CHAPTER 2 REVIEW AND THEORETICAL OF BACKGROUND

2.1 Theoretical Background

2.1.1 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:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \qquad -\infty < x < \infty \qquad (2.1)$$

Standard normal distribution is normal distribution that has $\mu = 0,\ \sigma^2 = 1$ and can be written by using the probability density function as:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \qquad -\infty < x < \infty$$

and the cumulative probability function:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{u^2}{2}\right) du. \qquad (2.2)$$

2.1.2 Exponential Distribution

The probability distribution of the exponential random variable is:

$$f(x) = \lambda e^{-\lambda x} \quad x \ge 0 \qquad (2.3)$$

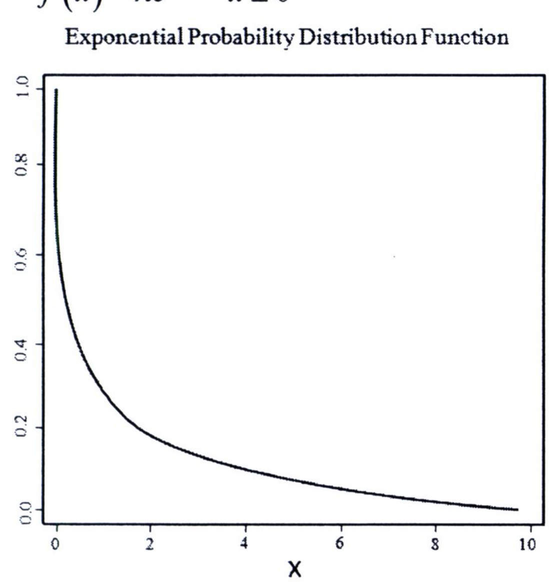


Figure 2.1 Exponential distribution.

where $\lambda > 0$ is a constant. A graph of the exponential distribution is shown in Figure 2.1 The mean and variance of the exponential distribution are:

$$\mu = \frac{1}{\lambda} \tag{2.4}$$

and

$$\sigma^2 = \frac{1}{\lambda^2} \tag{2.5}$$

respectively.

The cumulative exponential distribution is:

$$\begin{aligned} F(a) &= P\{x \le a\} \\ &= \int_0^a \lambda e^{-\lambda t} dt \\ &= 1 - e^{-\lambda a} \qquad a \ge 0 \end{aligned} \tag{2.6}$$

Figure 2.2 depicts the exponential cumulative distribution function.

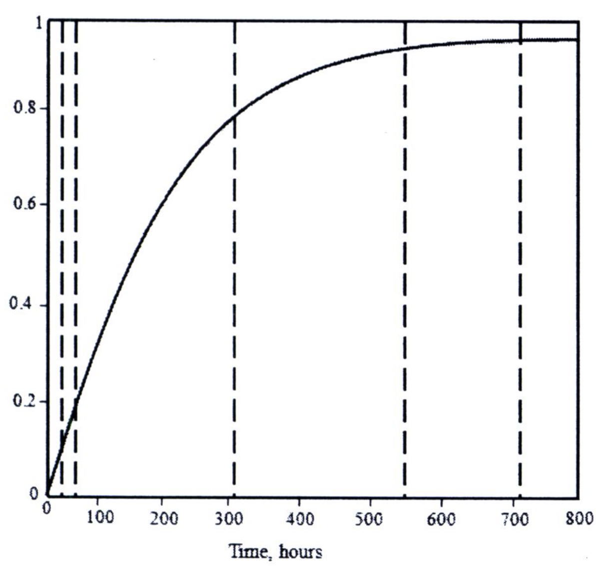


Figure 2.2 The exponential cumulative distribution functions.

In these applications, the parameter λ is called the failure rate of the system, and the mean of the distribution $\frac{1}{\lambda}$ is called the mean time to failure.

2.1.3 Principle of Quality Control Chart

The quality control chart is a graph representing quality measurement within certain time intervals. The quality measurements are computed from sets of samples. Coordinates in the quality control chart may be either sample means or sample ranges. Once all the coordinates are plotted, they will be joined by straight lines to reveal the quality pattern of the process. Three control limits, calculated from random samples, are also included in the chart. They are the Upper Control Limit (UCL) , the Center Limit (CL), and the Lower Control Limit (LCL). If all the points fall between the upper control limit and the lower control limit, the process is said to be in-control; otherwise, it is out-of-control. Figure 2.3 shows the typical control chart. The control limits are chosen and if the plotted points are within the control limits, the process is assumed to be in-control, and no action is required. However, if plotted points are not within the control limits, the process is assumed to be out-of-control, and the process required the investigation and corrective action. In this way the assignable cause or causes responsible for such a behavior are eliminated.

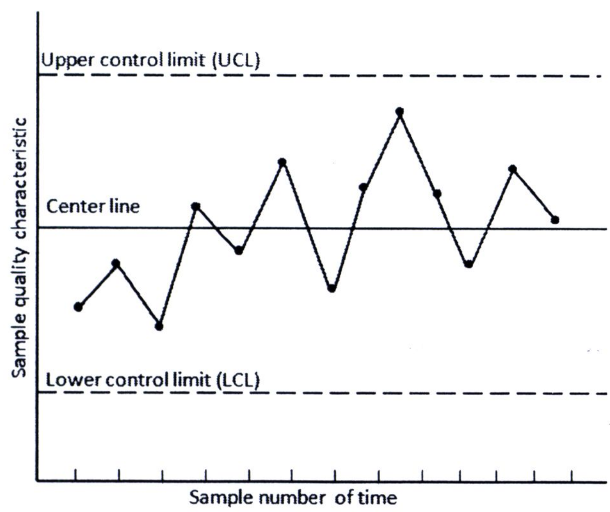


Figure 2.3 Typical control chart

During the procedure of detecting out-of-control states, it may be type I or type II errors. The probability of committing Type II error (β) is the indication that we assumed the process is in-control when it is really not in control. Similarly the probability of committing Type I error (α) is the indication of identifying process as out-of-control where it is really in-control.

A general model for a control chart is represented as follows:

Let $\bar{x}$ be a sample statistics that measures some quality characteristic of interest, and $\mu_{\bar{x}}$ is the mean and $\sigma_{\bar{x}}$ is the standard deviation of $\bar{x}$, then the centerline, the upper control limit, and the lower control limit become:

$$\begin{aligned} UCL &= \mu_{\bar{x}} + L\sigma_{\bar{x}} \\ Center\ Line &= \mu_{\bar{x}} \\ LCL &= \mu_{\bar{x}} - L\sigma_{\bar{x}} \end{aligned} \tag{2.7}$$

where L is the distance of the control limits from the centerline, expressed in standard deviation units. If the system of chance causes a variation in a quality characteristic that follows the normal distribution, then $\pm 3\sigma$ control limits indicates a possible 3 out of 1000 defective items. Whether the quality characteristic is normally distributed it is customary to base the control limits on a multiple is 3 and these outer limits are called 3-sigma limits or action limits. The inner limits, usually at 2σ, are called warning limits. When probability limits are used, the action limits are generally 0.001 limits and the warning limits are 0.025 limits.

If one or more points fall between the warning limits and the control limits, or very close to the warning limit, we should be suspicious that the process may not be operating properly. One possible action to take when this occurs is to increase the sampling frequency and /or the sample size so that more information about the process can be obtained quickly. Process control schemes that change the sample size and/or the sampling frequency depending on the position of the current sample value are call adaptive or variable sampling interval (or variable sample size, etc.) schemes.

The construction of a quality control chart is as follows:

1. Specify the control quality of the product: This is the initial step of making a control chart. The type of quality control chart that would be best matches the product being determined.
2. Limit the amount of data collection: The amount of data required depends on various factors: types of control charts, products, costs, and control methods. Example, using range and standard deviation to measure the distribution of 4-5 samples yields little difference, but the range varies more significantly when the sample size increases to more than 10.
3. Specify the time intervals for data collection: In general, the data might be collected once every 30 minutes, to once every hour. Alternatively, the data collection might be scheduled to be at certain time, e.g. during 10-11 a.m. every day. The frequency of data collection depends on the product manufacturing time and the number of samples being considered. There is no fixed rule. The more often we collect the data, the higher cost we have to pay. But more data allows better analysis of the manufacturing process.
4. Calculate the control bounds: If the deviation of the manufacturing process is average, 3-sigma control bound is typically used, which results in small number of defects. However in the case of more strict control or when the cost of defective products is high, the bound may be changed to 2.5-sigma, 2-sigma (warning), or 1-sigma.
5. Plot the control chart: Data points plotted in the control chart are connected by lines. Abnormality or outliner that represents an out-of-control condition can be observed if at least one of the followings occurs:

- One or more points lies outside the 3-sigma control bounds.

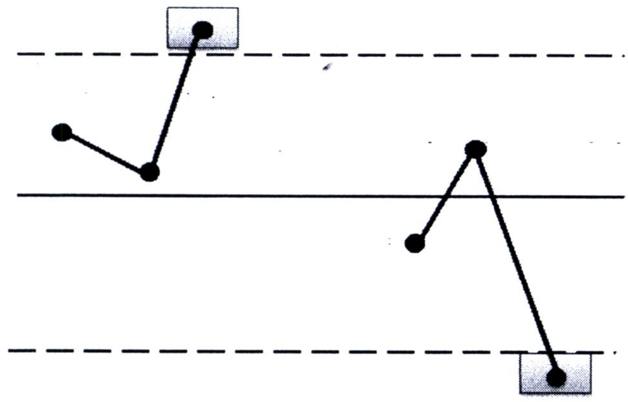


Figure 2.4 One or more points outside the control limits

- Seven or more consecutive points on one side of the centerline

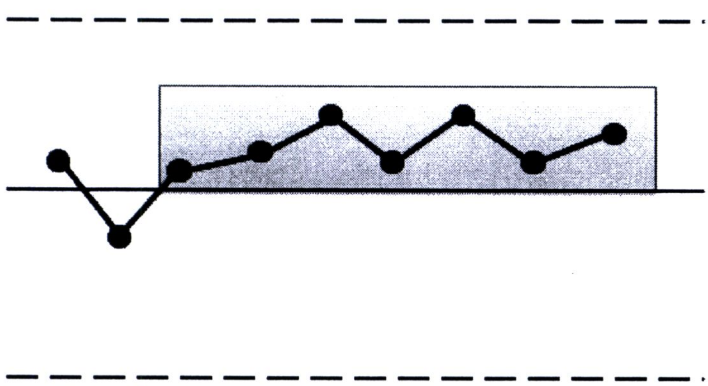


Figure 2.5 Seven or more consecutive points on one side of the centerline

- Six points in a row steadily increasing or decreasing.

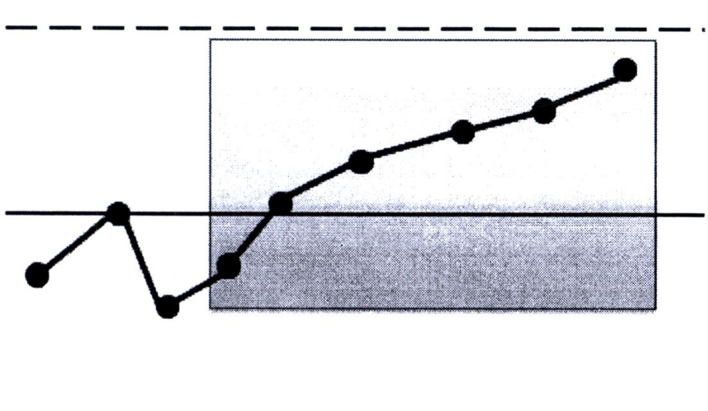


Figure 2.6 Six points in a row steadily increasing or decreasing

- Fourteen points alternating up and down.

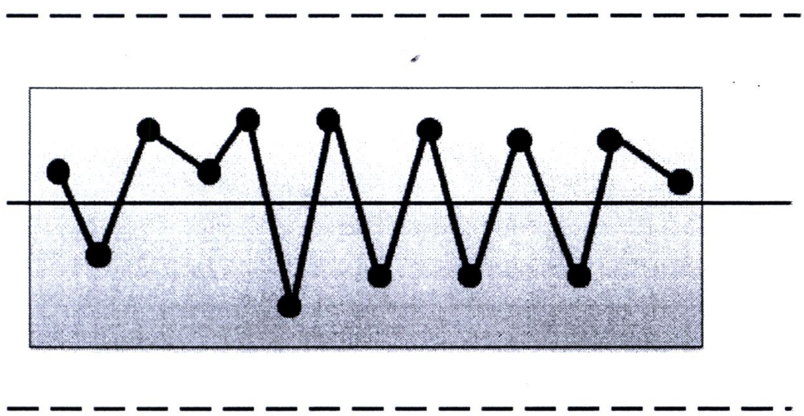


Figure 2.7 Fourteen points alternating up and down

- Two out of three consecutive points in the outer third of the control region.

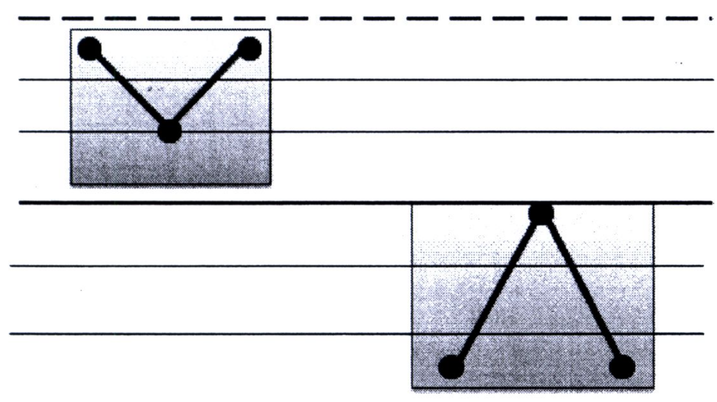


Figure 2.8 Two out of three consecutive points in the outer third of the control region

- Fifteen points in a row within the center third of the control region.

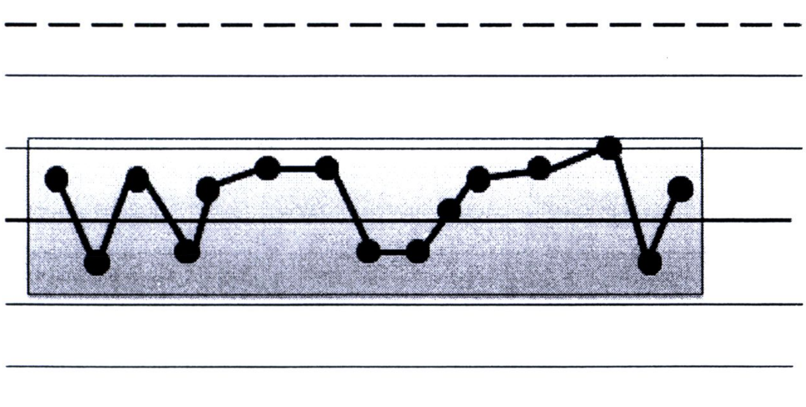


Figure 2.9 Fifteen points in a row within the center third of the control region

- Eight points on both sides of the centerline with none in the center third of the control region.

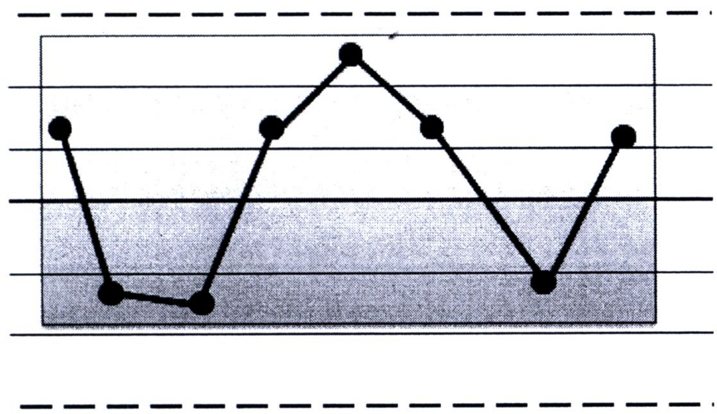


Figure 2.10 Eight points on both sides of the centerline with none in the center third of the control region

- Two out of three consecutive points lie outside 2-sigma warning bounds.
- Four out of five consecutive points lie outside 1-sigma or more from the middle bound

In all of these situations, only one side of middle bound is considered. Thus, a point above the warning sigma followed by a point below the warning sigma does not make an out-of-control phenomenon.

6. Develop the control chart and use it to improve the manufacturing process: when an out-of-control event happens, we need to adjust the control bounds. First, the causes of the out-of-control points are analyzed. Once a cause is identified, the corresponding point is deleted and new control bounds are re-located. A new control chart is then re-constructed using all the other points. (There may be new out-of-control points since the new bounds are thinner) This process is repeated until all the points are inside the control bounds. Then we may use these bounds to control the manufacturing process now and in the near future.

2.1.4 $\overline{X}$ Chart

The observations for the process variable X are assumed to be independent and normally distributed. When the process is in control, the mean and variance of X is μ_0 and $\sigma_0^{\ 2}$ respectively. The lower and upper control limits associated with the Shewhart $\overline{X}$ chart, (Montgomery, 2001) are:

$$\text{LCL}_{\overline{X}} = \mu_0 - L(\frac{\sigma_0}{\sqrt{n}})$$
$$\text{UCL}_{\overline{X}} = \mu_0 + L(\frac{\sigma_0}{\sqrt{n}}) \qquad (2.8)$$

Where L is the control limit parameter and n is the sample size. At any sampling instant t , the sample average $\overline{X}_t$ is compared against these limits, and if it is outside the limit, a search for an assignable cause is started. The rough guidelines for setting the $\overline{X}$ chart limits in practice are $n=4$ or 5, and $L=3$

$$LWL_{\overline{X}} = \overline{\overline{X}} - w\left(\frac{\sigma_0}{\sqrt{n}}\right)$$
$$UWL_{\overline{X}} = \overline{\overline{X}} + w\left(\frac{\sigma_0}{\sqrt{n}}\right) \tag{2.9}$$

where w is the warning limit coefficient of the $\overline{X}$ control chart.

Consider type I error of $\overline{X}$ control chart (See Appendix A.1)

$$\begin{aligned}\alpha &= P(\overline{\overline{X}} < LCL \mid \mu=\mu_0) + P(\overline{\overline{X}} > UCL \mid \mu=\mu_0) \\ &= 2\left[\Phi(-L)\right]\end{aligned} \tag{2.10}$$

where $\Phi(x)$ is the cumulative distribution function of normal distribution
α is Type I error probability

Consider type II error of $\overline{X}$ control chart (See Appendix A.2)

$$\begin{aligned}\beta &= P(\overline{\overline{X}} < UCL \mid \mu=\mu_0+\varepsilon) - P(\overline{\overline{X}} < LCL \mid \mu=\mu_0+\varepsilon) \\ &= \Phi\left(L - \frac{\varepsilon}{\frac{\sigma_0}{\sqrt{n}}}\right) - \Phi\left(-L - \frac{\varepsilon}{\frac{\sigma_0}{\sqrt{n}}}\right)\end{aligned} \tag{2.11}$$

β is Type II error probability

2.1.5 Shewhart's Variable Parameter $\overline{X}$ Control Chart

The average control chart by the Shewhart controls chart are likely to be use extensively, but it is convenient for only the data that has the normal distribution. So, for Shewhart's Variable Parameter $\overline{X}$ Control Chart when $i=1,2$ is given by

$$\text{Upper Control Limit} \quad UCL_i = \mu_0 + k_i \frac{\sigma'_x}{\sqrt{n_i}}$$
$$\text{Upper Warning Limit} \quad UWL_i = \mu_0 + w_i \frac{\sigma'_x}{\sqrt{n_i}} \tag{2.12}$$

$$\text{Lower Warning Limit } LWL_i = \mu_0 - w_i \frac{\sigma'_x}{\sqrt{n_i}}$$

$$\text{Lower Control Limit } LCL_i = \mu_0 - k_i \frac{\sigma'_x}{\sqrt{n_i}}$$

where μ_0 is the parameter of process average.

σ'_x is the parameter of process variation.

The estimator value parameter assessment of the equation (2.12) for each sample size n_i where $i = 1, 2$ can be as follows:

μ_0 have the estimator, which is given by $\bar{\bar{X}}_i = \frac{\sum_{j=1}^{k} \bar{x}_j}{m}$,

σ'_x have the estimator, which is given by $\sigma'_x = \frac{\bar{R}_i}{d_2}$,

d_2 is the constant value, depending on subgroup.

Where $\bar{R}_i = \frac{\sum_{j=1}^{k} R_j}{m}$

A_2 is the constant value, depending on the sample size for each subgroup.
Substituting all estimators in equation (2.12) would yield

$$\begin{aligned} UCL_i &= \bar{\bar{X}}_i + k_i \frac{A_2}{3} \bar{R}_i \\ UWL_i &= \bar{\bar{X}}_i + w_i \frac{A_2}{3} \bar{R}_i \\ LWL_i &= \bar{\bar{X}}_i - w_i \frac{A_2}{3} \bar{R}_i \\ LCL_i &= \bar{\bar{X}}_i - k_i \frac{A_2}{3} \bar{R}_i \end{aligned} \qquad (2.13)$$

2.1.6 Exponentially Weighted Moving-Average (EWMA) Chart

Mitra (1998) found that a geometric moving-average control chart, also known as an exponentially weighted moving-average (EWMA) chart, is based on this premise. One of the advantages of a exponential moving-average chart over $\bar{X}$ chart is that the former is more effective in detecting small changes in process parameters. The exponentially weighted moving-average at time step t is given by

$$Z_t = r\bar{X}_t + (1-r)Z_{t-1} \quad , \quad t = 1, 2, 3, \ldots \qquad (2.14)$$

Where r is a weight constant $(0 < r \le 1)$ and Z_0 is $\bar{\bar{X}}$

By using equation (2.14) repeatedly, we get:

$$
\begin{aligned}
Z_t &= r\overline{X}_t + (1-r)Z_{t-1} + r(1-r)^2 Z_{t-2} \\
&= r\overline{X}_t + (1-r)Z_{t-1} + r(1-r)^2 Z_{t-2} + \cdots + (1-r)^t Z_0
\end{aligned}
\tag{2.15}
$$

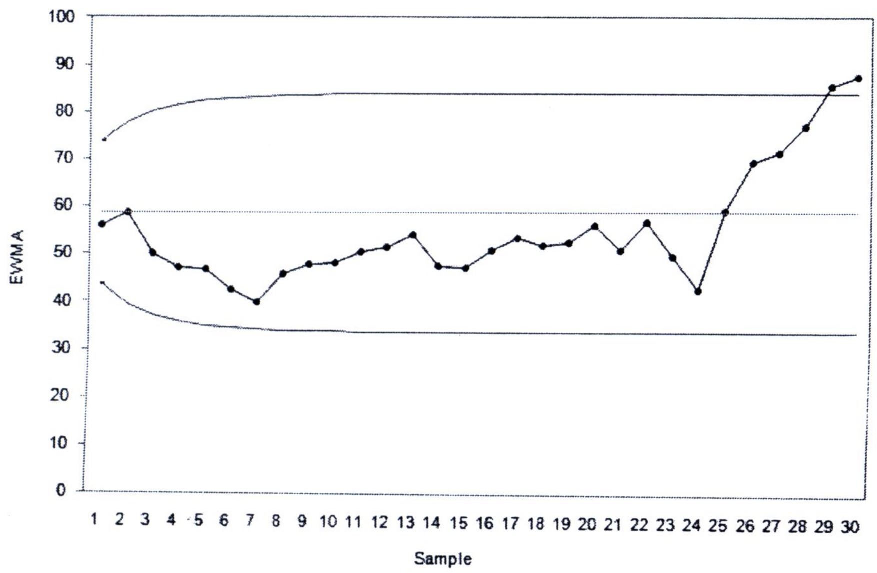


Figure 2.11 Exponentially weighted moving-average (EWMA) charts

The upper and lower control limits (denoted by LCL_{ewma} and UCL_{ewma}, respectively) for the EWMA chart are

$$
\begin{aligned}
UCL &= \overline{\overline{X}} + L\sigma_{\overline{Z}} \\
LCL &= \overline{\overline{X}} - L\sigma_{\overline{Z}}
\end{aligned}
\tag{2.16}
$$

where $\overline{\overline{X}}$ is the in-control value of the process mean and is usually identical to X_0, L is the control limit coefficient of the EWMA chart that determines the size of the critical region of the chart, and $\sigma_{\overline{Z}}$ is the asymptotic standard deviation of the sample statistic equal to:

$$
\sigma_{\overline{Z}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{r}{2-r}}
\tag{2.17}
$$

where σ is the standard deviation of the process characteristic and n is the sample size. The upper and lower warning limits (denoted by UWL and LWL, respectively) for the EWMA chart are:

$$
\begin{aligned}
UWL &= \overline{\overline{X}} + w\sigma_{\overline{Z}} \\
LWL &= \overline{\overline{X}} - w\sigma_{\overline{Z}}
\end{aligned}
\tag{2.18}
$$

where w is the warning limit coefficient of the EWMA chart. If the last sample point falls in the safe region (i.e., $LWL \le \bar{\bar{X}} \le UWL$), then take the next sample at the next fixed sampling time point (h_1). If the last sample point falls in the warning region (i.e., $UWL < \bar{\bar{X}} \le UCL$ or $LCL \le \bar{\bar{X}} < LWL$), then take the next sample using the short sampling time point (h_2). A search for the assignable cause is under taken when the sample point falls outside the control limits.

Consider type I error of EWMA control chart (see Appendix A.3)

$$\begin{aligned} \alpha &= P(\bar{\bar{X}} < LCL \mid \mu = \mu_0) + P(\bar{\bar{X}} > UCL \mid \mu = \mu_0) \\ &= 2\left[\Phi(-L)\right] \end{aligned} \tag{2.19}$$

where $\Phi(x)$ is cumulative distribution function of normal distribution

α is Type I error probability

Consider type II error of EWMA control chart (See Appendix A.4)

$$\begin{aligned} \beta &= P(\bar{\bar{X}} < UCL \mid \mu = \mu_0 + \varepsilon) - P(\bar{\bar{X}} < LCL \mid \mu = \mu_0 + \varepsilon) \\ &= \Phi\left(L - \frac{\varepsilon}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{r}{2-r}}} \right) - \Phi\left(-L - \frac{\varepsilon}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{r}{2-r}}} \right) \end{aligned} \tag{2.20}$$

β is Type II error probability

2.1.7 Multivariate EWMA (MEWMA) Chart

Lowry, et al. (1992) developed a multivariate extension of the univariate EWMA control chart introduced by Roberts (1959). The statistic of the multivariate EWMA control chart is defined as:

$$\vec{Z}_t = R\vec{X}_t + (1-R)\vec{Z}_{t-1} \qquad , t = 1,2,3,\ldots, \tag{2.21}$$

where $Z_0 = E\left(\vec{X}\right)$, $\vec{X} = \left(x_1, x_2, \ldots, x_n\right)$

and $R = \text{diag}(r_1, r_2, \ldots, r_p)$, $0 \le r_j \le 1$, $j = 1,2,\ldots,p$

When $r_1 \ne r_2 \ne \ldots \ne r_p$,

$$\vec{Z}_t = \begin{bmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & r_p \end{bmatrix} \bar{\vec{X}}_t + \begin{bmatrix} 1-r_1 & 0 & \cdots & 0 \\ 0 & 1-r_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1-r_p \end{bmatrix} \vec{Z}_{t-1} \quad , t \geq 1$$

If $r_1 = r_2 = ... = r_q = r$, the statistic can be simplified as:

$$\vec{Z}_t = r\vec{X}_t + (1-r)\vec{Z}_{t-1} \qquad , t \geq 1, \tag{2.22}$$

The chart gives an out-of-control signal when:

$$T_t^2 = \vec{Z}_t' \sum\nolimits_{Z_i}^{-1} \vec{Z}_i > h \tag{2.23}$$

where h is the upper control limit and $\sum_{Z_i}$ is the covariance matrix of Z_t. Note that there is no lower limit and the statistic Q_i is nonnegative since $\sum_{Z_i}$ is semi-positive definite matrix. The control limit h can be chosen by statistical design based on ARL_s requirement or economic design of the control chart. Prabhu and Runger (1997) modified the Markov chain approach to study the average run length performance of the MEWMA control chart.

Table 2.1 Average Run Length (ARL) of MEWMA $(p=2)$
(Prabhu and Runger, 1997)

	r			
d	0.2	0.4	0.6	0.8
0.0	201.00	199.00	200.00	200.00
0.5	35.10	51.90	73.60	95.5
1.0	10.10	13.20	19.30	28.1
1.5	5.50	5.74	7.24	10.30
2.0	3.80	3.54	3.86	4.75
2.5	2.91	2.55	2.53	2.75
3.0	2.42	2.04	1.88	1.91
	h			
	9.65	10.29	10.53	10.58

2.1.8 Expected Value

The expectation or mean of the random variable X, denoted by $E[X]$, Ross (1996) is defined by:

$$E[X] = \int_{-\infty}^{\infty} x dF(x) \tag{2.24}$$

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx \qquad \text{if } X \text{ is continuous}$$

$$= \int_{-\infty}^{\infty} x P\{X = x\} dx \qquad \text{if } X \text{ is discrete}$$

Provided the above integral exists.

2.1.9 Conditional Expectation

If X and Y are discrete random variables, the conditional probability mass function of X, given $Y = y$, is defined, for all y such that $P\{Y = y\} > 0$, Ross (1996) is defined by:

$$P\{X = x | Y = y\} = \frac{P\{X = x, Y = y\}}{P\{Y = y\}}. \tag{2.25}$$

The conditional distribution function of X given $Y = y$ is defined by:

$$F(x|y) = P\{X \leq x | Y = y\} \tag{2.26}$$

and the conditional expectation of X given $Y = y$, by:

$$E[X|Y = y] = \int x dF(x|y) = \int x dP\{X = x|Y = y\}. \tag{2.27}$$

If X and Y have a joint probability density function $f(x, y)$, the conditional probability density function of X, given $Y = y$, is defined for all y such that $f_y(y) > 0$ by:

$$f(x|y) = \frac{f(x, y)}{f_Y(y)}, \tag{2.28}$$

and the conditional probability distribution function of X, given $Y = y$, by:

$$F(x|y) = P\{X \leq x | Y = y\} = \int_{-\infty}^{x} f(x|y) dx. \tag{2.29}$$

The conditional expectation of X, given $Y = y$, is defined, in this case, by:

$$E[X|Y=y] = \int_{-\infty}^{\infty} xf(x|y)dx. \tag{2.30}$$

Thus all definitions are exactly as in the unconditional case, expected that all probabilities are now conditional on the event that $Y = y$.

Let us denote by $E[X|Y]$ that function of the random variable Y whose value at $Y = y$ is $E[X|Y = y]$. An extremely useful property of conditional expectation is that for all random variables X and Y.

$$E[X] = E[E[X|Y]] = \int E[X|Y = y]dF_Y(y) \tag{2.31}$$

when the expectation exist.

If Y is a discrete random variable, then equation (2.31) states:

$$E[X] = \sum E[X|Y=y]P\{Y=y\} \ , \tag{2.32}$$

while, if Y is continuous with density $f(y)$, then equation (2.31) says:

$$E[X] = \int_{-\infty}^{\infty} E[X|Y=y]f(y)dy \tag{2.33}$$

2.1.10 Mean and Variance

Assume that there are k characteristics to be measured in a given process and denote the measurements by $\underset{\sim}{X} = \left(x_1, x_2, ..., x_p\right)', \underset{\sim}{X} \sim N_p\left(\mu, \sum\right)$ which is a multivariate normal with mean vector $\underset{\sim}{\mu}$ and covariance $\sum$.

S is sample standard deviation

$$S = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(X_i - \bar{X}\right)', \tag{2.34}$$

D_x is the shift of $\underset{\sim}{\bar{X}}$

$$D_{\bar{x}} = \left(\underset{\sim}{\bar{X}} - \mu_0\right)' \sum\nolimits_0^{-1}\left(\underset{\sim}{\bar{X}} - \mu_0\right), \tag{2.35}$$

D_s is the variability of $\underset{\sim}{\bar{X}}$

$$D_s = \frac{1}{n}\sum_{i=1}^{n}\left[\left(\underset{\sim}{X}_i - \underset{\sim}{\bar{X}}\right)' \sum\nolimits_0^{-1}\left(\underset{\sim}{X}_i - \underset{\sim}{\bar{X}}\right)\right]. \tag{2.36}$$

2.1.11 Absorbing State; Absorbing Markov Chains (Ross, 1996)

Let $\{X_n, n=0,1,2,...\}$ is a stochastic process or Markov chain.
If $X_n = i$, then the process is said to be in state i at time n.
Suppose that whenever the process is in state i, there is a fixed probability P_{ij}
That it will next be in state j.
Suppose that,

$$P\{X_{n+1} = j | X_n = i, X_{n-1} = i-1, ..., X_1 = i_1, X_0 = i_0\} = P_{ij} \quad (2.37)$$

For all states $i_0, i_1, ..., i_{n-1}, i, j$ and all $n \geq 0$. Such a stochastic process is known as a Markov chain. Equation (2.37) may be interpreted as stating that, for a Markov chain, the conditional distribution of any future state X_{n+1}, given the past states $X_0, X_1, ..., X_{n-1}$, and the present state X_n, is independent of the past states and depends only on the present state. The value P_{ij} represents the probability that the process will, when in state i, next make a transition into state j.

$$P_{ij} \geq 0, \quad i, j \geq 0; \qquad \sum_{j=0}^{\infty} P_{ij} = 1, \quad i = 0,1,...$$

Let P denote the matrix of one-step transition probabilities P_{ij}, so that

$$P = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ \vdots & & & \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \vdots & \vdots & & \vdots \end{bmatrix}.$$

For any state i and j define f_{ij}^n to be the probability that, starting in i, the first transition into j occurs at time n. Formally,

$$f_{ij}^n = 0$$
$$f_{ij}^n = P\{X_n = j, X_k \neq j, k = 1, ..., n-1 | X_0 = i\}.$$

Let
$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^n.$$

Then f_{ij} denotes the probability of ever making a transition into state j, given that the process starts in i. (Note that for $i \neq j$, f_{ij} is positive if, and only if, j is accessible from i.) State j is said to be recurrent if $f_{ij} = 1$, and transient otherwise (State j is recurrent if, and only if, $\sum_{n=1}^{\infty} p_{ij}^n = \infty$).

Consider now a finite state Markov Chain and suppose that the states are numbered so that $T = \{1,2,...,t\}$ denotes the set of transient states. Let

$$Q = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1t} \\ P_{i1} & P_{i2} & \cdots & P_{it} \\ P_{t1} & P_{t2} & \cdots & P_{tt} \end{bmatrix}$$

Q is specifies only the transition probabilities from transient states into transient states, some of its row sums are less than 1 (for otherwise, T would be a closed class of states).

For transient states i and j, let m_{ij} denote the expected total number of time periods spent in state j given that the chain starts in state i. Conditioning on the initial transition yields:

$$\begin{aligned} m_{ij} &= \delta(i,j) + \sum_{k} P_{ik} m_{kj} \\ &= \delta(i,j) + \sum_{k=1}^{t} P_{ik} m_{kj} \end{aligned} \tag{2.38}$$

where $$\delta(i,j) = \begin{cases} 1 \ ; & i = j \\ 0 \ ; & otherwise \end{cases},$$

and where the final equality follows from the fact that $m_{kj} = 0$ when k is recurrent state Let M denote the matrix of values $m_{kj}, i, j = 1,...,t,$ that is,

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1t} \\ m_{i1} & m_{i2} & \cdots & m_{it} \\ m_{t1} & m_{t2} & \cdots & m_{tt} \end{bmatrix}$$

In matrix notation, equation (2.38) can be written as

$$M = I + QM$$

where I is the identity matrix of size t. As the preceding equation is equivalent to

$$(I - Q)M = I$$

We obtain, upon multiplying both sides by $(I + Q)^{-1}$, that

$$\begin{aligned} (I-Q)^{-1}(I-Q)M &= (I-Q)^{-1} I \\ M &= (I-Q)^{-1} \end{aligned} \tag{2.39}$$

That is, the quantities m_{ij}, $i \in T, j \in T$, can be obtained by inverting the matrix $(I-Q)$. (The existence of the inverse is easily established.)

For $i \in T, j \in T$, the quantity f_{ij}, equal to the probability of ever making a transition into state j given that the chain starts i, is easily determined from M.

$$\begin{aligned} m_{ij} &= E\left[\text{number of transitions into state } j \mid \text{start in i}\right] \\ &= m_{jj} f_{ij} \end{aligned}$$

where m_{ij} is the expected number of time periods spent in state j given that it is eventually entered from state i. Thus, we get:

$$f_{ij} = \frac{m_{ij}}{m_{jj}}$$

2.1.12 Control Chart Interpretation

The objective of performing multivariate SPC is to monitor process performance over time in order to detect any unusual events. It is essential to be able to track the cause of an out-of-control signal to maintain acceptable levels of quality and to allow for process improvements. However, the complexity of multivariate control charts and cross-correlation among variables makes it difficult to analyze assignable causes leading to the out-of-control signals. Several techniques have been developed that assist in the interpretation of out-of-control signals. Following the same sensitivity of the Shewhart $\overline{X}$ control chart, the Hotelling T^2 is more efficient in detecting larger process shifts. Mason and Young (1999) introduced a modification procedure for the T^2 control charts in order to enhance sensitivity toward detecting a small process shift.

A T^2 control chart is used primarily to monitor the mean vector of quality characteristics of a process. There are two versions of the T^2 chart, one for sub grouped data and the other for individual observations. They can be used not only in achieving a state of statistical control (Phase I) but also in maintaining control over the process (Phase II).

In some cases, the multivariate data can be grouped into rational subgroups, relying on properties of the production process that creates homogeneity within subgroups. When rational subgroups are present, a shift in the mean vector is presumed to be more likely to take place between subgroups (variability in the process over time) than within a subgroup (instantaneous process variability at a given time). This can be used to advantage by forming the sample covariance matrix for each subgroup, then averaging them to get an estimate of the process covariance matrix. The mean vectors for each subgroup can be examined for a shift, thus detecting assignable causes for the shift in the mean vector (Sullivan and Woodall, 1996).

Mason and Young (2001) studied the effectiveness of using the T^2 control charts for batch (sub grouped) processes. His study recommended that when the batch data are collected from the same multivariate normal distribution, T^2 statistic is recommended for detecting out-of-control signals. When the batch data are collected from multivariate normal distributions with different mean vectors, the translation of the different batches to a common origin again allows the usage of T^2 statistic to identify out-of-control signals. Translation to a common origin involves the subtraction of individual batch mean vectors from the corresponding batch observations.

However, sometimes the rational subgroup size is one, that is, data are structured only as individual observations, and process characteristics do not necessarily produce homogeneous subgroups of large size. In the case of individual observations, Sullivan and Woodall (1996) recommended using the sample mean vector and covariance matrix if any value of the T^2 statistic exceeds an upper control limit resulting in an out-of-control signal generated. In some industrial situations, such as chemical and process industries, it is either impractical or difficult to obtain a subgroup size of more than one unit, since these industries frequently have multiple quality characteristics that must be monitored. Therefore, the T^2 control chart with $n=1$ would be appropriate to use.

Mason, et al. (1997) presented a multivariate profile chart by superimposing an $\bar{X}$ chart of univariate statistics on top of the T^2 chart. By performing discrimination analysis, this allows the distinguishing of in-control conditions from out-of-control conditions to determine where assignable causes of variation are occurring. This analysis works by partitioning the multivariate control chart based on the contribution of each variable.

There are also graphical solutions to interpretation difficulty. Lowry and Montgomery (1995) proposed poly plots and multivariate control webs to superimpose univariate statistics on multivariate statistics in order for the user to test trends in individual statistics and realize how they affect other variables.

Jackson (1956) suggested that the multivariate control region be displayed as an ellipse for two variables $(p=2)$. However, when Jackson's control ellipse is used, the time sequence of the plotted points is lost. The results obtained from Jackson's control ellipse are exactly the same as those obtained from using the T^2 control chart. If an observation is outside the ellipse, it will also be above the control limit specified on the T^2 control chart. On the other hand, if an observation is inside the control ellipse, it will be below the control limit specified on the T^2 control chart. However, if an observation is exactly on the parameter of the ellipse, it will be exactly on the control limit line of the T^2 control chart. The results obtained by both methods are identical. Nevertheless, the T^2 control chart retains the time scale and summarizes the process condition by one value, while use of the control ellipse indicates pictorially the nature of the out-of-control conditions.

Figure 2.12 presents the control region for two variables with different levels of correlations. Here, it can be seen that when $(r=+0.8)$, the control ellipse is tilted to the right from the horizontal axis; on the other hand, when $(r=-0.8)$, the ellipse becomes tilted to the left from the horizontal axis. However, when $(r=0)$, the ellipse becomes a circle.

Jackson (1959) considered the case of investigating two or more $(p-1)$ related variables to analyze a multivariate process. The basic concept of the technique is to break up the T^2 statistic into a sum of its principal components, the linear portions of the original variables. Principal Component Analysis (PCA) is a reliable technique to interpret out-of-control signals, whereby components can be examined to understand why the process is out-of-control. This could be accomplished by expressing the T^2 statistic as the normalized principal component of the multi normal variables. Hence, when an out-of-control signal is received, components with abnormally high values are detected. Plots of these variables can be used to determine exactly what occurred in the original sets of data that contributed to the signal in the multivariate set of T^2 statistics (Mason, et al., 1997).

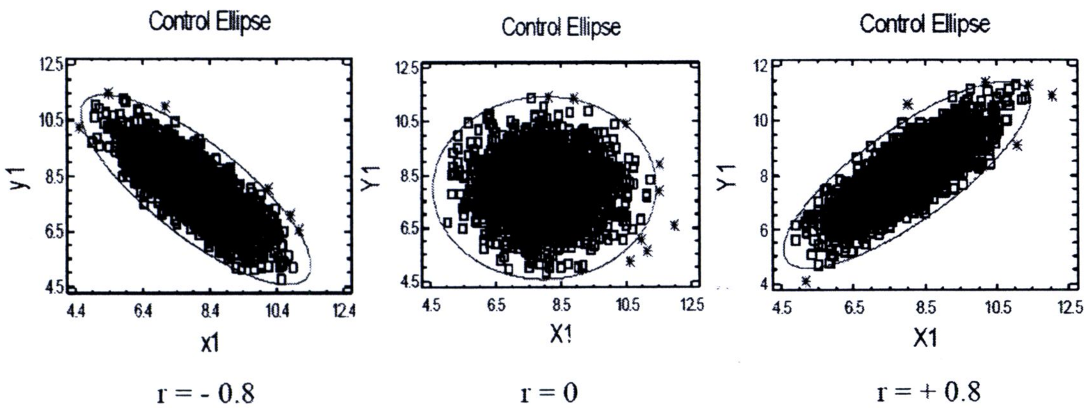


Figure 2.12 Ellipse control region (Source: Montgomery, 2001)

2.1.13 Average Run Length (*ARL*)

Mitra (1998) found that the average run length (ARL) is a measure of the expected number of consecutive samples taken until the sample statistic falls outside the control limits, and it is a function of the current process characteristics. To reduce the total cost, the ARL should be large when the process is in control, and it should be small when the process is out of control. The in-control ARL can be increased by widening the interval between the upper and lower control limits, but this would also cause the out-of-control ARL to increase, unless the sample size is increased as a counter-measure.

It can be shown that the in-control ARL for the Shewhart $\overline{X}$ chart is:

ARL_0 : The average run length during in-control period:

$$ARL_0 = \frac{1}{[2\Phi(-L)]} \tag{2.40}$$

and ARL_1 : The average run length during out-of-control period:

$$ARL_1 = \frac{1}{1 - \Phi(L - \frac{\varepsilon}{\sigma_{\bar{X}}}) + \Phi(-L - \frac{\varepsilon}{\sigma_{\bar{X}}})} \tag{2.41}$$

where $\Phi(\cdot)$ is the cumulative probability distribution function (cdf) for a standard normal variable.

ARL for the EWMA Control chart is:

ARL_0 : The average run length during in-control period

$$ARL_0 = \frac{1}{2[\Phi(-L)]} \tag{2.42}$$

and ARL_1 : The average run length during out-of-control period:

$$ARL_1 = \frac{1}{1 - \Phi(L - \frac{\varepsilon}{\sigma_{\bar{X}}\sqrt{\frac{r}{1-r}}}) + \Phi(-L - \frac{\varepsilon}{\sigma_{\bar{X}}\sqrt{\frac{r}{1-r}}})}. \tag{2.43}$$

2.1.14 Economic Model of Production Process Model (Magalhães and Epprecht, 2001: 191-200)

Existing expenditure function is in every single production hour. It bases on selection of optimum economic model value for parameters $n_1, n_2, h_1, h_2, w_1, w_2, k_1, k_2$ when developing expenditure function. The production process assumptions, utilized in expenditure function development, are as following:

Process Model

The following assumptions are the in-process product characteristic assumptions to be analyzed. The samples are assumed to be independent from each other and the initial production process will be under statistical control in which the $\bar{X}$ control chart equals to $\bar{\bar{X}}$ and standard deviation equals to σ'_x. Once a warning cause or nonconformity is existed, the mean value will shift from μ_0 to $\mu_0 + \delta\sigma'_x$ or $\mu_0 - \delta\sigma'_x$. While the process is still under control, the population is exponentially distributed with the mean value of $\frac{1}{\lambda}$ and not self reversible if any process change is existed. During process investigation, the probability of process continuation ability is an index variable δ_1 ($\delta_1 = 1$ if process is able to continue; $\delta_1 = 0$ if otherwise). The probability of process continuation ability during process repair or improvement is an index variable δ_2 ($\delta_1 = 2$ if process is able to continue; $\delta_2 = 0$ if otherwise). The μ, σ' and δ are assumed to be known in order to define parameters $n_1, n_2, h_1, h_2, w_1, w_2, k_1$ and k_2 of control chart.

The five production process expenditures caused by implementing economic model are as the following:

1. Expenditure caused by population control and sampling (C_{sam})
2. Expenditure caused by inspecting failure warning signal (C_{fa})
3. Expenditure caused by investigating for identifiable cause of nonconformity (C_r)
4. Expenditure caused by producing goods that is not conform to specifications while process is under control (C_{in})
5. Expenditure caused by producing goods that is not conform to specifications while process is not under control (C_{out})

Production Cycle

The production cycle is defined as production duration. Controlling of the production process is assumed to be constant at the beginning. Production cycle composes of two time periods which are: Under Control Period and Not Under Control Period. The details are described below:

1. Time period where production process is still under control (T_{in}): The time duration started from the beginning to the point where the warning cause is obviously identifiable.
2. Time period where production process is not under control (T_{out}): The time duration started from when the process starts changing until the failure warning is developed.
3. Analyzing period (T_a): Time period contributed to sample analysis and control chart result analysis.
4. Inspecting period (T_{ass}): Time period contributed to investigation of identifiable cause, once the production process is not under control.
5. Repairing period (T_r): Time period contributed to process repairing.

The Burdened Expenditure per One Production Cycle

The expenditure function is economically considered as per time unit expenditure function. $E(T)$ is expected value of production period duration and $E(C)$ is expected value of total expenses burdened in one production cycle. Hence, the expected value of total expenses per one time unit is:

$$ECTU = \frac{E(C)}{E(T)} \qquad (2.44)$$

The expected value of total expenses per one production cycle composes of the summation of all existed expenses while production process is both under and out of control. Hence, $E(C)$ composes of:

1. The expected value of expenditure per one production cycle due to the production of goods that is not conforms to specification while the production process is under control $E(C_{in})$ and out of control $E(C_{out})$. Hence,

$$E(C_{in})+E(C_{out})=\left(\frac{1}{\lambda}\right)C_0+C_1\left[AATS+E(T_a)+\delta_1 T_*+\delta_2 T_{**}\right] \quad (2.45)$$

Given C_0 and C_1 are hourly expenditure due to the production of goods that is not conform to specification while the production process is under control and out of control respectively. The mean value of time period while production process is under control is $\frac{1}{\lambda}$. The Adjusted Average Time to Signal $(AATS)$ is the expected value of time period since the production process starts changing until the failure warning signal $E(T_a)$ equals to $n'G$, where G is sampling time interval specified by control chart, n' is sample size while process is out of control, T_* is average time interval where the warning cause is detected and T_{**} is average time for process repairing.

2. The expected value of failure warning signal detection $E(C_{fa})$

$$E(C_{fa})=Y\,E(F) \quad (2.46)$$

where Y is expense caused by failure warning signal detection.
$E(F)$ is average number of independent failure warning signal.

Hence, expected number of failure warning signal $(E(F))$ is

$$E(F)=\left[\alpha_1 p_0+\alpha_2(1-p_0)\right]s \quad (2.47)$$

Where

$$\alpha_i=P(\bar{X}_i<LCL_i)+P(\bar{X}_i>UCL_i) \quad (2.48)$$
$$\alpha_i=P(Z<-k_i)+P(Z>k_i)$$
$$\alpha_i=2\Phi(-k_i)$$

where α_i is type I error
p_0 is the probability of small-size samples while the process is in control
$1-p_0$ is the probability of large-size samples while the process is out of control
$\Phi(\cdot)$ is the cumulative normal distribution.

From equation (2.49), the probability determination of p_0, which is a conditional probability could be as the following:

$$p_0 = P\left(LWL_i < \bar{X}_i < UWL_i \middle| LCL_i < \bar{X}_i < UCL_i\right) \tag{2.49}$$

$$p_0 = \frac{P\left(\mu_0 - w_i\frac{\sigma'}{\sqrt{n_i}} < \bar{X}_i < \mu_0 + w_i\frac{\sigma'}{\sqrt{n_i}}\right)}{P\left(\mu_0 - k_i\frac{\sigma'}{\sqrt{n_i}} < \bar{X}_i < \mu_0 + k_i\frac{\sigma'}{\sqrt{n_i}}\right)}$$
$$= \frac{P\left(-w_i < Z < w_i\right)}{P\left(-k_i < Z < k_i\right)}$$
$$= \frac{\Phi(w_i) - \Phi(-w_i)}{\Phi(k_i) - \Phi(-k_i)} \tag{2.50}$$

3. The expected value of expenses contributed to investing and repairing the cause of warning signal $E(C_r)$ is constant w

$$E(C_r) = w \tag{2.51}$$

4. The expected value of expenses contributed to sampling and controlling $E(C_{sam})$

$$E(C_{cam}) = (a + bn)s + (a + bn')s' \tag{2.52}$$

where a is the fixed expense (direct) per one sample.

b is the variable expense (indirect) per one sample.

n is the average sample size while process is in control.

n' is the average sample size while process is out of control.

s is the average number of sample specified by the control chart while process is in control.

s' is the average number of sample specified by the control chart while process is out of control.

N is a number of sample before process start changing.

Given $s = E(N)$ as average sampling point while process is under control.
Hence, warning cause existed between samples j and $j+1$ means the process average drifted from μ_0 to $\delta\sigma'$. When j is utilized prior to process change, which means $N = j$ similar to the existence of warning cause which is identifiable during the sampling interval t_j and t_{j+1}.
Hence:

$$E(N) = \sum_{j=0}^{\infty} jP(N = j) = \sum_{j=0}^{\infty} jP\left(t_j < T < t_{j+1}\right) \tag{2.53}$$

In this case, T is exponentially distributed with parameter λ. From the exponential distribution characteristics, we obtained:

$$E(N)=s=\frac{e^{-\lambda h_1}p_0+e^{-\lambda h_2}(1-p_0)}{1-e^{-\lambda h_1}p_0-e^{-\lambda h_2}(1-p_0)}$$

Hence, the expected value of total expenditure per one production cycle is obtained by integrating all equations from (2.45) to (2.53) which are:

$$E(C)=\frac{1}{\lambda}C_0+C_1\left[AATS+E(T_a)+\delta_1 T_*+\delta_2 T_{**}\right]+Y\,E(F)+w+(a+bn)s+(a+bn')s' \tag{2.54}$$

The means of each period $E(T)$ compose of:

1. The mean of the expected time that the process is in under control $E(T_{in})$ is

$$E(T_{in})=\frac{1}{\lambda}+(1-\delta_1)E(T_{fa}) \tag{2.55}$$

where $(1-\delta_1)E(T_{fa})$ is part of $E(T_{in})$

$\delta_1=0$, when the process is continuing during the warning

$\delta_1=1$, when the process cease during the warning

2. The mean of the expected time that the process is out of the action region $E(T_{out})$; where

$$E(T_{out})=AATS=E(R)+E(S) \tag{2.56}$$

$AATS$ is adjusted average time to signal.

If assigned t as the mean of expected time that starts since the cause of alarm occurs until the warning, with location between samples j and $j+1$

$$t=\frac{\int_{jh}^{(j+1)h}\lambda(t-jh_1)e^{-\lambda h}dt}{\int_{jh}^{(j+1)h}\lambda e^{-\lambda h}dt}=\frac{1-e^{-\lambda h}(1+\lambda h)}{\lambda(1-e^{-\lambda h})}. \tag{2.57}$$

The mean of the expected time since the cause of alarm occurs to the first sample to the new shift is produced $E(R)$ will be:

$$E(R)=E(E(R|A))\quad\text{and}\quad E(E(R|A))=E(E((h_i-t_i)|A)),$$

$$E(R)=\sum_{i=1}^{2}(h_i-t_i)\cdot P(A=h_i),$$

$$E(R)=\left\{h_1-\frac{1-e^{-\lambda h_1}(1+\lambda h_1)}{\lambda(1-e^{-\lambda h_1})}\right\}P(A=h_1)+\left\{h_2-\frac{1-e^{-\lambda h_2}(1+\lambda h_2)}{\lambda(1-e^{-\lambda h_2})}\right\}P(A=h_2). \quad (2.58)$$

As Reynolds (1988) hypothesized that $P(A=h_i)$ is proportionate to the length of the expected time A, thus, the chance of incidence will be:

$$P(A=h_1)=\frac{p_0h_1}{p_0h_1+(1-p_0)h_2}, \quad (2.59)$$

$$P(A=h_2)=\frac{(1-p_0)h_2}{p_0h_1+(1-p_0)h_2}. \quad (2.60)$$

$E(S)$ is the expected time between the first sample after the shift and the next warning, which $E(S)$ value depends on the position of B

$$E(S)=E(S|B=B_1)P(B=B_1)+E(S|B=B_2)P(B=B_2),$$
$$E(S)=E(T_1)P(B=B_1)+E(T_2)P(B=B_2). \quad (2.61)$$

When $$E(S|B=B_1)=E(T_1), \quad (2.62)$$

$$E(S|B=B_2)=E(T_2). \quad (2.63)$$

T_1 is the expected time since the first sample falls into the central region after the shift. T_2 is the expected time since the first sample falls into the warning region after the shift.

When B is the position of the first sample falling after the shift; if it falls in the central region, the symbol B_1 will be used; if it falls in the warning region, the symbol B_2 will be used; and if it falls in the action region (out of control), the symbol B_3 will be used. B value will depend on the length of the expected time h_i when shift occurs. Thus the probability of B_i will be:

$$\begin{aligned}P(B=B_1)&=P(B=B_1|A=h_1)P(A=h_1)+P(B=B_1|A=h_2)P(A=h_2)\\&=p_{11}P(A=h_1)+p_{21}P(A=h_2),\end{aligned} \quad (2.64)$$

$$\begin{aligned}P(B=B_2)&=P(B=B_2|A=h_1)P(A=h_1)+P(B=B_2|A=h_2)P(A=h_2)\\&=p_{12}P(A=h_1)+p_{22}P(A=h_2),\end{aligned} \quad (2.65)$$

and $$P(B=B_3)=1-P(B=B_1)-P(B=B_2), \tag{2.66}$$

when $$p_{i1}=P(B=B_1|A=h_i)=P\left(|U|<w_i \middle| U\sim N\left(\delta\sqrt{n_i},1\right)\right), \tag{2.67}$$

$$p_{i2}=P(B=B_2|A=h_i) \tag{2.68}$$

$$p_{i2}=P\left(\left(LCL_i<U<LWL_i\right)\cup\left(UWL_i<U<UCL_i\right)\middle|U\sim N\left(\delta\sqrt{n_i},1\right)\right),\ i=1,2$$

$$E(T_1)=E(M_1)E(V). \tag{2.69}$$

M_1 is a variable parameter of sample number that falls in the central region until warning, where M_1 has geometric distribution with the parameter $(1-p_1)$, when p_1 is the probability of sample falling in the central region, so:

$$p_1=p_{11}+p_{12}\sum_{i=1}^{\infty}p_{22}^{i-1}p_{21} \tag{2.70}$$

and $$\sum_{i=1}^{\infty}p_{22}^{i-1}p_{21}=\frac{p_{21}}{1-p_{22}} \tag{2.71}$$

when substitute $p_{ij}'s$ in a double.

$$E(M_1)=\frac{1}{1-p_1},$$

when substitute p_1 in a double ,so we will get

$$E(M_1)=\frac{1-p_{22}}{1-p_{22}-p_{11}+p_{11}p_{22}-p_{12}p_{21}}, \tag{2.72}$$

V is the length of the expected time during sample falling out of warning region, starting from the last sample in the central region, which the probability will be:

$$P(V=h_1)=p_{11}+p_{13}=1-p_{12}$$

$$\begin{aligned}P(V=h_1+ih_2)&=p_{12}p_{22}^{i-1}p_{21}+p_{12}p_{22}^{i-1}p_{23}\\&=p_{12}p_{22}^{i-1}(p_{21}+p_{23})\\&=p_{12}p_{22}^{i-1}(1-p_{22})\quad ;i=1,2,...\end{aligned} \tag{2.73}$$

$$E(V)=h_1+h_2\frac{p_{12}}{1-p_{22}} \tag{2.74}$$

If substitute $E(M_1), E(V)$ in $E(T_1)$, we will get:

$$E(T_1) = \frac{\left[h_1(1-p_{22}) + h_2 p_{12}\right]}{1 - p_{11} - p_{22} + p_{11}p_{22} - p_{12}p_{21}} \tag{2.75}$$

In the same way, it will be:

$$E(T_2) = \frac{\left[h_2(1-p_{11}) + h_1 p_{21}\right]}{1 - p_{11} - p_{22} + p_{11}p_{22} - p_{12}p_{21}} \tag{2.76}$$

3. The mean of the expected time to analyze sample and chart $E(T_a)$

when n is the mean of sample during the process is in control
n' is the mean of sample during the process is out of control
n_1 is the small size sample number which is used in analysis
n_2 is the large size sample number which is used in analysis
p_0 is the probability that the sample size will be small
$(1-p_0)$ is the probability that the sample size will be large
G is the mean of the expected time that uses in sample and chart analysis

$$p_0(\delta) = P\left(|Z| < w_i - \delta\sqrt{n_i} \;\middle\|\; |Z| < k_i - \delta\sqrt{n_i}\right) \qquad i = 1,2 \tag{2.77}$$

where

$$n = n_1 p_0 + n_2(1-p_0) \tag{2.78}$$

$$n' = n_1 p_0(\delta) + n_2\left(1 - p_0(\delta)\right) \tag{2.79}$$

$$E(T_a) = n'G \tag{2.80}$$

4. $E(T_{ass})$ is the expected time to assignable cause while the process is out of the action region; given

$$E(T_{ass}) = T_* \tag{2.81}$$

5. $E(T_r)$ is the expected time to repair; given

$$E(T_r) = T_{**} \tag{2.82}$$

The Expected Cycle Time is the summation of average time periods of each sub cycle time, which is:

$$\begin{aligned} E(T) &= E(T_{in}) + E(T_{out}) + E(T_a) + E(T_{ass}) + E(T_r) \\ &= \frac{1}{\lambda} + (1-\delta_1)E(T_{fa}) + AATS + n'G + T_* + T_{**} \end{aligned} \tag{2.83}$$

T_{*} and T_{**} are independent to each other or independent to process status (halted or continuing).
Hence, from equation (2.44), we obtained:

$$ECTU = \frac{\frac{1}{\lambda}C_0 + C_1\left[AATS + E\left(T_a\right) + \delta_1 T_{*} + \delta_2 T_{**}\right] + Y\,E\left(F\right) + w + \left(a + bn\right)s + \left(a + bn'\right)s'}{\frac{1}{\lambda} + \left(1 - \delta_1\right)E\left(T_{fa}\right) + AATS + n'G + T_{*} + T_{**}} \quad (2.84)$$

2.1.15 Use of Quality Loss Function in the Optimization Model

Serel (2008) found that the traditional formulation of economic design models, the costs due to nonconformities when the process is in control (C_I) and out of control (C_0) have been treated as constants. In recent years, influenced in part by the popularity of Taguchi methods in product design, the quality loss function concept has been incorporated into various statistical decision models where the cost due to poor quality needs to be estimated. In the traditional approach, the upper and lower specification limits are used to classify the quality of the process output as either acceptable or non-acceptable, and products falling outside the specification limits are considered to result in quality costs. In the loss function approach, the probability distribution describing the observations for the quality characteristic is explicitly taken into account in computing the costs, resulting from variation of the quality characteristic round its target. It is considered that cost of poor quality is incurred whenever the quality characteristic is not on its target; hence, products that are not produced on target incur cost, even though, they may conform to specification limits. Several researchers have applied the loss function approach in the economic design of $\overline{X}$ control charts.

2.1.16 Principle Genetic Algorithm

Genetic Algorithm (GA) is a search algorithm developed by Holland (1975) which is based on the mechanics of natural selection and genetics to search through decision space for optimal solutions. The metaphor underlying genetic algorithm is natural selection. In evolution, the problem that each species face is to search for beneficial adaptations to the complicated and changing environment. In other words, each species has to change its chromosome combination to survive in the living world. In Genetic Algorithm, a string represents a set of decisions (chromosome combination), that is a potential solution to a problem. Each string is evaluated on its performance with respect to the fitness function (objective function). The ones with better performance (fitness value) are more likely to survive than the ones with worse performance. Then the genetic information is exchanged between strings by crossover and perturbed by mutation. The result is a new generation with (usually) better survival abilities. This process is repeated until the strings in the new generation are identical, or certain termination conditions are met. A generic flow of Genetic Algorithm is given in Fig. 2.13. This algorithm is continued since the stopping criterion is reached. Genetic Algorithm is used in forming models to solve optimization problems. Readers can find more details of genetic algorithm in Goldberg (1989); Gen and Cheng (2000); Kaya (2009). Genetic Algorithm is different from other search procedures in the following ways (Chen, 2004):

1. Genetic Algorithm considers many points in the search space simultaneously, rather than a single point;
2. Genetic Algorithm works directly with strings of characters representing the parameter set, not the parameters themselves;
3. Genetic Algorithm uses probabilistic rules to guide their search, not deterministic rules. Because genetic algorithm considers many points in the search space simultaneously there is a reduced chance of converging to local optima.

In a conventional search, based on a decision rule, a single point is considered and that is unreliable in multimodal space. Genetic Algorithm consists of four main sections, Encoding, Selection, Reproduction, and Termination (Gen and Cheng, 2000; Goldberg, 1989; Mitchell, 1996).

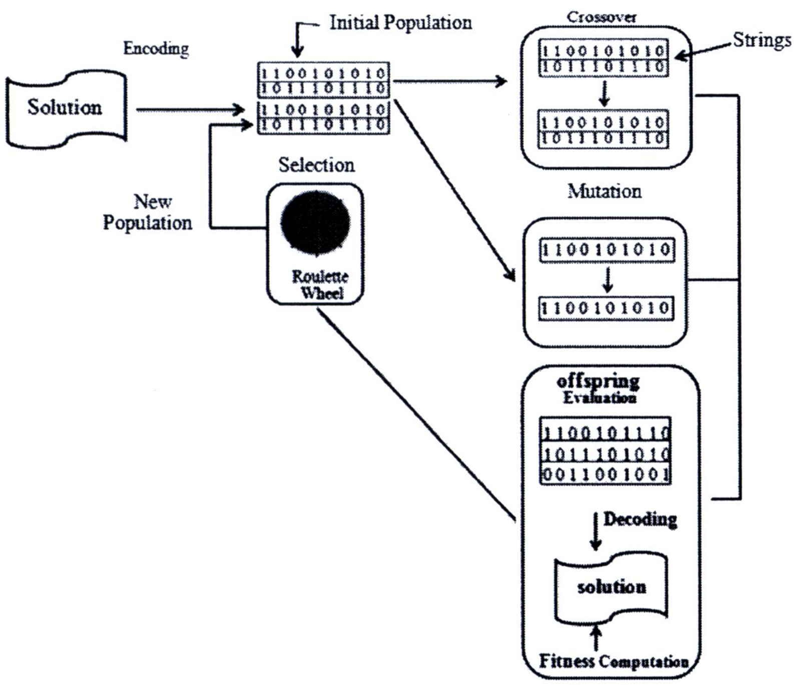


Figure 2.13 The fundamental cycle and operations of basic Genetic Algorithm (Gen and Cheng, 2000)

- Encoding

While using Genetic Algorithm, encoding a solution of a problem into a chromosome is very important. Various encoding methods have been created for particular problems to provide effective implementation of Genetic Algorithm for the last 10 years. According to what kind of symbol is used as the alleles of a gene, the encoding methods can be classified as follows (Gen and Cheng, 2000):

-Binary encoding,
-Real number encoding,
-Integer or literal permutation encoding,
-General data structure encoding

Kaya and Engin (2007) used ‘‘binary encoding” structure to determine the best parameters for Genetic Algorithm in their model. But, Kaya (2009) used ‘‘real number encoding" structure on the same model. And additionally a new chromosome representation was suggested to increase the effectiveness of Genetic Algorithm in that paper. This chromosome contained values on one gene, although the chromosome structure proposed by Kaya and Engin (2007) contained values on many genes.

- Selection

The fundamental idea behind Genetic Algorithm is mainly Darwinian natural selection. The selection leads the genetic search towards encouraging regions in the search space. During study work, many selection methods have been compared and examined. Common types of them are as follows (Gen and Cheng, 2000):

-Roulette wheel selection,
- (k+1)-selection,
-Tournament selection,
-Steady-state reproduction,
-Ranking and scaling,
-Sharing.

In this paper, ''Roulette wheel selection" structure is used as it is a well-known and the most used selection method.

- Recombination

Recombination operator is the most important tool for Genetic Algorithm. It makes the exchange of information acquired by the individuals and its broadcast to the next generation possible. The main two section of recombination is explained briefly as follows:

- Crossover

In Genetic Algorithm, crossover is a genetic operator used to vary the programming of a chromosome or chromosomes from one generation to the next. It is an analogy to reproduction and biological crossover, upon which Genetic Algorithm are based. In this study, five different crossover mechanisms are used and their performances are compared with each other. These are as follows (Kaya and Engin, 2007; Kaya, 2009):

-One-Point Crossover (OPX)
-Position Based Crossover (PBX)
-Order Crossover (OX)
-Partial-Mapped Crossover (PMX)
-Linear Order Crossover (LOX)

- Mutation operator

The premature convergence of a new generation can be prevented by the mutation operator. In this study, five different mutation mechanisms are used and their performances are compared with each other. These are as follows (Kaya and Engin, 2007; Kaya, 2009):

-Inversion Mutation
-Neighbor Exchange Mutation
-Reciprocal Exchange Mutation

2.1.17 Basic Concept in Design of Experiments (DOE) (Nutek, Inc.)

DOE is an experimental strategy in which effects of multiple factors are studied simultaneously by running tests at various levels of the factors. What levels should we take, how to combine them, and how many experiments should we run, are subjects of discussions in DOE.

Factors are variables (also think of as ingredients or parameters) that have direct influence on the performance of the product or process under investigation. Factors are of two types:

Discrete - assumes known values or status for the level.

Continuous - can assume any workable value for the factor levels.

Levels are the values or descriptions that define the condition of the factor held while performing the experiments.

If a factor is tested at two levels, you are forced to assume that the influence of the factor on the result is linear. When three or four levels of a factor are tested, it can indicate whether the factor has non-linear response or not. Factor behavior, that is whether it is linear or non-linear, plays important role in deciding whether to study three or four levels of the factor when the factor is of continuous type. The number of levels of a factor is limited to 2, 3, or 4 in our discussion.

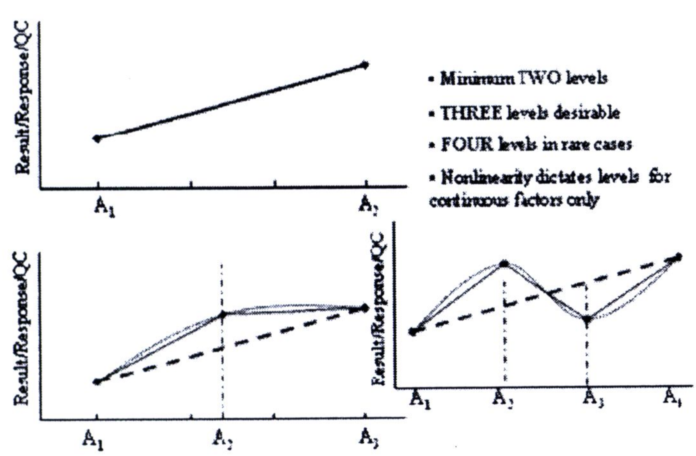


Figure 2.14 Levels of factors (Nutek, Inc.)

Desirable levels of factors for study (Figure 2.14)

- Minimum TWO levels
- THREE levels desirable
- FOUR levels in rare cases
- Nonlinearity dictates levels for continuous factors only

Consider two factors, A and B, at two levels each. They can be tested at four combinations

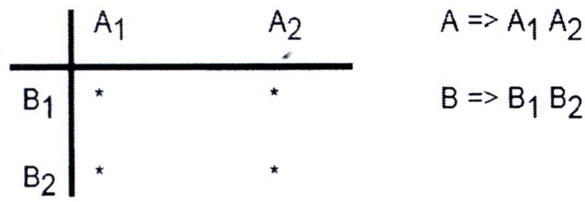


Figure 2.15 Four Experiments (Nutek, Inc.)

Figure 2.15 shows four experiments are: A_1B_1 A_1B_2 A_2B_1 A_2B_2

Likewise three factors A, B and C tested at 2-levels each requires 8 experiments
Factors: $A = A_1, A_2$ $B = B_1, B_2$ $C = C_1, C_2$
8 Experiments:

$A_1B_1C_1$ $A_1B_1C_2$ $A_1B_2C_1$ $A_1B_2C_2$
$A_2B_1C_1$ $A_2B_1C_2$ $A_2B_2C_1$ $A_2B_2C_2$

NOTATIONS
A(A1,A2) or $A^{\angle}$ represent 2-level factor

ONE 2-level factor offer TWO test conditions (A_1A_2)
TWO 2-level factors create FOUR $(2^2 = 4)$
test conditions $(A_1B_1, A_1B_2, A_2B_1, A_2B_2)$

THREE 2-level factors create
EIGHT $(2^3 = 8)$ possibilities.

$A_1B_1C_1$ $A_1B_1C_2$
$A_1B_2C_1$ $A_1B_2C_2$
$A_2B_1C_1$ $A_2B_1C_2$
$A_2B_2C_1$ $A_2B_2C_2$

Cond #	A	B	C
1	1	1	1
2	1	1	2
3	1	2	1
4	1	2	2
5	2	1	1
6	2	1	2
7	2	2	1
8	2	2	2

Figure 2.16 Notation and table shown here is a good way to express the full factorials conditions for a given set of factors included in the study (Nutek, Inc.)

From Figure 2.16 shows that:

1. 2-level factor offers TWO test conditions (A_1, A_2).
2. 2-level factors create FOUR ($2^2 = 4$ test conditions: A_1B_1 A_1B_2 A_2B_1 A_2B_2).
3. 2-level factors create EIGHT ($2^3 = 8$ test conditions: $A_1B_1C_1$ $A_1B_1C_2$ $A_1B_2C_1$ $A_1B_2C_2$ $A_2B_1C_1$ $A_2B_1C_2$ $A_2B_2C_1$ $A_2B_2C_2$) possibilities.

The total number of possible combinations (known as the full factorial) from a given number of factors all at 2-level can be calculated using the following formulas. The industrial practitioners, Taguchi constructed a set of special orthogonal arrays. Orthogonal arrays are a set of tables of numbers designated as $L_4, L_8, L_9, L_{12}, L_{16}$, etc. The smallest of the table, L_4, is used to design an experiment to study three 2-level factors

2.1.18 Principle Standard Orthogonal Arrays

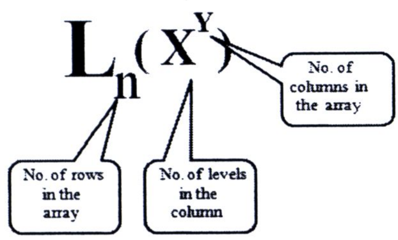
Mitra (1998) found that a product can be designed and manufactured based on a set of specifications demanded by the customer. Each specification has a required parameter value or values, which the manufactured product must be able to satisfy. Thus, the manufacturing process must be capable of producing the designed parameters, which is termed as the targeted value, according to the customer's specifications. Unfortunately in reality, manufacturing processes are far from ideal. Products manufactured tend to give a distribution that has a mean value slightly different from the targeted value. Thus, one of the main techniques used in Taguchi's quality control is to reduce the variation around the targeted value. According to Taguchi, the quality of a group of products can be improved by achieving its end product specifications distribution as close to the target value as possible. This concept can be realized by designing and building the quality into the product itself. Hence, Taguchi employs design experiments using specially constructed table, known as "Orthogonal Arrays (OA)" to treat the design process, such that the quality is build into the product during the product design stage.
Discussions of the various aspects of Orthogonal Arrays (OA) can be found in the following:

2.1.18.1 The Approach of Orthogonal Arrays

Mitra (1998) found that an experiment during the product design stages, involves the materials used in manufacturing the experimental product which affects the final quality outcome. Factors such as variations in the chemical ratio, the level of ingredients used, and how the product is formed together, will contribute to the variation in the targeted value of the final product.

Use this array (L-8) to design experiments with Seven 2-level factors

Trial #	A	B	C	D	E	F	G	Results
1	1	1	1	1	1	1	1	xx
2	1	1	1	2	2	2	2	xx
3	1	2	2	1	1	2	2	xx
4	1	2	2	2	2	1	1	xx
5	2	1	2	1	2	1	2	xx
6	2	1	2	2	1	2	1	xx
7	2	2	1	1	2	2	1	xx
8	2	2	1	2	1	1	2	xx



2-Level Arrays

$L_4(2^3)$

L-4 Orthogonal Array

Trial #	1	2	3
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

$L_8(2^7)$

$L_{12}(2^{11})$

$L_{16}(2^{15})$

3-Level Arrays

$L_9(3^4), L_{18}(2^1 3^7)$...

4-Level Arrays

$L_{16}(4^5)$...

Figure 2.17 Orthogonal arrays used to design experiments (Nutek, Inc.)

This Figure 2.17 describes that:

- The L-4 orthogonal array is intended to be used to design experiments with two or 2-level factors.
- There are a number of arrays available to design experiments with factors at 2, 3, and 4-level.
- The notations of the arrays indicate the size of the table (rows & columns) and the nature of its columns.

Orthogonal Arrays (OA) are a special set of, Completely Randomized Design constructed by Taguchi to lay out the product design experiments. By using this table, an orthogonal array of standard procedure can be used for a number of experimental situations. Consider Standard Orthogonal arrays by general table.

Standard Orthogonal Arrays (Mitra, 1998)

Table 2.2 General form orthogonal arrays

Orthogonal Array	Number of Rows	Maximum Number of Factors	Maximum Number of Columns at These Levels			
			2	3	4	5
L_4	4	3	3	-	-	-
L_8	8	7	7	-	-	-
L_9	9	4	-	4	-	-
L_{12}	12	11	11	-	-	-
L_{16}	16	15	15	-	-	-
L'_{16}	16	5	-	-	5	-
L_{18}	18	8	1	7	-	-
L_{25}	25	6	-	-	-	6
L_{27}	27	13	-	13	-	-
L_{32}	32	31	31	-	-	-
L'_{32}	32	10	1	-	9	-
L_{36}	36	23	11	12	-	-
L'_{36}	36	16	3	13	-	-
L_{50}	50	12	1	-	-	11
L_{54}	54	26	1	25	-	-
L_{64}	64	63	63	-	-	-
L'_{64}	64	21	-	-	21	-
L_{81}	81	40	-	40	-	-

Table 2.3 Orthogonal array of L_4 (2^3)

L_4 (2^3) Orthogonal Array

Experiment Number	Variable Settings		
	1	2	3
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

Table 2.4 Orthogonal array of L_8 (2^7)

L_8 (2^7) Orthogonal Array

Experiment Number	Variable Settings						
	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

Table 2.5 Orthogonal array of L_9 (3^4)

L_9 (3^4) Orthogonal Array

Experiment Number	Variable Settings			
	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

Table 2.6 Orthogonal array of L_{12} (2^{11})

L_{12} (2^{11}) Orthogonal Array

Experiment Number	Variable Settings										
	1	2	3	4	5	6	7	8	9	10	11
1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	2	2	2	2	2	2
3	1	1	2	2	2	1	1	1	2	2	2
4	1	2	1	2	2	1	2	2	1	1	2
5	1	2	2	1	2	2	1	2	1	2	1
6	1	2	2	2	1	2	2	1	2	1	1
7	2	1	2	2	1	1	2	2	1	2	1
8	2	1	2	1	2	2	2	1	1	1	2
9	2	1	1	2	2	2	1	2	2	1	1
10	2	2	2	1	1	1	1	2	2	1	2
11	2	2	1	2	1	2	1	1	1	2	2
12	2	2	1	1	2	1	2	1	2	2	1

Table 2.7 Orthogonal array of L_{16} (2^{15})

L_{16} (2^{15}) Orthogonal Array

Experiment Number	Variable Settings														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
3	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2
4	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1
5	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2
6	1	2	2	1	1	2	2	2	2	1	1	2	2	1	1
7	1	2	2	2	2	1	1	1	1	2	2	2	2	1	1
8	1	2	2	2	2	1	1	2	2	1	1	1	1	2	2
9	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
10	2	1	2	1	2	1	2	2	1	2	1	2	1	2	1
11	2	1	2	2	1	2	1	1	2	1	2	2	1	2	1
12	2	1	2	2	1	2	1	2	1	2	1	1	2	1	2
13	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1
14	2	2	1	1	2	2	1	2	1	1	2	2	1	1	2
15	2	2	1	2	1	1	2	1	2	2	1	2	1	1	2
16	2	2	1	2	1	1	2	2	1	1	2	1	2	2	1

Table 2.8 Orthogonal array of L'$_{16}$ (4^5)

L'$_{16}$ (4^5) Orthogonal Array

Experiment Number	Variable Settings				
	1	2	3	4	5
1	1	1	1	1	1
2	1	2	2	2	2
3	1	3	3	3	3
4	1	4	4	4	4
5	2	1	2	3	4
6	2	2	1	4	3
7	2	3	4	1	2
8	2	4	3	2	1
9	3	1	3	4	2
10	3	2	4	3	1
11	3	3	1	2	4
12	3	4	2	1	3
13	4	1	4	2	3
14	4	2	3	1	4
15	4	3	2	4	1
16	4	4	1	3	2

Table 2.9 Orthogonal array of L_{18} (2^1 x 3^7)

L_{18} (2^1 x 3^7) Orthogonal Array

Experiment Number	Variable Settings							
	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	1
2	1	1	2	2	2	2	2	2
3	1	1	3	3	3	3	3	3
4	1	2	1	1	2	2	3	3
5	1	2	2	2	3	3	1	1
6	1	2	3	3	1	1	2	2
7	1	3	1	2	1	3	2	3
8	1	3	2	3	2	1	3	1
9	1	3	3	1	3	2	1	2
10	2	1	1	3	3	2	2	1
11	2	1	2	1	1	3	3	2
12	2	1	3	2	2	1	1	3
13	2	2	1	2	3	1	3	2
14	2	2	2	3	1	2	1	3
15	2	2	3	1	2	3	2	1
16	2	3	1	3	2	3	1	2
17	2	3	2	1	3	1	2	3
18	2	3	3	2	1	2	3	1

Table 2.10 Orthogonal array of L_{25} (5^6)

L_{25} (5^6) Orthogonal Array

Experiment Number	Variable Settings					
	1	2	3	4	5	6
1	1	1	1	1	1	1
2	1	2	2	2	2	2
3	1	3	3	3	3	3
4	1	4	4	4	4	4
5	1	5	5	5	5	5
6	2	1	2	3	4	5
7	2	2	3	4	5	1
8	2	3	4	5	1	2
9	2	4	5	1	2	3
10	2	5	1	2	3	4
11	3	1	3	5	2	4
12	3	2	4	1	3	5
13	3	3	5	2	4	1
14	3	4	1	3	5	2
15	3	5	2	4	1	3
16	4	1	4	2	5	3
17	4	2	5	3	1	4
18	4	3	1	4	2	5
19	4	4	2	5	3	1
20	4	5	3	1	4	2
21	5	1	5	4	3	2
22	5	2	1	5	4	3
23	5	3	2	1	5	4
24	5	4	3	2	1	5
25	5	5	4	3	2	1

2.1.19 Principle of Multiple Linear Regressions

Walpole (2007) in most research problems where regression analysis is applied, more than one independent variable is needed in the regression model. The complexity of most scientific mechanisms is such that in order to be able to predict an important response, a multiple regression model is needed. When this model is linear in the coefficients, it is called a multiple regression analysis model. For the case of p independent variables $x_1, x_2, \ldots, x_p$, the mean of $Y|x_1, x_2, \ldots, x_p$ is given by the multiple linear regression model

$$\mu_Y|_{x_1, x_2, \ldots, x_p} = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p \tag{2.85}$$

and the estimated response is obtained from the sample regression equation:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p \tag{2.86}$$

where each regression coefficient β_i is estimated by $\hat{\beta}_i$ from the sample data using the method of least squares. As in the case of a single independent variable, the multiple linear regression models can often be an adequate representation of a more complicated structure within certain ranges of the independent variable.

Multiple linear regression models, particularly when the number of variable exceeds two, knowledge of matrix theory can facilitate the mathematical manipulations considerably. Suppose that the experimenter has p independent variable $x_1, x_2, \ldots, x_p$ and n observations $y_1, y_2, \ldots, y_n$, each of which can be expressed by the equation

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_p x_{1p} + \varepsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_p x_{2p} + \varepsilon_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_p x_{np} + \varepsilon_n \end{aligned} \tag{2.87}$$

Where the error terms are assumed to have the following properties:

1. $E(\varepsilon_i) = 0$
2. $Var(\varepsilon_i) = \sigma^2$ (constant); and (2.88)
3. $Cov(\varepsilon_j, \varepsilon_k) = 0, \; j \neq k$

In matrix notation, equation (2.87) becomes

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_r \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

And the specifications in equation (2.88) become

1. $E(\varepsilon) = 0$
2. $Cov(\varepsilon) = E(\varepsilon\varepsilon') = \sigma^2 I$

The errors of the estimation y_i by $\hat{y}_i$ are $y_i - \hat{y}_i = e_i$ (by using the method of least squares) (see Appendix A.7)

2.1.20 Analysis of Data

Mitra (1998) found that the statistical procedure used most often to analyze data is known as the analysis of variance (ANOVA). This technique determines the effects of the treatments, as reflected by their means, through an analysis of their variability. Details of this procedure are found in the listed reference (Box, et al., 1978). The total variability in the observation is partitioned into two components: the variation among the treatment means (also known as the treatment sum of squares) and the variation among the experimental units within treatments (also known as the error sum of squares). We have

$$\text{Total sum of square } (\text{SST}) = \text{Treatment sum of square } (\text{SSTR}) \\ + \text{Error sum of square } (\text{SSE})$$

The mean squares for treatments and for error are obtained by dividing the corresponding sum of squares by the appropriate number of degree of freedom. This number is 1 less than the number of observations in each source of variation. For a balanced design with p treatments, each with r replications, the total number of observations is rep. The total variability, therefore, has $(rep-1)$ degree of freedom. The number of degree of freedom for the treatments is $(p-1)$. For each treatment, there are r observations, so $(re-1)$ degrees of freedom apply toward the experimental error. The total number of degrees of freedom for the experimental error is, therefore, $p(re-1)$. We have the following notation:

p = Number of treatments

re = Number of replications for each treatment

y_{ij} = Response variable value of the j^{th} experimental unit that is assigned treatment i, $i = 1,2,3,\ldots,p$; $j = 1,2,3,\ldots,re$

$y_{i.}$ = Sum of the responses for the ith treatment; that is, $\sum_{j=1}^{r} y_{ij}$

$\bar{y}_{i.}$ = Mean response of the ith treatment; that is, $\dfrac{y_{i.}}{re}$

$y_{..}$ = Grand total of all observations; that is, $\sum_{i=1}^{p}\sum_{j=1}^{r} y_{ij}$

$\bar{\bar{y}}$ = Grand mean of all observations; that is, $\dfrac{y_{i..}}{rep}$

The notation, consisting of re observations for each of the p treatments, is shown in Table 2.11. The computations of the sum of squares are as follows. A correction factor C is first computed as:

$$C = \frac{y_{..}^2}{rep} \tag{2.89}$$

The total sum of square is:

$$\text{SST} = \sum_{i=1}^{p}\sum_{j=1}^{r} y_{ij} - C \tag{2.90}$$

The treatment sum of squares is determined form:

$$\text{SSTR} = \frac{\sum_{i=1}^{p} y_{i.}^2}{r} - C \tag{2.91}$$

Finally, the error sum of square is:

$$\text{SSE} = \text{SST} - \text{SSTR} \tag{2.92}$$

Next, the mean squares are found by dividing the sum of squares by the corresponding number of degree of freedom. So, the mean squares for treatment are:

$$\text{MSTR} = \frac{\text{SSTR}}{p-1} \tag{2.93}$$

The mean square error is given by:

$$MSE = \frac{SSE}{p(re-1)} \tag{2.94}$$

Test for Differences among Treatment Means

It is desirable to test the null hypothesis that the treatment means are equal against the alternative hypothesis that at least one treatment mean is different from the others. Denoting the treatment means by $\mu_1, \mu_2, \ldots, \mu_p$, we have the hypothesis

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_p$

H_1 : At least one μ_i is different from the others

The test procedure involves the $F-$Statistic, which is the ratio of the mean squares for treatment to the mean squares for error. The mean square error (MSE) is an unbiased estimate of σ^2, the variance of the experimental error. The test statistic is:

$$F = \frac{\text{MSTR}}{\text{MSE}} \tag{2.95}$$

Table 2.11 Notation for the Completely Design

	Replication					
Treatment	1	2	$\cdots$	*re*	Sum	Mean
1	y_{11}	y_{12}	$\cdots$	y_{1re}	$y_{1\cdot}$	$\bar{y}_1$
2	y_{21}	y_{22}	$\cdots$	y_{1re}	$y_{2\cdot}$	$\bar{y}_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
p	y_{p1}	y_{p2}	$\cdots$	y_{1re}	$y_{p\cdot}$	$\bar{y}_p$
					$y_{\cdot\cdot}$	$\bar{\bar{y}}$

with $(p-1)$ degrees of freedom in the numerator and $p(re-1)$ degrees of freedom in the denominator. For a chosen level of significance α, the critical value of F, which is found from the table value of F, is denoted by $F_{\alpha,(p-1),p(re-1)}$. If the computed test statistic $F > F_{\alpha,(p-1),p(re-1)}$, the null hypothesis is rejected, and we conclude that the treatment means are not all equal at the chosen level significance. This computational procedure is known as analysis of variance; it is shown in tabular format in Table 2.12.

Table 2.12 The analysis of variance (ANOVA) table

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F
Treatments	$p-1$	SSTR	$\text{MSTR} = \frac{\text{SSTR}}{p-1}$	$F = \frac{\text{MSTR}}{\text{MSE}}$
Error	$p(re-1)$	SSE	$MSE = \frac{SSE}{p(re-1)}$	
Total	$rep-1$	SST		

2.1 Literature reviews

The general purpose of SPC is useful in establishing and maintaining a state of statistical control and identifying special cause of variation. Two terms frequently used in SPC are Common cause and Special cause variations. In general terms, Common cause variation refers to the inherent natural variability in a process. Special cause variation is attributable to some assignable cause or change to the process which manifests itself in form of outliers, shifts or trends of some sort in data stream. Woodall (2000) states that differences in opinion exist about the purpose and scope of SPC strategy due to diversity of those working in quality field, including quality gurus and their followers, consultants, quality engineers, industrial engineers, professional practitioners, statisticians, managers, and others. In this section, the overall purpose and scope of SPC strategy are reviewed. Shewhart and his associates developed the Shewhart control charts during 1920's at Bell Telephone Laboratories. Shewhart (1931) defined maximum control as "condition reached when the chance cause fluctuations in a phenomenon produced by constant system of large number of chance causes in which no cause produces a predominating effect". He states that the primary purpose of control chart is to distinguish between two types of variation, Common cause and Special cause, in order to prevent over reaction or under reaction to the process. He considered Common cause of variation as set of causes attributable to inherent nature of the process that cannot be altered without changing the process itself, and also assignable cause of variation as unusual shocks and disruptions to process the causes of which can and should be removed. Some authors including Juran (1999) believed that control chart and test of hypothesis are very closely related. In this context there is an accept/reject decision based on the value of charted statistical and decision regions. Thus, a process is said to be in statistical control if the probability distribution representing quality characteristic is constant over time. The control chart is a useful tool for distinguishing between "in control" (stable) and "out of control" (unstable) operation in a case of an identically independently distributed (iid) data stream. Woodall (2000) stated that control charts are used to check process stability. In this context, a process is said to be in state of "statistical control" if the probability distribution representing the quality characteristic is constant over time. If there are some changes in the distribution, the process is said to be "out of control". Deming (1986) saw possibility of long term process improvement as being far more important than detection of changes. Deming clearly stated that meeting specification limits is not sufficient to ensure good quality and the variability of quality characteristic should be reduced such that "specifications are lost beyond horizon". Thus, for his goal of statistical process control corresponds to centering quality characteristic at target and continuously reducing variability. Deming strongly advocated the use of control charts but argued empathetically against hypothesis testing. Montgomery (2001) stated that SPC is a powerful collection of problem solving tools and useful in achieving process stability and improving capability through reduction of variability. Accordingly, the fundamental use of control chart is reduction of process variability, monitoring and surveillance of a process, estimation of product and process parameters. He stated that most important use of control chart is to improve the process by reducing variability. The process improvement activity using control chart is illustrated in Figure 2.18

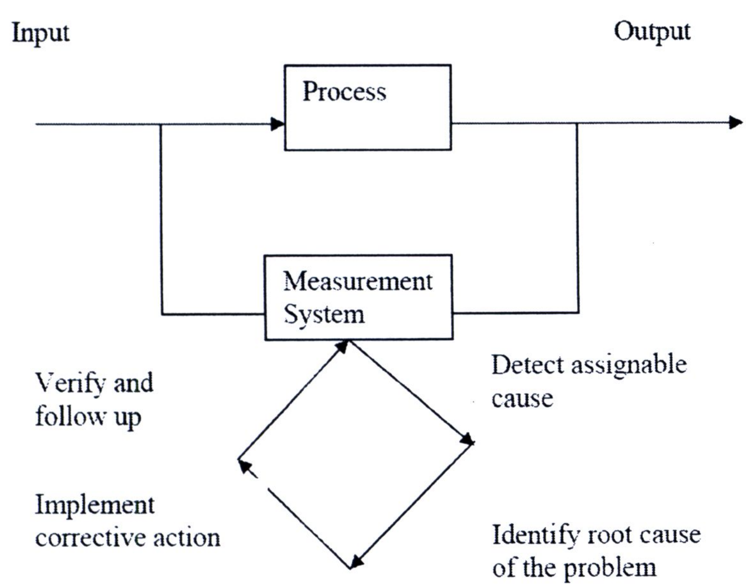


Figure 2.18 Process improvement using control chart (Montgomery, 2001)

Steiner and Mackay (2000) stated that there are three main uses of control charts:

1. To reduce the variation in an output characteristic by establishing a control chart to signal the change of an unidentified process input. The occurrence of the signal sets effort to identify this input.
2. To determine by when and by how much a process should be adjusted. A control chart is setup and adjustments are made only when a signal occurs.
3. To demonstrate process being stable and capable. The purpose here is to provide information to make decision regarding the receiving inspection.

Box and Luceno (1997) states that during a successful implementation of a control chart, practitioner needs to address three important decisions:

1. Is control chart an appropriate tool for application?
2. Which type of control chart to use?
3. Where should control limits be placed?

They indicated that the answer to the first question depends upon whether or not stable periods without changes in process mean or variance exist. If there is a stable variance but the process mean drifts, then automatic process control strategy should be considered as a means of reducing variability. They pointed out that answers to the second and third questions will depend on how these charts will be used; i.e. real time process monitoring, problem solving, assessment of process stability, nature of disturbance to be detected.

Control charts have been widely used for monitoring process stability and capability. Control charts are based on data, representing one or several quality-related characteristics of the product or service. If these characteristics are measurable on numerical scales, then variable control charts are used. On the often work, if the quality-related characteristics cannot be easily represented in numerical form, then attribute control charts are useful (Gulbay and Kahraman, 2006). Generally, they concentrated on if the process is ''under control'' or ''out of control''. However other quality constraints like quality cost, rate of errors, acceptance probability, consumer and producer risks, etc. must also be taken into account.

The quality chart that is averaged by exponential weighted method (Robert, 1959) is suitable for the data that has an average changed or slightly shifted. This shift could not be observed if Walter Shewhart's quality graph (Shewhart, 1942) is used in the estimation (Crowder, 1989; Lucus and Saccucci, 1990). Few studies attempted to establish the model in selection of parameters for using in the EWMA chart. Girschick and Rubin (1952) studied the selection of parameters for Shewhart $\bar{X}$ chart using the Duncan's model (1956). For Grant and Leavenworth (1996), they designed the selection of parameters using the DS EWMA charts (double sampling chart) for the EWMA control chart by considering with average run length (ARL). When the process is in the control region with $r = 0.75$, the DS EWMA chart will have more efficiency in defective detection than the normal EWMA chart. But if $r = 0.5$ or $r = 0.25$, the EWMA chart will be suitable to detect small shifts (a small change of mean) and the DS EWMA chart is good for large shift detection.

Estimation of ARL depended on r value using Simulation was studied by Roberts (1959). Robinson and Ho (1978) derived function of ARL value and established the table of ARL value. Later, Crowder (1987a) developed a method to convert ARL value into general term. Crowder (1987b, 1987c) and Saccucci, et al. (1990) calculated average run length (ARL) and property of the EWMA chart, and established the ARL value table for the EWMA chart. Montgomery and Mastrangelo (1991), Mastragelo and Montgomery (1995), and Mastrangelo and Brown (2000) studied the application of the moving centerline EWMA. Reynolds (1996a, 1996b) studied the EWMA chart by varying sampling number (n) and the width of control limit (k). Selection of quality control chart and fixation of parameters for the chart needs to concern the cost in process. Moreover, utilization of economic methodology together with the quality control chart would yield the lowest cost in production. For example, Magalhães and Epprecht (2001) used the economic design together with variable parameters for $\overline{X}$ chart and revealed that this results in lower cost than the use of constant parameters.

Lorenzen and Vance (1986) proposed a general method for determining the economic design of control charts. This method can be applied regardless of the statistic used. It is necessary to calculate only the average run-length of the statistics when assuming that the process is in-control and also assuming that the process is out-of-control in some specified manner. Alexander, et al. (1995) combined Duncan's cost model with the Taguchi loss function to develop a loss model for determining the three test parameters. This loss model explicitly considers the quality. Montgomery (2001) and Ho and Case (1994b) considered economic design of control charts for monitoring the process mean, which has been investigated extensively in the literature. Ho and Case (1994a) presented literature on control charts employing an EWMA type statistic. Several authors have explored the economic design of EWMA control charts to monitor the process mean. Park and Reynolds (1994) extended the traditional economic design of an EWMA chart to the case where the sampling interval and sample size may vary depending on the current chart statistic. Park and Reynolds (2008) considered IPC monitoring schemes by using an economic design approach under the inherent wandered of the process. It can be represented as an ARIMA (0,1,1) model. They consider a combination of two EWMA charts, with one EWMA statistic using the observed deviations from target, and the other EWMA statistic using the squared deviations from the target. They found that, if it is desirable to use only one control chart for simplicity, then the two EWMA control chart provides very good performance and would be preferable to using the EWMA chart. It is interesting to note that the EWMA chart is the standard. It would usually be considered for monitoring a process in the current setting, but the two EWMA control chart, actually have much better performances. Serel (2008) considered the case where the assignable cause changes only the process mean or dispersion. The economic design of EWMA mean charts was extended to the case where quality related costs are computed based on a loss function. They used the loss function to estimate. Serel and Moskowitz (2008) considered when the assignable causes lead to changes in both process mean and variance, simultaneous use of mean and dispersion charts is important for detecting the changes quickly. Joint economic design of EWMA charts for process mean and dispersion have been explored.

The use of a control chart requires the user to select several design parameters. For the fixed parameter control chart, these design parameters requiring predetermined, include the fixed sample size, control limits, and sampling interval length. One method of designing control chart based on an economic criterion is said the economic design. The usual approach to the economic design is to develop a cost model for a particular type of manufacturing process, and then derive the optimal parameters by minimizing the long-run expected cost per hour. The cost models have been widely used in determining the design parameters. The EWMA control chart design methods specify the optimal selection of variable parameters (sample size(n), the interval between samples(h), exponential weight used for each quality characteristic(r), the control limit for the EWMA process (k)in the chart. These decisions were based on statistical criteria, through restriction the probability of Type I or Type II errors. Using control charts is in fact economically motivated though the selection of design parameters did not original use cost information to motivate. The operator can control the cost of running and monitoring a process by ad hoc basis though cost tradeoffs are not explicitly used to choose chart parameters. So control charts became increasingly popularity, the idea of designing charts on the basis of cost tradeoffs, leading to economic design. To use the economic design and statistic criteria, the purpose is to minimize the average cost when a single out-of-control state (assignable cause) occurs. Duncan's cost model includes the cost of sampling and inspection, the cost of defective products, the cost of false alarms, the cost of searching for assignable caused, and the cost of process correction.

Stoumbos, et al. (2000) used the theory of a Variable Sampling Interval (VSI) and economic model to reduce the cost in production. For Chou, et al. (2008), they developed economic design of EMMA by Varying Sampling Intervals with sampling at Fixed Times (VSIFT) to estimate warning limit coefficient value(w), the control limit coefficient(k), and the exponential weight constant(r) for the lowest cost using the expense model of Lorenzen and Vance (1986). From the study, it was found that if the mean of the process has largely shifted from the target mean, in general this will reduce the sample size, resulting in a decreased number of sampling in each fixed interval and wider control limits. The increased frequent alarm will result in an increased sampling number and a wider control limits. The production cost of the defective products will increase when the production is out of the control limits, leading to a decreased sampling time. But if the cost per sampling unit is higher, this would lengthen the fixed sampling interval. Moreover, both the production cost of the defective products when the products are either in control limits or out of control limits, and the increased frequent alarm are a cause of higher total expenses. Fixation of parameters $(n, h, w, k, r)$for quality control chart is also as important as selection of the chart. This is because parameter value has effect on average run length(ARL), which is the mean of the points locating within the action region. Therefore, ARL value of the in control process and the cost in quality inspection will be higher than those of out of control.

There are several researchers studied for optimal parameter fixation. For example, Stoumbos and Reynolds (2001) used variable time intervals (h) and size of samples as parameters in the mix chart between EWMA chart and $\bar{X}$ chart. The results showed that the production cost was reduced. Besides, Marcela and Costa (2008) used double sampling for the EWMA chart and found that the number of warning was reduced when the average was changed without an increase of warning rate.

In this latter case a variable parameter sampling interval (VPSI) was considered by Runger and Pignatiello (1991); Amin and Miller (1993); Runger and Montgomery (1993); Reynold, et al. (1988, 1990) and Reynolds (1996a, 1996b). The variable sampling interval feather was extended to Cusum and EWMA charts (see Reynolds, et al., 1988, 1990 and Saccucci, et al., 1990). Recently, Baxley (1995) presented an application of EWMA chart with VPSI for Monsanto's nylon fiber plant in Pensacola, Florida. The size of samples was the second design parameter to be considered variable (see Prabhu, et al., 1993 ; Costa, 2008), subsequently, both parameters (sample size and sampling interval) were made variable (see Prabhu, et al., 1993; Costa, 1994; Rendtel, 1990) consider Cusum schemes with variable sampling intervals and sample size. Finally all design variable parameters were considered variable the economic design $\bar{X}$ chart with VSS was studied by Flaig (1991) and Park and Reynolds (1994), Park and Reynolds proposed an economic model for $\bar{X}$ chart with VSS when the process is subject to the occurrence of several assignable causes. Das, et al. (1997) developed a cost model for optimal dual sampling interval (DSI) policies with and without run rules. Also, Das, et al. (1997) proposed a further generalization of VSI policy for $\bar{X}$ chart in which the sampling intervals are treated as random variables and the sample sizes are considered a function of the sampling intervals. Magalhães and Epprecht (2001) proposed an economic design of a variable parameters $\bar{X}$ chart. Recently, Pongpullponsak and Charongrattanasakul (2009) proposed minimizing the cost of integrate systems approach to process control and maintenance model by EWMA control chart using genetic algorithm.

Generally there are two groups of SPC, i.e. Univariate Statistical Process Control (USPC) and Multivariate Statistical Process Control (MSPC) which are used for different scenarios. The univariate control chart has only one process output variable or quality characteristic measured and tested. One of the disadvantages of the USPC is that for a single process, there are many variables simultaneously. But there are many situations where control chart of two or more correlated quality characteristics are important, so the univariate quality control chart is not always the best method for monitoring correlated characteristics. This is because the correlations between variables result in degrading the statistics performance of these charts and the advancement in technology, complexity in product' process, customers' demand of higher quality, and competition in market , so it is necessary to use multivariate statistical process control. One common method of constructing multivariate control charts is based on Hotelling's statistics (Hotelling, 1974; Alt, 1985). This T^2 chart can be considered as the multivariate extension of the univariate Shewhart control charts, based on the monitoring of the means in independent samples. Multivariate CUSUM charts have been proposed by Woodall and Ncubed (1985) and Croisier (1988).

The first reference on MEWMA control charts corresponds to Lowry, et al. (1992) who define MEWMA as an extension of the univariate, only takes into account current process data, whereas MEWMA chart also includes past data, thereby it being more powerful to detect small changes in the process. Univariate systems only controlled one quality variable or characteristic. In multivariate systems a set of p interrelated variable will be controlled. Although MEWMA is more sensitive to small shifts than Hotelling, Hotelling T^2 is more sensitive in detecting a sudden change in the parameters, to take advantage of the power of control chart. Murphy (1987) proposed a method to identify the "out-of-control" variables based on discriminating between the process of being "in control" or "out-of-control". Murphy divided the complete set of variables into two subsets and then tried to determine which one of the subsets caused the "out-of-control" signal.

An extension of Murphy's work is Chua and Montgomery (1992). They proposed a three steps quality control process by using a MEWMA control chart, a backward selection algorithm a hyper plane method. A Multivariate Exponential Weighted Moving Average (MEWMA) control chart is established after every new observation a continuous basis until an out-of-control signal appears. If an out-of-control signal appears, then the backward selection algorithm and the hyper plane method are used to diagnose it. Linderman and Love (2000) presented the economic and economic statistical design procedure for MEWMA control chart based on Lorenzen-Vance cost function, with a lower bound on the in-control average run length (ARL_0) and upper bound for the out-of-control average run length (ARL_1). They solved this model by the Hooke and Jeeves's algorithm (1961) in which the ARL was estimated by simulation. By changing the method of ARL estimation from simulation to Markov chain approach, Molnau, et al. (1997) later used the same model and presented the results based on an experimental design.

Van Nuland (1992-1993) proposed a circle chart, this circle chart is only approximate and our simulations show that for n, the subgroup size, as large as 20, its actual false-alarm rate is 0.018 when the nominal rate is 0.05. This is very conservative and the resulting chart is therefore less sensitive in detecting shifts. Chao and Cheng (1996) developed a control chart, the Semicircle (SC) chart. This chart can jointly combine the detection of the mean shift and variability change into one single chart, and is simple to use and easy to understand. One of the most impressive features of the Semicircle chart is that it is easy to attribute an out-of-control signal to the cause of the mean shift or/and variability change. However, the SC chart is insensitive to small changes within a process. Combining the features of the Semicircle chart with the EWMA technique, the EWMA technique is directly applied to the static employed in the SC chart. Chen, et al. (2004) proposed a new EWMA Control Chart for Monitoring Both Location and Dispersion. This chart is very sensitive in detecting small changes within a process when a mean shift accompanies an increased variability change. It can simultaneously monitor both the process mean and the increased process variability, and detect the source and the direction of an out-of-control signal.

In this work, we develop an economic model of EWMA control chart with variable parameters by Lorenzen and Vance (1986). We used Costa Model to design variable parameters which vary in real time, based on current sample information. This model is developed, providing a cost function which represents the cost per time unit or controlling the quality of a process through a MEWMA control chart and Semi-circle MEWMA control chart. As the cost function is a function of the design parameters of the control chart, it provided a device for optimal selection of design parameters. So, we will consider the cost function of MEWMA control charts having all design parameters variable.