Chapter 7

Control charts for the spread

In the previous chapters, we considered control charts for the mean of individual serially correlated measurements. In Section 2.3, we mentioned that a control chart for the mean is in most cases accompanied by a control chart for the dispersion of the process outcomes. In this chapter, control charts for the spread of serially correlated data will be discussed. In the first section, we will focus on control charts for the spread of individual observations. In this case, it is not possible to estimate the spread on a given time point. The moving range control chart and so-called omnibus methods are discussed that are designed to circumvent this problem. In Sections 7.2 through 7.8, we will consider the case where the serially correlated data is sub grouped into samples, so that it becomes feasible to monitor estimators of the within-sample variance for a shift in the spread of the process.

As in the previous chapters, the discussion will be focused on a special case of serially correlated processes, viz. the AR(1) process. As argued earlier, the reason for only considering this type of serially correlated processes is that the AR(1) process is frequently encountered in practice. Furthermore, it may serve as an approximation to other time series models. However, the theory is easily adapted to include other time series models as well.

7.1 Control charts for the spread when $n = 1$

7.1.1 Introduction

In the previous chapters, we considered control charts for the mean of individual serially correlated measurements. There the variance of the process was assumed to be constant. The discussion focused on the efficiency of
detecting a shift in the mean of the process. In this section, we will assume that the mean of the process is constant. We only consider special causes that shift the variance of the process from $\sigma_Y^2$ to $b\sigma_Y^2$, where $b > 1$.

In the case of independently distributed individual observations, there is some controversy over the question of whether or not to use a separate chart for the spreading. There is a group of authors, such as Duncan (1986), Wheeler and Chambers (1990), Wetherill and Brown (1991) and Montgomery (1996), who recommend to use both a control chart for the mean, and a control chart for a measure of spread.

There are, however, other authors such as Nelson (1982, 1990) and Roes, Does and Schurink (1993) who argue that all the information is already contained within the $X$-chart. For example, when an increase in the variance of a process occurs, the limits on the control chart for the mean will be too tight. Hence, the probability of crossing one of the control limits of a control chart for the mean is increasing with $b > 1$. Therefore, the question has been raised whether a control chart for the spread adds extra power to a control chart for the mean when a shift in the spread of the process is to be detected.

For the case of individual observations, a chart for the spread could be based on the moving range ($MR$). The moving range at time $t$ is defined as

$$MR_t = |X_t - X_{t-1}|$$

where $\{X_t\}$ is the sequence of independently distributed individual observations that is to be monitored. Roes, Does, and Schurink (1993) computed the conditional probability (assuming independence of the observations) of observing a signal on the $MR$-chart, given that an $X$-chart for the mean does not signal. For selected shifts in the mean and shifts in the variance, the probability

$$P(MR_t \text{ signals } |E),$$

(7.1)

is considered, where $E$ is the event that the $X$-chart for the mean does not signal on time $t$ and time $t - 1$. In Table 7.1, the probabilities given by (7.1) are tabulated for various shifts in the mean and in the variance.

The values of $P(MR_t \text{ signals } |E)$ in Table 7.1 differ from those of Table 7 of Roes, Does, and Schurink (1993). Roes, Does, and Schurink (1993) neglected the lower control limit of the $MR$-chart in the computation of $P(MR_t \text{ signals } |E)$. The effect of ignoring the lower control limit is largest
7.1. CONTROL CHARTS FOR THE SPREAD WHEN $N = 1$

Table 7.1: Conditional probabilities of an out-of-control signal on the $MR$-chart.

| Shift in the mean | $P(MR_t \text{ signals } | E)$ | $P(X_t \text{ signals on } X \text{ chart})$ |
|-------------------|-------------------------------|---------------------------------|
| $\delta = 0.0$    | 0.00158                       | 0.0020                          |
| $\delta = 0.5$    | 0.00191                       | 0.0050                          |
| $\delta = 1.0$    | 0.00305                       | 0.0183                          |
| $\delta = 1.5$    | 0.00523                       | 0.0559                          |
| $\delta = 2.0$    | 0.00846                       | 0.1378                          |
| $\delta = 2.5$    | 0.01241                       | 0.2775                          |
| $\delta = 3.0$    | 0.01662                       | 0.4641                          |

| Shift in the spread | $P(MR_t \text{ signals } | E)$ | $P(X_t \text{ signals on } X \text{ chart})$ |
|---------------------|-------------------------------|---------------------------------|
| $b = 1.0$           | 0.00158                       | 0.0020                          |
| $b = 1.5$           | 0.00336                       | 0.0394                          |
| $b = 2.0$           | 0.00492                       | 0.1223                          |
| $b = 2.5$           | 0.00592                       | 0.2164                          |
| $b = 3.0$           | 0.00654                       | 0.3020                          |

for the in-control situations. Roes, Does, and Schurink (1993) report $P(MR_t \text{ signals } | E) = 0.00058$ for $\delta = 0.0$ and for $b = 1$. The difference becomes smaller for larger shifts in the mean or the variance.

The probabilities $P(MR_t \text{ signals } | E)$ are small for the out-of-control situations. Therefore, in Roes, Does and Schurink (1993) it is concluded that the contribution of the $MR$-chart to the power of discovering an out-of-control situation is small.

We will arrive at the same conclusion, using another argument. The fact that all the values of $P(MR_t \text{ signals } | E)$ are small could be the result of a poor design of the control charts, see Amin and Ethridge (1998). For example, the control limits of the $MR$-chart may be too wide. To properly assess the added value of the $MR$-chart, values of $P(MR_t \text{ signals } | E)$ for the out-of-control situations must be compared to the value of the same probability, evaluated in the in-control situation. It is the increase in power for detection of out-of-control situations, in comparison to the in-control situation, that makes a control chart sensitive in signalling special causes of variation.

The probability of an out-of-control signal on an $X$-chart when there
has been a shift in the mean of size $3\sigma_X$ is about 230 times as large as the in-control probability of an out-of-control signal. Given a shift in the mean of $3\sigma_X$, the conditional probability of a signal on the $MR$-chart is increased with a factor of about 10, compared to the in-control situation.

When the standard deviation of the process shifts from $\sigma_X$ to $3\sigma_X$, the probability of an out-of-control signal on the $X$-chart is increased with a factor of about 150, in comparison to the in-control situation. The conditional probability of an out of control signal on the $MR$-chart is increased by a factor smaller than 4.

Hence, in agreement with Roes, Does, and Schurink (1993), we conclude that the $MR$-chart adds very little to the power of an $X$-chart when there has been a shift in the mean, and more surprisingly, the contribution to the power is even less when there has been a shift in the process spread.

The approach of Roes, Does and Schurink (1993) was criticized by Adke and Hong (1997). They consider the probabilities

$$P[\text{MR chart signals between } t+1 \text{ and } t+n | \text{X chart does not signal between } t+1 \text{ and } t+n]$$

for various values of $n$. For $n = 2$, these probabilities are compared to those computed by Roes, Does, and Schurink (1993). The differences arise due to the fact that Roes, Does and Schurink (1993) assume that a shift in one of the process parameters occurs between $X_{t-1}$ and $X_t$, whereas Adke and Hong (1998) assume that both $X_{t-1}$ and $X_t$ are drawn from an out-of-control distribution. In Amin and Ethridge (1998) it was suggested that the differences are also caused by the fact that Adke and Hong used a two-sided $MR$-chart, whereas Roes, Does and Schurink (1993) essentially used a one-sided $MR$-chart. However, the probabilities computed by Adke and Hong (1997) are also based on a one-sided $MR$-chart. In Table 7.2, conditional probabilities of a signal on different $MR$-charts are tabulated for various shifts in the variance, for the case where only $X_t$ is out of control and for the case where both $X_{t-1}$ and $X_t$ are out of control.

The second column of Table 7.2 displays the probabilities of an out-of-control signal on a one-sided $MR$-chart when only $X_t$ is out of control. In the third column the probabilities corresponding to a two-sided $MR$-chart are tabulated for the case where only $X_t$ is out of control. The fourth column contains the probabilities of a signal on a one-sided $MR$-chart when both $X_{t-1}$ and $X_t$ are out of control. In the fifth column the probabilities
corresponding to a two-sided $MR$-chart are displayed for the case that both $X_{t-1}$ and $X_t$ are out of control.

### Table 7.2: Conditional probabilities of a signal on a $MR$-chart at time $t$, given that $X_{t-1}$ and $X_t$ do not signal on the $X$-chart.

<table>
<thead>
<tr>
<th>$b$</th>
<th>$X_t$ out of control</th>
<th>$X_{t-1}$ and $X_t$ out of control</th>
<th>$X_{t-1}$ and $X_t$ out of control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>one-sided $MR$</td>
<td>two-sided $MR$</td>
<td>one-sided $MR$</td>
</tr>
<tr>
<td>1.0</td>
<td>0.00058</td>
<td>0.00158</td>
<td>0.00058</td>
</tr>
<tr>
<td>1.5</td>
<td>0.00254</td>
<td>0.00336</td>
<td>0.00939</td>
</tr>
<tr>
<td>2.0</td>
<td>0.00420</td>
<td>0.00492</td>
<td>0.02257</td>
</tr>
<tr>
<td>2.5</td>
<td>0.00525</td>
<td>0.00592</td>
<td>0.03294</td>
</tr>
<tr>
<td>3.0</td>
<td>0.00590</td>
<td>0.00654</td>
<td>0.04009</td>
</tr>
</tbody>
</table>

The probabilities in Table 7.2 show that $P(MR_t < LCL_{MR})$ is not extremely small, and cannot be neglected when computing the probability of a signal on a two-sided $MR$-chart.

Let us consider the difference between the probability of a signal on the $MR$-chart in the in-control situation and the probability of a signal in an out-of-control situation. This difference, which is a measure for the added power of the $MR$-chart, becomes smaller when a lower control limit is added to a one-sided $MR$-chart. Therefore, when it is decided to use both an $X$-chart and an $MR$-chart to monitor independent individual observations, it is advisable to use a one-sided $MR$-chart.

The foregoing analysis can also be performed for serially correlated observations. However, the probabilities of an out-of-control signal depend on the state of the process. Instead of comparing single probabilities as in the i.i.d. case described above, functions of the state of the process must be compared. This does not provide a clear argument whether or not to use an $MR$-chart in addition to an $X$-chart in the case of serially correlated individual observations. In addition, for the case of independently distributed individual observations, we argued that using an $X$-chart alone is nearly as effective as using both an $X$-chart and an $MR$-chart. Amin and Ethridge (1998) arrive at the same conclusion, by comparing the ARL of the $X$-chart alone to the ARL of the combined $X$-$MR$ procedure. When the observations of the process are serially correlated, it is not to be expected either that a control chart based on moving ranges adds extra power to a control chart for the mean. Therefore, we advise to use only a control chart
for the mean to monitor a sequence of serially correlated observations.

However, Amin and Ethridge (1998) mention other considerations that justify the use of a combined procedure. For this reason, control charts for the spread of individual observations are briefly discussed in the next two subsections. In the sections following Section 7.2, we discuss control charts for the spread when the data is subgrouped first.

### 7.1.2 The moving range chart when \( n = 1 \)

In the previous subsection, we discussed the \( MR \)-chart for independently distributed individual observations. In this subsection, individual AR(1) observations \( \{Y_t\} \) are considered, with

\[
Y_t - \mu = \phi(Y_{t-1} - \mu) + \varepsilon_t \quad \text{for} \ t \in \mathbb{Z},
\]

where the \( \varepsilon_t \) are independently distributed as \( \varepsilon_t \sim \mathcal{N}(0, \sigma_{\varepsilon_t}^2) \). Note that the mean of the process \( \mu \) is not indexed by \( t \), since we assume that the process is in control with respect to the mean. Instead, to indicate that a shift in the spread may occur, the variance of the disturbances is indexed by \( t \). The moving range at time \( t \) is computed as

\[
MR_t = |Y_t - Y_{t-1}|.
\]

Suppose that the process is not only in control with respect to the mean, but also with respect to the spread (\( \sigma_{\varepsilon_t}^2 \) is constant for all \( t \)). Then \( Y_t - Y_{t-1} \) follows a normal distribution with expectation zero and variance \( 2(1 - \phi)\sigma_Y^2 \). For the expectation of \( MR_t \) we note that \( MR_t^2/(2(1 - \phi)\sigma_Y^2) \) follows a \( \chi^2 \) distribution with one degree of freedom. Next we are interested in evaluating the expectation of \( Z^\alpha \), where \( Z \) follows a \( \chi^2 \) distribution with one degree of freedom.

\[
E(Z^\alpha) = \int_0^\infty z^\alpha \frac{e^{-\frac{1}{2}z}}{\sqrt{2\pi}\Gamma(\frac{1}{2})} \, dz
\]

\[
= 2^\alpha \frac{\Gamma(\alpha + \frac{1}{2})}{\Gamma(\frac{1}{2})} \int_0^\infty \left( \frac{1}{2} \right)^{\alpha + \frac{1}{2}} z^{\alpha - \frac{1}{2}} e^{-\frac{1}{2}z} \, dz
\]

\[
= 2^\alpha \frac{\Gamma(\alpha + \frac{1}{2})}{\Gamma(\frac{1}{2})}.
\]
7.1. CONTROL CHARTS FOR THE SPREAD WHEN $N = 1$

Hence, using $\alpha = 1/2$, we obtain the following well-known result for $E(MR_t)$:

$$E(MR_t) = \frac{2}{\sqrt{\pi}} \sqrt{1-\phi} \sigma_Y,$$

see also Cryer and Ryan (1990). The factor $2/\sqrt{\pi}$ is usually denoted by $d_2(2)$. Note that in the case $\phi = 0$, $MR_t/d_2(2)$ is an unbiased estimator of $\sigma_Y$. However, for $\phi \neq 0$, $MR_t/d_2(2)$ is biased for $\sigma_Y$ by a factor of $\sqrt{1-\phi}$. For negative $\phi$, $\sigma_Y$ is overestimated, whereas for $\phi > 0$, $\sigma_Y$ is underestimated.

When the mean of $MR_t/d_2(2)$ is used as an estimator for $\sigma_Y$, the bias does not disappear. This explains the most commonly encountered problems when monitoring serially correlated data for a shift in the mean. When there is positive autocorrelation, $\sigma_Y$ is underestimated, and the control limits become too tight. Hence, false alarms are generated. For negative autocorrelation, the control limits are unnecessary wide, so that significant shifts in the process mean may go undetected (see for example Maragah and Woodall (1992)). In Subsection 7.5, we will see that estimating $\sigma_Y$ on the basis of $S^2$ also leads to underestimation for negative $\phi$, and to overestimation for positive $\phi$. Therefore, determining the width of the control limits based on $S^2$ will lead to the same complaints. However, the bias in $S^2$ disappears when the number of observations is sufficiently large.

For the variance of $MR_t$ we have

$$\text{Var} (MR_t) = E (|Y_t - Y_{t-1}|^2) - (E(MR_t))^2$$

$$= 2(1-\phi) \sigma_Y^2 - \frac{4}{\pi} (1-\phi) \sigma_Y^2$$

$$= 2(1-\phi) \left(1 - \frac{2}{\pi}\right) \sigma_Y^2.$$

For the covariance between $MR_t$ and $MR_{t+k}$ we firstly note that

$$\text{Cov} (Y_t - Y_{t-1}, Y_{t+k} - Y_{t+k-1})$$

$$= \text{Cov} (Y_t, Y_{t+k}) - \text{Cov} (Y_t, Y_{t+k-1})$$

$$- \text{Cov} (Y_{t-1}, Y_{t+k}) + \text{Cov} (Y_{t-1}, Y_{t+k-1})$$

$$= -\phi^{k-1} (1-\phi)^2 \sigma_Y^2.$$
so that we have for \( \rho_k \), the correlation between \( (Y_t - Y_{t-1}) \) and \( (Y_{t+k} - Y_{t+k-1}) \)

\[
\rho_k = \frac{1}{2} \phi^{k-1} (1 - \phi). \tag{7.2}
\]

With (A5) of Cryer and Ryan (1990) (see also page 92 of Johnson and Kotz (1972)) we obtain

\[
\text{Cov}(MR_t, MR_{t+k}) = \frac{4(1 - \phi)}{\pi} \left( \sqrt{1 - \rho_k^2} + \rho_k \arcsin(\rho_k) - 1 \right) \sigma_t^2,
\]

where \( \rho_k \) is given in (7.2). Thus, \( \{\rho_{MR,k}\} \), the autocorrelation function of the moving ranges of AR(1) data, can be computed as

\[
\rho_{MR,k} = \frac{2}{\pi - 2} \left( \sqrt{1 - \rho_k^2} + \rho_k \arcsin(\rho_k) - 1 \right),
\]

where \( \rho_k \) is given in (7.2). In Table 7.3, the autocorrelations between \( MR_t \) and \( MR_{t+k} \) are determined for lags \( k = 1, 2, \cdots, 10 \) for selected values of \( \phi \).

**Table 7.3:** Autocorrelations between \( MR_t \) and \( MR_{t+k} \) for selected values of \( \phi \).

<table>
<thead>
<tr>
<th>( \phi )</th>
<th>-0.9</th>
<th>-0.6</th>
<th>-0.3</th>
<th>0</th>
<th>0.3</th>
<th>0.6</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k = 1 )</td>
<td>0.8809</td>
<td>0.5989</td>
<td>0.3852</td>
<td>0.2239</td>
<td>0.1084</td>
<td>0.0352</td>
<td>0.0022</td>
</tr>
<tr>
<td>( k = 2 )</td>
<td>0.6928</td>
<td>0.2060</td>
<td>0.0334</td>
<td>0</td>
<td>0.0097</td>
<td>0.0126</td>
<td>0.0018</td>
</tr>
<tr>
<td>( k = 3 )</td>
<td>0.5507</td>
<td>0.0732</td>
<td>0.0030</td>
<td>0</td>
<td>0.0009</td>
<td>0.0045</td>
<td>0.0014</td>
</tr>
<tr>
<td>( k = 4 )</td>
<td>0.4401</td>
<td>0.0262</td>
<td>0.0003</td>
<td>0</td>
<td>0.0001</td>
<td>0.0016</td>
<td>0.0012</td>
</tr>
<tr>
<td>( k = 5 )</td>
<td>0.3529</td>
<td>0.0094</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0006</td>
<td>0.0009</td>
</tr>
<tr>
<td>( k = 6 )</td>
<td>0.2837</td>
<td>0.0034</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0008</td>
</tr>
<tr>
<td>( k = 7 )</td>
<td>0.2284</td>
<td>0.0012</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0006</td>
</tr>
<tr>
<td>( k = 8 )</td>
<td>0.1842</td>
<td>0.0004</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0005</td>
</tr>
<tr>
<td>( k = 9 )</td>
<td>0.1486</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0004</td>
</tr>
<tr>
<td>( k = 10 )</td>
<td>0.1201</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

For positive \( \phi \), the autocorrelations of successive moving ranges are small. This can be explained by noting that for values of \( \phi \) close to one, \( MR_t \) approximately equals \( |\varepsilon_t| \), and the disturbances are assumed to be
7.1. CONTROL CHARTS FOR THE SPREAD WHEN N = 1

independent. For strong positive autocorrelation, the moving ranges can be treated as statistics that are not serially correlated. This facilitates the design and the evaluation of the ARL of the MR-chart, since the ARL can then be approximated (given lower control limit LCL\textsubscript{MR} and upper limit UCL\textsubscript{MR}) by

$$\text{ARL}_{MR} \approx \frac{1}{1 - P(LCL_{MR} < MR_t < UCL_{MR})}.$$  

Since the distribution of MR\textsubscript{t} is not known, Monte Carlo simulation can be used to obtain the probability P(LCL\textsubscript{MR} < MR\textsubscript{t} < UCL\textsubscript{MR}). The control limits can be determined by requiring a certain in-control ARL and a certain out-of-control ARL.

In the foregoing, the MR-chart for individual serially correlated observations is discussed. However, the chart can also be applied to residuals of a fitted time series model. If the time series model is appropriate for the data, the residuals will be approximately uncorrelated. Furthermore, if the time series model and its parameters are known, a shift in the variance of the serially correlated data is transferred fully to a shift in the variance of the residuals, so that the main objection to the use of residuals charts for the mean does not apply to residuals charts for the spread. Therefore, the results of the MR-chart of residuals can be interpreted as in the case of independently distributed observations. In our view, the use of an MR-chart of residuals is to be preferred to the use of an MR-chart of correlated observations.

7.1.3 Omnibus control charts when n = 1

In Domangue and Patch (1991), a so-called omnibus EWMA control chart is proposed. The control chart is based on plotting statics

$$A_t = \lambda \left| \frac{Y_t - \mu}{\sigma_Y} \right|^{\alpha} + (1 - \lambda)A_{t-1}. \quad (7.3)$$

Domangue and Patch (1991) have evaluated the ARL behavior of the omnibus EWMA control chart for \( \alpha = 0.5 \) and \( \alpha = 2 \). The control chart is shown to react sensitively to (gradual or sudden) changes in the mean, and to (gradual or sudden) changes in the spread of the observations, and combinations thereof. In the same article, the performance of the omnibus EWMA control chart is compared to the behavior of other charts, such as various CUSUM schemes and combinations of X- and R-charts.
One of the CUSUM schemes considered is a so-called omnibus CUSUM chart, proposed by Hawkins (1981) and Healy (1987). It involves monitoring a sum

$$S_t = \max \left( 0, \left| \frac{Y_t - \mu}{\sigma_Y} \right|^\alpha - k + S_{t-1} \right),$$

where $k$ is the reference value. An out-of-control signal is given at time $t$ if $S_t > h$ for a decision interval $h$. This procedure can also be used to detect shifts in the mean or in the spread, or combinations thereof. In the ARL computations of Domangue and Patch (1991), it is shown that the performance of the omnibus EWMA control chart is comparable to the performance of the omnibus CUSUM control chart.

In MacGregor and Harris (1993), a special case of the omnibus EWMA control chart is considered, viz. $\alpha = 2$. The resulting chart is called the Exponentially Weighted Mean Square (EWMS) control chart. It is shown that the quantities $A_t$ are weighted sums of $\chi^2$ random variables. Furthermore, an approximating distribution function for $A_t$ is discussed. The statistics $A_t$ are approximately distributed as $\chi^2(\nu)/\nu$, where the number of degrees of freedom $\nu$ depends upon the exponential weighting parameter $\lambda$, the correlation structure of the $Y_t$’s and the degrees of freedom associated with $A_t$. MacGregor and Harris use this approximation to design the EWMS control chart.

In addition, MacGregor and Harris (1993) discuss a variant to the EWMS control chart, where $\mu$ is replaced by an estimate for $\mu_t$ in formula (7.3). They call the resulting statistic the Exponentially Weighted Moving Variance (EWMV). In the article, control limits for the EWMV chart are derived for the case of independent observations and for the case of autocorrelated observations. The EWMS chart is shown to respond both to changes in the mean and the variance. The EWMV chart only responds to changes in the variance.

In case of serially correlated data, the control charts of this subsection should preferably be applied to the residuals of an appropriate time series model.

### 7.2 Subgrouping serially correlated data

Let us again consider a sequence of observations $\{Y_t\}$, whose elements are serially correlated. In the previous section we discussed the difficulties that arise when such observations are to be monitored with a control chart for
7.2. SUBGROUPING SERIALLY CORRELATED DATA

the spread. In the remaining sections of this chapter we consider the case where these problems are circumvented by creating subgroups first.

Assume that from the sequence \( \{Y_t\} \), samples of size \( n \) are drawn. Between every two samples \( k \) realizations of \( \{Y_t\} \) are not observed. Figure 7.1 illustrates the way samples are drawn.

![Figure 7.1: The method of sampling.](image)

It is assumed that the underlying measurements \( Y_1, Y_2, \cdots \) (see Figure 7.1) are generated by an AR(1) model:

\[
Y_t - \mu_t = \phi(Y_{t-1} - \mu_{t-1}) + \varepsilon_t \quad \text{for } t \in \mathbb{Z},
\]

where the \( \varepsilon_t \) are independently distributed as \( \mathcal{N}(0, \sigma_{\varepsilon_t}^2) \). The index \( t \) in \( \mu_t \) and \( \sigma_{\varepsilon_t}^2 \) is used to indicate that a special cause of variation may shift the unconditional mean or variance of the process.

Furthermore, we will assume that the samples are rational subgroups (see Subsection 2.3.1); the samples are drawn in such a way that shifts in the mean or in the spread of the process only occur between samples, and not within samples. This allows us to assume that within each sample \( i \), the unconditional means \( \mu_{(n+k)(i-1)+1}, \cdots, \mu_{(n+k)(i-1)+n} \) are constant and equal to, say \( \mu_i^* \).

Likewise, the unconditional variances \( \sigma_{\varepsilon_{(n+k)(i-1)+1}}^2, \cdots, \sigma_{\varepsilon_{(n+k)(i-1)+n}}^2 \) are assumed to be constant within sample \( i \) and equal to, say \( \sigma_{\varepsilon_i^*}^2 \).

After creation of subgroups, it becomes feasible to construct control charts for the spread of the process. However, subgrouping the data has also consequences for the control chart of the mean of the process. Therefore, in Section 7.3, we will focus on control charts for the mean of samples of AR(1) observations. In Sections 7.4 through 7.7, control charts for the spread of subgrouped AR(1) data are discussed. In Section 7.8 the ARL performance of these control charts is evaluated.
7.3 Control charts for batch means

In this section we consider control charts based on the mean of samples of AR(1) data. The mean of the $i$th sample is denoted by $Y_i$. Note that the relation between $Y_i$ and the underlying observations is

$$Y_i = \frac{1}{n} \sum_{j=1}^{n} Y_{(i-1)(n+k)+j}.$$ 

Several authors have investigated the properties of the series $\{Y_i\}$. For the case $k = 0$, Kang and Schmeiser (1987) refer to Anderson (1976) to show that if $\{Y_i\}$ is a stationary ARMA$(p,q)$ process, then the sequence of sample means $\{\bar{Y}_i\}$ is a stationary ARMA$(p,q)$ process, with $\bar{q} = p - [(p - q)/n]$, where $\lfloor x \rfloor$ is the largest integer smaller than or equal to $x$, and $n$ is the sample size. For the case of AR(1) observations, $p = 1$, and $q = 0$, so that the sequence of sample means is an ARMA(1,1) process when the first element of a sample is drawn directly after the last element of the previous sample ($k = 0$). Alwan and Radson (1992b) consider the case $k > 0$, and show that the sequence of sample means is also an ARMA(1,1) process when there are $k$ unobserved AR(1) realizations between successive samples.

In the in-control situation, we have $\mu_i = \mu$ for samples $i = 1, 2, \ldots$, and $\sigma_{\varepsilon_i} = \sigma_{\varepsilon}$ for $i = 1, 2, \ldots$. Using the results of the cited references, the sequence of sample means of AR(1) observations obeys the following dynamic structure

$$\bar{Y}_i = (1 - \bar{\phi})\mu + \bar{\phi}\bar{Y}_{i-1} + \bar{\varepsilon}_i - \bar{\theta}\bar{\varepsilon}_{i-1} \quad (7.4)$$

where $\{\bar{\varepsilon}_i\}$ is a sequence of i.i.d. $\mathcal{N}(0, \sigma_{\varepsilon}^2)$ disturbances, $\bar{\phi}$ is the AR-parameter, and $\bar{\theta}$ is the MA-parameter.

It is now of interest how the parameters $\bar{\phi}$, $\bar{\theta}$ and $\sigma_{\varepsilon}^2$ of model (7.4) relate to the parameters $\phi$ and $\sigma_{\varepsilon}^2$ of the model for the underlying individual observations, and $n$ and $k$. In general, it is very difficult to determine the relationships between the parameters of an ARMA$(p,q)$ model and the parameters of the ARMA$(p,q)$ model of the aggregated time series. However, Alwan and Radson (1992b) tracked down closed form formulas for the case of underlying AR(1) observations. They show that

$$\bar{\phi} = \phi^{n+k},$$
7.3. CONTROL CHARTS FOR BATCH MEANS

where it is to be noted that \( \bar{\phi} \) does not depend on \( \sigma^2_\varepsilon \). The relation between \( \theta \) and \( \sigma^2_\varepsilon \) of the aggregated model and \( \phi, \sigma^2_\varepsilon, n, \) and \( k \) is derived as follows. The variance of \( Y_i \) is evaluated in terms of \( \phi, \sigma^2_\varepsilon, n, \) and \( k \):

\[
\text{Var}(Y_i) = \frac{\sigma^2_\varepsilon}{1 - \phi^2} \left[ 1 + \frac{2}{n} \sum_{j=1}^{n-1} (n-j)\phi^j \right] = \frac{\sigma^2_\varepsilon}{1 - \phi^2} \left[ n(1 + 2C_2) - 2\phi C_1 \right],
\]

where \( C_1 \) is defined as \( \sum_{j=1}^{n-1} j\phi^{j-1} \), and \( C_2 = \sum_{j=1}^{n-1} \phi^j \); notation borrowed from Alwan and Radson (1992b). This expression is equated to the variance of \( Y_i \), expressed in parameters of the aggregated model:

\[
\text{Var}(Y_i) = \frac{1 + \bar{\phi}^2 - 2\bar{\phi} \bar{\bar{\phi}}}{1 - \phi^2} \sigma^2_\varepsilon.
\]

Expressing the covariance between successive means in terms of \( \phi, \sigma^2_\varepsilon, n, \) and \( k \) leads to:

\[
\text{Cov}(Y_i, Y_{i+1}) = \frac{\sigma^2_\varepsilon}{1 - \phi^2} \frac{\phi^{n+k}}{n^2} \left[ n + \sum_{j=1}^{n-1} j\phi^{j-n} + \sum_{j=1}^{n} (n-j)\phi^j \right] = \frac{\sigma^2_\varepsilon}{1 - \phi^2} \frac{\phi^{n+k}}{n^2} \left[ n(1 + C_2) + \left( \frac{1}{\phi^{n-1} - \phi} \right) C_1 \right].
\]

Next, this expression is equated to

\[
\text{Cov}(Y_i, Y_{i+1}) = \frac{(1 - \bar{\phi}\bar{\bar{\phi}})(\bar{\phi} - \bar{\bar{\phi}})}{1 - \phi^2} \sigma^2_\varepsilon.
\]

Using \( \bar{\bar{\phi}} = \phi^{n+k} \), we have two equations in two unknowns. Solving for \( \bar{\theta} \) and \( \sigma^2_\varepsilon \) leads to the solutions

\[
\bar{\theta}_1 = \frac{-K_1(1 + \phi^{2(n+k)}) + 2K_2\phi^{n+k} + K_3}{2(K_2 - \phi^{n+k}K_1)},
\]

\[
\bar{\theta}_2 = \frac{-K_1(1 + \phi^{2(n+k)}) + 2K_2\phi^{n+k} - K_3}{2(K_2 - \phi^{n+k}K_1)},
\]

where, as in Alwan and Radson (1992b),

\[
K_1 = \frac{\sigma^2_\varepsilon}{1 - \phi^2} \frac{1}{n^2} C_3 \left[ n(1 + 2C_2) - 2\phi C_1 \right],
\]

\[
K_2 = \frac{1}{(\phi^{n-1} - \phi)} C_1,
\]

\[
K_3 = \frac{1}{\phi^{n-1} - \phi} C_2.
\]
CHAPTER 7. CONTROL CHARTS FOR THE SPREAD

\[ K_2 = \frac{\sigma_\varepsilon^2}{1 - \phi^2} \frac{\phi^{n+k}}{n^2} C_3 \left[ n(1 + C_2) + \left( \frac{1}{\phi^{n-1}} - \phi \right) C_1 \right], \]

\[ K_3 = \sqrt{(\phi^{2(n+k)} - 1) \left[ (2K_2 - \phi^{n+k}K_1)^2 - K_1^2 \right]}, \]

and

\[ C_3 = 1 - \phi^{2(n+k)}. \]

It can be shown that the two ARMA(1,1) models corresponding to \( \bar{\theta}_1 \) and \( \bar{\theta}_2 \) have the same autocorrelation function. However, only \( \bar{\theta}_1 \) satisfies the invertibility condition (see for example Harvey (1993)). The two solutions for \( \sigma_\varepsilon^2 \) corresponding to \( \bar{\theta}_1 \) and \( \bar{\theta}_2 \) are

\[ \sigma_{\varepsilon,1}^2 = \frac{K_1(1 + \phi^{2(n+k)}) - 2K_2\phi^{n+k} + K_3}{2 - 2\phi^{2(n+k)}} \]

\[ \sigma_{\varepsilon,2}^2 = \frac{K_1(1 + \phi^{2(n+k)}) - 2K_2\phi^{n+k} - K_3}{2 - 2\phi^{2(n+k)}}. \]

From these formulas it can be seen that \( \sigma_{\varepsilon}^2 \) does depend on all four parameters, and that \( \sigma_{\varepsilon}^2 \) is proportional to \( \sigma_{\varepsilon}^2 \). In Table 7.4, we computed the parameters \( \bar{\phi} \) and \( \bar{\theta} \) for selected values of \( \phi \) for the cases \( n = 5, k = 0, n = 5, k = 5 \), and \( n = 5, k = 10 \).

**Table 7.4:** ARMA(1,1) parameters \( \bar{\phi} \) and \( \bar{\theta} \)
as functions of \( \phi \), \( n \), and \( k \).

<table>
<thead>
<tr>
<th>( n = 5, k = 0 )</th>
<th>( n = 5, k = 5 )</th>
<th>( n = 5, k = 10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi )</td>
<td>( \bar{\phi} )</td>
<td>( \phi )</td>
</tr>
<tr>
<td>-0.9</td>
<td>-0.5905</td>
<td>0.0103</td>
</tr>
<tr>
<td>-0.6</td>
<td>-0.0778</td>
<td>0.0793</td>
</tr>
<tr>
<td>-0.3</td>
<td>-0.0024</td>
<td>0.0563</td>
</tr>
<tr>
<td>0.3</td>
<td>0.0024</td>
<td>-0.0735</td>
</tr>
<tr>
<td>0.6</td>
<td>0.0778</td>
<td>-0.1769</td>
</tr>
<tr>
<td>0.9</td>
<td>0.5905</td>
<td>-0.2457</td>
</tr>
</tbody>
</table>
7.4. THE MOVING RANGE CHART WHEN N > 1

The values for $\bar{\theta}$ differ from those presented in Table 1 of Alwan and Radson (1992b). This is due to an error in their expression for the covariance between $\bar{Y}_t$ and $\bar{Y}_{t-1}$; compare formula (7.6) to equation (13) of Alwan and Radson (1992b). As a result, $K_2$ is also erroneously defined. In addition, an error slipped into the definition of $K_1$; compare the definition above to the definition of $K_1$ in Appendix I of Alwan and Radson (1992b). This results in imaginary solutions for $\bar{\phi}$ in some cases. This does not happen if the computations are performed as indicated above.

Note from Table 7.4 that both the absolute value of $\bar{\phi}$ and the absolute value of $\bar{\theta}$ decrease considerably when $n$ or $k$ is increased. Hence, taking sample means reduces the serial correlation in the statistic that is to be monitored in the control chart. This effect of subgroup taking is also observed by Runger and Willemain (1995). These authors also consider the case of monitoring a weighted sample mean, where the weights are chosen in such a way that successive weighted means are uncorrelated. This procedure was introduced by Bischak, Kelton, and Pollock (1993).

From Table 7.4, it can also be concluded that the MA parameter $\bar{\theta}$ is much smaller throughout than the AR parameter $\bar{\phi}$, so that the series $\{\bar{Y}_t\}$ will, in most cases, be very similar to an AR(1) series. Alwan and Radson (1992b) have encountered this phenomenon frequently in practice.

The analysis above reduces monitoring sample means of AR(1) data to monitoring ARMA(1,1) data. This means that the control charts of Chapters 3 through 5 can be applied, provided that the control chart is adapted to account for ARMA(1,1) dependence in the data. Appendix A might offer guidance on which control chart to use, since the ARL behavior for some of these charts for ARMA(1,1) models is tabulated there.

In the following section, control charts for the spread of subgrouped serially correlated data are discussed.

7.4 The moving range chart when $n > 1$

In this subsection, we will investigate the properties of the $MR$-chart for AR(1) data that are grouped into subsamples. Consider again a sequence of AR(1) observations $\{Y_t\}$, generated by (2.3)

$$Y_t - \mu_t = \phi(Y_{t-1} - \mu_{t-1}) + \varepsilon_t \quad \text{for } t \in \mathbb{Z}. $$

Suppose that subgroups of size $n$ are formed in such a way that there are $k$ observations between two successive samples. Thus, if the first sample
consists of observations \( \{Y_1, Y_2, \ldots, Y_n\} \), the second sample consists of the observations \( \{Y_{n+k+1}, Y_{n+k+2}, \ldots, Y_{2n+k}\} \), and so on.

Within each sample, \( n - 1 \) moving ranges are computed. From the \( i \)th sample, which consists of the observations

\[
Y_{(n+k)(i-1)+1}, Y_{(n+k)(i-1)+2}, \ldots, Y_{(n+k)(i-1)+n},
\]

the moving ranges \( MR_{i,2}, \ldots, MR_{i,n} \) are computed as

\[
MR_{i,j} = |Y_{(n+k)(i-1)+j} - Y_{(n+k)(i-1)+j-1}| \quad \text{for } j = 2, \ldots, n.
\]

The \( i \)th average moving range is then computed as

\[
\overline{MR}_i = \frac{1}{n-1} \sum_{j=2}^{n} MR_{i,j}.
\]

These statistics are plotted in the \( \overline{MR} \) control chart, and compared to limits \( LCL_{\overline{MR}} \) and \( UCL_{\overline{MR}} \).

Using the results of Subsection 7.1.2, we obtain the following properties of \( \overline{MR}_i \). For the expected value of \( \overline{MR}_i \) we have

\[
E(\overline{MR}_i) = \frac{1}{n-1} \sum_{j=2}^{n} E(MR_{i,j})
\]

\[
= \frac{2}{\sqrt{n}} \sqrt{1 - \phi} \sigma_Y.
\]

Note that \( \overline{MR}_i/d_2(2) \) is biased as an estimator for \( \sigma_Y \) by a factor of \( \sqrt{1 - \phi} \), and that the bias does not disappear as the sample size increases. For the variance of \( \overline{MR}_i \) we have
7.4. THE MOVING RANGE CHART WHEN $N > 1$

\[
\text{Var}(MR_i) = \frac{1}{(n-1)^2} \left( \sum_{j=2}^{n} \text{Var}(MR_{i,j}) + 2 \sum_{j=2}^{n-1} \sum_{l=j+1}^{n} \text{Cov}(MR_{i,j}, MR_{i,l}) \right) = \frac{2(1-\phi)}{\pi(n-1)^2} \sigma_Y^2 \left( (n-1)(\pi - 2) + \sum_{l=1}^{n-2} 4(n-l-1) \left( \sqrt{1-\rho_l^2} + \rho_l \arcsin(\rho_l) - 1 \right) \right)
\]

where $\rho_l$ is defined in Equation (7.2) of Section 7.1.2. The expression for $\text{Var}(MR_i)$ of AR(1) data does not simplify much further as it does for the i.i.d. case, see Cryer and Ryan (1990). This is due to $\rho_l \neq 0$ for all $l$ if $\phi \neq 0$. However, the expression for $\text{Var}(MR_i)$ above is relatively easy to compute numerically in practical cases.

For the covariance between $MR_i$ and $MR_{i+m}$ we have

\[
\text{Cov}(MR_i, MR_{i+m}) = \frac{1}{(n-1)^2} \sum_{j=2}^{n} \sum_{l=2}^{n} \text{Cov}(MR_{i,j}, MR_{i+m,l}) = \frac{4(1-\phi)}{\pi(n-1)^2} \sigma_Y^2 \times \sum_{l=-n+2}^{n-2} (n-1-|l|) \left( \sqrt{1-\rho_q^2} + \rho_q \arcsin(\rho_q) - 1 \right)
\]

where $q = (n+k)m$, and $\rho_j$ is defined in Equation (7.2). The correlation between $MR_i$ and $MR_{i+m}$ is denoted by $\rho_{MR,m}$ and is defined as

\[
\rho_{MR,m} = \frac{\text{Cov}(MR_i, MR_{i+m})}{\text{Var}(MR_i)}.
\]
The correlations $\rho_{\sigma m}$ are extremely small for choices of $n$ and $k$ that will be used in practice (say 5 and 10 respectively). The correlations are the largest if the smallest possible sample size is chosen, which is $n = 2$, and when there are no unobserved observations between the samples, i.e. $k = 0$. Table 7.5 tabulates the correlations between successive moving ranges for $n = 2$, $k = 0$ for selected values of $\phi$ and the lag $m$ ranging from 1 to 10.

**Table 7.5:** Autocorrelations between $\overline{MR}_i$ and $\overline{MR}_{i+m}$ for selected values of $\phi$, with $n = 2$ and $k = 0$.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>$-0.9$</th>
<th>$-0.6$</th>
<th>$-0.3$</th>
<th>0</th>
<th>0.3</th>
<th>0.6</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 1$</td>
<td>0.6928</td>
<td>0.2060</td>
<td>0.0334</td>
<td>0</td>
<td>0.0097</td>
<td>0.0126</td>
<td>0.0018</td>
</tr>
<tr>
<td>$m = 2$</td>
<td>0.4401</td>
<td>0.0262</td>
<td>0.0003</td>
<td>0</td>
<td>0.0001</td>
<td>0.0016</td>
<td>0.0012</td>
</tr>
<tr>
<td>$m = 3$</td>
<td>0.2837</td>
<td>0.0034</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0008</td>
</tr>
<tr>
<td>$m = 4$</td>
<td>0.1842</td>
<td>0.0004</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0005</td>
</tr>
<tr>
<td>$m = 5$</td>
<td>0.1201</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0003</td>
</tr>
<tr>
<td>$m = 6$</td>
<td>0.0784</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
</tr>
<tr>
<td>$m = 7$</td>
<td>0.0513</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0001</td>
</tr>
<tr>
<td>$m = 8$</td>
<td>0.0336</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0001</td>
</tr>
<tr>
<td>$m = 9$</td>
<td>0.0220</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0001</td>
</tr>
<tr>
<td>$m = 10$</td>
<td>0.0144</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Note that the first five rows of Table 7.5 equal respectively the second, fourth, sixth, eighth and tenth row of Table 7.3. The correlations between successive $\overline{MR}_i$ decrease quickly with $k$ and also with $n$. For practical purposes, the successive correlations may be treated as uncorrelated statistics. As in the case for $n = 1$, this facilitates the design and the evaluation of the ARL of the $\overline{MR}$ control chart. The ARL will be approximated (given lower control limit $LCL_{\overline{MR}}$ and upper limit $UCL_{\overline{MR}}$) by

$$\text{ARL}_{\overline{MR}} \approx \frac{1}{1 - P(LCL_{\overline{MR}} < \overline{MR}_{i,j} < UCL_{\overline{MR}})}.$$ 

Again, the probability $P(LCL_{MR} < \overline{MR}_{i,j} < UCL_{MR})$ will be evaluated by Monte Carlo simulation. The control limits are determined by requiring a certain in-control ARL and a certain out-of-control ARL. In Subsection 7.8, the ARL curves of the moving range chart are compared to other schemes for monitoring the spread using subgrouped AR(1) observations.
7.5 The $S^2$-chart

In this subsection, we will discuss the $S^2$-chart for the spread of AR(1) observations. Consider a sequence of AR(1) observations $\{Y_t\}$, generated by (2.3)

$$Y_t - \mu_t = \phi(Y_{t-1} - \mu_{t-1}) + \varepsilon_t \quad \text{for } t \in \mathbb{Z}.$$  

Suppose that subgroups of size $n$ are formed in such a way that there are $k$ observations between two successive samples.

The sequence of observed sample variances $\{S_i^2\}$ is monitored by a lower control limit $LCL_{S^2}$ and an upper control limit $UCL_{S^2}$. The width of these limits depend on the standard deviation of the $S_i^2$, which is proportional to $\sigma_Y^2$. If the control limits are chosen such that in the in-control case $P(S_i^2 < LCL_{S^2}) = P(S_i^2 > UCL_{S^2})$, then the control limits will not be centered around $E(S_i^2)$ due to skewness of the distribution of $S_i^2$.

Consider the case where the process is in control with respect to the mean, that is, assume $\mu_t = \mu$ for all $t$. Then the expectation of the $i$th sample variance equals

$$E(S_i^2) = E \left( \frac{1}{n-1} \sum_{i=1}^{n} (Y_{(n+k)(i-1)+j} - \bar{Y}_i)^2 \right)$$

$$= E \left( \frac{1}{n-1} \left\{ \sum_{j=1}^{n} (Y_{(n+k)(i-1)+j} - \mu)^2 - n(\bar{Y}_i - \mu)^2 \right\} \right)$$

$$= \frac{n}{n-1} (\sigma_Y^2 - \text{Var}(\bar{Y}_i)),$$

where

$$\bar{Y}_i = \frac{1}{n} \sum_{j=1}^{n} Y_{(n+k)(i-1)n+j}.$$  

For the variance of $\bar{Y}_i$ we have from Anderson (1971)

$$\text{Var}(\bar{Y}_i) = \frac{\sigma_Y^2}{n} \left[ 1 + 2 \sum_{j=1}^{n-1} \left( 1 - \frac{j}{n} \right) \phi^j \right]$$

$$= \frac{\sigma_Y^2}{n^2 (1 - \phi)^2} \left( n - 2\phi - n\phi^2 + 2\phi^{n+1} \right),$$
so that we obtain

$$E(S_i^2) = \sigma_Y^2 \left( \frac{n}{n-1} \right) \left[ 1 - \frac{n - 2\phi - n\phi^2 + 2\phi^{n+1}}{n^2 (1 - \phi)^2} \right].$$

(7.6)

Hence, for finite samples, the sample variance is a biased estimator for $\sigma_Y^2$. The bias disappears for $n \to \infty$. In Figure 7.2, the behavior of the factor of $\sigma_Y^2$ in Equation (7.6) is plotted against the sample size $n$ for various $\phi$.

![Figure 7.2: Behavior of the bias factor in $E(S_i^2)$ against $n$ for various $\phi$.](image)

For the uncorrelated case $\phi = 0$, $S_i^2$ is an unbiased estimator of $\sigma_Y^2$. However, for negative $\phi$ the process variance is overestimated, whereas for $\phi > 0$ the variance is underestimated. Note that this bias does not disappear if the process variance is estimated as the average of a number of sample variances. For a common subgroup sample size of $n = 5$, the process is overestimated by 20% if $\phi = -0.9$, by 16% if $\phi = -0.6$, and by 10% if $\phi = -0.3$. With this sample size $\sigma_Y^2$ is underestimated by 15% if $\phi = 0.3$, by 40% if $\phi = 0.6$, and by 81% if $\phi = 0.9$! For large positive autocorrelation, the bias factor converges very slowly to one. Sample sizes larger than $n = 40$ are required to obtain a reasonable accurate estimate.
The bias in $S_i^2$ as an estimator for $\sigma_Y^2$ poses no direct problems for the construction of an $S^2$-chart for the spread. It is the first purpose of the control chart to detect changes in the process variance. This can very well be done using biased statistics, as long as the level of the control chart is not interpreted as the size of the variance of the process.

Next the covariance of $S_i^2$ and $S_{i+m}^2$ will be derived. Without loss of generality, we assume that $\mu_t = \mu = 0$. Define

$$Y = \begin{pmatrix} Y_{(n+k)(i-1)+1} \\ Y_{(n+k)(i-1)+2} \\ \vdots \\ Y_{(n+k)(i-1)+n} \\ Y_{(n+k)(i+m-1)+1} \\ Y_{(n+k)(i+m-1)+2} \\ \vdots \\ Y_{(n+k)(i+m-1)+n} \end{pmatrix}.$$  

The variance matrix of $Y$ is

$$\text{Var}(Y) = \begin{pmatrix} A & B \\ B' & A \end{pmatrix}$$

with

$$A = \frac{\sigma_x^2}{1 - \phi^2} \begin{pmatrix} 1 & \phi & \phi^2 & \ldots & \phi^{n-1} \\ \phi & 1 & \phi & \ldots & \phi^{n-2} \\ \phi^2 & \phi & 1 & \ldots & \phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \ldots & 1 \end{pmatrix},$$

and
\[
B = \frac{\sigma_x^2}{1 - \phi^2} \phi^{(n+k)m} \begin{pmatrix}
1 & \phi & \phi^2 & \cdots & \phi^{n-1} \\
1 & \phi & \phi^2 & \cdots & \phi^{n-2} \\
1 & \phi & 1 & \cdots & \phi^{n-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{1}{\phi^{n-1}} & \frac{1}{\phi^{n-2}} & \frac{1}{\phi^{n-3}} & \cdots & 1
\end{pmatrix}.
\]

Then \((n - 1)S_i^2\) can be written as \((PY)'PY = Y'PY\), where

\[
P = \begin{pmatrix}
M & 0 \\
0 & 0
\end{pmatrix},
\]

with \(0\) an \((n \times n)\) matrix with all elements equal to zero, and

\[
M = I_n - \frac{1}{n} \mathbf{u} \mathbf{u}' ,
\]

where \(I_n\) is the \((n \times n)\) identity matrix and \(\mathbf{u}\) is an \((n \times 1)\) vector with all elements equal to one. Analogously, \((n - 1)S_{i+m}^2\) can be written as \((QY)'QY = Y'QY\), where

\[
Q = \begin{pmatrix}
0 & 0 \\
0 & M
\end{pmatrix}.
\]

Using Corollary 4.1 from Magnus and Neudecker (1979), we have

\[
(n - 1)^2 \text{Cov}(S_i^2, S_{i+m}^2) = 2 \text{tr} \left[ P \text{Var}(Y) Q \text{Var}(Y) \right] = 2 \text{tr} \left( M B M B' M B M A \right) = 2 \text{tr} (M B (M B)^') = 2 \sum_{p=1}^{n} \sum_{q=1}^{n} (M B)^2_{pq}.
\]
where \((MB)_{pq}\) is the element in the \(p\)th row and \(q\)th column of \(MB\).

For \((n - 1)^2 \text{Var}(S_i^2) = (n - 1)^2 \text{Var}(S_{i+m})^2\) we have, using again Corollary 4.1 from Magnus and Neudecker (1979),

\[
\text{Var}(Y'PY) = 2 \text{tr} [(P \text{Var}(Y))^2] + 4 \mu^2 e_{2n}' \text{Var}(Y) Pe_{2n}
\]

\[
= 2 \text{tr} [(MA)^2].
\]

Hence for \(\rho_{S_i^2;m}\), the correlation coefficient between \(S_i^2\) and \(S_{i+m}^2\), we have

\[
\rho_{S_i^2;m} = \frac{\text{tr} [MB(BM)']}{\text{tr} [(MA)^2]}.
\]

The value of \(\rho_{S_i^2;m}\) is decreasing in \(n, k\), and \(m\), and is increasing with \(|\phi|\). However, for practical choices of \(n\) and \(k\), the values of the autocorrelation function \(\{\rho_{S_1^2}, \rho_{S_2^2}, \cdots\}\) are very close to zero. As an illustration, in Table 7.6 we tabulated the case where samples of size \(n = 5\) are drawn, and where there are \(k = 10\) unobserved AR(1) realizations between two samples. The correlations are tabulated for lags \(m = 1, \cdots, 5\) for the case where there is high negative autocorrelation, \(\phi = -0.8\), and for the case when there is high positive autocorrelation, \(\phi = 0.8\).

**Table 7.6**: Autocorrelations between \(S_i^2\) and \(S_{i+m}^2\) for \(n = 5, k = 10\).

<table>
<thead>
<tr>
<th>(m)</th>
<th>(\phi = -0.8)</th>
<th>(\phi = 0.8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00328040956151</td>
<td>0.00533538388005</td>
</tr>
<tr>
<td>2</td>
<td>0.00000406095034</td>
<td>0.00000660488533</td>
</tr>
<tr>
<td>3</td>
<td>0.00000000502721</td>
<td>0.00000000817645</td>
</tr>
<tr>
<td>4</td>
<td>0.0000000000622</td>
<td>0.00000000001012</td>
</tr>
<tr>
<td>5</td>
<td>0.00000000000001</td>
<td>0.00000000000001</td>
</tr>
</tbody>
</table>

From Table 7.6 it can be observed that the correlation between successive sample variances is minimal, even in the case of severe autocorrelation (either negative or positive). Therefore, the correlation between successive sample variances will be neglected in the construction of an \(S^2\)-chart.

Treating the sample variances as independent statistics allows us to determine the ARL for the \(S^2\)-chart assuming control limits \(LCL_{S^2}\) and \(UCL_{S^2}\) by
\[ ARL \approx \frac{1}{1 - P(LCL_{S^2} < S_i^2 < UCL_{S^2})}. \]

The control limits LCL_{S^2} and UCL_{S^2} are determined in the usual way by requiring a certain in-control ARL and a certain out-of-control ARL. To this end, the distribution of \( S_i^2 \) needs to be known. Schmid (1995a) remarks that the usual results for quadratic forms do not apply here, due to the fact that \( P \text{Var}(Y) \) and \( Q \text{Var}(Y) \) are not idempotent. To overcome this problem, Schmid (1995a) uses Laguerre expansion of the distribution function of \( S_i^2 \) to obtain the ARL of the \( S^2 \)-chart. Based on ARL requirements, he obtains upper control limits that account for serial correlation in the data.

The results of Schmid (1995a) show that, compared to the case \( \phi = 0 \), the control limits become wider for \( \phi < 0 \). This is to be expected in the light of Figure 7.2, where it is shown that \( S_i^2 \) overestimates \( \sigma_Y^2 \) for finite subsamples. Conversely, the control limits of the \( S^2 \)-chart are tighter for the case \( \phi > 0 \), due to underestimation of \( S_i^2 \).

In Section 7.8, we will simulate the ARL behavior of the \( S^2 \)-chart. This will be compared to the ARL behavior of the other charts for the spreading of subgrouped serially correlated data that are discussed in this chapter. Note that the \( S \) chart is not separately discussed. The properties of this control chart are comparable to the properties of the \( S^2 \)-chart.

### 7.6 The \( R \)-chart

In Alwan and Radson (1992a), the effect of positive first-order autocorrelation on the range of a subsample is investigated. The subsample range of sample \( i \) is defined as

\[
R_i = \max \left[ Y_{(n+k)(i-1)+1}, Y_{(n+k)(i-1)+2}, \ldots, Y_{(n+k)(i-1)+n} \right] + \min \left[ Y_{(n+k)(i-1)+1}, Y_{(n+k)(i-1)+2}, \ldots, Y_{(n+k)(i-1)+n} \right] - \min \left[ Y_{(n+k)(i-1)+1}, Y_{(n+k)(i-1)+2}, \ldots, Y_{(n+k)(i-1)+n} \right].
\]

In the case of independent observations, the mean of a subsample range increases with the sample size. Simulation studies that were performed by Alwan and Radson (1992a) show that the effect of positive autocorrelation on the mean of \( R_i \) is small, compared to the effect of the sample size \( n \). This indicates that \( R_i/d_2(n) \) is a biased estimator for \( \sigma_Y \), since the latter is increasing with \( |\phi| \). In Table 7.7, we simulated the mean of \( R_i/d_2(n) \) for \( \phi = -0.9, -0.6, -0.3, 0, 0.3, 0.6, 0.9 \), using sample sizes \( n = 2, 3, 4, 5, 6 \). The
number of unobserved AR(1) realizations between two samples was taken to be \( k = 0 \), since the results are not influenced by different values of \( k \).

Table 7.7: Simulated values for \( \mathbb{E}[R_i/d_2(n)] \) for selected values of \( \phi \) and \( n \).

<table>
<thead>
<tr>
<th>( \sigma_Y )</th>
<th>( n = 2 )</th>
<th>( n = 3 )</th>
<th>( n = 4 )</th>
<th>( n = 5 )</th>
<th>( n = 6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi = -0.9 )</td>
<td>2.2942</td>
<td>3.1552</td>
<td>2.4435</td>
<td>2.2230</td>
<td>2.1145</td>
</tr>
<tr>
<td>( \phi = -0.6 )</td>
<td>1.2500</td>
<td>1.5717</td>
<td>1.3881</td>
<td>1.3242</td>
<td>1.2933</td>
</tr>
<tr>
<td>( \phi = -0.3 )</td>
<td>1.0483</td>
<td>1.1975</td>
<td>1.1286</td>
<td>1.1040</td>
<td>1.0893</td>
</tr>
<tr>
<td>( \phi = 0 )</td>
<td>1.0000</td>
<td>1.0062</td>
<td>1.0011</td>
<td>0.9994</td>
<td>1.0009</td>
</tr>
<tr>
<td>( \phi = 0.3 )</td>
<td>1.0483</td>
<td>0.8772</td>
<td>0.9184</td>
<td>0.9418</td>
<td>0.9604</td>
</tr>
<tr>
<td>( \phi = 0.6 )</td>
<td>1.2500</td>
<td>0.7867</td>
<td>0.8587</td>
<td>0.9091</td>
<td>0.9458</td>
</tr>
<tr>
<td>( \phi = 0.9 )</td>
<td>2.2942</td>
<td>0.7288</td>
<td>0.8144</td>
<td>0.8914</td>
<td>0.9505</td>
</tr>
</tbody>
</table>

Table 7.7 shows that \( R_i/d_2(n) \) is biased for \( \sigma_Y \). For negative \( \phi \), \( \sigma_Y \) is overestimated in most cases. For positive \( \phi \), the bias is negative. Except for \( \phi = -0.9 \), the bias reduces with larger \( n \). For negative \( \phi \), the bias reduction is faster than for positive \( \phi \). However, the results for \( \phi = -0.9 \) show that the bias is positive for \( n = 2, 3 \), and \( \sigma_Y \) is underestimated for \( n = 4, 5, 6 \). Note that the entries in the column corresponding to \( n = 2 \) agree with the analytical results in the second column of Table 3.1.

Alwan and Radson (1992a) also performed simulation studies to determine the effect of AR(1) dependence in the observations on the correlation of successive subsample ranges. The results show that the simulated first-order autocorrelations are within the standard 95% confidence interval, for practical choices of \( n \) and \( k \), and \( \phi > 0 \). Especially for larger values of \( k \), the correlation between successive subsample ranges are indistinguishable from zero.

Utilizing these results, we decide to design the \( R \)-chart as if the \( R_i \)'s were independent. The ARL for the \( R \)-chart assuming control limits \( LCL_R \)
and UCL\(_R\) is then computed as

\[ ARL \approx \frac{1}{1 - P(LCL_R < R_i < UCL_R)}. \]

The probability in this expression will be evaluated by Monte Carlo simulation. The control limits LCL\(_R\) and UCL\(_R\) are again determined by making two requirements on ARL curve.

### 7.7 The residuals chart

Residuals charts have already been discussed in Chapters 3, 4 and 5. In those applications, residuals of a fitted time series model are monitored for a change in the mean of the correlated observations. However, a change in the variance of the correlated observations will also be transferred to the variance of residuals. This motivates the use of a control chart for the spread of residuals to monitor the variance of serially correlated observations.

Suppose that we have AR(1) observations available, generated by

\[ Y_t - \mu = \phi(Y_{t-1} - \mu) + \varepsilon_t \quad \text{for } t \in \mathbb{Z}. \]

The expectation of the observations is assumed to be constant and equal to \( \mu \) for all \( t \). The AR parameter \( \phi \) is also assumed to be constant. However, at time \( T \), the variance of the disturbances may change from, say, \( \sigma^2_{\varepsilon, 0} \) to \( \sigma^2_{\varepsilon, 1} \). The residual \( e_t \) is computed as

\[ (Y_t - \mu) - \phi(Y_{t-1} - \mu) = \varepsilon_t, \]

where the last equality only holds if \( \mu \) and \( \phi \) are known. A shift in the variance of the disturbances is thus transferred to a shift in the variance of the residuals of the same size. Hence, a residuals control chart for the spread is unaffected by the serial correlation in the data. Furthermore, since the residuals are (approximately) uncorrelated, the standard control charts for monitoring the variance of a process can be used and interpreted without further alterations. In the following section, the ARL behavior of an \( R \)-chart based on residuals is compared to the ARL performance of the \( \overline{MR} \)-chart, the \( S^2 \)-chart, and the \( R \)-chart, respectively.
7.8 ARL comparison

In this section, the ARL behavior of the control charts for the spreading of subgrouped data is evaluated. In Table 7.8(a), simulated ARL values of the $\overline{MR}$ chart for subgrouped AR(1) data are tabulated for $\phi = -0.9, -0.6, -0.3, 0, 0.3, 0.6, 0.9$. The values in the first row correspond to the case where the variance is in control, i.e., the variance of $\{Y_t\}$ is $\sigma_Y$. The entries in the second row are the simulated ARL values when the variance of $\{Y_t\}$ has shifted from from $\sigma_Y$ to $2\sigma_Y$. In the third row, the effect of a shift from $\sigma_Y$ to $3\sigma_Y$ on the ARL is simulated. Tables 7.8(b), (c) and (d) are set up similarly. Table 7.8(b) contains the simulated ARL values of the $S^2$-chart. In Table 7.8(c), the simulated ARL behavior of the $R$-chart for subgrouped AR(1) data is tabulated. The results of the simulations for an $R$-chart of residuals are presented in Table 7.8(d).

All entries of Table 7.8 are computed assuming that subgroups of size $n = 5$ are formed, and that there are $k = 10$ unobserved AR(1) realizations between successive samples. The values are based on 10,000 replications. The bracketed numbers are the simulated standard errors.

From the results of Table 7.8, we conclude that of the four control charts considered, the $R$-chart based on residuals has the best ARL performance for all values of $\phi$. Recall from Chapters 3, 4 and 5 that a shift in the mean of serially correlated data is, in general, not transferred to a shift of the same size in the mean of the residuals. A residuals chart for the spread of serially correlated data does not have such a disadvantage. If the appropriate time series model and its process parameters are known, a shift in the standard deviation of the process will be transferred to a shift of the same size in the standard deviation of the residuals.

Another argument in favor of the use of residuals-based control charts for the spread is that such charts are easy to design. If the fitted time series model is appropriate for the data, the residuals will be approximately uncorrelated, so that the control limits do not depend on the parameters of the model. The width of the control limits depends only on $\sigma_\varepsilon$, the standard deviation of the disturbances of the time series model, and the desired ARL behavior of the control chart. Hence, multipliers of $\sigma_\varepsilon$ that result in desired ARL behavior of the control chart have to be established only once. After that, control limits for monitoring the spread of any type of serially correlated process can be found by multiplying the appropriate multiplier with the standard deviation of the disturbances of the process.

In the light of the foregoing, we recommend to monitor the spread of a serially correlated process with a residuals-based control chart.
Table 7.8: Simulated ARL values of various control charts for the spread of subgrouped serially correlated data.

(a) $\overline{MR}$ chart of subgrouped data

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>-0.9</th>
<th>-0.6</th>
<th>-0.3</th>
<th>0.0</th>
<th>0.3</th>
<th>0.6</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b = 1$</td>
<td>374.8</td>
<td>376.9</td>
<td>367.7</td>
<td>367.6</td>
<td>360.8</td>
<td>359.2</td>
<td>374.7</td>
</tr>
<tr>
<td></td>
<td>(3.7)</td>
<td>(3.8)</td>
<td>(3.6)</td>
<td>(3.7)</td>
<td>(3.7)</td>
<td>(3.6)</td>
<td>(3.8)</td>
</tr>
<tr>
<td>$b = 2$</td>
<td>85.4</td>
<td>35.0</td>
<td>25.9</td>
<td>19.6</td>
<td>17.0</td>
<td>15.7</td>
<td>14.3</td>
</tr>
<tr>
<td></td>
<td>(0.9)</td>
<td>(0.3)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>$b = 3$</td>
<td>37.2</td>
<td>13.7</td>
<td>9.3</td>
<td>7.2</td>
<td>6.0</td>
<td>5.5</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>(0.4)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.0)</td>
<td>(0.0)</td>
</tr>
</tbody>
</table>

(b) $S^2$-chart of subgrouped data

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>-0.9</th>
<th>-0.6</th>
<th>-0.3</th>
<th>0.0</th>
<th>0.3</th>
<th>0.6</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b = 1$</td>
<td>359.6</td>
<td>377.1</td>
<td>375.7</td>
<td>383.1</td>
<td>366.0</td>
<td>366.8</td>
<td>364.4</td>
</tr>
<tr>
<td></td>
<td>(3.6)</td>
<td>(3.8)</td>
<td>(3.8)</td>
<td>(3.8)</td>
<td>(3.7)</td>
<td>(3.7)</td>
<td>(3.6)</td>
</tr>
<tr>
<td>$b = 2$</td>
<td>79.1</td>
<td>26.4</td>
<td>14.8</td>
<td>11.7</td>
<td>14.1</td>
<td>21.3</td>
<td>26.8</td>
</tr>
<tr>
<td></td>
<td>(0.8)</td>
<td>(0.3)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.2)</td>
<td>(0.3)</td>
</tr>
<tr>
<td>$b = 3$</td>
<td>33.3</td>
<td>9.2</td>
<td>5.0</td>
<td>4.1</td>
<td>4.9</td>
<td>7.3</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td>(0.3)</td>
<td>(0.1)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
</tbody>
</table>

(c) $R$-chart of subgrouped data

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>-0.9</th>
<th>-0.6</th>
<th>-0.3</th>
<th>0.0</th>
<th>0.3</th>
<th>0.6</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b = 1$</td>
<td>358.7</td>
<td>379.7</td>
<td>369.8</td>
<td>364.1</td>
<td>368.8</td>
<td>369.4</td>
<td>376.3</td>
</tr>
<tr>
<td></td>
<td>(3.6)</td>
<td>(3.8)</td>
<td>(3.7)</td>
<td>(3.6)</td>
<td>(3.6)</td>
<td>(3.7)</td>
<td>(3.7)</td>
</tr>
<tr>
<td>$b = 2$</td>
<td>57.3</td>
<td>22.4</td>
<td>15.2</td>
<td>12.8</td>
<td>14.8</td>
<td>20.2</td>
<td>26.5</td>
</tr>
<tr>
<td></td>
<td>(0.6)</td>
<td>(0.2)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.2)</td>
<td>(0.3)</td>
</tr>
<tr>
<td>$b = 3$</td>
<td>22.1</td>
<td>7.9</td>
<td>5.2</td>
<td>4.5</td>
<td>5.1</td>
<td>7.0</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.1)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
</tbody>
</table>

(d) $R$-chart of residuals of subgrouped data

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>-0.9</th>
<th>-0.6</th>
<th>-0.3</th>
<th>0.0</th>
<th>0.3</th>
<th>0.6</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b = 1$</td>
<td>370.8</td>
<td>370.8</td>
<td>363.9</td>
<td>364.1</td>
<td>365.8</td>
<td>362.5</td>
<td>367.4</td>
</tr>
<tr>
<td></td>
<td>(3.8)</td>
<td>(3.6)</td>
<td>(3.7)</td>
<td>(3.6)</td>
<td>(3.6)</td>
<td>(3.6)</td>
<td>(3.6)</td>
</tr>
<tr>
<td>$b = 2$</td>
<td>12.8</td>
<td>13.0</td>
<td>12.9</td>
<td>12.8</td>
<td>12.7</td>
<td>12.9</td>
<td>12.7</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>$b = 3$</td>
<td>4.5</td>
<td>4.4</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
</tr>
</tbody>
</table>