CHAPTER - I

INTRODUCTION
QUALITY is the main concern of any customer whereas PROFIT is the main interest of the seller or the producer. The above two terms namely QUALITY and PROFIT, play a very vital role in any industry. The term QUALITY may be defined as follows:

*The totality of features and characteristics of a product of service that bear on its ability to satisfy stated or implied needs of the user of that product.*
Thus the term QUALITY is mainly related to the satisfaction derived by the user, which is naturally measured only after using the product. In this sense it varies from person to person, time to time and material to material. For the purpose of analysis it may be measured through some chief characteristics, of that product. For instance, the quality of an electric bulb may be measured on the thickness of the filament or breaking strength of the shell if it is dropped from certain height or the number of hours it burns before it fails and so on. Similarly, the quality of cloth can be measured on the breaking strength of the yarn and so on. Further, in repetitive processes, there is a wide possibility of variation in the quality of the product and hence, there exists a necessity to control the same.

The quality of any product is mainly based on the raw material, technical know-how, design and the workmanship of the workers. On the other hand profit mainly, depends on the quality of the product that is maintained at a desired level by the combined effect of the designers, producers, researchers and workers. The
awareness to maintain the quality of an industrial output has been there, since the beginning of the industry and hence the concept of QUALITY CONTROL has a long history.

The concept of QUALITY CONTROL gives different meaning to different people. To the machine operator, sometimes it means an inconsistent putting-up and tearing-down of control charts. To the inspector, it can mean a set of sampling tables which he is to use cookbook style. To the design and production engineers, it can signify an infringement upon and annexation of their rights and duties. To a statistician, it may mean the application and solution of statistical formulae.

Thus quality control is the responsibility of every one in the enterprise. It can be defined as follows:

    Quality Control in its broadest sense refers to a spectrum of managerial methods for attempting to maintain the Quality of manufactured articles at a desired level.
Earlier, various measures were taken to control the quality which were laborious and time consuming. For instance, the long period of training for apprentices, factory inspection and research. The variation in quality can be basically classified into two categories namely (i) Chance Variation and (ii) Assignable Variation. The chance variation is the sum of the effect of the whole complex of random causes. In other words, this variation is small and the causes for this variation are numerous. The assignable variation is attributable to some special causes. Here the variation is large and causes for variation are very few in number and hence can be detected and eliminated from the process of production. If the chance variations are ordered in time or possibly on some other basis, they will behave in a random manner without showing cycles or runs or any other definite pattern. Further, it is also important to note that no specific variation to come can be predicted from the knowledge of past variation. Moreover, the variation produced by chance causes follow statistical laws. This paves the way for applying sophisticated statistical techniques to control the
quality. Sir Walter A. Shewhart of the Bell Telephone Laboratories was the first person who applied statistical methods to control the quality of an industrial output in the 1920's. This gave birth to a new field of statistics popularly known as STATISTICAL QUALITY CONTROL (SQC).

These new techniques were applied first in the Second World War, by the American Military to maintain the quality of war-material, food and medicine and they came out as quick, scientific and reliable methods to maintain the quality of the product. Since then, the appliational potentialities of these SQC techniques were spread like a Wildfire to all other fields of industry throughout the World. It is not an exaggeration to state that almost every industry now-a-days has its own STATISTICAL QUALITY CONTROL UNIT.

The chief objectives of these SQC techniques are (1) To produce quality relatively at a cheaper rate and (2) To optimize the profits. Numerous SQC techniques are developed by the researchers, suitable to various
types of products applicable at different stages of production starting from the designing state of a product, till it reaches the customer. These techniques can broadly be classified into the following five categories, namely:

(1) Lot acceptance sampling plans
(2) Rectifying inspection
(3) Control charts
(4) Reliability and
(5) Life testing.

Among these techniques control charts and rectifying inspection are mainly used for those products during the manufacturing stage, whereas lot acceptance sampling plans, reliability and life testing are used for the finished products. Life testing plays a prominent role in estimating the reliability of a unit/system. Agree [1] in 1957 defined reliability as follows:

Reliability is the probability of performing without failure a specified function under the given conditions for a specified period of time.
In life testing problems INSPECTION COST is another important concept which plays a predominant role particularly when the Testing involves destruction. This is because of the fact that, an increase in the inspection cost will automatically increase the cost of the unit. Thus due consideration is to be given to the inspection cost particularly when destruction is involved while testing the products, like T.V. colour Picture tubes, where the production cost itself is very high. This makes the necessity to terminate a life testing experiment in the middle, which results in obtaining a CENSORED DATA and one has to take decisions based on such data. Further details of censoring are discussed elaborately in the following section.

1.2. CENSORING IN LIFE TESTING EXPERIMENTS

Consider a life testing experiment of electric bulbs. A sample of n units is selected at random from all the bulbs produced in a batch or from a particular machine or from a shift. These selected n units are
placed for the test and the variable under consideration is the life time of the unit which is denoted by \( X \). The data regarding all the life times of \( n \) units say \( x_1, x_2, \ldots, x_n \) when recorded is called UN-CENSORED DATA or COMPLETE SAMPLE DATA. Estimation with complete sample is direct, simple and there are numerous methods of estimation available in literature.

On the other hand, one may want to terminate the test before complete information is recorded on all \( n \) items put for the test, for several reasons. One of the basic reasons may be that, the test is a destructive one so that the items on the test cannot be re-used. Hence, one has to terminate the test in the middle because one may not be willing to lose many products in testing particularly when the production cost of each unit is very high. Another reason to terminate test in the middle may be the time constraint and one cannot afford to wait indefinitely for a long time until all units put on test to fail. Thus there exist a natural motivation and necessity to develop the methods to estimate parameters based on the available data of those
experiments terminated in the middle. Such partial information is usually known as CENSORED DATA.

In the earlier example of testing the life times of electric bulbs if the test is terminated in the middle, we obtain a CENSORED DATA. Let \( r \) (\(<n\)) denote the number of bulbs failed up to the termination of the test and \( x_1, x_2, \ldots, x_r \) be the respective life times of these \( r \) bulbs. Since the remaining \((n-r)\) bulbs have not failed up to the termination of the test the full knowledge about their life times is not known. The data recorded namely \( x_1, x_2, \ldots, x_r \) is called CENSORED DATA because the exact life times of all the \( n \) bulbs are not known. Here we only know that \((n-r)\) bulbs, worked up to the termination of the test, which does not mean that the total life times of those \((n-r)\) bulbs are equal to the test termination time. Survival analysis is another one such field where Censored data occurs and one has to take decisions based on such partial information available to oneself.

Depending upon the nature of the criterion to terminate a test, we obtain the following four types of
censoring, namely:

(1) TYPE I CENSORING:

Let \( n \) items are put on a test and the test is terminated at a pre-determined time \( t_0 \), so that complete information on the first \( r \) ordered observations are recorded as follows:

\[ x(1) < x(2) < \cdots < x(r) \]

Here \( r \) is an integer valued random variable such that

\[ x(r) < t_0 < x(r+1) \]

Each of the remaining unobserved life times is known to be greater than \( t_0 \). Here \( t_0 \) is called test termination time and the type of censoring is called TYPE-I CENSORING. This type of sample is known as time censored sample.

(2) TYPE II CENSORING:

As in type I censoring, let \( n \) items are simultaneously put on a test and the test is terminated after a predetermined number \( k \) (or fraction \( k/n \)) of failures. Here we obtain the complete information on the first \( k \) ordered observations only and the data recorded is as follows:
Here also we know that the remaining \((n-k)\) items to be greater than \(x_{(k)}\). Here \(k\) or \((k/n)\) is a fixed constant. This type of sample is known as failed censored sample.

(3) **TYPE III ARBITRARY CENSORING AND RANDOM CENSORING**

In (1) and (2) it is assumed that all the \(n\) items are put on test simultaneously (or that the ordered observations are available). There may be a situation, however, where all items cannot be tested simultaneously. For example, each item may require certain time \(T_s\) to set up for the test and it may not be feasible to install all \(n\) items simultaneously for the test. Here the test starts at a time point \(t_0\) and will be terminated at time \(T\). Such type of censoring is more useful in clinical experiment, where patients enter for a service at different time points.

(4) **TYPE IV PROGRESSIVE CENSORING**:

Let \(n\) items be put on test in a life testing
experiment. At a predesignated censoring time $t_i$, a fixed number $C_i \geq 0$ of items are removed from the test ($i = 1, 2, \ldots, r; t_1 < t_2 < \ldots < t_r$). Finally at the time $t_r$ either a fixed number of $C_r$ items are removed from the test or testing is terminated with a random number $C_r$ items still functioning. Instead of fixed times $t_1, t_2, \ldots, t_r$, one could also carry out progressive censoring after specified number of failures. Thus after first $n_1$ failures an additional number $C_1$ may be removed, and the test be continued. After $n_2$ additional failures, $C_2$ further items may be removed and so on such that in $n$ steps, we remove $C_1 + C_2 + \ldots + C_n$ items from the test. Here the main aim of the progressive censoring is to reduce the testing time.

Now we proceed to give a brief literature review on the censored sampling in the following section.

1.3. LITERATURE REVIEW

It is already mentioned in the previous section
that the natural motivation and necessity to study the problems relating to the censored sampling arises in many practical fields of application like life testing, reaction-time studies, survival analysis and so on. In these experiments it is common practice to terminate the experiment before all items or specimens fail. A lot of research work is done relating to the censored sampling in the literature. A purposeful review of the literature is given below relevant to the present study.

of parameters from censored data. Kaplan and Meier [18] have suggested the application of non-parametric methods for censored data in 1958. The method of estimation is popularly known as Kaplan Meier estimator. After successful application of non-parametric estimations in censored data a lot of attention has been paid on this type of research work. Some of the noteworthy works here are Halprin [14], Gilber [10], Gehan [9], Mantel [20] and Eform [7]. In 1958 Mendenhall and Hader [22] discussed the estimation of parameters of mixed exponentially distributed failure time distribution from a censored failure data.

In 1961 Gupta and Groll [13] discussed the use of Gamma distribution in acceptance sampling plans with type I censored samples and presented a number of tables and charts for estimating the sample size necessary to guarantee a specified mean life, producers risk, confidence interval and so on. Wilk, Gnanadesikan and Huyett [35] in the year 1962 obtained MLE of the parameters of a Gamma distribution based on order statistics from a censored data. Roberts [27] in his
normal distribution from a single censored samples using maximum likelihood. A detailed discussion on the methods of estimation of parameters to exponential, gamma, Weibull and normal distributions with different types of censoring is given in his book [34] by Sinha, S.K., in 1986. Since then these techniques of estimating with censored data are applied in different fields of research wherever it is necessary. This dissertation is mainly based on those methods suggested by Sinha [34] in the case of exponential and Weibull distributions when the test is terminated at a predetermined time $t_0$.

1.4. MOTIVATION AND OBJECTIVES

It is already discussed in the previous section that different researchers have considered different methods of estimation suitable to various situations that arise in day to day life, Sinha [34] in 1986 has
given maximum likelihood estimates for the parameters of the distribution exponential, Weibull and other related distributions which have a lot of practical value in life testing experiments. This motivated us to demonstrate the utility of those results with a live data collected from Andhra Pradesh Lightings Limited (A.P.L.L) located at Anantapur, that manufactures electrical bulbs. The chief objectives of this dissertation are three fold namely:

(1) To demonstrate the utility of the results presented by Sinha relevant to type 1 censored samples in exponential and weibull distributions.

(2) To obtain the ML estimates of average life time of the bulbs of A.P.L.L:
   (a) Produced in different machines,
   (b) Produced in different shifts,
   (c) different types of bulbs with respect to their (Volts/Watts), (V/W) and to compare with each other, and

(3) To obtain the ML estimates of reliability of the bulbs of A.P.L.L of the above categories of a,b and c and to compare with each other.
The standard errors of the above estimates are also obtained in each case. With the above objectives, the data is collected from the records of A.P.L.L relating to the life testing experiments conducted by the SQC unit. The formulae used to estimate the above parameters are given in the following section for a ready reference.

1.5. ESTIMATION METHODS WITH TYPE I CENSORING

Let \( n \) represent the number of bulbs put for a life testing experiment and \( x_i \) represent the life time of \( i \)th bulb measured in hours. Let \( r \) represent the number of bulbs failed before the test termination time \( t_0 \). For the purpose of analysis first it is assumed the failure times follow an exponential distribution and the conditional probability density function of the failure time, given that the item has failed before \( t_0 \). This is given by

\[
h(x/\sigma) = \frac{1}{\sigma} e^{(-x/\sigma)} \frac{1 - e^{(-t_0/\sigma)}}{1 - e^{(-t_0/\sigma)}}, \quad 0 < x \leq t_0 \quad \ldots (1.5.1)
\]

\[
= 0, \text{ otherwise.}
\]
Estimation of the mean life time $\alpha$ can be done in the following two ways namely (i) Based on failure times and (ii) Based on the number of failures in the interval $[0, t_o]$.

Case : (i) Estimation of $\alpha$ based on failure times

Let $x(1), x(2), \ldots, x(r)$ represent the order statistics of $r$ items failed before $t_o$. The MLE of $\alpha$ denoted by $\hat{\alpha}$ is given by

$$\hat{\alpha} = \frac{\sum_{i=1}^{r} x(i) + (n-r)t_o}{r}, \quad r > 0 \quad \ldots (1.5.2)$$

$$= nt_o \quad \text{for } r = 0.$$  

The Bias of the estimate $\hat{\alpha}$ is given by

$$\text{Bias}(\hat{\alpha}) = \frac{tq}{np^2} \quad \ldots (1.5.3)$$

and S.E.($\hat{\alpha}$) is given by

$$\text{SE}(\hat{\alpha}) = \frac{\hat{\alpha}}{\sqrt{np}} \quad \ldots (1.5.4)$$

where

$$p = 1 - e^{-t_o/\hat{\alpha}} \quad \ldots (1.5.5)$$

$$q = 1 - p. \quad \ldots (1.5.6)$$

The MLE of the reliability $R(t)$ denoted by $\hat{R}(t)$ is given by:

$$\hat{R}(t) = e^{-t_o/\hat{\alpha}}. \quad \ldots (1.5.7)$$
Case (ii) : Based on the number of failures in the interval \([0, t_0]\).

Let \(r\) represents the number of items failed before \(t_0\) out of \(n\) items put for the test. This method gives good results when the proportion of failures i.e., \(r/n\) is in the \((0.2, 0.8)\). The estimated mean life time in this method is denoted by \(\hat{\sigma}'\) and is given by:

\[
\hat{\sigma}' = -t_0 \log(1-r/n) \quad \ldots(1.5.8)
\]

The variance and standard error of this estimate is given by

\[
\text{Var}(\hat{\sigma}') = \frac{\sigma^2 p}{nq(\log q)^2} \quad \ldots(1.5.9)
\]

\[
\text{S.E.}(\hat{\sigma}') = \frac{\hat{\sigma}' \sqrt{p}}{\sqrt{nq(\log q)^2}} \quad \ldots(1.5.10)
\]

The MLE of the reliability \(R(t)\) denoted by \(\hat{R}(t)\) is given by

\[
\hat{R}(t) = e^{-t_0/\hat{\sigma}'} \quad \ldots(1.5.11)
\]

A two parametric Weibull distribution is more appropriate and generalized version of an exponential distribution. Hence in the second phase of analysis it is assumed that the random variable under consideration
is a two parametric Weibull distribution, whose conditional probability density function of the failure time given that the item has failed before time \( t_o \) is given by

\[
f(x/\alpha, \alpha) = \frac{\alpha \left( \frac{x}{\alpha} \right)^{\alpha-1} \left( \frac{x}{\alpha} \right)^{\alpha}}{1 - e^{-\left( \frac{t_o}{\alpha} \right)^{\alpha}}} \quad \alpha > 0, \quad \alpha > 0, \quad 0 \leq x \leq t_o
\]

\[
= 0, \text{ otherwise.} \quad \ldots (1.5.12)
\]

\( \alpha \) = shape parameter

\( \sigma \) = scale parameter.

In 1960 Mendenhall and Lehman [23] studied the mean and variance of the estimator of \( \alpha \) for a two parametric Weibull distribution whose c.p.d.f is given in (1.5.12) when \( \alpha \) is known and the testing is terminated at a pre-determined time \( t_o \). The MLE of \( \alpha \) is given by

\[
\hat{\alpha} = \frac{\sum_{i=1}^{r} x_i^{\alpha} + (n-r)t_o^{\alpha}}{r} \quad \ldots (1.5.13)
\]

\[
E(\hat{\alpha}) = \alpha + t_o^{\alpha} \left[ nE(1/r) - 1/p \right] \quad \ldots (1.5.14)
\]

\[
Var(\hat{\alpha}) = t_o^{2\alpha} \left[ \frac{2^2}{nV(1/r)} \right] E(1/r) \quad \ldots (1.5.15)
\]

where

\[
q = e^{-t_o/\sigma} \quad \ldots (1.5.16)
\]
The MLE of the reliability $R(t)$ denoted by $\hat{R}(t)$ is given by

$\hat{R}(t) = e^{-t\alpha/\phi}$  \hspace{1cm} (1.5.21)

With the help of above formulae now we proceed to obtain in the following two chapters, the average life times, Reliabilities and their standard errors of the bulbs produced in A.P.L.L., Anantapur with a live data collected from it.