CHAPTER 2 THEORY AND LITERATURE REVIEW

2.1 Theoretical Background

2.1.1 Principle of quality control chart

Pongchavalit and Pongpullponsak (2007) found that quality control chart can be used to demonstrate the quality measured in a given time period. Measuring of quality calculates from sample set, where for quality control chart an average value or range of the samples may be used in calculation. Every value plotted in the graph appears in linear relationships which show the quality pattern of the process. For three control limits including upper control limit (UCL), central limit (CL), and lower control limit (LCL), they can be calculated from random sampling which have also been included in the chart. The process will be in control when every point falls between UCL and LCL, otherwise the process will be out of control (Figure 2.1).

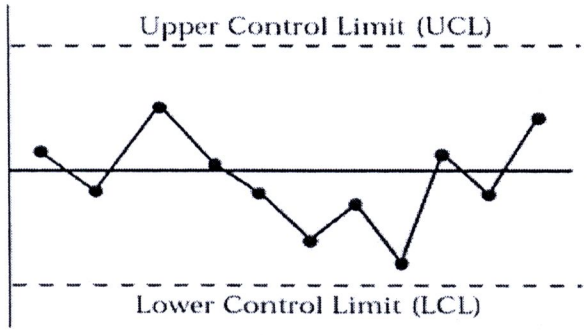

Figure 2.1 Typical Control Chart (Montgomery, 2001).

The steps for construction of a quality control chart are as below;

1. Determining the criterion needed to control or the objective of the control. This depends on demands of manufacturer and the control chart type selected. However, one needed to been controlled is the average of physical properties or quality of products.

2. Determining the sample number. The random sampling method used in the study will depend on the selected control chart type, cost in investigation, and volume of production.

3. Determining frequency of data collection. The frequency of random sampling is important in establishment of control chart. In general, there are 2 methods: 1) sampling based on certain interval period and 2) sampling in a given period.

4. Collecting the data. Data recording in a table will be different depending on the type of control chart where the data will then be used to calculate the control limits.

5. Calculating control limits. In establishment of quality control chart, the number of products that will be eliminated, which depends on the percentage of the data falling out of the control limits, has to be concerned. In this study, 3 σ was used. For control limits containing UCL, CL, and LCL, the data observed from the samples will be used in calculation.

6. Plotting and analysis the chart. After plotting the data on the control chart, distribution of each point on the chart will illustrate condition of the production. The process controller has to investigate the production process if the points on the control chart show as Figure 2.2-2.8.

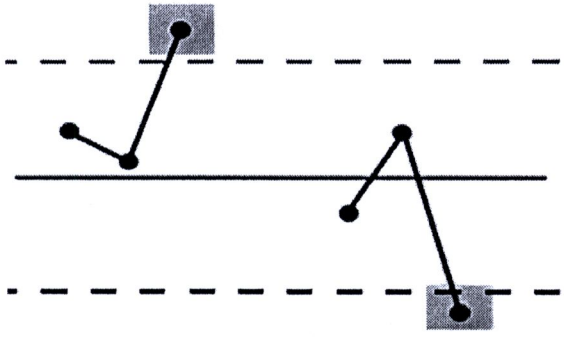

Figure 2.2 One or more points outside the control limits (Montgomery, 2001).

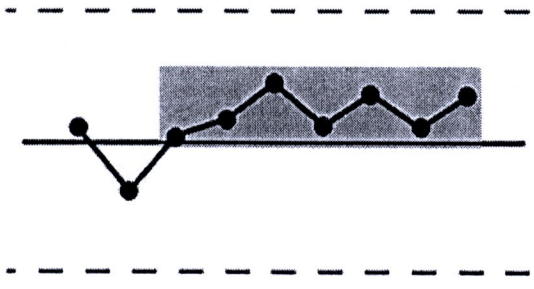

Figure 2.3 Seven or more consecutive points on one side of the centerline (Montgomery, 2001).

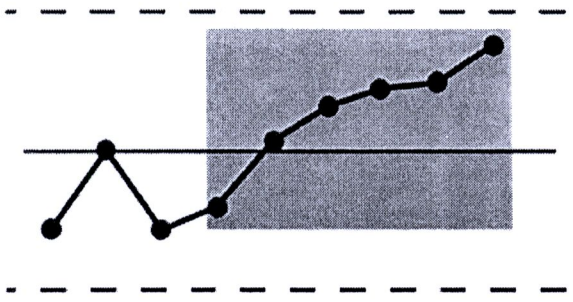

Figure 2.4 Six points in a row steadily increasing or decreasing (Montgomery, 2001).

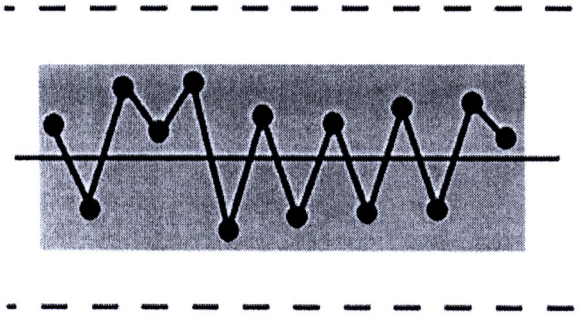

Figure 2.5 Fourteen points alternating up and down (Montgomery, 2001).

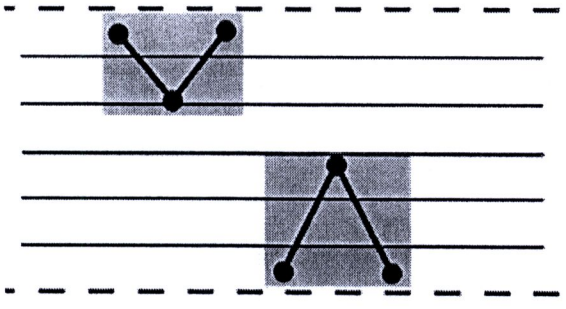

Figure 2.6 Two out of three consecutive points in the outer third of the control region (Montgomery, 2001).

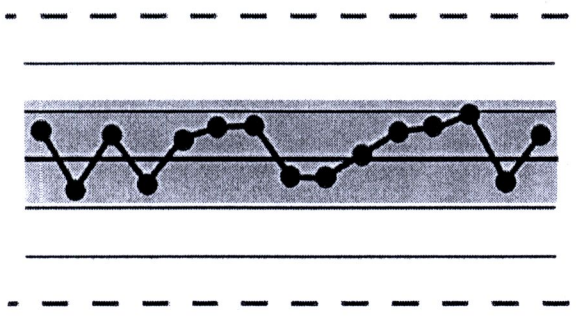

Figure 2.7 Fifteen points in a row within the center third of the control region (Montgomery, 2001).

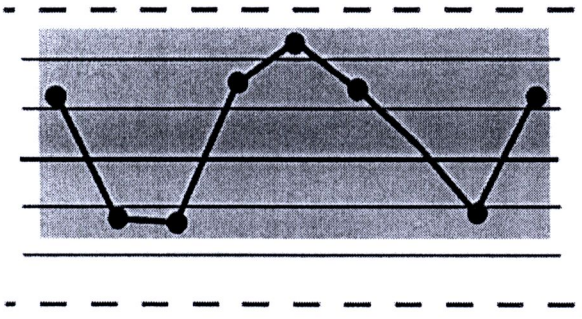

Figure 2.8 Eight points on both sides of the centerline with none in the center third of the control region (Montgomery, 2001).

7. Improving the control chart. Each point on the control chart that demonstrates abnormality will be deleted. The remaining points will be calculated for control limits before establishing a new chart. The improved control chart will be ready to use for controlling production process.

2.1.2 Expected Value

Expected value of average of random variable X, referred as $E[X]$. According to Ross (1996), it has been defined by

$$E[X] = \begin{cases} \int_{-\infty}^{\infty} xf(x)dx & \text{if } x \text{ is continuous} \\ \sum_{x} xP\left(X = x\right) & \text{if } x \text{ is discrete} \end{cases} \tag{2.1}$$

2.1.3 Normal distribution

Normal distribution is the most important probability distribution for statistical analysis because the occurrence of most events fit to this distribution pattern. The normal probability distribution of any random variable x can be defined as followed. Standard normal distribution is normal distribution with $\mu = 0, \sigma^2 = 1$ and can be

written by using the probability density function as

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{2.2}$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{u^2}{2}\right) du \tag{2.3}$$

2.1.4 Ranked set sampling

Ranked set sampling (RSS) has been proposed by McIntyre (1952). The samples obtained by this method will be ranked using other variables that relate to the variable of interest or the variable to be actual measurement. The steps in random ranked set sampling are described as below;

Step 1 random sampling for n sets, each set composes of n samples.

Step 2 ranking samples from the same set by using variables that relate to the variable of interest in performance.

Step 3 selecting samples for actual measurement, starting from the first smallest value of the 1^{st} set, then the second smallest value of 2^{nd} set until the last set; the maximum value will be selected.

Step 4 repeated step 1 through step 3 for cycles until obtaining the samples enough for actual measurement ($s = nr$).

To explain more for this method, assuming that 3 sample sets were random sampling to collect 3 samples/set and the sampling were repeated 4 cycles ($r = 4$). This can be concluded as a diagram below;

Round	Rank		
	1	2	3
1	@	♦	♦
	♦	@	♦
	♦	♦	@
2	@	♦	♦
	♦	@	♦
	♦	♦	@
3	@	♦	♦
	♦	@	♦
	♦	♦	@
4	@	♦	♦
	♦	@	♦
	♦	♦	@

Figure 2.9 Show the sample units for RSS.

From Figure 2.9, each row is sample unit that has been ranked into each sample set. The sample units marked with the symbol @ are selected to perform actual measurement. Therefore, the total of sample units that are random sampling from 4 cycles will be 36 samples but only 12 samples will be measured the criterion of interest. Moreover, in some circumstances, it might be difficult to make a decision on sample ranking. To avoid this situation, another characteristic that relates to the criterion of interest and requires lower costs to perform might be used in the ranking

instead. McIntyre (1952) proposed an estimator of population mean for ranked set sampling as below;

$$\bar{X}_{rss,j} = \frac{1}{n}\sum_{i=1}^{n} X_{(i:n)j}, \qquad j = 1, 2, ..., r \tag{2.4}$$

which is the unbiased estimator of μ and $X_{(i:n)j}$ is the i^{th} observed value in the j^{th} cycle which is chosen by using RSS.
Dell and Clustter (1972) demonstrated that variance of $\bar{X}_{rss,j}$ is

$$Var\left(\bar{X}_{rss,j}\right) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2_{(i:n)} \tag{2.5}$$

where $\sigma^2_{(i:n)}$ is variance of the i^{th} order statistic which is chosen by using RSS.

2.1.5 Median ranked set sampling

Median ranked set sampling (MRSS) was presented by Muttlak (1997). Using this method, the sample at the median of the sets is selected, if the set size is odd. If the set size is even, sample selection is from the $(n/2)^{th}$ order in the first half and the $((n+2)/2)^{th}$ order in the second half of the set. The method of MRSS can be concluded as following;

Step 1 select n sample units per set from the total n sample sets.

Step 2 allocate sample units into the set by using a variable related to a variable of interest in ranking.

Step 3 choose the sample units for actual measurement by selecting the smallest rank in the $((n+1)/2)^{th}$ order from the sample sizes with an odd number. For the sample sets with an even number, the smallest rank in the $(n/2)^{th}$ order of the first half and the smallest rank in the $((n+2)/2)^{th}$ order in the second half are chosen.

Step 4 repeated step 1 through step 3 for cycles until obtaining the samples enough for actual measurement ($s = nr$).
As seen in Figure 2.10, samples are randomly selected for 3 sets, where each set contains 3 sample units, then repeat this procedure for 4 times.

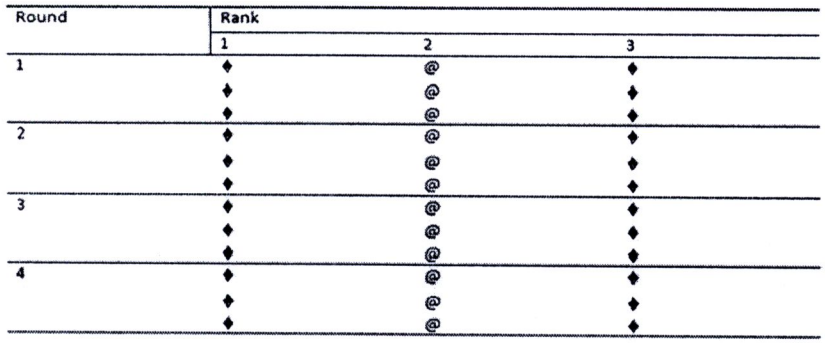

Figure 2.10 Show the sample units for MRSS in case I.

In the illustration, each row contains ranked sample units, and only the sample units marked as @ is chosen for actual measurement. Therefore, from the total 36 sample units selected for 4 cycles, only 12 units will be used for measuring a variable of interest. As seen in Figure 2.11, samples are randomly selected for 4 sets, where each set contains 4 sample units, then repeat this procedure for 4 times.

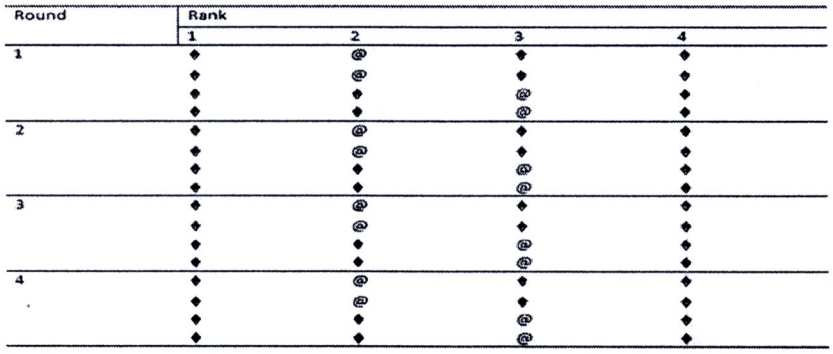

Figure 2.11 Show the sample units for MRSS in case II.

In the illustration, each row contains ranked sample units, and only the sample units marked as @ is chosen for actual measurement. Therefore, from the total 36 sample units selected for 4 cycles, only 12 units will be used for measuring a variable of interest. For the sample sizes with an odd number, given $X_{(i:m)j}$ is the $(n/2)^{th}$ order statistic of the i^{th} set from sample size n in the i^{th} cycle. In case of an even number, $X_{(i:m)j}$ is the $(n/2)^{th}$ order statistic of the i^{th} set from sample size $n(i = 1, 2, ..., L = n/2)$, and the $((n+2)/2)^{th}$ order statistic of the i^{th} set from sample size $n(i = L+1, L+2, ..., n)$, where Muttlak (1997) proposed an estimator of population mean for MRSS as below;

$$\bar{X}_{mrss,j} = \frac{1}{n}\sum_{i=1}^{n} X_{(i:m)j}, \qquad j = 1, 2, ..., r \tag{2.6}$$

where $X_{(i:m)j}$ is the i^{th} observed value in the j^{th} cycle which is chosen by using MRSS.
The variance of $\bar{X}_{mrss,j}$ is defined by

$$Var\left(\bar{X}_{mrss,j}\right) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2_{(i:m)} \tag{2.7}$$

where $\sigma^2_{(i:m)}$ is variance of the i^{th} order statistic which is chosen by using MRSS.

2.1.6 Ranked set sampling for multiple characteristics

Generally, in RSS only one characteristic is used in ranking samples, where ranking errors may easily occur. Thus it is recommended to consider multiple characteristics in ranking to reduce errors. Ridout (2003) described the procedure for selection of characteristics used in ranking below;

Step 1 select characteristics of interest, then designate as $f = 1, f = 2$ and so

on.

Step 2 let n_{fg} denote the number of samples ranked in the g^{th} set for the f^{th} characteristic, and V_f is the sample variance of the f^{th} characteristic in the sample number n_{fg}. For a balanced sampling, variance of the f^{th} characteristic equals zero ($V_f = 0$). When no ranking error occurs, a new set of size n is established by calculating the V_f values from each characteristic in the sample set to estimate the values of V_1 and V_2.

Step 3 select the sample that minimizes $V = V_1 + V_2$ from each sample set; in case of absence of ranking errors, select sample in the same way as MRSS; if there is more than one such sample, select one of them at random.

Step 4 repeated step 1 through step 3 for cycles until the desired sample size is obtained ($s = nr$).

From above to show procedure of ranked set sampling for multiple characteristics (RSSMC) as following, give the weight which is the variable of interest or the variable to be actual measurement, the high and width which are variable to may be relation with the weight by the high and width are variable of ranking. Follow the example;

Characteristic	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
High(A_1)	63.17	72.61	65.99	69.28	79.62
Width(A_2)	19.96	16.99	12.12	14.82	16.99
Weight	56.22	50.93	44.17	49.71[b]	51.43
V_1	48.54	6.13	17.14	0.73	90.03
V_2	14.33	0.66	16.46	1.83	0.66
V	62.87	6.79	33.60	2.57[c]	90.70

Note: [a] In each set, the data are obtained from the same method until we have n samples from n sets repeated for r rounds;
[b] is the sample selected by calculating from $V_f = (a_{fg} - \bar{a}_f)^2$;
[c] is a minimum value of $V_1 + V_2$ when A is the variable of characteristic used in ranking.

Considering RSSMC, a selected sample is the one with the smallest variance or the one that is the median of the samples ranked by multiple characteristics. Muttlak and AL-Sabah (2003) developed a control chart using the samples that are the median of MRSS data. Given $X_{(i:mc)j}$ is the order statistic that has the smallest sum variance in the f^{th} set with sample size n in the j^{th} cycle, where the RSS estimator for multiple characteristics could be calculated by

$$\bar{X}_{rssmc,j} = \frac{1}{n}\sum_{i=1}^{n} X_{(i:mc)j}, \qquad j = 1, 2, ..., r \tag{2.8}$$

where $X_{(i:mc)j}$ is the i^{th} observed value in the j^{th} cycle which is chosen by using RSSMC.

The variance of $\bar{X}_{rssmc,j}$ is defined by

$$Var\left(\bar{X}_{rssmc,j}\right) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2_{(i:mc)} \tag{2.9}$$

where $\sigma^2_{(i:mc)}$ is variance of the i^{th} order statistic which is chosen by using RSSMC.

2.1.7 Control chart for mean using simple random sampling

Let X_{ij} when $i = 1, 2, ..., n$ and $j = 1, 2, ..., r$ is the i^{th} sample unit in the j^{th} cycle of sample size n, and $X_{ij} \sim N(\mu, \sigma^2)$. When the population mean and variance are μ and σ^2, respectively, the Shewhart (1924) control chart for

$$\bar{X}_j = \frac{1}{n}\sum_{i=1}^{n} X_{ij}, \qquad j = 1, 2, ..., r \tag{2.10}$$

is defined by

$$\begin{aligned} UCL &= \mu + 3\frac{\sigma}{\sqrt{n}} \\ CL &= \mu \\ LCL &= \mu - 3\frac{\sigma}{\sqrt{n}} \end{aligned} \tag{2.11}$$

when UCL, CL and LCL are upper control limit, central limit, and lower control limit, respectively. After obtained the chart, the sample mean $\bar{X}_j, j = 1, 2, ..., r$ can be plotted into the upper control chart (Montgomery, 2001). However, in actual measurement, the mean μ and variance σ^2 are unknown so both μ and σ are estimated from the collected data, where the unbiased estimator for μ is

$$\bar{\bar{X}} = \frac{1}{r}\sum_{j=1}^{r} \bar{X}_j \tag{2.12}$$

but

$$\bar{S} = \frac{1}{r}\sum_{j=1}^{r} S_j \tag{2.13}$$

where

$$S_j = \left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_{ij} - \bar{X}_j\right)^2\right]^{1/2} \tag{2.14}$$

is an biased estimate for σ. We can use $\bar{S}/c_4$ as the unbiased estimate for σ where c_4 is calculated from

$$c_4 = \left(\frac{2}{n} - 1\right)^2 \frac{\Gamma(n-2)}{\Gamma[(n-1)/2]} \tag{2.15}$$

and the control chart for sample mean can be expressed as

$$\begin{aligned} UCL &= \bar{\bar{X}} + 3\frac{\bar{S}}{c_4\sqrt{n}} \\ CL &= \bar{\bar{X}} \\ LCL &= \bar{\bar{X}} - 3\frac{\bar{S}}{c_4\sqrt{n}} \end{aligned} \tag{2.16}$$

the sample means $\bar{X}_j, j = 1, 2, ..., r$ can now be plotted in the upper control chart (Montgomery, 2001).

2.1.8 Control chart for mean using ranked set sampling

The RSS mean $\bar{x}_{rss,j}$ at the j^{th} cycle can be plotted in the control chart based on RSS proposed by Salazar and Sinha (1997).

$$\begin{aligned} UCL &= \mu + 3\sigma_{\bar{X}_{rss}} \\ CL &= \mu \\ LCL &= \mu - 3\sigma_{\bar{X}_{rss}} \end{aligned} \tag{2.17}$$

when

$$\sigma_{\bar{X}_{rss}} = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2_{(i:n)}} \tag{2.18}$$

where the unbiased estimate for RSS (Takahashi and Wakimoto, 1968) is

$$\bar{X}_{rss} = \frac{1}{r}\sum_{j=1}^{r}\bar{X}_{rss,j} \tag{2.19}$$

The estimate for $\sigma_{\bar{X}_{rss}}$ suggested by Muttlak and AL-Sabah (2003) will be as

$$\hat{\sigma}_{\bar{X}_{rss}} = \left[\frac{1}{n}\hat{\sigma}^2_{rss} - \frac{1}{n^2}\sum_{i=1}^{n}\left(\bar{X}_{(i)} - \bar{X}_{rss}\right)^2\right]^{1/2} \tag{2.20}$$

when

$$\hat{\sigma}^2_{rss} = \frac{1}{nr-1}\sum_{i=1}^{n}\sum_{j=1}^{r}\left(X_{(i:n)j} - \bar{X}_{rss}\right)^2 \tag{2.21}$$

and

$$\bar{X}_{(i)} = \frac{1}{r}\sum_{j=1}^{r}X_{(i:n)j} \tag{2.22}$$

is the estimate for population mean of the i^{th} order statistic. Subsequently, the control chart can be constructed by using $\bar{X}_{rss}$ and $\hat{\sigma}_{\bar{X}_{rss}}$ as the equation below;

$$\begin{aligned} UCL &= \bar{X}_{rss} + 3\hat{\sigma}_{\bar{X}_{rss}} \\ CL &= \bar{X}_{rss} \\ LCL &= \bar{X}_{rss} - 3\hat{\sigma}_{\bar{X}_{rss}} \end{aligned} \tag{2.23}$$

2.1.9 Control chart for mean using median ranked set sampling

The MRSS mean $\bar{X}_{mrss,j}$ of the j^{th} cycle can be plotted in the control chart based on MRSS proposed by Muttlak (1997).

$$\begin{aligned} UCL &= \mu + 3\sigma_{\bar{X}_{mrss}} \\ CL &= \mu \\ LCL &= \mu - 3\sigma_{\bar{X}_{mrss}} \end{aligned} \tag{2.24}$$

where

$$\sigma_{\bar{X}_{mrss}} = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2_{(i:m)}} \tag{2.25}$$

In practical, the values of μ and $\sigma_{\bar{X}_{mrss}}$ are unknown so an unbiased estimator μ is calculated from MRSS data with normal distribution as shown below;

$$\bar{X}_{mrss} = \frac{1}{r}\sum_{i=1}^{r} \bar{X}_{mrss,j} \tag{2.26}$$

and the estimate for $\sigma_{\bar{X}_{mrss}}$ suggested by Muttlak and AL-Sabah (2003) is given by

$$\hat{\sigma}_{\bar{X}_{mrss}} = \left[\frac{1}{n(nr-1)}\sum_{j=1}^{r}\sum_{i=1}^{n}\left(X_{(i:m)j} - \bar{X}_{mrss}\right)^2\right]^{1/2} \tag{2.27}$$

where $X_{(i:m)j}$ is the estimator for population mean of the i^{th} order statistic. Thus $\bar{X}_{mrss}$ and $\hat{\sigma}_{\bar{X}_{mrss}}$ can be used to construct the control chart from MRSS as following;

$$\begin{aligned} UCL &= \bar{X}_{mrss} + 3\hat{\sigma}_{\bar{X}_{mrss}} \\ CL &= \bar{X}_{mrss} \\ LCL &= \bar{X}_{mrss} - 3\hat{\sigma}_{\bar{X}_{mrss}} \end{aligned} \tag{2.28}$$

2.1.10 Mean Vectors and Covariance Matrices

Suppose $X' = [X_1, X_2, ..., X_p]$ is a $p \times 1$ random vector. Then each element of X is a random variable with its own marginal probability distribution. The marginal means μ_i and variances σ_i^2 are defined as $\mu_i = E(X_i)$ and $\sigma_i^2 = E\left(X_i - \mu_i\right)^2, i = 1, 2, ..., p$ respectively. Specifically;

$$\mu_i = \begin{cases} \int_{-\infty}^{\infty} x_i f_i(x_i)dx_i & \text{if } x_i \text{ is a continuous random variable} \\ & \text{with probability density function } f_i(x_i) \\ \sum_{allx_i} x_i p_i(x_i) & \text{if } x_i \text{ is a discrete random variable} \\ & \text{with probability density function } p_i(x_i) \end{cases}$$

$$\sigma_i^2 = \begin{cases} \int_{-\infty}^{\infty} (x_i - \mu_i)^2 f_i(x_i)dx_i & \text{if } x_i \text{ is a continuous random variable} \\ & \text{with probability density function } f_i(x_i) \\ \sum_{allx_i} (x_i - \mu_i)^2 p_i(x_i) & \text{if } x_i \text{ is a discrete random variable} \\ & \text{with probability density function } p_i(x_i) \end{cases} \tag{2.29}$$

It will be convenient in later sections to denote the marginal variances by σ_{ii} rather than the more traditional σ_i^2, and consequently, we shell adopt this notation.
The behavior of any pair of random variables, such as X_i and X_k is described by their joint probability function, and a measure of the linear association between them is provided by the covariance

$$\sigma_{ik} = \begin{cases} \int_{-\infty}^{\infty} (x_i - \mu_i)^2 f_i(x_i)dx_i & \text{if } x_i, x_k \text{ is a continuous random variable} \\ & \text{with probability density function } f_{ik}(x_i, x_k) \\ \sum_{allx_i} (x_i - \mu_i)^2 p_i(x_i) & \text{if } x_i, x_k \text{ is a discrete random variable} \\ & \text{with probability density function } p_{ik}(x_i, x_k) \end{cases} \tag{2.30}$$

and μ_i and μ_k, $k = 1, 2, ...p$, are the marginal means. When $i = k$, the covariance becomes the marginal variances.
More generally, the collective behavior of the p random variables $X_1, X_2, ..., X_p$ or, equivalently, the random vector $X' = [X_1, X_2, ..., X_p]$, is described by a joint probability density function $f(x_1, x_2, ...x_p) = f(x)$. As we have already noted in this book $f(x)$ will often be the multivariate normal density function.
If the joint probability $P[X_i \leq x_i \text{ and } X_k \leq x_k]$ can be written as the product of the corresponding marginal probabilities, so that

$$P[X_i \leq x_i \text{ and } X_k \leq x_k] = P[X_i \leq x_i] P[X_k \leq x_k] \tag{2.31}$$

for all pairs of value x_i, x_k then X_i and X_k are said to be statistically independent. When X_i and X_k are continuous random variables with joint density $f_{ik}(x_i, x_k)$ and marginal densities $f_i(x_i)$ and $f_k(x_k)$, the independence condition becomes

$$f_{ik}(x_i, x_k) = f_i(x_i) f_k(x_k) \tag{2.32}$$

for all pairs (x_i, x_k).
The p continuous random variables $X_1, X_2, ..., X_p$ are mutually statistically independent if their joint density can be factored as

$$f_{12..p}(x_1, x_2, ..., x_p) = f_1(x_1) f_2(x_2) ... f_p(x_p) \tag{2.33}$$

for all p-tuples $(x_1, x_2, ..., x_p)$. Statistical independences has an important implication for covariance. The factorization in (2.33) implies that $Cov(X_i, X_k) = 0$. Thus

$$Cov(X_i, X_k) = 0 \qquad \text{if } X_i \text{ and } X_k \text{ are independent} \tag{2.34}$$

the converse of (2.34) is not true in general; there are situations where $Cov(X_i, X_k) = 0$, but X_i and X_k are not independent.
The means and covariances of the $p \times 1$ random vector X can be set out as matrices. The expected value of each element is contained in the vector of means $\mu = E(X)$, and the p variances σ_{ii} and the $p(p-1)/2$ distinct covariances $\sigma_{ik} (i < k)$ are contained in the symmetric variance-covariance matrix $\Sigma = E(X - \mu)(X - \mu)'$. Specifically;

$$E[X] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_p] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \mu \tag{2.35}$$

and $\Sigma = E(X - \mu)(X - \mu)'$

$$= E \left(\begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_p - \mu_p \end{bmatrix} \begin{bmatrix} X_1 - \mu_1, X_2 - \mu_2, \ldots, X_p - \mu_p \end{bmatrix} \right)$$

$$= E \begin{bmatrix} (X_1 - \mu_1)^2 & (X_1 - \mu_1)(X_2 - \mu_2) & \ldots & (X_1 - \mu_1)(X_p - \mu_p) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)^2 & \ldots & (X_2 - \mu_2)(X_p - \mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ (X_p - \mu_p)(X_1 - \mu_1) & (X_p - \mu_p)(X_2 - \mu_2) & \ldots & (X_p - \mu_p)^2 \end{bmatrix}$$

$$= \begin{bmatrix} E(X_1-\mu_1)^2 & E(X_1-\mu_1)(X_2-\mu_2) & \cdots & E(X_1-\mu_1)(X_p-\mu_p) \\ E(X_2-\mu_2)(X_1-\mu_1) & E(X_2-\mu_2)^2 & \cdots & E(X_2-\mu_2)(X_p-\mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_p-\mu_p)(X_1-\mu_1) & E(X_p-\mu_p)(X_2-\mu_2) & \cdots & E(X_p-\mu_p)^2 \end{bmatrix}$$

$$\Sigma = Cov(X) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} \tag{2.36}$$

2.1.11 Multivariate control chart (bivariate control chart)

Generally, the control chart that is widely used can control only one characteristic but in real situation it always has more than one characteristic to control. For example, both length and diameter of a pipeline are needed to be controlled at the same time to ensure that the length and diameter of the pipeline are acceptable for operating conditions. Unless both characteristics have been controlled dependently, the products might be unacceptable.

Controlling Several Related Quality Characteristics

Giving that two characteristics of production process are controlled under control protocol. If the average control charts of these two characteristics have been constructed independently, the control area can be drawn in a rectangular $ABCD$ as illustrating in Figure 2.12. The edge of the square will be UCL and LCL of both characteristics and shown in a geometric moving average. If bivariate observation of sample means $(\bar{X}_1, \bar{X}_2)$ has been plotted into the $ABCD$ area, the process seems to always be in control.

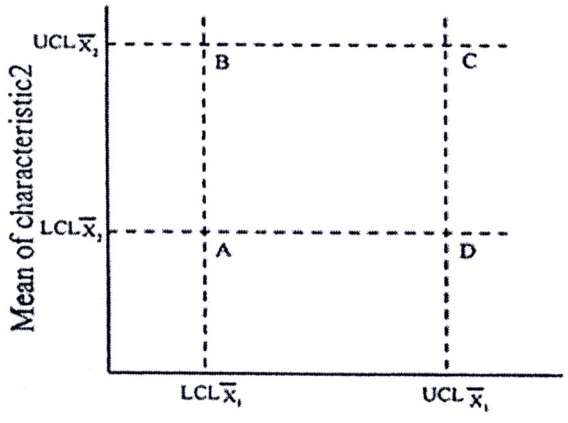

Figure 2.12 Rectangular control region when control charts are constructed independently.

As a result, using rectangular margin might be not right. Actually, the control area of two characteristics is naturally in an ellipse and the statistics of two characteristics are calculated using an elliptical function. If both characteristics are independent, the major axis and minor axis will be in parallel to the axis (Figure 2.12). If a pair of sample average $(\bar{X}_1, \bar{X}_2)$ falls within the elliptical area, the process will be in control. But if both characteristics have negative correlation, the elliptical shape will be similar to the ellipse B. Likewise, if both characteristics have positive correlation, the ellipse will look like the picture C.
As seen in Figure 2.13, if the variables have positive correlation and the control area is in a rectangular shape, it might be incorrectly used and might be wrong to analyze by drawing a picture. For instance, by using rectangular area, any points $(\bar{X}_1, \bar{X}_2)$ that fall into the area E or F will be concluded as out of control but in fact they are in control within the area G. On the other hand, if the points allocate inside the rectangular area but the process is actually out of control. The degree of correlation between variables have strongly influenced on incorrectly interpreting data. If the average chart of each characteristic was constructed dependently based on the basis of Type I error with the probability α and used it as rectangular control area, the probability of Type I error for combined control protocol will be

$$\alpha' = 1 - (1 - \alpha)^p \tag{2.37}$$

when p is the number of independently combined control variables, the probability of every sample mean p that is plotted within the rectangular area is $(1 - \alpha)^p$.

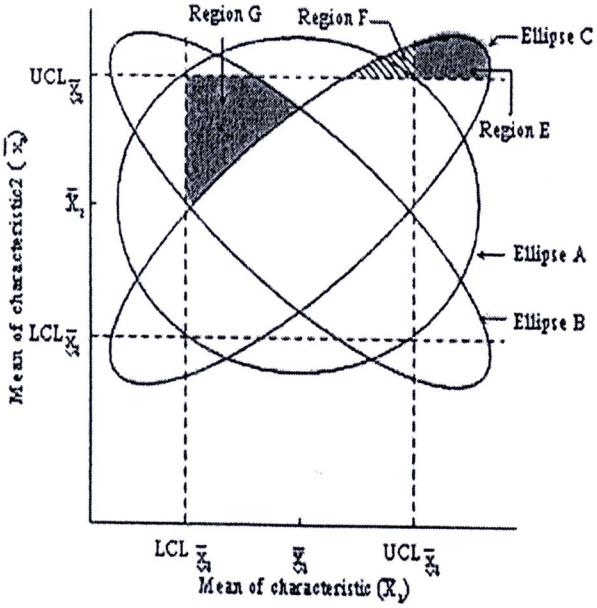

Figure 2.13 Elliptical control region.

The medium or the large values of p will have direct effect on an error of incorrect data interpreting. Giving that the control chart is constructed by using Type I error

with the probability of 0.0026. If there are 4 independent characteristics, the overall probability of Type I error $(\alpha)'$ for control process is

$$\alpha' = 1 - (0.9974)^4 = 0.0104 \tag{2.38}$$

In the case of dependent variables, it is difficult to observe the importance of Type I error. In practical, the control ellipse will be selected. Thus the probability of the sample mean that is plotted inside the elliptical area when the process is in control will be $(1 - \alpha)$, when α is the overall probability of Type I error.

Hotelling's Control Chart and Its Variation

Giving that two characteristics X_1 and X_2 have bivariate normal distribution. Assuming that characteristic means are $\bar{\bar{X}}_1$ and $\bar{\bar{X}}_2$, sample means are $\bar{X}_1$ and $\bar{X}_2$ variances are S_1^2 and S_2^2, and covariance between two variables is S_{12} for sample size n.

$$T^2 = \frac{n}{(S_1^2 S_2^2 - S_{12}^2)} \left[S_2^2(\bar{X}_1 - \bar{\bar{X}}_1)^2 + S_1^2(\bar{X}_2 - \bar{\bar{X}}_2)^2 - 2S_{12}(\bar{X}_1 - \bar{\bar{X}}_1)(\bar{X}_2 - \bar{\bar{X}}_2) \right] \tag{2.39}$$

Under this condition, the statistics has Hotelling's T^2 distribution with degrees of freedom 2 and $n = 1$, where the number 2 came from 2 variables to be considered and $n = 1$ is degree of freedom that relates to sample variance. If T^2 estimated using the equation (2.39) is more than $T^2_{\alpha,2,n-1}$, there will be at least one variable out of control. This step can be demonstrated by a graph, it was found that the equation (2.39) gives elliptical control area (Figure 2.13). if the variables are independent, covariance of these variables will be 0, and the control ellipse will be similar to the ellipse A and the combined control area will be shown by the area of the control ellipse A. Plotting of bivariate means $(\bar{X}_1, \bar{X}_2)$ inside this control ellipse, the condition of statistical control will be established. If two variables have positive correlation, then $S_{12} > 0$ and the control ellipse will be similar to the ellipse C. And if two variables have negative correlation, then $S_{12} < 0$ and the control area will be the same as the ellipse B.

Utilization of Hotelling's T^2 control ellipse step has 2 disadvantages. Firstly, the order of the plot points $(\bar{X}_1, \bar{X}_2)$ will be lost, implying that the running of graph plotting could not be monitored while the chart is still capable of doing other patterns. Secondly, construction of the control ellipse which contains more than two characteristics is a difficult job. To eliminate the 2_{nd} disadvantage, plotting of T^2 value obtained from the equation (2.39) on the control chart has to be based on a sample-by-sample basis, where it requires the time sequence of graph plotting. Subsequently, the control chart will have UCL as $T^2_{\alpha,2,n-1}$, when p is the number of the patterns that are not followed to the assumption, which receives from percentile points of the F-distribution by using the below relation.

$$T^2_{\alpha,2,n-1} = p\left(\frac{n-1}{n-p}\right) F_{\alpha,p,(n-p)} \tag{2.40}$$

When $F_{\alpha,p,(n-p)}$ demonstrates as points on F-distribution, the proportion on the right side is α with p degrees of freedom in the numerator and $(n - p)$ degrees of freedom in the denominator. If more than two characteristics are considered, T^2

value from the equation (2.39) can be written into a general form as below;

$$T^2 = n\left(\bar{X} - \bar{\bar{X}}\right)' S^{-1}\left(\bar{X} - \bar{\bar{X}}\right) \tag{2.41}$$

when $\bar{X}$ is the vector of sample means of characteristic p with sample size n, $\bar{\bar{X}}$ is the vector of target values for every characteristic, and S means the variance-covariance matrix of characteristic p. In practical, $\bar{\bar{X}}$ and S usually are the estimators of sample data when the process is in control. Under this condition, the control limit of the chart T^2 obtained from the equation (2.40) can be modified into (Alt, 1982).

$$UCL = \left(\frac{mnp - mp - np + p}{mn - m - p + 1}\right) F_{\alpha,p,(mn-m-p+1)} \tag{2.42}$$

When m is the number of samples, in which each sample has n size. Using the estimators $\bar{\bar{X}}$ and S, T^2 values of each m sample are calculated using the equation (2.41) before comparing with the UCL from the equation (2.42). If the T^2 value for the j^{th} sample is greater than UCL, it will be out of control point leading into inspection.

Step 1 calculating T_j^2 and the vector of the sample means by moving average, given

$$\bar{X}_j = \begin{bmatrix} \bar{X}_{1j} \\ \bar{X}_{2j} \\ \vdots \\ \bar{X}_{pj} \end{bmatrix}, \qquad j = 1, 2, ..., m \tag{2.43}$$

when $\bar{X}_{ij}$ is the sample means by moving average of the i^{th} characteristic for the j^{th} cycle, and can be estimated from

$$\bar{X}_{ij} = \frac{1}{n}\sum_{k=1}^{n} X_{ijk}, \qquad i = 1, 2, ..., p; j = 1, 2, ..., m \tag{2.44}$$

where x_{ijk} is the k^{th} observed value on the i^{th} characteristic in the j^{th} cycle.

Step 2 sample variance of the i^{th} characteristic in the j^{th} cycle by moving average, given

$$S_{ij}^2 = \frac{1}{n-1}\sum_{k=1}^{n}\left(X_{ijk} - \bar{X}_{ij}\right)^2, \qquad i = 1, 2, ..., p; j = 1, 2, ..., m \tag{2.45}$$

The covariance between characteristic i and characteristic h in the j^{th} cycle is calculated from

$$S_{ihj} = \frac{1}{n-1}\sum_{k=1}^{n}\left(X_{ijk} - \bar{X}_{ij}\right)\left(X_{hjk} - \bar{X}_{hj}\right), \qquad i \neq h; j = 1, 2, ..., m \tag{2.46}$$

The vector $\bar{\bar{X}}$ of target means of each characteristic in sample m is calculated by

$$\bar{\bar{X}}_i = \frac{1}{m}\sum_{j=1}^{m} \bar{X}_{ij}, \qquad i = 1, 2, ..., p \tag{2.47}$$

Step 3 the member of the variance-covariance matrix S in the equation (2.41) will be estimated from the mean of sample m.

$$S_i^2 = \frac{1}{m}\sum_{j=1}^{m} S_{ij}^2, \qquad i = 1, 2, ..., p \quad \text{and} \quad S_{ih} = \frac{1}{m}\sum_{j=1}^{m} S_{ihj}, \qquad i \neq h \quad (2.48)$$

Finally, the vector $\bar{\bar{X}}$ is estimated by using the member $\left(\bar{\bar{X}}_i\right)$, and the matrix S calculated as following;

$$S = \begin{bmatrix} S_1^2 & S_{12} & \dots & S_{1p} \\ & S_2^2 & \dots & S_{2p} \\ & & \ddots & \vdots \\ & & & S_p^2 \end{bmatrix} \quad (2.49)$$

In order to use in the equation (2.41) the matrix is converted into its inverse. When the number of preliminary samples m is large-say, $m > 100$ many practitioners use an approximate control limit, either

$$UCL = \chi_{\alpha,p}^2 \quad (2.50)$$

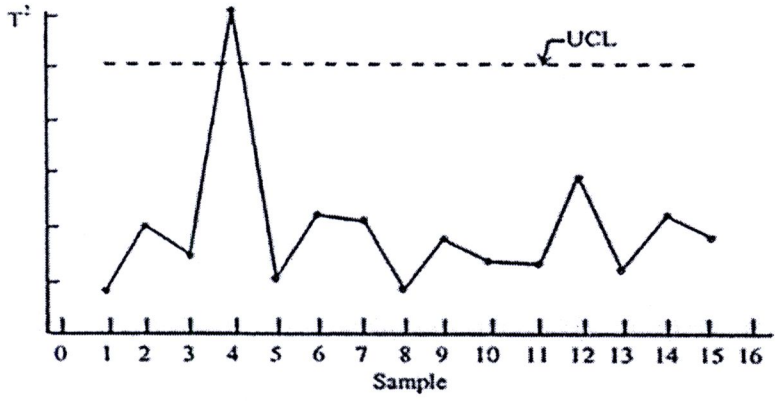

Figure 2.14 A Hotelling's T^2 chart.

Figure 2.14 shows the Hotelling's T^2 control chart constructed by using UCL from the equation (2.42). The plotting of T^2 value of each sample calculated from the equation (2.42) revealed that for 4 samples the T^2 value is higher than UCL which is out of control. The question is how to determine the method for using with 4.sample situation while the quality of products is out of control. Generally, the event containing two characteristics ($p = 2$) is complex. If the qualities of these two characteristics have highly positive correlation and the means of each characteristic are the expected values, it is recommended to maintain the similar relation in the process mean $\bar{\bar{X}}$. For example, in the j^{th} sample if $\bar{X}_{1j} > \bar{\bar{X}}_1$, then $\bar{X}_{2j} > \bar{\bar{X}}_2$ is the expected value. Similarly, if $\bar{X}_{1j} < \bar{\bar{X}}_1$, then $\bar{X}_{2j} < \bar{\bar{X}}_2$ is the expected value confirming that the mean of every characteristic moving in the same direction. In the contrast, if two characteristics have high negative relation and $\bar{X}_{1j} > \bar{\bar{X}}_1$, then

$\bar{X}_{2j} > \bar{\bar{X}}_2$. It means that this sample may have out of control point in the Hotelling's T^2 process. Using the method, it demonstrated that the bivariate process is out of control. This conclusion came from the use of 3σ as control limits of the chart constructed for each characteristic, where $\bar{X}_{1j} > \bar{\bar{X}}_1 + 3\sigma_{\bar{x}_i}$ or $\bar{X}_{2j} > \bar{\bar{X}}_2 + 3\sigma_{\bar{x}_i}$. But this can also occur to individual characteristic quality where the points are plotted inside UCL of the sub-control chart but the relation of the plotting values T^2 are greater than UCL on the combined control chart. Utilization of combined control chart with several characteristics requires investigation at the same time to gain the advantages. However, the individual control chart can detect only in some circumstances with the condition of being out of control when it is not a combined control chart.

Basically, inspection of changes in the process that contains several characteristics with positive relation requires more samples than that of the multi-characteristic process with negative relation. Besides, there are more than one cases of the multi-characteristic process with positive relation reporting for detection of big changes in the process. Normally, if the process inspected by the Hotelling's T^2 control chart is out of control, the individual control intervals will be calculated for individual sample characteristic. Thus if the probability of Type I error for combined control process is α, for the j^{th} sample the individual control intervals for the j^{th} characteristic will be;

$$\bar{\bar{X}}_i \pm t_{\alpha/2p,m(n-1)} S_i \sqrt{\frac{m-1}{mn}}, \qquad i = 1, 2, ..., p \tag{2.51}$$

where $\bar{\bar{X}}_i$ and S_i^2 obtained from the equations (2.47) and (2.48), respectively. If $\bar{X}_{ij}$ falls out of this range, the related characteristics should be inspected for lack of control.

2.1.12 Computing generalized inverses

We review some computational formulas for generalized inverses. The emphasis here is not on the development of formulas best suited for the numerical computation of generalized inverses on a computer. For instance, the most common method of computing the Moore-Penrose inverse of a matrix is through the computation of its singular value decomposition; that is, if $A = P_1 \triangle Q_1'$ is the singular value decomposition of A as give in Corollary 1 (Appendix 1), then A^+ can be easily computed via the formula $A^+ = Q_1 \triangle^{-1} P_1'$. The formulas provided here and in the problems are ones that, in some cases, may be useful for the computation of the generalized inverse of matrices of small size but, in most cases, are primarily useful for theoretical purposes.

Greville (1960) obtained an expression for the Moore-Penrose inverse of a matrix partitioned in the form $\left[\begin{array}{cc} B & c \end{array}\right]$, where, of course, the matrix B and the vector c have the same number of rows. This formula can be then used recursively to compute the Moore-Penrose inverse of an $m \times n$ matrix A. To see this, let a_j denote the j^{th} column of A and define $A_j = (a_1, ..., a_j$, So that A_j is the $m \times j$ matrix containing the first j column of A. Greville has shown that if we write $A_j = \left[\begin{array}{cc} A_{j-1} & a_j \end{array}\right]$, then

$$A_j^+ = \left[\begin{array}{c} A_{j-1}^+ - d_j b_j' \\ b_j' \end{array}\right] \tag{2.52}$$

where $d_j = A_{j-1}^{+} a_j$,

$$b_j' = \begin{cases} (c_j' c_j)^{-1} c_j', & \text{if } c_j \neq 0 \\ (1 + d_j' d_j)^{-1} d_j' A_{j-1}^{+}, & \text{if } c_j = 0 \end{cases} \tag{2.53}$$

and $c_j = a_j - A_{j-1} d_j$. Thus, $A^{+} = A_n^{+}$ can be computed by successively computing $A_2^{+}, A_3^{+}, ..., A_n^{+}$.

2.1.13 Average run length

Average run length is defined as the mean number of the last points that fall within control limits (UCL-LCL range), which is divided into 2 valves; ARL_0 and ARL_1. Both ARL_0 and ARL_1 can be calculated as describing below.
When the process is in control, it is evaluated by the ARL_0 which is estimated by

$$ARL_0 = \frac{1}{\hat{\alpha}} \tag{2.54}$$

where $\hat{\alpha}$ is the estimated value of probability that the data is out of control while the process is in control $[P(out|in)]$.
When the production process begins to change, it is evaluated by the ARL_1 which is estimated by

$$ARL_1 = \frac{1}{1 - \hat{\beta}} \tag{2.55}$$

where $\hat{\beta}$ is the estimated value of probability that the data is in control while the process is out of control $[P(in|out)]$.

2.2 Literature reviews

In this section, we review some of the works in the area of statistical quality control as well as the ranked set sampling and classified them into two separate groups.

2.2.1 Statistical quality control

Although quality control has been with us since when manufacturing began and competition accompanied manufacturing but, its scientific foundation with respect to how many sample units to inspect and what conclusion to draw from the result and the eventual extension to statistical quality control took place relatively late. The beginning of statistical quality control dates back to 1924, when Shewhart (1924) introduced his first control chart for the fractional nonconforming units. His first control chart monitors whether the nonconforming fraction of a product remains the control limits during the time of observation or not.
After over twenty five years from the original work of Shewhart (1924), Aroian and Levene (1950) proposed the first trials to determine the three decision parameters of a control chart namely; sample size, control limit, and time between sampling. With the aim of minimizing the number of product units when the process is out of control, they noted that the frequency of the false alarms which depends on the time interval between samples plays a greater role in the determination of the control limits than the probability of those false alarms per sample.
Weiler (1952), used sample size in constructing a model to minimize the average

amount of inspection before a process shift occurs. In his work, he had completely avoided the time interval between samples and the probability of detecting the effect of process shift. In other words, the average run length when the process is out of control was neglected.
Crowder (1987) presented a numerical procedure for the computation of average run lengths of a control chart using the combination of individual measurement and moving range chart based on two consecutive measurements. He supplied the exact expression for average run length in integral form and its approximation in numerical form. He also gave average run length values for several settings of control limits and shift in the process mean and standard deviation.
Salazar and Sinha (1997) constructed $\bar{X}$ control chart based on ranked set samplimg considering normal population and various shift values. Using visual comparison and Mote Carlo simulation for the computation of average run length, they show that ranked set sampling and median ranked set sampling, based control charts for means were considerably better in detecting a shift in process mean than that of the classical Shewhart $\bar{X}$ control chart with same sample size. In their work, they had considered both the cases where ranking can be and cannot be performed without error in ranking with equal and unequal allocations. In other words, perfect and imperfect ranking were considered.
Muttlak and AL-Sabah (2003) went further beyond the work of Salazar and Sinha (1997) by considering further modifications of ranked set sampling namely; extreme ranked set sampling, paired ranked set sampling and selected ranked set sampling. Using normal population and various shift values, they computed various average run length values with an aid of computer simulation and showed that all the control charts for means based on the above sampling techniques were better than those of classical Shewhart control charts.

2.2.2 Ranked set sampling

The method of ranked set sampling was first proposed by McIntyre (1952) in estimation of mean pasture yield. He noted that ranked set sampling is considerably more efficient in the estimation of a population mean than the standard simple random sampling. Although with no mathematical theory for McIntyre (1952) scheme over the next decade, Halls and Dell (1996) applied it on the estimation of forage yield. A major breakthrough in terms of necessary mathematical theory in support of McIntyre (1952)'s work were given by Takahashi and Wakimoto (1968). Through an independent work, they proved that the sample mean of the ranked set sampling is an unbiased estimator of the population mean with smaller variance as compared to sample mean of simple random sampling with same sample size.
In just about a year after the work of Takahashi and Wakimoto (1968), Takahashi (1969) this time around alone, reconsidered the problem in situation where the elements within each set are correlated. In this work, he proposed a model and an estimator of the population mean. The relative efficiencies of his estimators for some distribution were also computed. Takahashi (1969) went further with the modification of ranked set sampling by considering a situation where elements are randomly selected and measured before their position in a rank is determined.
Where the earlier works were assuming perfect ranking, Dell and Cluster (1972)

studied the case in which ranking may not be perfect. They showed that regardless of the error in ranking, mean of ranked set sampling is an unbiased estimator of the population mean and that the efficiency of the ranked set sampling estimator decrease with increasing ranking errors. Also noted in their work is that even with the error in ranking, the ranked set sampling estimator is still more efficient than that of the simple random sampling using same sample size. In other words

$$\frac{Var(\bar{X}_{srs})}{Var(\bar{X}_{rss})} \geq 1 \tag{2.56}$$

where $\bar{X}_{srs}$ and $\bar{X}_{rss}$ are the estimators of the population mean based on simple random sampling and ranked set sampling respectively. Equality holds in situation where judgment ranking is very poor to produce random sample.
Stokes (1977) studied a situation where the variable of interest X may not easily be measured or ordered but there is a concomitant variable Y which is correlated with the variable of interest X that can readily be ordered. A sampling method based on concomitant variable Y was proposed and observed that the precision of a population mean estimator depends on how strong the relationship between the X's and Y's She note that the mean estimator is equivalent to McIntyre (1952) estimator if the correlation coefficient $\rho = 1$ and equals simple random sampling estimator if $\rho = 0$.
Ridout and Cobby (1987) observed that apart from errors involved in ranking the variable of interest, another source of error due to non-random selection of set can arise in the practical implementation of ranked set sampling. The effects of such error on the relative efficiency of ranked set sampling estimators were studied and with an aid of example were able to show that the relative precision reduces more rapidly with increasing non-randomness in sampling as compared to errors in ranking the variable of interest.
Ridout (2003) selected samples by ranking individual criteria as $f = 1, f = 2$ respectively. Let n_{fg} is the sample number for giving the g^{th} order of the f^{th} criteria and let V_f is variance of sample n_{fg}. For balance sampling $V_f = 0$, giving the new set with the n size. Calculating new values from V_1, V_2 which were the values of each set sample, then choosing the minimum value of $V = V_1, V_2$ and random choosing, if there was more than one.

2.2.3 Multivariate control chart based on ranked set sampling

Pongpullponsak and Sontisamran (2010) proposed a model for ranked set sampling with multiple characteristics

$$\begin{aligned} UCL &= \bar{X}_{rssmc} + 3\hat{\sigma}_{\bar{X}_{rssmc}} \\ CL &= \bar{X}_{rssmc} \\ LCL &= \bar{X}_{rssmc} - 3\hat{\sigma}_{\bar{X}_{rssmc}} \end{aligned} \tag{2.57}$$

where $\hat{\sigma}_{\bar{X}_{rssmc}} \leq \hat{\sigma}_{\bar{X}_{rss}}$ In this work, using multivariate control charts for ranked set sampling with multiple characteristics and using Hotelling's control chart (1947) is a control chart that is used in only two variables in a chart.