# AN ANALYSIS OF SHEWHART QUALITY CONTROL CHARTS TO MONITOR BOTH THE MEAN AND VARIABILITY

by

KEITH JACOB BARRS

A Research Project Submitted to the Faculty

of the College of Graduate Studies

at Georgia Southern University

in Partial Fulfillment of the

Requirements of the Degree

Master of Science

Statesboro, Georgia

1996

**AN ANALYSIS OF SHEWHART QUALITY CONTROL CHARTS**

**TO MONITOR BOTH THE MEAN AND VARIABILITY**

**by**

**KEITH JACOB BARRS**

____________________________________
**Charles W. Champ, Chairperson**

____________________________________
**Richard J. Hathaway**

____________________________________
**John A. Rafter**

**Approved:**

____________________________________________________ ____________
**G. Lane Van Tassell** **Date**
**Associate Vice President for Academic Affairs**
**and Dean of Graduate Studies**

# ACKNOWLEDGMENTS

I would like to express my appreciation to Dr. Charles Champ for his guidance, patience and expertise throughout the completion of this project. Many thanks to my wife Sharon Morris Barrs for her help on word processing and answering questions about the computer. I would also like to thank Dr. Richard Hathaway and Dr. John Rafter for participating as committee members. But most of all, I would like to thank God for all things.

## ABSTRACT

When monitoring the mean of a continuous quality measure it is often recommended a separate chart be used to monitor the variability. These charts are traditionally designed separately. This project considers them together as a combined charting procedure and gives recommendations for their design. This is based on an average run length (ARL) analysis. The run length distribution is determined using two methods both based on a Markov chain approach.

TABLE OF CONTENTS

LIST OF TABLES

# Chapter 1. INTRODUCTION

The distribution of a quality characteristic often has more than one parameter. Each of these parameters may need to be monitored. This often leads to the use of more than one control chart for monitoring the process. In order to improve the sensitivity of a Shewhart chart to small shifts in the parameter of the quality characteristic being monitored, runs rules are added. A runs rule has the general form that if k out of the last m sample statistics fall in the interval (a,b) an out-of-control signal is given, where $k < m$ and $a < b$. A convenient notation for this runs rule is T (k,m,a,b).

Page (1955), the Western Electric Handbook (1956), Roberts (1957), Bissell (1978), Wheeler (1983), Coleman (1986), Palm (1990), and Champ, Lowry, and Woodall (1992) give discussions of runs rules. A complete description of how a Shewhart chart with runs rules can be expressed as a Markov chain was given by Champ (1986) and Champ and Woodall (1987). Similar, but not as complete, results using a Markov chain approach was obtained by Coleman (1986). Page (1955) and Bissell (1978) used a Markov chain approach for some simple combinations of runs rules. A Markov chain approach was used by Palm (1990) to obtain the percentage points of the run length for the Shewhart $\overline{X}$-chart. Champ, Lowry, and Woodall (1992) used a Markov chain to analyze the Shewhart R- and S-charts. Champ (1986) uses a Markov chain approach to analyze the run length distribution of the combined $\overline{X}$- and S- control charts used to monitor the mean and standard deviation of a normal distribution. Bragg and St. John (1991) used simulation to analyze the combined $\overline{X}$- and R- chart supplemented with runs rules.

In this research, we will investigate the use of individual Shewhart control charts with supplementary runs rules to monitor both the mean and standard deviation of a quality measure. Although separate charts are used, the charting procedure will be considered as a combined charting procedure. In Chapter 2, the Shewhart $\overline{X}$, R, and S quality control charts supplemented with runs rules are discussed. The procedures given in the literature for setting up these individual charts are given. Other combined mean and variability charts are discussed. In Chapter 3, the Markov chain approach of Champ and Woodall (1987) is used to develop two methods for analyzing the run length distribution of the combined $\overline{X}$- and R- (or S-) charts supplemented with runs rules. These methods provide simple ways to obtain run length properties of the combined chart when monitoring the mean and standard deviation of a normal distribution. Recommendations for selecting a combined charting procedure are given in Chapter 4.

# Chapter 2. COMBINED $\overline{X}$- AND R- (or S-) CHARTS

## 2.1 Introduction

The standard practice when monitoring the distribution of a continuous quality measure, X, is to monitor both the distribution mean and standard deviation. These two parameters of the distribution of the quality measure are also referred to as the process mean and standard deviation. The Shewhart $\overline{X}$- chart is commonly used to monitor the mean. To monitor the standard deviation, either the Shewhart R- or S- chart is used. According to Montgomery (1991), the R-chart is more widely used than the S-chart.

Although the main objective is usually to monitor the mean of a process, Montgomery (1991) among others, recommends using a chart to monitor the variability of the process. The discussion he gives, while being correct, provides some justification for the use of a separate control chart for monitoring the standard deviation. When there is a change in the mean only, it is more likely the $\overline{X}$-chart will signal a change in the process than either the R- or S-chart. These results are shown in Chapter 4. Also if there is only an increase in the standard deviation it is shown in Chapter 4 both the R- and S-charts are more likely to signal a change in the process than the $\overline{X}$-chart.

The concept of a process being in-control is discussed by Shewhart (1931) in terms of "natural" and "assignable" causes of variability in the process. Basically, a natural cause of variability is variability in the quality of the product designed or built into the process. Assignable causes of variability are (theoretically) removable without having to redesign the production process. In this research, it is assumed the causes of variability in the process are reflected in the mean, $\mu$, and standard deviation, $\sigma$, of the quality measure, X. When only natural causes of variability are present, the values of these parameters are taken to be $\mu = \mu_0$ and $\sigma = \sigma_0$ and the process is considered to be in-control. A process not in-control is referred to as an out-of-control process. Out-of-control can take place by either a shift in the mean or a shift in the standard deviation or both.

It is well known a shift in the standard deviation will have an effect on the distributions of the sample mean, range, and standard deviation. A shift in the mean will only effect the distribution of the sample mean. It is convenient to denote shifts in the mean relative to $\mu_0$ and $\sigma_0$. These values are defined as $\delta = (\mu - \mu_0)/(\sigma_0/\sqrt{n})$ and $\lambda = \sigma/\sigma_0$. The parameter $\delta$ is the change in the mean from $\mu_0$ measured in units equivalent to $\sigma_0/\sqrt{n}$, where n is the size of the sample in which the sample statistics, $\overline{X}$, R, and S are based. The parameter $\lambda$ represents change in the standard deviation, $\sigma$, relative to the in-control standard deviation, $\sigma_0$. In-control now corresponds to $\delta = 0$ and $\lambda = 1$. The process is out-of-control if $\delta \neq 0$ or $\lambda \neq 1$.

## 2.2 $\overline{X}$-, R-, and S- Charts

Again consider the quality measure X taken on the output of a repetitive production process. Information about the process is in the form of periodically observed independent random samples, $X_{t,1}, X_{t,2}, \ldots, X_{t,n}$ each of size n, where t = 1, 2, 3, ... . For each sample, the sample mean, $\overline{X}_t$, is computed and plotted against the sample number t, with t = 1, 2, 3, ... . The sample mean is defined by

$$\overline{X}_t = \frac{1}{n}\sum_{i=1}^{n} X_{t,i}\,.$$

Under the assumption sampling is from a normal distribution, the sample mean has a normal distribution with mean, $\mu$, and standard deviation, $\sigma/\sqrt{n}$. Even if the assumption of a normal population does not hold and n is "large," the distribution of the sample mean is approximately normal. This follows from the Central Limit Theorem. Burr (1967) found little difference in the properties of control charts based on the usually normal theory and their actual properties for various non-normal distributions. Schilling and Nelson (1976) studied control charts when the population was non-normal. For samples sizes of four or five, they found control charts based on normal theory perform about the same under a normal population as under various non-normal distributions.

If $X_{t,1}, X_{t,2}, \ldots, X_{t,n}$ are independent and identically distributed random variables (a random sample) from a normal distribution with mean, $\mu$, and standard deviation, $\sigma$, it follows

$$P[\mu_0 + a\frac{\sigma_0}{\sqrt{n}} < \bar{X}_t < \mu_0 + b\frac{\sigma_0}{\sqrt{n}}] = \Phi\left(\frac{b-\delta}{\lambda}\right) - \Phi\left(\frac{a-\delta}{\lambda}\right) \qquad (2.1)$$

where $\Phi(.)$ is the cumulative distribution function of a standard normal distribution. The value $\Phi\left(\frac{b-\delta}{\lambda}\right) - \Phi\left(\frac{a-\delta}{\lambda}\right)$ is the probability the sample mean falls in the interval $\left(\mu_0 + a\frac{\sigma_0}{\sqrt{n}}, \mu_0 + b\frac{\sigma_0}{\sqrt{n}}\right)$. Thus $1 - \Phi\left(\frac{b-\delta}{\lambda}\right) + \Phi\left(\frac{a-\delta}{\lambda}\right)$ is the probability the sample mean will fall outside this interval. As an example, for $\delta = 0$, $\lambda = 1$, $a = -3$, $b = +3$

$$\Phi(+3) - \Phi(-3) = 0.99865 - 0.00135 = 0.99730.$$

Thus, the probability the mean of a sample taken from an in-control process falls more than three standard errors, $\sigma/\sqrt{n}$, from the in-control mean, $\mu_0$, is 1-0.99730 = 0.0027. Hence, it is not very probable the sample mean will be more than three standard errors from the in-control mean and if this occurs this is taken as evidence against the hypothesis of an in-control process. With this in mind, Shewhart (1931) recommended placing, what he referred to as action lines or control limits, on the plot of the sample means against the sample number. The action lines are the horizontal lines drawn at the points, $\mu_0 - 3(\sigma_0/\sqrt{n})$ and $\mu_0 + 3(\sigma_0/\sqrt{n})$, on the vertical axis. As their name implies, a value of the sample mean falling outside these limits is a signal some corrective action to the process may need to be taken. Also, a horizontal center line is usually drawn at the point $\mu_0$ on the vertical axis.

The value b = -a = 3 is often used because Shewhart (1931) used it in various illustrations. This can be compared to the use of 5% for a level of significance following from comments made by Fisher (1932). It is common among the British to use b = -a = 3.09, since this would cause a signal on the average about one in five hundred times if the process is in-control. Control limits such as $\mu_0 - 3(\sigma_0/\sqrt{n})$ and $\mu_0 + 3(\sigma_0/\sqrt{n})$, are often referred to, respectively, as the lower control limit (LCL) and upper control limit (UCL).

Similarly the sample range, $R_t$, or the sample standard deviation, $S_t$, is also computed and plotted. The sample range is computed using

$$R_t = \max\{X_{t,1}, X_{t,2}, \ldots, X_{t,n}\} - \min\{X_{t,1}, X_{t,2}, \ldots, X_{t,n}\}$$

The standardize range, $W_t$, is defined to be $R_t / \sigma$. The distribution of the standardized range is described by Harter (1960). He gives tables of percentage points and moments for the standardize range based on samples from a normal population. A FORTRAN program to compute the cumulative distribution of the standardized range is given by Barnard (1978). These tables can be used to obtain the percentage points and moments of the distribution of any sample range based on a sample from a normal distribution as a function of the standard deviation, $\sigma$. Since $R_t = \sigma W_t$, then $E[R_t] = \sigma E[W_t]$, $V[R_t] = \sigma^2 V[W_t]$ and $R_\alpha = \sigma W_\alpha$, where $R_\alpha$ and $W_\alpha$ are the $\alpha^{th}$ percentage points, respectively, of the range and standardize range distributions. Commonly the notation $d_2 = E[W_t]$ and

$d_3^2 = V[W_t]$ is used. Harter (1960) gives the values for $d_2$ and $d_3^2$ for n = 1(1)50. The center line and the lower and upper control limits as recommended in Montgomery (1991) are

$$\begin{aligned} LCL &= (d_2 - 3d_3)\,\sigma_0 \\ CL &= d_2\sigma_0 \\ UCL &= (d_2 + 3d_3)\,\sigma_0, \end{aligned}$$

where LCL is chosen to be zero if $d_2 - 3d_3$ is less than zero. The lower and upper control limits can be chosen so that $P[R_t < LCL] = \alpha_1$ and $P[R_t > UCL] = \alpha_2$ with $\alpha_1 + \alpha_2 = \alpha$. Thus, LCL is $R_{\alpha_1}$ and the UCL is $R_{1-\alpha_2}$.

Suppose a sample size of n = 5 is to be used with $\sigma_0 = 2$. From Table 2 in Harter (1960), we find $d_2 = 2.3259239473$ and $d_3^2 = 0.7466376009$. Using the FORTRAN program given in Barnard (1978) (modified to perform calculations in double precision), the 2.5 and 97.5 percentage points of a standardized range were found to be $R_{0.025} = 0.84967$ and $R_{0.975} = 4.19703$. Thus, the center line and control limits for this chart are given by LCL = 3.1835, CL = 4.6518, and UCL = 11.9050 (rounded to four decimal places).

The sample standard deviation is defined by

$$S_t = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{t,i} - \overline{X}_t\right)^2}{n-1}}\,.$$

The sample standard deviation is a biased estimator of $\sigma$ with $E[S_t] = c_4\sigma$, where $c_4$ is a function of n and is given by

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

The standard deviation of the sample standard deviation, $S_t$, is given by $\sqrt{1-c_4^2}\,\sigma$. A Shewhart chart based on the sample standard deviation is defined by

$$\begin{aligned} LCL &= \left(c_4 - 3\sqrt{1-c_4^2}\right)\sigma_0 \\ CL &= c_4\sigma_0 \\ UCL &= \left(c_4 + 3\sqrt{1-c_4^2}\right)\sigma_0 \end{aligned}$$

An S-chart based on probability limits would have LCL as $S_{\alpha_1}$ and the UCL is $S_{1-\alpha_2}$ , where

$$S_\alpha = \sqrt{\frac{\chi^2_{n-1,\alpha}}{n-1}}\ \sigma_0$$

with $\chi^2_{n-1,\alpha}$ the $\alpha^{th}$ percentage point of a chi square distribution with n-1 degrees of freedom. Under the assumption sampling is from a normal distribution, it is a well-known fact the sample mean, $\overline{X}_t$, and sample standard deviation, $S_t$, are independent (see Bain and Engelhardt (1992)). A proof is given in Burroughs (1993) of the independence of the sample mean and range based on a random sample from a normal distribution.

Although little efficiency is lost by using the sample range as an estimator of the variability instead of the sample standard deviation of small sample sizes, this efficiency becomes more noticeable for $n \geq 10$. To see this, we first consider the two unbiased estimators, $R/d_2$ and $S/c_4$, of $\sigma$. The variances of these two estimators are given by

$V[R/d_2] = \left(d_3^2/d_2^2\right)\sigma^2$ and $V[S/c_4] = [\left(1-c_4^2\right)/c_4^2]\sigma^2$. For each sample size, n = 2(1)25, Table 2.1 gives the values of the constants $d_3^2/d_2^2$ and $\left(1-c_4^2\right)/c_4^2$. Further, the relative efficiency, re($R/d_2$, $S/c_4$) = V[$S/c_4$] / V[$R/d_2$], of $R/d_2$ to $S/c_4$ is also tabulated.

TABLE 2.1. Relative Efficiency of $R/d_2$ and $S/c_4$.

| n | $d_3^2/d_2^2$ | $\left(1-c_4^2\right)/c_4^2$ | re($R/d_2$, $S/c_4$) |
|---|---|---|---|
| 2 | 0.570796 | 0.570796 | 1.000000 |
| 3 | 0.275482 | 0.273240 | 0.991860 |
| 4 | 0.182628 | 0.178097 | 0.975189 |
| 5 | 0.138012 | 0.131768 | 0.954761 |
| 6 | 0.111964 | 0.104466 | 0.933035 |
| 7 | 0.094924 | 0.086498 | 0.911231 |
| 8 | 0.082911 | 0.073787 | 0.889947 |
| 9 | 0.073982 | 0.064324 | 0.869463 |
| 10 | 0.067077 | 0.057009 | 0.849897 |
| 11 | 0.061573 | 0.051185 | 0.831280 |
| 12 | 0.057078 | 0.046439 | 0.813597 |
| 13 | 0.053334 | 0.042497 | 0.796810 |
| 14 | 0.050164 | 0.039172 | 0.780872 |
| 15 | 0.047443 | 0.036328 | 0.765731 |
| 16 | 0.045079 | 0.033870 | 0.751333 |
| 17 | 0.043006 | 0.031723 | 0.737629 |
| 18 | 0.041171 | 0.029831 | 0.724570 |
| 19 | 0.039534 | 0.028153 | 0.712111 |
| 20 | 0.038064 | 0.026653 | 0.700211 |
| 21 | 0.036735 | 0.025304 | 0.688832 |
| 22 | 0.035528 | 0.024086 | 0.677938 |
| 23 | 0.034426 | 0.022980 | 0.667497 |
| 24 | 0.033416 | 0.021970 | 0.657480 |
| 25 | 0.032485 | 0.021046 | 0.647860 |

## 2.3 Estimators for $\mu_0$ and $\sigma_0$

If target values for $\mu_0$ and $\sigma_0$ are not given for a process, they must be estimated. The usual procedure is to base these estimates on a preliminary set of samples believed to be taken from a process when only natural causes of variability are present.

Consider the m independent random samples each of size, n, $X_{t,1}, X_{t,2}, \ldots, X_{t,n}$ from a normal distribution with mean, $\mu_0$ and standard deviation, $\sigma_0$, t = 1, 2, ... , m. The estimator commonly recommended in the literature for $\mu_0$ is

$$\bar{\bar{X}} = \frac{1}{m}\sum_{i=1}^{m}\bar{X}_t \, .$$

This estimator is an unbiased estimator of $\mu_0$ with standard error given by $\sigma_0/\sqrt{mn}$. Three other possible estimators for $\mu_0$ are the average of the sample medians, midranges, and trimmed means.

Possible estimators for $\mu_0$ are

$$\bar{R} = \frac{1}{m}\sum_{i=1}^{m} R_t \text{ (average of the sample ranges);}$$

$$\bar{S} = \frac{1}{m}\sum_{i=1}^{m} S_t \text{ (average of the sample standard deviations);}$$

$$\bar{V}^{1/2} = \sqrt{\frac{1}{m}\sum_{i=1}^{m} S_t^2} \text{ (root of the average of the sample variances); and}$$

$$\overline{MR} = \frac{1}{m}\sum_{t=1}^{m}\frac{\sum_{j=2}^{n}\left|X_{t,j-1} - X_{t,j}\right|}{n-1} \text{ (average of the moving ranges)}$$

Each of these estimators are biased estimators of $\sigma_0$. Dividing each by the appropriate function of m and n yields unbiased estimators of $\sigma_0$. These estimators are $\bar{R}/d_2$, $\bar{S}/c_4$, $\bar{V}^{1/2}/c_{4,m(n-1)}$, and $\overline{MR}/d_2$, where $d_2$ and $c_4$ are described in Section 2.2 and

$$c_{4,m(n-1)} = \sqrt{\frac{2}{m(n-1)}}\ \frac{\Gamma\left(\frac{m(n-1)+1}{2}\right)}{\Gamma\left(\frac{m(n-1)}{2}\right)}.$$

Burroughs (1993) shows under the assumptions of normality and independent random samples,

$$V[\bar{V}^{1/2}/c_{4,m(n-1)}] \le V[\bar{S}/c_4] \le V[\overline{MR}/d_2] \le V[\bar{R}/d_2].$$

Hence, the estimator, $\bar{V}^{1/2}/c_{4,m(n-1)}$, is the most efficient estimator of $\sigma_0$ under these assumptions.

## 2.4 Supplementary Runs Rules

In order to make a Shewhart chart more sensitive to small shifts in the mean, runs rules have been recommended. A runs rule causes the chart to signal if k out of the last m plotted statistics, Y, fall in the interval between E[Y] + a $\sqrt{V[Y]}$ and E[Y] + b $\sqrt{V[Y]}$, a < b. Runs rules suggested for the $\bar{X}$-chart are given in the Western Electric Handbook (1956). These rules are listed in Table 2.4.1. As with the $\bar{X}$-chart, supplementary runs rules can be used with the R-chart and the S-chart. Runs rules for the R-chart, see Table 2.4.2, are suggested in the Western Electric Handbook (1956). Note these rules are also expressed in terms of the standardized values of the plotted statistic. For example, the rule T(2,2,+2,+3) used with the R-chart is interpreted to mean a signal is given if 2 out of the last 2 sample ranges were between +2 and +3 standard errors above the in-control mean of the sampling distribution of the range. These rules or rules used with the $\overline{X}$ - chart may be applied to the S-chart. Champ, Lowry, and Woodall (1992) suggest some alternate rules for both the R- and S- charts.

Using rules to detect assignable causes of variability are discussed in the Western Electric Handbook (1956). Nelson (1984,1985) gives discussions on the use of runs rules for detecting an out-of-control process.

TABLE 2.4.1. Suggested Runs Rules for the $\overline{X}$ -chart

| | |
|---|---|
| 1. T(1,1,+3,+∞) | 5. T(8,8,-3,0) |
| 2. T(2,2,+2,+3) | 6. T(4,5,-3,-1) |
| 3. T(4,5,+1,+3) | 7. T(2,3,-3,-2) |
| 4. T(8,8,0,+3) | 8. T(1,1,-∞,3) |

TABLE 2.4.2. Suggested Runs Rules for the R-Chart

| | |
|---|---|
| 1. T(1,1,+3,+∞) | 5. T(10,10,-3,0) |
| 2. T(2,2,+2,+3) | 6. T(6,6,-3,-1) |
| 3. T(3,3,+1,+3) | 7. T(4,4,-3,-2) |
| 4. T(7,7,0,+3) | 8. T(1,1,-∞,-3) |

## 2.5 Other Simultaneous Mean and Variability Charts

As pointed out by Jennett and Welch (1939), a change in either the mean or the standard deviation of a process would alter the proportion defective articles being produced when goods are being manufactured to a specification. They suggest a monitoring scheme which plots the statistic $(U - \bar{X})/S$. The value, U, is an upper tolerance limit for the quality measure, X. If ease of calculations is needed, they suggest the scheme based on the statistic $(U - \bar{X})/W$, where W is an easily calculated estimator of the standard deviation such as the sample range.

Page (1955) considered using a single chart when changes in the standard deviation are relatively rare or unimportant. He suggested a set of runs rules for the $\bar{X}$-chart each of which was designed to detect a shift in the mean or variance. The runs rules he considered cause a signal if

(1) $T(1,1,-\infty,-b)$ or $T(1,1,b,\infty)$; or

(2) $T(m,m,-b,-a)$ or $T(m,m,a,b)$, $a < b$; or

(3) two out of the last m plotted statistics fall outside opposite warning lines.

The runs rules in (1) and (2) are used to detect an increase or decrease in the mean and the one in (3) would be useful in detecting an increase in the standard deviation. Page (1955) further suggested applying a runs rule of the form $T(m,m,-a,a)$ for detecting a decrease in the process standard deviation. He did not consider this rule in his analysis of the $\bar{X}$-chart supplemented with runs rules.

Chengalur, Arnold, and Reynolds (1989) investigated two control charting procedures monitoring both the mean and variance or standard deviation of a normal (population) distribution. One was the use of separate charts. They considered the $\bar{X}$-chart to monitor the mean and the $S^2$- chart to monitor the variance. The second procedure was based on the statistic proposed by Reynolds and Ghosh (1981). In each case the charts are modified to use variable sampling intervals. As they pointed out, the problem of monitoring both $\mu$

and $\sigma^2$ is analogous to the problem of simultaneously monitoring both the mean and variance. The hypothesis to be tested are

$$H_0: \mu = \mu_0 \text{ and } \sigma^2 = \sigma_0^2$$
$$H_a: \mu \neq \mu_0 \text{ or } \sigma^2 \neq \sigma_0^2 \tag{2.4.1}$$

or

$$H_0: \mu = \mu_0 \text{ and } \sigma^2 = \sigma_0^2$$
$$H_a: \mu \neq \mu_0 \text{ or } \sigma^2 > \sigma_0^2 \tag{2.4.2}$$

Their paper concentrated on the hypothesis (2.4.2).

The test statistic used to test $H_0: \mu = \mu_0$ and $\sigma^2 = \sigma_0^2$ verses $H_a: \mu \neq \mu_0$ or $\sigma^2 > \sigma_0^2$ was proposed by Reynolds and Ghosh (1981). It is based on the squared standardized deviations of the observations from the target value $\mu_0$ and has the form

$$Y_t = \sum_{j=1}^{n} \left( \frac{X_{t,j} - \mu_0}{\sigma_0} \right)^2$$

where t is the sample number with t = 1,2,3, ... . The null distribution of $Y_t$ is a chi square distribution with n degrees of freedom. When $\mu = \mu_0$ the test rejecting $H_0$ when $Y_t \geq \chi^2_{n,1-\alpha}$ is uniformly most powerful for detecting increases in $\sigma^2$. The value $\chi^2_{n,1-\alpha}$ is the (1 - $\alpha^{th}$) percentage point of a chi square distribution with n degrees of freedom. In general, this test is quite good at detecting changes in $\sigma^2$.

The expected value of $Y_t$ is n if the null hypothesis holds and is given by

$$E[Y_t] = \frac{n\sigma^2}{\sigma_0^2} + \left( \frac{\mu - \mu_0}{\sigma_0 / \sqrt{n}} \right)^2$$

if the alternative hypothesis holds. Under the assumption $H_0$ holds, the variance of $Y_t$ is 2n. Hence, the control limits for this charting procedure are

$$\begin{aligned} LCL &= \chi^2_{n,\alpha_1} \\ CL &= n \\ UCL &= \chi^2_{n,1-\alpha_2} \end{aligned}$$

White and Schroeder (1987) proposed a simultaneous control chart using resistant measures and a modified box plot display. This chart while being used to control the process mean and variability also displays information about the distribution and specifications of the quality measure. One can view this charting procedure as a combined median, $\tilde{X}$, chart and a interquartile range, IQR = $Q_3$ - $Q_1$, chart, where $Q_1$ and $Q_3$ are, respectively, the lower and upper sample quartiles. White and Schroder (1987) refer to IQR as the Q-spread. Based on the assumption the population of quality measures has a normal distribution, the control limits for the $\tilde{X}$-chart are given by

$$\begin{aligned} LCL &= \tilde{X}_{n,\alpha_1-r_1} \\ CL &= E[\tilde{X}] \\ UCL &= \tilde{X}_{n,1-r_1} \end{aligned}$$

where $\tilde{X}_{n,\gamma}$ is the $\gamma^{th}$ percentage point of the distribution of the median. The control limits for the IQR-chart are given by

$$\begin{aligned} LCL &= IQR_{n,\alpha_2-r_2} \\ CL &= E[IQR] \\ UCL &= IQR_{n,1-r_2} \end{aligned}$$

where $IQR_{n,\gamma}$ is the $\gamma^{th}$ percentage point of the distribution of the inner quartile range. Teichroew (1956) tabulates values for expected values of order statistics and the products of order statistics from a standard normal distribution. These tabulations can be used to determine numeric values for the expectations and variances in the control limit formulas for the $\tilde{X}$-chart and the IQR-chart . For the lower and upper percentage points for both charts, White and Schroeder (1987) chose respectively, -3 and +3.

Since the distribution of the sample mean is a function of the standard deviation as well as the mean, the performance of the $\overline{X}$-chart is affected by a shift in the standard deviation as well as the mean. When a combined charting procedure gives a signal it would be useful if the chart causing the signal was also an indication of which parameter, $\mu$ or $\sigma$, had shifted. With this in mind, we propose separate charts based on the statistics, $T_t = \left(\overline{X}_t - \mu_0\right) / \left(S_t / \sqrt{n}\right)$ and $S_t$. In general, the random variable, $T_t$, has a noncentral t-distribution with n-1 degrees of freedom and noncentrality parameter, $\delta/\lambda$. The distribution of $T_t$ has a central t-distribution if $\mu = \mu_0 (\delta = 0)$. A Shewhart chart based on $T_t$ is defined by

$$\begin{aligned} LCL &= t_{\alpha_2}; \\ CL &= 0 \\ UCL &= t_{1-\alpha_1} \end{aligned}$$

where $t_\alpha$ is the $\gamma^{th}$ percentage point of a central t-distribution with n-1 degrees of freedom.

As stated previously, the chart based on $S_t$ is not affected by a change in the mean. The chart based on $T_t$ is not affected by a change in $\sigma$ if there is no change in $\mu$. It is affected, though, if there is a change in both $\mu$ and $\sigma$ but only through the ratio $\delta/\lambda$. A relative large increase in the standard deviation relative to the shift in the mean would have only a small affect on the distribution of the statistic, $T_t$. On the other hand, this distribution would be affected by a relative large increase or decrease in the mean; or if the standard deviation decreases. A decrease in the standard deviation, although classified as an out-of-control condition, is associated with process improvement and needs to be detected. This condition would affect both the distributions of $T_t$ and $S_t$ possibility leading to a misinterpretation of a signal. It should be noted $T_t$ and $S_t$ are not independent.

In this section, we have reviewed some of the other proposed ways of monitoring the mean and variance simultaneously. In practice, as stated by Page (1955) and Chengular, Arnold, and Reynolds (1989), the typical approach is to use separate charts for each parameter and then to signal if either chart signals. Although this seems to be a reasonable approach, it may not be optimal for detecting certain types of process changes such as those involving a change in both $\mu$ and $\sigma$. The type of rectifying action to be taken and the changes likely to occur are important to consider before setting up any of these schemes. If a different rectifying action must be taken for a change in the mean as for a change in the standard deviation, Page (1955) suggested the control charting procedure should indicate the type of change in the process and not just its existence. In this research, we concentrate on analyzing the performance of separate charts to control both the mean and standard deviation of a normal distribution. Since the charts commonly used to control the mean and standard deviation are, respectively, the $\bar{X}$-chart and R- (or S-) chart, we analyze these two charts as a simultaneous control charting procedure. This research will provide an analysis of what is commonly done in practice and provide a foundation for further comparison of simultaneous control charting procedures.

# Chapter 3. EVALUATING THE RUN LENGTH DISTRIBUTION

## 3.1 Introduction

The run length of a quality control chart is the sample number in which the chart first gives a signal. Woodall and Ncube (1985) define the run length N of a multidimensional charting procedure as the minimum of $N_1, \ldots, N_p$, where $N_j$ is the run length of the $j^{th}$ chart monitoring the $j^{th}$ component of the mean. Similarly, Champ (1986) defined the run length for the combined $\overline{X}$- and R- (or S-) charts as $N = \min\{N_1, N_2\}$ where $N_1$ and $N_2$ are the run lengths of the $\overline{X}$- and R- (or S-) charts, respectively. Gan (1989) investigates the combined cumulative sum (CUSUM) mean chart and Shewhart variance chart, and the combined exponentially weighted moving average (EWMA) chart and Shewhart variance chart. He also defines the run length of these combined schemes as the minimum of the run lengths of the individual charts.

Various techniques have been used to evaluate the run length distribution of a control chart. Among these are the integral equation, Markov chain, and simulation methods. A procedure was developed by Champ and Woodall (1987) for representing a Shewhart chart supplemented with runs rules as a Markov chain. In this chapter, Champ and Woodall's (1987) method is used to develop procedures for evaluating the run length distribution of the combined $\overline{X}$- and R- (or S-) charts. The first of these methods uses Champ and Woodall's (1987) method to determine the run length distribution of the minimum run length from the run length distribution of the individual $\overline{X}$- and R- (or S-) charts. A modified version of Champ and Woodall's (1990) FORTRAN program was used to determine the run length distribution of the individual charts. The second method determines the Markov chain representation of the combined chart from the Markov chain representation of the individual charts found in Champ and Woodall (1987). The FORTRAN program of Champ and Woodall (1990) was modified to obtain the Markov chain representation of the combined chart.

### 3.2 Combined Run Length Distributions Approach

Champ and Woodall (1987) gave a procedure for representing a Shewhart chart supplemented with runs rules as a Markov chain. The FORTRAN program by Champ and Woodall (1990) calculates the average run length (ARL) for a Shewhart $\bar{X}$-chart supplemented with runs rules for various standardized shifts in the mean. A modified version of this program was used by Lowry, Champ, and Woodall (1994) to calculate the ARL and percentage points of both the R- and S-charts supplemented with runs rules. The program by Champ and Woodall (1990) can be modified to obtain the cumulative distribution function (CDF) of the run length of the $\bar{X}$-, R-, or S-charts supplemented with runs rules.

Let $N_1$ and $N_2$, respectively, denote the run lengths of the $\bar{X}$-chart and the R-chart (or S-chart). Further, denote the CDFs of $N_1$ and $N_2$, respectively, by $F_1(t)$ and $F_2(t)$. The CDF of the run length, $N = \min\{N_1, N_2\}$, of the combined $\bar{X}$- and R- (or S-) chart is determined by

$$\begin{aligned} F_N(t) &= P[N \le t] \\ &= 1 - P[N > t] \\ &= 1 - P[N_1 > t,\ N_2 > t] \\ &= 1 - P[N_1 > t]\,P[N_2 > t] \\ &= 1 - \{1 - P[N_1 \le t]\}\{1 - P[N_2 \le t]\} \\ &= 1 - [1 - F_1(t)][1 - F_2(t)], \end{aligned} \tag{3.2.1}$$

t=1, 2, 3, ... . The independence of $N_1$ and $N_2$ follows since $\bar{X}_t$ and $R_t$ (or $S_t$) are independent, t=1, 2, ... (see Chapter 2, Section 2.2).

It is convenient to define the function, $\bar{F}(\,)$, by $\bar{F}(t)$=1 - F(t), where F(t) is any CDF. Using this notation, we can express equation (3.2.1) as

$$F_N(t) = 1 - \bar{F}_1(t)\,\bar{F}_2(t)$$

or

$$\bar{F}_N(t) = \bar{F}_1(t)\,\bar{F}_2(t), \tag{3.2.2}$$

t=1, 2, ... . It follows from equations (3.2.1) and (3.2.2) the probability distribution function (pdf) of N can be expressed by

$$P(N = t) = [1 - F_1(t\text{-}1)][1 - F_2(t\text{-}1)] - [1 - F_1(t)][1 - F_2(t)] \tag{3.2.3}$$

or

$$P(N = t) = \bar{F}_N(t\text{-}1) - \bar{F}_N(t), \tag{3.2.4}$$

for t=1,2, ... , where we define $F_1(0)$=$F_2(0)$=0.

As a special case, consider the basic Shewhart $\bar{X}$- and R- (or S-) charts, that is, each chart is defined by the set of runs rules {T(1,1,-∞,-$b_L$), T(1,1,+$b_U$,+∞)}, where $b_L$ and $b_U$ are positive real numbers. It is well-known for this situation the run length, $N_i$, follows a geometric distribution with parameter $p_i$, the probability the chart signals at any sampling stage. The CDF of the run length $N_i$, as given in Bain and Engelhardt (1992), is $F_{N_i}(t) = 1 - q_i^t$, where $q_i$=1 - $p_i$, i=1,2. From equation (3.2.1), we see the CDF of the minimum is $F_N(t) = 1 - (q_1 q_2)^t$. It follows the run length, N, of the combined chart follows a geometric distribution with parameter $p = 1 - q_1 q_2$.

Various parameters of the run length are of interest. These include, among others, the mean, standard deviation, and percentage points of the run length. The mean is often referred to as the average run length (ARL). These parameters of the run length distribution of the combined Shewhart $\overline{X}$ - and R- (or S-) charts are given by

$$\mu_N = E[N] = \frac{1}{1 - q_1 q_2}, \tag{3.2.5}$$

$$\sigma_N = \frac{\sqrt{q_1 q_2}}{1 - q_1 q_2}, \text{ and} \tag{3.2.6}$$

$$N_\alpha = \left\lceil \frac{\log(1-\alpha)}{\log(q_1 q_2)} \right\rceil, \tag{3.2.7}$$

where $N_\alpha$ is the $\alpha^{th}$ percentage point and $\lceil\ \rceil$ is the ceiling function. A discussion of the derivation of the mean and standard deviation of a geometric distribution is given in Bain and Engelhardt (1992). The $\alpha^{th}$ percentage point is defined to be the smallest integer, $N_\alpha$, such that $P[N \le N_\alpha] \ge \alpha$. It follows $P[N \le N_\alpha] = 1 - P[N > N_\alpha] = 1 - (q_1 q_2)^{N_\alpha} = \alpha$. Solving this equation for the integer, $N_\alpha$, under the given condition it is the smallest integer such that $P[N \le N_\alpha] \ge \alpha$ yields equation (3.2.7).

These parameters of the minimum run length can be expressed in terms of the average run lengths, $E[N_1]$ and $E[N_2]$. Noting $E[N_i]=1/(1-q_i)$, for i=1, 2, these expressions are

$$\mu_N = E[N] = \frac{E[N_1]E[N_2]}{E[N_1] + E[N_2] - 1}, \tag{3.2.8}$$

$$\sigma_N = \sqrt{\frac{(E[N_1]-1)(E[N_2]-1)}{E[N_1] + E[N_2] - 1}}, \tag{3.2.9}$$

and

$$N_\alpha = \left\lceil \frac{\log(1-\alpha)}{\log\left(\frac{(E[N_1]-1)(E[N_2]-1)}{E[N_1]E[N_2]}\right)} \right\rceil, \tag{3.2.10}$$

If supplementary runs rules are applied individually to the $\overline{X}$-chart and the R-chart (or S-chart), the Markov chain approach of Champ and Woodall (1987) can be used to determine the run length properties of $N_1$ and $N_2$. Hence, the run length properties of N can be determined. As stated previously, the FORTRAN program of Champ and Woodall (1990) can be modified to obtain the CDF's of $N_1$ and $N_2$ for any value of t, t=1, 2, 3, ... . The CDF of N can then be determined for each t using either equation (3.2.1) or (3.2.2).

Although the values of the CDFs of $N_1$ and $N_2$ for each t can be obtained theoretically using the Markov chain approach, it may not be practical to obtain their values for large values of t. Woodall (1983) showed as t gets larger the tail probabilities, $P(N_i = t)$, i=1, 2, can be approximated by a geometric distribution. For large t, say for t greater than some value $t_i^*$ he shows there exists a $\lambda_i$ such that this approximation takes the form

$$P(N_i = t_i^* + k) \approx \lambda_i^k \, P(N_i = t_i^*), \tag{3.2.11}$$

i=1, 2 and k=1, 2, ... . Woodall (1983) gave two approximation for $\lambda_i$,

$$\hat{\lambda}_i = \frac{P\left(N_i = t_i^*\right)}{P\left(N_i = t_i^* - 1\right)} \tag{3.2.12}$$

and

$$\lambda_i' = \frac{1 - \sum_{t=1}^{t_i^*} P\left(N_i = t\right)}{1 - \sum_{t=1}^{t_i^*-1} P\left(N_i = t\right)} \tag{3.2.13}$$

Woodall (1983) suggested using the values of $\hat{\lambda}_i$ and $\lambda_i'$ to determine the value of $t_i^*$. For each t, determine the approximations $\hat{\lambda}_i$ and $\lambda_i'$ for $\lambda_i$ and choose $t_i^*$=t if these approximations are close, say $|\hat{\lambda}_i - \lambda'| < \varepsilon$ for $\varepsilon$ small. The values of $\hat{\lambda}_i$, $\lambda_i'$ and $t_i^*$, can be determined with a modification to the FORTRAN program of Champ and Woodall (1990) .

The CDF of the run length, $N_i$, now can be determined exactly for values of $1 \le t \le t_i^*$. For the values of $t=t_i^* + k$ the CDF of the run length, $N_i$, using equation (3.2.11) can be approximated by

$$\begin{aligned}
F_i\left(t_i^* + k\right) &= P[N_i \le t_i^* + k] \\
&= 1 - P[N_i > t_i^* + k] \\
&= 1 - \sum_{j=1}^{\infty} P[N_i = t_i^* + k + j] \\
&\approx 1 - \sum_{j=1}^{\infty} \lambda_i^{k+j} P\left(N_i = t_i^*\right) \\
&= 1 - \frac{\lambda_i^{k+1}}{1-\lambda_i} P[N_i = t_i^*] \\
&= \frac{1 - \lambda_i - \lambda_i^{k+1} P[N_i = t_i^*]}{1 - \lambda_i}
\end{aligned}$$

or

$$\bar{F}_i\left(t_i^* + k\right) \approx \frac{\lambda_i^{k+1}}{1-\lambda_i} P[N_i = t_i^*] \tag{3.2.14}$$

k=1, 2, 3, ... . Equation (3.2.14) cannot be used directly to obtain an approximate value for $F_i\left(t_i^* + k\right)$, since $\lambda_i$ is not known. Replacing $\lambda_i$ in the right hand side of this expression with either $\hat{\lambda}_i$ or $\lambda'_i$ yields an obtainable approximation for $F_i\left(t_i^* + k\right)$. Woodall (1983) recommends the use of $\lambda'_i$.

Wheeler (1983) evaluated the CDFs for various $\bar{X}$-charts supplemented with runs rules and for selected shifts in the mean of the process. He derived some closed formed expressions and used these expressions along with simulation to calculate tables of the CDFs for t=1(1)10. Using this same method, Coleman (1986) evaluated these charts and found an error in Wheeler's (1983) derivation of the CDF of the chart defined by the runs rules {T(1,1,3,∞), T(2,3,2,3), T(4,5,1,3)}.

For large ARLs, Woodall (1983) showed that

$$E\left(N_i\right) \approx \sum_{t=1}^{t_i^*} t\, P[N_i = t] + \lambda_i P[N_i = t_i^*][\frac{t_i^*}{1-\lambda_i} + \frac{1}{\left(1-\lambda_i\right)^2}] \tag{3.2.15}$$

and

$$E\left(N_i^2\right) \approx \sum_{t=1}^{t_i^*} t^2 P[N_i = t] + \lambda_i P[N_i = t_i^*][\frac{\left(t_i^*\right)^2}{1-\lambda_i} + \frac{2t_i^* - 1}{\left(1-\lambda_i\right)^2} + \frac{2}{\left(1-\lambda_i\right)^3}]. \tag{3.2.16}$$

Further he showed the $\alpha^{th}$ percentage point, $N_\alpha$, can be approximated by

$$N_{i,\alpha} \approx \left[ t_i^* - 1 + \ln[\left(1-\lambda_i\right) \frac{\sum_{t=1}^{t_i^*} P[N_i = t] - \alpha}{P[N_i = t_i^*]} + \lambda_i]/ \ln\left(\lambda_i\right) \right] \tag{3.2.17}$$

Note, the value of $N_{i,\alpha}$ should be determined exactly for $\alpha \le F_i\left(t_i^*\right)$ by searching for the value $N_{i,\alpha}$ such that $F_i\left(N_{i,\alpha} - 1\right) < \alpha$ and $F_i\left(N_{i,\alpha}\right) \ge \alpha$.

We now consider the geometric approximation of the tail probabilities for the run length distribution of the combined Shewhart $\overline{X}$- and R- (or S-) charts supplemented with runs rules. Assuming $t_1^* \le t_2^*$, the values of the CDFs of $N_1$ and $N_2$ are determined exactly using Champ and Woodall's (1987) procedure for t=1, 2, ... $t_1^*$. The CDF of N is then determined exactly by applying equation (3.2.1) for t=1, 2, ... $t_1^*$. For $t=t_i^*+1$, ... , $t_2^*$, the CDF of N can be determined approximately by

$$F_N\left(t_i^* + k\right) \approx 1 - \left(\frac{\lambda_i^{k+1}}{1-\lambda_i} P[N_i = t_i^*]\right)\left(1 - F_2\left(t_i^* + k\right)\right)$$

$$= 1 - \frac{\lambda_i^{k+1}}{1-\lambda_i} P[N_i = t_i^*]\bar{F}_2\left(t_i^* + k\right)$$

or

$$\bar{F}_N\left(t_i^* + k\right) \approx \frac{\lambda_i^{k+1}}{1-\lambda_i} P[N_i = t_i^*]\bar{F}_2\left(t_i^* + k\right) \tag{3.2.18}$$

for k=1, ... ,$t_2^*$ - $t_1^*$. For $t=t_2^*+1, t_2^*+2$, ... , or k=1, 2, 3, ... , the CDF becomes

$$F_N\left(t_2^* + k\right) \approx 1 - \left(\frac{\lambda_1^{t_2^* - t_1^* + k + 1}}{1-\lambda_1} P[N_1 = t_1^*]\right)\left(\frac{\lambda_2^{k+1}}{1-\lambda_2} P[N_2 = t_2^*]\right)$$

$$= 1 - \frac{\lambda_1^{t_2^* - t_1^* + k + 1}\lambda_2^{k+1}}{\left(1-\lambda_1\right)\left(1-\lambda_2\right)} P[N_1 = t_1^*]P[N_2 = t_2^*]$$

or

$$\bar{F}_N\left(t_2^* + k\right) \approx 1 - \frac{\lambda_1^{t_2^* - t_1^* + k + 1}\lambda_2^{k+1}}{\left(1-\lambda_1\right)\left(1-\lambda_2\right)} P[N_1 = t_1^*]P[N_2 = t_2^*]. \tag{3.2.19}$$

For the case where $t_1^* = t_2^*$, equation (3.2.19) becomes

$$F_N\left(t_2^* + k\right) \approx 1 - \frac{\left(\lambda_1 \lambda_2\right)^{k+1}}{\left(1-\lambda_1\right)\left(1-\lambda_2\right)} P[N_1 = t_1^*]P[N_2 = t_2^*],$$

or

$$\bar{F}_N\left(t_2^* + k\right) \approx \frac{\left(\lambda_1 \lambda_2\right)^{k+1}}{\left(1-\lambda_1\right)\left(1-\lambda_2\right)} P[N_1 = t_1^*]P[N_2 = t_2^*]. \tag{3.2.20}$$

For the case where $t_1^* \geq t_2^*$, $\bar{F}_N(t)$ is determined exactly using equation (3.2.2), for $t=1,2, \ldots, t_2^*$. The value of $\bar{F}_N\left(t_2^* + k\right)$ is approximated by

$$\bar{F}_N\left(t_2^* + k\right) \approx \frac{\lambda_2^{k+1}}{1-\lambda_2} \bar{F}_1\left(t_2^* + k\right) P[N_2 = t_2^*], \tag{3.2.21}$$

for $k=1, 2, \ldots, t_1^* - t_2^*$ ; and approximated by

$$\bar{F}_N\left(t_1^* + k\right) \approx \frac{\lambda_1^{k+1} \lambda_2^{t_1^* - t_2^* + k+1}}{\left(1-\lambda_1\right)\left(1-\lambda_2\right)} P[N_1 = t_1^*]P[N_2 = t_2^*], \tag{3.2.22}$$

for $k=t_1^* - t_2^*+1,\ t_1^* - t_2^*,+2, \ldots$ .

These tail approximations can now be used to obtain approximations for the mean, E[N], the variance, V[N], and the $\alpha^{th}$ percentage points, $N_\alpha$, of the combined run length distribution. The approximation to E[N] is derived in what follows. Assuming $t_1^* \leq t_2^*$,

$$E[N] = \sum_{t=1}^{\infty} t\ P[N = t]$$

$$= \sum_{t=0}^{\infty} P[N > t]$$

$$= \sum_{t=0}^{\infty} P[N_1 > t]\, P[N_2 > t]$$

$$= \sum_{t=0}^{t_1^*} \bar{F}_1(t)\bar{F}_2(t) + \sum_{t=t_1^*+1}^{t_2^*} \bar{F}_1(t)\bar{F}_2(t) + \sum_{t=t_2^*+1}^{\infty} \bar{F}_1(t)\bar{F}_2(t)$$

$$= \sum_{t=0}^{t_1^*} \bar{F}_1(t)\bar{F}_2(t) + \sum_{t=1}^{t_2^*-t_1^*} \bar{F}_1\left(t_1^* + t\right)\bar{F}_2\left(t_1^* + t\right) + \sum_{t=1}^{\infty} \bar{F}_1\left(t_1^* + t_2^* - t_1^* + t\right)\bar{F}_2\left(t_2^* + t\right)$$

$$\approx \sum_{t=0}^{t_1^*} \bar{F}_1(t)\bar{F}_2(t) + \frac{\lambda_1 f_1\left(t_1^*\right)}{1-\lambda_1} \sum_{t=1}^{t_2^*-t_1^*} \lambda_1^t \bar{F}_2\left(t_1^* + t\right)$$

$$+ f_1\left(t_1^*\right) f_2\left(t_2^*\right) \frac{\lambda_1^{t_2^*-t_1^*+1}\lambda_2}{\left(1-\lambda_1\right)\left(1-\lambda_2\right)} \sum_{t=1}^{\infty} \left(\lambda_1\lambda_2\right)^t$$

$$= \sum_{t=0}^{t_1^*} \bar{F}_1(t)\bar{F}_2(t) + \frac{\lambda_1 f_1\left(t_1^*\right)}{1-\lambda_1} \sum_{t=1}^{t_2^*-t_1^*} \lambda_1^t \bar{F}_2\left(t_1^* + t\right)$$

$$+ f_1\left(t_1^*\right) f_2\left(t_2^*\right) \frac{\lambda_1^{t_2^*-t_1^*+1}\lambda_2}{\left(1-\lambda_1\right)\left(1-\lambda_2\right)\left(1-\lambda_1\lambda_2\right)}. \tag{3.2.23}$$

For the case where $t_1^* \geq t_2^*$, we have

$$E[N] \approx \sum_{t=0}^{t_2^*} \bar{F}_1(t)\bar{F}_2(t) + \frac{\lambda_2 f_2\left(t_2^*\right)}{1-\lambda_2} \sum_{t=1}^{t_1^*-t_2^*} \lambda_2^t \bar{F}_1\left(t_2^* + t\right)$$

$$+ f_1\left(t_1^*\right) f_2\left(t_2^*\right) \frac{\lambda_1\lambda_2^{t_1^*-t_2^*+1}}{\left(1-\lambda_1\right)\left(1-\lambda_2\right)\left(1-\lambda_1\lambda_2\right)}. \tag{3.2.24}$$

It is well-known $V[N]=E[N^2] - (E[N])^2$. From this equation, in order to determine an approximation for $V[N]$, we need only find an approximation for the value $E[N^2]$. An approximation for $E[N^2]$ is derived in what follows.

$$E[N^2] = \sum_{t=1}^{\infty} t^2 \, P[N = t]$$

$$= \sum_{t=0}^{\infty} (2t+1) \, P[N > t]$$

$$= 2\sum_{t=1}^{\infty} t \, P[N > t] + \sum_{t=0}^{\infty} P[N > t]$$

$$= 2\Bigg[\sum_{t=1}^{t_1^*} t\bar{F}_1(t)\bar{F}_2(t) + \sum_{t=t_1^*+1}^{t_2^*} t\bar{F}_1(t)\bar{F}_2(t) + \sum_{t=t_2^*+1}^{\infty} t\bar{F}_1(t)\bar{F}_2(t)\Bigg] + E[N]$$

$$= 2\Bigg[\sum_{t=1}^{t_1^*} t\bar{F}_1(t)\bar{F}_2(t) + \sum_{t=1}^{t_2^*-t_1^*} (t_1^*+t)\bar{F}_1(t_1^*+t)\bar{F}_2(t_1^*+t)$$

$$+ \sum_{t=1}^{\infty} (t_2^*+t)\bar{F}_1(t_2^*+t)\bar{F}_2(t_2^*+t)\Bigg] + E[N]$$

$$\approx 2\Bigg[\sum_{t=1}^{t_1^*} t\bar{F}_1(t)\bar{F}_2(t) + \sum_{t=1}^{t_2^*-t_1^*} (t_1^*+t)\frac{\lambda_1^{t+1}}{1-\lambda_1} f_1(t_1^*)\bar{F}_2(t_1^*+t)$$

$$+ \sum_{t=1}^{\infty} (t_2^*+t)\frac{\lambda_1^{t_2^*-t_1^*+t+1}\lambda_2^{t+1}}{(1-\lambda_1)(1-\lambda_2)} f_1(t_1^*) f_2(t_2^*)\Bigg] + E[N]$$

$$= 2\Bigg[\sum_{t=1}^{t_1^*} t\bar{F}_1(t)\bar{F}_2(t) + \frac{t_1^* f_1(t_1^*)}{1-\lambda_1}\sum_{t=1}^{t_2^*-t_1^*} \lambda_1^{t+1}\bar{F}_2(t_1^*+t)$$

$$+\frac{f_1(t_1^*)}{1-\lambda_1}\sum_{t=1}^{t_2^*-t_1^*} t\lambda_1^{t+1}\bar{F}_2(t_1^*+t)$$

$$+\frac{t_2^*\lambda_1^{t_2^*-t_1^*}}{(1-\lambda_1)(1-\lambda_2)} f_1(t_1^*)f_2(t_2^*)\sum_{t=1}^{\infty}(\lambda_1\lambda_2)^{t+1}$$

$$\left.+\frac{\lambda_1^{t_2^*-t_1^*}}{(1-\lambda_1)(1-\lambda_2)} f_1(t_1^*)f_2(t_2^*)\sum_{t=1}^{\infty} t(\lambda_1\lambda_2)^{t+1}\right] + E[N]$$

$$=2\left[\sum_{t=1}^{t_1^*} t\bar{F}_1(t)\bar{F}_2(t)+\frac{t_1^* f_1(t_1^*)}{1-\lambda_1}\sum_{t=1}^{t_2^*-t_1^*}\lambda_1^{t+1}\bar{F}_2(t_1^*+t)\right.$$

$$+\frac{f_1(t_1^*)}{1-\lambda_1}\sum_{t=1}^{t_2^*-t_1^*} t\lambda_1^{t+1}\bar{F}_2(t_1^*+t)$$

$$+\frac{t_2^*\lambda_1^{t_2^*-t_1^*}(\lambda_1\lambda_2)^2}{(1-\lambda_1)(1-\lambda_2)(1-\lambda_1\lambda_2)} f_1(t_1^*)f_2(t_2^*)$$

$$\left.+\frac{\lambda_1^{t_2^*-t_1^*}(\lambda_1\lambda_2)^2}{(1-\lambda_1)(1-\lambda_2)(1-\lambda_1\lambda_2)} f_1(t_1^*)f_2(t_2^*)\right] + E[N]. \tag{3.2.25}$$

The value of $N_\alpha$ should be determined exactly for $\alpha \le F_N(t_1^*)$ by searching for the value $N_\alpha$ such that $F_N(N_\alpha - 1) < \alpha$ and $F_N(N_\alpha) \ge \alpha$. For the case where $F_N(t_1^*) < \alpha \le F_N(t_2^*)$, a search for an approximate value $N_\alpha$ should be made using

$$F_N(t) \approx 1 - \frac{\lambda_1^{t-t_1^*+1}}{1-\lambda_1} f_1(t_1^*)\bar{F}_2(t)$$

such that $F_N\left(N_\alpha - 1\right) < \alpha$ and $F_N\left(N_\alpha\right) \geq \alpha$. For $\alpha > F_N\left(t_2^*\right)$, an approximate value of $N_\alpha$ can be obtained by solving the following system of inequalities for $N_\alpha$:

$$F_N\left(t_2^* + N_\alpha - t_2^*\right) = 1 - \frac{\lambda_1^{t_2^* - t_1^* + N_\alpha - t_2^* + 1} \lambda_2^{N_\alpha - t_2^* + 1}}{\left(1 - \lambda_1\right)\left(1 - \lambda_2\right)} f_1\left(t_1^*\right) f_2\left(t_2^*\right) \geq \alpha$$

and

$$F_N\left(t_2^* + N_\alpha - 1 - t_2^*\right) = 1 - \frac{\lambda_1^{t_2^* - t_1^* + N_\alpha - t_2^*} \lambda_2^{N_\alpha - t_2^*}}{\left(1 - \lambda_1\right)\left(1 - \lambda_2\right)} f_1\left(t_1^*\right) f_2\left(t_2^*\right) < \alpha .$$

The solution to this system is given by the expression on the right-hand side of the approximation (3.2.25). This is an approximate value for $N_\alpha$.

$$N_\alpha \approx \left\lceil t_1^* - 1 + \frac{\log\left([\left(1 - \alpha\right)\left(1 - \lambda_1\right)\left(1 - \lambda_2\right)] / [f_1\left(t_1^*\right) f_2\left(t_2^*\right)]\right)}{\log\left(\lambda_1\right)} \right\rceil . \qquad (3.2.26)$$

A FORTRAN program is given in Appendix I implementing the combined run length distribution approach for the combined $\overline{X}$ - and R-chart.

### 3.3 Combined Markov Chain Approach

In this section, we develop the Markov chain representation of the combined Shewhart $\overline{X}$ - and R- (or S-) charts supplemented with runs rules. Using Champ and Woodall's (1987) procedure for a Markov chain representation can be determined for each of the individual charts. The FORTRAN program given by Champ and Woodall (1990) outputs a matrix containing information about the regions of the chart and state, next-state transitions.

Let $i_1$ and $i_2$ represent non-absorbing states, respectively, of the Markov chain representation of the Shewhart $\overline{X}$ - and R- (or S-) charts supplemented with runs rules. Further, let $h_1$ and $h_2$ represent the respective number of non-absorbing states of the

Markov chain representations. A non-absorbing state of the combined chart can be represented by the ordered pair $(i_1, i_2)$. Let i represent the number of this state. A convenient way of numbering these states is to let $i = (h_2 - 1)\ i_1 + i_2$. Since $\overline{X}$- and R- (or S-) are independent, the probability of making a transition from the non-absorbing state $(i_1, i_2)$ to the non-absorbing state $(j_1, j_2)$ is

$$p_{i,j} = P[i_1 \rightarrow j_1]\ P[i_2 \rightarrow j_2]. \tag{3.3.1}$$

We need only the matrix, **Q**, of probabilities given in equation (3.3.1) to evaluate the run length distribution of the combined chart.

The run length, $N^{(i)}$, of the combined chart starting in the non-absorbing state i, is defined as the number of the first sampling stage in which either one or both charts signals. For convenience we define the vector of run lengths, $\underset{\sim}{\mathbf{N}}$, by

$$\underset{\sim}{\mathbf{N}} = \left[N^{(1)}, N^{(2)}, \ldots N^{(h)},\right]'$$

where $h = h_1 h_2$. Using the results found in Brook and Evans (1972), the run length distributions is given by

$$P[\underset{\sim}{\mathbf{N}} = \underset{\sim}{\mathbf{t}}] = \mathbf{Q}^{t-1}(\mathbf{I} - \mathbf{Q})\underset{\sim}{\mathbf{1}} \tag{3.3.2}$$

t=1, 2, 3, ... , where $\underset{\sim}{\mathbf{1}}$ is a column vector of ones with the same column dimension as **Q** and $\underset{\sim}{\mathbf{t}} = [t,\ t, \ldots,\ t]' = t\underset{\sim}{\mathbf{1}}$. We define the vector $\underset{\sim}{\mathbf{F}} = \left(\underset{\sim}{\mathbf{t}}\right)$ by

$$\underset{\sim}{\mathbf{F}}\left(\underset{\sim}{\mathbf{t}}\right) = [F^{(1)}(t),\ F^{(2)}(t), \ldots,\ F^{(h)}(t)]' \tag{3.3.3}$$

where $F^{(i)}(t)$ is the CDF of the run length, $N^{(i)}$ i=1, 2, ... , h. Using equation (3.3.2), the vector of CDFs can be calculated by

$$\underset{\sim}{\mathbf{F}}\left(\underset{\sim}{\mathbf{t}}\right) = \left(\mathbf{I} - \mathbf{Q}^{\mathbf{t+1}}\right)\underset{\sim}{\mathbf{1}}, \tag{3.3.4}$$

t=1, 2, 3, ... . This expression has a similar form to the expression for the CDF of a geometric distribution.

As shown in Brook and Evans (1972), the vector of ARLs can be determined by the expression

$$\mu = E[\underset{\sim}{\mathbf{N}}] = \left(\mathbf{I}\text{-}\mathbf{Q}\right)^{-1}\underset{\sim}{\mathbf{1}}. \tag{3.3.5}$$

Determining the ARLs of the combined chart using equation (3.3.5) may not be practical since the number of non-absorbing states may be large. The most efficient method of evaluating the ARL of interest is to use equation (3.3.2) and Woodall's (1983) method for approximating the tail probabilities with a geometric distribution. The FORTRAN program given by Champ and Woodall (1990) was modified to evaluate the ARLs, standard deviations of the run length, and percentage points of the run length for the combined charting procedure. The modified program is listed in Appendix II along with an explanation of how to run the program.

## 3.4 Evaluating the Run Length Distribution When No Standards Are Given

In the previous sections, methods for determining the parameters of the run length distribution assumed the in-control mean and standard deviation were known. In this case the parameters, $\mu_0$ and $\sigma_0$, are referred to as target values or standards. In many situations no standards are given and the parameters, $\mu_0$ and $\sigma_0$, are estimated from a preliminary set of observations taken from the process when it is believed to be in-control. In this section, we develop a method for analyzing the run length distribution when no standards are given.

Reynolds, Ghosh and Hui (1981) investigate the run length properties of the $\overline{X}$-chart when the process variance is estimated from initial data. They use the standard deviation of the preliminary data as their estimate of the process variance. An expression for the ARL is given in the form of an integral equation which was evaluated numerically. In their evaluation of the ARL, it was shown how using small sample sizes to estimate the process variance increases the ARL and the variance of the run length. Also for some preliminary sample sizes the ARL will be infinite.

Ng and Case (1992) evaluated the run length properties of the $\overline{X}$-chart when both the in-control mean and standard deviation are estimated from preliminary data. They assume m preliminary samples each of size n are available from an in-control process. Their estimator for the in-control mean of the process is the mean of the sample means of the m preliminary samples. The mean of the ranges of these m samples was used as an estimate of the process in-control standard deviation. They expressed the ARL as an integral which they evaluated using numerical integration.

Let $\hat{\mu}_0$ and $\hat{\sigma}_0$, respectively, be estimators of $\mu_0$ and $\sigma_0$. Further, let the joint probability density function of $\hat{\mu}_0$ and $\hat{\sigma}_0$, be given by $f_{\hat{\mu}_0,\hat{\sigma}_0}(u,v \mid \mu_0,\sigma_0)$. For the case where $\hat{\mu}_0$ and $\hat{\sigma}_0$, are independent then

$$f_{\hat{\mu}_0,\hat{\sigma}_0}(u,v \mid \mu_0,\sigma_0) = f_{\hat{\mu}_0}(u \mid \mu_0,\sigma_0) f_{\hat{\sigma}_0}(v \mid \mu_0,\sigma_0).$$

Consider any parameter, $\xi = \xi(\mu_0,\sigma_0,\mu,\sigma)$ of the run length distribution. If $\mu_0$ and $\sigma_0$ are replaced with their estimators, then $\hat{\xi} = \hat{\xi}(\hat{\mu}_0,\hat{\sigma}_0,\mu,\sigma)$ is a random variable. Defining $\hat{\kappa} = \hat{\kappa}(\hat{\mu}_0,\hat{\sigma}_0,\mu,\sigma)$ such that the transformation of $(\hat{\xi},\hat{\kappa})$ of $(\hat{\mu}_0,\hat{\sigma}_0)$ is one-to-one then the pdf $\hat{\xi}$ can be determined from

$$f_{\hat{\xi}}\left(w \mid \mu_0, \sigma_0\right) = f_{\hat{\mu}_0}\left(u \mid \mu_0, \sigma_0\right) f_{\hat{\sigma}_0}\left(v \mid \mu_0, \sigma_0\right).$$

As can be seen, any parameter of the distribution of the random variable $\hat{\xi}$ is a parameter of the distribution of the run length.

In previous sections, the parameters, $\xi = \xi\left(\mu_0, \sigma_0, \mu, \sigma\right)$, of interest were the average run length, standard deviation of the run length, and percentage points of the run length. When no standards are given these values are random quantities, $\hat{\mu}_N = \hat{\mu}_N\left(\hat{\mu}_0, \hat{\sigma}_0, \mu, \sigma\right)$, $\hat{\sigma}_N = \hat{\sigma}_N\left(\hat{\mu}_0, \hat{\sigma}_0, \mu, \sigma\right)$, $\hat{N}_\alpha = \hat{N}_\alpha\left(\hat{\mu}_0, \hat{\sigma}_0, \mu, \sigma\right)$. Two parameters of interest of these random variables are their means and standard deviations. These parameters can be determined from

$$E[\hat{\xi}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w f_{\hat{\xi}}\left(w \mid \mu_0, \sigma_0\right)\, dw\,; \tag{3.4.1}$$

$$E[\hat{\xi}^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w^2 f_{\hat{\xi}}\left(w \mid \mu_0, \sigma_0\right) dw\,. \tag{3.4.2}$$

It now follows the average run length of the chart is given by $\mu_N = E[N] = E[\hat{\mu}_N]$ when no standards are given. Similarly, the standard deviation and percentage points of the run length distribution are, respectively, $\sigma_N = E[\hat{\sigma}_N]$ and $N_\alpha = E[\hat{N}_\alpha]$. A FORTRAN program is presently under development for determining the ARL, STDRL (standard deviation of the run length), and selected percentage points of the run length distribution for various standardized shifts in $\mu_0$ and $\sigma_0$, when no standards are given.

A commonly used procedure for determining parameters of the run length distribution is simulation. Conceptually this method is quite simple. For a given charting procedure, a large number, L, of run lengths are independently simulated from the run length distribution. This is viewed as a random sample, $RL_1, RL_2, \ldots, RL_L$, from the run length distribution of the chart. The $\overline{RL}$ is used as an estimator for $\mu_N$, the standard deviation of this sample, $S_{RL}$, is used as an estimator of $\sigma_N$, and functions of the order statistics, $RL_{(1)}, RL_{(2)}, \ldots, RL_{(L)}$, could be used as estimators of the percentage points, $N_\alpha$.

### 3.5 Evaluating the P[N=N$_1$] with P[N=N$_2$]

In this section, we give expressions for evaluating the probability the $\overline{X}$-chart signals on or before the R- (or S-) chart. That is, we determine $P[N=N_1]=P[N_1 \le N_2]$. Equivalent expressions will be given for the probability the run length of the combined chart is the run length of the R- (or S-) chart, $P[N=N_2]=P[N_2 \le N_1]$. With $E[N_1] = E[N_2]$, it will be demonstrated it is more likely the $\overline{X}$-chart will signal on or before the R- (or S-) chart for a shift in the mean. Further for an increase in the standard deviation of the process, it is more likely the R- (or S-) chart will signal first.

First, consider the case where the basic Shewhart $\overline{X}$-chart and R- (or S-) charts are combined to monitor the mean and standard deviation. As stated previously, the run lengths, $N_1$ and $N_2$, each have geometric distributions with respective parameters

$$p_1 = P[\overline{X} \le LCL_1] + P[\overline{X} \ge UCL_1]$$

and

$$p_2 = P[R(\text{or } S) \le LCL_2] + P[R(\text{or } S) \ge UCL_2].$$

It needs to be noted the upper and lower control limits for the $\overline{X}$-chart are functions of both $\mu_0$ and $\sigma_0$, whereas, the upper and lower control limits for the R- (and S-) chart are only functions of $\sigma_0$. Further the distribution of $\overline{X}$ is a function of $\mu$ and $\sigma$; and the distribution of R (and S) is a function of only $\sigma$. It follows the value, $p_1 = p_1(\mu_0, \sigma_0, \mu, \sigma)$, is a function of $\mu$ and $\sigma$ and the value, $p_2 = p_2(\sigma_0, \sigma)$, is a function of $\sigma$.

Now consider the probability the minimum run length is equal to the length of the $\bar{X}$-chart. Thus,

$$P[N = N_1] = P[N_1 \le N_2]$$

$$= \sum_{t=1}^{\infty} P[N_1 = t]P[N_2 \ge t]$$

$$= \sum_{t=1}^{\infty} p_1 q_1^{t-1} q_2^{t-1}$$

$$= \frac{p_1}{1 - q_1 q_2} \sum_{t=1}^{\infty} p_1 (q_1 q_2)^{t-1} (1 - q_1 q_2)$$

$$= \frac{p_1}{1 - (1 - p_1)(1 - p_2)}$$

$$= \frac{E[N_2]}{E[N_1] + E[N_2] - 1} . \tag{3.5.1}$$

It follows the probability the minimum run length is equal to the run length of the R- or (S-) chart is

$$P[N = N_2] = P[N_2 \le N_1]$$

$$= \frac{p_2}{p_1 + p_2 - p_1 p_2}$$

$$= \frac{E[N_1]}{E[N_1] + E[N_2] - 1} . \tag{3.5.2}$$

Also it is easy to show

$$P[N_1 < N_2] = \frac{E[N_2] - 1}{E[N_1] + E[N_2] - 1} \tag{3.5.3}$$

$$P[N_1 = N_2] = \frac{1}{E[N_1] + E[N_2] - 1} \tag{3.5.4}$$

$$P[N_1 > N_2] = \frac{E[N_1] - 1}{E[N_1] + E[N_2] - 1}\,. \tag{3.5.5}$$

It may be of interest to determine the conditions under which $P[N_1 \le N_2] > \frac{1}{2}$. For this to hold

$$\frac{E[N_2]}{E[N_1] + E[N_2] - 1} > \frac{1}{2}$$

or

$$E[N_2] > E[N_1] - 1.$$

Thus it is more likely the $\overline{X}$-chart signals on or before the R- (or S-) chart if the ARL of the R- (S-) chart exceeds one less than the ARL of the $\overline{X}$-chart. As will be illustrated in Chapter 4, this still holds approximately when runs rules are added to one or both charts.

Now, consider the combined chart where the Shewhart $\overline{X}$- and R- (or S-) charts are supplemented with runs rules. Assuming $t_1^* \le t_2^*$, the probability $\overline{X}$-chart signals on or before the R- (or S-) chart and is given by

$$P[N = N_1] = P[N_1 \le N_2]$$

$$=\sum_{t=1}^{\infty} P[N_1 = t, N_2 \geq t]$$

$$=\sum_{t=1}^{\infty} P[N_1 = t]P[N_2 \geq t]$$

$$=\sum_{t=1}^{t_1^*} P[N_1 = t]P[N_2 \geq t] + \sum_{t=t_1^*+1}^{t_2^*} P[N_1 = t]P[N_2 \geq t]$$

$$+ \sum_{t=t_2^*+1}^{\infty} P[N_1 = t]P[N_2 \geq t]$$

$$=\sum_{t=1}^{t_1^*} P[N_1 = t]P[N_2 \geq t] + \sum_{t=1}^{t_2^*-t_1^*} P[N_1 = t_1^* + t]P[N_2 \geq t_1^* + t]$$

$$+ \sum_{t=1}^{\infty} P[N_1 = t_1^* + t_2^* - t_1^* + t]P[N_2 \geq t_2^* + t]$$

$$\approx \sum_{t=1}^{t_1^*} P[N_1 = t]P[N_2 \geq t] + \sum_{t=1}^{t_2^*-t_1^*} \lambda_1^t P[N_1 = t_1^*]P[N_2 \geq t_1^* + t]$$

$$+ \sum_{t=1}^{\infty} \lambda_1^{t_2^*-t_1^*+t} P[N_1 = t_1^*]\frac{\lambda_2^t}{1-\lambda_2} P[N_2 = t_2^*]$$

$$=\sum_{t=1}^{t_1^*} P[N_1 = t]P[N_2 \geq t] + P[N_1 = t_1^*]\sum_{t=1}^{t_2^*-t_1^*} \lambda_1^t P[N_2 \geq t_1^* + t]$$

$$+ \frac{\lambda_1^{t_2^*-t_1^*}}{1-\lambda_2} P[N_1 = t_1^*]P[N_2 = t_2^*]\sum_{t=1}^{\infty} \left(\lambda_1 \lambda_2\right)^t$$

$$= \sum_{t=1}^{t_1^*} P[N_1 = t]P[N_2 \geq t] + P[N_1 = t_1^*] \sum_{t=1}^{t_2^* - t_1^*} \lambda_1^t P[N_2 \geq t_1^* + t]$$

$$+ \frac{\lambda_1^{t_2^* - t_1^* + 1} \lambda_2}{(1-\lambda_2)(1-\lambda_1\lambda_2)} P[N_1 = t_1^*]P[N_2 = t_2^*] \tag{3.5.6}$$

which follows from equation (3.2.14). If we assume $t_1^* \geq t_2^*$, then

$$P[N = N_1] = \sum_{t=1}^{t_2^*} P[N_1 = t]P[N_2 \geq t] + \sum_{t=t_2^*+1}^{t_1^*} P[N_1 = t]P[N_2 \geq t]$$

$$+ \sum_{t=t_1^*+1}^{\infty} P[N_1 = t]P[N_2 \geq t]$$

$$= \sum_{t=1}^{t_2^*} P[N_1 = t]P[N_2 \geq t] + \sum_{t=1}^{t_1^* - t_2^*} P[N_1 = t_2^* + t]P[N_2 \geq t_2^* + t]$$

$$+ \sum_{t=1}^{\infty} P[N_1 = t_1^* + t]P[N_2 \geq t_2^* + t_1^* - t_2^* + t]$$

$$\approx \sum_{t=1}^{t_2^*} P[N_1 = t]P[N_2 \geq t] + \sum_{t=1}^{t_1^* - t_2^*} P[N_1 = t_2^* + t] \frac{\lambda_2^t P[N_2 = t_2^*]}{1-\lambda_2}$$

$$+ \sum_{t=1}^{\infty} \lambda_1^t P[N_1 = t_1^*] \frac{\lambda_2^{t_1^* - t_2^* + t}}{1-\lambda_2} P[N_2 = t_2^*]$$

$$= \sum_{t=1}^{t_2^*} P[N_1 = t]P[N_2 \geq t] + \frac{P[N_2 = t_2^*]}{1-\lambda_2} \sum_{t=1}^{t_1^* - t_2^*} \lambda_2^t P[N_1 \geq t_2^* + t]$$

$$+ \frac{\lambda_2^{t_1^* - t_2^*} P[N_1 = t_1^*]P[N_2 = t_2^*]}{1-\lambda_2} \sum_{t=1}^{\infty} (\lambda_1\lambda_2)^t$$

$$= \sum_{t=1}^{t_2^*} P[N_1 = t]P[N_2 \geq t] + \frac{P[N_2 = t_2^*]}{1-\lambda_2} \sum_{t=1}^{t_1^*-t_2^*} \lambda_2^t P[N_1 \geq t_2^* + t]$$

$$+ \frac{\lambda_1 \lambda_2^{t_1^*-t_2^*+1} P[N_1 = t_1^*]P[N_2 = t_2^*]}{(1-\lambda_2)(1-\lambda_1\lambda_2)} . \tag{3.5.7}$$

The approximate probability the $\overline{X}$-chart signals on or before the R- (or S-) chart, assuming $t_1^* \leq t_2^*$, is given by

$$P[N_1 < N_2] \approx \sum_{t=1}^{t_1^*} P[N_1 = t]P[N_2 > t] + P[N_1 = t_1^*] \sum_{t=1}^{t_2^*-t_1^*} \lambda_1^t P[N_2 \geq t_1^* + t]$$

$$+ \frac{\lambda_1^{t_2^*-t_1^*+1} \lambda_2^2}{(1-\lambda_2)(1-\lambda_1\lambda_2)} P[N_1 = t_1^*]P[N_2 = t_2^*] \tag{3.5.8}$$

and assuming $t_2^* \leq t_1^*$,

$$P[N_1 < N_2] \approx \sum_{t=1}^{t_2^*} P[N_1 = t]P[N_2 > t] + \frac{P[N_2 = t_2^*]}{1-\lambda_2} \sum_{t=1}^{t_1^*-t_2^*} \lambda_2^{t-1} P[N_1 = t_2^* + t]$$

$$+ \frac{\lambda_1^2 \lambda_2^{t_1^*-t_2^*+1} P[N_1 = t_1^*]P[N_2 = t_2^*]}{(1-\lambda_2)(1-\lambda_1\lambda_2)} . \tag{3.5.9}$$

## 3.6 Methods for Evaluating Run Length Distributions of Other Combined Charts

Jennett and Welch (1939) proposed plotting the statistic, (U - $\bar{X}$)/S, or the statistic, $\left(U-\bar{X}\right)/W$, where W is an easily calculated estimator of the standard deviation such as the sample range. The value, U, is an upper tolerance limit for the quality measure, X. Although Jennett and Welch (1939) did not consider supplementing runs rules for this chart, rules could be added. Addition of runs rules to this chart is not considered in this research.

Consider the chart based on (U - $\bar{X}$)/S. To determine the control limits, we need the distribution of (U - $\bar{X}$)/S. To obtain the distribution of this statistic, first note U - $\bar{X}$ has a normal distribution with mean, U-$\mu$, and variance, $\sigma^2/n$. Now consider the following

$$\frac{U-\bar{X}}{S}=\sqrt{n}\,\frac{\dfrac{U-\bar{X}}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}/(n-1)}}=\sqrt{n}\,\frac{\dfrac{\left(U-\bar{X}\right)-(U-\mu)}{\sigma/\sqrt{n}}+\dfrac{(U-\mu)}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}/(n-1)}}$$

$$=\sqrt{n}\,\frac{Z+\dfrac{(U-\mu)}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{\chi^2_{(n-1)}}{(n-1)}}}=\sqrt{n}\;t_{n-1,\theta}$$

where $t_{n-1,\theta}$ denotes a random variable having a noncentral t-distribution with n-1 degrees of freedom and noncentrality parameter, $\theta=\dfrac{(U-\mu)}{\sigma/\sqrt{n}}$. We note here the statistic, $\left(U-\bar{X}\right)/\left(S/\sqrt{n}\right)$ has a noncentral t-distribution with n-1 degrees of freedom and noncentrality parameter, $\theta$.

The expected value of the random variable, $\dfrac{U-\bar{X}}{S}$, is given by

$$E[\frac{U-\bar{X}}{S}]\;=\;E[U-\bar{X}]\;E[S^{-1}]\text{(Since }\bar{X}\text{ and S are independent)}$$

$$=\left(U-\mu\right)E[S^{-1}]$$

$$= (U-\mu)E[(S^2)^{-1/2}].$$

It can be shown $S^2$ has a gamma distribution with parameters $\frac{2\sigma^2}{n-1}$ and $\frac{n-1}{2}$. Thus, $E[(S^2)^{-1/2}]$ is given by

$$E[(S^2)^{-1/2}] = \int_0^{\infty} v^{-1/2} \frac{1}{\Gamma\left(\frac{n-1}{2}\right)\left(\frac{2\sigma^2}{n-1}\right)^{\frac{(n-1)}{2}}} v^{\frac{n-1}{2}} e^{-v/2}\, dv$$

$$= (c_3\sigma)^{-1}, \text{ where } c_3 = \sqrt{\frac{2}{n-1}}\ \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)}.$$

Hence,

$$E[\frac{U-\bar{X}}{S}] = c_3^{-1}\frac{U-\mu}{\sigma}.$$

The variance of (U - $\bar{X}$)/S is given by

$$V[\frac{U-\bar{X}}{S}] = \left(\frac{n-1}{n-3} - \frac{1}{c_3}\right)\left(\frac{U-\mu}{\sigma}\right)^2 + \frac{n-1}{n(n-3)}$$

Define $\theta_0 = \frac{U-\mu_0}{\sigma_0}$, where $\mu_0$ and $\sigma_0$ are, respectively, the in-control values of $\mu$ and $\sigma$. The control limits for this charting procedure can now be expressed as

$$LCL = \sqrt{n}\ t_{n-1,\theta_0,\alpha_L}$$

$$
\begin{aligned}
\mathrm{CL} &= c_3^{-1}\theta_0 \\
\mathrm{UCL} &= \sqrt{n}\; t_{n-1,\theta_0,1-\alpha_U}
\end{aligned}
$$

where $t_{n-1,\theta_0,\gamma}$ is the $\gamma^{th}$ percentage point of a noncentral t-distribution with n - 1 degrees of freedom and noncentrality parameter, $\theta_0$. The probability this chart signals at any sampling stage is

$$
p = 1 - F_{t_{n-1,\theta}}\left(t_{n-1,\theta_0,1-\alpha_U}\right) + F_{t_{n-1,\theta}}\left(t_{n-1,\theta_0,1-\alpha_L}\right). \tag{3.6.1}
$$

Thus,

$$
\mu_N = \frac{1}{p};\ \sigma_N = \frac{\sqrt{1-p}}{p}\ ;\text{and } N_\alpha = \left\lceil \frac{\log(1-\alpha)}{\log(1-p)} \right\rceil
$$

where N denotes the run length of this charting procedure.

When changes in the standard deviation are relatively rare or unimportant, Page (1955) proposed a supplementing the $\overline{X}$-chart with certain runs rules. These runs rules were listed in Chapter 2. Using a difference equation approach, he obtained the following expression for the ARL of this charting procedure

$$
\mathrm{ARL} = \frac{\left(1 - rs - r^m - s^m + sr^m + s^m r\right)}{p_2 + rs(1 + p_0) + p_0\left(r^m + s^m - sr^m - s^m r\right)} \tag{3.6.2}
$$

where r (s) is the probability a sample point falls between the upper (lower) warning and action lines, $p_0$ the probability a sample point falls between the warning limits, and $p_2$ the probability a sample point falls outside the action limits.

The combined $\bar{X}$- and $S^2$- chart investigated by Chengular, Arnold, and Reynolds (1989) can be analyzed in similar manner as the combined $\bar{X}$- and S-chart. Their procedure, based on the statistic $\sum_{j=1}^{n}\left(X_{t,j}-\mu_0\right)^2/\sigma_0^2$, has a probability, p, of signaling at any sampling stage given by

$$p = F\left(\chi^2_{n,1-\alpha_1}\right) - F\left(\chi^2_{n,\alpha_1}\right) \tag{3.6.3}$$

where $F(\,)$ is the CDF of a noncentral chi square distribution with n degrees of freedom and noncentrality parameter, $\left[(\mu-\mu_0)/(\sigma_0/\sqrt{n})\right]^2$.

The simultaneous control chart proposed by White and Schroeder (1982) plotted on separate charts the median ($\tilde{X}$) and inner quartile range (IQR). The probability of a signal at any sampling stage is

$$p = 1 - P\left[\tilde{X}_{n,\alpha_1-\tau_1} < \tilde{X} < \tilde{X}_{n,1-\tau_1},\ IQR_{n,\alpha_2-\tau_2} < IQR < IQR_{n,1-\tau_2}\right]$$

$$= 1 - \int_{LCL_1}^{UCL_1}\int_{LCL_2}^{UCL_2} f_{\tilde{X},IQR}(u,v|\mu,\sigma)\,dv\,du\,. \tag{3.6.4}$$

The combined control chart based on the statistics $T = \left(\bar{X}-\mu_0\right)/\left(S/\sqrt{n}\right)$ and S, we proposed in Chapter 2 has probability, p, of giving a signal at any sampling stage of

$$p = 1 - P\left[-t_{n-1,1-\alpha_1} < T < t_{n-1,1-\alpha_1},\ \chi_{n-1,\alpha_2-\tau_2} < S < \chi_{n-1,1-\tau_2}\right]$$

$$= 1 - \int_{-t_{n-1,1-\alpha_1}}^{t_{n-1,1-\alpha_1}}\int_{\chi_{n-1,\alpha_2-\tau_2}}^{\chi_{n-1,1-\tau_2}} f_{T,S}(u,v|\mu,\sigma)\,dv\,du, \tag{3.6.5}$$

where $t_{n-1,\gamma}$ and $\chi_{n-1,\gamma}$ are the $\gamma^{th}$ percentage points, respectively, of the t- and $\chi$-distributions each with n-1 degrees of freedom. It can be shown the joint distribution is

$$f_{T,S}(u,v|\mu,\sigma) = \frac{1}{\sqrt{2\pi}(\sigma/v)} e^{-\frac{1}{2}\left(\frac{u-[\mu-\mu_0]/[v/\sqrt{n}]}{\sigma/v}\right)^2}$$

$$\times \frac{2}{\Gamma\left(\frac{n-1}{2}\right)\left(\frac{2\sigma^2}{n-1}\right)^{\frac{n-1}{2}}} v^{n-2} e^{-v^2/\left(\frac{2\sigma^2}{n-1}\right)} \tag{3.6.6}$$

for $-\infty < u < \infty$ and $v \le 0$; and zero, otherwise.

If runs rules are added the T- or S- chart the transition probabilities are computed by

$$p_{i,j} = P[i_1 \rightarrow j_1, i_2 \rightarrow j_2]$$

$$= P\left[a_1 < T < b_1, a_1 < S < b_1\right]$$

$$= \int_{a_1}^{b_1}\int_{a_2}^{b_2} f_{T,S}(u,v|\mu,\sigma)\, dv\, du \tag{3.6.7}$$

where $f_{T,S}(u,v|\mu,\sigma)$ is given in equation (3.6.6).

The purpose of this section is to give brief outlines of ways to evaluate the run length properties of various simultaneous control charting procedures. Although, no further work will be done in this research with these procedures, more work needs to be done to give a full comparison of all available simultaneous control charting procedures.

# Chapter 4. SELECTING A COMBINED CHARTING PROCEDURE

## 4.1 Introduction

In this chapter, we discuss the design of combined $\overline{X}$ - and R- (or S-) control charts supplemented with runs rules. The criteria used for selecting a chart is based on what will be referred to as the ARL criteria. This criteria chooses the best chart from a set of charts with the same in-control ARL as the one with the smallest out-of-control ARL for a given shift in the parameters. In the case that more than one of these charts exist, a non-statistical user defined simplicity of use criteria should then be used to determine the best chart. Another desirable characteristic of a combined charting procedure is to have a signal on the $\overline{X}$ -chart more likely if the mean shifts and a signal on the R- (or S-) more likely if the standard deviation shifts.

## 4.2 Effects of Shifts in the Standard Deviation

It appears for various combined charting procedures equation (3.2.8) gives a good approximation to E[N] when the process is in-control. This is illustrated in the examples to follow. Thus, equation (3.2.8) provides a simple formula for determining the ARL of the combined chart when the ARLs of the $\overline{X}$ - and R- (or S-) charts are given. More generally, if two of the three ARLs are given in equation (3.2.8) the other can be determined (at least approximately).

Consider the $\overline{X}$ - and R-charts defined, respectively, by the sets of runs rules

$$C_{(1)} = \{T(1,1,-\infty,-3),\ T(2,3,-3,-2), T(2,3,2,3),\ T(1,1,3,\infty)\}$$

and

$$C_{(2)} = \{T(1,1,-\infty,-2.233), T(4,5,-2.233,-1.005), T(4,5,1.004,3.537), T(1,1,3.537,\infty)\}.$$

The chart, $C_{(2)}$, was recommended in Lowry, Champ, and Woodall (1994) for samples of size 5. Table 4.2.1 lists the ARLs for the charts, $C_{(1)}$, $C_{(2)}$, and the combined chart $C_{(1,2)}$. The columns are labeled respectively, $ARL_{(1)}$, $ARL_{(2)}$, and $ARL_{(1,2)}$. Further, Table 4.2.1 contains a column headed $A\tilde{R}L_{(1,2)}$ giving an approximation to the ARL of the combined chart using equation (3.2.8). As can be seen, equation (3.2.8) does provide in general a good approximation.

As can be seen in Table 4.2.1, a change in either the mean or standard deviation has an impact on the average run length of the combined chart. This supports the usual recommendation to examine the R-chart first before the $\overline{X}$-chart in an attempt to determine the reason for a signal.

We recommend first selecting the in-control ARL of the combined charting procedure. The ARLs of the $\overline{X}$- and R- (or S-) charts supplemented with runs rules can then be selected to satisfy equation (3.2.8). These charts can then be designed separately at this point. A run length analysis of the combined charting procedure should be done so the practitioner knows what to expect, on the average, as to the run length performance of the combined chart.

For this particular example, the probability of the $\overline{X}$-chart signaling on or before the R-chart for no change in the standard deviation is generally less than one-half for shifts in the mean of up to about a 20% increase. For shifts in the mean of 30% or more of an increase with no shift in the standard deviation, the $\overline{X}$-chart is more likely to signal on or before the R-chart. In this example, the ARL of the $\overline{X}$-chart is significantly larger than the ARL of the R- chart. In the next example, both the $\overline{X}$- chart and the R-chart have the same in-control mean.

TABLE 4.2.1 Various shifts in δ and λ .

| $\delta$ | $\lambda$ | $ARL_{(1)}$ | $ARL_{(2)}$ | $ARL_{(1,2)}$ | $\tilde{ARL}_{(1,2)}$ | $P[N = N_1]$ |
|---|---|---|---|---|---|---|
| 0.00 | 1.00 | 225.44 | 166.41 | 96.58 | 95.98 | 0.42696 |
| 0.00 | 1.10 | 99.88 | 75.88 | 43.92 | 43.37 | 0.43690 |
| 0.00 | 1.20 | 53.65 | 32.26 | 20.87 | 20.38 | 0.38369 |
| 0.00 | 1.30 | 33.01 | 16.88 | 11.81 | 11.40 | 0.34932 |
| 0.00 | 1.40 | 22.41 | 10.49 | 7.71 | 7.37 | 0.33273 |
| 0.10 | 1.00 | 211.45 | 166.41 | 93.97 | 93.37 | 0.44289 |
| 0.10 | 1.10 | 95.67 | 75.88 | 43.13 | 42.57 | 0.44779 |
| 0.10 | 1.20 | 52.08 | 32.26 | 20.65 | 20.16 | 0.39109 |
| 0.10 | 1.30 | 32.32 | 16.88 | 11.73 | 11.32 | 0.35450 |
| 0.10 | 1.40 | 22.07 | 10.49 | 7.68 | 7.34 | 0.33652 |
| 0.20 | 1.00 | 177.56 | 166.41 | 86.76 | 86.15 | 0.48693 |
| 0.20 | 1.10 | 84.82 | 75.88 | 40.87 | 40.30 | 0.47859 |
| 0.20 | 1.20 | 47.84 | 32.26 | 20.02 | 19.51 | 0.41250 |
| 0.20 | 1.30 | 30.40 | 16.88 | 11.51 | 11.09 | 0.36968 |
| 0.20 | 1.40 | 21.09 | 10.49 | 7.59 | 7.24 | 0.34769 |
| 0.30 | 1.00 | 138.68 | 166.41 | 76.51 | 75.89 | 0.54963 |
| 0.30 | 1.10 | 71.05 | 75.88 | 37.52 | 36.94 | 0.52440 |
| 0.30 | 1.20 | 42.07 | 32.26 | 19.03 | 18.51 | 0.44579 |
| 0.30 | 1.30 | 27.66 | 16.88 | 11.16 | 10.72 | 0.39387 |
| 0.30 | 1.40 | 19.65 | 10.49 | 7.43 | 7.07 | 0.36574 |
| 0.40 | 1.00 | 104.46 | 166.41 | 65.00 | 64.41 | 0.61993 |
| 0.40 | 1.10 | 57.50 | 75.88 | 33.53 | 32.96 | 0.57891 |
| 0.40 | 1.20 | 35.90 | 32.26 | 17.77 | 17.24 | 0.48791 |
| 0.40 | 1.30 | 24.53 | 16.88 | 10.69 | 10.25 | 0.42558 |
| 0.40 | 1.40 | 17.93 | 10.49 | 7.23 | 6.86 | 0.38987 |

Consider a combined $\overline{X}$ - and R chart supplemented with the following runs rules:

$\overline{X}$ -chart: $C_{(1)} = \{T(1,1,-\infty,-3), T(2,3,-3,-2), T(2,3,2,3), T(1,1,3,\infty)\}$

R-chart: $C_{(2)} = \{T(1,1,-\infty,-2.233), T(4,5,-2.2330,-1.1105),$

$T(4,5,1.114,3.537), T(1,1,3.537,\infty)\}.$

Again let $C_{(1,2)}$ denote the combined chart. As can be seen from Table 4.2.2, the ARLs of these charts are approximately the same.

TABLE 4.2.2 Various shifts in δ and λ .

| δ | λ | $ARL_{(1)}$ | $ARL_{(2)}$ | $ARL_{(1,2)}$ | $\tilde{ARL}_{(1,2)}$ | $P[N = N_1]$ |
|---|---|---|---|---|---|---|
| 0.00 | 1.00 | 225.44 | 225.17 | 113.38 | 112.90 | 0.50170 |
| 0.00 | 1.10 | 99.88 | 94.46 | 49.28 | 48.80 | 0.49084 |
| 0.00 | 1.20 | 53.65 | 38.04 | 22.96 | 22.50 | 0.42296 |
| 0.00 | 1.30 | 33.01 | 19.05 | 12.71 | 12.32 | 0.37701 |
| 0.00 | 1.40 | 22.41 | 11.45 | 8.14 | 7.81 | 0.35209 |
| 0.10 | 1.00 | 211.45 | 225.17 | 109.78 | 109.30 | 0.51787 |
| 0.10 | 1.10 | 95.67 | 94.46 | 48.27 | 47.78 | 0.50185 |
| 0.10 | 1.20 | 52.08 | 38.04 | 22.69 | 22.23 | 0.43055 |
| 0.10 | 1.30 | 32.32 | 19.05 | 12.62 | 12.22 | 0.38234 |
| 0.10 | 1.40 | 22.07 | 11.45 | 8.10 | 7.77 | 0.35596 |
| 0.20 | 1.00 | 177.56 | 225.17 | 100.00 | 99.52 | 0.56174 |
| 0.20 | 1.10 | 84.82 | 94.46 | 45.43 | 44.94 | 0.53268 |
| 0.20 | 1.20 | 47.84 | 38.04 | 21.91 | 21.44 | 0.45244 |
| 0.20 | 1.30 | 30.40 | 19.05 | 12.36 | 11.95 | 0.39791 |
| 0.20 | 1.40 | 21.09 | 11.45 | 7.99 | 7.66 | 0.36736 |
| 0.30 | 1.00 | 138.68 | 225.17 | 86.53 | 86.06 | 0.62221 |
| 0.30 | 1.10 | 71.05 | 94.46 | 41.28 | 40.79 | 0.57776 |
| 0.30 | 1.20 | 42.07 | 38.04 | 20.70 | 20.23 | 0.48614 |
| 0.30 | 1.30 | 27.66 | 19.05 | 11.94 | 11.53 | 0.42260 |
| 0.30 | 1.40 | 19.65 | 11.45 | 7.82 | 7.48 | 0.38574 |
| 0.40 | 1.00 | 104.46 | 225.17 | 72.00 | 71.57 | 0.68737 |
| 0.40 | 1.10 | 57.50 | 94.46 | 36.45 | 35.98 | 0.63022 |
| 0.40 | 1.20 | 35.90 | 38.04 | 19.20 | 18.72 | 0.52823 |
| 0.40 | 1.30 | 24.53 | 19.05 | 11.40 | 10.97 | 0.45472 |
| 0.40 | 1.40 | 17.93 | 11.45 | 7.59 | 7.23 | 0.41020 |

For this example, the probability of the $\overline{X}$ -chart signaling on or before the R-chart for no change in the standard deviation is one-half or greater. When a shift of 30% to 40% increase in the standard deviation occurs the R-chart is more likely to signal even if shifts of 40% increase occur in the mean. It is very noticeable that the ARL for both charts goes down noticeably for shifts of this magnitude.

### 4.3 Selection of Control Limits for the R- and S-Charts

Lowry and Champ (1994) consider the problem of selecting the warning and action limits of the R- and S-charts. For each of these charts suggested in the literature, they found there is an interval of values for $\lambda$ in which the ARL of these charts is greater than the in-control ARL. Presently they are working on recommendations for selecting the warning and action limits for R- and S- charts supplemented with runs rules such that the ARL of the chart is a maximum when the process is in-control. They refer to such a chart as an unbiased charting procedure. They have obtained results for the basic S-chart.

Table 4.3.1 is a reproduction of the results found in their paper. The use of Table 4.3.1 can be illustrated using the example given in Champ and Lowry (1994). Suppose the in-control value of the standard deviation is $\sigma_0 = 2$ and an S-chart based on a sample of size n = 5 with an in-control ARL = 250 is to be used. From Table 4.3.1, we obtain the values for the lower and upper control limits, $b_L = 2.157$ and $b_U = 3.659$. Now using the method described in Chapter 2 and the value of $c_4 = 0.94$, the lower and upper control limits and center line for this unbiased chart are

$$\mathrm{LCL} = (0.94 - 2.157\sqrt{1-(0.94)^2}\,)\cdot(2) = 0.4082$$

$$\mathrm{CL} = (0.94)\cdot(2) = 1.8800$$

$$\mathrm{UCL} = (0.94 + 3.659\sqrt{1-(0.94)^2}\,)\cdot(2) = 4.3767.$$

As indicated in their paper, Champ and Lowry (1994) are working on a similar table to Table 4.3.1 for the R-chart. Also, they are working on a method for designing unbiased R- and S-charts supplemented with runs rules. If the R- (or S-) chart is designed to be an unbiased chart, then the combined $\overline{X}$ - and R- (or S-) chart will be unbiased.

TABLE 4.3.1:Values of $k_L$ and $k_U$

| | | ARL | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 50 | 100 | 125 | 200 | 250 | 500 | 1000 |
| 3 | $b_L$ | 1.630 | 1.722 | 1.732 | 1.770 | 1.785 | 1.822 | 1.848 |
| | $b_U$ | 3.305 | 3.651 | 3.756 | 3.972 | 4.071 | 4.367 | 4.647 |
| 4 | $b_L$ | 2.002 | 2.096 | 2.122 | 2.167 | 2.186 | 2.235 | 2.271 |
| | $b_U$ | 2.400 | 2.716 | 2.814 | 3.014 | 3.107 | 3.385 | 3.651 |
| 5 | $b_L$ | 1.856 | 2.002 | 2.043 | 2.123 | 2.157 | 2.253 | 2.332 |
| | $b_U$ | 2.951 | 3.269 | 3.367 | 3.567 | 3.659 | 3.933 | 4.194 |
| 6 | $b_L$ | 2.120 | 2.266 | 2.308 | 2.389 | 2.425 | 2.524 | 2.608 |
| | $b_U$ | 2.403 | 2.703 | 2.795 | 2.984 | 3.071 | 3.332 | 3.581 |
| 7 | $b_L$ | 1.946 | 2.118 | 2.168 | 2.266 | 2.309 | 2.433 | 2.542 |
| | $b_U$ | 2.812 | 3.118 | 3.212 | 3.404 | 3.492 | 3.757 | 4.008 |
| 8 | $b_L$ | 2.167 | 2.336 | 2.386 | 2.484 | 2.527 | 2.651 | 2.761 |
| | $b_U$ | 2.399 | 2.692 | 2.782 | 2.966 | 3.051 | 3.305 | 3.547 |
| 9 | $b_L$ | 1.997 | 2.183 | 2.237 | 2.346 | 2.394 | 2.534 | 2.659 |
| | $b_U$ | 2.736 | 3.034 | 3.125 | 3.312 | 3.399 | 3.657 | 3.902 |
| 10 | $b_L$ | 2.193 | 2.376 | 2.430 | 2.537 | 2.585 | 2.724 | 2.848 |
| | $b_U$ | 2.394 | 2.683 | 2.772 | 2.953 | 3.036 | 3.287 | 3.524 |

## 4.4 Concluding Remarks

This chapter contains a discussion of the design of combined $\overline{X}$- and R- (or S-) control charts supplemented with runs rules. The ARL criteria is used to select charts. It is important for the combined charting procedure to have a signal on the $\overline{X}$-chart more likely if the mean shifts and a signal on the R- (or S-) more likely if the standard deviation shifts. For shifts in the standard deviation it has been shown that equation (3.2.8) provides a simple formula for determining the ARL of the combined chart when the ARLs of the $\overline{X}$- and R- (or S-) charts are given. Listed in Table 4.2.1 are various shifts in the mean and standard deviation. The ARLs of the individual and combined charts as well as an approximated ARL for the combined chart (using equation 3.2.8) and the probability of the $\overline{X}$-chart signaling on or before the R-chart are listed. When setting up a combined control charting procedure, it is recommended to first select the in-control ARL of the combined charting procedure. $\overline{X}$- and R- (or S-) charts should be designed to have equal in-control ARL's such that the combined chart has the desired ARL.

# Chapter 5. CONCLUSION

## 5.1 General Conclusions

We have investigated the use of individual Shewhart control charts with supplementary runs rules to jointly monitor both the mean and standard deviation of a quality measure. Chapter 2, gave a general explanation of the Shewhart $\overline{X}$- and R- (or S-) quality control charts supplemented with runs rules. Procedures were given for setting up these individual charts to be taken as a joint control charting procedure. Other combined mean and variability charts were discussed as well. In Chapter 3, the Markov chain approach of Champ and Woodall (1987) was used to develop two methods for analyzing the run length distribution of the combined $\overline{X}$- and R- (or S-) charts supplemented with runs rules. These two methods provide a simple way to obtain run length properties of the combined chart when monitoring the mean and standard deviation of a normal distribution. The problem of setting up a $\overline{X}$- and R- (or S-) chart when no standards are given was discussed in Section 3.4. Methods were suggested for evaluating run length distributions of other combined charts. Recommendations for selecting a combined charting procedure were given in Chapter 4.

## 5.2 Areas of Further Research

This research has provided an analysis of what is commonly done in practice and provides a foundation for further comparison of simultaneous control charting procedures. Many other charting procedures are available and are being studied. Further study could include implementing an analytical and a simulation analysis of a charting procedure when no standards are given. A FORTRAN program is presently under development for determining the ARL, STDRL, and selected percentage points of the run length distribution when $\mu_0$ and $\sigma_0$, are estimated from preliminary data.

In the discussion of evaluating run length distributions of other combined charts Jennett and Welch (1939) proposed plotting an estimator of the standard deviation such as the sample range. Although supplementary runs rules for this chart were not considered, rules could be added which would need further analysis.

## REFERENCES


Bain, L.J., and Engelhardt, M. (1992), *Introduction to Probability and Mathematical Statistics*, PWS-Kent: Boston.

Barnard, J. (1978), "Probability Integral of the Normal Range." *Applied Statistics* **27**, 197-198.

Bissell, A.F. (1978), "An Attempt to Unify the Theory of Quality Control Procedures." *Bulletin in Applied Statistics* **5**, 113-128.

Bragg, D.J., and St. John R.C.(1991), "Joint $\overline{X}$ and R Charts under Shift in Mu and Sigma." ASQC Technical Conference, 547-550.

Brooks and Evans (1972), "An approach to the probability distribution of CUSUM run length," *Biometrika* **59**, 539-549.

Burr (1967), "The Effect of Non-Normality on Constants for $\overline{X}$ and R Charts," *Industrial Quality Control* **23**, 563-569.

Burroughs T.E., and Champ, C.W. (1993), "Analysis of Shewhart Charts with Run Rules when No Standards are Given, "Proceedings of the American Statistical association Section on Quality and Productivity,16-19.

Champ, C.W. (1986), "Exact Results for Shewhart Control Charts with Supplementary Runs Rules." unpublished Ph.D. dissertation, University of Southwestern Louisiana, Department of Statistics.

Champ, C.W., and Lowry, C.A. (1994), "Proceedings of the Decision Sciences Conference."

Champ, C.W., Lowry, C.A. and Woodall, W.H. (1992). "Supplementing the R-Chart and S-Chart with Runs Rules." unpublished manuscript.

Champ, C.W., and Woodall, W.H. (1987), "Exact Results for Shewhart Control Charts with Supplementary Runs Rules," *Technometrics* **29**, 393-399.

Champ, C.W., and Woodall, W.H. (1990), "Programs to Evaluate the Run Length Distribution of a Shewhart Control Chart with Supplementary Runs Rules," *Journal of Quality Technology* **22**, 68-73.

Chengalur, Arnold and Reynolds (1989), "Variable Sampling Intervals for Multiparameter Shewhart Charts," *Communications in Statistics* **18**, 1769-1792.

Coleman, D. (1986), "The Power Function for $\overline{X}$ Charts: Symbolic Manipulation of Recurrence Relations and Markov Chains." Technical report, RCA, Princeton, NJ.

Fisher (1932), Statistical Methods for Research Workers, 4th Ed. Oliver and Boyd, Edinburgh and London.

Gan, F.F. (1989), "Combined Cumulative Sum and Shewhart Variance Charts," *Journal of Statistical Computation and Simulation* **32**, 149-163.

Gan, F.F. (1989), "The Distribution of the Run Length of EWMA Schemes When Observations Are Exponentially Distributed," Research Report No. 376, Department of Mathematics, National University of Singapore.

Gan, F.F. (1989), "Combined Mean and Variance Control Charts," *ASQC Technical Conference*, 129-139.

Harter, H.L. (1960), "Tables of Range and Studentized Range," *Annals of Mathematical Statistics* **31**, 1122-1147.

Jennett, W.J., and Welch, B.L. (1939), "The Control of Proportion Defective as Judged by a Single Quality Characteristic Varying on a Continuous Scale," *Journal of the Royal Statistical Society* Series B 6, 80-88.

Lowry, C.A., Champ, C.W., and Woodall, W.H. (1994), "Supplementing the R-Chart and S-Chart with Runs Rules." submitted for publication.

Montgomery, D.C.(1991). *Introduction to Statistical Quality Control*, 2nd ed., John Wiley & Sons, Inc. New York.

Nelson, L. S.(1984) "The Shewhart Control Chart- Tests for Special Causes," *Journal of Quality Technology* **16**, 237-239.

Nelson, L. S.(1985) "Interpreting Shewhart $\overline{\mathrm{X}}$ Control Charts," *Journal of Quality Technology* **17**, 114-116.

Ng, C.H., and Case, K.E. (1992), "Control Limits and the ARL: Some Surprises," First *Industrial Engineering Research Conference Proceedings*, Institute of Industrial Engineers, 127-129.

Page, E. S. (1955), "Control Charts with Warning Lines," *Biometrika* **42**, 243-257.

Palm, A.C. (1990), "Tables of Run Length Percentiles for Determining the Sensitivity of Shewhart Control Charts for Averages with Supplementary Runs Rules," *Journal of Quality Technology* **22**, 289-298.

Reynolds, M.R. and Ghosh, B.K. (1981), "Designing Control Charts for Means and Variances," *ASQC Quality Congress Transactions- San Francisco*.

Reynolds, M.R., Ghosh, B.K. and Hui, Y.V. (1981), "Shewhart $\overline{\mathrm{X}}$-Charts with Estimated Process Variance," *Communications in Statistics-Theory and Methods* **10**, 1797-1822.

Roberts, S.W. (1957), "Properties of Control Chart Zone Tests," *The Bell System Technical Journal* **37**, 83-114.

Schilling, E.G., and Nelson, P.R. (1976), "The Effect of Non-Normality on the Control Limits of $\overline{\mathrm{X}}$-Charts," *Journal of Quality Technology* **8**, 183-188.

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

Teichroew, D.(1956),"Tables of Expected Values of Order Statistics and Products of Order Statistics." *Annals of Mathematical Statistics* **27**, 410-425.

Western Electric Company (1956), *Statistical Quality Control Handbook*. American Telephone and Telegraph Company Chicago, Ill.

Wheeler, D.J. (1983), "Detecting a Shift in Process Control: Tables of the Power Function for $\overline{X}$ Charts," *Journal of Quality Technology* **15**, 155-170.

White, E.M., and Schroeder, R. (1987), "A Simultaneous Control Chart," *Journal of Quality Technology* **19**, 1-10.

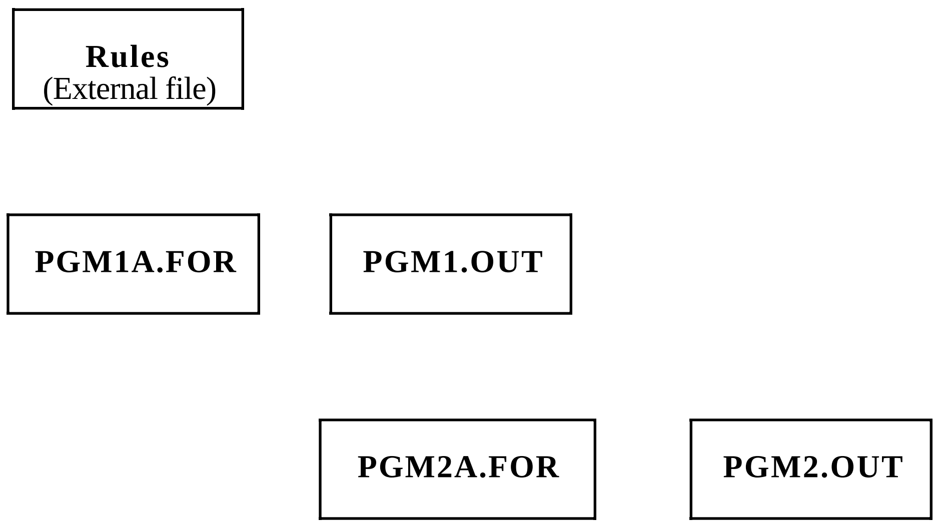
Woodall, W.H. (1983), "The Distribution of the Run Length of One-Sided CUSUM Procedures for Continuous Random Variables," *Technometrics* **25**, 295-301.

Woodall, W.H., and Ncube, M.M. (1984), "A Combined Shewhart-Cumulative Score Quality Control Chart," *Applied Statistics* **33**, 259-265.

Woodall, W.H., and Ncube, M.M. (1985), "Multivariate CUSUM Quality Control Procedures," *Technometrics* **27**, 285-292.

## Appendix I PROGRAMS FOR COMBINED RUN LENGTH DISTRIBUTION APPROACH

### Diagram of Programs for the Combined Run Length Distribution Approach




### RULES

### Description

Rules is an external data file that defines the charts. For the combined run length distribution approach a user would save two sets of runs rules defining two charts. As an example consider the following set of runs rules for the $\overline{X}$ - and R-chart

$\overline{X}$ : ={T(1,1,-9,-3), T(2,3,-3,-2),T(2,3,2,3), T(1,1,3,9)}

R: ={T(1,1,-9,-2.233),T(4,5,,-2.233,,-1.005),T(4,5,1.004,3.537), T(1,1,3.537,9)}. (Note: 9 is used as a practical value for infinity.)

**Program Description: PGM1A.FOR**

This program finds the Markov chain representations for the combined run length distributions approach. The rules defining the chart are read in from an external file called "RULES". Routines find the regions of the chart, the length of the state vector and the pointers to the end of the sub vector associated with each runs rule. The value of the present indicator variable is determined by runs rule and region combinations. The program then finds the state to state transitions by regions and sorts the states in ascending order of their base two representations. Any duplicate states are removed. Then the next-state transitions are numbered. Finally, the program outputs the state, next-state transition matrix.

**Program Listing**

```
C****************************************************************
C* PROGRAM FINDS THE MARKOV CHAIN REPRESENTATIONS            *
C* OF EACH OF THE SHEWHART CHARTS, ONE TO MONITOR            *
C* THE MEAN AND THE OTHER TO MONITOR THE STANDARD            *
C* DEVIATION.                                                *
C****************************************************************
C*
      CHARACTER*80 FMT
       INTEGER H,I,J,L,NS,NR,MR,CC,CK,CX,QH,SG,NT,
     &  NV,TMP,QQNS,K(20),M(20),D(20),PS(58),NX(58),
     &  X(20,10),Q(400,10),QQ(400),S(20)
      REAL A(20),B(20),R(41),TP
C*
C*
C*
C****************************************************************
C* THIS ROUTINE INPUTS THE RULES DEFINING THE CHART.         *
C****************************************************************
C*
      OPEN (60,FILE='PGM1.OUT',STATUS='OLD')
      OPEN (50,FILE='RULES',STATUS='OLD')
```

```
      DO 999 CC=1,2
      READ(50,50) FMT
   50 FORMAT(A80)
      READ(50,FMT) NT
      DO 101 I=1,NT
        READ(50,FMT) K(I),M(I),A(I),B(I)
  101 CONTINUE
C*
C***************************************************************
C* THIS ROUTINE FINDS THE REGIONS OF THE CHART.                *
C***************************************************************
C*
      R(1) = -9
      DO 201 I=1,NT
        R(2*I)   = A(I)
        R(2*I+1) = B(I)
  201 CONTINUE
      R(2*NT+2) = +9
      MR = 2*NT+1
      NR = MR
  202 CK = 0
      DO 204 J=1,MR
        IF (R(J).EQ.R(J+1).AND.J.LE.NR) THEN
          DO 203 L=J,NR
            R(L) = R(L+1)
  203     CONTINUE
          NR = NR-1
          CK = 1
        ENDIF
        IF (R(J).GT.R(J+1)) THEN
          TP = R(J)
          R(J) = R(J+1)
          R(J+1) = TP
          CK = 1
        ENDIF
  204 CONTINUE
      MR = MR-1
      IF (MR.GT.NR) MR=NR
      IF (CK.EQ.1.AND.MR.GE.1) GOTO 202
C*
C***************************************************************
C* THIS ROUTINE FINDS THE LENGTH OF THE STATE VECTOR           *
C* AND THE POINTERS TO  THE  END  OF  THE  SUBVECTOR           *
C* ASSOCIATED WITH EACH RUNS RULE.                             *
C***************************************************************
C*
      CK = 0
      NV = 0
      DO 301 I=1,NT
        NV = NV+M(I)-1
        IF (K(I).LT.M(I)) CK = 1
  301 CONTINUE
```

```
      IF (CK.EQ.0) NV = NV+1
      D(1) = M(1)-1
      DO 302 I=2,NT
        D(I) = D(I-1)+M(I)-1
        IF (M(I).EQ.1) D(I) = 0
  302 CONTINUE
C*
C*
C*
C*
C****************************************************************
C* THIS ROUTINE DETERMINES THE VALUE OF THE  PRESENT            *
C* INDICATOR  VARIABLE  BY  RUNS  RULE  AND   REGION            *
C* COMBINATION.                                                 *
C****************************************************************
C*
      DO 402 I=1,NT
        DO 401 J=1,NR
          X(I,J) = 0
          IF (A(I).LE.R(J).AND.R(J+1).LE.B(I))
     &       X(I,J) = 1
  401   CONTINUE
  402 CONTINUE
C*
C****************************************************************
C* THIS ROUTINE DETERMINES THE STATE TO STATE                   *
C* TRANSITIONS BY REGIONS.                                      *
C****************************************************************
C*
      QQ(1) = 0
      QQNS = 2**NV-1
      NS = 1
      H = 1
  500 QH = QQ(H)
      DO 501 L=1,NV
        PS(L) = QH-2*(QH/2)
        QH = QH/2
  501 CONTINUE
      DO 503 I=1,NT
        S(I) = 0
        IF (M(I).GT.1) THEN
          DO 502 L=D(I)-M(I)+2,D(I)
            S(I) = S(I)+PS(L)
  502     CONTINUE
        ENDIF
  503 CONTINUE
      DO 509 J=1,NR
        SG = 0
        DO 506 I=1,NT
          IF (SG.EQ.0) THEN
            IF (S(I)+X(I,J).GE.K(I)) THEN
              SG = 1
```

```
         ELSE
           IF (M(I).GT.1) NX(D(I)-M(I)+2)=X(I,J)
           IF (M(I).GT.2) THEN
            DO 504 L=D(I)-M(I)+3,D(I)
              NX(L) = PS(L-1)
504           CONTINUE
           ENDIF
         ENDIF
         IF (X(I,J).EQ.0.AND.M(I).GT.1) THEN
           TMP = S(I)-PS(D(I))+1
           L = D(I)
           CK = 0
505          IF (NX(L).EQ.1) THEN
             CK = 1
             IF (TMP.LT.K(I)) THEN
               NX(L) = 0
               TMP = TMP-1
               CK = 0
             ENDIF
           ENDIF
           L = L-1
           TMP = TMP+1
           IF (CK.EQ.0.AND.L.GE.D(I)-M(I)+2)
  &         GOTO 505
         ENDIF
       ENDIF
506    CONTINUE
      IF (SG.EQ.0) THEN
        QH = NX(1)
        DO 507 L=2,NV
          QH = QH+NX(L)*(2**(L-1))
507       CONTINUE
        CK = 0
        DO 508 L=1,NS
          IF (CK.EQ.0.AND.QH.EQ.QQ(L)) THEN
            Q(H,J) = QQ(L)
            CK = 1
          ENDIF
508       CONTINUE
        IF (CK.EQ.0) THEN
          NS = NS+1
          QQ(NS) = QH
          Q(H,J) = QH
        ENDIF
      ELSE
        Q(H,J) = QQNS
      ENDIF
509 CONTINUE
    H = H+1
    IF (H.LE.NS) GOTO 500
    NS = NS+1
    QQ(NS) = QQNS
```

```
     DO 510 J=1,NR
       Q(NS,J) = QQNS
  510 CONTINUE
C*
C*****************************************************************
C* THIS ROUTINE SORTS THE STATES IN ASCENDING ORDER              *
C* OF THEIR BASE TWO REPRESENTATIONS.                            *
C*****************************************************************
C*
  600 H = 0
     CK = 0
     DO 602 I=2,NS-H
       IF (QQ(I-1).GT.QQ(I)) THEN
         CK = 1
         TMP = QQ(I-1)
         QQ(I-1) = QQ(I)
         QQ(I) = TMP
         DO 601 J=1,NR
           TMP = Q(I-1,J)
           Q(I-1,J) = Q(I,J)
           Q(I,J) = TMP
  601     CONTINUE
       ENDIF
  602 CONTINUE
     H = H+1
     IF (CK.EQ.1) GOTO 600
C*
C*****************************************************************
C* THIS ROUTINE REMOVES ANY DUPLICATE STATES.                    *
C*****************************************************************
C*
  700 CK = 0
     I = 1
  701 H = I+1
  702 CX = 0
     DO 703 J=1,NR
       IF (Q(I,J).NE.Q(H,J)) CX=1
  703 CONTINUE
     IF (CX.EQ.0) THEN
       TMP = QQ(H)
       DO 705 L=1,H-1
         DO 704 J=1,NR
           IF (Q(L,J).EQ.TMP) Q(L,J)=QQ(I)
  704     CONTINUE
  705   CONTINUE
       DO 707 L=H,NS-1
         QQ(L) = QQ(L+1)
         DO 706 J=1,NR
           Q(L,J) = Q(L+1,J)
           IF (Q(L,J).EQ.TMP) Q(L,J)=QQ(I)
  706     CONTINUE
  707   CONTINUE
```

```
      NS = NS-1
      CK = 1
    ENDIF
    H = H+1
    IF (H.LT.NS) GOTO 702
    I = I+1
    IF (I.LT.NS-1) GOTO 701
    IF (CK.EQ.1) GOTO 700
C*
C*****************************************************************
C* THIS ROUTINE NUMBERS THE NEXT-STATE TRANSITIONS.              *
C*****************************************************************
C*
    DO 803 I=1,NS
      DO 802 J=1,NR
        IF (Q(I,J).LT.QQNS) THEN
          CK = 0
          L = 1
 801      IF (Q(I,J).EQ.QQ(L)) THEN
            Q(I,J) = L
            CK = 1
          ENDIF
          L = L+1
          IF (CK.EQ.0.AND.L.LT.NS) GOTO 801
        ELSE
          Q(I,J) = NS
        ENDIF
 802   CONTINUE
 803 CONTINUE
C*
C*****************************************************************
C* THIS ROUTINE OUTPUTS THE STATE, NEXT-STATE                    *
C* TRANSITION MATRIX.                                            *
C*****************************************************************
C*
 900 WRITE(60,961) NS,NR
 961 FORMAT(2(I4,1X))
    WRITE(60,962) (R(J),J=1,NR+1)
 962 FORMAT(10(F9.5))
    DO 964 I=1,NS
      WRITE(60,963) I,QQ(I),(Q(I,J),J=1,NR)
 963   FORMAT(20(1X,I4))
 964 CONTINUE
 999 CONTINUE
    CLOSE (50)
    CLOSE (60)
C*
    STOP
    END
```

**PGM1.OUT**.

The program PGM1A.FOR writes information to this external output file. It contains the state, next-state transition matrix for the combined run length distribution procedure.

**Program Description: PGM2A.FOR**

This program reads combined Markov chain representation from PGM1.OUT and calculates the average run lengths and standard deviations for various positive standardized shifts in the mean for the individual chart as well as the combined charts. The program also computes the probability the $\overline{X}$-chart is more likely to signal on or before the R-chart, i.e. $P[N = N_1]$. There are routines that input information about the chart. The program evaluates the run length distribution of the combined $\overline{X}$- and R-chart, where the statistics plotted are; $\frac{\overline{X} - \mu_0}{\sigma_0 / \sqrt{n}}$ and $r/d2 * \sigma_0$. Algorithm AS 126 Applied Statistics (1978) Vol. 27, No. 2 computes the probability of the normal range given T, the upper limit of integration, and N, the sample size. Algorithm AS 5 Applied Statistics (1968) Vol. 17, p.193 computes the lower tail area of non-central t-distribution. Algorithm ACM 291, Comm. ACM. (1966) Vol.9, p.684 evaluates the natural logarithm of $\Gamma$ (x). For x greater than zero this double precision function routine evaluates $\sqrt{2} * \Gamma\left(\left(m * (n-1) + 1\right) / 2\right)$, c4, $\sqrt{m * (n-1)} * \Gamma\left(m * (n-1) / 2\right)$ which computes the cumulative distribution function $P[y \le x]$ of a random variable Y having a Chi Squared distribution with n degrees of freedom using the function **dnml**. Largex is the largest value of x such that dexp(-x/2) is accurate for your particular machine.

**Program Listing**

```
C*************************************************************
C* COMBINED MEAN AND VARIABLITY CHARTS                       *
C*                                                           *
C* PROGRAM EVALUATES THE RUN LENGTH DISTRIBUTION OF          *
C* THE COMBINED XBAR AND R CHART                             *
C*                                                           *
C* THE STATISTICS PLOTTED ARE (XBAR-MU0)/(SIGMA0/SQRT(N))    *
C* AND R/(D2*SIGMA0)                                         *
C*************************************************************
C*
C*************************************************************
C* PROGRAM 2 (PGM 2)                                         *
C* THIS PROGRAM CALCULATES THE  AVERAGE  RUN  LENGTH         *
C* (ARL), THE STANDARD DEVIATION (STD), AND SELECTED         *
C* PERCENTAGE POINTS OF THE RUN LENGTH DISTRIBUTION.         *
C*************************************************************
C*
      INTEGER CC,CCC,CK(2),CV(31,9),I,II,
     &  IST,J,K,L1,L2,M,MAXN,MINN,NCP,ND(2),NS(2),
     &  NR(2),Q(2,200,10),QQ(2,200),SGN,STEP

      DOUBLE PRECISION ARL,ARL1,ARL2,CDF(2,101),
     &  CIV,CP(9),CUM(2),DCHISQ,DL,DNML,D2(25),
     &  D3(25),FCDF,L(2,200),LH,LP(2),P(2,10),
     &  PDF(2,101),PRNCST,P1EQ2,P1LE2,P1LT2,
     &  R(2,10),RNGPI,STD,U(2,200),ZA,ZB
C*
      WRITE(*,*) 'INPUT COMBINED CHART BY'
      WRITE(*,*) ' '
      WRITE(*,*) '(1) XBAR AND R'
      WRITE(*,*) '(2) XBAR AND S'
      WRITE(*,*) '(3) T AND S'
      READ(*,*) CCC
      WRITE(*,*) ' '
C*
      WRITE(*,*) 'INPUT SHIFTS IN MEAN CHART'
      READ(*,*) ND(1)
C*
      WRITE(*,*) 'INPUT SHIFTS IN DISPERSION CHART'
      READ(*,*) ND(2)
C*
      WRITE(*,*) 'INPUT DIRECTION OF SHIFT',
     &          ' IN DISPERSION CHART'
      WRITE(*,*) '(1) DECREASE'
      WRITE(*,*) '(2) INCREASE'
      READ(*,*) SGN
      SGN=2*SGN-3
C*
      WRITE(*,*) 'INPUT SAMPLE SIZE'
      READ(*,*) M
```

```
      IDF=M-1
      C4=CIV(1,M)
C*
      WRITE(*,*) 'INPUT MINIMUM N.'
      READ(*,*) MINN
      WRITE(*,*) 'INPUT MAXIMUM N.'
      READ(*,*) MAXN
C*
      STEP=1
      NCP=9
      CP(1)=0.01
      CP(2)=0.05
      CP(3)=0.10
      CP(4)=0.25
      CP(5)=0.50
      CP(6)=0.75
      CP(7)=0.90
      CP(8)=0.95
      CP(9)=0.99
C*
      D2(2)=1.1283791671D0
      D2(3)=1.6925687506D0
      D2(4)=2.0587507460D0
      D2(5)=2.3259289473D0
      D2(6)=2.5344127212D0
      D2(7)=2.7043567512D0
      D2(8)=2.8472006121D0
      D2(9)=2.9700263244D0
      D2(10)=3.0775054617D0
      D2(11)=3.1728727038D0
      D2(12)=3.2584552798D0
      D2(13)=3.3359803541D0
      D2(14)=3.4067631082D0
      D2(15)=3.4718268899D0
      D2(16)=3.5319827861D0
      D2(17)=3.5878839618D0
      D2(18)=3.6400637579D0
      D2(19)=3.6889630232D0
      D2(20)=3.7349501196D0
      D2(21)=3.7783358298D0
      D2(22)=3.8193846434D0
      D2(23)=3.8583234233D0
      D2(24)=3.8953481485D0
      D2(25)=3.9306292195D0
C*
      D3(2)=0.7267604553D0
      D3(3)=0.7891977107D0
      D3(4)=0.7740624738D0
      D3(5)=0.7466376009D0
      D3(6)=0.7191713092D0
      D3(7)=0.6942311313D0
      D3(8)=0.6721236717D0
```

```
      D3(9)=0.6525962151D0
      D3(10)=0.6352897762D0
      D3(11)=0.6198643117D0
      D3(12)=0.6060285277D0
      D3(13)=0.5935411244D0
      D3(14)=0.5822042445D0
      D3(15)=0.5718557265D0
      D3(16)=0.5623621426D0
      D3(17)=0.5536130572D0
      D3(18)=0.5455164487D0
      D3(19)=0.5379951043D0
      D3(20)=0.5309837904D0
      D3(21)=0.5244270274D0
      D3(22)=0.5182773314D0
      D3(23)=0.5124938181D0
      D3(24)=0.5070410861D0
      D3(25)=0.5018883188D0
C*
      DO 1 I=2,25
       D3(I) = DSQRT(D3(I))
1     CONTINUE
C*
C****************************************************************
C* THIS ROUTINE INPUTS INFORMATION ABOUT THE CHART.             *
C****************************************************************
C*
      OPEN (50,FILE='PGM1.OUT',STATUS='OLD')
      DO 55 CC=1,2
      READ(50,51) NS(CC),NR(CC)
   51 FORMAT(2(I4,1X))
      READ(50,52) (R(CC,J),J=1,NR(CC)+1)
   52 FORMAT(10(F9.5))
      DO 54 I=1,NS(CC)
          READ(50,53) II,IST,(Q(CC,I,J),J=1,NR(CC))
   53     FORMAT(20(1X,I4))
   54 CONTINUE
   55 CONTINUE
      CLOSE (50)
C*
C****************************************************************
C* THIS ROUTINE CALCULATES THE AVERAGE RUN LENGTHS              *
C* AND STANDARD DEVIATIONS OF THE CHART FOR VARIOUS             *
C* POSITIVE STANDARDIZED SHIFTS IN THE MEAN.                    *
C****************************************************************
C*
C*
      OPEN(60,FILE='PGM2.OUT',STATUS='OLD')
      DO 113 L1=0,ND(1),STEP
      DO 112 L2=0,ND(2),STEP
       DL=(L1/10.D0)/(1.D0+L2/10.D0)
       DO 101 J=1,NR(1)
        IF (CCC.EQ.1.OR.CCC.EQ.2) THEN
```

```
       ZA = (R(1,J)-L1/10.D0)/(1.D0+SGN*L2/10.D0)
       ZB = (R(1,J+1)-L1/10.D0)/(1.D0+SGN*L2/10.D0)
       P(1,J) = DNML(ZB)-DNML(ZA)
      ENDIF
      IF (CCC.EQ.3) THEN
       ZA=R(1,J)
       ZB=R(1,J+1)
       P(1,J)=PRNCST(ZB,IDF,DL,IFAULT)
     &         -PRNCST(ZA,IDF,DL,IFAULT)
      ENDIF
101   CONTINUE
     DO 102 J=1,NR(2)
      IF (CCC.EQ.1) THEN
       ZA=(D2(M)+R(2,J)*D3(M))/(1.D0+SGN*L2/10.D0)
       IF (ZA.LT.0.D0) ZA=0.D0
       ZB=(D2(M)+R(2,J+1)*D3(M))/(1.D0+SGN*L2/10.D0)
       IF (ZB.LT.0.D0) ZB=0.D0
       P(2,J) = RNGPI(ZB,M,IFAULT)
     &          -RNGPI(ZA,M,IFAULT)
      ENDIF
      IF (CCC.EQ.2.OR.CCC.EQ.3) THEN
       ZA=(C4+R(2,J)*DSQRT(1.D0-C4*C4))/(1.D0+SGN*L2/10.D0)
       IF (ZA.LT.0.D0) ZA=0.D0
       ZA=IDF*ZA*ZA
       ZB=(C4+R(2,J+1)*DSQRT(1.D0-C4*C4))/(1.D0+SGN*L2/10.D0)
       IF (ZB.LT.0.D0) ZB=0.D0
       ZB=IDF*ZB*ZB
       P(2,J)=DCHISQ(ZB,IDF)-DCHISQ(ZA,IDF)
      ENDIF
 102  CONTINUE
     N = 1
     DO 105 CC=1,2
     DO 104 I=1,NS(CC)-1
      U(CC,I) = 0.D0
      DO 103 J=1,NR(CC)
       IF (Q(CC,I,J).NE.NS(CC))
     &      U(CC,I)=U(CC,I)+P(CC,J)
 103   CONTINUE
      U(CC,I)=1.0D0-U(CC,I)
 104  CONTINUE
     PDF(CC,1) = U(CC,1)
     CUM(CC) = PDF(CC,1)
     CDF(CC,1) = CUM(CC)
     CK(CC) = 0
 105  CONTINUE
C*
 106  N=N+1
     DO 110 CC=1,2
     DO 108 I=1,NS(CC)-1
       L(CC,I) = 0.D0
       DO 107 J=1,NR(CC)
         IF (Q(CC,I,J).NE.NS(CC))
```

```
     &        L(CC,I)=L(CC,I)+P(CC,J)*U(CC,Q(CC,I,J))
  107    CONTINUE
  108   CONTINUE
        IF (U(CC,1).NE.0.D0.AND.CUM(CC).NE.1.D0) THEN
          LH = L(CC,1)/U(CC,1)
          LP(CC) = (1.D0-CUM(CC)-L(CC,1))/(1.D0-CUM(CC))
          TP = DABS(LH-LP(CC))
          IF (N.GT.MINN.AND.TP.LT.0.000001D0) CK(CC)=1
        ENDIF
        IF (N.GT.MAXN) CK(CC)=1
        DO 109 I=1,NS(CC)-1
          U(CC,I) = L(CC,I)
  109   CONTINUE
        PDF(CC,N) = L(CC,1)
        CUM(CC) = CUM(CC) + L(CC,1)
        CDF(CC,N) = CUM(CC)
  110   CONTINUE
        IF (CK(1).EQ.0.OR.CK(2).EQ.0) GOTO 106
C*
        ARL=1.D0
        ARL1=1.D0
        ARL2=1.D0
        P1LT2=0.D0
        P1EQ2=0.D0
        DO 111 I=1,N
          ARL=ARL+(1.D0-CDF(1,I))*(1.D0-CDF(2,I))
          ARL1=ARL1+(1.D0-CDF(1,I))
          ARL2=ARL2+(1.D0-CDF(2,I))
          P1LT2=P1LT2+PDF(1,I)*(1.D0-CDF(2,I))
          P1EQ2=P1EQ2+PDF(1,I)*PDF(2,I)
111     CONTINUE
        TP=(1.D0-LP(1))*(1.D0-LP(2))*(1.D0-LP(1)*LP(2))
        TP=LP(1)*LP(1)*LP(2)*LP(2)*PDF(1,N)*PDF(2,N)/TP
        ARL=ARL+TP
        TP=LP(1)/(1.D0-LP(1))
        TP=TP*TP
        ARL1=ARL1+TP*PDF(1,N)
        TP=LP(2)/(1.D0-LP(2))
        TP=TP*TP
        ARL2=ARL2+TP*PDF(2,N)
        TP=(1.D0-LP(2))*(1.D0-LP(1)*LP(2))
        TP=LP(1)*LP(2)*LP(2)*PDF(1,N)*PDF(2,N)/TP
        P1LT2=P1LT2+TP
        TP=LP(1)*LP(2)*PDF(1,N)*PDF(2,N)/(1.D0-LP(1)*LP(2))
        P1EQ2=P1EQ2+TP
        P1LE2=P1LT2+P1EQ2
        TP=(ARL1*ARL2)/(ARL1+ARL2-1)
        WRITE(*,60)  L1/10.D0,1.D0+SGN*L2/10.D0,
     &          ARL1,ARL2,ARL,TP,P1LE2
        WRITE(60,60) L1/10.D0,1.D0+SGN*L2/10.D0,
     &          ARL1,ARL2,ARL,TP,P1LE2
60      FORMAT(2(F4.2,1X),4(F8.2,1X),F7.5)
```

```
112   CONTINUE
113   CONTINUE
      CLOSE (60)
C*
  999 STOP
      END
C*
C*
C*
C*
C***************************************************************
C* ALGORITHM  AS  126  APPLIED  STATISTICS   (1978)            *
C* VOL. 27, NO. 2                                              *
C*                                                             *
C* COMPUTES THE PROBABILITY OF THE NORMAL RANGE                *
C* GIVEN T, THE UPPER LIMIT OF INTEGRATION, AND N,             *
C* THE SAMPLE SIZE.                                            *
C***************************************************************
C*
      DOUBLE PRECISION FUNCTION RNGPI(T,N,IFAULT)
      INTEGER N,I,IFAULT
      DOUBLE PRECISION A,B,C,DNML,G(8),H(8),RISF,
     &  T,X,XL,Y
C*
      DATA G(1),G(2),G(3),G(4),G(5),G(6),G(7),G(8)
     & /0.4947004675D0, 0.4722875115D0, 0.4328156012D0,
     &  0.3777022042D0, 0.3089381222D0, 0.2290083888D0,
     &  0.1408017754D0, 0.04750625492D0/
C*
      DATA H(1),H(2),H(3),H(4),H(5),H(6),H(7),H(8)
     & /0.01357622971D0, 0.03112676197D0, 0.04757925584D0,
     &  0.06231448563D0, 0.07479799441D0, 0.08457825969D0,
     &  0.09130170752D0, 0.09472530523D0/
C*
      RISF(X)=0.3989422804D0*EXP(-0.5D0*X*X)*
     &        (DNML(X)-DNML(X-T))**(N-1)
C*
      IFAULT=1
      RNGPI=0.D0
      IF (T.LE.0.D0.OR.N.LE.1) RETURN
      IFAULT=0
      XL=0.5D0*T
      A=0.5D0*(8.D0+XL)
      B=8.D0-XL
      Y=0.D0
      DO 1 I=1,8
        C=B*G(I)
        Y=Y+H(I)*(RISF(A+C)+RISF(A-C))
1     CONTINUE
      RNGPI=(2.D0*(DNML(XL)-0.5D0))**N+2.D0*B*Y*N
      IF (RNGPI.GT.1.D0) RNGPI=1.D0
      RETURN
```

```
      END
C*
      DOUBLE PRECISION FUNCTION DNML(X)
C*
C COMPUTES THE CUMULATIVE DISTRIBUTION FUNCTION
C P(Y<=X) OF A RANDOM VARIABLE Y HAVING A
C STANDARD NORMAL DISTRIBUTION.
C*
      DOUBLE PRECISION X,Y,S,RN,ZERO,ONE,ERF,SQRT2,PI
      DATA SQRT2,ONE/1.414213562373095D0,1.D0/
      DATA PI,ZERO/3.141592653589793D0,0.D0/
      Y=X/SQRT2
      IF (X.LT.ZERO) Y=-Y
      S=ZERO
      DO 1 N=1,37
      RN=N
      S=S+DEXP(-RN*RN/25.D0)/N*DSIN(2.D0*N*Y/5.D0)
1     CONTINUE
      S=S+Y/5.D0
      ERF=2.D0*S/PI
      DNML=(ONE+ERF)/2.D0
      IF (X.LT.ZERO) DNML=(ONE-ERF)/2.D0
      IF (X.LT.-8.3D0) DNML=ZERO
      IF (X.GT.8.3D0) DNML=ONE
      RETURN
      END
C*
C*
C*****************************************************************
C*                                                               *
C* ALGORITHM AS 5 APPL STATIST, 1968 VOL 17, P.193               *
C*                                                               *
C* COMPUTES LOWER TAIL AREA OF NON-CENTRAL                       *
C* T-DISTRIBUTION                                                *
C*                                                               *
C*****************************************************************
C*
      REAL*8 FUNCTION PRNCST(ST,IDF,D,IFAULT)
C*
C*****************************************************************
C ALGORITHM AS 5 APPL. STATIST. (1968) VOL. 17, P.193            *
C*                                                               *
C COMPUTES LOWER TAIL AREA OF NON-CENTRAL T-DISTRIBUTION*
C*****************************************************************
C*
      REAL*8 ST,D,G1,G2,G3,ZERO,ONE,TWO,HALF,EPS,EMIN,F,
     &  A,B,RB,DA,DRB,FMKM1,FMKM2,SUM,AK,FK,FKM1,
     &  ALNORM,TFN,ALOGAM,ZSQRT,ZEXP
C*
C CONSTANTS - G1 IS 1.0/SQRT(2.0*PI)
C           G2 IS 1.0/(2.0*PI)
C           G3 IS SQRT(2.0*PI)
```

```
C*
      DATA G1,G2,G3/0.3989422804,0.1591549431,2.5066282746/
      DATA ZERO,ONE,TWO,HALF,EPS,EMIN
     & /0.0,1.0,2.0,0.5,1.0E-6,12.5/
C*
      ZSQRT(A)=DSQRT(A)
      ZEXP(A)=DEXP(A)
C*
      F=IDF
      IF (IDF.GT.100) GOTO 50
      IFAULT=0
      IOE=MOD(IDF,2)
      A=ST/ZSQRT(F)
      B=F/(F+ST**2)
      RB=ZSQRT(B)
      DA=D*A
      DRB=D*RB
      SUM=ZERO
      IF (IDF.EQ.1) GOTO 30
      FMKM2=ZERO
      IF (ABS(DRB).LT.EMIN) FMKM2=A*RB*ZEXP(-HALF*DRB**2)
     &  *ALNORM(A*DRB,.FALSE.)*G1
      FMKM1=B*DA*FMKM2
      IF (ABS(D).LT.EMIN)
     &  FMKM1=FMKM1+B*A*G2*ZEXP(-HALF*D**2)
      IF (IOE.EQ.0) SUM=FMKM2
      IF (IOE.EQ.1) SUM=FMKM1
      IF (IDF.LT.4) GOTO 20
      IFM2=IDF-2
      AK=ONE
      FK=TWO
      DO 10 K=2,IFM2,2
      FKM1=FK-ONE
      FMKM2=B*(DA*AK*FMKM1+FMKM2)*FKM1/FK
      AK=ONE/(AK*FKM1)
      FMKM1=B*(DA*AK*FMKM2+FMKM1)*FK/(FK+ONE)
      IF (IOE.EQ.0) SUM=SUM+FMKM2
      IF (IOE.EQ.1) SUM=SUM+FMKM1
      AK=ONE/(AK*FK)
      FK=FK+TWO
10    CONTINUE
20    IF (IOE.EQ.0) GOTO 40
30    PRNCST=ALNORM(DRB,.TRUE.)+TWO*(SUM+TFN(DRB,A))
      RETURN
40    PRNCST=ALNORM(D,.TRUE.)+SUM*G3
      RETURN
C*
C NORMAL APPROXIAMTION - K IS NOT TESTED AFTER THE TWO CALLS
C OF ALOGAM, BECAUSE A FAULT IS IMPOSSIBLE WHEN F EXCEEDS 100
C*
50    IFAULT=1
      A=ZSQRT(HALF*F)*ZEXP(ALOGAM(HALF*(F-ONE),K)
```

```
     & -ALOGAM(HALF*F,K))*D
      PRNCST=ALNORM((ST-A)/ZSQRT(F*(ONE+D**2)
     & /(F-TWO)-A**2),.FALSE.)
      RETURN
      END
C*
      REAL*8 FUNCTION ALNORM(X,UPPER)
C*
C*************************************************************
C* ALGORITHM AS 66 APPL. STATIST. (1973) VOL.22, P.424      *
C*                                                           *
C* EVALUATES THE TAIL AREA OF THE STANDARDIZED NORMAL        *
C* CURVE FROM X TO INFINITY IF UPPER IS .TRUE. OR FROM       *
C* MINUS INFINITY TO X IF UPPER IS .FALSE.                   *
C*                                                           *
C*************************************************************

      REAL*8 LTONE,UTZERO,ZERO,HALF,ONE,CON,A1,A2,A3,
     &  A4,A5,A6,A7,B1,B2,B3,B4,B5,B6,B7,B8,B9,
     &  B10,B11,B12,X,Y,Z,ZEXP
      LOGICAL UPPER,UP
C*
C* LTONE AND UTZERO MUST BE SET TO SUIT THE PARTICULAR
C* COMPUTER (SEE INTRODUCTORY TEXT)
C*
      DATA LTONE,UTZERO /7.0,18.66/
      DATA ZERO,HALF,ONE,CON /0.0,0.5,1.0, 1.28/
      DATA A1,A2,A3,A4,A5,A6,A7
     &  /0.398942280444, 0.399903438504, 5.75885480458,
     &    29.8213557808,  2.62433121679, 48.6959930692,
     &    5.92885724438/
      DATA B1,B2,B3,B4,B5,B6,B7,B8,B9,B10,B11,B12
     &  /0.398942280385, 3.8052E-8, 1.00000615302,
     &   3.98064794E-4, 1.98615381364, 0.151679116635,
     &   5.29330324926, 4.8385912808, 15.1508972451,
     &   0.742380924027, 30.789933034, 3.99019417011/
C*
      ZEXP(Z)=DEXP(Z)
C*
      UP=UPPER
      Z=X
      IF (Z.GE.ZERO) GOTO 10
      UP=.NOT.UP
      Z=-Z
10    IF (Z.LE.LTONE.OR.UP.AND.Z.LE.UTZERO) GOTO 20
      ALNORM=ZERO
      GOTO 40
20    Y=HALF*Z*Z
      IF (Z.GT.CON) GOTO 30
C*
      ALNORM=HALF-Z*(A1-A2*Y/(Y+A3-A4/(Y+A5+A6/(Y+A7))))
      GOTO 40
```

```
C*
30   ALNORM=B1*ZEXP(-Y)/(Z-B2+B3/(Z+B4+B5/
   &      (Z-B6+B7/(Z+B8-B9/(Z+B10+B11/(Z+B12))))))
C*
40   IF (.NOT.UP) ALNORM=ONE-ALNORM
     RETURN
     END
C*
     REAL*8 FUNCTION TFN(X,FX)
C*
C*************************************************************
C* ALGORITHM AS 76 APPL. STATIST. (1974) VOL. 23, P. 455     *
C*                                                           *
C* CALCULATES THE T-FUNCTION OF OWEN, USING GAUSSIAN         *
C* QUADRATURE                                                *
C*************************************************************
C*
     REAL*8 U(5),R(5),X,FX,TP,TV1,TV2,TV3,TV4,ZERO,
    &  QUART,HALF,ONE,TWO,R1,R2,RT,XS,X1,X2,FXS,
    &  ZABS,ZEXP,ZLOG,ZSIGN,ZATAN
C*
     DATA U(1),U(2),U(3),U(4),U(5)
    &    /0.0744372, 0.2166977, 0.3397048,
    &     0.4325317, 0.4869533/
C*
     DATA R(1),R(2),R(3),R(4),R(5)
    &    /0.1477621, 0.1346334, 0.1095432,
    &     0.0747257, 0.0333357/
C*
     DATA NG,TP,TV1,TV2,TV3,TV4
    &    /5, 0.159155, 1.0E-35,
    &     15.0, 15.0, 1.0E-5/
C*
     DATA ZERO,QUART,HALF,ONE,TWO
    &    /0.0, 0.25, 0.5, 1.0, 2.0/
C*
     ZABS(X)=DABS(X)
     ZEXP(X)=DEXP(X)
     ZLOG(X)=DLOG(X)
     ZSIGN(X1,X2)=DSIGN(X1,X2)
     ZATAN(X)=DATAN(X)
C*
C TEST FOR X NEAR ZERO
C*
     IF (ZABS(X).GE.TV1) GOTO 5
     TFN=TP*ZATAN(FX)
     RETURN
C*
C TEST FOR LARGE VALUES OF ABS(X)
C*
5    IF (ZABS(X).GT.TV2) GOTO 10
C*
```

```
C TEST FOR FX NEAR ZERO
C*
      IF (ZABS(FX).GE.TV1) GOTO 15
10    TFN=ZERO
      RETURN
C*
C TEST WHETHER DABS(FX) IS SO LARGE THAT IT MUST BE
C TRUNCATED
C*
15    XS=-HALF*X*X
      X2=ZABS(FX)
      FXS=FX*FX
      IF (ZLOG(ONE+FXS)-XS*FXS.LT.TV3) GOTO 25
C*
C COMPUTATION OF TRUNCATION POINT BY NEWTON ITERATION
C*
      X1=HALF*X2
      FXS=QUART*FXS
20    RT=FXS+ONE
      X2=X1+(XS*FXS+TV3-ZLOG(RT))/(TWO*X1*
     &   (ONE/RT-XS))
      FXS=X2*X2
      IF (ZABS(X2-X1).LT.TV4) GOTO 25
      X1=X2
      GOTO 20
C*
C GAUSSIAN QUADRATURE
C*
25    RT=ZERO
      DO 30 I=1,NG
      R1=ONE+FXS*(HALF+U(I))**2
      R2=ONE+FXS*(HALF-U(I))**2
      RT=RT+R(I)*(ZEXP(XS*R1)/R1+ZEXP(XS*R2)/R2)
30    CONTINUE
      TFN=ZSIGN(RT*X2*TP,FX)
      RETURN
      END
C*
      REAL*8 FUNCTION ALOGAM(X,IFAULT)
C*****************************************************************
C* ALGORITHM ACM 291, COMM. ACM. (1966) VOL.9, P.684             *
C*                                                               *
C* EVALUATES NATURAL LOGARITHM OF GAMMA(X)                       *
C* FOR X GREATER THAN ZERO                                       *
C*****************************************************************
      REAL*8 A1,A2,A3,A4,A5,F,X,Y,Z,ZLOG,
     &  HALF,ZERO,ONE,SEVEN
C*
C* THE FOLLOWING CONSTANTS ARE ALOG(2PI)/2,
C* 1/160, 1/1260, 1/360 AND 1/12
C*
      DATA A1,A2,A3,A4,A5
```

```
     &  /0.918938533204673, 0.000595238095238,
     &   0.000793650793651, 0.002777777777778,
     &   0.083333333333333/
C*
      DATA HALF,ZERO,ONE,SEVEN
     &  /0.5, 0.0, 1.0, 7.0/
C*
      ZLOG(F)=DLOG(F)
C*
      ALOGAM=ZERO
      IFAULT=1
      IF (X.LE.ZERO) RETURN
      IFAULT=0
      Y=X
      F=ZERO
      IF (Y.GE.SEVEN) GOTO 30
      F=Y
10    Y=Y+ONE
      IF (Y.GE.SEVEN) GOTO 20
      F=F*Y
      GOTO 10
20    F=-ZLOG(F)
30    Z=ONE/(Y*Y)
      ALOGAM=F+(Y-HALF)*ZLOG(Y)-Y+A1
     &  +(((-A2*Z+A3)*Z-A4)*Z+A5)/Y
      RETURN
      END
C*
C*****************************************************************
C* THIS DOUBLE PRECISION FUNCTION ROUTINE COMPUTES               *
C*                                                               *
C*          SQRT(2)*GAMMA((M*(N-1)+1)/2)                         *
C*     C4 = ------------------------------------------           *
C*         SQRT(M*(N-1))*GAMMA(M*(N-1)/2)                        *
C*                                                               *
C*****************************************************************
C*
      DOUBLE PRECISION FUNCTION CIV(M,N)
C*
      INTEGER I,J,K,M,N
      DOUBLE PRECISION G1,G2,XK
C*
      K = M*(N-1)
      IF (M.EQ.0) K = N-1
      XK = K
C*
      G1 = 0.564189583D0
      G2 = 0.886226925D0
      IF (K.EQ.1) CIV = G1
      IF (K.GE.2) CIV = G2
      IF (K.GT.2) THEN
        IF ((K/2)*2.NE.K) THEN
```

```
      J = (K+1)/2
      CIV = G1
      DO 100 I=2,J
        CIV = ((2.D0*I-2.D0)/(2.D0*I-3.D0))*CIV
100     CONTINUE
     ELSE
      J = K/2
      CIV = G2
      DO 200 I=2,J
        CIV = ((2.D0*I-1.D0)/(2.D0*I-2.D0))*CIV
200     CONTINUE
     ENDIF
    ENDIF
C*
    CIV = DSQRT(2.D0)*CIV*DSQRT(XK)/XK
C*
    RETURN
    END
C*
C***************************************************************
C* COMPUTES  THE  CUMULATIVE  DISTRIBUTION  FUNCTION          *
C* P[Y<=X] OF A  RANDOM  VARIABLE  Y  HAVING  A  CHI          *
C* SQUARED DISTRIBUTION WITH N  DEGREES  OF  FREEDOM          *
C* USING THE FUNCTION DNML.                                   *
C***************************************************************
C*
    DOUBLE PRECISION FUNCTION DCHISQ(X,N)
    INTEGER N
    DOUBLE PRECISION A,Y,S,E,C,Z,X1,X,DNML
    DOUBLE PRECISION PI,ONE,ZERO,HALF,LARGEX
    LOGICAL BIGX
C*
C***************************************************************
C* LARGEX IS THE LARGEST VALUE OF X SUCH THAT                 *
C* DEXP(-X/2) IS ACCURATE FOR YOUR PARTICULAR                 *
C* MACHINE.                                                   *
C***************************************************************
C*
    DATA LARGEX,PI/174.99646D0,3.141592653589793D0/
    DATA ONE,ZERO,HALF/1.D0,0.D0,0.5D0/
    IF (X.LT.0.D0) WRITE(*,100) 'X',0
    IF (N.LT.1) WRITE(*,100) 'N',1
    IF (X.GE.0.D0.AND.N.GE.1) GOTO 10
    DCHISQ=-ONE
    RETURN
10   A=HALF*X
    I=MOD(N+1,2)
    BIGX=.FALSE.
    IF(X.GT.LARGEX) BIGX=.TRUE.
    Y=DEXP(-A)
    IF (N.EQ.1.OR.BIGX) Y=ZERO
    S=Y
```

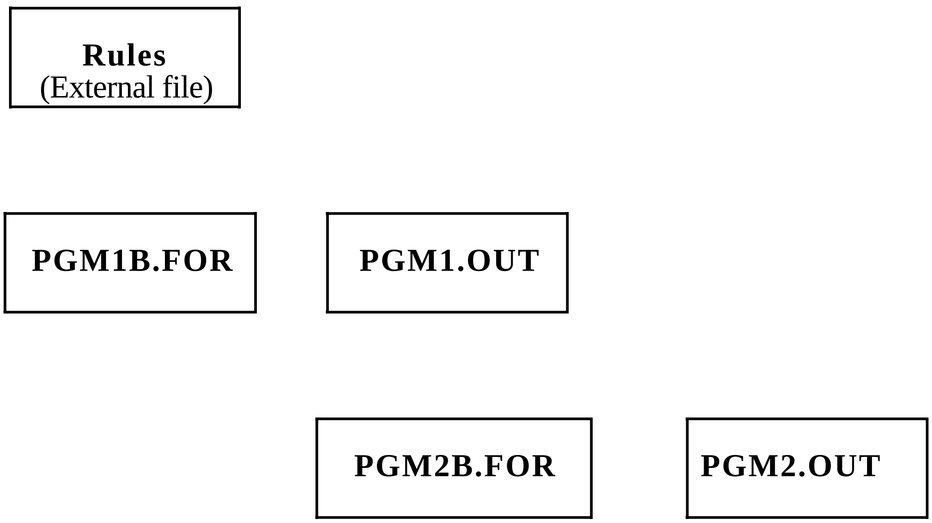
```
      IF (I.EQ.0) S=2*DNML(-DSQRT(X))
      IF (N.EQ.1) GOTO 30
      X1=HALF*(N-1)
      Z=(ONE+I)/2
      IF (.NOT.BIGX) GOTO 40
      E=ZERO
      IF (I.EQ.0) E=DLOG(DSQRT(PI))
      C=DLOG(A)
20    IF (Z.GT.X1) GOTO 30
      E=E+DLOG(Z)
      S=S+DEXP(C*Z-A-E)
      Z=Z+ONE
      GOTO 20
30    DCHISQ=ONE-S
      RETURN
40    E=ONE
      IF (I.EQ.0) E=ONE/DSQRT(PI*A)
      C=ZERO
50    IF (Z.GT.X1) GOTO 60
      E=E*A/Z
      C=C+E
      Z=Z+ONE
      GOTO 50
60    DCHISQ=ONE-C*Y-S
      RETURN
100   FORMAT(' IN THE FUNCTION DCHISQ(X,N) THE ',
     &'PARAMETER ',A1,' MUST BE >= ',I1)
      END
```

**PGM2.OUT**

The program outputs $\delta$ (the number of shifts in the mean), $\lambda$ (the number of shifts in the standard deviation), the average run lengths (ARLs) for the individual chart as well as the combined chart and $P[N = N_1]$, the probability the $\overline{X}$-chart is more likely to signal on or before the R-chart into an external data file called PGM2.OUT.

## Appendix II PROGRAMS FOR COMBINED MARKOV CHAIN APPROACH

### Diagram of Programs for the Combined Markov Chain Approach




### RULES

### Description

Rules is an external data file that defines the charts. For the combined Markov chain approach a user would save two sets of runs rules defining the two charts.

### Program Description: PGM1B.FOR

This program finds the states of a Markov chain representation of a Shewhart control chart with supplementary runs rules of the form T(k,m,a,b). The rules defining the chart are read in from an external file called "RULES". Routines find the regions of the chart, the length of the state vector and the pointers to the end of the sub vector associated with each runs rule. The value of the present indicator variable is determined by runs rule and region combinations. The program then finds the state to state transitions by regions and sorts the states in ascending order of their base two representations. Any duplicate states are removed. Then the next-state transitions are numbered. Finally, the program outputs the state, next-state transition matrix.

**Program Listing**

```
C***************************************************************
C* PROGRAM 1 (PGM 1B)                                          *
C* THIS PROGRAM FINDS THE STATES OF A MARKOV CHAIN             *
C* REPRESENTATION OF A SHEWHART CONTROL CHART WITH             *
C* SUPPLEMENTARY RUNS RULES OF THE FORM T(K,M,A,B).            *
C***************************************************************
C*
      CHARACTER*80 FMT
      INTEGER CC,DIM,DIMC
      PARAMETER(DIM=100,DIMC=200)
      INTEGER H,I,J,L,NS(2),NSC,NR(2),NRC,MR,CK,CX,
     &  QH,SG,NT,NV,TMP,QQNS,K(20),M(20),D(20),
     &  PS(58),NX(58),X(20,10),Q(2,DIM,10),QQ(DIM),
     &  QC(DIMC,50),S(20)
      REAL A(20),B(20),R(2,41),TP
C*
C***************************************************************
C* THIS ROUTINE INPUTS THE RULES DEFINING THE CHART.           *
C***************************************************************
C*
      OPEN (50,FILE='RULES',STATUS='OLD')
      DO 899 CC=1,2
      READ(50,50) FMT
   50 FORMAT(A80)
      READ(50,FMT) NT
      DO 101 I=1,NT
        READ(50,FMT) K(I),M(I),A(I),B(I)
  101 CONTINUE
C*
C***************************************************************
C* THIS ROUTINE FINDS THE REGIONS OF THE CHART.                *
C***************************************************************
C*
      R(CC,1) = -9
      DO 201 I=1,NT
        R(CC,2*I)   = A(I)
        R(CC,2*I+1) = B(I)
  201 CONTINUE
```

```
    R(CC,2*NT+2) = +9
    MR = 2*NT+1
    NR(CC) = MR
  202 CK = 0
    DO 204 J=1,MR
      IF (R(CC,J).EQ.R(CC,J+1).AND.J.LE.NR(CC)) THEN
       DO 203 L=J,NR(CC)
         R(CC,L) = R(CC,L+1)
  203     CONTINUE
       NR(CC) = NR(CC)-1
       CK = 1
      ENDIF
      IF (R(CC,J).GT.R(CC,J+1)) THEN
       TP = R(CC,J)
       R(CC,J) = R(CC,J+1)
       R(CC,J+1) = TP
       CK = 1
      ENDIF
  204 CONTINUE
    MR = MR-1
    IF (MR.GT.NR(CC)) MR=NR(CC)
    IF (CK.EQ.1.AND.MR.GE.1) GOTO 202
C*
C****************************************************************
C* THIS ROUTINE FINDS THE LENGTH OF THE STATE VECTOR            *
C* AND THE POINTERS TO THE  END OF THE SUBVECTOR                *
C* ASSOCIATED WITH EACH RUNS RULE.                              *
C****************************************************************
C*
    CK = 0
    NV = 0
    DO 301 I=1,NT
      NV = NV+M(I)-1
      IF (K(I).LT.M(I)) CK = 1
  301 CONTINUE
    IF (CK.EQ.0) NV = NV+1
    D(1) = M(1)-1
    DO 302 I=2,NT
      D(I) = D(I-1)+M(I)-1
      IF (M(I).EQ.1) D(I) = 0
  302 CONTINUE
C*
C****************************************************************
C* THIS ROUTINE DETERMINES THE VALUE OF THE PRESENT             *
C* INDICATOR  VARIABLE BY RUNS RULE AND REGION                  *
C* COMBINATION.                                                 *
C****************************************************************
C*
    DO 402 I=1,NT
      DO 401 J=1,NR(CC)
        X(I,J) = 0
        IF (A(I).LE.R(CC,J).AND.R(CC,J+1).LE.B(I))
```

```
   &      X(I,J) = 1
 401   CONTINUE
 402 CONTINUE
C*
C*****************************************************************
C* THIS ROUTINE DETERMINES THE STATE TO STATE                     *
C* TRANSITIONS BY REGIONS.                                        *
C*****************************************************************
C*
     QQ(1) = 0
     QQNS = 2**NV-1
     NS(CC) = 1
     H = 1
 500 QH = QQ(H)
     DO 501 L=1,NV
       PS(L) = QH-2*(QH/2)
       QH = QH/2
 501 CONTINUE
     DO 503 I=1,NT
       S(I) = 0
       IF (M(I).GT.1) THEN
         DO 502 L=D(I)-M(I)+2,D(I)
           S(I) = S(I)+PS(L)
 502     CONTINUE
       ENDIF
 503 CONTINUE
     DO 509 J=1,NR(CC)
       SG = 0
       DO 506 I=1,NT
         IF (SG.EQ.0) THEN
           IF (S(I)+X(I,J).GE.K(I)) THEN
             SG = 1
           ELSE
             IF (M(I).GT.1) NX(D(I)-M(I)+2)=X(I,J)
             IF (M(I).GT.2) THEN
               DO 504 L=D(I)-M(I)+3,D(I)
                 NX(L) = PS(L-1)
 504           CONTINUE
             ENDIF
           ENDIF
           IF (X(I,J).EQ.0.AND.M(I).GT.1) THEN
             TMP = S(I)-PS(D(I))+1
             L = D(I)
             CK = 0
 505           IF (NX(L).EQ.1) THEN
               CK = 1
               IF (TMP.LT.K(I)) THEN
                 NX(L) = 0
                 TMP = TMP-1
                 CK = 0
               ENDIF
             ENDIF
```

```
             L = L-1
             TMP = TMP+1
             IF (CK.EQ.0.AND.L.GE.D(I)-M(I)+2)
     &          GOTO 505
           ENDIF
         ENDIF
  506   CONTINUE
        IF (SG.EQ.0) THEN
          QH = NX(1)
          DO 507 L=2,NV
            QH = QH+NX(L)*(2**(L-1))
  507     CONTINUE
          CK = 0
          DO 508 L=1,NS(CC)
            IF (CK.EQ.0.AND.QH.EQ.QQ(L)) THEN
              Q(CC,H,J) = QQ(L)
              CK = 1
            ENDIF
  508     CONTINUE
          IF (CK.EQ.0) THEN
            NS(CC) = NS(CC)+1
            QQ(NS(CC)) = QH
            Q(CC,H,J) = QH
          ENDIF
        ELSE
          Q(CC,H,J) = QQNS
        ENDIF
  509 CONTINUE
      H = H+1
      IF (H.LE.NS(CC)) GOTO 500
      NS(CC) = NS(CC)+1
      QQ(NS(CC)) = QQNS
      DO 510 J=1,NR(CC)
        Q(CC,NS(CC),J) = QQNS
  510 CONTINUE
C*
C*****************************************************************
C* THIS ROUTINE SORTS THE STATES IN ASCENDING ORDER              *
C* OF THEIR BASE TWO REPRESENTATIONS.                            *
C*****************************************************************
C*
  600 H = 0
      CK = 0
      DO 602 I=2,NS(CC)-H
        IF (QQ(I-1).GT.QQ(I)) THEN
          CK = 1
          TMP = QQ(I-1)
          QQ(I-1) = QQ(I)
          QQ(I) = TMP
          DO 601 J=1,NR(CC)
            TMP = Q(CC,I-1,J)
            Q(CC,I-1,J) = Q(CC,I,J)
```

```
         Q(CC,I,J) = TMP
 601    CONTINUE
       ENDIF
 602 CONTINUE
     H = H+1
     IF (CK.EQ.1) GOTO 600
C*
C*
C****************************************************************
C* THIS ROUTINE REMOVES ANY DUPLICATE STATES.                   *
C****************************************************************
C*
 700 CK = 0
     I = 1
 701 H = I+1
 702 CX = 0
     DO 703 J=1,NR(CC)
       IF (Q(CC,I,J).NE.Q(CC,H,J)) CX=1
 703 CONTINUE
     IF (CX.EQ.0) THEN
       TMP = QQ(H)
       DO 705 L=1,H-1
         DO 704 J=1,NR(CC)
           IF (Q(CC,L,J).EQ.TMP) Q(CC,L,J)=QQ(I)
 704    CONTINUE
 705   CONTINUE
       DO 707 L=H,NS(CC)-1
         QQ(L) = QQ(L+1)
         DO 706 J=1,NR(CC)
           Q(CC,L,J) = Q(CC,L+1,J)
           IF (Q(CC,L,J).EQ.TMP) Q(CC,L,J)=QQ(I)
 706    CONTINUE
 707   CONTINUE
       NS(CC) = NS(CC)-1
       CK = 1
     ENDIF
     H = H+1
     IF (H.LT.NS(CC)) GOTO 702
     I = I+1
     IF (I.LT.NS(CC)-1) GOTO 701
     IF (CK.EQ.1) GOTO 700
C*
C****************************************************************
C* THIS ROUTINE NUMBERS THE NEXT-STATE TRANSITIONS.             *
C****************************************************************
C*
     DO 803 I=1,NS(CC)
       DO 802 J=1,NR(CC)
         IF (Q(CC,I,J).LT.QQNS) THEN
           CK = 0
           L = 1
 801         IF (Q(CC,I,J).EQ.QQ(L)) THEN
```

```
              Q(CC,I,J) = L
              CK = 1
            ENDIF
            L = L+1
            IF (CK.EQ.0.AND.L.LT.NS(CC)) GOTO 801
          ELSE
            Q(CC,I,J) = NS(CC)
          ENDIF
  802   CONTINUE
  803 CONTINUE
899   CONTINUE
      CLOSE (50)
C*
C**************************************************************
C* THIS ROUTINE CREATES THE STATE, NEXT-STATE                   *
C* TRANSITION MATRIX FOR THE COMBINED CHART                     *
C**************************************************************
C*
      NSC=(NS(1)-1)*(NS(2)-1)+1
      NRC=NR(1)*NR(2)
      DO 904 I1=1,NS(1)-1
       DO 903 I2=1,NS(2)-1
        I=(I1-1)*(NS(2)-1)+I2
        DO 902 J1=1,NR(1)
         DO 901 J2=1,NR(2)
          J=(J1-1)*NR(2)+J2
          QC(I,J)=(Q(1,I1,J1)-1)*(NS(2)-1)+Q(2,I2,J2)
          IF (Q(1,I1,J1).GE.NS(1)) QC(I,J)=NSC
          IF (Q(2,I2,J2).GE.NS(2)) QC(I,J)=NSC
901      CONTINUE
902     CONTINUE
903    CONTINUE
904   CONTINUE
      DO 906 J1=1,NR(1)
       DO 905 J2=1,NR(2)
        J=(J1-1)*NR(2)+J2
        QC(NSC,J)=NSC
905    CONTINUE
906   CONTINUE
C*
C**************************************************************
C* THIS ROUTINE OUTPUTS THE STATE, NEXT-STATE                   *
C* TRANSITION MATRIX.                                           *
C**************************************************************
C*
 9900 OPEN (60,FILE='PGM1.OUT',STATUS='OLD')
      WRITE(60,9961) NSC,NRC,NR(1),NR(2)
 9961 FORMAT(4(I4,1X))
      DO 9963 CC=1,2
       WRITE(60,9962) (R(CC,J),J=1,NR(CC)+1)
 9962  FORMAT(10(F9.5))
 9963 CONTINUE
```

```
     DO 9965 I=1,NSC
       WRITE(60,9964) I,(QC(I,J),J=1,NRC)
 9964   FORMAT(50(1X,I4))
 9965 CONTINUE
     CLOSE (60)
C*
     STOP
     END
```

**PGM1.OUT**

For the charts given in "**RULES**", the program PGM1B.FOR writes the Markov chain representation for each chart into this external output file.

**Program Description: PGM2B.FOR**

This program calculates the average run length (ARL), the standard deviation (STD), and selected percentage points of the run length distribution. The Markov chain representation of the chart is read in from an external file called "PGM1.OUT". A routine calculates the average run lengths and standard deviations of the chart for various positive standardized shifts in the mean. Various percentage points are then calculated. The program makes use of Algorithm AS 126 Applied Statistics (1978) Vol. 27, No. 2 which computes the probability of the normal range given T, the upper limit of integration, and N, the sample size. Subroutines included are a double precision function **dnml(x)** which computes the cumulative distribution function $P[y \leq x]$ of a random variable y having a standard normal distribution.

**Program Listing**

```
C****************************************************************
C* PROGRAM 2 (PGM 2) (STANDARDIZED REGIONS)                     *
C* THIS PROGRAM CALCULATES THE  AVERAGE  RUN LENGTH             *
C* (ARL), THE STANDARD DEVIATION (STD), AND SELECTED            *
C* PERCENTAGE POINTS OF THE RUN LENGTH DISTRIBUTION.            *
C****************************************************************
C*
      INTEGER CC,CV(11,11,9),DX,DV,I,IFAULT,II,J,J1,
```

```
     &  J2,K,N,NCP,NDX,NDV,NSC,M,NRC,NR(2),Q(100,50),
     &  QQ(100),SS,STEPX,STEPV
      DOUBLE PRECISION ARL(11,11),CDF(50),CP(9),CUM,
     &  DNML,D2(25),D3(25),FCDF,L(100),LH,LP,LX,LV,
     &  P(50),PCDF,PX(10),PV(10),R(2,10),RNGPI,
     &  STD(11,11),U(100),ZA,ZB
C*
      SS=1
      STEPX=1
      STEPV=1
      NCP=9
      CP(1)=0.01
      CP(2)=0.05
      CP(3)=0.10
      CP(4)=0.25
      CP(5)=0.50
      CP(6)=0.75
      CP(7)=0.90
      CP(8)=0.95
      CP(9)=0.99
C*
      D2(2)=1.1283791671D0
      D2(3)=1.6925687506D0
      D2(4)=2.0587507460D0
      D2(5)=2.3259289473D0
      D2(6)=2.5344127212D0
      D2(7)=2.7043567512D0
      D2(8)=2.8472006121D0
      D2(9)=2.9700263244D0
      D2(10)=3.0775054617D0
      D2(11)=3.1728727038D0
      D2(12)=3.2584552798D0
      D2(13)=3.3359803541D0
      D2(14)=3.4067631082D0
      D2(15)=3.4718268899D0
      D2(16)=3.5319827861D0
      D2(17)=3.5878839618D0
      D2(18)=3.6400637579D0
      D2(19)=3.6889630232D0
      D2(20)=3.7349501196D0
      D2(21)=3.7783358298D0
      D2(22)=3.8193846434D0
      D2(23)=3.8583234233D0
      D2(24)=3.8953481485D0
      D2(25)=3.9306292195D0
C*
      D3(2)=0.7267604553D0
      D3(3)=0.7891977107D0
      D3(4)=0.7740624738D0
      D3(5)=0.7466376009D0
      D3(6)=0.7191713092D0
      D3(7)=0.6942311313D0
```

```
      D3(8)=0.6721236717D0
      D3(9)=0.6525962151D0
      D3(10)=0.6352897762D0
      D3(11)=0.6198643117D0
      D3(12)=0.6060285277D0
      D3(13)=0.5935411244D0
      D3(14)=0.5822042445D0
      D3(15)=0.5718557265D0
      D3(16)=0.5623621426D0
      D3(17)=0.5536130572D0
      D3(18)=0.5455164487D0
      D3(19)=0.5379951043D0
      D3(20)=0.5309837904D0
      D3(21)=0.5244270274D0
      D3(22)=0.5182773314D0
      D3(23)=0.5124938181D0
      D3(24)=0.5070410861D0
      D3(25)=0.5018883188D0
C*
      DO 1 I=2,25
        D3(I) = DSQRT(D3(I))
1     CONTINUE
C*
      WRITE(*,*) 'INPUT # OF SHIFTS IN THE MEAN.'
      READ(*,*) NDX
      WRITE(*,*) 'INPUT # OF SHIFTS IN THE STDEV.'
      READ(*,*) NDV
      WRITE(*,*) 'INPUT SAMPLE SIZE.'
      READ(*,*) M
C*
C*****************************************************************
C* THIS ROUTINE INPUTS INFORMATION ABOUT THE CHART.              *
C*****************************************************************
C*
      OPEN(50,FILE='PGM1.OUT',STATUS='OLD')
      READ(50,51) NSC,NRC,NR(1),NR(2)
51    FORMAT(4(I4,1X))
      DO 53 CC=1,2
        READ(50,52) (R(CC,J),J=1,NR(CC)+1)
52      FORMAT(50(F9.5))
53    CONTINUE
      DO 55 I=1,NSC
        READ(50,54) II,(Q(I,J),J=1,NRC)
54      FORMAT(50(1X,I4))
55    CONTINUE
      CLOSE (50)
C*
C*
C*
C*
C*****************************************************************
C* THIS ROUTINE CALCULATES THE AVERAGE  RUN LENGTHS              *
```

```
C* AND STANDARD DEVIATIONS OF THE CHART FOR VARIOUS          *
C* POSITIVE STANDARDIZED SHIFTS IN THE MEAN.                 *
C*************************************************************
C*
     DO 116 DX=0,NDX,STEPX
     DO 115 DV=0,NDV,STEPV
C*
      DO 101 J1=1,NR(1)
        ZA=(R(1,J1)-DX/10.D0)/(1.D0+DV/10.D0)
        ZB=(R(1,J1+1)-DX/10.D0)/(1.D0+DV/10.D0)
        PX(J1)=DNML(ZB)-DNML(ZA)
101    CONTINUE
C*
      DO 102 J2=1,NR(2)
        ZA=(D2(M)+R(2,J2)*D3(M))/(1.D0+DV/10.D0)
        IF (ZA.LT.0.D0) ZA=0.D0
        ZB = (D2(M)+R(2,J2+1)*D3(M))/(1.D0+DV/10.D0)
        IF (ZB.LT.0.D0) ZB=0.D0
        PV(J2)=RNGPI(ZB,M,IFAULT)
    &              -RNGPI(ZA,M,IFAULT)
102     CONTINUE
C*
      DO 104 J1=1,NR(1)
        DO 103 J2=1,NR(2)
          J=(J1-1)*NR(2)+J2
          P(J)=PX(J1)*PV(J2)
103       CONTINUE
104    CONTINUE
C*
       DO 106 I=1,NSC-1
         U(I) = 0.D0
         DO 105 J=1,NRC
           IF (Q(I,J).NE.NSC) U(I)=U(I)+P(J)
105        CONTINUE
         U(I)=1.D0-U(I)
106     CONTINUE
       CUM = U(1)
       ARL(DX+1,DV+1) = CUM
       STD(DX+1,DV+1) = CUM
       CDF(0+1) = 0.D0
       CDF(1+1) = CUM
       CK = 0
       N = 1
107     N = N+1
       DO 109 I=1,NSC-1
         L(I) = 0.D0
         DO 108 J=1,NRC
           IF (Q(I,J).NE.NSC) L(I)=L(I)+P(J)*U(Q(I,J))
108        CONTINUE
109     CONTINUE
       IF (U(1).NE.0.0.AND.CUM.NE.1.0) THEN
         LH = L(1)/U(1)
```

```
        LP = (1-CUM-L(1))/(1-CUM)
        TP = DABS(LH-LP)
        IF (N.GT.9.AND.TP.LT.0.000001) CK=1
      ENDIF
      IF (N.GT.40) CK=1
      ARL(DX+1,DV+1) = ARL(DX+1,DV+1)+N*L(1)
      STD(DX+1,DV+1) = STD(DX+1,DV+1)+N*N*L(1)
      IF (CK.EQ.1) THEN
        TP = N/(1-LP) + 1/((1-LP)*(1-LP))
        ARL(DX+1,DV+1) = ARL(DX+1,DV+1)+LP*L(1)*TP
        TP = 1-LP
        TP = N*N/TP+(2*N-1)/(TP*TP)+2/(TP*TP*TP)
        STD(DX+1,DV+1) = STD(DX+1,DV+1)+LP*L(1)*TP
      ENDIF
      DO 110 I=1,NSC-1
        U(I) = L(I)
110     CONTINUE
      CUM = CUM + L(1)
      CDF(N+1) = CUM
      IF (CK.EQ.0) GOTO 107
      STD(DX+1,DV+1) = STD(DX+1,DV+1)-ARL(DX+1,DV+1)*ARL(DX+1,DV+1)
      STD(DX+1,DV+1) = DSQRT(STD(DX+1,DV+1))
C*
C***************************************************************
C* CALCULATION OF VARIOUS PERCENTAGE POINTS.                   *
C***************************************************************
C*
      I = 1
      DO 112 J=1,N
111       CK = 0
        IF (CDF(J+1).GE.CP(I)) THEN
          CV(DX+1,DV+1,I) = J
          I = I+1
          CK = 1
        ENDIF
        IF (CK.EQ.1.AND.I.LE.NCP) GOTO 111
112     CONTINUE
      FCDF = CDF(N+1)
      TP = CDF(N+1)-CDF(N)
      J = 1
113     IF (I.GT.NCP) GOTO 115
      PCDF = CDF(N+1)+LH*TP*(1-LH**J)/(1-LH)
114     CK = 0
      IF (PCDF.GE.CP(I)) THEN
        CV(DX+1,DV+1,I) = N+J
        I =I+1
        CK = 1
      ENDIF
      IF (CK.EQ.1.AND.I.LE.NCP) GOTO 114
      J = J+1
      GOTO 113
115   CONTINUE
```

```
116   CONTINUE
C*
C***************************************************************
C* THIS ROUTINE PRINTS THE AVERAGE RUN LENGTHS (ARL),         *
C* STANDARD DEVIATIONS, AND PERCENTAGE POINTS BY              *
C* STANDARDIZED SHIFTS IN THE MEAN.                           *
C***************************************************************
C*
200   WRITE(*,*) ARL(1,1)
      OPEN (60,FILE='PGM2.OUT',STATUS='OLD')
      WRITE(60,201) (CP(I),I=1,NCP)
201   FORMAT(40X,'PERCENTILES'/
     &       2X,'D',4X,'L',5X,'ARL',5X,'STD',
     &      4X,9(2X,F3.2))
      DO 204 DX=0,NDX,STEPX
      DO 203 DV=0,NDV,STEPV
        WRITE(60,202) DX/10.0,1.D0+SS*DV/10.D0,
     &   ARL(DX+1,DV+1),STD(DX+1,DV+1),
     &   (CV(DX+1,DV+1,I),I=1,NCP)
202     FORMAT(2(F4.2,1X),F7.2,1X,F7.2,3X,9(1X,I4))
203   CONTINUE
204   CONTINUE
      CLOSE (60)
999   STOP
      END
C*
C***************************************************************
C* ALGORITHM AS 126 APPLIED STATISTICS (1978)                 *
C* VOL. 27, NO. 2                                             *
C*                                                            *
C* COMPUTES THE PROBABILITY OF THE NORMAL RANGE               *
C* GIVEN T, THE UPPER LIMIT OF INTEGRATION,  AND  N,          *
C* THE SAMPLE SIZE.                                           *
C***************************************************************
C*
      DOUBLE PRECISION FUNCTION RNGPI(T,N,IFAULT)
      INTEGER N,I,IFAULT
      DOUBLE PRECISION A,B,C,DNML,G(8),H(8),RISF,
     &      T,X,XL,Y
C*
      DATA G(1),G(2),G(3),G(4),G(5),G(6),G(7),G(8)
     & /0.4947004675, 0.4722875115, 0.4328156012,
     &  0.3777022042, 0.3089381222, 0.2290083888,
     &  0.1408017754, 0.04750625492/
C*
      DATA H(1),H(2),H(3),H(4),H(5),H(6),H(7),H(8)
     & /0.01357622971, 0.03112676197, 0.04757925584,
     &  0.06231448563, 0.07479799441, 0.08457825969,
     &  0.09130170752, 0.09472530523/
C*
      RISF(X)=0.3989422804*EXP(-0.5*X*X)*
     &       (DNML(X)-DNML(X-T))**(N-1)
```

```
C*
      IFAULT=1
      RNGPI=0.D0
      IF (T.LE.0.D0.OR.N.LE.1) RETURN
      IFAULT=0
      XL=0.5D0*T
      A=0.5D0*(8.D0+XL)
      B=8.D0-XL
      Y=0.D0
      DO 1 I=1,8
        C=B*G(I)
        Y=Y+H(I)*(RISF(A+C)+RISF(A-C))
1     CONTINUE
      RNGPI=(2.D0*(DNML(XL)-0.5D0))**N+2.D0*B*Y*N
      IF (RNGPI.GT.1.D0) RNGPI=1.D0
      RETURN
      END
C*
C*
C*
      DOUBLE PRECISION FUNCTION DNML(X)
C*
C*
C*
C*****************************************************************
C* COMPUTES THE CUMULATIVE DISTRIBUTION FUNCTION               *
C* P(Y<=X) OF A RANDOM VARIABLE Y HAVING A                     *
C* STANDARD NORMAL DISTRIBUTION.                               *
C*****************************************************************
C*
      DOUBLE PRECISION X,Y,S,RN,ZERO,ONE,ERF,SQRT2,PI
      DATA SQRT2,ONE/1.414213562373095,1.D0/
      DATA PI,ZERO/3.141592653589793,0.D0/
      Y=X/SQRT2
      IF (X.LT.ZERO) Y=-Y
      S=ZERO
      DO 1 N=1,37
      RN=N
      S=S+DEXP(-RN*RN/25)/N*DSIN(2*N*Y/5)
1     CONTINUE
      S=S+Y/5
      ERF=2*S/PI
      DNML=(ONE+ERF)/2
      IF (X.LT.ZERO) DNML=(ONE-ERF)/2
      IF (X.LT.-8.3D0) DNML=ZERO
      IF (X.GT.8.3D0) DNML=ONE
      RETURN
      END
```

**PGM2.OUT**

The program PGM2B.FOR prints the average run lengths (ARLs), standard deviations, and percentage points by standardized shifts in the mean into this external data file.