Chapter – 4

Average Run Length of Control Chart for

Markov Dependent Sample Size
4.1. Introduction

Shewhart (1924) has introduced the $\bar{X}$ and $R$ charts in the late 1920s. These charts have found favour with practitioner when dealing with variables data to monitor the large process shift. Page (1954) has presented the CUSUM control chart which is a good alternative when small shifts are important. Robert (1959) has invented the EWMA control chart which is approximately equivalent to the CUSUM control chart in its ability to detect a small process shift. The EWMA can be viewed as a weighted average of all past and current observations. Santhakumaran et al. (1999) have proposed a modified exponentially weighted moving average control chart in which the weights are not a choice, but they are obtained from the nature of the process. So it performs well even in situations where the normality assumptions are violated to slight or moderate degree.

In designing a control chart of a production process larger sample sizes are used to detect smaller shifts. The process shift is relatively large, it is advisable to use small sample sizes. The most desirable situation is to take larger samples very frequently. This is not economically feasible.
A possible way is that either small samples are to be drawn frequently or larger samples are to be selected infrequently. An optimum way is that one who can mix both ways of taking large and small samples to detect all types of process shifts. Sivasamy, Santhakumaran and Subramanian (2000) have studied a control chart of a production process to detect all types of process shifts (large, moderate and small) by taking both the small and large samples under a switching rule of Markov dependent sampling. As comparing it with Shewhart control chart for variable sample size, the method of construction based on Markov dependent sampling is more sensitive than variable sample size chart. Also it has fixed control limits for all the sample sizes instead of transient control limits. It is quite easy to understand and detect out of control signals.

In this chapter, we propose some more impacts such as Markov dependent sampling with classified subgroups through ARL of the control chart are studied. In section 4.2, Markov dependent sampling is explained. Construction of $\bar{X}$ - chart for Markov dependent sample sizes are discussed in section 4.3. In section 4.4, the ARL curves for the $\bar{X}$ chart based on conventional subgroups and Markov dependent sample sizes are presented.
4.2 Markov Dependent Sampling

Suppose one wish to monitor a production process through the construction of control charts based on two different subgroup sizes. Denote the larger sample size as type – 1 and smaller size as type – 2. Each interval of time, select either type – 1 or type – 2 to monitor the process shifts. Start with type – 1, to determine the type of sample for the next inspection, one who uses a random experiment. If a random experiment is performed with two possible outcomes \( R_1 \) and \( \bar{R}_1 \) the probabilities of which are \( a \) and \( 1 - a \). If in this experiment \( R_1 \) occurs, the type – 2 sample is selected, otherwise type – 1 sample is selected. This rule is followed every time the previous selection was type – 1.

If the previous selection was type -2, as similar performs a random experiment with two possible outcomes \( R_2 \) and \( \bar{R}_2 \) the probabilities of which are \( b \) and \( 1 - b \). The type – 1 sample is selected if the event \( R_2 \) occurs in this experiment, otherwise type-2 is selected. Here if \( X_n \) denotes the type of selection at the \( n^{th} \) attempt then the sequence of sample types \( \{X_n; n = 0,1,2,\ldots\} \) forms a MC over the state space \{type – 1, type – 2\} with a unit step transition probability matrix \((tpm) P = (pij)\) of order \( 2 \times 2\).
where

\[ P_{11} = \Pr \{\text{type } -1 / \text{type } -1\} = 1 - a \]
\[ P_{12} = \Pr \{\text{type } -2 / \text{type } -1\} = a \]
\[ P_{21} = \Pr \{\text{type } -1 / \text{type } -2\} = b \]
\[ P_{22} = \Pr \{\text{type } -2 / \text{type } -2\} = 1 - b \]

Thus \( P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \)

The classical analysis of MC provides the stationary probabilities \( \pi_1, \pi_2 \) where \( \pi_j = \text{proportion of time type } -j \ (j = 1, 2) \) sample is being used. Let \( \Pi = (\pi_1, \pi_2) \) be a row vector, \( e = (1,1)' \) a column vector. Solving the system of equation \( \Pi P = \Pi \) and \( \Pi e = 1 \). Its solution is obtained as

\[ \pi_1 = b / (a + b) \]

and

\[ \pi_2 = a / (a + b). \]
4.3 Construction of Control Devices \( \bar{X} \) Chart

Let \( X \) be the study variable and \( X \sim N (\mu, \sigma^2) \). Let us use the following notations.

**Table 4.1**

Estimates of the control chart parameters

<table>
<thead>
<tr>
<th>Item</th>
<th>Type – 1</th>
<th>Type – 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>( n_1 )</td>
<td>( n_2 )</td>
</tr>
<tr>
<td>No. of samples</td>
<td>( k_1 )</td>
<td>( k_2 )</td>
</tr>
<tr>
<td>( i^{th} ) observation</td>
<td>( X_{1i} )</td>
<td>( X_{2i} )</td>
</tr>
<tr>
<td>Sample Mean</td>
<td>( \bar{X}<em>1 = \frac{1}{n_1} \sum</em>{i=1}^{n_1} X_{1i} )</td>
<td>( \bar{X}<em>2 = \frac{1}{n_2} \sum</em>{i=1}^{n_2} X_{2i} )</td>
</tr>
<tr>
<td>Over all Mean</td>
<td>( \bar{X}<em>1 = \frac{1}{k_1} \sum</em>{i=1}^{k_1} \bar{X}_{1i} )</td>
<td>( \bar{X}<em>2 = \frac{1}{k_2} \sum</em>{j=2}^{k_2} \bar{X}_{2j} )</td>
</tr>
<tr>
<td>Sample SD</td>
<td>( S_1 = \sqrt{\frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2} )</td>
<td>( S_2 = \sqrt{\frac{1}{n_2 - 1} \sum_{j=2}^{n_2} (X_{2j} - \bar{X}_2)^2} )</td>
</tr>
<tr>
<td>Over all SD</td>
<td>( S_1 = \frac{1}{k_1} \sum_{i=1}^{k_1} S_{1i} )</td>
<td>( S_2 = \frac{1}{k_2} \sum_{j=2}^{k_2} S_{2j} )</td>
</tr>
</tbody>
</table>

The statistic \( \bar{X} = p_1 \bar{X}_1 + p_2 \bar{X}_2 \) \hspace{1cm} (4.1)

is the best linear unbiased estimator (BLUE) of \( \mu \) since \( \bar{X}_1 \) and \( \bar{X}_2 \) are both unbiased estimators of \( \mu \) and are independent. Further,
\[ X \sim N \left( m \frac{p_1^2}{n_1} + \frac{p_2^2}{n_2} \sigma^2 \right) \]

\[ E(\hat{X}) = \mu \]

and S. E. (\(\hat{X}\)) = \(s \sqrt{\frac{p_1^2}{n_1} + \frac{p_2^2}{n_2}}\) \hspace{1cm} (4.2)

Thus, when the \(\mu\) and \(\sigma\) are known values,

\[ UCL = m + 3s \sqrt{\frac{p_1^2}{n_1} + \frac{p_2^2}{n_2}} \]

\[ CL = m \]

\[ LCL = m - 3s \sqrt{\frac{p_1^2}{n_1} + \frac{p_2^2}{n_2}} \] \hspace{1cm} (4.3)

On the other hand, when \(\mu\) and \(\sigma\) are not known, these are estimated from the observed \(k_1\) type -1 samples and \(k_2\) type -2 samples as described below.

\[ \hat{\mu} = \text{estimator of } \mu \]

\[ = \bar{X} \]

\[ = p_1 \bar{X}_1 + p_2 \bar{X}_2 \] \hspace{1cm} (4.4)

\[ \hat{\sigma} = \text{estimator of } \sigma \]

\[ = p_1 \frac{\hat{S}_1}{c_4(n_1)} + p_2 \frac{\hat{S}_2}{c_4(n_2)} \] \hspace{1cm} (4.5)

Each one of the sample standard derivations \(S_1\) and \(S_2\) is unbiased estimator of \(\sigma\). If \(X \sim N (\mu, \sigma^2)\), then \((n_1 - 1)S_1^2 / s^2 : c^2_{(n_1 - 1)}\) and
\[(n_2-1)S_2^2/s^2 \cdot c_4^2(n_2-1)\]. Further \(E(S_1) = c_4(n_1)\sigma\), i.e., \(S_1\) is actually the estimator of \(c_4(n_1)\sigma\) and \(E(S_2) = c_4(n_2)\sigma\), i.e., \(S_2\) is the estimator of \(c_4(n_2)\sigma\) and they are independent where \(c_4(n)\) is a constant that depends on any samples of size \(n\) whose value equals

\[
c_4(n) = \sqrt{\frac{2G}{G(n-1)}}
\]  

(4.6)

Since \(S_1/c_4(n_1)\) and \(S_2/c_4(n_2)\) are two unbiased estimators of \(\sigma\), the best linear estimator \(\hat{\sigma}\) of \(\sigma\) given by

\[
\hat{\sigma} = \frac{S_1}{c_4(n_1)} + \frac{S_2}{c_4(n_2)}
\]  

(4.7)

Thus from expression (4.7), it is seen that estimate of \(\sigma\) is to be computed as given in equation (4.5) from any observed MDS. The values of \(c_4(n)\) for different values of \(n\) are available in statistical tables. Using equations (4.4) and (4.5), one could obtain the control limits on \(\hat{x}\) chart as stated below.

\[
UCL = X + 3\left[p_1^2/n_1 + p_2^2/n_2\right]^{1/2} \{p_1 \hat{S}_1/c_4(n_1) + p_2 \hat{S}_2/c_4(n_2)\}
\]

\[
CL = X
\]  

(4.8)

\[
LCL = X - 3\left[p_1^2/n_1 + p_2^2/n_2\right]^{1/2} \{p_1 \hat{S}_1/c_4(n_1) + p_2 \hat{S}_2/c_4(n_2)\} \]
Let the constant be $A_3(n_1) = \frac{3}{\sqrt{n_1 c_4(n_1)}}$ and $A_3(n_2) = \frac{3}{\sqrt{n_2 c_4(n_2)}}$. Then the $\bar{x}$ chart control chart limits become

$$UCL = \bar{x} + \left\{ \sqrt{n_1} S_1 A_3(n_1) p_1 + \sqrt{n_2} S_2 A_3(n_2) p_2 \right\} \sqrt{\frac{p_1^2}{n_1} + \frac{p_2^2}{n_2}}$$

$$CL = \bar{x}$$

(4.9)

$$LCL = \bar{x} - \left\{ \sqrt{n_1} S_1 A_3(n_1) p_1 + \sqrt{n_2} S_2 A_3(n_2) p_2 \right\} \sqrt{\frac{p_1^2}{n_1} + \frac{p_2^2}{n_2}}$$

### 4.4 Average Run Length

The $ARL$ of a control chart at a given quality level is the average number of samples taken before action is taken. As in Montgomery (2001) for any control chart the $ARL$ is calculated as $ARL = \frac{1}{p}$ where $p$ is the probability that any point exceeds the control limits. We obtain the $ARL$ with respect to mean shift ($i.e., \delta = \mu + k\sigma$) for the control chart which is given in Table 4.1.

$ARL$ for $\bar{x}$-chart based on Markov dependent sample sizes is $1/(1-\beta)$

where

$b = f \left( 3 - k \cdot \frac{\bar{\sigma}}{\bar{\delta}} \right)$

and $f$ is the distribution function of $N(0,1)$. 
Table -4.1

ARL for the mean shift $\mu = \mu + k\sigma$

<table>
<thead>
<tr>
<th>$n_1$</th>
<th>$N_2$</th>
<th>.1</th>
<th>.3</th>
<th>.5</th>
<th>.7</th>
<th>.9</th>
<th>1.1</th>
<th>1.3</th>
<th>1.5</th>
<th>1.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>$\uparrow$</td>
<td>341.81</td>
<td>254.88</td>
<td>169.91</td>
<td>106.55</td>
<td>67.17</td>
<td>43.25</td>
<td>28.59</td>
<td>19.43</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>*</td>
<td>311.49</td>
<td>145.46</td>
<td>60.66</td>
<td>27.28</td>
<td>13.53</td>
<td>7.41</td>
<td>4.46</td>
<td>2.94</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>$\uparrow$</td>
<td>332.49</td>
<td>211.16</td>
<td>114.10</td>
<td>61.50</td>
<td>34.51</td>
<td>20.33</td>
<td>12.59</td>
<td>8.19</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>*</td>
<td>280.69</td>
<td>90.36</td>
<td>29.50</td>
<td>11.50</td>
<td>5.36</td>
<td>2.97</td>
<td>1.93</td>
<td>1.43</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>$\uparrow$</td>
<td>323.68</td>
<td>179.36</td>
<td>85.66</td>
<td>42.23</td>
<td>22.22</td>
<td>12.54</td>
<td>7.58</td>
<td>4.91</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>*</td>
<td>256.80</td>
<td>64.35</td>
<td>18.32</td>
<td>6.75</td>
<td>3.18</td>
<td>1.89</td>
<td>1.39</td>
<td>1.13</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>$\uparrow$</td>
<td>335.56</td>
<td>215.59</td>
<td>118.51</td>
<td>64.71</td>
<td>36.65</td>
<td>21.74</td>
<td>13.52</td>
<td>8.82</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>*</td>
<td>274.03</td>
<td>82.04</td>
<td>25.69</td>
<td>9.81</td>
<td>4.57</td>
<td>2.57</td>
<td>1.71</td>
<td>1.31</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>$\uparrow$</td>
<td>314.89</td>
<td>153.98</td>
<td>66.48</td>
<td>30.59</td>
<td>15.39</td>
<td>8.47</td>
<td>5.09</td>
<td>3.33</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>*</td>
<td>229.36</td>
<td>44.19</td>
<td>11.15</td>
<td>4.09</td>
<td>2.02</td>
<td>1.35</td>
<td>1.10</td>
<td>1.02</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>$\uparrow$</td>
<td>308.15</td>
<td>137.68</td>
<td>55.60</td>
<td>24.49</td>
<td>12.01</td>
<td>6.54</td>
<td>3.95</td>
<td>2.63</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>*</td>
<td>214.60</td>
<td>36.16</td>
<td>8.67</td>
<td>3.16</td>
<td>1.68</td>
<td>1.20</td>
<td>1.05</td>
<td>1.01</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>$\uparrow$</td>
<td>322.49</td>
<td>175.66</td>
<td>82.68</td>
<td>40.35</td>
<td>21.09</td>
<td>11.85</td>
<td>7.15</td>
<td>4.63</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>*</td>
<td>242.68</td>
<td>52.98</td>
<td>14.12</td>
<td>5.12</td>
<td>2.48</td>
<td>1.55</td>
<td>1.19</td>
<td>1.06</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>$\uparrow$</td>
<td>299.11</td>
<td>119.22</td>
<td>44.84</td>
<td>18.66</td>
<td>8.92</td>
<td>4.84</td>
<td>2.97</td>
<td>2.04</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>*</td>
<td>195.48</td>
<td>27.83</td>
<td>6.34</td>
<td>2.38</td>
<td>1.38</td>
<td>1.09</td>
<td>1.01</td>
<td>1.00</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>$\uparrow$</td>
<td>292.90</td>
<td>108.36</td>
<td>38.52</td>
<td>15.71</td>
<td>7.42</td>
<td>4.04</td>
<td>2.52</td>
<td>1.77</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>*</td>
<td>184.14</td>
<td>23.77</td>
<td>5.29</td>
<td>2.05</td>
<td>1.26</td>
<td>1.05</td>
<td>1.01</td>
<td>1.00</td>
</tr>
</tbody>
</table>

$\uparrow$ ARL for variable sample size   *- ARL for Markov dependent sample size
1. ARL Curve for Variables sample size when $n_1 = 1$, $n_2 = 5$

2. ARL Curve for MDS when $n_1 = 1$, $n_2 = 5$

---

1. ARL Curve for Variables sample size when $n_1 = 3$, $n_2 = 7$

2. ARL Curve for MDS when $n_1 = 3$, $n_2 = 7$
1. ARL Curve for Variables sample size when $n_1 = 7$, $n_2 = 6$

2. ARL Curve for MDS when $n_1 = 7$, $n_2 = 6$

1. ARL Curve for Variables sample size when $n_1 = 10$, $n_2 = 8$

2. ARL Curve for MDS when $n_1 = 10$, $n_2 = 8$
Figures 4.1, 4.2, 4.3 and 4.4 present these ARL curves for selected type -1 and type -2 sample sizes with \( \pi_1 = 5/9 \) and \( \pi_2 = 4/9 \) where the ARL is in terms of the expected number of samples taken in order to detect the shift. This is the out of control ARL (samples) when the process is in control. For example to illustrate the use of the curves, we may note that in order to detect a shift 1.1\( \sigma \) in variable sample size chart with \( n_1 = 6 \) and \( n_2 = 5 \) will require that approximately 9 samples be sampled but in Markov dependent sample size will require only two samples be sampled on the average. Also note that the ARL for MDS coverage's rapidly to one as the sample sizes \( n_1 \) and \( n_2 \) are slowly increasing as compared with the ARL for variable sample size.

_The Markov dependent sample size control chart can detect all types of production process shifts (small, moderate and large). The evaluation of the ARL involves a normal distribution with stationary probabilities of the selection of the sample types._