ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

SOME DEVELOPMENTS IN STATISTICAL QUALITY MONITORING AND QUALITY IMPROVEMENT

By

Mehdi Kiani

A THESIS

submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfillment of the requirements for
the degree of PhD in Statistics

Athens, Greece
June 2009

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

SOME DEVELOPMENTS IN STATISTICAL QUALITY MONITORING AND QUALITY IMPROVEMENT

By

Mehdi Kiani

A THESIS

submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfillment of the requirements for
the degree of PhD in Statistics

Athens, Greece
June 2009

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

Some Developments in Statistical Quality Monitoring and Quality Improvement

Mehdi Kiani

A THESIS

submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy, PhD, in Statistics

Athens, Greece
June 2009

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

Εξελίξεις στον Στατιστικό Ποιοτικό Ελεγχο και την Βελτίωση Διαδικασιών

Μέχτι Κιάνι

ΔΙΑΤΡΙΒΗ
Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Διδακτορικού Διπλώματος στη Στατιστική

Αθήνα
June 2009

DEDICATION

Στην Οικογένεια μου ανεξάντλητη πηγή αγάπης

To my family endless source of love

ACKNOWLEDGEMENTS

I want to thank supervisor of the current thesis, Professor John Panaretos for giving me the opportunity to work on my favourite field of Statistics. I want to also thank Assistant Professor Stelios Psarakis for his patience and understanding during the difficult times of this work. The support of friends and colleagues is truly invaluable. My family has helped me in every possible way and in every situation all the years of my life. The realization of this thesis would be impossible without them.

VITA

I was born in Tehran in 1972. I finished high school in June 1989 and in September 1990 I became a student in the University of Esfahan, Department of Statistics. I completed my studies in September 1994. I succeeded in the exams and became a postgraduate student of the Tarbiat Modares University, Department of Statistics. In September 1996, I received a Master's degree in Statistics and in March 2005 I was accepted as a Ph.D. student.

ABSTRACT

Mehdi Kiani

Some Developments in Statistical Quality Monitoring and Quality Improvement

June 2009

Statistical methods for quality control and quality improvement are commonly employed in industry (see, e.g. Montgomery (2005)) and are standard management practices in the service sector, where quality control serves as a means of improvement from the perspective of total quality (e.g. Stuart, et al. (1996)). This thesis deals with certain issues in both of these areas. In particular, in the first part of the thesis we address some problems related to statistical quality monitoring, specifically, to monitor the mean, the range and the standard deviation of a quality characteristic.

The Shewhart, the Bonferroni-adjustment and the analysis of means control charts are usually applied to monitor the mean of a quality characteristic. We establish a new control chart that is exactly based on normal and t distribution for known and unknown parameters, respectively. The advantages of the proposed approach against the Shewhart, the Bonferroni-adjustment, and the analysis of means approaches are presented in this thesis. Moreover, the Shewhart and the Bonferroni-adjustment R and S chart usually are applied to monitor the range and the standard deviation of a quality characteristic. We establish a new R and S chart and show the advantages of the proposed approach against the Shewhart and the Bonferroni-adjustment R and S control chart. To construct R control chart, the mean range (d_2) of the normal distribution function (NDF) should be evaluated. We present a formula for evaluation of the normal distribution function. Based on the proposed formula, we construct an extended table for the mean range of the NDF.

In the second part of the thesis, we address issues related to quality improvement. Specifically, we consider extensions for some response surface methods. These pertain to improving the estimation of the steepest ascent path that is the preliminary procedure to bring the users toward the vicinity of optimum response. To this aim, a data augmentation scheme and an extension to the modified Gram-Schmidt strategy are proposed. Using these methods, both the variance and the bias of the estimated path appear to decrease, considerably.

ΠΕΡΙΛΗΨΗ

Μέχτι Κιάνι

Εξελίξεις στον Στατιστικό Ποιοτικό Ελεγχο και την Βελτίωση Διαδικασιών

Ιόνιος 2009

Ο στατιστικός ποιοτικός έλεγχος και η βελτίωση της ποιότητας χρησιμοποιούνται συνήθως στη βιομηχανία (π.χ. Montgomery (2005)) και είναι τυπικές πρακτικές διαχείρισης στον τομέα των υπηρεσιών, όπου ο ποιοτικός έλεγχος λειτουργεί ως μέσο βελτίωσης της συνολικής ποιότητας (π.χ. Stuart, et al. (1996)). Η διατριβή αυτή, ασχολείται με συγκεκριμένα προβλήματα που αφορούν και τις δύο αυτές περιοχές. Συγκεκριμένα, στο πρώτο μέρος της διατριβής εξετάζουμε προβλήματα σχετικά με την παρακολούθηση μιας διαδικασίας ως προς την ποιότητα, ειδικότερα σε ότι αφορά τον έλεγχο της μέσης τιμής, του εύρους και της τυπικής απόκλισης ενός ποιοτικού χαρακτηριστικού.

Τα διαγράμματα Shewhart, της διόρθωσης Bonferroni όπως και το διάγραμμα της ανάλυσης των μέσων εφαρμόζονται συνήθως για να ελέγξουν τη μέση τιμή ενός ποιοτικού χαρακτηριστικού. Στη διατριβή αυτή προτείνουμε ένα νέο διάγραμμα ελέγχου το οποίο βασίζεται στην κανονική και t κατανομή, για γνωστές και άγνωστες παραμέτρους, αντίστοιχα. Εξετάζουμε τα πλεονεκτήματα της προτεινόμενης προσέγγισης σε σχέση με τη μέθοδο Shewhart, την διόρθωση Bonferroni, και την ανάλυση των μέσων τιμών. Επιπλέον, η μέθοδος Shewhart και η διόρθωση Bonferroni στα διαγράμματα R και S συνήθως εφαρμόζονται για να ελέγξουν το εύρος και την τυπική απόκλιση ενός ποιοτικού χαρακτηριστικού. Προτείνουμε ένα νέο διάγραμμα R και S και παρουσιάζουμε τα πλεονεκτήματα της προτεινόμενης προσέγγισης ενάντι της μεθόδου Shewhart και της διόρθωσης Bonferroni για το διάγραμμα ελέγχου R και S. Για την κατασκευή του διαγράμματος ελέγχου R, απαιτείται ο προσδιορισμός του μέσου εύρους (d_2) της κανονικής κατανομής (NDF). Παρουσιάζουμε έναν τύπο για

την αποτίμηση της συνάρτησης της κανονικής κατανομής με βάση τον οποίο κατασκευάζουμε έναν αναλυτικό πίνακα για το μέσο εύρος της NDF.

Στο δεύτερο μέρος της διατριβής, αντιμετωπίζουμε ζητήματα σχετικά με τη βελτίωση της ποιότητας. Συγκεκριμένα, εξετάζουμε κάποιες επεκτάσεις για συγκεκριμένες μεθόδους απόκρισης. Οι μέθοδοι αυτοί αφορούν τη βελτίωση της εκτίμησης της πιο απότομης ανοδικής πορείας η οποία είναι η προκαταρκτική διαδικασία η οποία θα οδηγήσει τους χρήστες προς την εγγύτητα της βέλτιστης απάντησης. Για το σκοπό αυτό, προτείνονται ένα σχήμα συμπλήρωσης στοιχείων και μια επέκταση στην τροποποιημένη μέθοδο Gram-Schmidt. Χρησιμοποιώντας αυτές τις μεθόδους, τόσο η διασπορά όσο και η μεροληψία της εκτιμηθείσας πορείας εμφανίζουν μία αξιοσημείωτη μείωση.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure page

1 Introduction

In this thesis, we focus on two main areas of statistical methods in quality improvement: statistical process control and response surface methodology.

Statistical process control is the collection of strategies for recognizing particular causes of variation, bringing a process into a state of control and for reducing variation about a target variable. It is extensively utilized in industry to retain manufacturing processes under control condition. The most valuable tool of statistical process control is control charts that we deal with it in chapters 2 to 5. These charts yield a graphical appearance of the process thus enabling any manager with or without statistical knowledge to immediately understand if the process is under control or not.

Statistical strategies for quality improvement require the use of more than just control charts. It would be difficult to keep a particular process characteristic in control condition without some information of the factors affecting that characteristic. Hence, in chapters 5 and 6, we look at response surface methodology as a subsection of statistically designed experiments.

Statistical procedures in quality improvement have had a long history. Walter A. Shewhart of the Bell Telephone Laboratories developed the statistical control-chart concept, in 1924. By the middle of the 1930s, statistical quality-control methods were in wide use at Western Electric, the manufacturing arm of the Bell System. For more details see Montgomery (2005), who presents a timeline of quality procedures from years 1700 through 2000s.

Control charts are the primary procedures of statistical process control. Some known and some newly proposed control charts are studied in the present thesis. The Shewhart, Bonferroni-adjustment and analysis of means (ANOM) control charts are typically applied to monitor the mean of a quality characteristic. The Shewhart and Bonferroni procedures are utilized to recognize special causes of variation in the production process, where the control limits are constructed by assuming a normal distribution for when the parameters (mean and standard deviation) are known and approximately normal distribution when the parameters are unknown. The ANOM method is an alternative to the analysis of variance method. It can be used to establish the mean control charts by applying the equicorrelated multivariate non-central t distribution. In this study, we establish new control charts, in phases I and II monitoring, based on normal and t distributions having as a cause a known (or unknown) parameter (standard deviation). Our proposed methods are at least as effective as the classical Shewhart methods and have some advantages.

The control limits of the average of a quality characteristic depend on the variability, the range or the standard deviation, of the production process. When the process variability is out of the control limits, the control limits on

the average quality characteristic will not have much meaning. The Shewhart and the Bonferroni-adjustment R and S charts usually are applied to monitor the range and the standard deviation of a quality characteristic. These charts are used to recognize the process variability, where the control limits are constructed by using approximately normal distribution for the known and unknown standard deviation parameter. In chapter 3, we establish new R and S charts that are based approximately on the normal distribution. The constant values to construct the new control limits are depended on both the sample group size (k) and the sample subgroup size (n) hence, the constant values for the Shewhart and Bonferroni approach are dependent only on the sample subgroup size (n). Additionally, the unknown standard deviation for the proposed approaches is estimated by a uniformly minimum variance unbiased estimator (UMVUE). This estimator has variance less than the one of the Shewhart and Bonferroni approach.

For constructing R control charts, the mean range (d_2) of the normal distribution function (NDF) must be evaluated. Chapter 4 presents a formula for the evaluation of the normal distribution function, $F(z)$, with greatest absolute error less than 4.02×10^{-14}. The proposed formula is based on the values $z \in (-\infty,+\infty)$ and is generated by applying polar integral. Furthermore, this chapter provides a precise evaluation and extended tables for the mean range of the normal distribution (d_2), using the proposed formula of the NDF. The value d_2 can be computed by existing commands of the available statistical software without requiring computer programming.

To improve the performance of a production process, we introduce an extension of the response surface methodology, available in the literature of statistical quality control. The response surface methodology is a heuristic that locally fits first-order models, and estimates the corresponding steepest ascent paths. The eventual objective of response surface methodology is to determine the optimum operating conditions for the factor space in which operating requirements are satisfied. Particularly, if the system is being investigated for the first time, starting conditions over operability region would often be not very close to an optimum. In these circumstances, the objective of the experimenter is to move rapidly and efficiently along a path of improvement toward the vicinity of optimum. The experimenters usually need a preliminary procedure, called steepest ascent, to bring them to a suitable point, optimum, and then they employ a more elaborate model as a second-order model to locate the optimum. In chapter 5, we develop a method that would apply more generally than methods currently available. We also extend the constrained path of steepest ascent, the confidence region for limited angle to true and estimated path, the generalized confidence region for the direction of the steepest ascent, and the confidence cone about the estimated path that is based on t and *Beta* distributions.

The parameters of the steepest ascent path are estimated by the ordinary least squares estimators. To improve the bias and the variance of the estimation of the parameters, we present strategies for specifying additional data to be included along with the data of a non-orthogonal design. The additional observations increase the magnitude of the information matrix X'X and the orthogonality of the design matrix. The new runs are created in a predefined

spherical or rectangular region. Optimum number of additional observations is presented in order to orthogonalize the design matrix X and optimize some function of the information matrix X'X. Comparisons of the results acquired with the proposed methods versus the most commonly used procedures for data augmentation are carried out. In addition, the advantages of the use of our techniques over the studied methods to solve the augmenting data problems are discussed.

In addition, to improve the bias and the variance of the estimations of the parameters of the steepest ascent path, we propose an extension of the modified Gram-Schmidt algorithm, for constructing an optimal design matrix. The proposed algorithm presents an orthogonal basis, in full working accuracy, for the space spanned by the columns of the original matrix. The method discussed makes use of this modified Gram-Schmidt strategy and employs Gaussian elimination.

In chapter 9, some final thoughts and a discussion for possible future research issues and generalizations are given for the different problems addressed.

2 A New Procedure to Monitor the Mean of a Quality Characteristic

CHAPTER OUTLINE

2-1. Introduction

The Shewhart, the Bonferroni-adjustment and the analysis of means control charts are common techniques for monitoring the process mean. Shewhart (1931) proposed a scheme for detecting out-of-control signals and shifts in the mean from its target value μ_0. Ott (1975), Rocke (1989), Ryan (1989), Chen (1997), Quesenberry (1997), Smith (1998), Maravelakis, et al., (2002), Champ and Jones (2004), Woodall et al. (2004), Montgomery (2005), and several other authors modified and extended the Shewhart control charts. The Shewhart procedure usually is based on at least 20 to 25 sample group sizes (k) and at least 4 to 6 sample subgroup sizes (n). This procedure with known mean and standard deviation parameters is based on a random variable that follows the normal distribution. When the mean and standard deviation are unknown the procedure is based on a statistic that follows approximately the normal distribution. The values of the subgroup averages ($\bar{X}_{i.} = \sum_{j=1}^{n} X_{ij} / n$) are plotted on the Shewhart control chart that includes the center line $E(\bar{X}_{i.})$ and the control limits $E(\bar{X}_{i.}) \pm Z_{\alpha/2}\sqrt{\text{var}(\bar{X}_{i.})}$, where the quality characteristics X_{ij} for $i = 1,2,...,k$ and $j = 1, 2,...,n$ (j^{th} observation in i^{th} subgroup) are assumed to be independent identically normally distributed with mean μ and variance σ^2.

Ryan (1989) introduced the Bonferroni-adjustment control limits as an alternative to the Shewhart approach. The control limits are given by $E(\bar{X}_{i.}) \pm Z_{\alpha/2k}\sqrt{\text{var}(\bar{X}_{i.})}$. In other words, to construct the Bonferroni control

limits the value α of the Shewhart control limits is replaced by the value α/k.

Ott (1967) introduced the ANOM control limits (see also Nedumaran and Pignatiello (2005)) for comparing a group of means in order to see if any one of them differs significantly from the overall mean. Schilling (1973) extended this scheme to what he called the ANOM for treatment effects or ANOME. Ott's procedure is carried out by comparing the sample mean values to the overall grand mean, about which decision lines have been constructed. If a sample mean lies outside these decision lines, it is declared to be significantly different from the grand mean. The main difference between the Bonferroni and ANOM control limits is that in the first the sample group and subgroup sizes (k,n) are usually as large as 20 or more $(k \geq 20)$, and 4 or more $(n \geq 4)$, respectively to compute the control limits, whereas in the second $k \geq 2$ and $n \geq 2$ is sufficient to compute the decision lines.

Ott's method is based on the multiple significance test proposed by Halperin et al. (1955). Later, Nelson (1982) obtained the exact critical points of $h_{(\alpha/2,k,\upsilon)}$, and used the decision lines $\bar{\bar{X}}_{..} \pm h_{(\alpha/2,k,\upsilon)} S_b \sqrt{(k-1)/(kn)}$, where the critical point $h_{(\alpha/2,k,\upsilon)}$ depends on k, $\upsilon = k(n-1)$ (degrees of freedom in S_b), and the significance level α, with,

$$S_b = \left(\sum_{i=1}^{k} \sum_{j=1}^{n} (X_{ij} - \bar{X}_{i.})^2 / (k(n-1)) \right)^{1/2} \quad ; \quad \bar{\bar{X}}_{..} = \sum_{i=1}^{k} \sum_{j=1}^{n} X_{ij} / (kn).$$

Some other applications of the ANOM for testing the interaction effects were investigated by Ramig (1983), Nelson (1988), Wludyka and Nelson (1997), and Budsaba et al. (2000). A full review of the ANOM technique is given by Rao (2005).

According to equicorrelated multivariate non-central t distribution for constructing the ANOM scheme, Tsai et al. (2005) examined a control chart for a random variable $W_i = (\bar{X}_{i.} - \bar{\bar{X}}_{..})$, with the center line 0, and the control limits $0 \pm t_{\alpha/2,\upsilon}\sqrt{\bar{V}\frac{k+1}{kn}}$, where

$\bar{V} = \sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)^2/(k(n-1))$ and $\upsilon = k(n-1)$.

This control chart was introduced by Yang and Hillier (1970).

In this chapter, we represent new control charts given by Kiani, Panaretos and Psarakis (2008). The purpose in phase I is to perceive the stability and variation in a process over time. We are concerned with ongoing monitoring to detect assignable causes in the process in phase II controlling. Useful recognitions of phase I and phase II applications have been studied already, for example, by Kang and Albin (2000), Woodall (2000), Hawkins et al. (2003), Woodall et al. (2004), Montgomery (2005), and Jensen et al. (2006).

The proposed control limits with known or unknown σ are established for random variables that follow the normal distribution and t distribution, exactly. Another property of the proposed methods is that the values of both sample group and subgroup sizes (k and n) for computing the control limits, need to be grater than 1.

When the parameters μ and σ are unknown, to construct the Shewhart and Bonferroni chart we often use the estimates of the parameters, whereas, as will be shown, for the proposed charts we do not use any estimators. The effects of the estimation on the performance of control charts have been studied by Ghosh et al. (1981), Quesenberry (1993), Chakraborti (2000), Maravelakis (2003), and Albers and Kallenberg (2004).

The chapter is organized as follows. In sections 2, 3, 4, we set out the Shewhart, Bonferroni, ANOM, and new control charts, respectively. The probability of a false alarm for the Shewhart and the strategy proposed here, as well as the probability of at least a false alarm for the Bonferroni and ANOM strategy are compared in section 5. The in-control average run lengths are described in sections 6 for the Shewhart and the proposed charts. In section 7, the results and some recommendations for constructing the control limits are presented.

2-2. The Shewhart and Bonferroni Control Chart

Assume that the random variables X_{ij}, for $i=1,2,...,k$ and $j=1,2,...,n$, which measures the quality of process, are independent normally distributed with mean μ and variance σ^2. The Shewhart control limits for this quality characteristic with known parameters and confidence $1-\alpha$ are $\mu \pm Z_{\alpha/2}\sigma/\sqrt{n}$, where the center line of control chart is μ. If the mean and standard deviation of the quality characteristic are unknown, they are estimated by the unbiased statistics $\bar{\bar{X}}_{..}$ and $\bar{S}/c_4$ where,

$$\bar{S} = \sum_{i=1}^{k} S_i / k \ ; \ S_i = (\sum_{j=1}^{n} (X_{ij} - \bar{X}_{i.})^2 / (n-1))^{1/2} \ ; \ c_4 = (\frac{2}{n-1})^{1/2} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}.$$

The random variable $S_i\sqrt{(n-1)}/\sigma$ is chi distributed with $n-1$ degrees of freedom. The mean and the standard deviation of the statistic S_i are $c_4\sigma$ and $\sigma\sqrt{1-c_4^2}$, respectively. (The constant value c_4 depends only on the sample subgroup size (n)).

The Shewhart chart with unknown parameters is constructed on the statistic $(\bar{X}_{i.} - \bar{\bar{X}}_{..}) / (\bar{S}/(c_4\sqrt{n}))$ in phase I and $(\bar{X}_{i.}^{f} - \bar{\bar{X}}_{..}) / (\bar{S}/(c_4\sqrt{n}))$ in phase II, where $\bar{X}_{i.}^{f}$ indicates a subgroup average for future observations. These statistics follow approximately the normal distribution for large sample sizes. As a consequence, the center line and the control limits for the Shewhart chart with unknown parameters are,

$$\begin{aligned} U\hat{C}L &= \bar{\bar{X}}_{..} + Z_{\alpha/2}\bar{S}/(c_4\sqrt{n}) \ ; \\ \hat{C}L &= \bar{\bar{X}}_{..} \ ; \\ L\hat{C}L &= \bar{\bar{X}}_{..} - Z_{\alpha/2}\bar{S}/(c_4\sqrt{n}) \ . \end{aligned} \qquad (1)$$

The unknown standard deviation σ can be also estimated by the unbiased statistic $\bar{R}/d_2$, where the statistic $\bar{R}$ is the average range and the constant value d_2 is the mean range of the standard normal variables. This statistic gives the Shewhart control limits as

$$\bar{\bar{X}}_{..} \pm Z_{\alpha/2}\bar{R}/(d_2\sqrt{n}) \ . \qquad (2)$$

Equation (2) is also based approximately on the normal distribution with large sample sizes.

The Bonferroni-adjustment control chart to improve the probability of one or more false alarms of the Shewhart chart was suggested by Ryan (1989). The Bonferroni-adjustment control limits with known and unknown parameters for retrospective monitoring in phase I are, respectively,

$$\mu \pm Z_{\alpha/2k}\,\sigma/\sqrt{n}\,, \tag{3}$$

$$\bar{\bar{X}}_{..} + Z_{\alpha/2k}\bar{S}/(c_4\sqrt{n})\,. \tag{4}$$

For constructing equations (3) and (4), the value α of Shewhart control limits is replaced by the value α/k.

2-3. The Analysis of Means Control limits

The analysis of means can be thought of as an alternative to the Bonferroni method, since it also considers a group of sample averages instead of one average at a time in order to determine whether any of the sample averages differ much from the overall mean. The construction of ours and the ANOM strategies are based on the t distribution, hence a brief description of the ANOM technique is presented here.

The random variables X_{ij} are iid normal variables with mean μ and variance σ^2. Therefore, in phase I, the correlated random variables $\bar{X}_{i.} - \bar{\bar{X}}_{..}$ and $\bar{X}_{i'.} - \bar{\bar{X}}_{..}$ for $i \neq i' = 1,2,...,k$, follow the normal distribution with mean 0 and variance $\sigma^2(k-1)/(kn)$, (see Appendix I).

Let $T_i = (\bar{X}_{i.} - \bar{\bar{X}}_{..})/S_{\bar{X}_{i.}-\bar{\bar{X}}_{..}}$. The ANOM control limits is based on the joint statistic $(T_1, T_2, ..., T_k)$ that is equicorrelated multivariate non-central t distributed with equicorrelations $\rho = -1/(k-1)$. The statistic T_i follows the t distribution with $k(n-1)$ degrees of freedom. Here,

$$S_{\bar{X}_{i.}-\bar{\bar{X}}_{..}} = \hat{\sigma}_{\bar{X}_{i.}-\bar{\bar{X}}_{..}} = S_b\sqrt{(k-1)/(kn)} \ ;$$

$$S_b^2 = \hat{\sigma}^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X}_{i.})^2/(k(n-1)).$$

Nelson (1982) defined the joint probability of T_i for $i = 1, 2, ..., k$ as $P\left[\bigcap_{i=1}^{k} |T_i| \le h_{(\alpha/2,k,v)}\right] = 1-\alpha$. Thus, $P[|T_i| \le h_{(\alpha/2,k,v)}] = 1-\alpha' \ge 1-\alpha$, such that α' is unknown and $\alpha' \le \alpha$. This probability results the ANOM control limits with center line $\bar{\bar{X}}_{..}$ and approximately the following limits,

$$UCL = \bar{\bar{X}}_{..} + h_{(\alpha/2,k,v)}S_b\sqrt{(k-1)/(kn)} \ ;$$

$$LCL = \bar{\bar{X}}_{..} - h_{(\alpha/2,k,v)}S_b\sqrt{(k-1)/(kn)} \ . \qquad (5)$$

Here, the exact critical values $h_{(\alpha/2,k,v)}$ depend on the desired level of significance (α), the sample sizes k, and the degrees of freedom $v = k(n-1)$.

Nelson (1982) and (1993) calculated the critical values $h_{(\alpha/2,k,v)}$ to satisfy

$$P[|T_1| \le h_{(\alpha/2,k,v)}, |T_2| \le h_{(\alpha/2,k,v)}, ..., |T_k| \le h_{(\alpha/2,k,v)}] = 1-\alpha \ .$$

The left side of this equation is,

$$K\int_0^\infty \int_0^\infty [g(sh,y,\rho)]^k s^{\upsilon-1} \exp\left[-(y^2+\upsilon s^2)/2\right] d_y d_s , \quad (6)$$

where,

$$g(sh,y,\rho) = 2\mathrm{Re}\left[\Phi\left\{\frac{sh-y\sqrt{\rho}}{\sqrt{1-\rho}}\right\}\right]-1;$$

$$\Phi(x+iy) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left[-(u+iy)^2/2\right] du;$$

$$K = 2\sqrt{\frac{2}{\pi}}\left[\frac{\upsilon}{2}\right]^{\upsilon/2} / \Gamma[\upsilon/2].$$

The function $\mathrm{Re}[\Phi(.)]$ is the real part of $\Phi(.)$, and $i=\sqrt{-1}$. Nelson (1993) numerically evaluated the double integral (6). The values $h_{\alpha/2,k,\upsilon}$ can be computed by replacing different values of the desired level α and the constants k and υ. Tables of the critical values $h_{\alpha/2,k,\upsilon}$ are given by Nelson (1993) for various values of k, υ and α.

2-4. A New Control Chart

As previously suggested, the new charts to monitor the mean quality characteristic with known or unknown parameter σ are based on the normal and the t distribution, respectively. The proposed charts are dependent only on the parameter σ.

In phase I, we have $\bar{X}_{i.} - \bar{\bar{X}}_{..} \sim N(0, \sigma^2(k-1)/(kn))$. Therefore, with known variance σ^2, the new control limits are

$$\bar{\bar{X}}_{..} \pm Z_{\alpha/2}\sigma\sqrt{(k-1)/(kn)},$$

(7) since, $P(|\bar{X}_{i.} - \bar{\bar{X}}_{..}|/[\sigma\sqrt{(k-1)/(kn)}] \le Z_{\alpha/2}) = 1-\alpha$. In this case, the center

line is $\bar{\bar{X}}_{..}$. For the construction of new control chart with unknown parameter σ^2, it is known that the random variables $(\bar{X}_{i.}-\bar{\bar{X}}_{..})/\sqrt{\sigma^2(k-1)/(kn)}$ and $k(n-1)S_b^2/\sigma^2=\sum_i^k\sum_j^n(X_{ij}-\bar{X}_{i.})^2/\sigma^2$ follow the standard normal distribution and the chi-square distribution, respectively. According to *Cochran's Theorem*, these random variables are independent. Therefore, the following statistic is t distributed with $k(n-1)$ degrees of freedom ($k>1$ and $n>1$),

$$T_i=\frac{(\bar{X}_{i.}-\bar{\bar{X}}_{..})/\sqrt{\sigma^2(k-1)/(kn)}}{\sqrt{\frac{k(n-1)S_b^2/\sigma^2}{k(n-1)}}}=\frac{\bar{X}_{i.}-\bar{\bar{X}}_{..}}{S_b\sqrt{(k-1)/(kn)}}\sim t\big(k(n-1)\big).$$

As a result, the new control chart with unknown variance is given by:

$$\begin{aligned} UCL &= \bar{\bar{X}}_{..}+t_{\alpha/2,k(n-1)}S_b\sqrt{(k-1)/(kn)} \\ CL &= \bar{\bar{X}}_{..} \\ LCL &= \bar{\bar{X}}_{..}-t_{\alpha/2,k(n-1)}S_b\sqrt{(k-1)/(kn)}\,, \end{aligned} \tag{8}$$

where, $P(|\bar{X}_{i.}-\bar{\bar{X}}_{..}|/[S_b\sqrt{(k-1)/(kn)}]\le t_{\alpha/2,k(n-1)})=1-\alpha$.

For controlling future subgroups, $\bar{X}_{i.}^f$, the variance of $\bar{X}_{i.}^f-\bar{\bar{X}}_{..}$ is evaluated to be $\sigma^2(k+1)/kn$. In phase II, the random variables $(\bar{X}_{i.}^f-\bar{\bar{X}}_{..})/(\sigma\sqrt{(k+1)/(kn)})$ and $(\bar{X}_{i.}^f-\bar{\bar{X}}_{..})/(S_b\sqrt{(k+1)/(kn)})$ follow the standard normal distribution and the t distribution, respectively. As a result, the proposed control limits, in phase II, with known and unknown σ are,

$$\bar{\bar{X}}_{..} \pm Z_{\alpha/2}\sigma\sqrt{(k+1)/(kn)}\,, \tag{9}$$

$$\bar{\bar{X}}_{..} \pm t_{\alpha/2,k(n-1)}S_b\sqrt{(k+1)/(kn)}\,. \tag{10}$$

Here, the sample group and subgroup sizes required to construct our proposed charts, with known and unknown parameter σ, are greater than 1, i.e. $k>1$ and $n>1$. The above control limits were first derived by Yang and Hillier (1970). This was pointed out to the author by one of the external examiners. The author was unaware of this reference.

2-5. The Performance of Retrospective Charts

Let the individual events G_i denote that the subgroup averages $\bar{X}_{i.}$ exceed the control limits of in control process. If these events are independent, then the sequence of trials comparing $\bar{X}_{i.}$ with UCL will be a sequence of Bernoulli trials and the overall occurrences of G_i will be a Binomial random variable with parameters k and $P(G_i)$. However, in the case of unknown parameters, these events for the Bonferroni and ANOM control limits are not independent. Hence a performance comparison between these charts for historical data in phase I is given based on a simulation study. We also use simulation to study the probability $P(G_i)$ for the estimated Shewhart chart, since the control limits are the approximations of true limits. For our proposed charts, in the case of the known and unknown parameter σ, the $P(G_i)$ can be easily evaluated theoretically.

For the Shewhart chart with known parameters the probability of at least a false signal is $1-(1-\alpha)^k$, since the events G_i follow the Binomial

distribution. Ryan (1989) showed that this probability is approximately equal to $k\alpha$. Hence, Ryan suggested the Bonferroni-adjustment scheme for the mean control limits, where the probability of one or more false alarms is improved to the desired value $\alpha \simeq 1-(1-\alpha/k)^k$, which is less than $1-(1-\alpha)^k$ for the Shewhart scheme. As already mentioned, the ANOM method is an alternative to the Bonferroni method, maintaining approximately the overall false alarm probability at the desired α. Nedumaran and Pignatiello (2005) compared this probability for the Bonferroni and ANOM procedures. The performance measure for these charts is the overall probability of a false signal. Based on their study, the actual probability of having at least one false alarm, using Monte Carlo simulation experiments (20,000 times), for the ANOM approach is slightly less than the one of the Bonferroni approach, and very close to the desired value α.

To compare the Shewhart scheme to our scheme, we use a performance measure the probability of a false alarm. In this case, the k subgroups of size n are generated (20,000 times) from a stable in-control *iid* normal process (see Appendix III). It should be noted that the FORTRAN program used by Nelson (1993) can also be adapted to evaluate the performance of the proposed scheme (8) when the variance is unknown. The estimated control limits are obtained according to (1) for the Shewhart strategy with unknown parameters and according to (7) and (8) for our strategies with known and unknown parameter. Table 2.1 shows the results of the estimated probability of a false alarm.

Table 2.1 Estimated probability of a false alarm, for intended $\alpha = 0.1$.

k	Approach/n	5	10	20
	(1)	0.0721	0.0783	0.0823
5	(7)	0.1101	0.0996	0.0947
	(8)	0.1091	0.0989	0.0982
	(1)	0.0801	0.0831	0.0868
15	(7)	0.0994	0.1023	0.0987
	(8)	0.0978	0.1039	0.0980
	(1)	0.0861	0.0918	0.0901
25	(7)	0.1026	0.0997	0.1031
	(8)	0.1062	0.0941	0.1063
	(1)	0.0939	0.0924	0.0966
35	(7)	0.1006	0.0989	0.1011
	(8)	0.1018	0.0972	0.1028

Table 2.1 *(Continued)* Estimated probability of a false alarm, for intended $\alpha = 0.01$.

k	Approach/n	5	10	20
	(1)	0.00762	0.00772	0.00881
5	(7)	0.01016	0.00994	0.00976
	(8)	0.01093	0.00979	0.00973
	(1)	0.00832	0.00853	0.00872
15	(7)	0.00992	0.01034	0.00991
	(8)	0.00988	0.01019	0.00981
	(1)	0.00870	0.00921	0.00911
25	(7)	0.01023	0.00989	0.01021
	(8)	0.01054	0.00932	0.01070
	(1)	0.00949	0.00928	0.00974
35	(7)	0.01009	0.00987	0.01010
	(8)	0.01014	0.00976	0.01022

Table 2.1 *(Continued)* Estimated probability of a false alarm, for intended $\alpha = 0.001$.

k	Approach/n	5	10	20
5	(1)	0.000759	0.000780	0.000871
	(7)	0.001027	0.000989	0.000980
	(8)	0.001094	0.000978	0.000975
15	(1)	0.000840	0.000861	0.000869
	(7)	0.000991	0.001032	0.000994
	(8)	0.000990	0.001018	0.000983
25	(1)	0.000871	0.000925	0.000917
	(7)	0.001020	0.000987	0.001023
	(8)	0.001051	0.000933	0.001072
35	(1)	0.000951	0.000924	0.000971
	(7)	0.001010	0.000984	0.001015
	(8)	0.001011	0.000973	0.001011

It can be concluded that the proposed new schemes, for small and large sizes k and n, perform better than the Shewhart scheme, in the sense that, the estimated false alarm probability of the proposed schemes is very close to the intended α. Indeed, in theory the desired α can be exactly attained applying the proposed schemes (7) and (8). However, because of

the small errors in simulation experiments and the fact that the random sample sizes are not large enough this cannot be achieved.

2-6. Average Run Length

The average run length (ARL) is the average number of subgroups that are plotted before a subgroup average indicates an out-of-control condition. The ARL can be calculated as $ARL = 1/p$, under the condition that the process observations are uncorrelated. Here, p is the probability that a point exceeds the control limits.

The average run length is considered for future subgroups, when the process is in control i.e. $\mu = \mu_0$, by plotting each subgroup on the control chart immediately after each sample is collected. Let the individual events G_i^f denote that the subgroup averages $\bar{X}_{i.}^f$ exceeds the control limits of the in control process.

In the case where the events G_i^f are independent, the sequence of trials, to compare $\bar{X}_{i.}^f$ with UCL, will be a sequence of Bernoulli trials and the run length between occurrences of G_i^f will be a Geometric random variable with probability $P(G_i^f)$. The in-control average run length will be $1/\alpha = 1/P(G_i^f)$,

$$P(G_i^f) = P(\bar{X}_{i.}^f \le LCL \text{ or } \bar{X}_{i.}^f \ge UCL \,|\, \mu = \mu_0).$$

Quesenberry (1993) suggested that the $P(G_i^f)$ for a classical Shewhart 3σ control chart in case of the known parameters is equal to $\alpha = 0.0027$, and with unknown parameters is approximately,

$$P(G_i^f) = 2[1-\Phi(3\{1+\frac{1}{k}[1+\frac{9(1-c_4^2)}{c_4^2}]\}^{-1/2})], \qquad (11)$$

where $\Phi(.)$ indicates the standard normal distribution function. Using equation (11), the $P(G_i^f)$ for the often recommended values $k=20$ and $n=4$ is 0.0048, which is greater than the intended $\alpha=0.0027$. Quesenberry (1993) recommended sample sizes of about $400/(n-1)$ to construct the classical Shewhart chart. Following this recommendation, for $m=133$ and $n=4$, the intended probability of a false alarm, i.e. 0.0027, will be obtained. As a result, the usual recommendations on the sample sizes are not sufficient to ensure that the Shewhart estimated control limits are close enough to the true limits. The $P(G_i^f)$ for the proposed methods (9) and (10) with known and unknown σ is equal to the desired value α for both small and large sample sizes.

The events G_i^f and $G_{i'}^f$, $i \neq i'$, for the Shewhart chart with known parameters are uncorrelated, since the control limits are the constant values and the subgroup averages $\bar{X}_{i.}^f$ and $\bar{X}_{i'.}^f$ are independent. Thus, the run length between occurrences of G_i^f is a Geometric random variable with probability $P(G_i^f)=\alpha$ and the ARL equals to $1/\alpha$. But, these events for the Shewhart chart with unknown parameters and the proposed chart with known and unknown parameter are not independent, since the random variables $\bar{X}_{i.}^f - U\hat{C}L$ and $\bar{X}_{i'.}^f - U\hat{C}L$ are not independent. The correlation between these random variables for the Shewhart method is,

$$corr(\bar{X}_{i.}^{f} - U\hat{C}L, \bar{X}_{i'.}^{f} - U\hat{C}L) = [1 + k\{1 + \frac{z_{\alpha/2}^2(1-c_4^2)}{c_4^2}\}^{-1}]^{-1},$$

while, for the proposed method with known parameter is $1/(k+1)$, and for unknown parameter is $[1+k\{1+(k+1)t_{\alpha/2,k(n-1)}^2(1-\psi^2)\}^{-1}]^{-1}$, (see Appendix II). In this case, ψ is an unbiasing factor to estimate σ, where $E(S_b/\psi)=\sigma$, $\text{var}(S_b/\psi)=\sigma^2(1-\psi^2)/\psi^2$. The statistic S_b is chi distributed with $k(n-1)$ degrees of freedom. Based on the raw moment function of chi distribution, ψ is,

$$\psi = \sqrt{\frac{2}{k(n-1)}}\Gamma\left(\frac{k(n-1)+1}{2}\right)/\Gamma\left(\frac{k(n-1)}{2}\right).$$

The correlations evaluated for the Shewhart and the proposed methods rely on k and n, where these are always positive. These correlations decrease when we use larger sample sizes k and n. As a consequence, for the Shewhart method with unknown parameters (1) and the proposed method with known and unknown parameter (9) and (10), the distribution of run length between occurrences of the events G_i^f is not a Geometric distribution. Hence, when the parameters are unknown, the ARL cannot be evaluated based on the mean of a Geometric distribution. To overcome this problem, the $ARL = 1/\alpha$ is estimated by the simulation experiments. The existence of correlation between the events G_i^f increases the ARL, making it greater than the intended ARL. Under these circumstances, the control limits (10) are not suitable for accomplishing the desired ARL.

Hence, we propose the following approximate control limits as an alternative for (10),

$$\bar{\bar{\bar{X}}}_{..} \pm t_{\alpha/2,k(n-1)} S_b \psi / \sqrt{n} \,. \tag{12}$$

Table 2.2 shows the results of simulation experiments for equations (1), (9), (10) and (12), (see Appendix IV). For each entry in Table 2.2, the mean control limits are computed corresponding to k samples of size n, and future samples are generated from an in control process until a subgroup average is found outside the control limits. The number of samples is one observation from the run length distribution. This procedure is replicated 20,000 times. Each table entry is the average of observations from the run length distribution.

Table 2.2 Estimated in-control ARL, for intended $\alpha = 0.1$.

k	Approach/n	5	10	20
5	(1)	7.27	7.81	8.84
	(9)	11.54	11.41	11.02
	(10)	16.80	16.21	15.41
	(12)	9.09	9.21	9.12
15	(1)	7.43	8.25	8.92
	(9)	10.31	11.02	10.06
	(10)	15.40	14.96	13.83
	(12)	9.46	9.49	10.01
25	(1)	8.31	9.19	8.83
	(9)	10.12	9.98	10.04
	(10)	13.76	14.03	13.89
	(12)	9.72	10.08	9.17
35	(1)	9.39	9.42	9.31
	(9)	10.64	10.14	10.11
	(10)	12.62	13.72	12.31
	(12)	10.19	9.90	9.94

Table 2.2 *(Continued)* Estimated in-control ARL, for intended $\alpha = 0.01$.

k	Approach/n	5	10	20
5	(1)	71.84	72.92	84.57
	(9)	111.74	110.10	111.21
	(10)	163.71	167.87	150.10
	(12)	92.26	91.91	91.93
15	(1)	74.51	81.72	85.01
	(9)	101.01	110.00	103.47
	(10)	152.31	140.89	138.92
	(12)	92.41	95.01	103.91
25	(1)	81.82	95.28	85.74
	(9)	103.65	94.67	101.44
	(10)	134.27	141.65	138.23
	(12)	98.56	102.76	98.98
35	(1)	97.28	96.01	91.48
	(9)	101.26	105.28	102.42
	(10)	123.85	137.96	128.54
	(12)	102.15	99.79	99.08

Table 2.2 *(Continued)* Estimated in-control ARL, for intended $\alpha = 0.001$.

k	Approach/n	5	10	20
5	(1)	717.94	783.51	858.18
	(9)	1213.04	1128.24	1194.95
	(10)	1582.19	1691.87	1558.74
	(12)	951.39	9175.09	963.53
15	(1)	778.36	8127.58	848.38
	(9)	1098.47	1181.27	1029.12
	(10)	1583.42	1426.19	1359.80
	(12)	931.25	969.48	1081.79
25	(1)	825.05	931.24	860.09
	(9)	1047.93	962.28	1104.28
	(10)	1389.28	1442.38	1362.86
	(12)	989.16	1032.19	989.54
35	(1)	974.29	971.38	921.93
	(9)	1040.19	1071.39	1031.75
	(10)	1267.86	1381.86	1271.19
	(12)	1031.03	993.45	995.84

As already mentioned, the probability $P(G_i^f)$, corresponding to (1), is approximated to be greater than the intended α. This indicates a reason to decrease the in control ARL for the Shewhart scheme. On the other hand, the correlation between the events G_i^f causes an increase of the ARL. Based on Table 2.2, it can be concluded that the ARL for the classical Shewhart scheme is less than the desired ARL. For the proposed limits (10) the ARL is greater than $1/\alpha$, although the $P(G_i)$ is exactly equal α. This is due to the correlation between the events G_i^f. According to simulation experiments, the performance of the proposed schemes (9) and (12), to achieve the intended in control ARL is more satisfactory than the one of the schemes (1) and (10). The probability of a false alarm for the scheme (9) is equal to α, and for the scheme (12) is relatively greater than α.

We have discussed the ARL for the existing method and the proposed one using simulation experiments. Discussion for analytical methods for evaluating the ARL can be found in Burroughs, Rigdon, and Champ (1995), Jones, Champ, and Rigdon (2001), and Jones, Champ, and Rigdon (2004).

2-7. Conclusion

It has been shown that the procedures suggested in this chapter, in both phases I and II, have three advantages over the classical Shewhart method: first the proposed scheme is established using small sample sizes; second the in-control ARL of the new procedure is very close to the

desired ARL; third the false alarm probability corresponding to the proposed methods equals the intended α.

It has been suggested in the literature to use the ANOM and the Bonferroni procedures to monitor historical data in phase I controlling. These methods maintain the overall false alarm probability approximately at a desired level α. The ANOM scheme performs better than the Bonferroni technique in achieving an overall probability of a false signal at the desired α.

We recommend using the proposed strategies if the individual occurrence of events G_i and G_i^f is required, and the ANOM strategy if the overall occurrence of events G_i is considered. The ANOM and the proposed methods are constructed on the statistic $\bar{X}_{i.} - \bar{\bar{X}}_{..}$ that includes more information than $\bar{X}_{i.}$ used for the Shewhart and Bonferroni methods. Moreover, the distribution function of $\bar{X}_{i.} - \bar{\bar{X}}_{..}$ relies only on the parameter σ, whereas, that of $\bar{X}_{i.}$ depends on both parameters μ and σ.

Appendix I:

$$Var(\bar{X}_{i.} - \bar{\bar{X}}_{..}) = Var(\bar{X}_{i.} - \textstyle\sum_i^k \bar{X}_{i.}/k)$$

$$= Var(\bar{X}_{i.}) - 2Cov(\bar{X}_{i.}, \bar{X}_{i.}/k) + Var(\textstyle\sum_i^k \bar{X}_{i.}/k)$$

$$= \frac{\sigma^2}{n} - \frac{2\sigma^2}{kn} + \frac{\sigma^2}{kn}$$

$$= \frac{\sigma^2(k-1)}{kn}$$

$$Cov(\bar{X}_{i.} - \bar{\bar{X}}_{..}, \bar{X}_{j.} - \bar{\bar{X}}_{..}) = Cov(\bar{X}_{i.}, \bar{\bar{X}}_{j.}) - Cov(\bar{X}_{i.}, \bar{\bar{X}}_{..}) - Cov(\bar{\bar{X}}_{..}, \bar{X}_{j.}) + Cov(\bar{\bar{X}}_{..}, \bar{\bar{X}}_{..})$$

$$= 0 - Cov(\bar{X}_{i.}, \sum_{i=1}^{k} \bar{X}_{i.}/k) - Cov(\sum_{j=1}^{k} \bar{X}_{j.}/k, \bar{X}_{j.}) + Var(\bar{\bar{X}}_{..})$$

$$= -1/k\, Var(\bar{X}_{i.}) - 1/k\, Var(\bar{X}_{j.}) + Var(\bar{\bar{X}}_{..})$$

$$= -\frac{\sigma^2}{kn} - \frac{\sigma^2}{kn} + \frac{\sigma^2}{kn}$$

$$= \frac{-\sigma^2}{kn}$$

$$Corr(\bar{X}_{i.} - \bar{\bar{X}}_{..}, \bar{X}_{j.} - \bar{\bar{X}}_{..}) = \frac{\left(Cov(\bar{X}_{i.} - \bar{\bar{X}}_{..}, \bar{X}_{j.} - \bar{\bar{X}}_{..})\right)}{\sqrt{Var(\bar{X}_{i.} - \bar{\bar{X}}_{..}) Var(\bar{X}_{j.} - \bar{\bar{X}}_{..})}} =$$

$$= \left(\frac{-\sigma^2}{kn}\right) / \left(\frac{\sigma^2(k-1)}{kn}\right)$$

$$= \frac{-1}{k-1}$$

Appendix II:

$$Cov(\bar{X}_{i.}^{f} - U\hat{C}L, \bar{X}_{i'.}^{f} - U\hat{C}L) = Var(U\hat{C}L)$$

$$= Var(\bar{\bar{X}}_{..} + t_{\alpha/2,k(n-1)} S_b \sqrt{(k+1)/kn})$$

$$= Var(\bar{\bar{X}}_{..}) + Var(t_{\alpha/2,k(n-1)} S_b \sqrt{(k+1)/kn}) + 0$$

$$= \sigma^2/kn + (t^2_{\alpha/2,k(n-1)}(k+1)/kn) Var(S_b)$$

$$= \sigma^2/kn + (t^2_{\alpha/2,k(n-1)}(k+1)/kn)(\sigma^2(1-\psi^2))$$

$$= (\sigma^2/kn)(1 + t^2_{\alpha/2,k(n-1)}(k+1)(1-\psi^2)).$$

Appendix III:

(Maple)

```
with(LinearAlgebra):
K:=15; N:=5; a:=0.1;
ns:=0; np:=0;
 if a=0.1 then z:=1.645 end if; if a=0.01 then z:=2.575 end if;
if a=0.1 and N=2 and K=5  then t:=2.015 end if;
if a=0.1 and N=2 and K=15 then t:=1.753 end if;
if a=0.1 and N=2 and K=25 then t:=1.708 end if;
if a=0.1 and N=5 and K=5  then t:=1.725 end if;
if a=0.1 and N=5 and K=15 then t:=1.671 end if;
if a=0.1 and N=5 and K=25 then t:=1.66 end if;
if a=0.1 and N=10 and K=5  then t:=1.810 end if;
if a=0.1 and N=10 and K=15 then t:=1.650 end if;
if a=0.1 and N=10 and K=25 then t:=1.645 end if;
if a=0.01 and N=2 and K=5  then t:=4.032 end if;
if a=0.01 and N=2 and K=15 then t:=2.947 end if;
if a=0.01 and N=2 and K=25 then t:=2.787 end if;
```

```
if a=0.01 and N=5 and K=5  then t:=2.845 end if;
if a=0.01 and N=5 and K=15 then t:=2.660 end if;
if a=0.01 and N=5 and K=25 then t:=2.630 end if;
if a=0.01 and N=10 and K=5  then t:=2.691 end if;
if a=0.01 and N=10 and K=15 then t:=2.580 end if;
if a=0.01 and N=10 and K=25 then t:=2.576 end if;
if N=2 then c4:=0.7979; end if;
if N=5 then c4:=0.9400; end if;
if N=10 then c4:=0.9727; end if;
with(Statistics):
for r from 1 to 200 do
C:=RandomMatrix(K,N);Xb:=RandomMatrix(K,1);
oneN:=RandomMatrix(N,1); oneK:=RandomMatrix(1,K);
for j from 1 to N do
oneN[j,1]:=1;
end do;
for i from 1 to K do
X := RandomVariable(Normal(6, 16));
C[i,1..N]:=Sample(X, N);
Xb[i,1]:= C[i,1..N].oneN[1..N,1]/N;
oneK[1,i]:=1;
end do;
C;
Xbb:=(oneK.Xb)/K;x22:=0;S2:=0;
for i from 1 to K do; x2:=0;
  for j from 1 to N do
     x2:=x2+(C[i,j]-Xb[i,1])^2;
  end do;S1:=sqrt(x2/(N-1));S2:=S2+S1;
  x22:=x22+x2;
```

```
end do; Sb:=sqrt(x22/(K*(N-1))); Sbb:=S2/K;
UCLS:=evalf(Xbb+z*Sbb/(c4*sqrt(N)), 10);LCLS:=evalf(Xbb-
z*Sbb/(c4*sqrt(N)),10);UCLp:=evalf(Xbb+t*Sb*sqrt((K-
1)/(K*N)),10);LCLp:=evalf(Xbb-t*Sb*sqrt((K-1)/(K*N)),10);
for i from 1 to K do;
if Xb[i,1]>UCLS[1,1] or Xb[i,1]<LCLS[1,1] then ns:=ns+1; end if;
if Xb[i,1]>UCLp[1,1] or Xb[i,1]<LCLp[1,1] then np:=np+1; end if;
end do;
end do;
ns; evalf(ns/(K*(r-1)),5); np; evalf(np/(K*(r-1)),5);
```

Appendix IV:

(Maple)

```
with(LinearAlgebra):
K:=5; N:=5; a:=0.1; R:=20; muu:=2; sig:=3;
ns:=0; np:=0; nsII:=0; npII:=0;
Xf:=RandomMatrix(1,N);Xbf:=RandomMatrix(1,1);Rs:=0; RpII:=0;RpkII:=0;
RpkkII:=0; RpI:=0;RpkI:=0; RpkkI:=0; Rsk:=0;
if a=0.1 then z:=1.645 end if; if a=0.01 then z:=2.575 end if;
if a=0.1 then z:=1.645 end if; if a=0.01 then z:=2.575 end if;
if a=0.1 and N=2 and K=5  then t:=2.015; end if;
if a=0.1 and N=2 and K=15 then t:=1.753; end if;
if a=0.1 and N=2 and K=25 then t:=1.708; end if;
if a=0.1 and N=5 and K=5  then t:=1.725; end if;
if a=0.1 and N=5 and K=15 then t:=1.671; end if;
if a=0.1 and N=5 and K=25 then t:=1.660; end if;
if a=0.1 and N=10 and K=5  then t:=1.810; end if;
if a=0.1 and N=10 and K=15 then t:=1.650; end if;
```

```
if a=0.1 and N=10 and K=25 then t:=1.645; end if;
if a=0.01 and N=2 and K=5  then t:=4.032; end if;
if a=0.01 and N=2 and K=15 then t:=2.947; end if;
if a=0.01 and N=2 and K=25 then t:=2.787; end if;
if a=0.01 and N=5 and K=5  then t:=2.845; end if;
if a=0.01 and N=5 and K=15 then t:=2.660; end if;
if a=0.01 and N=5 and K=25 then t:=2.630; end if;
if a=0.01 and N=10 and K=5  then t:=2.691; end if;
if a=0.01 and N=10 and K=15 then t:=2.580;  end if;
if a=0.01 and N=10 and K=25 then t:=2.576; end if;
ps:=evalf(sqrt(2/(K*(N-1)))*GAMMA((K*(N-1)+1)/2)/GAMMA((K*(N-
1))/2),10);
if N=2 then c4:=0.7979; end if;
if N=5 then c4:=0.9400; end if;
if N=10 then c4:=0.9727; end if;
with(Statistics):
for r from 1 to R do
  C:=RandomMatrix(K,N);Xb:=RandomMatrix(K,1);
  oneN:=RandomMatrix(N,1); oneK:=RandomMatrix(1,K);
  for j from 1 to N do
    oneN[j,1]:=1;
  end do;
  for i from 1 to K do
    X := RandomVariable(Normal(muu, sig));
    C[i,1..N]:=Sample(X, N);
    Xb[i,1]:= C[i,1..N].oneN[1..N,1]/N;
    oneK[1,i]:=1;
  end do;
  Xbb:=(oneK.Xb)/K;x22:=0;S2:=0;
```

```
for i from 1 to K do; x2:=0;
  for j from 1 to N do
  x2:=x2+(C[i,j]-Xb[i,1])^2;
  end do;
  S1:=sqrt(x2/(N-1));S2:=S2+S1;
  x22:=x22+x2;
end do;
Sb:=sqrt(x22/(K*(N-1))); Sbb:=S2/K;
UCLS:=evalf(Xbb+z*Sbb/(c4*sqrt(N)), 10);
LCLS:=evalf(Xbb-z*Sbb/(c4*sqrt(N)),10);
UCLpII:=evalf(Xbb+t*Sb*sqrt((K+1)/(K*N))*sqrt((K)/(K+1))*ps,10);
LCLpII:=evalf(Xbb-t*Sb*sqrt((K+1)/(K*N))*sqrt((K)/(K+1))*ps,10);
UCLpkII:=evalf(Xbb +z*sig*sqrt((K+1)/(K*N)),10);
LCLpkII:=evalf(Xbb -z*sig*sqrt((K+1)/(K*N)),10);
UCLpkkII:=evalf(muu +z*sig*sqrt((K+1)/(K*N))*sqrt((K)/(K+1))*ps,10);
LCLpkkII:=evalf(muu -z*sig*sqrt((K+1)/(K*N))*sqrt((K)/(K+1))*ps,10);
UCLpI:=evalf(Xbb+t*Sb*sqrt((K-1)/(K*N)),10);
LCLpI:=evalf(Xbb-t*Sb*sqrt((K-1)/(K*N)),10);
UCLpkI:=evalf(Xbb +z*sig*sqrt((K-1)/(K*N)),10);
LCLpkI:=evalf(Xbb -z*sig*sqrt((K-1)/(K*N)),10);
UCLpkkI:=evalf(muu +z*sig*sqrt((K-1)/(K*N)),10);
LCLpkkI:=evalf(muu -z*sig*sqrt((K-1)/(K*N)),10);
UCLSk:=evalf(muu +z*sig/sqrt(N),10);
LCLSk:=evalf(muu -z*sig/sqrt(N),10);
rs:=0;hs:=0; rpII:=0;hpII:=0; rpkII:=0;hpkII:=0; rpkkII:=0;hpkkII:=0;
rpI:=0;hpI:=0; rpkI:=0;hpkI:=0; rpkkI:=0;hpkkI:=0; rsk:=0;hsk:=0;
for ii from 1 by 1 while hs=0 do
Xf[1,1..N]:=Sample(X,N); Xbf:= Xf.oneN/N;
  if Xbf[1,1]<UCLS[1,1] and Xbf[1,1]>LCLS[1,1] then rs:=rs+1;
```

```
    else hs:=1;
  end if;
end do;
for ii from 1 by 1 while hpII=0 do
Xf[1,1..N]:=Sample(X,N); Xbf:= Xf.oneN/N;
  if Xbf[1,1]<UCLpII[1,1] and Xbf[1,1]>LCLpII[1,1] then rpII:=rpII+1;
    else hpII:=1;
  end if;
end do;
for ii from 1 by 1 while hpkII=0 do
Xf[1,1..N]:=Sample(X,N); Xbf:= Xf.oneN/N;
  if Xbf[1,1]<UCLpkII[1,1] and Xbf[1,1]>LCLpkII[1,1] then rpkII:=rpkII+1;
    else hpkII:=1;
  end if;
end do;
for ii from 1 by 1 while hpkkII=0 do
Xf[1,1..N]:=Sample(X,N); Xbf:= Xf.oneN/N;
  if Xbf[1,1]<UCLpkkII and Xbf[1,1]>LCLpkkII then rpkkII:=rpkkII+1;
    else hpkkII:=1;
  end if;
end do;
for ii from 1 by 1 while hpI=0 do
Xf[1,1..N]:=Sample(X,N); Xbf:= Xf.oneN/N;
  if Xbf[1,1]<UCLpI[1,1] and Xbf[1,1]>LCLpI[1,1] then rpI:=rpI+1;
    else hpI:=1;
  end if;
end do;
for ii from 1 by 1 while hpkI=0 do
Xf[1,1..N]:=Sample(X,N); Xbf:= Xf.oneN/N;
```

```
  if Xbf[1,1]<UCLpkI[1,1] and Xbf[1,1]>LCLpkI[1,1] then rpkI:=rpkI+1;
    else hpkI:=1;
  end if;
end do;
for ii from 1 by 1 while hpkkI=0 do
Xf[1,1..N]:=Sample(X,N); Xbf:= Xf.oneN/N;
  if Xbf[1,1]<UCLpkkI and Xbf[1,1]>LCLpkkI then rpkkI:=rpkkI+1;
    else hpkkI:=1;
  end if;
end do;
for ii from 1 by 1 while hsk=0 do
Xf[1,1..N]:=Sample(X,N); Xbf:= Xf.oneN/N;
  if Xbf[1,1]<UCLSk and Xbf[1,1]>LCLSk then rsk:=rsk+1;
    else hsk:=1;
  end if;
end do;
Rs:=Rs+rs;
RpII:=RpII+rpII;
RpkII:=RpkII+rpkII;
RpkkII:=RpkkII+rpkkII;
RpI:=RpI+rpI;
RpkI:=RpkI+rpkI;
RpkkI:=RpkkI+rpkkI;
Rsk:=Rsk+rsk;
for i from 1 to R do;
  Xf[1,1..N]:=Sample(X,N); Xbf:= Xf.oneN/N;
  if Xbf[1,1]>UCLS[1,1] or Xbf[1,1]<LCLS[1,1] then nsII:=nsII+1; end if;
  if Xbf[1,1]>UCLpII[1,1] or Xbf[1,1]<LCLpII[1,1] then npII:=npII+1; end if;
end do;
```

```
end do;
evalf(Rs/(R),5);
evalf(RpII/(R),5);
evalf(RpkII/(R),5);
evalf(RpkkII/(R),5);
evalf(RpI/(R),5);
evalf(RpkI/(R),5);
evalf(RpkkI/(R),5);
evalf(Rsk/(R),5);
evalf(nsII/(R^2),5);
evalf(npII/(R^2),5);
```

3 A New Procedure to Monitor the Range and Standard Deviation of a Quality Characteristic

CHAPTER OUTLINE

3-1. Introduction

The Shewhart and Bonferroni-adjustment control chart are common techniques for monitoring the process range and standard deviation of a quality characteristic. The Shewhart range and standard deviation control chart were introduced by Shewhart (1931). Ott (1975), Ryan (1989), Quesenberry (1997), Smith (1998) among others extended the Shewhart range and standard deviation control charts. The Shewhart procedure usually is based on sample group sizes (k) of at least 20 to 25 and on sample subgroup sizes (n) of at least 4 to 6. The Shewhart chart with known and unknown standard deviation parameter is based on a random variable that follows approximately the normal distribution.

In the case of the R chart, the values of the subgroup ranges (R_i) are plotted on a chart that includes the center line $E(R_i)$ and the following control limits

$$E(R_i) \pm Z_{\alpha/2}\sqrt{Var(R_i)}.$$

Here, the quality characteristics X_{ij} for $i = 1,2,...,k$ and $j = 1, 2,...,n$ (jth observation in ith subgroup) are supposed to be identically independently distributed according to the normal distribution with mean μ and variance σ^2, and $R_i = X_{i(n)} - X_{i(1)}$. Here, $X_{i(1)}$ and $X_{i(n)}$ are order statistics of the random variable X_{ij} for the ith subgroup while $E(R_i)$ and $\sqrt{Var(R_i)}$ are the mean and standard deviation of R_i.

It is well known (see e.g. Johnson, et al. (1994)), that the joint probability density function of $X_{i(1)}$ and $X_{i(n)}$ is given by

$$f_{1,n}(x,y)=\begin{cases} n(n-1)\,[F(y)-F(x)]^{n-2} f(x)f(y), & x<y \\ 0 & x\geq y. \end{cases}$$

Therefore, the joint probability density function of $X_{i(1)}$ and R would be

$$f_{1,R}(x,r)=n(n-1)\,[F(r+x)-F(x)]^{n-2} f(x)f(r+x).$$

As a result, the probability density function of R is obtained to be

$$f_R(r)=\int_{-\infty}^{+\infty} n(n-1)\,[F(r+x)-F(x)]^{n-2} f(x)f(r+x)\,dx,$$

where the functions $f(x)$ and $F(x)$ are respectively the probability density function and the cumulative density function of the normal random variable X with parameters (μ,σ^2). The mean range $\big(E(R)\big)$ can be evaluated to be

$$E(R)=\int_0^{+\infty}\int_{-\infty}^{+\infty} r\times n(n-1)\,[F(r+x)-F(x)]^{n-2} f(x)f(r+x)\,dx\,dr$$

$$=\int_{-\infty}^{+\infty}\left\{1-\big(F(x)\big)^n-\big(1-F(x)\big)^n\right\}dx \qquad (1)$$

(see, Hartley (1942), Pearson and Hartley (1970), Barnard (1978) and Johnson, et al. (1994)).

Let $Z_{ij}=\dfrac{X_{ij}-\mu}{\sigma}\overset{iid}{\sim}N(0,1)$. Then the cumulative distribution function of $Z_{ij}=Z$ is,

$$F(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{\frac{-t^2}{2}}dt \quad , \qquad -\infty<z<+\infty. \qquad (2)$$

Further, let R' denote the range of order statistics $Z_{h(1)}, Z_{h(2)},..., Z_{h(n)}$ for the hth subgroup ($h=1,2,...,k$). Then, using equations (1) and (2), the mean of R' is given by

$$E(R')=\int_{-\infty}^{+\infty}\left\{1-\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{\frac{-t^2}{2}}dt\right)^n-\left(1-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{\frac{-t^2}{2}}dt\right)^n\right\}dz\,.$$

The variance and the covariance of order statistics $Z_{h(1)}, Z_{h(2)}, ...Z_{h(n)}$ can be extended as (see Johnson et al. (1994))

$$Var(Z_{h(i)})=\frac{p_iq_i}{n+2}\{(F^{-1})_i'\}^2+\frac{p_iq_i}{(n+2)^2}\Big\{2(q_i-p_i)(F^{-1})_i'(F^{-1})_i''$$

$$+p_iq_i\{(F^{-1})_i'(F^{-1})_i'''+\frac{1}{2}[(F^{-1})_i'']^2\}\Big\}+...,$$

$$Cov(Z_{h(i)},Z_{h(j)})=\frac{p_iq_j}{n+2}\{(F^{-1})_i'(F^{-1})_j'\}+\frac{p_iq_i}{(n+2)^2}\Big\{(q_i-p_i)(F^{-1})_i''(F^{-1})_j'$$

$$+(q_j-p_j)(F^{-1})_j''(F^{-1})_i'+\frac{1}{2}p_iq_i(F^{-1})_j'(F^{-1})_i'''$$

$$+\frac{1}{2}p_jq_j(F^{-1})_i'(F^{-1})_j'''+\frac{1}{2}p_iq_j(F^{-1})_i''(F^{-1})_i''\Big\}+...,$$

where,

$p_i=i/(n+1)$;

$q_i=1-p_i$;

$(F^{-1})_i'=dF^{-1}/dy\,|_{y=r/(n+1)}$;

$(F^{-1})_i''=d^2F^{-1}/dy^2\,|_{y=r/(n+1)}$, etc.

Thus, the standard deviation of R' can be written

$$\sqrt{Var(R')} = \sigma\sqrt{Var(Z_{h(n)}) + Var(Z_{h(1)}) - 2Cov(Z_{h(n)}, Z_{h(1)})}\,.$$

Tables of the constant values $d_2 = E(R')$ and $d_3 = \sqrt{Var(R')}$ are given in Montgomery (2005), Ryan (1989), Quesenberry (1997) and others. Note that the parameters d_2 and d_3 are dependent only on the sample subgroup size (n). Furthermore, $R = X_{h(n)} - X_{h(1)}$ and $R' = Z_{h(n)} - Z_{h(1)}$ are related by $R = \sigma R'$. Consequently, the mean and standard deviation of R_i, $E(R_i)$ and $\sqrt{Var(R_i)}$, are obtained to be σd_2 and σd_3, respectively. If the standard deviation of the quality characteristic (σ) is unknown, the Shewhart and the Bonferroni-adjustment R chart, can be constructed using an unbiased estimate of σ that is given by the statistic $\bar{R}/d_2$, where $\bar{R}$ is the average range of the k preliminary samples.

In the case of the S chart, the values of the subgroup standard deviations $S_i = \sum_{j=1}^{n}(X_{ij} - \bar{X}_{i.})^2/(n-1)$, where $\bar{X}_{i.} = \sum_{i=1}^{n} X_{ij}/n$, are plotted on the chart. This chart includes the center line $E(S_i)$ and the control limits $E(S_i) \pm Z_{\alpha/2}\sqrt{Var(S_i)}$. Here, $E(S_i)$ and $\sqrt{Var(S_i)}$ are the mean and standard deviation of S_i, respectively.

Ryan (1989) introduced the Bonferroni-adjustment control limits as an alternative to the Shewhart approach. The Bonferroni-adjustment R and S control limits are given by $E(R_i) \pm Z_{\alpha/2k}\sqrt{Var(R_i)}$ and $E(S_i) \pm Z_{\alpha/2k}\sqrt{Var(S_i)}$.

Here, the value α of the Shewhart control limits is replaced by the value α/k to construct these control limits.

The new range and standard deviation control chart with known standard deviation (σ) are established similarly to the Shewhart and the Bonferroni control chart. When the standard deviation is unknown the proposed chart is estimated using a statistic with variance less than that of the Shewhart and the Bonferroni-adjustment chart. Furthermore, the constant value for the new chart with unknown standard deviation is dependent on the sample subgroup and group sizes (n,k) whereas the constant value of the Shewhart and Bonferroni chart is depended only on the sample subgroup size (n).

It should be suggested that the range and standard deviation control charts given in this chapter in both phases I and II are the same. Furthermore, often the Bonferroni R and S charts are used in phase I controlling for past data. Without lost of generality, the Bonferroni charts are also examined in phase II to recognize its performance in this phase.

In this chapter, the Shewhart and Bonferroni charts are presented in section 2, while, in section 3, the new charts are introduced. The in-control average run length (ARL_0) as well as the out-of-control average run length (ARL_1) for comparing these control charts are given in sections 4 and 5. Finally an example and some conclusions are presented in sections 6 and 7 respectively.

3-2. The Shewhart and Bonferroni R and S Charts

Suppose that the quality characteristics X_{ij} are identically independent distribution $N(\mu,\sigma^2)$, for $i=1,2,...,k$ and $j=1,2,...,n$. The statistics $R_i = X_{i(n)} - X_{i(1)}$ and $S_i = \sum_{j=1}^{n}(X_{ij}-\bar{X}_{i.})^2/(n-1)$ are used to construct the range and standard deviation charts, respectively. It is well known that $E(R_i)=\sigma d_2$, $\sqrt{Var(R_i)}=\sigma d_3$, $E(S_i)=\sigma c_4$ and $\sqrt{Var(S_i)}=\sigma\sqrt{1-c_4^2}$, where

$$c_4 = \sqrt{\frac{2}{n-1}}\Gamma\left[\frac{n}{2}\right]\Big/\Gamma\left[\frac{(n-1)}{2}\right].$$

The Shewhart R chart with a known or unknown parameter σ are based on the random variables $(R_i - E(R_i))/\sqrt{Var(R_i)}$ and $\left(R_i - \hat{E}(R_i)\right)/\sqrt{\hat{Var}(R_i)}$, respectively, and for the Shewhart S chart, $(S_i - E(S_i))/\sqrt{Var(S_i)}$ and $\left(S_i - \hat{E}(S_i)\right)/\sqrt{\hat{Var}(S_i)}$. Let $\hat{E}(R_i)=\bar{R}$, $\sqrt{\hat{Var}(R_i)}=d_3\bar{R}/d_2$, $\hat{E}(S_i)=\bar{S}$ and $\sqrt{\hat{Var}(S_i)}=\bar{S}\sqrt{1-c_4^2}/c_4$, where $\bar{R}=\sum_{i=1}^{k}R_i/k$ and $\bar{S}=\sum_{i=1}^{k}S_i/k$. These variables, for sample sizes as large as $k\geq 20$ and $n\geq 4$, follow approximately the standard normal distribution. These variables, for sample sizes as large as $k\geq 20$ and $n\geq 4$, follow approximately the standard normal distribution.

The Shewhart R and S control limits for a known parameter σ, with confidence $(1-\alpha)\%$, are given by $\sigma(d_2 \pm Z_{\alpha/2}d_3)$ and $\sigma(c_4 \pm Z_{\alpha/2}\sqrt{1-c_4^2})$, respectively. For these control limits, the center line is σd_2 and σc_4, where

the constant values d_2, d_3 and c_4 depend only on the sample subgroup size (n).

If the standard deviation of the quality characteristic is unknown, then it is estimated by the unbiased statistics $\bar{R}/d_2$ and $\bar{S}/c_4$ for Shewhart R and S chart, respectively. Then, the center line and the control limits for the Shewhart R chart with unknown parameter (σ) take the form

$$\begin{aligned} U\hat{C}L &= (\bar{R}/d_2)(d_2 + z_{\alpha/2}d_3); \\ \hat{C}L &= \bar{R}; \\ L\hat{C}L &= (\bar{R}/d_2)(d_2 - z_{\alpha/2}d_3), \end{aligned} \tag{3}$$

while, for the Shewhart S chart,

$$\begin{aligned} U\hat{C}L &= (\bar{S}/c_4)(c_4 + z_{\alpha/2}\sqrt{1-c_4^2}); \\ \hat{C}L &= \bar{S}; \\ L\hat{C}L &= (\bar{S}/c_4)(c_4 - z_{\alpha/2}\sqrt{1-c_4^2}). \end{aligned} \tag{4}$$

The Bonferroni-adjustment R and S control chart were suggested by Ryan (1989) in order to improve the probability of detecting one or more false alarms of the Shewhart chart. The Bonferroni-adjustment R and S control limits for known standard deviation parameter are given below

$$\begin{aligned} &\sigma(d_2 \pm Z_{\alpha/2k}d_3); \\ &\sigma(c_4 \pm Z_{\alpha/2k}\sqrt{1-c_4^2}). \end{aligned} \tag{5}$$

Furthermore, the Bonferroni R and S control limits with unknown parameter are

$$(\bar{R}/d_2)(d_2 \pm Z_{\alpha/2k} d_3) \tag{6}$$

$$(\bar{S}/c_4)(c_4 - z_{\alpha/2k}\sqrt{1-c_4^2}). \tag{7}$$

The center lines for the Bonferroni R chart with known and unknown standard deviation are σd_2 and $\bar{R}$, respectively, and for the Bonferroni S chart, σc_4 and $\bar{S}$.

3-3. The New R and S Chart

When the standard deviation is unknown, for constructing the new range and standard deviation charts, we need a good estimator of the standard deviation σ of the normal distribution $N \sim (\mu, \sigma^2)$. A brief presentation of some estimators of the standard deviation is given in subsection 3.1. A uniformly minimum variance unbiased (UMVU) estimator is suggested in section 3.2. The new R and S charts for both known and unknown standard deviation are presented in subsection 3.3.

3-3.1 A Brief Overview on the Estimation of the Standard Deviation

Markowitz (1968) suggested the use of the minimum mean-square-error estimator of σ given by $\hat{\sigma} = \sqrt{\sum_i^n (X_i - \bar{X})^2 / k}$, where

$$k = 2\left[\Gamma\left(\frac{n+1}{2}\right)\right]^2 / \left[\Gamma\left(\frac{n}{2}\right)\right]^2.$$

Prescott (1971a) introduced a linear estimator for the standard deviation defined as $\hat{\sigma} = a_n \bar{W}$. Here, a_n is the unbiasing factor and $\bar{W}$ is given by

$$\bar{W} = \left(\sum_{j=n-r+1}^{n} X_{i(j)} - \sum_{j=1}^{r} X_{i(j)} \right) / (3r) \qquad , \qquad i = 1,2,...,k .$$

Furthermore, $r = n/6$ is rounded up to the nearest whole number if $n/6$ is not an integer and $X_{i(1)} \le X_{i(2)} \le ... \le X_{i(n)}$ is an ordered sample of the normal distribution $N(\mu,\sigma^2)$. Prescott (1971b) proposed the use of another estimator for the standard deviation of the $N(\mu,\sigma^2)$ given by

$$\hat{\sigma} = \sum_{j=1}^{n} m_j X_{i(j)} / \sum_{j=1}^{n} m_j^2 .$$

In this case, $m_j = E\left((X_{i(j)} - \mu)/\sigma\right)$. This method does not take into consideration the covariance between the order statistics.

Vardeman (1999) considered using minimum mean-square-error estimator given, for the case of a single sample, by

$$\hat{\sigma} = \frac{R d_2}{d_2^2 + d_3^2} ; \quad \hat{\sigma} = \frac{S}{c_4} ,$$

where R and S are the range and standard deviation of the single sample, and the constants d_2, d_3 and c_4 are as introduced in previous sections. He also introduced a combination of several estimators for the case of r samples of possibly different sizes $n_1, n_2, ..., n_r$ with ranges $R_1, R_2, ..., R_r$ defined by $\hat{\sigma} = \gamma_1 R_1 + \gamma_2 R_2 + ... + \gamma_r R_r$, where

$$\gamma_i = \left(\sum_{i=1}^{r} \frac{d_2^2(n_i)}{d_3^2(n_i)} \right)^{-1} \frac{d_2(n_i)}{d_3^2(n_i)};$$

$$d_2(n_i) = E(R_i)/\sigma;$$

$$d_3(n_i) = \sqrt{Var(R_i)}/\sigma.$$

An analogous estimator was proposed for the case of r samples of possibly different sizes $n_1, n_2, ..., n_r$ with sample standard deviation estimators $S_1, S_2, ..., S_r$. The proposed estimators are $\hat{\sigma} = \gamma_1 S_1 + \gamma_2 S_2 + ... + \gamma_r S_r$ and $\hat{\sigma} = S_{pooled}/c_4(\upsilon+1)$. Here,

$$\gamma_i = \left(\sum_{i=1}^{r} \frac{c_4^2(n_i)}{c_5^2(n_i)} \right)^{-1} \frac{c_4(n_i)}{c_5^2(n_i)};$$

$$c_4(n_i) = E(S_i)/\sigma;$$

$$c_5(n_i) = \sqrt{Var(S_i)}/\sigma;$$

$$S_{pooled}^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2 + ... + (n_r-1)S_r^2}{(n_1-1)+(n_2-1)+...+(n_r-1)};$$

$$\upsilon = (n_1-1)+(n_2-1)+...+(n_r-1).$$

Some other estimators of the standard deviation have also been given by Glasser (1962), Khan (1968), Gurland (1971), Donatos (1989), Arnholt and Hebert (1995), Chen (1997), Watson (1997), Champ and Jones (2004). However, these authors did not employ UMVU estimators of the standard deviation of the normal distribution $N(\mu, \sigma^2)$.

3-3.2 An UMVU Estimator of the Standard Deviation

Let

$$S=\sqrt{\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij}-\bar{X}_{i.})^2/\big(k(n-1)\big)}\,.$$

It is well known that the random variable $k(n-1)S^2/\sigma^2$ is chi-square distributed. Then, the random variable $H=\sqrt{k(n-1)}S/\sigma$ follows the chi distribution with *k(n-1)* degrees of freedom. The probability density function and the rth raw moment of H are

$$P_H(h)=\frac{1}{2^{(k(n-1)/2)-1}\Gamma\big(k(n-1)/2\big)}e^{-h^2/2}h^{k(n-1)-1}\quad,\quad h>0 \tag{8}$$

$$E(H^r)=\frac{\sqrt{2^r}\,\Gamma\big((k(n-1)+r)/2\big)}{\Gamma\big(k(n-1)/2\big)}\,. \tag{9}$$

Moreover, the standard chi distribution (8) is in fact a standard gamma distribution with probability density function

$$P_H(h)=\frac{1}{\eta^{\zeta}\Gamma(\zeta)}e^{-(h-\gamma)/\eta}(h-\gamma)^{\zeta-1}\qquad,\qquad h>0,$$

where, the values ζ, η, and γ are $k(n-1)/2$, 2, and 0, respectively, with H replacing H^2. Assume further a constant value ψ to be the unbiasing factor of the standard deviation estimator, where

$$\psi=\left(\sqrt{\frac{2}{k(n-1)}}\,\Gamma\left(\frac{k(n-1)+1}{2}\right)\Big/\Gamma\left(\frac{k(n-1)}{2}\right)\right)\quad,\qquad kn<350 \text{ and}$$

$$\psi\approx\frac{4k(n-1)}{4k(n-1)+1}\qquad,\qquad kn\geq 350$$

The mean and variance of the statistic S are evaluated to be $\sigma\psi$ and $\sigma^2(1-\psi^2)$ respectively using equation (9). Thus, the statistic S/ψ is an unbiased estimator of the standard deviation (σ). Champ and Jones (2004) provide a good discussion about the unbiased estimators of the standard deviation (σ).

The constant value ψ depends on both the sample subgroup size (n) and the sample group size (k). The value of ψ with various sample sizes k and n is given in Table 3.10, for

$$k = 2(1)10,15(5)30,40,100,120,$$

and

$$n = 2(1)11,14,15,18,20(5)50,60(10)120.$$

The statistic S is an injective function of the complete sufficient statistic S^2 and the statistic S/ψ is an unbiased estimator of σ. Therefore according to the *Lehman-Scheffe* theorem, the statistic S/ψ is an UMVU estimator of σ (see Rohatgi (1984)). Then, the UMVU estimator S/ψ can be used for constructing the new range and standard deviation control chart with unknown standard deviation. In the sequel, we compare the range and standard deviation control charts that are based on the statistic $\bar{R}/d_2$ (for the Shewhart and Bonferroni approach) to those based on the statistic S/ψ (for the new approach). Table 3.1 shows that the variance of the statistic S/ψ, $\text{var}(S/\psi)=\sigma^2(1-\psi^2)/\psi^2$, is less than that of the statistics $\bar{R}/d_2$ and $\bar{S}/c_4$, i.e., $\text{var}(\bar{R}/d_2)=\sigma^2 d_3^2/(kd_2^2)$ and $\text{var}(\bar{S}/c_4)=\sigma^2(1-c_4^2)/(kc_4^2)$.

Table 3.1 The Variance of the statistics S/ψ, $\bar{R}/d_2$ and $\bar{S}/c_4$ with $\sigma^2=1$.

		k							
	n	2	5	10	15	20	25	60	120
$\sigma^2(1-\psi^2)/\psi^2$	2	0.27323	0.10440	0.05118	0.03392	0.02527	0.02015	0.00838	0.00413
	5	0.06427	0.02527	0.01267	0.00838	0.00620	0.00499	0.00209	0.00104
	10	0.02819	0.01117	0.00557	0.00374	0.00280	0.00228	0.00093	0.00046
	20	0.01328	0.00531	0.00277	0.00172	0.00132	0.00105	0.00044	0.00022
	25	0.01048	0.00413	0.00209	0.00139	0.00104	0.00083	0.00035	0.00017
$\sigma^2 d_3^2/(kd_2^2)$	2	0.28592	0.11437	0.05718	0.03812	0.02859	0.02287	0.00953	0.00477
	5	0.06899	0.02760	0.01380	0.00920	0.00690	0.00552	0.00230	0.00115
	10	0.03352	0.01341	0.00670	0.00447	0.00335	0.00268	0.00112	0.00056
	20	0.01905	0.00762	0.00381	0.00254	0.00190	0.00152	0.00063	0.00032
	25	0.01622	0.00649	0.00324	0.00216	0.00162	0.00130	0.00054	0.00027
$\sigma^2(1-c_4^2)/(kc_4^2)$	2	0.28537	0.11415	0.05707	0.03805	0.02854	0.02283	0.00951	0.00476
	5	0.06587	0.02635	0.01317	0.00878	0.00659	0.00527	0.00220	0.00110
	10	0.02846	0.01138	0.00569	0.00379	0.00285	0.00228	0.00095	0.00047
	20	0.01336	0.00534	0.00267	0.00178	0.00134	0.00107	0.00045	0.00022
	25	0.01056	0.00423	0.00211	0.00141	0.00106	0.00085	0.00035	0.00018

3-3.3 The New R and S Charts

The control limits for the average of a quality characteristic depend on the variability of the production process. While the process variability is outside the control limits, the control limits on the average quality characteristic will not have much meaning. Therefore, it is best that a range or standard deviation control limits is first set (see Montgomery (2005)). In this chapter the range and standard deviation control charts presented by Kiani, Panaretos and Psarakis (2008) is described in more details.

The quality characteristics X_{ij} for $i=1,2,...,k$ and $j=1,2,...,n$ are identically and independently normally distributed with mean μ and variance σ^2. The new range and standard deviation control charts with known standard deviation, like the Shewhart R and S control charts, are given by

$$UCL=\sigma(d_2+z_{\alpha/2}\,d_3);$$

$$CL=\sigma\,d_2;$$

$$LCL=\sigma(d_2-z_{\alpha/2}\,d_3)$$

$$UCL=\sigma(c_4+Z_{\alpha/2}\sqrt{1-c_4^2});$$

$$CL=\sigma c_4\ ;$$

$$LCL=\sigma(c_4-Z_{\alpha/2}\sqrt{1-c_4^2}).$$

To establish the proposed control charts with unknown standard deviation, we estimate $E(R_i)=\sigma d_2$, $\sqrt{Var(R_i)}=\sigma d_3$, $E(S_i)=\sigma c_4$ and $\sqrt{Var(S_i)}=\sigma\sqrt{1-c_4^2}$ using the UMVU estimators Sd_2/ψ, Sd_3/ψ, Sc_4/ψ and $S\sqrt{1-c_4^2}/\psi$, respectively. The resulting control limits for the proposed R chart with unknown standard deviation would be

$$\begin{aligned} U\hat{C}L &= (S/\psi)(d_2+z_{\alpha/2}d_3); \\ \hat{C}L &= (S/\psi)(d_2); \\ L\hat{C}L &= (S/\psi)(d_2-z_{\alpha/2}d_3), \end{aligned} \tag{10}$$

and the control limits for the proposed S chart are

$$\begin{aligned} U\hat{C}L &= (S/\psi)(c_4+z_{\alpha/2}\sqrt{1-c_4^2}); \\ \hat{C}L &= Sc_4/\psi; \\ L\hat{C}L &= (S/\psi)(c_4-z_{\alpha/2}\sqrt{1-c_4^2}). \end{aligned} \tag{11}$$

In sections 4 and 5, the in-control and out-of-control average run length for the Shewhart, Bonferroni and new R and S control charts are examined.

3-4. In-Control Average Run Length

The average run length is considered for future subgroups, in phase II. The in-control average run length (called ARL_0) is *the average number of subgroup ranges or standard deviations* that should be plotted before a subgroup range or standard deviation indicates an out-of-control condition. The ARL_0 can be calculated from $ARL_0=1/p$ under the condition that the process observations

are uncorrelated. Here, p is the probability that any point exceeds the control limits. The in-control average run length can be used to evaluate the performance of the control chart.

In this section, the average run length is considered for *the initial group and groups* 2,3,... with known and unknown parameter σ when the process is in control. Let the individual events G_i denote the subgroup range R_i or standard deviation S_i exceeds the control limits of the in control process ($R = R_0$ or $\sigma = \sigma_0$).

For the initial group of observations with unknown parameter σ, the events G_i and $G_{i'}$ for $i \neq i' = 1,2,...,k$ are not independent, since the statistics $R_i - U\hat{C}L$ and $R_j - U\hat{C}L$, or $S_i - U\hat{C}L$ and $S_j - U\hat{C}L$, for the Shewhart, the Bonferroni and the new charts are based on the same observations of the initial group.

In case the events G_i are independent, the sequence of trials, comparing R_i with $U\hat{C}L$ or S_i with $U\hat{C}L$ will be a sequence of Bernoulli trials and the run length between occurrences of G_i will be a Geometric random variable with probability $\alpha = P(G_i)$. Additionally, the in-control average run length would be $1/P(G_i)$ or $1/\alpha$ such that,

$$P(G_i) = P(R_i \leq L\hat{C}L \text{ or } R_i \geq U\hat{C}L \mid R = R_0) \tag{12}$$

or

$$P(G_i) = P(S_i \leq L\hat{C}L \text{ or } S_i \geq U\hat{C}L \mid \sigma = \sigma_0). \tag{13}$$

However, the statistics $R_i - \hat{UCL}$ and $R_j - \hat{UCL}$ or $S_i - \hat{UCL}$ and $S_j - \hat{UCL}$ for *the initial group with unknown parameter* are not independent events. Therefore, the in-control *ARL* for the initial group with unknown parameter can be not calculated.

For the initial group with known parameter, the correlation between random variables $R_i - UCL$ and $R_j - UCL$ or $S_i - UCL$ and $S_j - UCL$ can be obtained to be 0. Here, the *UCL* with known parameter is a constant value and the subgroup ranges R_i and R_j or standard deviations S_i and S_j are independent. Thus, the events G_i and G_j for the initial group with known parameter are uncorrelated.

For the groups 2,3,... *with known and unknown parameter*, the events G_i and G_j are uncorrelated, since the *UCL* with known and unknown parameter are based on the observations of initial group, while the R_i or S_i belong to groups 2,3,.... Thus, the correlation between random variables $R_i - UCL$ and $R_j - UCL$ or $S_i - UCL$ and $S_j - UCL$ can be obtained to be 0.

Based on the above, the sequence of the events $\{G_i\}$, for the initial group with known parameter and the groups 2,3,... with known and unknown parameter, would be Bernoulli trials and the run length between occurrences of G_i would be a Geometric random variable with probability $P(G_i)$. The probability $P(G_i)$ for both the Shewhart and the new approach is α, and for the Bonferroni-adjustment approach is α/k. As a result, the in-control

average run length (ARL_0) would be $1/P(G_i)=1/\alpha$ for the Shewhart and the new approach, and $1/P(G_i)=k/\alpha$ for the Bonferroni-adjustment approach. Thus the ARL_0 for the Shewhart and proposed chart ($1/\alpha$) is less than the ARL_0 for Bonferroni-adjustment chart (k/α, for $k \geq 2$). (See, also, Nedumaran and Pignatiello (2005) and Tsai et al. (2005)).

Now, we discuss the in-control average run length that based on *the average number of groups* before *a group* indicates an out-of-control condition. Here, the in-control average run length is called ARL_0'.

For the initial group with known parameter and the groups 2,3,... *with known and unknown parameter,* let the random variable Y denote the overall occurrences of events G_i for $i=1,2,...,k$. Then, this random variable should follow the Binomial distribution with probability distribution (for the Shewhart and the new approach) given by

$$P(Y=y)=\binom{k}{y}\alpha^y(1-\alpha)^{k-y} \quad , \quad y=0,1,2,...,k .$$

Therefore, the probability of one or more subgroup ranges or standard deviation falling out of the control limits (the probability of out-of-control condition for a group) for the Shewhart and the new approach is $P(Y \geq 1)=1-(1-\alpha)^k$. Ryan (1989) showed that the probability $1-(1-\alpha)^k$ is approximately equal with $k\alpha$. Thus, Ryan (1989) suggested the Bonferroni-adjustment approach for the control limits. In this case, the probability of one or more false alarm is improved to $1-(1-\alpha/k)^k$ that is less than $1-(1-\alpha)^k$. The probability distribution of Y, for the Bonferroni-adjustment approach,

follows the Binomial distribution with parameters $(k,\alpha/k)$. Therefore, the probability of one or more subgroup ranges or standard deviations falling out of the control limits for the Bonferroni-adjustment approach is given by $P(Y \geq 1) = 1-(1-\alpha/k)^k$. As a result, the ARL_0' is obtained to be $1/(1-(1-\alpha)^k)$ for the Shewhart and the new approach, and $1/(1-(1-\alpha/k)^k)$ for the Bonferroni-adjustment approach. That means $1/(1-(1-\alpha)^k) < 1/(1-(1-\alpha/k)^k$. Thus, the ARL_0' for the Shewhart and the new approach is less than that of the Bonferroni-adjustment approach. Consequently, the in-control average run length (ARL_0 , ARL_0') for the Bonferroni-adjustment approach is greater than the Shewhart and the new approach. In the next section we illustrate that the out-of-control average run length for the Bonferroni-adjustment approach is not so satisfactory. In other words, the power of the Bonferroni-adjustment control limits is considerably less than the one of the Shewhart and new approach.

3-5. Out-of-Control Average Run Length

The ability of the range and standard deviation control charts to detect shifts in process quality is described by the out-of-control average run length (ARL_1), in phase II controlling. The probabilities of detecting a one or more false alarms when the process is in control were improved by using the Bonferroni-adjustment approach. In this section, the power of control limits for the usual approach (Shewhart and Bonferroni) and the new approach are compared using ARL_1.

If the in-control value of the standard deviation shifts from σ_0 to $\sigma_1 = \lambda\sigma_0 > \sigma_0$, $(\lambda > 1)$ the probability of not detecting the range or standard deviation shift (β) is calculated by

$$\beta = P\left(LCL \le R_i \le UCL \middle| \sigma = \lambda\sigma_0\right) \qquad \text{(R Chart)} \tag{14}$$

or

$$\beta = P\left(LCL \le S_i \le UCL \middle| \sigma = \lambda\sigma_0\right) \qquad \text{(S Chart).} \tag{15}$$

These probabilities for both the Shewhart and proposed R and S charts with known parameter (σ) similarly are obtained to be

$$\beta = P\left(\frac{-Z_{\alpha/2}d_3 + d_2(1-\lambda)}{\lambda d_3} \le Z_i \le \frac{Z_{\alpha/2}d_3 + d_2(1-\lambda)}{\lambda d_3}\right) \qquad \text{(R Chart)} \tag{16}$$

and

$$\beta = P\left(\frac{-Z_{\alpha/2}\sqrt{1-c_4^2} + c_4(1-\lambda)}{\lambda\sqrt{1-c_4^2}} \le Z_i \le \frac{Z_{\alpha/2}\sqrt{1-c_4^2} + c_4(1-\lambda)}{\lambda\sqrt{1-c_4^2}}\right) \qquad \text{(S Chart)} \tag{17}$$

Meanwhile, the probability β for the Bonferroni-adjustment R and S chart with known parameter (σ) is

$$\beta = P\left(\frac{-Z_{\alpha/2k}d_3 + d_2(1-\lambda)}{\lambda d_3} \le Z_i \le \frac{Z_{\alpha/2k}d_3 + d_2(1-\lambda)}{\lambda d_3}\right) \qquad \text{(R Chart)} \tag{18}$$

and

$$\beta = P\left(\frac{-Z_{\alpha/2k}\sqrt{1-c_4^2} + c_4(1-\lambda)}{\lambda\sqrt{1-c_4^2}} \le Z_i \le \frac{Z_{\alpha/2k}\sqrt{1-c_4^2} + c_4(1-\lambda)}{\lambda\sqrt{1-c_4^2}}\right) \qquad \text{(S Chart).} \tag{19}$$

Usually, the parameter σ is unknown. In this case we obtain the probability β for the control limits with unknown parameter. We already showed that S/ψ is a UMVU estimator of the standard deviation. Let us call S/ψ by S_t. In order to compute the type II error (β) for the Shewhart, the Bonferroni

and the proposed approach with unknown parameter the standard deviation (σ_0) from equations (14) and (15) is estimated by S_t. Thus, the probability β for the Shewhart approach with unknown parameter is calculated using equations (3) and (14) for R chart and equations (4) and (15) for S chart

$$\beta = P\left(\frac{(\bar{R}/d_2)(d_2 - Z_{\alpha/2}d_3) - \lambda S_t d_2}{\lambda S_t d_3} \le Z_i \le \frac{(\bar{R}/d_2)(d_2 + Z_{\alpha/2}d_3) - \lambda S_t d_2}{\lambda S_t d_3}\right) \quad (20)$$

and

$$\beta = P\left(\frac{(\bar{S}/c_4)(c_4 - Z_{\alpha/2}\sqrt{1-c_4^2}) - \lambda S_t c_4}{\lambda S_t \sqrt{1-c_4^2}} \le Z_i \le \frac{(\bar{S}/c_4)(c_4 + Z_{\alpha/2}\sqrt{1-c_4^2}) - \lambda S_t c_4}{\lambda S_t \sqrt{1-c_4^2}}\right) \quad (21)$$

Similarly, the probability β for the Bonferroni-adjustment approach with unknown parameter is obtained by using equations (6) and (14) for R chart and equations (7) and (15) for S chart,

$$\beta = P\left(\frac{(\bar{R}/d_2)(d_2 - Z_{\alpha/2k}d_3) - \lambda S_t d_2}{\lambda S_t d_3} \le Z_i \le \frac{(\bar{R}/d_2)(d_2 + Z_{\alpha/2k}d_3/d_2) - \lambda S_t d_2}{\lambda S_t d_3}\right) \quad (22)$$

and

$$\beta = P\left(\frac{(\bar{S}/c_4)(c_4 - Z_{\alpha/2k}\sqrt{1-c_4^2}) - \lambda S_t c_4}{\lambda S_t \sqrt{1-c_4^2}} \le Z_i \le \frac{(\bar{S}/c_4)(c_4 + Z_{\alpha/2k}\sqrt{1-c_4^2}) - \lambda S_t c_4}{\lambda S_t \sqrt{1-c_4^2}}\right) \quad (23)$$

Also, the probability β for the new approach with unknown parameter is obtained using equations (10) and (14) for R chart and equations (11) and (15) for S chart as follow

$$\beta = P\left(\frac{(S/\psi)(d_2 - Z_{\alpha/2}d_3) - \lambda S_t d_2}{\lambda S_t d_3} \le Z_i \le \frac{(S/\psi)(d_2 + Z_{\alpha/2}d_3) - \lambda S_t d_2}{\lambda S_t d_3}\right) \quad (24)$$

and

$$\beta = P\left(\frac{(S/\psi)(c_4 - Z_{\alpha/2}\sqrt{1-c_4^2}) - \lambda S_t c_4}{\lambda S_t \sqrt{1-c_4^2}} \le Z_i \le \frac{(S/\psi)(c_4 + Z_{\alpha/2}\sqrt{1-c_4^2}) - \lambda S_t c_4}{\lambda S_t \sqrt{1-c_4^2}} \right) \quad (25)$$

The probability β with known standard deviation for various sample sizes n and coefficient λ is exhibited by the operating-characteristic (OC) curves. The OC curves are constructed according to the constant values of Table 3.2 and equations (16) and (17) for the Shewhart and the new R and S charts, respectively, and equations (18) and (19) for the Bonferroni-adjustment R and S chart.

Table 3.2 The constant values d_2, d_3 and c_4 to construct the OC curves

	n				
	2	5	10	20	25
d_2	1.128	2.326	3.078	3.735	3.931
d_3	0.853	0.864	0.797	0.729	0.708
c_4	0.7979	0.9400	0.9727	0.9869	0.9896

Figure 1 Operating characteristic curves for the Shewhart and new R chart (k=20 , a=0.01).

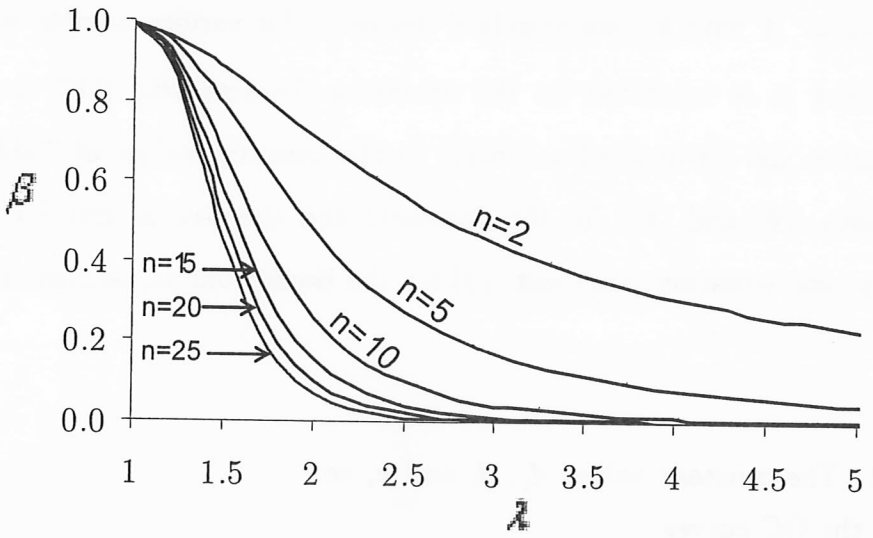


Figure 2 Operating characteristic curves for the Bonferroni R chart (k=20 , a=0.01).

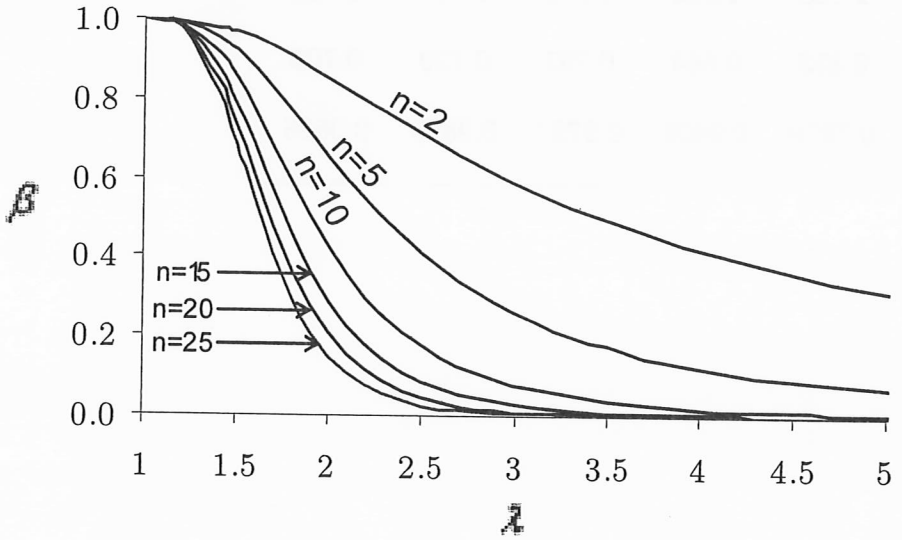


Figure 3 Operating characteristic curves for the Shewhart and new S chart (k=20 , a=0.01).

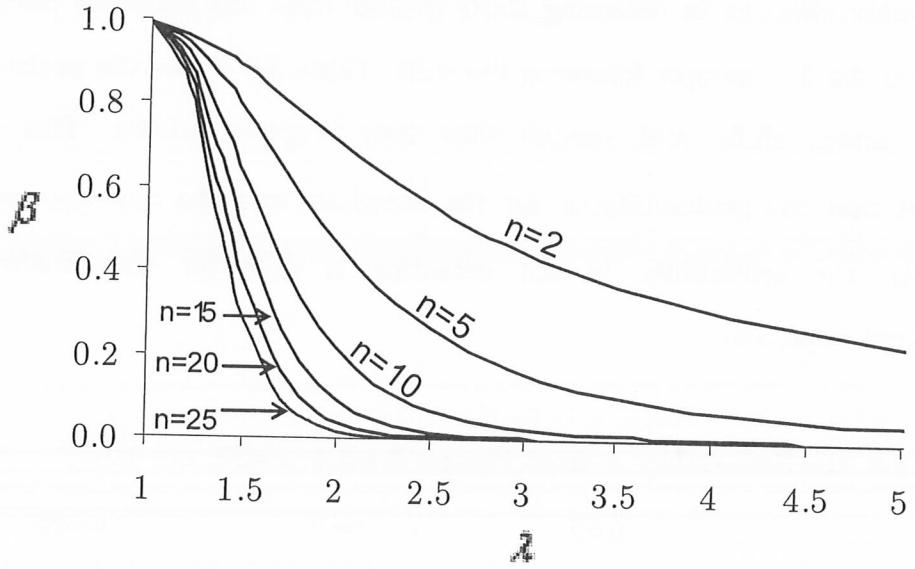
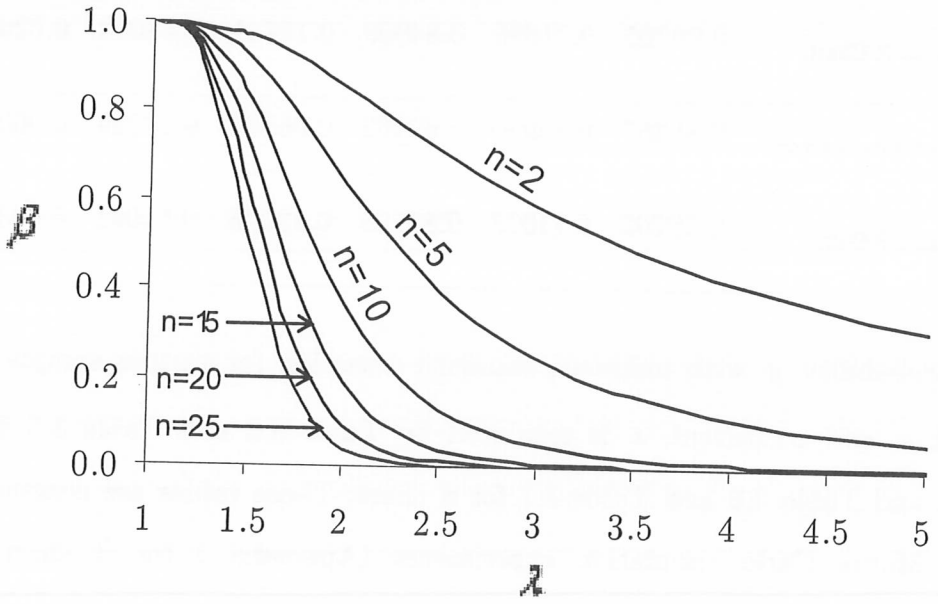


Figure 4 Operating characteristic curves for the Bonferroni S chart (k=20 , a=0.01).



Figures 3.1-3.4 indicate that the Shewhart and the new approach are considerably effective in detecting shifts greater than one standard deviation ($\lambda > 1$) on the first sample following the shift. Table 3.3 shows the probability β for various shifts and sample sizes from Figures 3.1-3.4. This table indicates that the probability β for the Shewhart and the new approach is less than the probability in not detecting a shift for the Bonferroni-adjustment approach.

Table 3.3 The Probability β from Figures 3.1-3.4 (k=20, $\alpha = 0.01$)

Approach	n=2		n=10		n=25	
	$\lambda = 1.5$	$\lambda = 2.5$	$\lambda = 1.5$	$\lambda = 2.5$	$\lambda = 1.5$	$\lambda = 2.5$
Shewhart-New R Chart	0.88363	0.55954	0.66503	0.09866	0.44671	0.01069
Bonferroni R Chart	0.96705	0.71096	0.84909	0.17741	0.68073	0.02625
Shewhart-New S Chart	0.88357	0.55930	0.62462	0.06863	0.28229	0.00098
Bonferroni S Chart	0.96702	0.71073	0.82199	0.13076	0.51091	0.00312

The probability β with unknown standard deviation for various sample sizes n and k, and coefficient λ is exhibited by Table 3.4 and Table 3.5 for R chart, and Table 3.6 and Table 3.7 for S chart. These tables are constructed using Monte Carlo simulation experiments (Appendix I for R chart and Appendix II for S chart) with constant values of Table 3.2 and equations (20),

(22), and (24) for the Shewhart, the Bonferroni-adjustment, and the new R chart, respectively, and equations (21), (23), and (25) for the Shewhart, the Bonferroni-adjustment, and the new S chart, respectively. If the subgroup range or standard deviation falls out of the control limits, a counter for that was increased by one. This procedure was replicated 10,000 times, and then the probability β was estimated by dividing the number of replications in which points exceeded the control limits with the total number of replications. As a result, the estimated probability of not detecting a range or a standard deviation shift for the Shewhart, the Bonferroni-adjustment and the new R and S chart are shown in the following tables 3.4-3.7.

Table 3.4 Estimated probability of not detecting a range shift (β) for $\alpha = 0.1$.

		n=2				n=5			
k	Approach/λ	1.5	2	2.5	3	1.5	2	2.5	3
10	Shewhart	0.763	0.673	0.517	0.424	0.596	0.315	0.170	0.080
	Bonferroni	0.898	0.767	0.650	0.551	0.804	0.504	0.288	0.163
	New	0.747	0.633	0.509	0.421	0.594	0.314	0.171	0.080
15	Shewhart	0.768	0.615	0.523	0.444	0.589	0.325	0.166	0.088
	Bonferroni	0.906	0.774	0.675	0.561	0.843	0.527	0.315	0.181
	New	0.760	0.609	0.524	0.441	0.579	0.322	0.170	0.084
20	Shewhart	0.767	0.639	0.526	0.445	0.592	0.309	0.171	0.088
	Bonferroni	0.910	0.779	0.582	0.593	0.839	0.546	0.328	0.203
	New	0.766	0.635	0.524	0.438	0.586	0.306	0.170	0.088
30	Shewhart	0.766	0.622	0.521	0.440	0.608	0.316	0.182	0.089
	Bonferroni	0.914	0.795	0.699	0.611	0.857	0.565	0.353	0.204
	New	0.762	0.619	0.516	0.439	0.608	0.315	0.180	0.089
40	Shewhart	0.765	0.629	0.523	0.446	0.601	0.523	0.171	0.020
	Bonferroni	0.923	0.807	0.705	0.613	0.875	0.593	0.367	0.228
	New	0.758	0.627	0.519	0.446	0.594	0.323	0.171	0.010
100	Shewhart	0.764	0.628	0.527	0.441	0.601	0.320	0.173	0.095
	Bonferroni	0.937	0.832	0.730	0.645	0.892	0.639	0.414	0.257
	New	0.763	0.627	0.526	0.438	0.602	0.317	0.174	0.095
120	Shewhart	0.769	0.628	0.526	0.451	0.596	0.321	0.175	0.100
	Bonferroni	0.940	0.836	0.741	0.649	0.901	0.653	0.419	0.263
	New	0.767	0.627	0.524	0.451	0.598	0.321	0.175	0.100

Table 3.4 (Continued) Estimated probability of not detecting a range shift (β) for $\alpha = 0.1$.

		n=10				n=25			
k	Approach/λ	1.5	2	2.5	3	1.5	2	2.5	3
10	Shewhart	0.430	0.117	0.024	0.008	0.224	0.012	0.000	0.000
	Bonferroni	0.687	0.249	0.078	0.015	0.464	0.048	0.001	0.000
	New	0.412	0.122	0.023	0.009	0.226	0.012	0.000	0.000
15	Shewhart	0.435	0.123	0.026	0.009	0.223	0.009	0.000	0.000
	Bonferroni	0.709	0.292	0.099	0.023	0.514	0.055	0.006	0.000
	New	0.434	0.125	0.026	0.009	0.226	0.010	0.000	0.000
20	Shewhart	0.447	0.128	0.024	0.006	0.218	0.013	0.001	0.000
	Bonferroni	0.742	0.317	0.097	0.036	0.546	0.066	0.004	0.000
	New	0.450	0.128	0.024	0.006	0.211	0.013	0.001	0.000
30	Shewhart	0.447	0.124	0.034	0.013	0.234	0.016	0.001	0.000
	Bonferroni	0.751	0.337	0.124	0.041	0.574	0.079	0.005	0.000
	New	0.442	0.124	0.033	0.013	0.233	0.016	0.001	0.000
40	Shewhart	0.445	0.131	0.034	0.006	0.224	0.012	0.001	0.000
	Bonferroni	0.778	0.349	0.123	0.043	0.582	0.083	0.008	0.000
	New	0.440	0.132	0.034	0.006	0.224	0.011	0.001	0.000
100	Shewhart	0.449	0.123	0.033	0.009	0.233	0.015	0.000	0.000
	Bonferroni	0.823	0.412	0.157	0.056	0.661	0.125	0.009	0.001
	New	0.451	0.123	0.032	0.009	0.234	0.014	0.000	0.000
120	Shewhart	0.447	0.130	0.033	0.011	0.237	0.013	0.000	0.000
	Bonferroni	0.830	0.420	0.163	0.059	0.666	0.126	0.011	0.001
	New	0.445	0.129	0.033	0.010	0.235	0.013	0.000	0.000

Table 3.5 Estimated probability of not detecting a range shift (β) for $\alpha = 0.01$.

		n=2				n=5			
k	Approach/ λ	1.5	2	2.5	3	1.5	2	2.5	3
	Shewhart	0.903	0.772	0.653	0.566	0.800	0.478	0.280	0.172
10	Bonferroni	0.942	0.862	0.742	0.636	0.920	0.625	0.387	0.240
	New	0.901	0.745	0.649	0.552	0.798	0.474	0.277	0.175
	Shewhart	0.898	0.742	0.658	0.572	0.820	0.490	0.289	0.166
15	Bonferroni	0.957	0.833	0.737	0.647	0.908	0.631	0.405	0.241
	New	0.896	0.730	0.656	0.570	0.808	0.480	0.280	0.168
	Shewhart	0.893	0.769	0.637	0.579	0.819	0.500	0.294	0.171
20	Bonferroni	0.941	0.837	0.739	0.650	0.907	0.649	0.410	0.257
	New	0.888	0.766	0.628	0.575	0.809	0.497	0.291	0.169
	Shewhart	0.895	0.767	0.643	0.558	0.809	0.491	0.312	0.170
30	Bonferroni	0.949	0.839	0.727	0.637	0.898	0.645	0.420	0.253
	New	0.893	0.762	0.638	0.555	0.801	0.487	0.312	0.169
	Shewhart	0.882	0.768	0.655	0.556	0.806	0.493	0.314	0.175
40	Bonferroni	0.936	0.843	0.733	0.646	0.899	0.637	0.406	0.259
	New	0.880	0.766	0.652	0.556	0.801	0.497	0.315	0.174
	Shewhart	0.885	0.762	0.649	0.568	0.802	0.514	0.311	0.182
100	Bonferroni	0.937	0.836	0.732	0.647	0.897	0.639	0.417	0.264
	New	0.884	0.761	0.650	0.566	0.801	0.512	0.313	0.181
	Shewhart	0.882	0.757	0.656	0.564	0.804	0.509	0.291	0.180
120	Bonferroni	0.944	0.832	0.738	0.646	0.899	0.635	0.407	0.259
	New	0.881	0.756	0.654	0.562	0.801	0.509	0.291	0.178

Table 3.5 (Continued) Estimated probability of not detecting a range shift (β) for $\alpha = 0.01$.

		n=10				n=25			
k	Approach/λ	1.5	2	2.5	3	1.5	2	2.5	3
	Shewhart	0.694	0.240	0.061	0.015	0.463	0.057	0.007	0.000
10	Bonferroni	0.842	0.396	0.123	0.037	0.653	0.107	0.013	0.000
	New	0.690	0.230	0.050	0.015	0.460	0.048	0.004	0.000
	Shewhart	0.701	0.271	0.081	0.020	0.471	0.076	0.002	0.000
15	Bonferroni	0.845	0.402	0.164	0.038	0.640	0.135	0.005	0.002
	New	0.696	0.264	0.079	0.019	0.467	0.066	0.002	0.000
	Shewhart	0.691	0.280	0.058	0.021	0.458	0.051	0.003	0.000
20	Bonferroni	0.817	0.422	0.142	0.046	0.666	0.110	0.005	0.000
	New	0.689	0.279	0.060	0.020	0.459	0.052	0.003	0.000
	Shewhart	0.690	0.266	0.087	0.038	0.465	0.044	0.003	0.000
30	Bonferroni	0.828	0.407	0.158	0.061	0.672	0.102	0.007	0.000
	New	0.682	0.264	0.085	0.033	0.461	0.044	0.003	0.000
	Shewhart	0.693	0.260	0.086	0.012	0.480	0.057	0.003	0.000
40	Bonferroni	0.836	0.415	0.157	0.051	0.647	0.127	0.009	0.000
	New	0.694	0.255	0.086	0.012	0.480	0.054	0.003	0.000
	Shewhart	0.691	0.279	0.087	0.028	0.476	0.054	0.003	0.000
100	Bonferroni	0.819	0.418	0.163	0.054	0.655	0.121	0.008	0.001
	New	0.692	0.281	0.088	0.028	0.469	0.053	0.002	0.000
	Shewhart	0.684	0.270	0.085	0.030	0.474	0.059	0.004	0.000
120	Bonferroni	0.824	0.414	0.152	0.059	0.657	0.117	0.011	0.001
	New	0.682	0.267	0.084	0.029	0.474	0.058	0.004	0.000

Table 3.6 Estimated probability of not detecting a standard deviation shift (β) for $\alpha = 0.1$.

		n=2				n=5			
k	Approach/λ	1.5	2	2.5	3	1.5	2	2.5	3
10	Shewhart	0.757	0.618	0.529	0.431	0.599	0.319	0.173	0.081
	Bonferroni	0.881	0.772	0.648	0.560	0.809	0.505	0.289	0.162
	New	0.745	0.598	0.523	0.425	0.597	0.317	0.172	0.080
15	Shewhart	0.749	0.627	0.532	0.447	0.591	0.328	0.175	0.086
	Bonferroni	0.904	0.773	0.669	0.567	0.844	0.529	0.319	0.180
	New	0.741	0.612	0.530	0.440	0.588	0.325	0.173	0.085
20	Shewhart	0.772	0.616	0.529	0.442	0.594	0.329	0.179	0.087
	Bonferroni	0.909	0.780	0.591	0.594	0.848	0.540	0.328	0.207
	New	0.768	0.602	0.527	0.437	0.589	0.328	0.176	0.085
30	Shewhart	0.763	0.623	0.530	0.444	0.606	0.330	0.181	0.080
	Bonferroni	0.916	0.804	0.703	0.617	0.861	0.569	0.352	0.208
	New	0.759	0.615	0.529	0.440	0.607	0.330	0.181	0.081
40	Shewhart	0.767	0.634	0.538	0.442	0.608	0.531	0.173	0.101
	Bonferroni	0.910	0.809	0.709	0.615	0.881	0.590	0.365	0.227
	New	0.761	0.630	0.521	0.441	0.599	0.329	0.172	0.098
100	Shewhart	0.766	0.631	0.531	0.445	0.609	0.329	0.172	0.098
	Bonferroni	0.936	0.841	0.741	0.648	0.899	0.641	0.418	0.253
	New	0.763	0.625	0.529	0.442	0.609	0.329	0.173	0.097
120	Shewhart	0.772	0.634	0.530	0.453	0.612	0.335	0.177	0.101
	Bonferroni	0.939	0.839	0.750	0.652	0.918	0.657	0.418	0.264
	New	0.769	0.632	0.530	0.453	0.611	0.333	0.177	0.100

Table 3.6 (Continued) Estimated probability of not detecting a standard deviation shift (β) for $\alpha = 0.1$.

		n=10				n=25			
k	Approach/ λ	1.5	2	2.5	3	1.5	2	2.5	3
	Shewhart	0.431	0.116	0.025	0.008	0.226	0.011	0.000	0.000
10	Bonferroni	0.686	0.246	0.078	0.016	0.467	0.049	0.001	0.000
	New	0.419	0.116	0.023	0.009	0.222	0.010	0.000	0.000
	Shewhart	0.437	0.120	0.029	0.009	0.223	0.010	0.000	0.000
15	Bonferroni	0.711	0.291	0.099	0.025	0.514	0.056	0.006	0.000
	New	0.437	0.118	0.027	0.009	0.226	0.010	0.000	0.000
	Shewhart	0.446	0.122	0.025	0.007	0.226	0.012	0.001	0.000
20	Bonferroni	0.743	0.314	0.096	0.035	0.467	0.065	0.004	0.000
	New	0.445	0.120	0.023	0.007	0.222	0.011	0.001	0.000
	Shewhart	0.448	0.126	0.035	0.012	0.234	0.015	0.001	0.000
30	Bonferroni	0.754	0.339	0.121	0.043	0.573	0.080	0.005	0.000
	New	0.444	0.123	0.032	0.012	0.234	0.014	0.001	0.000
	Shewhart	0.442	0.132	0.035	0.007	0.226	0.011	0.001	0.000
40	Bonferroni	0.777	0.342	0.124	0.045	0.467	0.085	0.007	0.000
	New	0.441	0.132	0.034	0.006	0.222	0.012	0.001	0.000
	Shewhart	0.447	0.127	0.031	0.009	0.233	0.016	0.000	0.000
100	Bonferroni	0.823	0.419	0.159	0.058	0.661	0.127	0.009	0.001
	New	0.456	0.124	0.031	0.008	0.234	0.017	0.000	0.000
	Shewhart	0.448	0.131	0.034	0.012	0.225	0.012	0.000	0.000
120	Bonferroni	0.834	0.423	0.165	0.058	0.462	0.129	0.010	0.001
	New	0.447	0.129	0.032	0.011	0.221	0.013	0.000	0.000

Table 3.7 Estimated probability of not detecting a standard deviation shift (β) for $\alpha = 0.01$.

		n=2				n=5			
k	Approach/ λ	1.5	2	2.5	3	1.5	2	2.5	3
	Shewhart	0.904	0.760	0.656	0.550	0.854	0.526	0.314	0.182
10	Bonferroni	0.956	0.840	0.748	0.636	0.942	0.684	0.448	0.258
	New	0.896	0.752	0.652	0.548	0.848	0.516	0.298	0.178
	Shewhart	0.908	0.761	0.654	0.553	0.852	0.528	0.318	0.185
15	Bonferroni	0.959	0.842	0.743	0.638	0.942	0.685	0.449	0.259
	New	0.896	0.754	0.651	0.550	0.847	0.519	0.299	0.180
	Shewhart	0.889	0.763	0.657	0.551	0.855	0.527	0.314	0.183
20	Bonferroni	0.959	0.841	0.745	0.638	0.946	0.687	0.447	0.259
	New	0.886	0.755	0.655	0.549	0.851	0.519	0.299	0.181
	Shewhart	0.895	0.762	0.655	0.556	0.853	0.525	0.319	0.184
30	Bonferroni	0.957	0.845	0.743	0.639	0.941	0.682	0.451	0.261
	New	0.891	0.751	0.650	0.552	0.846	0.516	0.305	0.179
	Shewhart	0.889	0.766	0.658	0.553	0.857	0.528	0.319	0.186
40	Bonferroni	0.955	0.841	0.750	0.632	0.944	0.687	0.450	0.263
	New	0.887	0.756	0.656	0.549	0.852	0.520	0.303	0.183
	Shewhart	0.890	0.762	0.652	0.556	0.855	0.529	0.313	0.181
100	Bonferroni	0.961	0.842	0.744	0.641	0.946	0.689	0.448	0.255
	New	0.888	0.758	0.650	0.551	0.851	0.521	0.297	0.182
	Shewhart	0.895	0.761	0.654	0.552	0.852	0.524	0.317	0.181
120	Bonferroni	0.957	0.843	0.747	0.639	0.940	0.683	0.449	0.257
	New	0.892	0.759	0.654	0.551	0.848	0.515	0.307	0.180

Table 3.7 (Continued) Estimated probability of not detecting a standard deviation shift (β) for $\alpha = 0.01$.

		n=10				n=25			
k	Approach/ λ	1.5	2	2.5	3	1.5	2	2.5	3
	Shewhart	0.634	0.214	0.038	0.020	0.264	0.010	0.000	0.000
10	Bonferroni	0.816	0.328	0.084	0.042	0.478	0.024	0.002	0.000
	New	0.628	0.214	0.037	0.020	0.258	0.011	0.000	0.000
	Shewhart	0.638	0.219	0.040	0.023	0.268	0.013	0.000	0.000
15	Bonferroni	0.819	0.332	0.088	0.044	0.483	0.026	0.003	0.000
	New	0.632	0.218	0.039	0.022	0.259	0.012	0.000	0.000
	Shewhart	0.635	0.215	0.039	0.021	0.262	0.015	0.000	0.000
20	Bonferroni	0.818	0.329	0.084	0.042	0.477	0.029	0.002	0.000
	New	0.629	0.214	0.039	0.021	0.254	0.016	0.000	0.000
	Shewhart	0.632	0.218	0.038	0.027	0.261	0.010	0.000	0.000
30	Bonferroni	0.815	0.332	0.085	0.049	0.477	0.026	0.002	0.000
	New	0.627	0.215	0.039	0.025	0.258	0.011	0.000	0.000
	Shewhart	0.637	0.212	0.043	0.022	0.266	0.015	0.000	0.000
40	Bonferroni	0.819	0.327	0.088	0.043	0.481	0.028	0.002	0.000
	New	0.634	0.214	0.042	0.023	0.265	0.013	0.000	0.000
	Shewhart	0.634	0.216	0.035	0.028	0.269	0.018	0.000	0.000
100	Bonferroni	0.819	0.329	0.080	0.043	0.480	0.030	0.003	0.000
	New	0.635	0.215	0.036	0.027	0.269	0.016	0.000	0.000
	Shewhart	0.639	0.214	0.045	0.022	0.267	0.012	0.000	0.000
120	Bonferroni	0.821	0.327	0.089	0.044	0.479	0.025	0.004	0.000
	New	0.639	0.214	0.043	0.023	0.267	0.012	0.000	0.000

The results in tables 3.4-3.7 show that with unknown standard deviation the new R and S chart perform better than both the Shewhart and the Bonferroni-adjustment R and S chart.

To construct the ARL_1, let the events W_i denote that the subgroup range or the standard deviation fall out of control limits, while the standard deviation of the quality characteristic is shifted from σ_0 to $\lambda\sigma_0$. The events W_i for the initial group with known parameter and the groups 2,3,... with known and unknown parameter are uncorrelated. Therefore, the sequence of events $\{W_i\}$ would be a sequence of Bernoulli events and the run length between occurrences of W_i would be a Geometric random variable with parameter $P(W_i) = (1-\beta)$. Consequently, the out-of-control average run length would be

$$ARL_1 = 1/(1-\beta). \tag{26}$$

Here, the probability β for known and unknown parameter σ is replaced by equations (16) and (20) for the Shewhart R chart and equations (17) and (21) for the Shewhart S chart, equations (18) and (22) for the Bonferroni-adjustment R chart and equations (19) and (23) for the Bonferroni S chart, equations (16) and (24) for the new R chart and equations (17) and (25) for the new R chart. Using the OC curves 3.1-3.4, and Table 3.3, for known parameter σ, the ARL_1 for the Shewhart and new approach is demonstrated to be less than the Bonferroni-adjustment approach for both R and S chart. For instance, using the probability β of Table 3.3 and the ARL_1 of equation (26), for $k = 20$, $\alpha = 0.01$, $n = 2$, and $\lambda = 1.5$, we have that the ARL_1 is equal 8.593 and 8.589 for the Shewhart and the proposed R and S chart, respectively, and equal 30.349 and 30.321 for the Bonferroni-adjustment R and S chart, respectively. Additionally, for unknown parameter σ, it can be

demonstrated, through Monte Carlo simulation experiments, that the ARL_1 for the new chart is slightly less than the Shewhart chart and considerably less than the Bonferroni-adjustment chart for both R and S chart. For instance, using the probability β of Table 3.5 and Table 3.7 and the ARL_1 of equation (26), for $k=20$, $\alpha=0.01$, $n=2$, and $\lambda=1.5$, we have that the ARL_1 is equal 9.346, 16.949, and 8.928 for the Shewhart, the Bonferroni, and the proposed R chart, respectively, and equal 9.009, 24.390, and 8.772 for the Shewhart, the Bonferroni, and the proposed S chart, respectively. It should be noted that another method for approximating the ARL can be found in Ng and Cases (1992).

3-6. Example

Twenty samples $(k=20)$ each of size four $(n=4)$ of piston rings for an automotive engine are produced by a forging process, have been taken when the process is in control (Table 3.8). Using the data of inside diameter, data setting up S, $\bar{R}$ and $\bar{S}$, the values of these statistics are calculated to be 0.01055, 0.0221 and 0.00988, respectively.

Table 3.8 Inside Diameter Measurements (mm) on Forged Piston Rings, (Montgomery (2005)).

Sa Nu	Observations				R_i	S_i
1	74.030	74.002	74.019	74.008	0.028	0.012
2	73.995	73.992	74.001	74.004	0.012	0.005
3	73.988	74.024	74.021	74.002	0.036	0.017
4	73.992	74.007	74.015	74.014	0.023	0.011
5	74.009	73.994	73.997	73.993	0.016	0.007
6	73.995	74.006	73.994	74.005	0.012	0.006
7	73.985	74.003	73.993	73.988	0.018	0.008
8	73.998	74.000	73.990	73.995	0.010	0.004
9	74.004	74.000	74.007	73.996	0.011	0.005
10	73.983	74.002	73.998	74.012	0.029	0.012
11	74.006	73.967	73.994	73.984	0.039	0.016
12	74.000	73.984	74.005	73.996	0.021	0.009
13	73.994	74.012	73.986	74.007	0.026	0.012
14	74.006	74.010	74.018	74.000	0.018	0.008
15	74.000	74.010	74.013	74.003	0.013	0.006
16	73.982	74.001	74.015	73.996	0.033	0.014
17	74.004	73.999	73.990	74.009	0.019	0.008
18	74.010	73.989	73.990	74.014	0.025	0.013
19	74.015	74.008	73.993	74.010	0.022	0.009
20	73.982	73.984	73.995	74.013	0.031	0.014

The values of the out-of-control average run length (Table 3.9) are given using equations (20), (22), and (24) into equation (26) for the Shewhart, the Bonferroni-adjustment, and the new R chart, respectively, and also equations (21), (23), and (25) into equation (26) for the Shewhart, the Bonferroni-adjustment, and the new S chart, respectively. Here, the constant values

ψ, d_2, d_3 and c_4 for $k=20$ and $n=4$, are 0.9958, 2.059, 0.88 and 0.9213, in the order mentioned.

Table 3.9 for fixed values $\alpha=0.1, 0.01$ and $\lambda=1.5,2,2.5,3$ shows that the out-of-control average run length for the proposed approach is less than the Shewhart and Bonferroni-adjustment approach for both R and S chart. The power of the control limits is $1-\beta$. As a result, it can be demonstrated that the new approach has maximum power of control limits compared with the Bonferroni-adjustment approach that has minimum power.

Table 3.9 Out-of-Control Average Run Length (k=20, n=4)

	α=0.1				α=0.01			
Approach / λ	1.5	2	2.5	3	1.5	2	2.5	3
Shewhart R chart	2.541	1.539	1.273	1.166	5.885	2.235	1.554	1.317
Bonferroni R chart	7.573	2.501	1.650	1.365	17.604	3.658	2.019	1.540
New R chart	2.463	1.517	1.263	1.161	5.538	2.174	1.531	1.305
Shewhart S chart	2.519	1.527	1.265	1.161	5.795	2.205	1.538	1.306
Bonferroni S chart	7.444	2.464	1.631	1.353	17.212	3.588	1.990	1.523
New S chart	2.444	1.507	1.256	1.156	5.467	2.148	1.517	1.296

3-7. Conclusion

It has been shown that, for an unknown standard deviation parameter, the suggested new range and standard deviation control chart has three advantages over the Shewhart and the Bonferroni-adjustment range and standard deviation control charts.

- The first is that the constant values to construct the control chart for the new approach are based on both sample subgroup size and sample group size;
- The second advantage is that, for a fixed value α, the ARL_1 for the new approach is less than the ARL_1 for the Shewhart and Bonferroni-adjustment approach and
- The third advantage is that the new approach is based on a statistic with variance less than that of the Shewhart and Bonferroni-adjustment approach.

Therefore, practitioners are advised to use the new approach for monitoring the variability of a quality characteristic.

Table 3.10 The constant value ψ for constructing the R and S control chart.

	k							
n	2	3	4	5	6	7	8	9
2	0.8862	0.9213	0.9400	0.9516	0.9594	0.9650	0.9693	0.9726
3	0.9400	0.9594	0.9693	0.9754	0.9794	0.9823	0.9845	0.9862
4	0.9594	0.9726	0.9794	0.9835	0.9862	0.9882	0.9896	0.9908
5	0.9693	0.9794	0.9845	0.9876	0.9896	0.9911	0.9922	0.9931
6	0.9754	0.9835	0.9876	0.9901	0.9917	0.9929	0.9937	0.9945
7	0.9794	0.9862	0.9896	0.9917	0.9931	0.9941	0.9948	0.9954
8	0.9823	0.9882	0.9911	0.9929	0.9941	0.9949	0.9955	0.9961
9	0.9845	0.9896	0.9922	0.9937	0.9948	0.9955	0.9961	0.9965
10	0.9862	0.9908	0.9931	0.9945	0.9954	0.9961	0.9965	0.9969
11	0.9876	0.9917	0.9937	0.9950	0.9958	0.9964	0.9969	0.9972
14	0.9904	0.9937	0.9952	0.9962	0.9968	0.9972	0.9976	0.9979
15	0.9911	0.9941	0.9955	0.9964	0.9970	0.9974	0.9978	0.9980
18	0.9927	0.9951	0.9963	0.9971	0.9976	0.9979	0.9982	0.9984
20	0.9934	0.9956	0.9967	0.9974	0.9978	0.9981	0.9984	0.9985
25	0.9948	0.9965	0.9974	0.9979	0.9983	0.9985	0.9987	0.9989
30	0.9957	0.9971	0.9978	0.9983	0.9986	0.9988	0.9989	0.9990
35	0.9963	0.9976	0.9982	0.9985	0.9988	0.9989	0.9991	0.9992
40	0.9968	0.9979	0.9984	0.9987	0.9989	0.9991	0.9992	0.9993
45	0.9972	0.9981	0.9985	0.9988	0.9991	0.9992	0.9993	0.9994
50	0.9974	0.9983	0.9987	0.9990	0.9991	0.9993	0.9994	0.9994
60	0.9979	0.9986	0.9990	0.9991	0.9993	0.9994	0.9995	0.9995
70	0.9982	0.9988	0.9991	0.9993	0.9994	0.9995	0.9995	0.9996
80	0.9984	0.9989	0.9992	0.9994	0.9995	0.9995	0.9996	0.9996
90	0.9986	0.9991	0.9993	0.9994	0.9995	0.9996	0.9996	0.9997
100	0.9987	0.9992	0.9994	0.9995	0.9996	0.9996	0.9997	0.9997
110	0.9989	0.9992	0.9994	0.9995	0.9996	0.9997	0.9997	0.9997
120	0.9989	0.9993	0.9995	0.9996	0.9996	0.9997	0.9997	0.9998

Table 3.10 (Continued) The constant value ψ for constructing the R and S control chart.

	k							
n	10	15	20	25	30	40	100	120
2	0.9754	0.9835	0.9876	0.9901	0.9917	0.9937	0.9975	0.9979
3	0.9876	0.9917	0.9937	0.9950	0.9958	0.9969	0.9988	0.9990
4	0.9917	0.9945	0.9958	0.9967	0.9972	0.9979	0.9992	0.9993
5	0.9937	0.9958	0.9969	0.9975	0.9979	0.9984	0.9994	0.9995
6	0.9950	0.9967	0.9975	0.9980	0.9983	0.9988	0.9995	0.9996
7	0.9958	0.9972	0.9979	0.9983	0.9986	0.9990	0.9996	0.9997
8	0.9964	0.9977	0.9983	0.9986	0.9988	0.9991	0.9996	0.9997
9	0.9969	0.9979	0.9984	0.9988	0.9990	0.9992	0.9997	0.9997
10	0.9972	0.9981	0.9986	0.9989	0.9991	0.9993	0.9997	0.9998
11	0.9975	0.9983	0.9988	0.9990	0.9992	0.9994	0.9998	0.9998
14	0.9980	0.9987	0.9990	0.9992	0.9994	0.9995	0.9998	0.9998
15	0.9983	0.9988	0.9991	0.9993	0.9994	0.9996	0.9998	0.9999
18	0.9985	0.9990	0.9993	0.9994	0.9995	0.9996	0.9999	0.9999
20	0.9986	0.9991	0.9993	0.9995	0.9996	0.9997	0.9999	0.9999
25	0.9990	0.9993	0.9995	0.9996	0.9997	0.9997	0.9999	0.9999
30	0.9991	0.9994	0.9996	0.9997	0.9997	0.9998	0.9999	0.9999
35	0.9993	0.9995	0.9996	0.9997	0.9998	0.9998	0.9999	0.9999
40	0.9994	0.9996	0.9997	0.9997	0.9998	0.9998	0.9999	0.9999
45	0.9994	0.9996	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000
50	0.9995	0.9997	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000
60	0.9996	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000	1.0000
70	0.9996	0.9998	0.9998	0.9999	0.9999	0.9999	1.0000	1.0000
80	0.9997	0.9998	0.9998	0.9999	0.9999	0.9999	1.0000	1.0000
90	0.9997	0.9998	0.9999	0.9999	0.9999	0.9999	1.0000	1.0000
100	0.9997	0.9998	0.9999	0.9999	0.9999	0.9999	1.0000	1.0000
110	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	1.0000	1.0000
120	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	1.0000	1.0000

Appendix I

$$k(n-1)S^2/\sigma^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij}-\bar{X}_{i.})^2/\sigma^2$$

$$\frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij}-\bar{\bar{X}}_{..})^2}{\sigma^2} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij}-\bar{X}_{i.})^2}{\sigma^2} + \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(\bar{X}_{i.}-\bar{\bar{X}}_{..})^2}{\sigma^2} - \frac{2\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij}-\bar{X}_{i.})(\bar{X}_{i.}-\bar{\bar{X}}_{..})}{\sigma^2}$$

Here, $\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij}-\bar{X}_{i.})(\bar{X}_{i.}-\bar{\bar{X}}_{..})$ is zero, thus

$$\frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij}-\bar{\bar{X}}_{..})^2}{\sigma^2} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij}-\bar{X}_{i.})^2}{\sigma^2} + \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(\bar{X}_{i.}-\bar{\bar{X}}_{..})^2}{\sigma^2}$$

$$= k(n-1)S^2/\sigma^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}(\bar{X}_{i.}-\bar{\bar{X}}_{..})^2/\sigma^2 .$$

According to *Cochran's Theorem* the random variable $k(n-1)S^2/\sigma^2$ is chi square distributed with $k(n-1)$ the degrees of freedom.

Appendix II

(S-Plus)

```
Time=20000 ; al=0.01 ;  n=5 ; lam=2.5  ; k=10
if  (al>=0.1 & al<=0.1) { z=1.6449 ;  if (k>=10 & k<=10)  z2= 2.576 ;
if (k>=15 & k<=15)  z2=2.713 ; if (k>=20 & k<=20)  z2= 2.807
if (k>=30 & k<=30)  z2=2.935 ; if (k>=40 & k<=40)  z2= 3.023 ;
if (k>=100 & k<=100)  z2=3.291 ; if (k>=120 & k<=120)  z2=3.341 }
```

```
if  (al>=0.01 & al<=0.01) { z=2.5758 ; if (k>=10 & k<=10)  z2= 3.291 ;
if (k>=15 & k<=15)  z2=3.403 ; if (k>=20 & k<=20)  z2= 3.481
if (k>=30 & k<=30)  z2=3.588 ; if (k>=40 & k<=40)  z2=3.662 ;
if (k>=100 & k<=100)  z2=3.891 ; if (k>=120 & k<=120)  z2=3.935 }
if  ( n>=2 & n<=2) {
d2=1.128 ; d3=0.853 ;
if (k>=10 & k<=10) psi= 0.9754 ;
if (k>=15 & k<=15)  psi= 0.9835  ;
if (k>=20 & k<=20) psi= 0.9876 ;
if (k>=25 & k<=25) psi= 0.9901 ;
if (k>=30 & k<=30) psi= 0.9917 ;
if (k>=40 & k<=40) psi= 0.9937 ;
if (k>=100 & k<=100) psi= 0.9975 ;
if (k>=120 & k<=120) psi= 0.9979}
if  ( n>=5 & n<=5) {
d2=2.326 ; d3=0.864
if (k>=10 & k<=10)  psi= 0.9937 ;
 if (k>=15 & k<=15)  psi= 0.9958 ;
if (k>=20 & k<=20)  psi= 0.9969 ;
if (k>=25 & k<=25)  psi= 0.9975
if (k>=30 & k<=30) psi= 0.9979 ;
if (k>=40 & k<=40) psi= 0.9984 ;
```

```
if (k>=100 & k<=100) psi= 0.9994 ;
if (k>=120 & k<=120) psi= 0.9995}
if  ( n>=10 & n<=10) {
d2=3.078 ; d3=0.797 ;
if (k>=10 & k<=10) psi= 0.9972 ;
if (k>=15 & k<=15)  psi= 0.9981 ;
if (k>=20 & k<=20) psi= 0.9986 ;
if (k>=25 & k<=25) psi= 0.9989 ;
if (k>=30 & k<=30)  psi= 0.9991 ;
if (k>=40 & k<=40) psi= 0.9993 ;
if (k>=100 & k<=100) psi= 0.9997 ;
if (k>=120 & k<=120) psi= 0.9998 }
if  ( n>=25 & n<=25) {
d2=3.931 ; d3=0.708 ;
if (k>=10 & k<=10) psi= 0.9990 ;
 if (k>=15 & k<=15)  psi= 0.9993 ;
if (k>=20 & k<=20) psi= 0.9995 ;
if (k>=25 & k<=25) psi= 0.9996 ;
if (k>=30 & k<=30) psi= 0.9997 ;
 if (k>=40 & k<=40) psi= 0.9997 ;
 if (k>=100 & k<=100) psi= 0.9999 ;
 if (k>=120 & k<=120) psi= 0.9999 }
```

```
nn1=0 ; nn2=0 ; nn3=0 ; Vs=0
b_matrix(rnorm(20*5,0,1),nrow=20,ncol=5)
   for ( h in 1:20){
      Vs_Vs+ (stdev(b[h,])^2)   }
   St_sqrt(Vs/20)/0.9969 ; Es_( sqrt(Vs/20)/0.9969)*d2 ;
Ss_( sqrt(Vs/20)/0.9969)*d3
for ( i in 1:Time) {
   a_matrix(rnorm(k*n,0,1),nrow=k,ncol=n)
   Rs=0 ;  Vs=0
   for ( h in 1:k){
      Rs_Rs+(max(a[h,])-min(a[h,]))
      Rbar_Rs/k
      Vs_Vs+ (stdev(a[h,])^2)
      Vbar_Vs/k   }
      LCL1_((Rbar/d2)*(d2-z*d3)-(lam*St*d2))/(lam* St *d3) ;
UCL1_((Rbar/d2)*(d2+z*d3)-(lam* St*d2))/(lam* St*d3)
      LCL2_((Rbar/d2)*(d2-z2*d3)-(lam*St*d2))/(lam* St *d3) ;
UCL2_((Rbar/d2)*(d2+z2*d3)-(lam* St*d2))/(lam* St*d3)
      LCL3_(( sqrt(Vbar)/psi)*(d2-z*d3)-(lam* St*d2))/(lam* St *d3) ; UCL3_((
sqrt(Vbar)/psi)*(d2+z*d3)-(lam* St*d2))/(lam* St *d3)
         for ( j in 1:k){
         Zi_((max(a[j,])-min(a[j,]))-(Es))/(Ss) ;  if (Zi<=UCL1 & Zi>=LCL1)
nn1=nn1+1  ; If (Zi<=UCL2 & Zi>=LCL2)  nn2=nn2+1
```

```
        if (Zi<=UCL3 & Zi>=LCL3)  nn3=nn3+1 } }
Beta1_(nn1/(k*Time)) ; Beta2_(nn2/(k*Time)) ; Beta3_(nn3/(k*Time))
print(Beta1) ; print(Beta2) ; print(Beta3)
```

Appendix III

(S-Plus)

```
Time=10000 ; al=0.01 ;  n=5 ; lam=2.5  ; k=10
if  (al>=0.1 & al<=0.1) { z=1.6449 ;
if (k>=10 & k<=10)  z2= 2.576 ;
if (k>=15 & k<=15)  z2=2.713 ;
 if (k>=20 & k<=20)  z2= 2.807
if (k>=30 & k<=30)  z2=2.935 ;
if (k>=40 & k<=40)  z2= 3.023 ;
if (k>=100 & k<=100)  z2=3.291 ;
if (k>=120 & k<=120)  z2=3.341 }
if  (al>=0.01 & al<=0.01) { z=2.5758 ;
if (k>=10 & k<=10)  z2= 3.291 ;
if (k>=15 & k<=15)  z2=3.403 ;
if (k>=20 & k<=20)  z2= 3.481
if (k>=30 & k<=30)  z2=3.588 ;
if (k>=40 & k<=40)  z2=3.662 ;
if (k>=100 & k<=100)  z2=3.891 ;
```

```
 if (k>=120 & k<=120)  z2=3.935 }
if  ( n>=2 & n<=2) {
c4=0.7979 ; c42=0.603 ;
if (k>=10 & k<=10) psi= 0.9754 ;
if (k>=15 & k<=15)  psi= 0.9835  ;
if (k>=20 & k<=20) psi= 0.9876 ;
if (k>=25 & k<=25) psi= 0.9901 ;
if (k>=30 & k<=30) psi= 0.9917 ;
if (k>=40 & k<=40) psi= 0.9937 ;
if (k>=100 & k<=100) psi= 0.9975 ;
if (k>=120 & k<=120) psi= 0.9979}
if  ( n>=5 & n<=5) {
c4=0.94 ; c42=0.381 ;
if (k>=10 & k<=10)  psi= 0.9937 ;
if (k>=15 & k<=15)  psi= 0.9958 ;
 if (k>=20 & k<=20)  psi= 0.9969 ;
 if (k>=25 & k<=25)  psi= 0.9975
if (k>=30 & k<=30) psi= 0.9979 ;
if (k>=40 & k<=40) psi= 0.9984 ;
if (k>=100 & k<=100) psi= 0.9994 ;
 if (k>=120 & k<=120) psi= 0.9995}
if  ( n>=10 & n<=10) {
```

```
c4=0.9727 ; c42=0.232 ;
if (k>=10 & k<=10) psi= 0.9972 ;
if (k>=15 & k<=15)  psi= 0.9981 ;
if (k>=20 & k<=20) psi= 0.9986 ;
if (k>=25 & k<=25) psi= 0.9989 ;
if (k>=30 & k<=30)  psi= 0.9991 ;
if (k>=40 & k<=40) psi= 0.9993 ;
if (k>=100 & k<=100) psi= 0.9997 ;
if (k>=120 & k<=120) psi= 0.9998 }
if  ( n>=25 & n<=25) {
c4=0.9895 ; c42=0.144 ;
if (k>=10 & k<=10) psi= 0.9990 ;
if (k>=15 & k<=15)  psi= 0.9993 ;
if (k>=20 & k<=20) psi= 0.9995 ;
if (k>=25 & k<=25) psi= 0.9996 ;
if (k>=30 & k<=30) psi= 0.9997 ;
if (k>=40 & k<=40) psi= 0.9997 ;
if (k>=100 & k<=100) psi= 0.9999 ;
if (k>=120 & k<=120) psi= 0.9999 }
nn1=0 ; nn2=0 ; nn3=0 ; Vs=0
b_matrix(rnorm(20*5,0,1),nrow=20,ncol=5)
  for ( h in 1:20){
```

```
    Vs_Vs+ (stdev(b[h,])^2)   }
  St_sqrt(Vs/20)/0.9969 ; Es_( sqrt(Vs/20)/0.9969)*c4 ; Ss_(
sqrt(Vs/20)/0.9969)*c42
for ( i in 1:Time) {
  a_matrix(rnorm(k*n,0,1),nrow=k,ncol=n)
  Sbar1=0 ;  Vs=0
  for ( h in 1:k){
    Sbar1_Sbar1+(stdev(a[h,]))
    Sbar_Sbar1/k
    Vs_Vs+ (stdev(a[h,])^2)
    Vbar_Vs/k  }
    LCL1_((Sbar/c4)*(c4-z*c42)-(lam*St*c4))/(lam* St *c42) ;
UCL1_((Sbar/c4)*(c4+z*c42)-(lam* St*c4))/(lam* St*c42)
    LCL2_((Sbar/c4)*(c4-z2*c42)-(lam*St*c4))/(lam* St *c42) ;
UCL2_((Sbar/c4)*(c4+z2*c42)-(lam* St*c4))/(lam* St*c42)
    LCL3_(( sqrt(Vbar)/psi)*(c4-z*c42)-(lam* St*c4))/(lam* St *c42) ; UCL3_((
sqrt(Vbar)/psi)*(c4+z*c42)-(lam* St*c4))/(lam* St *c42)
        for ( j in 1:k){
        Zi_((stdev(a[j,]))-(Es))/(Ss) ;  if (Zi<=UCL1 & Zi>=LCL1)  nn1=nn1+1
; if (Zi<=UCL2 & Zi>=LCL2)  nn2=nn2+1
        if (Zi<=UCL3 & Zi>=LCL3)  nn3=nn3+1 } }
Beta1_(nn1/(k*Time)) ; Beta2_(nn2/(k*Time)) ; Beta3_(nn3/(k*Time))
print(Beta1) ; print(Beta2) ; print(Beta3)
```

4

A Highly Accurate Approximation for the NDF and Mean Range of the Normal Distribution

CHAPTER OUTLINE

4-1. Introduction

The normal distribution function (NDF) plays a central role in statistical theory. This function numerically has been evaluated in the literature by using computer programming on values of $z \in (-\infty, +\infty)$. Theoretically, the NDF, $F(z)$, has been evaluated only for a subset of the values $z \in (0, +\infty)$ or $z \in (-9, +9)$, where the function $F(z)$ is

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-t^2}{2}} dt \quad , \qquad -\infty < z < +\infty . \tag{1}$$

The new evaluation of the function $F(z)$ that is proposed in this chapter is based on values of $z \in (-\infty, +\infty)$ with integration region $(-\infty, z)$ for $z \in (-\infty, +\infty)$ and is based on the error function $(erf(z))$. The integration region of the error function is $(0, z)$ for $z > 0$, or $(0, -z]$ for $z \le 0$. The error function is defined by

$$erf(z) = \int_{0}^{z} \frac{2}{\sqrt{\pi}} e^{\frac{-t^2}{2}} dt \quad , \qquad -\infty < z < +\infty . \tag{2}$$

The relation between the functions $F(z)$ and $erf(z)$ is

$$F(z) = \frac{1}{2}(1 - erf(-z/\sqrt{2})), \qquad -\infty < z < +\infty . \tag{3}$$

The proposed method for the evaluation of the NDF utilizes this relationship. The error function is first evaluated by using the polar integration. Then, the NDF is evaluated using equation (3). The associated greatest absolute error is less than 4.02×10^{-14} for $z \in (-\infty, +\infty)$.

In section 3-3 a new numerical evaluation of the NDF is present with 21 decimal places accuracy.

Suppose that the random variables Z_i are identically standard normal distributed for $i=1,2,...,n$, and the range of order statistics $Z_{(1)},Z_{(2)},...,Z_{(n)}$ is presented by $R=Z_{(n)}-Z_{(1)}$. The accurate evaluation for the mean range of the normal distribution ($E(R)=d_2$) is calculated using the proposed formula of the NDF, where

$$E(R)=\int_{-\infty}^{+\infty}\left\{1-\left(F(z)\right)^n-\left(1-F(z)\right)^n\right\}\,dz$$

The extended table of the previous equation for $n=2,(1)100,(20)440$ is constructed for the range control charts.

This chapter briefly discusses the existing theoretical and numerical evaluations of the NDF, in section 3-2. The new evaluation of the NDF and an application of this evaluation for computing the mean range are presented in sections 3-3, 3-4, respectively.

4-2. The Existing Formulas for the Evaluation of NDF

There are many formulas for the approximation of the NDF. Here we briefly present some of them. As already mentioned, most of these formulas were evaluated for a subset of the values of $z\in(0,+\infty)$ or $z\in(-9,+9)$.

Zelen and Severo (1964) suggested the following formulas (I-III), which hold for $z\in(0,+\infty)$:

A Highly Accurate Approximation for the NDF and Mean Range of the Normal Distribution

(I) $$F(z)=\int_0^z (a_0+a_2t^2+a_4t^4+a_6t^6)^{-1}dx \quad , \quad z\geq 0 ,$$

where $a_0=2.490895$, $a_2=1.466003$, $a_4=-0.024393$, and $a_6=0.178257$. The associated error, for $z\geq 0$, is less than 0.0027.

(II) $$F(z)=1-\frac{1}{2}(1+a_1z+a_2z^2+a_3z^3+a_4z^4)^{-4} \quad , \quad z\geq 0 ,$$

where,

$a_1=0.196854$; $a_2=0.115194$;

$a_3=0.000344$; $a_4=0.019527$; $a_5=0.937298$,

and the associated error, for $z\geq 0$, is less than 0.00025.

(III) $$F(z)=1-(a_1t+a_2t^2+a_3t^3)\frac{1}{\sqrt{2\pi}}e^{-z^2/2} \quad , \quad z\geq 0 .$$

Here,

$t=(1+0.33267z)^{-1}$; $a_1=0.4361836$;

$a_2=-0.1201676$; $a_3=0.937298$,

and the associated error, for $z\geq 0$, is less than 0.00001.

An approximate evaluation of the NDF for $z\geq 0$ with maximum absolute error 1.5×10^{-8} for $z=0$, is provided by the formula (see, e.g., Abramowitz and Stegun (1964))

$$F(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{\frac{-x^2}{2}}dx \approx 1-e^{\frac{-x^2}{2}}(b_1t+b_2t^2+b_3t^3+b_4t^4+b_5t^5)/\sqrt{2\pi}$$

where

$z\geq 0$ $t=1/(1+pz)$

$p=0.2316419$ $b_3=1.781477937$

$b_1 = 0.31938153$ $\qquad$ $b_4 = -1.821255978$

$b_2 = -0.356563782$ $\qquad$ $b_5 = 1.330274429$.

Hart (1966) suggested the following formula for $z > 2$,

$$F(z) = 1 - (z\sqrt{2\pi})^{-1}[\exp(-z^2/2)] \times \left(1 - \frac{(1+bz^2)^{1/2}}{(1+az^2)} \left\{ z\sqrt{\frac{\pi}{2}} + \left[\frac{1}{2}\pi z^2 + \frac{(1+bz^2)^{1/2}}{(1+az^2)} \exp(-z^2/2) \right]^{1/2} \right\}^{-1} \right)$$

This formula gives maximum absolute error 5.32×10^{-5} for $z = 1.04$. The constant values a and b are

$$a = \frac{1}{2\pi}\left[1 + (1 + 6\pi - 2\pi^2)^{1/2}\right] = 0.212024 ;$$

$$b = \frac{1}{2\pi}\left[1 + (1 + 6\pi - 2\pi^2)^{1/2}\right]^2 = 0.282455 .$$

Schucany and Gray (1968) introduced the following formula with maximum absolute error 1×10^{-5}

$$1 - F(z) = \left[(z^2 + 2)\sqrt{2\pi}\right]^{-1} \times z\left[\exp\left(\frac{-z^2}{2}\right)\right] \times \frac{-28 + 14z^2 + 6z^4 + z^6}{-4 - 20z^2 + 5x^4 + z^6} \qquad z > 2$$

Badhe (1976) presented the following two formulas:

$$F(z) = 1 - \frac{\frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}}{z}\left[1 - \frac{1}{(z^2+10)}\left(1 + \frac{1}{(z^2+10)}\left\{7 + \frac{1}{(z^2+10)}\right.\right.\right.$$

$$\left.\left.\left. \times \left[55 + \frac{1}{(z^2+10)}\left(445 + \frac{3745\left(8.5(z^2 - 0.4284639753z^{-2} + 1.240964109)^{-1} + 1\right)}{(z^2+10)}\right)\right]\right\}\right)\right]$$

with a maximum absolute error 5×10^{-8} for $z \geq 2$ and

A Highly Accurate Approximation for the NDF and Mean Range of the Normal Distribution

$$F(z)=\frac{1}{2}+x\left\{a+\frac{z^2}{32}\left[b+\frac{z^2}{32}\left(c+\frac{z^2}{32}\left\{d+\frac{z^2}{32}\left[e+\frac{z^2}{32}\left(f+\frac{z^2}{32}\left(g+h\frac{z^2}{32}\right)\right)\right]\right\}\right)\right]\right\}$$

for $0\le z\le 2$ with a maximum absolute error 2×10^{-9}

Here,

$a=0.3989422784$ $b=-2.127690079$

$c=10.2125662121$ $d=-38.8830314909$

$e=120.2836370787$ $f=-303.2973153419$

$g=575.073131917$ $h=-603.9068092058$.

Two further formulas for the NDF with maximum absolute error of 0.042% and 0.040% respectively were presented by Derenzo (1977).

$$F(z)=1-\frac{1}{2}\exp\left\{-\frac{(83z+351)z+562}{(703/x)+165}\right\} \quad \text{for} \quad 0<z\le 5.5,$$

$$F(z)=1-\frac{1}{\sqrt{2\pi}z}\exp\left(-\frac{z^2}{2}-\frac{0.94}{z^2}\right) \quad \text{for} \quad 5.5<z.$$

An approximation for the NDF with maximum absolute error 3×10^{-7} for $z\ge 0$ was introduced by Divgi (1979).

$$F(z)=(2\pi)^{-1/2}e^{-z^2/2}\sum_{j=0}^{10}a_jz^j,$$

where

$a_0=1.253313402$ $a_1=-0.9999673043$

$a_2=0.6262972801$ $a_3=-0.3316218430$

$a_4=0.1522723563$ $a_5=-5.982834993\times10^{-2}$

$a_6=1.91564935\times10^{-2}$ $a_7=-4.644960579\times10^{-3}$

$$a_8 = 7.771088713\times10^{-4} \qquad a_9 = -7.830823677\times10^{-5}$$

$$a_{10} = 3.534244658\times10^{-6}.$$

Moran (1980) suggested using the following series with maximum absolute error 1×10^{-9}.

$$F(z) = \frac{1}{2} + \frac{1}{\pi}\sum_{n=0}^{12}\left(n+\frac{1}{2}\right)^{-1} e^{\frac{-(n+(1/2))^2}{9}} \sin\left\{\frac{\sqrt{2}}{3}\left(n+\frac{1}{2}\right)z\right\}, \qquad |z| \le 7.$$

Lew (1981) introduced the following approximation with a maximum absolute error 0.002.

$$F(z) = \begin{cases} 1-(0.5-(2\pi)^{-1/2}(z-z^3/7) & , \quad 0 \le z \le 1 \\ (1+z)(2\pi)^{-1/2}e^{-z^2/2}/(1+z+z^2), & \quad 1 < z. \end{cases}$$

Bagby (1995) suggested that an approximation of the NDF with absolute error less than 3×10^{-5} could be

$$F(z) = 0.5 \pm 0.5\{1-(1/30)[7\exp(-z^2/2)+$$

$$16\exp(-z^2(2-\sqrt{2}))+(7+\pi z^2/4)\exp(-z^2)]\}^{1/2}.$$

Here, the minus sign refers to $z \le 0$.

The following approximation with maximum absolute error 3.1×10^{-4} of the NDF for $z \in (0,+\infty)$ was introduced by Waissi and Rossin (1996).

$$F(z) = 1/\left(1+\exp(\sum_{k=0}^{\infty} a_k z^{2k+1}\right),$$

where

$$a_1 z + a_2 z^3 + a_3 z^5 = 1.595208466z + 0.7412366556z^3 + 0.0007809431668z^5.$$

A Highly Accurate Approximation for the NDF and Mean Range of the Normal Distribution

Bryc (2002) presented a formula with maximum absolute error 1.9×10^{-5}.

$$F(z)=\frac{z^2+5.575192695z+12.77436324}{\sqrt{2\pi}z^3+14.38718147z^2+31.53531977z+2\times12.77436324}e^{-z^2/2}, \qquad z>0.$$

An approximation for the NDF, with greatest absolute error 2×10^{-6} was suggested by Shore (2004).

$$F(z)=[1+g(-z)-g(z)]/2,.$$

Here,

$$g(z)=\exp\{-\log(2)\exp[B[\exp(A+Cz)-1]+Dz]\},,$$

where,

$A=0.072880383$; $B=-2.36699428$;

$C=-0.40639929$; $D=0.19417768$.

Later, Shore (2005) improved the above approximation with a maximum absolute error 6×10^{-7} by proposing the formula

$$F(z)=[1+g(-z)-g(z)]/2 \quad , \qquad -9<z<9.$$

In this case,

$$g(z)=\exp\{-\log(2)\exp\{[\alpha/(\lambda/S_1)][(1+S_1z)^{(\lambda/S_1)}-1]+S_2z\}\} \quad , \qquad -9<z<9,$$

where,

$\lambda=-0.61228883$; $S_1=-0.11105481$;

$S_2=0.44334159$; $\alpha=-6.37309208$.

Some more approximations for the NDF have been suggested by Adams (1969), Hamaker (1978), Parsonson (1978), Heard (1979), Kennedy and Gentle (1980), Martynov (1981), Monahan (1981), Edgeman (1988), Pugh (1989), Revfeim (1990), Vedder (1993), Johnson et al (1994). In the sequel, we present some evaluations of the NDF based on computer programming.

Kerridge and Cook (1976) suggested using the following series for computing $F(z)$ on a computer ($z>0$) with absolute error less than 10^{-16}

$$F(z)=\int_0^z \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} = ze^{\frac{-z^2}{8}} \sum_{n=0}^{+\infty} \frac{1}{2n+1}\theta_{2n}(z/2), \qquad z>0,$$

where

$$\theta_n(z)=x^n H_n(z)/n!$$

$$\theta_{n+1}(z)=\frac{z^2(\theta_n-\theta_{n-1})}{n+1}, \qquad n=1,2,...$$

$$H_{n+1}(z)=zH_n(z)-nH_{n-1}(z), \qquad n=1,2,...$$

$$H_0(z)=1, \qquad H_1(z)=z.$$

Marsaglia (2004) provided the evaluation given bellow with absolute error less than 8×10^{-16} for $z\in(-\infty,+\infty)$

$$F(z)=0.5+(2\pi)^{-1/2}e^{-z^2/2}\left(z+\frac{z^3}{3}+\frac{z^5}{3.5}+\frac{z^7}{3.5.7}+\frac{z^9}{3.5.7.9}+...\right).$$

He provided the following function (using C) for the computation of $F(z)$

double Phi(double z)

*{Long double s = z, t = 0, b = z, q = z * z, i = 1;*

$$while(s!=t)\ s=(t=s)+(b*=q/(i+=2));$$

$$return\ 0.5+s*\exp(-0.5*q-0.91893853320467274178L);\ \}$$

In the next section, we present a simple formula for the numerical evaluation of the NDF with satisfactory accuracy for values of $z\in(-\infty,+\infty)$.

4-3. A New Evaluation of the NDF

In this section we propose evaluating the NDF using the relationship between the distribution function $F(z)$ and the error function $erf(z)$ given by (3). The integration region for the function $F(z)$ is $(-\infty,z)$ while that for the error function is $(0,z)$ and $(0,-z]$ for $z>0$ and $z\le 0$, respectively. Replacing z by $-z/\sqrt{2}$ in (2), we obtain

$$erf(-z/\sqrt{2})=\frac{2}{\sqrt{\pi}}\int_0^{-z/\sqrt{2}} e^{-t^2}dt\ ,\qquad -\infty<z<+\infty\ . \tag{4}$$

Let

$$I=\int_0^{-z/\sqrt{2}} e^{-t^2}dt\ . \tag{5}$$

Therefore I^2 is the following double integral that is based on the rectangular coordinates

$$I^2=\int_0^{-z/\sqrt{2}}\int_0^{-z/\sqrt{2}} e^{-(t_1^2+t_2^2)}dt_1dt_2\ . \tag{6}$$

The relationship between rectangular and polar coordinates can be found using the definition of trigonometric functions as

$$t_1=r\,Cos(\beta)\ ;\ t_2=r\,Sin(\beta)\ . \tag{7}$$

For $z \le 0$ *and* $r \ge 0$ *we have from equations (6) and (7),*

$$0 \le t_1 \le \frac{-z}{\sqrt{2}} \quad \Rightarrow \quad 0 \le r \le \frac{-z}{Cos(\beta)\sqrt{2}} \quad \Rightarrow \quad Cos(\beta) \ge 0$$

$$0 \le t_2 \le \frac{-z}{\sqrt{2}} \quad \Rightarrow \quad 0 \le r \le \frac{-z}{Sin(\beta)\sqrt{2}} \quad \Rightarrow \quad Sin(\beta) \ge 0. \tag{8}$$

Relation (8) gives $0 \le \beta \le \frac{\pi}{2}$ (Figure 4.1).

Also, for $z > 0$ *and* $r > 0$, *we obtain from equations* (6) *and* (7),

$$\frac{-z}{\sqrt{2}} \le t_1 < 0 \quad \Rightarrow \quad 0 < r \le \frac{-z}{Cos(\beta)\sqrt{2}} \quad \Rightarrow \quad Cos(\beta) \le 0$$

$$\frac{-z}{\sqrt{2}} \le t_2 < 0 \quad \Rightarrow \quad 0 < r \le \frac{-z}{Sin(\beta)\sqrt{2}} \quad \Rightarrow \quad Sin(\beta) \le 0 \tag{9}$$

Note that relation (9) implies $\pi \le \beta \le \frac{3\pi}{2}$ (Figure 4.2).

In what follows, equation (6) *will first be evaluated under the assumption that* $z \le 0$. Using relation (8):

$$-z/\sqrt{2} \ge t_1 \ge t_2 \ge 0 \quad \Rightarrow \quad Cos(\beta) \ge Sin(\beta) \ge 0.$$

Therefore,

$$0 \le \beta \le \frac{\pi}{4} \text{ and } 0 \le r \le \frac{-z}{Cos(\beta)\sqrt{2}}. \tag{10}$$

Similarly,

$$-z/\sqrt{2} \ge t_2 > t_1 \ge 0 \quad \Rightarrow \quad 0 \le Cos(\beta) < Sin(\beta).$$

Hence,

$$\frac{\pi}{4} < \beta \le \frac{\pi}{2} \text{ and } 0 \le r \le \frac{-z}{Sin(\beta)\sqrt{2}}. \tag{11}$$

Figure 4.1 Relationship between rectangular and polar coordinates

$(r \geq 0\,,\, z \leq 0\,,\, 0 \leq \beta \leq \pi/2)$

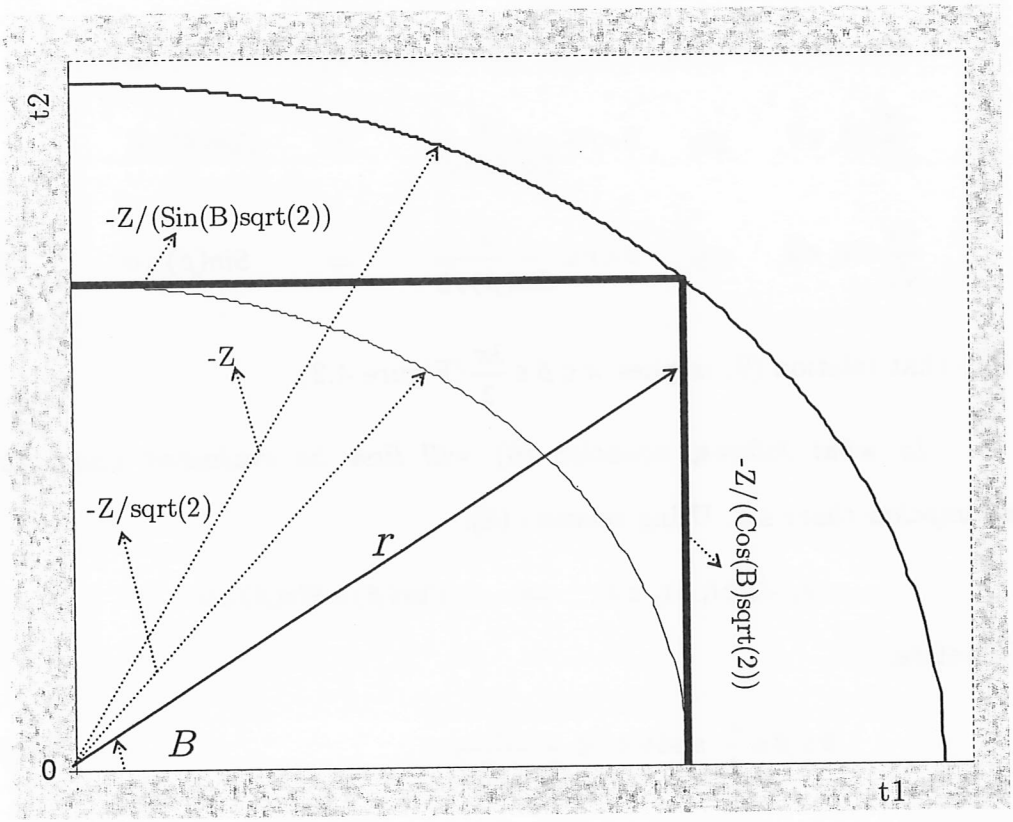
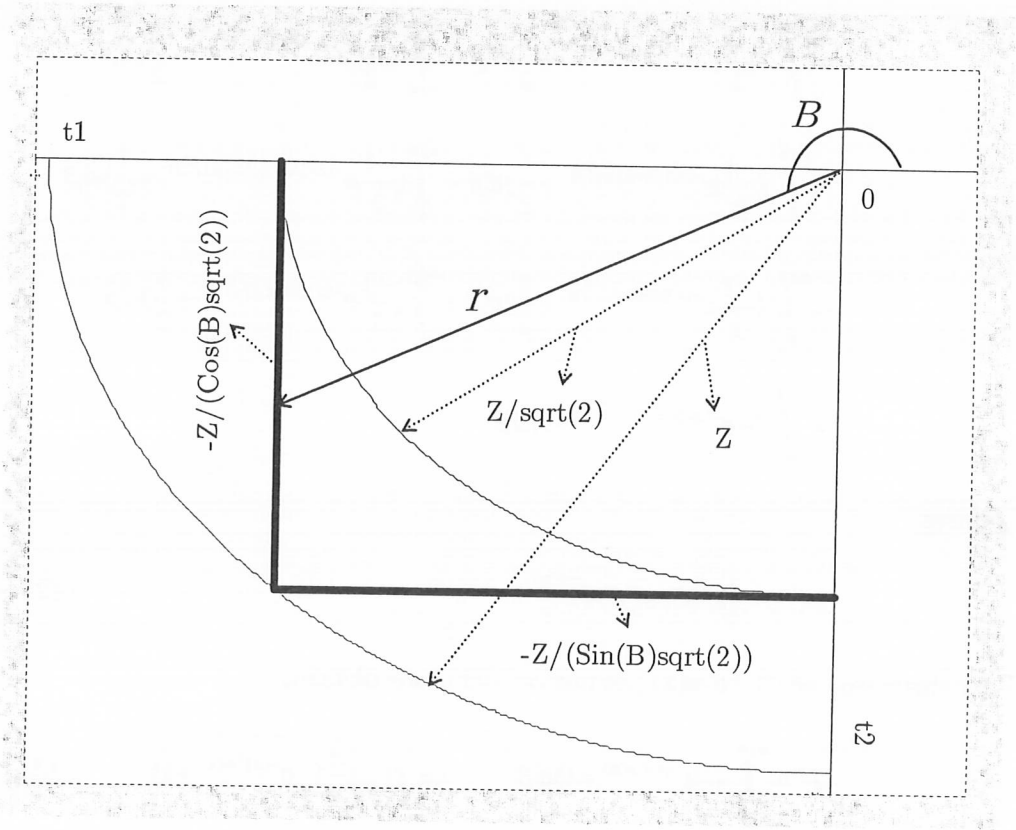


Figure 4.2 Relationship between rectangular and polar coordinates

$(r > 0\,,\, z > 0\,,\, \pi \le \beta \le 3\pi/2)$



Equation (6), for $z \le 0$, involving rectangular integrals is transformed to one involving polar integrals by using equation (7) and relations (10) and (11), thus

$$I^2 = \int_{\beta}\int_{r} re^{-r^2}\,dr\,d\beta$$

$$= \int_0^{\pi/4}\int_0^{-z/(Cos(\beta)\sqrt{2})} re^{-r^2}\,dr\,d\beta + \int_{\pi/4}^{\pi/2}\int_0^{-z/(Sin(\beta)\sqrt{2})} re^{-r^2}\,dr\,d\beta$$

$$= \int_0^{\pi/4} \frac{-1}{2}e^{-r^2}\Big)_0^{-z/(Cos(\beta)\sqrt{2})}\,d\beta + \int_{\pi/4}^{\pi/2} \frac{-1}{2}e^{-r^2}\Big)_0^{-z/(Sin(\beta)\sqrt{2})}\,d\beta$$

$$= \int_0^{\pi/4} (\frac{-1}{2}e^{-(z/(Cos(\beta)\sqrt{2}))^2} + \frac{1}{2})d\beta + \int_{\pi/4}^{\pi/2} (\frac{-1}{2}e^{-(z/(Sin(\beta)\sqrt{2}))^2} + \frac{1}{2})d\beta$$

$$= \int_0^{\pi/4} (\frac{-1}{2}e^{-(z/(Cos(\beta)\sqrt{2}))^2} + \frac{1}{2})d\beta + \int_0^{\pi/4} (\frac{-1}{2}e^{-(z/(Sin(\pi/2-\beta)\sqrt{2}))^2} + \frac{1}{2})d\beta$$

$$= \int_0^{\pi/4} (\frac{-1}{2}e^{-(z/(Cos(\beta)\sqrt{2}))^2} + \frac{1}{2})d\beta + \int_0^{\pi/4} (\frac{-1}{2}e^{-(z/(Cos(\beta)\sqrt{2}))^2} + \frac{1}{2})d\beta$$

$$= \int_0^{\pi/4} (-e^{-(z/(Cos(\beta)\sqrt{2}))^2} + 1)d\beta\,.$$

Letting,

$$\omega(\beta) = \left(1/(Cos(\beta)\sqrt{2})\right)^2. \tag{12}$$

Transforming $\omega(\beta)$ to $\omega(z)$, equation (21), we obtain,

$$I^2 = \int_0^{\pi/4} (-e^{-(z)^2\omega(z)} + 1)d\beta \qquad \Rightarrow I = \sqrt{\frac{\pi}{4}(-e^{-(z)^2\omega(z)} + 1)}\,. \tag{13}$$

Therefore, combining (5), (6), and (13), we have

$$erf(-z/\sqrt{2}) = 2I/\sqrt{\pi}\,, \qquad z \le 0\,. \tag{14}$$

Equation (6) can now be evaluated under the assumption that $z > 0$.

Using relation (9), we have

$$0 > t_1 \geq t_2 \geq -z/\sqrt{2} \quad \Rightarrow \quad 0 > Cos(\beta) \geq Sin(\beta).$$

Therefore,

$$\pi \leq \beta \leq \frac{5\pi}{4} \quad \text{and} \quad 0 < r \leq \frac{-z}{Cos(\beta)\sqrt{2}}. \tag{15}$$

Similarly,

$$0 > t_2 \geq t_1 \geq -z/\sqrt{2} \quad \Rightarrow \quad 0 > Sin(\beta) \geq Cos(\beta).$$

Hence,

$$\frac{5\pi}{4} < \beta \leq \frac{3\pi}{2} \quad \text{and} \quad 0 < r \leq \frac{-z}{Sin(\beta)\sqrt{2}}. \tag{16}$$

The $erf(-z/\sqrt{2})$, for $z > 0$, can be evaluated as

$$erf(-z/\sqrt{2}) = -2I/\sqrt{\pi} \qquad z > 0. \tag{17}$$

This is the result of the fact that,

$$erf(-z/\sqrt{2}) = \frac{2}{\sqrt{\pi}} \int_0^{-z/\sqrt{2}} e^{-t^2} dt$$

$$= -\frac{2}{\sqrt{\pi}} \int_{-z/\sqrt{2}}^{0} e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \int_{z/\sqrt{2}}^{0} e^{-t^2} dt \qquad (\textit{transforming } t \textit{ to } -t)$$

$$= -\frac{2}{\sqrt{\pi}} \int_0^{z/\sqrt{2}} e^{-t^2} dt = \frac{-2}{\sqrt{\pi}} \sqrt{\frac{\pi}{4}(-e^{-(z)^2\omega(z)} + 1)} = -2I/\sqrt{\pi} \ (\textit{Similarly to equation } 13).$$

Using equations (14) and (17), we obtain

$$erf(-z/\sqrt{2}) = Sign(-z)\left(2I/\sqrt{\pi}\right) = \begin{cases} -2I/\sqrt{\pi} & , \quad z > 0 \\ 2I/\sqrt{\pi} & , \quad z \leq 0 \end{cases} \tag{18}$$

where the function $Sign(-z)$ is defined by

$$Sign(-z) = \begin{cases} -1 & if \quad z > 0 \\ 1 & if \quad z \leq 0. \end{cases}$$

Equations (13) and (18) give

$$erf(-z/\sqrt{2}) = Sign(-z)\sqrt{1-e^{-z^2\omega}} \,. \tag{19}$$

Let us denote the proposed approximation of $F(z)$ by $F_A(z)$. Combining equations (3) and (19) yields

$$F_A(z) = \frac{1}{2}(1 - Sign(-z)\sqrt{1-e^{-z^2\omega(z)}}) \,. \tag{20}$$

The function $\omega(z)$ of equation (20) is approximated using equations (5) and (13),

$$\omega(z) = Ln\left(1 - \frac{4}{\pi}\left(\int_0^{-z/\sqrt{2}} e^{-t^2} dt\right)^2\right) / -z^2 \,. \tag{21}$$

To evaluate equation (21), the integral $\int_0^{-z/\sqrt{2}} e^{-t^2} dt$ is computed numerically, using Taylor expansion of this integral that is suitably convergent. The Taylor series of e^{-t^2} (21) is

$$e^{-t^2} = e^{-c^2} \sum_{k=0}^{+\infty} \frac{u^{(k)}(c)}{k!} (t-c)^k \,. \tag{22}$$

Note that c and t are in $[0,+\infty)$, $u^k(c) = f^{(k)}(c)/e^{-c^2}$, and

$$u^{(0)}(c) = 1 \;\; ; \;\; u^{(1)}(c) = -2c \,;$$

$$u^{(k)}(c) = -2(k-1)u^{(k-2)}(c) + (-2c)u^{(k-1)}(c) \qquad k \geq 2$$

The values of c are given below for various intervals of t.

$c = 1$ for $t \in [0,1.2)$; $c = 3$ for $t \in [1.2,3.2)$

$c = 5$ for $t \in [3.2,5.2)$; $c = 7$ for $t \in [5.2,7.2)$

$c = 9$ for $t \in [7.2,9.2)$; $c = 11$ for $t \in [9.2,+\infty)$

Based on the above, the maximum absolute error between the exact value of e^{-t^2} and its approximation using (22) would be -7.9×10^{-23} at $t=4.1$ (see Table 4.1).

Table 4.1: The error between exact and Taylor expansion values of e^{-t^2}

t	*Error*	*t*	*error*	*t*	*error*	*t*	*Error*	*t*	*error*	*t*	*Error*	*t*	*Error*
0.0	0.0E+00	1.6	0.0E+00	3.2	0.0E+00	4.8	-1.2E-25	6.4	-5.4E-33	8.0	0.0E+00	9.6	-3.3E-55
0.1	0.0E+00	1.7	0.0E+00	3.3	0.0E+00	4.9	-6.5E-26	6.5	0.0E+00	8.1	-1.8E-43	9.7	-4.8E-56
0.2	0.0E+00	1.8	0.0E+00	3.4	0.0E+00	5.0	0.0E+00	6.6	0.0E+00	8.2	-4.2E-44	9.8	-1.2E-56
0.3	0.0E+00	1.9	0.0E+00	3.5	0.0E+00	5.1	-8.1E-27	6.7	5.4E-35	8.3	-3.7E-45	9.9	0.0E+00
0.4	0.0E+00	2.0	0.0E+00	3.6	0.0E+00	5.2	-1.6E-27	6.8	2.3E-35	8.4	7.4E-46	10.0	0.0E+00
0.5	0.0E+00	2.1	0.0E+00	3.7	0.0E+00	5.3	-1.2E-27	6.9	-3.0E-36	8.5	0.0E+00	10.1	1.2E-59
0.6	0.0E+00	2.2	0.0E+00	3.8	0.0E+00	5.4	0.0E+00	7.0	0.0E+00	8.6	0.0E+00	10.2	-4.3E-60
0.7	0.0E+00	2.3	0.0E+00	3.9	0.0E+00	5.5	0.0E+00	7.1	0.0E+00	8.7	5.0E-48	10.3	-2.1E-61
0.8	0.0E+00	2.4	0.0E+00	4.0	0.0E+00	5.6	0.0E+00	7.2	0.0E+00	8.8	0.0E+00	10.4	-4.0E-62
0.9	0.0E+00	2.5	0.0E+00	**4.1**	**-7.9E-23**	5.7	0.0E+00	7.3	0.0E+00	8.9	-6.4E-50	10.5	0.0E+00
1.0	0.0E+00	2.6	0.0E+00	4.2	2.6E-23	5.8	-5.9E-30	7.4	1.8E-39	9.0	0.0E+00	10.6	-1.0E-63
1.1	0.0E+00	2.7	0.0E+00	4.3	0.0E+00	5.9	1.1E-30	7.5	0.0E+00	9.1	5.5E-51	10.7	8.1E-65
1.2	0.0E+00	2.8	0.0E+00	4.4	0.0E+00	6.0	0.0E+00	7.6	-3.0E-40	9.2	2.3E-52	10.8	0.0E+00
1.3	0.0E+00	2.9	0.0E+00	4.5	0.0E+00	6.1	1.2E-31	7.7	-5.2E-41	9.3	1.5E-52	10.9	1.4E-66
1.4	0.0E+00	3.0	0.0E+00	4.6	0.0E+00	6.2	-4.3E-32	7.8	0.0E+00	9.4	-2.0E-53	11.0	0.0E+00
1.5	0.0E+00	3.1	0.0E+00	4.7	-5.2E-25	6.3	0.0E+00	7.9	8.1E-43	9.5	0.0E+00	Inf	0.0E+00

If we round values of this table to 21 decimal places, we will have values of error equal to zero. This implies that the right side of (22) is an exact approximation of e^{-t^2} with 21 decimal places. Integrating (22) with respect to t from 0 to $-z/\sqrt{2}$, we have

$$\int_0^{-z/\sqrt{2}} e^{-t^2}dt = \int_0^{-z/\sqrt{2}} e^{-c^2}\sum_{k=0}^{+\infty}\frac{u^{(k)}(c)}{k!}(t-c)^k \qquad -\infty < z < +\infty$$

$$= Sign(-z)[e^{-c^2}\sum_{k=0}^{+\infty}\frac{u^{(k)}(c)}{(k+1)!}\left((-z/\sqrt{2}-c)^{k+1}-(-c)^{k+1}\right)] \qquad (23)$$

The right hand side of the previous equation, to (21) decimal places, is the approximation of $\int_0^{-z/\sqrt{2}} e^{-t^2}dt$ with error equal zero, for $-\infty < z < +\infty$.

Using equations (23) and (21) an empirically approximate value for ω is 0.623. This value gives maximum absolute error 0.0016 for $F(z)$ when $|z|=0.6$. For more precision, the value ω can be replaced by empirically approximate values of Table 4.2, using (24) or (27). The greatest absolute error of $F(z)$ using Table 4.2 is 47×10^{-5}, in equation (24) is less than 5×10^{-8}, and in equation (24) is less than 4.02×10^{-14}.

The right hand side of equation (23) indicates to an evaluation of $erf(-z/\sqrt{2})$, using this evaluation in equation (3) we will have a numerical approximation of the NDF with 21 decimal places accuracy.

Table 4.2: The value ω for the NDF

Z	ω	Maximum Absolute Error	Z	ω	Maximum Absolute Error
$0.00 \le \lvert z \rvert < 0.25$	0.6362	0.00001	$2.00 \le \lvert z \rvert < 2.25$	0.6025	0.0002
$0.25 \le \lvert z \rvert < 0.50$	0.6349	0.00009	$2.25 \le \lvert z \rvert < 2.50$	0.5960	0.0001
$0.50 \le \lvert z \rvert < 0.75$	0.6325	0.0002	$2.50 \le \lvert z \rvert < 2.75$	0.5900	0.00009
$0.75 \le \lvert z \rvert < 1.00$	0.6295	0.0003	$2.75 \le \lvert z \rvert < 3.00$	0.5841	0.000047
$1.00 \le \lvert z \rvert < 1.25$	0.6250	0.0004	$3.00 \le \lvert z \rvert < 3.25$	0.5785	0.00002
$1.25 \le \lvert z \rvert < 1.50$	0.6202	0.00047	$3.25 \le \lvert z \rvert < 3.50$	0.5733	0.00001
$1.50 \le \lvert z \rvert < 1.75$	0.6145	0.0004	$3.50 \le \lvert z \rvert < 3.75$	0.5685	0.000004
$1.75 \le \lvert z \rvert < 2.00$	0.6086	0.0003	$3.75 \le \lvert z \rvert$	0.5639	0.000002

$$\omega(z)=\begin{cases}0.0004254z^4+0.000001|z|^3-0.0095654z^2+0.0000003|z|+0.63661977 & 0\le|z|<1.05\\ 0.00009561z^4+0.0015564|z|^3-0.012371z^2+0.00228|z|+0.6359194 & 1.05\le|z|<1.97\\ -0.00039485z^4+0.00521546|z|^3-0.02265z^2+0.0151586|z|+0.6298524 & 1.97\le|z|<7\\ 0.467 & 7\le|z|\end{cases} \quad (24)$$

Let $e(z)=F_A(z)-F(z)$, where $F(z)$ and $F_A(z)$ are given by equations (1) and (20), respectively. From equation (24), we have

$$\frac{\partial\omega}{\partial z}=\begin{cases}0.0017016|z|^3+0.000003z^2-0.0191308|z|+0.0000003 & 0\le|z|<1.05\\ 0.00038244|z|^3+0.0046692z^2-0.024742|z|+0.00228 & 1.05\le|z|<1.97\\ -0.0015794|z|^3+0.01564638z^2-0.0453|z|+0.0151586 & 1.97\le|z|<7\\ 0 & 7\le|z|\end{cases} \quad (25)$$

The derivative of $e(z)$ is given by

$$\frac{\partial e(z)}{\partial z} = Sign(-z)(-\frac{1}{4}(2|z|\omega + \frac{\partial \omega(z)}{\partial z} z^2) e^{-z^2\omega(z)} / \sqrt{1-e^{-z^2\omega(z)}} + \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}) \qquad (26)$$

Substituting ω and $\frac{\partial \omega(z)}{\partial z}$ for equation (26) using equations (24) and (25), leads to the following signal table of $e(z)$, for $z \geq 0$.

Table 4.3: Signal table of the function $e(z)$ for $0 \leq z < 1.05$

z	0		0.66		0.93		<1.05
$\partial e(z)/\partial z$	0	+	0	-	0	+	<2.14E-09
$e(z)$	0	↗	4.19E-09	↘	-4.23E-09	↗	<1.21E-09

Table 4.3: (Continued for $1.05 \leq z < 1.97$)

z	1.05		1.13		1.36		1.65		1.88		<1.97
$\partial e(z)/\partial z$	5.84E-09	+	0	-	0	+	0	-	0	+	<4.14E-09
$e(z)$	-1.36E-08	↗	3.31E-08	↘	-3.89E-08	↗	2.99E-08	↘	-1.36E-08	↗	<1.88E-08

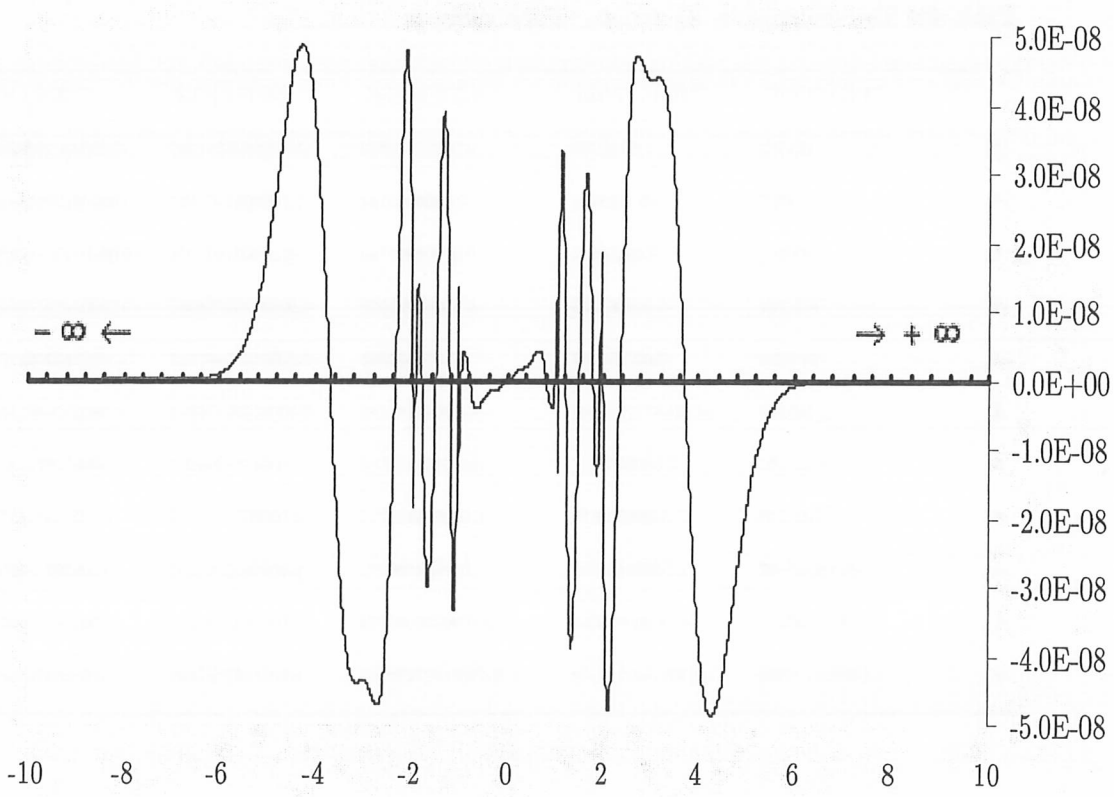
Table 4.3:(Continued for $1.97 \leq z < +\infty$)

z	1.97		2.14		2.7		3		3.14		4.3		$+\infty$
$\partial e(z)/\partial z$	-3.21E-09	-	0	+	0	-	0	+	0	-	0	+	0
$e(z)$	1.87E-08	↘	-1.36E-08	↗	4.72E-08	↘	4.36E-08	↗	4.39E-08	↘	-4.88E-08	↗	0

This table indicates that $|e(z)|< 5\times10^{-8}$ for $z \geq 0$. Similarly, we can derive a signal table of $e(z)$, for $z<0$, so that the maximum absolute error be less than 5×10^{-8}. Furthermore, this maximum absolute error is a global maximum. Therefore, the absolute value of $e(z)$, for each $z \in (-\infty,+\infty)$, is less than 5×10^{-8}. Figure 4.3 provides a plot of $e(z)$ for all z.

Figure 4.3: Error plot of approximation (20) with value $\omega(z)$ given by equation (24)

$(-\infty < z < +\infty)$



To further improve the maximum absolute error in $F(z)$, we introduce an empirically approximate value ω that is a polynomial of degree 10 as

$$\omega = A_1 z^{10} + A_2 |z^9| + A_3 z^8 + A_4 |z^7| + A_5 z^6 + A_6 |z^5| + A_7 z^4 + A_8 |z^3| + A_9 z^2 + A_{10} |z| + A_{11} \tag{27}$$

where the coefficients A_1 to A_{11} are given by Table 4.4.

Table 4.4 The coefficients A_1 to A_{11} of the value ω

	$0 \le \|z\| < 0.05$	$0.05 \le \|z\| < 0.2$	$0.2 \le \|z\| < 0.6$	$0.6 \le \|z\| < 1.3$	$1.3 \le \|z\| < 2$
A_1	-164311	0.137736	0.00000669889	-0.00000008241225	-0.00000018992895
A_2	59955	-0.1715751	-0.0000261114	0.0000015375137	0.000001900488007
A_3	-9319.7	0.09458524	0.00004402746	-0.000007633064	-0.000003559474385
A_4	816.244	-0.030361336	-0.00004721029	0.00001262788065	-0.000024170784487
A_5	-44.6684	0.006280987	0.000034698	-0.00001485887312	0.000124830559738
A_6	1.595165	-0.0008727685948	-0.000014186178	0.00001779174954	-0.00029708023354
A_7	-0.037152	0.0005065838	0.000428431084	0.0004118896607	0.00087182129003
A_8	0.0005748	-0.00000523618	-0.000000913721	0.0000057016368	-0.000437245711
A_9	-0.009569667	-0.00956401041	-0.009564100272	-0.0095659762953	-0.0092927825966
A_{10}	0.0000000287	-0.00000000502	-0.0000000096978	0.0000003187712	-0.0000979290111
A_{11}	0.636619772303	0.63661977241954	0.63661977270567	0.6366197463449	0.636635468821

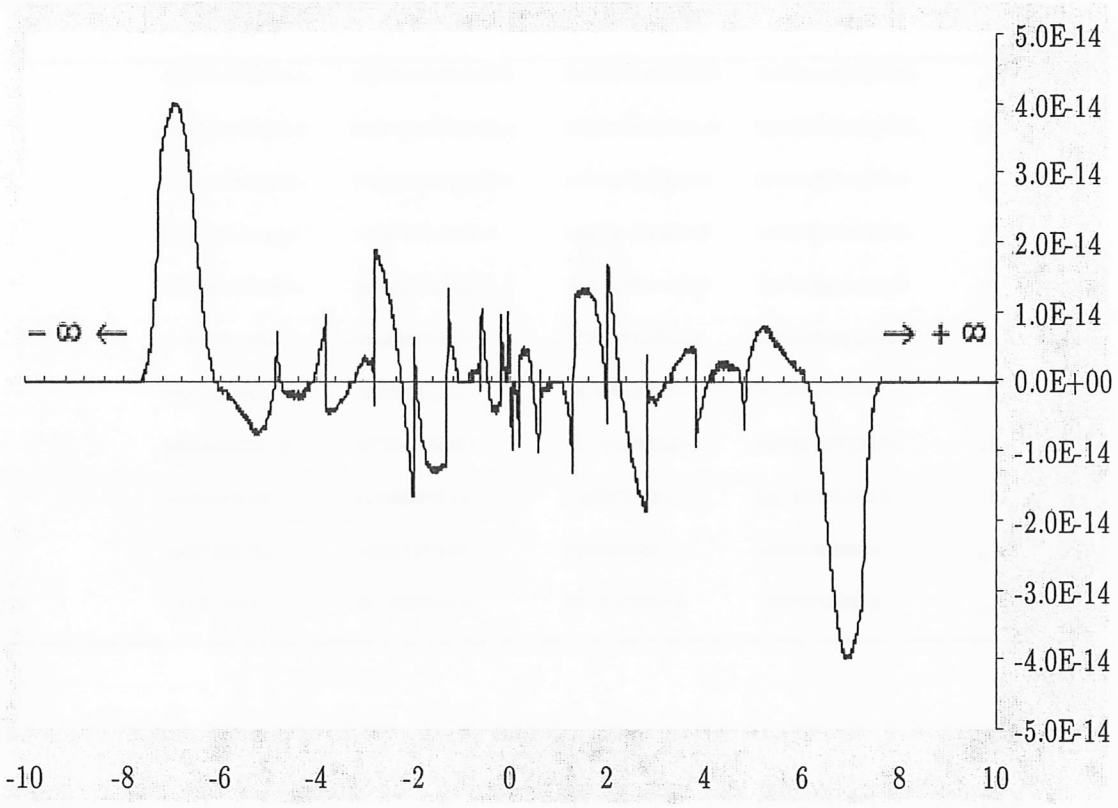
Table 4.4 (*continued*) The coefficients A_1 to A_{11} of the value ω

	$2 \leq \lvert z \rvert < 2.8$	$2.8 \leq \lvert z \rvert < 3.8$	$3.8 \leq \lvert z \rvert < 4.8$	$4.8 \leq \lvert z \rvert < 7.25$	$7.25 \leq \lvert z \rvert$
A_1	0.0000003131231703	-0.000000147576219	-0.0000004078129952	-0.00000476686175	0
A_2	-0.0000070000281146	0.000004891566529	0.0000170276673136	0.00025318775948	0
A_3	0.000066286888765	-0.00007132986144	-0.000319633235187	-0.0060468613807	0
A_4	-0.000342129007215	0.00059684421948	0.00355340826717	0.0855145832474	0
A_5	0.001044265987196	-0.00313107114309	-0.02592585496206	-0.7930288951203	0
A_6	-0.00202605872942	0.0105849777738	0.1298666026436	5.039109528662	0
A_7	0.002921707829	-0.023184927469	-0.453178626455	-22.21941097848	0
A_8	-0.0017726486535	0.0346075158335	1.090715959706	67.13242286399	0
A_9	-0.0091022041083	-0.041487748948	-1.73489396344	-133.004448684	0
A_{10}	0.000208670906	0.016600835847	1.618751155384	156.00174044	0
A_{11}	0.636488360887	0.63300431742	-0.04666376193	-81.6396720135	0.5265

Figure 4.4 shows an error plot of the NDF with absolute error less than 4.02×10^{-14}, replacing the values of A_1 to A_{11} of Table 4.1 back in equation (27) and then in equation (20).

Figure 4.4 Error plot of approximation (20) with the value of ω given by equation (27)

$(-\infty < z < +\infty)$



4-4. The Mean Range of the Normal Distribution

The probability and cumulative density function of the random variables $Z_i \overset{iid}{\sim} N(0,1)$ for $i = 1,2,...,n$ are

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} \qquad -\infty < z < +\infty$$

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-t^2}{2}} dt \qquad (28)$$

Let $R = Z_{(n)} - Z_{(1)}$, where $Z_{(1)}, Z_{(2)}, ..., Z_{(n)}$ are order statistics of the random variables Z_i. The joint probability density functions of $Z_{(1)}$ and $Z_{(n)}$ is given by

$$f_{1,n}(z,y) = \begin{cases} n(n-1)\,[F(y)-F(z)]^{n-2} f(z) f(y), & z < y \\ 0 & z \geq y \end{cases}$$

Therefore, the joint probability density function $Z_{(n)}$ and R would be

$$f_{1,R}(z,r) = n(n-1)\,[F(r+z)-F(z)]^{n-2} f(z) f(r+z) \qquad (29)$$

Using equation (29), the probability density function R is

$$f_R(r) = \int_{-\infty}^{+\infty} n(n-1)\,[F(r+z)-F(z)]^{n-2} f(z) f(r+z)\, dz \qquad r \geq 0$$

Johnson, et al (1994), evaluated the mean range $(E(R))$ as

$$E(R) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} r \times n(n-1)\,[F(r+z)-F(z)]^{n-2} f(z) f(r+z)\, dz\, dr$$

$$= \int_{-\infty}^{+\infty} \left\{1 - \left(F(z)\right)^n - \left(1-F(z)\right)^n\right\} dz \qquad (30)$$

The theoretical evaluations that were presented in section 3-2, are not suitable evaluations for computing equation (30). This is so because the theoretical approximation of the $F(z)$ are based on subsets of the values $0 < z < +\infty$ or $-9 < z < +9$, whereas $E(R)$ is based on the NDF with values $-\infty < z < +\infty$. Furthermore, the numerical evaluations of $F(z)$ in section 3-2, are not appropriate approximations to compute equation (30). The reason is that the existing commands, in statistical software such as MATLAB,

can not calculate this equation; therefore we should use a new computer program. Using equations (20) and (27) and the following commands in MATLAB, the mean range of the normal distribution is simply calculated,

$$F = @(z)\ \ 1 - F(z).\hat{}\,n - \left(1 - F(z)\right).\hat{}\,n\ ;$$

$$d2 = quad\left(F, b1, b2\right);$$

In these commands, $@(z)$ and *quad* are syntaxes to introduce the integral function and integral from $b1$ to $b2$, respectively, and *F(z)* is replaced by using equations (20) and (27).

Appendix I presents the commands used for computing d_2 with various values n. Additionally, Table 4.5 exhibits the mean range of the normal distribution for $n = 2\,(1)100,\quad 120\,(20)\,440$.

Table 4.5 The mean range of the normal distribution (d_2)

n	d_2	n	d_2	n	d_2	n	d_2
2	1.128399388	31	4.112947754	60	4.638575278	89	4.931322805
3	1.692583920	32	4.139357950	61	4.651140881	90	4.939415336
4	2.058770968	33	4.164837580	62	4.663476164	91	4.947407901
5	2.325954220	34	4.189447042	63	4.675588958	92	4.955302837
6	2.534427881	35	4.213241031	64	4.687486981	93	4.963102400
7	2.704374431	36	4.236269347	65	4.699177286	94	4.970808769
8	2.847220813	37	4.258577446	66	4.710666685	95	4.978424053
9	2.970049046	38	4.280206928	67	4.721961665	96	4.985950289
10	3.077530707	39	4.301195955	68	4.733068410	97	4.993389450
11	3.172900475	40	4.321579619	69	4.743992818	98	5.000743442
12	3.258470416	41	4.341390255	70	4.754740520	99	5.008014113
13	3.335996759	42	4.360657725	71	4.765316893	100	5.015203251
14	3.406780785	43	4.379409655	72	4.775727078	120	5.144189562
15	3.471845840	44	4.397671657	73	4.785975990	140	5.251203103
16	3.532003010	45	4.415453297	74	4.796068336	160	5.342459595
17	3.587905461	46	4.432804816	75	4.806008620	180	5.421871644
18	3.640086532	47	4.449732925	76	4.815801160	200	5.492099946
19	3.688987071	48	4.466256863	77	4.825450095	220	5.554982434
20	3.734975440	49	4.482394608	78	4.834959395	240	5.611867501
21	3.778362420	50	4.498162985	79	4.844332872	260	5.663767153
22	3.819412500	51	4.513577764	80	4.853574185	280	5.711459767
23	3.858338013	52	4.528653744	81	4.862686850	300	5.755557481
24	3.895363368	53	4.543404832	82	4.871674247	320	5.796546792
25	3.930645066	54	4.557844108	83	4.880539628	340	5.834824904
26	3.964332149	55	4.571983894	84	4.889286119	360	5.870703433
27	3.996555695	56	4.585835808	85	4.897916733	380	5.904481079
28	4.027431559	57	4.599410817	86	4.906434370	400	5.936372257
29	4.057062620	58	4.612719283	87	4.914841823	420	5.966570264
30	4.085540632	59	4.625771010	88	4.923141788	440	5.995240461

4-5. Conclusions

It has been shown that the new theoretical evaluation for the NDF has greatest absolute error less than 4.02×10^{-14}. In section 3-3 the new numerical evaluation of the NDF was presented with 21 decimal places accuracy. The proposed evaluation is based on the values $-\infty < z < +\infty$ while the existent evaluations are based on subset of the values $0 < z < +\infty$ or $-9 < z < 9$, theoretically, or the values $-\infty < z < +\infty$, numerically. The numerical evaluations have satisfying precision, but they are based on computer programming.

The new evaluation of the NDF with a good approximation is applied for calculating the mean range of the normal distribution with various sample sizes n.

Appendix I

(Matlab)

```
A=[-164311 0.137736 0.00000669889 -0.00000008241225
-0.00000018992895 0.0000003131231703 -0.000000147576219
-0.0000004078129952 -0.00000476686175 0
59955 -0.1715751 -0.0000261114 0.0000015375137
0.000001900488007 -0.0000070000281146 0.000004891566529
0.0000170276673136 0.00025318775948 0
-9319.7 0.09458524 0.00004402746 -0.000007633064
-0.000003559474385 0.000066286888765 -0.00007132986144
 -0.000319633235187 -0.0060468613807 0
816.244 -0.030361336 -0.00004721029 0.00001262788065
-0.000024170784487 -0.000342129007215 0.00059684421948
0.00355340826717 0.0855145832474 0
```

```
-44.6684 0.006280987 0.000034698 -0.00001485887312
0.000124830559738 0.001044265987196 -0.00313107114309
-0.02592585496206 -0.7930288951203 0
1.595165 -0.0008727685948 -0.000014186178 0.00001779174954
-0.00029708023354 -0.00202605872942 0.0105849777738
0.1298666026436 5.039109528662 0
-0.037152 0.0005065838 0.000428431084 0.0004118896607
0.00087182129003 0.002921707829 -0.023184927469
-0.453178626455 -22.21941097848 0
0.0005748 -0.00000523618 -0.000000913721 0.0000057016368
-0.000437245711 -0.0017726486535 0.0346075158335
1.090715959706 67.13242286399 0
-0.009569667 -0.00956401041 -0.009564100272
-0.0095659762953 -0.0092927825966 -0.0091022041083
-0.041487748948 -1.73489396344 -133.004448684 0
0.0000000287 -0.00000000502 -0.0000000096978
0.0000003187712 -0.0000979290111 0.000208670906
 0.016600835847 1.618751155384 156.00174044 0
0.636619772303 0.63661977241954 0.63661977270567
0.6366197463449 0.636635468821 0.636488360887
0.63300431742 -0.04666376193 -81.6396720135 0.5265];
B=[0 0.05 0.2 0.6 1.3 2 2.8 3.8 4.8 7.25
 0.05 0.2 0.6 1.3 2 2.8 3.8 4.8 7.25 9e+24];
z=1;
for n =2:30
    d2(n)=0;
    for j = 1:10
        s=z;
        F= @(z) 1-(0.5*(1-(sign(-s).*sqrt(1-exp(-
    ((A(1,j)*z.^12+A(2,j)*abs(z.^11)+A(3,j)*z.^10+A(4,j)*abs(z.^9)+A(5,j)*z.^8
    +A(6,j)*abs(z.^7)+A(7,j)*z.^6+A(8,j)*abs(z.^5)
    +A(9,j)*z.^4+A(10,j)*abs(z.^3)+A(11,j)*z.^2))))))).^n-(1-(0.5*(1-(sign(-
```

```
    s).*sqrt(1-exp(-
    ((A(1,j)*z.^12+A(2,j)*abs(z.^11)+A(3,j)*z.^10+A(4,j)*abs(z.^9)+A(5,j)*z.^8
    +A(6,j)*abs(z.^7)+A(7,j)*z.^6+
    A(8,j)*abs(z.^5)+A(9,j)*z.^4+A(10,j)*abs(z.^3)+A(11,j)*z.^2)))))).^n;
        d2(n)=d2(n)+quad(F,B(1,j),B(2,j))+ quad(F,-B(2,j),-B(1,j)) ;
    end
d2(n)
end
```

5

Extensions to the Strategy of the Steepest Ascent for Response Surface Methodology

CHAPTER OUTLINE

5-1. Introduction

In the previous chapters, we have suggested some tools of quality control. In what follows, we consider a major tool of quality improvement, which involves the set of activities used to ensure that the products meet requirements and are improved on a continuous basis. Quality improvement is often done on a project-by-project basis and involves teams led by personnel with specialized knowledge of statistical methods and experiments in applying them. For more details about statistical quality control and improvement, the interested reader is referred to Ryan (2000) and Montgomery (2005). To improve the performance of production process, we introduce an extension to response surface methodology (RSM).

RSM is a collection of mathematical and statistical techniques that are useful for the modeling and analysis of problems in which a response (output variable) of interest is influenced by several variables (input variables) and the objective is to optimize this response. The functional relationship between the response and the levels of the k input variables, ξ_1, ξ_2, ..., ξ_k, is given by,

$$y = \vartheta(\xi_1, \xi_2, ..., \xi_k) + \varepsilon ,$$

where, ε denotes the discrepancy between the observed response y and its mean, $\mathrm{E}(y) = \vartheta(\xi_1, \xi_2, ..., \xi_k)$. If the response is well modeled by a linear function of the independent variables, the approximating function is given by the first-order model,

$$y = \beta_1^o \xi_1 + \beta_2^o \xi_2 + ... + \beta_k^o \xi_k + \varepsilon^o . \qquad (1)$$

Otherwise, in many situations the second-order model is used,

$$y = \sum_{i=1}^{k} \beta_i^o \xi_i + \sum_{i=1}^{k} \beta_{ii}^o \xi_i^2 + \sum_{\substack{i=1,\, j=1 \\ i<j}}^{k} \beta_{ij}^o \xi_i \xi_j + \varepsilon^o \ . \qquad (2)$$

If the studied system were being investigated for the first time, starting conditions over operability region would often not be very close to an optimum. As a result, the use of model (2) frequently is not adequate under the starting conditions. In these circumstances, the objective of the experimenter is to move efficiently along a path of improvement toward the vicinity of optimum. We usually take as the path of improvement the line through the center of the interesting region to the fitted surface, where the direction of improvement is parallel to the coefficients vector of model (1).

It will usually be found that, after one or two applications of the steepest ascent, first-order effects will no longer dominate and the first-order approximation will be inadequate. At this stage, further the steepest ascent progress will not be possible and more sophisticated second-order methods should be applied. The vast literature of the RSM and the applications of second-order methods to accomplish optimum response are given by Khuri and Cornell (1996), Montgomery (2001), Myers and Montgomery (2002), Myers et al. (2002), Box and Draper (2007). To illustrate the method of the steepest ascent, examples corresponding to real data are given by Chang et al. (2006), Liu et al. (2007), Sun et al. (2007).

The goal of this chapter is to further develop the method of the steepest ascent. Attempts by researchers to use the steepest ascent process typically lead to difficult computations. Little effort has been made to employ this process, straightforwardly. The components of the steepest ascent path, which are based on a first-order model, are presented in section 5-2. Section 5-3 gives the path of the steepest ascent under some restriction in the direction of improvement.

Other important aspects are the developments of the confidence region for the direction of the steepest ascent, and the confidence cone about the estimated path, are presented in section 5-4.

5-2. Path of the Steepest Ascent

The steps along the path of the steepest ascent are proportional to the regression coefficients $\{\beta_j\}$, for $j=1,2,...,k$, of the first-order model,

$$y=\sum_{j=1}^{k}\beta_j x_j+\varepsilon. \tag{3}$$

In this model x_j for $j=1,2,...,k$ is the coded variable of the original ξ_j in (1), such that $x_j=(\xi_j-a_j)/b_j$ with

$$a_j=\left[\max(\xi_j)+\min(\xi_j)\right]/2;$$

$$b_j=\left[\max(\xi_j)-\min(\xi_j)\right]/2.$$

The coded variables are used to manipulate the model mathematically and the natural variables in running the process. The relation between β_j^o in the model (1) and β_j in the model (3) is $\beta_j^o=\beta_j/b_j$, for $j=1,2,...,k$.

5-2.1. True Path

For the coded variables in model (3), the directional vector to attain the optimum point is a vector parallel to $\underline{a}=(\beta_1,\beta_2,...,\beta_k)$ that is called true path of the steepest ascent. The unit length vector of $\underline{a}$ is

$$\underline{a}_U=\frac{\beta_1}{|\underline{a}|}\underline{i}_1+...+\frac{\beta_j}{|\underline{a}|}\underline{i}_j+...+\frac{\beta_k}{|\underline{a}|}\underline{i}_k,$$

where $\underline{i}_j$, for all j, is the k-dimension unit vector on the jth axis, and the $|\underline{a}|$ is the length of $\underline{a}$. Let $\underline{a}$ makes angles α_1, α_2, ... and α_k with the positive x_1, x_2, ..., and x_k axes. From the inner product $\underline{i}_j.\underline{a}=|\underline{i}_j||\underline{a}|\cos(\alpha_j)$, we have the direction cosines of $\underline{a}$ i.e. $\cos(\alpha_j)=\beta_j/|\underline{a}|$. Therefore, the cosine form of $\underline{a}_U$ will be

$$\underline{a}_U=\cos(\alpha_1)\underline{i}_1+\cos(\alpha_2)\,\underline{i}_2+...+\cos(\alpha_k)\,\underline{i}_k.$$

5-2.2. Estimated Path of the Steepest Ascent

When the parameters β_1, β_2, ..., β_k are unknown, they are estimated by the generalized least squares (GLS) or the ordinary least squares (OLS) estimators $\hat{\beta}_1$, $\hat{\beta}_2$, ..., $\hat{\beta}_k$. Then we get the estimated path of the steepest ascent,

$$\hat{\underline{a}}=(\hat{\beta}_1,\hat{\beta}_2,...,\hat{\beta}_k).$$

The first-order model (3) for N observations is exhibited by,

$$\underline{y}=X\underline{\beta}+\underline{\varepsilon}, \tag{4}$$

where, $\underline{y}$ indicates the $N\times1$ vector of responses, X the $N\times k$ matrix of independent variables, $\underline{\beta}$ the $k\times1$ vector of regression coefficients, $\underline{\varepsilon}$ the $N\times1$ vector of the random errors ($\mathrm{E}(\underline{\varepsilon})=0$ and $\mathrm{var}(\underline{\varepsilon})=\sigma^2V_N$). The GLS and OLS estimator of $\underline{\beta}$ (or equivalently $\underline{a}$) are, respectively,

$$\hat{\underline{\beta}}=(X'V^{-1}X)^{-1}X'V^{-1}\underline{y} \qquad (\text{for } \mathrm{var}(\underline{\varepsilon})=\sigma^2V_N), \tag{5}$$

$$\hat{\underline{\beta}}=(X'X)^{-1}X'\underline{y} \qquad (\text{for } \mathrm{var}(\underline{\varepsilon})=\sigma^2I_N), \tag{6}$$

(see Myers et al. (2002)).

Under the condition that the steepest ascent path has not been estimated precisely enough, the design matrix often requires to be orthogonalized by the data augmentation method or the Gram-Schmidt scheme. In the next chapter the orthogonalization of the design matrix will be discussed.

5-2.3. Origin Point of Path

Many authors, (see, for instance, Myers and Montgomery (2002)), argue that in the metric of the coded design variables the origin point of path is the design center i.e. $(0,0,...,0)$. In the sequel, a description is presented on how to assign the origin point.

Because of $\text{var}(\underline{\varepsilon}) = \sigma^2 V$, V must be $N \times N$ nonsingular and positive definite matrix. Additionally, there is a $N \times N$ nonsingular symmetric matrix A such that $V = A'A = AA$. For a point as $\underline{w} = (x_1, x_2, ..., x_k)'$ we have

$$\text{var}(\hat{y} \mid \underline{w}) = \underline{w}' \, \text{var}(\underline{\hat{\beta}}) \underline{w} = \underline{w}' (X' V^{-1} X)^{-1} \underline{w} \sigma^2 . \tag{7}$$

Let $A^{-1}X = Z$, then equation (7) will be

$$\text{var}(\hat{y} \mid \underline{w}) = \underline{w}' (Z'Z)^{-1} \underline{w} \sigma^2 . \tag{8}$$

The first and second order derivation of (8) are

$$\nabla^{(1)}_{\underline{w}} [\text{var}(\hat{y} \mid \underline{w})] = 2\underline{w}' (Z'Z)^{-1} \sigma^2 , \tag{9}$$

$$\nabla^{(2)}_{\underline{w}} [\text{var}(\hat{y} \mid \underline{w})] = 2(Z'Z)^{-1} \sigma^2 .$$

Since $Z'Z$ is a positive definite matrix in $\nabla^{(2)}_{\underline{w}} [\text{var}(\hat{y} \mid \underline{w})]$, $\text{var}(\hat{y} \mid \underline{w})$ is a concave function of $\underline{w}$. Based on equation (9), the $\text{var}(\hat{y} \mid \underline{w})$ has a global minimum in

point $\underline{w} = \underline{0}$. As a result, we can derive this point as the origin point on the path of the steepest ascent.

5-2.4. Points along the Path

In this subsection, the coordinates of a point on the path of the steepest ascent are defined in a manner different from that considered by Myers and Montgomery (2002) and Box and Draper (2007). As already mentioned, the point $\underline{w} = \underline{0}$ is the origin point. To determine other points along the path, the actual step size typically is given by the experimenter based on the process knowledge or other practical considerations. Let us define the step size by λ, for $0 < \lambda \leq 1$. The coordinates of the lth point, $\underline{w}_l$, along the path will be

$$\underline{w}_l = \left(\frac{l\beta_1}{\zeta}, \frac{l\beta_2}{\zeta}, ..., \frac{l\beta_k}{\zeta}\right) \quad \text{for} \quad l = 0,1,2,... \tag{10}$$

where, $\zeta = \max|\beta_j| / (\lambda |\underline{\hat{a}}|)$ for $j = 1,2,...,k$.

Experiments are conducted along the steepest ascent path until no further increase (decrease) in response is observed. Then a new first-order model may be fitted, a new path of the steepest ascent determined, and the procedure continues. Eventually, the experimenter will arrive in the vicinity of the optimum. This is usually indicated by lack of fit of a first-order model. At that time additional experiments are collected to obtain a more precise estimate of the optimum using the second-order model.

5-3. Constrained Path of the Steepest Ascent

Sometimes in experimental designs, we may have some constraint in the direction of the steepest ascent thus we cannot move forward very far in this direction. Under this limitation of the path, a researcher can proceed until

initial direction of the steepest ascent hits the constraining plane as P_c and then the new direction takes over. Box and Draper (2007) suggest the problem of the steepest ascent direction subject to the condition that the constraint is violated. We proceed by describing how, in more details, to attain the constrained path.

Figure 5.1 shows the initial and the new origin points O and O', the initial and the constrained steepest ascent path $\underline{a}$ and $\underline{a}_c$, and the constraining plane P_c. Denote

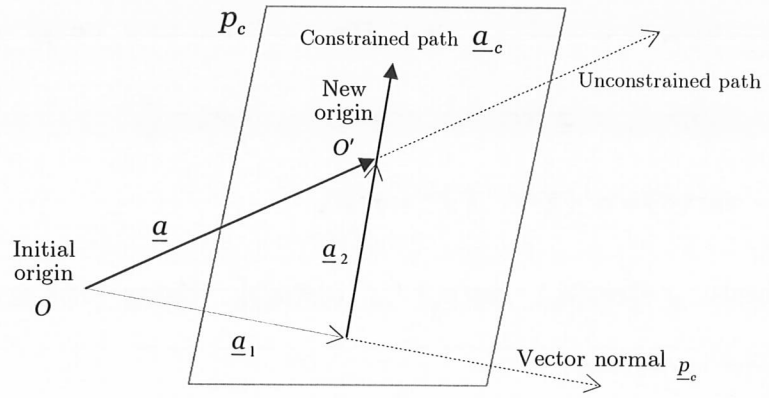
$\underline{a}_c = (a_{c1}, a_{c2}, ..., a_{ck})$,

P_c: $p_1 x_1 + p_2 x_2 + ... + p_k x_k = p_0$,

where, the vector normal of plane P_c is $\underline{p}_c = (p_1, p_2, ..., p_k)$, and for a given point $(x_{01}, x_{02}, ..., x_{0k})$ of P_c we have $p_0 = p_1 x_{01} + p_2 x_{02} + ... + p_k x_{0k}$. If the vector normal $\underline{p}_c$ is a nonzero vector, resolve the initial path of the steepest ascent, $\underline{a}$, into components $\underline{a}_1$ and $\underline{a}_2$, parallel and perpendicular vectors to $\underline{p}_c$, respectively. In the case, $\underline{a} = \underline{a}_1 + \underline{a}_2$; for $h > 0$, $\underline{a}_1 = h\underline{p}_c$; and the inner product $<\underline{a}_2, \underline{p}_c>$ is equal zero. Hence, $\underline{a}_2 = \underline{a} - \underline{a}_1 = \underline{a} - h\underline{p}_c$ and $<\underline{a}_2, \underline{p}_c> = <(\underline{a} - h\underline{p}_c), \underline{p}_c>$ $= <\underline{a}, \underline{p}_c> - h|\underline{p}_c|^2$. Since $<\underline{a}_2, \underline{p}_c> = 0$, $h = <\underline{a}, \underline{p}_c> / |\underline{p}_c|^2$. As a result,

$$\underline{a}_2 = \underline{a} - \frac{<\underline{a}, \underline{p}_c>}{|\underline{p}_c|^2} \underline{p}_c . \qquad (11)$$

Figure 5.1 The initial path of the steepest ascent and its modified path to a constraint



Using equation (11), the constrained unit length path will be

$$\underline{a}_{CU} = \frac{\underline{a}_2}{|\underline{a}_2|} = \left(\underline{a} - \frac{<\underline{a}, \underline{p}_c>}{|\underline{p}_c|^2} \underline{p}_c \right) \Big/ \left| \underline{a} - \frac{<\underline{a}, \underline{p}_c>}{|\underline{p}_c|^2} \underline{p}_c \right| = (a_{CU_1}, a_{CU_2}, ..., a_{CU_k}).$$

The coordinates of the new origin point O' is

$$O' = (\tau_0\beta_1, \tau_0\beta_2, ..., \tau_0\beta_k).$$

where, $\tau_0 = p_0 / <\underline{p}_c, \underline{a}>$, since each point on the vector $\underline{a}$ has the general coordinate $(\tau\beta_1, \tau\beta_2, ..., \tau\beta_k)$. Furthermore, the required point O' while it is on the initial path it also lies the constraining plane P_c. The specific value τ in the new origin point O' is called τ_0.

To acquire the coordinates of points on the constrained path, let us denote the step size by λ'. We define the coordinates of lth point along the $\underline{a}_{CU}$ by

$$\underline{u}_l = (\frac{la_{CU_1}}{\zeta'} + \tau_0\beta_1, \frac{la_{CU_2}}{\zeta'} + \tau_0\beta_2, ..., \frac{la_{CU_k}}{\zeta'} + \tau_0\beta_k) \qquad \text{for} \qquad l = 0,1,2,...,$$

where, $\zeta' = \max|a_{CU_j}|/\lambda'$ for $j = 1,2,...,k$.

5-4. Confidence Region for the Direction of the Steepest Ascent

According to model (4), under the condition that $\text{var}(\underline{\varepsilon}) = \sigma^2 V$, we have

$$\text{E}(\underline{\hat{\beta}}) = \underline{a}_{k\times 1} = |\underline{a}|\left[cos(\alpha_1) \quad cos(\alpha_2) \quad ... \quad cos(\alpha_k)\right]' \tag{12}$$

$$\text{var}(\underline{\hat{\beta}}) = \sigma^2(X'V^{-1}X)^{-1} = \sigma^2 W_k \tag{13}$$

Denote $\underline{\phi} = \left[cos(\alpha_1) \quad cos(\alpha_2) \quad ... \quad cos(\alpha_k)\right]'$. Using (12) and (13), the following model is derived

$$\underline{\hat{a}}_{k\times 1} = |\underline{a}|\underline{\phi} + \underline{\gamma}. \tag{14}$$

where, $\underline{\hat{a}}_{k\times 1}$ indicates the responses vector, $|\underline{a}|$ the parameter of model, $\underline{\phi}$ the regressors vector and $\underline{\gamma}$ the error vector ($\text{E}(\underline{\gamma}) = \underline{0}$ and $\text{var}(\underline{\gamma}) = W_k\sigma^2$). Corresponding to (14), the GLS estimator of $|\underline{a}|$ is

$$\widehat{|\underline{a}|} = (\underline{\phi}'W^{-1}\underline{\phi})^{-1}\underline{\phi}'W^{-1}\underline{\hat{a}}. \tag{15}$$

Under the normality assumption,

$$[\underline{\hat{a}}'W^{-1}\underline{\hat{a}} - (\widehat{|\underline{a}|}\underline{\phi})'W^{-1}\underline{\hat{a}}]/\sigma^2 \sim \chi^2(k-1), \tag{16}$$

$$[\underline{y}'V^{-1}\underline{y} - (X\underline{\hat{\beta}})'V^{-1}\underline{y}]/\sigma^2 \sim \chi^2(N-k). \tag{17}$$

Introducing equation (15) into (16), and (5) into (17), we have

$$\frac{[\underline{\hat{a}}'W^{-1}\underline{\hat{a}} - (\underline{\phi}'W^{-1}\underline{\phi})^{-1}(\underline{\phi}'W^{-1}\underline{\hat{a}})(\underline{\phi}'W^{-1}\underline{\hat{a}})]/(k-1)}{[\underline{y}'V^{-1}\underline{y} - (\underline{y}'V^{-1}X)(X'V^{-1}X)^{-1}(X'V^{-1}\underline{y})]/(N-k)} \sim F(k-1, N-k).$$

where, all the quantities are known except for the $\underline{\phi}$. Hence, a 100(1-α)% confidence region for the components of $\underline{\phi}$ i.e. $\cos(\alpha_j)$ is

$$\frac{(N-k)[\underline{\hat{a}}'W^{-1}\underline{\hat{a}}-(\underline{\phi}'W^{-1}\underline{\phi})^{-1}(\underline{\phi}'W^{-1}\underline{\hat{a}})(\underline{\phi}'W^{-1}\underline{\hat{a}})]}{(k-1)[\underline{y}'V^{-1}\underline{y}-(\underline{y}'V^{-1}X)(X'V^{-1}X)^{-1}(X'V^{-1}\underline{y})]} \le F_{\alpha,(k-1,N-k)}. \tag{18}$$

When $\mathrm{var}(\underline{\varepsilon})=\sigma^2 I$ and $X'X=I_k$ from (4), according to (18) the confidence region 100(1-α)% for the direction of the steepest ascent is

$$\sum_{j=1}^{k}\hat{\beta}_j^2-(k-1)MS_E F_{\alpha,(k-1,\,N-k)} \le [\sum_{j=1}^{k}\hat{\beta}_j\cos(\alpha_j)]^2, \tag{19}$$

where, MS_E indicates the estimator of σ^2. Thus the confidence region is a hypercone with apex at the origin and all points $(\cos(\alpha_1),\cos(\alpha_2),...,\cos(\alpha_k))$ have a unit distance away from the origin.

Associated with the confidence hypercone, define the limited semiplane angle between a line on the surface of the hypercone and its axis by θ. Substitute $\cos^2(\theta)=[\sum_1^k\hat{\beta}_j\cos(\alpha_j)]^2/\sum_1^k\hat{\beta}_j^2$, ($\sum_1^k\cos^2(\alpha_j)=1$), into (19), then the angle θ is evaluated by,

$$\theta=\arcsin[\frac{(k-1)MS_E F_{\alpha,(k-1,\,N-k)}}{\sum_{j=1}^{k}\hat{\beta}_j^2}]^{1/2}, \tag{20}$$

where, all quantities on the right-hand side in (20) are known.

Based on the angle θ, a fractional solid angle φ to determine what percentage of the possible directions was excluded by the 100(1-α)% confidence region (19) is given by Sztendur and Diamond (2002) and Box and Draper (2007). In what follows an extension of the current method to evaluate φ is presented.

In general, the surface area of a k-dimensional sphere of radius R, S_{sphere}, and the surface area of a k-dimensional cap of the sphere interior to the hypercone, S_{cap}, are

$$S_{sphere}(k,R)=\frac{2\pi^{k/2}R^{k-1}}{\Gamma(k/2)}, \tag{21}$$

$$S_{cap}=\int_{b}^{R} S_{sphere}(k-1,\sqrt{R^2-x_i^2})\times\sqrt{1+\left[\frac{\partial\left(\sqrt{R^2-x_i^2}\right)}{\partial x_i}\right]^2}\,dx_i\,. \tag{22}$$

where, x_i is a design variable and x_i-axis is revolving axis to obtain the surface area of the cap. The lower and upper limits of integration in (22) are $\cos(\theta)$ and 1.

To evaluate the fractional φ, let $u^2/(k-1)=x_i^2/(1-x_i^2)$ in equation (22). Using (21) and (22) we have

$$\varphi=\frac{1}{2}\int_{\sqrt{k-1}\cot(\theta)}^{+\infty}\frac{\Gamma(k/2)du}{\Gamma[(k-1)/2]\sqrt{(k-1)\pi}(1+u^2/(k-1))^{k/2}}\,. \tag{23}$$

Letting $u=1-x_i^2$ in (22), an alternative to the evaluation of φ is obtained as,

$$\varphi=\frac{\Gamma(1/2)}{4\sqrt{\pi}}\int_{0}^{\sin^2(\theta)}\frac{u^{(k-1)/2-1}(1-u)^{1/2-1}}{Beta(\frac{k-1}{2},\frac{1}{2})}du\,. \tag{24}$$

The right-hand side in (23) and (24) can be computed from the corresponding tables of t and $Beta$ distributions. As a consequence, according to the fractional φ, the appropriate direction is known with some considerable accuracy. The 100(1-α)% confidence region excludes about (1-φ)% of possible directions of advance.

5-5 An Example

A chemical engineer is interested in determining the operating conditions that maximize the yield of a process. Two controllable variables that influence the process yield are factor A (temperature) and factor B (time). She is currently operating the process with factor A of 210 deg F and factor B of 55 minutes, which result in yields of around 21 percent. Because it is unlikely that this region contains the optimum, she fits a first-order model and applies the method of the steepest ascent. In this case, the region of exploration for fitting the first-order model is (200, 220) deg F of factor A and (50, 60) minutes of factor B. If ξ_1 and ξ_2 denote the natural variables temperature and time, respectively, then the coded variables are $x_1 = (\xi_1 - 210)/10$ and $x_2 = (\xi_2 - 55)/5$.

The design used to fit first-order model is a 2^2 factorial augmented by five center points. Here, the center points are used for estimating the experimental error and for checking the adequacy of the first-order model. The experimental design is given in Table 5.1.

Table 5.1: Process Data for Fitting the First-Order Model (Simulation data).

ξ_1	ξ_2	x_1	x_2	y
200	50	-1	-1	20.1
220	50	1	-1	18.3
200	60	-1	1	25.2
220	60	1	1	23.0
210	55	0	0	20.3
210	55	0	0	21.6
210	55	0	0	21.0
210	55	0	0	22.2
210	55	0	0	20.2

Following the methods for two-level factorial designs, a first-order model with coded variables is fitted to the data of Table 5.1 by least squares

$$\hat{y} = 21.32 - 0.97x_1 + 2.48x_2 . \tag{25}$$

The first-order model with natural variables is

$$\hat{y} = 20.2 - 0.057\xi_1 + 0.246\xi_2 .$$

To check the inadequacy of the following second-order model, Table 5.2 is constructed,

$$y = \beta_0 + \sum_{i=1}^{2} \beta_i x_i + \sum_{i=1}^{2} \beta_{ii} x_i^2 + \beta_{12} x_1 x_2 + \varepsilon .$$

The analysis of variance in Table 5.2 indicates that the quadratic and interaction effects can be ignored.

Table 5.2 Analysis of Variance for the Second-Order Model

Source	Sum of Squares	d.f	Mean Square	F Value	p-value	
Model	28.368	3	9.456	14.962	0.0122	significant
A	3.803	1	3.803	6.017	0.0702	
B	24.503	1	24.503	38.770	0.0034	
AB	0.063	1	0.063	0.099	0.7689	not significant
Curvature	0.660	1	0.660	1.044	0.3646	not significant
Pure Error	2.528	4	0.632			
Total	31.556	8				

Therefore, the engineer constructs Table 5.3 to assign the significant effects in the first-order model.

Table 5.3 Analysis of Variance for the First-Order Model

Source	Sum of Squares	d.f	Mean Square	F Value	p-value	
Model	28.305	2	14.1525	26.12323	0.0011	significant
A	3.8025	1	3.8025	7.0188	0.0381	
B	24.5025	1	24.5025	45.22765	0.0005	
Residual	3.250556	6	0.541759			
Lack of Fit	0.722556	2	0.361278	0.571642	0.6048	not significant
Pure Error	2.528	4	0.632			
Total	31.55556	8				

This table indicates that the first-order model (25) is adequate. At this stage, the engineer computes points along the path $\underline{a} = (-0.97, 2.48)$, and observes the yields at these points until a decrease in response to be found.

The engineer decides to use a step size in the coded variable x_2 of 1 unit, i.e. $\lambda = 1$. This is equivalent to 5 minutes of reaction time as a step size in the natural variable ξ_2. Equation $\zeta = \max|\beta_j|/(\lambda|\underline{a}|)$ gives $\zeta = 0.931$. As a result, the coordinates of lth point, $\underline{w}_l$, on the estimated path will be

$$\underline{w}_l = \left(\frac{l\hat{\beta}_1}{\zeta|\underline{\hat{a}}|}, \frac{l\hat{\beta}_2}{\zeta|\underline{\hat{a}}|}\right) = \left(\frac{-0.97l}{0.931\times 2.663}, \frac{2.48l}{0.931\times 2.663}\right).$$

Table 5.4 shows the results in coded and natural variables with origin point (0,0).

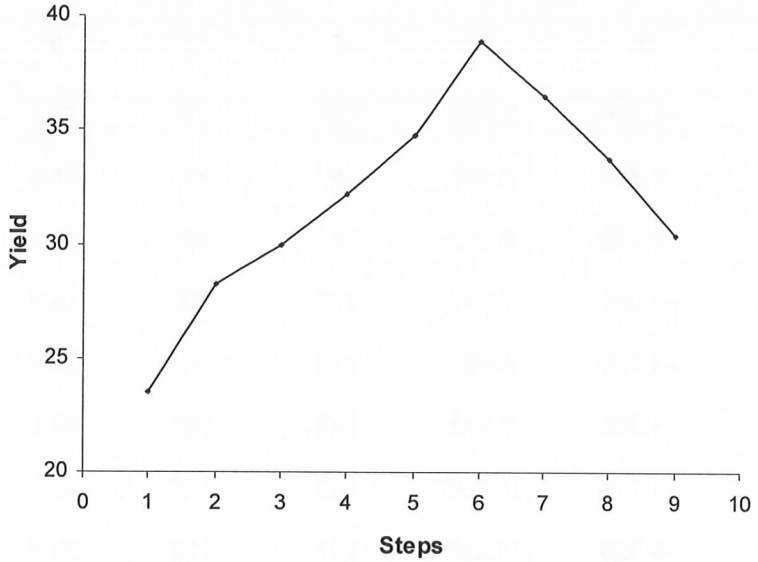
Table 5.4: The Steepest Ascent Experiments

Steps	x_1	x_2	ξ_1	ξ_2	y
Origin	0	0	210	55	-
1	-0.391	1	206	60	23.7
2	-0.782	2	202	65	28.2
3	-1.173	3	198	70	30.0
4	-1.565	4	194	75	32.2
5	-1.956	5	190	80	34.8
6	-2.347	6	187	85	**38.9**
7	-2.738	7	183	90	36.5
8	-3.129	8	179	95	33.6
9	-3.520	9	175	100	30.3

Based on the responses assembled in Table 5.4, Figure 5.2 is drawn. It shows the yield at each step along the path of the steepest ascent. Increases in response are observed through the sixth step; whereas, all steps beyond this

point result in a decrease in yield. Thus, a new first-order model should be fit in the general vicinity of the point (ξ_1, ξ_2)=(187, 85). The region of exploration of this first-order model is (177, 197) for ξ_1 and (80, 90) for ξ_2.

Figure 5.2: Yield versus steps along the path of the steepest ascent



To illustrate the constrained path of the steepest ascent, let us assume that we have the constraint $\xi_1 + \xi_2 \leq 270$ or, equivalently, $10x_1 + 5x_2 \leq 5$. Hence, the constraining plane is $10x_1 + 5x_2 = 5$, and its vector normal $\underline{p}_c = (10,5)$. As previously suggested, the initial path is $\hat{\underline{a}} = \hat{\underline{\beta}} = (-0.97, 2.48)$, thus the constrained unit length path will be $\hat{\underline{a}}_{CU} = (-0.4472, 0.8944)$. The value of p_0 on the constraining plane equals 5, so that $\tau_0 = 1.852$ and $O' = (-1.796, 4.593)$.

In these circumstances, the engineer uses a step size in the coded variable x_2 of 1 unit, on the constrained path. Equation $\zeta' = \max|a_{UC_j}|/\lambda'$ gives $\zeta' = 0.8944$, and $\underline{u}_l = (\frac{la_{CU_1}}{\zeta'} + \tau_0\hat{\beta}_1, \frac{la_{CU_2}}{\zeta'} + \tau_0\hat{\beta}_2)$ makes the coordinates of lth point

on the constrained path, $\underline{u}_l = (-0.5l - 1.796, l + 4.593)$. Responses along the constrained direction of the steepest ascent are reported in Table 5.5.

Table 5.5: The Steepest Ascent Experiments on the Constrained Path

Steps	x_1	x_2	ξ_1	ξ_2	y
Origin	-1.796	4.593	192	78	33.2
1	-2.296	5.593	187	83	34.4
2	-2.796	6.593	182	88	34.9
3	-3.296	7.593	177	93	35.6
4	-3.796	8.593	172	98	**36.7**
5	-4.296	9.593	167	103	36.1
6	-4.796	10.593	162	108	34.0
7	-5.296	11.593	157	113	33.7
8	-5.796	12.593	152	118	33.2
9	-6.296	13.593	147	123	31.23

Table 5.5 shows, increases in response through the fourth step. Steps beyond this point give a decrease in product. In this stage, a new first-order model should be fitted in the general vicinity of the point (ξ_1, ξ_2)=(172, 98). If the first-order effects will no longer dominate, then a second-order model should be used to accomplish optimum response.

In this example, using equation (20), the limited semiplane angle θ for $F_{0.1}(1, 6) = 3.776$ is equal to $\theta = 35^o$. Applying equation (23) or (24) the fractional solid angle φ with $\theta = 35$ equals to $\varphi = 0.0985$. Consequently, the 90%

confidence cone of path excludes about 90.15% of possible directions of the steepest ascent.

5-6. Conclusion

In this chapter, the objective of the experimenter is to move efficiently to the general vicinity of the optimum response. The method of the steepest ascent is used to move sequentially in the direction of the maximum increase in the response. The researchers have a real need to understand the geometry of this method. To this aim, some extensions to the steepest ascent process have been presented. We have proposed an alternative method for the determination of the origin point and other points along the unconstrained and constrained path. Finally, evaluations of the confidence region and confidence cone were provided for the direction of the steepest ascent.

6

The Augmentation of Existing Data for Improving the Path of Steepest Ascent

CHAPTER OUTLINE

6-1. Introduction

The main goal of this chapter is to improve the estimated path of steepest ascent in response surface problems, considered in chapter 5, to obtain the optimum response. Towards this purpose, the data augmentation procedure is used, in a first-order linear model including k factors. In particular, an extension to the optimal data augmentation method and its application in the steepest ascent methodology are presented. Without removing the original data, a few additional observations are selected by the proposed method to decrease the determinant of $(X'X)^{-1}$. The off-diagonal elements of $(X'X)^{-1}$ are equated to zero by these new observations. A quantity is also suggested to assess how well the experimental design is orthogonalized by the method of data augmentation.

Let us denote the input variables in the model by ξ_1, ξ_2, ..., ξ_k. Then, the functional relationship between the actual observed response y and the levels of the k inputs is given by $y = \vartheta(\xi_1,\xi_2,...,\xi_k)+\varepsilon$. Here, ε denotes the discrepancy between the observed response and its mean η given by

$$\eta = \vartheta(\xi_1,\xi_2,...,\xi_k).$$

If there is no curvature in the mean response (η), then it may be related to one input variable, by a response line, two input variables by a response plane or more than two input variables by a response hyperplane. Under this condition, the approximating function of η for k independent variables is denoted by the first-order model

$$\eta = \beta_0^o + \beta_1^o\xi_1 + \beta_2^o\xi_2 + ... + \beta_k^o\xi_k. \qquad (1)$$

If, on the other hand, there is curvature, in many situations a second-order model is used,

$$\eta = \beta_0^o + \sum_{i=1}^{k} \beta_i^o \xi_i + \sum_{i=1}^{k} \beta_{ii}^o \xi_i^2 + \sum_{\substack{i=1, j=1 \\ i<j}}^{k} \beta_{ij}^o \xi_i \xi_j \ ,$$

where, this model is flexible to exploit and philosophy of using them is based on a Taylor series analogy. Applying the second-order model with one, two or more input variables we can draw the response curve $\eta = \vartheta(\xi_1)$ in a two-dimensional plane, the response surface $\eta = \vartheta(\xi_1, \xi_2)$ in a three-dimensional space, or the response hypersurface $\eta = \vartheta(\xi_1, \xi_2, ..., \xi_k)$ in a k-dimensional hyperspace.

Let the independent variables ξ_i, $L_i \le \xi_i \le U_i$ for $i = 1,2,...,k$ be transformed to coded variables $x_i = [\xi_i - (L_i + U_i)/2]/[(U_i - L_i)/2]$. Then, for $k > 3$, the experimental region is a hypercube centered at the origin and bounded by the hyperplanes that orthogonally intersect the axes at -1 and 1. Without loss of generality, let us omit the coefficient β_0^o from model (1). Then, the first-order model of the actual response vector is

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} \ , \tag{2}$$

where, $\underline{Y}$ is the $N \times 1$ vector of responses, X is the $N \times k$ matrix of coded variables with linearly independent columns, $N \ge k$, $\underline{\beta} = (\beta_1, \beta_2, ..., \beta_k)$ is the $k \times 1$ vector of model coefficients and $\underline{\varepsilon}$ is the $N \times 1$ vector of errors ($\underline{\varepsilon} \sim N(\underline{0}, \sigma^2 I)$). The relation between β_j^o in model (1) and β_j in model (2) is $\beta_j^o = \beta_j / [(U_i - L_i)/2]$, for $j = 1,2,...,k$. In model (2), the ordinary least squares (OLS) estimator of $\underline{\beta}$ and its variance are

$$\underline{\hat{\beta}} = (X'X)^{-1} X'\underline{Y} \ , \tag{3}$$

$$\mathrm{var}(\underline{\hat{\beta}}) = \sigma^2 (X'X)^{-1} \ , \tag{4}$$

(see Myers et al. 2002). For reducing the variance and the bias of the estimated regression coefficients, we consider the technique of data augmentation. The OLS estimator is the most preferable since it provides an estimator with minimum variance in the category of unbiased estimators. However, in practice the hypothesized design matrix X is only a submatrix of the true design matrix. Hence, the corresponding OLS estimator of the parameter vector of the model is not accurate and a method is required to improve the design matrix, as much as possible, with a moderate number of new observations.

6-2. Steepest Ascent Method

When we are at a point on the response surface that is remote from the optimum, such as the current operating conditions, there is little curvature in the system and the first-order model will be appropriate. In these circumstances, the path of steepest ascent (descent) is often used to move sequentially in the direction of maximum increase (minimum decrease) in the response. We usually take as the steepest ascent path the line through the center of the experimental region and parallel to the normal vector to the fitted response surface. Therefore, the steps along the path are proportional to the components of $\underline{\beta}$ in model (2). In this chapter, we focus on the improvement of the estimated path of steepest ascent. For more details on the response surface methodology (RSM) and the steepest ascent process we refer the interested reader to the two books by Montgomery (2001), Myers and Montgomery (2002).

The data augmentation strategy improves the estimation of the steepest ascent path in situations where the path has not been estimated with adequate precision by using the original design. Applying this method, the column vectors

of the design matrix X are made orthogonal or nearly orthogonal, and eventually both the bias and the magnitude of the covariance matrix of the estimated path are decreased, significantly.

6-3. Optimality Criterion

Our purpose is to reduce the variance and the bias of the estimated regression coefficients, $\underline{\hat{\beta}}$. To attain these objectives, we utilize the most used optimality criterion for selecting a set of added points in the data augmentation schemes. This amounts to selecting the set of added points that would render a minimum value for the determinant of $(X'X)^{-1}$. It will be shown that the determinant $|(X'X)^{-1}|$ is minimized if the off-diagonal entries of $(X'X)^{-1}$, by adding new observations, are set equal to zero. Obviously, minimizing $|(X'X)^{-1}|$ leads to the minimization of $\text{var}(\underline{\hat{\beta}})$, in equation (4). On the basis of simulation experiments, such a minimization appears to lead to the minimization of the bias of $\underline{\hat{\beta}}$, in (3), as well. The minimization of $|(X'X)^{-1}|$ for augmenting a non-orthogonal design has been advocated by Evans (1979), Sztendur and Diamond (2002), and several other authors.

Brooks and Mickey (1961) show that the minimum number of observations required for estimating a gradient direction in the steepest ascent experiments is $N = k$, where k indicates the number of parameters in the model. A decreasing function of the number of trials (N) is given by,

$$\varphi(N) = (1/N)\text{E}[\cos(\theta)],$$

where, θ is the angle between true and estimated gradient, and $\mathrm{E}[\cos(\theta)]$ is the expected value of $\cos(\theta)$. See, also, the optimum seeking of the first-order multifactor designs discussed by Box (1952) and Box and Draper (2007). In this chapter we have the combined set of n original and m new observations. Based on Brooks and Mickey's suggestion the total number of observations i.e. $N = n + m$ should be equal or greater than k. Indeed, for the augmented design matrix when $N < k$, the corresponding information matrix would be singular. Hence, the evaluation of $\hat{\underline{\beta}}$ and $\mathrm{var}(\hat{\underline{\beta}})$, in equations (3) and (4), cannot be accomplished, and also the minimization criterion of $|(X'X)^{-1}|$ does not provide any information to determine which additional observations are the most informative ones.

6-4. Methods for Augmenting Existing Data

A very efficient procedure for augmenting data to predict the coefficients vector of an unplanned first-order model is introduced by Gaylor and Merrill (1968). They present a solution to augment new observations, simultaneously, to the existing data. In what follows, we consider an extension of this procedure and propose a quantity for judging whether one set of added observations based on this method produces an orthogonal design matrix or the user requires more observations to make an orthogonal or more nearly orthogonal design. The proposed quantity to recognize the orthogonality of the design matrix is the maximum value of $|(X'X)|$.

For the first-order model, the represented solution to the data augmentation problem is constructed, under the condition that the experimental region is a

rectangle. In a predefined experimental region $\aleph$, the proposed augmentation strategy is at least comparable to other existing procedures with region $\aleph$, for attaining the simultaneous objectives of orthogonalizing the original design, adding the new observations in the restricted experimental region, reducing the variance and the bias of the estimated coefficients vector, and maintaining the original data at the augmented design effectively.

Similar strategies to the proposed scheme to keep the original data at the improved design are the methods presented by Dykstra (1971) and Heiberger, et al. (1993). Dykstra's (1971) method is based on identifying a list of candidates to locate the additional data and recognizing the conditions in the experimental region where the variance of a predicted response is the largest and obtaining a new observation at those conditions. Dykstra's strategy and the proposed one produce the same augmented design matrices, because of the use of the equivalence of criteria in the construction of designs. In fact, the method given by Dykstra is based on the G-optimality criterion and the proposed one is constructed on the D-optimality criterion. The G-optimality criterion is one that minimizes the maximum variance of the predicted response over the experimental region and the D-optimality minimizes $|(X'X)^{-1}|$. See the equivalency of G and D-optimality criterion by Wynn (1970). In the method suggested by Heiberger, et al. (1993), the new observations in the eigenvector coordinates of the original cross-product matrix $X'X$ are augmented to the original design matrix to increase jointly the smaller eigenvalues of $X'X$. Although, this approach results in a fully orthogonal design matrix, in the case of additional observations of small size the new points are usually placed on the outside of the restricted region.

Other solutions to the data augmentation problem, keeping the original data, are Bayesian optimization methods. The interested researchers are referred to the approaches of Bayesian optimal design in Chaloner and Larntz (1989), Chaloner and Verdinelli (1995), Haines et al. (2003), and Ruggoo and Vandebroek (2004).

6-5. Maximum Value of the Determinant of $X'X$

In the sequel, we employ the widely used criterion for judging the performance of a data augmentation procedure, which is based on the maximization of $|(X'X)|$ or, equivalently, the minimization of $|(X'X)^{-1}|$. These amount to selecting the set of added points that would render a maximum value for the determinant of $X'X$. The criterion of maximizing the determinant $|(X'X)|$ eventually leads to an orthogonal design that is our desired design. Several other augmentation criteria for non-orthogonal designs consider maximizing the residual sum of squares, minimizing the multiple correlations, and minimizing the regression sum of squares of the independent variables (see e.g. Dykstra (1966), Box and Draper (2007) among others).

Theorem: Let V be a $N \times k$ design matrix with column vectors $\underline{v}_1$, $\underline{v}_2$, ..., $\underline{v}_j$, ..., $\underline{v}_k$. Let also $<\underline{v}_j, \underline{v}_{j'}> = c_{jj'}$ be the scalar product of $\underline{v}_j$ and $\underline{v}_{j'}$, and $\theta_{jj'}$ be the smallest angle between these vectors. If $j \neq j'$, then $c_{jj'}$'s are variables upon $\theta_{jj'}$'s, otherwise, if $j = j'$, then $c_{jj'}$'s or c_{jj}'s are fixed values (the squared lengths of $\underline{v}_j$'s). Let, $f_V(.) = f_V(c_{12}, c_{13}, ..., c_{jj'}, ..., c_{(k-1)k}) = f_V(c_{21}, c_{31}, ..., c_{j'j}, ..., c_{k(k-1)}) = |V'V|$, where, $c_{jj'} = c_{j'j}$ for all j and j'. Then, for the matrix V^* in the set of V's with squared lengths c_{jj},

for $j=1,2,...,k$, we have $f_{V^*}(.) \le \prod_{j=1}^{k} c_{jj}$. Moreover, $f_{V^*}(.) = \max_{V} f_{V^*}(.) = \prod_{j=1}^{k} c_{jj}$ or, equivalently,

$$|V^{*\prime}V^*| = \max_{V} |V^{*\prime}V^*| = \prod_{j=1}^{k} c_{jj}, \tag{5}$$

when $c_{jj'}$'s of the matrix V^* are equal to zeros, for $j \neq j' = 1,2,...,k$.

This theorem can be proved using induction on k (see Appendix I).

Corollary 1: Let us define $g_V(.) = g_V(c_{12}, c_{13}, ..., c_{jj'}, ..., c_{(k-1)k}) = g_V(c_{21}, c_{31}, ..., c_{j'j}, ..., c_{k(k-1)})$ $= |(V'V)^{-1}|$. Then, for the matrix V^* in the set of V's, we have $g_{V^*}(.) \ge (\prod_{j=1}^{k} c_{jj})^{-1}$. Moreover, $g_{V^*}(.) = \min_{V} g_{V^*}(.) = (\prod_{j=1}^{k} c_{jj})^{-1}$ or, equivalently, $|(V^{*\prime}V^*)^{-1}| = \min_{V} |(V^{*\prime}V^*)^{-1}|$ $(\prod_{j=1}^{k} c_{jj})^{-1}$, when $c_{jj'}$'s of the matrix V^* are equal to zeros, for $j \neq j' = 1,2,...,k$.

Corollary 2: We have $|V^{*\prime}V^*| = \max_{V} |V^{*\prime}V^*| = \prod_{j=1}^{k} c_{jj}$ or, $|(V^{*\prime}V^*)^{-1}| = \min_{V} |(V^{*\prime}V^*)^{-1}|$ $= (\prod_{j=1}^{k} c_{jj})^{-1}$ if the $\underline{v}_j$ and $\underline{v}_{j'}$, for $j \neq j' = 1,2,...,k$, are orthogonal vectors.

The well-known Gram-Schmidt orthogonalization is an alternate method to the proposed strategy to orthogonalize a given matrix. According to corollary (2), the exact minimum value of $|(X'X)^{-1}|$, i.e. $(\prod_{i=1}^{k} c_{ii})^{-1}$ is obtained by applying this method. However, using the Gram-Schmidt algorithm, we will have a transformed version of the original design matrix, and the responses should be created to correspond to the transformed matrix. Thereby, when the Gram-Schmidt process is exploited, the user may not be able to utilize some or all of original data to estimate the regression coefficients. The abandonment of the existing data, especially if the data were expensive or difficult to attain, is a disadvantage of the Gram-Schmidt method as compared to the augmenting data

scheme. The data augmentation strategy presented in this chapter provides guidance in the collecting a few additional observations in order to minimize or nearly minimize $|(X'X)^{-1}|$, without discarding the existing experimental observations. The algorithms of Gram-Schmidt to reduce the number of operations involved in the process and the instability of the process are given by Farebrother (1974), Giraud, et al. (2005), Smoktunowicz, et al. (2006), and Kiani (2008).

Other highly useful techniques for orthogonalizing a potential design are the exchange algorithms. These methods start with a nonsingular n-point design and next augment and eliminate one or more experiments in order to obtain increase in the determinant of the Fisher information matrix $X'X$. In the sense of removing the existing data these approaches are similar to the Gram-Schmidt method. The most commonly used exchange algorithms are presented by Mitchell (2000), and Heredia-Laugner et al. (2003).

6-6. Augmenting Observations in Design Space

In the current section, a sequential algorithm to augment a few data to the original data is offered in a k-dimensions coordinate system. In fact, our proposed solution is an extension to the simultaneous solution introduced by Gaylor and Merrill (1968), for allocating m new observations to the corners of a p-dimensional space, that would minimize $|(X'X)^{-1}|$. The use of our sequential method to solve the data augmentation problems is recommended, because of its attractive properties. It should be mentioned that Gaylor and Merrill's algorithm and the proposed one equivalently decrease the magnitude of $|(X'X)^{-1}|$.

Let us consider the first-order model is based on model (2),

$$\underline{Z} = U\underline{\beta} - \underline{\gamma}, \tag{6}$$

where, $\underline{Z} = \underline{Y} - \underline{1}'\underline{Y}\,\underline{1}/N$, $U = X - \underline{1}\,\underline{1}'X/N$, and $\underline{\gamma} = \underline{\varepsilon} - \underline{1}'\underline{\varepsilon}\,\underline{1}/N$, ($\underline{1}$ is the $N\times 1$ vector of ones). The Fisher information matrix $U'U$ is positive definite, hence, $|U'U| \neq 0$. As a consequence, the estimated path of steepest ascent or the OLS estimator of the regression coefficients ($\underline{\hat{\beta}}$) and its variance ($\mathrm{var}(\underline{\hat{\beta}})$) from the model (6) are

$$\underline{\hat{\beta}} = (U'U)^{-1}U'\underline{Z},$$

$$\mathrm{var}(\underline{\hat{\beta}}) = (U'U)^{-1}U'\,\mathrm{var}(\underline{Z})U(U'U)^{-1} = \left(\sigma^2(I - \frac{1}{N}I)\right)(U'U)^{-1}.$$

The solution for minimizing the magnitude of $(U'U)^{-1}$ gives the minimum magnitude of $\mathrm{var}(\underline{\hat{\beta}})$. Moreover, the maximization of $|(U'U)|$ provides the minimization of $|(U'U)^{-1}|$. For $i = 1,2,...,N = n+m$, let $(x_{i1}, x_{i2}, ..., x_{ij}, ..., x_{ik})$ denote to the coordinate of a point in the rectangular design region. Then, the diagonal and the off-diagonal elements of $U'U$ are

$$\begin{aligned} &\textstyle\sum_{i=1}^{N}(x_{ij} - \bar{x}_{.j})^2 && \text{for} \quad j = 1,2,...,k, \\ &\textstyle\sum_{i=1}^{N}(x_{ij} - \bar{x}_{.j})(x_{ij'} - \bar{x}_{.j'}) && \text{for} \quad j \neq j' = 1,2,...,k. \end{aligned} \tag{7}$$

Under the augmenting data condition, $\sum_{i=1}^{N}(x_{ij} - \bar{x}_{.j})^2$'s in (7) are (nearly) maximized to $\sum_{i=1}^{N}x_{ij}^2$'s and $\sum_{i=1}^{N}(x_{ij} - \bar{x}_{.j})(x_{ij'} - \bar{x}_{.j'})$'s are (nearly) equaled to zero, by adding the corner points in the rectangular region, where $x_{ij} = \pm 1$ and $\bar{x}_{.j} \approx 0$ for $i = n+1,...,N$ and $j = 1,2,...,k$. The maximization of $\sum_{i=1}^{N}(x_{ij} - \bar{x}_{.j})^2$'s provides the maximization of the diagonal entries of $X'X$, where $\max\sum_{i=1}^{N}x_{ij}^2 = \sum_{i=1}^{n}x_{ij}^2 + m$.

Furthermore, equating $\sum_{i=1}^{N}(x_{ij}-\bar{x}_{.j})(x_{ij'}-\bar{x}_{.j'})$'s to zero, we have the off-diagonal elements of $X'X$ equal to zero i.e.

$$\sum_{i}^{N} x_{ij}x_{ij'}=0 \qquad \text{for} \qquad j\neq j'=1,2,...,k. \tag{8}$$

That means the columns of design matrix X will be orthogonal vectors. As a result, based on corollary (2), the magnitude of $|(X'X)^{-1}|$ is minimized.

Using equation (8) we have

$$\sum_{i=1}^{n} x_{ij}x_{ij'}+\sum_{i=n+1}^{n+m} x_{ij}x_{ij'}=0 \qquad j\neq j'=1,2,...,k. \tag{9}$$

Further, the assumption $\bar{x}_{.j}=0$ gives

$$\sum_{i=1}^{n} x_{ij}+\sum_{i=n+1}^{n+m} x_{ij}=0 \qquad j=1,2,...,k. \tag{10}$$

Let us denote the number of new points for adding to the gth corner by n_g for $g=1,2,...,p=2^k$, where $\sum_{g=1}^{p} n_g=m$. Let us also denote the k-dimensions coordinates of a point in the gth corner by $(\omega_{g1},\omega_{g2},...,\omega_{gj},..,\omega_{gk})$, where $\omega_{gj}=\pm 1$ for $g=1,2,...,p$ and $j=1,2,...,k$. Applying equations (9) and (10), we have

$$\sum_{i=1}^{n} x_{ij}+\sum_{g=1}^{p}\omega_{gj}n_g=\sum_{i=1}^{n} x_{ij}x_{ij'}+\sum_{g=1}^{p}\omega_{gj}\omega_{gj'}n_g=0 \qquad j\neq j'=1,2,...,k.$$

To obtain a unique solution, for n_g, in the above equation, we choose to add other conditions which are independent of the existing conditions,

$$\sum_{i=1}^{n} x_{ij}x_{ij'}x_{ij''}+\sum_{i=n+1}^{n+m} x_{ij}x_{ij'}x_{ij''}=\sum_{i=1}^{n} x_{ij}x_{ij'}x_{ij''}x_{ij'''}+\sum_{i=n+1}^{n+m} x_{ij}x_{ij'}x_{ij''}x_{ij'''}$$

$$=\sum_{i=1}^{n} x_{ij}x_{ij'}x_{ij''}...x_{ij^k}+\sum_{i=n+1}^{n+m} x_{ij}x_{ij'}x_{ij''}...x_{ij^k}=0 \qquad j\neq j'\neq j''\neq ...\neq j^k=1,2,...,k,$$

such that, these conditions do not affect the properties desired. Furthermore, we already mentioned,

$$\sum_{g=1}^{p} n_g = m .$$

Under these circumstances, we will have p unknown values n_g for $g = 1,2,...,p$, and also p independent linear equations follow,

$$\sum_{i=1}^{n} x_{ij} + \sum_{g=1}^{p} \omega_{gj} n_g = \sum_{i=1}^{n} x_{ij} x_{ij'} + \sum_{g=1}^{p} \omega_{gj} \omega_{gj'} n_g = \sum_{i=1}^{n} x_{ij} x_{ij'} x_{ij''} + \sum_{g=1}^{p} \omega_{gj} \omega_{gj'} \omega_{gj''} n_g \qquad (11)$$

$$= ... = \sum_{i=1}^{n} x_{ij} x_{ij'} x_{ij''} ... x_{ij^k} + \sum_{g=1}^{p} \omega_{gj} \omega_{gj'} \omega_{gj''} ... \omega_{gj^k} n_g = 0 \qquad j \neq j' \neq j'' \neq ... \neq j^k = 1,2,...,k ,$$

and,

$$n_1 + n_2 + ... + n_p = m . \qquad (12)$$

Consequently, the number of points, n_g, that will be added to gth corner of the experimental region can be obtained by solving equations (11) and (12). These equations, for $g = 1,2,...,p$, give

$$n_g = \frac{1}{p}[m - \Psi_g], \qquad (13)$$

where,

$$\Psi_g = \sum_{j=1}^{k} \omega_{gj} \sum_{i=1}^{n} x_{ij} + \sum_{j \neq j'=1}^{k} \omega_{gj} \omega_{gj'} \sum_{i=1}^{n} x_{ij} x_{ij'} + ... + \sum_{j \neq j' \neq ... \neq j^{(k)}=1}^{k} \omega_{gj} \omega_{gj'} ... \omega_{gj^{(k)}} \sum_{i=1}^{n} x_{ij} x_{ij'} ... x_{ij^{(k)}} .$$

Meanwhile, we set the negative value n_g equal to zero and round the remaining values to integers.

As an alternative to the simultaneous method given by Gaylor and Merrill (1968), we propose a sequential generation of new points, in order to

i) overcome difficulties of the rounding rule problems for n_g,

ii) examine the optimality of the design after each set of new observations.

In each stage of data augmentation, first Ψ_g's is successively computed, then a new point is located in gth corner that its related Ψ_g is the smallest, for $g = 1,2,...,p$. Indeed, the algorithm presented by Gaylor and Merrill and our proposed one are similar in the concept of increasing the smaller Ψ_g's. Simulation experiments show these two procedures create the same optimal design.

6-7. Generalized and Nonlinear Models

As noted in section 6-2, the direction of steepest ascent is taken parallel to the normal vector to a fitted surface based on the first-order model. Hence we have exploited the model (2), where its response variable follows the normal distribution. The solution given in the previous section for augmenting data is applicable only in the case of the linear model (2). Under the situations that the responses vector is modeled by a generalized linear model and the goal is to find a solution for the augmentation data problem, we refer the user to the suggested methods by Dror and Steinberg (2006, 2008). See also Khuri et al. (2006), who provide a broad review of the topic for the model with a response variable that follows a distribution belonging to the exponential family.

The papers by Chaudhuri and Mykland (1993) and Sinha and Wiens (2002) present optimal data augmentation procedures for a nonlinear regression model. As noted, when we obtain a general vicinity of the optimum response using the steepest ascent path, then in most cases the second-order model is employed to find the levels of independent variables that optimize the predicted response. The nonlinear models can be used, in some particular cases, as an alternative to the second-order linear models.

However, to move efficiently along a path of improvement toward the vicinity of optimum, we have focused on the modification of the estimated coefficients vector of a linear model with the response variable of the normal distribution. In the sequel, an example is given to illustrate the method represented above and its application in the steepest ascent methodology.

6-8. An Example

A chemical engineer is interested in determining the operating conditions that maximize the yield of a process. Three controllable variables influence process yield: factor A (temperature), factor B (time) and factor C (pressure). She is currently operating the process with factor A of 110 deg F, factor B of 35 minutes, and factor C of 225 psi. Because it is unlikely that this region contains the optimum, she fits a first-order model and applies the method of steepest ascent. In this case, the region of exploration for fitting the first-order model is (100, 120) deg F of factor A, (20, 50) minutes of factor B and (200, 250) psi of factor C. If ξ_1, ξ_2 and ξ_3 denote the natural variables temperature, time and pressure, respectively, then the coded variables are defined by

$$x_1 = (\xi_1 - 110)/10 \quad ; \quad x_2 = (\xi_2 - 35)/15 \quad ; \quad x_3 = (\xi_3 - 225)/25 .$$

Please note the proposed method of augmenting data is constructed on the criterion of minimizing $|(X'X)^{-1}|$, where the solution attained is invariant under a change of variable scale. As a consequence, the linear transformation above does not ruin the generality of solution. Under these circumstances, for a set of simulation experiments the coded design matrix is

$$X'_{3\times 12}=\begin{bmatrix} 0.20 & 0.20 & 0.20 & -0.29 & 0.20 & 0.33 & 0.17 & 0.14 & -0.29 & 0.17 & 0.33 & -0.33 \\ -0.14 & -0.11 & 0.40 & -0.20 & -0.14 & -0.33 & 0.40 & -0.17 & 0.20 & -0.67 & -0.33 & 0.20 \\ 0.13 & 0.13 & -0.33 & 0.40 & 0.86 & 0.40 & -0.25 & -0.29 & 0.40 & 0.13 & 0.40 & 0.29 \end{bmatrix}.$$

For calculating n_g for $g=1,2,...,8$, tables 6.1, 6.2 and 6.3 are established,

Table 6.1: The values of x_{i1}, x_{i2}, ..., $x_{i1}x_{i2}x_{i3}$ for $i=1,2,...,12$.

	i												
	1	2	3	4	5	6	7	8	9	10	11	12	Sum
x_{i1}	0.20	0.20	0.20	-0.29	0.20	0.33	0.17	0.14	-0.29	0.17	0.33	-0.33	1.04
x_{i2}	-0.14	-0.11	0.40	-0.20	-0.14	-0.33	0.40	-0.17	0.20	-0.67	-0.33	0.20	-0.90
x_{i3}	0.13	0.13	-0.33	0.40	0.86	0.40	-0.25	-0.29	0.40	0.13	0.40	0.29	2.25
$x_{i1}x_{i2}$	-0.03	-0.02	0.08	0.06	-0.03	-0.11	0.07	-0.02	-0.06	-0.11	-0.11	-0.07	-0.36
$x_{i1}x_{i3}$	0.03	0.03	-0.07	-0.11	0.17	0.13	-0.04	-0.04	-0.11	0.02	0.13	-0.10	0.04
$x_{i2}x_{i3}$	-0.02	-0.01	-0.13	-0.08	-0.12	-0.13	-0.10	0.05	0.08	-0.08	-0.13	0.06	-0.63
$x_{i1}x_{i2}x_{i3}$	0.00	0.00	-0.03	0.02	-0.02	-0.04	-0.02	0.01	-0.02	-0.01	-0.04	-0.02	-0.19

Table 6.2: The values of ω_{g1}, ω_{g2}, ..., $\omega_{g1}\omega_{g2}\omega_{g3}$, for $g=1,2,...,8$.

	g							
	1	2	3	4	5	6	7	8
ω_{g1}	-1	-1	-1	-1	1	1	1	1
ω_{g2}	-1	-1	1	1	-1	-1	1	1
ω_{g3}	-1	1	-1	1	-1	1	-1	1
$\omega_{g1}\omega_{g2}$	1	1	-1	-1	-1	-1	1	1
$\omega_{g1}\omega_{g3}$	1	-1	1	-1	-1	1	-1	1
$\omega_{g2}\omega_{g3}$	1	-1	-1	1	1	-1	-1	1
$\omega_{g1}\omega_{g2}\omega_{g3}$	-1	1	1	-1	1	-1	-1	1

Table 6.3: The values of $\omega_{g1}\sum_{i=1}^{12}x_{i1}$, $\omega_{g2}\sum_{i=1}^{12}x_{i2}$, ..., $\omega_{g1}\omega_{g2}\omega_{g3}\sum_{i=1}^{12}x_{i1}x_{i2}x_{i3}$ for $g=1,2,...,8$.

	g							
	1	2	3	4	5	6	7	8
$\omega_{g1}\sum_{i=1}^{12}x_{i1}$	-1.04	-1.04	-1.04	-1.04	1.04	1.04	1.04	1.04
$\omega_{g2}\sum_{i=1}^{12}x_{i2}$	0.90	0.90	-0.90	-0.90	0.90	0.90	-0.90	-0.90
$\omega_{g3}\sum_{i=1}^{12}x_{i3}$	-2.25	2.25	-2.25	2.25	-2.25	2.25	-2.25	2.25
$\omega_{g1}\omega_{g2}\sum_{i=1}^{12}x_{i1}x_{i2}$	-0.36	-0.36	0.36	0.36	0.36	0.36	-0.36	-0.36
$\omega_{g1}\omega_{g3}\sum_{i=1}^{12}x_{i1}x_{i3}$	0.04	-0.04	0.04	-0.04	-0.04	0.04	-0.04	0.04
$\omega_{g2}\omega_{g3}\sum_{i=1}^{12}x_{i2}x_{i3}$	-0.63	0.63	0.63	-0.63	-0.63	0.63	0.63	-0.63
$\omega_{g1}\omega_{g2}\omega_{g3}\sum_{i=1}^{12}x_{i1}x_{i2}x_{i3}$	0.19	-0.19	-0.19	0.19	-0.19	0.19	0.19	-0.19
Ψ_g	-3.15	2.16	-3.35	0.19	-0.82	5.40	-1.68	1.25

Let the number m of added points to be defined $m=4$. Examination of Table 6.3 and equation (13) gives

$$n_1=0.89 \;;\; n_2=0.23 \;;\; n_3=0.92 \;;\; n_4=0.48,$$

$$n_5=0.60 \;;\; n_6=-0.17 \;;\; n_7=0.71 \;;\; n_8=0.34,$$

where $\sum_{g=1}^{8}n_g=4$. Using the rounding rule of n_g, we set

$$n_1=1 \;;\; n_2=0 \;;\; n_3=1 \;;\; n_4=0 \qquad (14)$$

$$n_5=1 \;;\; n_6=0 \;;\; n_7=1 \;;\; n_8=0.$$

Our proposed sequential method produces results similar to those obtained by (14), but the successive method is easier to use than the simultaneous one, as it overcomes difficulties associated with of the rounding rule in the determination of n_g's discussed in section 6-6. As Table 6.3 shows, in the first stage Ψ_3 is the

smallest among Ψ_g's, therefore the first new point is $(\omega_{31},\omega_{32},\omega_{33})=(-1,1,-1)$. The smallest Ψ_g for the second, third and forth stage are $\Psi_1=-4.15$, $\Psi_7=-3.71$, and $\Psi_5=-3.81$, respectively.

Let a batch of data that contains the new points located by the proposed method be refereed to as case 1, and a setting of the initial data as case 0. The results for these two cases and some other cases are given in Table 6.4.

Table 6.4: The initial design and the augmented designs, for $n=12$ and $m=4$.

Case	N	n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8	$\lvert(X'X)^{-1}\rvert$	$\lvert(X'X)^{-1}\rvert-\min\lvert(X'X)^{-1}\rvert$	$\bar{\theta}$
0	12	0	0	0	0	0	0	0	0	9.58E-01	3.08E-01	19.48
1	16	1	0	1	0	1	0	1	0	7.19E-03	1.31E-04	8.91
2	16	0	1	0	0	1	2	0	0	1.29E-02	5.85E-03	11.48
3	16	0	1	0	1	1	1	0	0	1.22E-02	5.14E-03	11.04
4	16	0	2	0	0	0	1	0	1	1.08E-02	3.70E-03	10.55

In this table, case 1 leads to a minimum value of $\lvert(X'X)^{-1}\rvert$, compared to cases 0, 2, 3 and 4. Additionally, a minimum value of $\lvert(X'X)^{-1}\rvert-\min\lvert(X'X)^{-1}\rvert$ is obtained when case 1 is used. As a result, if we construct the experimental design upon case 1, then we will achieve a nearly minimum value of the $\mathrm{var}(\underline{\hat{\beta}})$. In Table 6.4, the quantity $\min\lvert(X'X)^{-1}\rvert$ equals the value of the inverse determinant of the cross-product of design matrix as X^*, which is fully orthogonal. It should be noted that the points (rows) of X^* and X (the augmented design) have the equivalent distances from the origin point of the experimental region. Nevertheless, the points of X^* have been not restricted to a candidate set of

experimental points. The X^* points are located on a k-dimensional sphere of radius $\sqrt{k}$, for the coded variables.

The bias of steepest ascent path is considered by computing the smaller angle between true and estimated path, θ. Assume that the true path is $\underline{\beta}_p = (19, -37, 64)'$ and the experimental matrix is based on cases 0 to 4. Let the variance and the mean of error vector, possessing the normal distribution, be $400 \times I$ and $\underline{0}$, where the identical matrix I is 12 by 12 for case 0 and 16 by 16 for cases 1 to 4. Define $\hat{\underline{\beta}}_{0p}$, $\hat{\underline{\beta}}_{1p}$, ..., $\hat{\underline{\beta}}_{4p}$ the OLS estimators of $\underline{\beta}_p$ for cases 1 to 4. The smaller angle between true path and its estimate, $\hat{\underline{\beta}}_p$, is

$$\theta = \arccos\left(\frac{< \underline{\beta}_p, \hat{\underline{\beta}}_p >}{|\underline{\beta}_p \| \hat{\underline{\beta}}_p|}\right).$$

We obtained the value of θ, using a simulation data, and replicated this procedure 50,000 times (see Appendix II). Table 6.4 shows $\bar{\theta}$ that is the average value of θ. Case 1 leads to a minimum value of $\bar{\theta}$, while case 0 to a maximum one. Therefore, it can be concluded that the bias of estimated path from case 1 is less than other cases.

Although, the orthogonality of the original design was considerably increased with 4 additional points, however the quantity $|(X'X)^{-1}| - \min|(X'X)^{-1}|$ shows that the augmented design have no full orthogonal columns. Since the experimental region is a rectangular constrained region and the set of candidates for the new observations is the 8 corner points of a rectangle, then the construction of such orthogonal designs exactly is not often possible. However, one could attempt to improve the original design by adding more data to attain a closer to orthogonal

design. Table 6.5 is established in order to choose the proposed additional observations using the corner points. The measures of interest associated with the design matrix are calculated by the proposed sequential method to revise the design by additional points one at a time. Table 6.5 shows the results acquired for $m=4(2)16$.

Table 6.5: The initial design and the augmented designs, for $n=12$ and $m=4(2)16$.

m	N	n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8	$\lvert(X'X)^{-1}\rvert$	$\lvert(X'X)^{-1}\rvert-\min\lvert(X'X)^{-1}\rvert$	$\bar{\theta}$
0	12	0	0	0	0	0	0	0	0	9.58E-01	3.08E-01	19.48
4	16	1	0	1	0	1	0	1	0	7.19E-03	1.31E-04	8.91
6	18	1	0	1	1	1	0	1	1	2.76E-03	9.67E-05	7.15
8	20	1	1	2	1	1	0	1	1	1.28E-03	6.26E-06	6.20
10	22	2	1	2	1	1	1	1	1	7.34E-04	2.59E-05	6.00
12	24	2	1	2	1	2	1	2	1	4.34E-04	9.35E-07	5.79
14	26	2	1	2	2	2	1	2	2	2.86E-04	2.22E-06	5.39
16	28	2	2	3	2	2	1	2	2	1.96E-04	2.23E-07	4.94

Table 6.5 shows that the quantity $\lvert(X'X)^{-1}\rvert-\min\lvert(X'X)^{-1}\rvert$ or equivalently the non-orthogonality of the augmented design is significantly reduced when m increases from $m=4$ to $m=8$. This quantity remains stable about zero for $m=8$ through $m=16$. Table 6.5 also indicates that both the quantities $\lvert(X'X)^{-1}\rvert$ and $\bar{\theta}$ are decreased toward zero, sharply from $m=4$ to $m=8$ and slightly from $m=8$ to $m=16$. As a consequence, employing more data at the predefined list of candidate points we observe an increase in the orthogonality of the augmented design and a decrease in the variance and the bias of the estimated path. According to the design matrix in this example and other simulation designs, a large number of additional points is not recommended if creating the new observations is costly or

difficult. With a moderate number of new data (in this example $m=8$) a nearly orthogonal design is gained. In which cases, the steepest ascent path is estimated precisely enough.

6-9. Conclusion

We have explored the implications of the data augmentation strategy for additive models without discarding the existing observations. The minimum value of the inverse determinant of the information matrix $X'X$ has been employed, for judging the performance of the procedure proposed in this study. If the user is permitted to choose the m new observations, then, in an efficient manner, the original data are augmented by additional points either simultaneously or sequentially. The procedure selects points at the extremes of the experimental region, where the augmented design is nearly (or completely) orthogonalized, and the discrepancy between $|(X'X)^{-1}|$ and $\min|(X'X)^{-1}|$ is made close (or equal) zero. An application of the proposed data augmentation method has been given. Both bias and variance of the estimated path of steepest ascent is efficiently decreased by using this procedure.

Appendix I

Proof: For $k=2$, since $f_{V^*}(c_{12})=|V^{*'}V^*|=\prod_{j=1}^{k=2}c_{jj}-c_{12}^2$ is a convex function of the c_{12} with maximum value at $c_{12}=0$, then $\max_V f_{V^*}(c_{12})=f_{V^*}(0)$, i.e. $\max_V|V^{*'}V^*|=\prod_{j=1}^{k=2}c_{jj}$.

If we assume that equation (5) is true for k, i.e.

$$\max_V f_{V^*}(c_{12},c_{13},...,c_{jj'},...,c_{(k-1)k})=f_{V^*}(0,0,...,0)=\max_V|V^{*'}V^*|=\prod_{j=1}^{k}c_{jj},$$

then for $k+1$ we will have,

$$\max_{V} |V^{*'}V^{*}| = \max_{V} f_{V^{*}}(c_{12},c_{13},...,c_{jj'},...c_{(k-1)(k)},c_{(1)(k+1)},...,c_{(j)(k+1)},...,c_{(k)(k+1)})$$

$$= \max_{V} f_{V^{*}}(0,0,...,0,c_{(1)(k+1)},...,c_{(j)(k+1)},...,c_{(k)(k+1)}) . \qquad (15)$$

Function f in (15), can be represented by,

$$f_{V^{*}}(0,0,...,0,c_{(1)(k+1)},...,c_{(j)(k+1)},...,c_{(k)(k+1)}) = \prod_{j=1}^{k+1} c_{jj} - \sum_{j=1}^{k} c_{j(k+1)}^{2} \prod_{\substack{i^{(j)}=1,2,...,k \\ i^{(j)}\neq j}}^{k} c_{i^{(j)}i^{(j)}} . \qquad (16)$$

The first and second order partial derivatives of the function (16) upon $c_{j(k+1)}$, for $j=1,2,...,k$, are given by

$$\frac{\partial^{(1)} f_{V^{*}}(0,0,...,0,c_{(1)(k+1)},...,c_{(j)(k+1)},...,c_{(k)(k+1)})}{\partial c_{j(k+1)}} = -2c_{j(k+1)} \prod_{\substack{i^{(j)}=1,2,...,k \\ i^{(j)}\neq j}}^{k} c_{i^{(j)}i^{(j)}} , \qquad (17)$$

$$\frac{\partial^{(2)} f_{V^{*}}(0,0,...,0,c_{(1)(k+1)},...,c_{(j)(k+1)},...,c_{(k)(k+1)})}{\partial c_{j(k+1)}} = -2 \prod_{\substack{i^{(j)}=1,2,...,k \\ i^{(j)}\neq j}}^{k} c_{i^{(j)}i^{(j)}} < 0 . \qquad (18)$$

Equation (18) shows that the function $f_{V^{*}}(0,0,...,0,c_{(1)(k+1)},...,c_{(j)(k+1)},...,c_{(k)(k+1)})$ is a convex function, for all j, and because of (17), this function is maximized at $c_{j(k+1)}=0$, for $j=1,2,...,k$,

$$\max_{V} f_{V^{*}}(0,0,...,0,c_{(1)(k+1)},...,c_{(j)(k+1)},...,c_{(k)(k+1)}) = f_{V^{*}}(0,0,...,0) = \prod_{j=1}^{k+1} c_{jj} .$$

or, equivalently, for $k+1$, we have

$$|V^{*'}V^{*}| = \max_{V} |V^{*'}V^{*}| = \prod_{j=1}^{k+1} c_{jj} ,$$

where, the elements $c_{jj'}$ of $f_{V^{*}}(c_{12},c_{13},...,c_{jj'},...c_{(k-1)(k)},c_{(1)(k+1)},...,c_{(j)(k+1)},...,c_{(k)(k+1)})$ are equal to zero, for $j \neq j' = 1,2,...,k+1$. Hence equation (5) is true for all k. □

Appendix II

```
n=0;
k=50000;
Bp=[ 19 ; -37 ; 64];Y0=x0*Bp;
Y1=x1*Bp;t0=Bp'*Bp;
t1=Bp'*Bp; B0p=Bp;B1p=Bp;
for i=1:k
e=20*(2*(rand(18,1)-0.5)) / (sqrt(0.333));
 Y0(:,i)=x0*Bp+(e);
 Y1(:,i)=x1*Bp+(e);
B0p(:,i)=((x0'*x0)^-1)*x0'*Y0(:,i);
B1p(:,i)=((x1'*x1)^-1)*x1'*Y1(:,i);
t0(:,i)=abs(acos((Bp)'*(B0p(:,i))/ ( norm(Bp)*norm(B0p(:,i))) ));
t1(:,i)=abs (acos((Bp)'*(B1p(:,i))/ ( norm(Bp)*norm(B1p(:,i))) ) );
if t0(:,i) < t1(:,i)
 n=n+1;
end;
end
n/k
 180*mean(t0)/pi
180*mean(t1)/pi
```

7

Methods for the Construction of Optimal Designs in Regression Problems

CHAPTER OUTLINE

7-1. Introduction

Optimal designs in regression problems have received a lot of attention in the literature. The purpose of this chapter is to present a computational method for finding optimal designs in regression problems by augmenting the existing data by a moderate number of new data as an alternative method to that presented in the preceding chapter. The advantages of this method over the method introduced in the previous chapter are discussed.

If there is no correlation between the regressors, they are said to be orthogonal. When there are near linear dependencies between the regressors, it is said that we have multicollinearity and in such cases inferences based on the regression model can be misleading or erroneous. The data augmentation scheme is an efficient method to mitigate the problem of multicollinearity in the existing data.

Consider a linear regression model, including n original and l additional observations,

$$\underline{y}_{(n+l)\times 1} = X_{(n+l)\times(k)}\underline{\beta}_{(k\times 1)} + \underline{\varepsilon}_{(n+l)\times 1}, \quad \text{for} \quad l = 1,2,\ldots,$$

with $\underline{y}_{n+l} = \begin{pmatrix} \underline{y}_n \\ \underline{y}_l \end{pmatrix}$ denoting the response vector, whose first n elements are the original data, $\underline{y}_n$, and whose last l elements are the new data, $\underline{y}_l$; $X_{n+l} = \begin{pmatrix} X_n \\ X_l \end{pmatrix}$ denotes the design matrix with linearly independent columns, $n+l \geq k$; $\underline{\beta}$ denotes the coefficient vector; and $\underline{\varepsilon}_{n+l} = \begin{pmatrix} \underline{\varepsilon}_n \\ \underline{\varepsilon}_l \end{pmatrix}$ stands for the random error vector for which it is assumed that $\mathrm{E}(\underline{\varepsilon}_{n+l}) = \underline{0}$ and $\mathrm{var}(\underline{\varepsilon}_{n+l}) = \sigma^2 I_{n+l}$, I_{n+l} being the identity matrix.

The new observations $(\underline{y}_l, X_l)$, obtained by the simultaneous and sequential methods, will be considered to the aim of obtaining efficient estimators of the regression coefficients. We denote the number of additional points by l in a sequential method. If new observations are simultaneously generated, then the number of new points is indicated by m.

The least squares estimator of the parameter $\underline{\beta}$ and its variance are

$$\underline{\hat{\beta}} = (X_n' X_n)^{-1} X_n' \underline{y}_n \quad ; \quad \text{var}(\underline{\hat{\beta}}) = \sigma^2 (X_n' X_n)^{-1}.$$

Recall that the least squares estimator $\underline{\hat{\beta}}$ has minimum variance in the class of linear unbiased estimators of $\underline{\beta}$. Adding new observations will always improve an estimator. But we want to "optimally add" data to get the best improvement for a given number of additional observations. Adding new data causes a decrease of $|(X_n' X_n)^{-1}|$ and other beneficial results such as the reduction of mean squared error.

In the sequel, we concisely review the most commonly used data augmentation methods and the optimality criteria against which these are judged. We present the data augmentation algorithms in a spherical and a rectangular region. The proposed algorithms sequentially construct the lth new point so as to maximize the increase of the minimal eigenvalue of $X_{n+l-1}' X_{n+l-1}$. We also present a report to the performance of the proposed methods and the studied ones.

7-2. Alphabetic Optimality Criteria

Alphabetically optimal designs are used in the cases where standard or fractional factorial designs cannot be easily utilized. We briefly describe the most popular types of alphabetic optimality criteria: G, D, E, and A-optimality.

A G-optimal design is based on $\min_{\boldsymbol{x}_{n+l}\in\aleph} \max(\text{var}[\hat{y}(\boldsymbol{x}_{n+l})])$ criterion, where $\text{var}[\hat{y}(\boldsymbol{x}_{n+l})]$ indicates the variance of the predicted response at lth new point $\boldsymbol{x}_{n+l}$ on region $\aleph$, and equals to $\sigma^2 \boldsymbol{x}_{n+l}(X'_{n+l-1}X_{n+l-1})^{-1}\boldsymbol{x}'_{n+l}$. In other words, in utilizing the G-optimality criterion one assembles the additional observations to minimize the maximum of $\text{var}[\hat{y}(\boldsymbol{x}_{n+l})]$. This criterion has been considered by Smith (1918), Guest (1958) and Kiefer and Wolfowitz (1959), Song and Wong (1999).

Wald (1943) introduced the D-optimality criterion to maximize the determinant of $X'_{n+l}X_{n+l}$, or equivalently, to minimize the determinant of $(X'_{n+l}X_{n+l})^{-1}$. The concept of D-optimal criterion has been advocated by Mood (1946), De la Garza (1954), Hoel (1958), Kiefer (1961), Wynn (1970), Box and Draper (1971), Mitchell (2000), Li, et al. (2005), Dey and Mukerjee (2006), and several other authors.

In this chapter we consider the criterion of D-optimality when all k parameters are estimated. Sometimes, the experimenter is interested in only $s<k$ parameters, so that the other $k-s$ parameters are nuisance parameters. When interest is in $s<k$ parameters, G and D-optimality are defined by G_s and D_s-optimality, see Kiefer (1961), Karlin and Studden (1966), Atwood (1969).

Wynn (1970) presents the relationship between the maximization of the augmented information matrix and the variance of predicted response, which is proportional to the right-hand side of following equation,

$$\frac{|X'_{n+l}X_{n+l}|}{|X_{n+l-1}X_{n+l-1}|}-1=\boldsymbol{x}_{n+l}(X'_{n+l-1}X_{n+l-1})^{-1}\boldsymbol{x}'_{n+l}.$$

This mentions that the maximization of $|X'_{n+l}X_{n+l}|$, D-optimality, is equivalent to the maximization of $\mathrm{var}[\hat{y}(\boldsymbol{x}_{n+l})]$, G-optimality. The equivalency of G and D-optimality is suggested in general equivalence theorem by St. John and Draper (1975).

Ehrenfeld (1955) suggested that maximizing the minimum eigenvalue of $X'_{n+l}X_{n+l}$ be used as a design criterion, E-optimality. The relevance of E and D-optimality for problems of testing hypotheses and of estimation was indicated by Kiefer (1958). It was shown that the D-optimality is generally more meaningful. See also E-optimal designs by Moerbeek (2005).

Elfving (1952) and Chernoff (1953) considered the minimization of the trace of $(X'_{n+l}X_{n+l})^{-1}$ to obtain the A-optimal designs. The trace of $\sigma^2(X'_{n+l}X_{n+l})^{-1}$ is equal to the expected squared distance from the least squares estimator to the true parameter vector,

$$\mathrm{E}[(\underline{\beta}-\underline{\hat{\beta}})'(\underline{\beta}-\underline{\hat{\beta}})]=\sigma^2\mathrm{tr}(X'_{n+l}X_{n+l})^{-1}.$$

When the information matrix $X'_{n+l}X_{n+l}$ is ill-conditioned because of multicollinearity, the $\mathrm{tr}(X'_{n+l}X_{n+l})^{-1}$ will be large. By minimizing the $\mathrm{tr}(X'_{n+l}X_{n+l})^{-1}$, the multicollinearity in the experimental data is broken up and eventually a well-condition design matrix is gained. The trace of $(X'_{n+l}X_{n+l})^{-1}$ can

be given by the corresponding eigenvalues, $\lambda_{k,n+l} \geq \lambda_{k-1,n+l} \geq ... \geq \lambda_{1,n+l} \geq 0$, of the information matrix $X'_{n+l}X_{n+l}$, where $\mathrm{tr}(X'_{n+l}X_{n+l})^{-1} = \sum_{j=1}^{k} \lambda_{j,n+l}^{-1}$.

Our proposed optimal design subsumes alphabetically optimal designs. In the sense of these optimality criteria, in section 7-4 the advantages of the new method over the traditional design construction strategies are discussed.

7-3. Three Existing Methods

Gaylor and Merrill (1968) have extended the method proposed by Dykstra (1966), for the orthogonalization of unplanned experiments. Their objectives are to decrease the pairwise correlations among the regressors, and also to minimize the variances of regression coefficient estimates. Gaylor and Merrill present a solution to add additional observations, simultaneously, with the existing data. The solution is based on the first-order model such as

$$(y_i - \bar{y}) = \beta_1(\xi_{i1} - \bar{\xi}_1) + ... + \beta_j(\xi_{ij} - \bar{\xi}_j) + ... + \beta_k(\xi_{ik} - \bar{\xi}_k) + \varepsilon_i \quad \text{for} \quad i = 1,2,...,n+m,$$

where, $\bar{y}$ implies the mean of the responses, and $\bar{\xi}_j$ the mean of the natural independent variables ξ_{ij}, for $j = 1,2,...,k$. For more convenience, the linear transformation of the natural variables to coded ones, x_{ij}, is carried out. So, for an orthogonal design, we have

$$\mathrm{var}(\hat{y}) = \sigma^2[\frac{1}{n+m} + \sum_{j}^{k} \frac{(x_j - \bar{x}_j)^2}{\sum_{i}^{n+m}(x_{ij} - \bar{x}_i)^2}].$$

The maximum $\mathrm{var}(\hat{y})$ is minimized by adding points to $p = 2^k$ corners of the rectangular region, i.e. $x_j = \pm 1$, such that $\bar{x}_j = 0$ for all j. The number of points to be added to the gth corner, for $g = 1,2,...,2^k$, of a k-dimensions design is given by

$$n_g = \frac{1}{p}[m - \sum_{j=1}^{k} \omega_{gj} \sum_{i=1}^{n} x_{ij} - \sum_{j \neq j'=1}^{k} \omega_{gj}\omega_{gj'} \sum_{i=1}^{n} x_{ij}x_{ij'} - \ldots - \sum_{j \neq j' \neq \ldots \neq j^k=1}^{k} \omega_{gj}\omega_{gj'} \ldots \omega_{gj^k} \sum_{i=1}^{n} x_{ij}x_{ij'} \ldots x_{ij^k}],$$

In this case, ω_{gj}'s are equal to ± 1, and indicate to the components of the coordinates of the gth corner point in k-dimensions i.e. $(\omega_{g1}, \ldots, \omega_{gj}, \ldots, \omega_{gk})$.

Later, Dykstra (1971) presented a sequential approach for maximizing the determinant of the information matrix. The relationship between the determinant of the augmented information matrix, $|X'_{n+l}X_{n+l}|$, and the proportion of the variance of the estimated response, $\boldsymbol{x}_{n+l}(X'_{n+l-1}X_{n+l-1})^{-1}\boldsymbol{x}'_{n+l} = \mathrm{var}(\hat{y} \mid \boldsymbol{x}_{n+l})/\sigma^2$, is given

$$|X'_{n+l}X_{n+l}| = |X'_{n+l-1}X_{n+l-1} + \boldsymbol{x}'_{n+l}\boldsymbol{x}_{n+l}| = |X'_{n+l-1}X_{n+l-1}|[1 + \boldsymbol{x}_{n+l}(X'_{n+l-1}X_{n+l-1})^{-1}\boldsymbol{x}'_{n+l}].$$

The determinant of $X'_{n+l}X_{n+l}$ is maximized by taking as the next observation the point in the experimental region where $\mathrm{var}(\hat{y} \mid \boldsymbol{x}_{n+l})$ is greatest. In choosing specific combinations, for a first-order model the procedure will select points at the extremes of the experimental space, so that corner points of the rectangular region need be specified as candidates.

The algorithms introduced by Gaylor and Merrill (1968) and Dykstra (1971) produce D or G-optimal designs. In fact, the determinant of X'_nX_n is maximized by locating the new points in the design region, which minimize the maximum of the variance of the predicted response.

A simultaneous strategy to orthogonalize a design matrix is presented by Heiberger, Bhaumik, and Holland (1993). This method is based on the U-optimality criterion. The design $X^o_{n+m} = (X_n; X^o_m)$ is defined to be U-optimal with respect to the set of designs $X_{n+m} = (X_n; X_m)$ when X^o_m is chosen to make the $r = r(\delta^2)$ smallest eigenvalues of $X^{o'}_{n+m}X^o_{n+m}$ equal to each other and to a value $\lambda^o = \lambda^o(\delta^2)$. Here, X_m is a $m \times k$ augmented design with m new observations,

and δ^2 a quantity that equals to $\mathrm{tr}(X_m'X_m) = \mathrm{tr}(X_{n+m}'X_{n+m}) - \mathrm{tr}(X_n'X_n)$. For an ordered set of nonnegative numbers $\lambda_k' \geq \ldots \geq \lambda_2' \geq \lambda_1' \geq 0$, the relation between δ^2 and $\lambda^o(\delta^2)$ is

$$\delta^2 = \sum_{j=1}^{k^o} [\lambda^o(\delta^2) - \lambda_j'] \quad \text{or} \quad \lambda^o(\delta^2) = (\sum_{j=1}^{k^o} \lambda_j' + \delta^2)/(k - k^o + 1),$$

where, for any λ^o in interval $[\lambda_1', \lambda_k']$, $k^o = k^o(\lambda^o)$ is the index of the left endpoint of the subinterval $[\lambda_{k^o-1}', \lambda_{k^o}']$ in which the value λ^o appears.

In the algorithm suggested by Heiberger et al. (1993) the augmented matrix X_m is constructed under the restriction,

$$0 \leq \delta^2 \leq \sum_{j=1}^{k-1} (\lambda_{k,n} - \lambda_{j,n}),$$

where, $\lambda_{j,n}$'s, $\lambda_{k,n} \geq \lambda_{k-1,n} \geq \ldots \geq \lambda_{1,n} \geq 0$, are the eigenvalues of $X_n'X_n$. The upper bound of this restriction is concerned to another restriction,

$$\lambda_{k,n+m}(X_{n+m}'X_{n+m}) = \lambda_{k,n}(X_n'X_n),$$

where, $\lambda_{k,n+m}(X_{n+m}'X_{n+m})$ and $\lambda_{k,n}(X_n'X_n)$ are the largest eigenvalue of $(X_{n+m}'X_{n+m})$ and $(X_n'X_n)$, respectively.

Our procedure is similar in spirit to that of Heiberger et al. (1993), to solve the problem of augmenting the new sites to a potential design in the eigenvector coordinate system. A major difference between our work and that of Heiberger et al. (1993) is in the manner of collecting additional observations; we propose a sequential choice of observations, whilst they suggest a simultaneous one. The sequential algorithms are able to revise the measures of interest associated with the design after each batch of l observations, whereas the simultaneous ones are not. As will be seen, the proposed method places each new point in the best

location on the experimental reign, in order to attain an alphabetical optimal design.

7-4. Proposed Methods in Different Design Regions

With as few new observations as possible yet without selecting points outside the region $\aleph$ and discarding the original data in the design, the interpretation of our optimality criterion says the objectives of data augmentation are,

I) to maximize the determinant of crossproduct $X'_{n+l}X_{n+l}$, G/D-optimality,

II) to maximize the smallest eigenvalue of $X'_{n+l}X_{n+l}$, E-optimality,

III) to minimize the inverse trace of $X'_{n+l}X_{n+l}$, A-optimality,

IV) to maximize the proportion of $|X'_{n+l}X_{n+l}|$ to $\max|X'_{n+l}X_{n+l}|$, orthogonality.

Alphabetic optimality criteria in our purposes (I), (II) and (III) were described in the previous section. Purpose (IV) indicates a measure to assess the orthogonality of the augmented design. The $\max|X'_{n+l}X_{n+l}|$ in (IV) indicates to the crossproduct determination for a desired design, which is fully orthogonal and its columns have similar length to those of the augmented design X_{n+l}. The evaluation of this quantity is given. The optimality and orthogonality criteria of the design are acquired by successively bringing up the smallest eigenvalue of $X'_{n+l}X_{n+l}$, for all l.

Let the diagnostic proportions of orthogonality and D-optimality for the new point $\boldsymbol{x}_{n+l}$ be denoted by γ_l and τ_l, respectively, where,

$$\gamma_l = |X'_{n+l}X_{n+l}| / \max|X'_{n+l}X_{n+l}| \ ; \ \tau_l = |X'_{n+l}X_{n+l}| / |X'_{n+l-1}X_{n+l-1}| .$$

In the set of designs X_{n+l}, the design $X^*_{n+l^*} = [X_n; X^*_{l^*}]$ is said to be an optimal design when $X^*_{l^*}$ is selected so as to attain a desired level for the proportion γ_l or τ_l with an optimal number of required additional observations, l^*. In the senses

of E and A- optimality criteria of $X'_{n+l}X_{n+l}$, the $X^*_{n+l^*}$ is an optimal design, as well.

7-4-1. Optimum Number of New Observations

It is well-known, for each two given matrices X_{n+l_1} and X_{n+l_2}, $l_1 < l_2$, $|X'_{n+l_1}X_{n+l_1}| \leq |X'_{n+l_2}X_{n+l_2}|$, where $X_{n+l_2} = (X_{n+l_1}; X_{l_2-l_1})$. As a result, exploiting more new sites in the design, a better design is obtained, in the concept in the magnitude of its crossproduct determination. However, there are constraints, for example economic constraints, to produce large number of observations. Thus, a stopping rule is required to add sequential new rows to the X_{n+l}, such that it determines the optimum number of new observations (l^*) to satisfy the orthogonality and D-optimality of the design matrix.

Stopping rule for adding data: The process of generating the new points will be stopped whenever the proportion γ_l to be equal or nearly equal to the predefined γ, where $0 < \gamma \leq 1$.

Extensive simulation experiments show, for $\gamma = 1$, there is a new point such as $\boldsymbol{x}_{n+\tilde{l}}$, where for previous new points as $\boldsymbol{x}_{n+l'}$, $1 \leq l' < \tilde{l}$, the relevant proportions $\gamma_{l'}$ and $\tau_{l'}$ significantly (respectively) increase and decrease toward 1. For subsequent points as $\boldsymbol{x}_{n+l''}$, $l'' > \tilde{l}$, the related $\gamma_{l''}$'s remain stable about 1, and the related $\tau_{l''}$'s slightly decrease toward 1. Based on the stopping rule for adding data, for $\gamma = 1$ we have $l^* = \tilde{l}$.

The achievement of the proportion $\gamma = 1$ is usually desired for a design, the stability of $\gamma_{l''}$'s about 1 is a reason to choose the proportion γ to construct the stopping rule. Simulation experiments also indicate that for small γ, the stopping

rule performs satisfactory to recognize the optimum number of new points to obtain a design that it has the predefined proportion τ.

In the sequel, to evaluate $\max|X'_{n+l}X_{n+l}|$, we make use of the theorem shown in chapter 6. As a consequence, we have $|V^{*\prime}V^*| = \max_V |V^{*\prime}V^*| = \prod_{j=1}^{k} c_{jj}$ if the $\underline{v}_j$ and $\underline{v}_{j'}$, for $j \neq j' = 1,2,...,k$, are orthogonal vectors. As noted, $\max|X'_{n+l}X_{n+l}|$ is equal to the determinant of crossproduct of a desired experimental design, where the length of each column of X_{n+l} and that one of the desired design is equal. In the case, we compare the quantity $|X'_{n+l}X_{n+l}|$ and its intended value, $\max|X'_{n+l}X_{n+l}|$, to know how well the augmented designs have been orthogonally constructed.

7-4-2. A General Consideration

Let the matrix of normalized eigenvectors and the diagonal matrix of eigenvalues of $X'_{n+l-1}X_{n+l-1}$ be defined by Ξ_{n+l-1} and Λ_{n+l-1}, where,

$$X'_{n+l-1}X_{n+l-1} = \Xi_{n+l-1}\Lambda_{n+l-1}\Xi'_{n+l-1},$$

$$\Lambda_{n+l-1} = \mathrm{diag}(\lambda_{1,n+l-1},...,\lambda_{j,n+l-1},...,\lambda_{k,n+l-1}), \quad \text{for} \quad \lambda_{k,n+l-1} \geq ... \geq \lambda_{j,n+l-1} \geq ... \geq \lambda_{1,n+l-1} \geq 0.$$

Under these circumstances, define $E_{n+l-1} = X_{n+l-1}\Xi_{n+l-1}$ and $E_{n+l} = X_{n+l}\Xi_{n+l-1}$, the augmented design matrices relevant to X_{n+l-1} and X_{n+l}, in the coordinate system of the eigenvectors of $X'_{n+l-1}X_{n+l-1}$. Since $X'_{n+l-1}X_{n+l-1}$ are symmetric, its eigenvectors are orthogonal and then Ξ_{n+l-1}'s are invertible. Sequentially, for $l = 1,2,...$, the matrix E_{n+l} is created by adding a point to E_{n+l-1}. Then at each stage the augmented design in the original coordinate system is evaluated to be $X_{n+l} = E_{n+l}\Xi^{-1}_{n+l-1}$, where both E_{n+l} and Ξ^{-1}_{n+l-1} are known.

Our prime goal is to maximize the smallest eigenvalue of $X'_{n+l}X_{n+l}$. Let Λ_{n+l} denote to the diagonal matrix of eigenvalues of $X'_{n+l}X_{n+l}$. We have the

equivalency of Λ_{n+l} and the eigenvalues matrix of $E'_{n+l}E_{n+l}$, this means the smallest eigenvalues of $X'_{n+l}X_{n+l}$ and $E'_{n+l}E_{n+l}$ are equal.

In our algorithm, to maximize the smallest eigenvalue of $E'_{n+l}E_{n+l}$, the new points one at a time are based on the direction of the eigenvector associated to the smallest eigenvalue of matrices $X'_{n+l-1}X_{n+l-1}$, for all l. Therefore, adding such points we obtain E^*_{n+l}, where the smallest of the eigenvalues of $E^{*'}_{n+l}E^*_{n+l}$ is maximum in the class of crossproduct matrix $E'_{n+l}E_{n+l}$. As a result, the smallest eigenvalue of $X^{*'}_{n+l}X^*_{n+l}$ is maximum, as well. In this case, the matrices X^*_{n+l} are equal to the matrices $E^*_{n+l}\Xi^{*-1}_{n+l-1}$, for all l. Implicitly, the Ξ^*_{n+l-1}'s are the matrices of the normalized eigenvectors of $X^{*'}_{n+l-1}X^*_{n+l-1}$'s.

7-4.3. Creating the New Observations

As we noted earlier, the lth new point is created in the eigenvectors coordinate of $X'_{n+l-1}X_{n+l-1}$. This point is called $\boldsymbol{e}_{n+l}$, where the related point to $\boldsymbol{e}_{n+l}$ in the original coordinate system is $\boldsymbol{x}_{n+l}$. Our proposed new points are $\boldsymbol{e}^*_{n+l}$'s, where $\boldsymbol{e}^*_{n+l}=(R,0,...,0)$ and $R=\sqrt{k}$ is radius of spherical region for the centered data. Thus the lth new point in the original coordinates is $\boldsymbol{x}^*_{n+l}=\boldsymbol{e}^*_{n+l}\Xi^{*-1}_{n+l-1}$, such that

$$\boldsymbol{x}^*_{n+l}\boldsymbol{x}^{*'}_{n+l}=\boldsymbol{e}^*_{n+l}\boldsymbol{e}^{*'}_{n+l}=R^2 \quad \text{for} \quad l\geq 1.$$

Equation $\boldsymbol{x}^*_{n+l}\boldsymbol{x}^{*'}_{n+l}=R^2$, for all l, suggests that the new points do not exceed the prespecified spherical region with the radius R.

Let $E_{n+l}=[E_{n+l-1};\boldsymbol{e}_{n+l}]$ denote to the design matrix in the eigenvectors coordinate system. We propose E^*_{n+l} when all additional points as $\boldsymbol{e}_{n+l}$ be equal to $\boldsymbol{e}^*_{n+l}$. The matrices $\boldsymbol{e}^{*'}_{n+l}\boldsymbol{e}^*_{n+l}$ increase the first components of $E^{*'}_{n+l-1}E^*_{n+l-1}$'s by adding

R^2's. Let $\lambda^*_{j,n+l-1}$'s denote to the eigenvalues of $E^{*'}_{n+l-1}E^*_{n+l-1}$, where $\lambda^*_{k,n+l-1} \geq ... \geq \lambda^*_{j,n+l-1} \geq ... \geq \lambda^*_{2,n+l-1} \geq \lambda^*_{1,n+l-1} \geq 0$. Since, the matrix $E^{*'}_{n+l}E^*_{n+l}$ is a diagonal matrix, thus its smallest eigenvalue equals

$$\lambda^*_{1,n+l} = \min(\lambda^*_{1,n+l-1} + R^2, \lambda^*_{2,n+l-1}),$$

As a consequence, the smallest eigenvalue of $X^{*'}_{n+l}X^*_{n+l}$ is increased from $\lambda^*_{1,n+l-1}$ to $\lambda^*_{1,n+l}$, for all l, where $X^*_{n+1} = [X_n; \boldsymbol{x}^*_{n+1}]$ and, for $l \geq 2$, $X^*_{n+l} = [X^*_{n+l-1}; \boldsymbol{x}^*_{n+l}]$.

7-4-4. The Best New Point

The new point based on the direction of the eigenvector associated to the smallest eigenvalue of $X^{*'}_{n+l-1}X^*_{n+l-1}$ (or equivalently, the smallest eigenvalue of $E^{*'}_{n+l-1}E^*_{n+l-1}$) provides the best new point in the original coordinate system in the predefined spherical region. Let $\boldsymbol{e}^G_{n+l} = (\alpha_1 R^{(1)}, ..., \alpha_j R^{(j)}, ..., \alpha_k R^{(k)})$ be the common definition of the lth augmenting point in the eigenvectors coordinate of $X^{G'}_{n+l}X^G_{n+l}$ in a k-dimensional sphere of radius R. Here, for $j = 1,2,...,k$, $R \geq R^{(j)} \geq 0$, $-1 \leq \alpha_j \leq 1$, $\sum_{j=1}^{k} \alpha_j^2 = 1$, and X^G_{n+l} is the augmented matrix related to $\boldsymbol{e}^G_{n+l}$'s. Simulated experiments are generated, 100,000 times, corresponding to the $n \times k$ initial design matrix X_n. In addition, the values of α_j and $R^{(j)}$ for all j are randomly generated to construct the matrix X^G_{n+l}. All experiments conform the inequalities below,

i) $|X^{G'}_{n+l}X^G_{n+l}| \leq |X^{*'}_{n+l}X^*_{n+l}|$,

ii) $|X^{G'}_{n+l}X^G_{n+l}| / \max|X^{G'}_{n+l}X^G_{n+l}| \leq |X^{*'}_{n+l}X^*_{n+l}| / \max|X^{*'}_{n+l}X^*_{n+l}|$,

where, the matrix X^*_{n+l} is based on the proposed points $\boldsymbol{e}^*_{n+l}$. These inequalities indicate that the points $\boldsymbol{e}^*_{n+l}$ perform better than $\boldsymbol{e}^G_{n+l}$'s to maximize the

determinant of information matrix, in (i), and to orthogonalize the experimental design, in (ii). Moreover, the simulation experiments show that the design X^*_{n+l} are the best in the batch of designs X^G_{n+l} for maximizing $|X^{G'}_{n+l}X^G_{n+l}|/|X_nX_n|$, for maximizing the smallest eigenvalue of $X^{G'}_{n+l}X^G_{n+l}$, and for minimizing the inverse trace of $X^{G'}_{n+l}X^G_{n+l}$.

7-4-5. Additional Points in a Particular Region

If the proposed algorithm is studied in a spherical region with radius $R=\sqrt{k}$, it will be indicated by $KP^{(1)}$, where the new points are $\boldsymbol{x}^*_{n+l}$'s.

Let $x^*_{j,n+l}$ for $j=1,2,...,k$ denote to the jth element of $\boldsymbol{x}^*_{n+l}$. In the cases, where the design space is constrained to a hypercube centered at the origin and bounded by the hyperplanes that orthogonally intersect the axes at -1 and 1, we propose the new point to be defined by $\boldsymbol{x}^{*r}_{n+l}$, where its elements are,

$$x^{*r}_{j,n+l}=\begin{cases}+1 & x^*_{j,n+l}\geq 0\\ -1 & x^*_{j,n+l}<0\end{cases}.$$

In such situations, the proposed procedure is expressed by $KP^{(2)}$. It should be suggested that $\boldsymbol{x}^{*r}_{n+l}\boldsymbol{x}^{*r'}_{n+l}=R^2=k$, this means the required additional points are not spaced outside the predefined rectangular region. The $KP^{(2)}$ is an alternate method to that one by Gaylor and Merrill (1968) and Dykstra (1971), for solving the augmenting existing data problem in the rectangular region.

In order to carry out a comparison between the method presented by Heiberger et al. (1993) and the proposed one, the new points $\boldsymbol{x}^{*c}_{n+l}=\boldsymbol{e}^{*c}_{n+l}\Xi^{*c-1}_{n+l-1}$ will be used. In this case, $\boldsymbol{e}^{*c}_{n+l}=(R_L,0,0,...,0)$ and Ξ^{*c}_{n+l-1} is the normalized eigenvectors

matrix corresponding to the augmented matrix X^{*c}_{n+l-1}. Implicitly, R_L is equal to larger radius of elliptical region given in Heiberger et al. (1993).

As already mentioned, the $\lambda_{1,n}$ and $\lambda_{k,n}$ are smallest and largest eigenvalues of $X_n'X_n$, respectively. Based on these quantities, the larger radius of elliptical regions is evaluated to be $R_L=[(\lambda_{k,n}-\lambda_{1,n})/2]^{1/2}$, $\sqrt{3}[(\lambda_{k,n}-\lambda_{1,n})/2]^{1/2}/2$, $[(\lambda_{k,n}-\lambda_{1,n})/2]^{1/2}/2$ and so on, where the volume of R_L relies on the number set of new points. For a small set of additional observations, R_L is a large value (for example $R_L=[(\lambda_{k,n}-\lambda_{1,n})/2]^{1/2}$) and for a large set of new data R_L is a small value (for example $R_L=[(\lambda_{k,n}-\lambda_{1,n})/2]^{1/2}/2$). We denote the proposed method by $KP^{(3)}$, when the recommended new points $\boldsymbol{x}^{*c}_{n+l}$ is exploited in a spherical region with radius R_L.

7-5. Other Existing Methods

In sections 7-3 and 7-4, we considered the existing and the proposed methods, which are based on alphabetical optimality criteria. These methods hold the original data in the improved design and optimize some functions of the information matrix $X'_{n+m}X_{n+m}$.

In the alphabetically optimal design literature, there are papers that discuss the data augmentation in the exchange algorithms. These algorithms start with an original design matrix and then add and remove one or more experiments in order to acquire increase in the determinate of the information matrix. These methods are similar in concept to that of the existing methods in section 7-3 and our proposed ones, which optimize the original design by augmenting the new points. But, the algorithms studied differ from those existing in the literature in that the latter proceed with discarding some of the existing observations. We

refer the interested user to the exchange methods presented by Fedorov (1972), Mitchell (2000), and Heredia-Laugner et al. (2003).

In situations that a prior distribution of regression parameters ($\underline{\beta}$) corresponds to previous information, the Bayesian alphabetical optimality criteria are exploited to construct an optimal design. The objectives of the use of the Bayesian algorithms are to optimize the functions of the posterior variance matrix of the parameters vector $\underline{\beta}$, i.e. $\sigma^2(X_m'X_m+\Phi)^{-1}$, where the $k\times k$ matrix Φ is known. In this case, the prior variance matrix of $\underline{\beta}$ is $\sigma^2\Phi^{-1}$. In the present chapter, we focus on the alphabetical optimality criteria that have been discussed for non-Bayesian optimal design, in section 7-2. Considerations on the Bayesian alphabetical criteria and the Bayesian optimal designs are given by DeGroot and Goel (1979), and in the recent studies by Khuri et al. (2006), and Dror and Steinberg (2008).

7-6. Example and Comparison

Let us represent data given by Gaylor and Merrill (1968) in Table 7.1. The natural variables are denoted by ξ_1, ξ_2, and ξ_3, where the experimental region of interest is a cube limited by

$$-7\le\xi_1\le 6\ ;\ -7\le\xi_2\le 7\ ;\ -5\le\xi_3\le 12\,.$$

Then, the coded variables are defined by,

$$x_i=[\xi_i-(L_i+U_i)/2]/[(U_i-L_i)/2]\quad\text{for}\quad i=1,2,3\,,$$

such that, U_i and L_i are upper and lower bounds for ξ_i.

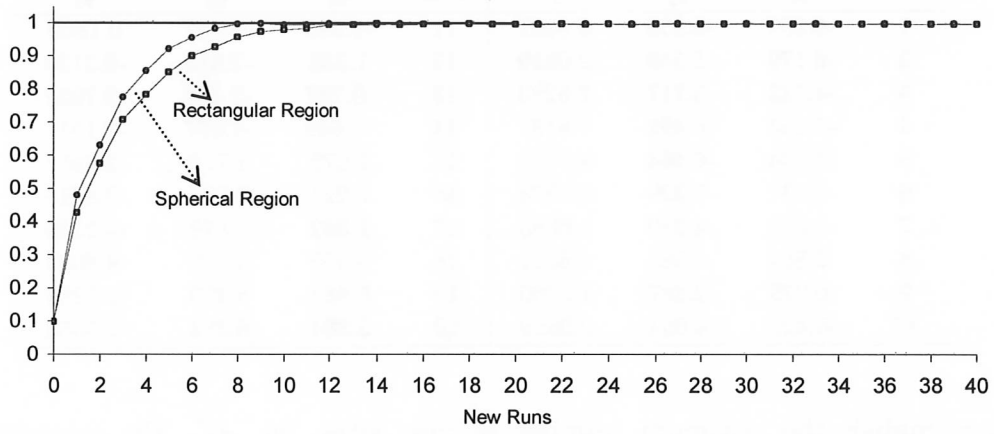
Table 7.1: The Natural Data by Gaylor and Merrill (1968).

Run	ξ_1	ξ_2	ξ_3	Run	ξ_1	ξ_2	ξ_3
1	-6.389	-5.330	6.0437	11	-1.593	-3.957	0.1869
2	-6.179	-5.549	9.0819	12	1.338	-2.613	-0.3136
3	-4.533	-5.717	7.6283	13	-0.787	-2.487	-2.7032
4	-5.293	-6.492	7.8131	14	-1.649	-1.077	-2.1917
5	-4.004	-6.464	3.0976	15	2.075	1.719	-2.2917
6	-2.631	-5.320	5.4978	16	2.224	0.946	-2.8516
7	-3.012	-4.080	1.0688	17	2.382	3.879	-4.2335
8	-2.864	-4.583	4.5822	18	3.350	3.510	-4.8033
9	-0.979	-2.887	1.0250	19	5.984	6.499	-2.1206
10	-0.420	-4.094	2.2669	20	3.384	6.383	-2.5354

To accomplish the optimum number of new sites, $m_{op} = l^*$, the stopping rule presented in the preceding section is employed. As a result, using the methods $KP^{(1)}$ and $KP^{(2)}$ we have $m_{op} = 8$ and $m_{op} = 12$, when the design space are a sphere with radius $R = \sqrt{3}$ and a rectangle limited by 1 and -1 in each dimension, respectively.

Figure 7.1 shows an increase of the orthogonality proportion i.e. γ_l, when there adding points exist. The proportion γ_l is significantly increased for $1 \le l \le 8$ and $1 \le l \le 12$, whereas for the small number of observations behind points 8 and 12, the proportions of orthogonality have slight increases toward 1. When $l \to \infty$ then $\gamma_l \simeq 1$.

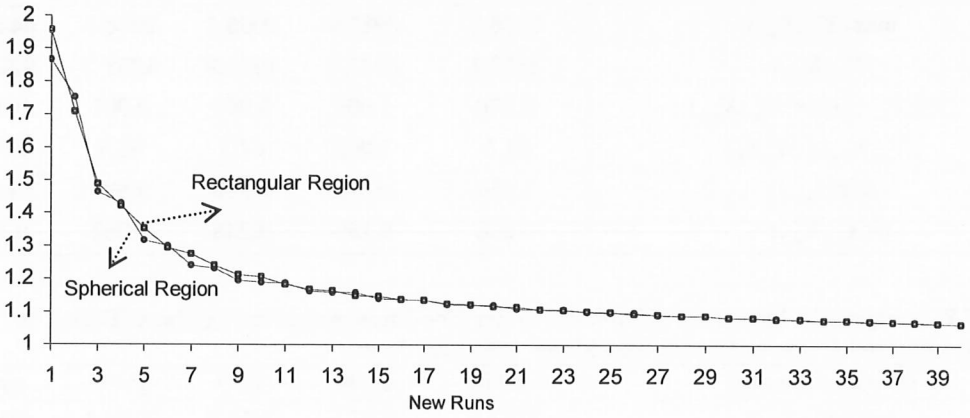
Figure 7.1: The Proportions of Orthogonality for Centered Data by Gaylor and Merrill (1968). (Number of New Data $l=1,2,...,40$).



The D-optimality criterion or the proportions τ_l are illustrated by Figure 7.2. The slope of curves is considerably sharp from the new observations 1 to 8 on the spherical region and from 1 to 12 on the rectangular region. The slopes are relatively smooth for next observations. When $l \to \infty$, $\tau_l \to 1$.

Based on these two figures, for the desired $\gamma = 1$ and $\tau = 1$, the optimum number $l^* = 8$ and $l^* = 12$ are determined for the spherical and rectangular region. As a result, the optimum numbers of adding points are similarly achieved either to orthogonalize the matrix X_{n+l} or to maximize the determinant of crossproduct $X'_{n+l}X_{n+l}$.

Figure 7.2: The proportions of $|X_{n+l}^{*'}X_{n+l}^{*}|$ to $|X_{n+l-1}^{*'}X_{n+l-1}^{*}|$ for Centered Data by Gaylor and Merrill (1968). (Number of New Data $l=1,2,...,40$).



According to the studied strategies by Gaylor and Merrill (1968), (*GM*), Dykstra (1971), (*Dy*), Heiberger, Bhaumik, and Holland (1993), (*HBH*), and our proposed strategies, comparisons to the measures of interest of the design matrix will be provided by using Table 7.2. The new experiments $\boldsymbol{x}_{n+l}^{*}$, $\boldsymbol{x}_{n+l}^{*r}$, and $\boldsymbol{x}_{n+l}^{*c}$ are generated by the proposed procedures ($KP^{(1)}$, $KP^{(2)}$, and $KP^{(3)}$), on the three distinct regions, which were suggested in section 7-4.

In this example, to construct the new point $\boldsymbol{x}_{n+l}^{*c}$, when the number of adding points is 4, 6, 8, 12, and 16, the larger radius of elliptical regions is calculated 2.7745, 2.4356, 1.9618, 1.7223, and 1.3873, respectively.

Table 7.2: Measures of Interest Under Condition Augmenting Data (Based on Original Centered Data by Gaylor and Miller 1968) for $l=m=4$.

Measure / Method	*GM / Dy*	*HBH*	$KP^{(1)}$	$KP^{(2)}$	$KP^{(3)}$
$\max\|X'_{n+l}X_{n+l}\|$	1227.7	3987.9	1227.9	1227.7	4917.8
$\|X'_{n+l}X_{n+l}\|$	962.4	3987.9	1049.3	962.4	4875.8
$\|X'_{n+l}X_{n+l}\|/\max\|X'_{n+l}X_{n+l}\|$	0.784	1.000	0.855	0.784	0.991
$\|X'_{n+l}X_{n+l}\|/\|X'_{n}X_{n}\|$	32.9	136.3	35.9	32.9	166.6
$\min(\lambda_{j,n+l})$	6.711	15.858	6.993	6.711	15.858
$\mathrm{tr}(X'_{n+l}X_{n+l})^{-1}$	0.328	0.189	0.312	0.328	0.178

Table 7.2 *(continued)*: Measures of Interest Under Condition Augmenting Data (Based on Original Centered Data by Gaylor and Miller 1968) for $l = m = 6$.

Measure / Method	*GM / Dy*	*HBH*	$KP^{(1)}$	$KP^{(2)}$	$KP^{(3)}$
$\max\lvert X'_{n+l}X_{n+l}\rvert$	2056.8	3987.9	2066.7	2056.8	6448.8
$\lvert X'_{n+l}X_{n+l}\rvert$	1800.7	3987.9	1974.9	1852.2	6309.3
$\lvert X'_{n+l}X_{n+l}\rvert / \max\lvert X'_{n+l}X_{n+l}\rvert$	0.876	1.000	0.956	0.901	0.978
$\lvert X'_{n+l}X_{n+l}\rvert / \lvert X'_nX_n\rvert$	61.5	136.3	67.5	63.3	215.6
$\min(\lambda_{j,n+l})$	8.459	15.858	9.993	9.693	15.858
$\mathrm{tr}(X'_{n+l}X_{n+l})^{-1}$	0.258	0.189	0.243	0.253	0.164

Table 7.2 *(continued)*: Measures of Interest Under Condition Augmenting Data (Based on Original Centered Data by Gaylor and Miller 1968) for $l = m = 8$.

Measure / Method	*GM / Dy*	*HBH*	$KP^{(1)}$	$KP^{(2)}$	$KP^{(3)}$
$\max\lvert X'_{n+l}X_{n+l}\rvert$	3192.4	3987.9	3211.1	3192.4	4917.8
$\lvert X'_{n+l}X_{n+l}\rvert$	3055.3	3987.9	3185.9	3055.3	4875.8
$\lvert X'_{n+l}X_{n+l}\rvert / \max\lvert X'_{n+l}X_{n+l}\rvert$	0.957	1.000	0.992	0.957	0.991
$\lvert X'_{n+l}X_{n+l}\rvert / \lvert X'_nX_n\rvert$	104.4	136.3	108.9	104.4	166.6
$\min(\lambda_{j,n+l})$	11.715	15.858	12.993	11.715	15.858
$\mathrm{tr}(X'_{n+l}X_{n+l})^{-1}$	0.210	0.189	0.205	0.210	0.178

Table 7.2 *(continued)*: Measures of Interest Under Condition Augmenting Data (Based on Original Centered Data by Gaylor and Miller 1968) for $l = m = 12$.

Measure / Method	*GM / Dy*	*HBH*	$KP^{(1)}$	$KP^{(2)}$	$KP^{(3)}$
$\max\lvert X'_{n+l}X_{n+l}\rvert$	6575.2	3987.9	6613.5	6575.2	6471.4
$\lvert X'_{n+l}X_{n+l}\rvert$	6548.3	3987.9	6612.5	6548.3	6469.9
$\lvert X'_{n+l}X_{n+l}\rvert / \max\lvert X'_{n+l}X_{n+l}\rvert$	0.996	1.000	1.000	0.996	1.000
$\lvert X'_{n+l}X_{n+l}\rvert / \lvert X'_nX_n\rvert$	223.8	136.3	226.0	223.8	221.1
$\min(\lambda_{j,n+l})$	17.564	15.858	18.462	17.564	18.259
$\mathrm{tr}(X'_{n+l}X_{n+l})^{-1}$	0.161	0.189	0.160	0.161	0.161

Table 7.2 *(continued)*: Measures of Interest Under Condition Augmenting Data (Based on Original Centered Data by Gaylor and Miller 1968) for $l = m = 16$.

Measure / Method	*GM / Dy*	*HBH*	$KP^{(1)}$	$KP^{(2)}$	$KP^{(3)}$
$\max\lvert X'_{n+l}X_{n+l}\rvert$	11760.0	3987.9	11794.9	11760.0	4939.5
$\lvert X'_{n+l}X_{n+l}\rvert$	11658.6	3987.9	11759.4	11727.0	4924.8
$\lvert X'_{n+l}X_{n+l}\rvert / \max\lvert X'_{n+l}X_{n+l}\rvert$	0.991	1.000	0.997	0.997	0.997
$\lvert X'_{n+l}X_{n+l}\rvert / \lvert X'_nX_n\rvert$	398.5	136.3	401.9	400.8	168.3
$\min(\lambda_{j,n+l})$	20.485	15.858	21.858	21.564	15.858
$\mathrm{tr}(X'_{n+l}X_{n+l})^{-1}$	0.133	0.189	0.132	0.132	0.177

As already suggested, with respect to the equivalency of D and G-optimality, the methods GM and Dy are equivalent. We also observed this equivalency for our example. The results associated with these methods are given in the column of Table 7.2 entitled GM/Dy. The recommendation is that in order to overcome the problems of rounding rule by the algorithm GM, its equivalent method i.e. Dy be used to establish a optimal design.

Table 7.2 implies that the measures of interest obtained by the procedure $KP^{(1)}$, which is based on the spherical region with radius $R=\sqrt{3}$, are more adequate than one of the methods GM/Dy. Furthermore, the rectangular method $KP^{(2)}$ is relatively more efficient to construct a augmented design than the methods GM/Dy.

The strategy HBH makes the augmented matrix X_{n+m} completely orthogonal i.e. $|X'_{n+m}X_{n+m}|/\max|X'_{n+m}X_{n+m}|=1$, for all m. For the small number of additional points, $m=4,6,8$, the elliptical strategy HBH appears to perform much better than the methods GM/Dy, $KP^{(1)}$ and $KP^{(2)}$. However, the HBH is not suitable for large numbers of new data, $m=12,16$, when compared to the schemes GM/Dy, $KP^{(1)}$ and $KP^{(2)}$, in the spirit of alphabetic optimality criteria. In this table, the G/D, E, and A-optimality criteria are mentioned in rows $|X'_{n+l}X_{n+l}|/|X'_nX_n|$, $\min(\lambda_{j,n+l})$, and $\mathrm{tr}(X'_{n+l}X_{n+l})^{-1}$, respectively. The spherical strategy $KP^{(3)}$ seems to perform better than the elliptical strategy HBH for either small or large number of m, but, it does not appear to produce better results than the spherical method $KP^{(1)}$ for the large m.

Consequently, if we have a restriction for the new observations in a predefined rectangular region, the strategy $KP^{(2)}$ is proposed, where the new sites are spaced in corner points. When the experimental region has been restricted, to add the new observations in a sphere, the strategy $KP^{(1)}$ is recommended, so that the new points are on the circumference of the k-dimensional sphere with $R=\sqrt{k}$. Moreover, $R=\sqrt{k}$ in the $KP^{(1)}$ is equal to the distance of corner points in the rectangular region from the origin point of the region, but for a small set of new points R_L in the method HBH is greater R. This indicates to a disadvantage of the HBH, where the new points in the HBH for small m are outside the predefined rectangular region. Finally, according to a comparison to the spherical strategies $KP^{(1)}$ and $KP^{(3)}$, it is concluded that the spherical strategy in region with large radius performs better than the method with small radius.

7-7. Conclusion

Some new methods have been introduced to augment the optimum number of the new points to orthogonalize the original design and to maximize the determinant of the crossproduct of independent regressors. These methods are also useful in maximizing the smallest eigenvalue of the information matrix and in minimizing the expected squared distance from the least squares estimator to the true parameter vector of model. When the experimenter is restricted to add the new observations in the predefined rectangular region, then we propose the strategy $KP^{(2)}$ for the construction of optimal designs. In case the experimenter is restricted to add the new observations in a spherical region, we suggest the use of the strategy $KP^{(1)}$. The spherical strategies in a region with a larger radius

appear to perform better than spherical strategies in a region with a smaller radius.

8

Extensions to Modified Gram-Schmidt Strategy and Its Application in Steepest Ascent Method

CHAPTER OUTLINE

8-1. Introduction

The present chapter introduces an efficient method (the Modified Gram-Schmidt (MGS) orthogonalization through Gaussian elimination) for solving linear squares problem as an alternative method to the data augmentation scheme, which was described in the preceding chapter. The experimenters often use interchanges in the process of Gaussian elimination, but the proposed algorithm is carried out without any interchanges of rows of the original design matrix X.

The proposed algorithm provides an orthogonal basis, in full working accuracy, for the space spanned by the columns of original design matrix. We present an application of the proposed MGS algorithm in the steepest ascent process. It is seen that operating our proposed method, the covariance and the bias of the estimated path of steepest ascent (SA) are efficiently decreased.

8-2. The Proposed Algorithm of the MGS Orthogonalization

Define jth nonzero column vector of matrix X in the Euclidean n-space by $\underline{x}_j = (x_{1j}, x_{2j}, ..., x_{nj})'$ for $j = 1,2,...,k$, where the columns are linear independent. Gaussian elimination is carried out on $X'X \mid X'$ by using the following algorithms (I) and (II), such that the algorithms are based on the MGS process. Under this situation, Gaussian elimination turns $X'X \mid X'$ into $R \mid Q'$ such that R is an invertible upper triangular matrix of $X'X$ and Q an orthonormal matrix ($Q'Q = I_k$) of original matrix X. Algorithm (I) is employed to construct an echelon form of the matrix $X'X$, i.e. the matrix R. This algorithm resolves the information matrix $X'X$ into $X'X = R'R$. Alternative strategies to the proposed algorithms are the singular vector QB (SVQB) factorization and the Cholesky QR (CholQR) factorization. These methods are based on the block orthogonalization

procedure, which have some interesting properties. However, these methods are not as stable as the MGS strategy. See the explicit descriptions on the SVQB and the CholQR factorization by Stathopoulos and Wu (2002).

Denote,

$$X'X=[c_{jj'}]=\begin{bmatrix}C_1\\C_2\\...\\C_k\end{bmatrix};\ X'=[x_{ji}]=\begin{bmatrix}X_1\\X_2\\...\\X_k\end{bmatrix};\ R=[r_{jj'}]=\begin{bmatrix}R_1\\R_2\\...\\R_k\end{bmatrix};\ Q'=[q_{ji}]=\begin{bmatrix}Q_1\\Q_2\\...\\Q_k\end{bmatrix},$$

for j and $j'=1,2,...,k$, and $i=1,2,...,n$.

Algorithm (I)

The entries of R are sequentially evaluated by the following scheme:

$$R_j=R_j^*/\left\|Q_j^*\right\|, \qquad \text{for} \qquad j=1,2,...,k,$$

where, $R_1^*=C_1$, $Q_1^*=X_1$, $R_j^*=C_j-\sum_{l=1}^{j-1}r_{lj}R_l$ and $Q_j^*=X_j-\sum_{l=1}^{j-1}r_{lj}Q_l^*/r_{ll}$ for $j=2,...,k$.

Algorithm (II)

The rows of matrix Q' are constructed the corresponding known components of matrices R and X, where

$$Q_1=X_1/r_{11},$$

$$Q_j=[X_j-\sum_{l=1}^{j-1}r_{lj}Q_l]/r_{jj} \qquad \text{for} \qquad j=2,...,k.$$

Indeed, Algorithm (II) is a backward solution that performs $Q=XR^{-1}$. Hence to prevent from doing operations in the second stage, the use of backward solution is preferable instead of its related algorithm. Algorithm (I) also involves more flops than the usual MGS method. As an alternative, the following program in Maple is based on the current algorithm of the MGS process, which includes fewer flops than one of the proposed algorithms. This program and that one

based on the algorithms above make the same Q and R matrices. The proposed algorithms (I) and (II) are simple and useful for the students who have a real need to understand the geometry of the MGS process.

```
# Maple code for the MGS algorithm.
with(LinearAlgebra):
n:=4; k:=3; X:=HilbertMatrix(k,n); Q:=X-X; V:=X-X;
Q[1,1..n]:=X[1,1..n]/VectorNorm(X[1,1..n],2);
for j from 2 to k do;
   S[1,1..n]:=Transpose(ZeroVector(n));V[j,1..n]:=X[j,1..n];
   for i from 1 to j-1 do;
      S[1,1..n]:=S[1,1..n]-(V[j,1..n].Transpose(Q[i,1..n])).Q[i,1..n];
   end do;
   V[j,1..n]:=X[j,1..n]+S[1,1..n];
   Q[j,1..n]:=V[j,1..n]/VectorNorm(V[j,1..n],2);
end do;
X:=Transpose(X); Q:=Transpose(Q); R:=Transpose(Q).X;
Qhat:=evalf(Q,16); Rhat:=evalf(R,16); # Qhat and Rhat are the approximations of Q and R.
```

Usually, the columns length of Q and X is not equal. Denote a homo-length MGS orthogonal form of the original X by Q_h. Defining $U = diag\,(\sqrt{c_{11}}, \sqrt{c_{22}}, \ldots, \sqrt{c_{kk}})$, this gives $Q_h = QU$, such that jth column length of X and Q_h, for all j, are equal.

8-3. The Accuracy of Proposed Method

The entries of R and Q are evaluated, using the proposed program. Extensive numerical experiments have shown that the lower components of R and the off-diagonal elements of $Q'Q$ are zero, exactly. This means to say that the proposed method quite strong to obtain the orthonormal matrix Q. However, all of the algorithms already designed work perfectly well within the Maple or Mathematica exact arithmetic. But, the existing algorithms were usually constructed in Matlab or Fortran in the finite-precision arithmetic through the

process computing, where the Q and R obtained in these cases are not full orthogonal and triangular, respectively.

Often, one or more reinforcement of the MGS scheme in the existing literature has been discussed to improve the orthogonality of vectors set, whereas, utilizing the proposed program any reinforcement is not required. Different schemes for the reinforcement, including reorthogonalization and iterative refinement, were described by Golub and Wilkinson (1966), Fletcher (1975), Björck (1978), Giraud, et al. (2005), and Smoktunowicz et al. (2006). All types of reinforcements are used for reducing the rounding errors in the computed Q and R. The rounding error problem is due to the constraints on the significant digits in the computer software utilized or/and the arithmetic operators, in particularly the division and square root operators ($/$ and $\sqrt{\ }$). Using the proposed program, the experimenters will be able to overcome the rounding error problem, such that its computations can be based on more than 100,000 significant digits and the exact arithmetic operators.

In the present chapter, define the floating point format or equivalently the approximations of Q and R by $\tilde{Q}$ and $\tilde{R}$. Once, the matrices Q and R were exactly evaluated according to real operations, if the $\tilde{Q}$ and $\tilde{R}$ are desired for future applications, then the floating point operations will be carried out only on the Q and R already evaluated. Under such conditions, the users only once encounter with the difficulties of rounding errors in order to accomplish the approximations $\tilde{Q}$ and $\tilde{R}$. Hence, it can be expected that using this strategy, we will have the maximum absolute value of the entries of $\tilde{Q}'\tilde{Q}-I_k$ and $\tilde{Q}\tilde{R}-X$ smaller than one of the strategies that carry out the floating point operations

from beginning to end of the process computing. Simulation experiments confirm the expectation mentioned for each $n \times k$ matrix X, particularly for large n and k. In one such case that the finite-precision arithmetic is applied thoroughly in the process calculation, the approximations are indicated by $\tilde{\tilde{Q}}$ and $\tilde{\tilde{R}}$.

8-4. QR Decomposition

In a general condition of the OLS problem, the normal equations $X'X\hat{\beta} = X'\underline{y}$ are ill-conditioned i.e. small errors in the calculations of the elements of $X'X$ can sometimes cause relatively large errors. When the columns of non-orthogonal matrix X are linearly independent, the OLS solution can often be computed more reliably through the QR decomposition of X. Substituting $X = QR$ into normal equations, $X'X\hat{\beta} = X'\underline{y}$ becomes $R\hat{\beta} = Q'\underline{y}$. Because of R is upper triangular, solving $R\hat{\beta} = Q'\underline{y}$ is very straightforward.

However, when the normal equations are based on the orthogonal Q_h given in section (8-2), instead of the original X, these equations are well-conditioned. In such case, since Q_h is orthogonal, thus a matrix decomposition is not required.

Assume that the n observations on the design X have already been taken and the normal equations are ill-conditioned. Moreover, assume that the statistician is not permitted to collect new observations on the orthogonal design Q_h. Under these circumstances, the use of the QR-factorization according to the proposed program is recommended to solve the ill-conditioned problem to attain the least squares estimators.

8-5. Covariance and Bias of the SA Path

The orthogonalization of a design matrix is used in the RSM to reduce the covariance and the bias of the estimated path. When we are at a point on the response surface that is remote from the optimum, such as the current operating conditions, there is little curvature in the system and the first order model will be appropriate. Our objective here is to lead the experimenter rapidly and efficiently along a path of improvement, path of SA, toward the general vicinity of the optimum.

As noted, previously the steps along the path are proportional to the regression coefficients of the first order model,

$$y = \beta_1 x_1 + \beta_2 x_2 + ... + \beta_j x_j + ... + \beta_k x_k + \varepsilon . \tag{1}$$

If the coefficients of this model are unknown, then they are estimated by the OLS estimators. Based on the coded variables x_j, the generalized form of model (1) are denoted by the model

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} , \tag{2}$$

where $\underline{y}$ is a $n \times 1$ vector of responses, X is a $n \times k$ matrix of coded variables, $\underline{\beta}$ is a $k \times 1$ vector of regression coefficients, and $\underline{\varepsilon}$ is a $n \times 1$ vector of errors ($\mathrm{E}(\underline{\varepsilon}) = \underline{0}$ and $\mathrm{cov}(\underline{\varepsilon}) = \sigma^2 I_k$). Based on the model (2), the ordinary least squares (OLS) estimator of $\underline{\beta}$ and its covariance are,

$$\underline{\hat{\beta}} = (X'X)^{-1} X'\underline{y} , \tag{3}$$

$$\mathrm{cov}(\underline{\hat{\beta}}) = \sigma^2 (X'X)^{-1} . \tag{4}$$

For the unknown coefficients vector $\underline{\beta}$, the OLS estimator and its covariance are given by equations (3) and (4). According to corollary (1) in

chapter (6), for the matrix X^* in the set of X's with squared lengths c_{jj} for $j=1,2,...,k$, the minimization of $|(X^{*\prime}X^*)^{-1}|$ is given by

$$|(X^{*\prime}X^*)^{-1}| = \min_X |(X^{*\prime}X^*)^{-1}| = (\prod_{j=1}^{k} c_{jj})^{-1}. \quad (5)$$

When the original matrix X is turned into $Q_h = X^*$, the original model $\underline{y} = X\underline{\beta} + \underline{\varepsilon}$ is replaced by

$$\underline{y}^* = Q_h\underline{\beta}^* + \underline{\varepsilon}^*, \quad (6)$$

where $E(\underline{\varepsilon}^*) = \underline{0}$ and $\text{var}(\underline{\varepsilon}^*) = \sigma^2 I_k$. Because of $|Q_h'Q_h| \geq |X'X|$, thus the Q_h is a D-optimal design. Utilizing the theorem and the corollary given in chapter (6), the magnitude of covariance matrix in favor of Q_h is equal to the minimum of the original one,

$$|\text{cov}(\underline{\hat{\beta}}^*)| = \sigma^2 |(Q_h'Q_h)^{-1}| = \sigma^2 / \prod_{j=1}^{k} c_{jj} = \sigma^2 \min |(X'X)^{-1}|.$$

As a result, $|\text{cov}(\underline{\hat{\beta}}^*)| \leq |\text{cov}(\underline{\hat{\beta}})|$. This means that the SA path based on the model (6) has the variation less than or equal to that one of the model (2).

The bias of the SA path is examined by computing the absolute value of the angle between the true and the estimated path,

$$\theta = |\arccos(< \underline{\beta}, \underline{\hat{\beta}} > / [|\underline{\beta}||\underline{\hat{\beta}}|])|. \quad (7)$$

Let us denote the angle θ for the paths based on the models (2) and (6) by θ_I and θ_o, respectively. The angles θ_I and θ_o will be compared using simulation experiments in the next section.

8-6. Example

An engineer is interested in obtaining more uniform fill heights in the bottles produced by her manufacturing process. The filling machine theoretically fills

each bottle to the correct target height, but in practice, there is variation around this target, and the bottler would like to understand better the sources of this variability and eventually reduce it.

The process engineer can control three variables during the filling process: the percent carbonation (factor A), the operating pressure in the filler (factor B), and the line speed (factor C). She is currently operating the process with factor A of 25 percent, factor B of 30 psi, and factor C of 205 bpm. Because it is unlikely that this region contains the optimum, she fits a first order model and applies the method of SA. In the case, the region of exploration for fitting the first order model is (20, 30) percent of factor A, (20, 40) psi of factor B and (190, 220) bpm of factor C.

If ξ_1, ξ_2 and ξ_3 denote the natural variables the percent carbonation, the operating pressure and the line speed, respectively, then the coded variables are,

$x_1 = (\xi_1 - 25)/5$; $x_2 = (\xi_2 - 30)/10$; $x_3 = (\xi_3 - 205)/15$.

She decides to run experiments in these three factors, with all 6 runs taken in random order. Assuming $\underline{\beta} = (1.35, 0.82, -4.1)'$, $\sigma^2 = 70.2$, $\text{cov}(\underline{\varepsilon}) = 70.2 I_k$ and $\text{E}(\underline{\varepsilon}) = \underline{0}$, the simulation experiments are created. The data that resulted from these experiments are shown in Table 8.1. Positive deviations are fill heights above the target, whereas negative deviations are fill heights below the target.

Table 8.1: Fill Height Runs for the Filling Process of Bottle (Initial Design).

Run	ξ_1	ξ_2	ξ_3	x_1	x_2	x_3	y
1	27	36	214	2/5	3/5	3/5	9.22
2	23	23	199	-2/5	-2/3	-2/5	14.59
3	26	34	208	1/5	2/5	1/5	0.13
4	24	31	199	-1/5	1/10	-2/5	2.95
5	27	36	211	2/5	3/5	2/5	-13.71
6	28	33	216	3/5	1/4	3/4	-14.34

Based on Table 8.1 we get,

$$X' = \begin{bmatrix} 2/5 & -2/5 & 1/5 & -1/5 & 2/5 & 3/5 \\ 3/5 & -2/3 & 2/5 & 1/10 & 3/5 & 1/4 \\ 3/5 & -2/5 & 1/5 & -2/5 & 2/5 & 3/4 \end{bmatrix} ; \; X'X = \begin{bmatrix} 23/25 & 287/300 & 113/100 \\ 287/300 & 5029/3600 & 1313/1200 \\ 113/100 & 1313/1200 & 577/400 \end{bmatrix}.$$

The proposed MGS method upon $[X'X \,|\, X']$ yields $[R \,|\, Q']$, where

$$R = \begin{bmatrix} \frac{\sqrt{23}}{5} & \frac{287\sqrt{23}}{1380} & \frac{113\sqrt{23}}{460} \\ 0 & \frac{\sqrt{765854}}{1380} & \frac{-279\sqrt{765854}}{1914635} \\ 0 & 0 & \frac{\sqrt{4246860218}}{332980} \end{bmatrix}, \; Q = \begin{bmatrix} \frac{2\sqrt{23}}{23} & \frac{127\sqrt{765854}}{382927} & \frac{24259\sqrt{4246860218}}{2123430109} \\ -\frac{2\sqrt{23}}{23} & -\frac{173\sqrt{765854}}{382927} & \frac{6807\sqrt{4246860218}}{2123430109} \\ \frac{\sqrt{23}}{23} & \frac{265\sqrt{765854}}{765854} & -\frac{2343\sqrt{4246860218}}{4246860218} \\ -\frac{\sqrt{23}}{23} & \frac{425\sqrt{765854}}{765854} & -\frac{30773\sqrt{4246860218}}{4246860218} \\ \frac{2\sqrt{23}}{23} & \frac{127\sqrt{765854}}{382927} & -\frac{9039\sqrt{4246860218}}{2123430109} \\ \frac{3\sqrt{23}}{23} & -\frac{258\sqrt{765854}}{382927} & -\frac{10347\sqrt{4246860218}}{2123430109} \end{bmatrix}.$$

In addition, the square root of diagonal elements of the information matrix $X'X$ gives $U = diag(\sqrt{23}/5, \sqrt{5029}/60, \sqrt{577}/20)$, therefore,

$$Q_h = QU = \begin{bmatrix} \frac{2}{5} & \frac{127\sqrt{3851479766}}{22975620} & \frac{24259\sqrt{24504383457860}}{42468602180} \\ -\frac{2}{5} & -\frac{173\sqrt{3851479766}}{22975620} & \frac{6807\sqrt{24504383457860}}{42468602180} \\ \frac{1}{5} & \frac{53\sqrt{3851479766}}{9190248} & -\frac{2343\sqrt{24504383457860}}{84937204360} \\ -\frac{1}{5} & \frac{85\sqrt{3851479766}}{9190248} & -\frac{30773\sqrt{24504383457860}}{84937204360} \\ \frac{2}{5} & \frac{127\sqrt{3851479766}}{22975620} & -\frac{9039\sqrt{24504383457860}}{42468602180} \\ \frac{3}{5} & -\frac{43\sqrt{3851479766}}{3829270} & -\frac{10347\sqrt{24504383457860}}{42468602180} \end{bmatrix}.$$

In this case, (a), which the Q_h is constructed according to real operations,

$$Q_h'Q_h = \begin{bmatrix} 23/25 & 0 & 0 \\ 0 & 5029/3600 & 0 \\ 0 & 0 & 577/400 \end{bmatrix}.$$

This indicates that the column vectors of Q_h are exactly orthogonal. (b) Employing the floating point operations upon the Q_h in order to obtain the approximation $\tilde{Q}_h$, we get,

$$\tilde{Q}_h'\tilde{Q}_h = \begin{bmatrix} 0.92 & 0.18e-15 & 0.6e-16 \\ 0.18e-15 & 1.397 & -0.88e-16 \\ 0.21e-16 & -0.88e-16 & 1.4425 \end{bmatrix},$$

where the significant digits of the calculations was defined equal to 16 i.e. $Digits := 16$.

(c) As an alternative for the method (b), if the approximation of Q_h to be computed, under the condition that the floating point operations to be carried out in thorough the process computing ($Digits := 16$) then,

$$\tilde{\tilde{Q}}_h'\tilde{\tilde{Q}}_h = \begin{bmatrix} 0.92 & 0.3e-14 & 0.149e-13 \\ 0.3e-14 & 1.399 & -0.211e-13 \\ 0.149e-13 & -0.211e-13 & 1.4425 \end{bmatrix}.$$

As a result, the strategy (b) that the floating point operations are only done on Q_h has relatively better performance than its alternative method i.e. (c). With regard to large design matrices that consisted of, for example, 20 rows and 20 columns the strategy (b) produces considerably more orthogonal matrix. For instance, to orthogonalize the Hilbert matrix of order 20, the scheme (b) produced ($Digits := 16$) the maximum absolute value of the off-diagonal entry of $\tilde{Q}_h'\tilde{Q}_h$ equal to $7.7e-17$ compared with $1.5e-1$ for the scheme (c).

Numerical experiments show the strategy (b) is well-stable, whereas the strategy (c) is not. To illustrate the performance of these strategies, Table 8.1 and Table 8.2 are established for the $n \times k$ matrices $X1$ and $X2$ with following Maple code.

 # Maple code for the matrix X1.

```
with(LinearAlgebra):
n:=100; k:=50; X1:= HilbertMatrix(k,n);
for i from k-1 to k do;
   for j from 1 to n do; X1[k-1,j]:=1; X1[i,j]:=j; end do;
end do; X1:=Transpose(X1);

# Maple code for the matrix X2.
with(LinearAlgebra): n:=100; k:=50; c:=10; d:=RandomMatrix(k,k);
for i from 1 to k do;
 for j from 1 to k do; if i=j then di[i,j]:=10^(c*((i-1)/(k-1))); else d[i,j]:=0; end if; end do;
end do;
o1:= RandomMatrix(k,k); (Q1,R1):=QRDecomposition(o1);
o2:= RandomMatrix(n,k); (Q2,R2) := QRDecomposition(o2);
X2:= Q1.d.Transpose(Q2); X2:=Transpose(X2);
```

In these tables, $\tilde{Q}$ and $\tilde{\tilde{Q}}$ denote to the orthogonal matrices obtained by the strategies (b) and (c), respectively. The entries of these tables suggest to the maximum absolute value of the off-diagonal elements of $\tilde{Q}'\tilde{Q}$ and $\tilde{\tilde{Q}}'\tilde{\tilde{Q}}$. As seen from the entries of both tables, in general the strategy (b) is an adequate method to orthogonalize the original matrix but the strategy (c) is not a desirable technique.

Table 8.2: The Maximum Absolute Value of the Off-Diagonal Elements of $\tilde{Q}'\tilde{Q}$ and $\tilde{\tilde{Q}}'\tilde{\tilde{Q}}$ for Matrix X1.

	$\tilde{Q}'\tilde{Q}$			$\tilde{\tilde{Q}}'\tilde{\tilde{Q}}$		
$n \backslash k$	10	25	50	10	25	50
50	8.435E-17	2.787E-17	3.051E-17	9.970E-01	1.000E+00	1.000E+00
75	6.783E-17	3.787E-17	8.457E-16	9.774E-01	9.936E-01	9.900E-01
100	7.083E-17	4.629E-17	5.986E-17	9.620E-01	9.999E-01	1.000E+00

Table 8.3: The Maximum Absolute Value of the Off-Diagonal Elements of $\tilde{Q}'\tilde{Q}$ and $\tilde{\tilde{Q}}'\tilde{\tilde{Q}}$ for Matrix X2.

		$\tilde{Q}'\tilde{Q}$			$\tilde{\tilde{Q}}'\tilde{\tilde{Q}}$		
c	$n \backslash k$	10	25	50	10	25	50
2	50	4.732E-17	9.064E-17	1.025E-16	4.572E-13	2.137E-13	1.528E-13
	75	5.168E-17	8.536E-17	4.874E-17	7.422E-12	2.620E-13	1.260E-13
	100	6.756E-17	4.419E-17	8.775E-17	1.212E-13	1.006E-13	3.112E-13
10	50	7.452E-17	8.641E-17	8.547E-17	9.984E-01	9.996E-01	9.989E-01
	75	6.665E-17	7.742E-17	9.425E-17	9.963E-01	9.910E-01	9.997E-01
	100	5.253E-17	6.466E-17	8.722E-17	1.000E+00	9.964E-01	9.992E-01

Recalling the model (2), for the initial design matrix X, we have

$$|\operatorname{cov}(\underline{\hat{\beta}})| = \sigma^2 |(X'X)^{-1}| = (70.2)\frac{9000000}{127541} = 4953.7,$$

while, for the orthogonal design Q_h in (6),

$$|\operatorname{cov}(\underline{\hat{\beta}}^*)| = \sigma^2 |(Q_h'Q_h)^{-1}| = (70.2)\frac{36000000}{66739859} = 37.9.$$

These show that the magnitude of the covariance matrix in the improved model (6) is significantly less than the original model (2).

Once again, the simulation experiments corresponding to the design Q_h are generated. Data are given in Table 8.4. Using Table 8.1 and Table 8.4, the θ_I and θ_o are calculated 84.75 and 70.02, respectively. This indicates that the bias of $\underline{\hat{\beta}}^*$ on model (6) is less than $\underline{\hat{\beta}}$ on the model (2).

Table 8.4: Fill Height Runs for the Filling Process of Bottle (Improved Design).

Run	ξ_1	ξ_2	ξ_3	x_1	x_2	x_3	y
1	27	33	218	$\frac{2}{5}$	$\frac{127\sqrt{3851479766}}{22975620}$	$\frac{24259\sqrt{24504383457860}}{42468602180}$	6.63
2	23	25	209	$-\frac{2}{5}$	$-\frac{173\sqrt{3851479766}}{22975620}$	$\frac{6807\sqrt{24504383457860}}{42468602180}$	8.86
3	26	34	204	$\frac{1}{5}$	$\frac{53\sqrt{3851479766}}{9190248}$	$-\frac{2343\sqrt{24504383457860}}{84937204360}$	10.12
4	24	36	196	$-\frac{1}{5}$	$\frac{85\sqrt{3851479766}}{9190248}$	$-\frac{30773\sqrt{24504383457860}}{84937204360}$	1.99
5	27	33	200	$\frac{2}{5}$	$\frac{127\sqrt{3851479766}}{22975620}$	$-\frac{9039\sqrt{24504383457860}}{42468602180}$	2.03
6	28	23	199	$\frac{3}{5}$	$-\frac{43\sqrt{3851479766}}{3829270}$	$-\frac{10347\sqrt{24504383457860}}{42468602180}$	7.81

Generally, to compare the bias of the estimated direction of optimum response based on the models (2) and (6), the results of simulation observations are reported. The performance measures θ_I and θ_o are computed for the designs X and Q_h given in Tables 8.1 and 8.2, respectively and this procedure is replicated 20,000 times. Let $\bar{\theta}_I$ and $\bar{\theta}_o$ to be the average values of θ_I and θ_o. The $\bar{\theta}_I$ and $\bar{\theta}_o$ are calculated 84.94 and 69.75, respectively. It can be demonstrated that the precision of the estimated path of the model (6) is higher than that of the initial model (2).

8-7. Conclusion

The implication of the MGS scheme through Gaussian elimination for the design orthogonalization has been introduced. The proposed method makes the QR decomposition of the original design matrix X in full working precision using real operations. When the floating point entries is desired, for all matrices in all simulation experiments we observed that the maximum absolute value of the

entries of $\tilde{Q}'\tilde{Q}-I$ is less than 8×10^{-d}, where d equals the predefined digits i.e. $Digits := d$. In the sequel, a D-optimal design the corresponding Q_h has been also offered, such that $|(Q_h'Q_h)^{-1}| = \min(|(X'X)^{-1}|)$. The estimated path of SA based on the design Q_h gives a better search direction in the RSM study than the estimated path on each $n\times k$ original design X.

9 Conclusions and Open Problems

CHAPTER OUTLINE

9-1. Conclusions

In this thesis we looked into some issues in quality monitoring and quality improvement. Specifically, as concerns the quality monitoring component we dealt with new control charts for monitoring the mean, the range and the standard deviation of a quality characteristic. The quality improvement issues we looked into were extensions of the strategy of steepest ascent in the area of response surface methodology.

In he first part of the thesis, in particular, it has been demonstrated that the proposed mean control charts, in both phase I and phase II, have three advantages over the classical Shewhart method: the proposed scheme is established using small sample sizes; the in-control *ARL* of the new procedure is very close to the desired *ARL*; and the false alarm probability corresponding to the proposed method equals the intended α. The use of the ANOM and the Bonferroni procedures to monitor historical data in phase I controlling are available in the literature. These methods maintain the overall false alarm probability approximately at a desired level α. The ANOM scheme performs better than the Bonferroni technique in achieving an overall probability of a false signal at the desired α. The use of the proposed strategies is recommended in the case, where the individual occurrence of events G_i and G_i^f is required, while the ANOM strategy is suggested when the overall occurrence of events G_i is considered. The ANOM and the proposed methods are constructed on the basis of the statistic $\bar{X}_{i.} - \bar{\bar{X}}_{..}$, which incorporates more information than $\bar{X}_{i.}$ used by the Shewhart and Bonferroni methods. Moreover, the distribution

function of $\bar{X}_{i.} - \bar{\bar{X}}_{..}$ depends only on the parameter σ, whereas, that of $\bar{X}_{i.}$ depends on both parameters μ and σ.

For an unknown standard deviation parameter, the proposed range and standard deviation control chart has also advantages over the Shewhart and Bonferroni-adjustment methods: the constant values to construct the control chart for the new approach are based on both the sample subgroup size and the sample group size; for some fixed value α, the ARL_1 for the new approach is less than that of the Shewhart and Bonferroni-adjustment scheme; and the new approach is based on a statistic with a variance less than that of the Shewhart and Bonferroni-adjustment strategies. An alternative evaluation of the normal distribution function was proposed and demonstrated to provide a satisfactory approximation to the theoretical values of it. This was applied for calculating the mean range of the normal distribution for various sample sizes n. In addition, the new evaluation for the normal distribution function has greatest absolute error less than 4.02×10^{-14}. The proposed evaluation is based on the values $-\infty < z < +\infty$ while the most of existent evaluations are based on subset of the values $0 < z < +\infty$ or $-9 < z < 9$.

The later part of the thesis refers to problems in response surface methodology contexts, where the initial estimate of the optimum operating conditions for the system often is far from the actual optimum. In such case, the method of steepest ascent is used in order to move sequentially along the steepest ascent path i.e. in the direction of the maximum increase in the response. Some extensions of the method of steepest ascent have been introduced. The extensions included the computation of the path of steepest

ascent, the path under some restriction, the confidence region for the direction of the steepest ascent, and the confidence cone about the estimated path. To lead the experimenter toward the general vicinity of the optimum response, an improvement of the path has been proposed. The improvement is based on the augmentation of the existing data strategy. Applying this method, the orthogonality of experimental matrix X is increased, such that the discrepancy between $|(XX)^{-1}|$ and $\min|(XX)^{-1}|$ is made equal or nearly equal to zero. Furthermore, it has been demonstrated that the augmenting data and the modified Gram-Schmidt strategies appear to lead to a quite satisfactory decrease in both the bias and the covariance matrix of the estimated path of steepest ascent.

9-2. Open Problems

Further research is needed to overcome the problem that the proposed control charts are only based on the information about the process contained in the past sample runs and they take no account of any information obtained by subsequent observations. This implies that the new control charts are of relatively less practical value in phase II monitoring charts. Indeed, the proposed charts are not very sensitive in monitoring the small shifts of the order of about 1.5σ or less. The classical Shewhart charts have also the same problems as mentioned for the proposed charts. Hence, in phase II, appropriate alternative charts to the proposed and the Shewhart ones may be constructed if one wishes to capture small shifts. The usual alternative methods to the Shewhart charts are the cumulative sum (called cusum) control chart, and the exponentially weighted moving average (called EWMA) control chart. It should

be noted, however, that the cusum and EWMA control charts are not as effective as the Shewhart chart or the proposed new charts in detecting large shifts. It would therefore be interesting to develop a new procedure that will overcome this shortcoming of the proposed charts.

A problem related to the construction of the R control charts is the choice of approximation for evaluating the mean range of the normal density function (NDF). Although, the proposed approximations have some advantages over the existing methods, they are not exact. More accurate approximations of the mean range of the NDF would therefore be worth developing.

In order to improve the performance of production process, in the second part of the thesis, we have introduced an extension to the steepest ascent methodology, which is available in the literature of the statistical quality control. As concerns the methods of data augmentation that have been suggested for improving the estimated path of steepest ascent with a path based on a first-order model, it is worth noting that often, after one or two applications of the steepest ascent method, first-order effects will no longer dominate and the first-order approximation will be inadequate. At this stage, a further application of the steepest ascent method will be of no help and, hence, more sophisticated second-order methods should be sought. Hence, an interesting problem in this area would be how to exploit the proposed methods of data augmentation for a second-order model to improve the estimation of its parameters.

Appendix Table I: Factors for Constructing Variables Control Charts (Montgomery (2001))

Observations in Sample, n	Chart for Averages: Factors for Control Limits			Chart for Standard Deviations: Factors for Center Line		Factors for Control Limits				Factors for Center Line		
	A	A_2	A_3	c_4	$1/c_4$	B_3	B_4	B_5	B_6	d_2	$1/d_2$	d_3
2	2.1213	1.8806	2.6586	0.7979	1.2533	0.0000	3.2664	0.0000	2.6063	1.128	0.8865	0.853
3	1.7321	1.0231	1.9545	0.8862	1.1284	0.0000	2.5684	0.0000	2.2761	1.693	0.5907	0.888
4	1.5000	0.7285	1.6281	0.9213	1.0854	0.0000	2.2662	0.0000	2.0879	2.059	0.4857	0.880
5	1.3416	0.5768	1.4273	0.9400	1.0638	0.0000	2.0889	0.0000	1.9635	2.326	0.4299	0.864
6	1.2247	0.4833	1.2872	0.9515	1.0510	0.0300	1.9700	0.0286	1.8744	2.534	0.3946	0.848
7	1.1339	0.4193	1.1819	0.9594	1.0423	0.1180	1.8820	0.1133	1.8055	2.704	0.3698	0.833
8	1.0607	0.3726	1.0991	0.9650	1.0363	0.1847	1.8153	0.1783	1.7517	2.847	0.3512	0.820
9	1.0000	0.3367	1.0317	0.9693	1.0317	0.2390	1.7610	0.2317	1.7069	2.970	0.3367	0.808
10	0.9487	0.3082	0.9753	0.9727	1.0281	0.2843	1.7157	0.2765	1.6689	3.078	0.3249	0.797
11	0.9045	0.2851	0.9273	0.9754	1.0252	0.3220	1.6780	0.3141	1.6367	3.173	0.3152	0.787
12	0.8660	0.2658	0.8859	0.9776	1.0229	0.3541	1.6459	0.3462	1.6090	3.258	0.3069	0.778
13	0.8321	0.2494	0.8496	0.9794	1.0210	0.3815	1.6185	0.3736	1.5852	3.336	0.2998	0.770
14	0.8018	0.2353	0.8173	0.9810	1.0194	0.4067	1.5933	0.3990	1.5630	3.407	0.2935	0.763
15	0.7746	0.2231	0.7886	0.9823	1.0180	0.4279	1.5721	0.4204	1.5442	3.472	0.2880	0.756
16	0.7500	0.2123	0.7626	0.9835	1.0168	0.4482	1.5518	0.4408	1.5262	3.532	0.2831	0.750
17	0.7276	0.2028	0.7391	0.9845	1.0157	0.4656	1.5344	0.4583	1.5107	3.588	0.2787	0.744
18	0.7071	0.1943	0.7176	0.9854	1.0148	0.4817	1.5183	0.4746	1.4962	3.640	0.2747	0.739
19	0.6882	0.1866	0.6979	0.9862	1.0140	0.4964	1.5036	0.4895	1.4829	3.689	0.2711	0.734
20	0.6708	0.1796	0.6797	0.9869	1.0133	0.5096	1.4904	0.5029	1.4709	3.735	0.2677	0.729
21	0.6547	0.1733	0.6629	0.9876	1.0126	0.5231	1.4769	0.5166	1.4586	3.778	0.2647	0.724
22	0.6396	0.1675	0.6472	0.9882	1.0119	0.5350	1.4650	0.5287	1.4477	3.819	0.2618	0.720
23	0.6255	0.1621	0.6327	0.9887	1.0114	0.5451	1.4549	0.5390	1.4384	3.858	0.2592	0.716
24	0.6124	0.1572	0.6191	0.9892	1.0109	0.5555	1.4445	0.5495	1.4289	3.895	0.2567	0.712
25	0.6000	0.1526	0.6063	0.9896	1.0105	0.5639	1.4361	0.5581	1.4211	3.931	0.2544	0.708

For $n > 25$: $A = \dfrac{3}{\sqrt{n}}$; $A_3 = \dfrac{3}{c_4\sqrt{n}}$; $c_4 \cong \dfrac{4(n-1)}{4n-3}$; $B_3 = 1 - \dfrac{3}{c_4\sqrt{2(n-1)}}$;

$B_4 = 1 + \dfrac{3}{c_4\sqrt{2(n-1)}}$; $B_5 = c_4 - \dfrac{3}{\sqrt{2(n-1)}}$; $B_5 = c_4 + \dfrac{3}{\sqrt{2(n-1)}}$.

Appendix Table II: ANOM Critical Values: Level of Significance 0.05 ($\alpha = 0.05$).

	Number of Means *(k)*								
df (v)	2	3	4	5	6	7	8	9	10
1	12.7	19.1	22.0	24.1	25.7	26.9	28.0	28.9	29.7
2	4.30	5.89	6.60	7.10	7.49	7.81	8.07	8.30	8.50
3	3.18	4.18	4.60	4.91	5.15	5.34	5.50	5.65	5.77
4	2.78	3.56	3.89	4.12	4.30	4.45	4.58	4.69	4.79
5	2.57	3.25	3.53	3.72	3.88	4.00	4.11	4.21	4.29
6	2.45	3.07	3.31	3.49	3.62	3.73	3.83	3.91	3.99
7	2.36	2.95	3.17	3.33	3.45	3.56	3.65	3.72	3.79
8	2.31	2.86	3.07	3.21	3.33	3.43	3.51	3.58	3.64
9	2.26	2.79	2.99	3.13	3.24	3.33	3.41	3.48	3.54
10	2.23	2.74	2.93	3.07	3.17	3.26	3.33	3.40	3.45
11	2.20	2.70	2.88	3.01	3.12	3.20	3.27	3.33	3.39
12	2.18	2.67	2.85	2.97	3.07	3.15	3.22	3.28	3.33
13	2.16	2.64	2.81	2.94	3.03	3.11	3.18	3.24	3.29
14	2.14	2.62	2.79	2.91	3.00	3.08	3.14	3.20	3.25
15	2.13	2.60	2.76	2.88	2.97	3.05	3.11	3.17	3.22
16	2.12	2.58	2.74	2.86	2.95	3.02	3.09	3.14	3.19
17	2.11	2.57	2.73	2.84	2.93	3.00	3.06	3.12	3.16
18	2.10	2.55	2.71	2.82	2.91	2.98	3.04	3.10	3.14
19	2.09	2.54	2.70	2.81	2.89	2.96	3.02	3.08	3.12
20	2.09	2.53	2.68	2.79	2.88	2.95	3.01	3.06	3.11
24	2.06	2.50	2.65	2.75	2.83	2.90	2.96	3.01	3.05
30	2.04	2.47	2.61	2.71	2.79	2.85	2.91	2.96	3.00
40	2.02	2.43	2.57	2.67	2.75	2.81	2.86	2.91	2.95
60	2.00	2.40	2.54	2.63	2.70	2.76	2.81	2.86	2.90
120	1.98	2.37	2.50	2.59	2.66	2.72	2.77	2.81	2.85
INF	1.96	2.34	2.47	2.56	2.62	2.68	2.72	2.76	2.80

Nelson, L. S. (1993).

Appendix Table II: ANOM Critical Values: Level of Significance 0.05 ($\alpha = 0.05$).

	Number of Means *(k)*								
df (v)	11	12	13	14	15	16	17	18	19
1	30.4	31.0	31.6	32.1	32.5	33.0	33.4	33.8	34.1
2	8.68	8.84	8.99	9.12	9.24	9.36	9.46	9.56	9.65
3	5.88	5.98	6.08	6.16	6.24	6.31	6.38	6.44	6.50
4	4.87	4.95	5.02	5.09	5.15	5.21	5.26	5.31	5.35
5	4.36	4.43	4.49	4.55	4.60	4.65	4.69	4.73	4.77
6	4.05	4.11	4.17	4.22	4.26	4.31	4.35	4.39	4.42
7	3.85	3.90	3.95	4.00	4.04	4.08	4.12	4.15	4.19
8	3.70	3.75	3.80	3.84	3.88	3.92	3.95	3.99	4.02
9	3.59	3.64	3.68	3.72	3.76	3.80	3.83	3.86	3.89
10	3.51	3.55	3.59	3.63	3.67	3.70	3.73	3.76	3.79
11	3.44	3.48	3.52	3.56	3.60	3.63	3.66	3.69	3.71
12	3.38	3.42	3.46	3.50	3.53	3.57	3.59	3.62	3.65
13	3.33	3.38	3.42	3.45	3.48	3.51	3.54	3.57	3.59
14	3.30	3.34	3.37	3.41	3.44	3.47	3.50	3.52	3.55
15	3.26	3.30	3.34	3.37	3.40	3.43	3.46	3.49	3.51
16	3.23	3.27	3.31	3.34	3.37	3.40	3.43	3.45	3.48
17	3.21	3.25	3.28	3.31	3.34	3.37	3.40	3.42	3.45
18	3.18	3.22	3.26	3.29	3.32	3.35	3.37	3.40	3.42
19	3.16	3.20	3.24	3.27	3.30	3.32	3.35	3.37	3.40
20	3.15	3.18	3.22	3.25	3.28	3.30	3.33	3.35	3.38
24	3.09	3.13	3.16	3.19	3.22	3.24	3.27	3.29	3.31
30	3.04	3.07	3.10	3.13	3.16	3.18	3.20	3.23	3.25
40	2.98	3.01	3.04	3.07	3.10	3.12	3.14	3.16	3.18
60	2.93	2.96	2.99	3.01	3.04	3.06	3.08	3.10	3.12
120	2.88	2.91	2.93	2.96	2.98	3.00	3.02	3.04	3.06
INF	2.83	2.86	2.88	2.90	2.93	2.95	2.97	2.98	3.00

Nelson, L. S. (1993).

Appendix Table III: ANOM Critical Values: Level of Significance 0.01 ($\alpha = 0.01$).

	Number of Means *(k)*								
df (v)	2	3	4	5	6	7	8	9	10
1	63.70	95.70	110	121	129	135	140	145	149
2	9.92	13.40	15.00	16.10	17.00	17.70	18.30	18.80	19.30
3	5.84	7.51	8.22	8.73	9.13	9.64	9.74	9.98	10.20
4	4.60	5.47	6.20	6.54	6.81	7.03	7.22	7.38	7.52
5	4.03	4.93	5.29	5.55	5.75	5.92	6.07	6.19	6.30
6	3.71	4.48	4.77	4.98	5.16	5.30	5.42	5.52	5.62
7	3.50	4.19	4.44	4.63	4.78	4.90	5.01	5.10	5.18
8	3.36	3.98	4.21	4.38	4.52	4.63	4.72	4.80	4.88
9	3.25	3.84	4.05	4.20	4.33	4.43	4.51	4.59	4.66
10	3.17	3.73	3.92	4.07	4.18	4.28	4.36	4.43	4.49
11	3.11	3.64	3.83	3.96	4.07	4.16	4.23	4.30	4.36
12	3.05	3.57	3.75	3.87	3.98	4.06	4.13	4.20	4.25
13	3.01	3.51	3.68	3.80	3.90	3.98	4.05	4.11	4.17
14	2.98	3.46	3.63	3.74	3.84	3.92	3.98	4.04	4.09
15	2.95	3.42	3.58	3.69	3.78	3.86	3.93	3.98	4.03
16	2.92	3.38	3.54	3.65	3.74	3.81	3.88	3.93	3.98
17	2.90	3.35	3.50	3.61	3.70	3.77	3.83	3.89	3.93
18	2.88	3.33	3.47	3.58	3.66	3.73	3.79	3.85	3.89
19	2.86	3.30	3.45	3.55	3.63	3.70	3.76	3.81	3.86
20	2.85	3.28	3.42	3.53	3.61	3.67	3.73	3.78	3.83
24	2.80	3.21	3.35	3.44	3.52	3.58	3.64	3.69	3.73
30	2.75	3.15	3.28	3.37	3.44	3.50	3.55	3.59	3.63
40	2.70	3.09	3.21	3.29	3.36	3.42	3.46	3.51	3.54
60	2.66	3.03	3.14	3.22	3.28	3.34	3.38	3.42	3.45
120	2.62	2.97	3.08	3.15	3.21	3.26	3.30	3.34	3.37
INF	2.58	2.91	3.01	3.08	3.14	3.19	3.22	3.26	3.29

Nelson, L. S. (1993).

Appendix Table III: ANOM Critical Values: Level of Significance 0.01 ($\alpha = 0.01$).

	Number of Means *(k)*								
df (v)	11	12	13	14	15	16	17	18	19
1	152	155	158	160	163	165	167	169	171
2	19.70	20.00	20.40	20.70	20.90	21.20	21.40	21.60	21.90
3	10.40	10.60	10.70	10.90	11.00	11.10	11.20	11.30	11.40
4	7.65	7.77	7.88	7.97	8.07	8.15	8.23	8.30	8.37
5	6.40	6.50	6.58	6.66	6.73	6.79	6.86	6.91	6.97
6	5.70	5.78	5.85	5.91	5.97	6.03	6.08	6.13	6.18
7	5.25	5.32	5.38	5.44	5.49	5.54	5.59	5.63	5.67
8	4.94	5.01	5.06	5.11	5.16	5.20	5.24	5.28	5.32
9	4.72	4.78	4.83	4.87	4.92	4.96	5.00	5.03	5.07
10	4.55	4.60	4.65	4.69	4.73	4.77	4.81	4.84	4.87
11	4.41	4.46	4.51	4.55	4.59	4.62	4.66	4.69	4.72
12	4.31	4.35	4.39	4.43	4.47	4.50	4.54	4.57	4.59
13	4.22	4.26	4.30	4.34	4.37	4.41	4.44	4.47	4.49
14	4.14	4.18	4.22	4.26	4.29	4.32	4.35	4.38	4.41
15	4.08	4.12	4.16	4.19	4.22	4.26	4.28	4.31	4.34
16	4.02	4.06	4.10	4.13	4.17	4.20	4.22	4.25	4.27
17	3.98	4.02	4.05	4.08	4.12	4.14	4.17	4.20	4.22
18	3.94	3.97	4.01	4.04	4.07	4.10	4.12	4.15	4.17
19	3.90	3.94	3.97	4.00	4.03	4.06	4.08	4.11	4.13
20	3.87	3.90	3.94	3.97	4.00	4.02	4.05	4.07	4.09
24	3.77	3.80	3.83	3.86	3.89	3.91	3.94	3.96	3.98
30	3.67	3.70	3.73	3.76	3.78	3.81	3.83	3.85	3.87
40	3.58	3.61	3.63	3.66	3.68	3.70	3.72	3.74	3.76
60	3.49	3.51	3.54	3.56	3.58	3.61	3.62	3.64	3.66
120	3.40	3.42	3.45	3.47	3.49	3.51	3.53	3.54	3.56
INF	3.32	3.34	3.36	3.38	3.40	3.42	3.44	3.45	3.47

Nelson, L. S. (1993).

References

Abramowitz, M., and Stegun, I. A. (1964). *Handbook of Mathematical Functions*, New York: Dover Publications, Inc.

Adams, A. G. (1969). Remark on Algorithm 304: Normal Curve Integral. *Communications of the Association for Computing Machinary*, Vol. 12, pp. 565-566.

Albers, W., and Kallenberg, W. C. M. (2004). Estimation in Shewhart Control Charts: Effects and Corrections. *Metrika,* Vol. 59, pp. 207-234.

Arnholt, A. T., and Hebert, J. L. (1995). Estimating the Mean With Known Coefficient of Variation. *The American Statistician*, Vol. 49, No. 4, pp. 367-369.

Atwood, C. L. (1969). Optimal and Efficient Design of Experiments. *The Annals of Mathematical Statistics*, Vol. 40, pp. 1570-1602.

Badhe, S. K. (1976). New Approximation of the Normal Distribution Function, *Communications in Statistics-Simulation and Computation,* Vol. 5, pp. 173-176.

Bagby, R. J. (1995). Calculating Normal Probabilities. *Amer. Math. Monthly.* Vol. 102, pp. 46-49.

Barnard, J. (1978). Algorithm AS 126: Probability Integral of the Normal Range. *Applied Statistics*, Vol. 27, pp. 197-198.

Björck, Å. (1978). Comment on the Iterative Refinement of Least-Squares Solutions. *Journal of American Statistical Association*, Vol. 73, No. 361, pp. 161-166.

Box, G. E. P. (1952). Multifactor Designs of First Order. *Biometrika*, Vol. 39, No. 1/2, pp. 49-57.

Box, G. E. P. and Draper, N. R. (2007). *Response Surfaces, Mixtures, Ridge Analysis*. 2nd edition. John Wiley, New York.

Box, G. E. P. and Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society,* Series B Vol. 13, No. 1, pp. 1–38.

Box, M. J. (1971). An Experimental Design Criterion for Precise Estimation of a Subset of the Parameters in a Nonlinear Model. *Biometrika*, Vol. 58, No. 1, pp. 149-153.

Box, M. J. and Draper, N. R. (1971). Factorial Designs, the $|X'X|$ Criterion, and Some Related Matters. *Technometrics*, Vol. 13, No. 4, pp. 731-742.

Brooks, S. H., and Mickey, M. R. (1961). Optimum Estimation of Gradient Direction in Steepest Ascent Experiments. *Biometrics*, Vol. 17, No. 1, pp. 48-56.

Bryc, W. (2002). A Uniform Approximation to the Right Normal Tail Integral. *Applied Mathematics and Computation,* Vol. 127, pp. 365–374.

Budsaba, K., Smith, C. E., and Riviere, J. E. (2000). Compass Plots: A Combination of Star Plot and Analysis of Means to Visualize Significant Interactions In Complex Toxicology Studies. *Toxicology Mechanisms and Methods*, Vol. 10, No. 4, pp. 313-332.

Burroughs, T. E., Rigdon, T. E., and Champ C. W. (1995). Analysis of the Shewhart S-Chart with Runs Rules When No Standards Are Given. *Proceedings of the Twenty-Sixth Annual Meeting of the Midwest Decision Seiences Istitute*, May 4-6, St. Louis, Missouri, 268-270.

Chakraborti, S. (2000). Run Length, Average Run Length and False Alarm Rate of Shewhart X-bar Chart: Exact Derivations by Conditioning. *Communications in Statistics: Simulation and Computation,* Vol. 29, pp. 62-81.

Chaloner, K., and Larntz, K. (1989). Optimal Bayesian Design Applied to Logistic Regression Experiments. *Journal of Statistical Planning and Inference*, Vol. 21, pp. 191-208.

Chaloner, K., and Verdinelli, I. (1995). Bayesian Experimental Design: A Review. *Statistical Science*, Vol. 10, No. 3, pp. 273-304.

Champ, C. W. and Jones, L. A. (2004). Designing Phase I Charts with Small Sample Sizes. *Quality and reliability Engineering International*, Vol. 20, pp. 497-510.

Chang, M. Y., Tsai, G. J., Houng, J. Y. (2006). Optimization of the Medium Composition for the Submerged Culture of Ganoderma Lucidum by Taguchi Array Design and Steepest Ascent Method. *Enzyme and microbial Technology,* Vol. 38, pp. 407-414.

Chaudhuri, P., and Mykland, P. A. (1993). Nonlinear Experiments: Optimal Design and Inference Based on Likelihood. *Journal of the American Statistical Association*, Vol. 88, No. 422, pp. 538-546.

Chen, G. (1997). The Mean and Standard Deviation of the Run Length Distribution of X-bar Charts with Control Limits Are Estimated. *Statistica Sinica,* Vol. 7, pp. 789-798.

Chernoff, H. (1953). Locally Optimum Designs for Estimating Parameters. *The Annals of Mathematical Statistics*, Vol. 24, pp. 586-602.

Daniel, J. W., Gragg, W. B., Kaufman, L., and Stewart, G. W. (1976). Reorthogonalization and Stable Algorithms for Updating the Gram-Schmidt QR factorization. *Math. Comp.* Vol. 30, pp. 772-795.

De la Garza, A. (1954). Spacing of Information in Polynomial Regression. *The Annals of Mathematical Statistics*, Vol. 25, pp. 123-130.

DeGroot, M. H., and Goel, P. K. (1979). Bayesian Estimation and Optimal Design in Partially Accelerated Life Testing. *Naval Res. Logist. Quart.,* Vol. 26, pp. 223-235.

Derenzo, S. E. (1977). Approximations for Hand Calculators Using Small Integral Coefficients, *Mathematics of Computation,* Vol. 31, pp. 214-225.

Dey, A., and Mukerjee, R. (2006). D-optimal Designs for Covariate Models. *Statistics*, Vol. 40, No. 4, pp. 297-305.

Divgi, D. R. (1979). Calculation of Univariate and Bivariate Normal Probability Function. *The Annals of Statistics*, Vol. 7, No. 4, pp. 903-910.

Donatos, G. S. (1989). A Monte Carlo Study of k-Class Estimators for Small Samples with Normal and Non-Normal Disturbances. *The Statistician*, Vol. 38, No. 1, pp. 11-20.

Dror, H. A., and Steinberg, D. M. (2006). Robust Experimental Designs for Multivariate Generalized Linear Models. *Technometrics*, Vol. 48, No. 4, pp. 520-529.

Dror, H. A., and Steinberg, D. M. (2008). Sequential Experimental Designs for Generalized Linear Models. *Journal of the American Statistical Association*, Vol. 21, No. 3, pp. 376-399.

Dykstra, O. J. (1966). The Orthogonalization of Undesigned Experiments. *Technometrics*, Vol. 8, No. 2, pp. 279-290.

Dykstra, O. Jr. (1971). The Augmentation of Experimental Data to Maximize $|X'X|$. *Technometrics*, Vol. 13, No. 3, pp. 682-688.

Edgeman, R. L. (1988). Normal Distribution Probabilities and Quantiles without Tables. *Mathematics of Computer Education*, Vol. 22, pp. 95-99.

Ehrenfeld, S. (1955). On the Efficiencies of Experimental Designs. *The Annals of Mathematical Statistics*, Vol. 26, pp. 247-255.

Elfving, G. (1952). Optimum Allocation in Linear Regression Theory. *The Annals of Mathematical Statistics*, Vol. 23, pp. 255-262.

Evans, J. W. (1979). Computer Augmentation of Experimental Designs to Maximize $|X'X|$. *Technometrics*, Vol. 21, No. 3, pp. 321-330.

Farebrother, R. W. (1974). Algorithm AS 76: Gram-Schmidt Regression. *Applied Statistics*, Vol. 23, No. 3, pp. 470-476.

Fedorov, V. V. (1972). *Theory of Optimal Experiments*, Trans. and ed. by W. J. Studden and E. M. Klimko. New York: Academic Press.

Fletcher, R. H. (1975). On the Iterative Refinement of Least Squares Solutions. *Journal of American Statistical Association*, Vol. 70, No. 349, pp. 109-112.

Gaylor, D. W., and Merrill, J. A. (1968). Augmenting Existing Data in Multiple Regression. *Technometrics*, Vol. 10, No. 1, pp. 73-81.

Ghosh, B. K., Reynolds, M. R., and Hui, Y. V. (1981). Shewhart X-bar Chart with Estimated Process Variance. *Communications in Statistics: Theory and Methods,* Vol. 10, pp. 1797-1822.

Giraud, L., Langou, J., Rozložník, M. and Eshof, J. V. D. (2005). Rounding Error Analysis of the Classical Gram-Schmidt Orthogonalization Process. *Numerische Mathematik*, Vol. 101, pp. 87-100.

Glasser, G. J. (1962). On Estimators for Variances and Covariances. *Biometrika*, Vol. 49, No. ½, pp. 259-262.

Golub, G. (1965). Numerical Methods for Solving Linear Least Squares Problems. *Numerische Mathematik*, Vol. 7, pp. 206-216.

Golub, G. H. and Wilkinson, J. H. (1966). Note on the Iterative Refinement of Least Squares Solution. *Numerische Mathematik*, Vol. 9, pp. 139-148.

Guest, P. G. (1958). The Spacing of Observations in Polynomial Regression. *The Annals of Mathematical Statistics*, Vol. 29, pp. 294-299.

Gurland, J., and Tripathi, R. C. (1971). A Simple Approximation for Unbiased of the Standard Deviation. *The American Statistician*, Vol. 25, No. 4, pp. 30-32.

Haines, L. M., Perevozskaya, I. and Rosenberger, W. F. (2003). Bayesian Optimal Designs for Phase I Clinical Trials. *Biometrics* Vol. 59, PP. 591–600.

Halperin, M., Greenhouse, S. W., Cornfield, J., and Zalokar, J. (1955). Tables of Percentage Points for the Studentized Maximum Absolute Deviation in Normal Samples. *Journal of American Statistical Association,* Vol. 50, pp. 185-195.

Hamaker, H. (1978). Approximating the Cumulative Normal Distribution and its Inverse. *Appl. Statist.* , Vol. 27, pp. 76-77.

Hart, R. G. (1966). A Close Approximation Related to the Error Function, *Mathematics of Computation*, Vol. 20, and pp 600-602.

Hartley, H. O. (1942). The Range in Random Samples. *Biometrika*, Vol. 32, pp. 334-348.

Hawkins, D. M., Qiu, P., and Kang, C. W. (2003). The Changepoint Model for Statistical Process Control. *Journal of Quality Technology*, Vol. 35, No. 4, pp. 355-366.

Healy, M. J. R. (1978). A Mean Difference Estimator of Standard Deviation in Symmetrically Censored Normal Samples. *Biometrika*, Vol. 65, No. 3, pp. 643-646.

Heard, T. J. (1979). Approximation to the Normal Distribution Function. *Mathematical Gazette*, Vol. 63, pp. 39-40.

Heiberger, R. M., Bhaumik, D. K., and Holland, B. (1993). Optimal Data Augmentation Strategies for Additive Models. *Journal of American Statistical Association*, Vol. 88, No. 423. pp. 926-938.

Hendrix, C. D. (1971). Empirical Optimization in Research and Development. *Proceedings of the Second Gordon Research Conference on Operation Research and Management Science*.

Heredia-Laugner, A., Carlyle, W. M., Montgomery, D. C., Borror, C. M., and Runger, G. C. (2003). Genetic Algorithms for the Construction of D-Optimal Designs. *Journal of Quality*

Technology, Vol. 35, No. 1, pp. 28-46.

Hill, W. J. and Hunter, W. G. (1974). Design of Experiments for Subsets of Parameters. *Technometrics*, Vol. 16, No. 3, pp. 425-434.

Jensen, W. A., Jones-Farmer, L. A., Champ, C. W., and Woodall, W. H. (2006). Effects of Parameter Estimation on Control Chart Properties: A Literature Review. *Journal of Quality Technology*, Vol. 38, No. 4, pp. 349-364.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions.* 2nd edition. Wiley, New York.

Jones, L. A., Champ, C. W., and Rigdon, S. E. (2001). Performance Analysis of Exponentially Weighted Moving Average Charts When Parameters Are Estimated. *Technometrics*, Vol. 43, pp. 156-167.

Jones, L. A., Champ, C. W., and Rigdon, S. E. (2004). The Run Length Distribution of the CUSUM Chart with Estimated Parameters. *Journal of Quality Technology*, Vol. 36, pp. 95-108.

Kang, L., and Albin, S. L. (2000). On-Line Monitoring When the Process Yields a Linear Profile. *Journal of Quality Technology*, Vol. 32, No. 4, pp. 418-426.

Karlin, S., and Studden, W. J. (1966), "Optimal Experimental Designs," *The Annals of Mathematical Statistics*, Vol. 37, pp. 783-815.

Kennedy, W. J., and Gentle, J. E. (1980). *Statistical Computing*. Marcel-Dekker: New York.

Kerridge, D. F., and Cook, G. W. (1976). Yet Another Series for the Normal Integral, *Biometrica*, Vol. 63, pp. 401-403.

Khan, R. A. (1968). A Note on Estimating the Mean of a Normal Distribution with Known Coefficient of Variation. *Journal of the American Statistical Association*, Vol. 63, No. 323, pp. 1039-1041.

Khuri, A. I. and Cornell, J. A. (1996). *Response Surfaces: Designs and Analysis*. 2nd edition. Dekker, New York.

Khuri, A. I., and Cornell, J. A. (1996). *Response Surfaces: Designs and Analysis*. 2nd edition. Dekker, New York.

Khuri, A. I., Mukherjee, B., Sinha, B. K., and Ghosh, M. (2006). Design Issues for Generalized Linear Models: A Review. *Statistical Science*, Vol. 21, No. 3, pp. 376-399.

Kiani, M. (2008). Extensions to Modified Gram-Schmidt Strategy and Its Application in Steepest Ascent Method. To appear in *Journal of Statistical Computation and Simulation.*

Kiani, M., Panaretos, J. and Psarakis, S. (2008). A New Procedure to Monitor the Rang and Standard Deviation of a Quality Characteristic. To appear in the *Quality and Quantity.*

Kiani, M., Panaretos, J. and Psarakis, S. (2008). A New Procedure to Monitor the Mean of a Quality Characteristic. To appear in the *Communications in Statistics* B, (*Simulation and Computation*).

Kiefer, J. (1958). On the Nonrandomized Optimality and Randomized Nonoptimality of Symmetrical Designs. *The Annals of Mathematical Statistics*, Vol. 29, pp. 675-699.

Kiefer, J. (1959). Optimum Experimental Designs. *Journal of the Royal Statistical Society,* Series B Vol. 21, No. 2, pp. 273-319.

Kiefer, J. (1961). Optimum Designs in Regression Problems, II. *The Annals of Mathematical Statistics*, Vol. 32, No. 1, pp. 298-325.

Kiefer, J. and Wolfowitz, J. (1959). Optimum Designs in Regression Problems. *The Annals of Mathematical Statistics*, Vol. 30, pp. 271-294.

Lew, R. A. (1981). An Approximation to the Cumulative Normal Distribution with Simple Coefficients. *Applied Statistics*, Vol. 30, No. 3, pp. 299-301.

Li, K., Lau, T., and Zhang, C. (2005). A Note on D-optimal designs for models with and without an intercept. *Statistical Papers*, Vol. 46, pp. 451-458.

Liu, C. H., Wu, M. L., Hwang, S. M. (2007). Optimization of Serum Free Medium for Cord Blood Mesenchymal Stem Cells. *Biochemical Engineering Journal*, Vol. 33, pp. 1-9.

Maravelakis, P. E. (2003). An Investigation of Some Characteristics of Univariat and Multivariat Control Charts. Ph.D. dissertation, Department of Statistics, Athens University of Economics and Business.

Maravelakis, P. E., Panaretos, J., and Psarakis, S. (2002). Effect of Estimation of the Process Parameter on the Control Limits of the Univariate Control Charts for Process Dispersion. *Communications in Statistics: Simulation and Computation.* Vol. 31, No. 3, pp. 443-461.

Markowitz, E. (1968). Minimum Mean-Square-Error Estimation of the Standard Deviation of the Normal Distribution. *The American Statistician*, Vol. 22, No. 3, pp. 26-26.

Marsaglia, G. (2004). Evaluating the Normal Distribution. *Journal of Statistical Software*, Vol. 11, No. 4, pp. 1-11.

Martynov, G. V. (1981). Evaluation of the Normal Distribution Function. *Journal of Soviet Mathematics,* Vol. 17, pp. 1857-1876.

Mitchell, T. J. (2000). An Algorithm for the Construction of "*D*-optimal" Experimental Designs. *Technometrics*, Vol. 42, No. 1, pp. 48-54.

Moerbeek, M. (2005). Robustness properties of *A*-, *D*-, and *E*-optimal designs for polynomial growth models with autocorrelated errors. *Computational Statistics & Data Analysis*, Vol. 48, No. 4, pp. 765-778.

Monahan, J. F. (1981). Approximation the Log of the Normal Cumulative. *Computer Science and Statistic: Proceeding of the Thirteenth Symposium on the Interface* W. F. Eddy (editor), 304-307, New York: Springer-Verlag.

Montgomery, D. C. (2001). *Design and Analysis of Experiments*. 5th edition. Wiley, New York.

Montgomery, D. C. (2005). *Introduction to Statistical Quality Control*, 5th edition. John Wiley, New York.

Mood, A. M. (1946). On Hotelling's Weighing Problem. *The Annals of Mathematical Statistics*, Vol. 17, pp. 432-446.

Moran, P. A. P. (1980). Calculation of the Normal Distribution Function, *Biometrika,* Vol. 67, pp. 675-676.

Myers, R. H. and Montgomery, D. C. (2002). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. John Wiley, New York.

Myers, R. H., Montgomery, D. C. and Vining, G. G. (2002). *Generalized Linear Models with Applications in Engineering and the Sciences*. John Wiley, New York.

Nedumaran, G., and Pignatiello, J. J. (2005). On Constructing Retrospective X-bar Control Chart Limits. *Quality and Reliability Engineering International*, Vol. 21, pp. 81-89.

Nelder, J. A. and Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, Vol. 7, pp. 308-313.

Nelson, L. S. (1973). A Sequential Simplex Procedure for Nonlinear Least Squares Estimation and Other Function Minimization Problems. 27^{th} *Annual Technical Conference Transactions, American Society for Quality Control.*

Nelson, P. R. (1982). Exact Critical Points for the Analysis of Means. *Communications in Statistics: Theory and Methods*, Vol. 11, pp. 699-709.

Nelson, P. R. (1988). Testing for Interactions Using the Analysis of Means. *Technometrics*, Vol. 30, pp. 53-61.

Nelson, P. R. (1993). Additional Using for the Analysis of Means and Extended Tables of Critical Values. *Technometrics,* Vol. 35, No. 1. pp. 61-71.

Ng, C. H., and Case, K. E. (1992). Control Limits and the ARL: Some Surprises. *Proceeding of the First Intustrial Engineering Research Conference Proceedings*, 21-29.

Ott, E. R. (1967). Analysis of Means-A Graphical Procedure. *Industrial Quality Control,* Vol. 24, pp. 101-109.

Ott, E. R. (1975). *Process Quality Control. McGraw-Hill Book Company*, New York.

Parsonson, S. L. (1978). An Approximation to the Normal Distribution Function. *Mathematical Gazette*, Vol. 62, pp. 118-121.

Pearson, E. S., and Hartley, H. O. (1970). *Biometrika Tables for Statisticians*, Vol. 1, Cambridge University press: London.

Plackett, R. L., and Burman, J. P. (1946). The Design of Optimum Multifactorial Experiments. *Biometrika*, Vol. 33, No. 4, pp. 305-325.

Prescott, P. (1971a). Use of a Simple Range-Type Estimator of Standard Deviation in Test of Hypotheses. *Biometrika*, Vol. 58, No. 2, pp. 333-340.

Prescott, P. (1971b). Distribution of the Normal Scores Estimator of the Standard Deviation of a Normal Population. *Biometrika*, Vol. 58, No. 3, pp. 631-636.

Pugh, G. A. (1989). Computing with Gaussian Distribution: A Survey of Algorithms, *ASQC Technical Conference Transactions,* Vol. 46, pp. 496-501.

Quesenberry, C. P. (1993). The Effect of Sample Size on Estimated Limits for X-bar and X Control Charts. *Journal of Quality Technology*, Vol. 25, pp. 237-247.

Quesenberry, C. P. (1997). *SPC Methods for Quality Improvement*. John Wiley & Sons.

Ramig, P. F. (1983). Applications of the Analysis of Means. *Journal of Quality Technology*, Vol. 15, pp. 19-25.

Rao, C. V. (2005). Analysis of Means-A Review. *Journal of Quality Technology*, Vol. 37, No. 4, pp. 308-315.

Revfeim. K. J. A. (1990). More Approximations for the Cumulative and Inverse Normal Distribution. *American Statistician*, Vol. 44, pp. 63-67.

Rice, J. R. (1966). Experiments on Gram-Schmidt Orthogonalization. *Math. Comp*. Vol. 20, pp. 325-328

Rocke, D. M. (1989). Robust Control Chars. *Technometrics*, Vol. 31, No. 2, pp. 173-184.

Rohatgi, V. K. (1984). *Statistical Inference*. John Wiley & sons, New York.

Ruggoo, A., and Vandebroek, M. (2004). Bayesian Sequential D-D Optimal Designs. *Computational Statistics and Data Analysis*, Vol. 47, pp. 655-673.

Ryan, T. P. (1989). *Statistical Methods for Quality Improvement.* John Wiley & Sons.

Ryan, T. P. (2000). *Statistical Methods for Quality Improvement.* John Wiley & Sons.

Schilling, E. G. (1973). A Systematic Approach to the Analysis of Means. Part I. Analysis of Treatment effects. *Journal of Quality Technology*, Vol. 5, pp. 93-108.

Schucany, W. R., and Gray, H. L. (1968). A New Approximation Related to the Error Function, *Mathematics of Computation*, Vol. 22, pp. 201-202.

Seegrist, D. W. (1973). Least Squares Analysis of Experimental Design Models by Augmenting the Data with Side Conditions. *Technometrics*, Vol. 15, No. 3, pp. 643-645.

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*, D. van Nostrand Co., New York.

Shore, H. (2004). Response Modeling Methodology (RMM)-Current Statistical Distributions, transformations and Approximations as Special Cases of RMM. *Commun. Statist. Theory Meth.* Vol. 33, No. 7, pp. 1491-1510.

Shore, H. (2005). Accurate RMM-Based Approximations for the CDF of the Normal Distribution. *Communications in Statistics-Theory and Methods,* Vol. 34, pp. 507-513.

Sinha, S., and Wiens, D. P. (2002). Robust Sequential Designs for Nonlinear Regression. *The Canadian Journal of Statistics*, Vol. 30, No. 4, pp. 601-618.

Smith, K. (1918). On the Standard Deviations of Adjusted and Interpolated of An Observed Polynomial Function and Its Constants and the Guidance They Give Towards A Proper Choice of the Distribution of Observations. *Biometrika*, Vol. 12, pp. 1-58.

Smith. G. M. (1998). *Statistical Process Control and Quality Improvement*. Prentice-Hall, Inc.

Smoktunowicz, A., Barlow, J. L. and Langou, J. (2006). A Note on the Error Analysis of Classical Gram-Schmidt. *Numerische Mathematik*. Vol. 105, pp. 299-313.

Song, D., and Wong, W. K. (1999). On The Construction of Grm-Optimal Designs. *Statistica Sinica,* Vol. **9**, pp. 263-272.

Spendley, W., Hext, G. R., and Himsworth, F. R. (1962). Sequential Application of Simplex Designs in Optimization and Evolutionary Operation. *Technometrics,* Vol. 4, pp. 441-461.

St. John, R. C., and Draper, N. R. (1975). D-optimality for Regression Designs: A Review. *Technometrics*, Vol. 17, No. 1, pp. 15-23.

Stathopoulos, A. and Wu, K. (2002). A Block Orthogonalization Procedure with Constant Synchronization Requirements. *SIAM Journal on Scientific Computing*, Vol. 23, No. 6, pp. 2165-2182.

Sun, Y., Han, B., Liu, W., Zhang, J., Gao, X. (2007). Substrate Induction and Statistical Optimization for the Production of Chitosanase from Microbacterium sp. OU01. *Bioresource Technology,* Vol. 98, pp. 1548-1553.

Sztendur, E. M., and Diamond, N. T. (2002). Extensions to Confidence Region Calculations for

the Path of Steepest Ascent. *Journal of Quality Technology,* Vol. 34, No. 3, pp. 289-296.

Tew, J. D. and Wilson, J. R. (1992). Validation of Simulation Analysis Methods for the Schruben-Margolin Correlation-Induction Strategy. *Operations Research*, Vol. 40, No. 1, pp. 87-103.

Tsai, T. R., Lin, J. J., Wu, S. J., and Lin, H. C. (2005). On Estimating Control Limits of X-bar Chart When the Number of Subgroups Is Small. *Int J Adv Manuf Technol,* Vol. 26, pp. 1312-1316.

Vardeman, S. B. (1999). A Brief Tutorial on the Estimation of the Process Standard Deviation. *IIE Transactions,* Vol. 31, pp. 503-507.

Vedder, J. D. (1993). An Invertible Approximation to the Normal-Distribution Function. *Comput. Stat. Data. An.*, Vol. 16, pp. 119-123.

Waissi, G., and Rossin, D. F. (1996). A Sigmoid Approximation to the Standard Normal Integral. *Applied Mathematics and Computation,* Vol. 77, pp. 91–95.

Wald, A. (1943). On the Efficient design of Statistical Investigation. *The Annals of Mathematical Statistics*, Vol. 14, pp. 134-140.

Watson, S. (1997). Evaluation of Semivariance Estimators Under Normal Conditions. *The Statistician,* Vol. 46, No. 4, pp. 495-503.

Wludyka, P. S., and Nelson, P. R. (1997). An Analysis-of-Means-Type Test for Variance from Normal Population. *Technometrics*, Vol. 39, pp. 274-285.

Woodall, W. H. (2000). Controversies and Contradictions in Statistical Process Control: with discussion. *Journal of Quality Technology*, Vol. 32, pp. 341-378.

Woodall, W. H., Spitzner, D. J., Montgomery, D. C., and Gupta, S. (2004). Using Control Charts to Monitor Process and Product Quality Profiles. *Journal of Quality Technology*, Vol. 36, No. 3, pp. 309-320.

Wynn, H. P. (1970). The Sequential Generation of *D*-optimum Experimental Designs. *The Annals of Mathematical Statistics*, Vol. 41, pp. 1655-1664.

Wynn, H. P. (1972). Results in the Theory and Construction of D-optimum Experimental Designs. *Journal of the Royal Statistical Society,* Series B Vol. 34, No. 1, pp. 133-147.

Yang, C-H, and Hillier, F. S. (1970). Mean and Variance Control Chart Limits Based on a Small Number of Subgroups. *Journal of Quality Technology,* Vol. 2, pp. 9-16.

Zelon, M., and Severo, N. C. (1964). Probability Function, *Handbook of Mathematical Functions,* M. Abramowitz and I. A. Stegun (editors), 925-995, U. S. Department of Commerce, *Applied Mathematics Series,* 55.