NONPARAMETRIC PROCEDURES FOR PROCESS CONTROL

by

Saad Taha Bakir

Dissertation submitted to the Graduate Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

M. R. Reynolds, Jr., Chairman

Klaus Hinkelmann

L. D. Lee

J. A. Nachlas

W. R. Pirie

August, 1977

Blacksburg, Virginia

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to for his interest, effort, and encouragement in directing this research. The author wishes to thank for co-reading this manuscript and for his constructive comments and suggestions. The author also wishes to thank , , and for their continued support and encouragement during the pursuit of his degree. He acknowledges for correcting one of the results in this research; and for his suggestions on the improvement of the presentation of the dissertation. Also, thanks are due to of the Department of Statistics, University of Wisconsin, for suggesting one of the main references used in this dissertation. The author is indebted to for her excellent typing of this manuscript.

Finally, the author expresses his deep gratitude to his wife, , for her patience and invaluable support.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table		Page

Table		Page

CHAPTER I

INTRODUCTION

1.1 The Problem

This dissertation is concerned with the development of new nonparametric procedures for the control chart problem where the purpose is to maintain the quality of the output of some process with respect to some characteristic--weight, length, percent defective, etc. This purpose is usually accomplished by observing the output of the process continuously in order to detect any change in quality as soon as it occurs in which case the mechanism of the process is modified in order to turn out product that conforms to the desired quality standard. This problem arises in industrial quality control and also in medical and environmental monitoring. For more discussion on control charts and their applications, see Duncan (1974), and Phillips (1969).

To state the problem more precisely, let $X_1, X_2, \ldots$ be a sequence of independent observations, taken at regular time intervals from the process. Assume that the observations have a symmetric continuous cumulative distribution function $F(X|\Delta)$, where Δ is the center of symmetry (= the median, or the mean if it exists). Suppose that it is known that Δ may deviate in time from a specified control value Δ_0. Then the problem is that of detecting any deviation of Δ from Δ_0 as quickly as possible after it occurs. In some situations, more than one observation, i.e., a sample of size more than one is taken at each time point.

In general, a control chart procedure may be described in the following way. At stage n ($n = 1, 2, \ldots$) the value X_n is observed and based on the values of $X_1, X_2, \ldots, X_n$, there are two possible decisions: (1) Decide that a significant deviation has occurred away from Δ_0 and stop sampling so that a rectifying action may be taken on the process. In this case it is said that the procedure "signals" to stop sampling, according to quality control terminology. (2) Decide that no deviation has occurred away from Δ_0 and continue sampling by observing X_{n+1} at stage n+1. If Δ remains at the control value Δ_0, the process is said to be "in-control". If Δ shifts away from Δ_0, the process is said to be "out-of-control".

It is desirable that a control chart procedure should signal to stop sampling quickly any time Δ shifts away from Δ_0. When Δ remains at the control value, it is desirable that the procedure should continue sampling since the process is operating properly. In the long run, however, there will be enough variation in the observations to the extent that most reasonable procedures will eventually signal to stop regardless of the true value of Δ; i.e., the probability of deciding $\Delta \neq \Delta_0$ and stopping the sampling is one for any value of Δ including $\Delta = \Delta_0$. For this reason, the concepts of type I and type II error probabilities in standard testing of statistical hypotheses are not, in general, meaningful for control chart procedures. Also, because a control chart procedure never stops sampling and accepts the hypothesis $\Delta = \Delta_0$, its structure does not conform with standard sequential testing of statistical hypotheses. The usual criterion for evaluating control chart procedures is the time it takes the procedure to signal when the process is in-control and when it is out-of-control. The time it takes for a control chart procedure to signal is called the "run length" and the quantity frequently used to characterize the distribution of the run length is its expected value, called the average run length (ARL). If Δ remains at the control value, it is desirable that the ARL be large so that the frequency of false signals is low. If Δ shifts away from the control value, the ARL, counted from the time of the shift, should be small. When comparing two control chart procedures, it is a standard practice to keep their average run lengths the same when the process is in-control and compare their average run lengths at specific shifts $\Delta \neq \Delta_0$. Let $ARL_1(\Delta)$ and $ARL_2(\Delta)$ denote the average run lengths of procedure 1 and procedure 2, respectively, when the median of the observations is Δ.Then procedure 1 is said to be more efficient at a given $\Delta = \Delta_1$ than procedure 2 if $ARL_1(\Delta_1) < ARL_2(\Delta_1)$ subject to the condition that $ARL_1(\Delta_0) \geq ARL_2(\Delta_0)$.

1.2 Objectives and Motivations

The objective of this dissertation is to develop nonparametric control chart procedures that are competitive with existing parametric procedures. This involves constructing procedures that are relatively simple to apply in practice and deriving exact or approximate expressions for the average run lengths of these procedures. The efficiency of the new nonparametric procedures will then be studied by comparing their average run lengths to the average run lengths of existing parametric procedures.

The main motivation for proposing nonparametric procedures is that most of the current control chart procedures are based on the assumption that the observations are normally distributed with known or estimated variance. However, in some practical situations the normality assumption cannot be maintained or there is not enough information to provide an estimate of the variance or the shape of the distribution. In such situations, a nonparametric procedure would be useful but there appears to be very few nonparametric procedures available in the literature on control charts.

Special attention has been given in this dissertation to the development of methods for computing the ARLs of the proposed procedures. The reason for this is that the ARL is an important criterion for designing and evaluating a control chart procedure and also for comparing several control chart procedures. While it is desirable to compute the ARL of a procedure exactly, it is frequently not possible to obtain exact analytic formula for it. Therefore, some attention has been given to the derivation of approximate analytic formulas for the ARLs of the proposed procedures.

1.3 Approach of the Investigation

The construction of the new nonparametric procedures is approached by proposing simple methods for ranking the observations and then using the ranks to construct test statistic for the procedures.

In Chapter IV, a method for ranking the observations within groups will be employed to develop two nonparametric control chart procedures. This method of ranking requires that groups of observations be obtained sequentially, or that the observations can be divided in groups of fixed size. In quality control applications, the observations are frequently taken in groups (samples of fixed size; see, for example, Duncan (1974)). If a process turns out individual observations that do not fall naturally into groups, then the observations have to be divided artificially into groups of the same size, say, g. For example, the first group can be made to contain the first g observations, the second group to contain the second g observations, and so on. The first procedure is introduced in Section 4.1 and it is a cumulative sum control chart-type procedure. Also, a method for computing the exact ARL of this procedure by a Markov chain approach is given in Section 4.1. Section 4.2 is devoted to the computation of tables that assist in the application of the procedure. The second procedure is introduced in Section 4.3 and it is a linear barrier type-procedure (see Chapter II). A method for computing its exact ARL by a Markov chain approach is also given in Section 4.3. In Section 4.4, several comparisons are made between the two proposed nonparametric procedures and two parametric procedures which are the Shewhart's procedure (see (2.1)) and the CSCC procedure (see (2.3)). The comparisons are accomplished by comparing exact values of the ARLs of the different procedures at various values of $\Delta \neq \Delta_0$, subject to the condition that ARLs of the compared procedures are the same when $\Delta = \Delta_0$.

In Chapter V, an alternative method of ranking the observations is proposed and is employed to construct two additional nonparametric control chart procedures. According to this method, each observation is ranked only with respect to a specified number of the observations immediately preceding it. These two procedures do not require the observations to be obtained in groups as was the case with the procedures of Chapter IV. In Section 5.1, some properties of the new ranks are investigated. In Section 5.2, a linear barrier procedure and a CSCC procedure based on this method of ranking are proposed. It was not possible to develop a method for computing the exact ARLs of these two procedures. A few values of the ARLs of the procedures were computed by simulation.

In Chapter VI, approximate analytic formulas are derived for the ARL of each of the four nonparametric control chart procedures that are proposed in this dissertation. The approach for obtaining these formulas was by approximating the probabilistic behavior of each procedure by the behavior of a Brownian motion process defined on the interval $(0, \infty)$. Also, Section 6.2 contains a modification for an approximate analytic formula that was originally derived by Reynolds (1975b) for the ARL of the parametric CSCC. This modification makes the approximation more accurate.

In Chapter VII, conclusions and recommendations are given together with some problems of interest for further research.

Finally, it was necessary to include in the dissertation an extra chapter, Chapter III, in which certain results concerning the Wilcoxon signed-rank statistic (defined in (3.1)) are derived. In Section 3.1, tables for the distribution of the Wilcoxon signed-rank statistic are compiled under normal shift alternatives. These tables are needed to compute the exact ARL for the procedures of Chapter IV. In Section 3.2, the exact variance of the Wilcoxon signed-rank statistic is derived because values of the variance are needed in Section 6.3 to compute approximate values of the ARLs of the procedures of Chapter IV.

CHAPTER II

REVIEW OF RELATED LITERATURE

2.1 Parametric Control Chart Procedures

The Shewhart's Procedure.

One of the earliest control chart procedures was the Shewhart control chart which was introduced by Shewhart (1931). A two-sided Shewhart control chart procedure signals at the first n for which

$$X_n \notin (\Delta_0 - k\sigma_X, \Delta_0 + k\sigma_X) \ , \qquad (2.1)$$

where σ_X^2 is the variance of the X_i's, which is assumed known, and $k > 0$ is a constant chosen to give a specific ARL when $\Delta = \Delta_0$. Otherwise, sampling is continued. A one-sided Shewhart procedure for detecting shifts in the positive direction signals at the first n for which $X_n \leq \Delta_0 + k\sigma_X$. For detecting shifts in the negative direction, a one-sided Shewhart procedure signals at the first n for which $X_n \leq \Delta_0 - k\sigma_X$. The constant k is usually taken to be equal to three when the observations are normally distributed. The run length of a Shewhart procedure is the number of observations taken before a signal occurs and it follows a geometric distribution. Thus, the ARL of a Shewhart procedure is the mean of this geometric distribution which can be computed easily. If $p(\Delta)$ denotes the probability that a Shewhart procedure does not signal at a given stage, then the ARL of the procedure at a specific Δ can be written as

$$\text{ARL}(\Delta) = [1 - p(\Delta)]^{-1} \ . \qquad (2.2)$$

It is seen that a Shewhart procedure uses the information in only one observation at a time and generally is not efficient for detecting small deviations in Δ from the control value. Various modifications of the Shewhart procedure, such as charts with warning limits will in general improve its performance, see Page (1955). The design of the Shewhart chart is usually based on the assumption that the observations are normally distributed with known variance.

Parametric Cumulative Sum Control Chart Procedure.

A more recent and, generally, more efficient procedure than Shewhart's procedure is the cumulative sum control chart (CSCC) which essentially uses the cumulative sum of the observations as the test statistic. The CSCC procedure was originally introduced by Page (1954) and a comprehensive account of it is contained in a monograph by van Dobben de Bruyn (1968). To describe the CSCC procedure, it can be assumed that $\Delta_0 = 0$, without loss of generality, since one can always work with the sequence $(X_i - \Delta_0)$ ($i = 1, 2, \ldots$) instead of the original sequence X_i $(i = 1, 2, \ldots)$. The CSCC procedure for detecting deviations in the positive direction signals to stop at the first n for which

$$\sum_{i=1}^{n} (X_i - k^+) - \min_{0 \leq m \leq n} \sum_{i=1}^{m} (X_i - k^+) \geq h^+ \sigma_X , \quad (2.3)$$

where $k^+ \geq 0$ (called reference value) and $h^+ > 0$ (called decision interval) are preassigned numbers and are considered as parameters of the CSCC procedure, and where $\sum_{i=1}^{0} \equiv 0$. The criterion for choosing

k^+ and h^+ is that the ARL of the resulting procedure be a minimum at a given $\Delta \neq 0$ subject to condition that the ARL be a specified value when $\Delta = 0$. To detect deviations in the negative direction, the CSCC procedure signals at the first n for which

$$\max_{0 \leq m \leq n} \sum_{i=1}^{m} (X_i + k^-) - \sum_{i=1}^{n} (X_i + k^-) \geq h^- \sigma_X , \qquad (2.4)$$

where $k^- \geq 0$ and $h^- > 0$ are preassigned parameters of this one-sided procedure. A two-sided CSCC procedure for detecting both positive and negative deviations in Δ from the control value signals at the first n for which either one of the above two inequalities is satisfied. A two-sided symmetric CSCC procedure can be constructed by taking $k^- = k^+$ and $h^- = h^+$.

Page (1954) established that a CSCC procedure is equivalent to a sequence of standard sequential tests of statistical hypotheses. The one-sided positive CSCC procedure, for example, is equivalent to a sequence of sequential statistical tests in which $\sum_{i=1}^{n} (X_i - k^+)$ is accumulated until either the sum is $\geq h^+ \sigma_X$ or ≤ 0. If $\sum_{i=1}^{n} (X_i - k^+) \geq h^+ \sigma_X$, the hypothesis that the process is in-control, i.e., $\Delta = \Delta_0 = 0$ is rejected and the CSCC procedure signals to stop. If $\sum_{i=1}^{n} (X_i - k^+) \leq 0$, the hypothesis that the process is in-control is accepted and a new test is started with the summation reset equal to zero before accumulating the next observation X_{n+1}. In contrast to Shewhart control charts, a CSCC procedure uses all past observations up to the current stage and is therefore expected to be more efficient, in terms of its ARL, than Shewhart control charts. Barnard (1959) devised a practical device called the V-mask, which is a V-shaped hole made in a piece of cardboard, to carry out a CSCC procedure in a simple manner.

In general, there is no simple way to determine the ARL of a CSCC procedure. A method for computing the exact ARL of a CSCC procedure by solving integral equations, of the form given in (2.5), is described in Chapter III of van Dobben de Bruyn (1968). Consider a positive sided CSCC procedure with parameters k and h and based on $X_1, X_2, \ldots$. Let $L(z|\Delta)$ denote the average run length of the procedure when the cumulative sum has value $0 \leq z \leq h$ given a specified value of Δ. Then $L(z|\Delta)$ satisfies the following integral equation

$$L(z|\Delta) = 1 + L(0|\Delta)G(-z|\Delta)+\int_0^h g(u-z|\Delta)L(u|\Delta)du \quad 0 \leq z \leq h \ , \qquad (2.5)$$

where $G(u|\Delta)$ and $g(u|\Delta)$ are the cumulative distribution and density functions of (X-k). Usually, interest is only in $L(0|\Delta)$, the ARL of a procedure starting at zero. However, as the above integral equation indicates, one has to solve the equation for all values of $0 \leq z \leq h$ in order to obtain $L(0|\Delta)$ and, in general, no analytic solution is possible. van Dobben de Bruyn used numerical methods to solve for $L(0|\Delta)$ and compiled tables for the ARL when the observations follow a normal distribution with known variance. These tables are exact up to a negligible error due to the numerical solution.

Usually it is desirable to have a simple, preferably analytic, way for computing the ARL of a control chart procedure. Accordingly, attention has been directed towards obtaining approximate, usually asymptotic, formulas for the ARL. Reynolds (1975b) derived one such asymptotic analytic formula for the ARL by approximating the behavior of the CSCC procedure by a Brownian motion process on the interval $(0, \infty)$. He reported that the approximation is not satisfactory unless some modification is made in the asymptotic formula and he gave one suggestion in this regard.

2.2 Nonparametric Control Chart Procedures

The Parent-Reynolds Signed Sequential Ranks Procedures.

Parent (1965) proposed a nonparametric control chart procedure based on the signed sequential ranks (SSR) of the observations. If R^+_{ij} denotes the rank of $|X_i|$ in the set $\{|X_1|, |X_2|, \ldots, |X_j|\}$, $i \leq j$, and sign (X_i) is defined to be 1 if $X_i \geq 0$ and -1 otherwise, then the signed sequential rank of X_i is defined to be $V_i = \text{sign}(X_i)R^+_{ii}$. Thus the signed sequential rank of a given observation is computed with respect to all observations preceding it and is different from the usual Wilcoxon signed-rank which is defined in (3.2). The test statistic used by Parent was $S_n = \sum_{i=1}^{n} V_i/i$, and his control chart procedure signals at the first n for which $S_n \notin (-a, a)$ for some pre-assigned positive number a; otherwise, sampling is continued. Reynolds (1972) proposed two nonparametric control chart procedures both based on the modified statistic $Z^+_n = \sum_{i=1}^{n} V_i/(i+1)$. The first may be termed as a linear barrier-type procedure and it signals at the first n for which

$$Z_n^+ \notin (-a, a) , \tag{2.6}$$

where a > 0 is determined so that the resulting procedure has a preassigned ARL when $\Delta = 0$. Otherwise, sampling is continued. The second, is a CSCC-type procedure and it signals at the first n for which

$$(Z_n^+ - nk) - \min_{0 \le m \le n} (Z_m^+ - mk) \ge h \text{ or } \max_{0 \le m \le n} (Z_m^+ + mk) - (Z_n^+ + nk) \ge h ,$$

where $k \geq 0$ and $h > 0$ are the parameters of this two-sided procedure and Z_0^+ is defined to be zero. Otherwise, sampling is continued. Again there is no simple way of computing the exact ARL of procedures based on signed sequential ranks. Asymptotic formulas for the ARLs of these procedures based on a Brownian motion approximation were derived by Reynolds (1972). He checked the accuracy of this Brownian motion approximation by comparing it to values of ARL obtained by simulation. It was found that the approximation is quite satisfactory in case of a linear barrier-type procedure. However, a modification must be made to improve the accuracy of the Brownian motion approximation in the case of a CSCC-type procedure. Also, Reynolds compared the average run lengths of these nonparametric signed sequential rank procedures with the average run lengths of the parametric CSCC procedure at different values of shifts in Δ away from the control value. It was found that the signed sequential rank procedures are comparable in efficiency to the parametric CSCC procedure for small shifts from the control value. However, the efficiency of these nonparametric procedures decreases for large shifts.

Some sequential nonparametric hypothesis testing procedures use ranking within groups to reduce the effort of ranking and to provide a sequence of independent test statistics. The idea of ranking the observations within groups was originally employed by Wilcoxon, et al. (1963) to develop a nonparametric sequential probability ratio test for the two-sample (equality of two populations) problem against Lehmann alternative hypotheses (see Lehmann (1953)). Also, Weed and Bradley (1971) used this same idea to develop a nonparametric sequential probability ratio test for the one-sample problem (testing that a population is symmetric about a given value) against Lehmann alternatives. The idea of ranking the observations within groups requires that groups of observations be obtained sequentially, or that the observations can be divided into groups of fixed size.

CHAPTER III

RESULTS CONCERNING THE WILCOXON SIGNED-RANK STATISTIC

The present chapter is devoted to the derivation of some results concerning the Wilcoxon signed-rank statistic which will be defined in equation (3.1). These results will be used to compute the average run lengths of the control chart procedures to be proposed in Chapter IV. In Section 3.1, tables will be compiled for the distribution of the Wilcoxon signed rank statistic when the observations come from a normal population. These tables will be used to compute exact values of the average run length as will be developed in Chapter IV. In Section 3.2, an analytic formula will be derived for the variance of the Wilcoxon signed-rank statistic. This formula will be used to compute approximate average run lengths of the proposed control chart procedures as will be developed in Chapter VI.

To define the Wilcoxon signed-rank statistic, let $X_1, X_2, \ldots, X_N$ be a random sample of size $N \geq 1$ from a population with a symmetric continuous cumulative distribution function $F(x|\Delta)$ where Δ is the center of symmetry (=the median, or the mean if it exists). Let $c(u)$ be equal to 1 or 0 according as u is $\geq$ or $<$ 0. Define

$$R_{jN} = \sum_{i=1}^{N} c(|X_j| - |X_i|) \ , \ 1 \leq j \leq N \quad .$$

In other words, R_{jN} is the rank of $|X_j|$ in the set $\{|X_1|, |X_2|, \ldots, |X_N|\}$. The Wilcoxon signed-rank statistic, which was originally introduced by Wilcoxon (1945), can be written in the form

$$SR_N = \sum_{j=1}^{N} Y_j \ , \tag{3.1}$$

where

$$Y_j = \text{sign}(X_j) R_{jN} \ . \tag{3.2}$$

The Y_j's will be called the Wilcoxon signed-ranks.

3.1 Distribution of the Wilcoxon Signed-Rank Statistic

When $\Delta = 0$, it is known in nonparametric statistics that the distribution of the Wilcoxon signed-rank statistic SR_N does not depend on $F(X|\Delta)$ and this distribution has been tabulated for samples of size up to $N = 15$, see Hollander and Wolfe (1973). When $\Delta \neq 0$, the distribution of SR_N depends on $F(X|\Delta)$.

To show how the tables of this section were compiled, a general method for computing the exact distribution of the Wilcoxon signed-rank statistic will now be outlined.

Define the random vector $\underset{\sim}{Z} = (Z_{(1)}, Z_{(2)}, \ldots, Z_{(N)})$ by

$$Z_{(j)} = \begin{cases} 1 & \text{if the jth order statistic among } |X_1|, \ldots, |X_N| \\ & \text{corresponds to a nonnegative observation,} \\ -1 & \text{otherwise.} \end{cases}$$

The Wilcoxon signed rank statistic can then be written in the form

$$SR_N = \sum_{j=1}^{N} jZ_{(j)} \ .$$

There are 2^N possible values $\underset{\sim}{z}$ of the random vector $\underset{\sim}{Z}$. The probability that $\underset{\sim}{Z}$ assumes a particular value $\underset{\sim}{z}$ will be denoted by $P(\underset{\sim}{Z} = \underset{\sim}{z}|\Delta)$. The probabilities $P(\underset{\sim}{Z} = \underset{\sim}{z}|\Delta)$ are usually called rank order probabilities. The probability that SR_N assumes a particular value ℓ is given by

$$\Pr(SR_N = \ell|\Delta) = \Sigma\, P(\underset{\sim}{Z} = \underset{\sim}{z}|\Delta) \;, \qquad (3.3)$$

where the summation extends over all values of $\underset{\sim}{z}$ for which $SR_N = \ell$. The possible values ℓ of SR_N are

$$-N(N+1)/2,\ -(N(N+1)/2)+2,\ \ldots,\ N(N+1)/2 \;.$$

Thus, the computation of the distribution of SR_N requires the determination of the rank order probabilities $P(\underset{\sim}{Z} = \underset{\sim}{z}|\Delta)$ for all possible values of $\underset{\sim}{Z}$. When the observations have a normal distribution with mean Δ and variance 1, Lever (1973) expressed the rank order probabilities in the form

$$P(\underset{\sim}{Z} = \underset{\sim}{z}|\Delta) = N! \underset{0<t_1<\ldots<t_N<\infty}{\int \cdots \int} \prod_{j=1}^{N} \phi(t_j - z_{(j)}\Delta)\,dt_j \;, \qquad (3.4)$$

where ϕ is the standard normal density function. The above expression remains valid for a normal distribution with mean μ and variance σ^2 provided that Δ is expressed in units of standard deviations, i.e., $\Delta = \mu/\sigma$. Under this normality assumption, numerical integration of multiple integrals was employed by Lever (1973) to compute $P(\underset{\sim}{Z} = \underset{\sim}{z}|\Delta)$ for $1 \le N \le 12$ and Δ = 0.0, 0.1, 0.2 (0.2) 1.0, 1.5, 2.0, 3.0. She reported that the results are accurate to 9 decimal places. In this dissertation, the values of the rank order probabilities as computed by Lever, are used to compute the distribution of SR_N in accordance with (3.3). The distribution of SR_N, thus computed, is given in Tables I, II, and III for N = 2, 6, 10; Δ = 0.2, 0.6, 1.0, 2.0, 3.0 where Δ is the mean of a normal distribution with variance 1. These last values of Δ were chosen to represent a range from a small to a large shift in the mean of the normal population.

Klotz (1963) computed the probabilities in (3.4) for $1 \leq N \leq 10$; $\Delta = 0\ (0.25)\ 1.5\ (0.5)\ 3.0$. He reported that the results were accurate to 4 decimal places. However,the computed values of the rank order probabilities were not published in his paper. Arnold (1965) computed the probabilities $P(\underset{\sim}{Z} = \underset{\sim}{z}|\Delta)$ with $1 \leq N \leq 10$ for shifts Δ = 0.25, 0.50, 1.0, 2.0 and 3.0 under the non-central t distribution with 1/2, 1, 2, and 4 degrees of freedom. The computed values of the rank order probabilities were used to present the power of the Wilcoxon signed-rank statistic; but, they were not published.

3.2 Variance of the Wilcoxon Signed-Rank Statistic

The exact variance of SR_N will now be derived and some computations will be included when the observations have normal, double exponential, and uniform distributions. An analytic formula for variance of SR_N will be needed in Chapter VI to compute a Brownian motion approximation to the ARL of a nonparametric control chart procedure based on SR_N.

Table I. Distribution of SR_N Under Normal Distribution with Mean Δ and Variance 1. Sample Size N = 2.

ℓ	$Pr(SR_N = \ell \mid \Delta)$				
	Δ = 0.20	Δ = 0.60	Δ = 1.00	Δ = 2.00	Δ = 3.00
3	0.335541816	0.526708536	0.707860980	0.955017305	0.997302026
1	.275809487	.275219508	.213489412	.042643828	.002686929
-1	.211626317	.122857180	.053478112	.001821299	.000009223
-3	0.177022395	0.075214770	.025171487	.000517569	.000001822

Table II. Distribution of SR_N Under Normal Distribution with Mean Δ and Variance 1. Sample Size N = 6.

ℓ	$\Pr(SR_N = \ell \mid \Delta)$				
	Δ = 0.20	Δ = 0.60	Δ = 1.0	Δ = 2.0	Δ = 3.0
21	0.037778085	0.146120474	0.35685899	0.871031223	0.991927896
19	.034821825	.107824960	.188287710	.101791128	.007912783
17	.032008144	.079814025	.103465245	.015889464	.000126967
15	.058869011	.120433588	.122649244	.008683712	.000031336
13	.053428183	.C86824935	.066253747	.001580473	.000000804
11	.072561309	.094792316	.056540339	.000650698	.000000158
9	.086156586	.089452944	.042312734	.000280911	.000000053
7	.078890165	.06628579	.024417298	.000060067	.000000002
5	.071845270	.049357214	.041333909	.000017842	0.0
3	.08090168	.045102978	.010667515	.000008568	0.0
1	.074592241	.035508672	.007247187	.000004565	0.0
-1	.065751046	.024107986	.003680577	.000000786	0.0
-3	.060589266	.018845360	.002440578	.000000350	0.0
-5	.043665604	.011033982	.001164566	.000000107	0.0
-7	.040023834	.008666071	.000805740	.000000062	0.0
-9	.036981399	.007012004	.000589726	.000000039	0.0
-11	.024632752	.003621693	.000226199	.000000004	0.0
-13	.014957756	.001854466	.000099534	.000000002	0.0
-15	.013717112	.001469095	.000070154	0.0	0.0
-17	.006365312	.000601741	.000026070	0.0	0.0
-19	.005927605	.000501483	.000020069	0.0	0.0
-21	0.005547338	0.000425509	0.000015949	0.0	0.0

Table III. Distribution of SR_N Under Normal Distribution with Mean Δ and Variance 1. Sample Size N = 10.

ℓ	$Pr(SR_N = \ell \mid \Delta)$				
	Δ = 0.20	Δ = 0.60	Δ = 1.0	Δ = 2.0	Δ = 3.0
55	0.004253371	0.040537017	0.177721459	0.794431040	0.986582725
53	.004037279	.033074132	.113654593	.136853339	.012955546
51	.003830798	.027147327	.075006680	.031927332	.000354260
49	.007274335	.045064863	.103596488	.021814135	.000100307
47	.006886937	.037003465	.069759711	.006249866	.000005354
45	.009785369	.045947306	.072668435	.003443726	.000001314
43	.012346028	.050749631	.067972684	.001862942	.000000433
41	.014542117	.051729304	.057487937	.000756708	.000000048
39	.016395306	.050393659	.046800101	.000330663	.000000010
37	.020508638	.054593903	.042789383	.000173929	.000000002
35	.023868792	.054728151	.036322131	.000089940	.000000001
33	.024828645	.047609746	.025699810	.000032995	.000000000
31	.027763856	.048751615	.023316718	.000017019	.000000000
29	.029931521	.045280614	.018170462	.000007710	.000000000
27	.032106324	.042804839	.014839286	.000004177	.000000000
25	.035443043	.041263369	.012363673	.000002485	.000000000
23	.036556994	.036900346	.009353569	.000001001	.000000000
21	.037373890	.032815878	.007123126	.000000488	.000000000
19	.039661866	.030693230	.005807489	.000000265	.000000000
17	.039751392	.026636904	.004310109	.000000123	.000000000

Table III (Continued). Distribution of SR_N Under Normal Distribution with Mean Δ and Variance 1. Sample Size N = 10.

ℓ	$Pr(SR_N = \ell \mid \Delta)$				
	Δ = 0.20	Δ = 0.60	Δ = 1.0	Δ = 2.0	Δ = 3.0
15	.040047563	.023618915	.003330912	.000000072	.000000000
13	.040257981	.021042505	.002637023	.000000042	.000000000
11	.039941933	.018164795	.001941565	.000000017	.000000000
9	.038681404	.015501887	.001449372	.000000005	.000000000
7	.038272995	.013411710	.001069779	.000000002	.000000000
5	.036791225	.011411227	.000818606	.000000002	.000000000
3	.034660518	.009420196	.000589309	.000000000	.000000000
1	.033471889	.008094029	.000450079	.000000000	.000000000
-1	.031439670	.006661578	.000334132	.000000000	.000000000
-3	.028851545	.005433629	.000235580	.000000000	.000000000
-5	.027124695	.004528863	.000174155	.000000000	.000000000
-7	.024832964	.003659076	.000124025	.000000000	.000000000
-9	.021320343	.002720497	.000079047	.000000000	.000000000
-11	.020243133	.002357637	.000064040	.000000000	.000000000
-13	.017886560	.001835527	.000044176	.000000000	.000000000
-15	.015898777	.001472043	.000032266	.000000000	.000000000
-17	.014066542	.001182842	.000023932	.000000000	.000000000
-19	.012223745	.000892543	.000015440	.000000000	.000000000
-21	.010283009	.000657183	.000010469	.000000000	.000000000
-23	.008858907	.000517521	.000007266	.000000000	.000000000
-25	.007580316	.000396066	.000005023	.000000000	.000000000
-27	.006062769	.000282665	.000003230	.000000000	.000000000

Table III(Continued). Distribution of SR_N Under Normal Distribution with Mean Δ and Variance 1. Sample Size N = 10.

ℓ	$Pr(SR_N = \ell \mid \Delta)$				
	Δ = 0.20	Δ = 0.60	Δ = 1.0	Δ = 2.0	Δ = 3.0
-29	.005090964	.000218643	.000002358	.000000000	.0000000000
-31	.004136243	.000157138	.000001503	.000000000	.0000000000
-33	.003331976	.000116914	.000001061	.000000000	.0000000000
-35	.002887254	.000094147	.000000817	.000000000	.0000000000
-37	.002149713	.000060008	.000000427	.000000000	.0000000000
-39	.001519355	.000038061	.000000246	.000000000	.0000000000
-41	.001196938	.000027229	.000000163	.000000000	.0000000000
-43	.000908877	.000018974	.000000106	.000000000	.0000000000
-45	.000649376	.000012520	.000000067	.000000000	.0000000000
-47	.000413178	.000007396	.000000037	.000000000	.0000000000
-49	.000394653	.000006583	.000000032	.000000000	.0000000000
-51	.000188942	.000002949	.000000013	.000000000	.0000000000
-53	.000181098	.000002657	.000000012	.000000000	.0000000000
-55	.000173837	.000002407	.000000010	.000000000	.0000000000

Although one can compute the variance of SR_N from the distribution of SR_N, this distribution is available only in the few cases mentioned in the last section. Thus, it would be desirable to have a formula for computing the variance of SR_N without the need to know its distribution. Another derivation for the variance of the Wilcoxon signed-rank statistic has recently appeared in Lehmann (1975).

Define the random variables ψ_{ij}(i, j = 1, 2, ..., N) by

$$\psi_{ij} = \begin{cases} 1 & |X_i| \leq X_j \ , \\ 0 & -|X_i| < X_j < |X_i| \ , \\ -1 & X_j \leq -|X_i| \ . \end{cases}$$

It can be seen that $Y_j = \sum_{i=1}^{N} \psi_{ij}$ and

$$SR_N = \sum_{j=1}^{N} \sum_{i=1}^{N} \psi_{ij} \ .$$

Reynolds (1975a) derived the following moments of $\{\psi_{ij}\}$.

$$\xi = E(\psi_{ij}) = 1/2 - \int_{-\infty}^{\infty} F(-x)\,dF(x) \quad , \tag{3.5}$$

$$\theta = E(\psi_{jj}) = 1 - 2F(0) \ , \tag{3.6}$$

and

$$E(\psi_{ij}\psi_{k\ell}) = \begin{cases} 1/2 & i = k \neq j = \ell \quad , \\ 1 & i = j = k = \ell \quad , \\ \gamma & i \neq j = k \neq \ell \neq i \ , \\ 2\gamma & i = k \neq j \neq \ell \neq i \ , \\ 1/3 & i \neq j = \ell \neq k \neq i \ , \\ 1/2 & i \neq j = k = \ell \quad , \\ 0 & i = \ell \neq j = k \quad , \\ \theta^2/2 & i = k = \ell \neq j \quad , \\ E(\psi_{ij})E(\psi_{k\ell}) & \text{otherwise} \ , \end{cases}$$

where

$$\gamma = \int_{-\infty}^{\infty} F(-x)^2 \, dF(x) + \xi - 1/3 \ . \tag{3.7}$$

$$\mathrm{Var}(SR_N) = \mathrm{Var}\left(\sum_{j=1}^{N} Y_j\right)$$

$$= \sum_{j=1}^{N} \mathrm{Var}(Y_j) + 2\sum_{j=1}^{N-1}\sum_{k=j+1}^{N} \mathrm{Cov}(Y_j, Y_k) .$$

$$\mathrm{Var}(Y_j) = \mathrm{Var}\left(\sum_{i=1}^{N} \psi_{ij}\right)$$

$$= \sum_{i=1}^{N} \mathrm{Var}(\psi_{ij}) + 2\sum_{i=1}^{N-1}\sum_{r=i+1}^{N} \mathrm{Cov}(\psi_{ij}, \psi_{rj}) .$$

It is seen that

$$\sum_{i=1}^{N} \mathrm{Var}(\psi_{ij}) = \sum_{i\neq j}^{N} \mathrm{Var}(\psi_{ij}) + \mathrm{Var}(\psi_{jj})$$

$$= \sum_{i\neq j}^{N} (E\psi_{ij}^2 - (E\psi_{ij})^2) + E\psi_{jj}^2 - (E\psi_{jj})^2$$

$$= (N - 1)(1/2 - \xi^2) + (1 - \Theta^2) .$$

Also,

$$\sum_{i=1}^{N-1}\sum_{r=i+1}^{N} \mathrm{Cov}(\psi_{ij}, \psi_{rj}) = \sum_{\substack{i=1\\ i\neq j}}^{N-1}\sum_{\substack{r=i+1\\ r\neq j}}^{N} \mathrm{Cov}(\psi_{ij}, \psi_{rj}) + \sum_{r=j+1}^{N} \mathrm{Cov}(\psi_{jj}, \psi_{rj})$$

$$+ \sum_{i=1}^{j-1} \mathrm{Cov}(\psi_{ij}, \psi_{jj})$$

$$=(1/2)(N-1)(N-2)(1/3-\xi^2)+(N-j)(1/2-\xi\theta)$$
$$+\ (j-1)(1/2-\xi\theta)$$
$$=(1/2)(N-1)(N-2)(1/3-\xi^2)+(N-1)(1/2-\xi\theta)\quad .$$

Thus

$$\mathrm{Var}(Y_j) = (N-1)(1/2-\xi^2)+(1-\theta^2)+(N-1)(N-2)(1/3-\xi^2)$$
$$+\ (N-1)(1-2\xi\theta)$$
$$=(1/6)(N+1)(2N+1)-(N-1)^2\xi^2-\theta^2-2(N-1)\xi\theta\ .$$

The covariance of Y_j and Y_k, for $j < k$, will now be determined.

$$\mathrm{Cov}(Y_j,\ Y_k) = \mathrm{Cov}(\sum_{i=1}^{N}\psi_{ij},\ \sum_{r=1}^{N}\psi_{rk}) = \sum_{i=1}^{N}\ \sum_{r=1}^{N}\mathrm{Cov}(\psi_{ij},\ \psi_{rk})\ .$$

Computation of the covariances in the last summation can be clarified by looking at Figure 1. The nonzero covariances occur along the diagonal, in the jth column, and in the kth row of Figure 1. The sum of the covariances along the diagonal is given by

$$\sum_{i=1}^{N}\mathrm{Cov}(\psi_{ij},\ \psi_{ik}) = \mathrm{Cov}(\psi_{jj},\ \psi_{jk}) + \mathrm{Cov}(\psi_{kj},\ \psi_{kk})$$
$$+\ \sum_{\substack{i\neq j\\ i\neq k}}^{N}\mathrm{Cov}(\psi_{ij},\ \psi_{ik})$$
$$=\ 2(\theta^2/2-\xi\theta)+(N-2)(2\gamma-\xi^2)\ .$$

The sum of the covariances in the jth column, excluding the diagonal element, is given by

$$\sum_{\substack{i=1\\ i\neq j}}^{N}\mathrm{Cov}(\psi_{ij},\psi_{jk}) = \mathrm{Cov}(\psi_{kj},\psi_{jk}) + \sum_{\substack{i\neq j\\ i\neq k}}^{N}\mathrm{Cov}(\psi_{ij},\psi_{jk})$$

$$= (0 - \xi^2) + (N - 2)(\gamma - \xi^2) \quad .$$

The sum of the covariances in the kth row, excluding the diagonal element, is given by

$$\sum_{\substack{r=1 \\ i \neq j,k}}^{N} \mathrm{Cov}(\psi_{kj}, \psi_{rk}) = (N - 2)(\gamma - \xi^2) \quad .$$

Thus,

$$\mathrm{Cov}(Y_j, Y_k) = \sum_{i=1}^{N} \sum_{r=1}^{N} \mathrm{Cov}(\psi_{ij}, \psi_{rk})$$

$$= 2(\theta^2/2 - \xi\theta) + (N - 2)(2\gamma - \xi^2) - \xi^2$$

$$+ (N - 2)(\gamma - \xi^2) + (N - 2)(\gamma - \xi^2)$$

$$= 4(N - 2)\gamma - (3N - 5)\xi^2 + \theta^2 - 2\xi\theta \quad .$$

Using these results, it follows that

$$\mathrm{Var}(SR_N) = \sum_{j=1}^{N} \mathrm{Var}(Y_j) + 2 \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} \mathrm{Cov}(Y_j, Y_k)$$

$$= N\,\mathrm{Var}(Y_j) + N(N - 1)\,\mathrm{Cov}(Y_j, Y_k)$$

$$= 4N(N - 1)(N - 2)\gamma - 2N(N - 1)(2N - 3)\xi^2$$

$$+ N(N - 2)\theta^2 - 4N(N - 1)\xi\theta + N(N+1)(2N+1)/6 \quad .$$

If the distribution function $F(X)$ is symmetric about zero then $\xi = \theta = \gamma = 0$ and

$$\mathrm{Var}(SR_N) = N(N+1)(2N+1)/6 \quad .$$

i \ r	1	2	3	...	j	...	k	...	N
1	(1j,1k)	(1j,2k)	(1j,3k)	...	(1j,jk)	...	(1j,kk)	...	(1j,Nk)
2	(2j,1k)	(2j,2k)	(2j,3k)	...	(2j,jk)	...	(2j,kk)	...	(2j,Nk)
3	(3j,1k)	(3j,2k)	(3j,3k)	...	(3j,jk)	...	(3j,kk)	...	(3j,Nk)
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
j	(jj,1k)	(jj,2k)	(jj,3k)	...	(jj,jk)	...	(jj,kk)	...	(jj,Nk)
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
k	(kj,1k)	(kj,2k)	(kj,3k)	...	(kj,jk)	...	(kj,kk)	...	(kj,Nk)
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
N	(Nj,1k)	(Nj,2k)	(Nj,3k)	...	(Nj,jk)	...	(Nj,kk)	...	(Nj,Nk)

Figure 1. Display of the covariances in $\sum_{i=1}^{N} \sum_{r=1}^{N} \mathrm{Cov}(\psi_{ij}, \psi_{rk})$ where only the subscripts are shown.

The mean of SR_N is given by

$$E(SR_N) = E \sum_{j=1}^{N} \sum_{i=1}^{N} \psi_{ij}$$

$$= E \sum_{j=1}^{N} \psi_{jj} + E \sum_{j=1}^{N} \sum_{\substack{i=1 \\ i \neq j}}^{N} \psi_{ij}$$

$$= N(N - 1)\xi + N\,\theta \ .$$

To compute numerical values of $Var(SR_N)$ and $E(SR_N)$, it is necessary to determine the values of ξ, θ, and γ. The values of ξ, θ, and γ will now be determined when the observations have a normal distribution, a double exponential distribution, or a uniform distribution.

Values of ξ, θ, γ Under a Normal Distribution

The cumulative distribution function of a normal distribution with mean Δ and variance 1 is given by

$$F(x) = (2\Pi)^{-1/2} \int_{-\infty}^{x} \exp[-\,(u - \Delta)^2/2]\ du \ .$$

Then, by formula (3.5),

$$\xi = 1/2 - \int_{-\infty}^{\infty} F(-x)dF(x) = 1/2 - Pr(X_1 + X_2 \leq 0)$$

$$= 1/2 - \Phi(-\sqrt{2}\ \Delta) \ ,$$

and by formula (3.6),

$$\theta = 1 - 2F(0) = 1 - 2\Phi(-\Delta) \ ;$$

where $\Phi(u)$ is the standard normal cumulative distribution function. γ, however, cannot be evaluated analytically because $\int_{-\infty}^{\infty} F(-x)^2 dF(x)$ does not admit an analytic solution when $F(x)$ is the normal cumulative distribution function. This integral was evaluated numerically by Simpson's rule using 2000 subdivisions over the interval [-10, 10]. Table IV gives values of ξ, θ, and γ for a normal distribution with mean Δ = 0, 0.2, 0.25, 0.5, 0.6, 1.0, 2.0, 3.0 and variance 1.

Values of ξ, θ, and γ Under a Double Exponential Distribution

The density function of a double exponential distribution with mean Δ and variance $2\phi^2$ is given by

$$f(x) = (2\phi)^{-1} \exp[-|x - \Delta|/\phi] ,$$

where $-\infty < x < \infty$, $-\infty < \Delta < \infty$, and $\phi > 0$. The cumulative distribution function is given by

$$F(x) = \begin{cases} (1/2) \exp[(x - \Delta)/\phi] & x \leq \Delta , \\ 1 - (1/2) \exp[-(x - \Delta)/\phi] & x > \Delta . \end{cases}$$

To compute $\xi = 1/2 - \int_{-\infty}^{\infty} F(-x)dF(x)$, we have

$$F(-x) = \begin{cases} (1/2) \exp[-(x + \Delta)/\phi] & x \geq -\Delta , \\ 1 - (1/2) \exp[(x + \Delta)/\phi] & x < -\Delta . \end{cases}$$

When $\Delta \geq 0$, it is seen that

$$\int_{-\infty}^{\infty} F(-x)dF(x) = (2\phi)^{-1}\int_{-\infty}^{-\Delta} (1 - (1/2)\exp[(x + \Delta)/\phi])\exp[(x - \Delta)/\phi]\, dx$$

$$+ (2\phi)^{-1} \int_{-\Delta}^{\Delta}(1/2)\exp[-(x + \Delta)/\phi]\exp[(x - \Delta)/\phi]\, dx$$

Table IV. Values of ξ, θ, and γ Under a Normal Distribution with Mean Δ and Variance 1.

Δ	ξ	θ	$\int_{-\infty}^{\infty} F(-x)^2 dF(x)$	γ
0.0	0.0	0.0	0.3333333	0.0
0.20	.1114116	.1585194	.2291721	.0072504
0.25	.1383306	.1974126	.2062665	.0112637
0.50	.2602168	.3829250	.1132022	.0400856
0.60	.3020594	.4514938	.0859443	.0546704
1.0	.4213190	.6826894	.0230664	.1110520
2.0	.4976580	.9544998	.0001743	.1644989
3.0	0.4999888	0.9973002	0.0000001	0.1666556
$\Delta \to \infty$	1/2	1	0	1/6

$$+ (2\phi)^{-1} \int_{\Delta}^{\infty}(1/2)\exp[-(x + \Delta)/\phi] \exp[-(x - \Delta)/\phi] \, dx$$

$$=(1/2)\exp[-2\Delta/\phi] -(1/8)\exp[-2\Delta/\phi]$$

$$+ (2\phi)^{-1} \Delta \exp[-2\Delta/\phi] +(1/8)\exp[-2\Delta/\phi]$$

$$= (2\phi)^{-1} (\phi + \Delta) \exp[-2\Delta/\phi] .$$

Similarly, when $\Delta < 0$, it can be found that

$$\int_{-\infty}^{\infty} F(-x)dF(x) = 1/2 + (2\phi)^{-1} (\phi + (\Delta - \phi) \exp[2\Delta/\phi])$$

Thus,

$$\xi = \begin{cases} 1/2 - (2\phi)^{-1} (\phi + \Delta)\exp[-2\Delta/\phi] & \Delta \geq 0 , \\ -(2\phi)^{-1} (\phi + (\Delta - \phi)\exp[2\Delta/\phi]) & \Delta < 0 . \end{cases}$$

Also,

$$\theta = 1 - 2F(0) = \begin{cases} 1 - \exp[-\Delta/\phi] & \Delta \geq 0 , \\ \exp[\Delta/\phi] & \Delta < 0 . \end{cases}$$

It remains now to evaluate γ. When $\Delta \geq 0$,

$$\int_{-\infty}^{\infty} F(-x)^2 dF(x) = (2\phi)^{-1} \int_{-\infty}^{-\Delta} (1 - (1/2)\exp[(x + \Delta)/\phi])^2 \exp[(x - \Delta)/\phi] \, dx$$

$$+ (2\phi)^{-1} \int_{-\Delta}^{\Delta}(1/4)\exp[-2(x + \Delta)/\phi] \exp((x - \Delta)/\phi] \, dx$$

$$+ (2\phi)^{-1} \int_{\Delta}^{\infty}(1/4)\exp[-2(x + \Delta)/\phi] \exp[-(x - \Delta)/\phi] \, dx$$

$$= (2\phi)^{-1} \int_{-\infty}^{-\Delta} (\exp[(x - \Delta)/\phi] - \exp[2x/\phi] +(1/4)\exp[(3x + \Delta)/\phi]) \, dx$$

$$+ (8\phi)^{-1} (\int_{-\Delta}^{\Delta} \exp[-(x + 3\Delta)/\phi] \, dx + \int_{\Delta}^{\infty} \exp[-(3x + \Delta)/\phi] \, dx$$

$$=(1/2)(\exp[-2\Delta/\phi] - (1/2)\exp[-2\Delta/\phi] + (1/12)\exp[-2\Delta/\phi])$$

$$+ (1/8)(\exp[-2\Delta/\phi] - \exp[-4\Delta/\phi]) + (1/24)\exp[-4\Delta/\phi]$$

$$= (1/12)(5\exp[-2\Delta/\phi] - \exp[-4\Delta/\phi]) \quad .$$

Similarly, when $\Delta < 0$ one can determine a formula for $\int_{-\infty}^{\infty} F(-x)^2 dF(x)$. However, such a formula is not needed for purposes of the dissertation since Δ is assumed to be nonnegative. Finally, γ is computed from (3.7).

Numerical values of ξ, θ, and γ are given in Table V.

Values of ξ, θ, and γ Under a Uniform Distribution

The density and cumulative distribution functions of a uniform distribution on the interval (α, β) are, respectively, given by

$$f(x) = \begin{cases} 1/(\beta - \alpha) & \alpha < x < \beta, \\ 0 & \text{otherwise} . \end{cases}$$

$$F(x) = \begin{cases} (x - \alpha)/(\beta - \alpha) & \alpha < x < \beta , \\ 0 & x < \alpha , \\ 1 & x > \beta . \end{cases}$$

It can be seen that

Table V. Values of ξ, θ, and γ Under a Double Exponential Distribution with Mean Δ and Variance 1.

Δ	ξ	θ	γ
0.0	0.0	0.0	0.0
0.20	.1356913	.2463620	.0121303
0.25	.1663018	.02978120	.0181546
0.50	.2924869	.5069301	.0555267
0.60	.3306545	.5719550	.0708664
1.0	.4286529	.7569933	.1196560
2.0	.4933127	.9408942	.1614349
3.0	0.4994587	0.9856304	0.1662114
$\Delta \to \infty$	1/2	1	1/6

$$\int_\alpha^\beta F(-x)dF(x) = \begin{cases} 0 & 0 \le \alpha < \beta , \\ \int_\alpha^\beta dx/(\beta - \alpha) & \alpha < \beta \le 0 , \\ \int_\alpha^{-\alpha} (-x - \alpha)dx/(\beta - \alpha)^2 & -\beta < \alpha \le 0 , \\ \int_\alpha^{-\beta} dx/(\beta - \alpha) + \int_{-\beta}^{\beta} (-x - \alpha)dx/(\beta - \alpha)^2 & \alpha < -\beta \le 0 . \end{cases}$$

Thus,

$$\xi = \begin{cases} 1/2 & 0 \le \alpha < \beta , \\ -1/2 & \alpha < \beta \le 0 , \\ 1/2 - 2\alpha^2/(\beta - \alpha)^2 & -\beta < \alpha \le 0 , \\ (3\beta^2 + \alpha^2 + 2\alpha\beta)/2(\beta - \alpha)^2 & \alpha < -\beta \le 0 . \end{cases}$$

The value of θ is given by

$$\theta = 1 - 2F(0) = \begin{cases} (\alpha + \beta)/(\beta - \alpha) & \alpha < 0 < \beta , \\ 1 & 0 < \alpha , \\ -1 & 0 > \beta . \end{cases}$$

Also,it is straightforward to show that

$$\int_\alpha^\beta F(-x)^2 dF(x) = \begin{cases} 0 & 0 \le \alpha < \beta , \\ 1 & \alpha < \beta \le 0 , \\ -8\alpha^3/3(\beta - \alpha)^3 & -\beta < \alpha \le 0 , \\ 2(\beta^3 + 3\alpha^2\beta)/3(\beta - \alpha)^3 & \\ \quad -(\beta + \alpha)/(\beta - \alpha) & \alpha < -\beta \le 0 . \end{cases}$$

The value of γ is obtained from (3.7).

Table VI gives numerical values of ξ, θ, and γ for a uniform distribution with mean Δ = 0.2, 0.25, 0.5, 0.6, 1.0, 2.0, 3.0, and variance 1.

Table VI. Values of ξ, θ, and γ Under a Uniform Distribution with Mean Δ and Variance 1.

α	$\beta=\alpha+\sqrt{12}$	Δ	ξ	θ	γ
-1.73205	1.73205	0.0	0.0	0.0	0.0
-1.53205	1.93205	0.20	.108803	.115470	.006156
-1.48205	1.98205	0.25	.133922	.144340	.009415
-1.23205	2.23205	.50	.247008	.289090	.033648
-1.13205	2.33205	0.60	.286410	.346410	.046143
-.73205	2.73205	1.0	.410684	.577590	.102517
0.26795	3.73205	2.0	.500000	1.000000	.166667
1.26795	4.73205	3.0	0.500000	1.000000	0.166667
		$\Delta\to\infty$	1/2	1	1/6

CHAPTER IV

NONPARAMETRIC CONTROL CHART PROCEDURES BASED ON WITHIN-GROUP RANKING

Two nonparametric control chart procedures based on ranking the observations within groups will be proposed in this Chapter. Several desirable properties justify proposing these two procedures. First, their application does not require the knowledge of the parent distribution or the variance of the observations. Second, they are simple to apply in the sense that little effort is required for computing the ranks of the observations. Third, the exact average run lengths of the procedures are simple to compute. Fourth, they perform satisfactorily when compared to parametric control chart procedures as will be shown in Section 4.4.

The first procedure, called a Cumulative Sum Control Chart-Grouped Signed rank (CSCC - GSR) procedure, is introduced and a method for computing its ARL is given in Section 4.1. In Section 4.2, the practical application of the CSCC - GSR procedure is investigated. The second procedure, called a Linear Barrier - Grouped Signed Rank (LB - GSR) is introduced and investigated in Section 4.3. Section 4.4 contains several comparisons, in terms of the ARL criterion, between the proposed nonparametric procedures and the parametric control chart procedures (i.e., the Shewhart's control charts and the parametric CSCC procedure).

Before introducing the two proposed procedures, the following notations and definitions are needed for the development of this Chapter.

Let $(X_{i1}, X_{i2}, \ldots, X_{ig})$ $i = 1, 2, \ldots$ be groups of independent observations taken sequentially on the output of some process. Let R^*_{ij} denote the rank of $|X_{ij}|$ in the set $\{|X_{i1}|, |X_{i2}|, \ldots, |X_{ig}|\}$ for $i = 1, 2, \ldots,$ and $j = 1, 2, \ldots, g$. For each $i = 1, 2, \ldots,$ define

$$U_{ij} = \text{sign}(X_{ij})R^*_{ij} \;, \; j = 1, 2, \ldots, g \quad .$$

It can be seen that the U_{ij} $(j = 1, 2, \ldots, g)$ are the Wilcoxon signed-ranks, as defined in (3.2), of the observations in the ith group. The U_{ij}'s will be called grouped signed-ranks (GSR's). The statistic that is proposed in this Chapter is

$$SR_n = \sum_{i=1}^{n} \sum_{j=1}^{g} U_{ij} \quad .$$

If $SR_{ig} = \sum_{j=1}^{g} U_{ij}$ then SR_n can be written as $SR_n = \sum_{i=1}^{n} SR_{ig}$ which is easily seen to be a sum of n independent Wilcoxon signed-rank statistics, each based on g observations.

It is well known in nonparametric statistics that the Wilcoxon signed-rank statistic can be used to test the null hypothesis that the distribution of a sequence of independent and identically distributed observations is symmetric about some specified value, Δ_0, against location alternatives (i.e., $\Delta \neq \Delta_0$). It is proposed, in this Chapter, that the statistic SR_n be similarly used to develop nonparametric control chart procedures. If the observations on the output of the process have a symmetric distribution about the control value $\Delta_0 = 0$ (without loss of generality), then the expected value of SR_n will be zero which indicates that the process is in-control. On the other hand, if the distribution of the observations is not symmetric about zero, then the expected value of SR_n will not be zero which indicates that the process is out-of-control.

For the sake of the nonparametric procedures of this dissertation, a process is said to be in-control if the distribution of the observations made on the process is symmetric about the control value. Otherwise, the process is said to be out-of-control.

4.1 A CSCC--GSR Procedure

Let $k \geq 0$ and $h > 0$ be specified constants the choice of which will be discussed in Section 4.2. The proposed one-sided CSCC--GSR procedure for detecting positive deviations from the control value $\Delta_0 = 0$ signals at the first n for which

$$\sum_{i=1}^{n} (SR_{ig} - k) - \min_{0 \leq m \leq n} \sum_{i=1}^{m} (SR_{ig} - k) \geq h \ .$$

As was mentioned in Section 2.1, this procedure is equivalent to a sequence of tests in which $\sum_{i=1}^{n} (SR_{ig} - k)$ is accumulated until either the sum is $\geq h$ or ≤ 0. The procedure signals at the first n for which $\sum_{i=1}^{n} (SR_{ig} - k) \geq h$. If $\sum_{i=1}^{n} (SR_{ig} - k) \leq 0$, then the sum is set equal to zero and sampling is continued by obtaining a new group of g observations and accumulating the resulting signed ranks. The one-sided procedure for detecting negative deviations signals at the first n for which

$$\max_{0\leq m\leq n} \sum_{i=1}^{m} (SR_{ig} + k) - \sum_{i=1}^{n} (SR_{ig} + k) \geq h .$$

Equivalently, this can be described as a procedure that signals to stop at the first n for which $\sum_{i=1}^{n} (SR_{ig} + k) \leq -h$; otherwise sampling is continued and the cumulative sum is set equal to zero whenever it becomes positive. A two-sided symmetric CSCC - GSR procedure can be constructed by employing the above two one-sided procedures simultaneously. Thus, the two-sided procedure signals at the first n for which either one of the one-sided procedures signals; otherwise, sampling is continued. To construct a two-sided asymmetric procedure, it is possible to choose constants $k^+ \geq 0$ and $h^+ > 0$ for the positive one-sided procedure; and $k^- \geq 0$ and $h^- > 0$ for the negative one-sided procedure where $k^+ \neq k^-$ and $h^+ \neq h^-$. One advantage of procedures based on ranking the observations within groups is that their average run lengths can be computed exactly (when k and h are integers) once the distribution of the Wilcoxon signed-rank statistic is determined. The method of computing the exact ARL of a CSCC - GSR procedure will now be developed.

Consider a one-sided CSCC - GSR procedure for detecting positive deviations from the control value $\Delta_0 = 0$. Define

$$S_n = \min\{h, \max\{0, S_{n-1} + SR_{ng} - k\}\} \text{and} \quad S_0 = 0 .$$

It can be seen that the one-sided CSCC - GSR procedure for detecting positive shifts signals at the first n for which $S_n = h$. Otherwise; sampling is continued. The possible values of SR_{ng}, n=1,2,..., are either the even numbers or odd numbers falling between -g(g+1)/2 and

$g(g+1)/2$ inclusive and depending on whether $g(g+1)/2$ is even or odd, respectively. If h and k are restricted to be nonnegative integers, then it can be seen that the sequence $\{S_n; n = 0, 1, 2, \ldots\}$ is a Markov chain with state space $\{0, 1, 2, \ldots, h\}$ and initial start at $S_0 = 0$. This result is immediate since the terms $(SR_{ng} - k)$, $n = 1, 2, \ldots$, are independently distributed. The state "h" is an absorbing state for the chain and it contains all values of the cumulative sum $\sum_{i=1}^{n} (SR_{ig} - k)$ which are equal to or greater than h. The state "0" contains all nonpositive values of the mentioned cumulative sum.

In applications, it may only be meaningful to restrict the chain to start at state "0". Theoretically, however, this restriction is unnecessary and a random start of the chain may be allowed. Therefore, a more general formulation can be obtained by setting $S_0 = j$ $(0 \leq j < h)$ according to some specified initial distribution of the chain. In this general formulation, it can be seen that the Markov chain $\{S_n; n = 0, 1, 2, \ldots\}$ is also homogeneous in time provided that the initial distribution assigns a positive probability to each state (except h) in the state space of the chain. This is true because the terms $(SR_{ng} - k)$, $n = 1, 2, \ldots$, are identically distributed. Any initial distribution may be chosen, for example, the discrete uniform distribution. Let $m' = (m_0, m_1, \ldots, m_{h-1})$, where m_j denotes the mean time to absorption given the chain started initially at state j $(j = 0, 1, 2, \ldots, h-1)$. In this context, the ARL of the CSCC - GSR procedure is simply gm_0 since the procedure always starts with $S_0=0$. Let the $(h+1)\times(h+1)$

matrix $P=||p_{ij}||$ denote the matrix of one-step transition probabilities of the Markov chain and let Q be an h×h submatrix of P denoting the one-step transition probabilities for the nonabsorbing (transient) states 0,1,2,...,h-1. Let I be the h×h identity matrix, and $\underset{\sim}{1}$ be the h×1 vector with all its elements being unity. Then, it is known, see Parzen (1962, page 241) that

$$\underset{\sim}{m} = (I - Q)^{-1} \underset{\sim}{1} .$$

The main effort in computing the ARL of a CSCC - GSR procedure is thus the determination of the matrix Q and inverting (I-Q). The matrix Q can be determined easily if the distribution of the Wilcoxon signed rank statistic based on g observations is available. The distribution of the Wilcoxon signed-rank statistic was discussed in Section 3.1 of this dissertation. If the observations are symmetrically distributed about zero, then the distribution of the Wilcoxon signed rank statistic does not depend on the parent distribution of the observations. Thus, the average run length of the procedure when the process is in control, call it ARL(0), does not depend on the distribution of the observations. However, if the process is out-of-control, (i.e., the observations are not symmetrically distributed about zero) then the average run length depends on the parent distribution of the observations. As a numerical illustration, consider a one-sided positive CSCC - GSR procedure based on groups of size g=4. The possible values of SR_{ig}, for any i, are the even numbers between -10 and 10. For simplicity take k=2, an even number, so that the cumulative $\sum_{i=1}^{n} (SR_{ig}-k)$ takes only even integral values, thus reducing the size of the state space for the Markov chain $\{S_n\}$. The ARL of the procedure will be computed assuming that the process remains in-control; i.e., the observations are symmetrically distributed about zero. In this case, the distribution of the Wilcoxon signed-rank statistic based on four observations is given by

SR_{ig}	-10	-8	-6	-4	-2	0	2	4	6	8	10
Probability	0.0625	.0625	.0625	.125	.125	.125	.125	.125	.0625	.0625	.0625
$SR_{ig} - k$	-12	-10	-8	-6	-4	-2	0	2	4	6	8

If h = 6, then the state space of the chain is {0,2,4,6} and the matrix of one-step transition probabilities is

$$P = \begin{array}{c} \\ 0 \\ 2 \\ 4 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 2 & 4 & 6 \end{array} \\ \begin{pmatrix} 0.6875 & 0.125 & 0.0625 & 0.125 \\ 0.5625 & 0.125 & 0.125 & 0.1875 \\ 0.4375 & 0.125 & 0.125 & 0.3125 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

and Q is the submatrix obtained by deleting the last row and column of P. Upon inverting I-Q, it can be found that

$$(I - Q)^{-1} = \begin{pmatrix} 5.44681 & 0.85106 & 0.51064 \\ 3.97163 & 1.78723 & 0.539007 \\ 3.29078 & 0.68085 & 1.47518 \end{pmatrix}$$

The mean times to absorption are thus given by

$$\underset{\sim}{m} = (1 - Q)^{-1} \cdot \underset{\sim}{1} = \begin{pmatrix} 6.8085 \\ 6.29786 \\ 5.44681 \end{pmatrix} = \begin{pmatrix} m_0 \\ m_2 \\ m_4 \end{pmatrix} .$$

It follows that the average run length ARL(0), when the process is in control, is equal to $m_0 = 6.8085$ in groups of size 5 which corresponds to 27.24 single observations. The values of m_2 and m_4

will be useful in evaluating the ARL when the process is out-of-control and do not serve a particular purpose when the process is in-control.

By the same method one can determine the average run length of a one-sided CSCC--GSR procedure for detecting negative deviations in the mean of the process from the control value $\Delta_0 = 0$. This procedure signals to stop if $\sum_{i=1}^{n} (SR_{ig} + k) \le -h$; otherwise, sampling is continued and the cumulative sum is set equal to zero whenever it becomes positive. Thus, the state space of the resulting Markov chain is $\{0, -1, -2, \ldots, -h\}$, with $-h$ being the absorbing state and 0 containing all nonnegative values of the cumulative sum. Then, the submatrix Q is obtained from the matrix of transition probabilities of the chain and (I-Q) is inverted to obtain the average run length of this one-sided procedure for detecting negative deviations in the mean of the process. Let ARL^+ and ARL^- denote the average run lengths of one-sided CSCC procedures for detecting positive and negative deviations, respectively. Then it can be shown (see van Dobben de Bruyn (1968)) that the ARL of the corresponding two-sided symmetric CSCC procedure is given by

$$ARL = (ARL^+)(ARL^-)/(ARL^+ + ARL^-) \quad .$$

Thus, the ARL of a two-sided symmetric CSCC--GSR procedure is simple to compute once the average run lengths of the corresponding one-sided procedures have been computed.

The above Markov chain approach has provided an exact method for computing average run lengths of CSCC--GSR procedures. The method required that the distribution of the Wilcoxon signed rank statistic be known and the inversion of a certain matrix. The values of k were restricted to be integers in order that the state space of the resulting Markov chain contain integral values. This restriction is not imperative since one may take k = 0.5, for example, and have a Markov chain with state space {0, 0.5, 1, 1.5, ..., h-0.5, h} for the positive one-sided procedure. However, this may result in a Q matrix of a large order, thus making the inversion of (I - Q) more difficult. Therefore, it would be desirable to restrict k to integral values only. In Chapter VI of this dissertation, an analytic asymptotic formula will be developed for procedures based on within group ranking. The purpose of such formulas is to provide a quick, although not exact, way for computing average run lengths. It is worth mentioning that the Markov chain resulting from a CSCC--GSR procedure has a finite number of non-absorbing states. This ensures that average run lengths of the procedure are finite. See Parzen (1962, page 239).

4.2 Application of the CSCC--GSR Procedure

To apply a CSCC--GSR procedure it is necessary to choose values for the parameters k and h. In control chart procedures, the usual criterion for selecting these parameters is to require that the combination k and h minimizes the ARL of the procedure when the process is out-of-control subject to the condition that the ARL be a specified value when the process is in-control. However, the state of being out-of-control is too general to uniquely determine the values of k and h. A common practice is to decide on a value Δ_1 for the deviation from $\Delta_0 = 0$ that is considered significant and is required to be detected quickly. Let $ARL(\Delta_1)$ denote the average run length of a CSCC control chart procedure when the deviation is Δ_1. Then the objective is to choose the combination of k and h that minimizes $ARL(\Delta_1)$ subject to the condition that ARL(0) is at least a preassigned value. ARL(0) serves somewhat the same role as the type I error probability in standard tests of statistical hypotheses. The value of Δ_1 should be considered as a simplification of a range of values for what is considered as a significant deviation in the value of Δ from the control value $\Delta_0 = 0$. A common method for the determination of k will now be outlined.

Johnson (1961) showed that the optimum value of k for the parametric CSCC procedure with normal observations is approximately $|\Delta_1|/2$ regardless of the specified value of ARL(0). Reynolds (1972) showed that if the cumulative sum used in any CSCC procedure can be approximated by a Brownian motion process on the interval $(0, \infty)$, then the optimum value of k is approximately $k = |\mu(\Delta_1)|/2$, regardless of the specified value of ARL(0), where $n\mu(\Delta_1)$ is the expectation of the cumulative sum corresponding to a deviation of Δ_1 from $\Delta_0 = 0$. In a CSCC--GSR procedure the cumulative sum $\sum_{i=1}^{n} (SR_{ig} - k)$ is a sum of independent and identically distributed terms. This cumulative sum can be approximated by a Brownian motion process on the interval $(0, \infty)$, see Section 6.3. Therefore, in a CSCC--GSR procedure the optimum value of k may be taken as $k = |\mu(\Delta_1)|/2$ where $\mu(\Delta_1)$ is the expectation of the Wilcoxon signed-rank statistic based on g observations. The mean of the Wilcoxon signed-rank statistic can be computed easily from results derived in section 3.2 of this dissertation. Table VII gives optimum values of k corresponding to different deviations Δ_1 in the mean of the observations made on the process when the parent distribution is normal (N), double exponential (D), and uniform (U). Each distribution was adjusted to have variance 1 and the values of k were computed using the approximation $k = \mu(\Delta_1)/2$. For cases where the variance of the observations is not 1, the values of Δ_1 and k are in units of standard deviation.

It can be seen from Table VII that the values of k do not differ much for the three distributions under consideration and hence optimum values of k do not depend much on the parent distribution of the observations. Also, it seems possible to classify these values of k into four classes: values optimum for detecting small shifts such as $\Delta_1 = 0.20$, values optimum for detecting medium shifts such as $\Delta_1 = 0.60$, values optimum for detecting large shifts such as $\Delta_1 = 1.00$, and values optimum for detecting very large shifts such as $\Delta_1 \geq 2.00$. For example, for groups of size 6 the values k = 2, k = 6, k = 8, and k = 10 may be taken as optimum values for detecting small, medium, large, and very large shifts in the mean of the observations from the control value. These values are common for the three distributions considered here.

Table VIII gives the values of $ARL^+(0)$, the average run length when the process is in-control, of a positive one-sided CSCC--GSR procedure for various values of k and h. The average run

Table VII. Approximate Optimum Values of k for a CSCC-GSR Procedure for Normal (N), Double Exponential (D), and Uniform (U) Observations.

Δ	Group Size											
	4				6				10			
	N	D	U	Common k	N	D	U	Common k	N	D	U	Common k
0.20	0.99	1.31	0.88	1	2.15	2.77	1.98	2	5.81	7.34	5.47	6
0.60	2.72	3.13	2.41	3	5.88	6.68	5.34	6	15.85	17.74	14.62	16
1.00	3.89	4.09	3.62	4	8.37	8.70	7.89	8	22.37	23.07	21.37	22
2.00	4.89	4.84	5.00	5	10.33	10.22	10.50	10	27.17	26.90	27.50	27
3.00	4.99	4.97	5.00	5	10.49	10.45	10.50	10	27.49	27.40	27.50	27

Table VIII. Values of $ARL^{+}(0)$ in Single Observations of a Positive One-Sided CSCC-GSR Procedure

Group Size	k	h = 2	4	6	8	10	12	14	16	18
g=2	1	8.00	32.00	96.00	256.00	644.00	1600.0			
g=4	0	9.14	12.19	17.07	25.60	35.16				
	2	12.80	19.69	27.24	40.70					
	4	21.32	30.69	53.17						
	6	32.00	63.96							
	8	64.00								
g=6	1	14.22	17.15	20.32	24.49	30.17	36.52	54.41	67.28	84.55
	3	17.45	21.04	26.00	33.28	42.02	51.80	65.92	80.57	100.66
	5	20.14	26.94	35.71	46.93	60.11	81.29	104.33	140.63	
	7	27.43	37.46	51.01	67.54	97.04	130.76	192.07		
	9	38.40	53.76	72.98	110.87	155.36	249.32			
	11	54.86	75.79	120.17	173.05	301.01				
	13	76.80	125.39	183.14	334.79					
	15	127.98	190.49	366.46						
	17	192.00	378.00							
	19	384.00								

Table VIII (Cntd). Values of $ARL^{+}(0)$ in Single Observations for a Positive One-Sided

CSCC-GSR Procedure Group Size g = 10.

k	h												
	2	4	6	8	10	12	14	16	18	20	22	24	26
5	26.0	28.7	31.6	35.0	38.8	43.1	47.9	53.3	59.3	65.9	73.4	81.5	90.3
7	28.8	31.9	35.5	39.6	44.2	49.5	55.5	62.3	69.9	78.6	88.2	99.0	111.0
13	40.6	46.1	52.6	60.3	69.2	79.7	92.4	106.9	124.1	144.8	168.5	198.5	233.6
21	72.6	85.7	102.2	121.8	146.4	177.3	216.1	269.4	337.3	420.1	530.5	677.9	870.0
23	86.1	103.0	123.5	149.3	182.1	223.6	281.8	357.1	450.1	577.0	750.3	983.7	1430.0
27	124.9	152.4	187.8	233.4	289.9	385.8	495.6	657.5	872.7	1183.3	1578.0	2330.0	3218.0

k	h											
	28	30	32	34	36	38	40	42	44	46	48	50
5	100.1	110.6	122.5	135.2	148.6	163.1	178.6	195.2	212.5	231.7	251.5	273.0
7	124.2	139.6	156.0	174.1	193.9	215.6	239.2	264.2	292.6	322.1	355.3	
13	273.0	320.7	377.2	442.7	515.8	609.3	709.0	833.4				
21	1102.0	1475.0	1850.0	2560.0								
23	1753.0	2314.0	3262.0									
27	50001.0											

lengths, $ARL^-(0)$, of a negative one-sided procedure are the same as those values of $ARL^+(0)$. Average run lengths, $ARL(0)$, of a two-sided symmetric procedure can be obtained through the relation

$$\begin{aligned} ARL(0) &= (ARL^+(0)\ ARL^-(0))/(ARL^+(0) + ARL^-(0)) \\ &= ARL^+(0)/2\ . \end{aligned}$$

The Markov chain approach was employed to obtain exact values of $ARL^+(0)$. Values of k were taken so that the state space of the resulting Markov chain is the set of even numbers $\{0, 2, 4, \ldots, h\}$ rather than being the set of all integers between 0 and k inclusive. This will halve the size of the matrices of transition probabilities. In order to allow for the possibility of stopping after sampling only one group of observations, the values of h were taken, except for groups of size 2, to satisfy $h \leq g(g + 1)/2 - k$. The values of $ARL^+(0)$ contained in Tables VIIIa and b are the same regardless of the parent distribution of the observations. This is an advantage of using a nonparametric procedure over using a parametric procedure where the average length depends on the parent distribution of the observations even when the process is in-control.

4.3 A Linear Barrier-Grouped Signed Rank (LB-GSR) Procedure

The proposed two-sided LB-GSR control chart procedure signals to stop sampling at the first n for which

$$SR_n = \sum_{i=1}^{n} SR_{ig} \notin (-a, a)$$

where a > 0 is a preassigned integer chosen to guarantee a specified value for ARL(0) of the procedure. Otherwise, sampling is continued. This procedure represents a Markov chain $\{SR_n; n = 0, 1, 2, \ldots\}$ with state space $\{-a, \ldots, -1, 0, 1, \ldots, a\}$ where SR_0 is defined to be zero. The states -a and a are absorbing states. Thus, the average run length of this procedure can be computed exactly using the Markov chain approach as developed in the last section. One reason for proposing the LB-GSR procedure is that it requires the determination of only one parameter, a, which makes it a little simpler to apply than a CSCC--GSR procedure. Also, it will be seen in the next section that the LB--GSR performs well for small shifts in the mean of the process. In Section 6.3 an approximate analytic formula based on a Brownian motion approximation will be developed for the ARL of a LB--GSR procedure.

One-sided LB--GSR procedures can also be constructed. A one-sided procedure for detecting positive shifts from the control value Δ_0 signals at the first n for which $SR_n \geq a$; otherwise, sampling is continued. A one-sided procedure for detecting negative shifts signals at the first n for which $SR_n \leq -a$. The difficulty with these one-sided procedures is that the state space of the resulting Markov chain contains an infinite number of nonabsorbing states. For example, the state space of a one-sided procedure for detecting positive shifts is $\{\ldots, -1, 0, 1, 2, \ldots, a\}$ with a being the absorbing state. This makes the computation of the ARL of these procedures by the method of last section impossible since the matrix Q will be of infinite order. Hence, there is no simple way to compute the exact average run lengths of the one-sided procedures. One-sided LB procedures are not efficient since SR_n may be far below zero when a shift occurs. In this case, it would take a long time for the procedure to signal.

4.4 Comparisons of Procedures

In this section various control chart procedures will be compared in terms of their average run lengths. In comparing the procedures, two cases will be considered. The first case is where more than one observation is taken at each of the time points. In this case, the sample means will be used for both the parametric CSCC and the Shewhart procedures. The Wilcoxon signed-rank statistic based on ranking the observations within groups will be used in CSCC--GSR and LB--GSR procedures. The second case is where only one observation is taken at each time point. In this case, the individual observations will be used for both the parametric CSCC and Shewhart procedures. Procedures based on grouped signed ranks would require that the observations be artifically grouped.

The usual criterion for comparing two control chart procedures is to adjust them so that they have the same average run length when the process is in-control and then compare their average run lengths at several points representing out-of-control states for the process. As was stated in the Section 1.1, procedure 1 is considered to be more efficient than procedure 2 at a shift Δ from

from the control value $\Delta_0=0$ if $ARL_1(\Delta)<ARL_2(\Delta)$ subject to the condition that $ARL_1(0) \geq ARL_2(0)$.

Tables IXa, b, c, d, e, f, g, h, and i display the ARL's of three one-sided control chart procedures for detecting positive shifts in in Δ from $\Delta_0 = 0$. These procedures are: the CSCC--GSR procedure based on groups of sizes six and ten, the parametric CSCC procedure based on single observations and on grouped observations, and the Shewhart procedure based on single and on grouped observations. The procedures will now be compared assuming that the observations have a normal distribution with mean Δ and variance 1. The Markov chain approach developed in Section 4.2 was employed to compute exact average run lengths of the CSCC--GSR procedure. Average run lengths of the parametric CSCC procedure were obtained by graphical interpolations in the van Dobben de Bruyn (1968) tables. Average run lengths of Shewhart procedure were computed using formula (2.1) in Chapter II. The parametric CSCC procedure has almost the same average run length for ungrouped and grouped observations except for large shift ($\Delta \geq 2.0$). The tables reveal that when all three procedures use grouped observations, the ARL of the CSCC--GSR procedure is only slightly larger than the ARL of the parametric CSCC and there is not much difference in the ARL's of the CSCC--GSR and Shewhart procedures for $\Delta \geq 0.6$. For large values of Δ, all three procedures take at least one group and the ARL is approximately equal to the group size g. When the CSCC--GSR procedure is compared to the parametric CSCC procedure and to Shewhart procedure both using single observations, it can be seen that the CSCC--GSR is roughly as efficient as the parametric CSCC procedure and more efficient than the Shewhart procedure for small and medium shifts. The parametric CSCC and Shewhart procedures using single observations are more efficient than the CSCC--GSR procedure when large shifts in Δ are considered. This last conclusion is expected since the CSCC--GSR procedure always takes at least one group to signal to stop. Thus, even under the normality assumption made on the observations, the CSCC--GSR procedure is competitive with the parametric procedures when observations are taken in groups. When the parametric procedures are used with single observations, the CSCC--GSR is competitive with them if one is interested primarily in detecting small shifts in Δ.

Tables Xa, b, d, c, and e display the ARL's of the two-sided LB--GSR procedure, the CSCC--GSR, the parametric CSCC, and the Shewhart procedure. Groups of size six and ten were considered and the observations are assumed to have a normal distribution with variance 1. Again, the Markov chain approach was employed to compute exact average run lengths for the LB--GSR. These tables indicate that the LB--GSR is very efficient for small deviations ($\Delta \leq 0.6$) in the mean of the process; but it is poor for large deviations. Actually, LB--GSR seems to be slightly more efficient than the CSCC--GSR procedure for shifts as small as $\Delta = 0.20$. However, it should only be concluded that the LB--GSR is as efficient as the CSCC--GSR for small shifts since the k value used in the CSCC--GSR is only approximately optimum.

Table IXa. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Optimal for Δ = 0.20 Group Size g=6 from $N(\Delta,1)$ Distribution

Δ	CSCC-GSR	Parametric CSCC		Shewhart	
		Ungrouped	Grouped	Ungrouped	Grouped
	k=3 h=18	k=0.11 h=6	k=.23 h=2	$X \geq 2.327$	$\sqrt{g}\ \bar{X} \geq 1.555$
0.0	101.0	100.0	101.0	100.0	100.0
0.20	39.3	35.5	36.6	59.8	41.8
0.60	15.3	13.0	13.8	23.8	12.9
1.0	10.4	7.6	8.8	10.8	7.4
2.0	6.8	3.8	6.1	2.7	6.0
3.0	6.0	2.6	6.0	1.3	6.0

Table IXb. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Optimal for Δ = 0.60 Group Size g = 6 from $N(\Delta,1)$ Distribution

Δ	CSCC-GSR	Parametric CSCC		Shewhart	
		Ungrouped	Grouped	Ungrouped	Grouped
	k=5 h=16	k=0.24 h=5.0	k=0.55 h=1.5	$X \geq 2.450$	$\sqrt{g}\ \bar{X} \geq 1.718$
0.0	140.6	140.0	141.0	140.0	140.0
0.20	50.3	42.1	46.8	81.8	54.7
0.60	16.6	13.5	14.4	31.1	14.9
1.0	10.6	7.4	8.4	13.6	7.8
2.0	6.8	3.5	6.0	3.1	6.0
3.0	6.0	2.4	6.0	1.4	6.0

Table IXc. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Optimal for Δ=1.0. Group Size g=6 from N(Δ,1) Distribution

Δ	CSCC-GSR	Parametric CSCC	Shewhart	
	k=9 h=12	Ungrouped k=0.45 h=4.0	Ungrouped $X \geq 2.652$	Grouped $\sqrt{g}\ \bar{X} \geq 1.977$
0.0	249.3	250.0	250.0	250.0
0.20	85.6	78.0	140.8	87.6
0.60	21.2	17.0	49.8	19.6
1.0	11.4	7.8	20.3	8.8
2.0	6.8	3.3	3.9	6.0
3.0	6.0	2.2	1.6	6.0

Table IXd. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Optimal for Δ=2.0. Group Size g=6 from N(Δ,1) Distribution

Δ	CSCC-GSR	Parametric CSCC	Shewhart	
	k=11 h=10	Ungrouped k=1.04 h=2.0	Ungrouped $X \geq 2.714$	Grouped $\sqrt{g}\ \bar{X} \geq 2.054$
0.0	301.0	300.0	300.8	300.0
0.20	107.9	130.0	167.5	101.9
0.60	24.7	32.0	57.9	21.5
1.0	12.0	10.5	23.1	9.2
2.0	6.8	2.8	4.2	6.0
3.0	6.0	1.6	1.6	6.0

Table IXe. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Optimal for $\Delta=3.0$. Group Size $g=6$ from $N(\Delta,1)$ Distribution

Δ	CSCC-GSR	Parametric CSCC	Shewhart	
		Ungrouped	Ungrouped	Grouped
	$k=11$ $h=10$	$k=1.49$ $h=1.3$	$X \geq 2.714$	$\sqrt{g}\ \bar{X} \geq 2.054$
0.0	301.0	300.0	300.8	300.0
0.20	107.9	155.0	167.5	101.9
0.60	24.7	44.5	57.9	21.5
1.0	12.0	16.0	23.1	9.2
2.0	6.8	3.1	4.2	6.0
3.0	6.0	1.5	1.1	6.0

Table IXf. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Optimal for Δ=0.20. Group Size g=10 from $N(\Delta,1)$ Distribution

Δ	CSCC-GSR	Parametric CSCC		Shewhart	
	k=5 h=50	Ungrouped k=0.08 h=10.0	Grouped k=.25 h=2.5	Ungrouped X>2.678	Grouped $\sqrt{g}\ \bar{X}$>1.787
0.0	272.5	270.0	270.0	270.0	270.5
0.20	67.6	58.0	62.0	151.4	80.7
0.60	24.8	19.8	21.7	53.0	18.4
1.0	18.4	11.8	13.6	21.4	10.9
2.0	12.0	5.8	10.0	4.0	10.0
3.0	10.1	3.9	10.0	1.6	10.0

Table IXg. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Optimal for Δ=0.60. Group Size g=10 from N(Δ,1) Distribution

Δ	CSCC-GSR k=15 h=40	Parametric CSCC	Shewhart	
		Ungrouped k=.31 h=8	Ungrouped $X \geq 3.122$	Grouped $\sqrt{g}\ \bar{X} \geq 2.366$
0.0	1113.5	1115.0	1113.5	1112.2
0.20	154.1	188.0	575.0	241.2
0.60	30.1	25.0	171.4	31.3
1.0	18.9	12.0	59.1	12.7
2.0	12.0	5.2	7.6	10.0
3.0	10.1	3.5	2.2	10.0

Table IXh. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Optimal for Δ=1.0. Group Size g=10 from N(Δ,1) Distribution

Δ	CSCC-GSR k=23 h=32	Parametric CSCC Ungrouped k=0.52 h=6	Shewhart Ungrouped X>3.425	Shewhart Grouped $\sqrt{g}\ \bar{X}$>2.740
0.0	3262.3	3260.0	3253.1	3255.1
0.20	381.2	450.0	1593.1	570.1
0.60	39.4	37.0	424.3	50.1
1.0	20.0	13.0	130.65	15.07
2.0	12.0	4.8	13.0	10.0
3.0	10.1	3.0	3.0	10.0

Table IXi. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Optimal for Δ=2.0. Groups Size g=10 from $N(\Delta,1)$ Distribution

Δ	CSCC-GSR	Parametric CSCC	Shewhart	
	k=27 h=28	Ungrouped k=1.1 h=3	Ungrouped $X \geq 3.540$	Grouped $\sqrt{g}\ \bar{X}>2.878$
0.0	5000.6	5000.0	4997.5	4997.5
0.20	607.3	1000.0	2387.2	809.6
0.60	48.1	118.0	609.3	61.2
1.0	20.9	22.0	180.4	16.4
2.0	12.0	4.1	16.2	10.0
3.0	10.1	2.2	3.4	10.0

Table Xa. Values of ARL(Δ) in Single Observations for Two-Sided Procedures Optimal for Δ=0.20. Group Size g=6 from N(Δ,1) Distribution

Δ	LB-GSR	CSCC-GSR	Parametric CSCC		Shewhart	
			Ungrouped	Grouped	Ungrouped	Grouped
	a = 21	k=3 h=18	k=0.11 h=6	k=.25 h=2	$\lvert X\rvert \geq 2.298$	$\sqrt{g}\ \bar{X} \geq 1.516$
0.0	46.5	50.3	50.0	51.5	46.4	46.4
0.20	31.5	35.1	32.8	35.3	41.3	34.3
0.60	14.3	15.3	13.0	14.4	21.4	12.4
1.0	10.2	10.4	7.6	8.8	10.2	7.3
2.0	6.8	6.8	3.8	6.1	2.6	6.0
3.0	6.0	6.0	2.6	6.0	1.3	6.0

Table Xb. Values of ARL(Δ) in Single Observations for Two-Sided Procedures Optimal for Δ=0.20. Group Size g=10 from N(Δ,1) Distribution

Δ	LB-GSR	CSCC-GSR	Parametric CSCC		Shewhart	
			Ungrouped	Grouped	Ungrouped	Grouped
	a = 55	k=5 h=46	k=.12 h=8	k=0.34 h=2	$\|X\| \geq 2.624$	$\sqrt{g}\|\bar{X}\| \geq 1.712$
0.0	114.2	115.9	115.0	115.0	115.0	115.1
0.20	56.7	60.2	51.4	55.4	99.535	66.8
0.60	23.6	23.2	17.0	18.5	45.2	17.4
1.0	18.3	16.4	10.0	12.2	19.1	10.8
2.0	12.0	10.4	4.9	10.0	3.8	10.0
3.0	10.1	10.0	3.3	10.0	1.6	10.0

Table Xc. Values of ARL(Δ) in Single Observations for Two-Sided Procedures Optimal for Δ=0.60. Group Size g=10 from N(Δ,1) Distribution

Δ	LB-GSR	CSCC-GSR	Parametric CSCC
	a = 55	k=15 h=22	Ungrouped k=0.31 h=5
0.0	114.2	109.2	115.0
0.20	56.7	61.2	58.3
0.60	23.6	18.8	16.0
1.0	18.3	11.8	8.2
2.0	12.0	10.0	3.6
3.0	10.1	10.0	2.5

Table Xd. Values of ARL(Δ) in Single Observations for Two-Sided Procedures Optimal for Δ=1.0. Group Size g=10 from N(Δ,1) Distribution

Δ	LB-GSR	CSCC-GSR	Parametric CSCC
	a = 55	k=23 h=12	Ungrouped k=0.45 h=4
0.0	114.2	111.8	115.0
0.20	56.7	67.3	68.3
0.60	23.6	19.1	16.0
1.00	18.3	11.5	7.9
2.00	12.0	10.0	3.3
3.00	10.0	10.0	2.2

Table Xe. Values of ARL(Δ) in Single Observations for Two-Sided Procedures Optimal for Δ=2.0. Group Size g=10 from N(Δ,1) Distribution

Δ	LB-GSR a = 55	CSCC-GSR k=27 h=8	Parametric CSCC Ungrouped k=0.98 h=2
0.0	114.2	116.7	115.0
0.20	56.7	71.2	85.3
0.60	23.6	19.8	26.0
1.0	18.3	11.5	9.8
2.0	12.0	10.0	2.7
3.0	10.1	10.0	1.6

All previous comparisons between the various control chart procedures were made under the assumption that the observations are normally distributed. For nonnormal observations, it is more difficult to make the desired comparisons because exact values of the average run lengths of the parametric CSCC and the CSCC--GSR are not available. The computation of the ARL of the parametric CSCC requires solving the integral equation (2.5) as mentioned in Chapter II. When the process is in-control, the ARL of a procedure based on GSRs is the same regardless of the type of distribution the observations have. When the process is out-of-control, the computation of the ARL of a procedure based on GSRs requires the knowledge of the distribution of the Wilcoxon signed-rank statistic under that particular distribution of the observations. There are no published results for the distribution of the Wilcoxon signed-rank statistic for distributions other than the normal.

In this dissertation, a limited simulation study is made to compare the various control chart procedures when the observations have a double exponential distribution. A reason for making the comparison under a double exponential distribution will now be stated. It is known that in the context of standard hypothesis testing against double exponential shift alternatives, the Wilcoxon signed-rank statistic is asymptotically more efficient (see Lehmann, 1975) than the parametric (i.e., the normal theory) tests. It is, therefore, anticipated that for quality control applications, the Wilcoxon signed-rank statistic will also be more efficient than the parametric CSCC procedure when the observations have a double exponential distribution. Double exponential random variables with variance 1 were generated on the computer from uniform random variables by using the probability integral transformations. The values Δ = 0.0, 0.2, 0.6, 1.0, 2.0, 3.0 were chosen to represent a range of small to large values for the mean of the double exponential distribution.

Tables XIa, b, c, d display the ARLs of the CSCC--GSR procedure the parametric CSCC, and the Shewhart procedures based on grouped and ungrouped observations. Values of the ARL of the CSCC--GSR and of the ungrouped parametric CSCC were obtained by simulation based on 300 runs. Values for the ungrouped Shewhart procedure are exact and were computed as indicated in Section 2.1. When the observations are grouped, the parametric CSCC and the Shewhart procedures are based on the means of the groups rather than on the individual observations. Thus, for groups of size six and ten, the distribution of the group means is approximately normal. Therefore, for grouped observations the ARL of the parametric CSCC and the Shewhart procedures were computed assuming a normal underlying distribution.

From the comparisons contained in Tables XIa, b, c, d it is possible to make the following conclusions. For small shifts ($\Delta \leq 0.2$), the CSCC--GSR is more efficient than both the parametric CSCC and the Shewhart procedures whether grouped or ungrouped. For medium shifts ($\Delta = 0.6$), the CSCC--GSR is almost as efficient as the ungrouped and grouped parametric CSCC, and the CSCC--GSR is better than Shewhart grouped or ungrouped (except when ARL_0 is small = 100). For large

Table XIa. Values of $ARL^{+}(\Delta)$ in Single Observations for Procedures Optimal for Δ=0.2. Group Size g=6 From Double Exponential Distribution.

Δ	CSCC--GSR	Parametric CSCC		Shewhart	
	k=3,h=18	Ungrouped k=0.11,h=6.0	Grouped k=.23,h=2	Ungrouped $X \geq 2.77$	Grouped $\sqrt{g}\ \bar{X} \geq 1.555$
0.0	101.0	101.5	101.0	100.5	100.0
0.2	31.0	34.6	36.6	75.8	41.8
0.6	13.6	13.5	13.8	43.0	12.9
1.0	9.8	7.5	8.82	24.4	7.4
2.0	7.1	3.8	6.1	5.9	6.0
3.0	6.0	2.7	6.0	1.6	6.0

Table XIb. Values of $ARL^{+}(\Delta)$ in Single Observations for Procedures Optimal for $\Delta=0.6$. Group Size g=6 From Double Exponential Distribution.

Δ	CSCC--GSR	Parametric CSCC		Shewhart	
	k=5,h=16	Ungrouped k=.24,h=5.0	Grouped k=.55,h=1.5	Ungrouped X > 3.01	Grouped $\sqrt{g}\ \bar{X}\geq 1.718$
0.0	140.6	137.0	141.0	141.2	140.0
0.2	39.2	51.8	46.8	106.4	54.68
0.6	14.2	13.7	14.4	60.4	14.92
1.0	9.7	7.0	8.4	34.3	7.8
2.0	6.9	3.5	6.1	8.3	6.0
3.0	6.0	2.3	6.0	2.03	6.0

Table XIc. Values of $ARL^{+}(\Delta)$ in Single Observations

for Procedures Optimal for Δ=0.2.

Group Size g=10 From Double Exponential Distribution.

Δ	CSCC--GSR	Parametric CSCC		Shewhart	
	k=5,h=46	Ungrouped k=.10,h=9.0	Grouped k=.34,h=2	Ungrouped $X \geq 3.36$	Grouped $\sqrt{g}\ \bar{X} \geq 1.715$
0.0	231.7	231.2	230.0	231.6	231.6
0.20	50.4	61.1	59.0	174.5	71.7
0.60	20.9	18.9	19.1	99.1	17.5
1.0	15.3	10.4	12.5	56.3	10.8
2.0	10.8	5.3	10.1	13.7	10.0
3.0	10.0	3.6	10.0	3.3	10.0

Table XId. Values of $ARL^{+}(\Delta)$ in Single Observations for Procedures Optimal for Δ=0.6. Group Size g=10 From Double Exponential Distribution.

Δ	CSCC--GSR	Parametric CSCC		Shewhart	
	k=15,h=40	Ungrouped k=.3,h=8.0	Grouped k=1.07,h=1.5	Ungrouped $X \geq 4.47$	Grouped $\sqrt{g}\ \bar{X} \geq 2.366$
0.0	1113.5	1119.6	1114.0	1113.0	1112.2
0.20	105.3	193.1	175.0	838.7	241.2
0.60	26.1	25.4	26.0	476.4	31.3
1.0	18.1	12.1	13.3	270.6	12.7
2.0	12.6	5.4	10.1	65.8	10.0
3.0	10.0	3.5	10.0	16.0	10.0

shifts ($\Delta \geq 1.0$), the CSCC--GSR is almost as efficient as the grouped parametric CSCC and the grouped Shewhart procedures. When compared with ungrouped parametric CSCC and ungrouped Shewhart, the CSCC--GSR is less efficient because it takes at least one group to signal no matter how large Δ is.

4.5 ARL For Shifts Occurring After the Procedure Has Started

In the discussion on the average run length for an out-of-control state, it was assumed that the process goes out of control at the time the control procedure is started. This assumption is usually made in the literature to make the determination of the ARL easier. In reality, however, the time at which the process goes out of control is unknown and will usually fall after the testing procedure has been running for some time. In this case, the time it takes to signal depends on the value of the test statistic at the time of the shift. For a CSCC type procedure the ARL can never be larger than its value when the test statistic is zero.

Suppose that for the first n_0 observations the process is in-control with $\Delta = 0$, and for observations taken after n_0, the value of Δ shifts to $\Delta > 0$. The computation of the ARL, counted from n_0, of a positive CSCC--GSR procedure will now be illustrated. At time n_0, the test statistic S_{n_0} may assume any integral value between 0 and h-1. Let τ be the time it takes the procedure to signal when a shift of $\Delta > 0$ occurs at time n_0. Then it can be seen that the average run length of the procedure is given by

$$E[\tau|S_{n_0} < h] = \sum_{j=0}^{h-1} E_\Delta[\tau|S_{n_0} = j]P(S_{n_0} = j)/P(S_{n_0} < h)$$

$$= \sum_{j=0}^{h-1} m_j \, p_{0j}(n_0) / \sum_{j=0}^{h-1} p_{0j}(n_0),$$

where $p_{0j}(n_0)$ is the n_0-step transition probability from state 0 to state j. The quantities $p_{0j}(n_0)$ are the elements of the first row of the matrix $(Q)^{n_0}$.

As a numerical illustration, consider a positive-sided CSCC--GSR procedure with parameters k=9 and h=12, and based on groups of size six. Suppose that after n_0=5 groups, a shift of Δ = 0.20 occurs in the mean of the observations which are assumed to be normally distributed with variance 1. The matrix Q (when Δ = 0) is given by

$$Q = \begin{array}{c} 0 \\ 2 \\ 4 \\ 6 \\ 8 \\ 10 \end{array} \begin{pmatrix} .84375 & .04688 & .03125 & .03125 & .01563 & .01563 \\ .78125 & .06250 & .04688 & .03125 & .03125 & .01563 \\ .71875 & .06250 & .06250 & .04688 & .03125 & .03125 \\ .65620 & .06250 & .06250 & .06250 & .04688 & .03125 \\ .57813 & .07813 & .06250 & .06250 & .06250 & .04688 \\ .50000 & .07813 & .07813 & .06250 & .06250 & .06250 \end{pmatrix}$$

The first row of Q^5 is

(.74861, .04580, .03279, .03128, .01825, .01672) .

The vector of mean absorption times (when Δ = 0.2) is

$(m_0, m_2, m_4, m_6, m_{10})$ = (14.3, 13.7, 13.1, 12.2, 11.3, 10.3)

in groups of size 6. Hence, the average run length when a shift of Δ = 0.20 occurs after five groups is

$$E(\tau|S_5 < 12) = \frac{12.5}{0.89342} = 14.0$$

in groups of size 6. When $n_0 = 10$, the first row of Q^{10} is

$$(.66193,\ .04050,\ .02899,\ .02764,\ .01614,\ .01479)\quad .$$

Hence, the average run length when a shift of $\Delta = 0.20$ occurs after ten groups is

$$E_{\Delta=0.20}\ (\tau \mid S_{10} < 12) = \frac{11.1}{0.78998} = 14.1$$

in groups of size 6. The above two average run lengths of 14.0 and 14.1 do not differ appreciably from the value 14.3 of the average run length computed when a shift of $\Delta = 0.20$ occurs right at the time the procedure is started.

When considering a shift of $\Delta = 1.0$ occurring at times $n_0 = 5$ and $n_0 = 10$, the vector of mean absorption times (expressed in groups of size 6) is

$$(m_0,\ m_2,\ m_4,\ m_6,\ m_8,\ m_{10}) = (1.9,\ 1.6,\ 1.4,\ 1.3,\ 1.2,\ 1.1)\quad .$$

Hence, the average run lengths (in groups of size 6) are

$$E_{\Delta=1.0}\ (\tau \mid S_5 < 12) = \frac{1.6}{0.89342} = 1.79\ ,$$

and

$$E_{\Delta=1.0}\ (\tau \mid S_{10} < 12) = \frac{1.4}{0.78998} = 1.77\ .$$

Again, it is seen that there is no appreciable change in the value of the average run length when shift occurs at the time the procedure has started, or at a later time. This conclusion is intuitive because if no shift occurs in the mean of the observations then the value of $\sum_{i=1}^{n} (SR_{ig} - k)$ is zero with high probability no matter how long the procedure has been running. Thus, the common practice of computing average run lengths assuming the process goes out-of-control at the time the procedure is started seems to be justified for the CSCC--GSR procedure.

CHAPTER V

A PROCEDURE BASED ON RANKING WITH RESPECT TO LAST M OBSERVATIONS

In this chapter, two more nonparametric control chart procedures are developed. The procedures are intended to serve cases where the observations cannot be taken in groups but it is still desirable to have a simple nonparametric procedure with little effort for ranking the observations.

5.1 Signed Sequential Ranks With Respect to Last M Observations

Let X_1, X_2, ... be a sequence of independent and identically distributed observations having a continuous cumulative distribution function $F(x)$ which is symmetric about Δ. Let $M \geq 1$ be a fixed integer. Define

$$R^+_{jM} = \text{rank of } |X_j| \text{ in the set } \{|X_{j-M+1}|,\ |X_{j-M+2}|,\ \ldots,\ |X_j|\} \text{ for } j \geq M ,$$

and

$$R^+_{jM} = \text{rank of } |X_j| \text{ in the set } \{|X_1|, |X_2|, \ldots, |X_j|\} \quad \text{for } j < M .$$

In other words, R^+_{jM} is the rank of $|X_j|$ in the set containing $|X_j|$ and the (M - 1) preceding X's in their absolute values. For example, suppose that M = 3. Then,

$$R_{1M}^{+} = \text{rank of } |X_1| \text{ in the set } \{|X_1|\} ,$$

$$R_{2M}^{+} = \text{rank of } |X_2| \text{ in the set } \{|X_1|, |X_2|\} ,$$

$$R_{3M}^{+} = \text{rank of } |X_3| \text{ in the set } \{|X_1|, |X_2|, |X_3|\} ,$$

$$R_{4M}^{+} = \text{rank of } |X_4| \text{ in the set } \{|X_2|, |X_3|, |X_4|\} ,$$

and

$$R_{5M}^{+} = \text{rank of } |X_5| \text{ in the set } \{|X_3|, |X_4|, |X_5|\} ,$$

and so on. This method of ranking the observations is a modification of the sequential ranking method of Parent, as described in Section 2.2, where an observation is ranked with respect to all observations preceding it in the sequence. It is simpler to rank an observation with respect to only the last (M - 1) observations rather than ranking it with respect to all preceding observations.

The signed sequential rank of X_j with respect to last M observations (SSRM) is defined by

$$T_j = \text{sign}(X_j)\ R_{jM}^{+} , \quad j = 1, 2, \ldots \quad . \tag{5.1}$$

An independence property of the T_j's should be noted. It can be seen that T_j and $T_{j'}$ are stochastically independent whenever $j' - j > M - 1$; i.e., whenever the distance between the pair is greater than M - 1. This is true since the sets of X's from which T_j and $T_{j'}$ are computed do not overlap. This independence property can be stated more generally by saying that sets $(T_i, T_{i+1}, \ldots, T_k)$

and $(T_{k+r}, T_{k+r+1}, \ldots, T_j)$ are stochastically independent whenever $r > M - 1$. This property will be used in Section 6.4 to derive approximate ARLs for the procedures to be developed in this Chapter. However, for $j' - j \leq M - 1$, T_j and $T_{j'}$ will in general be stochastically dependent even when the distribution function $F(x)$ of the observations is symmetric about zero. Parent (1965) established that the signed sequential ranks are independent if $F(x)$ satisfies the condition $F(-x) = F(0)(1 - F(x) + F(-x))$. This condition is satisfied if $F(x)$ is symmetric about zero. A counter example will now be given to show that this result is not valid for the SSRM. Suppose that $X_1, X_2, \ldots$ are independent with a common cumulative distribution function $F(x)$ that is symmetric about zero. Assuming that $M = 3$, it will be shown that

$$\Pr(T_4 = -2, T_5 = 3) \neq \Pr(T_4 = -2) \Pr(T_5 = 3) .$$

Lemma 2.2 of Parent (1965) established that sign(x) and $|X|$ are independent if and only if $F(-x) = F(0)(1 - F(x) + F(-x))$. It can be seen that if sign (x_j) and $|X_j|$ are independent then sign (x_j) and R^+_{jM} are independent since R^+_{jM} is a function of $|X_{j-M+1}|, |X_{j-M+2}|, \ldots, |X_j|$ only. Thus,

$$\begin{aligned}\Pr(T_4=-2)\Pr(T_5=3) &= \Pr(\mathrm{sign}(x_4)=-1)\Pr(R^+_{43}=2)\Pr(\mathrm{sign}(x_5)=1)\Pr(R^+_{53}=3) \\ &= F(0) \cdot 1/3 \cdot (1 - F(0)) \cdot 1/3 \\ &= 1/2 \cdot 1/3 \cdot 1/2 \cdot 1/3 = 1/36 .\end{aligned}$$

On the other hand,

$$\Pr(T_4=-2,\ T_5=3) = \Pr(\text{sign}(x_4)=-1,\ R^+_{43}=2,\ \text{sign}(x_5)=1,\ R^+_{53}=3)$$

$$= \Pr(\text{sign}(x_4)=-1)\ \Pr(\text{sign}(x_5=1))\Pr(R^+_{43}=2,\ R^+_{53}=3)$$

$$= (1/4)\Pr(R^+_{43}=2,\ R^+_{53}=3)$$

$$=(1/4)[\Pr(|x_2|<|x_4|<|x_3|<|x_5|)+\Pr(|x_3|<|x_4|<|x_2|<|x_5|)$$

$$+ \Pr(|x_3|<|x_4|<|x_5|<|x_2|)]$$

$$= 1/4 \cdot 3/4! = 1/32 \quad ,$$

since each ordering of the x's is equally likely. Hence T_4 and T_5 are stochastically dependent even when F(x) is symmetric about zero.

5.2 Control Chart Procedures Based on the SSRM's

The statistic to be used in this Chapter for constructing control chart procedures is defined by

$$TR_n = \sum_{j=1}^{n} \text{sign}(x_j)R^+_{jM} = \sum_{j=1}^{n} T_j \ . \tag{5.2}$$

It will be established in Section 6.4 that the expected value of TR_n is zero if the observations are symmetrically distributed about zero (i.e., the process is in-control). Thus, values of TR_n far away from zero may be taken as an indication that the observations are not symmetrically distributed about zero (i.e., the process is out-of-control). Accordingly, two control chart procedures will be constructed. First, a linear barrier procedure (LB--SSRM) is constructed that signals at the first n for which $TR_n \notin (-a, a)$ for some constant $a > 0$ which is chosen to provide a specified average run length for the procedure when the process is in-control. Second, one-sided or two-sided CSCC (CSCC--SSRM) procedures can be constructed. The one-sided procedure for detecting positive shifts in Δ from the control value $\Delta_0 = 0$ signals at the first n for which

$$\sum_{j=1}^{n} (T_j - k) - \min_{0 \le m \le n} \sum_{j=1}^{m} (T_j - k) \ge h , \qquad (5.3)$$

where $k \ge 0$ and $h>0$ are parameters of the procedure. Otherwise, sampling is continued. The one-sided CSCC for detecting negative shifts in Δ from $\Delta_0 = 0$ signals at the first n for which

$$\max_{0 \le m \le n} \sum_{j=1}^{m} (T_j + k) - \sum_{j=1}^{n} (T_j + k) \ge h . \qquad (5.4)$$

Otherwise, sampling is continued. The two-sided symmetric procedure signals at the first n for which either one of the above two inequalities is satisfied.

Now, attention will be devoted to the computation of the average run lengths of the procedures proposed above. Since the T_j's are stochastically dependent, the stochastic process $\{TR_n; n = 1, 2, \ldots\}$ does not form a Markov chain. This can be seen by noting that

$$\begin{aligned} &\Pr(TR_n = k_n | TR_1 = k_1, TR_2 = k_2, \ldots, TR_{n-M+1} = k_{n-M+1}, \\ &\qquad \ldots, TR_{n-1} = k_{n-1}) \\ &\quad = \Pr(TR_{n-1} + T_n = k_n | T_1 = k_1, T_1 + T_2 = k_2, \ldots, \\ &\qquad T_1 + T_2 + \ldots + T_{n-M+1} = k_{n-M+1}, \ldots, TR_{n-1} = k_{n-1}) \end{aligned}$$

$$= \Pr(T_n = k_n - k_{n-1} | T_1 = k_1, T_2 = k_2 - k_1, \ldots , T_{n-M+1} = k_{n-M+1} - k_{n-M}, \ldots , TR_{n-1} = k_{n-1})$$

$$\neq \Pr(T_n = k_n - k_{n-1} | TR_{n-1} = k_{n-1})$$

since T_n, $T_{n-1}, \ldots, T_{n-M+1}$ are in general stochastically dependent. However,

$$\Pr(T_n = k_n - k_{n-1} | TR_{n-1} = k_{n-1})$$

$$= \Pr(TR_n = k_n | TR_{n-1} = k_{n-1}) \; .$$

Hence,

$$\Pr(TR_n = k_n | TR_1 = k_1, \ldots, TR_{n-1} = k_{n-1})$$

$$\neq \Pr(TR_n = k_n | TR_{n-1} = k_{n-1})$$

which means that $\{TR_n ; n = 1, 2, \ldots\}$ does not form a Markov chain and the approach developed in Section 4.1 is not applicable to compute exact average run lengths of procedures based on the statistic TR_n. Moreover, it was not feasible to develop alternative approaches for computing the exact average run lengths of these procedures. In Section 6.4, approximate analytic formulas will be derived for the ARLs of the CSCC--SSRM and the LB--SSRM procedures.

Simulation studies were made to estimate few values of the ARLs of the CSCC--SSRM and the LB--SSRM procedures for the sake of comparisons with other control chart procedures. Table XII displays the ARLs of one-sided CSCC--SSRM and CSCC--GSR procedures for detecting positive shifts in the mean Δ of a normal distribution with variance 1. The ARLs of the CSCC--GSR procedures are exact and

Table XII. Values of $ARL^{+}(\Delta)$ in Single Observations for One-Sided Procedures Under $N(\Delta,1)$ Distribution.

	Optimal for $\Delta = 0.2$		Optimal for $\Delta = 0.6$	
Δ	CSCC--SSRM $M = 6$ $k=0.36, h=25.0$	CSCC--GSR $g = 6$ $k=3, h=18$	CSCC--SSRM $M = 6$ $k=1.0, h=18.4$	CSCC--GSR $g = 6$ $k=5, h=16$
0.0	99.3	101.0	137.2	140.6
0.20	40.7	39.3	52.3	50.3
0.60	17.7	15.3	19.4	16.6
1.0	13.3	10.4	13.4	10.6
2.0	11.0	6.8	11.1	6.8
3.0	10.9	6.0	11.0	6.0

Table XII (Contd). Values of $ARL^+(\Delta)$ in Single Observations for One-Sided Procedures Under $N(\Delta,1)$ Distribution.

	Optimal for $\Delta = 1.0$		Optimal for $\Delta \geq 2.0$	
Δ	CSCC--SSRM $M = 6$ k=1.4,h=17.3	CSCC--GSR $g = 6$ k=9,h=12	CSCC--SSRM $M = 6$ k=1.7,h=15.3	CSCC--GSR $g = 6$ k=11,h=10
0.0	233.7	249.3	285.8	301.0
0.20	79.4	85.6	96.3	107.9
0.60	24.2	21.2	27.3	24.7
1.0	15.3	11.4	16.0	12.0
2.0	12.1	6.8	12.5	6.8
3.0	12.0	6.0	12.3	6.0

were extracted from Section 4.4. The ARLs of the CSCC--SSRM procedures were estimated by simulation based on 500 iterations. This number of iterations was chosen so that each estimated ARL has a coefficient of variation of no more than 4 percent, an accuracy which is regarded as satisfactory for the present purpose. A moderate value $M = 6$ was chosen as a basis for computing the SSRMs. Values of k and h were chosen so that the CSCC--SSRM procedure is optimal for a specific shift $\Delta \neq 0$ and has a preassigned ARL when $\Delta = 0$. The optimum value of k was determined approximately from $k = (\mu_M/2)$ where $n\mu_M$ is the expected value of TR_n as given in (6.1). When comparing the average run lengths displayed in Table XII, it can be seen that the CSCC--SSRM procedure is, in general, less efficient than the CSCC--GSR procedure. However, when detecting small shifts (i.e., $\Delta = 0.2$), it seems that the two procedures have the same efficiency. When detecting large shifts (i.e., $\Delta \geq 2.0$), the ARL of the CSCC--SSRM procedure is twice the ARL of the CSCC--GSR procedure. Thus, whenever grouped observations can be obtained, the CSCC--GSR procedure should be preferred to the CSCC--SSRM procedure.

In Table XIII, a comparison is made between the LB--SSRM and the LB--GSR procedures under a normal distribution with variance 1. The ARLs of the LB--GSR procedure are exact. The ARLs of the LB--SSRM procedure were estimated by simulation based on 500 iterations, which resulted in a coefficient of variation of no more than 4 percent for each estimated ARL. The value $M = 6$ was again chosen for computing the SSRMs. The linear barriers (i.e., values of a) were determined so that the two procedures have the same ARL when $\Delta = 0$. It can be seen from Table XIII that the LB--SSRM is less efficient than the LB--GSR procedure especially when large shifts are considered.

Table XIII. Values of ARL(Δ) in Single Observations for Two-Sided Procedures Under N(Δ,1) Distribution.

Δ	LB--SSRM M = 6 a = 24	LB--GSR g = 6 a = 21
0.0	48.2	46.5
0.20	32.5	31.5
0.60	15.2	14.3
1.0	11.3	10.2
2.0	9.6	6.8
3.0	9.5	6.0

CHAPTER VI

APPROXIMATE FORMULAS FOR THE ARL

In the previous chapters, methods were given to compute exact average run lengths for some nonparametric control chart procedures proposed in this dissertation. These methods required the computation of transition probabilities and the inversion of certain matrices. Sometimes the transition probabilities are not simple to determine, or it may be cumbersome to invert the relevant matrices because of their sizes. Thus, exact values of average run lengths, although desirable, are not simple to compute. In such cases, it would be desirable to have approximations to the average run lengths that are simple to compute. This chapter is devoted to obtaining the desired approximate formulas for the average run lengths of several control chart procedures.

6.1 General Principles for Approximating Average Run Lengths

The main principle for obtaining approximations to the average run length of control chart procedures that are based on cumulative sums is the approximation of the cumulative sum by a Brownian motion process. More precisely, let ξ_1, ξ_2, ... be a sequence of random variables that are not necessarily independent or identically distributed. The ξ_i could be the observations themselves or functions of the observations. A control chart procedure that is based on cumulative sums is, in general, based on partial sums $S_n = \sum_i^n \xi_i$, $n = 0, 1, 2, \ldots$ with $S_0 = 0$. The sequence of partial sums $\{S_n\}$ forms a discrete time stochastic process which is to be approximated by a Brownian motion process. This approximation involves the following steps:

Step 1: The discrete time process $\{S_n\}$ is replaced by a continuous time process $\{X_N(t);\ 0 \leq t \leq 1\}$ that is defined, for each $N = 1, 2, \ldots$, by

$$X_N(t) = \begin{cases} S_n \text{ for } t = \frac{n}{N}, \ n = 0, 1, 2, \ldots, N. \\ S_{n-1} + (Nt-(n-1))\xi_n \text{ for } \frac{n-1}{N} < t < \frac{n}{N}, \ n = 0, 1, 2, \ldots, N. \end{cases}$$

In words, $X_N(t)$ is defined to coincide with S_n at points $t = \frac{n}{N}$, and for the remaining points $\frac{n-1}{N} < t < \frac{n}{N}$, the value of $X_N(t)$ is defined by linear interpolation between $\frac{n-1}{N}$ and $\frac{n}{N}$ for $n = 0, 1, 2, \ldots, N$. $X_N(t)$ is a polygonal function joining the vertices $\frac{n}{N}$ ($n = 0, 1, 2, \ldots, N$) and it is linear in each subinterval $[\frac{n-1}{N}, \frac{n}{N}]$. $X_N(t)$ can be written more concisely as

$$X_N(t) = S_{[Nt]} + (Nt-[Nt])\xi_{[Nt]+1}$$

$$= S_{[Nt]} + (Nt-[Nt])(S_{[Nt]+1} - S_{[Nt]});\ 0 \leq t \leq 1 ,$$

where [a] denotes the greatest integer less than or equal to a. By letting N vary, one gets a sequence of continuous time stochastic processes $\{X_N(t);\ 0 \leq t \leq 1\}$, $N = 1, 2, \ldots$.

Step 2: The sequence $\{X_N(t);\ 0 \leq t \leq 1\}$ of stochastic processes must be shown to converge to a Brownian motion process. More precisely, it has to be established that

$$(\text{Var } S_N)^{-1/2} (X_N(t) - E\, X_N(t)) \xrightarrow{D} W(t) \text{ as } N \to \infty \quad ;$$

where the symbol $\xrightarrow{D}$ indicates convergence in distribution, and $W(t)$ is the standardized Brownian motion process defined on $0 \leq t \leq 1$ and having mean 0 and variance t. Let $C[0, 1]$ denote the space of continuous real valued functions defined on the interval $[0, 1]$. Associate the space $C[0, 1]$ with the uniform topology which is generated by the metric

$$\rho(X, Y) = \sup_{0 \leq t \leq 1} |X(t) - Y(t)| \quad \text{for } X, Y \in C[0, 1] \quad .$$

The concept of convergence mentioned in this step must be taken to mean the convergence in distribution of a sequence $\{X_N(t)\}$, $N = 1, 2, \ldots$ of random functions in $C[0, 1]$ to a random function $X(t)$ in $C[0,1]$. According to the convention adopted by Billingsley (1968, page 23), a sequence $\{X_N(t)\}$ of random functions is said to converge in distribution to the random function $X(t)$, written as $X_N(t) \xrightarrow{D} X(t)$, if the probability measures P_N of the $X_N(t)$ converge weakly to the probability measure P of X, written as $P_N \Rightarrow P$. The definition of weak convergence of probability measures is given on page 2 of Billingsley. The theory of weak convergence in the space $C[0, 1]$ is contained in Chapter II of Billingsley (1968).

Theorems concerned with convergence to the Brownian motion process are called "functional central limit theorems". In this chapter two such theorems will be needed. For further reference, these theorems will now be stated. The following theorem is well known as Donsker's theorem, Donsker (1951).

Theorem 1. If the random variables ξ_i are independent and identically distributed with mean 0 and variance $\sigma^2 > 0$, then

$$[\mathrm{Var}\ S_N]^{-1/2}\ X_N(t) \xrightarrow{D} W(t) \quad \text{as } N \to \infty \quad .$$

To state the next theorem, the following definition (see, for example, page 166 of Billingsley (1968)) of an "m-dependent" sequence of random variables is needed.

Definition: A strictly stationary sequence $\xi_1, \xi_2, \ldots$ of random variables is said to be m-dependent if, for any $i \leq \ell < j$, the sets $\{\xi_i, \xi_{i+1}, \ldots, \xi_\ell\}$ and $\{\xi_{\ell+r}, \xi_{\ell+r+1}, \ldots, \xi_j\}$ are independent whenever $r > m$.

The statement of Theorem 2 below is a specialization of Theorem 20.1 on page 174 of Billingsley (1968).

Theorem 2. If $\xi_1, \xi_2, \ldots$ is an m-dependent $(m \geq 0)$ sequence, where the ξ_i have mean zero and positive finite variance, then

$$[\mathrm{Var}\ S_N]^{-1/2}\ X_N(t) \xrightarrow{D} W(t) \quad \text{as} \quad N \to \infty \quad .$$

Step 3. The average run length of a control chart procedure based on the cumulative sum $S_n = \sum_{i=1}^{n} \xi_i$ is the expected time that S_n hits the barrier specified by the control chart procedure. This average run length is approximated by the expected time that the approximating Brownian motion process hits that specific barrier.

For the purpose of stating the next two theorems, let $B(t)$ denote a Brownian motion process defined on the interval $[0, \infty)$ and having mean μt and variance $\sigma^2 t$.

Theorem 3. (Darling and Siegret (1953)). If the first passage time is defined by

$$\tau = \sup \{t > 0; -a < B(t) < a\} \quad ,$$

then

$$E(\tau) = \begin{cases} (a/\mu)(\exp[2\mu a/\sigma^2]-1)/(\exp[2\mu a/\sigma^2]+1) & \mu \neq 0 , \\ a^2/\sigma^2 & \mu = 0 . \end{cases}$$

Theorem 4. (Reynolds (1975b)). Define the first passage times

$$\tau^- = \sup\{t > 0; \max_{0 \le u \le t} B(u) - B(t) < h\} \quad ,$$

and

$$\tau^+ = \sup \{t > 0; B(t) - \min_{0 \le u \le t} B(u) < h\} \quad .$$

Then

$$E(\tau^-) = \begin{cases} -(h/\mu) + (\sigma^2/2\mu^2)(\exp[2\mu h/\sigma^2]-1) & \mu \neq 0 , \\ h^2/\sigma^2 & \mu = 0 , \end{cases}$$

and

$$E(\tau^{+}) = \begin{cases} (h/\mu) + (\sigma^2/2\mu^2)(\exp[-2\mu h/\sigma^2] - 1) & \mu \neq 0 \ , \\ h^2/\sigma^2 & \mu = 0 \ . \end{cases}$$

Theorems 3 and 4 provide approximations to the average run lengths of a linear barrier procedure and a CSCC procedure, respectively, as will be shown in the following sections.

6.2 Approximate ARL of a Parametric CSCC Procedure

Suppose that X_1, X_2, ... are independent and identically distributed observations (not necessarily normally distributed) with mean Δ and variance σ^2. Consider the one-sided parametric CSCC procedures based on the above observations and having parameters k and h. Let $ARL^{+}(\Delta)$ and $ARL^{-}(\Delta)$ denote the average run lengths of the one-sided procedures for detecting positive and negative shifts in the mean of the observations. Based on the Brownian motion approximation, Reynolds (1975b) derived the following asymptotic formulas:

$$ARL^{+}(\Delta) = \begin{cases} h/(\Delta-k)+\sigma^2(\exp[-2(\Delta-k)h/\sigma^2]-1)/(2(\Delta-k)^2) & \Delta-k \neq 0 \ , \\ h^2/\sigma^2 & (\Delta-k) = 0 \ . \end{cases}$$

$$ARL^{-}(\Delta) = \begin{cases} -h/(\Delta+k)+\sigma^2(\exp[2(\Delta+k)h/\sigma^2]-1)/(2(\Delta+k)^2) & (\Delta+k) \neq 0 \ , \\ h^2/\sigma^2 & (\Delta+k) = 0 \ . \end{cases}$$

Reynolds (1975b) compared the exact values of ARL^{+}, as tabulated in van Dobben de Bruyn (1968) for normal observations, with the

Brownian motion approximations. He found that the Brownian motion approximation underestimates the correct value of ARL^{+} and for small values of h the discrepancy in the approximation is serious. He suggested a modified Brownian motion approximation by using $h+\delta$, for some $\delta > 0$, instead of h in the formula of the approximation. He also suggested that δ be chosen so that the approximate ARL^{+} at $\mu-k = 0$ is equal to the exact value. However, this does not guarantee that the modified approximation will be good at values of $\mu-k \neq 0$. In this dissertation, a curve fitting analysis will be made to determine a formula for δ as a function of $(\Delta-k)$ and h. A SAS package computer program (PROCEDURE STEPWISE) was employed to determine a linear model fit to δ as

$$\delta = \beta_0+\beta_1(\Delta-k)+\beta_2(\Delta-k)^2+\beta_3|\Delta-k|+\beta_4\sqrt{|\Delta-k|}+\beta_5 h+\beta_6 h^2+\beta_7 h^3+\beta_8\sqrt{h}\ .$$

A linear model fit was chosen because it is simpler to handle than a nonlinear model. The input data for the independent variable $(\Delta-k)$ were eight values ($\Delta-k = -1.2, -0.8, -0.4, -0.2, 0.0, 1.0, 2.0, 4.0$) selected from the range of values considered in the van Dobben de Bruyn tables. The input data for the independent variable h were all the ten values ($h = 1.3, 1.5, 2.0, 2.5, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0$) considered in the van Dobben de Bruyn tables. The input data for the dependent variable δ were computed as follows. For each combination of values of $(\Delta-k)$ and h, δ was computed such that the modified Brownian motion approximation (i.e., using $h+\delta$ instead of h in approximating formula) gives the exact value of the ARL as given in the van Dobben de Bruyn tables. The SAS package program computed several "best" models according to the maximum R^2 improvement criterion. Models with few variables did not provide good fit and models with too many variables are too complicated to be useful in practice. Thus, from the possible models, the best three-variable model, as computed by the SAS program, will be adopted in this dissertation. This three-variable model is

$$\delta = 1.153517 + 0.060216(\Delta-k) + 0.056672(\Delta-k)^2 - 0.000072h^3 .$$

Table XIV displays the exact values of the ARL^+, as given in the van Dobben de Bruyn tables, with the values obtained from the modified Brownian motion approximation where δ is computed from the above three-variable model. For the range of values of $(\Delta-k)$ and h considered in Table XIV, it is seen that the modified approximation is quite satisfactory. No comparison is made with the exact values of ARL^+ when the observations are not normal because these values are not available.

6.3 Approximate ARL of Procedures Based on Grouped Signed-Ranks

The control chart procedures introduced in Chapter IV were based on the cumulative sum $SR_n = \sum_{i=1}^{n} SR_{ig}$ which is a sum of independent and identically distributed random variables SR_{ig}. From the results of Chapter III, we have, for any i = 1, 2 ...,

Table XIV. Values of ARL^+ for Parametric CSCC Procedure Obtained by Modified Brownian Motion Approximation. Under $N(\Delta, 1)$ Distribution.

Δ-k	h = 1.3		h = 2.0		h = 5.0		h = 10.0	
	Exact	Modified B. Motion Approx.	Exact	Modified B. Motion Approx.	Exact	Modified B. Motion Approx.	Exact	Modified B. Motion Approx.
-1.2	117	125.7	610	683.6	-	-	-	-
-0.8	35.4	35.0	114	114.3	14000	14253	-	-
-0.4	13.1	12.8	28.0	27.5	414	402.7	-	-
-0.2	8.7	8.5	15.9	15.7	104	102.3	820	979.9
0	6.1	6.0	10.0	9.94	38.1	37.8	126	122.8
1.0	2.04	2.07	2.74	2.77	5.75	5.76	10.7	10.69
2.0	1.27	1.28	1.58	1.63	3.11	3.12	5.62	5.6
4.0	1.00	0.868	1.02	1.04	1.86	1.79	3.05	3.02

$$E(SR_{ig}) = g(g-1)\xi + g\ \theta = \mu_g \text{ , say,}$$

$$\begin{aligned} Var(SR_{ig}) &= 4g(g-1)(g-2)\ \gamma - 2g(g-1)2g-3)\xi^2 \\ &+ g(g-2)\theta^2 - 4g(g-1)\ \xi\theta + g(g+1)(2g+1)/6 \\ &= \sigma_g^2 \text{ , say.} \end{aligned}$$

It follows that $E(SR_n) = n\mu_g$ and $Var(SR_n) = n\sigma_g^2$. For each N = 1, 2, ..., define

$$X_N(t) = SR_{[Nt]} + (Nt-[Nt])(SR_{[Nt]+1} - SR_{[Nt]}) \text{ ,}$$

$$\text{for } 0 \le t \le 1 \text{ .}$$

Then, by Theorem 1

$$(Var\ SR_N)^{-1/2}\ (X_N(t) - EX_N(t)) \overset{D}{\to} W(t) \quad \text{as } N \to \infty \text{ .}$$

Thus, SR_n can be approximated by a Brownian motion process with mean

$\mu_g t$ and variance $\sigma_g^2 t$. By Theorem 3 of the previous section, the approximate average run length of a symmetric two-sided LB--GSR procedure with parameter a is given by

$$\mathrm{ARL} = \begin{cases} (a/\mu_g)(\exp[2\mu_g a/\sigma_g^2]-1)/(\exp[2\mu_g a/\sigma_g^2]+1) & \mu_g \neq 0 \ , \\ a^2/\sigma_g^2 & \mu_g = 0 \ . \end{cases}$$

Table XV displays exact values of ARL of the symmetric two-sided LB--GSR procedure with values obtained by the Brownian motion approximation. There is indication that the Brownian motion approximation is satisfactory except for $\Delta=0$ and small shifts such as $\Delta = 0.20$.

A one-sided CSCC--GSR procedure for detecting negative deviations in the mean of the observations signals at the first n for which

$$\max_{0 \le m \le n} \sum_{i=1}^{m} (SR_{ig}+k) - \sum_{i=1}^{n} (SR_{ig}+k) \ge h \ .$$

Since

$$E \sum_{i=1}^{n} (SR_{ig}+k) = (\mu_g+k)\, n \ ,$$

and

$$\mathrm{Var} \sum_{i=1}^{n} (SR_{ig}+k) = \sigma_g^2\, n \ ,$$

then $\sum_{i=1}^{n} (SR_{ig}+k)$ can be approximated by a Brownian motion process with mean $(\mu_g+k)t$ and variance $\sigma_g^2\, t$. It follows from Theorem 4 that the average run length of this procedure is approximated by

Table XV. Values of ARL for Two-Sided LB--GSR Procedure Obtained by Brownian Motion Approximation. Under N(Δ, 1) Distribution.

Shift Δ	a = 21, g = 6		a = 55, g = 10	
	Exact	B. Motion Approx.	Exact	B. Motion Approx.
0.0	46.5	29.1	114.2	78.6
0.20	31.5	22.9	56.7	44.6
0.60	14.3	10.7	23.6	17.4
1.00	10.2	7.5	18.3	12.3
2.00	6.80	6.1	12.0	10.1
3.00	6.00	6.0	10.1	10.0

$$\mathrm{ARL}^- = \begin{cases} -h/(\mu_g+k) + \sigma_g^2(\exp[2(\mu_g+k)h/\sigma_g^2]-1)/(2(\mu_g+k)^2) & (\mu_g+k) \neq 0 , \\ h^2/\sigma_g^2 & (\mu_g+k) = 0 . \end{cases}$$

A one-sided CSCC--GSR procedure for detecting positive deviations in the mean of the observations signals at the first n for which

$$\sum_{i=1}^{n} (SR_{ig}-k) - \min_{0 \le m \le n} \sum_{i=1}^{m} (SR_{ig}-k) \ge h .$$

The sum $\sum_{i=1}^{n} (SR_{ig}-k)$ can be approximated by a Brownian motion process with mean $(\mu_g-k)t$ and variance $\sigma_g^2 t$. Thus the average run length of the procedure is approximated by

$$\mathrm{ARL}^+ = \begin{cases} h/(\mu_g-k) + \sigma_g^2(\exp[-2(\mu_g-k)h/\sigma_g^2]-1)/(2(\mu_g-k)^2) & (\mu_g-k) \neq 0 , \\ h^2/\sigma_g^2 & (\mu_g-k) = 0 . \end{cases}$$

The average run length for a symmetric two-sided CSCC--GSR procedure can be approximated by

$$\mathrm{ARL} = (\mathrm{ARL}^+)(\mathrm{ARL}^-)/(\mathrm{ARL}^+ + \mathrm{ARL}^-) .$$

In Table XVI, exact values of ARL^+ of a one-sided CSCC--GSR procedure were compared to Brownian motion approximation values. Group sizes g = 2, 6, 10 were considered as representing small, medium, and large sizes, respectively. When $\Delta \neq 0$ the average run lengths were computed under the assumption that the observations are normally distributed. The comparisons show that the Brownian motion approximations

Table XVI. Values of ARL^{+} in Single Observations for a One-Sided CSCC--GSR Procedure Obtained by Brownian Motion Approximation Under N(Δ, 1) Distribution.

Shift Δ	g = 2 k = 1 h = 10		g = 6 k = 3 h = 18		g = 10 k = 5 h = 50	
	Exact	B. Motion Approx.	Exact	B. Motion Approx.	Exact	B. Motion Approx.
0.0	648.0	248.0	101.0	33.1	272.5	105.2
0.20	149.0	87.8	39.3	19.0	67.6	40.8
0.60	29.5	26.4	15.3	10.1	24.8	17.1
1.00	15.8	15.2	10.4	7.5	18.4	12.3
2.00	10.5	10.4	6.8	6.1	12.0	10.1
3.00	10.0	10.0	6.0	6.0	10.1	10.0

underestimate the exact values of the average run lengths. For small values of Δ the approximation is very far from the correct value. This suggests a need to modify the approximation as described in the previous section. However, adequate data on the exact ARL^+ that include various values of k, h, and group sizes must be generated in order that a reliable curve fit can be achieved.

6.4 Approximate ARL for Procedures Based on SSRM

Control chart procedures based on ranking with respect to last M observations essentially use the statistic $TR_n = \sum_{j=1}^{n} T_j$ as defined in Chapter V. To obtain approximations to the average run lengths of such procedures, it must be established that $\{TR_n;\ n = 1, 2, \ldots\}$ can be approximated by a Brownian motion process. First, it is necessary to derive the mean and variance of TR_n as these will be used in the approximations. The quantities ψ_{ij} were defined and their properties were given in Section 3.2. It can be seen that

$$T_j = \begin{cases} \sum_{i=1}^{j} \psi_{ij} & j < M , \\ \\ \sum_{i=j-M+1}^{j} \psi_{ij} & j \geq M . \end{cases}$$

It follows that

$$TR_n = \sum_{j=1}^{n} T_j = \sum_{j=1}^{n} \sum_{i=j-M+1}^{j} \psi_{ij} .$$

Thus,

$$E(T_j) = \sum_{i=1}^{j} E\psi_{ij} = (j-1)E\psi_{ij} + E\psi_{jj}$$

$$= (j-1)\xi + \theta \qquad\qquad j < M \quad ,$$

and

$$E(T_j) = \sum_{i=j-M+1}^{j} E\psi_{ij} = (M-1)E\psi_{ij} + E\psi_{jj}$$

$$= (M-1)\xi + \theta \qquad\qquad j > M \quad .$$

It follows that for $n > M$

$$E(TR_n) = \sum_{j=1}^{M-1} E(T_j) + \sum_{j=M}^{n} E(T_j)$$

$$= \sum_{j=1}^{M-1} ((j-1)\xi+\theta) + (n-M+1)((M-1)\xi+\theta)$$

$$= (M-1)(M-2)\xi/2 + (M-1)\theta + (n-M+1)((M-1)\xi+\theta)$$

$$= n((M-1)\xi+\theta) - M(M-1)\xi/2$$

$$= n\mu_M + o(n) \quad ,$$

where

$$\mu_M = (M-1)\xi+\theta \quad . \tag{6.1}$$

The derivation of the asymptotic variance is somewhat involved and will now be given.

$$\mathrm{Var}(TR_n) = \sum_{j=1}^{n} \mathrm{Var}(T_j) + 2 \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \mathrm{Cov}(T_j, T_k)$$

$$= \sum_{j=M}^{n} \mathrm{Var}(T_j) + 2 \sum_{j=M}^{n-1} \sum_{k=j+1}^{n} \mathrm{Cov}(T_j, T_k)$$

$$+ o(n) \quad ,$$

since $\mathrm{Cov}(T_j, T_k) = 0$ for $k-j > M-1$. For $j \geq M$ it is seen that

$$\mathrm{Var}(T_j) = \sum_{i=j-M+1}^{j} \mathrm{Var}\ \psi_{ij} + 2 \sum_{i=j-M+1}^{j-1} \sum_{\ell=i+1}^{j} \mathrm{Cov}(\psi_{ij}, \psi_{\ell j})$$

$$= \mathrm{Var}\ \psi_{jj} + \sum_{i=j-M+1}^{j-1} \mathrm{Var}\ \psi_{ij} + \sum_{i=j-M+1}^{j-1} \mathrm{Cov}(\psi_{ij}, \psi_{jj})$$

$$+ \sum_{i=j-M+1}^{j-2} \sum_{\ell=i+1}^{j-1} \mathrm{Cov}(\psi_{ij}, \psi_{\ell j})$$

$$= (1-\theta^2) + (M-1)(1-2\xi^2)/2 + (M-1)(1-2\xi\theta)/2$$

$$+ (M-1)(M-2)(1/3 - \xi^2)/2$$

$$= (M-1)^2(1-3\xi^2)/3 + (M-1)(7-12\xi\theta)/6$$

$$+ (1-\theta^2) \quad .$$

It follows that

$$\sum_{j=M}^{n} \mathrm{Var}(T_j) = (n-M+1)\{(M-1)^2(1-3\xi^2)/3+(M-1)(7-12\xi\theta)/6$$

$$+ (1-\theta^2)\}$$

$$= n\{(M-1)^2(1-3\xi^2)/3+(M-1)(7-12\xi\theta)/6+(1-\theta^2)\}+o(n) \quad .$$

Since the sequence $T_M, T_{M+1}, \ldots$ is (M-1)-dependent, then

$$\mathrm{Cov}(T_s, T_{s+r}) = 0 \qquad s \geq M,\ r > M-1 \ .$$

For $s \geq M$ and $1 \leq r \leq M-1$, it can be shown that

$$\mathrm{Cov}(T_s, T_{s+r}) = \sum_{i=s-M+1}^{s} \sum_{\ell=s-M+r+1}^{s+r} \mathrm{Cov}(\psi_{is}, \psi_{\ell(s+r)})$$

$$= (M-(r+1))(2\gamma-\xi^2)+(M-1)(\gamma-\xi^2)$$

$$+ (\theta^2-2\xi\theta)/2 \ . \tag{6.2}$$

The covariances in $\sum_{j=M}^{n-1} \sum_{k=j+1}^{n} \mathrm{Cov}(T_j, T_k)$ are displayed in Figure 2.
From this figure it is seen that

$$\sum_{j=M}^{n-1} \sum_{k=j+1}^{n} \mathrm{Cov}(T_j, T_k) = \sum_{s=0}^{n-2M} \sum_{r=1}^{M-1} \mathrm{Cov}(T_{M+s}, T_{M+s+r})$$

$$+ \sum_{s=n-M+1}^{n-1} \sum_{r=1}^{n-s} \mathrm{Cov}(T_s, T_{s+r}) \ .$$

Using (6.2), it is seen that

$$\sum_{s=0}^{n-2M} \sum_{r=1}^{M-1} \mathrm{Cov}(T_{M+s}, T_{M+s+r}) = \sum_{s=0}^{n-2M} \sum_{r=1}^{M-1} \{(M-(r+1))(2\gamma-\xi^2)$$

$$+ (M-1)(\gamma-\xi^2)+(\theta^2-2\xi\theta)/2\}$$

$$= (n-2M+1) \sum_{r=1}^{M-1} \{(M-1)(3\gamma-2\xi^2)$$

$$- r(2\gamma-\xi^2) + (\theta^2-2\xi\theta)/2\}$$

$$= (n-2M+1)\{(M-1)^2(3\gamma-2\xi^2)$$

$$- (M/2)(M-1)(2\gamma-\xi^2)+(M-1)(\theta^2-2\xi\theta)/2\}$$

$$= n(M-1)\{(2M-3)\gamma-(3M-4)\xi^2/2+(\theta^2-\xi\theta)/2\}$$

$$- (2M-1)(M-1)\{(2M-3)\gamma-(3M-4)\xi^2/2+(\theta^2-2\xi\theta)/2\}$$

$$= n(M-1)\{(2M-3)\gamma-(3M-4)\xi^2/2+(\theta^2-\xi\theta)/2\}$$

$$+ o(n) \quad .$$

It can be shown that

$$\sum_{s=n-M+1}^{n-1} \sum_{r=1}^{n-s} \mathrm{Cov}(T_s, T_{s+r}) = M(M-1)^2(3\gamma-2\xi^2)/2+M(M-1)(\theta^2-2\xi\theta)/2$$

$$+ M^3(2\gamma-\xi^2)/6$$

$$= o(n) \quad .$$

It follows that

$$\mathrm{Var}(TR_n) = \sum_{j=M}^{n} \mathrm{Var}(T_j)+2\sum_{j=M}^{n-1} \sum_{k=j+1}^{n} \mathrm{Cov}(T_j, T_k)+o(n)$$

$$= n\{(M-1)^2(1-3\xi^2)/3+(M-1)(7-12\xi\theta)/6+(1-\theta^2)\}$$

$$+ 2n(M-1)\{(2M-3)\gamma-(3M-4)\xi^2/2+(\theta^2-2\xi\theta)/2\}$$

$$+ o(n)$$

$$= n\{(M-1)(4M-6)\gamma-(4M-5)(M-1)\xi^2+(M-2)\theta^2$$

j \ k	M+1	M+2	...	(2M-1)	2M	...	n-M+1	n-M+2	...	n-1	n
M	(M,M+1)	(M,M+2)	...	(M,2M-1)	0	...	0	0	...	0	0
M+1		(M+1,M+2)	...	(M+1,2M-1)	(M+1,2M)	...	0	0	...	0	0
⋮					. . .						
n-M							(n-M,n-M+1)	(n-M,n-M+2)	...	(n-M,n-1)	0
n-M+1								(n-M+1,n-M+2)	...	(n-M+1,n-1)	(n-M+1,n)
⋮					. . .						
n-2										(n-2,n-1)	(n-2,n)
n-1											(n-1,n)

Figure 2. Display of the Covariances in $\sum_{j=M}^{n-1} \sum_{k=j+1}^{n} \mathrm{Cov}(T_j, T_k)$

Where Only the Subscripts are Shown.

$$- 4(M-1)\xi\theta+(M+1)(2M+1)/6\}+o(n)$$

$$= n\, \sigma_M^2 + o(n) \;,$$

where $\sigma_M^2 = (M-1)(4M-6)\gamma-(4M-5)(M-1)\xi^2+(M-2)\theta^2-4(M-1)\xi\theta+(M+1)(2M+1)/6$.

If the observations are symmetrically distributed about zero, then

$$\mathrm{Var}(TR_N) = (M+1)(2M+1)/6 \;.$$

Theorem 2 will now be used to obtain the Brownian motion approximation to the sequence $\{TR_n\}$.

It is seen that the sequence $T_1, T_2, \ldots, T_M, T_{M+1}, \ldots$, is not strictly stationary since the random variables $T_1, T_2, \ldots, T_{M-1}$ are not identically distributed. For example, $ET_j = (j-1)\xi_1 + \xi_2$, for $j < M$, which depends on j. However, if the segment $T_1, T_2, \ldots, T_{M-1}$ is neglected, then the sequence $T_M, T_{M+1}, \ldots$ is strictly stationary as will be shown below. In Section 5.2 it was shwon that $(T_i, \ldots, T_\ell)$ and $(T_{\ell+r}, \ldots, T_j)$ are stochastically independent whenever $r > (M-1)$. It follows that the sequence is an (M-1)-dependent sequence as defined in Section 6.1. Strict stationarity of $T_M, T_{M+1}, \ldots$ can be argued as follows. Consider the two events $(T_{t_1} = L_1, T_{t_2} = L_2, \ldots, T_{t_r} = L_r)$ and $(T_{t_1+d} = L_1, T_{t_2+d} = L_2, \ldots, T_{t_r+d} = L_r)$ for any integers $d \geq 1$ and $r \geq 1$. Suppose that p of the integers $L_1, \ldots, L_r$ are positive. The event $(T_{t_1} = L_1, T_{t_2} = L_2, \ldots, T_{t_r} = L_r)$ corresponds to a number of orderings of absolute values of the observations in the set $E_1 = \{X_{t_1-M+1}, \ldots, X_{t_1}, \ldots, X_{t_1+r}\}$ such that p of the last r X's are positive. Similarly, the event $(T_{t_1+d} = L_1, T_{t_2+d} = L_2, \ldots, T_{t_r+d} = L_r)$

corresponds to a number of orderings of absolute values of the observations in the set $E_2 = \{X_{t_1+d-M+1}, \ldots, X_{t_1+d}, \ldots, X_{t_r+d}\}$ such that p of the last r X's are positive. Since the observations $X_1, X_2, \ldots$ are assumed to be independent and identically distributed and the sets E_1 and E_2 contain the same number of observations $(r+M-1)$ it follows that

$$\Pr(T_{t_1} = L_1, T_{t_2} = L_2, \ldots, T_{t_r} = L_r) = \Pr(T_{t_1+d} = L_1, T_{t_2+d} = L_2, \ldots, T_{t_r+d} = L_r) \quad .$$

Thus, the sequence $T_M, T_{M+1}, \ldots$ is strictly stationary. Now define $TR^*_n = \sum_{j=M}^{n} T_j$ and for each N = 1, 2, ... define

$$X^*_N(t) = TR^*_{[Nt]} + (Nt-[Nt])T_{[Nt]+1} \quad \text{for} \quad 0 \le t \le 1 \quad .$$

By Theorem 2 of Section 6.1 it follows that

$$(\text{Var } TR^*_N)^{-1/2} (X^*_N(t) - EX^*_N(t)) \xrightarrow{D} W(t) \quad \text{as} \quad N \to \infty \quad .$$

In the above discussion, the terms $T_1, T_2, \ldots, T_{M-1}$ were neglected. They have negligible effect as $N \to \infty$. Therefore the whole sum $TR_n = \sum_{j=1}^{n} T_j$ may be approximated by a Brownian motion process with mean $\mu_M t$ and variance $\sigma^2_M t$. Hence, an approximation to the ARL of a two-sided linear barrier procedure (-a, a) is given according to Theorem 3 by

$$\text{ARL} = \begin{cases} (a/\mu_M)(\exp[2\mu_M a/\sigma_M^2]-1)/(\exp[2\mu_M a/\sigma_M^2]+1) & \mu_M \neq 0 \quad , \\ a^2/\sigma_M^2 & \mu_M = 0 \quad . \end{cases}$$

No exact values of ARL for the above linear barrier procedure are available to check the accuracy of the Brownian motion approximation. The approximations were compared to values of ARL obtained by simulation based on 100 iterations and assuming a normal distribution with variance 1. This number of iterations gave a cofficient of variation of no more than 15 percent for each estimated ARL. The results are contained in Table XVII and they indicate that the Brownian motion approximations are satisfactory for linear barrier procedures.

For CSCC-type procedures based on ranking with respect to last M observations one can employ Theorem 4 of Section 6.1 to obtain the approximations to the average run lengths.

For a one-sided procedure, with parameters k and h, for detecting negative shifts, the approximate average run length is given by

$$\text{ARL}^- = \begin{cases} -h/(\mu_M+k)+\sigma_M^2(\exp[2(\mu_M+k)h/\sigma_M^2]-1)/(2(\mu_M+k)^2)] & \mu_M+k \neq 0 \quad , \\ h^2/\sigma_M^2 & \mu_M+k = 0 \quad . \end{cases}$$

For the one-sided procedure for detecting positive shifts, the approximate average run length is given by

$$\text{ARL}^+ = \begin{cases} h/(\mu_M-k)+\sigma_M^2(\exp[-2(\mu_M-k)h/\sigma_M^2]-1)/(2(\mu_M-k)^2) & \mu_M-k \neq 0 \quad , \\ h^2/\sigma_M^2 & \mu_M-k = 0 \quad . \end{cases}$$

Table XVII. Values of ARL Obtained by Brownian Motion Approximation and by Simulation for Two-Sided Linear Barrier Procedure Based on TR_n. Under $N(\Delta, 1)$ Distribution

Shift Δ	M = 2, a = 15		M = 6, a = 38		M = 10, a = 87	
	B. Motion Approx.	Simulation	B. Motion Approx.	Simulation	B. Motion Approx.	Simulation
0.0	90.0	106.69	100.00	108.34	196.60	228.51
0.25	43.45	47.72	42.05	47.19	60.32	68.30
0.50	23.32	24.77	22.57	24.87	31.93	37.77
1.00	13.59	14.48	13.62	16.19	19.44	24.49
2.00	10.33	10.92	11.04	13.58	16.01	20.66
3.00	10.02	10.52	10.87	13.25	15.83	20.42

Exact values for the ARL of CSCC procedures based on ranking with respect to last M observations are not available. To get reliable results using simulation would be expensive. Hence, it was not feasible to check the accuracy of the Brownian motion approximation.

CHAPTER VII

CONCLUSIONS AND PROBLEMS FOR FURTHER RESEARCH

Several conclusions and recommendations concerning the proposed nonparametric control chart procedures are made in Section 7.1. In Section 7.2, possible extensions to the present research are discussed and ideas for further research are suggested.

7.1 Conclusions

Four new nonparametric control chart procedures for detecting shifts in the center of symmetry Δ of a distribution from a control value $\Delta_0 = 0$ were proposed and investigated in this dissertation. Two of them, the CSCC--GSR and the LB--GSR, are based on the method of ranking the observations within groups. The other two procedures, the CSCC--SSRM and the LB--SSRM, are designed for the case where single observations are made on the output of the process and grouping is undesirable. A method, based on a Markov chain approach, was developed for computing the exact ARLs of the CSCC--GSR and the LB--GSR procedures. Analytic formulas, based on Brownian motion approximations, were derived for the ARLs of the four procedures.

All the proposed procedures share the following advantages.

(1) They are simple to apply in practice. At each stage of carrying out the procedures, the ranks of the observations and the test statistics are simple to compute.

(2) The application of the procedures does not require that the probability distribution or the variance of the observations be known.

(3) When the process is in-control, the ARL of each procedure does not depend on the parent distribution of the observations.

(4) It was possible to derive an approximate analytic formula for the ARL of each procedure. The CSCC--GSR and the LB--GSR procedures posses the further advantage that their ARLs can be computed exactly in a relatively simple way.

As far as the efficiency of the CSCC--GSR and the LB--GSR procedures is concerned, it is possible to draw the following conclusions.

(1) For detecting small shifts ($\Delta = 0.2$), the LB--GSR and the CSCC--GSR procedures proved to be very efficient. They are more efficient than the Shewhart's procedure (whether based on grouped or ungrouped observations) even under the normality assumption. They are very close in efficiency to the parametric CSCC procedure even under the normality assumption. A limited simulation study indicated that the CSCC--GSR procedure is more efficient than the parametric CSCC procedure when the observations have a double exponential distribution. The above claims can be substantiated by referring to Tables IXa, IXf, Xa, XIa, XIc.

(2) For detecting medium shifts ($\Delta = 0.6$ and 1.0), the LB--GSR and the CSCC--GSR procedures are more efficient than the ungrouped Shewhart's procedure even when the observations have a normal distribution. For detecting a shift of approximately $\Delta = 0.6$, the CSCC--GSR is as efficient as the grouped Shewhart procedure;

but it is less efficient than the parametric CSCC procedure, when the observations have a normal distribution. When the observations have a double exponential distribution with Δ = 0.6, the simulation study indicates that the CSCC--GSR procedure is more efficient than the Shewhart procedure and as efficient as the parametric CSCC procedure. The LB--GSR procedure is not efficient at medium shifts when the observations have a normal distribution. No results are available to check the efficiency of the LB--GSR procedure when the observations have a distribution other than the normal. The above conclusions can be substantiated by referring to Tables: IXb, IXc, IXg, XIb, XId.

(3) Finally, consider the case of detecting large shifts ($\Delta \geq 2.0$) and were observations are grouped. If the procedures (whether parametric or nonparametric) are designed so that it is possible to stop after one group then all will have ARL equal to the group size for every large shift. Thus, the LB--GSR and the CSCC--GSR procedures will have the same efficiency as the parametric procedures. When the observations are ungrouped, any nonparametric control chart procedure based on ranks will be inefficient when compared with parametric procedures at large values of Δ. A reason for this is that nonparametric procedures replace the values of the observations by ranks which are bounded in value no matter how large the shift Δ is and a number of observations is required before the sum of the ranks can reach a certain value. For example,for a CSCC procedure with k = 0 and

h = 21, a nonparametric procedure based on the Wilcoxon signed-ranks will need at least six observations (no matter how large Δ is) before the cumulative sum of the signed-ranks can reach the value of h.

7.2 Problems for Further Research

While this investigation has solved some problems of interest, it gave rise to other worthwhile problems that are left for further research.

Some ideas for further research are:

(1) More comparisons need to be performed between the CSCC--GSR and the parametric CSCC when the observations have distributions other than normal. Although such comparisons seem to be too straightforward, the computational effort involved is not trivial. The computation of the exact ARL of the parametric CSCC procedure requires solving integral equations which usually have to be solved numerically. The computation of the exact ARL of the CSCC--GSR procedure requires that the distribution of the Wilcoxon signed-rank statistic be first computed. Despite these difficulties, it is felt that this problem is worthwhile since, as a by-product, it is possible to study the effect of nonnormality on the ARL of the parametric CSCC procedure. Also, it is possible to use the distribution of the Wilcoxon signed-rank statistic for purposes other than computing the ARL of the CSCC--GSR procedure. For example, it can be used to study the exact power of the Wilcoxon signed-rank test.

(2) It is worthwhile to study the effect of group size on the efficiency of the LB--GSR and the CSCC--GSR procedures. Intuitively, it seems that increasing the group size has the effect of increasing the efficiency of the grouped signed-rank procedures. However, it should be noted that using groups of a large size will also make the ARL large (= group size) when considering large shifts.

(3) The performance of procedures based on the SSRMs needs more in-investigation. Extensive simulation studies need be made to obtain reliable values of the ARLs for various values of M.

(4) It is very desirable to improve the Brownian motion approximations to the ARLs of the several control chart procedures. If the Brownian motion approximations can be made to give reliable results, then this will result in a very simple and cheap way to compute the ARLs of the control chart procedures.

REFERENCES

Arnold, H. J. (1965). "Small Sample Power of the One Sample Wilcoxon Test for Non-Normal Shift Alternatives". Ann. Math. Statist. Vol. 36: 1767-1778.

Barnard, G. A. (1959). "Control Charts and Stochastic Processes". J. Roy. Statist. Soc. Ser. B. Vol. 21: 239-271.

Billingsley, P. (1968). Convergence of Probability Measures. John Wiley & Sons, New York.

van Dobben de Bruyn, C. S. (1968). Cumulative Sum Tests. Hafner, New York.

Donsker, M. D. (1951). "An Invariance Principle for Certain Probability Limit Theorems". Mem. Amer. Math. Soc., No. 6.

Duncan, A. J. (1974). Quality Control and Industrial Statistics. 4th ed. Richard D. Irwin, Inc.

Hollander, M. and Wolfe, D. A. (1973). Nonparametric Statistical Methods. John Wiley & Sons, New York.

Johnson, N. L. (1961). "A Simple Theoretical Approach to Cumulative Sum Control Charts". J. Amer. Statist. Assoc., Vol. 56: 835-840.

Klotz, J. H. (1963). "Small Sample Power and Efficiency for the One Sample Wilcoxon and Normal Scores Tests". Ann. Math. Statist., Vol. 34: 624-632.

Lehmann, E. L. (1975). Nonparametrics: Statistical Methods Based on Ranks. Holden-Day, Inc., San Francisco, Calif.

__________, (1953). "The Power of Rank Tests". Ann. Math. Statist., Vol. 24: 23-43.

Lever, C. S. (1973). "Rank Order Probabilities: One Sample Normal Shift Alternatives, Part A: Table of Rank Order Probabilities". Tech. Report No. ORNL-4887, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Page, E. S. (1954). "Continuous Inspection Schemes". Biometrika, Vol. 41: 100-114.

__________, (1955). "Control Charts with Warning Lines". Biometrika, Vol. 42: 243-257.

Parent, E. A., Jr. (1965). "Sequential Ranking Procedures". Tech. Report No. 80, Dept. of Statist., Stanford Univ.

Parzen, E. (1962). Stochastic Processes. Holden-Day, Inc., San Francisco, Calif.

Phillips, M. J. (1969). "A Survey of Sampling Procedures for Continuous Production". J. Roy. Statist. Soc. Ser. A, Vol. 132: 205-228.

Reynolds, M. R., Jr. (1972). "A Sequential Nonparametric Test for Symmetry with Applications to Process Control". Tech. Report No. 148, Dept. of Oper. Res. and Dept. of Statist., Stanford Univ.

__________, (1975a). "A Sequential Signed-Rank Test for Symmetry". Ann. Statist., Vol. 3: 382-400.

__________, (1975b). "Approximations to the Average Run Length in Cumulative Sum Control Charts". Technometrics, Vol. 17: 65-71.

Shewhart, W. A. (1931). Economic Control of Quality of Manufactured Product. van Nostrand, New York.

Weed, H. D., Jr. and Bradley, R. A. (1971). "Sequential One-Sample Grouped Signed Rank Tests for Symmetry: Basic Procedures". J. Amer. Statist. Assoc., Vol. 66: 321-326.

Wilcoxon, F. (1954). "Individual Comparisons by Ranking Methods". Biometrics, Vol. 1: 80-83.

Wilcoxon, F., Rhodes, L. J. and Bradley, R. A. (1963). "Two Sequential Two-Sample Grouped Rank Tests with Applications to Screening Experiments". Biometrics, Vol. 19: 58-84.

The two page vita has been removed from the scanned document. Page 1 of 2

The two page vita has been removed from the scanned document. Page 2 of 2

NONPARAMETRIC PROCEDURES FOR PROCESS CONTROL

by

Saad Taha Bakir

(ABSTRACT)

Three nonparametric control chart procedures are developed. The procedures are designed to detect any shift in the median of a sequence of observations from a specified control value.

The first two procedures require that groups of $g \geq 1$ observations be made sequentially on the output of the process. Then the Wilcoxon signed-ranks of the observations are computed within each group separately. The Wilcoxon signed-rank statistic SR_{ig} for the i^{th} group is computed as the sum of the signed ranks of the g observations in the i^{th} group. One of the two procedures employs a cumulative sum control chart-type stopping rule and it signals indicating a shift in the median of the process at the first n for which

$$\sum_{i=1}^{n} (SR_{ig} - k) - \min_{0 \leq m \leq n} \sum_{i=1}^{m} (SR_{ig} - k) \geq h ,$$

where $k \geq 0$ and $h > 0$ are parameters of the procedures, and where $\sum_{i=1}^{0} \equiv 0$.

The other procedure employs a linear barrier-type stopping rule and it signals at the first n for which

$$\sum_{i=1}^{n} SR_{ig} \notin (-a, a) ,$$

where $a > 0$ is a parameter of the procedure.

Based on the fact that $\{ \sum_{i=1}^{n} SR_{ig};\ n = 1, 2, \ldots\}$ forms a discrete time Markov chain, a method for determining the exact properties (average run lengths) of the procedures was developed.

The third procedure is proposed for situations where single observations, rather than grouped observations, are made on the output of the process. The procedure requires that an integer $M > 1$ be fixed apriori and the rank of an observation be computed only with respect to the preceding (M-1) observations. The procedure employs the sum of the signed ranks as a test statistic and a cumulative sum control chart-type stopping rule. It was not possible to determine the exact properties of the procedure through a Markov chain approach.

All the proposed procedures are simple to apply in practice since they require little effort in computing the ranks of the observations. Their application does not require that the distribution or the variance of the observations be known.

Several comparisons of the proposed procedures were made with other parametric control chart procedures. For normal observations and when a small shift in the mean is considered, there is indication that the proposed procedures perform nearly as good as the parametric procedures. For double exponential observations, some of the proposed nonparametric procedures perform better than the parametric procedures when a small shift in the mean is considered.