# A STUDY ON VARIOUS SINGLE VARIABLE CONTROL CHARTS

BY

LIM SEE KEN

**Project submitted in partial fulfillment of the requirements for the degree of Master of Science (Teaching of Mathematics)**

May 2009

## ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who gave me the possibility to complete this project. First of all, I would like to thank my project supervisor, Assoc. Prof. Michael Khoo Boon Chong from the School of Mathematical Sciences, Universiti Sains Malaysia, whose advice, stimulating suggestions and encouragements helped me in writing and editing this project.

At the same time, I am indebted to the library of Universiti Sains Malaysia for providing me with related reference books and journals for the completion of my project.

Last but not least, I would like to thank all those who have helped and supported me in one way or another throughout my study.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF APPENDICES

# SUATU KAJIAN TENTANG CARTA-CARTA KAWALAN BERPEMBOLEHUBAH TUNGGAL

## ABSTRAK

Teknik carta kawalan telah digunakan dengan luasnya di industri untuk mengawal proses bagi peningkatan kualiti. Kebiasaannya, dua carta kawalan diperlukan untuk mengawal min proses dan varians proses secara berasingan. Banyak usaha telah dilakukan untuk menghasilkan sejenis carta kawalan tunggal yang mampu mengawal min proses dan varians proses secara serentak. Dalam 15 tahun yang lalu, banyak carta kawalan tunggal telah dicadangkan untuk membolehkan kawalan serentak min dan varians proses. Kawalan serentak min dan varians proses adalah lebih bermakna oleh sebab dalam situasi sebenar, min dan varians proses boleh berubah secara serentak. Objektif projek ini adalah untuk membandingkan keberkesanan beberapa jenis carta kawalan tunggal, iaitu carta semibulatan, carta MaxEWMA dan carta MA tunggal, daripada segi purata panjang larian terkawal dan luar kawalan mereka.

# ABSTRACT

Control chart techniques have been widely used in industries to monitor a process for quality improvement. Usually, two control charts are required to monitor both the process mean and variability. Efforts have been made to use a single control chart to monitor both the process mean and variability at the same time. Numerous single variable control charts were suggested in the last 15 years to allow a simultaneous monitoring of both the process mean and variance. A simultaneous monitoring of the mean and variance is more meaningful as in a real situation; the mean and variance may shift at the same time. The objective of this project is to compare the performances of several single variable control charts, namely the semicircle chart, MaxEWMA chart and single MA chart, in terms of their in-control and out-of-control average run lengths (ARLs).

# CHAPTER 1

# INTRODUCTION

## 1.1 A Brief History of Statistical Quality Control

Issues on quality have existed since the rule of tribal chiefs, kings and pharoahs. An example of a quality issue in ancient times is found in the Code of Hammurabi, dating from 2150 B.C. Item 229 states that "If a builder has built a house for a man and his work is not strong, and the house falls in and kills the householder, that builder shall be slain." Phoenician inspectors eliminated any repeated violations of quality standards by chopping off the hands of the maker of a defective product (Gitlow, 1989). Inspectors accepted or rejected products and enforced government specifications. The emphasis was on equity of trade and complaint handling. In ancient Egypt (approximately 1450 B.C.), inspectors checked the squareness of stone blocks with a string as the stonecutter watched. This method was also used by the Aztecs in Central America (Gitlow, 1989).

Statistical methods and their application in quality improvement have had a long history. In 1924, Walter A. Shewhart of the Bell Telephone Laboratories developed the statistical control charts concept, which is often considered the formal beginning of statistical quality control. Towards the end of the 1920s, Harold F. Dodge and Harry G.

Romig, both from the Bell Telephone Laboratories, developed statistically based acceptance sampling as an alternative to 100% inspection. By the middle of the 1930s, statistical quality control methods were in wide use at Western Electric, the manufacturing arm of the Bell System. However, at that time, the importance of statistical quality control was not widely recognized by industries (Montgomery, 2005).

World War II saw a greatly expanded use and acceptance of statistical quality control concepts in manufacturing industries. Wartime experience made it apparent that statistical techniques were necessary to control and improve product quality. The American Society for Quality Control was formed in 1946. This organization promotes the use of quality improvement techniques for all types of products and services. It offers a number of conferences, technical publications, and training programs in quality assurance (Montgomery, 2005).

The next phase was the Total Quality Control period during the 1960s. An important feature during this period was the beginning of the gradual involvement of several departments and management personnel in the quality control process. A misconception prior to this period was that quality control is the responsibility of the inspection department. Oddly enough, this concept was accepted by many, the idea of which led to quality being considered someone else's "stepchild." People began to realize that each department had an important role in the production of a quality item (Miller, 1995).

The advent of the 1970s brought what Feigenbaum (1983) called the Total Quality Control Organization wide phase. This phase involved the participation of everyone in the company. Quality was associated with every individual. As this notion continued in the 1980s, it was termed by Feigenbaum (1983) the Total Quality System, which he defined as follows:

*A quality system is the agreed upon companywide and plant wide operating work structure, documented in effective, integrated technical and managerial procedures, for guiding the coordinated actions of the people, the machines, and the information of the company and plant in the best and most practical ways to assure customer quality satisfaction and economical costs of quality.*

In the 1980s, motivating slogans placed quality control in the limelight in the United States. Consumers were bombarded with advertisements relating to the high quality of the product and frequent comparisons were made with those of the competitor. These promotional efforts tried to point out certain product characteristics that were superior to that of similar products. Within the industry itself, an awareness of the importance of quality was beginning to evolve at all levels (Mitra, 1993).

With the continued growth of the use of computer in industry during the 1980s, an abundance of quality control software programs emerged in the market. The notion of a total quality system created an increased emphasis on vendor quality control, product design assurance, product quality audit, and other related areas. Industrial giants such as the Ford Motor Company and General Motors Corporation adopted the quality philosophy and made strides in the implementation of statistical quality control methods. They in turn continued to influence other companies to use quality control techniques. Thus, smaller companies who had previously not used statistical quality control methods were forced to adopt these methods in order to maintain their contracts. This process of requiring evidence of using quality control procedures will likely continue down to the smallest contractor or vendor. The 1990s will see the expanded use of quality control measures and increased attention to customers' needs. There will be no escape from the reality that the customer is the determinant of the level of quality. Industry will have to adjust to meet the needs of the customer.

Top management, responsible for the never-ending improvement of quality, must understand the relationship between quality and productivity and the benefits of improving quality. The quality environment of a firm is critical and stresses teamwork and communication. Deming's 14 points for management provide a guide to creating and establishing the quality environment through behavioral change and using statistical methods to continually improve the process (Del Castillo, 2002).

## 1.2 Objective of the Study

Numerous single variable control charts were suggested in the last 15 years to allow a simultaneous monitoring of both the process mean and variance. A simultaneous monitoring of the mean and variance is more meaningful as in a real situation; the mean and variance may shift at the same time. The objective of this project is to compare the performances of several single variable control charts, namely the semicircle, MaxEWMA and single MA charts, in terms of their in-control and out-of-control average run lengths (ARLs).

## 1.3 Organization of the Project

This section will discuss the organization of the project.

Chapter 1 provides a brief history of statistical quality control and focuses on quality improvement in any service or manufacturing organization. This chapter also discusses the different types of quality and the losses society incurs from the lack of quality in goods and services. The objective of this project is also presented in this chapter.

Chapter 2 highlights the uses of control charts. An overview on the structure of control charts will also is included in this chapter. This chapter also gives some discussion on average run length (ARL), a commonly used measure of performance for control charts.

Chapter 3 is mainly focusing on a review of the conventional Shewhart type charts for monitoring both the process mean and variance. A discussion on when and how to revise the limits of control charts is also given in this chapter.

Chapter 4 presents three types of single variables control charts, namely the semicircle, MaxEWMA and single MA charts. Procedures and calculations of the charts' statistics and control limits are shown.

Chapter 5 compares the performances of the three types of single variable charts considered in Chapter 4. Simulation studies are conducted using SAS to compute the ARLs of the charts. A discussion is also given to explain the performances of the charts.

Chapter 6 gives the conclusion of the study in this project. A brief discussion on the contributions of this project is given. Topics for further research are also suggested in this chapter.

# CHAPTER 2

# CONTROL CHARTS

## 2.1 Introduction

A process that has been defined and documented can be stabilized and then improved. In great measure this can be accomplished through the use of statistical control charts. These tools and methods must be used in an environment that provides a positive atmosphere for process improvement. W. E. Deming points out that "any attempt to use statistical techniques under conditions that rob the hourly worker of his pride in workmanship will lead to disaster" (Gitlow, 1989).

The Deming cycle is a method that can aid management in stabilizing a process and pursuing continuous and never-ending process improvement.

The Deming cycle operates by recognizing that problems (opportunities for improvement) in a process are determined by the difference between customer (internal and / or external) needs and process performance. If the difference is large, customer dissatisfaction may be high, but there is great opportunity for improvement. If the difference is small, the consequent opportunity for improvement is diminished. Nevertheless, it is always economical to continually attempt to decrease the difference between customer needs and process performance. Control charts are used to monitor and identify problems in a process.

## 2.2 Average Run Length (ARL)

Once one begins to contemplate alternative schemes for issuing out-of-control signals based on process-monitoring data, the need quickly arises to quantify what a given scheme might be expected to do. For example, if one is to choose intelligently between the Western Electric and Nelson sets of alarm rules introduced in the previous chapter, one needs some means of predicting the behavior of two kinds of monitoring schemes. The most effective means known for making this kind of prediction is the "Average Run Length (ARL)" notion.

The ARL can be used as a design criterion for control charts. ARL is the average (or expected) number of points that must be plotted before an out-of-control signal is given. If a process is in control, the ARL should be large, whereas if the process is out of control, the ARL should be small (Hogg, 1989). For the Shewhart-type charts, the ARL for an in-control process (Typically labeled $\mathrm{ARL}_o$) is given by

$$\mathrm{ARL}_o = \frac{1}{\alpha}, \tag{2.1}$$

where $\alpha$ is the false-alarm rate. For example, the $ARL_0$ for a $3\sigma$ Shewhart control chart applied to a process is $\frac{1}{0.0027} = 370.4$. In other words, for a stable in-control normally distributed process, we expect, on average, 370.4 plotted control points before an out-of-control signal is given.

If, however, the process has shifted, then the probability of an out-of-control signal increases as we would expect. The probability of not getting an out-of-control signal if the process has shifted is $\beta$. Thus, the probability of getting a signal is $1-\beta$; this probability is referred to as the power of the statistical procedure. It can be shown that the ARL for a Shewhart control chart when the process is out of control (typically labeled $ARL_1$) is

$$ARL_1 = \frac{1}{1-\beta}. \tag{2.2}$$

The ARL is a useful measure for the design of control charts. In particular, if one is interested in detecting mean shifts of a certain magnitude, the ARLs for different control charts can be computed and compared. Given two control charts that have the same in-control properties, the control chart which detects the out-of-control situation more quickly is more desirable. In terms of ARLs, for a given shift size, the control chart with a smaller out-of-control ARL is more desirable (Duncan, 1986).

## 2.3 The Structure of Control Charts

All control charts have a common structure. As shown in Figure 2.1, a control chart has a centerline representing the process average and upper and lower control limits that provide information on process variation (Shirland, 1993).

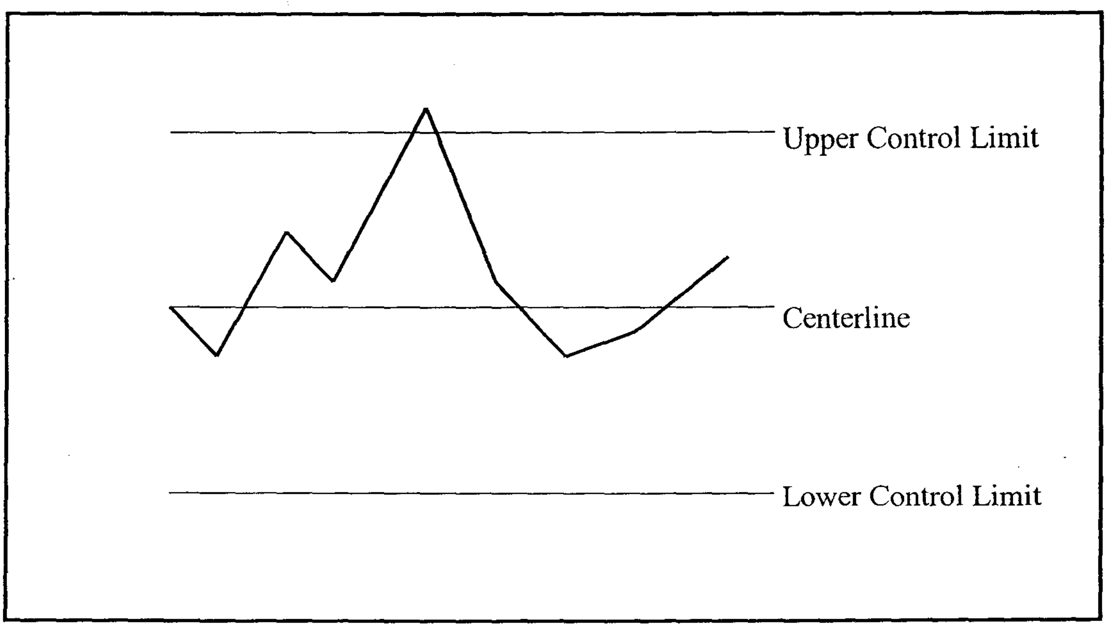


Figure 2.1 Structure of a Control Chart

Control charts are constructed by plotting samples, called subgroups, from a process. Control limits are based on the variation that occurs within the sampled subgroups. In this way, variation between the subgroups is intentionally excluded from the computation of the control limits; the common process variation becomes the variation on which we calculate the control limits. The computation of the control limits assume that there is no special cause of variation affecting the process. If a special cause of variation is present, a control chart, based solely on common variation, will highlight when and where the special cause has occurred. Consequently, a control chart makes possible the distinction between common and special variations and provides management and workers with a basis on which to take corrective actions on a process. The centerline of a control chart is taken to be the estimated mean of the sampling distribution (or the process average); the upper control limit is the mean plus three times the estimated standard deviation; and the lower control limit is the mean minus three times the estimated standard deviation. Subgroup means that behave nonrandom with respect to these control limits indicate of the presence of special causes of variation.

## 2.4 Uses of Control Charts

Control charts are statistical tools used to analyze and understand process variation. Control charts allow practitioners to understand the sources of variation in a process, and hence to manipulate and control those sources to reduce the difference between customers' needs and process performance. A desirable process is one that is stable and capable of improvement.

### 2.4.1 Stabilizing a Process with Control Charts

When the data consist of a series of fractions defective, the appropriate control chart is a $p$ chart. This is a depiction of the fraction of process data that has some attributes of interest.

When a practitioner determines that the causes of a variation are special causes, he or she should search for and eliminate the causes that may be attributable to a specific machine, worker, or group of workers, or a new batch of raw materials. After the special causes of variation have been identified and rectified, one will obtain a stable process that is in statistical control.

A stable process is a process that exhibits only common causes of variation resulting from inherent system limitations. The advantages of achieving a stable process are as follow:

1. Management knows the process capability and can predict performance, costs, and quality levels.
2. Productivity will be at the maximum, and costs will be minimized.
3. Management will be able to measure the effects of changes in the system with greater speed and reliability.

4. If management wants to alter specification limits, it will have the data to support its decision.

A stable process is a basic requirement for process-improvement efforts.

### 2.4.2 Improving a process with control charts

Once a process is stable, it has a known capability. A stable process may, nevertheless, produce an unacceptable number of defects (threshold state) and continue to do so as long as the system, as currently defined, remains the same. Management owns the system and must assume the ultimate responsibility for changing the system to reduce common variation and the difference between customer needs and process performance.

There are two areas for action to reduce the difference between customer needs and process performance. First, action may be taken to change the process average. This might include action to reduce the level of defects or process changes to increase production. Second, management can act to reduce the level of common causes of variation with the aim of achieving a never-ending improvement in a process. Procedures and inputs, such as the composition of the work force, training, supervision, materials, tools and machinery, and operational definitions, are the responsibilities of the management. The workers can only suggest changes; they cannot effect changes to the system.

## 2.5 Two Possible Mistakes in Using Control Charts

There are two different types of mistakes that the user of a control chart may make: over adjustment and under adjustment. Proper use of control charts will minimize the economic consequences of making either of these types of errors.

### 2.5.1 Over adjustment

The over adjustment error occurs when the user reacts to swings in the process data that are merely the result of common variation, such as adjusting a process downward if its past output is above average or adjusting a process upward if its past output is below average. In general, processes should not be adjusted on the basis of time-to-time observations but on the basis of information provided by a statistical control chart.

### 2.5.2 Under adjustment

Under adjustment, or lack of attention, results when a process is out-of-control and no effort is made to provide the necessary regulation. The process swings up and down in response to one or more special causes of variation, which may have compounding effects.

Avoiding both of these mistakes all of the time is an impossible task. That is, never adjusting the process – so that we never make the mistake of over adjusting – could result in several under adjustment. On the other hand, if we make very frequent adjustments to avoid the problem of under adjustment, we would probably be over adjusting. Control charts provide an economical means to minimize the loss that results from these two errors. Consequently, control charts provide management with information on when to take action on a process and when to leave it alone.

# CHAPTER 3

# VARIABLES CONTROL CHARTS

## 3.1 Introduction

A process in statistical control is not necessarily a good process which satisfies the required specifications. A process may be in statistical control, but the constant-cause variation may be so large that the process fails to meet the requirements that are demanded by the customer. Control charts provide information about how the process is running, not about how it should run.

Control charts help us detect unusual variation; they do not detect the reason. Special causes need to be detected and flagged, and control charts allow us to do this. However, we need to do more: Once we have identified unusual results, we need to find an explanation for them. Causes for the unusual events must be found through an investigation of the process. If the unusual measurements are undesirable, then one must make sure that the causes are prevented in the future. Of course, if they represent unusually good results, then one must make sure that the causes stay present.

Variables control charts are very important tools for process improvement. Variables control charts not only identify and differentiate between special and common causes of variation but also provide the data essential for process improvement.

$\overline{X}$ and $R$ charts use subgroups of size 2 to 10. When subgroup sizes are larger than 10, $\overline{X}$ and $S$ charts are used in place of $\overline{X}$ and $R$ charts. Both $\overline{X}$ and $R$ charts and $\overline{X}$ and $S$ charts are important tools in detecting variability. Variation of any kind must be continuously reduced if a firm aims to achieve process improvement.

## 3.2 The Design of Control Charts

The idea of a control chart is simple but ingenious. Shewhart recognized that it would be desirable to set limits based on the natural variation of any process so that variation within these limits is likely due to chance causes, but variation outside these limits would likely indicate a change in the process due to some special causes.

All control charts work on the basis that successive samples (also called subgroups) of a given size $n$ ($\geq 1$) are taken from a process at more or less regular intervals. From each of these samples, a number of sample statistics may be computed depending on the nature of the process measurements, for example, the sample mean, $\overline{X}$, the sample standard deviation, $S$ or the sample range, $R$. Different distribution parameters can potentially change over time; thus there is a need to look at different sample statistics, each dedicated to the monitoring of different aspects of the distribution (Alwan, 2000).

If the individual process measurements are generated from a random process with identical distribution over time, then any quantity based on the measurements will also be random with identical sampling distribution. Note that the distribution of the individual measurements and the distribution of the sample statistics are not necessarily the same. As an example, suppose that successive samples of size $n$ are drawn and the sample mean $\overline{X}$ is computed for each sample. If the individual measurements are randomly generated from a constant population over time with mean $\mu$ and standard deviation, $\sigma$, then the fluctuations of the sample means will be random and described by a constant sampling distribution over time with mean $\mu$ and standard deviation $\frac{\sigma}{\sqrt{n}}$. Furthermore, the central limit theorem tells us that this sampling distribution will be approximately (if not, exactly) normal, depending on how closely the underlying distribution for the individual measurements conforms to the normal distribution.

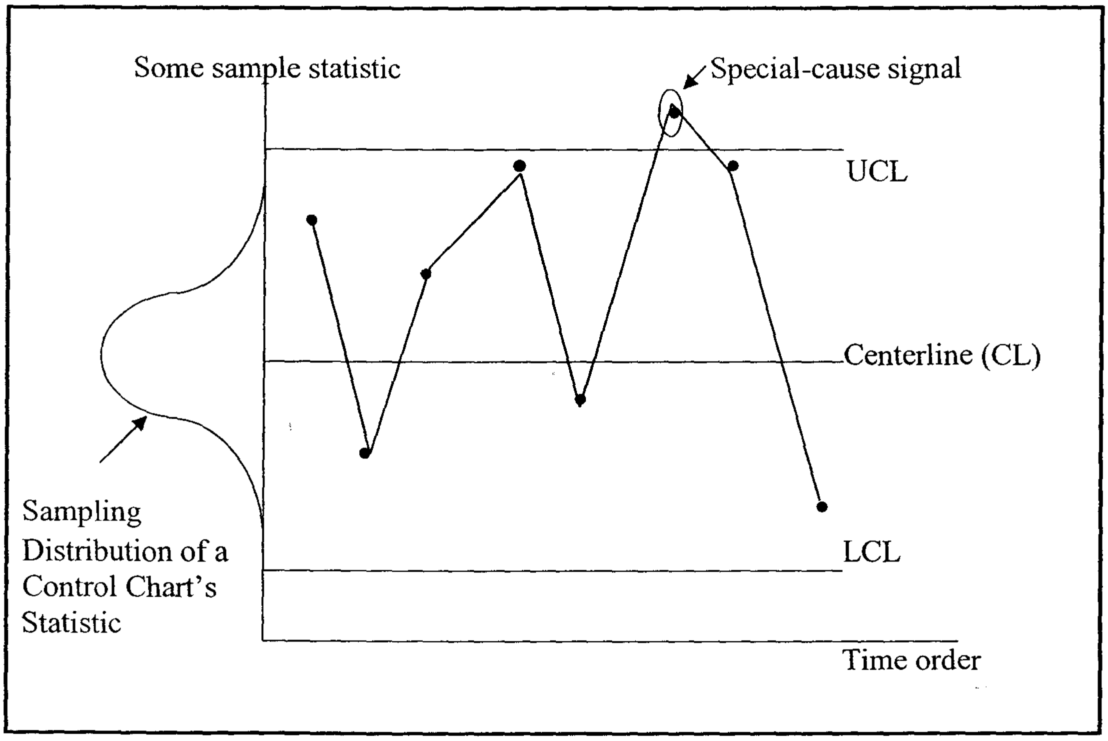


Figure 3.1 General structure of a Control Chart

The basic idea of a control chart is illustrated in Figure 3.1. A control chart displays process measures over time, whether individual readings or summary statistics of samples. The time period or sample number is conventionally laid out on the horizontal axis, with the data scale on the vertical axis. Superimposed on the time-series plot are typically three horizontal lines. One line is the Centerline (CL). Typically, the centerline represents the underlying mean of the sampling distribution for the sample statistic being plotted. In practice, this underlying mean is never known. Therefore, the centerline is taken as the average value of the control chart's observations.

The two other superimposed lines are called the Upper Control Limit (UCL) and Lower Control Limit (LCL). The idea is to place these limits so that if the process is in control, essentially all of the control chart's observations will fall within the limits. Given such a placement, it stands to reason that if any observation falls outside the control limits, there is reasonable evidence that a special cause is present. Most control charts are designed on the basis that not only is the sampling distribution symmetric but also it is well approximated by the normal distribution.

Normality is justified on the ground that it often arises from the cumulative effect of many small, independently acting, sources of variation. Since the dominant component of process variation, i.e., the common cause variation is viewed as a cumulative effect; normality would appear to be a reasonable assumption. Furthermore, due to the central limit theorem, many sample statistics are well approximated by the normal distribution, regardless of the underlying distribution for the individual measurements.

### 3.2.1 Mean and Standard Deviation of a Sample Mean

In statistical quality control, the data collected need to be described such that analysts can get a feel of the process or product characteristics. Such descriptions can be given through graphical methods or numerical measures. This section describes some of the commonly used numerical measures for deriving summary information from the observed values. Measures of central tendency tell something about the location of the observations and the value about which they are clustered. They help us to decide whether it is necessary to change the settings of process variables.

The mean is the average of the observations in a data set. In quality control, the mean is one of the most commonly used measures of estimation. It is easy to calculate and understand. The mean may be used to determine if, on the average, the process is operating around a desirable target value. The sample mean, or average (denoted by $\overline{X}$), is found by adding all the observations in a sample and dividing by the number of observations (or sample size, n) in that sample. If the $i$ th observation is denoted by $X_i$, then the sample mean is calculated as

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}. \tag{3.1}$$

The populations mean, μ is found by adding all of the data values in the population and dividing by the size of the population ($N$). It is calculated as

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N}. \tag{3.2}$$

The population mean is sometimes denoted as $E(X)$, the expected value of the random variable $X$. It is also called the mean of the probability distribution of $X$ (Mitra, 1993).

A widely used measure of dispersion in quality control is the range, which is the difference between the largest and smallest values in a data set. The range $R$ is defined as

$$R = X_L - X_S, \tag{3.3}$$

where $X_L$ is the largest observation and $X_S$ is the smallest observation.

The variance is a measure of the fluctuation of the observations around the mean. The larger the value, the greater the dispersion. The population variance, $\sigma^2$ is given by

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N}. \tag{3.4}$$

Here, $\mu$ is the population mean and $N$ represents the number of data points in the population. The sample variance $S^2$ is given by

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}, \tag{3.5}$$

where $\overline{X}$ represents the sample mean and $n$ is the number of observations in the sample. In most applications, the sample variance is calculated rather than the population variance because calculation of the latter is possible only by knowing every value in the population.

Observe that in calculating either the population variance or the sample variance, the deviation of each observation from the corresponding mean is obtained. A population variance $\sigma^2$ is a parameter, whereas a sample variance $S^2$ is an estimator, or a statistic. The value of $S^2$ may therefore change from sample to sample, whereas $\sigma^2$ should be constant. One desirable property of $S^2$ would be that even though it may not be equal to $\sigma^2$ for every sample, its value on the average would equal $\sigma^2$. This is known as the property of unbiasedness, where the mean or expected value of the estimator equals the corresponding parameter.

The standard deviation, like the variance, is a measure of the variability of the observations around the mean. It is equal to the positive square root of the variance. Thus, a standard deviation has the same unit of measurement as the observations and is easier to interpret than the variance. It is probably the most widely used measure of dispersion in quality control. The population standard deviation is given by

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}}. \qquad (3.6)$$

Similarly, the sample standard deviation $S$ is found as

$$S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}\,. \qquad (3.7)$$

As with the variance, the data set with the largest standard deviation will be identified as having the most variability about its average. If the probability distribution of a random variable is known, for instance, if it is normal distribution then information concerning the proportion of the observations that are within a certain number of standard deviations of the mean can be obtained.

### 3.2.2 How does a control chart works

A control chart is a graphical tool for monitoring the activity of an ongoing process. Control charts are sometimes referred to as Shewhart control charts, because Walter A. Shewhart first proposed their general theory. The value of the quality characteristic to be monitored is plotted along the vertical axis, while the horizontal axis represents the samples or subgroups (in order of time), from which the value of the quality characteristic is computed. Figure 3.2 shows a typical control chart.