Module – 2 Process Quality Improvement

What are the graphical and statistical techniques commonly used for understanding current state of quality? What are the process quality monitoring, control and improvement techniques?

Systematic solution approach to any quality improvement activity is critical and always emphasized by quality gurus (Juran, Deming, and Shewart). Various tools and techniques are commonly used to identify the critical control variables. The very basic techniques used in quality management is 7 QC Tools, which consist of Pareto Diagram, Process Flow Diagram, Cause and Effect Diagram, Check Sheets, Histogram, Run Charts, and Scatter diagram. Additional statistical tools used are hypothesis testing, regression analysis, ANOVA (Analysis of Variance), and Design of Experiment (DOE). In the following section, we will go through each and every technique in a greater detail.

7QC TOOLS

Pareto Diagram

Alfredo Pareto (1848-1923) conducted extensive studies of distribution of wealth in Europe. He found that there were a few people with a lot of money and majority of the people are having little money in their hand. This unequal distribution of wealth became an integral part of economic theory. Dr. Joseph Juran recognized this concept as a universal concept which can be applied to many other fields. He coined the phrase ‘vital few and useful many’.

A Pareto diagram is a graph that ranks data (on say types of defects) in descending order from left to right, as shown in Figure 2-2. In the diagram, data is classified as types of coating machines. Other possible data classifications include problem, complaints, causes, nonconformities types, and so forth. The vital few will come on the left of the diagram, and useful many are on the right. It is sometimes worthy to combine some of the useful many into one classification called "other". When this category is used, it is placed on the far right.

The vertical scale can be dollar value (or frequency), and percentage in each category is shown on top of each bar. In this case, Pareto diagrams were constructed for both frequency and dollar value. As can be seen from the figure, machine 35 has the greatest number of nonconformities,
but machine 51 has the greatest dollar value. Pareto diagrams can be distinguished from histograms (to be discussed) by the fact that horizontal scale of a Pareto diagram is categorical, whereas the scale for histogram is numerical or continuous.

**Figure 2-2** Simple Pareto Diagram

Pareto diagrams are used to identify the most important problem type. Usually, 75% of the problems are caused by 25% of the items. This fact is shown in the above figure, where coating machines 35 and 51 account for about 75% of the total non-conformities.

Actually, most important items could be identified by listing them in descending order. However, graph has an advantage of providing a visual impact, showing those vital few characteristics that need attention. Construction of a Pareto diagram is very simple. There are five steps involved:

**Step-1:** Determine method of classifying data: by problem, cause, nonconformity, and so forth.

**Step-2:** Decide if dollars (best), frequency, or both are to be used to rank the characteristics.

**Step-3:** Collect data for an appropriate time interval or use historical data.

**Step-4:** Summarize data and rank order categories from largest to smallest.

**Step-5:** Construct the diagram and find the vital few problem area.
The Pareto diagram is a powerful quality improvement tool to determine the most critical problem to be considered first. The diagram can also provide cumulative % information and given in many statistical software (say, MINITAB, [http://www.minitab.com/en-us/products/minitab/?WT.srch=1&WT.mc_id=SE001570](http://www.minitab.com/en-us/products/minitab/?WT.srch=1&WT.mc_id=SE001570)) as shown below.

**Figure 2-3** Pareto Diagram with Cumulative %

**Process Flow Diagram**

For many products and services, it may be useful to construct a process flow diagram. **Figure 2-4** shows a simple process flow diagram for order entry activity of a make-to-order company that manufactures gasoline filling station hose nozzles. These diagrams show flow of product or service as it moves through various processing stages. The diagram makes it easy to visualize the entire multistage process, identify potential trouble spots, waste activities, and locate control points. It answers the question, "Who is our next customer?" Improvements can be accomplished by changing (reengineering), reducing, combining, or eliminating process steps.
Standardized symbols may be used as recommended by industrial engineering and Lean Management (Value Stream Mapping, text book. In Six Sigma methodology, process mapping is done by SIPOC (Suppliers-Inputs-Process-Outputs-Customer).

**Cause and Effect Diagram**
A cause-and-effect (C&E) diagram is a picture composed of lines and symbols designed to represent a meaningful relationship between an effect (say Y) and its potential causes (say X). Potential causes (which have evidence) are not all possible causes that come up in brainstorming exercise. It was developed by Dr. Kaoru Ishikawa in 1968, and sometimes referred to as the ‘Ishikawa diagram’ or a ‘fish bone diagram’. C&E diagram is used to investigate either a "bad" effect and to take action to rectify the potential causes or a "good" effect and to learn those potential causes that are responsible for the effect. For every effect, there are likely to be numerous potential causes. **Figure 2-5** illustrates a simple C&E diagram with effect on right and causes on left. Effect is the quality characteristic that needs improvement. Causes are sometimes broken down into major sub causes related to work method, material, measurement, man (people), machinery (equipment), and environment (5M &
It is not necessary that every diagram will always have 5M and 1 E cause and can depend also on the problem type. There can be other major causes in case of service-type problem. Each major cause is further subdivided into numerous sub causes. For example, under work methods, we might have training, knowledge, ability, physical characteristics, and so forth. C&E diagrams are the means of picturing all these major and sub causes. The identified potential causes considered critical (say 1, 2, 3, 4 and 5 as given in the below diagram) may be further explored by experimentation to understand their impact on the house paint.

**Cause-and-effect Diagram of house Paint Peeling**

**Figure 2-5:** A Simple Cause and Effect Diagram

C & E diagrams are useful to

1) Identify potential causes and not all possible causes,
2) Analyze actual conditions for the purpose of product or service quality improvement
3) Eliminate conditions which cause nonconformities and customer complaints.
4) Statistical Experimentation, Decision-making and corrective-action activities.

C&E diagram can also be generated in MINITAB as shown below.
Check Sheets

Main purpose of check sheets in earlier days is to ensure that data was collected carefully and accurately by concerned personnel. Data is to be collected in such a manner that it can be quickly and easily used and analyzed. The form of check sheet is individualized for each situation and is designed by the project team. Figure 2-7 shows a check sheet for paint nonconformities for bicycles.

Check sheets can also be designed to show location of defects. For example, check sheet for bicycle paint non conformities could show an outline of a bicycle, with ‘X’ indicating location of nonconformities. Creativity plays a major role in design of a check sheet. It should be user-friendly and, whenever possible, include information on location.

Figure 2-6: Cause & Effect Diagram in MINITAB
Figure 2-7: A Typical Check Sheet

Histogram
Histogram provides variation information of the characteristic of interest, as illustrated by Figure 2-8. It suggests probability distribution shape of the sample observation and also indicates possible gap in the data. Horizontal axis in Figure 2-8 indicates scale of measurement and vertical axis represents frequency or relative frequency.
Histograms have certain identifiable characteristics, as shown in Figure 2-8. One characteristic of the distribution concerns symmetry or lack of symmetry of the data. Is the data equally distributed on each side of the center of measurement (e.g. Temperature), or it is skewed to right or left? Another characteristic concerns the kurtosis of the data. A final characteristic concerns number of modes, or peaks, in the data. There can be one mode, two modes (bi-modal) or multiple modes.

Histograms can also provide sufficient information about a quality problem to provide a basis for decision making without statistical analysis. They can also be compared in regard to location, spread, and shape. A histogram is like a snapshot of the process showing variation in the characteristic. Histograms can determine process capability, compare with specifications, suggest shape of the population, and indicate any discrepancies in the data. A typical histogram using MINITAB software is shown below.
Run Charts

A run chart, which is shown in Figure 2-10 is a very simple quality tool for analyzing process with respect to (w.r.t) time in development stage or, for that matter, when other charting techniques are not quite relevant. The important point is to draw a picture of the process w.r.t. time and let it "talk" to you. Plotting time oriented data points is a very effective way of highlighting any pattern observed w.r.t. time. This type of plotting should be done before doing histogram or any other statistical data analysis.
The horizontal axis in **Figure 2-10** is labeled as time (Day of the Week), and vertical axis of the graph represents measurement on variables of interest.

**Scatter Diagram**

The simplest way to determine if a relationship exists between TWO variables is to plot a scatter diagram. **Figure 2-11** shows a relationship between automotive speed and gas mileage. The figure indicates that as speed increases, gas mileage decreases or a negative relationship exist between the variables of interest. Automotive speed is plotted on x-axis and so-called independent variable. The independent variable is usually controllable. Here, gas mileage is on the y-axis and is the dependent or so-called response variable.

![Figure 2-11 Scatter Diagram](image

There are a few simple steps for constructing a scatter diagram. Data is collected as ordered pairs (x, y). The automotive speed is controlled and the gas mileage is measured. Horizontal and vertical scales are constructed with higher values on right for x-axis and on the top for y-axis. After the scales are labeled, data is plotted. Once the scatter diagram is complete, relationship or Pearson correlation (http://en.wikipedia.org/wiki/Correlation_coefficient) between two variables can be found out. In MINITAB, all relevant information can be derived using scatter plot
(GRAPH-Scatter Plot)/correlation/regression option and shown with an example in Figure 2-12, and Figure 2-13.

![Scatter Plot](image)

**Figure 2-12** Scatter Plot in MINITAB with positive correlation of 0.98

![Scatter Plot with Week or No Relationship](image)

**Figure 2-13** Scatter Plot in MINITAB with week correlation of 0.1

Week correlation will not imply ‘no’ relationship. There may be nonlinear relationship, which is not reflected by Pearson Correlation Coefficient. Few other graphical plots extensively used in Quality Data analysis are Control Chart, and Box Plot. These are discussed below.
Control Chart

Quality control is one approach that any organization adopts to detect defects and to take corrective actions. Quality control is employed to ensure the desired level of quality in the final goods and services. Quality control is about analysis of data for rectification of errors with respect to time. Walter Shewhart developed the control charts in 1924. It focuses on monitoring the performance of characteristic of interest over a period of time by looking at the variability in the data. There are two broad categories of control charts: control charts for attributes and control charts for variables. A variable control chart consists of a centre line (CL) that represents the mean value of the characteristic of interest. In addition, two other horizontal lines, namely the Upper Control Limit (UCL) and the Lower Control Limit (LCL), are also shown in the control chart. A typical variable control chart on mean and range of a characteristic, so-called X-bar and R is shown below.

![X-bar and R Control Chart for Paint Thickness characteristic](image)

**Figure 2-14** A Variable Control Chart

Sample mean (\( \bar{x} \)) chart monitors the accuracy and central tendency of the process output characteristic. Whereas, the sample range (R) chart monitor the variation of the characteristic, with respect to time. The calculation details on UCL, CL and LCL can be found in any Quality Management Book (Mitra, A, 2008; Montgomery, D.C., 2008). In attribute type of control chart (used to monitor number of defects or defectives), only
one chart is used to monitor deviation with respect to time. More details on control chart are given in below section on statistical technique.

**Box Plot**

Box plot provide a display on quartiles, and outliers for a given data set. If we need to compare variation of two data set (say two different service time), we may need the help of Box plot at initial stage before going into inferential statistics and hypothesis testing. A typical comparative Box Plot for two fast food restaurant is shown below.

![Box Plot of Two Restaurant Service Time](image)

**Figure 2-15** A Comparative Box Plot

Each box in the graph shows first quartile, second quartile and third quartile. The extension line (Whisker) beyond the box is minimum of 1.5*(Inter quartile Range) and extreme data point. ‘∗’ beyond the whisker is considered as outlier.

It is observed that although the service time median of Macnalds and KYC seems close, the variability of KYC data is much more than Macnalds. Thus it seems Macnalds is more consistent on service time than KYC.

In addition to the above chart, stem and leaf plot ([http://www.youtube.com/watch?v=cOl-d3BERkM](http://www.youtube.com/watch?v=cOl-d3BERkM)) and Multi-vari Chart ([http://en.wikipedia.org/wiki/Multi-vari_chart](http://en.wikipedia.org/wiki/Multi-vari_chart)) are also useful in certain situations.
This two are not discussed in detail and can be found in literatures, books, and web. In process quality improvement, not only the 7 QC tools and plots are important, but few statistical techniques are extensively used for inferential statistics and decision making. Few of them are discussed below.

**Statistical Techniques**

Few important statistical techniques, frequently used in quality improvement and decision making, include hypothesis testing, regression analysis, sampling technique, two sample t-test, Analysis of Variance (ANOVA), and Design of Experiment (DOE). These techniques are discussed briefly below.

**HYPOTHESIS TESTING**

Population parameters (say mean, variance) of any characteristic which are of relevance in most of the statistical studies are rarely known with certainty and thus estimated based on sample information. Estimation of the parameter can be a point estimate or an interval estimate (with confidence interval). However, many problems in engineering science and management require that we decide whether to accept or reject a statement about some parameter(s) of interest. The statement which is challenged is known as a null hypothesis, and the way of decision-making procedure is so-called hypothesis testing. This is one of the most useful techniques for statistical inference. Many types of decision-making problems in the engineering science can be formulated as hypothesis-testing problems. If an engineer is interested in comparing mean of a population to a specified value. These simple comparative experiments are frequently encountered in practice and provide a good foundation for the more complex experimental design problems that will be discussed subsequently. In the initial part of our discussion, we will discuss comparative experiments involving either one or two populations, and our focus is on testing hypothesis concerning the parameters of the population(s). We now give a formal definition of a statistical hypothesis.
Definition

A statistical hypothesis is a statement about parameter(s) of one or more populations.

For example, suppose that we are interested in burning rate of a solid propellant used to power aircrew escape systems. Now burning rate is a random variable that can be described by a probability distribution. Suppose that our interest focuses on mean burning rate (a parameter of this distribution). Specifically, we are interested in deciding whether or not the mean burning rate is 60 cm/s. We may express this formally as

\[ H_0 : \mu = 60 \text{ cm/s} \]
\[ H_1 : \mu \neq 60 \text{ cm/s} \]

The statement \( H_0 : \mu = 60 \text{ cm/s} \) is called the null hypothesis, and the statement \( H_1 : \mu \neq 60 \text{ cm/s} \) is so-called the alternative hypothesis. Since alternative hypothesis specifies values of \( \mu \) that could be either greater or less than 60 cm/s, it is called a two-sided alternative hypothesis. In some situations, we may wish to formulate a one-sided alternative hypothesis, as

\[ H_0 : \mu = 60 \text{ cm/s} \quad \text{Or} \quad H_0 : \mu = 60 \text{ cm/s} \]
\[ H_1 : \mu < 60 \text{ cm/s} \quad H_1 : \mu > 60 \text{ cm/s} \]

It is important to remember that hypotheses are always statements about the population or distribution under study, AND not statements about the sample. An experimenter generally believes the alternate hypothesis to be true. Hypothesis-testing procedures rely on using information in the random sample from the population of interest. Population (finite or infinite) information is impossible to collect. If this information is consistent with the null hypothesis then we will conclude that the null hypothesis is true: however if this information is inconsistent with null hypothesis, we will conclude that there is little evidence to support null hypothesis.

The structure of hypothesis-testing problems is generally identical in all engineering/science applications that are considered. Rejection of null hypothesis always leads to accepting alternative hypothesis. In our treatment of
hypothesis testing, null hypothesis will always be stated so that it specifies an exact value of the parameter (as in the statement \( H_0 : \mu = 60 \text{ cm/s} \)). The alternative hypothesis will allow the parameter to take on several values (as in the statement \( H_1 : \mu \neq 60 \text{ cm/s} \)). Testing the hypothesis involves taking a random sample, computing a test statistic from the sample data, and using the test statistic to make a decision about the null hypothesis.

**Testing a Statistical Hypothesis**

To illustrate the general concepts, consider the propellant burning rate problem introduced earlier. The null hypothesis is that the mean burning rate is 60 cm/s, and the alternative is that it is not equal to 60 cm/s. That is, we wish to test

\[
H_0 : \mu = 60 \text{ cm/s} \\
H_1 : \mu \neq 60 \text{ cm/s}
\]

Suppose that a sample of \( n=10 \) specimens is tested and that the sample mean burning rate \( \bar{x} \) is observed. The sample mean is an estimate of the true population mean \( \mu \). A value of the sample mean \( \bar{x} \) that falls close to the hypothesized value of \( \mu = 60 \text{ cm/s} \) is evidence that the true mean \( \mu \) is really 60 cm/s: that is, such evidence supports the null hypothesis \( H_0 \). On the other hand, a sample mean that is considerably different from 60 cm/s is evidence in support of the alternative hypothesis, \( H_1 \). Thus sample mean is the test statistic in this case.

Varied sample may have varied mean values. Suppose that if \( 58.5 \leq \bar{x} \leq 61.5 \), we will accept the null hypothesis \( H_0 : \mu = 60 \), and if either \( \bar{x} < 58.5 \) or \( 58.5 \bar{x} > 61.5 \), we will accept the alternative hypothesis \( H_1 : \mu \neq 60 \). The values of \( \bar{x} \) that are less than 58.5 and greater than 61.5 constitute the rejection region for the test, while all values that are in the interval 58.5 to 61.5 forms acceptance region. Boundaries between critical regions and acceptance region are so-called ‘critical values’. In our example the critical values are 58.5 and 61.5. Thus, we reject \( H_0 \) in favor of \( H_1 \) if the test statistic falls in the critical region and accept \( H_0 \) otherwise. This interval or region of acceptance is defined based on the concept of confidence interval and level of
significance for the test. More details on confidence interval and level of significance can be found in various web link (http://en.wikipedia.org/wiki/Confidence_interval; http://www.youtube.com/watch?v=iX0bKAeLbDo) and book (Montgomery and Runger, 2010).

Hypothesis decision procedure can lead to either of two wrong conclusions. For example, true mean burning rate of the propellant could be equal to 60 cm/s. However, for randomly selected propellant samples that are tested, we could observe a value of the test statistic, $X$, that falls into the critical region. We would then reject the null hypothesis $H_0$ in favor of the alternative, $H_1$, when, in fact $H_0$ is really true. This type of wrong conclusion is called a **Type I error**. This is a more serious mistake as compared to another error Type II explained below.

Now suppose that true mean burning rate is different from 60 cm/s, yet sample mean $\bar{X}$ falls in the acceptance region. In this case we would accept $H_0$ when it is false. This type of wrong conclusion is called a **Type II error**.

Thus, in testing any statistical hypothesis, four different situations determine whether final decision is correct or error. These situations are presented in **Table 2-1**.

Because our decision is based on random variables, probabilities can be associated with the Type I and Type II errors. The probability of making a Type I error is denoted by Greek letter $\alpha$. That is,

$$\alpha = P(\text{Type I error}) = P(\text{Reject } H_0 \mid H_0 \text{ is true})$$

Sometimes the Type I error probability is called the significance level ($\alpha$) or size of the test.

<table>
<thead>
<tr>
<th>Actual</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$ is true</td>
<td>Accept $H_0$</td>
</tr>
<tr>
<td>$H_0$ is false</td>
<td>Reject $H_0$</td>
</tr>
</tbody>
</table>

**Table 2-1** Type I and Type II Error

The steps followed in hypothesis testing are
(i) Specify the Null and Alternate hypothesis
(ii) Define level of significance($\alpha$) based on criticality of the experiment
(iii) Decide the type of test to be used (left tail, right tail etc.)
(iv) Depending on the test to be used, sample distribution, mean variance information, define the appropriate test statistic (z-test, t-test etc.)
(v) Considering the level of significance ($\alpha$), define the critical values by looking into standard statistical tables.
(vi) Decide on acceptance or rejection of null/ alternate hypothesis by comparing test statistic values calculated from samples, and as defined in step (iv), with standard value specified in statistical table.
(vii) Derive meaningful conclusion.

**Regression Analysis**

In many situations, two or more variables are inherently related, and it is necessary to explore nature of this relationship (linear or nonlinear). Regression analysis is a statistical technique for investigating the relationship between two or more variables. For example, in a chemical process, suppose that yield of a product is related to process operating temperature. Regression analysis can be used to build a model (response surface) to predict yield at a given temperature level. This response surface can also be used for further process optimization, such as finding the level of temperature that maximizes yield, or for process control purpose.

Let us look into **Table 2-2** with paired data collected on % Hydrocarbon levels (say x variable) and corresponding Purity % of Oxygen (say, y variable) produced in a chemical distillation process. The analyst is interested to estimate and predict the value of y for a given level of x, within the range of experimentation.
Table 2-2  Data Collected on % Hydrocarbon levels (x) and Purity % of Oxygen (y)

<table>
<thead>
<tr>
<th>Observation number</th>
<th>Hydrocarbon level x (%)</th>
<th>Purity y (%)</th>
<th>Observation number</th>
<th>Hydrocarbon level x (%)</th>
<th>Purity y (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.99</td>
<td>90.1</td>
<td>11</td>
<td>1.19</td>
<td>93.54</td>
</tr>
<tr>
<td>2</td>
<td>1.0</td>
<td>89.05</td>
<td>12</td>
<td>1.15</td>
<td>92.52</td>
</tr>
<tr>
<td>3</td>
<td>1.15</td>
<td>91.5</td>
<td>13</td>
<td>0.97</td>
<td>90.56</td>
</tr>
<tr>
<td>4</td>
<td>1.29</td>
<td>93.74</td>
<td>14</td>
<td>1.01</td>
<td>89.54</td>
</tr>
<tr>
<td>5</td>
<td>1.44</td>
<td>96.73</td>
<td>15</td>
<td>1.11</td>
<td>89.85</td>
</tr>
<tr>
<td>6</td>
<td>1.36</td>
<td>94.45</td>
<td>16</td>
<td>1.22</td>
<td>90.39</td>
</tr>
<tr>
<td>7</td>
<td>0.87</td>
<td>87.59</td>
<td>17</td>
<td>1.26</td>
<td>93.25</td>
</tr>
<tr>
<td>8</td>
<td>1.23</td>
<td>91.77</td>
<td>18</td>
<td>1.32</td>
<td>93.41</td>
</tr>
<tr>
<td>9</td>
<td>1.55</td>
<td>99.42</td>
<td>19</td>
<td>1.43</td>
<td>94.98</td>
</tr>
<tr>
<td>10</td>
<td>1.4</td>
<td>93.65</td>
<td>20</td>
<td>0.95</td>
<td>87.33</td>
</tr>
</tbody>
</table>

Scatter diagram is a firsthand visual tool to understand the type of relationship, and then regression analysis is recommended for developing any prediction model. Inspection of this scatter diagram (given in Figure 2-16) indicates that although no simple curve will pass exactly through all the points, there is a strong trend indication that the points lie scattered randomly along a straight line.

![Scatter Plot and Regression](scatter_plot.png)

**Figure 2-16** Scatter diagram and Trend of Relationship
Therefore, it is probably reasonable to assume that mean of the random variable \( Y \) is related to \( x \) by following straight-line relationship. This can be expressed as

\[
E(Y \mid x) = \mu_{y|x} = \beta_0 + \beta_1 x
\]

where, regression coefficient \( \beta_0 \) is so-called intercept and \( \beta_1 \) is the slope of the line. Slope and intercept is calculated based on a ordinary least square method. While the mean of \( Y \) is a linear function of \( x \), the actual observed value \( y \) does not fall exactly on a straight line. The appropriate way to generalize this to a probabilistic linear model is to assume that the expected value of \( Y \) is a linear function of \( x \). But for a particular value of \( x \), actual value of \( Y \) is determined by mean value from the linear regression model plus a random error term,

\[
Y = \beta_0 + \beta_1 x + \varepsilon
\]

where \( \varepsilon \) is the random error term. We call this model as simple regression model, because it has only one independent variable(\( x \)) or regressor. Sometimes a model like this will arise from a theoretical relationship. Many times, we do not have theoretical knowledge of the relationship between \( x \) and \( y \) and the choice of the model is based on inspection of a scatter diagram, such as we did with the oxygen purity data. We then think of the linear regression model as an empirical model with uncertainty (error).

The regression option in MINITAB can be used to get all the results of the model. The results as derived from MINITAB-REGRESSION option using oxygen purity data set is provided below.

The regression equation is

\[
\% \text{ Purity (y)} = 74.5 + 14.8 \% \text{ Hydrocarbon Level(x)}
\]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>74.494</td>
<td>1.666</td>
<td>44.73</td>
<td>0.000</td>
</tr>
<tr>
<td>% Hydrocarbon Level(x)</td>
<td>14.796</td>
<td>1.378</td>
<td>10.74</td>
<td>0.000</td>
</tr>
</tbody>
</table>

\( S = 1.13889 \quad \text{R-Sq = 86.5\%} \quad \text{R-Sq(adj) = 85.7\%} \)

**Analysis of Variance**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>149.55</td>
<td>149.55</td>
<td>115.30</td>
<td>0.000</td>
</tr>
</tbody>
</table>
We look into R-Sq value and if it is more than 70%, we assume the relationship is linear and y depends on x. More details on interpretation of other values are given in MINITAB help menu or readers can refer to any standard statistical or quality management book.

In this context, P-value and its interpretation is important from the context of hypothesis testing and regression analysis. General interpretation is that if the P-value is less than 0.05 (at 5% level of significance test) the NULL HYPOTHESIS is to be rejected. Reader may refer web (http://www.youtube.com/watch?v=lm_CagZXcv8; http://www.youtube.com/watch?v=TWmdzwAp88k) for more details on P-value and its interpretation.

**Common Abuses of Regression**

Regression is widely used and frequently misused. Care should be taken in selecting variables with which to construct regression equations and in determining form of a model. It is possible to develop statistical relationships among variables that are completely unrelated in a practical sense, For example, we might attempt to relate shear strength of spot welds with number of boxes of computer paper used by information systems group. A straight line may even appear to provide a good fit to the data, but the relationship is an unreasonable one. A strong observed association between variables does not necessarily imply that a causal relationship exists between those variables. Designed experimentation is the only way to prove causal relationships.

Regression relationships are also valid only for values of the regressor variable within range of original experimental/actual data. But it may be unlikely to remain so as we extrapolate. That is, if we use values of x beyond the range of observations, we become less certain about the validity of the assumed model and its prediction.
**Process quality monitoring and control**

Process quality is monitored by using acceptance sampling technique and control is achieved through Statistical process control (SPC) chart. Monitoring is essential so as to utilize full potential of the process. Statistical quality control is dating back to the 1920s. Dr. Walter A. Shewhart of the Bell Telephone Laboratories was one of the early pioneers of the field. In 1924 he wrote a memorandum showing a modern control chart, one of the basic tools of statistical process control. Dr. W. Edwards Deming and Dr. Joseph M. Juran have been instrumental in spreading statistical quality-control methods since World War II.

In any production process, regardless of how well-designed or carefully maintained it is, a certain amount of inherent or natural variability will always exist. This natural variability or "noise" is the cumulative effect of many small, essentially unavoidable causes. When the noise in a process is relatively small, we usually consider it an acceptable level of process performance. In the framework of statistical quality control, this natural variability is often called “chance cause variability". A process that is operating with only chance causes of variation is said to be in statistical control. In other words the chance causes are an inherent part of the process.

Other kinds of variability may occasionally be present in the output of a process. This variability in key quality characteristics usually arises from sources, such as improperly adjusted machines, operator errors, or defective raw materials. Such variability is generally large when compared to the background noise and it usually represents an unacceptable level of process performance. We refer these sources of variability that are not part of the chance cause pattern as assignable causes. A process that is operating in the presence of assignable causes is said to be out-of-control. There are varied statistical control chart to identify out of control signal. Typically any control chart will have an upper and lower control limit and a central line as given in Figure 2-17.
There is a close connection between control charts and hypothesis testing. Essentially the control chart is a test of the hypothesis that the process is in a state of statistical control.

There can be attribute (say monitoring defects or defective) control chart and variable control chart. Variable control chart example is given earlier. A ‘c’ attribute chart monitors defects. A ‘c’ chart is shown below, where number of defect data in engine assembly are collected over period of time. Thus at a particular time a sample engine assembly is selected and number of defects in the assembly is recorded and monitored.
Figure 2-18 A c-type Attribute Control Chart and limit lines

P-chart is used to monitor defectives.

Details on various types of control chart and their specific application in varied situation can be seen in many well known text books (Mitra, A, 2008; Montgomery, D.C., 2008) or web (http://www.youtube.com/watch?v=gTxaQkuv6sU).

The principles of control chart are based on acceptance sampling plans, provided by Harold F. Dodge and Harry G. Romig, who are employees of Bell System. Acceptance sampling plan is discussed in the following section.

Acceptance sampling plan

Acceptance sampling is concerned with inspection and decision making regarding product quality. In 1930’s and 1940’s, acceptance sampling was one of the major components of quality control, and was used primarily for incoming or receiving inspection.

A typical application of acceptance sampling is as follows: A company receives a shipment of product from its vendor. This product is often a component or raw material used in company's manufacturing process. A sample is taken from the lot, and some quality characteristic of the units in the sample is inspected referring to specification. On the basis of information in this sample, a decision is made regarding acceptance or rejection of the whole lot. Sometimes we refer
to this decision as lot sentencing. Accepted lots are put into production; rejected lots are returned to the vendor or may be subjected to some other lot-disposition action.

Although it is customary to think of acceptance sampling as a receiving inspection activity, there are other uses of sampling methods. For example, frequently a manufacturer will sample and inspect its own product at various stages of production. Lots that are accepted are sent forward for further processing, and rejected lots may be reworked or scrapped.

Three aspects of sampling include:

1) Purpose of acceptance sampling is to take decision on acceptance of lots, not to estimate the lot quality. Acceptance-sampling plans do not provide any direct form of quality control.

2) Acceptance sampling simply accepts and rejects lots. This is a post mortem kind of activity. Statistical process controls are used to control and systematically improve quality by reducing variability, but acceptance sampling is not.

3) Most effective use of acceptance sampling is not to "inspect quality into the product," but rather as an audit tool to ensure that output of a process conforms to requirements.

**Advantages and Limitations of Sampling Plan as compared to 100 % inspection**

In comparison with 100% inspection, acceptance sampling has following advantages.

(i) It is usually less expensive because there is less inspection.

(ii) There is less handling of product, hence reduced damage.

(iii) It is highly effective and applicable to destructive testing.

(iv) Fewer personnel are involved in inspection activities.

(v) It often greatly reduces amount of inspection error.

(vi) Rejection of entire lots as opposed to the simple return of defectives provides a stronger motivation to suppliers for quality improvement.

Acceptance sampling also has several limitations which include:

(i) There is risk of accepting "bad" lots and rejecting "good" lots, or Type I and Type II error.
(ii) Less information is usually generated about the product.

Types of Sampling Plans

There are a number of ways to classify acceptance-sampling plans. One major classification is by attributes and variables. Variables are quality characteristics that are measured on a numerical scale whereas attributes are quality characteristics are expressed on a "go, no-go" basis.

A single-sampling plan is a lot-sentencing procedure in which one sample units is selected at random from the lot, and disposition of the lot is determined based on information contained in that sample. For example, a single-sampling for attributes would consist of a sample size $n$ and an acceptance number $c$. The procedure would operate as follows: Select $n$ items at random from the lot. If there are fewer than $c$ defectives in the sample, accept the lot, and if there are more than $c$ defective in the sample, reject the lot.

Double-sampling plans are somewhat more complicated. Following an initial sample, a decision based on the information in that sample is made either to accept the lot, reject the lot or to take a second sample. If the second sample is taken, information from both first and second sample is combined in order to reach a decision whether to accept or reject the lot.

A multiple-sampling plan is an extension of the double-sampling concept, in that more than two samples may be required in order to reach a decision regarding the disposition of the lot. Sample sizes in multiple sampling are usually smaller than they are in either single or double sampling. The ultimate extension of multiple sampling is sequential sampling, in which units are selected from the lot one at a time, and following inspection of each unit, a decision is made either to accept the lot, reject the lot, or select another unit.

Random Sampling
Units selected for inspection from the lot should be chosen at random, and they should be representative of all the items in the lot. Random-sampling concept is extremely important in acceptance sampling and statistical quality control. Unless random samples are used, bias may be introduced. Say, suppliers may ensure that units packaged on the top of the lot are of extremely good quality, knowing that inspector will select sample from the top layer. This also helps in identifying any hidden factor during experimentation.

The technique often suggested for drawing a random sample is to first assign a number to each item in the lot. Then $n$ random numbers are drawn (from random number table or using excel/statistical software), where the range of these numbers is from 1 to the maximum number of units in the lot. This sequence of random numbers determines which units in the lot will constitute a sample. If products have serial or other code numbers, these numbers can be used to avoid process of actually assigning numbers to each unit. Details on different sampling plan can be seen in Mitra, A (2008).

**Process improvement Tools**

Acceptance sampling and statistical quality control techniques may not significantly reduce variability in the output. Process improvement by variation reduction is an important feature in quality management. There are varieties of statistical tools available for improving processes. Some of them are discussed below.

**ANOVA**

Many experiments involve more than two levels of a factor. Experimenter is interested to understand the influence of the factor on variability of output characteristic. In this case, Analysis of variance (ANOVA) is the appropriate statistical technique. This technique is explained with the help of following example.

Say, a product development engineer is interested in investigating tensile strength of a new synthetic fiber. The engineer knows from previous experience that the strength is affected by weight percent of cotton used in the blend of materials for the fiber. Furthermore, he suspects that increasing cotton content will increase the strength, at least initially. He also knows that cotton
content should range of 1 to 25 percent if final product is to have other quality characteristics that are desired. Engineer decides to test specimens at five levels of cotton weight percent: 5, 10, 15, and 20 percent. She also decides to test five specimens at each level of cotton content. This is an example of a single-factor (Cotton Weight %) experiment with a (level) = 5 of the factor and \( n = 5 \) replicates. The 25 runs should be made in random sequence.

**Table 2-3 Experimental Data**

<table>
<thead>
<tr>
<th>Cotton weight percentage</th>
<th>Experimental run number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
</tr>
</tbody>
</table>

The randomized test sequence is necessary to prevent the effects of unknown nuisance variables (or hidden factor influence), perhaps varying out of control during experiment, from contaminating the results. Balanced experiments with equal number of replicates are also preferred to minimize the experimental error.

This is so-called a single-factor analysis of variance model and known as fixed effects model as the factor level can be changed as required by the experimenter. Recall that \( y_{ij} \), represents the \( j \)th sample (\( j=1,..n \), replicate) output observations under \( i \)th treatment combination. Let \( \bar{y}_i \) represent the average of the observations under the \( i \)th treatment. Similarly, let \( \bar{y}_\cdot \) represent the grand total of all the observations and represent the grand average of all the observations. Expressed symbolically,

\[
y_i = \sum_{j=1}^{a} y_{ij} \quad \bar{y}_i = y_i/n \quad i=1,2,...,a
\]

\[
y_\cdot = \sum_{i=1}^{a} \sum_{j=1}^{a} y_{ij} \quad \bar{y}_\cdot = y_\cdot/N
\]

where, \( N \) or \( a \ast n \) is the total number of observations. The "dot" subscript notation used in above equations implies summation over the subscript that it replaces.

The appropriate hypotheses are
\[ H_0 : \mu_1 = \mu_2 = \cdots = \mu_n = 0 \]
\[ H_1 : \mu_i \neq \mu_j \] for at least one pair \((i, j)\)
Decomposition of the total sum of squares

The name analysis of variance is derived from a partitioning of total variability into its component parts. The total corrected sum of squares

\[ SS_T = \sum_{i=1}^{a} \sum_{j=1}^{n} \left( y_{ij} - \bar{y}_i \right)^2 \]

is used as a measure of overall variability in the data. Intuitively, this is reasonable because, if we were to divide SS, by the appropriate number of degrees of freedom (in this case, \( N - 1 \)), we would have the sample variance of the \( y \)'s. The sample variance is of course a standard measure of variability.

Note that the total corrected sum of squares \( SS_T \) may be written as

\[ SS_T = \sum_{i=1}^{a} \sum_{j=1}^{n} \left( y_{ij} - \bar{y}_i \right)^2 = \sum_{i=1}^{a} \sum_{j=1}^{n} \left[ \left( \bar{y}_i - \bar{y} \right) \left( y_{ij} - \bar{y}_i \right) \right]^2 \]

or,

\[ SS_T = \sum_{i=1}^{a} \sum_{j=1}^{n} \left( y_{ij} - \bar{y}_i \right)^2 = \sum_{i=1}^{a} \left( \bar{y}_i - \bar{y} \right)^2 + \sum_{i=1}^{a} \sum_{j=1}^{n} \left( y_{ij} - \bar{y}_i \right)^2 + 2 \sum_{i=1}^{a} \sum_{j=1}^{n} \left( \bar{y}_i - \bar{y} \right) \left( y_{ij} - \bar{y}_i \right) \]

and as it is proved that the third product term vanishes, we can rewrite the overall expression of \( SS_{\text{Total}} \) \( (SS_T) \) as

\[ SS_T = SS_{\text{Treatments}} + SS_E \]

Where, \( SS_{\text{Treatments}} \), is so-called the sum of squares due to treatments (i.e., between treatments), and \( SS_{\text{Error}} \), is called the sum of squares due to error (i.e., within treatments). There are an \( N \) total observations: thus \( SS_T \) has \( N - 1 \) degrees of freedom. If there are levels of the factor (and a treatment means), so \( SS_{\text{Treatments}} \), has \( a - 1 \) degrees of freedom. Finally, within any treatment there are \( n \) replicates providing \( n - 1 \) degrees of freedom with which to estimate the experimental error. Because there are \( a \) treatments, we have \( a(n - 1) = an - a = N - a \) degrees of freedom for error.
Statistical Analysis

The Analysis of variance table (Table 2-4) for the single-factor fixed effects model is given below

Table 2-4  ANOVA Table with formula

<table>
<thead>
<tr>
<th>Source of variance</th>
<th>Sum of squares</th>
<th>Degrees of freedom</th>
<th>Mean square</th>
<th>( F_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between treatments</td>
<td>( SS_{\text{Treatments}} = \sum_{i=1}^{a} \sum_{j=1}^{n} (\bar{y}<em>{ij} - \bar{y}</em>{..})^2 )</td>
<td>a-1</td>
<td>( MS_{\text{Treatments}} )</td>
<td>( F_0 = \frac{MS_{\text{Treatments}}}{MS_E} )</td>
</tr>
<tr>
<td>Error (within treatments)</td>
<td>( SS_{\text{Error}} = SS_T - SS_{\text{Treatments}} )</td>
<td>a(n-1)</td>
<td>( MS_E )</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( SS_T = \sum_{i=1}^{a} \sum_{j=1}^{n} (\bar{y}<em>{ij} - \bar{y}</em>{..})^2 )</td>
<td>an-1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Because the degrees of freedom for \( SS_{\text{Treatments}} \) and \( SS_{\text{Error}} \) add to \( N-1 \), the total number of degrees of freedom, Cochran’s theorem implies that \( SS_{\text{Treatments}} / \sigma^2 \) and \( SS_E / \sigma^2 \) are independently distributed chi-square random variables. Therefore, if the null hypothesis of no difference in treatment means is true, the ratio

\[
F_o = \frac{SS_{\text{Treatments}} / (a-1)}{SS_E / (N-a)} = \frac{MS_{\text{Treatments}}}{MS_E},
\]

is distributed as \( F \) with \( a - 1 \) and \( N - a \) degrees of freedom. Equation above is the test statistic for the hypothesis of no differences in treatment means.

From the expected mean squares we see that, in general, \( MS_E \) is an unbiased estimator of \( \sigma^2 \). Also, under the null hypothesis, \( MS_{\text{Treatments}} \) is an unbiased estimator of \( \sigma^2 \). However, if the null
hypothesis is false, the expected value of $MS_{\text{treatments}}$ is greater than $\sigma^2$. Therefore, under the alternative hypothesis, the expected value of the numerator of the test statistic (Equation given above for $F_o$) is greater than the expected value of the denominator, and we should reject $H_0$ on values of the test statistic that are too large. This implies an upper-tail and one-tail critical region. Therefore, we should reject $H_0$ and conclude that there are differences in the treatment means if

$$F_o > F_{a, a-1, N-a},$$

where, $F_o$ is computed from above equation. Alternatively, we can also use the p-value approach for decision making as provided by statistical softwares, say MINITAB, SAS.

Using MINITAB, we can obtain the following graphs and results for the above mentioned experiment on tensile strength:

![Boxplot of Tensile Strength](image)

**Figure 2-19** Box Plot of Data

From Box-plot it is observed that as cotton weight % increases tensile strength also improves. However, whether any two means are significantly different cannot be commented based on Box plot.
The residual plot confirms Normality assumption of error. Conclusion in case error is non-normal may be erroneous. Normality assumption can be tested by using Anderson-Darling test statistic value provided in MINITAB (Stat->Basic Statistics-> Normality Test).

**One-way ANOVA Analysis: Tensile Strength versus Weight % of cotton**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight percent of cotton</td>
<td>3</td>
<td>340.15</td>
<td>113.38</td>
<td>14.77</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>16</td>
<td>122.80</td>
<td>7.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>462.95</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

S = 2.770  R-Sq = 73.47%  R-Sq(adj) = 68.50%

**Individual 95% CIs For Mean Based on Pooled StDev**

<table>
<thead>
<tr>
<th>Level</th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
<td>10.000</td>
<td>3.162</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>15.800</td>
<td>3.114</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>16.800</td>
<td>1.924</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>21.600</td>
<td>2.702</td>
</tr>
</tbody>
</table>

Pooled St. Dev = 2.770

**Figure 2-21** ANOVA Results in MINITAB
ANOVA results given above confirm that changing the % weight influences tensile strength and it is linear (from R-square value).

For determining the best setting of % cotton weight, one can do the Fisher LSD comparison test as given in Figure 2-22.

Fisher 95% Individual Confidence Intervals

All Pairwise Comparisons among Levels of Weight percent of cotton

Simultaneous confidence level = 81.11%

Weight percent of cotton =  5 subtracted from:

<table>
<thead>
<tr>
<th>Weight percent of cotton</th>
<th>Lower</th>
<th>Center</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2.086</td>
<td>5.800</td>
<td>9.514</td>
</tr>
<tr>
<td>15</td>
<td>3.086</td>
<td>6.800</td>
<td>10.514</td>
</tr>
<tr>
<td>20</td>
<td>7.886</td>
<td>11.600</td>
<td>15.314</td>
</tr>
</tbody>
</table>

Weight percent of cotton = 10 subtracted from:

<table>
<thead>
<tr>
<th>Weight percent of cotton</th>
<th>Lower</th>
<th>Center</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>-2.714</td>
<td>1.000</td>
<td>4.714</td>
</tr>
<tr>
<td>20</td>
<td>2.086</td>
<td>5.800</td>
<td>9.514</td>
</tr>
</tbody>
</table>

Weight percent of cotton = 15 subtracted from:

<table>
<thead>
<tr>
<th>Weight percent of cotton</th>
<th>Lower</th>
<th>Center</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1.086</td>
<td>4.800</td>
<td>8.514</td>
</tr>
</tbody>
</table>

Figure 2-22 Fisher Comparison Test
The comparison tests confirm that cotton weight % 20 is significantly different from 15 %, and thus setting of 20 is suggested. Details on interpretation of comparison test and ANOVA is given in MINITAB example, MINITAB help, and any standard text book on Quality (Montgomery, D. C., 2014)

**Designed Experiment**

Statistical design of experiments refers to the process of planning an experiment so that appropriate data which can be analyzed by statistical methods can be collected, resulting in valid and objective conclusions. Statistical approach to experimental design is necessary if we wish to draw meaningful conclusions from the data. This helps in confirming any causal relationship. When problem involves data that are subject to experimental errors, statistical methodology is the only objective approach for analysis. Thus, there are two aspects to any experimental problem: design of experiment and statistical analysis of data.

Three basic principles of experimental design are replication, randomization and blocking (local control). By replication we mean a repetition of basic trial on different sample. In the metallurgical experiment, twice replication would consist of treating two specimens by oil quenching. Thus, if five specimens are treated in a quenching medium in different time point, we say that five replicates have been obtained.

Randomization is the cornerstone underlying use of statistical methods in experimental design. By randomization we mean that both allocation of experimental material and order in which individual runs or trials of experiment are to be performed are randomly determined. Statistical methods require that observations (or errors) be independently distributed random variables. Randomization usually validates this assumption. By properly randomizing an experiment, we also assist in "averaging out" the effects of extraneous (hidden) factors that may be present.

Blocking (or local control) nuisance variables is a design technique used to improve precision with which comparisons among factors of interest are made. For example, an experiment in a chemical process may require two batches of raw material to make all required runs. However, there could be differences between batches due to supplier-to-supplier variability, and if we are not specifically interested in this effect, we would think of different batches of raw material as a
nuisance factor. Generally, a block is a set of relatively homogeneous experimental conditions. There are many design options for blocking nuisance variables.

**Guidelines for designing experiment**

To use statistical approach in designing and analyzing an experiment, it is necessary for everyone involved in the experiment to have a clear idea in advance of exactly what is to be studied (objective of study), how data is to be collected, and at least an understanding of how this data is to be analyzed. Below section briefly discusses on outline and elaborate on some of the key steps in Design of Experiment (DOE). Remember that experiment can fail. However, it always provides some meaningful information.

**Recognition of a problem and its statement**- This may seem to be rather obvious point, but in practice it is often not simple to realize that a problem requiring experimentation exists, nor is it simple to develop a clear and generally accepted statement of problem. It is necessary to develop all ideas about objectives of experiment. Usually, it is important to solicit input from all concerned parties: engineering, quality assurance, manufacturing, marketing, management, customers (internal or external) and operating personnel.

It is usually helpful to prepare a list of specific problems or questions that are to be addressed by the experiment. A clear statement of problem often contributes substantially to better understanding of phenomenon being studied and final solution to the problem. It is also important to keep overall objective in mind.

**Choice of factors, levels, and range**- When considering factors that may influence performance of a process or system, experimenter usually finds that these factors can be classified as either potential design (x) factors or nuisance (z) factors. Potential design factors are those factors that experimenter may wish to vary during the experiment. Often we find that there are a lot of potential design factors, and some further classification of them is necessary. Some useful classification is design factors, held-constant factors, and allowed to-vary factors. The design factors are the factors actually selected for study in the experiment. Held-constant factors are variables that may exert some effect on the response, but for purposes of present experiment these factors are not of
interest, so they will be held at a specific level.

Nuisance (allowed to-vary) factors, on the other hand may have large effects that must be accounted for, yet we may not be interested in them in the context of the present experiment.

**Selection of the response variable**- In selecting response variable, experimenter should be certain that this variable really provides useful information about process under study. Most often, average or standard deviation (or both) of the measured characteristic will be response variable. It is critically important to identify issues related to defining responses of interest and how they are to be measured before conducting the experiment. Sometimes designed experiments are employed to study and improve the performance of measurement systems.

**Choice of experimental design**- If the pre-experimental planning activities mentioned above are done correctly, this step is relatively easy. Choice of design involves consideration of sample size (number of replicates) keeping in mind precision required for experiment, selection of a suitable run order for the experimental trials and determination of whether or not blocking restrictions are to be involved.

**Performing the experiment**- When running an experiment, it is vital to monitor the process carefully to ensure that everything is being done according to plan. Errors in experimental procedure or instrumental error during measurement at this stage will usually destroy experimental validity. Up-front planning is crucial to success. It is easy to underestimate the logistical and planning aspects of running a designed experiment in a complex manufacturing or research and development environment. Coleman and Montgomery (1993) suggest that prior to conducting experiment a few trial runs or pilot runs are often helpful.

**Statistical analysis of the data**- Statistical methods should be used to analyze the data so that results and conclusions are objective rather than judgmental in nature. If the experiment has been designed correctly statistical methods required are not elaborate. There are many excellent software packages (JMP, MINITAB, and DESIGN EXPERT) to assist in data analysis. Often we find that simple graphical methods play an important role in data analysis and interpretation.
Conclusions and recommendations- Once the data is analyzed, experimenter must draw practical conclusions from the results, and recommend a course of action. Graphical methods are often useful in this stage, particularly in presenting results to others. Follow-up runs and confirmation testing should also be performed to validate the conclusions from the experiment.

There are many design options for statistical experiment. For two factor experiment, the basic design is a two-way ANOVA. For higher number of factors, factorial design using orthogonal array is typically used. There are also central composite designs (CCD), extremely useful to identify higher order terms in the response surface model developed based on factorial design (http://www.youtube.com/watch?v=Z-uqadwwFsU). Three level Box–Behnken or BBD design (http://www.itl.nist.gov/div898/handbook/pri/section3/pri3362.htm) is also very useful in situation to identify quadratic terms and interaction in factorial design. Fractional factorial design is recommended if more than 8 factors are to be studied and there is a need to reduce the number of factors. This is also known as ‘screening experiment’. Sequential DOE or Response surface design is used to reach to global optimal setting in case of unimodal function. Taguchi’s method is a nonconventional approach used when there is little or no higher order interaction. Desirability function and dual response optimization may be used in case there are multiple y’s to be optimized simultaneously. Book by Montgomery, D.C. (2014) is an excellent reference to learn DOE techniques.