - \mu)^2]\ ,$$

when quadratic loss is used. The estimator d_0 is called a Bayes' rule if it minimizes the mean risk when the prior is $\pi(\mu)$, and $r(\pi,\ d_0(\underset{\sim}{X}))$ is called the Bayes' risk. When the loss function is quadratic, Hoel, Port, and Stone [37] show that the Bayes' rule is the mean of the posterior distribution of μ. To show that $\bar{X}$ is a Bayes' rule with respect to every prior for a quadratic loss in the class $\mathcal{D}_0$ of unbiased estimators, let $d(\underset{\sim}{X})$ be any other estimator in $\mathcal{D}_0$. Then, from property (i),

$$\begin{aligned} r(\pi,\ d(\underset{\sim}{X})) &= E_\mu[E_{\underset{\sim}{X}|\mu}(d(\underset{\sim}{X}) - \mu)^2] \\ &\geq \sigma_{\bar{X}}{}^2 = E_\mu[E_{\underset{\sim}{X}|\mu}(\bar{X} - \mu)^2] = r(\pi,\ \bar{X})\ , \qquad (3) \end{aligned}$$

where $\sigma_{\bar{X}}^2 = \sigma^2/n$. However, in general, when a quadratic loss function is used, there will be biased estimators which have less mean risk relative to a given prior than any unbiased estimator. In order to demonstrate this, consider the following example from Hoel, Port, and Stone. Let $X_1, X_2, \ldots, X_n$ denote a random sample from X, where $X \sim N(\mu, \sigma_I^2)$ with μ unknown and σ_I^2 known; and, let the prior density also be normal with mean β and variance α^2, both of which are specified. Then the Bayes' rule is given by

$$d_0(\underset{\sim}{X}) = (\sigma_I^2 \beta + \alpha^2 n\bar{X})/(\sigma_I^2 + \alpha^2 n)$$

with a Bayes' risk equal to

$$r(\pi, d_0(\underset{\sim}{X})) = (\alpha^2 \sigma_I^2)/(\sigma_I^2 + \alpha^2 n) .$$

Thus,

$$r(\pi, d_0(\underset{\sim}{X}))/r(\pi, \bar{X}) = [1 + (\sigma_I^2/n\, \alpha^2)]^{-1} .$$

Since $(\sigma_I^2/n\, \alpha^2) > 0$, $[1 + (\sigma_I^2/n\, \alpha^2)]^{-1} < 1$, and $r(\pi, d_0(\underset{\sim}{X})) \leq r(\pi, \bar{X})$. Thus the mean risk for the biased Bayes' estimator is less than the mean risk for the unbiased Bayes' estimator.

In order to show property (iii), recall that an estimator d_0 is said to be a minimax estimator in the class $\mathcal{D}$ of estimators if

$$\max_{\mu} R(\mu, d_0) = \min_{d \varepsilon \mathcal{D}} \max_{\mu} R(\mu, d) .$$

Since $\mathcal{D}$ is restricted to $\mathcal{D}_0$, the class of unbiased estimators, and a quadratic loss function is being used, $R(\mu, d) = \sigma_d^2$. By property (i),

$\sigma_{\bar{X}}^{2} \leq \sigma_{d}^{2}$ for every μ and for all $d \epsilon \mathcal{D}_0$, or $\sup_{\mu} \sigma_{\bar{X}}^{2} \leq \sup_{\mu} \sigma_{d}^{2}$ for all $d \epsilon \mathcal{D}_0$. That is, $\sup_{\mu} \sigma_{\bar{X}}^{2} = \min_{d \epsilon \mathcal{D}_0} \sup_{\mu} \sigma_{d}^{2}$, and property (iii) is established.

Properties (i), (ii), and (iii) essentially state that, in the class $\mathcal{D}_0$, $\bar{X}$ is an estimator with uniformly best risk for all μ and thus this estimator is a minimax estimator as well as a Bayes' rule, regardless of the chosen prior. As stated by Walker [66], $\bar{X}$ would "be eminently satisfactory from a minimal risk point of view."

Additional rationale for using $\bar{X}$ as the estimator for μ is provided by thinking of the elements in the sample as being generated from the following linear model:

$$\underbrace{\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}}_{\underset{\sim}{X}} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{A} \mu + \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}}_{\underset{\sim}{U}} \tag{4}$$

where the disturbances a_i are such that

$$E(\underset{\sim}{U}) = \underset{\sim}{0}, \quad E(\underset{\sim}{U}\,\underset{\sim}{U}^t) = \sigma_a^2 \, I_n \tag{5}$$

The assumptions stated in equations (4) and (5) specify what is known as the classical linear regression model with the exception that σ_a^2 is known. Under these conditions, the Gauss-Markov theorem states that the best linear unbiased estimator of μ is given by the least-squares estimator $(A^t A)^{-1} A^t \underset{\sim}{X}$, which reduces to $\bar{X}$, and the variance of this least-squares estimator is given by $\sigma_a^2 (A^t A)^{-1}$, which reduces to σ_a^2/n. A proof of the Gauss-Markov least squares theorem is given in Goldberger [29]. Thus, even though the least-squares estimator is identical with the maximum likelihood estimator, it was derived under different assumptions, the most important of which is the absence of any distributional assumptions concerning the a_i's and equivalently of the X_i's. In this absence, one cannot say $\bar{X}$ is normally distributed without reverting to the Central Limit Theorem. Note that, when the linear model is assumed as a process generator, the variance of the a_i's is identical with the variance of the X_i's, or $\sigma_a^2 = \sigma_I^2$.

Now that justification has been given for using $\bar{X}$ as an estimator for μ, it will be shown how $\bar{X}$ is used in maintaining statistical control of μ_0.

When there is only one quality characteristic, which is normally distributed, with standard values specified for the process mean and variance and successive random samples of size n are generated from this process. Shewhart [59] proposed that in order to maintain surveillance over μ_0 one should plot the successive values of $\bar{X}$ on a chart which has a central line (CL) and upper (UCL) and lower (LCL) control limits of the form:

$$\left.\begin{aligned} UCL &= \mu_0 + z_{\alpha/2}\ (\sigma_I/\sqrt{n}) \\ CL &= \mu_0 \\ LCL &= \mu_0 - z_{\alpha/2}\ (\sigma_I/\sqrt{n})\ . \end{aligned}\right\} \qquad (6)$$

The quantity $z_{\alpha/2}$ denotes the upper $\alpha/2$ percentage point of the standard normal random variable. Usually, $z_{\alpha/2} = 3.0$. A typical $\bar{X}$-chart is shown in Figure 1. If any $\bar{x}$'s plot above UCL or below LCL,

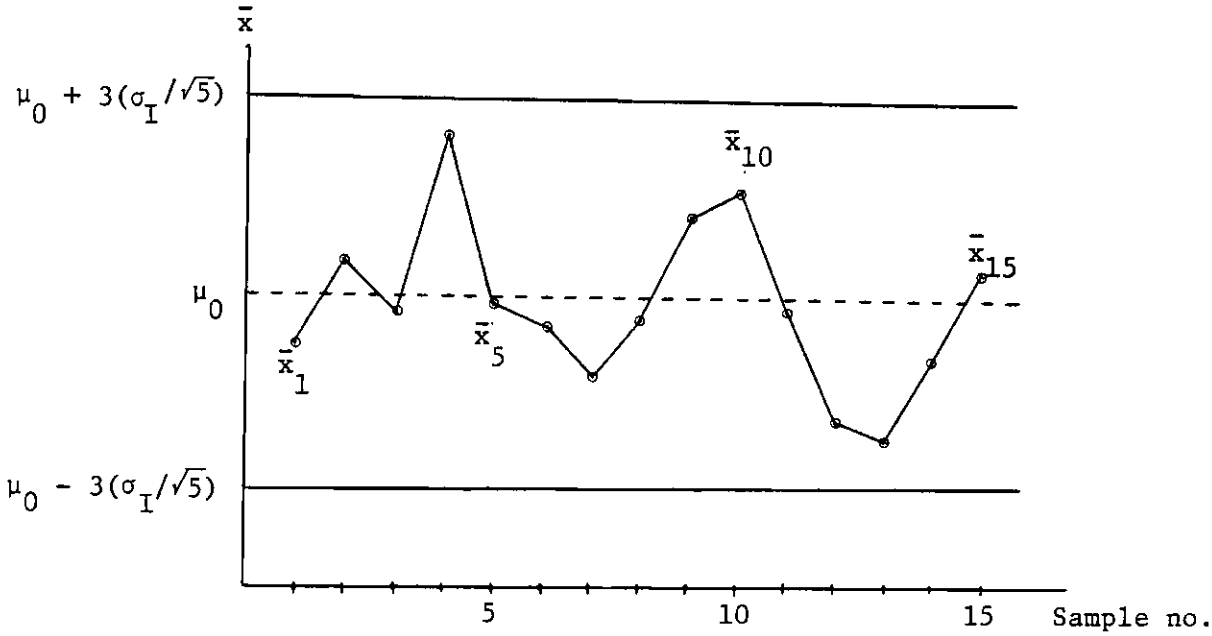

Figure 1. An $\bar{X}$-chart when n = 5

then a search is undertaken for any assignable causes.

The rationale behind the limits presented in equation (6) is that since $X \sim N(\mu_0, \sigma_I^2)$ when the process mean equals the nominal value μ_0 and $X_1, X_2, \ldots, X_n$ is a random sample from X, then $\bar{X} \sim N(\mu_0, \sigma_I^2/n)$ and

$$P[\mu_0 - z_{\alpha/2}(\sigma_I/\sqrt{n}) \leq \bar{X} \leq \mu_0 + z_{\alpha/2}(\sigma_I/\sqrt{n})] = 1 - \alpha . \tag{7}$$

Thus, the center line (CL) is set equal to $E(\bar{X})$ while the upper and lower control limits equal $E(\bar{X}) \pm k\sigma_{\bar{X}}$, where $k > 0$. Here $k = z_{\alpha/2}$.

For a single sample of size n, the control chart technique can also be viewed as a hypothesis testing problem. Namely, one is testing

$H_0:\mu = \mu_0$ vs. $H_1:\mu \neq \mu_0$ with known σ_I. The likelihood ratio test, details of which are given in Freund [25], yields the following critical region:

$$\begin{aligned} \omega = \{(x_1, x_2, \ldots x_n): \bar{x} < \mu_0 - z_{\alpha/2}(\sigma_I/\sqrt{n})\} \\ \cup \{(x_1, x_2, \ldots, x_n): \bar{x} > \mu_0 + z_{\alpha/2}(\sigma_I/\sqrt{n}\} \ , \end{aligned} \tag{8}$$

when $\mu = \mu_0$. Note that there is a one-to-one correspondence between the out-of-control region of the $\bar{X}$-chart and the critical region of the likelihood ratio test, given in equation (8). Thus, the control chart has associated with it the probability of Type I error, denoted by α, which is the probability of saying that the process mean has shifted from μ_0 when, in fact, it has not. In this instance, a search would be made for an assignable cause when none exists. When the control chart constant $z_{\alpha/2}$ is set equal to 3.0, $\alpha = 0.0027$ and only rarely would a search be made for a nonexistent assignable cause. Also inherent in the hypothesis testing viewpoint is the concept of the power of the control chart, denoted by $\pi(\mu_1)$, which is the probability of detecting that the process mean has shifted from μ_0 to a value μ_1. It is easily shown that

$$\pi(\mu_1) = \Phi(-z_{\alpha/2} - \delta(\sqrt{n}/\sigma_I)) + \Phi(-z_{\alpha/2} + \delta(\sqrt{n}/\sigma_I)) \ , \tag{9}$$

where $\delta = \mu_0 - \mu_1$ and Φ denotes the cumulative distribution function of the standard normal random variable. When the hypothesis testing viewpoint is adopted for successive samples of size n, the $\bar{X}$-control chart technique can be viewed as repeated tests of significance. That is, the decision maker is successively testing $H_0:\mu = \mu_0$ vs. $H_1:\mu \neq \mu_0$.

Inherent in the development of the likelihood ratio test critical region were the assumptions of process normality and independence of the sample elements, which again stresses the importance of these assumptions. In the next section, departures from the independence assumption will be investigated.

2.1.2 Dependent Observations

The development of the control limits, presented in equation (6), was based on the assumptions that $X \sim N(\mu_0, \sigma_I^2)$ and $X_1, X_2, \ldots, X_n$ was a random sample from X. This current section covers the development of control charts for the mean when the sampled elements are correlated. As a first step in this direction, we will consider what the effects are when one uses the control limits given in equation (6) when, in fact, the observations have a first-order serial correlation with the serial correlation coefficient donated by ρ. Many authors, starting at least as far back as Student [62], have reported the presence of such correlation in their successive measurements.

To investigate the effect of serial correlation, assume that the n sampled elements are jointly normal with $E(X_i) = \mu_0$, $\text{var}(X_i) = \sigma_c^2$, and $\text{Cov}(X_i, X_{i+j}) = \rho\sigma_c^2$ for j = 1 and 0 otherwise. If we let $\underset{\sim}{X}^t = [X_1, X_2, \ldots, X_n]$, then the joint density of X_i's is given by

$$f_{\underset{\sim}{X}^t}(\underset{\sim}{x}^t\,;\mu_0) = (2\pi)^{-n/2}|\Sigma_n|^{-1/2} \exp\{-(1/2)(\underset{\sim}{x} - \underset{\sim}{\mu})^t\, \Sigma_n^{-1}(\underset{\sim}{x} - \underset{\sim}{\mu})\}\ , \qquad (10)$$

where $\underset{\sim}{\mu}$ denotes $E(\underset{\sim}{X})$ and Σ_n denotes the $(n \times n)$ covariance matrix of the X_i's. Here

$$\underset{\sim}{\mu} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \mu_0 = \underset{\sim}{j}_n \, \mu_0 \quad \text{and} \quad \Sigma_n = \begin{bmatrix} \sigma_c^2 & \rho\sigma_c^2 & \cdots & 0 \\ \rho\sigma_c^2 & \sigma_c^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_c^2 \end{bmatrix}. \tag{11}$$

If P_n is the correlation matrix, then $P_n = D(1/\sigma_c)\, \Sigma_n\, D(1/\sigma_c)$ where $D(1/\sigma_c)$ is the $(n \times n)$ diagonal matrix with entries $(1/\sigma_c)$. Grenander and Rosenblatt [34] have shown that a necessary and sufficient condition for P_n to be positive definite is that $|\rho| < (2\cos[\pi/(n+1)])^{-1}$. Thus, all values of ρ in the interval $(-1, 1)$ are not possible. However, as noted by Scheffé [56], it follows that all values in $(-1/2, 1/2)$ are possible for all n. Scheffé has also shown that, under the conditions in equations (10) and (11),

$$\bar{X} \sim N(\mu_0,\ (\sigma_c^2/n)[1 + 2\rho(1 - n^{-1})]). \tag{12}$$

Thus, serial correlation affects only the dispersion of $\bar{X}$ and not its location. Equation (12) can be used to determine the true probability of Type I error, denoted by α_0, when one uses the control limits given by equation (6) assuming a nominal probability of Type I error denoted by α. Specifically, if $B = [1 + 2\rho(1 - n^{-1})]$, then

$$\alpha_0 = P[|(\bar{X} - \mu_0)\sqrt{n}/\sigma_c| > z_{\alpha/2}] = P[|(\bar{X} - \mu_0)\sqrt{n}/B\sigma_c| > z_{\alpha/2}/B]$$

$$= P[|Z| > z_{\alpha/2}/B] ,$$

where Z denotes the standard normal random variable. For $\alpha = .05$, Scheffé has prepared a table giving the effect of $\rho = (-0.4)(0.1)(+0.4)$

on α_0. The table indicates that for $\rho < 0$, $\alpha_0 < \alpha = 0.05$, while for $\rho > 0$, $\alpha_0 > \alpha = 0.05$. Obviously for $\rho = 0$, $\alpha_0 = \alpha$. Since Scheffé's table was prepared exclusively for $\alpha = 0.05$, and large n, Table 1 was prepared to indicate the effect of ρ on α_0 when the nominal level of significance is 0.0027 and n = 4 and 5, values frequently used by quality control decision makers. This table was prepared using Univac's PNORM subroutine in the MSFLIB Library. These results agree with those of Scheffé's in that, for $\rho < 0$, the true probability of Type I error (α_0) is less than the nominal value of .0027 while, for $\rho > 0$, the true probability is greater than the nominal value. Inspection of Table 1 also reveals that, for $\rho < 0$, α_0 decreases as n increases from 4 to 5, while, for $\rho > 0$, α_0 increases as n increases from 4 to 5.

Table 1. Values of Actual Significance Level (α_0) when Nominal Level is 0.0027

ρ	-0.4	-0.3	-0.2	-0.1	0.0	0.1	0.2	0.3	0.4
n=4	$.21 \cdot 10^{-5}$	$.52 \cdot 10^{-4}$	$.34 \cdot 10^{-3}$	0.0011	0.0027	0.0052	0.0085	0.0127	0.0177
n=5	$.57 \cdot 10^{-6}$	$.32 \cdot 10^{-4}$	$.27 \cdot 10^{-3}$	0.0011	0.0027	0.0053	0.0090	0.0137	0.0192

Equation (12) can also be used to derive a revised set of control limits when ρ is known from a large amount of past data or determined from experience with similar past processes. It immediately follows that

$$\left.\begin{aligned} UCL &= \mu_0 + z_{\alpha/2}(\sigma_c/\sqrt{n})[1 + 2\rho_0(1 - n^{-1})]^{1/2} \\ CL &= \mu_0 \\ LCL &= \mu_0 - z_{\alpha/2}(\sigma_c/\sqrt{n})[1 + 2\rho_0(1 - n^{-1})]^{1/2} , \end{aligned}\right\} \qquad (13)$$

where ρ_0 denotes the standard value of the serial correlation coefficient. The limits given by equation (13) differ from those derived under the assumption of independence by the factor $B = [1 + 2\rho_0(1 - n^{-1})]^{1/2}$, where $B = 1$ for $\rho_0 = 0$. Thus, if $\rho_0 = 0$, the control limits given by equation (13) are identical with those of equation (6) provided σ_c (the standard deviation of the correlated observations) equals σ_I (the standard deviation of the uncorrelated observations). Additional explanation of the relationship between σ_c and σ_I will be provided later in this section.

Padia [47] extended Scheffé's work by investigating samples which have a k^{th} order autocorrelative structure. If ρ_k denotes the last nonzero lag autocorrelation, then he has shown that

$$\text{Var }(\bar{X}) \doteq (\sigma_c^2/n)[1 + 2(\rho_1 + \rho_2 + \ldots + \rho_k)] \qquad (14)$$

to order $(1/n)$. Using equation (14), he determined the effect of various autocorrelative structures on the true probability of Type I error when testing $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$. Although one could use Padia's results in establishing new control limits, it would be preferable to have a more general approach. The approach that will be taken is to find the maximum likelihood estimator of μ, when there is any type of known correlative structure. Since a geometrical approach will be adopted, a few basic properties of n-dimensional Euclidean space, denoted by R^n, will be reviewed.

One of the most frequently occurring examples of a real vector space is $R^n = \{(x_1, x_2, \ldots, x_n) : x_i \text{ is real}\}$. Let $\underset{\sim}{e}_j$ denote the $(n \times 1)$ vector with a 1 in the j^{th} position and 0's elsewhere. Then, every vector $\underset{\sim}{x} \varepsilon R^n$ is such that $\underset{\sim}{x} = \sum_{j=1}^{n} x_j \underset{\sim}{e}_j$ and $\{\underset{\sim}{e}_1, \ldots, \underset{\sim}{e}_n\}$ is called the standard basis for R^n. Another operation that is frequently defined on R^n is the inner product of two vectors, denoted by $<\cdot, \cdot>$, where for $\underset{\sim}{x}, \underset{\sim}{y} \varepsilon R^n$, $<\underset{\sim}{x}, \underset{\sim}{y}>$ is defined to be $\underset{\sim}{x}^t \underset{\sim}{y} = \sum_{i=1}^{n} x_i y_i$. Since the inner product defined on R^n is nonnegative ($<\underset{\sim}{x}, \underset{\sim}{x}> \geq 0$ with $<\underset{\sim}{x}, \underset{\sim}{x}> = 0$ if and only if $\underset{\sim}{x} = \underset{\sim}{0}$), commutative ($<\underset{\sim}{x}, \underset{\sim}{y}> = <\underset{\sim}{y}, \underset{\sim}{x}>$), and linear ($<\underset{\sim}{x}_1, \alpha\underset{\sim}{x}_2 + \underset{\sim}{x}_3> = \alpha<\underset{\sim}{x}_1, \underset{\sim}{x}_2> + <\underset{\sim}{x}_1, \underset{\sim}{x}_3>$), the ordered pair $(R^n, <\cdot, \cdot>)$ is said to be an inner product space. In an inner product space, two vectors $\underset{\sim}{x}, \underset{\sim}{y} \varepsilon R_n$ are said to be orthogonal is $<\underset{\sim}{x}, \underset{\sim}{y}> = 0$. Thus, the basis vectors $\underset{\sim}{e}_j$ are mutually orthogonal. For $\underset{\sim}{x} \varepsilon R^n$, define the norm of $\underset{\sim}{x}$, denoted by $||\underset{\sim}{x}||$, to be $||\underset{\sim}{x}|| = (\sum_{i=1}^{n} x_i^2)^{1/2} = (\underset{\sim}{x}^t \underset{\sim}{x})^{1/2}$. Since the norm defined on R^n satisfies the properties (i) $||\underset{\sim}{x}|| = 0$ if and only if $\underset{\sim}{x} = \underset{\sim}{0}$, (ii) $||\alpha\underset{\sim}{x}|| = |\alpha| \; ||\underset{\sim}{x}||$, where α is a member of the reals, and (iii) $||\underset{\sim}{x}_1 + \underset{\sim}{x}_2|| \leq ||\underset{\sim}{x}_1|| + ||\underset{\sim}{x}_2||$, the ordered pair $(R^n, ||\cdot||)$ is called a normed vector space. The standard basis for R^n is said to be orthonormal since $||\underset{\sim}{e}_j|| = 1$. Since the square root of the inner product defines a norm for R^n, the ordered pair $(R^n, <\underset{\sim}{x}, \underset{\sim}{x}>^{1/2})$ is a normed vector space. For $\underset{\sim}{x}, \underset{\sim}{y} \varepsilon R^n$, define $d(\underset{\sim}{x}, \underset{\sim}{y}) = ||\underset{\sim}{x} - \underset{\sim}{y}|| = [\sum_{i=1}^{n} (x_i - y_i)^2]^{1/2}$. Since this real-valued function is nonnegative ($d(\underset{\sim}{x}, \underset{\sim}{y}) = 0$ if and only if $\underset{\sim}{x} = \underset{\sim}{y}$), commutative ($d(\underset{\sim}{x}, \underset{\sim}{y}) = d(\underset{\sim}{y}, \underset{\sim}{x})$), and satisfies the triangle inequality ($d(\underset{\sim}{x}, \underset{\sim}{z}) \leq d(\underset{\sim}{x}, \underset{\sim}{y}) + d(\underset{\sim}{y}, \underset{\sim}{z})$), the ordered pair (R^n, d) is called a metric space, d is called the metric, and $d(\underset{\sim}{x}, \underset{\sim}{y})$ is called the distance between $\underset{\sim}{x}$ and $\underset{\sim}{y}$. The metric defined above is frequently called the Euclidean metric. Additional details of these spaces can be found in Kasriel [41].

The Euclidean metric is very satisfactory for quite a few optimization problems. For example, in Section 2.1.1, the elements of the sample were independent, identical normally distributed random variables; and to find the maximum likelihood estimator of μ, it was required to find that value of μ which minimizes $Q(\mu) = \sum_{i=1}^{n} (x_i - \mu)^2 = ||\underset{\sim}{x} - \mu \underset{\sim}{j}_n||^2$. This is equivalent to finding the orthogonal projection of $\underset{\sim}{x}$ on L, the line generated by $\underset{\sim}{j}_n$, where this projection is merely some constant multiple of $\underset{\sim}{j}_n$, denoted by $c\underset{\sim}{j}_n$. It immediately follows that $\underset{\sim}{x} - c\underset{\sim}{j}_n$ is orthogonal to every vector in L. Specifically, $<\hat{\mu}\, \underset{\sim}{j}_n, \underset{\sim}{x} - \hat{\mu}\, \underset{\sim}{j}_n> = 0$, and $\hat{\mu} = <\underset{\sim}{j}_n, \underset{\sim}{x}>/<\underset{\sim}{j}_n, \underset{\sim}{j}_n> = \bar{x}$, as previously stated. However, it is sometimes convenient to use a non-orthonormal basis and it is necessary to modify the inner product defined on R^n. For example, let B be an $(n \times n)$ nonsingular, symmetric matrix and let $\underset{\sim}{w} = B\underset{\sim}{x}$. Then $<\underset{\sim}{x}, \underset{\sim}{x}>_I = \underset{\sim}{x}^t\underset{\sim}{x} = \underset{\sim}{w}^t(B^t)^{-1}\, B^{-1}\, \underset{\sim}{w} = \underset{\sim}{w}^t\, A\underset{\sim}{w} = <\underset{\sim}{w}, \underset{\sim}{w}>_A$ where $A = (BB^t)^{-1}$. Now if A is an $(n \times n)$ positive definite matrix and $<\underset{\sim}{w}, \underset{\sim}{w}>_A$ is defined to be $\underset{\sim}{w}^t\, A\underset{\sim}{w}$ for $\underset{\sim}{w} \epsilon R^n$, then $(R^n, <\cdot, \cdot>_A)$ is an inner product space, $(R^n, ||\cdot||)$ is a normed vector space with $||\underset{\sim}{w}|| = <\underset{\sim}{w}, \underset{\sim}{w}>_A$, and (R^n, d) is a metric space with $d(\underset{\sim}{w}_1, \underset{\sim}{w}_2) = ||\underset{\sim}{w}_1 - \underset{\sim}{w}_2||$. An additional explanation of inner product spaces in the metric of A is given in Timm [64].

The maximum likelihood estimator of μ will now be found.

Theorem 2.1: Let $X_1, X_2, \ldots, X_n$ be jointly normal with mean vector $\underset{\sim}{\mu}$ and covariance matrix Σ as given in equation (15):

$$\underset{\sim}{\mu} \begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \mu = \underset{\sim}{j}_n ; \quad \Sigma_n = \begin{bmatrix} \sigma_c^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_c^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_c^2 \end{bmatrix} \tag{15}$$

Then the maximum likelihood estimator of μ, denoted by $\hat{\mu}$ is given by

$$\hat{\mu} = (\underset{\sim}{X}^t \Lambda_n j_n)/(\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n) , \tag{16}$$

where $\Lambda_n = \Sigma_n^{-1}$.

Proof: The likelihood function, denoted by $L(\mu)$, is given by

$$L(\mu) = (2\pi)^{-(n/2)} |\Lambda_n|^{1/2} \exp \{-(1/2)(\underset{\sim}{x} - \mu \underset{\sim}{j}_n)^t \Lambda_n (\underset{\sim}{x} - \mu \underset{\sim}{j}_n)\} . \tag{17}$$

Since Λ_n is a positive definite matrix, $<\underset{\sim}{r}, \underset{\sim}{s}>_\Lambda = \underset{\sim}{r}^t \Lambda_n \underset{\sim}{s}$ is an inner product on R^n, $||\underset{\sim}{r}|| = <\underset{\sim}{r}, \underset{\sim}{r}>_\Lambda^{1/2}$ a norm, and $||\underset{\sim}{r} - \underset{\sim}{s}|| = <\underset{\sim}{r} - \underset{\sim}{s}, \underset{\sim}{r} - \underset{\sim}{s}>_\Lambda^{1/2}$ a metric for $\underset{\sim}{r}, \underset{\sim}{s} \varepsilon R^n$.

Now the likelihood function will be a maximum when the nonnegative quadratic form, $Q(\mu)$, in the exponent is a minimum, where $Q(\mu) = (\underset{\sim}{x} - \mu \underset{\sim}{j}_n)^t \Lambda_n (\underset{\sim}{x} - \mu \underset{\sim}{j}_n) = ||\underset{\sim}{x} - \mu \underset{\sim}{j}_n||^2$. Thus we wish to find that vector $\hat{\mu} \underset{\sim}{j}_n$, lying in the line L generated by $\underset{\sim}{j}_n$, which is closest to $\underset{\sim}{x}$. This is shown in Figure 2. But, the vector in L lying closest to $\underset{\sim}{x}$ is the projection of $\underset{\sim}{x}$ onto L, denoted by $P_L(\underset{\sim}{x})$. Since $\underset{\sim}{x} - P_L(\underset{\sim}{x})$ is orthogonal to every vector in L, $<P_L(\underset{\sim}{x}), \underset{\sim}{x} - P_L(\underset{\sim}{x})>_\Lambda = 0$ or $<\hat{\mu} \underset{\sim}{j}_n, \underset{\sim}{x} - \hat{\mu} \underset{\sim}{j}_n>_\Lambda = 0$. Thus, $\hat{\mu} <\underset{\sim}{j}_n, \underset{\sim}{x}>_\Lambda - \hat{\mu}^2 <\underset{\sim}{j}_n, \underset{\sim}{j}_n>_\Lambda = 0$ and $\hat{\mu} = <\underset{\sim}{j}_n, \underset{\sim}{x}>_\Lambda / <\underset{\sim}{j}_n, \underset{\sim}{j}_n>_\Lambda$.||

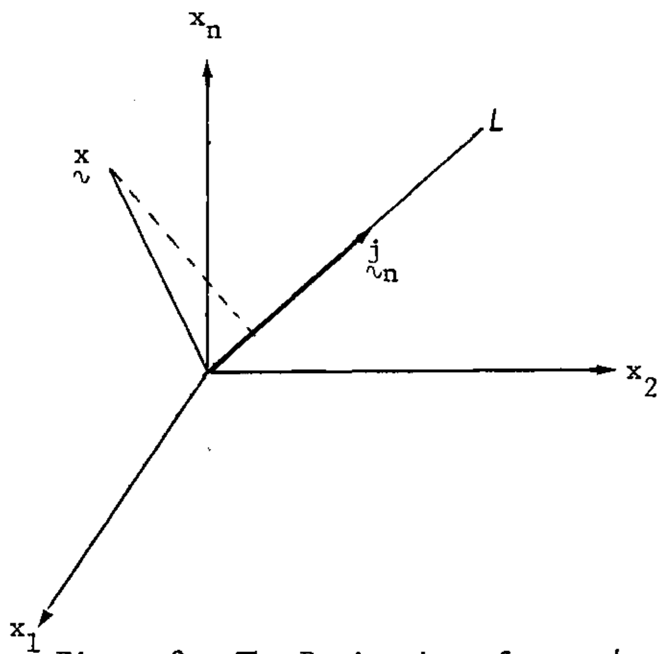

Figure 2. The Projection of $\underset{\sim}{x}$ on L.

The maximum likelihood estimator can also be written in summation notation:

$$\hat{\mu} = \sum_{j=1}^{n} \sum_{i=1}^{n} \lambda_{ij} X_{ij} / \sum_{j=1}^{n} \sum_{i=1}^{n} \lambda_{ij} = \sum_{j=1}^{n} [\sum_{i=1}^{n} \lambda_{ij} (\sum_{j=1}^{n} \sum_{i=1}^{n} \lambda_{ij})^{-1}] X_j , \quad (18)$$

where λ_{ij} are the entries in Λ_n. Thus, the numerator of $\hat{\mu}$ is merely the sum of all the X's where each X_j is weighted by the sum of the elements in the i^{th} row of Λ_n, and the denominator is the sum of all the entries in Λ_n. Since $\hat{\mu}$ is a linear combination of the X_j's, which are multivariate normal, then $\hat{\mu}$ is distributed as a univariate normal. The expected value and variance of $\hat{\mu}$ are obtained as follows:

$$E(\hat{\mu}) = (\underset{\sim}{j}_n^{\,t} \Lambda_n \underset{\sim}{j}_n)^{-1} [(E \underset{\sim}{X}^t) \Lambda_n \underset{\sim}{j}_n] = (\underset{\sim}{j}_n^{\,t} \Lambda_n \underset{\sim}{j}_n)^{-1} (\mu \underset{\sim}{j}_n)^t \Lambda_n \underset{\sim}{j}_n = \mu , \quad (19)$$

and

$$\text{Var}(\hat{\mu}) = E(\hat{\mu} - \mu)^2 = (\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)^{-2} E[(\underset{\sim}{X}^t \Lambda_n \underset{\sim}{j}_n - \mu \underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)^2]$$

$$= (\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)^{-2} E[\underset{\sim}{X}^t \Lambda_n j_n - \mu(\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)]^t[\underset{\sim}{X}^t \Lambda_n \underset{\sim}{j}_n - \mu(\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)]$$

$$= (\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)^{-2} \underset{\sim}{j}_n{}^t \Lambda_n [E(\underset{\sim}{X} - \mu \underset{\sim}{j}_n)(\underset{\sim}{X} - \mu \underset{\sim}{j}_n)^t] \Lambda_n \underset{\sim}{j}_n$$

$$= (\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)^{-2} \underset{\sim}{j}_n{}^t \Lambda_n \Sigma_n \Lambda_n \underset{\sim}{j}_n = (\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)^{-1}. \tag{20}$$

That $\hat{\mu}$ is an unbiased estimator of μ follows from equation (19). The above three properties of $\hat{\mu}$ can be combined by stating that

$$\hat{\mu} \sim N(\mu, (\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)^{-1}) \tag{21}$$

Let us now determine the risk properties of $\hat{\mu}$.

Let $\mathcal{D}_0$ denote the class of all unbiased estimators of μ when the sample is jointly normal with a mean vector and covariance matrix as given in equation (15). Then $\hat{\mu}$, the maximum likelihood estimator of μ given in Theorem 3.1, has the following properties in $\mathcal{D}_0$:

(i) $\hat{\mu}$ is the uniformly minimum variance estimator of μ;

(ii) $\hat{\mu}$ is a Bayesian estimator with respect to every prior when the loss function is quadratic; and,

(iii) $\hat{\mu}$ is a minimax estimator when the loss function is quadratic.

Statement (i) implies (ii) and (iii).

Property (i) can be established via several approaches. One approach makes use of the Cramér-Rao Lower Bound (CRLB), a precise statement of which can be found in Wilks [68]. Essentially, the Cramér-Rao inequality asserts that Var $(d(\underset{\sim}{X})) \geq 1/\text{Var}(W)$ where $d(\underset{\sim}{X})$ is any unbiased estimator for μ and $W = \partial(\ell n\, f_{\underset{\sim}{X}^t}(\underset{\sim}{X}^t; \mu))/\partial\mu$ where $f_{\underset{\sim}{X}^t}(\underset{\sim}{x}^t; \mu)$ is as stated in Equation (17). It can be shown that the regularity conditions of the Cramér-Rao inequality are met for unbiased estimators of finite variance in this case. Since $E(W) = 0$, $\text{Var}(W) = E(W^2)$. It follows from equation (17) that

$$\ell n\ f_{\underset{\sim}{X}^t}(\underset{\sim}{X}^t;\ \mu) = k - (1/2)(\underset{\sim}{X} - \mu\ \underset{\sim}{j}_n)^t\ \Lambda_n(\underset{\sim}{X} - \mu\ \underset{\sim}{j}_n)\ ,$$

where $k = -(n/2)\ \ell n\ (2\pi) + (1/2)\ \ell n\ |\Lambda_n|$, and

$$\partial(\ell n\ f_{\underset{\sim}{X}^t}(\underset{\sim}{X}^t;\ \mu))/\partial\mu = \underset{\sim}{X}^t\ \Lambda_n\ \underset{\sim}{j}_n - \mu\ \underset{\sim}{j}_n^{\ t}\ \Lambda_n\ \underset{\sim}{j}_n = \underset{\sim}{j}_n^{\ t}\ \Lambda_n\ \underset{\sim}{j}_n(\hat{\mu} - \mu) = W.$$

Now $E(W^2) = (\underset{\sim}{j}_n^{\ t}\ \Lambda_n\ \underset{\sim}{j}_n)^2\ E(\hat{\mu} - \mu)^2 = \underset{\sim}{j}_n^{\ t}\ \Lambda_n\ \underset{\sim}{j}_n$, and $\text{CRLB} = 1/E(W^2) = 1/\underset{\sim}{j}_n^{\ t}\ \Lambda_n\ \underset{\sim}{j}_n$, which equals the variance of $\hat{\mu}$. Hence $\hat{\mu}$ is a best estimator, in the minimum-variance sense, in $\mathcal{D}_0$. That is, its efficiency is 1 where the efficiency of an unbiased estimator is the ratio of the CRLB to its variance. Property (i) can also be shown by using the Lehmann-Scheffé Theorem, a precise statement of which can be found in Rohatgi [52]. Essentially this theorem asserts that if $d(\underset{\sim}{X})$ is an unbiased, complete, sufficient statistic for μ, then $d(\underset{\sim}{X})$ is the UMVUE of μ. Since the unbiasedness of $\hat{\mu}$ has been demonstrated in equation (19), we will now show that $\hat{\mu}$ is sufficient. This follows since $\hat{\mu}\ \underset{\sim}{j}_n = P_L(\underset{\sim}{x})$ and thus $\underset{\sim}{x} - \hat{\mu}\ \underset{\sim}{j}_n$ is orthogonal to every vector lying in L. Moreover, every vector in R^n decomposes uniquely into two orthogonal components, one lying in L and one in the orthogonal complement of L. Specifically,

$$\underset{\sim}{x} - \mu\ \underset{\sim}{j}_n = (\underset{\sim}{x} - \hat{\mu}\ \underset{\sim}{j}_n) + (\hat{\mu}\ \underset{\sim}{j}_n - \mu\ \underset{\sim}{j}_n).$$

Also recall that in an inner product space the Pythagorean property holds. Specifically,

$$||\underset{\sim}{x} - \mu \underset{\sim}{j}_n||^2 = ||\underset{\sim}{x} - \hat{\mu} \underset{\sim}{j}_n||^2 + (\hat{\mu} - \mu)^2 \; ||\underset{\sim}{j}_n||^2.$$

Thus,

$$f_{\underset{\sim}{X}^t}(\underset{\sim}{x}^t|\mu) = k\{\exp -(1/2)\|\underset{\sim}{x} - \hat{\mu} \underset{\sim}{j}_n\|^2\} \exp\{-(1/2)(\hat{\mu} - \mu)^2||\underset{\sim}{j}_n||^2\}$$

$$= g(\underset{\sim}{x}) \; h(\hat{\mu}, \mu) ,$$

and sufficiency is established by the factorization criterion (see Rohatgi [52]). As a next step in using the Lehmann-Scheffé theorem, recall that $\hat{\mu} \sim N(\mu, 1/\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n)$ and that this distribution is complete. That is,

$$(2\pi \; \sigma_{\hat{\mu}}^2)^{-1/2} \int_{-\infty}^{\infty} s(t) \exp\{-(t - \mu)^2/(2 \; \sigma_{\hat{\mu}}^2)\} \, dt = 0$$

for all μ implies $s(t) = 0$ almost everywhere, where $\sigma_{\hat{\mu}}^2 = 1/\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n$. Since $\hat{\mu}$ is sufficient and unbiased and since its distribution (which is normal) is complete with respect to the parameter μ, then the efficient property of $\hat{\mu}$ has also been established by the Lehmann-Scheffé theorem. That is, it possesses minimum variance in the class of unbiased estimators with finite variance.

Properties (ii) and (iii) are a direct result of property (i) and their proofs closely parallel the proofs in Section 2.1.1. Thus $\hat{\mu}$ is "eminently satisfactory from a minimal risk point of view."

Additional rationale for using $\hat{\mu}$ as the estimator for μ is provided by considering the same linear model given in equation (4), namely $\underset{\sim}{X} = A\mu + \underset{\sim}{U}$, where now the disturbances a_i are such that

$$E(\underset{\sim}{U}) = 0, \; E(\underset{\sim}{U} \, \underset{\sim}{U}^t) = \Sigma_n , \tag{22}$$

with known Σ_n as defined in equation (15). The conditions stated in equations (4) and (22) specify what is known as the generalized linear regression model with the exception that Σ_n is specified. Under these conditions, Aitken's generalized Gauss-Markov least-squares theorem asserts that the best (in a minimum variance sense) linear unbiased estimator of μ is given by the generalized least-squares estimator $(A^t\ \Sigma_n^{-1}\ A)^{-1}(A^t\ \Sigma_n^{-1}\ \underset{\sim}{X})$, which reduces to $(\underset{\sim}{j}_n^{\,t}\ \Lambda_n\ \underset{\sim}{X})/(\underset{\sim}{j}_n^{\,t}\ \Lambda_n\ \underset{\sim}{j}_n)$ since $A = \underset{\sim}{j}_n$ and $\Lambda_n = \Sigma_n^{-1}$. The variance of the generalized least-squares estimator is given by $(A^t\ \Sigma_n^{-1}\ A)^{-1}$, which reduces to $1/\underset{\sim}{j}_n^{\,t}\ \Lambda_n\ \underset{\sim}{j}_n$. Thus, the generalized least-squares estimator of μ, which was derived without any distributional assumptions about the a_i's and equivalently about the X_i's, is identical with the maximum likelihood estimator. Also note that improving a covariance structure of Σ_n upon the a_i's is equivalent to stating that the X_i's have a Σ_n covariance structure since $C(\underset{\sim}{X},\ \underset{\sim}{X}^t) = E(\underset{\sim}{X} - \underset{\sim}{\mu})(\underset{\sim}{X} - \underset{\sim}{\mu})^t = E(\underset{\sim}{j}_n\ \mu + \underset{\sim}{U} - \underset{\sim}{j}_n\ \mu)(\underset{\sim}{j}_n\ \mu + \underset{\sim}{U} - \underset{\sim}{j}_n\ \mu)^t = E(\underset{\sim}{U}\ \underset{\sim}{U}^t) = \Sigma_n$. The generalized least-squares estimator of μ could have also been obtained by transforming the disturbances by a nonsingular matrix R such that $C(R\underset{\sim}{U},\ (R\underset{\sim}{U})^t) = I$ and using ordinary least-squares on the transformed linear model $(R\underset{\sim}{X}) = (RA)\mu + (R\underset{\sim}{U})$ or $\underset{\sim}{X}^* = A^*\ \mu + \underset{\sim}{U}^*$. A detailed presentation of the generalized linear regression model and its equivalence to the classical linear regression model via a transformation can be found in Goldberger [29].

Control charts for the mean can now be constructed using the maximum likelihood estimator $\hat{\mu}$. When the process mean equals the nominal value μ_0, then $\hat{\mu} \sim N(\mu_0, 1/\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n)$ and

$$P[\mu_0 - z_{\alpha/2} \sqrt{1/\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n} \leq \hat{\mu} \leq \mu_0 + z_{\alpha/2} \sqrt{1/\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n}] = 1 - \alpha.$$

Thus, control chart limits for the mean in the presence of any type of known autocorrelative structure are given by

$$\left.\begin{aligned} UCL &= \mu_0 + z_{\alpha/2} \sqrt{1/\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n} \\ CL &= \mu_0 \\ LCL &= \mu_0 - z_{\alpha/2} \sqrt{1/\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n} , \end{aligned}\right\} \qquad (23)$$

where $\sqrt{1/\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n} = \sqrt{Var(\hat{\mu})}$. Thus, the control limits are of the form

$$E(\hat{\mu}) \pm z_{\alpha/2} \sqrt{Var(\hat{\mu})} .$$

As in the case of uncorrelated observations, the $\hat{\mu}$-control chart also has a power function, denoted by $\pi(\mu_1)$, where $\pi(\mu_1)$ is the probability of detecting that the process mean has shifted from μ_0 to μ_1. It is easily shown that

$$\pi(\mu_1) = \Phi(-z_{\alpha/2} - \delta(\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n)^{-1/2}) + \Phi(-z_{\alpha/2} + \delta(\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n)^{-1/2}) ,$$

where $\delta = \mu_0 - \mu_1$. The control limits presented in equation (23) reduce to those presented in equation (6) when the off-diagonal elements of Σ_n are zero and the diagonal elements are equal.

To gain additional insight into the nature of the control chart limits presented in equation (23), specifically consider the situation where the observations have a first-order serial correlation with mean vector $\underset{\sim}{\mu}$ and covariance matrix Σ_n as presented in equation (11). Let $n = 2, 3, 4$, and 5, sample sizes which occur frequently in practice. We first need to find $\Lambda_n = \Sigma_n^{-1}$. Since Σ_n is a diagonal matrix of type 2, its inverse exhibits certain properties which assists in its determination. See Greenberg and Sarhan [33]. For n = 2, 3, 4, 5, we arrive at the following set of Λ_n's:

$$\Lambda_2 = \frac{1}{\sigma_c^2(1-\rho^2)}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}, \quad \Lambda_3 = \frac{1}{\sigma_c^2(1-2\rho^2)}\begin{bmatrix} 1-\rho^2 & -\rho & \rho^2 \\ -\rho & 1 & -\rho \\ \rho^2 & -\rho & 1-\rho^2 \end{bmatrix}$$

$$\Lambda_4 = \frac{1}{\sigma_c^2(1-3\rho^2+\rho^4)}\begin{bmatrix} 1-2\rho^2 & -\rho(1-\rho^2) & \rho^2 & -\rho^3 \\ -\rho(1-\rho^2) & 1-\rho^2 & -\rho & \rho^2 \\ \rho^2 & -\rho & 1-\rho^2 & -\rho(1-\rho^2) \\ -\rho^3 & \rho^2 & -\rho(1-\rho^2) & 1-2\rho^2 \end{bmatrix} \tag{24}$$

$$\Lambda_5 = \frac{1}{\sigma_c^2(1-4\rho^2+3\rho^4)}\begin{bmatrix} 1-3\rho^2+\rho^4 & -\rho+2\rho^3 & \rho^2-\rho^4 & -\rho^3 & \rho^4 \\ -\rho+2\rho^3 & 1-2\rho^2 & -\rho+\rho^3 & \rho^2 & -\rho^3 \\ \rho^2-\rho^4 & -\rho+\rho^3 & 1-2\rho^2+\rho^4 & -\rho+\rho^3 & \rho^2-\rho^4 \\ -\rho^3 & \rho^2 & -\rho+\rho^3 & 1-2\rho^2 & -\rho+2\rho^3 \\ \rho^4 & -\rho^3 & \rho^2-\rho^4 & -\rho+2\rho^3 & 1-3\rho^2+\rho^4 \end{bmatrix}$$

From equation (24), we can easily calculate Var $(\hat{\mu})$, which is merely the reciprocal of the sum of the elements in Λ_n. These values are presented in Table 2. As a point of interest, Table 2 also contains Var $(\bar{X})$ as determined from equation (12). That Var $(\hat{\mu})$ < Var $(\bar{X})$ for $n \geq 3$ is not surprising since $\hat{\mu}$ is an efficient estimator for μ. To the quality control engineer, this means that the control limits are tighter when using $\hat{\mu}$ (see equation (23)) than when using $\bar{X}$ (see equation (13)), even though both control charts have the same center line μ_0.

Table 2. Comparison of Var $(\hat{\mu})$ and Var $(\bar{X})$

n	Var $(\hat{\mu})$	Var $(\bar{X})$
2	$\frac{\sigma_c^2}{2}(1+\rho)$	$\frac{\sigma_c^2}{2}(1+\rho)$
3	$\frac{\sigma_c^2}{3}\left(\frac{3-6\rho^2}{3-4\rho}\right)$	$\frac{\sigma_c^2}{3}\left(\frac{3+4\rho}{3}\right)$
4	$\frac{\sigma_c^2}{4}\left(\frac{2-6\rho^2+2\rho^4}{2-3\rho-\rho^2+\rho^3}\right)$	$\frac{\sigma_c^2}{4}\left(\frac{2+3\rho}{2}\right)$
5	$\frac{\sigma_c^2}{5}\left(\frac{5-20\rho^2+15\rho^4}{5-8\rho-6\rho^2+8\rho^3+\rho^4}\right)$	$\frac{\sigma_c^2}{5}\left(\frac{5+8\rho}{5}\right)$

In equations (4) and (22), impose the additional condition that $\underset{\sim}{U}$ is distributed as an n-variate normal. Then one can think of the sample elements as being generated from this linear model structure once Σ_n and μ are specified. However, for additional flexibility in investigating the specific nature of dependence among the observations in a sample, it is convenient to adopt the viewpoints and notation of autoregressive-moving average models, ARMA, as presented by Box and Jenkins [13] and Deutsch [18]. The mixed autoregressive-moving average model of order (p, q) is given by

$$\left.\begin{aligned} &\phi_p(B)\ \tilde{X}_i = \theta_q(B)\ a_i\ , \\ \text{where} \\ &a_i \sim \text{NID}(0,\ \sigma_a^{\ 2}). \end{aligned}\right\} \tag{25}$$

Since $\phi_p(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$, $\theta_q(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$, B is the backshift operator, and $\tilde{X}_i = X_i - \mu$, equation (25) can be rewritten as

$$\left.\begin{aligned} &\tilde{X}_i = \phi_1\ \tilde{X}_{i-1} + \ldots + \phi_p\ \tilde{X}_{i-p} + a_i - \theta_1\ a_{i-1} - \ldots - \theta_q\ a_{i-q}\ , \\ &a_i \sim \text{NID}\ (0,\ \sigma_a^{\ 2})\ . \end{aligned}\right\} \tag{26}$$

The model, given by equation (26), employs p + q + 2 parameters: $\phi_1, \ldots, \phi_p$, $\theta_1, \ldots, \theta_q$, $\sigma_a^{\ 2}$, and μ. The parameter μ is of primary interest to the quality control decision maker. Extensive investigation by Box and Jenkins have revealed that many physical processes can be adequately modeled when $p + q \leq 2$. In order to specifically show the one-to-one correspondence between different autocorrelative structures and ARMA models, let p = 0 and q = 1. In this instance, equation (26) reduces to

$$\left.\begin{aligned} &X_i = \mu + a_i - \theta_1\ a_{i-1}\ , \\ \text{where} \\ &a_i \sim \text{NID}\ (0,\ \sigma_a^{\ 2}) \end{aligned}\right\} \tag{27}$$

Such a model is called a first-order moving-average process, designated MA(1). It is easily shown that

$$\left.\begin{aligned} &E(X_i) = \mu,\ \mathrm{Var}\ (X_i) = (1 + \theta_1^{\,2})\ \sigma_a^{\,2}, \\ \text{and} \\ &\mathrm{Cov}(X_i, X_{i+k}) = -\theta_1\ \sigma_a^{\,2},\ k = 1 \\ &\qquad\qquad\qquad = 0 \qquad ,\ k > 1 \end{aligned}\right\} \tag{28}$$

If Σ_n denotes the $(n \times n)$ covariance matrix associated with the n sample elements generated from an MA(1) process, then Σ_n is the following type 2 diagonal matrix:

$$\Sigma_n = \sigma_a^2 \begin{bmatrix} (1+\theta_1^2) & -\theta_1 & 0 & \cdots & 0 \\ -\theta_1 & (1+\theta_1^2) & -\theta_1 & \cdots & 0 \\ 0 & -\theta_1 & (1+\theta_1^2) & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & (1+\theta_1^2) \end{bmatrix} \tag{29}$$

The covariance structure presented in equation (29) is identical with that presented in equation (11) for first-order serial correlation provided $\sigma_c^{\,2} = \sigma_a^{\,2}(1 + \theta_1^{\,2})$ and $\rho\sigma_c^{\,2} = -\theta_1\ \sigma_a^{\,2}$. Thus, if the observations have a first-order serial correlation with a specified ρ and $\sigma_c^{\,2}$, one can think of these observations as emanating from an MA(1) process with $\theta_1 = [-1 + (1 - 4\rho^2)^{1/2}]/2\rho$ and $\sigma_a^{\,2} = 2\rho^2\ \sigma_c^{\,2}/[1 + (1 - 4\rho^2)^{1/2}]$. Alternatively, if the observations are being generated from an MA(1) process with a known θ_1 and $\sigma_a^{\,2}$, this is equivalent to saying that they have a first-order serial correlation with $\sigma_c^2 = \sigma_a^2(1+\theta_1^2)$ and $\rho = -\theta_1/(1+\theta_1^2)$, where $\rho \varepsilon (-1/2, 1/2)$. Similar correspondence exists between other autocorrelative structures and ARMA (p,q) models. A more comprehensive coverage of ARMA models will be provided in Chapter III.

The relationship between the models for independent and dependent observations can now be stated. Recall that independent normal observations can be generated from model (1): $X_i = \mu + a_i$, $a_i \sim \text{NID}(0, \sigma_a^2)$, while first-order serially correlated observations can be generated from model (2): $X_i = \mu + a_i - \theta_1 a_{i-1}$, $a_i \sim \text{NID}(0, \sigma_a^2)$. Furthermore, let σ_I^2 denote Var (X_i) for model (1) and σ_c^2 denote Var (X_i) for model (2). Since $\sigma_I^2 = \sigma_a^2$ and $\sigma_c^2 = (1+\theta_1^2)\,\sigma_a^2$, it is obvious that $\sigma_c^2 \geq \sigma_I^2$ for constant σ_a^2 with $\sigma_I^2 = \sigma_c^2$ only when $\theta_1 = 0$ (which is equivalent to $\rho = 0$). However, it is possible to obtain both independent and correlated observations with $\sigma_I^2 = \sigma_c^2$. Let $\sigma_{a_I}^2$ denote the variance of the a_i for independent observations; let $\sigma_{a_c}^2$ denote the variance of the a_i for correlated observations. In order to have $\sigma_I^2 = \sigma_c^2$, we require $\sigma_{a_I}^2 = \sigma_{a_c}^2(1+\theta_1^2)$, or $\sigma_{a_I}^2/\sigma_{a_c}^2 = (1+\theta_1^2)$.

Before an example is presented, recall that for first-order serial correlation an $\bar{X}$-control chart would have as its limits $\mu_0 \pm z_{\alpha/2}$ $(\sigma_c/\sqrt{n})(1 + 2\rho[1 - n^{-1}])^{1/2}$. Also recall that for any type of autocorrelative structure the limits for a $\hat{\mu}$-control chart are given by $\mu_0 \pm z_{\alpha/2}$ $(\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n)^{-1/2}$. Furthermore, the limits for a $\hat{\mu}$-chart are always tighter than those of the correlated $\bar{X}$-chart. The relationship between these two control charts and the one developed for uncorrelated observations also needs to be explored. Recall that these limits were of the form $\mu_0 \pm z_{\alpha/2}$ $(\sigma_I/\sqrt{n})$, where μ_0 and σ_I may have been derived from a large amount of past data or past experiences with similar processes or selected to attain certain objectives. However, if the observations really have emanated from a first-order serially correlated process and the value of the process standard deviation has been obtained from watching such a process, then this process standard deviation is not σ_I. In fact, it is σ_c. Thus, if one assumes that there is no serial correlation when in fact there is, σ_I would be replaced by σ_c, and the control limits for $\bar{X}$ are actually given by $\mu_0 \pm z_{\alpha/2}(\sigma_c/\sqrt{n})$. Here, the correlated $\bar{X}$-chart will have tighter limits than those of the uncorrelated $\bar{X}$-chart when $\rho < 0$. To illustrate some of the previous comments, consider the following hypothetical example.

<u>Example 2.1</u>: From past experience with a process, it is determined that a first-order serial correlation exists between successive measurements with $\rho_0 = 0.47$ and that $\sigma_c^2 = 13.41$. This is equivalent to saying that the observations emanate from a first-order moving average process with $\theta_1 = -0.7$ and $\sigma_a = 3.0$. The nominal value of the process mean, μ_0, equals 30.0. Twenty samples, each of size 5, were generated from such a process and their values are given in Appendix A. To maintain control over the process mean, samples of size 5 will be taken every sampling interval, the $\hat{\mu}$-statistic will be calculated for each sample and plotted on a $\hat{\mu}$-control chart. The first step in constructing the $\hat{\mu}$-chart limits is to find the numerical entries in Λ_5 (see equation (24)):

$$\Lambda_5 = \begin{bmatrix} 0.1095 & -0.0743 & 0.0487 & -0.0294 & 0.0138 \\ -0.0743 & 0.1582 & -0.1037 & 0.0625 & -0.0294 \\ 0.0487 & -0.1037 & 0.1719 & -0.1037 & 0.0487 \\ -0.0294 & 0.0625 & -0.1037 & 0.1582 & -0.0743 \\ 0.0138 & -0.0294 & 0.0487 & -0.0743 & 0.1095 \end{bmatrix}$$

It immediately follows that $\underset{\sim}{j}_5^{\,t}\, \Lambda_5\, \underset{\sim}{j}_5 = 0.2251$ and $\sqrt{1/\underset{\sim}{j}_5^{\,t}\, \Lambda_5\, \underset{\sim}{j}_5} = \sqrt{4.4425} = 2.11$. If one chooses $\alpha = 0.0027$ (a traditional value), then $z_{\alpha/2} = 3.0$ and UCL $= \mu_0 + z_{\alpha/2} \sqrt{1/\underset{\sim}{j}_5^{\,t}\, \Lambda_5\, \underset{\sim}{j}_5} = 30.0 + (3.0)(2.11) = 36.33$ while LCL $= 23.67$. Instead of constructing a $\hat{\mu}$-chart, one could have set up a modified $\bar{X}$-chart with control limits defined in equation (13). In this instance, UCL $= \mu_0 + z_{\alpha/2}(\sigma_c/\sqrt{n})[1 + 2\rho_0(1 - n^{-1})]^{1/2} = 30 + (3.0)(3.66/\sqrt{5})[1 + 2(0.47)(4/5)]^{1/2} = 36.50$ while LCL $= 23.50$. If one was unaware of the presence of correlation, then the control chart limits for the traditional $\bar{X}$-chart are given by UCL $= \mu_0 + z_{\alpha/2}(\sigma_c/\sqrt{n}) = 30 + (3.0)(3.66/\sqrt{5}) = 34.91$ and LCL $= 25.09$. For each of the twenty samples, $\hat{\mu}_i$ and $\bar{x}_i$ were calculated (see Appendix A) and plotted on a control chart using all three sets of control limits. This is illustrated in Figure 3.

Inspection of Figure 3 reveals that the traditional $\bar{X}$-chart limits are tighter than those of the modified $\bar{X}$-chart and the $\hat{\mu}$-chart. The significance of this is demonstrated with sample No. 3 where $\bar{x}_3$ plots above the traditional $\bar{X}$-chart limits. Thus, if one were unaware of the presence of first-order correlation and used a traditional $\bar{X}$-control chart, one would search for nonexistent assignable causes more frequently than necessary. This is the case with $\bar{x}_3$.

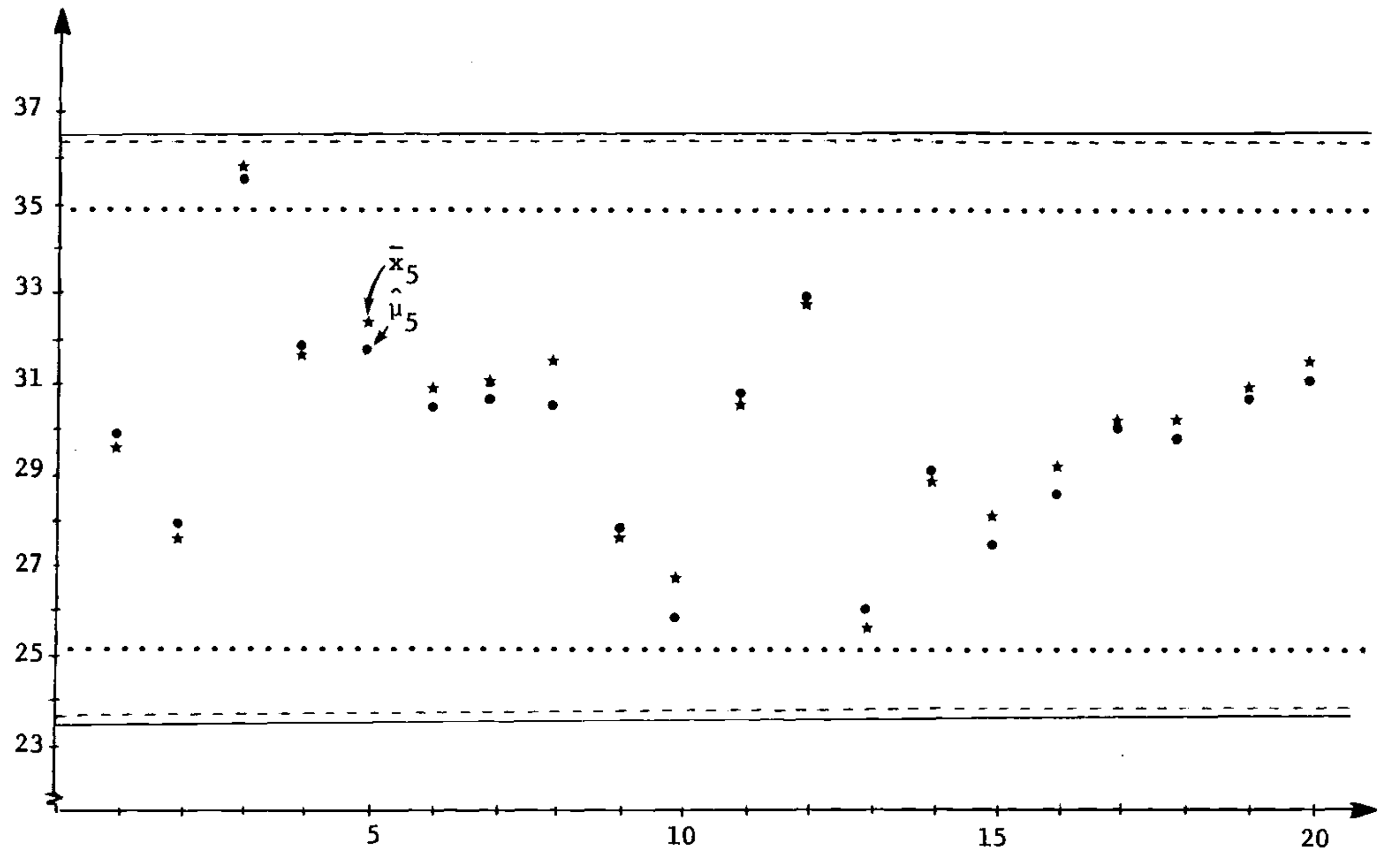

Figure 3. A Univariate $\hat{\mu}$ Control Chart with $\hat{\mu}$ Limits Designated by ---, Modified $\bar{X}$-Limits Designated by ___, and Traditional $\bar{X}$-Limits Designated by ... .

In this particular example, there is very little difference between the limits for the modified $\bar{X}$-chart and those of the $\hat{\mu}$-chart. However, this is not always the case. For example, if $\rho_0 = -0.47$ while $\sigma_c^2 = 13.41$, then the $\hat{\mu}$ limits are given by [27.82, 32.19], the modified $\bar{X}$ limits are given by [25.27, 34.73], and the traditional $\bar{X}$ limits are given by [25.09, 34.91], which are the same as before. In this case, the $\hat{\mu}$-chart limits are much tighter than those of either $\bar{X}$-chart. And, if the manufacturer were to use either $\bar{X}$-chart limits, it is the consumer who would suffer for an out-of-control process has less chance of being detected.

The control limits presented in equation (23) are valid for any type of correlative structure. For example, if the underlying process is AR(1) which is described by

$$\tilde{X}_i = \phi_1 \tilde{X}_{i-1} + a_i, \tag{30}$$

then

$$\left.\begin{aligned} &E(X_i) = \mu,\ \mathrm{Var}(X_i) = \sigma_a^2/(1 - \phi_1^2)\ , \\ &\mathrm{Cov}(X_i,\ X_{i+k}) = \sigma_a^2\ \phi_1^k/(1 - \phi_1^2),\ k \geq 1\ , \end{aligned}\right\} \tag{31}$$

and from equation (31) one can construct Σ_n, and Λ_n for the appropriate n.

The $(n \times n)$ covariance matric associated with equation (31) is a Toeplitz matrix while its inverse is a Jacobi matrix. The general forms of these matrices are given in Press [49] and Ray [51]. If one is calculating $\hat{\mu}$ for repeated samples, caution must be exercised in choosing the sampling intervals due to the nature of an AR(1) process. Equation (31) reveals that the covariance between observations decreases "exponentially," where this decrease is fairly rapid for small values of $|\phi_1|$. In this situation one would be fairly safe in disregarding the correlation between the X_i's belonging to adjoining sampling intervals. However, when $|\phi_1|$ is relatively large, the X_i's of adjoining intervals could be correlated and induce correlation between successive $\hat{\mu}$'s. Recall that $|\phi_1| < 1$ is required for stationarity.

In the univariate case, one does not have to adhere to the traditional format of the Shewhart chart by using the control limits and center line presented in equation (23). Since $\hat{\mu} \sim N(\mu_0, 1/\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n)$ when the process mean equals the nominal value μ_0, we see that

$$\left(\frac{\hat{\mu} - \mu_o}{1/\underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n} \right)^2 \sim \chi_1^2. \tag{32}$$

The control chart would now appear as in Figure 4, where $\chi^2_{1,\alpha}$ is such that $P(\chi_1^2 > \chi^2_{1,\alpha}) = \alpha$. The statistic plotted on the chart is $[(\hat{\mu} - \mu_0) \underset{\sim}{j}_n{}^t \Lambda_n \underset{\sim}{j}_n]^2$. Similar results will be obtained for the multivariate problem. One disadvantage in using such a chart is that runs above and below μ_0 can no longer be detected in the χ^2-chart.

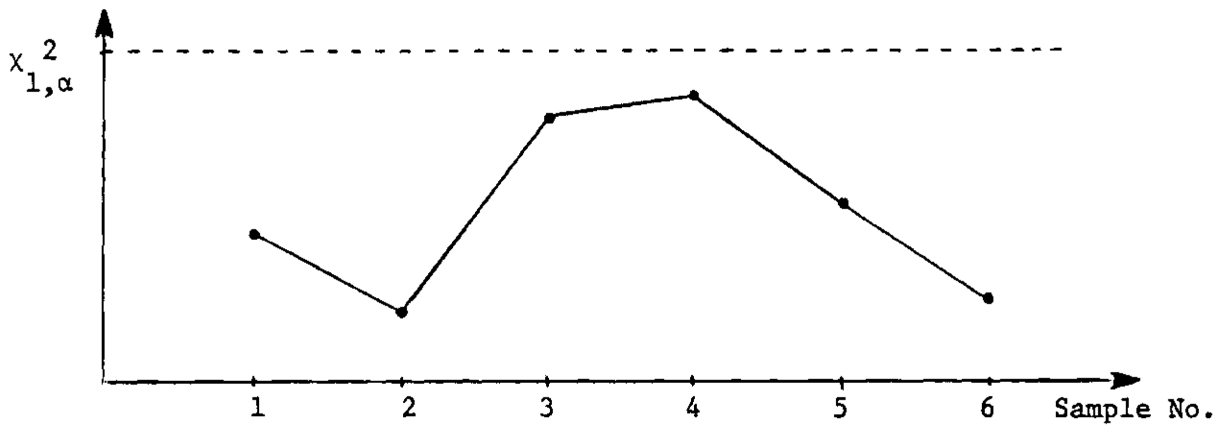

Figure 4. A χ^2 Control Chart for One Quality Characteristic

2.2 Multiple Quality Characteristics

2.2.1 Independent Observations

The general multivariate statistical quality control problem considers a repetitive process where each item is characterized by p quality characteristics, $X_1, X_2, \ldots, X_p$, which are random variables because of the chance causes inherent in the process. The probability law associated with $\underset{\sim}{X}^t = (X_1, X_2, \ldots, X_p)$ will be denoted by $f_{\underset{\sim}{X}^t}(\underset{\sim}{x}^t; \underset{\sim}{\mu}^t)$ where the (p x 1) population mean vector, denoted by $\underset{\sim}{\mu}$, is defined to be

$$\underset{\sim}{\mu}^t = E(\underset{\sim}{X}^t) = [E(X_1), \ldots, E(X_p)] = [\mu_1, \ldots, \mu_p] \tag{33}$$

and the (p x p) covariance matrix of $\underset{\sim}{X}$, denoted by Σ, is defined to be

$$\Sigma = \begin{bmatrix} V(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_p) \\ \vdots & \vdots & & \vdots \\ Cov(X_1, X_p) & Cov(X_2, X_p) & \cdots & V(X_p) \end{bmatrix} \tag{34}$$

When the scenario is a repetitive manufacturing operation, multiple measurements will be made on a sample of the successively manufactured items and it is desired that these multiple measurements behave as though they were obtained from a population having $f_{\underset{\sim}{X}}{}^t(x^t; \underset{\sim}{\mu}^t)$ as its probability distribution. In this section, interest is centered on the population mean vector $\underset{\sim}{\mu}$. When changes in the process cause the elements of $\underset{\sim}{\mu}$ to shift from their nominal values, denoted by $\underset{\sim}{\mu}_0$, it becomes necessary to detect these changes to insure a uniform quality product. Previous research by Alt [2] and Alt, Goode, and Wadsworth [6] have treated various aspects of this problem. Implicit in their work are the assumptions that

(i) the behavior of $\underset{\sim}{X}$ is adequately described by a p-variate normal distribution, namely,

$$f_{\underset{\sim}{X}^t}(\underset{\sim}{x}^t;\ \underset{\sim}{\mu}^t) = (2\pi)^{-p/2}|\Sigma|^{-1/2}\ \exp\{-(1/2)(\underset{\sim}{x} - \underset{\sim}{\mu})^t\ \Sigma^{-1}(\underset{\sim}{x} - \underset{\sim}{\mu})\}\ , \tag{35}$$

and

(ii) the sampled elements $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ behave as a random sample, that is,

$$f_{\underset{\sim}{X}_1^t,\ \underset{\sim}{X}_2^t, \ldots,\ \underset{\sim}{X}_n^t}(\underset{\sim}{x}_1^t, \ldots, \underset{\sim}{x}_n^t;\ \underset{\sim}{\mu}^t) = \prod_{i=1}^{n} f_{\underset{\sim}{X}^t}(\underset{\sim}{x}_i^t;\ \underset{\sim}{\mu}^t) \tag{36}$$

Under the above assumptions with Σ known, it is easily shown that the maximum likelihood estimator of $\underset{\sim}{\mu}$ is given by $\bar{\underset{\sim}{X}}$, the vector of sample means, where

$$\bar{\underset{\sim}{X}} \sim N_p(\underset{\sim}{\mu},\ \Sigma/n). \tag{37}$$

Details can be found in Press [49] and Anderson [9]. Although $\bar{\underset{\sim}{X}}$ has many desirable statistical properties, only its efficiency will be investigated here by using the multivariate version of the Cramér-Rao inequality.

For a fixed positive integer n, let $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ denote a sample of size n from a distribution that is one member of the family $\{f_{\underset{\sim}{X}^t}(\underset{\sim}{x}^t;\ \underset{\sim}{\theta}^t) : \underset{\sim}{\theta} \in \Omega\}$ where $\underset{\sim}{X}$ is (p x 1), $\underset{\sim}{\theta}$ is (r x 1), and Ω denotes the parameter space. Assume $f_{\underset{\sim}{X}^t}(\underset{\sim}{x}^t;\underset{\sim}{\theta}^t)$ satisfies certain regularity conditions. Let the (r x 1) vector $\underset{\sim}{d}$ be an unbiased estimator of $\underset{\sim}{\theta}$; and let the (r x 1) vector $\underset{\sim}{W}$ have as its components

$$W_i = \partial(\ell n\ f_{\underset{\sim}{X}_1^t, \ldots,\ \underset{\sim}{X}_n^t}(\underset{\sim}{X}_1^t, \ldots,\ \underset{\sim}{X}_n^t;\ \underset{\sim}{\theta}^t))/\partial\theta_i\ ,$$

for $i = 1, 2, \ldots, r$. Furthermore, let $I_{\underset{\sim}{\theta}}$ be the $(r \times r)$ matrix with $(i, j)^{th}$ entry given by $E(W_i W_j) = -E(\partial W_i / \partial \theta_j)$. If $\Sigma_{\underset{\sim}{d}}$ denotes the covariance matrix of $\underset{\sim}{d}$, then the generalized Cramér-Rao inequality states that $\Sigma_{\underset{\sim}{d}} - I_{\underset{\sim}{\theta}}^{-1}$ is positive semi-definite. A proof is given in Silvey [61]. For the specific problem at hand, $\underset{\sim}{\theta}$ is the $(p \times 1)$ vector $\underset{\sim}{\mu}$ since Σ is known, and $f_{\underset{\sim}{X}}{}^t(\underset{\sim}{x}^t; \underset{\sim}{\theta}^t)$ is the p-variate normal density which satisfies the regularity conditions. From equations (35) and (36), it follows that $\underset{\sim}{W} = n\, \Sigma^{-1} (\bar{\underset{\sim}{X}} - \underset{\sim}{\mu})$ and $\partial \underset{\sim}{W} / \partial \underset{\sim}{\mu}^t = -n\, \Sigma^{-1}$. Thus $I_{\underset{\sim}{\theta}} = -E(-n\, \Sigma^{-1}) = n\, \Sigma^{-1}$, and $I_{\underset{\sim}{\theta}}^{-1} = \Sigma/n$ is a "lower bound" for the variance-covariance matrix of an unbiased estimator of $\underset{\sim}{\mu}$. Since $\Sigma_{\bar{\underset{\sim}{X}}} = \Sigma/n$, $\Sigma_{\bar{\underset{\sim}{X}}} - I_{\underset{\sim}{\theta}}^{-1} = 0$ and the "lower bound" is attained in this case. For $p = 1$, I_{θ}^{-1} reduces to the well-known lower bound result of σ_I^2/n stated in Section 2.1.1. Since $\bar{\underset{\sim}{X}}$ is an unbiased sufficient statistic for $\underset{\sim}{\mu}$ and since its distribution is complete, the UMVUE property of $\bar{X}$ could have also been determined from the Lehmann-Scheffé Theorem.

Additional rationale for using $\bar{\underset{\sim}{X}}$ as the estimator for $\underset{\sim}{\mu}$ is provided by thinking of the sample elements $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ as being generated from the following linear model:

$$\underbrace{\begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & & \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{A} \underbrace{[\mu_1, \mu_2, \ldots, \mu_p]}_{B} + \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}}_{U} \quad (38)$$

where the $(n \times p)$ disturbance matrix U has the zero matrix as its expectation and the common variance-covariance matrix Σ within any row of U with zero covariance between rows of U. Thus, if $\underset{\sim}{X}^t = [\underset{\sim}{X}_1{}^t, \underset{\sim}{X}_2{}^t, \ldots, \underset{\sim}{X}_n{}^t]$, then

$$E(X) = AB,$$

and

$$C(\underset{\sim}{X}, \underset{\sim}{X}^t) = \begin{bmatrix} \Sigma & 0 & \ldots & 0 \\ 0 & \Sigma & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \Sigma \end{bmatrix} = I_n \otimes \Sigma \ . \tag{39}$$

The conditions stated in equations (38) and (39) specify what is known as the multivariate classical linear regression model. Goldberger [29] has shown that the least-squares estimator of B, found by minimizing the trace of $(X - AB)^t(X - AB)$, is given by $(A^t\ A)^{-1}\ A^t\ X$; which reduces to $\bar{\underset{\sim}{X}}^t$ for the A and X matrices stated in equation (38). Furthermore, the variance-covariance matrix of this estimator is given by $\Sigma \otimes (A^t\ A)^{-1}$, which reduces to Σ/n. Thus, even in the multivariate case, the least-squares estimator, which is the best linear unbiased estimator of $\underset{\sim}{\mu}$, is identical with the maximum likelihood estimator. Now that justification has been given for estimating $\underset{\sim}{\mu}$ by $\bar{\underset{\sim}{X}}$, we go on to study the important problem of testing whether the process mean has shifted from the nominal value $\underset{\sim}{\mu}_0$ and how this relates to $\bar{\underset{\sim}{X}}$.

Suppose $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ is a random sample of size n from a p-variate normal process with mean vector $\underset{\sim}{\mu}$ and known variance-covariance matrix Σ. The likelihood ratio test of $H_0: \underset{\sim}{\mu} = \underset{\sim}{\mu}_0$ vs. $H_1: \underset{\sim}{\mu} \neq \underset{\sim}{\mu}_0$ yields the following critical region

$$\omega = \{\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_n : n(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0)^t \, \Sigma^{-1}(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0) > \chi^2_{p,\alpha} \}, \tag{40}$$

when the null hypothesis is true. Thus, if the hypothesis testing viewpoint of one quality characteristic is generalized to multiple quality characteristics, the quality control engineer investigates statistical control of $\underset{\sim}{\mu}_0$ by taking a random sample of size n; computing $\bar{\underset{\sim}{x}}$ and determining whether

$$n(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0)^t \, \Sigma^{-1}(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0) > \chi^2_{p,\alpha} \quad . \tag{41}$$

If the inequality in (41) holds, then the decision maker would conclude that $\underset{\sim}{\mu}$ has shifted from $\underset{\sim}{\mu}_0$ and assignable causes would be sought. For successive samples of size n, the decision making process can be set up as a control chart similar in appearance to Figure 4 with $\chi^2_{1,\alpha}$ replaced by $\chi^2_{p,\alpha}$. On this chart, one plots the scalar quantities $n(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0)^t$ $\Sigma^{-1}(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0)$ for the successive samples and maintains control over $\underset{\sim}{\mu}_0$ by inspecting this χ^2-chart. Note that there is only an upper control limit, namely, UCL = $\chi^2_{p,\alpha}$, since the test statistic is a generalized measure of distance.

The decision rule presented in (41) can also be developed from a more intuitively appealing viewpoint, as presented in Anderson [9] and Press [49]. If the true process mean is $\underset{\sim}{\mu}$ while the nominal value is $\underset{\sim}{\mu}_0$, it is of interest to study how much $\underset{\sim}{\mu}$ deviates from $\underset{\sim}{\mu}_0$ or, equivalently, how much $\underset{\sim}{\mu} - \underset{\sim}{\mu}_0$ deviates from the zero vector. Since $\bar{\underset{\sim}{X}}$ is eminently qualified as an estimator for $\underset{\sim}{\mu}$, it seems reasonable to measure $\underset{\sim}{\mu} - \underset{\sim}{\mu}_0$ by using $\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}_0$, where $(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}_0) \sim N_p(\underset{\sim}{0}, \Sigma/n)$ when $\underset{\sim}{\mu} = \underset{\sim}{\mu}_0$. Since the deviations of each component of $\bar{\underset{\sim}{X}}$ from those of $\underset{\sim}{\mu}_0$ may be positive or negative and these deviations have differing variability, it is necessary to square these deviations and weight them by the reciprocal of their spread, which results in using the statistic $n(\underset{\sim}{\bar{X}} - \underset{\sim}{\mu}_0)^t \, \Sigma^{-1}(\underset{\sim}{\bar{X}} - \underset{\sim}{\mu}_0)$. It is easily shown that this statistic has a χ_p^2 distribution when the null hypothesis is true. Thus, this intuitive approach results in the same rule as that produced by the likelihood ratio test. This is not surprising since the intuitive approach is based on the sufficient statistic $\underset{\sim}{\bar{X}}$ and the maximum likelihood estimator is a function of this sufficient statistic, namely, the identity function. Furthermore, for testing $H_0:\underset{\sim}{\mu} = \underset{\sim}{\mu}_0$ vs. $H_1:\underset{\sim}{\mu} \neq \underset{\sim}{\mu}_0$, the likelihood ratio test is a function of every sufficient statistic for $\underset{\sim}{\mu}$, and hence of $\underset{\sim}{\bar{x}}$.

If a statistic does plot out of control on the χ^2-chart, the individual components of $\underset{\sim}{\mu}$ responsible for this need to be determined. One solution to this problem is obtained by using Sidak's inequality [60]: Let $\underset{\sim}{U}$ be distributed as a p-variate normal with $E(\underset{\sim}{U}) = \underset{\sim}{0}$, arbitrary variances and arbitrary correlations. Then, for any positive numbers $c_1, c_2, \ldots, c_p$,

$$P(|U_1| \leq c_1, \ldots, |U_p| \leq c_p) \geq \prod_{i=1}^{p} P(|U_i| \leq c_i). \tag{42}$$

For the specific problem at hand, $\underset{\sim}{\bar{X}}$ is distributed as a p-variate normal with unknown mean values $\mu_1, \ldots, \mu_p$ and known variances σ_i^2/n. To use Sidak's inequality, let $U_i = n^{1/2}(\bar{X}_i - \mu_i)/\sigma_i$, $i = 1, 2, \ldots, p$ and let $c_1 = c_2 = \ldots = c_p$ be such that $2\,\Phi(c_i) = 1 + (1 - \alpha)^{1/p}$. Then a rectangular confidence region for $\mu_1, \ldots, \mu_p$ with bounded confidence level $1 - \alpha$ is obtained by using the following individual confidence intervals:

$$[\bar{x}_i - c_i\ \sigma_i/\sqrt{n},\ \bar{x}_i + c_i\ \sigma_i/\sqrt{n}]\,. \tag{43}$$

As stated by Sidak ". . . we may always act as if all coordinates . . . were independent." The intervals generated by Sidak's inequality are shorter than those obtained using either the Bonferroni or Scheffé technique. Additional explanation of these latter two techniques and other aspects of the χ^2-chart, such as its power, can be found in Alt [2].

In this section, the development of the decision rule for maintaining control over $\underset{\sim}{\mu}_0$ was based on the assumptions of process normality and independence of the sample elements. Let us now determine how departures from this latter assumption affect the decision rule.

2.2.2 Dependent Observations

This section is concerned with the development of control charts for the mean vector when the sample elements are correlated and the quality of each item is determined by several, correlated characteristics. Thus, there is correlation across the sample elements as well as within each sample element. These statements can be formalized as follows.

Let the n sample elements be denoted by $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ where each $\underset{\sim}{X}_i$ is a $(p \times 1)$ vector. Let $\underset{\sim}{X}$ denote the $(np \times 1)$ vector of sample elements, where $\underset{\sim}{X}^t = [\underset{\sim}{X}_1{}^t, \underset{\sim}{X}_2{}^t, \ldots, \underset{\sim}{X}_n{}^t] = [X_{11}, \ldots, X_{p1}, X_{12}, \ldots X_{p2}, \ldots, X_{1n}, \ldots, X_{pn}]$. Let $\underset{\sim}{\mu}_{\underset{\sim}{X}}$ denote $E(\underset{\sim}{X})$, where this $(np \times 1)$ vector is given by

$$\underset{\sim}{\mu}_{\underset{\sim}{X}}{}^t = (\underset{\sim}{\mu}^t,\ \underset{\sim}{\mu}^t, \ldots, \underset{\sim}{\mu}^t)\,, \tag{44}$$

with $\underset{\sim}{\mu}^t = (\mu_1, \ldots, \mu_p)$ being the population mean vector of each $\underset{\sim}{X}_i$. If we let $A \otimes B$ denote the direct product (see Graybill [32]) of the

$(m_1 \times n_1)$ matrix A with the $(m_2 \times n_2)$ matrix B, then $A \otimes B$ is an $(m_1m_2 \times n_1n_2)$ matrix with entries (Ba_{ij}). Thus, equation (44) can be written as

$$\underset{\sim}{\mu}_{\underset{\sim}{X}} = (\underset{\sim}{j}_n \otimes I_p)\, \underset{\sim}{\mu}, \tag{45}$$

where I_p is the $(p \times p)$ identity matrix. Let $\Sigma_{\underset{\sim}{X}}$ denote the $(np \times np)$ covariance matrix of $\underset{\sim}{X}$. That is, $\Sigma_{\underset{\sim}{X}} = C(\underset{\sim}{X}, \underset{\sim}{X}^t) = E(\underset{\sim}{X} - \underset{\sim}{\mu}_{\underset{\sim}{X}})(\underset{\sim}{X} - \underset{\sim}{\mu}_{\underset{\sim}{X}})^t$. $\Sigma_{\underset{\sim}{X}}$ may be partitioned as follows:

$$\Sigma_{\underset{\sim}{X}} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \Sigma_{n1} & \Sigma_{n2} & \cdots & \Sigma_{nn} \end{bmatrix}, \tag{46}$$

where the $(p \times p)$ submatrix $\Sigma_{ij} = C(\underset{\sim}{X}_i, \underset{\sim}{X}_j^t) = E(\underset{\sim}{X}_i - \underset{\sim}{\mu})(\underset{\sim}{X}_j - \underset{\sim}{\mu})^t$. Now Σ_{ii} is a $(p \times p)$ symmetric matrix. However, in general, Σ_{ij} may not be a symmetric matrix and thus $\Sigma_{ij} \neq \Sigma_{ji}$. Note that $\Sigma_{ij}{}^t = \Sigma_{ji}$. Furthermore, $\Sigma_{\underset{\sim}{X}}$ is a symmetric, positive definite matrix. The maximum likelihood estimator of $\underset{\sim}{\mu}$ will now be found.

Theorem 2.2: Let $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ be jointly normal with mean vector $\underset{\sim}{\mu}_{\underset{\sim}{X}}$ and covariance matrix $\Sigma_{\underset{\sim}{X}}$ as given in equations (45) and (46), respectively. Then the maximum likelihood estimator of $\underset{\sim}{\mu}$, denoted by $\hat{\underset{\sim}{\mu}}$, is given by

$$\hat{\underset{\sim}{\mu}} = (B^t \Lambda_{\underset{\sim}{X}} B)^{-1} B^t \Lambda_{\underset{\sim}{X}} \underset{\sim}{X}, \tag{47}$$

where the $(np \times p)$ matrix B is defined to be

$$B = (\underset{\sim}{j}_n \otimes I_p) , \tag{48}$$

and $\underset{\sim}{\Lambda}_X = \underset{\sim}{\Sigma}_X^{-1}$

Proof: The likelihood function, denoted by $L(\underset{\sim}{\mu})$, is given by

$$L(\underset{\sim}{\mu}) = (2\pi)^{-(np/2)} \; |\underset{\sim}{\Lambda}_X|^{1/2} \exp\{-(1/2)\; Q(\underset{\sim}{\mu})\} , \tag{49}$$

where

$$Q(\underset{\sim}{\mu}) = [\underset{\sim}{x} - (\underset{\sim}{j}_n \otimes I_p)\underset{\sim}{\mu}]^t \; \underset{\sim}{\Lambda}_X[\underset{\sim}{x} - (\underset{\sim}{j}_n \otimes I_p)\; \underset{\sim}{\mu}] . \tag{50}$$

Now $Q(\underset{\sim}{\mu})$ can be expanded to yield the following:

$$Q(\underset{\sim}{\mu}) = \underset{\sim}{x}^t \; \underset{\sim}{\Lambda}_X \; \underset{\sim}{x} - 2\underset{\sim}{x}^t \; \underset{\sim}{\Lambda}_X(\underset{\sim}{j}_n \otimes I_p) \; \underset{\sim}{\mu} + \underset{\sim}{\mu}^t(\underset{\sim}{j}_n \otimes I_p)^t \; \underset{\sim}{\Lambda}_X(\underset{\sim}{j}_n \otimes I_p) \; \underset{\sim}{\mu} .$$

Recall that $\partial(A\underset{\sim}{\mu})/\partial\underset{\sim}{\mu} = A^t$ and $\partial(\underset{\sim}{\mu}^t \; A\underset{\sim}{\mu})/\partial\underset{\sim}{\mu} = 2 \; A\underset{\sim}{\mu}$. Thus,

$$\partial Q(\underset{\sim}{\mu})/\partial\underset{\sim}{\mu} = -2\underset{\sim}{x}^t \; \underset{\sim}{\Lambda}_x \; B + 2 \; B^t \; \underset{\sim}{\Lambda}_X \; B \; \underset{\sim}{\mu} .$$

Setting $\partial Q(\underset{\sim}{\mu})/\partial\underset{\sim}{\mu} = \underset{\sim}{0}$ yields equation (47). ||

As with the univariate case (Section 2.1.2), the result presented in equation (47) could have also been obtained using a geometrical approach. Since $\underset{\sim}{\Lambda}_X$ is a symmetric matrix, there exists an orthogonal matrix P such that $P^t \; \underset{\sim}{\Lambda}_X \; P$ is a diagonal matrix, $D(\lambda_i)$, whose elements are the eigenvalues of $\underset{\sim}{\Lambda}_X$. Furthermore, since $\underset{\sim}{\Lambda}_X$ is positive definite, every $\lambda_i > 0$. And, $P^t \; \underset{\sim}{\Lambda}_X \; P$ can be rewritten as $P^t \; \underset{\sim}{\Lambda}_X \; P = D(\sqrt{\lambda_i}) \; D(\sqrt{\lambda_i})$. Thus, by letting $\underset{\sim}{\Lambda}_X^{1/2} = PD(\sqrt{\lambda_i}) \; P^t$, we see that $\underset{\sim}{\Lambda}_X = \underset{\sim}{\Lambda}_X^{1/2} \; \underset{\sim}{\Lambda}_X^{1/2}$ where $\underset{\sim}{\Lambda}_X^{1/2}$ is a nonsingular symmetric matrix. This allows us to rewrite $Q(\underset{\sim}{\mu})$ as

$$Q(\underset{\sim}{\mu}) = (\underset{\sim}{x} - B\underset{\sim}{\mu})^t \, \underset{\sim}{\Lambda_X}^{1/2} \, \underset{\sim}{\Lambda_X}^{1/2} (\underset{\sim}{x} - B\underset{\sim}{\mu})$$

$$= (\underset{\sim}{\Lambda_X}^{1/2} \underset{\sim}{x} - \underset{\sim}{\Lambda_X}^{1/2} B\underset{\sim}{\mu})^t \, (\underset{\sim}{\Lambda_X}^{1/2} - \underset{\sim}{\Lambda_X}^{1/2} B\underset{\sim}{\mu})$$

$$= (\underset{\sim}{x}' - B'\underset{\sim}{\mu})^t \, (\underset{\sim}{x}' - B'\underset{\sim}{\mu}) = \langle \underset{\sim}{x}' - B'\underset{\sim}{\mu}, \underset{\sim}{x}' - B'\underset{\sim}{\mu} \rangle$$

$$= ||\underset{\sim}{x}' - B'\underset{\sim}{\mu}||^2,$$

where $\underset{\sim}{x}' = \underset{\sim}{\Lambda_X}^{1/2} \underset{\sim}{x}$ and $B' = \underset{\sim}{\Lambda_X}^{1/2} B$. The least squares problem is to find $\underset{\sim}{\mu}$ to minimize this or to find $\hat{\underset{\sim}{\mu}}$ such that $B'\hat{\underset{\sim}{\mu}}$ is the projection of $\underset{\sim}{x}'$ on the range of the linear transformation B'. This is equivalent to solving the normal equations

$$(B')^t \, B'\hat{\underset{\sim}{\mu}} = (B')^t \, \underset{\sim}{x}'$$

If $(B')^t \, B'$ is invertible, then

$$\hat{\underset{\sim}{\mu}} = [(B')^t \, B']^{-1} \, (B')^t \, \underset{\sim}{x}' .$$

Since $x' = \underset{\sim}{\Lambda_X}^{1/2} \underset{\sim}{x}$ and $B' = \underset{\sim}{\Lambda_X}^{1/2} B$, this reduces to

$$\hat{\underset{\sim}{\mu}} = [(\underset{\sim}{\Lambda_X}^{1/2} B)^t \, (\underset{\sim}{\Lambda_X}^{1/2} B)]^{-1} \, (\underset{\sim}{\Lambda_X}^{1/2} B)^t \, (\underset{\sim}{\Lambda_X}^{1/2} \underset{\sim}{x}) = (B^t \underset{\sim}{\Lambda_X} B)^{-1} \, B^t \, \underset{\sim}{\Lambda_X} \, \underset{\sim}{x} .$$

To gain further insight into equation (47), partition the (np x np) matrix $\underset{\sim}{\Lambda_X}$ as follows:

$$\underset{\sim}{\Lambda_X} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \cdots & \Lambda_{1n} \\ \Lambda_{21} & \Lambda_{22} & \cdots & \Lambda_{2n} \\ \vdots & \vdots & & \vdots \\ \Lambda_{n1} & \Lambda_{n2} & \cdots & \Lambda_{nn} \end{bmatrix},$$

where Λ_{ij} is a (p x p) matrix. Then $(B^t \underset{\sim}{\Lambda_X} B) = \sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ij}$, the sum of the n^2 submatrices in $\underset{\sim}{\Lambda_X}$. Furthermore, $B^t \underset{\sim}{\Lambda_X} \underset{\sim}{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ji} \underset{\sim}{x}_i$ and the value of each sample element is weighted by the sum of the (p x p) matrices in the i^{th} "column" of $\underset{\sim}{\Lambda_X}$. Thus, $\hat{\underset{\sim}{\mu}}$ can be written as

$$\hat{\underset{\sim}{\mu}} = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ij}\right)^{-1} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ji} \underset{\sim}{x}_i\right) \tag{51}$$

From equation (47), we see that $\hat{\underset{\sim}{\mu}} = A \underset{\sim}{x}$, where $A = (B^t \underset{\sim}{\Lambda_X} B)^{-1} B^t \underset{\sim}{\Lambda_X}$ has dimension (p x np). Hence $\hat{\underset{\sim}{\mu}}$ is distributed as a p-variate normal (see Rao [50]). The expected value vector and variance-covariance matrix of $\hat{\underset{\sim}{\mu}}$ are obtained as follows:

$$E(\hat{\underset{\sim}{\mu}}) = (B^t \underset{\sim}{\Lambda_X} B)^{-1} B^t \underset{\sim}{\Lambda_X}(E\underset{\sim}{X}) = (B^t \underset{\sim}{\Lambda_X} B)^{-1} B^t \underset{\sim}{\Lambda_X} B \underset{\sim}{\mu} = \underset{\sim}{\mu}, \tag{52}$$

and

$$\Sigma_{\hat{\underset{\sim}{\mu}}} = C(\hat{\underset{\sim}{\mu}}, \hat{\underset{\sim}{\mu}}^t) = E[(B^t \underset{\sim}{\Lambda_X} B)^{-1} B^t \underset{\sim}{\Lambda_X} \underset{\sim}{X} - \underset{\sim}{\mu}][(B^t \underset{\sim}{\Lambda_X} B)^{-1} B^t \underset{\sim}{\Lambda_X} \underset{\sim}{X} - \underset{\sim}{\mu}]^t$$

$$= E(B^t \underset{\sim}{\Lambda_X} B)^{-1}[B^t \underset{\sim}{\Lambda_X} \underset{\sim}{X} - B^t \underset{\sim}{\Lambda_X} B \underset{\sim}{\mu}][B^t \underset{\sim}{\Lambda_X} \underset{\sim}{X} - (B^t \underset{\sim}{\Lambda_X} B) \underset{\sim}{\mu}]^t (B^t \underset{\sim}{\Lambda_X} B)^{-1}$$

$$= (B^t \Lambda_{\underset{\sim}{X}} B)^{-1} \{B^t \Lambda_{\underset{\sim}{X}}[E(\underset{\sim}{X} - B \underset{\sim}{\mu})(\underset{\sim}{X} - B \underset{\sim}{\mu})^t] \Lambda_{\underset{\sim}{X}} B\} (B^t \Lambda_{\underset{\sim}{X}} B)^{-1}$$

$$= (B^t \Lambda_{\underset{\sim}{X}} B)^{-1} B^t \Lambda_{\underset{\sim}{X}} \Sigma_{\underset{\sim}{X}} \Lambda_{\underset{\sim}{X}} B(B^t \Lambda_{\underset{\sim}{X}} B)^{-1}$$

$$= (B^t \Lambda_{\underset{\sim}{X}} B)^{-1}$$

$$= [(\underset{\sim}{j}_n \otimes I_p)^t \Lambda_{\underset{\sim}{X}}(\underset{\sim}{j}_n \otimes I_p)]^{-1} \tag{53}$$

Equation (52) demonstrates that $\hat{\underset{\sim}{\mu}}$ is an unbiased estimator of $\underset{\sim}{\mu}$. The above three properties of $\hat{\underset{\sim}{\mu}}$ can be combined by stating that

$$\hat{\underset{\sim}{\mu}} \sim N_p(\underset{\sim}{\mu}, [(\underset{\sim}{j}_n \otimes I_p)^t \Lambda_{\underset{\sim}{X}}(\underset{\sim}{j}_n \otimes I_p)]^{-1}) . \tag{54}$$

The results presented in equation (54) reduce to those presented earlier in equation (21) for the univariate case. The efficiency of $\hat{\underset{\sim}{\mu}}$ will now be determined using the Cramér-Rao inequality, presented in Section 2.2.1.

From equations (49) and (50), it follows that $\underset{\sim}{W} = B^t \Lambda_{\underset{\sim}{X}} \underset{\sim}{X} - B^t \Lambda_{\underset{\sim}{X}} B \underset{\sim}{\mu}$ and $\partial W / \partial \underset{\sim}{\mu}^t = - B^t \Lambda_{\underset{\sim}{X}} B$. Thus $\underset{\sim}{I}_\theta = -E(-B^t \Lambda_{\underset{\sim}{X}} B) = B^t \Lambda_{\underset{\sim}{X}} B$, while $\underset{\sim}{I}_\theta^{-1} = (B^t \Lambda_{\underset{\sim}{X}} B)^{-1}$ is a "lower bound" for the variance-covariance matrix of an unbiased estimator of $\underset{\sim}{\mu}$. Since $\Sigma_{\hat{\underset{\sim}{\mu}}} = (B^t \Lambda_{\underset{\sim}{X}} B)^{-1}$, $\Sigma_{\hat{\underset{\sim}{\mu}}} - \underset{\sim}{I}_\theta^{-1} = 0$, and the "lower bound" is attained in this case. That $\hat{\underset{\sim}{\mu}}$ is a UMVUE can have just as well been shown using the Lehmann-Scheffé Theorem. Since the unbiasedness of $\hat{\underset{\sim}{\mu}}$ was shown in equation (52), the sufficiency property of $\hat{\underset{\sim}{\mu}}$ needs to be investigated next. Recall that $B \hat{\underset{\sim}{\mu}}$ was the projection of $\underset{\sim}{x}$ on the subspace V or R^{np} generated by B. Thus $\underset{\sim}{x} - B \hat{\underset{\sim}{\mu}}$ is orthogonal to every vector in V and lies in the orthocomplement of V, denoted by $V^{\perp}$. Moreover, every vector in R^{np} decomposes uniquely into two orthogonal components, one in V and the other in $V^{\perp}$. Specifically, $\underset{\sim}{x} - B\,\underset{\sim}{\mu} = (\underset{\sim}{x} - B\,\hat{\underset{\sim}{\mu}}) + (B\,\hat{\underset{\sim}{\mu}} - B\,\underset{\sim}{\mu})$, and, by the Pythagorean property,

$$||\underset{\sim}{x} - B\,\underset{\sim}{\mu}||^2 = ||\underset{\sim}{x} - B\,\hat{\underset{\sim}{\mu}}||^2 + ||B(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu})||^2 ,$$

where the norm is with respect to the metric matrix $\Lambda_{\underset{\sim}{X}}$. Thus,

$$f_{\underset{\sim}{X}^t}(\underset{\sim}{x}^t;\ \underset{\sim}{\mu}) = k \exp\{-(1/2)||\underset{\sim}{x} - B\,\hat{\underset{\sim}{\mu}}||^2\} \exp\{-(1/2)||B(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu})||^2\}$$

$$= g(\underset{\sim}{x})\ h\ (\hat{\underset{\sim}{\mu}},\ \underset{\sim}{\mu}) ,$$

and sufficiency is established. Since $\hat{\underset{\sim}{\mu}}$ is p-variate normal and the p-variate normal is a special case of the exponential family of distribution, which is complete, the density of $\hat{\underset{\sim}{\mu}}$ is complete. Thus, $\hat{\underset{\sim}{\mu}}$ is indeed a UMVUE of $\underset{\sim}{\mu}$.

Let us now show how the fact that $\hat{\underset{\sim}{\mu}} \sim N_p(\underset{\sim}{\mu},\ \Sigma_{\hat{\underset{\sim}{\mu}}})$ can be used to detect departures of $\underset{\sim}{\mu}$ from the nominal value $\underset{\sim}{\mu}_0$. Actually, interest is centered on how much $\underset{\sim}{\mu} - \underset{\sim}{\mu}_0$ deviates from the zero vector, where $\underset{\sim}{\mu} - \underset{\sim}{\mu}_0$ will be measured by $\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0$ since $\hat{\underset{\sim}{\mu}}$ is eminently qualified to estimate $\underset{\sim}{\mu}$. This results in focusing interest on the distributional properties of the statistic $(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)^t\ \Sigma_{\hat{\underset{\sim}{\mu}}}^{-1}(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)$ since the deviations need to be squared and weighted by the "reciprocal" of their variability. First we need to recall the noncentral chi-square random variable (see Graybill [31]):

If the $(p \times 1)$ vector $\underset{\sim}{W} \sim N_p(\underset{\sim}{a}, I_p)$, then $\underset{\sim}{W}^t \underset{\sim}{W}$ is distributed as a non-central chi-square random variable with p degrees of freedom and non-centrality parameter $\lambda = \underset{\sim}{a}^t \underset{\sim}{a}$, denoted by $\chi'^{\,2}_{p,\lambda}$.

When $\lambda = 0$, the noncentral chi-square random variable reduces to a regular chi-square random variable. A solution to the distributional problem can now be formulated.

Theorem 2.3: Since $\hat{\underset{\sim}{\mu}} \sim N(\underset{\sim}{\mu}, \Sigma_{\hat{\underset{\sim}{\mu}}})$, then $(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu})^t \Sigma_{\hat{\underset{\sim}{\mu}}}^{-1}(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0) \sim \chi'^{\,2}_{p,\lambda}$ where $\lambda = (\underset{\sim}{\mu} - \underset{\sim}{\mu}_0)^t \Sigma_{\hat{\underset{\sim}{\mu}}}^{-1}(\underset{\sim}{\mu} - \underset{\sim}{\mu}_0)$.

Proof: Since $\Sigma_{\hat{\underset{\sim}{\mu}}}$ is positive definite, there exists a nonsingular matrix R such that $\Sigma_{\hat{\underset{\sim}{\mu}}} = RR^t$. Let $\underset{\sim}{W} = R^{-1}(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)$. Then $E(\underset{\sim}{W}) = R^{-1}(\underset{\sim}{\mu} - \underset{\sim}{\mu}_0)$ and $C(\underset{\sim}{W}, \underset{\sim}{W}^t) = E(\underset{\sim}{W} - \underset{\sim}{\mu}_{\underset{\sim}{W}})(\underset{\sim}{W} - \underset{\sim}{\mu}_{\underset{\sim}{W}})^t = R^{-1} \; \Sigma_{\hat{\underset{\sim}{\mu}}}(R^{-1})^t = I_p$. Thus, $\underset{\sim}{W} \sim N_p(R^{-1}(\underset{\sim}{\mu} - \underset{\sim}{\mu}_0), I_p)$ and $\underset{\sim}{W}^t \underset{\sim}{W} \sim \chi'^{\,2}_{p,\lambda}$ where $\lambda = (\underset{\sim}{\mu} - \underset{\sim}{\mu}_0)^t(R^{-1})^t R^{-1}(\underset{\sim}{\mu} - \underset{\sim}{\mu}_0) = (\underset{\sim}{\mu} - \underset{\sim}{\mu}_0)^t \Sigma_{\hat{\underset{\sim}{\mu}}}^{-1}(\underset{\sim}{\mu} - \underset{\sim}{\mu}_0)$ and $\underset{\sim}{W}^t \underset{\sim}{W} = (\underset{\sim}{\mu} - \underset{\sim}{\mu}_0)^t \Sigma_{\hat{\underset{\sim}{\mu}}}^{-1}(\underset{\sim}{\mu} - \underset{\sim}{\mu}_0)$. ||

Note that when $E(\underset{\sim}{X}_i) = \underset{\sim}{\mu}_0$, for $i = 1, 2, \ldots, n$, then $\hat{\underset{\sim}{\mu}} \sim N_p(\underset{\sim}{\mu}_0, \Sigma_{\hat{\underset{\sim}{\mu}}})$, and

$$(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)^t \Sigma_{\hat{\underset{\sim}{\mu}}}^{-1}(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0) \sim \chi_p^2 . \tag{55}$$

Thus, to determine whether the process mean has shifted from the nominal value $\underset{\sim}{\mu}_0$, the decision maker would calculate $\hat{\underset{\sim}{\mu}}$ and determine whether

$$(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)^t \Sigma_{\hat{\underset{\sim}{\mu}}}^{-1}(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0) > \chi_{p,\alpha}^2 \quad . \tag{56}$$

If the above decision rule holds, then the decision maker would conclude that the process mean has shifted from $\underset{\sim}{\mu}_0$. For successive samples of size n, the test statistic is plotted on a chart with UCL = $\chi_{p,\alpha}^2$ similar to that of Figure 4, but with $\chi_{1,\alpha}^2$ replaced by $\chi_{p,\alpha}^2$.

When the decision maker computes $(\hat{\mu} - \mu_0)^t \Sigma_{\hat{\mu}}^{-1}(\hat{\mu} - \mu_0)$ and compares it with the control limit for successive samples of size n, he is merely performing repeated tests of significance. If one adopts this viewpoint, then the power of the test, denoted $\pi(\lambda)$, is given by

$$\pi(\lambda) = P[(\hat{\mu} - \mu_0)^t \Sigma_{\hat{\mu}}^{-1}(\hat{\mu} - \mu_0) > \chi_{p,\alpha}^{2} \mid \mu], \tag{57}$$

where $\pi(0) = \alpha$.

Let us determine the nature of the decision rule presented in equation (56) when the off-diagonal submatrices in Σ_X are zero, that is, $\Sigma_{ij} = 0$ for $i \neq j$. From equation (51), we see that $\hat{\mu} = (\sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ij})^{-1} (\sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ji}\, x_i)$. Since Σ_X is a block diagonal matrix, so is Λ_X. Furthermore, assuming that $\Lambda_{11} = \ldots = \Lambda_{nn} = \Lambda$, then $\sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ij} = n\Lambda = n\, \Sigma^{-1}$ and $(\sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ij})^{-1} = \Sigma/n$. Now $\sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ji}\, x_i = \sum_{i=1}^{n} \Lambda_{ii}\, x_i = \Lambda \sum_{i=1}^{n} x_i =$

$n\, \Lambda\, \bar{x} = n\, \Sigma^{-1}\, \bar{x}$ and $\hat{\mu} = (\Sigma/n)(n\, \Sigma^{-1}\, \bar{x}) = \bar{x}$. Also note that $\Sigma_{\hat{\mu}}^{-1} = \sum_{i=1}^{n} \sum_{j=1}^{n} \Lambda_{ij} = n\, \Lambda = n\, \Sigma^{-1}$. Thus, $(\hat{\mu} - \mu_0)^t\, \Sigma_{\hat{\mu}}^{-1}(\hat{\mu} - \mu_0) = n(\bar{x} - \mu_0)^t\, \Sigma^{-1}(\bar{x} - \mu_0)$. The conclusion is that the decision rule presented in equation (56) for multivariate, correlated observations reduces to that presented in equation (41) for multivariate, uncorrelated observations as should be the case since $\Sigma_{ij} = 0$ for $i \neq j$ implies uncorrelated observations.

When the decision maker reaches the conclusion that μ has shifted from μ_0, the determination of those components of μ responsible for this conclusion is of prime importance. One way to handle this problem is through the use of Sidak's inequality (see equation (42)). This results in the following set of simultaneous confidence intervals of bounded level $1 - \alpha$:

$$[\hat{\mu}_i - c_i \sqrt{\underset{\sim}{e}_i^t \, \Sigma_{\hat{\underset{\sim}{\mu}}} \, \underset{\sim}{e}_i} \, , \; \hat{\mu}_i + c_i \sqrt{\underset{\sim}{e}_i^t \, \Sigma_{\hat{\underset{\sim}{\mu}}} \, \underset{\sim}{e}_i} \,] \, ,$$

for $i = 1, 2, \ldots, p$, where $\underset{\sim}{e}_i^t = [0, \ldots, 0, 1, 0, \ldots, 0]$ with a 1 in the i^{th} position, $\hat{\mu}_i = \underset{\sim}{e}_i^t \hat{\underset{\sim}{\mu}}$, and c_i is such that $2\,\Phi(c_i) = 1 + (1 - \alpha)^{1/p}$.

In Section 2.1.2, it was pointed out how the ARMA models of order (p, q) could be used to represent different types of correlative structure for the sample elements. In the multivariate case, there is a similar correspondence. The generalization of univariate ARMA (p, q) models to the multivariate case is usually obtained by substituting vectors and matrices for the scalar quantities. For purposes of exploration, consider a bivariate, first-order moving average process, the model for which is presented in equation (58):

$$\left.\begin{aligned} \underset{\sim}{X}_t &= \underset{\sim}{\mu} - \Theta \; \underset{\sim}{a}_{t-1} + \underset{\sim}{a}_t \, , \\ \underset{\sim}{a}_t &\sim NID_2(\underset{\sim}{0}, \, \Sigma_{\underset{\sim}{a}}) \, . \end{aligned}\right\} \qquad (58)$$

The "$\underset{\sim}{a}_t \sim NID_2$" denotes that the $\underset{\sim}{a}_t$ are bivariate normally distributed random variables and that they are uncorrelated across time. For the bivariate case, $\underset{\sim}{X}_t$, $\underset{\sim}{\mu}$, and $\underset{\sim}{a}_t$ are each (2 x 1) vectors while Θ is a (2 x 2) matrix. Thus, the first part of equation (58) can be written as

$$\left.\begin{aligned} X_{1t} &= \mu_1 - \theta_{11}\, a_{1,t-1} - \theta_{12}\, a_{2,t-1} + a_{1t} \\ X_{2t} &= \mu_2 - \theta_{21}\, a_{1,t-1} - \theta_{22}\, a_{2,t-1} + a_{2t} \end{aligned}\right\} \tag{59}$$

Note that equation (58) describes the multivariate classical linear regression model when Θ is the zero matrix. In accordance with Fuller's [26] notation, let $\Gamma(h)$ denote the covariance matrix of $\underset{\sim}{X}_t$ and $\underset{\sim}{X}_{t+h}$. That is, $\Gamma(h) = E(\underset{\sim}{X}_t - \underset{\sim}{\mu})(\underset{\sim}{X}_{t+h} - \underset{\sim}{\mu})^t$ and let $\Gamma(-h) = E(\underset{\sim}{X}_t - \underset{\sim}{\mu})(\underset{\sim}{X}_{t-h} - \underset{\sim}{\mu})^t$. It follows that

$$\Gamma(h) = \left\{\begin{array}{ll} \Sigma_{\underset{\sim}{a}} + \Theta\, \Sigma_{\underset{\sim}{a}}\, \Theta^t & , h = 0 \\ -\Sigma_{\underset{\sim}{a}}\, \Theta^t & , h = 1 \\ -\Theta\, \Sigma_{\underset{\sim}{a}} & , h = -1 \\ 0 & , \text{otherwise} \end{array}\right\} \tag{60}$$

For a sample of size n from such a process, the covariance matrix of $\underset{\sim}{X}$ is as follows:

$$\Sigma_{\underset{\sim}{X}} = \begin{bmatrix} \Gamma(0) & \Gamma(1) & 0 & \cdots & 0 \\ \Gamma(-1) & \Gamma(0) & \Gamma(1) & \cdots & 0 \\ 0 & \Gamma(-1) & \Gamma(0) & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \Gamma(0) \end{bmatrix} \tag{61}$$

Thus the memory of a multivariate MA(1) model is only one period long, and the vectors of observations from such a process possess the multivariate analogue of first-order serial correlation. Note that $[\Gamma(1)]^t =$

$\Gamma(-1)$. For the univariate MA(1) process, it was required that $|\theta_1| < 1$ for purposes of invertibility. For the multivariate MA(1) process, the analogue is that the p = 2 roots of the determinantal equation

$$|Im - \Theta| = 0 \tag{62}$$

be less than one in absolute value (see Fuller [26]). Note that equation (61) is a special case of equation (46) with $\Sigma_{ii} = \Gamma(0)$, $\Sigma_{i-j} = \Gamma(1)$ if $i-j=-1$, $\Sigma_{i-j} = \Gamma(-1)$ if $i-j=1$, and $\Sigma_{|i-j|} = 0$ if $|i-j| > 1$.

One advantage in representing autocorrelation between the vectors of observations by using the multivariate analogue of ARMA models is that it facilitates the simulation of output from a process that meets the conditions of Theorem 2.2. Another more general advantage is that these models facilitate the study of the robustness of multivariate test procedures to departures from the independence assumption.

In order to demonstrate the decision rule presented in equation (56) and some of the other concepts in this section, consider the following example.

<u>Example 2.2</u> From past experience with a bivariate process, it is determined that the vectors of observations have a first-order serial correlation as exhibited in equation (61). From equation (60), we see that the components of $\Sigma_{\underset{\sim}{X}}$ are generally given as follows:

$$\Gamma(0) = \begin{bmatrix} -\theta_{11}\gamma_{11} - \theta_{12}\gamma_{21} + c^2_1 & -\theta_{11}\gamma_{12} - \theta_{12}\gamma_{22} + rc_1c_2 \\ -\theta_{21}\gamma_{11} - \theta_{22}\gamma_{21} + rc_1c_2 & -\theta_{21}\gamma_{12} - \theta_{22}\gamma_{22} + c^2_2 \end{bmatrix},$$

$$\Gamma(1) = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} = \begin{bmatrix} -c_1^2\,\theta_{11} - rc_1c_2\,\theta_{12} & -c_1^2\,\theta_{21} - rc_1c_2\,\theta_{22} \\ -rc_1c_2\,\theta_{11} - c_2^2\,\theta_{12} & -rc_1c_2\,\theta_{21} - c_2^2\,\theta_{22} \end{bmatrix} ,$$

$$\Gamma(-1) = \Gamma^t(1) ,$$

while

$$\Sigma_{\underset{\sim}{a}} = \begin{bmatrix} c_1^2 & rc_1c_2 \\ rc_1c_2 & c_2^2 \end{bmatrix}$$

Thus, the covariance structure of $\underset{\sim}{X}$ is specified when θ_{11}, θ_{12}, θ_{21}, θ_{22}, r, c_1, and c_2 are specified. Furthermore, θ_{11}, θ_{12}, θ_{21}, and θ_{22} must be fixed in such a way that equation (62) is satisfied. Specifically, we need to check that $|m_i| < 1$, $i = 1, 2$, where

$$m_i = [(\theta_{11} + \theta_{22}) \pm \sqrt{(\theta_{11} - \theta_{22})^2 + 4\,\theta_{12}\,\theta_{21}}]/2.$$

It should be noted that $|m_i| < 1$ for the triangular region shown in Figure 5. Thus, for any point in the triangular invertibility region,

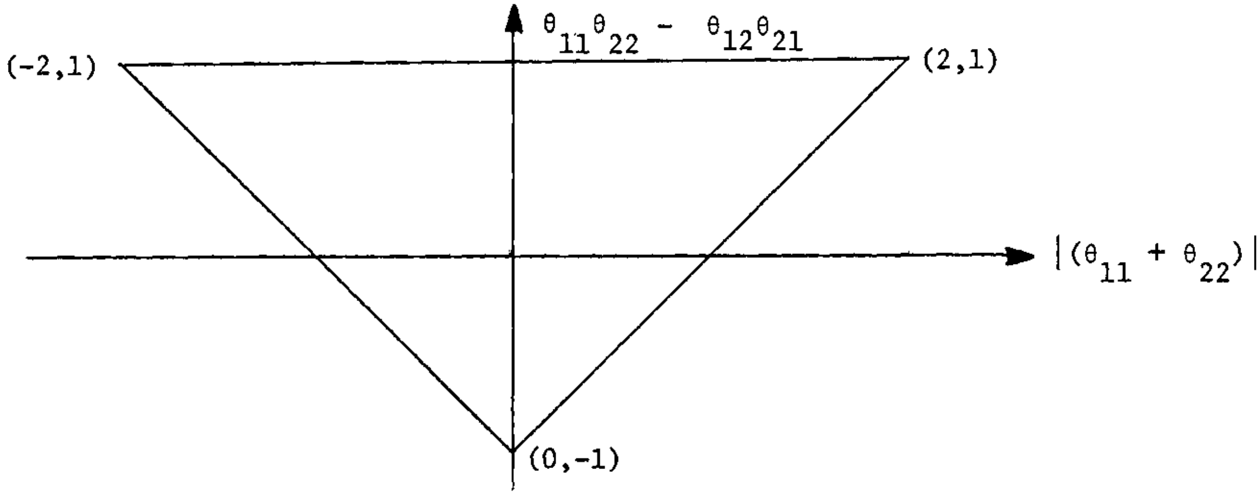

Figure 5. Invertibility Region for the Bivariate MA(1) Process.

there are many four-tuples $(\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22})$ corresponding to this point. For example, suppose we pick the point (0, -0.25). This implies $\theta_{11} + \theta_{22} = 0$ or $\theta_{11} = -\theta_{22}$. The other constraint is that $\theta_{11}\,\theta_{22} - \theta_{12}\,\theta_{21} = -0.25$, which reduces to $\theta^2_{22} + \theta_{12}\,\theta_{21} = 0.25$. If we let $\theta_{12} = \theta_{21}$, then the second constraint further reduces to $\theta^2_{22} + \theta^2_{12} = 0.25$. Thus the locus of points satisfying the second constraint is a circle with radius equal to $\sqrt{0.25}$ when we let $\theta_{12} = \theta_{21}$. To simulate output from such a process, many combinations of θ_{22} and θ_{12} could be selected. For example, four representative points are:

θ_{22}	θ_{12}
$\sqrt{2}/4$	$\sqrt{2}/4$
$-\sqrt{2}/4$	$\sqrt{2}/4$
$-\sqrt{2}/4$	$-\sqrt{2}/4$
$\sqrt{2}/4$	$-\sqrt{2}/4$

For illustrative purposes, let us look at $\theta_{22} = \sqrt{2}/4$ and $\theta_{12} = \sqrt{2}/4$, in which case $\theta_{21} = \sqrt{2}/4$ and $\theta_{11} = -\sqrt{2}/4$. Furthermore, r was set equal to zero, c_1 was set equal to one, and c_2 ranged from 1.0 to 4.0 in increments of 1.0. For convenience, $\mu_1 = \mu_2 = 0$. In total, 4 simulations were run with the same random number seed used in each run.

To gain further insight into the simulation procedure, let us rewrite equations (59) and (60) using the specific parameter values. Thus,

$$\Gamma(1) = \begin{bmatrix} \sqrt{2}/4 & -\sqrt{2}/4 \\ -c^2{}_2\ \sqrt{2}/4 & -\ c^2{}_2\sqrt{2}/4 \end{bmatrix} \ , \ \Gamma(-1) = \Gamma^t(1) \ ,$$

$$\Gamma(0) = \begin{bmatrix} (9/8) + (1/8)\ c^2{}_2 & (1/8)(c^2{}_2 - 1) \\ (1/8)(c^2{}_2 - 1) & c^2{}_2 + (1/8)(c^2{}_2 + 1) \end{bmatrix}$$

and

$$\Sigma_{\underset{\sim}{a}} = \begin{bmatrix} 1 & 0 \\ 0 & c^2{}_2 \end{bmatrix}$$

Furthermore,

$$X_{1t} = (\sqrt{2}/4)\ a_{1,t-1} - (\sqrt{2}/4)\ a_{2,t-1} + a_{1t}$$

$$X_{2t} = -\ (\sqrt{2}/4)\ a_{1,t-1} - (\sqrt{2}/4)\ a_{2,t-1} + a_{2t}$$

Thus, in these four runs, we are investigating the effect of increasing the variance of the second white noise generator.

Since we set $a_{1,0}$ and $a_{2,0}$ equal to zero, we discarded the first 100 $\underset{\sim}{X}_t$'s to overcome any transient effects. The first sample consisted of observations $\underset{\sim}{x}_{101}, \ldots, \underset{\sim}{x}_{110}$, from which $\hat{\underset{\sim}{\mu}}$ (as given in equation (51)) was calculated as well as $(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)^t(\Sigma_{\hat{\underset{\sim}{\mu}}})^{-1}(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)$, which was designated as SUM 1. The second sample consisted of observations $\underset{\sim}{x}_{151}, \ldots, \underset{\sim}{x}_{160}$. Again, both $\hat{\underset{\sim}{\mu}}$ and a value of the test statistic were computed. This was continued for a total of twenty samples. As a measure of comparison, $\bar{\underset{\sim}{x}}$

was also calculated for each sample as well as $n(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0)^t \Sigma^{-1}(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0)$, which was designated as SUM 2. SUM 2 ignores the autocorrelative structure since $\Sigma = \Gamma(0)$. The results are presented in Appendix B.

Note that for a fixed sample number, SUM 2 decreases as c_2 increases. This is apparent from examining the off-diagonal elements of $\Gamma(0)$ which become more negative as c_2 increases. Also, SUM 2 does not make use of $\Gamma(1)$. This decreasing behavior characteristic of SUM 2 becomes very important as we compare SUM 1 with SUM 2.

To investigate the effect of increasing c_2, we compare the magnitude of SUM 1 with SUM 2. For run 1, only 15% of the time was SUM 1 larger than SUM 2; for run 2, this increased to 35%; for run 3, this increased to 55%; and, for run 4, the figure is 65%. This increasing percentage is expected because of the decreasing behavior of SUM 2 discussed in the preceding paragraph. The significance of this is that as c_2 increases SUM 2 (the test statistic which ignores the autocorrelative structure) may fail to detect a shift in the population mean; however, for small c_2 SUM 2 will tend to indicate that a shift has occurred when, in fact, it has not. This concludes Example 2.2.

In this chapter, control charts for the mean were reviewed and developed for four different cases: (i) one quality characteristic, independent observations (ii) one quality characteristic, correlated observations (iii) multiple quality characteristics, independent observations, and (iv) multiple quality characteristics, correlated observations. While cases (i) and (iii) have been previously discussed in the literature, additional motivation for their use has been presented by demonstrating that the test statistic has favorable risk characteristics. The test statistics used in cases (ii) and (iv) also enjoy this property. Additional properties of the control procedure, such as the power and the relation to generalized least squares, were also presented.

CHAPTER III

ESTIMATION FOR THE MULTI-CONSEQUENCE INTERVENTION MODEL

In Chapter I, the concept of multiplicative empirical-stochastic models of order (p,d,q) x $(P,D,Q)_S$ was introduced and a synopisized list was presented of their enormous success in modeling a temporal sequence of occurrences for different scenarios. In Chapter II, a mathematical model was presented for both the univariate and multivariate ARMA (p,q) models, and it was explained how there is a relationship between these models and different types of autocorrelative structures. Chapter I also introduced the concept of an intervention model and the unique prespective it offers in evaluating an unplanned experiment with correlated observations for a change in the level of the underlying process.

Section 3.1 will elaborate upon earlier introductions to ARMA models with particular emphasis given to ARMA models of order (0,0,1) and (0,0,2). This section culminates with a full specification of the probability density function of a set of n observations from either an MA(1) or MA(2) process. Section 3.2 focuses on the estimation of the model parameters described in Section 3.1 via the technique of iterative, conditional least squares. Particular emphasis is given to the case where the treatment has altered not only the level of the series but also the values of the moving average parameters, which is designated the multi-consequence intervention model. Section 3.3 also addresses the estimation of these parameters but from the maximum likelihood viewpoint. The maximum likelihood estimates can be used to set up an asymptotic likelihood ratio test to investigate the hypothesis that the moving average parameters prior to the intevention are equal to those after the intervention. This section also shows why the maximum likelihood estimates may be different from the least squares estimates.

This Chapter concludes with an example for which both least squares and maximum likelihood estimates are obtained.

Selected portion of this Chapter appear in a paper by Alt, Deutsch, and Goode [4].

3.1 Description of MA(1) and MA(2) Models

3.1.1 Non-Intervention Situation

One very useful technique in modeling a temporal sequence of occurrences from a process is the multiplicative empirical-stochastic models proposed by Box and Jenkins [13]. The general form of these models of order (p,d,q) x $(P,D,Q)_S$ is given by

$$\left.\begin{aligned} &\phi_p(B)\Phi_P(B^S)\ \nabla^d \nabla_S^D\ Z_t = \theta_q(B)\ \Theta_Q\ (B^S)a_t, \\ &\qquad\qquad a_t \sim \mathrm{NID}\ (0,\sigma_a^2)\ , \end{aligned}\right\} \tag{63}$$

where $\phi_p(B)$ and $\Phi_P(B^S)$ are the nonseasonal and seasonal autoregressive operators, $\theta_q(B)$ and $\Theta_Q(B^S)$ are the nonseasonal and seasonal moving average operators, ∇^d and ∇_S^D are nonstationary and seasonal differencing operators, and S is the seasonal lag. For example, the multiplicative model of order (0,1,1) x $(0,1,1)_{12}$ is written as

$$\nabla \nabla_{12} Z_t = (1-\theta B) \ (1 - \Theta \ B^{12}) a_t, \tag{64}$$

where p=P=0, q=Q=1, d=D=1, and S=12. In equation (64), it is assumed that $a_t \sim NID(0,\sigma_a^2)$. By making use of the fact that $B^k Z_t = Z_{t-k}$ and $\nabla_k = (1-B)^k$, we see that equation (64) has the following equivalent representation:

$$Z_t - Z_{t-1} - Z_{t-12} + Z_{t-13} = a_t - \theta a_{t-1} - \Theta \ a_{t-12} + \theta \ \Theta \ a_{t-13}.$$

When there is no seasonal component (P=0, D=0, and Q=0), the multiplicative model reduces to the autoregressive integrated moving average (ARIMA) model of order (p,d,q), namely,

$$\phi_p(B) \ \nabla^d \ Z_t = \theta_q(B) \ a_t, \tag{65}$$

where quite frequently $\nabla^d Z_t$ is written as W_t. When d=1,2, the effect is to remove linear and quadratic trend, respectively, so that W_t is stationary in level. If no differencing is necessary, equation (65) reduces to

$$\phi_p(B) \ \tilde{Z}_t = \theta_q(B) \ a_t, \tag{66}$$

where $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, $\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$, and $\tilde{Z}_t = Z_t - \mu$ with μ denoting the process mean. This is frequently denoted the ARMA (p,q) model, where the weights $\phi_1, \dots, \phi_p$ and $\theta_1, \dots, \theta_q$ must satisfy certain stationarity-invertibility conditions. In this Chapter, we will be specifically concerned with the case where $\phi_1 = \dots = \phi_p = 0$. The notation MA(q) is used for such models. For example, the MA(1) model is given by

$$Z_t = \mu + a_t - \theta_1 a_{t-1} , \tag{67}$$

while the MA(2) model is given by

$$Z_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}. \tag{68}$$

The MA(1) and MA(2) models are invertible only if

$$-1 < \theta_1 < 1, \tag{69}$$

and

$$\left.\begin{array}{l} \theta_1 + \theta_2 < 1 \\ \theta_2 - \theta_1 < 1 \\ -1 < \theta_2 < 1 , \end{array}\right\} \tag{70}$$

respectively. No further restrictions are required for stationarity since, whatever the values of θ_1 and θ_2, equations (67) and (68) both define stationary processes. Since $a_t \sim NID(0,\sigma_a^2)$, it follows that for an MA(1) process

$$E(Z_t) = \mu , \tag{71}$$

$$Var(Z_t) = Var(\mu + a_t - \theta_1 a_{t-1}) = \sigma_a^2(1 + \theta_1^2), \tag{72}$$

and

$$\begin{aligned} Cov(Z_t, Z_{t+j}) &= E[(\mu + a_t - \theta_1 a_{t-1} - \mu)(\mu + a_{t+j} - \theta_1 a_{t+j-1} - \mu)] \\ &= E[a_t a_{t+j} - \theta_1 a_t a_{t+j-1} - \theta_1 a_{t-1} a_{t+j} + \theta_1^2 a_{t-1} a_{t+j-1}]. \end{aligned} \tag{73}$$

Thus

$$\left.\begin{aligned}\mathrm{Cov}(Z_t, Z_{t+j}) &= -\theta_1 \sigma_a^2 , \quad j=1 \\ &= 0 \quad , \quad j>1 ,\end{aligned}\right\} \tag{74}$$

and the memory of an MA(1) process is only one period long. The covariance matrix of the sample elements $Z_1, \ldots, Z_n$ from an MA(1) process, denoted by $\Sigma_{Z}^{(0,1)}$, is given by

$$\Sigma_{\underset{\sim}{Z}}^{(0,1)} = \sigma_a^2 \begin{bmatrix} 1+\theta_1^2 & -\theta_1 & 0 & \ldots & 0 \\ -\theta_1 & 1+\theta_1^2 & -\theta_1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1+\theta_1^2 \end{bmatrix} ; \tag{75}$$

and, the (nxl) expected value vector, denoted by $\underset{\sim}{\mu}_{\underset{\sim}{Z}}$, is given by

$$\underset{\sim}{\mu}_{\underset{\sim}{Z}} = [\mu, \mu, \ldots, \mu]^t = \mu \underset{\sim}{j}_n \quad , \tag{76}$$

where $\underset{\sim}{j}_n$ is the (nxl) vector all of whose entries are l's. Later on, it will prove convenient to adopt Box and Jenkins' notation and let $\Sigma_{\underset{\sim}{Z}}^{(0,1)} = \sigma_a^2 (M_n^{(0,1)})^{-1}$. Let $\underset{\sim}{Z}^t = [Z_1, Z_2, \ldots, Z_n]$ and let $\underset{\sim}{a}^t = [a_0, a_1 \ldots, a_n]$. Then $\underset{\sim}{Z} = C^{(0,1)} \underset{\sim}{a} + \underset{\sim}{\mu}_{\underset{\sim}{Z}}$, where $C^{(0,1)}$ is the following [n x (n+1)] matrix:

$$C^{(0,1)} = \begin{bmatrix} -\theta_1 & 1 & 0 & \ldots & 0 \\ 0 & -\theta_1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix} , \tag{77}$$

and it follows that $\underset{\sim}{Z}$ is distributed as an n-variate normal. To summarize, if $Z_1, Z_2, \ldots, Z_n$ emanate from an MA(1) process at n equispaced successive times, then

$$\underset{\sim}{Z} \sim N_n(\mu \, \underset{\sim}{j}_n^t, \Sigma_Z^{(0,1)}) \quad . \tag{78}$$

Similar results are obtained for an MA(2) process. Namely,

$$\underset{\sim}{\mu}_{\underset{\sim}{Z}} = \mu \, \underset{\sim}{j}_n \quad , \tag{79}$$

$$\Sigma_Z^{(0,2)} = \sigma_a^2 \begin{bmatrix} (1+\theta_1^2+\theta_2^2) & -\theta_1(1-\theta_2) & -\theta_2 & 0 & \cdots & 0 \\ -\theta_1(1-\theta_2) & (1+\theta_1^2+\theta_2^2) & -\theta_1(1-\theta_2) & -\theta_2 & \cdots & 0 \\ -\theta_2 & -\theta_1(1-\theta_2) & (1+\theta_1^2+\theta_2^2) & -\theta_1(1-\theta_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & (1+\theta_1^2+\theta_2^2) \end{bmatrix}$$

$$= \sigma_a^2 (M_n^{(0,2)})^{-1} \quad , \tag{80}$$

and $\underset{\sim}{Z} = C^{(0,2)} \, \underset{\sim}{a} + \underset{\sim}{\mu}_{\underset{\sim}{Z}}$, where $\underset{\sim}{a}^t = [a_{-1}, a_o, a_1, \ldots, a_n]$ and $C^{(0,2)}$ is the following [n x (n+2)] matrix:

$$C^{(0,2)} = \begin{bmatrix} -\theta_2 & -\theta_1 & 1 & 0 & \cdots & 0 \\ 0 & -\theta_2 & -\theta_1 & 1 & \cdots & 0 \\ 0 & 0 & -\theta_2 & -\theta_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & & 1 \end{bmatrix} \tag{81}$$

In summary, for an MA(2) model,

$$\underset{\sim}{Z} \sim N_n(\underset{\sim}{\mu}, \Sigma_{\underset{\sim}{Z}}^{(0,2)}) \quad .$$

The foregoing results immediately enable us to write down the probability density function, $f_{\underset{\sim}{Z}}$, of either MA process:

$$f_{\underset{\sim}{Z}^t}(\underset{\sim}{z}^t;\mu,\underset{\sim}{\theta}^t,\sigma_a^2)=(2\pi)^{-n/2}|\Sigma_{\underset{\sim}{Z}}^{(0,q)}|^{-1/2}\exp\{-(1/2)(\underset{\sim}{z}-\mu\underset{\sim}{j}_n)^t(\Sigma_{\underset{\sim}{Z}}^{(0,q)})^{-1}(\underset{\sim}{z}-\mu\underset{\sim}{j}_n)\}, \tag{82}$$

where $\underset{\sim}{\theta} = \theta_1$ for an MA(1) process and $\underset{\sim}{\theta} = [\theta_1,\theta_2]^t$ for an MA(2) process. Although the results were specifically developed for MA(1) and MA(2) processes, they are easily generalized to higher-order moving average processes.

3.1.2 Continuous Intervention Situation

As indicated in Chapter I, we will be primarily concerned with the continuous intervention situation where the treatment remains in effect at each time period after it has been introduced. For example, if we are monitoring the monthly occurrences of homicide for a particular city, an intervention might consist of a gun control law which remains in effect for a relatively long period of time after its introduction. Furthermore, we will assume that the intervention abruptly changes the level of the observations, although other types of level changes can be easily accommodated. To account for a possible change in level upon introducing an intervention after the $n_1^{\underline{th}}$ observation, consider the following modification of an MA(1) process:

$$\left.\begin{aligned} Z_t &= \mu + a_t - \theta_1 a_{t-1},\ t = 1,\ldots, n_1; \\ Z_t &= \mu + \delta + a_t - \theta_1 a_{t-1},\ t = n_1 + 1,\ldots, n_1 + n_2. \end{aligned}\right\} \tag{83}$$

We will assume $a_t \sim NID(0,\sigma_a^2)$ for $t=1,\ldots,n$, where $n=n_1+n_2$. This modified single consequence intervention model and its statistical analysis have been briefly considered by Glass, Willson, and Gottman [28]. We will further modify the intervention model of equation (83) to allow for the intervention affecting the process variability as well as the level. This multi-consequence intervention model has the following formulation:

$$\left.\begin{aligned} Z_t &= \mu + a_t - \theta_1 a_{t-1},\ t = 1,\ldots, n_1; \\ Z_t &= \mu + \delta + a_t - \gamma_1 a_{t-1},\ t = n_1 + 1,\ldots, n_1 + n_2. \end{aligned}\right\} \tag{84}$$

Thus, the model given in equation (84) differs from that presented in equation (83) since γ_1 has replaced θ_1 for $t=n_1+1,\ldots, n$. From equation (84), it follows that

$$E(Z_t) = \mu\ ,\ t = 1,\ldots, n_1$$

and

$$E(Z_t) = \mu + \delta, \; t = n_1 + 1, \ldots, n_1 + n_2 \quad ,$$

which is identical with the expected value of the single consequence intervention model presented in equation (83). If we partition the $(n \times 1)$ vector $\underset{\sim}{Z}$ into two components, namely, the $(n_1 \times 1)$ vector $\underset{\sim}{Z}_1 = [Z_1, \ldots, Z_{n_1}]^t$ and the $(n_2 \times 1)$ vector $Z_2 = [Z_{n_1+1}, \ldots, Z_n]^t$, then $E(\underset{\sim}{Z})$, denoted by $\underset{\sim}{\mu}_{\underset{\sim}{Z}}$, can be written as

$$\underset{\sim}{\mu}_{\underset{\sim}{Z}} = \begin{bmatrix} \underset{\sim}{\mu}_{\underset{\sim}{Z}_1} \\ \underset{\sim}{\mu}_{\underset{\sim}{Z}_2} \end{bmatrix} = \mu \; \underset{\sim}{j}_n + \delta \; \underset{\sim}{k}, \tag{85}$$

where

$$\underset{\sim}{k} = \begin{bmatrix} \underset{\sim}{0}_{n_1} \\ \underset{\sim}{j}_{n_2} \end{bmatrix} \quad . \tag{86}$$

Thus, the $(n \times 1)$ vector $\underset{\sim}{k}$ has 0's for its first n_1 entires followed by n_2 1's. It also follows from equation (84) that

$$\mathrm{Var}(Z_t) = \mathrm{Var}(\mu + a_t - \theta_1 a_{t-1}) = \sigma_a^2(1+\theta_1^2), \; t=1, \ldots, n_1,$$

while

$$\mathrm{Var}(Z_t) = \mathrm{Var}(\mu+\delta+a_t-\gamma_1 a_{t-1}) = \sigma_a^2(1+\gamma_1^2), \; t=n_1, \ldots, n_1+n_2 \; .$$

Furthermore,

$$\mathrm{Cov}(Z_t, Z_{t+1}) = - \theta_1 \; \sigma_a^2 \; , \; t=1, \ldots, n_1-1,$$

$$\mathrm{Cov}(Z_{n_1}, Z_{n_1+1}) = \mathrm{Cov}(Z_{n_1+1}, Z_{n_1}) = - \gamma_1 \, \sigma_a^2 \quad ,$$

and

$$\mathrm{Cov}(Z_t, Z_{t+1}) = -\gamma_1 \sigma_a^2, \ t=n_1+1,\ldots, n.$$

The above statements concerning the variance-covariance structure of $\underset{\sim}{Z}$ can be written in matrix form. Specifically, if $\Sigma_{\underset{\sim}{Z}}^{(0,1)}$ denotes (n x n) covariance matrix of $\underset{\sim}{Z}$, then

$$\Sigma_{\underset{\sim}{Z}}^{(0,1)} = \sigma_a^2 \left[\begin{array}{c|c} B_{\underset{\sim}{Z}_1}^{(0,1)} & (B_{21}^{(0,1)})^t \\ \hline B_{21}^{(0,1)} & B_{\underset{\sim}{Z}_2}^{(0,1)} \end{array}\right] \tag{87}$$

where

$$B_{\underset{\sim}{Z}_1}^{(0\ 1)} = \begin{bmatrix} (1+\theta_1^2) & -\theta_1 & \cdots & 0 & 0 \\ -\theta_1 & (1+\theta_1^2) & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & (1+\theta_1^2) & -\theta_1 \\ 0 & 0 & \cdots & -\theta_1 & (1+\theta_1^2) \end{bmatrix}$$

$$B_{\underset{\sim}{Z}_2}^{(0,1)} = \begin{bmatrix} (1+\gamma_1^2) & -\gamma_1 & \cdots & 0 & 0 \\ -\gamma_1 & (1+\gamma_1^2) & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & (1+\gamma_1^2) & -\gamma_1 \\ 0 & 0 & \cdots & -\gamma_1 & (1+\gamma_1^2) \end{bmatrix}$$

and

$$B_{21}^{(0,1)} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -\gamma_1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} .$$

Thus, $\Sigma_{\underset{\sim}{Z}}^{(0,1)}$ is a diagonal matrix of type 2 as are $B_{\underset{\sim}{Z}_1}^{(0,1)}$ and $B_{\underset{\sim}{Z}_2}^{(0,1)}$, which $B_{21}^{(0,1)}$ is the zero matrix except for the element in the northeast corner which is $-\gamma_1$. Furthermore, since $\underset{\sim}{Z} = C_I^{(0,1)} \underset{\sim}{a} + \underset{\sim}{\mu}_{\underset{\sim}{Z}}$, where $C_I^{(0,1)}$ is an $[n \times (n+1)]$ matrix similar to that presented in equation (77) except θ_1 in rows $n_1 + 1$ through rows $n_1 + n_2$ is replaced by γ_1, we see that $\underset{\sim}{Z}$ is distributed as an n-variate normal. This can be summarized by saying that for a first-order moving average intervention process, denoted $MA_I(1)$,

$$\underset{\sim}{Z} \sim N_n(\underset{\sim}{\mu}_{\underset{\sim}{Z}} , \Sigma_{\underset{\sim}{Z}}^{(0,1)}) \quad , \tag{88}$$

where $\underset{\sim}{\mu}_{\underset{\sim}{Z}}$ and $\Sigma_{\underset{\sim}{Z}}^{(0,1)}$ are presented in equation (85) and (87), respectively.

Let us now consider when the observations follow a second-order moving average model, and the intervention produces a constant effect starting with the $(n_1 + 1)^{th}$ observation. The model formulation is as follows:

$$\left.\begin{aligned} Z_t &= \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2},\ t = 1, \ldots, n_1 \quad ; \\ Z_t &= \mu + \delta + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2},\ t = n_1+1, \ldots, n_1+n_2 . \end{aligned}\right\} \tag{89}$$

We reformulate the single consequence model of equation (89) to a multi-consequence model by stating that

$$\left.\begin{aligned} Z_t &= \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}\ ,\ t = 1, \ldots, n_1\ ; \\ Z_t &= \mu + \delta + a_t - \gamma_1 a_{t-1} - \gamma_2 a_{t-2},\ t = n_1+1, \ldots, n_1+n_2. \end{aligned}\right\} \quad (90)$$

This reformulation not only allows for a change in level of the observations but also a change in their covariability. Since $a_t \sim \text{NID}(0,\sigma_a^2)$, it follows that

$$\underset{\sim}{\mu}_{\underset{\sim}{Z}} = \mu\, \underset{\sim}{j}_n + \delta\ \underset{\sim}{k}\ , \tag{91}$$

and

$$\Sigma_{\underset{\sim}{Z}}^{(0,2)} = \sigma_a^2 \left[\begin{array}{c|c} B_{\underset{\sim}{Z}_1}^{(0,2)} & (B_{21}^{(0,2)})^t \\ \hline B_{21}^{(0,2)} & B_{\underset{\sim}{Z}_2}^{(0,2)} \end{array}\right] \quad , \tag{92}$$

where

$$B_{\underset{\sim}{Z}_1}^{(0,2)} = \begin{bmatrix} (1+\theta_1^2+\theta_2^2) & -\theta_1(1-\theta_2) & -\theta_2 & \cdots & 0 & 0 \\ -\theta_1(1-\theta_2) & (1+\theta_1^2+\theta_2^2) & -\theta_1(1-\theta_2) & \cdots & 0 & 0 \\ -\theta_2 & -\theta_1(1-\theta_2) & (1+\theta_1^2+\theta_2^2) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (1+\theta_1^2+\theta_2^2) & -\theta_1(1-\theta_2) \\ 0 & 0 & 0 & \cdots & -\theta_1(1-\theta_2) & (1-\theta_1^2+\theta_2^2) \end{bmatrix} ,$$

(93)

$$B_{\underset{\sim}{Z}_2}^{(0,2)} = \begin{bmatrix} (1+\gamma_1^2+\gamma_2^2) & -\gamma_1(1-\gamma_2) & -\gamma_2 & \cdots & 0 & 0 \\ -\gamma_1(1-\gamma_2) & (1+\gamma_1^2+\gamma_2^2) & -\gamma_1(1-\gamma_2) & \cdots & 0 & 0 \\ -\gamma_2 & -\gamma_1(1-\gamma_2) & (1+\gamma_1^2+\gamma_2^2) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (1+\gamma_1^2+\gamma_2^2) & -\gamma_1(1-\gamma_2) \\ 0 & 0 & 0 & \cdots & -\gamma_1(1-\gamma_2) & (1+\gamma_1^2+\gamma_2^2) \end{bmatrix}, \tag{94}$$

and

$$B_{21}^{(0,2)} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -\gamma_2 & -\gamma_1+\theta_1\gamma_2 \\ 0 & 0 & \cdots & 0 & 0 & -\gamma_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix} . \tag{95}$$

Thus, for an $MA_I(2)$ model

$$\underset{\sim}{Z} \sim N_n(\underset{\sim}{\mu}_{\underset{\sim}{Z}}, \Sigma_{\underset{\sim}{Z}}^{(0,2)}) \quad , \tag{96}$$

where normality follows from the fact that $\underset{\sim}{Z} = C_I^{(0,2)} \underset{\sim}{a} + \underset{\sim}{\mu}_{\underset{\sim}{Z}}$ and $a_t \sim NID$.

3.2 Iterative, Conditional Least Squares (ICLS) Estimation

In the previous section, a detailed explanation was presented of the MA(1) and MA(2) models along with the modifications necessary to accommodate a multi-consequence intervention, that is, one which affects both the level and variability of the underlying process. In this section, we will be concerned with parameter estimation for the $MA_I(1)$ and $MA_I(2)$ models. Although we shall be primarily concerned with the estimation of μ and δ for each of these models, we shall see that both estimates are directly dependent upon the values of the moving-average parameters. Thus, we will use an iterative technique of searching on the moving-average parameters until those values are found which minimize the residual sum of squares. For this reason, the estimation technique is called iterative least squares. In order to provide a basis for estimation in the $MA_I(1)$ and $MA_I(2)$ cases, let us first consider the non-intervention MA(1) and MA(2) models.

3.2.1 Non-Intervention MA(q) Models

Let z_1, z_2, ..., z_n be n successive observations generated from the MA(1) process of equation (67), which can be rewritten as

$$a_t = z_t - \mu + \theta_1 a_{t-1} . \tag{97}$$

Box and Jenkins [13] suggest that μ can be replaced by $\bar{z} = n^{-1} \sum_{t=1}^{n} z_t$ where for "the sample sizes normally considered in time series analysis, this approximation will be adequate." Thus, equation (97) becomes

$$a_t = z_t - \bar{z} + \theta_1 a_{t-1} , \tag{98}$$

where θ_1 is the only unknown parameter. However, as Box and Jenkins point out, even when the z_t's are substituted into equation (98) and θ_1 is fixed, the a_t's still cannot be calculated recursively because a_1 depends on a_0 which is unknown. This difficulty is overcome by letting $a_0 = 0$, its marginal mean. Justification for this is given in Aigner [7]. Thus, conditional on $a_0 = 0$ and for a fixed θ_1, the a_t's in equation (98) can be recursively calculated. Actually, we are calculating $\hat{a}_t$'s, which are estimates of the unobservable a_t's. The objective is to find that value of θ_1 which minimizes

$$S_*(\theta_1) = \sum_{t=1}^{n} a_t^2(\theta_1 \mid a_0 = 0, \underset{\sim}{z}) = \sum_{t=1}^{n} (z_t - \bar{z} + \theta_1 a_{t-1})^2 \tag{99}$$

The asterisk subscript on S indicates that the sum of squares is conditional on $a_0 = 0$. This is further emphasized by the conditional notation of a_t, viz., $a_t(\theta_1 \mid a_0 = 0, \underset{\sim}{z})$. To assist in the search for θ_1, recall that $|\theta_1| < 1$ for invertibility purposes. Thus a table can be set up which lists θ_1 and $S_*(\theta_1)$ for the (-1,1) interval in whatever increments are desired. When a minimizing value of θ_1 is found, finer increments can be used over the reduced θ_1 neighborhood if so desired. Experience by other authors suggests that $S_*(\theta_1)$ is fairly well-behaved (unimodal) for large sample sizes.

One would proceed in a similar manner for the MA(2) process, where now

$$a_t = z_t - \mu + \theta_1 a_{t-1} + \theta_2 a_{t-2} \quad , \tag{100}$$

$$\begin{aligned} S_*(\theta_1,\theta_2) &= \sum_{t=1}^{n} a_t^2(\theta_1,\theta_2 \mid a_{-1} = a_0 = 0, \underset{\sim}{z}) \\ &= \sum_{t=1}^{n} (z_t - \bar{z} + \theta_1 a_{t-1} + \theta_2 a_{t-2}) \quad , \end{aligned} \tag{101}$$

and a grid search is performed to find those values of (θ_1,θ_2) which minimize $S_*(\theta_1,\theta_2)$. The extension of the ICLS estimation procedure to higher order MA(q) models is straightforward.

Box and Jenkins [13] give further justification for the ICLS procedure by relating it to a conditional likelihood function. Let $\underset{\sim}{a}^t = [a_1, a_2, \ldots, a_n]$. Then, since $a_t \sim \text{NID}(0,\sigma_a^2)$,

$$f(\underset{\sim}{a}^t) = (2\pi\sigma_a^2)^{-n/2} \exp\{-\underset{\sim}{a}^t \underset{\sim}{a}/2\sigma_a^2\}$$

and

$$\ln L_*(\underset{\sim}{\theta}^t, \sigma_a) = -(n/2)\ln 2\pi - (n/2)\ln \sigma_a^2 - S_*(\underset{\sim}{\theta}^t)/2\sigma_a^2 \quad ,$$

where L_* denotes the likelihood function conditional on $\underset{\sim}{a}_* = [a_o, a_{-1}, \ldots, a_{1-q}]^t = \underset{\sim}{0}^t$. Furthermore, since $\ln L_*$ depends on $\underset{\sim}{z}$ only through $S_*(\underset{\sim}{\theta})$, it follows that contours of $\ln L_*$ "for any fixed value of σ_a in the space of $(\underset{\sim}{\theta}, \sigma_a)$ are contours of S_*, that these maximum likelihood estimates are the same as the least squares estimates, and that in general we can, on the Normal assumption, study the behavior of the conditional likelihood by studying the conditional sum of squares function."

It is apparent from Box and Jenkins' write-up of the ICLS procedure that their primary interest is in obtaining values of the moving-average parameters with only a secondary interest in estimating μ . This emphasis is usually reversed for the intervention models. The adaptation of ICLS estimation to $MA_I(q)$ models is the topic of the next section.

3.2.2 $MA_I(q)$ Models

Statistical estimation of the intervention parameter δ and the process level was first reported by Box and Tiao [14]. Their results were exclusively for the ARIMA(0,1,1) model. See equation (65). The basic idea is to transform the n original observations to another set of variables amenable to statistical linear model analysis. Glass, Willson, and Gottman [28] extended their results to certain other ARIMA models by indicating the necessary transformation and providing examples. Because of their brief treatment of the single consequence $MA_I(1)$ model, we will further investigate this case before turning to the multi-consequence $MA_I(1)$ model, which has not been previously investigated.

3.2.2.1 Single-Consequence $MA_I(1)$ Model

The single consequence $MA_I(1)$ model was presented in equation (83), where it is postulated that the intervention abruptly changed the level of the series after the n_I^{th} observation. Before finding the necessary transformation, recall that the model $\underset{\sim}{Y} = X \underset{\sim}{\beta} + \underset{\sim}{a}$, with $\underset{\sim}{a} \sim N_n(\underset{\sim}{0}, \sigma^2 I)$, describes the classical normal linear regression model, details of which can be found in Goldberger [29]. In our case, $\underset{\sim}{Y}$ is an (n x 1) vector as is $\underset{\sim}{a}$, X is an (n x 2) matrix, and $\underset{\sim}{\beta} = [\mu, \delta]^t$. The transformation necessary to convert equation (83) into linear model form can be found by considering the first few z_t's. Specifically, $z_1 = \mu + a_1 - \theta_1 a_0$, where a_0 is unobtainable. However, if we let $a_0 = 0$, its marginal mean, then $z_1 = \mu + a_1$, which is linear model form. Thus, we let $y_1 = z_1$. Now $z_2 = \mu + a_2 - \theta_1 a_1$, where the $-\theta_1 a_1$ term prohibits z_2 from being in the desired format. However, if we multiply y_1 by θ_1 and add the result to z_2, the desired format is obtained. Namely, $y_2 = z_2 + \theta_1 y_1 = (1+\theta_1)\mu + a_2$. Similarly, $y_3 = z_3 + \theta_1 y_2 = (1+ \theta_1 + \theta_1^2)\mu + a_3$. In general,

$$y_t = (1+\theta_1 + \dots + \theta_1^{t-1}) \mu + a_t \quad , \tag{102}$$

for $t=1,\ldots, n_1$, where the required transformation is $y_t = z_t + \theta_1 y_{t-1}$ for $t=2,\ldots, n_1$. Since $z_{n_1+1} = \mu + \delta + a_{n_1+1} - \theta_1 a_{n_1}$, we see that $y_{n_1+1} = z_{n_1+1} + \theta_1 y_{n_1} = (1+\theta_1+\ldots+\theta_1^{n_1}) \mu + \delta + a_{n_1+1}$ is indeed in linear model format. Similarly. $y_{n_1+2} = (1+\theta_1+\ldots+\theta_1^{n_1+1}) \mu + (1+\theta_1) \delta + a_{n_1+2}$. In general

$$y_t = (1+\theta_1+\ldots+\theta_1^{t-1}) \mu + (1+\theta_1+\ldots+\theta_1^{t-(n_1+1)}) \delta + a_t, \qquad (103)$$

for $t=n_1 + 1,\ldots, n_1+n_2$. Equations (102) and (103) have the following matrix representation:

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n_1} \\ \hline y_{n_1+1} \\ y_{n_1+2} \\ \vdots \\ y_{n_1+n_2} \end{bmatrix}}_{Y} = \underbrace{\left[\begin{array}{cc} 1 & 0 \\ 1+\theta_1 & 0 \\ \vdots & \vdots \\ 1+\theta_1+\ldots+\theta_1^{n_1-1} & 0 \\ \hline 1+\theta_1+\ldots+\theta_1^{n_1} & 1 \\ 1+\theta_1+\ldots+\theta_1^{n_1+1} & 1+\theta_1 \\ \vdots & \vdots \\ 1+\theta_1+\ldots+\theta_1^{n_1+n_2-1} & 1+\theta_1+\ldots+\theta_1^{n_2-1} \end{array}\right]}_{X} \underbrace{\begin{bmatrix} \mu \\ \delta \end{bmatrix}}_{\beta} + \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_1} \\ \hline a_{n_1+1} \\ a_{n_1+2} \\ \vdots \\ a_{n_1+n_2} \end{bmatrix}}_{a} \qquad (104)$$

It immediately follows that $\hat{\underset{\sim}{\beta}} = (X^tX)^{-1} X^t\underset{\sim}{Y}$. At this point, Glass, Willson, and Gottman give a brief description of the iterative estimation procedure without giving specific formulae for $\hat{\mu}$ and $\hat{\delta}$. To fill this gap, let the elements of X^tX be denoted by c_{11}, c_{12}, and c_{22}. That is,

$$X^tX = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix} .$$

Now $c_{11} = 1 + (1+\theta_1)^2 + \ldots + (1+\theta_1+\ldots+\theta_1^{n_1-1})^2 + (1+\theta_1+\ldots+\theta_1^{n_1})^2 + \ldots + (1+\theta_1+\ldots+\theta_1^{n_1+n_2-1})^2$, and the individual terms are of the form $(1+\theta_1+\ldots+\theta_1^{i})^2$, for $i=0, 1, \ldots, n_1+n_2-1$. Recall that $\sum_{j=0}^{i} a^j = (1-a^{i+1})/(1-a)$. Thus

$$\left(\sum_{j=0}^{i} \theta_1^{j}\right)^2 = [(1-\theta_1^{i+1})/(1-\theta_1)]^2 = (1-2\theta_1^{i+1}))/(1-\theta_1)^2 ,$$

and

$$c_{11} = (1-\theta_1)^{-2} \sum_{i=0}^{n-1} (1-2\theta_1^{i+1} + \theta_1^{2(i+1)})$$

$$= (1-\theta_1)^{-2} (1-\theta_1^2)^{-1}[n(1-\theta_1^2) - 2\theta_1(1+\theta_1)(1-\theta_1^n) + \theta_1^2(1-\theta_1^{2n})] , \tag{105}$$

where $n = n_1 + n_2$. Proceeding in a similar fashion, we see that

$$c_{22} = (1-\theta_1)^{-2} (1-\theta_1^2)^{-1} [n_2(1-\theta_1^2) - 2\theta_1(1+\theta_1)(1-\theta_1^{2n_2}) + \theta_1^2(1-\theta_1^{2n_2})] \tag{106}$$

The calculations needed to obtain c_{12} are slightly more complex since the individual terms comprising c_{12} are of the form $(1+\theta_1+\ldots+\theta_1^{i})(1+\ldots+\theta_1^{n_1+i})$, for $i+0, 1, \ldots, n_2-1$. However, these individual terms can be rewritten as $(1+\theta_1+\ldots+\theta_1^{i})[(1+\theta_1+\ldots+\theta_1^{n_1-1})+\theta_1^{n_1}(1+\theta_1+\ldots+\theta_1^{i})]$

$= (1-\theta_1)^{-2}\ [(1-\theta_1^{\ i+1})(1-\theta_1^{\ n_1}) + \theta_1^{\ n_1}(1-\theta_1^{\ i+1})^2]$. Thus

$$c_{12} = (1-\theta_1)^{-2}(1-\theta_1^2)^{-1}[n_2(1-\theta_1^2) - \theta_1(1+\theta_1)(1-\theta_1^{\ n_2})(1+\theta_1^{\ n_1}) + \theta_1^{\ n_1+2}(1-\theta_1^{\ 2n_2})] \tag{107}$$

The calculation of $\hat{\underset{\sim}{\beta}}$ also depends upon the elements of the (2x1) vector $X^t\underset{\sim}{Y}$, denoted by s_{1Y} and s_{2Y}. The individual terms of s_{1Y} are of the form $(1+\theta_1+\ldots+\theta_1^{\ i-1})y_i$, for $i=1,2,\ldots,n_1+n_2$. Thus,

$$s_{1Y} = (1-\theta_1)^{-1} \sum_{i=1}^{n} (1-\theta_1^{\ i})y_i$$

$$= (1-\theta_1)^{-1}[(n_1+n_2)\bar{y}_{n_1+n_2} - \sum_{i=1}^{n_1+n_2} \theta_1^{\ i}\ y_i]. \tag{108}$$

where $\bar{y}_{n_1+n_2} = (n_1+n_2)^{-1} \sum_{i=1}^{n} y_i$. The second element of $X^t\underset{\sim}{Y}$, s_{2Y}, is the sum of individual terms of the form $(1+\theta_1+\ldots+\theta_1^{\ i-1})y_{n_1+i}$, for $i = 1,2,\ldots, n_2$. Thus,

$$s_{2Y} = (1-\theta_1)^{-1} \sum_{i=1}^{n_2} (1-\theta_1^{\ i})y_{n_1+i}$$

$$= (1-\theta_1)^{-1}(n_2\bar{y}_{n_2} - \sum_{i=1}^{n_2} \theta_1^{\ i}\ y_{n_1+i}) \quad , \tag{109}$$

where $\bar{y}_{n_2} = n_2^{-1} \sum_{i=1}^{n_2} y_{n_1+i}$. If we let c^{ij} denote the elements of $(X^tX)^{-1}$, where the elements of (X^tX) are given by equations (105)-(107), then

$$\hat{\underset{\sim}{\beta}} = \begin{bmatrix} \hat{\mu} \\ \hat{\delta} \end{bmatrix} = \begin{bmatrix} c^{11}\ s_{1Y} + c^{12}\ s_{2Y} \\ c^{12}\ s_{1Y} + c^{22}\ s_{2Y} \end{bmatrix} \quad . \tag{110}$$

Since c^{ij} and s_{iY} depend on θ_1, it may have been more appropriate to write the estimates of μ and δ as $\hat{\mu}(\theta_1)$ and $\hat{\delta}(\theta_1)$ to indicate that they are <u>conditional</u> least squares estimates. Since θ_1 is unknown, the estimates cannot be obtained. However, an ad hoc procedure has been suggested by Glass, Willson, and Gottman similar to the ICLS procedure of Box and Jenkins. Specifically, Let $\hat{\underset{\sim}{a}}$ denote the (nxl) vector of residuals or estimated errors. Then $\hat{\underset{\sim}{a}} = \underset{\sim}{y} - X\hat{\underset{\sim}{\beta}}$ where the values of $\hat{\underset{\sim}{a}}$ are contingent upon particular values of $\hat{\mu}$ and $\hat{\delta}$ which in turn are dependent upon θ_1. It seems reasonable to use that value of θ_1 which minimizes $S_*(\theta_1) = \sum_{t=1}^{n} \hat{a}_t^2 = \hat{\underset{\sim}{a}}^t \hat{\underset{\sim}{a}} = (\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})^t(\underset{\sim}{y}-X\hat{\underset{\sim}{\beta}})$, where minimizing $S_*(\theta_1)$ is equivalent to minimizing $\hat{\sigma}_a^2 = \hat{\underset{\sim}{a}}^t\hat{\underset{\sim}{a}}/(n-2)$, the estimated error variance. For that value of θ_1, $\hat{\mu}$ and $\hat{\delta}$ can be calculated from equation (110). The output format can be set up in table fashion with the following column headings: θ_1, $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}_a^2$, where the search for θ_1 is restricted to the interval (-1,1). One can then either perform tests of significance or construct confidence intervals for μ and δ by making use of the fact that both $(\hat{\mu}-\mu)/\hat{\sigma}_a(c^{11})^{1/2}$ and $(\hat{\delta}-\delta)/\hat{\sigma}_a(c^{22})^{1/2}$ are distributed as "pseudo" Student-t random variables with n-2 degrees of freedom. Actually, these quantities are T_{n-2} random variables only for known θ_1 as opposed to some fixed θ_1 which was found by searching on θ_1. Thus, one avenue of research is the true distribution of these quantities. Furthermore, $\hat{\mu}$ and $\hat{\delta}$ are correlated random variables and their joint confidence interval is elliptical. Thus, any condifence interval for μ or δ alone is merely a marginal one and the confidence level should be adjusted accordingly using some simultaneous procedure.

Before turning to the multi-consequence $MA_I(1)$ model, note that Box and Jenkins could have included μ as an additional parameter to be estimated instead of substituting $\bar{z}$ for μ, and they point this out. In this case, the (nxl) observation vector $\underset{\sim}{z}$ could have been transformed to an (nxl) vector $\underset{\sim}{y}$, via the transformation $y_1 = z_1$, and $y_t = z_t + \theta_1 y_{t-1}$, for t=2,...,n, and then ICLS could have been used to estimate μ and θ_1. However, their primary interest was in estimating θ_1 whereas the primary purpose of the transformation is to facilitate finding μ(and δ for the $MA_I(1)$ model) with θ_1 treated as a nuisance parameter. Regardless of whether the data is first transformed or not, both estimation approaches are iterative in that they search on θ_1, they are conditional in that $a_0=0$, and they both seek to find that value of θ_1 which minimizes $S_*(\theta_1) = \sum_{t=1}^{n} \hat{a}_t^2$. Thus, there is essentially no difference between the ICLS estimation technique of Box and Jenkins and that of Glass, Willson, and Gottman. Furthermore, any difference that does occur is a result of the basic difference in the philosophy of the traditional MA(1) model and that of the $MA_I(1)$ model.

3.2.2.2 Multi-Consequence $MA_I(1)$ Model

The multi-consequence $MA_I(1)$ model was presented in equation (84), where the assumption is that the introduction of a treatment after the n_1^{th} observation abruptly changed the level of the series by a linear additive effect δ and also altered the moving average parameter from θ_1 to γ_1. In order to transform the z_t's to y_t's, which are in statistical linear model form, we let $a_0=0$, $y_1=z_1$, and $y_t=z_t+\theta_1 y_{t-1}$, for t=2,..., n_1, while $y_t=z_t + \gamma_1 y_{t-1}$, for $t=n_1+1, \ldots, n_1+n_2$. Thus,

$$y_t = (1+\theta_1+\ldots+\theta_1^{t-1})\mu + a_t, \tag{111}$$

for $t=1,\ldots, n_1$, while

$$y_t = (1+\gamma_1+\ldots+\gamma_1^{t-n_1}+\theta_1\gamma_1^{t-n_1}+\ldots+\theta_1^{n_1-1}\gamma_1^{t-n_1})\mu$$
$$+(1+\gamma_1+\ldots+\gamma_1^{t-(n_1+1)})\delta+a_t \tag{112}$$

for $t=n_1+1,\ldots,n_1+n_2$. Equations (111) and (112) have the following matrix representation:

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n_1} \\ \hline y_{n_1+1} \\ y_{n_1+2} \\ \vdots \\ y_{n_1+n_2} \end{bmatrix}}_{\underset{\sim}{Y}} = \underbrace{\begin{bmatrix} 1 & 0 \\ (1+\theta_1) & 0 \\ \vdots & \vdots \\ (1+\theta_1+\ldots+\theta_1^{n_1-1}) & 0 \\ \hline [1+\gamma_1(1+\theta_1+\ldots+\theta_1^{n_1-1})] & 1 \\ [1+\gamma_1+\gamma_1^2(1+\theta_1+\ldots+\theta_1^{n_1-1})] & (1+\gamma_1) \\ \vdots & \vdots \\ [1+\ldots+\gamma_1^{n_2}(1+\theta_1+\ldots+\theta_1^{n_1-1})] & (1+\gamma_1+\ldots+\gamma_1^{n_2-1}) \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \mu \\ \delta \end{bmatrix}}_{\underset{\sim}{\beta}} + \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_1} \\ \hline a_{n_1+1} \\ a_{n_1+2} \\ \vdots \\ a_{n_1+n_2} \end{bmatrix}}_{\underset{\sim}{a}} \tag{113}$$

The elements of X^tX will be denoted by c_{11}, c_{12}, and c_{22}. After much tedious algebra, it can be shown that

$$c_{11} = (1-\theta_1)^{-2}[n_1-(1-\theta_1)^{-1}(2\theta_1)(1-\theta_1{}^{n_1})+(1-\theta_1^2)^{-1}\theta_1^2(1-\theta_1{}^{2n_1})]$$

$$+ (1-\gamma_1)^{-2}[n_2-(1-\gamma_1)^{-1}(2\gamma_1)(1-\gamma_1{}^{n_2})+(1-\gamma_1^2)^{-1}\gamma_1^2(1-\gamma_1{}^{2n_2})]$$

$$+ (1-\theta_1)^{-2}(1-\gamma_1^2)^{-1}(1-\theta_1{}^{n_1})^2\gamma_1^2(1-\gamma_1{}^{2n_2})$$

$$+ 2(1-\theta_1)^{-1}(1-\gamma_1)^{-1}(1-\theta_1{}^{n_1})[(1-\gamma_1)^{-1}\gamma_1(1-\gamma_1{}^{n_2})-(1-\gamma_1^2)^{-1}\gamma_1^2(1-\gamma_1{}^{2n_2})], \tag{114}$$

$$c_{22} = (1-\gamma_1)^{-2}(1-\gamma_1^2)^{-1}[n_2(1-\gamma_1^2)-2\gamma_1(1+\gamma_1)(1-\gamma_1{}^{n_2})+\gamma_1^2(1-\gamma_1{}^{2n_2})], \tag{115}$$

and

$$c_{12} = (1-\gamma_1)^{-2}[n_2-2(1-\gamma_1)^{-1}\gamma_1(1-\gamma_1{}^{n_2})+(1-\gamma_1^2)^{-1}\gamma_1^2(1-\gamma_1{}^{2n_2})]$$

$$+ (1-\theta_1)^{-1}(1-\gamma_1)^{-1}(1-\theta_1{}^{n_1})[(1-\gamma_1)^{-1}\ \gamma_1(1-\gamma_1{}^{n_2})-(1-\gamma_1^2)^{-1}\ \gamma_1^2(1-\gamma_1{}^{2n_2})]. \tag{116}$$

The results obtained in equations (114)-(116) were verified by letting $\gamma_1=\theta_1$ and observing that these results agreed with those presented in equations (105)-(107) for the single consequence $MA_I(1)$ model. Let s_{1Y} and s_{2Y} denote the elements of $X^t\underset{\sim}{Y}$. Then

$$s_{1Y} = (1-\theta_1)^{-1}(n_1\bar{y}_{n_1} - \sum_{i=1}^{n_1} \theta_1{}^i\ y_i)+(1-\gamma_1)^{-1}(n_2\bar{y}_{n_2} - \sum_{i=1}^{n_2} \gamma_1{}^i\ y_{n_1+i})$$

$$+ (1-\theta_1)^{-1}(1-\theta_1{}^{n_1}) \sum_{i=1}^{n_2} \gamma_1{}^i\ y_{n_1+i} \quad , \tag{117}$$

and

$$s_{2Y} = (1-\gamma_1)^{-1} \left(n_2\bar{y}_{n_2} - \sum_{i=1}^{n_2} \gamma_1^{\ i}\, y_{n_1+i}\right) , \tag{118}$$

where $\bar{y}_{n_1} = n_1^{-1} \sum_{i=1}^{n_1} y_i$ and $\bar{y}_{n_2} = n_2^{-1} \sum_{i=1}^{n_2} y_{n_1+i}$. These results also agree with those presented in equations (108) and (109) when θ_1 is substituted for γ_1. It follows from linear model theory that

$$\hat{\mu} = c^{11}\, s_{1Y} + c^{12}\, s_{2Y} \tag{119}$$

and

$$\hat{\delta} = c^{12}\, s_{1Y} + c^{22}\, s_{2Y} , \tag{120}$$

where c^{ij} denote the elements of $(X^tX)^{-1}$. Extending the ad hoc procedure of Glass, Willson, and Gottman to the multi-consequence model, we let $\hat{\underset{\sim}{a}} = \underset{\sim}{y} - X\hat{\underset{\sim}{\beta}}$, where the $\hat{\underset{\sim}{a}}$ vector is contingent upon particular values of $\hat{\mu}$ and $\hat{\delta}$ which in turn are contingent upon values of θ_1 and γ_1. Let $S_*(\theta_1,\gamma_1)$ be the sum of squared residuals or estimated errors for particular values of θ_1, γ_1, $\hat{\mu}$, and $\hat{\delta}$. That is, $S_*(\theta_1,\gamma_1) = \sum_{t=1}^{n} \hat{a}_t^{\,2} = \hat{\underset{\sim}{a}}^t\, \hat{\underset{\sim}{a}} = (\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})^t(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})$. It seems reasonable to find the (θ_1,γ_1) pair which minimizes $S_*(\theta_1,\gamma_1)$, where minimizing $S_*(\theta_1,\gamma_1)$ is equivalent to minimizing $\hat{\sigma}_a^2 = \hat{\underset{\sim}{a}}^t\, \hat{\underset{\sim}{a}}/(n-2)$. The search for the minimizing (θ_1,γ_1) pair can be restricted to the open unit square, that is, $(\theta_1,\gamma_1) \in \{(x_1,x_2)\colon\ 0 < x_i < 1,\ i = 1,2\}$. The output format associated with the search can be set up in table fashion with the following column headings: θ_1, γ_1, $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}_a$. Appendix C contains a listing of the computer program ICLSMAI(1) designed to find the optimal (θ_1,γ_1). After that (θ_1,γ_1) is selected which minimizes $\hat{\sigma}_a^2$, confidence intervals can be constructed or tests of significance can be performed for both μ and δ by making use of the fact that $(\hat{\mu}-\mu)/\hat{\sigma}_a(c^{11})^{1/2}$ and $(\hat{\delta}-\delta)/\hat{\sigma}_a(c^{22})^{1/2}$ are each distributed as pseudo-Student-t random variables with n-2 degrees of freedom. The "pseudo" prefix is necessitated by the fact that both ratios depend on the nuisance parameters (θ_1,γ_1). Furthermore, keep in mind that the true confidence region for (μ,δ) is elliptical in nature, even if (θ_1,γ_1) were known.

Although we will be primarily concerned with the maximum likelihood estimation of the parameters in the multi-consequence $MA_I(1)$ model, the ICLS estimates are useful as initial values for numerically determining the maximum likelihood estimates, and they also enable us to determine the closeness of these estimates to those obtained by the method of maximum likelihood. Let us now briefly consider the next higher order moving-average intervention model.

3.2.2.3 Single and Multi-Consequence $MA_I(2)$ Models

The single and multi-consequence $MA_I(2)$ models were presented in equations (89) and (90), respectively. The transformation necessary to convert the first n_1 observations of the single consequence $MA_I(2)$ model into linear model form is found by examining the first few z_t's. Specifically, $z_1 = \mu+a_1-\theta_1 a_0-\theta_2 a_{-1}$, where both a_0 and a_{-1} are unobtainable. However, if we let $a_0=a_{-1}=0$, then $z_1=\mu+a_1$, which is linear model form. Thus, $y_1=z_1$. Now $z_2=\mu+a_2-\theta_1 a_1-\theta_2 a_0=\mu+a_2-\theta_1 a_1$, since we set $a_0=0$, which is its marginal mean. We see that z_2 contains an unwanted $\theta_1 a_1$ term. However, if we multiply y_1 by θ_1 and add the result to z_2, the desired format is obtained. Namely, $y_2=z_2+\theta_1 y_1 = (\mu+a_2-\theta_1 a_1) + \theta_1(\mu+a_1)=(1+\theta_1)\mu+a_2$, which is indeed linear model form. Now

$z_3=\mu+a_3-\theta_1 a_2-\theta_2 a_1$, where the two terms $\theta_1 a_2$ and $\theta_2 a_1$ prohibit z_3 from being in linear model form. However, if we multiply y_2 by θ_1 and y_1 by θ_2 and add both of these terms to z_3, we obtain $y_3 = z_3 + \theta_1 y_2 + \theta_2 y_1 = [(1+\theta_1+\theta_1^2)+\theta_2]\mu+a_3$, which is linear model format. Similarly, we let $y_4 = z_4+\theta_1 y_3+\theta_2 y_2$, in which case $y_4=[(1+\theta_1+\theta_1^2+\theta_1^3)+\theta_2(1+2\theta_1)]\mu+a_4$. In general, the necessary transformation for the single consequence $MA_I(2)$ model is given by

$$\left.\begin{aligned} y_1 &= z_1 \\ y_2 &= z_2 + \theta_1 y_1 \\ y_t &= z_t + \theta_1 y_{t-1} + \theta_2 y_{t-2}, \end{aligned}\right\} \tag{121}$$

for $t=3,\ldots,n_1$. Note that the transformation given in equation (121) for the first n_1 observations of the single-consequence $MA_I(2)$ model is the same transformation used for the first n_1 transformations of the multi-consequence $MA_I(2)$ model since the models given in equation (89) and (90) are the same for the first n_1 observations. Furthermore, the necessary transformation on the second set of n_2 observations for the single consequence $MA_I(2)$ model is also given by

$$y_t = z_t + \theta_1 y_{t-1} + \theta_2 y_{t-2}, \tag{122}$$

for $t=n_1+1,\ldots, n_1+n_2$. This becomes obvious by examining the first few z_t's in this second set.

It is the transformations given in equations (121) and (122) that are of prime importance, for the transformed y_t's can be used as input to any standard regression package and the estimates of μ and δ are then easily obtained. However, in order to compare these y_t's with those prevously obtained for the single consequence $MA_I(1)$ model (see equation (102)), it would be nice to have a general expression for the y_t's. To accomplish this, expressions for $y_3,\ldots,y_8$ are obtained and rewritten as follows:

$$y_3 = [(1-\theta_1^3)(1-\theta_1)^{-1}]\mu + \theta_2\mu + a_3$$

$$y_4 = [(1-\theta_1^4)(1-\theta_1)^{-1}]\mu + \theta_2\mu[(d/d\theta_1)\{(1-\theta_1^3)(1-\theta_1)^{-1}\}] + a_4$$

$$y_5 = [(1-\theta_1^5)(1-\theta_1)^{-1}]\mu + \theta_2\mu[(d/d\theta_1)\{(1-\theta_1^4)(1-\theta_1)^{-1}\}] + \theta_2^2\mu + a_5$$

$$\begin{aligned} y_6 = {} & [(1-\theta_1^6)(1-\theta_1)^{-1}]\mu + \theta_2\mu[(d/d\theta_1)\{(1-\theta_1^5)(1-\theta_1)^{-1}\}] \\ & + (\theta_2^2\mu/2)\ [(d^2/d\theta_1^2)\{(1-\theta_1^4)(1-\theta_1)^{-1}\}] + a_6 \end{aligned}$$

$$\begin{aligned} y_7 = {} & [(1-\theta_1^7)(1-\theta_1)^{-1}]\mu + \theta_2\mu[(d/d\theta_1)\{(1-\theta_1^6)(1-\theta_1)^{-1}\}] \\ & + (\theta_2^2\mu/2)\ [d^2/d\theta_1^2)\{(1-\theta_1^5)(1-\theta_1)^{-1}\}] + \theta_2^3\mu + a_7 \end{aligned}$$

$$\begin{aligned} y_8 = {} & [(1-\theta_1^8)(1-\theta_1)^{-1}]\mu + \theta_2\mu[(d/d\theta_1)\{(1-\theta_1^7)(1-\theta_1)^{-1}\}] \\ & + (\theta^2\mu/2)[(d^2/d\theta_1^2)\{(1-\theta_1^6)(1-\theta_1)^{-1}\}] \\ & + (\theta_2^3\mu/3!)[(d^3/d\theta_1^3)\{(1-\theta_1^5)(1-\theta_1)^{-1}\} + a_8\ . \end{aligned}$$

Examination of the above equations suggests that a general expression for the first set of n_1 transformed observations from either the single or multi-consequence $MA_I(2)$ model is given by:

$$y_{2j} = [(1-\theta_1^{2j})(1-\theta_1)^{-1}]\mu + \theta_2\mu[(d/d\theta_1)\{(1-\theta_1^{2j-1})(1-\theta_1)^{-1}\}] + \ldots$$

$$+ [\theta_2^{\,j-1}\mu/(j-1)!][(d^{j-1}/d\theta_1^{\,j-1})\{(1-\theta_1^{\,j+1})(1-\theta_1)^{-1}\}] + a_{2j} ,$$

with

$$y_{2j+1} = [(1-\theta_1^{2j+1})(1-\theta_1)^{-1}]\mu + \theta_2\mu[(d/d\theta_1)\{(1-\theta_1^{2j})(1-\theta_1)^{-1}\}] + \ldots$$

$$+ [\theta_2^{\,j-1}\,\mu/(j-1)!][(d^{j-1}/d\theta_1^{\,j-1})\{(1-\theta_1^{\,j+2})(1-\theta_1)^{-1}\}] + \theta_2^j\mu + a_{2j+1} .$$

Note that when $\theta_2 = 0$, the above equations reduce to equation (102) which describes the n_1 transformed observations for the single-consequence $MA_I(1)$ model. This is predictable from examining equation (121) which reduces to $y_t = z_t + \theta_1\, y_{t-1}$ when $\theta_2 = 0$, which is the necessary transformation for the single-consequence $MA_I(1)$ model.

The only case that has not yet been considered is the transformation necessary on the n_2 observations after the treatment for the multi-consequence $MA_I(2)$ model. By examining the first few z_t's of equation (90), for $t=n_1+1,\ldots, n_1+n_2$, we see that the necessary transformation is given by

$$y_t = z_t + \gamma_1\, y_{t-1} + \gamma_2\, y_{t-2} \quad . \tag{123}$$

Thus, for the single-consequence $MA_I(2)$ model, the necessary transformation is given by equations (121) and (122), while for the multi-consequence $MA_I(2)$ model equations (121) and (123) describe the transformation. Once the transformation has been defined, the ICLS estimation procedure is straightforward. For the single-consequence $MA_I(2)$ model, it involves searching over (θ_1,θ_2) in the region given by equation (70) until that pair of values is found which minimizes $S_*(\theta_1,\theta_2) = \sum_{t=1}^{n} \hat{a}_t^2 = \hat{\underset{\sim}{a}}^t \hat{\underset{\sim}{a}} = (n-2)\hat{\sigma}_a^2$ where $\hat{\underset{\sim}{a}} = \underset{\sim}{y} - X\hat{\underset{\sim}{\beta}}$. The computer output associated with the search can be set up in a format with the following column headings: θ_1, θ_2, $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}_a^2$. For the minimizing (θ_1,θ_2) pair, additional statistical inference on μ and δ can be performed. One would proceed in a similar manner for the multi-consequence $MA_I(2)$ model where the search is now performed on the 4-tuple $(\theta_1,\theta_2,\gamma_1,\gamma_2)$ with the ordered pairs (θ_1,θ_2) and (γ_1,γ_2) each constrained to be in the triangular region described by equation (70).

The development of the necessary transformations and the application of the ICLS estimation procedure to higher order, single and multi-consequence $MA_I(q)$ models proceeds in a similar fashion. We will now investigate the maximum likelihood estimation of the parameters in the single and multi-consequence $MA_I(1)$ and $MA_I(2)$ models, where the ICLS estimates are used to provide initial estimates.

3.3 Maximum Likelihood Estimation

In this section, we present an algorithm for determining the exact likelihood function for single and multi-consequence $MA_I(1)$ and $MA_I(2)$ models for a given set of parameter values. It will be shown that while there are covenient analytical expressions for the maximum likelihood estimators of parameters μ and δ, no such expressions exist for the maximum likelihood estimators of the moving average parameters. However, this algorithm can be used to search the likelihood function over the permissible parameter space until those parameter values are found which maximize the likelihood function. Such values will be called the maximum likelihood estimates.

One reason that statistical inference for the pure moving average process is difficult stems from the fact that Arato [11] has shown that the dimensionality of the set of sufficient statistics is equal to the number of observations. That is, the number of sufficient statistics increases with the number of observations. He also shows that for a pure autoregressive process of order p the number of sufficient statistics is equal to $(p+1)(p+2)/2$. However, this is not to imply that the maximum likelihood estimates of a pure autoregressive process are easily obtained. As a matter of fact, when $p=1$, the maximum likelihood estimate of ϕ_1 is the solution to a cubic equation.

3.3.1 Maximum Likelihood Estimation of μ and δ

The single and multi-consequence $MA_I(q)$ models, $q=1,2$, were presented in equations (83)-(84) and (89)-(90), respectively. Each of these models shares a common facet in that the level of the series for the first n_1 observations equals μ while this level equals $\mu+\delta$ for observations $n_1+1,.., n_1+n_2=n$. In this section, we will obtain closed form expressions for the maximum likelihood estimates of μ and δ where these estimates are functions of the moving average parameters. Thus, these are conditional maximum likelihood estimates.

We will first consider the single consequence $MA_I(1)$ process. Let $\underset{\sim}{z} = [z_1,\ldots,z_{n_1}, z_{n_1+1},\ldots, z_{n_1+n_2}]^t$ be a sample of n observations generated from this process and let $\underset{\sim}{Z}$ be the (nx1) random vector assoicated with the vector of sample observations. Also, let $\underset{\sim}{a} = [a_0,a_1,\ldots,a_{n_1+n_2}]^t$ be an ((n+1) x 1) random vector where

$a_t \sim NID(0,\sigma_a^2)$. Thus, the joint distribution of $\underset{\sim}{a}$ equals

$$f(\underset{\sim}{a}^t;\sigma_a^2) = (2\pi\sigma_a^2)^{-(n+1)/2} \exp\{-\underset{\sim}{a}^t\underset{\sim}{a}/2\sigma_a^2\}. \tag{124}$$

Since $\underset{\sim}{Z} = C_I^{(0,1)}\ \underset{\sim}{a} + \underset{\sim}{\mu}_{\underset{\sim}{Z}}$, where $C_I^{(0,1)}$ is the [n x (n+1)] matrix presented in equation (77), it follows that $\underset{\sim}{Z} \sim N_n(\underset{\sim}{\mu}_{\underset{\sim}{Z}},\sigma_a^2\ C_I^{(0,1)}(C_I^{(0,1)})^t)$; and by definition, $\sigma_a^2\ C_I^{(0,1)}(C_I^{(0,1)})^t = \Sigma_{\underset{\sim}{Z}}^{(0,1)} = \sigma_a^2(M_n^{(0,1)})^{-1}$, where $\Sigma_{\underset{\sim}{Z}}^{(0,1)}$ is presented in equation (75). Thus,

$$f_{\underset{\sim}{Z}^t}(z^t;\mu,\delta,\theta_1,\sigma_a^2) = (2\pi\sigma_a^2)^{-n/2}|M_n^{(0,1)}|^{1/2}\exp\{-(\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}})^t M_n^{(0,1)}(\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}})/2\sigma_a^2\}, \tag{125}$$

where $\underset{\sim}{\mu}_{\underset{\sim}{Z}} = \mu\underset{\sim}{j}_n + \delta\underset{\sim}{k}$ and (nx1) vector $\underset{\sim}{k}$ is given in equation (86). In the logarithm of the likelihood function assoicated with equation (125), μ and δ appear only in the quadratic form

$$Q(\mu,\delta) = -\ (\underset{\sim}{z}-\mu\underset{\sim}{j}_n-\delta\underset{\sim}{k})^t\ M_n^{(0,1)}(\underset{\sim}{z}-\mu\underset{\sim}{j}_n-\delta\underset{\sim}{k})/2\sigma_a^2\ . \tag{126}$$

To find $\hat{\mu}$ and $\hat{\delta}$, the maximum likelihood estimates, note that

$$\begin{aligned} Q^*(\mu,\delta) &= -\ 2\sigma_a^2 Q(\mu,\delta)=((\underset{\sim}{z}-\delta\underset{\sim}{k})-\mu j_n)^t\ M_n^{(0,1)}((\underset{\sim}{z}-\delta\underset{\sim}{k}) - \mu\underset{\sim}{j}_n) \\ &= (\underset{\sim}{z}-\delta\underset{\sim}{k})^t M_n^{(0,1)}(\underset{\sim}{z}-\delta\underset{\sim}{k})-2\mu(\underset{\sim}{z}-\delta\underset{\sim}{k})^t M_n^{(0,1)}\ \underset{\sim}{j}_n+ \mu^2\underset{\sim}{j}_n^t\ M_n^{(0,1)}\ \underset{\sim}{j}_n\ , \end{aligned}$$

and

$$\partial Q^*(\mu,\delta)/\partial\mu = -\ 2(\underset{\sim}{z}-\delta\underset{\sim}{k})^t\ M_n^{(0,1)}\ \underset{\sim}{j}_n + 2\mu\ \underset{\sim}{j}_n^t\ M_n^{(0,1)}\ \underset{\sim}{j}_n \tag{127}$$

Also, note that

$$Q^*(\mu,\delta) = ((\underset{\sim}{z}-\mu\underset{\sim}{j}_n)-\delta\underset{\sim}{k})^t\ M_n^{(0,1)}((\underset{\sim}{z}-\mu\underset{\sim}{j}_n)-\delta\underset{\sim}{k})$$

$$= (\underset{\sim}{z}-\mu\underset{\sim}{j}_n)^t M_n^{(0,1)} (\underset{\sim}{z}-\mu\underset{\sim}{j}_n) - 2\delta\underset{\sim}{k}^t M_n^{(0,1)} (\underset{\sim}{z}-\mu\underset{\sim}{j}_n) + \delta^2 \underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{k}$$

and

$$\partial Q^*(\mu,\delta)/\partial\delta = -2\underset{\sim}{k}^t M_n^{(0,1)} (\underset{\sim}{z}-\mu\underset{\sim}{j}_n) + 2\delta \underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{k} . \tag{128}$$

When equations (127) and (128) are set equal to zero, we obtain the following pair of simultaneous equations:

$$\hat{\mu} \underset{\sim}{j}_n^t M_n^{(0,1)} \underset{\sim}{j}_n + \hat{\delta} \underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{j}_n = \underset{\sim}{z}^t M_n^{(0,1)} \underset{\sim}{j}_n$$

$$\hat{\mu} \underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{j}_n + \hat{\delta} \underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{k} = z^t M_n^{(0,1)} \underset{\sim}{k} ,$$

the solutions to which are given below:

$$\hat{\mu} = [(\underset{\sim}{z}^t M_n^{(0,1)} \underset{\sim}{j}_n) - \hat{\delta}(\underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{j}_n)]/\underset{\sim}{j}_n^t M_n^{(0,1)} \underset{\sim}{j}_n , \tag{129}$$

and

$$\hat{\delta} = \frac{(\underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{z})(\underset{\sim}{j}_n^t M_n^{(0,1)} \underset{\sim}{j}_n) - (\underset{\sim}{z}^t M_n^{(0,1)} \underset{\sim}{j}_n)(\underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{j}_n)}{(\underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{k})(\underset{\sim}{j}_n^t M_n^{(0,1)} \underset{\sim}{j}_n) - (\underset{\sim}{k}^t M_n^{(0,1)} \underset{\sim}{j}_n)^2} . \tag{130}$$

Note that, when $\delta = 0$ in the single-consequence model of equation (83), one only needs to estimate μ(assuming θ_1 is fixed) and equation (129) becomes $\hat{\mu} = \underset{\sim}{z}^t M_n^{(0,1)} \underset{\sim}{j}_n / \underset{\sim}{j}_n^t M_n^{(0,1)} \underset{\sim}{j}_n$, which is the result obtained in Chapter II (equation (16)) for estimating the mean of a univariate normal population when the sample elements are correlated. Thus, the quality control model presented in Chapter II is related to the intervection model presented in Chapter III. In one sense, the result presented in equation (16), which was also obtained by Dent [17], is more general than that obtained in equation (129) since it allows for any type of autocorrelative structure as opposed to that of an MA(q) process only. However, in another sense, the quality control model may appear to be more restrictive than the intervention model since a shift parameter is not specifically included. Equations (129)-(130) also point out that $\hat{\mu}$ and $\hat{\delta}$ are functions of the moving average parameter θ_1 since they depend on $M_n^{(0,1)} = \sigma_a^2(\Sigma_{\underset{\sim}{Z}}^{(0,1)})^{-1}$. However, these estimates are independent of σ_a^2.

The estimates given in equations (129)-(130) for the single consequence $MA_I(1)$ model are the same that would be obtained for the multi-consequence model with the exception that $\Sigma_{\underset{\sim}{Z}}^{(0,1)}$ is now given by equation (87), where $\Sigma_{\underset{\sim}{Z}}^{(0,1)} = \sigma_a^2(M_n^{(0,1)})^{-1}$. Thus, the estimates are functions of θ_1 and γ_1. Furthermore, the estimates of μ and δ for the multi-consequence $MA_I(2)$ are also of the same form with the exception that $\Sigma_{\underset{\sim}{Z}}^{(0,1)}$ is replaced by $\Sigma_{\underset{\sim}{Z}}^{(0,2)}$ as given in equation (92), where $\Sigma_{\underset{\sim}{Z}}^{(0,2)} = \sigma_a^2(M_n^{(0,2)})^{-1}$. The extension to higher-order MA processes is straightforward. Note that the main difficulty in obtaining $\hat{\mu}$ and $\hat{\delta}$, for fixed values of the moving average parameters, is the need to find $(\Sigma_{\underset{\sim}{Z}}^{(0,q)})^{-1}$ or equivalently $M_n^{(0,q)}$. This will be discussed in the following sections. Finally, note that equations (129)-(130) are valid for any type of ARMA(p,q) intervention process.

Let us now turn our attention to the estimation of the moving average parameters for the single and multi-consequence $MA_I(q)$ models, q=1,2.

3.3.2 Maximum Likelihood Estimation of Moving Average Parameters

This section addresses the maximum likelihood estimation of the moving average parameters for four specific cases: the single consequence $MA_I(1)$ model, the multi-consequence $MA_I(1)$ model, the single consequence $MA_I(2)$ model, and the multi-consequence $MA_I(2)$ model. The procedure used parallels that presented by Box and Jenkins [13], where Box and Jenkins treat the non-intervention moving average models and assumed $\mu = 0$. In order to handle the intervention model, their procedure needs to be modified for several reasons. First, the n observations for the intervention model are segmented into two groups where the moving average parameters may be different for each group. Second, we need to specifically include μ and δ since they are of prime interest in determining the effect of the intervention treatment. Third, we do not use the back-forecasting technique of Box and Jenkins to find estimates of $a_0, \ldots, a_{1-q}$ since this introduces a transient into the system, even though this effect may be small for large n. Instead, we use a least-squares estimate. Let us illustrate the procedure first for the single consequence $MA_I(1)$ model.

3.3.2.1 Single Consequence $MA_I(1)$ Model

The single consequence $MA_I(1)$ model was presented in equation (83). This can be rewritten as

$$\left.\begin{aligned} a_t &= Z_t - \mu + \theta_1\, a_{t-1},\ t = 1, \ldots, n_1 \\ &= Z_t - \mu - \delta + \theta_1\, a_{t-1},\ t = n_1 + 1, \ldots, n_1 + n_2\ , \end{aligned}\right\} \quad (131)$$

where $a_t \sim \mathrm{NID}(0, \sigma_a^2)$, $t=0, \ldots, n$. The joint distribution of $\underset{\sim}{a} = [a_0, a_1, \ldots, a_{n_1+n_2}]^t$ was given in equation (124), while the joint distribution of $\underset{\sim}{Z} = [Z_1, \ldots, Z_n]^t$ was given in equation (125). If equation (125) were to be interpreted as the likelihood function, then two basic problems exist in determining its value for a fixed set of parameter values, viz., finding $|M_n^{(0,1)}|$ and evaluating the quadratic form $Q(\mu,\delta)$ given in equation (126). These difficulties are overcome by making a transformation from the $[(n+1) \times 1]$ space of $\underset{\sim}{a}$ to the $[(n+1) \times 1]$ space of $\underset{\sim}{z}$ and a_*, where a_* denotes the preliminary value a_0. This transformation allows us to find the joint distribution of $\underset{\sim}{Z}$ and a_* as well as the conditional distribution of a_* given $\underset{\sim}{Z}$. The forms of these distributions enable us to overcome the above mentioned difficulties. The details of the transformation now follow.

From the model presented in equation (131), we can write down the following (n+1) equations, where the first equation is merely an identity:

$$
\begin{array}{l}
a_0 = a_0 \\
\hline
a_1 = Z_1 - \mu + \theta_1 a_0 \\
a_2 = Z_2 - \mu + \theta_1 a_1 \\
\quad \vdots \\
a_{n_1} = Z_{n_1} - \mu + \theta_1 a_{n_1-1} \\
\hline
a_{n_1+1} = Z_{n_1+1} - \mu - \delta + \theta_1 a_{n_1} \\
a_{n_1+2} = Z_{n_1+2} - \mu - \delta + \theta_1 a_{n_1+1} \\
\quad \vdots \\
a_{n_1+n_2} = Z_{n_1+n_2} - \mu - \delta + \theta_1 a_{n_1+n_2-1} .
\end{array}
$$

In the above system of equations, we substitute the expression for a_1 in that for a_2 and continue this substitution scheme until we have expressed $\underset{\sim}{a} = [a_0, \ldots, a_{n_1+n_2}]^t$ in terms of $\underset{\sim}{Z} = [Z_1, \ldots, Z_{n_1+n_2}]^t$ and $a_* = a_0$. Specifically,

$$a_0 = a_0$$

$$a_1 = Z_1 - \mu + \theta_1 a_0$$

$$a_2 = Z_2 + \theta_1 Z_1 - (1+\theta_1)\mu + \theta_1^2 a_0$$

$$\vdots$$

$$a_{n_1} = Z_{n_1} + \theta_1 Z_{n_1-1} + \ldots + \theta_1^{n_1-1} Z_1 - (1+\theta_1+\ldots+\theta_1^{n_1-1})\mu + \theta_1^{n_1} a_0$$

$$a_{n_1+1} = Z_{n_1+1} + \theta_1 Z_{n_1} + \ldots + \theta_1^{n_1} Z_1 - (1+\theta_1+\ldots+\theta_1^{n_1})\mu - \delta + \theta_1^{n_1+1} a_0$$

$$a_{n_1+2} = Z_{n_1+2} + \theta_1 Z_{n_1+1} + \ldots + \theta_1^{n_1+1} Z_1 - (1+\theta_1+\ldots+\theta_1^{n_1+1})\mu - (1+\theta_1)\delta + \theta_1^{n_1+2} a_0$$

$$\vdots$$

$$a_{n_1+n_2} = Z_{n_1+n_2} + \theta_1 Z_{n_1+n_2-1} + \ldots + \theta_1^{n_1+n_2-1} Z_1 - (1+\theta_1+\ldots+\theta_1^{n_1+n_2-1})\mu$$
$$-(1+\theta_1+\ldots+\theta_1^{n_2-1})\delta + \theta_1^{n_1+n_2} a_0 \tag{132}$$

The system of (n_1+n_2+1)equations in (132) has a matrix representation, namely,

$$\underset{\sim}{a} = L\,\underset{\sim}{Z} + \underset{\sim}{X}\,a_* - \underset{\sim}{b}\,\mu - \underset{\sim}{c}\,\delta, \tag{133}$$

where L is an $[(n_1+n_2+1)x(n_1+n_2)]$ matrix and $\underset{\sim}{X}$ is an $[(n_1+n_2+1)x1]$ vector as are $\underset{\sim}{b}$ and $\underset{\sim}{c}$. These are presented in equation (134), from which it is obvious that L, $\underset{\sim}{X}$, $\underset{\sim}{b}$, and $\underset{\sim}{c}$ are all functions of θ_1.

$$
L = \left[\begin{array}{cccc|cccc}
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\theta_1 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
\theta_1^{n_1-1} & \theta_1^{n_1-2} & \cdots & 1 & 0 & 0 & \cdots & 0 \\
\hline
\theta_1^{n_1} & \theta_1^{n_1-1} & \cdots & \theta_1 & 1 & 0 & \cdots & 0 \\
\theta_1^{n_1+1} & \theta_1^{n_1} & \cdots & \theta_1^2 & \theta_1 & 1 & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
\theta_1^{n_1+n_2-1} & \theta_1^{n_1+n_2-2} & \cdots & \theta_1^{n_2} & \theta_1^{n_2-1} & \theta_1^{n_2-2} & \cdots & 1
\end{array}\right],
$$

$$
\underset{\sim}{X} = \left[\begin{array}{c}
1 \\ \hline
\theta_1 \\ \theta_1^2 \\ \vdots \\ \theta_1^{n_1} \\ \hline
\theta_1^{n_1+1} \\ \theta_1^{n_1+2} \\ \vdots \\ \theta_1^{n_1+n_2}
\end{array}\right], \quad
\underset{\sim}{b} = \left[\begin{array}{c}
0 \\ \hline
1 \\ 1+\theta_1 \\ \vdots \\ 1+\theta_1+\ldots+\theta_1^{n_1-1} \\ \hline
1+\theta_1+\ldots+\theta_1^{n_1} \\ 1+\theta_1+\ldots+\theta_1^{n_1+1} \\ \vdots \\ 1+\theta_1+\ldots+\theta_1^{n_1+n_2-1}
\end{array}\right], \quad
\underset{\sim}{c} = \left[\begin{array}{c}
0 \\ \hline
0 \\ 0 \\ \vdots \\ 0 \\ \hline
1 \\ 1+\theta_1 \\ \vdots \\ 1+\theta_1+\ldots+\theta_1^{n_2-1}
\end{array}\right]
\tag{134}
$$

Recall that if the transformation from the (px1) vector $\underset{\sim}{U}$ to the (px1) vector $\underset{\sim}{V}$ is given by $\underset{\sim}{U}=B\underset{\sim}{V}$ where B is a nonsingular (pxp) matrix then the Jacobian, denoted by J, is the determinant of the matrix $\underset{\sim}{B}$. In our case, $\underset{\sim}{U}=\underset{\sim}{a}$, $\underset{\sim}{V}=[a_*|\underset{\sim}{z}^t]^t$, and B is the $[(n_1+n_2+1) \times (n_1+n_2+1)]$ matrix L^* where $L^*=[\underset{\sim}{X}|L]$. Thus, $|J|=1$, and by substituting equation (133) into equation (124) we see that the joint distribution of $\underset{\sim}{Z}$ and a_* is

$$f_{\underset{\sim}{Z}^t,a_*}(\underset{\sim}{z}^t,a_*;\mu,\delta,\ \theta_1,\sigma_a^2) = (2\pi\sigma_a^2)^{-(n+1)/2}\ \exp\{-S(\theta_1,a_*)/2\sigma_a^2\}, \tag{135}$$

where

$$S(\theta_1,a_*) = (L\underset{\sim}{z}+Xa_*-\underset{\sim}{b}\,\mu-\underset{\sim}{c}\delta)^t(L\underset{\sim}{z}+\underset{\sim}{X}a_*-\underset{\sim}{b}\,\mu-\underset{\sim}{c}\delta). \tag{136}$$

For convenience, we let $\underset{\sim}{d}=\underset{\sim}{b}\mu+\underset{\sim}{c}\delta$.

Define $\hat{a}_*$ to be the value of a_* which minimizes $S(\theta_1,a_*)$. To find $\hat{a}_*$, note that

$$S(\theta_1,a_*) = (L\underset{\sim}{z})^t(L\underset{\sim}{z}) + 2\ a_*\ \underset{\sim}{X}^t\ L\underset{\sim}{z} - 2\ \underset{\sim}{d}^t\ L\underset{\sim}{z} - 2a_*\underset{\sim}{d}^t\underset{\sim}{X} + a_*^2\underset{\sim}{X}^t\underset{\sim}{X}+\underset{\sim}{d}^t\underset{\sim}{d}$$

and

$$dS(\theta_1,a_*)/da_* = 2\ \underset{\sim}{X}^tL\underset{\sim}{z} - 2\underset{\sim}{d}^t\underset{\sim}{X} + 2a_*\ \underset{\sim}{X}^t\underset{\sim}{X}.$$

Setting this derivative equal to zero shows that $\hat{a}_*$ is the solution to the following normal equation:

$$\underset{\sim}{X}^t\underset{\sim}{X}\ \hat{a}_* = -\ \underset{\sim}{X}^t\ L\underset{\sim}{z} + \underset{\sim}{X}^t\ \underset{\sim}{d}. \tag{137}$$

To solve this equation, note that

$$\underset{\sim}{X}^t\underset{\sim}{X} = 1+\theta_1^2+\theta_1^4+\ldots+\theta_1^{2n_1}+\theta_1^{2n_1+2}+\ldots+\theta_1^{2(n_1+n_2)}$$

$$= (1-\theta_1^{2(n_1+n_2+1)})/(1-\theta_1^2). \tag{138}$$

Also note that $\underset{\sim}{X}^t L$ is a $[1 \times (n_1+n_2)]$ vector whose elements we will denote by $\ell_1, \ldots, \ell_{n_1+n_2}$, where

$$\ell_i = \theta_1^{\,i}(1-\theta_1^{2(n_1+n_2-i+1)})/(1-\theta_1^2) , \tag{139}$$

$i=1, \ldots, n_1+n_2$. Thus.

$$\underset{\sim}{X}^t L\underset{\sim}{z} = \sum_{i=1}^{n_1+n_2} \ell_i \, z_i . \tag{140}$$

Finally, note that

$$\underset{\sim}{X}^t\underset{\sim}{d} = \mu\, \theta_1(1-\theta_1)^{-1} \sum_{i=0}^{n_1+n_2-1} \theta_1^{\,i}(1-\theta_1^{\,i+1})$$

$$+ \delta\theta_1^{n_1+1}(1-\theta_1)^{-1} \sum_{i=0}^{n_2-1} \theta_1^{\,i}(1-\theta_1^{\,i+1}) . \tag{141}$$

From equation (137), we see that

$$\hat{a}_* = (-\underset{\sim}{X}^t L\underset{\sim}{z} + \underset{\sim}{X}^t\underset{\sim}{d})/(\underset{\sim}{X}^t\underset{\sim}{X}) , \tag{142}$$

where expressions for $\underset{\sim}{X}^t\underset{\sim}{X}$, $\underset{\sim}{X}^t L\,\underset{\sim}{z}$. and $\underset{\sim}{X}^t\underset{\sim}{d}$ are given in equations (138), (140), and (141), respectively.

Now $S(\theta_1, a_*)$ can be rewritten as follows:

$$S(\theta_1,a_*) = (L\underset{\sim}{z}+\underset{\sim}{X}\, a_*-\underset{\sim}{d})^t(L\underset{\sim}{z} + \underset{\sim}{X}\, a_*-\underset{\sim}{d})$$

$$= [(L\underset{\sim}{z}+\underset{\sim}{X}\,\hat{a}_*) - \underset{\sim}{X}(\hat{a}_*-a_*)-\underset{\sim}{d}]^t[(L\underset{\sim}{z}+\underset{\sim}{X}\hat{a}_*)-\underset{\sim}{X}(\hat{a}_*-a_*)-\underset{\sim}{d}]$$

$$= (L\underset{\sim}{z}+\underset{\sim}{X}\hat{a}_*)^t(L\underset{\sim}{z}+\underset{\sim}{X}\hat{a}_*)-2(\hat{a}_*-a_*)\underset{\sim}{X}^t(L\underset{\sim}{z}+\underset{\sim}{X}\hat{a}_*)-2\underset{\sim}{d}^t(L\underset{\sim}{z}+\underset{\sim}{X}\hat{a}_*)$$

$$+ (\hat{a}_* - a_*)^2 \; \underset{\sim}{X}^t \underset{\sim}{X} + 2(\hat{a}_* - a_*) \underset{\sim}{d}^t \underset{\sim}{X} + \underset{\sim}{d}^t \underset{\sim}{d}$$

$$= (L\underset{\sim}{z} + \underset{\sim}{X}\hat{a}_*)^t \; (L\underset{\sim}{z} + \underset{\sim}{X}\hat{a}_*) + (\hat{a}_* - a_*)^2 \underset{\sim}{X}^t \underset{\sim}{X} + \underset{\sim}{d}^t \underset{\sim}{d} - 2\underset{\sim}{d}^t (L\underset{\sim}{z} + \underset{\sim}{X}\hat{a}_*) \; ,$$

since $-2(\hat{a}_* - a_*)\underset{\sim}{X}^t (L\underset{\sim}{z} + \underset{\sim}{X}\hat{a}_*) + 2(\hat{a}_* - a_*)\underset{\sim}{d}^t \underset{\sim}{X}$

$$= -2(\hat{a}_* - a_*)[\underset{\sim}{X}^t L\underset{\sim}{z} + \underset{\sim}{X}^t \underset{\sim}{X}\hat{a}_* - \underset{\sim}{d}^t \; \underset{\sim}{X}] \; = \; 0 \; ,$$

from equation (137). Thus,

$$S(\theta_1, a_*) = \underbrace{[(L\underset{\sim}{z} + \underset{\sim}{X}\hat{a}_*) - \underset{\sim}{d}]^t [(L\underset{\sim}{z} + \underset{\sim}{X}\hat{a}_*) - \underset{\sim}{d}]}_{S(\theta_1)} + (a_* - \hat{a}_*)^2 \underset{\sim}{X}^t \underset{\sim}{X} \quad , \tag{143}$$

where $S(\theta_1)$ is a function of the observations but not of a_*. By definition,

$$f_{a_* | \underset{\sim}{Z}^t}(a_* | \underset{\sim}{z}^t;\; \mu, \delta,\; \theta_1, \sigma_a^2) = f_{\underset{\sim}{Z}^t, a_*}(\underset{\sim}{z}^t, a_*;\; \mu, \delta,\; \theta_1, \sigma_a^2) / f_{\underset{\sim}{Z}^t}(\underset{\sim}{z}^t;\; \mu, \delta, \theta_1, \sigma_a^2),$$

from which it follows that

$$f_{\underset{\sim}{Z}^t, a_*}(\underset{\sim}{z}^t, a_*;\; \mu, \delta,\; \theta_1, \sigma_a^2) = f_{\underset{\sim}{Z}^t}(\underset{\sim}{z}^t;\; \mu, \delta,\; \theta_1, \sigma_a^2) f_{a_* | \underset{\sim}{Z}^t}(a_* | \underset{\sim}{z}^t;\; \mu, \delta,\; \theta_1, \sigma_a^2), \tag{144}$$

where $f_{\underset{\sim}{Z}^t, a_*}$ is given in equation (135). Upon substituting equation (143) into equation (135) and then making use of equation (144), we see that

$$f_{a_* | \underset{\sim}{Z}^t}(a_* | \underset{\sim}{z}^t;\; \mu, \delta,\; \theta_1, \sigma_a^2) = (2\pi\sigma_a^2)^{-1/2} |\underset{\sim}{X}^t \underset{\sim}{X}|^{1/2} \; \exp\{-(a_* - \hat{a}_*)^2 (\underset{\sim}{X}^t \underset{\sim}{X}) / 2\sigma_a^2 \} \tag{145}$$

and

$$f_{\underset{\sim}{Z}^t}(z^t;\; \mu, \delta,\; \theta_1, \sigma_a^2) = (2\pi\sigma_a)^{-n/2} |\underset{\sim}{X}^t \underset{\sim}{X}|^{-1/2} \; \exp\{-S(\theta_1)/2\sigma_a^2\} \; , \tag{146}$$

where $S(\theta_1)$ is given in equation (143).

The following deductions can be made from the foregoing statements:

(i) From equation (145), we see that "$\hat{a}_*$ is the conditional expectation of a_*" given $\underset{\sim}{z}$ and $(\mu,\delta,\ \theta_1,\sigma_a^2) = \underset{\sim}{\xi}^t$. Denote $E(a_*|\underset{\sim}{z}^t,\ \underset{\sim}{\xi}^t)$ by $[a_*]$. Thus, $\hat{a}_* = [a_*]$. Since $\underset{\sim}{a} = L\underset{\sim}{z} + \underset{\sim}{X}a_* - \underset{\sim}{d}$, it follows that $[\underset{\sim}{a}] = L\underset{\sim}{z} + \underset{\sim}{X}[a_*] - \underset{\sim}{d}$ and that

$$S(\theta_1) = \sum_{t=0}^{n_1+n_2} [a_t]^2 \quad , \tag{147}$$

where $\hat{a}_*$ is obtained from equation (142).

(ii) By comparing equations (125) and (146), we see that

$$|\underset{\sim}{X}^t\underset{\sim}{X}|^{-1} = |M_n^{(0,1)}|$$

and

$$S(\theta_1) = (\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}})^t\ M_n^{(0,1)}(\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}}) \ .$$

Thus, an easy method for finding $|M_n^{(0,1)}|$ and evaluating the quadratic form has been provided. The determinant could have also been found by using a result of Rutherford's [53] or a later result of Shaman [58].

(iii) In order to compute $S(\theta_1) = \sum_{t=0}^{n_1+n_2} [a_t]^2$ for fixed θ_1, we let $[a_0] = \hat{a}_*$ and recursively calculate $[a_1]$ through $[a_{n_1+n_2}]$ from

$$[a_t] = z_t - \hat{\mu} + \theta_1[a_{t-1}], \tag{148}$$

for $t=1,2,\ldots,\ n_1$, while

$$[a_t] = a_t - \hat{\mu} - \hat{\delta} + \theta_1[a_{t-1}] \quad , \tag{149}$$

for $t=n_1+1,\ldots,\ n_1+n_2$.

These results are stated in the following theorem, which closes out this section.

Theorem 3.1: For the single consequence $MA_I(1)$ model, the unconditional likelihood is given by

$$L(\mu,\delta,\ \theta_1,\sigma_a^2|\underset{\sim}{z}^t) = (2\pi\sigma_a^2)^{-(n_1+n_2)/2}(\underset{\sim}{X}^t\underset{\sim}{X})^{-1/2}\exp\{-\sum_{t=0}^{n_1+n_2}[a_t]^2/2\sigma_a^2\}\ , \tag{150}$$

where $(\underset{\sim}{X}^t\underset{\sim}{X})$ is given in equation (138), $[\hat{a}_0] = a_*$ as given in equation (142); and $[a_t]$'s for $t=1,\ldots,\ n_1$ are given in equation (148) while, for $t=n_1+1,\ldots,\ n_1+n_2$, the $[a_t]$'s are given in equation (149).

3.3.2.2 Multi-Consequence $MA_I(1)$ Model

The multi-consequence $MA_I(1)$ model was presented in equation (84). This can be rewritten as

$$\left.\begin{aligned} a_t &= Z_t - \mu + \theta_1\ a_{t-1},\ t = 1,\ldots,\ n_1 \\ &= Z_t - \mu - \delta + \gamma_1\ a_{t-1},\ t = n_1+1,\ldots,\ n_1+n_2, \end{aligned}\right\} \tag{151}$$

where $a_t \sim NID(0,\sigma_a^2)$, $t=0,\ldots,\ n$. The joint distribution of $\underset{\sim}{a} = [a_0,\ a_1,\ldots,a_{n_1+n_2}]^t$ was given in equation (124), while the joint distribution of $\underset{\sim}{Z} = [Z_1,\ldots,Z_{n_1+n_2}]^t$ was given in equation (125) with the understanding that $\Sigma_{\underset{\sim}{Z}}^{(0,1)} = \sigma^2(M_n^{(0,1)})^{-1}$ is as presented in equation (87).

From the model presented in equation (151), we can write down the following (n+1) equations:

$$a_0 = a_0$$

$$a_1 = Z_1 - \mu + \theta_1 a_0$$

$$a_2 = Z_2 - \mu + \theta_1 a_1$$

$$\vdots$$

$$a_{n_1} = Z_{n_1} - \mu + \theta_1 a_{n_1-1}$$

$$a_{n_1+1} = Z_{n_1+1} - \mu - \delta + \gamma_1 a_{n_1}$$

$$a_{n_1+2} = Z_{n_1+2} - \mu - \delta + \gamma_1 a_{n_1+1}$$

$$\vdots$$

$$a_{n_1+n_2} = Z_{n_1+n_2} - \mu - \delta + \gamma_1 a_{n_1+n_2-1}$$

By successive substitution of a_1 for a_2 and so on, we can express $\underset{\sim}{a}$ in terms of $\underset{\sim}{Z}$ and $a_* = a_0$. Specifically,

$$a_0 = a_0$$

$$a_1 = Z_1 - \mu + \theta_1 a_0$$

$$a_2 = Z_2 + \theta_1 Z_1 - (1+\theta_1)\mu + \theta_1^2 a_0$$

$$\vdots$$

$$a_{n_1} = Z_{n_1} + \theta_1 Z_{n_1-1} + \ldots + \theta_1^{n_1-1} Z_1 - (1+\theta_1+\ldots+\theta_1^{n_1-1})\mu + \theta_1^{n_1} a_0$$

$$a_{n_1+1} = Z_{n_1+1} + \gamma_1 Z_{n_1} + \theta_1 \gamma_1 Z_{n_1-1} + \theta_1^2 \gamma_1 Z_{n_1-2} + \ldots + \theta_1^{n_1-1} \gamma_1 Z_1$$

$$- (1+\gamma_1+\theta_1\gamma_1+\ldots+\theta_1^{n_1-1}\gamma_1)\mu - \delta + \theta_1^{n_1} \gamma_1 a_0$$

$$a_{n_1+2} = Z_{n_1+2} + \gamma_1 Z_{n_1+1} + \gamma_1^2 Z_{n_1} + \gamma_1^2 \theta_1 Z_{n_1-1} + \theta_1^2 \gamma_1^2 Z_{n_1-2} + \ldots + \theta_1^{n_1-1} \gamma_1^2 Z_1$$

$$- (1+\gamma_1+\gamma_1^2+\theta_1\gamma_1^2 + \ldots + \theta_1^{n_1-1}\gamma_1^2)\mu - (1+\gamma_1^2)\delta + \theta_1^{2^{n_1}} \gamma_1^2 a_0$$

$$\vdots$$

$$\begin{aligned} a_{n_1+n_2} &= Z_{n_1+n_2}+\gamma_1 Z_{n_1+n_2-1}+\ldots+\gamma_1^{n_2-1} Z_{n_1+1}+\gamma_1^{n_2} Z_{n_1} \\ &+\theta_1\gamma_1^{n_2} Z_{n_1-1} + \theta_1^{2}\gamma_1^{n_2} Z_{n_1-2} +\ldots+ \theta_1^{n_1-1}\gamma_1^{n_2} Z_1 \\ &-(1+\gamma_1+\ldots+\gamma_1^{n_2}+\theta_1\gamma_1^{n_2}+\ldots+ \theta_1^{n_1-1}\gamma_1^{n_2})\mu \\ &-(1+\gamma_1+\ldots+\gamma_1^{n_2-1})\delta + \theta_1^{n_1}\gamma_1^{n_2} a_0 \end{aligned} \tag{152}$$

This system of (n_1+n_2+1) equations has the following matrix representation:

$$\underset{\sim}{a} = L\underset{\sim}{Z} + \underset{\sim}{X}\, a_* - \underset{\sim}{b}\mu - \underset{\sim}{c}\,\delta \ , \tag{153}$$

where

$$L = \left[\begin{array}{cccc|cccc} 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \theta_1 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \\ \theta_1^{n_1-1} & \theta_1^{n_1-2} & \ldots & 1 & 0 & 0 & \ldots & 0 \\ \hline \theta_1^{n_1-1}\gamma_1 & \theta_1^{n_1-2}\gamma_1 & \ldots & \gamma_1 & 1 & 0 & \ldots & 0 \\ \theta_1^{n_1-1}\gamma_1^2 & \theta_1^{n_1-2}\gamma_1^2 & \ldots & \gamma_1^2 & \gamma_1 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \theta_1^{n_1-1}\gamma_1^{n_2} & \theta_1^{n_1-2}\gamma_1^{n_2} & \ldots & \gamma_1^{n_2} & \gamma_1^{n_2-1} & \gamma_1^{n_2-2} & & 1 \end{array}\right]$$

$$
\underset{\sim}{X} = \begin{bmatrix} 1 \\ \theta_1 \\ \theta_1^2 \\ \vdots \\ \theta_1^{n_1} \\ \theta_1^{n_1}\ \gamma_1 \\ \theta_1^{n_1}\ \gamma_1^2 \\ \vdots \\ \theta_1^{n_1}\ \gamma_1^{n_2} \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} 0 \\ 1 \\ 1+\theta_1 \\ \vdots \\ 1+\theta_1+\ldots+\theta_1^{n_1-1} \\ 1+\gamma_1(1+\theta_1+\ldots+\theta_1^{n_1-1}) \\ 1+\ldots+\gamma_1^2(1+\theta_1+\ldots+\theta_1^{n_1-1}) \\ \vdots \\ 1+\ldots+\gamma_1^{n_2}(1+\theta_1+\ldots+\theta_1^{n_1-1}) \end{bmatrix}, \quad \underset{\sim}{c} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 1+\gamma_1 \\ \vdots \\ 1+\ldots+\gamma_1^{n_2-1} \end{bmatrix} \tag{154}
$$

In making the transformation $\underset{\sim}{a} = L^*[a_* \vdots \underset{\sim}{z}^t]^t$, where the $[(n_1+n_2+1) \times (n_1+n_2+1)]$ matrix $L^* = [\underset{\sim}{X} \vdots L]$, it is easily seen that $|J| = 1$. Bu substituting equation (153) into equation (124), we see that the joint distribution of $\underset{\sim}{Z}$ and a_* is

$$
f_{\underset{\sim}{Z}^t, a_*}(\underset{\sim}{z}^t,\ a_*;\ \mu,\delta,\ \theta_1,\gamma_1,\sigma_a^2) = (2\pi\sigma_a^2)^{-(n+1)/2}\ \exp\{-S(\theta_1,\gamma_1,a_*)/2\sigma_a^2\}, \tag{155}
$$

where

$$
S(\theta_1,\gamma_1,a_*) = (L\underset{\sim}{z}+\underset{\sim}{X}\ a_* - \underset{\sim}{b}\mu - \underset{\sim}{c}\delta)^t(L\underset{\sim}{z} + \underset{\sim}{X}\, a_* - \underset{\sim}{b}\mu - \underset{\sim}{c}\delta). \tag{156}
$$

For convenience, we let $\underset{\sim}{d} = \underset{\sim}{b}\mu + \underset{\sim}{c}\delta$.

Define $\hat{a}_*$ to be the value of a_* which minimizes $S(\theta_1,\gamma_1,a_*)$. By taking the derivative of $S(\theta_1,\gamma_1,a_*)$ with respect to a_* and setting this derivative equal to zero, we find that $\hat{a}_*$ is the solution to the following normal equation:

$$\underset{\sim}{X}^t \underset{\sim}{X}\, \hat{\underset{\sim}{a}}_* = - \underset{\sim}{X}^t\, L\underset{\sim}{z} + \underset{\sim}{X}^t\, \underset{\sim}{d}. \tag{157}$$

To solve this equation, note that

$$\underset{\sim}{X}^t \underset{\sim}{X} = (1-\theta_1^{2n_1})(1-\theta_1^2)^{-1} + \theta_1^{2n_1}(1-\gamma_1^{2(n_2+1)})(1-\gamma_1^2)^{-1} \ , \tag{158}$$

and, if ℓ_i denotes a general element of the $[1 \times (n_1+n_2)]$ matrix $\underset{\sim}{X}^t L$, then

$$\ell_i = \theta_1^{\,i}[(1-\theta_1^{2(n_1-i)})(1-\theta_1^2)^{-1} + \theta_1^{2(n_1-i)}(1-\gamma_1^{2(n_2+1)})(1-\gamma_1^2)^{-1}] \ , \tag{159a}$$

for $i=1,\ldots, n_1-1$,

$$\ell_{n_1} = \theta_1^{n_1}[(1-\gamma_1^{2(n_2+1)})(1-\gamma_1^2)^{-1}] \ , \tag{159b}$$

and

$$\ell_i = \theta_1^{n_1}\, \gamma_1^{i-n_1}[(1-\gamma_1^{2(n_1+n_2-i+1)})(1-\gamma_1^2)^{-1}] \ , \tag{159c}$$

for $i=n_1+1,\ldots, n_1+n_2$. Thus,

$$\underset{\sim}{X}^t\, L\underset{\sim}{z} = \sum_{i=1}^{n_1+n_2} \ell_i\, z_i \ . \tag{160}$$

Also note that

$$\begin{aligned}
\underset{\sim}{X}^t \underset{\sim}{d} = {} & \mu\theta_1(1-\theta_1)^{-1} \sum_{i=0}^{n_1-1} \theta_1^{\,i}(1-\theta_1^{\,i+1}) \\
& + \delta\theta_1^{n_1}\, \gamma_1(1-\gamma_1)^{-1} \sum_{i=0}^{n_2-1} \gamma_1^{\,i}\,(1-\gamma_1^{\,i+1}) \\
& + \mu\theta_1^{n_1}\, \gamma_1(1-\gamma_1)^{-1} \sum_{i=0}^{n_2-1} \gamma_1^{\,i}(1-\gamma_1^{\,i+1}) \\
& + \mu\theta_1^{n_1}\, \gamma_1(1-\theta_1^{n_1})(1-\theta_1)^{-1} \sum_{i=0}^{n_2-1} \gamma_1^{2i+1}
\end{aligned} \tag{161}$$

From equation (157), we see that

$$\hat{a}_* = (-\underset{\sim}{X}^t L\underset{\sim}{z} + \underset{\sim}{X}^t \underset{\sim}{d})/(\underset{\sim}{X}^t \underset{\sim}{X}) \tag{162}$$

where expressions for $\underset{\sim}{X}^t\underset{\sim}{X}$, $\underset{\sim}{X}^t L\underset{\sim}{z}$, and $\underset{\sim}{X}^t \underset{\sim}{d}$ are given in equations (158), (160), and (161), respectively. Note that when $\gamma_1 = \theta_1$, equations (158)-(161) are identical with equations (138)-(141) for the single consequence $MA_I(1)$ model.

By making use of equation (157), we see that $S(\theta_1,\gamma_1, a_*)$ can be rewritten as

$$S(\theta_1,\gamma_1, a_*)=\underbrace{[(L\underset{\sim}{z}+\underset{\sim}{X}\hat{a}_*)-\underset{\sim}{d}]^t[(L\underset{\sim}{z}+\underset{\sim}{X}\hat{a}_*)-\underset{\sim}{d}]}_{S(\theta_1,\gamma_1)} + (a_*-\hat{a}_*)^2 \underset{\sim}{X}^t \underset{\sim}{X}, \tag{163}$$

where $S(\theta_1,\gamma_1)$ is a function of the observations but not of a_*. Let $\underset{\sim}{\xi} = [\mu,\delta,\theta_1,\gamma_1,\sigma_a^2]^t$. Since

$$f_{\underset{\sim}{Z}^t,a_*}(\underset{\sim}{z}^t, a_*; \underset{\sim}{\xi}^t) = f_{\underset{\sim}{Z}^t}(\underset{\sim}{z}^t; \underset{\sim}{\xi}^t)f_{a_*|\underset{\sim}{Z}^t}(a_*|\underset{\sim}{z}^t; \underset{\sim}{\xi}^t) , \tag{164}$$

it follows from equations (155) and (163) that

$$f_{a_*|\underset{\sim}{Z}^t}(a_*|\underset{\sim}{z}^t; \underset{\sim}{\xi}^t) = (2\pi\sigma_a^2)^{-1/2}|\underset{\sim}{X}^t\underset{\sim}{X}|^{1/2} \exp\{-(a_*-\hat{a}_*)^2(X^tX)/2\sigma_a^2\} \tag{165}$$

and

$$f_{\underset{\sim}{Z}^t}(\underset{\sim}{z}^t; \underset{\sim}{\xi}^t) = (2\pi\sigma_a^2)^{-n/2}|\underset{\sim}{X}^t\underset{\sim}{X}|^{-1/2} \exp\{ -S(\theta_1,\gamma_1)/2\sigma_a^2\} , \tag{166}$$

where $S(\theta_1,\gamma_1)$ is given in equation (163).

As with the single consequence $MA_I(1)$ model, we deduce the following:

(i) $\hat{a}_*$ is the conditional expectation of a_* given $\underset{\sim}{z}$ and $\underset{\sim}{\xi}$. Also, $[a] = L\underset{\sim}{z} + \underset{\sim}{X}[a_*]-\underset{\sim}{d}$, where $[a_*]$ denotes $E(a_*|\underset{\sim}{z}^t,\underset{\sim}{\xi}^t)$. Thus,

$$S(\theta_1,\gamma_1) = \sum_{t=0}^{n_1+n_2} [a_t]^2 \quad . \tag{167}$$

(ii) $|\underset{\sim}{X}^t\underset{\sim}{X}|^{-1} = |M_n^{(0,1)}|$ and $S(\theta_1,\gamma_1) = (\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}})^t\, M_n^{(0,1)} (\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}})$.

(iii) In order to compute

$$S(\theta_1,\gamma_1) = \sum_{t=0}^{n_1+n_2} [a_t]^2 \quad ,$$

we let $[a_0] = \hat{a}_*$ and recursively calculate the first $n_1 [a_t]$'s from

$$[a_t] = z_t - \hat{\mu} + \theta_1 [a_{t-1}], \tag{168}$$

for $t=1,\ldots, n_1$, while the recursive relationship for the last $n_2 [a_t]$'s is given by

$$[a_t] = z_t - \hat{\mu} - \hat{\delta} + \gamma_1 [a_{t-1}] \quad , \tag{169}$$

for $t=n_1+1,\ldots, n_1+n_2$.

These results are stated in the following theorem.

Theorem 3.2: For the multi-consequence $MA_I(1)$ model, the unconditional likelihood function is given by

$$L(\mu,\delta,\theta_1,\gamma_1,\sigma_a^2|\underset{\sim}{z}^t)=(2\pi\sigma_a^2)^{-(n_1+n_2)/2} (\underset{\sim}{X}^t\underset{\sim}{X})^{-1/2} \exp\{- \sum_{t=0}^{n_1+n_2} [a_t]^2/2\sigma_a^2 \}, \tag{170}$$

where $[\underset{\sim}{X}^t\underset{\sim}{X}]$ is given in equation (158), $[a_0] = \hat{a}_*$ as given in equation (162), and $[a_t]$'s for $t=1,\ldots, n_1+n_2$ are given in equations (168) and (169). Since $\underset{\sim}{X}^t\underset{\sim}{X}$ is a scalar, the determinant symbol has been omitted.

3.3.2.3 Single and Multi-Consequence $MA_I(2)$ Models

Because of the rather complicated mathematical expressions that arise in trying to formulate the likelihood function for both the single and multi-consequence $MA_I(2)$ models, we will consider in detail only the single consequence $MA_I(2)$ model. The extension to the multi-consequence $MA_I(2)$ model is tedious but straightforward.

The single consequence $MA_I(2)$ model, which was presented in equation (89), can be rewritten as

$$\left.\begin{aligned} a_t &= Z_t - \mu + \theta_1\, a_{t-1} + \theta_2\, a_{t-2},\ t=1,\ldots,\ n_1 \\ &= Z_t - \mu - \delta + \theta_1\, a_{t-1} + \theta_2\, a_{t-2},\ t = n_1+1,\ldots,\ n_1+n_2, \end{aligned}\right\} \quad (171)$$

where $a_t \sim NID(0,\sigma_a^2)$ for $t = -1,0,1,\ldots,\ n_1+n_2$. Thus, the joint distribution of $\underset{\sim}{a} = [a_{-1},\ a_0,\ a_1,\ldots,\ a_{n_1+n_2}]^t$ equals

$$f(\underset{\sim}{a}^t;\ \sigma_a^2) = (2\pi\sigma_a^2)^{-(n+2)/2} \exp\{-\underset{\sim}{a}^t\ \underset{\sim}{a}/2\sigma_a^2\} \quad . \qquad (172)$$

The joint distribution of $\underset{\sim}{Z} = [Z_1,\ldots,\ Z_{n_1+n_2}]^t$ is n-variate normal since $\underset{\sim}{Z} = C_I^{(0,2)}\ \underset{\sim}{a} + \underset{\sim}{\mu}_{\underset{\sim}{Z}}$. Namely,

$$f_{\underset{\sim}{Z}^t}(\underset{\sim}{z}^t;\mu,\delta,\ \theta_1,\theta_2,\sigma_a^2)=(2\pi\sigma_a^2)^{-n/2}|M_n^{(0,2)}|^{1/2} \exp\{-(\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}})^t M_n^{(0,2)} (\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}})/2\sigma_a^2\}, \qquad (173)$$

where $\underset{\sim}{\mu}_{\underset{\sim}{Z}}$ is given in equation (91) and $\Sigma_{\underset{\sim}{Z}}^{(0,2)} = \sigma_a^2(M_n^{(0,2)})^{-1}$ has the same structure as $\Sigma_{\underset{\sim}{Z}}^{(0,2)}$ presented in equation (92) with the exception that γ_i needs to be replaced by θ_i for $i=1,2$.

From the model presented in equation (171), we can write down the following (n+2) equations:

$$\begin{aligned} &a_{-1} = a_{-1} \\ &a_0 = a_0 \\ &\overline{} \\ &a_1 = Z_1 - \mu + \theta_1 a_0 + \theta_2 a_{-1} \\ &a_2 = Z_2 - \mu + \theta_1 a_1 + \theta_2 a_0 \\ &\quad\vdots \end{aligned}$$

$$a_{n_1}=Z_{n_1}-\mu+\theta_1 a_{n_1-1}+\theta_2 a_{n_1-2}$$

— — — — — — — — — — — —

$$a_{n_1+1}=Z_{n_1+1}-\mu-\delta+\theta_1 a_{n_1}+\theta_2 a_{n_1-1}$$

$$a_{n_1+2}=Z_{n_1+2}-\mu-\delta+\theta_1 a_{n_1+1}+\theta_2 a_{n_1}$$

$$\vdots$$

$$a_{n_1+n_2}=Z_{n_1+n_2}-\mu-\delta+\theta_1 a_{n_1+n_2-1}+\theta_2 a_{n_1+n_2-2} \tag{174}$$

We now attempt to express $\underset{\sim}{a}$ in terms of $\underset{\sim}{Z}$ and $\underset{\sim}{a}_* = [a_{-1}, a_0]^t$. The first three equations are easily obtained:

$$a_{-1}=a_{-1}$$

$$a_0=a_0$$

$$a_1=Z_1-\mu+\theta_1 a_0+\theta_2 a_{-1}$$

By substituting this last expression for a_1 into the equation for a_2 in (174), we obtain

$$a_2=Z_2+\theta_1 Z_1-(1+\theta_1)\mu+(\theta_1^2+\theta_2)a_0+(\theta_1\theta_2)a_{-1} \quad .$$

Furthermore, by substituting the last two expressions for a_1 and a_2 into the equation for a_3 in (174), we obtain

$$a_3=Z_3+\theta_1 Z_2+(\theta_1^2+\theta_2)Z_1-(1+\theta_1+\theta_1^2+\theta_2)\mu$$
$$+ (\theta_1^3+2\theta_1\theta_2)a_0 + (\theta_1^2\theta_2+\theta_2^2)a_{-1} \quad .$$

By continuing this substitution scheme, we obtain the following expressions for a_4, a_5, a_6, and a_7:

$$a_4 = Z_4+\theta_1 Z_3+(\theta_1^2+\theta_2)Z_2+(\theta_1^3+2\theta_1\theta_2)Z_1-(1+\theta_1+\theta_1^2+\theta_1^3+\theta_2+2\theta_1\theta_2)\mu$$
$$+(\theta_1^4+3\theta_1^2\theta_2+\theta_2^2)a_0 + (\theta_1^3\theta_2+2\theta_1\theta_2^2)a_{-1}$$

$$a_5 = Z_5+\theta_1 Z_4+(\theta_1^2+\theta_2)Z_3+(\theta_1^3+2\theta_1\theta_2)Z_2+(\theta_1^4+3\theta_1^2\theta_2+\theta_2^2)Z_1$$
$$-(1+\theta_1+\theta_1^2+\theta_1^3+\theta_1^4+\theta_2+\theta_2^2+2\theta_1\theta_2+3\theta_1^2\theta_2)\mu$$
$$+(\theta_1^5+4\theta_1^3\theta_2+3\theta_1\theta_2^2)a_0+(\theta_1^4\theta_2+3\theta_1^2\theta_2^2+\theta_2^3)a_{-1}$$

$$a_6 = Z_6+\theta_1 Z_5+(\theta_1^2+\theta_2)Z_4+(\theta_1^3+2\theta_1\theta_2)Z_3+(\theta_1^4+3\theta_1^2\theta_2+\theta_2^2)Z_2+(\theta_1^5+4\theta_1^3\theta_2+3\theta_1\theta_2^2)Z_1$$
$$-(1+\theta_1+\ldots+\theta_1^5+\theta_2+\theta_2^2+2\theta_1\theta_2+3\theta_1\theta_2^2+4\theta_1^3\theta_2)\mu$$
$$+(\theta_1^6+5\theta_1^4\theta_2+6\theta_1^2\theta_2^2+\theta_2^3)a_0+(\theta_1^5\theta_2+4\theta_1^3\theta_2^2+3\theta_1\theta_2^3)a_{-1}$$

$$a_7 = Z_7+\theta_1 Z_6+(\theta_1^2+\theta_2)Z_5+(\theta_1^3+2\theta_1\theta_2)Z_4+(\theta_1^4+3\theta_1^2\theta_2+\theta_2^2)Z_3+(\theta_1^5+4\theta_1^3\theta_2+3\theta_1\theta_2^2)Z_2$$
$$+ (\theta_1^6+5\theta_1^4\theta_2+6\theta_1^2\theta_2^2+\theta_2^3)Z_1$$
$$-(1+\theta_1+\ldots+\theta_1^6+\theta_2+\theta_2^2+\theta_2^3+2\theta_1\theta_2+3\theta_1\theta_2^2+3\theta_1^2\theta_2+4\theta_1^3\theta_2+5\theta_1^4\theta_2+6\theta_1^2\theta_2^2)\mu$$
$$+(\theta_1^7+6\theta_1^5\theta_2+10\theta_1^3\theta_2^2+4\theta_1\theta_2^3)a_0+(\theta_1^6\theta_2+5\theta_1^4\theta_2^2+6\theta_1^2\theta_2^3+\theta_2^4)a_{-1}$$

At first glance, no discernible pattern is evident for the a_t's in terms of $\underset{\sim}{Z}$ and $\underset{\sim}{a}_*$. However, a recursive relationship for the elements of L, X, $\underset{\sim}{b}$, and $\underset{\sim}{c}$ is obtainable where $\underset{\sim}{a} = L\underset{\sim}{Z} + X\,\underset{\sim}{a}_* - \underset{\sim}{b}\mu = \underset{\sim}{c}\delta$. For the single consequence $MA_I(2)$ model, L is an $[(n_1+n_2+2)x(n_1+n_2)]$ matrix, X is an $[(n_1+n_2+2)x2]$ matrix, and $\underset{\sim}{b}$ and $\underset{\sim}{c}$ are both $[(n_1+n_2+2)x1]$ vectors.

Now L can be partitioned as follows:

$$L = \begin{bmatrix} 0_{2xn} \\ \hline L^* \end{bmatrix} ,$$

where 0_{2xn} is a$[2x(n_1+n_2)]$ matrix all of whse entries are zero while $L^* = [\ell^*_{ij}]$ is a$[(n_1+n_2)x(n_1+n_2)]$ lower triangular matrix with

$$\left.\begin{aligned} &\ell^*_{ii} = 1,\ i = 1,\ldots,\ n \\ &\ell^*_{i+1,i} = \ell^*_{i,i-1}\ ,\ i = 2,\ldots,\ n-1 \\ &\ell^*_{i+1,i-1} = \ell^*_{i,i-2}\ ,\ i = 3,\ldots,\ n-1 \\ &\qquad\vdots \\ &\ell^*_{n,2} = \ell^*_{n-1,1} \end{aligned}\right\} \qquad (175)$$

Equation (175) essentially states that the elements in the main diagonal of L^* are each equal to one, the elements in the first subdiagonal are equal to each other, the elements in the second subdiagonal are equal to each other, etc. To utilize the relationships expressed in (175), we generate the elements in the first column of L^* by the recursive relationship $\ell^*_{i,1} - \theta_1\ell^*_{i-1,1}+\theta_2\ \ell^*_{i-2,1}$, $i = 3,\ldots,\ n_1$ with the initial conditions $\ell^*_{1,1} = 1$ and $\ell^*_{2,1} = \theta_1$.

The (nx2) matrix X can be partitioned as follows:

$$X = \begin{bmatrix} I_2 \\ \hline X^* \end{bmatrix}$$

where I_2 is the (2x2) identity matrix and $X^* = [x^*_{ij}]$ is a $[(n_1+n_2)x2]$ matrix whose second column elements satisfy the recursive relationship $x^*_{i,2} = \theta_1 x^*_{i-1,2} + \theta_2 x^*_{i-2,2}$ for $i = 2,3,\ldots, n$ with the initial conditions $x^*_{0,2} = 1$ and $x^*_{1,2} = \theta_1$. The elements in the first column of X^* are such that $x^*_{i,1} = \theta_2 x^*_{i-1,2}$, for $i = 2,\ldots, n$ with $x^*_{1,1} = \theta_2$.

The $[(n+2)x1]$ $\underset{\sim}{b}$ and $\underset{\sim}{c}$ vectors can also be partitioned:

$$\underset{\sim}{b} = \begin{bmatrix} \underset{\sim}{0}_2 \\ \hline \underset{\sim}{b}^* \end{bmatrix} \quad , \qquad \underset{\sim}{c} = \begin{bmatrix} \underset{\sim}{0}_{n_1+2} \\ \hline \underset{\sim}{c}^* \end{bmatrix}$$

where $\underset{\sim}{b}^* = [b^*_i]$ is an (nx1) vector and $\underset{\sim}{c}^* = [c^*_i]$ is an (n_2x1) vector. The b^*_i's satisfy the recursive relationship $b^*_i = 1 + \theta_1 b^*_{i-1} + \theta_2 b^*_{i-2}$, for $i = 2,\ldots, n$, with the initial conditions that $b^*_0 = 0$ and $b^*_1 = 1$; and, the c^*_i's are such that $c^*_i = 1 + \theta_1 c^*_{i-1} + \theta_2 c^*_{i-2}$, for $i = 2,\ldots, n_2$, with $c^*_0 = 0$ and $c^*_1 = 1$. Thus $\underset{\sim}{a}$ can be expressed in terms of $\underset{\sim}{Z}$ and $\underset{\sim}{a}_*$ as follows:

$$\underset{\sim}{a} = L\underset{\sim}{Z} + X \underset{\sim}{a}_* - \underset{\sim}{b}\mu - \underset{\sim}{c}\delta , \qquad (176)$$

where recursive relationships have been presented for determining the elements of L, X, $\underset{\sim}{b}$, and $\underset{\sim}{c}$.

In making the transformation $\underset{\sim}{a} = L^*[\underset{\sim}{a}^t_* \mid \underset{\sim}{z}^t]^t$, where the $[(n_1+n_2+2)]$ matrix $L^* = [\ X \mid L\ ]$, it is easily seen that $|J| = 1$. By substituting equation (176) into equation (172), we see that the joint distribution of $\underset{\sim}{Z}$ and $\underset{\sim}{a}^*$ is

$$f_{\underset{\sim}{Z}^t,\underset{\sim}{a}^t_*}(\underset{\sim}{z}^t,\ \underset{\sim}{a}^t_*;\ \mu,\delta,\ \theta_1,\theta_2,\sigma^2_a)=(2\pi\sigma^2_a)^{-(n+2)/2} \exp\{-S(\theta_1,\theta_2,\ \underset{\sim}{a}_*)/2\sigma^2_a\} , \qquad (177)$$

where

$$S(\theta_1,\theta_2, \underset{\sim}{a}_*) = (L\underset{\sim}{z} + X\, \underset{\sim}{a}^* - \underset{\sim}{b}\mu - \underset{\sim}{c}\delta)^t(L\underset{\sim}{z} + X\, \underset{\sim}{a}^* - \underset{\sim}{b}\mu - \underset{\sim}{c}\delta). \quad (178)$$

For convenience, we let $\underset{\sim}{d} = \underset{\sim}{b}\mu + \underset{\sim}{c}\delta$.

Let $\hat{\underset{\sim}{a}}_*$ be the vector of values which minimizes $S(\theta_1,\theta_2, \underset{\sim}{a}_*)$. By taking the derivative of $S(\theta_1,\theta_2, \underset{\sim}{a}_*)$ with respect to $\underset{\sim}{a}_*$ and setting the resulting system of equations equal to the zero vector, we find that $\hat{\underset{\sim}{a}}_*$ is the solution to the following normal equations:

$$(X^tX)\, \hat{\underset{\sim}{a}}_* = - X^t\, L\underset{\sim}{z} + X^t\, \underset{\sim}{d}. \quad (179)$$

Since (X^tX) is nonsingular, we see that

$$\hat{\underset{\sim}{a}}_* = (X^tX)^{-1}\, (- X^t\, L\underset{\sim}{z} + X^t\underset{\sim}{d}). \quad (180)$$

By making use of equation (179), we find that $S(\theta_1,\theta_2, \underset{\sim}{a}_*)$ can be rewritten as

$$S(\theta_1,\theta_2, \underset{\sim}{a}_*)=\underbrace{[(L\underset{\sim}{z}+X\, \hat{\underset{\sim}{a}}_*)-\underset{\sim}{d}]^t[(L\underset{\sim}{z} + X\, \hat{\underset{\sim}{a}}_*)-\underset{\sim}{d}]}_{S(\theta_1,\theta_2)}+(\hat{\underset{\sim}{a}}_*-\underset{\sim}{a}_*)^tX^tX(\hat{\underset{\sim}{a}}_*-\underset{\sim}{a}_*) \quad , \quad (181)$$

where $S(\theta_1,\theta_2)$ is a function of the observations but not of $\underset{\sim}{a}_*$. Let $\underset{\sim}{\xi} = [\mu,\delta,\theta_1,\theta_2,\sigma_a^2]^t$. Since

$$f_{\underset{\sim}{Z}^t,\underset{\sim}{a}_*^t}(\underset{\sim}{z}^t, \underset{\sim}{a}_*^t; \underset{\sim}{\xi}^t) = f_{\underset{\sim}{Z}^t}(\underset{\sim}{z}^t; \underset{\sim}{\xi}^t)f_{\underset{\sim}{a}_*^t|\underset{\sim}{Z}^t}(\underset{\sim}{a}_*^t\, |\underset{\sim}{z}^t\; ; \underset{\sim}{\xi}^t) \quad , \quad (182)$$

it follows from equations (177) and (181) that

$$f_{a_*^t|\underset{\sim}{Z}^t}(\underset{\sim}{a}_*^t|\underset{\sim}{z}^t;\underset{\sim}{\xi}^t) = (2\pi\sigma_a^2)^{-1}|X^tX|^{1/2}\, \exp\{-(\underset{\sim}{a}_*-\hat{\underset{\sim}{a}}_*)^t(X^tX)(\underset{\sim}{a}_*-\hat{\underset{\sim}{a}}_*)/2\sigma_a^2\}$$
$$(183)$$

and

$$f_{\underset{\sim}{Z}^t}(\underset{\sim}{z}^t;\ \underset{\sim}{\xi}^t) = (2\pi\sigma_a^2)^{-n/2}|X^tX|^{-1/2}\exp\{-S(\theta_1,\theta_2)/2\sigma_a^2\}, \tag{184}$$

where $S(\theta_1,\theta_2)$ is given in equation (181).

Based on the foregoing statements, we can make the following deductions:

(i) From equation (183), we see that $\hat{\underset{\sim}{a}}_*$ is the conditional expectation of $\underset{\sim}{a}_*$ given $\underset{\sim}{z}$ and $\underset{\sim}{\xi}$. As with the $MA_I(1)$ models, let $[\underset{\sim}{a}_*]$ denote $E(\underset{\sim}{a}_*|\underset{\sim}{z},\ \underset{\sim}{\xi})$. Thus, $\hat{\underset{\sim}{a}}_* = [\underset{\sim}{a}_*]$, $[\underset{\sim}{a}] = L\underset{\sim}{z} + X[\underset{\sim}{a}_*]-\underset{\sim}{d}$, and

$$S(\theta_1,\theta_2) = \sum_{t=-1}^{n_1+n_2} [a_t]^2$$

(ii) $|X^tX|^{-1} = |M_n^{(0,2)}|$ and $S(\theta_1,\theta_2)=(\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}})^t\, M_n^{(0,2)}(\underset{\sim}{z}-\underset{\sim}{\mu}_{\underset{\sim}{Z}})$.

(iii) In order to compute $S(\theta_1,\theta_2) = \sum_{t=-1}^{n_1+n_2} [a_t]^2$, we let $[a_{-1},a_0]^t = \hat{\underset{\sim}{a}}_*$ and recursively calculate the first $n_1[a_t]$'s from

$$[a_t] = z_t - \hat{\mu} + \theta_1[a_{t-1}]+\theta_2[a_{t-2}] \quad , \tag{185}$$

for $t=1,\ldots,\ n_1$. The recursive relationship for the last $n_2[a_t]$'s is given by

$$[a_t] = z_t - \hat{\mu} - \hat{\delta} + \theta_1[a_{t-1}]+\theta_2[a_{t-2}] \quad , \tag{186}$$

for $t=n_1+1,\ldots,\ n_1+n_2$.

These results are stated in the following theorem.

<u>Theorem 3.3:</u> For the single consequence $MA_I(2)$ model, the unconditional

likelihood is given by

$$L(\mu,\delta,\theta_1,\theta_2,\sigma_a^2|\underset{\sim}{z})=(2\pi\sigma_a^2)^{-(n_1+n_2)/2}|X^tX|^{-1/2}\exp\{-\sum_{t=-1}^{n_1+n_2}[a_t]^2/2\sigma_a^2\} \tag{187}$$

Since (X^tX) is a (2x2) matrix, its determinant is easily evaluated once the elements of X have been recursively generated. The elements of $\hat{\underset{\sim}{a}}_* = [\underset{\sim}{a}_*]$ are also easily obtained.

In finding a facile computational form of the likelihood function for the $MA_I(2)$ model and higher order $MA_I(q)$ models, it appears that the main difficulty is in finding the elements of the matrices L and X and the vectors $\underset{\sim}{b}$ and $\underset{\sim}{c}$ which occur in the transformation $\underset{\sim}{a} = L\underset{\sim}{z} + X\,\underset{\sim}{a}_* - \underset{\sim}{b}\mu - \underset{\sim}{c}\delta$. Actually, upon inspecting Theorems 3.1-3.3, we see that only a function of the X matrix, (X^tX), explicitly appears, while L, X, $\underset{\sim}{b}$, and $\underset{\sim}{c}$ are implicitly used in calculating $\hat{\underset{\sim}{a}}_*$. However, estimates of a_0, $a_{-1},\ldots,$ a_{1-q} can be obtained by using a back-forecasting procedure outlined by Box and Jenkins [13]. Even though this approximation introduces "a transient into the system," its effect will "almost certainly be negligible by the time the beginning of the series is reached and thus will not affect the calculation of the a's." Thus, it is only the elements of the X matrix for which a recursive relationship needs to be determined, such as was done for the single consequence $MA_I(2)$ model. The recursive relationship for both $MA_I(1)$ models is obvious. Since it may be difficult to establish a recursive relationship to generate the elements of the X matrix for higher-order $MA_I(q)$ models, Box and Jenkins omit the $|X^tX|^{-1/2}$ term from the likelihood function and actually find unconditional least squares estimates by minimizing $\sum_{t=1-q}^{n}[a_t]^2$. These estimates are unconditional in the sense that estimates are obtained for the elements of $\underset{\sim}{a}_* = [a_0, a_{-1}, \ldots, a_{1-q}]$ rather than setting $\underset{\sim}{a}_* = \underset{\sim}{0}$. Box and Jenkins justify the omission of $|X^tX|^{-1/2}$ by stating that is of importance only for small n. However, there seems to be some disagreement on this point, and this is reported by Dent [17]. Furthermore, in intervention studies, the data bases are usually not that large and thus $|X^tX|^{-1/2}$ may play a major role.

3.3.3 Implementing the MLE Procedure

In Theorems 3.1-3.3, a computational form of the likelihood function was given for the $MA_I(1)$ and $MA_I(2)$ models when the parameter values are fixed. In this section, we discuss the finer points of implementing the computations with particular emphasis given to the multi-consequence $MA_I(1)$ model.

Section 3.3.1 focused on the MLE of μ and δ for a fixed set of moving-average parameters. Section 3.3.2 set forth a relatively easy way of evaluating the likelihood function where the case of computation was directed towards varying the moving-average parameters. The problem still remains of finding $\hat{\underset{\sim}{\xi}}$ which maximizes $L(\underset{\sim}{\xi}^t|\underset{\sim}{z}^t)$ where, for the multi-consequence $MA_I(1)$ model, $\underset{\sim}{\xi} = [\mu, \delta, \theta_1, \gamma_1, \sigma_a^2]^t$. Thus we wish to $\max\limits_{\xi} L(\underset{\sim}{\xi}^t|\underset{\sim}{z}^t)$. Now this maximization problem can be decomposed as follows:

$$\max_{\underset{\sim}{\xi}} L(\underset{\sim}{\xi}^t|\underset{\sim}{z}^t) = \max_{\theta_1,\gamma_1,\mu,\delta} [\max_{\sigma_a^2} L(\mu,\delta,\theta_1,\gamma_1,\sigma_a^2|\underset{\sim}{z}^t]$$

$$= \max_{\theta_1,\gamma_1} \{\max_{\mu,\delta}[\max_{\sigma_a^2} L(\mu,\delta,\theta_1,\gamma_1,\sigma_a^2|\ \underset{\sim}{z}^t)]\}.$$

Up to now, no particular reference has been made concerning the maximization of L with respect to σ_a^2. Taking the logarithm of equation (170), we find

$$\ln L(\underset{\sim}{\xi}^t|\underset{\sim}{z}^t) = -(n/2)\ln(2\pi)-(n/2)\ln(\sigma_a^2)-(1/2)(\underset{\sim}{X}^t\underset{\sim}{X}) -\{\sum_{t=0}^{n} [a_t]^2/2\sigma_a^2\}, \tag{188}$$

and

$$\partial \ln L/\partial\sigma_a^2 = - n/(2\sigma_a^2) + \sum_{t=0}^{n} [a_t]^2/(2(\sigma_a^2)^2) .$$

Setting this derivative equal to zero, we find that

$$\hat{\sigma}_a^2 = \sum_{t=0}^{n_1+n_2} [a_t]^2/(n_1+n_2) . \tag{189}$$

Thus, $\hat{\sigma}_a^2$ as presented in equation (189) is the maximum likelihood estimate of σ_a^2 for fixed μ, δ, θ_1, and γ_1. By making use of equation (189) in equation (170), we find that

$$\begin{aligned}\max_{\underset{\sim}{\xi}} L(\underset{\sim}{\xi}^t|\underset{\sim}{z}^t) &= \max_{\theta_1,\gamma_1,\mu,\delta} \{\max_{\sigma_a^2} L(\mu,\delta,\theta_1,\gamma_1,\sigma_a^2)\} \\ &= \max_{\theta_1,\gamma_1,\mu,\delta} (2\pi\hat{\sigma}_a^2)^{-(n_1+n_2)/2} (\underset{\sim}{X}^t\underset{\sim}{X})^{-1/2} \exp\{ - \sum_{t=0}^{n_1+n_2} [a_t]^2/2\hat{\sigma}_a^2\} \\ &= \max_{\theta_1,\gamma_1,\mu,\delta} (2\pi)^{-(n_1+n_2)/2} (\hat{\sigma}_a^2)^{-(n_1+n_2)/2} (\underset{\sim}{X}^t\underset{\sim}{X})^{-1/2}\exp\{-n/2\} .\end{aligned}$$

This last expression is equivalent to

$$\max_{\theta_1,\gamma_1,\mu,\delta} (\hat{\sigma}_a^2)^{-(n_1+n_2)/2} (\underset{\sim}{X}^t\underset{\sim}{X})^{-1/2} , \tag{190}$$

since $(2\pi)^{-n/2}$ and $e^{-n/2}$ are constants. By substituting equation (189) into (190), we can rewrite this as

$$\max_{\theta_1,\gamma_1,\mu,\delta} \quad \{\sum_{t=0}^{n_1+n_2} [a_t]^2/(n_1+n_2)\}^{-(n_1+n_2)/2} \; (\underset{\sim}{X}^t\underset{\sim}{X})^{-1/2} \quad .$$

In turn, this is equivalent to

$$\min_{\theta_1,\gamma_1,\mu,\delta} \quad \{ \sum_{t=0}^{n_1+n_2} [a_t]^2/(n_1+n_2)\}^{(n_1+n_2)/2} (\underset{\sim}{X}^t\underset{\sim}{X})^{1/2}$$

$$= \min_{\theta_1,\gamma_1} \quad \{ \min_{\mu,\delta} \; [\sum_{t=0}^{n_1+n_2} [a_t]^2/(n_1+n_2)]^{(n_1+n_2)/2} (\underset{\sim}{X}^t\underset{\sim}{X})^{1/2} \} \qquad (191)$$

Equation (191) clearly points out the difference between unconditional least squares (UCLS) estimation and maximum likelihood estimation. In UCLSE, one wishes to

$$\min_{\theta_1,\gamma_1,\mu,\delta} \quad \sum_{t=0}^{n_1+n_2} [a_t]^2 \quad ,$$

which is equivalent to

$$\min_{\theta_1,\gamma_1,\mu,\delta} \quad \{ \sum_{t=0}^{n_1+n_2} [a_t]^2/(n_1+n_2) \}^{(n_1+n_2)/2} \qquad (192)$$

Thus, UCLS estimation differs from ML estimation by the multiplicative effect of $(\underset{\sim}{X}^t\underset{\sim}{X})^{1/2}$.

Once that 4-tuple $(\hat{\mu},\hat{\delta},\hat{\theta}_1,\hat{\gamma}_1)$ is found which satisfies equation (191), $\hat{\sigma}_a^2$ is then found from equation (189).

The most difficult part of satisfying equation (191) is in finding $\hat{\mu}$ and $\hat{\delta}$ since this involves finding $M_n^{(0,1)}$, where $(M_n^{(0,1)})^{-1} = \Sigma_Z^{(0,1)}/\sigma_a^2$. Thus, for each (θ_1,γ_1) pair, it becomes necessary to compute another inverse. For a relatively large time series, $n_1 = n_2 = 300$, this exceeds the capacity of core storage. However, simplifications occur by making use of the patterned structure of $\Sigma_Z^{(0,1)}$ presented in equation (88). For notational convenience, we temporarily omit the (0,1) superscript. Thus,

$$\Sigma_{\underset{\sim}{Z}} = \sigma_a^2 \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right] = \sigma_a^2 \; M^{-1} .$$

It is well-known, e.g., see Anderson [10], that, provided the various inverses exist,

$$M = \begin{bmatrix} B_{11}^{-1} + B_{11} B_{12} C^{-1} B_{21} B_{11}^{-1} & -B_{11}^{-1} B_{12} C^{-1} \\ -C^{-1} B_{21} B_{11}^{-1} & C^{-1} \end{bmatrix} \tag{193}$$

where $C = B_{22} - B_{21} B_{11}^{-1} B_{12}$. Since $\Sigma_{\underset{\sim}{Z}}$ is positive definite, B_{11}^{-1} and C^{-1} exist. M could have also been given in a form which involves finding B_{22}^{-1}. For our specific problem, B_{11} and B_{22} are both tridiagonal matrices and thus further simplification results. Let $(B_{11}^{-1})_{i,j}$ denote the $(i,j)^{th}$ element of B_{11}^{-1}. Abraham and Weiss [1] show that

$$(B_{11}^{-1})_{i,j} = \frac{u(v^{2j}-1)}{v^{i+j}} \times \frac{[v^{2(n_1+1)} - v^{2i}]}{1 - v^{2(n_1+1)}} , \quad i \geq j$$

where $u = (1-\theta_1^2)^{-1}$ and $v = \theta_1$. Thus,

$$(B_{11}^{-1})_{i,j} = \frac{(1-\theta_1^{2j})}{(1-\theta_1^2)\theta_1^{i+j}} \times \frac{[\theta_1^{2i} - \theta_1^{2(n_1+1)}]}{1-\theta_1^{2(n_1+1)}}, \quad i \geq j = 1,\ldots, n_1 \tag{194}$$

An equivalent result was later given by Shaman [58]. In intervention studies, since n_1 is usually much larger than n_2, the significance of equation (194) is obvious.

From equation (193), we see that the only inverse which remains to be computed is $C^{-1} = (B_{22} - B_{21} B_{11}^{-1} B_{12})^{-1}$. Now $B_{21} B_{11}^{-1} B_{12} = B_{12}^t B_{11}^{-1} B_{12}$ is an $(n_2 x n_2)$ matrix whose entries are all zero except for the (1,1) entry, which is of the form $\gamma_1^2 (B_{11}^{-1})_{n_1,n_1}$. Using Abraham and Weiss' formula, we see that

$$\gamma_1^2 (B_{11}^{-1})_{n_1,n_1} = \gamma_1^2(1-\theta_1^{2n_1})/[1-\theta_1^{2(n_1+1)}]$$

Thus, C is also a tridiagonal matrix with

$$c_{11} = 1 + \gamma_1^2 - \gamma_1^2(1-\theta_1^{2n_1}) [1-\theta_1^{2(n_1+1)}]^{-1}$$

and the following general pattern:

$$C = \left[\begin{array}{c|ccccc} c_{11} & -\gamma_1 & 0 & 0 & \ldots & 0 \\ -\gamma_1 & (1+\gamma_1^2) & -\gamma_1 & 0 & \ldots & 0 \\ \hline 0 & -\gamma_1 & (1+\gamma_1^2) & -\gamma_1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1+\gamma_1^2 \end{array}\right] = \begin{bmatrix} c_{11} & \underset{\sim}{c}_{21}^t \\ \underset{\sim}{c}_{21} & C_{22} \end{bmatrix}$$

Now, assuming the various inverses exist,

$$C^{-1} = \begin{bmatrix} D^{-1} & -D^{-1} \underset{\sim}{c}_{21}^t C_{22}^{-1} \\ -C_{22}^{-1} \underset{\sim}{c}_{21} D^{-1} & C_{22}^{-1} + C_{22}^{-1} \underset{\sim}{c}_{21} D^{-1} \underset{\sim}{c}_{21}^t C_{22}^{-1} \end{bmatrix}, \tag{195}$$

where the (1x1) scalar $D = \underset{\sim}{c}_{11} - \underset{\sim}{c}_{21}^{t} C_{22}^{-1} \underset{\sim}{c}_{21}$. But C_{22} is an $[(n_2-1) \times (n_2-1)]$ tridiagonal matrix whose inverse elements are given by

$$(C_{22}^{-1})_{i,j} = \frac{(1-\gamma_1^{2j})}{(1-\gamma_1^2)\,\gamma_1^{i+j}} \times \frac{[\gamma_1^{2i} - \gamma_1^{2(n_1+1)}]}{1-\gamma_1^{2(n_1+1)}}, \quad i \geq j.$$

Thus, finding M has been reduced to finding B_{11}^{-1} and C_{22}^{-1}, where closed form expressions exist for generating the elements of these inverses.

Appendix D contains a listing of the computer program MLE MAI(1) designed to find the maximum likelihood estimators of the multi-consequence $MA_I(1)$ model.

3.3.4 Additional Statistical Inference

Although previous sections have discussed the determination of point estimates of the model parameters via the method of maximum likelihood, there are several inferential aspects that remain unanswered. For example, is the estimate of the shift statistically significant? Furthermore, in the multi-consequence intervention model, are the pre-intervention moving average parameters significantly different from the post-intervention parameters? We will answer these questions by specifically addressing the multi-consequence $MA_I(1)$ intervention model.

In the intervention models, it is the estimation of δ that is of prime importance. However, as can be seen from equation (130), $\hat{\delta}$ is contingent upon $M_n^{(0,1)}$ whose elements are θ_1 and γ_1. And, it remains to test

$$H_0:\ \theta_1 = \gamma_1 \quad \text{vs.} \quad H_1:\ \theta_1 \neq \gamma_1. \tag{196}$$

This problem is of sufficient importance in its own right without considering its influence on the estimation of δ. For, if the alternative hypothesis is true, the intervention treatment has affected the variance-covariance structure of the pre and post-observations. Actually, a one-sided alternative hypothesis may be in order since a decrease in the variability of the post-treatment observations seems plausible. To test the hypotheses stated in (196), we employ an asymptotic chi-squared test.

Let Ω denote the parameter space for the multi-consequence $MA_I(1)$ model. Then Ω is a subset of 5-dimensional space. Specifically,

$$\Omega = \{(\mu,\delta,\theta_1,\gamma_1,\sigma_a^2): -\infty<\mu<\infty,\ -\infty<\delta<\infty,\ -1<\theta_1<1, -1<\gamma_1<1, \sigma_a^2 > 0\}.$$

Let Ω_0 denote the parameter space when the null hypothesis is true. Thus,

$$\Omega_0 = \{(\mu,\delta,\theta_1,\gamma_1,\sigma_a^2):\ \theta_1=\gamma_1,\ -\infty<\mu<\infty,\ -\infty<\delta<\infty,\ \sigma_a^2 > 0\}.$$

Actually, Ω_0 defines the parameter space for the single consequence $MA_I(1)$ model. Let $L(\hat{\Omega}_0|\underset{\sim}{z}^t)$ denote the maximum value of the likelihood function found by using Theorem 3.1, and let $L(\hat{\Omega}|\underset{\sim}{z}^t)$ denote the maximum value of the likelihood function using Theorem 3.2. Define

$$\lambda(\underset{\sim}{Z}) = L(\hat{\Omega}_0|Z^t)/L(\hat{\Omega}|\underset{\sim}{Z}^t).$$

It can be shown that the distribution of $-2\,\ell n\lambda(\underset{\sim}{Z})$ converges to a χ_1^2 distribution when the null hypothesis is true. See Kendall and Stuart [43]. Thus, our decision rule is to reject H_0 when

$$-2\,\ell n\lambda(\underset{\sim}{z}) > \chi_{1,\alpha}^2 \quad . \tag{197}$$

Note that

$$- 2\ \ell n\lambda(\underset{\sim}{z}) = - 2[\ell n\ L(\hat{\Omega}_0|\underset{\sim}{z}^t) - \ell n\ L(\hat{\Omega}|\underset{\sim}{z}^t)]$$

$$= n\ \ell n(\hat{\sigma}_a^2)_0 + (\underset{\sim}{X}^t\underset{\sim}{X})_0 - n\ \ell n(\hat{\sigma}_a^2) - (\underset{\sim}{X}^t\underset{\sim}{X})\ , \qquad (198)$$

where the zero subscript indicates the values obtained under the null hypothesis $H_0 : \theta_1 = \gamma_1$. Thus, the decision rule stated in equation (197) can be restated as reject H_0 when

$$n\ \ell n(\hat{\sigma}_a^2)_0 + (\underset{\sim}{X}^t\underset{\sim}{X})_0 - n\ \ell n(\hat{\sigma}_a^2) - (\underset{\sim}{X}^t\underset{\sim}{X}) > \chi^2_{1,\alpha}\ . \qquad (199)$$

If the null hypothesis is rejected, one could then set up a pseudo t-test for testing $H_0 : \delta = 0$ vs. $H_1 : \delta \neq 0$ as described in Section 3.2.2.2 (Multi-Consequence $MA_I(1)$ Model); if the null hypothesis is not rejected, one would use the pseudo t-test described in Section 3.2.2.1 (Single-Consequence $MA_I(1)$ Model).

To illustrate the previous comments, consider the following example reported by Hall *et al* and used by Glass, Wilson, and Gottman [28].

Example 3.1: Figure 6 is a record of the daily number of "talk outs" of twenty-seven pupils in the second grade of an all-black urban proverty area school for a total time period of forty days. "Talk-outs" is a phrase describing the number of instances in which pupils talked to the teacher without first gaining permission such as occurs when the pupil raises his hand and talks to the teacher without being recognized. The number of "talk-outs" was recorded by the teacher on a hand held counter and a reliability check was made by an outside observer on two of these forty days. The first twenty days were denoted as the baseline period before the commencement of an intervention effect. Beginning on the first day of the fifth week (the 21st observation), the teacher initiated a program of systematic praise for those students who raised their hands and waited for recognition before talking. The teacher also allowed the students to choose a favorite activity such as working puzzles when the frequency of "talk-outs" was six or less.

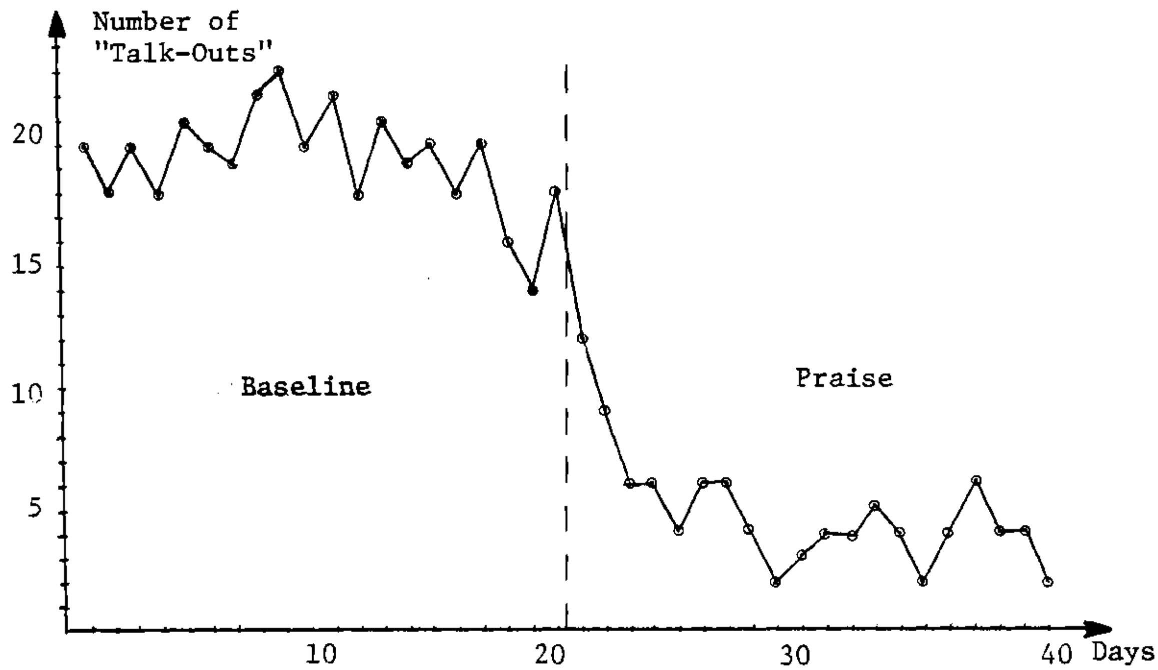

Figure 6. A Record of the Daily Number of "Talk-Outs".

A preliminary statistical analysis of the Hall *et al* data was conducted by Glass, Willson, and Gottman [28]. As a first step in identifying an appropriate model, seperate correlograms were computed for the pre-intervention and post-intervention series. As they state, "a single correlogram should not be computed without regard to possible intervention effects. The presence of an intervention effect can greatly increase autocorrelation coefficients." The first three estimated autocorrelations for both series as well as the average of both are given below:

	Lag		
	1	2	3
Pre-Intervention Autocorrelations	0.28	0.29	0.09
Post-Intervention Autocorrelations	0.46	0.12	0.07
Average Autocorrelations	0.37	0.20	0.08

The significance of these autocorrelations can be investigated using Bartlett's result for the variance of the kth estimated autocorrelation, denoted r_k. Bartlett's formula states that, assuming $\rho_v = 0$ for all $v > q$,

$$\mathrm{Var}(r_k) \simeq (1/N)[1 + 2\sum_{v=1}^{q} \rho_v^2] ,$$

for $k > q$. In practice, ρ_v is replaced by r_v. Furthermore, O.D. Anderson [8] states that r_k is approximately normally distributed for large N if $\rho_k = 0$. This allows us to compute standard error limits on the pre-I, post-I, and average autocorrelations stated above. These standard errors are given below as well as $r_k \pm 2\hat{\sigma}_{r_k}$. They are

	Lag		
	1	2	3
Pre-Intervention Limits	$\hat{\sigma}_{r_1} \simeq 0.22$ [-0.16, 0.72]	$\hat{\sigma}_{r_2} \simeq 0.24$ [-0.19, 0.77]	$\hat{\sigma}_{r_3} \simeq 0.26$ [-0.43, 0.61]
Post-Intervention Limits	$\hat{\sigma}_{r_1} \simeq 0.22$ [0.02, 0.90]	$\hat{\sigma}_{r_2} \simeq 0.27$ [-0.42, 0.66]	$\hat{\sigma}_{r_3} \simeq 0.27$ [-0.47,0.61]
Limits for Average	$\hat{\sigma}_{r_1} \simeq 0.16$ [0.05, 0.69]	$\hat{\sigma}_{r_2} \simeq 0.18$ [-0.16, 0.56]	$\hat{\sigma}_{r_3} \simeq 0.18$ [-0.28, 0.44]

As stated by Glass, Willson, and Gottman, "it is apparent that not even first differencing is required to remove the latter nonzero autocorrelations in the original data." Thus, one of the primary purposes of seperately calculating the pre-I and post-I autocorrelations is determining the stationarity of the series. How to combine the information from the pre-I and post-I autocorrelations is somewhat problematic. For example, the confidence intervals for the pre-I autocorrelations suggest a random process while those for the post-I autocorrelations suggest an $MA_I(1)$ process. Glass, Willson, and Gottman suggest averaging the pre-I and post-I autocorrelations. Using their suggestions, we tentatively identify the model as $MA_I(1)$. It now remains to test the two hypotheses that $\theta_1 = \gamma_1$ and $\delta = 0$.

Assuming a single consequence $MA_I(1)$ model and using <u>conditional</u> least squares estimation, Glass, Willson, and Gottman found that $\hat{\sigma}_a^2 = \hat{\underset{\sim}{a}}^t\hat{\underset{\sim}{a}}/(n-2)$ was minimized when $\hat{\theta}_1 = -.34$. At this value of θ_1, $\hat{\sigma}_a^2 = 4.47$, $\hat{\mu} = 19.24$ and $\hat{\delta} = -14.29$, which shows $\hat{\delta}$ to be significantly different from zero. Furthermore, in the region of optimal θ_1, the graph of $\hat{\sigma}_a^2$ versus θ_1 was fairly flat as was the graph of $\hat{\delta}$ versus θ_1.

When a multi-consequence $MA_I(1)$ model was assumed and <u>conditional</u> least squares estimation was employed (see computer program ICLSMAI(1) in Appendix C), it was found that $\hat{\theta}_1 = -0.28$, $\hat{\gamma}_1 = -0.65$, $\hat{\mu} = 19.07$, $\hat{\delta} = -13.94$, and $\hat{\sigma}_a^2 = 4.30$. Thus, while $\hat{\mu}$, $\hat{\delta}$, and $\hat{\sigma}_a^2$ for the multi-consequence model agree quite closely with the values for the single-consequence model, there appears to be considerable discrepancy in the values of the pre-I and post-I moving average parameters. To resolve this discrepancy, we employ maximum likelihood estimates and the likelihood ratio test.

Using the MLEMAI(1) program listed in Appendix D, it was found that $\hat{\theta}_1 = \hat{\gamma}_1 = -.25$, $\hat{\mu} = 19.26$, and $\hat{\delta} = -14.33$ are the maximum likelihood estimates under $H_0 : \theta_1 = \gamma_1$. These maximum likelihood estimates of μ and δ agree quite closely with both sets of least-squares estimates. Under the assumption that $\theta_1 \neq \gamma_1$, the maximum likelihood estimates are $\hat{\theta}_1 = -.19$, $\hat{\gamma}_1 = -.58$, $\hat{\delta} = 19.12$, and $\hat{\delta} = -14.06$. Again, there is close correspondence with the other cases for the estimates of μ and δ. Using the maximum likelihood estimates, we find that

$$-2 \ln \lambda(\underset{\sim}{z}) = 2.08,$$

and we would reject $H_0 : \theta_1 = \gamma_1$ only at the 14% level. Thus, if we do not reject $H_0 : \theta_1 = \gamma_1$, we employ the single-consequence $MA_I(1)$ model and find $\hat{\delta} = -14.29$ to be highly significant. Thus, there was a statistically significant decrease in the level of "talk-outs" commencing with the 21^{st} day when a reward system was initiated.

This concludes Chapter III.

CHAPTER IV

ECONOMIC ASPECTS OF CONTROL CHARTS FOR THE MEAN

This chapter extends the brief introduction to economic aspects of control charts presented in Chapter I. In this chapter, we will determine the constant in the control chart limit as well as the size of the sample to be selected. These will be chosen to minimize the average run length of an out of control process. Although we are considering a very simple type of economic control chart, it is perhaps the most valuable because of its ease of understanding and implementation.

In section 4.1, we review the work of Page [48] who laid the groundwork for this chapter by considering the case when the quality of output from a process is based on just one characteristic and the sample is random. In section 4.2, we extend the work of Page to the multivariate case with independent observations. That is, the quality of each item is determined by several characteristics and the vectors of observations are independent. In section 4.3, the quality of each item is dependent on only one characteristic; however, the observations are correlated. By comparing the results in the different sections, we are able to determine the influence of multiple quality characteristics and nonindependence on the parameters of interest.

4.1 One Quality Characteristic, Independent Observations

In section 2.1, it was shown that, when there is only one quality characteristic (X) which is normally distributed with standard values specified for the process mean (μ_0) and standard deviation (σ_I) and successive random samples of size n are generated from this process, the control chart limits are of the form $\mu_0 \pm B(\sigma_I/\sqrt{n})$, where $B = z_{\alpha/2} = 3.0$. In accordance with Page [48], let m denote the true value of the process mean which may vary from period to period. However, σ_I remains constant. Thus, $X \sim N(m, \sigma_I^2)$.

Let P(m) denote the probability that a given sample yields an $\bar{x}$ outside the control limits when m is the process mean. Then

$$P(m) = P(\bar{X} > \mu_0 + B\ \sigma_I/\sqrt{n}\ |m) + P(\bar{X} < \mu_0 - B\ \sigma_I/\sqrt{n}\ |m)$$

$$= P(Z > B + (\mu_0 - m)/(\sigma_I/\sqrt{n})) + P(Z < -B + (\mu_0 - m)/(\sigma_I/\sqrt{n})) \tag{200}$$

Let Y be a random variable denoting the number of samples up to and including the first one for which an $\bar{x}$ indicates an out of control process. Then Y is a geometric random variable with parameter P(m). Specifically,

$$p_Y(y) = P(m)[1 - P(m)]^{y-1},\ y = 1,2,\ldots$$
$$= 0 \qquad\qquad ,\ \text{otherwise}$$

It is well known that $E(Y) = 1/P(m)$.

Page defines the <u>average run length</u> (L) as the average number of articles inspected between two successive occasions when rectifying action is taken. For constant m,

$$L = nE(Y) = n/P(m),$$

which is the sample size per sample times the average number of samples up to and including the first one out of control.

Let L_0 denote the average run length when $m = \mu_0$. Since

$$P(\mu_0) = 2\ \Phi(-B),$$

it follows that

$$L_0 = n/[2\ \Phi(-B)]. \tag{201}$$

Let $k > 0$ be a value such that "a shift in the mean m of amount equal to or greater than $k\sigma_I$ is serious and we desire that such a shift should be detected as soon as possible after it has occurred." Define L_1 to be the average run length when $m = \mu_0 + k\sigma_I$. Since

$$P(\mu_0 + k\sigma_I) = \Phi(-B + k\sqrt{n}) + \Phi(-B - k\sqrt{n}),$$

it follows that

$$L_1 = n/[\Phi(-B + k\sqrt{n}) + \Phi(-B - k\sqrt{n})]. \tag{202}$$

Page provides two alternative schemes for determining B and n. The first of these is to choose that inspection scheme such that L_1 is minimized for some given large value of L_0 and fixed k. The second chooses that scheme such that L_0 is maximized for some given small value of L_1 and fixed k. We will concentrate only on the first scheme. By rewriting equation (201) as $n = 2\ L_0\ \Phi(-B)$ and substituting this result into equation (202), we see that

$$L_1 = \frac{2\ L_0\ \Phi(-B)}{\Phi(-B + k\sqrt{2\ L_0\ \Phi(-B)}) + \Phi(-B - k\sqrt{2\ L_0\ \Phi(-B)})} \tag{203}$$

The problem is to find B which minimizes L_1 for fixed L_0 and k. This B is then used to find n from the equation $n = 2\, L_0\, \Phi(-B)$. By using a computer search routine, Page constructed tables of n, B, and L_1 for L_0 = 2,000, 5,000, 10,000, 15,000, 20,000, 40,000, and 60,000 and k = (0.2)(0.1)(1.8), where (0.1) denotes the step size of k. Actually, Page's results were based on an upper control limit only. However, the values of n, B, and L_1 for two-sided control limits can be found from his tables by doubling the L_0 value. A computer program was written to duplicate Page's results, and, in order to facilitate later comparisons, the output is given in Table 3 for L_0 = 2,000, 5,000, 10,000, 15,000, 20,000, 40,000, and 60,000 with k = 0.2 (0.2) 1.8. These results correspond almost exactly with those of Page. Note that $B = \sqrt{\chi^2_{1,\alpha}}$. By inspecting the table, we see that, for a fixed L_0, n decreases as k increases. This is intuitively appealing since it says that a larger sample is needed to detect a small shift while a smaller sample will suffice for a large shift. The table also shows that, for a fixed k, n increases as L_0 increases. Thus, as the average run length of an in control process increases, a larger sample size is needed to detect a shift of given magnitude. Perhaps the most surprising result of Table 3 is that the control chart constant is quite frequently less than 3.0, the traditional value. For example, when L_0 = 5,000, it is only when $k \geq 1.0$ that $B \geq 3.0$. Thus, Page's scheme calls for tighter than usual control limits and larger than usual sample sizes to detect small shifts.

Table 3. Values of n, $\chi^2_{1,\alpha} = B^2$, and L_1 for fixed L_0 and k

One Characteristic, Independent Observations					
L_0	k	n	$\chi^2_{1,\alpha}$	B	L_1
2000	.20	114	3.623	1.903	192.3
2000	.40	44	5.245	2.290	68.5
2000	.60	24	6.311	2.512	36.1
2000	.80	15	7.149	2.674	22.6
2000	1.00	11	7.705	2.776	15.6
2000	1.20	8	8.284	2.878	11.5
2000	1.40	6	8.807	2.968	8.9
2000	1.60	5	9.140	3.023	7.0
2000	1.80	4	9.548	3.090	5.8
5000	.20	154	4.664	2.160	245.7
5000	.40	56	6.433	2.536	82.8
5000	.60	29	7.610	2.759	42.5
5000	.80	19	8.377	2.894	26.3
5000	1.00	13	9.068	3.011	18.0
5000	1.20	9	9.741	3.121	13.2
5000	1.40	7	10.199	3.194	10.1
5000	1.60	6	10.478	3.237	8.0
5000	1.80	5	10.827	3.290	6.5
10000	.20	187	5.528	2.351	287.8
10000	.40	65	7.405	2.721	93.8
10000	.60	34	8.579	2.929	47.5
10000	.80	21	9.459	3.076	29.1
10000	1.00	14	10.199	3.194	19.8
10000	1.20	11	10.635	3.261	14.4
10000	1.40	8	11.241	3.353	11.0
10000	1.60	6	11.774	3.431	8.7
10000	1.80	5	12.110	3.480	7.1
15000	.20	208	6.055	2.461	313.0
15000	.40	70	8.004	2.829	100.3
15000	.60	36	9.215	3.036	50.4
15000	.80	22	10.114	3.180	30.7
15000	1.00	15	10.827	3.290	20.8
15000	1.20	11	11.402	3.377	15.1
15000	1.40	9	11.774	3.431	11.5
15000	1.60	7	12.237	3.498	9.1
15000	1.80	6	12.520	3.538	7.4

Table 3. (Cont'd.)

One Characteristic, Independent Observations					
L_0	k	n	$\chi^2_{1,\alpha}$	B	L_1
20000	.20	222	6.449	2.539	331.1
20000	.40	74	8.425	2.903	105.0
20000	.60	38	9.642	3.105	52.5
20000	.80	23	10.555	3.249	31.9
20000	1.00	16	11.241	3.353	21.6
20000	1.20	12	11.774	3.431	15.7
20000	1.40	9	12.304	3.508	11.9
20000	1.60	7	12.763	3.573	9.4
20000	1.80	6	13.040	3.611	7.6
40000	.20	259	7.412	2.722	375.2
40000	.40	84	9.459	3.076	116.2
40000	.60	42	10.737	3.277	57.6
40000	.80	26	11.625	3.410	34.7
40000	1.00	18	12.304	3.508	23.4
40000	1.20	13	12.896	3.591	16.9
40000	1.40	10	13.361	3.655	12.8
40000	1.60	8	13.741	3.707	10.1
40000	1.80	6	14.364	3.790	8.2
60000	.20	281	7.998	2.828	401.3
60000	.40	89	10.094	3.177	122.8
60000	.60	45	11.360	3.370	60.5
60000	.80	27	12.304	3.508	36.4
60000	1.00	18	13.040	3.611	24.4
60000	1.20	13	13.606	3.689	17.6
60000	1.40	10	14.169	3.764	13.4
60000	1.60	8	14.582	3.819	10.5
60000	1.80	7	14.827	3.851	8.5

One can also use equations (201) and (202) to determine L_1 for traditional $B = 3.0$. Specifically, suppose we let $L_0 = 10{,}000$ and $B = 3.0$. From equation (201), we see that $n = 2(10{,}000)(.00135) = 27$. From equation (202) with $n = 27$, $B = 3.0$, and $k = 0.2$, we see that $L_1 = 1080$. From the tables with $k = 0.2$, we see that $L_1 = 288$, $n = 186$, and $B = 2.35$. Thus, the scheme to minimize L_1 results in a considerable reduction in the number of defective items produced before the out of control state is detected.

In arriving at equation (202), we looked at $P(\mu_0 + k\sigma_1)$ which is the probability of a given sample yielding an $\bar{x}$ outside the control limits when the true mean has shifted by k standard deviations from the nominal value. Thus, the measure of departure was in terms of process standard deviations. One could just as well have measured the departure using standard deviations of $\bar{X}$, viz., $\sigma_{\bar{X}}$ where $\sigma_{\bar{X}} = \sigma_I/\sqrt{n}$. In this case, $m = \mu_0 + (c\sigma_I/\sqrt{n})$ and the analogue to equation (202) is

$$L_1 = n/[\Phi(-B + c) + \Phi(-B - c)] \ ,$$

which is a simpler expression. However, the decision maker who sets up control charts may have a more difficult time interpreting departures expressed in terms of $\sigma_{\bar{X}}$ than when using σ_I. For this reason, we will not adopt this approach.

4.2 Multiple Quality Characteristics, Independent Observations

This section extends the results of the previous section by allowing the quality of each item to be governed by more than one quality characteristic. In order to do this, we need to recall certain results presented in Section 2.2.1.

Suppose $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ is a random sample of size n from a p-variate normal process with mean vector $\underset{\sim}{m}$ and known variance-covariance matrix Σ. Let $\underset{\sim}{\mu}_0$ denote the nominal value of the process mean. To maintain statistical control over $\underset{\sim}{\mu}_0$, the vector of sample means $(\bar{\underset{\sim}{x}})$ is calculated and it is necessary to determine whether $n(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0)^t \Sigma^{-1}(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0)$ exceeds the upper control limit $(\chi^2_{p,\alpha})$.

Page's procedure can be extended to this multivariate case by determining $P(\underset{\sim}{m})$ where

$$P(\underset{\sim}{m}) = P[n(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}_0)^t \Sigma^{-1}(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}_0) > \chi^2_{p,\alpha} \mid \underset{\sim}{m}] , \tag{204}$$

which is the probability that the statistic plots out of control when the true process mean is $\underset{\sim}{m}$. If $\underset{\sim}{X} \sim N_p(\underset{\sim}{m}, \Sigma)$ and hence $\bar{\underset{\sim}{X}} \sim N_p(\underset{\sim}{m}, \Sigma/n)$, then it follows that (see Alt [2])

$$n(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}_0)^t \Sigma^{-1}(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}_0) \sim \chi'^2_{p,\lambda}$$

where $\lambda = n(\underset{\sim}{m} - \underset{\sim}{\mu}_0)^t \Sigma^{-1}(\underset{\sim}{m} - \underset{\sim}{\mu}_0)$. Thus,

$$P(\underset{\sim}{m}) = P(\chi'^2_{p,\lambda} > \chi^2_{p,\alpha} \mid \underset{\sim}{m}) .$$

When $\underset{\sim}{m} = \underset{\sim}{\mu}_0$, $\lambda = 0$,

$$P(\underset{\sim}{\mu}_0) = P(\chi^2_p > \chi^2_{p,\alpha}) ,$$

and

$$L_0 = n/P(\chi^2_p > \chi^2_{p,\alpha}). \tag{205}$$

In order to measure departures of $\underset{\sim}{m}$ from $\underset{\sim}{\mu}_0$, it is necessary to account for the possible departure of each component of $\underset{\sim}{m}$. This is accomplished by introducing the (p x 1) vector $\underset{\sim}{\sigma}$ where

$$\underset{\sim}{\sigma}^t = [k_1\sigma_1, k_2\sigma_2, \ldots, k_p\sigma_p] ,$$

and letting $\underset{\sim}{m} = \underset{\sim}{\mu}_0 + \underset{\sim}{\sigma}$. We require that at least one $k_i > 0$. Thus, although we wish to simultaneously control the mean vector of several variables, it may be necessary to detect a shift in only one of these variables. When $\underset{\sim}{m} = \underset{\sim}{\mu}_0 + \underset{\sim}{\sigma}$, $\lambda = n\, \underset{\sim}{\sigma}^t\, \Sigma^{-1}\, \underset{\sim}{\sigma}$,

$$P(\underset{\sim}{\mu}_0 + \underset{\sim}{\sigma}) = P({\chi'}^2_{p,\lambda} > \chi^2_{p,\alpha}) ,$$

and

$$L_1 = n/P({\chi'}^2_{p,\lambda} > \chi^2_{p,\alpha}) . \tag{206}$$

The specific form of λ will become clearer when we let p = 2 and 3.

As with the univariate case, that inspection scheme will be chosen which minimizes L_1 for some given large value of L_0 and fixed $\underset{\sim}{k} = [k_1, \ldots, k_p]^t$. However, in the multivariate case, we must also fix $\underset{\sim}{\rho} = [\rho_{12}, \rho_{13}, \ldots, \rho_{p-1,p}]^t$ where ρ_{ij} denotes the correlation between quality characteristics X_i and X_j. By rewriting equation (205) as $n = L_0\, P(\chi_p^2 > \chi^2_{p,\alpha})$ and substituting this result into equation (206), we see that

$$L_1 = \frac{L_0\, P(\chi_p^2 > \chi^2_{p,\alpha})}{P({\chi'}^2_{p,\lambda} > \chi^2_{p,\alpha})} . \tag{207}$$

For fixed L_0, $\underset{\sim}{k}$, and $\underset{\sim}{\rho}$ we seek that $\chi^2_{p,\alpha}$ and n which minimizes L_1 as stated in equation (207). One difficulty in doing this is the need for evaluating the denominator of equation (207), which is the complementary cumulative noncentral chi-square distribution function evaluated at $\chi^2_{p,\alpha}$.

Let $Y = [\chi'^2_{\nu,\lambda}/(\nu + \lambda)]^h$. Sankaran [55] examined the cumulants of Y expressed in terms of the cumulants of $\chi'^2_{\nu,\lambda}$ as a power series in $(\nu + \lambda)^{-s}$. Sankaran chooses h so that the leading term in the third cumulant of Y vanishes. This results in Y being approximately normally distributed with

$$E(Y) = 1 + h(h - 1)(\nu + 2\lambda)(\nu + \lambda)^{-2}$$

$$- h(h - 1)(2 - h)(1 - 3h)(\nu + 2\lambda)^2(\nu + \lambda)^{-4}/2 ,$$

$$V(Y) = \frac{2h^2(\nu + 2\lambda)}{(\nu + \lambda)^2}\left[1 - \frac{(1 - h)(1 - 3h)(\nu + 2\lambda)}{2(\nu + \lambda)^2}\right],$$

and

$$h = 1 - (2/3)(\nu + \lambda)(\nu + 3\lambda)(\nu + 2\lambda)^{-2} .$$

An empirical comparison by Johnson and Kotz [40] shows that Sankaran's approximation is remarkably accurate for all values of λ. However, in reporting on Sankaran's approximation, there are several typographical errors in their equation for V(Y). Now

$$P(\chi'^2_{p,\lambda} > \chi^2_{p,\alpha}) = P[(\chi'^2_{p,\lambda}(p + \lambda)^{-1})^h > (\chi^2_{p,\alpha}(p + \lambda)^{-1})^h]$$

$$\approx 1 - \Phi[\{(\chi^2_{p,\alpha}(p + \lambda)^{-1})^h - E(Y)\}/\sqrt{V(Y)}] , \qquad (208)$$

and approximation (208) was used in the search routine.

The search routine used to find the minimum L_1 is a modified version of the success-failure method as described by Dixon. The basic idea is to let $n^{(1)} = 1$. Since L_0 is fixed and $n = n^{(1)}$ is now fixed, $\chi_{p,\alpha}^2$ can be determined from equation (205). By letting $n = n^{(1)}$ and $\chi_{p,\alpha}^2 = \chi_{p,\alpha}^{2\,(1)}$, a value of L_1 (denoted by $L_1^{(1)}$) can be determined from equation (206). Now let $n^{(2)} = n^{(1)} + \delta$, where δ is a positive integer greater than one, and determine $\chi_{p,\alpha}^{2\,(2)}$ from equation (205). In turn, $L_1^{(2)}$ is determined from equation (206). This procedure is continued until $L_1^{(k)} > L_1^{(k-1)}$. When this occurs, we set $n^{(k+1)} = n^{(k)} - 2\delta$ and evaluate $L_1^{(k+1)}$. We now set $n^{(k+2)} = n^{(k+1)} + (\delta/2)$ and evaluate $L_1^{(k+2)}$. This forward search is continued until some L_1 is greater than the previous one, at which time we go back to an earlier n and use a smaller δ in the forward search. Although time consuming, this eventually leads to a minimum value of L_1 for fixed L_0, $\underset{\sim}{k}$, and $\underset{\sim}{\rho}$. When the minimum value of L_1 is found, the output can be arranged in table format with column headings: L_0, $\underset{\sim}{\rho}^t$, $\underset{\sim}{k}^t$, n, $\chi_{p,\alpha}^2$, L_1, λ. It was this search routine that was used to generate the univariate results of Table 3 by letting $p = 1$, $\underset{\sim}{\rho} = \underset{\sim}{0}$, and $\underset{\sim}{k} = k_1$. When $p = 1$, the noncentrality parameter reduces to $\lambda = n\,k_1^2$.

The first case to be investigated is when $p = 2$. That is, there are two quality characteristics. In this instance,

$$\lambda = n(1 - \rho^2)^{-1}(k_1^2 - 2\rho\; k_1 k_2 + k_2^2). \tag{209}$$

Note that when $\rho = 0$ and $k_2 = 0$, $\lambda = n\,k_1^2$, which is the univariate noncentrality parameter. The minimization of L_1 was investigated for

L_0 = 5,000, 10,000, 20,000, and 40,000, ρ = (-0.8)(0.2)(+0.8), k_1 = (0.2)(0.2)(1.8) and k_2 = (0.0)(0.2)(1.8). The complete results are presented in Appendix E, with a few selected values shown in Table 4 to indicate the general pattern.

Inspection of Table 4 reveals several traits. One is that the sample size needed to detect departures of given magnitudes (fixed k_1 and k_2) is always larger for L_0 = 10,000 than for L_0 = 5,000. This result is intuitively appealing for we should expect that as L_0 increases a larger sample becomes necessary. We also note that reversing the roles of k_1 and k_2 for fixed ρ yields the same n, $\chi^2_{2,\alpha}$, and L_1. For example, when ρ = -0.8, k_1 = 0.2 and k_2 = 0.6, we get the same results as when k_1 = 0.6 and k_2 = 0.2. This occurs since the noncentrality parameter is symmetric in k_1 and k_2.

Upon first glancing at Table 4, it appears that, for a fixed L_0 and ρ, n decreases as k_1 and k_2 increase. Again, this is intuitively appealing, for the magnitude of the required sample size should indeed decrease as the magnitudes of the shifts which are important to detect increase. Usually, n is much larger for small k_1 and k_2 = 0.0 than for other values of k_2. The interpretation of k_2 = 0.0 is that it is important to detect a shift of zero magnitude in the second component, or an "infinitesimally small" shift. This accounts for the rather large sample sizes in this case. However, further inspection of Table 4 shows that it is not always true that n decreases as k_1 and k_2 increase for fixed L_0 and ρ. While this is true for $\rho \leq 0$ and also for ρ = 0.4 when k_1 is small, it is not true for the other values of k_1 and ρ = 0.4, nor

Table 4. Economic Parameters for Two Quality Characteristics, Independent Observations

			L_0 = 5,000			L_0 = 10,000		
	k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1^*	n	$\chi^2_{2,\alpha}$	L_1^*
$\rho = -0.8$	0.2	0.0	88	8.08	130 [1]	103	9.15	148 [1]
		0.2	32	10.10	45 [2]	36	11.25	50 [2]
		0.6	10	12.43	13 [3]	11	13.62	15 [3]
		1.0	5	13.81	6 [4]	5	15.20	7 [4]
	0.6	0.0	15	11.61	21 [5]	17	12.75	23 [5]
		0.2	10	12.43	13 [3]	11	13.62	15 [3]
		0.6	5	13.81	7 [6]	6	14.81	7 [6]
		1.0	3	14.81	4 [7]	3	16.22	4 [7]
	1.0	0.0	6	13.45	9 [8]	7	14.51	9 [8]
		0.2	5	13.81	6 [4]	5	15.20	7 [4]
		0.6	3	14.81	4 [7]	3	16.22	4 [7]
		1.0	2	15.65	3	2	17.00	3
$\rho = -0.4$	0.2	0.0	166	6.81	254 [9]	199	7.83	295 [9]
		0.2	77	8.35	112	89	9.44	127
		0.6	23	10.76	32 [10]	26	11.90	36 [10]
		1.0	11	12.24	15 [11]	12	13.45	17 [11]
	0.6	0.0	30	10.23	43 [12]	34	11.37	47 [12]
		0.2	23	10.76	32 [10]	26	11.90	36 [10]
		0.6	13	11.90	18 [13]	14	13.14	19 [13]
		1.0	8	12.87	10 [14]	9	14.02	11 [14]
	1.0	0.0	13	11.90	18 [13]	14	13.14	19 [13]
		0.2	11	12.24	15 [11]	12	13.45	17 [11]
		0.6	8	12.87	10 [14]	9	14.02	11 [14]
		1.0	5	13.81	7 [6]	6	14.81	8 [6]

Table 4. (Cont'd.)

			L_0 = 5,000			L_0 = 10,000		
	k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1^*	n	$\chi^2_{2,\alpha}$	L_1^*
ρ = 0.0	0.2	0.0	188	6.56	291	227	7.56	339
		0.2	113	7.58	169	133	8.64	194
		0.6	32	10.10	45 [2]	36	11.25	50 [2]
		1.0	15	11.61	20 [15]	16	12.87	22 [15]
	0.6	0.0	35	9.92	49 [16]	40	11.09	55 [16]
		0.2	32	10.10	45 [2]	36	11.25	50 [2]
		0.6	20	11.04	27	22	12.24	30
		1.0	12	12.06	16 [17]	13	13.29	17 [17]
	1.0	0.0	15	11.61	21 [5]	17	12.75	23 [5]
		0.2	15	11.61	20 [15]	16	12.87	22 [15]
		0.6	12	12.06	16 [17]	13	13.29	17 [17]
		1.0	8	12.87	11 [18]	9	14.02	12 [18]
ρ = +0.4	0.2	0.0	166	6.81	254 [9]	199	7.83	295 [9]
		0.2	145	7.08	221	173	8.11	254
		0.6	35	9.92	49 [16]	39	11.09	55 [16]
		1.0	14	11.75	20 [15]	16	12.87	22 [15]
	0.6	0.0	30	10.23	43 [12]	34	11.37	47 [12]
		0.2	35	9.92	49 [16]	39	11.09	55 [16]
		0.6	26	10.52	37	29	11.68	40
		1.0	14	11.75	20 [15]	16	12.87	22 [15]
	1.0	0.0	13	11.90	18 [13]	14	13.14	19 [13]
		0.2	14	11.75	20 [15]	16	12.87	22
		0.6	14	11.75	20 [15]	16	12.87	22 [15]
		1.0	11	12.24	15 [11]	12	13.45	17 [11]

Table 4. (Cont'd.)

	k_1	k_2	$L_0 = 5,000$ n	$\chi^2_{2,\alpha}$	L_1*	$L_0 = 10,000$ n	$\chi^2_{2,\alpha}$	L_1*
$\rho = +0.8$	0.2	0.0	88	8.08	130 [1]	103	9.15	148 [1]
		0.2	174	6.72	268	209	7.74	311
		0.6	24	10.68	33 [19]	27	11.82	36 [19]
		1.0	8	12.87	11 [18]	9	14.02	12 [18]
	0.6	0.0	15	11.61	21 [5]	17	12.75	23 [5]
		0.2	24	10.68	33 [19]	27	11.82	36 [19]
		0.6	32	10.10	45 [2]	36	11.25	50 [2]
		1.0	14	11.75	19 [20]	15	13.00	21 [20]
	1.0	0.0	6	13.45	9 [8]	7	14.51	9 [8]
		0.2	8	12.87	11 [18]	9	14.02	12 [18]
		0.6	14	11.75	19 [20]	15	13.00	21 [20]
		1.0	14	11.75	19 [20]	15	13.00	21 [20]

*The numerical superscripts indicate those entries which have the same values of the economic parameters n, $\chi^2_{2,\alpha}$, and L_1.

is it ever true when $\rho = 0.8$. Thus, for a relatively large positive correlation, the sample size needed to detect large positive shifts is larger than the sample sizes needed for smaller positive shifts. An explanation of this is provided by examining the noncentrality parameter λ, which is a generalized measure of distance of how far the true mean is from the nominal value. Fix $\rho = +0.4$ and $k_1 = 0.6$. When $k_2 = 0.2$, $\lambda = (n/.84)(.304)$; when $k_2 = 0.0$, $\lambda = (n/.84)(.360)$; and, when $k_2 = 0.6$, $\lambda = (n/.84)(.432)$. Inspection of Table 4 shows that, for the (k_1, k_2) pairs investigated, the largest sample size (35) occurred with the smallest value of the noncentrality parameter (.304), the next largest sample size (30) occurred with the next to the smallest value of the noncentrality parameter (.36), and the smallest sample size (26) occurred with the largest value of the noncentrality parameter. Thus, when the generalized measure of distance (λ) between the true mean and the nominal value is small, it is to be expected that a larger sample size will be needed to detect such a small shift. The results are summarized below.

(k_1, k_2)	λ	n
(0.6, 0.0)	(n/.84)(.360)	30
(0.6, 0.2)	(n/.84)(.304)	35
(0.6, 0.6)	(n/.84)(.432)	26

Let us now compare n for positive ρ with n for negative ρ. It is to be expected that both n's will be equal when $k_2 = 0$ since

$\lambda = n(1 - \rho^2)^{-1} k_1^2$ and the sign of ρ is lost through the squaring operation. However, for fixed k_1 and k_2, n is always much smaller for $\rho < 0$. However, this is not to imply that one should try to choose negatively correlated characteristics as opposed to positively correlated characteristics. The stated phenomenon occurs because we are looking at positive shifts ($k_1 > 0$, $k_2 > 0$) instead of negative shifts ($k_1 < 0$, $k_2 < 0$). Thus for $\rho < 0$ and $k_1 > 0$, $k_2 > 0$, the generalized distance measure (λ) is larger than for $\rho > 0$ and $k_1 > 0$, $k_2 > 0$. As the distance of the true mean from the nominal value increases, the sample size needed to detect this becomes smaller.

One additional topic of interest is how does the required sample size for two quality characteristics compare with the sample size for one quality characteristic (Table 3)? Some idea of this behavior is obtained by letting $\rho = 0.0$. Thus, $\lambda = n(k_1^2 + k_2^2)$. Now, when $k_2 = 0$, λ reduces to the univariate noncentrality parameter $n\ k^2$. However, the control limit will still be $\chi^2_{2,\alpha}$. Tables 3 and 4 show that, for $\rho = 0.0$, $k_2 = 0.0$ and fixed L_0 and k_1, the required sample size is larger for two quality characteristics than for one quality characteristic with this difference becoming smaller as k_1 increases. Furthermore, as soon as k_2 becomes positive, n for p = 2 is usually much smaller than for p = 1. Thus, an economical sample size is not an unusual result when two quality characteristics are used as opposed to one. As a final point of interest, note that the maximum n in Table 4 occurs for $\rho = 0.0$, $k_1 = 0.2$, and $k_2 = 0.0$. This is the one case where the required sample size for p = 1 (Table 3) is considerably smaller than for p = 2 (Table 4).

The next case to be investigated is when there are three quality characteristics (p = 3). Here,

$$\lambda = (n/\Delta)[k_1^2(1-\rho_{23}^2) + k_2^2(1-\rho_{13}^2) + k_3^2(1-\rho_{12}^2) + 2k_1k_2(\rho_{13}\rho_{23} - \rho_{12})$$

$$+ 2k_1k_3(\rho_{12}\rho_{23} - \rho_{13}) + 2k_2k_3(\rho_{12}\rho_{13} - \rho_{23})] \tag{210}$$

where

$$\Delta = 1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2\rho_{12}\rho_{13}\rho_{23} \; . \tag{211}$$

Note that, when $\rho_{13} = \rho_{23} = 0$ and $k_3 = 0$, equation (210) reduces to equation (209), which is the noncentrality parameter for two quality characteristics. The determination of that n and $\chi^2_{3,\alpha}$ which minimize L_1 was carried out for L_0 = 10,000, ρ_{12} = (-0.4)(0.4)(+0.4), ρ_{13} = (-04)(0.4)(+0.4), ρ_{23} = (-0.4)(0.4)(+0.4), k_1 = 0.2 (0.4) 1.0, k_2 = 0.0, 0.2 (0.4) 1.0, and k_3 = 0.0, 0.2 (0.4) 1.0. It was necessary to reduce the range of the ρ_{ij}'s, since, for given ρ_{12} and ρ_{13}, Kendall [42] has shown that ρ_{23} must lie in the range

$$\rho_{12}\,\rho_{13} \pm (1 - \rho_{12}^2 - \rho_{13}^2 + \rho_{12}^2\,\rho_{13}^2)^{1/2} \; .$$

Although additional values of ρ_{ij}'s could have been investigated, the limitations of space were another determining factor. Selected results are presented in Table 5.

Quite a few of the entries in Table 5 will be duplicates since they yield the same noncentrality parameter. These entries are denoted by superscript numerals in the L_1 column. For example, when ρ_{13} =

Table 5. Economic Parameters for Three Quality Characteristics, Independent Observations

$L_0 = 10{,}000$								
ρ_{12}	ρ_{13}	ρ_{23}	k_1	k_2	k_3	n	$\chi^2_{3,\alpha}$	L_1*
0.0	0.0	0.0	0.2	0.0	0.0	254	9.313	374
			0.2	0.2	0.0	149	10.478	213
			0.2	0.2	0.2	108	11.178	152
			0.6	0.6	0.6	17	15.140	23
-0.4	0.0	0.0	0.2	0.0	0.0	223	9.599	325 [1]
			0.2	0.2	0.0	99	11.367	140 [7]
			0.2	0.2	0.2	80	11.827	112
			0.6	0.6	0.6	13	15.706	17
+0.4	0.0	0.0	0.2	0.0	0.0	223	9.599	325 [1]
			0.2	0.2	0.0	194	9.904	281
			0.2	0.2	0.2	128	10.810	182
			0.6	0.6	0.6	21	14.691	28
-0.4	-0.4	0.0	0.2	0.0	0.0	189	9.961	274 [2]
			0.2	0.2	0.0	88	11.621	123 [3]
			0.2	0.2	0.2	56	12.594	77
			0.6	0.6	0.6	9	16.473	11
-0.4	+0.4	0.0	0.2	0.0	0.0	189	9.961	274 [2]
			0.2	0.2	0.0	88	11.621	123 [3]
			0.2	0.2	0.2	96	11.433	135 [4]
			0.6	0.6	0.6	15	15.405	20 [5]
+0.4	-0.4	0.0	0.2	0.0	0.0	189	9.961	274 [2]
			0.2	0.2	0.0	184	10.019	266 [6]
			0.2	0.2	0.2	96	11.433	135 [4]
			0.6	0.6	0.6	15	15.405	20 [5]

Table 5. (Cont'd.)

L_0 = 10,000								
ρ_{12}	ρ_{13}	ρ_{23}	k_1	k_2	k_3	n	$\chi^2_{3,\alpha}$	L_1^*
+0.4	+0.4	0.0	0.2	0.0	0.0	189	9.961	274 [2]
			0.2	0.2	0.0	184	10.019	266 [6]
			0.2	0.2	0.2	146	10.522	208
			0.6	0.6	0.6	24	14.407	32
-0.4	-0.4	-0.4	0.2	0.0	0.0	141	10.600	202 [11]
			0.2	0.2	0.0	53	12.712	73 [14]
			0.2	0.2	0.2	29	14.004	39
			0.6	0.6	0.6	4	18.194	6
-0.4	-0.4	+0.4	0.2	0.0	0.0	209	9.741	304 [7]
			0.2	0.2	0.0	99	11.367	140 [8]
			0.2	0.2	0.2	76	11.937	106 [9]
			0.6	0.6	0.6	12	15.874	16 [10]
-0.4	+0.4	-0.4	See entries for $\rho_{12} = \rho_{13} = -0.4$, $\rho_{23} = +0.4$					
+0.4	-0.4	-0.4	0.2	0.0	0.0	209	9.741	304 [7]
			0.2	0.2	0.0	158	10.351	227 [15]
			0.2	0.2	0.2	76	11.937	106 [9]
			0.6	0.6	0.6	12	15.874	16 [10]
+0.4	+0.4	-0.4	0.2	0.0	0.0	141	10.600	202 [11]
			0.2	0.2	0.0	194	9.904	281 [12]
			0.2	0.2	0.2	94	11.479	132 [13]
			0.6	0.6	0.6	15	15.405	20 [5]

Table 5. (Cont'd.)

$L_0 = 10{,}000$								
ρ_{12}	ρ_{13}	ρ_{23}	k_1	k_2	k_3	n	$\chi^2_{3,\alpha}$	L_1*
+0.4	-0.4	+0.4	See entries for $\rho_{12} = \rho_{13} = +0.4$, $\rho_{23} = -0.4$					
-0.4	+0.4	+0.4	0.2	0.0	0.0	141	10.600	202 [11]
			0.2	0.2	0.0	53	12.712	73 [14]
			0.2	0.2	0.2	94	11.479	132 [13]
			0.6	0.6	0.6	15	15.405	20 [5]
+0.4	+0.4	+0.4	0.2	0.0	0.0	209	9.741	304 [7]
			0.2	0.2	0.0	158	10.351	227 [15]
			0.2	0.2	0.2	172	10.166	248
			0.6	0.6	0.6	29	14.004	39

*The numerical superscripts indicate those entries which have the same values of the economic parameters n, $\chi^2_{2,\alpha}$, and L_1.

$\rho_{23} = 0$, $k_1 = 0.2$, and $k_2 = k_3 = 0.0$, we get the same economic parameters for $\rho_{12} = -0.4$ as we do for $\rho_{12} = +0.4$.

Table 5 shows that the largest n occurs when $\rho_{12} = \rho_{13} = \rho_{23} = 0$, $k_1 = 0.2$, and $k_2 = k_3 = 0$. In this case, λ reduces to the univariate noncentrality parameter. For p = 3, n = 254 (Table 5); for p = 2, n = 227 (Table 4); and, for p = 1, n = 187. Thus, the trivariate case does require more observations. Remember that $k_i = 0$ implies that it is important to detect an "infinitesimally" small shift in the i^{th} characteristic. However, as soon as k_1, k_2, and k_3 each equal 0.2, the trivariate sample size n = 108 is less than the bivariate sample size (n = 133) for $k_1 = k_2 = 0.2$ which, in turn, is less than the univariate sample size (n = 187) for $k_1 = 0.2$. Thus, even if the variables are uncorrelated, a χ_3^2 chart should be used instead of three univariate $\bar{X}$-charts if sample size economy is important.

Table 5 also seems to indicate that, for fixed ρ_{12}, ρ_{13}, and ρ_{23}, n decreases as k_1, k_2, and k_3 increase. However, this is true only for the first part of Table 5, for, when at lease two of the ρ_{ij}'s are positive, n first increases and then decreases as the k's increase. Thus, the required sample size does not always decrease as the magnitudes of the shifts that become important to detect increase in a positive direction. As in the bivariate case, this oddity is directly related to the value of the noncentrality parameter.

4.3 One Quality Characteristic, Correlated Observations

One of the cases considered in Section 2.1.2 assumed that the n elements of the sample were jointly normal and had a first-order serial correlation. See equations (10) and (11). If m dentoes the true value of the porcess mean, then $\bar{X} \sim N(m, (\sigma^2_c/n)[1 + 2\rho(1 - n^{-1})])$ and modified control limits for an $\bar{X}$-chart (see equation (13)) are of the form

$$\mu_0 \pm B(\sigma_c/\sqrt{n})[1 + 2\rho(1 - n^{-1})] \ ,$$

where μ_0 denotes the standard value of the porcess mean and B was substituted for $z_{\alpha/2}$. These limits were derived under the assumption that $m = \mu_0$. Although it was shown that $\hat{\mu}$ (the maximum likelihood estimator) is more efficient that $\bar{X}$, some justification for using $\bar{X}$ is provided by the r-dependent central limit theorem, where r-dependence means that X_t and X_{t+s} are autocorrelated only if $s \leq r$. The stationary r-dependent central limit theorem is stated by Kleijnen [44] as follows:

> Given an r-dependent strictly stationary sample X_1, $X_2, \ldots, X_t, \ldots, X_n$ with $E(X_t) = \mu$ and $E(|X_t|^3)$ existing, then the sample mean
>
> $$\bar{X} = \sum_{t=1}^{n} X_t/n$$
>
> is asymptotically normally distributed with mean μ and variance
>
> $$\text{Var }(\bar{X}) = (\sigma^2/n)[1 + 2 \sum_{s=1}^{n} (1 - \frac{s}{n}) \rho_s] \ .$$

Kleijnen points out that the variance is no asymptotic property but holds for any n. The primary importance of this theorem is that in the absence of process normality a correlated $\bar{X}$-chart can still be used.

Let P(m) denote the probability that a given sample yields an $\bar{x}$ outside the correlated control limits when m is the process mean. Then, for first-order serial correlation,

$$P(m) = P(\bar{X} > \mu_0 + B(\sigma_c/\sqrt{n}) \sqrt{1 + 2\rho(1 - n^{-1})} \,|\, m)$$

$$+ P(\bar{X} < \mu_0 - B(\sigma_c/\sqrt{n}) \sqrt{1 + 2\rho(1 - n^{-1})} \,|\, m)$$

$$= P(Z > B + a) + P(Z < -B + a) ,$$

where $a = (\mu_0 - m)/[(\sigma_c/\sqrt{n}) \sqrt{1 + 2\rho(1 - n^{-1})}]$. Continuing, we see that

$$P(m) = P(Z - a > B) + P(Z - a < -B) = P(|Z - a| > B) = P((Z - a)^2 > B^2)$$

Now $(Z - a) \sim N(-a, 1)$, and $(Z - a)^2 \sim \chi'^2_{1,\lambda}$ where $\lambda = a^2$. When $m = \mu_0$, $\lambda = 0$. and $P(\mu_0) = P(\chi_1^2 > B^2)$. When $m = \mu_0 + k\sigma_c$, $\lambda = n\, k^2/[1 + 2\rho(1 - n^{-1})]$, and $P(\mu_0 + k\sigma_c) = P(\chi'^2_{1,\lambda} > B^2)$. Note that B^2 is merely notation for $\chi^2_{1,\alpha}$.

Proceeding as we did in the earlier sections, we define

$$L_0 = n/P(\mu_0) = n/P(\chi_1^2 > \chi^2_{1,\alpha}) \tag{212}$$

and

$$L_1 = n/P(\mu_0 + k\sigma_c) = n/P(\chi'^2_{1,\lambda} > \chi^2_{1,\alpha}) , \tag{213}$$

where now

$$\lambda = nk^2/[1 + 2\rho(1 - n^{-1})] . \tag{214}$$

The above definition of λ allows us to investigate the effect of ρ on L_1. That inspection scheme will be chosen which minimizes L_1 for a large, fixed value of L_0. Specifically, we let L_0 = 5,000, 10,000, 20,000, 40,000, $\rho = (-0.4)(0.2)(+0.4)$, and $k = 0.2\ (0.2)(1.8)$. The reason for the restricted range of ρ was presented in Section 2.1.2. The results are presented in Table 6.

The results presented in Table 6 are identical with those presented in Table 3 (independent observations) when $\rho = 0$. We also see that, for a fixed L_0 and ρ, n decreases as k increases. This is true for both negative and positive ρ. The most surprising result in Table 6 is that, for fixed L_0 and k, n is always larger for positive ρ and smaller for negative ρ. This result is somewhat counterintuitive. Recall that, according to Nelson [46], a positive first-order serial correlation implies that a "higher-than-average observation tends to be followed by another higher-than-average observation, and although "there is no very long-lived persistence on one or the other side of the mean" the series should appear relatively smooth. However, for a negative first-order serial correlation, a higher-than-average observation tends to be followed by a lower-than-average observation, and the series should appear relatively choppy. Thus, in this latter case, we would expect that a larger sample size is needed to detect a shift of given magnitude since the shift is partially obscured by the choppy appearance of the

Table 6. Economic Parameters for One Quality Characteristic, First-Order Serial Correlation

		$L_0 = 5{,}000$				$L_0 = 10{,}000$			
ρ	k	n	$\chi^2_{1,\alpha}$	L_1	λ	n	$\chi^2_{1,\alpha}$	L_1	λ
-.4	.2	52	6.565	73.47	9.86	59	7.579	82.35	11.25
-.4	.4	19	8.377	25.60	13.33	21	9.459	27.91	14.89
-.4	.6	11	9.374	14.22	16.20	12	10.478	15.29	17.89
-.4	.8	8	9.956	9.57	19.94	8	11.241	10.23	19.94
-.4	1.0	6	10.478	7.12	22.27	6	11.774	7.58	22.27
-.4	1.2	5	10.827	5.67	25.92	5	12.110	5.95	25.92
-.4	1.4	4	11.241	4.66	27.22	4	12.520	4.92	27.22
-.4	1.6	3	11.774	4.09	25.60	4	12.520	4.27	35.56
-.4	1.8	3	11.774	3.44	32.40	3	13.040	3.61	32.40
-.2	.2	108	5.277	166.77	7.22	129	6.183	192.59	8.62
-.2	.4	38	7.125	55.32	10.23	44	8.111	62.00	11.82
-.2	.6	20	8.284	28.49	12.21	23	9.292	31.48	14.01
-.2	.8	13	9.068	17.74	14.25	14	10.199	19.45	15.32
-.2	1.0	9	9.741	12.31	15.63	10	10.827	13.40	17.29
-.2	1.2	7	10.199	9.14	17.72	8	11.241	9.94	20.11
-.2	1.4	6	10.478	7.20	20.88	6	11.774	7.70	20.88
-.2	1.6	5	10.627	5.86	23.04	5	12.110	6.21	23.04
-.2	1.8	4	11.241	4.52	23.82	4	12.520	5.14	23.82
.0	.2	154	4.664	245.67	6.20	187	5.528	287.81	7.52
.0	.4	56	6.433	82.78	9.12	65	7.405	93.79	10.56
.0	.6	29	7.610	42.53	10.80	34	8.579	47.50	12.60
.0	.8	19	8.377	26.25	12.80	21	9.459	29.05	14.08
.0	1.0	13	9.066	17.95	14.00	14	10.199	19.77	15.00
.0	1.2	9	9.741	13.16	14.40	11	10.635	14.39	17.28
.0	1.4	7	10.199	10.07	15.68	8	11.241	10.99	17.64
.0	1.6	6	10.478	7.97	17.92	6	11.774	8.73	17.92
.0	1.8	5	10.827	6.50	19.44	5	12.110	7.07	19.44
.2	.2	194	4.270	315.65	5.58	239	5.102	373.74	6.87
.2	.4	71	6.013	103.15	8.26	84	6.946	123.45	9.75

Table 6. (Cont'd.)

		$L_0 = 5,000$				$L_0 = 10,000$			
ρ	k	n	$\chi^2_{1,\alpha}$	L_1	λ	n	$\chi^2_{1,\alpha}$	L_1	λ
.2	.6	38	7.125	55.84	10.10	44	8.111	62.77	11.65
.2	.8	24	7.953	34.49	11.56	27	9.000	38.42	12.93
.2	1.0	17	8.579	23.57	13.06	19	9.642	26.09	14.49
.2	1.2	12	9.215	17.18	13.67	14	10.199	18.94	15.73
.2	1.4	9	9.741	13.11	14.41	11	10.635	14.42	17.21
.2	1.6	7	10.199	10.34	15.17	8	11.241	11.33	17.00
.2	1.8	6	10.478	8.34	16.89	7	11.488	9.16	19.20
.4	.2	230	3.982	379.40	5.14	286	4.791	453.05	6.39
.4	.4	86	5.676	131.92	7.77	102	6.600	151.46	9.20
.4	.6	46	6.783	68.51	9.49	53	7.772	77.35	10.89
.4	.8	29	7.610	42.42	10.83	33	8.634	47.46	12.25
.4	1.0	20	8.284	29.02	11.92	23	9.292	32.26	13.58
.4	1.2	15	8.807	21.17	13.17	17	9.845	23.42	14.76
.4	1.4	11	9.374	16.15	13.57	13	10.333	17.80	15.74
.4	1.6	9	9.741	12.70	14.68	10	10.827	13.99	16.30
.4	1.8	7	10.199	10.26	15.25	8	11.241	11.27	17.04

Table 6. (Cont'd.)

		L_0 = 20,000				L_0 = 40,000			
ρ	k	n	$\chi^2_{1,\alpha}$	L_1	λ	n	$\chi^2_{1,\alpha}$	L_1	λ
-.4	.2	67	8.606	91.34	12.84	74	9.691	100.37	14.24
-.4	.4	23	10.555	30.20	16.46	25	11.698	32.54	18.03
-.4	.6	13	11.625	16.36	19.60	14	12.763	17.41	21.32
-.4	.8	9	12.304	10.83	22.86	9	13.542	11.48	22.86
-.4	1.0	7	12.763	8.00	26.67	7	14.078	8.40	26.67
-.4	1.2	5	13.361	6.30	25.92	6	14.364	6.64	32.07
-.4	1.4	4	13.741	5.23	27.22	5	14.700	5.45	35.28
-.4	1.6	4	13.741	4.38	35.56	4	15.106	4.55	35.56
-.4	1.8	3	14.364	3.84	32.40	3	15.612	4.11	32.40
-.2	.2	150	7.149	218.91	10.02	173	8.142	245.63	11.56
-.2	.4	50	9.140	68.74	13.42	55	10.232	75.49	14.76
-.2	.6	25	10.404	34.49	15.21	28	11.488	37.53	17.01
-.2	.8	16	11.241	21.14	17.45	17	12.409	22.86	18.51
-.2	1.0	11	11.934	14.49	18.95	12	13.040	15.56	20.61
-.2	1.2	8	12.520	10.68	20.11	9	13.542	11.40	22.50
-.2	1.4	6	13.040	8.31	20.88	7	14.078	8.82	24.12
-.2	1.6	5	13.361	6.64	23.04	6	14.364	7.11	27.27
-.2	1.8	4	13.741	5.52	23.82	5	14.700	5.86	29.16
.0	.2	222	6.449	331.12	8.92	259	7.412	375.17	10.40
.0	.4	74	8.425	104.96	12.00	84	9.459	116.21	13.60
.0	.6	38	9.642	52.50	14.04	42	10.737	57.55	15.48
.0	.8	23	10.555	31.86	15.36	26	11.625	34.73	17.28
.0	1.0	16	11.241	21.58	17.00	18	12.304	23.41	19.00
.0	1.2	12	11.774	15.66	18.72	13	12.896	16.90	20.16
.0	1.4	9	12.304	11.91	19.60	10	13.361	12.82	21.56
.0	1.6	7	12.763	9.39	20.48	8	13.741	10.08	23.04
.0	1.8	6	13.040	7.62	22.68	6	14.364	8.20	22.68

Table 6. (Cont'd.)

		$L_0 = 20,000$				$L_0 = 40,000$			
ρ	k	n	$\chi^2_{1,\alpha}$	L_1	λ	n	$\chi^2_{1,\alpha}$	L_1	λ
.2	.2	287	5.995	433.65	8.24	337	6.941	494.85	9.67
.2	.4	98	7.912	138.97	11.35	110	8.966	154.70	12.72
.2	.6	50	9.140	69.75	13.19	56	10.199	76.76	14.73
.2	.8	31	10.014	42.36	14.76	34	11.128	46.35	16.13
.2	1.0	21	10.737	28.63	15.92	23	11.852	31.18	17.35
.2	1.2	15	11.360	20.71	16.76	17	12.409	22.47	18.81
.2	1.4	12	11.774	15.71	18.61	13	12.896	16.99	20.01
.2	1.6	9	12.304	12.32	18.82	10	13.361	13.28	20.65
.2	1.8	7	12.763	9.94	19.20	8	13.741	10.66	21.51
.4	.2	346	5.666	529.31	7.72	409	6.595	607.46	9.12
.4	.4	118	7.579	171.32	10.62	135	8.593	191.47	12.13
.4	.6	61	8.777	86.32	12.49	69	9.818	95.34	14.09
.4	.8	38	9.642	52.53	14.03	42	10.737	57.64	15.45
.4	1.0	26	10.333	35.51	15.25	29	11.423	38.81	16.92
.4	1.2	19	10.922	25.70	16.36	21	12.020	27.98	17.96
.4	1.4	14	11.488	19.48	16.83	16	12.520	21.15	19.01
.4	1.6	11	11.934	15.27	17.72	12	13.040	16.54	19.14
.4	1.8	9	12.304	12.28	18.84	10	13.361	13.27	20.63

series. However, the results of Table 6 indicate otherwise. The phenomenon of a larger sample size for positive ρ is partly explained by examining the behavior of λ, as presented in equation (214). Recall that λ is a generalized measure of distance of how far the true mean is from the nominal value. For fixed k and n, λ is larger for negative ρ than it is for positive ρ. Thus, since the generalized distance is smaller for positive ρ, we need a larger sample size to detect this smaller shift.

One would also use the r-dependent central limit theorem to determine the economic parameters for second and higher-order serial correlation.

We now continue our investigation of average run lengths for one quality characteristic in the presence of correlated observations by looking at the $\hat{\mu}$-chart. In this case, control limits were of the form $\mu_0 \pm B\sqrt{1/\underset{\sim}{j}^t{}_n \Lambda_n \underset{\sim}{j}_n}$ where $B = z_{\alpha/2}$ and Λ_n is the inverse of the variance-covariance matrix of the elements of the sample. Recall that this chart is valid for detecting departures from μ_0 when there is <u>any</u> type of autocorrelative structure. Under these conditions,

$$P(m) = P(\hat{\mu} > \mu_0 + B\sqrt{1/\underset{\sim}{j}^t{}_n \Lambda_n \underset{\sim}{j}_n}|m) + P(\hat{\mu} < \mu_0 - B\sqrt{1/\underset{\sim}{j}^t{}_n \Lambda_n \underset{\sim}{j}_n}|m)$$

$$= P(Z > B + a) + P(Z < -B + a)$$

$$= P(\chi'^2_{1,\lambda} > B^2) ,$$

where $\lambda = a^2 = (\mu_0 - m)^2 (\underset{\sim}{j}^t{}_n \Lambda_n \underset{\sim}{j}_n)$ and B^2 is notation for $\chi^2_{1,\alpha}$. Since $P(\mu_0) = P(\chi_1{}^2 > \chi^2_{1,\alpha})$, it follows that

$$L_0 = n/P(\chi_1^{\,2} > \chi_{1,\alpha}^{\,2}) . \tag{215}$$

Let $k > 0$ be such that a shift in the mean m of amount equal to or greater than $k\sigma_c$ is serious. When $m = \mu_0 + k\sigma_c$, $P(\mu_0 + k\sigma_c) = P(\chi'^{\,2}_{1,\lambda} > \chi_{1,\alpha}^{\,2})$ and

$$L_1 = n/P(\chi'^{\,2}_{1,\lambda} > \chi_{1,\alpha}^{\,2}) , \tag{216}$$

where

$$\lambda = (k\sigma^2_{\ c})(\underset{\sim}{j}^t_{\ n}\ \Lambda_n\ \underset{\sim}{j}_n) \tag{217}$$

In looking at equation (217) for the noncentrality parameter, two difficulties seem to arise. First, the noncentrality parameter seems to be a function of $\sigma^2_{\ c}$. Secondly, the noncentrality parameter involves the calculation of an inverse since $\Lambda_n = \Sigma_n^{-1}$. To investigate the specific nature of these difficulties, let us consider first-order serial correlation.

The variance-covariance matrix of first-order serial correlation was presented in equation (11) and can be written as

$$\Sigma_n = \sigma^2_{\ c} \begin{bmatrix} 1 & \rho & 0 & \cdots & 0 \\ \rho & 1 & \rho & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = \sigma^2_{\ c}\ C_n ,$$

by factoring out the $\sigma^2_{\ c}$. It immediately follows that $\Lambda_n = \Sigma_n^{-1} = (1/\sigma^2_{\ c})\ C_n^{\ -1}$ and

$$\lambda = k(\underset{\sim}{j}^t_n\, C_n^{-1}\, \underset{\sim}{j}_n) , \tag{218}$$

This allays our first concern that λ depends on σ^2_c. However, equation (218) very definitely points out that, for fixed k and ρ, λ cannot be determined until C_n^{-1} has been obtained. This chore is somewhat facilitated by the work of Abraham and Weiss [1]. Let $(C_n^{-1})_{ij}$ denote the $(i,j)^{th}$ element of C_n^{-1}. They have shown that

$$(C_n^{-1})_{ij} = a\ b, \tag{219}$$

where

$$a = \frac{\dfrac{1}{\sqrt{1-4\rho^2}}\left[\left(\dfrac{\sqrt{1-4\rho^2}-1}{2\rho}\right)^{2j} - 1\right]}{\left(\dfrac{\sqrt{1-4\rho^2}-1}{2\rho}\right)^{i+j}}$$

and

$$b = \frac{\left(\dfrac{\sqrt{1-4\rho^2}-1}{2\rho}\right)^{2(n+1)} - \left(\dfrac{\sqrt{1-4\rho^2}-1}{2\rho}\right)^{2i}}{1 - \left(\dfrac{\sqrt{1-4\rho^2}-1}{2\rho}\right)^{2(n+1)}} ,$$

for $i \geq j = 1,\ldots, n$.

In order to find that n and $\chi^2_{1,\alpha}$ which minimize L_1 by the success-failure search procedure, we fix L_0 and let $n = n^{(1)}$, in which case $\chi^{2\ (1)}_{1,\alpha}$ is determined from equation (215). Since ρ is fixed and n is temporarily fixed at $n^{(1)}$, $C_{n(1)}^{-1}$ is obtained using equation (219), and $\lambda^{(1)}$ is determined in accordance with equation (218). This immediately enables us to evaluate $L_1^{(1)}$ from equation (216).

We now set $n = n^{(2)}$ and proceed as outlined above. Again, we must find a $C_{n^{(2)}}^{-1}$. Because each iteration requires finding an inverse and each line of output requires many iterations, we will not pursue this further at this time.

In this chapter, we have determined the economic parameters (n and the control chart constant) by using the scheme of minimizing L_1 for a large fixed value of L_0. This was done for three cases, where the first case merely reviewed Page's work for one quality characteristic and independent observations. The second case extended Page's scheme to two and three quality characteristics with independent observations. One general conclusion was that a χ^2_p-chart requires smaller sample sizes than p $\bar{X}$-charts. The third case extended Page's scheme for one quality characteristic by allowing first-order serial correlation between the observations. In general, negatively correlated observations yielded the most favorable result.

CHAPTER V

THE MULTIVARIATE, MULTI-CONSEQUENCE INTERVENTION MODEL

In Chapter III, we introduced the multi-consequence intervention model for a univariate time series where the observations occur at n equispaced epochs. The "multi-consequence" terminology refers to the fact that the intervention may have affected the variability-covariability of the process. In Chapter II, we presented the concept of a vector-valued or multivariate ARMA model to represent various types of correlation structure across the vectors of observations. This chapter extends the material presented in Chapters II and III by considering the multivariate, multi-consequence intervention model, its properties, and its estimation.

The need for such a model becomes obvious where one realizes that the introduction of an intervention in some geographic region has the potential to affect not only the occurrences of that particular region but also the occurrences of contiguous regions. For example, when Connecticut instituted its speeding crackdown in 1955 (see Glass [27]), the monthly fatalities per 100,000,000 miles may have also been affected for the states of New York, Massachusetts, Rhode Island, and New Jersey. To assess the simultaneous impact of the speeding crackdown, it may have been prudent to monitor $\underset{\sim}{Z}_t = [Z_{1t}, Z_{2t}, Z_{3t}, Z_{4t}, Z_{5t}]^t$, where Z_{it} represents the monthly fatalities per 100 million miles for each of the five states. Correlation may exist within $\underset{\sim}{Z}_t$, such as between Z_{it} and $Z_{i't}$, as well as across the $\underset{\sim}{Z}_t$'s.

This chapter will specifically address the p-variate multi-consequence intervention model for a first-order moving average process. The extension to higher order moving average processes is obvious.

Selected portions of this chapter appear in a paper by Alt and Deutsch [3].

5.1 Properties of the Multivariate, Multi-Consequence Intervention Model

5.1.1 Model Description

The bivariate, first-order moving average process was presented in equation (59), namely,

$$\begin{aligned} Z_{1t} &= \mu_1 - \theta_{11}a_{1,t-1} - \theta_{12}a_{2,t-1} + a_{1t} \\ Z_{2t} &= \mu_2 - \theta_{21}a_{1,t-1} - \theta_{22}a_{2,t-1} + a_{2t} \;. \end{aligned}$$

Note that, when $\theta_{12} = \theta_{21} = 0$, each equation describes a univariate MA(1) process. In order to accommodate a constant, continuous intervention effect commencing with the $(n_1 + 1)$ th observation, we modify equation (59) as follows:

$$\left.\begin{aligned} Z_{1t} &= \mu_1 - \theta_{11}a_{1,t-1} - \theta_{12}a_{2,t-1} + a_{1t} \\ Z_{2t} &= \mu_2 - \theta_{21}a_{1,t-1} - \theta_{22}a_{2,t-1} + a_{2t} \;, \end{aligned}\right\} \quad (220a)$$

for $t = 1,\ldots, n_1$, and

$$\left.\begin{aligned} Z_{1t} &= \mu_1 + \delta_1 - \theta_{11}a_{1,t-1} - \theta_{12}a_{2,t-1} + a_{1t} \\ Z_{2t} &= \mu_2 + \delta_2 - \theta_{21}a_{1,t-1} - \theta_{22}a_{2,t-1} + a_{2t} \;, \end{aligned}\right\} \quad (220b)$$

for $t = n_1 + 1,\ldots, n_1 + n_2$. Equations (220a) and (220b) have the following matrix representation:

$$\left.\begin{aligned} \underset{\sim}{Z}_t &= \underset{\sim}{\mu} - \Theta\underset{\sim}{a}_{t-1} + \underset{\sim}{a}_t, \quad t = 1,\ldots, n_1 \\ \underset{\sim}{Z}_t &= \underset{\sim}{\mu} + \underset{\sim}{\delta} - \Theta\underset{\sim}{a}_{t-1} + \underset{\sim}{a}_t, \quad t = n_1{+}1,\ldots, n_1{+}n_2, \end{aligned}\right\} \quad (221)$$

where

$$\underset{\sim}{Z}_t = \begin{bmatrix} Z_{1t} \\ Z_{2t} \end{bmatrix}, \quad \underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \underset{\sim}{\delta} = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}, \quad \Theta = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}.$$

It is assumed that $\underset{\sim}{a}_t \sim NID_2\ (\underset{\sim}{0},\ \Sigma_{\underset{\sim}{a}})$; that is, the $\underset{\sim}{a}_t$'s are bivariate normally distributed random variables and they are uncorrelated across time. Note that the matrix formulation presented in equation (221), even though specifically developed for the bivariate case, also represent the multivariate or p-variate, single consequence intervention model. In the p-variate case, $\underset{\sim}{Z}_t$, $\underset{\sim}{\mu}$, and $\underset{\sim}{\delta}$ are each (p x 1) vectors and Θ is a (p x p) matrix.

Equation (221), which represents the single consequence intervention model for a multivariate, first-order moving average process, can be used as a basis in the formulation of the <u>multivariate, multi-consequence intervention model</u>, hereafter designated by $MMA_I(1)$. Let us first consider the bivariate case. For t = 1,..., n_1, the model is identical with that presented in equation (220a). However, since the intervention may have affected the post-intervention moving average parameters as well as the level, equation (220b) becomes modified as follows:

$$\left.\begin{aligned} Z_{1t} &= \mu_1 + \delta_1 - \psi_{11}a_{1,t-1} - \psi_{12}a_{2,t-1} + a_{1t} \\ Z_{2t} &= \mu_2 + \delta_2 - \psi_{21}a_{1,t-1} - \psi_{22}a_{2,t-1} + a_{2t}. \end{aligned}\right\} \quad (222)$$

Thus, the $MMA_I(1)$ model has the following matrix formulation:

$$\left.\begin{aligned} \underset{\sim}{Z}_t &= \underset{\sim}{\mu} - \Theta\underset{\sim}{a}_{t-1} + \underset{\sim}{a}_t,\ t = 1,\ldots,n_1 \\ \underset{\sim}{Z}_t &= \underset{\sim}{\mu} + \underset{\sim}{\delta} - \Psi\underset{\sim}{a}_{t-1} + \underset{\sim}{a}_t,\ t = n_1+1,\ldots,n_1+n_2 \end{aligned}\right\} \quad (223)$$

The matrix formulation of $MMA_I(q)$, q > 1, models is straightforward.

Let us now investigate the properties of the $MMA_I(1)$ model.

5.1.2. Properties of $MMA_I(1)$ Model

Since $E(\underset{\sim}{a}_t) = \underset{\sim}{0}$, it follows that

and
$$\left.\begin{aligned} E(\underset{\sim}{Z}_t) &= \underset{\sim}{\mu},\ t = 1,\ldots, n_1 \\ E(\underset{\sim}{Z}_t) &= \underset{\sim}{\mu} + \underset{\sim}{\delta},\ t = n_1 + 1,\ldots, n_1 + n_2 \ . \end{aligned}\right\} \quad (224)$$

Note that this is also the expectation of the single consequence intervention model, described in equation (221).

Let $\Gamma(h)$ denote the (p x p) covariance matrix of $\underset{\sim}{Z}_t$ and $\underset{\sim}{Z}_{t+h}$ prior to the intervention. That is, $\Gamma(h) = E[(\underset{\sim}{Z}_t - E(\underset{\sim}{Z}_t))(\underset{\sim}{Z}_{t+h} - E(\underset{\sim}{Z}_{t+h}))^t]$ and $\Gamma(-h) = E[(\underset{\sim}{Z}_t - E(\underset{\sim}{Z}_t))(\underset{\sim}{Z}_{t-h} - E(\underset{\sim}{Z}_{t-h}))^t]$. Now, for $t = 1,\ldots, n_1 - 1$,

$$\begin{aligned} \Gamma(h) &= E[(\underset{\sim}{Z}_t - \underset{\sim}{\mu})(\underset{\sim}{Z}_{t+h} - \underset{\sim}{\mu})^t] \\ &= E[(\underset{\sim}{\mu} - \Theta\underset{\sim}{a}_{t-1} + \underset{\sim}{a}_t - \underset{\sim}{\mu})(\underset{\sim}{\mu} - \Theta\, \underset{\sim}{a}_{t+h-1} + \underset{\sim}{a}_{t+h} - \underset{\sim}{\mu})^t] \\ &= E(\Theta\, \underset{\sim}{a}_{t-1}\underset{\sim}{a}_{t+h-1}^t\Theta^t - \Theta\underset{\sim}{a}_{t-1}\underset{\sim}{a}_{t+h}^t - \underset{\sim}{a}_t\underset{\sim}{a}_{t+h-1}^t\Theta^t + \underset{\sim}{a}_t\underset{\sim}{a}_{t+h}^t) \ . \end{aligned}$$

And, it immediately follows that, for $t = 1,\ldots, n_1 - 1$,

$$\Gamma(h) = \left\{\begin{array}{ll} \underset{\sim}{\Sigma}_a + \Theta\underset{\sim}{\Sigma}_a\, \Theta^t & , h = 0 \\ -\underset{\sim}{\Sigma}_a\, \Theta^t & , h = 1 \\ -\Theta\underset{\sim}{\Sigma}_a & , h = -1 \\ 0 & , \text{otherwise} \end{array}\right\} \quad (225a)$$

Furthermore,

$$\begin{aligned} E[(\underset{\sim}{Z}_{n_1} - \underset{\sim}{\mu})(\underset{\sim}{Z}_{n_1+1} - \underset{\sim}{\mu} - \underset{\sim}{\delta})^t] &= E[(-\Theta\underset{\sim}{a}_{n_1-1} + \underset{\sim}{a}_{n_1})(-\Psi\underset{\sim}{a}_{n_1} + \underset{\sim}{a}_{n_1+1})^t] \\ &= E(-\underset{\sim}{a}_{n_1}\underset{\sim}{a}_{n_1}^t\, \Psi^t) \\ &= -\underset{\sim}{\Sigma}_a\, \Psi^t \ . \end{aligned} \quad (225b)$$

Let $\Gamma^*(h)$ denote the (p x p) covariance matrix of $\underset{\sim}{Z}_t$ and $\underset{\sim}{Z}_{t+h}$ after the intervention. Then, for $t = n_1 + 1,\ldots, n_1 + n_2 = n$,

$$\Gamma^*(h) = \begin{cases} \Sigma_a + \Psi\Sigma_a\Psi^t & , h = 0 \\ -\Sigma_a\Psi^t & , h = 1 \\ -\Psi\Sigma_a & , h = -1 \\ 0 & , \text{otherwise} \end{cases} \tag{225c}$$

The expected value and covariance properties of the $MMA_I(1)$ model can be written in an alternate format. To do this, let the n sample elements be denoted by $Z_1, \ldots, Z_{n_1}, Z_{n_1+1}, \ldots Z_{n_1+n_2}$ where each Z_i is a (p x 1) vector. Let Z denote the (np x 1) vector of sample elements, where $Z^t = [Z_1^t, \ldots, Z_{n_1}^t, Z_{n_1+1}^t, \ldots, Z_n] = [Z_{11}, Z_{p1}, \ldots, Z_{1n_1}, \ldots, Z_{pn_1}, Z_{1,n_1+1}, \ldots, Z_{p,n_1+1}, \ldots, Z_{1,n}, \ldots, Z_{p,n}]$. Let μ_Z denote $E(Z)$, where this (np x 1) vector is given by

$$\mu_Z = [\mu^t, \ldots, \mu^t, (\mu + \delta)^t, \ldots, (\mu + \delta)^t] \ , \tag{226}$$

with $\mu^t = [\mu_1, \ldots, \mu_p]$. As in Chapter II, let $A \otimes B$ denote the direct product of the matrices A and B. Thus, equation (226) can be written as

$$\mu_Z = (j_n \otimes I_p)\ \mu + (k_n \otimes I_p)\ \delta \ , \tag{227}$$

where k_n is an (n x 1) vector which has 0's for its first n_1 entries followed by n_2 1's. If Σ_Z denotes the (np x np) covariance matrix of Z, then Σ_Z may be partitioned as follows:

$$\left[\begin{array}{ccccc|ccccc}
\Gamma(0) & \Gamma(1) & 0 \cdots & 0 & & 0 & 0 & 0 & \cdots & 0 \\
\Gamma(-1) & \Gamma(0) & \Gamma(1) \cdots & 0 & & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 \cdots & \Gamma(0) & & \Gamma^*(1) & 0 & 0 & \cdots & 0 \\
\hline
0 & 0 & 0 \cdots & \Gamma^*(-1) & & \Gamma^*(0) & \Gamma^*(1) & 0 & \cdots & 0 \\
0 & 0 & 0 \cdots & 0 & & \Gamma^*(-1) & \Gamma^*(0) & \Gamma^*(1) & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 \cdots & 0 & & 0 & 0 & 0 & \cdots & \Gamma^*(0)
\end{array}\right] \tag{228}$$

This patterned structure illustrates that the memory of a $MMA_I(1)$ model is only one period long.

Note that the expected value and covariance structure of the $MMA_I(1)$ model presented in equations (227) and (228) is very similar to that of the quality control model presented in Section 2.2.2, when there is a first-order serial correlation. One of the biggest differences is in their philosophy. In the intervention model we wish to test, based on a single sample of size n, whether the intervention has shifted the level from $\underset{\sim}{\mu}$ to $\underset{\sim}{\mu} + \underset{\sim}{\delta}$ commencing with the (n_1+1)th observation, and our primary interest is in the magnitudes and directions of the components of the shift vector $\underset{\sim}{\delta}$. In the quality control model, we wish to test whether the process mean remains at a nominal value $\underset{\sim}{\mu}_0$ in repeated samples of size n. Furthermore, in the multivariate intervention model, initial interest is also centered on the possible change in covariance structure accompanying the introduction of the intervention.

The distribution of $\underset{\sim}{Z}$ will now be investigated. Let

$$\underset{\sim}{a}^t = [\underset{\sim}{a}_o^t, \underset{\sim}{a}_1^t, \ldots, \underset{\sim}{a}_{n_1}^t, \underset{\sim}{a}_{n_1+1}^t, \ldots, \underset{\sim}{a}_{n_1+n_2}^t] ,$$

where $\underset{\sim}{a}_i^t = [a_{1i}, a_{2i}, \ldots, a_{pi}]$. Thus, $\underset{\sim}{a}$ is an $((n+1)\ p \times 1)$ vector. Since $\underset{\sim}{Z} = C\, \underset{\sim}{a} + \underset{\sim}{\mu}_z$, where C is an $(np \times (n+1)p)$ matrix, it follows that $\underset{\sim}{Z}$ is distributed as an np-variate normal. These properties can be summarized by saying that

$$\underset{\sim}{Z} \sim N_{np}\ (\underset{\sim}{\mu}_{\underset{\sim}{Z}}, \underset{\sim}{\Sigma}_{\underset{\sim}{Z}}) \tag{229}$$

where $\underset{\sim}{\mu}_{\underset{\sim}{Z}}$ is given in equation (227) and $\underset{\sim}{\Sigma}_{\underset{\sim}{Z}}$ is given in equation (228).

In order to satisfy invertibility conditions, constraints must be placed on the elements of the Θ and Ψ matrices. Specifically, we require that the p roots of the determinantal equation

$$|Im - \Theta| = 0 \tag{230a}$$

be less than one in absolute value and also that the p roots of the determinantal equation

$$|Im - \Psi| = 0 \tag{240b}$$

be less than one in absolute value. Thus, similar to the univariate multi-consequence $MA_I(1)$ model, two sets of invertibility conditions are required. For p = 2, the invertibility region of each component of the $MMA_I(1)$ model is shown in Figure 5 (see Chapter II).

Additional aspects of non-intervention vector-valued time series can be found in Hannan [36].

5.2 Least Squares Estimation of the $MMA_I(1)$ Model

In the previous section, a detailed explanation was presented of the multivariate, multi-consequence, first-order moving average intervention model ($MMA_I(1)$) and its properties. In this section, we will be concerned with parameter estimation for this model with primary concern directed towards the estimation of $\underset{\sim}{\mu}$ and $\underset{\sim}{\delta}$. The least squares estimates of $\underset{\sim}{\mu}$ and $\underset{\sim}{\delta}$ are obtained by transforming the original $\underset{\sim}{Z}_t$'s to $\underset{\sim}{Y}_t$'s which are amenable to statistical linear model analysis. We shall see that the least square estimates of $\underset{\sim}{\mu}$ and $\underset{\sim}{\delta}$ are directly dependent upon Θ and Ψ. As in the univariate case, we employ an iterative technique of searching on the elements of Θ and Ψ until those values are found which minimize the residual sum of squares of the $\underset{\sim}{Y}_t$'s. However, before demonstrating the least squares estimation procedure, it may be helpful to review the basic concepts of the multivariate classical linear regression model.

5.2.1 Multivariate Linear Regression Model

In this section, we review some of the concepts of the multivariate, multiple linear regression model. A more detailed treatment of this material can be found in Goldberger [29].

The multivariate linear regression model has the following matrix formulation:

$$Y = X\,B + E\ , \tag{231}$$

where

$$Y = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{bmatrix} = [\underset{\sim}{Y}_1,\ \underset{\sim}{Y}_2, \ldots,\ \underset{\sim}{Y}_p]\ ,$$

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} = [\underset{\sim}{x}_1,\ \underset{\sim}{x}_2, \ldots,\ \underset{\sim}{x}_k]\ ,$$

$$B = \begin{bmatrix} \beta_{11} & \beta_{12} & & \beta_{1p} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2p} \\ \vdots & \vdots & & \vdots \\ \beta_{k1} & \beta_{k2} & & \beta_{kp} \end{bmatrix} = [\underset{\sim}{\beta}_1,\ \underset{\sim}{\beta}_2, \ldots,\ \underset{\sim}{\beta}_p],$$

and

$$E = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} \cdots \varepsilon_{1p} \\ \varepsilon_{21} & \varepsilon_{22} \cdots \varepsilon_{2p} \\ \vdots & \vdots \quad\quad \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} \cdots \varepsilon_{np} \end{bmatrix} = [\underset{\sim}{\varepsilon}_1, \underset{\sim}{\varepsilon}_2, \ldots, \underset{\sim}{\varepsilon}_p] \ .$$

This formulation arises when there are n observations on each of p variables $Y_1, \ldots, Y_p$ for each of k variables $x_1, \ldots, x_k$. Each row of Y corresponds to a single joint observation. Equation (231) imples that for each $\underset{\sim}{Y}_j$, j = 1,..., p, there is a relation or model of the form

$$\underset{\sim}{Y}_j = X\underset{\sim}{\beta}_j + \underset{\sim}{\varepsilon}_j \ ,$$

which is merely the univariate multiple regression model. Thus, each column of Y refers to one of the p relations.

Let the (kxp) matrix $\hat{B}$ denote the estimate of B. Let $W = E^tE$. Thus, W is a (pxp) matrix with entries $w_{ij} = \underset{\sim}{\varepsilon}_i^t \underset{\sim}{\varepsilon}_j$. In order to estimate B, we minimize the trace of W denoted by trW. Goldberger [29] shows that this is equivalent to minimizing $|n^{-1}W|$, the generalized error variance. He also points out that minimizing trW is equivalent to fitting each of the p relations ($\underset{\sim}{Y}_j = X\underset{\sim}{\beta}_j + \underset{\sim}{\varepsilon}_j$) by the least-squares criterion, which leads to $(X^tX)\underset{\sim}{b}_j = X^t\underset{\sim}{Y}_j$, j=1,..., p. The resulting normal equation for all p relations are

$$X^tX\hat{B} = X^tY \ ,$$

and the estimate of B is the (kxp) matrix

$$\hat{B} = (X^tX)^{-1} X^tY \ , \tag{232}$$

if the rank of X is k. We again point out that the columns of $\hat{B}$ could have been generated by fitting p univariate multiple regression relations.

The point estimates of the elements of the B matrix, given in equation (232), can be used as a starting point in the derivation of corresponding interval estimates. However, in order to derive interval estimates, it is necessary to impose a distributional property on the error matrix E. Let $\underset{\sim}{\varepsilon}_i^t$ denote the i th row of E, for i = 1,..., n. We assume that $\underset{\sim}{\varepsilon}_i^t \sim N_p(\underset{\sim}{0}^t, \Sigma)$. We also adopt the univariate multiple regression concept that the errors are independent across observations, or $\underset{\sim}{\varepsilon}_j \sim N_n(\underset{\sim}{0}, \sigma_j^2 I_n)$ for j = 1,..., n. In order to combine the row and column distributional assumptions, let $\underset{\sim}{E}$ denote the (nxp) matrix E written in dictionary form. Then $\underset{\sim}{E} \sim N_{np}(\underset{\sim}{0}, I_n \otimes \Sigma)$. It immediately follows that the expectation and variance-covariance matrix of $\hat{\underset{\sim}{B}}$, which denotes $\hat{B}$ written in dictionary form, are $\underset{\sim}{\mu}_{\hat{\underset{\sim}{B}}} = \underset{\sim}{B}$ and $C(\hat{\underset{\sim}{B}}, \hat{\underset{\sim}{B}}^t) = (X^tX)^{-1} \otimes \Sigma$. Let c^{ij} denote the $(i,j)\underline{th}$ element of $(X^tX)^{-1}$. Thus, $c^{ii}\ \Sigma$ is the variance-covariance matrix of $\hat{\underset{\sim}{\beta}}_i^t$, which is the i th row of $\hat{B}$, and $\sigma_j^2(X^tX)^{-1}$ is the variance-covariance matrix of $\hat{\underset{\sim}{\beta}}_j$, which is the j th column of $\hat{B}$. It also follows that the elements of $\hat{B}$ are normally distributed since they are linear combinations of the elements of Y. The estimate of Σ, denoted by $\hat{\Sigma}$, is given by

$$\hat{\Sigma} = (Y-X\hat{B})^t\ (Y-X\hat{B})/(n-k).$$

It follows from the above statements that a $100(1-\alpha)\%$ confidence interval for β_{ij}, i=1,..., k, j=1,..., p, is given by $\hat{\beta}_{ij} \pm t_{1-\alpha/2,n-k}\hat{\sigma}_j(c^{ii})^{\frac{1}{2}}$ where $\hat{\sigma}_j$ is the square root of the j th diagonal element of $\hat{\Sigma}$. Additional types of confidence intervals can be constructed as the need arises.

The multivariate linear regression model provides the basis for estimating $\underset{\sim}{\mu}$ and $\underset{\sim}{\delta}$ in the $MMA_I(1)$ model. As with the univariate case of

Chapter III, it is necessary to transform the $MMA_I(1)$ model into the format of a multivariate linear regression model.

5.2.2 Least Squares Estimation Procedure

As it now stands, the $MMA_I(1)$ model presented in equation (223) is not in a multivariate linear regression format. The transformation necessary to convert equation (223) into linear model format can be found by considering the first few $\mathbf{z}_t$'s. Specifically, $\mathbf{z}_1 = \boldsymbol{\mu} + \mathbf{a}_1 - \Theta\mathbf{a}_0$. Thus, $\mathbf{z}_1$ depends on both a current and previous error vector. An obvious way to alleviate this is to let $\mathbf{a}_0 = \mathbf{0}$, its marginal mean. Upon doing this, $\mathbf{z}_1 = \boldsymbol{\mu} + \mathbf{a}_1$, and we let $\mathbf{y}_1 = \mathbf{z}_1$. Now $\mathbf{z}_2 = \boldsymbol{\mu} + \mathbf{a}_2 - \Theta\mathbf{a}_1$, and it also depends on both a current and previous error vector. However, if we multiply $\mathbf{y}_1$ by Θ and add $\Theta\mathbf{y}_1$ to $\mathbf{z}_2$, we obtain $\mathbf{z}_2 + \Theta\mathbf{y}_1 =$ $= (\boldsymbol{\mu} + \mathbf{a}_2 - \Theta\mathbf{a}_1) + (\Theta\boldsymbol{\mu} + \Theta\mathbf{a}_1) = \boldsymbol{\mu} + \Theta\boldsymbol{\mu} + \mathbf{a}_2$. Thus, we let $\mathbf{y}_2 = \mathbf{z}_2 + \Theta\mathbf{y}_1$. Furthermore, since $\mathbf{z}_3 = \boldsymbol{\mu} - \Theta\mathbf{a}_2 + \mathbf{a}_3$, we let $\mathbf{y}_3 = \mathbf{z}_3 + \Theta\mathbf{y}_2$. In general, for $t = 1,\ldots, n_1$, we employ the transformation

$$\mathbf{y}_t = \mathbf{z}_t + \Theta\, \mathbf{y}_{t-1}, \tag{233}$$

in which case

$$\mathbf{y}_t = (I + \Theta + \ldots + \Theta^{t-1})\, \boldsymbol{\mu} + \mathbf{a}_t .$$

For $t = n_1+1,\ldots, n_1 + n_2$, we again employ the transformation presented in equation (233), which results in

$$\mathbf{y}_t = (I + \Theta + \ldots + \Theta^{t-1})\, \boldsymbol{\mu} + (I + \Theta + \ldots + \Theta^{t-(n_1+1)})\, \boldsymbol{\delta} + \mathbf{a}_t .$$

By transforming the original $\mathbf{z}_t$'s into $\mathbf{y}_t$'s using equation (233), the $\mathbf{y}_t$'s when put in the appropriate format are amenable to statistical linear model analysis. Although specific formulae could be developed for the elements of $\hat{\underset{\sim}{\mu}}$ and $\hat{\underset{\sim}{\delta}}$, this will not be done at this time. However, it is important to note that $\hat{\underset{\sim}{\mu}}$ and $\hat{\underset{\sim}{\delta}}$ are conditional least squares estimates in that they are dependent on the elements of Θ and Ψ. To find the elements of $\hat{\underset{\sim}{\mu}}$ and $\hat{\underset{\sim}{\delta}}$, we search over the Ψ and Θ matrices until we find that pair (Ψ,Θ) which minimizes the squared residuals of the transformed variates.

In Chapter V, we have proposed a multivariate, multi-consequence intervention model, determined its properties, and outlined an estimation procedure for its parameters.

CHAPTER VI

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The theme of this research has been the development and investigation of the properties of tests for location in the presence of correlated observations. This theme has been investigated for the quality control scenario (monitoring a process by using repeated samples) as well as the interrupted time series quasi-experiment setting (detecting a shift in the level of a single sample).

However, this research has also investigated inferential problems concerning the process variability of the ITSQE. Another topic that was explored was the economics of sampling in the presence of correlated observations.

A more thorough summary of the results is presented in Section 6.1, followed by conclusions in Section 6.2, and recommendations for future research in Section 6.3.

6.1 Summary of Results

This section contains the results of this dissertation. The section in which the result was first presented is given in parentheses.

6.1.1 Chapter II. Control Charts for Correlated Observations

1. When there is only one quality characteristic with a standard value specified for the process mean and the autocorrelative structure among the observations is known, the maximum likelihood

estimator of μ is given by

$$\hat{\mu} = (\underset{\sim}{X}^t \Lambda_n \underset{\sim}{j}_n)/(\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n).$$

The estimator of μ was derived under the assumption that the observations are obtained from an n-dimensional multivariate normal process (Section 2.1.2).

2. Under the above conditions,

$$\hat{\mu} \sim N(\mu, (\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n)^{-1}),$$

which demonstrates that μ is unbiased (Section 2.1.2).

3. Also, $\hat{\mu}$ is the uniformly minimum variance estimator of μ. This implies that, in the class of unbiased estimators, μ is a Bayesian estimator with respect to every prior when the loss function is quadratic; and, μ is a minimax estimator when the loss function is quadratic (Section 2.1.2).

4. Control limits for the process mean are given by

$$\mu_0 \pm z_{\alpha/2} \sqrt{1/\underset{\sim}{j}_n^t \Lambda_n \underset{\sim}{j}_n} \ .$$

where $E(\hat{\mu}) = \mu_0$ (Section 2.1.2).

5. The estimator $\hat{\mu}$ is also the generalized least-squares estimator, which is a known result (Section 2.1.2).

6. In the presence of serial correlation of degree r, justification for using $\bar{X}$ is provided by the r-dependent central limit theorem (Section 4.3).

7. When there are multiple quality characteristics with standard values specified for the process mean vector and the covariance

structure within each vector of observations as well as among the vectors of observations is known, the maximum likelihood estimator of μ is given by

$$\hat{\underset{\sim}{\mu}} = [(\underset{\sim}{j}_n \otimes I_p)^t \Lambda_{\underset{\sim}{X}}(\underset{\sim}{j}_n \otimes I_p)]^{-1} (\underset{\sim}{j}_n \otimes I_p)^t \Lambda_{\underset{\sim}{X}} \underset{\sim}{X}.$$

The estimator was derived under the assumption that the observations are obtained from an np-multivariate normal process (Section 2.2.2).

8. Under the previously stated conditions,

$$\hat{\underset{\sim}{\mu}} \sim N_p(\underset{\sim}{\mu}, [(\underset{\sim}{j}_n \otimes I_p)^t \Lambda_{\underset{\sim}{X}}(\underset{\sim}{j}_n \otimes I_p)]^{-1}) \quad ,$$

showing that $\hat{\underset{\sim}{\mu}}$ is unbiased (Section 2.2.2).

9. Also, $\hat{\underset{\sim}{\mu}}$ is the uniformly minimum variance estimator of μ (Section 2.2.2).

10. If $(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)^t \Sigma_{\hat{\mu}}^{-1} (\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0) > \chi^2_{p,\alpha}$, the conclusion is that the process mean has shifted from the nominal value μ_0 (Section 2.2.2).

<u>6.1.2 Chapter III. Estimation for the Multi-Consequence Intervention Model</u>

11. A multi-consequence intervention model was proposed for first and second-order moving average processes (Section 3.1.2). For a first-order moving average process, the model is

$$Z_t = \mu + a_t - \theta_1 a_{t-1}, \; t = 1,\ldots,n_1$$

$$Z_t = \mu + \delta + a_t - \gamma_1 a_{t-1}, \; t = n_1 + 1,\ldots,n_1 + n_2 \; .$$

The expected value and covariance structure of these models was considered in detail (Section 3.1.2).

12. Specific formulas were developed for conditional least squares estimators of μ and δ for both the single and multi-consequence $MA_I(1)$ models (Sections 3.2.2.1 and 3.2.2.2), and a computer program was written to accomplish this (Appendix C).
13. A procedure was indicated for obtaining conditional least squares estimates of μ and δ for both the single and multi-consequence $MA_I(2)$ models (Section 3.2.2.3).
14. Explicit expressions were obtained for the maximum likelihood estimators of μ and δ for any $MA_I(q)$ model. These estimators are for fixed values of the moving average parameters (Section 3.3.1).
15. An algorithm was developed for calculating the unconditional likelihood function of the single and multi-consequence $MA_I(1)$ models for a given set of parameter values (Section 3.3.2.1 and 3.3.2.2).
16. A procedure was indicated for calculating the unconditional likelihood function of the single and multi-consequence $MA_I(2)$ models for a given set of parameter values (Section 3.3.2.3).
17. Explicit instructions were given for implementing the maximum likelihood estimation algorithm (Section 3.3.3), and a computer program was written to accomplish this (Appendix D) for the multi-consequence $MA_I(1)$ model.

18. A likelihood ratio test was proposed to test $H_0:\theta_1 = \gamma_1$ for the $MA_I(1)$ model (Section 3.3.4). The outcome of this test influences the statistical inferential procedure to be used for δ.

6.1.3 Chapter IV. Economic Aspects of Control Charts for the Mean.

19. Page's scheme for the determination of the sample size and control chart constant was extended to the case of two quality characteristics and independent observations (Section 4.2).
20. Page's scheme was also investigated for three quality characteristics and independent observations (Section 4.2),
21. Using a modified $\bar{X}$-chart with justification provided by the r-dependent central limit theorem, Page's scheme was employed to determine the sample size and control chart constant needed for one quality characteristic when the observations have a first-order serial correlation.

6.1.4 Chapter V. The Multivariate Multi-Consequence Intervention Model

22. The univariate multi-consequence intervention model of Chapter III was extended to include vector-valued moving average processes. That is, at each epoch of time, the sample element is a vector $\underset{\sim}{Z}_t = [Z_{1t}, Z_{2t}, \ldots, Z_{pt}]^t$ where the elements comprising $\underset{\sim}{Z}_t$ may be correlated. The model for a bivariate, first-order moving average, multi-consequence intervention model is

$$Z_{1t} = \mu_1 - \theta_{11} a_{1,t-1} - \theta_{12} a_{2,t-1} + a_{1t}$$

$$Z_{2t} = \mu_2 - \theta_{21} a_{1,t-1} - \theta_{22} a_{2,t-1} + a_{2t}, \quad t = 1,\ldots,n_1$$

$$Z_{1t} = \mu_1 + \delta_1 - \psi_{11} a_{1,t-1} - \psi_{12} a_{2,t-1} + a_{1t}$$

$$Z_{2t} = \mu_2 + \delta_2 - \psi_{21} a_{1,t-1} - \psi_{22} a_{2,t-1} + a_{2t}, \quad t = n_1+1,\ldots,n_1+n_2$$

In matrix notation, this becomes

$$\underset{\sim}{Z}_t = \underset{\sim}{\mu} - \Theta \underset{\sim}{a}_{t-1} + \underset{\sim}{a}_t, \quad t = 1,\ldots\ldots,n_1$$

$$\underset{\sim}{Z}_t = \underset{\sim}{\mu} + \underset{\sim}{\delta} - \Psi \underset{\sim}{a}_{t-1} + \underset{\sim}{a}_t, \quad t = n_1 + 1,\ldots, n_1 + n_2$$

The expected value and covariance properties of this model were explored (Section 5.1).

23. A procedure was indicated for obtaining conditional least squares estimates of the level and shift parameters for the bivariate, first-order moving average, multi-consequence intervention model (Section 5.2).

6.2 Conclusions

This section contains conclusions arising from this research.

1. Whether there be one or multiple quality characteristics, the maximum likelihood estimator of the process mean is valid for any type of autocorrelative structure and is the uniformly minimum variance unbiased estimator of the process mean.
2. The multi-consequence intervention models offer a new type of flexibility for modeling the interrupted time series quasi experiment (ITSQE) which will also reduce the residual variance.

3. One should always test the equality of the pre-intervention and post-intervention moving average parameters since the estimates of the process level and shift are contingent upon them.
4. When the process quality depends on several quality characteristics and there are independent vectors of observations, the use of a single χ_p^2-chart instead of p $\bar{X}$-charts generally decreases the sample size that needs to be selected.
5. When the process quality depends on only one quality chracteristic and the observations are correlated, the presence of negative autocorrelation results in the selection of a smaller sample size.

6.3 Recommendations for Future Research

Some perceptions on future reserach are:

1. The concept of the multi-consequence intervention model needs to be extended to pure autoregressive processes and autoregressive-moving average processes. Needless to say, the maximum likelihood and least squares estimation procedures also need to be extended.
2. There is a need to consider nonstationary multiconsequence intervention models and their estimation because of their proven applicability.
3. The maximum likelihood technique of parameter estimation needs to be extended to the multivariate intervention model.
4. Economic parameters need to be determined for the $\hat{\mu}$-chart.

APPENDIX A

DATA FOR EXAMPLE 2.1

Appendix A

This is the data for Example 2.1. There are twenty samples, each of size 5, from a univariate process with first-order serial correlation equal to 0.47, variance equal to 13.41 and process mean equal to 30.0.

Sample Number	x_1	x_2	x_3	x_4	x_5	$\bar{x}$	$\hat{\mu}$
1	26.149	25.392	28.910	32.657	35.011	29.624	29.934
2	28.231	24.838	27.843	28.561	28.623	27.619	28.059
3	35.435	39.090	39.970	32.520	31.688	35.741	35.590
4	33.580	31.120	35.358	31.376	27.246	31.736	31.870
5	33.713	33.469	32.428	33.817	28.548	32.395	31.789
6	27.596	31.966	34.198	30.636	30.042	30.888	30.594
7	28.767	30.783	34.496	31.857	28.872	30.955	30.678
8	27.030	27.533	35.621	39.204	28.235	31.525	30.515
9	30.723	29.506	27.992	24.209	24.849	27.456	27.731
10	22.982	29.768	29.875	26.188	24.082	26.579	25.808
11	31.330	33.887	27.948	26.095	32.754	30.403	30.672
12	34.693	34.548	31.036	30.853	32.080	32.642	32.659
13	23.057	21.952	26.784	26.991	28.965	25.550	26.037
14	33.141	32.665	25.366	23.228	28.808	28.642	29.071
15	32.323	32.008	25.988	25.908	23.633	27.972	27.550
16	26.982	34.313	29.025	26.148	28.597	29.013	28.423
17	35.402	33.550	28.186	27.215	26.206	30.112	30.035
18	26.745	26.942	31.729	34.890	30.728	30.206	29.820
19	27.987	32.290	35.043	28.617	29.900	30.767	30.799
20	28.182	31.098	29.858	33.297	34.615	31.410	31.072

APPENDIX B

DATA FOR EXAMPLE 2.2

Appendix B

This is the data for Example 2.2. There are four simulation runs. For each run, twenty samples were generated where each sample consisted of ten (2x1) vectors of observations. For each sample, $\hat{\underset{\sim}{\mu}} = [\hat{\mu}_1, \hat{\mu}_2]^t$ was calculated as well as $\bar{\underset{\sim}{x}} = [\bar{x}_1, \bar{x}_2]^t$. Furthermore, for each sample, the test statistic (denoted SUM 1)

$$(\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)^t (\Sigma_{\hat{\mu}})^{-1} (\hat{\underset{\sim}{\mu}} - \underset{\sim}{\mu}_0)$$

was calculated as well as the statistic (denoted SUM 2)

$$(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0)^t \, \Sigma_{ii}^{-1} \, (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_0),$$

which completely ignores the correlative structure. For each run,

$$\theta_{11} = \sqrt{2}/4, \; \theta_{12} = \theta_{21} = \theta_{22} = \sqrt{2}/4$$

$$r = 0.0$$

$$c_1 = 1.0$$

$$\mu_1 = \mu_2 = 0.0$$

The four runs were obtained by letting $c_2 = 1.0\ (1.0)\ 4.0$.
The results are as follows.

Run Number 1. $c_2 = 1.0$						
Sample Number	$\hat{\mu}_1$	$\bar{x}_1$	$\hat{\mu}_2$	$\bar{x}_2$	SUM 1	SUM 2
1	.544	.530	-.177	-.151	1.603	2.428
2	-.026	.039	.042	-.040	0.035	0.250
3	-.744	-.767	.173	.160	3.194	4.915
4	-.106	-.175	-.004	.052	0.107	0.265
5	-.052	-.103	-.001	.019	0.024	0.088
6	.280	.238	-.070	-.039	0.446	0.466
7	-.057	-.036	.061	.069	0.065	0.048*
8	-.024	.043	.023	-.026	0.009	0.020
9	-.267	-.156	.092	.018	0.385	0.197*
10	.252	.287	-.052	-.112	0.378	0.761
11	-.303	-.371	.065	.097	0.542	1.180
12	-.479	-.508	.143	.131	1.257	2.200
13	-.347	-.335	.079	.059	0.700	0.925
14	.264	.297	-.037	-.052	0.462	0.727
15	-.452	-.414	.144	.082	1.110	1.423
16	-.165	-.161	.049	.068	0.149	0.245
17	-.200	-.206	.071	.039	0.216	0.353
18	-.664	-.662	.188	.189	2.433	3.789
19	.597	.779	-.088	-.183	2.312	5.125
20	-.126	-.042	.106	.027	0.194	0.020*

*The asterisk denotes SUM 1 is greater than SUM 2.

Run Number 2. $c_2 = 2.0$						
Sample Number	$\hat{\mu}_1$	$\bar{x}_1$	$\hat{\mu}_2$	$\bar{x}_2$	SUM 1	SUM 2
1	.532	.530	-.253	-.151	1.307	1.891
2	-.013	.039	.091	-.040	0.059	0.015*
3	-.753	-.766	.203	.160	2.984	3.874
4	-.116	-.174	-.083	.052	0.242	0.206*
5	-.058	-.103	.022	.019	0.016	0.069
6	.282	.238	-.084	-.039	0.408	0.369*
7	-.040	-.036	.096	.069	0.053	0.021*
8	-.018	.043	.042	-.026	0.010	0.014
9	-.259	-.156	.127	.018	0.307	0.156*
10	.256	.287	.005	-.112	0.480	0.578
11	-.308	-.371	.048	.097	0.563	0.923
12	-.474	-.508	.205	.131	1.054	1.723
13	-.351	-.335	.064	.059	0.710	0.731
14	.275	.297	.048	-.052	0.683	0.574*
15	-.444	-.414	.213	.082	0.908	1.123
16	-.163	-.161	.065	.068	0.127	0.184
17	-.194	-.206	.132	.039	0.176	0.279
18	-.660	-.662	.236	.189	2.131	2.952
19	.620	.779	.017	-.183	2.851	4.027
20	-.102	-.042	.251	.027	0.361	0.014*

*The asterisk denotes SUM 1 is greater than SUM 2.

Run Number 3. $c_2 = 3.0$						
Sample Number	$\hat{\mu}_1$	$\bar{x}_1$	$\hat{\mu}_2$	$\bar{x}_2$	SUM 1	SUM 2
1	.521	.530	-.328	-.151	1.019	1.400
2	-.005	.039	.139	-.039	0.085	0.010*
3	-.752	-.767	.232	.160	2.699	2.873
4	-.128	-.175	-.160	.052	0.381	0.152*
5	-.056	-.103	.043	.019	0.012	0.051
6	.281	.238	-.098	-.039	0.363	0.274*
7	-.032	-.036	.132	.069	0.056	0.013*
8	-.014	.043	.061	-.026	0.012	0.010*
9	-.253	-.156	.161	.018	0.240	0.116*
10	.263	.287	.060	-.112	0.597	0.426*
11	-.311	-.371	.032	.097	0.577	0.683
12	-.466	-.508	.266	.131	0.840	1.276
13	-.354	-.335	.049	.059	0.715	0.542*
14	.287	.297	.131	-.052	0.913	0.426*
15	-.434	-.414	.281	.082	0.702	0.833
16	-.161	-.161	.080	.068	0.105	0.135
17	-.185	-.206	.191	.039	0.132	0.207
18	-.653	-.662	.283	.189	1.821	2.184
19	.637	.779	.120	-.183	3.350	2.984*
20	-.079	-.042	.395	.027	0.531	0.010*

*The asterisk denotes SUM 1 is greater than SUM 2.

Run Number 4. $c_2 = 4.0$						
Sample Number	$\hat{\mu}_1$	$\bar{x}_1$	$\hat{\mu}_2$	$\bar{x}_2$	SUM 1	SUM 2
1	.512	.530	-.385	-.151	0.801	1.028
2	.001	.039	.176	-.040	0.106	0.007*
3	-.749	-.767	.254	.160	2.474	2.111*
4	-.137	-.175	-.219	.052	0.488	0.112*
5	-.055	-.103	.060	.019	0.008	0.038
6	.279	.238	-.108	-.039	0.328	0.201*
7	-.027	-.036	.159	.069	0.060	0.009*
8	-.012	.043	.076	-.026	0.014	0.008*
9	-.249	-.156	.187	.018	0.190	0.085*
10	.269	.287	.102	-.112	0.690	0.312*
11	-.314	-.371	.019	.097	0.587	0.502*
12	-.459	-.508	.313	.131	0.675	0.937
13	-.356	-.335	.038	.059	0.719	0.399*
14	.297	.297	.195	-.052	1.091	0.313*
15	-.426	-.414	.334	.082	0.546	0.612
16	-.159	-.161	.092	.068	0.088	0.099
17	-.179	-.206	.236	.039	0.099	0.152
18	-.648	-.662	.319	.189	1.585	1.603
19	.650	.779	.200	-.183	3.730	2.192*
20	-.062	-.042	.505	.027	0.665	0.007*

*The asterisk denotes SUM 1 is greater than SUM 2.

APPENDIX C

LISTING OF COMPUTER PROGRAM ICLSMAI (1)

```
 PROGRAM SUMSQ(INPUT,OUTPUT)
*****************************************************************
    *   *******
*        INPUT:      LOWER LIMIT, UPPER LIMIT, STEP SIZE
    *   FOR THETA1
*                    LOWER LIMIT, UPPER LIMT,  STEP SIZE
    *   FOR THETA2
*        OUTPUT:      FOR EACH THETA1, AND THETA2 :MUHAT,
    *   DELTAHAT,VARIA
*****************************************************************
    *   *********
 DIMENSION W(600),Y(600),S(2,2),SI(2,2)
+,
+T(2),BETA(2),X(600,2),XBETA(600),AHAT(200)
+,UL(2,2),IPR(2),R(2,2)
 REAL L1,L2
 READ 8001,N1,N2
 8001 FORMAT(3I6)
 READ 8020,L1,U1,STEP1
 READ 8020,L2,U2,STEP2
 8020 FORMAT(F14.6)
*********** INITIALISATION CONTROLS FOR INTERMEDIATE
    *   PRINTS
 KEYW=1
 KEYY=1
 KEYS=1
 KEYSI=1
 KEYT=1
 KEYB=1
 KEYX=1
 KEYB=1
 KEYXB=1
 KEYA=1
 N1PN2=N1+N2
 READ 8020,(W(I),I=1,N1PN2)
 IF(KEYW.NE.1) GOTO 10
 PRINT 9010
 9010 FORMAT(/10X,*--- VALUES OF W---*/)
 PRINT 9015,(W(I),I=1,N1PN2)
 9015 FORMAT(30X,10(F9.3))
 10 CONTINUE
******** DO LOOP OVER THETA1AND DO LOOP OVER THETA2
 NSTEP1= IFIX((U1-L1)/STEP1)+1
 THETA1=L1
 DO 4000 LOOP1=1,NSTEP1
 NSTEP2= IFIX((U2-L2)/STEP2)+1
 THETA2=L2
 DO 5000 LOOP2=1,NSTEP2
********COMPUTATION OF Y VECTOR
 Y(1)=W(1)
```

```
 DO 20 I=2,N1
 20 Y(I)=W(I)+THETA1*Y(I-1)
 N1P1=N1+1
 DO 30 I=N1P1,N1PN2
 30 Y(I)=W(I)+THETA2*Y(I-1)
*************** PRINT Y IF NEEDED
 IF(KEYY.NE.1) GOTO 40
 PRINT 9020
 9020 FORMAT(//10X,* VALUES OF Y*/)
 PRINT 9025,(Y(I),I=1,N1)
 9025 FORMAT(5X,10(F9.4))
 PRINT 9026
 9026 FORMAT(*---------------------------*)
 PRINT 9025,(Y(I),I=N1P1,N1PN2)
*************** PRINTING OF Y OVER
********** CALCULATION OF SS11
 40 TERM1=1./(1.-THETA1)**2
 T12=THETA1*THETA1
 T1=THETA1
 T2=THETA2
 T1N1=T1**N1
 T12N1=T1**(2*N1)
 T22=T2*T2
 T2N2=T2**N2
 T22N2=T2**(2*N2)
 VAL1=FLOAT(N1)
 VAL2=2.*T1*(1.-T1N1)/(1.-T1)
 VAL3=T12*(1.-T12N1)/(1.-T12)
 TERM2=VAL1-VAL2+VAL3
*******************************************************
 TERM3= 1./(1.-T2)**2
******TERM4
 VAL1=FLOAT(N2)
 VAL2= 2.*THETA2*(1.-T2N2)/(1.-T2)
 VAL3=T22*(1.-T22N2)/(1.-T22)
 TERM4=VAL1-VAL2+VAL3
*********************************************************
 *  *
 VAL1=((1.-T1N1)/(1.-T1))**2
************************** TERM5
 VAL2=T22
 VAL3= (1.-T22N2)/(1.-T22)
 TERM5= VAL1*VAL2*VAL3
*********************************************************
 *  *****
 TERM6=2.*(1.-T1N1)/((1.-T1)*(1.-T2))
******** TERM7
 VAL1=THETA2*(1.-T2N2)/(1.-T2)
 VAL2=T22*(          1.-T22N2)/(1.-T22)
 TERM7=VAL1-VAL2
```

```
S(1,1)= TERM1*TERM2+TERM3*TERM4+TERM5+TERM6*TERM7
******* CALCULATION OF SS22
TERM1= 1./(1.-T2)**2
VAL1=FLOAT(N2)
VAL2= 2.* T2*(1.-T2N2)/(1.-T2)
VAL3=T22*(1.-T22N2)/(1.-T22)
TERM2=VAL1-VAL2+VAL3
S(2,2)= TERM1*TERM2
********* CALCULATION OF SS12
TERM1=1./(1.-T2)**2
VAL1=FLOAT(N2)
VAL2=2.*T2*(1.-T2N2)/(1.-T2)
VAL3=(T2**2)*(1.-T22N2)/(1.-T2**2)
TERM2=VAL1-VAL2+VAL3
TERM3=(1.-T1N1)/((1.-T1)*(1.-T2))
VAL1=T2*(1.-T2N2)/(1.-T2)
VAL2=T22*(1.-T22N2)/(1.-T22)
TERM4=VAL1-VAL2
S(1,2)= TERM1*TERM2+TERM3*TERM4
S(2,1)=S(1,2)
*********TO COMPUTE SS1Y AND SS2Y
********* FIRST COMPUTE SUMS
SUM1=0.
SUM2=0.
SUM3=0.
SUM4=0.
DO 50 I=1,N1
SUM1=SUM1+Y(I)
50 SUM2=SUM2+Y(I)*(T1**I)
DO 60 I=1,N2
SUM3=SUM3+Y(N1+I)
60 SUM4=SUM4+Y(N1+I)*(T2**I)
********* CALCULATION OF T(1)
TERM1=1./(1.-T1)
TERM2=SUM1-SUM2
TERM3=1./(1.-THETA2)
TERM4=SUM3-SUM4
TERM5=(1.-T1N1)/(1.-THETA1)
TERM6=SUM4
T(1)=TERM1*TERM2+TERM3*TERM4+TERM5*TERM6
***************** CALCULATION OF T(2)
T(2)=TERM3*JERM4
********** PRINT S AND T IF NEEDED
IF(KEYS.NE.1) GOTO 70
TERM2=SUM3-SUM4
PRINT 9030
9030 FORMAT(/10X,*S MATRIX*/)
PRINT 9035,((S(I,J),J=1,2),I=1,2)
9035 FORMAT(10X,2F14.6)
70 IF(KEYT.NE.1) GOTO 80
```

```
PRINT 9038
9038 FORMAT(/10X,* T VECTOR*)
PRINT 9035, T(1),T(2)
*****************************************************************
    *  ********
************* HERE INSMRT STATEMENTSFOR INVERSION
L=0
CALL INVITR(S,UL,2,2,IPR,R,SI,D1,L,DX,KD)
************ PRINT INVERSE IF POSSIBLE
IF(KEYSI.NE.1) GOTO 80
PRINT 9040
9040 FORMAT(/10X,* INVERSE OF S*/)
PRINT 9035,((SI(I,J),J=1,2),I=1,2)
******** CALCULATION OF BETA VECTOR
80 BETA(1)= SI(1,1)*T(1)+SI(1,2)*T(2)
BETA(2)=SI(2,1)*T(1)+SI(2,2)*T(2)
******* PRINT BETA VECTOR IF NEEDED
IF(KEYB.NE.1) GOTO 90
PRINT 9050
9050 FORMAT(/10X,*   BETA VECTOR*/)
PRINT 9035,BETA(1),BETA(2)
******** TO FORMULATE X MATRIX
90 X(1,1)=1.
DO 100 I=2,N1
X(I,2)=0.
100 X(I,1)=X(I-1)+THETA1**(I-1)
C=(1.-T1N1)/(1.-T1)
X(N1+1,1)=1.+C*THETA2
X(N1+1,2)=1.
DO 110 I=2,N2
SUM=1.
IM1=I-1
DO 120 J=1,IM1
120 SUM=SUM+(THETA2**J)
X(N1+I,2)=SUM
X(N1+I,1)=SUM+(THETA2**I)*C
110 CONTINUE
************** PRINT X MATRIX IF NEEDED
IF(KEYX.NE.1) GOTO 130
PRINT 9060
9060 FORMAT(/10X,*X MATRIX*/)
PRINT 9035,(X(I,1),X(I,2),I=1,N1PN2)
************* TO FIND PRODUCT OF BETA AND  X
130 DO150 I=1,N1PN2
150 XBETA(I)=X(I,1)*BETA(1)+X(I,2)*BETA(2)
*********** PRINT XBETA IF NEEDED
IF(KEYXB.NE.1) GOTO 160
PRINT 9070
9070 FORMAT(/10X,* PRODUCT OF X AND BETA*)
PRINT 9035,(XBETA(I),I=1,N1PN2)
```

```
******* CALCULATION OF AHAT
160 DO 180 I=1,N1PN2
180 AHAT(I)= Y(I)-XBETA(I)
*****************PRINT AHAT IF NEEDED
IF(KEYA.NE.1) GOTO 190
PRINT 9080
9080 FORMAT(10X,*AHAT VECTOR*)
PRINT 9035,(AHAT(I),I=1,N1PN2)
******** CALCULATION OF SA SQUARED
SUM=0.
190 DO 195 I=1,N1PN2
195 SUM=SUM+AHAT(I)*AHAT(I)
VAR=SUM/FLOAT(N1+N2-2)
**********0PRINT ALL VALUES
PRINT 9090,T1,T2,BETA(1),BETA(2),VAR
9090 FORMAT(/*  THETA1=*,F14.6,*THETA2=*,F14.6,5X,
     *  *MUHAT=*,
+F14.6,5X,*DELTAHAT*,F14.6,5X,*VAR=*,F14.6)
THETA2=T2+STEP2
5000 CONTINUE
THETA1=THETA1+STEP1
4000 CONTINUE
STOP
END
```

APPENDIX D

LISTING OF COMPUTER PROGRAM MLEMAI (1)

```
 PROGRAM POWELL(INPUT,OUTPUT)
*  MAXIMUM LIKELIHOOD ESTIMATION PROGRAM
* INPUT:  1.TIME SERIES DATA... READ INTO ARRAY Z1
*          2.STARTING VALUES FOR THETA1,GAMA1
* OUTPUT:   OPTIMUM VALUES OF THE VARIABLES
*           THETA1,GAMA1,MUHAT,DELTAHAT
*ATTACH SUBROUTINE INVITR FROM MSFLIB BEFORE RUNNING
 *  THE PROGRAM
*************************************************************
 *  **********
 DIMENSION X(4),W(28),E(4),Z1(40)
 COMMON N1,N2,Z(40),X,F
+,NITER
 DATA (Z1(I),I=1,40)/
+20.,18.,20.,18.,21.,20.,19.,22.,23.,20.,
+22.,18.,21.,19.,20.,16.,20.,15.,14.,13.,
+12.,9.,6.,6.,4.,6.,6.,4.,2.,3.,
+4.,4.,5.,4.,2.,4.,6.,4.,4.,2./
 NITER=0
 N=2
 N1=20
 N2=20
 N1PN2=N1+N2
 IPRINT=2
 DO 103 K9=1,40
 103 Z(K9)=Z1(K9)
 MAXIT=500
 ESCALE=0.5
 READ 5,X(1)
 READ 5,X(2)
 5 FORMAT(F14.6)
 6993 FORMAT(F15.6)
 E(1)=0.05
 E(2)=0.05
 E(3)=0.05
 E(4)=0.05
 NW=N*(N+3)
 CALL BOTM(X,E,N,EF,ESCALE,IPRINT,MAXIT,W,NI,NO,NW)
* NON LINEAR OPTIMIZATION USING POWELLS ALGORITHM
*ADOPTED FROM MIZE & KUESTER:OPTIMIZATION TECHNIQUES
 *  WITH FORTRAN I
*
 PRINT 1
 1 FORMAT(//,5X,*VALUES OF THE VARIABLES*)
 DO 100 J=1,N
 PRINT 2, J,X(J)
 2 FORMAT(/5X,*X(*,I2,*) = *,E16.8)
 100 CONTINUE
 PRINT 3, EF
 3 FORMAT(//,5X,*OPTIMUM VALUE OF F = *,E16.8)
```

```
END
*
*
SUBROUTINE BOTM(X,E,N,EF,ESCALE,IPRINT,MAXIT,W,NI,NC,
*  NW)
DIMENSION X(N),W(NW),E(N)
COMMON JUMK(2),FJUNK(40),FXXX(4),F
*
DDMAG=0.1*ESCALE
SCER=0.05/ESCALE
JJ=N*(N+1)
JJJ=JJ+N
K=N+1
NFCC=1
IND=1
INN=1
DO 4 I=1,N
W(I)=ESCALE
DO 4 J=1,N
W(K)=0.
IF(I-J) 4,3,4
*
3 W(K)=ABS(E(I))
4 K=K+1
ITREC=1
ISGRAD=2
CALL CALCFX
FKEEP=2.*ABS(F)
5 ITONE=1
FP=F
SUM=0.
IXP=JJ
DO 6 I=1,N
IXP=IXP+1
6 W(IXP)=X(I)
IDIRN=N+1
ILINE=1
7 DMAX=W(ILINE)
DACC=DMAX*SCER
DMAG=AMIN1(DDMAG,0.1*DMAX)
DMAG=AMAX1(DMAG,20.*DACC)
DDMAX=10.*DMAG
GOTO (70,70,71), ITONE
*
*
70 DL=0.
D=DMAG
FPREV=F
IS=5
FA=FPREV
```

```
DA=DL
8 DD=D-DL
DL=D
58 K=IDIRN
DO 9 I=1,N
X(I)=X(I)+DD*W(K)
9 K=K+1
CALL CALCFX
NFCC=NFCC+1
GOTO (10,11,12,13,14,96),IS
14 IF(F-FA) 15,16,24
16 IF(ABS(D)-DMAX)17,17,18
17 D=D+D
GOTO 8
18 PRINT 19
19 FORMAT(5X,*MAXIMUM CHANGE DOES NOT ALTER FUNCTION*)
GOTO 20
15 FB=F
DB=D
GOTO 21
24 FB=FA
DB=DA
FA=F
DA=D
21 GOTO (83,23),ISGRAD
23 D=DB+DB-DA
*
IS=1
GOTO 8
83 D=0.5*(DA+DB-(FA-FB)/(DA-DB))
IS=4
IF((DA-D)*(D-DB))25,8,8
25 IS=1
IF(ABS(D-DB)-DDMAX)8,8,26
26 D=DB+SIGN(DDMAX,DB-DA)
IS=1
DDMAX=DDMAX+DDMAX
DDMAG=DDMAG+DDMAG
IF(DDMAG.GE.1.0E+60) DDMAG=1.0E+60
IF(DDMAX-DMAX) 8,8,27
27 DDMAX=DMAX
GOTO 8
13 IF(F-FA) 28,23,23
28 FC=FB
DC=DB
29 FB=F
DB=D
GOTO 30
12 IF(F-FB) 28,28,31
31 FA=F
```

```
DA=D
GOTO 30
11 IF(F-FB) 32,10,10
32 FA=FB
DA=DB
GOTO 29
71 DL=1.
DDMAX=5.
FA=FP
DA=-1.
FB=FHOLD
DB=0.
D=1.
10 FC=F
DC=D
30 A=(DB-DC)*(FA-FC)
B=(DC-DA)*(FB-FC)
IF((A+B)*(DA-DC))33,33,34
33 FA=FB
DA=DB
FB=FC
DB=DC
GOTO 26
34 D=0.5*(A*(DB+DC)+B*(DA+DC))/(A+B)
DI=DB
FI=FB
IF(FB-FC)44,44,43
43 DI=DC
FI=FC
44 GOTO(86,86,85),ITONE
85 ITONE=2
GOTO 45
86 IF(ABS(D-DI)-DACC) 41,41,93
93 IF(ABS(D-DI)-0.03*ABS(D)) 41,41,45
45 IF((DA-DC)*(DC-D))47,46,46
46 FA=FB
DA=DB
FB=FC
DB=DC
GOTO 25
47 IS=2
IF((DB-D)*(D-DC)) 46,6,8
48 IS=3
GOTO 8
41 F=FI
D=DI-DL
DD=SQRT((DC-DB)*(DC-DA)*(DA-DB)/(A+B))
DO 49 I=1,N
X(I)=X(I)+D*W(IDIRN)
W(IDIRN)=DD*W(IDIRN)
```

```
49 IDIRN=IDIRN+1
W(ILINE)=W(ILINE)/DD
ILINE=ILINE+1
IF(IPRINT-1) 51,50,51
50 PRINT 52,ITREC,NFCC,F,(X(I),I=1,N)
52 FORMAT(/*ITERATION*,I5,I15,*FUNCTION VALUES*,10X,
+*F=*,E15.8,2(E16.8,2X))
GOTO (51,53),IPRINT
51 GOTO (55,38), ITONE
55 IF(FPREV-F-SUM) 94,95,95
95 SUM=FPREV-F
JIL=ILINE
94 IF(IDIRN-JJ) 7,7,84
84 GOTO (92,72), IND
92 FHOLD=F
IS=6
IXP=JJ
DO 59 I=1,N
IXP=IXP+1
59 W(IXP)=X(I)-W(IXP)
DD=1.
GOTO 58
96 GOTO (112,87), IND
112 IF(FP-F) 37,37,91
91 D=2.*(FP+F-2.*FHOLD)/(FP-F)**2
IF(D*(FP-FHOLD-SUM)**2-SUM) 87,37,37
87 J=JIL*N+1
IF(J-JJ) 60,60,61
60 DO 62 I=J,JJ
K=I-1
62 W(K)=W(I)
DO 97 I=JIL,N
97 W(I-1)=W(I)
61 IDIRN=IDIRN-N
ITONE=3
K=IDIRN
IXP=JJ
AAA=0.
DO 67 I=1,N
IXP=IXP+1
W(K)=W(IXP)
IF(AAA-ABS(W(K)/E(I))) 66,67,67
66 AAA=ABS(W(K)/E(I))
67 K=K+1
DDMAG=1.
W(N)=ESCALE/AAA
ILINE=N
GOTO 7
37 IXP=JJ
AAA=0.
```

```
F=FHOLD
DO 99 I=1,N
IXP=IXP+1
X(I)=X(I)-W(IXP)
IF( AAA*ABS(E(I))-ABS(W(IXP))) 98,99,99
98 AAA=ABS(W(IXP)/E(I))
99 CONTINUE
GOTO 72
38 AAA=AAA*(1.+DI)
GOTO (72,106), IND
72 IF(IPRINT-2) 53,50,50
53 GOTO ( 109,88), IND
109 IF(AAA-0.1) 20,20,76
76 IF(F-FP) 35,78,78
78 PRINT 80
80 FORMAT(5X,*ACCURACY LIMITED BY ERRORS IN F*)
GOTO 20
88 IND=1
35 DDMAG=0.4*SQRT(ABS(FP-F))
IF(DDMAG.GE.1.0E+60) DDMAG=1.0E+60
ISGRAD=1
108 ITREC=ITREC+1
IF(ITREC-MAXIT) 5,5,81
81 PRINT 82,MAXIT
82 FORMAT(I5,*ITERATIONS COMPLETED BY BOTM*)
IF(F-FKEEP) 20,20,110
110 F=FKEEP
DO 111 I=1,N
JJJ=JJJ+1
111 X(I)=W(JJJ)
GOTO 20
101 JIL=1
FP=FKEEP
IF(F-FKEEP) 105,76,104
104 JIL=2
FP=F
F=FKEEP
105 IXP=JJ
DO 113 I=1,N
IXP=IXP+1
K=IXP+N
GOTO (114,115),JIL
114 W(IXP)=W(K)
GOTO 113
115 W(IXP)=X(I)
X(I)=W(K)
113 CONTINUE
JIL=2
GOTO 92
106 IF(AAA-0.1) 20,20,107
```

```
 20 EF=F
 RETURN
 107 INN=1
 GOTO 35
 END
 SUBROUTINE CALCFX
 DIMENSION Z(40),MUZ(40),MI(40,40),M(40,40),J(40),K(40)
     *   ,
+MZ(40),MJ(40),MK(40),L(40),AIT(40),
+UL(40,40),IPR(40),R(40,40)
+,AT(40)
 REAL MUZ,MU,MI,M,J,K,MJHAT,MZ,MK,MJ,KMZ,JMJ,KMJ
+,KMK
+,L
 COMMON N1,N2,Z,T1,G1,MJHAT,DHAT
+,F
+,NITER
***** FURMULATION OF MI
***** INITIALISE ALL ENTRIES TO ZERO
 INDEXM=0
 INDEXJ=0
 INDEXD=0
 INDEXL=0
 INDEXS=0
 INDEXXD=0
 N1PN2=N1+N2
 N1P1=N1+1
 N1M1=N1-1
 T12=T1*T1
 G12=G1*G1
 MU=MUHAT
 D=DHAT
 N2M1=N2-1
 DO 5 I=1,N1PN2
 DO 5 J1=1,N1PN2
 5 MI(I,J1)=0.
***************************** BLOCK 1
***************************** MAIN DIAGONAL
 DO 10 I=1,N1
 10 MI(I,I)=1.+T12
******************************* SUB DIAGONALS
 DO 20 I=2,N1
 MI(I,I-1)=-T1
 20 MI(I-1,I)= -T1
************* FORMAT TO PRINT TEN ELENENTS IN EACH ROW
 5000 FORMAT(5X,10F10.5)
**************** FORMAT O PRINT ONE ELEMENT IN A ROW
 6000 FORMAT(5X,F10.5)
************************** BLOCK2
************************** MAIN DIAGONAL AND
     *   SUBDIAGONALS
```

```
DO 30 I=N1P1,N1PN2
MI(I,I)=1.+G12
MI(I,I-1)= -G1
30 MI(I-1,I)=-G1
********* TAKE INVERSE OF MI TO FORMULZTE M
L3=0
CALL INVITR(MI,UL,N1PN2,N1PN2,IPR,R,M,D1,L3,DX,KD)
************** PRINT M,MI IF NEEDED
IF(INDEXM.NE.1) GOTO 31
    PRINT 1000
1000 FORMAT(/10X,*MI MATRIX*)
PRINT 5000,((MI(I,JI),JI=1,N1PN2),I=1,N1PN2)
PRINT 1010
1010 FORMAT(/10X,*M MATRIX*)
PRINT 5000,((M(I,JI),JI=1,N1PN2),I=1,N1PN2)
******* NOW FORMULATE SIGMA
31 CONTINUE
DO 38 I=1,N1PN2
DO 38 JI=1,N1PN2
38 CONTINUE
*********** TO COMPUTE MUHAT AND DELTAHAT
DO 60 I=1,N1PN2
MJ(I)=0.
DO 70 J1=1,N1PN2
70 MJ(I)= MJ(I)+M(I,J1)
60 CONTINUE
******** FORMULATE J VECTOR AND K VECTOR
DO 65 I=1,N1PN2
65 J(I)=1.
DO 75 I=1,N1
75 K(I)=0.
DO 76 I=N1P1,N1PN2
76 K(I)=J(I)
*************** PRINT J AND K IF NEEDED
IF(INDEXJ.NE.1) GOTO 79
PRINT 1040,(J(I),I=1,N1PN2)
1040 FORMAT(/10X,*K MATRIX*/(10X,F14.6))
*********** COMPUTE MZ N MK)
79 CONTINUE
DO 80 I=1,N1PN2
MZ(I)=0.
MK(I)=0.
DO 90 J1=1,N1PN2
MZ(I)=MZ(I)+M(I,J1)*Z(J1)
MK(I)=MK(I)+M(I,J1)*K(J1)
90 CONTINUE
80 CONTINUE
************** COMPUTE KMZ,KMJ,ZMJ,KMK,JMJ
KMZ=0.
DMJ=0.
```

```
ZMJ=0.
KMK=0.
JMJ=0.
KMJ=0.
DO 100 I=1,N1PN2
KMZ=KMZ+K(I)*MZ(I)
KMJ=KMJ+K(I)*MJ(I)
ZMJ=ZMJ+Z(I)*MJ(I)
KMK=KMK+K(I)*MK(I)
JMJ=JMJ+J(I)*MJ(I)
100 CONTINUE
DHAT=(KMZ*JMJ-ZMJ*KMJ)/(KMK*JMJ-KMJ*KMJ)
MUHAT=(ZMJ-DHAT*KMJ)/JMJ
MU=MUHAT
D=DHAT
********* PRINT DHAT AND MUHAT IF NEEDED
IF( INDEXD.NE.1) GOTO 101
PRINT 1050,DHAT,MUHAT
1050 FORMAT(/10X,*DHAT=*,F14.6,10X,*MUHAT=*,F14.6)
******* FORMULATE L VECTOR
101 CONTINUE
DO 110 I=1,N1
TX01=1.-T12**(N1-I)
TX02=1.-T12
TX03=T12**(N1-I)
TX04=1.-G12**(N2+1)
TX05=1.-G12
110 L(I)=(T1**I)*((TX01/TX02)+(TX03*TX04/TX05))
DO 120 I=N1P1,N1PN2
L(I)=(T1**N1)*(G1**(I-N1))
L(I)=L(I)*(1.-G1**(2*(N1PN2-I+1)))/(1.-G12)
120 CONTINUE
******* PRINT L VECTOR IF NEEDED
IF(INDEXL.NE.1) GOTO 121
PRINT 1060,(L(I),I=1,N1PN2)
1060 FORMAT(10X,*L VECTOR*/(10X,F14.6))
121 CONTINUE
****** TO, FIND XTD=TERM1+TERM2+TERM3+TERM4
******* CALCULATE SUM1,SUM2,SUM3
SUM1=0.
SUM2=0.
SUM3=0.
DO 130 I1=1,N1
I=I1-1
130 SUM1=SUM1+(T1**I)*(1.-T1**(I+1))
DO 140 I1=1,N2
I=I1-1
SUM2=SUM2+(G1**I)*(1.-G1**(I+1))
SUM3=SUM3+G1**(2*I+1)
140 CONTINUE
```

```
******* PRINT SUMS IF NEEDED
 IF(INDEXS.NE.1) GOTO 141
 PRINT 1070,SUM1,SUM2,SUM3
 1070 FORMAT(/10X,*SUM1=*,F14.6,5X,*SUM2=*,F14.6,*SUM3=
*   *,F14.6)
 141 CONTINUE
********** MU=MUHAT
 TERM1=MU*T1/(1.-T1)
 TERM1=TERM1*SUM1
 TERM2=DHAT*(T1**N1)*G1/(1.-G1)
 TERM2=TERM2*SUM2
 TERM3=MU*(T1**N1)*G1/(1.-G1)
 TERM3=TERM3*SUM2
 TERM4=MU*(T1**N1)*G1*(1.-T1**N1)/(1.-T1)
 TERM4=TERM4*SUM3
 XTD=TERM1+TERM2+TERM3+TERM4
***** TO COMPUTE XTX
 XL=0.
 DO 150 I=1,N1PN2
 150 XL=XL+L(I)*Z(I)
*********   CALCULATION OF XTX
 XTX=(1.-T1**(2*N1))/(1.-T12)
 XTXC=XTX+(T1**(2*N1))*(1.-G1**(2*(N2+1)))/(1.-G12)
 XTX=XTXC
 AHAT=(XTD-XL)/XTX
******************** COMPUTATION OF A
 IF(INDEXXD.NE.1) GOTO 151
 PRINT 1080,XTD,XTX,AHAT
 1080 FORMAT(/10X,*XTD=*,F14.6,5X,*XTX=*,F14.6,5X,
*   *AHAT=*,F14.6)
 151 CONTINUE
 AT(1)=Z(1)-MUHAT+T1*AHAT
 DO 160 I=2,N1
 160 AT(I)=Z(I)-MU+T1*AT(I-1)
 DO 170 I=N1P1,N1PN2
 170 AT(I)=Z(I)-MU-DHAT+G1*AT(I-1)
******** TO CALCULATE F
 F=AHAT*AHAT
 DO 180 I=1,N1PN2
 F=F+(AT(I)*AT(I)/N1PN2)
 180 CONTINUE
 F=(N1PN2/2)*ALOG(F)+.5*ALOG(XTX)
 PRINT 183,T1,G1,MUHAT,DHAT,F
 183 FORMAT(5F14.6)
 NITER=NITER+1
 IF(NITER.EQ.5)NITER=0
 RETURN
 END
```

APPENDIX E

ECONOMIC PARAMETERS FOR TWO QUALITY CHARACTERISTICS, INDEPENDENT OBSERVATIONS

Two Characteristics, Independent Observations

		$\rho=-.80$											
		$L_0=5{,}000$			$L_0=10{,}000$			$L_0=20{,}000$			$L_0=40{,}000$		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
.20	0.00	88	8.08	130.1	103	9.15	147.6	118	10.27	165.1	133	11.41	182.6
	.20	32	10.10	45.1	36	11.25	50.0	41	12.39	54.9	45	13.58	59.7
	.40	16	11.49	22.3	18	12.64	24.5	20	13.81	28.6	22	15.01	28.7
	.60	10	12.43	13.4	11	13.62	14.6	12	14.81	15.7	13	16.06	16.9
	.80	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	1.00	5	13.81	6.4	5	15.20	7.0	6	16.22	7.5	6	17.54	8.0
	1.20	4	14.25	4.9	4	15.65	5.2	4	17.00	5.7	5	17.87	6.0
	1.40	3	14.81	3.8	3	16.22	4.1	3	17.54	4.5	4	18.25	4.7
	1.60	2	15.65	3.2	3	16.22	3.4	3	17.54	3.6	3	18.98	3.7
	1.80	2	15.65	2.5	2	17.00	2.7	2	18.25	2.9	2	19.74	3.2
.40	0.00	29	10.30	41.3	33	11.42	45.7	37	12.58	50.0	41	13.76	54.4
	.20	16	11.49	22.3	18	12.64	24.5	20	13.81	26.6	22	15.01	28.7
	.40	10	12.43	13.7	11	13.62	14.9	12	14.81	16.1	13	16.06	17.4
	.60	7	13.14	9.2	8	14.25	10.0	8	15.65	10.8	9	16.78	11.5
	.80	5	13.81	6.7	6	14.81	7.2	6	16.22	7.7	6	17.54	8.3
	1.00	4	14.25	5.0	4	15.65	5.4	5	16.57	5.9	5	17.87	6.2
	1.20	3	14.81	3.9	3	16.22	4.3	4	17.00	4.6	4	18.25	4.8
	1.40	2	15.65	3.3	3	16.22	3.5	3	17.54	3.6	3	18.98	3.9
	1.60	2	15.65	2.6	2	17.00	2.8	2	18.25	3.0	3	18.98	3.3
	1.80	2	15.65	2.3	2	17.00	2.4	2	18.25	2.5	2	19.74	2.6
.60	0.00	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	.20	10	12.43	13.4	11	13.62	14.6	12	14.81	15.7	13	16.08	16.9
	.40	7	13.14	9.2	8	14.25	10.0	8	15.65	10.8	9	16.78	11.5
	.60	5	13.81	6.7	6	14.81	7.3	6	16.22	7.8	6	17.54	8.4
	.80	4	14.25	5.1	4	15.65	5.5	5	16.57	5.9	5	17.87	6.3
	1.00	3	14.81	4.0	3	16.22	4.4	4	17.00	4.7	4	18.25	4.9
	1.20	3	14.81	3.4	3	16.22	3.5	3	17.54	3.7	3	18.98	3.9
	1.40	2	15.65	2.7	2	17.00	2.9	2	18.25	3.1	3	18.98	3.3
	1.60	2	15.65	2.3	2	17.00	2.4	2	18.25	2.5	2	19.74	2.7
	1.80	2	15.65	2.1	2	17.00	2.2	2	18.25	2.2	2	19.74	2.3
.80	0.00	9	12.64	12.5	10	13.81	13.6	11	15.01	14.7	12	16.22	15.8
	.20	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	.40	5	13.81	6.7	6	14.81	7.2	6	16.22	7.7	6	17.54	8.3
	.60	4	14.25	5.1	4	15.65	5.5	5	16.57	5.9	5	17.87	6.3
	.80	3	14.81	4.0	3	16.22	4.4	4	17.00	4.7	4	18.25	4.9
	1.00	3	14.81	3.4	3	16.22	3.5	3	17.54	3.7	3	18.98	4.0
	1.20	2	15.65	2.7	2	17.00	2.9	2	18.25	3.2	3	18.98	3.4
	1.40	2	15.65	2.3	2	17.00	2.4	2	18.25	2.6	2	19.74	2.8
	1.60	2	15.65	2.1	2	17.00	2.2	2	18.25	2.2	2	19.74	2.3
	1.80	1	17.00	1.8	1	18.25	2.0	2	18.25	2.1	2	19.74	2.1
1.00	0.00	6	13.45	8.5	7	14.51	9.1	8	15.65	9.9	8	17.00	10.5
	.20	5	13.81	6.4	5	15.20	7.0	6	16.22	7.5	6	17.54	8.0
	.40	4	14.25	5.0	4	15.65	5.4	5	16.57	5.9	5	17.87	6.2

		ρ=-.80 (continued)											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
	.60	3	14.81	4.0	3	16.22	4.4	4	17.00	4.7	4	18.25	4.9
	.80	3	14.81	3.4	3	15.22	3.5	3	17.54	3.7	3	18.98	4.0
	1.00	2	15.65	2.7	2	17.00	2.9	2	18.25	3.2	3	18.98	3.4
	1.20	2	15.65	2.3	2	17.00	2.4	2	18.25	2.6	2	19.74	2.8
	1.40	2	15.65	2.1	2	17.00	2.2	2	18.25	2.3	2	19.74	2.4
	1.60	1	17.00	1.8	1	18.25	2.0	2	18.25	2.1	2	19.74	2.1
	1.80	1	17.00	1.5	1	18.25	1.6	1	19.74	1.8	1	20.90	2.0
1.20	0.00	5	13.81	6.2	5	15.20	6.6	6	16.22	7.2	6	17.54	7.6
	.20	4	14.25	4.9	4	15.65	5.2	4	17.00	5.7	5	17.87	8.0
	.40	3	14.81	3.9	3	16.22	4.3	4	17.00	4.6	4	18.25	4.8
	.60	3	14.81	3.4	3	16.22	3.5	3	17.54	3.7	3	18.98	3.9
	.80	2	15.65	2.7	2	17.00	2.9	2	18.25	3.2	3	18.98	3.4
	1.00	2	15.65	2.3	2	17.00	2.4	2	18.25	2.8	2	19.74	2.8
	1.20	2	15.65	2.1	2	17.00	2.2	2	18.25	2.3	2	19.74	2.4
	1.40	1	17.00	1.8	1	18.25	2.1	2	18.25	2.1	2	19.74	2.2
	1.60	1	17.00	1.5	1	18.25	1.6	1	19.74	1.8	1	20.90	2.0
	1.80	1	17.00	1.3	1	18.25	1.4	1	19.74	1.5	1	20.90	1.6
1.40	0.00	4	14.25	4.7	4	15.65	5.0	4	17.00	5.4	4	18.25	5.8
	.20	3	14.81	3.8	3	16.22	4.1	3	17.54	4.5	4	18.25	4.7
	.40	2	15.65	3.3	3	16.22	3.5	3	17.54	3.6	3	18.98	3.9
	.60	2	15.65	2.7	2	17.00	2.9	2	18.25	3.1	3	18.98	3.3
	.80	2	15.65	2.3	2	17.00	2.4	2	18.25	2.6	2	19.74	2.8
	1.00	2	15.65	2.1	2	17.00	2.2	2	18.25	2.3	2	19.74	2.4
	1.20	1	17.00	1.8	1	18.25	2.1	2	18.25	2.1	2	19.74	2.2
	1.40	1	17.00	1.5	1	18.25	1.6	1	19.74	1.9	1	20.90	2.1
	1.60	1	17.00	1.3	1	18.25	1.4	1	19.74	1.5	1	20.90	1.6
	1.80	1	17.00	1.2	1	18.25	1.2	1	19.74	1.3	1	20.90	1.4
1.60	0.00	3	14.81	3.7	3	16.22	3.9	3	17.54	4.2	4	18.25	4.6
	.20	2	15.65	3.2	3	16.22	3.4	3	17.54	3.6	3	18.98	3.7
	.40	2	15.65	2.6	2	17.00	2.8	2	18.25	3.0	3	18.98	3.3
	.60	2	15.65	2.3	2	17.00	2.4	2	18.25	2.5	2	19.74	2.7
	.80	2	15.65	2.1	2	17.00	2.2	2	18.25	2.2	2	19.74	2.3
	1.00	1	17.00	1.8	1	18.25	2.0	2	18.25	2.1	2	19.74	2.1
	1.20	1	17.00	1.5	1	18.25	1.6	1	19.74	1.8	1	20.90	2.0
	1.40	1	17.00	1.3	1	18.25	1.4	1	19.74	1.5	1	20.90	1.6
	1.60	1	17.00	1.2	1	18.25	1.2	1	19.74	1.3	1	20.90	1.4
	1.80	1	17.00	1.1	1	18.25	1.1	1	19.74	1.2	1	20.90	1.2
1.80	0.00	2	15.65	3.0	3	16.22	3.3	3	17.54	3.5	3	18.98	3.6
	.20	2	15.65	2.5	2	17.00	2.7	2	18.25	2.9	2	19.74	3.2
	.40	2	15.65	2.3	2	17.00	2.4	2	18.25	2.5	2	19.74	2.6
	.60	2	15.65	2.1	2	17.00	2.2	2	18.25	2.2	2	19.74	2.3
	.80	1	17.00	1.8	1	18.25	2.0	2	18.25	2.1	2	19.74	2.1
	1.00	1	17.00	1.5	1	18.25	1.6	1	19.74	1.8	1	20.90	2.0
	1.20	1	17.00	1.3	1	18.25	1.4	1	19.74	1.5	1	20.90	1.8
	1.40	1	17.00	1.2	1	18.25	1.2	1	19.74	1.3	1	20.90	1.4
	1.60	1	17.00	1.1	1	18.25	1.1	1	19.74	1.2	1	20.90	1.2
	1.80	1	17.00	1.0	1	18.25	1.1	1	19.74	1.1	1	20.90	1.1

Two Characteristics, Independent Observations

		ρ=-.60											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
.20	0.00	136	7.21	205.8	161	8.26	236.6	188	9.33	267.7	214	10.46	298.8
	.20	56	8.98	80.5	64	10.10	90.3	73	11.22	100.0	81	12.40	109.7
	.40	29	10.30	39.9	32	11.49	44.1	36	12.64	48.3	40	13.81	52.5
	.60	17	11.37	23.7	19	12.53	26.9	21	13.71	28.3	23	14.89	30.5
	.80	12	12.06	15.8	13	13.29	17.2	14	14.51	18.6	15	15.77	20.0
	1.00	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	1.20	6	13.45	8.5	7	14.51	9.2	8	15.65	9.9	8	17.00	10.6
	1.40	5	13.81	6.6	6	14.81	7.2	6	16.22	7.7	6	17.54	8.2
	1.60	4	14.25	5.3	4	15.65	5.8	5	16.57	6.2	5	17.87	6.6
	1.80	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
.40	0.00	47	9.33	66.9	54	10.44	74.7	60	11.51	82.5	67	12.78	90.3
	.20	29	10.30	39.9	32	11.49	44.1	36	12.54	48.3	40	13.81	52.5
	.40	18	11.25	25.0	20	12.43	27.4	22	13.62	29.9	24	14.81	32.2
	.60	12	12.06	16.9	14	13.14	18.4	15	14.38	19.9	16	15.65	21.5
	.80	9	12.64	12.1	10	13.81	13.1	11	15.01	14.2	12	16.22	15.2
	1.00	7	13.14	9.1	7	14.51	9.9	3	15.65	10.6	9	16.78	11.3
	1.20	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
	1.40	4	14.25	5.7	5	15.20	6.1	5	16.57	6.6	5	17.87	7.0
	1.60	4	14.25	4.7	4	15.65	5.0	4	17.00	5.4	4	18.25	5.8
	1.80	3	14.81	3.9	3	16.22	4.2	4	17.00	4.6	4	18.25	4.8
.60	0.00	24	10.68	33.8	27	11.82	37.2	30	13.00	40.7	33	14.19	44.1
	.20	17	11.37	23.7	19	12.53	26.0	21	13.71	28.3	23	14.89	30.5
	.40	12	12.06	16.9	14	13.14	18.4	15	14.38	19.9	16	15.65	21.5
	.60	9	12.64	12.4	10	13.81	13.4	11	15.01	14.5	12	16.22	15.6
	.80	7	13.14	9.4	8	14.25	10.2	8	15.65	11.0	9	16.78	11.7
	1.00	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
	1.20	4	14.25	6.0	5	15.20	6.4	5	16.57	6.9	6	17.54	7.3
	1.40	4	14.25	4.9	4	15.65	5.2	4	17.00	5.6	5	17.87	6.0
	1.60	3	14.81	4.1	3	16.22	4.4	4	17.00	4.7	4	18.25	4.9
	1.80	3	14.81	3.5	3	16.22	3.7	3	17.54	3.9	3	18.98	4.2
.80	0.00	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	.20	12	12.06	15.8	13	13.29	17.2	14	14.51	18.6	15	15.77	20.0
	.40	9	12.64	12.1	10	13.81	13.1	11	15.01	14.2	12	16.22	15.2
	.60	7	13.14	9.4	8	14.25	10.2	8	15.65	11.0	9	16.78	11.7
	.80	6	13.45	7.5	6	14.81	8.1	7	15.91	8.7	7	17.26	9.3
	1.00	5	13.81	6.1	5	15.20	6.5	5	16.57	7.0	6	17.54	7.5
	1.20	4	14.25	5.0	4	15.65	5.4	4	17.00	5.8	5	17.87	6.1
	1.40	3	14.81	4.2	3	16.22	4.6	4	17.00	4.8	4	18.25	5.1
	1.60	3	14.81	3.6	3	16.22	3.8	3	17.54	4.1	3	18.98	4.4
	1.80	2	15.65	3.1	3	16.22	3.4	3	17.54	3.5	3	18.98	3.7
1.00	0.00	10	12.43	14.0	11	13.62	15.3	13	14.65	16.5	14	15.91	17.7
	.20	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	.40	7	13.14	9.1	7	14.51	9.9	8	15.85	10.6	9	16.78	11.3

		ρ=−.60 (continued)											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
	.60	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
	.80	5	13.81	6.1	5	15.20	6.5	5	16.57	7.0	6	17.54	7.5
	1.00	4	14.25	5.0	4	15.65	5.4	5	16.57	5.9	5	17.87	6.2
	1.20	3	14.81	4.2	4	15.65	4.6	4	17.00	4.9	4	18.25	5.1
	1.40	3	14.81	3.6	3	16.22	3.9	3	17.54	4.2	4	18.25	4.5
	1.60	2	15.65	3.2	3	16.22	3.4	3	17.54	3.8	3	18.98	3.8
	1.80	2	15.65	2.7	2	17.00	2.9	2	18.25	3.2	2	18.98	3.4
1.20	0.00	8	12.87	10.2	8	14.25	11.1	9	15.41	11.9	10	16.57	12.7
	.20	6	13.45	8.5	7	14.51	9.2	8	15.65	9.9	8	17.00	10.6
	.40	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
	.60	4	14.25	6.0	5	15.20	6.4	5	16.57	6.9	6	17.54	7.3
	.80	4	14.25	5.0	4	15.65	5.4	4	17.00	5.8	5	17.87	6.1
	1.00	3	14.81	4.2	4	15.65	4.6	4	17.00	4.9	4	18.25	5.1
	1.20	3	14.81	3.7	3	16.22	3.9	3	17.54	4.2	4	18.25	4.5
	1.40	2	15.65	3.3	3	16.22	3.4	3	17.54	3.6	3	18.98	3.8
	1.60	2	15.65	2.8	2	17.00	3.0	2	18.25	3.3	3	18.98	3.4
	1.80	2	15.65	2.4	2	17.00	2.6	2	16.25	2.7	2	19.74	3.0
1.40	0.00	6	13.45	7.8	6	14.81	8.4	7	15.91	9.0	8	17.00	9.7
	.20	5	13.81	6.6	6	14.81	7.2	6	16.22	7.7	6	17.54	8.2
	.40	4	14.25	5.7	5	15.20	6.1	5	16.57	6.6	5	17.87	7.0
	.60	4	14.25	4.9	4	15.65	5.2	4	17.00	5.6	5	17.87	6.0
	.80	3	14.81	4.2	3	16.22	4.6	4	17.00	4.8	4	18.25	5.1
	1.00	3	14.81	3.6	3	16.22	3.9	3	17.54	4.2	4	18.25	4.5
	1.20	2	15.65	3.3	3	16.22	3.4	3	17.54	3.6	3	18.98	3.8
	1.40	2	15.65	2.8	2	17.00	3.0	3	17.54	3.3	3	18.98	3.4
	1.60	2	15.65	2.4	2	17.00	2.6	2	18.25	2.8	2	19.74	3.0
	1.80	2	15.65	2.2	2	17.00	2.3	2	18.25	2.4	2	19.74	2.8
1.60	0.00	5	13.81	6.2	5	15.20	6.6	6	16.22	7.2	6	17.54	7.6
	.20	4	14.25	5.3	4	15.65	5.8	5	16.57	6.2	5	17.87	6.6
	.40	4	14.25	4.7	4	15.65	5.6	4	17.00	5.4	4	18.25	5.8
	.60	3	14.81	4.1	3	16.22	4.4	4	17.00	4.7	4	18.25	4.9
	.80	3	14.81	3.6	3	16.22	3.8	3	17.54	4.1	3	18.98	4.4
	1.00	2	15.65	3.2	3	16.22	3.4	3	17.54	3.6	3	18.98	3.8
	1.20	2	15.65	2.8	2	17.00	3.0	2	18.25	3.3	3	18.98	3.4
	1.40	2	15.65	2.4	2	17.00	2.6	2	18.25	2.8	2	19.74	3.0
	1.60	2	15.65	2.3	2	17.00	2.3	2	18.25	2.5	2	19.74	2.8
	1.80	2	15.65	2.1	2	17.00	2.2	2	18.25	2.3	2	19.74	2.4
1.80	0.00	4	14.25	5.0	4	15.65	5.4	4	17.00	5.8	5	17.87	6.1
	.20	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
	.40	3	14.81	3.9	3	16.22	4.2	4	17.00	4.6	4	18.25	4.8
	.60	3	14.81	3.5	3	16.22	3.7	3	17.54	3.9	3	18.98	4.2
	.80	2	15.65	3.1	3	16.22	3.4	3	17.54	3.5	3	18.98	3.7
	1.00	2	15.65	2.7	2	17.00	2.9	2	18.25	3.2	3	18.98	3.4
	1.20	2	15.65	2.4	2	17.00	2.6	2	18.25	2.7	2	19.74	3.0
	1.40	2	15.65	2.2	2	17.00	2.3	2	18.25	2.4	2	19.74	2.6
	1.60	2	15.65	2.1	2	17.00	2.2	2	18.25	2.3	2	19.74	2.4
	1.80	1	17.00	2.0	2	17.00	2.1	2	18.25	2.1	2	19.74	2.2

Two Characteristics, Independent Observations

		ρ=-.40											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
.20	0.00	166	6.81	254.3	199	7.83	294.7	232	8.91	335.3	267	10.02	376.2
	.20	77	8.35	112.3	89	9.44	126.8	102	10.56	141.4	114	11.71	156.0
	.40	39	9.71	55.3	44	10.85	61.5	50	11.98	67.7	55	13.18	73.8
	.60	23	10.76	32.4	26	11.90	35.7	29	13.07	39.0	32	14.25	42.3
	.80	16	11.49	21.3	17	12.75	23.4	19	13.91	25.4	21	15.10	27.4
	1.00	11	12.24	15.1	12	13.45	16.5	14	14.51	17.9	15	15.77	19.2
	1.20	8	12.87	11.4	9	14.02	12.3	10	15.20	13.3	11	16.39	14.2
	1.40	7	13.14	8.9	7	14.51	9.6	8	15.65	10.3	9	16.78	11.0
	1.60	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
	1.80	4	14.25	5.9	5	15.20	6.3	5	16.57	6.7	6	17.54	7.2
.40	0.00	58	8.91	83.8	67	10.01	94.0	76	11.14	104.3	85	12.31	114.5
	.20	39	9.71	55.3	44	10.85	61.5	50	11.98	67.7	55	13.18	73.8
	.40	25	10.60	35.4	29	11.68	39.0	32	12.87	42.7	35	14.07	46.3
	.60	17	11.37	23.8	19	12.53	26.1	21	13.71	28.4	23	14.89	30.6
	.80	12	12.06	16.9	14	13.14	18.5	15	14.38	20.0	17	15.53	21.6
	1.00	9	12.64	12.6	10	13.81	13.7	11	15.01	14.8	12	16.22	15.9
	1.20	7	13.14	9.8	8	14.25	10.6	9	15.41	11.4	9	16.78	12.2
	1.40	6	13.45	7.8	6	14.81	8.4	7	15.91	9.0	8	17.00	9.7
	1.60	5	13.81	6.4	5	15.20	6.9	6	16.22	7.5	6	17.54	7.8
	1.80	4	14.25	5.3	4	15.65	5.7	5	16.57	6.1	5	17.87	8.5
.60	0.00	30	10.23	42.6	34	11.37	47.1	38	12.53	51.6	42	13.71	56.2
	.20	23	10.76	32.4	26	11.90	35.7	29	13.07	39.0	32	14.25	42.3
	.40	17	11.37	23.8	19	12.53	26.1	21	13.71	28.4	23	14.89	30.6
	.60	13	11.90	17.6	14	13.14	19.2	16	14.25	20.8	17	15.53	22.5
	.80	10	12.43	13.4	11	13.62	14.5	12	14.81	15.7	13	16.06	16.9
	1.00	8	12.87	10.4	9	14.02	11.3	9	15.41	12.2	10	16.57	13.1
	1.20	6	13.45	8.3	7	14.51	9.0	7	15.91	9.7	8	17.00	10.4
	1.40	5	13.81	6.8	6	14.81	7.4	6	16.22	7.9	7	17.22	8.4
	1.60	4	14.25	5.7	5	15.20	6.1	5	16.57	6.5	5	17.87	7.0
	1.80	4	14.25	4.8	4	15.65	5.1	4	17.00	5.5	5	17.87	5.9
.80	0.00	19	11.14	26.1	21	12.33	28.6	23	13.53	31.2	26	14.65	33.7
	.20	16	11.49	21.3	17	12.75	23.4	19	13.91	25.4	21	15.10	27.4
	.40	12	12.06	16.9	14	13.14	18.5	15	14.38	20.0	17	15.53	21.6
	.60	10	12.43	13.4	11	13.62	14.5	12	14.81	15.7	13	16.06	16.9
	.80	8	12.87	10.7	9	14.02	11.6	10	15.20	12.5	10	16.57	13.4
	1.00	6	13.45	8.7	7	14.51	9.3	8	15.65	10.1	8	17.00	10.8
	1.20	5	13.81	7.1	6	14.81	7.7	6	16.22	8.3	7	17.26	8.8
	1.40	4	14.25	6.0	5	15.20	6.4	5	16.57	6.9	6	17.54	7.3
	1.60	4	14.25	5.0	4	15.65	5.4	4	17.00	5.8	5	17.87	6.1
	1.80	3	14.81	4.3	4	15.65	4.7	4	17.00	4.9	4	18.25	5.2
1.00	0.00	13	11.90	17.7	14	13.14	19.4	16	14.25	21.0	17	15.53	22.6
	.20	11	12.24	15.1	12	13.45	16.5	14	14.51	17.9	15	15.77	19.2
	.40	9	12.64	12.6	10	13.81	13.7	11	15.01	14.8	12	16.22	15.9

		ρ=-.40 (continued)											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
	.60	8	12.87	10.4	9	14.02	11.3	9	15.41	12.2	10	16.57	13.1
	.80	6	13.45	8.7	7	14.51	9.3	8	15.65	10.1	8	17.00	10.8
	1.00	5	13.81	7.2	6	14.81	7.8	6	16.22	8.4	7	17.26	8.9
	1.20	5	13.81	6.1	5	15.20	6.5	5	16.57	7.1	6	17.54	7.5
	1.40	4	14.25	5.2	4	15.65	5.6	5	16.57	6.0	5	17.87	6.3
	1.60	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
	1.80	3	14.81	3.8	3	16.22	4.1	3	17.54	4.5	4	18.25	4.7
1.20	0.00	10	12.43	12.9	11	13.62	14.1	12	14.81	15.2	13	16.06	16.3
	.20	8	12.87	11.4	9	14.02	12.3	10	15.20	13.3	11	16.39	14.2
	.40	7	13.14	9.8	8	14.25	10.6	9	15.41	11.4	9	16.78	12.2
	.60	6	13.45	8.3	7	14.51	9.0	7	15.91	9.7	8	17.00	10.4
	.80	5	13.81	7.1	6	14.81	7.7	6	16.22	8.3	7	17.26	8.8
	1.00	5	13.81	6.1	5	15.20	6.5	5	16.57	7.1	6	17.54	7.5
	1.20	4	14.25	5.2	4	15.65	5.6	5	16.57	6.0	5	17.87	6.4
	1.40	3	14.81	4.6	4	15.65	4.9	4	17.00	5.2	4	18.25	5.5
	1.60	3	14.81	3.9	3	16.22	4.3	4	17.00	4.6	4	18.25	4.8
	1.80	3	14.81	3.5	3	16.22	3.7	3	17.54	3.9	3	18.98	4.3
1.40	0.00	7	13.14	9.9	8	14.25	10.7	9	15.41	11.5	10	16.57	12.4
	.20	7	13.14	8.9	7	14.51	9.6	8	15.65	10.3	9	16.78	11.0
	.40	6	13.45	7.8	6	14.81	8.4	7	15.91	9.0	8	17.00	9.7
	.60	5	13.81	6.8	6	14.81	7.4	6	16.22	7.9	7	17.26	8.4
	.80	4	14.25	6.0	5	15.20	6.4	5	16.57	6.9	6	17.54	7.3
	1.00	4	14.25	5.2	4	15.65	5.6	5	16.57	6.0	5	17.87	6.3
	1.20	3	14.81	4.6	4	15.65	4.9	4	17.00	5.2	4	18.25	5.5
	1.40	3	14.81	4.0	3	16.22	4.3	4	17.00	4.6	4	18.25	4.8
	1.60	3	14.81	3.5	3	16.22	3.8	3	17.54	4.0	3	18.98	4.3
	1.80	2	15.65	3.2	3	16.22	3.4	3	17.54	3.6	3	18.98	3.8
1.60	0.00	6	13.45	7.8	6	14.81	8.5	7	15.91	9.1	8	17.00	9.7
	.20	5	13.81	7.1	6	14.81	7.6	6	16.22	8.3	7	17.26	8.8
	.40	5	13.81	6.4	5	15.20	6.9	6	16.22	7.4	6	17.54	7.8
	.60	4	14.25	5.7	5	15.20	6.1	5	16.57	6.5	5	17.87	7.0
	.80	4	14.25	5.0	4	15.65	5.4	4	17.00	5.8	5	17.87	6.1
	1.00	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
	1.20	3	14.81	3.9	3	16.22	4.3	4	17.00	4.6	4	18.25	4.8
	1.40	3	14.81	3.5	3	16.22	3.8	3	17.54	4.0	3	18.98	4.3
	1.60	2	15.65	3.2	3	16.22	3.4	3	17.54	3.6	3	18.98	3.8
	1.80	2	15.65	2.8	2	17.00	3.1	3	17.54	3.3	3	18.98	3.4
1.80	0.00	5	13.81	6.3	5	15.20	6.8	6	16.22	7.4	6	17.54	7.8
	.20	4	14.25	5.9	5	15.20	6.3	5	16.57	6.7	6	17.54	7.2
	.40	4	14.25	5.3	4	15.65	5.7	5	16.57	6.1	5	17.87	6.5
	.60	4	14.25	4.8	4	15.65	5.1	4	17.00	5.5	5	17.87	5.9
	.80	3	14.81	4.3	4	15.65	4.7	4	17.00	4.9	4	18.25	5.2
	1.00	3	14.81	3.8	3	16.22	4.1	3	17.54	4.5	4	18.25	4.7
	1.20	3	14.81	3.5	3	16.22	3.7	3	17.54	3.9	3	18.98	4.3
	1.40	2	15.65	3.2	3	16.22	3.4	3	17.54	3.6	3	18.98	3.8
	1.60	2	15.65	2.8	2	17.00	3.1	3	17.54	3.3	3	18.98	3.4
	1.80	2	15.65	2.5	2	17.00	2.7	2	18.25	2.9	2	19.74	3.2

Two Characteristics, Independent Observations

		ρ=-.20											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
.20	0.00	183	6.62	281.8	220	7.63	327.8	258	8.70	374.2	298	9.80	420.9
	.20	96	7.91	141.7	112	8.98	161.0	128	10.10	180.5	145	11.24	199.9
	.40	48	9.29	68.8	55	10.41	76.9	62	11.55	84.9	69	12.72	93.0
	.60	28	10.37	39.6	32	11.49	43.8	36	12.64	47.9	39	13.86	52.1
	.80	19	11.14	25.7	21	12.33	28.2	23	13.53	30.7	25	14.73	33.2
	1.00	13	11.90	18.1	15	13.00	19.7	16	14.25	21.4	18	15.41	23.0
	1.20	10	12.43	13.4	11	13.62	14.6	12	14.81	15.7	13	16.06	16.9
	1.40	8	12.87	10.4	9	14.02	11.3	9	15.41	12.2	10	16.57	13.0
	1.60	6	13.45	8.3	7	14.51	9.0	7	15.91	9.7	8	17.00	10.3
	1.80	5	13.81	6.8	6	14.81	7.4	6	16.22	7.9	7	17.26	8.4
.40	0.00	65	8.68	93.6	74	9.81	105.2	84	10.94	116.9	95	12.09	128.6
	.20	48	9.29	68.8	55	10.41	76.9	62	11.55	84.9	69	12.72	93.0
	.40	32	10.10	45.1	36	11.25	50.0	41	12.38	54.9	45	13.58	59.7
	.60	22	10.85	30.2	24	12.06	33.2	27	13.21	36.3	30	14.38	39.2
	.80	15	11.61	21.3	17	12.75	23.3	19	13.91	25.2	21	15.10	27.3
	1.00	12	12.06	15.7	13	13.29	17.1	14	14.51	18.5	15	15.77	19.9
	1.20	9	12.64	12.0	10	13.81	13.0	11	15.01	14.1	12	16.22	15.1
	1.40	7	13.14	9.5	8	14.25	10.3	8	15.65	11.1	9	16.78	11.8
	1.60	6	13.45	7.7	6	14.81	8.3	7	15.91	8.9	7	17.26	9.6
	1.80	5	13.81	6.4	5	15.20	6.9	6	16.22	7.4	6	17.54	7.8
.60	0.00	34	9.98	47.7	38	11.14	52.8	43	12.28	58.0	47	13.49	63.2
	.20	28	10.37	39.6	32	11.49	43.8	36	12.64	47.9	39	13.86	52.1
	.40	22	10.85	30.2	24	12.06	33.2	27	13.21	36.3	30	14.38	39.2
	.60	16	11.49	22.6	18	12.64	24.8	20	13.81	26.9	23	14.89	29.0
	.80	13	11.90	17.1	14	13.14	18.7	15	14.38	20.2	17	15.53	21.8
	1.00	10	12.43	13.2	11	13.62	14.4	12	14.81	15.5	13	16.06	16.7
	1.20	8	12.87	10.5	9	14.02	11.4	9	15.41	12.3	10	16.57	13.1
	1.40	6	13.45	8.5	7	14.51	9.2	8	15.65	9.9	8	17.00	10.6
	1.60	5	13.81	7.0	6	14.81	7.6	6	16.22	8.1	7	17.26	8.7
	1.80	4	14.25	5.9	5	15.20	6.3	5	16.57	6.8	6	17.54	7.3
.80	0.00	21	10.94	29.2	24	12.06	32.2	26	13.29	35.1	29	14.44	37.9
	.20	19	11.14	25.7	21	12.33	28.2	23	13.53	30.7	25	14.73	33.2
	.40	15	11.61	21.3	17	12.75	23.3	19	13.91	25.2	21	15.10	27.3
	.60	13	11.90	17.1	14	13.14	18.7	15	14.38	20.2	17	15.53	21.8
	.80	10	12.43	13.7	11	13.62	14.9	12	14.81	16.1	13	16.06	17.4
	1.00	8	12.87	11.1	9	14.02	12.0	10	15.20	13.0	11	16.39	13.9
	1.20	7	13.14	9.1	7	14.51	9.8	8	15.65	10.6	9	16.78	11.3
	1.40	6	13.45	7.5	6	14.81	8.1	7	15.91	8.7	7	17.26	9.3
	1.60	5	13.81	6.3	5	15.20	6.8	6	16.22	7.3	6	17.54	7.8
	1.80	4	14.25	5.4	4	15.65	5.9	5	16.57	6.2	5	17.87	6.6
1.00	0.00	15	11.61	19.9	16	12.87	21.8	18	14.02	23.6	20	15.20	25.5
	.20	13	11.90	18.1	15	13.00	19.7	16	14.25	21.4	18	15.41	23.0
	.40	12	12.06	15.7	13	13.29	17.1	14	14.51	18.5	15	15.77	19.9

		ρ=-.20 (continued)											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
	.60	10	12.43	13.2	11	13.62	14.4	12	14.81	15.5	13	16.06	16.7
	.80	8	12.87	11.1	9	14.02	12.0	10	15.20	13.0	11	16.39	13.9
	1.00	7	13.14	9.3	8	14.25	10.1	8	15.65	10.8	8	16.78	11.6
	1.20	6	13.45	7.8	6	14.81	8.5	7	15.91	9.1	8	17.00	9.7
	1.40	5	13.81	6.6	5	15.20	7.2	6	16.22	7.7	6	17.54	8.2
	1.60	4	14.25	5.7	5	15.20	6.1	5	16.57	6.5	5	17.87	7.0
	1.80	4	14.25	4.9	4	15.65	5.3	4	17.00	5.7	5	17.87	6.0
1.20	0.00	11	12.24	14.5	12	13.45	15.8	13	14.65	17.1	14	15.91	18.4
	.20	10	12.43	13.4	11	13.62	14.6	12	14.81	15.7	13	16.06	16.9
	.40	9	12.64	12.0	10	13.81	12.0	11	15.01	14.1	12	16.33	15.1
	.60	8	12.87	10.5	9	14.02	11.4	9	15.41	12.3	10	16.57	13.1
	.80	7	13.14	9.1	7	14.51	9.8	8	15.65	10.6	9	16.78	11.3
	1.00	6	13.45	7.8	6	14.81	8.5	7	15.91	9.1	8	17.00	9.7
	1.20	5	13.81	6.7	6	14.81	7.3	6	16.22	7.8	6	17.54	8.4
	1.40	4	14.25	5.9	5	15.20	6.3	5	16.57	6.7	6	17.54	7.2
	1.60	4	14.25	5.1	4	15.65	5.5	5	16.57	5.9	5	17.87	6.2
	1.80	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
1.40	0.00	8	12.87	11.1	9	14.02	12.0	10	15.20	13.0	11	16.39	13.9
	.20	8	12.87	10.4	9	14.02	11.3	9	15.41	12.2	10	16.57	13.0
	.40	7	13.14	9.5	8	14.25	10.3	8	15.65	11.1	9	16.78	11.8
	.60	6	13.45	8.5	7	14.51	9.2	8	15.65	9.9	8	17.00	10.6
	.80	6	13.45	7.5	6	14.81	8.1	7	15.91	8.7	7	17.26	9.3
	1.00	5	13.81	6.6	5	15.20	7.2	6	16.22	7.7	6	17.54	8.2
	1.20	4	14.25	5.9	5	15.20	6.3	5	16.57	6.7	6	17.54	7.2
	1.40	4	14.25	5.1	4	15.65	5.5	5	16.57	6.0	5	17.87	6.3
	1.60	3	14.81	4.6	4	15.65	4.9	4	17.00	5.2	4	18.25	5.5
	1.80	3	14.81	4.0	3	16.22	4.3	4	17.00	4.7	4	18.25	4.9
1.60	0.00	7	13.14	8.8	7	14.51	9.5	8	15.65	10.2	8	17.00	11.0
	.20	6	13.45	8.3	7	14.51	9.0	7	15.91	9.7	8	17.00	10.3
	.40	6	13.45	7.7	6	14.81	8.3	7	15.91	8.9	7	17.26	9.6
	.60	5	13.81	7.0	7	14.81	7.6	6	16.22	8.1	7	17.26	8.7
	.80	5	13.81	6.3	5	15.20	6.8	6	16.22	7.3	6	17.54	7.8
	1.00	4	14.25	5.7	5	15.20	6.1	5	16.57	8.5	5	17.87	7.0
	1.20	4	14.25	5.1	4	15.65	5.5	5	16.57	5.9	5	17.87	6.2
	1.40	3	14.81	4.6	4	15.65	4.9	4	17.00	5.2	4	18.25	5.5
	1.60	3	14.81	4.0	3	16.22	4.4	4	17.00	4.7	4	18.25	4.9
	1.80	3	14.81	3.6	3	16.22	3.9	3	17.54	4.2	4	18.25	4.5
1.80	0.00	5	13.81	7.2	6	14.81	7.7	6	16.22	8.3	7	17.26	8.8
	.20	5	13.81	6.8	6	14.81	7.4	6	16.22	7.9	7	17.26	8.4
	.40	5	13.81	6.4	5	15.20	6.9	6	16.22	7.4	6	17.54	7.8
	.60	4	14.25	5.9	5	15.20	6.3	5	16.57	6.8	6	17.54	7.3
	.80	4	14.25	5.4	4	15.65	5.9	5	16.57	6.2	5	17.87	6.6
	1.00	4	14.25	4.9	4	15.65	5.3	4	17.00	5.7	5	17.87	6.0
	1.20	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
	1.40	3	14.81	4.0	3	16.22	4.3	4	17.00	4.7	4	18.25	4.9
	1.60	3	14.81	3.6	3	16.22	3.9	3	17.54	4.2	4	18.25	4.5
	1.80	3	14.81	3.4	3	16.22	3.5	3	17.54	3.7	3	18.98	4.0

Two Characteristics, Independent Observations

		$\rho = 0.0$											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
.20	0.00	188	6.56	290.7	226	7.57	338.6	267	8.63	387.0	308	9.73	435.6
	.20	113	7.58	169.3	133	8.64	193.5	454	9.73	217.8	175	10.86	242.1
	.40	56	8.98	80.5	64	10.10	90.3	73	11.22	100.0	81	12.40	109.7
	.60	32	10.10	45.1	36	11.25	50.0	41	12.38	54.9	45	13.58	59.7
	.80	21	10.94	28.7	23	12.15	31.6	26	13.29	34.5	28	14.51	37.3
	1.00	15	11.61	19.9	16	12.87	21.8	18	14.02	23.7	20	15.20	25.5
	1.20	11	12.24	14.7	12	13.45	16.0	13	14.65	17.3	14	15.91	18.6
	1.40	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	1.60	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	1.80	6	13.45	7.4	6	14.81	7.9	7	15.91	8.5	7	17.26	9.0
.40	0.00	67	8.62	96.7	77	9.73	108.9	87	10.87	121.1	98	12.02	133.2
	.20	56	8.98	80.5	64	10.10	90.3	73	11.22	100.0	81	12.40	109.7
	.40	38	9.76	54.5	44	10.85	60.5	49	12.02	66.6	54	13.21	72.7
	.60	26	10.52	36.1	29	11.68	39.9	32	12.87	43.6	36	14.02	47.3
	.80	18	11.25	25.0	20	12.43	27.4	22	13.62	29.9	24	14.81	32.2
	1.00	13	11.90	18.1	15	12.00	19.8	16	14.25	21.5	18	15.41	23.2
	1.20	10	12.43	13.7	11	13.62	14.9	12	14.81	16.1	13	16.06	17.4
	1.40	8	12.87	10.7	9	14.02	11.6	10	15.20	12.6	10	16.57	13.5
	1.60	6	13.45	8.7	7	14.51	9.3	8	15.65	10.0	8	17.00	10.7
	1.80	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
.60	0.00	35	9.92	49.3	40	11.04	54.7	44	12.24	60.1	49	13.41	65.5
	.20	32	10.10	45.1	36	11.25	50.0	41	12.38	54.9	45	13.58	59.7
	.40	26	10.52	36.1	29	11.68	39.9	32	12.87	43.6	36	14.02	47.3
	.60	20	11.04	27.4	22	12.24	30.1	25	13.37	32.8	27	14.58	35.4
	.80	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	1.00	12	12.06	15.8	13	13.29	17.2	14	14.51	18.6	15	15.77	20.1
	1.20	9	12.64	12.4	10	13.81	13.4	11	15.01	14.5	12	16.22	15.6
	1.40	7	13.14	9.9	8	14.25	10.7	9	15.41	11.6	10	16.57	12.4
	1.60	8	13.45	8.1	7	14.51	8.8	7	15.91	9.4	8	17.00	10.1
	1.80	5	13.81	6.7	6	14.81	7.3	6	16.22	7.8	6	17.54	8.4
.80	0.00	22	10.85	30.3	24	12.06	33.3	27	13.21	36.3	30	14.38	39.3
	.20	21	10.94	28.7	23	12.15	31.6	26	13.29	34.5	28	14.51	37.3
	.40	18	11.25	25.0	20	12.43	27.4	22	13.62	29.9	24	14.81	32.2
	.60	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	.80	12	12.06	16.7	14	13.14	18.2	15	14.38	19.7	16	15.65	21.2
	1.00	10	12.43	13.4	11	13.62	14.6	12	14.81	15.8	13	16.06	17.0
	1.20	8	12.87	10.9	9	14.02	11.8	10	15.20	12.8	11	16.39	13.7
	1.40	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	1.60	6	13.45	7.5	6	14.81	8.1	7	15.91	8.7	7	17.26	9.3
	1.80	5	13.81	6.3	5	15.20	6.8	6	16.22	7.3	6	17.54	7.8
1.00	0.00	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	.20	15	11.61	19.9	16	12.87	21.8	18	14.02	23.7	20	15.20	25.5
	.40	13	11.90	18.1	15	13.00	19.8	16	14.25	21.5	18	15.41	23.2

		ρ= 0.0 (continued)											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
	.60	12	12.06	15.8	13	13.29	17.2	14	14.51	18.6	15	15.77	20.1
	.80	10	12.43	13.4	11	13.62	14.6	12	14.81	15.8	13	16.06	17.0
	1.00	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	1.20	7	13.14	9.5	8	14.25	10.3	8	15.65	11.1	9	16.78	11.8
	1.40	6	13.45	8.0	7	14.51	8.7	7	15.91	9.3	8	17.00	10.0
	1.60	5	13.81	6.8	6	14.81	7.4	6	16.22	7.9	7	17.26	8.4
	1.80	4	14.25	5.9	5	15.20	6.3	5	16.57	6.7	6	17.54	7.2
1.20	0.00	11	12.24	15.0	12	13.45	16.4	13	14.65	17.7	15	15.77	19.1
	.20	11	12.24	14.7	12	13.45	16.0	13	14.65	17.3	14	15.91	18.6
	.40	10	12.43	13.7	11	13.62	14.9	12	14.81	16.1	13	16.06	17.4
	.60	9	12.64	12.4	10	13.81	13.4	11	15.01	14.5	12	16.22	15.6
	.80	8	12.87	10.9	9	14.02	11.8	10	15.20	12.8	11	16.39	13.7
	1.00	7	13.14	9.5	8	14.25	10.3	8	15.65	11.1	9	16.78	11.8
	1.20	6	13.45	8.2	7	14.51	8.9	7	15.91	9.6	8	16.00	10.2
	1.40	5	13.81	7.1	6	15.81	7.6	6	16.22	8.2	7	17.26	8.8
	1.60	5	13.81	6.2	5	15.20	6.6	6	16.22	7.2	6	17.54	7.6
	1.80	4	14.25	5.3	4	15.65	5.8	5	16.57	6.2	5	17.87	6.5
1.40	0.00	9	12.64	11.5	9	14.02	12.5	10	15.20	13.5	11	16.39	14.4
	.20	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	.40	8	12.87	10.7	9	14.02	11.6	10	15.20	12.6	10	16.57	13.5
	.60	7	13.14	9.9	8	14.25	10.7	9	15.41	11.6	10	16.57	12.4
	.80	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	1.00	6	13.45	8.0	7	14.51	8.7	7	15.91	9.3	8	17.00	10.0
	1.20	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
	1.40	5	13.81	6.3	5	15.20	6.7	6	16.22	7.3	6	17.54	7.7
	1.60	4	14.25	5.5	5	15.20	6.0	5	16.57	8.3	5	17.87	8.8
	1.80	4	14.25	4.9	4	15.65	5.2	4	17.00	5.6	5	17.87	8.0
1.60	0.00	7	13.14	9.1	7	14.51	9.9	8	15.65	10.6	9	16.78	11.3
	.20	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	.40	6	13.45	8.7	7	14.51	9.3	8	15.65	10.0	8	17.00	10.7
	.60	6	13.45	8.1	7	14.51	8.8	7	15.91	9.4	8	17.00	10.1
	.80	6	13.45	7.5	6	14.81	8.1	7	15.91	8.7	7	17.26	9.3
	1.00	5	13.81	6.8	6	14.81	7.4	6	16.22	7.9	7	17.26	8.4
	1.20	5	13.81	6.2	5	15.20	6.6	6	16.22	7.2	6	17.54	7.6
	1.40	4	14.25	5.5	5	15.20	6.0	5	16.57	6.3	5	17.87	6.8
	1.60	4	14.25	4.9	4	15.65	5.3	4	17.00	5.7	5	17.87	6.0
	1.80	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
1.80	0.00	6	13.45	7.4	6	14.81	8.0	7	15.91	8.6	7	17.26	9.2
	.20	6	13.45	7.4	6	14.81	7.9	7	15.91	8.5	7	17.26	9.0
	.40	5	13.81	7.1	6	14.81	7.6	8	16.22	8.2	7	17.26	8.8
	.60	5	13.81	6.7	6	14.81	7.3	6	16.22	7.8	6	17.54	8.4
	.80	5	13.81	6.3	5	15.20	6.8	6	16.22	7.3	6	17.54	7.8
	1.00	4	14.25	5.9	5	15.20	6.3	5	16.57	6.7	6	17.54	7.2
	1.20	4	14.25	5.3	4	15.65	5.8	5	16.57	6.2	5	17.87	6.5
	1.40	4	14.25	4.9	4	15.65	5.2	4	17.00	5.6	5	17.87	6.0
	1.60	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
	1.80	3	14.81	4.0	3	16.22	4.3	4	17.00	4.7	4	18.25	4.9

Two Characteristics, Independent Observations

		$\rho=+0.20$											
		$L_0=5{,}000$			$L_0=10{,}000$			$L_0=20{,}000$			$L_0=40{,}000$		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
.20	0.00	183	6.62	281.8	220	7.63	327.8	258	8.70	274.2	398	9.80	420.9
	.20	130	7.30	195.6	154	8.35	224.6	178	9.44	253.7	203	10.57	282.9
	.40	62	8.78	89.9	72	9.87	101.0	81	11.02	112.1	91	12.17	123.2
	.60	34	9.98	48.6	39	11.09	53.9	44	12.24	59.2	48	13.45	64.5
	.80	22	10.85	30.2	24	12.06	33.2	27	13.21	36.3	30	14.38	39.2
	1.00	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	1.20	11	12.24	15.0	12	13.45	16.4	13	14.65	17.7	15	15.77	19.1
	1.40	9	12.64	11.5	9	14.02	12.5	10	15.20	13.4	11	16.39	14.4
	1.60	7	13.14	9.0	7	14.51	9.8	8	15.65	10.5	9	16.78	11.3
	1.80	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
.40	0.00	65	8.68	93.6	74	9.81	105.2	84	10.94	116.9	95	12.09	128.6
	.20	62	8.78	89.9	72	9.87	101.0	81	11.02	112.1	91	12.17	123.2
	.40	45	9.42	63.4	51	10.56	70.7	57	11.71	78.0	63	12.91	85.3
	.60	30	10.23	41.5	33	11.42	45.9	37	12.58	50.3	41	13.76	54.7
	.80	20	11.04	28.0	23	12.15	30.8	25	13.37	33.6	28	14.51	36.3
	1.00	15	11.61	19.9	16	12.87	21.8	18	14.02	23.6	20	15.20	25.5
	1.20	11	12.24	14.8	12	13.45	16.1	13	14.65	17.4	14	15.91	18.8
	1.40	8	12.87	11.4	9	14.02	12.4	10	15.20	13.4	11	16.39	14.3
	1.60	7	13.14	9.1	7	14.51	9.8	8	15.65	10.6	9	16.78	11.3
	1.80	6	13.45	7.4	6	14.81	8.0	7	15.91	8.6	7	17.26	9.1
.60	0.00	34	9.98	47.7	38	11.14	52.8	43	12.28	58.0	47	13.49	63.2
	.20	34	9.98	48.6	39	11.09	53.9	44	12.24	59.2	48	13.45	64.5
	.40	30	10.23	41.5	33	11.42	45.9	37	12.58	50.3	41	13.76	54.7
	.60	23	10.76	32.0	26	11.90	35.2	29	13.07	38.5	32	14.25	41.7
	.80	17	11.37	23.9	19	12.53	26.3	21	13.71	38.6	23	14.89	30.8
	1.00	13	11.90	18.1	15	13.00	19.7	16	14.25	21.4	18	15.41	23.0
	1.20	10	12.43	13.9	11	13.62	15.1	12	14.81	16.4	14	15.91	17.6
	1.40	8	12.87	11.0	9	14.02	11.9	10	15.20	12.8	11	16.39	13.8
	1.60	7	13.14	8.9	7	14.51	9.6	8	15.65	10.3	9	16.78	11.1
	1.80	5	13.81	7.3	6	14.81	7.8	7	15.91	8.5	7	17.26	9.0
.80	0.00	21	10.94	29.2	24	12.06	32.2	26	13.29	35.1	29	14.44	37.9
	.20	22	10.85	30.2	24	12.06	33.2	27	13.21	36.3	30	14.38	39.2
	.40	20	11.04	28.0	23	12.15	30.8	25	13.37	33.6	28	14.51	36.3
	.60	17	11.37	23.9	19	12.53	26.3	21	13.71	28.6	23	14.89	30.8
	.80	14	11.75	19.5	16	12.87	21.3	17	14.13	23.1	19	15.30	25.0
	1.00	12	12.06	15.7	13	13.29	17.1	14	14.51	18.5	15	15.77	19.9
	1.20	9	12.64	12.6	10	13.81	13.7	11	15.01	14.8	12	16.22	15.9
	1.40	8	12.87	10.2	8	14.25	11.1	9	15.41	11.9	10	16.57	12.8
	1.60	6	13.45	8.4	7	14.51	9.1	8	15.65	9.8	8	17.00	10.5
	1.80	5	13.81	7.0	6	14.81	7.6	6	16.22	8.1	7	17.26	8.7
1.00	0.00	15	11.61	19.9	16	12.87	21.8	18	14.02	23.6	20	15.20	25.5
	.20	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	.40	15	11.61	19.9	16	12.87	21.8	18	14.02	23.6	20	15.20	25.5

		ρ=+0.20 (continued)											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
	.60	13	11.90	18.1	15	13.00	19.7	16	14.25	21.4	18	15.41	23.0
	.80	12	12.06	15.7	13	13.29	17.1	14	14.51	18.4	15	15.77	19.9
	1.00	10	12.43	13.2	11	13.62	14.4	12	14.81	15.5	13	16.06	16.7
	1.20	8	12.87	11.1	9	14.02	12.0	10	15.20	13.0	11	16.39	13.9
	1.40	7	13.14	9.3	8	14.25	10.1	8	15.65	10.8	9	16.78	11.6
	1.60	6	13.45	7.8	6	14.81	8.5	7	15.91	9.1	8	17.00	9.7
	1.80	5	13.81	6.6	5	15.20	7.2	6	16.22	7.7	6	17.54	8.2
1.20	0.00	11	12.24	14.5	12	13.45	15.8	13	14.65	17.1	14	15.91	18.4
	.20	11	12.24	15.0	12	13.45	16.4	13	14.65	17.7	15	15.77	19.1
	.40	11	12.24	14.8	12	13.45	16.1	13	14.65	17.4	14	15.91	18.8
	.60	10	12.43	13.9	11	13.62	15.1	12	14.81	16.4	14	15.91	17.6
	.80	9	12.64	12.6	10	13.81	13.7	11	15.01	14.8	12	16.22	15.9
	1.00	8	12.87	11.1	9	14.02	12.0	10	15.20	13.0	11	16.39	13.9
	1.20	7	13.14	9.6	8	14.25	10.4	9	15.41	11.3	9	16.78	12.0
	1.40	6	13.45	8.3	7	14.51	9.0	7	15.91	9.7	8	17.00	10.3
	1.60	5	13.81	7.2	6	14.81	7.7	6	16.22	8.3	7	17.26	8.8
	1.80	5	13.81	6.2	5	15.20	6.6	6	16.22	7.2	6	17.54	7.6
1.40	0.00	8	12.87	11.1	9	14.02	12.0	10	15.20	13.0	11	16.39	13.9
	.20	9	12.64	11.5	9	14.02	12.5	10	15.20	13.4	11	16.39	14.4
	.40	8	12.87	11.4	9	14.02	12.4	10	15.20	13.4	11	16.39	14.3
	.60	8	12.87	11.0	9	14.02	11.9	10	15.20	12.8	11	16.39	13.8
	.80	8	12.87	10.2	8	14.25	11.1	9	15.41	11.9	10	16.57	12.8
	1.00	7	13.14	9.3	8	14.25	10.1	8	15.65	10.8	9	16.78	11.6
	1.20	6	13.45	8.3	7	14.51	9.0	7	15.91	9.7	8	17.00	10.3
	1.40	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
	1.60	5	13.81	6.5	5	15.20	7.0	6	16.22	7.5	6	17.54	8.0
	1.80	4	14.25	5.7	5	15.20	6.1	5	16.57	8.5	5	17.87	7.0
1.60	0.00	7	13.14	8.8	7	14.51	9.5	8	15.65	10.2	8	17.00	11.0
	.20	7	13.14	9.0	7	14.51	9.8	8	15.65	10.5	9	16.78	11.3
	.40	7	13.14	9.1	7	14.51	9.8	8	15.65	10.6	9	16.78	11.3
	.60	7	13.14	8.9	7	14.51	9.6	8	15.65	10.3	9	16.78	11.1
	.80	6	13.45	8.4	7	14.51	9.1	8	15.65	9.8	8	17.00	10.5
	1.00	6	13.45	7.8	6	14.81	8.5	7	15.91	9.1	8	17.00	9.7
	1.20	5	13.81	7.2	6	14.81	7.7	8	16.22	8.3	7	17.26	8.8
	1.40	5	13.81	6.5	5	15.20	7.0	6	16.22	7.5	6	17.54	8.0
	1.60	4	14.25	5.8	5	15.20	6.3	5	16.57	6.7	6	17.54	7.2
	1.80	4	14.25	5.2	4	15.65	5.6	5	16.57	6.0	5	17.87	6.3
1.80	0.00	5	13.81	7.2	6	14.81	7.7	6	16.22	8.3	7	17.26	8.8
	.20	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
	.40	6	13.45	7.4	6	14.81	8.0	7	15.91	8.6	7	17.26	9.1
	.60	5	13.81	7.3	6	14.81	7.8	7	15.91	8.5	7	17.26	9.0
	.80	5	13.81	7.0	6	14.81	7.6	6	16.22	8.1	7	17.26	8.7
	1.00	5	13.81	6.6	5	15.20	7.2	6	16.22	7.7	6	17.54	8.2
	1.20	5	13.81	6.2	5	15.20	6.6	6	16.22	7.2	6	17.54	7.6
	1.40	4	14.25	5.7	5	15.20	6.1	5	16.57	6.5	5	17.87	7.0
	1.60	4	14.25	5.2	5	15.65	5.6	5	16.57	6.0	5	17.87	6.3
	1.80	4	14.25	4.8	4	15.65	5.1	4	17.00	5.4	4	17.87	5.8

Two Characteristics, Independent Observations

		ρ=+0.40											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
.20	0.00	166	6.81	254.3	199	7.83	294.7	232	8.91	335.3	267	10.02	376.2
	.20	145	7.08	220.7	173	8.11	254.4	201	9.20	288.4	230	10.32	322.4
	.40	66	8.65	95.8	76	9.76	107.8	87	10.87	119.8	97	12.04	131.9
	.60	35	9.92	49.1	39	11.09	54.5	44	12.24	59.9	49	13.41	65.2
	.80	21	10.94	29.6	24	12.06	32.6	27	13.21	35.5	29	14.44	38.4
	1.00	14	11.75	19.8	16	12.87	21.7	18	14.02	23.5	19	15.30	25.4
	1.20	11	12.24	14.3	12	13.45	15.5	13	14.65	16.7	14	15.91	18.0
	1.40	8	12.87	10.7	9	14.02	11.7	10	15.20	12.6	10	16.57	13.5
	1.60	6	13.45	8.4	7	14.51	9.1	8	15.65	9.8	8	17.00	10.5
	1.80	5	13.81	6.8	6	14.81	7.4	6	16.22	7.9	7	17.26	8.4
.40	0.00	58	8.91	83.8	67	10.01	94.0	76	11.14	104.3	85	12.31	114.5
	.20	66	8.65	95.8	76	9.76	107.8	87	10.87	119.8	97	12.04	131.9
	.40	50	9.21	72.1	58	10.30	80.6	65	11.45	89.1	72	12.64	97.6
	.60	33	10.04	46.1	37	11.20	51.0	41	12.38	56.0	46	13.53	61.0
	.80	22	10.85	30.0	24	12.06	33.0	27	13.21	36.0	29	14.44	38.9
	1.00	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.5
	1.20	11	12.24	15.0	12	13.45	16.3	13	14.65	17.6	15	15.77	19.0
	1.40	8	12.87	11.4	9	14.02	12.3	10	15.20	13.3	11	16.39	14.2
	1.60	7	13.14	8.9	7	14.51	9.6	8	15.65	10.3	9	16.78	11.1
	1.80	5	13.81	7.2	7	14.81	7.7	6	16.22	8.3	7	17.26	8.8
.60	0.00	30	10.23	42.6	34	11.37	47.1	38	12.53	51.6	42	13.71	56.2
	.20	35	9.92	49.1	39	11.09	54.5	44	12.24	59.9	49	13.41	65.2
	.40	33	10.04	46.1	37	11.20	51.0	41	12.38	56.0	46	13.53	61.0
	.60	26	10.52	36.5	29	11.68	40.2	33	12.81	44.0	36	14.02	47.8
	.80	20	11.04	26.9	22	12.24	29.6	24	13.45	32.2	26	14.65	34.8
	1.00	14	11.75	19.8	16	12.87	21.7	18	14.02	23.5	19	15.30	25.4
	1.20	11	12.24	14.9	12	13.45	16.2	13	14.65	17.5	15	15.77	18.9
	1.40	9	12.64	11.5	9	14.02	12.5	10	15.20	13.5	11	16.39	14.4
	1.60	7	13.14	9.1	7	14.51	9.9	8	15.65	10.6	9	16.78	11.3
	1.80	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
.80	0.00	19	11.14	26.1	21	12.33	28.6	23	13.53	31.2	26	14.65	33.7
	.20	21	10.94	29.6	24	12.06	32.6	27	13.21	35.5	29	14.44	38.4
	.40	22	10.85	30.0	24	12.06	33.0	27	13.21	36.0	29	14.44	38.9
	.60	20	11.04	26.9	22	12.24	29.6	24	13.45	32.2	26	14.65	34.8
	.80	16	11.49	22.3	18	12.64	24.4	20	13.81	36.5	22	15.01	28.6
	1.00	13	11.90	17.7	14	13.14	19.4	16	14.25	21.0	17	15.53	22.6
	1.20	10	12.43	14.0	11	13.62	15.3	13	14.65	16.5	14	15.91	17.7
	1.40	8	12.87	11.2	9	14.02	12.1	10	15.20	13.1	11	16.39	14.0
	1.60	7	13.14	9.0	7	14.51	9.8	8	15.65	10.5	9	16.78	11.2
	1.80	6	13.45	7.4	6	14.81	8.0	7	15.91	8.6	7	17.26	9.1
1.00	0.00	13	11.90	17.7	14	13.14	19.4	16	14.25	21.0	17	15.53	22.6
	.20	14	11.75	19.8	16	12.87	21.7	18	14.02	23.5	19	15.30	25.4
	.40	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4

		$\rho=+0.40$ (continued)												
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
	.60	14	11.75	19.8	16	12.87	21.7	18	14.02	23.5	19	15.30	25.4
	.80	13	11.90	17.7	14	13.14	19.4	16	14.25	21.0	17	15.53	22.6
	1.00	11	12.24	15.1	12	13.45	16.5	14	14.51	17.9	15	15.77	19.2
	1.20	9	12.64	12.6	10	13.81	13.7	11	15.01	14.8	12	16.22	15.9
	1.40	8	12.87	10.4	9	14.02	11.3	9	15.41	12.2	10	16.57	13.1
	1.60	6	13.45	8.7	7	14.51	9.3	8	15.65	10.1	8	17.00	10.8
	1.80	5	13.81	7.2	6	14.81	7.8	6	16.22	8.4	7	17.26	8.9
1.20	0.00	10	12.43	12.9	11	13.62	14.1	12	14.81	15.2	13	16.06	16.3
	.20	11	12.24	14.3	12	13.45	15.5	13	14.65	16.7	14	15.91	18.0
	.40	11	12.24	15.0	12	13.45	16.3	13	14.65	17.6	15	15.77	19.0
	.60	11	12.24	14.9	12	13.45	16.2	13	14.65	17.5	15	15.77	18.9
	.80	10	12.43	14.0	11	13.62	15.3	13	14.65	16.5	14	15.91	17.7
	1.00	9	12.64	12.6	10	13.81	13.7	11	15.01	14.8	12	16.22	15.9
	1.20	8	12.87	11.0	9	14.02	11.9	10	15.20	12.9	11	16.39	13.8
	1.40	7	13.14	9.5	8	14.25	10.3	8	15.65	11.1	9	16.78	11.8
	1.60	6	13.45	8.1	7	14.51	8.7	7	15.91	9.4	8	17.00	10.0
	1.80	5	13.81	6.9	6	14.81	7.4	6	16.22	8.0	7	17.26	8.5
1.40	0.00	7	13.14	9.9	8	14.25	10.7	9	15.41	11.5	10	16.57	12.4
	.20	8	12.87	10.7	9	14.02	11.7	10	15.20	12.6	10	16.57	13.5
	.40	8	12.87	11.4	9	14.02	12.3	10	15.20	13.3	11	16.39	14.2
	.60	9	12.64	11.5	9	14.02	12.5	10	15.20	13.5	11	16.39	14.4
	.80	8	12.87	11.2	9	14.02	12.1	10	15.20	13.1	11	16.39	14.0
	1.00	8	12.87	10.4	9	14.02	11.3	9	15.41	12.2	10	16.57	13.1
	1.20	7	13.14	9.5	8	14.25	10.2	8	15.65	11.1	9	16.78	11.8
	1.40	6	13.45	8.4	7	14.51	9.1	8	15.65	9.8	8	17.00	10.5
	1.60	6	13.45	7.4	6	14.81	8.0	7	15.91	8.6	7	17.26	9.1
	1.80	5	13.81	6.4	5	15.20	7.0	6	16.22	7.5	6	17.54	7.9
1.60	0.00	6	13.45	7.8	6	14.81	8.5	7	15.91	9.1	8	17.00	9.7
	.20	6	13.45	8.4	7	14.51	9.1	8	15.65	9.8	8	17.00	10.5
	.40	7	13.14	8.9	7	14.51	9.6	8	15.65	10.3	9	16.78	11.1
	.60	7	13.14	9.1	7	14.51	9.9	8	15.65	10.6	9	16.78	11.3
	.80	7	13.14	9.0	7	14.51	9.8	8	15.65	10.5	9	16.78	11.2
	1.00	6	13.45	8.7	7	14.51	9.3	8	15.65	10.1	8	17.00	10.8
	1.20	6	13.45	8.1	7	14.51	8.7	7	15.91	9.4	8	17.00	10.0
	1.40	6	13.45	7.4	7	14.81	8.0	7	15.91	8.6	7	17.26	9.1
	1.60	5	13.81	6.6	6	14.81	7.2	6	16.22	7.7	6	17.54	8.2
	1.80	4	14.25	6.0	5	15.20	6.4	5	16.57	6.9	6	17.54	7.3
1.80	0.00	5	13.81	6.3	5	15.20	6.8	6	16.22	7.4	6	17.54	7.8
	.20	5	13.81	6.8	6	14.81	7.4	6	16.22	7.9	7	17.26	8.4
	.40	5	13.81	7.2	6	14.81	7.7	6	16.22	8.3	7	17.26	8.8
	.60	6	13.45	7.4	6	14.81	7.9	7	15.91	8.8	7	17.26	9.1
	.80	6	13.45	7.4	6	14.81	8.0	7	15.91	8.6	7	17.26	9.1
	1.00	5	13.81	7.2	6	14.81	7.8	6	16.22	8.4	7	17.26	8.9
	1.20	5	13.81	6.9	6	14.81	7.4	6	16.22	8.0	7	17.26	8.5
	1.40	5	13.81	6.4	5	15.20	7.0	6	16.22	7.5	6	17.54	7.9
	1.60	4	14.25	6.0	5	15.20	6.4	5	16.57	6.9	6	17.54	7.3
	1.80	4	14.25	5.4	4	15.65	5.9	5	16.57	6.2	5	17.87	6.6

Two Characteristics, Independent Observations

		ρ=+0.60											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
.20	0.00	136	7.21	205.8	161	8.26	236.6	188	9.33	267.7	214	10.46	298.8
	.20	160	6.88	244.9	191	7.92	283.3	224	8.98	322.1	257	10.10	361.0
	.40	66	8.65	95.5	76	9.76	107.5	86	10.90	119.5	97	12.04	131.4
	.60	32	10.10	45.1	36	11.25	50.0	41	12.38	54.9	45	13.58	59.7
	.80	19	11.14	26.0	21	12.33	28.6	23	13.53	21.1	26	14.65	33.7
	1.00	13	11.90	17.0	14	13.14	18.6	15	14.38	20.1	17	15.53	21.7
	1.20	9	12.64	12.0	10	13.81	13.1	11	15.01	14.1	12	16.22	15.1
	1.40	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	1.60	5	13.81	7.0	6	14.81	7.5	6	16.22	8.1	7	17.26	8.6
	1.80	4	14.25	5.6	5	15.28	6.1	5	16.57	6.4	5	17.87	6.9
.40	0.00	47	9.33	66.9	54	10.44	74.7	60	11.61	82.5	67	12.78	90.3
	.20	66	8.65	95.5	76	9.76	107.5	86	10.90	119.5	97	12.04	131.4
	.40	56	8.98	80.5	64	10.10	90.3	73	11.22	100.0	81	12.40	109.7
	.60	35	9.92	49.0	39	11.09	54.4	44	12.24	59.8	49	13.41	65.1
	.80	22	10.85	29.9	24	12.06	32.9	27	13.21	35.8	29	14.44	38.8
	1.00	14	11.75	19.6	16	12.87	21.4	18	14.02	23.2	19	15.30	25.0
	1.20	10	12.43	13.7	11	13.62	14.9	12	14.81	16.1	13	16.06	17.4
	1.40	8	12.87	10.2	8	14.25	11.0	9	15.41	11.8	10	16.57	12.7
	1.60	6	13.45	7.8	6	14.81	8.4	7	15.91	9.1	8	17.00	9.7
	1.80	5	13.81	6.2	5	15.20	6.7	6	16.22	7.2	6	17.54	7.6
.60	0.00	24	10.68	33.8	27	11.82	37.2	20	13.00	40.7	33	14.19	44.1
	.20	32	10.10	45.1	36	11.25	50.0	41	12.38	54.9	45	13.58	59.7
	.40	35	9.92	49.0	39	11.09	54.4	44	12.24	59.8	59	13.41	65.1
	.60	29	10.30	40.8	33	11.42	45.2	37	12.58	49.5	40	13.81	53.8
	.80	21	10.94	29.4	24	12.06	32.3	26	13.29	35.3	29	14.44	38.1
	1.00	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	1.20	11	12.24	14.8	12	13.45	16.2	13	14.65	17.5	14	15.91	18.8
	1.40	8	12.87	11.0	9	14.02	12.0	10	15.20	12.9	11	16.39	13.9
	1.60	6	13.45	8.5	7	14.51	9.2	8	15.65	9.9	8	17.00	10.6
	1.80	5	13.81	6.7	6	14.81	7.3	6	16.22	7.8	6	17.54	8.4
.80	0.00	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	.20	19	11.14	26.0	21	12.33	28.6	23	13.53	31.1	26	14.65	33.7
	.40	22	10.85	29.9	24	12.06	32.9	27	13.21	35.8	29	14.44	38.8
	.60	21	10.94	29.4	24	12.06	32.3	26	13.29	35.3	29	14.44	38.1
	.80	18	11.25	25.0	20	12.43	27.4	22	13.62	29.9	24	14.81	32.2
	1.00	14	11.75	19.6	16	12.87	21.4	18	14.02	23.2	19	15.30	25.0
	1.20	11	12.24	14.9	12	13.45	16.3	13	14.65	17.6	15	15.77	19.0
	1.40	9	12.64	11.5	9	14.02	12.5	10	15.20	13.5	11	16.39	14.4
	1.60	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	1.80	5	13.81	7.2	6	14.81	7.7	6	16.22	8.3	7	17.26	8.8
1.00	0.00	10	12.43	14.0	11	13.62	15.3	13	14.65	16.5	14	15.91	17.7
	.20	13	11.90	17.0	14	13.14	18.6	15	14.38	20.1	17	15.53	21.7
	.40	14	11.75	19.6	16	12.87	21.4	18	14.02	23.2	19	15.30	25.0

		ρ=+0.60 (continued)											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
	.60	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	.80	14	11.75	19.6	16	12.87	21.4	18	14.02	23.2	19	15.20	25.0
	1.00	13	11.90	17.0	14	13.14	18.6	15	14.38	20.1	17	15.53	21.7
	1.20	10	12.43	14.0	11	13.62	15.3	13	14.65	16.5	14	15.91	17.7
	1.40	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	1.60	7	13.14	9.1	7	14.51	9.9	8	15.65	10.6	9	16.78	11.3
	1.80	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
1.20	0.00	8	12.87	10.2	8	14.25	11.1	9	15.41	11.9	10	16.57	12.7
	.20	9	12.64	12.0	10	13.81	13.1	11	15.01	14.1	12	16.22	15.1
	.40	10	12.43	13.7	11	13.62	14.9	12	14.81	16.1	13	16.06	17.4
	.60	11	12.24	14.8	12	13.45	16.2	13	14.65	17.5	14	15.91	18.8
	.80	11	12.24	14.9	12	13.45	16.3	13	14.65	17.6	15	15.77	19.0
	1.00	10	12.43	14.0	11	13.62	15.3	13	14.65	16.5	14	15.91	17.7
	1.20	9	12.64	12.4	10	13.81	13.4	11	15.91	14.5	12	16.22	15.6
	1.40	8	12.87	10.5	9	14.02	11.4	9	15.41	12.4	10	16.57	13.2
	1.60	7	13.14	8.8	7	14.51	9.5	8	15.65	10.3	9	16.78	11.0
	1.80	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
1.40	0.00	6	13.45	7.8	6	14.81	8.4	7	15.91	9.0	8	17.00	9.7
	.20	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	.40	8	12.87	10.2	8	14.25	11.0	9	15.41	11.8	10	16.57	12.7
	.60	8	12.87	11.0	9	14.02	12.0	10	15.20	12.9	11	16.39	13.9
	.80	9	12.64	11.5	9	14.02	12.5	10	15.20	13.5	11	16.39	14.4
	1.00	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	1.20	8	12.87	10.5	9	14.02	11.4	9	15.41	12.4	10	16.57	13.2
	1.40	7	13.14	9.4	8	14.25	10.2	8	15.65	11.1	9	16.78	11.8
	1.60	6	13.45	8.2	7	14.51	8.9	7	15.91	9.6	8	17.00	10.2
	1.80	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
1.60	0.00	5	13.81	6.2	5	15.20	6.6	6	16.22	7.2	6	17.54	7.6
	.20	5	13.81	7.0	6	14.81	7.5	6	16.22	8.1	7	17.26	8.6
	.40	6	13.45	7.8	6	14.81	8.4	7	15.91	9.1	8	17.00	9.7
	.60	6	13.45	8.5	7	14.51	9.2	8	15.65	9.9	8	17.00	10.6
	.80	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	1.00	7	13.14	9.1	7	14.51	9.9	8	15.65	10.6	9	16.78	11.3
	1.20	7	13.14	8.8	7	14.51	9.5	8	15.65	10.3	9	16.78	11.0
	1.40	6	13.45	8.2	7	14.51	8.9	7	15.91	9.6	8	17.00	10.2
	1.60	6	13.45	7.5	6	14.81	8.1	7	15.91	8.7	7	17.26	9.3
	1.80	5	13.81	6.6	5	15.20	7.2	6	16.22	7.7	6	17.54	8.2
1.80	0.00	4	14.25	5.0	4	15.65	5.4	4	17.00	5.8	5	17.87	6.1
	.20	4	14.25	5.6	5	15.20	6.1	5	16.57	6.4	5	17.87	6.9
	.40	5	13.81	6.2	5	15.20	6.7	6	16.22	7.2	6	17.54	7.6
	.60	5	13.81	6.7	6	14.81	7.3	6	16.22	7.8	6	17.54	8.4
	.80	5	13.81	7.2	6	14.81	7.7	6	16.22	8.3	7	17.26	8.8
	1.00	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
	1.20	6	13.45	7.4	6	14.81	7.9	7	15.91	8.6	7	17.26	9.1
	1.40	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
	1.60	5	13.81	6.6	5	15.20	7.2	6	16.22	7.7	6	17.54	8.2
	1.80	5	13.81	6.1	5	15.20	6.5	5	16.57	7.1	6	17.54	7.5

Two Characteristics, Independent Observations

		ρ=+0.80											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
.20	0.00	88	8.08	130.1	103	9.15	147.6	118	10.27	165.1	133	11.41	182.6
	.20	174	6.72	268.2	209	7.74	311.4	245	8.80	354.9	283	9.90	398.7
	.40	56	8.98	80.5	64	10.10	90.3	73	11.22	100.0	81	12.40	109.7
	.60	24	10.68	33.0	27	11.82	36.4	30	13.00	39.8	33	14.19	43.1
	.80	13	11.90	17.9	15	13.00	19.6	16	14.25	21.2	18	15.41	22.9
	1.00	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	1.20	6	13.45	7.8	6	14.81	8.5	7	15.91	9.1	8	17.00	9.7
	1.40	4	14.25	5.8	5	15.20	6.2	5	16.57	6.6	6	17.54	7.1
	1.60	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
	1.80	3	14.81	3.5	3	16.22	3.8	3	17.54	4.0	3	18.98	4.3
.40	0.00	29	10.30	41.3	33	11.42	34.7	37	12.58	50.0	41	13.76	54.4
	.20	56	8.98	80.5	64	10.10	90.3	73	11.22	100.0	81	12.40	109.7
	.40	61	8.81	88.7	71	9.90	99.7	80	11.04	110.6	90	12.19	121.6
	.60	34	9.98	47.4	38	11.14	52.5	43	12.28	57.7	47	13.40	62.8
	.80	18	11.25	25.0	20	12.43	27.4	22	13.62	29.9	24	14.81	32.2
	1.00	11	12.24	15.0	12	13.45	16.3	13	14.65	17.7	15	15.77	19.0
	1.20	7	13.14	10.0	8	14.25	10.8	9	15.41	11.6	10	16.57	12.5
	1.40	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
	1.60	4	14.25	5.3	4	15.65	5.8	5	16.57	6.1	5	17.87	6.5
	1.80	3	14.81	4.1	3	16.22	4.5	4	17.00	4.8	4	18.25	5.0
.60	0.00	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	.20	24	10.68	33.0	27	11.82	36.4	30	13.00	39.8	33	14.19	43.1
	.40	34	9.98	47.4	38	11.14	52.5	43	12.28	57.7	47	13.49	62.8
	.60	32	10.10	45.1	36	11.25	50.0	41	12.38	54.9	45	13.58	59.7
	.80	22	10.85	30.1	24	12.06	33.1	27	13.21	36.1	30	14.38	39.1
	1.00	14	11.75	18.8	15	13.00	20.6	17	14.13	22.3	18	15.41	24.1
	1.20	9	12.64	12.4	10	13.81	13.4	11	15.01	14.5	12	16.22	15.6
	1.40	6	13.45	8.7	7	14.51	9.3	8	15.65	10.1	8	17.00	10.8
	1.60	5	13.81	6.4	5	15.20	6.9	6	16.22	7.4	6	17.54	7.8
	1.80	4	14.25	4.9	4	15.65	5.2	4	17.00	5.6	5	17.87	6.0
.80	0.00	9	12.64	12.5	10	13.81	13.6	11	15.01	14.7	12	16.22	15.8
	.20	13	11.90	17.9	15	13.00	19.6	16	14.25	21.2	18	15.41	22.9
	.40	18	11.25	25.0	20	12.43	27.4	22	13.62	29.9	24	14.81	32.2
	.60	22	10.85	30.1	24	12.06	33.1	27	13.21	36.1	30	14.38	39.1
	.80	20	11.04	27.7	22	12.24	30.4	25	13.37	33.1	27	14.58	35.8
	1.00	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	1.20	11	12.24	14.4	12	13.45	15.7	13	14.65	17.0	14	15.91	18.3
	1.40	8	12.87	10.2	8	14.25	11.1	9	15.41	11.9	10	16.57	12.8
	1.60	6	13.45	7.5	6	14.81	8.1	7	15.91	8.7	7	17.26	9.3
	1.80	4	14.25	5.7	5	15.20	6.1	5	16.57	6.5	5	17.87	7.0
1.00	0.00	6	13.45	8.5	7	14.51	9.1	8	15.65	9.9	8	17.00	10.5
	.20	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	.40	11	12.24	15.0	12	13.45	16.3	13	14.65	17.7	15	15.77	19.0

		ρ=+0.80 (continued)											
		L_0=5,000			L_0=10,000			L_0=20,000			L_0=40,000		
k_1	k_2	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1	n	$\chi^2_{2,\alpha}$	L_1
	.60	14	11.75	18.8	15	13.00	20.6	17	14.13	22.3	18	15.41	24.1
	.80	15	11.61	20.6	17	12.75	22.6	18	14.02	24.5	20	15.20	26.4
	1.00	14	11.75	18.8	15	13.00	20.6	17	14.13	22.3	18	15.41	24.1
	1.20	11	12.24	15.0	12	13.45	16.3	13	14.65	17.7	15	15.77	19.0
	1.40	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	1.60	6	13.45	8.5	7	14.51	9.1	8	15.65	9.9	8	17.00	10.5
	1.80	5	13.81	6.4	5	15.20	7.0	6	16.22	7.5	6	17.54	8.0
1.20	0.00	5	13.81	6.2	5	15.20	6.6	6	16.22	7.2	6	17.54	7.6
	.20	6	13.45	7.8	6	14.81	8.5	7	15.91	9.1	8	17.00	9.7
	.40	7	13.14	10.0	8	14.25	10.8	9	15.41	11.6	10	16.57	12.5
	.60	9	12.64	12.4	10	13.81	13.4	11	15.01	14.5	12	16.22	15.6
	.80	11	12.24	14.4	12	13.45	15.7	13	14.65	17.0	14	15.91	18.3
	1.00	11	12.24	15.0	12	13.45	16.3	13	14.65	17.7	15	15.77	19.0
	1.20	10	12.43	13.7	11	13.62	14.9	12	14.81	16.1	13	16.06	17.4
	1.40	8	12.87	11.4	9	14.02	12.4	10	15.20	13.4	11	16.39	14.3
	1.60	7	13.14	9.0	7	14.51	9.8	8	15.65	10.5	9	16.78	11.3
	1.80	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
1.40	0.00	4	14.25	4.7	4	15.65	5.0	4	17.00	5.4	4	18.25	5.8
	.20	4	14.25	5.8	5	15.20	6.2	5	16.57	6.6	6	17.54	7.1
	.40	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
	.60	6	13.45	8.7	7	14.51	9.3	8	15.65	10.1	8	17.00	10.8
	.80	8	12.87	10.2	8	14.25	11.1	9	15.41	11.9	10	16.57	12.8
	1.00	8	12.87	11.3	9	14.02	12.3	10	15.20	13.2	11	16.39	14.2
	1.20	8	12.87	11.4	9	14.02	12.4	10	15.20	13.4	11	16.39	14.3
	1.40	8	12.87	10.5	9	14.02	11.4	9	15.41	12.3	10	16.57	13.1
	1.60	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	1.80	6	13.45	7.4	6	14.81	8.0	7	15.91	8.6	7	17.26	9.1
1.60	0.00	3	14.81	3.7	3	16.22	3.9	3	17.54	4.2	4	18.25	4.6
	.20	3	14.81	4.5	4	15.65	4.8	4	17.00	5.1	4	18.25	5.4
	.40	4	14.25	5.3	4	15.65	5.8	5	16.57	6.1	5	17.87	6.5
	.60	5	13.81	6.4	5	15.20	6.9	6	16.22	7.4	6	17.54	7.8
	.80	6	13.45	7.5	6	14.81	8.1	7	15.91	8.7	7	17.26	9.3
	1.00	6	13.45	8.5	7	14.51	9.1	8	15.65	9.9	8	17.00	10.5
	1.20	7	13.14	9.0	7	14.51	9.8	8	15.65	10.5	9	16.78	11.3
	1.40	7	13.14	9.0	7	14.51	9.7	8	15.65	10.4	9	16.78	11.2
	1.60	6	13.45	8.3	7	14.51	9.0	7	15.91	9.7	8	17.00	10.3
	1.80	5	13.81	7.3	6	14.81	7.8	6	16.22	8.5	7	17.26	9.0
1.80	0.00	2	15.65	3.0	3	16.22	3.3	3	17.54	3.5	3	18.98	3.6
	.20	3	14.81	3.5	3	16.22	3.8	3	17.54	4.0	3	18.98	4.3
	.40	3	14.81	4.1	3	16.22	4.5	4	17.00	4.8	4	18.25	5.0
	.60	4	14.25	4.9	4	15.65	5.2	4	17.00	5.6	5	17.87	6.0
	.80	4	14.25	5.7	5	15.20	6.1	5	16.57	6.5	5	17.87	7.0
	1.00	5	13.81	6.4	5	15.20	7.0	6	16.22	7.5	6	17.54	8.0
	1.20	5	13.81	7.1	6	14.81	7.6	6	16.22	8.2	7	17.26	8.8
	1.40	6	13.45	7.4	6	14.81	8.0	7	15.91	8.6	7	17.26	9.1
	1.60	5	13.81	7.3	6	14.81	7.8	6	16.22	8.5	7	17.26	9.0
	1.80	5	13.81	6.7	6	14.81	7.3	6	16.22	7.8	6	17.54	8.4

BIBLIOGRAPHY

1. Abraham, P. B. and Weiss, G., "On the Calculation of Analytic Functions of Cyclic Matrices," Journal of Mathematical Physics, Vol. 3, pp. 340-345, 1962.

2. Alt, F. B., "Aspects of Multivariate Quality Control Charts," unpublished Master's Thesis, Georgia Institute of Technology, 1973.

3. Alt, F. B. and Deutsch, S. J., "Statistical Estimation of Intervention Effects for Vector-Valued Processes," presented at the Criminal Justice Evaluation Methods Session of the National ORSA/TIMS Meeting, Nov., 1977.

4. Alt, F. B., Deutsch, S. J., and Goode, J. J., "Estimation for the Multi-Consequence Intervention Model," to appear in American Statistical Association, Proceedings of the Statistical Computing Section, 1977.

5. Alt, F. B., Deutsch, S. J., and Walker, J. W., "Control Charts for Correlated, Multivariate Observations," Annual Technical Conference Transactions, American Society for Quality Control, pp. 360-369, 1977.

6. Alt, F. B., Goode, J. J., and Wadsworth, H. M., "Small Sample Probability Limits for the Mean of a Multivariate Normal Process," Annual Technical Conference Transactions, American Society for Quality Control, pp. 170-176, 1976.

7. Aigner, D. J., "A Compendium on Estimation of the Autoregressive-Moving Average Model from Time Series Data," International Economic Review, Vol. 12, pp. 348-371, 1971.

8. Anderson, O. D., Time Series Analysis and Forecasting, Butterworth and Co., Ltd., Kent, England, 1976.

9. Anderson, T. W., An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, Inc., New York, 1958.

10. Anderson, T. W., An Introduction to the Statistical Analysis of Time Series, John Wiley & Sons, Inc., New York, 1970.

11. Arato, M., "On the Sufficient Statistics for Stationary Gaussian Random Processes," Theor. Probability Appl., Vol. 6, pp. 199-201, 1961.

12. Basu, J. P., Odell, P. L., and Lewis, T. O., "The Effects of Intraclass Correlation on Certain Significance Tests When Samping from Multivariate Normal Population," Communications in Statistics, Vol. 3, pp. 899-908, 1974.

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

14. Box, G. E. P. and Tiao, G. C., "A Change in Level of a Non-Stationary Time Series," Biometrika, Vol. 52, pp. 181-192, 1965.

15. Campbell, D. T. and Stanley, J. C., "Experimental and Quasi-Experimental Designs for Research on Teaching," in Gage, N. L. (editor), Handbook of Research and Teaching, Rand McNally, Chicago, 1963.

16. Daniel, H. E., "The Effect of Departures from Ideal Conditions Other than Non-Normality on the t and z Tests of Significance, Proc. Camb. Phil. Soc., Vol. 34, pp. 321-328, 1938.

17. Dent, W., "Computation of the Exact Likelihood Function of an ARIMA Process," Tech. Rep. 45, Department of Statistics, University of Iowa, 1975.

18. Deutsch, S. J., " Time Series Modeling," mimeographed lecture notes for Georgia Institute of Technology course number ISyE 642: Applied Time Series, 1974.

19. Deutsch, S. J., "Probabilistic Modeling of Crime Incidence," presented at the ORSA/TIMS National Meeting, Models and Criminal Justice Session, 1974.

20. Deutsch, S. J., "Stochastic Modeling of Crime Rates," Technical Note #1, Dept. of Justice - L.E.A.A., Grant #75 NI-99-0091, 1975.

21. Deutsch, S. J., and Alt, F. B., "The Effect of Massachusetts' Gun Control Law on Gun-Related Crimes in Boston," Evaluation Quarterly, Vol. 1., No. 4, 1977.

22. Deutsch, S. J. and Rardin, R. L., "Preliminary Results from Univariate Modeling of Index Crimes in Selected Major Cities," Technical Note #4, Dept. of Justice - L.E.A.A., Grant #75 NI-99-0091, 1975.

23. Deutsch, S. J. and Rardin, R. L., "Completion of Univariate Modeling, Technical Note #9, Dept. of Justice - L.E.A.A., Grant #75 NI-99-0091, 1976.

24. Duncan, A. J., Quality Control and Industrial Statistics, 4th ed., Richard D. Irwin, Inc., Homewood, Illinois, 1974.

25. Freund, J. E., Mathematical Statistics, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962.

26. Fuller, W. A., Introduction to Statistical Time Series, John Wiley & Sons, Inc., New York, 1976.

27. Glass, G. V., "Analysis of Data on the Connecticut Speeding Crack-down as a Time-Series Quasi-Experiment," Law and Society Review, Vol. 3, pp. 55-76, 1968.

28. Glass, G. V., Willson, V. L. and Gottman, J. M., Design and Analysis of Time-Series Experiments, Colorado Associated University Press, Boulder, Colorado, 1975.

29. Goldberger, A. S., Econometric Theory, John Wiley & Sons, Inc., New York, 1964.

30. Grant, E. L. and Leavenworth, R. S. Statistical Quality Control, 4th ed., McGraw-Hill, New York, 1972.

31. Graybill, F. A., Introduction to Linear Statistical Models, Volume I, McGraw-Hill, New York, 1961.

32. Graybill, F. A., Introduction to Matrices with Applications in Statistics, Wadsworth Publishing Co., Belmont, Calif., 1969.

33. Greenberg, B. G. and Sarhan, A. E., "Matrix Inversion, Its Interest and Application in Analysis of Data," Journal of the American Statistical Association, Vol. 54, pp. 755-766, 1959.

34. Grenander, U. and Rosenblatt, M., Statistical Analysis of Stationary Time Series, John Wiley & Sons, Inc., New York, 1956.

35. Guttman, I., Wilks, S. S. and Hunter, J. S., Introductory Engineering Statistics, 2nd ed., John Wiley & Sons, Inc., New York, 1971.

36. Hannan, E. J., Multiple Time Series, John Wiley & Sons, Inc., New York, 1970.

37. Hoel, P. G., Port, S. C. and Stone, C. J., Introduction to Statistical Theory, Houghton Mifflin Co., Boston, 1971.

38. Johnson, R. A. and Bagshaw, M., "The Effect of Serial Correlation on the Performance of CUSUM Tests," Technometrics, Vol. 16, pp. 103-112, 1974.

39. Johnson, R. A. and Bagshaw, M., "The Effect of Serial Correlation on the Performances of CUSUM Tests, II," Technometrics, Vol. 17, pp. 73-80, 1975.

40. Johnson, N. L. and Kotz, S., Continuous Univariate Distributions-2 , Houghton Mifflin Co., Boston, 1970.

41. Kasriel, R. H., Undergraduate Topology, W. B. Saunders Co., Philadelphia, 1971.

42. Kendall, M. G., The Advanced Theory of Statistics, Vol. 1, 5th ed., Hafner Publishing Co., New York, 1952.

43. Kendall, M. G. and Stuart, A., The Advanced Theory of Statistics, Volume 2, Hafner Publishing Co., New York, 1967.

44. Kleijnen, J. P., Statistical Techniques in Simulation, Pt. 2, Marcel Dekker, Inc., New York, 1975.

45. Leuthold, R. M., MacCormick, A. J., Schmitz, A. and Watts, D. G., "Forecasting Daily Hog Prices and Quantities: A Study of Alternative Forecasting Techniques," Journal of the American Statistical Association, Vol. 65, pp. 90-107, 1970.

46. Nelson, C. R., Applied Time Series Analysis for Managerial Forecasting, Holden-Day, Inc., San Francisco, 1973.

47. Padia, W. L., "Effect of Autocorrelation on Probability Statements about the Mean," Masters Thesis, Laboratory of Educational Research, University of Colorado, 1973.

48. Page, E. S., "Control Charts for the Mean of a Normal Population," J. R. Statist. Soc., B., Vol. 16, pp. 131-135, 1954.

49. Press, S. J., Applied Multivariate Analysis, Holt, Rinehart, and Winston, Inc., New York, 1972.

50. Rao, C. R., Linear Statistical Inference and Its Applications, 2nd ed., John Wiley & Sons, Inc., New York, 1973.

51. Ray, W. D., "The Inverse of a Finite Toeplitz Matrix," Technometrics, Vol. 12, pp. 153-156, 1970.

52. Rohatgi, V. K., An Introduction to Probability Theory and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1976.

53. Rutherford, D. E., "Some Continuant Determinants Arising in Physics and Chemistry-II," Proc. Roy. Soc. Edin., A. Vol. LXIII, pp. 232-241, 1951.

54. Saboia, J. L. M., "Autoregressive Integrated Moving Average (ARIMA) Models for Birth Forecasting," Journal of the American Statistical Association, Vol. 72, pp. 264-270, 1977.

55. Sankaran, M., "Approximations to the Non-Central Chi-Square Distribution," Biometrika, Vol. 50, pp. 199-204, 1963.

56. Scheffé, H., The Analysis of Variance, John Wiley & Sons, Inc., New York, 1959.

57. Schilling, E. G. and Nelson, P. R., "The Effect of Non-Normality on the Control Limits of $\bar{X}$ Charts," Journal of Quality Technology, Vol. 8, No. 4, pp. 183-188, 1976.

58. Shaman, P.,"On the Inverse of the Covariance Matrix of a First Order Moving Average," Biometrika, Vol. 56, pp. 595-600, 1969.

59. Shewhart, W. A., Economic Control of Quality of Manufactured Product, D. Van Nostrand Co., Inc., New York, 1931.

60. Sidak, Z., "Rectangular Confidence Regions for the Means of Multivariate Normal Populations," Journal of the American Statistical Association, Vol. 62, pp. 626-633, 1967.

61. Silvey, S. D., Statistical Inference, Penguin Books, Baltimore, Maryland, 1970.

62. Student, "Errors of Routine Analysis," Biometrika, Vol. 19, pp. 151-164, 1927.

63. Thompson, H. E. and Tiao, G. C., "Analysis of Telephone Data: A Case Study of Forecasting Seasonal Time Series," Bell Journal of Economics and Management Science, Vol. 2, pp. 515-541, 1971.

64. Timm, N. H., Multivariate Analysis with Applications in Education and Psychology, Brooks/Cole Publishing Co., Monterey, Calif., 1975.

65. Tucker, H. G., An Introduction to Probability and Mathematical Statistics, Academic Press, New York, 1962.

66. Walker, J. W., "Basic Ideas for Making Statistical Decisions," mimeographed lecture notes for Georgia Institute of Technology course number MA 416: Mathematical Statistics-I, 1974.

67. Walsh, J. E., "Concerning the Effect of Intraclass Correlation on Certain Significance Tests," Annals of Mathematical Statistics, Vol. 18, pp. 88-96, 1947.

68. Wilks, S. S., Mathematical Statistics, John Wiley & Sons, Inc., New York, 1962.

VITA

Francis Alt was born in Baltimore, Maryland, where he attended elementary school. After graduating from Loyola High School (Towson, Maryland) in 1959, he worked full time for five years and attended Loyola Evening College as an undergraduate mathematics major on a part-time basis. He received his B.E.S. degree in Industrial Engineering and Operations Research from the Johns Hopkins University in 1967.

In Fall, 1967, he entered the graduate program in the School of Industrial and Systems Engineering at Georgia Institute of Technology and was employed as a part-time graduate teaching assistant. In Summer, 1971, he attended the Southern Region Education Board Graduate Summer Session in Statistics at Clemson University. In Fall, 1971, he joined the faculty of the School of Industrial and Systems Engineering as a part-time Instructor. The M.S. degree was awarded in March, 1974. The doctoral dissertation defense was passed on August 5, 1977, and the Ph.D. degree will be awarded in December, 1977.

In Fall, 1977, he joined the faculty of the Department of Management Science and Statistics, University of Maryland at College Park, as an Assistant Professor.

He is married to the former Judith Densmore, and they have one daughter, Elizabeth.